

APPLICATIONS IN TIME-FREQUENCY SIGNAL PROCESSING

Edited by

Antonia Papandreou-Suppappola



CRC PRESS

APPLICATIONS IN TIME-FREQUENCY SIGNAL PROCESSING

THE ELECTRICAL ENGINEERING AND APPLIED SIGNAL PROCESSING SERIES

Edited by Alexander Pouliarikas

*The Advanced Signal Processing Handbook:
Theory and Implementation for Radar, Sonar,
and Medical Imaging Real-Time Systems*

Stergios Stergiopoulos

The Transform and Data Compression Handbook
K.R. Rao and P.C. Yip

Handbook of Multisensor Data Fusion
David Hall and James Llinas

Handbook of Neural Network Signal Processing
Yu Hen Hu and Jenq-Neng Hwang

Handbook of Antennas in Wireless Communications
Lal Chand Godara

Noise Reduction in Speech Applications
Gillian M. Davis

Signal Processing Noise
Vyacheslav P. Tuzlukov

Digital Signal Processing with Examples in MATLAB®
Samuel Stearns

Applications in Time-Frequency Signal Processing
Antonia Papandreou-Suppappola

Forthcoming Titles

Propagation Data Handbook for Wireless Communication System Design
Robert Crane

The Digital Color Imaging Handbook
Gaurav Sharma

Smart Antennas
Lal Chand Godara

Pattern Recognition in Speech and Language Processing
Wu Chou and Bing Huang Juang

Nonlinear Signal and Image Processing: Theory, Methods, and Applications
Kenneth Barner and Gonzalo R. Arce

Forthcoming Titles (*continued*)

Soft Computing with MATLAB®

Ali Zilouchian

Signal and Image Processing Navigational Systems

Vyacheslav P. Tuzlukov

Wireless Internet: Technologies and Applications

Apostolis K. Salkintzis and Alexander Poularikas

APPLICATIONS IN TIME-FREQUENCY SIGNAL PROCESSING

Edited by

Antonia Papandreou-Suppappola

*Arizona State University
Tempe, Arizona*



CRC PRESS

Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Papandreou-Suppappola, Antonia

Applications in time-frequency signal processing / Antonia Papandreou-Suppappola.

p. cm. (Electrical engineering and applied signal processing series)

Includes bibliographical references and index.

ISBN 0-8493-0065-7 (alk. paper)

1. Signal processing. 2. Time-series analysis. 3. Frequency spectra. I. Title. II. Series

TK5102.9 .P37 2002

621.382'2—dc21

200273356

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the authors and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming, and recording, or by any information storage or retrieval system, without prior permission in writing from the publisher.

All rights reserved. Authorization to photocopy items for internal or personal use, or the personal or internal use of specific clients, may be granted by CRC Press LLC, provided that \$1.50 per page photocopied is paid directly to Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 USA. The fee code for users of the Transactional Reporting Service is ISBN ISBN 0-8493-0065-7/03/ \$0.00+\$1.50. The fee is subject to change without notice. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

The consent of CRC Press LLC does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific permission must be obtained in writing from CRC Press LLC for such copying.

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation, without intent to infringe.

Visit the CRC Press Web site at www.crcpress.com

© 2003 by CRC Press LLC

No claim to original U.S. Government works

International Standard Book Number ISBN 0-8493-0065-7

Library of Congress Card Number 200273356

Printed in the United States of America 1 2 3 4 5 6 7 8 9 0

Printed on acid-free paper

Dedication

To Seth and Saul, my family

Preface

Tea Afar

Signals! Chirps, spikes, or slides?
In time, obscured, their content hides
from my perplexed, inquisitive eyes.
Of what are these recordings comprised?
Off I trudge through an unwieldy maze:
I measure the strength, the height, the phase.
Call the reserves! An army of Crays!
Surely they'll dispel this haze.
But crunching the entire night through
the thrashing's for naught: I've nothing new
but heaps of numbers and nary a clue.
An idea comes forth. My spirit's anew!
I'll call my friend, Jean Fourier!
He tries his best but cannot say
what has briefly passed through this way
and what is here for good, to stay.
So, I ask you all, despondently,
is there hope for lost souls like me?
For I have found I need to see
at once, both time and frequency.

— Seth Bowen Suppappola

The area of time–frequency signal processing is befitting to signals and systems that are *nonstationary* or *time-varying*, implying that their spectral characteristics change with time. The time-variation occurs naturally in many real-world events, including the propagation of waves in different mediums, biological or biomedical occurrences and seismic activities. It can also be found in technological areas including radar and sonar processing, wireless communications, mechanical use and military jamming operations. Some examples of time-varying signals include: waveforms from electrocardiograms, seizures in newborns, temporomandibular joint clicking sounds, nuclear magnetic resonance signals, postural sway responses, speech, music, sounds made by bats for echolocation, dolphin and whale whistles, bird songs, mechanical sounds like the changing of pitch in the whistle of a passing train, acoustic sounds from structure failure, impulse responses of wireless communication channels, wideband modulating or jamming waveforms, car-engine knocks, vibrational signals, underwater acoustic sounds and many more.

The Fourier transform is of limited use for the analysis of time-varying signals as it does not provide easily accessible information about the time localization of given signal frequencies. Suitable time-varying processing tools are time–frequency representations or TFRs. These are functions of both time and frequency as they map a one-dimensional signal in time to a two-dimensional function in time and frequency. TFRs are potentially capable of displaying the temporal localization of the spectral components of a signal along its instantaneous frequency or group delay. They can preserve important signal information such as energy and temporal or spectral moments. They can also satisfy many properties such as regularity or invertibility (to synthesize a signal without any loss of information) and covariance to signal transformations (such as time shifts, frequency shifts, nonlinear group delay shifts, and scale changes). The successful use of TFRs in many real-world applications has been demonstrated in wireless communications, speech and image processing, acoustics, pattern recognition, machine monitoring, detection, estimation, classification and sonar or radar processing.

This book is intended for a varied audience that includes practicing engineers in industry or government labs and university researchers. The sometimes very theoretical concepts in time–frequency signal processing can make it difficult for a novice practicing engineer to implement. Therefore, the objectives of this book are (a) to study some specific time–frequency methods from authors who are known authorities and leading experts on the matter; (b) to investigate the importance and signal processing capabilities of the methods as they apply to specific applications; and (c) to provide algorithms for these techniques to ensure ease of implementation and usage. This is achieved when the learning process on time–frequency methodologies is simplified by realizing in practice what is established in theory. Thus, for the successful use of time–frequency methods, some fundamental theory, together with some practical knowledge on what to use the method for, how to use it, when to use it, and why to use it, is essential.

The book consists of ten chapters that present a number of time–frequency applications for processing time-varying signals arising from underground to underwater, through the atmosphere and space. The chapters are outlined as follows:

- [Chapter 1](#), my contribution, is intended to provide a tutorial overview on time–frequency signal processing from both a theoretical and a practical perspective. It can also be viewed as an introduction to processing methods for time-varying signals so that novice and expert users alike would benefit from the book material contained herein. Specifically, the chapter surveys classical and recently developed TFRs, discusses their real-world applications, demonstrates the significance of matching a signal’s time-variation with different TFR attributes, and provides an extended bibliography on time–frequency applications.
- [Chapter 2](#), by A.R. Lindsey, L. Zhao and M.G. Amin, is on jamming interference excision by estimating the jammer’s instantaneous frequency and using time–frequency subspace projection techniques. The authors apply their technique to suppress interference in global positioning systems (GPS) for direct sequence spread spectrum wireless communications.

- [Chapter 3](#), by P. Loughlin and L. Cohen, concentrates on the theoretical development and implementation algorithms of positive TFRs that preserve important information on signal energy densities or marginals. Due to their nonnegativity, these TFRs can also provide correct instantaneous frequency as well as other conditional moment information.
- [Chapter 4](#), by D. Groutage, D. Bennink, P. Loughlin and L. Cohen, uses the aforementioned positive TFRs in real-world applications for the analysis and classification of acoustic scattering and propagation.
- [Chapter 5](#), by P. Flandrin, F. Auger and E. Chassande-Mottin, provides a practical perspective to the reassignment time–frequency method as a post-processing enhancement tool to increase the readability of existing TFRs. The authors provide efficient implementation algorithms and freeware software for their methods, and real-world application examples.
- [Chapter 6](#), by G. Matz and F. Hlawatsch, presents time–frequency-based designs of linear time-varying filters for nonstationary statistical processing of systems such as wireless communication channels. The authors provide discrete-time implementation algorithms for their techniques, and discuss possible online capabilities for real-time processing.
- [Chapter 7](#), by W. J. Williams, examines reduced interference distributions or RIDs that were designed to suppress cross terms of classical TFRs while preserving their desirable properties. The author provides discrete implementation and fast computational algorithms with software for the RIDs, and demonstrates their applicability in numerous real-world scenarios.
- [Chapter 8](#), by P. Steeghs, R. Baraniuk and J.E. Odegard, investigates the use of time–frequency methodologies such as adaptive techniques or TFRs of decomposed signals for processing seismic data or acoustic waveforms from the earth’s subsurface. The authors provide real-world examples for seismic analysis and attribute extraction.
- [Chapter 9](#), by B. Boashash and M. Mesbah, presents various time–frequency methods for the automatic detection and classification of EEG seizures in newborns. The authors compare various TFR-based seizure detectors and conclude that the B distribution approach is the most successful one, as it matches the time–frequency nature of these waveforms.
- [Chapter 10](#), by J. Droppo and L. Atlas, examines classification techniques based on the use of class-dependent time–frequency features. The features are obtained by designing kernels that can highly separate different classes of signals. The performance of these techniques is investigated using real-world classification applications for speech recognition.

Note that each chapter uses its own mathematical notation following the authors' adaptation from their publications. As a result, each chapter should be treated as its own entity with newly defined notation and abbreviations.

Antonia Papandreou-Suppappola
Arizona State University

Acknowledgments

I would like to thank all the authors of this notable collection of manuscripts. The contributors are some of the most distinguished researchers in the area of time-frequency signal processing, and their recognized expertise is testament to the high quality of the material contained in this work. To them, I express my gratitude for their subscribing to the notion of this book and readily agreeing to participate. In alphabetical order, these experts are Profs. Moeness Amin, Les Atlas, Richard Baraniuk, Boualem Boashash, Leon Cohen, Patrick Flandrin, Franz Hlawatsch, Patrick Loughlin, and William J. Williams. I would particularly like to express my gratitude to Prof. Alexander D. Poularikas, editor of the CRC Press Electrical Engineering Signal Processing Series, for his invitation to edit this book and his valuable advice and guidance throughout its preparation. My thanks also to Ms. Nora Konopka, CRC Press acquisitions editor, who was very instrumental in bringing this publication into fruition. Furthermore, I offer my deep gratitude to Dr. Seth B. Suppappola for his scientific input and technical support, and for his valuable insight drawn from his experience as a practicing engineer in industry and for the Navy.

Editor

Antonia Papandreou-Suppappola received her Ph.D. in Electrical Engineering in May 1995 at the University of Rhode Island, where she then held a research faculty position with Navy support. She joined Arizona State University as an assistant professor in August 1999.

Dr. Papandreou-Suppappola's current research interests are in the areas of Time-Frequency Signal Processing, Signal Processing for Wireless Communications, Integrated Sensing and Processing, and Estimation and Detection theory. Her extensive publication record includes more than 50 refereed journal articles, book chapters, and conference papers. She has received research funding from numerous agencies including the National Science Foundation (NSNF), the Defense Advanced Research Projects Agency (DARPA), the Office of Naval Research (ONR), and the Naval Undersea Warfare Center (NUWC), and is a recipient of the NSF CAREER Award. She is a member of the IEE Signal Processing, Communications, and Women in Engineering societies.

Contributors

Moeness Amin Department of Electrical and Computer Engineering, Villanova University, Villanova, Pennsylvania

Les Atlas Department of Electrical Engineering, University of Washington, Seattle, Washington

F. Auger Université de Nantes, GE44-IUT de Saint Nazaire, France

Richard G. Baraniuk Department of Electrical and Computer Engineering, Rice University, Houston, Texas

David Bennink ManTech System Engineering Co., Bremerton, Washington

Boualem Boashash Signal Processing Research Center, Queensland University of Technology Brisbane, Queensland, Australia

E. Chassande-Mottin Observatoire de la Cote d'Azur Equipe ILGA Nice, France

Leon Cohen Department of Physics and Astronomy, Hunter College, City University of New York, New York

James Droppo Microsoft Research, Redmond, Washington

Patrick Flandrin Laboratoire de Physique, École Normale Supérieure de Lyon, Lyon, France

Dale Groutage Naval Surface Warfare Center, Bremerton, Washington

Franz Hlawatsch Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Vienna, Austria

Alan R. Lindsey Air Force Research Lab/IFGC, Rome, New York

Patrick Loughlin University of Pittsburgh, Pittsburgh, Pennsylvania

Gerald Matz Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Vienna, Austria

Mostefa Mesbah Signal Processing Research Center, Queensland University of Technology Brisbane, Queensland, Australia

Jan Erik Odegard Department of Electrical and Computer Engineering, Rice University, Houston, Texas

Antonia Papandreou-Suppappola Telecommunications Research Center, Department of Electrical Engineering, Arizona State University, Tempe, Arizona

Philippe Steeghs TNO Physics and Electronics Laboratory, The Hague, Netherlands

William J. Williams University of Michigan, Ann Arbor, Michigan

Liang Zhao Department of Electrical and Computer Engineering, Villanova University, Villanova, Pennsylvania

Contents

- 1 Time-Frequency Processing: Tutorial on Principles and Practice.....1**
Antonia Papandreou-Suppappola
- 2 Interference Excision via Time-Frequency Distributions85**
Alan R. Lindsey, Liang Zhao and Moeness Amin
- 3 Positive Time-Frequency Distributions121**
Patrick Loughlin and Leon Cohen
- 4 Positive Time-Frequency Distributions and Acoustic Echoes.....163**
Dale Groutage, David Bennink, Patrick Loughlin and Leon Cohen
- 5 Time-Frequency Reassignment: From Principles to Algorithms179**
Patrick Flandrin, F. Auger and E. Chassande-Mottin
- 6 Linear Time-Frequency Filters: On-line Algorithms and Applications.....205**
Gerald Matz and Franz Hlawatsch
- 7 Discrete Reduced Interference Distributions273**
William J. Williams
- 8 Time-Frequency Analysis of Seismic Reflection Data307**
Philippe Steeghs, Richard G. Baraniuk and Jan Erik Odegard
- 9 Time-Frequency Methodology for Newborn EEG Seizure Detection339**
Boualem Boashash and Mostefa Mesbah
- 10 Quadratic Time-Frequency Features for Speech Recognition371**
James Droppo and Les Atlas

1

Time–Frequency Processing: Tutorial on Principles and Practice

Antonia Papandreou-Suppappola

Arizona State University

CONTENTS

1.1	Introduction	1
1.2	Time-varying Signals and their Representation	8
1.3	Quadratic Time–Frequency Representations	16
1.4	Matched Time–Frequency Processing for Various Applications	33
1.5	Applications in Time–Frequency Signal Processing	40
1.6	Concluding Remarks	44
	Acknowledgments	45
	References	46
	Appendix A: Acronyms in Alphabetical Order	82
	Appendix B: Mathematical Notation in Alphabetical Order	84

1.1 Introduction

The area of signal processing is founded on a rigorous mathematical exposition that is important for a profound understanding of the subject. The application of these mathematical concepts and techniques is necessitated by the continuous developments in many technologically advanced fields that require the processing of signals to extract important information. Signals convey information to represent measured streams of real-world application-dependent data such as remote-sensing satellite waveforms or seismic waves. For a practical application, a signal can be processed in a multitude of ways to extract specific information that cannot easily be obtained in the time domain. The processing of such signals forms the basis of many applications including analysis, synthesis, filtering, characterization or modeling, modulation, detection, estimation, classification, suppression, cancellation, equalization, coding and synchronization. A classical tool to accommodate this processing is the Fourier

transform (FT)* that is widely used to extract frequency information from the time domain signal. However, although successful in a wide range of applications, Fourier theory often possesses intrinsic limitations that depend on the signal to be processed.

1.1.1 Demand for time–frequency processing techniques

The purpose of this tutorial is to aid many signal processing practitioners to comprehend, utilize, conjecture and prove useful the theory on extracting information from signals that are nonstationary or time varying (TV). These signals have frequency content and properties that change with time. This class of signals is very common in real-world occurrences and, as such, it is very important to be able to process the signals as accurately as possible. TV signals include the following: the impulse response of a wireless communications channel, radar and sonar acoustic waves, seismic acoustic waves, biomedical signals such as the electrocardiogram (ECG) or neonatal seizures, biological signals such as bat or dolphin echolocation sounds, vocals in speech, notes in music, engine noise, shock waves in fault structures and jamming interference signals.

The FT does provide the overall frequency information present in a signal. However, it is of limited use for the analysis of TV signals because it does not provide easily accessible information about the signal spectral localization over short periods of time. Some suitable processing tools for these signals are transformations that provide information about the time–frequency (TF) content of the signals. The transformations can be one-dimensional (1-D) TV transforms (TWTs) with basis functions having joint TF characteristics. Specifically, the TWTs could match the instantaneous frequency (IF) of a signal and be functions of the rate of change of the IF. A TWT example is the linear matched signal transform (LMST) that enables the analysis of linear frequency-modulated (FM) chirp signals by presenting them as localized peaks in the linear FM (LFM) rate domain [1, 2].

These transformations can also be two-dimensional (2-D) time–frequency representations (TFRs) of actual time and frequency. The concept of processing in the TF domain dates as early as 1932 when Wigner introduced the Wigner distribution (WD) in the context of quantum mechanics as a function of position and momentum [3]. The WD was reinvented by Ville in 1948 in the context of signals as a function of time and frequency [4]. In 1946, the sound spectrograph was introduced to represent speech signals for visual interpretation [5]. This relates to the short-time FT (or its squared magnitude, the spectrogram) because it assumes that the signal is stationary over a small duration of time such that the FT of consecutive time segments provides an adequate signal representation [6]. In 1966, Cohen provided an overall class of TFRs based on the WD and emphasized its importance in signal processing [7]. The WD [3, 4, 7–12] and the spectrogram [6, 13] are most often used for TF analysis due to their simplicity in theory and implementation. Also, many other TFRs have been proposed as referenced in various books on TF analysis [14–38].

*A summary of the acronyms used in this chapter can be found in Appendix A.

Both 1-D and 2-D transformations can provide important information about TV signals. Note, however, that an accurate signal representation is only obtained when the transformation matches the TF characteristics of the signal.

This chapter attempts to provide a response to the multifaceted question on the application of TF: *Which technique does one use and why? How does one use and implement the technique?* Before describing the mathematical methodology of TF techniques and attempting to answer this question, we demonstrate their significance in a specific application example simply by mentioning the tool and providing its desirable properties for use in this application. This emphasizes the notion of carrying out in practice what is advocated in principle.

1.1.2 Demonstration: time–frequency techniques for wireless communication applications

1.1.2.1 Challenges in wireless communications

With the wireless revolution, the communications research area is rapidly growing and confronting many challenges in advanced system design to improve the functionality and performance of mobile technology. These challenges include the escalating market demands for new and improved wireless applications involving more complex and dynamic channels. Such channels, however, suffer from numerous physical impairments and require higher data rates to provide universal access and channel bandwidths to extend over a wider range of frequencies. Another challenge is the need to improve multiple access techniques to accommodate more users in the available bandwidth while reducing undesirable channel distortions. As the wireless channel is TV due to the constantly changing physical characteristics of the medium, these demands introduce various distortions such as increased levels of multiple access interference and multipath and Doppler spreading [39–41]. Various 1-D processing methods have been proposed to counteract these problems, but they mostly work relatively well for low data rate transmissions. These challenges give rise to the need for TV signal processing techniques.

TF methods have been introduced to provide advanced methodologies for improvement of wireless technology. Specifically, TFRs were used to suppress TV jamming interference in military wireless communications [2,42–50]; to provide pulse-shaping techniques in multicarrier communications [51–54]; to provide TF-based transfer function characterizations of TV systems [55–60] without time-invariant approximations [39, 61]; to process TV Wiener filters [62, 63]; to enhance receiver performance using TF-based estimation and detection techniques [64–67]; to employ TV signals like LFM chirps in modulation, demodulation, spreading and equalization schemes [54, 68–74]; to characterize wireless systems and combat fading distortion effects achieving improved receiver performance [51, 57, 60, 66]; and to increase multipath and Doppler diversity in multiple access schemes [75–77].

1.1.2.2 Wideband interference suppression application

To illustrate the applicability of TFRs in real-world applications, we next show how they can be used to suppress wideband jamming interference in a direct sequence spread-spectrum (DSSS) wireless communications system [39, 78, 79]. In such a system, narrowband information is spread using a pseudonoise (PN) sequence before transmission over a slow fading, frequency-selective wireless channel. Because the transmission signal is now wideband due to the spreading, it is less susceptible to distortion when intentionally jammed with a narrowband interference signal, for example, in radar or sonar military applications. When the jamming interference power is high, additional suppression techniques need to be employed [80]. These signal processing techniques aim to suppress the interference from the received data by transforming it to a different domain where the characteristics of the interference are more evident than in the time domain. One possible method to suppress broadband LFM interference is by taking the WD of the interfered signal, 2-D masking the WD of the LFM interference and then using the WD least-squares synthesis technique to obtain an estimate of the interference [44]. The masking is possible because the WD of an LFM chirp is well defined over a linear support region in the TF plane with slope equal to the FM chirp rate. The estimated interference is then subtracted from the received data in the time domain before detection at the receiver. Another possible method for achieving the suppression of the LFM interference is by taking the LMST of the received signal, thresholding the LMST to remove the highly localized interference components and then taking the inverse LMST of the interference suppressed received signal [2, 81]. This method is effective because the LMST of an LFM chirp results in a peak in the FM rate domain.

1.1.2.2.1 Graphic demonstration Both the WD and the LMST procedures are demonstrated graphically. The DSSS system described earlier is shown in Figure 1.1 [78]. At the transmitter, the modulated information signal $b(t)$ is spread when multiplied with the PN sequence $p(t)$. By propagating through an additive white Gaussian noise (AWGN) channel, the product $m(t)$ is corrupted both by an LFM interference $i(t)$ and the zero-mean AWGN $n(t)$. Without the mitigation scheme, the corrupted signal $y(t)$ is the received signal $r(t)$. When $y(t)$ is processed with the WD mitigation

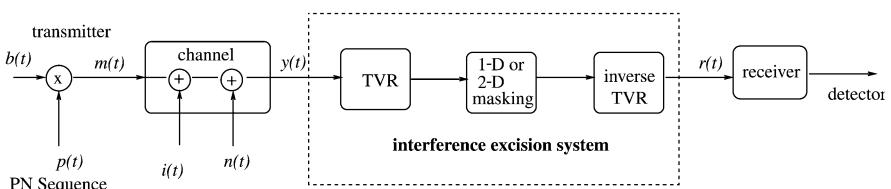
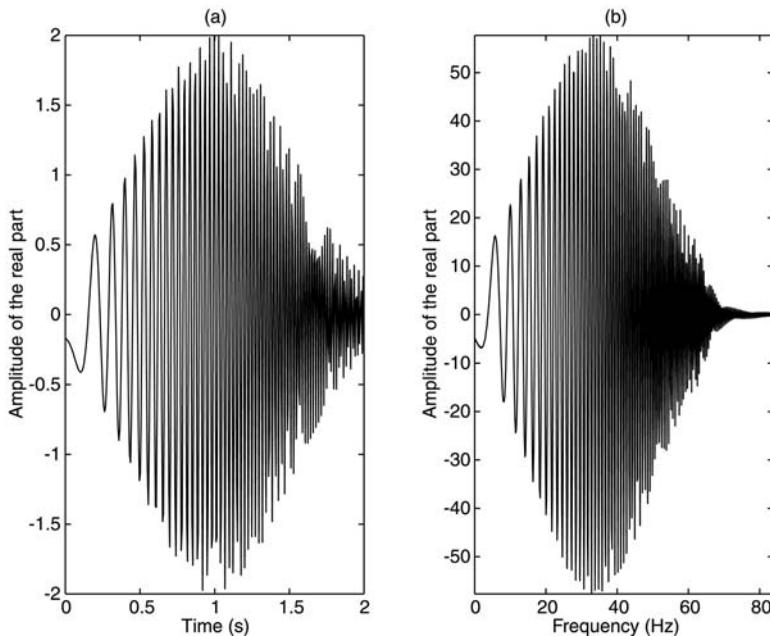


FIGURE 1.1

A DSSS interference suppression system using TV representations (TVRs) such as the 2-D WD or the 1-D LMST together with their inverse transformations.

**FIGURE 1.2**

An LFM chirp interference with FM rate 33.6 Hz^2 in (a) the time domain and (b) the frequency domain. The sampling frequency used was 168 Hz.

system, for example, to obtain an estimate of the interference $i_{\text{est}}(t)$, the interference suppressed signal is $r(t) = m(t) + n(t) + (i(t) - i_{\text{est}}(t))$.

Figure 1.2(a) and (b) shows the LFM interference in the time domain and its FT (which is also a chirp) in the frequency domain, respectively. The sampling frequency of the interference is 168 Hz and its chirp rate is 33.6 Hz^2 . Figure 1.3(a) provides the image plot of the WD of this LFM interference. Because the WD is a function of both time and frequency, an image plot can provide information on time from the horizontal axis and on frequency from the vertical axis. The darker TF regions on the image plot represent higher WD amplitudes. We can observe that frequency depends linearly on time as the WD of the LFM chirp is a line in the TF plane. This information, however, cannot be easily extracted from Figure 1.2(a) and (b). We also note that the LFM chirp is wideband but does not occupy the entire TF plane. On the other hand, the spread signal has noiselike characteristics and, together with noise, they extend throughout the TF plane. In Figure 1.3(c), we demonstrate the WD of the LFM chirp interference together with the transmitted spread data and the noise. Note that the jammer-to-signal ratio (JSR) is 2.5 dB, and the signal-to-noise ratio (SNR) is 25.3 dB. The TF mask used in estimating the interference is shown in Figure 1.3(b). It consists of ones around the TF region of the interference and is zero everywhere else. Note that the interference can be clearly seen in the TF domain. After the multiplication of the mask with the WD of the corrupted signal, followed by the inverse WD of

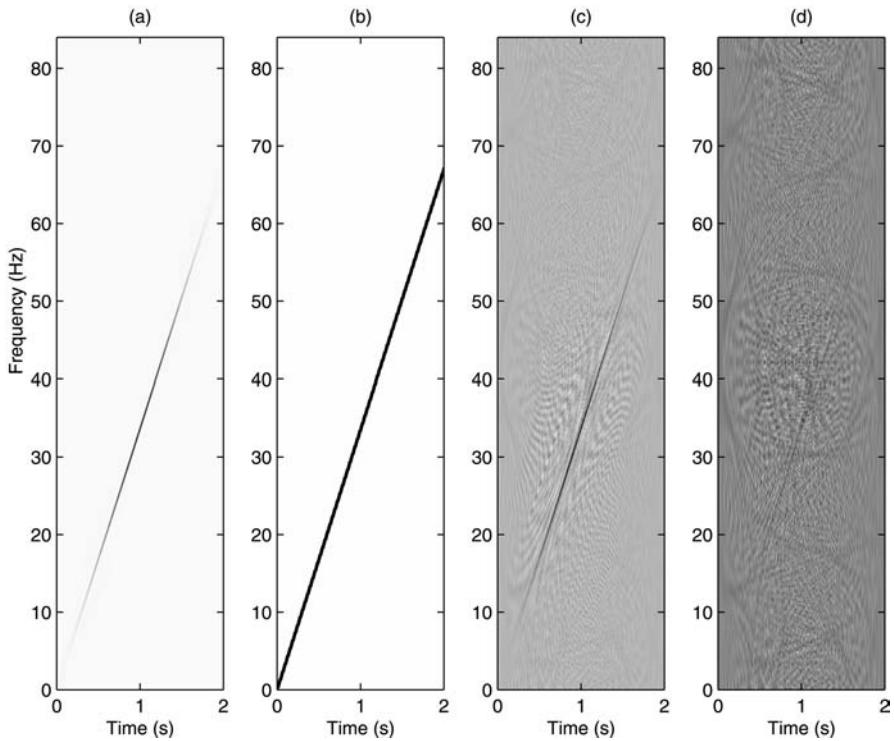


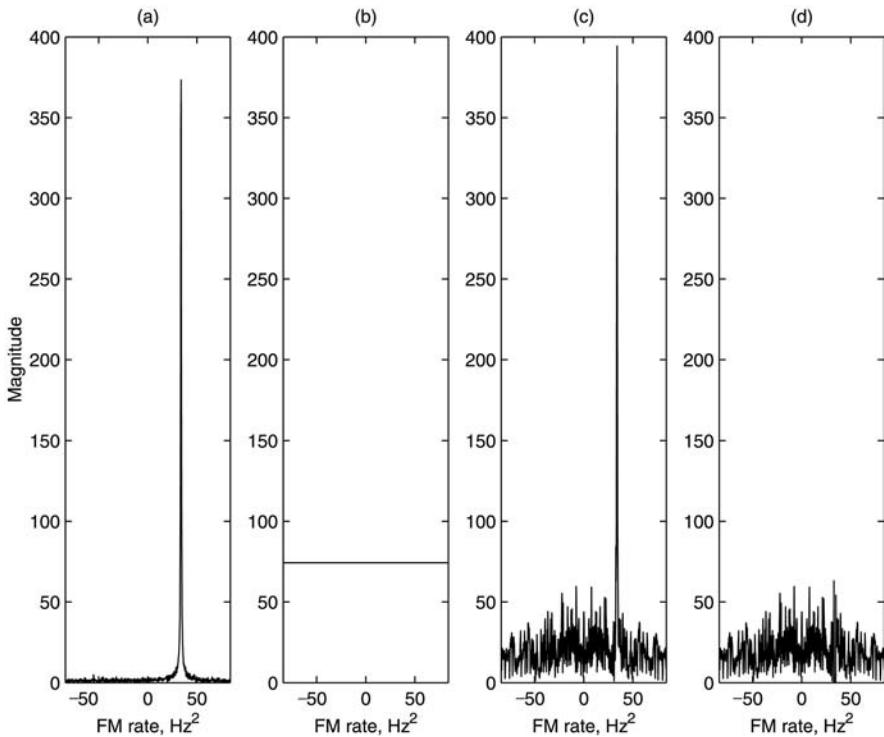
FIGURE 1.3

The WD of (a) the LFM chirp interference in Figure 1.2, (c) the interference added to the noisy spread information signal and (d) the noisy spread information after the estimated interference is suppressed. (b) The 2-D time-frequency (TF) mask used for the suppression.

the product, we obtain an estimate of the interference that we then subtract from the corrupted signal. It can be clearly seen that the WD of the interference suppressed received signal shown in Figure 1.3(d), does not contain the LFM chirp interference. Note that performance curves for this technique can be found in [44].

The same received data with the prescribed noise and interference is also processed using the LMST as shown in Figure 1.4. The 1-D LMST of the LFM interference is plotted in Figure 1.4(a). As expected, the transform is highly localized along the FM rate axis at the correct FM rate, 33.6 Hz^2 , of the interference. As it can be observed from the LMST of the interfered signal in Figure 1.4(c), the characteristics of the spread data are very different from those of the interference. The thresholding used to suppress the interference is given in Figure 1.4(b), and the LMST of the estimated interference suppressed signal is given in Figure 1.4(d). Just as with the WD, the LMST procedure also successfully reduces the interference, and performance curves as compared with the WD method are provided in [2, 49].

1.1.2.2.2 Implementation issues. The WD can be implemented using the MATLAB^R function `tfrwv.m` that is a part of the freeware Time-Frequency Toolbox

**FIGURE 1.4**

The LMST of (a) the LFM chirp interference in Figures 1.2 and 1.3, (c) the interference added to the noisy spread information signal and (d) the interference suppressed noisy spread information. (b) The 1-D thresholding used for the suppression.

(TFT) software [82]. This toolbox, created by F. Auger, P. Flandrin, P. Gonçalvès and O. Lemoine, contains functions for many other TFRs, and can be presently downloaded from <http://crttsn.univ-nantes.fr/~auger/tfb.html>. The TFT toolbox is continuously maintained and updated with new MATLAB versions. The algorithms for computing the WD synthesis and the LMST can be found in [83–85] and [49], respectively.

1.1.3 Importance of theoretical concepts

Thus far, we have not used any mathematical formulations to support the previously mentioned applications. The sometimes very theoretical concepts in TF signal processing are strenuous when a practicing engineer wants to implement a TFR that, based on its properties, is desirable for the particular application at hand. However, some fundamental theoretical knowledge is essential to successfully use the TFRs and obtain desirable information. A superficial understanding cannot allow the fine-tuning of relevant TFR parameters to achieve the best results for the application.

Thus, in processing a TV signal, we need to transform it to a different domain as well as choose the ideal tool in that domain to obtain better insight into important signal characteristics.

It is natural to ask why the WD was chosen in the preceding simulations and not another TFR such as the spectrogram. In this case, we require a TFR that is invertible so that the interference signal can be recovered using synthesis methods. In addition, the TFR must be highly localized along the TF structure of the jamming interference. The spectrogram is a smoothed WD that depends on windowing techniques, and thus does not satisfy these two requirements. On the other hand, the WD is invertible [83–85] and is very highly localized for signals with constant or linear TF structures. For example, the WD of an LFM chirp can be computed analytically to yield a Dirac delta function that is nonzero only for values of frequency that are linearly related to time. Note that the WD also cannot be applied successfully when the jamming interference consists of multiple LFM chirps. Theoretically, it can be shown that the WD includes oscillatory cross terms (CTs) between the individual WDs (or autoterms) of any two signals [12, 86]. These CTs are important because they contain energy information about the signal, and thus need to be captured by the 2-D mask. However, it is difficult to design a TF mask to encompass both autoterms and CTs. The LMST, on the other hand, is a linear transform and, as a result, does not suffer from CTs. The LMST of multiple LFM chirps simply results in multiple peaks (at the different FM rates) that can all be thresholded simultaneously. Furthermore, suppose that the wireless application is in a dispersive medium and the interference is no longer linear in the TF plane. In that case, neither the WD nor the LMST can effectively reduce the interference and other processing tools need to be investigated based on the properties they satisfy.

Because it is evident that some theory is necessary for notable processing, we first introduce various types of representations of TV signals in Section 1.2, and then concentrate on quadratic TFRs in Section 1.3. Section 1.4 discusses the importance of matching TFRs with the signals to be processed, and applications of TFRs in various research areas are provided in Section 1.5. Finally, Section 1.6 provides some overall concluding remarks on the use of TF processing techniques as a response to the question posed in the last paragraph of Section 1.1.1.

1.2 Time-varying Signals and their Representation

1.2.1 Time-varying signal models

TV signals (signals with time-dependent spectral content) are very commonly found in nature or are generated for technological use. In both cases, it is important to understand the characteristics of these signals, and process them with the appropriate tools to extract relevant information. For example, it is desirable to model the signals used when bats echolocate or dolphins whistle; to derive an accurate TV impulse

response that can improve the performance of a wireless communications channel; and to design ideal signals for use in radar or sonar applications.

A general model for any TV signal represented in the time domain is given by[†] $x(t) = a(t) e^{j2\pi b \eta(\frac{t}{t_r})}$ where $a(t)$ denotes the TV amplitude of the signal and $\eta(t/t_r)$ denotes its TV phase function. In the frequency domain, a TV spectrum can be represented as $G(f) = R(f) e^{-j2\pi c \xi(\frac{f}{f_r})}$, with amplitude spectrum $R(f)$ and phase spectrum $\xi(f/f_r)$. Note that $a(t)$ and $\eta(t/t_r)$ are not the FT pairs of $R(f)$ and $\xi(f/f_r)$, respectively. Here, t_r and f_r correspond to positive time and frequency reference values, respectively, for normalization and unit precision. For example, t_r can be an arbitrary point in time that corresponds to the temporal origin.

Two examples of TV signals are the time domain FM waveform and its dual[‡] frequency domain time-modulated (TM)[§] waveform. These are defined, respectively, as:

$$x(t) = \sqrt{|\nu(t)|} e^{j2\pi b \eta(t/t_r)} \quad (1.1)$$

$$G(f) = \sqrt{|\tau(f)|} e^{-j2\pi c \xi(f/f_r)} \quad (1.2)$$

Here, $\nu(t) = \frac{d}{dt} \eta(t/t_r)$ is the IF of the FM signal, and $\tau(f) = \frac{d}{df} \xi(f/f_r)$ is the group delay (GD) of the TM signal. An example of an FM signal is the LFM chirp interference (encountered in Section 1.1.2) that is defined as in Equation (1.1), $x(t) = \sqrt{2|t|/t_r^2} e^{j2\pi b \text{sgn}(t)|t/t_r|^2}$, with phase function $\eta(t/t_r) = \text{sgn}(t)|t/t_r|^2$. Here, $\text{sgn}(t)$ is the sign (± 1) of t , and b is the FM rate of the chirp. LFM chirps are important in many applications. For example, they are used for radar or sonar echo ranging as they provide more accurate range and Doppler resolution estimates than tones [87]. They are also used in TV modulation schemes in frequency-hopped code division multiple access (FH-CDMA) systems to improve performance [54, 70, 71, 88, 89] over frequency-shift keying (FSK) signals [78, 79, 90]. This is because chirps are bandwidth efficient and possess inherent immunity against multipath and Doppler fading [88, 89]. Another TV waveform is the hyperbolic TM signal that is defined in Equation (1.2) with GD $\tau(f) = 1/f$ and logarithmic phase spectrum $\xi(f/f_r) = \ln(f/f_r)$ for $f > 0$. This is a dispersive signal by nature because its high-frequency components are delayed in time by a larger amount than its low-frequency components. Hyperbolic TM signals are wideband Doppler-invariant signals (invariant to expansions or compressions) similar to the signals used by bats for echolocation [91]. Note that for both LFM and hyperbolic TM signals, the specific TV amplitude ensures that the signals are orthonormal [92].

The IF curve of FM signals and the GD curve of TM signals can provide important information about the time variation of the signals. Thus, it is important to design

[†]A summary of our adapted notation is provided in Appendix B.

[‡]By dual, we imply interchanging the signal x in time with X in frequency and the time t with the frequency f .

[§]Note that by a TM signal, we imply that the signal group delay changes with frequency. By using duality, an FM signal has time-dependent IF.

transformations that can provide high localization along these curves. For example, because GD is a measure of the time delay introduced in each sinusoidal component of the signal at frequency f , an ideal TFR should preserve this change in GD or frequency-dependent time shift $\tau(f)$ [1, 93]. These TV signal transformations may be 1-D such as the IF and the GD, or 2-D such as the linear, quadratic, adaptive and expansion-based TFRs. In addition, higher order TFRs also exist such as the *positive Cohen distributions* [94–96], the third-order WD [97, 98], the higher order WD [99] or other higher order TFRs [100–105]. Next, we consider different signal transformations and investigate their applicability in processing TV signals.

1.2.2 Fourier signal transformation

The classical FT analysis tool is successful in analyzing signals by providing their spectral content. It is defined in the frequency domain as $X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt$, thus representing the signal $x(t)$ as a weighted sum of sinusoids with different frequencies [106]. The FT provides the frequencies whose presence in the time domain signal is significant. However, it does not demonstrate temporal localization of the frequencies, and thus does not provide at which times these frequencies occur. Thus, TV signals cannot be successfully analyzed using classical Fourier theory. This is demonstrated by the FT of the LFM chirp in [Figure 1.2\(b\)](#) that does not provide straightforward information of the signal linear dependence of frequency on time.

1.2.3 Time-varying signal transformations of one output variable

1.2.3.1 Instantaneous frequency and group delay transforms

The IF and GD are 1-D signal representations that can properly provide information on a single frequency (time) component that occurred at a particular single time (frequency) point. Specifically, they provide meaningful information for monocomponent[¶] TV signals that can be expressed in the form of Equation (1.1). The IF, $\nu_x(t) = b\nu(t) = b\frac{d}{dt}\eta(t/t_r)$, of the signal $x(t)$ in Equation (1.1) can be obtained as the derivative of the instantaneous phase that is given as $\eta(t/t_r) = \frac{1}{2\pi b} \arg\{x(t)\}$ [4, 106]. Here, $\arg\{Ae^{j\theta}\} = \theta$ for A real. The IF assumes that at each time only a single frequency exists. The GD, $\tau_G(f) = c\tau(f) = c\frac{d}{df}\xi(f/f_r)$, of the signal spectrum $G(f)$ in Equation (1.2) is the derivative of the phase spectrum $\xi(f/f_r) = -\frac{1}{2\pi c} \arg\{G(f)\}$. The GD can be interpreted as the time delay or distortion introduced by a linear time-invariant system at each frequency [106, 107]. Specifically, consider a signal propagating through a system with dispersive characteristics (nonlinear GD) whose spectrum is $G(f) = R(f)e^{-j2\pi c_1 \xi_1(f/f_r)}$. If $G(f)$ is passed through an all-pass, dispersive system with output $Y(f) = e^{-j2\pi c_2 \xi_2(f/f_r)} G(f)$, then the GD change is proportional to the derivative of $\xi_2(f/f_r)$. Thus, the GD of the output signal is given by $\tau_Y(f) = [c_2 \frac{d}{df} \xi_2(f/f_r) - c_1 \frac{d}{df} \xi_1(f/f_r)]$. When the system's response is

[¶]By monocomponent, we mean that the signal cannot be further simplified to additional terms in the form of Equation (1.1) except possibly in infinite signal expansions.

matched to the input signal's TF characteristics, that is, if $\xi_1 = \xi_2$, then the GD is $[(c_2 - c_1) \frac{d}{df} \xi_1(f/f_r)]$. Thus, the GD of dispersive systems can be used to estimate the system response by estimating the shift c_2 of the input parameter c_1 .

Both the IF and the GD signal transforms correspond to single curves in the TF plane that provide only one frequency per time point. When a signal consists of two or more components, then the IF or the GD do not provide the correct spectral variation. For example, consider a signal $x(t) = x_1(t) + x_2(t)$ where $x_1(t)$ and $x_2(t)$ are LFM chirps. The first LFM chirp is the same as the one used as interference in Figure 1.2, and the other chirp has a different initial frequency and a chirp rate of -16.8 Hz^2 . Figure 1.5(a) shows the sum of the IFs, $\nu_{x_1}(t) + \nu_{x_2}(t)$, of the two LFM chirps. As expected, we obtain a highly localized representation of the linear dependence of frequency on time for each chirp. This is the representation that we would like to have such that at time point t , we expect to observe two different frequencies. However, in general, we do not have the individual terms separately but superimposed together. The IF, $\nu_x(t) = \nu_{x_1+x_2}(t)$, of the signal sum $x(t)$ is shown

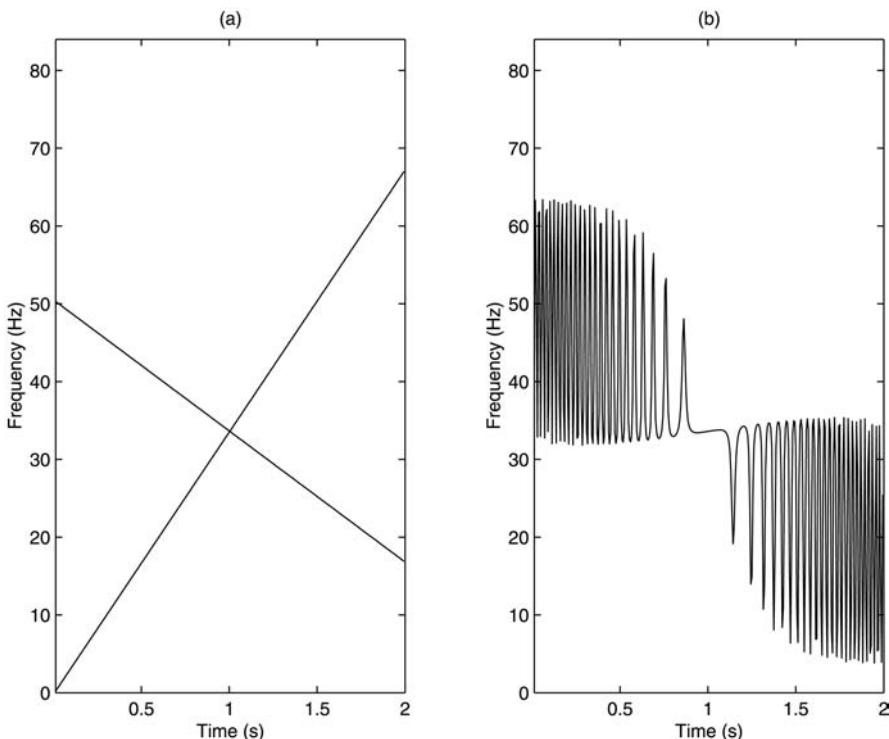


FIGURE 1.5

The analysis signal is the sum of two LFM chirps with FM rates 33.6 Hz^2 , as in Figure 1.2(a), and -16.8 Hz^2 . (a) The sum of the IFs of each LFM chirp, and (b) the IF of the sum of the two LFM chirps. The sampling frequency used for both signals is 168 Hz.

in [Figure 1.5\(b\)](#). We can see that the transform fails to provide the correct time-dependent frequency localization of $x(t)$. Note also that the IF and GD transforms do not preserve the energy information of the analysis signal. Thus, the IF and the GD cannot successfully represent multicomponent signals. However, they can provide ideal representations for monocomponent signals with nonlinear TF structures.

1.2.3.2 Matched signal transforms

Just as the FT localizes a sinusoid at its fundamental frequency, the matched signal transform (MST) was designed to localize TV signals with nonlinear phase at their TM rate [1, 2, 49, 92, 93, 108, 109]. This localization is a result of choosing the basis nonlinear characteristic function $\xi(f/f_r)$ of the transform to match, in TF structure, the phase function of the TV signal to be processed. We define the MST of a signal spectrum $G(f)$ as:

$$\mathfrak{N}_G^{(\xi)}(c) = \int_{f \in \varphi} G(f) \sqrt{|\tau(f)|} e^{j2\pi c \xi(f/f_r)} df \quad (1.3)$$

where φ contains the values of f in the domain of $\xi(f/f_r)$, and $\tau(f) = \frac{d}{df} \xi(f/f_r)$. We obtain different MSTs for analyzing different types of signals by choosing a matched, differentiable and one-to-one function $\xi(f/f_r)$. Examples of the MST include the *Mellin* or *hyperbolic* transform that is matched to hyperbolic TF characteristics $\tau_G(f) = 1/f$ [109, 110], and has been used to analyze Doppler-invariant signals; the κ th *power transform* matched to signals with GD $\tau(f) = \frac{\kappa}{f_r} |f/f_r|^{\kappa-1}$, $\kappa \neq 0$ [92, 93, 111]; and the exponential transform that is matched to exponential GD structures $\tau(f) = \frac{\kappa}{f_r} e^{\kappa f/f_r}$ [93, 112]. The MST localization property states that the MST of the TM signal in [Equation \(1.2\)](#) provides a highly localized representation because it results in a Dirac delta function at the TM rate of the signal. Specifically:

$$G(f) = \sqrt{|\tau(f)|} e^{-j2\pi c_0 \xi(\frac{f}{f_r})} \Rightarrow \mathfrak{N}_G^{(\xi)}(c) = \delta(c - c_0)$$

provided that the phase spectrum of the signal is equal to the characteristic function $\xi(f/f_r)$ of the MST.

The MST can be computed efficiently using the fast FT because it can be shown to be equivalent to the inverse FT of a nonlinearly sampled or warped version of the signal. Specifically, [Equation \(1.3\)](#) can also be expressed as $\mathfrak{N}_G^{(\xi)}(c) = \int_{f \in \hbar} Z(f) e^{j2\pi(c/f_r)f} df / \sqrt{f_r}$ where $Z(f) = (\mathcal{W}_\xi G)(f) = |f_r \tau(f_r \xi^{-1}(f/f_r))|^{-1/2} G(f_r \xi^{-1}(f/f_r))$ is the warped signal, \mathcal{W}_ξ is the warping operator, $\xi^{-1}(f/f_r)$ is the inverse function of $\xi(f/f_r)$ and \hbar denotes the domain of $\xi^{-1}(f/f_r)$.

Note that we can also define a dual MST, $\Upsilon_x^{(\eta)}(b) = \int_{t \in \varphi} x(t) \sqrt{|\nu(t)|} e^{-j2\pi b \eta(t/t_r)} dt$ [2, 108], that localizes a time domain signal $x(t)$ with phase function $\eta(t/t_r)$ and IF $\nu(t) = \frac{d}{dt} \eta(t/t_r)$. Here, φ contains the values of t in the domain of the MST characteristic function $\eta(t/t_r)$. In particular, the dual MST of the FM signal $x(t) = \sqrt{|\nu(t)|} e^{j2\pi b_0 \eta(\frac{t}{t_r})}$ in [Equation \(1.1\)](#) is a Dirac delta function $\Upsilon_x^{(\eta)}(b) = \delta(b - b_0)$. [Table 1.1](#) provides some dual MST transforms based on their

TABLE 1.1Examples of the Dual MST Based on Its Characteristic Function $\eta(t/t_r)$

Characteristic Function	Definition	Transform
$\eta(t/t_r) = t/t_r$	$\Upsilon_x(b) = X(b/t_r)$ $= \int_{-\infty}^{\infty} x(t) e^{-j2\pi \frac{b}{t_r} t} dt$	Fourier
$\eta(t/t_r) = \ln(t/t_r),$ $t > 0$	$\Upsilon_x(b) = \int_0^{\infty} x(t) \frac{1}{\sqrt{t}} \times e^{-j2\pi b \ln(t/t_r)} dt$	Dual Mellin
$\eta(t/t_r) = \text{sgn}(t) t/t_r ^2$	$\Upsilon_x(b) = \int_{-\infty}^{\infty} x(t) \sqrt{ 2t/t_r^2 } \times e^{-j2\pi b \text{sgn}(t) t/t_r ^2} dt$	Linear
$\eta(t/t_r) = \text{sgn}(t) t/t_r ^{\kappa},$ $\kappa \neq 0$	$\Upsilon_x(b) = \int_{-\infty}^{\infty} x(t) \sqrt{ \frac{\kappa}{t_r} \left \frac{t}{t_r}\right ^{\kappa-1}} \times e^{-j2\pi b \text{sgn}(t) t/t_r ^{\kappa}} dt$	Power
$\eta(t/t_r) = e^{(\kappa t/t_r)},$ $\kappa \neq 0$	$\Upsilon_x(b) = \int_{-\infty}^{\infty} x(t) \sqrt{ \frac{\kappa}{t_r} } e^{(\kappa t/t_r)} \times e^{-j2\pi b e^{(\kappa t/t_r)}} dt$	Exponential

characteristic function $\eta(t/t_r)$. An example of the dual MST is the linear MST or LMST in Section 1.1.2.2 [2, 108] with $\nu(t) = \frac{2}{t_r} |t/t_r|$. The LMST, defined as $\Upsilon_x(b) = \int_{-\infty}^{\infty} x(t) \sqrt{|2t/t_r^2|} e^{-j2\pi b \text{sgn}(t)|t/t_r|^2} dt$, provides localization for an LFM chirp at its FM rate, that is:

$$x(t) = \sqrt{|2t/t_r^2|} e^{j2\pi b_0 \text{sgn}(t)|t/t_r|^2} \Rightarrow \Upsilon_x(b) = \delta(b - b_0) \quad (1.4)$$

This is an important result for detection and estimation applications [2, 108], especially for radar and sonar where LFM chirps are used. In [113, 114], a chirp FT was defined in the continuous time domain that is different from the LMST. Specifically, this transform does not provide ideal localization as in Equation (1.4) because it is not unitary [115]. The discrete version of the chirp FT, however, was shown to be a constant value when it matched the characteristics of a discrete chirp signal [113].

1.2.4 Time–frequency representations of two output variables

TFRs demonstrate how the frequency content of a signal changes with time, and are thus ideal for processing TV signals. They are designed to map a 1-D signal $x(t)$ with FT $X(f)$ into a 2-D function of time and frequency denoted equivalently as $T_x(t, f)$

or $T_X(t, f)$. TFRs can manifest temporal localization of spectral components [7, 92, 116–119] in a similar way that a musical score represents the notes to be played at different times. As a result, they have been successfully used in many TV signal processing application areas such as wireless communications, speech, radar, sonar, acoustic and image processing. The theoretical development of TFRs can be found in many manuscripts including [14–18, 29–34, 92, 116, 117, 120–123].

As different signals are matched to specific TFR properties in a particular application, their analysis requires different TFRs. TFRs have been classified in a variety of methods based on diverse criteria. One classification is based on the specific TFR dependence on the processing signal as outlined next.

1.2.4.1 Linear TFRs

A linear TFR $T_x(t, f)$ satisfies the linearity superposition principle that states that if $x(t) = \alpha x_1(t) + \beta x_2(t)$ is a linear combination of $x_1(t)$ and $x_2(t)$, then the TFR of the sum must satisfy $T_x(t, f) = \alpha T_{x_1}(t, f) + \beta T_{x_2}(t, f)$ where α and β are complex coefficients. For this principle to hold, the TFR must depend only on the analysis signal and some signal independent window function h_T , that is, $T_x(t, f) = \int_{-\infty}^{\infty} x(\tau) h_T(\tau; t, f) d\tau$. Two linear TFRs that have been used in many applications are the short-time Fourier transform (STFT) [124–130] and the wavelet transform [29, 31, 32, 35, 121, 131–134]. The STFT is the only linear TFR that preserves both time and frequency shifts on the analysis signal, an important property for speech, image processing and filterbank decoding applications. The wavelet transform preserves scale changes (compressions or expansions) on the signal, an important property for multiresolution analysis applications such as detection of singularities or edges in images. Note that as these TFRs are based on windowing techniques, their TF resolution depends on the choice of window characteristics.

1.2.4.2 Quadratic TFRs

Any quadratic TFR (QTFR) can be expressed as $T_x(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t_1) x^*(t_2) K_T(t_1, t_2; t, f) dt_1 dt_2$ where K_T is a signal-independent function that characterizes the QTFR. QTFRs satisfy the quadratic superposition principle as a QTFR of $x(t) = \alpha x_1(t) + \beta x_2(t)$ satisfies $T_x(t, f) = |\alpha|^2 T_{x_1}(t, f) + |\beta|^2 T_{x_2}(t, f) + 2\Re[\alpha\beta^* T_{x_1, x_2}(t, f)]$. The term T_{x_1, x_2} is the cross QTFR or CT of $x_1(t)$ and $x_2(t)$, and $\Re[\alpha]$ denotes the real part of α . For QTFRs, windowing techniques are not required because the objective is to form energy distributions so that the signal energy, also a quadratic representation, can be distributed in the TF plane. However, windowing techniques are often used to suppress CTs that may impede processing because they are oscillatory. Two important QTFRs include the WD [3, 4, 8–12] and its [4] smoothed version, the spectrogram [6, 13]. Note that QTFRs often overcome the TF resolution problem that limits the linear TFRs [116].

1.2.4.3 Adaptive TFRs

Some TFRs were proposed to adapt to the signal TF changes. In particular, an adaptive TFR can be obtained by estimating some pertinent parameters of a signal-dependent function at different time intervals. Such TFRs provide highly localized representations without suffering from QTFR CTs. The trade-off is that these TFRs may not satisfy some desirable properties such as energy preservation. Examples of adaptive TFRs include the high resolution TFR [135], the signal-adaptive optimal-kernel TFRs [136, 137], the optimal radially Gaussian TFR [138] and Cohen’s nonnegative distribution [94]. Reassigned TFRs also adapt to the signal by employing other QTFRs of the signal such as the spectrogram, the WD or the scalogram [139, 140]. For example, the reassigned spectrogram moves spectrogram values toward local centroids of the WD to obtain a very high localized TFR without any CTs.

1.2.4.4 TFRs based on signal expansions

Some TFRs are obtained from the expansion of a TV signal into an infinite number of TF-shifted versions of a weighted elementary atom. For example, for the Gabor expansion, the atom is chosen to be a Gaussian signal because it is the most concentrated signal in both time and frequency [33, 34, 141–147]. Thus, the signal is expanded as a weighted summation of TF-shifted Gaussians where the weight coefficients determine how strong a component the signal has at that TF point. Another example uses the matching pursuit iterative algorithm that decomposes a signal into a linear combination of elementary functions selected from a complete and redundant dictionary. This is achieved using successive approximations of the signal with orthogonal projections on dictionary elements [148]. A dictionary of Gaussian atoms with all possible TF shifts and scale changes was used in [148], and a quadratic TFR was obtained by summing the WDs of all selected elements in the expansion. This modified WD is free of CTs, and it can be shown to preserve signal energy, TF shifts and scale changes [148–151]. Note that parsimonious expansions based on the matching pursuit principle but with matched dictionary elements were also proposed in [152–155].

1.2.5 Time–frequency representation implementation

Computational algorithms for many of these TFRs have been designed and disseminated on the World Wide Web (WWW). This was important in demonstrating that these signal representations can be very successful tools for processing both signals and systems in real-life applications. Some TFR software can be presently found at the following WWW sites:

- TFT toolbox, <http://crttsn.univ-nantes.fr/~auger/tftb.html> by F. Auger, P. Flan-drin, P. Gonçalvès and O. Lemoine
- Time–frequency signal analysis (TFSA), <http://www.sprc.qut.edu.au/~sprc/tfsa/index.html> by B. Boashash
- Rice DSP software, <http://www-dsp.rice.edu/software/TFA/> by R. Baraniuk

- Discrete time–frequency laboratory (DTFL), <http://www.eecs.umich.edu/~wjjw/dtfl.html> by W. J. Williams
- Time–frequency analysis software, <http://www.mathworks.com/matlabcentral/fileexchange/Category.jsp> under *DiscretTFRs* by J. O’Neill
- Joint time–frequency analysis (JTFA), http://www.ni.com/analysis/joint_time.htm by S. Qian and D. Chen, National Instruments Corporation
- Wavelab, <http://www-stat.stanford.edu/~wavelab/> by D. Donoho
- Time–frequency toolbox in C, <http://www.sigproc.eng.cam.ac.uk/~md283/toolbox/Cifibeng.html> by M. Davy and E. Roy
- LastWave, <http://wave.cmap.polytechnique.fr/soft/LastWave/index.html> by E. Bacry

Other TFR software can also be found at Web sites of many TF processing researchers. For example, software for the S-transform [156] can presently be found at ftp://ftp.univ-nantes.fr/pub/universite/iutstnazaire/tftb/contribs/stockwell/s_transform.html; and for the reassignment TFRs at <http://www.aei-potsdam.mpg.de/~eric/ecm/theses/ecm98.html> by E. Chassande-Mottin or in the TFT toolbox.

1.3 Quadratic Time–Frequency Representations

Although linear, adaptive and expansion-based TFRs have been used in many applications, for the rest of the chapter we are going to concentrate on the widely used QTFRs. QTFRs have been used extensively to analyze, modify and synthesize TV signals. For example, the well-known WD [3, 4, 7–12, 116] has been used to analyze speech data as it preserves any possible TF shifts of the speech [23, 116]. This is an important property because the 2-D transformation should be able to detect a change of pitch or formant frequencies as time changes. The many other desirable properties of the WD, such as conservation of energy and marginal quantities such as the signal instantaneous power and spectral energy density [3, 9, 116, 122, 157], have also made the WD useful in other applications such as optics, acoustics, bioengineering, image processing, oceanography, signal synthesis [83–85, 158] and TV filtering [116]. Note that in presenting the various QTFRs, we answer the question in the last paragraph of Section 1.1.1 by providing specific important properties, implementation methods and simulation examples.

1.3.1 Multitude of quadratic time–frequency representations

The WD and the spectrogram (SPEC) are two of the earliest QTFRs to be developed and are still used in many applications based on their properties and their computational simplicity. Since then, many other QTFRs have been developed and published in the literature. By no means exhaustively, we provide next an alphabetical list of

some of these QTFRs: α -WD [159]; affine smoothed pseudo-WD [160, 161]; Akroyd distribution [162]; Altes–Marinovich Q-distribution (QD) [163, 164]; Bertrand P_k -distributions [165, 166]; Born–Jordan distribution [7, 86]; Butterworth distribution (BUD) [167]; Choi–Williams distribution (CWD) [168]; cone–kernel distribution [169]; D -Flandrin distribution [170]; exponential smoothed pseudo-WD [112, 171]; exponential WD [112]; generalized exponential distribution [167]; hyperbogram [109]; Levin distribution [172]; Margenau–Hill distribution [173]; minimum variance kernel distribution [174]; multiform tiltable distribution [175]; Page distribution [176]; passive or active Unterberger distributions [177]; power smoothed pseudo-WD [111, 178]; power WD [111]; powergram [179]; pseudo-Wigner distribution (PWD) [111]; reassigned TFRs [139, 140, 180, 181]; reassigned spectrogram [139, 140]; reassigned scalogram [182]; reduced interference distribution [183]; Rihaczek distribution [184]; S-method [185]; scalogram [161]; smoothed pseudo-QD [109, 186]; smoothed pseudo-Wigner distribution (SPWD) [86]; SPEC [5, 6, 13]; and WD [3, 4, 7–12]. Note that other QTFR lists can be found, for example, in [15, 16, 116, 117].

1.3.2 Quadratic time–frequency classification based on properties

Although many QTFRs have been proposed in the literature, no one QTFR exists that can be used effectively in all possible applications. This is because different QTFRs are best suited for analyzing signals with specific types of properties and TF structures. For example, the WD does not support nonlinear signal structures, and thus cannot successfully analyze a hyperbolic TM signal.

The TV signal application areas require processing tools that are well matched to the signal properties. As a result, although a QTFR can provide a time-dependent spectral characterization of a TV signal that is processed, one must choose the QTFR that can provide meaningful and accurate information about the signal. To assist the user in selecting the appropriate analysis tool, QTFRs are often classified based on the various properties (some of which are listed in Table 1.2) they satisfy. A QTFR is then chosen that satisfies properties desirable for a given application. For example, if a signal needs to be synthesized after processing in the TF plane, then the chosen QTFR should conserve energy information. On the other hand, if it is important to extract the GD of a signal from its QTFR, then the GD preservation property must be satisfied by the chosen QTFR. Thus, the choice of a QTFR depends on the set of properties it satisfies, and whether that set of properties is essential to the particular application under consideration.

A classification that has been found useful is based on covariance properties. In particular, a QTFR is said to satisfy a covariance property if the QTFR preserves (or is covariant to) transformations on the signal. Some important covariance properties that a QTFR $T_X(t, f)$ may satisfy include:

Constant time shift, or constant group delay shift (GDS):

$$T_{\mathcal{S}_\tau X}(t, f) = T_X(t - \tau, f) \text{ with } (\mathcal{S}_\tau X)(f) = e^{j2\pi\tau f} X(f) \quad (1.5)$$

TABLE 1.2Properties of a QTFR $T_X(t, f)$ of a Signal $x(t)$

Property Name	Property Definition
Time shift covariance	$(S_\tau X)(f) = e^{j2\pi\tau f} X(f)$ $\Rightarrow T_{S_\tau X}(t, f) = T_X(t - \tau, f)$
Frequency shift covariance	$(\mathcal{M}_\nu X)(f) = X(f - \nu)$ $\Rightarrow T_{\mathcal{M}_\nu X}(t, f) = T_X(t, f - \nu)$
Real valuedness	$T_X(t, f) = T_X^*(t, f)$
Energy conservation	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T_X(t, f) dt df = \int_{-\infty}^{\infty} X(f) ^2 df$
Time marginal	$\int_{-\infty}^{\infty} T_X(t, f) df = x(t) ^2$
Frequency marginal	$\int_{-\infty}^{\infty} T_X(t, f) dt = X(f) ^2$
Time localization	$x(t) = \delta(t - t_0) \Rightarrow T_X(t, f) = \delta(t - t_0)$
Frequency localization	$X(f) = \delta(f - f_0) \Rightarrow T_X(t, f) = \delta(f - f_0)$
Scale change covariance	$(\mathcal{C}_a X)(f) = \frac{1}{\sqrt{ a }} X(f/a)$ $\Rightarrow T_{\mathcal{C}_a X}(t, f) = T_X(at, f/a)$
Moyal's formula (unitarity [115])	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T_{X_1}(t, f) T_{X_2}(t, f) dt df$ $= \left \int_{-\infty}^{\infty} X_1(f) X_2^*(f) df \right ^2$
Group delay	$\frac{\int_{-\infty}^{\infty} t T_X(t, f) dt}{\int_{-\infty}^{\infty} T_X(t, f) dt} = \tau_X(f) = -\frac{1}{2\pi} \frac{d}{df} \arg X(f)$
Instantaneous frequency	$\frac{\int_{-\infty}^{\infty} f T_X(t, f) df}{\int_{-\infty}^{\infty} T_X(t, f) df} = \nu_x(t) = \frac{1}{2\pi} \frac{d}{dt} \arg x(t)$
Finite time support	$x(t) = 0 \text{ for } t \notin [t_1, t_2]$ $\Rightarrow T_X(t, f) = 0 \text{ for } t \notin [t_1, t_2]$
Finite frequency support	$X(f) = 0 \text{ for } f \notin [f_1, f_2]$ $\Rightarrow T_X(t, f) = 0 \text{ for } f \notin [f_1, f_2]$
Hyperbolic GDS covariance	$(\mathcal{H}_c X)(f) = e^{-j2\pi c \ln \frac{f}{f_r}} X(f)$ $\Rightarrow T_{\mathcal{H}_c X}(t, f) = T_X\left(t - \frac{c}{f}, f\right)$
Dispersive GDS covariance	$(\mathcal{D}_c X)(f) = e^{-j2\pi c \xi(f/f_r)} X(f)$ $\Rightarrow T_{\mathcal{D}_c X}(t, f) = T_X(t - c\tau(f), f)$

Note: Different applications require the preservation of different properties.

Linear GDS:

$$\begin{aligned} T_{\mathcal{L}_c X}(t, f) &= T_X \left(t - c \frac{2}{f_r} \left| \frac{f}{f_r} \right|, f \right) \text{ with} \\ (\mathcal{L}_c X)(f) &= e^{j2\pi c \operatorname{sgn}(f)|f/f_r|^2} X(f) \end{aligned} \quad (1.6)$$

Hyperbolic GDS:

$$\begin{aligned} T_{\mathcal{H}_c X}(t, f) &= T_X \left(t - c/f, f \right) \text{ with} \\ (\mathcal{H}_c X)(f) &= e^{j2\pi c \ln \frac{f}{f_r}} X(f), f > 0 \end{aligned} \quad (1.7)$$

Power GDS:

$$\begin{aligned} T_{\mathcal{P}_c X}(t, f) &= T_X \left(t - c \frac{\kappa}{f_r} \left| \frac{f}{f_r} \right|^{\kappa-1}, f \right) \text{ with} \\ (\mathcal{P}_c X)(f) &= e^{j2\pi c \operatorname{sgn}(f)|\frac{f}{f_r}|^\kappa} X(f) \end{aligned} \quad (1.8)$$

Exponential GDS:

$$\begin{aligned} T_{\mathcal{E}_c X}(t, f) &= T_X \left(t - c \frac{\kappa}{f_r} e^{\kappa f/f_r}, f \right) \text{ with} \\ (\mathcal{E}_c X)(f) &= e^{j2\pi c e^{\kappa f/f_r}} X(f) \text{ and } \tau(f) = \frac{d}{df} \xi \frac{f}{f_r} \end{aligned} \quad (1.9)$$

Dispersive GDS:

$$\begin{aligned} T_{\mathcal{D}_c X}(t, f) &= T_X(t - c\tau(f), f) \text{ with} \\ (\mathcal{D}_c X)(f) &= e^{j2\pi c \xi(\frac{f}{f_r})} X(f) \end{aligned} \quad (1.10)$$

Frequency shift:

$$\begin{aligned} T_{\mathcal{M}_\nu X}(t, f) &= T_X(t, f - \nu) \text{ with} \\ (\mathcal{M}_\nu X)(f) &= X(f - \nu) \end{aligned} \quad (1.11)$$

Scale change:

$$T_{\mathcal{C}_a X}(t, f) = T_X(at, f/a) \text{ with } (\mathcal{C}_a X)(f) = \frac{1}{\sqrt{|a|}} X(f/a) \quad (1.12)$$

The preceding notation, for example $T_{S_\tau X}(t, f)$, stands for $T_Y(t, f)$ with $Y(f) = (S_\tau X)(f)$. Note that the hyperbolic GDS operator $(\mathcal{H}_c X)(f)$ causes a nonlinear time shift, $t = c/f$, of the GD of the signal that is inversely proportional to frequency. On the other hand, the constant time shift operator $(S_\tau X)(f)$ causes a constant time shift $t = \tau$ that does not depend on frequency. In Equation (1.10), $(\mathcal{D}_c X)(f)$ is the dispersive GDS operator that causes a dispersive time shift $t = c\tau(f)$ on the signal. Here, $\xi(f/f_r)$ represents a nonlinear phase change of $X(f)$ and $\tau(f) = \frac{d}{df}\xi(f/f_r)$ is its derivative. QTFRs that preserve constant time shifts are well matched to signals with linear phase whereas QTFRs that preserve nonconstant time shifts are well matched to signals with nonlinear phase.

Two well-known classes of QTFRs are Cohen’s class of constant time shift and constant frequency shift covariant QTFRs [7, 11, 16, 122, 187], and the affine class of constant time shift covariant and scale change covariant QTFRs [16, 116, 161, 165, 188]. These two QTFR classes are formed by grouping together QTFRs that satisfy the indicated pair of covariance properties. Both of these classes preserve constant time delays, and are thus expected to work best when the analysis signal has linear TF characteristics. Other classes of QTFRs were proposed to ideally match signals with nonlinear or dispersive TF characteristics. For example, the hyperbolic class QTFRs [92, 109, 186, 189] are covariant to hyperbolic GDS and scale changes on the signal, and are best used to analyze signals with dispersive hyperbolic GD. Other covariant QTFR classes include the power classes [111, 178, 179, 189, 190], the exponential class [112, 171], many unitarily equivalent QTFR classes [191–193] and the displacement covariant classes [194–197]. These QTFR classes satisfy different covariance properties, and, as a result, are useful in different types of applications.

Next, we discuss in more detail some classes of QTFRs based on their specific TF structures.

1.3.3 Cohen’s class quadratic time–frequency representations

1.3.3.1 Theoretical formulation

Cohen’s class of signal-independent kernel functions contains all QTFRs, $T_X^{(C)}(t, f)$, that satisfy the time shift covariance in Equation (1.5) and the frequency shift covariance in Equation (1.11) [7, 11, 122]. Both covariance properties are important in applications where the signal needs to be analyzed at all TF points with fixed TF resolution. As a result, QTFRs in Cohen’s class exhibit analysis characteristics that do not change with time and frequency, and are useful for constant bandwidth analysis such as for speech, narrowband Doppler systems and multipath environments.

Any Cohen’s class QTFR can be expressed in two equivalent expressions:

$$T_X^{(C)}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_T^{(C)}(t - \hat{t}, f - \hat{f}) \text{WD}_X(\hat{t}, \hat{f}) d\hat{t} d\hat{f} \quad (1.13)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi_T^{(C)}(\tau, \nu) \text{AF}_X(\tau, \nu) e^{j2\pi(t\nu - f\tau)} d\tau d\nu \quad (1.14)$$

in terms of the 2-D signal-independent kernel functions $\psi_T^{(C)}(t, f)$ and $\Psi_T^{(C)}(\tau, \nu)$. In Equation (1.13), a Cohen’s class QTFR of a signal $x(t)$ with FT $X(f)$ is the result of a 2-D convolution of its kernel $\psi_T^{(C)}(t, f)$ with the WD, $\text{WD}_X(t, f)$, of the signal that is defined as:

$$\begin{aligned}\text{WD}_X(t, f) &= \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{j2\pi f \tau} d\tau \\ &= \int_{-\infty}^{\infty} X\left(f + \frac{\nu}{2}\right) X^*\left(f - \frac{\nu}{2}\right) e^{j2\pi t\nu} d\nu\end{aligned}\quad (1.15)$$

In Equation (1.14), a Cohen’s class QTFR can be written as the 2-D FT of the product of its kernel $\Psi_T^{(C)}(\tau, \nu)$ and the Woodward ambiguity function (AF) of the signal. The AF is the 2-D FT of the WD in Equation (1.15) [15],

$$\begin{aligned}\text{AF}_X(\tau, \nu) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{WD}_X(t, f) e^{j2\pi \tau f} e^{-j2\pi t\nu} dt df \\ &= \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi \nu t} dt\end{aligned}\quad (1.16)$$

and it can be considered as the correlation of the signal with a TF-shifted version of itself. Because Equations (1.13) and (1.14) are equivalent, it can be shown that $\psi_T^{(C)}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi_T^{(C)}(\tau, \nu) e^{-j2\pi \tau f} e^{j2\pi \nu t} d\tau d\nu$ [16, 116, 187]. Note that once the kernel is fixed, it uniquely characterizes a QTFR from Cohen’s class.

1.3.3.2 Cohen’s class QTFR examples

Some QTFRs of Cohen’s class are provided in the first column of [Table 1.3](#). Their formulation can be obtained by substituting their kernels (given in the second column) in the QTFR formulations in Equations (1.13) or (1.14). For example, the SPEC can be obtained by substituting its kernel $\psi_{\text{SPEC}}^{(C)}(t, f) = \text{WD}_{\Gamma}(-t, -f)$ in Equation (1.13). Here, $\text{WD}_{\Gamma}(t, f)$ is the WD in Equation (1.15) of an analysis window $\Gamma(f)$. The WD can be similarly obtained by inserting $\Psi_{\text{WD}}^{(C)}(\tau, \nu) = 1$ in Equation (1.14), thus demonstrating its FT relation to the AF. Other members of Cohen’s class include the α -WD that equals to the WD for $\alpha = 0$ and to the Rihaczek distribution for $\alpha = 1/2$; the CWD; the generalized exponential distribution; the BUD; the PWD; the SPWD; the Born–Jordon distribution; the Levin distribution; the Page distribution; the cone-kernel distribution; the reduced interference distributions; and the multiform tiltable distribution.

Cohen’s class QTFRs satisfy many desirable properties depending on their defining kernel. In the first column of [Table 1.4](#), we provide a list of desirable QTFR properties (as defined in [Table 1.2](#)) that are guaranteed to hold for a Cohen’s class QTFR if its kernel satisfies the kernel constraints in the second column of Table 1.4. The actual properties that different Cohen’s class QTFRs satisfy are listed in the third column of [Table 1.3](#). Note, for example, that as the WD has the simple kernel $\Psi_{\text{WD}}^{(C)}(\tau, \nu) = 1$, it satisfies all the properties listed in [Table 1.4](#), and is a highly localized QTFR. However,

TABLE 1.3Kernels and Properties (from [Table 1.4](#)) of Some Cohen’s Class QTFRs

Cohen’s Class QTFR	Kernels	Properties Satisfied
WD	$\psi_{\text{WD}}^{(C)}(t, f) = \delta(t)\delta(f)$ $\Psi_{\text{WD}}^{(C)}(\tau, \nu) = 1$	CP ₁ –CP ₁₃
α -WD	$\psi_{\alpha\text{-WD}}^{(C)}(t, f) = \frac{1}{ \alpha } e^{j2\pi \frac{tf}{\alpha}}$ $\Psi_{\alpha\text{-WD}}^{(C)}(\tau, \nu) = e^{j2\pi \alpha \tau \nu}$ if $ \alpha < \frac{1}{2}$	CP ₁ , CP ₂ , CP ₄ –CP ₁₀ , CP ₁₃ and CP ₁₄
Choi–Williams distribution (GWD)	$\Psi_{\text{CWD}}^{(C)}(\tau, \nu) = e^{-\frac{(2\pi\tau\nu)^2}{\sigma}}$	CP ₁ –CP ₉ , CP ₁₁ , CP ₁₂
Generalized exponential distribution (GED)	$\Psi_{\text{GED}}^{(C)}(\tau, \nu) = e^{-(\frac{\tau}{\tau_0})^{2M}(\frac{\nu}{\nu_0})^{2N}}$	CP ₁ –CP ₈ , CP ₉ if $M = N$ CP ₁₁ if $M > \frac{1}{2}$, CP ₁₂ if $N > \frac{1}{2}$
Butterworth distribution (BUD)	$\Psi_{\text{BUD}}^{(C)}(\tau, \nu) = \frac{1}{1+(\frac{\tau}{\tau_0})^{2M}(\frac{\nu}{\nu_0})^{2N}}$	CP ₁ –CP ₈ , CP ₉ if $M = N$, CP ₁₁ if $M > \frac{1}{2}$, CP ₁₂ if $N > \frac{1}{2}$
Spectrogram (SPEC)	$\psi_{\text{SPEC}}^{(C)}(t, f) = \text{WD}_{\Gamma}(-t, -f)$ $\Psi_{\text{SPEC}}^{(C)}(\tau, \nu) = \text{AF}_{\Gamma}(-\tau, -\nu)$	CP ₁ –CP ₃ , CP ₄ if $\int_{-\infty}^{\infty} \gamma(t) ^2 dt = 1$
Pseudo-WD (PWD)	$\psi_{\text{PWD}}^{(C)}(t, f) = \delta(t)\text{WD}_{\Gamma}(0, f)$ $\Psi_{\text{PWD}}^{(C)}(\tau, \nu) = \gamma(\tau/2) \times \gamma^*(-\tau/2)$	CP ₁ –CP ₃ , CP ₁₃ , CP ₄ , CP ₅ and CP ₇ if $ \gamma(0) = 1$, CP ₁₁ if $\gamma(0) = 1$
Smoothed pseudo-WD (SPWD)	$\psi_{\text{SPWD}}^{(C)}(t, f) = h(t)\text{WD}_{\Gamma}(0, f)$ $\Psi_{\text{SPWD}}^{(C)}(\tau, \nu) = H(\nu)\gamma(\tau/2) \times \gamma^*(-\tau/2)$	CP ₁ , CP ₂ , CP ₃ if $h(t) \in R$, CP ₄ if $H(0) \gamma(0) ^2 = 1$

Note: The analysis windows $\gamma(t) \leftrightarrow \Gamma(f)$ and $h(t) \leftrightarrow H(f)$ are FTs pairs. The WD and the AF are defined in Equations (1.15) and (1.16), respectively.

TABLE 1.4

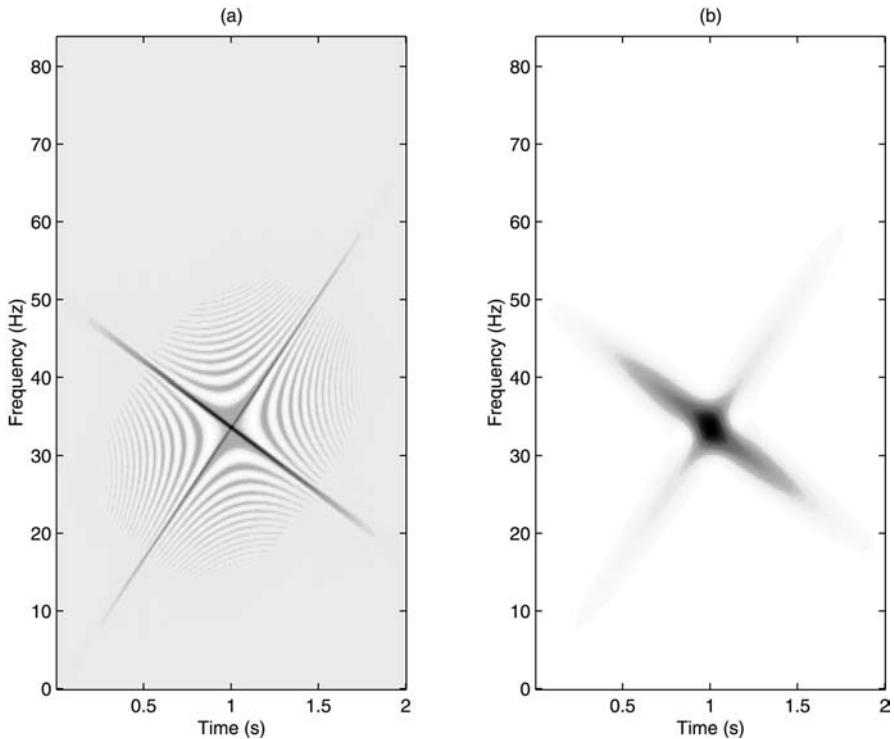
Desirable Properties and Corresponding Kernel Constraints for Cohen's Class QTFRs

Cohen's Class Property	Cohen's Class Kernel Constraint
CP ₁ time shift covariance	Any $\Psi_T^{(C)}(\tau, \nu)$
CP ₂ frequency shift covariance	Any $\Psi_T^{(C)}(\tau, \nu)$
CP ₃ real valuedness	$\Psi_T^{(C)}(\tau, \nu) = \Psi_T^{(C)*}(-\tau, -\nu)$
CP ₄ energy conservation	$\Psi_T^{(C)}(0, 0) = 1$
CP ₅ time marginal	$\Psi_T^{(C)}(0, \nu) = 1$
CP ₆ frequency marginal	$\Psi_T^{(C)}(\tau, 0) = 1$
CP ₇ time localization	$\Psi_T^{(C)}(0, \nu) = 1$
CP ₈ frequency localization	$\Psi_T^{(C)}(\tau, 0) = 1$
CP ₉ scale change covariance	$\Psi_T^{(C)}(\tau, \nu) = \Psi_T^{(C)}(\tau\nu)$
CP ₁₀ Moyal's formula (unitarity)	$ \Psi_T^{(C)}(\tau, \nu) = 1$
CP ₁₁ instantaneous frequency	$\Psi_T^{(C)}(0, \nu) = 1$ and $\frac{\partial}{\partial \tau} \Psi_T^{(C)}(\tau, \nu) \Big _{\tau=0} = 0$
CP ₁₂ group delay	$\Psi_T^{(C)}(\tau, 0) = 1$ and $\frac{\partial}{\partial \nu} \Psi_T^{(C)}(\tau, \nu) \Big _{\nu=0} = 0$
CP ₁₃ finite time support	$\int_{-\infty}^{\infty} \Psi_T^{(C)}(\tau, \nu) e^{j2\pi t \nu} d\nu = 0$ for $ \frac{t}{\tau} > \frac{1}{2}$
CP ₁₄ finite frequency support	$\int_{-\infty}^{\infty} \Psi_T^{(C)}(\tau, \nu) e^{-j2\pi f \tau} d\tau = 0$ for $ \frac{f}{\nu} > \frac{1}{2}$

it suffers from CTs when used to analyze signals with multiple components or signals with nonlinear TF characteristics. This is demonstrated in [Figure 1.6\(a\)](#) with the WD of the sum of the two crossing LFM chirps used in [Figure 1.5](#). It can be seen that half way between any two points in the TF plane, the WD produces oscillatory CTs that could provide misleading information about the signal TF structure. The SPEC of the same signal in [Figure 1.6\(b\)](#) is a smoothed WD that suppresses CTs. This, however, comes at the expense of loosing many properties (see [Table 1.3](#)) as well as the high resolution of the WD autotermcomponents. Specifically, the maximum temporal (spectral) resolution that can be achieved by the SPEC is based on the duration (bandwidth) of its smoothing window [116].

1.3.3.3 Implementation algorithms

Because any Cohen's class QTFR can be obtained by the 2-D FT of the product of the AF and the QTFR kernel $\Psi_T^{(C)}(\tau, \nu)$ in Equation (1.14), the QTFR can be efficiently implemented using 2-D fast FT algorithms. The AF can be computed using the TFT

**FIGURE 1.6**(a) The WD and (b) the spectrogram of the sum of the two LFM chirps in [Figure 1.5](#).

toolbox [82]. By recalling that the WD suffers from CTs when analyzing multicomponent signals, the kernel $\Psi_T^{(C)}(\tau, \nu)$ is often chosen to be a low-pass function. This is because the CTs are mapped away from the origin in the AF domain whereas the desirable autoterms are mapped to the origin. Thus, a 2-D low-pass function reduces CTs while preserving some of the desirable properties of the WD. Note that computational algorithms for many Cohen's class QTFRs can be found in the TFT toolbox [82]. In [198], an overview of possible errors due to aliasing is provided for discrete formulations of Cohen's class QTFRs.

1.3.4 Affine class quadratic time–frequency representations

1.3.4.1 Theoretical formulation

The affine class [16, 116, 161, 165, 188] contains all QTFRs $T_X^{(A)}(t, f)$ that satisfy the constant time shift covariance property in Equation (1.5) and the scale change covariance property in Equation (1.12). The scale change covariance is important for multiresolution analysis (similar to the analysis offered by the wavelet transform [116, 121]), for self-similar signals [199], for scale-covariant systems [200], for image

enlargement or compression and in the context of the wideband Doppler effect [201]. Many affine class QTFRs (like the scalogram that is the squared magnitude of the wavelet transform) have constant-Q TF analysis characteristics where the analysis bandwidth is proportional to the analysis frequency. This constant-Q analysis offers an alternative to the constant-bandwidth analysis achieved by QTFRs in Cohen's class.

Any affine class QTFR can be expressed as:

$$T_X^{(A)}(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi_T^{(A)}\left(f(t - \hat{t}), -\frac{\hat{f}}{f}\right) \text{WD}_X(\hat{t}, \hat{f}) d\hat{t} d\hat{f} \quad (1.17)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi_T^{(A)}\left(f\tau, \frac{\nu}{f}\right) \text{AF}_X(\tau, \nu) e^{j2\pi t\nu} d\tau d\nu \quad (1.18)$$

which is in terms of the 2-D FT pair kernels $\psi_T^{(A)}(c, b) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Psi_T^{(C)}(\zeta, \beta) e^{-j2\pi\zeta b} e^{j2\pi\beta c} d\zeta d\beta$ [116]. Note that any one kernel uniquely characterizes an affine QTFR $T^{(A)}(t, f)$ in Equations (1.17) and (1.18).

1.3.4.2 Affine class QTFR examples

Some members of the affine class are summarized in Table 1.5 together with their corresponding kernel functions. The WD in Equation (1.15) as well as the α -WD are members of the affine class and Cohen's class because they satisfy all three properties shared by both classes: time shift covariance in Equation (1.5), frequency shift covariance in Equation (1.11) and scale change covariance in Equation (1.12). Other affine QTFRs include the scalogram [121, 161], the affine smoothed pseudo-WD (ASPWD), the unitary Bertrand P_0 -distribution [165, 166, 170, 202], the Flandrin D -distribution [170], and the passive (PUD) and active (AUD) Unterberger distributions [165, 170]. An important family of affine QTFRs is given by the Bertrand P_κ -distributions with varying parameter κ [165, 166]. In addition to having scale and time shift covariance, the Bertrand P_κ -distributions also preserve frequency-dependent nonlinear time shifts with κ th order power TF characteristics [166, 179].

Any one of these QTFRs can be formulated by substituting their kernel from Table 1.5 in Equations (1.17) or (1.18). For example, the scalogram can be obtained by substituting its kernel $\psi_{\text{SCAL}}^{(A)}(c, b) = \text{WD}_\Theta(-c, -b)$ in Equation (1.17). This kernel is actually the WD of the analysis wavelet function $\Theta(f)$. Note that Equation (1.17) describes a 2-D affine convolution or smoothing operation that results in QTFRs that could possibly be used for multiresolution processing. Thus, the scalogram is actually obtained by 2-D smoothing the WD of the signal with the WD of the wavelet function [16, 160, 161].

An affine class QTFR can satisfy additional desirable properties provided that the kernel of the QTFR satisfies certain constraints associated with these properties. A list of QTFR properties and kernel constraints is provided in Table 1.6, and the third column of Table 1.5 lists the properties satisfied by specific affine QTFRs. Based on the multiresolution application at hand, the desirable properties narrow down the

TABLE 1.5

Kernels and Properties (from Table 1.6) of Some Affine Class QTFRs

Affine QTFR	Kernels	Properties Satisfied
WD	$\psi_{\text{WD}}^{(A)}(c, b) = \delta(c)\delta(b+1)$ $\Psi_{\text{WD}}^{(A)}(\zeta, \beta) = e^{-j2\pi\zeta}$	AP ₁ –AP ₁₃
α -WD	$\psi_{\alpha\text{-WD}}^{(A)}(c, b) = \frac{1}{ \alpha } e^{j2\pi\frac{c(b+1)}{\alpha}}$ $\Psi_{\alpha\text{-WD}}^{(A)}(\zeta, \beta) = e^{j2\pi\alpha\zeta\beta} e^{-j2\pi\zeta}$	AP ₁ , AP ₂ , AP ₄ –AP ₉ , AP ₁₃ AP ₁₁ and AP ₁₂ if $ \alpha < \frac{1}{2}$
Scalogram (SCAL)	$\psi_{\text{SCAL}}^{(A)}(c, b) = \text{WD}_{\Theta}(-c, -b)$ $\Psi_{\text{SCAL}}^{(A)}(\zeta, \beta) = \text{AF}_{\Theta}(-\zeta, -\beta)$	AP ₁ –AP ₃ AP ₄ if $\int_{-\infty}^{\infty} \Theta(f) ^2 \frac{df}{ f } = 1$
Affine smoothed pseudo WD (ASPWD)	$\psi_{\text{ASPWD}}^{(A)}(c, b) = s(c)H(b)$ $\Psi_{\text{ASPWD}}^{(A)}(\zeta, \beta) = h(\zeta)S(\beta)$	AP ₁ , AP ₂ AP ₃ if $H(b), s(c) \in R$ AP ₄ if $S(0)$ $\times \int_{-\infty}^{\infty} H(b) \frac{db}{ b } = 1$
Bertrand	$\Phi_{P_0}^{(A)}(b, \beta) = G(\beta)\delta(b - F(\beta))$	AP ₁ –AP ₅ , AP ₇
P_0 -distribution	with $F(\beta) = -\frac{\beta}{2}\coth(\frac{\beta}{2})$ and $G(\beta) = \frac{\beta/2}{\sinh(\beta/2)}$	AP ₉ –AP ₁₁ , AP ₁₄
Flandrin D -distribution	$\Phi_D^{(A)}(b, \beta) = G(\beta)\delta(b - F(\beta))$ with $F(\beta) = -1 - (\frac{\beta}{4})^2$ and $G(\beta) = 1 - (\frac{\beta}{4})^2$	AP ₁ –AP ₇ , AP ₁₀
Passive	$\Phi_{\text{PUD}}^{(A)}(b, \beta) = G(\beta)\delta(b - F(\beta))$	AP ₁ –AP ₇ , AP ₁₀ , AP ₁₁
Unterberger (PUD)	with $F(\beta) = -\sqrt{1 + (\frac{\beta}{2})^2}$ and $G(\beta) = \frac{1}{\sqrt{1 + (\frac{\beta}{2})^2}}$	
Active	$\Phi_{\text{AUD}}^{(A)}(b, \beta) = G(\beta)\delta(b - F(\beta))$	AP ₁ –AP ₅ , AP ₇
Unterberger (AUD)	with $F(\beta) = -\sqrt{1 + (\frac{\beta}{2})^2}$ and $G(\beta) = 1$	AP ₈ , AP ₁₀ , AP ₁₁
Bertrand P_{κ} -distributions	$\Phi_{P_{\kappa}}^{(A)}(b, \beta) = \int_{-\infty}^{\infty} \delta(-b + \frac{\beta}{2} - \lambda_{\kappa}(u)) \times \delta(-b - \frac{\beta}{2} - \lambda_{\kappa}(-u)) \mu(u) du$	AP ₁ , AP ₂ , other properties depending on $\mu(u)$

Note: Here, $\Phi_T^{(A)}(b, \beta) = \int_{-\infty}^{\infty} \Psi_T^{(A)}(\zeta, \beta) e^{-j2\pi b \zeta} d\zeta$, the window functions $s(c) \leftrightarrow S(\beta)$ and $h(\zeta) \leftrightarrow H(b)$ are FT pairs, and $\Theta(f)$ is a wavelet function. From [165], $\mu(u)$ is a real and even weighting function and $\lambda_{\kappa}(u) = (\kappa \frac{e^{-u}-1}{e^{-\kappa u}-1})^{\frac{1}{\kappa-1}}$ for $\kappa \neq 0, 1$.

TABLE 1.6

Desirable Properties and Corresponding Kernel Constraints for Affine Class QTFRs

Affine Class QTFR Property	Affine Class Kernel Constraint
AP ₁ scale change covariance	Any $\Phi_T^{(A)}(b, \beta)$
AP ₂ time shift covariance	Any $\Phi_T^{(A)}(b, \beta)$
AP ₃ real valuedness	$\Phi_T^{(A)}(b, \beta) = \Phi_T^{(A)*}(b, -\beta)$
AP ₄ energy conservation	$\int_{-\infty}^{\infty} \Phi_T^{(A)}(b, 0) \frac{db}{ b } = 1$
AP ₅ frequency marginal	$\Phi_T^{(A)}(b, 0) = \delta(b+1)$
AP ₆ time marginal	$\int_{-\infty}^{\infty} \Phi_T^{(A)}(b, \alpha b) \frac{db}{ b } = 1, \forall \alpha$
AP ₇ frequency localization	$\Phi_T^{(A)}(b, 0) = \delta(b+1)$
AP ₈ time localization	$\int_{-\infty}^{\infty} \Phi_T^{(A)}(b, \beta) db = 1, \forall \beta$
AP ₉ Moyal's formula (unitarity [439])	$\int_{-\infty}^{\infty} \Phi_T^{(A)*}(b\beta, \alpha\beta) \Phi_T^{(A)}(\beta, \alpha\beta) d\beta = \delta(b-1), \forall \alpha$
AP ₁₀ group delay	$\Phi_T^{(A)}(b, 0) = \delta(b+1)$ and $\frac{\partial}{\partial \beta} \Phi_T^{(A)}(b, \beta) _{\beta=0} = 0$
AP ₁₁ finite frequency support	$\Phi_T^{(A)}(b, \beta) = 0$ for $\left \frac{b+1}{\beta}\right > \frac{1}{2}$
AP ₁₂ finite time support	$\phi_T^{(A)}(c, \zeta) = 0$ for $\left \frac{c}{\zeta}\right > \frac{1}{2}$
AP ₁₃ frequency shift covariance	$\Psi_T^{(A)}(\zeta, \beta) = \Psi_T^{(A)}(\zeta\beta) e^{-j2\pi\zeta}$
AP ₁₄ hyperbolic GDS covariance	$\Phi_T^{(A)}(b, \beta) = G_T(\beta) \delta(b + \frac{\beta}{2} \coth \frac{\beta}{2})$ where $G_T(\beta)$ is a 1-D kernel function

Note: $\Phi_T^{(A)}(b, \beta) = \int_{-\infty}^{\infty} \Psi_T^{(A)}(\zeta, \beta) e^{-j2\pi b \zeta} d\zeta$ and $\phi_T^{(A)}(c, \zeta) = \int_{-\infty}^{\infty} \Psi_T^{(A)}(\zeta, \beta) e^{j2\pi \beta c} d\beta$.

choice of the affine QTFR. For example, if it is important to process a signal without the presence of CTs irrespective of whether energy is conserved, then the scalogram can be chosen instead of the WD or the unitary Bertrand P_0 -distribution.

1.3.4.3 Implementation algorithms

The affine QTFR formulation in Equation (1.18) is not a simple 2-D multiplication operation like the corresponding formulation of a Cohen's class QTFR in Equation (1.14) unless the output frequency f is fixed. As a result, the kernel variables

change for different output frequency values, thus achieving multiresolution processing. MATLAB algorithms to compute many affine QTFRs can be found in the TFT toolbox [82], and some implementation algorithms are also given in [110].

1.3.5 Dispersive class quadratic time–frequency representation

1.3.5.1 Theoretical formulation

Many TV signals we encounter, such as dolphin and whale whistles, radar and sonar waveforms, and shock waves in fault structures, have nonlinear TF characteristics. To improve the processing of such signals, QTFRs that satisfy the dispersive GDS covariance property were designed in [1, 92, 112, 190, 203]. This property, defined in Equation (1.10), is important for analyzing signals propagating through systems with dispersive characteristics or, equivalently, with nonlinear GD functions. If the signal spectrum $X(f)$ is passed through an all-pass dispersive system with output $Y(f) = e^{-j2\pi\xi(f/f_r)} X(f)$, then the change in GD, $\tau(f) = \frac{d}{df}\xi(f/f_r)$, is proportional to the derivative of the one-to-one phase function $\xi(f/f_r)$. Because GD is a measure of the time delay introduced in each sinusoidal component of the signal at frequency f , the ideal QTFR $T_X(t, f)$ should preserve this change in GD or frequency-dependent time shift $\tau(f)$ as indicated in Equation (1.10). Different dispersive GDS can be obtained by fixing $\xi(f/f_r)$ in Equation (1.10). In particular, we obtain (1) linear GDS $\tau(f) = (2/f_r)|f/f_r|$ in Equation (1.6) [1]; (2) hyperbolic GDS $\tau(f) = 1/f$ in Equation (1.7) [109]; (3) κ th power GDS $\tau(f) = \frac{\kappa}{f_r}|f/f_r|^{\kappa-1}$ in [179]; or (4) exponential GDS $\tau(f) = \frac{\kappa}{f_r} e^{\kappa f/f_r}$ in [112].

Dispersive GDS covariant, or simply dispersive, QTFR classes are unitarily equivalent [192–197] to known QTFR classes because they can be obtained by warping existing time shift covariant classes such as Cohen’s class or the affine class. We can obtain dispersive class QTFRs $T^{(D)}$ by warping Cohen’s class QTFRs $T = T^{(C)}$ or affine class QTFRs $T = T^{(A)}$ using^{||} [1, 92, 93, 112, 204]:

$$T_X^{(D)}(t, f) = T_{\mathcal{W}_\xi X} \left(\frac{t}{f_r \tau(f)}, f_r \xi \left(\frac{f}{f_r} \right) \right)$$

$$\text{with } (\mathcal{W}_\xi X)(f) = \left| f_r \tau \left(f_r \xi^{-1} \left(\frac{f}{f_r} \right) \right) \right|^{-1/2} X \left(f_r \xi^{-1} \left(\frac{f}{f_r} \right) \right) \quad (1.19)$$

The class formulations in terms of the signal and kernel functions can be found in [204]. For example, a dispersive class obtained by warping Cohen’s class QTFRs can be formulated by inserting Equation (1.13) in Equation (1.19) with $T = T^{(C)}$. The dispersively warped version of the WD (DWD) is a member of the dispersive class, and is defined as $DWD_X(t, f) = WD_{\mathcal{W}_\xi X} \left(\frac{t}{f_r \tau(f)}, f_r \xi(f/f_r) \right)$.

The dispersive QTFRs always satisfy Equation (1.10) for a given function $\xi(f/f_r)$. Due to the warping, the time shift covariance in Equation (1.5) is transformed onto the

^{||}The function $\xi(f/f_r)$ for the dispersive QTFRs in Equation (1.19) is the same as the characteristic function of the MST in Equation (1.3).

dispersive GDS covariance as the time shift operator \mathcal{S}_{c/f_r} transforms to the dispersive GDS operator $\mathcal{D}_c = \mathcal{W}_\xi^{-1} \mathcal{S}_{c/f_r} \mathcal{W}_\xi$ in Equation (1.10) where \mathcal{W}_ξ^{-1} is the inverse warping operator. These QTFRs also satisfy another property depending on $\xi(f/f_r)$ as well as the effect of the warping on the frequency shift covariance in Equation (1.11) or on the scale change covariance in Equation (1.12). A list of properties that different dispersive QTFRs satisfy is provided in [93]. For example, the DWD is highly localized for the TM signal $G(f)$ in Equation (1.2). Specifically, the DWD is a Dirac delta function at GD $\tau(f)$ of the signal, i.e., $DWD_G(t, f) = |\tau(f)|\delta(t - c\tau(f))$. Thus, the DWD is ideally matched to TM signals when the $\tau(f)$ in the DWD formulation matches the GD of the signal. Note that a dual dispersive class can be similarly obtained to match the dispersive FM signals in Equation (1.1) by preserving dispersive IF shifts [205].

1.3.5.2 Dispersive class QTFR examples

Different dispersive QTFR classes can be obtained simply by choosing the function $\xi(f/f_r)$ or its derivative $\tau(f)$ in Equation (1.19). Once the function is fixed, the resulting QTFRs can be useful in analyzing signals whose GD characteristics are the same or approximately the same as $\tau(f)$. Some examples of these classes include the linear chirp class (warped affine class) with linear GD; the hyperbolic class (warped Cohen's class) with hyperbolic GD; the κ th power class (warped affine class) with κ th order power GD; and the exponential class (warped affine class) with exponential GD. Table 1.7 provides dispersive QTFR classes together with their defining characteristic functions $\xi(f/f_r)$ and $\tau(f)$, and the GDS covariance property. Also included are the name of the classes and some important class members. Next, we consider in more detail some of these classes and the signals that these QTFRs are best matched to.

1.3.5.2.1 Hyperbolic class When the characteristic function is logarithmic, $\xi(f/f_r) = \ln(f/f_r)$, and the class warped is Cohen's class in Equation (1.19), then the resulting QTFRs are matched to signals with hyperbolic GD $\tau(f) = 1/f$, $f > 0$. These QTFRs are called members of the hyperbolic class [92, 109, 186, 190, 204], and are covariant to hyperbolic GDS in Equation (1.7) and scale changes in Equation (1.12). The latter property follows because warping the frequency shift operator \mathcal{M} in Equation (1.12) results in the scale operator $\mathcal{C}_a = \mathcal{W}_\xi^{-1} \mathcal{M}_{\ln a} \mathcal{W}_\xi$ in Equation (1.12). Some examples of hyperbolic QTFRs include the Altes–Marinovich QD [163, 164], the unitary Bertrand P_0 -distribution [166], and smoothed versions of the QD [109]. Following Equation (1.19) with $\xi(f/f_r) = \ln(f/f_r)$, the QD is the warped version of the WD:

$$\begin{aligned} QD_X(t, f) &= WD_Y(tf/f_r, \ln(f/f_r)) \\ &= f \int_{-\infty}^{\infty} X(fe^{\beta/2}) X^*(fe^{-\beta/2}) e^{j2\pi tf\beta} d\beta \end{aligned} \quad (1.20)$$

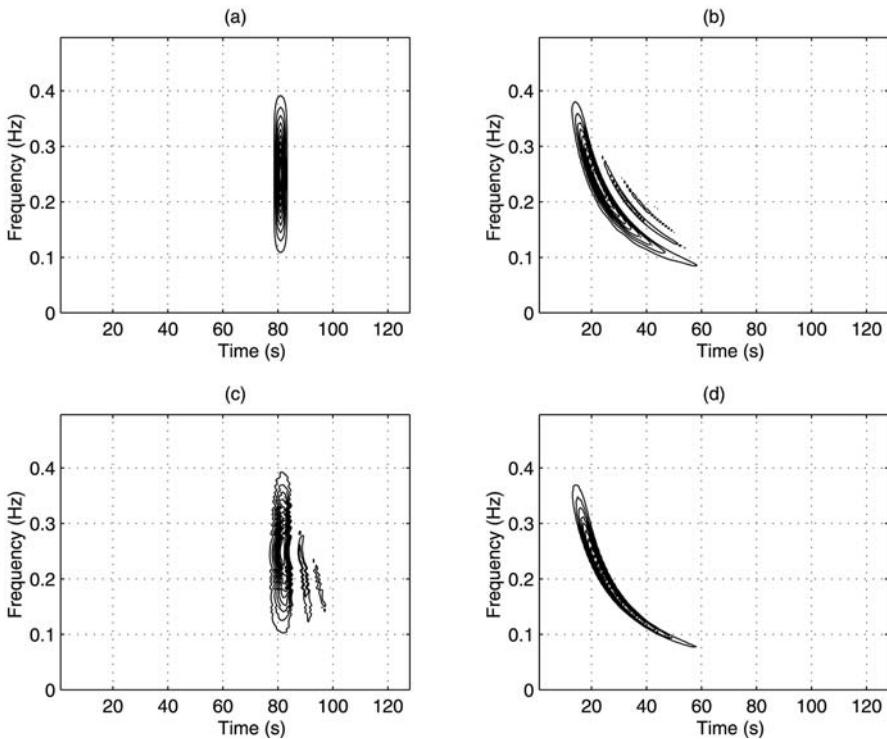
TABLE 1.7

Various Dispersive QTFR Classes with Their Characteristic Functions and Their Derivatives, Important Covariance Property, Class Name and Some Important QTFR Members

Characteristic Function $\xi(\frac{f}{f_r})$	GDS $\tau(f) = \frac{d}{df} \xi(\frac{f}{f_r})$	Covariance Property	QTFR Class	QTFR Members
$\frac{f}{f_r}$	$1/f_r$	Constant GDS	Cohen's class	WD and spectrogram
$\ln \frac{f}{f_r}$	$1/f$	Hyperbolic GDS	Hyperbolic class	QD and hyperbogram
$\operatorname{sgn}(f) f/f_r ^2$	$\frac{2}{f_r} f/f_r $	Linear GDS	Linear chirp class	LWD and chirrogram
$\operatorname{sgn}(f) f/f_r ^\kappa$	$\frac{\kappa}{f_r} \times f/f_r ^{\kappa-1}$	Power GDS	κ th Power class	Power WD and powergram
$e^{\kappa f/f_r}$	$\frac{\kappa}{f_r} e^{\kappa f/f_r}$	Exponential GDS	Exponential class	EWD and exponogram

for $f > 0$ where $Y(f) = \sqrt{e^{f/f_r}} X(f_r e^{f/f_r})$. Just like the WD, its unitary equivalent QTFR, the QD, satisfies many desirable properties but suffers from CTs when processing multicomponent signals. Hyperbolic QTFRs are ideal for analyzing Doppler-invariant signals or systems, signals with logarithmic phase functions similar to the signals used by bats for echolocation or dolphins for communication [206], and self-similar random signals like fractals or fractional Brownian motion [199]. Bat echolocation signals, for example, were shown to have logarithmic phase characteristics in [91, 207]. Note that hyperbolic QTFRs provide an alternative framework for constant-Q analysis from the affine class because both classes satisfy the scale change covariance property. If a particular application involves signals with hyperbolic characteristics, and it is not essential to provide the exact reference frequency or time location of the signal, then the QD is preferred over the widely used WD. If it is also important to preserve constant time shifts on the signal, then the unitary Bertrand P_0 -distribution is preferred over the QD because it is a member of both the affine class and the hyperbolic class, and thus preserves both constant and hyperbolic GDS.

The matching of hyperbolic QTFRs to signals with hyperbolic GD vs. the matching of Cohen's or affine QTFRs to signals with constant GD is demonstrated in Figure 1.7. Specifically, when the analysis signal is an impulse — or constant TM signal in Equation (1.2) with $\xi(f/f_r) = f/f_r$ — its WD in Figure 1.7 provides a highly localized QTFR. On the other hand, its QD in Figure 1.7 suffers from inner

**FIGURE 1.7**

TF analysis of an impulse (constant TM signal) using (a) the WD and (c) the QD, and a hyperbolic TM signal using (b) the WD and (d) the QD. The sampling frequency used is 1 Hz.

CTs. The reverse occurs when the analysis signal is a hyperbolic TM signal — the signal in Equation (1.2) with $\xi(f/f_r) = \ln(f/f_r)$. The WD of this signal in Figure 1.7 is not well matched to hyperbolic GD characteristics and thus results in inner CTs. The QD of the same signal in Figure 1.7 is ideally matched to the hyperbolic TF signal and thus it provides a highly localized representation along the hyperbolic GD function without any inner CTs. This follows because hyperbolic class QTFRs satisfy the property of Equation (1.10) with $\xi(f/f_r) = \ln(f/f_r)$ and hyperbolic GDS $\tau(f) = 1/f$, $f > 0$. Note that when the signal GD does *not* match the dispersive GDS $\tau(f)$ of the QTFR, significant distortion may occur that could impede analysis [208, 209]. This is demonstrated by analyzing the impulse using the hyperbolic QD in Figure 1.7, and the hyperbolic TM signal using Cohen's WD in Figure 1.7.

1.3.5.2.2 Power classes When the affine class is warped with $\xi(f/f_r) = \text{sgn}(f)|f/f_r|^\kappa$ and $\tau(f) = (\kappa/f_r)|f/f_r|^{\kappa-1}$ in Equation (1.19), the resulting QTFRs

are members of the κ th power class, $\kappa \neq 0$ [1, 92, 111, 178, 179, 186]. The covariance in Equation (1.10) simplifies to the κ th power GDS covariance in Equation (1.8), and the scale covariance of the affine class is preserved. The κ th power GDS covariance makes the κ th power class QTFRs useful for analyzing signals whose TF localization is related to the κ th power law geometry in the TF plane. In addition, power QTFRs can be used in multiresolution analysis applications because they are scale covariant. Power QTFR members include the power WD, the power Bertrand P_0 -distribution, and the powergram [179]. The power WD is a warped version of the WD given as:

$$\begin{aligned} P_\kappa \text{WD}_X(t, f) = & |f/\kappa| \int_{-\infty}^{\infty} X(f \xi_\kappa^{-1}(1 + \beta/2)) X^*(f \xi_\kappa^{-1}(1 - \beta/2)) e^{j2\pi \frac{tf}{\kappa}\beta} \\ & \times |1 - \beta^2/4|^{\frac{1-\kappa}{2\kappa}} d\beta \end{aligned} \quad (1.21)$$

where $\xi_\kappa^{-1}(f/f_r) = \text{sgn}(f)|f/f_r|^{1/\kappa}$. As a result, it satisfies many desirable properties including energy conservation and power GD localization. Just like the WD, the power WD suffers from CTs when used to analyze multicomponent signals. These CTs can be reduced by smoothing the power WD even though some of the desirable properties of the power WD are no longer satisfied. Unlike the WD, the power WD does not have inner CTs when a signal has power GD characteristics [111]. The P_κ -distributions [165, 166] satisfy the κ th power GDS covariance of Equation (1.8) in addition to the scale covariance and the constant time shift covariance. Hence, they are simultaneously members of the affine class and the κ th power class.

Note that the affine class is actually the power class with $\kappa = 1$. When $\kappa = 2$, the corresponding power class is the *linear chirp class* that is well matched to signals with linear TF characteristics. Two QTFRs from the linear chirp class are the linearly warped WD (LWD) and the chirrogram. These are obtained when the WD and the spectrogram, respectively, are warped as in Equation (1.19) with quadratic characteristic function $\xi(f/f_r) = \text{sgn}|(f)/f_r|^2$. The LWD provides high localized representations when analyzing linear TM signals. Note, however, that the WD also satisfies this property as it is matched to any TF structures that do not curve. On the other hand, the chirrogram has a definite TF resolution advantage over the spectrogram when analyzing multicomponent signals with linear characteristics. This is because the smoothing operation of the chirrogram is performed along lines of any slope in the TF plane whereas the smoothing of the spectrogram is only along horizontal or vertical lines [209].

Power QTFRs are ideal for signals that propagate through linear systems with specific power GD characteristics such as when a wave propagates through a dispersive medium [111]. Other signals that are matched to κ th power QTFRs include the dispersive propagation of a shock wave in a steel beam ($\kappa = 1/2$) [210, 211]; transionospheric signals measured by satellites ($\kappa = -1$) [212]; acoustical waves reflected from a spherical shell immersed in water [213]; some cetacean mammal whistles [206, 214], and diffusion equation based waveforms ($\kappa = 1/2$) [215] (e.g., waves from uniform distributed RC transmission lines [216]). Furthermore, power laws can be used to roughly approximate more complicated GD functions.

1.3.5.3 Implementation algorithms

Dispersive QTFRs could be computationally intensive when implemented directly using numerical integration as in Equation (1.21). An alternative implementation scheme is based on the warping formulation in Equation (1.19) that allows the use of existing efficient algorithms for computing Cohen’s class or affine class QTFRs. Such algorithms were used in [217] for hyperbolic QTFRs, and in [111, 178] for power QTFRs. Specifically, for computing dispersively warped versions of the WD, the algorithms consist of three steps as demonstrated in Equation (1.19): (1) warp the signal $X(f)$, (2) compute the WD of the warped signal and (3) transform the TF axis according to $(t, f) \rightarrow (\frac{t}{f_r \tau(f)}, f_r \xi(f/f_r))$. Note that the increased computational complexity of the dispersive QTFRs is the trade-off for the improved performance in analyzing signals with matched dispersive GD characteristics.

1.4 Matched Time–Frequency Processing for Various Applications

For successful TF analysis, it is advantageous to match the specific time shift of a QTFR in Equation (1.10) with (changes in) the GD of the signal. In some applications, signals with known GD, $\tau(f) = \frac{d}{df} \xi(f/f_r)$, need to be processed. As a result, a matched QTFR can be designed as in Equation (1.19) with a characteristic function $\xi(f/f_r)$. When the signal GD is not known *a priori*, some preprocessing is necessary before designing a well-matched QTFR. A rough GD estimate can be obtained by fitting a curve through the spectrogram of the signal or by using one of the many proposed algorithms to estimate GD (or IF) characteristics [218–221]. Because the phase function of the signal needs to be one-to-one for designing its matched QTFR by appropriately warping the WD or its smoothed versions, approximations of the GD function can also be used. Next, we demonstrate the importance of matching TV signals with QTFRs using various simulations.

1.4.1 Constant and linear time–frequency structures

The SPEC is most often used in applications due to its ease of implementation and its relation to the fast FT. Just like the WD, it is well matched to signals with constant TF support. However, because the SPEC is inherently either a short-time window technique or a narrowband filtering technique, its TF resolution depends on whether the SPEC window is matched to the signal structure [135]. Figure 1.8 shows that the SPEC can cause distortion when used to analyze two closely spaced LFM chirps and its window does not match the signal TF structure. The WD of this signal is shown in Figure 1.8 to yield very good TF localization but with CTs between the two signal components. In Figure 1.8, a multiform, tiltable exponential distribution [175] is used to reduce the CTs while preserving the WD TF localization. Figure 1.8

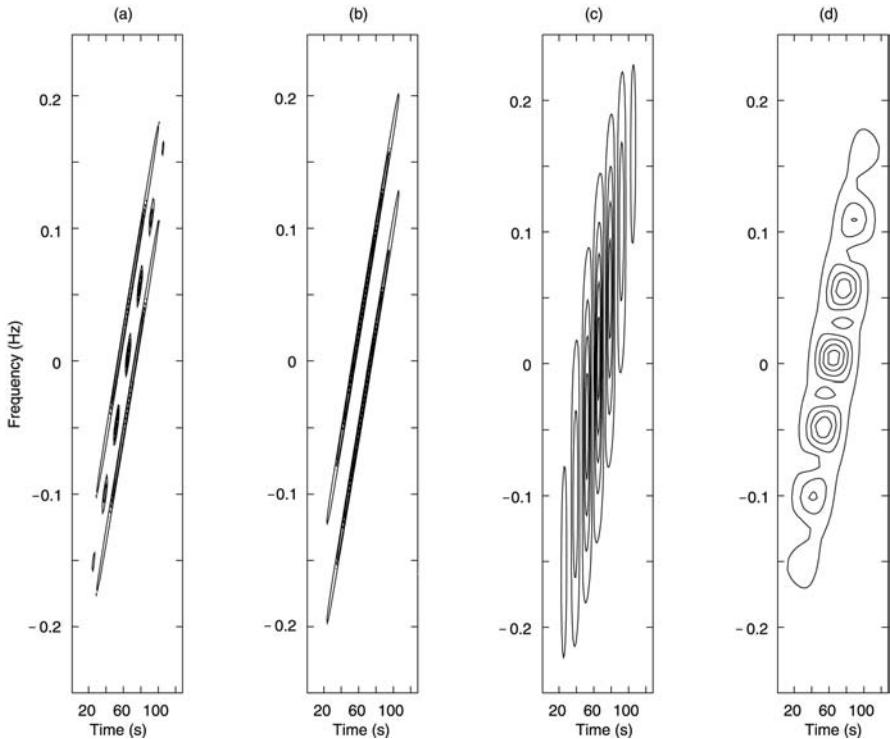


FIGURE 1.8

The analysis signal is the sum of two closely spaced LFM chirps. (a) WD, (b) multiform, tiltable exponential distribution, (c) spectrogram with a short duration window (wideband analysis) and (d) spectrogram with a long duration window (narrowband analysis). The sampling frequency used is 1 Hz.

computes the SPEC using a wideband Hanning window resulting in a temporal localized representation that could be misinterpreted as the representation of seven short duration transients thus leading to erroneous information. The narrowband analysis in Figure 1.8 using a SPEC with a long-duration Hanning window gives the misleading impression that several tones, well-localized in frequency, are present in the signal. Thus, use of the SPEC with windows that are not matched to the signal TF structure in this example can cause distortion in the TF plane [135].

1.4.2 Constant and hyperbolic time–frequency structures

To demonstrate the possible mismatch between constant and hyperbolic TF analysis, we consider a simulated signal that has three hyperbolic TM components, $X_i(f) = e^{-j2\pi c_i \ln f} / \sqrt{f}$ where $c_i = 3 + 4i$ for $i = 0, 1, 2$ and $f > 0$ — the signal in Equation (1.2) with $f_r = 1$ Hz and $\xi(f) = \ln f$ — and two TF-shifted Gaussian components. Only hyperbolic QTFRs (covariant to $1/f$ hyperbolic GDS) like the smoothed QDs match the hyperbolic components whereas only Cohen's QTFRs (covariant to

constant GDS) like the smoothed WDs match the Gaussian components (with constant support in the TF plane). [Figure 1.9\(a\)](#) shows the *ideal* signal representation for comparison with the QTFRs. The ideal representation can be obtained as the sum of the IFs of each signal component that, unless a method exists to separate the various signal components, is in general unattainable. Both the WD in Figure 1.9(b) and the QD in Figure 1.9(c) suffer from oscillatory CTs between any pair of the five signal components. Figure 1.9(d) shows that the SPWD is not as successful at removing oscillatory CTs. The smoothed pseudo QD (SPQD) in Figure 1.9(e) is well matched to the three hyperbolic TM signals because the smoothing window is matched to their TF structure and is thus better at reducing CTs. Due to local smoothing with negative CTs, the amplitude of the hyperbolic TM signals in the SPWD and that of the Gaussian signals in the SPQD are significantly reduced. Also, there is distortion in the TF support of the smoothed QTFRs: the hyperbolic components in the SPWD now appear constant and the Gaussian components in the SPQD are no longer vertical.

1.4.3 Constant and exponential time–frequency structures

In the following example, we demonstrate the advantage of using exponential class QTFRs to analyze signals with exponential GD. The test signal consists of the sum of three exponential TM signals, $X(f) = \sqrt{\kappa} e^{\kappa f/2} (e^{-j2\pi 0.3e^{\kappa f}} + e^{-j2\pi 1.8e^{\kappa f}} + e^{-j2\pi 3.8e^{\kappa f}})$ for $\kappa = 2.3$. Note that each signal component corresponds to the signal in Equation (1.2) with $f_r = 1$ Hz and $\xi(f) = e^{\kappa f}$. [Figure 1.10](#) shows the ideal TFR of this signal consisting of the sum of the GDs of each of the three components. [Figure 1.11](#) compares six different QTFRs of this signal: (1) two affine QTFRs, the WD and the ASPWD; (2) two exponential QTFRs, the exponential WD (EWD) and the exponential smoothed pseudo-WD (ESPWD) with $\kappa = 2.3$; and (3) the EWD and ESPWD with $\kappa = 1$. Note that κ th exponential QTFRs are well matched for analyzing exponential TM signals because the exponential GDS of the QTFRs is matched to the exponential GD of the signal provided that the power parameter κ is the same for both the QTFRs and the signals. This is demonstrated in Figure 1.11(c) where the EWD with $\kappa = 2.3$, chosen to match the power parameter κ of the signal, shows well-concentrated CTs along the exponential GD curve. On the other hand, the WD and the EWD with $\kappa = 1$ have more complicated CT structures in Figure 1.11(a) and 1.11(e), respectively. By using a smoothing component along the exponential GD curve, the ESPWD with $\kappa = 2.3$ reduces the CTs with only moderate loss of TF resolution in Figure 1.11(d). However, because the ASPWD smoothing is not matched to the signal exponential TF characteristics, it does not succeed in removing all the CTs in Figure 1.11(b). The ESPWD with $\kappa = 1$ in Figure 1.11(f) removes more CTs than the ASPWD in Figure 1.11(b). However, the autoterms of the ESPWD with $\kappa = 1$ are distorted because its exponential GDS is not matched to the $\kappa = 2.3$ exponential TF characteristics of the signal, and a mismatch occurs. Note, however, that choosing the wrong parameter κ of the ESPWD in Figure 1.11(f) yields better results than choosing an affine QTFR that is covariant to constant time shifts as in Figure 1.11(b).

1.4.4 Constant and power time–frequency structures

To demonstrate the advantage of power class QTFRs when the analysis signal has power dispersive characteristics, consider a windowed version of the TM signal,

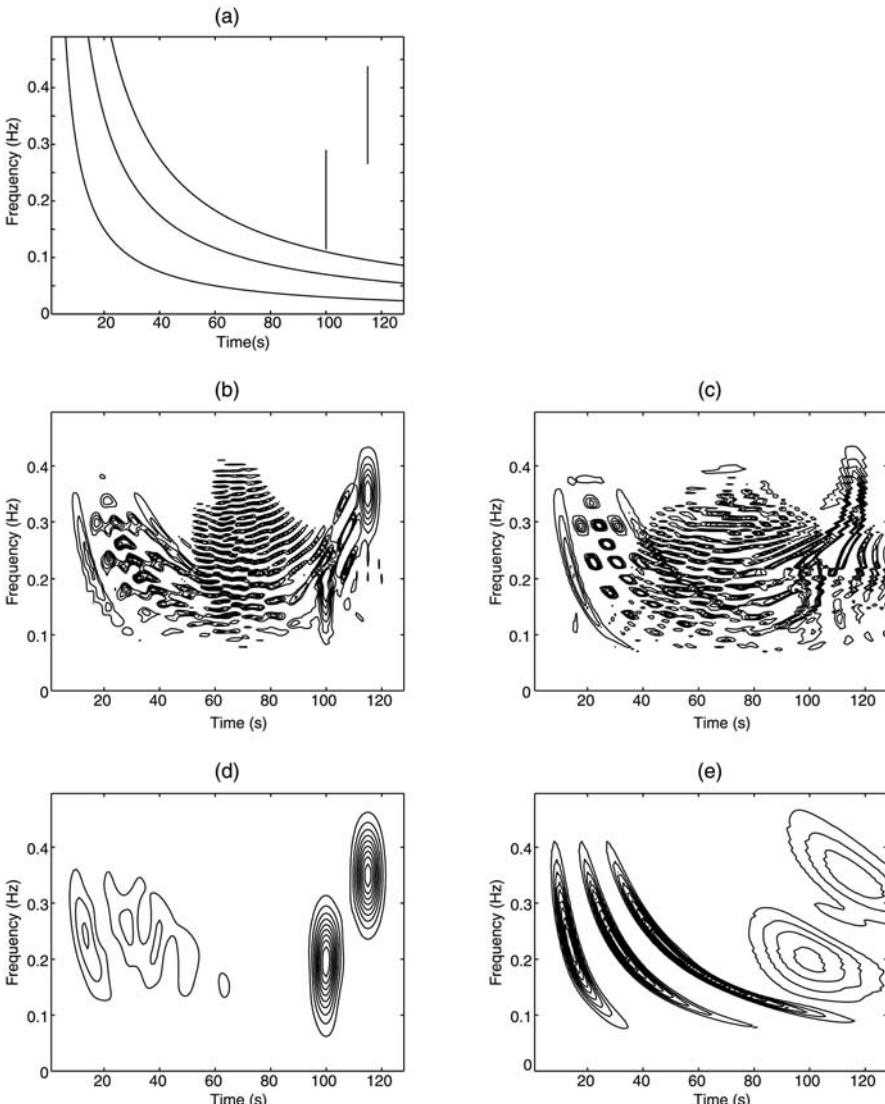
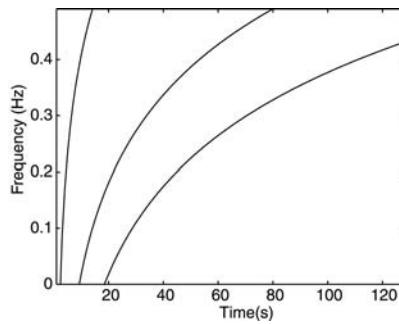
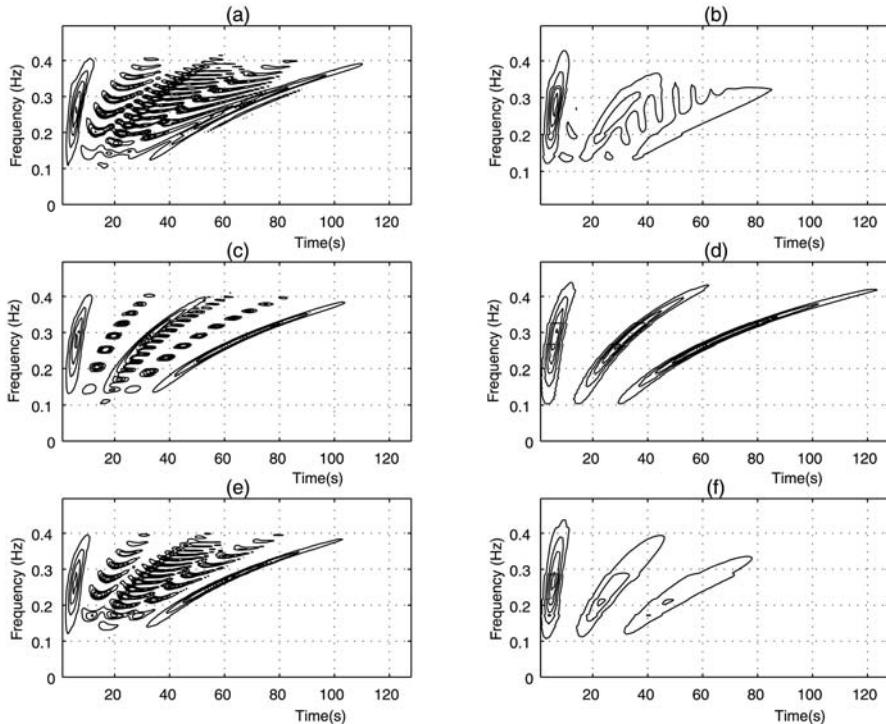


FIGURE 1.9

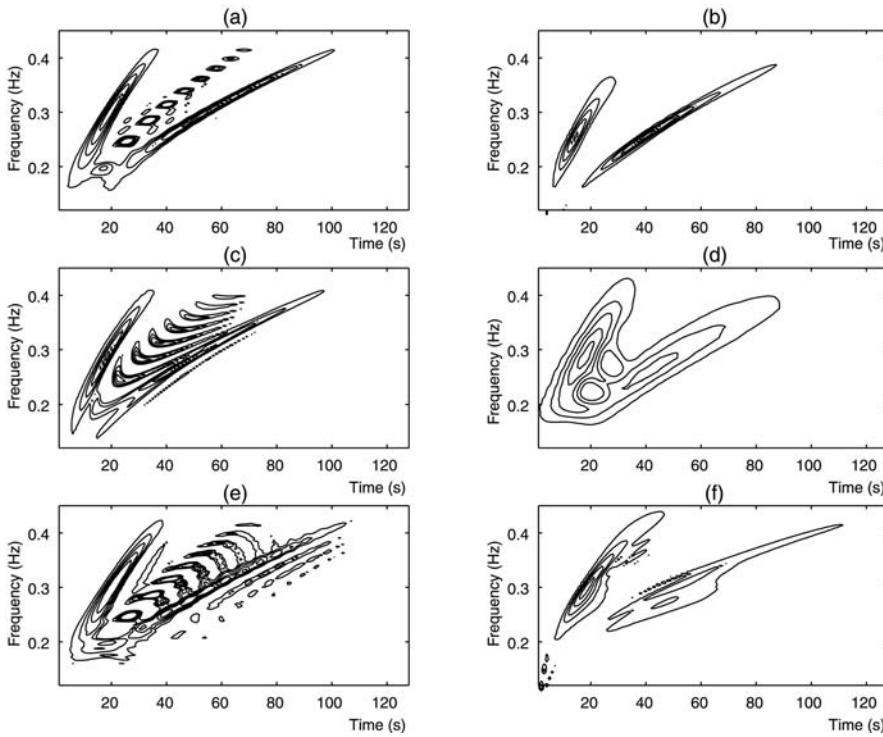
The analysis signal is the sum of three hyperbolic TM signals and two TF shifted Gaussian signals. (a) Ideal TFR, (b) WD, (c) QD, (d) SPWD and (e) SPQD. Note that a mismatch occurs between the hyperbolic GD and the WD constant time shifts in (b) and (d), and between the Gaussian component constant GDs and the QD hyperbolic GDSs in (c) and (e). The sampling frequency used is 1 Hz.

**FIGURE 1.10**

The *ideal* TFR of three exponential TM signals with $\kappa = 2.3$ consists of the sum of the GD function of each signal component.

**FIGURE 1.11**

The analysis signal is the sum of three exponential TM signals with $\kappa = 2.3$. On the right-hand side are QTFRs that are smoothed pseudo versions of the QTFRs on the left-hand side. The first row shows QTFRs from the affine class, the second row shows QTFRs from the exponential class with $\kappa = 2.3$ and the third row shows QTFRs from the exponential class with $\kappa = 1$. (a) WD, (b) ASPWD, (c) EWD with $\kappa = 2.3$, (d) ESPWD $\kappa = 2.3$, (e) EWD with $\kappa = 1$ and (f) ESPWD with $\kappa = 1$. The sampling frequency used was 1 Hz.

**FIGURE 1.12**

QTFR analysis of a signal consisting of two windowed power TM components with power parameter $\kappa = 3$ and sampling frequency 1 Hz. (a) Power WD with $\kappa = 3$, (b) PSPWD with $\kappa = 3$, (c) WD (power WD with $\kappa = 1$), (d) ASPWD (PSPWD with $\kappa = 1$), (e) power WD with $\kappa = 4$ and (f) PSPWD with $\kappa = 4$.

$X(f) = \sqrt{3f^2} e^{-j2\pi cf^3}$, that is given by Equation (1.2) with $f_r = 1$ Hz, $\xi(f) = f^3$ and $f > 0$. Figure 1.12(a) and (b) shows the results obtained with the power WD and a power smoothed pseudo-WD (PSPWD) with a very short analysis window [111]. Both QTFRs have power parameter $\kappa = 3$ that is matched to the power parameter of the signal. The power WD in Figure 1.12(a) has very good TF concentration but large CTs [178]. These CTs are effectively suppressed in the PSPWD in Figure 1.12(b) with hardly any loss of TF concentration. Also shown — in Figure 1.12(c) and (d) — are the results obtained with the WD and the ASPWD that are both members of the affine class (which is the power class with $\kappa = 1$). The WD in Figure 1.12(c) is not matched to the power signal, displaying complicated CTs. The ASPWD in Figure 1.12(d) does not remove all the CTs and has a larger loss of TF concentration than does the PSPWD in Figure 1.12(b). Even though all QTFRs in Figure 1.12 are scale covariant, the results of the two QTFRs with $\kappa = 3$ in Figure 1.12(a) and (b) are better than those of the corresponding two affine QTFRs in Figure 1.12(c) and (d) because the former two are optimally matched to the $\kappa = 3$ power law GD of the signal.

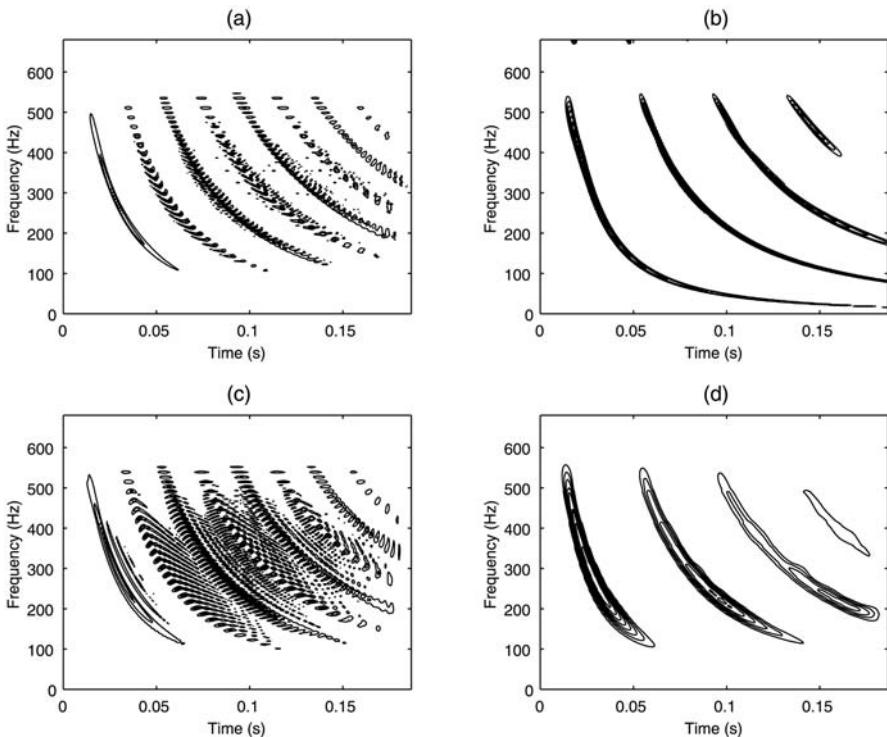


FIGURE 1.13

Power class analysis of a bandpass-filtered segment of the measured impulse response of a steel beam. (a) Power WD with $\kappa = 0.35$, (b) PSPWD with $\kappa = 0.35$, (c) WD and (d) ASPWD. The sampling frequency of the data is 4096 Hz.

Figure 1.12(e) and 1.12(f) shows the results obtained when analyzing the preceding signal using the power WD and the PSPWD with power parameter $\kappa = 4$ (which is different from the power parameter of the signal). The PSPWD in Figure 1.12(f) has better CT removal and better TF concentration along the true GD than the ASPWD in Figure 1.12(d) because the power parameter mismatch in Figure 1.12(f) is smaller than in Figure 1.12(d) [178, 208].

1.4.5 Power time–frequency structures with real data

To demonstrate the use of power QTFRs in the analysis of real data signals, Figure 1.13 shows two power QTFRs with $\kappa = 0.35$ and two affine QTFRs of the measured impulse response of a steel beam with rectangular cross section** [210, 211]. The impulse response was obtained by lightly tapping one end of the steel beam in the

**This impulse response was obtained from D. Newland and J. Woodhouse. It was measured in a laboratory experiment that the latter conducted at Cambridge University.

direction orthogonal to the flat side of the beam. Bending waves traveled along the beam until they were reflected at the free end back and forth thereby producing a series of echoes with increasing dispersion. The QTFRs in Figure 1.13 display a bandpass-filtered segment of the impulse response with duration of 190 ms, bandwidth of 440 Hz and sampling frequency of 4096 Hz. The PSPWD with $\kappa = 0.35$ in Figure 1.13(b) shows better resolution and CT suppression than the other three QTFRs. Note that the specific value of $\kappa = 0.35$ was chosen empirically to match the TF curvature of the primary reflection [111].

1.5 Applications in Time–Frequency Signal Processing

TFRs have been extensively used in many applications to process (analyze, estimate, detect, classify, characterize or model) TV signals and systems. As a result, many papers have been published in different research areas to demonstrate the successful results obtained when a matching TFR was used to process an application-dependent signal or system. Next, we present some of these applications and provide the references of some relevant papers.^{††}

1.5.1 Time-varying signal processing for wireless communications

1.5.1.1 Time-varying channel characterization

TF techniques have been successfully applied in wireless communication applications to improve system performance and to meet major challenges that arise from the escalating demand for new higher data rate wireless technologies. In particular, to characterize the TV nature of the channel, especially when its spectral characteristics vary rapidly with time, one benefits from using TF methods to describe how the spectral content of a system changes with time [87]. This characterization provides important information about different types of signal distortion that the channel causes such as multipath and Doppler spread, and it is based on TF symbol and spreading function representations [30, 55–57, 60, 75, 205, 222–227]. The spreading function provides an important interpretation of a narrowband channel output as a weighted superposition of multipath and Doppler shifts on the input signal [55, 56, 87, 224]. Its 2-D FT, the Weyl symbol [55, 56, 224], can be interpreted as the transfer function of a TV narrowband system. Specifically, the Weyl symbol is a TFR that is used to characterize TV systems as well as nonstationary random processes [56, 224].

^{††}Although many other papers have been published in these areas, we cannot list all of them due to space limitations.

1.5.1.2 TF receivers and channel diversity

When a wireless communications channel spreads the transmitted signal both in time and frequency, the resulting distortion can affect detection performance at the receiver [39]. In practice, broadband signals such as spread-spectrum waveforms are used to reduce these TV effects, but they do not fully exploit the advantage of using 2-D models because they do not consider time and frequency spread jointly [61]. The TF formulations exploit the 2-D nature of the wireless channel, and can allow for alternative detector formulations for single or multiuser systems [53, 65, 66, 76, 77, 228–235]. Some of these receiver structures can be shown to improve diversity over fast varying frequency-selective channels. For example, it was shown in [77] that a TF RAKE receiver in a DS-CDMA system outperformed the traditional RAKE receiver designed for multipath diversity [79] and provided diversity in both multipath and Doppler.

1.5.1.3 Time-varying modulation schemes

LFM chirps are used in radar and sonar applications due to their important properties. Specifically, LFM chirps provide accurate estimates for range and Doppler [87], and they have an inherent immunity against Doppler and multipath fading [89]. These TV broadband signals have also been used in communication applications because they are resistant against jamming and they do not require synchronization [90]. For example, LFM chirps have been used for TV broadband modulation [236] and have been shown to perform better than narrowband modulation schemes [54, 68, 70, 73, 74, 88, 89, 237].

1.5.1.4 Time-varying jamming interference mitigation

When the intentional or unintentional interference is TV, such as an LFM chirp jammer, TFRs have been used to suppress it. The evolutionary spectrum and the Hough transform were used to estimate the jammer IF in [238, 239], and adaptive TFR techniques, such as the adaptive Wigner–Hough transform, were adaptively used in [45, 240]. In [44], WD synthesis techniques were used to estimate the LFM jammer (see [Figure 1.3](#)); in [42], TFR-based adaptive filtering was used to suppress both LFM and frequency hopping jammers; and in [43], IF-based interference mitigation techniques were used. The short-time FT was used in [241], and subspaced projection techniques were applied in [242, 243] as precorrelation methods for suppressing linear FM interference in global positioning system (GPS) receivers. The authors in [47, 244] based their mitigation of LFM chirp interference on the chirplet TF decomposition, whereas [46] used similar expansions based on adaptive overdetermined frame representations. The fractional FT that is well matched to LFM chirps was applied in [48], and the LMST transform was successfully used in [2, 49] as shown in [Figure 1.4](#). Note that TF techniques have also been used to reject narrowband interference [245].

1.5.2 Biomedical signal processing

TF analysis techniques have been used in many biomedical applications as extensively documented in [246] and in the book titled *Time Frequency and Wavelets in Biomedical Signal Processing* [37]. Other overview manuscripts on biomedical applications of TFRs include [247–253]. TF methodologies have been used to analyze and synthesize aortic and pulmonary components of the second heart sound [254]; to characterize atrial fibrillation in the surface ECG [255]; to detect ischemia-related depolarization changes in the QRS complex of ECG waveforms [256]; to design noninvasive techniques for early detection of focal brain ischemic injuries [257]; to extract features from vibroarthrographic signals emitted by human knee joints [258]; to detect electroencephalogram (EEG) seizures (such as in newborns) [259–263] or other processing of EEG signals [248, 264–271]; and to design automatic diagnostic systems for breast cancer [272]. Various biomedical waveforms were analyzed using different types of TFRs including electromyography (EMG) potentials waveforms [273–277]; temporomandibular joint (TMJ) clicking sounds [278–280]; blood flow [281–285]; heart murmurs [286, 287]; phonocardiogram sounds [288, 289]; nuclear magnetic resonance (NMR) signals [290, 291]; electrogastrographic signals [292]; electrocorticogram signals in temporal lobe epilepsy [293]; electroenterogram signals [294]; otoacoustic emissions [295–297]; fatal breathing movements [298]; postural sway responses [299]; and heart valve sounds [300].

1.5.3 Biological signal processing

Most biological signals are TV and TFRs have been used to study their properties and to analyze or classify them [301–303]. For example, bats use acoustical energy to detect their targets in a manner similar to sonar signal processing because they compare their emission with its echo to determine the range and shape of their target. It is thought that the bat's auditory system may be doing a frequency-to-time transformation to sort overlapping echoes and different studies have shown that the bat sonar signals have hyperbolic TF characteristics [207]. Bat sonar signals were processed using TF techniques in [304–308]. Other biological signal TF processing was applied to marine mammal sounds in [206, 309–314], dog heart sounds in [315, 316] and cat auditory evoked potentials in [317].

1.5.4 Detection, estimation and classification

In most of the previously mentioned applications, we often need to detect, estimate or classify TV signals. Although 1-D techniques could be applied, the TV nature of the signals calls for TF techniques to achieve higher performance. It was shown in [67] that some TF detectors are equivalent to 1-D classical detectors (e.g., a 1-D optimal matched filter vs. a 2-D optimal WD-based matched filter), whereas others can be obtained based on Cohen's class TF shift covariant QTFRs. Although some of these detectors are not optimal, the corresponding receiver structures may be more advantageous in different situations such as when the received signal has unknown parameters. In the presence of various forms of distortion, 2-D detectors matched to

the transmitted TV signal may still demonstrate localization in the TF plane whereas 1-D (time or Fourier-based) detectors may not. In such a case, the detection could be performed by 2-D thresholding techniques in the TF plane. This is particularly useful for noncoherent receivers with unknown system parameters because the parameters could be extracted directly from the localized QTFRs. For example, if the WD is used to detect an LFM chirp [318], the receiver structure is quite simplified because the WD of a chirp is a line in the TF plane. Furthermore, if the LFM chirp has unknown TF shift parameters, then, because the WD is covariant to TF shifts, it can be used to first estimate the delay in time and frequency by finding the new location of the chirp in the TF plane.

An important consideration for designing TF receivers in a particular application is that the TFR used ought to match the transmitted signal. As a result, when detecting TV signals with nonlinear GD, the characteristics of the 2-D receivers should also have nonlinear characteristics for compatibility [203, 319]. For example, when detecting a hyperbolic TM signal in AWGN, a detector can be designed using the QD. By assuming that the TM rate is unknown, then the estimator–correlator detector [320, 321] first estimates the TM rate by integrating the QD over hyperbolic curves in the TF plane and then keeping the maximum result. The hyperbolic TM signal with the estimated TM rate could be detected by comparing the correlation result with some threshold [203, 322]. Dispersive TV signals in noise can also be estimated and detected using the 1-D MST transforms. For example, it was shown in [108] that signals with dispersive IF such as power FM signals can be estimated and detected (when completely unknown) using the power MST. This method provides a higher probability of detection when compared with the energy detector that is used when the signal is completely unknown.

Other types of TF detectors have also been designed and published in the literature [13, 30, 63, 154, 319, 323–332]. Some applications that use TF detectors include the following: structural damage or fault detection [333–337], speech events and pitch detection [338–344], radar detection [345–353], detection of underwater signals [354, 355], mines detection [356], seismic detection [357–359], detection of knocks in car engines [360, 361], detection of abrupt changes [362, 363], detection of venous air embolism [364], detection of echoes [365], rotating machinery monitoring [366], detection of radio-astronomical observations [367], frequency tracking [368] and detection of vibrational signals [369]. TF detectors have also been applied to improve the performance of wireless communication receiver systems for fast fading, frequency-selective channels [65, 76, 228–235].

TF methods are also important in the estimation [13, 30, 108, 218, 319, 326, 370–377] and classification of TV signals [13, 152, 378–389]. For example, in wireless communications, TF methodologies were used to estimate the channel under various distortion conditions [390–392], the parameters of DSSS signals [373] and the parameters of multiuser OFDM synchronization [393]. TF estimation was also used in speech [394], in radar [395], in sonar [396], for sensor arrays [397], for motion parameter estimation [398, 399], for acoustic emission of Formula 1 engines [400], for mobile velocity estimation [401] and for IF estimation [218–221, 374, 402, 403]. Classification TF techniques were used for radar signals [404–406], for music

[407–411], for minelike targets [412, 413], for underwater sonar sounds [414–417], for human physiological data [418], for tool-wear monitoring [419] and for speech [344, 420]. Note that other references on TF detection, estimation and classification can be found in biomedical and biological applications.

1.5.5 Radar and sonar processing

TFRs have been used extensively in radar and sonar applications. In both fields, due to the nature of the TV mediums, LFM chirps are used for processing as they are bandwidth efficient. As a wideband TV signal, the LFM chirp and its TF delayed echoes can be successfully processed in the TF plane. For example, constant time shifts are important in shallow water sonar signal processing. Due to boundary interactions, a bottom bounce path may be received several milliseconds after the direct path. Thus, a QTFR analyzing the received signal must preserve the delay associated with the difference in path lengths. As we have seen earlier, Cohen's class QTFRs like the WD preserve these delays. In radar, when an LFM chirp is transmitted, its distorted reflection could have a different chirp rate. As a result, the processing QTFR should preserve possible shifts in the LFM rate. Examples of such QTFRs include the linear chirp class QTFRs such as the LWD and the chirrogram.

In [20, 421], TF processing methods are presented that are specifically adapted to radar applications. In particular, these methods look at signature diagnostics and automatic target recognition radar applications such as target feature extraction from the radar signal [20, 422]. Other applications of TFRs in radar include ISAR imaging and motion compensation [423, 424], SAR imaging of moving targets [425, 426], classification of radar signals [427] and analysis of radar signals [20, 424, 428–432]. In sonar processing, TF processing has been used for beamforming [433], matched field processing [434], acoustic transients sonar processing [435–438] and acoustic backscatter classification [416].

1.6 Concluding Remarks

The large amount of published scientific literature testifies to the fact that TF processing is essential for the ideal characterization of TV signals and systems. Because many different types of TF techniques exist, it is important to search for the one that is most pertinent to the present application. Although the WD and the spectrogram QTFRs are often the easiest to use, they do not always provide an accurate characterization of the real data. The spectrogram, for example, could be used to obtain an overall characterization of the TV signal structure, and then the information could be used to invest in another QTFR that is well matched to the data for further processing that requires information that is not provided by the spectrogram.

Section 1.1.1 states that this chapter attempts to provide a response to the question: *Which (TF) technique does one use and why? How does one use and implement the*

technique? Concentrating on various quadratic TFRs, we discuss their usefulness, providing definitions, important properties, implementation methods and simulation examples that emphasize the importance of matching the technique to the analysis signal. In general, the choice of a QTFR in a particular application depends on many factors such as the relevance of the properties satisfied by the QTFR, the computational cost and speed of the QTFR, and the trade-off in using the QTFR. For example, if an application requires that no information is lost after the signal is processed, then the chosen QTFR must preserve that information (such as using a unitary QTFR to preserve energy information). If the real data to be processed has a specific nonlinear GD or is expected to propagate in a medium with nonlinear GD, then the chosen QTFR must preserve shifts of the nonlinear GD curves in the TF plane. If these shifts are by a constant amount of time and frequency, then Cohen's class QTFRs are an adequate choice. On the other hand, hyperbolic class QTFRs are preferred when the shifts are along hyperbolic GD curves. If different resolution is required for different frequencies, then scale covariant QTFRs such as affine QTFRs can provide an adequate match. If no properties are important other than to obtain high localization along the signal IF, then adaptive QTFRs could be chosen. The reassignment TFR output, for example, follows the IF curve of a nonlinear TV signal to provide high resolution. When processing multicomponent signals in practical applications, smoothed QTFRs are preferred as they reduce the effect of unwanted cross terms. However, to avoid TF distortion, it is important to match the signal TF dependence not only with the QTFR characteristics but also with the smoothing characteristics of the QTFR. For example, when analyzing the whistles of a school of dolphins and the whistle frequencies increase with time in a power fashion [206], then a smoothed power QTFR could be selected because the nonlinear smoothing can follow the power IF curve in the TF plane.

The mismatched examples provided in Section 1.4 demonstrate that if the real data consists of components with various TF structures, then it is difficult to obtain an ideal QTFR to match every single component. In most applications though, there is a dominant TF structure that needs to be extracted from a signal that has multiple TF characteristics. For example, in monitoring acoustic emissions from concrete structures, the warning signals that indicate future catastrophic failures have a distinct signature that is different from all other possible acoustic sounds in the vicinity. In such cases, a QTFR must be chosen that is well matched to the desirable TF signatures. The chosen QTFR could be such that the other components suffer only a minimum loss of TF resolution or, on the contrary, such that information about the other components is mostly lost because it is of no interest to the application.

Acknowledgments

The author would like to express her gratitude to those who contributed to this work. Specifically, many thanks go to her graduate students Ms. Moushumi Zaman and Mr. Srinivasa Kalyan Machineni for their help in completing the bibliography,

Mr. Dongwon Park for his help with some of the simulations and Ms. Shen Hao for proofreading the chapter. This work was supported in part by the National Science Foundation grants CCR-0134002 and EIA-0074663.

References

- [1] A. Papandreou-Suppappola, Generalized Time-Shift Covariant Quadratic Time-Frequency Representations with Arbitrary Group Delays, in Proceedings of the 29th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, October 1995, pp. 553–557.
- [2] S. Chetwani and A. Papandreou-Suppappola, Time-Varying Interference Suppression in Communication Systems Using Time-Frequency Signal Transforms, in Proceedings of the 34th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, October 2000, pp. 460–464.
- [3] E.P. Wigner, On the quantum correction for thermo-dynamic equilibrium, *Phys. Rev.*, 40, 749–759, 1932.
- [4] J. Ville, Théorie et applications de la notion de signal analytique, *Câbles et Transmission*, 2A, 61–74, 1948 (translated into English by I. Selin, RAND Corporation Report T-92, Santa Monica, CA, August 1958).
- [5] W. Koenig, H.K. Dunn and L.Y. Lacy, The sound spectrograph, *J. Acoust. Soc. Am.*, 18, 19–49, 1946.
- [6] J.B. Allen and L.R. Rabiner, A unified approach to STFT analysis and synthesis, in Proceedings of the IEEE, 65, 1558–1564, November 1977.
- [7] L. Cohen, Generalized phase-space distribution functions, *J. Math. Phys.*, 7, 781–786, 1966.
- [8] E.P. Wigner, Quantum-mechanical distribution functions revisited, in *Perspectives in Quantum Theory*, W. Yourgrau and A.V. der Merwe, Eds., Dover Publications, New York, 1971.
- [9] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, The Wigner distribution — A tool for time-frequency signal analysis, Part I. Continuous-time signals, *Philips J. Res.*, 35, 217–250, 1980.
- [10] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, The Wigner distribution — A tool for time-frequency signal analysis, Part II. Discrete-time signals, *Philips J. Res.*, 35, 276–300, 1980.
- [11] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, The Wigner distribution — A tool for time-frequency signal analysis, Part III. Relations with other time-frequency signal transformations, *Philips J. Res.*, 35, 372–389, 1980.

- [12] F. Hlawatsch and P. Flandrin, The interference structure of the Wigner distribution and related time-frequency signal representations, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W.F.G. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997.
- [13] R.A. Altes, Detection, estimation, and classification with spectrograms, *J. Acoust. Soc. Am.*, 67, 1232–1246, April 1980.
- [14] B. Boashash, Ed., *Time-Frequency Signal Analysis — Methods and Applications*, Longman-Cheshire, Melbourne, Australia, 1992.
- [15] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.
- [16] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999 (translated from French, *Temps-fréquence*, Hermès, Paris, 1993).
- [17] W.F.G. Mecklenbräuker and F. Hlawatsch, Eds., *The Wigner Distribution — Theory and Applications in Signal Processing*, Elsevier, Amsterdam, 1997.
- [18] S. Qian and D. Chen, *Joint Time-Frequency Analysis: Methods and Applications*, Prentice Hall, Upper Saddle River, NJ, 1996.
- [19] R. Carmona, H.L. Hwang, and B. Torrésani, *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms with an Implementation in S*, Academic Press, New York, 1998.
- [20] V.C. Chen and H. Ling, *Time-Frequency Transforms for Radar Imaging and Signal Analysis*, Artech House, 2002.
- [21] K. Grochenig and K.U. Grochenig, *Foundations of Time-Frequency Analysis*, Applied Numerical and Harmonic Analysis, Birkhäuser, Basel, 2000.
- [22] S. Qian, *Introduction to Time-Frequency and Wavelet Transforms*, Prentice Hall, New York, 2001.
- [23] M.D. Riley, *Speech Time-Frequency Representation*, Dordrecht, Kluwer Academic, 1988.
- [24] L. Cohen and P. Loughlin, Eds., *Recent Developments in Time-Frequency Analysis*, Vol. 9, Kluwer Academic, Dordrecht, 1998 (Special Issue of Multi-dimensional Systems and Signal Processing).
- [25] B. Boashash, Ed., *Time-Frequency Signal Analysis and Processing*, Prentice Hall, New York, 2002.
- [26] L. Debnath, Ed., *Wavelet Transforms and Time-Frequency Signal Analysis*, Birkhäuser, Boston, 2001.
- [27] R. Tolimieri, M. An, and J.J. Benedetto, Eds., *Time-Frequency Representations*, Applied and Numerical Harmonic Analysis Series, Birkhäuser and Springer Verlag, 1998.

- [28] A. Mertins, *Signal Analysis: Wavelets, Filter Banks, Time-Frequency Transforms and Applications*, John Wiley & Sons, New York, 1999.
- [29] I. Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, 1992.
- [30] F. Hlawatsch, *Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation*, Kluwer, Boston, 1998.
- [31] C.K. Chui, *An Introduction to Wavelets*, Academic Press, New York, 1992.
- [32] S.G. Mallat, *Wavelet Tour of Signal Processing*, 2nd ed., Academic Press, New York, 1999.
- [33] H.G. Feichtinger and T. Strohmer, Eds., *Gabor Analysis and Algorithms: Theory and Applications*, Springer, 1998.
- [34] H. Feichtinger and T. Strohmer, Eds., *Advances in Gabor Analysis*, Birkhäuser, 2001.
- [35] R.R. Coifman, Y. Meyer, and V. Wickerhauser, *Wavelet Analysis and Signal Processing*, Yale University, New Haven, CT, 1991.
- [36] J.M. Combes, A. Grossman, and P. Tchamitchian, Eds., *Wavelets, Time-Frequency Methods and Phase Space*, Springer-Verlag, Berlin, 1989.
- [37] M. Akay, Ed., *Time Frequency and Wavelets in Biomedical Signal Processing*, IEEE Press Series in Biomedical Engineering, John Wiley & Sons, New York, 1997.
- [38] L. Stanković, Z. Uskoković, and S. Stanković, Time-frequency signal analysis, Podgorica, Montenegro, 1999 (research monograph).
- [39] T.S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, 1996.
- [40] H.V. Poor and G.W. Wornell, Eds., *Wireless Communications: Signal Processing Perspectives*, Prentice Hall, Upper Saddle River, NJ, 1998.
- [41] A.F. Molisch, Ed., *Wideband Wireless Digital Communications*, Prentice Hall, Upper Saddle River, NJ, 2001.
- [42] M.G. Amin, Interference mitigation in spread spectrum communication systems using time-frequency distributions, *IEEE Trans. Signal Process.*, 45, 90–101, January 1997.
- [43] C. Wang and M.G. Amin, Performance analysis of instantaneous frequency-based interference excision techniques in spread spectrum communications, *IEEE Trans. Signal Process.*, 46, 70–82, January 1998.
- [44] S.R. Lach, M.G. Amin, and A.R. Lindsey, Broadband Nonstationary Interference Excision for Spread Spectrum Communications Using Time-Frequency

Synthesis, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Seattle, WA, May 1998, pp. 3257–3260.

- [45] S. Barbarossa and A. Scaglione, Adaptive time-varying cancellation of wideband interferences in spread spectrum communications based on time-frequency distributions, *IEEE Trans. Signal Process.*, 47, pp. 957–965, April 1999.
- [46] M.L. Kramer and D.L. Jones, Nonstationary Interference Suppression Using Adaptive Overdetermined Frame Representations, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1, May 1998, pp. 23–27.
- [47] A. Bultan and A.N. Akansu, A Novel Time-Frequency Exciser in Spread Spectrum Communications for Chirp-Like Interference, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Seattle, WA, May 1998, pp. 3265–3268.
- [48] O. Akay and G.F. Boudreaux-Bartels, Broadband Interference Excision in Spread Spectrum Communication Systems via Fractional Fourier Transform, in Proceedings of the 32nd Asilomar Conference on Signals, Systems, and Computers, 1, 832–837, November 1998.
- [49] S. Chetwani, Matched Signal Transforms and Their Application in Wireless Communication Systems, Master's thesis, Arizona State University, Tempe, AZ, May 2001.
- [50] D. Park and A. Papandreou-Suppappola, Synthesis of Dispersive Signals and Applications in Wireless Communications, in Proceedings of the IEEE Transactions on Acoustics, Speech and Signal Processing, Orlando, FL, May 2002.
- [51] W. Kozek and A.F. Molisch, Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels, *IEEE J. Selected Areas Commun.*, 16, 1579–1589, October 1998.
- [52] A.R. Lindsey, Wavelet packet modulation for orthogonally multiplexed communication, *IEEE Trans. Signal Processing*, 45, 1336–1339, 1997.
- [53] T.A. Kadous, K. Liu, and A.M. Sayeed, Optimal Time-Frequency Signaling for Rapidly Time-Varying Channels, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Salt Lake City, UT, May 2001.
- [54] C. Gupta and A. Papandreou-Suppappola, Wireless CDMA Communications using Time-Varying Signals, in Proceedings of the IEEE International Symposium on Signal Processing and its Applications, 1, 242–245, August 2001 (invited paper).
- [55] G.B. Folland, *Harmonic Analysis in Phase Space*, Princeton University Press, Princeton, NJ, 1989.

- [56] W. Kozek, On the Generalized Weyl Correspondence and its Application to Time-Frequency Analysis of Linear Time-Varying Systems, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, Canada, 1992, pp. 167–170.
- [57] W. Kozek and A.F. Molisch, On the Eigenstructure of Underspread WSSUS Channels, in Proceedings of the IEEE Workshop on Signal Processing Applications in Wireless Communications, Paris, France, 1997, pp. 325–328.
- [58] G. Matz and F. Hlawatsch, 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 Wavelet Time-Frequency Analysis*, 39, 4041–4071, August 1998.
- [59] B.G. Iem, A. Papandreou-Suppappola, and G.F. Boudreaux-Bartels, Classes of smoothed Weyl symbols, *IEEE Signal Process. Lett.*, 7, 186–188, July 2000.
- [60] B.G. Iem, A. Papandreou-Suppappola, and G.F. Boudreaux-Bartels, Wideband Weyl symbols for dispersive time-varying processing of systems and random signals, *IEEE Trans. Signal Process.*, 50, 1077–1090, May 2002.
- [61] B. Friedlander and A. Zeira, Detection of broadband signals in frequency and time dispersive channels, *IEEE Trans. Signal Process.*, 44, 1613–1622, July 1996.
- [62] G. Matz and F. Hlawatsch, Robust Time-Varying Wiener Filters: Theory and Time-Frequency Formulation, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 401–404.
- [63] F. Hlawatsch, G. Matz, H. Kirchauer, and W. Kozek, Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation, *IEEE Trans. Signal Process.*, 48, 1417–1432, May 2000.
- [64] H. Artés, G. Matz, and F. Hlawatsch, An Unbiased Scattering Function Estimator for Fast Time-Varying Channels, in Proceedings of the IEEE Workshop on Signal Processing Advances in Wireless Communications, Annapolis, MD, May 1999, pp. 411–414.
- [65] J.G. Gallaire and A.M. Sayeed, Interference-Resistant Detection Based on Time-Frequency Subspaces, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 389–392.
- [66] A.M. Sayeed, Canonical Time-Frequency Processing for Broadband Signaling Over Dispersive Channels, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 369–372.
- [67] P. Flandrin, A time-frequency formulation of optimum detection, *IEEE Trans. Signal Process.*, 36, 1377–1384, September 1988.

- [68] T. Kumagai and K. Kobayashi, A New Group Demodulator with Multi-Symbol Chirp Fourier Transform for Mobile Communication Systems, in Proceedings of the IEEE International Conference on Universal Personal Communications, November 1995.
- [69] Y. Tsai and J. Chang, The Feasibility of Combating Multipath Interference by Chirp Spread Spectrum Techniques over Rayleigh and Rician Fading Channels, in Proceedings of the IEEE International Conference on Spread Spectrum Techniques and Applications 1, 282–286, July 1994.
- [70] C. Gupta, Multi-Rate Chirp Frequency-Hopped Code Division Multiple Access for Wireless Channels, Master’s thesis, Arizona State University, Tempe, AZ, May 2001.
- [71] C. Gupta and A. Papandreou-Suppappola, Multiple access signal processing techniques for wireless communications using multirate chirps, *IEEE Signal Process. Lett.*, [submitted].
- [72] S. Barbarossa and R. Torti, Chirped-OFDM for Transmission Over Time-Varying Channels with Linear Delay/Doppler Spreading, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Salt Lake City, UT, May 2001.
- [73] A.G. Orozco-Lugo and D.C. McLernon, Blind Channel Equalization Using Chirp Modulating Signals, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 2721–2724, June 2000.
- [74] J.Q. Pinkney, A.B. Sesay, S. Nichols, and R. Behin, A Robust High Speed Indoor Wireless Communications System Using Chirp Spread Spectrum, in Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering, 1, 84–89, May 1999.
- [75] A.M. Sayeed and B. Aazhang, Communications over multipath fading channels: a time-frequency perspective, in *Wireless Communications: CDMA versus TDMA*, Kluwer Academic, Dordrecht, 1997, pp. 73–98.
- [76] A.M. Sayeed, A. Sendonaris, and B. Aazhang, Multiuser detection in fast-fading multipath environments, *IEEE J. Selected Areas Commun.*, 1691–1701, December 1998.
- [77] A.M. Sayeed and B. Aazhang, Joint multipath-Doppler diversity in mobile wireless communications, *IEEE Trans. Commun.*, 47, 123–132, January 1999.
- [78] S. Haykin, *Communication Systems*, 4th ed., John Wiley & Sons, New York, 2001.
- [79] J.G. Proakis, *Digital Communications 4th ed.*, McGraw–Hill, Boston, MA, 2001.
- [80] L.B. Milstein, Interference rejection techniques in spread spectrum communications, in Proceedings of the IEEE, 76, 657–671, June 1988.

- [81] S. Chetwani and A. Papandreou-Suppappola, Unitary warped transforms for time-varying interference suppression in communication systems, *IEEE Signal Process. Lett.* [submitted].
- [82] F. Auger, P. Flandrin, P. Gonçalvès and O. Lemoine, Time-Frequency Toolbox for MATLAB, User's Guide and Reference Guide, 1996; available at <http://iut-saint-nazaire.univ-nantes.fr/~auger/ftfb.html>.
- [83] G.F. Boudreaux-Bartels and T.W. Parks, Time-varying filtering and signal estimation using Wigner distribution synthesis techniques, *IEEE Trans. Acoust., Speech, Signal Process.*, 34, 442–451, June 1986.
- [84] G.F. Boudreaux-Bartels, Time-varying signal processing using Wigner distribution synthesis techniques, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W.F.G. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997.
- [85] F. Hlawatsch and W. Krattenthaler, Signal synthesis algorithms for bilinear time-frequency signal representations, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W.F.G. Mecklenbräuker and F. Hlawatsch, Eds., Amsterdam, The Netherlands: Elsevier, 1997.
- [86] P. Flandrin, Some Features of Time-Frequency Representations of Multi-component Signals, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, San Diego, CA, March 1984, pp. 41B.4.1–41B.4.4.
- [87] L.J. Ziomek, *Underwater Acoustics: A Linear Systems Theory Approach*, Academic Press, Orlando, FL, 1985.
- [88] S.E. El-Khamy, S.E. Shaaban, and E.A. Tabet, Efficient Multiple-Access Communications Using Multi-User Chirp Modulation Signals, in Proceedings of the IEEE International Symposium on Spread Spectrum Techniques and Applications, 3, 1209–1213, September 1996.
- [89] S. El-Khamy, S.E. Shaaban, and E.A. Tabet, Frequency Hopped-Chirp Modulation (FH-CM) for Multi-User Signaling in Multipath Dispersive Media, in Proceedings of the IEEE International Symposium on Antennas Propagation, 1, 396–399, July 1999.
- [90] M.R. Winkler, Chirp signals for communications, in Proceedings of the IEEE WESCON Convention, 1962.
- [91] R.A. Altes and E.L. Titlebaum, Bat signals as optimally Doppler tolerant waveforms, *J. Acoust. Soc. Am.*, 48, 1014–1020, October 1970.
- [92] A. Papandreou-Suppappola, F. Hlawatsch, and G.F. Boudreaux-Bartels, Quadratic time-frequency representations with scale covariance and generalized time-shift covariance: a unified framework for the affine, hyperbolic, and power classes, *Digital Signal Process. a Rev. J.*, 8, 3–48, January 1998.

- [93] A. Papandreou-Suppappola, R.L. Murray, B.G. Iem, and G.F. Boudreault-Bartels, Group delay shift covariant quadratic time-frequency representations, *IEEE Trans. Signal Process.*, 49, 2549–2564, November 2001.
- [94] L. Cohen and T.E. Posch, Positive time-frequency distribution functions, *IEEE Trans. Acoust. Speech, Signal Process.*, 33, 31–37, February 1985.
- [95] P. Loughlin, J. Pitton, and L.E. Atlas, Construction of positive time-frequency distributions, *IEEE Trans. Signal Process.*, 42, 2697–2705, October 1994.
- [96] M. Colas, G. Gelle, and G. Delaunay, Sliced Higher Order Time-Frequency Representations for Detection and Classification: Acoustical Diagnosis Application to Faulty Fan Belts, in Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, 1999, pp. 183–186.
- [97] N.L. Gerr, Introducing a third-order Wigner distribution, in Proceedings of the IEEE, 76, 290–292, March 1988.
- [98] A. Swami, Third-order Wigner distribution: Definition and Properties, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Toronto, Canada, May 1991, pp. 3081–3084.
- [99] J.R. Fonollosa and C.L. Nikias, Wigner higher order moment spectra: definition, properties, computation and application to transient signal analysis, *IEEE Trans. Signal Process.*, 41, 245–266, January 1993.
- [100] D.C. Shin and C.L. Nikias, Complex ambiguity functions using non-stationary higher order cumulant estimates, *IEEE Trans. Signal Process.*, 43, 2649–2664, November 1995.
- [101] C.L. Nikias, Higher-order spectral analysis, in *Advances in Spectral Analysis and Array Processing*, Vol. 1, S. Haykin, Ed., Prentice Hall, Englewood Cliffs, NJ, 1991, pp. 326–365.
- [102] A.V. Dandawate and G.B. Giannakis, Differential delay-Doppler estimation using second- and higher-order ambiguity functions, IEE in Proceedings of the Radar Signal Processing, 140, 410–418, December 1993.
- [103] B. Ristic, G. Roberts, and B. Boashash, Higher-Order Scale Spectra and Higher-Order Time-Scale Distributions, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1995, pp. 1577–1580.
- [104] R.L. Murray, A. Papandreou-Suppappola, and G.F. Boudreault-Bartels, New Higher Order Spectra and Time-Frequency Representations for Dispersive Signal Analysis, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 4, May, 1998, pp. 2305–2308.
- [105] R.L. Murray, A. Papandreou-Suppappola, and G.F. Boudreault-Bartels, A New Class of Affine Higher Order Time-Frequency Representations, in Proceedings

of the IEEE International Conference on Acoustics, Speech and Signal Processing, Phoenix, AZ, March 1999, pp. 1613–1616.

- [106] A. Papoulis, *Signal Analysis*, McGraw-Hill, New York, 1977.
- [107] D. Preis, Phase distortion and phase equalization in audio signal processing — A tutorial review, *J. Audio Eng. Soc.*, 30, 774–794, November 1982.
- [108] C. Annivas, Time-Frequency Detection and Estimation Methodologies for Non-linear Time-Varying Signals, Master’s thesis, Arizona State University, Tempe, AZ, March, 2002.
- [109] A. Papandreou, F. Hlawatsch, and G.F. Boudreaux-Bartels, 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*, 41, 3425–3444, December 1993.
- [110] J.P. Ovarlez, J. Bertrand, and P. Bertrand, Computation of Affine Time-Frequency Distributions Using the Fast Mellin Transform, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, San Francisco, CA, 1992, pp. 117–120.
- [111] F. Hlawatsch, A. Papandreou-Suppappola, and G.F. Boudreaux-Bartels, The power classes — quadratic time-frequency representations with scale covariance and dispersive time-shift covariance, *IEEE Trans. Signal Process.*, 47, 3067–3083, November 1999.
- [112] A. Papandreou-Suppappola and G.F. Boudreaux-Bartels, The Exponential Class and Generalized Time-Shift Covariant quadratic time-frequency representations, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 429–432.
- [113] X.-G. Xia, Discrete chirp-Fourier transform and its application to chirp rate estimation, *IEEE Trans. Signal Process.*, 48, 3122–3133, November 2000.
- [114] P. Fan and X.-G. Xia, A Modified Discrete Chirp-Fourier Transform Scheme, in Proceedings of the 5th International Conference on Signal Processing, 1, 57–60, 2000.
- [115] F. Hlawatsch, Regularity and unitarity of bilinear time-frequency signal representations, *IEEE Trans. Inf. Theory*, 38, 82–94, January 1992.
- [116] F. Hlawatsch and G.F. Boudreaux-Bartels, Linear and quadratic time-frequency signal representations, *IEEE Signal Process. Mag.*, 9, 21–67, April 1992.
- [117] G.F. Boudreaux-Bartels, Mixed time-frequency signal transformations, in *The Transforms and Applications Handbook*, A.D. Poularikas, Ed., CRC Press, Boca Raton, FL, 1996.

- [118] Proceedings IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, 1992, 1994, 1996, 1998.
- [119] S. Qian and D. Chen, Joint time-frequency analysis, *IEEE Signal Process. Mag.*, 16, 52–67, March 1999.
- [120] C.S. Burrus, R.A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, Prentice-Hall, Upper Saddle River, NJ, 1998.
- [121] O. Rioul and M. Vetterli, Wavelets and signal processing, *IEEE Signal Process. Mag.*, 8, 14–38, October 1991.
- [122] L. Cohen, Time-frequency distribution — A review, in Proceedings of the IEEE, 77, 941–981, July 1989.
- [123] L.F. Chaparro and A. Akan, Eds., *Franklin Institute*, 337, July 2000, Special Issue on Time-Frequency Signal Analysis and Its Applications.
- [124] S.H. Nawab and T.F. Quatieri, Short-time Fourier transform, in *Advanced Topics in Signal Processing*, J.S. Lim and A.V. Oppenheim, Eds., Prentice Hall, Englewood Cliffs, NJ, 1988.
- [125] L.R. Rabiner and R.W. Schafer, *Digital Processing of Speech Signals*, Prentice Hall, Englewood Cliffs, NJ, 1978.
- [126] J.B. Allen, Short-term spectral analysis, synthesis, and modification by discrete Fourier transform, *IEEE Trans. Acoust. Speech, Signal Process.*, 25, 235–238, June 1977.
- [127] J.B. Allen, D.A. Berkley, and J. Blauert, Multimicrophone signal processing techniques to remove room reverberation from speech signals, *J. Acoust. Soc. Am.*, 62, 912–915, October 1977.
- [128] J.B. Allen and L.R. Rabiner, A unified approach to STFT analysis and synthesis, in Proceedings of the IEEE, 65, 1558–1564, November 1977.
- [129] M.R. Portnoff, Time-frequency representations of digital signals and systems based on short-time Fourier analysis, *IEEE Trans. Acoust. Speech Signal Process.*, 28, 55–69, February 1980.
- [130] R.E. Crochiere and L.R. Rabiner, *Multi-Rate Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1983.
- [131] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inf. Theory*, 36, 961–1005, September 1990.
- [132] I. Daubechies, The wavelet transform: a method for time-frequency localization, in *Advances in Spectrum Analysis and Array Processing*, S. Haykin, Ed., Prentice Hall, Englewood Cliffs, NJ, 1991, pp. 366–417.
- [133] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Analysis Mach. Intelligence*, 11, 674–693, July 1989.

- [134] Y. Meyer, Orthonormal wavelets, in *Wavelets, Time-Frequency Methods and Phase Space*, J.M. Combes, A. Grossman, and P. Tchamitchian, Eds., Springer-Verlag, Berlin, 1989, pp. 21–37.
- [135] D.L. Jones and T.W. Parks, A high resolution data-adaptive time-frequency representation, *IEEE Trans. Acoust. Speech Signal Process.*, 38, 2127–2135, 1990.
- [136] R.G. Baraniuk and D.L. Jones, A signal-dependent time-frequency representation: optimal kernel design, *IEEE Trans. Signal Process.*, 41, 1589–1602, April 1993.
- [137] D.L. Jones and R.G. Baraniuk, An adaptive optimal-kernel time-frequency representation, *IEEE Trans. Signal Process.*, 43, 2361–2371, October 1995.
- [138] R.G. Baraniuk and D.L. Jones, Signal-dependent time-frequency analysis using a radially Gaussian kernel, *Signal Process.*, 32, 263–284, June 1993.
- [139] P. Flandrin, F. Auger, and E. Chassande-Mottin, Time-frequency reassignment: from principles to algorithms, in *Applications in Time-Frequency Signal Processing*, A. Papandreou-Suppappola, Ed., CRC Press, Boca Raton, FL, 2002.
- [140] A. Auger, P. Flandrin, and E. Chassande-Mottin, Time-frequency reassignment, in *Time-Frequency Signal Analysis and Processing*, B. Boashash, Ed., Prentice Hall, New York, 2002.
- [141] D. Gabor, Theory of Communication, *J. IEE*, 93(III), 429–457, November 1946.
- [142] M.J. Bastiaans, Gabor’s expansion of a signal into Gaussian elementary signals, in Proceedings of the IEEE, 68, 538–539, 1980.
- [143] C.W. Helstrom, An expansion of a signal in Gaussian elementary signals, *IEEE Trans. Inf. Theory*, IT-12, 81–82, January 1966.
- [144] L.K. Montgomery and I.S. Reed, A generalization of the Gabor-Helstrom transform, *IEEE Trans. Inf. Theory*, IT-13, 344–345, April 1967.
- [145] G.A. Nelson, Signal analysis in time and frequency using Gaussian wavefunctions, in J.O. Scanlan, Ed., NATO Advanced Studies Institute on Network and Signal Theory, 1972, pp. 454–460.
- [146] A.J.E.M. Janssen, Gabor representation of generalized functions, *J. Math. Analysis Appl.*, 83, 377–394, 1981.
- [147] J. Wexler and S. Raz, Discrete Gabor expansions, *Signal Process.*, 21, 207–220, 1990.
- [148] S.G. Mallat and Z. Zhang, Matching pursuits with time-frequency dictionaries, *IEEE Trans. Signal Process.*, 41, 3397–3415, December 1993.

- [149] S. Qian and D. Chen, Decomposition of the Wigner distribution and time-frequency distribution series, *IEEE Trans. Signal Process.*, 42, 2836–2842, October 1994.
- [150] A. Bultan, A four-parameter atomic decomposition of chirplets, *IEEE Trans. Signal Process.*, 47, 731–745, March 1999.
- [151] M.R. McClure and L. Carin, Matching pursuits with a wave-based dictionary, *IEEE Trans. Signal Process.*, 45, 2912–2927, December 1997.
- [152] A. Papandreou-Suppappola and S.B. Suppappola, Analysis and classification of time-varying signals with multiple time-frequency structures, *IEEE Signal Process. Lett.*, March 2002.
- [153] A. Papandreou-Suppappola and S.B. Suppappola, Adaptive Time-Frequency Representations for Multiple Structures, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, PA, August 2000, pp. 579–583.
- [154] K. Ghartey, A. Papandreou-Suppappola, and D. Cochran, On the Use of Matching Pursuit Time-Frequency Techniques for Multiple-Channel Detection, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 3201–3204, May 2001.
- [155] S.P. Ebenezer, A. Papandreou-Suppappola, and S.B. Suppappola, Matching Pursuit Classification for Time-Varying Acoustic Emissions, in Proceedings of the 35th Annual Asilomar Conference on Signals, Systems and Computers, 1, 715–719, November 2001.
- [156] R.G. Stockwell, L. Mansinha, and R.P. Lowe, Localization of the complex spectrum: the S-Transform, *IEEE Trans. Signal Process.*, 44, 998–1001, April 1996.
- [157] P. Flandrin and W. Martin, Sur les Conditions Physiques Assurant l’Unicité de la Représentation de Wigner-Ville Comme Représentation Temps-Fréquence, in Neuvième Coll. sur le Traitement du Signal et ses Applications, Nice, France, May 1983.
- [158] G.F. Boudreax-Bartels, Time-Frequency Signal Processing Algorithms: Analysis and Synthesis Using Wigner Distributions, Ph.D. thesis, Rice University, Houston, TX, 1983.
- [159] A.J.E.M. Janssen, On the locus and spread of pseudo-density functions in the time-frequency plane, *Philips J. Res.*, 37(3), 79–110, 1982.
- [160] P. Flandrin and O. Rioul, Affine Smoothing of the Wigner-Ville Distribution, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Albuquerque, NM, April 1990, pp. 2455–2458.
- [161] O. Rioul and P. Flandrin, Time-scale energy distributions: a general class extending wavelet transforms, *IEEE Trans. Signal Process.*, 40, 1746–1757, July 1992.

- [162] M.H. Ackroyd, Short-time spectra and time-frequency energy distributions, *J. Acoust. Soc. Am.*, 50, 1229–1231, 1970.
- [163] R.A. Altes, Wide-band, proportional-bandwidth Wigner-Ville analysis, *IEEE Trans. Acoust. Speech Signal Process.*, 38, 1005–1012, June 1990.
- [164] N.M. Marinovich, The Wigner Distribution and the Ambiguity Function: Generalizations, Enhancement, Compression and Some Applications, Ph.D. thesis, City University of New York, 1986.
- [165] J. Bertrand and P. Bertrand, Affine time-frequency distributions, in *Time-Frequency Signal Analysis — Methods and Applications*, B. Boashash, Ed., Longman-Cheshire, Melbourne, Australia, 1992, pp. 118–140.
- [166] J. Bertrand and P. Bertrand, A class of affine Wigner functions with extended covariance properties, *J. Math. Phys.*, 33, 2515–2527, 1992.
- [167] A. Papandreou and G.F. Boudreaux-Bartels, Generalization of the Choi-Williams distribution and the Butterworth distribution for time-frequency analysis, *IEEE Trans. Signal Process.*, 41, 463–472, January 1993.
- [168] H.I. Choi and W.J. Williams, Improved time-frequency representation of multi-component signals using exponential kernels, *IEEE Trans. Acoust. Speech Signal Process.*, 37, 862–871, June 1989.
- [169] Y. Zhao, L.E. Atlas, and R.J. Marks, The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals, *IEEE Trans. Acoust. Speech Signal Process.*, 38, 1084–1091, July 1990.
- [170] P. Flandrin, Sur une Classe Générale d’Extensions Affines de la Distribution de Wigner-Ville, in 13ème Colloque sur le Traitement du Signal et ses Applications, Juan-les-Pins, France, September 1991.
- [171] A. Papandreou-Suppappola, B.G. Iem, R.L. Murray, and G.F. Boudreaux-Bartels, Properties and Implementation of the Exponential Class of Quadratic Time-Frequency Representations, in Proceedings of the 13th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 1996, pp. 237–241.
- [172] M.J. Levin, Instantaneous spectra and ambiguity functions, *IEEE Trans. Inf. Theory*, IT-13, 95–97, 1967.
- [173] H. Margenau and R.W. Hill, Correlation between measurements in quantum theory, *Prog. Theor. Phys.*, 26, 722–738, 1961.
- [174] M. Amin, Minimum variance time-frequency distribution kernels for signals in additive noise, *IEEE Trans. Signal Process.*, 44, 2352–2356, September 1996.
- [175] A.H. Costa and G.F. Boudreaux-Bartels, Design of time-frequency representations using a multiform, tiltable exponential kernel, *IEEE Trans. Signal Process.*, 43, 2283–2301, October 1995.

- [176] C.H. Page, Instantaneous power spectra, *J. Appl. Phys.*, 23, 103–106, 1952.
- [177] A. Unterberger, The calculus of pseudo-differential operators of Fuchs type, *Commun. Partial Differential Equations*, 9, 1179–1236, 1984.
- [178] A. Papandreou-Suppappola, F. Hlawatsch, and G.F. Boudreux-Bartels, Power Class Time-Frequency Representations: Interference Geometry, Smoothing, and Implementation, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 193–196.
- [179] F. Hlawatsch, A. Papandreou, and G.F. Boudreux-Bartels, The Power Classes of Quadratic Time-Frequency Representations: A Generalization of the Affine and Hyperbolic Classes, in Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 1993, pp. 1265–1270.
- [180] F. Auger and P. Flandrin, The Why and How of Time-Frequency Reassignment, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, October 1994, pp. 197–200.
- [181] F. Auger and P. Flandrin, Improving the readability of time-frequency and time-scale representations by the reassignment method, *IEEE Trans. Signal Process.*, 43, 1068–1089, May 1995.
- [182] E. Chassande-Mottin and P. Flandrin, Reassigned Scalograms and singularities, in *European Congress of Mathematics: Progress in Mathematics*, C. Casacuberta, Ed., vol. 202, Birkhäuser, Basel, 2000, pp. 582–590.
- [183] W.J. Williams and J. Jeong, Reduced interference time-frequency distributions, in *Time-Frequency Signal Analysis—Methods and Applications*, B. Boashash, Ed., Longman-Cheshire, Melbourne, Australia, 1992.
- [184] A.W. Rihaczek, Signal energy distribution in time and frequency, *IEEE Trans. Inf. Theory*, IT-14(3), 369–374, 1968.
- [185] L. Stankovic, S. Stankovic, and Z. Uskokovic, S-methods with Variable Convolution Window, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 185–188.
- [186] F. Hlawatsch, A. Papandreou-Suppappola and G.F. Boudreux-Bartels, The hyperbolic class of quadratic time-frequency representations, Part II. Sub-classes, intersection with the affine and power classes, regularity, and unitarity, *IEEE Trans. Signal Process.*, 45, 303–315, February 1997.
- [187] F. Hlawatsch, Duality and classification of bilinear time-frequency signal representations, *IEEE Trans. Signal Process.*, 39, 1564–1574, July 1991.
- [188] P. Flandrin and P. Goncalvès, Geometry of affine time-frequency distributions, *Appl. Computational Harmonic Analysis*, 3, 10–39, January 1996.

- [189] A. Papandreou-Suppappola, Quadratic time-frequency representations preserving constant and dispersive time shifts, in *Wavelets and Signal Processing*, L. Debnath, Ed., Birkhäuser, Boston, 2002.
- [190] A. Papandreou, F. Hlawatsch, and G.F. Boudreux-Bartels, A Unified Framework for the Scale Covariant Affine, Hyperbolic, and Power Class Time-Frequency Representations Using Generalized Time-Shifts, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Detroit, MI, May 1995.
- [191] R.G. Baraniuk, Warped Perspectives in Time-Frequency Analysis, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, October 1994, pp. 528–531.
- [192] R.G. Baraniuk and D.L. Jones, Unitary equivalence: a new twist on signal processing, *IEEE Trans. Signal Process.*, 43, 2269–2282, October 1995.
- [193] R.G. Baraniuk, Covariant time-frequency representations through unitary equivalence, *IEEE Signal Process. Lett.*, 3, 79–81, March 1996.
- [194] F. Hlawatsch and H. Bölcseki, Displacement-covariant time-frequency distributions based on conjugate operators, *IEEE Signal Process. Lett.*, 3, 44–46, February 1996.
- [195] F. Hlawatsch and T. Twaroch, Covariant (α, β) , Time-Frequency, and (a, b) Representations, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 437–440.
- [196] A.M. Sayeed and D.L. Jones, A canonical covariance-based method for generalized joint signal representations, *IEEE Signal Process. Lett.*, 3, 121–123, April 1996.
- [197] A.M. Sayeed and D.L. Jones, A Simple Covariance-Based Characterization of Joint Signal Representations of Arbitrary Variables, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 433–436.
- [198] A.H. Costa and G.F. Boudreux-Bartels, An overview of aliasing errors in discrete-time formulations of time-frequency representations, *IEEE Trans. Signal Process.*, 47, 1463–1474, May 1999.
- [199] P. Flandrin, Scale-Invariant Wigner Spectra and Self-Similarity, in Proceedings of the European Signal Processing Conference, Barcelona, Spain, September 1990, pp. 149–152.
- [200] C. Braccini and G. Gambardella, Form-invariant linear filtering: theory and applications, *IEEE Trans. Acoust. Speech Signal Process.*, 34, 1612–1628, December 1986.

- [201] J.M. Speiser, Wide-band ambiguity function, *IEEE Trans. Inf. Theory*, 13, 122–123, 1967.
- [202] R.G. Shenoy and T.W. Parks, Wide-band ambiguity functions and affine Wigner distributions, *Signal Process.*, 41, 339–363, 1995.
- [203] A. Papandreou-Suppappola, New Classes of Quadratic Time-Frequency Representations with Scale Covariance and Generalized Time-Shift Covariance: Analysis, detection, and estimation, Ph.D. thesis, University of Rhode Island, Kingston, RI, 1995.
- [204] A. Papandreou-Suppappola, Time-frequency representations covariant to frequency-dependent time shifts, in *Time-Frequency Signal Analysis and Processing*, B. Boashash, Ed., Prentice Hall, New York, 2002.
- [205] A. Papandreou-Suppappola, B.G. Iem, and G.F. Boudreux-Bartels, Time-frequency symbols for statistical signal processing, in *Time-Frequency Signal Analysis and Processing*, B. Boashash, Ed., Prentice Hall, New York, 2002.
- [206] A. Papandreou-Suppappola and L.T. Antonelli, Use of Quadratic Time-Frequency Representations to Analyze Cetacean Mammal Sounds, Technical rep. 11,284, Naval Undersea Warfare Center, Newport, RI, December, 2001.
- [207] J.A. Simmons, P.A. Saillant, and S.P. Dear, Through a bat's ear, *IEEE Spectrum*, 29, 46–48, March 1992.
- [208] A. Papandreou-Suppappola and G.F. Boudreux-Bartels, Distortion that Occurs When the Signal Group Delay Does not Match the Time-Shift Covariance of a Time-Frequency Representation, in Proceedings of the 30th Annual Conference on Information Sciences and Systems, Princeton, NJ, March 1996, pp. 520–525.
- [209] A. Papandreou-Suppappola and G.F. Boudreux-Bartels, The Effect of Mismatching Analysis Signals and Time-Frequency Representations, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 149–152.
- [210] P. Guillemain and P. White, Wavelet Transforms for the Analysis of Dispersive Systems, in Proceedings of the IEEE UK Symposium on Applications of Time-Frequency and Time-Scale Methods, University of Warwick, Coventry, UK, August 1995, pp. 32–39.
- [211] D.E. Newland, Time-frequency and time-scale analysis by harmonic wavelets, in *Signal Analysis and Prediction*, A. Prochazka, Ed., Birkhäuser, Boston, 1998, chap. 1.
- [212] M.J. Freeman, M.E. Dunham, and S. Qian, Trans-ionospheric Signal Detection by Time-Scale Representation, in Proceedings of the IEEE UK Symposium on Applications of Time-Frequency and Time-Scale Methods, University of Warwick, Coventry, UK, August 1995, pp. 152–158.

- [213] J.P. Sessarego, J. Sageloli, P. Flandrin, and M. Zakharia, Time-frequency Wigner-Ville analysis of echoes scattered by a spherical shell, in *Wavelets, Time-Frequency Methods and Phase Space*, J.M. Combes, A. Grossman, and P. Tchamitchian, Eds., Springer-Verlag, Heidelberg, 1989, pp. 147–153.
- [214] A. Papandreou-Suppappola and S.B. Suppappola, Sonar echo ranging using signals with non-linear time-frequency characteristics, *IEEE Trans. Signal Process.* [submitted].
- [215] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill, New York, 1953.
- [216] V. Székely, Distributed RC networks, in *The Circuits and Filters Handbook*, W.K. Chen, Ed., CRC Press/IEEE Press, Boca Raton, FL, 1995, pp. 1203–1221.
- [217] K.G. Canfield and D.L. Jones, Implementing Time-Frequency Representations for non-Cohen Classes, in Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 1993, pp. 1464–1468.
- [218] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal, in Proceedings of the IEEE, 8, 520–568, 1992.
- [219] V. Katkovnik and L. Stankovic, Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length, *IEEE Trans. Signal Process.*, 46, 2315–2325, September 1998.
- [220] Z.M. Hussain and B. Boashash, Adaptive Instantaneous Frequency Estimation of Multi-Component FM Signals, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 657–660, June 2000.
- [221] V. Katkovnik, Nonparametric estimation of instantaneous frequency, *IEEE Trans. Inf. Theory*, 43, 183–189, January 1997.
- [222] P.A. Bello, Characterization of randomly time-variant linear channels, *IEEE Trans. Commun.*, 11, 360–393, 1963.
- [223] W. Kozek, On the transfer function calculus for underspread LTV channels, *IEEE Trans. Signal Process.*, 45, 219–223, January 1997.
- [224] R.G. Shenoy and T.W. Parks, The Weyl correspondence and time-frequency analysis, *IEEE Trans. Signal Process.*, 42, 318–331, February 1994.
- [225] B.G. Iem, A. Papandreou-Suppappola, and G.F. Boudreaux-Bartels, New concepts in Narrowband and Wideband Weyl Correspondence Time-Frequency Techniques, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 3, 1573–1576, May 1998.
- [226] B.G. Iem, A. Papandreou-Suppappola, and G.F. Boudreaux-Bartels, Wideband Time-Frequency Weyl Symbol and its Generalization, in Proceedings of

- the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 29–32.
- [227] A. Bircan, S. Tekinay, and A.N. Akansu, Time-Frequency and Time-Scale Representation of Wireless Communication Channels, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 373–376.
 - [228] S. Bhashyam, A.M. Sayeed, and B. Aazhang, Time-selective signaling and reception for communication over multipath fading channels, *IEEE Trans. Commun.*, 48, 83–94, January 2000.
 - [229] T.L. Nguyen and B. Senadji, Detection of Frequency Modulated Signals in Rayleigh Fading Channels Based on Time-Frequency Distributions, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2, 729–732, June 2000.
 - [230] O. Ertug and B. Baykal, Time-Frequency Decorrelating Decision-Feedback Multiuser Detector over Fast-Fading Multipath Channels for Synchronous DS-CDMA Systems, in Proceedings of the IEEE 3rd Workshop on Signal Processing Advances in Wireless Communications, March 2001, pp. 170–173.
 - [231] B. Ertug and B. Baykal, Space-time-frequency multiuser detection over fast-fading multipath channels for synchronous DS-CDMA systems, in Proceedings of the IEEE Vehicular Technological Conference, 2, 1503–1507, 2001.
 - [232] B. Barkat and S. Attallah, A Study of Time-Frequency Based Detectors for FSK Modulated Signals in a Flat Fading Channel, in Proceedings of the IEEE Signal Processing Workshop on Statistical Signal Processing, 2001, pp. 206–209.
 - [233] S. Burley, M. Darnell, and P. Clark, Enhanced PN Code Tracking and Detection Using Wavelet Packet Denoising, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 365–368.
 - [234] E. Bejjani, J.C. Belfiore, and P. Leclair, Coherent Detection for Transmission over Severely Time and Frequency Dispersive Multipath Fading Channels, in Proceedings of the IEEE Symposium on Information Theory, 1995.
 - [235] D.J. Van Wyk and L.P. Linde, A Discrete Pseudo Wigner Distribution Spread Spectrum Detection Procedure Employing Complex Spreading Sequences, in Proceedings of the IEEE South African Symposium on Communications and Signal Processing, October 1994, pp. 33–40.
 - [236] G.R. Cooper and C.D. McGillem, *Modern Communications and Spread Spectrum*, McGraw-Hill, New York, 1986.
 - [237] S.E. Carter and D.C. Malocha, SAW device implementation of a weighted stepped chirp code signal for direct sequence spread spectrum communications

- systems, *IEEE Trans. Ultrasonics, Ferroelectrics Frequency Control*, 47, 967–973, July 2000.
- [238] R. Suleesathira, L.F. Chaparro, and A. Akan, Discrete evolutionary transform for time-frequency signal analysis, *J. Franklin Inst.*, 337, 347–364, July 2000.
- [239] R. Suleesathira and L. Chaparro, Interference Mitigation in Spread Spectrum Using Discrete Evolutionary Hough Transforms, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 2821–2824, June 2000.
- [240] S. Barbarossa, A. Scaglione, S. Spalletta, and S. Votini, Adaptive Suppression of Wideband Interferences in Spread-Spectrum Communications Using the Wigner-Hough Transform, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 3861–3864, April 1997.
- [241] X. Ouyang and M. Amin, Short-time Fourier transform receiver for nonstationary interference excision in direct sequence spread spectrum communications, *IEEE Trans. Signal Process.*, 49, 851–863, April 2001.
- [242] L. Zhao, M.G. Amin, and A.R. Lindsey, Subspace Projection Techniques for Anti-FM Jamming GPS Receivers, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, August 2000, pp. 529–533.
- [243] Y. Zhang and M.G. Amin, Array processing for nonstationary interference suppression in DS/SS communications using subspace projection techniques, *IEEE Trans. Signal Process.*, 49, 3005–3014, December 2001.
- [244] A. Bultan, M.V. Tazebay, and A.N. Akansu, A Comparative Performance Study of Excisers for Various Interference Classes, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 381–384.
- [245] A. Ranheim, Narrowband interference rejection in direct-sequence spread-spectrum system using time-frequency decomposition, IEE Proceedings on Communications, 142, 393–400, December 1995.
- [246] G.F. Boudreaux-Bartels and R.L. Murray, Time-frequency signal representations for biomedical signals, in *Biomedical Engineering Handbook*, J.D. Bronzino, Ed., CRC Press, Boca Raton, FL, June 1995.
- [247] IEE Colloquium on Time-Frequency Analysis of Biomedical Signals, 1997, digest no. 1997/006.
- [248] C. Bigan and W.S. Woolfson, Time-frequency analysis of short segments of biomedical data, in Proceedings of the IEE Science, Measurement and Technology, 147, 368–373, November 2000.
- [249] J.H. Dripps, An Introduction to Time-Frequency Methods Applied to Biomedical signals, in Proceedings of the IEE Colloquium on Time-Frequency Analysis of Biomedical Signals, 1997, pp. 1/1–1/4, digest no. 1997/006.

- [250] J.A. Crowe, The Wavelet Transform and Its Application to Biomedical Signals, in Proceedings of the IEE Colloquium on Time-Frequency Analysis of Biomedical Signals, 1997, pp. 2/1–2/3, digest no. 1997/006.
- [251] M. Akay and C. Mello, Wavelets for Biomedical Signal Processing, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 6, 2688–2691, 1997.
- [252] M. Unser, Wavelets, Statistics, and Biomedical Applications, in Proceedings of the IEEE Signal Processing Workshop on Statistical Signal and Array Processing, 1996, pp. 244–249.
- [253] W.J. Williams, Reduced interference distributions: biological applications and interpretations, in Proceedings of the IEEE, 84, 1264–1280, 1996.
- [254] X. Jingping, L.G. Durand, and P. Pibarot, Extraction of the aortic and pulmonary components of the second heart sound using a nonlinear transient chirp signal model, *IEEE Trans. Biomed. Eng.*, 48, 277–283, March 2001.
- [255] M. Stridh, L. Sornmo, C.J. Meurling, and S.B. Olsson, Characterization of atrial fibrillation using the surface ECG: time-dependent spectral properties, *IEEE Trans. Biomed. Eng.*, 48, 19–27, January 2001.
- [256] B. Gramatikov and V. Iyer, Ischemia-Related Depolarization Changes Detected in Body-Surface Ambulatory ECG Recordings, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 2, 912–914, July 2000.
- [257] J. Zhang, J. Liu, C. Zheng, W. Tao, N.V. Thakor, and A. Xie, Noninvasive early detection of focal cerebral ischemia, *IEEE Eng. Med. Biol. Mag.*, 19, 74–81, November 2000.
- [258] S. Krishnan, R.M. Rangayyan, G.D. Bell, and C.B. Frank, Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology, *IEEE Trans. Biomed. Eng.*, 47, 773–783, June 2000.
- [259] M. Roessgen, A.M. Zoubir, and B. Boashash, Seizure detection of newborn EEG using a model-based approach, *IEEE Trans. Biomed. Eng.*, 45, 673–685, June 1998.
- [260] P. Celka, B. Boashash, and P. Colditz, Preprocessing and time-frequency analysis of newborn EEG seizures, in Proceedings of the IEEE Engineering in Medicine and Biology Magazine, 20, 30–39, September–October 2001.
- [261] B. Boashash, M. Mesbah, and P. Colditz, Newborn EEG Seizure Pattern Characterization Using Time-Frequency Analysis, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2, 1041–1044, May 2001.

- [262] S. Blanco, S. Kochen, O.A. Rosso, and P. Salgado, Applying time-frequency analysis to seizure EEG activity, in Proceedings of the IEEE Engineering in Medicine and Biology Magazine, 16, 64–71, January–February 1997.
- [263] F. Wendling, M.B. Shamsollahi, J.M. Badier, and J.J. Bellanger, Time-frequency matching of warped depth-EEG seizure observations, *IEEE Trans. Biomed. Eng.*, 46, 601–605, May 1999.
- [264] M. Shen and F. Fenglin, The Analysis of Dynamic EEG Signals by Using Wavelet Packets Decomposition, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 85–88.
- [265] X. Yan, S. Haykin, and R.J. Racine, Multiple window time-frequency distribution and coherence of EEG using Slepian sequences and Hermite functions, *IEEE Trans. Biomed. Eng.*, 46, 861–866, July 1999.
- [266] S. Wu, Y. Qian, Z. Gao, and J. Lin, A novel method for beat-to-beat detection of ventricular late potentials, *IEEE Trans. Biomed. Eng.*, 48, 931–935, August 2001.
- [267] J. Millet-Roig, J.J. Rieta-Ibanez, E. Vilanova, A. Mocholi, and F.J. Chorro, Time-frequency analysis of a single ECG: to discriminate between ventricular tachycardia and ventricular fibrillation, *Comput. Cardiol.*, 711–714, 1999.
- [268] P. Lander and E.J. Berbari, Time-frequency plane Wiener filtering of the high-resolution ECG: background and time-frequency representations, *IEEE Trans. Biomed. Eng.*, 44, 247–255, April 1997.
- [269] A. Bousbia-Salah, A. Belouchrani, and A. Cichocki, Application of Time-Frequency Distributions to the Independent Component Analysis of ECG Signals, in Proceedings of the IEEE International Symposium on Signal Processing Applications, 1, 238–241, 2001.
- [270] R.H. Clayton, R.W.F. Campbell, and A. Murray, Time-Frequency Analysis of Ventricular Arrhythmias, in Proceedings of the IEE Colloquium on Time-Frequency Analysis of Biomedical Signals, 1997, pp. 6/1–6/4, digest no. 1997/006.
- [271] S. Kadambe, R. Murray, and G.F. Boudreaux-Bartels, Wavelet transform-based QRS complex detector, *IEEE Trans. Biomed. Eng.*, 46, 838–848, July 1999.
- [272] G. Georgiou and F.S. Cohen, A Hope for a Cheap and Reliable Automatic Diagnostic System for Breast Cancer, in Proceedings of the IEEE Ultrasonics Symposium, 2, 1311–1314, 2000.
- [273] P. Bonato, S.H. Roy, M. Knaflitz, and C.J. De Luca, Time-frequency parameters of the surface myoelectric signal for assessing muscle fatigue during cyclic dynamic contractions, *IEEE Trans. Biomed. Eng.*, 48, July 2001.

- [274] S. Karlsson, J.Yu, and M. Akay, Enhancement of spectral analysis of myoelectric signals during static contractions using wavelet methods, *IEEE Trans. Biomed. Eng.*, 46, 670–684, June 1999.
- [275] S. Karlsson, J. Yu, and M. Akay, Time-frequency analysis of myoelectric signals during dynamic contractions: a comparative study, *IEEE Trans. Biomed. Eng.*, 47, 228–238, February 2000.
- [276] M. Gutierrez, M. Rebelo, and S. Furui, A Multiresolution Approach for Computing Myocardial Motion, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 89–92.
- [277] C.S. Pattichis and M.S. Pattichis, Time-scale analysis of motor unit action potentials, *IEEE Trans. Biomed. Eng.*, 46, 1320–1329, November 1999.
- [278] Y. Sungyub, J.R. Boston, T.E. Rudy, C.M. Greco, and J.K. Leader, Time-frequency analysis of temporomandibular joint (TMJ) clicking sounds using radially Gaussian kernels, *IEEE Trans. Biomed. Eng.*, 48, 936–939, August 2001.
- [279] D. Djurdjanovic, S.E. Widmalm, W.J. Williams, C.K.H. Koh, and K.P. Yang, Computerized classification of temporomandibular joint sounds, *IEEE Trans. Biomed. Eng.*, 47, 977–984, August 2000.
- [280] C. Zheng, S.E. Widmalm, and W.J. Williams, New Time-Frequency Analyses of EMG and TMJ Sound Signals, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 2, 741–742, 1989.
- [281] K. Sunagawa, H. Kanai, and M. Tanaka, Simultaneous Measurement of Blood Flow and Arterial Wall Vibrations in Radial and Axial Directions, in Proceedings of the IEEE Ultrasonics Symposium, 2, 1541–1544, 2000.
- [282] B.S. Krongold, A.M. Sayeed, M.A. Moehring, J.A. Ritcey, M.P. Spencer, and D.I. Jones, Time-scale detection of microemboli in flowing blood with Doppler ultrasound, *IEEE Trans. Biomed. Eng.*, 46, 1081–1089, September 1999.
- [283] A. Stefanovska, M. Bracic, and H.D. Kverno, Wavelet analysis of oscillations in the peripheral blood circulation measured by laser Doppler technique, *IEEE Trans. Biomed. Eng.*, 46, 1230–1239, October 1999.
- [284] B.T. Tan, Y.K. Lew, C.H.K. Koh, and Y.C. Soh, Measurement of Blood Pressure Using Wavelet Transforms, in Proceedings of the International Conference on Information Communications and Signal Processing, 3, 1358–1362, 1997.
- [285] X. Long, J.W. Lee, and V.C. Roberts, Dynamic Spectral Analysis of Arterial Doppler Blood Flow Signals Using Time-Frequency Representations, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 2, 873–874, 1996.

- [286] P.R. White, W.B. Collis, and A.P. Salmon, Time-Frequency Analysis of Heart Murmurs in Children, in Proceedings of the IEE Colloquium on Time-Frequency Analysis of Biomedical Signals, 1997, pp. 3/1–3/4, digest no. 1997/006.
- [287] A. Haghghi-Mood and J.N. Torry, Time-Frequency Analysis of Systolic Murmurs, in Proceedings of the IEE Colloquium on Time-Frequency Analysis of Biomedical Signals, 1997, pp. 7/1–7/2, digest no. 1997/006.
- [288] H.Sava and L.G. Durand, Automatic Detection of Cardiac Cycle Based on an Adaptive Time-Frequency Analysis of the Phonocardiogram, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 3, 1316–1319, 1997.
- [289] X. Zhang, L. Durand, L. Senhadji, H.C. Lee, and J.L. Coatrieux, Time-frequency scaling transformation of the phonocardiogram based on the matching pursuit method, *IEEE Trans. Biomed. Eng.*, 45, 972–979, August 1998.
- [290] O.A. Ahmed and M.M. Fahmy, NMR signal enhancement via a new time-frequency transform, *IEEE Trans. Med. Imaging*, 20, 1018–1025, October 2001.
- [291] Y. Lu, S. Joshi, and J.M. Morris, Noise reduction for NMR FID signals via Gabor expansion, *IEEE Trans. Biomed. Eng.*, 44, 512–528, June 1997.
- [292] W. Qiao, H.H. Sun, W.Y. Chey, and K.Y. Lee, Continuous Wavelet Analysis as an Aid in the Representation and Interpretation of Electrogastrographic Signals, in Proceedings of the Biomedical Engineering Conference, 1996, pp. 140–141.
- [293] H.P. Zaveri, W.J. Williams, L.D. Iasemidis, and J.C. Sackellares, Time-frequency representation of electrocorticograms in temporal lobe epilepsy, *IEEE Trans. Biomed. Eng.*, 39, 502–509, May 1992.
- [294] J.L. Martinez-de-Juan, F.J. Saiz, J.L. Ponce, M. Meseguer, and S. Sancho Fornos, Retrieval of the Small Intestinal Pressure from Time-Frequency Analysis of the Electroenterogram, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 3, 1505–1508, 1998.
- [295] G. Tognola, F. Grandori, and P. Ravazzani, Wavelet analysis of click-evoked otoacoustic emissions, *IEEE Trans. Biomed. Eng.*, 45, 686–697, June 1998.
- [296] M. Li, H.G. McAllister, N.D. Black, and T.A. De Perez, Perceptual time-frequency subtraction algorithm for noise reduction in hearing aids, *IEEE Trans. Biomed. Eng.*, 48, 979–988, September 2001.
- [297] J. Cheng, Signal Processing Approaches on Otoacoustic Emissions, in Proceedings of the International Conference on Signal Processing, 2, 1612–1615, 1998.

- [298] J.A. Moczko, J. Jezewski, and A. Gacek, Detection of Fetal Breathing Movements with Joint Time-Frequency Analysis of Cardiotocogram Records, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 3, 1501–1504, 1998.
- [299] S. Yoo, P.J. Loughlin, M.S. Redfern, and J.M. Furman, Postural Sway Response to Chirp Moving Scene Perturbations, in Proceedings of the Joint BMES/EMBS Conference, 2, 973, 1999.
- [300] P.M. Bentley, P.M. Grant, and J.T.E. McDonnell, Time-frequency and time-scale techniques for the classification of native and bioprosthetic heart valve sounds, *IEEE Trans. Biomed. Eng.* 45, 125–128, January 1998.
- [301] D.J. Nelson, Cross-Spectral Methods for Processing Biological Signals, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, 2000, pp. 400–404.
- [302] D.J. Nelson, P. Loughlin, G. Cristobal, and L. Cohen, Time-Frequency Methods for Biological Signal Estimation, in Proceedings of the International Conference on Pattern Recognition, 3, 110–114, 2000.
- [303] M.L. Brown, W.J. Williams, and A.O. Hero, Non-Orthogonal Gabor Representation of Biological Signals, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 4, 305–308, 1994.
- [304] P. Flandrin, Time-frequency processing of bat sonar signals, in *Animal Sonar—Processes and Performance*, P.E. Nachtigall and P.W.B. Moore, Eds., Plenum Press, New York, 1988, pp. 797–802.
- [305] B. Ristic and B. Boashash, Scale Domain Analysis of a Bat Sonar Signal, in Proceedings IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, October 1994, pp. 373–376.
- [306] L. Jin, L.Z. Biao, S.C. Chan, and C.K. Chan, Wideband Ambiguity Function of South-China FM Bat Sonar Signals, in Proceedings of the IEEE Conference on Computer, Communication, Control and Power Engineering, 3, 310–313, 1993.
- [307] H. Zou, Q. Dai, R. Wang, and Y. Li, Parametric TFR via windowed exponential frequency modulated atoms, *IEEE Signal Process. Lett.*, 8, 140–142, May 2001.
- [308] S.K. Chittajallu, M. Palakal, and K. Kohrt, Computational model of signal-encoding in the mammalian peripheral auditory system, in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 2, 1303–1307, 1994.
- [309] D. Groutage, J. Schempp, and L. Cohen, Characterization and Analysis of Marine Mammal Sounds Using Time-Frequency and Time-Prony Techniques, in Proceedings of the Oceans Engineering, 1, I/253–I/258, 1994.

- [310] Q.Q. Huynh, L.N. Cooper, N. Intrator, and H. Shouval, Classification of underwater mammals using feature extraction based on time-frequency analysis and BCM theory, *IEEE Trans. Signal Process.*, 46, 1202–1207, May 1998.
- [311] Z.-H. Michalopoulou, Underwater Transient Signal Processing: Marine Mammal Identification, Localization, and Source Signal Deconvolution, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1, 503–506, April 1997.
- [312] P.L. Tyack, W.J. Williams, and G. Cunningham, Time-Frequency Fine Structure of Dolphin Whistles, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, Canada, 1992, pp. 17–20.
- [313] G. Roberts, A.M. Zoubir, and B. Boashash, Time-Frequency Classification Using a Multiple Hypotheses Test: an Application to the Classification of Humpback Whale Signals, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1, 563–566, April 1997.
- [314] P.J. Loughlin, J.W. Pitton, and B. Hannaford, Fast Approximations to Positive Time-Frequency Distributions, with Applications, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2, 1009–1012, May 1995.
- [315] N. Lovell and K. Lee, Time Frequency Analysis of Canine Heart Sounds, in Proceedings of the First Joint BMES/EMBS Conference, 2, 1999.
- [316] D. Chen, L.-G. Durand, and H.C. Lee, Application of the Cone-Kernel Distribution to Study the Effect of Myocardial Contractility in the Genesis of the First Heart Sound in Dog, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 2, 959–960, 1995.
- [317] J. Raz, Wavelet Models of Event-Related Potentials, in Proceedings International Conference of the IEEE Engineering in Medicine and Biology Society, 1, 191–192, 1994.
- [318] S.M. Kay and G.F. Boudreux-Bartels, On the Optimality of the Wigner Distribution for Detection, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Tampa, FL, March 1985, pp. 27.2.1–27.2.4.
- [319] A. Papandreou, G.F. Boudreux-Bartels, and S.M. Kay, Detection and Estimation of Generalized Chirps Using Time-Frequency Representations, in Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers, 1, 50–54, October/November 1994.
- [320] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, Upper Saddle River, NJ, 1993.
- [321] S.M. Kay, *Fundamentals of Statistical Signal Processing*, Vol. 2, *Detection Theory*, Prentice Hall, Upper Saddle River, NJ, 1998.

- [322] A. Papandreou, S.M. Kay, and G.F. Boudreux-Bartels, The Use of Hyperbolic Time-Frequency Representations for Optimum Detection and Parameter Estimation of Hyperbolic Chirps, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, October 1994, pp. 369–372.
- [323] A.M. Sayeed and D.L. Jones, Optimal detection using bilinear time-frequency and time-scale representations, *IEEE Trans. Signal Process.*, 43, 2872–2883, December 1995, also, corrections in the *IEEE Trans. Signal Process.*, March 1997, pp. 761–762.
- [324] S. Ramprashad, T.W. Parks, and R. Shenoy, Signal modeling and detection using cone classes, *IEEE Trans. Signal Process.*, 44, 329–338, February 1996.
- [325] A.M. Sayeed and D.L. Jones, Optimal Quadratic Detection Using Bilinear Time-Frequency and Time-Scale Representations, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, October 1994, pp. 365–368.
- [326] A.M. Sayeed and D.L. Jones, Optimum quadratic detection and estimation using generalized joint signal representations, *IEEE Trans. Signal Process.*, 44, 3031–3043, December 1996.
- [327] G. Matz and F. Hlawatsch, Time-Frequency Formulation and Design of Optimal Detectors, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 213–216.
- [328] M. Wang, A.K. Chan, and C.K. Chui, Linear frequency-modulated signal detection using radon-ambiguity transform, *IEEE Trans. Signal Process.*, 46, 571–586, March 1998.
- [329] E.J. Zalubas, J.C. O'Neill, W.J. Williams, and I.A.O. Hero, Shift and Scale Invariant Detection, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 3637–3640, April 1997.
- [330] W.J. Pielemeier, G.H. Wakefield, and M.H. Simoni, Time-frequency analysis of musical signals, in Proceedings of the IEEE, 84, 1216–1230, September 1996.
- [331] M.L. Fowler and L.H. Sibul, A unified formulation for detection using time-frequency and time-scale methods, in Proceedings of the 25th Asilomar Conference on Signals, Systems and Computers, 1, 637–642, November 1991.
- [332] T.P. Krauss, T.W. Parks, and R.G. Shenoy, Detection of Time-Frequency Concentrated Transient Signals, in Proceedings of the IEEE International Symposium on Information Theory, 1993, p. 88.
- [333] H. Ye, S.X. Ding, and G. Wang, Integrated design of fault detection systems in time-frequency domain, *IEEE Trans. Automatic Control*, 47, 384–390, February 2002.

- [334] A. Karasaridis, M. Maalej, S. Pantazopoulou, and D. Hatzinakos, Time-Frequency Analysis of Sensor Data for Detection of Structural Damage in Instrumented Structures, in Proceedings of the 13th International Conference on Digital Signal Processing, 2, 1997, 817–820.
- [335] T. Brotherton, T. Pollard, and D. Jones, Applications of Time-Frequency and Time-Scale Representations to Fault Detection and Classification, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, Canada, 1992, pp. 95–98.
- [336] S. Pon Varma, A. Papandreou-Suppappola, and S.B. Suppappola, Detecting Faults in Structures Using Time-Frequency Techniques, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 6, 3593–3596, May 2001.
- [337] M. Davy, H. Cottreau, and C. Doncarli, Loudspeaker Fault Detection Using Time-Frequency Representations, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 3329–3332, May 2001.
- [338] J.L. Navarro-Mesa, E. Lleida-Solano, and A. Moreno-Bilbao, A new method for epoch detection based on the Cohen’s class of time frequency representations, *IEEE Signal Process. Lett.*, 8, 225–227, August 2001.
- [339] H. Nam, H.-S. Kim, Y. Kwon, and S.-I. Yang, Speaker Verification System Using Hybrid Model with Pitch Detection by Wavelets, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 153–156.
- [340] L. Janer, New Pitch Detection Algorithm Based on Wavelet Transform, in Proceedings IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 165–168.
- [341] K. Wang and D.M. Goblirsch, Extracting Dynamic Features Using the Stochastic Matching Pursuit Algorithm for Speech Event Detection, in Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding, 1997, pp. 132–139.
- [342] Q. Zhao, Q. Gao, and H. Chi, Detection of Spectral Transition for Speech Perception Based on Time-Frequency Analysis, in Proceedings of the International Conference on Information Communications and Signal Processing, 1, 522–526, 1997.
- [343] R.M. Nickel and W.J. Williams, Using the Time Frequency Structure of Pitch Periods to Improve Speaker Verification Systems, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 145–148.
- [344] J. Fang, J. McLaughlin, L. Owsley, L. Atlas, and J. Sachs, Quadratic Detectors for Feature Extraction in Text-Independent Speaker Authentication, in Pro-

ceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, October 1994, pp. 644–647.

- [345] D.C. Braunreiter, H.-W. Chen, M.L. Cassabaum, J.G. Riddle, A.A. Samuel, J.F. Scholl, and H.A. Schmitt, On the Use of Space-Time Adaptive Processing and Time-Frequency Data Representations for Detection of Near-Stationary Targets in Monostatic Clutter, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, 2000, pp. 472–475.
- [346] T.K. Bhattacharya, G. Jones, and D. DiFilippo, Time frequency based detection scheme for missile approach warning system MAWS, in Proceedings of the Radar (Conference Publication Number 449), 539–543, 1997.
- [347] C. Xu, Z. Chai, S. Shu, and M. Zhu, SAR Detection of Moving Targets Using Approximate Wavelet Transform and Time-Frequency Analysis, in Proceedings of the IEEE International Symposium on Circuits and Systems, 4, 2561–2563, 1997.
- [348] J.P. Stephens, Advances in signal processing technology for electronic warfare, *IEEE Aerospace Electron. Syst. Mag.*, 11, 31–38, November 1996.
- [349] M.P. Fargues and W.A. Brooks, Comparative study of time-frequency and time-scale transforms for ultra-wideband radar transient signal detection, in Proceedings of the IEE Radar, Sonar Navigation, 142, 236–242, October 1995.
- [350] S. Haykin and T. Bhattacharya, Wigner-Ville Distribution: an Important Functional Block for Radar Target Detection in Clutter, in Proceedings of the 28th Asilomar Conference on Signals, Systems and Computers, 1, 68–72, 1994.
- [351] H.N. Kritikos and J. Teti, Radar Scattering: a Time-Frequency Perspective, in Proceedings of the Microwave Optoelectronal Conference, 2, 483–487, August 1997.
- [352] J. Xin and N.M. Bilgutay, Detection and Resolution of Multiple Targets Using Time-Frequency Techniques, in Proceedings of the IEEE Ultrasonics Symposium, 2, 1133–1137, 1994.
- [353] B.W. Gillespie and L.E. Atlas, Optimization of Time and Frequency Resolution for Radar Transmitter Identification, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 3, 1341–1344, March 1999.
- [354] X. Dermin and G. Lei, Wavelet Transform and Its Application to Autonomous Underwater Vehicle Control System Fault Detection, in Proceedings of the International Symposium on Underwater Technology, 2000, pp. 99–104.
- [355] P.M. Oliveira and V. Barroso, Data Driven Underwater Transient Detection Based on Time-Frequency Distributions, in Proceedings of the MTS/IEEE OCEANS 2000 Conference on Exhibition, 2, 1059–1063, September 2000.

- [356] E. Duflos, P. Hervy, F. Nivelle, S. Perrin, and P. Vanheeghe, Time-Frequency Analysis of Ground Penetrating Radar Signals for Mines Detection Applications, in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 1, 520–525, 1999.
- [357] B. Leprettre, N. Martin, F. Glangeaud, and J.-P. Navarre, Three-component signal recognition using time, time-frequency, and polarization information—application to seismic detection of avalanches, *IEEE Trans. Signal Process.*, 46, 83–102, January 1998.
- [358] T.A. de Perez, M. Juanatey, M.C. Stefanelli, M. Abondano, and A. Viloria, Earthquake Epicenter Location Using Time-Frequency Representations, in Proceedings of the 2nd IEEE International Caracas Conference on Devices, Circuits and Systems, 1998, pp. 217–221.
- [359] C. Rivero-Moreno and B. Escalante-Ramirez, Seismic Signal Detection with Time-Frequency Models, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 345–348.
- [360] G. Matz and F. Hlawatsch, Time-Frequency Methods for Signal Detection with Application to the Detection of Knock in Car Engines, in Proceedings of the IEEE-SP Workshop on Statistical Signal and Array Processing, Portland, OR, September 1998, pp. 196–199.
- [361] B. Samimy and G. Rizzoni, Mechanical signature analysis using time-frequency signal processing: application to internal combustion engine knock detection, in Proceedings of the IEEE, 84, 1330–1343, September 1996.
- [362] H. Laurent and C. Doncarli, Abrupt Changes Detection in the Time-Frequency Plane, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 285–288.
- [363] H. Laurent, E. Hitti, and M.-F Lucas, Abrupt Changes Detection in the Time-Scale and in the Time-Frequency Planes: A Comparative Study, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 501–504.
- [364] R. Schlag, U. Korell, B. Siegmund, B. Pfeiffer, and M. Pandit, Automatic Detection System of Venous Air Embolism Employing Signal Processing Methods, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 4, 1925–1928, May 1998.
- [365] P. Daponte, G. Fazio, and A. Molinaro, Detection of echoes using time-frequency analysis techniques, *IEEE Trans. Instrum. Meas.*, 45, 30–40, February 1996.
- [366] J.M. Spanjaard, P.J. Sherman, L.B. White, and S. Lau, Periodic Autoregressive Time-Frequency Analysis for Monitoring of Rotating Machinery with Variable

- Period, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 465–468.
- [367] A. Leshem, A.-J. van der Veen, and E. Deprettere, Detection and Blanking of GSM Interference in Radio-Astronomical Observations, in Proceedings of the 2nd IEEE Workshop on Signal Processing Advances in Wireless Communications, Annapolis, MD, May 1999, pp. 374–377.
- [368] R.M. Nickel and W.J. Williams, High Resolution Frequency Tracking via Non-negative Time-Frequency Distributions, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, 2000, pp. 612–615.
- [369] H. Lao and S. Zein-Sabatto, Analysis of Vibration Signal's Time-Frequency Patterns for Prediction of Bearing's Remaining Useful Life, in Proceedings of the 33rd Southeastern Symposium on System Theory, 2001, pp. 25–29.
- [370] G. Matz and F. Hlawatsch, Minimax Robust Time-Frequency Filters for Non-stationary Signal Estimation, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Phoenix, AZ, March 1999, pp. 1333–1336.
- [371] Z.M. Hussain and B. Boashash, Adaptive Instantaneous Frequency Estimation of Multi-Component FM Signals, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2, 657–660, June 2000.
- [372] M.G. Amin, Time-frequency spectrum analysis and estimation for non-stationary random processes, in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash, Ed., Longman-Cheshire, Melbourne, Australia, 1992, pp. 208–232.
- [373] S. Barbarossa and A. Scaglione, Parameter Estimation of Spread Spectrum Frequency-Hopping Signals Using Time-Frequency Distributions, in Proceedings of the 1st IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications, 1997, pp. 213–216.
- [374] Z.M. Hussain and B. Boashash, Multicomponent if Estimation: a Statistical Comparison in the Quadratic Class of Time-Frequency Distributions, in Proceedings of the IEEE International Symposium on Circuits and Systems, 2, 109–112, 2001.
- [375] K. Ghartey, D. Cochran, and A. Papandreou-Suppappola, Multi-Channel Signal Detection Using Time-Varying Estimation Techniques, in Proceedings of the IEEE 6th International Symposium on Signal Processing and its Applications, 2, 577–580, August 2001.
- [376] A.B. Gershman and M.G. Amin, Coherent Wideband DOA Estimation of Multiple FM Signals Using Spatial Time-Frequency Distributions, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 3065–3068, June 2000.

- [377] A. Gershman, M. Pesavento, and M. Amin, Estimating parameters of multiple wideband polynomial-phase sources in sensor arrays, *IEEE Trans. Signal Process.*, 49, 2924–2934, December 2001.
- [378] M. Davy, C. Doncarli, and G.F. Boudreux-Bartels, Improved optimization of time-frequency-based signal classifiers, *IEEE Signal Process. Lett.*, 8, 52–57, February 2001.
- [379] M. Davy and C. Doncarli, Optimal Kernels of Time-Frequency Representations for Signal Classification, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 581–584.
- [380] B.W. Gillespie and L.E. Atlas, Optimizing time-frequency kernels for classification, *IEEE Trans. Signal Process.*, 49, 485–496, March 2001.
- [381] D. Korosec and C. Doncarli, Enhanced Signal Classification Scheme Using a Selected Information in the Ambiguity Domain, in Proceedings of the 33rd Asilomar Conference on Signals, Systems and Computers, 1, 710–712, 1999.
- [382] P. Wellig and G.S. Moschytz, Classification of Time-Varying Signals Using Time-Frequency Atoms, in Proceedings of the 1st Joint BMES/EMBS Conference, 2, 953, 1999.
- [383] H. Ketterer, F. Jondral, and A.H. Costa, Classification of Modulation Modes Using Time-Frequency Methods, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 2471–2474, March 1999.
- [384] G. Roberts, A. Zoubir, and B. Boashash, Non-stationary Gaussian signal classification, in Proceedings of the Digital Signal Processing and its Applications, 2, 526–529, 1996.
- [385] L. Atlas, L. Owsley, J. McLaughlin, and G. Bernard, Automatic Feature-Finding for Time-Frequency Distributions, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, June 1996, pp. 333–336.
- [386] I. Vincent, C. Doncarli, and E. Le Carpentier, Nonstationary Signals Classification Using Time-Frequency, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, October 1994, pp. 233–236.
- [387] I. Moraitakis and M. Fargues, Feature Extraction of Intra-Pulse Modulated Signals Using Time-Frequency Analysis, in Proceedings Military Communication Conference, 2, 737–741, 2000.
- [388] S. Barbarossa and O. Lemoine, Analysis of nonlinear FM signals by pattern recognition of their time-frequency representation, *IEEE Signal Proc. Lett.*, 3, 112–115, April 1996.

- [389] A. Belouchrani and M.G. Amin, Time-frequency MUSIC, *IEEE Signal Process. Lett.*, 6, 109–110, May 1999.
- [390] A. Bultan and R. Haddad, Channel Estimation in Noisy Conditions Using Time-Frequency Domain Filtering, in Proceedings of the 33rd Asilomar Conference on Signals, Systems and Computers, 2, 1642–1646, 1999.
- [391] X.-G. Xia, Channel Estimation in Low SNR Environments Using Chirp Signals and Joint Time-Frequency Filters, in Proceedings of the IEEE 1st Workshop on Multimedia Signal Processing, 1997, pp. 383–388.
- [392] M.J. Fernandez-Getino Garcia, O. Edfors, and J.M. Paez-Borrallo, Joint Channel Estimation and Peak-to-Average Power Reduction in Coherent OFDM: a Novel Approach, in Proceedings of the IEEE Vehicular Technology Conference, 2, 815–819, 2001.
- [393] H. Bolcskei, Blind High-Resolution Uplink Synchronization of OFDM-Based Multiple Access Schemes, in Proceedings of the 2nd IEEE Workshop on Signal Processing Advances in Wireless Communications, 1999, pp. 166–169.
- [394] A. Klapuri, Pitch Estimation Using Multiple Independent Time-Frequency windows, in Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 1999, pp. 115–118.
- [395] A. Quinquis and E. Radoi, Accurate Estimation of the Time Delay in Radar Target Range Profile Reconstruction Using Time-Frequency and Superresolution Algorithms, in Proceedings of the 6th IEEE International Conference on Electronic Circuits and Systems, 3, 1449–1452, 1999.
- [396] N. Ma and D. Vray, Bottom Backscattering Coefficient Estimation from Wideband Chirp Sonar Echoes by Chirp Adapted Time-Frequency Representation, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 4, 2461–2464, May 1998.
- [397] A.B. Gershman, M. Pesavento, and M.G. Amin, Estimating the Parameters of Multiple Wideband Chirp Signals in Sensor Arrays, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, 2000, pp. 461–471.
- [398] S. Stanković and I. Djurović, Motion parameter estimation by using time-frequency representations, *Electron. Lett.*, 37, 1446–1448, November 2001.
- [399] A. Scaglione and S. Barbarossa, Estimating motion parameters using parametric modeling based on time-frequency representations, in Proceedings of the Radar 97 (Conference Publication Number 449), 280–284, 1997.
- [400] P. Azzoni, D. Moro, and G. Rizzoni, Time-Frequency Signal Analysis of the Acoustic Emission of Formula 1 Engines, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 441–444.

- [401] G. Azemi, B. Senadji, and B. Boashash, Estimation of the Velocity of Mobile Units in Micro-Cellular Systems Using the Instantaneous Frequency of the Received Signal, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Salt Lake City, UT, May 2001.
- [402] I. Djurovic and L. Stankovic, Influence of high noise on the instantaneous frequency estimation using quadratic time-frequency distributions, *IEEE Signal Process. Lett.*, 7, 317–319, November 2000.
- [403] B. Barkat and B. Boashash, Instantaneous frequency estimation of polynomial FM signals using the peak of the PWVD: statistical performance in the presence of additive Gaussian noise, *IEEE Trans. Signal Process.*, 47, 2480–2490, September 1999.
- [404] K.-T. Kim, I.-S. Choi, and H.-T. Kim, Efficient radar target classification using adaptive joint time-frequency processing, *IEEE Trans. Antennas Propagation*, 48, 1789 – 1801, December 2000.
- [405] P. Lombardo, Estimation of Target Motion Parameters from Dual-Channel SAR, in Proceedings of the IEEE National Radar Conference, 1997, pp. 13–18.
- [406] P.K. Kumar and K.M.M. Prabhu, Classification of Radar Returns Using Wigner-Ville Distribution, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 6, 3105–3108, May 1996.
- [407] C. Delfs and F. Jondral, Classification of Piano Sounds Using Time-Frequency Signal Analysis, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 3, 2093–2096, April 1997.
- [408] D.R. Franklin and J.F. Chicharo, Paganini-a Music Analysis and Recognition Program, in Proceedings of the 5th International Symposium on Signal Processing and its Applications, 1, 107–110, 1999.
- [409] I. Arroabarren, M. Zivanovic, J. Bretos, A. Ezcurra, and A. Carlosena, Measurement of Vibrato in Lyric Singers, in Proceedings of the IEEE Conference on Instrumentation Measurement Technology, 3, 1529–1534, 2001.
- [410] R. Keren, Y.Y. Zeevi, and D. Chazan, Multiresolution Time-Frequency Analysis of Polyphonic Music, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 565–568.
- [411] V. Pierson and P.-O. Amblard, Studying the Performances of Linear Signature Detector on a Time-Frequency Representation, in Proceedings of the IEEE Digital Signal Processing Workshop, 1996, pp. 423–426.
- [412] H.C. Strifors, A. Gustafsson, S. Abrahamson, and G.C. Gaunaurd, Fuzzy-Cluster Representation of Time-Frequency Signatures as a Means for Automatic Classification of Buried Mine-Like Targets, in Proceedings of the

IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 597–600.

- [413] B. Barkat, A.M. Zoubir, and C.L. Brown, Application of Time-Frequency Techniques for the Detection of Anti-personnel Landmines, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, PA, August 2000, pp. 594–597.
- [414] S. Kadambe and T. Adali, Application of Cross-Term Deleted Wigner Representation CDWR for Sonar Target Detection/Classification, in Proceedings of the 32nd Asilomar Conference on Signals, Systems and Computers, 1, 822–826, November 1998.
- [415] F. Lari and A. Zakhori, Automatic Classification of Active Sonar Data Using Time-Frequency Transforms, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, Canada, 1992, pp. 21–24.
- [416] S.S. Abeysekera, Underwater Shell Type Target Classification Using Time-Frequency Features of Acoustic Backscatter Signals, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 593–596.
- [417] G. Nicq and M. Brussieux, a Time-Frequency Method for Classifying Objects at Low Frequencies, in Proceedings of the OCEANS (MTS/IEEE) Conference on Exhibition, 1, 148–152, 1998.
- [418] N. Darvish and R.I. Kitney, Time-Frequency and Time-Scale Methods in the Detection and Classification of Non-stationarities in Human Physiological Data, in Proceedings Conference Record of the 28th Asilomar Conference on Signals, Systems and Computers, 2, 1085–1158, 1994.
- [419] B. Gillespie and L. Atlas, Data-Driven Time-Frequency Classification Techniques Applied to Tool-Wear Monitoring, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2, 11649–11652, June 2000.
- [420] J. Droppo and L. Atlas, Application of Classifier-Optimal Time-Frequency Distributions to Speech Analysis, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 585–588.
- [421] V.C. Chen and H. Ling, Joint time-frequency analysis for radar signal and image processing, *IEEE Signal Process. Mag.*, 16, 81–93, March 1999.
- [422] E.J. Rothwell, K.M. Chen, and D.P. Nyquist, An adaptive-window-width short-time Fourier transform for visualization of radar target substructure resonances, *IEEE Trans. Antennas Propagation*, 46, 1393–1395, September 1998.

- [423] Z. Bao, C. Sun, and M. Xing, Time-frequency approaches to ISAR imaging of maneuvering targets and their limitations, *IEEE Trans. Aerospace Electron. Syst.*, 37, 1091–1099, July 2001.
- [424] L.C. Trintinalia and H. Ling, Joint time-frequency ISAR using adaptive processing, *IEEE Trans. Antennas Propagation*, 45, 221–227, February 1997.
- [425] R. Fiedler and R. Jansen, Joint Time-Frequency Analysis of SAR Data, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, PA, August 2000, pp. 480–484.
- [426] Y. Ding, N. Xue, and J. Munson, An Analysis of Time-Frequency Methods in SAR Imaging of Moving Targets, in Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop, Cambridge, MA, March 2000, pp. 221–225.
- [427] S.-H. Yoon, B. Kim, and Y.-S. Kim, Helicopter classification using time-frequency analysis, *Electron. Lett.*, 36, 1871–1872, October 2000.
- [428] H.N. Kritikos and J. Teti, A time-frequency analysis method for radar scattering, *IEEE Trans. Microwave Theory Tech.*, 46, 257–260, March 1998.
- [429] S.L. Marple, Large Dynamic Range Time-Frequency Signal Analysis with Application to Helicopter Doppler Radar Data, in Proceedings of the 6th International Symposium on Signal Processing and its Applications, 1, 260–263, August 2001.
- [430] V.C. Chen and S. Qian, Joint time-frequency transform for radar range-Doppler imaging, *IEEE Trans. Aerospace Electron. Syst.*, 34, 486–499, April 1998.
- [431] V.C. Chen, Analysis of Radar Micro-Doppler with Time-Frequency Transform, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, PA, August 2000, pp. 463–466.
- [432] A. Hero, H. Messer, J. Goldberg, D.J. Thomson, M.G. Amin, G. Giannakis, A. Swami, J.K. Tugnait, A. Nehorai, A.L. Swindlehurst, J.-F. Cardoso, L. Tong, and J. Krolik, Highlights of statistical signal and array processing, *IEEE Signal Process. Mag.*, 15, 21–64, September 1998.
- [433] R. Imai, Y. Hashimoto, K. Kikuchi, and S. Fujii, High-resolution beamforming by the Wigner-Ville distribution method, *IEEE J. Oceanic Eng.*, 25, 105–110, January 2000.
- [434] R.M. Dizaji, N.B. Chapman, and R.L. Kirlin, Time-Frequency Matched Field Processing for Time Varying Sources, in Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop, Cambridge, MA, March, 2000, pp. 117–120.
- [435] P. Loughlin, D. Groutage, and R. Rohrbaugh, Time-Frequency Analysis of Acoustic Transients, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 3, 2125–2128, April 1997.

- [436] G.P. Zvara, Real time time-frequency active sonar processing: a SIMD approach, *IEEE J. Oceanic Eng.*, 18, 520–528, October 1993.
- [437] T. Brotherton, T. Pollard, R. Barton, A. Krieger, and S.L. Marple, Application of Time-Frequency and Time-Scale Analysis to Underwater Acoustic Transients, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, Canada, 1992, pp. 513–516.
- [438] R. Baraniuk, D. Jones, T. Brotherton, and S.L. Marple, Applications of Adaptive Time-Frequency Representations to Underwater Acoustic Signal Processing, in Proceedings of the 25th Asilomar Conference on Signals, Systems and Computers, 2, 1109–1113, November 1991.
- [439] F. Hlawatsch, A. Papandreou, and G.F. Boudreaux-Bartels, Regularity and Unitality of Affine and Hyperbolic Time-Frequency Representations, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 3, 245–248, April 1993.

Appendix A: Acronyms in Alphabetical Order

1-D	One-dimensional
2-D	Two-dimensional
AF	Ambiguity function
ASPWD	Affine smoothed pseudo-Wigner distribution
AUD	Active Unterberger distribution
AWGN	Additive white Gaussian noise
BUD	Butterworth distribution
CDMA	Code division multiple access
CT	Cross term
CWD	Choi–Williams distribution
DS-CDMA	Direct sequence code division multiple access
DSSS	Direct sequence spread spectrum
DWD	Dispersive Wigner distribution
ECG	Electrocardiogram
EEG	Electroencephalogram
EMG	Electromyography
ESPWD	Exponential smoothed pseudo–Wigner distribution
EWD	Exponential Wigner distribution
FH-CDMA	Frequency-hopped code division multiple access
FM	Frequency modulation
FSK	Frequency shift keying
FT	Fourier transform
GD	Group delay
GDS	Group delay shift
GED	Generalized exponential distribution
GPS	Global positioning system
IF	Instantaneous frequency
ISAR	Inverse synthetic aperture radar
JSR	Jamming-to-Signal ratio
LFM	Linear frequency modulated
LMST	Linear matched signal transform
LWD	Linearly warped Wigner distribution
MST	Matched signal transform
NMR	Nuclear magnetic resonance
PN	Pseudo noise
PSPWD	Power smoothed pseudo-Wigner distribution
PUD	Passive Unterberger distribution
PWD	Pseudo Wigner distribution
QD	Altes–Marinovich Q-distribution
QTFR	Quadratic time–frequency representation

SAR	Synthetic aperture radar
SCAL	Scalogram
SNR	Signal-to-noise ratio
SPEC	Spectrogram
SPQD	Smoothed pseudo Q-distribution
SPWD	Smoothed pseudo-Wigner distribution
TF	Time–frequency
TFR	Time–frequency representation
TFT	Time–frequency toolbox
TM	Time modulation
TMJ	Temporomandibular joint
TV	Time varying
TVR	Time-varying representation
TVT	Time-varying transform
WD	Wigner distribution
WWW	World Wide Web

Appendix B: Mathematical Notation in Alphabetical Order

\mathcal{C}_a	Scale change operator
\mathcal{D}_c	Dispersive group delay shift operator
\mathcal{H}_c	Hyperbolic group delay shift operator
\mathcal{L}_c	Linear group delay shift operator
\mathcal{M}_ν	Frequency shift operator
\mathcal{P}_c	Power group delay shift operator
\mathcal{S}_τ	Time shift (or constant group delay shift) operator
\mathcal{W}_ξ	Warping operator
\mathcal{W}_ξ^{-1}	Inverse of the warping operator
$\text{AF}_X(\tau, \nu)$	Ambiguity function
$K_T(t_1, t_2; t, f)$	Generalized form of a 4-D kernel for QTFRs
$P_\kappa \text{WD}_X(t, f)$	κ th power WD
$\text{QD}_X(t, f)$	Altes–Marinovich Q-distribution
$\Re[\alpha]$	Real part of α
$T_{x_1, x_2}(t, f)$	Cross quadratic time–frequency representation
$T_x(t, f) = T_X(t, f)$	Time–frequency representation
$T_X^{(A)}(t, f)$	Affine class QTFR
$T_X^{(C)}(t, f)$	Cohen’s class QTFR
$T_X^{(D)}(t, f)$	Dispersive class QTFR
$\text{WD}_X(t, f)$	Wigner distribution
$X(f)$	FT or spectrum of the signal $x(t)$
$\text{sgn}(f)$	Sign of f
f	Frequency variable
f_r	Positive normalization frequency
$h_T(\tau; t, f)$	Generalized form of a 3-D kernel for linear TFRs
t	Time variable
t_r	Positive reference time
$x(t)$	Time domain signal
$\Upsilon_x^{(\eta)}(b)$	Dual matched signal transform
$\Phi_T^{(A)}(b, \beta)$	Kernel in the affine QTFR class
$\Psi_T^{(A)}(\zeta, \beta)$	Kernel in the affine QTFR class
$\Psi_T^{(C)}(\tau, \nu)$	Kernel in Cohen’s QTFR class
$\eta(t/t_r)$	Signal phase function or characteristic basis function
$\lambda(u)$	Parameter function of the Bertrand P_κ -distributions
$\mu(u)$	Parameter function of the Bertrand P_κ -distributions
$\nu_x(t)$	Instantaneous frequency of $x(t)$
$\xi(f/f_r)$	Phase spectrum or characteristic basis function
$\xi^{-1}(f/f_r)$	Inverse of the characteristic basis function
$\tau_X(f)$	Group delay of $x(t)$
$\psi_T^{(A)}(c, b)$	Kernel in the affine QTFR class
$\psi_T^{(C)}(t, f)$	Kernel in Cohen’s QTFR class
$\aleph_X^{(\xi)}(c)$	Matched signal transform

Interference Excision via Time–Frequency Distributions: Applications to Global Positioning System Antijam

Alan R. Lindsey

Air Force Research Lab/IFGC

Liang Zhao and Moeness Amin

Villanova University

CONTENTS

2.1	Overview	85
2.2	Global Positioning System Signal Structure	88
2.3	Interference Mitigation in Direct Sequence Spread-spectrum Systems	89
2.4	Subspace Projection Techniques	91
2.5	Conclusions	117
	Acknowledgment	118
	References	118

2.1 Overview

The global positioning system (GPS) is a satellite-based, worldwide, all-weather navigation and timing system [10, 11, 18]. It was initiated by the Department of Defense primarily for the U.S. military to provide precise estimates of position, velocity and time for all resources with a compatible receiver. However, the civilian applications of GPS are growing at an astonishing rate, a phenomenon not completely understood during the design phase of the system. GPS has found applications in land transportation, civil aviation, maritime commerce, surveying and mapping, construction, mining, agriculture, Earth sciences, electric power systems, telecommunications and outdoor recreational activities. It is not an exaggeration to suggest that the GPS has transformed our lives in unmistakeable and irreversible ways. We can now plant fields with automatic steering controls, survey and map large plots of land with centimeter accuracy, enter a destination into a handheld computer with a GPS receiver attached and get precise maps and directions, dig and bulldoze and trench and fill

to specifications without transits or levels, and innumerable other everyday activities that rely on position or velocity or time information. Although these uses were certainly dreamed about in the early days of the GPS standup, they were by no means considered essential, and certainly were not anticipated to come about so quickly or to have such a strong foothold.

The GPS is composed of three segments — space, control and receiver:

1. The *space segment* consists of 24 satellites, each of which continuously transmits a ranging signal that includes the navigation message.
2. The *control segment* tracks each satellite and periodically uploads to the satellite its prediction of future satellite positions and satellite clock time corrections.
3. The *user receiver*, which is our main focus in the time–frequency (TF) interference rejection problem, tracks the ranging signals of selected satellites and calculates its three-dimensional (3-D) position, velocity and local time.

Each satellite transmits data packets, updated periodically, that contain information on packet transmission time, satellite position, trajectories, etc. By comparing the transmitting time and the receiving time of one satellite signal, the user can acquire the knowledge of its range from the satellite. With the knowledge of the satellite position, the user is known to be on the surface of a sphere centered around the satellite with a radius equal to the range. Therefore, with the accurate knowledge of the range from three satellites and their position, the user's position can be calculated as the intersection point of the three spheres. However, all receivers are not created equal. Different models use different quality components and no two electrical components behave exactly the same, so the clocks, especially on consumer models, are not precisely accurate. The range calculation mentioned, actually called pseudorange, does not account for this. However, a signal from an additional satellite allows the receiver to solve for all four unknowns — the three position coordinates and user clock bias. [Figure 2.1](#) depicts this concept of GPS navigation.

The ever-increasing reliance on GPS for navigation and guidance has created a growing awareness of the need for adequate protection against both intentional and unintentional noise in the GPS frequency band — the former is generally referred to as *jamming* and the latter, *interference*. Although interference is typically benign in motive, consisting of nothing more than innocent electromagnetic radiation from licensed transmitters that happen to have characteristics in conflict with the extremely low-power GPS signal, jamming is generally considered hostile and is directed mainly at military users and commercial aviation. This is not always the case, however, as a recent incident in France that denied GPS to a commercial flight during unannounced jamming exercises can attest. In any case, jamming is addressed in this chapter — and assumptions about the nature and characteristics of the interfering signals are driven by this presumption.

Jamming is a procedure that attempts to block reception of a desired signal by the intended receiver. In general terms, it is high power signal that occupies the same

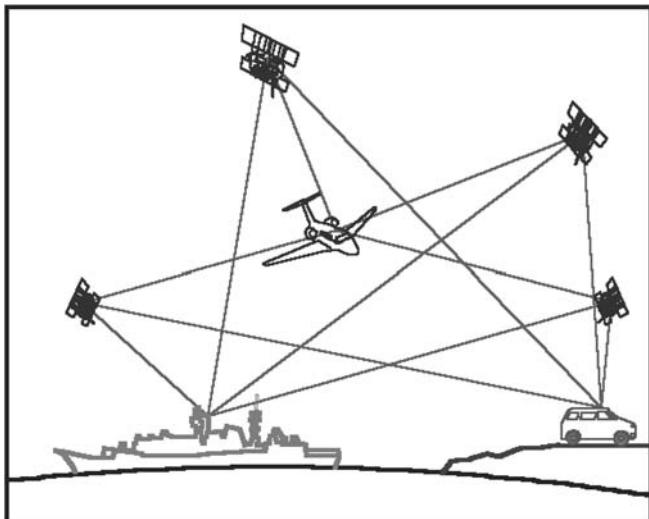


FIGURE 2.1
GPS navigation.

space, time slot or frequency spectrum as the desired signal, making reception by the intended receiver difficult or impossible. Designers of military as well as commercial communication systems have, through the years, developed numerous antijamming techniques to counter these threats. As these techniques become effective for interference removal and mitigation, jammer designers develop increasingly sophisticated signals that are difficult to mitigate. These “smart” jammers typically incorporate temporal nonstationarities or spatial diversity that complicates the identification and subsequent excision process.

All jammer mitigation strategies share one quality in common: dimensional isolation. In layman terms this amounts to “find out the jammer’s parameters and do not use them.” This works to the advantage of the receiver. Sheer economics mandates that no practical signal can be everywhere in space using all frequencies at all times. Thus, every jammer has its weaknesses — the compromise that had to be made between bandwidth and power and space. If it is very high power, it probably had to compromise on frequency or space coverage. If it is transmitted from multiple points in space, the compromise might be low power or a duty cycle or a narrowband frequency. If it is a broadband signal (not instantaneously narrowband), then it probably emanates from a specific point in space. Of course, putting a jammer in between the receiver and the transmitted signal and having it emit powerful broadband signals continuously would probably be very effective against one satellite. However, GPS has 24, more than 4 (and as many as 10) of which are always in view anywhere in the world; thus again, the advantage is to the receiver.

Furthermore, it is understood that many diverse means of mitigating the jamming signal are available including notched filters [6], direct jammer synthesis [13],

amplitude domain processing [1] and space-time processing [17], along with myriad solutions oriented around radiation pattern control and adaptive antennas [20]. In addition, several very good general papers exist on the GPS antijam problem — both of a broad overview nature [19] or assessment focused [14]. However, it is not the purpose of this chapter to rewrite these very good sources. Instead, the intent is to provide the reader with an introduction to a specific application of TF representations — that of GPS interference mitigation. Although a thorough tutorial on even this specific topic could fill volumes as well. To this end, the content focuses on one method, “orthogonal subspace projection” of utilizing TFRs for interference mitigation in GPS.

The task of the receiver is to identify the dimensionality of the jammer (space, time, frequency or combination of these) and then isolate it from that of the desired signal. This process, in most cases, involves some degradation of the transmitted signal. For instance, a jammer coming from a certain direction can be isolated in space via array processing methods that place a null in the direction of the jammer. The array pattern is designed such that the desired signal, coming from a different direction, is received. However, the major loss factors in this case are side lobes of the radiation pattern that still allow some jammer power as well as compromise in the receiving lobes, depending on the nearness of the jammer to the transmitter in space. If a jammer has a duty cycle less than 100%, it is possible to isolate it by only receiving the desired signal during off times. Of course the loss factor in this situation is the off time for the receiver. Similarly, jammers with narrowband frequency ranges can be isolated with various schemes that amount to identifying the jammer spectrum and filtering the received signal from those frequencies. Again, losses occur in ignoring the part of the desired signal that occupies the filtered band. For the purposes of this chapter, only spectral processing is considered.

2.2 Global Positioning System Signal Structure

GPS system is a code division multiple access (CDMA) system that employs BPSK-modulated direct sequence spread-spectrum (DSSS) signals. Signals from all satellites are transmitted at the same carrier frequencies: link 1 (L_1) at a frequency of 1575.42 MHz and link 2 (L_2) at 1227.6 MHz. The navigation data are transmitted at a symbol rate of 50 b/sec. It is spread by a coarse acquisition (C/A) code and a precision (P) code. The C/A code provides standard positioning service (resolution equal to or less than 100 m in the horizontal plane). It is available to all users worldwide. The C/A code is a Gold sequence with a chip rate of 1.023 MHz and a period of 1023 chips (i.e., its period is 1 msec, and there are 20 periods of the spreading code within each data bit). The P code can provide precise positioning service (equal to or less than 22 m in the horizontal plane), and mainly intended for military users. It is a pseudorandom code at the rate of 10.23 MHz and with a period of 1 week. These two

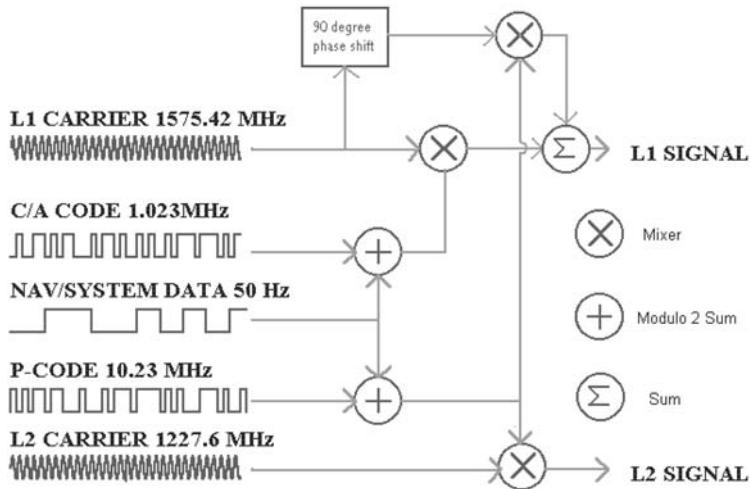


FIGURE 2.2
GPS signal generation diagram.

spreading codes are multiplexed in quadrature phases as shown in Figure 2.2. The carrier L_1 is modulated by both C/A code and P code, whereas the carrier L_2 is only modulated by P code.

The C/A code is of interest in this chapter, the peak power spectral density of which exceeds that of the P code by 13 dB [18]. The GPS signal is usually very weak, with signal-to-noise ratio (SNR) roughly -14 to -20 dB [8, 15] and jammer-to-signal ratio (JSR) easily larger than 40 dB. Due to the high JSR, and the noiselike spread-spectrum characteristic of the satellite signals, a frequency modulated (FM) instantaneously narrowband jammer can have a clear signature in the TF domain. It is noted here that because the P code is very weak compared with the C/A code, noise and jammer, its presence is ignored in subsequent analysis.

2.3 Interference Mitigation in Direct Sequence Spread-spectrum Systems

Spread-spectrum systems are implicitly able to provide a certain degree of protection against jammers or unintentional interferers. However, the jammer power may be made to be much stronger than the desired signal power, and the spreading gain may be insufficient to overcome this power deficit. For this reason, many methods have been proposed for interference suppression in DSSS communications. It is not within the scope of this chapter to detail these efforts, because that has been done very well

and very thoroughly in the published literature. Two sources that deal primarily with TF methods applied to general DSSS interference mitigation are [3] and [4], both of which have extensive reference lists that cover the topic in great detail for the reader interested in deeper understanding.

Applications of these mitigation techniques to the GPS system are much more focused, however. Because the parameters for the GPS signal are publicly disseminated and well understood, they need to be customized and perhaps modified to current techniques for the specific case. This is the story for the developments considered in this chapter. For instance, the GPS signal frequencies are known; thus jammers are most effective when the power is concentrated around those frequencies. However, concentrating jammer frequencies is an advantage to the receiver because identification and characterization is simplified.

Bilinear time–frequency distributions (TFDs) for signal power localization in the TF plane have motivated several new effective approaches for excision of nonstationary jammers. In the following, a new and effective precorrelation filter using projection techniques is developed to reject nonstationary jammers that can be accurately characterized by an instantaneous frequency parameter.

The BPSK-modulated DSSS signal may be expressed as:

$$s(t) = \sum_i I_i b_i(t - iT_b) \quad (2.1)$$

where $I_i \in \{-1, 1\} \forall i$ represents the binary data information sequence and T_b is the bit interval, which is 20 msec in the case of GPS systems. The i th binary information signal $b_i(t)$ is further decomposed as a superposition of L time-translated spreading chips $p(n)$, pulse shaped by a unit–energy function $q(t)$ of duration of τ_c , which is 1/1023 msec in the case of C/A code. Accordingly:

$$b_i(t) = \sum_{n=1}^L p_i(n)q(t - n\tau_c) \quad (2.2)$$

The signal for one data bit at the receiver, after demodulation, and sampling at chip rate, becomes:

$$x(n) = p(n) + w(n) + j(n) \quad 1 \leq n \leq L \quad (2.3)$$

where $p(n)$ is the chip sequence, $w(n)$ is the complex white noise and $j(n)$ is the jamming signal. The variables in Equation (2.3) can be written in the vector form:

$$\mathbf{x} = \mathbf{p} + \mathbf{w} + \mathbf{j} \quad (2.4)$$

where:

$$\begin{aligned} \mathbf{x} &= [x(1) \ x(2) \ x(3) \ \cdots \ x(L)]^T \\ \mathbf{p} &= [p(1) \ p(2) \ p(3) \ \cdots \ p(L)]^T \\ \mathbf{w} &= [w(1) \ w(2) \ w(3) \ \cdots \ w(L)]^T \\ \mathbf{j} &= [j(1) \ j(2) \ j(3) \ \cdots \ j(L)]^T \end{aligned}$$

All terms are complex except for the real-valued pseudonoise (PN) sequence.

“The method presented herein focuses on recovering the navigation data. It assumes perfect alignment between the received C/A code and the receiver C/A code. Therefore, it necessitates the unbiased preservation of the peak location of the correlation of the above codes because this location is required to determine the pseudo range of the receiver for each satellite. It can be shown that the proposed projection technique causes insignificant bias to the correlation peak location, specifically when the interference is modeled by an FM signal.”

2.4 Subspace Projection Techniques

2.4.1 Periodic signals in the time–frequency domain

For GPS C/A code, the PN sequence is periodic. The PN code of length 1023 repeats itself 20 times within 1 symbol of the 50 bps navigation data. Consequently, it is no longer of a continuous spectrum in the frequency domain, but instead of spectral lines. The case is the same for periodic jammers. Figures 2.4 and 2.3 show the effect of periodicity of the signal and the jammer on their respective power distribution over time and frequency, using the Wigner–Ville distribution. In both figures, a PN sequence of length 32 samples that repeats 8 times is used. A nonperiodic chirp jammer of a 50 dB JSR is added in Figure 2.3. A periodic chirp jammer of 50 dB JSR with the same period as the C/A code is included in Figure 2.4.

Note that the chosen value of 50 dB JSR has a practical significance. The spread-spectrum systems, with their interference immunity can tolerate, in a typical GPS C/A code receiver, a narrowband interference of approximately 40 dB JSR. However, field tests show that jammer strength often exceeds that value due to the weakness of

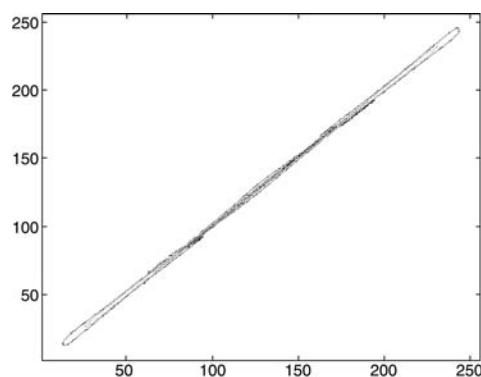
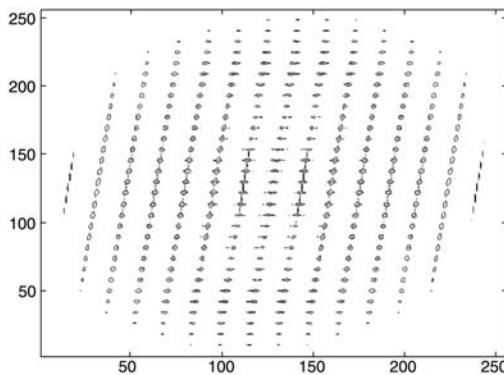


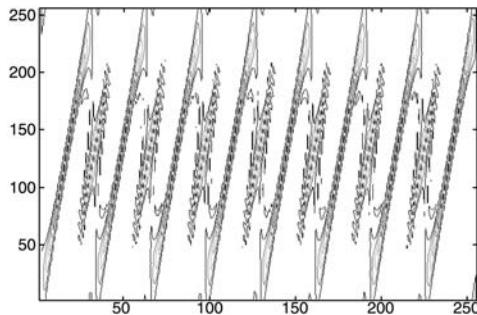
FIGURE 2.3

Periodic signal corrupted by a nonperiodic jammer in the time–frequency (TF) domain.

**FIGURE 2.4**

Periodic signal corrupted by a periodic jammer in the TF domain.

the signal. SNR in both figures are -20 dB, which is also close to its practical value [7, 8, 15]. Due to high JSR, the jammer is dominant in both figures. By comparing Figure 2.3, with its sharp instantaneous frequency (IF) line, to Figure 2.4, with no apparent feature whatsoever, it is clear that the periodicity of the jammer brings more difficulty to IF estimation using TFD. This problem can be solved by applying a short data window when using pseudo-Wigner–Ville distribution. Note that the window length should be less than the jammer period. Figure 2.5 shows the result of applying a window of length 31 to the same data used in Figure 2.3. The horizontal discrete harmonic line artifacts of the long window are effectively processed out leaving a much clearer picture of the true IF information.

**FIGURE 2.5**

Periodic signal corrupted by a periodic jammer in the TF domain (short window).

2.4.2 Symbol-period jammers

The concept of subspace projection for instantaneously narrowband jammer suppression is to remove the jammer components from the received data by projecting it onto the subspace that is orthogonal to the jammer subspace, as illustrated in Figure 2.6.

Once the IF of the nonstationary jammer is estimated from the TF domain, or by using any other IF estimator [7, 9, 12], the interference signal vector \mathbf{j} in Equation (2.4) can be constructed, up to ambiguity in phase and possibly in amplitude. This ambiguity is not detrimental though, as the reader can see.

In the proposed interference excision approach, the data vector is partitioned into Q blocks, each of length P (i.e., $L = PQ$). For the GPS C/A code, $Q = 20$, $P = 1023$ and all Q blocks are identical. Block-processing provides the flexibility to discard the portions of the data bit, over which significant errors occur in the IF estimates. The orthogonal projection method makes use of the fact that, in each block, the jammer has a one-dimensional (1-D) subspace J in the P -dimensional space V , which is spanned by the received data vector. The interference can be removed from each block by projecting the received data on the corresponding orthogonal subspace G of the interference subspace J . The subspace J is estimated using the IF information. Because the FM jammer signals are uniquely characterized by their IFs, the i th FM jammer samples in the k th block can be expressed as:

$$u_k(i) = \frac{1}{\sqrt{P}} \exp[j\phi_k(i)] \quad (2.5)$$

The projection matrix for k th block is given by:

$$\mathbf{V}_k = \mathbf{I} - \mathbf{u}_k \mathbf{u}_k^H \quad (2.6)$$

The vector \mathbf{u}_k is the unit norm basis vector in the direction of the interference vector of the k th block, and superscript H denotes vector or matrix Hermitian. The result

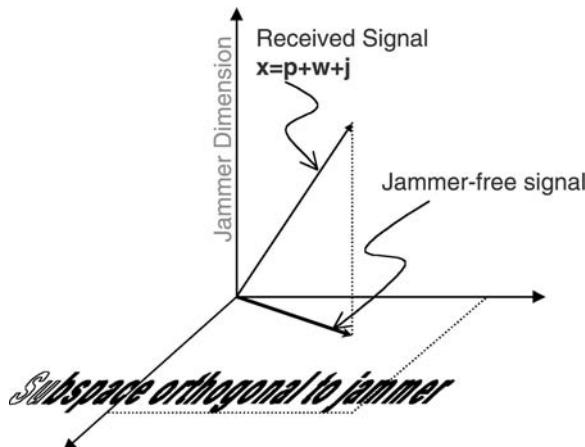


FIGURE 2.6

Interference mitigation via projection method.

of the projection over the k th data block is:

$$\bar{\mathbf{x}}_k = \mathbf{V}_k \mathbf{x}_k \quad (2.7)$$

where \mathbf{x}_k is the input data vector. By using the three different components that make up the input vector in Equation (2.4), the output of the projection filter \mathbf{V}_k can be written as:

$$\bar{\mathbf{x}}_k = \mathbf{V}_k(\mathbf{p}_k + \mathbf{w}_k + \mathbf{j}_k) \quad (2.8)$$

The noise is assumed to be complex white Gaussian with zero mean:

$$E[w(n)] = 0, E[w(n)^* w(n+l)] = \sigma^2 \delta(l), \forall l \quad (2.9)$$

where $\delta(\cdot)$ is the Dirac delta function. By assuming total interference excision through the projection operation, then:

$$\begin{aligned} \mathbf{V}_k \mathbf{j}_k &= \mathbf{0} \\ \bar{\mathbf{x}}_k &= \mathbf{V}_k \mathbf{p}_k + \mathbf{V}_k \mathbf{w}_k \end{aligned} \quad (2.10)$$

The decision variable y_r is the real part of y , which is obtained by correlating the filter output $\bar{\mathbf{x}}_k$ with the corresponding k th block of the receiver PN sequence and summing the results over the Q blocks. That is:

$$y = \sum_{k=0}^{Q-1} \bar{\mathbf{x}}_k^H \mathbf{p}_k \quad (2.11)$$

The PN code is periodic so $\mathbf{p}_k = \mathbf{p} \forall k$, and the preceding variable can be written in terms of the constituent signals as:

$$y = \sum_{k=0}^{Q-1} \mathbf{p}^T \mathbf{V}_k \mathbf{p} + \sum_{k=0}^{Q-1} \mathbf{w}^H \mathbf{V}_k \mathbf{p} \underline{\Delta} y_1 + y_2 \quad (2.12)$$

where y_1 and y_2 are the contributions of the PN and noise sequences to the decision variable, respectively. In common spread-spectrum systems, y_1 is considered as a random variable. However, in GPS system, due to the fact that each satellite is assigned a fixed Gold code [11, 18], and that the Gold code is the same for every navigation data symbol, y_1 can no longer be treated as a random variable, but instead as a deterministic value. This is a key difference between the GPS system and other spread-spectrum systems. The value of y_1 is given by:

$$\begin{aligned} y_1 &= \sum_{k=0}^{Q-1} \mathbf{p}^T \mathbf{V}_k \mathbf{p} = \sum_{k=0}^{Q-1} \mathbf{p}^T (\mathbf{I} - \mathbf{u}_k \mathbf{u}_k^H) \mathbf{p} \\ &= \sum_{k=0}^{Q-1} (\mathbf{p}^T \mathbf{p} - \mathbf{p}^T \mathbf{u}_k \mathbf{u}_k^H \mathbf{p}) = QP - \sum_{k=0}^{Q-1} (\mathbf{p}^T \mathbf{u}_k \mathbf{u}_k^H \mathbf{p}) \end{aligned} \quad (2.13)$$

Define:

$$\beta_k = \frac{\mathbf{p}^T \mathbf{u}_k}{\sqrt{P}} \quad (2.14)$$

as the complex correlation coefficient between the PN sequence vector \mathbf{p} and the jammer vector \mathbf{u} . β reflects the component of the signal that is in the jammer subspace, and represents the degree of resemblance between the signal sequence and the jammer sequence. Because the signal is a PN sequence, and the jammer is a nonstationary FM signal, the correlation coefficient is typically very small. With the preceding definition, y_1 can be expressed as:

$$y_1 = P(Q - \sum_{k=0}^{Q-1} |\beta_k|^2) \quad (2.15)$$

From Equation (2.15), it is clear that y_1 is a real value, which is the result of the fact that the projection matrix \mathbf{V} is Hermitian. With the noise assumptions in Equation (2.9), y_2 is complex white Gaussian with zero mean. Therefore:

$$\begin{aligned} \sigma_{y_2}^2 &= \sigma E[|y_2|^2] = E \left[\left(\sum_{k=0}^{Q-1} \mathbf{w}^H \mathbf{V}_k \mathbf{p} \right)^H \left(\sum_{k=0}^{Q-1} \mathbf{w}^H \mathbf{V}_k \mathbf{p} \right) \right] \\ &= \sum_{k=0}^{Q-1} \sum_{l=0}^{Q-1} \mathbf{p}^T \mathbf{V}_k E[\mathbf{w}_k \mathbf{w}_l^H] \mathbf{V}_l \mathbf{p} = \sum_{k=0}^{Q-1} \mathbf{p}^T \mathbf{V}_k E[\mathbf{w}_k \mathbf{w}_k^H] \mathbf{V}_k \mathbf{p} \\ &= \sigma^2 \sum_{k=0}^{Q-1} \mathbf{p}^T \mathbf{V}_k \mathbf{V}_k \mathbf{p} = \sigma^2 \sum_{k=0}^{Q-1} \mathbf{p}^T \mathbf{V}_k \mathbf{p} = \sigma^2 y_1 \end{aligned} \quad (2.16)$$

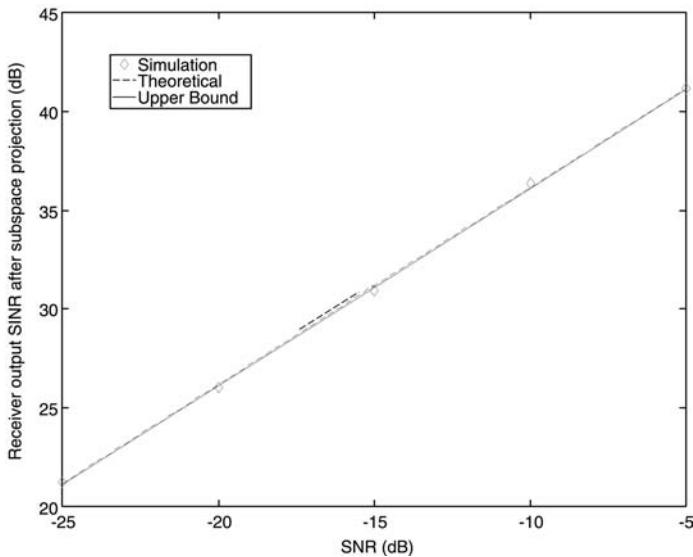
Because the decision variable y_r is the real part of y , then:

$$y_r = y_1 + \operatorname{Re}[y_2] \quad (2.17)$$

where $\operatorname{Re}\{y_2\}$ denotes the real part of y_2 and is real white Gaussian with zero mean and variance $\frac{1}{2}\sigma_{y_2}^2$. Therefore, the signal-to-interference-and-noise ratio (SINR) is:

$$\begin{aligned} SINR &= \frac{y_1^2}{\operatorname{var}\{\operatorname{Re}\{y_2\}\}} = \frac{y_1^2}{\frac{1}{2}\sigma_{y_2}^2} = \frac{2y_1}{\sigma^2} \\ &= 2P(Q - \sum_{k=0}^{Q-1} |\beta_k|^2) / \sigma^2 \end{aligned} \quad (2.18)$$

In the absence of jammers, no excision is necessary, and the SINR (in this case SNR) of the receiver output becomes $2PQ/\sigma^2$, which represents the upper bound for the anti-jamming performance. Clearly, $\frac{2P}{\sigma^2} \sum_{k=0}^{Q-1} |\beta_k|^2$ is the reduction in the receiver performance caused by the proposed jammer suppression techniques. This quantity reflects the energy of the power of the signal component that is in the jammer subspace. If the jammer and spread-spectrum signals are orthogonal (i.e., their correlation

**FIGURE 2.7**

Receiver SINR vs. SNR.

coefficient $|\beta| = 0$), then interference suppression is achieved with no loss in performance. However, as stated earlier, in the general case, $|\beta|$ is often very small; thus, the projection technique can excise FM jammers effectively with only insignificant signal loss. The lower bound of SINR is zero and corresponds to $|\beta| = 1$. This case requires the jammer to assume the C/A code (i.e., identical and synchronous with actual one). Figure 2.7 depicts the theoretical SINR in Equation (2.18), its upper bound and estimated values using computer simulation. The SNR assumes 5 different values $[-25, -20, -15, -10, -5]$ dB. In this figure, the signal is the Gold code of satellite SV 1, and the jammer is a periodic chirp FM signal with normalized frequency 0 to 0.5 and has the same period as the C/A code. For this case, the correlation coefficient $|\beta|$ is very small, $|\beta| = 0.038$. The JSR used in the computer simulation is set to 50 dB. Due to the large computation involved, 1000 realizations for each SNR value are simulated. Figure 2.7 demonstrates that the theoretical value of SINRs is almost the same as the upper bound, and both values are very close to the simulation result. In the simulation as well as in the derivation of Equation (2.18), exact knowledge of the jammer IF is assumed. Inaccuracies in the IF estimation can have an effect on the receiver performance [5].

2.4.3 General periodic jammers

The derivation of the receiver performance in the previous section implicitly assumes that the jammer period, T_j , is equal to the data symbol length, T_g , of the GPS signal (i.e., $T_j = T_g = PQ$). This is not as concocted as it sounds, because there are good reasons for a jammer to synchronize with the symbol period. However, an extremely subtle change in the performance results when general periodic jammers

are considered, having its origins in the nonergodicity of the underlying random process of computing SINR for an unknown but deterministic jammer. In the previous analysis, the decision variable does not change from symbol to symbol because the jammer is assumed to be periodic over each symbol. In the general case, the jammer is still assumed to be fractionally periodic (i.e., inside symbol boundaries). In this case the decision variable for successive symbols is different, leading to the necessity for statistical analysis in developing the SINR relationship.

By considering the more general case, the jammer is presumed to be a periodic signal with $T_j = \frac{N}{M} P$, where N and M are integers and relatively prime. Recall that $P = 1023$ is the period of the C/A spreading code. From the preceding definition, M jammer periods extend over N blocks of the GPS signal, and, as such, different segments of the jammer signature can infringe on different blocks in the symbol. This framework also applies, without loss of generality, to the case where the jammer period extends over multiple data symbols as well. However, it is now necessary to add the subscript i to associate the receiver decision variables with the i th symbol. It must be stressed that the projection process still applies to each of the Q blocks in the symbol separately, and as such, the dimensionality does not change. The projection matrix \mathbf{V} , which used to depend only on k , now also depends on the symbol index i . It is straightforward to show that the correlation output at the receiver is:

$$y_i = \sum_{k=0}^{Q-1} \mathbf{p}^T \mathbf{V}_{ik} \mathbf{p} + \sum_{k=0}^{Q-1} \mathbf{w}^H(n) \mathbf{V}_{ik} \mathbf{p} \Delta y_{i1} + y_{i2} \quad (2.19)$$

The decision variable is the real part of y_i . It can be shown that:

$$y_{i1} = P(Q - \sum_{k=0}^{Q-1} |\beta_{ik}|^2) \quad (2.20)$$

On the other hand, the correlation output due to the noise y_{i2} is a complex Gaussian zero-mean random variable, and its variance can be readily obtained as:

$$\sigma_{y_{i2}}^2 = \sigma^2 P(Q - \sum_{k=0}^{Q-1} |\beta_{ik}|^2) \quad (2.21)$$

Because there are N symbols for every M jammer periods, and $M \neq N$, both variables y_{i1} and y_{i2} assume different values over N consecutive GPS symbols. The jammer can then be cast as symbol dependent, assuming N distinct waveforms. In this case, one simple measure of the receiver performance is to average the SINR over N consecutive symbols, that is:

$$\begin{aligned} SINR_{av} &= E[SINR_i] \\ &= \sum_{i=1}^N \Pr(SINR_i | J_i) \Pr(J_i) \end{aligned}$$

$$= \frac{1}{N} \sum_{i=1}^N SINR_i \quad (2.22)$$

where $SINR_i$ and $SINR_{av}$ denote the receiver SINR over the i th symbol and the average receiver SINR, respectively. In Equation (2.22), $SINR_i$ is treated as a discrete random variable that takes N possible values with equal probability. $J_i (i = 1, \dots, N)$ are the segments of the jammer signal over N consecutive symbols. In Equation (2.22) $\Pr(x)$ denotes the probability of the event x and $\Pr(J_i) = 1/N$. The $SINR_i$ is:

$$SINR_i = \frac{2P \left(Q - \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)}{\sigma^2} \quad (2.23)$$

Accordingly:

$$SINR_{av} = \frac{2P \left(Q - \frac{1}{N} \sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)}{\sigma^2} \quad (2.24)$$

The expression in Equation (2.24) although simple to calculate, smoothes out high and low SINR values. In this regard, the average value in Equation (2.24) does not properly penalize poor or reward good receiver performance. Further, it is difficult to establish a relationship between the receiver $SINR_{av}$ and its basic electrical rhythm (BER). Most importantly, the expression in Equation (2.24) does not account for the self-noise term that reflects the level of signal distortion produced by the induced correlation of the code chips as a result of the excision process. Hence, a more proper way to measure the receiver performance is to deal with y_1 as a random variable. In this case, the average receiver SINR is referred to as \overline{SINR} to distinguish it from Equation (2.18). Assert that symbol “1” is transmitted and contaminated by one of N possible jammer signals occurring with the same probability. In this case, the mean value and the variance of the correlator output due to the GPS signal can be derived as:

$$\begin{aligned} E[y_1] &= \sum_{i=1}^N E[y_1|J_i] P_r(J_i) = \frac{1}{N} \sum_{i=1}^N E[y_1|J_i] \\ &= \frac{\sum_{i=1}^N \left[P \left(Q - \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right) \right]}{N} = PQ \left(1 - \frac{1}{NQ} \sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right) \end{aligned} \quad (2.25)$$

$$\begin{aligned} \sigma_{y_1}^2 &= E[y_1^2] - E^2[y_1] \\ &= \frac{1}{N} \sum_{i=1}^N P^2 \left(Q - \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)^2 - P^2 Q^2 \left(1 - \frac{1}{NQ} \sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)^2 \\ &= \frac{P^2}{N} \sum_{i=1}^N \left(\sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)^2 - \frac{P^2}{N^2} \left(\sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)^2 \end{aligned} \quad (2.26)$$

Similarly, the average noise power is:

$$\begin{aligned}\sigma_{y_2}^2 &= E[y_2^2] = \sum_{i=1}^N E[y_2^2|J_i] P_r(J_i) \\ &= \frac{1}{N} \sum_{i=1}^N E[y_2^2|J_i] = \frac{\sum_{i=1}^N \sigma^2 P \left(Q - \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)}{N} \\ &= \sigma^2 PQ \left(1 - \frac{1}{NQ} \sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)\end{aligned}\quad (2.27)$$

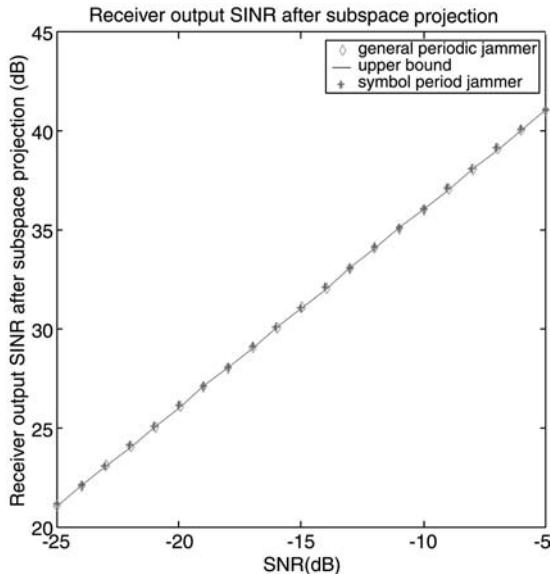
From Equations (2.25) to (2.27) and (2.18), the average *SINR* is given by:

$$\begin{aligned}\overline{SINR} &= \frac{E^2[y_1]}{\sigma_{y_1}^2 + \frac{1}{2} \sigma_{y_2}^2} \\ &= \frac{PQ^2 \left(1 - \frac{1}{NQ} \sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)^2}{\frac{P}{N} \sum_{i=1}^N \left(\sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)^2 - \frac{P}{N^2} \left(\sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)^2 + \frac{1}{2} \sigma^2 Q \left(1 - \frac{1}{NQ} \sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)}\end{aligned}\quad (2.28)$$

This expression represents the *SINR* of the receiver implementing subspace projections to remove a periodic jammer and is also valid for the case in which the jammer assumes N possible waveforms. In the case that the jammer has the same period as the GPS data symbol, then $N = 1$, and Equation (2.28) simplifies to Equation (2.18). By comparing Equation (2.28) to Equations (2.24) and (2.18), it is clear that Equation (2.28) includes the self-noise component $\sigma_{y_1}^2$ that arises due to the differences in the distortion effects of interference excision on the GPS signal over N symbols. In the absence of jamming, no excision is necessary, and the *SINR* of the receiver output becomes $2PQ/\sigma^2$, which represents the upper bound of the interference suppression performance. Moreover, if the jammer and the spreading codes are orthogonal (i.e., $\beta_{ik} = 0$), the interference suppression is also achieved with no loss in optimum receiver performance. It is noted, however, that the values of the cross-correlation coefficient $|\beta_{ik}|$ between the PN sequence signal and the nonstationary FM jammer are typically very small. This allows the proposed projection technique to excise FM jammers effectively with insignificant signal loss. Computer simulations show that ranges from 0 to 0.14. With these values, the self-noise $\sigma_{y_1}^2$ is negligible compared with the Gaussian noise for the low SNR conditions that often prevail in GPS environment. In this case, Equation (2.28) can be simplified to the following:

$$\overline{SINR} \approx \frac{E^2[y_1]}{\frac{1}{2} \sigma_{y_2}^2} = \frac{2PQ \left(1 - \frac{1}{NQ} \sum_{i=1}^N \sum_{k=0}^{Q-1} |\beta_{ik}|^2 \right)}{\sigma^2}\quad (2.29)$$

which is similar to Equation (2.18) and has the same form as Equation (2.24). Therefore, $SINR_{av}$ and \overline{SINR} approximately yield the same performance measure.

**FIGURE 2.8**

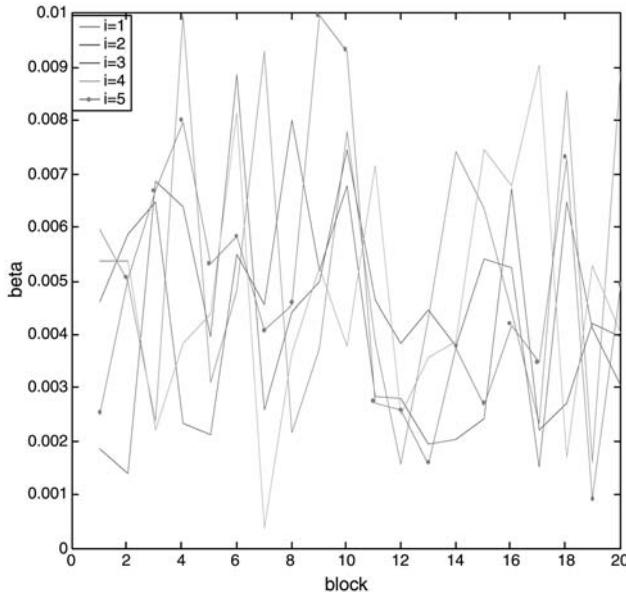
Output SINR vs. SNR (general periodic jammer).

Figure 2.8 plots the receiver SINR vs. SNR according to Equation (2.28) for the two cases of $N = 1$ and $N/M = 5/3$. In both cases, the normalized start and end frequencies of the chirp jammer are 0 and 0.5, respectively. The SNR values range from -25 to -5 dB, and the GPS signal is the Gold code of satellite SV 1. It is clear that the period of the jammer has little effect on the result of interference suppression performance, because both SINR curves are very close to the upper bound. These figures clearly show that the SINR change linearly with the input SNR, which can be easily recognized from Equation (2.29). Figure 2.9 shows the $|\beta_{ik}|$ values for the underlying example. It is evident from this figure that there is no clear pattern in the variation of the cross-correlation coefficients. The range values of $|\beta_{ik}|$ do not change greatly over different symbols. Experiments with different N , M , chirp rate and satellite signals have given similar results.

2.4.4 Discussions on cross-correlation coefficients

The cross-correlation coefficient is defined in the time domain in Equation (2.14). Moreover, the Parseval theorem of discrete Fourier transform (DFT) provides a clear link between the time domain correlation and the frequency domain correlation:

$$\sum_{n=0}^{P-1} p^*[n]u[n] = \frac{1}{N} \sum_{k=0}^{P-1} X^*[k]U[k] \quad (2.30)$$

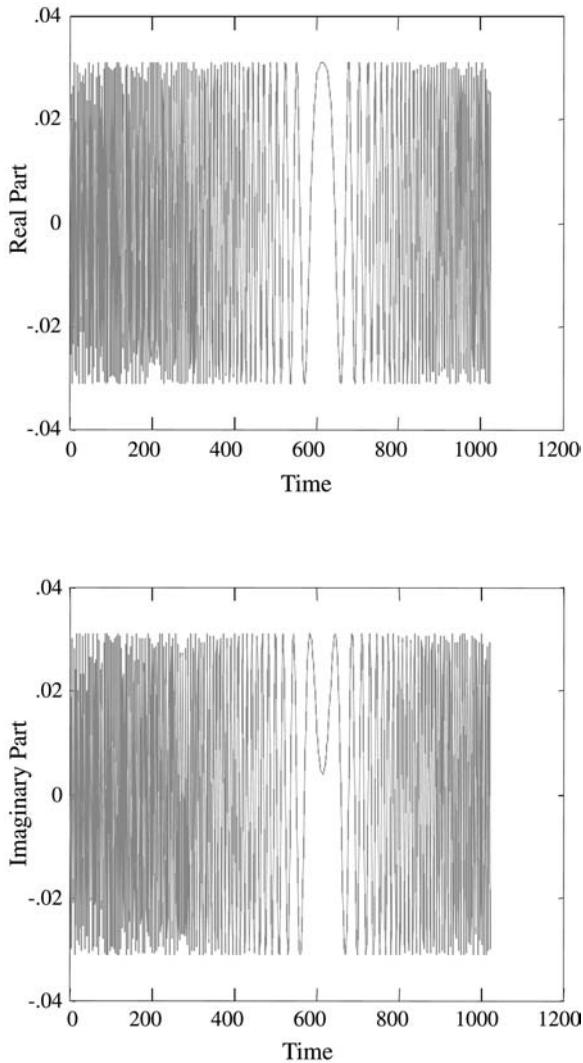
**FIGURE 2.9**

$|\beta_{ik}|$ for different symbols and blocks.

where $X[k]$ and $U[k]$ are the DFT of \mathbf{p} and \mathbf{u}_k , respectively. The conjugate sign on $p[n]$ can be dropped because the Gold code is real. Insights can be gained into the jamming excision problem using frequency domain analysis.

Figure 2.10 shows the real and imaginary parts of a normalized chirp signal with length $N = 1023$, and frequency approximate range $[-0.6\pi, 0.4\pi]$. This figure depicts how the IF of the chirp signal changes. Figure 2.11 shows the amplitude and the phase of the fast Fourier transform (FFT) of the same chirp signal, with the frequency range clearly evident. The amplitude in frequency over $[-0.6\pi, 0.4\pi]$ is flat as a result of the linear change of the frequency over this band. Outside the signal frequency band, the amplitude drops, indicating small power presence. Therefore, only $U[k]$ within the chirp frequency range contributes to the cross-correlation coefficient. One might mistakenly conclude that the narrower the chirp frequency range, the smaller the cross-correlation coefficient becomes. This is not the case, however, because the amplitude of the flat spectrum is larger for narrower bands, with all chirp signals assuming equal power. Figure 2.12 through Figure 2.14 depict the frequency domain of chirp signals with length $N = 1023$ and frequency ranges of $[0, 0.2\pi]$, $[-\pi, \pi]$ and $[-\pi, -0.9666\pi]$, respectively. These figures show the inverse relationships between the spectrum amplitude and its frequency range. It is noted that Figure 2.11 may be typical for those jammers that contaminate the entire GPS signal band. On the other hand, Figure 2.14 is representative of slowly varying jammers that might sweep through different blocks.

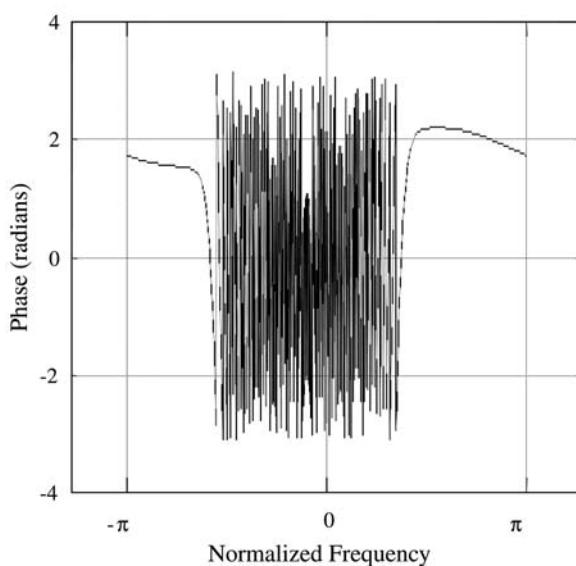
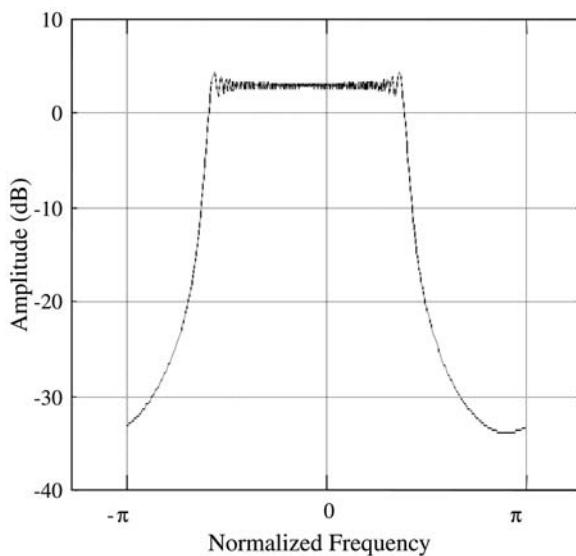
Figure 2.15 shows the magnitude and phase of the Gold code of satellite 1 in the frequency domain. Both plots have no spectral coherence due to the fact that

**FIGURE 2.10**

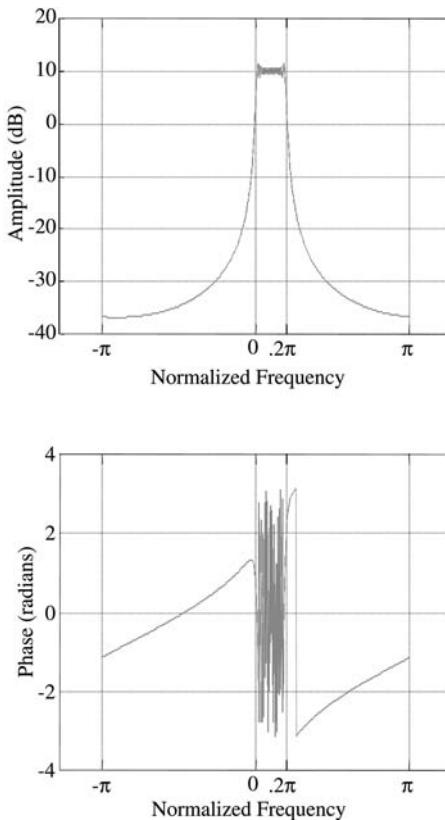
Time domain chirp over $[0.6\pi, 0.4\pi]$.

Gold code is a pseudorandom code that appears like white noise. The component at frequency 0 is very small because the Gold codes chosen for the GPS satellites are all balanced codes. An observation of the Gold code in the frequency domain is that its spectral lines differ greatly, as much as 20 to 30 dB.

Within the class of sinusoidal jammers, the jammers that lock into the strong Gold code spectral lines naturally are more harmful than those having frequency contents that coincide with weak spectral lines. For a normalized sinusoidal jammer vector \mathbf{u} ,

**FIGURE 2.11**

FFT of complex chirp over $[-0.6\pi, 0.4\pi]$.

**FIGURE 2.12**

Chirp signal over $[0, 0.2\pi]$.

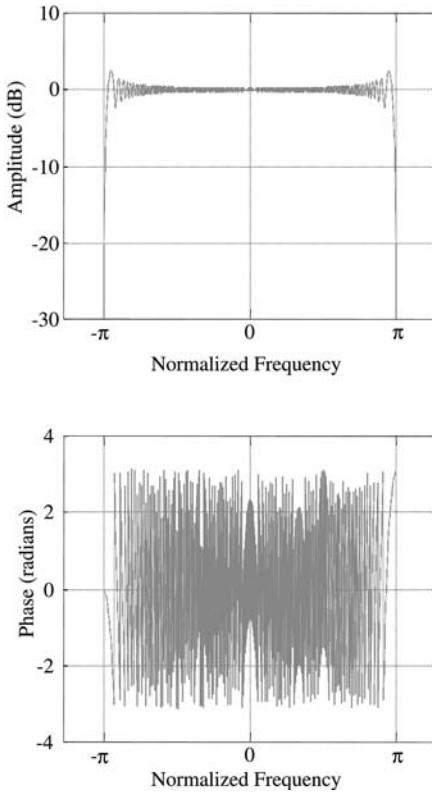
where $u(n) = \frac{1}{\sqrt{P}} \exp(jn\omega)$ for $n = 1, 2, \dots, P$, the cross-correlation coefficient is:

$$\beta = \mathbf{p}^T \mathbf{u} = \frac{\sum_{n=0}^{P-1} p(n) \exp(jn\omega)}{\sqrt{P}} \quad (2.31)$$

where the numerator is the discrete-time FT of the Gold code. Suppose $\omega = \frac{2\pi}{P}k$, then Equation (2.31) becomes:

$$\beta = \frac{X(k)}{\sqrt{P}} \quad (2.32)$$

where $X(k)$ is the k th spectral line of the Gold code. Therefore, the cross-correlation coefficient takes its maximum value when the k th spectral line of the Gold code is the strongest component. The cross-correlation coefficients $|\beta|$ corresponding to the strong spectral lines for each satellite can then be determined. It is found that the largest value of $|\beta|$ is 0.09

**FIGURE 2.13**

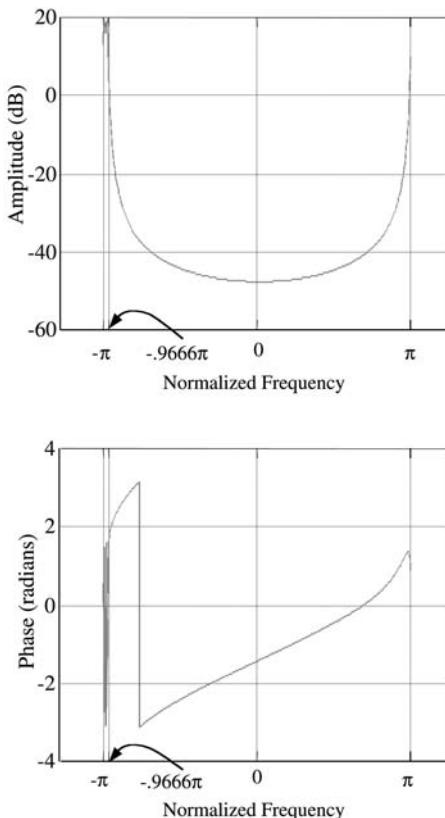
Chirp signal over $[-\pi, \pi]$.

From the SINR expression in Equation (2.18), it is straightforward to show that the receiver SINR for the sinusoidal case is given by:

$$\begin{aligned} SINR &= \frac{2P}{\sigma^2} \left(Q - \sum_{m=0}^{Q-1} |\beta_m|^2 \right) \\ &= \frac{2P}{\sigma^2} \left(Q - \frac{1}{P} \sum_{m=0}^{Q-1} |X(k_m)|^2 \right) \end{aligned} \quad (2.33)$$

Sinusoidal jammers are only special cases of chirp jammers. For chirp jammers, the simple relation guiding the range of the cross-correlation coefficient given by Equation (2.32) is no longer applicable. Thus, extensive numerical search methods must be carried out in case of chirp jammers to determine the range and properties of the cross-correlation coefficient. However, it must be clarified that this search is strictly for analysis purposes only and not part of the real-time receiver operation.

The jammer considered is of length $N = P = 1023$, and it is characterized by its starting frequency f_s and its ending frequency f_e . The search is carried out by setting

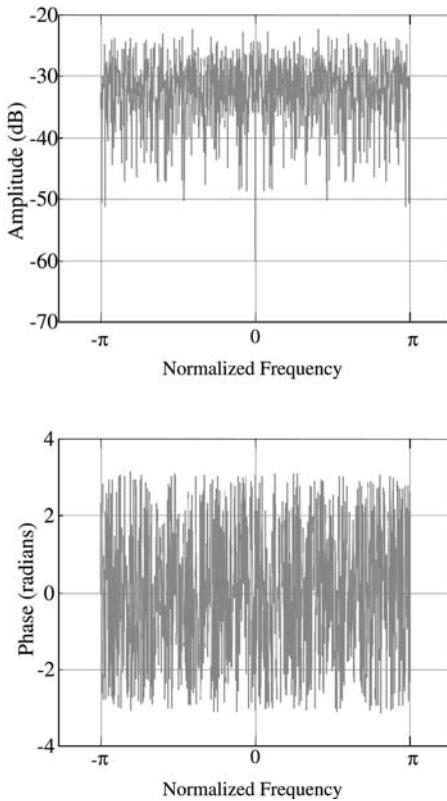
**FIGURE 2.14**

FFT of complex chirp over $[-\pi, 0.9666\pi]$.

f_s from $-\pi$ to π with a search step of 0.001π , and the procedure is followed for f_e . Thus, the search of largest $|\beta|$ and smallest $|\beta|$ covers all possible starting and ending frequencies with an incremental step of 0.001π . The search is performed for every satellite signal. It is noted that $|\beta|$ takes the same value for both jammers with start and end frequencies (f_s, f_e) and $(-f_s, -f_e)$.

The largest value of $|\beta|$ obtained from the search is 0.1446 for satellite 22 corresponding to normalized jammer frequencies $(f_s, f_e) = (-0.051, 0.464)$. The minimum value of $|\beta|$ obtained is 0.000008 for satellite 12 corresponding to the jammer frequencies $(f_s, f_e) = (0.021, -0.459)$. By comparing with the values of $|\beta|$ for the sinusoidal jammers, it is clear that chirp jammers with specific frequency ranges can yield higher values of $|\beta|$ (i.e., cause more signal distortion).

If the chirp jammer frequencies (f_s, f_e) are assumed to be uniformly distributed over the searched numerical values, then the expected value of $|\beta|$ and $|\beta|^2$ can be easily obtained. These values are $E[|\beta|] = 0.028$ and $E[|\beta|^2] = 0.001$, respectively.

**FIGURE 2.15**

Gold code in frequency domain.

The expected value of Equation (2.33) yields the average receiver SINR for all chirp jammers:

$$\begin{aligned}
 E[SINR] &= E\left[2P\left(Q - \sum_{m=0}^{Q-1} |\beta_m|^2\right)\right] / \sigma^2 \\
 &= 2P\left(Q - E\left[\sum_{m=0}^{Q-1} |\beta_m|^2\right]\right) / \sigma^2 \\
 &= 2PQ\left(1 - E\left[|\beta_m|^2\right]\right) / \sigma^2
 \end{aligned} \tag{2.34}$$

Here the SINR lower bound can be obtained by replacing all $|\beta_m|$ in Equation (2.34) with its maximum possible value obtained from the search results. Compared with the upper bound $2PNQ/\sigma^2$, it is noted that the worst jamming case only degrades the receiver SINR by $-10 \log(1 - |\beta|_{max}^2) = 0.09$ dB.

2.4.5 In the presence of several satellite signals

When the received signal is assumed to include the signals from other satellites in view, Equation (2.4) can be rewritten as:

$$\mathbf{x} = \sum_{m=1}^M A_m \mathbf{p}_m + \mathbf{w} + \mathbf{j} \quad (2.35)$$

where the first term incorporates the spreading codes of all satellites in view, including the signal of interest, taken to be \mathbf{p}_1 for convenience. A_m is the amplitude of the m th satellite signal ($A_1 = 1$), and M is the number of satellites in sight, which is typically around 10. The spreading codes from different satellites, \mathbf{p}_2 through \mathbf{p}_M are assumed to be unsynchronized and can therefore be considered as the other Gold codes with time shifts. No Doppler effect is considered. Similar to Equation (2.12), the correlation value after projection and despread can be written in terms of the constituent signals as:

$$\begin{aligned} y &= \sum_{k=0}^{Q-1} \mathbf{p}_1^T \mathbf{V}_k \mathbf{p}_1 + \sum_{k=0}^{Q-1} \mathbf{w}_k^T \mathbf{V}_k \mathbf{p}_1 + \sum_{m=2}^M \sum_{k=0}^{Q-1} A_m \mathbf{p}_m^T \mathbf{V}_k \mathbf{p}_1 \\ &\triangleq y_1 + y_2 + y_3 \end{aligned} \quad (2.36)$$

where y_3 is the contribution of the other satellite signals to the correlator output. Due to the fact that \mathbf{p}_1 is a fixed code, and used over each navigation data symbol of the respective signal, y_1 takes a deterministic value. The value of y_3 is given by:

$$\begin{aligned} y_3 &= \sum_{m=2}^M \sum_{k=0}^{Q-1} A_m \mathbf{p}_m^T (I - \mathbf{u}_k \mathbf{u}_k^H) \mathbf{p}_1 \\ &= \sum_{m=2}^M A_m \left(\sum_{k=0}^{Q-1} \mathbf{p}_m^T \mathbf{p}_1 - \sum_{k=0}^{Q-1} \mathbf{p}_m^T \mathbf{u}_k \mathbf{u}_k^H \mathbf{p}_1 \right) \end{aligned} \quad (2.37)$$

Let:

$$\alpha_{mn} = \frac{\mathbf{p}_m^T \mathbf{p}_n}{P} \quad (2.38)$$

denote the cross-correlation coefficient between spreading code \mathbf{p}_m and \mathbf{p}_n , and:

$$\beta_{mk} = \frac{\mathbf{p}_m^T \mathbf{u}_k}{\sqrt{P}} \quad (2.39)$$

represent the cross-correlation coefficient between the m th satellite signal and the jammer vector over the k th block. Thus:

$$\begin{aligned} y_3 &= \sum_{m=2}^M A_m \left(Q P \alpha_{m1} - P \sum_{k=0}^{Q-1} \beta_{mk} \beta_{1k}^* \right) \\ &= P \left(Q \sum_{m=2}^M A_m \alpha_{m1} - \sum_{m=2}^M A_m \sum_{k=0}^{Q-1} \beta_{mk} \beta_{1k}^* \right) \end{aligned} \quad (2.40)$$

$$\text{Re}[y_3] = PQ \sum_{m=2}^M A_m \alpha_{m1} - P \sum_{m=2}^M A_m \text{Re} \left[\sum_{k=0}^{Q-1} \beta_{mk} \beta_{1k}^* \right] \quad (2.41)$$

Due to the property of the Gold code, the cross-correlation coefficient $\alpha_{m1}, \forall m$, takes on three different values with the following probability [18]:

$$\alpha_{m1} = \begin{cases} -\frac{65}{1023} & .125 \\ -\frac{1}{1023} & .75 \\ \frac{63}{1023} & .125 \end{cases} \quad (2.42)$$

The mean value of the preceding random variable α is $-1/1023$. Let $\eta_m = \text{Re} \left[\sum_{k=0}^{Q-1} \beta_{mk} \beta_{1k}^* \right]$. Then Equation (2.41) is rewritten:

$$\text{Re}[y_3] = PQ \sum_{m=2}^M A_m \alpha_{m1} - P \sum_{m=2}^M A_m \eta_m \quad (2.43)$$

$$\begin{aligned} E[\text{Re}^2[y_3]] &= E \left[P^2 Q^2 \left(\sum_{m=2}^M A_m \alpha_{m1} \right)^2 - 2P^2 Q \sum_{m=2}^M A_m \alpha_{m1} \sum_{m=2}^M A_m \eta_m \right. \\ &\quad \left. + P^2 \left(\sum_{m=2}^M A_m \eta_m \right)^2 \right] \end{aligned} \quad (2.44)$$

The first term of Equation (2.44) is:

$$\begin{aligned} E \left[P^2 Q^2 \sum_{m=2}^M \sum_{n=2}^M A_m A_n \alpha_{m1} \alpha_{n1} \right] &= P^2 Q^2 \sum_{m=2}^M \sum_{n=2}^M A_m A_n E[\alpha_{m1} \alpha_{n1}] \\ &= P^2 Q^2 \left(\sum_{m=n=2}^M A_m^2 E[\alpha_{m1}^2] + \sum_{m=2}^M \sum_{\substack{n=2 \\ n \neq m}}^M A_m A_n E[\alpha_{m1}] E[\alpha_{n1}] \right) \\ &= P^2 Q^2 \left(\sum_{m=n=2}^M A_m^2 E[\alpha^2] + E^2[\alpha_{m1}] \sum_{m=2}^M \sum_{\substack{n=2 \\ n \neq m}}^M A_m A_n \right) \\ &= PQ^2 \sum_{m=2}^M A_m^2 + Q^2 \sum_{m=2}^M \sum_{\substack{n=2 \\ n \neq m}}^M A_m A_n \end{aligned} \quad (2.45)$$

which makes use of the knowledge that:

$$E[\alpha^2] = \frac{1025}{1023^2} \approx \frac{1}{P} \quad (2.46)$$

$$E[\alpha] = \frac{1}{1023} = -\frac{1}{P} \quad (2.47)$$

To calculate the second term of Equation (2.44), it is important to obtain $E[\eta_m]$, which can be approximated by zero, because:

$$E[\beta_{mk}] = E\left[\frac{\mathbf{p}_m^T \mathbf{u}_k}{\sqrt{P}}\right] = \frac{E[\mathbf{p}_m^T] \mathbf{u}_k}{\sqrt{P}} = \frac{-\frac{1}{P}[1, 1, 1, \dots, 1] \mathbf{u}_k}{\sqrt{P}} \approx 0 \quad (2.48)$$

$$\begin{aligned} E[\eta_m] &= E\left[\operatorname{Re}\left[\sum_{k=0}^{Q-1} \beta_{mk} \beta_{1k}^*\right]\right] = \operatorname{Re}\left[E\left[\sum_{k=0}^{Q-1} \beta_{mk} \beta_{1k}^*\right]\right] \\ &= \operatorname{Re}\left[\sum_{k=0}^{Q-1} E[\beta_{mk}] \beta_{1k}^*\right] \approx 0 \end{aligned} \quad (2.49)$$

As such, the second term of Equation (2.44) can be given by:

$$-E\left[2P^2Q \sum_{m=2}^M A_m \alpha_{m1} \sum_{m=2}^M A_m \eta_m\right] = -2PQ \sum_{m=2}^M A_m \sum_{m=2}^M A_m E[\eta_m] \approx 0 \quad (2.50)$$

The third term of (2.44) can be obtained as:

$$\begin{aligned} E\left[P^2 \left(\sum_{m=2}^M A_m \eta_m\right)^2\right] &= P^2 E\left[\sum_{m=2}^M \sum_{n=2}^M A_m \eta_m A_n \eta_n\right] \\ &= P^2 \sum_{m=2}^M \sum_{n=2}^M A_m A_n E[\eta_m \eta_n] \\ &= P^2 \sum_{m=2}^M A_m^2 E[\eta_m^2] + \sum_{m=2}^M \sum_{\substack{n=2, \\ m \neq n}}^M A_m A_n E[\eta_m \eta_n] \\ &= P^2 \sum_{m=2}^M A_m^2 E[\eta_m^2] + \sum_{m=2}^M \sum_{\substack{n=2, \\ m \neq n}}^M A_m A_n E[\eta_m] E[\eta_n] \\ &= P^2 \sum_{m=2}^M A_m^2 E[\eta_m^2] \end{aligned} \quad (2.51)$$

Because η_m is in the same order of the square of β_{mk} , and $|\beta_{mk}| < 1$, the value of η_m^2 must be smaller than η_m . Thus, from Equation (2.49), $E[\eta_m^2]$ can also be approximated by zero. Therefore, the expression in Equation (2.51) has a negligible value, and Equation (2.44) can be approximated by its first term:

$$E[\operatorname{Re}^2[y_3]] \approx E\left[P^2 Q^2 \left(\sum_{m=2}^M A_m \alpha_{m1}\right)^2\right]$$

$$= PQ^2 \sum_{m=2}^M A_m^2 + Q^2 \sum_{m=2}^M \sum_{\substack{n=2 \\ n \neq m}}^M A_m A_n \quad (2.52)$$

The receiver SINR performance becomes:

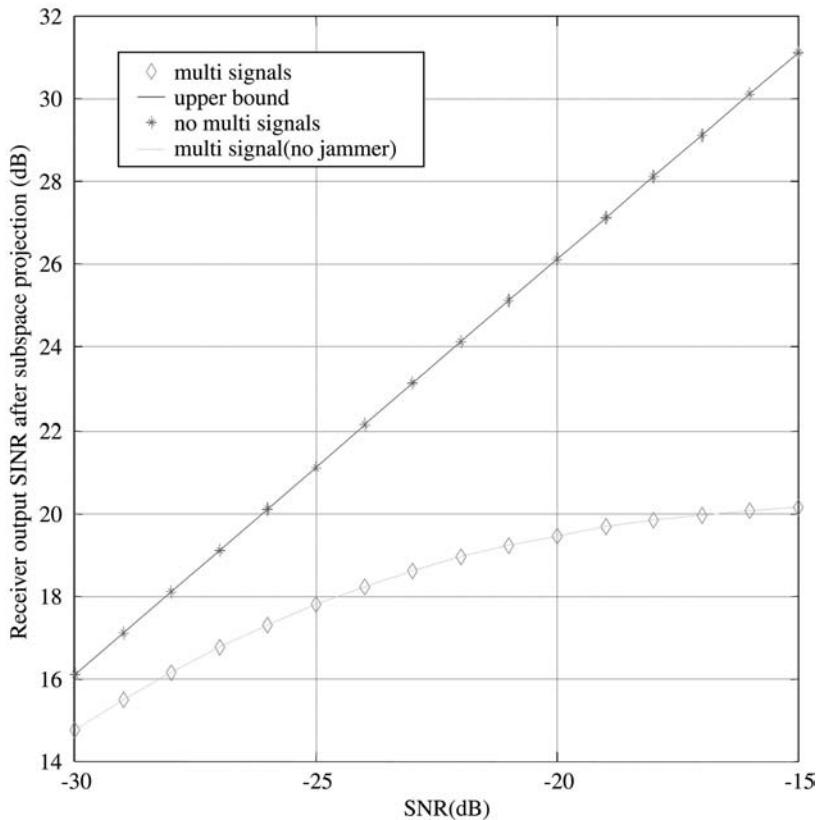
$$\begin{aligned} SINR &= \frac{y_1^2}{\frac{1}{2}\sigma_{y_2}^2 + E[\text{Re}^2[y_3]]} \\ &\approx \frac{y_1^2}{\frac{1}{2}\sigma_{y_2}^2 + PQ^2 \sum_{m=2}^M A_m^2 + Q^2 \sum_{m=2}^M \sum_{\substack{n=2 \\ n \neq m}}^M A_m A_n} \\ &= \frac{P^2 \left(1 - \frac{1}{Q} \sum_{k=0}^{Q-1} |\beta_{1k}|^2 \right)^2}{\frac{1}{2}P \left(1 - \frac{1}{Q} \sum_{k=0}^{Q-1} |\beta_{1k}|^2 \right) + PQ^2 \sum_{m=2}^M A_m^2 + Q^2 \sum_{m=2}^M \sum_{\substack{n=2 \\ n \neq m}}^M A_m A_n} \end{aligned} \quad (2.53)$$

The difference between Equation (2.53) and the *SINR* expression in Equation (2.28) comes from the multiple access interference when the signals of the other satellites are taken into consideration. As discussed earlier, the values of cross-correlation coefficient are typically very small, and their effect in the reduction of the receiver performance is negligible compared with the multiple access interference. [Figure 2.16](#) depicts the *SINR* expression in Equations (2.53) and (2.28) as a function of the input SNR. In the simulation presented, the number of satellites in sight is assumed to be $M = 10$, and the amplitude A_m of all satellite signals is set to 1. This figure suggests that the degradation caused by the other satellite signals is more severe than the loss due to the projection operation. [Figure 2.17](#) depicts Equation (2.53) as a function of the number of satellites in sight. The input SNR is set at -20 dB. This figure demonstrates the effect of the multiple access interference that increases directly with the number of satellites in sight.

2.4.6 Subspace array processing

Compared with the subspace projection techniques in the single-sensor case, the use of multisensor arrays greatly increases the dimension of the available signal subspace. This use allows both the distinctions in the spatial and temporal signatures of the GPS signals from those of the interferers to play equal roles in suppressing the jammer with minimum distortion of the desired signal. In this section, the applicability of multisensor subspace projection techniques for suppressing nonstationary jammers in GPS receivers is examined. It is noted that notations in this section are slightly different from the previous discussions due to more factors considered.

In GPS, the PN sequence of length $P(1023)$ repeats itself $Q(20)$ times within 1 symbol of the 50 bps navigation data. Discrete-time form is used, where all the signals

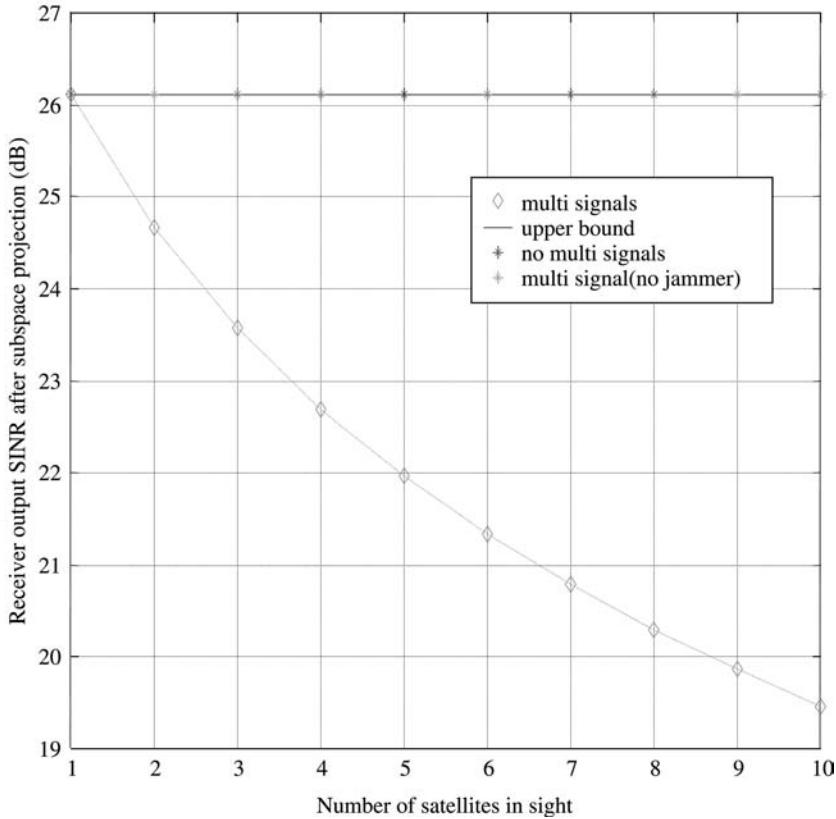
**FIGURE 2.16**

Receiver output SINR vs. SNR in the presence of multiple satellite signals.

are sampled at the chip rate of the C/A code. The subject N sensor antenna array receives signals in a communication channel restricted to flat fading. In the proposed interference excision approach, the PNQ sensor output samples are partitioned into Q blocks, each of P chips and PN samples. The jammer can be consecutively removed from the 20 blocks that constitute 1 symbol. This is achieved by projecting the received data in each block on the corresponding orthogonal subspace of the jammer. The jammer-free signal is then correlated with the replica PN sequence on a symbol-by-symbol basis. First, consider the subspace projection within each block. The array output vector at the k th sample is given by:

$$\begin{aligned}\mathbf{x}(k) &= \mathbf{x}_s(k) + \mathbf{x}_u(k) + \mathbf{w}(k) \\ &= c(k)\mathbf{h} + \sum_{i=1}^U A_i u_i(k) \mathbf{a}_i + \mathbf{w}(k)\end{aligned}\quad (2.54)$$

where \mathbf{x}_s , \mathbf{x}_u and \mathbf{w} are the signal, the jammer and the white Gaussian noise contributions, respectively; \mathbf{h} is the signal spatial signature; and $c(k)$ is the spreading PN

**FIGURE 2.17**

Receiver SINR performance vs. the number of signals in sight.

sequence. All U jammers are considered as normalized instantaneously narrowband FM signals with constant amplitude, so that $\frac{u_i(k)=1}{\sqrt{P}\exp[j\varphi_i(k)]}$ is the i th jammer frequency signature. A_i and \mathbf{a}_i are the amplitude and spatial signature, respectively. Furthermore, the channels are normalized and $\|\mathbf{h}\|_F^2 = N$ and $\|\mathbf{a}\|_F^2 = N$, where $\|\cdot\|_F^2$ is the Frobenius norm of a vector.

The noise vector $\mathbf{w}(k)$ is zero-mean, temporally and spatially white:

$$\begin{aligned} E[\mathbf{w}(k)\mathbf{w}^H(k+l)] &= 0 \\ E[E[\mathbf{w}(k)\mathbf{w}^H(k+l)]] &= \sigma^2 \delta(l) \mathbf{I}_N \end{aligned} \quad (2.55)$$

where σ^2 is the noise power, and \mathbf{I}_N is the $N \times N$ identity matrix. Using P sequential array vector samples within the block yields the following $PN \times 1$ vector:

$$\begin{aligned} \mathbf{X} &= [\mathbf{x}^T(1) \ \mathbf{x}^T(2) \ \dots \ \mathbf{x}^T(L)]^T \\ &= \mathbf{X}_s + \mathbf{X}_u + \mathbf{W} \end{aligned} \quad (2.56)$$

The vector \mathbf{X}_u consists of the U jammer signals, and is expressed as:

$$\mathbf{X}_u = \sum_{i=1}^U A_i \mathbf{U}_i \quad (2.57)$$

with:

$$\begin{aligned} \mathbf{U}_i &= [u_i(1) \ u_i(2) \ \dots \ u_i(P)]^T \otimes \mathbf{a}_i \\ &= \mathbf{u} \otimes \mathbf{a}_i \end{aligned} \quad (2.58)$$

where \otimes denotes the Kronecker product. Therefore:

$$\mathbf{U} = [\mathbf{U}_1 \ \mathbf{U}_2 \ \dots \ \mathbf{U}_U] \quad (2.59)$$

spans the jammer subspace. Its orthogonal subspace projection matrix is given by:

$$\mathbf{V} = \mathbf{I}_{PN} - \mathbf{U}(\mathbf{U}^H \mathbf{U})^{-1} \mathbf{U}^H = \mathbf{I}_{PN} - \mathbf{U} \mathbf{U}^H \quad (2.60)$$

Application of this projection to the received signal yields:

$$\mathbf{V}\mathbf{X} \triangleq \mathbf{X}_\perp = \mathbf{V}\mathbf{X}_s + \mathbf{V}\mathbf{W} \quad (2.61)$$

with the jammer component removed entirely. The signal vector \mathbf{X}_s can be rewritten as:

$$\mathbf{X}_s = [c(1) \ c(2) \ \dots \ c(P)]^T \otimes \mathbf{h} \triangleq \mathbf{q} \quad (2.62)$$

where the vector \mathbf{q} represents the spatial–temporal signature of the GPS signal. The result of despreading in the subspace-projection-based array system over one block is:

$$y = \mathbf{q}^H \mathbf{X}_\perp = \mathbf{q}^H \mathbf{V}\mathbf{q} + \mathbf{q}^H \mathbf{V}\mathbf{W} \triangleq y_1 + y_2 \quad (2.63)$$

where y_1 and y_2 are the contributions of the PN and the noise sequences to the decision variable, respectively. For simplification, it is assumed that the jammers share the same period as the GPS data symbol. In GPS systems, due to the fact that each satellite is assigned a fixed Gold code, and that the Gold code is the same for every navigation data symbol, y_1 is a deterministic value, as opposed to the analysis in many spread-spectrum applications where it is treated as a random variable. The value of y_1 is then given by:

$$\begin{aligned} y_1 &= \mathbf{q}^H \mathbf{V}\mathbf{q} = \mathbf{q}^H (\mathbf{I}_{LN} - \mathbf{U} \mathbf{U}^H) \mathbf{q} \\ &= \mathbf{q}^H \mathbf{q} - \mathbf{q}^H [\mathbf{U}_1 \ \mathbf{U}_2 \ \dots \ \mathbf{U}_U] \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \vdots \\ \mathbf{U}_U \end{bmatrix} \mathbf{q} \\ &= PN - \mathbf{q}^H \left(\sum_{i=1}^U \mathbf{U}_i \mathbf{U}_i^H \right) \mathbf{q} \\ &= PN - \sum_{i=1}^U \mathbf{q}^H \mathbf{U}_i \mathbf{U}_i^H \mathbf{q} \end{aligned} \quad (2.64)$$

where:

$$\mathbf{q}^H \mathbf{U}_i = (\mathbf{c} \otimes \mathbf{h})^H (\mathbf{u}_i \otimes \mathbf{a}_i) = (\mathbf{c}^H \mathbf{u}_i)(\mathbf{h}^H \mathbf{a}_i) \quad (2.65)$$

Define:

$$\alpha_i = \frac{\mathbf{h}^H \mathbf{a}_i}{N} \quad (2.66)$$

as the spatial cross-correlation coefficient between the signal and the i th jammer, and:

$$\beta_i = \frac{\mathbf{c}^T \mathbf{u}_i}{\sqrt{P}} \quad (2.67)$$

as the temporal cross-correlation coefficient between the PN sequence and the i th jammer vector. Therefore:

$$\mathbf{q}^H \mathbf{U}_i = \sqrt{P} N \alpha_i \beta_i \quad (2.68)$$

$$y_1 = PN - PN \sum_{i=1}^U |\alpha_i|^2 |\beta_i|^2 = PN \left(1 - \sum_{i=1}^U |\alpha_i|^2 |\beta_i|^2 \right) \quad (2.69)$$

From Equation (2.69), it is clear that y_1 is real, which is due to the Hermitian property of the projection matrix \mathbf{V} . With the assumptions in Equation (2.55), y_2 is complex Gaussian with zero mean. Therefore:

$$E[y] = y_1 = PN \left(1 - \sum_{i=1}^U |\alpha_i|^2 |\beta_i|^2 \right) \quad (2.70)$$

$$\begin{aligned} \text{Var}[y] &= \text{Var}[y_2] = E[|y_2|^2] \\ &= E[\mathbf{q}^H \mathbf{V} \mathbf{W} \mathbf{W}^H \mathbf{V}^H \mathbf{q}] = \mathbf{q}^H \mathbf{V} E[\mathbf{W} \mathbf{W}^H] \mathbf{V} \mathbf{q} \\ &= \sigma^2 \mathbf{q}^H \mathbf{V} \mathbf{V} \mathbf{q} = \sigma^2 \mathbf{q}^H \mathbf{V} \mathbf{q} \\ &= \sigma^2 y_1 = \sigma^2 PN \left(1 - \sum_{i=1}^U |\alpha_i|^2 |\beta_i|^2 \right) \end{aligned} \quad (2.71)$$

Equations (2.70) and (2.71) are derived for only one block of the signal symbol. In the following the subscript m is added to identify y with block m ($m = 1, 2, \dots, Q$). Note that the meaning of these subscripts is different from that in the discussion of general periodic jammers. The output of the symbol-level despread is obtained by summing all Q blocks, as:

$$y = \sum_{m=1}^Q y_m \quad (2.72)$$

Because y_m are Gaussian random variables with zero mean, y is also a zero-mean Gaussian random variable. The decision variable y_r is the real part of y :

$$y_r = \text{Re}[y] \quad (2.73)$$

The expected value of y_r is:

$$E[y_r] = E[y] = \sum_{m=1}^Q y_m$$

$$\begin{aligned}
&= PN \left(Q - \sum_{m=1}^Q \sum_{i=1}^U |\alpha_{mi}|^2 |\beta_{mi}|^2 \right) \\
&= PN \left[Q - \sum_{i=1}^U \left(|\alpha_i|^2 \sum_{m=1}^Q |\beta_{mi}|^2 \right) \right]
\end{aligned} \tag{2.74}$$

where α_{mi} and β_{mi} are the spatial and temporal cross-correlation coefficients between the signal and the i th jammer over block m . Because the changes in the spatial signatures of the signals and jammers are very small compared with the period of one block (1 msec), α_{mi} can be simplified to α_i . The variance of y_r is:

$$\begin{aligned}
\sigma_{y_r}^2 &= \frac{1}{2} \sigma_y^2 = \frac{1}{2} \sum_{m=1}^Q \sigma_{y_m}^2 \\
&= \frac{1}{2} \sigma^2 PN \left[Q - \sum_{i=1}^U \left(|\alpha_i|^2 \sum_{m=1}^Q |\beta_{mi}|^2 \right) \right]
\end{aligned} \tag{2.75}$$

Therefore, the receiver SINR expression after projection and despreading is given by:

$$\begin{aligned}
SINR &= \frac{E^2[y]}{\text{Var}[y]} \\
&= \frac{2PN \left[Q - \sum_{i=1}^U \left(|\alpha_i|^2 \sum_{m=1}^Q |\beta_{mi}|^2 \right) \right]}{\sigma^2}
\end{aligned} \tag{2.76}$$

The temporal and spatial coefficients appear as multiplicative products in Equation (2.76). This implies that the spatial and temporal signatures play equal roles in the receiver performance. In the absence of the jammers, no excision is necessary, and the SINR of the receiver output becomes $2PNQ/\sigma^2$, which represents the upper bound of the interference suppression performance. Clearly, the term:

$$2PN \sum_{i=1}^U \left(|\alpha_i|^2 \sum_{m=1}^Q |\beta_{mi}|^2 \right) \tag{2.77}$$

in Equation (2.76) is the reduction in the receiver performance caused by the proposed interference suppression technique. It reflects the energy of the signal component that is in the jammer subspace. It is important to note that if the jammers and the DSSS signal are orthogonal, either in spatial domain ($\alpha_i = 0$) or in temporal domain ($\beta_{mi} = 0$), the interference excision is achieved with no loss in performance. In the general case, β_{mi} are very small and much smaller than α_i . Therefore, the difference in the temporal signatures of the incoming signals allow the proposed projection technique to excise FM jammers effectively with only insignificant signal loss. The spatial cross-correlation coefficients α_i are fractional values and, as such, further reduce the undesired term in Equation (2.77).

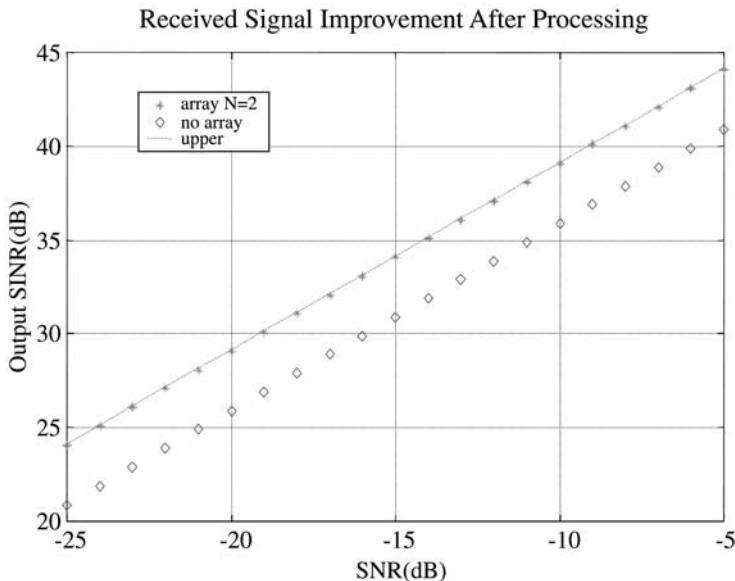


FIGURE 2.18
Output SINR vs. SNR.

Using arrays improves these interference suppression algorithms in several ways. First, the employment of an antenna array can lead to more accurate IF estimation of the jammers [16]. Second, in comparison with the result of the single-sensor case:

$$SINR = \frac{2P \left(Q - \sum_{i=1}^U \sum_{m=1}^Q |\beta_{mi}|^2 \right)}{\sigma^2} \quad (2.78)$$

multisensor receivers, at minimum, improve $SINR$ by the array gain. This is true, independent of the underlying fading channels and scattering environment. As discussed earlier, compared with the temporal signatures, the contributions of spatial cross-correlation coefficients α_i in influencing the receiver SINR are, in most cases, insignificant.

A snapshot of simulation results, encapsulated in Figure 2.18, is presented illustrating the properties of the performance of the proposed projection technique. Two uncorrelated chirp jammers are considered. The angle of arrival (AOA) of the satellite signal and the jammers are 5° , 40° and 60° , respectively. A two-element array is considered with half-wavelength spacing. The satellite PN sequence is the Gold code of satellite SV 1, and the normalized frequency of the jammers are from 0.01 to 0.2 and from 0.5 to 0.3, respectively. The $SINR$ of the single sensor case is also plotted for comparison. The array gain is evident in Figure 2.18, plotting output SINR vs. input SNR. In this example, $|\alpha_1| = 0.643$, $|\alpha_2| = 0.340$, $|\beta_1|$ is in the range from 0.0049 to 0.0604, and $|\beta_2|$ is in the range from 0.0067 to 0.0839. With the preceding values, the term in Equation (2.77) is far less than $2PNQ$, which allows $SINR$ in Equation (2.76) to be very close to its upper bound.

2.5 Conclusions

Bilinear TFDs for signal power localization in the TF plane have motivated several new effective approaches for interference excisions, including the new and effective projection techniques for the interference excisions in DSSS systems. In the proposed technique, the IF of the jammer is estimated and used to construct the jammer subspace, from which interference excision is achieved by projecting the received data onto the subspace orthogonal to the jammer.

The growing importance of GPS demands adequate interference excision and antijamming capabilities. Because GPS is a CDMA system that employs BPSK-modulated DSSS signals, the subspace projection technique is a potential candidate for GPS antijamming problems. However, as is always the case when transitioning a general theory to an application with specific parameters, certain exploitations are available—in this case the periodicity of the C/A code requires special treatment in the TF domain processing, but also affords some nice simplification in the analysis. In this chapter, techniques based on subspace projections are applied to GPS antijamming, with both single and multiple receivers. The performances of these techniques, as measured by received SINR, are derived, showing the effective interference excision results.

Acknowledgment

This work has been supported by the Air Force Research Laboratory, grant no. F30602-00-1-0515.

References

- [1] R. Abimoussa and R. Landry, Jr., Anti-Jamming Solution to Narrowband CDMA Interference Problem, 2000 Canadian Conference on Electrical and Computer Engineering, Halifax, NS, Canada, March 2000.
- [2] M.G. Amin, Interference mitigation in spread-spectrum communication systems using time-frequency distributions, *IEEE Trans. Signal Process.*, 45(1), 90–102, January 1997.
- [3] M.G. Amin and A. Akansu, Time-frequency for interference excision in spread-spectrum communications, Highlights of Signal Processing for Communications, Celebrating a Half Century of Signal Processing, G. Giannakis, Ed., *IEEE SP Mag.*, 16(2), March 1999.

- [4] M.G. Amin and A.R. Lindsey, Time-Frequency for Interference Mitigation in Spread Spectrum Communication Systems, in *Time-Frequency Applications*, Prentice Hall, New York, 2001.
- [5] M.G. Amin, R.S. Ramineni, and A.R. Lindsey, Subspace-Based Interference Excision for Spread Spectrum Communication, in Proceedings of the ICASSP 2000, Istanbul, Turkey, June 5–9, 2000.
- [6] M. Amin, C. Wang, and A. Lindsey, Optimum interference excision in spread spectrum communications using open loop adaptive filters, *IEEE Trans. Signal Process.*, 47(7), 1966–1976, July 1999.
- [7] B. Boashash, Estimating and interpreting the instantaneous frequency of a signal, Parts 1 and 2, in Proceedings of the IEEE, 80(12), December 1990.
- [8] M.S. Braasch and A.J. Van Dierendonck, GPS receiver architectures and measurements, in Proceedings of the IEEE, January 1999.
- [9] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.
- [10] P. Enge and P. Misra, Scanning the issue/technology, in Proceedings of the IEEE, January 1999.
- [11] E.D. Kaplan, Ed., *Understanding GPS: Principles and Applications*, Artech House Publishers, 1996.
- [12] S. Kay, A fast and accurate single frequency estimator, *IEEE Trans. Acoust., Speech, Signal Process.*, 37(12), December 1979.
- [13] S. Lach, M.G. Amin, and A.R. Lindsey, Broadband nonstationary interference excision in spread spectrum communication systems using time-frequency synthesis techniques, *IEEE J. Selected Areas Commun.*, April 1999.
- [14] R. Landry, Jr., V. Calmettes, and M. Bousquet, Impact of Interference on a Generic GPS Receiver and Assessment of Mitigation Techniques, in Proceedings of the ISSSTA'98 International Symposium on Spread Spectrum Techniques and Applications, Sun City, South Africa, September 1998.
- [15] R. Landry, Jr., P. Mouyon, and D. Lekaim, Interference mitigation in spread spectrum systems by wavelet coefficients thresholding, *Eur. Trans. Telecommunications*, 9, 191–202, March–April 1998.
- [16] W. Mu, M.G. Amin, and Y. Zhang, Bilinear signal synthesis in array processing, *IEEE Trans. Signal Process.* [submitted].
- [17] W. Myrick, M.D. Zoltowski, and J.S. Goldstein, Exploiting Conjugate Symmetry in Power-Minimization-Based Pre-Processing for GPS: Reduced Complexity and Smoothness, in Proceedings of the 2000 IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 5, June 2000, p. 2833–2836.
- [18] B.W. Parkinson and J.J. Spilker, Jr., Eds., *Global Positioning System: Theory and Applications*, American Institute of Aeronautics and Astronautics, 1996.

- [19] A. Pinker, D. Walker, and C. Smith, Jamming the GPS Signal, in Proceedings of the 55th Annual Meeting of the Institute of Navigation, Cambridge, MA, June 28–30, 1999, pp. 829–837.
- [20] M.W. Rosen and M.S. Braasch, Low-Cost GPS Interference Mitigation Using Single Aperture Cancellation Techniques, in Proceedings of the Institute of Navigation National Technical Meeting, Long Beach, CA, January 1998.
- [21] D.M. Upton, T.N Upadhyay, and J. Marchese, Commercial-off-the-Shelf (COTS) GPS Interference Canceller and Test Results, in Proceedings of the Institute of Navigation National Technical Meeting, Long Beach, CA, January 1998.
- [22] C. Wang and M.G. Amin, Performance analysis of instantaneous frequency based interference excision techniques in spread spectrum communications, *IEEE Trans. Signal Processing*, 46(1), 70–83, January 1998.
- [23] Y. Zhang and M.G. Amin, Blind beamforming for suppression of instantaneously narrowband signals in DS/SS communications using subspace projection techniques, in Proceedings of the SPIE: Digital Wireless Communication II, Vol. 4045, Orlando, FL, April 2000.
- [24] L. Zhao, M.G. Amin, and A.R. Lindsey, Subspace projection techniques for anti-FM jamming GPS receivers, in Proceedings of the 10th IEEE Workshop on Statistical Signal and Array Processing, August 2000, pp. 529–533.
- [25] L. Zhao, M.G. Amin, and A.R. Lindsey, Subspace Array Processing for the Suppression of FM Jamming in GPS Receivers, in Proceedings of the Asilomar Conference on Signals, Systems, and Computers, Monterey, CA, October 2000.

3

Positive Time–Frequency Distributions

Patrick Loughlin

University of Pittsburgh

Leon Cohen

City University of New York

CONTENTS

3.1	Introduction	121
3.2	Positive Distributions: Brief History, Formulation and Relation to Bilinear Distributions	123
3.3	Joint Densities and Conditional Densities	132
3.4	Joint vs. Conditional Densities, Time-varying Spectrum and Strong Support	135
3.5	Uncertainty Principle, Positive Distributions and Wigner Distribution	139
3.6	Construction of Positive Distributions	144
3.7	Collection of Examples and Plausibility of Results	147
3.8	Conclusions	155
	Acknowledgments	155
	References	155

3.1 Introduction

In 1932, Wigner introduced a remarkable viewpoint that has influenced many fields [101]. His aim was to produce a function, a joint density of time and frequency, which satisfies the marginals and would give insight as to the relationships between the time and frequency variables, and which could be used as a practical calculational tool.* In particular, Wigner sought a joint function of time and frequency, $P(t, \omega)$, that satisfies the marginal conditions:

$$\int P(t, \omega) d\omega = |s(t)|^2 \quad (3.1)$$

*Of course, Wigner was addressing the quantum case but the mathematics and conditions are the same.

$$\int P(t, \omega) dt = |S(\omega)|^2 \quad (3.2)$$

where $s(t)$ is the signal and $S(\omega)$ is its Fourier transform (FT):

$$s(t) = \frac{1}{\sqrt{2\pi}} \int S(\omega) e^{j\omega t} d\omega \quad (3.3)$$

$$S(\omega) = \frac{1}{\sqrt{2\pi}} \int s(t) e^{-j\omega t} dt \quad (3.4)$$

The marginals $|s(t)|^2$ and $|S(\omega)|^2$ are the energy temporal density (or instantaneous power) and the energy spectral density, respectively. Wigner gave what is now called the Wigner distribution:

$$W(t, \omega) = \frac{1}{2\pi} \int s^*(t - \frac{1}{2}\tau) s(t + \frac{1}{2}\tau) e^{-j\tau\omega} d\tau \quad (3.5)$$

Wigner pointed out that it was not manifestly positive and said, “It may be shown, however, that there does not exist any expression $[P(t, \omega)]$ which is bilinear in $[s(t)]$, satisfies [the marginals] and is everywhere (for all values $[t, \omega]$) positive” [101].[†] Wigner subsequently provided a proof of this statement several years later [102], where again he was very clear on the fact that he was referring to bilinear functions of the signal.

About 10 years after Wigner, the spectrographic analyzer was developed at Bell Labs [48]. This analog device gave a measure of the signal energy over time in particular frequency bands by bandpass filtering the signal, and then squaring the output. This device is mathematically equivalent to the squared magnitude of the short-time Fourier transform (STFT), called the spectrogram, which is closely related to the time–frequency (TF) expansion of a signal given by Gabor at about the same time [36]. The formulation for the spectrogram is:

$$P_{SP}(t, \omega) = \left| \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} s(\tau) h(\tau - t) d\tau \right|^2 \quad (3.6)$$

where $h(t)$ is the window function of the STFT, or in terms of the filter-bank interpretation, it is the impulse response of a low-pass filter that is then modulated to different bands by the complex exponential.

With the advent of digital computing and the fast Fourier transform (FFT), the spectrogram became the standard method for analyzing the time-varying frequency content of a signal. One of the basic reasons for the success of the spectrogram is because it is manifestly positive and hence can be readily interpreted and manipulated in the standard ways of ordinary densities.

Although the spectrogram has the desirable property that it is manifestly positive, it has limitations as a method for describing the time-varying spectral content of signals.

[†]Wigner, again, addressed wave functions instead of signals. We have substituted signal analysis phrases (in square brackets) for quantum phrases where appropriate.

Among them are:

1. It introduces a window into the analysis, and hence the results are window dependent. Often one cannot disentangle the effects of the window to obtain the inherent information and characteristics of the signal.
2. The spectrogram suffers from a TF resolution trade-off, meaning that if we want fine time resolution, we have to make the window narrow. Doing so, however, artificially modifies the signal into a short duration signal with a spectrum that may have little or no relation to the original signal.
3. Related to the previous point, as we vary the window from very narrow in time to very narrow in frequency (broad in time), we obtain very different representations of the same signal. Which one is, in some sense, the more accurate or “correct” representation of the inherent time-varying spectral structure of the signal?
4. The spectrogram does not satisfy the marginals, and hence we do not generally obtain correct results for expectation values (averages) such as the duration and bandwidth of the signal.

Because of these limitations, and for other reasons as well, there was great interest in developing alternatives to the spectrogram for TF analysis. Important developments include the Choi–Williams and Zhao–Atlas–Marks distributions [5, 105]. However, all of these distributions are bilinear and hence they are not manifestly positive, which introduces difficulties of its own.

There are positive distributions that satisfy the marginals, are easily constructed and in addition produce sensible results. Our aim in this chapter is threefold. First, we address some of the main issues concerning TF distributions and in particular positive distributions. Second, we discuss how the positive distributions are constructed. Third, by way of illustrative examples, we discuss the properties of these distributions. ‡

3.2 Positive Distributions: Brief History, Formulation and Relation to Bilinear Distributions

Henceforth when we say *positive distribution* we mean distributions that satisfy the marginals and are nonnegative. Accordingly, the spectrogram, which is nonnegative, is not a positive distribution, because it does not satisfy the marginals.

As mentioned earlier, Wigner pointed out that nonnegative bilinear distributions satisfying the marginals do not exist. Subsequently, however, most investigators

‡For more on the applications of positive distributions in signal analysis, we refer the interested reader to the accompanying chapter by Groutage et al., as well as to the several references listed at the end of this chapter.

just assumed that positive distributions simply do not exist; that is, they failed to appreciate that Wigner's statement applied only to bilinear distributions. Also, many investigators presented philosophical arguments for their nonexistence and often the uncertainty principle was invoked in a vague way. However, as early as 1966 one of the authors argued that it is clear that positive distributions *do* exist and gave one example [7]. In addition, it was pointed out that something strange is going on because if there was a philosophical or physical reason for their nonexistence, how is it possible that exceptions are present even with the Wigner distribution itself? (See Section 3.5.)

Subsequently, Cohen and Zaparovanny [26] and Cohen and Posch [24] gave an explicit formulation by which any number of positive distributions could be obtained. Indeed, they showed that it was easy to generate an infinite number of them for any signal, just as in the bilinear case. Of course, these distributions clearly are not bilinear in the signal, and in fact they are at least quadrilinear. Having a formulation that gives all possible positive distributions for a signal is both a curse and a blessing. The blessing is that it is likely that at least one of the distributions is reasonable; the challenge is figuring out which one. That can also be said of the bilinear distributions, and indeed it was not until many years after their formulation in 1966 that people began to understand the issues of kernel design and how that impacted on the bilinear distribution obtained and the reasonableness of the result. Likewise with the positive distributions: The formulation opened up the possibilities, and subsequently methods began to emerge for their practical construction. The first such method was developed by Loughlin, Pitton and Atlas and since then there has been considerable interest in the development and application of the positive distributions.

Positive distributions can be formulated in the following way:[§]

$$P(t, \omega) = |S(\omega)|^2 |s(t)|^2 \Omega(u(t), v(\omega)) \quad (3.7)$$

where Ω is a positive function called the kernel that characterizes the particular distribution. As with the bilinear distributions there are many possible choices for the kernel. It can be obtained in the following way. First choose any positive function of two variables, $\Omega(u, v)$, such that it satisfies:

$$\begin{aligned} \int_0^1 \Omega(u, v) dv &= 1 \\ \int_0^1 \Omega(u, v) du &= 1 \end{aligned} \quad (3.8)$$

It is easy to generate such functions [15, 24]. Then, for u and v substitute the functions of time and frequency $u(t)$ and $v(\omega)$ given by:

$$\begin{aligned} u(t) &= \int_{-\infty}^t |s(t')|^2 dt' \\ v(\omega) &= \int_{-\infty}^{\omega} |S(\omega')|^2 d\omega' \end{aligned} \quad (3.9)$$

[§]We assume that the energy of the signal is normalized to one. If not, the right-hand side of Equations (3.7) and (3.9) should be divided by the total energy, E .

That this formulation produces distributions that are nonnegative is obvious, and it is readily shown that it satisfies the marginals as well [24]. Other properties, such as shift invariance, scale invariance and finite support, may be found in [15, 24, 56].

One property that we highlight here, and return to again later, is that of the conditional moments, namely, the instantaneous frequency and group delay properties. If we write the signal and its spectrum as:

$$\begin{aligned} s(t) &= A(t)e^{j\varphi(t)} \\ S(\omega) &= B(\omega)e^{j\psi(\omega)} \end{aligned} \tag{3.10}$$

it is widely held that the first conditional moments in time and frequency of a TF distribution should equal the instantaneous frequency and group delay of the signal:

$$\begin{aligned} \langle \omega \rangle_t &= \int \omega P(\omega|t) d\omega = \varphi'(t) \\ \langle t \rangle_\omega &= \int t P(t|\omega) d\omega = -\psi'(\omega) \end{aligned} \tag{3.11}$$

where $\varphi'(t)$ is the instantaneous frequency of the signal, $-\psi'(\omega)$ is the group delay, and $P(\omega|t)$ and $P(t|\omega)$ are the conditional distributions defined in the standard way (see [Section 3.3.1](#)). The Wigner distribution always yields these results, but the positive distributions do not normally give this result.[¶] However, we point out that sometimes the positive distributions do give this result, and it is of interest to learn under what conditions that is the case. We return to these properties later, but note here that when the positive distributions do not yield a widely held property of (bilinear) TF distributions, one can take one of two views: that the positive distributions are not appropriate, or that perhaps the widely held properties need to be reconsidered.

3.2.1 Relation between the positive distribution formulation and the bilinear formulation

All TF distributions may be obtained from [7, 15]:

$$C(t, \omega) = \frac{1}{4\pi^2} \iiint s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) \phi(\theta, \tau) e^{-j\theta t - j\tau\omega + j\theta u} du d\tau d\theta \tag{3.12}$$

where $\phi(\theta, \tau)$ is a two-dimensional function called the kernel that characterizes a specific distribution. If the kernel is not a functional of the signal, then one obtains

[¶]This point was first made by Claasen and Mecklenbrauker [6], who showed that a distribution that is both positive and has finite support (see [Section 3.4.2](#)) cannot generally yield the instantaneous frequency and group delay of the signal for its first conditional moments. Accordingly, they argued against the requirement that the distribution is to be nonnegative.

the bilinear distributions and that is the most common way that Equation (3.12) is used. However, the kernel may be taken to be a functional of the signal. Indeed, if we want to obtain a positive distribution, then the kernel must be functionally dependent on the signal. If the kernel is not a functional of the signal, one may produce nonnegative distributions, for example, spectrograms, but they cannot satisfy the marginals.

Equation (3.12) may be used to study the positive distributions, and in fact it is useful to do so, for example, to show that product kernels are not necessary for scale invariance [56]. The relations between the kernel of the bilinear formulation and that of the positive formulation may be found in [24]. Although the positive distributions can be obtained from the bilinear formulation with signal-dependent kernels, the formulation given by Equation (3.7) is usually preferred, particularly for implementation, because it makes apparent and explicit the positive nature of the distributions, and because the signal dependence of kernels in the bilinear formulation can be difficult to work with.

It is nevertheless interesting to consider the type of kernels in the bilinear formulation that can produce a positive distribution. These kernels can be very different than the kernels we have become accustomed to with the bilinear distributions, in particular, with regard to the important developments by Williams concerning kernel design. For bilinear distributions, the kernel is typically taken to be a function that tapers to zero away from the axes so as to reduce the so-called cross terms, and it is one along the axes so that the correct marginals are obtained. However, kernels for the positive distributions, although they equal one along the axes as required for the marginals, do not necessarily taper to zero away from the axes and in fact they may diverge to infinity. We give a specific example of such a kernel, for a finite-duration tone:

$$s(t) = \frac{1}{\sqrt{T}} \operatorname{rect}\left(\frac{t}{T}\right) e^{j\omega_0 t} \quad (3.13)$$

where:

$$\operatorname{rect}\left(\frac{t}{T}\right) = \begin{cases} 1, & \text{if } -T/2 \leq t \leq T/2 \\ 0 & \text{otherwise} \end{cases} \quad (3.14)$$

For this signal, the kernel:

$$\phi(\theta, \tau) = \frac{\operatorname{sinc}(\theta T / 2\pi)}{\operatorname{sinc}(\theta(T - |\tau|) / 2\pi)} \quad (3.15)$$

yields, on insertion into Equation (3.12), the following positive distribution [56]:

$$P(t, \omega) = \frac{1}{2\pi} \operatorname{rect}\left(\frac{t}{T}\right) \operatorname{sinc}^2(T(\omega - \omega_0) / 2\pi) \quad (3.16)$$

We point out that while this kernel is unfamiliar and perhaps troubling in some respects (e.g., it diverges to infinity), it nonetheless satisfies the marginal conditions, $\phi(0, \tau) = \phi(\theta, 0) = 1$, and it yields a perfectly reasonable and well-behaved distribution for this signal. Note, too, that the kernel is indeed signal dependent, in that it depends explicitly on the duration T of the signal and further it yields a positive distribution for signals only of this form. Here again we see that, as in the case of

the conditional moment properties briefly mentioned earlier, the positive distributions can bring forth surprising results that are contrary to the conventional wisdom. For example, perhaps something is to be gained by using kernels that do not taper to zero.^{||}

3.2.2 Nonnegative bilinear distributions

The question of how far one can force bilinear distributions to be positive has been discussed by many investigators. Of course, it was Wigner who stated in his first paper that nonnegative bilinear distributions that satisfy the marginals do not exist. Hence if we are going to generate nonnegative bilinear distributions, they cannot satisfy the marginals. One approach is to simply ask what kernels when substituted into Equation (3.12) produce nonnegative distributions. Another equivalent approach is to choose a fixed distribution, such as the Wigner distribution, and consider the following:

$$W_{SM}(t, \omega) = \int L(t - t', \omega - \omega') W(t', \omega') dt' d\omega' \quad (3.17)$$

where $L(t, \omega)$ is a smoothing function. Other approaches have also been devised. The basic result of these approaches is that the resulting nonnegative distribution is a spectrogram or a sum of spectrograms [4, 41–44, 46, 49, 53, 54, 73, 76–78].

3.2.3 Significance of bilinearity

It is important to address the issue as to why there is a preference, indeed almost a requirement by some, for bilinearity of the distribution. One argument is that since the instantaneous power and the spectral density are bilinear in the signal, the joint density should also be bilinear. This argument has been analyzed in detail and has been shown to be false [19]. In fact, any legitimate joint distribution of any two variables is at least bilinear in its marginals. Thus, it follows that a proper TF distribution should also be bilinear *in its marginals*; and hence because the marginals themselves are each bilinear in the signal, the distribution is at least quadrilinear in the signal.

Another reason for the emphasis on bilinear distributions is that the bilinear distributions give a simple formulation for the calculation of averages. This is particularly so for the quantum case, from which TF distributions arose. In quantum mechanics, which is the most successful description we have of nature, the methodology of the theory is cast in Hilbert space formalism that involves taking expectations by placing operators between functions, and this method has a straightforward transcription for bilinear distributions. In particular, if we want to find the average value of the operator $\mathcal{G}(\mathcal{T}, \mathcal{W})$ we do this by calculating:

^{||}Along these lines, we point out that divergent windows for the spectrogram, which produce divergent kernels, have been previously considered. Their advantage is that one obtains a spectrogram with potentially desirable instantaneous frequency and instantaneous bandwidth properties [22, 27].

$$\langle \mathcal{G}(\mathcal{T}, \mathcal{W}) \rangle = \int s^*(t) \mathcal{G}(\mathcal{T}, \mathcal{W}) s(t) dt \quad (3.18)$$

Equation (3.18) is the average value of the physical quantity corresponding to the operator \mathcal{G} .

Now part of the attractiveness of joint representations is that one can calculate the average in the standard manner, that is, by phase–space integration with respect to a distribution $C(t, \omega)$, which may be a simpler calculation:

$$\langle \mathcal{G}(\mathcal{T}, \mathcal{W}) \rangle = \iint g(t, \omega) C(t, \omega) dt d\omega \quad (3.19)$$

For the two procedures to give the same answer, which is required if we are going to use the phase–space approach to calculate the average, the relationship between the function g and the operator \mathcal{G} is [7]:

$$\mathcal{G}(\mathcal{T}, \mathcal{W}) = \iint \gamma(\theta, \tau) \phi(\theta, \tau) e^{j\theta\mathcal{T} + j\tau\mathcal{W}} d\theta d\tau \quad (3.20)$$

where:

$$\gamma(\theta, \tau) = \frac{1}{4\pi^2} \iint g(t, \omega) e^{-j\theta t - j\tau\omega} dt d\omega \quad (3.21)$$

and $\phi(\theta, \tau)$ is the kernel of the distribution. Notice that Equation (3.18) is inherently bilinear if we assume that the operator is signal independent (which is usually the case). Also, in the proof of Equation (3.19) the distribution $C(t, \omega)$ is given by Equation (3.12) with the assumption that the kernel is signal independent, and hence the distribution is bilinear. Hence, obtaining averages by Equation (3.19) so that they equal the averages given by Equation (3.18) has been well developed for the case where $C(t, \omega)$ is bilinear. This formalism has not been developed for the positive distributions. Of course, perhaps it can be developed or another method can be developed, but thus far no one has done it.

We point out, however, that if the operator is of the form:

$$\mathcal{G}(\mathcal{T}, \mathcal{W}) = f(\mathcal{T}) + h(\mathcal{W}) \quad (3.22)$$

then *any* distribution $C(t, \omega)$ that satisfies the marginals, including the positive distributions, will give the correct averages as per Equations (3.18) and (3.19), where:

$$g(t, \omega) = f(t) + h(\omega) \quad (3.23)$$

That is, if *and only if* the distribution satisfies the marginals are we assured that for any arbitrary functions f and h can we obtain the correct average from the distribution:

$$\begin{aligned} \langle f(\mathcal{T}) + h(\mathcal{W}) \rangle &= \int s^*(t) [f(\mathcal{T}) + h(\mathcal{W})] s(t) dt \\ &= \iint [f(t) + h(\omega)] C(t, \omega) dt d\omega \end{aligned} \quad (3.24)$$

In fact, this equation is the condition that Wigner imposed in his original paper. Although the first integral is indeed bilinear in the signal, the second integral does

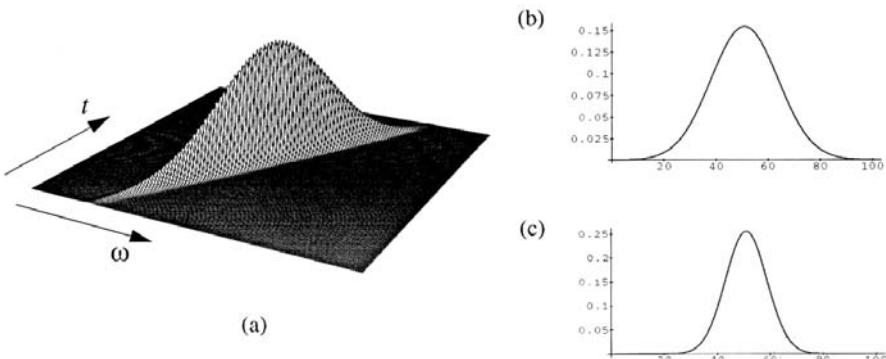
not need to be. In particular, Equation (3.24) is satisfied by the positive distributions, which are not bilinear, because all that is required is that the distribution yield the correct marginals. Restriction to bilinear distributions may be a convenience, but it is not a necessity. In fact, such a restriction forces the distribution to be negative so that it can yield the correct marginals and hence the correct averages. This has led to many discussions about the meaning of negative values and invocations of the uncertainty principle; however, as we have seen, there is no *a priori* requirement for the distribution to be bilinear, only that it satisfies the marginals. If we take $C(t, \omega)$ to be a positive distribution, all such discussions are moot. (We return to the uncertainty principle and its bearing on TF distributions later in this chapter.)

3.2.4 The cross-term issue

Bilinear distributions produce so-called *cross terms* and it has been a very important development in the field to understand these cross terms and to develop bilinear distributions that minimize them. The general view is that cross terms arise because the distribution is a nonlinear, specifically bilinear, function of the signal. A major advance in mitigating the cross terms was made by Williams, who developed the concepts needed to generate bilinear distributions that reduce the cross items while simultaneously preserving desirable properties of the distributions, particularly the marginals [5,47]. It would seem that if we have a cross-term problem with bilinear distributions, would we not have a much greater problem with the positive distributions, because they are even more nonlinear in the signal? The answer is not necessarily because it is possible for the nonlinearities to conspire to give something simple. (Even in the bilinear case this is true: consider, for example, the spectrogram, which like the Wigner distribution is bilinear in the signal, but unlike the Wigner distribution it does not suffer from the same cross-term problem.) Ultimately, of course, no amount of debate can answer this question. However, the positive distributions we present give simple, clear and sensible results.

Furthermore the positive distributions force us to reexamine the whole issue of cross terms. Consider that in the case of the energy density $|s(t)|^2$, or the energy spectral density $|S(\omega)|^2$, one does not usually go on about cross terms. Of course they “exist” in the sense that if we add two signals together, we get constructive and destructive interference, which we could label the cross terms. However, we do not, because the signal (and spectrum) is what it is, and we have come to understand the physical phenomena underlying the interference of waves. So why does the issue arise with TF distributions? Perhaps it is because the first distributions developed, such as the Wigner distribution, gave very insightful results for some signals but for other signals the results were far from clear and often troubling, in that energy appeared in the TF plane where none would be expected; for example, during time intervals when the signal was zero. These peculiarities were difficult to explain, and much attention was focused on understanding them. Ultimately, they came to be known as cross terms, and how to eliminate them became the dominant viewpoint in the field.

One is forced to wonder whether the issue would have even arisen if the positive distributions had been formulated first. To illustrate, let us consider a positive

**FIGURE 3.1**

A positive joint TFD and its marginals for a linear chirp signal, $s(t) \sim e^{-\alpha t^2/2} e^{j\omega_0 t + \beta t^2/2}$. Note that, like the marginal densities, the joint distribution is unimodal.

distribution of a linear chirp:

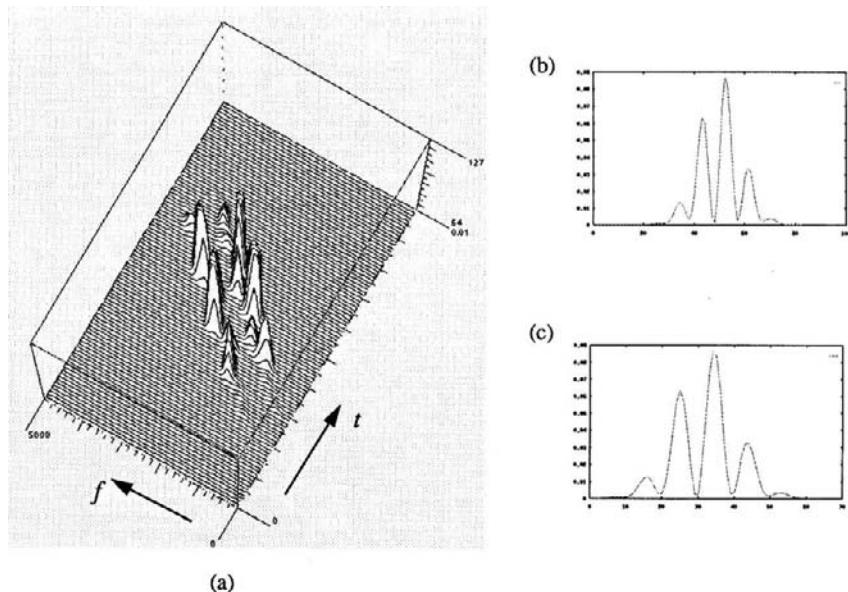
$$s(t) = (\alpha/\pi)^{1/4} e^{-\alpha t^2/2 + j\beta t^2/2 + j\omega_0 t} \quad (3.25)$$

which is plotted along with its marginal densities in Figure 3.1. Now let us add two chirps together. The positive time-frequency distribution (TFD) of this signal, along with its marginals, is shown in Figure 3.2. In the first case, namely, the single chirp, note that the marginals are unimodal, as is the joint distribution. In the second case, which is the sum of two chirps, the chirps constructively and destructively interfere, so that now the instantaneous power and the energy spectral density are multimodal, as shown. This phenomenon is well known and understood. So why should it not be that the joint distribution would also be multimodal, as is the positive distribution shown in Figure 3.2? Where are the so-called cross terms? If we do not raise the question for the marginal densities, why raise them for the joint distribution, in particular, for the case of the positive distribution shown?

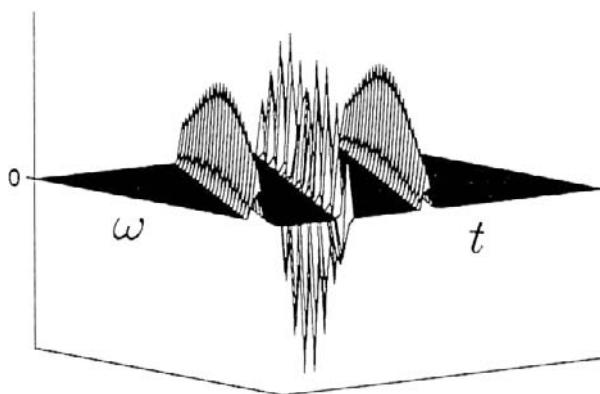
Now contrast this result with the Wigner distribution of the two chirps, which is shown in Figure 3.3: We see something very different and peculiar, namely, the large oscillating component midway between the two chirps. This component, which swings positive and negative, lies in the TF plane where no true signal component exists. This peculiarity of the Wigner distribution gave rise to the notion of cross terms, and because it is obviously an undesirable feature of the distribution, efforts in the development of the bilinear distributions were focused on eliminating or reducing these problems. However, as shown above, such peculiarities do not arise with the positive distributions.

3.2.5 Random signals

Just as with the bilinear distributions, the positive distributions can also be applied to random signals. In particular, as with the Wigner–Ville spectrum [72], which is

**FIGURE 3.2**

A positive joint distribution and its marginals for the sum of two chirps. Note that the marginal densities are multimodal, reflecting the destructive and constructive interference of the waves. A joint distribution should also reflect this physical phenomenon, as does the positive distribution.

**FIGURE 3.3**

The Wigner distribution for the sum of two chirps. Unlike the positive distribution, the Wigner distribution exhibits peculiarities, namely, a large oscillating term midway between the two chirps. This oddity of the Wigner distribution is called the cross term.

the expected value of the Wigner distribution of a random signal, one may define the positive distribution for random signals as:

$$\overline{P(t, \omega)} = E[|S(\omega)|^2 |s(t)|^2 \Omega(u(t), v(\omega))] \quad (3.26)$$

where here E denotes the expected value.

Note that the marginal densities are $E[|s(t)|^2]$ and $E[|S(\omega)|^2]$. If the random signal is stationary, the frequency marginal equals the power spectral density times a delta function, $E[|S(\omega)|^2] = P(\omega')\delta(\omega - \omega')$, where here $P(\omega)$ denotes the usual power spectrum (i.e., the FT of the autocorrelation function of the signal). Of course, when we are fortunate enough to be dealing with stationary and ergodic signals, we can evaluate the expected value by time averages; and in fact for that fortuitous case, there is no real need for TF analysis. For nonstationary random signals, there is no ergodic theorem and hence the expected value must either be approximated by local TF averaging, or as ensemble averages if multiple realizations are available. Further considerations of positive distributions for random signals can be found in [33, 83–86]. In particular, Pitton has developed a multiwindow method for estimating these TF distributions.

3.3 Joint Densities and Conditional Densities

The fundamental quantity of densities of two variables is the joint density. However, from the joint density one derives other quantities and densities that often are more fruitful for many reasons. Of particular importance are the conditional densities given by:

$$\begin{aligned} P(\omega | t) &= \frac{P(t, \omega)}{P(t)} \\ P(t | \omega) &= \frac{P(t, \omega)}{P(\omega)} \end{aligned} \quad (3.27)$$

where $P(t)$ and $P(\omega)$ are the marginal distributions:

$$\begin{aligned} P(t) &= \int P(t, \omega) d\omega \\ P(\omega) &= \int P(t, \omega) dt \end{aligned} \quad (3.28)$$

Because one gets the correct marginals for the positive distributions, we have:

$$\begin{aligned} P(t) &= |s(t)|^2 \\ P(\omega) &= |S(\omega)|^2 \end{aligned} \quad (3.29)$$

The conditional densities are one-dimensional (1-D) densities, are properly normalized and can be treated as any other 1-D density. However, they are densities for a fixed value of the other variable and therefore they have that other variable as a parameter. For example $P(\omega|t)$ is the density of frequency for a given time. Therefore, even though it is a 1-D density, one can plot it as a function of frequency vs. time by plotting it for successive values of time. Even though we may plot it this way, it is nevertheless important to appreciate the distinction between the joint density and the conditional density, a point that we return to in Section 3.4.

3.3.1 Conditional means, standard deviations, and proper or realizable quantities

From the 1-D conditional distributions one defines the conditional means by:

$$\langle \omega \rangle_t = \int \omega P(\omega|t) d\omega \quad (3.30)$$

$$\langle t \rangle_\omega = \int t P(t|\omega) dt \quad (3.31)$$

and the conditional standard deviations are defined by:

$$\sigma_{\omega|t}^2 = \int (\omega - \langle \omega \rangle_t)^2 P(\omega|t) d\omega \quad (3.32)$$

$$\sigma_{t|\omega}^2 = \int (t - \langle t \rangle_\omega)^2 P(t|\omega) dt \quad (3.33)$$

A very important fundamental and practical issue is whether quantities defined from joint densities, such as the conditional moments above, are proper and can be interpreted consistently and in the usual way. We must distinguish between “proper” and “correct.” We illustrate what we mean by example.

Consider the conditional standard deviation in ordinary probability. It is defined as a manifestly positive quantity, and it gives a measure of the spread of the distribution about its mean. However, that is not the case with the Wigner distribution and almost all other bilinear distributions, where it is possible for the conditional variance to be negative and hence for the conditional standard deviation to be complex valued. Therefore, although we may manipulate the bilinear distributions in the same manner as ordinary distributions, and calculate expectation values in the same way, often one does not know how to interpret the results. Of course, one can take the viewpoint that when the result is obviously wrong, such as when the conditional variance is negative, we simply do not bother trying to use it. However, what about the cases where the answer is not obviously wrong? With the bilinear distributions, how can we be sure that the answer is at least plausible? Because these distributions go negative, there is always the question of whether the results are reasonable and proper; hence we can never really focus on the true issue, which is whether the moments are correct. Again, we must distinguish between proper and correct. Proper moments are those that can be interpreted in the usual way, and do not have the nagging underlying doubt

as to whether they are plausible. That of course does not mean the answer is correct, only that it is possible. This question does not arise with the positive distributions. If one uses positive distributions, then one is assured that the moments are internally consistent and that they can be interpreted in the standard way. Hence, one can concentrate on the issue of whether the moments are correct.

3.3.2 Positive distributions, first conditional moments and instantaneous frequency: an example

It is illustrative to continue the previous discussions with a specific example. Consider two tones:

$$s(t) = A_1 e^{j\omega_1 t} + A_2 e^{j\omega_2 t} \quad (3.34)$$

$$= A(t) e^{j\varphi(t)} \quad (3.35)$$

where the amplitudes A_1 and A_2 are taken to be constants and ω_1 and ω_2 are positive. The spectrum of this signal consists of two delta functions at ω_1 and ω_2 :

$$S(\omega) = A_1 \delta(\omega - \omega_1) + A_2 \delta(\omega - \omega_2) \quad (3.36)$$

Solving for the phase and amplitude of the signal, we have:

$$\varphi(t) = \arctan \frac{A_1 \sin \omega_1 t + A_2 \sin \omega_2 t}{A_1 \cos \omega_1 t + A_2 \cos \omega_2 t} \quad (3.37)$$

$$A^2(t) = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\omega_2 - \omega_1)t \quad (3.38)$$

and taking the derivative of the phase we obtain the instantaneous frequency of this signal:

$$\varphi'(t) = \frac{1}{2}(\omega_2 + \omega_1) + \frac{1}{2}(\omega_2 - \omega_1) \frac{A_2^2 - A_1^2}{A^2(t)} \quad (3.39)$$

It is straightforward to verify that for unequal strength tones, the instantaneous frequency of the two-tone signal ranges outside the bandwidth [17, 70, 71]. Therefore, even though the first conditional spectral moment $\langle \omega \rangle_t$ of many bilinear distributions equals the instantaneous frequency, interpreting this conditional moment in the usual way, namely, as the average of one quantity (frequency) for a particular value of the other (time), yields the paradox that the supposed average exceeds the range of values of the quantity. The problem arises, of course, because the bilinear distributions are not proper. In particular, they take on negative values and hence the conditional moments cannot be interpreted in the usual way.

No such difficulty arises with the positive distributions, and in fact considering the conditional moments of the positive distributions has given rise to new ideas concerning instantaneous frequency. Because the first conditional moment of a positive distribution never ranges outside the spectral support of the signal, it usually does not equal the instantaneous frequency, although surprisingly for some signals it does.

This has led to the question of when is the instantaneous frequency equal to the average frequency at each time or, equivalently, under what conditions will the instantaneous frequency not range outside the spectral support of the signal? The answer is not that the signal must be monocomponent as is a common view, but rather when the signal exhibits a particular symmetry in the TF plane [74, 81]. Thus, there are multi-component signals for which the instantaneous frequency can indeed be interpreted as the average frequency at each time. Of course, most signals do not possess such symmetry, and accordingly the instantaneous frequency and the average frequency at each time are two different quantities.

These considerations concerning the conditional moments of positive TF distributions have led to new ideas about the average frequency of a signal at a given time, and new ideas on the amplitude and phase of a signal [69].

3.4 Joint vs. Conditional Densities, Time-varying Spectrum and Strong Support

3.4.1 Is $P(t,\omega)$ a time-varying spectrum?

It is important to appreciate the difference between the joint density $P(t,\omega)$ and the conditional density $P(\omega|t)$. Of course, they are related and we can get either one from the other by appropriate normalization by the marginal. However, that normalization leads to a critical and important distinction between the joint density and the conditional density. In particular, it has bearing on the interpretation of the density and specifically with regard to viewing TF distributions as the “time-varying spectrum” of the signal. Principally, despite a common view to the contrary, the distribution $P(t,\omega)$ is *not* the time-varying spectrum of the signal. Rather, it is the conditional distribution $P(\omega|t)$ that is more appropriately viewed as the time-varying spectrum. The reason is because it is from the conditional distribution that we directly obtain information about frequency at a particular time, such as the average frequency at a given time $\langle \omega \rangle_t$ and the spread in frequency at a particular time $\sigma_\omega(t)$ — see Equations (3.30) and (3.32).

Also, very often the joint distribution is zero where one might not expect a time-varying spectrum to be zero. This has caused some to conclude that there is something wrong with the positive distributions. Here again, the problem is not with the distribution, but rather with an incorrect presumption, namely, that the joint distribution is the time-varying spectrum. Often these difficulties in interpretation are readily resolved when we consider the appropriate density and appreciate that the joint and conditional densities are different, albeit related, functions.

To illustrate, let us consider the joint and conditional positive distributions of a two-tone signal:

$$s(t) = s_1(t) + s_2(t) \quad (3.40)$$

$$= A(t)[e^{j\omega_1 t} + e^{j\omega_2 t}] \quad (3.41)$$

where:

$$A(t) = e^{-\frac{t^2}{2\sigma^2}} \quad (3.42)$$

is a Gaussian amplitude that we apply to make the signal finite energy so that we can calculate the positive distribution. The instantaneous power of this signal is:

$$|s(t)|^2 = 4e^{-\frac{t^2}{\sigma^2}} \cos^2 \left[\frac{(\omega_1 - \omega_2)}{2} t \right] \quad (3.43)$$

For simplicity, let us take the correlationless positive distribution:

$$P(t, \omega) = |s(t)|^2 |S(\omega)|^2 / E \quad (3.44)$$

where E is the signal energy. In addition, we assume that $|\omega_1 - \omega_2| \gg 0$ and $\sigma^2 \gg 0$, for which case:

$$|S(\omega)|^2 \sim \sigma^2 [e^{-\sigma^2(\omega-\omega_1)^2} + e^{-\sigma^2(\omega-\omega_2)^2}] \quad (3.45)$$

Calculating the signal energy from the spectral density we find that $E \sim \sigma$, and therefore we obtain for the joint distribution:

$$P(t, \omega) \sim \sigma e^{-t^2/\sigma^2} \cos^2 \left[\frac{(\omega_1 - \omega_2)}{2} t \right] [e^{-\sigma^2(\omega-\omega_1)^2} + e^{-\sigma^2(\omega-\omega_2)^2}] \quad (3.46)$$

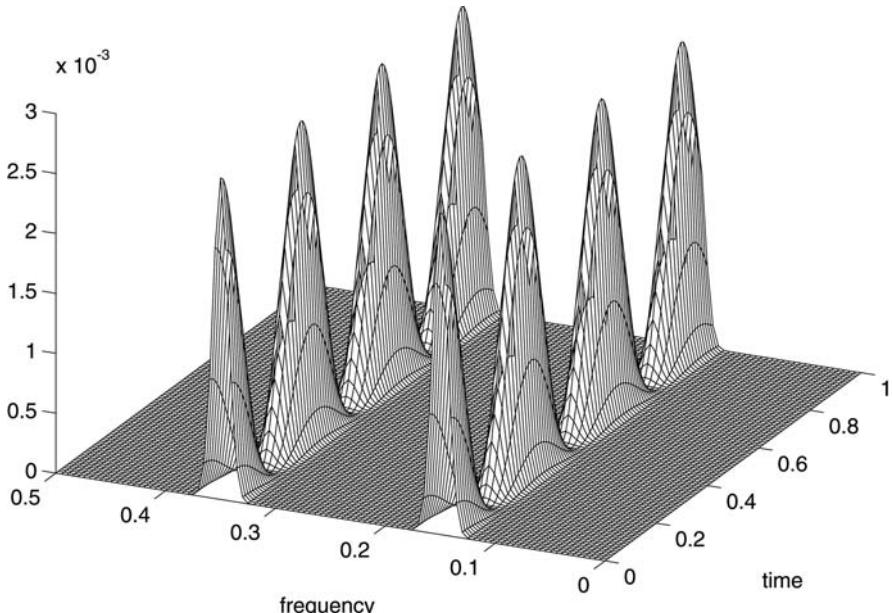


FIGURE 3.4

A joint positive distribution, Equation 3.46, for two complex tones. Note that, like the example of the two chirps, the joint distribution reflects the instantaneous power of the two tones, which because of the destructive and constructive interference of the waves, exhibits periodic peaks and troughs. Compare the joint distribution with the conditional distribution of the tones, shown in [Figure 3.5](#).

which is plotted in [Figure 3.4](#). Note that the joint distribution has peaks and valleys in time, which correspond to the maximum and minimum values of the instantaneous power of the signal. Many people believe that these variations in the instantaneous power of a signal should not be reflected as amplitude variations in the time-varying spectrum; hence, if we view the joint distribution as the time-varying spectrum, we may be compelled to conclude that there is something wrong with the positive distributions. Again, the problem is not with the distribution, but rather with the intent of viewing it as the time-varying spectrum.

Consider now the conditional distribution for this signal, which is given by:

$$P(\omega | t) \sim \sigma [e^{-\sigma^2(\omega - \omega_1)^2} + e^{-\sigma^2(\omega - \omega_2)^2}] \quad (3.47)$$

and is plotted in [Figure 3.5](#). Note that the amplitude variations corresponding to the instantaneous power of the signal that appeared in the joint distribution do not appear in the conditional distribution. Indeed, in this case, the conditional distribution appears to be a reasonable representation of the time-varying spectrum of this two-tone signal.

One reason that conditional distributions have not been considered to any significant degree previously in TF is that for bilinear distributions, the conditional distributions

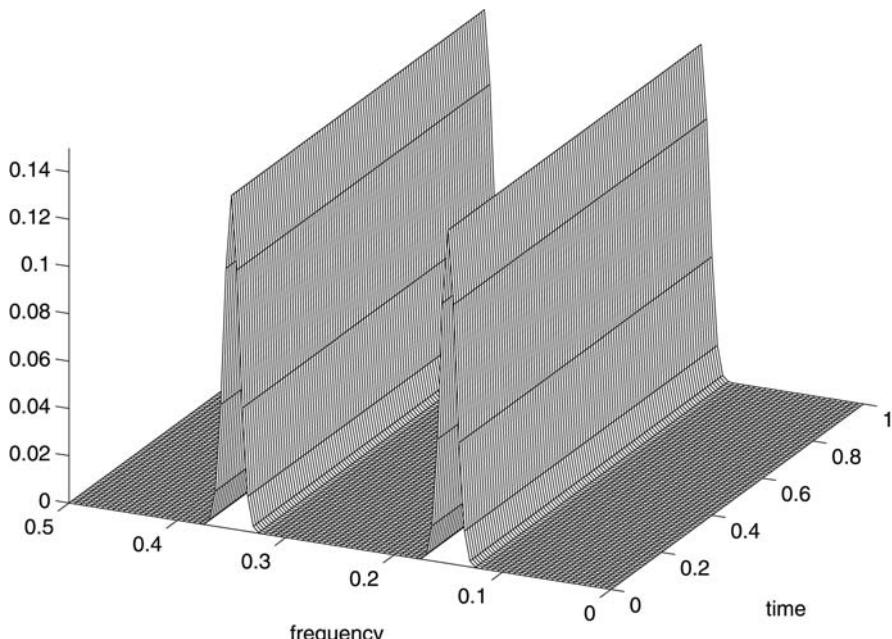


FIGURE 3.5

The time-conditional distribution, Equation 3.47, for two complex tones. Note that the conditional distribution, and not the joint distribution, is closer to our intuition of what a time-varying spectrum should look like — for two tones, it exhibits a spectral peak at each tone, which is constant over time.

are often problematic. In particular, because the cross term appears in the TF plane where no signal energy exists, when we divide the distribution by the marginal density to obtain the conditional distribution, we encounter infinities at the cross-term locations. Hence, it is usually difficult to compute conditional bilinear distributions and usually they cannot be sensibly interpreted. No such problems arise with the positive distributions.

3.4.2 Strong finite support

Let us return to considerations of the joint distribution, and, in particular, the fact that it exhibits peaks and valleys that are coincident with the instantaneous power of the signal (and also with the spectral density of the signal). That is, when the energy temporal density is large at a particular time, so is the joint distribution at that time; and when it is low, the joint distribution is low at that time. Likewise in frequency: if there is little power at a particular frequency in the spectral density, then the joint distribution reflects that fact.

This property is summarized in the “support” property of the distribution, and, in particular, the strong finite support property, so named to contrast it to the usual (weak) finite support property [16, 64]. The (weak) finite support property is well known and was touted as an advantage of the Wigner distribution over the spectrogram. It states that if a signal (or its spectrum) is zero outside some interval, then the distribution should also be zero outside the interval. The Wigner distribution satisfies this property, but the spectrogram does not because of the distortions caused by the windowing inherent to that method.

Most would agree that this property, namely, weak finite support, is a reasonable requirement; and we point out that, like the Wigner distribution, the positive distributions also satisfy this finite support property. Let us take our considerations further. Suppose now that a finite-duration signal is also zero during brief intervals over its duration. For example, suppose that we record someone counting numbers, “one” ... “two” ... etc., and suppose further that in between each spoken digit, the speaker pauses briefly so that no sound is made. Thus when we look at the instantaneous power, we see that it is large during the utterance of each digit, and it is zero in between each sound. Further, it is zero before and after the speaker started and ended the counting. Is it not reasonable to expect that the distribution should be zero not only before and after the speaker started counting but also during the brief intervals between sounds, when nothing was being said? (Let us ignore, as is done when making the usual argument for finite support, that some noise will occur in our recording, so that we can focus on the essence of the argument, namely, that when no signal power exists, is it not reasonable to expect that the distribution reflect this fact?) This is the essence of the strong finite support property. The Wigner distribution and the spectrogram do not satisfy this property. Thus, for example, in the recorded speech example, the Wigner distribution is zero prior to the start and after the end of the entire utterance, but it is not zero during the intervals of silence.

Similar considerations apply in the spectral domain as well. Suppose we have a sound that ranges between 300 to 3000 Hz, but the spectral density indicates that no

sound was produced between 1000 and 1100 Hz. The positive distribution will reflect this fact, but the Wigner distribution will not be zero over this spectral band.

These properties are referred to as *strong finite time support* and *strong finite frequency support*, respectively. Strong support implies the usual (weak) support condition, but the converse is not true. These properties are important not because the positive distributions satisfy them (in fact, bilinear distributions exist that satisfy the strong support conditions [64]), but because they are physically relevant. How can we proclaim to have a joint density of time and frequency that contradicts the individual densities, when it is the individual densities that (we hope) we all agree on? If no signal power exists at certain times or frequencies, that fact should be reflected by the joint TF density so as not to have a contradiction. Arguments against this point can usually be traced to a preference to interpret the joint density as the time-varying spectrum, which as we have already discussed is not the correct viewpoint and points again to the distinction between the joint density and the conditional density.

3.5 Uncertainty Principle, Positive Distributions and Wigner Distribution

Very often one hears the argument that the uncertainty principle implies that we cannot have a manifestly positive distribution, and further that the negative values of the Wigner distribution are a consequence of the uncertainty principle. Whether people have a mathematical or philosophical argument in mind is usually not clear. In either case it is a curious statement to make because even for the Wigner distribution an exception exists. Hence, how can there be an exception if it is a fact, philosophical or otherwise, that the uncertainty principle precludes positive distributions? There cannot be an exception if the uncertainty principle is paramount, as indeed it is. These arguments are wrong: the uncertainty principle has no bearing on the existence of positive distributions. We attempt to clarify the issue in this section.

The uncertainty principle depends on the marginals *only*, that is, on $|s(t)|^2$ and $|S(\omega)|^2$, and, critically, the fact that the underlying functions $s(t)$ and $S(\omega)$ are FT pairs. To see this, recall the definitions of spread (variance) in time and frequency:

$$\sigma_t^2 = \int (t - \langle t \rangle)^2 |s(t)|^2 dt \quad (3.48)$$

$$\sigma_\omega^2 = \int (\omega - \langle \omega \rangle)^2 |S(\omega)|^2 d\omega \quad (3.49)$$

From only these definitions, we obtain the uncertainty principle by multiplying the two equations together, making use of the fact that $s(t)$ and $S(\omega)$ are FT pairs, and applying the Schwarz inequality, by which the uncertainty principle follows [17]:

$$\sigma_t \sigma_\omega \geq \frac{1}{2} \quad (3.50)$$

Therefore, any distribution that satisfies the marginals satisfies the uncertainty principle. There are an infinite number of distributions, positive and otherwise, that satisfy the marginals. Because the positive distributions satisfy the marginals, they satisfy the uncertainty principle. The Wigner distribution also satisfies the uncertainty principle, and the fact that it has negative values is a peculiarity of the distribution, not of the uncertainty principle.

Incidentally, for the signal:

$$s(t) = (\alpha/\pi)^{1/4} e^{-\alpha t^2/2 + j\beta t^2/2 + j\omega_0 t} \quad (3.51)$$

the Wigner distribution is:

$$W(t, \omega) = \frac{1}{\pi} e^{-\alpha t^2 - (\omega - \beta t - \omega_0)^2 / \alpha} \quad (3.52)$$

which is manifestly positive. Does this distribution satisfy the uncertainty principle? Yes, because the Wigner distribution satisfies the marginals for any signal. Therefore, we see that even for the Wigner distribution there is a case where the distribution is positive and also satisfies the uncertainty principle. This shows that the issue of whether the distribution goes negative is not related to the uncertainty principle, even for the Wigner distribution.

If a distribution does not yield the correct marginals, the product of the variances may or may not satisfy the uncertainty product of the signal. The spectrogram, for example, does not satisfy the marginals and it does not yield the uncertainty product of the signal. Note that this is not a violation of the uncertainty principle, because as soon as we forgo the marginal conditions — the underlying functions of which are FT pairs — all bets are off. It is important, again, to appreciate that the uncertainty principle constrains the variances of functions that are FT pairs.

3.5.1 Local variances: is there an uncertainty principle?

Very often when people speak of the uncertainty principle (which is a global result), they intuitively really want to refer to something local. For example, it is often believed that one cannot have simultaneously good time resolution and good frequency resolution because the uncertainty principle precludes it. However, that is not what the uncertainty principle states. Instead, the uncertainty principle for signals prevents a signal from being both short duration and narrowband. However, those are global concepts, and for TF we are interested in knowing the frequency at a particular time, etc. What can we say about that? Can one develop “local” uncertainty principles? The answer is yes, and indeed the positive distributions clarify the issue considerably [58, 60, 62]. Moreover, the results we present have nothing to do with TF per se but are true for any joint distribution. We nevertheless illustrate our point for time and frequency.

Of course, one must define what is meant by “resolution” when making claims about what the uncertainty principle limits. However, that is usually not done and thus considerations of the uncertainty principle and its bearing, if any, on resolution are

usually ambiguous. Therefore, let us stick to variance as our measure. In particular, for our measure of local resolution we consider the *conditional* or local variances, which we denote by $\sigma_{t|\omega}^2$ and $\sigma_{\omega|t}^2$, which are the variance in time at a particular frequency, and the variance in frequency at a particular time, respectively. Is there an uncertainty product for the local variances? We address this question as follows.

It can be shown that for *any* distribution of two variables $P(x, y)$ the local variances and global variances are related by [17]:

$$\sigma_y^2 = \int \sigma_{y|x}^2 P(x) dx + \int (\langle y \rangle_x - \langle y \rangle)^2 P(x) dx \quad (3.53)$$

$$\sigma_x^2 = \int \sigma_{x|y}^2 P(y) dy + \int (\langle x \rangle_y - \langle x \rangle)^2 P(y) dy \quad (3.54)$$

Note that all the preceding terms are nonnegative. We emphasize that these equations are mathematical identities for any joint distribution. Let us therefore write the identical equations for TF distributions:

$$\sigma_\omega^2 = \int \sigma_{\omega|t}^2 |s(t)|^2 dt + \int (\langle \omega \rangle_t - \langle \omega \rangle)^2 |s(t)|^2 dt \quad (3.55)$$

$$\sigma_t^2 = \int \sigma_{t|\omega}^2 |S(\omega)|^2 d\omega + \int (\langle t \rangle_\omega - \langle t \rangle)^2 |S(\omega)|^2 d\omega \quad (3.56)$$

where we have assumed that the TF distribution satisfies the marginals because that is a necessary and sufficient condition to satisfy the usual uncertainty principle.

As just mentioned, for a proper distribution, such as the positive distributions, all the preceding terms are nonnegative. It follows immediately, that if we multiply the two equations together, we observe that the left-hand side is the product of the global variances, and the right-hand side is the product of the average local variances, plus other terms *all of which are positive*. Hence, the product of the average local variances is *upper bounded* by the global variances. That is, if we define the average of the local variances as:

$$\langle \sigma_{\omega|t}^2 \rangle = \int \sigma_{\omega|t}^2 |s(t)|^2 dt \quad (3.57)$$

$$\langle \sigma_{t|\omega}^2 \rangle = \int \sigma_{t|\omega}^2 |S(\omega)|^2 d\omega \quad (3.58)$$

then we obtain the following inequality relation [58, 60, 62]:

$$0 \leq \langle \sigma_{\omega|t}^2 \rangle \langle \sigma_{t|\omega}^2 \rangle \leq \sigma_\omega^2 \sigma_t^2 \quad (3.59)$$

This result is true for any proper joint distribution and has nothing specifically to do with TF distributions. The fact that the global variances may or may not be lower bounded has no bearing on Equation (3.59). Of course, this equation does not generally follow for bilinear TF distributions, because for these distributions, the conditional variances are not necessarily positive and therefore we cannot manipulate the preceding equations in the same way to obtain the inequalities in Equation (3.59).

Is there a lower limit? In the inequality expression above, we have indicated that the lower bound on the product of the average local variances may be identically zero. We now show that this is indeed so by way of example [62].

Consider the example of the Wigner distribution of the chirp signal, which is given by Equation (3.52). For this particular case the Wigner distribution is a positive distribution, that is, it is a proper distribution, and accordingly we may calculate and manipulate its global and local variances as we did above for proper distributions. The average of the local variances of the Wigner distribution of this signal are:

$$\begin{aligned}\langle \sigma_{\omega|t}^2 \rangle &= \frac{\alpha}{2} \\ \langle \sigma_{t|\omega}^2 \rangle &= \frac{\alpha}{2(\alpha^2 + \beta^2)}\end{aligned}\quad (3.60)$$

and their product is therefore:

$$\langle \sigma_{\omega|t}^2 \rangle \langle \sigma_{t|\omega}^2 \rangle = \frac{1}{4} \frac{\alpha^2}{\alpha^2 + \beta^2} \quad (3.61)$$

The global variances for this case are:

$$\begin{aligned}\sigma_t^2 &= \frac{1}{2\alpha} \\ \sigma_\omega^2 &= \frac{\alpha^2 + \beta^2}{2\alpha}\end{aligned}\quad (3.62)$$

and their product satisfies the standard uncertainty principle:

$$\sigma_t^2 \sigma_\omega^2 = \frac{1}{4} \frac{\alpha^2 + \beta^2}{\alpha^2} \geq \frac{1}{4} \quad (3.63)$$

Note too that the inequality between the average local variances and the global variances is also satisfied; that is, we have:

$$\langle \sigma_{\omega|t}^2 \rangle \langle \sigma_{t|\omega}^2 \rangle = \frac{1}{4} \frac{\alpha^2}{\alpha^2 + \beta^2} \leq \sigma_t^2 \sigma_\omega^2 = \frac{1}{4} \frac{\alpha^2 + \beta^2}{\alpha^2} \quad (3.64)$$

Note further that α and β are signal parameters that can be any value we choose. Let us consider signals for which we keep β fixed and consider what happens when we take the limit as $\alpha \rightarrow 0$. We have that:

$$\begin{aligned}\sigma_t^2 \sigma_\omega^2 &\rightarrow \infty \\ \text{as : } \alpha &\rightarrow 0\end{aligned}\quad (3.65)$$

This is a sensible result because as α approaches zero we are making the chirp of longer and longer duration.

Now consider the average local variance product for the same limit:

$$\begin{aligned}\langle \sigma_{\omega|t}^2 \rangle \langle \sigma_{t|\omega}^2 \rangle &= \frac{1}{4} \frac{\alpha}{\alpha^2 + \beta^2} \rightarrow 0 \\ \text{as : } \alpha &\rightarrow 0\end{aligned}\quad (3.66)$$

We see that although the product of the global variances goes to infinity, the product of the average local variances goes to zero. This, too, is a sensible result because as α gets smaller and smaller, the amplitude of the chirp is becoming constant and accordingly less and less amplitude modulation occurs so that the instantaneous bandwidth of the chirp is getting smaller and smaller.

We point out that for this case, it so happens that the local variances are also equal to the expressions given in Equation (3.60). That is, for the Wigner distribution of the chirp, the local variances themselves are:

$$\begin{aligned}\sigma_{\omega|t}^2 &= \frac{\alpha}{2} \\ \sigma_{t|\omega}^2 &= \frac{\alpha}{2(\alpha^2 + \beta^2)}\end{aligned}\tag{3.67}$$

and hence:

$$\sigma_{\omega|t}^2 \sigma_{t|\omega}^2 = \frac{1}{4} \frac{\alpha}{\alpha^2 + \beta^2}\tag{3.68}$$

Therefore we have:

$$\sigma_{\omega|t}^2 \sigma_{t|\omega}^2 \rightarrow 0\tag{3.69}$$

as :

$$\alpha \rightarrow 0$$

Accordingly, not only the average local variances can be made as close to zero as we want but also the local variances themselves can be made as close to zero as we want (at least for this example). Of course, for any signal, if the average of the local variance is zero, then it must be that the local variance itself is zero for any proper distribution, because the variance is a nonnegative quantity. In addition, it is certainly conceivable that for other signals, even when the average of the local variance is not zero, the local variance itself could be zero at particular times or frequencies, for which the local uncertainty product at specific times or frequencies would be identically zero.

The inequality we have derived that bounds the product of the average of the local variances is exact, and demonstrates that no uncertainty product applies to the local variances. This is a very simple and straightforward result, and it has nothing specifically to do with TF.^{**} That is, in general, the local variances are upper bounded by the global variances and lower bounded by zero for any legitimate joint density. The fact that the global variances themselves may be lower bounded is irrelevant in terms of this bound on the local variances. Although the usual uncertainty principle places a lower bound on the product of the global variances, it does not place a lower bound on the local variances.

^{**}In fact, consideration of the bilinear distributions would not have led us to these results, because the local variances of the bilinear distributions are not nonnegative and hence we cannot interpret and manipulate them as we have done to obtain the inequality relation in Equation (3.59). It was through considerations of the local variances of positive distributions that we were led to this result.

3.5.2 Local uncertainty for the spectrogram

Why then does the notion persist that the uncertainty principle precludes simultaneously good resolution in time and frequency? One reason is that the spectrogram has dominated development and influenced our thinking to such a strong degree, and for the spectrogram it is indeed true that the uncertainty principle limits what can be done.^{††} However, that is a limitation of the spectrogram. It is the windowing operation of the spectrogram and the uncertainty principle of the window that cause frequency resolution to degrade as we improve time resolution with the spectrogram. However, the window is an operation that we have introduced in the analysis, and no such problem arises with the positive distributions as we have just shown. Further, the spectrogram does not satisfy the marginals, and it is therefore not a proper distribution. What we find from the positive distributions is that in fact there is no inherent limit on the local variances, and we can do better than the spectrogram.

3.6 Construction of Positive Distributions

An infinite number of positive distributions exists for any signal, each characterized by a particular choice for the kernel $\Omega(u(t), v(\omega))$. This situation is analogous to the bilinear distributions, which are characterized by the kernel $\phi(\theta, \tau)$. As it was with the bilinear distributions, the issue now with the positive distributions is how to choose and construct them in some reasonable or systematic way. The first practical procedure for accomplishing this objective was developed by and Loughlin, Pitton and Atlas [55, 63, 65]. Subsequently, other methods have been developed for constructing positive distributions. The common approach is to use a constrained optimization procedure to obtain the distribution directly, rather than trying to explicitly obtain the kernel $\Omega(u, v)$. This approach, namely, to obtain the distribution directly, is in fact common to all the methods to date, with the exception of [100]. We briefly review the main approaches and refer the interested reader to the original papers for the details [28–30, 34, 37, 65, 75, 82–84, 87–89, 93, 95–97].

3.6.1 Maximum entropy

With the maximum entropy approach, or more generally minimum cross entropy, the basic idea is to find a positive distribution that is closest in a minimum cross entropy sense to an initial guess, or “prior distribution,” subject to the marginal constraints, and possibly additional constraints such as conditional or joint moments. Thus, we

^{††}We refer the reader to [4, 18, 20, 31, 80, 103, 104] for detailed considerations of the uncertainty principle as it pertains to the spectrogram.

seek the distribution $P(t, \omega)$ that minimizes the integral:

$$\Delta H = \iint P(t, \omega) \log\left(\frac{P(t, \omega)}{P_0(t, \omega)}\right) dt d\omega \quad (3.70)$$

where $P_0(t, \omega)$ is the prior distribution. Loughlin, Pitton and Atlas took the spectrogram, or combinations of spectrograms, as the prior, and then solved for the non-negative distribution that is closest to the prior in a minimum cross-entropy measure, subject to the marginal and possibly other constraints [65]. In practice, the solution is obtained iteratively. Additional contributions to the method, including generalized marginal constraints, fast algorithms and the design of informative priors (i.e., priors in addition to spectrograms), have also been made [34, 37, 96]. The maximum entropy approach has been applied to many real-world signals such as analysis of machine faults [59, 91, 92], speech [88] and biomedical signals [3, 67, 68].

Additional constraints such as instantaneous frequency and group delay constraints may also be imposed, but it is often the case that these constraints are inconsistent with the requirements that the distribution is nonnegative and yields the correct marginals, as discussed previously. Even when these or other constraints are consistent, from a practical point of view it is usually difficult to impose them via maximum entropy because one must iteratively solve nonlinear equations. Accordingly, other optimization criteria have been investigated for implementing positive distributions.

3.6.2 Least squares

With the success of the maximum entropy approach in terms of demonstrating the relative ease by which positive distributions could be obtained, and further because the results obtained were very reasonable and often exhibited greater clarity and detail than the bilinear distributions, improvements and modifications quickly followed, including more efficient practical methods of implementation. Some of these we have already noted in the previous section, such as the fast algorithm of Groutage [37] and the consideration of constraints beyond the marginals. In this section, we summarize methods that use least-squares measures as the optimality criterion, rather than maximum entropy and minimum cross entropy.

Least squares is a widely used optimization measure, largely because it is mathematically tractable and lends itself to efficient methods of solution and practical implementation. Contributors to the development of iterative least-squares techniques for constructing positive TF distributions include Nickel, Sang and Williams [75], Sang, Williams and O’Neil [93]; Emresoy and Loughlin [28, 29]; and Pitton [83, 84]. (Least squares has also been used to obtain distributions that are not necessarily positive [97].) The idea is similar to the approach of maximum entropy, except that we seek the distribution $P(t, \omega)$ closest to the prior in a least-squares sense. That is, we seek the distribution that minimizes the integral:

$$\iint [P(t, \omega) - P_0(t, \omega)]^2 dt d\omega \quad (3.71)$$

subject to positivity of $P(t, \omega)$, the marginal constraints, and possibly additional constraints such as joint and/or conditional moments, among others. From a practical standpoint, it is easier to impose these latter constraints via least squares than via maximum entropy. For example, to impose an instantaneous frequency constraint, in maximum entropy one must solve the equation [65]:

$$\int \omega P_i(t, \omega) \exp(\lambda(t) \omega) d\omega = \varphi'(t) |s(t)|^2 \quad (3.72)$$

for $\lambda(t)$ and then apply the iteration:

$$P_{i+1}(t, \omega) = P_i(t, \omega) \exp(\lambda(t) \omega) \quad (3.73)$$

Conversely, the least-squares solution to the same problem yields the iteration [29]:

$$P_{i+1}(t, \omega) = P_i(t, \omega) + \lambda(t) \omega \quad (3.74)$$

where:

$$\lambda(t) = \frac{[\varphi'(t) |s(t)|^2 - \int \omega' P_i(t, \omega') d\omega']}{\int \omega'^2 d\omega'} \quad (3.75)$$

where the integration over frequency in the denominator term is over the spectral support of the TF distribution.

The results obtained with the least-squares approaches are similar to those obtained with maximum entropy, particularly for weighted least squares where of course we use the same prior distribution and constraints for both approaches [28]. Again, the main advantage of (weighted) least-squares over entropy-based optimization is numerical efficiency and the ease by which constraints such as conditional moments can be implemented.

3.6.3 Deconvolution

A third approach that has been developed, first introduced by Pitton, Atlas and Loughlin [87–89] and then modified by Emresoy and El-Jaroudi [30], makes use of the known convolution relationship between TF distributions. In particular, any two bilinear distributions are related by [6, 15, 17]:

$$P_1(t, \omega) = \iint B(t', \omega') P_2(t - t', \omega - \omega') dt' d\omega' \quad (3.76)$$

A similar relationship has been derived for the positive distributions, and, in particular it has been shown that a spectrogram can be expressed as a “blurred” positive distribution [87–89]. Accordingly, borrowing from deblurring methods in image processing, one can develop deconvolution techniques to iteratively deblur a spectrogram and obtain a positive TF distribution [88].

3.6.4 Gabor expansion technique

A recent interesting approach has been given by Pedersen [82]. Since the Wigner distribution of a Gabor logon [36] (a complex tone with a Gaussian amplitude) is a positive TF distribution, the idea is to expand the signal and its FT in terms of normalized Gabor logons and then combine the expansions in such a way so as to preserve nonnegativity and closely approximate the marginals. This method yields a nonnegative distribution with the correct time marginal, and a frequency marginal that is very close to the true frequency marginal (Pedersen notes that in numerical implementations, the differences “are negligible” [82]). How to obtain a distribution with specified conditional moments using this technique is an open question at this time. Nevertheless, because the Gabor expansion-based technique is not iterative, unlike the constrained optimization techniques, it may provide a more numerically efficient method for generating positive TF distributions. Other techniques for fast approximations of positive TF distributions have also been considered [66]; and as the application of these distributions continues to grow, it is likely that more efficient methods will continue to emerge (although we point out that the optimization procedures, particularly the least-squares methods, are not all that computationally intensive).

3.7 Collection of Examples and Plausibility of Results

We now show, by way of several examples with comparisons to the Wigner distribution and a spectrogram, that the results given by the positive distributions are plausible and reasonable and satisfy our intuition. As noted previously, the positive distributions have been applied to many real signals in a variety of areas [57]. However, in this section we concentrate on synthetic signals because we are interested in bringing forth the salient features of these distributions and therefore it is important to be certain of the signal we are investigating. We refer the interested reader to the previously cited references for applications to real signals, as well as to the accompanying chapter by Groutage et al. [38].

In Figures 3.6 to 3.9 we present a number of examples. We refer the reader to the figure captions for additional comments.

Two tone bursts — This example, shown in Figure 3.6, highlights the strong finite support properties. The signal consists of a low-frequency tone, followed by an interval where the signal is zero and then followed by a high-frequency tone. Note in Figure 3.6 that the positive distribution exhibits strong finite support, meaning it is zero when the signal is zero and when the spectrum is zero. This is not the case for either the Wigner distribution or the spectrogram. (In Figures 3.6 to 3.9, the distributions are plotted as gray-scale images, where black is highest intensity and white is lowest intensity. For the Wigner distribution, which goes negative, only the positive portion of the distribution is plotted. Along the top of each plot is the real

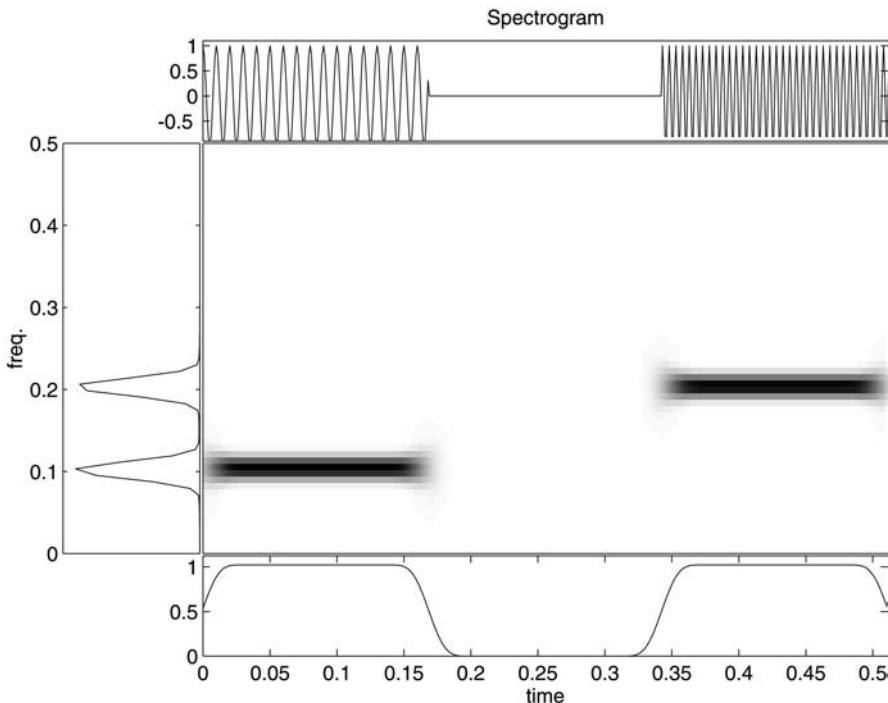


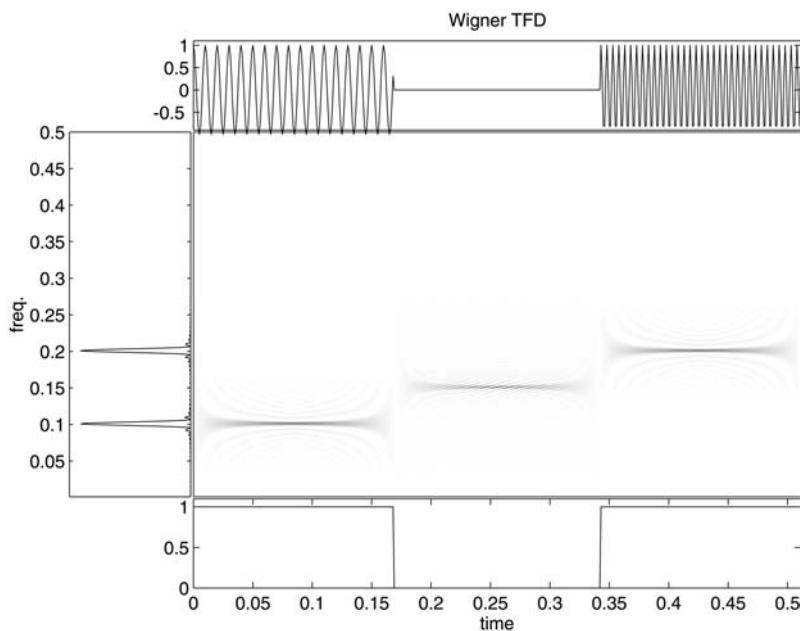
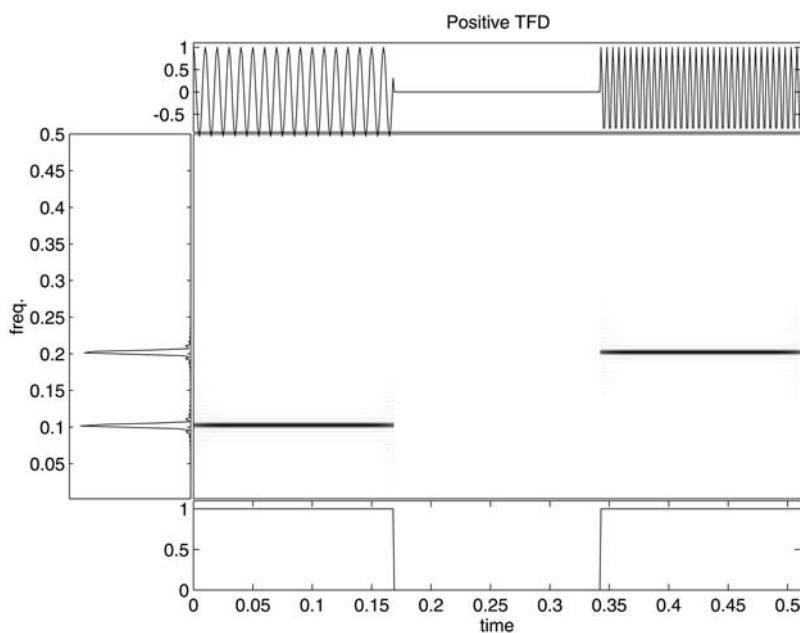
FIGURE 3.6a

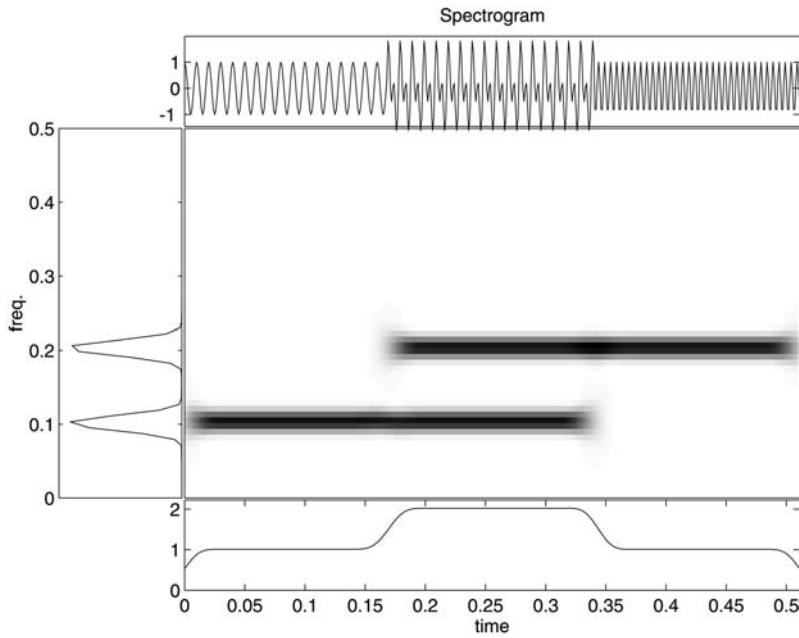
The signal consists of two finite-duration complex tones at two different times. The distributions shown are (a) spectrogram, (b) Wigner distribution and (c) positive distribution.

part of the signal, to the left is the frequency marginal of the distribution, and along the bottom panel is the time marginal.)

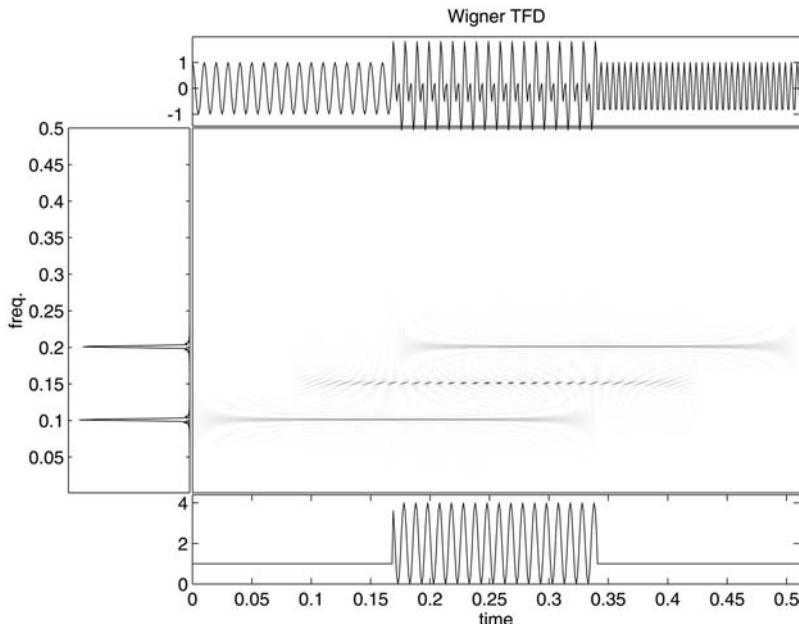
Two overlapping tones — In this example, the signal is similar to that of the previous example, except now the two tones overlap partially in time. Accordingly, destructive and constructive interference of the waves occurs during this interval, as reflected in the instantaneous power. This example, plotted in Figure 3.7, highlights this property of positive distributions, as well as the strong support and marginal properties. For Figure 3.7, note that during the interval when the signals overlap, there is constructive and destructive interference in the instantaneous power (shown in the bottom panel of the Wigner and positive distribution plots) of the signal. Accordingly, this physical phenomenon must be reflected in the joint distribution. The positive distribution exhibits this event in a manner consistent with the time marginal, and shows spectral content only at the frequencies of the signal. That is not the case with the Wigner distribution or the spectrogram.

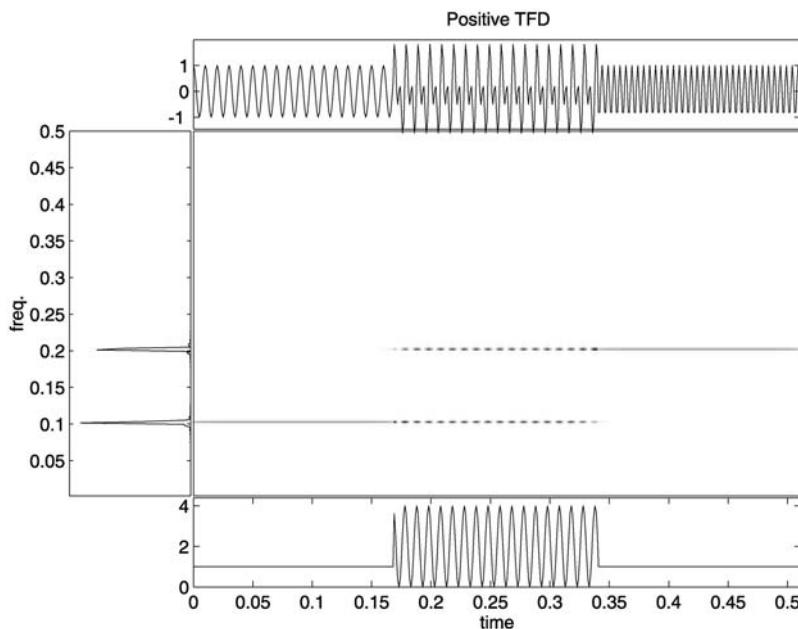
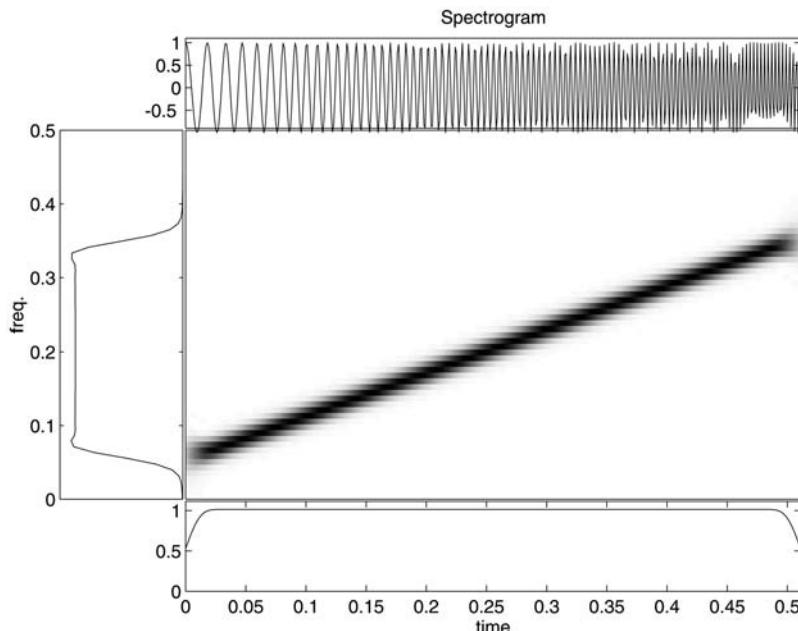
Linear FM chirp — In this example the instantaneous frequency of the signal increases linearly with time. Accordingly, the spectrum is broad, but in the TF plane, one expects energy concentration along the instantaneous frequency of this signal.

**FIGURE 3.6b****FIGURE 3.6c**

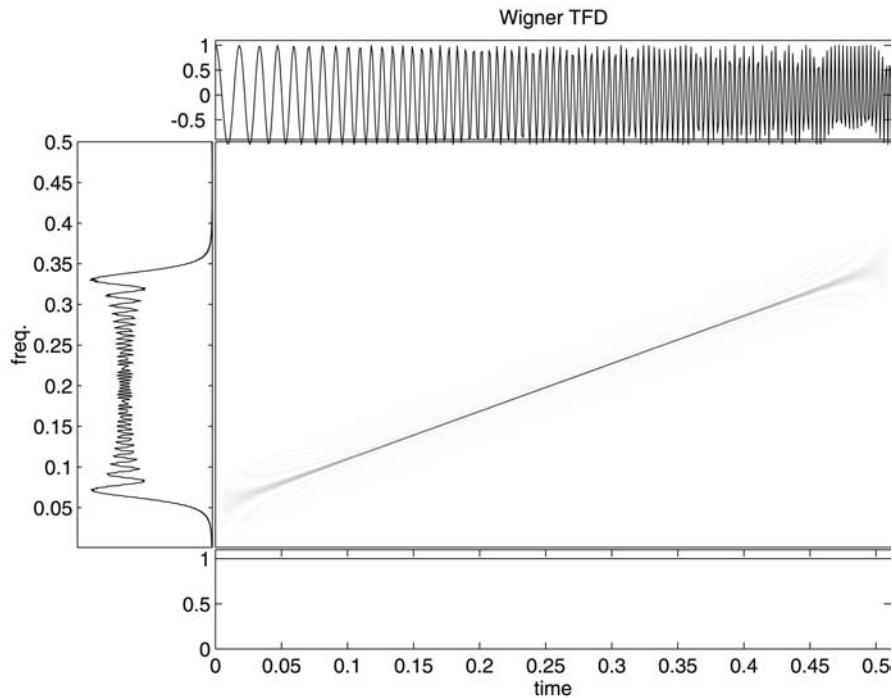
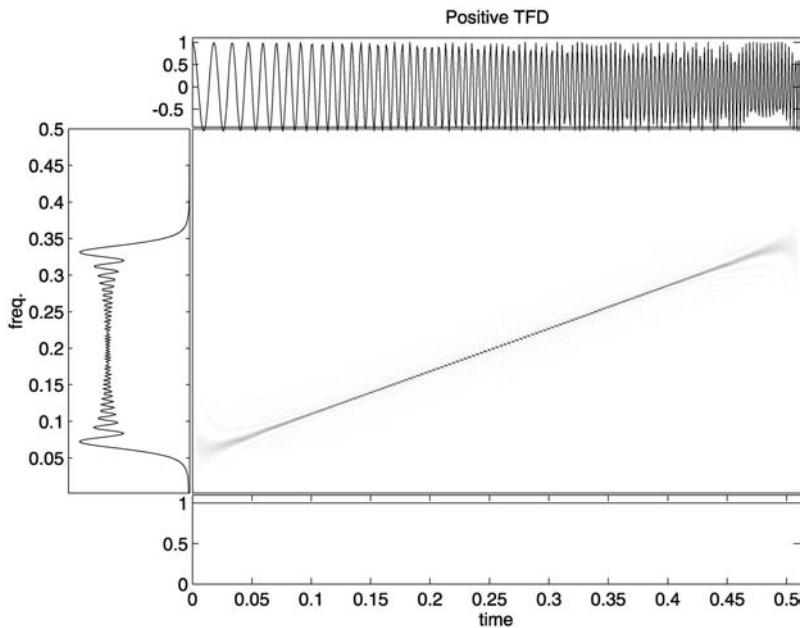
**FIGURE 3.7a**

The signal is similar to that shown in [Figure 3.6](#), except that now the two tones overlap partially in time. The distributions are (a) spectrogram, (b) Wigner distribution and (c) positive distribution.

**FIGURE 3.7b**

**FIGURE 3.7c****FIGURE 3.8a**

The signal is a finite-duration, complex linear FM chirp with constant amplitude. (a) spectrogram, (b) Wigner distribution, (c) positive distribution.

**FIGURE 3.8b****FIGURE 3.8c**

That is the case with the Wigner distribution and the positive distribution shown in Figure 3.8, but the spectrogram is broader, again due to the distortions caused by the window. Although it goes negative (because the chirp is finite duration), the Wigner distribution in Figure 3.8 exhibits energy highly concentrated along the instantaneous frequency of the chirp, and the conditional mean frequency of the distribution equals the instantaneous frequency of the signal. Like the Wigner distribution of this signal, the positive distribution also exhibits high concentration along the instantaneous frequency, and its conditional mean frequency is equal to the instantaneous frequency of the signal. Unlike the Wigner distribution, the positive distribution is nonnegative while simultaneously satisfying the marginals.

Tone + Impulse — In this example the signal consists of a narrowband tone and a broadband impulse, as shown in Figure 3.9. This example demonstrates the ability

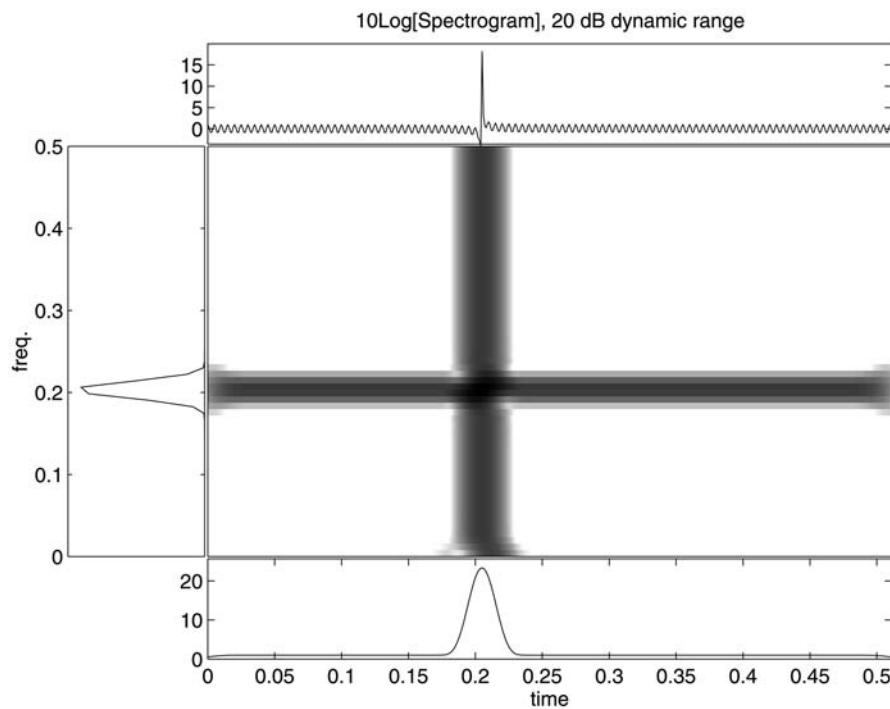


FIGURE 3.9a

The signal consists of a complex tone plus an impulse. The Wigner distribution in (b) and the positive distribution in (c) both simultaneously resolve the impulse and the tone, unlike the spectrogram in (a), which suffers from a TF resolution trade-off due to the windowing inherent to that technique. (TFD plots here are log-amplitude gray-scale, spanning a range of approximately 30 dB for the positive distribution, and 20 dB for the spectrogram and Wigner distribution.)

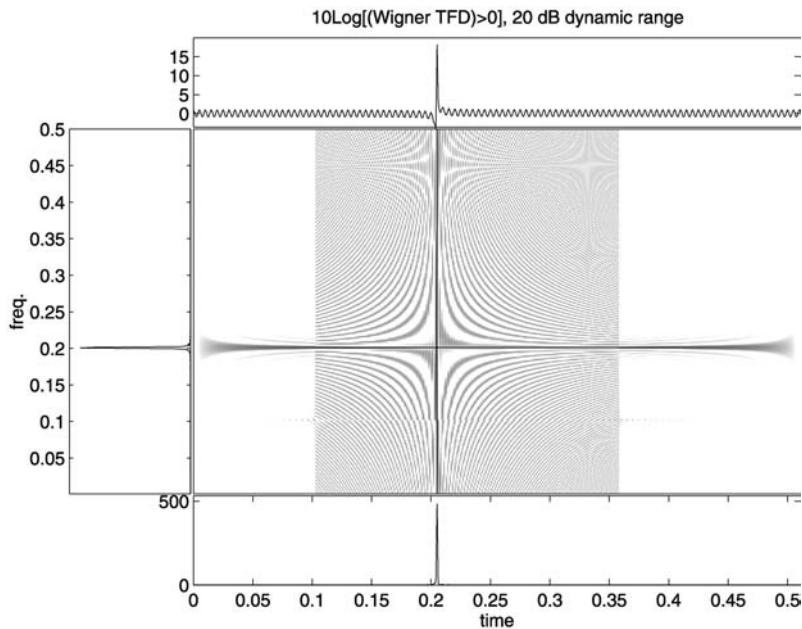


FIGURE 3.9b

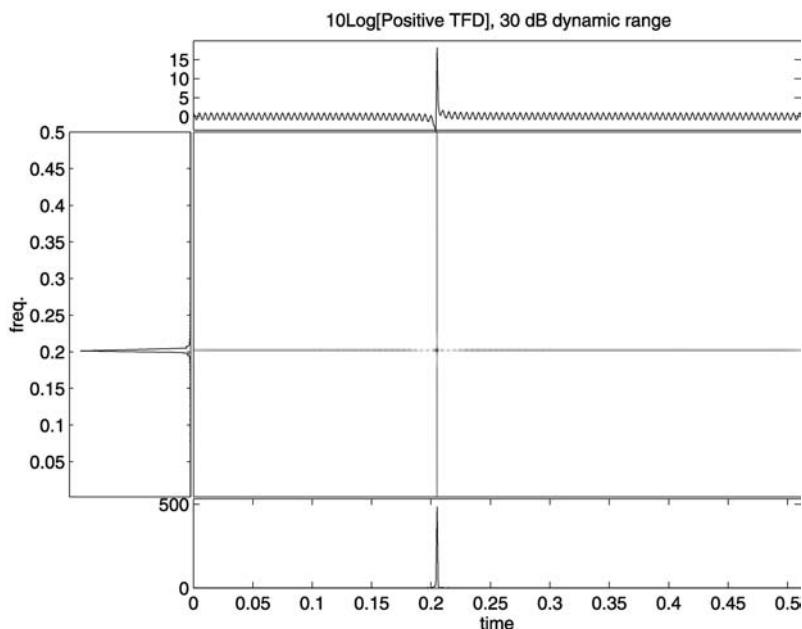


FIGURE 3.9c

of the positive distributions, and the Wigner distribution, to simultaneously resolve narrowband and broadband signal components, unlike the spectrogram, which is limited by the window inherent to that method.

3.8 Conclusions

In this chapter we reviewed the theory and methods of implementation of the positive TF distributions. A number of issues were discussed and clarified, including the question of bilinearity, cross terms and bearing of the uncertainty principle on the existence of such distributions and on the local resolution of the distribution. Examples were presented to illustrate the properties of these distributions.

Acknowledgments

The authors have been fortunate to collaborate with many of the contributors to the development and application of positive TF distributions. We gratefully acknowledge the following individuals for many enlightening discussions and contributions: Les Atlas, Gary Bernard, Dale Groutage, Jim Pitton, Ted Posch, Bob Rohrbaugh and Bill Williams.

The work of P. Loughlin was supported by the Office of Naval Research (N00014-98-1-0680), the National Science Foundation (MIP-9624089), the National Institutes of Health (R01-DC04435) and Boeing Commercial Airplane. The work of L. Cohen was supported by the Office of Naval Research and the NSA HBCU/MI program.

References

- [1] S. Basu and V.V. Lybas, *Ann. Fond. L. Broglie*, 8, 271, 1982.
- [2] F. Bopp, La Mechanique Quantique Est–Elle Une Mechanique Statistique Classique Particuliere?, *Ann. L'Inst. H. Poincare*, 15, 81–112, 1956.
- [3] J. Bulgrin, B. Rubal, T. Posch, and J. Moody, Comparison of Binomial, ZAM and minimum Cross-Entropy Time-Frequency Distributions of Intracardiac Heart Sounds, in Proceedings of the 28th Asilomar Conference, 1994, pp. 383–387.

- [4] N.D. Cartwright, *Physica*, 83A, 210, 1976.
- [5] H. Choi and W. Williams, Improved time-frequency representation of multi-component signals using exponential kernels, *IEEE Trans. ASSP*, 37(6), 862–871, 1989.
- [6] T. Claassen and W. Mecklenbrauker, The Wigner distribution — A tool for time-frequency analysis, Part I–III, *Phil. J. Res.*, 35(3–6), 217–250, 276–300, 372–389, 1980.
- [7] L. Cohen, Generalized phase–space distribution functions, *J. Math. Phys.*, 7, 781–786, 1966.
- [8] L. Cohen, Can quantum mechanics be formulated as a classical probability theory?, *Philos. Sci.*, 33, 317, 1966.
- [9] L. Cohen, Quantization problem and variational principle in the phase space formulation of quantum mechanics, *J. Math. Phys.*, 17, 1863, 1976.
- [10] L. Cohen, Local kinetic energy in quantum mechanics, *J. Chem. Phys.*, 70, 788, 1979.
- [11] L. Cohen, Representable local kinetic energy, *J. Chem. Phys.*, 80, 4277, 1984.
- [12] L. Cohen, Distribution functions for multivariate marginals, *J. Math. Phys.*, 25, 2402, 1984.
- [13] L. Cohen, Properties of positive time-frequency distribution functions, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 27(6), 1, 1985.
- [14] L. Cohen, A critical review of the fundamental ideas of joint time-frequency distributions, in Proceedings of the IEEE International Conference on Systems and Circuits, 1, 42, 1986.
- [15] L. Cohen, Time-frequency distributions — a review, in Proceedings of the IEEE, 77, 941–981, 1989.
- [16] L. Cohen, A primer on time-frequency analysis, in *New Methods in Time-Frequency Analysis*, B. Boashash, Ed., Longman-Cheshire, Sydney, 1992.
- [17] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, New York, 1995.
- [18] L. Cohen, Uncertainty principles of the short-time Fourier transform, in Proceedings of the SPIE, 2563, 80–90, 1995.
- [19] L. Cohen, Time-Varying Spectral Analysis — New Concepts and Unresolved Problems, in Proceedings IEEE Symposium on Applications of Time-Frequency and Time-Scale Methods, 1995, pp. 3–10.
- [20] L. Cohen, The uncertainty principle for windowed wave functions, *Optics Commun.*, 179, 221–229, 2000.

- [21] L. Cohen, The uncertainty principle for the short-time Fourier transform and wavelet transform in *Wavelet Transforms and Time-Frequency Signal Analysis*, L. Debnath, Ed., Birkhauser, Basel, 2001, pp. 217–232.
- [22] L. Cohen and C. Lee, Local bandwidth and optimal windows for the short-time Fourier transform, in Proceedings of the SPIE Advanced Algorithms, Architectures and Signal Processing IV, 1152, 401–425, 1989.
- [23] L. Cohen and C. Lee, Instantaneous bandwidth for signals and spectrogram, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2451, 1990.
- [24] L. Cohen and T. Posch, Positive time-frequency distribution functions, *IEEE Trans. ASSP*, 33, 31–38, 1985.
- [25] L. Cohen and T.E. Posch, Generalized ambiguity functions, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 27(6) 1, 1985.
- [26] L. Cohen and Y. Zaparovanny, Positive quantum joint distributions, *J. Math. Phys.*, 21, 794–796, 1980.
- [27] K. Davidson and P. Loughlin, Compensating for window effects in the calculation of spectrographic instantaneous bandwidth, *IEEE Trans. Biomed. Eng.*, 47(4), 556–558, 2000.
- [28] M. Emresoy and P. Loughlin, Weighted least-squares implementation of Cohen–Posch TFDs, *IEEE Trans. Signal Process.*, 46(3), 753–757, 1998.
- [29] M. Emresoy and P. Loughlin, Weighted least squares Cohen–Posch time-frequency distributions with specified conditional and joint moment constraints, *IEEE Trans. Signal Process.*, 47(3), 893–900, 1999.
- [30] M. Emresoy and A. El-Jaroudi, Evolutionary spectrum estimation by positivity constrained deconvolution, *IEEE Trans. Signal Process.*, 47(3), 889–893, 1999.
- [31] B.-G. Englert and K. Wodkiewicz, *Phys. Rev. A*, 51, R2661, 1995.
- [32] E. Entralgo, V. Kuryshkin, and Y.-I. Zaparovanny, in *Microphysical Reality and Quantum Formalism*, Kluwer Academic, Dordrecht, 1988.
- [33] P. Flandrin, On the positivity of the Wigner-Ville spectrum, *Signal Process.*, 11, 187–189, 1986.
- [34] J. Fonollosa, Positive time-frequency distributions based on joint marginal constraints, *IEEE Trans. Signal Process.*, 44(8), 2086–2091, 1996.
- [35] A. Francos and Porat, Parametric Estimation of Multicomponent Signals Using Minimum Cross Entropy Time-Frequency distributions, in IEEE-SP Proceedings of the International Symposium on Time-Frequency/Time-Scale Analysis, 1996, pp. 321–324.

- [36] D. Gabor, Theory of communication, *IEE J. Commun. Eng.*, 93, 429–457, 1946.
- [37] D. Groutage, A fast algorithm for computing minimum cross-entropy positive time-frequency distributions, *IEEE Trans. Signal Process.*, 45(8), 1954–1970, 1997.
- [38] D. Groutage, D. Bennink, P. Loughlin, and L. Cohen, Positive Time-Frequency Distributions and Acoustic Echoes, this volume.
- [39] R. Hudson, When is the Wigner quasi-probability density non-negative?, *Rep. Math. Phys.*, 6, 249–252, 1974.
- [40] K. Husimi, in Proceedings of the Phys. Math. Soc. Japan, 22, 264, 1940.
- [41] A.J.E.M. Janssen, Positivity of weighted Wigner distributions, *SIAM J. Math. Anal.*, 12, 752–758, 1981.
- [42] A. J. E. M. Janssen, Positivity properties of phase–plane distribution functions, *J. Math. Phys.*, 25, 2240–2252, 1984.
- [43] A.J.E.M. Janssen and T.A.C.M. Claasen, On positivity of time-frequency distributions, *IEEE Trans. Acoust. Speech Signal Process.*, 33, 1029–1032, 1985.
- [44] A.J.E.M. Janssen, Bilinear phase–plane distribution functions and positivity, *J. Math. Phys.*, 26, 1986–1994, 1985.
- [45] A.J.E.M. Janssen, A note on “positive time-frequency distributions,” *IEEE Trans. Acoust. Speech Signal Process.*, 35, 701–703, 1987 (see also author’s reply, *ibid.*).
- [46] A.J.E.M. Janssen, Positivity of time-frequency distribution functions, *Signal Process.*, 14, 243–252, 1988.
- [47] J. Jeong and W. Williams, Kernel design for reduced interference distributions, *IEEE Trans. Signal Process.*, 40(2), 402–412, 1992.
- [48] W. Koenig, H. Dunn, and L. Lacy, The sound spectrograph, *J. Acoust. Soc. Am.*, 18(1), 19–49, 1946.
- [49] V. Kuryshkin, *Inst. H. Poincare*, 22(1), 81, 1972.
- [50] V. Kuryshkin, *Int. J. Theor. Phys.*, 7, 451, 1972.
- [51] V. Kuryshkin, Some problems of quantum mechanics possessing a non-negative phase–space distribution functions, *Int. J. Theor. Phys.*, 7, 451–466, 1973.
- [52] V. Kuryshkin, Uncertainty principle and the problems of joint coordinate–momentum probability density in quantum mechanics, in *The Uncertainty Principle and the Foundations of Quantum Mechanics*, John Wiley & Sons, New York, 1977, pp. 61–83.

- [53] V. Kuryshkin, I. Lybas, and Y.-I. Zaparovanny, *Ann. Fond. L. de Broglie*, 5, 105, 1980.
- [54] V. Kuryshkin and Y.-I. Zaparovanny, *C. R. Acad. Sci. (Paris)*, 2, 17, 1984.
- [55] P. Loughlin, Time-Frequency Energy Density Functions: Theory and Synthesis, Ph.D. dissertation, University of Washington, Seattle, 1992.
- [56] P. Loughlin, Comments on scale invariance of time-frequency distributions, *IEEE Signal Proc. Lett.*, 2(12), 217–218, 1995.
- [57] P. Loughlin, Cohen-Pesch (positive) time-frequency distributions: development and applications, *Appl. Signal Process.*, 4, 122–130, 1997.
- [58] P. Loughlin, Moments and maximum entropy densities in time-frequency, in Proceedings of the SPIE, 3461, 110–119, 1998.
- [59] P. Loughlin and G. Bernard, Cohen-Pesch (positive) time-frequency distributions and their application to machine vibration analysis, *Mech. Syst. Signal Proc.*, 11(4), 561–576, 1997.
- [60] P. Loughlin and K. Davidson, Positive Local Variances of Time-Frequency Distributions and Local Uncertainty, in Proceedings of the IEEE-SP International Symposium Time-Frequency and Time-Scale Analysis, 1998, pp. 541–544.
- [61] P. Loughlin, D. Groutage, and R. Rohrbaugh, Time-frequency analysis of acoustic transients, in Proceedings of the IEEE ICASSP'97, 3, 2125–2128, 1997.
- [62] P. Loughlin, J. Pitton, and L. Atlas, Proper Time-Frequency Energy Distributions and the Heisenberg Uncertainty Principle, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, 1992, pp. 151–154.
- [63] P. Loughlin, J. Pitton, and L. Atlas, An information-theoretic approach to positive time-frequency distributions, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing,'92, 5, 125–128, 1992.
- [64] P. Loughlin, J. Pitton, and L. Atlas, Bilinear time-frequency representations: new insights and properties, *IEEE Trans. Signal Process.*, 41(2), 750–767, 1993.
- [65] P. Loughlin, J. Pitton, and L. Atlas, Construction of positive time-frequency distributions, *IEEE Trans. Signal Process.*, 42(10), 2697–2705, 1994.
- [66] P. Loughlin, J. Pitton, and B. Hannaford, Approximating time-frequency density functions via optimal combinations of spectrograms, *IEEE Signal Process. Lett.*, 1(12), 199–202, 1994.
- [67] P. Loughlin, M. Redfern, and J. Furman, Time-varying characteristics of visually induced postural sway, *IEEE Trans. Rehab. Eng.*, 4(4), 416–424, 1996.

- [68] P. Loughlin and M. Redfern, Spectral characteristics of visually-induced postural sway in healthy elderly and healthy young subjects, *IEEE Trans. Rehab. Eng.*, 9(1), 24–30, 2001.
- [69] P. Loughlin and B. Tacer, On the amplitude- and frequency-modulation decomposition of signals, *J. Acoust. Soc. Am.*, 100(3), 1594–1601, 1996.
- [70] P. Loughlin and B. Tacer, Instantaneous frequency and the conditional mean frequency of a signal, *Signal Process.*, 60(2), 153–162, 1997.
- [71] L. Mandel, Interpretation of instantaneous frequencies, *Am. J. Phys.*, 42, 840–846, 1974.
- [72] W. Martin and P. Flandrin, Wigner–Ville spectral analysis of nonstationary processes, *IEEE Trans. Acoust. Speech Signal. Process.*, 33(6), 1461–1470, 1985.
- [73] G. Mourguès, M.R. Feix, J.C. Andrieux, and P. Bertrand, Not necessary but sufficient conditions for the positivity of generalized Wigner functions, *J. Math. Phys.*, 26, 2554–2555, 1985.
- [74] W. Nho and P. Loughlin, When is instantaneous frequency the average frequency at each time?, *IEEE Signal Process. Lett.*, 6(4), 78–80, 1999.
- [75] R. Nickel, T. Sang, and W. Williams, A new signal adaptive approach to positive time-frequency distributions with suppressed interference terms, *IEEE Proc. ICASSP'98*, 3 1777–1780, 1998.
- [76] A.H. Nuttall, Wigner Distribution Function: Relation to Short–Term Spectral estimation, Smoothing, and Performance in Noise, Naval Underwater Systems Center, Technical report p. 8225, 1998.
- [77] A.H. Nuttall, The Wigner Distribution Function with Minimum Spread, Naval Underwater Systems Center, Technical report 8317, 1988.
- [78] R.F. O’Connell and E.P. Wigner, Quantum–mechanical distribution functions: conditions for uniqueness, *Phys. Lett.*, 83A, 145–148, 1981.
- [79] R.F. O’Connell, The Wigner distribution function — 50th birthday, *Found. Phys.*, 13, 83–92, 1983.
- [80] R.F. O’Connell and E.P. Wigner, *Phys. Lett.*, 85A, 121, 1981.
- [81] P. Oliveira and V. Barroso, Instantaneous frequency of multicomponent signals, *IEEE Signal Process. Lett.*, 6(4), 81–83, 1999.
- [82] F. Pederson, A Gabor expansion-based positive time-dependent power spectrum, *IEEE Trans. Signal Process.*, 47(2), 587–590, 1999.
- [83] J. Pitton, Linear and quadratic methods for positive time-frequency distributions, *IEEE Proc. ICASSP'97*, 5, 3649–3652, 1997.

- [84] J. Pitton, Positive time-frequency distributions via quadratic programming, *Multidimensional Syst. and Signal Process.*, 9(4), 439–445, 1998.
- [85] J. Pitton, The Statistics of Time-Frequency Analysis, *J. Franklin Inst.*, 337, 379–388, 2000.
- [86] J. Pitton, A spectral generating theorem for nonstationary stochastic signals, in Proceedings of the SPIE Advanced Signal Processing Algorithms, Architectures, and Implementation, 4116, 25–32, 2000.
- [87] J. Pitton, L. Atlas, and P. Loughlin, Deconvolution for Positive time-frequency distributions, in Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, 1450–1454, 1993.
- [88] J. Pitton, L. Atlas, and P. Loughlin, Applications of positive time-frequency distributions to speech processing, *IEEE Trans. Speech Audio Proc.*, 2(4), 554–566, 1994.
- [89] J. Pitton, P. Loughlin, and L. Atlas, Positive time-frequency distributions via maximum entropy deconvolution of the evolutionary spectrum, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing '93, 4, 436–439, 1993.
- [90] R. Rohrbaugh, Application of time-frequency analysis to machinery condition assessment, in Proceedings of the 27th Asilomar Conference on Signals, Systems and Computers, 2, 1455–1458, 1993.
- [91] R. Rohrbaugh, Time-frequency analysis of a motor-generator, in Proceedings of the SPIE Advanced Algorithms, Architectures and Signal Processing VI, 2846, 240–250, 1996.
- [92] R. Rohrbaugh and L. Cohen, Time-Frequency Analysis of a cam-Operated Pump, in Proceedings of the 49th Meeting Society Machinery Failure Prevention Technology, 1995, pp. 349–361.
- [93] T. Sang, W. Williams, and J. O'Neil, An Algorithm for Positive Time-Frequency Distributions, IEEE-SP Proceedings of the International Symposium on Time-Frequency/Time-Scale Analysis, 1996, pp. 165–168.
- [94] M.O. Scully and L. Cohen, Quasi-probability distributions for arbitrary operators, in *The Physics of Phase Space*, Y.S. Kim and W.W. Zachary, Eds., Springer-Verlag, New York, 1987.
- [95] S. Shah, A. El-Jaroudi, P. Loughlin, and L. Chaparro, Signal synthesis and positive time-frequency distributions, *J. Franklin Inst.*, 337, 317–328, 2000.
- [96] S. Shah, P. Loughlin, L. Chaparro, and A. El-Jaroudi, Informative priors for minimum cross-entropy positive time-frequency distributions, *IEEE Signal Process. Lett.*, 4(6), 176–177, 1997.

- [97] R. Streifel, Synthesis of Time-Frequency Representations by the Method of Projections onto Convex Sets, Masters thesis, Department of Electrical Engineering, University of Washington, Seattle, 1991.
- [98] B. Tacer and P. Loughlin, Time-scale energy density functions, *IEEE Trans. Signal Process.*, 44(5), 1310–1314, 1996.
- [99] B. Tacer and P. Loughlin, Nonstationary signal classification using the joint moments of time-frequency distributions, *Pattern Recognition*, 31(11), 1998.
- [100] H. Tohyama, S. Kikkawa, and K. Ohara, On an optimum positive time-frequency distribution by the least squares method, *Trans. Inst. Electron. Inform. Commun. Eng.*, J75-A1(3), 661–663, 1992 (in Japanese); S. Kikkawa and H. Tohyama, A least squares method for generating positive time-frequency distributions [unpublished English manuscript].
- [101] E.P. Wigner, On the quantum correction for thermodynamic equilibrium, *Phys. Rev.*, 40, 749–759, 1932.
- [102] E.P. Wigner, Quantum-mechanical distribution functions revisited, in *Perspectives in Quantum Theory*, W. Yourgrau and A. van der Merwe, Eds., MIT Press, Cambridge, MA, 1971, pp. 25–36.
- [103] K. Wodkiewicz, *Phys. Rev. Lett.*, 52, 1064, 1984.
- [104] K. Wodkiewicz, *Phys. Lett.*, A124, 207, 1987.
- [105] Y. Zhao, L. Atlas, and R. Marks, II, The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals, *IEEE Trans. ASSP*, 38(7), 1084–1091, 1990.

4

Positive Time–Frequency Distributions and Acoustic Echoes

Dale Groutage

Naval Surface Warfare Center

David Bennink

ManTech System Engineering Co.

Patrick Loughlin

University of Pittsburgh

and

Leon Cohen

City University of New York

CONTENTS

4.1	Introduction	163
4.2	Time–Frequency Distributions	164
4.3	Acoustic Scattering and Acoustic Transients	166
4.4	Feature Extraction and Classification from Time–Frequency Distributions ..	173
4.5	Conclusions	175
	References	175

4.1 Introduction

A number of studies have investigated acoustic scattering using time–frequency distributions [6, 7, 9, 14–17, 20–24]. The advantage of using time–frequency methods for acoustic scattering is that the elastic responses of the insonified object can be seen more clearly in the time–frequency plane as contrasted to examining the time or frequency response. Most of the previous studies have used bilinear distributions such as the spectrogram or Wigner distribution. In this chapter we demonstrate applications of the positive distributions to study acoustic scattering and acoustic transients. Comparisons to spectrograms are provided to highlight advantages provided by the positive time–frequency distributions (TFDs). Details on the theory and methods for positive

TFDs have been provided in a companion chapter in this volume, but nonetheless for the sake of continuity we summarize the basic ideas in this chapter before proceeding with examples of their application to acoustic signals.

4.2 Time–Frequency Distributions

TFDs are particularly useful in analyzing a signal when the spectral properties of a signal are changing in time. When frequencies are changing in time the time series itself or its Fourier transform (FT) does not give a full picture of how the spectral properties are changing. A number of mechanisms cause frequencies to change and we discuss the main ones here. Of prime importance is when the physical parameters governing a system are themselves not constant, in time or otherwise. For example, if the spring constant of a simple harmonic oscillator is changing in time or is position dependent, then the frequency of the oscillation will change. Of course, almost all musical instruments take advantage of changing physical parameters to produce changing frequencies. Another type of common situation is when the signal is going through a filter that is changing. Dispersion, both structural (or geometric) and dispersion due to the media, is yet another important mechanism that induces changing frequencies in a signal as it propagates, particularly in acoustics.

Another type of situation for which time–frequency methods are useful is when the signal is composed of many frequencies that are due to different sources. Time–frequency analysis is often able to separate the frequencies in time and hence to show that different sources exist. Now the “sources” can be those that are producing the signal or can be scattering sources where each scatterer has different scattering spectral properties. We illustrate this with an example of some importance. Consider the scattering from objects that have internal structure such as submarines. Submarines are composed of many substructures such as ballast tanks, missile-launch tubes, bulkheads and pressure tanks, each having its own dispersive and propagating properties. When a submarine is insonified by a sound source, each substructure scatters sound. Scattered sound, or echoes, from all the submarine substructures combine to form a total scattered echo, or signature. A joint time–frequency energy density can provide a means to find the individual contributions of each major substructure as they contribute to the overall scattering. From the joint time–frequency energy density, it is sometimes possible to identify the major contributions of individual scatterers that come from internal substructures. If indeed one can identify these internal structures, they would be an important signature. This is one of the motivating aims of the recent interest in time–frequency analysis for the study of elastic scattering. For an overview of these ideas and applications, we refer the reader to the review articles by [Dragonette et al.](#) and [Gaunaurd and Strifors](#) [6, 9] as well as the other articles noted earlier in Section 4.1 and listed at the end of the chapter.

4.2.1 Types of time-frequency distributions

The development of time-frequency distributions has a long history and we refer the reader to the standard review articles and books on the subject for a thorough understanding of the field and its development (e.g., [2, 3, 8]). However, we provide a simple discussion as to the types of distributions currently available. First, we note that the basic idea has always been to have a joint density of time and frequency $P(t, \omega)$ that shows how the energy of the signal is distributed in time and frequency. This joint density should satisfy some basic properties and one of the most important is that the marginals are satisfied. The marginal densities are $|s(t)|^2$ and $|S(\omega)|^2$ where $s(t)$ is the signal and $S(\omega)$ is its FT:

$$s(t) = \frac{1}{\sqrt{2\pi}} \int S(\omega) e^{j\omega t} d\omega \quad (4.1)$$

$$S(\omega) = \frac{1}{\sqrt{2\pi}} \int s(t) e^{-j\omega t} dt \quad (4.2)$$

The time and frequency marginals are also called the instantaneous power and the energy density spectrum, respectively. The marginal conditions are:

$$\int P(t, \omega) d\omega = |s(t)|^2 \quad (4.3)$$

$$\int P(t, \omega) dt = |S(\omega)|^2 \quad (4.4)$$

It was Wigner who gave, in 1932, the first distribution that satisfied the marginal conditions; this distribution is now commonly called the Wigner distribution. The Wigner distribution is bilinear in the signal. Although it satisfies the important condition of the marginals, it takes on negative values for almost all signals, which renders its interpretation as a true joint density problematic. Negative values in a time-frequency distribution are common to all bilinear distributions that satisfy the marginals, and it was Wigner himself who proved that bilinear distributions that satisfy the marginal conditions cannot be manifestly positive for all signals.

In the 1940s the spectrogram was developed, which is also bilinear and is constructed in such a way that it is manifestly positive. (It is the squared magnitude of the short-time Fourier transform (STFT) of the signal.) The spectrogram does not satisfy the marginals and the degree to which it approximates them depends on the window function.* To have a distribution that is manifestly positive and also satisfies the marginal conditions one must explore nonbilinear distributions, that is, distributions that are more nonlinear in the signal than bilinear distributions. Such distributions have been developed and explicitly constructed, and their calculation

*The windowing operation also introduces other difficulties, such as a trade-off in resolution between time and frequency, and the entanglement of signal and window properties that cannot usually be separated. It should be emphasized that these difficulties, in particular the resolution trade-off, are inherent to this particular method and are not a fundamental limitation of TFDs.

is very straightforward [4, 5, 19]. These are the so-called “positive distributions.” See [Chapter 3](#) in this volume by Loughlin and Cohen for additional details on the theory and methods for obtaining positive TFDs.

4.3 Acoustic Scattering and Acoustic Transients

In this section we give examples of a positive TFD of the acoustic scattering response from an underwater cylindrical shell, and of a transient signal from an underwater vehicle, with comparison to the results obtained using spectrograms.

4.3.1 Example 1: elastic cylindrical shell

This example illustrates analysis of data from the acoustic scattering response of an elastic cylindrical shell. The incident angle for the insonification energy is 75 degrees with respect to the centerline of the cylinder. The scattering response of the cylindrical shell is illustrated in Figure 4.1. This signal contains two distinct parts; one is the specular response and the other is the resonant response. The specular response is that portion of the signal in the time period roughly from 0.2 to 0.4 normalized time

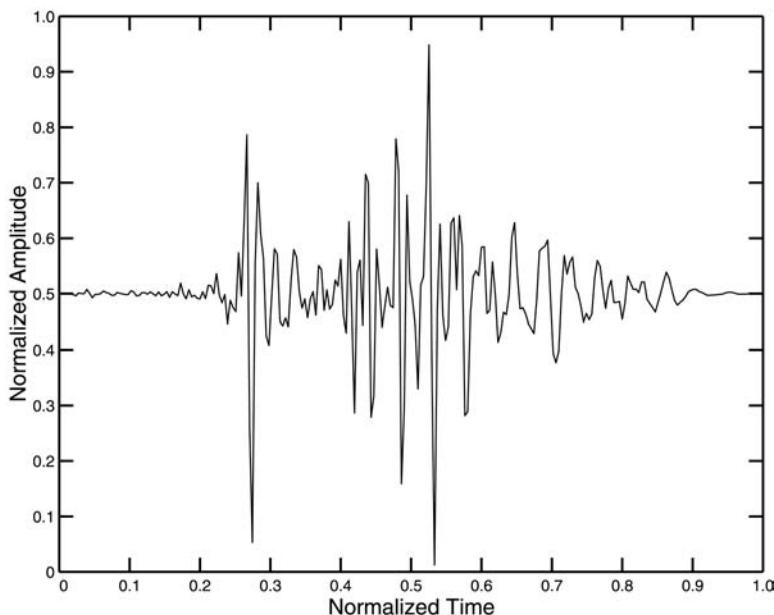
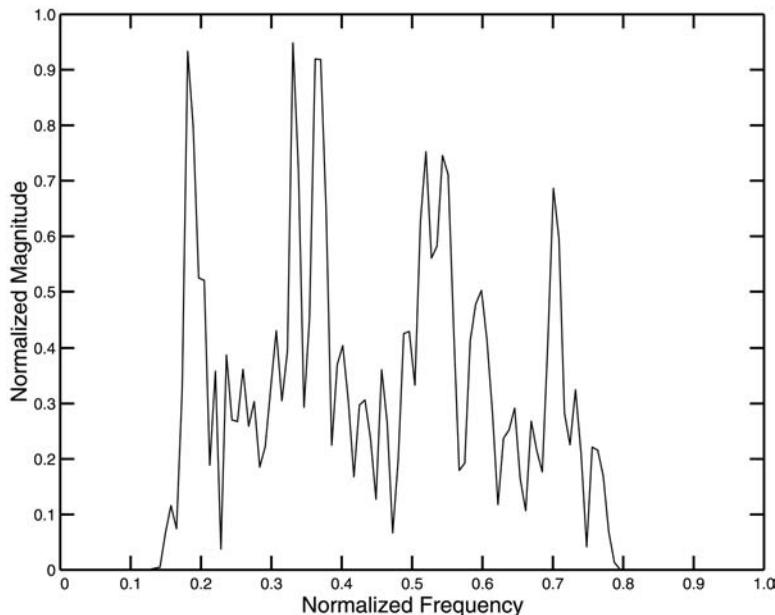


FIGURE 4.1

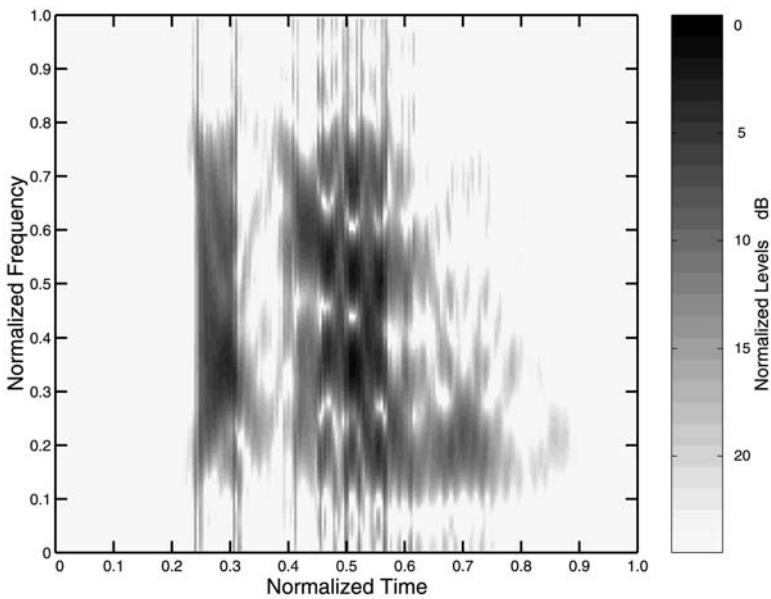
Scattering response of cylindrical shell.

**FIGURE 4.2**

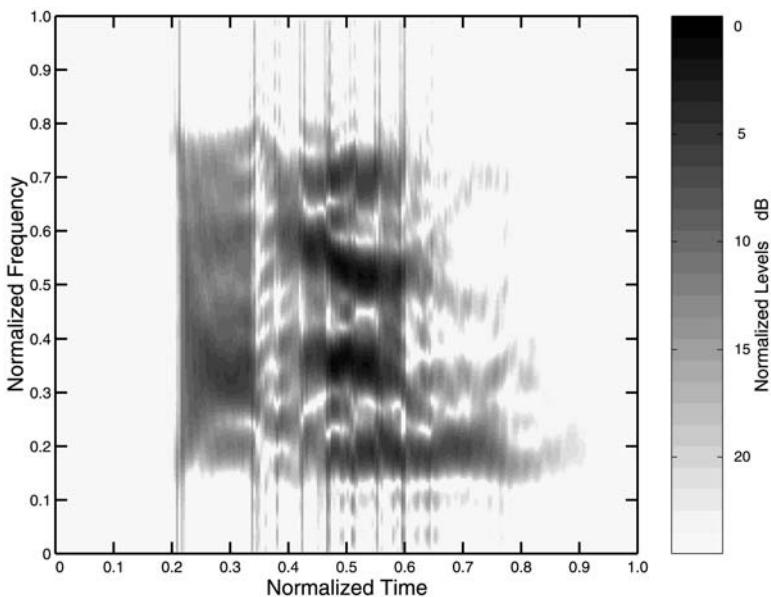
Normalized energy density spectrum for scattering response.

units (NTU) and the resonant portion of the response is in the time period from 0.4 to 1.0 NTU. Figure 4.2 presents the normalized energy density spectrum for the scattering response. It shows the energy bands for the entire signal and where in frequency the energy is concentrated. However, where in time the energy is concentrated cannot be obtained by the energy density spectrum. This information is important when analyzing the acoustic scattering response of a structure such as an elastic cylinder.

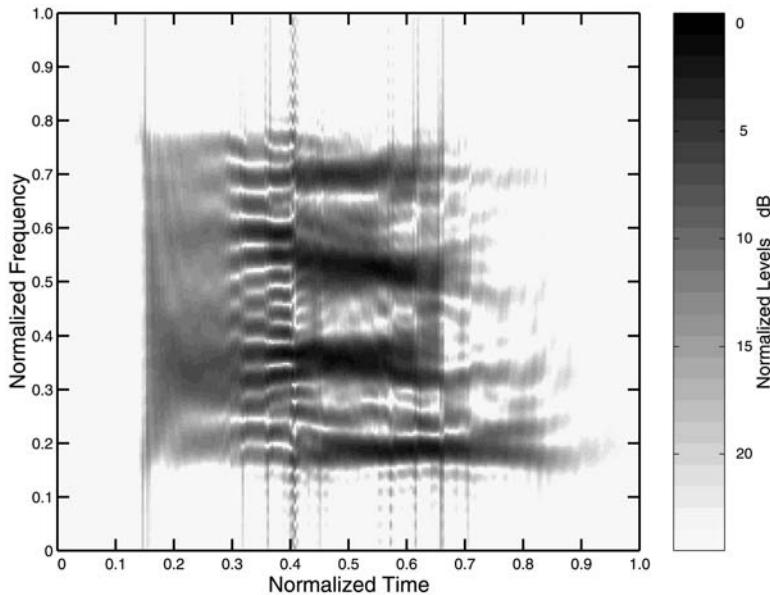
Figures 4.3 through 4.5 present spectrograms of the scattering response signal for window lengths of 16, 32, and 64 data points, respectively (the signal length is 256 samples). The fast Fourier transform (FFT) size for all these spectrograms is zero padded to 512 points. Each spectrogram presents a different view of the joint time-frequency energy distribution, each smearing energy in a different manner over time and frequency, depending on the window size. The spectrogram in Figure 4.3 clearly separates in time the two responses and illustrates that the bulk of the energy associated with the specular response occurs before 0.3 NTU. It also indicates that the resonant response contains roughly three regions of distinct energy concentrations, namely, between 0.4 and 0.45 NTU, between 0.45 and 0.5 NTU, and between 0.5 and 0.55 NTU. Furthermore, the resonant response appears to build in strength with the highest concentrations of energy released by the structure after 0.5 NTU and then tapering off as time proceeds. The resonant portion of the response represents the energy that the structure stores and releases as a modal response. Because the broadband spectrogram smears energy in frequency, resolution in frequency is very poor.

**FIGURE 4.3**

Spectrogram of scattering response signal for window length of 16 data points.

**FIGURE 4.4**

Spectrogram of scattering response signal for window length of 32 data points.

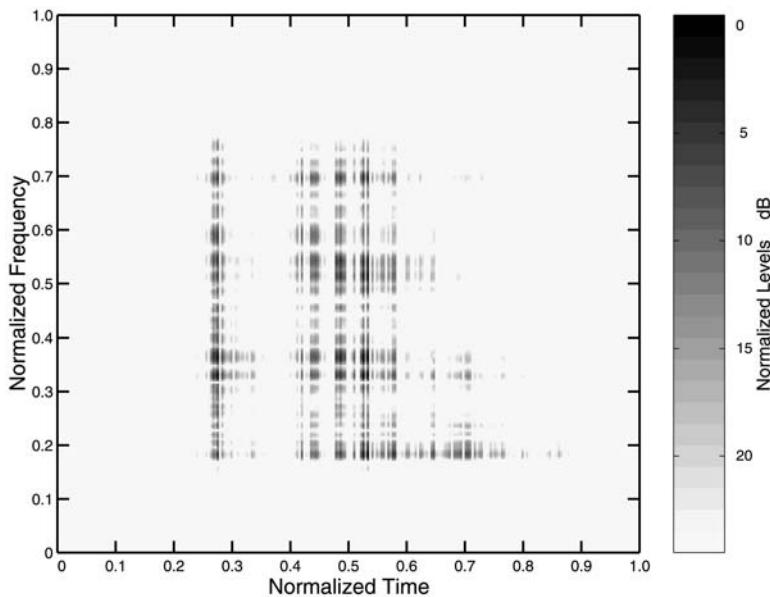
**FIGURE 4.5**

Spectrogram of scattering response signal for window length of 64 data points.

As the window length increases, time resolution degrades, as a consequence of the time–frequency resolution trade-off inherent to the spectrogram. In the spectrogram of Figure 4.4, frequency resolution is higher than that in Figure 4.3, and shows clearly that a number of distinct modes are present in the resonant portion of the response. These occur in normalized frequency, roughly between 0.2 and 0.3, between 0.3 and 0.4, between 0.5 and 0.6, and the last between 0.6 and 0.7. Also note that it appears that bands of energy also show up in the specular portion of the response with one between normalized frequency 0.3 and 0.4, and one between 0.5 and 0.6. As the window length gets longer (Figure 4.5), the frequency resolution increases; however, the resolution in time simultaneously decreases, and it is becoming more difficult to ascertain that two distinct temporal parts of the signal are present, let alone where in time they occur.

The narrowband (long window) spectrogram in Figure 4.5 reveals six strong modes and possibly even more of a weaker nature. This spectrogram, however, has the worst characteristic for smearing energy across time. Also note that the spectrogram generates what appears to be interference during the time period between the specular and resonant portions of the signal. This is clearly an artifact and is a result of the spectrogram putting energy in an incorrect portion of the time–frequency plane, owing to the windowing operation inherent to this technique for time–frequency analysis.

Figure 4.6 shows the minimum cross-entropy positive TFD [19] constructed by an efficient algorithm [10]. This TFD illustrates that clearly two distinct portions of the response exist, one that occurs between 0.25 and 0.3 NTU and the other starting at just after 0.4 NTU and continuing until 0.9 NTU. Simultaneously, the modes of the

**FIGURE 4.6**

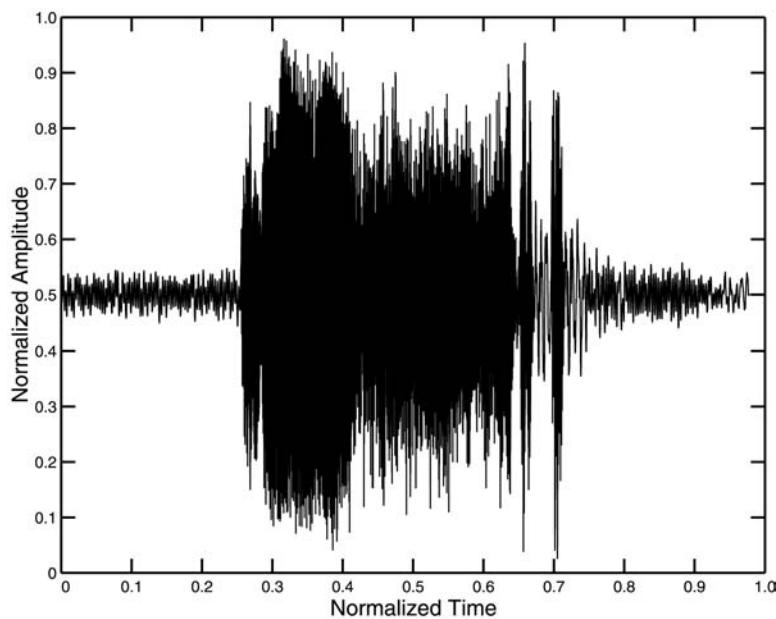
Minimum cross-entropy positive TFD.

resonant response are clearly evident. The energy builds up in the resonant response until it reaches a maximum around 0.53 NTU and then dies down. Another way of viewing this phenomenon is that the stored energy in the structure is released in small amounts at first and then as time progresses, larger amounts are released until the maximum amount of energy is released at approximately 0.53 NTU. Following the maximum release, a “die down” in the release of energy by the structure persists until about 0.9 NTU. Also distinct bands of energy occur over different frequency regions in the specular response.

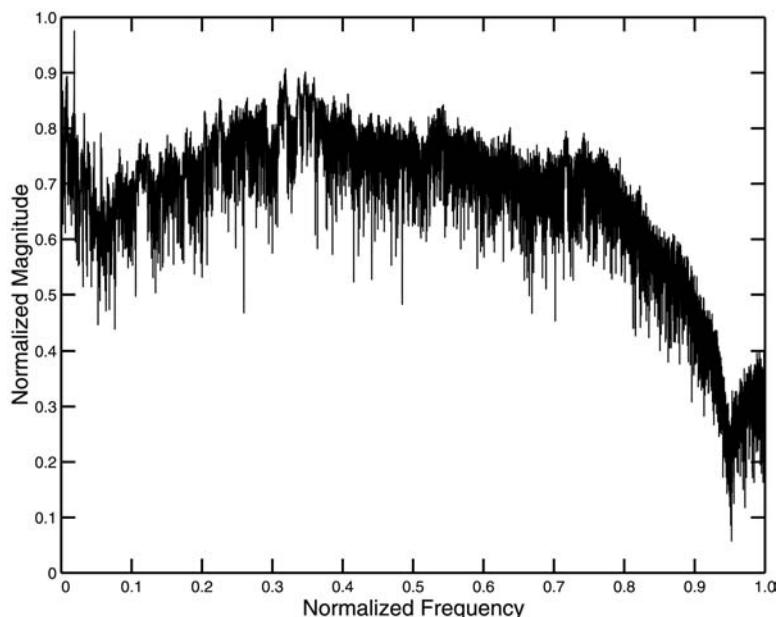
4.3.2 Example 2: underwater vehicle

This example provides analysis of actual data for an acoustic transient signal from an underwater vehicle. The normalized time series for the transient signal is presented in [Figure 4.7](#). Inspection of this figure indicates that the signal contains at least five distinct events in time. The first is of short duration and is at the commencement of the transient signal, which starts at approximately 0.25 NTU and ends at approximately 0.275 NTU. The second event starts at approximately 0.275 NTU and ends at 0.41 NTU. The third event starts at 0.41 NTU and ends at 0.625 NTU. The fourth and fifth events are of short duration, the fourth starting at approximately 0.625 NTU and ending at approximately 0.66 NTU and the fifth event starting at approximately 0.7 NTU and ending at approximately 0.72 NTU.

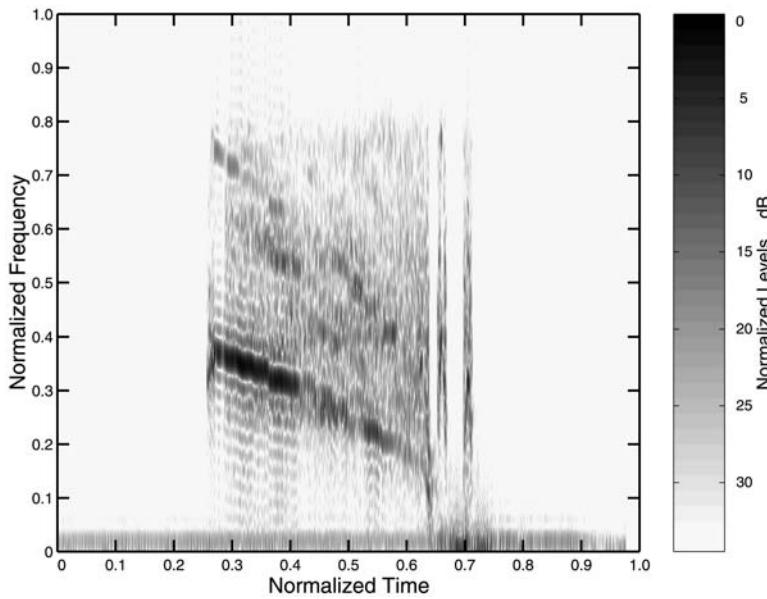
[Figure 4.8](#) shows the normalized energy density spectrum for the acoustic transient signal from the underwater vehicle. It does not reveal any information as to when in time particular elements of the spectral energy occur.

**FIGURE 4.7**

Normalized time series for transient signal.

**FIGURE 4.8**

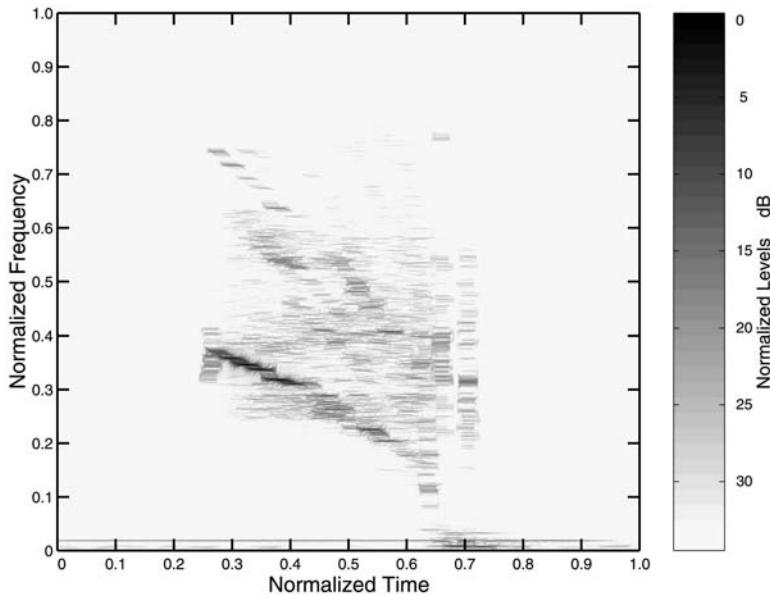
Normalized energy density spectrum for acoustic transient signal from an underwater vehicle.

**FIGURE 4.9**

Plot of a broadband (short window) spectrogram of transient signal time series.

Figure 4.9 is a plot of a broadband (short window) spectrogram of the transient signal time series. This spectrogram was generated using a window of length 64 data points, with a fast Fourier transform (FFT) zeropadded to 2048 points. The five events of the signal are clearly illustrated in Figure 4.9. One interesting feature is what appears to be a linear down-sweep frequency modulation starting at the beginning of the transient signal at approximately 0.39 NFU and sweeping down to approximately 0.3 NFU at the end of the second time event which ends at approximately 0.41 NTU. Also harmonically related down-sweep features appear at higher frequencies. This spectrogram provides relatively high resolution in time, but frequency resolution is poor with energy smeared across the entire band up to the Nyquist frequency.

Figure 4.10 is a narrowband (long window) spectrogram of this signal, generated using a window of size 1024 data points (zeropadded to a 2048-point FFT). Note that the start and end times of the five distinct events of the signal are smeared in time, indicative of the poor time resolution of the narrow band spectrogram. This spectrogram does a better job of identifying the frequency components of this transient signal. However, the down-sweep frequency modulation discussed earlier appears now to have a “twisted rope” structure to it. This marked difference in the structure of this frequency modulation (FM) as revealed by the two different spectrograms highlights a difficulty with the spectrogram: different window sizes can give very different representations of the same event, and one cannot always be certain which representation is in some sense more accurate and which is more corrupted by artifacts

**FIGURE 4.10**

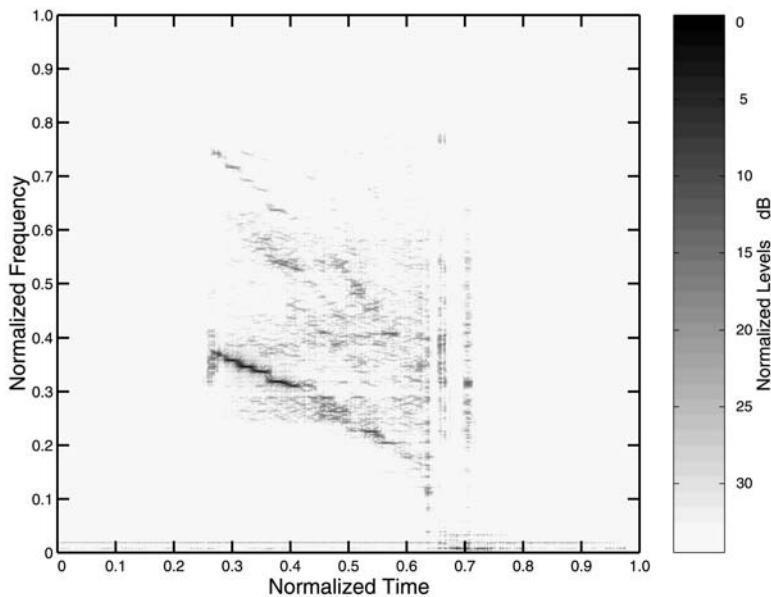
Narrowband spectrogram for acoustic transient signal from underwater vehicle

induced by the windowing operation. In this particular case, the length of the spectrogram window is generating this twisted-rope artifact, as becomes clear when we examine this signal using a positive time–frequency distribution.

In Figure 4.11, we plot a positive distribution, which provides simultaneously high resolution in both time and frequency. Only in the positive distribution of this figure are the discrete frequency steps (which are indeed associated with the physical mechanism that generated the signal) uniquely discernible in the down-sweep frequency modulation feature of the signal (feature in Figure 4.11 between normalized frequencies 0.2 and 0.4). Thus, the twisted-rope feature seen in the previous spectrogram occurs because the relatively long window mixes the different frequencies from consecutive steps.

4.4 Feature Extraction and Classification from Time–Frequency Distributions

In this section we discuss an application of positive TFDs for feature extraction and classification. The basic aim is to classify individual short duration transient signals into respective classes, that is, automatically classify sonar signals as belonging to one of ten classes.

**FIGURE 4.11**

Positive distribution for acoustic transient signal from underwater vehicle.

Groutage and Bennink have developed a method of moments for deriving features from density functions derived from a singular value decomposition (SVD) of energy density TFDs [11]. They have more recently developed a modification of this method, called the transformed SVD (TSVD) method. The transformation rotates the original SVD basis vectors to a new orientation in the span of their vector space. This transformation leads to a new matrix decomposition that minimizes the number of basis vectors required to associate with a given time–frequency energy density concentration. Details of the method are given in [12, 13].

4.4.1 Sonar signal classification

The TSVD method has been applied to classification of acoustic transients signals collected on sonar systems. Approximately 1000 acoustic transients that come from 10 different classes were processed via positive TFDs, from which features were then extracted via the TSVD method. The classifier was developed by researchers at the Naval Underwater Warfare Center in Rhode Island [1].

Table 4.1 presents classification results in the form of a confusion matrix. The diagonal entries of the matrix represent transients that are classified correctly. The off-diagonal data represent misclassification of data. Out of 963 transients, 926 were classified correctly, and 37 misclassified, yielding 96.2% correct classification performance.

TABLE 4.1

Confusion Matrix for Classification of Ten Classes of Acoustic Transient Data Recorded Aboard Submarines

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Class 10
Class 1	346	1	15	1	0	0	0	0	0	0
Class 2	0	73	0	0	0	0	0	0	0	0
Class 3	0	0	35	0	2	0	0	0	1	0
Class 4	0	0	3	81	1	0	0	0	0	0
Class 5	3	0	7	0	95	0	0	0	0	0
Class 6	0	0	0	0	0	9	0	2	0	0
Class 7	0	0	0	0	0	0	39	0	0	0
Class 8	0	0	0	0	0	0	1	28	0	0
Class 9	0	0	0	0	0	0	0	0	207	0
Class 10	0	0	0	0	0	0	0	0	0	13

4.5 Conclusions

In this chapter, we provide a review of the application of positive TFDs to study acoustic echoes and acoustic scattering. As noted, many phenomena in acoustics are frequency dependent and time–frequency analysis offers a powerful approach to understanding the mechanisms involved. In addition, we have shown that effective classification schemes can be developed based on the time–frequency properties.

References

- [1] P. Baggensstoss, Class-specific feature sets in classification, *IEEE Trans. Signal Process.*, 47(12), 3428–3432, 1999.
- [2] L. Cohen, Time-frequency distributions — a review, *IEEE Proc.*, 77, 941, 1989.
- [3] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.
- [4] L. Cohen and Y. Zaparovanny, Positive quantum joint distributions, *J. Math. Phys.*, 21(4), 794–796, 1980.
- [5] L. Cohen and T. Posch, Positive time-frequency distribution functions, *IEEE Trans. Acoust., Speech Signal Process.*, ASSP-33(1), 31–38, 1985.
- [6] L. Dragonette, D. Drumheller, C. Gaumond, D. Hughes, et al., The application of two-dimensional signal transformations to the analysis and synthesis of structural excitations observed in acoustical scattering, in Proceedings of the IEEE, 84(9), 1249–1263, 1996.

- [7] D. Drumheller, D. Hughes, B. O'Connor, and C. Gaumond, Identification and synthesis of acoustic scattering components via the wavelet transform, *J. Acoust. Soc. Am.*, 97(6), 3649–3656, 1995.
- [8] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, 1999.
- [9] G. Gaunaurd and H. Strifors, Signal analysis by means of time-frequency (Wigner-type) distributions — Applications to sonar and radar echoes,in Proceedings of the IEEE, 84(9), 1231–1248, 1996.
- [10] D. Groutage, A fast algorithm for computing minimum cross-entropy positive Time-frequency distributions, *IEEE Trans. Signal Process.*, 45(8), 1954–1970, 1997.
- [11] D. Groutage and D. Bennink, Feature Sets For Non-stationary Signals Derived from Moments of the Singular Value Decomposition of Cohen–Posch (Positive Time-Frequency) Distributions, in *IEEE Trans. Signal Process.*, May 2000.
- [12] D. Groutage and D. Bennink, Principal Features for Non-Stationary Signals Derived from Moments Of The Singular Value Decomposition of Cohen–Posch (Positive Time-Frequency) Distributions, in Proceedings of the SPIE, San Diego, CA, August 2000.
- [13] D. Groutage and D. Bennink, A New Matrix Decomposition Yields Optimum Principal Features for Time-Frequency Distributions, in Proceedings of the IEEE 10th SP Workshop on Statistical Signal and Array Processing, August 2000.
- [14] D. Hughes, Backscattering of Sound by Spherical Shells in Water, Ph.D. dissertation, Washington State University, Pullman, WA, 1992.
- [15] D. Hughes and P. Marston, Local temporal variance of Wigner's distribution function as a spectroscopic observable: lamb wave resonances of a spherical shell, *J. Acoust. Soc. Am.*, 94(1), 499–505, 1993.
- [16] D. Hughes, C. Gaumond, L. Dragonette, and B. Houston, Synthesized wave packet basis for monostatic scattering from a randomly ribbed, finite cylindrical shell, *J. Acoust. Soc. Am.*, 97(3), 1399–1408, 1995.
- [17] D. Hughes and L. Cohen, Instantaneous bandwidth and local duration in acoustic scattering, *Appl. Signal Process.*, 3(2), 68–77, 1996.
- [18] L. Kinsler, A. Frey, A. Coppens, and J. Sanders, *Fundamentals of Acoustics*, John Wiley & Sons, New York, 1982.
- [19] P. Loughlin, J. Pitton, and L. Atlas, Construction of positive time-frequency distributions, *IEEE Trans. Signal Process.*, 42(10), 2697–2705, 1994.
- [20] W. Prosser, M. Seale, and B. Smith, Time-frequency analysis of the dispersion of Lamb modes, *J. Acoust. Soc. Am.*, 105(5), 2669–2676, 1999.

- [21] C.E. Rosenthal and L. Cohen, Time-Frequency Analysis of Acoustic Scattering, in Proceedings of the 24th Asilomar Conference on Signals, Systems and Computers, 1991, pp. 1061–1067.
- [22] J. Sessarego, J. Sageloli, P. Flandrin, and M. Zakharia, Time-frequency analysis of signals related to scattering problems in acoustics, Part I Wigner–Ville analysis of echoes scattered by a spherical shell, in *Wavelets, Time-Frequency Methods, and Phase Space*, J. Combes, A. Grossman, and P. Tchamitchian, Eds., Springer-Verlag, Berlin, 1989, pp. 147–153,
- [23] N. Yen, Time-frequency representation of acoustic signals by means of Wigner distribution function: Implementation and interpretation, *J. Acoust. Soc. Am.*, 81(6), 1841–1850, 1987.
- [24] N. Yen, L. Dragonette, and S. Numrich, Time-frequency analysis of acoustic scattering from elastic objects, *J. Acoust. Soc. Am.*, 87(6), 2359–2370, 1990.

5

Time–Frequency Reassignment: From Principles to Algorithms

P. Flandrin*

Ecole Normale Supérieure de Lyon

F. Auger*

Université de Nantes

E. Chassande-Mottin*

Observatoire de la Côte d'Azur

CONTENTS

5.1	Introduction	179
5.2	Reassigned Time–Frequency Distributions	180
5.3	Reassignment in Action	190
5.4	Real Case Studies	199
5.5	Conclusions	201
	References	201

5.1 Introduction

Time–frequency analysis (TF) is a field that has experienced a number of qualitative and quantitative changes during the last two decades. Whereas most of classical signal processing studies of the 1970s were aimed at stationary signals and processes, many efforts were devoted to less idealized situations during the 1980s, and the idea of TF progressively emerged as a new paradigm for nonstationarity. It is now well recognized that many signal processing problems can be advantageously phrased in a TF language, and the issue may no longer be designing brand new methods from scratch, but instead in adequately using some of the many tools that we have at our disposal, or in improving them for specific tasks. In some sense, the purpose of this chapter has to be understood from this second generation perspective, because what is discussed here essentially builds on the methods that have already been extensively

*All authors are also members of the GdR-PRC *ISIS* of the CNRS.

studied and used. New advances nevertheless are to be provided, thanks to fresh interpretations that have been made possible by recent developments in TF analysis.

This chapter is devoted to *reassignment* a technique that was first introduced in the 1970s in a somehow restricted framework, with a scope that has been substantially enlarged, thanks to the new developments that modern TF analysis has experienced. If the concerns are what reassignment is and what it is good for, the explanation is at least twofold. First, reassignment can be viewed as a postprocessing technique aimed at improving readability of TF distributions (*exploratory signal analysis*). Second, reassignment can be used as a way of coding useful TF information, so as to be part of decision schemes (*signal processing*). We concentrate mainly on the first aspect, referring the interested reader to more comprehensive treatments concerning the second one [11].

More precisely, the chapter is organized as follows. Section 5.2 first motivates the usefulness of reassignment by stressing how it permits to overcome the localization and interference trade-off that is usually observed in classical TF analysis. The reassignment principle then is to be detailed in the simplest case of the spectrogram, and some examples illustrate different facets of the technique. Section 5.3 focuses on reassignment in action: starting from the spectrogram case, efficient algorithmic issues are discussed, as well as extensions to more general situations, including time-scale distributions such as the scalogram (squared wavelet transform). Finally, Section 5.4 points out a number of real-world situations where reassignment may be of effective usefulness, both in exploratory data analysis and in signal manipulation.

The point of view adopted in this chapter is mostly practical, with as little theory as needed (for more fundamental aspects, the interested reader is referred to [2] or [11]). Throughout the text, and for a sake of illustration, an extensive use is made of MATLAB routines that are part of a freeware toolbox [3], downloadable from:

<http://iut-saint-nazaire.univ-nantes.fr/~auger/tfib.html>

The specific procedures used are available from the universal resource locator (URL):

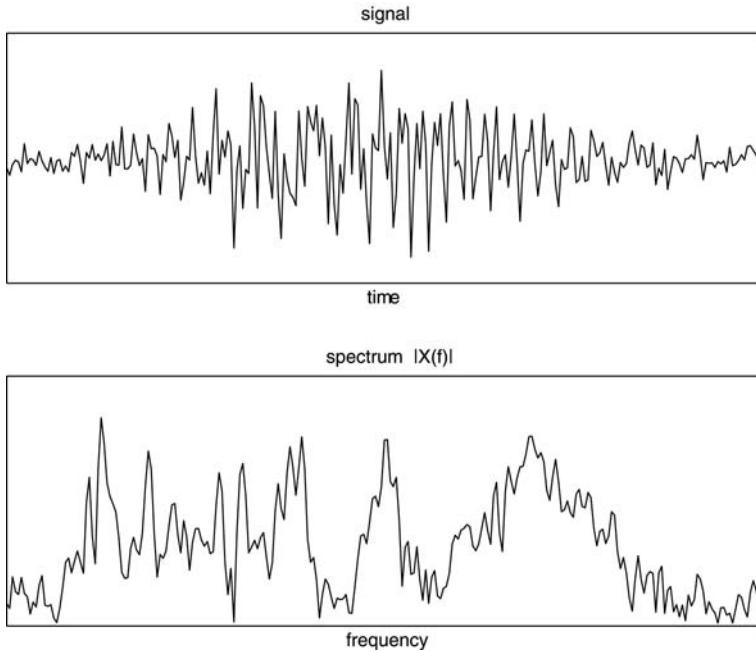
<http://iut-saint-nazaire.univ-nantes.fr/~auger/publis/CRC.html>

and they must be explicitly considered as part of the chapter, because they allow the readers not only to reproduce the figures used as illustrations but also to make their own variations on their production, so as to become more familiar with the proposed tools.

5.2 Reassigned Time–Frequency Distributions

5.2.1 Motivation example

TF tools are extensively used for exploratory signal analysis. To address some of the problems that conventional TF tools are faced with, let us first consider the example of [Figure 5.1](#), produced by running **model1plot**.

**FIGURE 5.1**

A three-component signal embedded in noise. From a point of view of exploratory data analysis, neither the waveform (top) nor its spectrum (bottom) adequately reveals the actual structure of the analyzed signal — Equation (5.1).

In fact, looking at the waveform or at its spectrum does not allow for a simple interpretation of the inner structure of the signal \mathbf{x} , which consists in a linear combination of 3 different amplitude-modulated–frequency-modulated (AM–FM) components embedded in noise with a 20-dB signal-to-noise ratio (SNR), generated according to:

$$N = 256; \text{fmins}=0.05; \text{fmaxs}=0.2; \quad (5.1)$$

`[x, fxs, fxl, dt] = model1(N, fmins, fmaxs, 0.45, 0.25, N/3, 0.25, 20);`

In such a situation, a much clearer insight would be gained by an explicit description of the time-frequency structure of each of the components. The relevance of such a mixed description is supported by [Figure 5.2](#), obtained by running **model1TF**. Indeed, the simplified model displayed in the top left subplot of Figure 5.2 makes apparent the coexistence of two FM components (one sinusoidal and one linear, of respective instantaneous frequencies fxs and fxl) and one logon (Gaussian wave packet of effective duration dt). The purpose of exploratory TF analysis is therefore to produce a picture as close to this idealized model as possible, given the observed three-component signal, of *a priori* unknown structure.

Because of its many theoretical properties [16], the Wigner–Ville distribution (WVD) could be thought of as the most appropriate tool, but it appears (top right

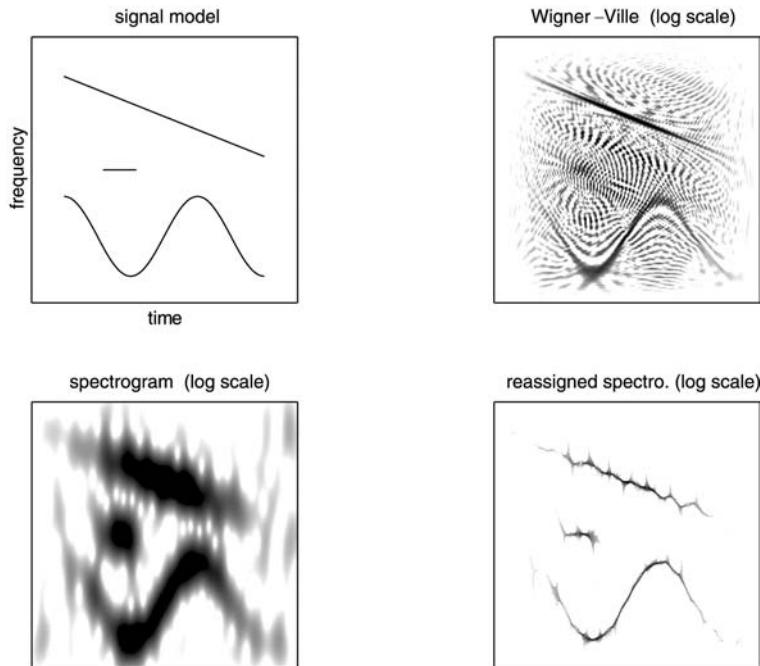


FIGURE 5.2

A three-component signal embedded in noise. The idealized time–frequency model is compared with three images produced by the Wigner–Ville distribution (WVD), a spectrogram, and a reassigned spectrogram. Whereas the readability of the WVD (top right) is hampered by oscillatory interference terms and the spectrogram (bottom left) suffers from a poor resolution, the reassigned spectrogram (bottom right) provides the user with a time–frequency picture which is almost identical to the idealized model (top left).

of Figure 5.2) that, whereas individual components are rather sharply described, the overall readability is hampered by cross-components interference phenomena [23]. As is well-known, a spectrogram (bottom left of Figure 5.2) contains much less crossterms, but this image cleaning is obtained at the expense of a smearing of the TF signatures of the individual components. Overcoming this trade-off between localization and interference is precisely one of the “*raisons d’être*” of reassignment, whose application results in a picture (bottom right of Figure 5.2) that is almost identical to the idealized model we started with.

5.2.2 Reassignment basics

To understand how the miracle of Figure 5.2 occurs, it is worth going back to basic definitions and properties of classical TF distributions.

5.2.2.1 From spectrograms to Wigner–Ville

First, a spectrogram $S_x^h(t, \omega)$ is usually defined as:

$$S_x^h(t, \omega) := |F_x^h(t, \omega)|^2 \quad (5.2)$$

where $F_x^h(t, \omega)$ stands for the short-time Fourier transform (STFT):

$$F_x^h(t, \omega) := \int_{-\infty}^{+\infty} x(s) h^*(s-t) e^{-i\omega s} ds \times e^{it\omega/2} \quad (5.3)$$

Those definitions are explicitly dependent on some short-time window $h(t)$, aimed at limiting the evaluation of the Fourier transform (FT) to some specified neighborhood of the current date t . The window introduced this way may be thought of as a measurement device, with the consequence that the produced distribution jointly depends on the signal and on the window. In particular, STFTs and spectrograms are faced with a TF resolution trade-off, because the shorter the duration of $h(t)$, the better their resolution in time, but the larger the bandwidth of the spectrum $H(\omega)$ and, henceforth, the poorer their resolution in frequency.

Instead of fixing arbitrarily $h(t)$, one may try to make it depend adaptively on the analyzed signal. Thinking of the STFT as a linear filtering operation, the powerful concept of matched filtering suggests the intuitive choice $h(t) = x_-(t) := x(-t)$ (i.e., to take as window the time-reversed version of the analyzed signal). It then follows that:

$$F_x^{x_-}(t, \omega) = W_x(t/2, \omega/2) / 2 \quad (5.4)$$

where:

$$W_x(t, \omega) := \int_{-\infty}^{+\infty} x(t+s/2) x^*(t-s/2) e^{-is\omega} ds \quad (5.5)$$

is nothing but the WVD [16].

By definition, the WVD does present the advantage of depending only on the signal. Its role is central in TF analysis and it possesses a number of theoretical properties [12, 16] among which one can mention its ability to be perfectly localized in the case of linear FM signals. These appealing features have, however, to be paid at some price because, for example, a WVD cannot be positive everywhere, thus forbidding a local density interpretation. Furthermore, in accordance with the elementary identity $(a+b)^2 = a^2 + b^2 + 2ab$, the fully quadratic nature of the WVD is known to create spurious cross terms, characterized by oscillating contributions located midway in between any two interacting components [23].

5.2.2.2 Localization vs. interference

Because interference terms of the WVD are by nature oscillatory, one can think of reducing them by applying some low-pass smoothing, but such a smoothing operation has also the negative effect of spreading out localized signal terms. This trade-off between joint localization and the importance of interference terms is clearly apparent when comparing (as in [Figure 5.2](#)) the WVD and a spectrogram, properly interpreted

as a smoothed WVD. Indeed, instead of using its usual definition, as in Equation (5.2), a spectrogram can be equivalently expressed as [16]:

$$S_x^h(t, \omega) = \iint_{-\infty}^{+\infty} W_x(s, \xi) W_h(s-t, \xi-\omega) \frac{ds d\xi}{2\pi} \quad (5.6)$$

thus making explicit the fact that it results from the smoothing of the signal WVD by the window WVD.

5.2.2.3 Reassignment principle

An important consequence of the smoothing formula in Equation (5.6) is that the value that a spectrogram takes at a given point (t, ω) of the plane cannot be considered as pointwise, but instead results from the summation of a continuum of WVD contributions within some TF domain defined as the essential TF support of the short-time window. In other words, a whole distribution of values is summarized by a single number, and this number is assigned to the geometric center of the domain over which the distribution is considered.

By reasoning with a mechanical analogy, the situation is as if the total mass of an object were assigned to its geometric center, an arbitrary point that — except in the very specific case of an homogeneous distribution over the domain — has no reason to suit the actual distribution.

A much more meaningful choice is to assign the total mass to the center of gravity of the distribution within the domain, and this is precisely what reassignment does: at each TF point (t, ω) where a spectrogram value is computed, one also computes the two quantities:

$$\hat{t}_x(t, \omega) := \frac{1}{S_x^h(t, \omega)} \iint_{-\infty}^{+\infty} s W_x(s, \xi) W_h(s-t, \xi-\omega) \frac{ds d\xi}{2\pi} \quad (5.7)$$

$$\hat{\omega}_x(t, \omega) := \frac{1}{S_x^h(t, \omega)} \iint_{-\infty}^{+\infty} \xi W_x(s, \xi) W_h(s-t, \xi-\omega) \frac{ds d\xi}{2\pi} \quad (5.8)$$

which define the local centroids of the WVD distribution W_x , as seen through the TF window W_h centered in (t, ω) . The spectrogram value is then moved from the point (t, ω) where it has been computed to this centroid $(\hat{t}_x(t, \omega), \hat{\omega}_x(t, \omega))$, leading to define the reassigned spectrogram as:

$$\check{S}_x^h(t, \omega) := \iint_{-\infty}^{+\infty} S_x^h(s, \xi) \delta(t - \hat{t}_x(s, \xi), \omega - \hat{\omega}_x(s, \xi)) \frac{ds d\xi}{2\pi} \quad (5.9)$$

Conceptually, reassignment can be considered a two-step process: (1) a smoothing, whose main purpose is to rub out oscillatory interferences, but whose drawback is to smear localized components; (2) a squeezing, whose effect is to refocus the contributions that survived the smoothing.

5.2.2.4 Some historical comments

As has been sketched earlier, the presentation of the reassignment principle is a modern way [2] of revisiting a much older idea [25, 27].

The argument used by Kodera, De Villedary and Gendrin in their seminal papers [25, 27] was not related to TF smoothing, but to phase. In fact, the centroids as in Equations (5.7) and (5.8), used in reassignment happen to be related to the phase of the STFT, an information that is discarded when considering a spectrogram as a squared STFT. More precisely, denoting by $\varphi(t, \omega)$ the phase of the STFT, as in Equation (5.3), and using the simplified notation $\partial_u \varphi = \partial \varphi / \partial u$, local centroids, as in Equations (5.7) and (5.8) can be shown [25, 27] to be equivalently given by:

$$\hat{t}_x(t, \omega) = \frac{t}{2} - \partial_\omega \varphi(t, \omega) \quad (5.10)$$

$$\hat{\omega}_x(t, \omega) = \frac{\omega}{2} + \partial_t \varphi(t, \omega) \quad (5.11)$$

These quantities can be interpreted as the local instantaneous frequency (IF) and group delay (GD) of the analyzed signal, as filtered within the TF domain defined by the TF window W_h centered in (t, ω) . From this interpretation, the rationale for reassignment (initially referred to as a “modified moving window method” [25, 27]) was to favor energy concentrations in the vicinity of local IFs and GDs.

Although introduced in the mid-1970s and applied with success in geophysics, the reassignment technique retained almost no attention of the TF community the next 20 years. Apart from the fact that the methodology has been little publicized per se, the reason is certainly that it had been introduced much before the problems it addressed were extensively studied (during the mid-1980s), and that its applicability remained computationally limited for a while.

Not until the mid-1990s did the idea of reassignment surface again [2]. Algorithmic improvements were proposed and, thanks to the many developments that occurred in the field, the scope of the technique was considerably enlarged, far beyond only the spectrogram case. In parallel, other related techniques were developed independently (e.g., the “ridge and skeleton” method [5, 13, 19, 20], the “instantaneous frequency density” [17], the “differential spectral analysis” [18] or the “synchrosqueezing” technique [29]).

5.2.3 Toy examples

5.2.3.1 Linear chirps

It is well known [16] that an idealized linear FM signal has a WVD that is perfectly localized in the TF plane, namely:

$$x(t) = \exp\{\imath(\omega_0 t + \beta t^2/2)\} \Rightarrow W_x(t, \omega) = \delta(\omega - (\omega_0 + \beta t)) \quad (5.12)$$

for any modulation rate β .

Therefore, it may be concluded that reassigned spectrograms automatically inherit this perfect localization, because the centroid of a line distribution necessarily belongs

to its (localized) support. It is worth noting that, geometrically, TF lines include pure tones ($\beta = 0$) and impulses ($\beta = \infty$) as limiting cases. In all cases, any reassigned spectrogram (i.e., whatever the window $h(t)$ it is computed with) is theoretically guaranteed to be perfectly localized.

If we turn to the more realistic situation of a linear chirp, defined as:

$$x(t) = \exp\{-\pi\gamma t^2\} \times \exp\{\imath(\omega_0 t + \beta t^2/2)\} \quad (5.13)$$

explicit calculations can be conducted, but they are not reproduced in this chapter (the interested reader is referred to [27] or [11]). We instead investigate through examples how a reassigned spectrogram jointly depends on the analyzed chirp and on the short-time window. The proposed numerical experiment consists in generating an **Nx**-points chirp **x**, sweeping from **fmin** to **fmax**, embedded in an **-points signal **y**:**

```
x = gchirp(Nx,fmin,fmax); Lx = length(x);
y = zeros(Ny,1); y(Ny/2-Lx/2:Ny/2+Lx/2-1) = x;
```

By running:

```
sprspplot(y, dt, Nb, Nh, c);
```

this signal **y** is then analyzed by a spectrogram (with a Gaussian window of length **Nh**) and its reassigned version, with both TF distributions computed every **dt** time samples on a frequency grid of **Nb** bins, and displayed either in color ($c = 1$) or in black and white ($c = 0$). A typical example corresponding to the set of parameters:

```
Nx = 16; fmin = 0; fmax = 0.5; Ny = 256; % signal
Nh = 255; dt = 1; Nb = 128; c = 0; % TFD and display
```

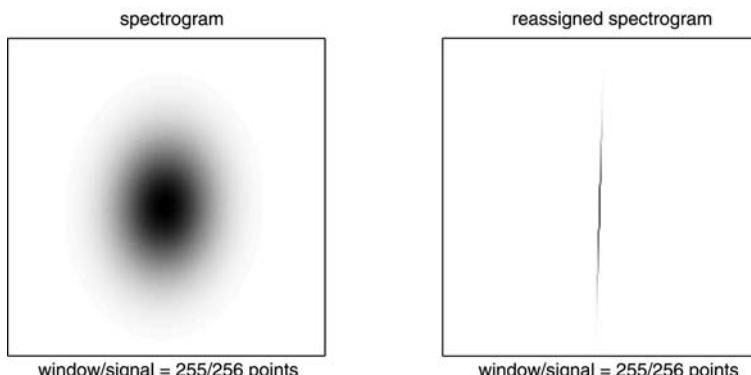
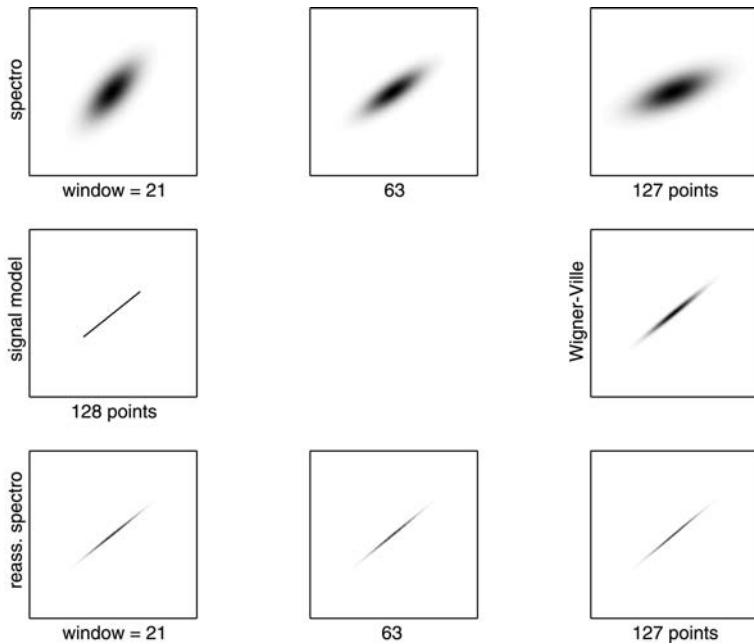


FIGURE 5.3

Reassigning a short chirp. In the case of a very short transient signal (the effective support of the chirp **x** considered here is **Nx = 16** points, for an observation **y** of total length **Ny = 256** points) analyzed with a substantially larger window **h** (**Nh = 255** points), the TF localization is dramatically increased by reassignment.

**FIGURE 5.4**

Chirp rate vs. analysis window. When analyzing a chirp signal of fixed chirp rate, the result given by conventional spectrograms (top row) heavily depends on the length of the short-time window. This interdependence between signal and window characteristics is dramatically reduced when replacing spectrograms by their reassigned counterparts (bottom row), leading to sharply localized distributions approaching the idealized model sketched in the left subplot of the middle row. For a sake of comparison, the WVD is displayed in the right subplot of the middle row.

is given in [Figure 5.3](#). It clearly evidences the drastic squeezing effect of reassignment in a situation of a very transient chirp ($N_x = 16$) that cannot be considered as quasimonochromatic with respect to the length ($N_h = 255$) of the analyzing window. Variations on this example are left to the reader.

A more comprehensive comparison is given in Figure 5.4, obtained by running:

```
varwindow(Nx, Nh1, Nh2, Nh3, Nb, c);
```

with the specific values:

```

Nx = 128; % signal
Nh1 = 21; Nh2 = 63; Nh3 = 127; % windows
Nb = 128; c = 0; % TFD and display

```

Whereas, in such a case, conventional spectrograms happen to be poorly localized and to depend heavily on the length of the chosen short-time window, their reassigned counterparts prove to always provide fairly similar, correctly localized results (the skeletonized model and the WVD are also given for comparison).

5.2.3.2 Nonlinear FM signals

It has been argued in Section 5.2.3.1 that the reassignment process ends up with a perfect localization in the case of linear FM signals. Conceptually, a similar benefit is expected to be obtained as long as the analyzed signal locally behaves as a linear FM, locality referring to the TF support of the smoothing window. This interdependence between the window length and the local modulation rate can be dynamically observed by running:

```
quasilinear(Nx, Nh, Nb, Nhp, c);
```

where **Nx** stands for the length of an **Nhp** half-periods sinusoidal FM signal, and **Nh** stands for the length of the short-time window used in computing its reassigned spectrogram over **Nb** frequency bins.

Setting for instance the parameter values to:

```
Nx = 256; Nhp = 21; % signal  
Nh = 19; % window  
Nb = 128; c = 0; % TFD and display
```

leads, at time **t** = **ceil(15*Nx/21)**, to the snapshot given in Figure 5.5. What is evidenced by this image is that (slight) departures from localization are observed

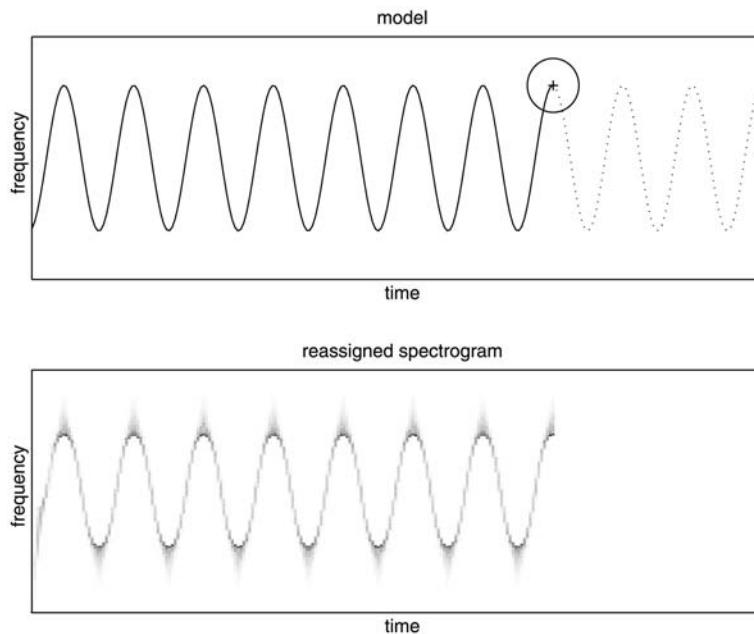


FIGURE 5.5

Local quasilinearity and localization. Due to the local action of reassignment, the perfect localization property that holds for linear FM signals carries over to locally quasilinear situations. In this example of a sinusoidal FM signal, localization is (slightly) degraded at those points where the instantaneous frequency trajectory cannot be considered as quasilinear within the TF smoothing window (represented by the ellipse).

only in those regions where the local IF trajectory of the model cannot be considered as quasilinear within the TF smoothing window centered at the point of interest.

5.2.3.3 Localization vs. resolution

Whereas reassigning a spectrogram has been shown to end up with a sharply localized distribution, care has to be taken in interpreting this behavior in terms of resolution (i.e., in the ability of separating closely spaced components). In fact, when more than one component is seen within the TF smoothing window, a beating effect occurs and results in interference fringes, thus preventing the consideration of reassignment as some superresolution process. This can be illustrated by running:

```
resol(Nx, Nh, Nb, c);
```

in which a series of crossing linear FM signals of length **Nx** and of various chirp rates is analyzed over **Nb** frequency bins, with the same short-time window of length **Nh**. Choosing for instance:

```
Nx = 128; % signal
Nh = 31; % window
Nb = 128; c = 0; % TFD and display
```

leads to the result displayed in Figure 5.6. Interference effects (emphasized by using a logarithmic scale for the amplitude) are clearly apparent in those regions where the two chirps are simultaneously present within the TF smoothing window.

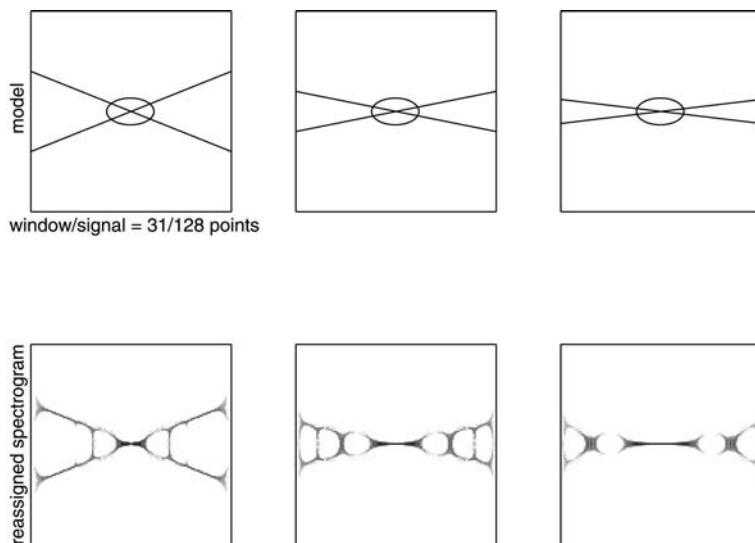


FIGURE 5.6

Localization and resolution. When more than one component is seen within the TF smoothing window (represented by the ellipse), a beating effect occurs and results in interference fringes. This behavior is illustrated here by two crossing linear FM signals of various chirp rates.

A similar effect can be dynamically observed by visualizing the animation **model1TFmovie.gif**, in which a signal of the **model1** type is considered, with a logon whose center follows a TF trajectory crossing (back and forth) the linear FM component.

5.3 Reassignment in Action

In Section 5.2.3, we show that reassignment yields significantly improved TF representations. The present section, devoted to computer algorithms, shows that the computational cost paid for this improvement is not as heavy, and does not prevent the use of this method in practical cases.

5.3.1 Algorithms for spectrograms

5.3.1.1 Reassignment operators deduced from phase differences

As mentioned in Section 5.2.2.4, the reassignment operators $\hat{t}_x(t, \omega)$ and $\hat{\omega}_x(t, \omega)$ have been initially deduced from the phase of the STFT. These expressions have been used in the first implementations of the reassignment principle [26]. For a signal sampled with a sampling period T_s and a frequency axis obtained with a fast Fourier transform (FFT) of length N , the partial derivatives of the phase were replaced by first-order differences:

$$\partial_t \varphi[n, m] \approx \frac{\varphi[n+1, m] - \varphi[n-1, m]}{2T_s} \quad (5.14)$$

$$\approx \frac{1}{2T_s} \arg(F_x^h[n+1, m] F_x^h[n-1, m]^*) \quad (5.15)$$

$$\partial_\omega \varphi[n, m] \approx \frac{\varphi[n, m+1] - \varphi[n, m-1]}{(4\pi/NT_s)} \quad (5.16)$$

$$\approx \frac{NT_s}{4\pi} \arg(F_x^h[n, m+1] F_x^h[n, m-1]^*) \quad (5.17)$$

where $F_x^h[n, m]$ and $\varphi[n, m]$ are, respectively, the STFT value and its phase at time $t = nT_s$ and angular frequency $\omega = \frac{2\pi m}{NT_s}$. However, this approach suffers from the phase unwrapping problem [33], and yields only approximated values of the reassignment operators.

5.3.1.2 More reliable algorithm

Expressions in Equations (5.15) and (5.17) suggest relating the partial derivatives of the phase to the partial derivatives of the STFT. Defining $G_x^h(t, \omega) = F_x^h(t, \omega) e^{-i\omega t/2} = |F_x^h(t, \omega)| e^{i(\varphi - \omega t/2)}$ leads to:

$$\partial_t G_x^h(t, \omega) = - \int_{-\infty}^{+\infty} x(s) \frac{dh^*}{dt}(s-t) e^{-i\omega s} ds = -G_x^{dh}(t, \omega) \quad (5.18)$$

$$= \partial_t |F_x^h(t, \omega)| e^{i(\varphi - \omega t/2)} + i(\partial_t \varphi - \omega/2) G_x^h(t, \omega) \quad (5.19)$$

$$\Rightarrow \hat{\omega}_x(t, \omega) = \omega - \text{Im} \left\{ \frac{G_x^{dh}(t, \omega)}{G_x^h(t, \omega)} \right\} \quad (5.20)$$

where $\text{Im}\{Z\}$ stands for the imaginary part of the complex number Z , and dh is the derivative of the analysis window $dh(t) = \frac{dh}{dt}(t)$. The partial derivative of the phase with respect to the angular frequency leads to a similar expression for $\hat{t}_x(t, \omega)$:

$$\partial_\omega G_x^h(t, \omega) = -i(G_x^{th}(t, \omega) + t G_x^h(t, \omega)) \quad (5.21)$$

$$= \partial_\omega |F_x^h(t, \omega)| e^{i(\varphi - \omega t/2)} + i(\partial_\omega \varphi - t/2) G_x^h(t, \omega) \quad (5.22)$$

$$\Rightarrow \hat{t}_x(t, \omega) = t + \text{Re} \left\{ \frac{G_x^{th}(t, \omega)}{G_x^h(t, \omega)} \right\} \quad (5.23)$$

with $th(t) = t h(t)$.

Approximated phase differentiations can then be avoided by computing two additional STFTs. These STFTs use the same signal values as $F_x^h(t, \omega)$, and only differ by their analysis windows. As a consequence, these new expressions do not imply a drastic increase of computational complexity, and can be computed in parallel. It should also be noted that if $h(t) = e^{-\alpha t^2}$, a choice that corresponds to the so-called *Gabor spectrogram*, then $\frac{dh}{dt}(t) = -2\alpha th(t)$ and $G_x^{dh}(t, \omega) = -2\alpha G_x^{th}(t, \omega)$. Only one supplementary STFT is required, which makes the reassignment process still faster. This particular case is implemented in the function **tfrrgab** of the TF toolbox [3].

For sampled signals and discrete Fourier transforms (DFTs) of length N , discrete time and discrete frequency versions of the reassignment operators are defined as:

$$\hat{n}_x[n, m] := \frac{1}{T_s} \hat{t}_x \left(nT_s, \frac{2\pi m}{NT_s} \right) = n + \text{Re} \left\{ \frac{G_x^{T_s^{-1}th}[n, m]}{G_x^h[n, m]} \right\} \quad (5.24)$$

$$\hat{m}_x[n, m] := \frac{NT_s}{2\pi} \hat{\omega}_x \left(nT_s, \frac{2\pi m}{NT_s} \right) = m - \frac{N}{2\pi} \text{Im} \left\{ \frac{G_x^{T_s dh}[n, m]}{G_x^h[n, m]} \right\} \quad (5.25)$$

with $T_s dh[n] = T_s \frac{dh}{dt}(nT_s)$ and $T_s^{-1}th[n] = n h(nT_s)$.

The computation of the reassigned spectrogram of a signal \mathbf{x} at the time instants given in a row vector \mathbf{t} , as done in the function **tfrrsp** of the TF toolbox [3], can then be achieved in two steps.

5.3.1.2.1 The first step computes $G_x^h[n, m]$ (**tfr**), $G_x^{T_s^{-1}th}[n, m]$ (**tf2**) and $G_x^{T_s dh}[n, m]$ (**tf3**):

```
[xrow,xcol]=size(x); [trow,tcol]=size(t);
[hrow,hcol]=size(h); Lh=(hrow-1)/2;

% create and initialize three arrays to zero
% these arrays have as many rows as frequency bins,
% and as many columns as time intants
tfr=zeros(N,tcol); tf2=zeros(N,tcol); tf3=zeros(N,tcol);

% compute th and dh
Th=h.*[-Lh:Lh]'; Dh=dwindow(h);

% compute the three STFTs
for icol=1:tcol,
    ti= t(icol);
    tau=-min([round(N/2)-1,Lh,ti-1]): ...
        +min([round(N/2)-1,Lh,xrow-ti]);
    indices= rem(N+tau,N)+1;

    % normalization to preserve energy
    norm_h=norm(h(Lh+1+tau));
    tfr(indices,icol)=x(ti+tau).*conj( h(Lh+1+tau))/norm_h;
    tf2(indices,icol)=x(ti+tau).*conj( Th(Lh+1+tau))/norm_h;
    tf3(indices,icol)=x(ti+tau).*conj( Dh(Lh+1+tau))/norm_h;
end;

% final call to the FFT function
tfr=fft(tfr); tf2=fft(tf2); tf3=fft(tf3);

avoidw=find(tfr~=0);
tf2(avoidw)=round(real(tf2(avoidw)./tfr(avoidw)/Dt));
tf3(avoidw)=round(imag(N*tf3(avoidw)./tfr(avoidw)/(2.0*pi)));
tfr=abs(tfr).^2;
```

The expression of **indices** allows the values of the time lag $\tau = \dots, -2, -1, 0, 1, 2, \dots$ to be located at the indices $\dots, N-1, N, 1, 2, 3, \dots$ as required by the **FFT** function. In addition, when $F_x^h[n, m]$ is zero valued, $S_x^h[n, m]$ is of course also zero valued. Reassignment operators can neither be computed nor be used when $S_x^h[n, m]$ is zero. This explains the use of the variable **avoidw**.

5.3.1.2.2 The second step iteratively creates the reassigned spectrogram by adding $S_x^h[n, m]$, the value of the spectrogram at the point $(nT_s, \frac{2\pi m}{NT_s})$ to the value of $\check{S}_x^h[\hat{n}, \hat{m}]$:

```

rtfr= zeros(N,tcol);
Ex=mean(abs(x(min(t):max(t))).^2); Threshold=1.0e-6*Ex;
for icol=1:tcol, % for all time samples
    for jcol=1:N, % for all frequency bins
        if abs(tfr(jcol,icol))>Threshold,
            icolhat= icol + tf2(jcol,icol);
            icolhat=min(max(icolhat,1),tcol);
            % not smaller than 1, not greater than tcol

            jcolhat=jcol - tf3(jcol,icol);
            jcolhat=rem(rem(jcolhat-1,N)+N,N)+1;
            % must be inside [1,N]

            rtfr(jcolhat,icolhat)= rtfr(jcolhat,icolhat)...
                + tfr(jcol,icol) ;
        end;
    end;
end;

```

It should be emphasized that unlike the first step, which greatly benefits from the vectorization capabilities of MATLAB, this second step is a slower process under the MATLAB environment, because of the two embedded loops. Faster implementations written in the American National Standards Institute (ANSI) C language are also available at:

www-sigproc.eng.cam.ac.uk/~md283/toolbox/Ctfibeng.html

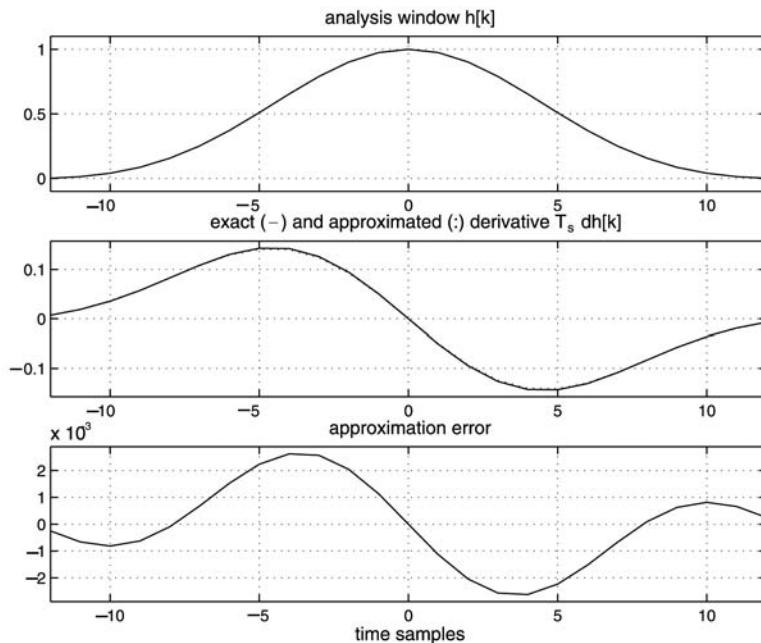
5.3.1.3 Computing window derivative

The first step of this algorithm requires, however, an evaluation of the derivative of the analysis window function $h(t)$. An exact expression of this derivative could be obtained from the analytical expression of $h(t)$. This solution is too restrictive, because it does not allow someone to use self-made windows, whose derivatives may be difficult to get analytically. This is the reason why instead of the exact expression we use an accurate approximation, obtained by a centered first-order derivative. For this, $h(t)$ is modeled on its finite length support $[-T/2, T/2]$ as:

$$h(t) = \begin{cases} \alpha + \beta t + h_0(t) & \text{for } |t| \leq T/2 \\ 0 & \text{for } |t| > T/2 \end{cases} \quad (5.26)$$

where $\alpha = (h(T/2) + h(-T/2))/2$ and $\beta = (h(T/2) - h(-T/2))/T$ are the offset and the slope of a linear trend, respectively, and $h_0(t)$ is continuous and zero valued at the window edges $-T/2$ and $T/2$. Under these conditions, $\frac{dh_0}{dt}(t)$ can be accurately approximated by a first-order difference, and $\frac{dh}{dt}(t)$ can be deduced according to:

$$\frac{dh}{dt}(t) = \begin{cases} \alpha \delta(t+T/2) - \alpha \delta(t-T/2) + \beta + \frac{dh_0}{dt}(t) & \text{for } |t| \leq T/2 \\ 0 & \text{for } |t| > T/2 \end{cases} \quad (5.27)$$

**FIGURE 5.7**

Window derivative approximation. For a Blackman window of 25 points, the difference between the exact and the approximated derivative is very weak.

To evaluate the quality of this approximation, the function **about.dh** compares the exact derivative of a Blackman window [21] with its approximation derived from Equation (5.27). For **Nh=25**, the results are presented in Figure 5.7, showing a maximum error of 3×10^{-3} , leading to a relative error of less than 2%. This error still decreases when the window length increases.

The framework presented here is the background of the function **dwindow**, which essentially reduces to:

```
function Dh=dwindow(h);
% h is a row vector
[hrow,hcol]=size(h); Lh=(hrow-1)/2;

% compute the discontinuities on the border
step_height=(h(1)+h(hrow))/2;
ramp=(h(hrow)-h(1))/(hrow-1);

% add zeros at both sides
h2=[0;h-step_height-ramp*(-Lh:Lh).' ;0];
```

```
Dh=(h2(3:hrow+2)-h2(1:hrow))/2 + ramp; % first order
difference
Dh(1) =Dh(1) +step_height; % add discontinuities
Dh(hrow)=Dh(hrow)-step_height;
```

This function allows the analysis window to be given by a vector of numerical values, instead of chosen among a set of predefined possibilities (Hamming, Hann, Kaiser-Bessel, ... [21]). Thanks to these functions **tfrrsp** and **dwindow**, the reassigned spectrogram of a signal can be easily obtained by a few instructions, such as in the ones included in the function **sprspplot(x, dt, Nb, Nh, c)** presented in Section 5.2.3.1:

```
h = window(Nh,'Gauss'); % Gaussian window
Nx = max(size(x)); t = 1:dt:Nx; % select time samples
[sp, rsp]=tfrrsp(x,t,2*Nb,h,1); % compute the
reassigned spectro
```

The sensitivity of the reassigned spectrogram to the choice of the window shape or length can therefore be easily studied by simply changing the parameters of the function **window**.

5.3.2 Reassigning other time–frequency distributions

5.3.2.1 Reassigned smoothed pseudo-Wigner–Ville distributions

As mentioned in Section 5.2.2.1, spectrograms suffer from a difficult trade-off between time and frequency resolutions, which make them irrelevant for several real case applications. Hopefully, other TF representations exist, which can also be reassigned. For instance, the smoothed pseudo-Wigner–Ville distribution (SPWV) [16] is a very flexible tool, which allows an independent adjustment of the time and frequency smoothings:

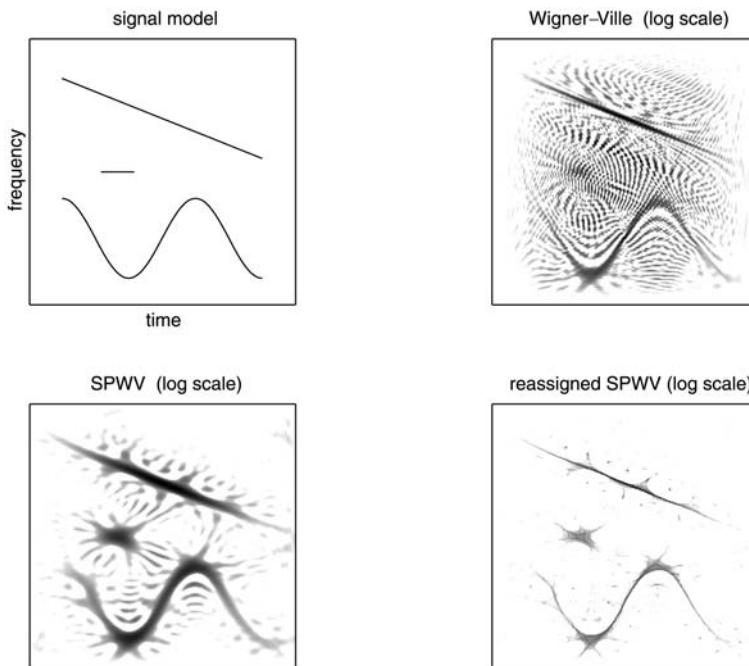
$$\text{SPWV}_x^{g,h}(t,\omega) := \int \int_{-\infty}^{+\infty} g(s-t) H(\xi - \omega) W_x(s, \xi) \frac{ds d\xi}{2\pi} \quad (5.28)$$

where g and h are two even smoothing windows with $h(0) = G(0) = 1$. In that case, expressions in Equations (5.7) and (5.8) become:

$$\hat{t}_x(t,\omega) = t + \frac{\text{SPWV}_x^{g,th}(t,\omega)}{\text{SPWV}_x^{g,h}(t,\omega)} \quad (5.29)$$

$$\hat{\omega}_x(t,\omega) = \omega - i \frac{\text{SPWV}_x^{dg,h}(t,\omega)}{\text{SPWV}_x^{g,h}(t,\omega)} \quad (5.30)$$

These new expressions show that reassigning the SPWV distribution only requires two additional SPWV with particular smoothing windows. The resulting algorithm, presented in Figure 2 of Reference [2] and implemented in the function **tfrrspwv** of the TF toolbox [3], has the same organization in two steps than the algorithm of the reassigned spectrogram:

**FIGURE 5.8**

SPWV reassignment. This figure is similar to [Figure 5.2](#), but the spectrogram and its reassigned version have been replaced by the SPWV and its reassigned version, computed for exactly the same signal realization. The improvement brought by this kind of representation is clearly evidenced.

- The first one computes the three SPWVs.
- The second one iteratively creates the reassigned smoothed pseudo-Wigner–Ville distribution $\hat{P}_x^{g,h}(t,\omega)$ by adding $P_x^{g,h}[n,m]$ to $\hat{P}_x^{g,h}[\hat{n},\hat{m}]$.

The relevance of this representation is supported by Figure 5.8, obtained by running **model1TF2**; this new figure shows two new representations of exactly the same signal as the one used in Figure 5.2. The SPWV (bottom left) yields a better localization of the signal components than the spectrogram, and therefore its reassigned version (bottom right) is still closer to the idealized model and less sensitive to noise than the reassigned spectrogram. In this script, the SPWV and its reassigned version are simply obtained by the instructions:

```
g = window( 21,'Kaiser') ;
h = window(121,'Kaiser') ;
[spwv,rspwv] = tfrrspwv(x,t,N,g,h,1) ;
```

Improvements can easily be looked for by simply changing the definitions of **g** and **h**.

5.3.2.2 Reassigned scalograms

The reassignment principle can also be applied to time-scale representations of the affine class [32]. A widely used member of this class is the scalogram, which is the squared modulus of the continuous wavelet transform:

$$\text{SC}_x^h(t, a) := \left| \text{CW}_x^h(t, a) \right|^2 \quad (5.31)$$

$$\text{with } \text{CW}_x^h(t, a) := \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(u) h^* \left(\frac{u-t}{a} \right) du \quad (5.32)$$

where $h(t)$ (called the *mother wavelet*) is the impulse response of a bandpass filter of central frequency ω_0 , and a is a scale parameter, related to a frequency variable by the relationship $a = \omega_0/\omega$.

As the expression in Equation (5.6), the scalogram results from an affine smoothing of the WVD:

$$\text{SC}_x^h(t, a) = \int \int_{-\infty}^{+\infty} W_x(s, \xi) W_h \left(\frac{s-t}{a}, a\xi \right) \frac{ds d\xi}{2\pi} \quad (5.33)$$

As evidenced by this expression, $\text{SC}_x^h(t, a)$ can be interpreted as the summation of a whole set of energy measures $W_x(s, \xi)$ contained within a TF domain delimited by $W_h(\frac{s-t}{a}, a\xi)$. Instead of assigning this number to the geometric center of this domain, which does not depend on the analyzed signal, it seems more relevant to assign it to the center of gravity, defined by:

$$\hat{t}_x(t, a) = \frac{1}{\text{SC}_x^h(t, a)} \int \int_{-\infty}^{+\infty} s W_x(s, \xi) W_h \left(\frac{s-t}{a}, a\xi \right) \frac{ds d\xi}{2\pi} \quad (5.34)$$

$$\hat{\omega}_x(t, a) := \frac{\omega_0}{\hat{a}_x(t, a)} = \frac{1}{\text{SC}_x^h(t, a)} \int \int_{-\infty}^{+\infty} \xi W_x(s, \xi) W_h \left(\frac{s-t}{a}, a\xi \right) \frac{ds d\xi}{2\pi} \quad (5.35)$$

The resulting reassigned scalogram, defined as:

$$\check{\text{SC}}_x^h(t', a') := \int \int_{-\infty}^{+\infty} \text{SC}_x^h(t, a) \delta(t - \hat{t}_x(t, a), a' - \hat{\omega}_x(t, a)) \frac{a'^2 dt da}{a^2} \quad (5.36)$$

benefits both from the smoothing performed by the mother wavelet, and from the reassignment, which refocuses the scalogram on the squeezed signal description given by the WVD.

From a computational point of view, the local centroids can efficiently be computed by means of two additional wavelet transforms, using two particular mother wavelets:

$$\hat{t}_x(t, a) = t + \mathcal{R}e \left\{ \frac{a \text{CW}_x^{th}(t, a)}{\text{CW}_x^h(t, a)} \right\} \quad (5.37)$$

$$\hat{\omega}_x(t, a) = \frac{\omega_0}{\hat{a}_x(t, a)} = -\mathcal{I}m \left\{ \frac{\text{CW}_x^{dh}(t, a)}{\text{CW}_x^h(t, a)} \right\} \quad (5.38)$$

Several mother wavelet functions can be used. One of them is the Morlet wavelet:

$$h(t) = \frac{1}{\sqrt{T}} e^{-\frac{t^2}{2T^2}} e^{i\omega_0 t} \quad (5.39)$$

In that case, $dh(t) = \frac{th(t)}{T^2} + i\omega_0 h(t)$ and $\text{CW}_x^{dh}(t, a) = -\frac{1}{T^2} \text{CW}_x^{th}(t, a) - i\omega_0 \text{CW}_x^h(t, a)$. As for the Gabor spectrogram, only $\text{CW}_x^h(t, a)$ and $a\text{CW}_x^{th}(t, a)$ need to be computed, because the expression in Equation (5.35) becomes:

$$\hat{\omega}_x(t, a) = \frac{\omega_0}{\hat{a}_x(t, a)} = \frac{\omega_0}{a} + \frac{1}{a^2 T^2} \mathcal{I}m \left\{ \frac{a \text{CW}_x^{th}(t, a)}{\text{CW}_x^h(t, a)} \right\} \quad (5.40)$$

Computer algorithms (as the one used in the function **tfrmsc**) can therefore be deduced from the discrete-time versions of the following expressions:

$$\text{CW}_x^h(t, a) = \sqrt{\frac{|\omega|}{\omega_0 T}} \int_{-\infty}^{+\infty} x(t+\tau) e^{-\frac{\omega^2 \tau^2}{2\omega_0^2 T^2}} e^{-i\omega\tau} d\tau \quad (5.41)$$

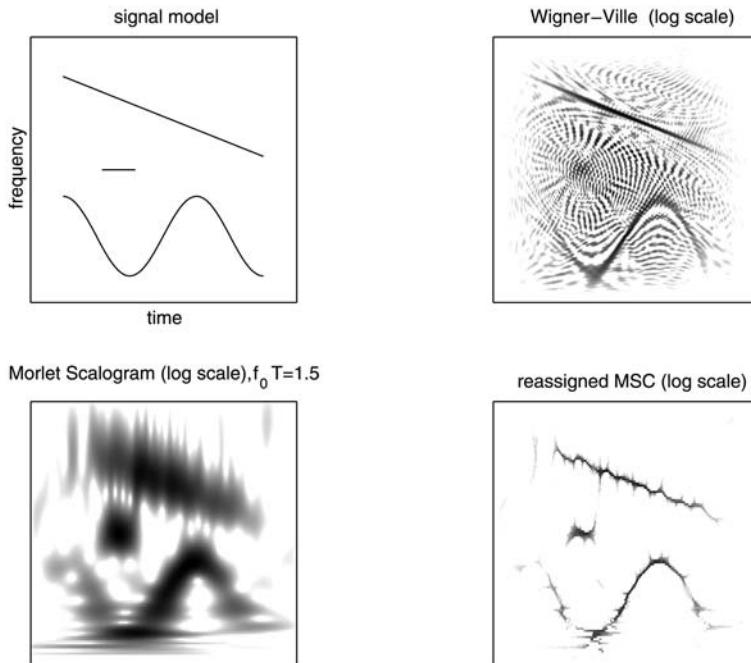


FIGURE 5.9

Scalogram reassignment. This figure is similar to Figure 5.2, but the spectrogram and its reassigned version have been replaced by the Morlet scalogram and its reassigned version, computed for exactly the same signal realization. The use of this kind of representation for the analyzed signal does not seem to be relevant.

$$a \text{CW}_x^{th}(t, a) = \sqrt{\frac{|\omega|}{\omega_0 T}} \int_{-\infty}^{+\infty} x(t + \tau) \tau e^{-\frac{\omega^2 \tau^2}{2\omega_0^2 T^2}} e^{-i\omega\tau} d\tau \quad (5.42)$$

with $a = \omega_0/\omega$. Because the Gaussian analyzing window used in these expressions depends on the frequency (or scale) parameter, FFT algorithms can no longer be used, resulting in much slower algorithms. It should be emphasized that ω_0 and T only appear through their product $\omega_0 T$, which is the only degree of freedom of this representation. Increasing this parameter improves the frequency resolution and reduces the time resolution.

The representations obtained with the Morlet scalogram and its reassigned version for the signal taken as example through this chapter are presented on [Figure 5.9](#). Because the sinusoidal FM requires a small time resolution at low frequencies, the linear FM component at higher frequencies has a very broad frequency localization. This illustrates the fact that a proper choice of a TF representation can only be deduced from a study of the signal structure. The reassigned version, however, remains very close to the signal model.

5.4 Real Case Studies

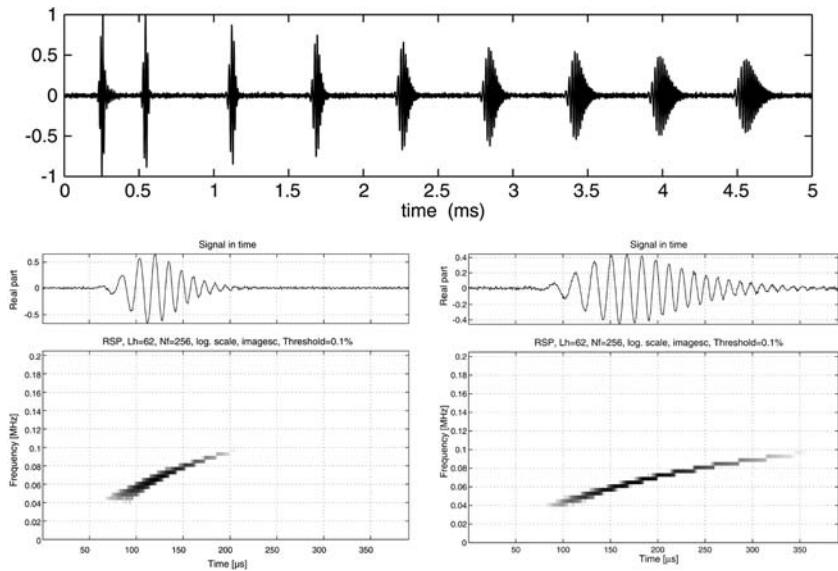
The principles and the algorithms presented in Sections 5.2 and 5.3 can be used efficiently in real case situations. This section briefly presents two applications chosen among many possible ones.

The first one deals with nondestructive testing of metallic cables for suspension bridges. For this, a short wave is emitted by a magnetostrictive transducer, and propagated through a cable under test. The reflections on the normal and defective edges are then received by a sensor [1]. The resulting signal is shown on top of [Figure 5.10](#), whereas reassigned spectrograms in the bottom part give a close look to the fifth and ninth reflected waves, showing precisely the dispersion of the wave through the metallic cable.

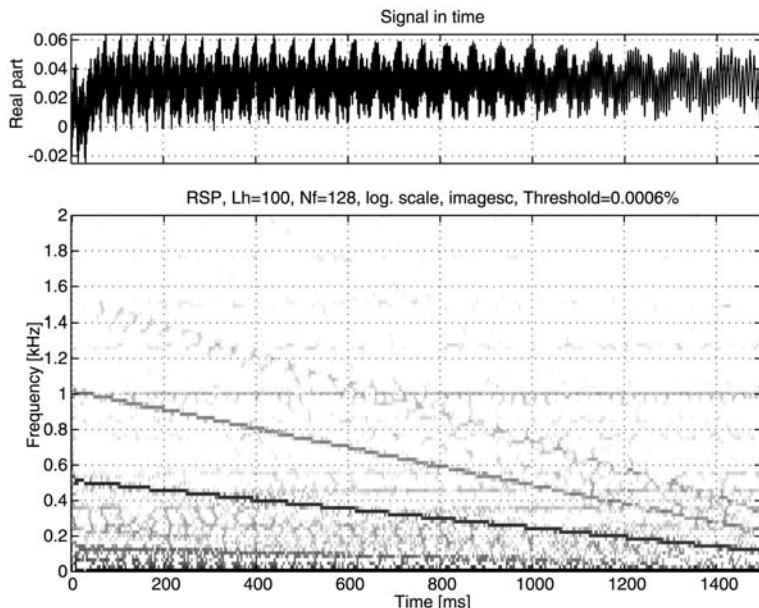
The second application deals with sensorless control of direct current (DC) motors. The electrical contacts and disconnections made by the brushes on the commutator create low-amplitude components in the stator current signal, with frequencies related to the rotor speed [4]. The IFs of these components can be used to estimate the motor speed and to avoid the use of a mechanical sensor in a speed control loop. To illustrate the feasibility of this approach, [Figure 5.11](#) shows the reassigned spectrogram of the stator current during a motor stopping. This representation clearly shows three signal components with almost linearly varying frequency. The higher the frequency is, the higher the chirping rate and therefore the better the estimation of the deceleration, but also the weaker the signal components.

Finally, a very interesting application in musical signal synthesis can also be found at:

<http://www.cerlsoundgroup.org/Loris>

**FIGURE 5.10**

Mechanical waves propagating through a metallic cable. This figure shows the signal received by a sensor (top), and the TF analysis of the fifth (left) and ninth (right) reflections, from which an accurate study of the wave dispersion along the metallic cable can be done.

**FIGURE 5.11**

Rotor speed estimation from the stator current. This figure shows the TF analysis (reassigned spectrogram) of the current signal of a DC motor, evidencing AC components whose instantaneous frequencies are related to the rotor speed.

5.5 Conclusions

Reassignment is a process having the goal of building readable TF representations, by first rubbing out oscillatory interferences, and then squeezing the remaining energy contributions to refocus them on the signal components. Reassignment is a very general principle with efficiency that has been rigorously proved, and that applies to a wide class of TF distributions, yielding an easy-to-read TF analysis, especially at high SNR. Efficient algorithms are available, thus allowing its practical use in effective signal analysis applications.

Beyond TF analysis, useful information on the signal structure can also be extracted from the reassignment operators, and used in a signal processing application. On the one hand, it can be applied to chirp detection, because we have shown [10] that reassignment can be used as the key element of an optimal detector. On the other hand, it can be used as the background of a TF partitioning algorithm for multicomponent signals [8, 11].

References

- [1] J.-C. Aime, M. Brissaud, and L. Laguerre, Generation and Detection of Elastic Guided Waves with Magnetoelastic Device for the Nondestructive Evaluation of Steel Cables and Bars, in Proceedings of the 15th World Conference on Nondestructive Testing, Rome, Italy, 2000.
- [2] F. Auger and P. Flandrin, Improving the readability of time-frequency and time-scale representations by the reassignment method, *IEEE Trans. Signal Process.*, SP-43(5), 1068–1089, 1995.
- [3] F. Auger, P. Flandrin, P. Gonçalvès, and O. Lemoine, Time-frequency toolbox for MATLAB, user’s guide and reference guide, available at <http://iut-saint-nazaire.univ-nantes.fr/~auger/tftb.html>.
- [4] F. Auger and M. Hilairet, Suivi de Raies Spectrales avec un Faible Coût de Calcul (in French), in Proceedings of the 17ème Colloque GRETSI, Vannes, France, 1999, pp. 909–912.
- [5] R. Carmona, H.L. Hwang, and B. Torrésani, *Practical Time-Frequency Analysis*, Academic Press, San Diego, CA, 1998.
- [6] E. Chassande-Mottin, F. Auger, and P. Flandrin, Supervised Time-Frequency Reassignment, in Proceedings of the IEEE International Symposium on Time-Frequency and Time-Scale Analysis, Paris, France, 1996, pp. 517–520.

- [7] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin, Differential reassignment, *IEEE Signal Process. Lett.*, SPL-4(10), 293–294, 1997.
- [8] E. Chassande-Mottin, F. Auger, I. Daubechies, and P. Flandrin, Partition du Plan Temps-Fréquence et Réallocation (in French), in Proceedings of the 16ème Colloque GRETSI, Grenoble, France, 1997, pp. 1447–1450.
- [9] E. Chassande-Mottin, F. Auger, and P. Flandrin, On the statistics of spectrogram reassignment, *Multidim. Syst. Signal Process.*, 9(4), 355–362, 1998.
- [10] E. Chassande-Mottin and P. Flandrin, On the time-frequency detection of chirps, *Appl. Comp. Harm. Anal.*, 6(9), 252–281, 1999.
- [11] E. Chassande-Mottin, F. Auger, and P. Flandrin, Time-frequency/time-scale reassignment, in *Wavelets and Signal Processing*, L. Debnath, Ed., Birkhäuser, Boston, to appear 2002.
- [12] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1995.
- [13] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrésani, Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies, *IEEE Trans. Inf. Theory*, IT-38(2), 644–673, 1992.
- [14] I. Djurovic and L.J. Stankovic, Time-frequency representation based on the reassigned S-method, *Signal Process.*, 77(1), 115–120, 1999.
- [15] K. Fitz, L. Haken, and P. Christensen, Transient preservation under Transformation in an Additive Sound Model, in Proceedings of the International Computer Music Conference, Berlin (D), 2000, pp. 392–395.
- [16] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.
- [17] D. Friedman, Instantaneous Frequency Distribution vs. Time : An Interpretation of the Phase Structure of Speech, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Tampa, FL, 1985, pp. 1121–1124.
- [18] V. Gibiat, F. Wu, P. Perio, and S. Chantreuil, Analyse spectrale différentielle (A.S.D.), *C. R. Acad. Sci. Paris, sér. II*, 294, 633–636, 1982.
- [19] P. Guillemain and R. Kronland-Martinet, Horizontal and Vertical Ridges associated to Continuous Wavelet Transforms, in Proceedings of the IEEE International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, BC, 1992, pp. 63–66.
- [20] P. Guillemain and R. Kronland-Martinet, Characterization of acoustic signals through continuous linear time-frequency representations, in Proceedings of the IEEE, 84(4), 561–587, 1996.

- [21] F.J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, in *Proceedings of the IEEE*, 66(1), 51–83, 1978.
- [22] F. Hlawatsch and G.F. Boudreaux-Bartels, Linear and quadratic time-frequency signal representations, *IEEE Signal Process. Mag.*, 21–67, 1992.
- [23] F. Hlawatsch and P. Flandrin, The interference structure of Wigner distributions and related time-frequency signal representations, in *The Wigner Distribution: Theory and Applications in Signal Processing*, W. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1998, pp. 59–133.
- [24] J.-M. Innocent and B. Torrésani, Wavelets and binary coalescences detection, *Appl. Comp. Harm. Anal.*, 4(2), 113–116, 1997.
- [25] K. Kodera, C. De Villedary, and R. Gendrin, A new method for the numerical analysis of nonstationary signals, *Phys. Earth Plan. Int.*, 12, 142–150, 1976.
- [26] K. Kodera, Analyse Numérique de Signaux Géophysiques Non-Stationnaires, Thèse de doctorat, Université de Paris VI, France, 1976.
- [27] K. Kodera, R. Gendrin, and C. De Villedary, Analysis of time-varying signals with small BT values, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-26(1), 64–76, 1978.
- [28] R.J. McAulay and T.F. Quatieri, Speech analysis–synthesis based on a sinusoidal representation, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-34(4), 744–754, 1986.
- [29] S. Maes, The Synchrosqueezed Representation Yields a New Reading of the Wavelet Transform, in *Proceedings of the SPIE 95 on OE/Aerospace Sensing and Dual Use Photonics*, Orlando, FL, 1995, pp. 532–559.
- [30] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, New York, 1998.
- [31] C. Richard and R. Lengellé, Joint recursive implementation of time-frequency representations and their modified version by the reassignment method, *Signal Process.*, 60(2), 163–179, 1997.
- [32] O. Rioul and P. Flandrin, Time-scale energy distributions: a general class extending wavelet transforms, *IEEE Trans. Signal Process.*, SP-40(7), 1746–1757, 1992.
- [33] J.-M. Tribolet, A new phase unwrapping algorithm, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-25(2), 170–177, 1977.

6

Linear Time–Frequency Filters: On-line Algorithms and Applications

Gerald Matz and Franz Hlawatsch

Vienna University of Technology

CONTENTS

6.1	Introduction	205
6.2	Explicit Time–Frequency Filter Design I: Zadeh Filter	208
6.3	Explicit Time–Frequency Filter Design II: Weyl Filter	210
6.4	Explicit Time–Frequency Filter Design III: Simulation Results and Applications	224
6.5	Implicit Time–Frequency Filter Design I: Short-time Fourier Transform Filter	230
6.6	Implicit Time–Frequency Filter Design II: Gabor Filter	243
6.7	Implicit Time–Frequency Filter Design III: Simulation Results and Applications	256
6.8	Conclusions	262
	Acknowledgment	265
	References	265

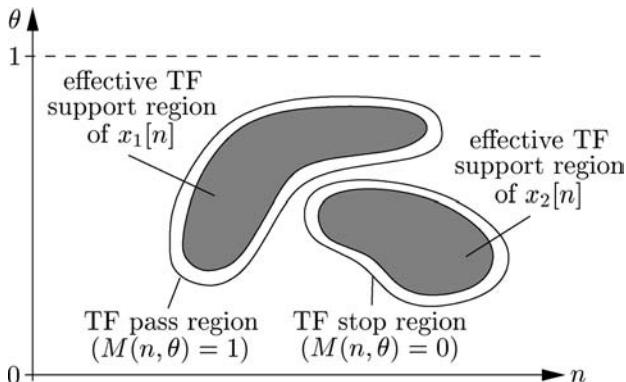
6.1 Introduction

Linear time-varying (LTV) filters [1–7] have important applications including non-stationary statistical signal processing (signal detection and estimation, spectrum estimation, etc.) and communications over time-varying channels (interference excision, channel modeling, estimation, equalization, etc.). LTV filters are particularly useful for weighting, suppressing or separating nonstationary signal components.

The input–output relation of a discrete-time LTV filter (LTV system) \mathbf{H} reads

$$y[n] = (\mathbf{H}x)[n] = \sum_{n'=-\infty}^{\infty} h[n, n'] x[n'] = \sum_{m=-\infty}^{\infty} g[n, m] x[n-m], \quad (6.1)$$

where $g[n, m] = h[n, n - m]$ and $h[n, n'] = g[n, n - n']$ are two different definitions

**FIGURE 6.1**

Simple example motivating the concept of TF filters. The signal consists of two effectively TF-disjoint components $x_1[n]$ and $x_2[n]$. The filter is supposed to pass $x_1[n]$ and suppress $x_2[n]$.

of the impulse response of \mathbf{H} (note that $n, n' \in \mathbb{Z}$ denote absolute time whereas $m \in \mathbb{Z}$ denotes time delay).

6.1.1 Time–frequency filters

The nonstationary nature of input signal $x[n]$, output signal $y[n]$ and LTV filter \mathbf{H} suggests the use of time–frequency (TF) representations for designing and, possibly, also implementing LTV filters. Here, we consider a concept of *linear TF filters* that is based on a TF representation of the filter or of the input signal. This concept can be motivated by the simple example in Figure 6.1. The filter input signal $x[n]$ consists of two components $x_1[n]$ and $x_2[n]$ that are effectively nonoverlapping in the TF domain. We desire the filter \mathbf{H} to pass $x_1[n]$ and suppress $x_2[n]$ so that the filter output signal is $y[n] = x_1[n]$. This specification of \mathbf{H} can be expressed by a TF weight function $M(n, \theta)$ (θ denotes normalized frequency) that is 1 on the effective TF support region of $x_1[n]$ and 0 on the effective TF support region of $x_2[n]$.

More generally, any other desired TF weighting (TF-dependent attenuation or amplification) can be described by some real-valued TF weight function $M(n, \theta) \in \mathbb{R}$. Thus, *the LTV filter H is specified by a prescribed TF weight function $M(n, \theta)$* . The two different general approaches to designing linear TF filters based on a TF weight function follow:

- *Explicit design [8–14].* The LTV filter \mathbf{H} is calculated such that a TF representation of \mathbf{H} is equal to or best approximates the TF weight function $M(n, \theta)$. The filtering is then performed in the time domain according to Equation (6.1).
- *Implicit design [8, 9, 15–25].* The LTV filter \mathbf{H} is designed *implicitly* during the filtering, which is an analysis–weighting–synthesis procedure. First (analysis), a linear TF representation of the input signal $x[n]$ is calculated. Second (weighting), this TF representation is multiplied by the TF weight function $M(n, \theta)$.

Third (synthesis), the output signal $y[n]$ is calculated in a linear manner from the TF function obtained in the second step. Because all processing steps are linear, the overall procedure amounts to an LTV filter.*

6.1.2 Underspread filters

The LTV filters discussed here are supposed to perform a TF weighting; in this context, significant time shifts or frequency shifts are undesired. The TF shifts introduced by an LTV filter \mathbf{H} are characterized by the *spreading function* [28–37]

$$S_{\mathbf{H}}(m, \nu) \triangleq \sum_{n=-\infty}^{\infty} h[n, n-m] e^{-j2\pi\nu n} = \sum_{n=-\infty}^{\infty} g[n, m] e^{-j2\pi\nu n}. \quad (6.2)$$

Here, m and ν denote time shift (time lag, delay) and frequency shift (frequency lag, Doppler shift), respectively. The input–output relation in Equation (6.1) can be reformulated as

$$y[n] = \sum_{m=-\infty}^{\infty} \int_0^1 S_{\mathbf{H}}(m, \nu) x[n-m] e^{j2\pi\nu n} d\nu, \quad (6.3)$$

whereby the filter output signal $y[n]$ is represented as a weighted superposition of TF shifted versions $x[n-m] e^{j2\pi\nu n}$ of the input signal $x[n]$; the weight function in this superposition is precisely the spreading function $S_{\mathbf{H}}(m, \nu)$.

An LTV filter \mathbf{H} is said to be *underspread* if $S_{\mathbf{H}}(m, \nu)$ is well concentrated about the origin of the (m, ν) plane [31, 32, 34–36]. In view of Equation (6.3), this indicates that the filter introduces only small TF shifts. In contrast, LTV filters introducing large TF shifts are termed *overspread*. As mentioned earlier, TF filters as considered here are supposed to be essentially underspread.

6.1.3 Style and organization of this chapter

Practical applications often require real-time operation or involve arbitrarily long signals. Thus, the emphasis of our presentation is on efficient *on-line implementations* for all TF filters considered. The term *on-line* implies that a fixed number of operations per signal sample and a fixed amount of memory are required, independently of the total signal length; furthermore, at a given time instant n_0 , the input signal $x[n]$ and the TF weight function $M(n, \theta)$ have to be known only within a local, fixed-length time interval about n_0 . Our presentation also includes discussions of the choice of design parameters and estimates of algorithm complexity.

The rest of this chapter is organized as follows. Sections 6.2 to 6.4 discuss two explicit TF filter designs, namely, the Zadeh filter (Section 6.2) and the Weyl filter (Section 6.3). Two implicit TF filter designs are discussed next, namely, the

*An implicit design using *quadratic* TF signal representations, such as the Wigner distribution or a smoothed Wigner distribution, has also been studied [26, 27]. Because it results in a nonlinear filter, it is not considered here.

(multiwindow) short-time Fourier transform (STFT) filter in Section 6.5 and the (multiwindow) Gabor filter in Section 6.6. Simulation results and selected applications illustrating the specific advantages and limitations of the various TF filter methods are presented in Sections 6.4 and 6.7. Finally, an overall comparison of the various TF filters is provided in the concluding Section 6.8.

6.2 Explicit Time–Frequency Filter Design I: Zadeh Filter

In this section and the next two sections, we consider the explicit design of TF filters. We start with an especially simple design scheme.

6.2.1 Zadeh function and Zadeh filter

6.2.1.1 Zadeh function

For a discrete-time LTV filter \mathbf{H} , Zadeh’s time-varying transfer function (briefly called *Zadeh function* hereafter) is a TF representation defined as [3, 28–32, 34, 38, 39]

$$Z_{\mathbf{H}}(n, \theta) \triangleq \sum_{m=-\infty}^{\infty} h[n, n-m] e^{-j2\pi\theta m} = \sum_{m=-\infty}^{\infty} g[n, m] e^{-j2\pi\theta m}.$$

The spreading function in Equation (6.2) is the two-dimensional (2-D) Fourier transform (FT) of the Zadeh function, that is:

$$S_{\mathbf{H}}(m, \nu) = \sum_{n=-\infty}^{\infty} \int_0^1 Z_{\mathbf{H}}(n, \theta) e^{-j2\pi(\nu n - m\theta)} d\theta. \quad (6.4)$$

The impulse response $g[n, m]$ can be reobtained from $Z_{\mathbf{H}}(n, \theta)$ through the inversion formula:

$$g[n, m] = \int_0^1 Z_{\mathbf{H}}(n, \theta) e^{j2\pi\theta m} d\theta. \quad (6.5)$$

If \mathbf{H} is underspread as defined in Subsection 6.1.2, the Zadeh function $Z_{\mathbf{H}}(n, \theta)$ can be interpreted as an (approximate) time-varying frequency response or TF transfer function of \mathbf{H} , in the sense that $Z_{\mathbf{H}}(n, \theta)$ describes the TF weighting effected by \mathbf{H} [31, 34–36]. Furthermore, because of the FT relation in Equation (6.4), the concentration of $S_{\mathbf{H}}(m, \nu)$ in the underspread case implies that the Zadeh function of an underspread filter is a smooth (lowpass-type) function. Conversely, the Zadeh function of an overspread filter has oscillatory (bandpass-type) components corresponding to components of $S_{\mathbf{H}}(m, \nu)$ away from the origin.

6.2.1.2 Zadeh filter

The Zadeh filter design is based on the transfer function interpretation of $Z_{\mathbf{H}}(n, \theta)$ in the underspread case. Given a prescribed TF weight function $M(n, \theta)$ of an LTV

filter \mathbf{H} to be designed, we simply set the Zadeh function[†] of \mathbf{H} equal to $M(n, \theta)$, that is

$$Z_{\mathbf{H}_Z}(n, \theta) \triangleq M(n, \theta). \quad (6.6)$$

The LTV filter \mathbf{H}_Z defined by this equation can be called *Zadeh filter* [8–10]. This design is justified if $M(n, \theta)$, and thus $Z_{\mathbf{H}_Z}(n, \theta)$, is smooth so that \mathbf{H}_Z is underspread.

The impulse response of the Zadeh filter is obtained using the inversion formula (6.5):

$$g_Z[n, m] = \int_0^1 M(n, \theta) e^{j2\pi\theta m} d\theta. \quad (6.7)$$

The Zadeh filter is then implemented according to Equation (6.1). It is easily shown that for $M(n, \theta) \equiv 1$, the Zadeh filter output signal $y[n] = (\mathbf{H}_Z x)[n]$ is equal to the input signal $x[n]$.

6.2.2 On-line implementation of Zadeh filter

For a practical implementation of the Zadeh filter design, we define a discrete-frequency TF weight function $M[n, k]$ as a frequency-sampled version of $M(n, \theta)$:

$$M[n, k] \triangleq M\left(n, \frac{k}{K}\right), \quad k = 0, 1, \dots, K-1$$

Here, K is the number of frequency samples (assumed even for simplicity) and $1/K$ is the frequency sampling period. K must be chosen large enough so that $M(n, \theta)$ is sampled sufficiently densely; this requires a larger K for faster variations of $M(n, \theta)$ in the θ direction.

In accordance with this frequency discretization, Equation (6.7) is replaced by the inverse discrete Fourier transform (IDFT) relation:

$$g_Z[n, m] = \frac{1}{K} \sum_{k=0}^{K-1} M[n, k] e^{j2\pi \frac{k}{K} m}, \quad m = -\frac{K}{2}, -\frac{K}{2} + 1, \dots, \frac{K}{2} - 1 \quad (6.8)$$

Note that this results in a finite-delay impulse response (i.e., $g_Z[n, m]$ has finite m -support). By using the general input–output relation in Equation (6.1), the output signal can now be written as

$$y[n] = \sum_{m=-K/2}^{K/2-1} g_Z[n, m] x[n-m]. \quad (6.9)$$

Thus, the computation of $y[n]$ involves the input signal samples $x[n-K/2+1], x[n-K/2+2], \dots, x[n+K/2]$ located about the current time point n . By defining the

[†]The Zadeh function is a linear TF representation of \mathbf{H} . Whereas this linearity is not necessary for the design to result in a linear filter, it simplifies the filter on-line implementation of the filter. An explicit design of linear TF filters based on a quadratic TF filter representation has been presented in [11].

length- K vectors $\mathbf{g}_n \triangleq (g_Z[n, K/2-1] \ g_Z[n, K/2-2] \cdots g_Z[n, -K/2])^T$ and $\mathbf{x}_n \triangleq (x[n-K/2+1] \ x[n-K/2+2] \cdots x[n+K/2])^T$ (where T stands for transposition), the input–output relation in Equation (6.9) can be written as the vector inner product (dot product)

$$y[n] = \mathbf{g}_n^T \mathbf{x}_n. \quad (6.10)$$

The on-line implementation of the Zadeh filter \mathbf{H}_Z can thus be summarized at each time instant n as follows:

1. Determine \mathbf{g}_n as the length- K IDFT of $M[n, k]$ with respect to k ; (see Equation (6.8)).
2. Calculate the output signal sample $y[n]$ according to Equation (6.9) or, equivalently, Equation (6.10).

This on-line Zadeh filter satisfies $y[n] = x[n]$ for $M[n, k] \equiv 1$. Because Equation (6.9) involves the future input samples $x[n+1], x[n+2], \dots, x[n+K/2]$, causal operation of the Zadeh filter necessitates the introduction of a time delay of $K/2$ samples. At each time instant n , the algorithm requires an IDFT of length K and an inner product of two length- K vectors. Thus, the computational complexity of the on-line Zadeh filter is $\mathcal{O}(K + K \log K)$ operations per output sample. The delay of $K/2$ implies that the $K/2$ past impulse response vectors $\mathbf{g}_{n-1}, \mathbf{g}_{n-2}, \dots, \mathbf{g}_{n-K/2}$ have to be stored. In total, $K^2/2 + K$ memory locations are required ($K/2$ length- K impulse response vectors \mathbf{g}_i and the length- K input signal vector \mathbf{x}_n). Additional memory may be required by the fast Fourier transform (FFT) unless an in-place algorithm is used.

Simulation results illustrating the performance and application of the on-line Zadeh filter are presented in Section 6.4.

6.3 Explicit Time–Frequency Filter Design II: Weyl Filter

The *Weyl filter* is obtained by using the Weyl symbol (instead of the Zadeh function) as the TF filter representation underlying the explicit TF filter design. In certain cases, the Weyl filter features better performance than the Zadeh filter (see Subsections 6.4.2 and 6.4.4). This comes at the expense of an increased theoretical and algorithmic complexity.

6.3.1 Weyl symbol

The discrete-time Weyl symbol is defined as [10, 30–37, 40–43]

$$L_{\mathbf{H}}(n, \theta) \triangleq 2 \sum_{m=-\infty}^{\infty} \tilde{g}[n, m] e^{-j4\pi\theta m} \quad (6.11)$$

with

$$\tilde{g}[n, m] \triangleq h[n+m, n-m].$$

The symmetrical structure of $\tilde{g}[n, m]$ causes the Weyl symbol to have important advantages over the Zadeh function. On the other hand, in the discrete-time setting considered here, this symmetry also causes two (related) problems. First, $L_{\mathbf{H}}(n, \theta)$ is 1/2-periodic with respect to θ , which causes aliasing effects unless the filter \mathbf{H} is a halfband filter as explained presently. Second, only values of $h[n_1, n_2]$ with $n_1 + n_2$ even enter in the definition of $L_{\mathbf{H}}(n, \theta)$ (because $n_1 = n+m$ and $n_2 = n-m$ entail $n_1 + n_2 = 2n$). This means that n_1 and n_2 are either both even or both odd, corresponding to the impulse response samples $h[2l, 2l'] = \tilde{g}[l+l', l-l']$ and $h[2l+1, 2l'+1] = \tilde{g}[l+l'+1, l-l']$ with $l, l' \in \mathbb{Z}$; the remaining impulse response samples $h[2l, 2l'+1]$ and $h[2l+1, 2l']$ have no effect on $L_{\mathbf{H}}(n, \theta)$. Thus, the correspondence between $L_{\mathbf{H}}(n, \theta)$ and $h[n, n']$ is not one-to-one unless the class of filters \mathbf{H} is suitably restricted *a priori*.

Inverting the FT in Equation (6.11) yields

$$\tilde{g}[n, m] = \int_0^{1/2} L_{\mathbf{H}}(n, \theta) e^{j4\pi\theta m} d\theta, \quad (6.12)$$

which allows procurement of the impulse response samples $h[2l, 2l'] = \tilde{g}[l+l', l-l']$ and $h[2l+1, 2l'+1] = \tilde{g}[l+l'+1, l-l']$. The remaining samples $h[2l, 2l'+1]$ and $h[2l+1, 2l']$ are left unspecified.

If \mathbf{H} is an underspread halfband filter, the Weyl symbol $L_{\mathbf{H}}(n, \theta)$ is a smooth function that, moreover, can be interpreted as an (approximate) TF transfer function of \mathbf{H} [10, 31, 32, 34–36].

6.3.1.1 Halfband filters

The concept of *halfband filters* is essential for a characterization of aliasing in the Weyl symbol. For an arbitrary but fixed center frequency θ_0 , let us define the *halfband subspace* \mathcal{H} as the linear space of signals $x[n]$ whose FT $X(\theta)$ is supported in the halfband $[\theta_0 - 1/4, \theta_0 + 1/4]_1$ (here, $[\theta_1, \theta_2]_1 \triangleq \bigcup_{k=-\infty}^{\infty} [\theta_1 + k, \theta_2 + k]$ denotes the 1-periodic continuation of a frequency interval $[\theta_1, \theta_2]$). That is, $x[n] \in \mathcal{H}$ if and only if $X(\theta) = 0$ for $\theta \notin [\theta_0 - 1/4, \theta_0 + 1/4]_1$. Special cases of halfband subspaces are the space of analytic signals ($\theta_0 = 1/4$) and the space of signals oversampled by a factor of two ($\theta_0 = 0$).

We now consider the following three special types of LTV filters:

- An *input halfband filter* picks up input signal components only in a given halfband subspace \mathcal{H} . That is, if $x[n] \notin \mathcal{H}$, then the output signal $y[n]$ is zero. (Mathematically speaking, the null space [5, 6] of the filter is the complement of the halfband subspace \mathcal{H} .)

- An *output halfband filter* is an LTV filter with output signals lying in the halfband subspace \mathcal{H} (i.e., $y[n] \in \mathcal{H}$ for all $x[n]$). (Mathematically speaking, the range [5, 6] of the filter is \mathcal{H} .)
- Finally, an input–output halfband filter is both an input halfband filter and an output halfband filter for the same halfband subspace \mathcal{H} . The linear space of all halfband filters is denoted as \mathcal{H}^2 .

It can be shown that the Weyl symbol is a nonaliased TF representation of an LTV filter \mathbf{H} if and only if \mathbf{H} is a halfband filter. This complicates, but does not really restrict, the practical application of the Weyl symbol. Indeed, a fullband filter can always be associated to an equivalent halfband filter that is obtained from the fullband filter either by means of oversampling by a factor of 2 or, in the case of a real-valued filter, by conversion to the associated “analytic filter”. For underspread filters (having small frequency shifts), the preceding three filter types — input halfband filter, output halfband filter and halfband filter — are approximately equivalent; for example, an underspread input halfband filter is approximately also an output halfband filter and, thus, a halfband filter.

6.3.2 Weyl filter

As mentioned earlier, the Weyl symbol of an underspread halfband filter \mathbf{H} can be interpreted as an (approximate) TF transfer function of \mathbf{H} . This transfer function interpretation motivates a TF filter design where the Weyl symbol of the filter is set equal to the specified TF weight function $M(n, \theta)$ [8–10, 12]:

$$L_{\mathbf{H}_W}(n, \theta) \triangleq M(n, \theta). \quad (6.13)$$

Because $L_{\mathbf{H}_W}(n, \theta)$ has frequency period 1/2, the TF model function $M(n, \theta)$ must be specified on a *halfband* frequency interval $[\theta_0 - 1/4, \theta_0 + 1/4]$, where θ_0 is some fixed center frequency. Hereafter, $M(n, \theta)$ is considered periodic in θ with period 1/2. Any LTV filter \mathbf{H}_W that satisfies Equation (6.13) is called a *Weyl filter*. The Weyl filter design is justified only if \mathbf{H}_W is underspread, which requires that $M(n, \theta)$ is sufficiently smooth.

Inserting Equation (6.13) into Equation (6.12) yields

$$\tilde{g}_W[n, m] = \int_0^{1/2} M(n, \theta) e^{j4\pi\theta m} d\theta. \quad (6.14)$$

This allows us to obtain $h_W[2l, 2l'] = \tilde{g}_W[l + l', l - l']$ and $h_W[2l + 1, 2l' + 1] = \tilde{g}_W[l + l' + 1, l - l']$, whereas $h_W[2l, 2l' + 1]$ and $h_W[2l + 1, 2l']$ are left unspecified. Thus, the Weyl filter \mathbf{H}_W is not uniquely defined. We next consider three alternative design constraints that uniquely define \mathbf{H}_W but do not result in an exact halfband filter in general. Modified Weyl filter designs that result in a halfband filter are presented in Subsection 6.3.3. On-line implementations of all these Weyl filters are discussed in Subsection 6.3.4.

6.3.2.1 Minimum-energy Weyl filter

The simplest design constraint that can be used for removing the ambiguity of the Weyl filter design rule in Equation (6.13) is that of minimum energy. The energy of \mathbf{H}_W is defined as

$$E_{\mathbf{H}_W} \triangleq \sum_{n=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} |h_W[n, n']|^2.$$

Because Equation (6.13) does not specify $h_W[2l, 2l' + 1]$ and $h_W[2l + 1, 2l']$, $E_{\mathbf{H}_W}$ is obviously minimized by choosing $h_W[2l, 2l' + 1] = h_W[2l + 1, 2l'] = 0$. The resulting *minimum-energy Weyl filter* maps even (odd) samples of the input signal to even (odd) samples of the output signal:

$$\begin{aligned} y[2l] &= \sum_{l'=-\infty}^{\infty} h_W[2l, 2l'] x[2l'], \\ y[2l+1] &= \sum_{l'=-\infty}^{\infty} h_W[2l+1, 2l'+1] x[2l'+1]. \end{aligned} \quad (6.15)$$

This is shown in [Figure 6.2\(a\)](#). With $h_W[2l, 2l'] = \tilde{g}_W[l + l', l - l']$ and $h_W[2l + 1, 2l' + 1] = \tilde{g}_W[l + l' + 1, l - l']$, Equation (6.15) is equivalent to the input–output relation:

$$y[n] = \sum_{m=-\infty}^{\infty} \tilde{g}_W[n-m, m] x[n-2m], \quad (6.16)$$

where $\tilde{g}_W[n, m]$ is given by Equation (6.14). The minimum-energy Weyl filter is a central component of other Weyl filter designs, to be discussed next.

6.3.2.2 Input halfband Weyl filter

Alternatively, we may constrain the Weyl filter \mathbf{H}_W to be an input halfband filter as further defined earlier. It can be shown that this design constraint specifies the missing impulse response samples $h_W[2l, 2l' + 1]$ and $h_W[2l + 1, 2l']$ via the following relations with the known samples $h_W[2l, 2l']$ and $h_W[2l + 1, 2l' + 1]$:

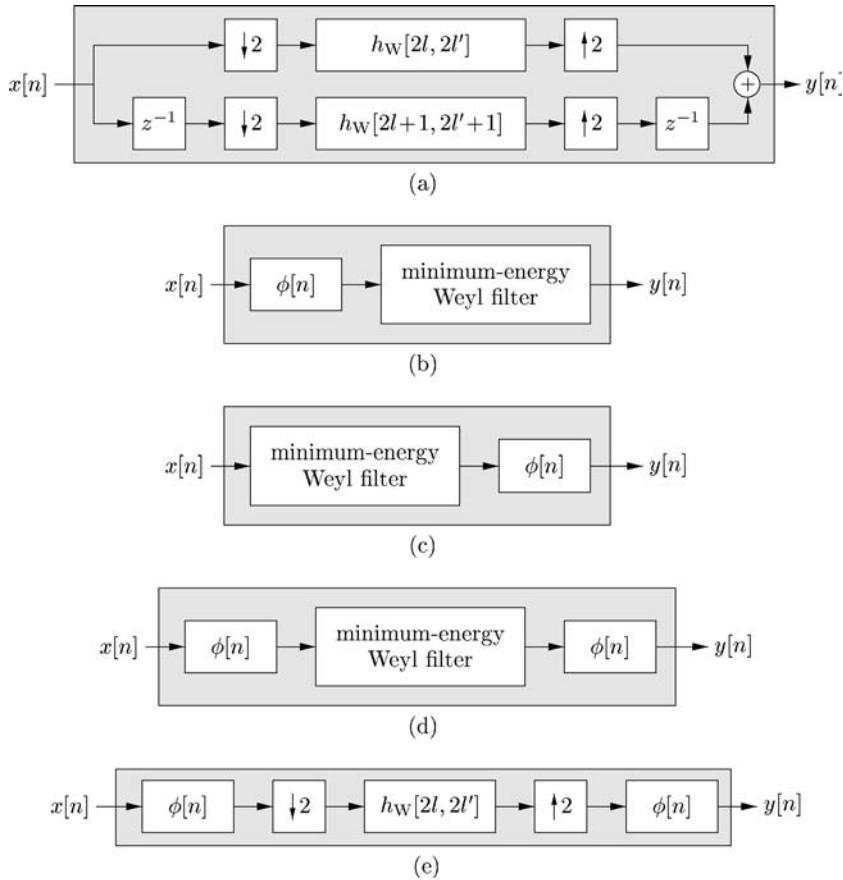
$$h_W[2l, 2l' + 1] = \sum_{i=-\infty}^{\infty} \phi[2l' + 1 - 2i] h_W[2l, 2i], \quad (6.17a)$$

$$h_W[2l + 1, 2l'] = \sum_{i=-\infty}^{\infty} \phi[2l' - 2i - 1] h_W[2l + 1, 2i + 1], \quad (6.17b)$$

where $\phi[n]$ is the impulse response of an ideal low-pass or bandpass filter with (total) bandwidth $1/2$ and center frequency θ_0 , that is

$$\phi[n] \triangleq \frac{\sin(\pi n/2)}{\pi n} e^{j2\pi\theta_0 n}. \quad (6.18)$$

The resulting special Weyl filter is called *input halfband Weyl filter*.

**FIGURE 6.2**

Weyl filters and variations. (a) Minimum-energy Weyl filter; (b) input halfband Weyl filter; (c) output halfband Weyl filter; (d) halfband Weyl filter; (e) approximate halfband Weyl filter. The symbol $\downarrow 2$ denotes subsampling by 2 (i.e., retaining only the even-indexed samples), $\uparrow 2$ denotes upsampling by 2 (i.e., inserting a zero between two adjacent samples) and z^{-1} denotes the unit delay.

The interpolation relations in Equation (6.17) can alternatively be formulated as a halfband restriction of the input signal. The input halfband Weyl filter can thus be implemented as follows:

1. Halfband restriction of the input signal:

$$x_{\mathcal{H}}[n] = \sum_{n'=-\infty}^{\infty} \phi[n-n'] x[n'] \in \mathcal{H}. \quad (6.19)$$

2. Application of the minimum-energy Weyl filter to $x_{\mathcal{H}}[n]$; see Equation (6.16):

$$y[n] = \sum_{m=-\infty}^{\infty} \tilde{g}_{\mathcal{W}}[n-m, m] x_{\mathcal{H}}[n-2m]. \quad (6.20)$$

The filtering operation in Equation (6.20) only uses the impulse response samples $\tilde{g}_{\mathcal{W}}[n, m]$ that are specified by the original design rule in Equation (6.14). The overall filter structure is shown in [Figure 6.2\(b\)](#).

6.3.2.3 Output halfband Weyl filter

As a third alternative, we may constrain $\mathbf{H}_{\mathcal{W}}$ to be an output halfband filter. This can be shown to specify the unknown impulse response samples as

$$h_{\mathcal{W}}[2l, 2l'+1] = \sum_{i=-\infty}^{\infty} \phi[2l-2i-1] h_{\mathcal{W}}[2i+1, 2l'+1], \quad (6.21a)$$

$$h_{\mathcal{W}}[2l+1, 2l'] = \sum_{i=-\infty}^{\infty} \phi[2l+1-2i] h_{\mathcal{W}}[2i, 2l']. \quad (6.21b)$$

The resulting special Weyl filter is called *output halfband Weyl filter*.

The relations in Equation (6.21) can alternatively be formulated as a halfband restriction of the output signal. Accordingly, the output halfband Weyl filter can be implemented as follows:

1. Application of the minimum-energy Weyl filter to the input signal $x[n]$; see Equation (6.16):

$$\tilde{y}[n] = \sum_{m=-\infty}^{\infty} \tilde{g}_{\mathcal{W}}[n-m, m] x[n-2m]. \quad (6.22)$$

2. Halfband restriction of the resulting intermediate output signal $\tilde{y}[n]$:

$$y[n] = \sum_{n'=-\infty}^{\infty} \phi[n-n'] \tilde{y}[n'] \in \mathcal{H}. \quad (6.23)$$

The filtering operation in Equation (6.22) only uses the impulse response samples $\tilde{g}_{\mathcal{W}}[n, m]$ that are specified by Equation (6.14). The overall filter structure is illustrated in [Figure 6.2\(c\)](#).

6.3.3 Halfband Weyl filter

In this subsection, we present two modified Weyl filter designs that result in halfband filters and thus solve the aliasing problem associated with the Weyl filter design rule in Equation (6.13). The on-line implementation of these filters is discussed in Subsection 6.3.4.

6.3.3.1 Definition and calculation of halfband Weyl filter

Let \mathbf{H} be a halfband filter for a halfband subspace \mathcal{H} with center frequency θ_0 . The halfband filter property $\mathbf{H} \in \mathcal{H}^2$ is incompatible with the Weyl filter design rule $L_{\mathbf{H}_W}(n, \theta) \triangleq M(n, \theta)$ in Equation (6.13): in general, no Weyl filter \mathbf{H}_W exists that is a halfband filter. Therefore, we replace the Weyl filter definition $L_{\mathbf{H}_W}(n, \theta) \equiv M(n, \theta)$ by the constrained minimization

$$\mathbf{H}_{HW} \triangleq \arg \min_{\mathbf{H} \in \mathcal{H}^2} \|M - L_{\mathbf{H}}\|,$$

where $\|M - L_{\mathbf{H}}\|^2 = \sum_{n=-\infty}^{\infty} \int_0^{1/2} |M(n, \theta) - L_{\mathbf{H}}(n, \theta)|^2 d\theta$. We term \mathbf{H}_{HW} *halfband Weyl filter*. According to its definition, \mathbf{H}_{HW} is the halfband filter whose Weyl symbol is closest to $M(n, \theta)$. It can be shown that \mathbf{H}_{HW} can be obtained as follows:

1. The TF weight function $M(n, \theta)$ is made “halfband consistent” via the convolution [27]

$$\widetilde{M}(n, \theta) = \sum_{n'=-\infty}^{\infty} \frac{\sin(2\pi\Omega(\theta)(n-n'))}{\pi(n-n')} M(n', \theta), \quad (6.24)$$

which is a low-pass filtering in the time direction with θ -dependent cutoff frequency $\Omega(\theta) \triangleq 1/2 - 2|\theta - \theta_0|$. In particular, the cutoff frequency is 0 at the halfband edges $\theta = \theta_0 \pm 1/4$ and 1/2 at the halfband center $\theta = \theta_0$.

2. The inverse Weyl transform in Equation (6.12) is applied to $\widetilde{M}(n, \theta)$:

$$\tilde{g}_{HW}[n, m] = \int_0^{1/2} \widetilde{M}(n, \theta) e^{j4\pi\theta m} d\theta.$$

The function $\tilde{g}_{HW}[n, m]$ is all we need for implementing the halfband Weyl filter (see presently). From $\tilde{g}_{HW}[n, m]$, we can immediately obtain $h_{HW}[2l, 2l'] = \tilde{g}_{HW}[l + l', l - l']$ and $h_{HW}[2l + 1, 2l' + 1] = \tilde{g}_{HW}[l + l' + 1, l - l']$. If desired, the remaining impulse response samples $h_{HW}[2l, 2l' + 1]$ and $h_{HW}[2l + 1, 2l']$ can be determined via the interpolation relation

$$h_{HW}[n, n'] = \sum_{l=-\infty}^{\infty} \sum_{l'=-\infty}^{\infty} \phi[n-2l] \phi[n'-2l'] h_{HW}[2l, 2l']. \quad (6.25)$$

(Such a relation holds for any halfband filter; it shows that the impulse response $h[n, n']$ of a halfband filter is completely determined by the even-indexed samples $h[2l, 2l']$.) We emphasize, however, that the impulse response samples $h_{HW}[2l + 1, 2l' + 1]$, $h_{HW}[2l, 2l' + 1]$ and $h_{HW}[2l + 1, 2l']$ are not required for implementing the halfband Weyl filter.

The preceding method for calculating \mathbf{H}_{HW} is instructive because it shows that halfband consistency of $M(n, \theta)$ corresponds to a specific smoothness with respect to n . An alternative method for calculating \mathbf{H}_{HW} is as follows:

1. Calculate the impulse response samples $h_W[2l, 2l']$ and $h_W[2l+1, 2l'+1]$ of the conventional Weyl filter by means of Equation (6.14) (note that the original TF weight function $M(n, \theta)$ is used).
2. Calculate $h_{HW}[n, n']$ from $h_W[2l, 2l']$ and $h_W[2l+1, 2l'+1]$ by means of the 2-D halfband restriction

$$\begin{aligned} h_{HW}[n, n'] = & \sum_{l=-\infty}^{\infty} \sum_{l'=-\infty}^{\infty} \phi[n-2l] \phi[n'-2l'] h_W[2l, 2l'] \\ & + \sum_{l=-\infty}^{\infty} \sum_{l'=-\infty}^{\infty} \phi[n-2l-1] \phi[n'-2l'-1] h_W[2l+1, 2l'+1]. \end{aligned} \quad (6.26)$$

(In fact, it suffices to calculate $h_W[2l, 2l']$; the remaining impulse response samples are not required for implementing the halfband Weyl filter.)

For $M(n, \theta) \equiv 1$ and $x[n] \in \mathcal{H}$, the halfband Weyl filter satisfies $(\mathbf{H}_{HW}x)[n] = x[n]$.

6.3.3.2 Implementation of the halfband Weyl filter

The 2-D halfband restriction of the impulse response in Equation (6.26) can be reformulated as a halfband restriction of both the input signal and the output signal. The halfband Weyl filter can thus be implemented as follows.

1. Halfband restriction of the input signal; see Equation (6.19):

$$x_{\mathcal{H}}[n] = \sum_{n'=-\infty}^{\infty} \phi[n-n'] x[n'] \in \mathcal{H}. \quad (6.27)$$

2. Application of the minimum-energy Weyl filter to $x_{\mathcal{H}}[n]$; see Equation (6.16):

$$\tilde{y}[n] = \sum_{m=-\infty}^{\infty} \tilde{g}_W[n-m, m] x_{\mathcal{H}}[n-2m]. \quad (6.28)$$

3. Halfband restriction of the output signal; see Equation (6.23):

$$y[n] = \sum_{n'=-\infty}^{\infty} \phi[n-n'] \tilde{y}[n'] \in \mathcal{H} \quad (6.29)$$

Similarly to the input and output halfband Weyl filters, the filtering operation in Equation (6.28) only uses the impulse response samples $\tilde{g}_W[n, m]$ that are specified by Equation (6.14). The overall filter structure is illustrated in [Figure 6.2\(d\)](#). Note that this implementation is efficient because it avoids the computationally intensive “halfband projection” Equation (6.24) or (6.26). The on-line implementation of the halfband Weyl filter is considered in Subsection 6.3.4.

6.3.3.3 Approximate halfband Weyl filter

An even more efficient — though only approximate — implementation of the halfband Weyl filter can be developed. We start by inserting Equation (6.25) into Equation (6.1), which yields the following exact implementation of the halfband Weyl filter:

1. Calculation of the even-indexed samples of the halfband-restricted input signal; see Equation (6.27):

$$x_{\mathcal{H}}[2l] = \sum_{n=-\infty}^{\infty} \phi[2l-n] x[n]. \quad (6.30)$$

2. LTV filtering of $x_{\mathcal{H}}[2l]$ using the even-indexed impulse response samples $h_{HW}[2l, 2l']$:

$$\tilde{y}[2l] = \sum_{l'=-\infty}^{\infty} h_{HW}[2l, 2l'] x_{\mathcal{H}}[2l'] = \sum_{m=-\infty}^{\infty} \tilde{g}_{HW}[2l-m, m] x_{\mathcal{H}}[2(l-m)]. \quad (6.31)$$

3. Halfband interpolation of the output signal:

$$y[n] = \sum_{l=-\infty}^{\infty} \phi[n-2l] \tilde{y}[2l]. \quad (6.32)$$

This “subsampled” scheme appears to be considerably more efficient than the previous algorithm in Equations (6.27) to (6.29). Unfortunately, this is not true because $h_{HW}[2l, 2l']$ is used instead of $h_W[2l, 2l']$, and thus the computationally intensive halfband projection in Equation (6.24) or (6.26) still has to be performed. Omitting this halfband projection, that is, using $h_W[2l, 2l']$ instead of $h_{HW}[2l, 2l']$ in Equation (6.31), yields an *approximate* implementation of the halfband Weyl filter with unrivaled efficiency. This approximate halfband Weyl filter, hereafter denoted $\tilde{\mathbf{H}}_{HW}$, is defined by the three-step procedure in Equations (6.30) to (6.32) with the LTV filtering in Equation (6.31) replaced by

$$\tilde{y}[2l] = \sum_{l'=-\infty}^{\infty} h_W[2l, 2l'] x_{\mathcal{H}}[2l'] = \sum_{m=-\infty}^{\infty} \tilde{g}_W[2l-m, m] x_{\mathcal{H}}[2(l-m)]. \quad (6.33)$$

Note that this requires $\tilde{g}_W[n, m]$ only for n, m both even and both odd. The overall filter structure is illustrated in [Figure 6.2\(e\)](#). The approximate halfband Weyl filter

$\tilde{\mathbf{H}}_{HW}$ is particularly efficient because the LTV filtering in Equation (6.33) runs at half the sampling rate and only uses the impulse response samples that are specified by the original Weyl filter design rule in Equation (6.14). The on-line implementation of $\tilde{\mathbf{H}}_{HW}$ is considered in Subsection 6.3.4.

Because of the halfband prefiltering in Equation (6.30) and the postinterpolation in Equation (6.32), $\tilde{\mathbf{H}}_{HW}$ is a halfband filter just as the halfband Weyl filter \mathbf{H}_{HW} , but it is of course different from \mathbf{H}_{HW} unless $M(n, \theta)$ happens to be halfband consistent beforehand. However, $M(n, \theta)$ is often “almost halfband consistent,” and then $\tilde{\mathbf{H}}_{HW}$ is a good approximation to \mathbf{H}_{HW} . Indeed, this is the case if $M(n, \theta)$ is sufficiently smooth with respect to n , especially near the halfband edges. Thus, if the time variations of $M(n, \theta)$ are not too fast (which usually is true if \mathbf{H}_{HW} is underspread), the approximate halfband Weyl filter $\tilde{\mathbf{H}}_{HW}$ is similar to \mathbf{H}_{HW} .

6.3.4 On-line implementation of (halfband) Weyl filters

We now discuss the efficient on-line implementation of the various Weyl filters and of the halfband Weyl filter (both exact and approximate version).

6.3.4.1 On-line implementation of minimum-energy Weyl filter

The minimum-energy Weyl filter is the basic Weyl filter version. It also represents the algorithmic core of other Weyl filter designs, including the halfband Weyl filter but not the approximate halfband Weyl filter. For an on-line formulation of the input–output relation in Equation (6.16), we define the discrete-frequency TF weight function $M[n, k]$ by sampling $M(n, \theta)$ on the fundamental frequency period $[0, 1/2]$:

$$M[n, k] \triangleq M\left(n, \frac{k}{2K}\right), \quad k = 0, 1, \dots, K-1$$

Here, K is assumed to be even for simplicity, and large enough so that $M(n, \theta)$ is sampled sufficiently densely. In accordance with this frequency discretization, the inverse Weyl transform in Equation (6.14) is replaced by the IDFT relation:

$$\tilde{g}_W[n, m] = \frac{1}{K} \sum_{k=0}^{K-1} M[n, k] e^{j2\pi \frac{k}{K} m}, \quad m = -\frac{K}{2}, -\frac{K}{2} + 1, \dots, \frac{K}{2} - 1 \quad (6.34)$$

Because of the finite m -support of $\tilde{g}_W[n, m]$, Equation (6.16) simplifies to:

$$y[n] = \sum_{m=-K/2}^{K/2-1} \tilde{g}_W[n-m, m] x[n-2m]. \quad (6.35)$$

This involves the input samples $x[n-K+2], x[n-K+4], \dots, x[n+K]$ about the current time point n . By defining the length- K vectors $\tilde{\mathbf{g}}_n \triangleq (\tilde{g}_W[n-K/2+1, K/2-1], \tilde{g}_W[n-K/2+2, K/2-2], \dots, \tilde{g}_W[n+K/2, -K/2])^T$ and $\mathbf{x}_n \triangleq (x[n-K+2]$

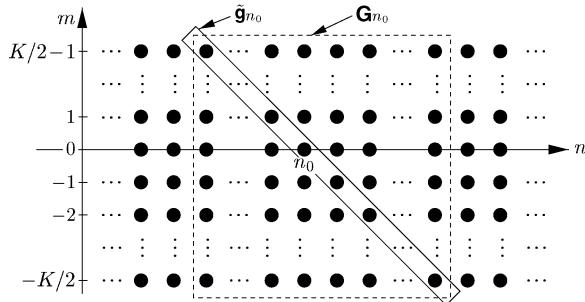
**FIGURE 6.3**

Illustration of the matrix \mathbf{G}_n . The circles represent the samples of $\tilde{g}_W[n, m]$ (corresponding to the impulse response samples $h_W[2l, 2l']$ and $h_W[2l+1, 2l'+1]$). The diagonal $\tilde{\mathbf{g}}_{n_0}$ contains the samples $\tilde{g}_W[n_0 - m, m]$ used by the minimum-energy Weyl filter at time n_0 ; see Equations (6.35) and (6.36).

$x[n - K + 4] \cdots x[n + K]^T$, the input–output relation in Equation (6.35) can be rewritten as

$$y[n] = \tilde{\mathbf{g}}_n^T \mathbf{x}_n. \quad (6.36)$$

The vector $\tilde{\mathbf{g}}_n$ is the main diagonal of a $K \times K$ matrix \mathbf{G}_n whose i th column ($i = 1, 2, \dots, K$) is the length- K IDFT with respect to k of the TF weight function slice $M[n - K/2 + i, k]$; compare with Equation (6.34). This is depicted in Figure 6.3. The on-line implementation of the minimum-energy Weyl filter can then be summarized at each time instant n as follows:

1. Update the matrix \mathbf{G}_n by adding a new column — calculated as the IDFT of $M[n + K/2, k]$ with respect to k according to Equation (6.34) — and removing the oldest column; then obtain the vector $\tilde{\mathbf{g}}_n$ as the main diagonal of \mathbf{G}_n .
2. Compute the output signal sample $y[n]$ according to Equation (6.35) or (6.36).

Because \mathbf{x}_n contains the future input samples $x[n + 2], x[n + 4], \dots, x[n + K]$, causal operation requires the introduction of a time delay of K samples. The computational complexity of the on-line minimum-energy Weyl filter (including calculation of $\tilde{\mathbf{g}}_n$) is as follows. At each time instant n , we perform an IDFT of length K and an inner product of two length- K vectors. This requires $\mathcal{O}(K + K \log K)$ operations per output sample, which is equivalent to the on-line Zadeh filter discussed in Sub-section 6.2.2. Because of the delay of K samples, we have to store the K past vectors $\tilde{\mathbf{g}}_{n-1}, \tilde{\mathbf{g}}_{n-2}, \dots, \tilde{\mathbf{g}}_{n-K}$. Thus, a total of $K^2 + 2K$ memory locations are required (K length- K vectors $\tilde{\mathbf{g}}_i$ and $2K$ samples contained in[†] \mathbf{x}_n and \mathbf{x}_{n-1}). These memory

[†]We need to store both \mathbf{x}_n and \mathbf{x}_{n-1} because each vector contains only every second sample of $x[n]$.

requirements can, however, be reduced as follows. By substituting $n + K - i$ for n in Equation (6.35), we obtain

$$y[n+K-i] = \sum_{m=-K/2}^{K/2-1} \tilde{g}_W[n+K-i-m, m] x[n+K-i-2m]. \quad (6.37)$$

Consider now the following “partial sums” of Equation (6.37):

$$s_i[n] \triangleq \sum_{m=K/2-i}^{K/2-1} \tilde{g}_W[n+K-i-m, m] x[n+K-i-2m], \quad i = 1, 2, \dots, K.$$

We have $y[n] = s_K[n]$ (i.e., the output signal $y[n]$ is obtained as the K th partial sum). Furthermore, the partial sums can be calculated in a recursive manner according to:

$$s_i[n] = \begin{cases} s_{i-1}[n-1] + \tilde{g}_W[n+\frac{K}{2}, \frac{K}{2}-i] x[n+i], & i = 2, 3, \dots, K \\ \tilde{g}_W[n+\frac{K}{2}, \frac{K}{2}-1] x[n+1], & i = 1. \end{cases}$$

For causal operation, the definition and recursive calculation of the partial sums have to be modified by introducing a delay of K samples. Accordingly, at time n we have to store the $K/2$ past length- K impulse response slices $\tilde{g}_W[n-K/2, m], \tilde{g}_W[n-K/2+1, m], \dots, \tilde{g}_W[n-1, m]$ (with $m = -K/2, -K/2+1, \dots, K/2-1$) in addition to the length- K partial-sum vector ($s_1[n] s_2[n] \dots s_K[n]$) and the length- K input signal vector ($x[n-K+1] x[n-K+2] \dots x[n]$). This amounts to a total of $K^2/2 + 2K$ memory locations, which is about half the memory required for the direct implementation described earlier. The computational complexity (number of operations) is not reduced, however.

6.3.4.2 On-line implementation of input halfband, output halfband and halfband Weyl filter

According to Subsection 6.3.2, the input (output) halfband Weyl filter is the minimum-energy Weyl filter preceded (succeeded) by a halfband restriction of the input (output) signal. Similarly, according to Equations (6.27) to (6.29), the halfband Weyl filter is the minimum-energy Weyl filter both preceded and succeeded by a halfband restriction. Thus, these three filters can use the on-line implementation of the minimum-energy Weyl filter discussed earlier. We merely require additional FIR or IIR filters[§] that approximate the halfband restriction, that is, the idealized low-pass or bandpass filter defined by Equation (6.18).

6.3.4.3 On-line implementation of approximate halfband Weyl filter

Finally, we discuss the efficient on-line implementation of the approximate halfband Weyl filter $\tilde{\mathbf{H}}_{HW}$. The LTV filtering in Equation (6.33) involves $\tilde{g}_W[n, m]$ only for

[§]Here, group delay distortions must be kept small because they affect the TF structure of the signal.

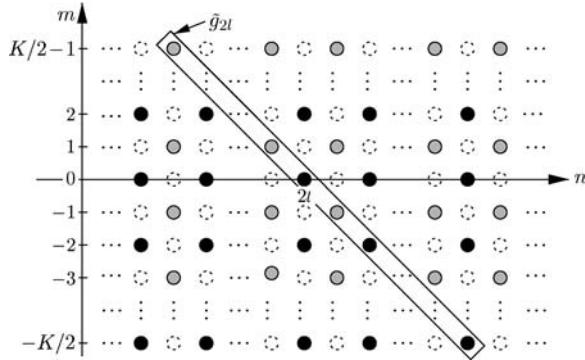


FIGURE 6.4

Illustration of the matrices $\mathbf{G}_{2l}^{(e)}$ and $\mathbf{G}_{2l}^{(o)}$. The black circles represent the elements of $\mathbf{G}_{2l}^{(e)}$ (corresponding to the impulse response samples $\tilde{g}_W[2l-2i, 2i]$), and the shaded circles represent the elements of $\mathbf{G}_{2l}^{(o)}$ (corresponding to $\tilde{g}_W[2l-2i-1, 2i+1]$). The diagonal \tilde{g}_{2l} contains the samples $\tilde{g}_W[2l-m, m]$ used by the approximate halfband Weyl filter at time $n = 2l$; see Equations (6.40) and (6.41). The dotted circles indicate the samples of $\tilde{g}_W[n, m]$ that do not need to be computed (these correspond to $h_W[2l+1, 2l'+1]$).

n, m both even and both odd (see Figure 6.4). However, the IDFT in Equation (6.34) yields all samples $\tilde{g}_W[n, m]$. This inefficiency can be avoided by using the following decimation-in-frequency type IDFT relations [44]. For simplicity, we assume that K is a multiple of 4. Then, for $n = 2l$ even, there is

$$\tilde{g}_W[2l, 2i] = \frac{2}{K} \sum_{k=0}^{K/2-1} M_e[2l, k] e^{j2\pi \frac{k}{K/2} i}, \quad i = -\frac{K}{4}, -\frac{K}{4}+1, \dots, \frac{K}{4}-1, \quad (6.38)$$

with $M_e[n, k] \triangleq M[n, k] + M[n, k+K/2]$. For $n = 2l+1$ odd one has

$$\tilde{g}_W[2l+1, 2i+1] = \frac{2}{K} \sum_{k=0}^{K/2-1} M_o[2l+1, k] e^{j2\pi \frac{k}{K/2} i}, \quad i = -\frac{K}{4}, -\frac{K}{4}+1, \dots, \frac{K}{4}-1, \quad (6.39)$$

with $M_o[n, k] \triangleq (M[n, k] - M[n, k+K/2]) e^{j2\pi \frac{k}{K}}$. Thus, at each time instant n only a length- $K/2$ IDFT is needed instead of the length- K IDFT in Equation (6.34).

Because of the finite m -support of $\tilde{g}_W[n, m]$, the LTV filtering operation in Equation (6.33) reduces to

$$\tilde{y}[2l] = \sum_{m=-K/2}^{K/2-1} \tilde{g}_W[2l-m, m] x_{\mathcal{H}}[2(l-m)], \quad (6.40)$$

which involves the subsampled, halfband-restricted input samples $x_{\mathcal{H}}[2(l-K/2+1)], x_{\mathcal{H}}[2(l-K/2+2)], \dots, x_{\mathcal{H}}[2(l+K/2)]$ about the current time point $n = 2l$.

By using the length- K vectors $\tilde{\mathbf{g}}_{2l} \triangleq (\tilde{g}_W[2l - K/2 + 1, K/2 - 1] \tilde{g}_W[2l - K/2 + 2, K/2 - 2] \cdots \tilde{g}_W[2l + K/2, -K/2])^T$ and $\mathbf{x}_{\mathcal{H},2l} \triangleq (x_{\mathcal{H}}[2(l - K/2 + 1)] x_{\mathcal{H}}[2(l - K/2 + 2)] \cdots x_{\mathcal{H}}[2(l + K/2)])^T$, Equation (6.40) can be written as

$$\tilde{y}[2l] = \tilde{\mathbf{g}}_{2l}^T \mathbf{x}_{\mathcal{H},2l}. \quad (6.41)$$

The vector $\tilde{\mathbf{g}}_{2l}$ can be interpreted as follows. Define the $K/2 \times K/2$ matrix $\mathbf{G}_{2l}^{(e)}$ whose i th column ($i = 1, 2, \dots, K/2$) is the length- $K/2$ IDFT with respect to k of $M_e[2(l - K/4 + i), k]$; compare with Equation (6.38). Similarly, define the $K/2 \times K/2$ matrix $\mathbf{G}_{2l}^{(o)}$ whose i th column is the length- $K/2$ IDFT with respect to k of $M_o[2(l - K/4 + i) - 1, k]$; see Equation (6.39). Then, the even-indexed elements of $\tilde{\mathbf{g}}_{2l}$ are the main diagonal of $\mathbf{G}_{2l}^{(e)}$ and the odd-indexed elements of $\tilde{\mathbf{g}}_{2l}$ are the main diagonal of $\mathbf{G}_{2l}^{(o)}$; compare with Figure 6.4. The on-line implementation of the approximate halfband Weyl filter can thus be summarized at each subsampled time instant $n = 2l$ as follows:

1. Update the matrices $\mathbf{G}_{2l}^{(e)}$ and $\mathbf{G}_{2l}^{(o)}$ by adding a new column (calculated by means of length- $K/2$ FFTs according to Equations (6.38) and (6.39)) and removing the oldest column; then obtain the vector $\tilde{\mathbf{g}}_{2l}$ from the main diagonals of $\mathbf{G}_{2l}^{(e)}$ and $\mathbf{G}_{2l}^{(o)}$ as explained above.
2. Calculate an approximation to the halfband-decimated input signal sample $x_{\mathcal{H}}[2l]$ in Equation (6.30) using an FIR or IIR filter [17].
3. Compute the intermediate output signal sample $\tilde{y}[2l]$ according to Equation (6.40) or (6.41).
4. Calculate an approximation to the halfband-interpolated output signal sample $y[n]$ in Equation (6.32) using an FIR or IIR interpolation filter [17].

For causality, we have to introduce a time delay of K samples, plus additional delays caused by the halfband predecimation and postinterpolation. At each subsampled time instant $n = 2l$, the algorithm requires two length- $K/2$ IDFTs and an inner product of two length- K vectors. Thus, the computational complexity (including the calculation of $\tilde{\mathbf{g}}_{2l}$ but not the predecimation and postinterpolation) is $\mathcal{O}(K/2 + K/2 \log(K/2))$ operations per output sample, which is significantly less than for the exact on-line halfband Weyl filter. Because of the delay of K samples, we need to store the $K/2$ past vectors $\tilde{\mathbf{g}}_{2l-2}, \tilde{\mathbf{g}}_{2l-4}, \dots, \tilde{\mathbf{g}}_{2l-K}$. The total number of memory locations is $K^2/2 + K$ (i.e., $K/2$ length- K vectors $\tilde{\mathbf{g}}_{2l}$ and the length- K vector $\mathbf{x}_{\mathcal{H},2l}$). These memory requirements can again be reduced by a recursive calculation. By substituting $l + K/2 - i$ for l in Equation (6.40), we obtain

$$\tilde{y}\left[2\left(l + \frac{K}{2} - i\right)\right] = \sum_{m=-K/2}^{K/2-1} \tilde{g}_W\left[2\left(l + \frac{K}{2} - i\right) - m, m\right] x_{\mathcal{H}}\left[2\left(l + \frac{K}{2} - i - m\right)\right]. \quad (6.42)$$

We now use the following “partial sums” of Equation (6.42):

$$s_i[2l] \triangleq \sum_{m=K/2-2i}^{K/2-1} \tilde{g}_W \left[2\left(l+\frac{K}{2}-i\right)-m, m \right] x_H \left[2\left(l+\frac{K}{2}-i-m\right) \right]$$

where $i = 1, 2, \dots, K/2$. The intermediate output signal in Equation (6.40) is obtained as the $K/2$ th partial sum $\tilde{y}[2l] = s_{K/2}[2l]$. Furthermore, the partial sums can be calculated recursively as

$$s_i[2l] = \begin{cases} s_{i-1}[2(l-1)] + \tilde{g}_W \left[2l+\frac{K}{2}, \frac{K}{2}-2i \right] x_H[2(l+i)] \\ \quad + \tilde{g}_W \left[2l+\frac{K}{2}-1, \frac{K}{2}-2i+1 \right] \\ \quad \times x_H[2(l+i-1)], & i = 2, 3, \dots, K/2 \\ \tilde{g}_W \left[2l+\frac{K}{2}, \frac{K}{2}-2 \right] x_H[2(l+1)] \\ \quad + \tilde{g}_W \left[2l+\frac{K}{2}-1, \frac{K}{2}-1 \right] x_H[2l], & i = 1. \end{cases}$$

For causal operation, we again have to introduce a delay of K samples. Accordingly, at time $2l$ we have to store the $K/2$ past length- $K/2$ impulse response slices $\tilde{g}_W[2l-K/2-1, m], \tilde{g}_W[2l-K/2, m], \dots, \tilde{g}_W[2l-2, m]$, the length- $K/2$ partial-sum vector $(s_1[2l] \ s_2[2l] \ \dots \ s_{K/2}[2l])$, and the length- $K/2$ input vector $(x_H[2(l-K/2)] \ x_H[2(l-K/2+1)] \ \dots \ x_H[2(l-1)])$. This amounts to $K^2/4 + K$ memory locations, which is about half the memory required previously.

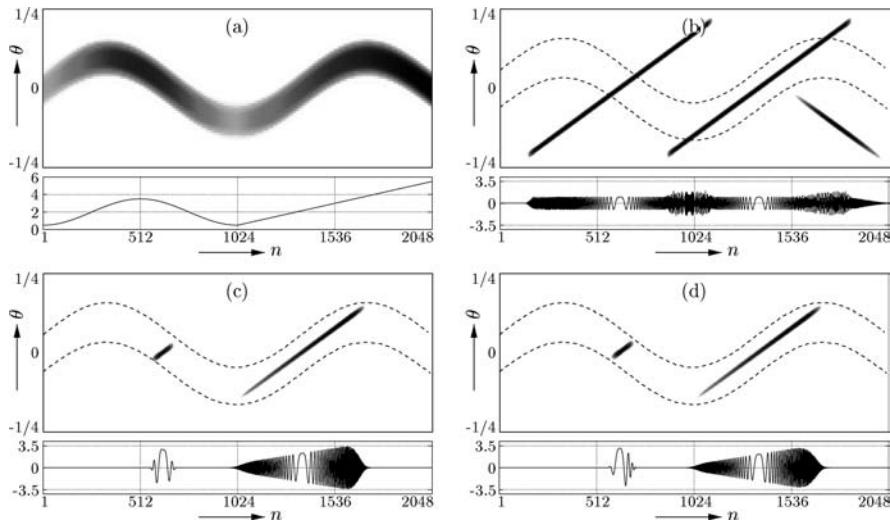
6.4 Explicit Time–Frequency Filter Design III: Simulation Results and Applications

This section presents simulation results for the on-line implementations of the Zadeh filter \mathbf{H}_Z (see [Subsection 6.2.2](#)) and of the approximate halfband Weyl filter $\tilde{\mathbf{H}}_{HW}$ (see [Subsection 6.3.4](#)). The applications considered include the separation and denoising of chirp signals, the generation of nonstationary random processes and FM demodulation. We demonstrate advantages of $\tilde{\mathbf{H}}_{HW}$ over \mathbf{H}_Z and problems caused by non-smooth (oscillatory) TF weight functions.

6.4.1 Time-varying bandpass filter

We first design an LTV bandpass filter with time-varying center frequency and gain. The TF weight function $M(n, \theta)$ (discretized with $K = 64$ frequency bins) is shown in [Figure 6.5\(a\)](#). The input signal $x[n]$ of length 2048 samples consists of three linear

^TThese slices have length $K/2$ because for any given slice, m assumes only even or odd values; see Equations (6.38) and (6.39) and [Figure 6.4](#).

**FIGURE 6.5**

Time-varying bandpass filtering. (a) TF weight function (lower part: time-varying gain); (b) input signal (upper part: SPWD, lower part: real part); (c) output signal of Zadeh filter; (d) output signal of approximate halfband Weyl filter. In (b) to (d), the dashed lines indicate the TF pass region defined by the TF weight function.

chirp components. A smoothed pseudo-Wigner distribution (SPWD) [45–47] and the real part of $x[n]$ are shown in Figure 6.5(b).

The output signals of \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ are shown in Figure 6.5(c) and (d). It is seen that both filters succeed in implementing the desired TF weighting as specified by $M(n, \theta)$: they extract and amplify the central part of the first chirp component, multiply the central part of the second chirp component by a linearly increasing gain, and suppress the third chirp component. The similarity of the two output signals (the normalized energy of their difference is 0.036 or -14.37 dB) is due to the fact that $M(n, \theta)$ is smooth and thus both filters are underspread [9].

6.4.2 Separation and denoising of chirp signals

Our next example shows that the halfband Weyl filter $\tilde{\mathbf{H}}_{HW}$ is superior to the Zadeh filter \mathbf{H}_Z in the case of obliquely oriented (chirpy) TF weight functions. The input signal $x[n]$ of length 256 samples consists of two closely spaced, windowed chirp components $x_i[n] = w_i[n - n_i] e^{j[\pi c(n - n_i)^2 + 2\pi\theta_i(n - n_i)]}$ with chirp rate $c = 0.7 / 256 = 2.73 \cdot 10^{-3}$; see Figure 6.6(a). The filtering task is to suppress the upper chirp component while passing the lower chirp component; the latter is shown in Figure 6.6(b). The TF weight function $M(n, \theta)$ (discretized with $K = 256$ frequency bins) is depicted in Figure 6.6(c). From the filter output signals shown in Figure 6.6(d) and (e), it is seen that $\tilde{\mathbf{H}}_{HW}$ suppresses the upper component much better than \mathbf{H}_Z . Indeed, the signal-to-noise ratio (SNR) improvements achieved by \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ are

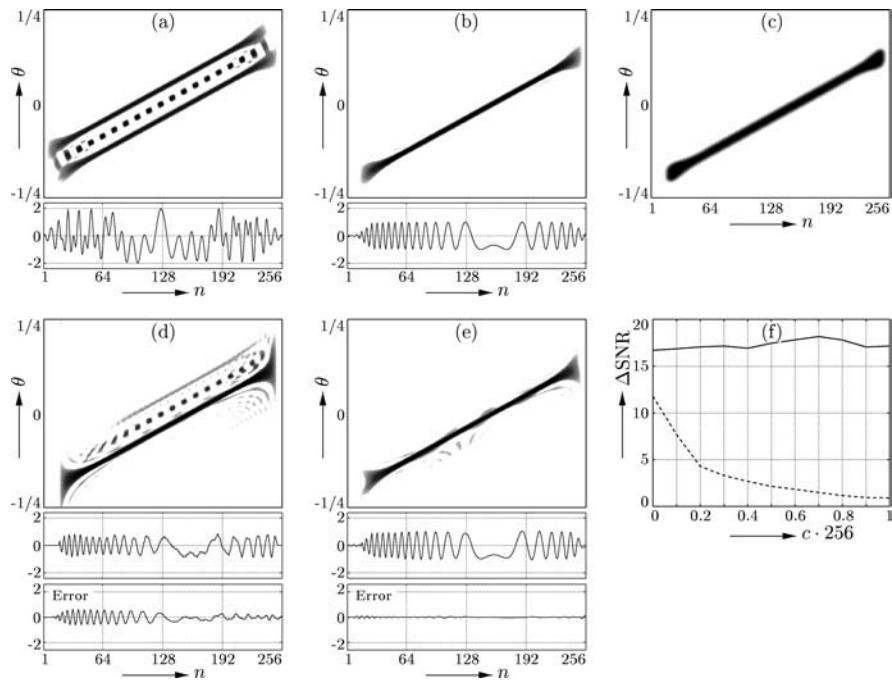
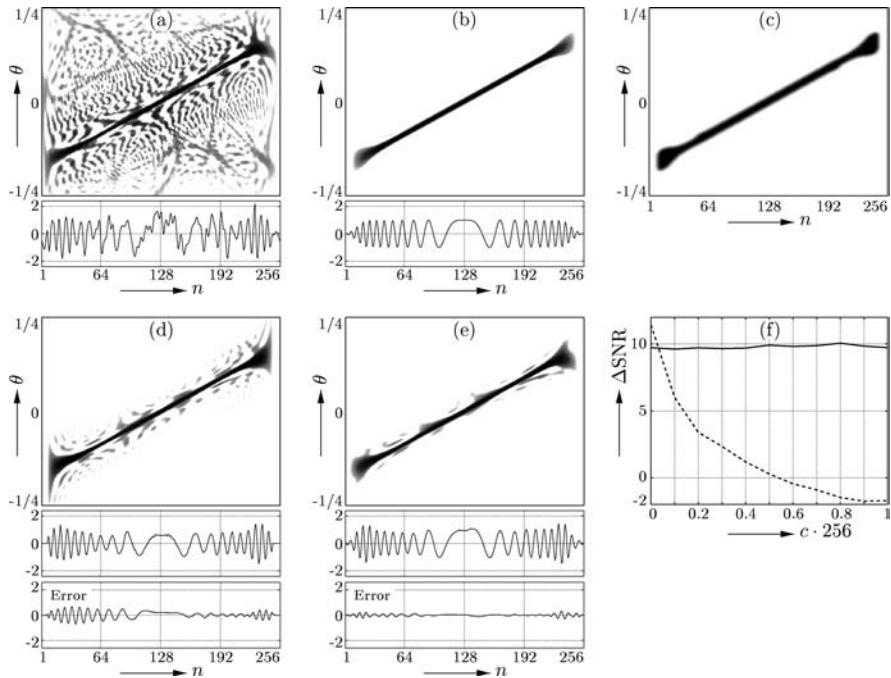


FIGURE 6.6

Separation of chirp signals. (a) Input signal (upper part: SPWD, lower part: real part); (b) desired signal component; (c) TF weight function; (d) output signal of Zadeh filter; (e) output signal of approximate halfband Weyl filter; (f) SNR improvement in dB vs. chirp rate parameter $c \cdot 256$ (solid line: approximate halfband Weyl filter, dashed line: Zadeh filter). The bottom parts of (d) and (e) show the respective error signal (filter output signal minus desired signal).

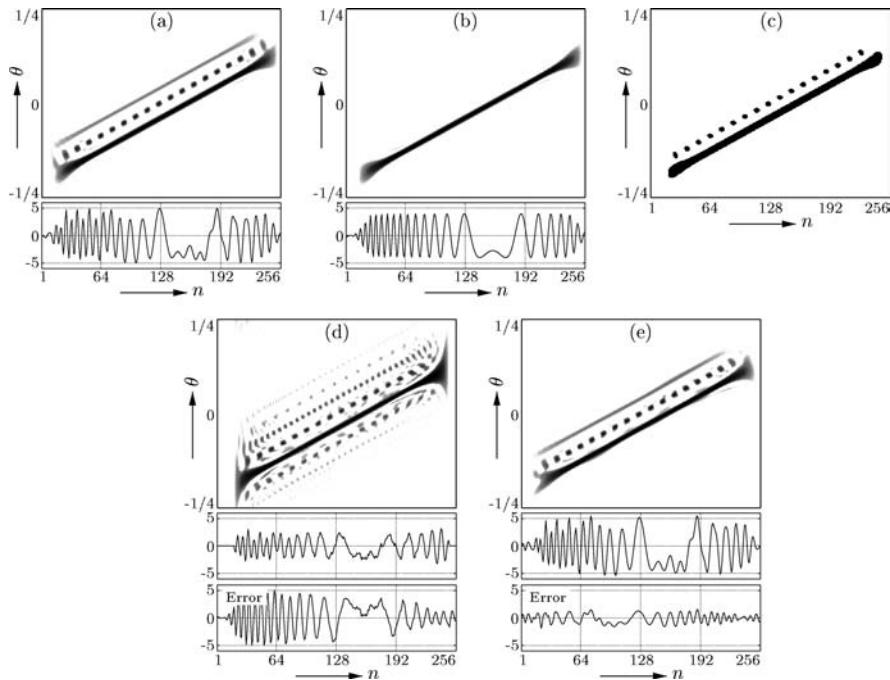
1.49 and 18.14 dB, respectively. (The definition of the SNR is based on the deviation from the desired lower chirp component.) This poor performance of \mathbf{H}_Z is due to the oblique orientation of the TF weight function $M(n, \theta)$. To further analyze this effect, we repeated this experiment for various chirp rates c between 0 and 1/256. The resulting SNR improvements are shown in Figure 6.6(f). Whereas the performance of $\tilde{\mathbf{H}}_{HW}$ is almost independent of c , the performance of \mathbf{H}_Z degrades rapidly with increasing c .

Figure 6.7 shows analogous results for a noise-contaminated input signal $x[n] = s[n] + w[n]$, where $s[n]$ is a chirp signal with chirp rate $c = 0.7/256$ and $w[n]$ is halfband-restricted white noise of variance 1/2 (corresponding to an SNR of $\frac{\|s\|^2}{256} / E\{|w[n]|^2\} = 3$ dB). The TF weight function $M(n, \theta)$ was discretized with $K = 256$ frequency bins. It is seen that $\tilde{\mathbf{H}}_{HW}$ suppresses the noise much better than \mathbf{H}_Z : $\tilde{\mathbf{H}}_{HW}$ achieves an SNR improvement of 9.87 dB whereas \mathbf{H}_Z even degrades the SNR by 0.91 dB. It is furthermore seen that, again, the SNR improvement of $\tilde{\mathbf{H}}_{HW}$ is independent of the chirp rate c whereas that of \mathbf{H}_Z rapidly decreases with increasing c .

**FIGURE 6.7**

Denoising of a chirp signal. (a) Input signal (upper part: SPWD, lower part: real part); (b) desired signal component; (c) TF weight function; (d) output signal of Zadeh filter; (e) output signal of approximate halfband Weyl filter; (f) SNR improvement in dB vs. chirp rate parameter $c \cdot 256$ (solid line: approximate halfband Weyl filter, dashed line: Zadeh filter).

We finally consider a nonsmooth TF weight function with oscillatory components. The input signal $x[n]$ consists of two chirp components — see Figure 6.8(a) — of which the lower (stronger) one — see Figure 6.8(b) — is to be extracted. The TF weight function $M[n, k]$ used $K = 256$ frequency bins; it was obtained by thresholding an SPWD of $x[n]$ (i.e., all SPWD values above threshold were set equal to 1 in $M[n, k]$). It is seen from Figure 6.8(c) that $M[n, k]$ contains oscillatory components resulting from SPWD cross-terms [45–47] above threshold. These components cause both \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ to be overspread (i.e., to introduce undesired TF shifts). Indeed, both output signals — see Figure 6.8(d) and (e) — contain one or more parasitic components in addition to the desired component. These parasitic components are not the results of the upper (weak) chirp component being passed by the filter (note that $M[n, k]$ is zero on the corresponding TF support region); instead, they are TF-shifted versions of the desired lower component. The output SNR of \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ was 11.12 dB and 1.56 dB, respectively, worse than the input SNR. We conclude that oscillatory TF weight functions lead to undesired TF shift effects in \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$. Furthermore, \mathbf{H}_Z is again inferior to $\tilde{\mathbf{H}}_{HW}$ due to the chirpy TF weight function.

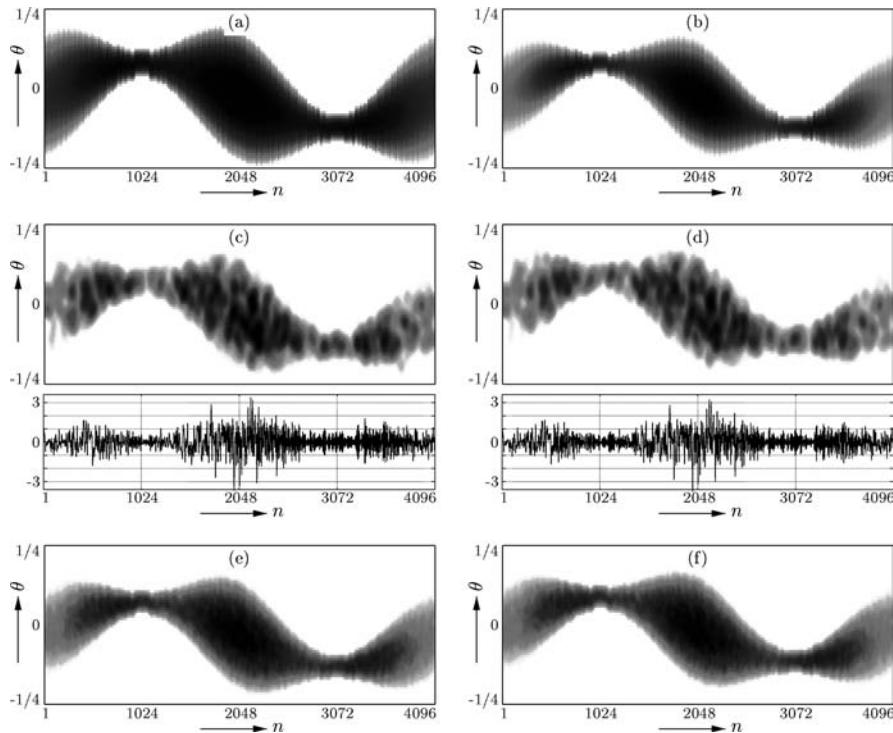
**FIGURE 6.8**

Separation of chirp signals using a nonsmooth (oscillatory) TF weight function. (a) Input signal (upper part: SPWD, lower part: real part); (b) desired signal component; (c) TF weight function; (d) output signal of Zadeh filter; (e) output signal of approximate halfband Weyl filter.

6.4.3 Generation of nonstationary random processes

Next, we consider the use of TF filters for generating a nonstationary random process $y[n]$ according to $y[n] = (\mathbf{H}x)[n]$, where $x[n]$ is stationary white noise with unit variance. The correlation function of $y[n]$ is determined by the innovations system [48] \mathbf{H} according to $E\{y[n_1]y^*[n_2]\} = \sum_{n=-\infty}^{\infty} h[n_1, n]h^*[n_2, n]$. For \mathbf{H} we used the on-line TF filters \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$. The TF weight function $M(n, \theta)$ (discretized with $K = 64$ frequency bins) is shown in Figure 6.9(a). Due to the construction of $y[n]$, $M(n, \theta)$ determines the mean TF energy distribution of $y[n]$. In fact, one can expect that a suitably defined time-varying power spectrum of $y[n]$ (e.g., the Wigner–Ville spectrum [45, 49–53]) is approximately equal to the squared TF weight function $M^2(n, \theta)$ depicted in Figure 6.9(b) [35, 36, 52, 54].

Figure 6.9(c) and (d) shows the output signals of \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ obtained for the same realization of the white noise input signal $x[n]$. The two output signals are seen to be very similar. Indeed, their normalized mean square difference (averaged over 100 realizations) was 0.057 or -12.4 dB. Furthermore, we estimated the Wigner–Ville spectra of the nonstationary random processes $(\mathbf{H}_Z x)[n]$ and $(\tilde{\mathbf{H}}_{HW} x)[n]$ by averaging the SPWDs of 100 realizations. From Figure 6.9(e) and (f), the two estimated spectra are seen to be practically identical (their normalized difference energy was 0.064

**FIGURE 6.9**

Generation of a nonstationary random process using Zadeh and Weyl innovations filters. (a) TF weight function; (b) squared TF weight function; (c) output signal of Zadeh filter (upper part: SPWD, lower part: real part); (d) output signal of approximate halfband Weyl filter; (e) estimated Wigner–Ville spectrum of $(\mathbf{H}_Z x)[n]$; (f) estimated Wigner–Ville spectrum of $(\tilde{\mathbf{H}}_{HW} x)[n]$.

or -11.91 dB); moreover, as expected, they are very similar to $M^2(n, \theta)$ in Figure 6.9(b). Again, as in Subsection 6.4.1, we can conclude that in the case of smooth TF weight functions, the Zadeh filter and Weyl filter yield very similar results.

6.4.4 FM demodulation

In our last application example, we consider a wideband frequency modulation (FM) communication system [55] operating at low SNR. The idea is to replace a conventional time-invariant receive bandpass filter by an adaptive TF filter to achieve better noise suppression. The FM transmit signal was formed by frequency-modulating a speech signal with bandwidth 11 kHz using modulation index 5. The resulting FM signal is shown in Figure 6.10(a); its effective analog bandwidth^{||} is 110 kHz. The FM signal was corrupted by white noise with an SNR of -3 dB. The

^{||}The FM modulation and subsequent signal processing were done in discrete time based on a sampled version of the speech signal. The effective analog bandwidth is stated to establish a relation to real-world situations.

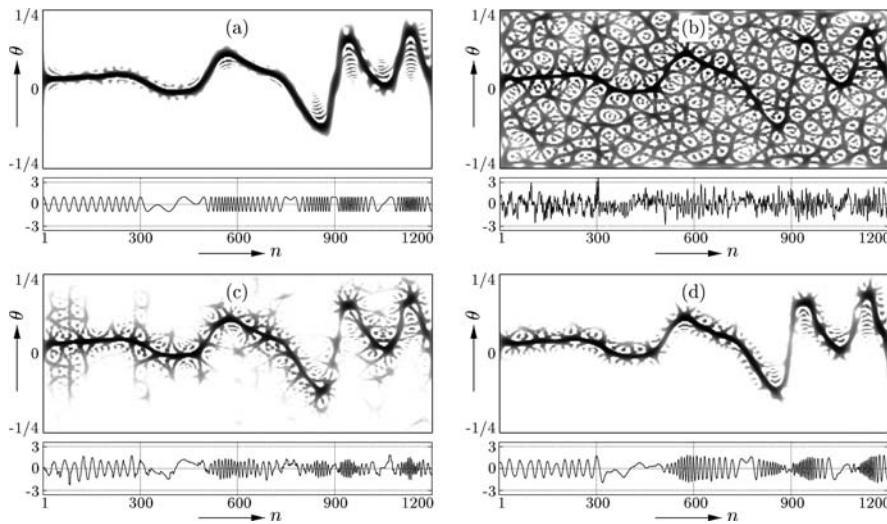


FIGURE 6.10

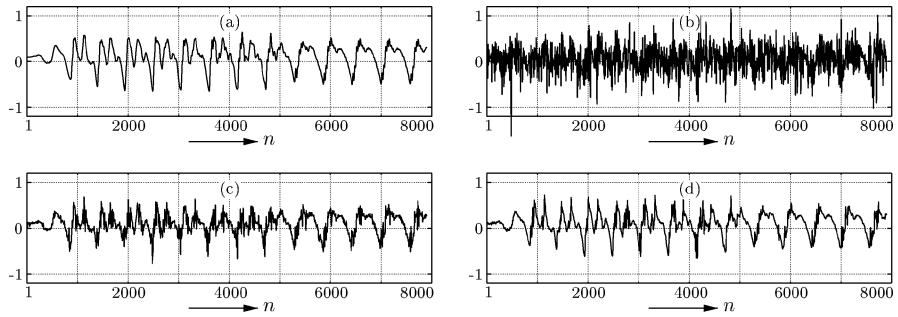
Adaptive TF filtering of an FM signal. (a) Noiseless FM signal (upper part: SPWD, lower part: real part); (b) to (d) filtered noisy FM signals: (b) using time-invariant bandpass filter, (c) using adaptive Zadeh filter, (d) using adaptive approximate halfband Weyl filter.

noisy FM signal was then filtered by (1) a time-invariant bandpass filter of bandwidth 110 kHz; (2) an adaptive version of \mathbf{H}_Z ; and (3) an adaptive version of $\tilde{\mathbf{H}}_{HW}$. Finally, each filter output signal was demodulated by means of the MATLAB function **ademodce** [56].

The TF weight function $M[n, k]$ used $K = 64$ frequency bins. At each time instant n , it was derived from the noisy FM signal as $M[n, k_0[n]] = 1$, $M[n, k_0[n] \pm 1] = 1/2$, and $M[n, k] = 0$ for $k \notin \{k_0[n] - 1, k_0[n], k_0[n] + 1\}$, where $k_0[n]$ is the discrete frequency location of the maximum of an SPWD of the noisy FM signal at time n . This means that \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ were designed from the noisy FM signal in an adaptive, on-line manner without using prior information about the TF structure of the clean FM signal.

The filter output signals are shown in Figure 6.10(b) to (d). The SNR improvements achieved by \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ are better by 2.65 and 4.46 dB, respectively, than the SNR improvement achieved by the time-invariant bandpass filter. Note that $\tilde{\mathbf{H}}_{HW}$ performs better than \mathbf{H}_Z because large parts of the TF weight function $M[n, k]$ are obliquely oriented (compare with Subsection 6.4.2).

The resulting demodulated signals are shown for an extended signal duration of 8000 samples in Figure 6.11. It is seen that due to the low input SNR of -3 dB, FM demodulation using the output of the time-invariant bandpass filter produces very poor results (the demodulator output SNR is only -0.96 dB). The results achieved with \mathbf{H}_Z and $\tilde{\mathbf{H}}_{HW}$ are seen to be much better (the demodulator output SNRs are 6.06 and 8.56 dB, respectively).

**FIGURE 6.11**

Results of FM demodulation. (a) Original speech signal; (b) to (d) demodulated signals: (b) using time-invariant bandpass filter, (c) using adaptive Zadeh filter, (d) using adaptive approximate halfband Weyl filter.

6.5 Implicit Time–Frequency Filter Design I: Short-time Fourier Transform Filter

In this section and the next two sections, we consider the implicit TF design of LTV filters. Here, as explained in Subsection 6.1.1, an LTV filter is implemented — and thereby designed in an implicit manner — by means of an analysis-weighting-synthesis procedure that involves a linear TF signal representation. In this section, the TF signal representation is chosen as the *short-time Fourier transform* (STFT). The resulting LTV filter is termed *STFT filter*.

6.5.1 Short-time Fourier transform and Short-time Fourier transform filter

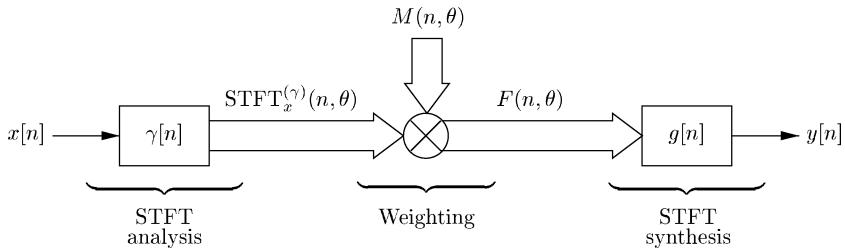
6.5.1.1 STFT

The STFT of a signal $x[n]$ is defined as [15, 16, 45, 47, 57]

$$\text{STFT}_x^{(\gamma)}(n, \theta) \triangleq \sum_{n'=-\infty}^{\infty} x[n'] \gamma^*[n'-n] e^{-j2\pi\theta(n'-n)}, \quad (6.43)$$

where $\gamma[n]$ is a suitable analysis window. The squared magnitude of the STFT can be interpreted as a measure of the amount of energy that the signal $x[n']$ has about the TF analysis point (n, θ) . The STFT can be inverted as [15, 16, 20, 45, 47, 57]

$$x[n] = \sum_{n'=-\infty}^{\infty} \int_0^1 \text{STFT}_x^{(\gamma)}(n', \theta) g[n-n'] e^{j2\pi\theta(n-n')} d\theta, \quad (6.44)$$

**FIGURE 6.12**

Block diagram of an STFT filter.

provided that the synthesis window $g[n]$ and the analysis window $\gamma[n]$ satisfy the perfect-reconstruction condition

$$\langle g, \gamma \rangle \triangleq \sum_{n=-\infty}^{\infty} g[n] \gamma^*[n] = 1. \quad (6.45)$$

If $g[n] = \gamma[n]$, then Equation (6.45) simply means that $\gamma[n]$ is normalized (i.e., $\|\gamma\|^2 = \sum_{n=-\infty}^{\infty} |\gamma[n]|^2 = 1$). Note that Equation (6.45) can always be achieved by a simple scaling of either window unless $\langle g, \gamma \rangle = 0$.

6.5.1.2 STFT filter

The STFT filter procedure consists of the following three steps that are illustrated in Figure 6.12 [8, 9, 15, 16, 18–20]:

- Analysis.* The STFT of the input signal $x[n]$ is calculated according to Equation (6.43).
- Weighting.* The STFT of $x[n]$ is multiplied by the prescribed TF weight function $M(n, \theta)$:

$$F(n, \theta) \triangleq M(n, \theta) \text{STFT}_x^{(\gamma)}(n, \theta).$$

The resulting TF function $F(n, \theta)$ corresponds in a certain sense to the desired STFT of the filter output signal $y[n]$; however, in general $F(n, \theta)$ is not a valid STFT of any signal.

- Synthesis.* Finally, the filter output signal $y[n]$ is calculated by applying the STFT inversion relation** in Equation (6.44) to $F(n, \theta)$

$$y[n] = \sum_{n'=-\infty}^{\infty} \int_0^1 F(n', \theta) g[n-n'] e^{j2\pi\theta(n-n')} d\theta.$$

**This may appear somewhat arbitrary because $F(n, \theta)$ is not a valid STFT. However, if Equation (6.45) is satisfied, then $y[n]$ is the signal whose STFT (with analysis window $g[n]$) best matches $F(n, \theta)$ in the least-squares sense [20].

Because all three processing steps are linear, the overall procedure amounts to an LTV filter — hereafter called STFT filter and denoted \mathbf{H}_S — that is designed implicitly during the filtering. Note that the STFT filter depends not only on the TF weight function $M(n, \theta)$ but also on the windows $\gamma[n]$ and $g[n]$. The impulse response of \mathbf{H}_S is obtained as

$$h_S[n_1, n_2] = \sum_{n=-\infty}^{\infty} \int_0^1 M(n, \theta) g[n_1 - n] \gamma^*[n_2 - n] e^{j2\pi\theta(n_1 - n_2)} d\theta.$$

If the windows satisfy Equation (6.45), then the STFT filter satisfies the perfect-reconstruction property, that is, $(\mathbf{H}_S x)[n] = x[n]$ for $M(n, \theta) \equiv 1$.

6.5.1.3 Comparison of STFT filter with Zadeh filter

It can be shown [8, 9, 35] that the Zadeh function of \mathbf{H}_S is given by the convolution relation^{††}

$$\begin{aligned} Z_{\mathbf{H}_S}(n, \theta) &= M(n, \theta) \ast R_{g, \gamma}(n, \theta) \\ &= \sum_{n'=-\infty}^{\infty} \int_0^1 M(n', \theta') R_{g, \gamma}(n - n', \theta - \theta') d\theta', \end{aligned} \quad (6.46)$$

where $R_{g, \gamma}(n, \theta) \triangleq \sum_{m=-\infty}^{\infty} g[n] \gamma^*[n - m] e^{-j2\pi\theta m} = g[n] \Gamma^*(\theta) e^{-j2\pi\theta n}$ is the cross Rihaczek distribution of $g[n]$ and $\gamma[n]$ [45–47, 58–60]. By comparing Equation (6.46) with Equation (6.6), we see that the STFT filter \mathbf{H}_S associated to the TF weight function $M(n, \theta)$ is equivalent to the Zadeh filter \mathbf{H}_Z associated to the modified TF weight function $\widetilde{M}(n, \theta) \triangleq M(n, \theta) \ast R_{g, \gamma}(n, \theta)$. For typical windows $\gamma[n]$ and $g[n]$, $R_{g, \gamma}(n, \theta)$ is a smooth function that is concentrated about $n = 0$ and $\theta = 0$. Thus, $\widetilde{M}(n, \theta)$ is a smoothed version of $M(n, \theta)$. If $M(n, \theta)$ is itself sufficiently smooth and if the windows $\gamma[n]$, $g[n]$ are chosen appropriately, then $\widetilde{M}(n, \theta) \approx M(n, \theta)$, which means that the STFT filter \mathbf{H}_S associated to $M(n, \theta)$ approximates the Zadeh filter \mathbf{H}_Z associated to the same TF weight function $M(n, \theta)$. This is further considered in Subsection 6.5.2.

6.5.1.4 Underspread property of the STFT filter

Because $Z_{\mathbf{H}_S}(n, \theta)$ is a smoothed version of $M(n, \theta)$, it tends to be a smooth function even if $M(n, \theta)$ is not smooth. From our discussion in Subsection 6.1.2, it then follows that \mathbf{H}_S is inherently an *underspread* filter. This is a major difference from the Weyl filter and the Zadeh filter that is underspread only if $M(n, \theta)$ is smooth. This result

^{††}A similar relation exists for the Weyl symbol of \mathbf{H}_S . However, because the Weyl symbol is a nonaliased TF representation only for halfband filters, this relation is not considered here.

can also be inferred from the following expression for the spreading function — see Equation (6.2) — of the STFT filter:

$$S_{\mathbf{H}_S}(m, \nu) = \widehat{M}(m, \nu) A_{g,\gamma}(m, \nu) \quad (6.47)$$

where $\widehat{M}(m, \nu) \triangleq \sum_{n=-\infty}^{\infty} \int_0^1 M(n, \theta) e^{-j2\pi(\nu n - m\theta)} d\theta$ is the FT of $M(n, \theta)$ and $A_{g,\gamma}(m, \nu) \triangleq \sum_{n=-\infty}^{\infty} g[n] \gamma^*[n-m] e^{-j2\pi\nu n}$ is the cross ambiguity function of $g[n]$ and $\gamma[n]$ [45–47, 59, 61, 62]. For usual windows $g[n]$ and $\gamma[n]$, $A_{g,\gamma}(m, \nu)$ is well concentrated about the origin. Hence, Equation (6.47) implies that also the spreading function of \mathbf{H}_S is well concentrated about the origin. This again shows that STFT filters are inherently underspread.

6.5.2 Choice of Windows

We now discuss the choice of the STFT analysis window $\gamma[n]$ and the STFT synthesis window $g[n]$. Under the assumption that $M(n, \theta)$ is reasonably smooth, it makes sense to desire that the STFT filter \mathbf{H}_S approximates the Zadeh filter^{††} \mathbf{H}_Z associated to $M(n, \theta)$ — recall from Subsection 6.2.1 that the Zadeh filter design is appropriate only if $M(n, \theta)$ is smooth. Thus, we desire that $Z_{\mathbf{H}_S}(n, \theta) \stackrel{!}{\approx} M(n, \theta)$ or, taking the FT,

$$S_{\mathbf{H}_S}(m, \nu) \stackrel{!}{\approx} \widehat{M}(m, \nu).$$

By comparing with Equation (6.47), we see that this requires

$$A_{g,\gamma}(m, \nu) \stackrel{!}{\approx} \begin{cases} 1, & (m, \nu) \in \mathcal{S}_{\widehat{M}} \\ \text{arbitrary}, & (m, \nu) \notin \mathcal{S}_{\widehat{M}} \end{cases} \quad (6.48)$$

where $\mathcal{S}_{\widehat{M}}$ is the effective support region of $\widehat{M}(m, \nu)$ or a suitably simplified (e.g., rectangular or elliptical) region that contains the effective support region of $\widehat{M}(m, \nu)$. Outside $\mathcal{S}_{\widehat{M}}$, $A_{g,\gamma}(m, \nu)$ is effectively arbitrary because it is multiplied by $\widehat{M}(m, \nu) \approx 0$ in Equation (6.47). The design rule in Equation (6.48) is illustrated in Figure 6.13. Note that the effective support region $\mathcal{S}_{\widehat{M}}$ of $\widehat{M}(m, \nu)$ is all we need to know for designing the windows $\gamma[n]$ and $g[n]$. If $M(n, \theta)$ is smooth, then $\widehat{M}(m, \nu)$ is well concentrated about the origin, which means that $\mathcal{S}_{\widehat{M}}$ is small.

The perfect-reconstruction condition in Equation (6.45) can be reformulated as $A_{g,\gamma}(0, 0) = 1$, which is consistent with the design rule in Equation (6.48). Note that Equation (6.48) means that $A_{g,\gamma}(m, \nu)$ should not decay too fast as we move away from the origin $(0, 0)$.

^{††}We use the Zadeh filter because the Weyl filter would imply a restriction to halfband filters. However, the design rules to be presented can be modified such that they apply to the Weyl filter. This modification essentially consists of replacing $S_{\mathbf{H}_S}(m, \nu)$ and $A_{g,\gamma}(m, \nu)$ with symmetrical versions.

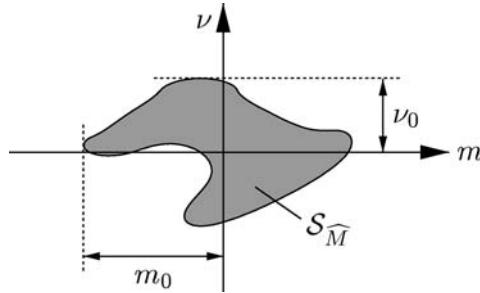
**FIGURE 6.13**

Illustration of the design rule in Equation (6.48). The cross-ambiguity function $A_{g,\gamma}(m,\nu)$ should be approximately 1 on $\mathcal{S}_{\widehat{M}}$. Also shown are the maximum time shift $m_0 \triangleq \max_{(m,\nu) \in \mathcal{S}_{\widehat{M}}} |m|$ and the maximum frequency shift $\nu_0 \triangleq \max_{(m,\nu) \in \mathcal{S}_{\widehat{M}}} |\nu|$ of $\mathcal{S}_{\widehat{M}}$.

6.5.2.1 Equal analysis and synthesis windows

Let us first assume $g[n] = \gamma[n]$ with $\gamma[n]$ real valued, even symmetrical and normalized ($\|\gamma\| = 1$). If $\mathcal{S}_{\widehat{M}}$ is sufficiently small ($M(n,\theta)$ is sufficiently smooth), then the deviation of $A_\gamma(m,\nu) \triangleq A_{\gamma,\gamma}(m,\nu)$ from 1 can be approximated within $\mathcal{S}_{\widehat{M}}$ as $1 - A_\gamma(m,\nu) \approx 2\pi^2(\Theta_\gamma^2 m^2 + N_\gamma^2 \nu^2)$ (see [63]). Here, $N_\gamma^2 \triangleq \sum_{n=-\infty}^{\infty} n^2 \gamma^2[n]$ and $\Theta_\gamma^2 \triangleq \frac{1}{\pi^2} \int_{-1/2}^{1/2} \sin^2(\pi\theta) \Gamma^2(\theta) d\theta$ are measures of the effective duration and bandwidth of $\gamma[n]$, respectively. The maximum deviation within $\mathcal{S}_{\widehat{M}}$ can then be approximated by

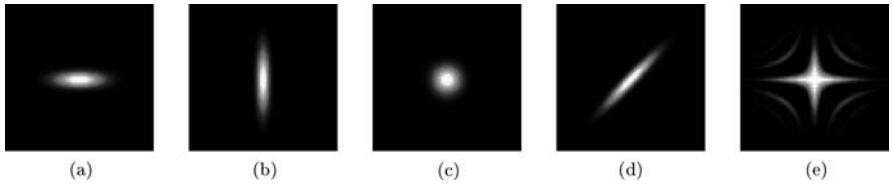
$$d \triangleq 2\pi^2(\Theta_\gamma^2 m_0^2 + N_\gamma^2 \nu_0^2), \quad (6.49)$$

where $m_0 \triangleq \max_{(m,\nu) \in \mathcal{S}_{\widehat{M}}} |m|$ and $\nu_0 \triangleq \max_{(m,\nu) \in \mathcal{S}_{\widehat{M}}} |\nu|$ denote the maximum time shift and maximum frequency shift within the region $\mathcal{S}_{\widehat{M}}$ respectively, (see Figure 6.13).

According to Equation (6.48), we have to choose $\gamma[n]$ such that d is small. Unfortunately, due to the uncertainty principle, N_γ and Θ_γ cannot both be arbitrarily small. For a given duration-bandwidth product $N_\gamma \Theta_\gamma = \mu$, it easily follows from Equation (6.49) that the ratio N_γ / Θ_γ minimizing d and the resulting minimum d are given by:

$$\frac{N_\gamma}{\Theta_\gamma} = \frac{m_0}{\nu_0}, \quad d_{\min} = 4\pi^2 \mu m_0 \nu_0 \quad (6.50)$$

The first relation allows to match N_γ and Θ_γ to the shape of $\mathcal{S}_{\widehat{M}}$. For example, if $\mathcal{S}_{\widehat{M}}$ is widely spread out in the m direction and well concentrated in the ν direction, we should choose a window with a large duration N_γ and a small bandwidth Θ_γ . The second relation shows that windows with a small duration-bandwidth product $\mu = N_\gamma \Theta_\gamma$ are advantageous, and that the resulting deviation is smaller for a smaller value of $m_0 \nu_0$ (i.e., for a smaller $\mathcal{S}_{\widehat{M}}$). If $\mathcal{S}_{\widehat{M}}$ is obliquely oriented, even better matching results can be achieved by using a window with linear chirp modulation. Figure 6.14(a) to (d) shows the deviation $|1 - A_\gamma(m,\nu)|$ for some windows.

**FIGURE 6.14**

Deviation $|1 - A_{g,\gamma}(m, \nu)|$ for special choices of windows. (a) to (d) Equal analysis and synthesis windows with (a) large duration, (b) small duration, (c) medium duration, (d) linear chirp modulation; (e) different analysis and synthesis windows with $\gamma[n]$ of large duration and $g[n]$ of small duration. White corresponds to 0, black corresponds to 1, horizontal axis is time shift m , vertical axis is frequency shift ν , and the origin $(m, \nu) = (0, 0)$ is at the center of each plot.

6.5.2.2 Different analysis and synthesis windows

Next, we consider different analysis and synthesis windows. If they are real valued and even symmetrical and satisfy Equation (6.45), one can show

$$|1 - A_{g,\gamma}(m, \nu)| \approx \sqrt{\left[2\pi^2(\Theta_{g,\gamma}^2 m^2 + N_{g,\gamma}^2 \nu^2)\right]^2 + 2\pi^2 \kappa_{g,\gamma}^2 m^2 \nu^2}$$

for $(m, \nu) \in \mathcal{S}_{\widehat{M}}$, provided that the support region $\mathcal{S}_{\widehat{M}}$ is sufficiently small. Here, $N_{g,\gamma}^2 \triangleq \sum_{n=-\infty}^{\infty} n^2 g[n] \gamma[n]$, $\Theta_{g,\gamma}^2 \triangleq \frac{1}{\pi^2} \int_{-1/2}^{1/2} \sin^2(\pi\theta) G(\theta) \Gamma(\theta) d\theta$, and $\kappa_{g,\gamma} \triangleq \int_{-1/2}^{1/2} \sin(2\pi\theta) G'(\theta) \Gamma(\theta) d\theta$, with $G'(\theta)$ the derivative of $G(\theta)$. Now $N_{g,\gamma}$ and $\Theta_{g,\gamma}$ can simultaneously be made small by choosing a short (long) analysis window and a long (short) synthesis window, so that both $g[n] \gamma[n]$ and $G(\theta) \Gamma(\theta)$ are short. However, in this case $\kappa_{g,\gamma}$ is large, resulting in significant values of the term $2\pi^2 \kappa_{g,\gamma}^2 m^2 \nu^2$ away from the m and ν axes. An example is shown in Figure 6.14(e).

Because of the limited design freedom in designing $A_{g,\gamma}(m, \nu)$, accurate approximation of the Zadeh filter by an STFT filter requires a small $\mathcal{S}_{\widehat{M}}$, that is, a smooth $M(n, \theta)$. This limitation can be removed by the multiwindow STFT filter considered in Subsections 6.5.4 and 6.5.5.

6.5.3 On-line implementation of short-time Fourier transform filter

For an efficient on-line implementation of the STFT filter, we assume that $\gamma[n]$ and $g[n]$ are supported in the intervals $[-L_\gamma/2, L_\gamma/2 - 1]$ and $[-L_g/2, L_g/2 - 1]$, respectively, with the window lengths L_γ and L_g assumed even. We use the discrete STFT, which is derived from the STFT in Equation (6.43) by a sampling of the frequency variable θ using K frequency bins. Here, K is assumed an even number with $K \geq \max\{L_\gamma, L_g\}$, so that the support intervals of both $\gamma[n]$ and $g[n]$ are contained in the interval $[-K/2, K/2 - 1]$. Thus, the number of frequency bins is at least as large as the window lengths. This frequency sampling is also applied to the TF weight function $M(n, \theta)$, which yields the discrete TF weight function

$M[n, k] \triangleq M(n, k/K)$. Here, K must be chosen large enough so that $M(n, \theta)$ is sampled sufficiently densely. We then propose the following discrete-frequency version of the three-step procedure in Subsection 6.5.1.

1. *Analysis.* The discrete STFT of the input signal $x[n]$ is calculated:

$$\begin{aligned} \text{STFT}_x^{(\gamma)}[n, k] &\triangleq \text{STFT}_x^{(\gamma)}\left(n, \frac{k}{K}\right) \\ &= \sum_{n'=n-L_\gamma/2}^{n+L_\gamma/2-1} x[n'] \gamma^*[n'-n] e^{-j2\pi \frac{k}{K}(n'-n)} \\ &= \sum_{m=-K/2}^{K/2-1} x_n^{(\gamma)}[m] e^{-j2\pi \frac{k}{K}m}, \quad k = 0, 1, \dots, K-1, \end{aligned} \quad (6.51)$$

$$(6.52)$$

where $x_n^{(\gamma)}[m] \triangleq x[n+m] \gamma^*[m]$. Thus, the STFT analysis step at time n amounts to computing the length- K DFT of the vector $(x_n^{(\gamma)}[-K/2] \ x_n^{(\gamma)}[-K/2+1] \ \dots \ x_n^{(\gamma)}[K/2-1])$. This involves the local input signal samples $x[n-L_\gamma/2], x[n-L_\gamma/2+1], \dots, x[n+L_\gamma/2-1]$.

2. *Weighting.* The discrete STFT of $x[n]$ is multiplied by $M[n, k]$:

$$F[n, k] \triangleq M[n, k] \text{STFT}_x^{(\gamma)}[n, k], \quad k = 0, 1, \dots, K-1.$$

This amounts to the element-wise product of the two length- K vectors $(M[n, 0] \ M[n, 1] \ \dots \ M[n, K-1])$ and $(\text{STFT}_x^{(\gamma)}[n, 0] \ \text{STFT}_x^{(\gamma)}[n, 1] \ \dots \ \text{STFT}_x^{(\gamma)}[n, K-1])$.

3. *Synthesis.* The output sample $y[n]$ is obtained as the inverse discrete STFT of $F[n, k]$:

$$y[n] = \frac{1}{K} \sum_{n'=n-L_g/2+1}^{n+L_g/2} \sum_{k=0}^{K-1} F[n', k] g[n-n'] e^{j2\pi \frac{k}{K}(n-n')}.$$

This can be formulated as an overlap-add scheme [15–17]:

$$y[n] = \sum_{n'=n-L_g/2+1}^{n+L_g/2} y_{n'}^{(g)}[n-n'] = \sum_{m=-L_g/2}^{L_g/2-1} y_{n-m}^{(g)}[m]. \quad (6.53)$$

there, $y_{n'}^{(g)}[m]$ is a windowed version of the IDFT of $F[n', k]$ with respect to k ,

$$y_{n'}^{(g)}[m] \triangleq f_{n'}[m] g[m]$$

where

$$f_{n'}[m] \triangleq \frac{1}{K} \sum_{k=0}^{K-1} F[n', k] e^{j2\pi \frac{k}{K} m}.$$

For a recursive implementation of this overlap-add scheme, let:

$$s_i[n] \triangleq \sum_{m=L_g/2-i}^{L_g/2-1} y_{n+L_g-i-m}^{(g)}[m], \quad i = 1, 2, \dots, L_g,$$

denote partial sums of Equation (6.53). The output sample $y[n]$ is obtained as the L_g th partial sum, $y[n] = s_{L_g}[n]$. Furthermore, the partial sums can be calculated recursively as

$$s_i[n] = \begin{cases} s_{i-1}[n-1] + y_{n+L_g/2}^{(g)}[\frac{L_g}{2}-i], & i = 2, 3, \dots, L_g \\ y_{n+L_g/2}^{(g)}[\frac{L_g}{2}-1], & i = 1. \end{cases}$$

This on-line STFT filter satisfies the perfect-reconstruction property provided that the windows $\gamma[n]$ and $g[n]$ satisfy the perfect-reconstruction condition in Equation (6.45). However, because at time n the future input samples $x[n+1], x[n+2], \dots, x[n+(L_\gamma + L_g)/2]$ are used, a time delay of $(L_\gamma + L_g)/2$ samples has to be introduced. At any time n , the STFT analysis, weighting and synthesis steps require $\mathcal{O}(L_\gamma + K \log K)$, $\mathcal{O}(K)$, and $\mathcal{O}(L_g + K \log K)$ operations, respectively, making a total of $\mathcal{O}(L_\gamma + L_g + K + 2K \log K)$ operations per signal sample. Because the length- L_γ input vector $(x[n-L_\gamma+1] \ x[n-L_\gamma+2] \ \dots \ x[n])$ and the length- L_g partial-sum vector $(s_1[n] \ s_2[n] \ \dots \ s_{L_g}[n])$ need to be stored, $L_\gamma + L_g$ memory locations are required.

Substantial simplifications of this algorithm can be achieved by the special choices $\gamma[n] = \delta[n]$ or $g[n] = \delta[n]$ (the latter choice leads to the filterbank summation method [15–17]).

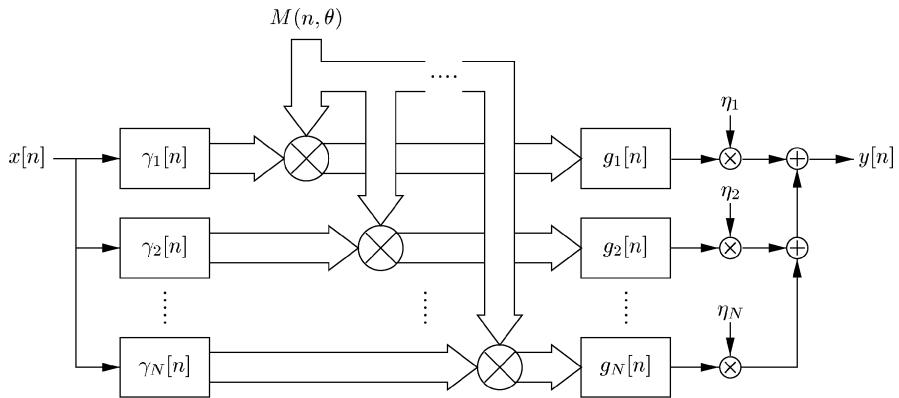
6.5.4 Multiwindow short-time Fourier transform filter

A generalization of the STFT filter \mathbf{H}_S with increased flexibility is depicted in Figure 6.15. The output signal is a linear combination (weighted sum) of the output signals of N STFT filters $\mathbf{H}_S^{(i)}$, $i = 1, 2, \dots, N$, with different windows $\gamma_i[n]$, $g_i[n]$ but the same TF weight function $M(n, \theta)$. The resulting *multiwindow STFT filter* is given by [8, 9, 12, 64]

$$\mathbf{H}_{MS} \triangleq \sum_{i=1}^N \eta_i \mathbf{H}_S^{(i)},$$

where the constants η_i are the branch weights* associated to the individual STFT filters $\mathbf{H}_S^{(i)}$. Using a larger number N of STFT filters yields increased flexibility of

*Note that these branch weights could equally well be incorporated in the analysis or synthesis windows.

**FIGURE 6.15**

Block diagram of a multiwindow STFT filter.

design at the expense of increased computational complexity. Note that the STFT filter \mathbf{H}_S is a special case with $N = 1$.

The impulse response of the multiwindow STFT filter \mathbf{H}_{MS} is given by

$$h_{MS}[n_1, n_2] = \sum_{n=-\infty}^{\infty} \int_0^1 M(n, \theta) p[n_1 - n, n_2 - n] e^{j2\pi\theta(n_1 - n_2)} d\theta,$$

with

$$p[n, n'] \triangleq \sum_{i=1}^N \eta_i g_i[n] \gamma_i^*[n']. \quad (6.54)$$

It is convenient to interpret $p[n, n']$ as the impulse response of an LTV system \mathbf{P} ; this system will be called the *prototype system*. Note that the prototype system \mathbf{P} subsumes the influence of the windows $\gamma_i[n]$, $g_i[n]$ and the branch weights η_i on \mathbf{H}_{MS} . It can be shown that perfect reconstruction ($(\mathbf{H}_{MS} x)[n] = x[n]$ for $M(n, \theta) \equiv 1$) is achieved if and only if

$$\sum_{n=-\infty}^{\infty} p[n, n] = 1$$

or equivalently

$$\sum_{i=1}^N \eta_i \langle g_i, \gamma_i \rangle = 1. \quad (6.55)$$

Perfect reconstruction of the individual STFT filter branches $\mathbf{H}_S^{(i)}$ (that is, $\langle g_i, \gamma_i \rangle = 1$ for $i = 1, 2, \dots, N$, — see Equation (6.45)) — together with the normalization $\sum_{i=1}^N \eta_i = 1$ is a sufficient but not a necessary condition for perfect reconstruction of \mathbf{H}_{MS} .

6.5.4.1 Comparison of multiwindow STFT filter with Zadeh filter

The Zadeh function of \mathbf{H}_{MS} is obtained as [8, 9, 64]

$$Z_{\mathbf{H}_{\text{MS}}}(n, \theta) = M(n, \theta) * Z_{\mathbf{P}}(n, \theta) = \sum_{n'=-\infty}^{\infty} \int_0^1 M(n', \theta') Z_{\mathbf{P}}(n-n', \theta-\theta') d\theta'. \quad (6.56)$$

By comparing with Equation (6.6), we see that the multiwindow STFT filter \mathbf{H}_{MS} using TF weight function $M(n, \theta)$ is equivalent to the Zadeh filter \mathbf{H}_Z using the modified TF weight function $\tilde{M}(n, \theta) \triangleq M(n, \theta) * Z_{\mathbf{P}}(n, \theta)$. Depending on the choice of the windows $\gamma_i[n], g_i[n]$ and of the branch weights η_i , $Z_{\mathbf{P}}(n, \theta)$ can assume quite different shapes. In particular, if $Z_{\mathbf{P}}(n, \theta)$ is well concentrated and peaked about the origin, we obtain $\tilde{M}(n, \theta) \approx M(n, \theta)$. Here, \mathbf{H}_{MS} is approximately equal to the Zadeh filter \mathbf{H}_Z using the original TF weight function $M(n, \theta)$, and hence it is underspread only if $M(n, \theta)$ is sufficiently smooth. The choice of \mathbf{P} such that \mathbf{H}_{MS} approximates \mathbf{H}_Z is discussed in Subsection 6.5.5. If, on the other hand, $Z_{\mathbf{P}}(n, \theta)$ is more spread out and smooth, Equation (6.56) expresses a substantial smoothing of $M(n, \theta)$. Here, $Z_{\mathbf{H}_{\text{MS}}}(n, \theta)$ is a smooth function even if $M(n, \theta)$ is not smooth, and thus the multiwindow STFT filter \mathbf{H}_{MS} is inherently underspread just as the STFT filter.

6.5.4.2 On-line implementation of multiwindow STFT filter

The on-line implementation of \mathbf{H}_{MS} essentially amounts to the individual on-line implementations of the N STFT filters $\mathbf{H}_S^{(i)}$ (see Subsection 6.5.3). When using FFT length $K \geq \max_{i=1,2,\dots,N}\{L_{\gamma_i}, L_{g_i}\}$, we require $\mathcal{O}(\sum_{i=1}^N(L_{\gamma_i} + L_{g_i}) + N(K + 2K \log K))$ operations per signal sample, plus $\mathcal{O}(N)$ operations for calculating the weighted sum of the N STFT filter output signals. The number of memory locations required is $\max_{i=1,2,\dots,N}\{L_{\gamma_i}\} + \sum_{i=1}^N L_{g_i}$ (we have to store an input vector of length $\max_{i=1,2,\dots,N}\{L_{\gamma_i}\}$ and N partial-sum vectors of respective lengths L_{g_i} — see Subsection 6.5.3). Note that the multiwindow STFT filter is ideally suited to parallel processing.

6.5.5 Choice of prototype system

Next, we provide guidelines and methods for the choice (design) of the prototype system \mathbf{P} . We recall from Equation (6.54) that \mathbf{P} involves the design parameters $\gamma_i[n], g_i[n], \eta_i$ and N . Our discussion is based on the following generalization[†] of Equation (6.47)

$$S_{\mathbf{H}_{\text{MS}}}(m, \nu) = \widehat{M}(m, \nu) S_{\mathbf{P}}(m, \nu), \quad (6.57)$$

where again $\widehat{M}(m, \nu) = \sum_{n=-\infty}^{\infty} \int_0^1 M(n, \theta) e^{-j2\pi(\nu n - m\theta)} d\theta$.

[†]To appreciate the relation with Equation (6.47), note that the spreading function of \mathbf{P} can be expressed in terms of the cross-ambiguity functions of the windows as $S_{\mathbf{P}}(m, \nu) = \sum_{i=1}^N \eta_i A_{g_i, \gamma_i}(m, \nu)$.

Under the assumption that $M(n, \theta)$ is reasonably smooth, we desire that \mathbf{H}_{MS} approximates the Zadeh filter (compare with Subsection 6.5.2). This means that $Z_{\mathbf{H}_{\text{MS}}}(n, \theta) \stackrel{!}{\approx} M(n, \theta)$ or, equivalently,

$$S_{\mathbf{H}_{\text{MS}}}(m, \nu) \stackrel{!}{\approx} \widehat{M}(m, \nu).$$

By comparing with Equation (6.57), we see that we have to choose \mathbf{P} such that (see Figure 6.13)

$$S_{\mathbf{P}}(m, \nu) \stackrel{!}{\approx} \begin{cases} 1, & (m, \nu) \in \mathcal{S}_{\widehat{M}} \\ \text{arbitrary}, & (m, \nu) \notin \mathcal{S}_{\widehat{M}} \end{cases} \quad (6.58)$$

which generalizes Equation (6.48). The support region $\mathcal{S}_{\widehat{M}}$ is small because $M(n, \theta)$ was assumed smooth. Because the perfect-reconstruction condition in Equation (6.55) implies $S_{\mathbf{P}}(0, 0) = 1$, the design rule in Equation (6.58) means that $S_{\mathbf{P}}(m, \nu)$ should not decay too fast as we move away from the origin $(0, 0)$. We next propose two different methods for designing a \mathbf{P} according to Equation (6.58).

6.5.5.1 First design method

Let $\tilde{S}(m, \nu)$ be a function that is 1 on $\mathcal{S}_{\widehat{M}}$ and rolls off to 0 outside $\mathcal{S}_{\widehat{M}}$. We can interpret $\tilde{S}(m, \nu)$ as the spreading function of an ideal prototype system $\widetilde{\mathbf{P}}$, that is, $\tilde{S}(m, \nu) = S_{\widetilde{\mathbf{P}}}(m, \nu)$. By using Equation (6.2), the impulse response of $\widetilde{\mathbf{P}}$ is obtained as

$$\tilde{p}[n, n'] = \int_0^1 \tilde{S}(n - n', \nu) e^{j2\pi\nu n} d\nu.$$

To derive the multiwindow STFT parameters $\gamma_i[n]$, $g_i[n]$ and η_i from $\tilde{p}[n, n']$, we use the singular value decomposition [5, 6, 65] of $\widetilde{\mathbf{P}}$,

$$\tilde{p}[n, n'] = \sum_{i=1}^{\infty} \sigma_i u_i[n] v_i^*[n'].$$

Here, $\sigma_i \geq 0$, $u_i[n]$ and $v_i[n]$ are the singular values, left singular functions and right singular functions of $\widetilde{\mathbf{P}}$, respectively. Comparing with Equation (6.54) suggests the use of $\eta_i = \sigma_i$, $\gamma_i[n] = v_i[n]$ and $g_i[n] = u_i[n]$. Unfortunately, this would generally require infinitely many STFT filter branches ($N = \infty$). Therefore, for a given, finite number N of STFT filter branches, we use for the prototype system \mathbf{P} the rank- N system that minimizes $\|\tilde{S} - S_{\mathbf{P}}\|$ subject to the perfect-reconstruction constraint $\sum_{n=-\infty}^{\infty} p[n, n] = 1$. It can be shown that this \mathbf{P} is given by

$$p[n, n'] = \frac{1}{\kappa} \sum_{i=1}^N \sigma_i u_i[n] v_i^*[n']$$

with:

$$\kappa \triangleq \sum_{i=1}^N \sigma_i \langle u_i, v_i \rangle.$$

This results in the multiwindow STFT filter parameters:

$$\eta_i = \frac{\sigma_i}{\kappa}, \quad \gamma_i[n] = v_i[n], \quad g_i[n] = u_i[n] \quad i = 1, 2, \dots, N$$

The choice of N is governed by a trade-off between a small approximation error $\|\tilde{S} - S_{\mathbf{P}}\|$ (obtained with N large) and a small computational complexity of \mathbf{H}_{MS} (obtained with N small).

6.5.5.2 Second design method

Next, we propose an alternative design method that uses a family of standard windows (sampled Hermite functions). Let us choose $S_{\widehat{M}}$ as the elliptical region defined by $\rho^2 \triangleq (m/m_0)^2 + (\nu/\nu_0)^2 \leq 1$, where the semiaxes m_0 and ν_0 are such that $S_{\widehat{M}}$ contains the effective support of $\widehat{M}(m, \nu)$. The design rule in Equation (6.58) can then be written as

$$S_{\mathbf{P}}(m, \nu) \stackrel{!}{\approx} \begin{cases} 1, & \rho^2 \leq 1 \\ \text{arbitrary}, & \rho^2 > 1. \end{cases} \quad (6.59)$$

We design \mathbf{P} such that $S_{\mathbf{P}}(m, \nu)$ has elliptical symmetry. By transferring continuous-time results [14, 60, 63] into the discrete-time domain via a sampling argument, it can be shown that elliptical symmetry of $S_{\mathbf{P}}(m, \nu)$ is obtained (up to negligible aliasing errors) if the analysis and synthesis windows are equal and given by the first N sampled Hermite functions $H_i[n], i = 1, 2, \dots, N$, [14, 42, 60, 63]:

$$\gamma_i[n] = g_i[n] = H_i[n] \triangleq \frac{1}{\sqrt{2^{i-3/2}(i-1)!}\Delta} \overline{H}_{i-1}\left(\sqrt{2\pi} \frac{n}{\Delta}\right) e^{-\pi(n/\Delta)^2}. \quad (6.60)$$

Here, $\overline{H}_i(\alpha)$ is the Hermite polynomial of order i :

$$\overline{H}_i(\alpha) \triangleq (-1)^i e^{\alpha^2} \frac{d^i}{d\alpha^i} e^{-\alpha^2}, \quad i = 0, 1, \dots$$

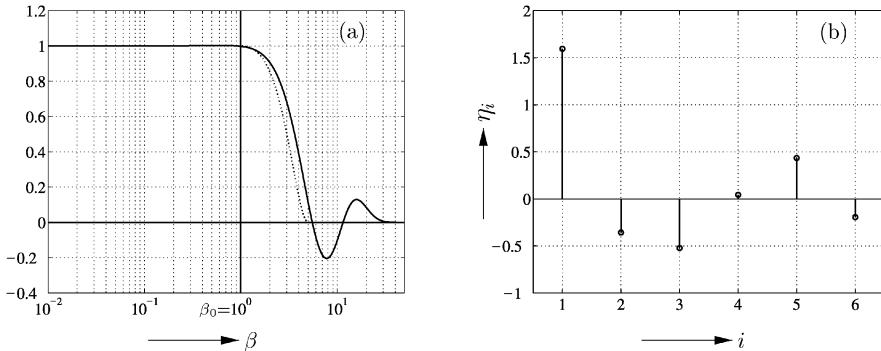
and $\Delta = \sqrt{m_0/\nu_0}$ is a scaling parameter. Apart from aliasing errors that are negligible as long as N is not too large,[‡] the spreading function of \mathbf{P} is then given by (see [14, 42, 60, 63])

$$S_{\mathbf{P}}(m, \nu) = \sum_{i=1}^N \eta_i L_{i-1}(\pi m_0 \nu_0 \rho^2). \quad (6.61)$$

Here, $L_i(\beta), \beta > 0$, denotes the orthonormal Laguerre functions defined as:

$$L_i(\beta) \triangleq \overline{L}_i(\beta) e^{-\beta/2}$$

[‡]The effective duration and bandwidth of $H_i[n]$ are $\sqrt{(i+1)/\pi}\Delta$ and $\sqrt{(i+1)/\pi}/\Delta$, respectively. When using the $H_i[n]$ up to $i = N$, the maximum effective bandwidth $\sqrt{(N+1)/\pi}/\Delta$ must be $< 1/2$ to avoid aliasing errors.

**FIGURE 6.16**

Example of a multiwindow STFT prototype system design using the second design method with parameters $N = 6$ and $\beta_0 = \pi m_0 \nu_0 = 1$. (a) Target profile (dotted line) and weighted least-squares approximation (solid line); (b) resulting branch weights η_i .

with

$$\bar{L}_i(\beta) \triangleq \sum_{j=0}^i \binom{i}{j} \frac{(-\beta)^j}{j!}, \quad i = 0, 1, \dots$$

and the branch weights η_i remain to be determined. The expression in Equation (6.61) shows that $S_{\mathbf{P}}(m, \nu)$ has indeed elliptical symmetry. The design rule in Equation (6.59) now becomes

$$\sum_{i=1}^N \eta_i L_{i-1}(\beta) \stackrel{!}{\approx} \begin{cases} 1, & \beta \leq \beta_0 \triangleq \pi m_0 \nu_0 \\ \text{arbitrary}, & \beta > \beta_0. \end{cases} \quad (6.62)$$

Because for N finite there will not exist any branch weight sequence η_i such that $\sum_{i=1}^N \eta_i L_{i-1}(\beta)$ is exactly 1 for $\beta \leq \beta_0$, we adopt a weighted least-squares approach. We choose a target profile that is 1 for $\beta \leq \beta_0$ and performs a smooth roll-off from 1 to 0 for $\beta > \beta_0$. We then calculate the η_i by means of a weighted least-squares fit [65, 66] of $\sum_{i=1}^N \eta_i L_{i-1}(\beta)$ to that target profile. The fact that $\sum_{i=1}^N \eta_i L_{i-1}(\beta)$ is arbitrary for $\beta > \beta_0$ can be taken into account by attaching very small weights to the associated error terms.[§] An example of this prototype design for $m_0 \nu_0 = 1/\pi$ (i.e., $\beta_0 = 1$) and $N = 6$ is shown in Figure 6.16.

We can conclude from this discussion that if $M(n, \theta)$ is reasonably smooth, \mathbf{P} can be designed such that the resulting multiwindow STFT filter \mathbf{H}_{MS} provides a good approximation to the Zadeh filter. In a similar manner, \mathbf{P} can be designed such that \mathbf{H}_{MS} approximates the halfband Weyl filter. This latter approximation is verified experimentally in Subsection 6.7.2.

[§]Using an ordinary least-squares fit on the interval $[0, \beta_0]$ was observed to lead to numerical instabilities.

6.6 Implicit Time–Frequency Filter Design II: Gabor Filter

A major problem with the STFT filter or the multiwindow STFT filter is its large computational complexity, which is caused by the redundancy of the (discrete-frequency) STFT as a signal representation. Indeed, one signal sample $x[n]$ is represented by K STFT samples $\text{STFT}_x^{(\gamma)}(n, k/K)$ ($k = 0, 1, \dots, K - 1$) or, in the case of the multiwindow STFT, even by NK samples. This problem can be alleviated or altogether avoided by a subsampling of the discrete-frequency STFT. This leads to the *Gabor transform* (or *Gabor expansion*) [25, 37, 47, 67, 68] and, in turn, to an implicit TF filter design method that is termed *Gabor filter* [9, 16, 21–25].

The Gabor transform is equivalent to a uniform DFT filterbank [17, 69–71], and its mathematical analysis can be based on frame theory [37, 67, 72–74]. However, we provide a self-contained discussion that does not presuppose familiarity with filterbanks or frame theory.

6.6.1 Gabor transform and Gabor filter

6.6.1.1 Gabor transform

The Gabor transform $\text{GT}_x^{(\gamma)}[l, k]$ is a sampled STFT, that is, $\text{STFT}_x^{(\gamma)}(n, \theta)$ evaluated on a TF lattice $n = lL, \theta = k/K$ with $l \in \mathbb{Z}$ and $k = 0, 1, \dots, K - 1$:

$$\begin{aligned} \text{GT}_x^{(\gamma)}[l, k] &\triangleq \text{STFT}_x^{(\gamma)}\left(lL, \frac{k}{K}\right) \\ &= \sum_{n=-\infty}^{\infty} x[n] \gamma^*[n-lL] e^{-j2\pi \frac{k}{K}(n-lL)}. \end{aligned} \quad (6.63)$$

The discrete-time, discrete-frequency STFT in Equation (6.51) is a special case of the Gabor transform with $L = 1$ (i.e., no time subsampling). The *oversampling factor* $\xi \triangleq K/L$ is a measure of the overall density of the TF lattice $(lL, k/K)$; a larger ξ corresponds to a denser TF lattice and, in turn, a higher computational complexity. The cases $K = L$ ($\xi = 1$), $K > L$ ($\xi > 1$), and $K < L$ ($\xi < 1$) are termed *critical sampling*, *oversampling* and *undersampling*, respectively [37, 67].

The Gabor transform can be inverted as

$$x[n] = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} \text{GT}_x^{(\gamma)}[l, k] g[n-lL] e^{j2\pi \frac{k}{K}(n-lL)}, \quad (6.64)$$

provided that the analysis window $\gamma[n]$ and the synthesis window $g[n]$ are *dual* or, equivalently, *biorthogonal* with respect to the adjoint TF lattice $(k'K, l'/L)$ with

$k' \in \mathbb{Z}$, $l' = 0, 1, \dots, L - 1$ [37, 67]. This biorthogonality can be formulated in terms of the cross-ambiguity function $A_{g,\gamma}(m, \nu)$ (see Subsection 6.5.1) as

$$A_{g,\gamma}\left(k'K, \frac{l'}{L}\right) = \sum_{n=-\infty}^{\infty} g[n] \gamma^*[n - k'K] e^{-j2\pi \frac{l'}{L} n} = \frac{L}{K} \delta[k'] \delta[l']. \quad (6.65)$$

It can be shown that this biorthogonality condition cannot be satisfied in the case of undersampling ($K < L$ or $\xi < 1$). The design of dual (biorthogonal) windows is discussed in [37, 67]. However, in the context of TF filters, approximate biorthogonality often is sufficient.

6.6.1.2 Gabor filter

The Gabor filter is an implicit TF filter design scheme that is essentially analogous to the STFT filter. It consists of the following three steps:

1. *Analysis.* The Gabor transform of the input signal $x[n]$ is calculated according to Equation (6.63).
 2. *Weighting.* The Gabor transform of $x[n]$ is multiplied by the discrete TF weight function $M[l, k] \triangleq M(lL, k/K)$ (i.e., the time-subsampled, frequency-sampled version of $M(n, \theta)$):
- $$F[l, k] \triangleq M[l, k] \text{GT}_x^{(\gamma)}[l, k], \quad l \in \mathbb{Z}, \quad k = 0, 1, \dots, K-1.$$
3. *Synthesis.* The filter output signal $y[n]$ is calculated by applying the Gabor inversion relation in Equation (6.64) to $F[l, k]$:

$$y[n] = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} F[l, k] g[n - lL] e^{j2\pi \frac{k}{K} (n - lL)}. \quad (6.66)$$

This procedure defines an LTV filter that is called Gabor filter and denoted \mathbf{H}_G . The Gabor filter depends not only on the TF weight function $M[l, k]$ but also on the TF lattice parameters L, K and on the analysis and synthesis windows $\gamma[n]$ and $g[n]$. The impulse response of \mathbf{H}_G is

$$h_G[n, n'] = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} M[l, k] g[n - lL] \gamma^*[n' - lL] e^{j2\pi \frac{k}{K} (n - n')}.$$

If the biorthogonality condition in Equation (6.65) is satisfied, then \mathbf{H}_G satisfies the perfect-reconstruction property, that is, $(\mathbf{H}_G x)[n] = x[n]$ for $M[l, k] \equiv 1$.

6.6.1.3 Comparison of Gabor filter with Zadeh filter

The Zadeh function of the Gabor filter is obtained as

$$Z_{\mathbf{H}_G}(n, \theta) = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} M[l, k] R_{g,\gamma}\left(n - lL, \theta - \frac{k}{K}\right), \quad (6.67)$$

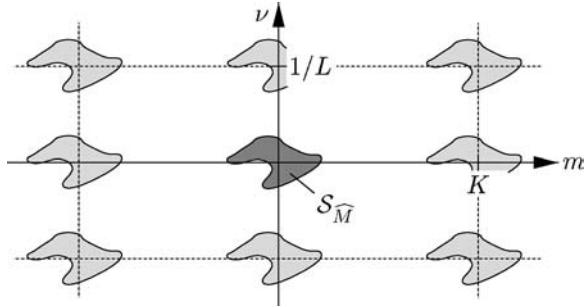
**FIGURE 6.17**

Illustration of the periodic function $\widehat{M}_p(m, \nu)$ and the design rule in Equation (6.69). $A_{g,\gamma}(m, \nu)$ should be approximately L/K on $S_{\widehat{M}}$ and approximately 0 on S_p , the union of all lightly shaded regions.

where $R_{g,\gamma}(n, \theta)$ is the cross Rihaczek distribution of $g[n]$ and $\gamma[n]$ (see Subsection 6.5.1). By comparing with Equation (6.6), we see that the Gabor filter \mathbf{H}_G associated to $M(n, \theta)$ is equivalent to the Zadeh filter \mathbf{H}_Z associated to $\widetilde{M}(n, \theta) \triangleq \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} M[l, k] R_{g,\gamma}(n - lL, \theta - k/K)$, which can be interpreted as a 2-D interpolation of the subsampled TF weight function $M[l, k] = M(lL, k/K)$. This subsampling and interpolation may lead to undesired aliasing and reconstruction effects that are analyzed in the following.

6.6.1.4 Gabor filter and underspread property

The spreading function of \mathbf{H}_G is given by

$$S_{\mathbf{H}_G}(m, \nu) = \frac{K}{L} \widehat{M}_p(m, \nu) A_{g,\gamma}(m, \nu), \quad (6.68)$$

where

$$\widehat{M}_p(m, \nu) \triangleq \sum_{k'=-\infty}^{\infty} \sum_{l'=0}^{L-1} \widehat{M}\left(m - k'K, \nu - \frac{l'}{L}\right).$$

We see that $\widehat{M}_p(m, \nu)$ is derived from $\widehat{M}(m, \nu)$ through a 2-D periodization based on the adjoint TF lattice (see Figure 6.17); this periodization is the main difference from the STFT filter relation (6.47). If $\gamma[n]$ and $g[n]$ are well behaved, $A_{g,\gamma}(m, \nu)$ is concentrated about the origin (see Subsection 6.6.2). According to Equation (6.68), the same is then true for $S_{\mathbf{H}_G}(m, \nu)$, which means that \mathbf{H}_G is underspread in this case. We also see from Equation (6.68) and Figure 6.17 that the perfect-reconstruction condition in Equation (6.65), $A_{g,\gamma}(k'K, l'/L) = \frac{L}{K} \delta[k'] \delta[l']$, is instrumental in suppressing the image components $\widehat{M}(m - k'K, \nu - l'/L)$ with $k' \neq 0, l' \neq 0$ that are contained in $\widehat{M}_p(m, \nu)$. If these image components are not suppressed through the multiplication by $A_{g,\gamma}(m, \nu)$ in Equation (6.68), they lead to components in $S_{\mathbf{H}_G}(m, \nu)$ that correspond to undesired TF shifts by $(m, \nu) \approx (k'K, l'/L)$. These TF shifts are incompatible with the underspread property of \mathbf{H}_G .

6.6.2 Choice of lattice parameters and windows

Next, we consider the choice of the TF lattice parameters L , K and the windows $\gamma[n]$, $g[n]$. Under the assumption that $M(n, \theta)$ is reasonably smooth, we desire that the Gabor filter \mathbf{H}_G approximates the Zadeh filter, that is, $Z_{\mathbf{H}_G}(n, \theta) \stackrel{!}{\approx} M(n, \theta)$ or, equivalently,

$$S_{\mathbf{H}_G}(m, \nu) \stackrel{!}{\approx} \widehat{M}(m, \nu).$$

By comparing with Equation (6.68), we see that this can be satisfied if and only if

$$A_{g,\gamma}(m, \nu) \approx \begin{cases} L/K, & (m, \nu) \in \mathcal{S}_{\widehat{M}} \\ 0, & (m, \nu) \in \mathcal{S}_p \\ \text{arbitrary,} & \text{otherwise,} \end{cases} \quad (6.69)$$

where $\mathcal{S}_{\widehat{M}}$ is the effective support region of the fundamental component $\widehat{M}(m, \nu)$ and \mathcal{S}_p is the union of the support regions of all image components $\widehat{M}(m - k'K, \nu - l'/L)$ with $k' \neq 0$, $l' \neq 0$ (see Figure 6.17). Note the difference from the STFT filter design rule in Equation (6.48). As a by-product, the design rule in Equation (6.69) ensures that

$$A_{g,\gamma}\left(k'K, \frac{l'}{L}\right) \approx \frac{L}{K} \delta[k'] \delta[l'].$$

This means that the perfect-reconstruction (biorthogonality) condition in Equation (6.65) is approximately satisfied. Conversely, biorthogonal windows do not necessarily satisfy Equation (6.69): they satisfy $A_{g,\gamma}(m, \nu) = L/K$ at $(m, \nu) = 0$ but not necessarily on all of $\mathcal{S}_{\widehat{M}}$, and they satisfy $A_{g,\gamma}(m, \nu) = 0$ at $(m, \nu) = (k'K, l'/L)$ with $k' \neq 0$, $l' \neq 0$ but not necessarily on all of \mathcal{S}_p . We next show how to choose TF lattice parameters and windows according to the design rule in Equation (6.69).

6.6.2.1 Choice of TF lattice parameters

The TF lattice parameters L and K determine the oversampling factor $\xi = K/L$ that, in turn, determines the computational complexity of the filter. Choosing L larger or K smaller results in a coarser TF sampling lattice and thus a smaller computational complexity, but also yields a denser adjoint lattice, that is, more closely spaced translates $\widehat{M}(m - k'K, \nu - l'/L)$ in $\widehat{M}_p(m, \nu)$. This may cause the translates to overlap, which means that the sampling $M[l, k] = M(lL, k/K)$ results in aliasing. Here, the support regions $\mathcal{S}_{\widehat{M}}$ and \mathcal{S}_p overlap and the design rule in Equation (6.69) cannot be satisfied.

For further analysis, let us consider a TF weight function $M(n, \theta)$ whose FT $\widehat{M}(m, \nu)$ is effectively supported within a rectangular region $\mathcal{S}_{\widehat{M}} = [-m_0, m_0] \times [-\nu_0, \nu_0]$ of area $A = 4m_0\nu_0$. In this case, aliasing in $\widehat{M}_p(m, \nu)$ is prevented if and only if

$$\frac{1}{L} > 2\nu_0 \text{ and } K > 2m_0. \quad (6.70)$$

This can be satisfied only if $K/L > 4m_0\nu_0 = A$ (i.e., if the oversampling factor satisfies $\xi > A$). Hence, a smoother TF weight function (smaller A) allows for a

smaller ξ . If A is very small, ξ may be significantly smaller than 1 (undersampling). However, this may make it very difficult to satisfy the design rule in Equation (6.69) by appropriate choice of the windows $\gamma[n]$ and $g[n]$.

Once that the oversampling factor $\xi = K/L$ (i.e., the product of K and $1/L$) has been chosen, it remains to choose the ratio of K and $1/L$. A choice that is matched to the support region $\mathcal{S}_{\widehat{M}} = [-m_0, m_0] \times [-\nu_0, \nu_0]$ is given by [35, 75]

$$\frac{K}{1/L} = KL \stackrel{!}{\approx} \frac{m_0}{\nu_0} \quad (6.71)$$

(usually, this can be satisfied only approximately because $K, L \in \mathbb{Z}$). This leads to the rules

$$L = \text{round} \left\{ \sqrt{\frac{m_0}{\xi \nu_0}} \right\}, \quad K = \text{round} \left\{ \sqrt{\frac{\xi m_0}{\nu_0}} \right\}, \quad (6.72)$$

where $\text{round}\{\cdot\}$ denotes the rounding operation. The resulting separation (gap) between the effective support regions of the translates $\widehat{M}(m - k'K, \nu - l'/L)$ is then approximately $(\sqrt{\frac{\xi}{m_0 \nu_0}} - 2)m_0$ in the m direction and $(\sqrt{\frac{\xi}{m_0 \nu_0}} - 2)\nu_0$ in the ν direction.

6.6.2.2 Choice of windows

Next, we consider the choice of the windows $\gamma[n]$ and $g[n]$. We assume that the TF lattice parameters L and K have been chosen such that aliasing is prevented, that is, the support regions $\mathcal{S}_{\widehat{M}}$ and \mathcal{S}_p do not overlap. According to our design rule in Equation (6.69), we have to choose the windows $\gamma[n]$ and $g[n]$ such that $A_{g,\gamma}(m, \nu) \approx L/K$ on $\mathcal{S}_{\widehat{M}}$ (this means that the fundamental component $\widehat{M}(m, \nu)$ is preserved in Equation (6.68)), but $A_{g,\gamma}(m, \nu) \approx 0$ on \mathcal{S}_p ; this means that the image components $\widehat{M}(m - k'K, \nu - l'/L)$ with $k' \neq 0, l' \neq 0$ are suppressed in Equation (6.68).

The design of windows satisfying $A_{g,\gamma}(m, \nu) \approx L/K$ on $\mathcal{S}_{\widehat{M}}$ was discussed in the context of the STFT filter (see Subsection 6.5.2; note that it suffices to scale one of the windows to obtain $A_{g,\gamma}(m, \nu) \approx L/K$ instead of $A_{g,\gamma}(m, \nu) \approx 1$). If the windows are well behaved, $A_{g,\gamma}(m, \nu)$ decays to zero as we move sufficiently far away from the origin. This means that our second condition, $A_{g,\gamma}(m, \nu) \approx 0$ on \mathcal{S}_p , also is satisfied provided that the adjoint TF lattice is not too dense (i.e., the oversampling factor $\xi = K/L$ is not too small). The design of windows satisfying the preceding conditions is easier if broader gaps occur between the translates $\widehat{M}(m - k'K, \nu - l'/L)$ and, thus, between the support regions $\mathcal{S}_{\widehat{M}}$ and \mathcal{S}_p .

We can conclude from this discussion that the Zadeh filter can easily and accurately be approximated by a Gabor filter provided that the TF weight function $M(n, \theta)$ is reasonably smooth and the oversampling factor $\xi = K/L$ is not too small. These conditions are significantly relaxed in the case of the multiwindow Gabor filter discussed in Subsections 6.6.4 and 6.6.5.

6.6.3 On-line implementation of Gabor filter

For an on-line implementation of the Gabor filter, we assume that $\gamma[n]$ and $g[n]$ are supported in the intervals $[-L_\gamma/2, L_\gamma/2 - 1]$ and $[-L_g/2, L_g/2 - 1]$, respectively, with even window lengths L_γ and L_g . As an important difference from the on-line STFT filter in Subsection 6.5.3, the number of frequency bins K is allowed to be smaller than the window lengths (indeed, this is the usual case); this effectively corresponds to a subsampling with respect to frequency. Let $Q_\gamma \triangleq 2\lceil \frac{L_\gamma}{2K} \rceil$ be the smallest even number $\geq L_\gamma/K$ and $Q_g \triangleq 2\lceil \frac{L_g}{2L} \rceil$ the smallest even number $\geq L_g/L$. An efficient on-line implementation of the Gabor filter is then given as follows[¶] [17].

1. *Analysis.* The Gabor transform in Equation (6.63) of the input signal $x[n]$ is calculated at time $n = lL$:

$$\begin{aligned} \text{GT}_x^{(\gamma)}[l, k] &= \sum_{n=lL-L_\gamma/2}^{lL+L_\gamma/2-1} x[n] \gamma^*[n-lL] e^{-j2\pi \frac{k}{K}(n-lL)} \\ &= \sum_{n=lL-Q_\gamma K/2}^{lL+Q_\gamma K/2-1} x[n] \gamma^*[n-lL] e^{-j2\pi \frac{Q_\gamma k}{Q_\gamma K}(n-lL)}. \end{aligned} \quad (6.73)$$

This could be implemented as a DFT with length $Q_\gamma K$ followed by a subsampling by the factor Q_γ . However, we can rewrite Equation (6.73) as

$$\text{GT}_x^{(\gamma)}[l, k] = \sum_{m=0}^{K-1} \tilde{x}_l^{(\gamma)}[m] e^{-j2\pi \frac{k}{K} m}, \quad k = 0, 1, \dots, K-1,$$

with the prealiased length- K sequence:

$$\tilde{x}_l^{(\gamma)}[m] \triangleq \sum_{i=-Q_\gamma/2}^{Q_\gamma/2-1} x_l^{(\gamma)}[m+iK], \quad m = 0, 1, \dots, K-1,$$

where $x_l^{(\gamma)}[m] \triangleq x[lL+m] \gamma^*[m]$. Thus, the DFT length has been reduced by the factor Q_γ . This prealiasing amounts to breaking up the locally windowed segment $x_l^{(\gamma)}[m]$ of the input signal $x[n]$ around lL into Q_γ blocks of length K each and adding these blocks as shown in Figure 6.18. Note that the Gabor analysis step at time instant $n = lL$ involves the input signal samples $x[lL-L_\gamma/2], x[lL-L_\gamma/2+1], \dots, x[lL+L_\gamma/2-1]$.

2. *Weighting.* The Gabor transform samples are multiplied by $M[l, k] = M(lL, k/K)$:

$$F[l, k] \triangleq M[l, k] \text{GT}_x^{(\gamma)}[l, k], \quad k = 0, 1, \dots, K-1.$$

[¶]For the cases of critical sampling and integer oversampling (i.e., $K/L \in \mathbb{N}$), alternative efficient implementations based on the polyphase representation can be obtained [17].

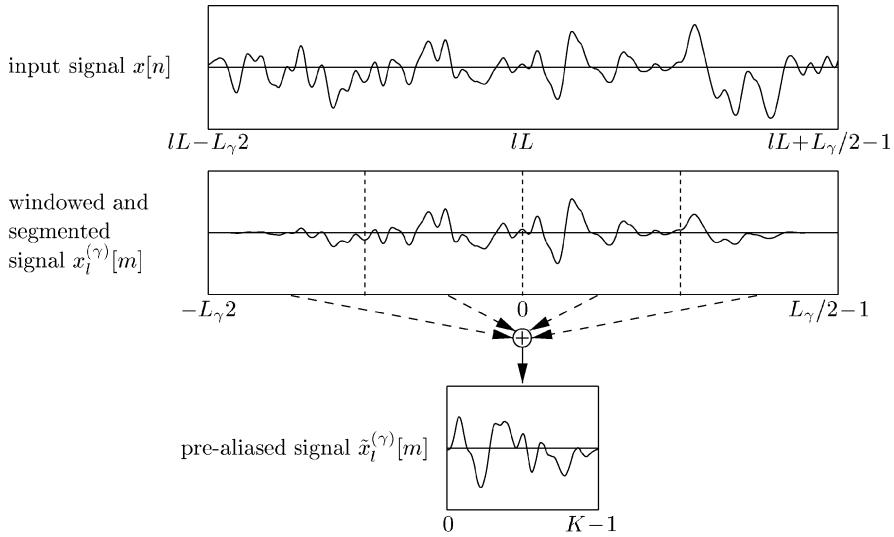
**FIGURE 6.18**

Illustration of the prealiasing (windowing–segmentation–summation) procedure used for Gabor analysis. The analysis window length is $L_\gamma = 4K$, corresponding to $Q_\gamma = 4$ blocks.

This amounts to the element-wise product of the two length- K vectors $(M[l, 0] M[l, 1] \cdots M[l, K-1])$ and $(\text{GT}_x^{(\gamma)}[l, 0] \text{GT}_x^{(\gamma)}[l, 1] \cdots \text{GT}_x^{(\gamma)}[l, K-1])$.

3. *Synthesis.* The Gabor synthesis step at time $n = lL$ produces a segment of the output signal $y[n]$ of length L , specifically, $y[n]$ for $n = lL, lL+1, \dots, (l+1)L-1$, and it involves $F[l', k]$ (and, thus, the TF weight function $M[l', k]$) for $l' = l - Q_g/2 + 1, l - Q_g/2 + 2, \dots, l + Q_g/2$.

The Gabor synthesis relation in Equation (6.66) can be rewritten as

$$y[n] = \sum_{l'} y_{l'}^{(g)}[n - l'L], \quad (6.74)$$

where

$$y_{l'}^{(g)}[n] \triangleq K f_{l'}[n] g[n]$$

with

$$f_{l'}[n] \triangleq \frac{1}{K} \sum_{k=0}^{K-1} F[l', k] e^{j2\pi \frac{k}{K} n},$$

and the summation limits are considered presently. In Equation (6.74), $y_{l'}^{(g)}[n - l'L]$ is the component of the output signal $y[n]$ that is contributed by $F[l', k]$. Because $y_{l'}^{(g)}[n]$ is supported within $[l'L - L_g/2, l'L + L_g/2 - 1]$, only Q_g such output signal components overlap. Indeed, within the interval $[lL, (l+1)L-1]$ associated to time lL , $y[n]$ is made up by the components

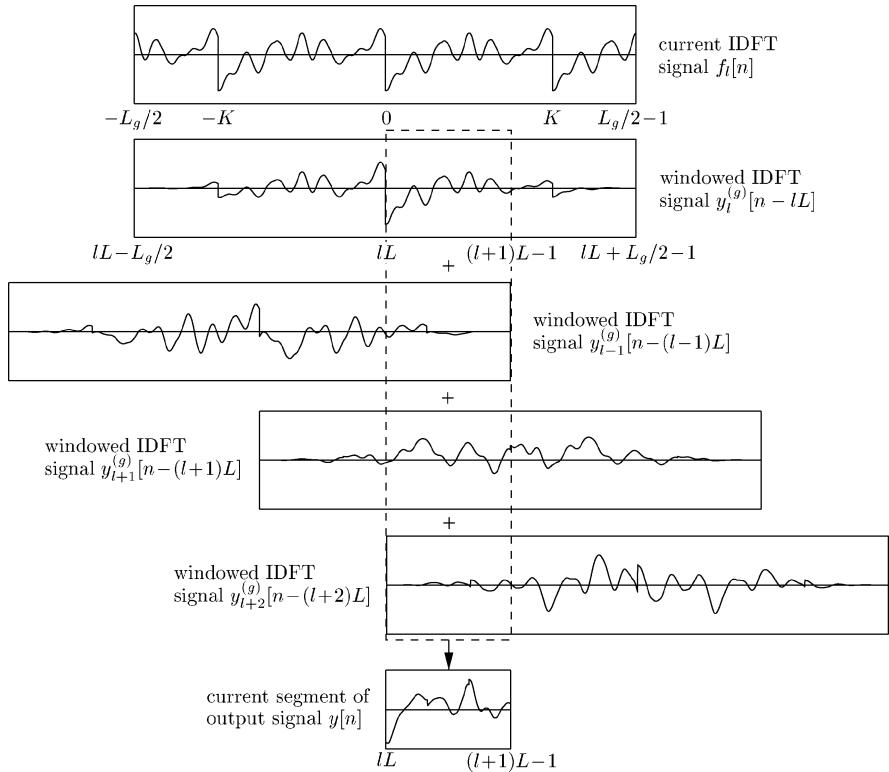
**FIGURE 6.19**

Illustration of the overlap-add procedure used for Gabor synthesis. The synthesis window length is $L_g = 4L$, corresponding to $Q_g = 4$ output signal components $y_l^{(g)}[n]$ added.

$y_{l'}^{(g)}[n - l'L]$ with $l' = l - Q_g/2 + 1, l - Q_g/2 + 2, \dots, l + Q_g/2$, so that Equation (6.74) becomes:

$$y[lL+i] = \sum_{l'=l-Q_g/2+1}^{l+Q_g/2} y_{l'}^{(g)}[lL+i-l'L] = \sum_{l'=-Q_g/2}^{Q_g/2-1} y_{l-l'}^{(g)}[i+l'L] \quad (6.75)$$

with $i = 0, 1, \dots, L-1$. (Note that for $L = 1$, this reduces to Equation (6.53).) This synthesis procedure can be viewed as an overlap-add scheme [16, 17] (see Figure 6.19). In particular, let us consider the right-hand side of Equation (6.75) extended to the larger interval $i = 0, 1, \dots, Q_g L - 1$:

$$s_i[l] \triangleq \sum_{l'=-Q_g/2}^{Q_g/2-1} y_{l-l'}^{(g)}[i+l'L], \quad i = 0, 1, \dots, Q_g L - 1.$$

Note that $s_i[l]$ equals $y[lL + i]$ for $i = 0, 1, \dots, L - 1$ but not for $i = L, L + 1, \dots, Q_g L - 1$. Furthermore, $s_i[l]$ can be calculated recursively as:

$$s_i[l] = \begin{cases} s_{i+L}[l-1] \\ + y_{l+Q_g/2}^{(g)} \left[i - \frac{Q_g L}{2} \right], & i = 0, 1, \dots, (Q_g - 1)L - 1 \\ y_{l+Q_g/2}^{(g)} \left[i - \frac{Q_g L}{2} \right], & i = (Q_g - 1)L, (Q_g - 1)L + 1, \dots, Q_g L - 1 \end{cases}$$

which corresponds to adding the new overlapping output signal component. The current output signal block is finally obtained as $y[lL + i] = s_i[l]$ for $i = 0, 1, \dots, L - 1$.

Because Equation (6.75) involves future values $F[l + 1, k], F[l + 2, k], \dots, F[l + Q_g/2, k]$ which, in turn, involve the input $x[n]$ up to time $n = lL + L_\gamma/2 + Q_g L/2$, causal implementation of this algorithm requires the introduction of a time delay of $L_\gamma/2 + Q_g L/2$ samples. At a given time instant $n = lL$ (i.e., every L samples), the Gabor analysis, weighting and synthesis steps use $\mathcal{O}(L_\gamma + K \log K)$, $\mathcal{O}(K)$, and $\mathcal{O}(L_g + K \log K)$ operations, respectively. The total number of operations per output signal sample is thus approximately $\mathcal{O}\left(\frac{K}{L}(2 \log K + 1) + \frac{L_\gamma + L_g}{L}\right)$, which is essentially proportional to the oversampling factor $\xi = K/L$. As long as ξ is not too large, the on-line implementation of the Gabor filter is highly efficient. The memory requirements are similarly favorable: we have to store L_γ input signal samples and the length- $Q_g L$ vector $(s_0[l] \ s_1[l] \ \dots \ s_{Q_g L - 1}[n])$, which requires $L_\gamma + Q_g L \approx L_\gamma + L_g$ memory locations.

6.6.4 Multiwindow Gabor filter

In analogy to the multiwindow STFT filter in Subsections 6.5.4 and 6.5.5, the *multiwindow Gabor filter* \mathbf{H}_{MG} is defined as a weighted sum of N Gabor filters $\mathbf{H}_G^{(i)}$ with different windows $\gamma_i[n], g_i[n]$ but the same TF lattice parameters L, K and TF weight function $M(n, \theta)$ [9]:

$$\mathbf{H}_{\text{MG}} \triangleq \sum_{i=1}^N \eta_i \mathbf{H}_G^{(i)}.$$

Note that the Gabor filter \mathbf{H}_G is a special case with $N = 1$. Using a larger number N of Gabor filters allows reduction of the TF sampling density K/L (see [76–78]). The oversampling factor is now defined as $\xi \triangleq NK/L$, and the cases of critical sampling, oversampling and undersampling are defined by $NK = L$ ($\xi = 1$), $NK > L$ ($\xi > 1$) and $NK < L$ ($\xi < 1$), respectively.

The impulse response of the multiwindow Gabor filter \mathbf{H}_{MG} is obtained as

$$h_{\text{MG}}[n, n'] = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} M[l, k] p[n - lL, n' - lL] e^{j2\pi \frac{k}{K}(n - n')},$$

with $p[n, n'] = \sum_{i=1}^N \eta_i g_i[n] \gamma_i^*[n']$ as in Equation (6.54). We again interpret $p[n, n']$ as the impulse response of a “prototype system” \mathbf{P} . It can be shown [37, 76–78] that

perfect reconstruction (that is, $(\mathbf{H}_{MG}x)[n] = x[n]$ for $M[l, k] \equiv 1$) is satisfied if and only if

$$S_{\mathbf{P}}\left(k'K, \frac{l'}{L}\right) = \frac{L}{K} \delta[k'] \delta[l'], \quad (6.76)$$

which generalizes (6.65). If Equation (6.76) is met, the window sets $\{\gamma_i[n]\}$ and $\{\eta_i g_i[n]\}$ are said to be *dual* [37, 76–78]. Perfect reconstruction cannot be obtained in the undersampled case ($\xi < 1$). Individual duality (biorthogonality) of each window pair $\gamma_i[n], g_i[n]$ (that is, each window pair individually satisfies Equation (6.65)) together with $\sum_{i=1}^N \eta_i = 1$ is sufficient but not necessary for Equation (6.76) to hold. In fact, pairwise duality is possible only for $K/L \geq 1$, corresponding to an oversampling factor $\xi \geq N$. Frame-theoretical methods for the design of dual window sets are discussed in [76–78]. There, it is also shown that a larger oversampling factor ξ yields more freedom in the design of dual window sets and, thus, of the prototype system \mathbf{P} . However, in the context of TF filters, approximate perfect reconstruction (duality) is usually sufficient.

6.6.4.1 Comparison of multiwindow Gabor filter with Zadeh filter

The Zadeh function of the multiwindow Gabor filter \mathbf{H}_{MG} is given by:

$$Z_{\mathbf{H}_{MG}}(n, \theta) = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} M[l, k] Z_{\mathbf{P}}\left(n - lL, \theta - \frac{k}{K}\right)$$

which generalizes Equation (6.67). By comparing with Equation (6.6), we see that \mathbf{H}_{MG} is the Zadeh filter associated to the TF weight function $\widehat{M}(n, \theta) \triangleq \sum_{l=-\infty}^{\infty} \sum_{k=0}^{K-1} M[l, k] Z_{\mathbf{P}}(n - lL, \theta - k/K)$. This can be interpreted as a 2-D interpolation of the subsampled TF weight function $M[l, k] = M(lL, k/K)$. The implications of this subsampling and interpolation are analyzed in Subsection 6.6.5.

6.6.4.2 On-line implementation of multiwindow Gabor filter

The on-line implementation of the multiwindow Gabor filter \mathbf{H}_{MG} essentially reduces to the individual on-line implementations of the N Gabor filters $\mathbf{H}_G^{(i)}$ (see Subsection 6.6.3). This makes \mathbf{H}_{MG} ideally suited to parallel processing. The computational complexity of \mathbf{H}_{MG} does not have to be significantly higher than that of a single-window Gabor filter because \mathbf{H}_{MG} typically uses a reduced TF sampling density K/L . The overall number of operations required per output signal sample can be estimated as $\mathcal{O}\left(\frac{NK}{L}(2 \log K + 1) + \frac{N}{L}(\bar{L}_\gamma + \bar{L}_g)\right)$, with $\bar{L}_\gamma \triangleq \frac{1}{N} \sum_{i=1}^N L_{\gamma_i}$ and $\bar{L}_g \triangleq \frac{1}{N} \sum_{i=1}^N L_{g_i}$. This is essentially proportional to the oversampling factor $\xi = NK/L$. The memory requirements, on the other hand, are higher than in the single-window case. In particular, whereas the relevant segment of the input signal

^{||}For consistency with the established notion of duality, we temporarily incorporate the branch weights η_i in one of the window sets.

(see [Subsection 6.6.3](#)) is essentially the same for all filter branches and needs to be stored only once, the overlap and add procedure requires L_{g_i} memory locations for each branch. This amounts to a total of $\max_{i=1,2,\dots,N}\{L_{\gamma_i}\} + \sum_{i=1}^N L_{g_i}$ memory locations.

6.6.5 Choice of lattice parameters and prototype system

Yet to be considered is the choice of the TF lattice parameters L, K and the prototype system \mathbf{P} involving the design parameters $\gamma_i[n], g_i[n], \eta_i$, and N . Our discussion generalizes that in [Subsection 6.6.2](#). It is based on the following generalization of [Equation \(6.68\)](#):

$$S_{\mathbf{H}_{\text{MG}}}(m, \nu) = \frac{K}{L} \widehat{M}_p(m, \nu) S_{\mathbf{P}}(m, \nu) \quad (6.77)$$

where again $\widehat{M}_p(m, \nu) = \sum_{k'=-\infty}^{\infty} \sum_{l'=0}^{L-1} \widehat{M}(m - k'K, \nu - l'/L)$.

Under the assumption that $M(n, \theta)$ is reasonably smooth, we desire that \mathbf{H}_{MG} approximates the Zadeh filter, that is, $Z_{\mathbf{H}_{\text{MG}}}(n, \theta) \stackrel{!}{\approx} M(n, \theta)$ or, equivalently

$$S_{\mathbf{H}_{\text{MG}}}(m, \nu) \stackrel{!}{\approx} \widehat{M}(m, \nu).$$

By comparing with [Equation \(6.77\)](#), we see that this can be satisfied if and only if (see [Figure 6.17](#)):

$$S_{\mathbf{P}}(m, \nu) \approx \begin{cases} L/K, & (m, \nu) \in \mathcal{S}_{\widehat{M}} \\ 0, & (m, \nu) \in \mathcal{S}_p \\ \text{arbitrary} & \text{otherwise,} \end{cases} \quad (6.78)$$

with $\mathcal{S}_{\widehat{M}}$ and \mathcal{S}_p as in [Subsection 6.6.2](#). This design rule generalizes [Equation \(6.69\)](#); also note the difference from the multiwindow STFT filter design rule in [Equation \(6.58\)](#). As a by-product, [Equation \(6.78\)](#) ensures that

$$S_{\mathbf{P}}\left(k'K, \frac{l'}{L}\right) \approx \frac{L}{K} \delta[k'] \delta[l'], \quad (6.79)$$

which means that the perfect-reconstruction condition in [Equation \(6.76\)](#) is approximately satisfied. Conversely, however, [Equation \(6.76\)](#) does not necessarily imply [Equation \(6.78\)](#) (see the discussion in [Subsection 6.6.2](#)). We next show how to choose TF lattice parameters and a prototype system satisfying [Equation \(6.78\)](#).

6.6.5.1 Choice of TF lattice parameters

The choice of the TF lattice parameters L and K is analogous to the single-window case discussed in [Subsection 6.6.2](#). The only difference is that the product of K and $1/L$, denoted $\zeta \triangleq K/L$ hereafter, is no longer equal to the oversampling factor $\xi = NK/L$. Thus, all relevant results from [Subsection 6.6.2](#) also apply to the multiwindow Gabor filter if ξ is formally replaced with ζ . In particular, once that the TF lattice

density $\zeta = K/L$ has been chosen such that $\zeta > 4m_0\nu_0$, Equation (6.72) yields $L = \text{round}\left\{\sqrt{\frac{m_0}{\zeta\nu_0}}\right\}$ and $K = \text{round}\left\{\sqrt{\frac{\zeta m_0}{\nu_0}}\right\}$. The separation (gap) between the translates $\widehat{M}(m - k'K, \nu - l'/L)$ is then approximately $(\sqrt{\frac{\zeta}{m_0\nu_0}} - 2)m_0$ in the m direction and $(\sqrt{\frac{\zeta}{m_0\nu_0}} - 2)\nu_0$ in the ν direction.

6.6.5.2 Choice of prototype system

Under the assumption that L and K have been chosen such that aliasing is prevented (see previous section) a prototype system \mathbf{P} satisfying Equation (6.78) can be designed using an extension of the second design method of Subsection 6.5.5. We define $\mathcal{S}_{\widehat{M}}$ as the smallest elliptical region that contains the effective support region of $\widehat{M}(m, \nu)$. Let m_0 and ν_0 be the semiaxes of this elliptical region. Furthermore, we use a second, larger elliptical region \mathcal{S}_1 with semiaxes $m_1 = \mu m_0 > m_0$ and $\nu_1 = \mu \nu_0 > \nu_0$ that contains $\mathcal{S}_{\widehat{M}}$ but no part of \mathcal{S}_p . The analysis and synthesis windows are chosen as the first N sampled Hermite functions, that is, $\gamma_i[n] = g_i[n] = H_i[n]$ for $i = 1, 2, \dots, N$, with scaling parameter $\Delta = \sqrt{m_0/\nu_0}$ (see Equation (6.60)). Our design rule in Equation (6.78) then leads to the following modification of Equation (6.62):

$$\sum_{i=1}^N \eta_i L_{i-1}(\beta) \stackrel{!}{\approx} \begin{cases} L/K, & \beta \leq \beta_0 \\ \text{rolloff}, & \beta_0 < \beta \leq \beta_1 \\ 0, & \beta > \beta_1, \end{cases} \quad (6.80)$$

where $\beta_0 \triangleq \pi m_0 \nu_0$, $\beta_1 \triangleq \pi m_1 \nu_1$. Thus, the calculation of the branch weights η_i again corresponds to a curve-fitting problem that can be solved using a weighted least-squares method; the largely arbitrary shape of $\sum_{i=1}^N \eta_i L_{i-1}(\beta)$ in the rolloff interval $(\beta_0, \beta_1]$ can be taken into account by attaching very small weights to the associated approximation errors. This design is easier for a larger oversampling factor $\xi = NK/L$: a larger K/L increases the separation between $\mathcal{S}_{\widehat{M}}$ and \mathcal{S}_p and a larger N increases the overall design freedom in designing \mathbf{P} .

We can conclude that for $M(n, \theta)$ reasonably smooth, the prototype system \mathbf{P} can be designed such that the Zadeh filter is accurately approximated by a multiwindow Gabor filter.

6.6.5.3 Design example

We present a design example for a TF weight function characterized by $m_0 = 5$ and $\nu_0 = 2 \cdot 10^{-4}$. We use the TF lattice density $\zeta = K/L = 1/2$ and the matched lattice parameters $L = \text{round}\left\{\sqrt{\frac{m_0}{\zeta\nu_0}}\right\} = 224$ and $K = L/2 = 112$. The separation (gap) between the translates is then about $20.36 m_0$ in the m direction and $20.36 \nu_0$ in the ν direction. Thus, aliasing is prevented. For the design of the prototype system \mathbf{P} , we use Hermite function windows $H_i[n]$ as discussed earlier. The results of the curve-fitting problem in Equation (6.80) with $\beta_0 = \pi m_0 \nu_0 = \pi 10^{-3}$ and $\beta_1 = 4\pi 10^{-1}$ are shown in Figure 6.20 for $N = 1, 2$, and 5 (i.e., oversampling factors $\xi = 1/2, 1$,

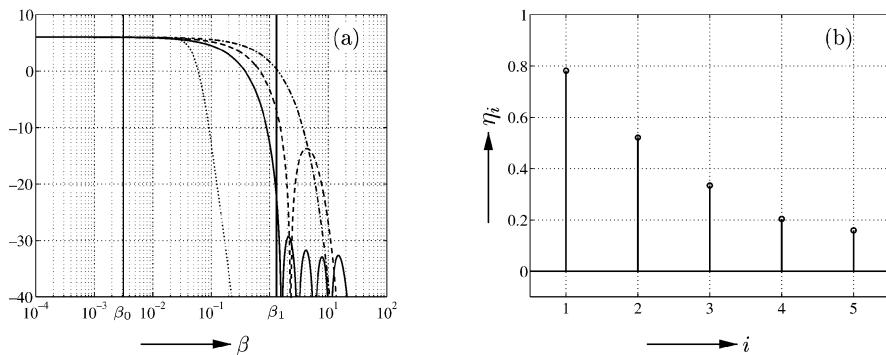


FIGURE 6.20

Examples of a multiwindow Gabor prototype system design using parameters $N = 1, 2$ or 5 , $\beta_0 = \pi m_0 \nu_0 = \pi 10^{-3}$ and $\beta_1 = 4\pi 10^{-1}$. (a) Target profile (dotted line) and weighted least-squares approximations obtained for $N = 1$ (dash-dotted line), $N = 2$ (dashed line) and $N = 5$ (solid line), all shown in dB; (b) resulting branch weights η_i for $N = 5$.

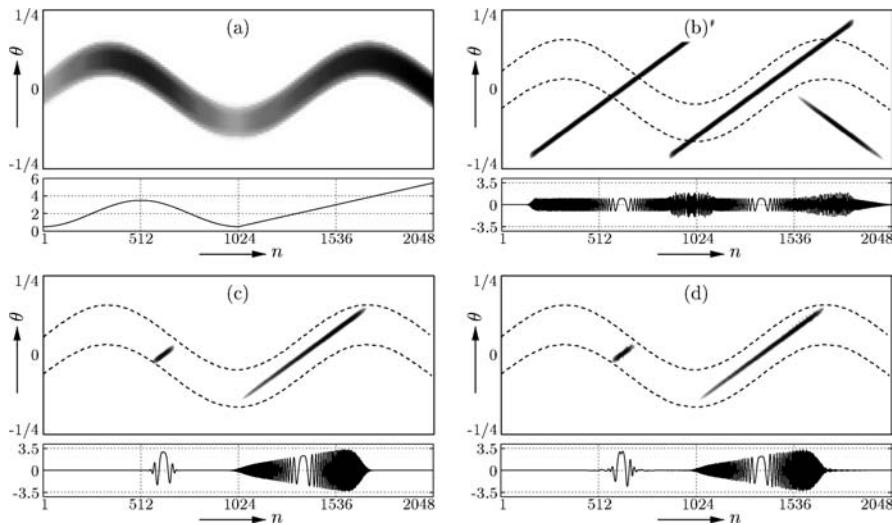
and $5/2$). It is seen that good approximation results (including approximate perfect reconstruction) are achieved already for moderate values of N .

6.7 Implicit Time–Frequency Filter Design III: Simulation Results and Applications

In this section, we present simulation results for the on-line STFT and Gabor filters and their multiwindow extensions. The applications considered include the separation and denoising of FM signals and adaptive nonstationary speech enhancement. Among other results, we demonstrate that the STFT filter has relatively poor TF selectivity but it is advantageous for nonsmooth TF weight functions containing oscillatory components.

6.7.1 Time-varying bandpass filter

We first reconsider the TF weight function $M(n, \theta)$ and input signal $x[n]$ from Sub-section 6.4.1 (see Figure 6.5), shown again in Figure 6.21(a) and (b) for convenience. The output signals obtained with an on-line STFT filter \mathbf{H}_S and an on-line Gabor filter \mathbf{H}_G are shown in Figure 6.21(c) and (d). The on-line STFT filter uses $K = 64$ frequency bins and equal analysis and synthesis windows ($g[n] = \gamma[n]$) of Hanning type and length 48. The on-line Gabor filter uses TF lattice parameters $L = 16$ and $K = 32$ (hence, oversampling factor $\xi = 2$), a low-pass synthesis window of length 128 (designed using the MATLAB function **fir1** [79]), and the minimum dual analysis window (calculated as discussed in [37] and truncated to length 256). It is seen

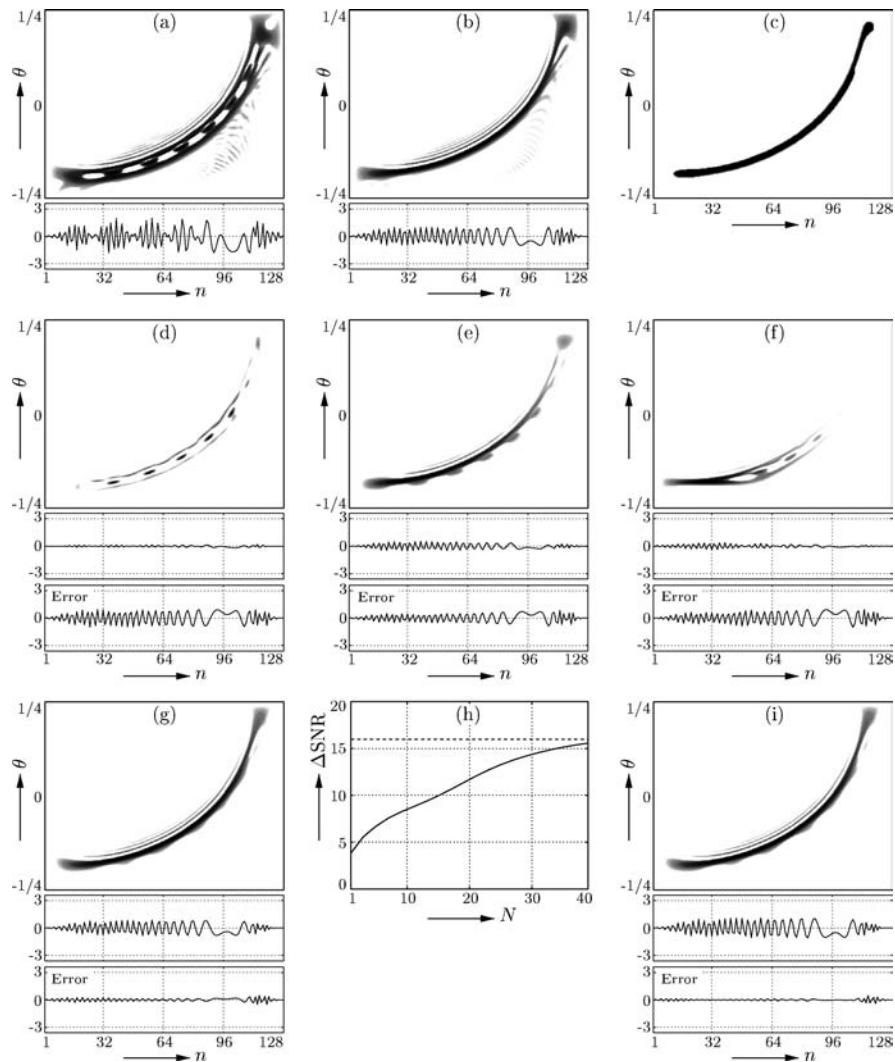
**FIGURE 6.21**

Time-varying bandpass filtering (see Figure 6.5). (a) TF weight function (lower part: time-varying gain); (b) input signal (upper part: SPWD, lower part: real part); (c) output signal of STFT filter; (d) output signal of Gabor filter. In (b) to (d), the dashed lines indicate the TF pass region defined by the TF weight function.

that both filters succeed in implementing the desired TF weighting. Their output signals are quite similar and, moreover, also similar to those of the Zadeh filter and halfband Weyl filter that were shown in Figure 6.5(c) and (d). This similarity is due to the smoothness of $M(n, \theta)$ which causes all TF filters to be underspread.

6.7.2 Separation and denoising of frequency-modulated signals

We next present an example where the smoothing of the TF weight function that is inherent to the STFT filter — see Equation (6.46) — results in a loss of TF resolution (or TF selectivity). The input signal of length 128 consists of two closely spaced FM signals, as shown in Figure 6.22(a). The filtering task is to suppress the lower FM component while passing the upper FM component (the latter is shown in Figure 6.22(b)). The TF weight function was discretized using $K = 128$ frequency bins; it is shown in Figure 6.22(c). We implemented three different on-line STFT filters using equal analysis and synthesis windows of Hanning type and lengths 4, 32 and 128, respectively. As shown in Figure 6.22(d) to (f), all filters pass significant parts of the lower (undesired) FM component and attenuate the upper (desired) FM component. (The attenuation results from the smoothing of the TF weight function inherent to STFT filters.) As a consequence, the respective SNR improvements were only 0.72, 4.01 and 1.47 dB. The STFT filter with medium window length 32 performs relatively best; this approximately agrees with the optimum effective window duration according to Equation (6.50) that was obtained as $N_\gamma = 28$.

**FIGURE 6.22**

Separation of FM signals. (a) Input signal (upper part: SPWD, lower part: real part); (b) desired signal component; (c) TF weight function; (d) to (f) output signals of STFT filters using (d) a short window, (e) a medium-length window, and (f) a long window; (g) output signal of multiwindow STFT filter with $N = 21$ branches; (h) SNR improvement (in dB) of multiwindow STFT filter vs. N (for comparison, the dashed line shows the SNR improvement of the approximate halfband Weyl filter); (i) output signal of approximate halfband Weyl filter. The bottom parts of (d) to (g) and (i) show the respective error signal (filter output signal minus desired signal).

Next, by using the first design method in Subsection 6.5.5, we designed multiwindow STFT filters \mathbf{H}_{MS} with various numbers N of STFT filter branches. The output signal obtained with $N = 21$ is depicted in Figure 6.22(g); the SNR improvement was 12.04 dB. Figure 6.22(h) shows that the SNR improvement grows with increasing N and, for relatively large values of N , approaches that of the halfband Weyl filter $\tilde{\mathbf{H}}_{\text{HW}}$ (15.98 dB). The output signal of $\tilde{\mathbf{H}}_{\text{HW}}$ is shown in Figure 6.22(i) for comparison. In general, \mathbf{H}_{MS} performs worse than $\tilde{\mathbf{H}}_{\text{HW}}$ and its computational and memory requirements are higher. We can conclude that STFT filters and multiwindow STFT filters with small N are not well suited to applications requiring very high TF selectivity.

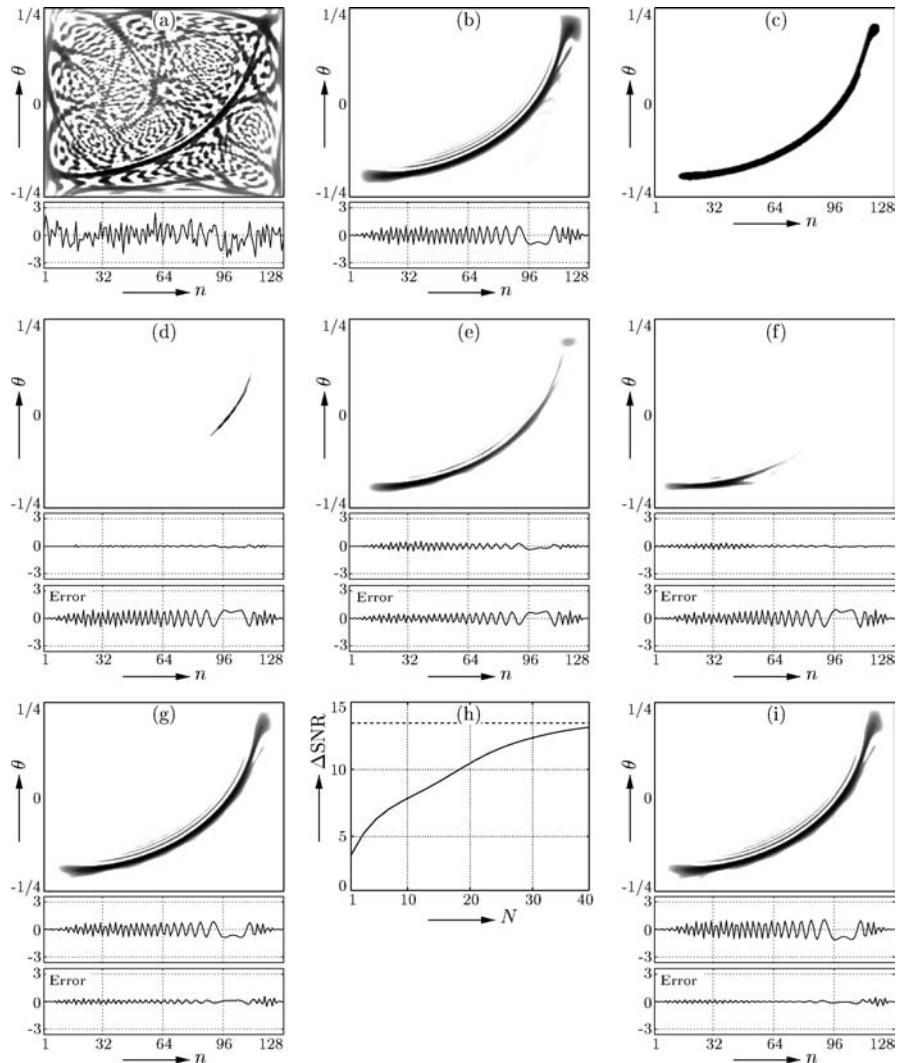
In Figure 6.23, analogous results are shown for a noisy input signal $x[n] = s[n] + w[n]$, where $s[n]$ is an FM signal and $w[n]$ is halfband-restricted white noise. The SNR is $\frac{\|s\|^2}{128} / \mathbb{E}\{|w[n]|^2\} = 0$ dB. STFT filters with $K = 128$ and equal analysis and synthesis windows of length 4, 32 and 128 achieved SNR improvements of 1.81, 5.27 and 2.63 dB, respectively. A multiwindow STFT filter with $N = 21$ achieved 12.9 dB, which is close to the 13.56 dB obtained with the halfband Weyl filter $\tilde{\mathbf{H}}_{\text{HW}}$. Again, due to the high TF selectivity required, STFT filters and multiwindow STFT filters with small N have significantly poorer performance than $\tilde{\mathbf{H}}_{\text{HW}}$.

STFT filters can be advantageous in the case of a nonsmooth TF weight function $M[n, k]$ with oscillatory components. Let us reconsider the final chirp separation example from Subsection 6.4.2 (see Figure 6.8). The input $x[n]$, the desired component and the TF weight function $M[n, k]$ are shown again in Figure 6.24(a) to (c) for convenience. Figure 6.24(d) shows the output signal of an STFT filter with a Hanning analysis and synthesis window of length 55. The undesired chirp component is well suppressed, without significant parasitic TF shift effects as those produced by the Zadeh filter \mathbf{H}_Z and halfband Weyl filter $\tilde{\mathbf{H}}_{\text{HW}}$; see Figure 6.8(d) and (e). However, the desired component is attenuated, which causes an SNR degradation by 4.95 dB. This is better than \mathbf{H}_Z (SNR degradation by 11.12 dB) but still poorer than $\tilde{\mathbf{H}}_{\text{HW}}$ (1.56 dB). Motivated by the oblique orientation of $M[n, k]$, we next used an STFT filter with a Hanning window of length 128 modulated by a linear chirp signal with chirp rate $c = 0.7/256$. The output signal is shown in Figure 6.24(e). The filter now achieves an SNR improvement of 2.7 dB, which is 4.26 dB better than $\tilde{\mathbf{H}}_{\text{HW}}$. Figure 6.24(f) shows that even better results can be obtained with chirpy analysis and synthesis windows of different lengths (here chosen as 256 and 16, respectively; the chirp rates of the windows are as before). This filter achieves an SNR improvement of 6.82 dB.

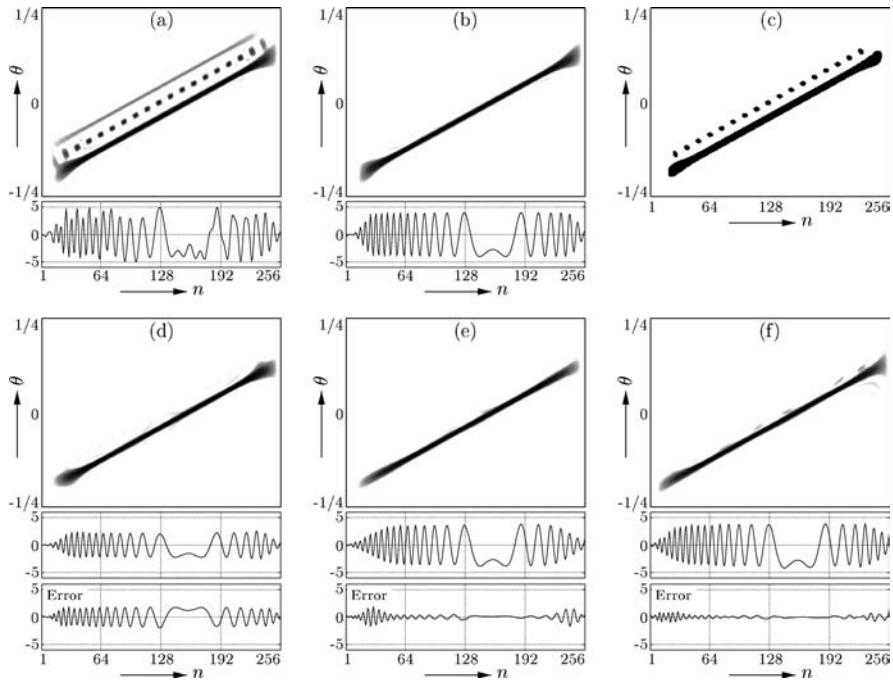
6.7.3 Adaptive speech enhancement

We finally consider the enhancement (denoising) of noisy speech. The speech signal $s[n]$ of length 8000 samples and its noisy version $x[n] = s[n] + w[n]$ are shown in Figure 6.25(a) and (b). The noise $w[n]$ is white; the SNR is 0 dB. The noise variance σ_w^2 is assumed to be known.

We use a multiwindow Gabor filter because it is computationally efficient and allows for a stable on-line estimation of signal statistics. The number of Gabor

**FIGURE 6.23**

Denoising of an FM signal. (a) Input signal (upper part: SPWD, lower part: real part); (b) desired signal component; (c) TF weight function; (d) to (f) output signals of STFT filters using (d) a short window, (e) a medium-length window, and (f) a long window; (g) output signal of multiwindow STFT filter with $N = 21$ branches; (h) SNR improvement (in dB) of multiwindow STFT filter vs. N (for comparison, the dashed line shows the SNR improvement of the approximate halfband Weyl filter); (i) output signal of approximate halfband Weyl filter.

**FIGURE 6.24**

Separation of chirp signals using a nonsmooth (oscillatory) TF weight function (see Figure 6.8). (a) Input signal (upper part: SPWD, lower part: real part); (b) desired signal component; (c) TF weight function; (d) to (f) output signals of STFT filters using (d) equal analysis and synthesis windows of length 55, (e) equal chirped analysis and synthesis windows of length 128, (f) different chirped analysis and synthesis windows of lengths $L_\gamma = 256$ and $L_g = 16$.

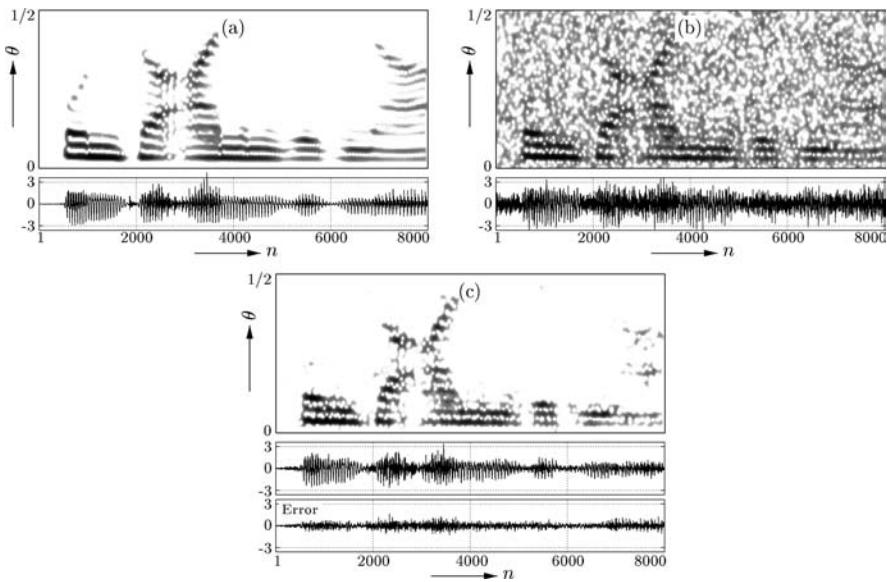
filter branches is $N = 10$. The lattice parameters are $L = 96$ and $K = 48$; thus, the oversampling factor is $\zeta = NK/L = 5$. The analysis and synthesis windows $\gamma_i[n] = g_i[n]$ are the first 10 Hermite functions designed for** $m_0 = 16$ and $\nu_0 = 2 \cdot 10^{-3}$ (compare with Subsections 6.5.5 and 6.6.5). The Hermite functions were truncated to length 256. The branch weights are equal $\eta_i = 1/10$ for $i = 1, 2, \dots, 10$.

By modeling $s[n]$ as a nonstationary random process and adopting a TF Wiener filter approach (see [12, 64, 80, 81]), we construct the TF weight function as:

$$M[l, k] = \frac{\overline{W}_s(lL, k/K)}{\overline{W}_s(lL, k/K) + \sigma_w^2} = 1 - \frac{\sigma_w^2}{\overline{W}_x(lL, k/K)}.$$

Here, $\overline{W}_s(lL, k/K)$ and $\overline{W}_x(lL, k/K) = \overline{W}_s(lL, k/K) + \sigma_w^2$ denote the (subsampled) Wigner–Ville spectra [45, 49–53] of $s[n]$ and $x[n] = s[n] + w[n]$, respectively.

**This choice of m_0 and ν_0 represents an educated guess because the true values are not known *a priori*.

**FIGURE 6.25**

Speech enhancement using a multiwindow Gabor filter. (a) Noiseless speech signal (upper part: SPWD, lower part: real part); (b) noisy speech signal; (c) filtered (enhanced) noisy speech signal.

Because $\overline{W}_x(lL, k/K)$ is unknown, we replace it with the multiwindow estimate [82, 83]:

$$\widehat{\overline{W}}_x\left(lL, \frac{k}{K}\right) \triangleq \frac{1}{N} \sum_{i=1}^N |\text{GT}_x^{(\gamma_i)}[l, k]|^2.$$

(Note that the $\text{GT}_x^{(\gamma_i)}[l, k]$ are calculated as part of the filter procedure.) Thus, $M[l, k]$ is designed in a signal-adaptive, on-line manner without prior knowledge about the statistics of the clean speech signal. The filter output is shown in Figure 6.25(c); the SNR improvement is 6.4 dB. The filter (including estimation) requires only about 120 operations per signal sample.

6.8 Conclusions

TF filters are a special concept of LTV filters: the design of the LTV filter is based on a TF representation of the filter or of the input signal. This is theoretically appropriate and practically useful in situations where the desired filter characteristic is most naturally specified in the TF domain by means of a TF weight function.

In this chapter, we considered two different approaches to such a TF design of an LTV filter \mathbf{H} . The *explicit* design consists of matching a suitable TF representation of \mathbf{H} to the prescribed TF weight function. The *implicit* design is actually a TF implementation of the filter: a linear TF representation of the input signal is multiplied by the TF weight function and subsequently the output signal is synthesized in a linear manner. We discussed the *Zadeh filter* and (various versions of) the *Weyl filter* as two explicit TF filter design schemes. Furthermore, we studied the *STFT filter*, the *multiwindow STFT filter*, the *Gabor filter*, and the *multiwindow Gabor filter* as implicit TF filter designs. Major properties of the TF filters studied can be summarized as follows:

- All filters allow on-line implementations and are thus suited to real-time applications and applications involving arbitrarily long signals.
- If the TF weight function is reasonably smooth (in particular, if it does not contain oscillatory components), then all filters tend to be *underspread* and to perform similarly. Specifically, the (multiwindow) STFT and Gabor filters can be designed such that they approximate the Zadeh filter or the Weyl filter. The underspread property means that the filters do not introduce undesired TF shifts.
- If, on the other hand, the TF weight function contains significant oscillatory components, then the Zadeh and Weyl filters are *overspread* (i.e., they introduce undesired TF shifts). However, the STFT filter and, if properly designed, also the Gabor filter are still underspread due to an inherent smoothing of the TF weight function.
- Because of this inherent smoothing, the STFT and Gabor filters have limited TF selectivity (TF resolution). Thus, they are not well suited to applications such as the separation of signal components that are very closely spaced in the TF plane.
- The Weyl filter is essentially restricted to halfband signals. In the case where fullband signals are to be processed, this requires an additional preprocessing (e.g., interpolation by two) and postprocessing (e.g., decimation by two).
- The Weyl filter is superior to the Zadeh filter when the TF weight function contains obliquely oriented (chirpy) components.
- The multiwindow STFT and Gabor filters overcome certain design restrictions of the single-window case. For example, the multiwindow STFT filter can have high TF selectivity. The multiwindow filters are furthermore attractive for nonstationary statistical signal processing because they allow a simple on-line estimation of signal statistics.
- Whereas explicit TF filter designs (Zadeh filter and Weyl filter) are fairly straightforward, implicit designs are more demanding because certain design

TABLE 6.1

Complexity (Computational Cost and Memory Requirements) of Various On-line TF Filters

Filter Type	Operations Per Sample	Memory Locations
Zadeh filter	$\mathcal{O}(K + K \log K)$	$K^2/2 + K$
Minimum-energy Weyl filter	$\mathcal{O}(K + K \log K)$	$K^2/2 + 2K$
Approximate halfband Weyl filter	$\mathcal{O}(K/2 + K/2 \log(K/2))$	$K^2/4 + K$
Multiwindow STFT filter	$\mathcal{O}\left(N(K + 2K \log K + \bar{L}_\gamma + \bar{L}_g)\right)$	$\max_{i=1,2,\dots,N} \{L_{\gamma_i}\} + N\bar{L}_g$
Multiwindow Gabor filter	$\mathcal{O}\left(\frac{NK}{L}(2 \log K + 1) + \frac{N}{L}(\bar{L}_\gamma + \bar{L}_g)\right)$	$\max_{i=1,2,\dots,N} \{L_{\gamma_i}\} + N\bar{L}_g$

Note: The input halfband, output halfband and halfband Weyl filters have the complexity of the minimum-energy Weyl filter plus additional complexity caused by the halfband restriction of the input or output signal. The complexity of the STFT filter and Gabor filter is that of the multiwindow STFT filter and multiwindow Gabor filter, respectively, with $N = 1$. Also, $\bar{L}_\gamma = \frac{1}{N} \sum_{i=1}^N L_{\gamma_i}$ and $\bar{L}_g = \frac{1}{N} \sum_{i=1}^N L_{g_i}$.

parameters (e.g., TF lattice parameters and analysis and synthesis windows) need to be chosen. The design methods we presented require prior knowledge or an educated guess of the effective support region of the FT of the TF weight function.

- The (multiwindow) STFT filter tends to be more complex than the other filters (in particular, Zadeh filter and Weyl filter). The (multiwindow) Gabor filter has minimum complexity as long as the oversampling factor is not too high. Table 6.1 provides a comparison of the computational cost and memory requirements of the various on-line TF filters.

Clearly, we can conclude from this summary that no single TF filter design is “best.” The choice of a TF filter depends on the specific application. Some questions to be asked in this context are: Which computational complexity can be tolerated? Is high TF selectivity required? Does the TF weight function contain obliquely oriented components? Is the TF weight function reasonably smooth or does it contain oscillatory components whose detrimental effects should be suppressed? Is an on-line estimation of signal statistics required?

A linear TF filter that has not been considered in this chapter is the *TF projection filter* [8, 13, 14]. The TF weight characteristic of this filter is restricted to passing and suppressing signal components; no other weights are possible. On the other hand, the TF projection filter has extremely high TF selectivity in that it can realize a very sharp

cutoff between pass and stop regions. An on-line implementation and applications of the TF projection filter are reported in [84].

Acknowledgment

This work was supported by FWF grant P11904-TEC.

References

- [1] L.E. Franks, *Signal Theory*, Prentice Hall, Englewood Cliffs, NJ, 1969.
- [2] T. Kailath, *Linear Systems*, Prentice Hall, Englewood Cliffs, NJ, 1980.
- [3] L.A. Zadeh, Frequency analysis of variable networks, in Proceedings of the IRE, 76, 291–299, March 1950.
- [4] T.A.C.M. Claassen and W.F.G. Mecklenbräuker, On stationary linear time-varying systems, *IEEE Trans. Circuits Syst.*, 29, 169–184, 1982.
- [5] A.W. Naylor and G.R. Sell, *Linear Operator Theory in Engineering and Science*, Springer, New York, 2nd ed., 1982.
- [6] I.C. Gohberg and M.G. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, American Mathematical Society, Providence, RI, 1969.
- [7] P. Dewilde and A.-J. van der Veen, *Time-Varying Systems and Computations*, Kluwer, Boston, 1998.
- [8] W. Kozek and F. Hlawatsch, A Comparative Study of Linear and Nonlinear Time-Frequency Filters, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, BC, October 1992, pp. 163–166.
- [9] F. Hlawatsch and G. Matz, Linear time-frequency filters, in *Time-Frequency Signal Analysis and Processing*, B. Boashash, Ed., Prentice Hall, Englewood Cliffs, NJ, 2002.
- [10] W. Kozek, Time-frequency signal processing based on the Wigner-Weyl framework, *Signal Process.*, 29, 77–92, October 1992.

- [11] F. Hlawatsch and G. Matz, Quadratic time-frequency analysis of linear time-varying systems, in *Wavelet Transforms and Time-Frequency Signal Analysis*, L. Debnath, Ed., Birkhäuser, Boston, MA, 2001, chap. 9.
- [12] F. Hlawatsch, G. Matz, H. Kirchauer, and W. Kozek, Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation, *IEEE Trans. Signal Process.* 48(5), 1417–1432, May 2000.
- [13] F. Hlawatsch and W. Kozek, Time-frequency projection filters and time-frequency signal expansions, *IEEE Trans. Signal Process.*, 42, 3321–3334, December 1994.
- [14] F. Hlawatsch, *Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation*, Kluwer, Boston, 1998.
- [15] S.H. Nawab and T.F. Quatieri, Short-time Fourier transform, in *Advanced Topics in Signal Processing*, J.S. Lim and A.V. Oppenheim, Eds., Prentice Hall, Englewood Cliffs, NJ, 1988, pp. 289–337.
- [16] M.R. Portnoff, Time-frequency representation of digital signals and systems based on short-time Fourier analysis, *IEEE Trans. Acoust. Speech Signal Process.*, 28, 55–69, February 1980.
- [17] R.E. Crochiere and L.R. Rabiner, *Multirate Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1983.
- [18] I. Daubechies, Time-frequency localization operators: a geometric phase space approach, *IEEE Trans. Inf. Theory*, 34, 605–612, July 1988.
- [19] R. Bourdier, J.F. Allard, and K. Trumpf, Effective frequency response and signal replica generation for filtering algorithms using multiplicative modifications of the STFT, *Signal Process.* 15, 193–201, September 1988.
- [20] M.L. Kramer and D.L. Jones, Improved Time-Frequency Filtering Using an STFT Analysis-Modification-Synthesis Method, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, PA, October 1994, pp. 264–267.
- [21] S. Farkash and S. Raz, Time variant filtering via the Gabor representation, in Proceedings of the EUSIPCO-90, L. Torres, E. Masgrau, and M.A. Lagunas, Eds., Elsevier, Amsterdam, 1990, pp. 509–512.
- [22] S. Farkash and S. Raz, Linear systems in Gabor time-frequency space, *IEEE Trans. Signal Process.*, 42(3), 611–617, March 1994.
- [23] H.G. Feichtinger and K. Nowak, A first survey on Gabor multipliers, in *Advances in Gabor Analysis*, H.G. Feichtinger and T. Strohmer, Eds., Birkhäuser, Boston, MA, 2002.

- [24] X.-G. Xia and S. Qian, Convergence of an iterative time-variant filtering based on discrete Gabor transform, *IEEE Trans. Signal Process.*, 47(10), 2894–2899, October 1999.
- [25] S. Qian and D. Chen, *Joint Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1996.
- [26] G.F. Boudreaux-Bartels, Time-varying signal processing using Wigner distribution synthesis techniques, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997, pp. 269–317.
- [27] F. Hlawatsch and W. Krattenthaler, Signal synthesis algorithms for bilinear time-frequency signal representations, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997, pp. 135–209.
- [28] P.A. Bello, Characterization of randomly time-variant linear channels, *IEEE Trans. Commun. Syst.*, 11, 360–393, 1963.
- [29] A. Gersho, Characterization of time-varying linear systems, in Proceedings of the IEEE, 51, 238, January 1963.
- [30] W. Kozek, On the Generalized Weyl Correspondence and Its Application to Time-Frequency Analysis of Linear Time-Varying Systems, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Victoria, Canada, October 1992, pp. 167–170.
- [31] G. Matz and F. Hlawatsch, 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*, 39, 4041–4071, August 1998.
- [32] G. Matz and F. Hlawatsch, Time-frequency transfer function calculus of linear time-varying systems, in *Time-Frequency Signal Analysis and Processing*, B. Boashash, Ed., Prentice Hall, Englewood Cliffs, NJ, 2002.
- [33] R.G. Shenoy and T.W. Parks, The Weyl correspondence and time-frequency analysis, *IEEE Trans. Signal Process.*, 42(2), 318–331, February 1994.
- [34] W. Kozek, On the transfer function calculus for underspread LTV channels, *IEEE Trans. Signal Process.*, 45(1), 219–223, January 1997.
- [35] W. Kozek, Matched Weyl-Heisenberg Expansions of Nonstationary Environments, Ph.D. thesis, Vienna University of Technology, March 1997.
- [36] G. Matz, A Time-Frequency Calculus for Time-Varying Systems and Non-stationary Processes with Applications, Ph.D. thesis, Vienna University of Technology, November 2000.
- [37] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, 2001.

- [38] N.-C. Huang and J.K. Aggarwal, On linear shift-variant digital filters, *IEEE Trans. Circuits Syst.*, 27, 672–679, August 1980.
- [39] N.-C. Huang and J.K. Aggarwal, Synthesis and implementation of recursive linear shift-variant digital filters, *IEEE Trans. Circuits Syst.*, 30, 29–36, January 1983.
- [40] G.B. Folland, *Harmonic Analysis in Phase Space*, Vol. 122, *Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ, 1989.
- [41] J. Ramanathan and P. Topiwala, Time-frequency localization via the Weyl correspondence, *SIAM J. Matrix Analysis Appl.*, 24(5), 1378–1393, 1993.
- [42] A.J.E.M. Janssen, Wigner weight functions and Weyl symbols of non-negative definite linear operators, *Philips J. Res.*, 44, 7–42, 1989.
- [43] M.W. Wong, *Weyl Transforms*, Springer, New York, 1998.
- [44] A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1975.
- [45] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, San Diego, CA, 1999.
- [46] F. Hlawatsch and P. Flandrin, The interference structure of the Wigner distribution and related time-frequency signal representations, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997, pp. 59–133.
- [47] F. Hlawatsch and G.F. Boudreux-Bartels, Linear and quadratic time-frequency signal representations, *IEEE Signal Process. Mag.*, 9, 21–67, April 1992.
- [48] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed., McGraw-Hill, New York, 1991.
- [49] W. Martin and P. Flandrin, Wigner-Ville spectral analysis of nonstationary processes, *IEEE Trans. Acoust. Speech Signal Process.*, 33, 1461–1470, December 1985.
- [50] P. Flandrin and W. Martin, The Wigner-Ville spectrum of nonstationary random signals, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997, pp. 211–267.
- [51] M. Amin, Time-frequency spectrum analysis and estimation for non-stationary random processes, in *Advances in Spectrum Estimation*, B. Boashash, Ed., Longman Cheshire, Melbourne, 1992, pp. 208–232.
- [52] G. Matz and F. Hlawatsch, Time-Varying Spectra for Underspread and Over-spread Nonstationary Processes, in Proceedings of the 32nd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 1998, pp. 282–286.

- [53] G. Matz and F. Hlawatsch, Time-varying power spectra of nonstationary random processes, in *Time-Frequency Signal Analysis and Processing*, B. Boashash, Ed., Prentice Hall, Englewood Cliffs, NJ, 2002.
- [54] G. Matz, F. Hlawatsch, and W. Kozek, Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes, *IEEE Trans. Signal Process.*, 45, 1520–1534, June 1997.
- [55] A.B. Carlson, *Communication Systems*, 3rd ed., McGraw-Hill, New York, 1986.
- [56] <http://www.mathworks.com/access/helpdesk/help/toolbox/comm/comm.shtml>.
- [57] J.B. Allen and L.R. Rabiner, A unified approach to short-time Fourier analysis and synthesis, in *Proceedings of the IEEE*, 65(11), 1977, 1558–1564.
- [58] A.W. Rihaczek, Signal energy distribution in time and frequency, *IEEE Trans. Inf. Theory*, 14, 369–374, May 1968.
- [59] T.A.C.M. Claassen and W.F.G. Mecklenbräuker, The Wigner distribution — A tool for time-frequency signal analysis; Part III. Relations with other time-frequency signal transformations, *Philips J. Res.*, 35(6), 372–389, 1980.
- [60] A.J.E.M. Janssen, Positivity and spread of bilinear time-frequency distributions, in *The Wigner Distribution — Theory and Applications in Signal Processing*, W. Mecklenbräuker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997, pp. 1–58.
- [61] P. Flandrin, Ambiguity function, in *Time-Frequency Signal Analysis and Processing*, B. Boashash, Ed., Prentice Hall, Englewood Cliffs, NJ, 2002.
- [62] P.M. Woodward, *Probability and Information Theory with Application to Radar*, Pergamon Press, London, 1953.
- [63] C.H. Wilcox, The synthesis problem for radar ambiguity functions, in *Radar and Sonar, Part I*, R.E. Blahut, W. Miller, Jr., and C.H. Wilcox, Eds., Springer, New York, 1991, pp. 229–260.
- [64] W. Kozek, H.G. Feichtinger, and J. Scharinger, Matched Multiwindow Methods for the Estimation and Filtering of Nonstationary Processes, in *Proceedings of the IEEE ISCAS-96*, Atlanta, GA, May 1996, pp. 509–512.
- [65] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
- [66] C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems*, Prentice Hall, Englewood Cliffs, NJ, 1974.
- [67] H.G. Feichtinger and T. Strohmer, Eds., *Gabor Analysis and Algorithms: Theory and Applications*, Birkhäuser, Boston, 1998.
- [68] D. Gabor, Theory of communication, *J.IEE*, 93(3), 429–457, 1946.

- [69] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, NJ, 1993.
- [70] H. Bölskei and F. Hlawatsch, Oversampled modulated filter banks, in *Gabor Analysis and Algorithms: Theory and Applications*, H.G. Feichtinger and T. Strohmer, Eds., Birkhäuser, Boston, 1998, pp. 295–322.
- [71] H. Bölskei, F. Hlawatsch, and H.G. Feichtinger, Oversampled FIR and IIR DFT filter banks and Weyl-Heisenberg frames, in Proceedings of the IEEE ICASSP-96, 3, 1391–1394, May 1996.
- [72] R.J. Duffin and A.C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Am. Math. Soc.*, 72, 341–366, 1952.
- [73] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.
- [74] R.M. Young, *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1980.
- [75] W. Kozek, H.G. Feichtinger, and T. Strohmer, Time-Frequency Synthesis of Statistically Matched Weyl-Heisenberg Prototype Signals, in Proceedings of the IEEE International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, October 1994, pp. 417–420.
- [76] Y.Y. Zeevi, M. Zibulski, and M. Porat, Multi-window Gabor schemes in signal and image representations, in *Gabor Analysis and Algorithms: Theory and Applications*, H.G. Feichtinger and T. Strohmer, Eds., Birkhäuser, Boston, 1998, pp. 381–407 chap. 12.
- [77] M. Zibulski and Y.Y. Zeevi, Discrete multiwindow Gabor-type transforms, *IEEE Trans. Signal Process.*, 45, 1428–1442, June 1997.
- [78] M. Zibulski and Y.Y. Zeevi, Analysis of multiwindow Gabor-type schemes by frame methods, *Appl. Computational Harmonic Analysis*, 4, 188–221, April 1997.
- [79] <http://www.mathworks.com/access/helpdesk/help/toolbox/signal/signal.shtml>.
- [80] G. Matz and F. Hlawatsch, Minimax robust nonstationary signal estimation based on a p -point uncertainty model, *J. Franklin Inst.*, 337, 403–419, July 2000.
- [81] G. Matz, F. Hlawatsch, and A. Raidl, Signal-Adaptive Robust Time-Varying Wiener Filters: Best Subspace Selection and Statistical Analysis, in Proceedings of the IEEE ICASSP-2001, Salt Lake City, UT, May 2001, pp. 3945–3948.
- [82] W. Kozek and K. Riedel, Quadratic Time-Varying Spectral Estimation for Underspread Processes, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, October 1994, pp. 460–463.

- [83] M. Bayram and R. Baraniuk, Multi-Window Spectrum Estimates for Nonstationary Random Processes, in Proceedings of the IEEE ICASSP-97, Munich, Germany, May 1997, pp. 173–176.
- [84] G. Matz and F. Hlawatsch, Time-frequency projection filters: on-line implementation, subspace tracking, and application to interference suppression, in Proceedings of the IEEE ICASSP-2002, Orlando, FL, May 2002, pp. 1213–1216.

Discrete Reduced Interference Distributions

William J. Williams
University of Michigan

CONTENTS

7.1	Introduction	273
7.2	Reduced Interference Distributions	274
7.3	Discrete Time–Frequency Distributions	277
7.4	Applications	289
7.5	Conclusions	300
	Acknowledgments	300
	References	300
	Appendix A — Generalized Discrete Time–Frequency Distribution Scripts	304
	Appendix B — Kernel Scripts	305

7.1 Introduction

Digital signal processing (DSP) methodologies have fostered great advances in many areas. The advent of digital computing resources in the research laboratory setting together with the popularization of the Cooley–Tukey fast Fourier transform (FFT) has been credited with the rapid growth of signal processing, particularly DSP. As with any body of theory and application, many simplifying assumptions in DSP promote the teaching of the subject and ease the application. Many of the most powerful and effective algorithms in signal processing start with the assumption of stationarity. In addition, the deterministic portion of the signal is often assumed to be composed of complex exponentials that are the solutions to linear time-invariant (LTI) differential equations. These assumptions are often valid enough to yield good results when the signals and systems involved result from engineering design that often assures compliance with these conceptualizations. Signals of biological and physical origin often do not comply with these assumptions, however, resulting in disappointment when conventional techniques are used. Newly emerging techniques of time-frequency (TF) analysis can provide new insights into the nature of many signals. The Wigner distribution and the spectrogram are well-known time-frequency

distributions (TFDs). Since 1990, a number of new TFDs have been introduced, including the reduced interference distributions (RIDs). This volume covers a number of these advances. This chapter describes some results using discrete RIDs. First, some of the theory in terms of continuous mathematics is briefly presented. Next, discrete methods for realizing RIDs are introduced. Care is taken to fully explain the discrete theory in sufficient detail to understand the computational algorithms used in demonstrating some informative results. MATLAB® scripts are provided at the end of this chapter to enable the readers to try the methods for themselves. Finally, actual research results using these techniques are presented along with interpretations of the new and meaningful insights provided by discrete RID analysis.

7.2 Reduced Interference Distributions

A brief treatment of ambiguity function concepts is helpful in understanding these concepts. Greater detail is available in several books on TF analysis [1, 2]. Let $F(\omega)$ be the FT of the signal $f(t)$:

$$F(\omega) = F[f(t)] = \int_{-\infty}^{\infty} f(t)e^{-j\omega t} dt \quad (7.1)$$

and:

$$f(t) = F^{-1}[F(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{j\omega t} d\omega \quad (7.2)$$

Let $R_f(t, \tau)$ be the instantaneous autocorrelation of a complex signal $f(t)$, defined as:

$$R_f(t, \tau) = f(t + \tau/2)f^*(t - \tau/2) \quad (7.3)$$

where f^* denotes the complex conjugate of f . The Wigner distribution (WD) of $f(t)$ is defined as the Fourier transform (FT) of $R_f(t, \tau)$ with respect to the lag variable τ :

$$W_f(t, \omega) = F_\tau[f(t + \tau/2)f^*(t - \tau/2)] = F_\tau[R_f(t, \tau)] \quad (7.4)$$

The symmetrical ambiguity function (AF) is defined as the inverse Fourier transform (IFT) of $R_f(t, \tau)$ with respect to the first variable:

$$A_f(\theta, \tau) = F_t^{-1}[f(t + \tau/2)f^*(t - \tau/2)] = F_t^{-1}[R_f(t, \tau)] \quad (7.5)$$

Thus, $W_f(t, \omega)$ and $A_f(\theta, \tau)$ are related by the two-dimensional (2-D) FT:

$$W_f(t, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_f(\theta, \tau) e^{-(t\theta + \omega\tau)} d\theta d\tau \quad (7.6)$$

These relationships may be combined with Equation (7.4) to show that $C_f(t, \omega, \phi)$, Cohen's equation, can be found by:

$$C_f(t, \omega, \phi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\theta, \tau) A_f(\theta, \tau) e^{-j(t\theta + \omega\tau)} d\theta d\tau \quad (7.7)$$

Although the WD may be found from the symmetrical ambiguity function by means of a double FT, any member of Cohen's class of distributions may be found by first using the kernel $\phi(\theta, \tau)$ to alter the symmetrical ambiguity function and then carrying out the double FT. The generalized ambiguity function, $\phi(\theta, \tau)A_f(\theta, \tau)$ [3] greatly aids one in clearly seeing the effect of the kernel in determining $C_f(t, \omega, \phi)$.

The WD has a unity valued kernel such that the generalized ambiguity function is identical to the ambiguity function and its TF representation preserves both the autoterms and the cross terms. The kernels of the spectrogram and the RID emphasize the autoterms and suppress the cross terms, but in very different ways.

7.2.1 Exponential distribution

The exponential distribution (ED) developed by Choi and Williams is an attempt [4] to improve on the WD. It has a kernel $\phi(\theta, \tau) = \exp(-\frac{\theta^2\tau^2}{\sigma^2})$, and it proves to be quite effective in suppressing the interferences while retaining high resolution. Its performance has been compared with those of the spectrogram and the WD in a variety of environments. The σ parameter may be varied over a range of values to obtain different trade-offs between cross-term suppression and high autoterm TF resolution. In fact, as σ becomes very large the ED kernel approaches the WD kernel. This provides the best resolution but the cross terms become large and approach WD cross terms in size.

7.2.2 Reduced interference distribution kernels

A brief description of RIDs and RID kernels is appropriate at this point. A comprehensive review of the RID is available elsewhere [5, 6]. RID requirements and properties are discussed in comparison with the WD. The unity value of the WD kernel guarantees the desirable properties of the WD. However, it is not necessary to require the kernel to be unity for all θ and τ to maintain most of its desirable properties. It is sufficient to ensure that the kernel is unity along $\theta = 0$ and $\tau = 0$ and that the kernel is such that $\phi^*(\theta, \tau) = \phi(-\theta, -\tau)$, the latter property ensuring realness. The RID kernel is cross shaped and acts as a low-pass filter in both θ and τ . This ensures that the time and frequency marginals are preserved and the instantaneous frequency and group delay properties are retained. In addition, it is often quite desirable for a distribution to possess the time and frequency support property. This property ensures that the distribution does not extend beyond the support of the signal in time or the support of its FT in frequency. The time and frequency support property may be maintained for the RID by ensuring that:

$$\psi(t, \tau) = \int_{-\infty}^{\infty} \phi(\theta, \tau) e^{-j\theta t} d\theta = 0$$

if

$$|\tau| < 2|t| \quad (7.8)$$

This forms a cone-shaped region in t, τ . The WD obviously satisfies this support property because the FT of unity is an impulse function, clearly staying within the t, τ limits. The form of the kernel in θ, ω is also cone shaped, ensuring the frequency support property of Equation (7.8). The ED can be brought into the RID requirements by imposing an RID window as suggested earlier. The RID is not a truly new distribution because the Born–Jordan kernel [7], $\phi(\theta, \tau) = \text{sinc}(\theta\tau)$ meets all the RID requirements. The RID possesses almost all the desirable properties of the WD except for its unitary property, $|\phi(\theta, \tau)|=1$ for all θ, τ . Zhao, Atlas and Marks [8] suggest a cone-shaped kernel for nonstationary signal analysis, but they impose different restrictions not in common with RID properties.

RIDs may be designed using a very simple procedure.

Step 1. Design a primitive function $h(t)$ that is real and satisfies the following:

R1: $h(t)$ has a unit area, that is, $\int_{-\infty}^{\infty} h(t)dt = 1$.

R2: $h(t)$ is a symmetrical with time, that is, $h(-t) = h(t)$.

R3: $h(t)$ has support on $[-1/2, 1/2]$, that is, $h(t) = 0$ for $|t| > 1/2$.

R4: $h(t)$ tapers smoothly toward both ends so that its frequency response has little high-frequency content. That is, $|H(\theta)| \ll 1$ for $|\theta| \gg 0$, where $H(\theta)$ is the FT of $h(t)$.

Step 2. Take the FT of $h(t)$, that is:

$$H(\theta) = \int_{-\infty}^{\infty} h(t)e^{-j\theta t} dt$$

Step 3. Replace θ by $\theta\tau$ in $H(\theta)$.

The primitive function, $h(t)$ may be considered to be a window or impulse response of a filter. It is interesting to note that $\delta(t)$ is the RID primitive for the WD kernel. The FT of $\delta(t)$ is 1, but this represents an extreme RID with no interference-reducing characteristics. Good RID kernels attenuate rapidly with distance away from the τ and θ axes. This is demonstrated graphically in the next section.

The RID has the following integral expression that is important to discrete considerations later in this chapter:

$$RID_f(t, \omega; h) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{|\tau|} h\left(\frac{u-t}{\tau}\right) f(u + \tau/2) f^*(u - \tau/2) e^{-j\tau\omega} du d\tau \quad (7.9)$$

For computation, the generalized autocorrelation function is:

$$R'_f(t, \tau; h) = \int_{-\infty}^{\infty} \frac{1}{|\tau|} h\left(\frac{u-t}{\tau}\right) f(u + \tau/2) f^*(u - \tau/2) du \quad (7.10)$$

and:

$$RID_f(t, \omega; h) = \int_{-\infty}^{\infty} R'_f(t, \tau; h) e^{-j\tau\omega} d\tau \quad (7.11)$$

is carried out. The RID exhibits an interesting scale property. This means that the TF representation of scaled signals do not change in overall shape with RID representation. They inversely expand and shrink in the time and frequency directions just as the FT dictates.

7.2.3 Types of time–frequency distributions

Just as in one-dimensional (1-D) signal processing, the time and frequency variables may be continuous or discrete. However, for TFDs, four possibilities exist for the TF variable pairs and consequently, four types of TFDs [9, 10]. Table 7.1 delineates the four types of TFDs, the characteristics of the time and frequency variables and the type of transform associated with each type of TFD.

TABLE 7.1
Four Types of TFDs

Type	Time Domain	Frequency Domain	Transform
I	t, aperiodic	ω , aperiodic	Fourier
II	n, aperiodic	ω , periodic	DTFT
III	t, periodic	k, aperiodic	Fourier series
IV	n, periodic	k, periodic	DFT

Type I TFDs have held center stage in much of the literature and RIDs were developed under those considerations. However, type IV TFDs are the main focus for the balance of this chapter, because both discrete time and discrete frequency are required for practical, computer-based computation. Thus, some time needs to be taken with type IV formulations to understand some of the benefits and pitfalls of the fully discrete forms.

7.3 Discrete Time–Frequency Distributions

FDs distributions are presented in a continuous form for theoretical development and discussion of properties. However, one usually wishes to utilize a discrete form of the distribution for computational convenience using a digital computer. Claasen and

Mecklenbräuker [11] developed discrete forms of the WD:

$$W_x(n, \omega) = 2 \sum_{k=-\infty}^{\infty} e^{-j2\omega k} x(n+k) \cdot x^*(n-k) \quad (7.12)$$

If discrete values of the local autocorrelation, $R_x(n, m)$, were available for all integer n and m , then one could express the discrete form of the WD as the discrete-time Fourier transform (DTFT) of $R_x(n, m)$, or:

$$W_x(n, \omega) = \sum_{m=-\infty}^{\infty} e^{-j\omega m} R_x(n, m) \quad (7.13)$$

Notice that in (7.12), as k takes on the values $0, 1, 2, \dots$ the discrete local autocorrelations $R_x(n, 0) = x(n)x^*(n)$, $R_x(n, 2) = x(n+1)x^*(n-1)$, $R_x(n, 4) = x(n+2)x^*(n-2)$ are evaluated. Discrete values two samples apart are thus correlated. Local autocorrelation values for odd integer spacings are not available; thus, the discrete form of the WD is formed from the even integer correlation lag values. This means that the local autocorrelation is undersampled by a factor of two compared with the sequence $x(n)$ and aliasing may occur, even for adequately sampled signals. The analytical form of the signal presents no problem with aliasing, however, because only half the period of the DTFT spectrum is occupied. If aliasing is a problem with the discrete WD and not with the original sequence $x(n)$, then additional points may be interpolated for $x(n)$ to fill in the missing correlation values required. This is not a trivial matter and requires considerable care in understanding and utilizing discrete TFDs.

7.3.1 Generalized discrete time–frequency distributions

Shift-invariant bilinear discrete TFDs are specified by a discrete kernel, and can be rewritten in the rotated form of:

$$\begin{aligned} TFR_x(n, \omega; \psi) = & \sum_{n_1} \sum_{n_2} [x(n+n_1)e^{-j\omega(n+n_1)}] \psi(-\frac{n_1+n_2}{2}, n_1 - n_2) \\ & \times [x(n+n_2)e^{-j\omega(n+n_2)}]^* \end{aligned} \quad (7.14)$$

The discrete form of the WD kernel in the (n, m) domain is, with \check{Z} the even integers and \hat{Z} the odd integers:

$$\begin{aligned} \Psi(n, m) &= \delta(n), m \in \check{Z} \\ \Psi(n, m) &= 0, m \in \hat{Z} \end{aligned} \quad (7.15)$$

This results in a kernel in the (θ, m) domain by means of a DTFT along the n direction:

$$\begin{aligned} \Phi(\theta, m) &= \sum_{n \in \check{Z}} \Psi(n, m) e^{-j2\pi\theta n} = 1, m \in \check{Z} \\ \Phi(\theta, m) &= 0, m \in \hat{Z} \end{aligned} \quad (7.16)$$

This result does not provide much insight. It is certainly not unity for all m , but only for even m . This means that the ambiguity function is downsampled by a factor of 2. Instead of periodic in 2π as would be the case for full sampling, the result must be considered to be periodic in π in the θ direction. In the rotated form, for Equation (7.14), one would have an infinite sequence of 1's along a 45 degree line and 0's everywhere else. For a finite $N \times N$ matrix, one would have a matrix with all 0's, except along the antidiagonal as expressed for an 9×9 example in Equation (7.17). This rotated form is very useful for realizing the discrete TFD. The discrete LAF can be formed by taking the outer product of two 9 point sequences $x(n)(u(n) - u(n-9))$ and $x^*(-n)(u(-n) - u(-n+9))$ for a 9×9 rotated LAF matrix. The 0 time, 0 correlation lag point is at the location of the middle 1 in Equation (7.17). The discrete time variable n ranges along the diagonal and the discrete lag variable m ranges along the antidiagonal. However, the lag values are $[-8, -6, -4, -2, 0, 2, 4, 6, 8]$.

The kernel resulting from this is:

$$K = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7.17)$$

The product of the kernel matrix, K , and the LAF matrix is:

$$G = K.*LAF \quad (7.18)$$

and $.*$ denotes element-by-element multiplication of the matrices as is the convention in MATLAB. Next, the diagonals of G are summed to form a nine-element vector, $X(n)$. Finally, a frequency slice at location n is produced by:

$$TFD(k,n) = FFT_9(IFFTSHIFT(X(n))) \quad (7.19)$$

where $FFT_9(\bullet)$ denotes a nine-point FFT and $IFFTSHIFT(\bullet)$ denotes the MATLAB action of ifftshift, which places the vector in proper form for using the FFT, by rotating the elements of the vector such that the zero time value is the first element. This set of steps is exactly that which is carried out in the script provided in Appendix A.

Unfortunately, this can produce aliasing by undersampling in lag, because the odd lag values are not available along the antidiagonal of $G(n)$. This problem has

prompted a lot of attention and a variety of solutions have been proposed for discrete versions of Cohen's class [11–16]. It is claimed that the only absolutely foolproof method for dealing with this is to sample the signal at a rate twice that normally used to ensure no aliasing (four times the Nyquist frequency) or interpolate the sampled sequence to fill in the missing values. However, one simple alternative that works quite well in many cases has been proposed by Jeong and Williams. An example is provided in terms of the discrete TFD [15] and later dubbed the “quasi-Wigner” TFD [9]. Here, the missing values are filled in by using the odd lag values available to either side of the main antidiagonal and averaging them. The quasi-WD kernel is given in Equation (7.20):

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7.20)$$

The equivalent of Equation (7.14) is:

$$\Psi(n, m) = \delta(n), m \in \check{Z}, = \frac{\delta(n - \frac{1}{2}) + \delta(n + \frac{1}{2})}{2}, m \in \hat{Z} \quad (7.21)$$

This results in:

$$\begin{aligned} \Phi(\theta, m) &= 1, m \in \check{Z} \\ \Phi(n, m) &= \cos(\theta), m \in \hat{Z} \end{aligned} \quad (7.22)$$

The kernel is periodic with 2π in this case. This not a pretty picture, but the aliasing situation has been improved. The kernel is near unity close to $\Phi(0, 0)$, thus, ambiguity functions that have their main components in that region can be treated in a manner similar to the continuous Wigner kernel.

The binomial kernel [17], defined in the (m, n) domain is:

$$\begin{aligned} \psi(n, 0) &= \delta(n) \\ \psi(n, \pm 1) &= \frac{1}{2}\delta(n + \frac{1}{2}) + \frac{1}{2}\delta(n - \frac{1}{2}) \\ \psi(n, \pm 2) &= \frac{1}{4}\delta(n + 1) + \frac{1}{2}\delta(n) + \frac{1}{4}\delta(n - 1) \\ \psi(n, |m|) &= \psi(n, |m| - 1) * \psi(n, 1) \end{aligned} \quad (7.23)$$

where $*$ denotes discrete convolution. This process generates the well-known binomial coefficients. The correlation shift index m is assumed to take on the values $-\infty, \dots, -1, 0, 1, \dots, \infty$ and the time shift index, n is assumed to take on the values $-\infty, \dots, -1, -0.5, 0, 0.5, 1, \dots, \infty$.

The results for the binomial kernel and the quasi-Wigner kernel are the same for $m = 0$ and $m = \pm 1$. This provides an easy starting point for finding the ambiguity domain form of the binomial kernel: $\psi(n, 2)$ is the convolution of $\psi(n, 1)$ with itself and $\psi(n, 3)$ is the convolution of $\psi(n, 1)$ with $\psi(n, 2)$. In the ambiguity domain, convolution becomes multiplication. Thus, generally the result for lag m in the ambiguity domain is known and:

$$\Phi(\theta, m)_{\text{binomial}} = \cos^{|m|}(\theta) \quad (7.24)$$

Unlike the Wigner and quasi-Wigner result, nice convergence is realized for the binomial kernel. The ambiguity domain kernel is shown later in [Figure 7.3\(d\)](#).

This kernel has the characteristics of the continuous RID kernel. It is unity along the θ and m axes, allowing it to approximate the RID properties, and it falls off rapidly away from the axes, providing the interference suppression. This kernel is periodic in 2π , which means that the period halving and subsequent aliasing caused by the discrete Wigner kernel has been avoided. The kernel takes care of the odd lag values while at the same time providing the desirable binomial kernel properties.

The rotated and truncated matrix form of this kernel in the (n, m) domain is given in Equation (7.25). To enforce RID properties, the sum of values along m for $\Psi(n, m)$ must be one. It can be seen that this is accomplished for the binomial kernel:

$$BN = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{4}{16} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{3}{8} & \frac{6}{16} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{3}{8} & \frac{4}{16} & 0 \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ 0 & \frac{4}{16} & \frac{3}{8} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{6}{16} & \frac{3}{8} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{4}{16} & \frac{1}{8} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7.25)$$

Another RID kernel has an easily realized equivalent. This is the Born–Jordan kernel suggested by Cohen in his classical paper that established Cohen’s class [18].

The discrete Born–Jordan kernel is:

$$\begin{aligned}\Psi(n, m)_{\text{Born–Jordan}} &= \frac{1}{|m|+1} \sum_{k=-\frac{m}{2}}^{\frac{m}{2}} \delta(n-k), \quad m, n \in \check{Z} \\ \Psi\left(\frac{n}{2}, m\right)_{\text{Born–Jordan}} &= \frac{1}{|m|+1} \sum_{k=-\frac{m}{2}}^{\frac{m}{2}} \delta(n-k), \quad m, n \in \check{Z}\end{aligned}\quad (7.26)$$

which is consistent with the continuous form.

The ambiguity form of this can easily be found, because the summations in Equation (7.26) represent a rectangular function whose DTFT is a periodic sinc function or Dirichlet function. It is well known that the DTFT of this sequence is $\frac{\sin((m+1)\omega/2)}{\sin(\omega/2)}$. The result is:

$$\Phi(\omega, m) = \frac{1}{|m|+1} \frac{\sin((|m|+1)\omega/2)}{\sin(\omega/2)} \quad (7.27)$$

An example of the rotated matrix for the Born–Jordan kernel given in Equation (7.28) is:

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & 0 \\ \frac{1}{5} & \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ 0 & \frac{1}{5} & \frac{1}{4} & \frac{1}{3} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & \frac{1}{4} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7.28)$$

A modified version of the binomial kernel provides additional flexibility. The generalized binomial kernel is defined to be:

$$\begin{aligned}\psi(n, 0) &= \delta(n) \\ \psi(n, \pm 1) &= \frac{1}{2}\delta(n + \frac{1}{2}) + \frac{1}{2}\delta(n - \frac{1}{2}) \\ \psi(n, \pm 2) &= b\delta(n+1) + a\delta(n) + b\delta(n-1) \\ \psi(n, |m|) &= \psi(n, |m|-2) * \psi(n, \pm 2), \quad m > 2\end{aligned}\quad (7.29)$$

where the constraint $a + 2b = 1$ or $b = \frac{a-1}{2}$ is imposed. The resulting TFD retains the discrete RID characteristics, but has the advantage of more flexibility in that adjusting the value of a provides a smooth transition from the binomial TFD to the quasi-Wigner TFD as a ranges from $\frac{1}{2}$ to 1. TFDs outside of this range can also be achieved by choosing $a < \frac{1}{2}$.

7.3.1.1 Alias-free characteristics

These issues are complicated and viewed differently by different workers in the field; thus, a comprehensive discussion does not serve the purposes of this chapter. However, a few comments are appropriate for practical usage. The quasi-Wigner kernel provides little smoothing and the resulting TFD is thus more subject to aliasing than the TFD resulting from the binomial kernel. Appropriate cautions are mentioned [15] in terms of limiting the frequency range of the signals analyzed to less than half the sampling frequency.

Morris and Wu provide additional insights for discrete TFDs and define two types of alias-free discrete members of Cohen's class [19]. They call these *type I alias free* and *type II alias free*. Their type I alias free class is essentially the ambiguity function, group delay, time–frequency distribution (AF-GDTFD) defined by Jeong and Williams, but with a slightly different formulation. Their type II alias-free class specifically requires increasing the sampling rate to reduce or eliminate aliasing. These results apply to discrete-time continuous frequency formulations. The use of both discrete-time and discrete-frequency fall under the type IV all-discrete class (see [Table 7.1](#)), which is the usual computational situation.

O'Neill and Williams develop these concepts further [9] and demonstrate that the simple periodic repetition of spectra along the frequency axis that results when continuous signals are sampled (type II in [Table 7.1](#)) and potentially result in aliasing when spectra overlap must be carefully reconsidered in the case of discrete TFDs. Discrete TFDs exist on a complicated periodic surface. An added complication of these periodic surfaces is that cross terms can occur between periodic repetitions of the discrete TFDs. Others [20] interpret these cross terms as aliased terms and provide a number of examples to support this viewpoint. However, for purposes of this chapter, the idea that the phenomena can more generally be interpreted in terms of periodic repetitions and possible cross terms between these repetitions is retained. Stanković and Djurović [21] have provided additional commentaries on this issue. The interested reader is referred to the referenced literature for further elaboration of these ideas.

A further elaboration of the concept of the type IV class of TFDs in terms of the WD has been provided [10] and it is claimed that there is a unique extension of the continuous WD to the type IV class of TFDs. In addition to other considerations, this requires an odd length of samples to be processed for each frequency slice and resulting LAF. This idea has been followed in this chapter as well, for all discrete TFDs provided.

The continuous RIDs enjoy ten valuable characteristics [6, 22]. The type II TFDs that approximate RIDs also enjoy similar characteristics [15, 19]. Type IV RID-like TFDs only approximate many of these characteristics. However, with oversampling of the signal and long analysis sequences, true RID characteristics are approached.

7.3.1.2 Analytical signal

The analytical signal features prominently in bilinear TF analysis. Boashash [23] distinguishes the WD from the Wigner–Ville distribution (WVD) on the basis of the

analytical signal. If one uses a real $x(t)$, the WD is obtained. However, if one converts the real signal into an analytical signal by:

$$z(t) = x(t) + j\mathcal{H}[x(t)] \quad (7.30)$$

where \mathcal{H} is the Hilbert transform, then the WVD is obtained by using $z(t)$. Generally, this is a good idea. The negative frequency half of the spectrum is zero for the analytical signal. Thus, cross-term activity caused by that part of the spectrum is eliminated. If the analytical signal is used in forming the type II and type IV TFDs, both cross-term activity and aliasing due to negative portions of the spectrum are avoided. However, if the signal is near base band, distortion in the TFD may result and this approach is not recommended [5, 6, 24].

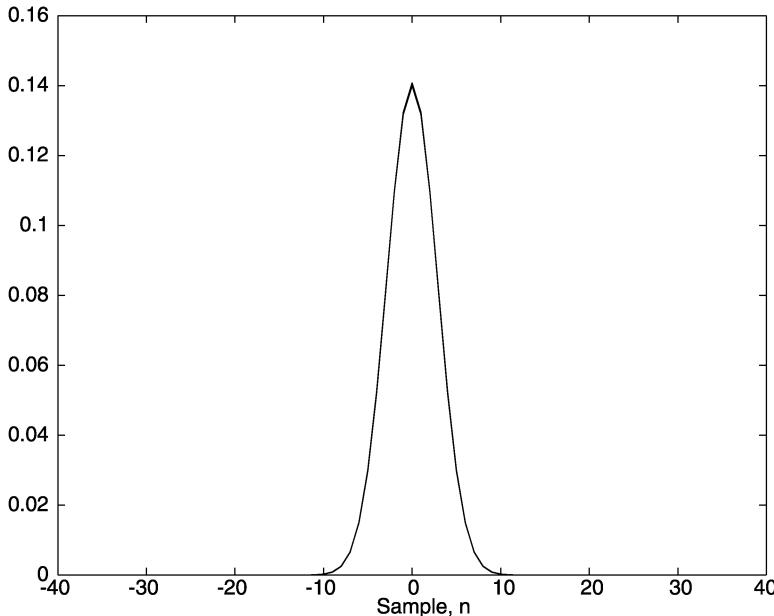
7.3.1.3 Scale properties of discrete kernels

The scale property of the continuous RIDs as expressed in Equations (7.9) and (7.10) is very desirable in preserving scale covariance. Properly designed discrete RIDs approximate this property well. The discrete Born–Jordan TFD can be seen to do this perfectly in the discrete case. Each finite impulse response (FIR) filter at each lag (m) value is a time-scaled and properly weighted version of the others. The binomial kernel is not as exact, but closely approximates the property. Figure 7.1 compares a binomial kernel slice at $m = 65$ downsampled by 2 and compared with the slice at $m = 33$, with weighting properly adjusted. One can see that the match is almost exact. This does not hold as well for small values of m , but quickly improves for larger m . Examples of the effects on discrete TFDs follow later in the chapter.

7.3.2 Fast algorithms for discrete reduced interference distributions

Several fast algorithms for discrete RIDs have been reported. One of the criticisms of TFDS, in general, and RIDs, in particular, involves the computational times required that are somewhat larger than for other types of representations. Barry was the first to address this problem with his algorithm for a fast version of the Choi and Williams ED [25]. This was not a fully discrete algorithm, but a discrete approximation of a continuous TFD. O’Hair and Suter [26] developed a Zak transform-based algorithm for the binomial TFD. Their algorithm is attractive in terms of analyzing a band of frequencies efficiently. Cunningham and Williams developed a spectrogram method that allows TFDs to be expanded in terms of weighted spectrograms, using windows derived from an eigendecomposition of the kernel [27]. These methods were further developed and extended by using wavelet-like windows. Methods along a similar path have also been reported by Le and Glesner [28]. Finally, optimum windows that allow a smaller number of windows to be used in forming the spectrograms, thus reducing the computational burden, have been reported [29, 30].

Amin and colleagues have made a number of contributions to the approximation of TFDs via kernel representations, including some of the key concepts [31]. The article closest to the discrete concepts in this chapter describes a methodology using

**FIGURE 7.1**

Comparison of binomial kernel slices at $m = 33$ and $m = 65$ showing almost perfect overlap.

a “powers-of-two” approximation of the kernel [32] such that the kernel coefficients can be represented by powers of two. This is very efficient and provides excellent approximations in many cases.

The development of these methods has brought the computation of discrete fixed kernel TFDs into or close to the realm of practicality for many applications. The MATLAB scripts provided with this chapter are not nearly as fast and efficient as is possible. However, they perform reasonably and were designed to be exact and easily understood as direct applications of the concepts presented in this chapter.

7.3.2.1 Windowing

Truncation of the kernel matrix at some value of m is required for practical computation. That value was chosen to be $|m| = 128$ in the examples provided in this chapter. This effectively provides rectangular windowing with all its liabilities. Essentially, the frequency marginal is now the conventional periodogram with its strong side lobes. Better spectral shaping can be achieved. Additional windowing may be injected by multiplying the sequence $X(n)$ by a window just before applying the FFT in Equation (7.19). Because this is well covered in introductory textbooks, it is not discussed further here. One should note that the frequency marginal is altered by the additional windowing, however. Windows are needed to compute the spectrogram and this windowing imposes an uncertainty effect. However, the purpose of the window with the spectrogram is to provide a “snapshot” of a limited signal segment. This is not the

case for high-resolution TFDs such as those discussed in this chapter. In that case, the window is limited in length for computational convenience, but it could be much longer to achieve better frequency resolution.

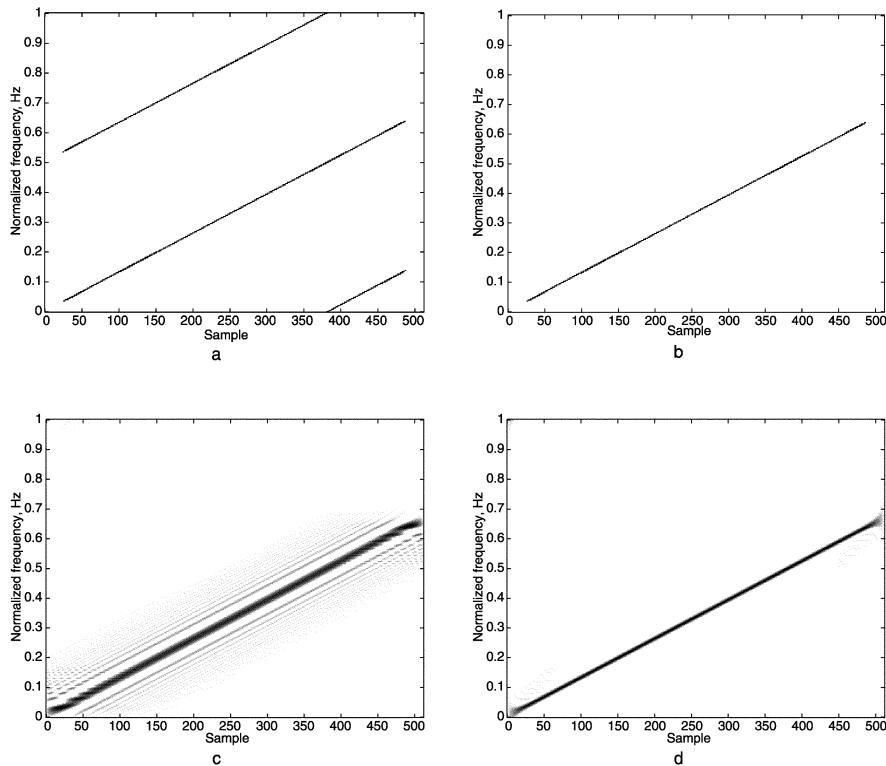
7.3.2.2 Nonnegativity

The quasi-Wigner TFD, the discrete Born–Jordan TFD and the binomial TFD exhibit negative values, caused by interference terms. Although this may seem to be disturbing, it often does not carry serious consequences. In fact, conventional spectral analysis yields negative spectral values for several conventional windows. This is rightly regarded as a numerical artifact. Nonnegative TFDs with proper marginals can be achieved [33], but with the loss of other desirable properties and a higher computational burden. One can often gain additional insight by applying several techniques on a given signal before deciding on the method to use in the final analysis.

7.3.3 Comparisons of group delay, time–frequency distributions

A comparison of group delay, time–frequency distributions (GDTFDs) is appropriate at this point. The first comparison of the performance of the GDTFD using a series of kernels is provided using a maximum $|m|$ value of 128. [Figure 7.2](#) shows the result for a rising chirp (in the analytical form) that exceeds the Nyquist frequency for a short time at the end. The conventional (undersampled) Wigner kernel promotes aliasing as seen in [Figure 7.2\(a\)](#). The discrete quasi-Wigner kernel provides an alias-free result as shown in [Figure 7.2\(b\)](#). The binomial kernel provides an unaliased result, with slight broadening of the chirp in [Figure 7.2\(c\)](#). The discrete Born–Jordan kernel provides an unaliased result with more broadening and distinct side-lobe activity in [Figure 7.2\(d\)](#). The kernels in the (θ, m) domain are shown in [Figure 7.3](#). Black denotes a value close to one and white denotes a value close to zero. The Wigner kernel — [Figure 7.3\(a\)](#) — shows the undersampled structure. A continuous Wigner kernel would be one everywhere. The quasi-Wigner kernel shows the desired unity values along $m = 0$ and $\theta = 0$, which is an RID requirement ([Figure 7.3b](#)). The discrete Born–Jordan kernel is shown in [Figure 7.3\(c\)](#) and the binomial kernel is shown in [Figure 7.3\(d\)](#). Both have a distinct RID structure, but the Born–Jordan kernel exhibits distinct side-lobe activity along hyperbolic curves. All these kernels are periodic in 1 Hz, except for the discrete Wigner kernel, which is actually periodic in 0.5 Hz. The discrete Born–Jordan kernel has a smaller “island” around the origin, which means that there is less intersection of the chirp ambiguity function than is the case for the binomial kernel. Thus, less of the chirp ambiguity is low-pass filtered and broadened in the TF domain.

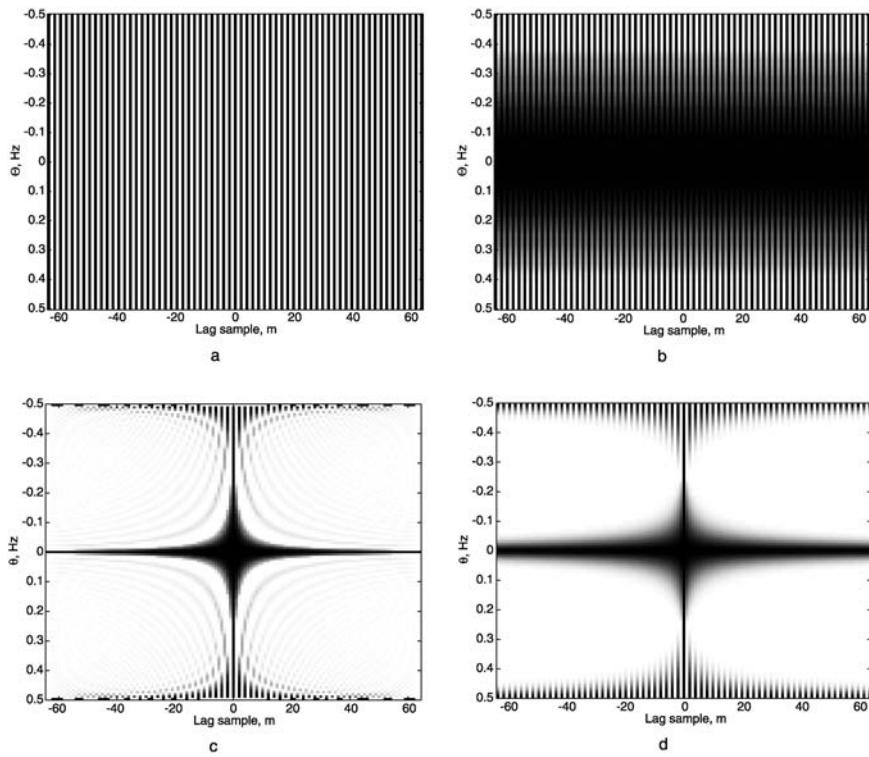
A very convenient test signal has been found to be of use in making comparisons. The test signal is composed of a bottlenose dolphin click obtained from the William Watkins group at Woods Hole Oceanographic Institution. Dolphins apparently use clicks to probe the environment. These signals, unlike simulated signals, exhibit the interesting irregularities common in naturally occurring signals. Simulated signals too often concentrate on limited aspects of signals and are not realistic for testing

**FIGURE 7.2**

Results for discrete TFDs. a.Wigner, b.quasi-Wigner, c.Born–Jordan, d.binomial TFD.

the ability of TF analysis in revealing previously unknown aspects of signals. The basic dolphin click was altered in two ways and combined with the unaltered signal to form a test signal. One alteration simply involved shifting the click in frequency. The other alteration consisted of downsampling the signal by a factor of two and then multiplying by $\sqrt{2}$ to restore the proper energy value. These signals were then concatenated to form the result shown in [Figure 7.4](#).

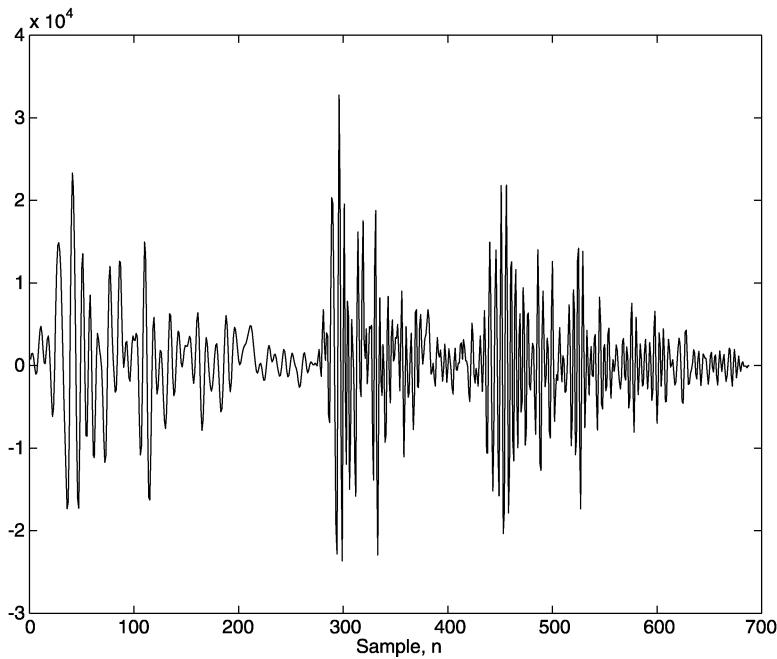
The discrete TFD results using a maximum $|m| = 128$ window are shown in [Figure 7.5](#). The spectrogram result is badly smeared and emphasizes the tonal components — [Figure 7.5\(a\)](#). The quasi-Wigner — [Figure 7.5\(b\)](#) result shows the click structure, but with considerable interference. The discrete Born–Jordan result — [Figure 7.5\(c\)](#) shows the click structure clearly with two tonal components and two impulsive or rapidly rising chirp components. The binomial result — [Figure 7.5\(d\)](#) — is similar to the discrete Born–Jordan result, but less “blocky.” The discrete Born–Jordan result exhibits some superimposed fluctuation due to side-lobe activity, but little interference between click examples. Both the quasi-Wigner result and the binomial result exhibit interference activity between the last two clicks. There is

**FIGURE 7.3**

Kernels for discrete TFDs. a.Wigner, b.quasi-Wigner, c.Born-Jordan, d.binomial TFD.

little evident interference between the first two clicks. This is likely due to the limited value of m used, preventing the inclusion of the two clicks at the same time.

The discrete RID results exhibit time-shift, frequency-shift and scale covariance. The basic signature remains the same for time shift and frequency shift. The scaled click compresses in time and expands in frequency just as required by FT properties. The spectrogram result is smeared and the scaled portion does not resemble the other parts. The quasi-Wigner result exhibits fair scale covariance, but the clicks are not as similar in structure as are the true discrete RID results. Cohen's class, in general, guarantees time-shift and frequency-shift covariance for continuous TFDs with fixed kernels. The continuous Wigner TFD and the continuous RIDs also exhibit scale covariance because the kernels are product kernels [5]. The product kernel idea cannot be extended exactly to the discrete domain, but the discrete RIDs exhibit nice scale covariance, nevertheless. Scale covariance is a feature of wavelet transforms. However, frequency covariance is not. Thus RIDs enjoy all three covariances, which is convenient for design of optimum kernels for a given signal class and for pattern recognition under variations of time shift, frequency shift and scale.

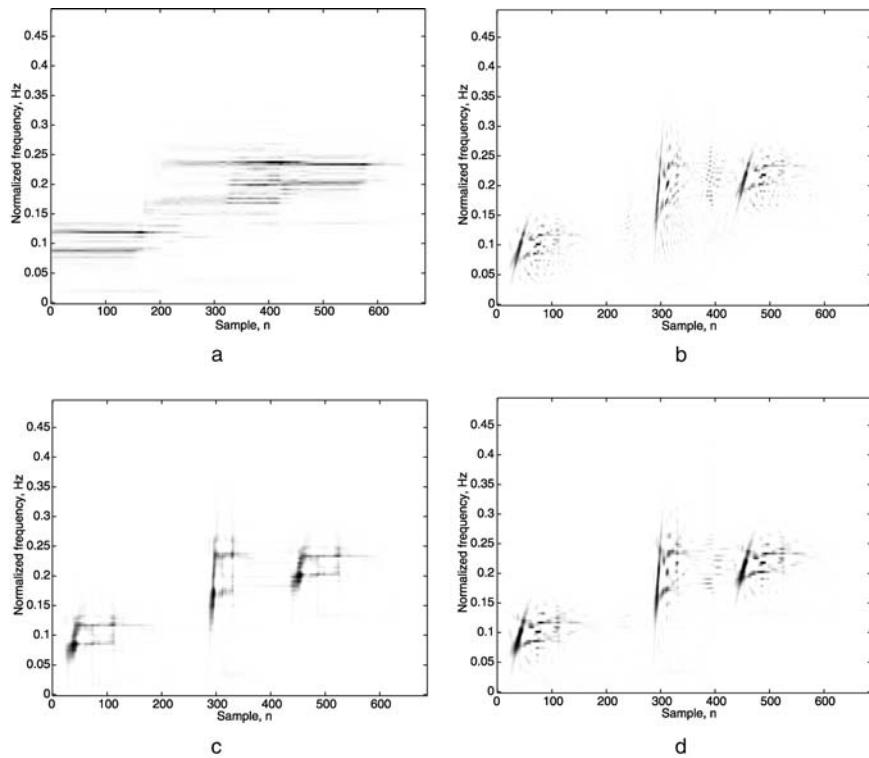
**FIGURE 7.4**

Original dolphin click, scaled and time-shifted version and frequency-shifted and time-shifted version.

The quasi-Wigner, discrete Born–Jordan and binomial TFDs have approximately correct frequency marginals. The marginal obtained by summing over all samples of n compares very well to the periodogram result using a rectangular data window. An example of the frequency marginal result is provided in Figures 7.6 and 7.12

7.4 Applications

It is hoped that sufficient theoretical and computational details have been made available to readily support a number of applications of the methods. Several examples are now presented. This is not a comprehensive review of applications. A more comprehensive treatment is available [5] and a comprehensive review of biological and biomedical applications is in preparation by the author. The intent in this chapter is to show the power of the discrete TFD methods and provide the reader with some means of replicating such results using the MATLAB scripts in Appendix A and Appendix B.

**FIGURE 7.5**

Results for three dolphin clicks. a. Spectrogram, b. quasi-Wigner, c. Born–Jordan, d. binomial TFD.

7.4.1 Underwater acoustics

Much of the work reported in this chapter was supported by the Office of Naval Research under an accelerated acoustic transients initiative. It is critical to identify and understand both natural and man-made sounds in the ocean. To this end, methods were developed to better characterize marine mammal sounds in cooperation with the William Watkins group at Woods Hole Oceanographic Institution. Watkins provided classical examples of the failures of the spectrogram in properly representing marine mammal sounds in his 1966 article on biological acoustics [34]. As an expert observer of marine mammal sounds, Watkins was dissatisfied with presently available methods for characterizing marine mammal sounds. RID methods were developed for his group and have been in use there for several years. Dolphins have an uncanny ability to detect and identify underwater objects, even those buried in mud. The dolphin click structure presented here is only one example of their complex sound-making ability. However, the structure could serve to determine range and Doppler of a target. The two impulsive structures could serve to determine range and the two tones, Doppler effects. Single tones and single impulses would potentially provide

ambiguous results. Time delay and scaling of return echoes would be available from these clicks, however, as well as frequency shifts. In addition to understanding what uses marine mammals make of their complex sound structures, it is of interest to possibly identify individual marine mammals by their sound signatures. Such signatures have been used to identify individual animals [35]. Methods have been developed to provide an invariant representation of signals that have undergone time shifts, frequency shifts and scale changes [36]. Individual sperm whales have been identified using these techniques [35].

7.4.2 Speech

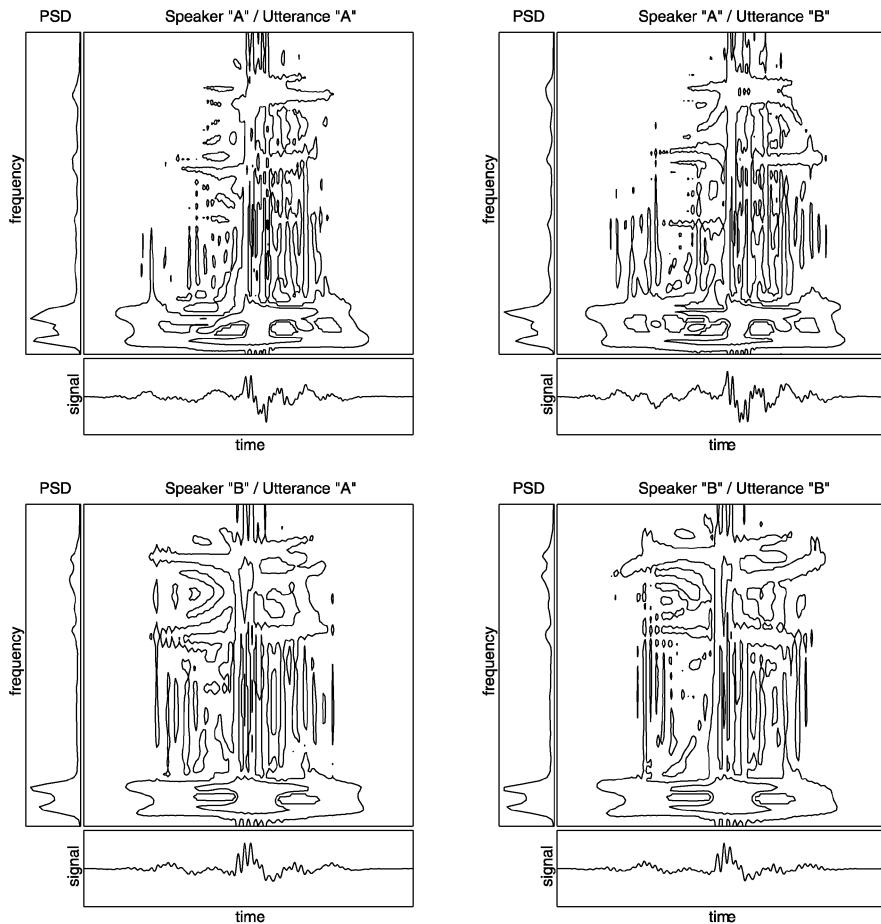
Human speaker identification and verification is an increasingly important issue. Time-frequency signatures of speakers have been used for this purpose [36, 37]. In contrast to present methods for speech analysis, TF-based methods can focus on individual pitch periods of speech. This is a microspeech structure method in contrast to present methods, which focus on the macrostructure of speech. These results are presently competitive with macromethods and provide independent information about the speaker, potentially leading to improved accuracy. [Figure 7.6](#) compares the binomial TFD signatures for two speakers for two different utterances. One can see that the TFDs for the same speaker for different utterances have similar structure. These signatures appear to be highly individualized.

7.4.3 Brain signals

TF methods are useful in characterizing epileptic seizures [38]. Spike and slow-wave components in electroencephalography (EEGs) are related to certain types of epilepsy [5], and these signals exhibit unique TF signatures. Spectrograms provide blurred representations of these phenomena, but the binomial RID provides a dramatic view of their activity. [Figure 7.7](#) shows the binomial TFD of a series of spike and slow waves. One can see that the energy trajectories rapidly change from low frequency to high frequency and then return just as rapidly to low frequencies. A simple third-order differential equation with time-varying coefficients can model this phenomenon well [5]. One can extract a slice of the TFD at high frequencies to form a time series that exhibits a strong impulsive structure. This form is useful in identifying the spikes and studying their occurrences in terms of a point process conceptualization.

7.4.4 Temporomandibular joint sounds

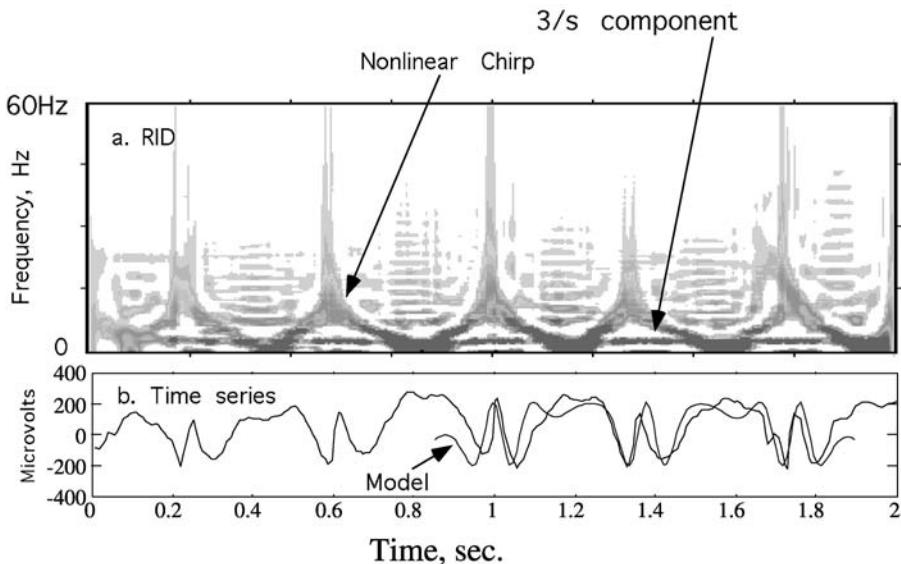
Many people have experienced the popping or clicking of a joint. Sometimes the temporomandibular joint (TMJ) clicks or pops when the jaw is moved. In many cases, this is a rare and harmless event. However, it may be indicative of a potentially serious condition. Such activity may be associated with a great deal of pain, perhaps even triggering headaches or other distressing symptoms. Even if one does not experience pain, these joint sounds may herald the development of more serious conditions. There is almost always someone in the audience with personal knowledge of pain and

**FIGURE 7.6**

Binomial TFDs for two speakers. Top, speaker A utterances A and B; bottom, speaker B utterances A and B. (From W.J. Williams, *Proc. IEEE*, 84[9], 1264–1280, 1996. With permission.)

suffering from these problems, in terms of themselves, friends or family members when TF results are presented on these TMJ sounds. Previous work aimed at analyzing these signals has been spotty and mixed in results. Usually, the clinician simply listens to the sound and tries to report the results in some objective way. However, this method is highly subjective and often couched in terms such as likening the sound to “a dry stick breaking” or “the crunching of dry snow.” Arthritic joints, in particular, often produce creaking or crunching sounds. This is often called *crepitus*.

TFD is very useful in the analysis of these sounds [39]. The results of the TMJ click analysis using high-resolution TFD techniques are shown in Figure 7.8. The WD and the binomial TFD show similar results when used to analyze the click. The binomial TFD shows a very broadband component of very short time duration, followed by

**FIGURE 7.7**

Spike and slow-wave example. a. Binomial TFD and b. time series. The time series is shown with a simple model result. (From W.J. Williams, *Proc. IEEE*, 84[9], 1264–1280, 1996. With permission.)

a rather prolonged low-frequency component and a shorter-duration component at a higher frequency. The initial broadband component is produced by the discontinuity at the start of the click. The signal (plotted along the front margin) rises rapidly to a very sharp peak. The TFD of an impulse, $\delta(t)$ should produce a knife-edge aligned along frequency. The FT of $\delta(t - t_o)$ is a constant with frequency, but exists only at the time of the impulse.

The pad that cushions the action of the condyle as it moves within the joint capsule is thought to be damaged in this example. Instead of moving smoothly over the pad, the condyle may push the damaged pad tissue ahead of it. Finally, the condyle slips over the bunched pad tissue, producing a click. The TF results seem to reflect these events well.

Five distinct types of TMJ sounds have been identified during extensive research studies using RID techniques in the analysis. Several types of these clicks have been successfully classified using pattern recognition methods applied to the TFD signatures [40].

7.4.5 Heart sounds and muscle sounds

Heart sounds have long been utilized in diagnosis, using the simple stethoscope. Barry and Cole first used TFDs to study muscle sounds produced during contraction [41]. Wood, Barry and colleagues [42, 43], have utilized TF techniques extensively in the study of heart sounds. They have extensively explored the use of TF techniques as

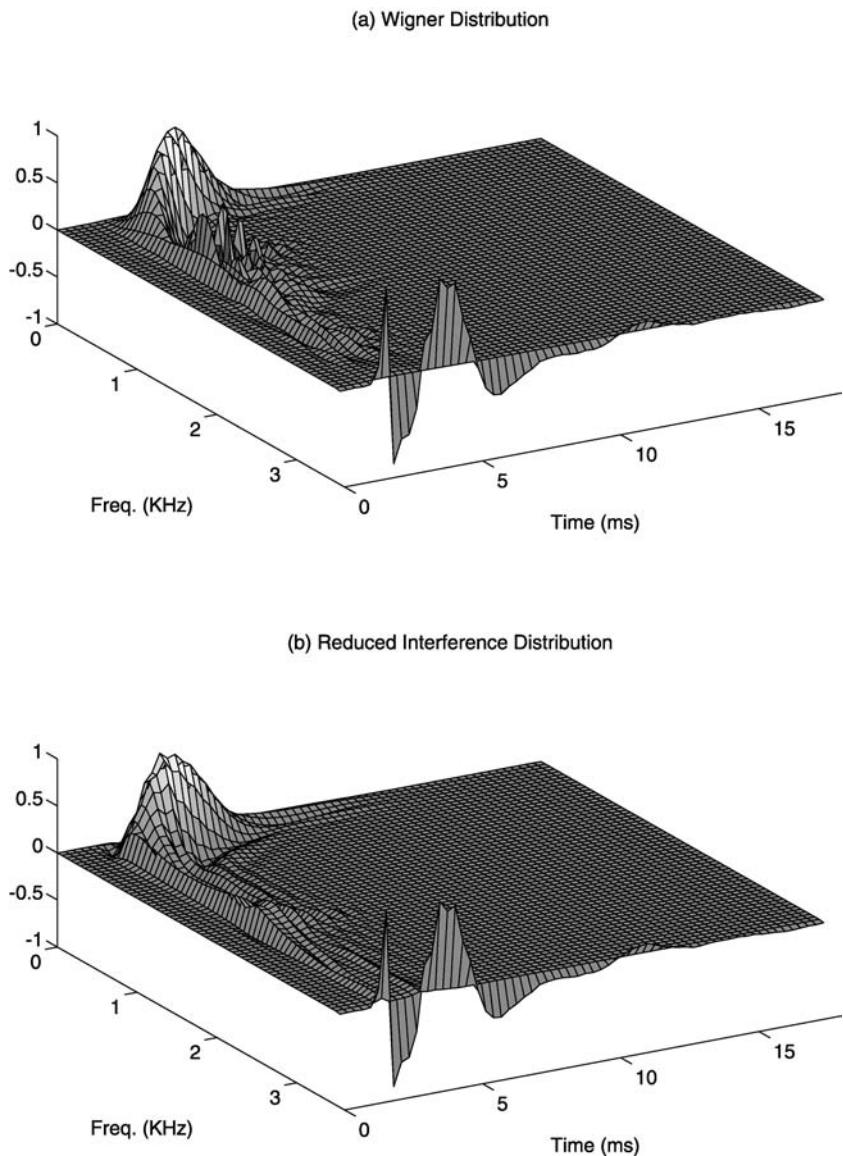


FIGURE 7.8

TMJ click analysis using a long window (a) quasi-Wigner (b) RID (binomial TFD). The signal is appended at the front of the mesh plots as a reference. (From W.J. Williams, *Proc. IEEE*, 84[9], 1264–1280, 1996. With permission.)

tools for the analysis of heart sounds. The binomial TFD has been found to be quite useful in these studies. Heart valve action produces sharp impulses and contracting heart muscle produces chirps. These features can be observed in TF results.

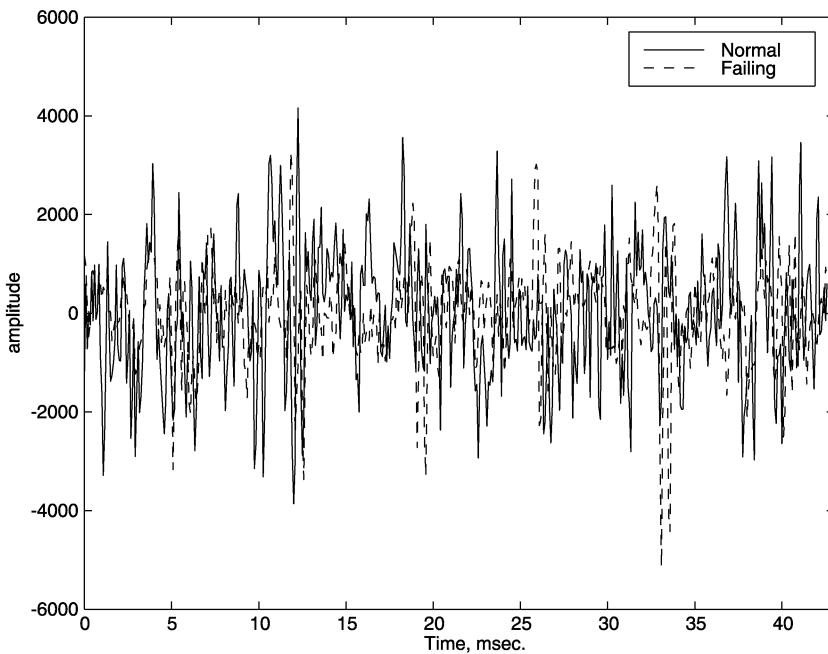
7.4.6 Machine signals

TF analysis has a potentially large role in machine monitoring and diagnosis. Atlas, Bernard and Narayanan have pioneered such methods [44]. Binomial TFDs have been found to be useful in detecting the impending failure of bearings [45]. The impulsive structures found in EEGs have a similar analog in machine monitoring. When chips or defects form in bearings, they can be identified by extracting a slice from the TFD at a high frequency. The impulsive structure of these slices is highly indicative of a developing fault. The example presented here is for a ball bearing with three balls rotating at 3600 r/min. Accelerometer signals were examined at the beginning of the test and just before failure to examine the character of these signals in normal and failing conditions. It is often difficult to see the precursors of failure in the time series, or in the power spectrum of these signals, though some hints are available. The revolutions per second (r/sec) value of the signal rises rapidly at failure, but it is difficult to predict failure far ahead of time using this measure. Impulsive components begin to appear in the signal near failure, but these components are difficult to sort out from other signal components. [Figure 7.9](#) shows time series for the vertical axis for a bearing in the normal condition and in the failing condition. The failing condition is defined to be 120 sec before failure in this case. One cannot see an obvious difference in the raw signals. The binomial TFD provides a dramatic picture of the differences in signal structure in the two conditions, however, as shown in [Figure 7.10](#). The binomial distribution obtained during the normal state of the bearing shows some structure, but the energy patterns as a function of time and frequency are somewhat variable. As the bearing approaches failure, an increase occurs in the vertical structuring of the binomial TFD as well, indicating increasing impulsive components. Energy becomes more concentrated into discrete components during failure. Our research has shown that bearing failure can be predicted well before the actual failure results using TFD methods.

7.4.7 Research in physics

7.4.7.1 High energy plasma discharges

High energy microwave device discharges are very well characterized using the binomial TFD and its derivatives. These methods have provided new insights into the nature of the phenomena not previously available and have greatly helped in understanding the physics of these phenomena [46]. The binomial TFD of the output from a coaxial gyrotron oscillator and the accompanying voltage fluctuations on the diode reveal that the frequency modulation (FM) on the radio frequency (RF) output is directly correlated with the voltage fluctuations. This result is shown in [Figure 7.11](#).

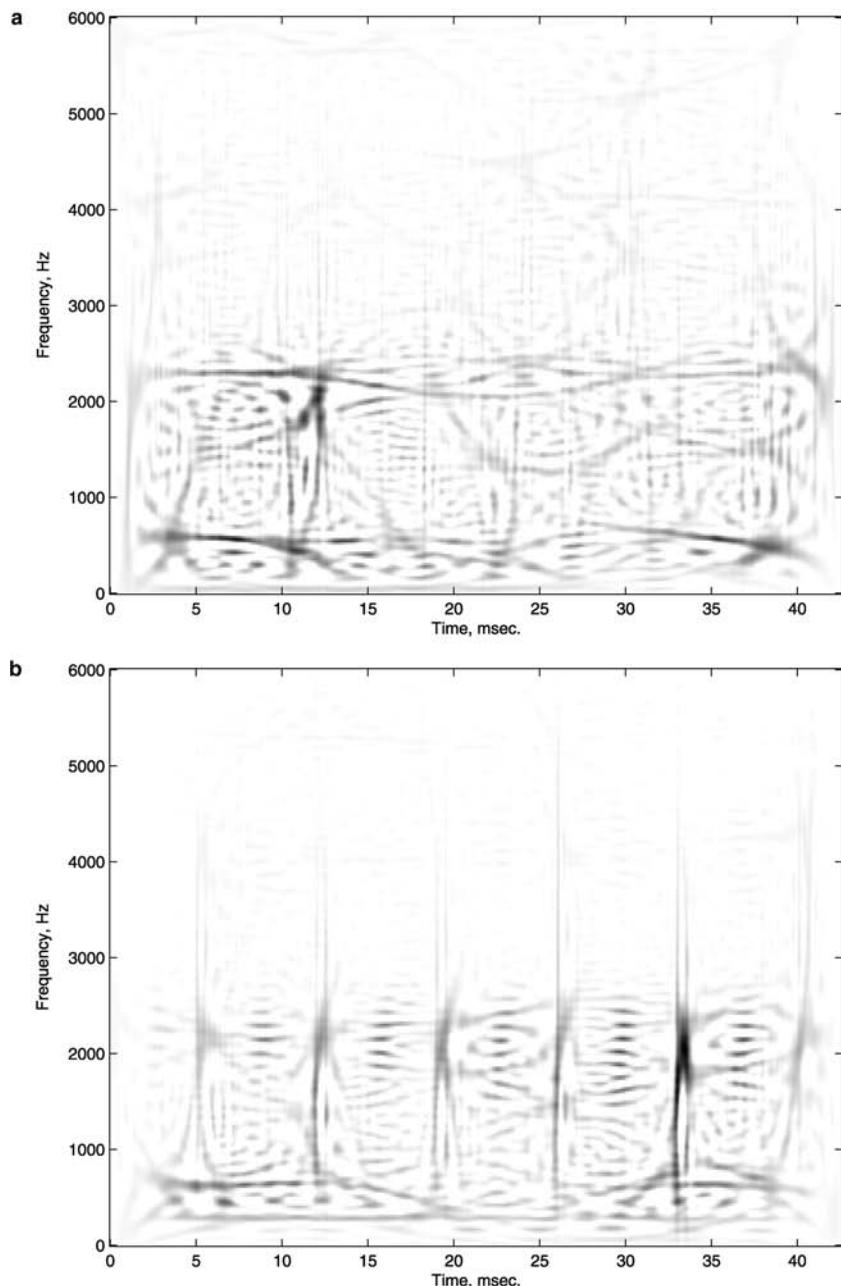
**FIGURE 7.9**

Comparison of a normal and failing bearing signal.

This technique is under extension and is opening up new ways of investigating pulse shortening, mode competition, noise and unwanted frequencies in RF generation.

7.4.7.2 Turbulent air flow

Turbulent air flow produces numerous brief vortices that form and dissipate rapidly. It seems to be widely accepted by the turbulence community that the intermittency observed in fully turbulent flows is closely related to the existence of intense vorticity events, localized in time and space, also known as coherent structures. Classical methods have commonly provided a view of the power spectra of these effects. The binomial TFD has revealed the precise time and nature of these phenomena and revealed their relationships across scales in an important study of turbulence [47]. By using TF distributions for the analysis of the scattered acoustic signals, one can show how the legibility of these signals is significantly improved as illustrated in Figure 7.12. It is important to note that preserving the power spectra in terms of the frequency marginals adds to the credibility of the TF representation as far as the physics community is concerned.

**FIGURE 7.10**

Comparison of a.normal and b.failing bearing Binomial TFDs using a 512-point window.

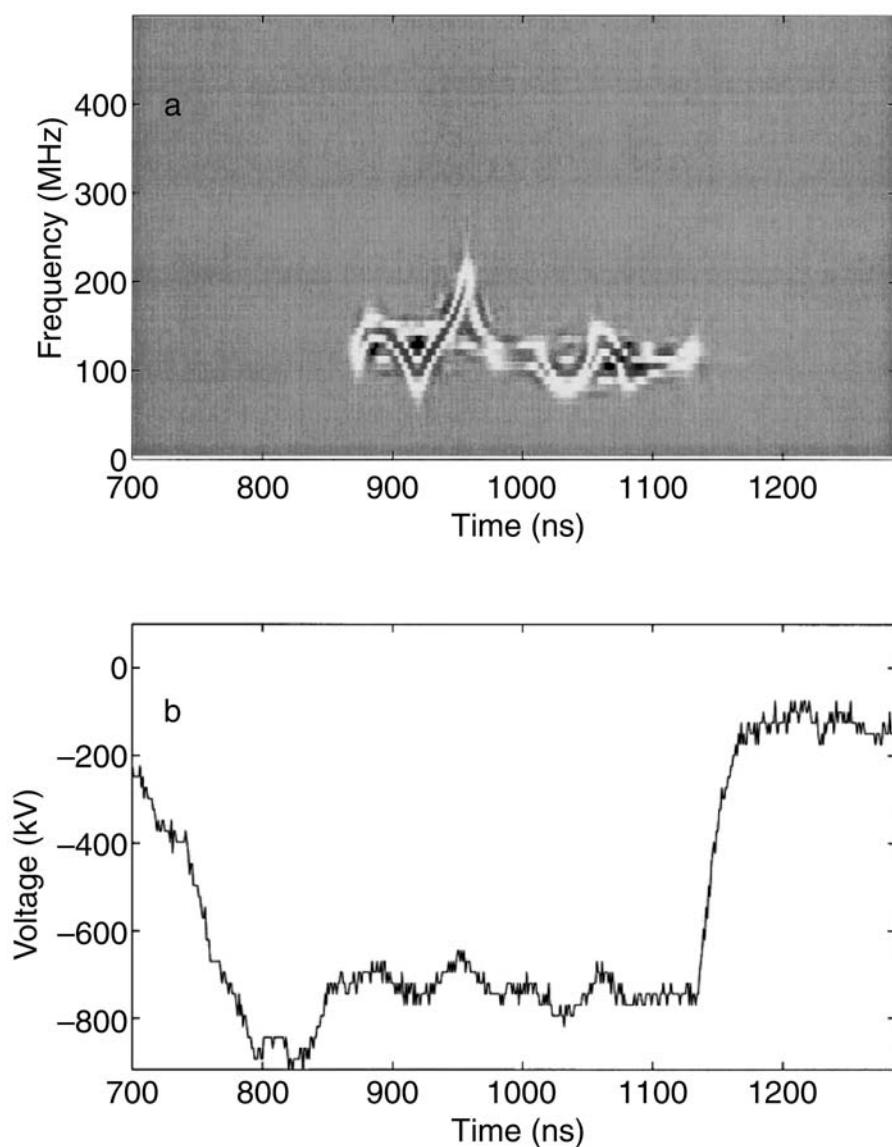
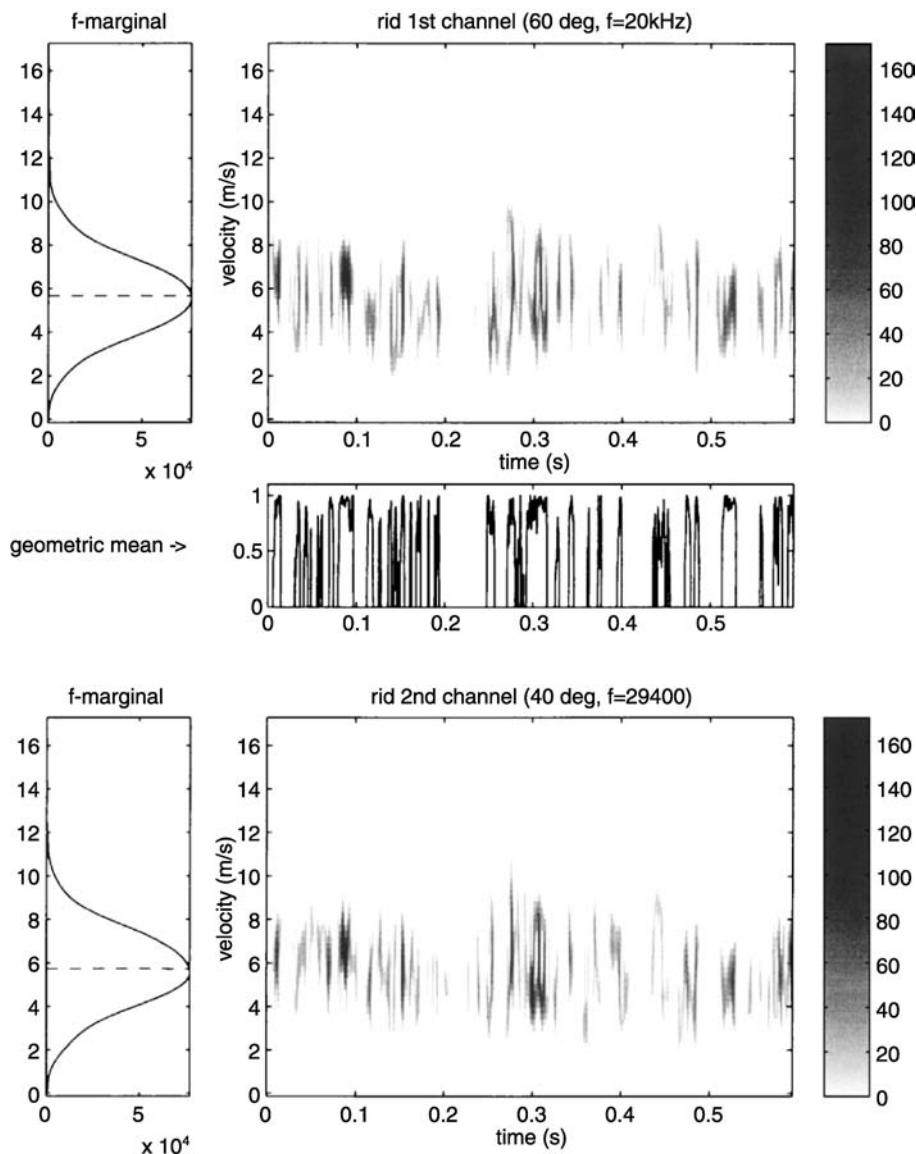


FIGURE 7.11

a. Binomial TFD and b. diode voltage fluctuation for a high energy microwave device during discharge.
(From C.W. Peters et al., *Phys. Rev. E*, 58[5], 6880–6883, 1998. With permission.)

**FIGURE 7.12**

Turbulence results using binomial TFD. Top, RID for 20 KHz with frequency marginal to the left. Bottom, RID for 29.4 KHz with frequency marginal to the left. Middle, geometric mean of the two signals. (From C. Baudet, O. Michel and W.J. Williams, *Phys. D*, 128[1], 1–17, 1999. With permission.)

7.5 Conclusions

The purpose of this chapter is to introduce the concepts of discrete time reduced interference distributions, and provide easily understandable mathematical treatments and useful MATLAB scripts for computation of the discrete TFDs so that the reader can repeat the examples and try the methods on their favorite application. Much more efficient and comprehensive software is available for potential power users of these techniques via www.QuantumSignal.com. TF analysis methods are slowly finding their way into many areas. It is hoped that this chapter aids and speeds that process.

Acknowledgments

This research was supported in part by grants from the Rackham School of Graduate Studies and the Office of Naval Research, National Science Foundation: grant no. BCS-9110571 and ONR grants no. N00014-89-J-1723, N00014-90-J-1654 and N000014-97-1-0072. Many thanks go to graduate students and colleagues without whose contributions this work would not have been possible. The author is also grateful for the careful reading of the manuscript by Eugene Zalubas.

References

- [1] L. Cohen, *Time-Frequency Signal Analysis*, Prentice Hall, New York, 1995.
- [2] P. Flandrin, *Time-Frequency Time Scale Analysis*, Academic Press, San Diego, CA, 1999.
- [3] L. Cohen and T. Posch, Generalized ambiguity functions, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 85, 1025–1028, 1985.
- [4] H.I. Choi and W.J. Williams, Improved time-frequency representation of multi-component signals using exponential kernels, *IEEE Trans. Acoust. Speech Signal Process.*, 37(6), 862–871, 1989.
- [5] W.J. Williams, Reduced interference distributions: biological applications and interpretations, in Proceedings of the IEEE, 84(9), 1264–1280, 1996.

- [6] W.J. Williams, The reduced interference distribution, in *Time-Frequency and Wavelets in Biomedical Signal Processing*, M. Akay, Ed., IEEE Press, Piscataway, NJ, 1997, pp. 3–39.
- [7] L. Cohen, Time-frequency distributions — a review, in Proceedings of the IEEE, 77(7), 941–981, 1989.
- [8] Y. Zhao, L.E. Atlas, and R.J. Marks, II, The use of cone-shaped kernels for generalised time-frequency representations of non-stationary signals, *IEEE Trans. Acoust. Speech Signal Process.*, 38(7), 1084–1091, 1990.
- [9] J.C. O'Neill and W.J. Williams, Shift covariant time-frequency distributions of discrete signals, *IEEE Trans. Signal Process.*, 47(1), 133–146, 1999.
- [10] J.C. O'Neill, P. Flandrin, and W.J. Williams, On the existence of discrete Wigner distributions, *IEEE Signal Process. Lett.*, 6(12), 1999.
- [11] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, The Wigner distribution — a tool for time-frequency signal analysis — Part II. Discrete time signals, *Philips J. Res.*, 35, 276–300, 1980.
- [12] T.A.C.M. Classen and W.F.G. Mecklenbräuker, The aliasing problem in discrete-time Wigner distributions, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-31(5), 1067–1072, 1983.
- [13] F. Peyrin and R. Prost, A unified definition for the discrete-time discrete-frequency, and discrete-time/frequency Wigner distributions, *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-34(4), 858–867, 1986.
- [14] A.H. Nuttall, Alias-free discrete-time Wigner distribution function and complex ambiguity function, Naval Underwater Systems Centre, April 14, 1989, Technical report 8533.
- [15] J. Jeong and W.J. Williams, Alias-free generalized discrete-time time-frequency distributions, *IEEE Trans. Signal Process.*, 40(11), 2757–2765, 1992.
- [16] E.C. Bekir, A contribution to the unaliased discrete-time Wigner distribution, *J. Acoust. Soc. Am.*, 93(1), 363–371, 1993.
- [17] W.J. Williams and J. Jeong, Reduced interference time-frequency distributions, in *Time-Frequency Signal Analysis: Methods and Applications*, B. Boashash, Ed., Longman Cheshire, Melbourne, Australia, 1991, chap. 3.
- [18] L. Cohen, Generalized phase-space distribution functions, *J. Math. Phys.*, 7, 781–786, 1966.
- [19] J.M. Morris and D. Wu, On alias-free formulations of discrete Cohen's class distributions, *IEEE Trans. Signal Process.*, 44(6), 1355–1364, 1996.
- [20] A.H. Costa and G.F. Boudreault-Bartels, An overview of aliasing errors in discrete-time formulations of time-frequency distributions, *IEEE Trans. Signal Process.*, 47(5), 1463–1474, 1999.

- [21] L.J. Stanković and I. Djurović, A note on “an overview of aliasing errors in discrete-time formulations of time-frequency representations,” *IEEE Trans. Signal Process.*, 40(1), 257–259, 2001.
- [22] J. Jeong and W.J. Williams, Kernel design for reduced interference distributions, *IEEE Trans. Signal Process.*, 40(2), 402–412, 1992.
- [23] B. Boashash, Time-frequency signal analysis, in *Advances in Spectrum Estimation*, S. Haykin, Ed., Prentice Hall, New York, 1989, pp. 418–517.
- [24] W.J. Williams, Reduced interference time-frequency distributions: scaled decompositions and interpretations, in *Wavelet Transforms and Time-Frequency Signal Analysis*, L. Debnath, Ed., Birkhäuser, Boston, 2001, pp. 381–417.
- [25] D.T. Barry, Fast calculation of the Choi–Williams distribution, *IEEE Trans. Signal Process.*, 40(2), 450–455, 1992.
- [26] J.R. O’Hair and B.W. Suter, The Zak transform and decimated time-frequency distributions, *IEEE Trans. Signal Process.*, 44(5), 1099–1110, 1996.
- [27] G.S. Cunningham and W.J. Williams, Fast implementations of discrete time-frequency distributions, *IEEE Trans. Signal Process.*, 42(6), 1496–1508, 1994.
- [28] T. Le and M. Glesner, A flexible and approximate computing approach for time-frequency distributions, *IEEE Trans. Signal Process.*, 48(4), 1193–1196, 2000.
- [29] W.J. Williams and S. Aviyente, Optimum window time-frequency distribution decompositions, in IEEE Proceedings of the 32nd Asilomar Conference on Signals, Systems and Computers, 1, 817–821, 1998.
- [30] S.S. Aviyente and W.J. Williams, Discrete scale vectors and decomposition of time-frequency kernels, *Adv. Signal Process. Algorithms, Architectures, Implementations X, Proc. SPIE*, 4116, 100–109, 2000.
- [31] M.G. Amin, Spectral decomposition of time-frequency distribution kernels, *IEEE Trans. Signal Process.*, 42(5), 1156–1165, 1994.
- [32] G.T. Venkatesan and M.G. Amin, Time-frequency distribution kernel design over a discrete powers-of-two space, *IEEE Signal Process. Lett.*, 3(12), 305–306, 1996.
- [33] P.J. Loughlin, J.W. Pitton, and L.E. Atlas, Construction of positive time-frequency distributions, *IEEE Trans. Signal Process.*, 42(10), 2679–2705, 1994.
- [34] W.A. Watkins, The harmonic interval fact or artefact in spectral analysis of pulse trains, *Mar. Bio-acoustics*, 2, 15–43, 1966.
- [35] E. Zalubas, J. O’Neill, W. Williams, and A. Hero, Shift and scale invariant detection, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5, 3637–3640, 1997.

- [36] W.J. Williams, E.J. Zalubas, R.M. Nickel, and A.O. Hero, III, Scale and translation invariant methods for enhanced time-frequency pattern recognition, *Multi-dimensional Syst. Signal Process.*, 9(4), 465–473, 1998.
- [37] R.M. Nickel and W.J. Williams, On local time-frequency features of speech and their employment in speaker verification, *J. Franklin Inst.*, 337(4), 469–481, 2000.
- [38] H.P. Zaveri, W.J. Williams, L.D. Iasemidis, and J.C. Sackellares, Time-frequency representations of electrocorticograms in temporal lobe epilepsy, *IEEE Trans. Biomed. Eng.*, 39(5), 502–509, 1992.
- [39] W.J. Williams, S.E. Widmalm, and C.S. Zheng, Time-frequency distributions of tmj sounds, *J. Oral Rehabil.*, 18, 403–412, 1991.
- [40] D. Djurdjanovic, S.E. Widmalm, W.J. Williams, C.K.H. Koh, and K.P. Yang, Computerized classification of temporomandibular joint sounds, *IEEE Trans. Biomed. Eng.*, 47(8), 977–984, 2000.
- [41] D.T. Barry and N.M. Cole, Muscle sounds occur at the resonant frequency of skeletal muscle, *IEEE Trans. Biomed. Eng.*, 37, 525–531, 1990.
- [42] J.C. Wood and D.T. Barry, Quantification of first heart sound frequency dynamics across the human chest wall, *Med. Biol. Eng. Comput.*, 32, S71–S78, 1994.
- [43] J.C. Wood, A.J. Buda, and D.T. Barry, Time-frequency transforms: a new approach to first heart sound frequency analysis, *IEEE Trans. Eng. Med. Biol.*, 39(7), 728–739, 1992.
- [44] L.E. Atlas, G.D. Bernard, and S.B. Narayanan, Applications of time-frequency analysis to signals from manufacturing and machine monitoring sensors, in Proceedings of the IEEE, 84(9), 1319–1329, 1996.
- [45] W.J. Williams, C.K.H. Koh, and J. Ni, Bearing monitoring using reduced interference time-frequency distributions, in *Life Extension of Aging Machinery and Structures*, H.C. Pusey and S.C. Pusey, Eds., The Vibration Institute, 1995, pp. 315–326.
- [46] C.W. Peters, R.L. Jaynes, Y.Y. Lau, R.M. Gilgenbach, W.J. Williams, J.M. Hochman, W.E. Cohen, J.I. Rintamaki, D.E. Vollers, and T.A. Spencer, Time-frequency analysis of modulation of high-power microwaves by electron-beam voltage fluctuations, *Phys. Rev. E*, 58(5), 6880–6883, 1998.
- [47] C. Baudet, O. Michel, and W.J. Williams, Detection of coherent vorticity structures using time-scale resolved acoustic spectroscopy, *Phys. D*, 128(1), 1–17, 1999.

Appendix A — Generalized Discrete Time–Frequency Distribution Scripts

gdtfdo.m

```
function q=gdtfdo(x,N,kr)
%x is the signal vector in row form, N is the window size,
kr is the kernel matrix (APPENDIX B)
%Find the size of the input signal. The signals is a row vector with n samples.
[m, n] = size(x);
%Pad the signal with zeros to have room at the ends.
xx=[zeros(1,N) x zeros(1,N)];
%Run a window along the signal and compute a frequency slice for each n point.
for k=1:n
%Grab a backward segment from the sample point.
tem1=fliplr(xx(k+1:k+N));
%Grab a forward segment from the sample point and form the LAF matrix in proper
%orientation.
tem2=xx(N+k:2*N+k-1);
uu=fliplr(tem1)'*tem2;
td=kr.*uu;
for r=1:N
%Pick off correlation values along each lag sample.
%Apply the kernel to the LAF vector.
aa(r)=sum(diag(td,-N+r));
end
aaa=[aa fliplr(conj(aa(2:N)))];
tf=real(fft(aaa));q(:,k)=real(tf)';
end
```

Appendix B — Kernel Scripts

BornJordan.m

```
function k=BornJordan(N)
%k=BornJordan(N)
%N is the size of the kernel
kk=diag((1/N)*ones(1,N));
for n=1:N-1
a=diag((1/n)*ones(1,n),-N+n);
kk=kk+a;
end
k=kk;
```

binomial.m

```
function k=binom(N)
%k=binom(N), N is the size of the kernel
k=zeros(N,N);
k(1,N)=1;
pair=[.5 .5];
v=1;
for m=2:N
v=conv(v,pair);
u=diag(v,N-m);
k=k+u;
end
k=fliplr(k);
```

quasi-Wigner.m

```
function k=quasi-wigner(N)
k1=eye(N);
halfs=.5*ones(1,N-1);
km=diag(halfs,-1);
kp=diag(halfs,1);
k=flipud(k1+km+kp);
```

Copyright Notice

These MATLAB scripts are the sole property of the author, William J. Williams, and may be freely used for educational and individual explorations. Their use is also permitted for such research and publication as long as this copyright is acknowledged. These scripts remain the sole property of the author, including any alterations and revisions. They may not be used for commercial purposes without permission of the author. In addition, certain aspects of the methods are the subject of patents presently owned by the University of Michigan and the appropriate authority for commercial use and licensing must be obtained before such use.

8

Time–Frequency Analysis of Seismic Reflection Data

Philippe Steeghs

TNO Physics and Electronics Laboratory

Richard G. Baraniuk and Jan Erik Odegard

Rice University

CONTENTS

8.1	Introduction	307
8.2	Seismic Sequence Analysis	309
8.3	Time–Frequency Representations for Seismic Signal Analysis	312
8.4	Seismic Attribute Extraction	317
8.5	Hybrid Linear and Quadratic Time–Frequency Seismic Attributes	322
8.6	Three-dimensional Seismic Attribute Extraction	325
8.7	Conclusions	335
	References	336

8.1 Introduction

In this chapter we apply quadratic time–frequency (TF) representations to the analysis of seismic reflection data. Seismic imaging of the earth’s subsurface is an essential technique in exploring for oil and gas accumulations. A seismic image is obtained by probing the subsurface with acoustic waves. An example of a seismic cross section is shown in [Figure 8.1\(b\)](#). The horizontal axis is the spatial location at the surface, and the vertical axis is time. Each column in this image represents a recording of the reflected wave amplitude as a function of time at the corresponding surface location.

When seismic waves propagate through the subsurface, energy is reflected back toward the surface at acoustic impedance contrasts. The strength of this impedance contrast is called the *reflectivity*. The seismic cross section of [Figure 8.1\(b\)](#) is a representation of the (band-limited) reflectivity of the subsurface. If the seismic wave velocity is known, then the time axis can be converted to depth. In the image of [Figure 8.1\(b\)](#), a two-way (down and up) travel time of 1 sec corresponds roughly

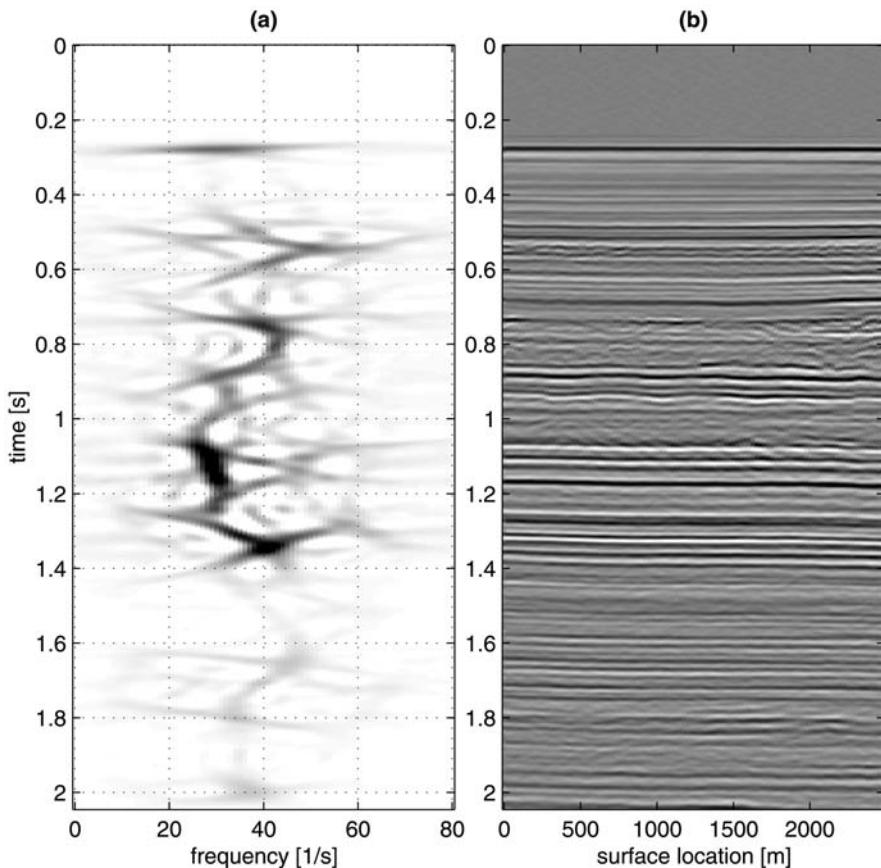


FIGURE 8.1

(a) Time–frequency representation of the leftmost seismic trace (column) of the seismic cross-section image shown in (b).

to a depth of 1 km below the surface. When a rapid change in acoustic impedance occurs at a certain depth, this generally implies a change in the composition of the rocks. As a result, it is possible to make an educated guess of the geologic structure of the subsurface, based on the seismic image.

In the following we show how TF analysis of seismic images can help to identify and classify sequences of seismic reflections. The characteristics of such sequences can provide a hint of the geologic processes that resulted in the observed reflectivity.

Figure 8.1(a) is a TF representation of the first vertical time trace of the seismic section of Figure 8.1(b). Changes in the TF patterns coincide with boundaries of seismic sequences. The location and distinctive character of the sequences is more pronounced in the TF representation than in the time signal, illustrating the added value of TF analysis for seismic sequence interpretation.

The analysis and classification of seismic waveforms is called *seismic attribute analysis*. Seismic attributes are features extracted from a seismic image that elucidate signal characteristics that are relevant for the geologic interpretation of the reflectivity image. Until the last few years, the emphasis has been on one-dimensional (1-D) analysis along the vertical (time) axis. This is because the vertical axis corresponds to the chronology of geologic events — rocks that were formed longer ago in geologic time are found below younger sediments. Nowadays, three-dimensional (3-D) seismic data acquisition has become the standard in the petroleum industry. The result is that the volume of data analyzed has grown tremendously. Paper plots of seismic cross sections have been replaced by gigabytes of image data that are interpreted on computer screens or, increasingly, in immersive visualization environments. Seismic attributes are increasingly important for detection and visualization of subsurface structures that are hidden in these enormous volumes of data. In conjunction with this development, there also has been a shift toward full 3-D feature extraction and analysis. In the last part of this chapter we discuss 3-D seismic data analysis with extensions of our TF analysis methods to higher dimensions.

8.2 Seismic Sequence Analysis

The frequency content of a seismic reflection record is primarily dependent on the bandwidth of the outgoing seismic source pulse and the absorption characteristics of the subsurface. Variations within this band are primarily the result of changes in the timing of seismic reflections. A reflection sequence can be described with attributes such as the continuity, amplitude and frequency of the reflections. The TF representation of a seismic section brings forward characteristics of the seismic sequence that are not easily observed in the time domain or frequency domain alone. We are mainly interested in localizing strong transitions in frequency characteristics over time, because these transitions indicate where changes occur in the geologic circumstances under which the rocks were formed. The more gradual time variations of frequency content in-between these transitions are also of interest, because these may provide clues for relating the signal characteristics to the geologic process that resulted in a certain subsurface structure.

Our model of a seismic sequence is that of a layered earth, where each of the layers is bounded by seismic impedance discontinuities. Typically, a sequence consists of a stack of layers that in turn is bounded at the top and base by a major discontinuity. We start with an analysis of the TF representation of such a “generic sequence.”

Figure 8.1(a) illustrates a TF representation of the leftmost column of the seismic section to its right — Figure 8.1(b). We observe that the seismic signal is clearly a multicomponent signal. At about $t = 0.3$ sec a strong impulselike component exists; this is a reflection from the strong contrast between water and sea bottom. The seismic response below the sea-bottom reflection is a complicated interference pattern,

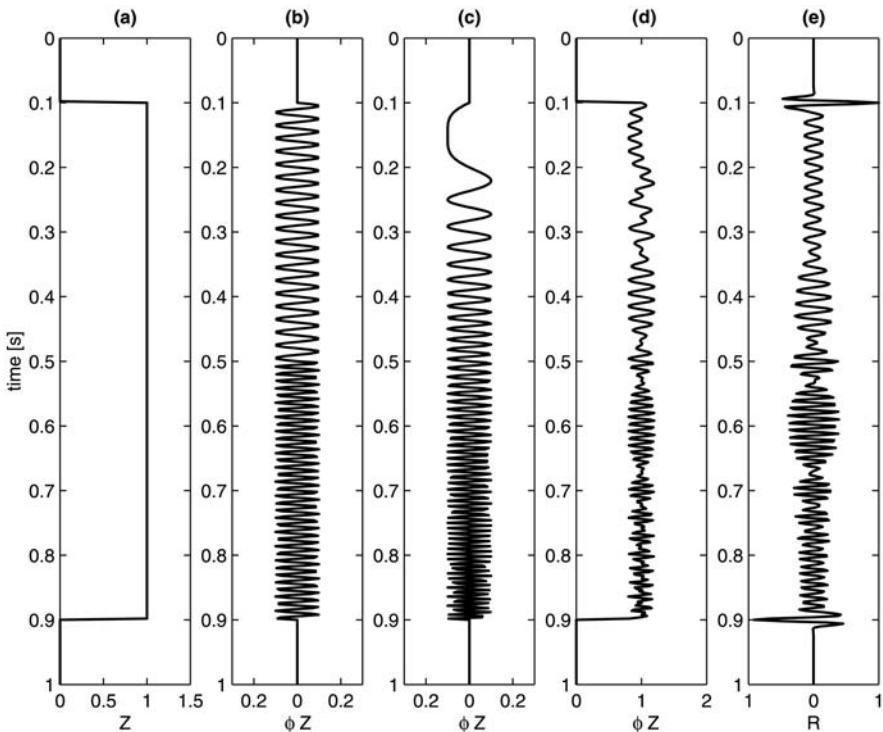
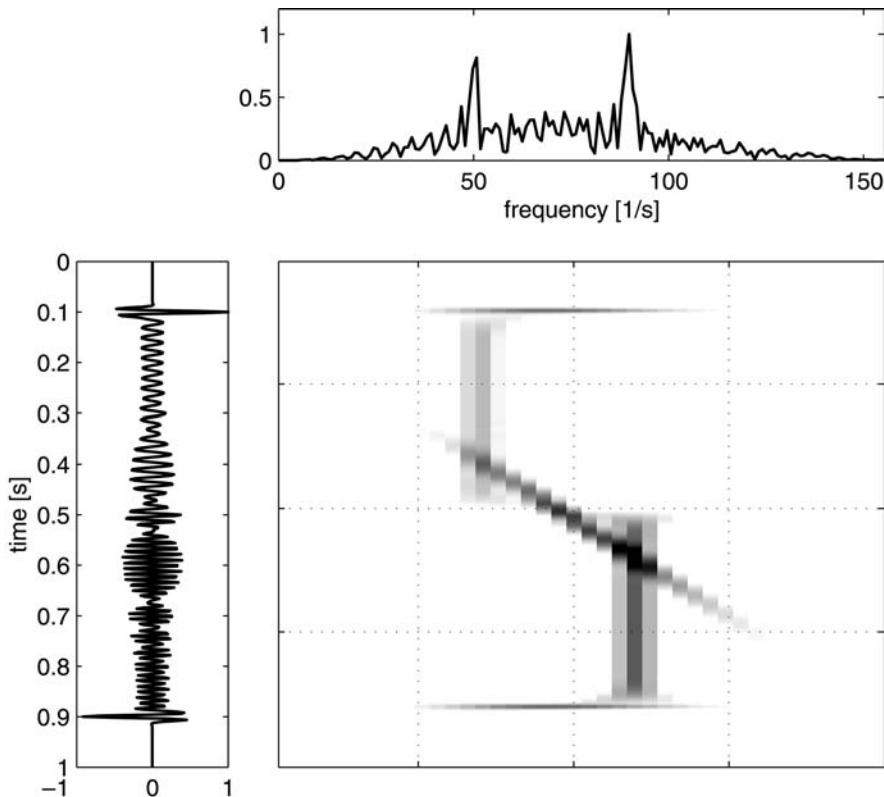


FIGURE 8.2

Synthetic model of a seismic sequence with building blocks: (a) layer of unit impedance, (b) harmonic impedance fluctuation (a cosine) and (c) component with increasing impedance as a function of depth (a chirp), with (d) sum of the three components and (e) seismic reflection response.

resulting from reflections at rock interfaces that cannot be resolved individually. We can subdivide the seismic signal into seismic sequences on the basis of its distinctive TF patterns, which is clearly much easier than subdividing based on the time signal itself.

Figure 8.2 illustrates the composition of a generic seismic TF sequence. We have constructed an impedance model of the subsurface that consists of three types of components. The first component is a rectangular boxcar function — Figure 8.2(a)— of unit impedance. This homogeneous impedance is perturbed with a purely harmonic impedance variation (a cosine, Figure 8.2[b]). We have introduced a frequency change at $t = 0.5$ sec. The third component is a perturbation with a linearly increasing frequency (a linear chirp, Figure 8.2[c]). The overall, superposition impedance model is shown in Figure 8.2(d). The simplest model for the seismic response $u(t)$ to this impedance variation is the reflectivity of the impedance function $r(t)$ convolved with a seismic source signal $s(t)$:

**FIGURE 8.3**

Idealized TF representation of the seismic response to the sequence of Figure 8.2. The time signal is shown to the left of the TFR and the power spectrum above. The TF representation has been obtained by summing the Wigner distributions of the seismic responses to the sequence components of Figure 8.2(a) to (c).

$$u(t) = \int_{-\infty}^{+\infty} r(t') s(t-t') dt' \quad (8.1)$$

The seismic response of the sequence model of Figure 8.2(d) is shown in Figure 8.2(e).

Figure 8.3 shows an idealized TF representation of the reflectivity function $u(t)$ of Figure 8.2(e). Next to the time and frequency axes of the TF representation, we plot the time signal and the Fourier power spectrum. This idealized TF representation was created by adding the Wigner distributions of the signal components. The components of the impedance model can be easily discerned in the TF image. The top and the base of the sequence give rise to an impulse-line TF pattern, like the sea-bottom reflector in Figure 8.1. The harmonic and chirping perturbations of the impedance can also be easily located in the TF plane.

This simple example demonstrates the value of TF analysis over time-domain analysis for seismic sequences. However, the TF representation of the measured seismic data in [Figure 8.1](#) shows that seismic reflection signals are clearly multicomponent signals. Moreover, in the preceding example we synthesized an ideal TF representation by adding the TF representations of the components. For measured data we cannot separate the components beforehand, which makes choosing the proper TF representation an important issue.

8.3 Time–Frequency Representations for Seismic Signal Analysis

If we weight the data $u(t)$ at each time t with a window function $w(t)$, we obtain the modified signal [1]:

$$u_t(\tau) = u(\tau) w(\tau - t) \quad (8.2)$$

The sliding window Fourier transform (FT) is then given by:

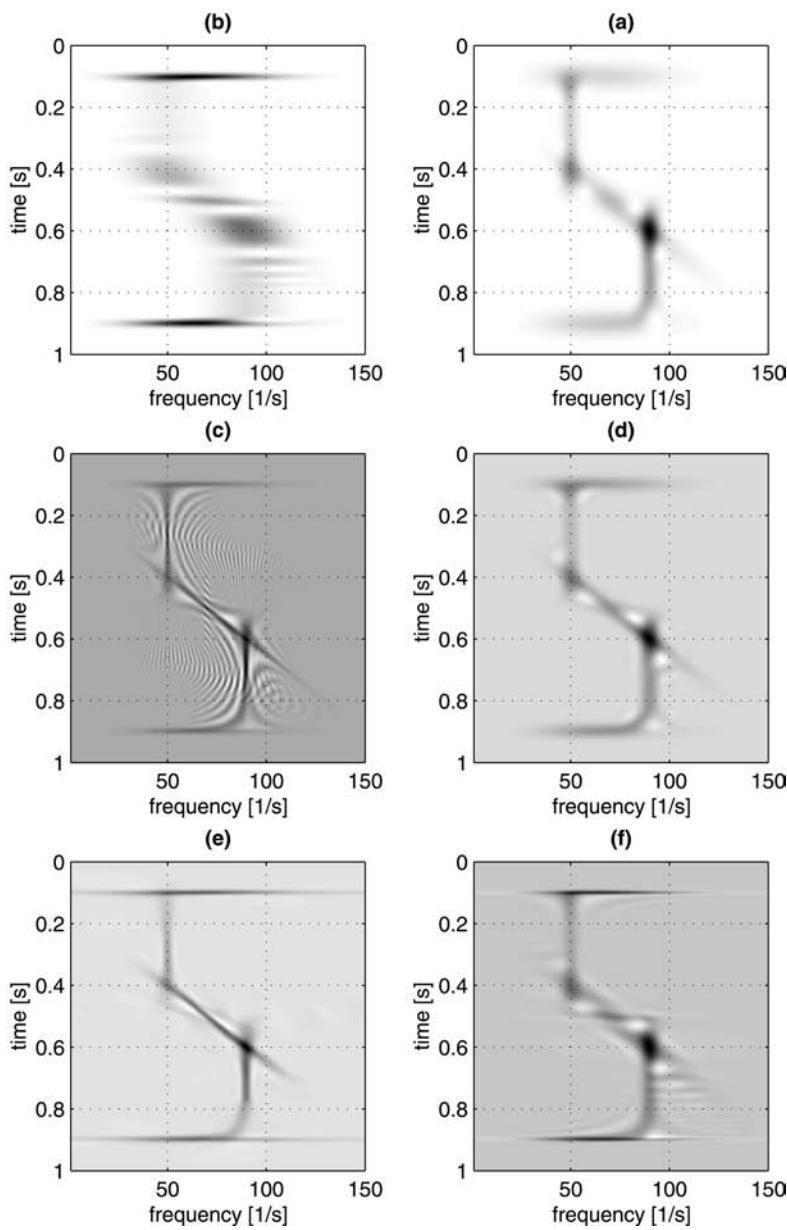
$$\hat{u}_t(f) = \int_{-\infty}^{+\infty} u_t(\tau) e^{-j2\pi f\tau} d\tau \quad (8.3)$$

Repeating this procedure for each time t and taking the squared modulus of $\hat{u}_t(f)$, we obtain the spectrogram TF representation:

$$S(t, f) = |\hat{u}_t(f)|^2 \quad (8.4)$$

The localization of the spectrogram is strongly influenced by the choice of window function $w(t)$. The windowed signal $u_t(\tau)$ and the local spectrum are an FT pair. Consequently, their breadth in time and frequency are linked by the uncertainty principle. [Figure 8.4](#) illustrates this property of the spectrogram. We analyze the synthetic seismic signal of [Figure 8.2\(e\)](#) with both a short time window ([Figure 8.4\[a\]](#)) and a long window ([Figure 8.4\[b\]](#)).^{*} In the short-window spectrogram we observe that the signal components are well localized in time but poorly localized in frequency. If we increase the window size, frequency localization is improved, but time localization deteriorates. This makes the spectrogram less suited for seismic data analysis, because we require accurate localization of both the sequence boundaries (time localization) and the changing patterns within each sequence (frequency localization).

^{*}Note that all the TF representations that we show in this chapter have been computed using the complex-valued analytical signal corresponding to the real-valued signal obtained in the seismic measurement. There are several motivations for using the analytical signal. For instance, its complex valuedness allows the decomposition in amplitude and phase. Also, the absence of negative frequencies in its Fourier spectrum has certain advantages with respect to both computational efficiency and interpretation of the TF analysis results [1].

**FIGURE 8.4**

TF representations of the seismic response to the sequence model of Figure 8.2: (a) spectrogram with a short window, (b) spectrogram with a long window, (c) Wigner distribution, (d) smoothed pseudo-Wigner distribution, (e) AOK representation and (f) cone-kernel representation.

An effective way to minimize window effects is to match the analysis window to the signal. For certain types of signals, excellent results can be achieved by using the reversed signal as an analysis window, which yields the Wigner distribution:

$$W(t, f) = \int_{-\infty}^{+\infty} u(t + \tau/2) u^*(t - \tau/2) e^{-j2\pi f\tau} d\tau \quad (8.5)$$

where the asterisk denotes complex conjugation.

[Figure 8.4\(c\)](#) displays the Wigner distribution of the synthetic seismic trace. The energy localization in the TF plane has improved considerably. We can clearly observe a much sharper TF localization compared with the spectrogram. However, oscillating ridges now appear between signal components at (t, f) locations where no energy is expected. These so-called *cross terms* result from the quadratic nature of the Wigner distribution. Cross term interference complicates the interpretation of the Wigner distribution of measured seismic signals.

Cross terms can be largely suppressed by smoothing the Wigner distribution over time and frequency to obtain a new representation:

$$P(t, f) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Psi(t', f') W(t - t', f - f') dt' df' \quad (8.6)$$

The smoothing function $\Psi(t, f)$ is called the *kernel* of the representation. Each particular kernel results in a different TF representation, whose properties can be derived directly from the kernel. [Figure 8.4\(d\)](#) shows the smoothed pseudo-Wigner distribution, which uses a separable kernel $\Psi(t, f) = h(t) G(f)$. In this example we have used Gaussian-shaped time and frequency windows for $h(t)$ and $G(f)$. We see that the cross terms are largely suppressed without overly compromising the time and frequency resolution.

The class of TF representations obtained by smoothing the Wigner distribution is called *Cohen's class* [1]. The spectrogram in Equation (8.4) is also a member of this class, with a kernel that is the Wigner distribution of the analysis window $w(t)$. By taking the 2-D Fourier transform of Equation 8.6, the convolution becomes a weighting operation on the characteristic function of TF representation:

$$M(\nu, \tau) = \Psi(\nu, \tau) A(\nu, \tau) \quad (8.7)$$

Here $M(\nu, \tau)$ is the characteristic function (FT of $P(t, f)$), $\Psi(\nu, \tau)$ is the FT of the kernel, and $A(\nu, \tau)$ is the ambiguity function (FT of the Wigner distribution).

In search of a smoothing kernel adapted to the properties of seismic signals, we aim to maximize the suppression of the cross terms while compromising other desirable properties of the TF representation as little as possible. For instance, there is a trade-off between the degree of cross-term suppression (maximal smoothing) and preserving resolution (minimal smoothing).

The cone–kernel:

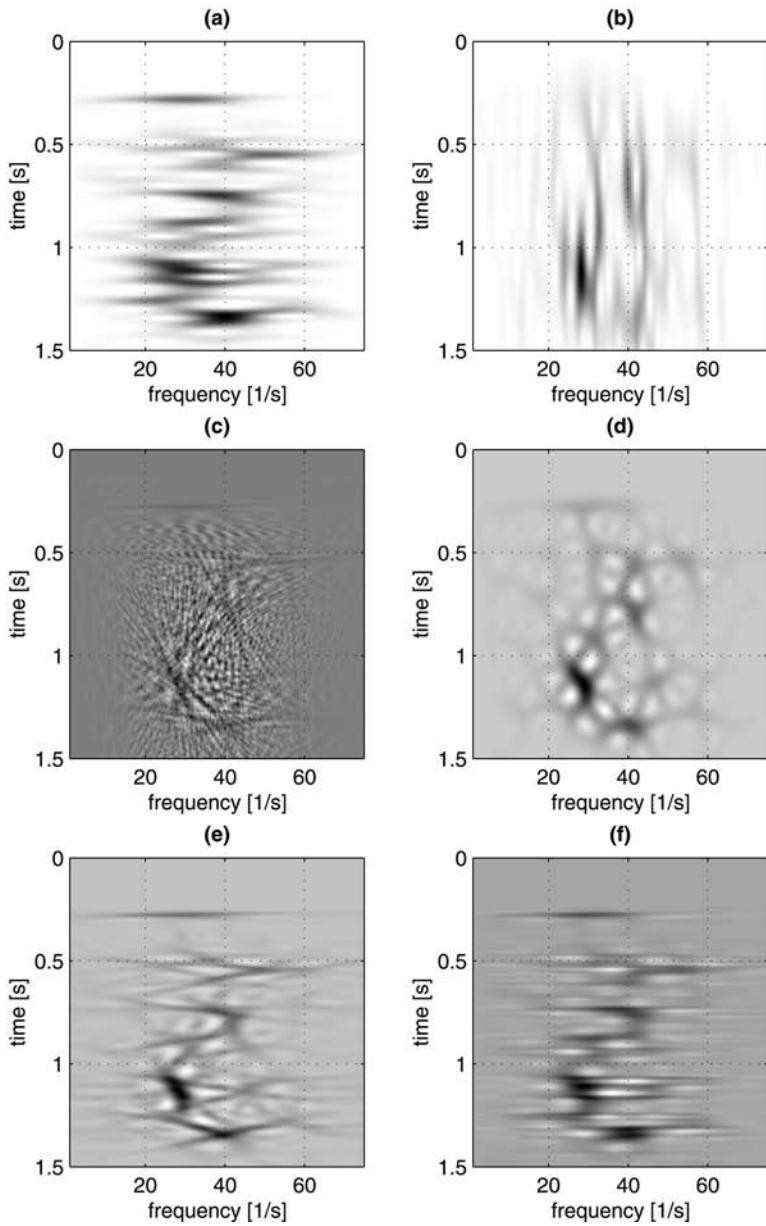
$$\Psi(\nu, \tau) = w(\tau) |\tau| \frac{\sin(2\pi a\nu\tau)}{2\pi a\nu\tau} \quad (8.8)$$

appears well suited to seismic data analysis [2]. Here $w(\tau)$ is a real and symmetrical window function, and a is a dimensionless constant — typically $a = 1/2$ is taken. The cone-kernel TF representation is depicted in [Figure 8.4\(f\)](#). Because the kernel is zero along the frequency-shift (ν) axis, impulse-like signals are suppressed, and components that are parallel to the time axis (harmonics) are emphasized. Thus, this kernel does not smooth impulses, as does the spectrogram, but suppresses them outright. This results in a very effective representation for accurately localizing frequency transitions. However, the chirping component in the signal cannot be very well distinguished in the TF image, because it is not parallel to the time axis.

The location and amplitude of the cross terms depend on the characteristics of the signal. Consequently, this trade-off can only be optimized by adapting the shape of the smoothing kernel to the characteristics of the signal under analysis. The adaptive optimum kernel (AOK) TF representation [3] of the seismic trace is shown in [Figure 8.4\(e\)](#). The AOK representation adapts the shape of the kernel at each time instant t to maximally concentrate energy in the TF plane while suppressing cross terms. In the AOK representation of [Figure 8.4\(e\)](#), the cross terms have been largely suppressed, with only minimal resolution deterioration compared with the Wigner distribution. Note that both the short duration signal components (pulses) and long duration components (tones) are localized equally well. This ability to sharply localize energy in both time and frequency makes the AOK very well suited for local spectral analysis of seismic signals.

[Figure 8.5](#) shows the TF representations of the seismic field data of [Figure 8.1](#). [Figure 8.5\(a\)](#) and [\(b\)](#) shows the result of spectrogram analyses with a short and long window. It would be very hard to make a meaningful subdivision of the seismic signal on the basis of the spectrograms. The short-window spectrogram shows only the individual reflections, whereas the long-window spectrogram resolves only the individual frequencies and fails to localize the sequence transitions. The Wigner distribution of [Figure 8.5\(c\)](#) is difficult to interpret, because it is overwhelmed by cross terms. Smoothing the Wigner distribution results in an improvement, but it is still difficult to distinguish individual components. The cone kernel and AOK representations appear to perform the best in terms of separating and localizing seismic sequence patterns in the data — [Figure 8.5\(e\)](#) and [\(f\)](#). The cone-kernel TF representation appears well suited for separating the seismic signal into distinct sequences because of the accurate localization of frequency transitions. The AOK representation gives a more balanced picture of the TF properties of the seismic signal. Strong impulse-like signals, such as the sea-bottom reflection, are more clearly imaged; and also components with an increasing or decreasing frequency as a function of time can be observed.

Up to now we have focused on the TF representation of a single seismic trace, for example, a column from the seismic section of [Figure 8.1\(b\)](#). However, seismic field data consist of large numbers of traces. It would be impossible for a seismic interpreter to inspect the TF representation of each individual trace from a seismic survey. For this reason, it is desirable to “summarize” a TF representation with a limited number of features. Localization of features in time (and space) is crucial, because there is a one-to-one mapping between the time coordinate and the depth

**FIGURE 8.5**

TF representations of the leftmost seismic trace of the seismic cross section of Figure 8.1(b).
 (a) Spectrogram with a short window, (b) spectrogram with a long window, (c) pseudo-Wigner distribution,
 (d) smoothed pseudo-Wigner distribution, (e) AOK representation TFR and (f) cone-kernel representation.

below the surface of the earth. Feature extraction from the TF representation is the subject of the next sections.

8.4 Seismic Attribute Extraction

Seismic attributes aid the interpretation of seismic images by elucidating salient signal characteristics. Traditionally, complex-trace analysis via the Hilbert transform has been used for attribute extraction [4]. The standard complex-trace attributes are the instantaneous amplitude (reflection strength), phase and frequency. Complex-trace attributes are still widely used in seismic signal analysis.

In complex-trace notation, the seismic trace is an analytical signal given by:

$$u^a(t) = a(t) e^{j\phi(t)} \quad (8.9)$$

where $a(t)$ is the instantaneous amplitude and $\phi(t)$ is the instantaneous phase. The complex-trace instantaneous frequency $g(t)$ is the derivative of the instantaneous phase:

$$g(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} \quad (8.10)$$

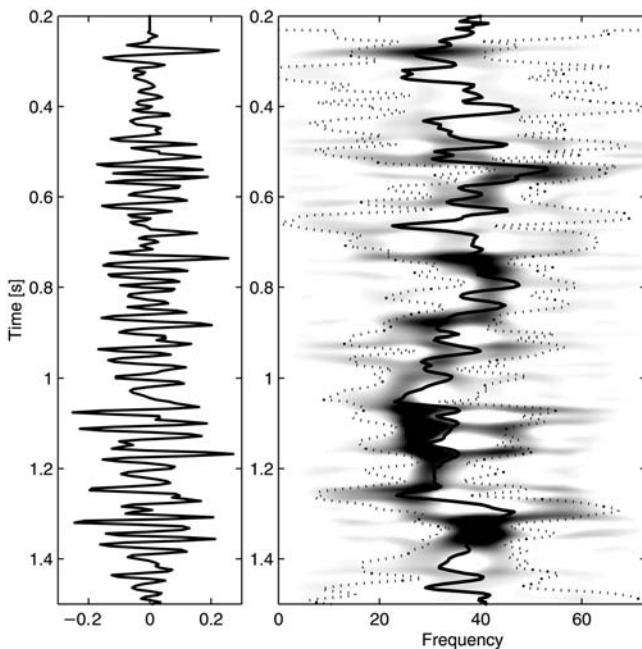
The instantaneous frequency can be estimated directly using a discretized version of Equation (8.10). Unfortunately, this estimate is highly susceptible to noise. A more robust estimate of the instantaneous frequency can be formed by considering the relationship between the Wigner distribution and the instantaneous frequency [1,5]:

$$g(t) = \frac{\int_{-\infty}^{+\infty} f W(t, f) df}{E(t)} \quad (8.11)$$

where $E(t) = \int_{-\infty}^{+\infty} W(t, f) df$ is the time marginal. Hence, the mean frequency is the first moment of the TF representation, normalized by the energy. Direct utilization of Equation (8.11) leads to high-variance estimates for noisy signals. However, bias can be traded against variance by replacing the Wigner distribution with a smoothed TF representation $P(t, f)$ from Cohen's class in Equation (8.6).

[Figure 8.6](#) illustrates the use of a TF representation for seismic attribute extraction. Figure 8.6 shows a seismic trace and its cone–kernel TF representation. In TF we plot the mean frequency $g(t)$ (solid line) and the standard deviation around the average (dotted line) as a function of time. The standard deviation corresponds to an instantaneous bandwidth. Seismic attribute images can be generated by plotting these time curves in the same manner as the amplitude data in [Figure 8.1\(b\)](#).

[Figure 8.7](#) plots the mean frequency $g(t)$ of each vertical trace in the seismic data of Figure 8.1(b) as extracted from four different TF representations: spectrogram, pseudosmoothed Wigner distribution, AOK representation and cone–kernel representation. Dark values correspond to high attribute values; a linear scaling has been used.

**FIGURE 8.6**

(Left) Seismic trace and (right) cone–kernel TF representation with (solid line) instantaneous frequency and (dashed lines) instantaneous bandwidth attributes.

The image of the spectrogram mean frequency in Figure 8.7(a) is somewhat blurred, making it difficult to localize the sequence transitions and relate the attribute features to the amplitude data. The smoothed pseudo-Wigner result of Figure 8.7(b) clearly shows the major transitions and gives detailed and very well localized frequency information. In the AOK result, we observe more gradual changes in frequency content, similar to the spectrogram attribute image of Figure 8.7(a). However, in the AOK result the temporal transitions are also clearly visible. This combination of eliciting both more gradual and instantaneous features makes the AOK representation the best of these four representations for attribute extraction. However, for large data volumes, the computational effort involved in the AOK could become too burdensome for practical application. The cone-kernel result of Figure 8.7(d) shows that this TFR is a computationally more attractive alternative to adaptive TF analysis.

Besides the mean frequency, a great number of other signal attributes can be extracted from a TF representation. A straightforward extension is to calculate the higher order moments. For instance, we can extract an instantaneous bandwidth attribute by calculating the variance around the mean frequency $g(t)$:

$$B(t)^2 = \frac{\int_{-\infty}^{+\infty} (f - g(t))^2 P(t, f) df}{E(t)} \quad (8.12)$$

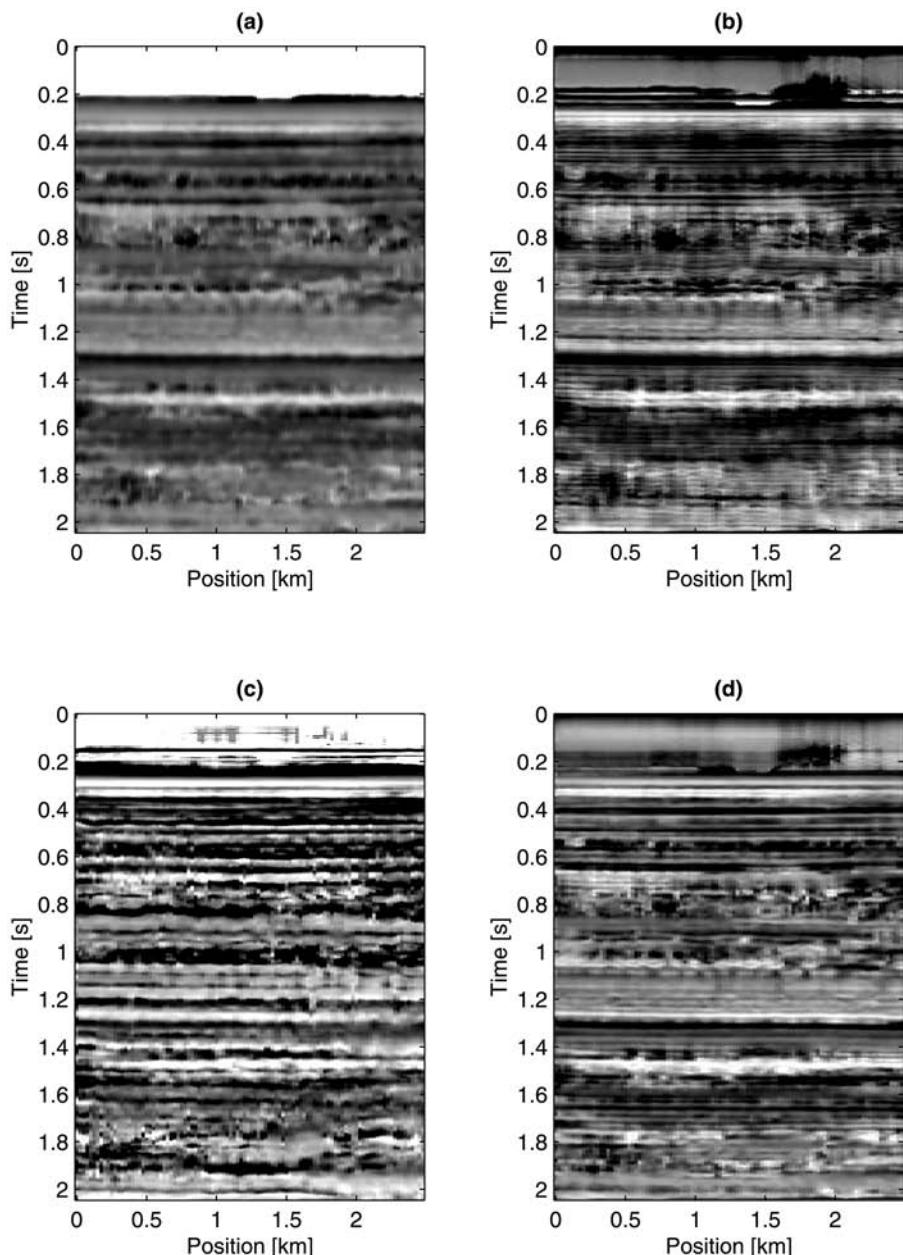


FIGURE 8.7

Seismic attribute images of the seismic cross section of Figure 8.1(b). The mean frequency attribute extracted using the (a) spectrogram, (b) smoothed pseudo-Wigner distribution, (c) AOK representation and (d) cone-kernel representation. Dark values represent high frequencies.

Other statistics measures for characterizing density functions include the skew and kurtosis [6,7]. Skew measures the deviation of the density function from a normal (Gaussian) distribution and is computed from the third moment around the mean:

$$s(t) = \frac{\int_{-\infty}^{+\infty} (f - g(t))^3 P(t, f) df}{B(t)^3 E(t)} \quad (8.13)$$

A positive skew signifies an asymmetrical distribution with a tail extending out toward positive frequencies. A normal distribution has zero skew.

Kurtosis measures the peakedness of the distribution and is related to the fourth-order moment [7]:

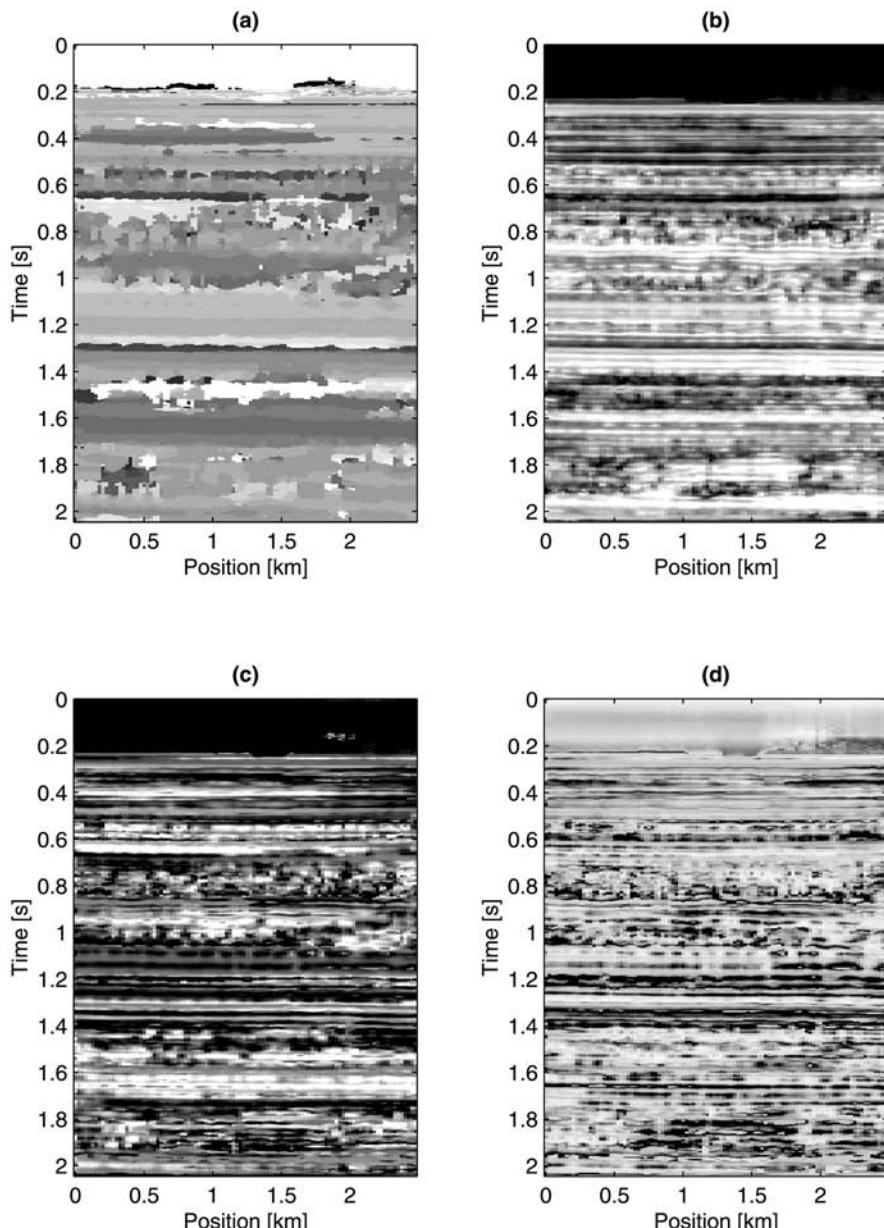
$$k(t) = \frac{\int_{-\infty}^{+\infty} (f - g(t))^4 P(t, f) df}{B(t)^4 E(t)} - 3 \quad (8.14)$$

where the term -3 makes the value 0 for a normal distribution. Density functions with a positive kurtosis have a more sharply peaked shape than a Gaussian. Negative kurtosis signifies a distribution that is flatter than a normal distribution.

[Figure 8.8](#) shows four additional attribute images for the seismic section of [Figure 8.1\(b\)](#) extracted from the cone-kernel representation. Figure 8.8(a) plots the peak frequency as a function of time. The peak frequency is very well suited for segmenting the section into a limited number of large-scale seismic sequences. However, the peak frequency image lacks the fine detail that can be observed in the mean frequency images of [Figure 8.7](#). The sequences boundaries we would define on the basis of the peak frequency image coincide with changes in instantaneous bandwidth values in Figure 8.8(b). Some boundaries (at $t = 0.65$ sec and $t = 1.25$ sec) that are less pronounced in the mean frequency image are clearly visible in the instantaneous bandwidth image.

The instantaneous skew is less sensitive to the transitions in TF patterns, but brings forward the characteristics of the reflectivity between the sequence boundaries. Several intervals can be characterized by their distinctive skew values. For example, between $t = 1$ sec and $t = 1.2$ sec we have relatively high skew, whereas around $t = 1.6$ sec we observe low skew values. High values of kurtosis coincide with more chaotic and noisy reflection patterns in the amplitude data. The laterally continuous reflections around $t = 0.4$ sec result in consistently low kurtosis values. Around $t = 0.8$ sec, there is a region with more chaotic reflection patterns. Here, we observe also relatively high kurtosis values. Right above $t = 1.4$ sec we see two reflectors with a high kurtosis value. In the AOK TF representation of Figure 8.1(a), we also observe that the energy is highly concentrated around a single frequency. This is typical for the seismic response to a “thin bed”—a rock layer that is thin with respect to the seismic wavelength ($< \lambda/4$), with a strong reflectivity at the top and the base of the layer [8]. The individual reflections from the top and base of the layer cannot be resolved, but the frequency of the reflection “tunes” to the frequency that corresponds to the (time) thickness of the layer.

These examples demonstrate the utility of TF representations for seismic interpretation and attribute extraction. However, the computational effort of calculating

**FIGURE 8.8**

Seismic attribute images of the seismic cross section of Figure 8.1(b), computed using the cone-kernel representation. (a) Peak frequency, (b) instantaneous bandwidth, (c) instantaneous skew and (d) instantaneous kurtosis.

a representation for each seismic trace may be a serious impediment to using this method on large seismic data volumes. A typical seismic data volume consists of several million traces, each of approximately 1000 samples. Fortunately, for TF representations with a fixed kernel, it is in many cases possible to calculate the attribute moments directly without calculating the highly redundant TF representation itself. In the next section we present an efficient method for robustly calculating the signal attributes using the principles of hybrid linear and quadratic TF analysis.

8.5 Hybrid linear and Quadratic Time–Frequency Seismic Attributes

In this section, we present a computationally efficient method for estimating moment-based TF attributes that avoids the calculation of a quadratic TF representation altogether. Our approach exploits the TF distribution series concept developed by Qian, Morris and Chen [9–11] and requires only a sparse linear signal decomposition. We present a general procedure for calculating a range of attributes, but focus on instantaneous frequency and bandwidth estimation for concreteness.

Qian, Morris and Chen proposed a method for TF analysis that uses both a linear TF basis representation (the Gabor transform) and a quadratic TF representation (the Wigner distribution) to generate a signal-adaptive nonlinear TF representation. The Gabor transform decomposes a signal $u(t)$ in terms of logons [12] or TF atoms [12]:

$$u(t) = \sum_{(m,n)} c_{m,n} \phi_{m,n}(t), \quad (m,n) \in \mathbb{Z}^2 \quad (8.15)$$

$$c_{m,n} = \left\langle u, \tilde{\phi}_{m,n} \right\rangle = \int_{-\infty}^{\infty} u(t) \tilde{\phi}_{m,n}^*(t) dt \quad (8.16)$$

The synthesis atoms are generated by TF shifting a prototype atom $w(t)$ by discrete step sizes T and F :

$$\phi_{m,n}(t) = e^{j2\pi m F t} w(t - nT) \quad (8.17)$$

The $\tilde{\phi}(t)$ are dual atoms derived from $\phi(t)$ [9, 12]. A natural choice for the prototype atom $w(t)$ is the Gaussian, because it has optimal concentration and localization in TF and a strictly positive Wigner distribution [1]. For Equation (8.15) to be a stable representation with Gaussian atoms requires a mild oversampling in the TF plane ($TF < 1$).

The auto-Wigner distribution $W_u(t, f)$ of a signal $u(t)$ is defined in Equation (8.5); the cross-Wigner distribution between two signals $u(t)$ and $v(t)$ is defined as:

$$W_{uv}(t, f) = \int_{-\infty}^{\infty} u(t + \tau/2) v^*(t - \tau/2) e^{-j2\pi f \tau} d\tau \quad (8.18)$$

By inserting Equation (8.16) in Equation (8.5) and using Equation (8.18), the auto-Wigner distribution can be decomposed as:

$$\begin{aligned} W_u(t, f) &= \sum_{(m,n)} |c_{m,n}|^2 W_{\phi_{m,n}}(t, f) \\ &\quad + \sum_{(m,n) \neq (m',n')} c_{m,n} c_{m',n'}^* W_{\phi_{m,n}, \phi_{m',n'}}(t, f) \end{aligned} \quad (8.19)$$

This expression identifies two distinct contributions to the Wigner distribution. The first summation in Equation (8.19) corresponds to a linear sum of the (strictly positive) auto-Wigner distributions of the TF atoms; it provides an approximate description of the signal TF behavior without capturing many details. The second summation involves all the cross-Wigner distributions between different atoms in the Gabor decomposition. The cross-Wigner distributions in this summation may take on both positive and negative values.

The key observation is this [9, 10]: the cross-Wigner distributions between closely spaced atoms ((m, n) close to (m', n')) generally refine the TF representation of the signal, whereas the cross-Wigner distributions between distant atoms ((m, n) far from (m', n')) generate global interference terms that hamper interpretation.

Qian, Morris and Chen generate a TF distribution series — which we term a *hybrid TF representation* [11] — by retaining all the auto-Wigner distribution terms in the first summation of Equation (8.19), but only those cross-Wigner distributions arising from closely spaced atoms. Due to the elliptical or circular symmetry of the Gaussian atom, the Euclidean distance metric is the most natural measure of atom separation; the l_1 metric (Manhattan distance) $d[(m, n), (m', n')] = |m - m'| + |n - n'|$ is a good approximation with a lower computational complexity. Given a threshold distance δ , second summation terms are included in the final representation only if the distance between the interacting atoms is less than δ . The hybrid method thus produces a δ -parameterized class of TF representations:

$$\begin{aligned} \widetilde{W}_u^{(\delta)}(t, f) &= \sum_{(m,n)} |c_{m,n}|^2 W_{\phi_{m,n}}(t, f) \\ &\quad + \sum_{0 < d[(m,n), (m',n')] < \delta} c_{m,n} c_{m',n'}^* W_{\phi_{m,n}, \phi_{m',n'}}(t, f) \end{aligned} \quad (8.20)$$

The parameter δ controls the trade-off between component resolution and cross-component interference. By a suitable selection of δ , high resolution TF representations are generated; moreover, they can be determined at a much lower computational expense than signal-adaptive quadratic TF representations offering similar performance (e.g., the AOK). The auto- and cross-Wigner distributions in Equation (8.20) are signal independent and hence can be analytically computed and stored in memory. The computational demands of the Gabor transform are also much lower than those of quadratic TF representations. By replacing the Gabor decomposition with a wavelet transform, we can generate high-resolution, low-interference time-scale representations [11].

We now detail how the hybrid TF representation approach can be adopted to robustly estimate the instantaneous frequency and instantaneous bandwidth. Our approach requires the much lower computational expense of determining a linear, mildly oversampled Gabor representation of the signal.

By substituting Equation (8.19), we can rewrite the numerator of Equation (8.11) as:

$$\sum_{(m,n,m',n')} c_{m,n} c_{m',n'}^* \int_{-\infty}^{\infty} f W_{\phi_{m,n}\phi_{m',n'}}(t, f) df \quad (8.21)$$

When we use a Gaussian of variance σ^2 for the Gabor synthesis atom, closed-form expressions can be developed for the integrals in this summation.[†] By applying the same expansion to the denominator, we can express Equation (8.11) as:

$$g(t) = \frac{\sum_{m,n,m',n'} c_{m,n} c_{m',n'}^* V_{m,n,m',n'}(t)}{\sum_{m,n,m',n'} c_{m,n} c_{m',n'}^* A_{m,n,m',n'}(t)} \quad (8.22)$$

with:

$$A_{m,n,m',n'}(t) = \frac{1}{2\pi} \exp[-j2\pi(n-n')Ft] \exp\left[-\frac{(t-mT)^2 + (t-m'T)^2}{2\sigma^2}\right] \quad (8.23)$$

and:

$$V_{m,n,m',n'}(t) = \left[\frac{-j(m-m')T}{4\pi\sigma^2} - \frac{(n+n')F}{2} \right] A_{m,n,m',n'}(t) \quad (8.24)$$

If all (m, n, m', n') combinations are included in the estimate, then this formula is equivalent to the (noise-sensitive, high-variance) Wigner distribution-based estimate in Equation (8.11). By adopting the hybrid TF representation approach (retaining only a subset of the terms when $(m, n) \neq (m', n')$), we can reduce the variance with the introduction of some bias. The performance is of a level similar to that obtained with the AOK TF representation; however, the computational expense is much reduced, because the $A_{m,n,m',n'}(t)$ and $V_{m,n,m',n'}(t)$ are signal independent and can be pre-computed and stored in memory. The primary cost of the algorithm is then that of computing the (barely oversampled) Gabor transform coefficients $c_{m,n}$.

The second-order moment of the hybrid TF representation provides a similarly computationally efficient estimate for the instantaneous bandwidth:

$$B^2(t) = \frac{\sum_{m,n,m',n'} c_{m,n} c_{m',n'}^* D_{m,n,m',n'}(t)}{\sum_{m,n,m',n'} c_{m,n} c_{m',n'}^* A_{m,n,m',n'}(t)} - g^2(t) \quad (8.25)$$

[†]Similar analytical expressions are possible for other windows, including, for example the square, triangle, raised cosine, and Jones multiplexed windows.

with:

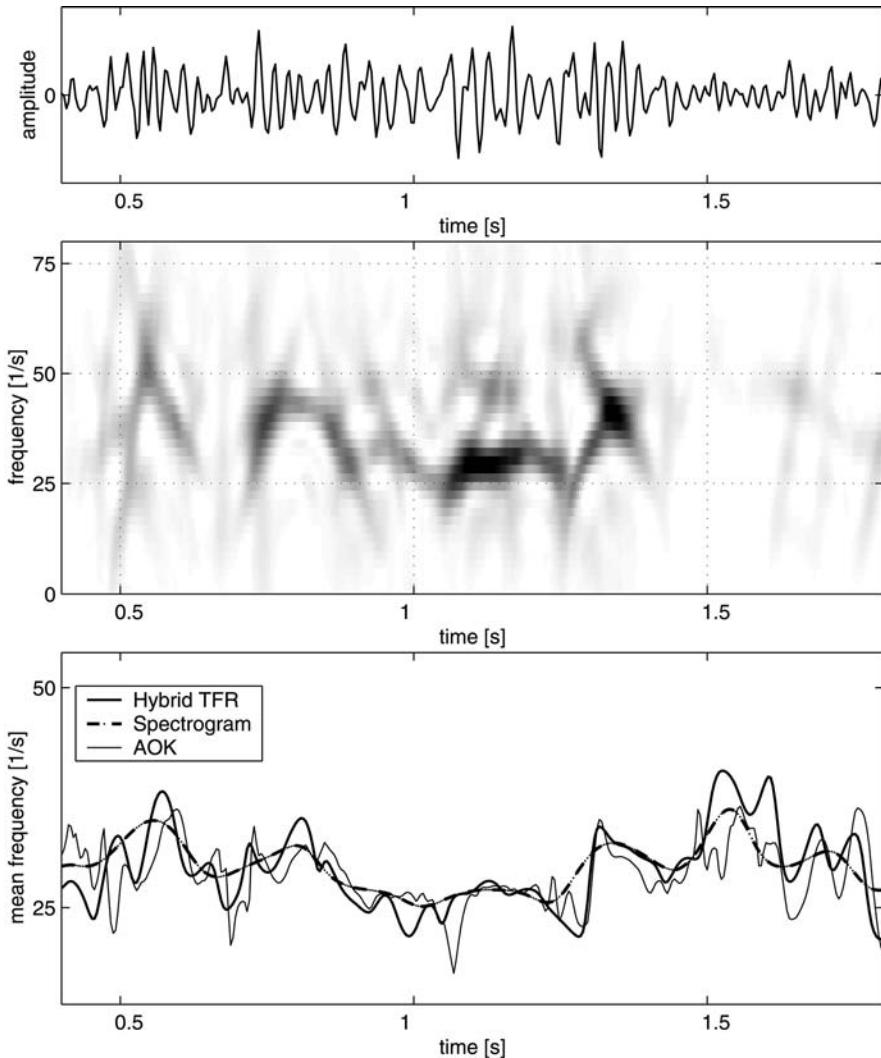
$$\begin{aligned}
 D_{m,n,m',n'}(t) = & \frac{1}{(4\pi)^2} \exp \left[-\frac{(t-mT)^2 + (t-m'T)^2}{2\sigma^2} \right] \exp[-j2\pi(n-n')Ft] \\
 & \times \left[-\frac{(T(m-m'))^2}{\sigma^4} + \frac{2+4\pi jTF(m-m')(n+n')}{\sigma^2} \right. \\
 & \left. + 4\pi^2 (F(n+n'))^2 \right]
 \end{aligned} \tag{8.26}$$

We can provide no strict guidelines concerning the choice of the distance threshold δ . If the hybrid estimate is formed without cross terms (i.e., with $\delta = 0$), then attribute performance is very similar to the spectrogram. We have observed that Manhattan distance thresholds δ between 1 and 3 provide good results in our experiments. Setting $\delta = 1$ results in more smoothing and hence more robust estimates in very noisy environments. The choice of $\delta = 3$ reduces the bias of estimates and is appropriate when there is little noise and a single dominant signal component.

It is straightforward to extend this method to other higher order moments such as instantaneous skew, Equation (8.13), and instantaneous kurtosis, Equation (8.14), in a similar fashion, but the third- and fourth-order terms generate more complicated expressions that we do not provide here.

[Figure 8.9](#) compares the instantaneous frequency estimates from three different TF representations for the seismic trace of [Figure 8.1\(b\)](#). The fast hybrid instantaneous frequency estimate is very close to that obtained using the much more computationally expensive AOK representation. The time resolution of the spectrogram instantaneous frequency estimate is significantly poorer than the other two methods.

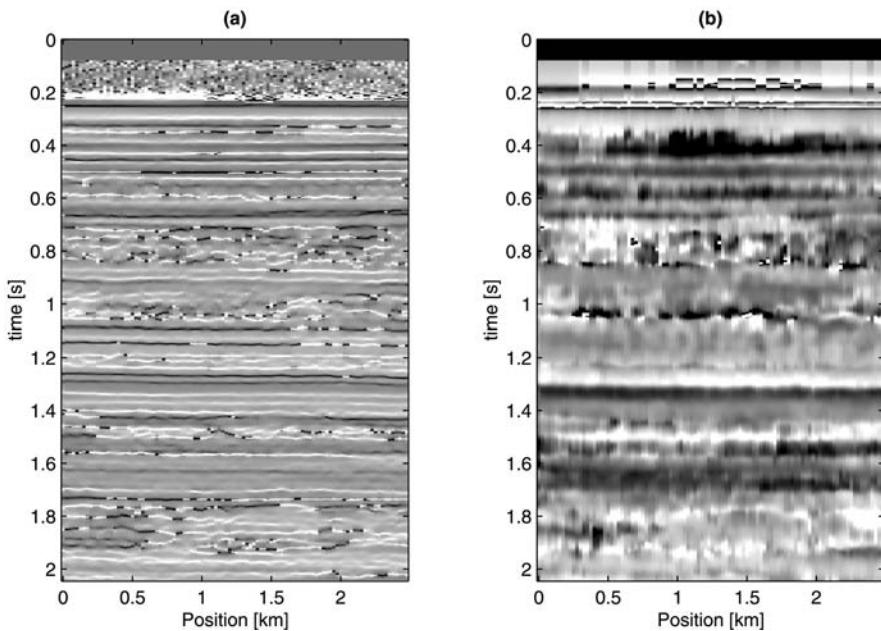
[Figure 8.10 \(a\)](#) and (b) shows complex-trace instantaneous frequency of Equation (8.10) vs. the hybrid instantaneous frequency estimate of Equation (8.22) for the seismic section of Figure 8.1(b). The wild fluctuations indicate that the complex-trace instantaneous frequency estimate is very noise sensitive. The sharp peaks in the complex-trace instantaneous frequency estimate obscure the frequency trends of interest for a seismic sequence analysis. By comparing Figure 8.10 with [Figure 8.8](#), we see that the detail and robustness of the hybrid instantaneous frequency estimate is again close to that of the AOK-based estimate. For visual interpretation, the hybrid instantaneous frequency provides a greatly enhanced image as compared with the complex-trace instantaneous frequency or the spectrogram average. This estimation algorithm potentiates the TF-based interpretation of large seismic data volumes, which up to now have been severely inhibited by the large computational effort involved in calculating TF representations.

**FIGURE 8.9**

(Top) seismic trace, (middle) AOK TF representation, and (bottom) mean frequency of the seismic trace.

8.6 Three-dimensional Seismic Attribute Extraction

Currently, 3-D seismic data acquisition is the standard in the seismic industry. The added value that the 3-D view of geology provides is far greater than the extra cost of data acquisition. However, a consequence of the more detailed subsurface picture

**FIGURE 8.10**

(a) Hilbert transform instantaneous frequency and (b) Hybrid linear and quadratic TF instantaneous frequency.

is that the volumes of seismic data oil company geologists need to sift through and to interpret have grown tremendously. To facilitate the interpretation of these large 3-D data volumes, a number of novel techniques for seismic feature extraction and visualization are under development. Extensions of seismic attribute techniques to higher dimensions is a logical step for improving the extraction and processing of geologic information from 3-D seismic data volumes.

The success of seismic volume attribute extraction has demonstrated that the extra information in 3-D data can only be fully exploited if the feature extraction is also fully 3-D, that is, we must move from 1-D time trace attributes to 3-D volume attributes. Volume-dip [13–15] and coherence [16, 17] attributes illustrate how rapid changes in the characteristics of the 3-D signal generally indicate a geologic discontinuity. This one-to-one relation of signal characteristics with geologic features underlies the added value of seismic attribute images. In this section we introduce a seismic volume attribute extraction technique that extends the 1-D attributes discussed in the previous section.

The starting point of our volume attribute extraction procedure is an analysis of the seismic data with a 3-D local Radon transformation. The Radon transformation is used in a wide range of seismic processing and analysis applications. The main reason for its popularity is that it decomposes the data into its plane-wave components [18–20]. For seismic interpretation purposes, a local Radon transformation would be preferred;

often it is not sufficient to know which plane-wave components are present in the data, but also where they occur. Local Radon transformations have been defined and applied for different purposes [21–23]. These local Radon representations are all based on applying the classical “slant stack” to windowed portions of the data. However, in the same way the spectrogram is a member of Cohen’s class of TF representations, the sliding-window Radon spectrum is just one choice from an infinite number of local Radon transforms. The Wigner–Radon representation we propose as a candidate for the local slant-stack power spectrum emerges naturally from the definition of the local wavenumber-frequency power spectrum.

Let $u(\mathbf{x}, t)$ denote the 3-D seismic signal, with $\mathbf{x} = \{x, y\}$ the spatial coordinate vector and t the time coordinate. The Radon transform of $u(\mathbf{x}, t)$ is defined as:

$$\check{u}(\mathbf{p}, \tau) = \int_{-\infty}^{+\infty} u(\mathbf{x}, \tau + \mathbf{p} \cdot \mathbf{x}) d\mathbf{x} \quad (8.27)$$

The vector $\mathbf{p} = \{p_x, p_y\}$ represents the slopes of the signal with respect to the (x, y) plane in the x and y directions, and τ is the intercept with the time axis. In this formulation the Radon transformation essentially sums the signal along lines with slopes \mathbf{p} and intercept τ . This interpretation explains the term *slant stack* that is widely used in exploration seismology for this quantity.

The Radon transform of a signal is closely related to its FT. The FT of Equation (8.27) with respect to intercept time τ is given by:

$$\check{u}(\mathbf{p}, f) = \int_{-\infty}^{+\infty} \hat{u}(\mathbf{x}, f) e^{j2\pi f \mathbf{p} \cdot \mathbf{x}} d\mathbf{x} \quad f > 0 \quad (8.28)$$

where $\hat{u}(\mathbf{x}, f)$ denotes the temporal FT of $u(\mathbf{x}, t)$. In all practical cases $u(t)$ is either real valued or a complex-valued analytical signal. In both cases we can restrict the analysis to positive temporal frequency f . Equation (8.28) is a 3-D FT of $u(\mathbf{x}, t)$. It follows that we can obtain the temporal frequency domain representation of the Radon transform from the 3-D FT $\tilde{u}(\mathbf{k}, f)$ of the signal $u(\mathbf{x}, t)$:

$$\check{u}(\mathbf{p}, f) = \tilde{u}(f\mathbf{p}, f) = \tilde{u}(\mathbf{k}, f) \quad f > 0 \quad (8.29)$$

where $\mathbf{k} = \{k_x, k_y\} = f\mathbf{p}$ is the spatial frequency vector. This relation leads us to the definition of a local Radon power spectrum that is based on the multidimensional Wigner distribution of the signal u .

The 3-D instantaneous autocorrelation function is defined as:

$$R(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = u(\mathbf{x} + \boldsymbol{\xi}/2, t + \tau/2) u^*(\mathbf{x} - \boldsymbol{\xi}/2, t - \tau/2) \quad (8.30)$$

where $\boldsymbol{\xi} = \{\xi_x, \xi_y\}$ and τ are the space and time-shift variables, respectively. The 3-D FT over the space shifts $\boldsymbol{\xi}$ and time shift τ yields the Wigner distribution:

$$W(\mathbf{x}, t; \mathbf{k}, f) = \int \int_{-\infty}^{+\infty} R(\mathbf{x}, t; \boldsymbol{\xi}, \tau) e^{j2\pi(\mathbf{k}\boldsymbol{\xi} - f\tau)} d\boldsymbol{\xi} d\tau \quad (8.31)$$

We can now obtain a local Radon power spectrum $Q(\mathbf{x}, t; \mathbf{p}, \tau)$ by invoking the relation in Equation 8.29 followed by an inverse FT:

$$Q(\mathbf{x}, t; \mathbf{p}, \tau) = \int_0^\infty W(\mathbf{x}, t; \mathbf{p}, f) e^{j2\pi f\tau} df \quad (8.32)$$

For seismic attribute extraction we do not perform this inverse FT but instead take the average of the local Radon power spectrum with frequency:

$$Q(\mathbf{x}, t; \mathbf{p}, \tau = 0) = \int_0^\infty W(\mathbf{x}, t; \mathbf{p}, f) df \quad (8.33)$$

The two basic seismic attributes that we extract from this spectrum are the local average slopes \bar{p}_x and \bar{p}_y of the signal. The average slope in the x direction is given by:

$$p_x(\mathbf{x}, t) = \frac{\iint p_x Q(\mathbf{x}, t; \mathbf{p}) d\mathbf{p}}{\iint Q(\mathbf{x}, t; \mathbf{p}) d\mathbf{p}} \quad (8.34)$$

In seismic interpretation, the parameters dip and azimuth are commonly used to describe local geometric features. The volume dip is the modulus of the local slopes, given by:

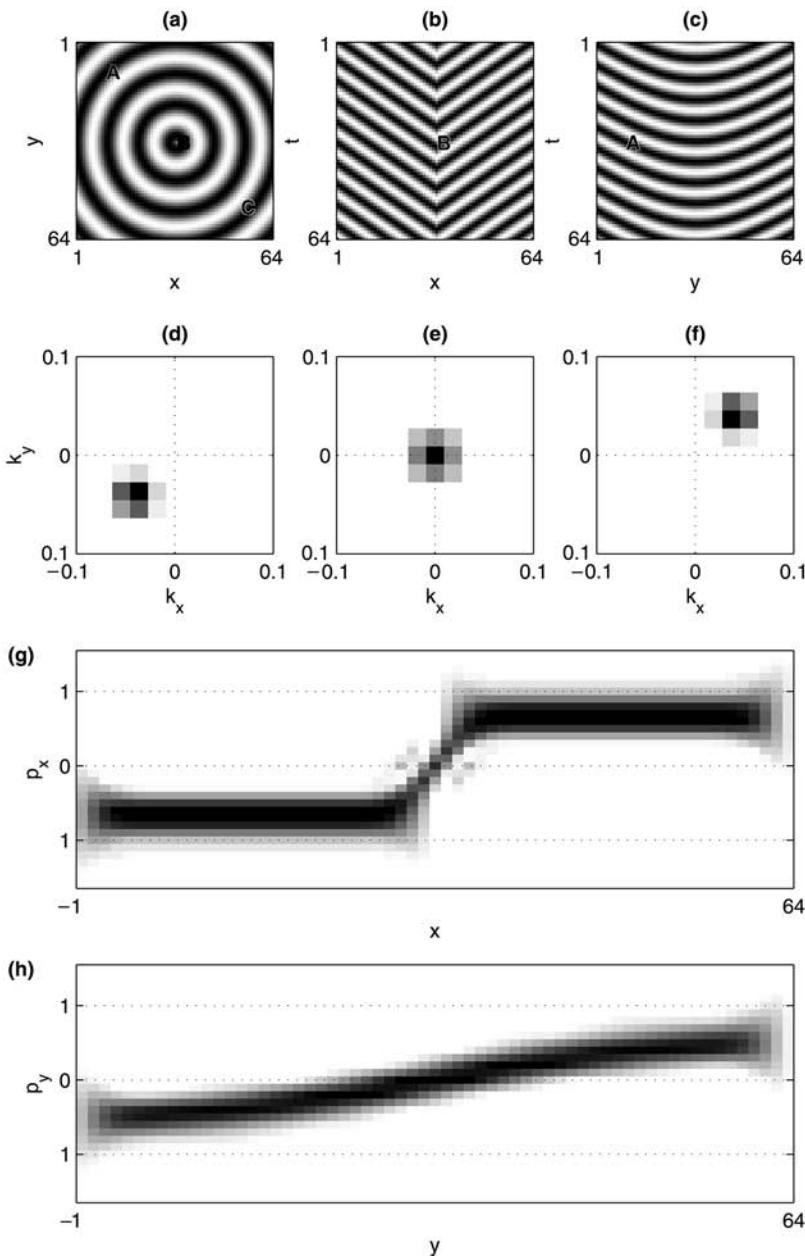
$$|\bar{p}| = (\bar{p}_x^2 + \bar{p}_y^2)^{1/2} \quad (8.35)$$

The azimuth angle is the direction of the signal in the (horizontal) (x, y) plane. The local average azimuth α is given by:

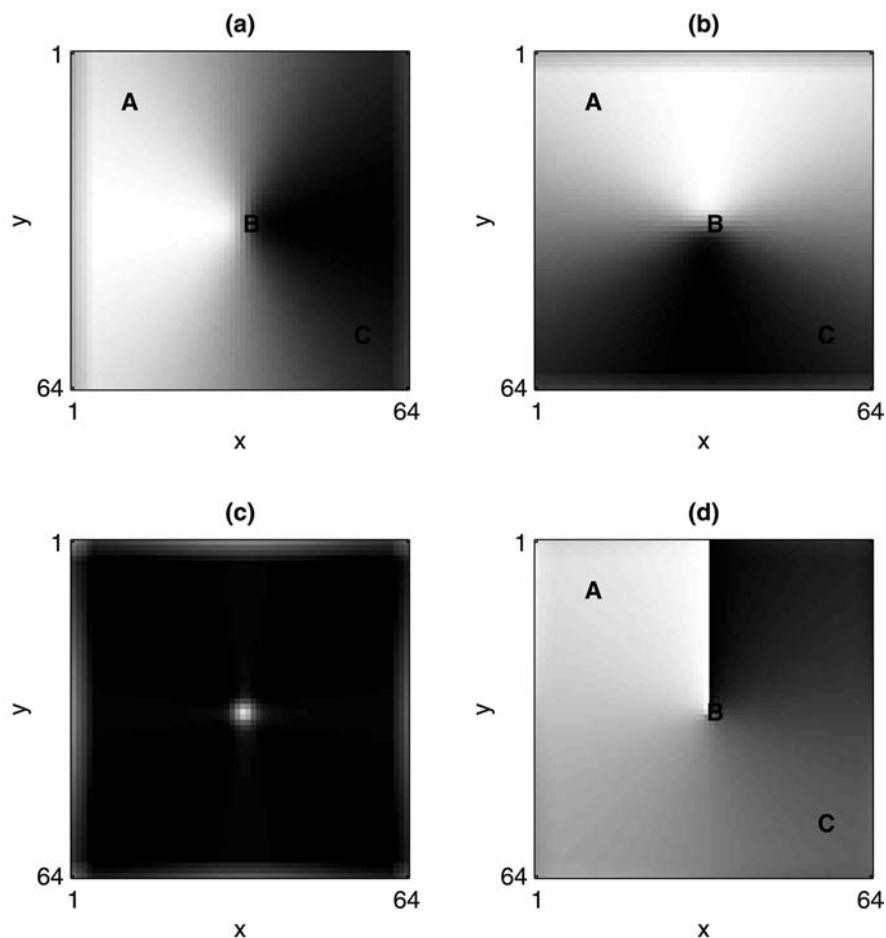
$$\alpha = \tan^{-1} \left(\frac{\bar{p}_x}{\bar{p}_y} \right) \quad (8.36)$$

With the relation of the local Radon power spectrum to the Wigner distribution, we have all the tools and techniques that have been developed for Cohen's class TF representations at our disposal. In a practical implementation, we first compute a multidimensional TF representation and then obtain the local Radon representation by interpolating the wavenumber-frequency spectrum onto a grid of (\mathbf{p}, f) coordinates. This has the advantage that efficient fast Fourier transform (FFT)-based computational algorithms for TF analysis can be used. A disadvantage is that the interpolation step introduces errors; using more sophisticated interpolation methods greatly increases the computational load. Alternatively, the local Radon spectrum can be computed by directly performing a slant stack of the local autocorrelation function. For the 3-D sliding-window slant stack, a fast recursive algorithm has been developed, which is based on a recursive implementation of the discrete Fourier transform (DFT) [13, 24].

[Figures 8.11](#) and [8.12](#) illustrate the volume attribute extraction procedure. The data volume is a cube with dimensions $(n_x, n_y, n_t) = (64, 64, 64)$. The signal is

**FIGURE 8.11**

(a) Time slice of 3-D data cube; (b) (x, t) cross section through 3-D data volume; (c) (y, t) cross section through 3-D data volume; (d) to (f) constant temporal frequency slices through the 3-D pseudo-Wigner distribution at locations denoted A, B and C in slices (a) to (c), respectively; (g) (x, p_x) spectrum along a line of constant y through B; (h) (y, p_y) spectrum along a line of constant x through A.

**FIGURE 8.12**

Volume attributes of the time slice of Figure 8.11(a). (a) Local average slope in the x direction, (b) local average slope in the y direction, (c) modulus slope (volume dip) and (d) azimuth angle.

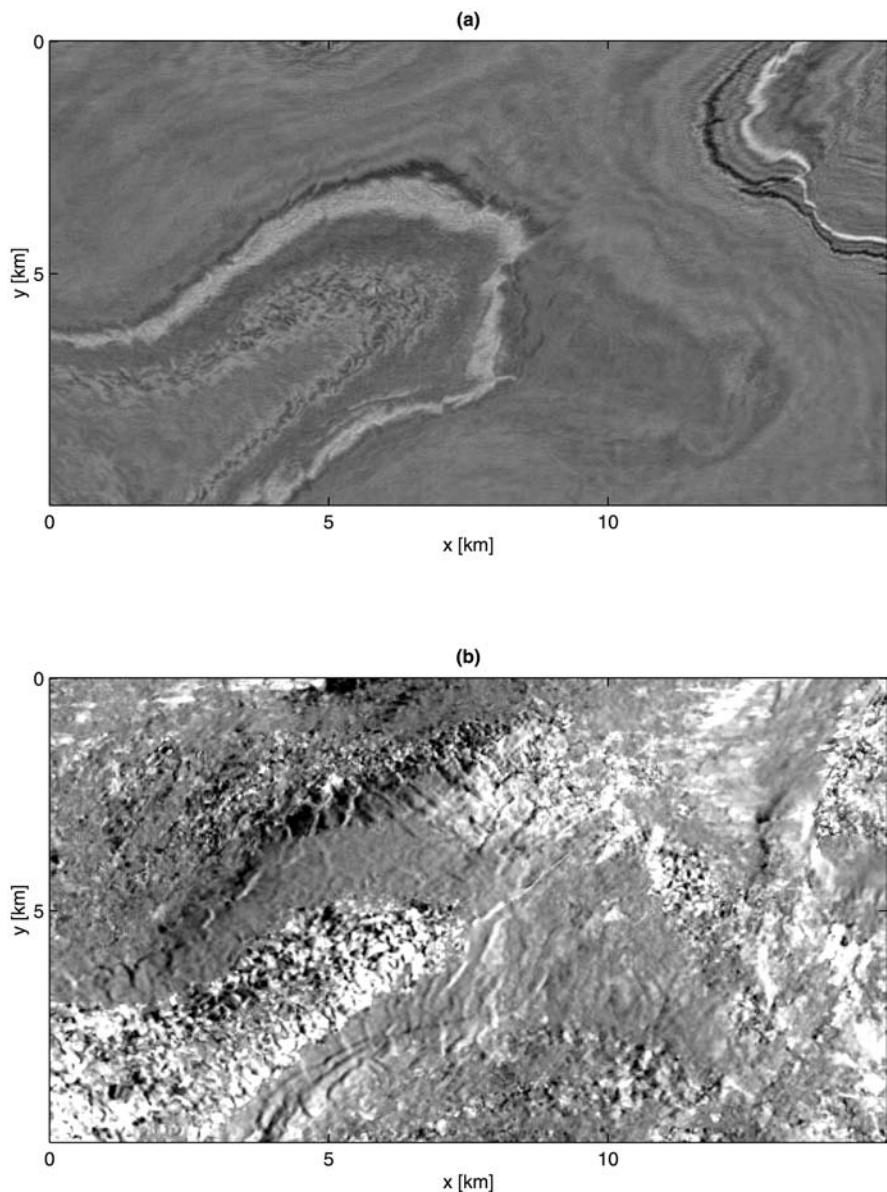
3-D complex exponential with a constant temporal frequency f_0 . In the horizontal direction, the spatial frequency is constant in the radial direction $r = (x^2 + y^2)^{1/2}$. Figure 8.11(a) to (c) shows three slices through this volume. In a seismic data volume, Figure 8.11(a) would represent a time slice, whereas Figure 8.11(b) and (c) would be vertical cross sections or seismic lines. Figure 8.11(c) to (e) shows three slices through the 3-D pseudo-Wigner distribution of the data cube at the locations that are denoted A, B, and C in the data slices. The slices are taken at frequency f_0 and

show the localization of the spatial frequencies in the horizontal, (x, y) plane. The next step in the attribute extraction procedure is the computation of the local Radon representation $Q(\mathbf{x}, t; \mathbf{p})$. Figure 8.11(f) shows the p_x marginal of this local Radon spectrum along a line of constant y and t , cutting through point B in the data cube. The (x, p_x) spectrum is obtained by summing the Radon spectrum over p_y . Note that the slope p_y of the signal does not change along this line through the data volume. The image shows how the energy is concentrated at a single constant slope and changes sign at the center of the data cube (B). Figure 8.11(g) shows the (y, p_y) spectrum of the signal along a line of constant x and t that cuts through point A. The slope of the signal varies gradually with spatial location y .

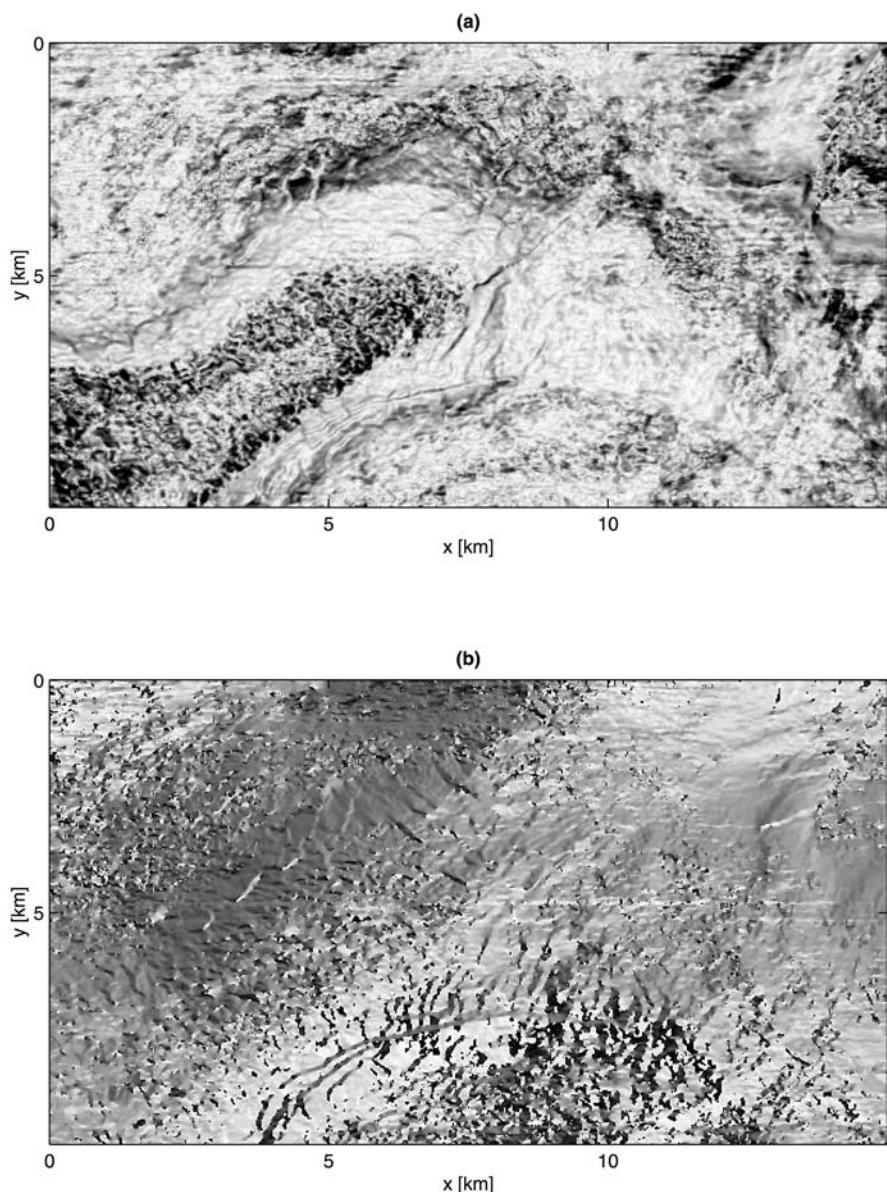
Figure 8.12 shows the attributes extracted from the local Radon spectrum. The average local slopes in the x and y directions are shown in Figure 8.12(a) and (b). The modulus of the slope (volume dip) and azimuth angle are shown in Figure 8.12(c) and (d). The volume dip is constant, except at the center of the time slice (B), where the sign change occurs. The azimuth indicates the direction of the signal with respect to “north” and is constant along radial lines originating from the center of the slice.

Results from a real seismic data volume are given in Figures. 8.13 and 8.14. Figure 8.13(a) shows a time slice through a seismic data cube from the North Sea. The spatial sample spacing is 25 m in x and y and the time sampling is 4 msec. The average slope attributes were extracted from a 3-D smoothed pseudo-Wigner distribution. The size of the autocorrelation cube was $(m_x, m_y, m_t) = (7, 7, 7)$, corresponding to 150 m in the horizontal directions and 24 msec in the vertical direction. The smoothing kernel was three samples in all directions. Figure 8.13(b) shows the slope of the signal in the y direction. White values denote negative slopes, indicating that the signal locally slopes upward in time and consequently toward lower depths. In this way the slope attribute image gives an impression of the 3-D geometric properties of the seismic reflections, and thus the geologic structure. Figure 8.14(a) shows the average volume dip, which gives an indication of the steepness of the seismic reflections. Signal discontinuities, such as faults or chaotic reflections, are indicated by high volume dip values (shown in black). The azimuth attribute is shown in Figure 8.14(b) and highlights changes in the direction of the slope, giving a view of the general structure but also indicating small-scale faulting.

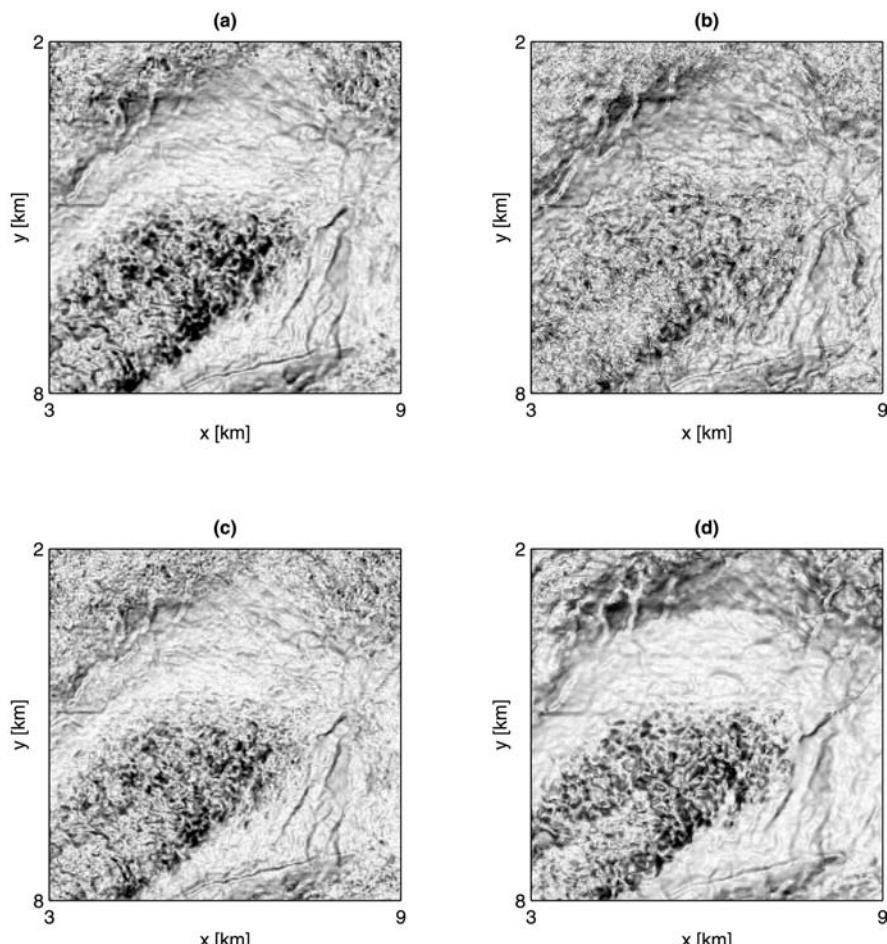
In Figure 8.15 we compare the results obtained with four different Local Radon representations. The volume dip map of a part of the time slice of Figure 8.13(a) is shown. Figure 15(a) shows the result that is obtained using the 3-D spectrogram. The pseudo-Wigner distribution and smoothed pseudo-Wigner distribution results are shown in Figure 8.15(b) and (d), respectively. Figure 15(c) shows the attribute image obtained using a 3-D version of the modified Wigner distribution of Stanković et al. [25]. The smoothed pseudo-Wigner distribution and spectrogram results are very similar. However, in the smoothed pseudo-Wigner result transitions in reflector slope can be more accurately localized. The pseudo-Wigner distribution is more susceptible to noise, which masks several geologically interesting features in the attribute image. The 3-D modified Wigner distribution results in an attribute image that mixes the properties of the pseudo-Wigner distribution and spectrogram images.

**FIGURE 8.13**

(a) Seismic time slice and (b) local average slope in the x direction. Downward slopes are black.

**FIGURE 8.14**

(a) Volume dip and (b) azimuth attributes of the time slice of [Figure 8.13\(a\)](#).

**FIGURE 8.15**

Comparison of attribute images, obtained using different local Radon spectra, of the time slice of Figure 8.13(a). Volume dip images obtained using (a) spectrogram, (b) pseudo-Wigner distribution, (c) modified Wigner distribution and (d) smoothed pseudo-Wigner distribution.

8.7 Conclusions

In this chapter, we give an overview of a number of applications of TF representations in seismic data processing, from the analysis of seismic sequences to efficient attribute extraction to 3-D attributes for volumetric data. TF representations provide a new physically relevant domain from which to extract information on the earth's subsurface. As the accuracy of TF attributes increases and their computational cost

decreases, they are poised to play a central role in the seismic interpretation process. Finally, we note that it is also possible to define instantaneous attributes in terms of time-scale representations such as the wavelet transform; promising results in this direction have been developed [26–28].

References

- [1] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, New York, 1995.
- [2] Y. Zhao, L. Atlas, and R. Marks, The use of cone-shaped kernels for generalized time frequency representations of nonstationary signals, *IEEE Trans. Signal Process.*, 38, 1084–1091, 1990.
- [3] D. Jones and R. Baraniuk, An adaptive optimal kernel time-frequency representation, *IEEE Trans. Signal Process.*, 43(10), 2361–2371, 1995.
- [4] T. Taner, F. Koehler and R. Sheriff, Complex seismic trace analysis, *Geophysics*, 44, 1041–1063, 1979.
- [5] P. Flandrin, *Time-frequency/Time-scale Analysis*, Academic Press, San Diego, CA, 1999.
- [6] D. Childers, *Probability and Random Processes*, Irwin, 1997.
- [7] P. Loughlin and K. Davidson, Instantaneous kurtosis, *IEEE Signal Process. Lett.*, 7, 156–159, June 2000.
- [8] J. Robertson and H. Nogami, Complex seismic trace analysis of thin beds, *Geophysics*, 49(4), 344–352, 1984.
- [9] S. Qian and J.M. Morris, Wigner distribution decomposition and cross-terms deleted representation, *Signal Process.*, 27, 125–144, May 1992.
- [10] S. Qian and D. Chen, Decomposition of the Wigner distribution and time-frequency distribution series, *IEEE Trans. Signal Process.*, 42, 2836–2842, October 1994.
- [11] M. Pasquier, P. Gonçalvès and R. Baraniuk, Hybrid linear/bilinear time-scale analysis, *IEEE Trans. Signal Process.*, 47, 254–259, January 1999.
- [12] D. Gabor, Theory of communication, *J. IEE*, 93, 429–444, 1946.
- [13] P. Steeghs, J. Fokkema and G. Diephuis, Local Radon Power Spectra for 3-D Seismic Attribute Extraction, SEG 68th Annual International Meeting, Society of Exploration Geophysicists, 1998, pp. 645–648.
- [14] P. Steeghs, A Fast Algorithm for Dip and Azimuth Computations, in SEG 69th Annual International Meeting, Society of Exploration Geophysicists, 1999, pp. 1146–1149.

- [15] P. Steeghs, I. Overeem and S. Tigrek, Seismic volume attribute analysis of the Cenozoic succession in the L08 block (Southern North Sea), *Global Planetary Change*, 27, 245–262, 2000.
- [16] K. Marfurt, R. Kirlin S. Farmer and M. Bahorich, 3-D seismic attributes using a semblance-based coherency algorithm, *Geophysics*, 63, 1150–1176, 1998.
- [17] A. Gersztenkorn and K. Marfurt, Eigenstructure-based coherence computations as an aid in 3-D structural and stratigraphic mapping, *Geophysics*, 64, 1468–1479, 1999.
- [18] P. Stoffa, P. Buhl, J. Diebold and F. Wenzel, Direct mapping of seismic data to the domain of intercept time and ray parameter — a plane wave decomposition, *Geophysics*, 46, 255–267, 1981.
- [19] R. Phinney, K.R. Chowdhury and L. Frazer, Transformation and analysis of record sections, *J. Geophys. Res.*, 86, 359–377, 1981.
- [20] S. Treitel, P. Gutowski and D. Wagner, Plane wave decomposition of seismograms, *Geophysics*, 47, 175–1401, 1982.
- [21] G. McMechan and R. Ottolini, A direct observation of a $p - \tau$ curve in a slant stacked wavefield, *Bull. Seismol. Soc. Am.*, 70, 775–789, 1983.
- [22] B. Milkereit, Decomposition and inversion of seismic data — An instantaneous slowness approach, *Geophys. Prospecting*, 35, 875–894, 1987.
- [23] A. Copeland, G. Ravichandran and M. Trivedi, Localized Radon transform-based detection of ship wakes in SAR images, *IEEE Trans. Geosci. Remote Sensing*, 33, 33–45, 1995.
- [24] A. Papoulis, *Signal Analysis*, McGraw-Hill, New York, 1977.
- [25] S. Stanković, L. Stanković and Z. Usković, On the local frequency, group shift, and cross-terms in some multi-dimensional time-frequency distributions: a method for multi-dimensional time-frequency analysis, *IEEE Trans. Signal Process.*, 43, 1719–1724, 1995.
- [26] R.L.C. van Spaendonck and R. Baraniuk, Directional Scale Analysis for Seismic Interpretation, in *SEG 69th Annual International Meeting*, Society of Exploration Geophysicists, 1999, pp. 1844–1847.
- [27] R.L.C. van Spaendonck, F.C.A. Fernandes, M. Coates and C.S. Burrus, Non-Redundant, Directionally Selective, Complex Wavelets, in IEEE Proceedings of the International Conference on Image Processing Extended Abstracts II, IEEE, 2000, pp. 379–382.
- [28] F.C.A. Fernandes, R.L.C. van Spaendonck, M. Coates and C.S. Burrus, Directional complex-wavelet processing, in Proceedings of the SPIE 2000, 4119, 61–61, 2000.

Time–Frequency Methodology for Newborn Electroencephalographic Seizure Detection

Boualem Boashash and Mostefa Mesbah

Queensland University of Technology

CONTENTS

9.1	Introduction	339
9.2	Seizure Criteria and Time–Frequency Methodology	342
9.3	Seizure Patterns	353
9.4	Electroencephalogram Background Patterns	357
9.5	Time–Frequency Matched Detector	359
9.6	Discussion and Conclusions	366
	Acknowledgments	367
	References	367

9.1 Introduction

Techniques previously designed for electroencephalographic (EEG) seizure detection in the newborn have been relatively inefficient due to their incorrect assumption of local stationarity of the EEG. To overcome the problem raised by the proven nonstationarity of the EEG signal, current methods are extended to a time–frequency (TF) approach [8, 10]. This allows the analysis and characterization of the different newborn EEG patterns, the first step toward an automatic TF seizure detection and classification. An in-depth analysis of the previously proposed autocorrelation and spectrum seizure detection techniques identified the detection criteria that can be readily extended to the TF domain. We present the various patterns of observed TF seizure signals and relate them to current specialist knowledge of seizures. In particular, initial results indicate that a quasilinear instantaneous frequency (IF) can be used as a critical feature of the EEG seizure characteristics. These findings led to propose a TF-based seizure detector. This detector performs a two-dimensional (2D) correction between the EEG signal and a reference template selected as a model of the EEG seizure in TF domain.

9.1.1 Seizures and the electroencephalographic signal

Seizure may be the most frequent, and often the only, clinical sign of central nervous system dysfunction in the neonate [34]. Unlike adult seizure, the clinical manifestations of newborn seizure are subtle and hence require the constant attention of a medical specialist for diagnosis.

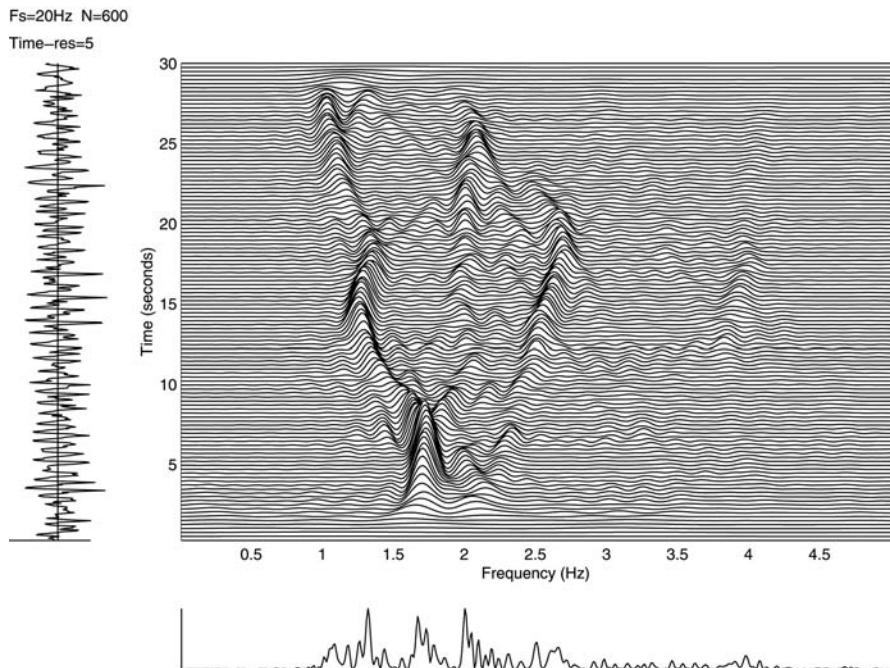
EEG has become a successful means of seizure detection in adults. This usually involves identifying sharp repetitive waveforms that indicate the onset of seizure. In adults, these EEG seizures are easily recognizable against a low amplitude random background characteristic of normal brain activity. The problem of detecting seizure in newborn babies, however, is complicated by a number of factors [28]. First, healthy newborn EEG signals representing normal brain activity often contain patterns such as spurious waveforms and sharp spikes. These characteristics, which would otherwise be detected as seizure in adults, are simply the result of extra electrical activity produced by the immature brain as it continues to form. Seizures, however, still usually appear in the EEG data as repetitive waveforms and the problem lies in discerning the healthy spikes from those formed from seizures. Second, visual symptoms of seizure, such as muscle spasms, rapid eye movement and drooling, are much subtler in newborns and may be easily missed. These visual indicators are also natural movements common to all newborn babies. Third, physical activity of babies in the intensive care environment is often subdued by medication to prevent injuries caused by unpredictable movements. This also reduces the chance of seizure detection using visual signs altogether. For these reasons, the EEG may be the only alternative for seizure detection in newborn.

9.1.2 Nonstationary and multicomponent characteristics of the electroencephalogram

There are three published methods for EEG seizure detection in newborns. The technique of Roessgen, Zoubir and Boashash [28] is a parametric approach based on a nonlinear estimation of 11 model parameters for detection. The two other methods are nonparametric. The spectrum technique introduced by Gotman et al. [16] uses frequency analysis to determine the changes in the dominant peak of the spectrum of short epochs of EEG data. The autocorrelation technique introduced by Liu [20] performs analysis in the time domain using short epochs of EEG data.

All three techniques are based on the assumption that the EEG signals are locally stationary. However, a closer examination of these signals clearly shows that they exhibit significant nonstationary and multicomponent features [9, 24] ([Figure 9.1](#)). Therefore, the basic assumption on which these three methods are based is essentially invalid and at best only an approximation.

To take this fact into account, we extend the autocorrelation and spectrum techniques to the nonstationary case by introducing both time variation and frequency variation parameters in the basic equations on which these techniques are based. This extension leads to a TF formulation in the $(t - f)$ domain or a time-varying autocorrelation formulation in the time-lag domain $(t - \tau)$.

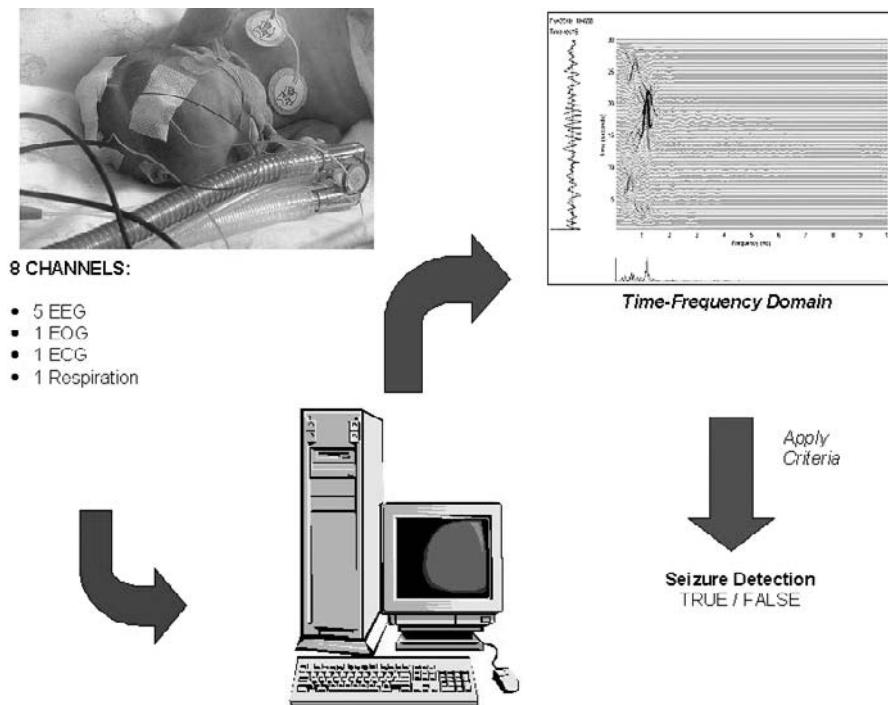
**FIGURE 9.1**

Time–frequency (TF) representations of newborn EEG seizure signal using the B distribution. This EEG clearly shows the multicomponent and nonstationary nature of the seizure.

Having taken into account the nonstationary aspect of the EEG signals, we still need to deal with the other aspect, namely, the multicomponent behavior. This requires a selection of an appropriate time-frequency distribution (TFD) that is capable of handling multicomponent signals. Once this has been done, a calibration process is undertaken to translate the seizure criteria from the spectrum and autocorrelation domains to the TF domain. This initially involves the reproduction of the seizure detection criteria used by the two previously mentioned methods and map them in a joint TF domain. Features in the joint $t - f$ domain characterizing seizure are then identified and a detection process based on those features is constructed and tested. Once we have a better understanding of the processes involved for each method and determined which components are most suited to a TF representation, we combine them in a TF-based detection process as illustrated in [Figure 9.2](#).

9.1.3 Electroencephalographic acquisition

The electrical signals produced in the brain are monitored in a noninvasive manner by measuring variations in potential on the scalp. The EEG signals are measured using strategically placed small electrodes on the scalp. One electrode, usually at the base of the skull, acts as a reference (ground) signal, and various channels of data are

**FIGURE 9.2**

TF-based seizure detection process.

created by measuring the voltage differences between neighboring electrodes. Due to the size of most newborn babies' heads, only 5 channels of EEG have been recorded in each session using the 10 to 20 International System of Electrode Placement. The sampling rate used for recording the EEG data was 256 Hz. For the sake of artifact detection, three auxiliary signals representing electro-oculogram (EOG), electrocardiogram (ECG) and respiration were also recorded. Data used in this study were collected at the Royal Women's Hospital Perinatal Intensive Care Unit in Brisbane, Australia.

9.2 Seizure Criteria and Time–Frequency Methodology

9.2.1 Basic definitions and relationships

The autocorrelation function of a random nonstationary process $x(t)$ may be written as:

$$R_x(t, \tau) = E[x(t)x(t + \tau)] \quad (9.1)$$

where τ represents a time difference, or lag and E is the expectation operator. If the process is at least wide-sense stationary (WSS), the autocorrelation function becomes simply a function of τ , that is:

$$R_x(t, \tau) = R_x(\tau) \quad (9.2)$$

For the case of stationary signals, Wiener–Khintchine theorem states that the power spectral density of $x(t)$ can be obtained by simply taking the Fourier transform (FT) of Equation (9.2). In other words:

$$S_x(f) = \underset{\tau \rightarrow f}{FT} \{R_x(\tau)\} \quad (9.3)$$

where $R_x(\tau)$ is defined in Equation (9.2). This equation is valid only for stationary signals. For the case of nonstationary signals, Equation (9.3) gives averaged values over the variations of the signal spectral contents and thus significant information may be lost.

For nonstationary signals, the Wiener–Khintchine theorem can be extended so that the time-varying power spectral density $S_x(t, f)$ of a signal $x(t)$ is also related to the time-varying autocorrelation function $R_x(t, \tau)$ of the same signal by the following FT relation:

$$S_x(t, f) = \underset{\tau \rightarrow f}{FT} \{R_x(t, \tau)\} \quad (9.4)$$

or more explicitly:

$$S_x(t, f) = \int_{-\infty}^{\infty} R_x(t, \tau) e^{-j2\pi f \tau} d\tau \quad (9.5)$$

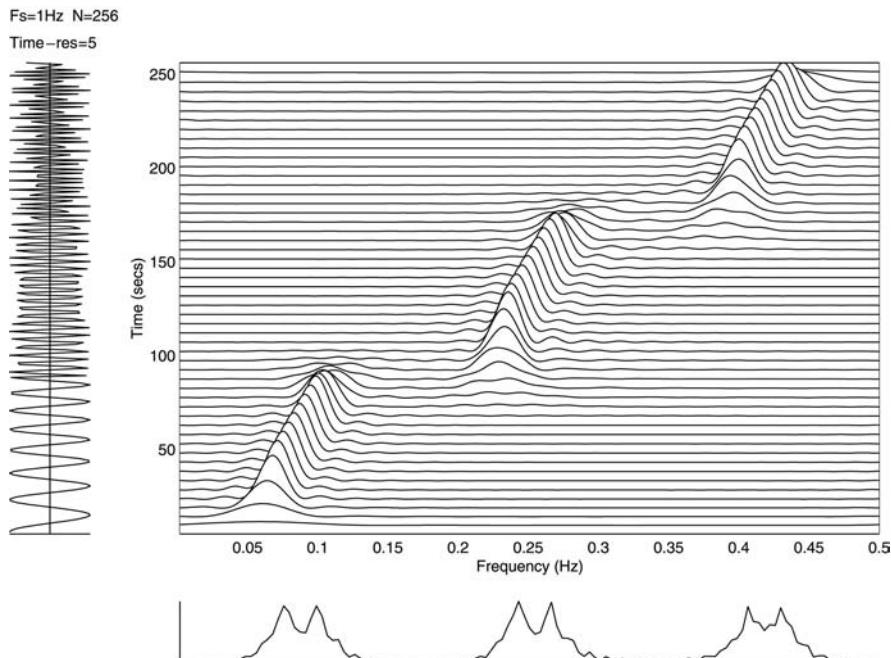
The expression $S_x(t, f)$ given by Equation (9.5) reduces to the Wigner–Ville spectrum [5]. For practical reasons, $x(t)$ is replaced by its analytical associate $z(t)$. Estimates of $S_z(t, f)$ can be expressed by the quadratic class of TFDs [5]:

$$\rho_z(t, f) = W_z(t, f) \underset{t, f}{**} \gamma(t, f) \quad (9.6)$$

where $\gamma(t, f)$ is a 2-D kernel window that is application dependent, $W_z(t, f)$ is the Wigner–Ville distribution (WVD) of $z(t)$, and $\underset{t, f}{**}$ indicates a double convolution in both time and frequency. By expanding Equation (9.6), we obtain [5]:

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\nu, \tau) z(u + \tau/2) z^*(u - \tau/2) e^{j2\pi(\nu t - \nu u - f\tau)} du d\nu d\tau \quad (9.7)$$

where the superscript * stands for complex conjugate. The kernel filter $g(\nu, \tau)$ characterizes a particular TF distribution and is related to $\gamma(t, f)$ via a double Fourier transformation. Equation (9.7) is the general formulation of the quadratic class of

**FIGURE 9.3**

TF representation of a three-component signal.

TFDs in terms of the Doppler-lag kernel $g(\nu, \tau)$ [6]. The TFD $\rho_z(t, f)$ can be further expressed in the following form:

$$\rho_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(t-u, \tau) z(u + \tau/2) z^*(u - \tau/2) e^{-j2\pi f \tau} du d\tau \quad (9.8)$$

where $G(t, \tau)$ is the inverse FT of filter kernel $g(\nu, \tau)$. A shorter version of Equation (9.8) is given by:

$$\rho_z(t, f) = FT_{\tau \rightarrow f} \{ K_z(t, \tau)^* G(t, \tau) \} \quad (9.9)$$

where $K_z(t, \tau) = z(t + \tau/2) z^*(t - \tau/2)$ is the signal kernel. This equation leads to the easiest discrete-time version of TFDs for implementation purpose [5].

A typical TF plot of a nonstationary multicomponent frequency-modulated (FM) signal is shown in Figure 9.3. The representation on the left is the time-domain representation. From this representation we observe that the signal significantly changes its periodicity characteristics at time $t = 100$ sec and time $t = 175$ sec. The representation at the bottom is the magnitude spectrum of the signal. It shows that this signal consists of three disjoint frequency bands. The main plot is the TF representation of the same signal. It not only combines the information from both time and frequency representations but also preserves the phase information, disregarded

or unavailable in the other two representations. This allows the joint TF representation to clearly show that the signal consists of three components characterized by:

1. Frequency 0.05 to 0.1 Hz during time $t = 0$ to 100 sec
2. Frequency 0.25 to 0.3 Hz during time $t = 100$ to 175 sec
3. Frequency 0.4 to 0.45 Hz during time $t = 175$ to 250 sec

The TF representation shows not only the number of components present in the signal but also how they vary in time and frequency, when they start and finish and whether they coexist at the same time or same frequency. For this particular example, the TF representation indicates the linear FM nature of all three components, which cannot be deduced from either time or frequency representations only.

9.2.2 Time–frequency distribution selection

The EEG data collected thus far show that neonatal EEG seizures are highly nonstationary and occasionally multicomponent, and are mostly concentrated in the band of frequency (0 to 5) Hz [8–10]. These factors must be taken into account when selecting an optimal TF distribution, because each TFD is more suited to representing signals with particular characteristics. For example, the WVD is known to be optimal for monocomponent linear FM signals but performs poorly for multicomponent or highly nonlinear FM signals.

Because neonatal EEG signals are nonstationary and occasionally multicomponent, the selected TFD needs to have good spectral resolution and reduced cross terms. For this sake, the performances of several distributions have been compared to find an optimal TF representation of real neonatal EEG data. The comparison study included the spectrogram, Wigner–Ville (WV), Choi–Williams (CW), B (BD), Zalto–Atlas–Marks (ZAM), Born–Jordan (BJ) and Rihaczek–Margenau distributions [8]. The performances of the resulting TFDs have been compared using an objective quantitative measure criterion [7]. Based on this criterion along with extensive experimental results, the B distribution BD with the smoothing parameter β equals to 0.01 — see Equation (9.10) — has been found to be the most suitable TF representation of the newborn EEG. The BD is defined in terms of its time-lag kernel — see Equation (9.8) — which is given by:

$$G(t, \tau) = \left(\frac{|\tau|}{\cosh^2(t)} \right)^\beta \quad (9.10)$$

where β ($0 < \beta \leq 1$) is a parameter that controls the sharpness of the cutoff of the 2D filter in the time-lag domain (t, τ) . [Figure 9.4](#) shows the TFDs of a 30-sec sample of real newborn EEG data using the B, CW, ZAM and WV distributions and clearly illustrates the superiority of B distribution in terms of resolution and cross-term reduction.

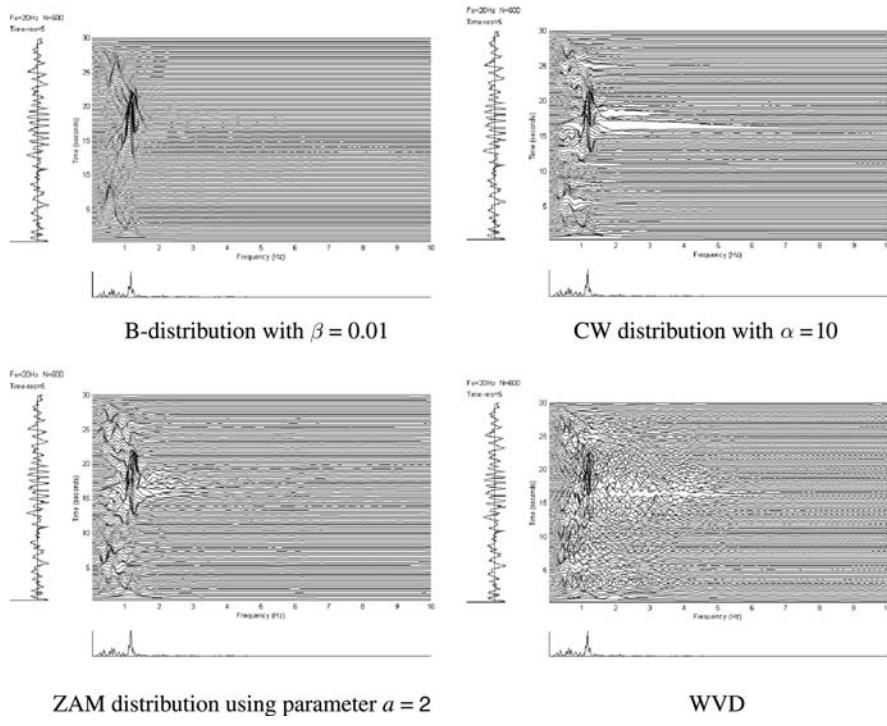


FIGURE 9.4

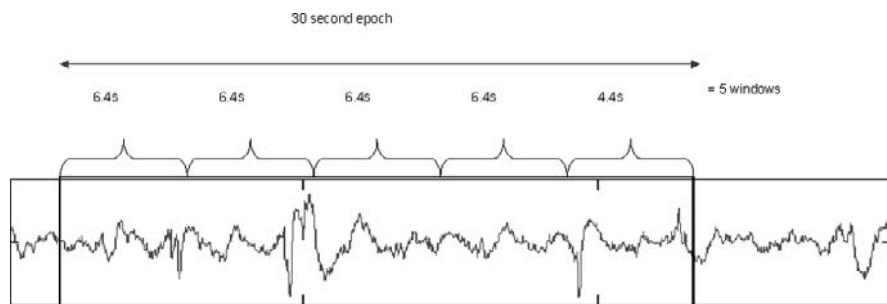
TF representations of a real epoch of newborn EEG using different TFDs.

9.2.3 From time domain to time–frequency domain

Seizure detection using the autocorrelation method relies on the assumption that the essential characteristic in newborn seizure EEG is periodicity. The periodicity of the autocorrelation of short epochs of EEG data is quantified (scored) using a nonlinear rule-based algorithm. In this technique, an epoch consisting of 30 sec of data is divided into 5 windows (see Figure 9.5). Depending on the autocorrelation function of each window, up to four primary periods (T_1, \dots, T_4) are calculated for each window in an epoch as shown in Figure 9.6. These times correspond to the moment centers of the first four peaks of the autocorrelation function. The windows are then scored whereby more evenly spaced primary periods are allocated larger scores. After each window in an epoch is scored, a second rule-based scheme is used to classify each epoch as positive or negative. If an epoch is classified as positive in two or more channels, this epoch is considered to contain seizure [20].

9.2.3.1 Seizure criteria for the autocorrelation method

A single window is declared as seizure positive by the autocorrelation method if the following criteria are met [8]:

**FIGURE 9.5**

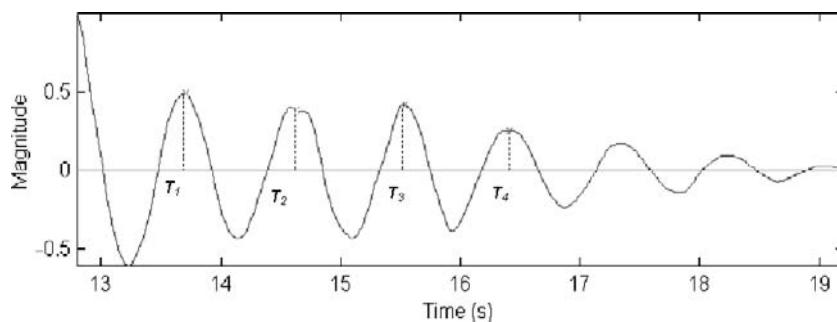
Epoch and window definitions according to the autocorrelation method.

1. At least four periods exist within the positive half of the autocorrelation function.
2. The differences between the ratios of each moment center to the first and the nearest integer are less than 0.150.
3. The total score obtained by summing all moment center scores is greater than or equal to 12 (out of a maximum of 15).

9.2.3.2 Mapping of autocorrelation seizure criteria to time–frequency domain

An EEG signal within a given window is considered seizure if in the TF domain a continuous spectral peak exists within that window and meets the following criteria:

1. All frequencies within the spectral line are greater than 0.625 Hz within 6.4-sec windows or greater than 0.909 Hz within 4.4-sec windows.

**FIGURE 9.6**

Autocorrelation function for one window.

2. The length of continuous dominant spectral peak within the window is greater than 3 sec.

A continuous spectral peak is defined as adjoining peaks within the TF array above a selected threshold of one fifth the maximum array value. The threshold was chosen experimentally.

Criterion 1 is a direct translation to the frequency domain of the first criterion described earlier in the autocorrelation domain. The second criterion has been deduced by observing several TF representations of seizure positive windows as defined by the autocorrelation method. The scoring system proposed by the autocorrelation method focuses on identifying periodic regions of data in the lag domain. Periodic regions are clearly identified in TF representations by a dominant spectral peak occurring for a certain time interval.

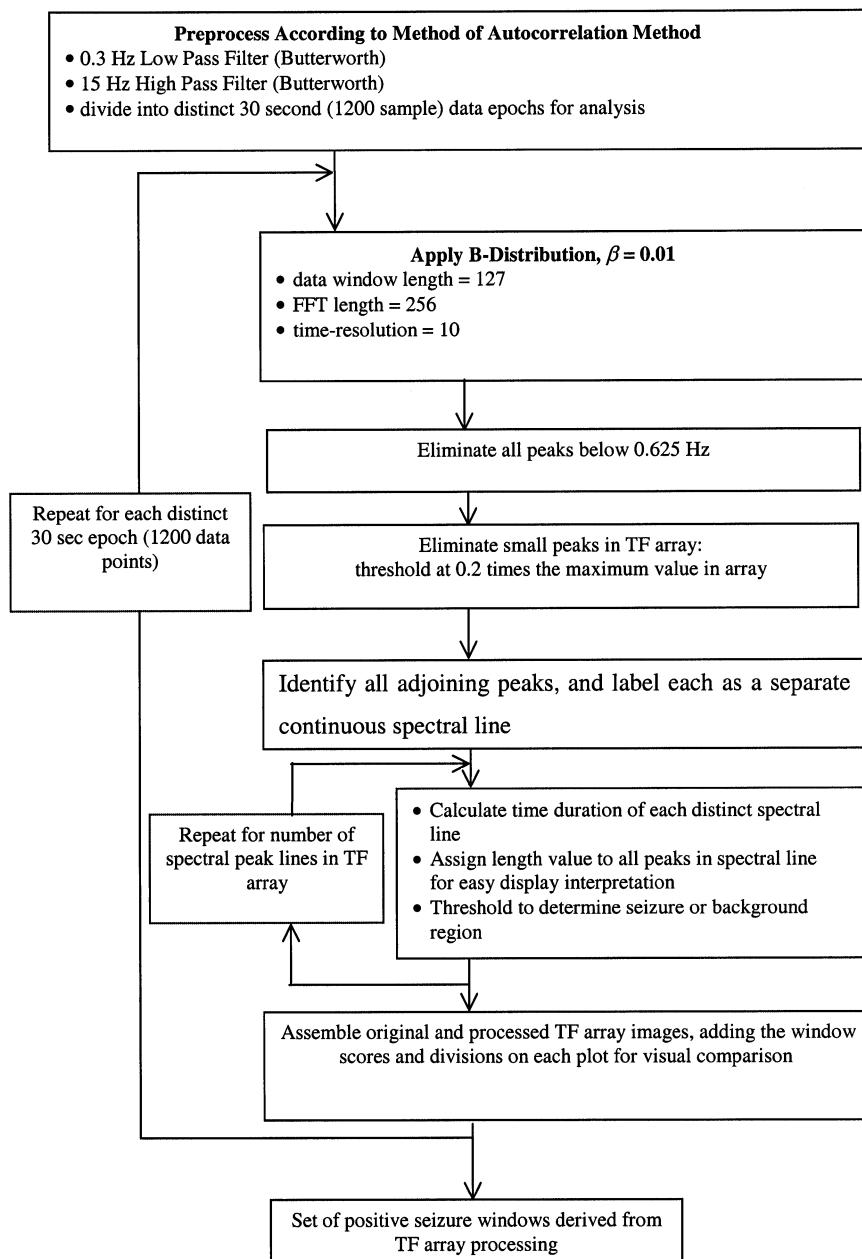
9.2.3.3 Results and discussion

Extraction of the seizure criteria listed earlier in the TF domain has been successfully calibrated with the autocorrelation method. Peak detection techniques have been employed to simplify the extraction process, resulting in a detection array illustrating positions and lengths of continuous spectral lines within each epoch. [Figure 9.7](#) shows the algorithm flowchart used in the calibration of this method.

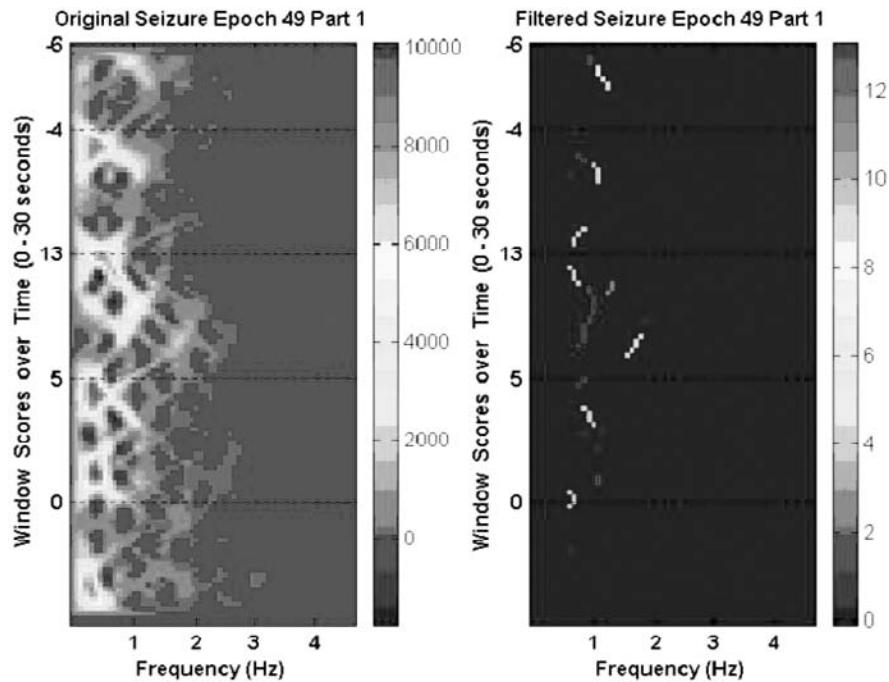
Good results have been obtained using TF algorithm to detect individual seizure windows of real neonatal EEG in the TF domain. Approximately 75% of windows detected as seizure positive by the autocorrelation method are detected using TF criteria listed earlier. The results of applying the TF-based detection method are summarized in [Figure 9.8](#). In this figure, the original epoch refers to the raw TF array produced from preprocessed EEG data. Image of this array appears on the left side of the figure. The image is also divided into four distinct 6.4-sec windows and one 4.4-sec window as defined by the autocorrelation method. Window scores attributed to these windows are displayed at the end of each window division for easy comparison between the TF information and the corresponding score allocated by autocorrelation method.

9.2.4 From frequency domain to time–frequency domain

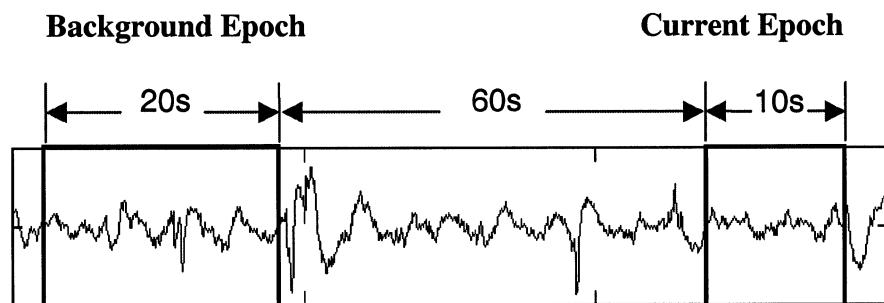
The spectrum analysis method has also been used to detect periodic discharges in the EEG signal. As shown in [Figure 9.9](#), a background epoch is defined as a 20-sec segment of EEG finishing 60 sec before the start of the current 10-sec epoch under investigation. From the frequency spectrum of each 10-sec epoch, the following features are extracted: the frequency of the dominant spectral peak, the width of the dominant spectral peak and the ratio of the power in the dominant spectral peak to that of the background spectrum in the same frequency band.

**FIGURE 9.7**

Implementation and calibration of the TF extension of time-domain method.

**FIGURE 9.8**

The mapping of the features from autocorrelation to TF. Seizure windows are defined by the presence of high value lines. As the length of each spectral line is denoted by its color, windows with dark lines should indicate scores higher or equal to 12.

**FIGURE 9.9**

Epoch and window definitions according to the spectrum method.

9.2.4.1 Seizure criteria in the spectral domain

The 10-sec epoch of EEG data is considered seizure positive if any of the 3 following criteria are met [16]:

	<i>Dominant Frequency</i>	<i>Half-Bandwidth</i>	<i>Power Ratio</i>
Criterion 1	0.5–1.5 Hz	≤ 0.6 Hz	3–4
Criterion 2	1.5–10 Hz	≤ 0.6 Hz	2–4
Criterion 3	1.5–10 Hz	≤ 1 Hz	4–80

To limit the number of false alarms, a seizure detection is discounted if the epoch is largely nonstationary, if a large amount of alternating current (AC) power noise is present or if it appears that an EEG lead has been disconnected.

9.2.4.2 Mapping of spectrum criteria to time–frequency domain

The preceding criteria pertaining to frequency and bandwidth are clearly observable in the TF domain. That is, each spectra containing a dominant peak that meets either of the criteria:

1. Frequency in the range 0.5 to 1.5 Hz and width ≤ 0.6 Hz
2. Frequency in the range 1.5 to 10 Hz and width ≤ 1 Hz

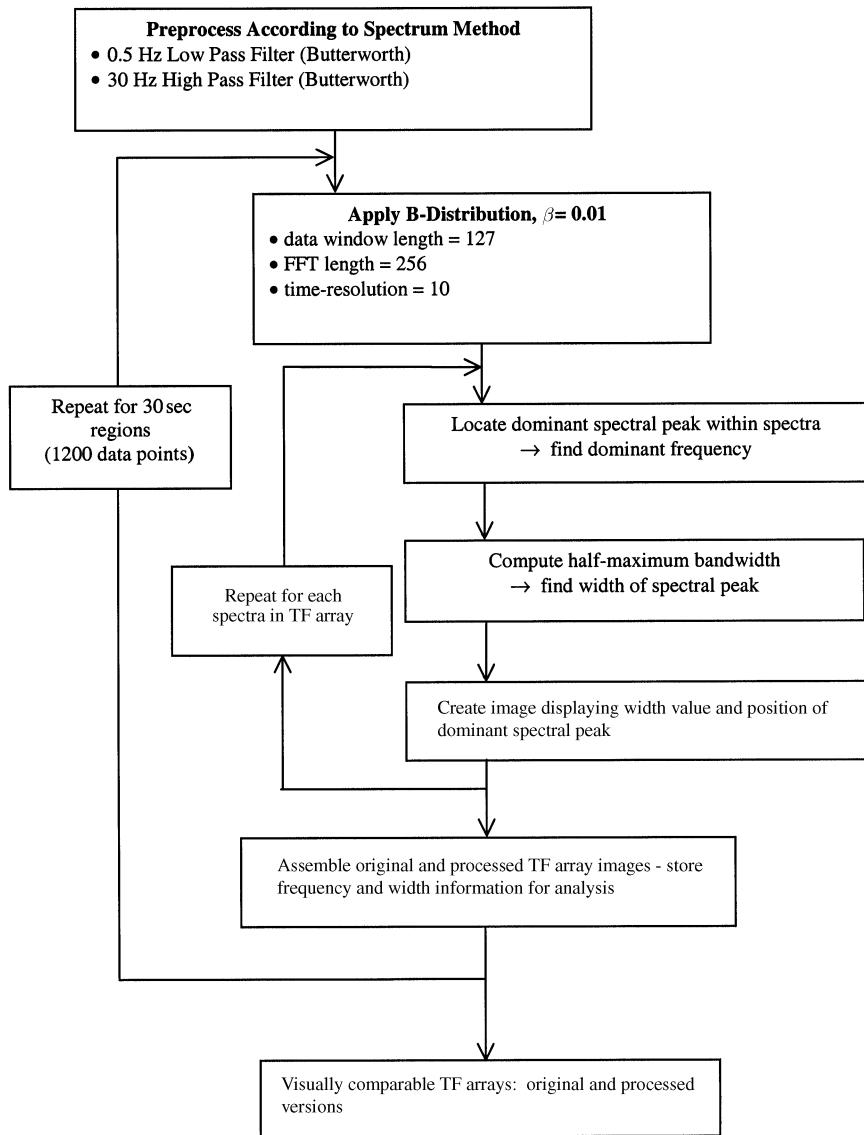
may be considered for further seizure detection pertaining to power ratio. Disregarding the power ratio criteria, the second criterion given in the preceding table becomes a subset of the third criterion.

9.2.4.3 Results and discussion

The calibration of the two TF criteria identified earlier is illustrated in [Figure 9.10](#). This essentially extracts frequency and width information, the results of which are visible in the plots shown in [Figure 9.11](#) for a 30-sec epoch. Data are presented by highlighting the dominant frequency with a color indicating the width of the spectral peak. Boxed sections of the array indicate EEG regions detected as containing seizure by the conventional spectrum method.

9.2.5 Time–frequency patterns

Calibration of the TF method using both autocorrelation and spectrum method has been successfully performed. This calibration led to consideration of a fully integrated TF detection method by combining the different TF seizure features. The visual analysis of these EEG TF features indicates that there are two distinguishable classes of TF EEG patterns: seizure and background. Further, the seizure patterns can

**FIGURE 9.10**

Implementation and calibration of TF extension of frequency-based method.

be characterized in terms of the dominant spectral peak as either a linear FM or a piecewise linear FM whereas the background patterns exhibit a low-frequency burst of activities or activities with no clearly defined pattern. These visual observations correlate well with the clinical information related to the different patterns found in

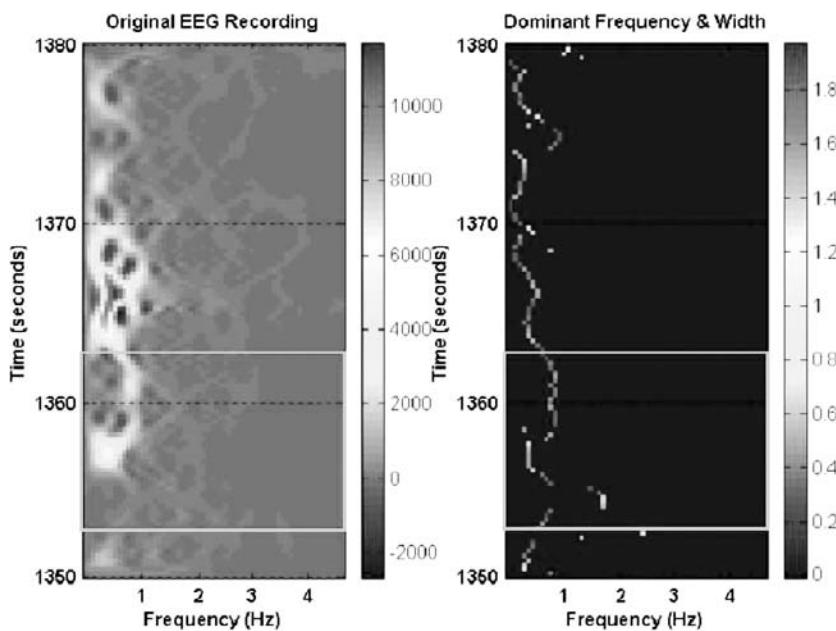


FIGURE 9.11

Mapping of the features from spectrum to TF. The color of the dominant frequency is related to the width of the spectral peak. Boxed sections indicate the presence of seizure as detected by the spectrum method.

EEG [21]. A few representative TF patterns that characterize each of the subclasses defined so far are given next.

9.3 Seizure Patterns

As mentioned earlier, the following patterns characterize the different patterns found thus far in the EEG signals in terms of dominant spectral peak.

9.3.1 Linear frequency-modulated patterns

A number of the EEG seizures analyzed in the TF domain can be approximated by linear FM (LFM) with either fixed or time-varying amplitudes. These patterns can be further classified into the following subclasses:

9.3.1.1 Class A (linear frequency-modulated patterns with a quasi-constant frequency)

Figure 9.12 shows a seizure that has a linear FM behavior with an almost constant frequency. The amplitude of the TF seizure pattern increases at the onset and decreases

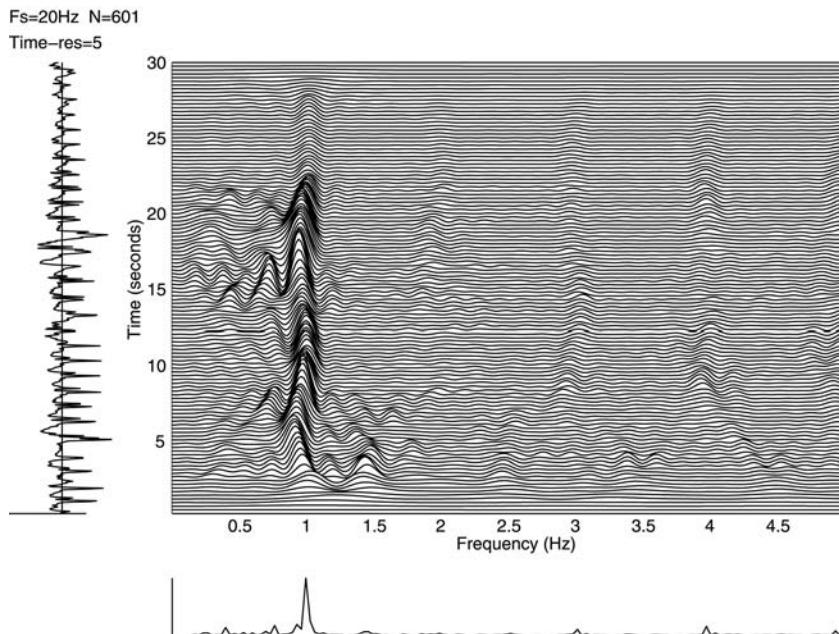


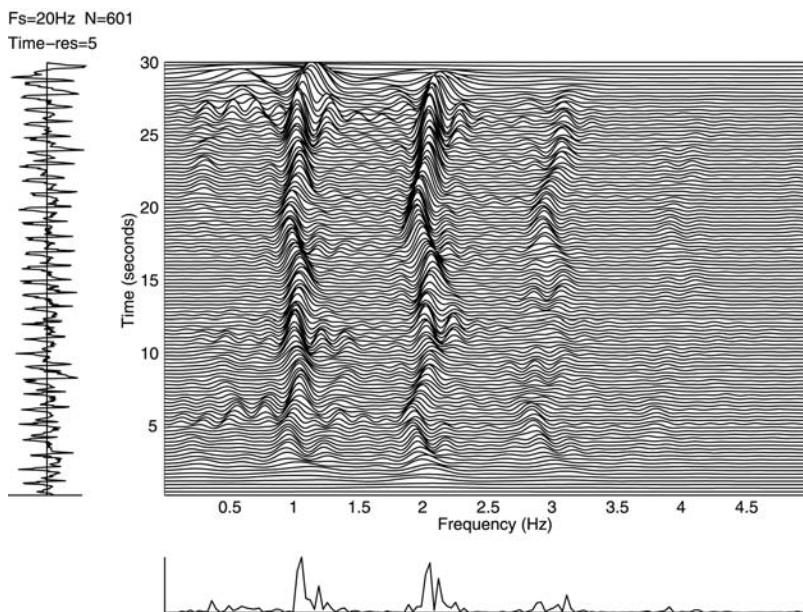
FIGURE 9.12

TF representation of a seizure that shows a linear FM behavior with constant frequency.

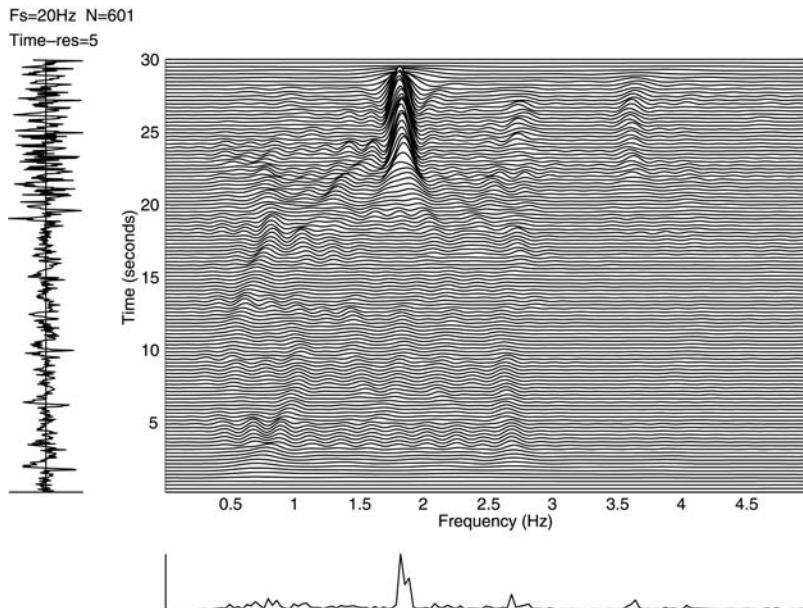
toward the end. These observations are in agreement with the seizure definition found in medical literature [21]. In this figure, we can clearly see the dominant peak (frequency domain) and the repetitive pattern of the spikes and sharp waves (time domain) characterizing the newborn seizures. A major advantage of the TF representation is that we can easily distinguish the seizure from the background. An interesting phenomenon that was observed is the existence of subharmonics in a number of TF representations of EEG seizure as illustrated in, for example, [Figure 9.13](#). This characteristic of the newborn EEG seizure has also been reported by Lombroso [21].

9.3.1.2 Class B (short linear frequency-modulated patterns with a Quasi-constant frequency)

[Figure 9.14](#) shows a seizure pattern that differs from those in class A in one important aspect, the duration. A disagreement exists between researchers about what constitutes a seizure. The duration of rhythmic discharges that characterizes a seizure is highly variable, from as short as 1 sec to as long as 30 min [25]. To consider an EEG signal as a seizure, some researchers require that it must last at least 10 sec [12, 16]; others require a minimum of 20 sec [23], whereas a third group does not specify a time limit [19–21]. Due to this disagreement, these controversial seizures were classified as a separate category.

**FIGURE 9.13**

TF representation of a seizure that shows subharmonics.

**FIGURE 9.14**

TF representation of a short duration harmonic discharge "seizure."

9.3.1.3 Class C (linear frequency-modulated patterns with a decreasing frequency)

Figure 9.15 of this class can be characterized by its IF that tends to decrease with time. This frequency-decreasing behavior is widely accepted as a characteristic of the EEG seizure [21, 29]. The advantage of the TF representation over those of the time or frequency is that the linear FM character is easily recognizable. A similar behavior can be seen in the work of Franaszczuk on adult seizures originating from the mesial temporal lobe during the periods of organized rhythmic activity [14].

9.3.2 Piecewise linear frequency-modulated patterns

A pattern of this class is shown in [Figure 9.16](#). It can be approximated by piecewise linear FM. Something similar can be observed in one of the TF representations produced by Franaszczuk in his study of the adult seizure originating from the mesial temporal lobe [14].

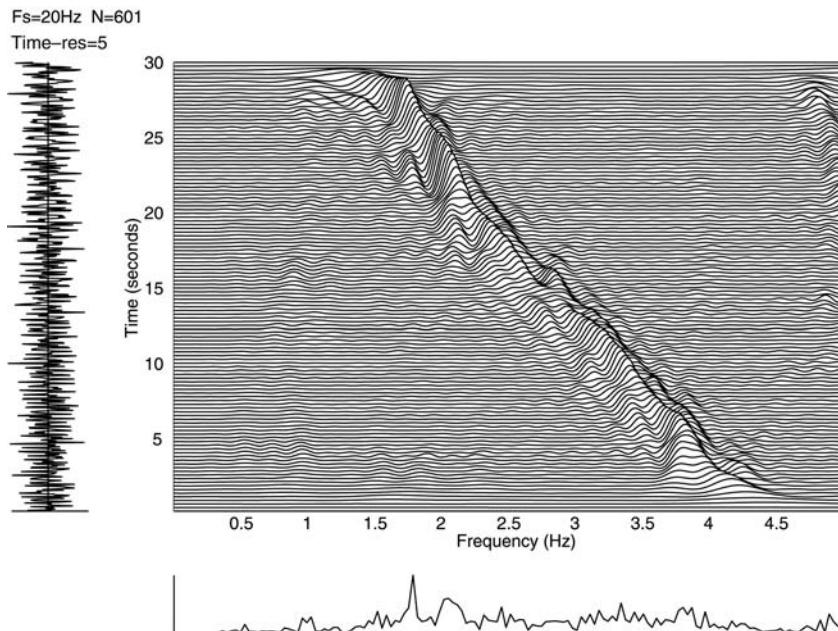
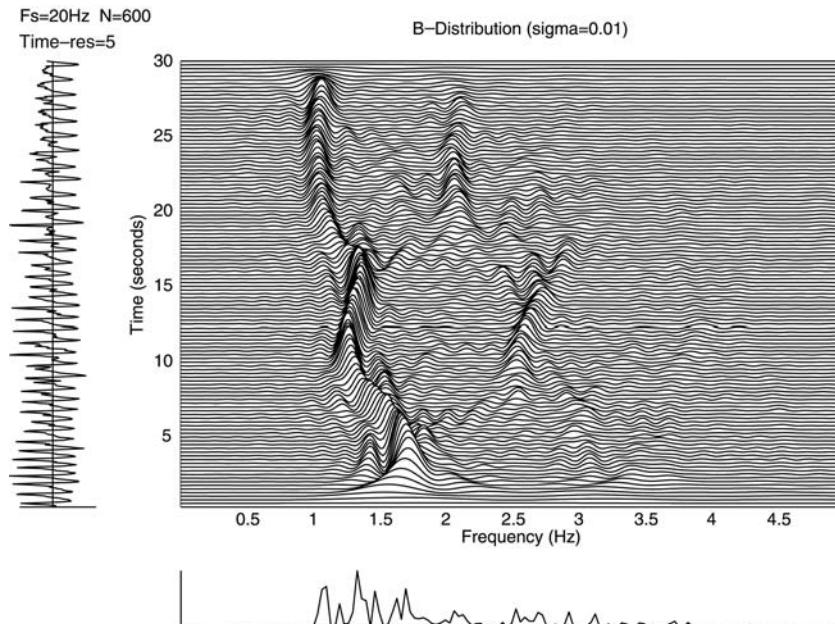


FIGURE 9.15

TF representation that shows an LFM behavior with decreasing frequency.

**FIGURE 9.16**

TF representation of an EEG seizure that shows a piecewise LFM behavior.

9.4 Electroencephalogram Background Patterns

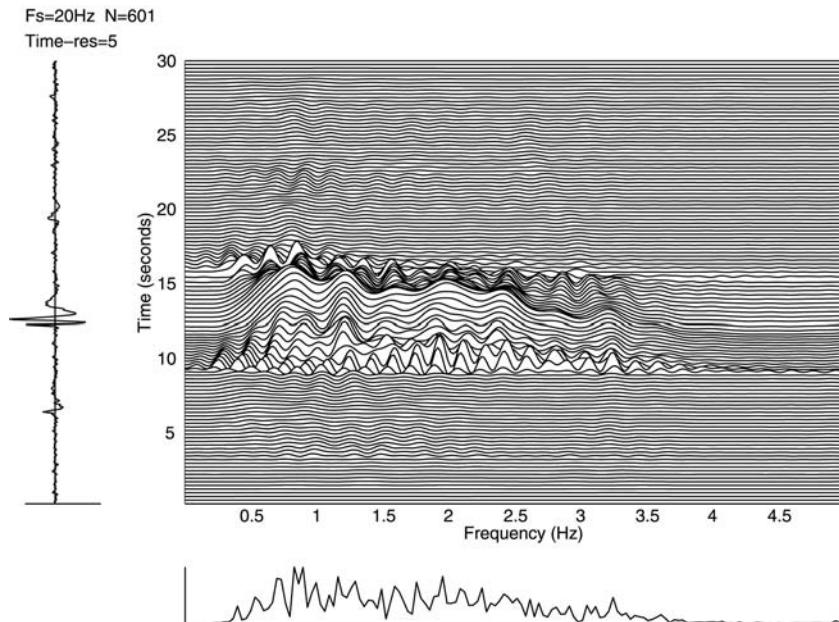
By background, we mean any signal that is not classified as seizure. Two distinct patterns have been observed while analyzing the EEG signals: burst suppression activity and irregular activity with no clear pattern.

9.4.1 Class E (burst suppression)

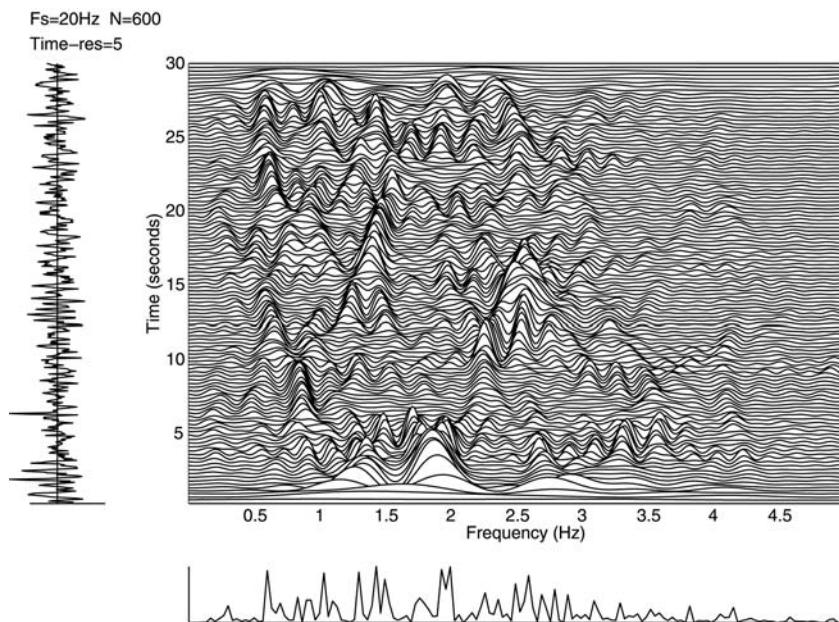
[Figure 9.17](#) is an example of this class. It shows instance of burst activities. These are short period signals with a high energy lasting for few seconds and usually occurring at frequencies below 8 Hz. These features are characteristic of burst suppression as defined in the literature. Burst suppression consists of burst of high voltage activity lasting 1 to 10 sec and composed of various patterns (δ [0 to 4 Hz] and θ [4 to 8 Hz] with superimposed and intermixed spikes, sharp waves and faster activity) followed by a marked background attenuation [29].

9.4.2 Class F (activities lacking a specific pattern)

[Figure 9.18](#) is an example of EEG epochs lacking a well-defined and consistent pattern. These types of activities do not seem to be constrained within the low-frequency band characterizing the seizure signals.

**FIGURE 9.17**

TF representation of burst suppression.

**FIGURE 9.18**

TF representation of an EEG lacking specific and consistent pattern.

9.5 Time-Frequency Matched Detector

The matched filter (which essentially reduces to a correlator receiver) is very popular for constructing detectors and classifiers. For known deterministic signals in white Gaussian noise (WGN), the correlator receiver is known to be optimal in the Neyman Pearson sense for detection [32] and in a Bayesian sense for classification [31]. To construct a detector, the output of a correlator is compared with a threshold. The threshold is chosen such that the probability of a false alarm is maintained.

Traditionally, the correlator receiver is implemented in time domain as a one-dimensional (1-D) correlation between the received noisy signal $x(t)$ and a reference signal $s(t)$. For random signals, this can be done using the corresponding spectral representations. To extend this detector to handle nonstationary signals, the 1-D correlation is replaced by a 2-D correlation involving the TF distribution of $x(t)$ and $s(t)$, as shown in Figure 9.19. The resulting test statistic is given by:

$$T(x) = \iint \rho_x(t, f) \rho_s(t, f) dt df \quad (9.11)$$

This type of detector has been implemented using different quadratic TF distributions such as the spectrogram [1, 33], the reassigned spectrogram [33], the WVD and cross-WVD [2–4, 13, 18, 30], and generalized time-shift covariant quadratic TFD [26–27]. It has also been implemented in the ambiguity domain by replacing the TFDs by the auto- and cross-ambiguity functions [33].

By using Moyal's formula [17], we get:

$$\iint \rho_x(t, f) \rho_s(t, f) dt df = \left| \int x(t) s^*(t) dt \right|^2 \quad (9.12)$$

This equality is only valid when the filter kernel $g(\nu, \tau)$ in Equation (9.8) is unimodular, that is:

$$|g(\nu, \tau)| = 1 \quad (9.13)$$

for all ν and τ .

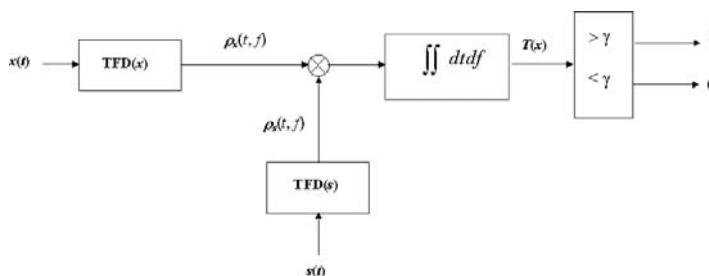


FIGURE 9.19

TF matched receiver.

This is the case, for example, for the WV, the Rihaczek [13] and the generalized time-shift quadratic TFDs [27]. Equation (9.12) is an alternative interpretation of the correlator receiver in terms of a correlation of the TFDs.

It has been shown [2, 13] that for the case of deterministic signal in noise (even WGN) that the TF-based correlator is suboptimal. This suboptimality can be partly explained by the nonlinearity of the quadratic TFDs that accentuates the effects of noise by introducing artifacts.

To use a correlator receiver, it is usually required that the wave shape of the reference signal (or other related information such as its TFD) as well as the noise statistics are known. For the case of EEG seizure detection, the wave shape of the EEG seizure is generally unknown. In Section 9.3, however, we have shown that the EEG seizure can be characterized by a linear or a piecewise linear FM. Based on this conclusion, we propose to construct a TF-based matched filter (i.e., essentially a correlator). To do this, a representative TF distribution of a linear FM $\rho_{ref}(t, f)$ is selected to serve as template (reference). The correlator statistic $T(z)$ used is the 2-D cross-correlation between the noisy signal and the reference signal TFDs [2], that is:

$$T(z) = \iint \rho_{ref}(t, f) \rho_z^*(t, f) dt df \quad (9.14)$$

where z is the analytical signal corresponding to the EEG signal to be analyzed.

9.5.1 Implementation of the time–frequency-matched detector

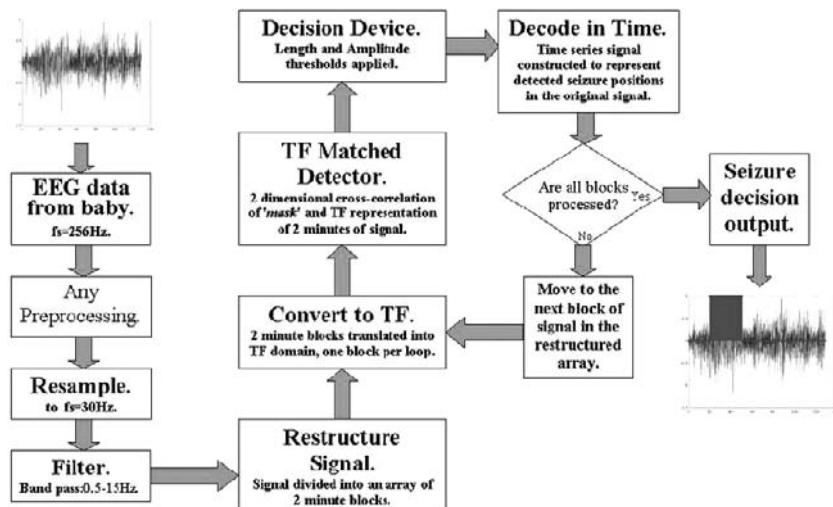
A working model of the TF-matched detector as shown in [Figure 9.20](#) has been implemented. A description of its main components are described in the next section.

9.5.1.1 Preprocessing

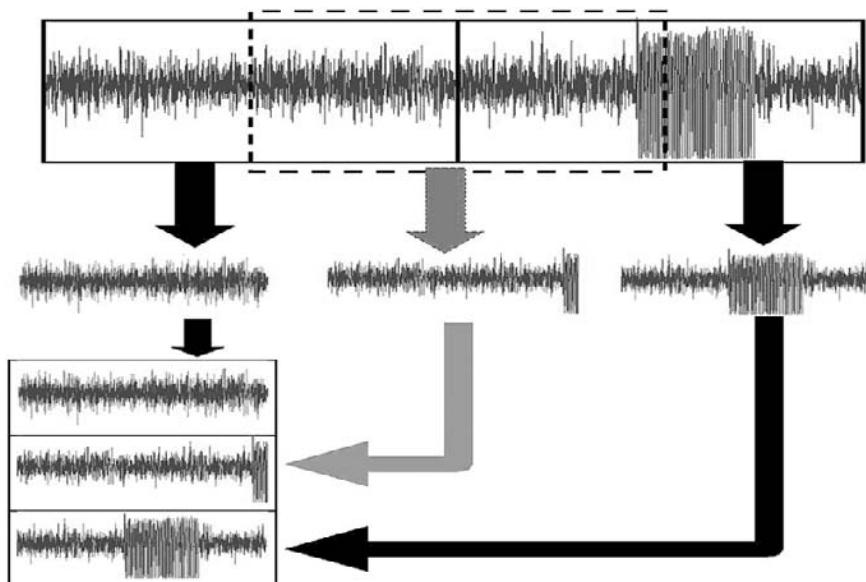
This stage can include any preprocessing of the EEG signal collected from the newborn's scalp. Some major reasons for preprocessing the EEG signal include artifact (such as ECG, EOG and EMG) removal, noise filtering and resampling the signal to comply with detector input specifications. In the present implementation of this detector, the process of artifact removal has not been included.

9.5.1.2 Signal restructuring

Because it is neither desirable nor practical to work with a very long EEG signal, we found it necessary to segment the EEG into an array of signals of fixed length. After some preliminary testing, we found a length of 2 min to be suitable for performing the cross-correlation. Shorter signal lengths led to higher rates of miss detections and false alarms. Once the full input EEG signal is divided into blocks of 2-min duration, each block is stored as a row of the newly formed array of signals. In the development of this detector, a protocol of 50% overlap of each block was adopted, as displayed in [Figure 9.21](#).

**FIGURE 9.20**

Block diagram of TF-matched detector.

**FIGURE 9.21**

Segmenting the EEG signal in blocks of 2-min length.

9.5.1.3 Detection loop

The detection loop is executed until all the blocks of the input EEG signal have been processed. An offset value is maintained, giving a precise location in the original signal where abnormal events are detected.

9.5.1.4 Cross-correlation implementation

This forms the critical step in the effectiveness of this TF seizure detector. The cross-correlation between the input TF array and the template (mask) is obtained using the 2-D cross-correlation function given by Equation (9.14). The most crucial process is the choice of the template. This problem is discussed later. We found it convenient to normalize TF distributions of both the reference signal and the EEG signal.

9.5.1.5 Amplitude and length criteria

Ideally there would be one peak value in the output of the cross-correlation array, with its amplitude determining the presence or absence of seizure. This proved a little unreliable, and it was decided to search for a sequential series of values over the amplitude threshold. This proved very successful, and a minimum ridge length of 20 sec over the amplitude threshold was classed as a seizure. The 20-sec length adopted is larger than the minimum 10-sec length of EEG seizure adopted by many neurologists [12]. This length can be changed at will. An example of the correlation output, following the application of a threshold is shown in Figure 9.22, where the ridges of length equal or greater than 20 sec indicate the presence of seizure.

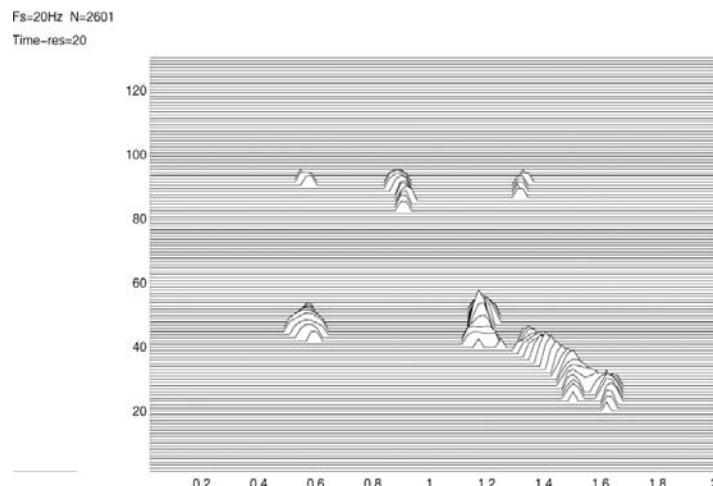
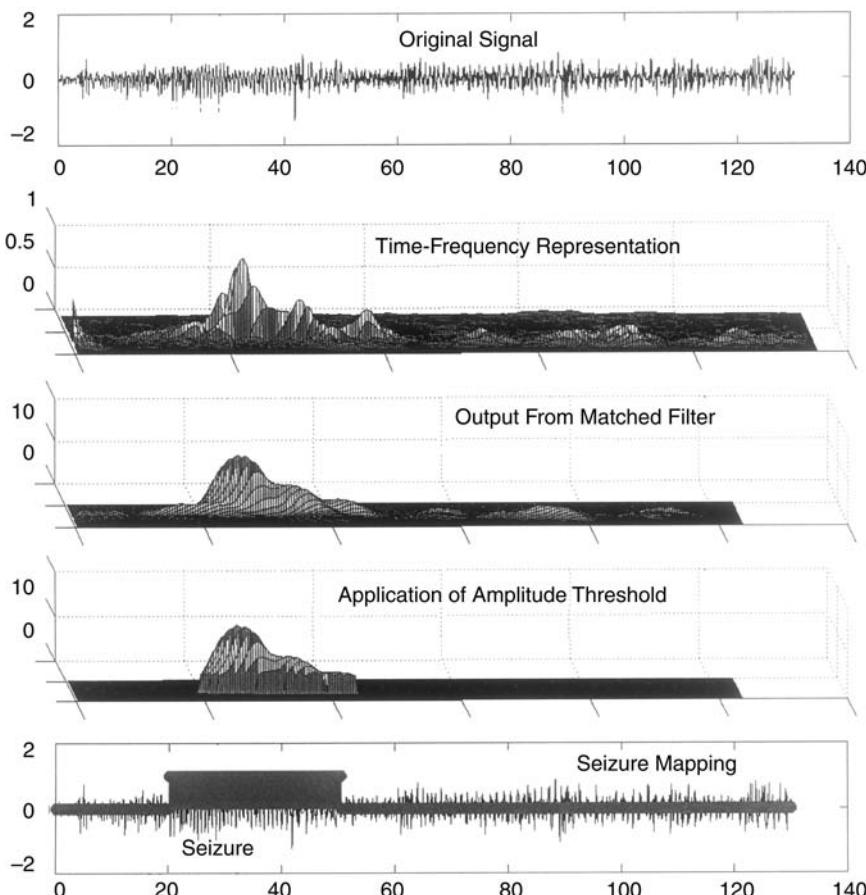


FIGURE 9.22

Amplitude threshold applied to output of the matched filter.

**FIGURE 9.23**

Different steps of seizure localization.

9.5.1.6 Map seizure decision to real-time location

This stage simply ties all the independent decisions on each block of processed signal (remapping any seizure decision to a time series function) of equivalent length to the input EEG signal. This output waveform consists of ones or zeros, where one indicates the presence of seizure at the corresponding time. Figure 9.23 summarizes the five stages from input signal to seizure indicated on the output signal.

9.5.2 Experimental setup

To validate and calibrate the detection method, simulated data generated by the EEG model proposed by Roessgen, Zoubir and Boashash [28] is used.

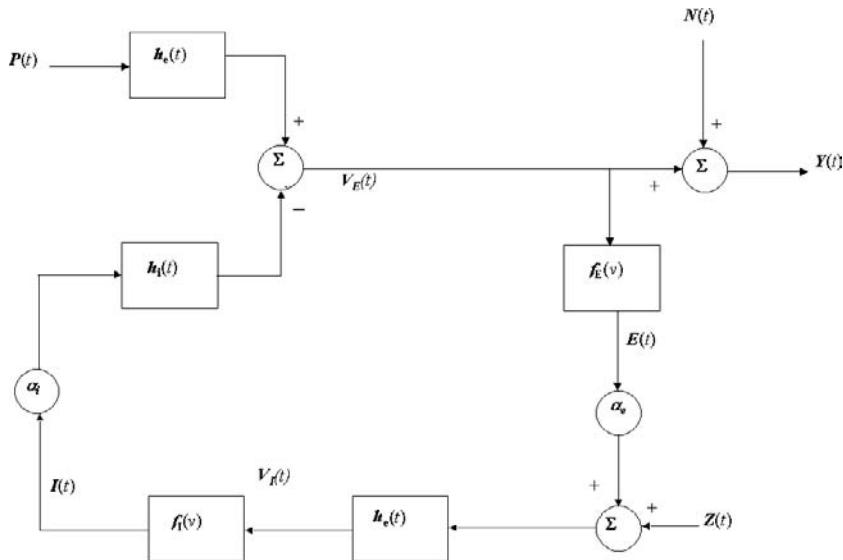


FIGURE 9.24
EEG model.

9.5.2.1 Electroencephalogram model

Roessgen extended the model of the α rhythm of the thalamus proposed by da Silva [22] by introducing a seizure input waveform $Z(t)$ as shown in Figure 9.24. The model considers a population of interconnected neurons driven by a random input — a WGN $P(t)$ — that is assumed to originate from deeper brain structure such as thalamus and brain stem. The parameters of the model reflect some physiological characteristics such as neuronal interconnectivity, synaptic pulse response and excitation threshold. For an in-depth discussion of this model, see [28].

To account for the nonstationarities of the EEG, the periodic sawtooth signal $Z(t)$ of period f_c used by Roessgen was replaced by the following more realistic quasi-periodic nonstationary signal that fits a linear FM law with a time-varying amplitude:

$$S_{LFM}(t) = Z(t)e^{j\pi\alpha t^2} \quad (9.15)$$

The typical range for α is $[-0.07$ to $0]$. The initial frequency f_0 is typically in the range of 0.5 to 5 Hz. These values were reported in [15, 21]. A signal-to-background ratio (SBR) was introduced to account for the fact that the seizure and background components of the newborn EEG are of almost similar magnitudes. The SBR in decibels is defined as:

$$SBR = 10 \log \left(\frac{\sigma_s^2}{\sigma_b^2} \right) \quad (9.16)$$

where σ_s^2 and σ_b^2 represent the power in the seizure and the background components of the EEG signal, respectively.

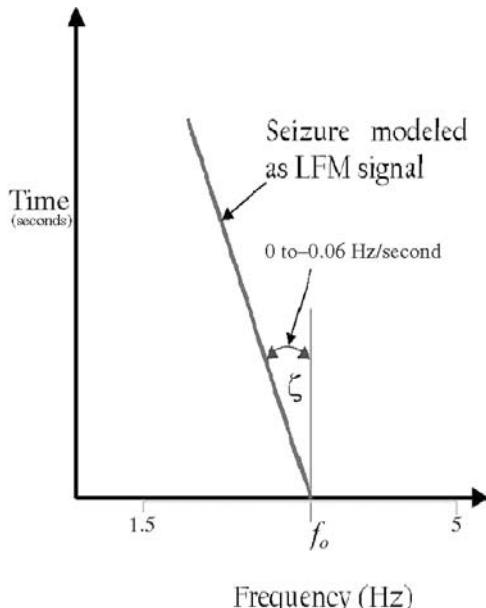


FIGURE 9.25
Template model.

9.5.2.2 Template selection

The reference template to be used for detection is the TF distribution of a linear FM (LFM, Figure 9.25). The duration of the LFM signal is set at 20 sec as discussed earlier. We have chosen the slope of the LFM to be negative because the majority of the seizures encountered at this stage are characterized by a frequency that is either constant or decreasing with time.

To find the best slope of the template ridge (see Figure 9.25), that is, the one that corresponds to the best detection rate for the different types of seizures, a testing stage was necessary. A number of templates with different ridge slopes were created. These templates were tested using the seizure detection system presented earlier (see Figure 9.20) and a number of EEG signals produced by the model in Figure 9.24. These signals were generated using different combinations of the model parameters. It was found that the template with $\zeta = -0.05 \text{ Hz/sec}$ gives the best results.

9.5.2.3 Threshold selection

To select an optimum threshold that realizes the good compromise between the rate of good detections and the rate of false alarms, a similar test to the one previously discussed was performed. Among the different thresholds tested, the ones that have a value around 4 resulted in a good compromise.

9.5.3 Time–frequency detector evaluation

To check its reliability and to validate its effectiveness, the TF detector is tested using synthetic data generated by the EEG model (see [Figure 9.24](#)). The synthetic signal contains a mixture of randomly placed seizures and background. For this investigation, 1000 synthetic EEG signals were used, with the parameter selection varied as follows:

- SBR = random [0 to 20 dB]
- SNR = random [0 to 20 dB]
- α = random [−0.06 to 0]
- f_c = random [1.5 to −5]

The average detection rate was 99.1 whereas the false alarm rate was 0.4%. These good results were expected because the template adopted is well adapted to the EEG model that only produces noisy linear FM signals with time-varying amplitudes. Other EEG models could be able to generate different types of EEG-like seizures such as piecewise LFM. For those signals, a template such as the one proposed by Celka, Boashash and Colditz [11] would be used.

9.6 Discussion and Conclusions

This chapter shows that the TF domain is the preferable basis from which to develop a complete EEG seizure detection scheme. The proposed TF approach was successfully calibrated using existing methods based on autocorrelation detection criteria and spectral estimation methods

The initial patterns obtained by a TF analysis of EEG seizure signals specifically confirm that EEG seizures in newborn are well characterized by a linear FM or piecewise linear FMs. These results are encouraging and further analysis on other data sets is currently in progress to refine and extend these findings. The characterization of nonstationary EEG signals in the TF domain is the first step toward an automatic method of seizure detection and classification that use powerful tools of TF signal processing [6]. In this direction and based on the findings thus far obtained, we proposed a TF detector based on cross-correlating the TF representation of the EEG signal with a TF reference template. The design of the template takes into account the TF characteristics of the EEG seizure thus far extracted. The performance of this TF detector was tested on synthetic signals. The results obtained were very encouraging (99.1% detection rate and 0.4% false alarm rate). Current work not yet finalized at time of publication is concentrating on using real newborn EEG signals.

Acknowledgments

This work is funded by the Australian Research Council (ARC). The authors wish to thank Professor Paul Colditz of the Royal Women's Hospital in Brisbane for providing us access to the Perinatal Research Centre and Dr. Chris Burke of the Royal Children Hospital for the interpretation of the EEG data.

References

- [1] R.A. Altes, Detection, estimation, and classification with spectrogram, *J. Acoust. Soc. Am.*, 67(4), 1232–1246, April 1980.
- [2] B. Boashash and P. O'Shea, Time-Frequency Analysis Applied to Signaturing of Underwater Acoustic Signals, International Conference of Acoustics, Speech, and Signal Processing, New York, April 11–14, 1988.
- [3] B. Boashash and P. O'Shea, Application of Wigner–Ville distribution to the identification of machine noise, in Proceedings of the SPIE, 975, 209–220, August 1988.
- [4] B. Boashash and P. O'Shea, A methodology for detection and classification of some underwater acoustic signals using Time-Frequency analysis techniques, *IEEE Trans. ASSP*, 38(11), 1829–1844, November 1990.
- [5] B. Boashash, Time-Frequency signal analysis, in S. Haykin, Ed., *Advances in Spectral Estimation and Array Processing*, Prentice Hall, New York, 1991, pp. 418–517.
- [6] B. Boashash, Ed., *Time-Frequency Signal Analysis: Methods and Applications*, Longman Cheshire/Wiley, Melbourne/New York, 1992.
- [7] B. Boashash and V. Sucic, A Resolution Performance Measure for Quadratic Time-Frequency Distributions, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, PA, 2000, pp. 584–588.
- [8] B. Boashash, H. Carson, and M. Mesbah, Detection of Seizures in Newborns Using Time-Frequency Analysis of EEG Signals, in Proceedings of the IEEE Workshop on Statistical Signal and Array Processing, Pocono Manor, PA, 2000, pp. 564–568.
- [9] B. Boashash, M. Mesbah, and P. Colditz, Newborn EEG Seizure Pattern Characterization Using Time-Frequency Analysis, in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (on CDROM), Salt Lake City, UT, May 2001.

- [10] B. Boashash and M. Mesbah, A Time-Frequency approach for newborn seizure detection: rationale and relationship to stationary methods, *IEEE Eng. Med. Biol. Mag.*, [in press].
- [11] P. Celka, B. Boashash, and P. Colditz, Pre-processing and Time-Frequency analysis of newborn EEG seizures, *IEEE Eng. Med. Biol. Mag.*, [in press].
- [12] R. Clancy, Interictal sharp EEG transients in neonatal seizures, *J. Child Neurol.*, 4, 30–38, 1989.
- [13] P. Flandrin, A Time-Frequency formulation of optimum detection, *IEEE Trans. Acoust. Speech Signal Process.*, 36(9), September 1988.
- [14] P.J. Franaszczuk, G.K. Bergy, P.J. Durka, and H.M. Eisenberg, Time-Frequency analysis using matching pursuit algorithm applied to seizure originating from the mesial temporal lobe, *Electroencephalogr. Clin. Neurophysiol.*, 106, 513–521, 1998.
- [15] J. Gotman, J. Zhang, J. Rosenblatt, and R. Gotman, Automatic seizure detection in newborns and infants, in Proceedings of the IEEE Engineering in Medicine and Biology Society, 2, 913–914, 1995.
- [16] J. Gotman, J. Zhang, J. Rosenblatt, and R. Gottesman, Evaluation of an automatic seizure detection method for the newborn EEG, *Electroencephalogr. Clin. Neurophysiol.*, 103, 363–369, 1997.
- [17] A.J.E.M. Janssen, Positivity and spread of Time-Frequency distributions, in *The Wigner Distribution: Theory and Applications in Signal Processing*, W. Meclenbrauker and F. Hlawatsch, Eds., Elsevier, Amsterdam, 1997, pp. 1–58.
- [18] S. Kay and G.F. Boudreaux-Bartels, On the Optimality of the Wigner Distribution for Detection, in Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, March 1985, pp. 1017–1020.
- [19] P. Kellaway and R.A. Hratchovy, Status epilepticus in newborns: a perspective on neonatal seizures, in *Advances of Neurology*, A.V. Delgado-Escueta, C.G. Westerlain, D.M. Treiman, and R.J. Porter, Eds., Vol. 34, Raven Press, New York, 1983, pp. 93–99.
- [20] A. Liu, J.S. Hahn, G.P. Heldt, and R.W. Coen, Detection of Neonatal Seizures through Computerized EEG Analysis, *Electroencephalogr. Clin. Neurophysiol.*, 82, 30–37, 1992.
- [21] C.T. Lombroso, Neonatal EEG polygraphy in normal and abnormal newborns, in *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, E. Niedermeyer and F.H. Lopes da Silva, Eds., Williams & Wilkins, Baltimore, 1993, pp. 802–875.
- [22] F.H. Lopes da Silva, Dynamics of EEG as signals of neuronal population: models and theoretical considerations, in *Electroencephalography: Basic*

Principles, Clinical Applications, and Related Fields, E. Niedermeyer and F.H. Lopes da Silva, Eds., Williams & Wilkins, Baltimore, 1993.

- [23] C.N. McCutchen, R. Coen, and V.J. Iragui, Periodic lateralized epileptiform discharges in asphyxiated neonates, *Electroencephalogr. Clin. Neurophysiol.*, 61, 210–217, 1985.
- [24] M. Mesbah, B. Boashash, and P. Colditz, Analysis of Newborn EEG Time-Frequency Patterns, in Proceedings of the 3rd Australian Workshop on Signal Processing and its Applications (on CDROM), Brisbane, Australia, December 2000.
- [25] A.J. Olievera, M.L. Nunes, and J.C. da Costa, Polysomnography in neonatal seizures, *Clinical Neurophysiol.*, 111, Suppl. 2, S74–S80, 2000.
- [26] A. Papandreou, S. Kay, and G. Faye Boudreax-Bartels, The use of Hyperbolic Time-Frequency Representations for Optimum Detection and Parameter Estimation of Hyperbolic Chirps, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, October 25–28, 1994, pp. 369–372.
- [27] A. Papandreou, G.F. Boudreux-Bartels, and S.M. Kay, Detection and estimation of generalized chirps using Time-Frequency representations, in 28th Annual Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 1994, pp. 50–54.
- [28] M. Roessgen, A. Zoubir, and B. Boashash, Seizure detection of newborn EEG using a model based approach, *IEEE Trans. Biomed. Eng.*, 45(6), June 1998.
- [29] J.C. Rowe, G.L. Holmes, J. Hafford, S. Robinson, A. Philipps, T. Rosenkrantz, and J. Raye, Prognostic value of the electroencephalogram in term and preterm infants following neonatal seizures, *Electroencephalogr. Clin. Neurophysiol.*, 60, 183–196, 1985.
- [30] B. Samimi and G. Rizzoni, Time-Frequency Analysis for Improved Detection of Internal Combustion Engine Knock, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Philadelphia, October 25–28, 1994, pp. 178–181.
- [31] L.L. Scharf, *Statistical Signal Processing — Detection, Estimation, and Time Series Analysis*, Addison-Wesley, Reading, MA, 1991.
- [32] H.L. Van Trees, *Detection, Estimation, and Modulation Theory*, John Wiley & Sons, New York, 1968.
- [33] S.P. Varma, A. Papandreou-Suppappola, and S.B. Suppappola, Detecting Faults in Structures Using Time-Frequency Techniques, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (on CDROM), Salt Lake City, UT, May 2001.
- [34] J.J. Volpe, Neonatal seizures, in *Neurology of the Newborn*, W.B. Saunders, Philadelphia, 1995, pp. 172–207.

10

Quadratic Time–Frequency Features for Speech Recognition

James Droppo

Microsoft Research

Les Atlas

University of Washington

CONTENTS

10.1	Introduction	371
10.2	Background and Definitions	372
10.3	Class-Dependent Time–Frequency Representations	376
10.4	Illustrative Example: Underwater Transient Identification	380
10.5	Kernel Design with Fisher’s Discriminant Metric	383
10.6	Application to Speech Recognition	385
10.7	English Alphabet Recognition	390
10.8	Conclusions	393
	References	394

10.1 Introduction

Quadratic time–frequency representations (TFRs) can be uniquely characterized by an underlying function called a kernel. In earlier time–frequency (TF) research, kernels for a number of properties, such as finite-time support and minimal quadratic interference, have been derived. Although some of the resulting TFRs may offer advantages for classification of certain types of signals, the goal of accurate classification is rarely an explicit goal of kernel design. For explicit classification, it is not necessarily desirable to represent the energy distribution of a signal in time and frequency accurately. In fact, such a representation may be in conflict with the goal of classification, which is to generate a TFR that is maximally separable from TFRs from different classes. (Here, the term *class* is used to refer to any grouping, arbitrary or otherwise, of similar data.) Given that many classification problems, such as

speech recognition, conventionally begin with test signals that represent examples of different classes, it is possible to design TFRs that intentionally highlight differences between classes.

We begin by reviewing the concept of autoambiguity functions. This review serves as the foundation for our class-dependent TFR approach, which, as shown through an example, designs a kernel that maximally separates classes of training signals. We describe two techniques for class-dependent kernel design where the latter technique is based on a standard linear discriminant.

To validate our approach, performance advantages are demonstrated on two speech recognition tasks: phoneme recognition and spoken English alphabet recognition. For both tasks our class-dependent TFR features, when added to a system containing more static conventional mel-frequency cepstral (MFC) features, markedly reduce the error rate in recognition of test signals.

10.2 Background and Definitions

We begin with a discrete, periodic sequence $x[n]$ with period N . Its discrete Fourier transform (DFT) is given by:

$$X[k] = \sum_{n=0}^{N-1} x[n] \exp(-2j\pi nk/N) \quad (10.1)$$

Including an arbitrary kernel function $\varphi[\eta, \tau]$, a general discrete autoambiguity representation is:

$$A[\eta, \tau] = \varphi[\eta, \tau] \sum_{n=0}^{N-1} x^*[n - \tau] x[n] \exp(2j\pi\eta n/N) \quad (10.2)$$

A two-dimensional (2-D) DFT relates this autoambiguity representation to a general TFR:

$$P[n, k] = \sum_{\eta=0}^{N-1} \exp(-2j\pi\eta n/N) \sum_{\tau=0}^{N-1} \exp(-2j\pi\tau k/N) A[\eta, \tau] \quad (10.3)$$

Members of this general class of DTF representations are generated with specific kernel functions $\varphi[\eta, \tau]$, including the discrete Rihaczek TFR [18], the discrete Margeneau–Hill distribution and the spectrogram.

Given a discrete time signal $x[n]$, operator theory [15] dictates that an acceptable root distribution is the discrete time, discrete frequency Rihaczek representation [18], where $\varphi[\eta, \tau] = 1$ in Equation (10.2):

$$P_R[n, k] = x^*[n] \exp(2j\pi nk/N) X[k] \quad (10.4)$$

By taking the 2-D DFT, this Rihaczek representation is expressed as the specific autoambiguity function:

$$A_R[\eta, \tau] = \sum_{n=0}^{N-1} x^*[n] x[n + \tau] \exp(2j\pi n\eta/N) \quad (10.5)$$

Can a root representation other than the Rihaczek be used? Yes. Because one TFR can be derived from any other with application of the appropriate transforming kernel, any TFR related to the Rihaczek representation through an invertible kernel may serve as an initial, base representation. For our eventual purposes of classification, a class-discriminating kernel can vary with the base TFR chosen, due to the varying amounts of TF similarity between the signal classes.

The representation described by Equation (10.5) contains a lot of information about the signal $x[n]$. Along $\eta = 0$, the function $A_R[\eta, \tau]$ represents the stationary auto-correlation of the signal $x[n]$, with τ taking its familiar role as the time lag variable. Along $\tau = 0$, the function $A_R[\eta, \tau]$ represents the spectrum of the instantaneous energy of the signal $x[n]$. This leads us to refer to η as the modulation frequency variable, because it relates to frequency components of the envelope of the signal. In general, a point (η, τ) refers to how severely a specific correlation coefficient is modulated.

10.2.1 Example illustrations

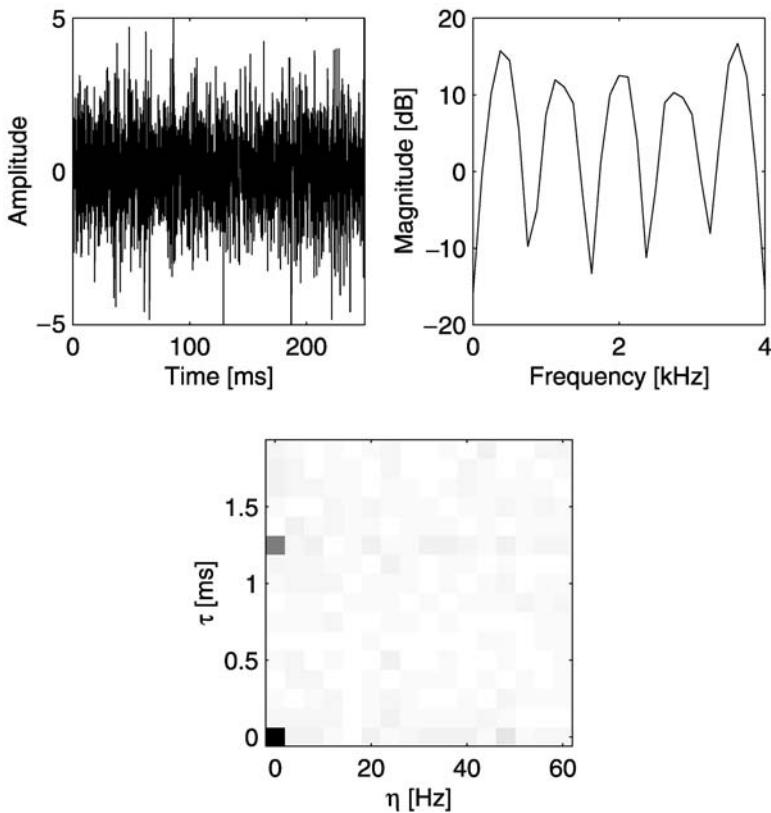
Although representations in time and frequency usually make intuitive sense to the untrained viewer, representations in η and τ , as in the autoambiguity function in Equation (10.2), tend to be less intuitive. Autoambiguity functions capture both static spectral information and evolutionary spectral information in a small number of points. We illustrate some key concepts with some synthesized test signals.

In general, stationary noise has no extent beyond $\eta = 0$, white noise has no extent beyond $\tau = 0$, stationary white noise is concentrated at the origin and more interesting structures appear when $\eta \neq 0$ and $\tau \neq 0$. Time modulations increase the function extent of a signal autoambiguity in η , and the coloration of spectra is responsible for extent in τ .

The lag variable τ is related to time correlations in the signal. In fact, along the lag variable τ for a kernel $\varphi[0, \tau] = 1$, Equation (10.2) reduces to the stationary autocorrelation of the signal:

$$A[0, \tau] = \sum_{n=0}^{N-1} x^*[n - \tau] x[n] \quad (10.6)$$

As a result, values along the $\eta = 0$ axis are interpreted as the autocorrelation of the signal $x[n]$. Consequently, any stationary signal can concentrate its energy within this region. [Figure 10.1](#) is one example of a signal with this behavior. The signal $x[n]$ was generated by passing stationary white Gaussian noise (WGN) through a linear

**FIGURE 10.1**

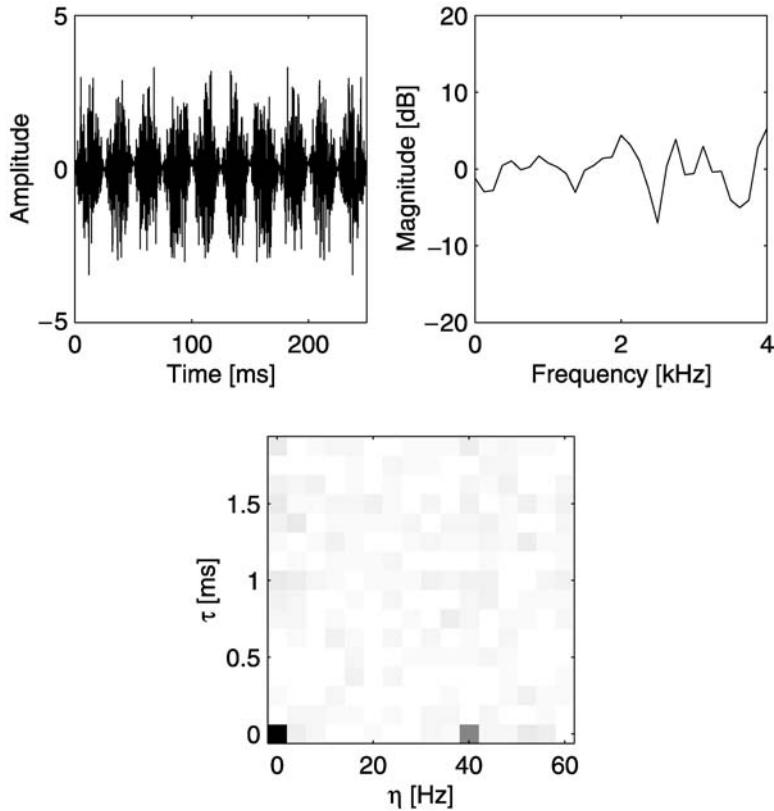
Colored Gaussian noise.

time-invariant finite impulse response filter. The filtering operation does not change the stationarity of the signal, although it may change its amplitude.

The modulation frequency variable η is related to time modulations of the signal. Along the modulation frequency variable η for a kernel $\varphi[\eta, 0]$, Equation (10.2) reduces to the DFT of the squared magnitude of the original signal:

$$A[\eta, 0] = \sum_{n=0}^{N-1} |x[n]|^2 \exp(2j\pi\eta n/N) \quad (10.7)$$

As a result, values along the $\tau = 0$ axis show how the modulation “envelope” of the signal $x[n]$ is changing with time. A signal with little correlation between its samples can concentrate its energy in this region. [Figure 10.2](#) is one example of a signal with this behavior. The signal was generated by modulating stationary WGN with a $\cos(t)$ function. This process does not change the correlation of the individual samples, but does make the signal quite nonstationary.

**FIGURE 10.2**

Modulated white Gaussian noise.

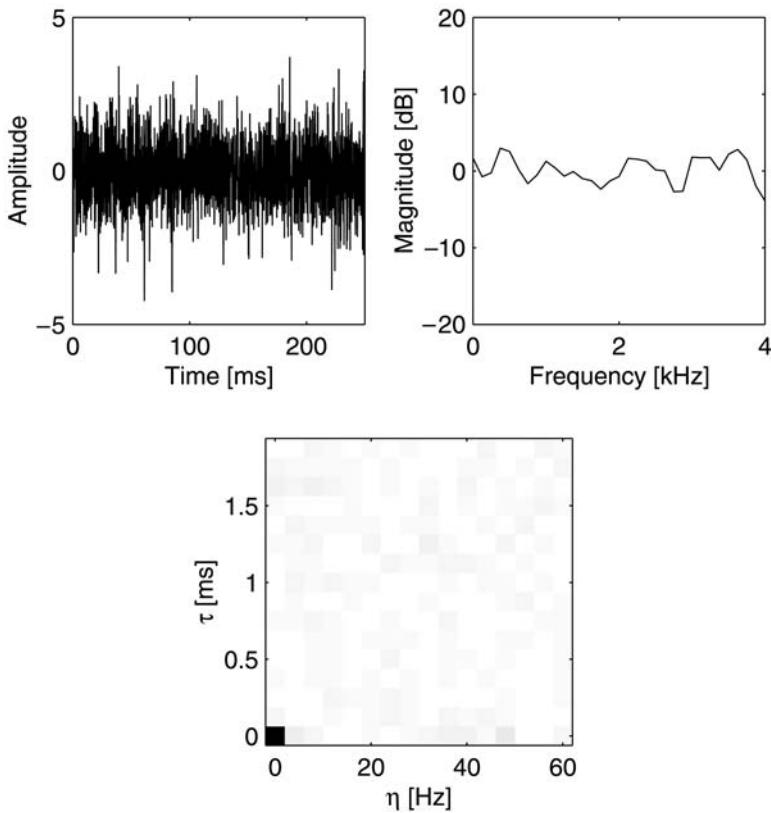
If a signal is both stationary and white (uncorrelated), such as stationary WGN, it can concentrate all its energy at the origin. Such a case is shown in Figure 10.3.

More generally, one can interpret each sample in the (η, τ) plane as the output of a nonstationary sinusoidal filter bank, where distance from the origin is analogous to bandwidth, and the angle from the η axis represents the linear chirp rate of the filter bank. To see this, consider a kernel function:

$$\varphi[\eta, \tau] = \delta[\eta - a] \delta[\tau - b] \text{ where} \\ \delta[n] = \begin{cases} 1 & n = 0 \\ 0 & \text{otherwise} \end{cases} \quad (10.8)$$

In the (n, k) plane, this kernel becomes a complex exponential, with parameters controlled by a and b :

$$\Phi[n, k] = \exp(2j\pi(an + bk)/N) \quad (10.9)$$

**FIGURE 10.3**

White Gaussian noise.

10.3 Class-Dependent Time–Frequency Representations

Most previous work in TF representations concentrated on generating images that humans could examine and extract relevant information. This research tended to focus on developing representations that had desirable properties. These properties included, but were not limited to, satisfying marginal equations — strictly positive or real — and having reduced cross terms [4], resulting in representations such as the Rihaczek, Wigner–Ville (WV), Choi–Williams (CW) and Zhao–Atlas–Marks (ZAM) (cone kernel), among others [4]. Generally, these properties were chosen because they produced visually appealing representations or preserved certain information in the original signal.

The traditional disadvantage to using any of these representations as features for automatic classifiers is the deluge of information. Whereas stationary spectral features

produce vectors with at most a few dozen dimensions, a real length N signal can potentially produce a representation with $N(N + 1)/2$ linearly independent dimensions. For a 250-msec signal sampled at 16 kHz, this amounts to no less than 8 million dimensions.

There have been several attempts in the literature to reduce this information into salient features, including taking autocorrelation coefficients of the representation in time or frequency to achieve a time-shift or frequency-shift invariant representation of the data [20], and finding 2-D moments of the TF representation [19]. Although some of these representations may offer advantages in classification for certain types of signals, they cannot necessarily generalize to arbitrary discrimination and classification problems.

Class-dependent TF representations use a discriminative approach to kernel design. The design goal for class-dependent kernels is maximum separability of the classes, and therefore higher classification accuracy. Classification performance is the primary concern in the kernel design procedure.

If TFRs such as those alluded to earlier are used, we have to hope that if our signal is well represented, then it will be well classified. Though this seems like a reasonable proposition, a representation may well bear a great deal of information that all signals under study share as well as information unique to each individual example signal. Such information is irrelevant to the classification, but most past approaches leave it up to a subsequent classifier to “screen out” these details.

The class-dependent kernel method isolates that portion of the TF structure that is useful for classification. That is, when possible, a closed-form solution is derived for an optimally discriminating kernel. This signal is not signal dependent, but signal class dependent.

Other researchers are using this idea to generate TF representations, and excitement is building as experiments continue to show promising results. Those few methods that propose to optimize the kernel for classification constrain the form of the kernel to Gaussian functions with symmetries [6, 9]. Our approaches, as described next, are nonparametric and hence less constrained with symmetries.

This section presents an overview of the commonalities among different methods proposed. Following this is one section for each of two methods, starting with the more basic Euclidean distance method and concluding with an informative–theoretical measure for kernel design.

10.3.1 Class-dependent kernel method

Given a labeled set of data from several classes, we seek to generate a kernel function that emphasizes classification relevant details present in the representation. That is, we choose a single kernel function that forms the best separation between data from different classes.

Given a discrete-time signal $x[n]$, the general autoambiguity representation is:

$$A[\eta, \tau] = \varphi[\eta, \tau] \sum_{n=0}^{N-1} x^*[n] x[n + \tau] \exp(2j\pi\eta n/N) \quad (10.10)$$

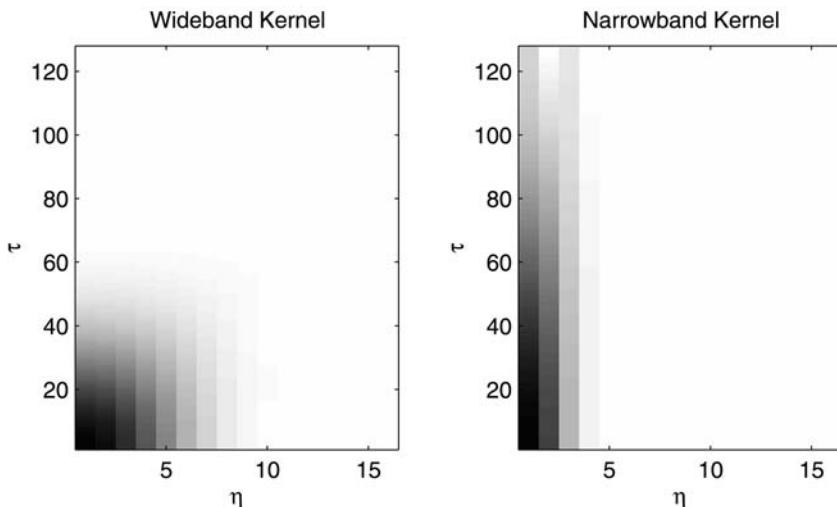


FIGURE 10.4
Spectrogram kernels.

In this section, the design freedom inherent in the kernel formulation is leveraged to reduce the dimensionality of the representation, without degrading classification performance.

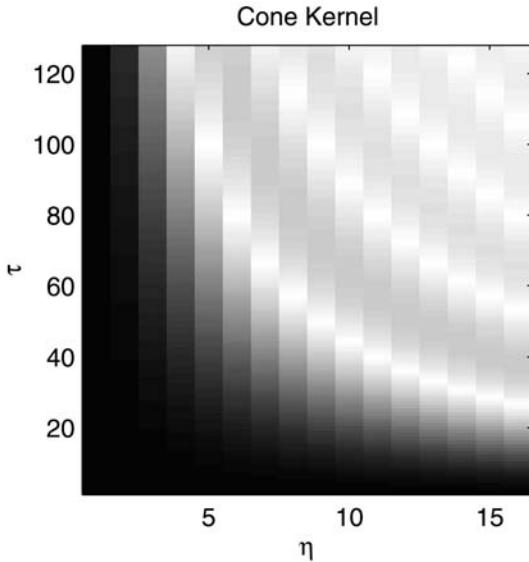
10.3.2 Traditional kernel design

One can get an intuition about kernel design by studying the properties of traditional kernels. Among the set of traditional noninformation preserving kernel functions are the spectrogram kernel and the cone kernel.

The spectrogram kernel is dependent on a window function $w[n]$, usually rectangular or tapered. The kernel is the conjugate of the base TFR of the window function:

$$\varphi[\eta, \tau] = \sum_{n=0}^{N-1} w[n] w^*[n + \tau] \exp(2j\pi\eta n/N) \quad (10.11)$$

A plot of this kernel for two different windows is shown in Figure 10.4. In general, these are low-pass kernels. That is, they tend to concentrate energy near the origin, and have limited extent in η and τ . Furthermore, the extent in η is proportional to the inverse of the extent in τ . This trade-off is the infamous TF resolution trade-off inherent in the spectrogram and is governed by the length of window functions. A larger extent in τ corresponds to a narrowband spectrogram. Fine frequency resolution is achieved, at the expense of the time resolution. Conversely, a larger extent in η corresponds to a wideband spectrogram. Fine time resolution results, but frequency resolution is sacrificed.

**FIGURE 10.5**

Cone (Zhao–Atlas–Marks) kernel.

The cone kernel [21], Figure 10.5, was formulated to overcome some of the problems in choosing a global TF trade-off. The extent in η is a function of the extent in τ . This means that features that rely on the fine time structure of the signal are estimated with greater precision in time, and features that correspond with longer lags are averaged over more data.

The cone kernel shares one important property with the spectrogram kernel. It, too, is low pass. In fact, the great majority of kernel design procedures produce kernels that have been low pass. In retrospect, this makes sense. The kernel is smoothing the representation. The information that is conventionally kept is the low-pass gross structure and not the fine structure, which may contain noise.

10.3.3 Kernel design with euclidean distances

Our first attempt focused on finding the kernel that maximally separates the auto-ambiguity representations $A_1[\eta, \tau]$ and $A_2[\eta, \tau]$ of two signals. This goal of maximizing the Euclidean distance between the two representations in the (η, τ) plane simplifies to:

$$\hat{\varphi}[\eta, \tau] = \underset{\varphi[\eta, \tau]}{\operatorname{argmax}} \sum_{\eta, \tau} |\varphi[\eta, \tau] (A_1[\eta, \tau] - A_2[\eta, \tau])|^2 \quad (10.12)$$

For Equation (10.12) to have a unique solution, the kernel is constrained to have unit energy:

$$\sum_{\eta, \tau} |\varphi[\eta, \tau]|^2 = 1 \quad (10.13)$$

In this case, the problem collapses to that of maximizing a real symmetrical quadratic form subject to the geometric constraint, Equation (10.13). The well-known solution to this problem is to choose the eigenvector associated with the largest eigenvalue. The eigenvectors of this problem are:

$$\varphi^{(i,j)}[\eta, \tau] = \delta[\eta - i]\delta[\tau - j] \quad (10.14)$$

with associated eigenvalues:

$$\lambda^{(i,j)} = |A_1[i, j] - A_2[i, j]|^2 \quad (10.15)$$

The best kernel is the eigenvector that is associated with the maximal eigenvalue. That is, the kernel consists of one nonzero point, at the location where $A_1[\eta, \tau]$ and $A_2[\eta, \tau]$ are maximally separated.

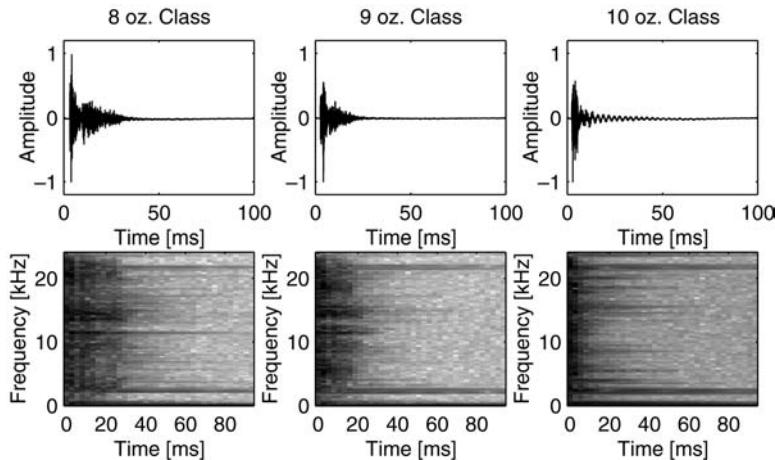
Extensions were made to this simple technique to make it more robust and to allow it to discriminate between two classes of signals instead of only a pair of signals. To make it more robust, we tried choosing not just the single point where $A_1[\eta, \tau]$ and $A_2[\eta, \tau]$ were maximally separated, but also the second or third top contenders. To handle classes of signals, the mean of each class was substituted for the representations $A_1[\eta, \tau]$ and $A_2[\eta, \tau]$.

Although neither of the preceding practical extensions was theoretically deep, we were able to capture one key idea from this method. Instead of classifying on the entire TF representation, we could do the classification in a subspace of the entire (η, τ) plane. This subspace could be chosen so that it was good for discrimination. This concept of class-dependent subspaces is illustrated in the following example.

10.4 Illustrative Example: Underwater Transient Identification

The first tests of our TF representation in a classification system were with simulated underwater transients. The data were intended to represent short duration, transient-like passive sonar acoustic events, with both time and frequency structure. Figure 10.6 shows three typical time series from our data set. Each signal was generated by suspending a 12-oz glass bottle from the ceiling, and tapping it with a small hammer. Several classes of data were generated by filling the bottle with water in 13 steps, adding about 1 oz at each step. As the water level increased, it was hoped that the changing resonance and impulse response of the bottle would make classification possible.

Many signals from each class were recorded on a digital audiotape. Later, the data were converted to analog and redigitized at a sampling rate of 48 kHz. From each class 30 consecutive seconds of data were digitized. All complete transients from each class were extracted with a 100-msec window. This window was placed to contain 10 msec of relative silence followed by the entire transient.

**FIGURE 10.6**

One example from each class, time series and spectrogram.

When played at the original sampling rate of 48 kHz, it is possible for a human listener to make rough classification estimates. When played at a rate of 8 kHz, a human listener can on average discriminate between adjacent water levels (classes), but not with complete accuracy.

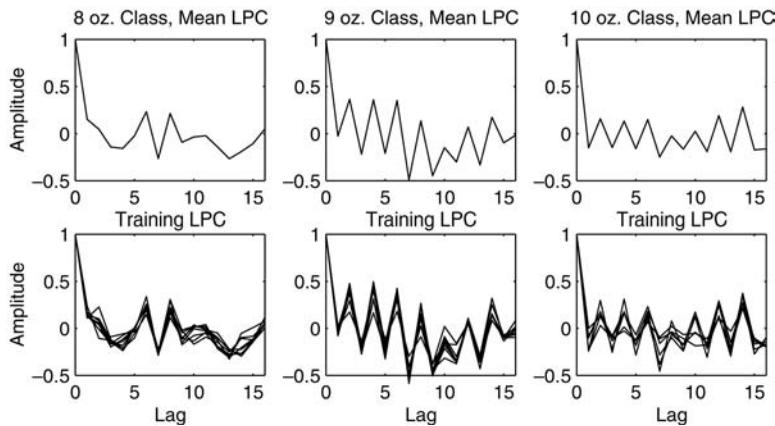
Three data sets were chosen for automatic classification, representing approximately 8, 9 and 10 oz. of water in the bottle. For each class 16 independent examples were available. Half were reserved for testing data, and half were used as training data.

Two classifiers were designed. The first classifier was based on the method of linear predictive coefficients (LPCs) or, equivalently, an autoregressive model. The second classifier was based on our class-dependent kernel, as described in the previous section.

10.4.1 Linear predictive coefficient method

Linear predictive coefficients (LPCs) are good at capturing the spectral content of a signal, but ignore any time domain information that may be present. LPCs assume that the observed signal was generated by passing either white noise or a periodic impulse train through a purely recursive discrete-time linear filter. The LPCs are an approximation of the coefficients of this filter.

For each training signal in each class, the first 16 LPCs were generated. Its average LPC vector then represented each class. Figure 10.7 shows the mean LPC vector for each of the three classes, along with plots of the vectors used for training. Each signal in the test set was classified into one of the three classes by forming a new LPC feature, and determining which class mean was closest (in the Euclidean sense) to the test data.

**FIGURE 10.7**

Linear predictive coefficients (AR(15)).

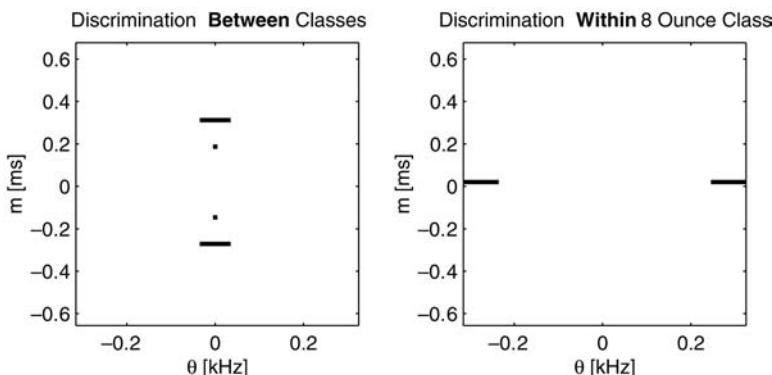
The LPC classifier was able to perfectly classify all the test data. We interpret this result as indicating that spectral information was sufficient for classification.

10.4.2 Class-dependent kernel method

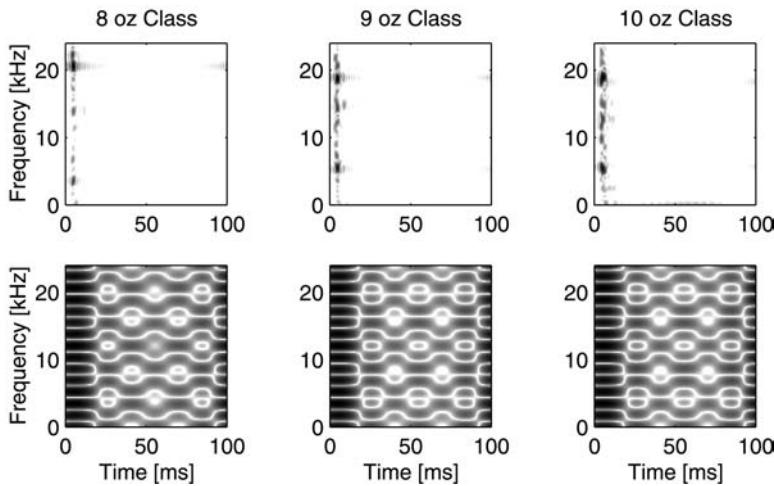
Theoretically, our class-dependent kernel should be able to find any time-variant behavior in the signals important for classification. The LPC contain only time-invariant spectral information, and do not have this luxury.

A kernel was generated to maximally separate the three classes from each other. That is, we would expect our class-dependent kernel to generate distributions for each class that are far apart according to our Euclidean distance measure.

The kernel found by our procedure is shown in the left half of Figure 10.8. By design, the kernel only takes on values of 1.0 and 0.0. The dark areas in the figure

**FIGURE 10.8**

Class-dependent kernels.

**FIGURE 10.9**

Example Rihaczek (top) and class-dependent (bottom) distributions.

represent the only nonzero regions in the kernel. This kernel focuses most of its energy along the $\eta = 0$ line in the ambiguity plane. Such a kernel tends to emphasize the time-invariant spectral content of a signal. This indicates that, for this classification task, stationary frequency information is more important. This agrees with our intuition that the difference between the classes is in the resonance of the bottle.

Figure 10.9 shows examples of TF representations before and after smoothing with a class-dependent kernel, for one signal in each class. From these images, it is apparent that the kernel emphasizes the frequency differences while at the same time smoothing the signals along the time axis.

Our class-dependent kernel method was also used to generate a kernel that would optimally discriminate between signals within a single class. That is, the training data consisted of each signal from the training data for one class, separately. The resulting kernel is reproduced in the right half of Figure 10.8. It has a concentration of energy along the $\tau = 0$ axis. In contrast to our previous example, this kernel tends to emphasize the time domain information in the signal. This indicates that within the class, the spectral information was uniform and not useful for classification. The relevant information exists mainly in the time domain. To summarize, the only within-class differences were the way the bottle was tapped from signal to signal, a difference which class-dependent kernels could, if desired, be designed to highlight.

10.5 Kernel Design with Fisher’s Discriminant Metric

A more sophisticated design goal for the class-dependent kernel is to find a subspace of the entire TF representation space, in which the distribution of the training data is

suitable for classification. Given a set of training examples $A_i^{(c)}$, indicating training example number i from class c , this subspace can be identified via a discrimination metric.

The previous Euclidean classifier defined by Equation (10.12) implicitly chooses its subspace using a discrimination metric computed from the means of each class:

$$D[\eta, \tau] = |\mu_1[\eta, \tau] - \mu_2[\eta, \tau]|^2 \quad (10.16)$$

In its place, a linear Fisher's discriminant can be used to identify dimensions suitable for classification [1, 14]. For N classes, and for each unique (η, τ) point, a Fisher's linear discriminant is computed. This computation is based on the mean $\mu_i[\eta, \tau]$ and variance $\sigma_i^2[\eta, \tau]$ of each class i and each coordinate $[\eta, \tau]$.

$$D[\eta, \tau] = \frac{\sum_{i=1}^N \sum_{j=i+1}^N |\mu_i[\eta, \tau] - \mu_j[\eta, \tau]|^2}{\sum_{i=1}^N \sigma_i^2[\eta, \tau]} \quad (10.17)$$

The numerator of Equation (10.17) is similar to Equation (10.16), and measures the variance of the data between classes. The denominator of Equation (10.17) measures the variance of the data within each class. Large discriminant values indicate dimensions along which the training data have large variances between classes, and small variances within each class.

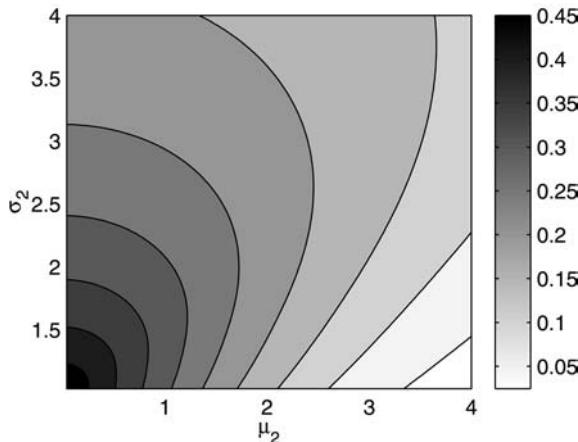
This $D[\eta, \tau]$ value is then used to rank order the points, from highest to lowest. In theory, large values can occur at coordinates where classes have both a large separation in the mean and a low within-class variance. Once a task-dependent subspace has been defined, a simple classifier can be built within that subspace.

As more points are used, the classes are modeled in a larger space. As a result, they are allowed to become farther apart. Unfortunately, as more points are added, the distributions of the classes are allowed to increasingly overlap. With the first few points, the former effect dominates. As more points are added, the rate at which new information arrives decreases.

The former effect dominates as the first few dimensions are added, and then the latter effect dominates as more points are added. The key here is to add dimensions that enhance the separation of the data, while minimizing the overlap, or variance, of the classes. This approach has been successfully applied to radar transmitter identification [8], helicopter gearbox fault classification [8] and speech phoneme discrimination [7].

A kernel function can be specified in either plane. The class-dependent kernel $\varphi[\eta, \tau]$ operates as a multiplicative mask on $A[\eta, \tau]$ in the autoambiguity plane. Its 2-D DFT $\Phi[n, k]$ operates convolutionally on the discrete time, discrete frequency Rihaczek TF representation $P_R[n, k]$. Namely, the class-dependent kernel smooths the TF representation to optimally (in the sense of maximizing Fisher's linear discriminant) separate the classes.

One problem observed in the use of Fisher's discriminant metric, is that it does not fully capture the merit of the individual dimensions chosen. [Figure 10.10](#) illustrates

**FIGURE 10.10**

Pairwise Gaussian probability of error. The gray scale shows a range from zero probability of error (white) to almost a chance level (black).

the probability of error of choosing between two Gaussian random variables. The first random variable is assumed to have zero mean and unit variance. (This assumption is without a loss of generality because if it were not the case, a linear scale and time shift would make it so.) The second random variable has the mean μ_2 and a variance σ_2 , which are axes of Figure 10.10. Figure 10.10 also shows the probability of error, given that the observation was equally likely to have come from only one of the two Gaussian sources.

Figure 10.10 can be interpreted as one of merit for choosing the (x, y) pair to discriminate the classes. Regions of low probability are to be preferred over regions of higher probability. If all pairs of variables fall such that $\mu > \sigma$, Fisher's metric ranks them similarly to using Figure 10.10. However, the two metrics disagree in the region $\mu < \sigma$. In particular, two random variables with the same mean can be more discriminable than two whose means differ. Fisher's discriminant metric ignores this effect.

10.6 Application to Speech Recognition

This section illustrates how the theoretical results and approach presented earlier can be applied to acoustic signal classification, in general, and to speech recognition, in particular. Two speech recognition tasks are presented. In the first section that follows some results are presented on isolated phone recognition. The classifier's task is to discriminate among several sets of phones, extracted from continuously spoken English. The TF features do not perform as well as traditional mel frequency

cepstral coefficient (MFCC) features, but when both types of features are combined, the overall classification rate is better than either system alone.

The next section extends this result to classification of spoken letters taken from the American English alphabet. This task is more difficult than isolated phone recognition, because phone boundaries are unknown in the test set, and a more sophisticated method of combining the features must be employed. Again, when both types of features are combined, the overall classification rate is substantially better than either system alone.

10.6.1 Isolated phone recognition

For the isolated phone recognition task, individual phones were extracted from continuous speech, and subsequently classified. The phones were grouped into several confusable sets, where within each set phone-based speech recognition systems usually experience confusion. This allowed us to experiment with small data sets and classification problems, which while fairly well labeled, were still known to cause problems with conventional systems.

All these phone recognition experiments were performed on data from the TIMIT corpus. This corpus contains continuous speech from 460 native English speakers, organized into 8 dialects.

In the TIMIT corpus, the core test set consists of 24 speakers, 2 males and 1 female from each of the 8 dialect regions. There are a total of 168 speakers in the full test set, and a total of 462 speakers in the training set.

10.6.2 Mel-frequency cepstral coefficients and hidden markov model phone recognition system

A conventional speaker-independent speech recognition system was constructed with the HTK modeling tool kit. This system provided a baseline for all improvements and modifications. This baseline system used standard MFCC. The coefficients were generated with a window length of 25 msec and a window skip of 5 msec. The static MFCC and energy terms were supplemented with standard δ and $\delta-\delta$ coefficients, for a total of 36 coefficients. The acoustic models consisted of the 46 context-independent phones.

Three-state hidden Markov models (HMMs) were used to classify the phonemes. Each HMM state contained an eight-mixture diagonal covariance continuous density Gaussian probability distribution function. Initial HMM parameters were estimated from a subset of the training data using the tool **HInit**. Each state initially consisted of one Gaussian mixture with a diagonal covariance.

Embedded training was performed with the tool **HERest**. The entire set of training data updates the set of models at once using Baum–Welch reestimation. New Gaussian mixtures were added one at a time, followed by two passes of embedded training.

By using a fully trained model, isolated phone recognition can be performed on segments extracted from the full test set, or continuous phone recognition can be performed on entire utterances.

TABLE 10.1

Vowel Performance of the HMM Recognition System, with Confusion Matrix and Error Rate per Vowel

	aa	ae	ah	ao	ax	ay	Error Rate (%)
aa	504	29	42	116	5	150	40.4
ae	20	634	26	2	0	90	17.9
ah	143	102	305	47	36	227	64.5
ao	125	1	10	581	0	44	23.6
ax	89	44	292	82	770	158	46.3
ay	44	13	18	2	0	609	11.2

The phone, word and sentence-level recognition rate for the system was comparable with similar systems referenced in the literature. Lee and Hon [12] reported a phone recognition rate of 58.77% using context-independent (CI) phones and no language model. Our baseline system, with the parameters cited earlier, has a phone recognition rate of 52.64%.

Tables 10.1 and 10.2 show the confusion matrices for two confusable test sets. Here, a set of vowels and a set of consonants were extracted from the test set. A fully trained model was used to perform isolated recognition on each set, constrained to choose within the confusable set.

Each row of the tables corresponds with the correct transcription, and each column corresponds with the recognizer output. For example, 29 utterances of /aa/ were misrecognized as /ae/, and the overall error rate for that phone was 40.4%.

As expected, this baseline system already performed reasonably well. The overall error rate within this set is 36.5%. Among the vowels chosen, the pair {/aa/, /ao/} was the most troublesome.

The average error rate within the set of consonants (Table 10.2) is 24.6%, considerably lower than for the vowels. It is reasonable to assume that this is due, at least in

TABLE 10.2

Consonant Performance of the HMM Recognition System, with Confusion Matrix and Error Rate per Consonant

	b	d	g	k	p	t	Error Rate (%)
b	557	33	47	3	218	28	37.1
d	24	404	80	45	39	249	52.0
g	10	18	276	120	12	16	38.9
k	3	6	37	1022	42	94	15.1
p	25	5	0	17	840	70	12.2
t	1	59	6	41	54	1206	11.8

part, to the fit of the model to the data. The consonants generally exhibit piecewise stationary behavior, whereas the vowel sounds can continuously change over time.

10.6.3 Time–frequency isolated phone recognition

As a first step toward building a full recognition system, the class-dependent technique was explored on a less complex task. The recognition task was reduced to isolated recognition of phones extracted from continuous speech, so that it would be easier to examine the behavior of the new techniques.

The class-dependent phone recognition system, by itself, did not produce stellar results. However, in the next section, it is shown that the information in the class-dependent TF representation can be used to supplement the preceding MFCC features to build a recognition system that outperforms one built on either feature set in isolation.

The class-dependent system was built using custom software, written in the C programming language, glued together with MATLABTM script files. The only features used in this system were TF representations. These representations were generated with a window length of 250 msec and a window skip of 50 msec. Increasing the window length and skip by a factor of ten was possible, and desirable, because of the nature of the feature set. A subset (mask) of the ambiguity plane representation for a particular task was chosen as described previously, using the class-conditional method. A single mixture Gaussian with full covariance was used to classify the data.

It is notable that we get good results even when no state information is present in the classifier. We would expect the quality of results to be directly related to the stationarity of the actual phone.

Recognition performance was tested on several separate discrimination tasks, as summarized in Table 10.3. The goal was to show that the class-dependent kernel functioned appropriately when presented with actual speech data, and that the TF representations contained enough information to build a speech recognition system.

A comparison of performance for three of the classes is presented in [Table 10.4](#). The HMM recognition system outperformed the system built on class-dependent TF representations in two out of the three test cases.

For the vowels test, the class-dependent kernel chose autocorrelation coefficients along the $\eta = 0$ axis. This is a confirmation of the knowledge that these sounds only

TABLE 10.3
Confusable Acoustic Classes for Phone Classification

Class	Members
Vowels	/aa/, /ae/, /ah/, /ao/, /ax/, /ay/
Glides	/l/, /r/, /w/
Consonants	/b/, /d/, /g/, /k/, /p/, /t/
Consonant–vowel pairs	/ae/, /n/, /ae/-/n/, /n/-/ae/

TABLE 10.4
Isolated Phone Error Rate (95% Confidence)

	HMM Recognizer (%)	Class-dependent TFR (%)
Vowels	36.9 (± 0.9)	49.7 (± 1.0)
Glides	15.5 (± 1.0)	25.9 (± 1.2)
C-V pairs	36.3 (± 2.8)	31.5 (± 2.7)

differ in their stationary spectra, but were found automatically. The class-dependent result, therefore, is identical to training on four fixed lags of the autocorrelation function. The HMM recognizer produces approximately 26% fewer errors than the class-dependent TFR.

For the glides test, the class-dependent kernel chose both stationary autocorrelation coefficients and points along $\eta = 1$, which correspond to modulations around 6.7 Hz. The class-dependent kernel classifies the data according not only to stationary spectral information but also to how that spectrum changes with time. For this test, the HMM recognizer produces approximately 40% fewer errors than the class-dependent TFR.

In the third test, discrimination among /æ/, /n/, /æ/-/n/ and /n/-/æ/, the class-dependent kernel did not use any stationary spectral information. It used 5 points in the ambiguity plane, at τ values that corresponded to 6.7 and 13.3 Hz. In this case, our TF approach outperformed the HMM- and MFCC-based recognizer. The class-dependent TFR produced approximately 13% fewer errors than the class-dependent TFR. The confidence intervals for this test are greater than the first two, because there are fewer occurrences of specific phone pairs than individual phones in the database.

10.6.4 Hybrid system phone recognition

A hybrid system was constructed that combined the outputs of both the MFCC and class-dependent TF representation features. First, the set of recognition likelihoods for each model were normalized across the set of all possible utterances:

$$p_{\text{norm}}[k] = \frac{p[k]}{\sum_l p[l]} \quad (10.18)$$

Next, a mixing parameter α was used to combine the likelihoods from the two classifiers. The mixing parameter was chosen to minimize the error after resubstitution with the training set:

$$p_{\text{mix}}[k] = p_{\text{norm}}^{(1)}[k] + p_{\text{norm}}^{(2)}[k] \quad (10.19)$$

The results of this mixing were quite good, and motivated us in building a larger speech recognition system. For vowels, the errors were reduced by approximately 6%.

For consonants, the error rate actually went up slightly with a deterioration of consonant performance of much less than 1%.

10.7 English Alphabet Recognition

The next set of tasks involved recognizing isolated American English letters. This task was chosen for two reasons. First, the recognizer could employ a flat language model, and the recognition accuracy is a function of only the acoustic modeling. Second, smaller numbers of acoustic classes and amounts of data could result in faster turnaround on experiments. Both of these reasons make English alphabet recognition an ideal task for prototyping and demonstrating the power of the class-dependent TFR.

The ISOLET corpus [5] is used for the experiments. The task consists of recognizing isolated letters drawn from the 26-letter American English alphabet. This corpus contains a total of 150 speakers, which are conventionally divided into a set of 120 training speakers and 30 testing speakers. Each speaker is associated with 52 utterances, consisting of 2 utterances of each letter. This makes for a total corpus size of 7800 utterances, equally divided among the letters of the alphabet and the gender of the speaker.

Cole reported a 4.68% speaker-independent error rate (73/1560 errors) on this task [5]. The system consisted of a rule-based segmenter and neural network classifier. Other authors have also attacked this task, with varying results. The previous best appears to be by Loizou and Spanias, with a speaker-independent error rate of 2.63% (41/1560 errors) [13]. This system incorporated context-dependent phone HMMs, as well as new feature representations for improved stop consonant discrimination, and subspace approaches for improved nasal discrimination.

Microsoft's WhisperTM speech recognizer was used as the baseline system in our experiments. It has the flexibility to take different acoustic and language models. Here we used a set of HMMs with continuous-density output distributions consisting of Gaussian mixture densities. A more complete description of the Whisper speech recognition system can be found in [10].

Three separate model configurations were tested. All three were composed of speaker-independent phone models. The simplest configuration used CI and gender-independent (GI) models. We made the system progressively more complex by first adding context-dependent (CD) and then gender-dependent (GD) models. A standard dictionary was used to map letters to corresponding phone models. There were 27 CI phones in this dictionary, and 53 CD phones.

MFCCs were chosen as the feature representation for the baseline system. It has very good performance and is currently widely used in many state-of-the-art speech recognition systems. The static MFCC coefficients were generated with a window length of 25 msec and a window skip of 10 msec. These are augmented by full δ and half of the $\delta-\delta$ coefficients, for a total of 33 coefficients.

The class-dependent TF features were used both as the only features for recognition, and as an alternative feature stream during decoding. These class-dependent features were generated with a window length of 100 msec and a window skip of 10 msec. This is a shorter window skip than we used in the isolated phone recognition system. The change was necessary to synchronize the MFCC and TF features in time. The TF features were chosen as described previously, using the class-dependent method.

A scheme for multiple feature decoding was incorporated to improve recognition performance in the hybrid system with class-dependent and MFCC features. During training, one set of models is trained for each feature stream. In the decoder, the search space was augmented with an additional feature stream dimension. The decoder considered all the features and found the word sequence with maximum likelihood over all models, states and feature streams. For more details see [11].

We only deal with isolated recognition in the ISOLET task. Given that there is only one letter in each utterance, the framework can be simplified to:

$$L^* = \operatorname{argmax}_L \sum_{i \in \Psi} p(L|F_i) p(F_i|A) \quad (10.20)$$

In this equation, Ψ is the set of all feature streams, A is the waveform data for speech signal and F_i is the feature associated with a given feature stream. The term $p(F_i|A)$ serves as a weighting factor to weight contributions from different feature streams. In our hybrid system with multiple features, the weights are class dependent.

10.7.1 Experimental results

Recognition was performed using MFCC features only, using class-dependent TF features only, and both features together in a hybrid system.

For each system, the number of parameters (Gaussian mixtures) was adjusted to maximize recognition accuracy. The goal was not to compare systems with equivalent numbers of parameters, but instead to demonstrate the maximum performance of the system for a given feature set.

The recognition systems based solely on MFCC features performed as well as, or better than, the best published results on this task. The results are summarized in Table 10.5. The recognition systems used CI or CD phones, and either GI or GD models.

TABLE 10.5
Baseline MFCC Performance

System	Mixtures	Errors	Utterances	Error Rate (%)
CI–GI	12	79	1560	5.06
CD–GI	8	50	1560	3.21
CD–GD	8	38	1560	2.44

TABLE 10.6

Class-Dependent TFR Performance

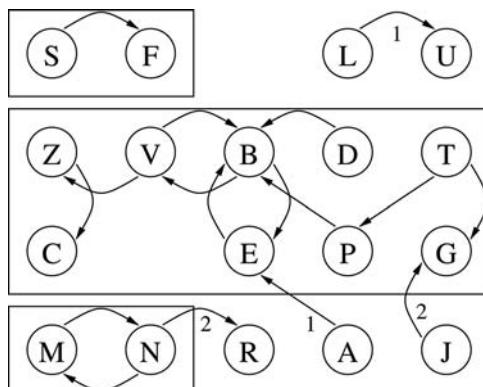
System	Mixtures	Errors	Utterances	Error Rate (%)
CI–GI	12	73	1560	4.68
CD–GI	8	64	1560	4.10
CD–GD	8	58	1560	3.72

The baseline system errors (Figure 10.11) were consistent with those expected. Out of 38 errors, 32 occurred within 3 confusable letter classes: the nasals {M, N}, the fricatives {S, F} and the infamous E-set {B, C, D, E, G, P, T, V, Z}. If all these within-class errors were eliminated, only six errors would remain, a significant improvement over even the best result.

Features for the class-dependent TFR system were designed to eliminate the within-class confusion typical of the MFCC system. That is, the discriminant was computed to discriminate within classes of phone models {/s/, /f/}, {/m/, /n/} and {/b/, /s/, /d/, /gh/, /p/, /t/, /v/, /z/}.

The class-dependent TFR recognition system accuracy fell between the initial results by Cole et al.[5] and the result by Loizou and Spanias [13]. The actual error rates for this system are presented in Table 10.6. Although the class-dependent TFR system did a better job of not confusing letters within confusable sets, too many between-class errors were introduced.

It is notable that the performance increase in moving to a CD and GD system is not nearly as great as it was for the MFCC system. The current hypothesis is that because the class-dependent TFR feature selection algorithm is both CI and GI,

**FIGURE 10.11**

MFCC system substitution errors. Each box represents a confusable set.

TABLE 10.7

Hybrid MFCC and Class-dependent TFR System Performance

System	Errors	Utterances	Error Rate (%)
CD TFR	58	1560	3.72
MFCC	38	1560	2.44
Hybrid	25	1560	1.60
Improvement of hybrid over MFCC		13 Fewer	Dropped 34.2

context and gender information is smoothed in the feature space, and not recoverable by modifying the stochastic model.

10.7.2 Hybrid system

As mentioned previously, if all the within-class errors were eliminated, but the between-class errors remained, the performance of the MFCC system would be phenomenal. The hybrid system we built is a two-stage recognition system that refines the classification accuracy within each confusable class by using multiple features. This two-stage approach has the potential of eliminating within-class errors, but neither increases nor decreases the between-class errors.

The first stage uses only MFCC features for recognition. As discussed earlier, this stage makes few between-class errors. The second stage relabels data within each class using class-dependent TFR features tuned for discrimination within that class together with MFCC features. It is built from the models generated for the CD, GD, MFCC and class-dependent TFR systems. The decoder for this stage used the multiple feature stream approach described previously. The performance of the hybrid system is shown in Table 10.7.

Most of the improvement came from eliminating within-class errors in the fricative set and the E-set. The hybrid system, as expected, performed better than either the MFCC or the class-dependent TFR systems in isolation. Overall, the error rate was reduced by 34.2%. Our error rate of 1.60% has almost 40% fewer errors than the previous best-published result on the ISOLET task [13].

10.8 Conclusions

By using the concepts of operator theory, we have been able to forge a direct connection between a discrete, finite-length input signal, and its discrete-time, discrete-frequency TFR. With our approach, we represent each TFR as the 2-D circular convolution of a kernel with the Rihaczek distribution of the signal.

The class-dependent kernel was developed to extract from the TF representation those dimensions useful for classification. It is important to note that the kernel we obtain for optimal separation maximizes the TF difference given the original distribution (the Rihaczek). If the two signal classes have very dissimilar Rihaczek TFRs, then our method can find very little room for improvement over the already easy to separate classes.

As shown, these techniques can be combined to provide TF features for input to a classifier. These features are not limited, like conventional cepstra, to static spectral estimates and their time differences. The design approach also intentionally suppresses extraneous information.

We also show potential application of class-dependent TF features to speech recognition. It can also be used together with traditional features such as MFCC to improve performance. Experiments demonstrate that class-dependent TF features, in combination with the multiple feature stream decoder, are very effective, especially for discrimination within classes. Together with MFCC, we achieve a 1.6% error rate, very high performance on ISOLET alphabet recognition and significantly higher than MFCC alone.

References

- [1] L. Atlas, J. Droppo, and J. McLaughlin, Optimizing time-frequency distributions for automatic classification, in Proceedings of the SPIE, 3162, 161–171, 1997.
- [2] T.A.C.M. Classen and W.F.G. Mecklenbräuker, The Wigner distribution-a tool for time-frequency signal analysis. Part I. Continuous-time signals, *Philips J. Res.*, 35, 217–250, 1980.
- [3] L. Cohen, Generalized phase-space distribution functions, *J. Math. Phys.*, 7, 781–786, 1966.
- [4] L. Cohen, *Time-Frequency Analysis*, Prentice Hall Signal Processing Series, Englewood Cliffs, NJ, 1995.
- [5] R. Cole, M. Fanty, Y. Muthusamy, and M. Gopalakrishnan, Speaker-independent recognition of spoken English letters, in Proceedings of the International Joint Conference on Neural Networks, 2, 45–51, 1990.
- [6] M. Davy and C. Doncarli, Optimal Kernels of Time-Frequency Representations for Signal Classification, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, PA, October 1998, pp. 581–584.

- [7] J. Droppo and L. Atlas, Application of Classifier-Optimal Time-Frequency Distributions to Speech Analysis, in Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, Pittsburgh, October 1998, pp. 585–588.
- [8] B. Gillespie and L. Atlas, Optimizing time-frequency kernels for classification, *IEEE Trans. Signal Process.*, 49(3), 485–496, 2001.
- [9] C. Heitz, Optimum time-frequency representations for the classification and detection of signals, *Appl. Signal Process.*, 3, 124–143, 1995.
- [10] X. Huang, A. Acero, F. Alleva, M.Y. Hwang, L. Jiang, and M. Mahajan, Microsoft Windows Highly Intelligent Speech Recognizer: Whisper, in Proceedings of the 1995 IEEE International Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, May 1995.
- [11] L. Jiang and X.D. Juang, Unified decoding and feature representation for improved speech recognition, in Proceedings of the Eurospeech 99, 3, 1331–1334, 1999.
- [12] K.-F. Lee and H.-W. Hon, Speaker-independent phone recognition using hidden Markov models, *IEEE Trans. Acoust. Speech Signal Process.*, 37(11), 1641–1648, November 1989.
- [13] P. Loizou and A. Spanias, High-performance alphabet recognition, *IEEE Trans. Speech Audio Process.*, 4(6), 430–445, November 1996.
- [14] J. McLaughlin, J. Droppo, and L. Atlas, Class-dependent, discrete time-frequency distributions via operator theory, in Proceedings of the 1997 IEEE International Conference on Acoustics, Speech and Signal Processing, 3, 2045–2048, 1997.
- [15] S.B. Narayanan, J. McLaughlin, L. Atlas, and J. Droppo, An Operator Theory Approach to Discrete Time-Frequency Distributions, in Proceedings IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, 1997, pp. 521–524.
- [16] A.H. Nutall, Alias-free Wigner distribution functions and complex ambiguity functions for discrete-time samples, Naval Underwater Systems Center, April 14 1989, Technical report 8533.
- [17] M. Richman, T. Parks, and R. Shenoy, Features of a Discrete Wigner Distribution, in 1996 IEEE Digital Signal Process Workshop Proceedings, 1996, pp. 427–430.
- [18] A.W. Rihaczek, Signal energy distribution in time and frequency, *IEEE Trans. Inf. Theory*, 14, 369–374, 1968.
- [19] B. Tacer and P. Loughlin, Time-frequency based classification, in Proceedings of the SPIE — International Society Optical Engineering, 2846, 186–192, August 1996.

- [20] E.J. Zalubas, J.C. Oneill, W.J. Williams, and A.O. Hero, III, Shift and Scale Invariant Detection, in Proceedings of the 1997 IEEE International Conference on Acoustics, Speech and Signal Processing, Munich, Germany, April 21–24, 1997, pp. 3637–3640.
- [21] Y. Zhao, L.E. Atlas, and R.J. Marks, The use of cone-shaped kernels for generalized time-frequency representations of nonstationary signals, *IEEE Trans. Acoust. Speech Signal Process.*, 38, 1084–1091, 1990.