Time-Frequency Analysis

Concepts and Methods

Edited by
Franz Hlawatsch
François Auger
Time-Frequency Analysis
This page intentionally left blank
Time-Frequency Analysis

Concepts and Methods

Edited by
Franz Hlawatsch
François Auger
# Contents

Preface .......................................................... 13

**FIRST PART. FUNDAMENTAL CONCEPTS AND METHODS** ............ 17

Chapter 1. Time-Frequency Energy Distributions: An Introduction ... 19
Patrick FLANDRIN

  1.1. Introduction ............................................. 19
  1.2. Atoms .................................................. 20
  1.3. Energy .................................................. 21
      1.3.1. Distributions .................................... 22
      1.3.2. Devices .......................................... 22
      1.3.3. Classes .......................................... 23
  1.4. Correlations ........................................... 26
  1.5. Probabilities .......................................... 27
  1.6. Mechanics ............................................. 29
  1.7. Measurements .......................................... 29
  1.8. Geometries ............................................ 32
  1.9. Conclusion ............................................. 33
  1.10. Bibliography .......................................... 34

Chapter 2. Instantaneous Frequency of a Signal ....................... 37
Bernard PICINBONO

  2.1. Introduction ............................................. 37
  2.2. Intuitive approaches .................................... 38
  2.3. Mathematical definitions ................................ 40
      2.3.1. Ambiguity of the problem ......................... 40
      2.3.2. Analytic signal and Hilbert transform ............ 40
      2.3.3. Application to the definition of instantaneous frequency .... 42
      2.3.4. Instantaneous methods ............................ 45
  2.4. Critical comparison of the different definitions .............. 46
      2.4.1. Interest of linear filtering ....................... 46
Chapter 3. Linear Time-Frequency Analysis I: Fourier-Type Representations

Rémi GRIBONVAL

3.1. Introduction .................................. 61
3.2. Short-time Fourier analysis ......................... 62
  3.2.1. Short-time Fourier transform ..................... 63
  3.2.2. Time-frequency energy maps ..................... 64
  3.2.3. Role of the window .......................... 66
  3.2.4. Reconstruction/synthesis ....................... 71
  3.2.5. Redundancy ................................ 71
3.3. Gabor transform; Weyl-Heisenberg and Wilson frames ................. 71
  3.3.1. Sampling of the short-time Fourier transform .......... 71
  3.3.2. Weyl-Heisenberg frames ........................ 72
  3.3.3. Zak transform and “critical” Weyl-Heisenberg frames .... 74
  3.3.4. Balian-Low theorem ........................... 75
  3.3.5. Wilson bases and frames, local cosine bases .......... 75
3.4. Dictionaries of time-frequency atoms; adaptive representations ... 77
  3.4.1. Multi-scale dictionaries of time-frequency atoms ...... 77
  3.4.2. Pursuit algorithm ............................ 78
  3.4.3. Time-frequency representation ................... 79
3.5. Applications to audio signals .................... 80
  3.5.1. Analysis of superimposed structures ................ 80
  3.5.2. Analysis of instantaneous frequency variations ........ 80
  3.5.3. Transposition of an audio signal ................ 82
3.6. Discrete algorithms .......................... 82
  3.6.1. Fast Fourier transform ........................ 83
  3.6.2. Filter banks: fast convolution ................... 83
  3.6.3. Discrete short-time Fourier transform ............ 85
  3.6.4. Discrete Gabor transform ....................... 86
3.7. Conclusion .................................. 86
Chapter 4. Linear Time-Frequency Analysis II: Wavelet-Type Representations

Thierry BLU and Jérôme LEBRUN

4.1. Introduction: scale and frequency ........................................... 94
4.2. Continuous wavelet transform ................................................ 95
  4.2.1. Analysis and synthesis .................................................. 95
  4.2.2. Multiscale properties ................................................... 97
4.3. Discrete wavelet transform .................................................. 98
  4.3.1. Multi-resolution analysis ............................................... 98
  4.3.2. Mallat algorithm .......................................................... 104
  4.3.3. Graphical representation .............................................. 106
4.4. Filter banks and wavelets ................................................... 107
  4.4.1. Generation of regular scaling functions .............................. 108
  4.4.2. Links with approximation theory ..................................... 111
  4.4.3. Orthonormality and bi-orthonormality/perfect reconstruction .. 112
  4.4.4. Polyphase matrices and implementation ............................. 114
  4.4.5. Design of wavelet filters with finite impulse response .......... 114
4.5. Generalization: multi-wavelets ............................................ 116
  4.5.1. Multi-filter banks ....................................................... 116
  4.5.2. Balancing and design of multi-filters ................................ 118
4.6. Other extensions .............................................................. 121
  4.6.1. Wavelet packets ......................................................... 121
  4.6.2. Redundant transformations: pyramids and frames ................ 122
  4.6.3. Multi-dimensional wavelets ........................................... 123
4.7. Applications ................................................................. 124
  4.7.1. Signal compression and denoising ................................... 124
  4.7.2. Image alignment .......................................................... 125
4.8. Conclusion .............................................................................. 125
4.9. Acknowledgments ................................................................... 126
4.10. Bibliography ......................................................................... 126

Chapter 5. Quadratic Time-Frequency Analysis I: Cohen’s Class ........ 131

François AUGER and Éric CHASSANDE-MOTTIN

5.1. Introduction ............................................................................. 131
5.2. Signal representation in time or in frequency ......................... 132
  5.2.1. Notion of signal representation ....................................... 132
  5.2.2. Temporal representations ............................................... 133
  5.2.3. Frequency representations ............................................. 134
  5.2.4. Notion of stationarity ..................................................... 135
  5.2.5. Inadequacy of monodimensional representations ............... 136
Chapter 5. Time-Frequency Analysis

5.3. Representations in time and frequency

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.1. “Ideal” time-frequency representations</td>
<td>137</td>
</tr>
<tr>
<td>5.3.2. Inadequacy of the spectrogram</td>
<td>140</td>
</tr>
<tr>
<td>5.3.3. Drawbacks and benefits of the Rihaczek distribution</td>
<td>142</td>
</tr>
<tr>
<td>5.4. Cohen’s class</td>
<td>142</td>
</tr>
<tr>
<td>5.4.1. Quadratic representations covariant under translation</td>
<td>142</td>
</tr>
<tr>
<td>5.4.2. Definition of Cohen’s class</td>
<td>143</td>
</tr>
<tr>
<td>5.4.3. Equivalent parametrizations</td>
<td>144</td>
</tr>
<tr>
<td>5.4.4. Additional properties</td>
<td>145</td>
</tr>
<tr>
<td>5.4.5. Existence and localization of interference terms</td>
<td>148</td>
</tr>
<tr>
<td>5.5. Main elements</td>
<td>155</td>
</tr>
<tr>
<td>5.5.1. Wigner-Ville and its smoothed versions</td>
<td>155</td>
</tr>
<tr>
<td>5.5.2. Rihaczek and its smoothed versions</td>
<td>157</td>
</tr>
<tr>
<td>5.5.3. Spectrogram and S transform</td>
<td>158</td>
</tr>
<tr>
<td>5.5.4. Choi-Williams and reduced interference distributions</td>
<td>158</td>
</tr>
<tr>
<td>5.6. Conclusion</td>
<td>159</td>
</tr>
<tr>
<td>5.7. Bibliography</td>
<td>159</td>
</tr>
</tbody>
</table>

Chapter 6. Quadratic Time-Frequency Analysis II: Discretization of Cohen’s Class

Stéphane Grassin

6.1. Quadratic TFRs of discrete signals

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1.1. TFRs of continuous-time deterministic signals</td>
<td>167</td>
</tr>
<tr>
<td>6.1.2. Sampling equation</td>
<td>167</td>
</tr>
<tr>
<td>6.1.3. The autocorrelation functions of the discrete signal</td>
<td>168</td>
</tr>
<tr>
<td>6.1.4. TFR of a discrete signal as a function of its generalized ACF</td>
<td>169</td>
</tr>
<tr>
<td>6.1.5. Discussion</td>
<td>171</td>
</tr>
<tr>
<td>6.1.6. Corollary: ambiguity function of a discrete signal</td>
<td>172</td>
</tr>
<tr>
<td>6.2. Temporal support of TFRs</td>
<td>173</td>
</tr>
<tr>
<td>6.2.1. The characteristic temporal supports</td>
<td>173</td>
</tr>
<tr>
<td>6.2.2. Observations</td>
<td>175</td>
</tr>
<tr>
<td>6.3. Discretization of the TFR</td>
<td>176</td>
</tr>
<tr>
<td>6.3.1. Meaning of the frequency discretization of the TFR</td>
<td>176</td>
</tr>
<tr>
<td>6.3.2. Meaning of the temporal discretization of the TFR</td>
<td>176</td>
</tr>
<tr>
<td>6.3.3. Aliased discretization</td>
<td>177</td>
</tr>
<tr>
<td>6.3.4. “Non-aliased” discretization</td>
<td>179</td>
</tr>
<tr>
<td>6.4. Properties of discrete-time TFRs</td>
<td>180</td>
</tr>
<tr>
<td>6.4.1. Discrete-time TFRs</td>
<td>181</td>
</tr>
<tr>
<td>6.4.2. Effect of the discretization of the kernel</td>
<td>182</td>
</tr>
<tr>
<td>6.4.3. Temporal inversion</td>
<td>182</td>
</tr>
<tr>
<td>6.4.4. Complex conjugation</td>
<td>183</td>
</tr>
<tr>
<td>6.4.5. Real-valued TFR</td>
<td>183</td>
</tr>
<tr>
<td>6.4.6. Temporal moment</td>
<td>183</td>
</tr>
</tbody>
</table>
Chapter 7. Quadratic Time-Frequency Analysis III: The Affine Class and Other Covariant Classes

Paulo Gonçalvès, Jean-Philippe Ovarlez and Richard Baraniuk

7.1. Introduction .................................................. 193
7.2. General construction of the affine class ......... 194
  7.2.1. Bilinearity of distributions ....................... 194
  7.2.2. Covariance principle ............................... 195
  7.2.3. Affine class of time-frequency representations .... 198
7.3. Properties of the affine class ....................... 201
  7.3.1. Energy ............................................. 201
  7.3.2. Marginals ......................................... 202
  7.3.3. Unitarity ......................................... 202
  7.3.4. Localization ..................................... 203
7.4. Affine Wigner distributions .......................... 206
  7.4.1. Diagonal form of kernels ......................... 206
  7.4.2. Covariance to the three-parameter affine group .... 209
  7.4.3. Smoothed affine pseudo-Wigner distributions .... 211
7.5. Advanced considerations .............................. 216
  7.5.1. Principle of tomography .......................... 216
  7.5.2. Operators and groups ............................. 217
7.6. Conclusions ............................................. 222
7.7. Bibliography ............................................. 223

SECOND PART. ADVANCED CONCEPTS AND METHODS .................. 227

Chapter 8. Higher-Order Time-Frequency Representations .......... 229

Pierre-Olivier Amblard

8.1. Motivations ............................................. 229
8.2. Construction of time-multifrequency representations .... 230
  8.2.1. General form and desirable properties .......... 230
  8.2.2. General classes in the symmetric even case .... 231
  8.2.3. Examples and interpretation ..................... 236
  8.2.4. Desired properties and constraints on the kernel .. 237
  8.2.5. Discussion ....................................... 239
10.6. Properties of the spectra of underspread processes 291
  10.6.1. Approximate equivalences 292
  10.6.2. Approximate properties 295
10.7. Estimation of time-varying spectra 296
  10.7.1. A class of estimators 296
  10.7.2. Bias-variance analysis 297
  10.7.3. Designing an estimator 299
  10.7.4. Numerical results 300
10.8. Estimation of non-stationary processes 302
  10.8.1. TF formulation of the optimum filter 303
  10.8.2. TF design of a quasi-optimum filter 304
  10.8.3. Numerical results 305
10.9. Detection of non-stationary processes 306
  10.9.1. TF formulation of the optimum detector 309
  10.9.2. TF design of a quasi-optimum detector 310
  10.9.3. Numerical results 311
10.10. Conclusion 313
10.11. Acknowledgements 315
10.12. Bibliography 315

Chapter 11. Non-stationary Parametric Modeling 321
Corinne MAILHES and Francis CASTANIÉ

11.1. Introduction 321
11.2. Evolutionary spectra 322
  11.2.1. Definition of the “evolutionary spectrum” 322
  11.2.2. Properties of the evolutionary spectrum 324
11.3. Postulate of local stationarity 325
  11.3.1. Sliding methods 325
  11.3.2. Adaptive and recursive methods 326
  11.3.3. Application to time-frequency analysis 328
11.4. Suppression of a stationarity condition 329
  11.4.1. Unstable models 329
  11.4.2. Models with time-varying parameters 332
  11.4.3. Models with non-stationary input 340
  11.4.4. Application to time-frequency analysis 346
11.5. Conclusion 348
11.6. Bibliography 349

Chapter 12. Time-Frequency Representations in Biomedical Signal Processing 353
Lotfi SENHADJI and Mohammad Bagher SHAMSOLLAHI

12.1. Introduction 353
12.2. Physiological signals linked to cerebral activity 356
12.2.1. Electroencephalographic (EEG) signals ................. 356
12.2.2. Electrocorticographic (ECoG) signals ................. 359
12.2.3. Stereoelectroencephalographic (SEEG) signals .... 359
12.2.4. Evoked potentials (EP) .................................. 362
12.3. Physiological signals related to the cardiac system ... 363
12.3.1. Electrocardiographic (ECG) signals .................... 363
12.3.2. R-R sequences ........................................... 365
12.3.3. Late ventricular potentials (LVP) ....................... 367
12.3.4. Phonocardiographic (PCG) signals ..................... 369
12.3.5. Doppler signals ......................................... 372
12.4. Other physiological signals ................................ 372
12.4.1. Electrogastrographic (EGG) signals .................... 372
12.4.2. Electromyographic (EMG) signals ..................... 373
12.4.3. Signals related to respiratory sounds (RS) .......... 374
12.4.4. Signals related to muscle vibrations .................. 374
12.5. Conclusion ............................................... 375
12.6. Bibliography ................................................ 376

Chapter 13. Application of Time-Frequency Techniques to Sound Signals: Recognition and Diagnosis ................................................................. 383
Manuel DAVY

13.1. Introduction ................................................. 383
13.1.1. Decision .................................................. 384
13.1.2. Sound signals ......................................... 384
13.1.3. Time-frequency analysis as a privileged decision-making tool 384
13.2. Loudspeaker fault detection ............................... 386
13.2.1. Existing tests ........................................... 386
13.2.2. A test signal .......................................... 388
13.2.3. A processing procedure ............................... 389
13.2.4. Application and results ................................ 391
13.2.5. Use of optimized kernels ............................. 395
13.2.6. Conclusion ................................................. 399
13.3. Speaker verification ........................................ 399
13.3.1. Speaker identification: the standard approach .... 399
13.3.2. Speaker verification: a time-frequency approach .... 403
13.4. Conclusion .................................................. 405
13.5. Bibliography ................................................ 406

List of Authors .................................................. 409

Index ............................................................... 413
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 esti-
mation and detection of underspread processes is discussed.

Chapter 11, “Non-stationary Parametric Modeling”, considers non-stationary ran-
dom 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 Sig-
nals: Recognition and Diagnosis”, proposes a time-frequency technique for super-
vised 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 representa-
tion 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 con-
tribute to a deepened understanding and appreciation of this fascinating subject, which
is still witnessing considerable developments.

We would like to thank J.-P. Ovarlez for important contributions during the ini-
tial phase of this work, G. Matz for helpful assistance and advice, and, above all,
P. Flandrin for his eminent role in animating research on time-frequency analysis.

Vienna and Saint Nazaire, June 2008.
FIRST PART

Fundamental Concepts and Methods
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

Chapter written by Patrick FLANDRIN.
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

\[
x(t_0) = \int_{-\infty}^{\infty} \hat{x}(f) e_f(t_0) \, df,
\]

where we use the notation \( e_f(t) := \exp(j2\pi ft) \). If the first decomposition favors the temporal description, the second (in which \( \hat{x}(f) \) 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

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

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 $g_{t,f}(.) \to \delta_{t}(.)$, perfect localization in frequency and no localization in time when $g_{t,f}(.) \to e_{f}(.)$). 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

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

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 $g_{t,f}(s) := g(s - t) \exp(j2\pi fs)$ 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

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

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

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

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):

$$S^0_x(t, f) := \left| \int_{-\infty}^{\infty} x(s) g^*(s-t) e^{-j2\pi fs} \, ds \right|^2. \quad (1.5)$$

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

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

(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

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

This is easily verified to equal

\[ R_x(t, f) = x(t) \, \hat{x}(f) \, e^{-j2\pi ft} \]

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

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

through the introduction of a specific, suitably chosen function \( \phi_{d-D}(\tau, \xi) \) (thus, Page’s distribution is obtained by choosing \( \phi_{d-D}(\tau, \xi) = \exp(-j\pi|\tau|\xi) \) and Rihaczek’s distribution by taking \( \phi_{d-D}(\tau, \xi) = \exp(-j\pi\tau\xi) \)). 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 \textit{a priori} specification of this function. In particular, the simplest choice, namely \( \phi_{d-D}(\tau, \xi) = 1 \), leads to the definition

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

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).

\textbf{Covariances.} The second possibility of constructing classes of solutions consists of imposing a structure that is very general \textit{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 \( K(s, s'; t, f) \) \textit{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 \textit{covariance} constraints in order to

---

3. The subscript d-D is short for “delay-Doppler”; note that \( \tau \) is a delay variable and \( \xi \) 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
\[ \rho_{Hx}(t, f) = (\tilde{H}\rho_x)(t, f) \]
be satisfied, where \( H : L^2(\mathbb{R}) \to L^2(\mathbb{R}) \) represents a transformation operator acting on the signal (and \( \tilde{H} : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2) \) 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
\[ \tilde{x}(t) := x(t - \tau) e^{j2\pi f t} \Rightarrow \rho_{\tilde{x}}(t, f) = \rho_x(t - \tau, f - \xi) \]
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
\[ \phi_{d-D}(\tau, \xi) := \int_{-\infty}^{\infty} K_0(t + \frac{\tau}{2}, t - \frac{\tau}{2}) e^{j2\pi f t} dt. \]
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]
\[ \Omega_x(t, f) = \frac{f}{f_0} \int_0^\infty \int_0^\infty \pi(\xi, \zeta) \hat{x}\left(\frac{\zeta - \xi/2}{f_0/f}\right) \hat{x}^*\left(\frac{\zeta + \xi/2}{f_0/f}\right) e^{-j2\pi(f/f_0)\xi t} d\xi d\zeta, \]
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
\[ \pi(\xi, \zeta) = \frac{\xi/2f_0}{\sinh(\xi/2f_0)} \delta\left(\zeta - \frac{\xi}{2} \coth \frac{\xi}{2f_0}\right). \]
The usual definition of the unitary Bertrand distribution is [BER 92]

\[ B_x(t, f) := f \int_{-\infty}^{\infty} \sqrt{\lambda(u)\lambda(-u)} \hat{x}(f\lambda(u)) \hat{x}^* (f\lambda(-u)) e^{-j2\pi uf} du, \quad (1.9) \]

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 \( \hat{\gamma}_x(\tau) \) of the random correlation function, which can be provided by a deterministic correlation:

\[ \hat{\Gamma}_x(f) = \int_{-\infty}^{\infty} w(\tau) \hat{\gamma}_x(\tau) e^{-j2\pi f\tau} d\tau. \quad (1.10) \]

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 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 remark-
able 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

\[ \rho_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_{d-D}(\tau, \xi) A_x(\tau, \xi) e^{j2\pi(\xi t - \tau f)} d\tau d\xi , \]

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 \(\phi_{d-D}(\tau, \xi)\), which is \textit{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 \(\phi_{d-D}(\tau, \xi)\) [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 \textit{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 character-
istic function” of time and of frequency [VIL 48].

**Distribution function.** Page’s definition mentioned previously in (1.6) can be inter-
preted 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):

\[ \rho_x(t) := |x(t)|^2, \quad \rho_x(f) := |\hat{x}(f)|^2, \]

or also, for unit-energy signals, as probability densities relative to the time and fre-
quency 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:
\[
\int_{-\infty}^{\infty} \rho_x(t, f) \, dt = \rho_x(f), \quad \int_{-\infty}^{\infty} \rho_x(t, f) \, df = \rho_x(t). \tag{1.11}
\]

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 \(\phi_{d-D}(0, \xi) = \phi_{d-D}(\tau, 0) = 1\)).

**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* \(t_x(f)\)
\[
\int_{-\infty}^{\infty} t \rho_x(t|f) \, dt = t_x(f) := -\frac{1}{2\pi} \frac{d}{df} \arg \hat{x}(f)
\]
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 Co-
hen’s class, namely [FLA 99a]
\[
\frac{\partial \phi_{d-D}(\tau, \xi)}{\partial \xi} \bigg|_{\xi=0} = \frac{\partial \phi_{d-D}(\tau, \xi)}{\partial \tau} \bigg|_{\tau=0} = 0.
\]
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 character-
ized 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 distri-
bution. However, this viewpoint encounters a definition difficulty due to the pos-
sible 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