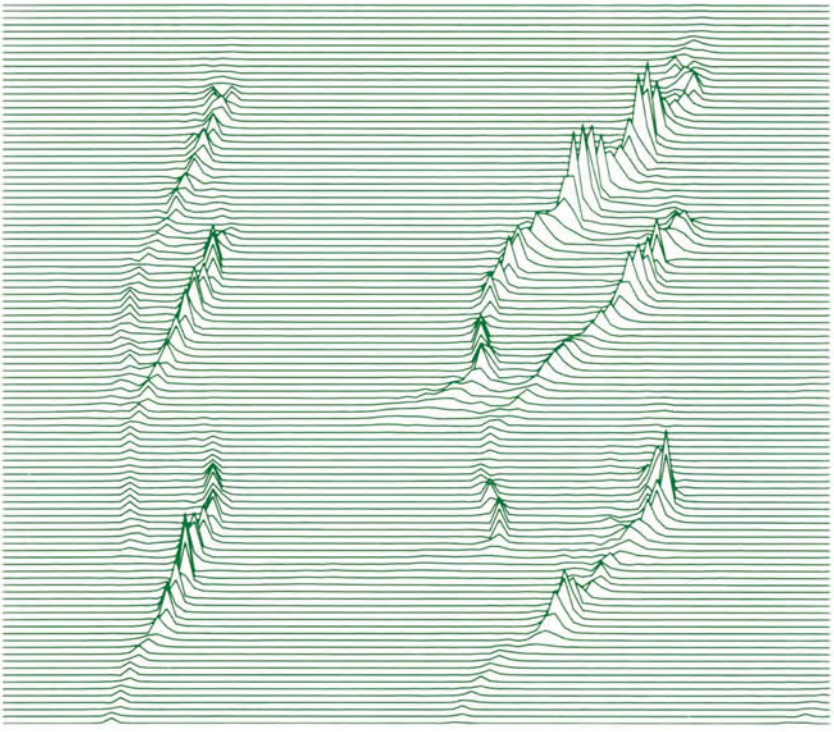
ELSEVIER

Time Frequency Signal Analysis and Processing

A Comprehensive Reference



Edited by
Boualem Boashash

Time Frequency Signal Analysis and Processing

A Comprehensive Reference

Full catalogue information on all books, journals and electronic products can be found on the Elsevier homepage at: http://www.elsevier.com

ELSEVIER PUBLICATIONS OF RELATED INTEREST

JOURNALS:

Advanced Engineering Informatics
Annual Reviews in Control
Automatica
Biomedicine and Pharmacotherapy
Computers and Electrical Engineering
Computers in Industry
Control Engineering Practice
Engineering Applications of AI
Image and Vision Computing
Journal of the Franklin Institute
Measurement
Mechatronics
Robotics and Autonomous Systems
Robotics and Computer-Integrated Manufacturing
Sensors and Actuators A: Physical
Sensors and Actuators B: Chemical
Signal Processing

Time Frequency Signal Analysis and Processing

A Comprehensive Reference

Edited by

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

2003

ELSEVIER

Amsterdam – Boston – Heidelberg – London – New York – Oxford
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo

ELSEVIER Ltd
The Boulevard, Langford Lane
Kidlington, Oxford OX5 1GB, UK

First edition 2003

Library of Congress Cataloging in Publication Data
A catalog record from the Library of Congress has been applied for.

British Library Cataloguing in Publication Data
A catalogue record from the British Library has been applied for.

ISBN: 0-08-044335-4

♾ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in Hungary.

To my sons

Issam Alam and Jamil Mehdi

and my late wife

Theresa Adele

This Page Intentionally Left Blank

Preface

Time-Frequency Signal Analysis and Processing (TFSAP) is a collection of theory and algorithms used for analysis and processing of non-stationary signals, as found in a wide range of applications including telecommunications, radar, and biomedical engineering. This book brings together the main knowledge of TFSAP, from theory to applications, in a user-friendly reference suitable for both expert and non-expert readers.

The **contents** of the book include:

1. a *comprehensive tutorial introduction* to TFSAP, accessible to anyone who has taken a first course in signals and systems;

2. more specialized theory and algorithms, concisely presented by some of the leading authorities on the respective topics; and

3. studies of *key applications*, written by leading researchers, showing how to use TFSAP methods to solve practical problems.

The **motivation** for producing this book was twofold:

- My original and widely used decade-old tutorial on TFSAP [1] needed updating in two respects. First, some of the advances of the last decade are sufficiently fundamental to warrant inclusion in an introductory treatment, while others are sufficiently important to demand coverage in any comprehensive review of TFSAP. Second, new applications have widened the range of disciplines interested in TFSAP, and thus reduced the common background knowledge that may be expected of readers. Part I of this book addresses these needs.

- The need for a standard language of discourse became apparent in 1990 while I was editing the 23 contributions to the first comprehensive book in the field [2]. These seminal contributions to TFSAP led to further developments throughout the 1990s, including some significant advances in practical methods suitable for non-stationary signals. These efforts continued apace as this book was being written. Such rapid progress produced a variety of new terminologies and notations that were in need of standardization and inclusion in an updated reference book.

The **organization** of this book uses five Parts, each Part including several Chapters, and each Chapter comprising several Articles. Part I introduces the basic concepts while Parts II to V cover more advanced or specialized areas.

Part I defines and explains the basic concepts of TFSAP, intuitively derives a variety of well-known time-frequency distributions (TFDs), and then reduces them to a common form. This leads to the general treatment of quadratic TFDs in Chapter 3, which should be regarded as the core of the book and as a prerequisite for the later chapters.

Part II gives more details on some fundamental topics of TFSAP, such as TFD design and signal analysis in the (t, f) plane.

Part III describes specialized techniques used in implementation, measurement and enhancement of TFDs.

Part IV presents the key statistical techniques for TFSAP of noisy signals, including a full treatment of detection and classification methods.

Part V describes a representative selection of TFSAP applications, encompassing telecommunications, radar, sonar, power generation, image quality, automotive applications, machine condition monitoring, and biomedical engineering.

Usability is enhanced by an updated *consolidated bibliography* (alphabetical by author) and a *two-level index* (which also serves as a dictionary of abbreviations).

Under the standard review procedure used for this book, each Article had two (usually external) reviewers concentrating on scientific rigor and accuracy, plus two anonymous internal reviewers concentrating on clarity and consistency.

Acknowledgments are due to a number of people who made possible the completion of this book. Foremost among them are my two sons, who aided me to continue this work during and after my wife's final illness, thus contributing to my sense of balance and purpose during this difficult period. I thank all authors and reviewers, and the organizers of the Special Sessions on TFSAP at ISSPA conferences, for their expertise, timely effort and professionalism, and for facilitating the exchange of ideas between contributors to this book. I thank my research students and the SPRC staff for valuable assistance. In particular, Gavin Putland assisted with the technical editing of portions of Part I and was responsible for the final mix-down of the authors' LaTeX™ and PostScript™ files.

Boualem Boashash,
Editor.

References

[1] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[2] B. Boashash, ed., *Time-Frequency Signal Analysis: Methods and Applications.* Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

Contents

List of Contributors

K. Abed-Meraim, Ecole Nationale Superieure des Telecommunication (Telecom Paris), Dept. TSI (Signal & Image Processing), 46 rue Barrault, 75634 Paris, France.

O. Akay, Dokuz Eylül University, Department of Electrical and Electronics Engineering, Izmir, Turkey.

T. Alieva, Technische Universiteit Eindhoven, Faculteit Elektrotechniek, 5600 MB Eindhoven, Netherlands.

M. G. Amin, Villanova University, Department of Electrical and Computer Engineering, Villanova, PA 19085, USA.

F. Auger, Centre for Research and Technology Transfer (CRTT), Boulevard de l'Universite, 44602 Saint Nazaire, Cedex 44602, France.

S. Aviyente, The University of Michigan, Department of Electrical Engineering and Computer Science, Ann Arbor, MI 48109, USA.

R. Baraniuk, Rice University, Department of Electrical & Computer Engineering, Houston, TX 77005-1892, USA.

S. Barbarossa, University of Rome "La Sapienza", INFOCOM Department, 00184 Rome, Italy.

B. Barkat, Nanyang Technological University, School of Electrical and Electronic Engineering, Singapore 63978.

V. Barroso, Instituto Superior Tecnico, Instituto de Sistemas e Robotica/DEEC, Lisboa, Portugal.

M. J. Bastiaans, Technische Universiteit Eindhoven, Faculteit Elektrotechniek, 5600 MB Eindhoven, Netherlands.

A. Beghdadi, L2TI-Institut Galilee, Universite Paris 13, 93430 Villetaneuse, France.

A. Belouchrani, Ecole Nationale Polytechnique, EE Department, Algiers, Algeria.

M. Benidir, LSS-Supelec, Universite de Paris-Sud, Gif-sur-Yvette, France.

J. Bertrand, LPTMC, University Paris VII, 75251 Paris, Cedex 05, France.

P. Bertrand, LPTMC, University Paris VII, 75251 Paris, Cedex 05, France.

B. Boashash, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Qld 4001, Australia.

J. F. Böhme, Ruhr-Universitaet Bochum, Department of Electrical Engineering and Information Science, 44780 Bochum, Germany.

G. F. Boudreaux-Bartels, University of Rhode Island, Department of Electrical and Computer Engineering, Kingston RI 02881, USA.

S. Carstens-Behrens, Robert Bosch GmbH, 70049 Stuttgart, Germany.

V. Chandran, EESE, Queensland University of Technology, Brisbane, Qld 4001, Australia.

E. Chassande-Mottin, Max Planck Institut fur Gravitationphysik, Albert Einstein Institut, D-14424 Golm, Germany.

P. Colditz, Perinatal Research Centre, Royal Women's Hospital, Brisbane, 4029, Australia.

I. Djurovic, Kyoto Institute of Technology, Department of Mechanical and Systems Engineering, Matsugasaki, Sakyo Kyoto 606-8585, Japan.

A. El-Jaroudi, University of Pittsburgh, Department of Electrical Engineering, Pittsburgh, PA 15261, USA.

M. K. Emresoy, University of Pittsburgh, Department of Electrical Engineering, Pittsburgh, PA 15261, USA.

B. G. Ferguson, Defence Science and Technology Organisation, Pyrmont, NSW 2009, Australia.

P. Flandrin, Laboratoire de Physique, (UMR 5672 CNRS), Ecole Normale Superieure de Lyon, 69364 Lyon Cedex 07, France.

G. J. Frazer, Defence Science and Technology Organisation, Surveillance Systems Division, Edinburgh, SA 5111, Australia.

W. M. Grady, The University of Texas at Austin, Department of Electrical and Computer Engineering, Austin, TX 78712-1084, USA.

F. Hlawatsch, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, A-1040 Vienna, Austria.

Z. M. Hussain, RMIT, School of Electrical and Computer Systems Engineering, Melbourne, Victoria 3001, Australia.

B.-G. Iem, Kangnung National University, Electronics Department, Kangwon-do 210-702, Republic of Korea.

R. Iordache, Signal Processing Laboratory, Tampere University of Technology, FIN-33101 Tampere, Finland.

D. L. Jones, University of Illinois at Urbana-Champaign, Department of Electrical and Computer Engineering, Urbana IL 61801, USA.

G. Jones, Sicom Systems Ltd., Fonthill, Ontario L0S 1E0, Canada.

S. M. Joshi, Lucent Technologies, Alameda CA 94501, USA.

V. Katkovnik, Kwangju Institute of Science and Technology, Department of Mechatronics, Puk-gu, Kwangju 500-712, Republic of Korea.

A. R. Leyman, Center for Wireless Communication, National University of Singapore, 10 Kent Ridge Crescent, Singapore, 119260.

A. R. Lindsey, U.S. Air Force Research Laboratory, Rome, NY 13441, USA.

Nguyen Linh-Trung, Ecole Nationale Superieure des Telecommunication (Telecom Paris), Dept. TSI (Signal & Image Processing), 46 rue Barrault, 75634 Paris, France.

K. W. Lo, Defence Science and Technology Organisation, Pyrmont, NSW 2009, Australia.

L. Marple Jr., ORINCON Corporation, San Diego, CA 92121, USA.

J. Mathew, CIEAM, Queensland University of Technology, Brisbane, Qld 4001, Australia.

G. Matz, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, A-1040 Vienna, Austria.

M. Mesbah, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Qld 4001, Australia.

J. M. Morris, University of Maryland Baltimore County, Computer Science and Electrical Engineering Department, Catonsville, MD 21250, USA.

P. M. Oliveira, Escola Naval, DAEI, Base Naval de Lisboa, Alfeite, 2800 Almada, Portugal.

P. O'Shea, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Qld 4001, Australia.

A. Papandreou-Suppappola, Telecommunications Research Center, Arizona State University, Department of Electrical Engineering, Tempe AZ 85287-7206, USA.

B. Picinbono, Laboratoire des Signaux et Systemes, Supelec, Ecole Superieure d'Electricite, 91192 Gif-sur-Yvette, France.

E. J. Powers, The University of Texas at Austin, Department of Electrical and Computer Engineering, Austin, TX 78712-1084, USA.

G. R. Putland, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Qld 4001, Australia.

S. Qian, National Instruments Corp., Austin TX 78759, USA.

A. M. Rao, University of Illinois at Urbana-Champaign, Digital Signal Processing Group, Urbana IL 61801, USA.

A. M. Sayeed, University of Wisconsin-Madison, Department of Electrical and Computer Engineering, Madison, WI 53706, USA.

Y. J. Shin, The University of Texas at Austin, Department of Electrical and Computer Engineering, Austin, TX 78712-1084, USA.

LJ. Stanković, University of Montenegro, Elektrotehnicki Fakultet, 81000 Podgorica, Montenegro, Yugoslavia.

S. Stanković, University of Montenegro, Elektrotehnicki Fakultet, 81000 Podgorica, Montenegro, Yugoslavia.

V. Sucic, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Qld 4001, Australia.

S. B. Suppappola, Acoustic Technologies, Mesa, AZ 85204, USA.

G. Tauboeck, Telecommunications Research Center Vienna (FTW), A-1040 Vienna, Austria.

A. J. van Leest, Technische Universiteit Eindhoven, Faculteit Elektrotechniek, 5600 MB Eindhoven, Netherlands.

W. J. Williams, The University of Michigan, Department of Electrical Engineering and Computer Science, Ann Arbor, MI 48109, USA.

X.-G. Xia, University of Delaware, Department of Electrical and Computer Engineering, Newark DE 19716, USA.

Y. Zhang, Villanova University, Department of Electrical and Computer Engineering, Villanova, PA 19085, USA.

Part I

Introduction to the Concepts of TFSAP

This Page Intentionally Left Blank

Chapter 1

Time-Frequency Concepts[0]

Time-frequency signal analysis and processing (TFSAP) concerns the analysis and processing of signals with time-varying frequency content. Such signals are best represented by a **time-frequency distribution (TFD)**, which is intended to show how the energy of the signal is distributed over the two-dimensional time-frequency space. Processing of the signal may then exploit the features produced by the concentration of signal energy in two dimensions (time *and* frequency) instead of only one (time *or* frequency).

The first chapter begins the introductory tutorial which constitutes Part I of the book. This tutorial updates the one given in [1] by including recent advances, refining terminology, and simplifying both the presentations of concepts and formulations of methodologies. Reading the three chapters of Part I will facilitate the understanding of the later chapters.

The three sections of Chapter 1 present the key concepts needed to formulate time-frequency methods. The first (Section 1.1) explains why time-frequency methods are preferred for a wide range of applications in which the signals have time-varying characteristics or multiple components. Section 1.2 provides the signal models and formulations needed to describe temporal and spectral characteristics of signals in the time-frequency domain. It defines such basic concepts as analytic signals, the Hilbert transform, the bandwidth-duration product and asymptotic signals. Section 1.3 defines the key quantities related to time-frequency methods, including the instantaneous frequency (IF), time-delay (TD) and group delay (GD).

[0]Author: **Boualem Boashash**, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Australia. Reviewers: K. Abed-Meraim, A. Beghdadi, M. Mesbah, G. Putland and V. Sucic.

1.1 The Need for a Time-Frequency Distribution (TFD)

The two classical representations of a signal are the time-domain representation $s(t)$ and the frequency-domain representation $S(f)$. In both forms, the variables t and f are treated as mutually exclusive: to obtain a representation in terms of one variable, the other variable is "integrated out". Consequently, each classical representation of the signal is **non-localized** w.r.t. the excluded variable; that is, the frequency representation is essentially averaged over the values of the time representation at *all* times, and the time representation is essentially averaged over the values of the frequency representation at *all* frequencies.

In the time-frequency distribution, denoted by $\rho(t,f)$, the variables t and f are *not* mutually exclusive, but are present together. The TFD representation is **localized** in t *and* f.

1.1.1 Representation of Three Real-Life Signals

The usefulness of representing a signal as a function of both time *and* frequency is illustrated by considering three signals of practical importance:

1. **Sinusoidal FM signal:** Monophonic television sound, like monophonic FM radio, is transmitted on a frequency-modulated carrier. If the audio signal is a pure tone of frequency f_{m} (the modulating frequency), then the frequency of the carrier is of the form

$$f_i(t) = f_{\mathrm{c}} + f_{\mathrm{d}} \cos[2\pi f_{\mathrm{m}} t + \phi] \tag{1.1.1}$$

   where t is time, $f_i(t)$ is the frequency modulation law (FM law), f_{c} is the mean (or "center") carrier frequency, f_{d} is the peak frequency deviation, and ϕ allows for the phase of the modulating signal. The amplitude of the carrier is constant.

2. **Linear FM signal:** Consider a sinusoidal signal of total duration T, with constant amplitude, whose frequency increases from f_0 to $f_0 + B$ at a constant rate $\alpha = B/T$. If the origin of time is chosen so that the signal begins at $t = 0$, the FM law may be written

$$f_i(t) = f_0 + \alpha t \; ; \quad 0 \le t \le T \; . \tag{1.1.2}$$

   In an electronics laboratory, such a signal is called a **linear frequency sweep**, and might be used in an automated experiment to measure the frequency response of an amplifier or filter. In mineral exploration, a linear FM signal is called a **chirp** or **Vibroseis** signal, and is used as an acoustic "ping" for detecting underground formations [2,3].

3. **Musical performance:** A musical note consists of a number of "components" of different frequencies, of which the lowest frequency is called the fundamental and the remainder are called overtones [4, p. 270]. These components are

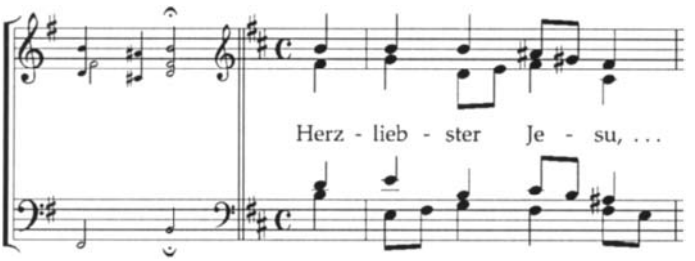


Fig. 1.1.1: An example of musical notation [5]. Roughly speaking, the horizontal dimension is time and the vertical dimension is frequency.

present during a specific time interval and may vary in amplitude during that interval. In modern musical notation, each note is represented by a "head". The vertical position of the head (together with other information such as the clef and key signature) indicates the *pitch*, i.e. the frequency of the most prominent component (usually the fundamental). The horizontal position of the head in relation to other symbols indicates the starting time, and the duration is specified by the shading of the head, attached bars, dots, stem and flags, and tempo markings such as *Allegro*. The choice of instrument—each instrument being characterized by its overtones and their relationships with the fundamental—is indicated by using a separate stave for each instrument or group of instruments, or a pair of staves for a keyboard instrument. Variations in amplitude are indicated by dynamic markings such as *mp* and *crescendo*. Fig. 1.1.1 illustrates the system. By scanning a set of staves vertically, one can see which fundamentals are present on which instruments at any given time. By scanning the staves horizontally, one can see the times at which a given fundamental is present on a given instrument.

Each of the three above signals has a time-varying frequency or time-varying "frequency content". Such signals are referred to as **non-stationary**.

The three examples described above are comprehensible partly because our sense of hearing readily interprets sounds in terms of variations of frequency or "frequency content" with time. However, conventional representations of a signal in the time domain or frequency domain do not facilitate such interpretation, as shown below.

1.1.2 Time-Domain Representation

Any signal can be described naturally as a function of time, which we may write $s(t)$. This representation leads immediately to the **instantaneous power**, given by $|s(t)|^2$, which shows how the energy of the signal is distributed over time; the total signal energy is

$$E = \int_{-\infty}^{\infty} |s(t)|^2 \, dt. \tag{1.1.3}$$

But the time-domain description has limitations, as may be seen by applying it to the above three examples:

1. The **sinusoidal FM signal** whose frequency satisfies Eq. (1.1.1) may be written

$$s_1(t) = A\cos\left(2\pi f_c t + \tfrac{f_d}{f_m}\sin[2\pi f_m t + \phi] + \psi\right) \tag{1.1.4}$$

where A is the amplitude and ψ a phase offset; the fraction f_d/f_m is called the **modulation index** and is equal to the peak phase deviation (in radians) from $2\pi f_c t$. This equation, by itself, does not clearly show how the frequency varies with time. If we imagine a graph of $s_1(t)$ vs. t, it would give the impression of an oscillating frequency, but determining the frequency as a function of time from the graph would be a tedious and imprecise exercise.

2. The **linear FM signal** whose frequency satisfies Eq. (1.1.2) may be written

$$s_2(t) = A\,\mathrm{rect}\left[\tfrac{t-T/2}{T}\right]\cos\left(2\pi[f_0 t + \tfrac{\alpha}{2}t^2] + \psi\right) \tag{1.1.5}$$

where, again, A is the amplitude and ψ a phase offset. The rect function is a rectangular pulse of unit height and unit duration, centered on the origin of time; that is,

$$\mathrm{rect}\,\tau = \begin{cases} 1 & \text{if } |\tau| \le 1/2\,; \\ 0 & \text{otherwise.} \end{cases} \tag{1.1.6}$$

Hence the rect[...] factor in Eq. (1.1.5) is equal to unity for $0 \le t \le T$, and zero elsewhere. But again it is not immediately apparent why Eq. (1.1.5) has the required FM law.

The graph of $s_2(t)$ vs. t is shown on the left side of Fig. 1.1.2(a), for $\psi = 0$, $T = 64\,\mathrm{s}$, $f_0 = 0.1\,\mathrm{Hz}$ and $\alpha = (3/640)\,\mathrm{Hz\,s^{-1}}$. Although the graph gives a clear impression of a steadily increasing frequency, the exact FM law is not clearly and readily visible.

3. A **musical performance** can be represented as (for example) an air pressure curve at a particular point in space. Each such curve is a time-varying pressure, and may be converted by a microphone and amplifier into an electrical signal of the form $s_3(t)$. Indeed, music is routinely recorded and broadcast in this way. However, the function $s_3(t)$ is nothing like the form in which a composer would write music, or the form in which most musicians would prefer to read music for the purpose of performance. Neither is it of much use to a recording engineer who wants to remove noise and distortion from an old "vintage" recording. Musical waveforms are so complex that a graph of $s_3(t)$ vs. t would be almost useless to musicians and engineers alike.

These three examples show that the time-domain representation tends to obscure information about frequency, because it assumes that the two variables t and f are mutually exclusive.

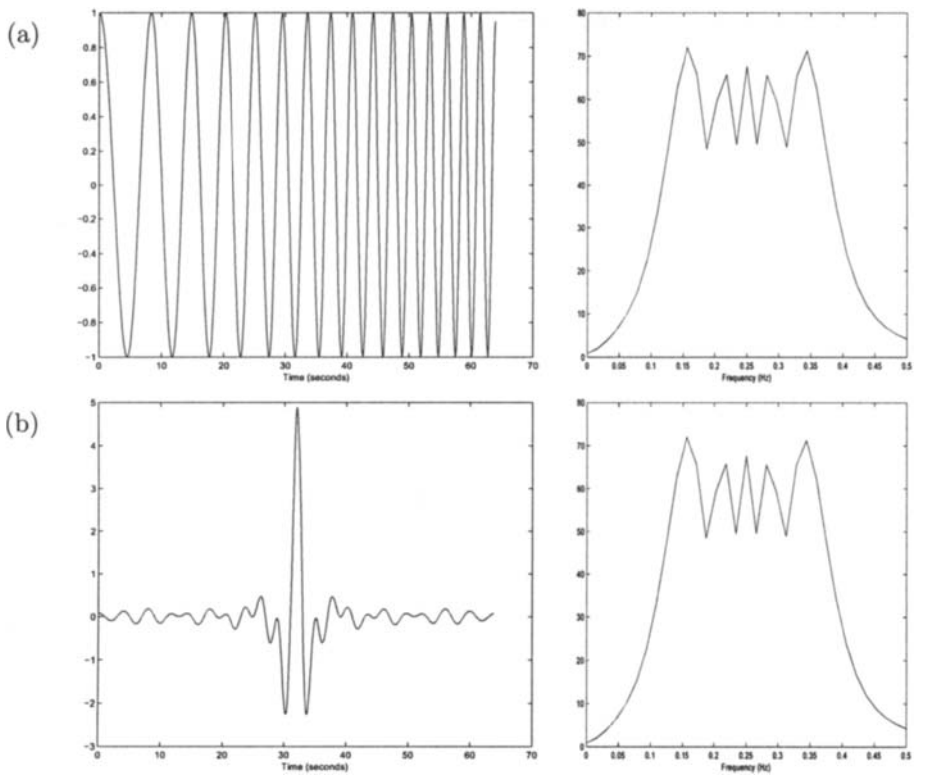


Fig. 1.1.2: The importance of phase: (a) time-domain representation (left) and magnitude spectrum (right) of a linear FM signal [Eq. (1.1.5)] with duration 64 seconds, starting frequency 0.1 Hz and finishing frequency 0.4 Hz; (b) time-domain representation and magnitude spectrum of another signal with the same magnitude spectrum as that in part (a). All plots use a sampling rate of 8 Hz.

1.1.3 Frequency-Domain Representation

Any practical signal $s(t)$ can be represented in the frequency domain by its Fourier transform $S(f)$, given by

$$S(f) = \underset{t \to f}{\mathcal{F}}\{s(t)\} \triangleq \int_{-\infty}^{\infty} s(t)\, e^{-j2\pi ft} dt. \tag{1.1.7}$$

For convenience, the relation between $s(t)$ and $S(f)$ may be written "$s(t) \underset{t \to f}{\longleftrightarrow} S(f)$" or simply "$s(t) \leftrightarrow S(f)$". The Fourier transform (FT) is in general complex; its magnitude is called the **magnitude spectrum** and its phase is called the **phase spectrum**. The square of the magnitude spectrum is the **energy spectrum** and shows how the energy of the signal is distributed over the frequency domain; the

total energy of the signal is

$$E = \int_{-\infty}^{\infty} |S(f)|^2 \, df = \int_{-\infty}^{\infty} S(f)\, S^*(f)\, df \tag{1.1.8}$$

where the superscripted asterisk (*) denotes the complex conjugate. Although the representation $S(f)$ is a function of frequency only—time having been "integrated out"—the FT is a complete representation of the signal because the signal can be recovered from the FT by taking the inverse Fourier transform (IFT):

$$s(t) = \underset{t \leftarrow f}{\mathcal{F}^{-1}} \{S(f)\} = \int_{-\infty}^{\infty} S(f)\, e^{j2\pi f t} df. \tag{1.1.9}$$

But the "completeness" of the FT representation does not make it convenient for all purposes, as may be seen by considering the same three examples:

1. **Sinusoidal FM signal:** If $\phi = \psi = 0$ in Eq. (1.1.4), the expression for $s_1(t)$ can be expanded into an infinite series as

$$s_1(t) = A \sum_{n=-\infty}^{\infty} J_n(\beta) \cos 2\pi (f_c + n f_m) t \tag{1.1.10}$$

where β is the modulation index ($\beta = f_d / f_m$) and J_n denotes the Bessel function of the first kind, of order n [6, p. 226]. In the frequency domain, this becomes an infinite series of delta functions; one of these (the **carrier**) is at the mean frequency f_c, and the remainder (the **sideband** tones) are separated from the carrier by multiples of the modulating frequency f_m. Although the number of sideband tones is theoretically infinite, the *significant* ones[1] may be assumed to lie between the frequencies $f_c \pm (f_d + f_m)$ or, more conservatively, $f_c \pm (f_d + 2f_m)$. This information is essential if one is designing a tuning filter to isolate the TV audio carrier or separate one FM channel from adjacent channels. But it is inadequate if one is designing a modulator or demodulator, because its connection with the FM law is even more obscure than that of Eq. (1.1.4).

2. The magnitude spectrum of the **linear FM signal**: Eq. (1.1.5) is shown on the right side of Fig. 1.1.2(a), for $\psi = 0$, $T = 64\,\text{s}$, $f_0 = 0.1\,\text{Hz}$ and $\alpha = (3/640)\,\text{Hz}/s$. The graph shows that magnitude is significant in the band corresponding to the frequency sweep range ($0.1\,\text{Hz} < f < 0.4\,\text{Hz}$), and furthermore that the energy is mostly confined to that band. However, the graph fails to show that the frequency is increasing with time. In other words, the

[1] For more detail, see Carlson [6], pp. 220–37. The above description considers only positive frequencies; similar comments apply to the negative frequencies. Theoretically, the lower sideband tones belonging to the positive-frequency carrier extend into the negative frequencies, while the corresponding sideband tones belonging to the negative-frequency carrier extend into the positive frequencies; but such "aliased" components are negligible if f_c and f_d are appropriately chosen.

magnitude spectrum tells us what frequencies are present in the signal, but not the "times of arrival" of those frequencies (the latter information, which we call **time delay**, is encoded in the phase spectrum).

The shortcomings of the magnitude spectrum may be seen even more clearly in Fig. 1.1.2(b), which shows a signal whose magnitude spectrum (right) is identical to that of the linear FM signal in Fig. 1.1.2(a), but whose appearance in the time domain (left) is very different from a linear FM.[2]

3. Similarly, the **musical performance** has a magnitude spectrum which tells us what frequencies are present, but not *when* they are present; the latter information is again encoded in the phase. The magnitude spectrum may exhibit as many as 120 peaks corresponding to the notes of the chromatic scale in the audible frequency range, and the relative heights of those peaks may tell us something about the tonality of the music (or whether it is tonal at all), but the timing of the notes will not be represented in the magnitude spectrum and will not be obvious from the phase spectrum.

These three examples show that the frequency-domain representation "hides" the information about timing, as $S(f)$ does not mention the variable t.

1.1.4 Joint Time-Frequency Representation

As the conventional representations in the time domain or frequency domain are inadequate in the situations described above, an obvious solution is to seek a representation of the signal as a *two-variable* function or distribution whose domain is the two-dimensional (t, f) space. Its constant-t cross-section should show the frequency or frequencies present at time t, and its constant-f cross-section should show the time or times at which frequency f is present. Such a representation is called a **time-frequency representation** (**TFR**) or **time-frequency distribution** (**TFD**).

As an illustration of what is desired, Fig. 1.1.3 shows one particular TFD of the linear FM signal of Eq. (1.1.5), for $\psi = 0$, $T = 64\,\mathrm{s}$, $f_0 = 0.1$ and $\alpha = (3/640)\,\mathrm{Hz\,s^{-1}}$. The TFD not only shows the start and stop times and the frequency range, but also clearly shows the variation in frequency with time. This variation may be described by a function $f_i(t)$, called the **instantaneous frequency** (**IF**). A signal may have more than one IF; for example, Fig. 1.1.4 shows a TFD of a sum of two linear FM signals, each of which has its own IF.[3] These IF features are not apparent in conventional signal representations.

Non-stationary signals for which a TFD representation may be useful occur not only in broadcasting, seismic exploration and audio, from which our three exam-

[2]The signal in part (b) of Fig. 1.1.2 was obtained from that in part (a) by taking the FFT, setting the phase to zero, taking the inverse FFT and shifting the result in time. It is *not* the product of a sinc function and a cosine function.

[3]N.B.: In Fig. 1.1.4 and in all subsequent graphs of TFDs, the labels on axes are similar to those in Fig. 1.1.3, but may be more difficult to read because of space constraints.

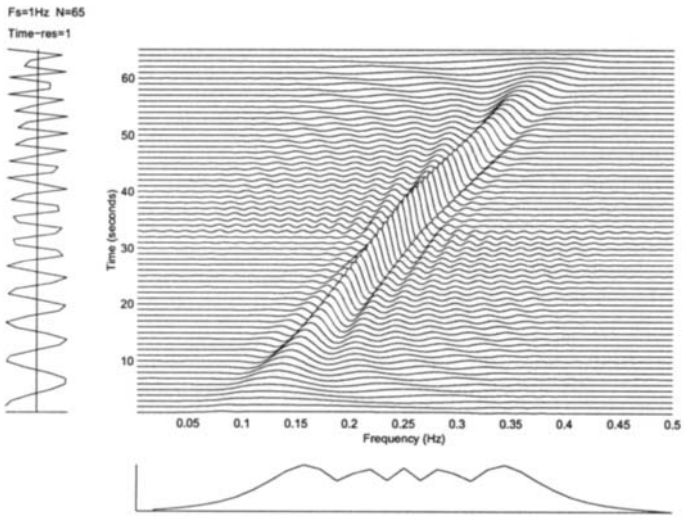


Fig. 1.1.3: A time-frequency representation of a linear FM signal [Eq. (1.1.5)] with duration 65 samples, starting frequency 0.1 and finishing frequency 0.4 (sampling rate 1 Hz). The time-domain representation appears on the left, and the magnitude spectrum at the bottom; this pattern is followed in all TFD graphs in Part I of this book.

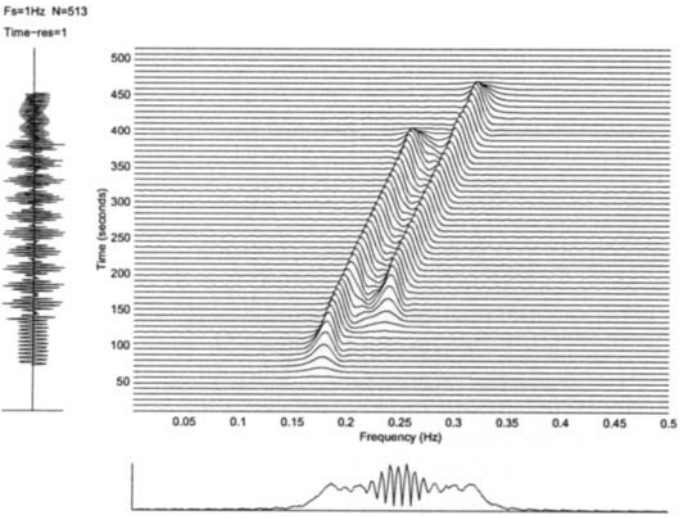


Fig. 1.1.4: A time-frequency representation of two linear FM signals with close parallel FM laws, duration 512 samples, frequency separation 0.04 (sampling rate 1 Hz).

ples are taken, but also in numerous other engineering and interdisciplinary fields such as telecommunications, radar, sonar, vibration analysis, speech processing and

medical diagnosis. **Time-frequency signal processing (TFSP)** is the processing of signals by means of TFDs.

1.1.5 Desirable Characteristics of a TFD

The use of a TFD for a particular purpose is inevitably based on particular assumptions concerning the properties of TFDs. To ensure that the conclusions of the analysis are sound, these assumptions must be identified and verified. Generalizing somewhat, we may say that the desirable properties of a TFD are those on which the most likely applications of TFDs depend. The following applications illustrate some likely uses of TFDs:

- Analyze the raw signal in the (t, f) domain so as to identify its characteristics such as time variation, frequency variation, number of components, relatives amplitudes, etc.
- Separate the components from each other and from the background noise by filtering in the (t, f) domain.
- Synthesize the filtered TFD in the time domain.
- Analyze specific components separately:
  - Track the instantaneous amplitude;
  - Track the instantaneous frequency;
  - Track the instantaneous bandwidth (spread of energy about the IF).
- Choose a mathematical model of the signal, showing clearly the significant characteristics, such as the IF.

These applications can be carried out using a TFD with the following properties:

1(a) The TFD is real (because energy is real).

1(b) The integral of the TFD over the entire (t, f) plane is the total energy of the signal (so that energy is conserved).

1(c) The integral over a rectangular region of the (t, f) plane, corresponding to a finite bandwidth and finite time interval, is approximately the energy of the signal in that bandwidth over that interval, provided that the bandwidth and interval are sufficiently large (see Section 1.2.5 for details).

2 For a monocomponent FM signal, the peaks of the constant-time cross-sections of the TFD should give the IF law which describes the signal FM law.

3 For a multicomponent FM signal, the dominant peaks of the TFD should reflect the components' respective FM laws; and the TFD should resolve any close components, as in Fig. 1.1.4.

A simple measure of the resolution of a TFD is the concentration of energy about the IF law(s) of a signal [Article 7.4 and Chapter 10].

The linear FM signal, because of its simple FM law, is a convenient test signal for verifying the first four properties. Property 3 may be tested by a sum of linear FM signals.

The above properties have further implications. For example, the resolution property (3) helps to ensure robustness in the presence of noise, while the last two properties (2 and 3) together allow discernment of the multiple IFs of multicomponent signals.

1.2 Signal Formulations and Characteristics in the (t, f) Domain

1.2.1 Signal Models used in (t, f) Methods

To represent signals such as the linear FM, several types of signal models are commonly used in TFSAP. The choice of model depends on the number and nature of the parameters needed to describe the signal. For example, a single sinusoid with constant frequency and normalized amplitude and phase is described by the equation

$$s_4(t) = \cos 2\pi f_c t \ , \tag{1.2.1}$$

in which the only parameter is the frequency f_c. If the amplitude and phase are significant in the application, two more parameters are needed. A linear combination of such signals may be written in the form

$$s_5(t) = \sum_{k=1}^{M} a_k \cos(2\pi f_k t + \psi_k) \ , \tag{1.2.2}$$

which involves $3M$ parameters.

Because $s_4(t)$ and $s_5(t)$ contain terms (or "components") of constant amplitude, frequency and phase, they are completely and clearly described by Fourier transforms; no time dependence is required. But a sinusoidal FM signal or chirp signal requires a TFD. A musical audio signal also calls for a form of TFD (as is suggested by the notation in which music is written), and the TFD should clearly resolve the multiple components.

Further difficulties are raised by more complex signals such as

$$s_6(t) = \left(\sum_{k=1}^{M} a_k(t)\, e^{j2\pi \int_0^t f_k(\tau)\, d\tau} \right) + w(t) \tag{1.2.3}$$

where $a_k(t)$ is the *time-varying* amplitude of the k^{th} component, $f_k(t)$ is the *time-varying* frequency of the k^{th} component, and $w(t)$ is additive noise. The analysis of such a signal requires the ability not only to distinguish the *time-varying* components from each other in spite of their varying amplitudes and frequencies, but also

to separate them from the noise. Such comments still apply if the amplitude of the k^{th} component is a multiplicative noise factor $m_k(t)$, as in the signal

$$s_7(t) = \left(\sum_{k=1}^{M} m_k(t)\, e^{j2\pi \int_0^t f_k(\tau)\, d\tau} \right) + w(t). \tag{1.2.4}$$

In this case, an even more specialized analysis will be required.

1.2.2 Analytic Signals

It is well known that a signal $s(t)$ is real if and only if (iff)

$$S(-f) = S^*(f)\ , \tag{1.2.5}$$

where $S(f)$ is the Fourier transform of $s(t)$. In other words, a real signal is one that exhibits **Hermitian symmetry** between the positive-frequency and negative-frequency components, allowing the latter to be deduced from the former. Hence the negative-frequency components of a real signal *may be eliminated from the signal representation without losing information.* In the case of a real lowpass signal, removal of negative frequencies has two beneficial effects. First, it halves the total bandwidth, allowing the signal to be sampled at half the usual Nyquist rate without aliasing [7,8]. Second, it avoids the appearance of some interference terms generated by the interaction of positive and negative components in quadratic TFDs (to be treated in detail in Section 3.1.2).

Definition 1.2.1: *A signal $z(t)$ is said to be* **analytic** *iff*

$$Z(f) = 0 \quad \textit{for} \quad f < 0\ , \tag{1.2.6}$$

where $Z(f)$ is the Fourier transform of $z(t)$.

In other words, an analytic signal contains *no negative frequencies*; it may have a spectral component at zero frequency (DC).

Theorem 1.2.1: *The signal*

$$z(t) = s(t) + jy(t)\ , \tag{1.2.7}$$

where $s(t)$ and $y(t)$ are real, is analytic with a real DC component, if and only if

$$Y(f) = (-j \operatorname{sgn} f)\, S(f) \tag{1.2.8}$$

where $S(f)$ and $Y(f)$ are the FTs of $s(t)$ and $y(t)$, respectively, and where

$$\operatorname{sgn} \xi \triangleq \begin{cases} -1 & \textit{if } \xi < 0\ ; \\ 0 & \textit{if } \xi = 0\ ; \\ +1 & \textit{if } \xi > 0\ . \end{cases} \tag{1.2.9}$$

Proof: Take the FT of Eq. (1.2.7) and use Eq. (1.2.6). ∎

1.2.3 Hilbert Transform; Analytic Associate

If the Fourier transforms of $s(t)$ and $y(t)$ are related according to Eq. (1.2.8), we say that $y(t)$ is the **Hilbert transform** of $s(t)$, and we write

$$y(t) = \mathcal{H}\{s(t)\}. \tag{1.2.10}$$

Hence we may restate Theorem 1.2.1 as follows: *A signal is analytic with a real DC component if and only if its imaginary part is the Hilbert transform of its real part.*

By invoking the "if" form of Theorem 1.2.1 and restating the sufficient condition in terms of Eq. (1.2.7), we may now see the practical significance of the theorem and the Hilbert transform: Given a real signal $s(t)$, we can construct the complex signal

$$z(t) = s(t) + j\mathcal{H}\{s(t)\} \tag{1.2.11}$$

and know that $z(t)$ is analytic. This $z(t)$ is called the analytic signal "corresponding to" or "associated with" the real signal $s(t)$. In this book, for convenience, we shall usually call $z(t)$ the **analytic associate** of $s(t)$.

By taking the IFT of Eq. (1.2.8) and applying Eq. (1.2.10), we arrive at the following concise definition of the Hilbert transform:

Definition 1.2.2: *The* **Hilbert transform** *of a signal $s(t)$, denoted by $\mathcal{H}\{s(t)\}$, is*

$$\mathcal{H}\{s(t)\} = \underset{t\leftarrow f}{\mathcal{F}^{-1}}\left\{(-j\,\mathrm{sgn}\,f)\,\underset{t\rightarrow f}{\mathcal{F}}\{s(t)\}\right\}. \tag{1.2.12}$$

where $\mathcal{F}\{\ldots\}$ denotes the Fourier transform.

In other words, the Hilbert transform of $s(t)$ is evaluated as follows:

1. Take the Fourier transform $S(f)$ of $s(t)$;
2. Multiply $S(f)$ by $-j$ for positive f, by $+j$ for negative f, and by zero for $f = 0$;
3. Take the inverse Fourier transform.

According to step 2 of the above procedure, a Hilbert transformer introduces a phase lag of 90 degrees (as $-j = e^{-j\pi/2}$), producing a signal in **quadrature** to the input signal. The effect is well illustrated by the following result, which is easily verified using Definition 1.2.2 and a table of transforms:

Example 1.2.1: *If f_0 is a positive constant, then*

$$\mathcal{H}\{\cos(2\pi f_0 t)\} = \sin(2\pi f_0 t) \tag{1.2.13}$$

$$\mathcal{H}\{\sin(2\pi f_0 t)\} = -\cos(2\pi f_0 t). \tag{1.2.14}$$

It would be convenient if the pattern of Example 1.2.1 were applicable to modulated signals; so that for example, we could say

$$\mathcal{H}\{a(t)\cos\phi(t)\} = a(t)\sin\phi(t)\ , \tag{1.2.15}$$

which would imply that the analytic associate of the real signal $s(t) = a(t)\cos\phi(t)$ is

$$\begin{aligned} z(t) &= a(t)\cos\phi(t) + j\mathcal{H}\{a(t)\cos\phi(t)\} \\ &= a(t)\cos\phi(t) + ja(t)\sin\phi(t) \\ &= a(t)\,e^{j\phi(t)}. \end{aligned} \tag{1.2.16}$$

The condition under which Eq. (1.2.15) holds is that the variation of $a(t)$ is sufficiently slow to ensure "spectral disjointness", i.e. to avoid overlap between the spectrum of $a(t)$ and the spectrum of $\cos\phi(t)$.

Eq. (1.2.8) indicates that the transfer function of a Hilbert transformer is $-j\,\mathrm{sgn}\,f$. The corresponding impulse response is

$$\mathop{\mathcal{F}^{-1}}_{t\leftarrow f}\{-j\,\mathrm{sgn}\,f\} = \frac{1}{\pi t}. \tag{1.2.17}$$

Using this result and applying the convolution property to Eq. (1.2.12), we obtain a definition of the Hilbert transform in the time domain:

$$\mathcal{H}\{s(t)\} = s(t) * \frac{1}{\pi t} \tag{1.2.18}$$

$$= \frac{1}{\pi}\,\mathrm{p.v.}\left\{\int_{-\infty}^{\infty}\frac{s(\tau)}{t-\tau}\,d\tau\right\}\ , \tag{1.2.19}$$

where p.v.$\{\ldots\}$ denotes the Cauchy **principal value** of the improper integral [9], given in this case by[4]

$$\lim_{\delta\to 0}\left[\int_{-\infty}^{t-\delta}\frac{s(\tau)}{t-\tau}\,d\tau + \int_{t+\delta}^{\infty}\frac{s(\tau)}{t-\tau}\,d\tau\right]. \tag{1.2.20}$$

1.2.4 Duration, Bandwidth, BT Product

1.2.4.1 Finite Duration and Finite Bandwidth

In real life, signals are observed for finite periods of time and processed by devices with finite usable bandwidths. If a signal $s(t)$ has the Fourier transform $S(f)$, the **duration** of the signal is the smallest range of times outside which $s(t) = 0$, while the **bandwidth** of the signal is the smallest range of frequencies outside which $S(f) = 0$. These definitions, as we shall see, lead to the conclusion that a finite

[4] In practice Eq. (1.2.18) is rarely used, because the FT properties make it easier to work in the frequency domain.

duration implies infinite bandwidth and vice versa. In practice, however, any useful signal must have a beginning and an end (finite duration) and its FT must be within the frequency range of the measuring/processing equipment (finite bandwidth).[5] In practice, therefore, the strict definitions need to be relaxed in some way.

A **time-limited signal**, of **duration** T centered at time $t = 0$, can be expressed as

$$s_T(t) = s(t)\,\mathrm{rect}[t/T]\ , \tag{1.2.21}$$

where the subscript T indicates the duration [see Eq. (1.1.6)].

The FT of $s_T(t)$ is

$$S_T(f) = S(f) \underset{f}{*} T\,\mathrm{sinc}\,fT \tag{1.2.22}$$

where $\underset{f}{*}$ denotes convolution in frequency. Thus the bandwidth of $S_T(f)$ is infinite.

If, in order to avoid the effects of discontinuities, we replace $\mathrm{rect}[t/T]$ with a smoother window $w(t)$ of the same duration T, we can write

$$s_T(t) = s(t)\,\mathrm{rect}[t/T]\,w(t)\ , \tag{1.2.23}$$

whose Fourier transform still involves a convolution with $\mathrm{sinc}\,fT$, giving an infinite bandwidth.

In analogy to the time-limited case, a **band-limited signal**, of bandwidth B centered at the origin, can be expressed in the frequency domain as

$$S_B(f) = S(f)\,\mathrm{rect}[f/B] \tag{1.2.24}$$

In the time domain, the signal is given by

$$s_B(t) = s(t) \underset{t}{*} B\,\mathrm{sinc}\,Bt \tag{1.2.25}$$

which has an infinite duration. Thus, under the "usual" definitions of duration and bandwidth, a finite bandwidth implies infinite duration.

1.2.4.2 Effective Bandwidth and Effective Duration

If there is no finite bandwidth containing *all* the energy of the signal, there may still be a finite bandwidth containing *most* of the energy. Hence, for example, a bandwidth containing 99% of the signal energy might be accepted as a useful measure of the signal bandwidth. If the nominated fraction of the signal energy were confined between the frequencies $f_{\min}$ and $f_{\max}$, the bandwidth would be $B = f_{\max} - f_{\min}$.

A less arbitrary but more relaxed measure of bandwidth is the so-called **effective bandwidth** B_e, defined by

$${B_e}^2 = \frac{1}{E_s}\int_{-\infty}^{\infty} f^2 |S(f)|^2\,df \tag{1.2.26}$$

[5] In practice, all acquisition and measuring systems are low-pass.

where $S(f)$ is the FT of the signal, and E_s is the total energy of the signal, given by

$$E_s = \int_{-\infty}^{\infty} |S(f)|^2 \, df. \tag{1.2.27}$$

B_e^2 is the **second moment** of $|S(f)|^2$ w.r.t. frequency, about the origin ($f = 0$). For brevity, we call B_e^2 the "second moment of the signal w.r.t. frequency". As an aid to remembering the definition, note that if f were a random variable and $|S(f)|^2$ were its probability density function (p.d.f.), then we would have $E_s = 1$, so that B_e^2 would be the variance of f if the mean of f were zero. Thus the effective bandwidth B_e is analogous to the standard deviation of f. When we consider how little of a typical probability distribution falls within one standard deviation, we realize that the effective bandwidth is only a mathematical construction and should *not* be used as an estimate of the bandwidth required for accurate measurement and processing of the signal.

The definition of duration, like that of bandwidth, can be relaxed so as to obtain both finite bandwidth and finite duration. For example, one could define the duration T where a nominated fraction of the signal energy occurs. The so-called **effective duration** T_e, defined by

$$T_e{}^2 = \frac{1}{E_s} \int_{-\infty}^{\infty} t^2 |s(t)|^2 \, dt \, , \tag{1.2.28}$$

is the second moment of $|s(t)|^2$ w.r.t. time, about the origin ($t = 0$). For brevity, we refer to T_e^2 as "the second moment of the signal w.r.t. time." The definitions of effective duration and effective bandwidth were proposed in 1946 by Gabor [10].[6]

Slepian [11] has proposed definitions of duration and bandwidth based on the accuracy of the detecting and measuring apparatus. Let us define a "time truncation" of $s(t)$ as a signal $\hat{s}(t)$ that satisfies

$$\hat{s}(t) = \begin{cases} 0 & \text{if } t < t_1 \\ s(t) & \text{if } t_1 \le t \le t_2 \\ 0 & \text{if } t > t_2 \end{cases} \tag{1.2.29}$$

where $t_2 > t_1$, so that the duration of $\hat{s}(t)$ is $t_2 - t_1$. Let a "frequency truncation" be defined similarly in the frequency domain. Then the **Slepian duration** of $s(t)$ is the duration of the shortest time-truncation $\hat{s}(t)$ that the apparatus cannot distinguish from $s(t)$, while the **Slepian bandwidth** of $s(t)$ is the bandwidth of the most narrow-band frequency-truncation $\hat{S}(f)$ that the apparatus cannot distinguish from $S(f)$.

Given suitable definitions of B and T, the **bandwidth-duration product** BT is self-explanatory. The BT product also has a practical significance: because a

[6] Dennis Gabor (1900–1979), a Hungarian-born electrical engineer who settled in Britain, is best known for the invention of holography (1947–8), for which he was awarded the Nobel Prize for Physics in 1971.

signal of total bandwidth B can be reconstructed from samples at the sampling rate B, the total number of samples required to represent the signal—i.e. the number of **degrees of freedom** in the signal—is equal to BT. In other words, BT is a measure of the *information richness* of the signal. This implies that there may be little point in attempting to extract large amounts of information from a signal with a small BT product.

Example 1.2.2: *For the Gaussian signal*

$$s(t) = e^{-\alpha^2 t^2}, \tag{1.2.30}$$

the effective duration is

$$T_e = \frac{1}{2\alpha}. \tag{1.2.31}$$

The FT of $s(t)$ *is*

$$S(f) = \frac{\sqrt{\pi}}{\alpha} e^{-\pi^2 f^2/\alpha^2}, \tag{1.2.32}$$

so that the effective bandwidth is

$$B_e = \frac{\alpha}{2\pi}. \tag{1.2.33}$$

From Eqs. (1.2.31) and (1.2.33),

$$B_e T_e = \frac{1}{4\pi}. \tag{1.2.34}$$

It can be shown that the Gaussian signal is the only signal for which this equality holds, and that for all other signals $B_e T_e > \frac{1}{4\pi}$ *[10].*

1.2.5 Asymptotic Signals

For signals with significant information content, it is desired to know not only the overall bandwidth B, but also the distribution of energy through the bandwidth, e.g. the frequencies present, their relative amplitudes, and the times during which they are significant. Similarly we may want to know not only the overall duration T, but also the distribution of energy throughout the duration, e.g. the times during which the signal is present, the relative amplitudes at those times, and the significant frequencies present during those times. Such signals may be modeled by the class of asymptotic signals.

Definition 1.2.3: *A signal* $s(t)$ *is* **asymptotic** *iff it has the following properties:*

(a) The duration T*, as defined by Slepian, is finite;*

(b) The bandwidth B*, as defined by Slepian, is finite;*

(c) The product BT *is large (e.g. greater than 10);*

(d) The amplitude is bounded so that

$$\int_{-\infty}^{\infty} |s(t)|^2 dt \tag{1.2.35}$$

is finite.

Asymptotic signals allow useful approximations for deriving simple signal models (e.g. to express analytic signals).

1.2.6 Monocomponent vs. Multicomponent Signals

A **monocomponent** signal is described in the (t, f) domain by one single "ridge", corresponding to an elongated region of energy concentration. Furthermore, interpreting the crest of the "ridge" as a graph of IF vs. time, we require the IF of a monocomponent signal to be a single-valued function of time. Fig. 1.1.3 shows an example of a monocomponent signal.

Such a monocomponent signal has an analytic associate of the form

$$z(t) = a(t)\, e^{j\phi(t)} \ , \tag{1.2.36}$$

where $\phi(t)$ is differentiable, and $a(t)$ is real and positive (being the instantaneous amplitude). If $s(t)$ itself is real and asymptotic, it can be expressed as

$$s(t) = a(t) \cos \phi(t). \tag{1.2.37}$$

A **multicomponent** signal may be described as the sum of two or more monocomponent signals. Fig. 1.1.4 shows an example of a multicomponent signal (composed of two components).

1.3 Instantaneous Frequency and Time-Delay

1.3.1 Instantaneous Frequency (IF)

Definition 1.3.1: *The* **instantaneous frequency** *of a monocomponent signal is*

$$f_i(t) = \frac{1}{2\pi} \phi'(t) \tag{1.3.1}$$

where $\phi(t)$ is the instantaneous phase of the signal.

This formulation is justified below by considering first a constant-frequency signal, then a variable-frequency signal.

Consider the amplitude-modulated signal

$$x(t) = a(t) \cos(2\pi f_c t + \psi) \tag{1.3.2}$$

where f_c and ψ are constant. As t increases by the increment $1/f_c$, the argument of the cosine function increases by 2π and the signal passes through one cycle. So the

period of the signal is $1/f_c$, and the frequency, being the reciprocal of the period, is f_c. The same signal can be written as

$$x(t) = a(t)\cos\phi(t) \tag{1.3.3}$$

where

$$\phi(t) = 2\pi f_c t + \psi\ , \tag{1.3.4}$$

from which we obtain

$$f_c = \frac{1}{2\pi}\phi'(t). \tag{1.3.5}$$

Although the left-hand side of this equation (the frequency) has been assumed constant, the right-hand side would be variable if $\phi(t)$ were a nonlinear function. So let us check whether this result can be extended to a *variable frequency.*

Consider a signal whose analytic associate is

$$z(t) = a(t)\, e^{j\phi(t)} \tag{1.3.6}$$

where $a(t)$ and $\phi(t)$ are real and $a(t)$ is positive; then $a(t)$ is called the **instantaneous amplitude** and $\phi(t)$ is called the **instantaneous phase.** Let $\phi(t)$ be evaluated at $t = t_1$ and $t = t_2$, where $t_2 > t_1$. By the mean value theorem of elementary calculus, if $\phi(t)$ is differentiable, there exists a time instant t between t_1 and t_2 such that

$$\phi(t_2) - \phi(t_1) = (t_2 - t_1)\,\phi'(t). \tag{1.3.7}$$

Let p_i be the period of one particular oscillation of $z(t)$, and let $f_i = 1/p_i$. If $t_2 = t_1 + p_i$, then $\phi(t_2) = \phi(t_1) + 2\pi$, so that Eq. (1.3.7) becomes

$$2\pi = p_i\phi'(t)\ ; \tag{1.3.8}$$

that is,

$$f_i = \frac{\phi'(t)}{2\pi}. \tag{1.3.9}$$

Now t is an instant during a cycle of oscillation and f_i is the frequency of that oscillation, suggesting that the right-hand side be defined as the **instantaneous frequency** (**IF**) at time t, as in Definition 1.3.1 above.

Comparing Eqs. (1.2.37) and (1.3.6), we see that

$$s(t) = \mathrm{Re}\{z(t)\}\,. \tag{1.3.10}$$

where $\mathrm{Re}\{\ldots\}$ denotes the real part. Now let us define

$$y(t) \triangleq \mathrm{Im}\{z(t)\} = a(t)\sin\phi(t) \tag{1.3.11}$$

where $\mathrm{Im}\{\ldots\}$ denotes the imaginary part.

Using Eq. (1.3.1), we can easily confirm that the signals described by Eqs. (1.1.4) and (1.1.5) have the IFs given in Eqs. (1.1.1) and (1.1.2) respectively.

Definition 1.3.1 is strictly meaningful only for a monocomponent signal, while a multicomponent signal ought to have a separate IF for each component. In particular, note that the IF of the sum of two signals is *not* the sum of their two IFs.

The IF is a detailed description of the frequency characteristics of a signal. This contrasts with the notion of "average frequency" defined next.

Definition 1.3.2: *The* **average frequency** *of a signal is*

$$f_0 = \frac{\int_0^\infty f|S(f)|^2\,df}{\int_0^\infty |S(f)|^2\,df}. \tag{1.3.12}$$

where $S(f)$ is the Fourier transform of the signal.

In other words, f_0 is the **first moment** of $|S(f)|^2$ w.r.t. frequency; that is, we define the "average" frequency as if $|S(f)|^2$ were the probability density function of the frequency.

Notice that if $|S(f)|^2$ were replaced by a TFD, f_0 would become a function of time, suggesting that perhaps the first moment of a TFD w.r.t. frequency is a measure of IF. The conditions under which this is true will be explained in due course (and summarized in the table on p. 75). For the moment, we simply note that any reasonable time-frequency representation of a signal should contain information about the IF laws of the components. In particular, it would be most convenient if the *crests* of the ridges in the (t, f) domain represented the IF laws.

1.3.2 IF and Time Delay (TD)

The IF of a signal indicates the dominant frequency of the signal at a given time. Let us now seek a **dual** or "inverse" of the IF, indicating the dominant time when a given frequency occurs.

If $z(t)$ is an analytic signal with the Fourier transform

$$Z(f) = A\,\delta(f - f_i) \tag{1.3.13}$$

where A is in general complex, the dominant frequency is f_i.

Taking the inverse FT of $Z(f)$ gives

$$z(t) = Ae^{j2\pi f_i t}. \tag{1.3.14}$$

The instantaneous phase of $z(t)$, denoted by $\phi(t)$, is

$$\phi(t) = \arg z(t) = 2\pi f_i t + \arg A \tag{1.3.15}$$

so that

$$f_i = \frac{1}{2\pi}\,\phi'(t). \tag{1.3.16}$$

Although this result has been obtained for constant frequency, the right-hand of Eq. (1.3.16) is also valid for a variable-frequency signal, as explained earlier.

Now let us repeat the argument with the time and frequency variables interchanged. The signal

$$z(t) = a\,\delta(t - \tau_{\mathrm{d}}) \ , \tag{1.3.17}$$

where a is in general complex, is an impulse at time τ_{d}. If we ask what is the "delay" of this signal, the only sensible answer is τ_{d}. The FT of this signal is

$$Z(t) = a\,e^{-j2\pi f \tau_{\mathrm{d}}} \tag{1.3.18}$$

and its phase, denoted by $\theta(f)$, is

$$\theta(f) = \arg Z(f) = -2\pi f \tau_{\mathrm{d}} + \arg a \tag{1.3.19}$$

so that

$$\tau_{\mathrm{d}} = -\frac{1}{2\pi}\,\theta'(f). \tag{1.3.20}$$

Again, although this result has been obtained for a constant τ_{d} , the right-hand side of Eq. (1.3.20) is well defined even if $\theta'(f)$ varies with f.

Definition 1.3.3: *If $z(t)$ is an analytic signal with the Fourier transform $Z(f)$, then the* **time delay** *(TD) of $z(t)$, denoted by $\tau_{\mathrm{d}}(f)$, is*

$$\tau_{\mathrm{d}}(f) = -\frac{1}{2\pi}\,\theta'(f) \ , \tag{1.3.21}$$

where $\theta(f) = \arg Z(f)$

Notice that the definitions of τ_{d} and f_i are similar, except that time and frequency are interchanged and Eq. (1.3.21) has an extra minus sign; hence we say that time delay is the **dual** of instantaneous frequency.[7]

Seeing that the instantaneous frequency $f_i(t)$ is a function assigning a frequency to a given time, whereas the time delay $\tau_{\mathrm{d}}(f)$ is a function assigning a time to a given frequency, we may well ask whether the two functions are inverses of each other. Clearly they are not *always* inverses, because the IF function may not be invertible. So let us restrict the question to **invertible** signals, i.e. monocomponent signals whose IFs are monotonic functions of time.

One example of an invertible signal is the generalized Gaussian signal, i.e. a linear FM signal with a Gaussian envelope. Let such a signal peak at time t_0, with peak amplitude A, center frequency f_{c} , sweep rate α and decay constant β, and suppose

[7]The term "time delay" is well established, but tautological. The term "instantaneous frequency" is a quantity with the dimensions of frequency modified by an adjective indicating localization in time. The dual of this term should therefore be a quantity with the dimensions of time modified by an adjective indicating localization in frequency, e.g. "frequency-localized delay" or, more succinctly, "frequency delay".

that the instantaneous frequency is positive while the envelope is significant. Then the analytic associate of the signal is

$$z(t) = A \exp\left(j2\pi\left[f_c[t-t_0] + \tfrac{\alpha+j\beta}{2}[t-t_0]^2\right]\right) \tag{1.3.22}$$

and using Eq. (1.3.16) with $a(t) = A\exp\left(-\pi\beta(t-t_0)^2\right)$, its IF is

$$f_i(t) = f_c + \alpha[t-t_0]. \tag{1.3.23}$$

To find the inverse function of $f_i(t)$, we simply solve for t, obtaining

$$t = t_0 + \frac{f_i(t) - f_c}{\alpha}\ , \tag{1.3.24}$$

which suggests that the time delay $\tau_d(f)$ of $z(t)$ can be estimated by

$$\hat{\tau_d}(f) = t_0 + \frac{f - f_c}{\alpha}. \tag{1.3.25}$$

In general τ_d and $\hat{\tau_d}$ are not equal, but converge as the BT product increases. As an example, Fig. 1.3.1 shows the IF (solid line) and TD (dotted line) of the generalized Gaussian signal $z(t)$ for two values of B_eT_e, where the subscript e means "effective". Note that the two curves are closer for the larger B_eT_e.

1.3.3 Mean IF and Group Delay (GD)

Let $z(t)$ be a bandpass analytic signal with center frequency f_c. Let its Fourier transform be

$$Z(f) = M(f-f_c)\, e^{j\theta(f)} \tag{1.3.26}$$

where the magnitude $M(f-f_c)$ and phase $\theta(f)$ are real. If the signal has **linear phase** in the support of $Z(f)$, i.e. if $\theta(f)$ is a linear function of f wherever $Z(f)$ is non-zero, we can let

$$\theta(f) = -2\pi\tau_p f_c - 2\pi\tau_g[f-f_c] \tag{1.3.27}$$

where τ_p and τ_g are real constants with the dimensions of time. Eq. (1.3.26) then becomes

$$\begin{aligned} Z(f) &= M(f-f_c)\, e^{-j(2\pi\tau_p f_c + 2\pi\tau_g[f-f_c])} && (1.3.28)\\ &= e^{-j2\pi f_c\tau_p}\, M(f-f_c)\, e^{-j2\pi\tau_g[f-f_c]}. && (1.3.29) \end{aligned}$$

Taking the IFT of $Z(f)$, we find

$$z(t) = m(t-\tau_g)\, e^{j2\pi f_c[t-\tau_p]} \tag{1.3.30}$$

where $m(t)$ is the IFT of $M(f)$. Now because $M(f)$ is real, $m(t)$ is Hermitian [i.e. $m(-t) = m^*(t)$], so that $|m(t)|$ is even. Hence τ_g is the time about which the envelope function is symmetrical; for this reason, τ_g is called the **group delay** [12, pp. 123–124]. The phase of the oscillatory factor is $-2\pi f_c\tau_p$, wherefore τ_p is called the **phase delay**. These observations together with Eq. (1.3.27) lead to the following definition:

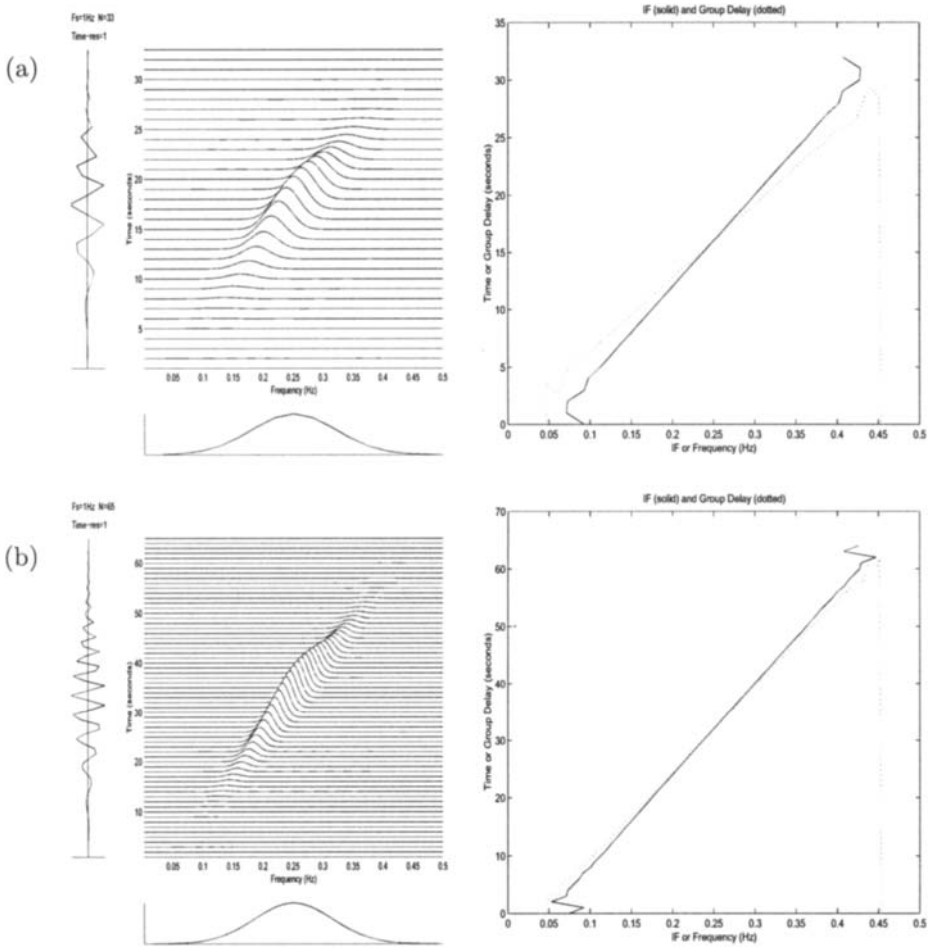


Fig. 1.3.1: Instantaneous frequency and time delay for a linear FM signal with a Gaussian envelope: (a) total duration $T = 33$ secs, $B_e T_e = 0.1806$; (b) $T = 65$ secs, $B_e T_e = 0.3338$. For each signal, the left-hand graph shows the time trace, spectrum and TFD [Wigner-Ville distribution, defined later in Eq. (2.1.17)], while the right-hand graph shows the IF (solid line) and time delay (dotted line). The vertical dotted segments are caused by truncation of the frequency range (to avoid finite-precision effects).

Definition 1.3.4: *If an analytic signal has the Fourier transform*

$$Z(f) = |Z(f)|\, e^{j\theta(f)} \;, \tag{1.3.31}$$

then the **group delay** *(GD) of the signal is*

$$\tau_{\mathrm{g}}(f) = -\frac{1}{2\pi}\,\theta'(f) \tag{1.3.32}$$

and the **phase delay** *of the signal is*

$$\tau_{\rm p}(f) = -\frac{1}{2\pi f}\,\theta(f). \tag{1.3.33}$$

Eq. (1.3.32) is found by differentiating Eq. (1.3.27) w.r.t. f, and Eq. (1.3.33) is found by putting $f_{\rm c} = f$ in Eq. (1.3.27). Whereas Eq. (1.3.27) assumes linear phase, the above definition is meaningful whether the phase is linear or not.

Further, Eq. (1.3.32) is the same as Eq. (1.3.21). But the physical interpretations are different; the time delay applies to an impulse, whereas the group delay applies to the envelope of a narrowband signal.

Now consider the dual of the above argument. Let $z(t)$ be a time-limited signal centered on $t = t_{\rm c}$, and let

$$z(t) = a(t - t_{\rm c})\, e^{j\phi(t)} \tag{1.3.34}$$

where $a(t)$ and $\phi(t)$ are real. If the signal has **constant instantaneous frequency** in the support of $z(t)$, i.e. if $\phi(t)$ is a linear function of t wherever $z(t)$ is non-zero, we can let

$$\phi(t) = 2\pi f_0 t_{\rm c} + 2\pi f_m[t - t_{\rm c}] \tag{1.3.35}$$

where f_0 and f_m are real constants with the dimensions of frequency. Eq. (1.3.34) then becomes

$$z(t) = a(t - t_{\rm c})\, e^{j(2\pi f_0 t_{\rm c} + 2\pi f_m[t - t_{\rm c}])} \tag{1.3.36}$$

$$= e^{j2\pi f_0 t_{\rm c}}\, a(t - t_{\rm c})\, e^{j2\pi f_m[t - t_{\rm c}]}. \tag{1.3.37}$$

Taking the FT of $z(t)$, we find

$$Z(f) = A(f - f_m)\, e^{-j2\pi[f - f_0]t_{\rm c}} \tag{1.3.38}$$

where $A(f)$ is the FT of $a(t)$. Now because $a(t)$ is real, $A(f)$ is Hermitian, so that $|A(f)|$ is even. Hence f_m is the frequency about which the amplitude spectrum is symmetrical; for this reason, f_m is called the **mean IF**. Differentiating Eq. (1.3.35) w.r.t. t leads to the following definition:

Definition 1.3.5: *For the signal*

$$z(t) = |z(t)|\, e^{j\phi(t)}\ , \tag{1.3.39}$$

the **mean IF** *is*

$$f_m(t) = \frac{1}{2\pi}\,\phi'(t). \tag{1.3.40}$$

Thus the mean IF is the same as the IF defined earlier [Eq. (1.3.1)], but the physical interpretations are different. The IF has been derived for a tone (and earlier for a modulated sinusoid), whereas the mean IF applies to the spectrum of a short-duration signal.

1.3.4 Relaxation Time, Dynamic Bandwidth

For a linear FM signal, the instantaneous phase $\phi(t)$ is quadratic. So $\phi(t)$ can be expanded in a Taylor series about $t = t_0$:

$$\phi(t) = \phi(t_0) + \phi'(t_0)[t-t_0] + \tfrac{1}{2}\phi''(t_0)[t-t_0]^2 \tag{1.3.41}$$

$$= \phi(t_0) + 2\pi f_i(t_0)[t-t_0] + \tfrac{1}{2}2\pi f_i'(t_0)[t-t_0]^2. \tag{1.3.42}$$

The **relaxation time** T_r, as defined by Rihaczek [13, p. 374], is the duration over which the instantaneous phase deviates no more than $\pi/4$ from linearity. That is,

$$\left|\tfrac{1}{2}2\pi f_i'(t_0)T_r^2/4\right| = \pi/4. \tag{1.3.43}$$

Solving this equation leads to the following definition:

Definition 1.3.6: *The* **relaxation time** *of a signal is*

$$T_r(t) = \left|\frac{df_i(t)}{dt}\right|^{-1/2} \tag{1.3.44}$$

where $f_i(t)$ is the instantaneous frequency.

The dual of relaxation time, known as **dynamic bandwidth**, is the bandwidth over which the phase spectrum, assumed to be a quadratic function of frequency, deviates no more than $\pi/4$ from linearity. The result is as follows:

Definition 1.3.7: *The* **dynamic bandwidth** *of a signal is*

$$B_d(f) = \left|\frac{d\tau_d(f)}{df}\right|^{-1/2} \tag{1.3.45}$$

where $\tau_d(f)$ is the time delay.

As the relaxation time is a measure of the time needed to observe significant variations in IF, so the dynamic bandwidth is a measure of the bandwidth needed to observe significant variations in time delay.

1.4 Summary and Discussion

Clear rationales are developed for justifying the need to use joint time-frequency representations for non-stationary signals such as FM signals.

An understanding of the concept of instantaneous frequency and its dual, time delay, is necessary for the interpretation of TFDs. The use of the analytic associate of a given real signal, rather than the signal itself, is useful for reducing the required sampling rate, and essential for obtaining an unambiguous instantaneous frequency.

The analytic associate is obtained via the Hilbert transform. The notions of BT product and monocomponent signals are introduced. For monocomponent asymptotic signals, analytic associates can be written using an exponential form. The next chapter introduces various formulations of time-frequency distributions and demonstrates the importance of using the analytic associate.

References

[1] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[2] B. Boashash, "Wigner analysis of time-varying signals—Its application in seismic prospecting," in *Proc. European Signal Processing Conf. (EUSIPCO-83)*, pp. 703–706, Nürnberg, September 1983.

[3] B. Boashash and B. Escudie, "Wigner-Ville analysis of asymptotic signals and applications," *Signal Processing*, vol. 8, pp. 315–327, June 1985.

[4] T. D. Rossing, *The Science of Sound.* Reading, MA: Addison-Wesley, 2nd ed., 1990.

[5] J. S. Bach (1685-1750), *St Matthew Passion*, nos. 2–3 (ed. A. Van Ryckeghem). *GMD Music Archive*, http://www.gmd.de/Misc/Music/scores/Introduction.html, 2000. Continued by *Werner Icking Music Archive*, http://icking-music-archive.sunsite.dk/ .

[6] A. B. Carlson, *Communication Systems.* Tokyo: McGraw-Hill, 2nd ed., 1975.

[7] B. Boashash, "Note on the use of the Wigner distribution for time-frequency signal analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 36, pp. 1518–1521, September 1988.

[8] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner Distribution—A tool for time-frequency signal analysis," *Philips J. of Research*, vol. 35, pp. 217–250 (Part 1), 276–300 (Part 2) & 372–389 (Part 3), 1980.

[9] F. B. Hildebrand, *Advanced Calculus for Engineers.* New York: Prentice-Hall, 1949.

[10] D. Gabor, "Theory of communication," *J. IEE*, vol. 93(III), pp. 429–457, November 1946.

[11] D. Slepian, "On bandwidth," *Proc. IEEE*, vol. 64, pp. 292–300, March 1976.

[12] A. Papoulis, *Signal analysis.* New York: McGraw-Hill, 1977.

[13] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Information Theory*, vol. 14, pp. 369–374, May 1968.

This Page Intentionally Left Blank

Chapter 2

Heuristic Formulation of Time-Frequency Distributions[0]

Having established the basic signal formulations in the first chapter, we now turn to the problem of representing signals in a joint time-frequency domain. Given an analytic signal $z(t)$ obtained from a real signal $s(t)$, we seek to construct a time-frequency distribution $\rho_z(t,f)$ to represent precisely the energy, temporal and spectral characteristics of the signal. We choose the symbol ρ_z in the expectation that the TFD will represent an "energy density of z" in the (t,f) plane. We would also like the constant-t cross-section of $\rho_z(t,f)$ to be some sort of "instantaneous spectrum" at time t.

In this chapter we examine a variety of *ad hoc* approaches to the problem, namely the Wigner-Ville distribution (Section 2.1), a time-varying power spectral density called the Wigner-Ville Spectrum (2.2), localized forms of the Fourier Transform (2.3), filter banks (2.4), Page's instantaneous power spectrum (2.5), and related energy densities (2.6). Finally (in Section 2.7), we show how all these distributions are related to the first-mentioned Wigner-Ville distribution, thus setting the scene for the more systematic treatment in the next chapter.

The various distributions are illustrated using a linear FM asymptotic signal. The linear FM signal [Eq. (1.1.5)] is regarded as the most basic test signal for TFDs because it is the simplest example of a signal whose frequency content varies with time. It is clearly monocomponent, and is asymptotic if its BT product is large. The minimum requirement for a useful TFD is that it clearly shows the IF law of an asymptotic linear FM signal, giving a reasonable concentration of energy about the IF law (which, for an asymptotic signal, is equivalent to the TD law).

---

[0]Author: **Boualem Boashash**, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Australia. Reviewers: K. Abed-Meraim, A. Beghdadi, M. Mesbah, G. Putland and V. Sucic.

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

2.1 Method 1: The Wigner-Ville Distribution

For a monocomponent signal, it is reasonable to expect that the TFD should take the form of a knife-edge ridge whose crest is a graph of the IF law in the (t, f) plane. Mathematically we idealize the "knife edge" as a delta function w.r.t. frequency.

2.1.1 Knife-Edge IF Indication

Noting that ρ_z is a function of frequency f and represents a kind of spectrum, we may reasonably require ρ_z to be the FT of some function related to the signal. Let us call this function the **signal kernel** and give it the symbol K_z. So the signal kernel can be written as $K_z(t,\tau)$, and the TFD is

$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\{K_z(t,\tau)\}. \tag{2.1.1}$$

2.1.2 Formulation of the Signal Kernel

To find a suitable form for $K_z(t,\tau)$, for simplicity, let us first consider the case of the unit-amplitude monocomponent FM signal

$$z(t) = e^{j\phi(t)} \tag{2.1.2}$$

whose instantaneous frequency is

$$f_i(t) = \frac{\phi'(t)}{2\pi}. \tag{2.1.3}$$

We would like the TFD of $z(t)$ at any given time to be a unit delta function at the instantaneous frequency, so that the "instantaneous spectrum" reduces to the ordinary Fourier transform in the constant-frequency case; that is, we want

$$\rho_z(t,f) = \delta(f - f_i(t)). \tag{2.1.4}$$

Substituting this into Eq. (2.1.1) and taking the IFT, we obtain

$$\begin{aligned} K_z(t,\tau) &= \underset{\tau\leftarrow f}{\mathcal{F}^{-1}}\{\delta(f - f_i(t))\} = e^{j2\pi f_i(t)\tau} \\ &= e^{j\phi'(t)\tau}. \end{aligned} \tag{2.1.5}$$

Because $\phi'(t)$ in Eq. (2.1.5) is not directly available, we write

$$\phi'(t) = \lim_{\tau\to 0} \frac{\phi(t+\frac{\tau}{2}) - \phi(t-\frac{\tau}{2})}{\tau}, \tag{2.1.6}$$

and use the approximation

$$\phi'(t) \approx \frac{1}{\tau}\left[\phi(t+\tfrac{\tau}{2}) - \phi(t-\tfrac{\tau}{2})\right] \tag{2.1.7}$$

which is called the **central finite-difference** (CFD) approximation [1,2]. Substituting Eq. (2.1.7) into Eq. (2.1.5) and using Eq. (2.1.2) gives the signal kernel

$$\begin{aligned} K_z(t,\tau) &= e^{j\phi(t+\frac{\tau}{2})}\, e^{-j\phi(t-\frac{\tau}{2})} \\ &= z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\ . \end{aligned} \tag{2.1.8}$$

2.1.3 The Wigner Distribution

Substituting Eq. (2.1.8) into Eq. (2.1.1), we obtain

$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\big\{z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\big\} \tag{2.1.9}$$

$$= \int_{-\infty}^{\infty} z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\, e^{-j2\pi f\tau}\, d\tau. \tag{2.1.10}$$

Eq. (2.1.10) is given the symbol $\mathcal{W}_z(t,f)$, and is called the **Wigner distribution** (WD) in honor of its discoverer,[1] who derived it in 1932 in a quantum-mechanical context [3].

The approximation in Eq. (2.1.7) is exact if ϕ is a linear function of t, i.e. if $\phi'(t)$ is constant; it is also exact if $\phi(t)$ is quadratic [4, p. 298], i.e. if $\phi'(t)$ is linear. Thus the WD gives an "unbiased" estimate of the IF for a complex linear FM signal.

The constant-frequency real signal

$$s(t) = \cos 2\pi f_c t \tag{2.1.11}$$

leads to the signal kernel

$$\begin{aligned} K_s(t,\tau) &= s(t+\tfrac{\tau}{2})\, s(t-\tfrac{\tau}{2}) \\ &= \cos 2\pi f_c(t+\tfrac{\tau}{2})\, \cos 2\pi f_c(t-\tfrac{\tau}{2}) \\ &= \tfrac{1}{2}\cos 2\pi f_c\tau + \tfrac{1}{2}\cos 2\pi 2f_c t. \end{aligned} \tag{2.1.12}$$

Taking Fourier transforms w.r.t. τ gives the WD

$$\begin{aligned} \mathcal{W}_s(t,f) &= \tfrac{1}{4}\delta(f-f_c) + \tfrac{1}{4}\delta(f+f_c) \\ &\quad + \tfrac{1}{2}[\cos 2\pi 2f_c t]\, \delta(f). \end{aligned} \tag{2.1.13}$$

The terms in $\delta(f \mp f_c)$ are naturally expected and arise because $s(t)$ may be expressed as a sum of complex sinusoids at frequencies $\pm f_c$. The term in $\delta(f)$ is an artifact arising because the **nonlinearity** of the WD causes interaction between the positive- and negative-frequency terms.

[1] E. P. Wigner (1902–1995) was born in Hungary, studied chemical engineering in Germany and eventually settled in the United States, where he specialized in mathematical physics. He was a joint winner of the 1963 Nobel Prize for Physics for his many contributions to particle physics, including his law of conservation of parity and his work on the strong nuclear force.

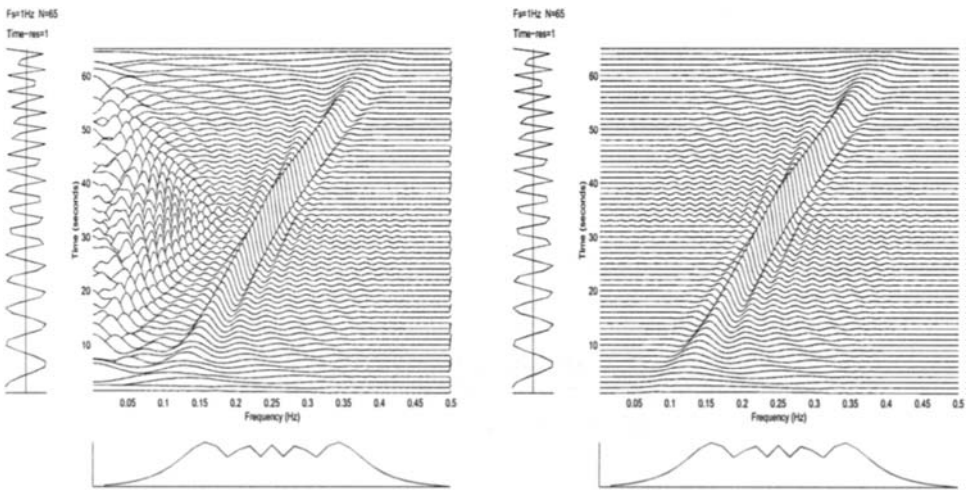


Fig. 2.1.1: Wigner distribution (left) and Wigner-Ville distribution (right) of a real linear FM signal [Eq. (1.1.5)] with duration 65 samples, starting frequency 0.1 and finishing frequency 0.4 (sampling rate 1 Hz). Note the low-frequency artifacts in the Wigner distribution, caused by interaction between positive and negative frequencies.

By a similar argument, we find that the non-windowed linear FM signal

$$s(t) = A\cos\left(2\pi[f_0 t + \tfrac{\alpha}{2}t^2]\right) \tag{2.1.14}$$

leads to the signal kernel

$$\begin{aligned} K_s(t,\tau) &= \tfrac{1}{2}A^2\cos 2\pi f_i(t)\tau \\ &\quad + \tfrac{1}{2}A^2\cos 2\pi\left[\tfrac{\alpha\tau^2}{4} + 2f_0 t + \alpha t^2\right] \end{aligned} \tag{2.1.15}$$

where $f_i(t) = f_0 + \alpha t$. Taking Fourier transforms w.r.t. τ gives the WD

$$\begin{aligned} W_s(t,f) &= \tfrac{1}{4}A^2\delta(f - f_i(t)) + \tfrac{1}{4}A^2\delta(f + f_i(t)) \\ &\quad + \tfrac{1}{2}A^2 \underset{\tau\to f}{\mathcal{F}}\left\{\cos 2\pi\left[\tfrac{\alpha\tau^2}{4} + 2f_0 t + \alpha t^2\right]\right\}. \end{aligned} \tag{2.1.16}$$

The terms in $\delta(f \mp f_i(t))$ are naturally expected, while the last term in the signal kernel gives rise to a continuum of artifacts in the WD (see Fig. 2.1.1).

These artifacts, which greatly diminish the usefulness of the WD for real signals, are removed by modifying the WD with the analytic signal in accordance with the following.

2.1.4 The Wigner-Ville Distribution

Definition 2.1.1: *The* **Wigner-Ville distribution (WVD)** *of a signal $s(t)$, denoted by $W_z(t,f)$, is defined as the WD of its analytic associate, i.e.*

$$W_z(t,f) = \underset{\tau \to f}{\mathcal{F}} \left\{ z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2}) \right\} \tag{2.1.17}$$

where $z(t)$ is the analytic associate of $s(t)$.

The name "Wigner-*Ville* distribution", as opposed to "Wigner distribution", emphasizes the use of the analytic signal [5] and recognizes the contribution of Ville [6], who derived the distribution in a signal-processing context in 1948. Noting that a signal can have a time-dependent frequency content, Ville sought an "instantaneous spectrum" having the attributes of an energy density (property 1 in Section 1.1.5) and satisfying the so-called **marginal conditions**:[2]

- the integral of the TFD $\rho_z(t,f)$ w.r.t. frequency is the instantaneous power $|z(t)|^2$;
- the integral of the TFD $\rho_z(t,f)$ w.r.t. time is the energy spectrum $|Z(f)|^2$.

By analogy with the conditional moments of a p.d.f., and using known relationships between the moments of a p.d.f. and its characteristic function, Ville was able to show that the distribution now known as the WVD had the desired properties [7, pp. 946–7]. Using Eq. (2.1.8), we obtain

$$W_z(t,f) = \underset{\tau \to f}{\mathcal{F}} \{ K_z(t,\overset{\bullet}{\tau}) \}. \tag{2.1.18}$$

The signal kernel $K_z(t,\tau)$ is also called the **instantaneous autocorrelation function** (IAF) of $z(t)$.

Furthermore, all TFDs in this Chapter, unless otherwise stated, are defined from the analytic associate of the signal, not from the real or "raw" signal.

2.1.4.1 The WVD of a Linear FM Signal

Eq. (2.1.2) describes a constant-amplitude, infinite-duration signal. We can allow for non-constant amplitude and finite duration using the form

$$z(t) = a(t)\, e^{j\phi(t)} \tag{2.1.19}$$

where $a(t)$ is real. For this signal we find

$$K_z(t,\tau) = K_a(t,\tau)\, e^{j\psi(t,\tau)} \tag{2.1.20}$$

[2]The name "marginal" can be explained with reference to the discrete-time, discrete-frequency case: if the TFD were written as a two dimensional array of discrete energies, each energy corresponding to a discrete time (vertical axis) and a discrete frequency (horizontal axis), then the sum over time for each frequency could be written in the horizontal "margin" of the array, and the sum over frequency for each time could be written in the vertical "margin".

where

$$K_a(t,\tau) = a(t+\tfrac{\tau}{2})\, a(t-\tfrac{\tau}{2}) \tag{2.1.21}$$

$$\psi(t,\tau) = \phi(t+\tfrac{\tau}{2}) - \phi(t-\tfrac{\tau}{2}). \tag{2.1.22}$$

If $\phi(t)$ is quadratic (i.e. if $f_i(t)$ is linear), then the CFD approximation is exact and gives $\psi(t,\tau) = \phi'(t)\tau = 2\pi f_i(t)\tau$, so that

$$K_z(t,\tau) = K_a(t,\tau)\, e^{j2\pi f_i(t)\tau}. \tag{2.1.23}$$

So $K_z(t,\tau)$, considered as a function of τ, has a constant frequency equal to $f_i(t)$; this reduction is called **dechirping**. If we let $\mathcal{W}_a(t,f) = \underset{\tau\to f}{\mathcal{F}}\{K_a(t,\tau)\}$, so that $\mathcal{W}_a(t,f)$ is the WD of $a(t)$, and take the FT of Eq. (2.1.23) w.r.t. τ, we obtain

$$W_z(t,f) = \mathcal{W}_a(t,f) \underset{f}{*} \delta(f - f_i(t)) \tag{2.1.24}$$

$$= \mathcal{W}_a(t, f - f_i(t)). \tag{2.1.25}$$

Now $K_a(t,\tau)$ is real and even in τ. Hence $\mathcal{W}_a(t,f)$ is real and even in f, so that $W_z(t,f)$ is real and symmetrical about $f = f_i(t)$.

Thus, even for a *finite*-duration linear FM signal, such as the one analyzed in Fig. 1.1.3, we expect the WVD to give a clear indication of the IF law. In fact, the TFD shown in Fig. 1.1.3 *is* the WVD of a finite duration linear FM signal.

2.1.4.2 The WVD in terms of the Spectrum

The variables t, f and τ are called time, frequency and lag, respectively. We now introduce the variable ν, which represents frequency *shift* just as τ represents time shift; accordingly, ν will be called **Doppler**.[3]

Let

$$k_z(\nu,f) = \underset{t\to\nu}{\mathcal{F}}\{W_z(t,f)\}. \tag{2.1.26}$$

Writing out the definitions of the FT and the WVD and taking invariant factors inside the integral signs, we obtain

$$k_z(\nu,f) = \iint z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\, e^{-j2\pi(f\tau+\nu t)} dt\, d\tau \tag{2.1.27}$$

where the integrals are from $-\infty$ to ∞. If we write

$$x = t + \tfrac{\tau}{2}\ ;\quad y = t - \tfrac{\tau}{2} \tag{2.1.28}$$

and solve for t and τ, obtaining

$$t = \tfrac{1}{2}(x+y)\ ;\quad \tau = x - y\ , \tag{2.1.29}$$

[3]The well-known "Doppler effect" is actually a frequency *scaling*. But when the effect is used to measure velocity, the scaling factor is usually close to unity and the bandwidth of interest is usually narrow. Under these conditions, the frequency scaling is well approximated by a frequency *shift* proportional to the velocity.

then the use of the Jacobian yields

$$dt\,d\tau = dx\,dy. \tag{2.1.30}$$

With these substitutions, we find that Eq. (2.1.27) can be factored into

$$\begin{aligned} k_z(\nu, f) &= \int_{-\infty}^{\infty} z(x)\, e^{-j2\pi[f+\nu/2]x} dx \cdot \int_{-\infty}^{\infty} z^*(y)\, e^{j2\pi[f-\nu/2]y} dy \\ &= Z(f+\tfrac{\nu}{2})\, Z^*(f-\tfrac{\nu}{2}) \end{aligned} \tag{2.1.31}$$

where $Z(f) = \mathcal{F}\{z(t)\}$. Noting that $k_z(\nu, f)$ has a similar form to $K_z(t, \tau)$ in the Doppler-frequency domain, we describe $k_z(\nu, f)$ as the "spectral autocorrelation function".

Substituting Eq. (2.1.31) into Eq. (2.1.26) and taking the IFT yields an expression for the WVD in terms of $Z(f)$:

$$W_z(t, f) = \int_{-\infty}^{\infty} Z(f+\tfrac{\nu}{2})\, Z^*(f-\tfrac{\nu}{2})\, e^{j2\pi\nu t} d\nu. \tag{2.1.32}$$

2.1.4.3 Effects of Time- and Frequency-Limiting

A practical signal $z(t)$ is often expected to be both time-limited and band-limited, despite theoretical constraints. Let us assume that $z(t)$ is windowed in both time and frequency.

For the time windowing, we can replace $z(t)$ by

$$z_w(t) = z(t)\, w(t - t_0). \tag{2.1.33}$$

The WVD of $z_w(t)$ is

$$W_{z_w}(t, f) = \int_{-\infty}^{\infty} z(t+\tfrac{\tau}{2})\, w(t-t_0+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\, w^*(t-t_0-\tfrac{\tau}{2})\, e^{-j2\pi f\tau} d\tau. \tag{2.1.34}$$

Putting $t_0 = t$ gives

$$W_{z_w}(t, f) = \int_{-\infty}^{\infty} g_2(\tau)\, K_z(t, \tau)\, e^{-j2\pi f\tau} d\tau \tag{2.1.35}$$

where

$$g_2(\tau) = w(\tfrac{\tau}{2})\, w^*(-\tfrac{\tau}{2}). \tag{2.1.36}$$

If $g_2(\tau) = 0$ for $|\tau| > T/2$, then the limits of integration in Eq. (2.1.35) may be changed to $\pm T/2$. Notice that $W_{z_w}(t, f)$ differs from the WVD of $z(t)$ in that the IAF is multiplied by $g_2(\tau)$ before being Fourier-transformed w.r.t. τ. This $g_2(\tau)$ is thus the effective lag window corresponding to the sliding time window $w(t - t_0)$. For the frequency windowing, we can replace $Z(f)$ by

$$Z_H(f) = Z(f)\, H(f - f_0). \tag{2.1.37}$$

Using Eq. (2.1.32), we can take the WVD corresponding to $Z_H(f)$ and put $f_0 = f$, obtaining in a similar manner to the above:

$$W_{Z_H}(t,f) = \int_{-\infty}^{\infty} G_1(\nu)\, Z(f+\tfrac{\nu}{2})\, Z^*(f-\tfrac{\nu}{2})\, e^{j2\pi\nu t} d\nu \tag{2.1.38}$$

where

$$G_1(\nu) = H(\tfrac{\nu}{2})\, H^*(-\tfrac{\nu}{2}). \tag{2.1.39}$$

If $G_1(\nu) = 0$ for $|\nu| > B/2$, then the limits of integration in Eq. (2.1.38) may be changed to $\pm B/2$. Notice that $W_{Z_H}(t,f)$ differs from the WVD of $z(t)$ in that the spectral autocorrelation function is multiplied by $G_1(\nu)$ before being inverse-Fourier-transformed w.r.t. ν. This $G_1(\nu)$ is the effective Doppler window corresponding to the sliding frequency window $H(f - f_0)$.

The effects of time windowing and frequency windowing may be combined as multiplication by the factor

$$g(\nu,\tau) = G_1(\nu)\, g_2(\tau) \tag{2.1.40}$$

in the (ν,τ) domain. If we consider time-limiting alone, the resulting TFD is given by Eq. (2.1.35) and is called the **windowed WVD** [8–10]. If we consider band-limiting alone, the resulting TFD is given by Eq. (2.1.38). We shall call this the **filtered WVD**. The window shapes of G_1 and g_2 should be selected to ensure the properties 1 to 3 in Section 1.1.5 are verified.

2.2 Method 2: Time-Varying Power Spectral Density

2.2.1 Spectra of Non-Stationary Random Processes

If $z(t)$ is a complex random signal, its symmetrical **autocorrelation** function is defined as

$$\mathcal{R}_z(t,\tau) = \mathcal{E}\{z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\} \tag{2.2.1}$$

where $\mathcal{E}\{\ldots\}$ denotes the expected value. If $z(t)$ is wide-sense stationary, then $\mathcal{R}_z(t,\tau)$ is independent of t, and the limit

$$\mathcal{S}_z(f) \triangleq \lim_{T\to\infty} \mathcal{E}\left\{ \frac{1}{T} \left| \underset{t\to f}{\mathcal{F}}\{z(t)\ \mathrm{rect}(t/T)\} \right|^2 \right\} \tag{2.2.2}$$

is called the **power spectral density (PSD)** of the random process $z(t)$, and describes the distribution of signal power over the frequencies. The PSD is related to the autocorrelation function by the **Wiener-Khintchine theorem**, which states that

$$\mathcal{S}_z(f) = \underset{\tau\to f}{\mathcal{F}}\{\mathcal{R}_z(t,\tau)\}. \tag{2.2.3}$$

If $z(t)$ is not wide-sense stationary, the right-hand side of this equation depends on t, so that the left-hand side also depends on t, suggesting the generalization

$$\mathcal{S}_z(t,f) \triangleq \underset{\tau \to f}{\mathcal{F}}\{\mathcal{R}_z(t,\tau)\}. \tag{2.2.4}$$

This $\mathcal{S}(t,f)$ may be called the time-varying spectrum or **evolutive spectrum** (ES), and is interpreted as the "time-dependent PSD" of the non-stationary signal. The right-hand side of Eq. (2.2.4) may be expanded as

$$\underset{\tau \to f}{\mathcal{F}}\{\mathcal{R}_z(t,\tau)\} = \int_{-\infty}^{\infty} \mathcal{E}\left\{z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\right\} e^{-j2\pi f\tau} d\tau \tag{2.2.5}$$

$$= \mathcal{E}\left\{\int_{-\infty}^{\infty} z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2}) e^{-j2\pi f\tau} d\tau\right\} \tag{2.2.6}$$

$$\mathcal{S}_z(t,f) = \mathcal{E}\{W_z(t,f)\}. \tag{2.2.7}$$

Eq. (2.2.7) shows that the expected value of the WVD is the FT of the time-dependent autocorrelation function [11]; that is, the ES is the expected value of the WVD. For this reason, the ES is also called the **Wigner-Ville spectrum**.

If $z(t)$ is deterministic, Eq. (2.2.7) reduces to

$$\mathcal{S}_z(t,f) = W_z(t,f). \tag{2.2.8}$$

2.2.2 Estimating the Wigner-Ville Spectrum

Eq. (2.2.7) refers to an ensemble average of the random process $W_z(t,f)$. If we have only one realization of this process, we may be able to assume that $W_z(\theta,f)$ is locally ergodic on the interval $t - \Delta/2 < \theta < t + \Delta/2$, where Δ is positive and independent of t. First we calculate an estimate of $W_z(t,f)$ for the local values of $z(t)$, yielding

$$\widehat{W}_z(t,f) = \underset{\tau \to f}{\mathcal{F}}\{g_2(\tau)\, K_z(t,\tau)\} \tag{2.2.9}$$

$$= G_2(f) \underset{f}{*} W_z(t,f) \tag{2.2.10}$$

where $g_2(\tau)$, the effective analysis window, is real, even and time-limited, and $G_2(f) = \mathcal{F}\{g_2(\tau)\}$.

Then, to obtain an estimate of $\mathcal{S}_z(t,f)$, we replace the ensemble average over all realizations of $W_z(t,f)$ by a time average of $\widehat{W}_z(t,f)$ over the interval Δ. We can calculate such an average using a real, even weighting function $g_1(t)$, with the properties

$$g_1(t) \begin{cases} > 0 & \text{if } |t| \le \Delta/2 \\ = 0 & \text{otherwise} \end{cases} \tag{2.2.11}$$

$$\int_{-\infty}^{\infty} g_1(t)\, dt = 1\ . \tag{2.2.12}$$

The resulting estimate of $\mathcal{S}_z(t,f)$ is

$$\widehat{\mathcal{S}}_z(t,f) = \int_{-\infty}^{\infty} g_1(\theta - t)\, \widehat{W}_z(\theta, f)\, d\theta. \tag{2.2.13}$$

Using the evenness of g_1 and substituting from Eq. (2.2.10), we have

$$\widehat{\mathcal{S}}_z(t,f) = g_1(t) \underset{t}{*} \widehat{W}_z(t,f) = g_1(t) \underset{t}{*} G_2(f) \underset{f}{*} W_z(t,f). \tag{2.2.14}$$

This can be written

$$\widehat{\mathcal{S}}_z(t,f) = \gamma(t,f) \underset{(t,f)}{**} W_z(t,f) \tag{2.2.15}$$

where

$$\gamma(t,f) = g_1(t)\, G_2(f) \tag{2.2.16}$$

and the double asterisk denotes **double convolution**.

Eq. (2.2.15) defines a class of estimates for time-varying spectra obtained by a double convolution of the WVD with a 2D filter. This will be used in the next chapter to define quadratic TFDs with specific properties.

2.3 Method 3: Windowed Fourier Transform (STFT, Spectrogram & Gabor Transform)

2.3.1 STFT and Spectrogram

Consider a signal $s(\tau)$ and a real, even window $w(\tau)$, whose FTs are $S(f)$ and $W(f)$ respectively. To obtain a localized spectrum of $s(\tau)$ at time $\tau = t$, multiply the signal by the window $w(\tau)$ centered at time $\tau = t$, obtaining

$$s_w(t,\tau) = s(\tau)\, w(\tau - t)\ , \tag{2.3.1}$$

and then take the FT w.r.t. τ, obtaining

$$F_s^w(t,f) = \underset{\tau \to f}{\mathcal{F}} \{ s(\tau)\, w(\tau - t) \}. \tag{2.3.2}$$

$F_s^w(t,f)$ is called the **short-time Fourier transform (STFT)**.

The squared magnitude of the STFT, denoted by $S_s^w(t,f)$, is called the **spectrogram**:

$$S_s^w(t,f) = |F_s^w(t,f)|^2 \tag{2.3.3}$$

$$= \left| \underset{\tau \to f}{\mathcal{F}} \{ s(\tau)\, w(\tau - t) \} \right|^2 \tag{2.3.4}$$

$$= \left| \int_{-\infty}^{\infty} s(\tau)\, w(\tau - t)\, e^{-j2\pi f \tau} d\tau \right|^2 . \tag{2.3.5}$$

In the notation $S_s^w(t,f)$, the upper-case S stands for "spectrogram", while the subscript stands for the signal and the superscript for the filter.

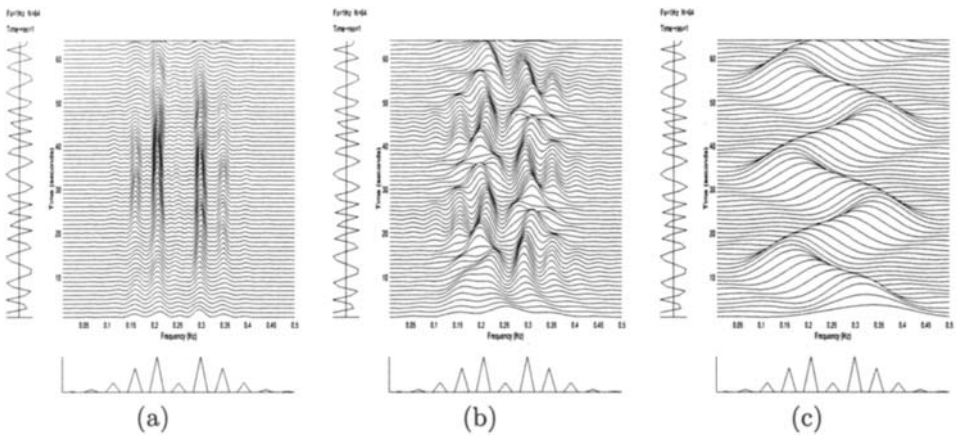


Fig. 2.3.1: Spectrogram of a sinusoidal FM signal [Eq. (1.1.4)] with 65 samples (sampling rate 1 Hz), $A = 1$, $f_c = 1/4$, $f_m = 3/64$, $f_d = 6/64$, rectangular window of length Δ samples: (a) $\Delta = 63$; (b) $\Delta = 23$; (c) $\Delta = 7$.

The observation window $w(\tau)$ allows localization of the spectrum in time, but also smears the spectrum in frequency[4] in accordance with the "uncertainty relationship" [12], leading to a trade-off between time resolution and frequency resolution. The problem is illustrated in Fig. 2.3.1, which shows the spectrogram of a sinusoidal FM signal for a rectangular window of three different lengths. If the window is long compared with the modulating signal, the frequency resolution is sufficient to show the sideband tones (the "multicomponent aspect" [13] of the signal), but the time resolution is insufficient to show the FM law (the "monocomponent aspect"). If the window is short compared with the modulating signal, the time resolution is sufficient to show the FM law but the frequency resolution is insufficient to show the sideband tones.

The spectrogram is nonlinear; but the nonlinearity is introduced only in the final step (taking the squared magnitude) and therefore does not lead to undesirable artifacts present in other TFDs. This freedom from artifacts, together with simplicity, robustness and ease of interpretation, has made the spectrogram a popular tool for speech analysis (resolution of speech into phonemes and formants) since its invention in 1946 [14].

2.3.2 Optimal Window Length of the Spectrogram

The spectrogram involves a compromise between time resolution and frequency resolution: a longer window provides less localization in time and more discrimination in frequency.

[4] **Smearing** is caused by the convolution operation. If an image (a function of two coordinates) is convolved with a confusion pattern (another function of the same two coordinates), the result is a blurred image. If the confusion pattern is a line, we tend to describe the blurring as a "smearing".

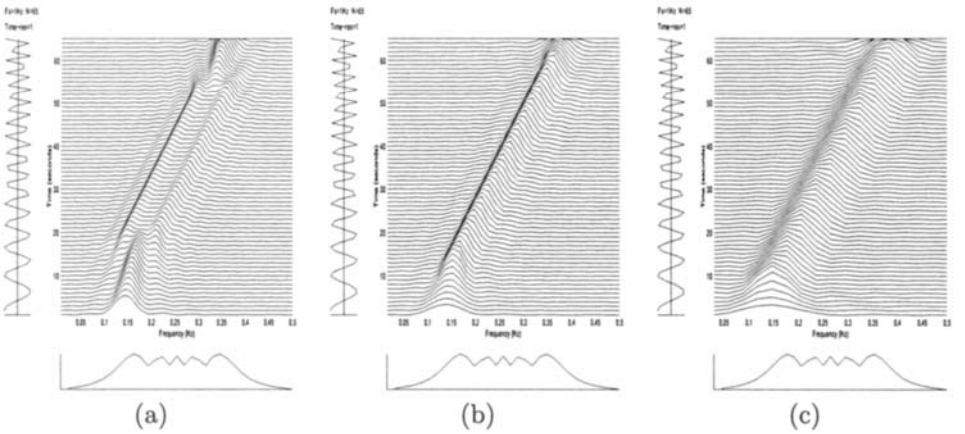


Fig. 2.3.2: Spectrogram of a linear FM signal [Eq. (1.1.5)] with duration 65, starting frequency 0.1 and finishing frequency 0.4, for a rectangular window of length Δ: (a) $\Delta = 33$; (b) $\Delta = 21$; (c) $\Delta = 11$. The optimal window length according to Eq. (2.3.6) is 20.7.

The purpose of the window is to obtain a time-slice of the signal during which the spectral characteristics are nearly constant. If the window is too long, it fails to capture the most rapid variations of spectral content. If it is too short, it smears the TFD in the frequency dimension without a commensurate improvement in detail in the time dimension. The more rapidly the spectral content changes, the shorter the window must be.

Hence, for a monocomponent signal of constant amplitude, the optimal window duration is inversely related to the rate of change of the IF. More precisely, if the window is rectangular and has total duration Δ, it can be shown [15, 16] that the value

$$\Delta = \sqrt{2} \left| \frac{df_i(t)}{dt} \right|^{-1/2} \tag{2.3.6}$$

is optimal in the sense that it minimizes the half-height width of the resulting ridge in the (t, f) plane. This optimal Δ is proportional to the relaxation time T_r; compare it with Eq. (1.3.44). For a linear FM signal, the optimal window duration simplifies to $\Delta = \sqrt{2T/B}$, where T is the signal duration and B is the signal bandwidth [17]. Fig. 2.3.2 shows the spectrogram of a chirp signal (the same signal as in Fig. 1.1.3) for a rectangular window of three different lengths, one of which ($\Delta = 21$) is optimal according to Eq. (2.3.6).

Even for the optimal window length, the spectrogram is not a delta function describing the IF law. The use of this optimal window is inconvenient because it requires knowledge of the IF, and this knowledge might be obtainable only by some sort of time-frequency analysis. Moreover, if the IF law is nonlinear, the optimal window duration varies with time. In the case of the sinusoid FM signal of Fig. 2.3.1, the optimal window length is time-varying if the signal is considered as a modulated

carrier. Although it is possible to vary the spectrogram window length with time and even with frequency, such procedures have a cost in computational efficiency [18]. A recent iterative algorithm for matching the spectrogram window to the estimated IF, so that the spectrogram of a monocomponent signal is concentrated along the IF law, is described in [19] and in Article 10.1.

2.3.3 STFT vs. Gabor Transform

In 1946, while studying the requirements for efficient signal transmission, Gabor [12] noted that the (t, f) plane can be divided into an array of rectangles using a bank of filters, each of which is switched on for a single interval of time and passes a single band of frequencies. Each rectangle was called a **logon**, and its dimensions were called the **decay time** and the **tuning width**. Gabor noted that the dimensions must satisfy the Heisenberg uncertainty relation

$$\Delta t \Delta f \geq \frac{1}{4\pi} \tag{2.3.7}$$

where Δt and Δf are the effective duration and bandwidth of the logon [12]. Gabor showed this relationship to be "at the root of the fundamental principle of communication" [12], in that it puts a lower limit on the spread of a signal in time and frequency.[5] For the minimum elemental area, which is obtained in the case of a complex Gaussian signal, Eq. (2.3.7) becomes an equality as in Eq. (1.2.34).

In Gabor's representation, each logon is assigned a complex coefficient $c_{n,k}$, where n is the time index and k the frequency index. The signal $s(t)$ is expanded in the doubly infinite series

$$s(t) = \sum_{n,k} c_{n,k}\, \psi_{n,k}(t) \tag{2.3.8}$$

where the sum is over all integers n and k, and $\psi_{n,k}(t)$ is a function centered about time $n\Delta t$ and frequency $k\Delta f$. To find the coefficients $c_{n,k}$, let $h_{n,k}(t)$ and $\psi_{n,k}(t)$ be related by

$$\int_{-\infty}^{\infty} h^*_{m,l}(t)\, \psi_{n,k}(t)\, dt = \begin{cases} 1 & \text{if } m=n \text{ and } l=k\ ; \\ 0 & \text{otherwise}\ . \end{cases} \tag{2.3.9}$$

[5] Eq. (2.3.7) is perhaps best known for its corollary in quantum mechanics. If we write x (position) for t, and k (wave number) for $2\pi f$, we obtain

$$\Delta x \Delta k \geq 1/2.$$

De Broglie's "matter wave" relation may be written $k = 2\pi p/h$, where p is the momentum and h is Planck's constant. Making this substitution for k in the above equation, we obtain

$$\Delta x \Delta p \geq \frac{h}{4\pi}\ ,$$

which, for some readers, will be more familiar than Eq. (2.3.7).

In other words, let $h_{m,l}(t)$ be orthogonal to every $\psi_{n,k}(t)$ except $\psi_{m,l}(t)$, or, equivalently, let $\psi_{m,l}(t)$ be orthogonal to every $h_{n,k}(t)$ except $h_{m,l}(t)$; functions related in this way are called **dual functions**. Multiplying Eq. (2.3.8) by $h^*_{m,l}(t)$ and integrating w.r.t. t, we obtain an expression for $c_{n,k}$:

$$c_{n,k} = \int_{-\infty}^{\infty} s(\tau)\, h^*_{n,k}(\tau)\, d\tau. \tag{2.3.10}$$

If we choose

$$h_{n,k}(\tau) = w(\tau - n\Delta t)\, e^{j2\pi k \Delta f \tau} \tag{2.3.11}$$

where w denotes a real Gaussian function, then Eq. (2.3.10) becomes

$$c_{n,k} = \int_{-\infty}^{\infty} s(\tau)\, w(\tau - n\Delta t)\, e^{-j2\pi k \Delta f \tau} d\tau \tag{2.3.12}$$

$$= \underset{\tau \to k\Delta f}{\mathcal{F}} \{s(\tau)\, w(\tau - n\Delta t)\}. \tag{2.3.13}$$

Expression (2.3.13) is known as the **Gabor transform**. When $\psi_{n,k}(t)$ is chosen as the dual function of $h_{n,k}(t)$, Eq. (2.3.8) gives the **inverse Gabor transform**. Eq. (2.3.13) has the same form as Eq. (2.3.2) except that t and f are discretized, wherefore the Gabor transform has been described as a sampled STFT. Suitable selections of ψ and h need to be made for the TFD to verify the properties 1 to 3 in Section 1.1.5, at least approximately.

Lerner [20] extended Gabor's work by allowing the elements of the (t, f) plane to be non-rectangular. Helstrom [21] generalized the expansion by replacing the discrete elementary cell weighting with a continuous function $\xi(\tau, t, f)$. Wavelet theory [22] was developed later as a further extension of Gabor's work, but choosing the partitions of the (t, f) plane so as to give constant-Q filtering [23, 24].

2.4 Method 4: Filtered Function of Time

2.4.1 Filter Banks and the Sonograph

Whereas the spectrogram is conceived as a function of frequency with dependence on the timing of a window, the sonograph is conceived as a function of time with dependence on the tuning of a filter. Consider a signal $s(t)$ with spectrum $S(\nu)$, and a lowpass filter with the real impulse response $h(t)$ and the transfer function $H(\nu)$. To extract the bandpass "component" of $s(t)$ at frequency $\nu = f$, we shift the filter function in frequency so that it is centered at $\nu = f$, multiply the signal spectrum by the shifted filter function, and take the inverse FT w.r.t. ν, obtaining

$$B^H_s(t, f) = \underset{t \leftarrow \nu}{\mathcal{F}^{-1}} \{S(\nu)\, H(\nu - f)\}. \tag{2.4.1}$$

where the "B" stands for "bandpass". This signal may be understood as the complex output of a bandpass filter with center frequency $+f$; the input $s(t)$ may be reconstructed as the sum of the outputs of such **filter banks** whose transfer functions add up to unity within the bandwidth of $s(t)$.

The squared magnitude of $B_s^H(t,f)$ is denoted by $S_s^{(H)}(t,f)$ and is called the **sonograph** or **sonogram**:

$$S_s^{(H)}(t,f) = |B_s^H(t,f)|^2 \tag{2.4.2}$$

$$= \left| \mathcal{F}_{t \leftarrow \nu}^{-1} \{S(\nu)\, H(\nu - f)\} \right|^2 \tag{2.4.3}$$

$$= \left| \int_{-\infty}^{\infty} S(\nu)\, H(\nu - f)\, e^{j2\pi\nu t} d\nu \right|^2 . \tag{2.4.4}$$

For the sonograph, the optimal bandwidth of the band-pass filter is related to the time delay in the same way that the optimal window length for the spectrogram is related to the instantaneous frequency.

Like the spectrogram, the sonograph is nonlinear, but the nonlinearity is introduced only in the final step and does not lead to artifacts. Filter banks—the practical realization of the sonograph—have long been used in music broadcasting, recording and even home entertainment, although the frequency division is not necessarily linear; for example, graphic equalizers and analyzers usually have logarithmic frequency division.

2.4.2 Equivalence to Spectrogram

Theorem 2.4.1: *The spectrogram and sonograph are equal if the window function of the spectrogram is real and even and equal to the impulse response of the sonograph filter for* $f = 0$.

Proof: Applying the inverse convolution property to Eq. (2.4.1), we obtain

$$B_s^H(t,f) = s(t) \underset{t}{*} h(t) e^{j2\pi f t} \tag{2.4.5}$$

$$= \int_{-\infty}^{\infty} s(\tau)\, h(t-\tau)\, e^{j2\pi f(t-\tau)} d\tau \tag{2.4.6}$$

$$= e^{j2\pi f t} \underset{\tau \to f}{\mathcal{F}} \{s(\tau)\, h(t-\tau)\} \tag{2.4.7}$$

which yields

$$S_s^{(H)}(t,f) = \left| \underset{\tau \to f}{\mathcal{F}} \{s(\tau)\, h(t-\tau)\} \right|^2 . \tag{2.4.8}$$

Comparing Eqs. (2.3.4) and (2.4.8), we see that the spectrogram and sonograph are the same if

$$h(t) = w(-t) , \tag{2.4.9}$$

which is the case if $w(t)$ is even and equal to $h(t)$. The condition that $w(t)$ be real is redundant in the proof, but is assumed in the definition of the spectrogram. ■

2.5 Method 5: Instantaneous Power Spectra

2.5.1 Page Distribution

This approach attempts to define a kind of "running spectrum" by considering the variations of the signal spectrum as time t increases.

Given a signal $s(t)$, let us define the **running transform** $S_-(t,f)$ as the FT of the signal s up to time t. To do this, we first define the "auxiliary signal" $s_t(\theta)$ as

$$s_t(\theta) = \begin{cases} s(\theta) & \text{if } \theta \le t \\ 0 & \text{if } \theta > t\,. \end{cases} \tag{2.5.1}$$

Then the running transform is simply

$$S_-(t,f) = \underset{\theta\to f}{\mathcal{F}}\{s_t(\theta)\} = \int_{-\infty}^{t} s(\theta)\, e^{-j2\pi f\theta} d\theta. \tag{2.5.2}$$

As the normal energy spectrum is the squared magnitude of the FT, so the **running energy spectrum** up to time t, denoted by $e_s(t,f)$, is the squared magnitude of the running transform:

$$e_s(t,f) = |S_-(t,f)|^2 = S_-(t,f)\, S_-^*(t,f). \tag{2.5.3}$$

Differentiating each expression in this equation w.r.t. time, and denoting the time-derivative of $e_s(t,f)$ by $\mathrm{P}_s(t,f)$, we obtain

$$\mathrm{P}_s(t,f) = \frac{\partial}{\partial t}\Big[|S_-(t,f)|^2\Big] \tag{2.5.4}$$

$$= S_-(t,f)\,\frac{\partial}{\partial t}[S_-^*(t,f)] + S_-^*(t,f)\,\frac{\partial}{\partial t}[S_-(t,f)]\,. \tag{2.5.5}$$

This $\mathrm{P}_s(t,f)$, being the time-derivative of a time-dependent energy spectrum, may be understood as a kind of *time-dependent power spectrum.* It is now known as the **Page distribution** in honor of its discoverer [25].

By substituting the right-hand expression of Eq. (2.5.2) into Eq. (2.5.4), we obtain

$$\mathrm{P}_s(t,f) = \frac{\partial}{\partial t}\left[\left|\int_{-\infty}^{t} s(\theta)\, e^{-j2\pi f\theta} d\theta\right|^2\right] \tag{2.5.6}$$

which is the usual definition of the Page distribution. Using Eq. (2.5.2) to evaluate the partial derivatives in Eq. (2.5.5), we obtain the alternative expression

$$\mathrm{P}_s(t,f) = 2\,\mathrm{Re}\big\{s^*(t)\, S_-(t,f)\, e^{j2\pi ft}\big\} \tag{2.5.7}$$

or, substituting from Eq. (2.5.2) and writing $\tau = t - \theta$,

$$\mathrm{P}_s(t,f) = 2\,\mathrm{Re}\left\{\int_0^{\infty} s^*(t)\, s(t-\tau)\, e^{j2\pi f\tau}\, d\tau\right\}. \tag{2.5.8}$$

If $s(t)$ is real as Page assumed [25], Eq. (2.5.8) becomes

$$\mathrm{P}_s(t,f) = 2\int_0^{\infty} s(t)\, s(t-\tau)\cos(2\pi f\tau)\, d\tau. \tag{2.5.9}$$

The Page distribution of a time-limited linear FM signal is shown in Fig. 2.7.1(c). Notice that the distribution can take *negative values*, which are inconsistent with the notion of an energy *distribution*, but perfectly consistent with the notion of an energy *gradient* as defined in Eq. (2.5.6). As negative values compensate for earlier spurious positive values caused by the truncation of $s(t)$ to produce $s_t(\theta)$, the presence of negative values implies that energy is delocalized in the (t, f) domain.

2.6 Method 6: Energy Density

2.6.1 Rihaczek's Complex Energy Density

In search of a TFD localized in both time and frequency, Rihaczek [26] considered the energy of a complex deterministic signal over finite ranges of t and f, and allowed those ranges to become infinitesimal, obtaining what he called a **complex energy density**.

Here we offer a simpler derivation than that given by Rihaczek. The energy of a complex signal $z(t)$, with Fourier transform $Z(f)$, is

$$\begin{aligned} E &= \int_{-\infty}^{\infty} |z(t)|^2 dt = \int_{-\infty}^{\infty} z(t)\, z^*(t)\, dt = \int_{-\infty}^{\infty} z(t) \int_{-\infty}^{\infty} Z^*(f)\, e^{-j2\pi f t} df\, dt \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{R}_z(t, f)\, dt\, df \end{aligned} \tag{2.6.1}$$

where $\mathrm{R}_z(t, f)$, the energy density function, is defined by

$$\mathrm{R}_z(t, f) = z(t)\, Z^*(f)\, e^{-j2\pi f t}. \tag{2.6.2}$$

This $\mathrm{R}_z(t, f)$ is the **Rihaczek distribution** (RD). If we express $Z(f)$ and hence $Z^*(f)$ in terms of $z(\lambda)$ and use the substitution $\tau = t - \lambda$, we obtain the alternative form

$$\mathrm{R}_z(t, f) = \int_{-\infty}^{\infty} z(t)\, z^*(t - \tau)\, e^{-j2\pi f \tau} d\tau. \tag{2.6.3}$$

A distribution equivalent to Rihaczek's was derived earlier, in the context of quantum mechanics, by J. G. Kirkwood [27], so that the RD is also called the **Kirkwood-Rihaczek distribution** [28, p. 26].

From Eq. (2.6.2) it is easily verified that

$$\int_{-\infty}^{\infty} \mathrm{R}_z(t, f)\, df = |z(t)|^2 \tag{2.6.4}$$

$$\int_{-\infty}^{\infty} \mathrm{R}_z(t, f)\, dt = |Z(f)|^2. \tag{2.6.5}$$

That is, the RD satisfies the **marginal conditions** [see the discussion following Eq. (2.1.17)].

Integrating Eq. (2.6.4) w.r.t. t and Eq. (2.6.5) w.r.t. f, we obtain respectively

$$\int_{t_1}^{t_2} \int_{-\infty}^{\infty} \mathrm{R}_z(t,f)\, df\, dt = \int_{t_1}^{t_2} |z(t)|^2 dt \tag{2.6.6}$$

$$\int_{f_1}^{f_2} \int_{-\infty}^{\infty} \mathrm{R}_z(t,f)\, dt\, df = \int_{f_1}^{f_2} |Z(f)|^2 df. \tag{2.6.7}$$

The right-hand side of Eq. (2.6.6) is the energy in the time interval between t_1 and t_2, while the right-hand side of Eq. (2.6.7) is the energy in the frequency band between f_1 and f_2. Together, the two equations indicate that $\mathrm{R}_z(t,f)$ can be interpreted as an energy density over an arbitrary time interval *and* an arbitrary frequency band.

In Eq. (2.6.2), the RD has the signal $z(t)$ as a factor. It follows that the RD is zero at those times when $z(t)$ is zero; this property is called **strong time support**. Similarly we see that the RD is zero at those frequencies for which the spectrum $Z(f)$ is zero; this property is called **strong frequency support**.

2.6.2 Levin's Real Energy Density

The **Levin distribution (LD)** is simply the *real part* of the RD. It follows that the LD, like the RD, has strong time support and strong frequency support. Taking the real parts of Eqs. (2.6.4) to (2.6.5), we further conclude that the LD satisfies the marginal conditions and their corollaries.

Let the LD of the complex signal $z(t)$ be denoted by $\mathrm{L}_z(t,f)$. Taking the real part of Eq. (2.6.2), we obtain the definition

$$\mathrm{L}_z(t,f) = \mathrm{Re}\left\{ z(t)\, Z^*(f)\, e^{-j2\pi ft} \right\}. \tag{2.6.8}$$

Taking the real part of Eq. (2.6.3) yields the alternative expression

$$\mathrm{L}_z(t,f) = \mathrm{Re}\left\{ \int_{-\infty}^{\infty} z(t)\, z^*(t-\tau)\, e^{-j2\pi f\tau} d\tau \right\}. \tag{2.6.9}$$

If $z(t)$ is replaced by a real signal $s(t)$, Eq. (2.6.9) reduces to the cosine form obtained by Levin [29].

Historically, the distribution obtained by Levin was a modification of the Page distribution and a precursor of the RD. But it was first discovered in a quantum-mechanical context by Margenau and Hill [30]; so it is also called the **Margenau-Hill distribution** [28, p. 26].

2.6.3 Windowed Rihaczek and Levin Distributions

Because energy is a real quantity, the real part of the RD is more significant than the imaginary part or the magnitude. Hence a "plot of the RD" is usually a plot of the real part, i.e. the LD. Fig. 2.7.1(d) shows such a plot for a time-limited linear

FM signal. Although Eqs. (2.6.6) and (2.6.7) indicate that the energy of the TFD is well localized in the time and frequency dimensions separately, i.e. in strips parallel to the f and t axes, it does not follow that the energy is well localized in both dimensions at once. Indeed Fig. 2.7.1(d) shows that the peaks of the TFD are not confined to the IF law, but show many spurious features. The WVD of this signal [Fig. 2.7.1(a)] is much cleaner. Because the RD/LD performs so poorly on such a simple signal, it must be regarded as only of theoretical interest.

By comparison with the RD, the spectrogram is remarkably free of artifacts. Recall that the spectrogram is the squared magnitude of the STFT. So one way to reduce artifacts in the RD is to introduce the STFT as a factor [31] instead of the spectrum $Z(f)$, in Eq. (2.6.2). The resulting distribution, which might be called the **windowed Rihaczek distribution**, is

$$\rho_z(t,f) = z(t)\Big[\underset{\tau\to f}{\mathcal{F}}\{z(\tau)\,w(\tau-t)\}\Big]^* e^{-j2\pi ft} \tag{2.6.10}$$

where w is the window function of the STFT. As the real part of the RD is the LD, we might as well designate the real part of the windowed RD as the **windowed Levin distribution**. A windowed LD of a time-limited linear FM signal is shown in Fig. 2.7.1(e); note the reduction in artifacts compared with the conventional LD.

From the form of Eq. (2.6.10), we see that the windowed RD and the windowed LD have strong time support. Other properties of these distributions will emerge in later sections.

2.7 Relationship between TFDs

So far, we have considered six different approaches to defining a TFD. All the approaches seem natural and reasonable, and yet lead paradoxically to at least five different definitions of a TFD, all of them *quadratic* in the signal. Using the "signal kernel" approach [Eq. (2.1.1)], we may write for any TFD $\rho_z(t,f)$:

$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\{R_z(t,\tau)\}. \tag{2.7.1}$$

where $R_z(t,\tau)$ is found simply by taking the inverse FT of $\rho_z(t,f)$ w.r.t. f. In the case of the WVD, $R_z(t,\tau)$ is the instantaneous autocorrelation function (IAF), denoted by $K_z(t,\tau)$. In other cases we shall call $R_z(t,\tau)$ the **smoothed IAF**. The reason for this term will become apparent as we relate $R_z(t,\tau)$ to $K_z(t,\tau)$ for each of the TFDs that we have defined.

2.7.1 Spectrogram

From Eq. (2.3.4), the spectrogram with window function w can be rewritten in the generalized notation as

$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\{z(\tau)\,w(\tau-t)\}\Big[\underset{\tau\to f}{\mathcal{F}}\{z(\tau)\,w(\tau-t)\}\Big]^*. \tag{2.7.2}$$

Taking the IFT ($f \rightarrow \tau$) of both sides gives

$$\begin{aligned} R_z(t,\tau) &= [z(\tau)\, w(\tau-t)] \underset{\tau}{*} [z^*(-\tau)\, w^*(-\tau-t)] \\ &= \int_{-\infty}^{\infty} z(\lambda)\, w(\lambda-t)\, z^*(\lambda-\tau)\, w^*(\lambda-\tau-t)\, d\lambda \\ &= \int_{-\infty}^{\infty} z(u+\tfrac{\tau}{2})\, w(u-t+\tfrac{\tau}{2})\, z^*(u-\tfrac{\tau}{2})\, w^*(u-t-\tfrac{\tau}{2})\, du \end{aligned} \tag{2.7.3}$$

where $\lambda = u+\frac{\tau}{2}$ is the dummy variable in the convolution. Exploiting the evenness of w, this can be written

$$\begin{aligned} R_z(t,\tau) &= \int_{-\infty}^{\infty} w^*(t-u+\tfrac{\tau}{2})\, w(t-u-\tfrac{\tau}{2})\, z(u+\tfrac{\tau}{2})\, z^*(u-\tfrac{\tau}{2})\, du \\ &= G(t,\tau) \underset{t}{*} K_z(t,\tau) \end{aligned} \tag{2.7.4}$$

where

$$G(t,\tau) = w^*(t+\tfrac{\tau}{2})\, w(t-\tfrac{\tau}{2}) \tag{2.7.5}$$

and $K_z(t,\tau)$ is given by Eq. (2.1.8). $G(t,\tau)$ is called the **time-lag kernel.**[6] Eq. (2.7.4) defines the time-lag kernel as that which must be convolved in time with the IAF to obtain the smoothed IAF; the word "smoothed" refers to the convolution. We shall adopt this definition for *all* TFDs given by Eq. (2.7.1).

Eq. (2.7.5) gives the time-lag kernel for the spectrogram. As w is real, we may interchange w and w^*, leading to the conclusion that the time-lag kernel for the spectrogram is simply the IAF of the observation window function.

2.7.2 Wigner-Ville Distribution

The "smoothed IAF" for the WVD is

$$R_z(t,\tau) = K_z(t,\tau) = \delta(t) \underset{t}{*} K_z(t,\tau) \tag{2.7.6}$$

so that

$$G(t,\tau) = \delta(t). \tag{2.7.7}$$

In this trivial case, the "smoothing" makes no difference.

2.7.3 Rihaczek Distribution

Eq. (2.6.3) may be written

$$\rho_z(t,f) = \underset{\tau\rightarrow f}{\mathcal{F}}\{z(t)\, z^*(t-\tau)\}. \tag{2.7.8}$$

[6]The term "kernel" was used in this sense by Claasen and Mecklenbräuker [32]. To minimize the risk of confusion between the "time-lag kernel" $G(t,\tau)$ and the "signal kernel" $K_z(t,\tau)$, the latter is usually called the IAF in this book.

Taking the IFT gives

$$\begin{aligned} R_z(t,\tau) &= z(t)\,z^*(t-\tau) && (2.7.9)\\ &= \delta(t-\tfrac{\tau}{2}) \underset{t}{*} [z(t+\tfrac{\tau}{2})\,z^*(t-\tfrac{\tau}{2})] && (2.7.10)\\ &= \delta(t-\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau) && (2.7.11) \end{aligned}$$

so that

$$G(t,\tau) = \delta(t-\tfrac{\tau}{2}). \qquad (2.7.12)$$

2.7.4 Levin Distribution

Using Eq. (2.7.8) as the definition of the Rihaczek distribution, and taking the real part, we obtain for the Levin distribution

$$\begin{aligned} \rho_z(t,f) &= \mathrm{Re}\Big\{ \underset{\tau\to f}{\mathcal{F}} \{z(t)\,z^*(t-\tau)\} \Big\} && (2.7.13)\\ &= \tfrac{1}{2} \underset{\tau\to f}{\mathcal{F}} \{z(t)\,z^*(t-\tau)\} + \tfrac{1}{2} \Big[\underset{\tau\to f}{\mathcal{F}} \{z(t)\,z^*(t-\tau)\} \Big]^*. && (2.7.14) \end{aligned}$$

Taking the IFT, we obtain

$$\begin{aligned} R_z(t,\tau) &= \tfrac{1}{2}\,z(t)\,z^*(t-\tau) + \tfrac{1}{2}\,z^*(t)\,z(t+\tau) && (2.7.15)\\ &= \tfrac{1}{2}\delta(t-\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau) + \tfrac{1}{2}\delta(t+\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau)\\ &= \tfrac{1}{2}[\delta(t+\tfrac{\tau}{2}) + \delta(t-\tfrac{\tau}{2})] \underset{t}{*} K_z(t,\tau) && (2.7.16) \end{aligned}$$

so that

$$G(t,\tau) = \tfrac{1}{2}[\delta(t+\tfrac{\tau}{2}) + \delta(t-\tfrac{\tau}{2})]. \qquad (2.7.17)$$

2.7.5 Windowed Rihaczek Distribution

Eq. (2.6.10) can be written

$$\begin{aligned} \rho_z(t,f) &= z(t) \underset{\tau\to f}{\mathcal{F}} \{z^*(-\tau)\,w^*(-\tau-t)\}\, e^{-j2\pi ft} && (2.7.18)\\ &= z(t) \underset{\tau\to f}{\mathcal{F}} \{z^*(t-\tau)\,w^*(-\tau)\}. && (2.7.19) \end{aligned}$$

Taking the IFT gives

$$\begin{aligned} R_z(t,\tau) &= z(t)\,z^*(t-\tau)\,w^*(-\tau) && (2.7.20)\\ &= [\delta(t-\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau)]\,w^*(-\tau) && (2.7.21)\\ &= [w^*(-\tau)\,\delta(t-\tfrac{\tau}{2})] \underset{t}{*} K_z(t,\tau) && (2.7.22) \end{aligned}$$

so that

$$G(t,\tau) = w^*(-\tau)\,\delta(t-\tfrac{\tau}{2}). \qquad (2.7.23)$$

Because w is real and even, this reduces to

$$G(t,\tau) = w(\tau)\,\delta(t-\tfrac{\tau}{2}). \tag{2.7.24}$$

Comparing Eqs. (2.7.11) and (2.7.22), we see that the smoothed IAFs of the RD and the windowed RD differ by the factor $w^*(-\tau)$; that is, the distributions differ by a windowing operation in the lag domain prior to Fourier transformation from τ to f [31].

2.7.6 Windowed Levin Distribution

From Eq. (2.7.19), the real part of the windowed RD is

$$\tfrac{1}{2}\,z(t)\underset{\tau\to f}{\mathcal{F}}\{z^*(t-\tau)\,w^*(-\tau)\} + \tfrac{1}{2}\,z^*(t)\Big[\underset{\tau\to f}{\mathcal{F}}\{z^*(t-\tau)\,w^*(-\tau)\}\Big]^*. \tag{2.7.25}$$

Taking the IFT, we find that the smoothed IAF of the windowed LD is

$$\begin{aligned} R_z(t,\tau) &= \tfrac{1}{2}\,z(t)\,z^*(t-\tau)\,w^*(-\tau) + \tfrac{1}{2}\,z^*(t)\,z(t+\tau)\,w(\tau) \\ &= \tfrac{1}{2}\,[\delta(t-\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau)]\,w^*(-\tau) + \tfrac{1}{2}\,[\delta(t+\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau)]\,w(\tau) \\ &= \tfrac{1}{2}\,[w(\tau)\,\delta(t+\tfrac{\tau}{2}) + w^*(-\tau)\,\delta(t-\tfrac{\tau}{2})] \underset{t}{*} K_z(t,\tau) \end{aligned} \tag{2.7.26}$$

so that

$$G(t,\tau) = \tfrac{1}{2}\,[w(\tau)\,\delta(t+\tfrac{\tau}{2}) + w^*(-\tau)\,\delta(t-\tfrac{\tau}{2})] \tag{2.7.27}$$

or, because w is real and even,

$$G(t,\tau) = \tfrac{1}{2}\,w(\tau)\,[\delta(t+\tfrac{\tau}{2}) + \delta(t-\tfrac{\tau}{2})]. \tag{2.7.28}$$

Comparing Eqs. (2.7.16) and (2.7.26), and noting that w is real and even, we see that the smoothed IAFs of the LD and the windowed LD differ by the factor $w(\tau)$; that is, the distributions differ by a windowing operation in the lag domain [31].

2.7.7 Page Distribution

Rewriting Eq. (2.5.8) using the unit step function $u(t)$, we obtain for the Page distribution

$$\rho_z(t,f) = 2\,\mathrm{Re}\left\{\int_{-\infty}^{\infty} z^*(t)\,z(t-\lambda)\,u(\lambda)\,e^{j2\pi f\lambda}d\lambda\right\}. \tag{2.7.29}$$

With the substitution $\tau = -\lambda$, this becomes

$$\rho_z(t,f) = 2\,\mathrm{Re}\left\{\underset{\tau\to f}{\mathcal{F}}\{z^*(t)\,z(t+\tau)\,u(-\tau)\}\right\} \tag{2.7.30}$$

$$= \underset{\tau\to f}{\mathcal{F}}\{z^*(t)\,z(t+\tau)\,u(-\tau)\} + \Big[\underset{\tau\to f}{\mathcal{F}}\{z^*(t)\,z(t+\tau)\,u(-\tau)\}\Big]^*. \tag{2.7.31}$$

Taking the IFT, we obtain

$$\begin{aligned} R_z(t,\tau) &= z^*(t)\,z(t+\tau)\,u(-\tau) + z(t)\,z^*(t-\tau)\,u(\tau) && (2.7.32)\\ &= u(-\tau)[\delta(t+\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau)] \;+\; u(\tau)[\delta(t-\tfrac{\tau}{2}) \underset{t}{*} K_z(t,\tau)] && (2.7.33)\\ &= [u(-\tau)\,\delta(t+\tfrac{\tau}{2}) + u(\tau)\,\delta(t-\tfrac{\tau}{2})] \underset{t}{*} K_z(t,\tau) && (2.7.34) \end{aligned}$$

so that

$$\begin{aligned} G(t,\tau) &= u(-\tau)\,\delta(t+\tfrac{\tau}{2}) + u(\tau)\,\delta(t-\tfrac{\tau}{2}) && (2.7.35)\\ &= \delta(t-|\tfrac{\tau}{2}|). && (2.7.36) \end{aligned}$$

2.7.8 Relationship between the WVD and Other TFDs

By taking the FT of Eq. (2.7.4) w.r.t. τ and using specific forms for $G(t,\tau)$, all the considered TFDs can be written in the same form as Eq. (2.2.15):

$$\rho_z(t,f) = \gamma(t,f) \underset{(t,f)}{**} W_z(t,f) \tag{2.7.37}$$

where

$$\gamma(t,f) = \underset{\tau\to f}{\mathcal{F}}\{G(t,\tau)\} \tag{2.7.38}$$

is the TFD time-frequency kernel.

This then suggests that all the TFDs naturally introduced so far can be considered to be smoothed WVDs. This observation led to the design or rediscovery of several other TFDs that are briefly discussed next.

2.7.9 Other Popular TFDs

Having derived some TFDs by intuitive methods and then determined their time-lag kernels, let us now define a few more TFDs directly in terms of their time-lag kernels.

Name	$G(t,\tau)$
Windowed WVD (w-WVD)	$\delta(t)\,w(\tau)$
Sinc *or* Born-Jordan (BJ)	$\frac{1}{\lvert 2\alpha\tau\rvert}\,\mathrm{rect}\,\frac{t}{2\alpha\tau}$
Exponential *or* Choi-Williams (CW)	$\frac{\sqrt{\pi\sigma}}{\lvert\tau\rvert}e^{-\pi^2\sigma t^2/\tau^2}$
Windowed sinc *or* Zhao-Atlas-Marks (ZAM)	$w(\tau)\,\mathrm{rect}\,\frac{t}{2\tau/a}$
B-distribution (BD)	$\lvert\tau\rvert^{\beta}\cosh^{-2\beta}t$
Modified B-distribution (MBD)	$\frac{\cosh^{-2\beta}t}{\int_{-\infty}^{\infty}\cosh^{-2\beta}\xi\,d\xi}$

In the above table, the parameters α, σ, a and β are real and positive, and $w(\tau)$ is a window function. The CW distribution was defined in [33] and the ZAM distribution

in [34]. The BJ distribution, called the "sinc distribution" in [28, p. 26], was defined in [7,35] using an operational rule of Born and Jordan [36, p. 873]. Observe that the BJ distribution is a special case of the ZAM with $a = 1/\alpha$ and $w(\tau) = a/|2\tau|$. These distributions are illustrated for a linear FM signal in Fig. 2.7.1, parts (f)–(i). The BD, w-WVD, and MBD kernels are "separable", while the w-WVD kernel is also "Doppler-independent" and the MBD kernel is also "lag-independent"; these concepts are discussed later, especially in Article 5.7. In Fig. 2.7.1 parts (j)–(l) we have respectively plotted a Doppler-independent, a lag-independent, and a separable-kernel TFD of a linear FM signal. We shall see [e.g. in Table 3.3.3 on p. 77] that the separable and Doppler (lag)-independent kernels have been introduced in order to obtain certain desired TFD properties.

For convenience, Table 2.7.1 collects and tabulates the definitions of the TFDs that have been derived heuristically in this chapter, and of the six TFDs that have just been defined in terms of their time-lag kernels.

2.8 Summary and Discussion

Constructing a quadratic TFD from the *analytic associate* of a given real signal, rather than from the signal itself, avoids spurious terms caused by interference between positive-frequency and negative-frequency components.

Every TFD that we have derived heuristically is *quadratic in the signal*; that is, if the signal is scaled by a factor k, the TFD is scaled by a factor k^2. This is to be expected because

(a) Each TFD considered so far is related to some sort of *energy density*. For example, the Page distribution is defined as a gradient of energy, hence accommodating for the negative values that occur in the TFD.

(b) The signal has been assumed to be an effort variable or a flow variable (examples of effort variables are voltage, force, and pressure; the corresponding flow variables are current, linear velocity, and volume velocity), *and* power is proportional to the product of an effort variable and the corresponding flow variable, hence (in a linear system) to the *square* of the effort variable or of the flow variable.

(c) We have seen in Section 2.7 that every TFD considered in this chapter can be written as the FT of a smoothed IAF [Eq. (2.7.1)], which is the convolution of an auxiliary function (the time-lag kernel filter $G(t,\tau)$) and the ordinary IAF $K_z(t,\tau)$, which in turn is *quadratic* in the signal.

Eq. (2.7.4) effectively defines a TFD in terms of its time-lag kernel and is the key to the theory and design of quadratic TFDs, as will be detailed in the next chapter.

References

[1] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals," *Proc. IEEE*, vol. 80, pp. 520–538, April 1992.

Table 2.7.1: Special forms of selected quadratic TFDs (3rd column), together with their time-lag kernels (2nd column). The window $w(\tau)$ is assumed to be real and even. Integrals, unless otherwise noted, are from $-\infty$ to ∞. The forms involving double integrals are obtained by direct substitution of the time-lag kernel into Eq. (3.2.9). The w-WVD may also be so obtained. Other forms are quoted from this chapter.

Distribution	$G(t,\tau)$	$\rho_z(t,f)$
Wigner-Ville	$\delta(t)$	$\int_{-\infty}^{\infty} z(t+\frac{\tau}{2})\, z^*(t-\frac{\tau}{2})\, e^{-j2\pi f\tau} d\tau$
Levin	$\frac{1}{2}\big[\delta(t+\frac{\tau}{2})+\delta(t-\frac{\tau}{2})\big]$	$\mathrm{Re}\big\{z(t)\, Z^*(f)\, e^{-j2\pi ft}\big\}$
Born-Jordan	$\frac{1}{\lvert 2\alpha\tau\rvert}\,\mathrm{rect}\,\frac{t}{2\alpha\tau}$	$\int_{-\infty}^{\infty}\int_{t-\lvert\alpha\tau\rvert}^{t+\lvert\alpha\tau\rvert} \frac{1}{2\alpha\tau}\, z(u+\frac{\tau}{2})\, z^*(u-\frac{\tau}{2})\, e^{-j2\pi f\tau} du\, d\tau$
Modified B	$\frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty}\cosh^{-2\beta}\xi\, d\xi}$	$\iint \frac{\cosh^{-2\beta}(t-u)}{\int\cosh^{-2\beta}\xi\, d\xi}\, z(u+\frac{\tau}{2})\, z^*(u-\frac{\tau}{2})\, e^{-j2\pi f\tau} du\, d\tau$
w-WVD	$\delta(t)\, w(\tau)$	$\int_{-\infty}^{\infty} w(\tau)\, z(t+\frac{\tau}{2})\, z^*(t-\frac{\tau}{2})\, e^{-j2\pi f\tau} d\tau$
w-Levin	$\frac{w(\tau)}{2}\big[\delta(t+\frac{\tau}{2})+\delta(t-\frac{\tau}{2})\big]$	$\mathrm{Re}\Big\{z(t)\Big[\int_{-\infty}^{\infty} z(\tau)\, w(\tau-t)\, e^{-j2\pi f\tau} d\tau\Big]^*\, e^{-j2\pi ft}\Big\}$
ZAM	$w(\tau)\,\mathrm{rect}\,\frac{t}{2\tau/a}$	$\int_{-\infty}^{\infty}\int_{t-\lvert\frac{\tau}{a}\rvert}^{t+\lvert\frac{\tau}{a}\rvert} w(\tau)\, z(u+\frac{\tau}{2})\, z^*(u-\frac{\tau}{2})\, e^{-j2\pi f\tau} du\, d\tau$
Rihaczek	$\delta(t-\frac{\tau}{2})$	$z(t)\, Z^*(f)\, e^{-j2\pi ft}$
w-Rihaczek	$w(\tau)\,\delta(t-\frac{\tau}{2})$	$z(t)\Big[\int_{-\infty}^{\infty} z(\tau)\, w(\tau-t)\, e^{-j2\pi f\tau} d\tau\Big]^*\, e^{-j2\pi ft}$
Page	$\delta(t-\lvert\frac{\tau}{2}\rvert)$	$\frac{\partial}{\partial t}\Big[\Big\lvert\int_{-\infty}^{t} z(\tau)\, e^{-j2\pi f\tau} d\tau\Big\rvert^2\Big]$
Choi-Williams	$\frac{\sqrt{\pi\sigma}}{\lvert\tau\rvert} e^{-\pi^2\sigma t^2/\tau^2}$	$\iint \frac{\sqrt{\pi\sigma}}{\lvert\tau\rvert} e^{\frac{-\pi^2\sigma(t-u)^2}{\tau^2}}\, z(u+\frac{\tau}{2})\, z^*(u-\frac{\tau}{2})\, e^{-j2\pi f\tau} du\, d\tau$
B	$\lvert\tau\rvert^{\beta}\cosh^{-2\beta} t$	$\iint \frac{\lvert\tau\rvert^{\beta}}{\cosh^{2\beta}(t-u)}\, z(u+\frac{\tau}{2})\, z^*(u-\frac{\tau}{2})\, e^{-j2\pi f\tau} du\, d\tau$
Spectrogram	$w(t+\frac{\tau}{2})\, w(t-\frac{\tau}{2})$	$\Big\lvert\int_{-\infty}^{\infty} z(\tau)\, w(\tau-t)\, e^{-j2\pi f\tau} d\tau\Big\rvert^2$

[2] B. Boashash, P. J. O'Shea, and M. J. Arnold, "Algorithms for instantaneous frequency estimation: A comparative study," in *Proc. SPIE: Advanced Signal-Processing Algorithms, Architectures, and Implementations*, vol. 1348, pp. 126–148, Soc. of Photo-optical Instrumentation Engineers, San Diego, 10–12 July 1990.

[3] E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physics Review*, vol. 40, pp. 749–759, June 1932.

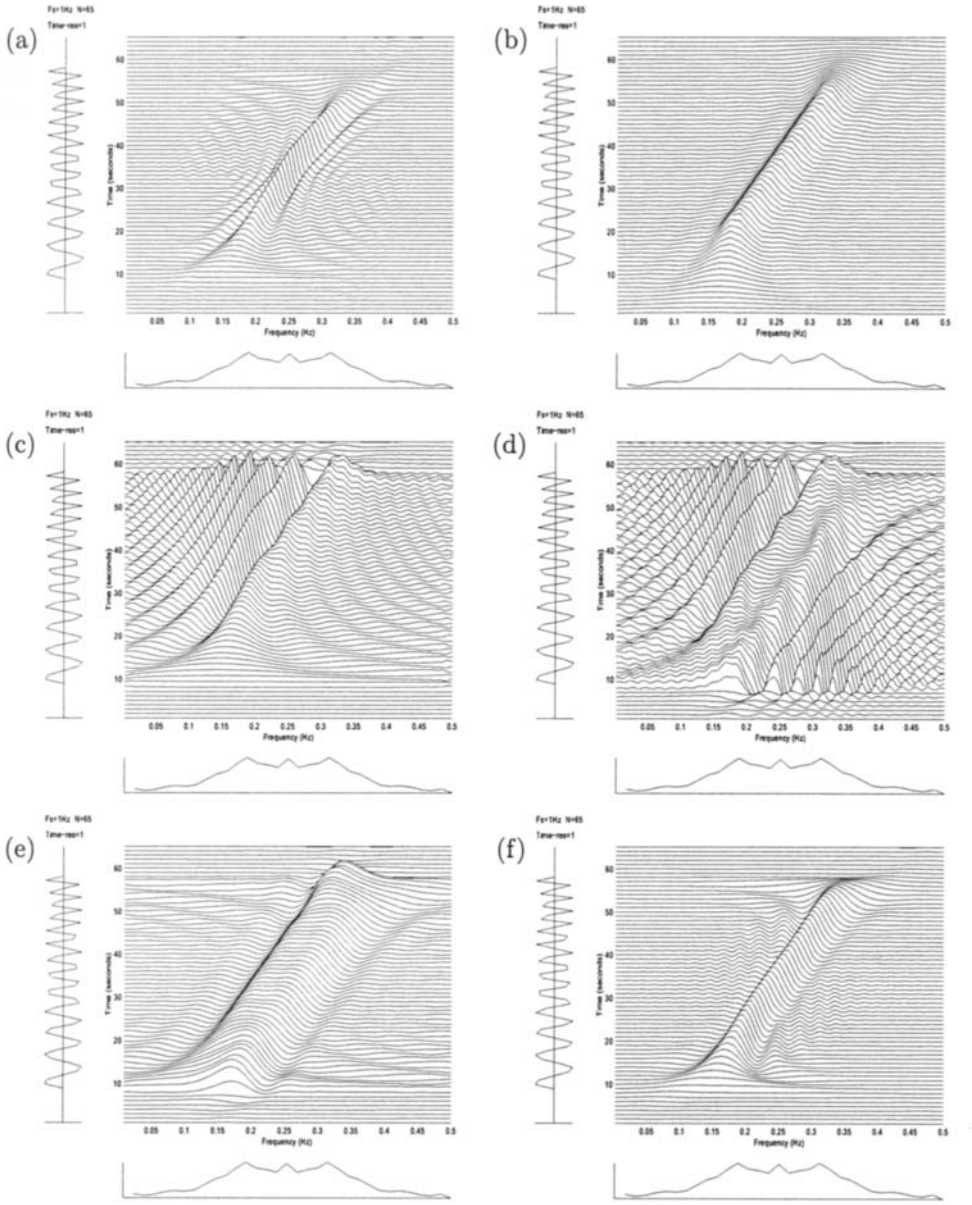


Fig. 2.7.1: TFDs of a linear FM signal with displayed duration 65 samples (sampling rate 1 Hz), unit amplitude from sample 9 to sample 57, zero amplitude elsewhere, frequency range 0.1 (at sample 1) to 0.4 (at sample 65): (a) Wigner-Ville; (b) Spectrogram, 21-point rectangular window; (c) Page; (d) Levin; (e) Windowed Levin, 19-point Hamming window; (f) Choi-Williams, $\sigma = 4$;... [continued].

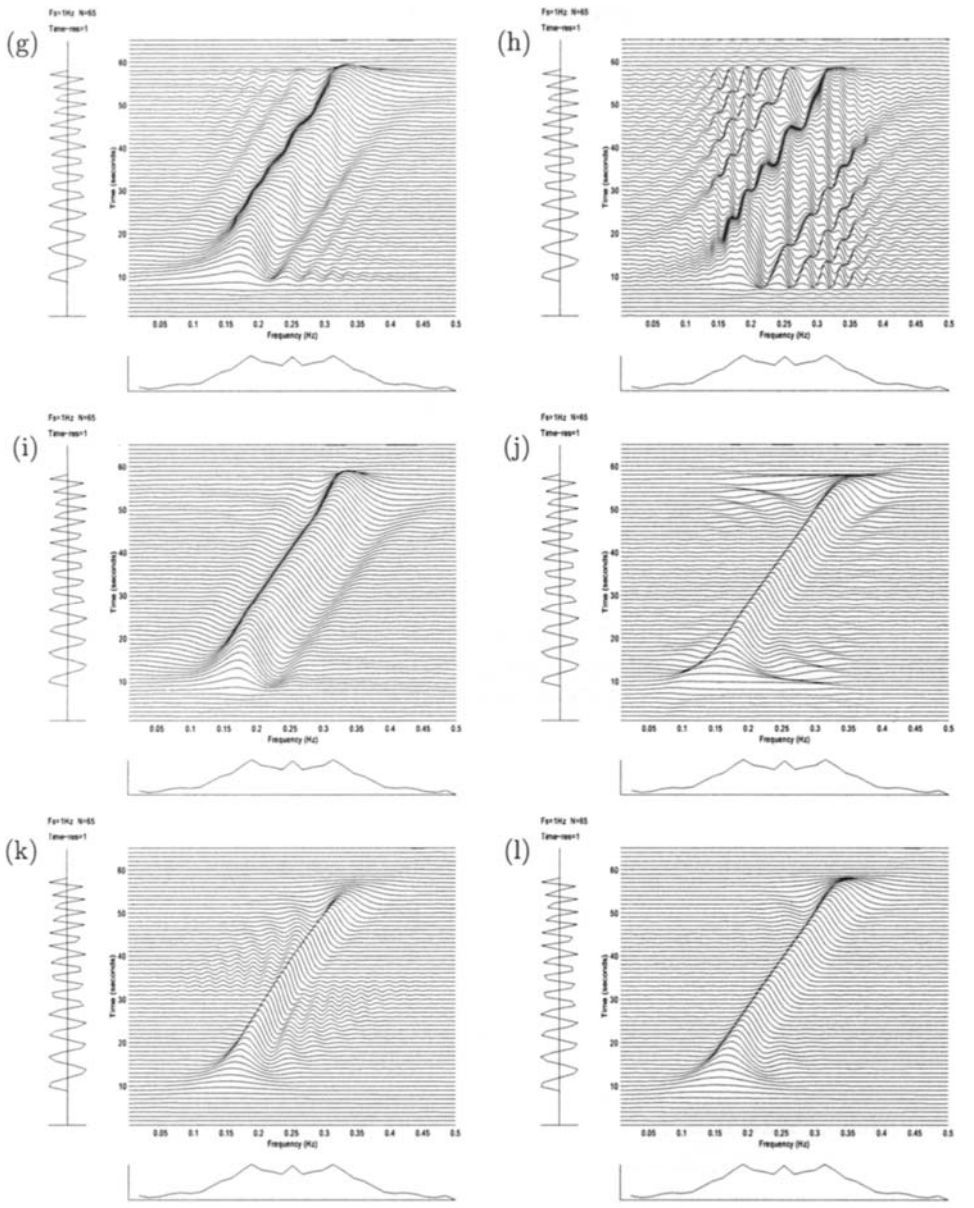


Fig. 2.7.1: [continuing] ...(g) Born-Jordan, $\alpha = 1/2$; (h) Zhao-Atlas-Marks, $a = 2$, $g_2(\tau) = 1$; (i) Zhao-Atlas-Marks, $a = 2$, $g_2(\tau) = 19$-point Hamming window; (j) Doppler-independent, $g_2(\tau) = 29$-point Hamming window; (k) Lag-independent, $g_1(t) = 9$-point Hanning; (l) Separable kernel, $g_1(t) = 9$-point Hanning, $g_2(\tau) = 29$-point Hamming.

[4] S. D. Conte and C. de Boor, *Elementary numerical analysis: An algorithmic approach.* Tokyo: McGraw-Hill, 3rd ed., 1980.

[5] B. Boashash, "Note on the use of the Wigner distribution for time-frequency signal analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 36, pp. 1518–1521, September 1988.

[6] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmissions*, vol. 2A, no. 1, pp. 61–74, 1948. In French. English translation: I. Selin, *Theory and applications of the notion of complex signal*, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).

[7] L. Cohen, "Time-frequency distributions—A review," *Proc. IEEE*, vol. 77, pp. 941–981, July 1989. Invited paper.

[8] B. Boashash and P. J. Black, "An efficient real-time implementation of the Wigner-Ville distribution," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 35, pp. 1611–1618, November 1987.

[9] B. Boashash and H. J. Whitehouse, "High resolution Wigner-Ville analysis," in *Eleventh GRETSI Symp. on Signal Processing and its Applications*, pp. 205–208, Nice, France, 1–5 June 1987.

[10] H. J. Whitehouse, B. Boashash, and J. M. Speiser, "High-resolution processing techniques for temporal and spatial signals," in *High-resolution methods in underwater acoustics* (M. Bouvet and G. Bienvenu, eds.), ch. 4, pp. 127–176, Berlin: Springer, 1991.

[11] W. Martin, "Time-frequency analysis of random signals," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'82)*, vol. 3, pp. 1325–1328, Paris, 3–5 May 1982.

[12] D. Gabor, "Theory of communication," *J. IEE*, vol. 93(III), pp. 429–457, November 1946.

[13] G. R. Putland and B. Boashash, "Can a signal be both monocomponent and multicomponent?," in *Third Australasian Workshop on Signal Processing Applications (WoSPA 2000)*, Brisbane, Australia, 14–15 December 2000. Paper no. 32.

[14] W. Koenig, H. K. Dunn, and L. Y. Lacy, "The sound spectrograph," *J. Acoustical Soc. of America*, vol. 18, no. 1, pp. 19–49, 1946.

[15] J. Imberger and B. Boashash, "Application of the Wigner-Ville distribution to temperature gradient microstructure: A new technique to study small-scale variations," *J. of Physical Oceanography*, vol. 16, pp. 1997–2012, December 1986.

[16] L. R. O. Storey, "An investigation of whistling atmospherics," *Phil. Trans. Roy. Soc.*, vol. A246, pp. 113–141, 1953.

[17] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[18] D. L. Jones and T. W. Parks, "A high-resolution data-adaptive time-frequency representation," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 2127–2135, December 1990.

[19] M. K. Emresoy and A. El-Jaroudi, "Iterative instantaneous frequency estimation and adaptive matched spectrogram," *Signal Processing*, vol. 64, pp. 157–65, January 1998.

[20] J. Kay and R. Lerner, *Lectures in Communications Theory.* McGraw-Hill, 1961.

[21] C. W. Helstrom, "An expansion of a signal in Gaussian elementary signals," *IEEE Trans. Information Theory*, vol. 12, pp. 81–82, January 1966.

[22] I. Daubechies, "The wavelet transform: A method for time-frequency localization," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 8, pp. 366–417, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[23] Y. Meyer, *Wavelets: Algorithms and applications.* Philadelphia, PA: Soc. for Industrial and Applied Mathematics, 1993. Translated and revised by Robert D. Ryan. Original French title: *Ondelettes et algorithems concurrents.*

[24] S. G. Mallat, *A Wavelet Tour of Signal Processing.* San Diego / London: Academic Press, 2nd ed., 1999.

[25] C. H. Page, "Instantaneous power spectra," *J. of Applied Physics*, vol. 23, pp. 103–106, January 1952.

[26] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Information Theory*, vol. 14, pp. 369–374, May 1968.

[27] J. G. Kirkwood, "Quantum statistics of almost classical ensembles," *Physics Review*, vol. 44, pp. 31–37, 1933.

[28] L. Cohen, "Introduction: A primer on time-frequency analysis," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 1, pp. 3–42, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[29] M. J. Levin, "Instantaneous spectra and ambiguity functions," *IEEE Trans. Information Theory*, vol. 10, pp. 95–97, January 1964.

[30] H. Margenau and R. N. Hill, "Correlation between measurements in quantum theory," *Progress of Theoretical Physics*, vol. 26, pp. 722–738, 1961.

[31] G. R. Putland. personal communications, Signal Processing Research Centre, Queensland University of Technology, 2001.

[32] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner Distribution—A tool for time-frequency signal analysis; Part 3: Relations with other time-frequency signal transformations," *Philips J. of Research*, vol. 35, no. 6, pp. 372–389, 1980.

[33] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 37, pp. 862–871, June 1989.

[34] Y. Zhao, L. E. Atlas, and R. J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[35] L. Cohen, "Generalized phase-space distribution functions," *J. of Mathematical Physics*, vol. 7, pp. 781–786, May 1966.

[36] M. Born and P. Jordan, "Zur quantenmechanik," *Zeitschrift für Physik.*, vol. 34, pp. 858–888, 1925.

This Page Intentionally Left Blank

Chapter 3

Theory of Quadratic TFDs[0]

The quadratic time-frequency distributions (TFDs) introduced in the last chapter represent the majority of the methods used in practical applications that deal with non-stationary signals. In this chapter, which completes the introductory tutorial, we show that these particular quadratic TFDs belong to a general class of TFDs whose design follows a common procedure, and whose properties are governed by common laws. This quadratic class may be considered as the class of *smoothed* Wigner-Ville distributions (WVDs), where the "smoothing" is described in the (t, f) domain by convolution with a "time-frequency kernel" function $\gamma(t, f)$, and in other domains by multiplication and/or convolution with various transforms of $\gamma(t, f)$. The generalized approach allows the definition of new TFDs that are better adapted to particular signal types, using a simple and systematic procedure as opposed to the *ad hoc* methods of Chapter 2.

The first section (3.1) extends Section 2.1 by enumerating in detail the key properties and limitations of the WVD. Thus it motivates the introduction of general quadratic TFDs (Section 3.2) and prepares for the discussion of their properties (Section 3.3). In Section 3.2, using Fourier transforms from lag to frequency and from time to Doppler (frequency shift), the quadratic TFDs and their kernels are formulated in four different but related two-dimensional domains. One of these, namely the Doppler-lag (ν, τ) domain, leads to the definition of the "ambiguity function" and allows the smoothing of the WVD to be understood as a filtering operation. In Section 3.3, the list of properties of the WVD is supplemented by mentioning some desirable TFD properties not shared by the WVD. The various TFD properties are then expressed in terms of constraints on the kernel, so that TFD design is reduced to kernel design. Three tables are provided showing the kernel properties equivalent to various TFD properties, the kernels of numerous popular TFDs in the various two-dimensional domains, and the properties of those same TFDs.

---

[0]Author: **Boualem Boashash**, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Australia. Reviewers: K. Abed-Meraim, A. Beghdadi, M. Mesbah, G. Putland and V. Sucic.

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

3.1 The WVD

3.1.1 Properties of the WVD

We have seen that the WVD has the simplest time-lag kernel (see Eq. (2.7.7)), and that each of the other TFDs can be written as a filtered WVD using a specific time-lag kernel filter (see Eq. (2.7.37)). In this sense the WVD is the basic or prototype TFD and the other TFDs are variations thereon. Moreover, we see in Fig. 2.7.1 that the WVD gives the sharpest indication of the IF law of a linear FM signal. These are some of the reasons why the WVD is the most widely studied TFD and deserves a further detailed description of its properties, as listed below.

- **Realness (RE):** $W_z(t,f)$ is real for all z, t and f.

- **Time-shift invariance** (also called **time covariance**): A time shift in the signal causes the same time shift in the WVD; that is, if

$$z_r(t) = z(t - t_0) \ , \tag{3.1.1}$$

  then

$$W_{z_r}(t,f) = W_z(t - t_0, f). \tag{3.1.2}$$

- **Frequency-shift invariance** (also called **frequency covariance**): A frequency shift in the signal causes the same frequency shift in the WVD; that is, if

$$z_r(t) = z(t)\, e^{j2\pi f_0 t} \ , \tag{3.1.3}$$

  then

$$W_{z_r}(t,f) = W_z(t, f - f_0). \tag{3.1.4}$$

- **Time marginal (TM):** Integration of the WVD over frequency gives the instantaneous power:

$$\int_{-\infty}^{\infty} W_z(t,f)\, df = |z(t)|^2. \tag{3.1.5}$$

- **Frequency marginal (FM):** Integration of the WVD over time gives the energy spectrum:

$$\int_{-\infty}^{\infty} W_z(t,f)\, dt = |Z(f)|^2. \tag{3.1.6}$$

- **Global energy:** Integration of the WVD over the entire (t,f) plane yields the signal energy E_z:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_z(t,f)\, dt\, df = E_z. \tag{3.1.7}$$

- **Instantaneous frequency (IF):** For an analytic signal, the first moment [i.e. the mean] of the WVD w.r.t. frequency is the IF:

$$\frac{\int_{-\infty}^{\infty} f W_z(t,f)\,df}{\int_{-\infty}^{\infty} W_z(t,f)\,df} = \frac{1}{2\pi}\frac{d}{dt}[\arg z(t)]\,. \tag{3.1.8}$$

- **Time delay (TD):** The first moment of the WVD w.r.t. time is the TD:

$$\frac{\int_{-\infty}^{\infty} t\, W_z(t,f)\,dt}{\int_{-\infty}^{\infty} W_z(t,f)\,dt} = -\frac{1}{2\pi}\frac{d}{df}[\arg Z(f)]\,. \tag{3.1.9}$$

  where $Z(f)$ is the FT of $z(t)$.

- **Time support (TS):** The time support of $W_z(t,f)$ is limited by the duration of $z(t)$; that is, if $z(t) = 0$ for $t < t_1$ and for $t > t_2$, then $W_z(t,f) = 0$ for $t < t_1$ and for $t > t_2$.

- **Frequency support (FS):** The frequency support of $W_z(t,f)$ is limited by the bandwidth of $z(t)$; that is, if $\mathcal{F}\{z(t)\} = Z(f) = 0$ for $f < f_1$ and for $f > f_2$, then $W_z(t,f) = 0$ for $f < f_1$ and for $f > f_2$.

- **Convolution invariance:** The WVD of the time-convolution of two signals is the time-convolution of the WVDs of the two signals; that is, if

$$z_3(t) = z_1(t) \underset{t}{*} z_2(t)\ , \tag{3.1.10}$$

  then

$$W_{z_3}(t,f) = W_{z_1}(t,f) \underset{t}{*} W_{z_2}(t,f). \tag{3.1.11}$$

- **Modulation invariance:** The WVD of the frequency-convolution of two signals is the frequency-convolution of the WVDs of the two signals; that is, if

$$z_3(t) = z_1(t)\, z_2(t)\ , \tag{3.1.12}$$

  then

$$W_{z_3}(t,f) = W_{z_1}(t,f) \underset{f}{*} W_{z_2}(t,f). \tag{3.1.13}$$

- **Invertibility:** If $W_z(t,f)$ is the WVD of the signal $z(t)$, it may be shown [1, pp. 223–4] that

$$\int_{-\infty}^{\infty} W_z(t/2,f)\, e^{j2\pi ft} df = z(t)\, z^*(0). \tag{3.1.14}$$

  Putting $t = 0$ in this result yields the magnitude, but not the phase, of $z(0)$. Hence a signal $z(t)$ may be recovered from its WVD *up to a phase factor.*

- **Inner-product invariance:** The WVD is a **unitary** transformation; that is, it preserves inner products:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{z_1}(t,f)\, W_{z_2}(t,f)\, dt\, df = \left|\int_{-\infty}^{\infty} z_1(t)\, z_2^*(t)\, dt\right|^2 . \tag{3.1.15}$$

The above properties are not independent; for example, it is a trivial exercise to show that either of the marginals (TM or FM) implies the global energy condition.

A more comprehensive list of the WVD properties is given in [1]. These either follow directly from the definition, or are proven in [2].

A notable omission from the above list is **positivity**; the WVD can assume negative values, and indeed does so for almost every signal.

Applicability: Many properties of the WVD are desirable in applications; for example, realness is consistent with the notion of energy density in as much as energy is real, while convolution invariance and modulation invariance make the WVD partly compatible with linear filtering theory. Nevertheless, the WVD was not applied to real-life problems until the late 1970s, when it was implemented for the purposing of processing linear FM signals used in seismic prospecting [3].

3.1.2 Limitations of the WVD

Despite its many desirable properties, the WVD has some drawbacks. It may assume large negative values. Further, because it is **bilinear** in the signal rather than linear, it suffers from spurious features called **artifacts** or **cross-terms**, which appear midway between true signal components in the case of multicomponent signals as well as non-linear mono- and multicomponent FM signals [see Article 4.2].

3.1.2.1 Nonlinear Monocomponent FM Signals and Inner Artifacts

Consider a monocomponent signal $z(t) = a(t)\, e^{j\phi(t)}$. In the case of a *linear* FM signal, the WVD gives an accurate representation of the IF law (Fig. 1.1.3) because the CFD approximation to $\phi'(t)$ is exact, so that the signal kernel is a dechirped function of τ [see Eqs. (2.1.19) to (2.1.25)]. In the case of a *nonlinear* FM signal, the CFD approximation is *not* exact and the signal kernel (i.e. the IAF) is *not* dechirped. This results in the formation of **inner artifacts**, which arise from "within" a single component.

An example of a nonlinear FM signal is the **hyperbolic FM** signal with rectangular amplitude $a(t)$ and the phase given by

$$\phi(t) = \frac{2\pi f_0}{\alpha} \ln|1+\alpha t|. \tag{3.1.16}$$

The IF is

$$f_i(t) = \frac{\phi'(t)}{2\pi} = \frac{f_0}{1+\alpha t}. \tag{3.1.17}$$

For the time interval $0 \le t \le T$, the starting frequency is $f_0 = f_i(0)$. The finishing frequency is $f_{max} = f_i(T) = f_0/(1+\alpha T)$.

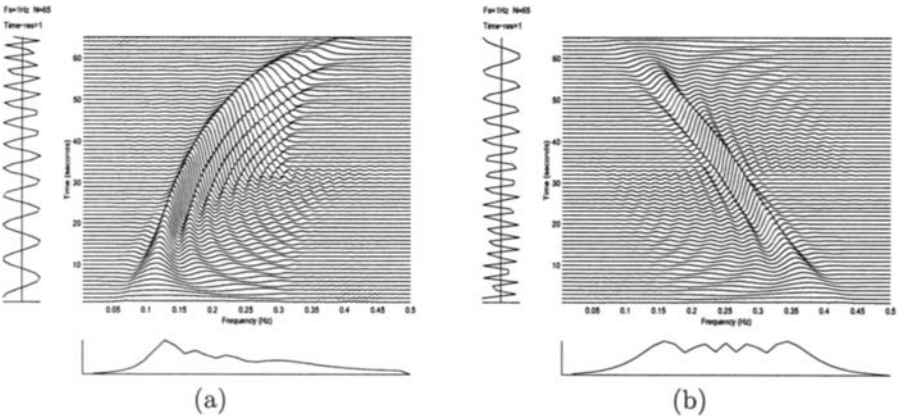


Fig. 3.1.1: WVD of (a) hyperbolic FM signal with starting frequency 0.1 and finishing frequency 0.4 Hz; (b) linear FM signal with starting frequency 0.4 and finishing frequency 0.1 Hz. Both plots are for a duration of 65 samples (sampling rate 1 Hz).

Fig. 3.1.1(a) shows the WVD of a hyperbolic FM signal with duration 65 samples, $f_0 = 0.1$ and $f_{max} = 0.4$. While the crest of the WVD seems to be a reasonable approximation to the IF law, the energy concentration is poorer. The many spurious ridges are the inner artifacts. They alternate in sign as we move *normal* to the IF law in the (t, f) plane; this is a characteristic feature of inner artifacts.

For comparison, Fig. 3.1.1(b) shows the WVD of the linear FM signal with the same duration and the same frequency limits (with falling frequency). Note the superior energy concentration for the linear FM case, and the more attenuated artifacts distributed on both sides of the main ridge.

Artifacts caused by nonlinear FM laws can be reduced by windowing the IAF in τ before taking the FT, leading to the **windowed WVD**. This procedure, however, causes a loss of frequency resolution [4].

3.1.2.2 Multicomponent Signals and Outer Artifacts

If $z(t)$ is a multicomponent signal, the algebraic expansion of $K_z(t,\tau)$ contains cross-product terms which, when Fourier-transformed, give rise to spurious features in the WVD. These are the **outer artifacts** or **cross-terms**. To explain these, consider the signal

$$z(t) = z_1(t) + z_2(t) \tag{3.1.18}$$

where $z(t)$, $z_1(t)$ and $z_2(t)$ are analytic. Expanding the IAF, we obtain

$$K_z(t,\tau) = K_{z_1}(t,\tau) + K_{z_2}(t,\tau) + K_{z_1 z_2}(t,\tau) + K_{z_2 z_1}(t,\tau) \tag{3.1.19}$$

where $K_{z_1 z_2}(t,\tau)$ and $K_{z_2 z_1}(t,\tau)$ are the "signal cross-kernels" or instantaneous cross-correlation functions (e.g. $K_{z_1 z_2}(t,\tau) = z_1(t+\frac{\tau}{2})\, z_2^*(t-\frac{\tau}{2})$). Taking FTs of

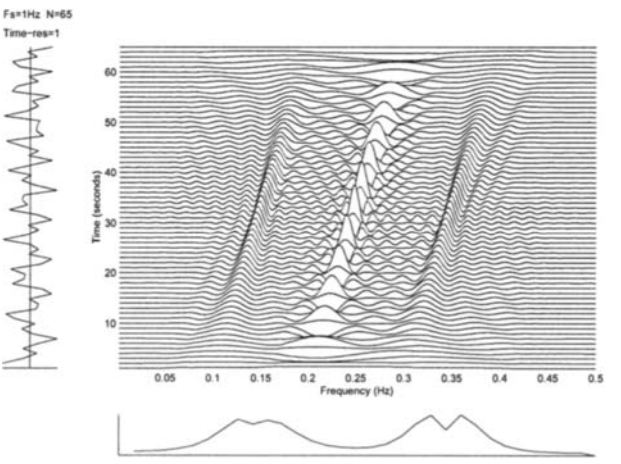


Fig. 3.1.2: The WVD of the sum of two linear FM signals with frequency ranges 0.1–0.2 and 0.3–0.4 Hz, unit amplitudes, and a duration of 65 samples (sampling rate 1 Hz).

Eq. (3.1.19) w.r.t. τ [using Eq. (2.1.18)], we find

$$W_z(t,f) = W_{z_1}(t,f) + W_{z_2}(t,f) + 2\mathrm{Re}\{W_{z_1 z_2}(t,f)\} \tag{3.1.20}$$

where $W_{z_1}(t,f)$ and $W_{z_2}(t,f)$ are the WVDs of $z_1(t)$ and $z_2(t)$, and $W_{z_1 z_2}(t,f)$ is the **cross-Wigner-Ville distribution** (XWVD) of $z_1(t)$ and $z_2(t)$, defined by

$$W_{z_1 z_2}(t,f) = \int_{-\infty}^{\infty} z_1(t+\tfrac{\tau}{2})\, z_2^*(t-\tfrac{\tau}{2})\, e^{-j2\pi f\tau} d\tau. \tag{3.1.21}$$

Thus the WVD of the sum of two signals is not just the sum of the signals' WVDs, but also of their XWVDs. If $z_1(t)$ and $z_2(t)$ are monocomponent signals, $W_{z_1}(t,f)$ and $W_{z_2}(t,f)$ are the auto-terms, while $2\mathrm{Re}\{W_{z_1 z_2}(t,f)\}$ is a cross-term.

Fig. 3.1.2 shows the WVD of the sum of two parallel linear FM signals. There seem to be three components rather than two; the "extra" component at the mean frequency of the expected components has a large oscillating amplitude, and occurs in a region of the (t,f) plane where we expect no energy at all. Fig. 3.1.3(a) shows the WVD of the sum of two FM signals crossing over in frequency. A large number of undulations appear in addition to the two main ridges. In both examples, the cross-terms alternate in sign as we move *parallel* to the expected features in the (t,f) plane; this is a characteristic feature of cross-terms.

3.1.2.3 Suppression of Cross-Terms

Cross-terms can make the WVD difficult to interpret, especially if the components are numerous or close to each other, and the more so in the presence of noise; cross-terms between signal components and noise exaggerate the effects of noise and

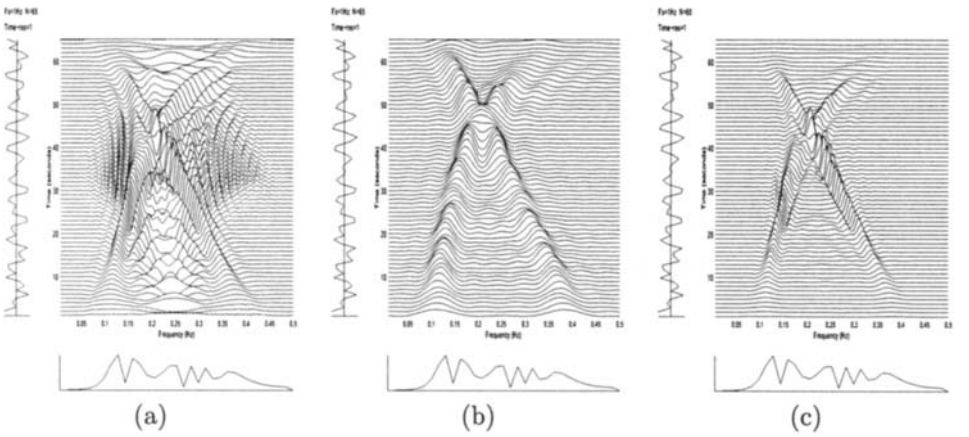


Fig. 3.1.3: Suppression of cross-terms in the sum of a rising hyperbolic FM signal (frequency range 0.1 to 0.4 Hz) and a falling linear FM signal (frequency range 0.4 to 0.1 Hz), with unit amplitudes and a duration of 65 samples (sampling rate 1 Hz): (a) WVD; (b) spectrogram with 21-point rectangular window; (c) masked WVD, being the product of (a) and (b).

cause rapid degradation of performance as the SNR decreases. For such reasons, cross-terms are often regarded as the fundamental limitation on the applicability of quadratic time-frequency methods, and the desire to suppress them has led to several approaches such as:

1. If we multiply the WVD by the spectrogram, we obtain the so-called **masked WVD** [5], which combines the cross-term suppression of the spectrogram with the high resolution of the WVD. To illustrate the effect, Fig. 3.1.3 shows the WVD, spectrogram and masked WVD of the sum of two crossed FM signals; observe that the masked WVD is "cleaner" than the WVD and has better resolution than the spectrogram.

2. Eq. (3.1.21) indicates that the XWVD of $z_1(t)$ and $z_2(t)$ is linear in $z_2(t)$. If $z_1(t)$ is a reference signal and $z_2(t)$ is the signal under analysis, there will be no cross-terms between components of $z_2(t)$. This observation inspired efforts to displace the WVD with the XWVD in relevant areas of application. A reference signal is readily available in optimal detection [6], sonar and radar [7], and seismic exploration [8]. In applications where reference signals are not available, a filtered version of the observed signal is used as a reference signal. The filtering procedure may use the IF as a critical feature of the signal; for example, the data-dependent TFDs defined in [9] use, as a reference signal, the signal component that maximizes the energy concentration in the TFD.

3. All quadratic TFDs in Chapter 2 can be identified by a time-lag kernel, including the spectrogram which suppresses cross-terms. By analyzing the properties of time-lag kernels, we can define and design TFDs that attenuate cross-terms.

A quadratic TFD in which the cross-terms are attenuated relative to the auto-terms is often called a **reduced-interference distribution (RID)**.

This chapter concentrates on the RID approach for defining quadratic TFDs with the property of cross-terms suppression.

3.2 Formulations of Quadratic TFDs

3.2.1 Time-Lag Formulations and Other Domain Definitions

Given an analytic signal $z(t)$, the instantaneous autocorrelation function (IAF) was defined in Eq. (2.1.8) as

$$K_z(t,\tau) = z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2}) \tag{3.2.1}$$

The IAF $K_z(t,\tau)$ is a function of 2 time variables: the actual time t and the time lag τ. By taking the dual domain of t and τ in frequency, we obtain the frequency variables ν and f. This allows for four domains of representation:

$$(t,\tau)$$
$$(t,f)$$
$$(\nu,f)$$
$$(\nu,\tau).$$

Starting with the IAF, we define the Wigner-Ville distribution (WVD) by taking the FT ($\tau \rightarrow f$):

$$W_z(t,f) = \underset{\tau\rightarrow f}{\mathcal{F}}\{K_z(t,\tau)\}. \tag{3.2.2}$$

The FT ($t \rightarrow \nu$) of the WVD defines the spectral autocorrelation function (SAF) [Eq. (2.1.26)] as

$$k_z(\nu,f) = \underset{t\rightarrow \nu}{\mathcal{F}}\{W_z(t,f)\}. \tag{3.2.3}$$

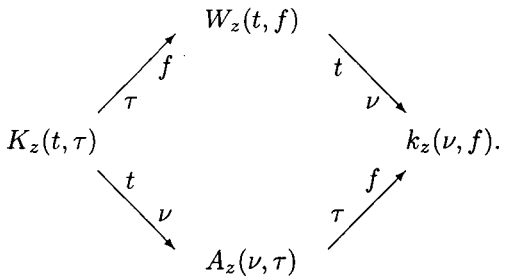
The FT ($t \rightarrow \nu$) of the IAF $K_z(t,\tau)$ equals the IFT ($\tau \leftarrow f$) of the SAF $k_z(\nu,f)$, and defines the symmetrical **ambiguity function** (AF) as:

$$A_z(\nu,\tau) = \underset{t\rightarrow \nu}{\mathcal{F}}\{K_z(t,\tau)\} = \underset{\tau\leftarrow f}{\mathcal{F}^{-1}}\{k_z(\nu,f)\} \tag{3.2.4}$$

where $\mathcal{F}\{\ldots\}$ represents the forward (direct) Fourier transform, and $\mathcal{F}^{-1}\{\ldots\}$ its inverse.[1]

[1]Many authors define the symmetrical ambiguity function using an *inverse* transform from t to ν, with the result that the arrows in the "diamond diagrams" [e.g. Eq. (3.2.5)] point upwards instead of to the right (this affects only the lower-left and upper-right arrows). Here we choose our definitions so that transforms from a time variable (t or τ) to a frequency variable (ν or f) are always *forward*.

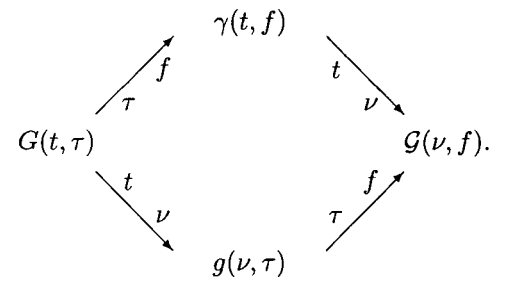
If we represent Fourier transformations by arrows labeled with the participating variables, Eqs. (3.2.2) to (3.2.4) may be combined into the single graphical equation



(3.2.5)

This single graphical representation of several equations is very useful in that it links several known methods by simple FT operations. The knowledge of FT properties then allows to use characteristics of a method in one domain and transfer them to another domain. We can so relate radar methods in the Doppler-lag domain to (t,f) methods and to cyclostationary methods.

Now let us represent similarly the TFD kernel in these four domains by taking various FTs of the **time-lag kernel** $G(t,\tau)$:



(3.2.6)

We refer to $g(\nu,\tau)$ as the **Doppler-lag kernel**, to $\mathcal{G}(\nu,f)$ as the **Doppler-frequency kernel**, and to $\gamma(t,f)$ as the **time-frequency kernel**.

3.2.2 Time-Frequency Formulation

We have defined the smoothed IAF as

$$R_z(t,\tau) = G(t,\tau) \underset{t}{*} K_z(t,\tau) \tag{3.2.7}$$

where $G(t,\tau)$ is the time-lag kernel. Following Eq. (2.7.1), we define a class of **quadratic TFDs** (which are smoothed WVDs) as

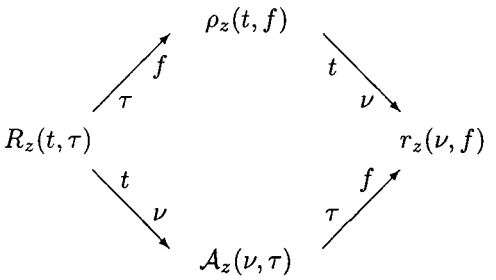
$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\{R_z(t,\tau)\}. \tag{3.2.8}$$

Substituting Eq. (3.2.7) into Eq. (3.2.8), writing out the convolution and the transform, and substituting from Eq. (3.2.1), we obtain

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(t-u,\tau)\, z(u+\tfrac{\tau}{2})\, z^*(u-\tfrac{\tau}{2})\, e^{-j2\pi f\tau}\, du\, d\tau \tag{3.2.9}$$

which defines the general class of quadratic TFDs in terms of the signal and the time-lag kernel. The class of TFDs of this form is called the **quadratic class.**[2]

Eq. (3.2.8) is included in the graphical equation



(3.2.10)

which assigns symbols to the various FTs of the smoothed IAF [11, p. 436]. By analogy with Eqs. (3.2.3) and (3.2.4), we call $r_z(\nu,f)$ the "generalized SAF" and $\mathcal{A}_z(\nu,\tau)$ the "generalized ambiguity function" (GAF)[3] or more precisely **filtered ambiguity function**; note the distinction between the normal A in Eq. (3.2.5) and the calligraphic $\mathcal{A}$ in Eq. (3.2.10).

Using Eq. (3.2.7) and the lower two arrows in Eq. (3.2.10), together with the convolution properties of the FT, we obtain in sequence

$$\mathcal{A}_z(\nu,\tau) = g(\nu,\tau)\, A_z(\nu,\tau) \tag{3.2.11}$$

$$r_z(\nu,f) = \mathcal{G}(\nu,f) \underset{f}{*} k_z(\nu,f) \tag{3.2.12}$$

Then, using Eq. (3.2.7) and the upper left arrow in Eq. (3.2.10), together with the convolution properties, we obtain

$$\rho_z(t,f) = \gamma(t,f) \underset{(t,f)}{**} W_z(t,f) \tag{3.2.13}$$

where the double asterisk denotes double convolution in t and f (cf. [11], pp. 437, 475).

Eq. (3.2.13) defines the quadratic class of TFDs in terms of the WVD and the time-frequency kernel. For this reason we regard the WVD as the "basic" or "prototype" quadratic TFD, and all other quadratic TFDs as filtered versions thereof.

---

[2]The **quadratic class** as defined here satisfies the time-shift- and frequency-shift-invariance properties. Most current authors seem to equate "**Cohen's class**" with the quadratic class. In addition, some authors apply the term "quadratic time-frequency representation" or "QTFR" to all time-frequency representations that are quadratic in the signal (whether they have the form of Eq. (3.2.9) or not) so that "Cohen's class" becomes a subset of the "QTFRs". However, the class originally defined by Cohen in a quantum-mechanical context [10] differs from the quadratic class [Eq. (3.2.9)] in that the marginal conditions must be satisfied, while the kernel may depend on the signal (in which case the TFD is *not* quadratic in the signal). We shall see (e.g. in Table 3.3.3) that many useful quadratic TFDs do not satisfy the marginals, so that they are not members of "Cohen's class" as originally defined.

[3]The term **generalized ambiguity function** is also used with a different meaning in connection with polynomial TFDs; see Article 5.5.

Eq. (3.2.13) also suggests a method of suppressing cross-terms in the WVD. Because inner artifacts alternate as we move in the frequency direction, they can be attenuated by choosing a $\gamma(t,f)$ with a sufficient spread in the frequency direction. Similarly, because cross-terms (outer artifacts) alternate as we move in the time direction, they can be attenuated by choosing a $\gamma(t,f)$ with a sufficient spread in the time direction.

Eqs. (3.2.10) and (3.2.11) show how one quadratic TFD can be transformed to another by a linear transformation.

3.2.3 Doppler-Lag Formulation and TFD Design

Substituting for $K_z(t,\tau)$ in Eq. (3.2.4) and writing out the transform, we obtain

$$A_z(\nu,\tau) = \int_{-\infty}^{\infty} z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\, e^{-j2\pi\nu t} dt \tag{3.2.14}$$

$$= \int_{-\infty}^{\infty} z(t+\tfrac{\tau}{2}) \left[z(t-\tfrac{\tau}{2})\, e^{j2\pi\nu t}\right]^* dt. \tag{3.2.15}$$

The expression in square brackets can be obtained by delaying $z(t+\frac{\tau}{2})$ in time by τ and shifting it in frequency by ν, indicating that $A_z(\nu,\tau)$ is the correlation of the signal with a time-delayed and frequency-shifted version of itself. This correlation is well known in radar theory as the Sussman ambiguity function [12]; the name "ambiguity" arises from the equivalence between time-shifting and frequency-shifting for linear FM signals, which are frequently used in radar. Hence the Doppler-lag (ν,τ) domain is also called the **ambiguity domain.**

Eq. (3.2.10) indicates that the quadratic TFD is a 2D FT (half inverse, half forward) of the filtered ambiguity function. Writing out the transforms gives

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\nu,\tau)\, A_z(\nu,\tau)\, e^{j2\pi(\nu t - f\tau)} d\nu\, d\tau. \tag{3.2.16}$$

Then writing the dummy u for t in Eq. (3.2.14) and substituting the result into Eq. (3.2.16) gives

$$\rho_z(t,f) = \iiint g(\nu,\tau)\, z(u+\tfrac{\tau}{2})\, z^*(u-\tfrac{\tau}{2})\, e^{j2\pi(\nu t - \nu u - f\tau)} du\, d\nu\, d\tau \tag{3.2.17}$$

where the integrals are from $-\infty$ to ∞; this defines the quadratic TFD in terms of the Doppler-lag kernel $g(\nu,\tau)$.

From the lower left side of Eq. (3.2.6), the relationship between the time-lag kernel and the Doppler-lag kernel is

$$g(\nu,\tau) = \underset{t\to\nu}{\mathcal{F}}\{G(t,\tau)\}. \tag{3.2.18}$$

Using this equation, the time-lag kernels determined in Section 2.7 may be converted to Doppler-lag form. For example, using Eq. (2.7.7), we find that the Doppler-lag kernel for the WVD is $g(\nu,\tau) = \underset{t\to\nu}{\mathcal{F}}\{\delta(t)\} = 1$.

Eq. (3.2.13) and its 2D FT Eq. (3.2.16) express that a quadratic TFD can be designed using basic filter design principles. As in 1D filter design, the filter specifications are best given in the domain where the filtering operation is expressed as a multiplication as in Eq. (3.2.11), rather than in the dual domain with a convolution as in Eq. (3.2.13).

If the Doppler-lag kernel $g(\nu,\tau)$ has the separable form $G_1(\nu)\,g_2(\tau)$, multiplication by this kernel may include the combined effect of time windowing and frequency windowing (see the discussion of Eq. (2.1.40)).

We can use the same process to define a filter which attenuates cross-terms in quadratic TFDs. The cross-terms in the (t,f) domain tend to be highly oscillatory, so that the corresponding terms in the dual (ν,τ) domain tend to be far from the origin (high-pass). The auto-terms in the (t,f) domain tend to be smooth and well delineated, so that the corresponding terms in the dual (ν,τ) domain are concentrated about the origin or "pass through" the origin (low-pass). This behavior is well known in the field of radar [13,14]. Hence the cross-terms in the ambiguity domain can be "filtered out" by selecting a kernel filter[4] $g(\nu,\tau)$ that deemphasizes information far from the origin in the Doppler-lag domain.

Various authors [10,15,16] have shown that other desirable TFD properties are equivalent to constraints on the kernel filter, and we shall see that most of these constraints are conveniently expressed in the Doppler-lag domain. For all the above reasons, *the design of the TFD kernel filter is usually performed in the Doppler-lag domain.* Often, the resulting kernel is then described in the time-lag domain for ease of TFD implementation, as given by Eq. (3.2.9).

3.2.4 Doppler-Frequency Formulation

From Eqs. (3.2.10) and (3.2.12) we have

$$\rho_z(t,f) = \underset{t\leftarrow\nu}{\mathcal{F}^{-1}} \left\{ \mathcal{G}(\nu,f) \underset{f}{*} k_z(\nu,f) \right\} \tag{3.2.19}$$

or, writing out the transform and convolution,

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{G}(\nu,f-\eta)\,k_z(\nu,\eta)\,e^{j2\pi\nu t}d\eta\,d\nu \tag{3.2.20}$$

Writing η for f in Eq. (2.1.31) and substituting the result into Eq. (3.2.20), we obtain

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{G}(\nu,f-\eta)\,Z(\eta+\tfrac{\nu}{2})\,Z^*(\eta-\tfrac{\nu}{2})\,e^{j2\pi\nu t}d\eta\,d\nu. \tag{3.2.21}$$

This defines quadratic TFDs in terms of the Doppler-frequency kernel $\mathcal{G}(\nu,f)$ and the signal spectrum. The interest of this formulation is that TFDs of narrow-band signals expressed by their spectra may be more efficiently computed in this form.

[4]The term "kernel filter" is used to reinforce the idea that designing quadratic TFDs essentially reduces to filter design with specific constraints.

3.2.5 Examples of Simple TFD Formulations

3.2.5.1 Separable Kernels

A simple way to design kernel filters for quadratic TFDs is to consider the special case of a separable kernel (see Article 5.7):

$$g(\nu,\tau) = G_1(\nu)\, g_2(\tau). \tag{3.2.22}$$

If we let

$$G_1(\nu) = \underset{t\to\nu}{\mathcal{F}}\{g_1(t)\} \tag{3.2.23}$$

and

$$G_2(f) = \underset{\tau\to f}{\mathcal{F}}\{g_2(\tau)\}, \tag{3.2.24}$$

then Eq. (3.2.11) becomes

$$\mathcal{A}_z(\nu,\tau) = G_1(\nu)\, g_2(\tau)\, A_z(\nu,\tau). \tag{3.2.25}$$

Then, using Eqs.(3.2.10), (3.2.23), and(3.2.24), we obtain

$$\rho_z(t,f) = g_1(t) \underset{t}{*} W_z(t,f) \underset{f}{*} G_2(f). \tag{3.2.26}$$

Eq. (3.2.25) shows that the design of the kernel filter $G_1(\nu)\, g_2(\tau)$ is greatly simplified as the 2D filtering operation is replaced by two successive 1D filtering operations. Equivalently, in Eq. (3.2.26), the two convolutions can be evaluated in either order, indicating that the Doppler-dependent and lag-dependent factors in the separable kernel $g(\nu,\tau)$ lead to separate convolutions in time and frequency, respectively.

A separable kernel is not to be confused with a **product kernel**, which is a function of the product $\nu\tau$, e.g.

$$g(\nu,\tau) = g_3(\nu\tau). \tag{3.2.27}$$

3.2.5.2 Doppler-Independent Kernels

A **Doppler-independent** (DI) kernel is a special case of a separable kernel obtained by putting

$$G_1(\nu) = 1 \tag{3.2.28}$$

in Eqs. (3.2.22) and (3.2.23), which then become

$$g(\nu,\tau) = g_2(\tau) \tag{3.2.29}$$

$$g_1(t) = \delta(t). \tag{3.2.30}$$

Making these substitutions in Eqs. (3.2.25) and (3.2.26), we obtain

$$\mathcal{A}_z(\nu,\tau) = g_2(\tau)\, A_z(\nu,\tau) \tag{3.2.31}$$

$$\rho_z(t,f) = G_2(f) \underset{f}{*} W_z(t,f). \tag{3.2.32}$$

The last result shows that a DI kernel is obtained by applying only 1D filtering and causes smearing of the WVD in the frequency direction *only*.

3.2.5.3 Lag-Independent Kernels

A **lag-independent** (LI) kernel is another special case of the separable kernel, obtained by putting

$$g_2(\tau) = 1 \tag{3.2.33}$$

in Eqs. (3.2.22) and (3.2.24), which then become

$$g(\nu, \tau) = G_1(\nu) \tag{3.2.34}$$

$$G_2(f) = \delta(f). \tag{3.2.35}$$

Making these substitutions in Eqs. (3.2.25) and (3.2.26), we obtain

$$\mathcal{A}_z(\nu, \tau) = G_1(\nu)\, A_z(\nu, \tau) \tag{3.2.36}$$

$$\rho_z(t, f) = g_1(t) \underset{t}{*} W_z(t, f). \tag{3.2.37}$$

The last result shows that an LI kernel is obtained by applying only 1D filtering and causes smearing of the WVD in the time direction *only*.

The kernel of the WVD is $g(\nu, \tau) = 1$, which is both Doppler-independent and lag-independent. The windowed WVD kernel, however, is Doppler-independent.

Article 5.7 provides an in-depth treatment of such separable kernels, including examples of quadratic TFDs obtained by this simple kernel filter design procedure.

3.3 Properties of Quadratic TFDs

3.3.1 Desirable Properties

Some quadratic TFDs verify desirable properties that are not shared by the WVD, and vice versa. Later we shall relate the properties of a TFD to the constraints of its kernel, and tabulate properties for selected TFDs. But first we discuss some properties that have been promoted [17–19] as fundamental for a wide range of applications.

1. **Concentration of local energy:** The energy in a certain region R in the (t, f) plane, denoted by E_{z_R}, is given by the integral of the TFD over the region R; for example, if R is the region within a time interval Δt and a frequency band Δf, the energy within R is

$$E_{z_R} = \int_{\Delta t} \int_{\Delta f} \rho_z(t, f)\, df\, dt. \tag{3.3.1}$$

2. **IF/TD visualization:** The TFD of a monocomponent signal directly depicts the instantaneous frequency $f_i(t)$ and time delay $\tau_d(f)$ as a range of peaks along the curve representing the FM law in the (t, f) plane. That is, if $z(t)$ is a monocomponent signal and $\rho_z(t, f)$ is its TFD, then

$$\max_f \rho_z(t, f) = f_i(t). \tag{3.3.2}$$

3. **Reduced interference (RI):** The TFD attenuates or suppresses inner and outer artifacts (cross-terms) relative to the signal components (auto-terms).

Another condition considered by some authors is **positivity**, or more precisely **non-negativity (NN)**, defined as:

$$\rho_z(t,f) \geq 0 \quad \forall \quad z,t,f. \tag{3.3.3}$$

Among quadratic TFDs, only *sums of spectrograms* possess NN. This means that the Doppler-lag kernel of a "positive" TFD is a sum of ambiguity functions, which makes the NN property incompatible with both the IF property (see Table 3.3.1) and IF visualization [20]. Hence NN is usually considered non-essential because its cost is excessive.

A subclass $\mathcal{P}$ of quadratic TFDs comprises those TFDs which satisfy the realness, time marginal, frequency marginal, instantaneous frequency, time support and frequency support properties. Researchers have shown much interest in TFDs of this class, and especially RIDs of this class. To some extent the design of RIDs is the art of "improving" on the spectrogram by

- sacrificing NN,
- improving resolution, and
- retaining sufficient reduced-interference ability for the application.

The result is usually a compromise between the spectrogram and the WVD, involving a time-frequency kernel filter less extensive than that of the spectrogram. The compromise may involve sacrificing one or both of the marginal properties (time or frequency). Although the marginals are critical in the field in quantum physics [10], they seem to be less important in signal processing. As evidence of this we may cite the following:

- The spectrogram does not satisfy the marginals and yet has always been regarded as a very useful TFD.
- Attempts to improve on the spectrogram are most often motivated by the need for higher resolution (see e.g. [21]) rather than any desire to satisfy the marginals.
- If the marginal conditions do not hold, it is still possible that the integral forms thereof [Eqs. (2.6.6) and (2.6.7)] are approximately true over sufficiently broad time intervals or frequency bands.

For these reasons it is suggested that $\mathcal{P}$ is not necessarily appropriate for signal-processing applications, and an alternative class $\mathcal{P}'$, which comprises reduced-interference distributions satisfying the realness, global and local energy, IF visualization and components resolution properties, would be more relevant. Note that the properties of class $\mathcal{P}'$ are the same as the properties listed at the end of Section 1.1.5.

3.3.2 TFD Properties & Equivalent Kernel Constraints

Using some properties listed in [22], the relevant kernel constraints were adapted for separable, DI and LI kernels. The results are collected in Table 3.3.1.

The proofs of these kernel constraints use Eq. (3.2.13), which states that the quadratic TFD is the WVD convolved in t and f with the time-frequency kernel $\gamma(t,f)$:

$$\rho_z(t,f) = \gamma(t,f) \underset{(t,f)}{**} W_z(t,f).$$

Using this relation, we can find *sufficient* conditions under which certain properties of the WVD carry over to $\rho_z(t,f)$. For example:

- Realness holds if $\gamma(t,f)$ is real.
- Because 2D convolution is shift-invariant, time- and frequency-shift invariance hold for any fixed $\gamma(t,f)$.
- Time support and time extent hold for a DI kernel, which does not redistribute the WVD in time, whereas frequency support and frequency extent hold for an LI kernel, which does not redistribute the WVD in frequency.
- The IF moment property holds for a DI kernel if $G_2(f)$ has a first moment (mean frequency) of zero, so that convolution w.r.t. f does not change the first moment of the WVD w.r.t. f. Similarly, the GD property holds for an LI kernel if $g_1(t)$ has a first moment (mean time) of zero, so that convolution w.r.t. t does not change the first moment of the WVD w.r.t. t.

Some further proofs of kernel constraints are given by Cohen [23].

3.3.3 Examples of TFDs with Specific Properties

For the TFDs defined so far, Table 3.3.2 lists their kernels in various domains, and Table 3.3.3 shows their properties.

Note: A time-lag kernel satisfying the time-support constraint is described as a **butterfly function** [11] or **cone-shaped kernel** [24]; that is, the nonzero values of the kernel are confined to the interior of a two-dimensional cone in the (t,τ) plane (see the entry in the "General" column of Table 3.3.1). The Born-Jordan and ZAM distributions have kernels of this type.

Inspection of Table 3.3.1 confirms that the WVD satisfies all of the listed properties except reduced interference (RI). The reduced-interference capabilities of separable kernels warrant special attention. Because the inner artifacts alternate as we move normal to the components, they also alternate in the frequency direction and can be suppressed by convolution with a sufficiently long $G_2(f)$ [see Eq. (3.2.32)], which corresponds to a sufficiently short $g_2(\tau)$. This is possible for a DI kernel but not an LI kernel. Because the cross-terms (outer artifacts) alternate as we move parallel to the components, they also alternate in the time direction and can be suppressed by convolution with a sufficiently long $g_1(t)$ [see Eq. (3.2.37)], which corresponds to a sufficiently short $G_1(\nu)$. This is possible for an LI kernel but not a

Table 3.3.1: TFD properties and associated constraints on the Doppler-lag kernel for general, separable, Doppler-independent and lag-independent kernels. The asterisk in "WVD only*" means that the WVD may be multiplied by a scale factor. **Abbreviations for properties:** NN = non-negativity; RE = realness; TI = time-shift invariance (or time covariance); FI = frequency-shift invariance (or frequency covariance); TM = time marginal; FM = frequency marginal; IF = instantaneous frequency; TD = time delay; TS = time support; FS = frequency support; RI = RID capability.

Property	KERNEL CONSTRAINTS			
	General —	Separable $g(\nu,\tau) = G_1(\nu)\, g_2(\tau)$	DI $g(\nu,\tau) = g_2(\tau)$	LI $g(\nu,\tau) = G_1(\nu)$
NN:	$G(t,\tau)$ is a sum of IAFs.	$g_1(t)\, g_2(\tau)$ is a sum of IAFs.	Never	Never
RE:	$g(\nu,\tau) = g^*(-\nu,-\tau)$.	$G_1(\nu)\, g_2(\tau) = G_1^*(-\nu)\, g_2^*(-\tau)$.	$G_2(f)$ is real.	$g_1(t)$ is real.
TI:	$g(\nu,\tau)$ does not depend on t.	$G_1(\nu)\, g_2(\tau)$ does not depend on t.	$g_2(\tau)$ does not depend on t.	$G_1(\nu)$ does not depend on t.
FI:	$g(\nu,\tau)$ does not depend on f.	$G_1(\nu)\, g_2(\tau)$ does not depend on f.	$g_2(\tau)$ does not depend on f.	$G_1(\nu)$ does not depend on f.
TM:	$g(\nu,0) = 1 \ \ \forall\nu$	$G_1(\nu)\, g_2(0) = 1 \ \ \forall\nu$	$g_2(0) = 1$	WVD only
FM:	$g(0,\tau) = 1 \ \ \forall\tau$	$G_1(0)\, g_2(\tau) = 1 \ \ \forall\tau$	WVD only	$G_1(0) = 1$
IF:	$g(\nu,0) = \text{const.}$ $\frac{\partial g}{\partial \tau}\big\vert_{\tau=0} = 0 \ \ \forall\nu.$	$G_1(\nu)\, g_2(0) = \text{const.}$ $g_2'(0) = 0.$	$g_2'(0) = 0$	WVD only*
TD:	$g(0,\tau) = \text{const.}$ $\frac{\partial g}{\partial \nu}\big\vert_{\nu=0} = 0 \ \ \forall\tau.$	$G_1(0)\, g_2(\tau) = \text{const.}$ $G_1'(0) = 0.$	WVD only*	$G_1'(0) = 0$
TS:	$G(t,\tau) = 0$ if $\vert\tau\vert < 2\vert t\vert$.	DI only	Always	WVD only*
FS:	$\mathcal{G}(f,\nu) = 0$ if $\vert\nu\vert < 2\vert f\vert$.	LI only	WVD only*	Always
RI:	Unrestricted	Unrestricted	Inner x-terms	Outer x-terms

DI kernel. A general separable kernel with sufficiently short $G_1(\nu)$ and $g_2(\tau)$ can therefore attenuate both kinds of artifacts; a TFD with such a kernel is shown in Fig. 2.7.1 part (l) for a linear FM signal.

Properties of separable kernels and their special cases, such as the B-distribution and Modified B-distribution, are discussed further and illustrated on examples in Article 5.7. Design of RIDs is discussed more generally in Article 5.2.

Table 3.3.2: Kernels of selected TFDs in the time-lag, Doppler-lag, and (where possible) Doppler-frequency domains. The window $w(t)$ is assumed to be real and even. Its FT and AF are $W(f)$ and $A_w(\nu,\tau)$, respectively. The prefix "w-" means "windowed".

	KERNEL		
Distribution	$G(t,\tau)$	$g(\nu,\tau)$	$\mathcal{G}(\nu,f)$
Wigner-Ville	$\delta(t)$	1	$\delta(f)$
Levin	$\frac{1}{2}\left[\delta(t+\frac{\tau}{2})+\delta(t-\frac{\tau}{2})\right]$	$\cos(\pi\nu\tau)$	$\frac{1}{2}\left[\delta(f+\frac{\nu}{2})+\delta(f-\frac{\nu}{2})\right]$
Born-Jordan	$\frac{1}{\lvert 2\alpha\tau\rvert}\,\mathrm{rect}\,\frac{t}{2\alpha\tau}$	$\mathrm{sinc}\,(2\alpha\nu\tau)$	$\frac{1}{\lvert 2\alpha\nu\rvert}\,\mathrm{rect}\,\frac{f}{2\alpha\nu}$
Modified B	$\frac{\cosh^{-2\beta}t}{\int_{-\infty}^{\infty}\cosh^{-2\beta}\xi\,d\xi}$	$\frac{\lvert\Gamma(\beta+j\pi\nu)\rvert^2}{\Gamma^2(\beta)}$	$\frac{\lvert\Gamma(\beta+j\pi\nu)\rvert^2}{\Gamma^2(\beta)}\,\delta(f)$
w-WVD	$\delta(t)\,w(\tau)$	$w(\tau)$	$W(f)$
w-Levin	$\frac{w(\tau)}{2}\left[\delta(t+\frac{\tau}{2})+\delta(t-\frac{\tau}{2})\right]$	$w(\tau)\cos(\pi\nu\tau)$	$\frac{1}{2}\left[W(f+\frac{\nu}{2})+W(f-\frac{\nu}{2})\right]$
ZAM	$w(\tau)\,\mathrm{rect}\,\frac{t}{2\tau/a}$	$w(\tau)\frac{a}{2\lvert\tau\rvert}\,\mathrm{sinc}\,\frac{2\nu\tau}{a}$	
Rihaczek	$\delta(t-\frac{\tau}{2})$	$e^{-j\pi\nu\tau}$	$\delta(f+\frac{\nu}{2})$
w-Rihaczek	$w(\tau)\,\delta(t-\frac{\tau}{2})$	$w(\tau)\,e^{-j\pi\nu\tau}$	$W(f+\frac{\nu}{2})$
Page	$\delta(t-\lvert\frac{\tau}{2}\rvert)$	$e^{-j\pi\nu\lvert\tau\rvert}$	$\frac{1}{2}\left[\delta(f+\frac{\nu}{2})+\delta(f-\frac{\nu}{2})\right]$ $+\,j\nu/[2\pi(f^2-\nu^2/4)]$
Choi-Williams	$\frac{\sqrt{\pi\sigma}}{\lvert\tau\rvert}e^{-\pi^2\sigma t^2/\tau^2}$	$e^{-\nu^2\tau^2/\sigma}$	$\frac{\sqrt{\pi\sigma}}{\lvert\nu\rvert}e^{-\pi^2\sigma f^2/\nu^2}$
B	$\lvert\tau\rvert^{\beta}\cosh^{-2\beta}t$	$\frac{\lvert\tau\rvert^{\beta}\,\lvert\Gamma(\beta+j\pi\nu)\rvert^2}{2^{1-2\beta}\,\Gamma(2\beta)}$	
Spectrogram	$w(t+\frac{\tau}{2})\,w(t-\frac{\tau}{2})$	$A_w(\nu,\tau)$	$W(f+\frac{\nu}{2})\,W(f-\frac{\nu}{2})$

3.4 Summary, Discussion and Conclusions

This chapter ends the tutorial introduction to TFSAP constituted by Part I of this book. In essence, the main message and findings of this chapter are as follows. For a monocomponent linear FM signal, the WVD is optimal for energy concentration about the IF and for unbiased estimation of the IF. If a signal has nonlinear frequency modulation and/or multiple components, the WVD suffers from inner artifacts and/or outer artifacts (cross-terms), respectively; in either case, some form of reduced interference quadratic TFD (RID) is to be preferred over the WVD. The design of RIDs is best undertaken by designing the desired kernel filter in the ambiguity domain, and using Fourier transforms to see the effects in the time-lag and time-frequency domains. To be a useful tool for practical applications, quadratic TFDs are expected to be real, to satisfy the global and local energy requirements, and to resolve signal components while reflecting the components' IF laws through the peaks of their dominant ridges in the (t,f) plane. Several RIDs were designed

Table 3.3.3: Properties of the TFDs whose kernels are listed in Table 3.3.2. The window $w(t)$ is assumed to be real and even. An exclamation (!) means that the property is always *satisfied*, while an asterisk (*) means that the property is satisfied subject to normalization of the window (for TM) or the value of the parameter (for TS and FS). Comments on the kernel type are in parentheses (thus).

	PROPERTY						
Distribution	RE	TM	FM	IF	TD	TS	FS
Wigner-Ville	!	!	!	!	!	!	!
Levin (product kernel)	!	!	!	!	!	!	!
Born-Jordan (product kernel)	!	!	!	!	!	*	*
Modified B (LI kernel)	!		!		!		!
w-WVD (DI kernel)	!	*		!		!	
w-Levin	!	*		!		!	
ZAM	!	*		!		*	
Rihaczek (product kernel)		!	!			!	!
w-Rihaczek		*				!	
Page	!	!	!	!		!	
Choi-Williams (product kernel)	!	!	!	!	!		
B (separable kernel)	!						
Spectrogram	!						

using simple separable kernels, demonstrating the procedure to construct quadratic TFDs that meet the above requirements. Special-purpose quadratic TFDs can be easily designed to meet the specifications of particular applications (as we shall see in several articles in the following chapters). In general, however, the WVD and various RIDs are the most useful TFDs; the spectrogram, which has been widely used, is at best subsumed by quadratic RIDs (of which it is a special case), and at worst made obsolete by them.

The remaining four Parts of the book elaborate on these issues, discuss advanced design methods for TFDs, and present a wide selection of methodologies, algorithms, and applications that demonstrate how to use time-frequency methods in practice. In particular, questions and issues such as how to select a TFD for a particular application, how to implement it and how best to apply it are covered. These chapters constituting the remaining four Parts include a number of Articles which tend to cover more advanced and detailed material, complementing and supplementing the tutorial introduction of Chapter 1.

Part II of the book gives more details on some fundamental topics of TFSAP

such as TFD design and signal analysis in the (t, f) plane. It includes two chapters (4 and 5). Chapter 4 presents some advanced concepts for TF signal analysis, TF signal processing and TF system analysis. Chapter 5 presents a number of methods for designing TFDs, including a treatment of the ambiguity function.

Part III of the book describes specialized techniques used in implementation, measurement and enhancement of TFDs. It includes three chapters (6, 7 and 8). Chapter 6 deals with the implementation and realization of TFDs; in particular, the formulation of discrete-time quadratic TFDs is presented and computation issues are described. Chapter 7 presents quality measures for TFDs and methods for performance enhancement. Chapter 8 describes methods and algorithms for multi-sensor and time-space processing used in applications such as sonar and telecommunications.

Part IV presents the key statistical techniques for TFSAP of random signals. It includes four chapters (9 to 12). Chapter 9 presents time-frequency methods for random processes and noise analysis; Chapter 10 describes methods for instantaneous frequency estimation; Chapter 11 deals with the field of time-frequency synthesis and filtering, including time-varying filter design; and Chapter 12 presents time-frequency methods for signal detection, classification and estimation.

Part V describes a representative selection of TFSAP applications encompassing a wide range of fields and industries. It includes three chapters (13 to 15). Chapter 13 presents time-frequency applications in telecommunications; Chapter 14 describes time-frequency methods in radar, sonar and acoustics; and chapter 15 details a number of time-frequency methods for diagnosis and monitoring used in a wide range of diverse applications.

A detailed description of the contents of each chapter can be found in the table of contents of the book and in the introduction to each chapter. For a more detailed search of a topic needed, e.g. for the explanation of an advanced concept, the reader is referred to the detailed index at the end of the book.

For various reasons such as space limitations and publication delays, a number of recent references could not be included in the relevant chapters. They are sufficiently important to be listed here for the sake of completeness. They include new theoretical developments on a range of topics discussed elsewhere in this book such as TFDs with complex argument [25], rotated (t, f) kernel filters [26], Gabor analysis [27], detection [28], spectrogram segmentation [29], time-frequency plane decomposition in atoms [30], the issue of positivity [31, 32], IF estimation [33, 34], feature extraction [35], an illustration of the concept of BT product [36], polyspectra [37] and Volterra series [38].

References

[1] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner Distribution—A tool for time-frequency signal analysis; Part 1: Continuous-time signals," *Philips J. of Research*, vol. 35, no. 3, pp. 217–250, 1980.

[2] N. G. de Bruijn, "A theory of generalized functions, with applications to Wigner distri-

bution and Weyl correspondence," *Nieuw Archief voor Wiskunde (3)*, vol. 21, pp. 205–280, 1973.

[3] B. Bouachache, "Representation temps-frequence," Tech. Rep. 373/78, Soc. Nat. ELF-Aquitaine, Pau, France, 1978. 56 pp.

[4] B. Ristic, *Some aspects of signal dependent and higher-order time-frequency and time-scale analysis of non-stationary signals.* PhD thesis, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Australia, 1995.

[5] B. Barkat and B. Boashash, "Higher order PWVD and Legendre based time-frequency distribution," in *Proc. Sixth IEEE Internat. Workshop on Intelligent Signal Processing and Communication Systems (ISPACS'98)*, vol. 2, pp. 532–536, Melbourne, Australia, 5–6 November 1998.

[6] B. V. K. Vijaya Kumar and C. W. Carroll, "Performance of Wigner distribution function based detection methods," *Optical Engineering*, vol. 23, pp. 732–737, November-December 1984.

[7] H. H. Szu, "Two-dimensional optical processing of one-dimensional acoustic data," *Optical Engineering*, vol. 21, pp. 804–813, September–October 1982.

[8] P. J. Boles and B. Boashash, "Application of the cross-Wigner-Ville distribution to seismic data processing," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 20, pp. 445–466, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[9] D. L. Jones and T. W. Parks, "A high-resolution data-adaptive time-frequency representation," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 2127–2135, December 1990.

[10] L. Cohen, "Generalized phase-space distribution functions," *J. of Mathematical Physics*, vol. 7, pp. 781–786, May 1966.

[11] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[12] S. M. Sussman, "Least-squares synthesis of radar ambiguity functions," *IRE Trans. Information Theory*, vol. 8, pp. 246–254, April 1962.

[13] G. F. Boudreaux-Bartels, *Time-frequency signal processing algorithms: Analysis and synthesis using Wigner distribution.* PhD thesis, Rice University, 1983.

[14] P. Flandrin, "Some features of time-frequency representations of multicomponent signals," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'84)*, vol. 3, pp. 41B.4.1–41B.4.4, San Diego, 19–21 March 1984.

[15] L. Cohen, "Introduction: A primer on time-frequency analysis," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 1, pp. 3–42, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[16] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, pp. 402–412, February 1992.

[17] B. Boashash, "Wigner analysis of time-varying signals—Its application in seismic prospecting," in *Proc. European Signal Processing Conf. (EUSIPCO-83)*, pp. 703–706, Nürnberg, September 1983.

[18] B. Boashash and B. Escudie, "Wigner-Ville analysis of asymptotic signals and applications," *Signal Processing*, vol. 8, pp. 315–327, June 1985.

[19] B. Bouachache, *Representation temps-frequence.* Thesis for diplome de docteur ingenieur, Institut National Polytechnique de Grenoble, France, May 1982. 164 pp. In French.

[20] B. Boashash, L. B. White, and J. Imberger, "Wigner-Ville analysis of non-stationary random signals (with application to turbulent microstructure signals)," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'86)*, vol. 4, pp. 2323–2326, Tokyo, 7–11 April 1986.

[21] E. F. Velez and H. Garudadri, "Speech analysis based on smoothed Wigner-Ville distribution," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 15, pp. 351–374, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[22] W. J. Williams and J. Jeong, "Reduced interference time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 3, pp. 74–97, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[23] L. Cohen, "Time-frequency distributions—A review," *Proc. IEEE*, vol. 77, pp. 941–981, July 1989. Invited paper.

[24] Y. Zhao, L. E. Atlas, and R. J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[25] L. Stanković, "Time-frequency distributions with complex argument," *IEEE Trans. Signal Processing*, vol. 50, pp. 475–486, March 2002.

[26] M. J. Bastiaans, T. Alieva, and L. Stanković, "On rotated time-frequency kernels," *IEEE Signal Processing Letters*, vol. 9, pp. 378–381, November 2002.

[27] H. G. Feichtinger and T. Strohmer, *Advances in Gabor Analysis.* Birkhaüser, 2002.

[28] C. Richard, "Time-frequency-based detection using discrete-time discrete-frequency Wigner distributions," *IEEE Trans. Signal Processing*, vol. 50, pp. 2170–2176, September 2002.

[29] C. Hory, N. Martin, and A. Chehikian, "Spectrogram segmentation by means of statistical features for non-stationary signal interpretation," *IEEE Trans. Signal Processing*, vol. 50, pp. 2915–2925, December 2002.

[30] R. Adelino and F. da Silva, "Atomic decomposition with evolutionary pursuit," *Digital Signal Processing: A Review Journal*, vol. 13, pp. 317–337, April 2003.

[31] M. Davy and A. Doucet, "Copulas: A new insight into positive time-frequency distributions," *IEEE Signal Processing Letters*, vol. 10, pp. 215–218, July 2003.

[32] L. Knockaert, "A class of positive isentropic time-frequency distributions," *IEEE Signal Processing Letters*, vol. 9, pp. 22–25, January 2002.

[33] L. Angrisani and M. D'Arco, "A measurement method based on a modified version of the chirplet transform for instantaneous frequency estimation," *IEEE Trans. Instrumentation & Measurement*, vol. 51, pp. 704–711, August 2002.

[34] A. Papandreou-Suppappola and S. B. Suppappola, "Analysis and classification of time-varying signals with multiple time-frequency structures," *IEEE Signal Processing Letters*, vol. 9, pp. 92–95, March 2002.

[35] E. Grall-Maes and P. Beauseroy, "Mutual information-based feature extraction on the time-frequency plane," *IEEE Trans. Signal Processing*, vol. 50, pp. 779–790, April 2002.

[36] N. Stevenson, E. Palmer, J. Smeathers, and B. Boashash, "The BT product as a signal dependent sample size estimate in hypothesis testing: An application to linear/nonlinear discrimination in bandwidth limited systems," in *Proc. Seventh Internat. Symp. on Signal Processing and its Applications (ISSPA'03)*, pp. 551–554, Paris, 1–4 July 2003.

[37] A. Hanssen and L. L. Scharf, "A theory of polyspectra for nonstationary stochastic processes," *IEEE Trans. Signal Processing*, vol. 51, pp. 1243–1252, May 2003.

[38] N. Sang-Won and E. J. Powers, "Volterra series representation of time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 51, pp. 1532–1537, July 2003.

This Page Intentionally Left Blank

Part II

Fundamental Principles of TFSAP

This Page Intentionally Left Blank

Chapter 4

Time-Frequency Signal and System Analysis

This chapter extends the material described in Chapter 1. It presents advanced key principles underlying and justifying the use of time-frequency methods in signal and system analysis problems and applications. The topic is covered in nine focused articles with appropriate internal cross-referencing to this and other chapters. The issues and techniques described in this chapter were selected to provide different facets and perspectives to allow for a deeper insight into the foundations of the field.

The chapter begins with an alternative presentation of the concepts of analytic signal and IF with emphasis on the physical constraints leading to the mathematical model that is used to define the analytic signal (Article 4.1). Then, the key issue of cross-terms generation is described in detail for a better understanding of their localization in the time-frequency domain for a given signal. The quadratic superposition principle is used to explain the mechanism generating the cross-terms and the subsequent trade-off between cross-term reduction and increased localization (4.2). This is followed by an examination of the covariance property of TFDs for important signal transformations like (t, f) shifts or scaling, an important characteristic for signal processing applications (4.3). Another key aspect of time-frequency methods is that the (t, f) uncertainty relations determine the issue of lower bounds in achievable (t, f) resolution (4.4). Using methods such as coordinate change methods, we can also define joint distributions of other variables than t and f that may be better suited for specific applications (4.5). Having calculated and plotted the time-frequency representation of a given signal, we wish to make precise measurements and estimate signal parameters directly from the (t, f) plane. Formulations of measures such as spread measures are provided (4.6). Time-frequency methods can also be used to describe linear time-varying input-output relationships (4.7). The relationships between time-frequency methods such as the WVD and the fractional FT is described using the Radon-Wigner transform (4.8). Finally, a decomposition of the WVD via the Gabor expansions is able to link the spectrogram and the WVD (4.9).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

4.1 ANALYTIC SIGNAL & INSTANTANEOUS FREQUENCY[0]

4.1.1 The Problem

Instantaneous frequency (IF) is the basic concept used in numerous everyday systems like frequency modulation transmitters or receivers. However, even if this concept is easy to introduce intuitively [see Article 1.3], its rigorous mathematical definition leads to numerous difficulties.

In order to understand this point let us start from the very beginning. A purely monochromatic signal $x(t) = a\cos(\omega t + \phi)$ has an amplitude a, an angular frequency ω and an initial phase ϕ. Its *instantaneous phase* $\phi(t)$ is $\omega t + \phi$, which is a linear function of time, and the frequency is the derivative of the phase. It is obvious that such a signal cannot transmit information, and for this purpose some modulation procedure is required. By multiplying the carrier monofrequency signal $\cos(\omega t)$ by the positive function $m(t)$ we obtain the signal $m(t)\cos(\omega t)$, and it is quite natural to admit that its *instantaneous amplitude* is $m(t)$. Similarly, considering the signal $a\cos[\phi(t)]$, it is natural to say that its *instantaneous amplitude* is a and its *instantaneous phase* is $\phi(t)$. Its IF is obtained by differentiation of $\phi(t)$.

Even if the previous intuitive definitions appear quite natural and are widely used in practical applications dealing with signal modulation, we immediately note that they cannot be satisfactory. Indeed starting from a given signal $x(t)$ it is possible to introduce an infinite number of pairs $[a(t), \phi(t)]$ such that

$$x(t) = a(t)\cos[\phi(t)] \; . \tag{4.1.1}$$

This leads to the conclusion that the definitions given previously, even though they are widely used, are not coherent and must be reformulated. Indeed starting from a pair of functions $[a(t), \phi(t)]$ one can introduce by (4.1.1) a well-defined signal $x(t)$. But we also have to solve the inverse problem, i.e. to obtain a well-defined pair from the real signal $x(t)$, and this cannot be done without additional conditions.

There are numerous ways to realize this task. However for various reasons that cannot be discussed here, the most classical procedure makes use of the *analytic signal* (AS), which introduces the concept of *canonical pair*. The purpose of this article is to clearly introduce these concepts and to explore some of their consequences concerning the definitions and properties of IF presented in Article 1.3.

4.1.2 Analytic Signal and Canonical Pair

4.1.2.1 Mathematical Definitions

As said before, the problem is to realize a one-to-one correspondence between a real signal $x(t)$ and its canonical pair. This pair defines the instantaneous amplitude and phase of the signal and differentiating with respect to the time of the phase

[0]Author: **Bernard Picinbono**, Laboratoire des Signaux et Systèmes, Supélec, Plateau de Moulon, 91190 Gif-sur-Yvette, France (bernard.picinbono@lss.supelec.fr). Reviewers: M. Benidir and P. Flandrin.

yields the IF. The principle of the method is to realize a one-to-one correspondence between $x(t)$ and a complex signal $z(t)$, and to deduce the pair of functions as being the modulus and the argument of $z(t)$. The procedure makes use of the *analytic signal* (AS) defined as follows.

Let $X(f)$ be the Fourier transform (FT) of the signal $x(t)$. As $x(t)$ is real, $X(f)$ satisfies the Hermitian symmetry $X(f) = X^*(-f)$ which means that the negative frequencies in the spectrum do not get new information with respect to the positive frequencies. This is the basis of the definition of the AS which is deduced from $x(t)$ by suppression of these negative frequencies.

More precisely, let $Z(f)$ be the function defined by

$$Z(f) = 2u(f)X(f), \tag{4.1.2}$$

where $u(.)$ is the unit-step signal. The AS $z(t)$ of $x(t)$ is defined as being the IFT of $Z(f)$. Conversely we have $x(t) = (1/2)[z(t) + z^*(t)]$. Indeed the FT of $z^*(t)$ is $Z^*(-f)$ and using (4.1.2) and the Hermitian symmetry of $X(f)$ yields that the FT of $(1/2)[z(t) + z^*(t)]$ is $X(f)[u(f) + u(-f)] = X(f)$. This shows the one-to-one correspondence between any real signal $x(t)$ and its AS $z(t)$.

Another way to introduce the AS is to make use of the *Hilbert Transform* (HT). To any real signal it is possible to associate its HT $y(t) = \mathrm{H}[x(t)]$ defined as being the IFT of $Y(f) = -j\mathrm{Sg}(f)X(f)$, where Sg(.) means the sign function. It is clear that $y(t)$ is also real and furthermore it results from (4.1.2) that $z(t) = x(t) + jy(t)$, which means that the HT $y(t)$ of $x(t)$ is the imaginary part of its AS.

It is worth pointing out that $z(t)$ and $y(t)$ are deduced from $x(t)$ by linear filters with frequency responses $2u(f)$ and $-j\mathrm{Sg}(f)$, respectively.

As $z(t)$ cannot be a real function because its Fourier transform $Z(f)$ is zero for $f < 0$, it can be written as

$$z(t) = a(t)\exp[j\phi(t)]\,, \tag{4.1.3}$$

where the phase $\phi(t)$ is defined modulo 2π and $a(t)$ is non-negative. As a conclusion, using the AS makes it possible to associate with any real signal a unique pair $[a(t), \phi(t)]$ called in what follows its *canonical pair* [1].

It results from the definition that any complex signal the FT of which is zero for negative frequencies is an AS, and more precisely the AS of its real part. The class of AS is therefore defined by this property of the FT. A direct consequence of this definition is that the sum and the product of two ASs is still an AS. This is evident for the sum because FT is a linear operation. This comes from an obvious calculation of convolution for the product.

There is a long literature on the AS starting from the original papers [2,3]. There are some questions concerning the physical meaning of the AS, and some of them are mentioned and discussed in the review paper [4]. Furthermore it is shown in [5] that, starting from some *a priori* physical assumptions, the only possible definition of the instantaneous amplitude and phase is the one given just above. However it is worth pointing out that other physical conditions lead to other definitions that

are not discussed below [6]. Nevertheless the use of the AS introduces the only definition allowing easy calculations and interpretations.

Once the definition is given, the question which immediately follows is to characterize a canonical pair. Indeed, there is no reason for a given pair $[a(t), \phi(t)]$ to be canonical, and therefore amplitude and phase of signals cannot be arbitrary.

4.1.2.2 Characterization of a Canonical Pair

It results from our definitions that a pair of real functions $[a(t), \phi(t)]$ is canonical if and only if the complex function $z(t)$ defined by (4.1.3) is an AS which means that its FT $Z(f)$ is zero for negative frequencies. This is apparently a weak condition, but this intuition is wrong. On the contrary, the complex functions having this property are rather specific on a mathematical point of view, and the physical consequences of these properties will be analyzed in what follows.

As the characterization of an AS is deduced from spectral conditions (zero FT for negative frequencies) it is tempting to translate the same idea for $a(t)$ and $\phi(t)$. This can only be partially realized.

However, as the most general cases for applications of amplitude and frequency modulation concern narrow-band signals, it is appropriate to separate the contribution of the carrier frequency and to use, instead of (4.1.1), a signal in the form

$$x(t) = a(t)\cos[\omega_0 t + \phi(t)]\ , \tag{4.1.4}$$

In this case the complex signal $w(t) = a(t)\exp[j\phi(t)]$ is called the *complex envelope* of $x(t)$. Therefore the question is to know under which conditions on $[a(t), \phi(t)]$ it is possible to ensure that $w(t)\exp(j\omega_0 t)$ is an AS.

There is an elementary answer [7] introducing the concept of asymptotic AS discussed below. The idea is that any physical signal $w(t)$ is approximately band-limited. So, as multiplication by $\exp(j\omega_0 t)$ introduces a translation of the spectrum towards the high frequencies, the FT of $w(t)\exp(j\omega_0 t)$ becomes approximately zero for negative frequencies when ω_0 increases. This means that any signal in this form becomes approximately an AS for a sufficiently high value of the carrier frequency.

This reasoning based on spectral considerations is not without interest, but requires a more detailed mathematical analysis. For this purpose let us first consider the case of amplitude modulation. The corresponding pair is $[a(t), \omega_0 t + \phi]$. It is easy to verify that this pair of functions is canonical if and only if $a(t)$ is a positive band-limited signal, which means that its FT $A(f)$ is zero for $|f| > B$, where $B = \omega_0/2\pi$. This is a very simple spectral characterization of a canonical pair and this shows that for amplitude modulation the general reasoning presented just before can be rigorously applied.

It is therefore tempting to try to use spectral methods for the characterization of more general pairs of functions $[a(t), \phi(t)]$. Unfortunately this is more difficult.

Saying that $a(t)\exp[j\phi(t)]$ is an AS is equivalent to saying that the Hilbert transform of $a(t)\cos[\phi(t)]$ is equal to $a(t)\sin[\phi(t)]$. It is therefore appropriate to use the so-called Bedrosian theorem [8] dealing with the Hilbert transform of a product

of two real functions $x_1(t)$ and $x_2(t)$. A very simple derivation of this theorem and some extensions can be found in [9]. The main result is as follows: let $X_1(f)$ and $X_2(f)$ be the FTs of $x_1(t)$ and $x_2(t)$ respectively. If $x_1(t)$ and $x_2(t)$ are low and high frequency (B) signals respectively, or if $X_1(f) = 0$ for $f > B$ and $X_2(f) = 0$ for $f < B$, then

$$H[x_1(t)x_2(t)] = x_1(t)H[x_2(t)] \; . \tag{4.1.5}$$

A direct application of this result is that if $a(t)$ is a low-frequency signal (B) and $\cos[\phi(t)]$ a high-frequency signal (B), or if their spectra do not overlap, then

$$H\{a(t)\cos[\phi(t)]\} = a(t)H\{\cos[\phi(t)]\} \; . \tag{4.1.6}$$

However this does not at all imply that

$$H\{\cos[\phi(t)]\} = \sin[\phi(t)] \; , \tag{4.1.7}$$

as frequently admitted. In fact this equation would imply that $z(t) = \exp[\,j\phi(t)]$ is an AS or that the pair $[1, \phi(t)]$ is canonical. We shall see that this requires very specific properties of the structure of the phase $\phi(t)$ which cannot be characterized by spectral considerations only, as for amplitude modulation.

It results from this discussion that amplitude modulation can easily be treated by spectral considerations, but that this is no longer the case for frequency modulation. As this article is mainly devoted to IF, we shall consider signals with constant amplitude, or phase signals more in detail.

4.1.3 Phase Signals, Regular Case

4.1.3.1 General Structure

Phase signals (PSs) are the basic elements of phase or frequency modulation and this shows their importance. They are *real signals* with constant instantaneous amplitude. They are characterized by a canonical pair $[1, \phi(t)]$ or by a unique function $\phi(t)$ such that $\exp[\,j\phi(t)]$ is an AS. This means that (4.1.7) is satisfied. For such signals all the information is contained in the instantaneous phase (or frequency).

The condition that $\exp[\,j\phi(t)]$ is an AS requires very specific properties on the phase $\phi(t)$. These properties have been analyzed in the framework of the study of analytic functions and for example in Chapter 17 of [10].

The most general structure of the AS of a *regular phase signal* is

$$z(t) = b(t)\exp[\,j(\omega_0 t + \theta)] \; , \tag{4.1.8}$$

where θ is arbitrary, ω_0 is non-negative and $b(t)$ is a Blaschke function defined by

$$b(t) = \prod_{k=1}^{N} \frac{t - z_k}{t - z_k^*} \; , \quad z_k \in P_+ \; , \tag{4.1.9}$$

P_+ being the half-plane of the complex plane defined by $\mathrm{Im}(z) > 0$. The quantity ω_0 is the carrier angular frequency and can be equal to zero. The expression *regular* (or nonsingular) especially means that the number N of factors in the product is finite. The interpretation of (4.1.9) is very simple. In order to be an AS the function $b(t)$ for complex values of t must have all its poles in the lower half-plane $\mathrm{Im}(z) < 0$ because, as seen below, this ensures that its FT is zero for negative frequencies. In order to have a modulus equal to one each pole must be associated with a corresponding zero symmetric to this pole with respect to the real axis. This procedure is well known in filter theory: stable phase filters have the same number of poles and zeros and these zeros are symmetric of the poles with respect to the imaginary axis. The stability and causality conditions imply that all the poles are in the left half-plane of the complex plane.

It is obvious that $|b(t)| = 1$, which implies that $|z(t)| = 1$. Let us now explain why $z(t)$ is an AS. For this we must analyze the structure of the FT $B(f)$ of $b(t)$. As N is finite, $b(t)$ is a rational function in t. If all the z_k's are distinct we can write

$$b(t) = 1 + \sum_{k=1}^{N} \frac{c_k}{t - z_k^*} , \tag{4.1.10}$$

where $c_k = \lim_{t \to z_k^*} (t - z_k^*) b(t)$. As a consequence we have

$$B(f) = \delta(f) + \sum_{k=1}^{N} C_k(f) , \tag{4.1.11}$$

where $C_k(f)$ is the FT of $c_k(t - z_k^*)^{-1}$. Because of the localization of z_k^* in the complex plane, we deduce that $C_k(f) = 0$ for $f < 0$, which implies that $B(f) = 0$ for $f < 0$, and ensures that $b(t)$ is an AS. Finally, as $\omega_0 > 0$, $z(t)$ also is an AS. The reasoning can be extended without difficulty when some poles z_k are no longer distinct.

Of course the phase of $b(t)$ is

$$\phi_b(t) = \mathrm{Arg}[b(t)] \, , \mathrm{mod}(2\pi) , \tag{4.1.12}$$

and, as a result, we can say that any phase signal can be written as (4.1.1) where $a(t) = 1$ and $\phi(t)$ must have the form

$$\phi(t) = \theta + \omega_0 t + \phi_b(t) \, , \mathrm{mod}(2\pi) . \tag{4.1.13}$$

In practice the continuity of the phase leads to suppressing the term $\mathrm{mod}(2\pi)$ and this convention is adopted in all that follows.

This most general phase is defined by N complex parameters z_k and 2 real parameters ω_0 and θ. Furthermore it is obvious that the phase $\phi_b(t)$ is the sum of the N phases of the factors appearing in the product (4.1.9). Let $\phi_k(t)$ be the phase of $(t - z_k)(t - z_k^*)^{-1}$. This gives

$$\phi(t) = \theta + \omega_0 t + \sum_{k=1}^{N} \phi_k(t). \tag{4.1.14}$$

By introducing the real and imaginary parts of z_k, or $z_k = a_k + jb_k$, one obtains

$$\phi_k(t) = 2\arctan\frac{b_k}{a_k - t}\ , -\pi/2 \leq \phi_k(t) \leq \pi/2\ . \tag{4.1.15}$$

In conclusion the signal $\cos[\phi(t)]$ is regularly phase modulated only if its phase takes the form (4.1.13), and, as this is generally not the case, its amplitude is not constant.

4.1.3.2 Properties of Regular Phase Signals

PSs have numerous specific properties analyzed in [1]. We shall only outline the most significant. The first point to note is that PSs do not have a finite energy and, as a consequence, the calculation of their FT requires some care, often due to the presence of Dirac distributions characterizing spectral lines.

In reality (4.1.8) and (4.1.11) show that the AS of a PS has only one spectral line located at the carrier frequency $f_0 = \omega_0/2\pi$ and its FT is zero for frequencies smaller than f_0. As a consequence PSs are high frequency (f_0) signals. Furthermore the functions $C_k(f)$ appearing in (4.1.11) that are zero for negative frequencies, are also bounded and cannot be equal to zero for positive frequencies. This means that a PS cannot be a low-frequency signal except when it is monochromatic.

It is now interesting to study the structure of the IF of a PS. It is obtained by differentiating the instantaneous phase. The most general form of this phase is given by (4.1.14) and (4.1.15), and differentiating this equation yields the angular IF

$$\omega(t) = \omega_0 + \omega_m(t) = \omega_0 + 2\sum_{k=1}^{N}\frac{b_k}{b_k^2 + (a_k - t)^2}\ . \tag{4.1.16}$$

As the coefficients b_k are positive, because of the localization of the zeros z_k, we deduce that the IF $\omega(t)$ is always greater than ω_0. This is another illustration of the fact that the FT of $x(t)$ is zero for $f < f_0$ and this also shows that the IF $\omega(t)$ belongs to the frequency domain of this FT.

It is worth pointing out that the information carried by the IF of a phase signal is entirely in the term $\omega_m(t)$ where the index m means the modulation term. We note that this function tends to zero when $|t| \to \infty$. This means that $\omega_m(t)$ cannot be a periodic function, and this is related to the fact that a regular phase signal cannot have spectral lines, except those coming from the carrier frequency f_0.

Furthermore we note that $\omega_m(t)$ is a rational function in t. The polynomials appearing in the numerator and the denominator have the degrees $2N-2$ and $2N$, respectively. As N is arbitrary, we deduce that by using the $2N$ parameters a_k and b_k it is possible to approximate a large class of functions. The most limiting constraint on these functions comes from the necessary behavior for $|t| \to \infty$. In fact $\omega_m(t)$ decreases at infinity in $|t|^{-2}$, which is a strong restriction on the IF.

4.1.4 Singular and Asymptotic Phase Signals

4.1.4.1 Infinite Products

The simplest example of a Blaschke function with an infinite value of N is when the poles of (4.1.9) are $z_k = kT + jb$, with $T = 2\pi/\omega$ and $b = (1/\omega)\ln(1/a)$ with $0 < a < 1$. In this case it can be shown that

$$b(z) = \prod_{k=-\infty}^{+\infty} \frac{z - z_k}{z - z_k^*} = \frac{a - \mathrm{e}^{j\omega z}}{1 - a\mathrm{e}^{j\omega z}} \,. \tag{4.1.17}$$

This means that the function

$$b(t) = \frac{a - \mathrm{e}^{j\omega t}}{1 - a\mathrm{e}^{j\omega t}} \tag{4.1.18}$$

is the AS of a phase signal. This can immediately be verified because $|z(t)| = 1$ and the FT is zero for negative frequencies. In order to verify this point it suffices to note that $b(t)$ is periodic, and by using the geometric series one sees that the Fourier coefficients F_n are zero for $n < 0$ and furthermore we have $F_0 = a$ and $F_n = -(1 - a^2)a^{n-1}$.

The corresponding IF is

$$\omega(t) = \omega_0 + \omega \frac{1 - a^2}{1 - 2a\cos(\omega t) + a^2} \,. \tag{4.1.19}$$

Note the analogy with (4.1.16). Most of the properties of PSs indicated previously are verified, except those using the point that N is finite. In this singular case phase signals contain only spectral lines, but their number is infinite. There are various examples of phase signals that can be generated from (4.1.18) by products of AS of various phase signals of the previous form.

4.1.4.2 Poles on the Real Axis

By using arguments that cannot be presented here (see theorem (17-15) in [10]) one can show that the function $z(t) = \exp(-j/t)$ is an AS. This means that the signal $x(t) = \cos(1/at)$, $a > 0$, is a PS and its Hilbert transform is $-\sin(1/at)$. These signals are called *hyperbolic phase chirps* and have a strongly oscillatory behavior in the neighborhood of the origin of time. It is interesting to note that the explicit expression of its FT can be obtained.

4.1.4.3 Asymptotic Phase Signals

We have seen that if $w(t)$ is band-limited, the signal $w(t)\exp(j\omega_0 t)$ becomes an AS for ω_0 sufficiently high. This situation can appear for completely different reasons an we shall discuss the particular example of the so-called *parabolic phase chirp*. Let us therefore consider the signal $x(t) = \cos[\omega_0 t + a^2 t^2]$. It is not a PS, because

its phase does not have the structure analyzed above. This means that its AS is not $v(t) = \exp[\omega_0 t + a^2 t^2]$. Indeed this function is not an AS because its FT is $V(f) = k \exp[-jb(f - f_0)^2]$, which is not zero for negative frequencies. Furthermore it does not decrease for large values of $|f|$. As this decreasing behavior is the argument indicated in [7] and mentioned above to say that $w(t) \exp(j\omega_0 t)$ can be an AS for $\omega_0 \to \infty$, this argument cannot be used here. However the character of asymptotic AS is nevertheless true, but for completely different reasons. The FT $V(f)$ satisfies $|V(f)|^2 = |k|^2$, but makes very rapid oscillations for negative frequencies and large values of f_0, in such a way that even if it is never vanishing, its mean value in any frequency interval is approximately zero. In other words $v(t)$ can never be considered as an AS, as defined previously, but in any integration in the frequency domain it has the same properties as an AS. This reasoning can be extended to other cases that cannot be presented here.

4.1.5 Summary and Conclusions

The analytic signal is the most convenient theoretical means to introduce the concept of instantaneous frequency of signals. However its mathematical structure introduces some constraints, especially for phase signals used in frequency modulation. Most of these constraints are presented and discussed.

References

[1] B. Picinbono, "On Instantaneous Amplitude and Phase of Signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 552–560, March 1997.

[2] D. Gabor, "Theory of communication," *J. IEE*, vol. 93(III), pp. 429–457, November 1946.

[3] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmissions*, vol. 2A, no. 1, pp. 61–74, 1948. In French. English translation: I. Selin, *Theory and applications of the notion of complex signal*, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).

[4] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals," *Proc. IEEE*, vol. 80, pp. 520–538, April 1992.

[5] D. Vakman, "On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency," *IEEE Trans. Signal Processing*, vol. 44, pp. 791–797, April 1996.

[6] P. Loughlin and B. Tacer, "On the amplitude and frequency-modulation decomposition of signals," *J. Acoustical Soc. of America*, vol. 100, pp. 1594–1601, September 1996.

[7] A. H. Nuttall, "On the quadrature approximation to the Hilbert transform of modulated signals," *Proc. IEEE*, vol. 54, pp. 1458–1459, 1966.

[8] E. Bedrosian, "A product theorem for Hilbert transforms," *Proc. IEEE*, vol. 51, pp. 686–689, 1963.

[9] B. Picinbono and W. Martin, "Représentation des signaux par amplitude et phase instantanées," *Ann. Télécommunications*, vol. 38, pp. 179–190, 1983.

[10] W. Rudin, *Real and complex analysis*. New York: McGraw-Hill, 1987.

4.2 CROSS-TERMS & LOCALIZATION IN QUADRATIC TIME-FREQUENCY DISTRIBUTIONS [0]

4.2.1 Identifying Cross-Terms

4.2.1.1 Cross-Terms in Spectrum Analysis

Let $x(t)$ and $y(t)$ be any two signals, and a and b any two complex numbers. If $X(f) := \mathcal{F}_{t\to f}\{x(t)\}$ stands for the Fourier transform of $x(t)$—with a similar notation for $y(t)$—, we then have

$$|\mathcal{F}_{t\to f}\{ax(t)+by(t)\}|^2 = |a|^2|X(f)|^2 + |b|^2|Y(f)|^2 + 2\mathrm{Re}\{ab^*X(f)Y^*(f)\}. \quad (4.2.1)$$

This means that the spectrum energy density of the sum of two signals doesn't reduce to the sum of the individual densities (unless the signals are spectrally disjoint), but satisfies a quadratic superposition principle involving a third term, referred to as a *cross-term*. While this situation is, mathematically speaking, a necessary by-product of the quadratic nature of the transformation and of the elementary identity $(a+b)^2 = a^2 + b^2 + 2ab$, it also corresponds to a physical reality in terms of interfering waves. This can be illustrated by the simple example where a delayed replica is superimposed to a given waveform. Assuming the model $y(t) = x(t-\tau)$, with $a = b = 1$, we get:

$$|\mathcal{F}_{t\to f}\{x(t)+x(t-\tau)\}|^2 = 4\,(\cos\pi\tau f)^2\,|X(f)|^2, \quad (4.2.2)$$

and the existence of a cross-term is instrumental in the creation, in the composite spectrum, of "fringes" whose periodicity in f is controlled by the delay τ: it is in fact the basis of interferometry.

4.2.1.2 Cross-Terms, from Spectrograms to the Wigner-Ville Distribution

Switching to time-frequency distributions, similar considerations apply naturally to spectrograms defined by $S_x(t,f) := |F_x(t,f)|^2$, with

$$F_x(t,f) := \mathcal{F}_{s\to f}\{x(s)\,h^*(s-t)\}, \quad (4.2.3)$$

where $h(t)$ is a short-time analysis window. We get in this case

$$S_{ax+by}(t,f) = |a|^2 S_x(t,f) + |b|^2 S_y(t,f) + 2\mathrm{Re}\{ab^*F_x(t,f)F_y^*(t,f)\}, \quad (4.2.4)$$

and cross-terms show up as long as the (essential) time-frequency supports of the short-time Fourier transforms $F_x(t,f)$ and $F_y(t,f)$ overlap (see Figure 4.2.1). A similar situation is observed with the scalograms (i.e. the squared modulus of the wavelet transform).

---

[0]Author: **Patrick Flandrin**, Laboratoire de Physique (UMR 5672 CNRS), *Ecole Normale Supérieure de Lyon*, 46 allée d'Italie, 69364 Lyon Cedex 07, France (flandrin@ens-lyon.fr). Reviewers: F. Hlawatsch and W. J. Williams.

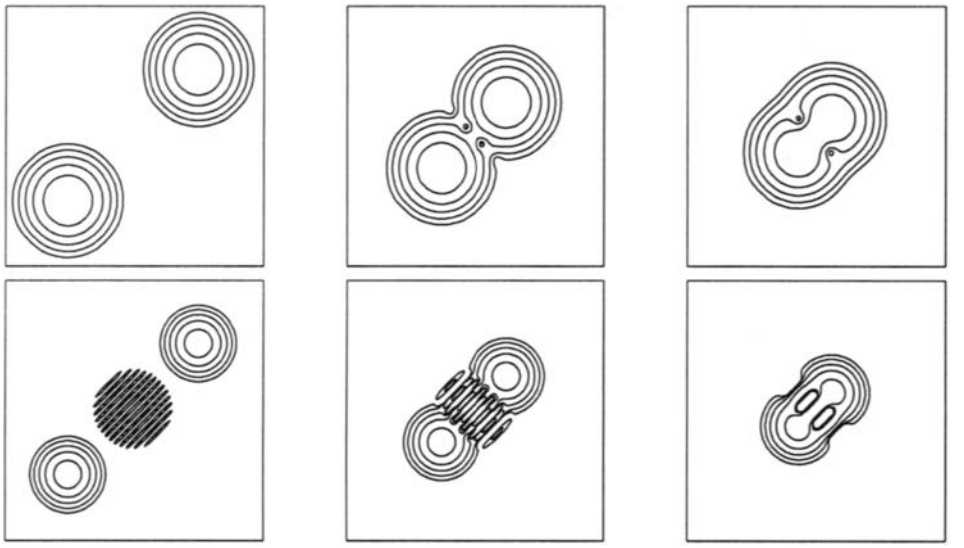


Fig. 4.2.1: Cross-terms in the time-frequency analysis of two shifted Gaussian wave packets. In the spectrogram case (top row), cross-terms are only present when the effective supports of the individual contributions overlap, and they fade out when the distance between the components increases. In the Wigner-Ville case (bottom row), cross-terms persistently exist as an oscillating contribution located midway between the two components, with an increasing oscillation rate when the inter-components distance increases. (In all diagrams, time is horizontal, frequency is vertical, and isocontours of the (positive parts of the) distributions are displayed.)

As squared linear transforms, spectrograms happen however to be a restricted version of quadratic transforms. In fact, they only correspond to a special case of the more general Cohen's class,[1] which consists of all quadratic time-frequency energy distributions covariant to time and frequency shifts, and whose most prominent member is the Wigner-Ville distribution (WVD) [1]:

$$W_x(t,f) := \mathcal{F}_{\tau\to f}\{x(t+\tau/2)x^*(t-\tau/2)\}. \tag{4.2.5}$$

The WVD itself satisfies a quadratic superposition principle according to which

$$W_{ax+by}(t,f) = |a|^2 W_x(t,f) + |b|^2 W_y(t,f) + 2\mathrm{Re}\{ab^* W_{x,y}(t,f)\}, \tag{4.2.6}$$

with

$$W_{x,y}(t,f) := \mathcal{F}_{\tau\to f}\{x(t+\tau/2)y^*(t-\tau/2)\}. \tag{4.2.7}$$

As compared to spectrograms, the situation is, however, drastically changed in the sense that increasing the (time-frequency) distance between the components no

[1] That is, the quadratic class; see p. 68n.

longer implies a vanishing of the cross-term controlled by (4.2.7). More precisely, if we let

$$x_{\pm}(t) := x_0(t \mp \Delta t/2) \exp\{\pm i2\pi(\Delta f/2)t\}, \tag{4.2.8}$$

we get (see Figure 4.2.1):

$$W_{x_+ + x_-}(t,f) = W_{x_0}(t-\Delta t/2, f-\Delta f/2) + W_{x_0}(t+\Delta t/2, f+\Delta f/2) + I(t,f), \tag{4.2.9}$$

with

$$I(t,f) = 2W_{x_0}(t,f)\cos[2\pi(t\Delta f - f\Delta t)]. \tag{4.2.10}$$

This simplified situation exemplifies the general rules which control the *interference geometry* of the WVD [2,3], whose cross-terms:

1. are located *midway* between the interacting components;
2. *oscillate* proportionally to the inter-components' time-frequency distance;
3. have a direction of oscillation *orthogonal* to the straight line connecting the components.

4.2.2 Reducing Cross-Terms

In the case of multicomponent signals of the form

$$x(t) = \sum_{n=1}^{N} x_n(t), \tag{4.2.11}$$

WVD cross-terms are created in between any two components, leading to a combinatorial proliferation that quickly hampers readability. Cross-terms reduction is therefore an important issue in many applications. Since cross-terms appear between any two contributions, at positive as well as negative frequencies, a first recipe is to compute WVDs on analytic signals rather than on real-valued ones. The reason is that the recourse to the analytic signal forces spectrum contributions to be zero on the real half-line of negative frequencies, thus suppressing, *de facto*, cross-terms stemming from interactions between negative and positive frequencies.

4.2.2.1 The Ambiguity Domain Interpretation

A further way of reducing WVD cross-terms amounts to identifying them as oscillating contributions (as opposed to the smoother variations attached to "signal terms"), thus suggesting to make use of a low-pass smoothing. The spectrogram case discussed above is in fact a first instance of such a procedure, since it is well-known (see, e.g., [1]) that:

$$S_x(t,f) = \int\int_{-\infty}^{+\infty} W_x(s,\xi)\, W_h(s-t, \xi - f)\, ds\, d\xi, \tag{4.2.12}$$

with the low-pass nature of the window $h(t)$ carrying over to its WVD $W_h(t,f)$. More generally, replacing in (4.2.12) $W_h(t,f)$ by some (almost) arbitrary function $\gamma(t,f)$ defines Cohen's class [1]:

$$\rho_x(t,f;\gamma) := \int\!\!\int_{-\infty}^{+\infty} W_x(s,\xi)\,\gamma(s-t,\xi-f)\,ds\,d\xi, \tag{4.2.13}$$

and offers a versatile framework for achieving cross-terms reduction.

Among the many possible parameterizations of Cohen's class, one that proves most powerful for the considered problem operates in the so-called *ambiguity plane*, obtained by Fourier duality from the time-frequency plane (cf. Section 3.2.1 and Article 5.1). Starting from the general form (4.2.13) of Cohen's class with (time-frequency) kernel $\gamma(t,f)$, we obtain by a 2D Fourier transformation that

$$\rho_x(t,f;\gamma) \rightleftharpoons \mathcal{A}_x(\nu,\tau;g) := g(\nu,\tau)\,A_x(\nu,\tau), \tag{4.2.14}$$

with $g(\nu,\tau)$ the 2D inverse Fourier transform of $\gamma(t,f)$ and

$$A_x(\nu,\tau) := \mathcal{F}_{t\to-\nu}\{x(t+\tau/2)x^*(t-\tau/2)\} \tag{4.2.15}$$

the *ambiguity function* (AF) of $x(t)$: in the ambiguity plane, Cohen's class is nothing but a weighted ambiguity function.

In the case of the two-component signal (4.2.8), we have

$$A_{x_++x_-}(\nu,\tau) = 2A_{x_0}(\nu,\tau)\cos[\pi(\nu\Delta t+\tau\Delta f)] + J(\nu,\tau), \tag{4.2.16}$$

with

$$J(\nu,\tau) = A_{x_0}(\nu+\Delta f,\tau-\Delta t) + A_{x_0}(\nu-\Delta f,\tau+\Delta t). \tag{4.2.17}$$

In accordance with the interpretation of the AF as a time-frequency correlation function (cf. Article 5.1), it thus appears that signal terms are mainly concentrated in the vicinity of the origin of the ambiguity plane, while cross-terms are (symmetrically) located at a distance from the origin which is equal to the time-frequency distance between the interacting components.

4.2.2.2 Kernel Design for Reduced Interference

The above AF interpretation [4] offers an alternative distinction between "signal terms" and "cross-terms", and is the clue for a possible reduction of the latter: according to (4.2.14), it suffices to choose a kernel function $g(\nu,\tau)$ so as to keep the AF as unchanged as possible in the vicinity of the origin, while suppressing as much as possible contributions off the origin. In a first approximation, the kernel function has therefore to be of *low-pass* type, thus guaranteeing that it equivalently operates a smoothing in the time-frequency plane. In this respect, a versatile solution is to make use of a *separable* kernel function of the form $\gamma(t,f) = k(t)\,H(f)$. The associated distribution is referred to as a "smoothed pseudo-WVD" [1], and it allows qualitatively for a smooth transition between the WVD (no smoothing, i.e., $\gamma(t,f) =$

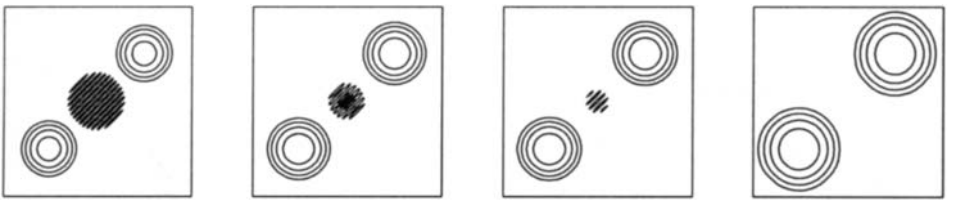


Fig. 4.2.2: Cross-terms reduction by time-frequency smoothing. Using a smoothed pseudo-WVD with a separable kernel allows for a continuous transition between a WVD (left) and a spectrogram (right).

$\delta(t)\,\delta(f)$) and spectrograms (lower bounded smoothing, $\gamma(t,f) = W_h(t,f)$), see Figure 4.2.2.

More elaborate approaches have been developed, either by incorporating specific constraints to be satisfied by "reduced interference distributions" (see [5] and Article 5.2) or by making the kernel signal-dependent (see [6] and Article 5.3), so as to reject at best cross-terms contributions located off the origin in the ambiguity plane.

4.2.3 Cross-Terms and Localization

As presented so far, cross-terms could be thought of as being specific of "multicomponent" situations. However, the notion of a multicomponent signal—as it is modeled in (4.2.11)—is somehow artificial, in the sense that any signal can always be, at will, split into arbitrary parts. This is especially the case for *frequency modulated* signals that intuition (and physics) would like to identify as monocomponent, even if any arbitrary chopping is expected to create cross-terms between the so-created "components". In such a situation, cross-terms are referred to as "inner interference terms", in contrast to "outer interference terms" resulting from the interaction between contributions attached to time-frequency domains that are essentially disconnected [2,3].

4.2.3.1 Localization from Interference

Iterating *ad infinitum* the procedure of chopping a signal into more and more elementary parts leads to a pointwise application of the WVD cross-term construction rules enounced before. In a nutshell, if we admit to describe a chirp signal by means of its instantaneous frequency trajectory in the time-frequency plane, inner cross-terms are created midway of the chord between any two points located on this trajectory, with a possible reduction when applying suitable smoothings (see Figure 4.2.3).

While a more refined description of this mechanism can be obtained on the basis of stationary phase arguments, a companion viewpoint is given by Janssen's

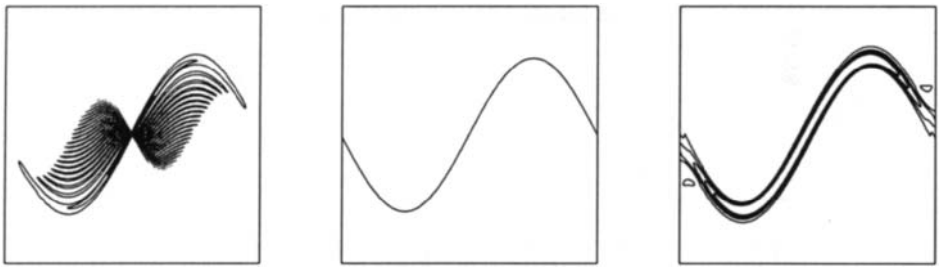


Fig. 4.2.3: Inner interference terms in the time-frequency analysis of a chirp signal. In the WVD case (left), cross-terms appear midway between any two points belonging to the instantaneous frequency trajectory (middle). As for the "outer" interference terms of Figure 4.2.2, these "inner" interference terms can be reduced by using a smoothed pseudo-WVD (right).

interference formula:

$$W_x^2(t,f) = \int\!\!\int_{-\infty}^{+\infty} W_x(t+\tau/2, f+\nu/2)\, W_x(t-\tau/2, f-\nu/2)\, d\tau\, d\nu, \qquad (4.2.18)$$

according to which a non-zero value of the WVD at a given time-frequency point results from the superposition of all values of the very same distribution that are symmetrically located with respect to the considered point. This means that the WVD is indeed the result of some "holographic" construction that, in some sense, makes impossible a clear-cut distinction between "signal terms" and "cross-terms".

As a further consequence, *localization* of the WVD for linear chirps [1] can also be viewed as a by-product of interference. The geometric explanation of this well-known property is as follows: if a signal is to have a WVD which is perfectly localized on a given time-frequency curve of the plane, it is then necessary—by application of the interference geometry principle—that all of the mid-points between any two points located on the curve belong themselves to the same curve. It follows that *linear* chirps are the only admissible solutions in the WVD case, since straight lines are the only curves of the plane that are the geometric locus of all of their mid-points (see Figure 4.2.4). This principle ("a signal term is nothing but the emergence of coherent cross-terms") is very general and can be extended to other types of quadratic time-frequency distributions (e.g., affine) equipped with suitably modified geometries, thus justifying geometrically their localization on nonlinear chirps [7].

4.2.3.2 The Interference/Localization Trade-Off

In ordinary Fourier analysis, and therefore in spectrogram analysis, there exists a necessary trade-off between the resolution in time and the resolution in frequency. In the more general framework of Cohen's class, the situation has to be interpreted differently, with a new trade-off between the *joint* resolution (in both time and frequency) on the one hand and the level of cross-terms on the other hand. This clearly

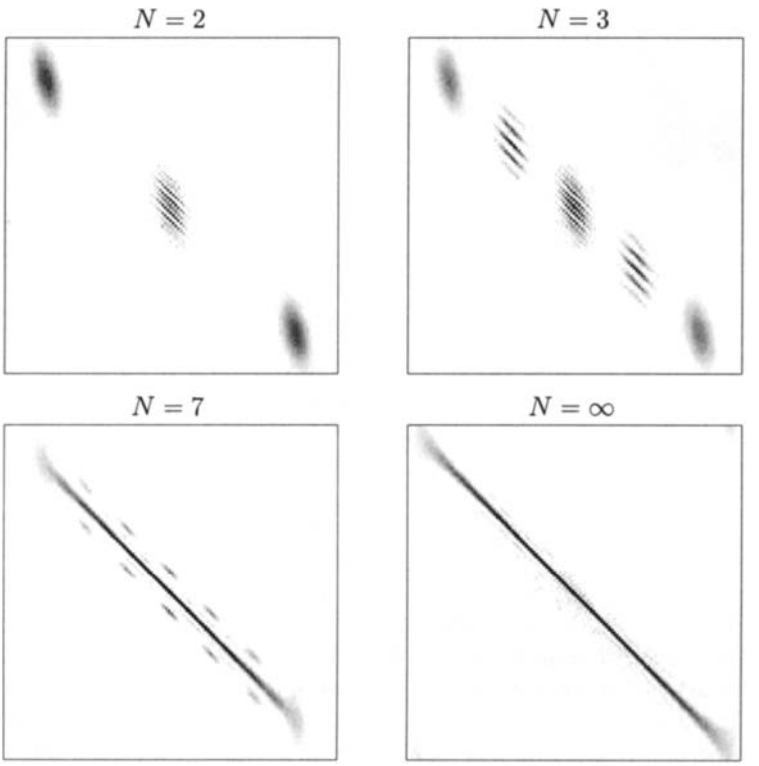


Fig. 4.2.4: Localization from interference. The localization property of the WVD on straight lines of the time-frequency plane can be seen as resulting from a constructive interference process. Cross-terms of the WVD being located midway between any two interacting components (see Figure 4.2.1), it is here shown how an increasing number N of aligned wave packets creates an increasing number of cross-terms that are aligned too. In the limit where $N \to \infty$, this leads to a perfect localization of the distribution along the line, which is the geometric locus of all of its mid-points.

appears in the examples of Figures 4.2.2 and 4.2.3, where a fading out of cross-terms (be they "inner" or "outer") is achieved only at the expense of spreading out signal terms, i.e., a resolution loss. It is worth noting that overcoming such a trade-off is possible to some extent, by using reassignment techniques (cf. Article 7.2).

4.2.4 Summary and Conclusions

Quadratic time-frequency distributions satisfy a quadratic superposition principle which necessarily results in the existence of cross-terms. On the one hand, such cross-terms may be considered as troublesome since, in the case of a signal composed of many (time-frequency disjoint) components, their combinatorial proliferation hampers readability; on the other hand, in the case of individual components like chirps, the interferential generation of cross-terms may guarantee an improved

localization as compared to linear transforms. Understanding the mechanism underlying cross-terms is therefore a key for trading-off interference reduction and increased localization in quadratic time-frequency distributions.

References

[1] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[2] F. Hlawatsch, "Interference terms in the Wigner distribution," in *Proc. Internat. Conf. on Digital Signal Processing*, pp. 363–367, Florence, Italy, 5–8 September 1984.

[3] 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.), pp. 59–133, Amsterdam: Elsevier, 1997.

[4] P. Flandrin, "Some features of time-frequency representations of multicomponent signals," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'84)*, vol. 3, pp. 41B.4.1–41B.4.4, San Diego, 19–21 March 1984.

[5] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 37, pp. 862–871, June 1989.

[6] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Trans. Signal Processing*, vol. 41, pp. 1589–1602, April 1993.

[7] P. Flandrin and P. Gonçalvès, "Geometry of affine time-frequency distributions," *Applied & Computational Harmonic Analysis*, vol. 3, pp. 10–39, January 1996.

4.3 THE COVARIANCE THEORY OF TIME-FREQUENCY ANALYSIS [0]

4.3.1 The Covariance Principle

Many important classes of linear and bilinear/quadratic time-frequency representations (TFRs) can be defined by a *covariance property*. For example, the family of *short-time Fourier transforms* [1,2]

$$F_x^h(t,f) = \int_{-\infty}^{\infty} x(t')\, h^*(t'-t)\, e^{-j2\pi f t'}\, dt' \tag{4.3.1}$$

(where t and f denote time and frequency, respectively, $x(t)$ is the signal under analysis, and $h(t)$ is a function that does not depend on $x(t)$) consists of all linear TFRs L that are *covariant* to time-frequency (TF) shifts according to

$$L_{\mathbf{S}_{\tau,\nu}x}(t,f) = e^{-j2\pi(f-\nu)\tau}\, L_x(t-\tau, f-\nu)\,. \tag{4.3.2}$$

Here, $\mathbf{S}_{\tau,\nu}$ is the TF shift operator defined as $\big(\mathbf{S}_{\tau,\nu}x\big)(t) = x(t-\tau)\, e^{j2\pi\nu t}$. Thus, among all linear TFRs, the short-time Fourier transform is axiomatically defined by the TF shift covariance property (4.3.2). Similarly, *Cohen's class* (with signal-independent kernel $h(t_1,t_2)$), given by [1]

$$C_{x,y}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1)\, y^*(t_2)\, h^*(t_1-t, t_2-t)\, e^{-j2\pi f(t_1-t_2)}\, dt_1 dt_2\,, \tag{4.3.3}$$

comprises all bilinear/quadratic TFRs B that are covariant to TF shifts according to

$$B_{\mathbf{S}_{\tau,\nu}x,\mathbf{S}_{\tau,\nu}y}(t,f) = B_{x,y}(t-\tau, f-\nu)\,. \tag{4.3.4}$$

Similar covariance-based interpretations and definitions can be given for many other important classes of linear TFRs (e.g., wavelet transform, hyperbolic wavelet transform, and power wavelet transform [2–4]) as well as bilinear/quadratic TFRs (e.g., affine, hyperbolic, and power classes [1,3–6]; see also Articles 5.6, 7.1, and 15.3).

In this article, we present a unified *covariance theory of TF analysis* that allows the systematic construction of covariant TFRs [7–9]. (See [10] for a much more detailed treatment.) Covariance properties are important in TF analysis since specific unitary signal transformations often occur in practice—e.g., time shifts and frequency shifts as described by the TF shift operator $\mathbf{S}_{\tau,\nu}$ correspond to the delays and Doppler shifts, respectively, encountered in radar and mobile communications.

[0] Authors: **F. Hlawatsch**, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Gusshausstrasse 25/389, A-1040 Vienna, Austria (email: fhlawats@pop.tuwien.ac.at, web: http://www.nt.tuwien.ac.at/dspgroup/time.html) and **G. Tauböck**, Telecommunications Research Center Vienna (FTW), Tech Gate Vienna, Donau-City-Strasse 1, A-1220 Vienna, Austria (email: tauboeck@ftw.at, web: http://www.ftw.at). Reviewers: J.-P. Ovarlez and A. Papandreou-Suppappola.

4.3.2 Time-Frequency Displacement Operators

A key element of our covariance theory is the concept of *TF displacement operators* (DOs), which are operators that displace signals in the TF plane. The DO concept generalizes the TF shift operator $\mathbf{S}_{\tau,\nu}$, for which the TF displacements are simple shifts (translations), to other types or geometries of TF displacements.

Group fundamentals. In what follows, we will need some fundamentals of groups. A set $\mathcal{G}$ together with a binary operation $\star$ that maps $\mathcal{G}\times\mathcal{G}$ to $\mathcal{G}$ is called a *group* if it satisfies the following properties:

1. There exists an *identity element* $g_0 \in \mathcal{G}$ such that $g \star g_0 = g_0 \star g = g$ for all $g \in \mathcal{G}$.

2. To every $g \in \mathcal{G}$, there exists an *inverse element* $g^{-1} \in \mathcal{G}$ such that $g \star g^{-1} = g^{-1} \star g = g_0$.

3. Associative law: $g_1 \star (g_2 \star g_3) = (g_1 \star g_2) \star g_3$ for all $g_1, g_2, g_3 \in \mathcal{G}$.

If, in addition, $g_1 \star g_2 = g_2 \star g_1$ for all $g_1, g_2 \in \mathcal{G}$, the group is called *commutative* or *Abelian*. An elementary example of a commutative group is $(\mathbb{R}, +)$ for which $g_1 \star g_2 = g_1 + g_2$, $g_0 = 0$, and $g^{-1} = -g$. Two groups $(\mathcal{G}, \star)$ and $(\mathcal{H}, \diamond)$ are said to be *isomorphic* if there exists an invertible mapping $\psi : \mathcal{G} \to \mathcal{H}$ such that $\psi(g_1 \star g_2) = \psi(g_1) \diamond \psi(g_2)$ for all $g_1, g_2 \in \mathcal{G}$.

Definition and examples of DOs. We are now ready to give a formal definition of DOs. A DO is a family of unitary operators $\mathbf{D}_{\alpha,\beta}$ indexed by a 2-D "displacement parameter" (α, β) that belongs to some group $(\mathcal{D}, \circ)$. This operator family $\{\mathbf{D}_{\alpha,\beta}\}_{(\alpha,\beta)\in\mathcal{D}}$ is supposed to satisfy the following two properties:

1. A displacement by the group identity parameter (α_0, β_0) is no displacement, i.e.,

$$\mathbf{D}_{\alpha_0,\beta_0} = \mathbf{I},$$

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

2. A displacement by (α_1, β_1) followed by a displacement by (α_2, β_2) is equivalent (up to a phase factor) to a single displacement by $(\alpha_1, \beta_1) \circ (\alpha_2, \beta_2)$, i.e.,

$$\mathbf{D}_{\alpha_2,\beta_2}\mathbf{D}_{\alpha_1,\beta_1} = e^{j\phi(\alpha_1,\beta_1;\alpha_2,\beta_2)}\,\mathbf{D}_{(\alpha_1,\beta_1)\circ(\alpha_2,\beta_2)} \quad \forall\ (\alpha_1,\beta_1),(\alpha_2,\beta_2) \in \mathcal{D}, \tag{4.3.5}$$

   with $\phi(\alpha_1, \beta_1; \alpha_2, \beta_2)$ being a continuous function.

More precisely, stated in mathematical terms, a DO $\mathbf{D}_{\alpha,\beta}$ is an *irreducible and faithful projective representation* of a group $(\mathcal{D}, \circ)$; the function $e^{j\phi(\alpha_1,\beta_1;\alpha_2,\beta_2)}$ is known as the *cocycle*. For $e^{j\phi(\alpha_1,\beta_1;\alpha_2,\beta_2)} \equiv 1$, $\mathbf{D}_{\alpha,\beta}$ is a *unitary* group representation [10].

Two basic examples of a DO are the following:

- The TF shift operator $\mathbf{S}_{\tau,\nu}$. Here, $(\alpha,\beta)=(\tau,\nu)$ and $(\mathcal{D},\circ)$ is the commutative group $(\mathbb{R}^2,+)$ with operation $(\tau_1,\nu_1)\circ(\tau_2,\nu_2)=(\tau_1+\tau_2,\nu_1+\nu_2)$; furthermore $(\tau_0,\nu_0)=(0,0)$, $(\tau,\nu)^{-1}=(-\tau,-\nu)$, and $\phi(\tau_1,\nu_1;\tau_2,\nu_2)=-2\pi\nu_1\tau_2$.

- The TF scaling/time shift operator $\mathbf{R}_{\sigma,\tau}$ defined as $\big(\mathbf{R}_{\sigma,\tau}x\big)(t)=\frac{1}{\sqrt{|\sigma|}}\,x\big(\frac{t-\tau}{\sigma}\big)$. Here, $(\alpha,\beta)=(\sigma,\tau)$, $(\mathcal{D},\circ)$ is the noncommutative *affine group* with $\mathcal{D}=\mathbb{R}^+\times\mathbb{R}$ and group operation $(\sigma_1,\tau_1)\circ(\sigma_2,\tau_2)=(\sigma_1\sigma_2,\tau_1\sigma_2+\tau_2)$; furthermore $(\sigma_0,\tau_0)=(1,0)$, $(\sigma,\tau)^{-1}=(1/\sigma,-\tau/\sigma)$, and $\phi(\sigma_1,\tau_1;\sigma_2,\tau_2)\equiv 0$.

Additional structure of DOs. The interpretation that a DO $\mathbf{D}_{\alpha,\beta}$ performs TF displacements motivates certain topological assumptions which can be shown [10] to imply that $(\mathcal{D},\circ)$ is a *simply connected 2-D Lie group*. This, in turn, can be shown to have the following two important consequences [10]:

1. The group $(\mathcal{D},\circ)$ underlying $\mathbf{D}_{\alpha,\beta}$ is either isomorphic to the group $(\mathbb{R}^2,+)$ underlying $\mathbf{S}_{\tau,\nu}$ or isomorphic to the affine group underlying $\mathbf{R}_{\sigma,\tau}$ (see the examples above).

2. The DO $\mathbf{D}_{\alpha,\beta}$ is *separable* (possibly up to a phase factor) in the following sense: there exists an isomorphism $\psi:(\alpha,\beta)\to(\alpha',\beta')$ such that the parameter-transformed DO $\mathbf{D}_{\psi^{-1}(\alpha',\beta')}$ (briefly written as $\mathbf{D}_{\alpha',\beta'}$ hereafter) can be factored as [10]

$$\mathbf{D}_{\alpha',\beta'}=e^{j\mu(\alpha',\beta')}\,\mathbf{B}_{\beta'}\mathbf{A}_{\alpha'}\,. \tag{4.3.6}$$

Here, $\mathbf{A}_{\alpha'}$ and $\mathbf{B}_{\beta'}$ (termed *partial DOs*) are indexed by 1-D displacement parameters $\alpha'\in(\mathcal{A},\bullet)$ and $\beta'\in(\mathcal{B},*)$, respectively, where $(\mathcal{A},\bullet)$ and $(\mathcal{B},*)$ are *commutative* groups that are isomorphic to $(\mathbb{R},+)$. For example,

$$\mathbf{S}_{\tau,\nu}=\mathbf{F}_\nu\mathbf{T}_\tau \qquad\text{and}\qquad \mathbf{R}_{\sigma,\tau}=\mathbf{T}_\tau\mathbf{C}_\sigma\,,$$

with the time-shift operator $\mathbf{T}_\tau$, frequency-shift operator $\mathbf{F}_\nu$, and TF scaling operator $\mathbf{C}_\sigma$ defined as $(\mathbf{T}_\tau x)(t)=x(t-\tau)$, $(\mathbf{F}_\nu x)(t)=x(t)\,e^{j2\pi\nu t}$, and $(\mathbf{C}_\sigma x)(t)=\frac{1}{\sqrt{|\sigma|}}\,x\big(\frac{t}{\sigma}\big)$, respectively.

4.3.3 Covariant Signal Representations: Group Domain

We shall now discuss the construction of TF representations that are covariant to a given DO $\mathbf{D}_{\alpha,\beta}$. This construction is a two-stage process: first, we construct covariant signal representations that are functions of the displacement parameter (i.e., the group variables) (α,β). Subsequently (in Sections 4.3.4 and 4.3.5), we will convert these covariant (α,β) representations into covariant TF representations.

Covariance in the group domain. A linear (α,β) representation $L_x(\alpha,\beta)$ is called *covariant to a DO* $\mathbf{D}_{\alpha,\beta}$ if

$$L_{\mathbf{D}_{\alpha',\beta'}x}(\alpha,\beta) = e^{j\phi\left((\alpha,\beta)\circ(\alpha',\beta')^{-1};\alpha',\beta'\right)}\, L_x\big((\alpha,\beta)\circ(\alpha',\beta')^{-1}\big)\,, \tag{4.3.7}$$

for all signals $x(t)$ and for all $(\alpha,\beta),(\alpha',\beta')\in\mathcal{D}$ [10]. Similarly, a bilinear/quadratic (α,β) representation $B_{x,y}(\alpha,\beta)$ is called covariant to a DO $\mathbf{D}_{\alpha,\beta}$ if

$$B_{\mathbf{D}_{\alpha',\beta'}x,\,\mathbf{D}_{\alpha',\beta'}y}(\alpha,\beta) = B_{x,y}\big((\alpha,\beta)\circ(\alpha',\beta')^{-1}\big)\,, \tag{4.3.8}$$

for all signal pairs $x(t),y(t)$ and for all $(\alpha,\beta),(\alpha',\beta')\in\mathcal{D}$ [9,10]. Note that the "linear" covariance property (4.3.7) differs from the "bilinear" covariance property (4.3.8) in that it contains a phase factor.

For example, for $\mathbf{S}_{\tau,\nu}$ the covariance properties (4.3.7) and (4.3.8) become

$$\begin{aligned} L_{\mathbf{S}_{\tau',\nu'}x}(\tau,\nu) &= e^{-j2\pi(\nu-\nu')\tau'}\, L_x(\tau-\tau',\nu-\nu') \\ B_{\mathbf{S}_{\tau',\nu'}x,\,\mathbf{S}_{\tau',\nu'}y}(\tau,\nu) &= B_{x,y}(\tau-\tau',\nu-\nu')\,, \end{aligned}$$

which are seen to be identical to (4.3.2) and (4.3.4), respectively. For $\mathbf{R}_{\sigma,\tau}$, we obtain

$$\begin{aligned} L_{\mathbf{R}_{\sigma',\tau'}x}(\sigma,\tau) &= L_x\Big(\frac{\sigma}{\sigma'},\frac{\tau-\tau'}{\sigma'}\Big) \\ B_{\mathbf{R}_{\sigma',\tau'}x,\,\mathbf{R}_{\sigma',\tau'}y}(\sigma,\tau) &= B_{x,y}\Big(\frac{\sigma}{\sigma'},\frac{\tau-\tau'}{\sigma'}\Big). \end{aligned}$$

Construction of covariant (α,β) representations. It can be shown [10] that all linear (α,β) representations covariant to a DO $\mathbf{D}_{\alpha,\beta}$ as defined in (4.3.7) are given by

$$L_x(\alpha,\beta) = \langle x,\mathbf{D}_{\alpha,\beta}h\rangle = \int_{\mathcal{I}} x(t)\,(\mathbf{D}_{\alpha,\beta}h)^*(t)\,dt\,, \tag{4.3.9}$$

where $h(t)$ is an arbitrary function and $\mathcal{I}$ is the time interval on which $\mathbf{D}_{\alpha,\beta}$ is defined. Similarly, all bilinear/quadratic (α,β) representations covariant to a DO $\mathbf{D}_{\alpha,\beta}$ as defined in (4.3.8) are given by [9,10]

$$B_{x,y}(\alpha,\beta) = \langle x,\mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}\,y\rangle = \int_{\mathcal{I}}\int_{\mathcal{I}} x(t_1)\,y^*(t_2)\,\big[\mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}\big]^*(t_1,t_2)\,dt_1dt_2\,, \tag{4.3.10}$$

where $\mathbf{H}$ is an arbitrary linear operator and $\big[\mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}\big](t_1,t_2)$ denotes the kernel of the composed operator $\mathbf{D}_{\alpha,\beta}\mathbf{H}\mathbf{D}_{\alpha,\beta}^{-1}$. The equations (4.3.9) and (4.3.10) provide canonical expressions for all covariant linear and bilinear/quadratic (α,β) representations.

For example, for $\mathbf{S}_{\tau,\nu}$ these expressions yield

$$\begin{aligned} L_x(\tau,\nu) &= \int_{-\infty}^{\infty} x(t)\,h^*(t-\tau)\,e^{-j2\pi\nu t}\,dt \\ B_{x,y}(\tau,\nu) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1)\,y^*(t_2)\,h^*(t_1-\tau,t_2-\tau)\,e^{-j2\pi\nu(t_1-t_2)}\,dt_1dt_2\,, \end{aligned}$$

which are seen to be the short-time Fourier transform in (4.3.1) and Cohen's class in (4.3.3), respectively. Similarly, for $\mathbf{R}_{\sigma,\tau}$ we obtain time-scale versions of the wavelet transform and the affine class [1]:

$$L_x(\sigma,\tau) = \frac{1}{\sqrt{|\sigma|}} \int_{-\infty}^{\infty} x(t)\, h^*\Big(\frac{t-\tau}{\sigma}\Big)\, dt$$

$$B_{x,y}(\sigma,\tau) = \frac{1}{|\sigma|} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1)\, y^*(t_2)\, h^*\Big(\frac{t_1-\tau}{\sigma}, \frac{t_2-\tau}{\sigma}\Big)\, dt_1 dt_2\,.$$

4.3.4 The Displacement Function

The covariant (α,β) representations constructed above can be converted into covariant TF representations (ultimately, we are interested in TF representations and not in (α,β) representations). This conversion uses a mapping $(\alpha,\beta) \rightarrow (t,f)$ that is termed the *displacement function* (DF) since it describes the TF displacements performed by a DO $\mathbf{D}_{\alpha,\beta}$ in terms of TF coordinates [10].

The DF concept is based on the following reasoning. If a signal $x(t)$ is TF localized about some TF point (t_1,f_1), then the transformed ("displaced") signal $(\mathbf{D}_{\alpha,\beta}\, x)(t)$ will be localized about some other TF point (t_2,f_2) that depends on (t_1,f_1) and α,β. We can thus write

$$(t_2,f_2) = e_{\mathbf{D}}(t_1,f_1;\alpha,\beta)\,,$$

with some function $e_{\mathbf{D}}(t,f;\alpha,\beta)$ that will be called the *extended DF* of the DO $\mathbf{D}_{\alpha,\beta}$. For $\mathbf{S}_{\tau,\nu}$ and $\mathbf{R}_{\sigma,\tau}$, for example, it can easily be argued that the extended DF is given by

$$e_{\mathbf{S}}(t,f;\tau,\nu) = (t+\tau, f+\nu)\,, \qquad e_{\mathbf{R}}(t,f;\sigma,\tau) = \Big(\sigma t+\tau, \frac{f}{\sigma}\Big)\,. \tag{4.3.11}$$

Construction of the extended DF. In general, the extended DF cannot be found "by inspection," and therefore we need a systematic method for constructing the extended DF of a given DO $\mathbf{D}_{\alpha,\beta}$ [10]. The expression[1] $\mathbf{D}_{\alpha,\beta} = e^{j\mu(\alpha,\beta)}\, \mathbf{B}_\beta \mathbf{A}_\alpha$ (see (4.3.6)) states that $\mathbf{D}_{\alpha,\beta}$ is, up to a phase factor, the composition or series connection of $\mathbf{A}_\alpha$ and $\mathbf{B}_\beta$. Hence, $e_{\mathbf{D}}(t,f;\alpha,\beta)$ can be obtained by composing the extended DF of the partial DO $\mathbf{A}_\alpha$, $e_{\mathbf{A}}(t,f;\alpha)$, and the extended DF of the partial DO $\mathbf{B}_\beta$, $e_{\mathbf{B}}(t,f;\beta)$, according to

$$e_{\mathbf{D}}(t,f;\alpha,\beta) = e_{\mathbf{B}}\big(e_{\mathbf{A}}(t,f;\alpha);\,\beta\big)\,. \tag{4.3.12}$$

Using this expression, the task of constructing $e_{\mathbf{D}}(t,f;\alpha,\beta)$ reduces to the task of constructing $e_{\mathbf{A}}(t,f;\alpha)$ and $e_{\mathbf{B}}(t,f;\beta)$. We will explain the construction of $e_{\mathbf{A}}(t,f;\alpha)$ [10]; the construction of $e_{\mathbf{B}}(t,f;\beta)$ is of course analogous.

[1] For simplicity of notation, we assume that the parameter transformation $\psi : (\alpha,\beta) \rightarrow (\alpha',\beta')$ described in Section 4.3.2 has already been performed, and we write α,β instead of α',β'.

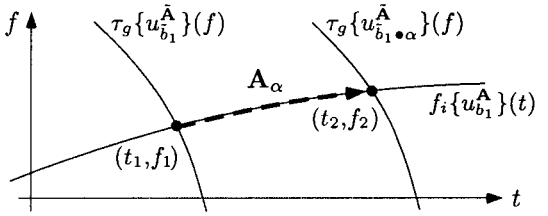


Fig. 4.3.1: Construction of the extended DF of $\mathbf{A}_\alpha$.

We first recall the "definition" of $e_{\mathbf{A}}(t,f;\alpha)$: If $x(t)$ is localized about (t_1,f_1), then $(\mathbf{A}_\alpha x)(t)$ will be localized about $(t_2,f_2) = e_{\mathbf{A}}(t_1,f_1;\alpha)$. In order to find (t_2,f_2), we consider the (generalized) eigenfunctions $u_b^{\mathbf{A}}(t)$ of $\mathbf{A}_\alpha$. These are defined by $\big(\mathbf{A}_\alpha u_b^{\mathbf{A}}\big)(t) \propto u_b^{\mathbf{A}}(t)$ and indexed by a parameter $b \in (\tilde{\mathcal{A}},\tilde{\bullet})$, where $(\tilde{\mathcal{A}},\tilde{\bullet})$ is again a commutative group isomorphic to $(\mathbb{R},+)$. The TF locus of $u_b^{\mathbf{A}}(t)$ is characterized by the instantaneous frequency $f_i\{u_b^{\mathbf{A}}\}(t)$ or the group delay $\tau_g\{u_b^{\mathbf{A}}\}(f)$, whichever exists.[2] Here, e.g., we assume existence of $f_i\{u_b^{\mathbf{A}}\}(t)$. Let us choose b_1 such that the TF curve defined by $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$ passes through (t_1,f_1), i.e.,

$$f_i\{u_{b_1}^{\mathbf{A}}\}(t_1) = f_1\,. \tag{4.3.13}$$

This is shown in Fig. 4.3.1. Now since $\big(\mathbf{A}_\alpha u_b^{\mathbf{A}}\big)(t) \propto u_b^{\mathbf{A}}(t)$, $\mathbf{A}_\alpha$ preserves the TF locus of $u_{b_1}^{\mathbf{A}}(t)$. Therefore, *under the action of* $\mathbf{A}_\alpha$, *all TF points on the curve* $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$—*including* (t_1,f_1)—*are mapped again onto TF points on* $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$. Hence, $(t_2,f_2) = e_{\mathbf{A}}(t_1,f_1;\alpha)$ must lie on $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$ (see Fig. 4.3.1), i.e., there must be

$$f_i\{u_{b_1}^{\mathbf{A}}\}(t_2) = f_2\,. \tag{4.3.14}$$

In order to find the exact position of (t_2,f_2) on the TF curve defined by $f_i\{u_{b_1}^{\mathbf{A}}\}(t)$, we use the fact that to any partial displacement operator $\mathbf{A}_\alpha$ there exists a *dual operator* $\tilde{\mathbf{A}}_{\tilde{\alpha}}$ with $\tilde{\alpha} \in (\tilde{\mathcal{A}},\tilde{\bullet})$ that is defined by the "almost commutation relation" $\tilde{\mathbf{A}}_{\tilde{\alpha}}\mathbf{A}_\alpha = e^{j2\pi\psi_{\mathcal{A}}(\alpha)\psi_{\tilde{\mathcal{A}}}(\tilde{\alpha})}\,\mathbf{A}_\alpha\tilde{\mathbf{A}}_{\tilde{\alpha}}$ [10,11]. For example, the dual operator of $\mathbf{T}_\tau$ is $\mathbf{F}_\nu$ and vice versa. Let $u_{\tilde{b}}^{\tilde{\mathbf{A}}}(t)$ with $\tilde{b} \in (\mathcal{A},\bullet)$ denote the (generalized) eigenfunctions of $\tilde{\mathbf{A}}_{\tilde{\alpha}}$ and assume, e.g., that the group delay $\tau_g\big\{u_{\tilde{b}}^{\tilde{\mathbf{A}}}\big\}(f)$ exists. Let us choose $\tilde{b}_1$ such that the TF curve defined by $\tau_g\big\{u_{\tilde{b}_1}^{\tilde{\mathbf{A}}}\big\}(f)$ passes through (t_1,f_1), i.e. (see Fig. 4.3.1)

$$\tau_g\big\{u_{\tilde{b}_1}^{\tilde{\mathbf{A}}}\big\}(f_1) = t_1\,. \tag{4.3.15}$$

Now assuming suitable parameterization of $u_{\tilde{b}}^{\tilde{\mathbf{A}}}(t)$, it can be shown [10] that

$$\big(\mathbf{A}_\alpha u_{\tilde{b}}^{\tilde{\mathbf{A}}}\big)(t) \propto u_{\tilde{b}\bullet\alpha}^{\tilde{\mathbf{A}}}(t) \qquad \text{for all } \alpha,\tilde{b} \in (\mathcal{A},\bullet)\,. \tag{4.3.16}$$

[2]The instantaneous frequency of a signal $x(t)$ is defined as $f_i\{x\}(t) = \frac{1}{2\pi}\frac{d}{dt}\arg\{x(t)\}$; it exists if $\arg\{x(t)\}$ is differentiable and $x(t) \neq 0$ almost everywhere. The group delay of $x(t)$ is defined as $\tau_g\{x\}(f) = -\frac{1}{2\pi}\frac{d}{df}\arg\{X(f)\}$ with $X(f) = \int_{-\infty}^{\infty} x(t)\,e^{-j2\pi ft}dt$; it exists if $\arg\{X(f)\}$ is differentiable and $X(f) \neq 0$ almost everywhere.

Thus, $\mathbf{A}_\alpha$ *maps all TF points on* $\tau_g\{u_{\tilde{b}_1}^{\tilde{\mathbf{A}}}\}(f)$*—including* (t_1, f_1)*—onto TF points on* $\tau_g\{u_{\tilde{b}_1 \bullet \alpha}^{\tilde{\mathbf{A}}}\}(f)$. So $(t_2, f_2) = e_{\mathbf{A}}(t_1, f_1; \alpha)$ must lie on $\tau_g\{u_{\tilde{b}_1 \bullet \alpha}^{\tilde{\mathbf{A}}}\}(f)$ (see Fig. 4.3.1), i.e.,

$$\tau_g\{u_{\tilde{b}_1 \bullet \alpha}^{\tilde{\mathbf{A}}}\}(f_2) = t_2 \,. \tag{4.3.17}$$

The construction of $e_{\mathbf{A}}$ can now be summarized as follows (see Fig. 4.3.1):

1. For any given (t_1, f_1), we calculate associated eigenfunction parameters $b_1 \in (\tilde{\mathcal{A}}, \tilde{\bullet})$ and $\tilde{b}_1 \in (\mathcal{A}, \bullet)$ as the solutions to (4.3.13) and (4.3.15), respectively:

$$f_i\{u_{b_1}^{\mathbf{A}}\}(t_1) = f_1 \,, \qquad \tau_g\{u_{\tilde{b}_1}^{\tilde{\mathbf{A}}}\}(f_1) = t_1 \,. \tag{4.3.18}$$

2. The extended DF $e_{\mathbf{A}}$ is defined by the identity $(t_2, f_2) \equiv e_{\mathbf{A}}(t_1, f_1; \alpha)$, where (t_2, f_2) is obtained as the solution to the system of equations (4.3.14), (4.3.17):

$$f_i\{u_{b_1}^{\mathbf{A}}\}(t_2) = f_2 \,, \qquad \tau_g\{u_{\tilde{b}_1 \bullet \alpha}^{\tilde{\mathbf{A}}}\}(f_2) = t_2 \,. \tag{4.3.19}$$

A similar construction of $e_{\mathbf{A}}$ can be used if, e.g., $\tau_g\{u_b^{\mathbf{A}}\}(f)$ and $f_i\{u_{\tilde{b}}^{\tilde{\mathbf{A}}}\}(t)$ exist instead of $f_i\{u_b^{\mathbf{A}}\}(t)$ and $\tau_g\{u_{\tilde{b}}^{\tilde{\mathbf{A}}}\}(f)$. An example for this case will be provided in Section 4.3.6.

After construction of the extended DF of $\mathbf{A}_\alpha$ as detailed above, the extended DF of $\mathbf{B}_\beta$ is constructed by means of an analogous procedure, and finally the extended DF of $\mathbf{D}_{\alpha,\beta}$ is obtained by composing $e_{\mathbf{A}}$ and $e_{\mathbf{B}}$ according to (4.3.12).

The DF. The above discussion has shown how to construct the *extended* DF $e_{\mathbf{D}}(t, f; \alpha, \beta)$. We go on to define the DF $d_{\mathbf{D}}(\alpha, \beta)$ by fixing t, f in $e_{\mathbf{D}}(t, f; \alpha, \beta)$:

$$d_{\mathbf{D}}(\alpha, \beta) \triangleq e_{\mathbf{D}}(t_0, f_0; \alpha, \beta) \qquad \text{with } t_0, f_0 \text{ arbitrary but fixed.}$$

The DF is a mapping $(\alpha, \beta) \to (t, f)$, i.e., from the displacement parameter (or group) domain to the TF domain. If the inverse DF $d_{\mathbf{D}}^{-1}(t, f)$ exists, then it can be shown [10] that the extended DF can be written as

$$e_{\mathbf{D}}(t, f; \alpha, \beta) = d_{\mathbf{D}}\big(d_{\mathbf{D}}^{-1}(t, f) \circ (\alpha, \beta)\big) \,. \tag{4.3.20}$$

Examples. Application of the construction explained above to the DOs $\mathbf{S}_{\tau,\nu}$ and $\mathbf{R}_{\sigma,\tau}$ yields the extended DFs $e_{\mathbf{S}}(t, f; \tau, \nu) = (t + \tau, f + \nu)$ and $e_{\mathbf{R}}(t, f; \sigma, \tau) = \big(\sigma t + \tau, \frac{f}{\sigma}\big)$. Note that this agrees with (4.3.11). Corresponding DFs are obtained by setting, e.g., $t = f = 0$ in $e_{\mathbf{S}}(t, f; \tau, \nu)$ and $t = 0$, $f = f_0 \neq 0$ in $e_{\mathbf{R}}(t, f; \sigma, \tau)$:

$$d_{\mathbf{S}}(\tau, \nu) = (\tau, \nu) \,, \qquad d_{\mathbf{R}}(\sigma, \tau) = \Big(\tau, \frac{f_0}{\sigma}\Big) \,.$$

A further example will be discussed in detail in Section 4.3.6.

4.3.5 Covariant Signal Representations: Time-Frequency Domain

In Section 4.3.3, we derived covariant linear and bilinear/quadratic signal representations that were functions of the displacement parameter (α, β). Using the inverse DF mapping $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$, these covariant (α, β) representations can now be converted into covariant TF representations (TFRs).

Covariance in the TF domain. A linear TFR $\tilde{L}_x(t, f)$ is called *covariant to a DO* $\mathbf{D}_{\alpha,\beta}$ if

$$\tilde{L}_{\mathbf{D}_{\alpha',\beta'}x}(t, f) = e^{j\phi\left(d_{\mathbf{D}}^{-1}(t,f)\circ(\alpha',\beta')^{-1};\alpha',\beta'\right)}\, \tilde{L}_x\left(e_{\mathbf{D}}(t, f; (\alpha', \beta')^{-1})\right), \tag{4.3.21}$$

for all $x(t)$ and for all $(\alpha', \beta') \in \mathcal{D}$ [10]. Similarly, a bilinear/quadratic TFR $\tilde{B}_{x,y}(t, f)$ is called covariant to a DO $\mathbf{D}_{\alpha,\beta}$ if

$$\tilde{B}_{\mathbf{D}_{\alpha',\beta'}x,\,\mathbf{D}_{\alpha',\beta'}y}(t, f) = \tilde{B}_{x,y}\left(e_{\mathbf{D}}(t, f; (\alpha', \beta')^{-1})\right), \tag{4.3.22}$$

for all $x(t), y(t)$ and for all $(\alpha', \beta') \in \mathcal{D}$ [10]. With (4.3.20), it is seen that these covariance properties are simply the (α, β)-domain (group-domain) covariance properties (4.3.7) and (4.3.8) with the transformation $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$.

For example, for $\mathbf{S}_{\tau,\nu}$ the covariance properties (4.3.21) and (4.3.22) become

$$\begin{aligned}
\tilde{L}_{\mathbf{S}_{\tau',\nu'}x}(t, f) &= e^{-j2\pi(f-\nu')\tau'}\, \tilde{L}_x(t-\tau', f-\nu') \\
\tilde{B}_{\mathbf{S}_{\tau',\nu'}x,\,\mathbf{S}_{\tau',\nu'}y}(t, f) &= \tilde{B}_{x,y}(t-\tau', f-\nu').
\end{aligned}$$

These relations are equivalent to (4.3.2) and (4.3.4), respectively. For $\mathbf{R}_{\sigma,\tau}$, we obtain

$$\begin{aligned}
\tilde{L}_{\mathbf{R}_{\sigma',\tau'}x}(t, f) &= \tilde{L}_x\Big(\frac{t-\tau'}{\sigma'}, \sigma' f\Big) \\
\tilde{B}_{\mathbf{R}_{\sigma',\tau'}x,\,\mathbf{R}_{\sigma',\tau'}y}(t, f) &= \tilde{B}_{x,y}\Big(\frac{t-\tau'}{\sigma'}, \sigma' f\Big).
\end{aligned}$$

Construction of covariant TF representations. It has been observed above that the TF covariance properties (4.3.21) and (4.3.22) are equivalent to the (α, β)-domain covariance properties (4.3.7) and (4.3.8), respectively, apart from the transformation $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$. From this equivalence, it follows that all covariant linear TFRs $\tilde{L}_x(t, f)$ are obtained from corresponding covariant linear (α, β) representations $L_x(\alpha, \beta)$ as given by (4.3.9) simply by setting $(\alpha, \beta) = d_{\mathbf{D}}^{-1}(t, f)$. Consequently, all linear TFRs covariant to a DO $\mathbf{D}_{\alpha,\beta}$ are given by

$$\tilde{L}_x(t, f) = L_x(\alpha, \beta)\Big|_{(\alpha,\beta)=d_{\mathbf{D}}^{-1}(t,f)} = \langle x, \mathbf{D}_{d_{\mathbf{D}}^{-1}(t,f)}h\rangle = \int_{\mathcal{I}} x(t')\,(\mathbf{D}_{d_{\mathbf{D}}^{-1}(t,f)}h)^*(t')\,dt', \tag{4.3.23}$$

where $h(t)$ is an arbitrary function and $\mathcal{I}$ is the time interval on which $\mathbf{D}_{\alpha,\beta}$ is defined. Similarly, all covariant bilinear/quadratic TFRs $\tilde{B}_{x,y}(t, f)$ are obtained

from corresponding covariant bilinear/quadratic (α,β) representations $B_{x,y}(\alpha,\beta)$ as given by (4.3.10) by setting $(\alpha,\beta) = d_{\mathbf{D}}^{-1}(t,f)$. Thus, all covariant bilinear/quadratic TFRs are given by

$$\tilde{B}_{x,y}(t,f) = B_{x,y}(\alpha,\beta)\Big|_{(\alpha,\beta)=d_{\mathbf{D}}^{-1}(t,f)} = \langle x, \mathbf{D}_{d_{\mathbf{D}}^{-1}(t,f)}\mathbf{H}\mathbf{D}^{-1}_{d_{\mathbf{D}}^{-1}(t,f)} y\rangle$$

$$= \int_{\mathcal{I}}\int_{\mathcal{I}} x(t_1)\, y^*(t_2)\, \left[\mathbf{D}_{d_{\mathbf{D}}^{-1}(t,f)}\mathbf{H}\mathbf{D}^{-1}_{d_{\mathbf{D}}^{-1}(t,f)}\right]^*(t_1,t_2)\, dt_1 dt_2\,, \qquad (4.3.24)$$

where $\mathbf{H}$ is an arbitrary linear operator. The equations (4.3.23) and (4.3.24) provide canonical expressions for all covariant linear and bilinear/quadratic TFRs [10].

For example, the classes of all linear and bilinear/quadratic TFRs covariant to $\mathbf{S}_{\tau,\nu}$ follow from (4.3.23) and (4.3.24) as

$$\tilde{L}_x(t,f) = \int_{-\infty}^{\infty} x(t')\, h^*(t'-t)\, e^{-j2\pi f t'}\, dt'$$

$$\tilde{B}_{x,y}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1)\, y^*(t_2)\, h^*(t_1-t, t_2-t)\, e^{-j2\pi f(t_1-t_2)}\, dt_1 dt_2\,;$$

they are seen to be the short-time Fourier transform in (4.3.1) and Cohen's class in (4.3.3), respectively. Similarly, the classes of all linear and bilinear/quadratic TFRs covariant to $\mathbf{R}_{\sigma,\tau}$ are obtained as

$$\tilde{L}_x(t,f) = \sqrt{\left|\frac{f}{f_0}\right|}\int_{-\infty}^{\infty} x(t')\, h^*\left(\frac{f}{f_0}\,(t'-t)\right) dt'$$

$$\tilde{B}_{x,y}(t,f) = \left|\frac{f}{f_0}\right|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(t_1)\, y^*(t_2)\, h^*\left(\frac{f}{f_0}\,(t_1-t), \frac{f}{f_0}\,(t_2-t)\right) dt_1 dt_2\,;$$

they are TF versions of the wavelet transform [2] and the affine class [1,5]. Thus, the short-time Fourier transform, the wavelet transform, Cohen's class, and the affine class have all been obtained by means of the systematic construction provided by covariance theory.

4.3.6 Example: Hyperbolic Wavelet Transform and Hyperbolic Class

So far, we have considered the elementary DOs $\mathbf{S}_{\tau,\nu}$ and $\mathbf{R}_{\sigma,\tau}$ as illustrative examples. Let us now apply our covariance theory to a situation that is a little less elementary. We consider the DO $\mathbf{V}_{\sigma,\gamma}$ defined by

$$(\mathbf{V}_{\sigma,\gamma}x)(t) = (\mathbf{G}_\gamma \mathbf{C}_\sigma x)(t) = \frac{1}{\sqrt{\sigma}}\, x\left(\frac{t}{\sigma}\right) e^{j2\pi\gamma \ln(t/t_0)}\,, \qquad t>0,\ \sigma>0,\ \gamma\in\mathbb{R}\,.$$

Here, $\mathbf{C}_\sigma$ is the TF scaling operator defined by $(\mathbf{C}_\sigma x)(t) = \frac{1}{\sqrt{\sigma}}\, x\left(\frac{t}{\sigma}\right)$ with $\sigma > 0$; furthermore, $\mathbf{G}_\gamma$ is the *hyperbolic frequency-shift operator* defined by

$(\mathbf{G}_\gamma x)(t) = x(t)\, e^{j2\pi\gamma \ln(t/t_0)}$ with $t_0 > 0$ arbitrary but fixed. Note that $\mathbf{C}_\sigma$ and $\mathbf{G}_\gamma$ are dual operators since $\mathbf{G}_\gamma \mathbf{C}_\sigma = e^{j2\pi\gamma \ln\sigma}\, \mathbf{C}_\sigma \mathbf{G}_\gamma$. Comparison of the relation $\mathbf{V}_{\sigma_2,\gamma_2}\mathbf{V}_{\sigma_1,\gamma_1} = e^{-j2\pi\gamma_1 \ln\sigma_2}\, \mathbf{V}_{\sigma_1\sigma_2,\gamma_1+\gamma_2}$ with (4.3.5) shows that (σ,γ) belongs to the commutative group $(\mathbb{R}^+ \times \mathbb{R}, \circ)$ with group law $(\sigma_1,\gamma_1)\circ(\sigma_2,\gamma_2) = (\sigma_1\sigma_2, \gamma_1+\gamma_2)$, identity element $(1,0)$, and inverse elements $(\sigma,\gamma)^{-1} = (1/\sigma, -\gamma)$. This group is isomorphic to the group $(\mathbb{R}^2, +)$. Furthermore, we see that the cocycle phase function is given by $\phi(\sigma_1,\gamma_1;\sigma_2,\gamma_2) = -2\pi\gamma_1 \ln\sigma_2$.

We now begin our construction of TFRs covariant to the DO $\mathbf{V}_{\sigma,\gamma}$. In the (σ,γ) domain, the covariance properties (4.3.7) and (4.3.8) read as

$$L_{\mathbf{V}_{\sigma',\gamma'}x}(\sigma,\gamma) = e^{-j2\pi(\gamma-\gamma')\ln\sigma'}\, L_x\Big(\frac{\sigma}{\sigma'}, \gamma-\gamma'\Big)$$
$$B_{\mathbf{V}_{\sigma',\gamma'}x,\mathbf{V}_{\sigma',\gamma'}y}(\sigma,\gamma) = B_{x,y}\Big(\frac{\sigma}{\sigma'}, \gamma-\gamma'\Big),$$

and the covariant linear and bilinear/quadratic (σ,γ) representations are obtained from (4.3.9) and (4.3.10) as

$$L_x(\sigma,\gamma) = \frac{1}{\sqrt{\sigma}} \int_0^\infty x(t)\, h^*\Big(\frac{t}{\sigma}\Big)\, e^{-j2\pi\gamma\ln(t/t_0)}\, dt\,, \qquad \sigma > 0$$
$$B_{x,y}(\sigma,\gamma) = \frac{1}{\sigma} \int_0^\infty \int_0^\infty x(t_1)\, y^*(t_2)\, h^*\Big(\frac{t_1}{\sigma}, \frac{t_2}{\sigma}\Big)\, e^{-j2\pi\gamma\ln(t_1/t_2)}\, dt_1 dt_2\,, \qquad \sigma > 0\,.$$

Next, we construct the DF of $\mathbf{V}_{\sigma,\gamma} = \mathbf{G}_\gamma \mathbf{C}_\sigma$. We first consider the extended DF of $\mathbf{C}_\sigma$. Although clearly $e_{\mathbf{C}}(t,f;\sigma) = (\sigma t, f/\sigma)$, we shall derive $e_{\mathbf{C}}(t,f;\sigma)$ using the systematic construction procedure from Section 4.3.4. The eigenfunctions of $\mathbf{C}_\sigma$ are $u_\gamma^{\mathbf{C}}(t) = \frac{1}{\sqrt{t}}\, e^{j2\pi\gamma\ln(t/t_0)}$, with instantaneous frequency $f_i\{u_\gamma^{\mathbf{C}}\}(t) = \gamma/t$. The eigenfunctions of the dual operator $\mathbf{G}_\gamma$ are given by $u_s^{\mathbf{G}}(t) = \delta(t-s)$, with group delay $\tau_g\{u_s^{\mathbf{G}}\}(f) \equiv s$. (It can be verified that (4.3.16) is satisfied: $\big(\mathbf{C}_\sigma u_s^{\mathbf{G}}\big)(t) = \sqrt{\sigma}\,\delta(t-\sigma s) \propto u_{\sigma s}^{\mathbf{G}}(t)$.) Thus, (4.3.18) becomes $\gamma_1/t_1 = f_1$ and $s_1 = t_1$, which yields the eigenfunction parameters associated to the TF point (t_1,f_1) as $\gamma_1 = t_1 f_1$, $s_1 = t_1$. Similarly, (4.3.19) becomes $\gamma_1/t_2 = f_2$ and $s_1\sigma = t_2$, which yields $t_2 = s_1\sigma = \sigma t_1$ and $f_2 = \gamma_1/(s_1\sigma) = f_1/\sigma$. Hence, the extended DF of $\mathbf{C}_\sigma$ is finally obtained as

$$e_{\mathbf{C}}(t_1,f_1;\sigma) = (t_2,f_2) = \Big(\sigma t_1, \frac{f_1}{\sigma}\Big).$$

The extended DF of $\mathbf{G}_\gamma$ is obtained similarly. The eigenfunctions of $\mathbf{G}_\gamma$ are $u_s^{\mathbf{G}}(t) = \delta(t-s)$, with group delay $\tau_g\{u_s^{\mathbf{G}}\}(f) \equiv s$. The eigenfunctions of the dual operator $\mathbf{C}_\sigma$ are $u_\gamma^{\mathbf{C}}(t) = \frac{1}{\sqrt{t}}\, e^{j2\pi\gamma\ln(t/t_0)}$, with instantaneous frequency $f_i\{u_\gamma^{\mathbf{C}}\}(t) = \gamma/t$. (We verify that (4.3.16) is satisfied: $\big(\mathbf{G}_{\gamma'} u_\gamma^{\mathbf{C}}\big)(t) = \frac{1}{\sqrt{t}}\, e^{j2\pi\gamma\ln(t/t_0)}\, e^{j2\pi\gamma'\ln(t/t_0)} \propto u_{\gamma+\gamma'}^{\mathbf{C}}(t)$.) Thus, (4.3.18) (with the roles of instantaneous frequency and group delay as well as time and frequency interchanged) becomes $s_1 = t_1$ and $\gamma_1/t_1 = f_1$, which yields the eigenfunction parameters $s_1 = t_1$, $\gamma_1 = t_1 f_1$. Similarly, (4.3.19) (with the same interchange of roles) becomes $s_1 = t_2$ and $(\gamma_1+\gamma)/t_2 = f_2$, whence

$t_2 = s_1 = t_1$ and $f_2 = (\gamma_1 + \gamma)/s_1 = f_1 + \gamma/t_1$. Hence, the extended DF of $\mathbf{G}_\gamma$ is obtained as

$$e_{\mathbf{G}}(t_1, f_1; \gamma) = (t_2, f_2) = \Big(t_1, f_1 + \frac{\gamma}{t_1}\Big).$$

The extended DF of $\mathbf{V}_{\sigma,\gamma} = \mathbf{G}_\gamma \mathbf{C}_\sigma$ can now be calculated by composing $e_{\mathbf{C}}(t, f; \sigma)$ and $e_{\mathbf{G}}(t, f; \gamma)$ according to (4.3.12), which yields

$$e_{\mathbf{V}}(t, f; \sigma, \gamma) = e_{\mathbf{G}}\big(e_{\mathbf{C}}(t, f; \sigma); \gamma\big) = \Big(\sigma t, \frac{f + \gamma/t}{\sigma}\Big).$$

Finally, the DF (and inverse DF) of $\mathbf{V}_{\sigma,\gamma}$ follow upon setting $t = t_0 > 0$ and $f = 0$:

$$d_{\mathbf{V}}(\sigma, \gamma) = e_{\mathbf{V}}(t_0, 0; \sigma, \gamma) = \Big(\sigma t_0, \frac{\gamma}{\sigma t_0}\Big), \qquad d_{\mathbf{V}}^{-1}(t, f) = \Big(\frac{t}{t_0}, tf\Big).$$

With the DF at our disposal, we are ready to pass from the (σ, γ) domain into the TF domain. The TF covariance properties (4.3.21) and (4.3.22) become

$$\tilde{L}_{\mathbf{V}_{\sigma',\gamma'}x}(t, f) = e^{-j2\pi(tf - \gamma')\ln\sigma'}\, \tilde{L}_x\Big(\frac{t}{\sigma'}, \sigma'\Big(f - \frac{\gamma'}{t}\Big)\Big)$$

$$\tilde{B}_{\mathbf{V}_{\sigma',\gamma'}x, \mathbf{V}_{\sigma',\gamma'}y}(t, f) = \tilde{B}_{x,y}\Big(\frac{t}{\sigma'}, \sigma'\Big(f - \frac{\gamma'}{t}\Big)\Big),$$

and the covariant linear and bilinear/quadratic TFRs are obtained from (4.3.23) and (4.3.24) as

$$\tilde{L}_x(t, f) = \sqrt{\frac{t_0}{t}} \int_0^\infty x(t')\, h^*\Big(t_0 \frac{t'}{t}\Big)\, e^{-j2\pi t f \ln(t'/t_0)}\, dt', \quad t > 0$$

$$\tilde{B}_{x,y}(t, f) = \frac{t_0}{t} \int_0^\infty \int_0^\infty x(t_1)\, y^*(t_2)\, h^*\Big(t_0 \frac{t_1}{t}, t_0 \frac{t_2}{t}\Big)\, e^{-j2\pi t f \ln(t_1/t_2)}\, dt_1 dt_2, \quad t > 0.$$

These TFRs are analogous to (respectively) the hyperbolic wavelet transform and the hyperbolic class introduced in [4], the difference being that in [4] the hyperbolic time-shift operator was used instead of the hyperbolic frequency-shift operator $\mathbf{G}_\gamma$.

4.3.7 Summary and Conclusions

Time-frequency representations (TFRs) that are covariant to practically important signal transformations—like time and frequency shifts, time-frequency scaling (dilation/compression), or dispersive time and frequency shifts—are of great relevance in applications. We have presented a unified and coherent *covariance theory of TFRs* that allows the systematic construction of TFRs covariant to two-parameter transformations. We note that a much more detailed and mathematically rigorous discussion with many additional references is provided in [10], where also the extension to groups not isomorphic to $(\mathbb{R}, +)$ is outlined. Furthermore, relations of covariance theory with the principle of unitary equivalence are discussed in [10, 12] (cf. also Article 4.5).

References

[1] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[2] S. G. Mallat, *A Wavelet Tour of Signal Processing.* San Diego: Academic Press, 1st ed., 1998.

[3] 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 Processing*, vol. 47, pp. 3067–3083, November 1999.

[4] 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 Processing*, vol. 41, pp. 3425–3444, December 1993. Special Issue on Wavelets and Signal Processing.

[5] J. Bertrand and P. Bertrand, "Affine time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 5, pp. 118–140, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[6] F. Hlawatsch, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels, "The hyperbolic class of quadratic time-frequency representations—Part II: Subclasses, intersection with the affine and power classes, regularity, and unitarity," *IEEE Trans. Signal Processing*, vol. 45, pp. 303–315, February 1997.

[7] F. Hlawatsch and H. Bölcskei, "Unified theory of displacement-covariant time-frequency analysis," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 524–527, Philadelphia, PA, 25–28 October 1994.

[8] F. Hlawatsch and T. Twaroch, "Covariant (α, β), time-frequency, and (a, b) representations," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 437–440, Paris, 18–21 June 1996.

[9] A. M. Sayeed and D. L. Jones, "A canonical covariance-based method for generalized joint signal representations," *IEEE Signal Processing Letters*, vol. 3, pp. 121–123, April 1996.

[10] F. Hlawatsch, G. Tauböck, and T. Twaroch, "Covariant time-frequency analysis," in *Wavelets and Signal Processing* (L. Debnath, ed.), Boston: Birkhäuser, 2003.

[11] A. M. Sayeed and D. L. Jones, "Integral transforms covariant to unitary operators and their implications for joint signal representations," *IEEE Trans. Signal Processing*, vol. 44, pp. 1365–1377, June 1996.

[12] R. G. Baraniuk, "Covariant time-frequency representations through unitary equivalence," *IEEE Signal Processing Letters*, vol. 3, pp. 79–81, March 1996.

4.4 UNCERTAINTY IN TIME-FREQUENCY ANALYSIS [0]

In 1946, Gabor [1] introduced the Heisenberg uncertainty relations to the signal processing community, with the simple but powerful inequality

$$\sigma_t \cdot \sigma_f \geq \frac{1}{4\pi}, \tag{4.4.1}$$

where σ_t and σ_f are the time and frequency standard deviations of the signal $s(t)$. That is (we will assume unit energy signals, centered at zero time and frequency):

$$\sigma_t = \int_{-\infty}^{\infty} t^2 |s(t)|^2 \, dt, \tag{4.4.2}$$

$$\sigma_f = \int_{-\infty}^{\infty} f^2 |S(f)|^2 \, df, \tag{4.4.3}$$

$S(f)$ being the Fourier Transform of $s(t)$. The impact of this result in the field of spectrum estimation was immediate, since it seems to imply the existence of bounds on the achievable frequency resolution when analyzing finite length segments of data. Satisfaction of (4.4.1) implies that, for any given (non-zero) σ_t, there will be a minimum σ_f. However, some care must be exercised when following this line of reasoning, since standard deviations are not, in general, acceptable measures of broadness, and do not bear an easy relation with the concept of resolution. As an example, consider a long rectangular window. This window has a spectrum with a very thin main lobe and is, thus, capable of high frequency resolution. However, the standard deviation of its power spectrum is ∞, as can be seen using $S(f) = \sin(af)/(af)$ in (4.4.3). This example shows that (4.4.1), being a limit on standard deviations, does not, *de per si*, necessarily limit the achievable frequency resolution. High resolution waveforms may have large standard deviations. Other limitations of the use of standard deviations as measures of broadness can be found in [2] and [3]. We will introduce a different measure of broadness, which avoids these limitations.

With the development of joint time-frequency analysis, the issue of uncertainty and/or frequency resolution has to be rethought. In time-frequency analysis, we often deal with a single bivariate function $\rho_s(t, f)$, desirably representing the distribution of energy in the time-frequency plane. For acceptance of a given $\rho_s(t, f)$ as a valid time-frequency distribution (TFD), one often requires satisfaction of the time and frequency marginal conditions [see p. 33]:

$$\int_{-\infty}^{\infty} \rho_s(t, f) \, df = |s(t)|^2 \tag{4.4.4}$$

$$\int_{-\infty}^{\infty} \rho_s(t, f) \, dt = |S(f)|^2. \tag{4.4.5}$$

The Heisenberg-Gabor relations (4.4.1) will imply limits to the joint behavior of the marginals, $|s(t)|^2$ and $|S(f)|^2$. But they do not imply limits on $\rho_s(t, f)$ at other time-frequency locations [4]. Do such limits exist? Are there uncertainty

[0]Authors: **Paulo M. Oliveira**, Escola Naval, DAEl, Base Naval de Lisboa, Alfeite, 2800 Almada, Portugal (pmonica@mail.telepac.pt) and **Victor Barroso**, Instituto Superior Técnico ISR/DEEC, Torre Norte, Piso 7 Av. Rovisco Pais, 1049-001, Lisboa, Portugal (vab@isr.ist.utl.pt).
Reviewers: M. G. Amin and X.-G. Xia.

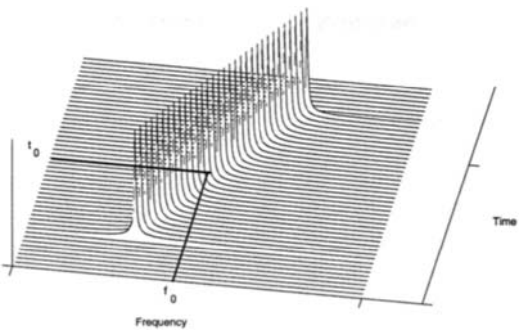


Fig. 4.4.1: Example of a thin $\rho_s(t,f)$ with broad marginals.

limits *within* the time-frequency plane, or can we achieve infinite time-frequency resolution, under the constraint that the marginals must satisfy (4.4.1)? To fully appreciate the scope of the question, let us consider an hypothetical $\rho_s(t,f)$, constituted by an infinitely narrow time-frequency ridge (see Fig. 4.4.1). Even though the marginals are broad, and, for any given σ_t can always be made to satisfy (4.4.1) by adjusting the frequency slope of the ridge, such a $\rho_s(t,f)$ possesses an infinite time-frequency *concentration* capability, which allows us to know the exact frequency that the signal occupied at any precise time (f_0 and t_0, in the example figure). The question to be answered thus becomes: is such a $\rho_s(t,f)$ physically acceptable? Is infinite time-frequency *concentration* capability an achievable goal? And is infinite time-frequency *resolution* (the ability to separate two closely spaced components) also an achievable goal? Or are there limits (certainly other than (4.4.1)) to the concentration and/or resolution of any physically acceptable $\rho_s(t,f)$?

4.4.1 The Time-Frequency Plane

Several attempts have been made to obtain uncertainty relations applicable to regions of the time-frequency plane other than the marginals. Each one of these attempts starts by defining a concentration/dispersion measure in the time-frequency plane (Δ_{tf}). Examples of such measures are [5], [6] (for arbitrary T):

$$\Delta_{tf} = \iint \left(\frac{t^2}{T^2} + T^2 f^2\right) \rho_s(t,f)\, dt\, df, \tag{4.4.6}$$

$$\Delta_{tf} = \iint \left(1 - e^{-(t^2/T^2 + T^2 f^2)}\right) \rho_s(t,f)\, dt\, df. \tag{4.4.7}$$

$$\Delta_{tf} = \iint t^2 f^2\, \rho_s(t,f)\, dt\, df, \tag{4.4.8}$$

$$\Delta_{tf} = \iint t^2 (f - f_i(t))^2\, \rho_s(t,f)\, dt\, df. \tag{4.4.9}$$

All these measures turn out to imply a particular set of uncertainty relations when $\rho_s(t,f)$ is taken as being the Wigner-Ville distribution (p. 33) or, more generally, one of the bilinear distributions from Cohen's class (p. 68n and refs. [5,6]). A different and promising approach can also be found in [6], where the Slepian-Landau-Pollack approach is extended to ellipsoidal regions (axis-aligned) in the time-frequency plane. However, careful analysis of these measures will show that none of them is still an appropriate measure of the concentration capabilities of a general $\rho_s(t,f)$. Let us take (4.4.6), as an example. When applied to the Wigner-Ville distribution, it implies that $\Delta_{tf} \geq \frac{1}{2\pi}$ [6]. However, this distribution is known to generate infinitely thin ridges in the time-frequency plane for linear chirps. This shows that a given $\rho_s(t,f)$ can have the infinite time-frequency location capabilities of Fig. 4.4.1, even if it is lower bounded by measure (4.4.6). Furthermore, the known uncertainty relations implied by (4.4.6)-(4.4.9) have been obtained for particular time-frequency distributions, and cannot be interpreted as general physical laws.

To avoid this restriction of the results to the scope of the particular time-frequency tool being used to derive them, we will take a different approach, based on the general concept of spectral information. This approach leads to general results that are not specific to any one time-frequency distribution.

4.4.2 Information and Spectral Estimation

In the following discussion, let us assume that we have a second degree *ensemble view*. That is: we will assume that, by observing the signal at time t_1, we apprehend the expected value of the ensemble of signals ($E\{s(t_1)\}$), and not only the value of one particular realization; by also observing the signal at time t_2, we now not only apprehend the value of $E\{s(t_2)\}$, but also $R_s(t_1,t_2) = E\{s(t_1)s^*(t_2)\}$. Since we are, under this assumption, directly observing expected values, and not mere realizations of the process, we may, in what follows, safely ignore the practical aspects of real estimators, such as variance. This idealization will enable us to focus only on the aspects which are, for our purposes, most relevant.

• **Stationary signals.** From this idealized point of view, let us now consider the estimation of the power spectrum of a stationary signal. To perform the estimate, one must extract the needed information out of the signal. But how much information must we collect? How much observation time do we need? At first, increasing the observation time will provide better and better estimates, in a process converging to the true power spectrum. But, after convergence (assuming it ever happens), will further increases in observation time provide more information about the power spectrum? It clearly doesn't. Once this convergence process is essentially completed, no further observation time is needed, because no further information is required. The total amount of information that must be obtained for proper spectral estimation will be referred to as the *spectral complexity of the signal* (C_s). The point to be made here is that, if the observation time available to us is less than the time needed to reach C_s, our spectral estimate will be degraded, since we will be missing part of the information needed for a correct estimate. On

the other hand, if the observation time available is greater than the time needed to reach C_s, the last part of the signal will be informationless. Denoting by $D(t)$ the density of information contained in the signal (the function $D(t)$ depends on the type of estimator being used, an issue which we will not pursue here), and by $I(t)$ the amount of collected information, the collection procedure can be summarized as follows (assuming that the observation started at t_0):

$$I(t) = \int_{t_0}^{t} D(\xi)\, d\xi \leq \int_{t_0}^{t_0+\tau_R} D(\xi)\, d\xi = C_s, \tag{4.4.10}$$

τ_R being the time needed to reach complexity. This immediately raises the question of how to quantify C_s. There are good reasons (some of which will become apparent) to use a measure of spectral narrowness as a measure of spectral complexity. While any acceptable measure of narrowness is basically equivalent, leading to a rescaled version of the quantity, we must be careful to avoid the pitfalls of global measures of width such as the standard deviation, as previously discussed. Let us then define

$$C_s = \sqrt{\max_f \left(-\frac{\partial^2}{\partial f^2} P_s(f) \right)}, \tag{4.4.11}$$

where $P_s(f) = E\{|S(f)|^2\}$ is the power spectrum density of the signal. This measure, based on the Hessian of the power spectrum, relates directly to the idea of frequency resolution, and is closely related with the definition of Fisher information.

From (4.4.11), the spectral complexity of a sinusoid is infinite. Even though an impulsive spectrum may seem simple, we must note that the perfect localization capability needed for a proper estimate does imply an infinite amount of information. Hence the infinite C_s. Also, we must consider the fact that the spectrum of a pure sinusoid will always become narrower and narrower with increasing observation time, without ever stabilizing. Additional observation time will always improve our spectral estimate and, thus, provide new information. Complexity will never be reached for finite observation times. This is thus the high end of spectral complexity, where all observation time becomes useful and brings additional information. In the low end, we have white noise. An instantaneous ensemble observation fully characterizes its very low complexity (zero, in fact) power spectrum. Further observation of the noise ensemble will not contribute with any new information concerning its spectrum.

We must now address the next logical question: how do we determine τ_R? The amount of information needed to estimate the power spectrum is clearly the same amount of information needed to estimate its inverse Fourier Transform, the autocorrelation function. Hence, we only need to observe the signal for the amount of time needed to determine all (relevant) lags of its autocorrelation function. The time to reach complexity is thus the time support of the autocorrelation function. This is a very gratifying conclusion, since the spectral complexity (4.4.11) and the time support of the autocorrelation function $R_s(\tau)$ can be shown to be, in fact,

directly related to each other through yet another "uncertainty relation":

$$\frac{1}{C_s} \cdot \sigma_R \geq \frac{1}{2\pi}, \tag{4.4.12}$$

where $\sigma_R^2 = \int_{-\infty}^{\infty} \tau^2 \, |R_s(\tau)| \, d\tau$. In general, signals with narrowband components (and, thus, of high spectral complexity) require a high collection time τ_R (the limit is again a phase randomized sinusoid, known to have an autocorrelation function with infinite time support [7, p. 33]). Signals without narrowband components (hence, of low spectral complexity) have small collection times (in the limiting case of white noise, the autocorrelation function has zero time support). In any case, observing the ensemble for periods longer than τ_R is not useful, since no additional information about the power spectrum will be obtained.

• **Non-stationary signals.** Let us now consider non-stationary signals, and assume that we desire to estimate the power spectrum of a non-stationary signal at time t_1. This instantaneous spectrum will have a given amount of spectral complexity ($C_s^{t_1}$), and to properly estimate it, we need to collect this very same amount of information about the spectrum (or the autocorrelation function) at time t_1. But to represent time t_1, all we have is $s(t_1)$ itself, and no finite amount of spectral information can be extracted from an instantaneous value of the signal, since it would imply the acceptance of an infinite information density. On the other hand, information collected at times other than t_1 will only be useful if and only if it is correlated with the spectral information at time t_1. In the previous case of stationary signals, the spectral information at any time was totally correlated with the spectral information at any other time. In the non-stationary case, however, we must weight the information collected at times t other than t_1 with the non-unitary correlation factor that determines how useful the collected information is for estimates at time t_1 (we will denote this weighting factor by $u^{t_1}(t - t_1)$). We now have to distinguish between useful past and future ($u^{t_1}(t - t_1) \neq 0$), and non-useful past and future ($u^{t_1}(t - t_1) \approx 0$). The collection procedure (4.4.10), using superscripts to denote the particular time for which the spectrum estimate is intended, becomes:

$$\begin{aligned} I^{t_1}(t) &= \int_{t_0}^{t} D(\xi) \, u^{t_1}(\xi - t_1) \, d\xi, \qquad \text{with} \\ I^{t_1}(t) &\leq C_s^{t_1} = \int_{-\infty}^{\infty} D(\xi) \, u^{t_1}(\xi - t_1) \, d\xi. \end{aligned} \tag{4.4.13}$$

The factor $u^{t_1}(t - t_1)$ is thus just formalizing the fact that observing a non-stationary signal *now* does not tell us much concerning the spectrum of the signal a fortnight ago, and may even be misleading, and decrease our state of knowledge (in the case of negative $u^{t_1}(t - t_1)$). The exception lies, of course, in the case of signals with deterministic and known frequency dynamics, since in these cases the information collected at any time can always be made useful, by taking the dynamics into proper account. Knowledge of the frequency dynamics thus makes

the utility factor constant and unitary, bringing the case of non-stationary signals to the very same situation one encounters with stationary signals.

As an example, consider the estimation of the power spectrum of a constant amplitude linear chirp (let us assume a positive chirping rate), with uniformly distributed random phase:

$$s(t) = e^{j(\alpha t^2 + \theta)}.$$

Its autocorrelation function is easily seen to be:

$$R_s(t, t-\tau) = R_s(t, \tau) = e^{-j(\alpha\tau^2 - 2\alpha\tau t)}.$$

To determine $u(t)$, we can now determine how correlated are the autocorrelation functions at times t_1 and t_2, with $t_2 > t_1$. Due to the infinite energy of these autocorrelation functions, in the computation of their correlation factor we will limit the integration region to an arbitrarily large region ($-l$ to l) centered at the zeroth lag. That is:

$$\begin{aligned} u^{t_1}(t_2 - t_1) &= R_R(t_2, t_1, l) = \\ &= \frac{1}{2l} \int_{-l}^{l} R_s(t_2, \xi)\, R_s^*(t_1, \xi)\, d\xi. \end{aligned}$$

This means that, in our case,

$$u^{t_1}(t_2 - t_1) = \frac{\sin\left(2\alpha l(t_2 - t_1)\right)}{2\alpha l(t_2 - t_1)}. \tag{4.4.14}$$

The inclusion of $u^{t_1}(t_2 - t_1)$ in (4.4.13) (in this case, a sinc function) will limit the amount of collectible information relative to time t_1 and, thus, will set an upper bound on the achievable spectral complexity and, hence, the achievable frequency resolution.

This is thus the answer we have been trying to obtain. There are limits to the achievable frequency resolution within the Time-Frequency plane, due to the fact that the period of time during which one can collect information concerning the spectrum at a given time is decreasing as the spectral dynamics increase. For increasing dynamics (α, in our example) the useful neighborhood (and, hence, the amount of useful information) will continuously decrease, and so will the achievable spectral complexity and, consequently, the achievable spectral resolution. This namely means that the faster a chirp moves, the broader it becomes in the t-f plane. This predicted broadening of the power spectrum (and consequent decrease in resolution capabilities) with the increase of the chirping rate is, in fact, observed in most bilinear TFDs. To illustrate it, we computed the Margenau-Hill distribution of a cubic chirp. The results can be seen in Figure 4.4.2, were the broadening of the main lobe of this Time-Frequency distribution with the increase in chirp rate is clearly apparent.

Let us now try to determine, for the case of the chirp, the best observation time. From (4.4.14), we see that the best strategy is to limit the observation time to the

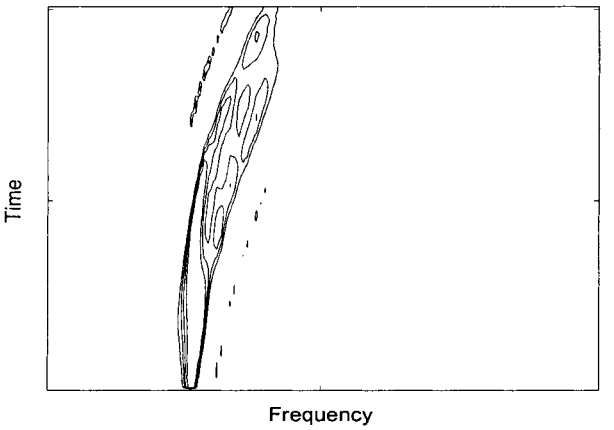


Fig. 4.4.2: Cubic Chirp. Margenau-Hill distribution.

main lobe of the sinc function, since we will then avoid negative values of $u^{t_1}(t-t_1)$. That is, observe the signal between $t_1 - \tau$ and $t_1 + \tau$, where $\tau = \pi/2\alpha l$. But this implies that τ is the maximum lag of the observed autocorrelation function and, thus, $l = \tau$. Thus we conclude that the best observation time for a linear chirp is

$$\tau = \sqrt{\frac{\pi}{2\alpha}} = \sqrt{\frac{1}{2\frac{\partial}{\partial t}\left[f_i(t)\right]}}, \tag{4.4.15}$$

$f_i(t)$ being the chirp instantaneous frequency (in this particular case, we may safely identify the concept of instantaneous frequency with the derivative of the phase function). With hindsight, it is now interesting to observe that: *(a)* (4.4.15) is the optimum observation time for short-time Fourier analysis of a linear chirp; *(b)* (4.4.15) is the effective time support of the optimum data independent smoothing window to use with the Wigner-Ville distribution [8]; *(c)* it is also basically the same quantity defined by Rihaczek as the signal's "relaxation time" [9]. These separate results can now easily be understood as particular manifestations of (4.4.15).

A last comment must be made, concerning the use of models. Assuming a model for the frequency dynamics, such as the linear model implicit in the Wigner-Ville Distribution (or higher order models in the Polynomial Wigner-Ville Distribution [10]), is an attempt to increase the size of what we called the "useful neighborhood", by projecting all collected spectral information to the time of interest, according to the spectral dynamics of the assumed model. If, by inspiration or mere chance, the assumed model is, in fact, the correct one, we will overcome the limits imposed by the non-stationarity, and fall within the traditional, stationary uncertainty relations, as previously discussed. If, on the other hand, the model is incorrect, we must be prepared to pay for the wrong assumption. We may have apparently improved our frequency resolution, but we must pay for it with bias and artifacts. Another, more

subtle, type of assumption, is made whenever we arbitrarily decide that the "true" distribution is the one maximizing some chosen criteria. It may or may not be a sensible, supported assumption. It is an assumption, nonetheless. It may get us a better frequency resolution, but only if verified by the signal under analysis. This is exactly the type of trade-off one finds in parametric spectrum estimation.

4.4.3 Summary and Conclusions

This article addresses the issue of determining if there are lower bounds to the achievable time-frequency resolution within the Time-Frequency plane. The approach based on the informational aspects of the estimation achieves results independently of the specific joint power spectrum. The limits to the achievable time-frequency resolution are a direct consequence of the spectral dynamics of the signal and the implied decorrelation of the power spectrum from one moment to the next. Increasing spectral dynamics imply decreasing time-frequency resolution capabilities. The optimum observation time depends on the spectral dynamics of the signal and is tool-independent. Article 4.6 further explores these issues.

References

[1] D. Gabor, "Theory of communication," *J. IEE*, vol. 93(III), pp. 429–457, November 1946.

[2] J. Hilgevoord and J. Uffink, "The mathematical expression of the uncertainty principle," in *Microphysical Reality and Quantum Formalism* (A. van der Merwe, G. Tarozzi, and F. Selleri, eds.), no. 25–26 in FUNDAMENTAL THEORIES OF PHYSICS, pp. 91–114, Dordrecht: Kluwer, 1988. Proc. of the Conf. at Urbino, Italy, 25 Sep. to 3 Oct., 1985.

[3] P. M. Oliveira and V. Barroso, "Uncertainty in the time-frequency plane," in *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pp. 607–611, Pocono Manor, PA, 14–16 August 2000.

[4] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[5] N. G. de Bruijn, "A theory of generalized functions, with applications to Wigner distribution and Weyl correspondence," *Nieuw Archief voor Wiskunde (3)*, vol. 21, pp. 205–280, 1973.

[6] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[7] A. D. Whalen, *Detection of Signals in Noise.* Academic Press, 1971.

[8] J. C. Andrieux, R. Feix, G. Mourgues, P. Bertrand, B. Izrar, and V. T. Nguyen, "Optimum smoothing of the Wigner-Ville distribution," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 35, pp. 764–769, June 1987.

[9] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Trans. Information Theory*, vol. 14, pp. 369–374, May 1968.

[10] B. Boashash and P. J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 216–220, January 1994.

4.5 GENERALIZED TFRs VIA UNITARY TRANSFORMS [0]

Despite the broad applicability of time-frequency representations, there exist situations where an analysis in terms of time and frequency coordinates is not appropriate (see [1–3], for example). These problems require joint distributions of other variables.

Joint distributions generalize single-variable distributions that measure the energy content of some physical quantity in a signal. Given a quantity a represented by the Hermitian (self-adjoint) operator $\mathcal{A}$, we obtain the density $|(\mathbb{F}_{\mathcal{A}} s)(a)|^2$ measuring the "a content" of the signal s simply by squaring the projection of s onto the formal eigenfunctions[1] $\mathbf{u}_a^{\mathcal{A}}$ of $\mathcal{A}$ [5]

$$|(\mathbb{F}_{\mathcal{A}}\, s)(a)|^2 \;=\; \left| \int s(x) \left[\mathbf{u}_a^{\mathcal{A}}(x)\right]^* dx \right|^2. \tag{4.5.1}$$

Classical examples of single variable densities include the time density $|(\mathbb{F}_{\mathcal{T}} s)(t)|^2 = |s(t)|^2$ and frequency density $|(\mathbb{F}_{\mathcal{F}} s)(f)|^2 = |S(f)|^2$ obtained by projecting onto the Dirac eigenfunctions of the time operator $(\mathcal{T}s)(x) = x\, s(x)$ and the sinusoidal eigenfunctions of the frequency operator $(\mathcal{F}s)(x) = \frac{1}{j2\pi}\dot{s}(x)$. (We will use both S and $\mathbb{F}s$ to denote the Fourier transform of the signal s.)

Joint distributions attempt to measure the simultaneous signal energy content of multiple physical quantities. Given two quantities a and b, a joint distribution $(\mathbf{P}_{a,b}\, s)(a,b)$ measuring the joint a-b content in the signal s has as marginals the respective $\mathcal{A}$ and $\mathcal{B}$ energy densities[2]

$$\int (\mathbf{P}_{a,b}\, s)(a,b)\, db \;=\; |(\mathbb{F}_{\mathcal{A}}\, s)(a)|^2 \tag{4.5.2}$$

$$\int (\mathbf{P}_{a,b}\, s)(a,b)\, da \;=\; |(\mathbb{F}_{\mathcal{B}}\, s)(b)|^2. \tag{4.5.3}$$

The Wigner distribution from Cohen's class of time-frequency distributions [5] supplies a classical example of a joint distribution that marginalizes to the time and frequency densities.

Many different constructions have been proposed for generating joint distributions. The various approaches fall into two broad categories: *general methods* and *coordinate change methods.* General methods can create distributions for every

[0]Author: **Richard Baraniuk**, Department of Electrical and Computer Engineering, Rice University, 6100 Main Street, Houston, TX 77005, USA (richb@rice.edu; http://www.dsp.rice.edu).
Reviewers: P. Flandrin and F. Auger.

[1]When the operator $\mathcal{A}$ is Hermitian, the eigenequation is merely algebraic, and the eigenfunctions are actually tempered distributions. More rigorously, we could employ projection-valued measures for the eigenfunctions [4].

[2]Alternatively, we can define joint distributions in terms of their covariance properties under certain unitary transformations. For more details, see [6, 7].

possible pairing of physical quantities by working from first principles [5,8,9]. Coordinate change methods, on the contrary, sacrifice some flexibility for simplicity by bootstrapping existing distributions into new contexts using signal or axis transformations [1,2,8,10,11].

In the following, we review these two categories of joint distribution constructions and address their linkage.

4.5.1 Three Approaches to Joint Distributions

4.5.1.1 Cohen's General Method

Given two operators $\mathcal{A}$ and $\mathcal{B}$ representing two arbitrary physical quantities a and b, Cohen forms the joint distribution of a and b as (see [5,8] for more details)

$$(\mathbf{P}_{a,b}\, s)(a,b) \;=\; \iiint s^*(x)\, \left(e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})}s\right)(x)\, e^{-j2\pi(\alpha a+\beta b)}\, dx\, d\alpha\, d\beta. \qquad (4.5.4)$$

Cohen's construction is general, but it requires that we solve a sometimes complicated operator equation to express the exponentiated operator $e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})}$. The time-frequency case is well understood; using the time and frequency operators $\mathcal{T}$ and $\mathcal{F}$ yields Cohen's class of time-frequency distributions [5, p. 136].

4.5.1.2 Axis Transformation Method

Joint a–b distributions are easily obtained when we can functionally relate these variables to time and frequency by $a = \alpha^{-1}(t)$ and $b = \beta^{-1}(f)$. In this special case, we can derive an a–b distribution simply by warping the coordinates of a time-frequency distribution $\mathbf{P}_{t,f}$ [8]

$$(\mathbf{P}_{a,b}\, s)(a,b) \;=\; \left|\dot{\alpha}(a)\,\dot{\beta}(b)\right|\; (\mathbf{P}_{t,f}\, s)\,[\alpha(a),\beta(b)]\,. \qquad (4.5.5)$$

It is easily verified that all $\mathbf{P}_{a,b}s$ obtained in this way correctly marginalize to $|(\mathbb{F}_{\mathcal{A}}s)(a)|^2 = |\dot{\alpha}(a)|\;|(\mathbb{F}_{\mathcal{T}}s)(\alpha(a))|^2 = |\dot{\alpha}(a)|\;|s(\alpha(a))|^2$ and $|(\mathbb{F}_{\mathcal{B}}s)(b)|^2 = |\dot{\beta}(b)|\;|(\mathbb{F}_{\mathcal{F}}s)(\beta(b))|^2 = |\dot{\beta}(b)|\;|S(\beta(b))|^2$.

Example: Distributions $\mathbf{P}_{t,r}s$ of time t and "inverse frequency" r (represented by the operator $\mathcal{R} = \frac{f_0}{\mathcal{F}}$) can be constructed from Cohen's class time-frequency distributions (i.e. the quadratic class as defined on p. 68) through the change of variable $r = \frac{f_0}{f}$ [5, p. 238]. The resulting class of distributions, obtained as

$$(\mathbf{P}_{t,r}\, s)(t,r) \;=\; \frac{f_0}{r^2}\,(\mathbf{P}_{t,f}\, s)\left(t, \frac{f_0}{r}\right), \qquad (4.5.6)$$

marginalizes to time $|s(t)|^2$ and inverse frequency $|(\mathbb{F}_{\mathcal{R}}s)(r)|^2 = \frac{f_0}{r^2}\,|S(f_0/r)|^2$.

4.5.1.3 Signal Transformation Method (Unitary Equivalence)

Joint a–b distributions are also easily generated when the quantities a and b are *unitarily equivalent* to time and frequency [1,10], with

$$\mathcal{A} = \mathbf{U}^{-1}\mathcal{T}\mathbf{U}, \qquad \mathcal{B} = \mathbf{U}^{-1}\mathcal{F}\mathbf{U}, \tag{4.5.7}$$

and $\mathbf{U}$ a unitary transformation. In this case, a joint a–b distribution can be obtained by preprocessing a time-frequency distribution with $\mathbf{U}$

$$(\mathbf{P}_{a,b}\, s)(a,b) = (\mathbf{P}_{t,f}\, \mathbf{U}s)(a,b). \tag{4.5.8}$$

The signal transformation $\mathbf{U}$ can be interpreted as "rotating" the coordinates of the time-frequency distribution to the new variables.

All $\mathbf{P}_{a,b}s$ obtained in this way correctly marginalize to $|(\mathbb{F}_{\mathcal{A}}s)(a)|^2 = |(\mathbb{F}_{\mathcal{T}}\mathbf{U}s)(a)|^2 = |(\mathbf{U}s)(a)|^2$ and $|(\mathbb{F}_{\mathcal{B}}s)(b)|^2 = |(\mathbb{F}_{\mathcal{F}}\mathbf{U}s)(b)|^2 = |(\mathbb{F}\mathbf{U}s)(b)|^2$ [1].

Example: Define the logarithmic time operator $(\mathcal{L}s)(x) = \log(x)\, s(x)$, and define the Mellin operator $\mathcal{H} = \frac{1}{2}(\mathcal{T}\mathcal{F} + \mathcal{F}\mathcal{T})$. (Cohen refers to $\mathcal{H}$ as the "scale" operator [5,12].) These operators are unitarily equivalent to $\mathcal{T}$ and $\mathcal{F}$ as in (4.5.7) with $(\mathbf{U}s)(x) = e^{x/2}\, s(e^x)$. Therefore, we can construct distributions marginalizing to logarithmic time $|(\mathbb{F}_{\mathcal{L}}s)(l)|^2 = |e^{l/2}\, s(e^l)|^2$ and Mellin transform

$$|(\mathbb{F}_{\mathcal{H}}s)(m)|^2 = \left| \int s(x)\, e^{-j2\pi m \log x}\, |x|^{-1/2}\, dx \right|^2 \tag{4.5.9}$$

simply by preprocessing the signal by $\mathbf{U}$ before computing a time-frequency distribution [1,2,10,11].

4.5.2 Linking Signal and Axis Transformations

Due to the individual limitations of the axis and signal transformation methods [13], only general operator methods can generate joint distributions for all possible operator pairs. However, when used in *tandem,* the axis and signal transformation methods yield a powerful method for generating a large number of joint distributions [13]. By executing first a unitary preprocessing transformation on the signal and then an axis warping transformation on the distribution, we can remap time-frequency distributions to a large class of different joint distributions.

Consider two variables a and b whose Hermitian operator representations $\mathcal{A}$ and $\mathcal{B}$ satisfy the following two conditions:

1. $\mathcal{A}$ and $\mathcal{B}$ can be related to the time and frequency operators as

   $$\mathcal{A} = \mathbf{Z}^{-1}\mathcal{T}\mathbf{Z}, \qquad \mathcal{B} = \mathbf{U}^{-1}\mathcal{F}\mathbf{U}, \tag{4.5.10}$$

   with $\mathbf{Z}$ and $\mathbf{U}$ unitary transformations.

2. The composition $\mathbf{V} = \mathbf{Z}\mathbf{U}^{-1}$ is an axis warping operator of the form

   $$(\mathbf{V}g)(x) = |\dot{v}(x)|^{1/2}\, g[v(x)], \tag{4.5.11}$$

   with v a smooth, 1–1 function.

In this case, a joint a–b distribution $(\mathbf{P}_{a,b}\, s)(a,b)$ can be constructed from any time-frequency distribution $(\mathbf{P}_{t,f}\, s)(t,f)$ through [13]

$$(\mathbf{P}_{a,b}\, s)(a,b) \;=\; |\dot{v}(a)|\;(\mathbf{P}_{t,f}\,\mathbf{U}s)\,[v(a),b]\,. \tag{4.5.12}$$

The interpretation of (4.5.12) is simple: The transformation $\mathbf{U}$ rotates the (t,f) coordinates of the time-frequency distribution to the new coordinates (u,b). The transformation $\mathbf{V}$ then warps the rotated time axis u to align it with the quantity a.

In other words, if $\mathcal{A}$ and $\mathcal{B}$ relate to $\mathcal{T}$ and $\mathcal{F}$ as in (4.5.10) and (4.5.11), then we can obtain a large class of a–b distributions using a simple three-step procedure: [3]

1. transform the signal $s \mapsto \mathbf{U}s$;
2. compute a conventional time-frequency distribution of the transformed signal;
3. warp the remapped time axis of the resulting distribution.

The advantage of the double transformation method is that it breaks the severe restrictions placed on the quantities a and b by both the axis and signal transformation methods described in Section 4.5.1. By allowing the choice of both $\mathbf{U}$ and $\mathbf{Z}$, we gain access to a much larger class of distributions. However, completely free choice is still not possible, because $\mathbf{U}$ and $\mathbf{Z}$ must have the structural property of canceling modulo the warping operator $\mathbf{V}$.

4.5.3 Examples of Linked Signal/Axis Transformations

Linked signal/axis transformations are especially useful for transforming time-frequency distributions to distributions of time versus a "warped frequency" variable. In this case, we set $\mathbf{Z}$ to the identity operator, choose $\mathbf{U}$ to be a warping operator based around the axis warping function v^{-1}, and set $\mathbf{V} = \mathbf{U}^{-1}$ to warp the rotated time axis back the time variable. The resulting distributions marginalize to time

$$\int (\mathbf{P}_{t,b}\, s)(t,b)\, db \;=\; |s(t)|^2 \tag{4.5.13}$$

and the integral transform

$$\int (\mathbf{P}_{t,b}\, s)(t,b)\, dt \;=\; \left| \int s(x)\, e^{-j2\pi b v(x)}\, |\dot{v}(x)|^{1/2}\, dx \right|^2 \tag{4.5.14}$$

[3] If we define the 2-d transformation $\overline{\mathbf{V}}$ as the area-preserving change of variables

$$(\overline{\mathbf{V}}G)(x,y) \;=\; |\dot{v}(x)|\, G\,[v(a),b]\,,$$

then we can write (4.5.12) in the standard form $\mathbf{P}_{a,b} = \overline{\mathbf{V}}\mathbf{P}_{t,f}\mathbf{U}$ found in [1]. However, whereas [1] emphasized using $\overline{\mathbf{V}}$ only to warp both axes of $\mathbf{P}_{a,b}\, s$ back to indicate time and frequency, in this article we exploit a range of different $\overline{\mathbf{V}}$.

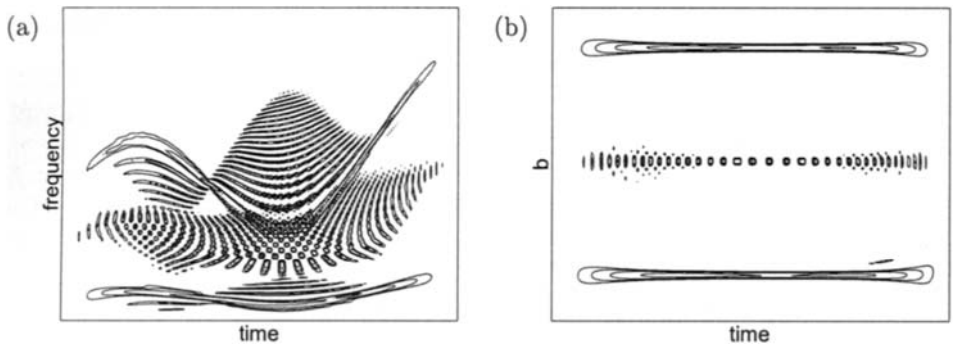


Fig. 4.5.1: (a) Wigner time-frequency distribution of a synthetic test signal composed of two frequency modulated tones. (b) Warped Wigner distribution with time and "composite linear/sinusoidal instantaneous frequency" (variable "b" in (4.5.14)) marginals. The variable b measures the variation of the linear/sinusoidal instantaneous frequency path in time-frequency.

that projects the signal onto the eigenfunctions of the operator $\mathbf{U}^{-1}\mathcal{T}\,\mathbf{U}$ [1]

$$\mathbf{u}_b^{\mathbf{U}^{-1}\mathcal{T}\,\mathbf{U}}(x) \;=\; (\mathbf{U}^{-1}\mathbf{u}_b^{\mathcal{T}})(x) \;=\; e^{j2\pi b v(x)}\,|\dot{v}(x)|^{1/2}. \tag{4.5.15}$$

Each choice of v results in a different warped frequency transform matched to a distinct class of instantaneous frequency characteristics. The Fourier, Mellin, and chirp transforms result from the choices $v(x) = x$, $v(x) = \log x$, and $v(x) = |x|^c \operatorname{sgn}(x)$, respectively.

To continue the example of Section 4.5.1.3, applying the warp $\mathbf{U}^{-1}$ to the logarithmic time axis of the logarithmic time versus Mellin distribution remaps that axis back to true time. The resulting distributions lie in Cohen's class of time-Mellin distributions (time-scale in his terminology) [5, 12]. This class contains the Marinovich-Altes (warped Wigner) distribution [11, 14]. It is important to note that this class is out of the grasp of either signal or axis warping alone.

In Figure 4.5.1, we show two distributions of a signal consisting of two components concentrated along composite linear/sinusoidal instantaneous frequencies. Since the Wigner time-frequency distribution does not match signals of this type as well as sinusoids, impulses, and linear chirps, it exhibits copious cross-components. Prewarping the signal to account for the form of the signal yields a postwarped distribution that marginalizes to time and "composite linear/sinusoidal instantaneous frequency" content (corresponding to variable "b" in (4.5.14)) and therefore better matches the signal.

Reversing the rôles of time and frequency in the warping procedure will yield frequency versus warped time distributions that match different classes of group delay (dispersion) characteristics.

4.5.4 Summary and Conclusions

Signal and axis transformations provide a simple framework for studying certain distributions of variables beyond time-frequency and time-scale. When applicable, warping results turn the theory of joint distributions of arbitrary variables into an easy exercise of coordinate transformation.

References

[1] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: A new twist on signal processing," *IEEE Trans. Signal Processing*, vol. 43, pp. 2269–2282, October 1995.

[2] 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 Processing*, vol. 41, pp. 3425–3444, December 1993. Special Issue on Wavelets and Signal Processing.

[3] M. Coates and W. J. Fitzgerald, "Time-frequency signal decomposition using energy mixture models," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, vol. 2, pp. 633–636, Istanbul, 5–9 June 2000.

[4] G. B. Folland, *Harmonic Analysis in Phase Space*. No. 122 in ANNALS OF MATHEMATICS STUDIES, Princeton, NJ: Princeton University Press, 1989.

[5] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[6] O. Rioul and P. Flandrin, "Time-scale energy distributions: A general class extending wavelet transforms," *IEEE Trans. Signal Processing*, vol. 40, pp. 1746–1757, July 1992.

[7] J. Bertrand and P. Bertrand, "A class of affine Wigner functions with extended covariance properties," *J. of Mathematical Physics*, vol. 33, pp. 2515–2527, July 1992.

[8] M. O. Scully and L. Cohen, "Quasi-probability distributions for arbitrary operators," in *The Physics of Phase Space: Nonlinear Dynamics and Chaos, Geometric Quantization, and Wigner Functions* (Y. S. Kim and W. W. Zachary, eds.), no. 278 in LECTURE NOTES IN PHYSICS, pp. 253–263, Berlin: Springer, 1987. Proc. First Internat. Conf. on the Physics of Phase Space, University of Maryland, College Park, MD, 20–23 May 1986.

[9] R. G. Baraniuk, "Beyond time-frequency analysis: Energy densities in one and many dimensions," *IEEE Trans. Signal Processing*, vol. 46, pp. 2305–2314, September 1998.

[10] R. G. Baraniuk and L. Cohen, "On joint distributions of arbitrary variables," *IEEE Signal Processing Letters*, vol. 2, pp. 10–12, January 1995.

[11] R. A. Altes, "Wideband, proportional-bandwidth Wigner-Ville analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1005–1012, June 1990.

[12] L. Cohen, "The scale representation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3275–3292, December 1993.

[13] R. G. Baraniuk, "Joint distributions of arbitrary variables made easy," *J. of Multidimensional Systems & Signal Processing*, vol. 9, pp. 341–348, October 1998. Special issue on time-frequency analysis.

[14] G. Eichmann and N. M. Marinovich, "Scale-invariant Wigner distribution," in *Proc. SPIE: Analog Optical Processing and Computing*, vol. 519, pp. 18–25, Soc. of Photo-optical Instrumentation Engineers, Cambridge, MA, 25–26 October 1984.

4.6 SIGNAL MEASURES IN THE TIME-FREQUENCY PLANE[0]

4.6.1 Time-Frequency Analysis

The goal of time-frequency analysis has primarily been to characterize and visualize the behavior of nonstationary signals. This is achieved by distilling both the amplitude and phase information of a signal's time series to present an image of the variation of the frequency content of the signal with respect to time. The resulting time-frequency distribution (TFD), which has many incarnations, desires to show this information with the highest possible resolution (energy concentration) and integrity (correct boundaries in time and frequency). The one aspect of the TFD that has not been well understood (and thus utilized), however, is the derivation of useful measures and parameters from within the time-frequency plane itself. We seek here to provide an introduction to measurements in the time-frequency domain.

4.6.2 Density Distributions and Energy Distributions

Of the large number of TFDs that have been developed, most fall into and between two classes - the density distributions and the energy distributions. They may alternately be called (in a non-strict sense) the unsmoothed and smoothed TFDs. Density distributions are of the form

$$\rho_{\gamma_z}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_z(t-t_1, f-f_1) e^{\frac{j2\pi t_1 f_1}{\gamma}}\, dt_1\, df_2, \quad 0 \le \gamma \le \frac{1}{2}\,. \tag{4.6.1}$$

where $W_z(t,f)$ is the Wigner-Ville Distribution (WVD) and is expressed as

$$W_z(t,f) = \int_{-\infty}^{\infty} z(t+\frac{\tau}{2}) z^*(t-\frac{\tau}{2}) e^{-j2\pi f\tau}\, d\tau\,. \tag{4.6.2}$$

The WVD is a member of the density class for $\gamma \to 0$ (which may be seen by taking the double Fourier transform of $e^{\frac{j2\pi t_1 f_1}{\gamma}}$ and then applying the limit).

The other form, an energy distribution, is better known as a spectrogram,

$$|F_z^h(t,f)|^2 = S_z^h(t,f) = \Big|\int_{-\infty}^{\infty} z(\tau) h^*(t-\tau) e^{-j2\pi f\tau}\, d\tau\Big|^2 \tag{4.6.3}$$

where $h(t)$ is the analyzing window. Energy distributions are obtained from density TFDs by smoothing with the (density) TFD of the window function. Many distributions fall in between the two classes, e.g. the reduced interference distribution (RID) [1], and the Zhao-Atlas-Marks distribution (ZAM) [2]. In addition, there are many distributions that are signal dependent/adaptive (an example is provided later in this work and further dealt with in Article 5.3 of this book).

---

[0]Author: **Graeme Jones**, Sicom Systems Ltd., 67 Canboro Road, 2nd Floor Fonthill Ontario, L0S 1E0 Canada (gjones@sicomsystems.com). Reviewers: D. L. Jones and W. J. Williams.

A question that naturally arises here is why such a distinction between TFDs should reasonably be made? The answer lies in our conventions of signal representation. For a typical one-dimensional signal, the (complex) time series tell us much useful information about the composition of the signal (e.g. its phase), but it is often a poor way to visualize the signal itself. Power and spectral density plots are much more useful for such things. Examination of a signal in two dimensions raises similar concerns. The rest of the article looks to density TFDs as providing the fine organizational data about the signal, while viewing an energy TFD (spectrogram) as a convenient tool for visualizing it.

4.6.3 Signal Measures in Time-Frequency

To extend concepts of one-dimensional signals to two dimensions, we should find some natural extension of the way we characterize and measure them. At a very basic level, we often separate a signal into its amplitude and phase, and then pursue further measurements. For example, the instantaneous frequency is a common way to describe the phase variation. By providing a straight-forward extension to one-dimensional mathematical expressions of interest,

$$\begin{aligned}\int_{-\infty}^{\infty} f^m |Z(f)|^2 df &= \int_{-\infty}^{\infty} [\frac{1}{z(t)}(\frac{1}{2\pi j})^m \frac{d^m}{dt^m} z(t)]\, |z(t)|^2 dt \\ &= \int_{-\infty}^{\infty} F^m(t)|z(t)|^2\, dt \end{aligned} \tag{4.6.4}$$

and

$$\begin{aligned}\int_{-\infty}^{\infty} t^n |z(t)|^2 dt &= \int_{-\infty}^{\infty} [\frac{1}{Z(f)}(\frac{-1}{2\pi j})^n \frac{d^n}{df^n} Z(f)]\, |Z(f)|^2 df \\ &= \int_{-\infty}^{\infty} T^n(f)|Z(f)|^2\, df\ , \end{aligned} \tag{4.6.5}$$

much is revealed about how measures in time-frequency can actually be derived. It is from expressions such as these above that widely used quantities, like the instantaneous frequency and group delay ($F^1(t)$ and $T^1(f)$ in the equations above), have been defined.

4.6.3.1 Extension of the Fourier Relations to Two Dimensions

Providing two-dimensional relations in the form of equations 4.6.4 and 4.6.5 is straight-forward. Standard Fourier theory yields an expression of the form:

$$\begin{aligned}&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n f^m |v(t,f)|^2\, dt\, df \\ &= (\frac{-1}{2\pi j})^n (\frac{1}{2\pi j})^m \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} V^*(t_1,f_1) \frac{\partial^m}{\partial t_1^m}\frac{\partial^n}{\partial f_1^n} V(t_1,f_1)\, dt_1\, df_1 \end{aligned} \tag{4.6.6}$$

where

$$V(t_1, f_1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} v(t, f) e^{-j2\pi t f_1} e^{j2\pi f t_1} \, dt \, df \, . \tag{4.6.7}$$

We now desire to incorporate within the above expression the classes of TFDs. Since equation 4.6.6 is based on an energy quantity (the absolute square of $v(t, f)$) it makes sense to replace this with a spectrogram. Using Fourier relations for TFDs in two dimensions (see [3] for the complete derivation), local parameters in two dimensions can be derived as:

$$T^n F^m{}_\gamma(t, f) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n f^m \rho_{\gamma_z}(t_1, f_1) \rho_{\gamma_h}(t - t_1, f - f_1) dt_1 df_1}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_{\gamma_z}(t_1, f_1) \rho_{\gamma_h}(t - t_1, f - f_1) \, dt_1 \, df_1} \tag{4.6.8}$$

$$= \frac{1}{|F_z^h(t, f)|^2} (\frac{-1}{2\pi j})^n (\frac{1}{2\pi j})^m \frac{\partial^m}{\partial \tau^m} \frac{\partial^n}{\partial \nu^n} [e^{-j\pi\nu t} e^{j\pi\tau f} F_z^h(t + (\frac{1}{2} - \gamma)\tau, f + (\frac{1}{2} + \gamma)\tau) F_z^{h*}(t - (\frac{1}{2} + \gamma)\tau, f - (\frac{1}{2} - \gamma)\tau)] \, |_{(\nu,\tau)=(0,0)} \tag{4.6.9}$$

where the short-time Fourier transform (STFT) has been defined here as:

$$F_z^h(t, f) = \int_{-\infty}^{\infty} z(\frac{t}{2} + \tau) h^*(\frac{t}{2} - \tau) e^{-j2\pi f \tau} \, d\tau \, . \tag{4.6.10}$$

In the limit $h(t) \to \delta(t)$, the conventional one-dimensional quantities are extracted. For example, with $n = 1$, the following relation is produced for the above limit:

$$T^0 F^m{}_\gamma(t) = \frac{\int_{-\infty}^{\infty} f^m \rho_{\gamma_z}(t, f) df}{\int_{-\infty}^{\infty} \rho_{\gamma_z}(t, f) \, df} \, . \tag{4.6.11}$$

When $\gamma = 0$ and $m = 1$, the actual instantaneous frequency ($F^1(t)$ of equation 4.6.4) is returned. This observation reinforces the notion that these quantities are a two-dimensional extension of the one-dimensional local measures.

4.6.3.2 Generation of Local Measurements in Time-Frequency

We now concentrate on results based on the WVD ($\gamma = 0$ in the above equations) due to the fact that it is real. The theoretical relation of 4.6.8 prompts the following general expression for local time-frequency measures

$$T^n F^m{}_\gamma(t, f) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t^n f^m W_z(t_1, f_1) L_{(t,f)}(t_1, f_1) dt_1 df_1}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_z(t_1, f_1) L_{(t,f)}(t_1, f_1) \, dt_1 \, df_1} \, . \tag{4.6.12}$$

L should be a function in time-frequency space that is concentrated or localized in some way. This necessitates the use of a window function whose WVD has strong localizing properties. The next section reveals how the WVD of a Gaussian function is an ideal choice for this task, and also presents a special property of Gaussian-based local measures that ultimately lead to their successful interpretation and application.

4.6.4 Properties and Interpretation of Local Measurements in Time-Frequency

This section builds upon the quantities advanced as local time-frequency measures from the previous sections. Here we focus our attention on the set of first and second order measures, which we seek to associate with notions of signal extent and spread. In classical time-series analysis, the measures of extent for Fourier pairs are bounded by the Fourier (Heisenberg) uncertainty principle. There are a number of different ways of expressing these limits, of which two are given below [4]:

$$\langle (t^2 - \bar{t})^2 \rangle \langle (f^2 - \bar{f})^2 \rangle - (\langle tf \rangle - \bar{t}\bar{f})^2 \geq \frac{1}{16\pi^2} \tag{4.6.13}$$

and [5]

$$\langle (t - \bar{t})^2 \rangle + \langle (f - \bar{f})^2 \rangle \geq \frac{1}{2\pi} \tag{4.6.14}$$

where $\langle\rangle$ represents the normalized average operator, and $\bar{t}$ and $\bar{f}$ are the time and frequency (magnitude squared) signal averages. These uncertainty principles provide bounds on how a signal is spread in the time and the frequency domain, and are satisfied with equality for Gaussian functions.

When TFDs are considered, no such equivalent relations are apparent. Some very interesting results (which relate to the concepts of spread and extent) can be revealed, however, when Gaussian windows are utilized. To develop such results in time-frequency, it is useful to employ orthogonal Hermite function expansions of the time signals, to derive expressions for the local moments. For example, Hermite function properties [6] allow expression of the scalar time operator as

$$\begin{aligned} tP_{mn}(t,f) = \frac{j}{8\pi} \quad & ((2m+2)^{\frac{1}{2}} P_{m+1,n}(t,f) - (2m)^{\frac{1}{2}} P_{m-1,n}(t,f) \\ & + (2n)^{\frac{1}{2}} P_{m,n-1}(t,f) - (2n+2)^{\frac{1}{2}} P_{m,n+1}(t,f)) \end{aligned} \tag{4.6.15}$$

where $P_{mn}(t,f)$ is the (cross) WVD of an mth and nth order Hermite function. Other expressions for powers of t and f follow. By using relations such as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_{mn}(t,f) P_{qr}(t,f)\, dt\, df = (h_m, h_q)(h_n, h_r) \tag{4.6.16}$$

(the left hand side shows inner products of Hermite functions in $L^2(\mathcal{R})$), as well as the orthogonality of the Hermite functions

$$(h_n, h_m) = 0 \;\; \forall m \neq n\,, \tag{4.6.17}$$

the local moments of a TFD may be derived. A very significant result is available when the WVD of a Gaussian signal, $h(t) = e^{-\pi t^2}$ (a zeroth-order Hermite function), is used as the localizing window to calculate the local parameters of equation 4.6.12. It is stated thus:

$$[T^2F^0(t,f) - T^1F^0(t,f)^2] + [T^0F^2(t,f) - T^0F^1(t,f)^2] = \frac{1}{4\pi}\;;\;\; \forall\, z(t),\; (h(t) = e^{-\pi t^2}) \tag{4.6.18}$$

This result has been confirmed in experimental work, and a complete theoretical proof was given by the author in [3] (also showing the result holds for any linear FM Gaussian window). The implication of this simple result is quite profound. It may be observed that it is similar in form to the addition uncertainty principle of equation 4.6.14. This lends strong support to the notion that the Gaussian window is a credible localizing function. A remarkable property forthcoming from the above result is that it holds even if one of the time or frequency local spread measures is negative.

One reason that measures in the time-frequency plane as well as conventional instantaneous quantities have been cited as non-intuitive is that they often yield values that do not appear related to the physical reality (e.g. a negative instantaneous frequency for a signal that has no negative frequency content). The relation given above shows a logical balance exists within this measurement framework. The measures do not just possess nice theoretical properties, however. To support this, the next section applies these local time-frequency measures to improve the resolution of the spectrogram.

4.6.5 Application of Local Time-Frequency Measures to Energy Distributions (Spectrograms)

Energy distributions are an important tool for displaying time-frequency signal content. Their main disadvantage is that they can often skew the actual time-frequency energy content, since they employ a window that necessarily possesses its own time-frequency signature. A window function that is circularly symmetric in the time-frequency plane is one solution, but this provides mediocre resolution in all directions at the expense of any smearing.

A more appropriate solution would be to adapt the analyzing window so that it captures the local time-frequency character of the signal (see [7–9] for some novel examples). There are a number of ways to adapt the window to the local time-frequency behavior. A Taylor series expansion of the signal's TFD about the region of interest is one approach, but there is no guarantee that the local window generated from the approximation will produce a valid energy distribution.

Instead of matching the first few terms of a Taylor series, let us match the first couple of local moments, or in this case the bandwidth or spread measures. The quantities of interest are the local time bandwidth, LTB $(T^2F^0(t,f) - T^1F^0(t,f)^2)$, the local frequency bandwidth, LFB $(T^0F^2(t,f) - T^0F^1(t,f)^2)$, and the local cross time-frequency product LTF $(T^1F^1(t,f))$. After applying a matching algorithm (described below), the resulting adaptive energy distribution would be of the form:

$$S_z^{h_M}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_z(t_1,f_1)W_h(t-t_1,f-f_1,t,f)\,dt_1\,df_1 \quad (4.6.19)$$

$$= \Big|\int_{-\infty}^{\infty} z(\tau)h_f(t,\tau)e^{-j2\pi f\tau}\,d\tau\Big|^2 \quad (4.6.20)$$

where $W_h(t_1,f_1,t,f)$ is the WVD of the location-dependent window function.

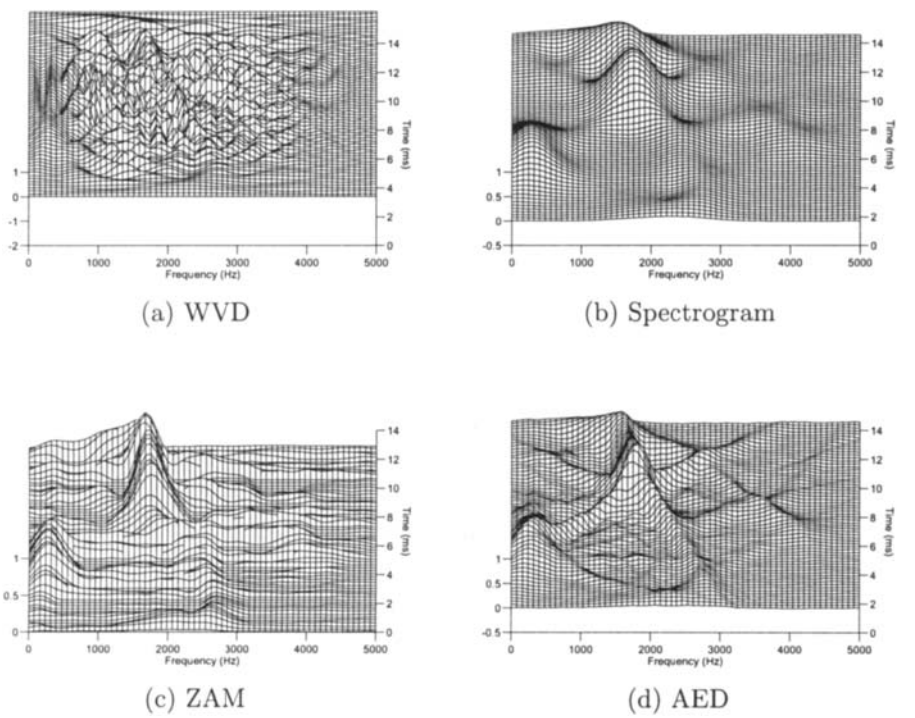


Fig. 4.6.1: Time-Frequency Distributions of a Humpback Whale Sound

The matching process is simple, and yields excellent results. It uses the spread measures to determine the most appropriate values for α and β (of a Gaussian window, $h(t) = e^{-\pi\alpha t^2} e^{j\pi\beta t^2}$) at each point of interest. An iterative approach is used, relying on the local uncertainty result of equation 4.6.18. A full explanation of the algorithm appears in [3]. The following section provides an example of this window matching technique.

4.6.6 Example Result for an Adaptive Energy Distribution

The algorithm briefly described in the previous section will now be exercised. This example comprises a highly non-stationary real signal - a sample from a Humpback whale song. Its density TFD (Fig. 4.6.1(a)) is uninterpretable. Although the circularly symmetric spectrogram of Fig. 4.6.1(b) improves the visualization, the information is highly smoothed. For comparison, a ZAM distribution is provided (Fig. 4.6.1(c)), which is in the class of TFDs that fall in between the density and energy distributions. The final figure, Fig. 4.6.1(d), displays the adaptive energy

distribution (AED), revealing a level of detail not seen in the previous two plots.

4.6.7 Summary and Conclusions

A primer on the use of quantities derived from the time-frequency plane has been presented here. It was shown that basic Fourier relations, used to define measures such as instantaneous frequency, can be extended (albeit with some care) to time-frequency distributions. The tools for deriving and analyzing these local time-frequency measures were introduced, and an interesting property of the second-order measures was presented. An example was provided to demonstrate potential applications of these local measures in the time-frequency plane. Article 4.4 complements this material with a discussion on the uncertainty principle in time-frequency analysis.

References

[1] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 37, pp. 862–871, June 1989.

[2] Y. Zhao, L. E. Atlas, and R. J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[3] G. Jones and B. Boashash, "Generalized instantaneous parameters and window matching in the time-frequency plane," *IEEE Trans. Signal Processing*, vol. 45, pp. 1264–1275, May 1997.

[4] C. W. Helstrom, *The Statistical Theory of Signal Detection.* New York: Oxford University Press, 1968.

[5] T. W. Parks and R. G. Shenoy, "Time-frequency concentrated basis functions," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'90)*, pp. 2459–2462, Albuquerque, NM, 3–6 April 1990.

[6] S. Thangavelu, *Lectures on Hermite and Laguerre Expansions.* Princeton, NJ: Princeton University Press, 1993.

[7] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Trans. Signal Processing*, vol. 41, pp. 1589–1602, April 1993.

[8] R. N. Czerwinski and D. L. Jones, "Adaptive cone-kernel time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1715–1719, July 1995.

[9] D. L. Jones and R. G. Baraniuk, "A simple scheme for adapting time-frequency representations," *IEEE Trans. Signal Processing*, vol. 42, pp. 3530–3535, December 1994.

4.7 TIME-FREQUENCY TRANSFER FUNCTION CALCULUS OF LINEAR TIME-VARYING SYSTEMS [0]

4.7.1 Linear Time-Varying Systems

Due to their generality, *linear time-varying* (LTV) systems (which can equivalently be viewed as linear operators [1]) have important advantages over linear time-invariant (LTI) systems. Applications of LTV systems include mobile communications (see Articles 9.5, 13.2, and 13.3), machine monitoring (see Articles 15.2 and 15.6), and nonstationary statistical signal processing (see Articles 9.2, 9.4, 12.1, and 12.4). An LTV system $\mathbf{H}$ maps the input signal $x(t)$ to an output signal $y(t)$ according to

$$y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} h(t,t')\,x(t')\,dt', \tag{4.7.1}$$

where $h(t,t')$ is the kernel (impulse response) of $\mathbf{H}$. LTI systems and their dual, *linear frequency-invariant* (LFI) systems, are special cases of LTV systems. For an LTI system, $y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} g(t-t')\,x(t')\,dt'$ and thus $h(t,t') = g(t-t')$. For an LFI system, $y(t) = (\mathbf{H}x)(t) = w(t)\,x(t)$ and thus $h(t,t') = w(t)\,\delta(t-t')$.

For LTI and LFI systems, there exist physically intuitive and numerically efficient analysis and design methods that are based on the spectral transfer function $G(f) = \int_{-\infty}^{\infty} g(\tau)\,e^{-j2\pi f\tau}\,d\tau$ and on the temporal transfer function $w(t)$, respectively. Unfortunately, a similar transfer function does not exist for LTV systems in general. However, in this article we will show that for the important class of *underspread* LTV systems, the *generalized Weyl symbol* constitutes an approximate time-frequency (TF) transfer function. We note that other TF symbols are discussed in Article 9.2.

4.7.2 The Generalized Weyl Symbol

The *generalized Weyl symbol* (GWS) of an LTV system $\mathbf{H}$ is a family of linear TF representations defined as [2]

$$L_{\mathbf{H}}^{(\alpha)}(t,f) \triangleq \int_{-\infty}^{\infty} h^{(\alpha)}(t,\tau)\,e^{-j2\pi f\tau}d\tau$$

with

$$h^{(\alpha)}(t,\tau) \triangleq h\Big(t+\Big(\frac{1}{2}-\alpha\Big)\tau, t-\Big(\frac{1}{2}+\alpha\Big)\tau\Big), \tag{4.7.2}$$

where α is a real-valued parameter. The GWS reduces to the ordinary *Weyl symbol* [3–6] for $\alpha = 0$, to Zadeh's *time-varying transfer function* [7] for $\alpha = 1/2$, and to

[0]Authors: **G. Matz** and **F. Hlawatsch**, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Gusshausstrasse 25/389, A-1040 Vienna, Austria (email: g.matz@ieee.org, fhlawats@pop.tuwien.ac.at, web: http://www.nt.tuwien.ac.at/dspgroup/time.html). Reviewers: S. Barbarossa and A. Papandreou-Suppappola.

Bello's *frequency-dependent modulation function* [8] (also known as *Kohn-Nirenberg symbol* [3]) for $\alpha = -1/2$:

$$L_{\mathbf{H}}^{(0)}(t,f) = \int_{-\infty}^{\infty} h\Big(t+\frac{\tau}{2}, t-\frac{\tau}{2}\Big)\, e^{-j2\pi f\tau}\, d\tau$$

$$L_{\mathbf{H}}^{(1/2)}(t,f) = \int_{-\infty}^{\infty} h(t,t-\tau)\, e^{-j2\pi f\tau}\, d\tau$$

$$L_{\mathbf{H}}^{(-1/2)}(t,f) = \int_{-\infty}^{\infty} h(t+\tau,t)\, e^{-j2\pi f\tau}\, d\tau\,.$$

In what follows, α will be considered fixed.

The GWS $L_{\mathbf{H}}^{(\alpha)}(t,f)$ is a *linear* TF representation of the LTV system $\mathbf{H}$. It contains all information about $\mathbf{H}$ since the kernel of $\mathbf{H}$ can be recovered from the GWS:

$$h(t,t') = \int_{-\infty}^{\infty} L_{\mathbf{H}}^{(\alpha)}\Big(\Big(\frac{1}{2}+\alpha\Big)t + \Big(\frac{1}{2}-\alpha\Big)t', f\Big)\, e^{j2\pi f(t-t')}\, df\,.$$

Also, the input-output relation (4.7.1) can be reformulated in terms of the GWS. This reformulation becomes especially simple for $\alpha = \pm 1/2$:

$$y(t) = \int_{-\infty}^{\infty} L_{\mathbf{H}}^{(1/2)}(t,f)\, X(f)\, e^{j2\pi ft}\, df, \qquad Y(f) = \int_{-\infty}^{\infty} L_{\mathbf{H}}^{(-1/2)}(t,f)\, x(t)\, e^{-j2\pi ft}\, dt.$$

For a rank-one system with impulse response $h(t,t') = u(t)\,u^*(t')$, $L_{\mathbf{H}}^{(\alpha)}(t,f)$ reduces to the generalized Wigner distribution [9] of the signal $u(t)$. Other interesting properties of the GWS can be found in [2,5,10] and for $\alpha = 0$ in [3,4,6].

Next, we consider the GWS of some simple specific systems. The results obtained suggest that (under appropriate assumptions to be discussed later) the GWS can be interpreted as a "TF transfer function" that characterizes the "TF weighting" produced by the LTV system $\mathbf{H}$, i.e., the way in which a component of the input signal $x(t)$ located about some TF point (t,f) is attenuated ($|L_{\mathbf{H}}^{(\alpha)}(t,f)| < 1$), amplified ($|L_{\mathbf{H}}^{(\alpha)}(t,f)| > 1$), or passed without attenuation or amplification ($|L_{\mathbf{H}}^{(\alpha)}(t,f)| = 1$) by $\mathbf{H}$.

- The GWS of the identity operator $\mathbf{I}$ with kernel $h(t,t') = \delta(t-t')$ is given by $L_{\mathbf{I}}^{(\alpha)}(t,f) \equiv 1$ (i.e., no attenuation/amplification anywhere in the TF plane).
- The GWS of the *TF shift operator* $\S_{\nu,\tau}^{(\alpha)}$ defined by[1] $(\S_{\nu,\tau}^{(\alpha)}x)(t) = x(t-\tau)\, e^{j2\pi\nu t}\, e^{j2\pi\nu\tau(\alpha-1/2)}$ is a two-dimensional complex sinusoid,[2] $L_{\S_{\nu,\tau}^{(\alpha)}}^{(\alpha)}(t,f) =$

[1]The parameter α in $\S_{\nu,\tau}^{(\alpha)}$ corresponds to the infinitely many ways of defining a joint TF shift by combining time shifts and frequency shifts. In particular, $\alpha = 1/2$ corresponds to first shifting in time and then shifting in frequency, whereas $\alpha = -1/2$ corresponds to first shifting in frequency and then in time.

[2]Note that the GWS parameter α is chosen equal to the parameter α in $\S_{\nu,\tau}^{(\alpha)}$.

$e^{j2\pi(\nu t-\tau f)}$, and thus $\left|L^{(\alpha)}_{\S^{(\alpha)}_{\nu,\tau}}(t,f)\right| = 1$ (i.e., no attenuation/amplification anywhere in the TF plane).

- The GWS of an LTI system with kernel $h(t,t') = g(t-t')$ reduces to the ordinary transfer function $G(f)$ for all t, i.e., $L^{(\alpha)}_{\mathbf{H}}(t,f) \equiv G(f)$.
- The GWS of an LFI system with kernel $h(t,t') = w(t)\,\delta(t-t')$ reduces to the temporal transfer function $w(t)$ for all f, i.e., $L^{(\alpha)}_{\mathbf{H}}(t,f) \equiv w(t)$.

The last two examples show the GWS's consistency with the conventional transfer functions.

In what follows, we shall investigate the conditions under which the GWS can be interpreted as a TF transfer function. The key element in this investigation is an analysis of the *TF shifts* produced by $\mathbf{H}$. For this analysis, we need another linear TF representation of $\mathbf{H}$, to be discussed next.

4.7.3 The Generalized Spreading Function

Besides a TF-dependent weighting, LTV systems can also introduce TF shifts of various input signal components. Here, the TF shift operator $\S^{(\alpha)}_{\nu,\tau}$ mentioned above—with ν denoting frequency (Doppler) shift and τ denoting time shift/delay—is an elementary example. A joint description of the TF shifts introduced by a linear system $\mathbf{H}$ is given by the *generalized spreading function* (GSF) [2]

$$S^{(\alpha)}_{\mathbf{H}}(\nu,\tau) \triangleq \int_{-\infty}^{\infty} h^{(\alpha)}(t,\tau)\, e^{-j2\pi\nu t}\, dt\,,$$

with $h^{(\alpha)}(t,\tau)$ as in (4.7.2). Like the GWS, the GSF $S^{(\alpha)}_{\mathbf{H}}(\nu,\tau)$ is a *linear* TF representation of the LTV system $\mathbf{H}$, and it contains all information about $\mathbf{H}$ since the kernel of $\mathbf{H}$ can be recovered from the GWS. The input-output relation (4.7.1) can be reformulated in terms of the GSF according to

$$y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S^{(\alpha)}_{\mathbf{H}}(\nu,\tau)\,(\S^{(\alpha)}_{\nu,\tau}x)(t)\,d\nu\,d\tau\,.$$

This represents the output signal $y(t) = (\mathbf{H}x)(t)$ as a weighted superposition of TF shifted versions $(\S^{(\alpha)}_{\nu,\tau}x)(t) = x(t-\tau)\,e^{j2\pi\nu t}\,e^{j2\pi\nu\tau(\alpha-1/2)}$ of the input signal $x(t)$. The (ν,τ)-dependent weights in this superposition are given by the GSF, thus establishing the GSF's interpretation as a "TF shift distribution" of $\mathbf{H}$. The extension of $S^{(\alpha)}_{\mathbf{H}}(\nu,\tau)$ about the origin of the (ν,τ) plane indicates the amount of TF shifts caused by $\mathbf{H}$. In particular, a large extension of $S^{(\alpha)}_{\mathbf{H}}(\nu,\tau)$ in the ν direction indicates large frequency/Doppler shifts (equivalently, fast time variation) and a large extension of $S^{(\alpha)}_{\mathbf{H}}(\nu,\tau)$ in the τ direction indicates large time shifts (delays).

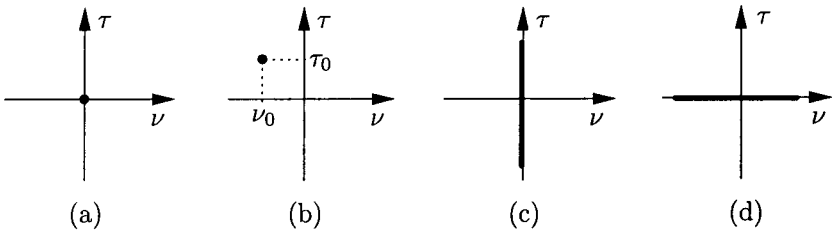


Fig. 4.7.1: Schematic representation of the GSF magnitude of some (classes of) linear systems: (a) identity operator, (b) TF shift operator, (c) LTI system, (d) LFI system.

GSFs with different α values differ merely by a phase factor, i.e.,

$$S_{\mathbf{H}}^{(\alpha_2)}(\nu,\tau) = S_{\mathbf{H}}^{(\alpha_1)}(\nu,\tau)\, e^{j2\pi(\alpha_1-\alpha_2)\nu\tau}\,.$$

Therefore, the GSF magnitude is independent of α, $\left|S_{\mathbf{H}}^{(\alpha_1)}(\nu,\tau)\right| = \left|S_{\mathbf{H}}^{(\alpha_2)}(\nu,\tau)\right|$, and thus we may simply write $|S_{\mathbf{H}}(\nu,\tau)|$. The GSF is related to the GWS by a two-dimensional Fourier transform,

$$S_{\mathbf{H}}^{(\alpha)}(\nu,\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L_{\mathbf{H}}^{(\alpha)}(t,f)\, e^{-j2\pi(\nu t-\tau f)}\, dt\, df. \tag{4.7.3}$$

Further properties of the GSF are described in [2,5,10].

Again, it is instructive to consider a few examples (see Fig. 4.7.1):

- The GSF of the identity operator $\mathbf{I}$ is given by $S_{\mathbf{I}}^{(\alpha)}(\nu,\tau) = \delta(\nu)\delta(\tau)$, which is zero for $(\nu,\tau) \neq (0,0)$ (i.e., neither frequency shifts nor time shifts).
- The GSF of a TF shift operator $\S_{\nu_0,\tau_0}^{(\alpha)}$ is obtained as $S_{\S_{\nu_0,\tau_0}^{(\alpha)}}^{(\alpha)}(\nu,\tau) = \delta(\nu-\nu_0)\,\delta(\tau-\tau_0)$, which is zero for $(\nu,\tau) \neq (\nu_0,\tau_0)$ (i.e., no frequency and time shifts other than by ν_0 and τ_0, respectively).
- The GSF of an LTI system $\mathbf{H}$ with kernel $h(t,t') = g(t-t')$ is given by $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau) = \delta(\nu)g(\tau)$ (i.e., only time shifts whose distribution is characterized by the impulse response $g(\tau)$).
- The GSF of an LFI system $\mathbf{H}$ with $h(t,t') = w(t)\delta(t-t')$ is given by $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau) = W(\nu)\delta(\tau)$, where $W(\nu)$ is the Fourier transform of $w(t)$ (i.e., only frequency shifts whose distribution is characterized by $W(\nu)$).

4.7.4 Underspread LTV Systems

Using the GSF, we now define the class of *underspread* LTV systems [5,10] for which, as we will see in Section 4.7.5, the GWS acts as a "TF transfer function." Conceptually, an LTV system is underspread if its GSF is well concentrated about

the origin of the (ν, τ) plane, which indicates that the system introduces only small TF shifts, i.e., the system's time variations are slow and/or its memory is short. In contrast, systems introducing large TF shifts are termed *overspread* [5,10].

There are two alternative mathematical characterizations of the GSF extension and, in turn, of underspread systems. The first characterization [10] requires that the support of the GSF $S_{\mathbf{H}}^{(\alpha)}(\nu, \tau)$ is confined to a compact region $\mathcal{G}_{\mathbf{H}}$ about the origin of the (ν, τ) plane, i.e., $|S_{\mathbf{H}}(\nu, \tau)| = 0$ for $(\nu, \tau) \notin \mathcal{G}_{\mathbf{H}}$. Let $\nu_{\mathbf{H}} \triangleq \max_{(\nu,\tau)\in\mathcal{G}_{\mathbf{H}}} |\nu|$ and $\tau_{\mathbf{H}} \triangleq \max_{(\nu,\tau)\in\mathcal{G}_{\mathbf{H}}} |\tau|$ denote the maximum frequency shift and time shift, respectively, introduced by the system $\mathbf{H}$. We define the *Doppler-delay spread* of $\mathbf{H}$ as $\sigma_{\mathbf{H}} \triangleq 4\nu_{\mathbf{H}}\tau_{\mathbf{H}}$, which is the area of the rectangle $[-\nu_{\mathbf{H}}, \nu_{\mathbf{H}}] \times [-\tau_{\mathbf{H}}, \tau_{\mathbf{H}}]$ enclosing $\mathcal{G}_{\mathbf{H}}$. Underspread LTV systems are then defined by the condition $\sigma_{\mathbf{H}} \ll 1$.

Unfortunately, the GSF of practical LTV systems rarely has compact support. An alternative, much more flexible characterization of the GSF extension and of underspread systems that does not require a compact support assumption is based on the normalized *weighted GSF integrals*[3] [5]

$$m_{\mathbf{H}}^{(\phi)} \triangleq \frac{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(\nu,\tau)\,|S_{\mathbf{H}}(\nu,\tau)|\,d\nu\,d\tau}{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |S_{\mathbf{H}}(\nu,\tau)|\,d\nu\,d\tau}.$$

Here, $\phi(\nu, \tau)$ is a nonnegative weighting function that satisfies $\phi(\nu, \tau) \geq \phi(0, 0) = 0$ and penalizes GSF contributions lying away from the origin. We also define the *GSF moments* $m_{\mathbf{H}}^{(k,l)} \triangleq m_{\mathbf{H}}^{(\phi_{k,l})}$ as special cases of $m_{\mathbf{H}}^{(\phi)}$ using the weighting functions $\phi_{k,l}(\nu, \tau) = |\nu|^l |\tau|^k$ with $k, l \in \mathbb{N}_0$. Thus, without the assumption of compact GSF support, a system $\mathbf{H}$ can now be considered to be underspread if suitable GSF integrals/moments are "small." While this is not a clear-cut definition of underspread systems, it has the advantage of being more flexible than the previous definition that was based on the area of the (assumedly) compact support of the GSF. It can be shown that if the GSF of $\mathbf{H}$ does have compact support with maximum frequency shift $\nu_{\mathbf{H}}$ and maximum time shift $\tau_{\mathbf{H}}$, then $m_{\mathbf{H}}^{(k,l)} \leq \nu_{\mathbf{H}}^l \tau_{\mathbf{H}}^k$ and, in particular, $m_{\mathbf{H}}^{(k,k)} \leq (\sigma_{\mathbf{H}}/4)^k$. Thus, LTV systems that are underspread in the compact-support sense can be considered as a special case of the extended underspread framework based on weighted GSF integrals and moments.

Examples of various types of underspread systems are illustrated in Fig. 4.7.2. It should be noted that the concept of underspread systems is not equivalent to that of slowly time-varying (quasi-LTI) systems. A quasi-LTI system (i.e., small $m_{\mathbf{H}}^{(0,l)}$) may be overspread if its memory is very long (i.e., very large $m_{\mathbf{H}}^{(k,0)}$), and a system with faster time-variations (i.e., larger $m_{\mathbf{H}}^{(0,l)}$) may be underspread if its memory is short enough (i.e., very small $m_{\mathbf{H}}^{(k,0)}$). Finally, note that according to (4.7.3), the GWS of an underspread LTV system is a *smooth* function.

[3] Other definitions of weighted GSF integrals and moments can be found in [5].

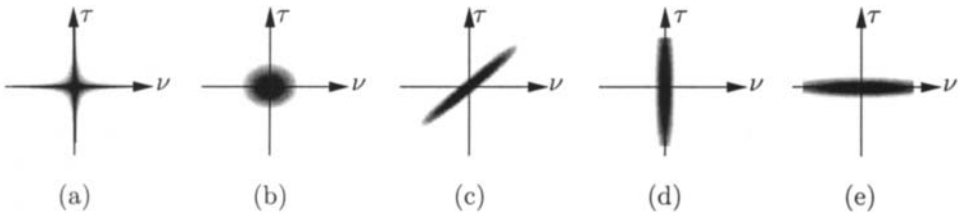


Fig. 4.7.2: GSF magnitude of (a) an underspread system with small $m_{\mathbf{H}}^{(1,1)}$, (b) an underspread system with small $m_{\mathbf{H}}^{(1,0)} m_{\mathbf{H}}^{(0,1)}$, (c) a "chirpy" underspread system [5], (d) a quasi-LTI system (slowly time-varying; small $m_{\mathbf{H}}^{(0,1)}$), (e) a quasi-LFI system (short memory; small $m_{\mathbf{H}}^{(1,0)}$).

4.7.5 Time-Frequency Transfer Function Calculus

For *underspread* LTV systems as defined in the previous section, the GWS acts as an approximate "TF transfer function" that generalizes the spectral (temporal) transfer function of LTI (LFI) systems. Indeed, if specific weighted GSF integrals $m_{\mathbf{H}}^{(\phi)}$ and/or moments $m_{\mathbf{H}}^{(k,l)}$ are small, one can show the validity of several transfer function approximations [5, 10], some of which are discussed in the following. Applications of these approximations include time-varying spectral analysis, linear TF filter design, and detection/estimation of nonstationary random processes (for references see Section 4.7.6).

Adjoint system. For an LTI system $\mathbf{H}$ with transfer function $G(f)$, the transfer function of the *adjoint* $\mathbf{H}^+$ is $G^*(f)$. A similar correspondence holds for LFI systems. For general LTV systems, the GWS of the adjoint $\mathbf{H}^+$ (with kernel $h^+(t,t') = h^*(t',t)$ [1]) is not equal to the conjugate of the GWS of $\mathbf{H}$ unless $\alpha = 0$. However, for an underspread LTV system $\mathbf{H}$ this is approximately true:

$$L_{\mathbf{H}^+}^{(\alpha)}(t,f) \approx \left[L_{\mathbf{H}}^{(\alpha)}(t,f)\right]^*. \tag{4.7.4}$$

Indeed, it can be shown [5] that the associated approximation error is upper bounded as

$$\left|L_{\mathbf{H}^+}^{(\alpha)}(t,f) - \left[L_{\mathbf{H}}^{(\alpha)}(t,f)\right]^*\right| \;\leq\; 4\pi\,|\alpha|\,\|S_{\mathbf{H}}\|_1\, m_{\mathbf{H}}^{(1,1)}, \tag{4.7.5}$$

with $\|S_{\mathbf{H}}\|_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |S_{\mathbf{H}}(\nu,\tau)|\,d\nu\,d\tau$. Thus, for an underspread system where $m_{\mathbf{H}}^{(1,1)}$ is small, the approximation (4.7.4) will be quite good. It follows from (4.7.4) that the GWS of a self-adjoint, underspread LTV system is approximately real-valued:

$$\mathbf{H}^+ = \mathbf{H} \quad\Longrightarrow\quad \left[L_{\mathbf{H}}^{(\alpha)}(t,f)\right]^* \approx L_{\mathbf{H}}^{(\alpha)}(t,f)\,, \quad \Im\left\{L_{\mathbf{H}}^{(\alpha)}(t,f)\right\} \approx 0\,.$$

In addition, if the underspread LTV system $\mathbf{H}$ is positive (semi-)definite [1], then $\Re\left\{L_{\mathbf{H}}^{(\alpha)}(t,f)\right\} \approx L_{\mathbf{H}}^{(\alpha)}(t,f)$ is approximately nonnegative.

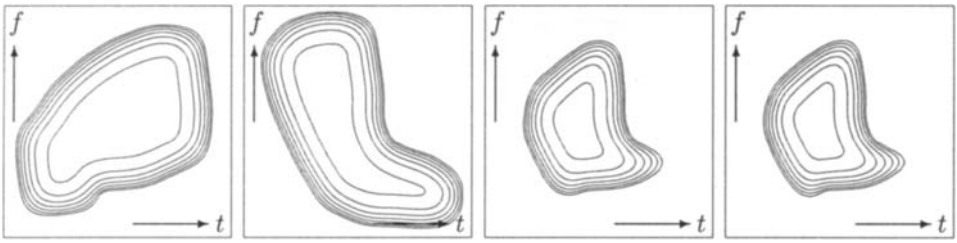


Fig. 4.7.3: Transfer function approximation of the Weyl symbol (GWS with $\alpha = 0$) for a composition of underspread systems: (a) Weyl symbol of $\mathbf{H}_1$, (b) Weyl symbol of $\mathbf{H}_2$, (c) Weyl symbol of $\mathbf{H}_2\mathbf{H}_1$, and (d) product of the individual Weyl symbols $L^{(0)}_{\mathbf{H}_1}(t,f)$ and $L^{(0)}_{\mathbf{H}_2}(t,f)$. Note the similarity of (c) and (d). In this and subsequent simulations, time duration is 128 samples and normalized frequency ranges from $-1/4$ to $1/4$.

Composition of systems. The transfer function of a composition (series connection) of two LTI systems with transfer functions $G_1(f)$ and $G_2(f)$ is given by the product $G_1(f)\,G_2(f)$. An analogous result holds for the composition of two LFI systems. This composition property of transfer functions is the cornerstone of many signal processing techniques. Unfortunately, a similar composition property no longer holds true for the GWS in the case of general LTV systems. However, the GWS of the composition $\mathbf{H}_2\mathbf{H}_1$ of two jointly underspread [5] LTV systems $\mathbf{H}_1$ and $\mathbf{H}_2$ is approximately equal to the product of their GWSs, i.e.,

$$L^{(\alpha)}_{\mathbf{H}_2\mathbf{H}_1}(t,f) \approx L^{(\alpha)}_{\mathbf{H}_1}(t,f)\,L^{(\alpha)}_{\mathbf{H}_2}(t,f). \tag{4.7.6}$$

An upper bound on the associated approximation error (similar to (4.7.5)) can again be provided [5]. Combining (4.7.4) with (4.7.6), we furthermore obtain

$$L^{(\alpha)}_{\mathbf{H}\mathbf{H}^+}(t,f) \approx L^{(\alpha)}_{\mathbf{H}^+\mathbf{H}}(t,f) \approx \left|L^{(\alpha)}_{\mathbf{H}}(t,f)\right|^2.$$

Fig. 4.7.3 shows an example illustrating the approximation (4.7.6) for $\alpha = 0$. In this example, the maximum normalized error $\max_{t,f}\left|L^{(0)}_{\mathbf{H}_2\mathbf{H}_1}(t,f) - L^{(0)}_{\mathbf{H}_1}(t,f)\,L^{(0)}_{\mathbf{H}_2}(t,f)\right| / \max_{t,f}\left|L^{(0)}_{\mathbf{H}_2\mathbf{H}_1}(t,f)\right|$ is 0.045, which means that the approximation is quite good.

Approximate eigenfunctions and eigenvalues. The response of an LTI system with transfer function $G(f)$ to a complex sinusoid $e^{j2\pi f_0 t}$ is $G(f_0)\,e^{j2\pi f_0 t}$, and the response of an LFI system with temporal transfer function $w(t)$ to a Dirac impulse $\delta(t-t_0)$ is given by $w(t_0)\,\delta(t-t_0)$. Hence, complex sinusoids and Dirac impulses are the eigenfunctions of LTI and LFI systems, respectively, with the eigenvalues given by corresponding values of the transfer function, $G(f_0)$ and $w(t_0)$. In contrast, the eigenfunctions of general LTV systems are not localized or structured in any sense. However, for underspread LTV systems, well TF localized functions are approximate eigenfunctions and the GWS $L^{(\alpha)}_{\mathbf{H}}(t,f)$ constitutes an approximate eigenvalue

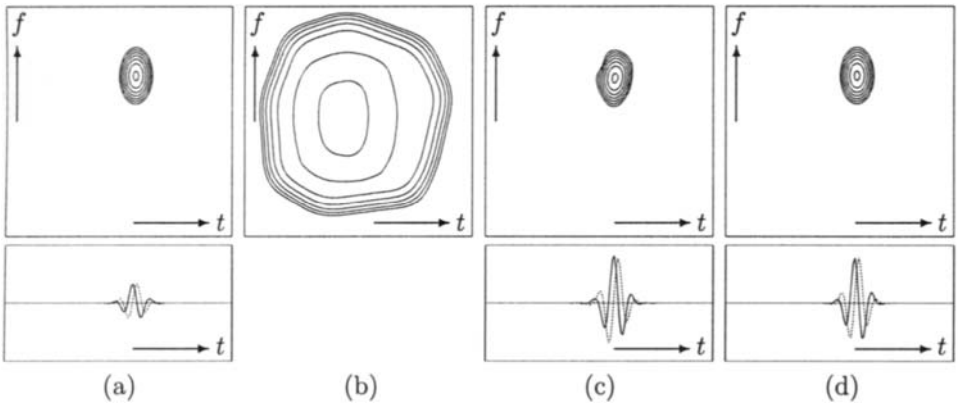


Fig. 4.7.4: Eigenfunction/eigenvalue approximation of the Weyl symbol (GWS with $\alpha = 0$) of an underspread LTV system: (a) Wigner distribution [12] (top) and real and imaginary parts (bottom) of input signal $s_{t_0,f_0}(t)$, (b) Weyl symbol of $\mathbf{H}$, (c) output signal $(\mathbf{H}s_{t_0,f_0})(t)$, and (d) input signal $s_{t_0,f_0}(t)$ multiplied by $L_{\mathbf{H}}(t_0, f_0)$. Note the similarity of (c) and (d).

distribution over the TF plane (for related results see [11] and Article 13.3). Indeed, consider the following family of signals,

$$s_{t_0,f_0}(t) = s(t - t_0)\, e^{j2\pi f_0 t},$$

where $s(t)$ is a signal well TF localized about $t = 0$ and $f = 0$. Evidently, $s_{t_0,f_0}(t)$ is well TF localized about $t = t_0$ and $f = f_0$. For an underspread LTV system $\mathbf{H}$, one can show [5]

$$\left(\mathbf{H}\, s_{t_0,f_0}\right)(t) \;\approx\; L_{\mathbf{H}}^{(\alpha)}(t_0, f_0)\, s_{t_0,f_0}(t)\,, \tag{4.7.7}$$

i.e., the signals $s_{t_0,f_0}(t)$ are "approximate eigenfunctions" of $\mathbf{H}$, with the associated "approximate eigenvalues" given by the GWS values $L_{\mathbf{H}}^{(\alpha)}(t_0, f_0)$. Eq. (4.7.7) shows that a signal well TF localized about some TF point is passed by $\mathbf{H}$ nearly undistorted; it is merely weighted by the GWS value at that point, thus corroborating the GWS's interpretation as a *TF transfer function*. An example illustrating the approximation (4.7.7) for the case $\alpha = 0$ is shown in Fig. 4.7.4. In this example, the normalized error $\left\|\mathbf{H}s_{t_0,f_0} - L_{\mathbf{H}}^{(0)}(t_0, f_0)\, s_{t_0,f_0}\right\|^2 / \left\|\mathbf{H}s_{t_0,f_0}\right\|^2$ (with $\|x\|^2 = \int_{-\infty}^{\infty} |x(t)|^2\, dt$) is 0.06.

Approximate uniqueness of the GWS. For an underspread LTV system $\mathbf{H}$, it can furthermore be shown [5] that the GWS is only weakly dependent on the parameter α. That is,

$$L_{\mathbf{H}}^{(\alpha_1)}(t, f) \approx L_{\mathbf{H}}^{(\alpha_2)}(t, f)$$

for moderate values of $|\alpha_1 - \alpha_2|$. This means that the TF transfer function provided by the GWS is approximately *unique*. In particular, the Weyl symbol $L_{\mathbf{H}}^{(0)}(t, f)$, Zadeh's time-varying transfer function $L_{\mathbf{H}}^{(1/2)}(t, f)$, and Bello's frequency-dependent

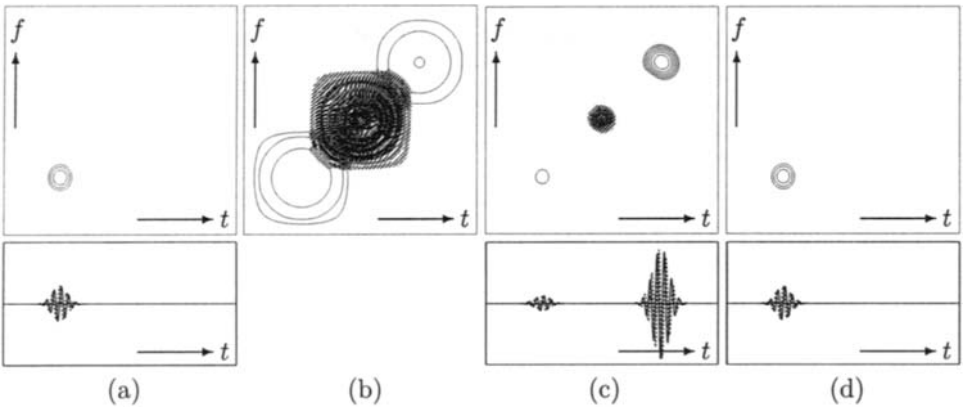


Fig. 4.7.5: Violation of the eigenfunction/eigenvalue approximation of the Weyl symbol in the case of an overspread LTV system: (a) Wigner distribution (top) and real and imaginary parts (bottom) of input signal $s_{t_0,f_0}(t)$, (b) Weyl symbol of $\mathbf{H}$, (c) output signal $(\mathbf{H}s_{t_0,f_0})(t)$, and (d) input signal $s_{t_0,f_0}(t)$ multiplied by $L_{\mathbf{H}}(t_0,f_0)$. This figure should be compared with Fig. 4.7.4; note that now the signals in (c) and (d) are very different. The overspread character of $\mathbf{H}$ is indicated by the rapid oscillation of the Weyl symbol in (b) and by the fact—evident upon comparison of (a) and (c)—that the signal $s_{t_0,f_0}(t)$ is partly TF shifted by the system $\mathbf{H}$.

modulation function $L_{\mathbf{H}}^{(-1/2)}(t,f)$ will be approximately equivalent for an underspread LTV system $\mathbf{H}$.

Discussion. As mentioned before, the above approximate relations (more can be found in [5, 10]) extend analogous (exact) relations satisfied by the conventional transfer function of LTI/LFI systems. In this sense, the GWS is an approximate TF transfer function of underspread LTV systems. As a mathematical underpinning of these approximations, explicit upper bounds on the associated approximation errors have been developed [5,10]. These bounds are formulated in terms of the GSF parameters $m_{\mathbf{H}}^{(\phi)}$, $m_{\mathbf{H}}^{(k,l)}$ or $\sigma_{\mathbf{H}}$ defined in Section 4.7.4. If specific such parameters are small (indicating that $\mathbf{H}$ is underspread), the upper bounds on specific approximation errors are small and thus the respective approximation will be good.

On the other hand, we caution that the above approximations and, thus, the GWS's interpretation as a TF transfer function are *not* valid for overspread LTV systems. This is illustrated in Fig. 4.7.5.

4.7.6 Summary and Conclusions

While general linear, time-varying systems are fairly difficult to work with, we have shown that for the practically important class of *underspread* systems a very simple and intuitively appealing time-frequency transfer function calculus can be developed. Indeed, the generalized Weyl symbol can be used as an approximate *time-frequency transfer function* in a similar way as the conventional transfer function of time-invariant systems. Applications of this time-frequency transfer function

calculus to nonstationary signal analysis and processing are considered in Articles 9.4, 11.1, and 12.4. Finally, an extension to *random* systems/channels is considered in Article 9.5.

References

[1] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science.* New York: Springer, 2nd ed., 1982.

[2] W. Kozek, "On the generalized Weyl correspondence and its application to time-frequency analysis of linear time-varying systems," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 167–170, Victoria, BC, 4–6 October 1992.

[3] G. B. Folland, *Harmonic Analysis in Phase Space.* No. 122 in ANNALS OF MATHEMATICS STUDIES, Princeton, NJ: Princeton University Press, 1989.

[4] W. Kozek, "Time-frequency signal processing based on the Wigner-Weyl framework," *Signal Processing*, vol. 29, pp. 77–92, October 1992.

[5] 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. of Mathematical Physics*, vol. 39, pp. 4041–4070, August 1998. Special Issue on Wavelet and Time-Frequency Analysis.

[6] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 318–331, February 1994.

[7] L. A. Zadeh, "Frequency analysis of variable networks," *Proc. IRE*, vol. 3A-8, pp. 291–299, March 1950.

[8] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Communication Systems*, vol. 11, pp. 360–393, December 1963.

[9] 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.), pp. 59–133, Amsterdam: Elsevier, 1997.

[10] W. Kozek, "On the transfer function calculus for underspread LTV channels," *IEEE Trans. Signal Processing*, vol. 45, pp. 219–223, January 1997.

[11] S. Barbarossa and A. Scaglione, "On the capacity of linear time-varying channels," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'99)*, pp. 2627–2630, Phoenix, AZ, 15–19 March 1999.

[12] W. Mecklenbräuker and F. Hlawatsch, eds., *The Wigner Distribution — Theory and Applications in Signal Processing.* Amsterdam: Elsevier, 1997.

4.8 WIGNER DISTRIBUTION AND FRACTIONAL FOURIER TRANSFORM [0]

4.8.1 Time-Frequency Representations

A Fourier transformation maps a one-dimensional time signal $x(t)$ into a one-dimensional spectrum denoted (in the notation to be explained in this article) by $X_{\pi/2}(f)$. A shortcoming of the Fourier transform (FT) is that it does not give an obvious indication of the time location of the spectral components, which is an important characteristic of non-stationary or time-varying signals. To describe such a signal we may use a time-frequency representation (TFR), which maps a one-dimensional time signal into a two-dimensional function of time *and* frequency. In this article we consider the fractional FT, which belongs to the class of *linear* TFRs, and establish its connection to the Wigner distribution (WD), which is one of the most widely used *quadratic* TFRs in electrical engineering. In particular we use the Radon-Wigner transform (RWT), which relates projections of TFRs to the squared modulus of the fractional FT, and we show that all frequently used moments of the WD can be obtained from the RWT in an optimal way. Moreover, we introduce some fractional FT moments that can be useful for signal analysis.

4.8.2 Wigner Distribution and Ambiguity Function

The *Wigner distribution* is defined (see Section 2.1.3) as

$$\begin{aligned} W_x(t,f) &= \int_{-\infty}^{\infty} x(t+\tfrac{1}{2}\tau)\, x^*(t-\tfrac{1}{2}\tau)\, \exp(-j2\pi f\tau)\, d\tau \\ &= \int_{-\infty}^{\infty} X_{\pi/2}(f+\tfrac{1}{2}\nu)\, X^*_{\pi/2}(f-\tfrac{1}{2}\nu)\, \exp(j2\pi\nu t)\, d\nu, \end{aligned} \tag{4.8.1}$$

where $x(t)$ is a time signal and $X_{\pi/2}(f)$ its FT. The WD is always real-valued, but not necessarily positive; it preserves time and frequency shifts, and satisfies the marginal properties, which means that the frequency and time integrals of the WD, $\int W_x(t,f)\, df$ and $\int W_x(t,f)\, dt$, correspond to the signal's instantaneous power $|x(t)|^2$ and its spectral energy density $|X_{\pi/2}(f)|^2$, respectively. The WD can roughly be considered as the signal's energy distribution over the time-frequency plane, although the uncertainty principle prohibits the interpretation as a point time-frequency energy density.

If in Eqs. (4.8.1) the integrations are carried out over the common variable (t or f) instead of over the difference variable (τ or ν), we get the *ambiguity function* $A_x(\tau,\nu)$, which is related to the WD by means of a Fourier transformation [1]:

$$A_x(\tau,\nu) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x(t,f)\, \exp[-j2\pi(\nu t - f\tau)]\, dt\, df. \tag{4.8.2}$$

[0]Authors: **Tatiana Alieva** and **Martin J. Bastiaans**, Technische Universiteit Eindhoven, Faculteit Elektrotechniek, Postbus 513, 5600 MB Eindhoven, Netherlands (M.J.Bastiaans@tue.nl).
Reviewers: G. Faye Boudreaux-Bartels and LJubiša Stanković.

The ambiguity function (AF) is another widely-used quadratic TFR.

4.8.3 Fractional Fourier Transform

The *fractional Fourier transform* of a signal $x(t)$ is defined as [2,3]

$$X_\alpha(u) = \mathcal{R}^\alpha\left[x(t)\right](u) = \int_{-\infty}^{\infty} K(\alpha,t,u)\, x(t)\, dt, \tag{4.8.3}$$

where the kernel $K(\alpha,t,u)$ is given by

$$K(\alpha,t,u) = \frac{\exp(j\frac{1}{2}\alpha)}{\sqrt{j\sin\alpha}}\, \exp\left(j\pi\frac{(t^2+u^2)\cos\alpha - 2ut}{\sin\alpha}\right) = K(\alpha,u,t). \tag{4.8.4}$$

The fractional FT with parameter α can be considered as a generalization of the ordinary FT; thus the fractional FT for $\alpha = \frac{1}{2}\pi$ and $\alpha = -\frac{1}{2}\pi$ reduces to the ordinary and inverse FT, respectively. For $\alpha = 0$ the fractional FT corresponds to the identity operation, $X_0(u) = \mathcal{R}^0[x(t)](u) = x(u)$, and for $\alpha = \pm\pi$ to the axis reversal operation, $X_{\pm\pi}(u) = \mathcal{R}^{\pm\pi}[x(t)](u) = x(-u)$. With respect to the parameter α, the fractional FT is continuous, periodic [$\mathcal{R}^{\alpha+2\pi n} = \mathcal{R}^\alpha$, with n an integer] and additive [$\mathcal{R}^\alpha\mathcal{R}^\beta = \mathcal{R}^{\alpha+\beta}$], and has the symmetry relation $\mathcal{R}^\alpha[x^*(t)](u) = \{\mathcal{R}^{-\alpha}[x(t)](u)\}^*$. The inverse fractional FT can thus be written as

$$x(t) = \mathcal{R}^{-\alpha}\left[X_\alpha(u)\right](t) = \int_{-\infty}^{\infty} K(-\alpha,t,u)\, X_\alpha(u)\, du. \tag{4.8.5}$$

Since the Hermite-Gauss functions $\Psi_n(t) = (2^{n-1/2}n!)^{-1/2}\exp(-\pi t^2)\, H_n(\sqrt{2\pi}t)$, with $H_n(t)$ the Hermite polynomials, are eigenfunctions of the fractional FT with eigenvalues $\exp(-jn\alpha)$, and since they compose a complete orthonormal set, it is possible to write the fractional FT kernel in the alternative form

$$K(\alpha,t,u) = \sum_{n=0}^{\infty} \exp(-jn\alpha)\, \Psi_n(t)\, \Psi_n(u). \tag{4.8.6}$$

The important property of the fractional FT, which allows us to establish a connection between it and the WD, the AF, and other members of Cohen's class [1] of quadratic TFRs, is that a fractional Fourier transformation produces a *rotation* of these functions in the time-frequency plane [2]:

$x(t)$	$\longleftrightarrow$	$W_x(t,f)$ and $A_x(\tau,\nu)$
$\downarrow$ fractional FT		$\downarrow$ rotation of WD and AF
$X_\alpha(t) = \mathcal{R}^\alpha[x]$	$\longleftrightarrow$	$W_{X_\alpha}(t,f) = W_x(t\cos\alpha - f\sin\alpha, t\sin\alpha + f\cos\alpha)$ and $A_{X_\alpha}(\tau,\nu) = A_x(\tau\cos\alpha - \nu\sin\alpha, \tau\sin\alpha + \nu\cos\alpha)$.

Table 4.8.1: Fractional Fourier transform properties

Linearity:

$$\mathcal{R}^{\alpha}\left[ax(t)+by(t)\right](u)=a\mathcal{R}^{\alpha}\left[x(t)\right](u)+b\mathcal{R}^{\alpha}\left[y(t)\right](u)$$

Parseval's equality:

$$\int_{-\infty}^{\infty} x(t)\,y^{*}(t)\,dt=\int_{-\infty}^{\infty} X_{\alpha}(u)\,Y_{\alpha}^{*}(u)\,du$$

Shift theorem (real τ):

$$\mathcal{R}^{\alpha}\left[x(t-\tau)\right](u)=X_{\alpha}(u-\tau\cos\alpha)\,\exp\left[j\pi\sin\alpha\left(\tau^{2}\cos\alpha-2u\tau\right)\right]$$

Modulation theorem (real ν):

$$\mathcal{R}^{\alpha}\left[x(t)\,\exp(j2\pi\nu t)\right](u)=X_{\alpha}(u-\nu\sin\alpha)\,\exp\left[-j\pi\cos\alpha\left(\nu^{2}\sin\alpha-2u\nu\right)\right]$$

Scaling theorem (real c and β, where $\tan\beta=c^{2}\tan\alpha$):

$$\mathcal{R}^{\alpha}\left[x(ct)\right](u)=\sqrt{\frac{\cos\beta}{\cos\alpha}}\,\exp\left(j\frac{\alpha-\beta}{2}\right)\exp\left[j\pi u^{2}\cot\alpha\left(1-\frac{\cos^{2}\beta}{\cos^{2}\alpha}\right)\right]\times X_{\beta}\left(\frac{u}{c}\frac{\sin\beta}{\sin\alpha}\right),$$

Hence, we conclude that $W_{X_{\alpha}}(u,v)=W_{x}(t,f)$, where the coordinates (u,v) in the rotated frame are related to (t,f) via the matrix relationship

$$\begin{pmatrix} u \\ v \end{pmatrix}=\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} t \\ f \end{pmatrix}. \tag{4.8.7}$$

A similar relation holds for the AF.

The main properties of the fractional FT are listed in Table 4.8.1, and the fractional FT of some common functions are given in Table 4.8.2.

4.8.4 Fractional Power Spectrum and Radon-Wigner Transform

If we introduce the *fractional power spectrum* $|X_{\alpha}(t)|^{2}$ as the squared modulus of the corresponding fractional FT, we obtain that these fractional power spectra are the projections of the WD upon a direction at an angle α in the time-frequency plane [4],

$$|X_{\alpha}(t)|^{2}=\int_{-\infty}^{\infty} W_{X_{\alpha}}(t,f)\,df=\int_{-\infty}^{\infty} W_{x}(t\cos\alpha-f\sin\alpha,\,t\sin\alpha+f\cos\alpha)\,df, \tag{4.8.8}$$

Table 4.8.2: Fractional Fourier transforms of some common functions

$x(t)$	$X_\alpha(u)$
$\delta(t-\tau)$	$\frac{\exp(j\frac{1}{2}\alpha)}{\sqrt{j\sin\alpha}}\exp\left[j\pi\frac{(\tau^2+u^2)\cos\alpha-2u\tau}{\sin\alpha}\right]$
$\exp(j2\pi t\nu)$	$\frac{\exp(j\frac{1}{2}\alpha)}{\sqrt{\cos\alpha}}\exp\left[-j\pi(\nu^2+u^2)\tan\alpha+j2\pi u\nu\sec\alpha\right]$
$\exp\left(jc\pi t^2\right)$	$\frac{\exp(j\frac{1}{2}\alpha)}{\sqrt{\cos\alpha+c\sin\alpha}}\exp\left(j\pi u^2\frac{c-\tan\alpha}{1+c\tan\alpha}\right)$
$H_n(\sqrt{2\pi}t)\exp(-\pi t^2)$	$H_n(\sqrt{2\pi}u)\exp(-\pi u^2)\exp(-jn\alpha)$, H_n are the Hermite polynomials
$\exp\left(-c\pi t^2\right)\quad(c\geq 0)$	$\frac{\exp(j\frac{1}{2}\alpha)}{\sqrt{\cos\alpha+jc\sin\alpha}}\exp\left[\pi u^2\frac{j(c^2-1)\cot\alpha-c\csc^2\alpha}{c^2+\cot^2\alpha}\right]$

and that they are related to the AF by a Fourier transformation:

$$|X_\alpha(t)|^2=\int_{-\infty}^{\infty}A_x(f\sin\alpha,-f\cos\alpha)\exp(-j2\pi ft)\,df. \tag{4.8.9}$$

The set of fractional power spectra for the angles $\alpha\in[0,\pi)$ is called the *Radon-Wigner transform*, because it defines the Radon transform [5] of the WD. The WD can be obtained from the RWT by applying an inverse Radon transformation. Note also that the AF can be reconstructed from the RWT by a simple inverse Fourier transformation, see Eq. (4.8.9), and that other members of Cohen's class of TFRs can be constructed subsequently.

The RWT can be considered as a quadratic TFR of $x(t)$, which has very advantageous properties. It is positive, invertible up to a constant phase factor, and ideally combines the concepts of the instantaneous power $|x(t)|^2$ and the spectral energy density $|X_{\pi/2}(f)|^2$. The association of the RWT with the power distributions allows its direct measurement in optics and quantum mechanics, which opens new perspectives for optical signal processing and quantum state characterization.

4.8.5 Fractional Fourier Transform Moments

The application of the different TFRs often depends on how informative their moments are and how easily these moments can be measured or calculated. The established connection between the WD and the RWT permits to find an optimal

way for the calculation of the known WD moments and to introduce *fractional FT moments* that can be useful for signal analysis.

By analogy with time and frequency moments [6],

$$\int_{-\infty}^{\infty} t^n \left|x(t)\right|^2 dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t^n W_x(t,f)\, dt\, df$$

$$\int_{-\infty}^{\infty} f^n \left|X_{\pi/2}(f)\right|^2 df = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^n W_x(t,f)\, dt\, df = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t^n W_{X_{\pi/2}}(t,f)\, dt\, df,$$

the fractional FT moments can be introduced:

$$\frac{1}{(4\pi j)^m}\int_{-\infty}^{\infty} t^n \left(\frac{\partial}{\partial t_1} - \frac{\partial}{\partial t_2}\right)^m X_\alpha(t_1)\, X_\alpha^*(t_2)\Big|_{t_1=t_2=t} dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t^n f^m W_{X_\alpha}(t,f)\, dt\, df.$$

The *zero-order* fractional FT moment E,

$$E = \int_{-\infty}^{\infty} \left|X_\alpha(t)\right|^2 dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{X_\alpha}(t,f)\, dt\, df = \int_{-\infty}^{\infty} \left|x(t)\right|^2 dt, \qquad (4.8.10)$$

is invariant under fractional Fourier transformation, which expresses the energy conservation law of a unitary transformation, also known as Parseval's relation.

The normalized *first-order* fractional FT moment m_α,

$$m_\alpha = \frac{1}{E}\int_{-\infty}^{\infty} t\left|X_\alpha(t)\right|^2 dt = \frac{1}{E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t\, W_{X_\alpha}(t,f)\, dt\, df, \qquad (4.8.11)$$

is related to the center of gravity of the fractional power spectrum. One can write the simple connection

$$m_\alpha = m_0 \cos\alpha + m_{\pi/2}\sin\alpha \qquad (4.8.12)$$

between the first-order fractional FT moments. It is easy to see that the pair $(m_\alpha, m_{\alpha+\pi/2})$ is connected to $(m_0, m_{\pi/2})$ through the rotation transformation (4.8.7), and that $m_\alpha^2 + m_{\alpha+\pi/2}^2 = m_0^2 + m_{\pi/2}^2$ is invariant under fractional Fourier transformation. The fractional domain corresponding to the zero-centered fractional power spectrum can be found as $\tan\alpha = -m_0/m_{\pi/2}$.

The normalized *second-order central* fractional FT moment p_α,

$$\begin{aligned} p_\alpha &= \frac{1}{E}\int_{-\infty}^{\infty} (t-m_\alpha)^2 \left|X_\alpha(t)\right|^2 dt = \frac{1}{E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (t-m_\alpha)^2\, W_{X_\alpha}(t,f)\, dt\, df \\ &= \frac{1}{E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t^2\, W_{X_\alpha}(t,f)\, dt\, df - m_\alpha^2, \end{aligned} \qquad (4.8.13)$$

is related to the effective width of the signal in the fractional FT domain. The

normalized *mixed* second-order central fractional FT moment μ_α is given by

$$\begin{aligned}\mu_\alpha &= \frac{1}{E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(t-m_\alpha)\,(f-m_{\alpha+\pi/2})\,W_{X_\alpha}(t,f)\,dt\,df \\ &= \frac{1}{E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t\,f W_{X_\alpha}(t,f)\,dt\,df - m_\alpha m_{\alpha+\pi/2} \\ &= \frac{1}{4\pi jE}\int_{-\infty}^{\infty}\left[\frac{\partial X_\alpha(t)}{\partial t}\,X_\alpha^*(t) - X_\alpha(t)\,\frac{\partial X_\alpha^*(t)}{\partial t}\right] t\,dt - m_\alpha m_{\alpha+\pi/2}.\end{aligned} \tag{4.8.14}$$

The following relationships between the second-order fractional FT moments hold:

$$\begin{aligned}p_\alpha &= p_0\cos^2\alpha + p_{\pi/2}\sin^2\alpha + \mu_0\sin 2\alpha \\ \mu_\alpha &= -\tfrac{1}{2}(p_0 - p_{\pi/2})\sin 2\alpha + \mu_0\cos 2\alpha.\end{aligned} \tag{4.8.15}$$

In general all second-order moments p_α and μ_α can be obtained from any three second-order moments p_α taken for three different angles α from the region $[0,\pi)$. The mixed moment μ_0, for instance, can be expressed as $\mu_0 = -\frac{1}{2}(p_0+p_{\pi/2})+p_{\pi/4}$.

From Eqs. (4.8.15) we conclude that the sum of the signal widths in the time and the frequency domain is invariant under fractional Fourier transformation:

$$p_\alpha + p_{\alpha+\pi/2} = p_0 + p_{\pi/2}. \tag{4.8.16}$$

Note also that the fractional domain corresponding to the extremum signal width p_α can be found by solving the equation $\tan 2\alpha = 2\mu_0/(p_0 - p_{\pi/2})$, i.e., $\mu_\alpha = 0$. Due to the invariance relationship (4.8.16), the solution of this equation corresponds to the domain with the smallest p_α and the largest $p_{\alpha+\pi/2}$, or vice versa.

For the product of the signal widths we find

$$p_\alpha p_{\alpha+\pi/2} = p_0 p_{\pi/2} + \tfrac{1}{4}\left[(p_0 - p_{\pi/2})^2 - 4\mu_0^2\right]\sin^2 2\alpha - \tfrac{1}{2}\mu_0\,(p_0 - p_{\pi/2})\sin 4\alpha, \tag{4.8.17}$$

which expression is, in general, not invariant under fractional Fourier transformation; invariance does occur, for instance, in the case of eigenfunctions of the Fourier transformation, $g(t)$, say, for which $G_{\alpha+\pi/2}(t) = \exp(-jn\frac{1}{2}\pi)\,G_\alpha(t)$ [cf. the eigenvalues $\exp(-jn\alpha)$ in Eq. (4.8.6)], in which case $p_\alpha = p_0 = p_{\pi/2}$ and $\mu_\alpha = 0$. Note that, due to the uncertainty principle, we have $p_\alpha p_{\alpha+\pi/2} \geq \frac{1}{4}$. The fractional FT domain where the product $p_\alpha p_{\alpha+\pi/2}$ has an extremum value can be found by solving the equation $\tan 4\alpha = 4\mu_0(p_0 - p_{\pi/2})/[(p_0 - p_{\pi/2})^2 - 4\mu_0^2]$.

From Eqs. (4.8.15) we finally conclude the following property for the mixed moment μ_α:

$$\mu_\alpha = \tfrac{1}{2}(p_{\alpha+\pi/4} - p_{\alpha-\pi/4}) = -\mu_{\alpha\pm\pi/2}. \tag{4.8.18}$$

The fractional FT moments may be helpful in the search for the most appropriate fractional domain to perform a filtering operation; in the special case of noise that is equally distributed throughout the time-frequency plane, for instance, the fractional domain with the smallest signal width is then evidently the most preferred one.

Instead of *global* moments, which we considered above, one can consider *local* fractional FT moments, which are related to such signal characteristics as the instantaneous power and instantaneous frequency (for $\alpha = 0$) or the spectral energy density and group delay (for $\alpha = \frac{1}{2}\pi$) in the different fractional FT domains.

The local frequency in the fractional FT domain with parameter α is defined as

$$U_{X_\alpha}(t) = \frac{\int_{-\infty}^{\infty} f\, W_{X_\alpha}(t,f)\, df}{\int_{-\infty}^{\infty} W_{X_\alpha}(t,f)\, df} = \frac{1}{2\,|X_\alpha(t)|^2} \int_{-\infty}^{\infty} \left.\frac{\partial |X_\beta(\tau)|^2}{\partial \beta}\right|_{\beta=\alpha} \operatorname{sgn}(\tau - t)\, d\tau. \tag{4.8.19}$$

The local frequency $U_{X_\alpha}(t)$ is related to the phase $\varphi_\alpha(t) = \arg X_\alpha(t)$ of the fractional FT $X_\alpha(t)$ through $U_{X_\alpha}(t) = (1/2\pi) d\varphi_\alpha(t)/dt$. This implies that the derivative of the fractional power spectra with respect to the angle α defines the local frequency in the fractional domain, and that it can be used for solving the phase retrieval problem by measuring intensity functions only.

We finally mention the relationship between the central local fractional second-order moment and the instantaneous power in the fractional FT domain:

$$V_{X_\alpha}(t) = \frac{\int_{-\infty}^{\infty} [f - U_{X_\alpha}(t)]^2\, W_{X_\alpha}(t,f)\, df}{\int_{-\infty}^{\infty} W_{X_\alpha}(t,f)\, df} = -\tfrac{1}{4} \frac{d^2 \ln |X_\alpha(t)|^2}{dt^2}. \tag{4.8.20}$$

We conclude that all frequently used moments of the WD can be obtained from the RWT.

4.8.6 Applications

The fractional FT and the WD are applied in such diverse fields as quantum mechanics, optics, and signal processing [1, 4–7].

The wide application of the fractional FT in optics is based on the fact that – in the paraxial approximation of the scalar diffraction theory – it describes the optical field evolution during propagation through a quadratic refractive index (lens-like) medium. The RWT, associated with the intensity distributions, is used in particular for the reconstruction of the WD and subsequently of the complex field amplitude (in the case of coherent light) or the two-point correlation function (in the case of partially coherent light).

In signal processing the RWT was primarily developed for detection and classification of multi-component linear FM signals in noise [7]. Since the fractional FT of the chirp-type signal $\exp(-j\pi t^2/\tan\beta)$ reads

$$\exp(j\tfrac{1}{2}\alpha)\, \sqrt{\sin\beta/\sin(\beta-\alpha)}\, \exp[-j\pi u^2/\tan(\beta-\alpha)],$$

it becomes proportional to a Dirac-function $\delta(u)$ for $\alpha \to \beta$, and it can be detected as a local maximum on the RWT map. Analogously, in order to remove chirp-type

noise, a notch filter, which minimizes the signal information loss, can be placed at the proper point of the corresponding fractional FT domain [4].

Instead of performing, as usual, filtering operations in the frequency or the time domain, it can be done in a more appropriate fractional domain, for instance, the one that corresponds to the best signal/noise time-frequency separation [4].

The complexity of computation for the fractional FT is $O(N \log N)$ [8], where N is the time-bandwidth product of the signal.

4.8.7 Summary and Conclusions

We have described the relationship between the fractional Fourier transform and the Wigner distribution by using the Radon-Wigner transform, which is a set of projections of the Wigner distribution as well as a set of squared moduli of the fractional Fourier transform. We have introduced the concept of fractional Fourier transform moments and have proposed a way for the calculation of the well-known global and local moments of the Wigner distribution, based on the knowledge of a few fractional power spectra. The application of the results in optics and signal processing has been discussed briefly. Article 5.8 further explores the relationship between the fractional Fourier transform and quadratic TFDs.

References

[1] G. F. Boudreaux-Bartels, "Mixed time-frequency signal transformations," in *The Transforms and Applications Handbook* (A. D. Poularikas, ed.), ch. 12, Boca Raton, FL: CRC Press, 1st ed., 1996.

[2] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Processing*, vol. 42, pp. 3084–3091, November 1994.

[3] H. M. Ozaktas, M. A. Kutay, and D. Mendlovic, "Introduction to the fractional Fourier transform and its applications," in *Advances in Imaging & Electron Physics—Vol. 106* (P. W. Hawkes, B. Kazan, and T. Mulvey, eds.), ch. 4, pp. 239–291, San Diego: Academic Press, February 1999.

[4] A. W. Lohmann, D. Mendlovic, and Z. Zalevsky, "Fractional transformations in optics," in *Progress in Optics—Vol. 38* (E. Wolf, ed.), ch. 4, pp. 263–342, Amsterdam: Elsevier, 1998.

[5] S. R. Deans, "Radon and Abel transforms," in *The Transforms and Applications Handbook* (A. D. Poularikas, ed.), ch. 8, Boca Raton, FL: CRC Press, 1st ed., 1996.

[6] W. Mecklenbräuker and F. Hlawatsch, eds., *The Wigner Distribution —Theory and Applications in Signal Processing*. Amsterdam: Elsevier, 1997.

[7] J. C. Wood and D. T. Barry, "Tomographic time-frequency analysis and its application toward time-varying filtering and adaptive kernel design for multicomponent linear FM signals," *IEEE Trans. Signal Processing*, vol. 42, pp. 2094–2104, 1994.

[8] H. M. Ozaktas, O. Arıkan, M. A. Kutay, and G. Bozdağı, "Digital computation of the fractional Fourier transform," *IEEE Trans. Signal Processing*, vol. 44, pp. 2141–2150, September 1996.

4.9 GABOR SPECTROGRAM[0]

4.9.1 Power Spectrum

For more than a century, the Fourier transform based power spectrum has been one of the most popular and powerful tools for engineers/scientists to analyze and characterize the natural events whose features are not obvious in the time domain. The success of the application of the Fourier-transform-based power spectrum largely hinges on whether the associated signal's frequency contents evolve over time. As illustrated in Section 1.1.1, the conventional Fourier transform is only suitable for a signal whose frequencies do not change during the period of observation. It is inadequate for signals whose frequencies are not constant, such as those encountered in Doppler phenomena and seismology. To overcome this drawback, we may extend the classical time-independent power spectrum to a time-dependent power spectrum.

4.9.2 Gabor Spectrogram

For a signal $s(t) = a(t)e^{j\phi(t)}$ with a Fourier transform $S(f)$, intuitively, a desirable time-dependent power spectrum $\rho_s(t, f)$ should possess the following properties:

$$\int_{-\infty}^{+\infty} \rho_s(t, f)dt = |S(f)|^2 \tag{4.9.1}$$

$$\int_{-\infty}^{+\infty} \rho_s(t, f)df = |s(t)|^2 = |a(t)|^2 \tag{4.9.2}$$

$$\frac{\int_{-\infty}^{+\infty} f\rho_s(t, f)df}{\int_{-\infty}^{+\infty} \rho_s(t, f)df} = \frac{\int_{-\infty}^{+\infty} f\rho_s(t, f)df}{|a(t)|^2} = \phi'(t) \tag{4.9.3}$$

where the first derivative of the phase, $\phi'(t)$, called the (angular) instantaneous frequency, describes how the signal's frequencies evolve over time [see Article 1.3]. It is worth noting that properties (4.9.1) to (4.9.3) are all related to the *average* of $\rho_s(t, f)$.

Over the years, many approaches have been proposed to evaluate the time-dependent spectrum. Among them, the most widely used may be the spectrogram based on the STFT (short-time Fourier transform). It is simple, but it cannot simultaneously meet the properties (4.9.1) to (4.9.3). On the other hand, the Wigner-Ville distribution satisfies conditions (4.9.1) to (4.9.3), but its application has been limited due to the so-called crossterm interference. One effective approach to reducing the crossterm interference is to apply a two-dimensional lowpass filter to the Wigner-Ville distribution. The resulting class of representations is called Cohen's class (or the quadratic class; see the note on p. 68). It has been found that a certain type

[0]Author: **Shie Qian**, National Instruments Corp., Austin, TX 78759, USA (shie.qian@ni.com).
Reviewers: X.-G. Xia and Flemming Munk.

of two-dimensional filter can rearrange crossterms without destroying the desirable properties possessed by the Wigner-Ville distribution (see [1] and Article 4.2).

As an alternative, we can also first write the signal in terms of the Gabor expansion and then take the Wigner-Ville distribution. The resulting Wigner-Ville distribution then is a superposition of concentrated, symmetrical, and oscillating two-dimensional Gaussian functions. The energy (computed by averaging a two-dimensional Gaussian function) contained in each individual two-dimensional Gaussian function is inversely proportional to the rate of its oscillation. The higher the oscillation, the smaller the energy (or average). Since the desirable properties are all computed by averaging $\rho_s(t,f)$, the highly oscillated Gaussian functions have limited influence on the properties (4.9.1) to (4.9.3). On the other hand, the highly oscillated Gaussian functions have significant contribution to the crossterm interference. Consequently, we can remove the highly oscillated terms, without significantly altering desirable properties, to lessen the crossterm interference. The resulting representation was named the Gabor spectrogram to distinguish it from the spectrogram computed from the STFT [2,3]. In what follows, we shall give a more detailed discussion about the Gabor spectrogram.

For a given signal $s(t)$, the Wigner-Ville distribution is defined as

$$W_s(t,f) = \int_{-\infty}^{+\infty} s(t+\frac{\tau}{2})s^*(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau \tag{4.9.4}$$

When $s(t)$ is the time- and frequency-shifted Gaussian function $h_{m,n}(t)$, i.e.,

$$s(t) = h_{m,n}(t) = (\alpha\pi)^{-0.25}\exp\{-\frac{(t-mT)^2}{2\alpha} + j2\pi n\Omega t\} \tag{4.9.5}$$

the corresponding Wigner-Ville distribution has the form

$$W_h(t,f) = 2\exp\{-\frac{(t-mT)^2}{\alpha} - \alpha(2\pi)^2(f-n\Omega)^2\} \tag{4.9.6}$$

which is non-negative and symmetric about the point $(mT, n\Omega)$. The contour plot of (4.9.6) forms conic ellipses. Because expression (4.9.6) is a WVD, it satisfies conditions (4.9.1) to (4.9.3). Since at any time instant there is only one frequency tone, we can consider the signal in (4.9.5) as a monotone signal.

As mentioned in the preceding section, the major deficiency of the Wigner-Ville distribution is the crossterm interference. For example, when $s(t)$ contains multiple monotone signals, such as multiple Gaussians with different time- and frequency-shifts, then we will observe some components which are not expected. In this case, the resulting Wigner-Ville distribution still possesses the properties (4.9.1) to (4.9.3), but the unwanted crossterm can destroy the features in which we are interested. Since the crossterm always occurs midway between two monotone signals, one way of reducing the crossterm is to decompose the analyzed signal as monotone components.

For example, we can decompose an arbitrary signal $s(t)$ as the sum of monotone time- and frequency-shifted Gaussian elementary functions $h_{m,n}(t)$ in (4.9.5), via the Gabor expansion, i.e.,

$$s(t) = (\alpha\pi)^{-0.25} \sum_m \sum_n c_{m,n} h_{m,n}(t) \tag{4.9.7}$$

where $c_{m,n}$ is traditionally named the Gabor coefficient. It can be computed by the STFT with a dual function of $\gamma(t)$. If the dual window function $\gamma(t)$ is concentrated at $(0,0)$, the origin of the joint time-frequency plane, then the Gabor coefficients $c_{m,n}$ describe the signal's behavior in the vicinity of $(mT, n\Omega)$.

Taking the Wigner-Ville distribution of both sides of (4.9.7) yields

$$W_s(t,f) = \sum_{m,m'} \sum_{n,n'} c_{m,n} c^*_{m',n'} W_{h,h'}(t,f) \tag{4.9.8}$$

The function $W_{h,h'}(t,f)$ denotes the cross-Wigner-Ville distribution of elementary functions $h_{m,n}(t)$ and $h_{m',n'}(t)$. It is a two-dimensional Gaussian function,

$$\begin{aligned} W_{h,h'}(t,f) \;=\; & 2\exp\{-\alpha^{-1}(t - \frac{m+m'}{2}T)^2 - \alpha(2\pi)^2(f - \frac{n+n'}{2}\Omega)^2\} \\ & \times \exp\{j2\pi[(n-n')\Omega t - (m-m')T(f - \frac{n+n'}{2}\Omega)]\} \end{aligned} \tag{4.9.9}$$

Its amplitude is symmetrical and concentrated about the point

$$(\frac{m+m'}{2}T, \frac{n+n'}{2}\Omega).$$

Moreover, it is also oscillating. While the rate of oscillation in the frequency domain is determined by $(m-m')T$, the rate of oscillation in the time domain is dependent on $(n-n')\Omega$. Therefore, the further $h_{m,n}(t)$ and $h_{m',n'}(t)$ are apart, the higher the oscillation. If the Wigner-Ville distribution is thought of as the signal's energy distribution, then (4.9.9) can be considered as an *energy atom*. Equation (4.9.8) shows that the Wigner-Ville distribution is a linear combination of an infinite number of energy atoms. Finally, $W_{h,h'}(t,f)$ and $W_{h',h}(t,f)$ form a pair of complex conjugates, i.e.,

$$W_{h,h'}(t,f) + W_{h',h}(t,f) = 2Re[W_{h,h'}(t,f)] \tag{4.9.10}$$

Recall that all desired properties are related to the average of the Wigner-Ville distribution. Hence, the slowly oscillating energy atoms would have more contribution to the desired properties than that of fast-oscillating energy atoms. This is because the average of energy atoms is inversely proportional to the rate of oscillation. The higher the oscillation, the smaller the energy (or average). Moreover, we also know that the rate of oscillation is completely determined by the distance between two elementary functions. Hence, the energy atom whose corresponding

elementary functions $h_{m,n}(t)$ are close to each other dominates the desirable properties. The term whose elementary functions are further apart causes crossterm interference, but has limited contribution to the useful features.

Based on these observations, we re-group (4.9.9) into a new representation, named as the Gabor spectrogram, i.e.,
$GS_D(t,f) =$

$$\sum_{|m-m'|+|n-n'|\leq D} c_{m,n}c^*_{m',n'} 2\exp\{-\alpha^{-1}(t-\frac{m+m'}{2}T)^2 - \alpha(2\pi)^2(f-\frac{n+n'}{2}\Omega)^2\}$$

$$\exp\{j2\pi[(n-n')Tt - (m-m')T(\omega - \frac{n+n'}{2}\Omega)]\} \qquad (4.9.11)$$

where the parameter D denotes the order of the Gabor spectrogram. Note that the Gabor spectrogram (4.9.11) can be effectively computed by separable two-dimensional multi-rate filters [4].

When $D = 0$, (4.9.11) reduces to

$$GS_0(t,f) = 2\sum_{m,n} |c_{m,n}|^2 \exp(-\alpha^{-1}(t-mT)^2 - \alpha(2\pi)^2(f-n\Omega)^2) \qquad (4.9.12)$$

which is non-negative and has no crossterm interference. However, it does not have the desired properties (4.9.1) to (4.9.3) either. As the order D increases, however, the Gabor spectrogram is more and more similar to the Wigner-Ville distribution. In other words, it is closer to the properties (4.9.1) to (4.9.3). Meanwhile, crossterms would be introduced. Since the higher oscillated energy atoms are directly related to the crossterms but have limited contribution to the desired properties, by adjusting the order of the Gabor spectrogram we could find a good compromise between crossterm interference and the desired properties. For example, when $D = 4$, the error between the Gabor spectrogram based results (4.9.1) to (4.9.3) and the expected values are usually less than 1%, whereas the crossterms are negligible in most applications.

4.9.3 Numerical Simulations

Figure 4.9.1 shows the 4th order Gabor spectrogram of a bat sound. Corresponding 0th and 2nd order Gabor spectrograms are plotted in Figure 4.9.2. As shown, the time-frequency resolution improves as the order increases. The Gabor spectrogram eventually converges to the Wigner-Ville distribution. Compared to the Gabor spectrogram, the STFT-based spectrogram in Figure 4.9.3 has poor time-frequency resolution but without the crossterm interference that appears in the Wigner-Ville distribution.

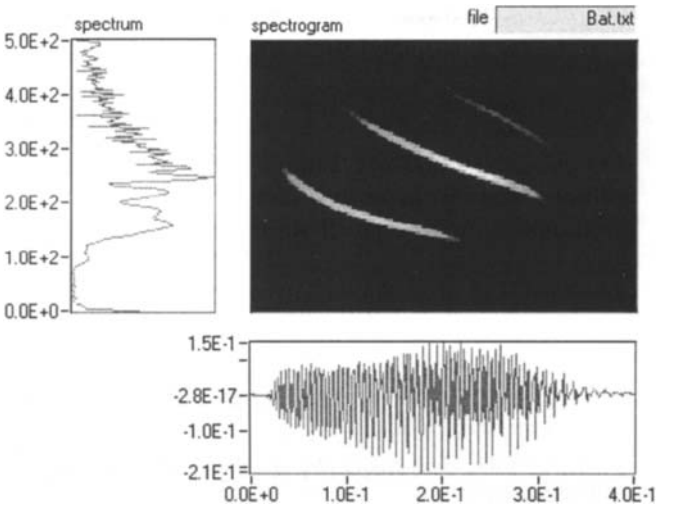


Fig. 4.9.1: Fourth order Gabor spectrogram of bat sound.

Fig. 4.9.2: Zeroth (left) and second (right) order Gabor spectrogram of bat sound.

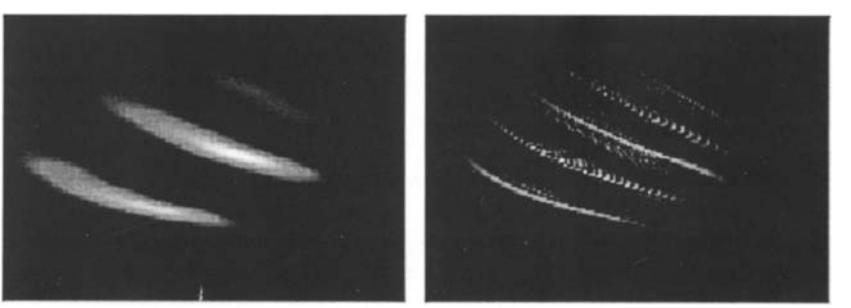


Fig. 4.9.3: STFT-based spectrogram (left) and Wigner-Ville distribution of bat sound.

4.9.4 Summary and Conclusions

As an alternative to the STFT and the ordinary Wigner-Ville distribution, a Wigner-Ville distribution decomposed via the Gabor expansion is introduced. The resulting representation is known as the Gabor spectrogram. One of its major advantages is that by altering the order, we can easily balance the resolution and crossterm interference. The Gabor spectrogram is closer to the STFT-based spectrogram for lower orders, but converges to the Wigner-Ville distribution as the order tends to infinity. Article 6.3 further explores these notions.

References

[1] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] S. Qian and D. Chen, *Joint Time-Frequency Analysis: Methods & Applications.* Upper Saddle River, NJ: Prentice-Hall, 1996.

[3] S. Qian and D. Chen, "Joint time-frequency analysis," *IEEE Signal Processing Magazine*, vol. 16, pp. 52–65, March 1999.

[4] S. Qian, Y. Rao, and D. Chen, "A fast Gabor spectrogram," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, vol. 2, pp. 653–656, Istanbul, 5–9 June 2000.

[5] D. Gabor, "Theory of communication," *J. IEE*, vol. 93(III), pp. 429–457, November 1946.

[6] S. Qian, *Introduction to Time-frequency and Wavelet Transforms.* Englewood Cliffs, NJ: Prentice-Hall, 2002.

[7] J. Wexler and S. Raz, "Discrete Gabor expansions," *Signal Processing*, vol. 21, pp. 207–221, November 1990.

[8] E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," *Physics Review*, vol. 40, pp. 749–759, June 1932.

Chapter 5

Design of Time-Frequency Distributions

This chapter describes in detail specific examples of design of Time-Frequency Distributions (TFDs), extending the material described in Chapters 2 and 3. This topic is covered in eight articles with appropriate cross-referencing.

Ambiguity functions are traditionally used in radar and sonar. As duals of TFDs by a 2D Fourier transform, they are shown to be a key basis for TFD kernel design methodologies (Article 5.1). One of the first TFDs constructed on this basis is the Gaussian kernel distribution (or Choi-Williams distribution), a TFD designed for its reduced interference properties (5.2). To better adapt to the signal under analysis and achieve higher resolution and concentration performance, we can design adaptive TFDs using optimization theory (5.3). TFDs can also be designed and adapted to a specific class of signals such as polynomial FM signals, leading to the formulation of polynomial WVDS (5.4). The design of such methods can be related mathematically to the theory of polynomial derivatives, thus allowing for a revised design procedure (5.5). Another example of TFDs adapted to a particular class of signals is that of TFDs adapted to signals with dispersive group delay (5.6). Another step by step methodology for the design of a specific TFD is provided using separable kernels (5.7). Finally, the class of generalized marginal TFDs is related to the fractional Fourier transform, allowing further design possibilities (5.8).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

5.1 AMBIGUITY FUNCTIONS[0]

5.1.1 The Radar/Sonar Problem

Let us consider the typical radar/sonar problem in which the detection of a target (and the estimation of its relative range d and velocity v with respect to the emitter/receiver) is achieved from the analysis of the returning echo $r(t)$ associated to a given emitted waveform $x(t)$. Assuming a perfect reflection in the echo formation process and a constant radial velocity between the emitter/receiver and the target, $r(t)$ can be modeled as an attenuated replica of $x(t)$, up to a (range encoding) round trip delay, a (velocity encoding) modification due to the Doppler effect and some observation noise. On the basis of various criteria (maximum likelihood, Neyman-Pearson strategy, maximum contrast,...), it is known [1] that a basic ingredient for solving the detection problem is a measure of (linear) similarity, in the L^2-sense of a correlation, between the signal to detect and the actual echo ("matched filter" principle). Given the assumed model, it is therefore natural to compare the received echo with a battery of templates $(\mathbf{T}_{d',v'}x)(t)$, where $\mathbf{T}_{d',v'}$ stands for the range-velocity transformation attached to the candidate pair (d',v'), so that estimates of d and v can be inferred from:

$$(\hat{d},\hat{v}) := \arg\max_{(d',v')} |\langle r, \mathbf{T}_{d',v'}x\rangle| \,. \tag{5.1.1}$$

As far as the deterministic part of the above inner product is concerned, the ideal situation would be to deliver zero values for all range-velocity pairs except for $(d',v') \equiv (d,v)$. This, however, proves not to be achievable (as justified below), leading to a joint determination of range and velocity which is intrinsically ambiguous: this is the reason why a quantity of the type $\langle x, \mathbf{T}_{d,v}x\rangle$ is loosely referred to as an *ambiguity function* (AF).

5.1.2 Definitions of Ambiguity Functions

In order to be more specific in defining AFs, care has to be taken to physical considerations about the Doppler effect, which accounts for a time stretching of the returning echo.

Narrowband ambiguity functions. In the general case, the Doppler factor expresses as $\eta := (c+v)/(c-v)$, where c stands for the celerity of the propagating waves in the considered medium. In the radar case, the celerity of electromagnetic waves is $c = 3\times 10^8$ m/s and, even if we assume a relative target velocity as large as $v = 3,600$ km/h, we end up with a Doppler factor such that $\eta - 1 = 6.66\times 10^{-6} \ll 1$, thus justifying the approximation $\eta \approx 1+2v/c$. It follows that, if the emitted signal is of the form $x(t) := \tilde{x}(t)\exp\{i2\pi f_0 t\}$, where $\tilde{x}(t)$ is a complex envelope that

[0]Author: **Patrick Flandrin**, Laboratoire de Physique (UMR 5672 CNRS), *Ecole Normale Supérieure de Lyon*, 46 allée d'Italie, 69364 Lyon Cedex 07, France (flandrin@ens-lyon.fr). Reviewers: J. P. Ovarlez and A. Papandreou-Suppappola.

is narrowband with respect to the carrier f_0, the deterministic part $r_d(t)$ of the returning echo $r(t)$ admits the approximation $r_d(t) \propto x(t-\tau)\,\exp\{i2\pi\nu t\}$, with τ a round trip delay such that $\tau := 2d/v$ and ν a Doppler *shift* such that $\nu := 2\pi f_0 v/c$. The corresponding inner product

$$\langle x, r_d\rangle \propto \int_{-\infty}^{+\infty} x(t)\, x^*(t-\tau)\, e^{-i2\pi\nu t}\, dt \tag{5.1.2}$$

is therefore proportional to a quantity referred to as the *narrowband* AF of $x(t)$. Whereas this formulation is the one initially introduced by Woodward [2], it often proves useful to rather make use of an inverse Fourier transform and to adopt the following *symmetrized* definition:

$$A_x(\nu,\tau) := \int_{-\infty}^{+\infty} x\,(t+\tau/2)\; x^*\,(t-\tau/2)\; e^{i2\pi\nu t}\, dt. \tag{5.1.3}$$

Wideband ambiguity functions. The above definitions (5.1.2)–(5.1.3) are based on approximations that may prove not to be relevant in contexts different from radar. This is especially the case in airborne sonar, where the celerity of acoustic waves is $c = 340$ m/s, thus leading to $\eta \approx 1.2$ for relative radial velocities $v \approx 100$ km/h. A similar situation can also be observed (although to a smaller extent) in underwater sonar, where the sound celerity in water $c = 1500$ m/s and typical relative velocities $v \approx 2.6$ m/s lead to $\eta \approx 1.034$. In such cases, the previous approximation of a Doppler shift is no longer valid for wideband signals, and the more general form

$$\tilde{A}_x(\eta,\tau) := \sqrt{\eta}\int_{-\infty}^{+\infty} x(t)\, x^*(\eta(t-\tau))\, dt \tag{5.1.4}$$

has to be preferred as a definition of a *wideband* AF [3].

Such a wideband definition naturally reduces to the narrowband one when the analyzed signal is narrowband.

Ambiguity functions and time-frequency distributions. AFs can be viewed as two-variable generalizations of correlation functions. In this respect they are duals of energy distributions. As seen in Sections 3.2.1 and 3.2.3, it follows from the definition (5.1.3) that

$$\int\!\!\int_{-\infty}^{+\infty} A_x(\nu,\tau)\, e^{-i2\pi(\nu t+\tau f)}\, d\nu\, d\tau = W_x(t,f), \tag{5.1.5}$$

where $W_x(t,f)$ is the Wigner-Ville distribution (WVD).[1] More generally, the whole Cohen class of quadratic time-frequency distributions can be obtained as the 2D Fourier transform of weighted (narrowband) AFs $g(\nu,\tau)\, A_x(\nu,\tau)$ (see, e.g., [5]).

[1] It is worth noting that the symmetrized AF (5.1.3) has in fact been pioneered by J. Ville [4] as a form of "time-frequency characteristic function".

Similarly, a properly symmetrized version of the wideband ambiguity function (5.1.4) can be shown [6] to be in Mellin-Fourier duality with a wideband time-frequency distribution, referred to as *Altes' Q-distribution* $Q_x(t,f)$. More precisely, if we let

$$\hat{A}_x(\eta,\tau) := \int_{-\infty}^{+\infty} x(\eta^{-1/2}(t+\tau/2))\, x^*(\eta^{+1/2}(t-\tau/2))\, dt = \tilde{A}_x(\eta,\eta^{-1/2}\tau), \quad (5.1.6)$$

we have [6]

$$\int_0^{+\infty}\int_{-\infty}^{+\infty} \hat{A}_x(\eta,\tau)\, e^{-i2\pi f\tau}\, \eta^{i2\pi t-1}\, d\eta\, d\tau = Q_x(t,f), \quad (5.1.7)$$

with the warping equivalence $Q_x(t,f) \equiv W_{\check{x}}(t,\log f)$, if $\check{X}(f) := X(e^f)$.

Another interesting connection can be pointed out between AFs and linear time-frequency (or time-scale) representations. In fact, the right-hand side of (5.1.2) can be viewed as the *short-time Fourier transform* of $x(t)$, with window $h(t) := x(t)$; in the same respect, (5.1.4) is nothing but the *wavelet transform* of $x(t)$, with wavelet $\psi(t) := x(t)$ (and scale $a := 1/\eta$). In both cases, the AF is exactly identical to the *reproducing kernel* of the corresponding linear transform [5].

5.1.3 Properties of Narrowband Ambiguity Functions

Invariances and covariances. Whereas members of Cohen's class[2] are *covariant* with respect to time and frequency shifts, the squared modulus of the AF (a quantity referred to as the *ambiguity surface* (AS)) is *invariant* to such transformations (i.e., $|A_y(\nu,\tau)|^2 = |A_x(\nu,\tau)|^2$ for any shifted version $y(t) := x(t-\theta)\exp\{i2\pi\xi t\}$ of a given signal $x(t)$). In a similar way, the AF inherits—by Fourier duality—from a number of properties satisfied by the WVD, such as covariance with respect to dilations, rotations or chirp modulations [5,7].

Cross-sections. As it has been mentioned, the narrowband AF can be seen as a *correlation function* with respect to time and frequency shifts. As such, it is hermitian symmetric: $A_x(-\nu,-\tau) = A_x^*(\nu,\tau)$, and it satisfies the inequality

$$|A_x(\nu,\tau)| \leq |A_x(0,0)| = \|x\|_2^2. \quad (5.1.8)$$

Although this interpretation cannot be pushed too far (in particular, the AF is not a non-negative definite quantity, since its 2D Fourier transform—namely, the WVD—can attain negative values), cross-sections of the AF are meaningful 1D correlation functions, since we have:

$$A_x(0,\tau) = \int_{-\infty}^{+\infty} x\,(t+\tau/2)\; x^*\,(t-\tau/2)\; dt \quad (5.1.9)$$

[2]That is, the quadratic class; see p. 68n.

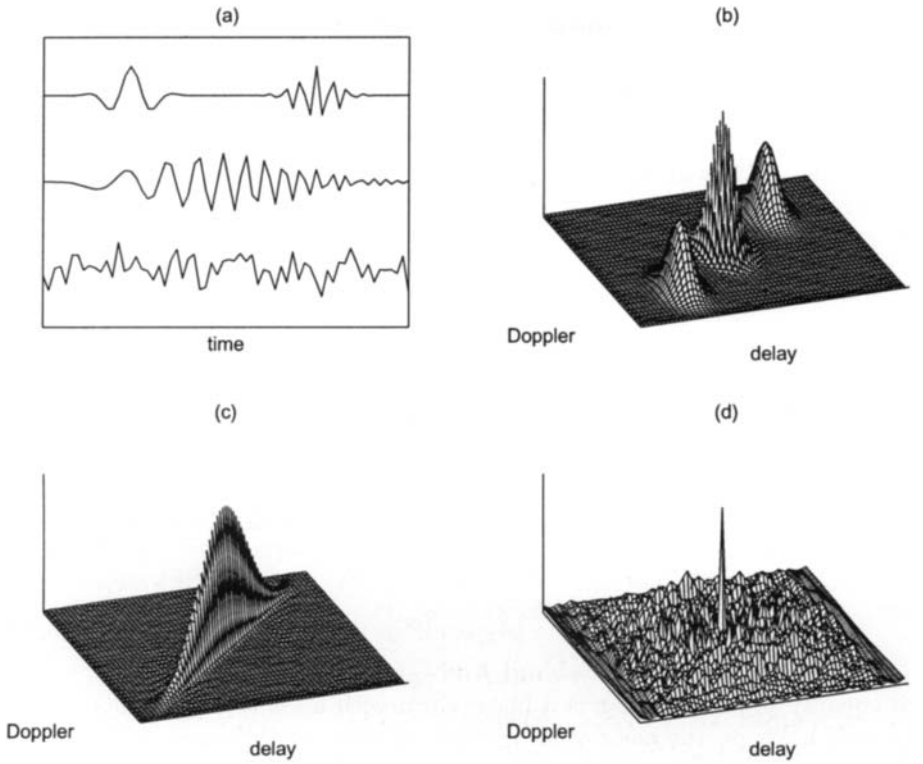


Fig. 5.1.1: Ambiguity functions as time-frequency correlation functions. Subplots (b) to (d) display the ambiguity surfaces attached respectively to the three signals plotted in (a), namely (from top to bottom): two Gabor logons, a linear chirp and a sample of white Gaussian noise. All three surfaces share the common property of attaining their maximum value at the origin of the plane, with values off the origin that reveal correlations in the signal structure, with respect to both time and frequency shifts (referred to as delay and Doppler).

and

$$A_x(\nu, 0) = \int_{-\infty}^{+\infty} X\,(f + \nu/2)\; X^*\,(f - \nu/2)\; df. \tag{5.1.10}$$

This idea of a time-frequency correlation function (which is illustrated in Figure 5.1.1) is instrumental in the design of reduced interference distributions within Cohen's class (cf. Article 4.2).

Volume invariance and self-transformation. Using Parseval's relation and Moyal's formula [5], we readily get that, for any two signals x and y,

$$\int\!\!\int_{-\infty}^{+\infty} A_x(\nu, \tau)\, A_y^*(\nu, \tau)\, d\nu\, d\tau = \left| \int_{-\infty}^{+\infty} x(t)\, y^*(t)\, dt \right|^2 . \tag{5.1.11}$$

Setting $y \equiv x$ in this equation, it follows that

$$\int\int_{-\infty}^{+\infty} |A_x(\nu,\tau)|^2 \, d\nu \, d\tau = \|x\|_2^4, \tag{5.1.12}$$

i.e., that the AS has an *invariant volume* that is only fixed by the signal's energy. More remarkably, (5.1.12) is just a special case of Siebert's *self-transformation* property [8]:

$$\int\int_{-\infty}^{+\infty} |A_x(\nu,\tau)|^2 \, e^{i2\pi(\nu t+\tau f)} \, d\nu \, d\tau = |A_x(f,t)|^2, \tag{5.1.13}$$

from which it can be inferred that an AF is a highly structured function.

Uncertainty principles. If we combine the "correlation" inequality (5.1.8) and the "volume invariance" property (5.1.12), it is clear that an AS cannot be perfectly concentrated at the origin of the plane. This limitation, that is sometimes referred to as the *radar uncertainty relation* [9], admits a more precise L^p-norm formulation ($p > 2$) as follows [10]:

$$I_x(p) := \int\int_{-\infty}^{+\infty} |A_x(\nu,\tau)|^p \, d\nu \, d\tau \leq \frac{2}{p} \, \|x\|_2^{2p}, \tag{5.1.14}$$

with equality if and only if x is a linear chirp with a Gaussian envelope. A similar result holds on the basis of an *entropic* measure of sharpness, leading to the inequality

$$S_x := -\int\int_{-\infty}^{+\infty} |A_x(\nu,\tau)|^2 \, \log |A_x(\nu,\tau)|^2 \, d\nu \, d\tau \geq 1 \tag{5.1.15}$$

for unit energy signals.

The common interpretation of those inequalities is that an AS cannot be zero everywhere except at the origin of the plane. In the case where all of the AS is supposed to be concentrated around the origin, it has necessarily to extend over a domain (whose area is non-zero) which defines the joint accuracy of any delay-Doppler measurement [1, 11]. However, ASs which are more sharply peaked at the origin can be found, provided that non-zero values are accepted somewhere off the origin in the plane: for unit energy signals, ASs with null values except in $(0, 0)$ can be obtained on convex domains whose *clear area* cannot however be greater than 4 [12]. An example is given in Figure 5.1.2.

Delay/Doppler estimation. The best achievable performance in joint estimation of delay and Doppler is bounded. The actual Cramér-Rao bounds on variances and covariances can be derived from the Fisher information matrix of the problem, whose terms can themselves be expressed as partial derivatives of the AS, in the case of additive white Gaussian noise [1]. Since the AS is basically the maximum likelihood estimator for delay and Doppler, and since this estimator can be shown to

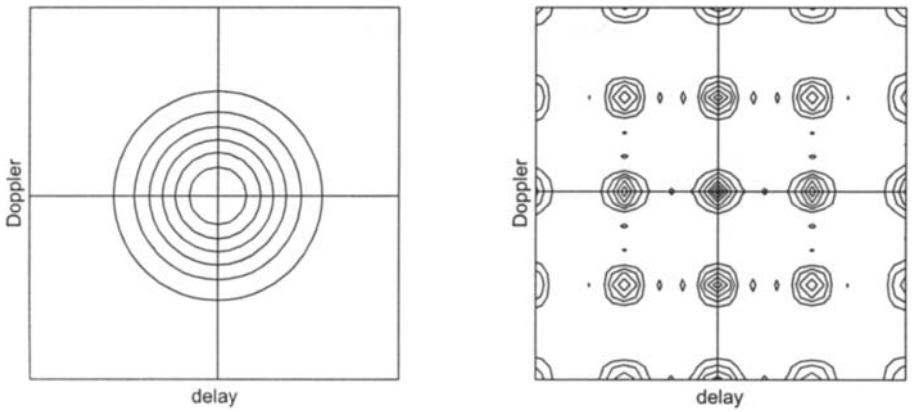


Fig. 5.1.2: Sharpening the central peak of an ambiguity surface. Whereas a single Gaussian pulse has an ambiguity surface whose central peak cannot have an effective area $\mathcal{A}$ smaller than a limit fixed by the "radar uncertainty principle" (left), a signal defined by the superposition of a number of replicae of such a pulse, periodically shifted in time and frequency, may guarantee a sharper central peak (right). This has however to be paid at the price of auxiliary peaks off the origin, with a "clear area" of the order of $\mathcal{A}$.

be asymptotically efficient, it thus follows that the AS geometry is a direct indicator of the expected accuracy in the estimation. Roughly speaking, variances in delay and Doppler estimation are given by the effective widths of the central peak of the AS.

Signal design. In active problems, in which the emitted signal can be freely chosen (up to a certain extent), an important issue is to design waveforms with a prescribed AF (or AS), so that some desired performance can be guaranteed. From a purely theoretical point of view, a signal is entirely determined (up to a pure phase term) by its AF, since we can invert the definition (5.1.3) according to:

$$x(t) = \frac{1}{x^*(0)} \int_{-\infty}^{+\infty} A_x(\nu, \tau)\, e^{-i\pi\nu\tau}\, d\nu. \tag{5.1.16}$$

Unfortunately, as it has been said before, an AF is a highly structured function and an arbitrary 2D function has in general no reason to be admissible, i.e., to be the actual AF of some signal. Different approaches have been proposed to overcome this limitation. One can first think of looking for the signal $\hat{x}(t)$ whose AF approaches at best a given time-frequency function $F(\nu, \tau)$, according, e.g., to a L^2-distance [13]:

$$\hat{x}(t) = \arg\min_x \int\int_{-\infty}^{+\infty} |A_x(\nu, \tau) - F(\nu, \tau)|^2\, d\nu\, d\tau. \tag{5.1.17}$$

One can also rely on the physical interpretation of the AF as a time-frequency correlation function and promote waveforms with adapted time-frequency characteristics. In this respect, a large bandwidth (resp., a long duration) is required

for an accurate estimation of delay (resp., Doppler). The simultaneous consideration of these two design principles advocates the use of *chirp* signals with a large bandwidth-duration (BT) product (see [11] and Section 1.2.4).

5.1.4 Remarks on Wideband Ambiguity Functions

In many respects, properties of the wideband AF can be seen as natural generalizations of the narrowband case (although some properties, such as volume invariance, may no longer be satisfied), reducing to them in the narrowband limit.

In parallel with what has been previously mentioned in the narrowband case, the best achievable performance in the joint estimation of delay and Doppler can be expressed, in the wideband case, in terms of geometrical properties of the wideband AS [14].

A companion problem is that of *Doppler tolerance*, which consists in obtaining an unbiased estimate of delay in the presence of any unknown Doppler [15]. Doppler acting as a stretching on the emitted signal, the condition for no bias can be translated into the fact that the effective time-frequency structure of the emitted waveform is invariant under stretching. It turns out that the hyperbola is the only curve of the plane which is invariant under dilation/compression transformations: in terms of chirps, assumed to be conveniently described on the plane by a time-frequency skeleton, this justifies [5] the use of *logarithmic* phases, i.e., of hyperbolic chirps resembling those commonly observed in natural sonar systems (bats) [15].

Finally, it must be pointed out that computing an AF proves more involved in the wideband case than in the narrowband case. Efficient solutions, based on the Mellin transform, have been proposed in [14].

5.1.5 Summary and Conclusions

Ambiguity functions measure the degree of similarity that exists between a signal and its delayed and dopplerized versions: as such, they are central in the evaluation and design of radar/sonar systems. Since they are basically (two-dimensional) time-frequency correlation functions, ambiguity functions also offer a dual perspective on quadratic time-frequency distributions that proves especially useful for kernel design issues.

References

[1] H. L. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. I: "Detection, Estimation, and Linear Modulation Theory". New York: Wiley, 1968. Reprinted 2001.

[2] P. M. Woodward, *Probability and Information Theory with Applications to Radar*. London: Pergamon, 1953.

[3] E. J. Kelly and R. P. Wishner, "Matched filter theory for high-velocity, accelerating targets," *IEEE Trans. Military Electronics*, vol. 9, pp. 56–59, 1965.

[4] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmissions*, vol. 2A, no. 1, pp. 61–74, 1948. In French. English translation: I. Selin, *Theory*

and applications of the notion of complex signal, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).

[5] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[6] R. A. Altes, "Wideband, proportional-bandwidth Wigner-Ville analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1005–1012, June 1990.

[7] A. Papoulis, *Signal analysis.* New York: McGraw-Hill, 1977.

[8] W. M. Siebert, "Studies of Woodward's uncertainty function," *Quarterly Progress Report* (MIT Electronics Research Lab, Cambridge, MA), pp. 90–94, 1958.

[9] D. E. Vakman, *Sophisticated Signals and the Uncertainty Principle in Radar.* New York: Springer, 1968. Translated by K. N. Trirogoff; edited by E. Jacobs.

[10] E. H. Lieb, "Integral bounds for radar ambiguity functions and Wigner distributions," *J. of Mathematical Physics*, vol. 31, pp. 594–599, March 1990.

[11] A. W. Rihaczek, *Principles of high-resolution radar.* New York: McGraw-Hill, 1969. Reprinted Los Altos, CA: Peninsula Publishing, 1985.

[12] R. Price and E. M. Hofstetter, "Bounds on the volume and height distributions of the ambiguity function," *IEEE Trans. Information Theory*, vol. 11, pp. 207–214, 1965.

[13] S. M. Sussman, "Least-squares synthesis of radar ambiguity functions," *IRE Trans. Information Theory*, vol. 8, pp. 246–254, April 1962.

[14] J. Bertrand, P. Bertrand, and J. P. Ovarlez, "The Mellin transform," in *The Transforms and Applications Handbook* (A. D. Poularikas, ed.), ch. 11, pp. 829–885, Boca Raton, FL: CRC Press, 1st ed., 1996.

[15] R. A. Altes and E. L. Titlebaum, "Bat signals as optimally Doppler tolerant waveforms," *J. Acoustical Soc. of America*, vol. 48, pp. 1014–1020, October 1970.

5.2 REDUCED INTERFERENCE TIME-FREQUENCY DISTRIBUTIONS [0]

5.2.1 Nonstationarity, Resolution and Interference

Much background material required for understanding the present article may be found in Chapter 3 and Articles 4.2 and 5.1. These contributions can considerably augment or replace the citations in the present article, which have a historical importance for the topic at hand. This article presents the concept and design methodology for Reduced Interference Distributions or RIDs. It has been quite difficult to handle nonstationary signals such as chirps satisfactorily using conceptualizations based on stationarity. The spectrogram represents an attempt to apply the Fourier transform for a short-time analysis window, within which it is hoped that the signal behaves reasonably according to the requirements of stationarity. By moving the analysis window along the signal, one hopes to track and capture the variations of the signal spectrum as a function of time. However, the spectrogram often presents serious difficulties when used to analyze rapidly varying signals. If the analysis window is made short enough to capture rapid changes in the signal it becomes impossible to resolve frequency components of the signal which are close in frequency during the analysis window duration (see Sections 2.3.1 and 2.3.2).

The Wigner distribution (WD) has been employed as an alternative to overcome this shortcoming. The WD was first introduced in the context of quantum mechanics and revived for signal analysis by Ville [1] and for this reason is sometimes called the Wigner-Ville distribution (WVD). In this book, the WVD is assumed to use the analytic form of the signal (see Sections 2.1.3 and 2.1.4). The WD has many important and interesting properties [1]. It provides a high resolution representation in time and in frequency for a nonstationary signal such as a chirp. However, its energy distribution is not non-negative and it often possesses severe cross terms, or interference terms, between components in different (t, f) regions, potentially leading to confusion and misinterpretation. An excellent discussion on the geometry of interferences has been provided by Flandrin in Article 4.2. RIDs effect a desirable compromise, which retains a number of desirable TFD properties, yet provides reduced cross-term interference, making then a good choice for the analysis of nonstationary signals.

Both the spectrogram and the WD are members of the quadratic class of distributions [1]. For these distributions a time shift in the signal is reflected as an equivalent time shift in the (t, f) distribution and a shift in the frequency of the signal is reflected as an equivalent frequency shift in the (t, f) distribution. The spectrogram, the WD and the RID all have this property. Different distributions in the quadratic class can be obtained by selecting different kernel functions (see Article 2.7).

---

[0]Author: **William J. Williams**, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor MI 48109, USA (wjw@eecs.umich.edu). Reviewers: B. Boashash and P. Flandrin.

Desirable properties of a distribution and associated kernel requirements were extensively investigated by Claasen and Mecklenbräuker in the early 1980s [1]. Boashash provided a comprehensive study of the WVD in the early 1990s [2]. Comprehensive comparisons of TFDs, including the newer methods, were provided in the early 1990s [3].

Choi and Williams introduced one of the earliest "new" distributions [4], which they called the Exponential Distribution or ED. This new distribution overcomes several drawbacks of the spectrogram and WD, providing high resolution with suppressed interferences [4,5]. The method used in developing the ED is very general and is the basis of this article.

5.2.2 The Reduced Interference Distribution

5.2.2.1 Ambiguity Function Relationships

The key to understanding (t, f) relationships and manipulations is a thorough understanding of the ambiguity domain. Let $Z(f)$ be the FT of the signal $z(t)$:

$$Z(f) = F[z(t)] = \int z(t)e^{-j2\pi ft}dt \tag{5.2.1}$$

and

$$z(t) = F^{-1}[Z(f)] = \int Z(f)e^{j2\pi ft}df \tag{5.2.2}$$

Let $R_z(t,\tau)$ be the *instantaneous autocorrelation* of a complex signal $z(t)$; that is,

$$R_z(t,\tau) = z(t+\tau/2)z^*(t-\tau/2) \tag{5.2.3}$$

where z^* denotes the complex conjugate of z. The Wigner distribution of $z(t)$ is the Fourier Transform (FT) of $R_z(t,\tau)$ with respect to the lag variable τ:

$$\begin{aligned} W_z(t,f) &= F_\tau[z(t+\tau/2)z^*(t-\tau/2)] \\ &= F_\tau[K_z(t,\tau)]. \end{aligned} \tag{5.2.4}$$

Similarly, but with a different physical meaning, the symmetrical ambiguity function (AF) is defined as the Inverse Fourier transform (IFT) of $R_z(t,\tau)$ with respect to the first variable:

$$\begin{aligned} A_z(\nu,\tau) &= F^{-1}_{t\to\nu}[z(t+\tau/2)z^*(t-\tau/2)] \\ &= F^{-1}_{t\to\nu}[R_z(t,\tau)]. \end{aligned} \tag{5.2.5}$$

Thus, $W_z(t,f)$ and $A_z(\nu,\tau)$ are related by the two-dimensional (2-D) FT:

$$W_z(t,f) = \iint A_z(\nu,\tau)e^{-j2\pi(t\nu+f\tau)}d\nu\, d\tau. \tag{5.2.6}$$

Equation 5.2.6 may be altered with a kernel to show that $\rho_z(t,f,g)$ may be found by

$$\rho_z(t,f,g) = \iint g(\nu,\tau)A_z(\nu,\tau)e^{-j2\pi(t\nu+f\tau)}d\nu\,d\tau \tag{5.2.7}$$

which is an expression of the quadratic class of distributions, in general. Thus, while the WD may be found from the symmetric ambiguity function by means of a double Fourier transform, any member of the quadratic class of distributions may be found by first multiplying the kernel $g(\nu,\tau)$ by the symmetric ambiguity function and then carrying out the double Fourier transform. The generalized ambiguity function, $g(\nu,\tau)A_z(\nu,\tau)$ [1], is a key concept in (t,f) which aids one in clearly seeing the effect of the kernel in determining $\rho_z(t,f,g)$.

The kernel for the WD is unity, so the generalized ambiguity function is identical to the ambiguity function, and its (t,f) representation (the double Fourier transform) preserves both the auto-terms and the cross-terms. The kernels of the spectrogram and the RID emphasize the auto-terms and deemphasize the cross terms, but in very different ways.

5.2.2.2 The Exponential Distribution

Each TFD discussed is valuable under certain conditions. The ED, developed by Choi and Williams [4], is an attempt to improve on the WD. It has a kernel $g(\nu,\tau) = \exp(\frac{-\nu^2\tau^2}{\sigma})$, and it proves to be quite effective in suppressing the interferences while retaining high resolution. A comparison of the kernels of the spectrogram, the WD and the ED are shown in Fig. 5.2.1.

Interference terms tend to lie away from the axes in the ambiguity plane, while autoterms tend to lie on the axes. The spectrogram kernel attenuates everything away from the $(0,0)$ point, the WD kernel passes everything, and the ED kernel passes everything on the axes and attenuates away from the axes. Thus, the ED generally attenuates interference terms. This provides its reduced interference characteristic. The spectrogram reduces interference also, but at a cost to the autoterms. Reduced interference may be achieved while maintaining a number of very nice mathematical properties with just a few constraints. The ED is just one example of the RID class. A methodology which guarantees a RID yet allows a considerable flexibility in design is discussed in the following section.

5.2.3 Kernel Selection for RID

A more formal description of RIDs is appropriate at this point. RID requirements and properties will be 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 ν in order to maintain most of its desirable properties. It is sufficient to ensure that the kernel is unity along $\nu = 0$ and $\tau = 0$ and that the kernel is such that $g^*(\nu,\tau) = g(-\nu,-\tau)$, the later property ensuring realness. The RID kernel is cross-shaped and acts as a low pass filter in both ν and τ. The spectrogram has the virtue of suppressing cross-terms as does

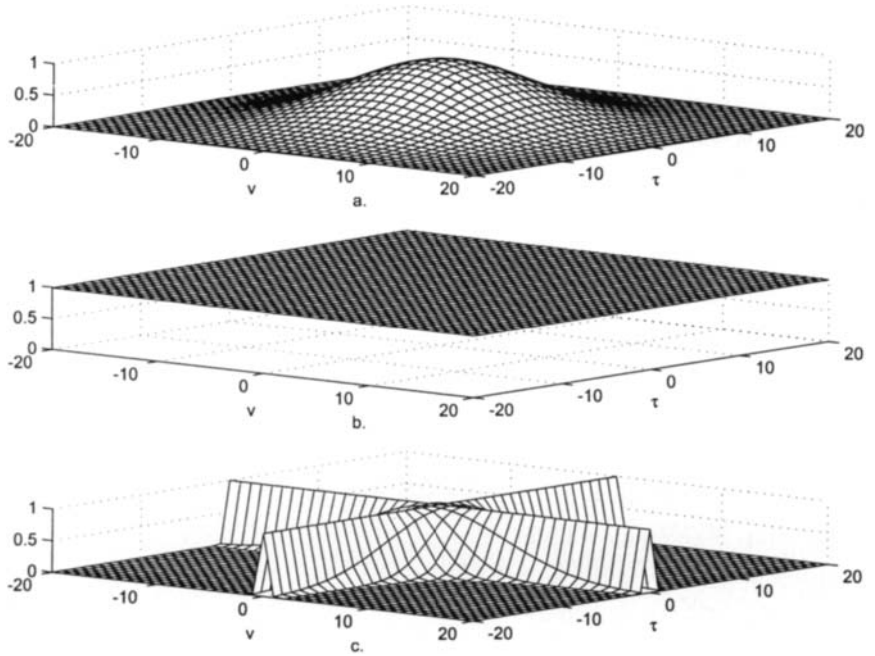


Fig. 5.2.1: Comparison of (a) spectrogram, (b) WD and (c) ED kernels in the ambiguity plane

the RID, and has the further advantage of being non-negative which is not the case for the WD and the RID. The RID possesses almost all of the desirable properties of the WD except for its unitary property ($|g(\nu,\tau)|$=1 for all ν,τ).

It is 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 Fourier transform in frequency. The time and frequency support property may be maintained for the RID by insuring that

$$\psi(t,\tau) = \int g(\nu,\tau)e^{-j2\pi\nu t}d\nu = 0 \text{ if } |\tau| < 2|t|. \tag{5.2.8}$$

This forms a "cone-shaped" region in (t,τ). The WD obviously satisfies this support property since the Fourier transform of unity is an impulse function, clearly staying within the (t,τ) limits. The form of the kernel in (ν,f) is also cone shaped, ensuring the frequency support property. Zhao, Atlas and Marks [6] suggest a cone shaped kernel for nonstationary signal analysis, but they impose restrictions such that time support only is ensured. The ED can be brought into the RID requirements by imposing a RID window as suggested above. The RID is not a totally new distribution since the Born-Jordan kernel, $g(\nu,\tau) = \text{sinc}(\nu\tau) = \frac{\sin(\pi\nu\tau)}{\pi\nu\tau}$, meets all of the RID requirements [1].

Table 5.2.1: Distribution properties and associated kernel requirements.

P0. nonnegativity : $\rho_z(t,f;g) \geq 0 \; \forall t,f$
R0. $g(\nu,\tau)$ is a positively weighted sum of ambiguity functions.
P1. realness : $\rho_z(t,f;g) \in R$
R1. $g(\nu,\tau) = g^*(-\nu,-\tau)$
P2. time shift : $s(t) = z(t-t_0) \Rightarrow \rho_s(t,f;g) = \rho_z(t-t_0,z;g)$
R2. $g(\nu,\tau)$ does not depend on t.
P3. frequency shift : $s(t) = z(t)e^{j2\pi f_0 t} \Rightarrow \rho_s(t,f;g) = \rho_z(t,f-f_0;g)$
R3. $g(\nu,\tau)$ does not depend on f.
P4. time marginal : $\int \rho_z(t,f)df = z(t)z^*(t)$
R4. $g(\nu,0) = 1 \; \forall \nu$
P5. frequency marginal : $\int \rho_z(t,f;g)dt = Z(f)Z^*(f)$
R5. $g(0,\tau) = 1 \; \forall \tau$
P6. instantaneous frequency : $\frac{\int f\rho_z(t,f;g)df}{\int \rho_z(t,f;g)df} = f_i(t)$
R6. R4 and $\frac{\partial g(\nu,\tau)}{\partial \tau}|_{\tau=0} = 0 \; \forall \nu$
P7. group delay : $\frac{\int t\rho_z(t,f;g)dt}{\int \rho_z(t,f;g)dt} = t_g(f)$
R7. R5 and $\frac{\partial g(\nu,\tau)}{\partial \nu}|_{\nu=0} = 0 \; \forall \tau$
P8. time support : $z(t) = 0$ for $|t| > t_c \Rightarrow \rho_z(t,f;g) = 0$ for $|t| > t_c$
R8. $\psi(t,\tau) \stackrel{\Delta}{=} \int g(\nu,\tau)e^{-j2\pi\nu t}d\nu = 0$ for $|\tau| < 2|t|$
P9. frequency support : $Z(f) = 0$ for $|f| > f_c \Rightarrow \rho_z(t,f;g) = 0$ for $|f| > f_c$
R9. $\int g(\nu,\tau)e^{j2\pi f\tau}d\tau = 0$ for $|\nu| < 2|f|$
P10. Reduced interference
R10. $g(\nu,\tau)$ is a 2-D low pass filter type.
P11. Scale: $s(t) = \frac{1}{\sqrt{|a|}}z(at) \Rightarrow \rho_s(t,f;g) = \rho_z(at,f/a;g)$
R11. $g(\nu,\tau) = g(\nu\tau)$ is a product kernel.

Table 5.2.1 summarizes desirable properties (P) and related kernel requirements (R) of distributions. No known, practical distribution with a fixed kernel (not dependent on the signal or time) is able to meet all of these requirements.

Table 5.2.2 illustrates how several common distributions satisfy the desirable properties. Two distributions whose kernels meet the RID requirements (Born-Jordan and windowed ED) are included. Note that the RID exhibits time-shift, frequency-shift and scale covariance [7]. Some prefer to call this *invariance* because the TFD signature does not change in shape under these transformations. Fig. 5.2.2 illustrates the ED for a succession of signals consisting of a dolphin click, its 2:1 scaled and energy-corrected version, and its frequency-shifted version. This shows the time-shift, frequency-shift and scale covariance of the ED. Any true RID would exhibit similar covariance properties.

5.2.3.1 Design Procedures for Effective RID Kernels

There is much more that can be done in terms of kernel design. It is possible to bring much of the work that has been done on windows and digital filters to bear in designing effective RID kernels [8]. The starting point for the kernel design is to consider a primitive function $h(t)$. This function is designed to have unit area ($\int h(t)dt = 1$) and is symmetrical; i.e. $h(-t) = h(t)$. It is limited such that $h(t) = 0$ for $|t| > 1/2$, and is tapered so that it has little high frequency content; that is, it is a low-pass impulse response. Then the kernel is

$$g(\nu, \tau) = H(\nu\tau) \tag{5.2.9}$$

where $H(\nu)$ is the Fourier transform of $h(t)$. It can be shown that $H(\nu\tau)$ satisfies the RID requirements. Desirable characteristics of $h(t)$ have been the subject of intensive study in terms of windows in time domain terms or filters in frequency domain terms. All of this knowledge can be used to select effective RID kernels. Note that in the (t, τ) domain the form of the RID kernel is $\frac{1}{|\tau|}h(\frac{t}{|\tau|})$. In order to compute the RID, one convolves this form of the kernel with $K(t, \tau)$ along the t dimension and then Fourier transforms that result. Thus the kernel is scaled in t and normalized by $\frac{1}{|\tau|}$ for each τ. This is a wavelet-like characteristic. The RID thus exhibits time-shift, frequency-shift and scale covariance—unlike wavelets, which commonly exhibit only scale and time-shift covariance, and many other time-frequency distributions which exhibit only time-shift and frequency-shift covariance.

5.2.3.2 Optimum and Adaptive RIDs

The RID constraints are mild enough to permit a considerable amount of flexibility in RID kernel design. Several TFDs which are full fledged RIDs have been proposed. In other cases some of the properties in Table 5.2.1 have been abandoned or relaxed in order to achieve a more specific effect for a narrow class of signals. Space does not permit a full discussion of these variations on the theme, but a more comprehensive discussion is available elsewhere [9]. RID kernels may also be adapted to minimize the entropy of the resulting TFD with nice results [10]

5.2.3.3 Discrete RID Kernels

Requirements for discrete forms of the RID are similar to those of the discrete WD. The discrete RID may be formed by

$$\mathrm{RID}_z(n, f) = \sum_{m=-\infty}^{\infty} K_z(n, m) *_n \psi(n, m)e^{-j2\pi fm} \tag{5.2.10}$$

where

$$g(m, f) = \sum_{n=-\infty}^{\infty} \psi(n, m)e^{-j2\pi fn} \tag{5.2.11}$$

Table 5.2.2: Comparison of properties for several distributions. The Born-Jordan distribution and the windowed ED are RIDs. $A_w(\nu,\tau)$ is the ambiguity function of the spectrogram window $w(t)$. $W(u)$ is the Fourier transform of $\mathrm{rect}(t)$.

Distribution	$g(\nu,\tau)$	P0	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11
Wigner	1		x	x	x	x	x	x	x	x	x		x
Rihaczek	$e^{j\pi\nu\tau}$			x	x	x	x			x	x		x
Exponential (ED)	$e^{-\nu^2\tau^2/2\sigma}$		x	x	x	x	x	x	x			x	x
Spectrogram	$A_w(\nu,\tau)$	x	x	x	x							x	
Born-Jordan	$\frac{\sin(\nu\tau/2)}{\nu\tau/2}$		x	x	x	x	x	x	x	x	x	x	x
Windowed-ED	$e^{-u^2/2\sigma} * W(u)\|_{u=\nu\tau}$			x	x	x	x	x	x	x	x	x	x
Cone (ZAM)	$g(\tau)\|\tau\|\frac{\sin(a\nu\tau)}{a\nu\tau}$		x	x	x	x				x		x	

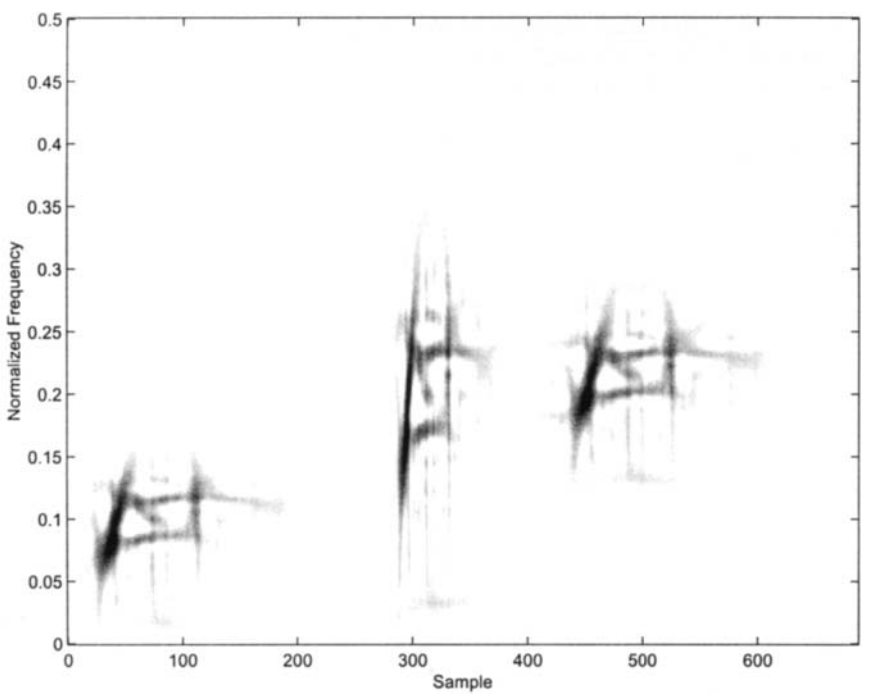


Fig. 5.2.2: ED with original, scaled and frequency shifted dolphin click

is the discrete RID kernel.

The discrete RID may thus be conveniently formed by obtaining the local autocorrelation $K_z(n,m)$, convolving it with $\psi(n,m)$ along n, and taking the DTFT of the result with respect to m. A fully discrete form of the kernel is more desirable. A very convenient discrete RID kernel has been discovered based on the binomial distribution [5]. The form of the kernel is

$$G(n,m) = \frac{1}{2^{|m|}}\binom{|m|}{k}\delta(n - |m| + k). \tag{5.2.12}$$

The correlation shift index m takes the values $-\infty \ldots -1, 0, 1, \ldots \infty$, and the time shift index n takes the values $-\infty \ldots -1, -.5, 0, .5, 1, \ldots \infty$. It can be shown that the signal structure of the discrete local autocorrelation and the discrete form of the kernel can be easily formulated to include the half-integers [11].

5.2.4 Comparisons of TFD Results

A few comparisons of TFD results are warranted. The spectrogram, the discrete WD and the binomial distribution (BD) just described are chosen for this purpose. A synthetic signal composed of a low-frequency, frequency-modulated sinusoid (warble) and a rising chirp serves to illustrate some of the differences between these TFDs. The alias-free form of the discrete TFD was used in each case [11]. The sampling frequency was assumed to be one sample per second.

One can readily see that the chirp is evident in the spectrogram of Fig. 5.2.3(a), but rather smeared in frequency as one might expect when using a 512 point window. The warble is a continuous smear with little evident fine structure. The WD result is shown in Fig. 5.2.3(b).

A 512 point analysis window was used for the WD. The resolution for the WD is dramatically better than is the case for the spectrogram, however. One can see that the chirp is very well resolved and the frequency-modulated "warble" is also well resolved. However, there is an interlaced band of cross-term activity between the chirp and the warble. Finally, the BD result is shown in Fig. 5.2.3(c). Here, the warble and the chirp are well resolved with little evident cross-term activity between them, but there is still a bit of interlaced cross-term activity within the sinusoidal result.

5.2.5 Summary and Conclusions

The RID approach provides an advantage over the WD and the spectrogram. It may be efficiently computed, particularly in its discrete forms. The RID retains almost all of the desirable properties of the WD, but with considerably reduced interference. The RID has proved to be a useful investigative tool in many applications, revealing previously unobservable signal characteristics in many cases. Despite the constraints which guarantee the RID properties, the kernel retains a great deal of flexibility and may be tailored to specific applications with considerable benefit.

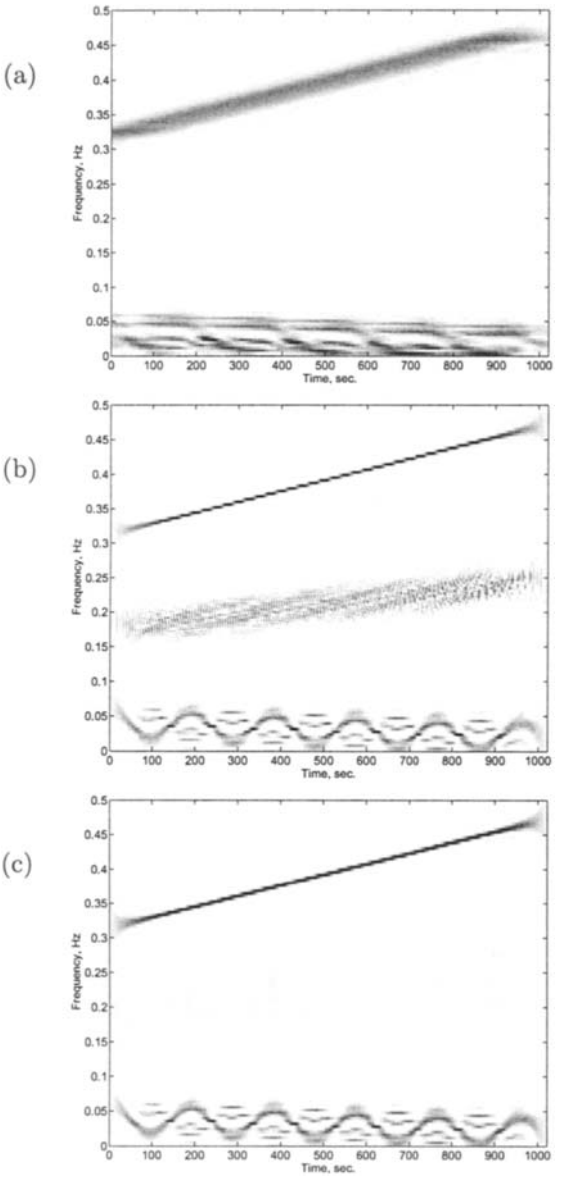


Fig. 5.2.3: Comparison of Binomial TFDs for a 512pt. window: (a) Spectrogram, (b) WVD, (c) Binomial TFD. From [12].

References

[1] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[3] P. J. Loughlin, J. W. Pitton, and L. E. Atlas, "Bilinear time-frequency representations: New insights and properties," *IEEE Trans. Signal Processing*, vol. 41, pp. 750–767, 1993.

[4] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multi-component signals using exponential kernels," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 37, pp. 862–871, June 1989.

[5] W. J. Williams and J. Jeong, "Reduced interference time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 3, pp. 74–97, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[6] Y. Zhao, L. E. Atlas, and R. J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[7] F. Hlawatsch and R. L. Urbanke, "Bilinear time-frequency representations of signals: The shift-scale invariant class," *IEEE Trans. Signal Processing*, vol. 42, pp. 357–366, 1994.

[8] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, pp. 402–412, February 1992.

[9] W. J. Williams, "Reduced interference distributions: Biological applications and interpretations," *Proc. IEEE*, vol. 84, pp. 1264–1280, September 1996.

[10] W. J. Williams and T.-H. Sang, "Adaptive RID kernels which minimize time-frequency uncertainty," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 96–99, Philadelphia, PA, 25–28 October 1994.

[11] J. Jeong and W. J. Williams, "Alias-free generalised discrete-time time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 40, pp. 2757–2765, November 1992.

[12] W. J. Williams, "Recent advances in time-frequency representations: Some theoretical foundation," in *Time Frequency and Wavelets in Biomedical Signal Processing* (M. Akay, ed.), ch. 1, New York: IEEE/Wiley, 1998.

[13] L. Stanković, "Auto-term representation by the reduced interference distributions: A procedure for kernel design," *IEEE Trans. Signal Processing*, vol. 44, pp. 1557–1563, June 1996.

5.3 ADAPTIVE TIME-FREQUENCY ANALYSIS [0]

The lack of a single time-frequency representation (TFR) that is "best" for all applications has resulted in a proliferation of TFRs, each corresponding to a different, fixed mapping from signals to the time-frequency plane. A major drawback of all fixed mappings is that, for each mapping, the resulting time-frequency representation is satisfactory only for a limited class of signals. Adaptive TFRs tune their representation to each signal to offer good performance for a large class of signals.

A natural way to approach tuning is via *optimization*; this requires an objective function or performance measure that expresses what we "want to see" and a class of representations to choose from. Below we overview several optimized TFRs. Broadly speaking, the approaches break into adaptive TFRs based on the linear short-time Fourier transform (STFT) and those based on the quadratic Wigner distribution. When properly tuned, a high precision representation results, as we see in Figure 5.3.1.

5.3.1 Adaptive Short-Time Fourier Transforms

The simplicity, efficiency, robustness, and generally good performance of the short-time Fourier transform (STFT) make it the primary tool for routine time-frequency analysis (see Section 2.3.1). However, in some cases, the spectrogram (its squared magnitude) shows considerably inferior concentration relative to more sophisticated quadratic time-frequency representations. Jones and Parks [1] have shown that inferior resolution in the spectrogram is primarily due to *mismatch* between the window and signal components; that is, the spectrogram performs best when the duration and orientation of the analysis window match those of the local signal components. When these are not known a priori or when they vary between multiple components in a complex signal, an adaptive-window STFT can adjust the window parameters over time and frequency to obtain near-optimal performance for all components.

An adaptive-window STFT of the signal x takes the form

$$S(t,\omega) \;=\; \int x(\tau)\, w_{t,\omega}(\tau - t)\, e^{-j\omega\tau} d\tau, \tag{5.3.1}$$

where the only difference between the adaptive-window STFT and the conventional STFT is the potential variation of the window w with time and frequency. A family of unit-energy Gaussian windows of the form

$$w_{t,\omega}(\tau) \;=\; (-2\,\mathrm{Re}[c_{t,\omega}]/\pi)^{1/4} \exp\bigl[c_{t,\omega}(\tau - t)^2\bigr]\, e^{-j\omega\tau} \tag{5.3.2}$$

[0] Authors: **Richard Baraniuk**, Department of Electrical and Computer Engineering, Rice University, 6100 Main Street, Houston, TX 77005, USA (richb@rice.edu, http://www.dsp.rice.edu), and **Douglas L. Jones**, Department of Electrical and Computer Engineering and Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801, USA (jones@dsp.csl.uiuc.edu). Reviewers: P. Flandrin and F. Auger.

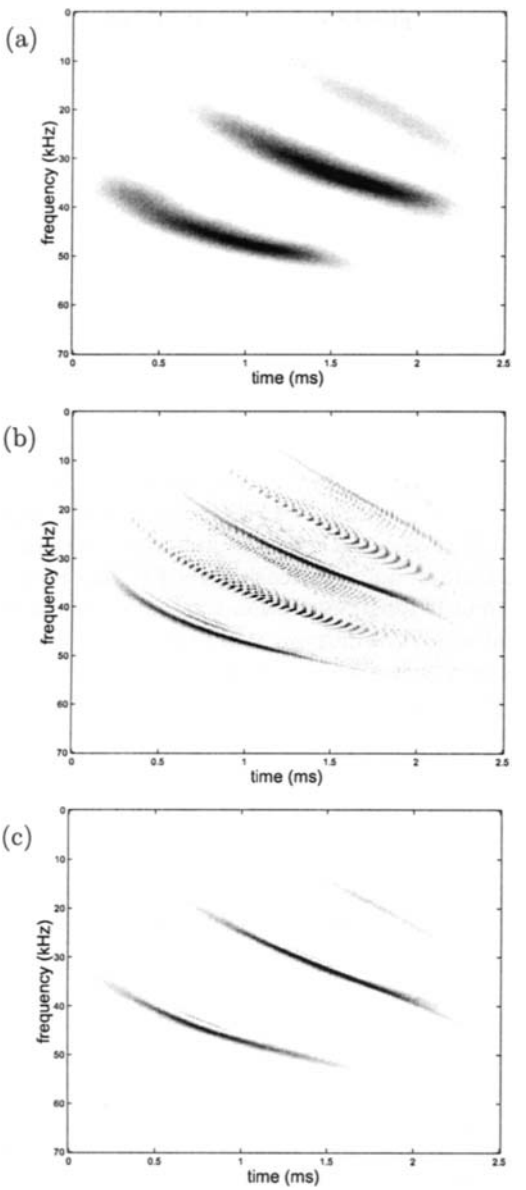


Fig. 5.3.1: Time-frequency analysis of 2.5 ms of an echolocation pulse emitted by the large brown bat, *Eptesicus fuscus*. Pictured are the (a) spectrogram, (b) Wigner distribution, and (c) adaptive radially Gaussian optimal-kernel distribution (Section 5.3.2.3). The 1/0 optimal-kernel distribution looks very similar. (Thanks to Curtis Condon, Ken White, and Al Feng of the Beckman Institute of the University of Illinois for this data and for permission to use it in this publication.)

is adopted in [2], where the real part of the window parameter $c_{t,\omega}$ controls the effective time-duration of the window, and the imaginary part determines the chirp rate (or time-frequency tilt).

The window parameter function $c(t,f)$ is adapted at each time-frequency location (t,f) to maximize a measure of local concentration

$$\max_{c} \frac{\iint |S(u,v)\,\kappa(t-u,f-v)|^{2p}\,du\,dv}{\left(\iint |S(u,v)\,\kappa(t-u,f-v)|^{p}\,du\,dv\right)^{2}}, \tag{5.3.3}$$

with $\kappa(t,f)$ a time-frequency window that tapers to zero away from $(t,f)=(0,0)$ and $p \geq 1$ ([2] takes $p=2$). Maximizing this concentration measure is equivalent to maximizing several other measures of sharpness, focus, peakiness, or kurtosis or minimizing a measure of Rényi entropy [3].

A fairly efficient algorithm in [2] makes extensive use of FFTs both to compute the local concentration measure and the necessary STFTs. In essence, the algorithm computes in parallel a number of STFTs with fixed windows spanning the range of acceptable window parameters, computes the local concentration metric for each window at each time-frequency location via FFT-based fast convolution, and then selects at each time-frequency location the value corresponding to the STFT with the optimal window. The computational cost is approximately two orders of magnitude greater than a single fixed-window STFT.

Support of both time and frequency adaptivity leads to a relatively expensive algorithm. Often, adaptation of the window with frequency is unnecessary. For such cases, a very efficient adaptive window spectrogram has been developed [4]. This method adjusts one or more window parameters (most commonly, window duration) with time to best match the local signal structure. The same local concentration measure (5.3.3) can be optimized. The efficient algorithm computes a few STFTs spanning the range of the adapting window parameter(s), computes the resulting concentration metric for each parameter, and interpolates this data to estimate the exact value of the parameter maximizing the concentration. A window with that optimal value is then applied at the current time to compute the optimal-window STFT using a single FFT.

This simplified time-adaptive-window STFT performs well on many signals and often provides nearly all of the potential benefits of adaptive time-frequency analysis at a cost only a few times that of a conventional fixed-window spectrogram. A similar procedure creates an adaptive-wavelet continuous wavelet transform [4].

The adaptive spectrogram representation for speech signals developed by Glinski adapts the window based on a segmentation (provided by the user) of the signal into pitch periods [5].

5.3.2 Adaptive Quadratic Representations

Quadratic TFRs from Cohen's class can be interpreted as filtered versions of the Wigner distribution. Taking Fourier transforms, this corresponds to weighting the

ambiguity function (AF, Article 5.1) $A(\theta,\tau)$ of the signal by a two-dimensional kernel function $\phi(\theta,\tau)$.

In an adaptive quadratic TFR, we adapt the form of the kernel to match the shape of the signal's components in the ambiguity domain. In particular, we typically desire that ϕ be close to 1 in the vicinity of the signal's auto-components and close to 0 in the vicinity of the signal's cross-components. Again optimization is a key tool.

5.3.2.1 1/0 Kernel Method

Given a signal and its AF, Baraniuk and Jones define the optimal 1/0 kernel as the real, non-negative function ϕ_{opt} that solves the following optimization problem [6]:

$$\max_{\phi} \iint |A(\theta,\tau)\,\phi(\theta,\tau)|^2\,d\theta\,d\tau \tag{5.3.4}$$

subject to

$$\phi(0,0) \;=\; 1 \tag{5.3.5}$$

$$\phi(\theta,\tau) \text{ is radially nonincreasing} \tag{5.3.6}$$

$$\iint |\phi(\theta,\tau)|^2\,d\theta\,d\tau \;\leq\; \alpha, \qquad \alpha \geq 0. \tag{5.3.7}$$

The radially nonincreasing constraint (5.3.6) can be expressed explicitly as

$$\phi(r_1,\psi) \;\geq\; \phi(r_2,\psi) \qquad \forall\, r_1 < r_2, \;\; \forall\, \psi, \tag{5.3.8}$$

where r_i and ψ correspond to the polar coordinates radius and angle, respectively.

The constraints (5.3.5)–(5.3.7) and performance measure (5.3.4) are formulated so that the optimal kernel passes auto-components and suppresses cross-components. The constraints force the optimal kernel to be a lowpass filter of fixed volume α; maximizing the performance measure encourages the passband of the kernel to lie over the auto-components (see Section 3.2.3 and Articles 4.2, 5.1 and 5.2). Both the performance measure and the constraints are insensitive to the orientation angle and aspect ratio (scaling) of the signal components in the (θ,τ) plane. Analysis in [6] shows that the optimal kernel solving (5.3.4)–(5.3.7) takes on essentially only the values of 1 and 0.

By controlling the volume under the optimal kernel, the parameter α controls the tradeoff between cross-component suppression and smearing of the auto-components. Reasonable bounds are $1 \leq \alpha \leq 5$. At the lower bound, the optimal kernel shares the same volume as a spectrogram kernel, while at the upper bound, the optimal kernel smooths only slightly. In fact, as $\alpha \to \infty$, the optimal-kernel distribution converges to the Wigner distribution of the signal.

Analyzing a signal with this optimal-kernel TFR entails a three-step procedure: (1) compute the AF of the signal; (2) solve the linear program (5.3.4)–(5.3.7) in variables $|\phi|^2$ (a fast algorithm is given in [7]); (3) Fourier transform the AF-kernel product $A(\theta,\tau)\,\phi_{\text{opt}}(\theta,\tau)$ to obtain the optimized TFR $C_{\text{opt}}(t,f)$.

5.3.2.2 Radially Gaussian Kernel Method

Although the 1/0 kernel is optimal according to the criteria (5.3.4)–(5.3.7), its sharp cutoff may introduce ringing (Gibbs phenomena) into the optimized TFR, especially for small values of the volume parameter α. For an alternative, direct approach to smooth optimal kernels, explicit smoothness constraints can be appended to the kernel optimization formulation (5.3.4)–(5.3.7). In [8], Baraniuk and Jones constrain the kernel to be Gaussian along radial profiles:

$$\phi(\theta,\tau) \;=\; e^{-\frac{(\theta t_0)^2+(\tau/t_0)^2}{2\sigma^2(\psi)}} \tag{5.3.9}$$

with t_0 a dimension parameter (typically $t_0 = 1\text{s}$).

The term $\sigma(\psi)$ represents the dependence of the Gaussian spread on radial angle ψ. Any kernel of the form (5.3.9) is bounded and radially nonincreasing and, furthermore, smooth if σ is smooth. Since the shape of a radially Gaussian kernel is completely parameterized by this function, finding the optimal, radially Gaussian kernel for a signal is equivalent to finding the optimal function σ_{opt} for the signal. A hybrid gradient/Newton ascent algorithm solving the (nonlinear) system (5.3.4), (5.3.7), (5.3.9) is detailed in [8].

See Figure 5.3.1 for an example on a bat chirp signal.

5.3.2.3 Adaptive Optimal Kernel (AOK) Method

While the 1/0 and radially Gaussian TFRs generally perform well, they are block-oriented techniques that design only one kernel for the entire signal. For analyzing signals with characteristics that change over time; for real-time, on-line operation; or for very long signals, adaptive signal-dependent TFRs are required.

Adaptation of the kernel to track the local signal characteristics over time requires that the kernel optimization procedure consider only the local signal characteristics. An ambiguity-domain design procedure such as the radially Gaussian kernel optimization technique described above does not immediately admit such time localization, since the AF includes information from all times and frequencies in the signal. This difficulty has been surmounted, however, by the development of a time-localized, or short-time, AF [9]. Application of the radially Gaussian kernel optimization procedure to the short-time AF localized at time t_0 produces an optimal kernel $\phi_{\text{opt}}(\theta,\tau;t_0)$ and an optimal-kernel TFR frequency slice $C_{\text{opt}}(t_0,f)$ at time t_0. Since the algorithm alters the kernel at each time to achieve optimal local performance, better tracking of signal changes results.

5.3.2.4 Adaptive Cone-Kernel TFR

The cone-kernel quadratic TFR of Zhao, Atlas, and Marks [10] has many desirable properties, including relatively good time-frequency resolution, good cross-term suppression, and global time-support and time-marginal preservation along with high computational efficiency. However, the performance and inner time-support behav-

ior depend on a cone-length parameter, and mismatch between this parameter and signal component duration can lead to poor resolution and cross-terms [11].

An adaptive cone kernel distribution has been developed that adjusts the cone-length parameter over time to best match the local signal structure according to a short-time energy metric [11]. This technique in effect computes cone-kernel representations with several cone lengths in parallel and determines the optimal cone length at each time (reminiscent of Section 5.3.1). A very efficient recursive algorithm allows the parallel computations to be performed with very little overhead relative to a single cone-length TFR. The net result is an adaptive algorithm with the same order of complexity as the standard short-time Fourier transform. The technique has shown excellent performance on many signals.

5.3.2.5 Adaptive Reduced-Interference TFR

Adaptive TFRs based on other approaches have also been developed. In [12], Krishnamachari and Williams design an adaptive kernel via a projection-based approach using generalized (chirp) time-frequency marginals (see Article 5.2). Cross-term-reduced marginals are generated from short-time Fourier transforms that are relatively free of cross terms. A cross-term-reduced distribution is then reconstructed from these marginals via tomographic principles.

Nickel, Sang, and Williams [13] develop an "iterated projections distribution" by iteratively extracting the negative part of the current TFR, estimating the cross-term from it via a signal-dependent mask function that adaptively identifies cross-term regions in the ambiguity domain, and subtracting the current cross-term estimate from the current TFR estimate. The procedure iterates the above steps until sufficient convergence is achieved. The method shows excellent time-frequency concentration and cross-term suppression in the examples in [13].

5.3.3 Summary and Conclusions

The drive for higher TFR performance (resolution, concentration, etc.) will be satisfied only by adaptive techniques. Optimal TFRs rest on a well-grounded mathematical theory, and new performance metrics or required properties result in new optimization formulations and new TFRs. Moreover, many nonlinear TFRs, such as the reassignment method (Article 7.2), can be regarded as signal-dependent, optimized TFRs. The adaptive TFR concept has also been extended to random signals (Article 12.1). Finally, the same adaptivity principles can be applied to time-scale representations and continuous wavelet transforms.

References

[1] D. L. Jones and T. W. Parks, "A resolution comparison of several time-frequency representations," *IEEE Trans. Signal Processing*, vol. 40, pp. 413–420, February 1992.

[2] D. L. Jones and T. W. Parks, "A high-resolution data-adaptive time-frequency representation," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 2127–2135, December 1990.

[3] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Trans. Information Theory*, vol. 47, pp. 1391–1409, May 2001.

[4] D. L. Jones and R. G. Baraniuk, "A simple scheme for adapting time-frequency representations," *IEEE Trans. Signal Processing*, vol. 42, pp. 3530–3535, December 1994.

[5] S. C. Glinski, "Diphone speech synthesis based on a pitch-adaptive short-time Fourier transform," Master's thesis, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1981.

[6] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Trans. Signal Processing*, vol. 41, pp. 1589–1602, April 1993.

[7] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Fast algorithm for optimal kernel design," *IEEE Trans. Signal Processing*, vol. 42, pp. 134–146, January 1994.

[8] R. G. Baraniuk and D. L. Jones, "Signal-dependent time-frequency analysis using a radially Gaussian kernel," *Signal Processing*, vol. 32, pp. 263–284, June 1993.

[9] D. L. Jones and R. G. Baraniuk, "Adaptive optimal-kernel time-frequency representation," *IEEE Trans. Signal Processing*, vol. 43, pp. 2361–2371, October 1995.

[10] Y. Zhao, L. E. Atlas, and R. J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[11] R. N. Czerwinski and D. L. Jones, "Adaptive cone-kernel time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1715–1719, July 1995.

[12] S. Krishnamachari and W. J. Williams, "Adaptive kernel design in the generalized marginals domain for time-frequency analysis," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, vol. 3, pp. 341–344, Adelaide, Australia, 19–22 April 1994.

[13] R. M. Nickel, T.-H. Sang, and W. J. Williams, "A new signal adaptive approach to positive time-frequency distributions with suppressed interference terms," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98)*, vol. 3, pp. 1777–1780, Seattle, 12–15 May 1998.

5.4 POLYNOMIAL WIGNER-VILLE DISTRIBUTIONS [0]

5.4.1 Polynomial FM Signals

A complex signal can be written in the form

$$z(t) = a(t)\, e^{j\phi(t)}, \tag{5.4.1}$$

where the amplitude $a(t)$ and instantaneous phase $\phi(t)$ are real. Let us define

$$f_z(t) = \frac{\phi'(t)}{2\pi}. \tag{5.4.2}$$

This $f_z(t)$ is the instantaneous frequency (IF) of $z(t)$ *if* $z(t)$ *is analytic.* In this Article we simply use Eqs. (5.4.1) and (5.4.2) as definitions of $z(t)$ and $f_z(t)$, *without* assuming that $z(t)$ is analytic. Results concerning $f_z(t)$ will also be valid for the IF [usually written $f_i(t)$] when $z(t)$ is analytic.

The factor $a(t)$ allows amplitude modulation, phase inversion and time limiting. If $\phi(t)$ is a polynomial function of degree p, so that $f_z(t)$ is a polynomial of degree $p-1$, then $z(t)$ is a **polynomial-phase** or **polynomial FM** signal. If $p > 2$, then $f_z(t)$ is nonlinear, so that $z(t)$ is an example of a **nonlinear FM** signal.

Such nonlinear FM signals occur both in nature and in man-made applications [1]. For example, the sonar systems of some bats use *hyperbolic* and *quadratic* FM signals for echo-location. Some radar systems use *quadratic* FM pulse compression signals. Earthquakes and underground nuclear tests may generate nonlinear FM seismic signals in some long-propagation modes. The altitude and speed of an aircraft may be estimated from the nonlinear IF of the engine noise reaching the ground. Nonlinear FM signals also appear in communications, astronomy, telemetry and other disciplines. As these examples suggest, the problem of estimating the IF of a nonlinear FM signal is of some practical importance.

For a deterministic linear FM signal, the Wigner-Ville distribution (WVD) gives an unbiased estimate of the IF. To obtain the same property with higher-order polynomial FM signals, an extension of the WVD called the **polynomial Wigner-Ville distribution (PWVD)** was defined [2,3]. If the instantaneous phase $\phi(t)$ is a polynomial of degree not exceeding p, then the IF estimate based on a PWVD of order p is unbiased [4].

5.4.2 Principles of Formulation of Polynomial WVDs

We seek a Time-Frequency Distribution of the form

$$P_z(t,f) = \underset{\tau \to f}{\mathcal{F}}\{R_z(t,\tau)\} \tag{5.4.3}$$

[0]Authors: **Boualem Boashash** and **Gavin R. Putland**, Signal Processing Research Centre, Queensland University of Technology, GPO Box 2434, Brisbane, Q 4001, Australia (b.boashash@qut.edu.au, g.putland@qut.edu.au). Reviewers: B. Barkat and LJ. Stanković.

where $R_z(t,\tau)$, called the **signal kernel**,[1] somehow depends on $z(t)$. If $a(t) = 1$, then, for the best possible representation of the IF law, we would like $P_z(t,f)$ to be equal to $\delta(f - f_z(t))$. Making this substitution in Eq. (5.4.3) and taking inverse FTs, we find

$$R_z(t,\tau) = e^{j2\pi f_z(t)\tau} = e^{j\phi'(t)\tau}. \tag{5.4.4}$$

So, while $z(t)$ may have a variable frequency with respect to t, we want $R_z(t,\tau)$ to have a constant frequency w.r.t. τ, namely $f_z(t)$. That is, we want the signal kernel to **dechirp** the signal, yielding a sinusoid of constant frequency, for which the FT is the optimal representation and gives a delta function at $f_z(t)$.

To estimate $\phi'(t)$ in Eq. (5.4.4), we use a central finite-difference (CFD) approximation of the form

$$\phi'(t) \approx \widehat{\phi}'(t) = \frac{1}{\tau} \sum_{l=1}^{q} b_l[\phi(t+c_l\tau) - \phi(t-c_l\tau)] \tag{5.4.5}$$

and choose the dimensionless[2] real coefficients b_l and c_l so that the approximation is exact if $\phi(t)$ is a polynomial of degree p. Let that polynomial be

$$\phi(t) = \sum_{i=0}^{p} a_i t^i \tag{5.4.6}$$

so that

$$\phi'(t) = \sum_{i=1}^{p} i a_i t^{i-1}. \tag{5.4.7}$$

A polynomial of degree p remains a polynomial of degree p if the argument is shifted and scaled. If we shift the time origin so that $t = 0$ in Eqs. (5.4.5) to (5.4.7), we see that if $\phi(t)$ contains only even-power terms, both $\phi'(t)$ and $\widehat{\phi}'(t)$ are zero, so that the even-power terms do not introduce any errors into the estimate. So we may assume that p is even and consider only the odd-power terms in $\phi(t)$. There are $p/2$ such terms, hence $p/2$ degrees of freedom in the coefficients a_i, suggesting that the CFD estimate can be made exact by using only $p/2$ finite differences in Eq. (5.4.5), with uniform sampling intervals. But we shall retain the generality of the sampling instants because we can do so without further algebraic complexity, and because the extra degrees of freedom turn out to be useful. With $q = p/2$, Eq. (5.4.5) becomes

$$\phi'(t) = \frac{1}{\tau} \sum_{l=1}^{p/2} b_l[\phi(t + c_l\tau) - \phi(t - c_l\tau)] \tag{5.4.8}$$

[1] This notation is consistent with the convention that $R_z(t,\tau)$ [Article 2.7] is a generalization of $K_z(t,\tau)$ [Section 2.1.2]. But here, as we shall see, the generalization is in a new direction.

[2] We could allow c_l to have the dimensions of time and dispense with the symbol τ; however, retaining τ will emphasize the correspondence between the PWVD and the ordinary WVD.

where the use of ϕ' instead of $\widehat{\phi'}$ asserts the exactness of the estimate. Substituting this into Eq. (5.4.4), and renaming the signal kernel as $R_z^{(p)}(t,\tau)$ to acknowledge the dependence on p, we obtain

$$R_z^{(p)}(t,\tau) = \prod_{l=1}^{p/2} \left[e^{j\phi(t+c_l\tau)}\, e^{-j\phi(t-c_l\tau)} \right]^{b_l} . \tag{5.4.9}$$

Then, substituting from Eq. (5.4.1) with $a(t) = 1$, we find

$$R_z^{(p)}(t,\tau) = \prod_{l=1}^{p/2} \left[z(t+c_l\tau)\, z^*(t-c_l\tau) \right]^{b_l} . \tag{5.4.10}$$

The resulting TFD, denoted by $W_z^{(p)}(t,f)$ and given by Eq. (5.4.3) as

$$W_z^{(p)}(t,f) = \underset{\tau\to f}{\mathcal{F}} \left\{ R_z^{(p)}(t,\tau) \right\} , \tag{5.4.11}$$

is called a **polynomial Wigner distribution** (or **polynomial WD**) of order p.

Thus we arrive at a general definition: *A polynomial WD of order p of the signal $z(t)$ is a function $W_z^{(p)}(t,f)$ given by Eqs. (5.4.10) and (5.4.11), such that the coefficients b_l and c_l satisfy Eq. (5.4.8) when $\phi(t)$ is a polynomial of degree not exceeding p.* In the special case in which $z(t)$ is analytic, the polynomial WD becomes the **polynomial Wigner-Ville distribution (PWVD)**.

If we put $p=2$, $c_1=\frac{1}{2}$ and $b_1=1$, then $R_z^{(p)}(t,\tau)$ reduces to $z(t+\frac{\tau}{2})\, z^*(t-\frac{\tau}{2})$, which is the instantaneous autocorrelation function (IAF), denoted by $K_z(t,\tau)$. We might therefore describe $R_z^{(p)}(t,\tau)$ as a **polynomial IAF** or **higher-order IAF**.

Eq. (5.4.10) shows that for any $z(t)$, the polynomial IAF is Hermitian in τ. It follows that the polynomial WD is *real*.

5.4.3 IF Estimates with Zero Deterministic Bias

When the IF is a polynomial of degree not exceeding $p-1$, the PWVD of order p gives an *unbiased* estimate of the IF law, as is shown by the following result.

Theorem 5.4.1: *If $z(t)$ and $f_z(t)$ are given by Eqs. (5.4.1) and (5.4.2), where $\phi(t)$ is a polynomial of degree not exceeding p, and if $W_z^{(p)}(t,f)$ satisfies the general definition of a pth-order polynomial WD of $z(t)$, then $W_z^{(p)}(t,f)$ is symmetrical in f about $f=f_z(t)$.*

Proof: Substituting Eq. (5.4.1) into Eq. (5.4.10) and simplifying, we find

$$R_z^{(p)}(t,\tau) = R_a^{(p)}(t,\tau) \exp\left(j \sum_{j=1}^{p/2} b_l [\phi(t+c_l\tau) - \phi(t-c_l\tau)] \right) \tag{5.4.12}$$

where $R_a^{(p)}(t,\tau)$, as the notation suggests, is the polynomial IAF for $a(t)$, given by

$$R_a^{(p)}(t,\tau) = \prod_{l=1}^{p/2} \left[a(t+c_l\tau)\, a(t-c_l\tau) \right]^{b_l} . \tag{5.4.13}$$

Because $\phi(t)$ is a polynomial function of degree not exceeding p, and because $W_z^{(p)}(t,f)$ is a pth-order polynomial WD, Eq. (5.4.8) is applicable, so that Eq. (5.4.12) becomes

$$R_z^{(p)}(t,\tau) = R_a^{(p)}(t,\tau)\, e^{j\phi'(t)\,\tau} = R_a^{(p)}(t,\tau)\, e^{j2\pi f_z(t)\,\tau}. \tag{5.4.14}$$

Taking Fourier transforms ($\tau \rightarrow f$), we find

$$W_z^{(p)}(t,f) = W_a^{(p)}(t,f) \underset{f}{*} \delta(f - f_z(t)) = W_a^{(p)}(t, f - f_z(t)) \tag{5.4.15}$$

where $W_a^{(p)}(t,f)$ is the corresponding polynomial WD of $a(t)$:

$$W_a^{(p)}(t,f) = \underset{\tau \rightarrow f}{\mathcal{F}} \left\{ R_a^{(p)}(t,\tau) \right\}. \tag{5.4.16}$$

From Eq. (5.4.13) we see that $R_a^{(p)}(t,\tau)$ is real and even in τ. Hence, from Eq. (5.4.16), $W_a^{(p)}(t,f)$ is real and even in f. Then, from Eq. (5.4.15), $W_z^{(p)}(t,f)$ is real and symmetrical in f about $f = f_z(t)$. ■

Because a symmetrical distribution is symmetrical about its first moment, Theorem 5.4.1 has the following corollary: *If $z(t)$ has polynomial phase of degree not exceeding p, and if $W_z^{(p)}(t,f)$ is a pth-order polynomial WD of $z(t)$, then the first moment of $W_z^{(p)}(t,f)$ w.r.t. f is equal to $f_z(t)$.*

Being unbiased for higher-degree polynomial FM signals, PWVDs can solve problems that quadratic TFDs cannot [2]. PWVDs also give optimal frequency resolution in the sense that they are FTs of maximal-length polynomial IAFs derived from the full-length signal.

5.4.4 Calculation of Coefficients

Applying Eq. (5.4.6) in Eq. (5.4.8) gives

$$\phi'(t) = \frac{1}{\tau} \sum_{l=1}^{p/2} \left[b_l \sum_{i=0}^{p} a_i \left[(t + c_l\tau)^i - (t - c_l\tau)^i \right] \right]. \tag{5.4.17}$$

Eqs. (5.4.7) and (5.4.17) give two expressions for $\phi'(t)$. Equating these expressions, and shifting and scaling the time variable so that $t = 0$ and $\tau = 1$, we obtain

$$a_1 = \sum_{l=1}^{p/2} \left[b_l \sum_{i=0}^{p} a_i \left[{c_l}^i - (-c_l)^i \right] \right] = 2 \sum_{l=1}^{p/2} \left[\sum_{i=1,3,\ldots}^{p-1} a_i\, {c_l}^i \right] b_l. \tag{5.4.18}$$

In the left-hand and right-hand expressions of this equation, the coefficient of a_i is zero for all even values of i, justifying the decision to consider only odd values of i. Equating coefficients of a_1 gives

$$\frac{1}{2} = \sum_{l=1}^{p/2} c_l b_l \ , \tag{5.4.19}$$

and equating coefficients of a_3 , a_5 ,... , a_{p-1} gives

$$0 = \sum_{l=1}^{p/2} {c_l}^i \, b_l \; ; \quad i = 3, 5, \ldots, p-1. \tag{5.4.20}$$

The last two equations can be written in matrix form as

$$\begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_{p/2} \\ {c_1}^3 & {c_2}^3 & {c_3}^3 & \cdots & {c_{p/2}}^3 \\ {c_1}^5 & {c_2}^5 & {c_3}^5 & \cdots & {c_{p/2}}^5 \\ \vdots & \vdots & \vdots & & \vdots \\ {c_1}^{p-1} & {c_2}^{p-1} & {c_3}^{p-1} & \cdots & {c_{p/2}}^{p-1} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{p/2} \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.4.21}$$

and solved algebraically or numerically.

Eq. (5.4.21) is underdetermined: $p/2$ equations in p unknowns leave $p/2$ degrees of freedom. Various ways of exploiting the degrees of freedom give rise to various forms of the PWVD. Here we mention two forms described in [2, p. 217].

If we decide that we want uniform sampling intervals, we choose the values of c_l and accept the resulting values of b_l , which in general turn out to be fractions. The resulting form of the PWVD is called the **fractional-powers form** or **Form I**, of which an example is given in [3, p. 550]. The need to compute fractional powers is a cause of inefficiency.

Alternatively, if we decide that the polynomial IAF must contain only positive integer powers, we choose the values of b_l and accept the resulting values of c_l , which in general give non-uniform sampling intervals. The resulting form of the PWVD is called the **integer-powers form** or **Form II**. In a discrete-time implementation, non-uniform sampling requires interpolation, but this is more efficient than computing non-integer powers. The number of interpolations required can be reduced by time-scaling [2, p. 218].

(There is also a "complex-time form" or "Form III" based on an analytic extension of the signal with a complex time argument; this is described in [5], with further details in [6].)

Notice that given one solution to Eq. (5.4.21), we can obtain another solution by changing the sign of any c_l and of the corresponding b_l. Thus, from one solution, we can always find another solution in which all the indices b_l are positive, ensuring that the polynomial IAF does not become unbounded as the signal approaches zero.

5.4.5 Examples

In the trivial case for which $p = 2$ (quadratic phase, linear FM), Eq. (5.4.21) reduces to $c_1 b_1 = 1/2$. If we fix the sampling points by choosing $c_1 = 1/2$, that leaves $b_1 = 1$. Alternatively, if we choose $b_1 = 1$ for unit powers, that leaves $c_1 = 1/2$. Thus we have a degenerate case in which Forms I and II overlap. Substituting for

p, b_1 and c_1 in Eq. (5.4.8), we obtain the simplest possible CFD approximation:

$$\phi'(t) = \frac{1}{\tau}\left[\phi(t+\tfrac{\tau}{2}) - \phi(t-\tfrac{\tau}{2})\right]. \tag{5.4.22}$$

The same substitutions in Eq. (5.4.10) yield the polynomial IAF

$$R_z^{(2)}(t,\tau) = z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2}). \tag{5.4.23}$$

This is just the ordinary IAF $K_z(t,\tau)$, which when substituted into Eq. (5.4.11) yields the ordinary (quadratic) Wigner distribution. So, for $p = 2$, Theorem 5.4.1 says that the Wigner distribution is symmetrical about $f_z(t)$ if $z(t)$ is a quadratic-phase (i.e. linear FM) signal. This confirms that the WVD gives an unbiased estimate of the IF for deterministic linear FM signals.

In the case for which $p = 4$ (quartic phase, cubic FM), Eq. (5.4.21) reduces to the 2×2 system

$$\begin{aligned} c_1 b_1 + c_2 b_2 &= 1/2 \\ {c_1}^3 b_1 + {c_2}^3 b_2 &= 0 \end{aligned} \quad . \tag{5.4.24}$$

Form I: If we take $c_1 = 1/4$ and $c_2 = -1/2$, Eqs. (5.4.24) become a linear system with solutions $b_1 = 8/3$, $b_2 = 1/3$. Substituting these values into Eq. (5.4.10) gives

$$R_z^{(4)}(t,\tau) = \left[z(t+\tfrac{\tau}{4})\, z^*(t-\tfrac{\tau}{4})\right]^{\frac{8}{3}} \left[z(t-\tfrac{\tau}{2})\, z^*(t+\tfrac{\tau}{2})\right]^{\frac{1}{3}} \tag{5.4.25}$$

where the superscript "(4)" indicates order 4.

Form II: If we take $b_1 = 2$ and $b_2 = 1$, Eqs. (5.4.24) become a nonlinear system with solutions

$$c_1 = \left[2(2-2^{1/3})\right]^{-1} \approx 0.6756 \;\; ; \;\; c_2 = -2^{1/3} c_1 \approx -0.8512 \, . \tag{5.4.26}$$

Substituting for b_1 and b_2 in Eq. (5.4.10) gives

$$R_z^{(4)}(t,\tau) = \left[z(t+c_1\tau)\, z^*(t-c_1\tau)\right]^2 \, z(t+c_2\tau)\, z^*(t-c_2\tau). \tag{5.4.27}$$

where c_1 and c_2 are as in Eqs. (5.4.26). This is one of an infinite number of Form-II solutions. More details on the design procedure are given in [1,2,7] and in Article 5.5 (next).

Fig. 5.4.1 illustrates the effectiveness of the PWVD in suppressing artifacts caused by errors in the CFD estimate of the IF. The IF of the test signal is

$$f_z(t) = f_c + f_d\, C_5(t/T) \;\; ; \;\; 0 \le t \le T \tag{5.4.28}$$

where f_c is the center frequency, f_d is the minimax frequency deviation, and C_5 is the Chebyshev polynomial of degree 5. Trace (a) shows the ordinary Wigner-Ville distribution (WVD), which clearly cannot handle the nonlinear IF law. The Form-II PWVD shown in trace (b) was computed by the "*TFSA*" software toolbox, developed in-house by the Signal Processing Research Centre, QUT. It is of 6th order, so that its signal kernel exactly dechirps the 5th-degree IF law. The superiority of the PWVD is evident, as is its symmetry about the IF law.

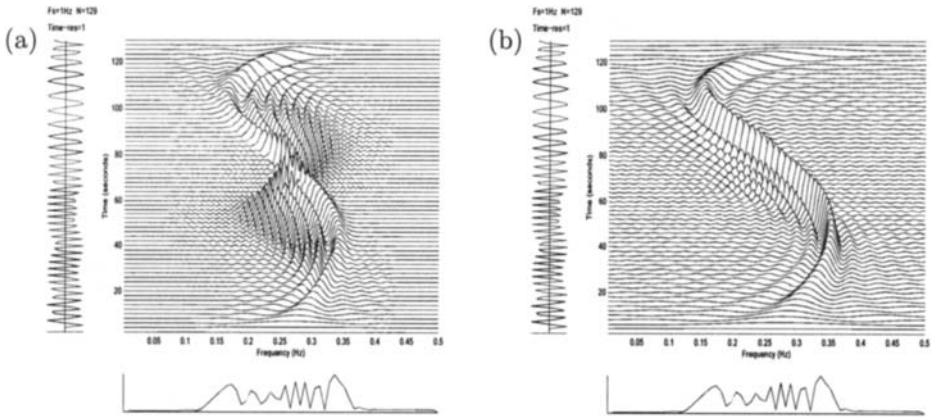


Fig. 5.4.1: Time-frequency representations of a 5th-degree Chebyshev polynomial FM signal [Eq. (5.4.28)] with center frequency $f_c = 0.25$ Hz, minimax frequency deviation $f_d = 0.1$ Hz, duration $T = 128$ seconds: (a) WVD; (b) 6th-order Form-II PWVD. Both TFDs are unwindowed. Each plot shows time vertically (range 0 to 128 s; resolution 2 s) and frequency horizontally (range 0 to 0.5 Hz), and has the time trace at the left and the magnitude spectrum at the bottom.

5.4.6 Multicomponent Signals and Polynomial TFDs

The use of TFDs for multicomponent signal analysis requires a reliable method of suppressing cross-terms (see Section 3.1.2 and Articles 4.2 and 5.2).

For *quadratic* TFDs, the mechanisms of generation and suppression of cross-terms are well understood. In the WVD, each cross-term is generated midway between the interacting components and alternates at a rate proportional to the separation between the components (see Article 4.2). Other quadratic TFDs may be obtained from the WVD by 2D low-pass filtering, which attenuates the cross-terms because of their alternating (high-pass) character.

For higher-order *polynomial* TFDs, there is the added difficulty that *cross-terms do not necessarily alternate*, so that it may not be possible to suppress cross-terms entirely by convolving them with a simple smoothing function in the (t, f) plane. However, if polynomial TFDs are implemented according to the **S-method**, as described in Article 6.2, the generation of cross-terms can be avoided [8].

5.4.7 Summary and Conclusions

A PWVD of degree p is derived from a p^{th}-order CFD approximation to the derivative of the instantaneous phase. It reduces to the ordinary WVD if $p = 2$. It is real and symmetrical about the IF law of an FM signal whose instantaneous phase is a polynomial of degree not exceeding p (i.e. whose IF is a polynomial of degree not exceeding $p-1$), even if that signal is also amplitude-modulated.

This topic is developed further in Article 5.5 (next). More properties of PWVDs are given in [1]. Some implementation issues are discussed in [4] and [7]. The effect of noise is considered in Article 10.4.

References

[1] B. Boashash and B. Ristic, "Polynomial time-frequency distributions and time-varying higher order spectra: Application to the analysis of multicomponent FM signal and to the treatment of multiplicative noise," *Signal Processing*, vol. 67, pp. 1–23, May 1998.

[2] B. Boashash and P. J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 216–220, January 1994.

[3] L. Stanković and S. Stanković, "An analysis of instantaneous frequency representation using time-frequency distributions—Generalized Wigner distribution," *IEEE Trans. Signal Processing*, vol. 43, pp. 549–552, February 1995.

[4] B. Boashash, "Time-Frequency Signal Analysis: Past, present and future trends," in *Digital Control and Signal Processing Systems and Techniques* (C. T. Leondes, ed.), no. 78 in CONTROL AND DYNAMIC SYSTEMS: ADVANCES IN THEORY AND APPLICATIONS, ch. 1, pp. 1–69, San Diego: Academic Press, 1996.

[5] S. Stanković and L. Stanković, "Introducing time-frequency distribution with a 'complex-time' argument," *Electronics Letters*, vol. 32, pp. 1265–1267, July 1996.

[6] S. Stanković and L. Stanković, "Approach to the polynomial Wigner distributions," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 153–156, Paris, 18–21 June 1996.

[7] B. Barkat and B. Boashash, "Design of higher order polynomial Wigner-Ville distributions," *IEEE Trans. Signal Processing*, vol. 47, pp. 2608–2611, September 1999.

[8] L. Stanković, S. Stanković, and I. Djurović, "Architecture for realization of the cross-terms free polynomial Wigner-Ville distribution," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. III, pp. 2053–2056, Munich, 21–24 April 1997.

5.5 DESIGN OF POLYNOMIAL TFDs, WITH APPLICATIONS [0]

5.5.1 Decompositions of Polynomial Derivatives

In many technological applications such as radar, sonar and communications, signals can be modeled as higher-order polynomial phase signals (PPSs) with constant or slowly time-varying amplitude (see Section 5.4.1). The Wigner-Ville Distribution (WVD) can process PPSs having polynomials of degree less than or equal to two. In the case of higher degrees, methods based on higher-order statistics seem to be necessary to process such signals [1,2]. In [3], the Polynomial Phase Transform (PPT) was introduced for the estimation of constant amplitude PPSs. The kernel of this transform is based on an expression of a polynomial derivative using the principle of the finite difference. In [4, 5] an exact decomposition of the derivatives of any order of a polynomial $\phi(t)$ is proposed in terms of shifted versions of this polynomial, i.e., $\phi(t-t_0), ..., \phi(t-t_n)$. This decomposition is used to design time-frequency distributions that generalize the classical WVD and Ambiguity Function (AF). The relationships between these distributions, the PPT and the Polynomial Wigner-Ville Distribution (PWVD) are discussed. The estimation of the coefficients appearing in the model of a PPS affected by multiplicative and additive noise using the high-order ambiguity function is discussed in [6]. The analysis of the performances of the proposed estimators is studied in details in [5] and [6]. In this article, we recall the main results concerning the use of polynomials in the design of higher order time frequency distributions. Let $t_0, \ldots, t_Q$ denote $Q+1$ arbitrary distinct real numbers, $\mathcal{R}_N[t]$ the set of all polynomials of degree $\leq N$ and $\phi^{(\ell)}(t)$ the derivative of order ℓ of the polynomial ϕ. Consider the decomposition of $\phi^{(\ell)}(t)$, in the form :

$$\phi^{(\ell)}(t) = \sum_{k=0}^{Q} \alpha_k^{\ell} \phi(t - t_k) \qquad \forall t, \quad (\ell = 0, 1, \ldots, N) \tag{5.5.1}$$

where, for a given ℓ, the coefficients α_k^{ℓ} depend only on the t_k's and are solution of the linear system given in the Appendix. Then the 3 following properties are equivalent [4,5] :
(i)–At least one polynomial ϕ of degree N satisfies (5.5.1)
(ii)–All the polynomials of $\mathcal{R}_N[t]$ satisfy (5.5.1).
(iii)–The polynomial ϕ satisfies the following identity

$$\tau^{\ell} \phi^{(\ell)}(t) = \sum_{k=0}^{Q} \alpha_k^{\ell} \phi(t - t_k \tau) \quad \forall \tau \quad \forall t \quad (\ell = 0, 1, \ldots, N). \tag{5.5.2}$$

---

[0]Author : **Messaoud Benidir**, LSS-Supelec, Université de Paris-Sud, France (benidir@lss.supelec.fr). Reviewers : B. Barkat and B. Picinbono.

The existence of the α_k^ℓ's and their expressions in terms of the t_k are discussed in [4]. In particular, for

$$\phi(t) = \sum_{i=0}^{N} a_i t^i, \quad a_N \neq 0 \tag{5.5.3}$$

it is shown that the smallest integer Q for which the decomposition (5.5.1) is possible is $Q = N$ if $\ell = N$ and $Q = N - 1$ if $\ell < N$ assuming in the last case that

$$t_0 + \ldots + t_{N-1} = 0. \tag{5.5.4}$$

For instance, if $\ell = N$, one can choose $Q = N$ and (5.5.2) becomes

$$\sum_{k=0}^{N} \alpha_k^N \, \phi(t - \tau t_k) = \tau^N N! a_N \quad \forall t. \tag{5.5.5}$$

If $\ell = N - 1$, assuming condition (5.5.4), one can choose $Q = N - 1$ and (5.5.2) becomes

$$\sum_{k=0}^{N-1} \alpha_k^{N-1} \, \phi(t - \tau t_k) = \tau^{N-1}[(N-1)! a_{N-1} + N! a_N t] \quad \forall t \quad \forall \tau. \tag{5.5.6}$$

Example 5.5.1: Particular parameters t_k's and their associated α_k^ℓ's

Consider the important case $\ell = Q$ and

$$t_k = k + c, \quad k = 0, \ldots, Q \tag{5.5.7}$$

where c is an arbitrary parameter for $Q = N$ and $c = -Q/2$ for $Q = N - 1$ in order to satisfy (5.5.4). In both cases $Q = N$ and $Q = N - 1$, the α_k^Q's appearing in (5.5.1) are given by

$$\alpha_k^Q = (-1)^k C_Q^k \triangleq (-1)^k c_k, \quad k = 0, \ldots, Q \tag{5.5.8}$$

where the C_Q^k denote the number of combinations.

5.5.2 Design of Time-Frequency Distributions

We consider $Q + 1$ arbitrary real and distinct parameters : $t_0, \ldots, t_Q$, and denote by $\alpha_0^\ell, \ldots, \alpha_Q^\ell$ the corresponding coefficients appearing in (5.5.2). We introduce the following kernel

$$K_{Q+1}^\ell(z, t, \tau) \triangleq \prod_{k=0}^{Q} (z(t - t_k\tau))^{\alpha_k^\ell} \tag{5.5.9}$$

where $z(t)$ is a PPS defined by

$$z(t) = b(t)e^{j\phi(t)}, \quad \phi(t) = \sum_{i=0}^{N} a_i t^i \quad \text{with} \quad N > 1. \tag{5.5.10}$$

Let $z(t)$ be a constant amplitude PPS with $b(t) = b_0$ and $N \leq Q+1$. Using property (5.5.2) and relations (5.5.5), (5.5.6), the kernel $\mathcal{K}^{\ell}_{Q+1}$ can be written in the following forms according to the value of ℓ :

$$\begin{aligned}
\mathcal{K}^{\ell}_{Q+1}(z,t,\tau) &= \exp\left[j\sum_{k=0}^{Q}\alpha_k^{\ell}\phi(t-t_k\tau)\right] \\
&= \exp\left[j\phi^{(\ell)}(t)\right] \quad \text{for} \quad \ell = 0,1,\ldots,N \\
&= \exp\left[j\tau^{N-1}[(N-1)!a_{N-1} + N!a_N t]\right] \quad \text{for} \quad \ell = N-1 \\
&= \exp\left[j\tau^N N! a_N\right] \quad \text{for} \quad \ell = N
\end{aligned} \tag{5.5.11}$$

The Fourier Transform of the kernel $\mathcal{K}^{\ell}_{Q+1}$ with respect to t or τ leads to general time-frequency distributions that will be analyzed in the following sections.

5.5.2.1 Generalized Ambiguity Function

The Generalized Ambiguity Function (GAF) is defined by :

$$z(t) \rightarrow \mathcal{A}^{\ell}_{Q+1}(z,\theta,\tau) \triangleq \int_0^T \mathcal{K}^{\ell}_{Q+1}(z,t,\tau)\, e^{-j\theta t} dt. \tag{5.5.12}$$

where the kernel $\mathcal{K}^{\ell}_{Q+1}$ is given by (5.5.9).

In the following we consider the case $\ell = N-1$ and introduce the notation $\mathcal{A}^{N-1}_{Q+1} \triangleq \mathcal{A}_{Q+1}$. For a PPS given by (5.5.10) with a constant amplitude $b(t) = b_0$, one has [5]

$$\mathcal{A}_{Q+1}(z,\theta,\tau) = T\exp\left(j\left[(N-1)!a_{N-1}\tau^{N-1} + \frac{\Phi}{2}T\right]\right)\operatorname{sinc}\left(\frac{\Phi}{2}T\right)$$

where $\Phi \triangleq N!a_N\tau^{N-1} - \theta$ and

$$\arg\max_{\theta} |\mathcal{A}_{Q+1}(z,\theta,\tau)| = N!a_N\tau^{N-1}. \tag{5.5.13}$$

The function $\mathcal{A}_{Q+1}$ is thus independent of the parameters t_k and it can be seen as a generalization of the PPT defined in [3,7] by :

$$P_{Q+1}(z,\theta,\tau) \triangleq \int_0^T \prod_{k=0}^{Q} \left(z^{\$k}(t-k\tau)\right)^{c_k} e^{-j\theta t} dt, \quad c_k \triangleq C_Q^k$$

where $z^{\$k} = z$ if k is even and z^* if k is odd. For instance, if the t_k's are given by (5.5.7), taking into account (5.5.2.1), one can easily deduce from the properties of the PPT [3] that, for a constant amplitude PPS given by (5.5.10), one has in both cases $Q = N$ and $Q = N - 1$

$$\mathcal{A}_{Q+1}(z,\theta,\tau) = b_0^{-2^{N-1}} \exp(j\Psi) P_N(z,\theta,\tau) \tag{5.5.14}$$

where $\Psi \triangleq \frac{N!(N-1)}{2} a_N \tau^N$ and

$$\arg\max_{\theta} |\mathcal{A}_{Q+1}(z,\theta,\tau)| = \arg\max_{\theta} |P_N(z,\theta,\tau)|. \tag{5.5.15}$$

The discrete version of $\mathcal{A}_{Q+1}$, denoted by $\underline{\mathcal{A}}_{Q+1}$, takes the form [5] :

$$\underline{\mathcal{A}}_{Q+1}(z,\theta,\tau) \triangleq \sum_{n=N_i}^{N_f} \prod_{k=0}^{N-1} z(n - t_k\tau)^{\alpha_k^Q} \exp(-jn\Delta\theta)$$

where Δ is the sampling period. The result (5.5.13) shows that the GAF has the ability to detect constant amplitude PPSs, similarly to the ability of the FT to detect pure frequencies. This gives a method for determining the degree, N, of the polynomial phase and its higher-order coefficient a_N [3, 5, 7]. This method may be iterated and allows the determination of $a_{N-1}, \ldots, a_1$. To obtain a_0 it is sufficient to calculate the phase of the demodulated signal $z(t)\exp\{-j\sum_{i=1}^{N} a_i t^i\}$.

Example 5.5.2: Estimation of the phase using the GAF

Consider a third-order PPS with the phase

$$\Phi(t) = a_3 t^3 + a_2 t^2 + a_1 t + a_0, \quad a_3 = 0.0063, a_2 = 0.2513, a_1 = 7.854, a_0 = 1.0003.$$

The spectrum of the signal affected by an additive noise is shown in Fig. 5.5.1(a). Applying (5.5.13), one obtains [5] the results given in Fig. 5.5.1(b), Fig. 5.5.1(c) and Fig. 5.5.1(d).

From the peak of $|\underline{\mathcal{A}}_3(z,\theta,\tau)|$, Fig. 5.5.1(b) reveals a peak in the neighborhood of 15.1 and this leads to an estimator $\widehat{a}_{3,A} = 0.0063$ of a_3.

From the peak of $|\underline{\mathcal{A}}_2(z_1(t)\},\theta,\tau)|$, $z_1(t) = z(t)\exp(-j\hat{a}_{3,A}t^3)$, Fig. 5.5.1(c) reveals a peak giving an estimator $\widehat{a}_{2,A} = 0.2513$, of a_2.

From the peak of $|\underline{\mathcal{A}}_1(z_2(t)\},\theta,\tau)|$, $z_2(t) = z_1(t)\exp(-j\hat{a}_{2,A}t^2)$, Fig. 5.5.1(d) reveals a peak giving an estimator $\widehat{a}_{1,A} = 7.854$, of a_1.

5.5.2.2 Generalized Wigner Distribution

The Generalized Wigner Distribution (GWD) is defined by [5] :

$$z(t) \rightarrow \mathcal{W}^{\ell}_{Q+1}(z,t,\omega) = \int_0^T \mathcal{K}^{\ell}_{Q+1}(z,t,\tau)\ \ell\ \tau^{\ell-1}\ e^{-j\omega\tau^{\ell}} d\tau \quad \ell \neq 0$$

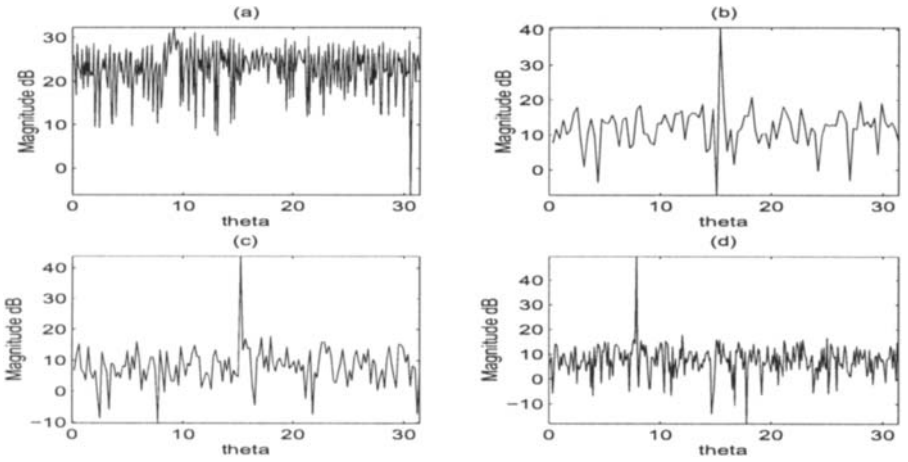


Fig. 5.5.1 Estimation of coefficients a_3, a_2, a_1 : (b) yields $\widehat{a}_3$, (c) yields $\widehat{a}_2$, (d) yields $\widehat{a}_1$.

where the kernel $\mathcal{K}^{\ell}_{Q+1}$ *is given by (5.5.9).* For $\ell = 1$, one obtains the PWVD class over the interval $[0, T]$ [8].

For a constant amplitude PPS given by (5.5.10), the GWD is independent of the parameters t_k for $Q \geq N$ and one has [5] :

$$\mathcal{W}^{\ell}_{Q+1}(z,t,\omega) = T^{\ell} \exp\left(j\frac{\phi^{(\ell)}(t)-\omega}{2}T^{\ell}\right)\operatorname{sinc}\left(\frac{\phi^{(\ell)}(t)-\omega}{2}T^{\ell}\right) \tag{5.5.16}$$

$$\arg\max_{\omega} |\mathcal{W}^{\ell}_{Q+1}(z,t,\omega)| = \phi^{(\ell)}(t). \tag{5.5.17}$$

This shows that $|\mathcal{W}^{\ell}_{Q+1}(z,t,\omega)|$ is independent of the parameters t_k and it produces a maximum at $\omega = \phi^{(\ell)}(t)$ in the plane (t,ω).

Application of the particular GWD $\mathcal{W}^{N}_{Q+1}$ If the signal $z(t)$ is a constant amplitude PPS given by (5.5.10) and $\ell = N$, one can choose $Q = N$ and the notation $\mathcal{W}_{N+1} \triangleq \mathcal{W}^{N}_{Q+1}$ is introduced. In this case, the result (5.5.17) gives :

$$\arg\max_{\omega} |\mathcal{W}_{N+1}(z,t,\omega)| = N! a_N. \tag{5.5.18}$$

That is $|\mathcal{W}_{N+1}(z,t,\omega)|$ produces a maximum at $\omega = N! a_N$ in the plane (t,ω). This leads to a procedure that allows us to determine the degree of the phase and the value of its higher-order coefficient [5]. This procedure may be iterated and one can determine the other coefficients $a_{N-1}, \ldots, a_1$. Finally to obtain a_0, it is sufficient to calculate the phase of $z(t)\exp\{-j\sum_{i=1}^{N} a_i t^i\}$. According to the result above, the distribution $\mathcal{W}_{N+1}$ can be computed by choosing arbitrary parameters t_k. For

instance, if the t_k are given by (5.5.7) then the α_k are given by (5.5.8). The discrete version of $\mathcal{W}_{N+1}$ is defined by :

$$\underline{\mathcal{W}}_{N+1}(z,n,\omega) = N \sum_{m=1}^{N_f'} \prod_{k=0}^{N} (z^{\$k}(n-t_k m))^{\alpha_k^Q} \, (m\Delta)^{N-1} \, e^{-j(m\Delta)^N \omega}. \qquad (5.5.19)$$

5.5.3 Estimation of the Phase of a PPS

Consider the problem of estimating the phase of a PPS signal

$$y(t) = z(t) + w(t), \quad z(t) = b(t)e^{j\phi(t)}$$

where ϕ is of degree N assumed to be known, $w(t)$ is a complex circular $\mathcal{N}(0,\sigma_w^2)$ white noise and $b(t)$ a constant amplitude or a multiplicative Gaussian white noise. Introducing the notation $x_n = x(n\Delta)$ where Δ is the sampling period, the discrete-version of this model is given by :

$$y_n = b_n e^{j\phi_n} + w_n, \quad \phi_n = \sum_{i=0}^{N} a_i (n\Delta)^i \quad \text{for} \quad n = 1, ..., N_e. \qquad (5.5.20)$$

5.5.3.1 Case of Constant Amplitude and Additive Noise

Consider the case of a constant amplitude $b(t) = b_0$. Using the GAF and the GWD, two iterative algorithms to estimate the polynomial phase are proposed in [5].

The GAF-based estimator We take $Q = N - 1 = \ell$ and the t_k's given by (5.5.7). In the discrete version of the GAF, $\underline{\mathcal{A}}_N$ (5.5.2.1), one choose $t_k = k - \frac{N-1}{2}$, $N_i = 1 + \frac{N-1}{2}\tau$, $N_f = N_e - \frac{N-1}{2}\tau$, $c_k \stackrel{\Delta}{=} C_{N-1}^k$ and τ satisfying $0 < \tau < \frac{N_e-1}{N-1}$. It is shown [5] that the first-order moment of the GAF is given by

$$E\left[\underline{\mathcal{A}}_N(y,\theta,\tau)\right] = \underline{\mathcal{A}}_N(z,\theta,\tau), \forall\tau \qquad (5.5.21)$$

and, for the complex signal $z(n)$ given by (5.5.10), we have

$$\arg\max_{\theta} \left| E\left[\underline{\mathcal{A}}_N(y,\theta,\tau)\right]\right| = N!(\Delta\tau)^{N-1} a_N. \qquad (5.5.22)$$

This shows that $|E\left[\underline{\mathcal{A}}_N(y,\theta,\tau)\right]|$ produces a maximum at $\theta = N!(\Delta\tau)^{N-1}a_N$ and a_N can thus be determined using (5.5.22). The remaining coefficients can be computed via a recursive algorithm in the same way as in Example 5.5.2 [5].

Property 5.5.1: *[5] It is always possible to choose τ in order to ensure that the variables $y(n - t_k\tau)$ and $y(m - t_l\tau)$ are independent for $n \neq m$. For such values of τ, the variance of $\underline{\mathcal{A}}_N(y,\theta,\tau)$ is given by :*

$$\text{var}\left\{\underline{\mathcal{A}}_N(y,\theta,\tau)\right\} = b_0^{2^N} \, (N_e - (N-1)\tau) \;\; \mathcal{G}(\text{SNR}) \qquad (5.5.23)$$

where $\mathrm{SNR} = b_0^2/\sigma_w^2$ *and*

$$\mathcal{G}(\mathrm{SNR}) = \prod_{k=0}^{N-1} \left[\sum_{i=0}^{c_k} C_{c_k}^{i^2} i! \left(\frac{1}{\mathrm{SNR}} \right)^i \right] - 1. \tag{5.5.24}$$

The result (5.5.23) shows that the variance of $\underline{A}_N(y,\theta,\tau)$ depends on the degree N of the polynomial phase, the number of samples N_e, the SNR and the delay parameter τ. It does not however depend on the coefficients of the polynomial phase.

Denoting by $\underline{P}_N(y,\theta,\tau)$ the discrete version of the PPT and according to the result of the second-order moment of $\underline{P}_N(y,\theta,\tau)$ (given by Eq. (4.7) in [3]), it is shown that [5] :

$$\mathrm{var}\,\{\underline{A}_N(y,\theta,\tau)\} = \mathrm{var}\,\{\underline{P}_N(y,\theta,\tau)\} \qquad \text{for } \tau = \frac{N_e}{N} \tag{5.5.25}$$

For a large number of samples, the estimator defined by

$$\widehat{a}_{N,A} = \frac{1}{N!(\tau\Delta)^{N-1}} \arg\max_{\theta} |\underline{A}_N(y,\theta,\tau)| \tag{5.5.26}$$

is unbiased and its variance can be approximated by [5]

$$\mathrm{var}\{\widehat{a}_{N,A}\} \approx \frac{6}{(N!)^2 \tau^{2N-2} L^3 \Delta^{2N}}\ \mathcal{G}_N(\mathrm{SNR}) \tag{5.5.27}$$

where $L = N_e - (N-1)\tau$. Denoting by $\mathrm{var}\{\widehat{a}_{N,P}\}$ the variance of the estimator of a_N derived from the peak of the PPT, it is shown [5] that :

$$\mathrm{var}\{\widehat{a}_{N,A}\} = \mathrm{var}\{\widehat{a}_{N,P}\} \quad \text{for} \quad \tau = \frac{N_e}{N} \tag{5.5.28}$$

and for high SNR, this variance can be approximated by

$$\mathrm{var}\{\widehat{a}_{N,A}\} \approx \frac{6C_{2N-2}^{N-1}}{\mathrm{SNR}(N!)^2 \tau^{2N-2} L^3 \Delta^{2N}}. \tag{5.5.29}$$

The GWV distribution based estimator Let us now consider the estimator of a_N, based on the GWV distribution. We take $Q = N = \ell$ and the t_k's given by (5.5.7). The parameters appearing in the discrete version $\underline{\mathcal{W}}_N$ (5.5.19) are $t_k = k+c$, $N_f' = \frac{N_e-1}{N+c}$, $c_k' = C_N^k$ and $0 < c < N_e - 1 - N$. As it is the case for the estimator $\underline{A}_N(y,\theta,\tau)$, it is shown [5] that for any value of c, one has

$$E\,[\underline{\mathcal{W}}_N(y,n,\omega)] = \underline{\mathcal{W}}_N(z,n,\omega) \tag{5.5.30}$$

$$\arg\max_{\omega} |E\,[\underline{\mathcal{W}}_N(y,n,\omega)]\,| = N! a_N. \tag{5.5.31}$$

Thus coefficient a_N of ϕ can be determined via (5.5.31) and the other coefficients by using the same procedure as in the GAF. In addition, similar results to those proposed in Property 5.5.1 are established in [5]. Denoting by $\mathcal{N}_{\mathrm{A}}$ and $\mathcal{N}_{\mathrm{W}}$ the minimal value of N_e to be treated respectively with the GAF and $\underline{\mathcal{W}}_N$, one has

$$\mathcal{N}_{\mathrm{W}} \approx 2N\,\mathcal{N}_{\mathrm{A}} \quad \text{for} \quad \mathrm{SNR} \gg 1 \quad \text{and} \quad \mathcal{N}_{\mathrm{W}} \gg \mathcal{N}_{\mathrm{A}} \quad \text{for} \quad \mathrm{SNR} \ll 1.$$

The result above shows that, for low and high SNR, the GAF needs less samples than the GWD to ensure an accuracy estimation, in particular for high SNR it needs $2N$ time less samples than the GWD.

5.5.3.2 Case of Multiplicative and Additive Noises

Assume now in the model (5.5.20) that b_n is a real multiplicative white noise independent of w_n. One takes $\tau_N = N_e/N$ and consider the particular version of the GAF denoted $\mathcal{P}_N$. If the multiplicative noise is zero mean, $\mathcal{A}_N$ is not able to estimate a_N. In order to solve this problem, it is possible to apply $\mathcal{A}_N$ to the signal square y_n^2 [6,9]. In this case of non-zero mean multiplicative noise, one can establish the following [6] :

$$E\left[\mathcal{P}_N(y,\theta,\tau_N)\right] = \left(\prod_{k=0}^{N-1} m_{c_k}\right) \mathcal{P}_N(e^{j\phi_n},\theta,\tau_N) \tag{5.5.32}$$

where $m_p = E[b_n^p] \;\; \forall p \in \mathbb{N}$ and for $m = N-1, ..., 1$

$$\arg\max_{\theta} \left|E\left[\mathcal{P}_m(y_{n,m},\theta,\tau_m)\right]\right| = m! a_m (\tau_m \Delta)^{m-1} \tag{5.5.33}$$

$$\arg\left\{E\left[y_{n,0}\right]\right\} = a_0 \tag{5.5.34}$$

where $y_{n,m} = y_n e^{-j\sum_{i=m+1}^{N} a_i (n\Delta)^i}$ and $\tau_m = N_e/m$. These recursions allows us to determine the coefficients $\{a_i\}_{i=0}^{N}$. Denoting by $\mathrm{Re}\{x\}$ the real part of x, the mean m_b of b_n can be estimated as follows

$$\widehat{m}_b = \frac{1}{N_e} \mathrm{Re}\left\{\sum_{n=1}^{N_e} y_n e^{-j\sum_{i=0}^{N} \widehat{a}_i (n\Delta)^i}\right\} \tag{5.5.35}$$

and the variance by

$$\mathrm{var}\left\{\mathcal{P}_N(y,\theta,\tau_N)\right\} = \frac{N_e}{N}\left(\prod_{k=0}^{N-1} m_{c_k}^2\right) \mathcal{H}(\sigma_w) \tag{5.5.36}$$

where

$$\mathcal{H}(\sigma_w) = \prod_{k=0}^{N-1}\left[\sum_{i=0}^{c_k} C_{c_k}^{i^2} \frac{m_{2c_k-2i}}{m_{c_k}^2} i! \sigma_w^{2i}\right] - 1. \tag{5.5.37}$$

The variance of $\mathcal{P}_N$ depends on the degree N of the phase, the number N_e of samples, the moments of b_n and σ_w^2. In order to ensure that the estimation of a_N will occur in the exact frequency bin, we should have

$$\frac{\left(\prod_{k=0}^{N-1} m_{c_k}^2\right) |\mathcal{P}_N(e^{j\phi_n}, N!a_N(\tau_N\Delta)^{N-1}, \tau_N)|^2}{\mathrm{var}\{\mathcal{P}_N(y,\theta,\tau_N)\}} \gg 1. \tag{5.5.38}$$

In spectral analysis, it is enough to have this ratio greater than 25 and this leads to the following condition

$$N_e \geq 25\, N\, \mathcal{H}(\sigma_w) \tag{5.5.39}$$

which gives the lower bound for the number of samples to use, in order to ensure an accurate estimation of a_N at given values of degree N, variance σ_w^2 and the moments of the noise b_n.

5.5.4 Summary and Conclusions

Decompositions of polynomial derivatives are used to design time-frequency distributions. The first one, referred to as the GAF, is very close to the the PPT. The second one, referred to as the GWD, appears as a generalization of the PWVD. Applications to PPSs affected by multiplicative and additive noise are proposed.

Additive noise is further considered in Article 10.4. A more intuitive introduction to PWVDs is given in Article 5.4.

5.5.5 Appendix

We consider $Q+1$ arbitrary real and distinct parameters : $t_0,\ldots,t_Q$, and denote by $\alpha_0^\ell,\ldots,\alpha_Q^\ell$ one solution of the following Vandermonde system :

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ t_0 & t_1 & \ldots & t_Q \\ \vdots & \vdots & & \vdots \\ t_0^\ell & t_1^\ell & \ldots & t_Q^\ell \\ \vdots & \vdots & & \vdots \\ t_0^N & t_1^N & \ldots & t_Q^N \end{pmatrix} \begin{pmatrix} \alpha_0^\ell \\ \alpha_1^\ell \\ \vdots \\ \alpha_\ell^\ell \\ \vdots \\ \alpha_Q^\ell \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ (-1)^\ell \ell! \\ \vdots \\ 0 \end{pmatrix}. \tag{5.5.40}$$

The existence of the solution of the system above is discussed in the the general case in [4]. For a given k $(k = 0,\ldots,Q)$, we denote by $\sigma_Q^p(t_{l\neq k})$, $(p = 0,\ldots,Q)$, the classical symmetric polynomials of the Q parameters : $t_0,\ldots,t_{k-1},t_{k+1},\ldots,t_Q$ defined by $\sigma_Q^0(t_{l\neq k}) \stackrel{\Delta}{=} 1$ and

$$\sigma_Q^p(t_{l\neq k}) \stackrel{\Delta}{=} \sum_{\{t_0,\ldots,t_{k-1},t_{k+1},\ldots,t_Q\}} t_{i_1}\ldots t_{i_p}, \quad (p = 1,\ldots,Q)$$

If $N \leq Q+1$, the system above has a solution given by

$$\alpha_k^\ell = \ell! \frac{\sigma_Q^{Q-\ell}(t_{l\neq k})}{\prod_{i\neq k}(t_i - t_k)}, \quad (k = 0, \ldots, Q) \tag{5.5.41}$$

assuming that the t_k's satisfy condition (5.5.4) in the case $\ell < N$ and $N = Q+1$.

References

[1] R. Kumaresan and S. Verma, "On estimating the parameters of chirp signals using rank reduction techniques," in *Proc. 21st Asilomar Conf. on Signals, Systems, and Computers*, pp. 555–558, Pacific Grove, CA, 2–4 November 1987.

[2] A. H. Nuttall, "Efficient evaluation of polynomials and exponentials of polynomials for equispaced arguments," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 35, pp. 1486–1487, October 1987.

[3] S. Peleg and B. Porat, "Estimation and classification of polynomial-phase signals," *IEEE Trans. Information Theory*, vol. 37, pp. 422–430, March 1991.

[4] M. Benidir, "Characterization of polynomial functions and application to time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 45, pp. 1351–1354, May 1997.

[5] M. Benidir and A. Ouldali, "Polynomial phase signal analysis based on the polynomial derivatives decompositions," *IEEE Trans. Signal Processing*, vol. 47, pp. 1954–1965, July 1999.

[6] A. Ouldali and M. Benidir, "Statistical analysis of polynomial phase signals affected by multiplicative and additive noise," *Signal Processing*, vol. 78, pp. 19–42, October 1999.

[7] S. Peleg and B. Friedlander, "Discrete polynomial-phase transform," *IEEE Trans. Signal Processing*, vol. 43, pp. 1901–1914, August 1995.

[8] B. Boashash and P. J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 216–220, January 1994.

[9] B. Boashash and B. Ristic, "Analysis of FM signals affected by Gaussian AM using the reduced Wigner-Ville trispectrum," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'93)*, vol. IV, pp. 408–411, Minneapolis, 27–30 April 1993.

[10] B. Boashash and B. Ristic, "A time-frequency perspective of higher-order spectra as a tool for non-stationary signal analysis," in *Higher-Order Statistical Signal Processing* (B. Boashash, E. J. Powers, and A. M. Zoubir, eds.), ch. 4, pp. 111–149, Melbourne/N.Y.: Longman/Wiley, 1995.

5.6 TIME-FREQUENCY REPRESENTATIONS COVARIANT TO GROUP DELAY SHIFTS [0]

5.6.1 Group Delay Shift Covariance Property

Quadratic time-frequency representations (QTFRs) [1–5] are potentially capable of displaying the temporal localization of the spectral components of a nonstationary signal (see Chapter 3). However, no single 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 time-frequency (TF) structures. In order to assist the user in selecting the appropriate analysis tool, QTFRs are often classified based on the various properties they satisfy including *covariance* properties. A QTFR is said to satisfy a covariance property if the QTFR preserves, or is covariant to, certain TF changes on the signal. For example, for a signal $x(t)$ with Fourier transform $X(f)$, Cohen's-class QTFRs $T_X^{(C)}(t,f)$ (with signal-independent kernels; i.e. the quadratic class as defined on p. 68) are covariant to constant time shifts [2]. Specifically,

$$Y(f) = (\mathcal{Y}_\tau X)(f) = e^{-j2\pi\tau f} X(f) \quad \Rightarrow \quad T_Y^{(C)}(t,f) = T_X^{(C)}(t-\tau, f), \qquad (5.6.1)$$

where $\mathcal{Y}_\tau$ is the constant time shift operator.[1] Constant time shifts are important, for example, 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. Cohen's class QTFRs also preserve constant frequency shifts $T_Y^{(C)}(t,f) = T_X^{(C)}(t,f-\nu)$ when $Y(f) = (\mathcal{Q}_\nu X)(f) = X(f-\nu)$. On the other hand, affine class QTFRs $T_X^{(A)}(t,f)$ preserve constant time shifts as in (5.6.1) and scale changes (dilations) on the signal, i.e., $T_Y^{(A)}(t,f) = T_X^{(A)}(at, f/a)$ when $Y(f) = (\mathcal{C}_a X)(f) = X(f/a)/\sqrt{|a|}$ (see [3,6] and Article 7.1).

In some applications, it is important to preserve signal time shifts caused by the signal propagating through systems with dispersive TF characteristics or, equivalently, with nonlinear group delay[2] . A dispersive system is one which delays, in time, different frequencies by different amounts. For example, in underwater acoustics, backscattering from immersed targets such as spherical shells may result in dispersive waves [3], thus leading to echoes with frequency-dependent group delay. If a QTFR preserves times shifts by an amount equal to the change in group delay in a signal or system, then the corresponding QTFR property is referred to as *group*

---

[0] Author: **Antonia Papandreou-Suppappola**, Telecommunications Research Center, Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-7206 USA (papandreou@asu.edu). Reviewers: P. Flandrin and D. L. Jones.

[1] Table 5.6.1 summarizes various operators used in this article and their effect on signals.

[2] The group delay of a signal $x(t)$, with Fourier transform $X(f) = a(f)\, e^{-j2\pi\,\vartheta(f/f_r)}$, is the derivative of the phase modulation $\vartheta(f/f_r)$ of the signal, i.e., $r(f) = \frac{d}{df}\vartheta(f/f_r)$. Here, $a(f) \geq 0$ is the signal's amplitude modulation, and $f_r > 0$ is a reference frequency.

Table 5.6.1: Summary of some commonly used operators in this article together with their effect on a signal $x(t)$. Some of these operators depend on a differentiable one-to-one function $\Lambda(b)$. Here, $\Lambda'(b) = \frac{d}{db}\Lambda(b)$, $\Lambda^{-1}(\Lambda(b)) = b$, $\mathrm{sgn}(b)$ yields the sign (± 1) of b, and $f_r > 0$ is a reference frequency. Also, $\wp$ and $\aleph$ are the domain and range, respectively, of $\Lambda(\cdot)$.

Operator Name	Operator	Effect of Operator
Identity	$\mathcal{I}$	$X(f), \quad f \in \Re$
Scaling	$\mathcal{C}_a$	$\frac{1}{\sqrt{\lvert a\rvert}} X(\frac{f}{a})$
Constant frequency shift	$\mathcal{Q}_\nu$	$X(f-\nu)$
Constant time shift	$\mathcal{Y}_\tau$	$e^{-j2\pi\tau f} X(f)$
Hyperbolic time shift	$\mathcal{D}_c^{(\Lambda)}$, $\Lambda(b) = \ln b$	$e^{-j2\pi c \ln(f/f_r)} X(f), \quad f \in \Re^+$
Power time shift	$\mathcal{D}_c^{(\Lambda)}$, $\Lambda(b) = \mathrm{sgn}(b)\lvert b\rvert^\kappa$	$e^{-j2\pi c\, \mathrm{sgn}(f) \lvert \frac{f}{f_r}\rvert^\kappa} X(f), \quad f \in \Re$
Exponential time shift	$\mathcal{D}_c^{(\Lambda)}$, $\Lambda(b) = e^{\kappa b}$	$e^{-j2\pi c\, e^{\kappa f/f_r}} X(f), \quad f \in \Re$
Group delay shift	$\mathcal{D}_c^{(\Lambda)}$	$e^{-j2\pi c\, \Lambda(\frac{f}{f_r})} X(f), \quad f \in \wp$
Hyperbolic warping	$\mathcal{U}_\Lambda$, $\Lambda(b) = \ln b$	$e^{\frac{f}{2f_r}} X(f_r e^{f/f_r}), \quad f \in \Re$
Power warping	$\mathcal{U}_\Lambda$, $\Lambda(b) = \mathrm{sgn}(b)\lvert b\rvert^\kappa$	$\frac{1}{\sqrt{\lvert\kappa\rvert}} \lvert f/f_r\rvert^{\frac{1-\kappa}{2\kappa}} X\Big(f_r\, \mathrm{sgn}(f) \lvert\frac{f}{f_r}\rvert^{\frac{1}{\kappa}}\Big), \quad f \in \Re$
Exponential warping	$\mathcal{U}_\Lambda$, $\Lambda(b) = e^{\kappa b}$	$\sqrt{\frac{f_r}{\lvert\kappa\rvert f}}\, X\Big(\frac{f_r}{\kappa} \ln \frac{f}{f_r}\Big), \quad f \in \Re^+$
Dispersive warping	$\mathcal{U}_\Lambda$	$\Big\lvert \Lambda'\Big(\Lambda^{-1}\Big(\frac{f}{f_r}\Big)\Big)\Big\rvert^{-\frac{1}{2}} \cdot X\Big(f_r \Lambda^{-1}\Big(\frac{f}{f_r}\Big)\Big), \quad f \in \aleph$

delay shift (GDS) *covariance* (also called generalized time shift in [7]). If $X(f)$ is passed through an allpass dispersive system with output $Y(f) = e^{-j2\pi\Lambda(f/f_r)} X(f)$, the change in group delay, $\tau(f) = \frac{d}{df}\Lambda(f/f_r)$, is proportional to the derivative of the phase function $\Lambda(f/f_r)$. Since group delay is a measure of the time delay introduced in each sinusoidal component of the signal at frequency f, the ideal QTFR $T_Y(t,f)$ should preserve this frequency-dependent, group delay change $\tau(f)$ [7,8].

Property Definition

The effect of the GDS operator[3] on a signal $x(t)$ is given by

$$Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c\, \Lambda(\frac{f}{f_r})} X(f), \quad f \in \wp \tag{5.6.2}$$

[3]One could also define the *dual* to this operator, i.e., $(\mathcal{D}_c^{(\xi)} x)(t) = e^{j2\pi c\, \xi(t/t_r)} x(t)$, $t_r > 0$, to represent time-varying instantaneous frequency shifts (see Article 9.2).

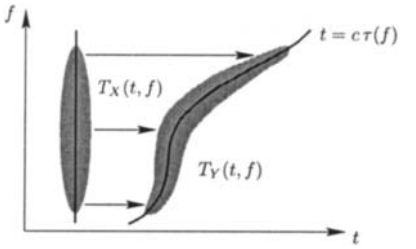


Fig. 5.6.1: A QTFR preserves a dispersive change in a signal's group delay $c\tau(f)$ as the QTFR is shifted along the time axis by an amount equal to $c\tau(f)$. Here, $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f)$ as in (5.6.2).

where $\wp$ is the domain of $\Lambda(\cdot)$ and $c \in \Re$ (see [7–9] and Article 15.3). A QTFR T is GDS covariant if the QTFR of the output, T_Y, corresponds to the QTFR of the input, T_X, shifted in time by an amount equal to the change in group delay, $c\tau(f) = c\frac{d}{df}\Lambda(\frac{f}{f_r})$, that is introduced in (5.6.2). For $f \in \wp$, the property states

$$Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c\,\Lambda(\frac{f}{f_r})} X(f) \;\Rightarrow\; T_Y(t,f) = T_X\big(t - c\tau(f), f\big) \quad (5.6.3)$$

when $\Lambda(b)$ is a differentiable one-to-one function. Fig. 5.6.1 demonstrates the QTFR transformation in (5.6.3) due to the signal transformation in (5.6.2) that results in a nonlinear change in group delay. The parameter c expresses the amount of dispersion or nonlinear time modulation on the signal as shown in Fig. 5.6.2. The GDS operator $\mathcal{D}_c^{(\Lambda)}$ in (5.6.2) is unitarily equivalent to the constant time shift operator $\mathcal{Y}_{c/f_r}$ in (5.6.1) [see [7–10] and Articles 4.5 & 15.3],

$$\mathcal{D}_c^{(\Lambda)} = \mathcal{U}_\Lambda^{-1} \mathcal{Y}_{c/f_r} \mathcal{U}_\Lambda \quad (5.6.4)$$

where $(\mathcal{U}_\Lambda^{-1}(\mathcal{U}_\Lambda X))(f) = X(f)$, and the dispersive warping operator is given by [8]

$$Z(f) = (\mathcal{U}_\Lambda X)(f) = \big|\Lambda'\big(\Lambda^{-1}(f/f_r)\big)\big|^{-\frac{1}{2}}\; X\big(f_r \Lambda^{-1}(f/f_r)\big)\;,\; f \in \aleph. \quad (5.6.5)$$

Here, $\aleph$ denotes the range of $\Lambda(\cdot)$. The unitary warping operator [8] in (5.6.5) preserves inner products, i.e., $\int_{f\in\aleph} (\mathcal{U}_\Lambda X)(f)\,(\mathcal{U}_\Lambda X)^*(f)\,df = \int_{f\in\wp} X(f)\,X^*(f)\,df$. Note that since the GDS covariance in (5.6.3) follows directly from the constant time shift covariance in (5.6.1) via (5.6.4), it experiences the same problems as the constant time shift covariance when the time origin is not known. An example of such a case is the estimation of the range of a target in an active sonar application.

Specific Examples of the GDS Covariance Property

The GDS covariance property in (5.6.3) simplifies to a *particular* covariance property (satisfied by a different class of QTFRs) when the differentiable one-to-one function $\Lambda(b)$ and the time shift $\tau(f) = \frac{d}{df}\Lambda(\frac{f}{f_r})$ are fixed. Depending on the choice of $\Lambda(b)$, the time shift $\tau(f)$ may be constant, linear, or nonlinear (dispersive). For example,

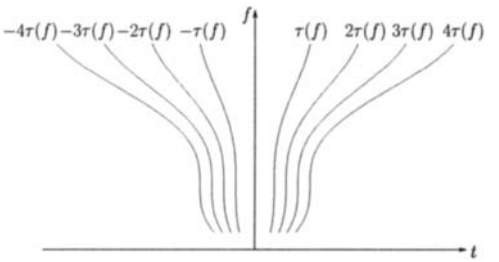


Fig. 5.6.2: Various group delay curves $t = c\tau(f)$ corresponding to positive and negative changes in the parameter c in Equation (5.6.2).

- for $\Lambda(b) = b$ (with $\wp = \aleph = \Re$), the GDS covariance in (5.6.3) simplifies to the constant time shift, $\tau(f) = \frac{1}{f_r}$, covariance in (5.6.1) of Cohen's class or the affine class [4].
- for $\Lambda(b) = \ln b$, $b > 0$ (with $\wp = \Re_+$ and $\aleph = \Re$), the covariance in (5.6.3) simplifies to the (dispersive) hyperbolic time shift, $\tau(f) = \frac{1}{f}$, $f > 0$, covariance of the hyperbolic class [11].
- for $\Lambda(b) = \text{sgn}(b)\,|b|^\kappa$, $\kappa \neq 0$ (with $\wp = \aleph = \Re$), the covariance in (5.6.3) simplifies to the (possibly dispersive) κth power time shift, $\tau(f) = \frac{d}{df}\Lambda(\frac{f}{f_r}) = \frac{\kappa}{f_r}|\frac{f}{f_r}|^{\kappa-1}$, covariance of the κth power class (see [9] and Article 15.3). The GDS is constant for $\kappa = 1$, and linearly dependent on frequency for $\kappa = 2$.
- for $\Lambda(b) = e^{\kappa b}$ (with $\wp = \Re$ and $\aleph = \Re_+$), the covariance in (5.6.3) simplifies to the (dispersive) κth exponential time shift, $\tau(f) = \frac{\kappa}{f_r}e^{\kappa f/f_r}$, covariance of the κth exponential class [7,8].

Matched Signal Analysis

For successful TF analysis, it is advantageous to match the GDS of a QTFR in (5.6.3) with the signal's group delay. Thus, QTFRs that satisfy (5.6.3) are ideally suited to analyze signals with group delay equal to that GDS. The TF geometry underlying the GDS covariance property is related to the *generalized impulse* function defined as

$$I_c^{(\vartheta)}(f) = \sqrt{|r(f)|}\, e^{-j2\pi c\,\vartheta(\frac{f}{f_r})}, \quad f \in \wp, \tag{5.6.6}$$

where $\wp$ is the domain of the phase function $\vartheta(\frac{f}{f_r})$. The group delay $r(f) = \frac{d}{df}\vartheta(\frac{f}{f_r})$ reflects the dispersion characteristics of the class which is covariant to the GDS $\tau(f)$ in (5.6.3) only when $r(f) = \tau(f)$ (or, equivalently, when $\vartheta(b) = \Lambda(b)$ in (5.6.3)).

5.6.2 Classes of GDS Covariant QTFRs

As shown in (5.6.4), the GDS operator in (5.6.2) can be acquired by unitarily warping the time shift operator in (5.6.1) [8, 10]. Thus, GDS covariant QTFRs

can be obtained by appropriately warping dispersively constant time shift covariant QTFRs, such as Cohen's class or affine class QTFRs, using

$$T_X^{(D\,class)}(t,f) \;=\; T_Z^{(class)}\Big(\frac{t}{f_r\tau(f)}\,,\, f_r\,\Lambda\Big(\frac{f}{f_r}\Big)\Big)\,,\quad f\in\wp\,, \tag{5.6.7}$$

where $\wp$ is the domain of $\Lambda(\frac{f}{f_r})$ and $Z(f) = (\mathcal{U}_\Lambda X)(f)$ is the dispersively warped signal in (5.6.5). The superscript *class* indicates which QTFR class undergoes the warping in (5.6.7), and *D class* indicates the resulting QTFR class with possible dispersive characteristics. For example, with *class* = *C*, $T^{(C)}$ is a Cohen's class QTFR that can be warped to yield a new QTFR $T^{(DC)}$ (see below). A more specific example follows. Consider the Wigner distribution (WD) [2–4], defined as $W_Z(t,f) = \int_{-\infty}^{\infty} Z(f+\frac{\nu}{2})Z^*(f-\frac{\nu}{2})\,e^{j2\pi t\nu}\,d\nu$ [see Subsection 2.1.4.2], which is a member of Cohen's class. By warping the WD as in (5.6.7), one obtains the GDS covariant version of the WD or frequency dispersively warped WD (DWD) [7,8]

$$W_X^{(\Lambda)}(t,f) = W_Z\Big(\frac{t}{f_r\tau(f)}\,,\, f_r\,\Lambda\Big(\frac{f}{f_r}\Big)\Big)\,,\quad f\in\wp \tag{5.6.8}$$

which is a specific QTFR example of $T^{(DC)}$. For a fixed $\Lambda(b)$, this transformation can be achieved in three steps. First, the signal is warped as $Z(f) = (\mathcal{U}_\Lambda X)(f)$, and then the WD of the warped signal $Z(f)$ is computed. Lastly, the TF axes are transformed for correct TF localization using $t \to t/(f_r\tau(f))$ and $f \to f_r\Lambda(f/f_r)$. Note that this method ensures that the GDS covariant QTFRs, including the DWD, always satisfy (5.6.3) for a given differentiable one-to-one function $\Lambda(b)$.

GDS covariant QTFRs from Cohen's class

The GDS covariant Cohen's class QTFRs, $T_X^{(DC)}(t,f)$, are obtained by dispersively warping Cohen's QTFRs, $T_X^{(C)}(t,f)$, using (5.6.7) with *class* = *C* and a warping function $\Lambda(b)$ chosen to give the desired GDS covariance in (5.6.3). Note that Cohen's class QTFRs are given by [2]

$$T_X^{(C)}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \gamma_T(t-\hat{t},f-\hat{f})\,W_X(\hat{t},\hat{f})\,d\hat{t}\,d\hat{f}$$

where $W_X(t,f)$ is the WD. The kernel $\gamma_T(t,f)$ uniquely characterizes the Cohen's class QTFR $T^{(C)}$ as well as its corresponding warped version $T^{(DC)}$ in (5.6.7) [7,8]. Due to the warping in (5.6.5), the constant time shift and frequency shift covariances that all Cohen's class QTFRs satisfy are transformed into the GDS and warped frequency shift covariances, respectively, that all DC QTFRs satisfy [7, 8]. An important member of the DC class that satisfies many desirable properties is the DWD, $W_X^{(\Lambda)}(t,f)$, in (5.6.8). For example, it satisfies a specific set of marginal properties, and provides a highly concentrated representation for the generalized impulse $I_c^{(\vartheta)}(f)$ provided $\vartheta(b) = \Lambda(b)$ in (5.6.6). Specifically, for $X(f) = I_c^{(\vartheta)}(f)$, $W_X^{(\Lambda)}(t,f) = |\tau(f)|\,\delta(t-c\tau(f))$. Other QTFR members and their properties can be found in [7,8].

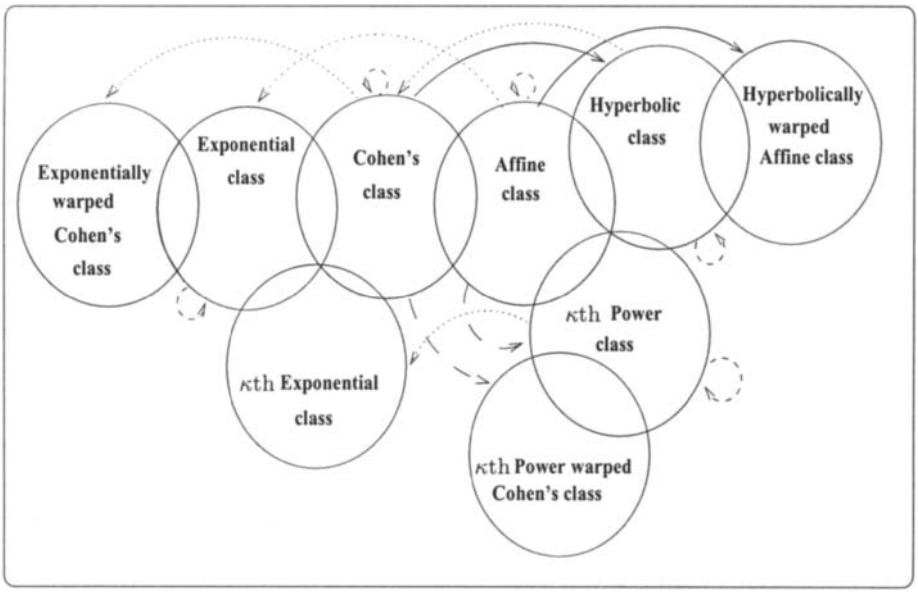


Fig. 5.6.3: Some GDS covariant QTFR classes. An arrow points from a QTFR class being warped to a new QTFR class. Different types of warping are demonstrated depending on $\Lambda(b)$ in (5.6.7). They include hyperbolic warping with $\Lambda(b) = \ln b$ (____), power warping with $\Lambda(b) = \text{sgn}(b)\,|b|^{\kappa}$ (— — —), or exponential warping with $\Lambda(b) = e^{\kappa b}$ (. . .). The warping (- - -) maps a QTFR class back to itself ("self-mapping"). For example, the hyperbolic class in (5.6.9) is obtained by warping Cohen's class using the hyperbolic warping with $\Lambda(b) = \ln b$ (____) in (5.6.7).

Different DC classes can be obtained, suitable in different applications, simply by choosing a differentiable one-to-one function $\Lambda(b)$ to match a signal's TF characteristics. Once $\Lambda(b)$ is fixed, it can be replaced in Equation (5.6.7) to obtain the corresponding formulation of the specified class; in Equation (5.6.8) to obtain the DWD; and in Equation (5.6.3) to obtain a fixed GDS (possibly dispersive) covariance property. Some class examples obtained by choosing the function $\Lambda(b)$ are summarized below, and are also pictorially demonstrated in Fig. 5.6.3.

Cohen's class: When $\Lambda(b) = b$ and $\tau(f) = 1/f_r$, the warping in (5.6.7) with $\wp = \Re$ maps Cohen's class back to itself since $(\mathcal{U}_\Lambda X)(f) = (\mathcal{I}X)(f) = X(f)$ in (5.6.5) becomes the (non-dispersive) identity operator. The GDS covariance property in (5.6.3) simplifies to the constant time shift covariance in (5.6.1). Some members of Cohen's class include the WD and the spectrogram [2].

Hyperbolic class: When $\Lambda(b) = \ln b$ and $\tau(f) = 1/f$, $f > 0$, the DC class in (5.6.7) (with $\wp = \Re_+$ and $class = C$) is the hyperbolic class [11]. Thus, any hyperbolic class QTFR, $T_X^{(H)}(t,f)$, $f > 0$, can be obtained by warping a corresponding Cohen's class QTFR, $T_X^{(C)}(t,f)$, using the transformation

$$T_X^{(H)}(t,f) = T_Z^{(C)}\left(\frac{tf}{f_r},\, f_r \ln\frac{f}{f_r}\right) \tag{5.6.9}$$

where $Z(f) = (\mathcal{U}_\Lambda X)(f) = e^{\frac{f}{2f_r}}\, X(f_r\, e^{\frac{f}{f_r}})$ in (5.6.5). The GDS covariance in (5.6.3) simplifies to the hyperbolic time shift covariance $T_Y^{(H)}(t,f) = T_X^{(H)}(t - c/f, f)$ where $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c \ln \frac{f}{f_r}}\, X(f)$. The warped frequency shift covariance simplifies to scale covariance [7,8]. These two covariance properties defining the hyperbolic class are important for the analysis of Doppler-invariant signals similar to the signals used by bats for echolocation, and for the analysis of self-similar random processes. Members of the hyperbolic class include the Altes Q-distribution [12] (the DWD in (5.6.8) with $\Lambda(b) = \ln b$), the unitary Bertrand P_0-distribution ([6] and Article 7.1), and the hyperbologram [11].

Other class examples include the power warped Cohen's class [7,8] when $\Lambda(b) = \operatorname{sgn}(b)\,|b|^\kappa$ in (5.6.7), and the exponentially warped Cohen's class [7,8] when $\Lambda(b) = e^b$ in (5.6.7) (see Fig. 5.6.3).

GDS covariant QTFRs from the affine class

The GDS covariant affine class QTFRs, $T_X^{(DA)}(t,f)$, are obtained by warping the affine class QTFRs $T_X^{(A)}(t,f)$ using (5.6.7) with *class* $= A$. The affine class QTFRs are given by (see [3,6] and Article 7.1)

$$T_X^{(A)}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_T\Big(f(t-\hat{t}), -\hat{f}/f\Big)\, W_X(\hat{t},\hat{f})\, d\hat{t}\, d\hat{f}$$

where $W_X(t,f)$ is the WD and the kernel $\psi_T(a,b)$ uniquely characterizes $T^{(A)}$ and its corresponding warped version $T^{(DA)}$. The warping transforms the two covariance properties defining the affine class, the constant time shift covariance and scale covariance, into two new covariance properties. These properties, which define the DA class, are the GDS covariance in (5.6.3), and the warped scale covariance [7,8]. The latter property simplifies to known covariance properties based on $\Lambda(b)$ as shown below. The DWD in (5.6.8) is a member of both the DC and the DA classes since the WD is a member of both Cohen's class and the affine class. Other QTFR members and their desirable properties can be found in [7,8].

Various QTFR classes useful in different applications are obtained by fixing $\Lambda(b)$ in (5.6.7). Some examples are listed below and are also demonstrated in Fig. 5.6.3.

Affine class: The affine class is an example of a "self-mapping" since, when $\Lambda(b) = b$ and $\tau(f) = 1/f_r$, the warping in (5.6.7) maps the affine class back to itself. Some important affine QTFRs include the WD, the Bertrand P_κ-distributions ([6] and Article 7.1), and the scalogram [3,4].

κth Power class: When $\Lambda(b) = \operatorname{sgn}(b)\,|b|^\kappa$ and $\tau(f) = \frac{\kappa}{f_r}|\frac{f}{f_r}|^{\kappa-1}$, $\kappa \neq 0$, the DA in (5.6.7) (with $\wp = \Re$ and *class* $= A$) is the κth power class (see [7,9] and Article 15.3). Different power classes are obtained by varying κ, and the affine class is obtained when $\kappa = 1$. The κth power QTFRs, $T^{(\kappa P)}$, are obtained by warping corresponding members of the affine class (cf. (5.6.7))

$$T_X^{(\kappa P)}(t,f) = T_Z^{(A)}\Big((t/\kappa)\,|f/f_r|^{1-\kappa}\,,\; f_r \operatorname{sgn}(f)\,|f/f_r|^\kappa\Big)$$

where $Z(f) = (\mathcal{U}_\Lambda X)(f) = |\kappa|^{-\frac{1}{2}} |f/f_r|^{\frac{1-\kappa}{2\kappa}} X\Big(f_r \operatorname{sgn}(f) |f/f_r|^{\frac{1}{\kappa}}\Big)$ in Equation (5.6.5). The GDS covariance in (5.6.3) simplifies to the κth power time shift covariance property $T_Y^{(\kappa P)}(t,f) = T_X^{(\kappa P)}(t - c\frac{\kappa}{f_r}|\frac{f}{f_r}|^{\kappa-1}, f)$ where $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c \operatorname{sgn}(f) |\frac{f}{f_r}|^\kappa} X(f)$. The warped scale covariance simplifies to scale covariance [7,8]. These two properties defining the κth power class are important for analyzing signals propagating through systems with power-law dispersive TF characteristics. Members of the κth power class include the power WD, the Bertrand P_κ-distributions ([6] and Article 7.1), and the powergram ([9] and Article 15.3).

κth Exponential class: The κth exponential class QTFRs, $T^{(\kappa E)}$, [7,8] are obtained when $\Lambda(b) = e^{\kappa b}$ and $\tau(f) = \frac{\kappa}{f_r} e^{\kappa f/f_r}$, $\kappa \neq 0$, in the DA formulation in (5.6.7) (with $\wp = \Re$ and *class* $= A$). Any such QTFRs can be written as warped affine QTFRs, $T_X^{(\kappa E)}(t,f) = T_Z^{(A)}\Big(\frac{1}{\kappa} t\, e^{-\kappa f/f_r},\, f_r\, e^{\kappa f/f_r}\Big)$, where $Z(f) = (\mathcal{U}_\Lambda X)(f) = \sqrt{\frac{f_r}{|\kappa| f}}\, X(\frac{f_r}{\kappa} \ln \frac{f}{f_r})$ in (5.6.5). These QTFRs satisfy the κth exponential GDS covariance given as $T_Y^{(\kappa E)}(t,f) = T_X^{(\kappa E)}\Big(t - c\frac{\kappa}{f_r} e^{\kappa f/f_r}, f\Big)$ where $Y(f) = (\mathcal{D}_c^{(\Lambda)} X)(f) = e^{-j2\pi c\, e^{\kappa f/f_r}} X(f)$. This is an important property for analyzing signals passing through exponentially dispersive systems. The warped scale covariance simplifies to constant frequency shift covariance [7]. Some important class members include the κth exponential WD (the DWD in (5.6.8) with $\Lambda(b) = e^{\kappa b}$), and the κth Cohen-Bertrand P_0-distribution [7]. When $\kappa = 1$, the κth exponential class yields the exponential class [7,8] in Fig. 5.6.3.

Another DA class example is the hyperbolically warped affine class [7,8] obtained when $\Lambda(b) = \ln b$ in (5.6.7) (see Fig. 5.6.3).

5.6.3 Simulation Example

When using QTFRs to analyze signals in real life applications, some pre-processing or a priori knowledge is necessary to identify the TF characteristics of the signal so as to match them to the GDS covariance of the QTFR. For example, signals with linear TF characteristics, such as Dirac impulses $X(f) = e^{-j2\pi c f/f_r}/\sqrt{f_r}$ or dolphin clicks, are best analyzed using Cohen's or affine class QTFRs as these QTFRs preserve constant time shifts $\tau(f) = 1/f_r$ that equal the constant group delay $r(f) = 1/f_r$ of the signal. Signals with hyperbolic TF characteristics, such as hyperbolic impulses (the signal in (5.6.6) with $\vartheta(b) = \ln b$ and $r(f) = 1/f$) or Doppler-invariant signals such as the signals used by bats for echolocation, are best analyzed using hyperbolic QTFRs since they preserve hyperbolic time shifts $\tau(f) = 1/f$. Fig. 5.6.4 analyzes the sum of two windowed hyperbolic impulses given (before windowing) as $X(f) = \frac{1}{\sqrt{f}}\Big(e^{-j2\pi 3 \ln \frac{f}{f_r}} + e^{-j2\pi 7 \ln \frac{f}{f_r}}\Big)$, $f > 0$. As expected, the hyperbolic Altes Q-distribution [12] (the DWD in (5.6.8) with $\Lambda(b) = \ln b$) in Fig. 5.6.4(c) results in high TF concentration along the two hyperbolae $t = 3/f$ and $t = 7/f$. However, it also results in cross terms (CTs) along the mean hyperbola

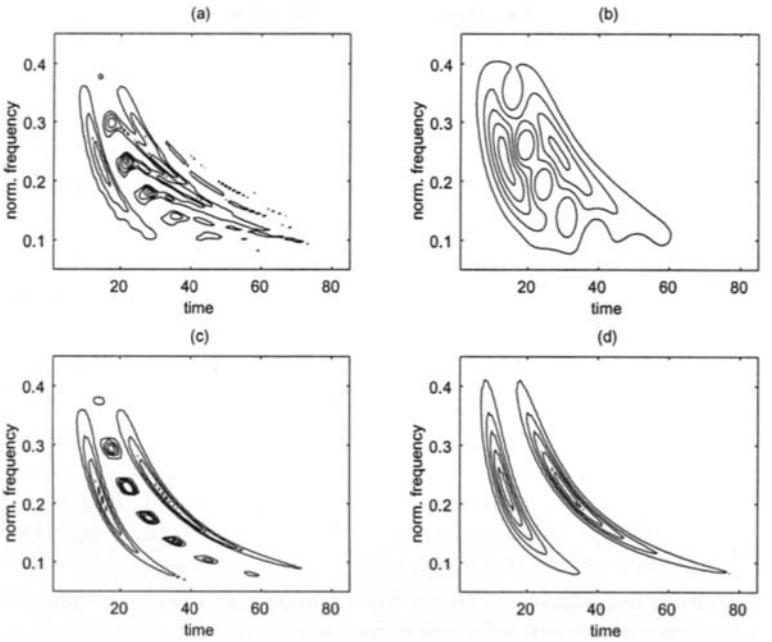


Fig. 5.6.4: TF analysis of the sum of two windowed hyperbolic impulses. The first row shows QTFRs from Cohen's class, and the second row shows QTFRs from the hyperbolic class in (5.6.9). (a) WD, (b) smoothed pseudo WD, (c) Altes Q-distribution (QD), (d) smoothed pseudo QD.

$t = (3+7)/(2f)$ [11]. The smoothed pseudo Altes Q-distribution (the hyperbolically warped version of the smoothed pseudo WD) in Fig. 5.6.4(d) removes the CTs with some loss of TF resolution. Cohen's class QTFRs, such as the WD in Fig. 5.6.4(a) and the smoothed pseudo WD in Fig. 5.6.4(b), are not well-matched to hyperbolic impulses. The WD results in complicated CTs between the two signal components as well as inner interference terms [4]. In comparison to the smoothed pseudo Altes Q-distribution in Fig. 5.6.4(d), the smoothed pseudo WD in Fig. 5.6.4(b) has a larger loss of TF resolution, and it is not as successful at removing all the CTs. Other examples, including real data simulations, can be found in [8,9] and Article 15.3.

5.6.4 Summary and Conclusions

This article emphasizes the importance of matching a QTFR to the analysis signal for successful processing. Specifically, it presents new classes of QTFRs that are ideal when the dispersive group delay or changes in group delay of a signal match the frequency-varying time shift covariance property of the QTFRs. These QTFRs are important in analyzing signals propagating through systems with dispersive, nonlinear characteristics including hyperbolic, power or exponential. The QTFR classes considered here are obtained based on covariance properties, and can thus be

considered as special cases of the covariant QTFRs in Article 4.3. The contributions of F. Hlawatsch and G. F. Boudreaux-Bartels to the development of this work may been seen in [9,11] and in Article 15.3.

References

[1] B. Boashash, ed., *Time-Frequency Signal Analysis: Methods and Applications.* Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[2] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[3] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[4] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[5] A. Papandreou-Suppappola, ed., *Applications in Time-Frequency Signal Processing.* Boca Raton, FL: CRC Press, 2002.

[6] J. Bertrand and P. Bertrand, "Affine time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 5, pp. 118–140, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[7] A. Papandreou-Suppappola, "Generalized time-shift covariant quadratic time-frequency representations with arbitrary group delays," in *Proc. 29th Asilomar Conf. on Signals, Systems, and Computers*, pp. 553–557, Pacific Grove, CA, October–November 1995.

[8] A. Papandreou-Suppappola, R. L. Murray, B. G. Iem, and G. F. Boudreaux-Bartels, "Group delay shift covariant quadratic time-frequency representations," *IEEE Trans. Signal Processing*, vol. 49, pp. 2549–2564, November 2001.

[9] 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 Processing*, vol. 47, pp. 3067–3083, November 1999.

[10] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: A new twist on signal processing," *IEEE Trans. Signal Processing*, vol. 43, pp. 2269–2282, October 1995.

[11] 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 Processing*, vol. 41, pp. 3425–3444, December 1993. Special Issue on Wavelets and Signal Processing.

[12] R. A. Altes, "Wideband, proportional-bandwidth Wigner-Ville analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1005–1012, June 1990.

5.7 DESIGN OF HIGH-RESOLUTION QUADRATIC TFDs WITH SEPARABLE KERNELS [0]

5.7.1 RIDs and Quadratic TFDs

Reduced-interference distributions (RIDs) smooth out the unwanted cross-terms in the time-frequency (t, f) domain [see Chapter 3 and Article 5.2]. The spectrogram is the best-known RID, but not the only one; it is well known that a time-frequency distribution (TFD) is a RID if its kernel has a two-dimensional low-pass characteristic in the Doppler-lag (ν, τ) domain [1]. A suitably chosen **separable kernel**, i.e. a kernel with the form $G_1(\nu)\, g_2(\tau)$, can meet this requirement while giving higher time-frequency resolution than the spectrogram; for example, the "smoothed WVD" discussed in [2] has a separable kernel. It has also been shown that if the kernel is *independent of lag* (i.e. a function of Doppler alone in the Doppler-lag domain, or of time alone in the time-lag domain), then the resulting TFD can exhibit fine frequency resolution and high attenuation of cross-terms [3,4].

This article explores the properties of TFDs with separable kernels, including lag-independent kernels, characterizes the signals for which such kernels can be recommended, and gives examples of RID designs using such kernels. It builds on the argument presented in Sections 3.2.1 and 3.2.2 (pp. 66–69).

5.7.2 Separable Kernel Formulations

In the Doppler-lag domain, a **separable kernel** has the form

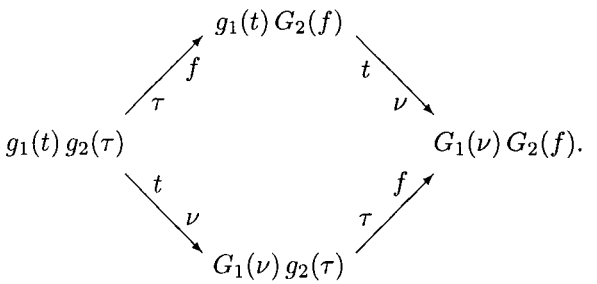
$$g(\nu, \tau) = G_1(\nu)\, g_2(\tau). \tag{5.7.1}$$

If we let

$$G_1(\nu) = \underset{t \to \nu}{\mathcal{F}}\{g_1(t)\} \tag{5.7.2}$$

$$G_2(f) = \underset{\tau \to f}{\mathcal{F}}\{g_2(\tau)\}, \tag{5.7.3}$$

then the relationships shown in the graphical Eq. (3.2.6) become



(5.7.4)

[0]Authors: **Boualem Boashash** and **Gavin R. Putland**, Signal Processing Research Centre, Queensland University of Technology, GPO Box 2434, Brisbane, Q 4001, Australia (b.boashash@qut.edu.au, g.putland@qut.edu.au). Reviewers: Z. M. Hussain and V. Sucic.

Substituting Eq. (5.7.1) into Eq. (3.2.11), we obtain the filtered ambiguity function

$$\mathcal{A}_z(\nu,\tau) = G_1(\nu)\, g_2(\tau)\, A_z(\nu,\tau) \tag{5.7.5}$$

(also called the "generalized ambiguity function" by some authors). Then, using the convolution properties and the notations of the graphical Eq. (3.2.10), we find

$$r_z(\nu,f) = G_1(\nu)\, G_2(f) \underset{f}{*} k_z(\nu,f) \tag{5.7.6}$$

$$R_z(t,\tau) = g_2(\tau)\, g_1(t) \underset{t}{*} K_z(t,\tau) \tag{5.7.7}$$

$$\rho_z(t,f) = g_1(t) \underset{t}{*} W_z(t,f) \underset{f}{*} G_2(f) \tag{5.7.8}$$

where $K_z(t,\tau) = z(t+\frac{\tau}{2})\, z^*(t-\frac{\tau}{2})$ is the (unsmoothed) IAF, $k_z(\nu,f)$ is its 2DFT (the spectral autocorrelation function), $W_z(t,f)$ is the WVD, $r_z(\nu,f)$ is the smoothed spectral autocorrelation function, $R_z(t,\tau)$ is the smoothed IAF, and $\rho_z(t,f)$ is the quadratic TFD.

Eq. (5.7.7) shows that the effect of the lag-dependent factor on the TFD is simply "lag windowing", i.e. multiplication by the same factor in the (t,τ) domain before transforming to the (t,f) domain. In Eq. (5.7.8), the two convolutions are associative (i.e. can be performed in either order), so that we may consider the Doppler-dependent and lag-dependent factors as leading to separate convolutions in time and frequency, respectively.

A **Doppler-independent** (DI) kernel is a special case of a separable kernel obtained by putting

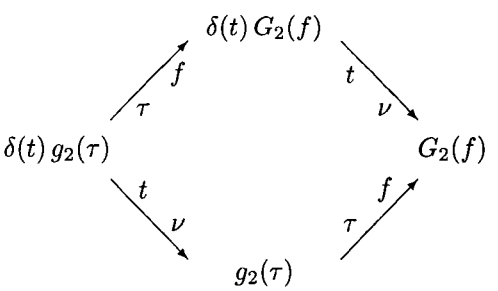
$$G_1(\nu) = 1 \tag{5.7.9}$$

in Eqs. (5.7.1) and (5.7.2), which then become

$$g(\nu,\tau) = g_2(\tau) \tag{5.7.10}$$

$$g_1(t) = \delta(t). \tag{5.7.11}$$

Making these substitutions in Eqs. (5.7.4) to (5.7.8), we obtain



(5.7.12)

$$\mathcal{A}_z(\nu,\tau) = g_2(\tau)\, A_z(\nu,\tau) \tag{5.7.13}$$

$$r_z(\nu,f) = G_2(f) \underset{f}{*} k_z(\nu,f) \tag{5.7.14}$$

$$R_z(t,\tau) = g_2(\tau)\, K_z(t,\tau) \tag{5.7.15}$$

$$\rho_z(t,f) = G_2(f) \underset{f}{*} W_z(t,f). \tag{5.7.16}$$

As seen in Eq. (5.7.12), a "Doppler-independent" kernel is indeed independent of Doppler in all four domains. In the Doppler-lag domain it is a function of lag alone. Eq. (5.7.16) shows that a DI kernel causes smoothing (or "smearing") of the WVD in the frequency direction only. The graphical Eq. (3.2.10) defined the notation

$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\{R_z(t,\tau)\}. \tag{5.7.17}$$

Substituting Eq. (5.7.15) into Eq. (5.7.17) gives

$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\{g_2(\tau)\,K_z(t,\tau)\} \tag{5.7.18}$$

which shows that a quadratic TFD with a DI kernel is a *windowed WVD*; the "windowing" is applied in the lag direction before Fourier transformation from lag to frequency.

A **lag-independent** (LI) kernel is another special case of a separable kernel, obtained by putting

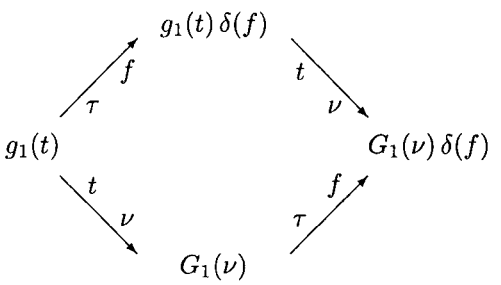
$$g_2(\tau) = 1 \tag{5.7.19}$$

in Eqs. (5.7.1) and (5.7.3), which then become

$$g(\nu,\tau) = G_1(\nu) \tag{5.7.20}$$

$$G_2(f) = \delta(f). \tag{5.7.21}$$

Making these substitutions in Eqs. (5.7.4) to (5.7.8), we obtain



(5.7.22)

$$\mathcal{A}_z(\nu,\tau) = G_1(\nu)\,A_z(\nu,\tau) \tag{5.7.23}$$

$$r_z(\nu,f) = G_1(\nu)\,k_z(\nu,f) \tag{5.7.24}$$

$$R_z(t,\tau) = g_1(t) \underset{t}{*} K_z(t,\tau) \tag{5.7.25}$$

$$\rho_z(t,f) = g_1(t) \underset{t}{*} W_z(t,f). \tag{5.7.26}$$

The last result shows that an LI kernel causes smoothing of the WVD in the time direction only.

As seen in Eq. (5.7.22), a "lag-independent" kernel is indeed independent of lag in all four domains. In the time-lag domain it is a function of time alone; for this reason, such kernels have been called "time-only kernels" [3,4].

The WVD kernel is $g(\nu,\tau) = 1$, which is both Doppler-independent and lag-independent; it may be regarded as DI with $g_2(\tau) = 1$ or as LI with $G_1(\nu) = 1$.

Table 5.7.1: TFD properties and associated kernel requirements for separable, Doppler-independent and lag-independent kernels. *Explanation of properties*: **Time marginal:** The integral of the TFD over frequency is the instantaneous power. **Freq. marginal:** The integral of the TFD over time is the energy spectrum. **Inst. Freq.:** The IF is the first moment of the TFD w.r.t. frequency. **Time delay:** The time delay is the first moment of the TFD w.r.t. time. **Time support:** If the non-zero values of the signal are confined to a certain time interval, so are the non-zero values of the TFD. **Freq. support:** If the non-zero values of the spectrum are confined to a certain frequency range, so are the non-zero values of the TFD. "WVD only*" (with asterisk) means a WVD multiplied by an arbitrary constant.

PROPERTY	KERNEL CONSTRAINTS		
	Separable $g(\nu,\tau) = G_1(\nu)\, g_2(\tau)$	DI $G_1(\nu) = 1$	LI $g_2(\tau) = 1$
Realness	$G_1(\nu)\, g_2(\tau)$ $= G_1^*(-\nu)\, g_2^*(-\tau)$.	$G_2(f)$ is real.	$g_1(t)$ is real.
Time marginal	$G_1(\nu)\, g_2(0) = 1 \ \ \forall\nu$	$g_2(0) = 1$	WVD only
Freq. marginal	$G_1(0)\, g_2(\tau) = 1 \ \ \forall\tau$	WVD only	$G_1(0) = 1$
Inst. freq.	$G_1(\nu)\, g_2(0) = \text{const.}$ $g_2'(0) = 0$.	$g_2'(0) = 0$	WVD only*
Time delay	$G_1(0)\, g_2(\tau) = \text{const.}$ $G_1'(0) = 0$.	WVD only*	$G_1'(0) = 0$
Time support	DI only	Always	WVD only*
Freq. support	LI only	WVD only*	Always
RID potential	Unrestricted	Inner artifacts	Cross-terms

5.7.3 Properties

Table 5.7.1 is extracted from Table 3.3.1 on p. 75. The properties of time-shift invariance and frequency-shift invariance are omitted, being common to all quadratic TFDs. Positivity is omitted because, in practice, the design of a separable kernel involves a deliberate sacrifice of non-negativity in favor of higher (t, f) resolution.

Let us define a **proper** DI or LI kernel as one that is *non*-constant (so that the resulting TFD is *not* a WVD, with or without amplitude scaling). Similarly, let us define a **proper** separable kernel as one that is neither DI nor LI. Then Table 5.7.1 shows that a TFD with a DI kernel can satisfy the realness, time marginal, time support and instantaneous frequency (IF) properties; but no proper DI kernel satisfies the frequency marginal, frequency support or time delay property. Similarly, a TFD with an LI kernel can satisfy the realness, frequency marginal, frequency support and time delay properties; but no proper LI kernel satisfies the time marginal, time support or IF property.

The reduced-interference property ("RID potential") requires further explanation. The WVD may contain interference terms of two kinds. **Inner artifacts** or "inner interference terms" [see Article 4.2] are caused by nonlinear frequency modulation laws, and cause the WVD to alternate as we move normal to the expected feature(s) in the (t, f) plane. In the case of a multicomponent signal, **Cross-terms** or "outer interference terms" [see Article 4.2] are caused by cross-product terms in the IAF $K_z(t, \tau)$, and cause the WVD to alternate as we move parallel to the expected features in the (t, f) plane.

Thus the inner artifacts alternate as we move in the frequency direction, and may therefore be suppressed by convolution with a sufficiently long $G_2(f)$ [see Eq. (5.7.16)], which corresponds to a sufficiently short $g_2(\tau)$. This is possible for a DI kernel but not an LI kernel. Similarly, the cross-terms alternate as we move in the time direction, and may therefore be suppressed by convolution with a sufficiently long $g_1(t)$ [see Eq. (5.7.26)], which corresponds to a sufficiently short $G_1(\nu)$. This is possible for an LI kernel but not a DI kernel. A proper separable kernel causes convolution in both time and frequency and can suppress both kinds of interference terms.

For an LI kernel, the suppression of cross-terms is facilitated if the components are of slowly-varying frequencies, so that the cross-terms extend (and alternate) approximately in the time direction. Furthermore, the loss of frequency resolution caused by convolution with $g_1(t)$ is proportional to the rate of change of the IF; for constant frequency, the components run parallel to the time axis, so that there is no loss of resolution apart from that caused by the time-variation of frequency resolution in the WVD.

5.7.4 Design Examples of Separable-Kernel TFDs

Early experience suggested that the kernel of a RID must exhibit a *two-dimensional* low-pass characteristic in the Doppler-lag domain (see e.g. [1], pp. 79–81). The **B-distribution (BD)** defined in [5] has the separable time-lag kernel

$$G_{\mathrm{B}}(t, \tau) = |\tau|^{\beta} \cosh^{-2\beta} t \tag{5.7.27}$$

where β is a positive real parameter that controls the degree of smoothing. This kernel is low-pass in the Doppler dimension but *not* in the lag dimension. But, paradoxically, the B-distribution has shown impressive reduced-interference properties for certain signals [5].

The best results from the BD are consistently obtained for small positive values of β, for which the lag-dependent factor is nearly constant apart from a "slot" at $\tau = 0$. Moreover, any desired lag-dependence can be introduced later by windowing prior to Fourier transformation from τ to f, as is often done for computational economy or improved time resolution. Accordingly, the BD was modified in [3, 4] by making the "lag-dependent" factor *exactly* constant. The resulting **modified B-distribution (MBD)** had an LI kernel. This inspired a more thorough inquiry into the properties of LI kernels, and the findings explained the behavior of the BD,

whose kernel may fairly be described as "nearly LI" for small positive β.

The time-lag kernel of the MBD is

$$G_{\mathrm{MB}}(t,\tau) = g_\beta(t) = \frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} \xi \, d\xi} \tag{5.7.28}$$

where β is a positive real parameter and the denominator is for normalization. The graph of $g_\beta(t)$ vs. t is a bell-shaped curve whose spread is inversely related to β.

If the signal is a tone or sum of tones, we can obtain closed-form expressions showing that the MBD has optimal concentration about the IF law. In particular, the MBD of a single tone is a delta function of frequency, while the MBD of the sum of two tones comprises two delta functions plus a cross-term whose amplitude is controlled by β (see [6] and Article 10.3). For most signals there is no closed-form expression for the MBD, so we must resort to numerical computations with discretized variables. For discrete time n and discrete lag m, the MBD kernel becomes

$$G_{\mathrm{MB}}(n,m) = g_\beta(n) = \frac{\cosh^{-2\beta} n}{\sum_i \cosh^{-2\beta} i}. \tag{5.7.29}$$

The following numerical examples include one MBD for comparison. They also include TFDs with separable kernels using two types of window functions that are well known in digital signal processing and spectral analysis. The M-point **Hamming** function, where M is odd, is

$$\mathrm{hamm}_M(i) = 0.54 + 0.46 \cos \tfrac{2\pi i}{M} \; ; \quad -\tfrac{M-1}{2} \le i \le \tfrac{M-1}{2} \tag{5.7.30}$$

where i is discrete time or lag. The L-point **Hanning** function, where L is odd, is

$$\mathrm{hann}_L(i) = 0.5 + 0.5 \cos \tfrac{2\pi i}{L} \; ; \quad -\tfrac{L-1}{2} \le i \le \tfrac{L-1}{2}. \tag{5.7.31}$$

In the numerical examples given below, the Hamming function is preferred in the lag direction and the Hanning in the time direction, in order to minimize ripples in the (t,f) domain; and whenever a proper separable kernel is compared with DI and LI kernels, it has the same lag factor as the DI kernel and the same Doppler factor (or time factor) as the LI kernel.

5.7.5 Results and Discussion

Some numerical computations of separable-kernel TFDs, and of other TFDs for purposes of comparison, are presented in the following graphs. Complete specifications of signals and kernels are given in the figure captions for ease of reference. Each graph includes the TFD (main panel, with time axis vertical), time plot (left panel) and magnitude spectrum (bottom panel).

Fig. 5.7.1 shows two TFDs of a signal comprising a tone (constant frequency) and a chirp (linearly increasing frequency). Part (a) shows the WVD, with the prominent cross-term. Part (b) shows the effect of an LI kernel, which smoothes

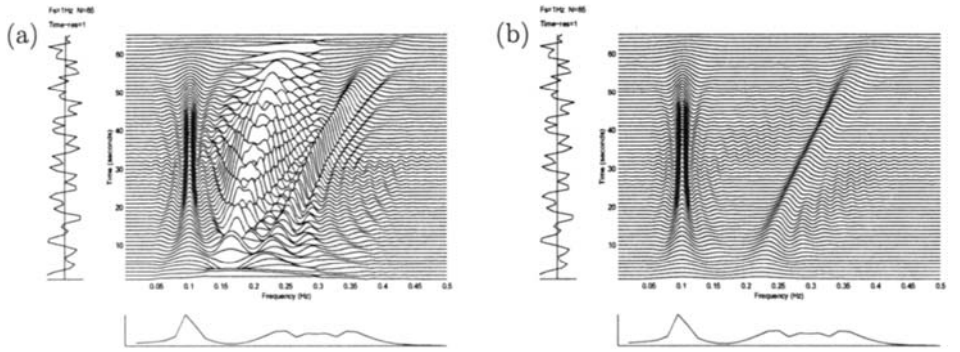


Fig. 5.7.1: TFDs of the sum of a tone (frequency 0.1) and a linear FM signal (frequency range 0.2 to 0.4), unit amplitudes, duration 65 samples, sampling rate 1 Hz: (a) WVD; (b) lag-independent, $G(n,m) = \text{hann}_{15}(n)$.

the WVD in the time direction, suppressing the oscillatory cross-term. For the tone, this smoothing causes only a slight loss of frequency resolution, and only because the frequency resolution of the WVD varies with time. For the chirp, the loss of frequency resolution is greater because the direction of smoothing is not parallel to the IF law; that is, the smoothing is "across" as well as "along" the component.

Fig. 5.7.2 compares six TFDs of a two-component signal with slowly-varying frequencies; the lower-frequency component is a pure tone. In Fig. 5.7.3, the tone is replaced by a faster-varying sinusoidal FM signal (nonlinear FM). In each figure, part (a) shows the WVD, while part (b) shows the effect of a DI kernel, part (c) an LI kernel, and part (d) a separable kernel combining the lag and time functions of parts (b) and (c). Note that (b) and (d) are related by the same time-smoothing as (a) and (c), while (c) and (d) are related by the same frequency-smoothing as (a) and (b). In each figure, part (f) shows a spectrogram for comparison with the separable-kernel TFD.

In Fig. 5.7.2, a cross-term is prominent in the WVD (a) and is not suppressed by the DI kernel (b). It is suppressed by the Hanning LI kernel (c) and the MBD kernel (e), which is also LI. Both LI kernels are bell-shaped functions of time, and their parameters have been chosen to give similar degrees of smoothing in time, accounting for the similarity between graphs (c) and (e). The proper separable kernel (d) has lower frequency resolution, but less ripple about the IF laws, than the LI kernels. The spectrogram (f) gives the best suppression of artifacts and the lowest resolution. Comparing Fig. 5.7.1(b) and Fig. 5.7.2(c), we see that the faster-varying IF, which causes a faster-varying cross-term frequency and lower minimum beat frequency, needs a longer time-smoothing function for adequate suppression of cross-terms.

In Fig. 5.7.3, both cross-terms and inner artifacts are visible. The cross-terms appear as "rough terrain" between the components, while the inner artifacts appear as spurious ridges in the sinusoidal FM component; both types are prominent in the WVD (a). The DI kernel (b) is effective against the inner artifacts. The LI

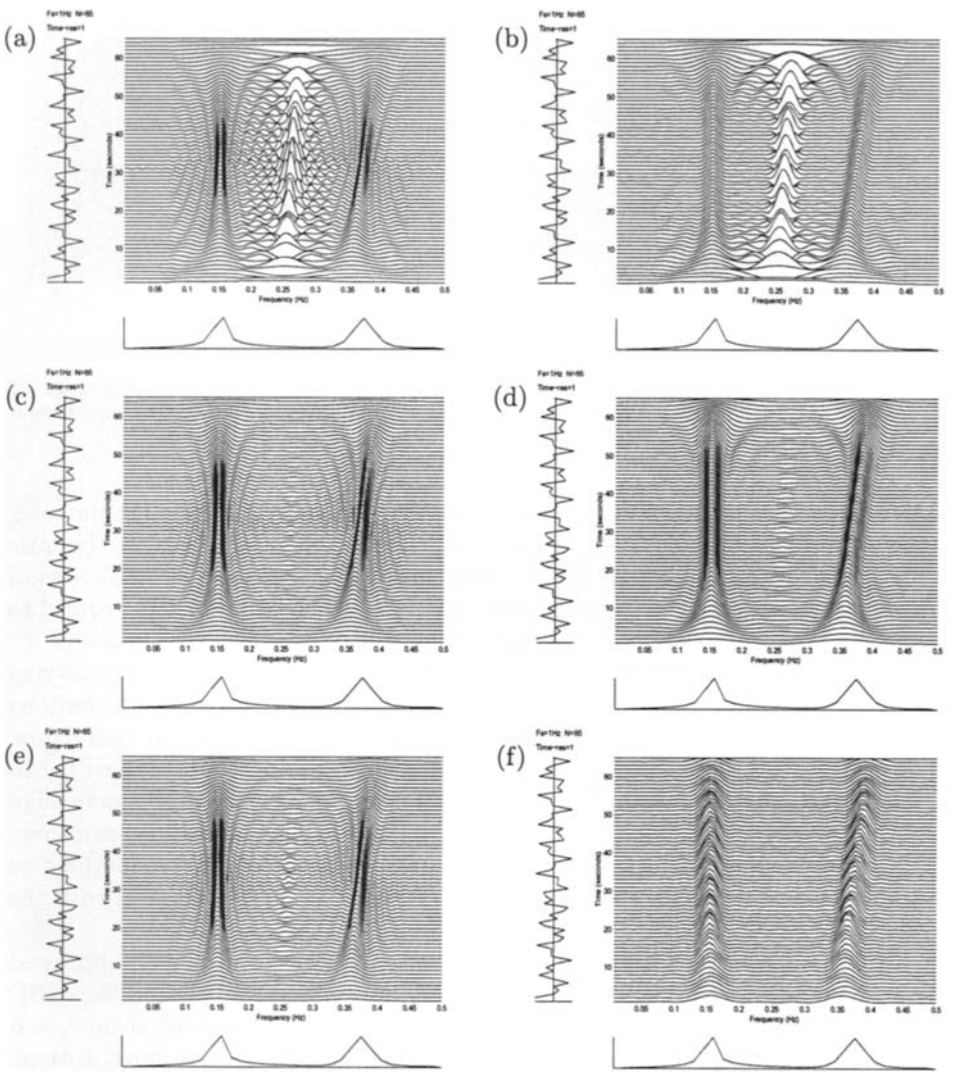


Fig. 5.7.2: TFDs of the sum of a tone (frequency 0.15) and a linear FM signal (frequency range 0.35 to 0.4), unit amplitudes, duration 65 samples, sampling rate 1 Hz: (a) WVD; (b) Doppler-independent, $G(n,m) = \mathrm{hamm}_{47}(m)$; (c) lag-independent, $G(n,m) = \mathrm{hann}_{11}(n)$; (d) separable, $G(n,m) = \mathrm{hann}_{11}(n)\ \mathrm{hamm}_{47}(m)$; (e) modified B, $\beta = 0.2$; (f) spectrogram, 35-point rectangular window.

kernel (c) is effective against the cross-terms. The separable kernel (d) is effective against both. The well-known Choi-Williams distribution (e), with the parameter σ visually optimized, spreads the inner artifacts in time and the cross-terms in frequency but does not satisfactorily suppress either. In this example the separable-kernel TFD (d) appears to offer a better compromise between resolution and cross-

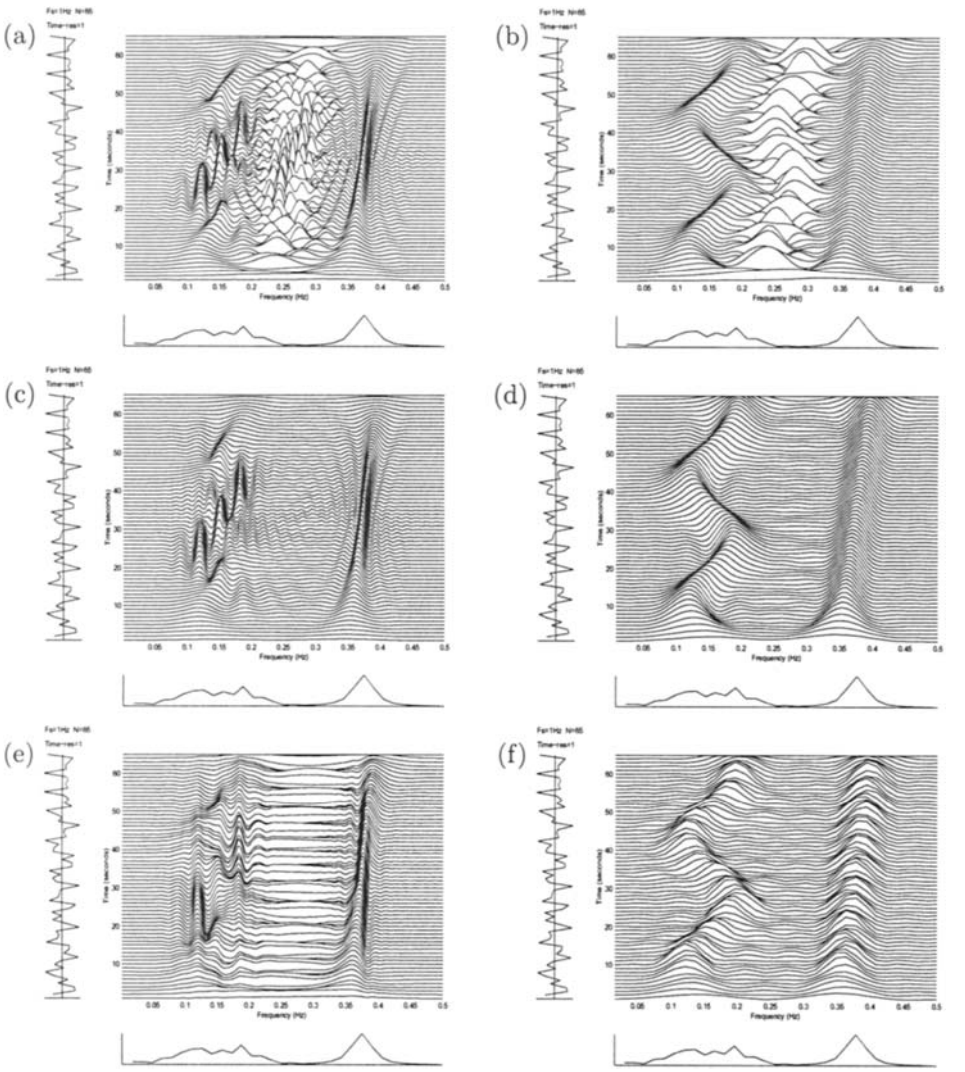


Fig. 5.7.3: TFDs of the sum of a sinusoidal FM signal (frequency 0.15±0.05, 2 cycles of modulation) and a linear FM signal (frequency range 0.35 to 0.4), unit amplitudes, duration 65 samples, sampling rate 1 Hz: (a) WVD; (b) Doppler-independent, $G(n,m) = \mathrm{hamm}_{23}(m)$; (c) lag-independent, $G(n,m) = \mathrm{hann}_{11}(n)$; (d) separable, $G(n,m) = \mathrm{hann}_{11}(n)\ \mathrm{hamm}_{23}(m)$; (e) Choi-Williams, $\sigma = 1$; (f) spectrogram, 17-point rectangular window.

term suppression than either the spectrogram (f) or the CWD (e), neither of which has a separable kernel. Comparing Fig. 5.7.2(d) with Fig. 5.7.3(d), we may confirm that a more nonlinear IF requires a shorter lag window (for the suppression of inner artifacts), giving coarser frequency resolution.

5.7.6 Summary and Conclusions

A separable kernel gives separate control of the frequency-smoothing and time-smoothing of the WVD: the lag-dependent factor causes a convolution in the frequency direction in the (t, f) plane, while the Doppler-dependent factor causes a convolution in the time direction. A Doppler-*in*dependent (DI) kernel smoothes the WVD in the frequency direction only, reducing the inner artifacts and preserving the time marginal. A lag-*in*dependent (LI) kernel smoothes the WVD in the time direction only, reducing the cross-terms and preserving the frequency marginal.

For an LI kernel, slower variations in component frequencies allow easier suppression of cross-terms and higher frequency resolution in the TFD; such kernels should therefore be considered for representing multicomponent signals with slowly-varying instantaneous frequencies. For a multicomponent signal with at least one highly nonlinear IF law, a lag-dependent factor is also needed to suppress the inner artifacts.

The separable-kernel approach allows a complete understanding and appraisal of the properties and behavior of the smoothed WVD [2], the B-distribution [5] and the modified B-distribution [3,4]. It also enables the construction of high-resolution quadratic TFDs using the classical smoothing functions commonly encountered in digital signal processing and spectral analysis.

References

[1] W. J. Williams and J. Jeong, "Reduced interference time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 3, pp. 74–97, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[2] E. F. Velez and H. Garudadri, "Speech analysis based on smoothed Wigner-Ville distribution," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 15, pp. 351–374, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[3] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, vol. II, pp. 657–660, Istanbul, 5–9 June 2000.

[4] Z. M. Hussain and B. Boashash, "Multi-component IF estimation," in *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pp. 559–563, Pocono Manor, PA, 14–16 August 2000.

[5] B. Barkat and B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis," *IEEE Trans. Signal Processing*, vol. 49, pp. 2232–2239, October 2001.

[6] Z. M. Hussain and B. Boashash, "Design of time-frequency distributions for amplitude and IF estimation of multicomponent signals," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, vol. 1, pp. 339–342, Kuala Lumpur, 13–16 August 2001.

5.8 FRACTIONAL FOURIER TRANSFORM AND GENERALIZED-MARGINAL TFDs[0]

5.8.1 Fractional Fourier Transform

The fractional Fourier transform (FRFT) is a rotation of the time-frequency plane. For a real α, the fractional Fourier transform $\mathbf{F}_\alpha$ with angle α is defined by

$$(\mathbf{F}_\alpha s)(u) =$$

$$\begin{cases} \sqrt{\frac{1-j\cot\alpha}{2\pi}} e^{j\frac{u^2}{2}\cot\alpha} \int_{-\infty}^{\infty} s(t) e^{j\frac{t^2}{2}\cot\alpha} e^{-jut\csc\alpha} dt, & \text{if } \alpha \text{ is not a multiple of } \pi, \\ s(u), & \text{if } \alpha \text{ is a multiple of } 2\pi, \\ s(-u), & \text{if } \alpha+\pi \text{ is a multiple of } 2\pi. \end{cases} \tag{5.8.1}$$

Also, one can see that $\mathbf{F}_{2n\pi}$ for an integer n is the identity transformation and $\mathbf{F}_{\frac{\pi}{2}}$ is the traditional Fourier transform $\mathbf{F}$. Moreover, the following rotation property holds:

$$\mathbf{F}_{\alpha+\beta} = \mathbf{F}_\alpha \mathbf{F}_\beta.$$

For more details, see, for example, [1, 2]. A numerical example of the FRFT of a signal is illustrated in Fig. 5.8.1.

With the FRFT, it was proved in [2, 3] that a rotation of a Wigner-Ville distribution is still a Wigner-Ville distribution as explained below.

Let $W_s(t,\omega)$ denote the Wigner-Ville distribution of a signal $s(t)$, i.e.,

$$W_s(t,\omega) = \int s(t+\frac{\tau}{2}) s^*(t-\frac{\tau}{2}) e^{-j\omega\tau} d\tau. \tag{5.8.2}$$

Let α be an angle and $(\tilde{t},\tilde{\omega})$ be a rotation of (t,ω) with angle α:

$$\begin{cases} \tilde{t} &= t\cos\alpha + \omega\sin\alpha, \\ \tilde{\omega} &= -t\sin\alpha + \omega\cos\alpha, \end{cases} \tag{5.8.3}$$

and

$$\tilde{W}_s(\tilde{t},\tilde{\omega}) = W_s(\tilde{t}\cos\alpha - \tilde{\omega}\sin\alpha, \tilde{t}\sin\alpha + \tilde{\omega}\cos\alpha).$$

Then (see e.g. [2])

$$\tilde{W}_s(\tilde{t},\tilde{\omega}) = \int (\mathbf{F}_\alpha s)(\tilde{t}+\frac{\tau}{2}) (\mathbf{F}_\alpha s)^*(\tilde{t}-\frac{\tau}{2}) e^{-j\tilde{\omega}\tau} d\tau. \tag{5.8.4}$$

The equation (5.8.4) tells us that the rotation $\tilde{W}_s(\tilde{t},\tilde{\omega})$ of the Wigner-Ville distribution $W_s(t,\omega)$ of a signal s is the Wigner-Ville distribution of the signal $\mathbf{F}_\alpha s$. It was

[0]Author: **Xiang-Gen Xia**, Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA (xxia@ee.udel.edu). Reviewers: Paulo M. Oliveira and Shie Qian.

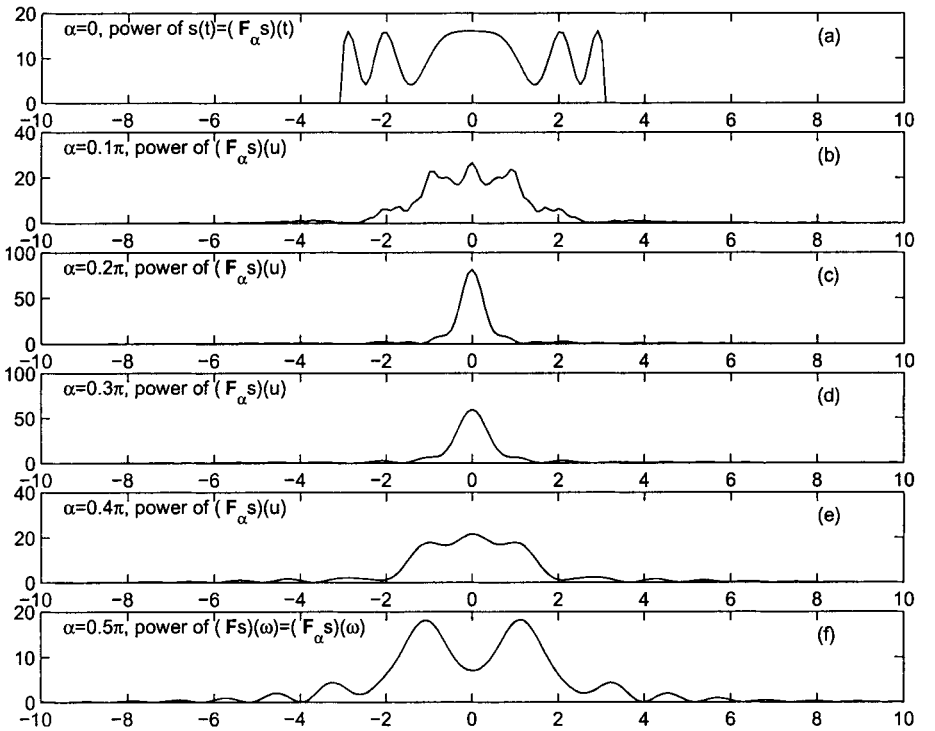


Fig. 5.8.1: The Fractional Fourier transform of a signal $s(t)$ in (a) for (a) $\alpha = 0\pi$; (b) $\alpha = 0.1\pi$; (c) $\alpha = 0.2\pi$; (d) $\alpha = 0.3\pi$; (e) $\alpha = 0.4\pi$; (f) $\alpha = 0.5\pi$.

also proved in [3] that a rotation of a Radon-Wigner distribution is also a Radon-Wigner distribution by using the FRFT technique. It is known that a Wigner-Ville distribution satisfies the conventional marginal properties, i.e.,

$$\int \tilde{W}_s(\tilde{t}, \tilde{\omega}) d\tilde{\omega} = |(\mathbf{F}_\alpha s)(\tilde{t})|^2, \tag{5.8.5}$$

and

$$\int \tilde{W}_s(\tilde{t}, \tilde{\omega}) d\tilde{t} = |\mathbf{F}(\mathbf{F}_\alpha s)(\tilde{\omega})|^2 = |(\mathbf{F}_{\alpha+\pi/2} s)(\tilde{\omega})|^2. \tag{5.8.6}$$

5.8.2 Generalized-Marginal Time-Frequency Distribution

For joint time-frequency (TF) distributions, one often imposes the marginal properties: the integrals of a TF distribution along the time t and the frequency ω are the powers of the signal in the frequency and the time domains, respectively. Sat-

isfaction of the time and the frequency marginals in Cohen's class[1] is equivalent to the kernel property $\phi(0,\tau) = \phi(\theta,0) = 1$ for all real θ and τ, where $\phi(\theta,\tau)$ is the Doppler-lag kernel[2] (see [4] and Table 3.3.1). This fact will later be seen to be just a special case of a more general property.

It is normally accepted that, when signals are chirp signals, the TF distributions should be concentrated on lines in the TF plane. The question is the following. If we have some prior information about a signal, can we take advantage of it in the design of a TF distribution? Or, can we impose some requirements on a TF distribution along these lines? If so, how? The generalized-marginal TF distributions proposed in [5] give an answer to these questions.

5.8.2.1 Generalized Marginals

Generalized marginals are the marginals beyond the usual time and frequency marginals. Let $P_s(t,\omega)$ be a TF distribution of a signal $s(t)$. Let L_α denote the straight line

$$\omega\cos\alpha - t\sin\alpha = 0\,, \tag{5.8.7}$$

i.e. the line on the time-frequency plane through the origin with angle α. Let $L_\alpha(u)$ denote the line

$$\omega\cos\alpha - t\sin\alpha = u\,, \tag{5.8.8}$$

i.e. the general member of the family of lines parallel to L_α, parameterized by the real number u, where u is the signed perpendicular distance from L_α. We call $P_s(t,\omega)$ a **generalized-marginal time-frequency distribution**, if the line integrals of $P_s(t,\omega)$ along the lines $L_{\alpha_k}(u)$, for $k = 1,2,...,N$, are the powers of the FRFT with angles $\alpha_k + \pi/2$, $k = 1,2,...,N$, of the signal s, respectively. In other words,

$$\int_{L_{\alpha_k}(u)} P_s(t,\omega)dx = |(\mathbf{F}_{\alpha_k+\pi/2}s)(u)|^2,\ \ k = 1,2,...,N, \tag{5.8.9}$$

where the dummy variable x is on the line $L_{\alpha_k}(u)$. This is written more simply as

$$\int_{L_{\alpha_k}} P_s(t,\omega)dx = |\mathbf{F}_{\alpha_k+\pi/2}s|^2,\ \ k = 1,2,...,N.$$

It is clear that, when $\alpha_1 = 0, \alpha_2 = \pi/2$, and $N = 2$, the above generalized-marginals are the conventional marginals. Also, the angles α_k may be chosen to be close to the angles of chirp signals in the TF plane.

5.8.2.2 Generalized-Marginal Time-Frequency Distributions in Cohen's Class

In this subsection, we study TF distributions with kernels $\phi(\theta,\tau)$ in Cohen's class [4], which are generalized-marginal TF distributions (5.8.9). We show that a TF

[1] That is, the quadratic class; see p. 68n.

[2] In this article, $\theta = 2\pi\nu$ and $\omega = 2\pi f$.

distribution with kernel $\phi(\theta,\tau)$ in Cohen's class is a generalized-marginal one (5.8.9) if and only if its kernel $\phi(\theta,\tau)$ is equal to 1 on the lines that are perpendicular to L_{α_k}, $k=1,2,...,N$, and pass through the origin. This implies that the Wigner-Ville distribution satisfies all the generalized-marginals and it is the only one in Cohen's class with this property.

A TF distribution for a signal $s(t)$ in Cohen's class is defined by

$$P_s(t,\omega)=\frac{1}{4\pi^2}\int\int A_s(\theta,\tau)e^{-j\theta t-j\tau\omega}d\theta d\tau,$$

where $A_s(\theta,\tau)$ is the generalized ambiguity function of the signal $s(t)$ with a kernel $\phi(\theta,\tau)$:

$$A_s(\theta,\tau)=\phi(\theta,\tau)\int s(u+\frac{\tau}{2})s^*(u-\frac{\tau}{2})e^{j\theta u}du.$$

The TF distribution $P_s(t,\omega)$ can be also written as

$$P_s(t,\omega)=\int\int\int e^{-j\theta t-j\tau\omega+j\theta u}\phi(\theta,\tau)s(u+\frac{\tau}{2})s^*(u-\frac{\tau}{2})d\theta d\tau du. \qquad (5.8.10)$$

Then, $P(t,\omega)$ is a generalized-marginal TF distribution if and only if the following holds. Let

$$\begin{array}{rcl}\tilde{t} &=& t\cos\alpha+\omega\sin\alpha,\\ \tilde{\omega} &=& -t\sin\alpha+\omega\cos\alpha,\end{array} \quad \text{or} \quad \begin{array}{rcl}t &=& \tilde{t}\cos\alpha-\tilde{\omega}\sin\alpha,\\ \omega &=& \tilde{t}\sin\alpha+\tilde{\omega}\cos\alpha.\end{array}$$

Then, the condition (5.8.9) is equivalent to

$$\int P_s(\tilde{t}\cos\alpha_k-\tilde{\omega}\sin\alpha_k,\tilde{t}\sin\alpha_k+\tilde{\omega}\cos\alpha_k)d\tilde{t}=|(\mathbf{F}_{\alpha_k+\pi/2}s)(\tilde{w})|^2,\ \ k=1,2,...,N. \qquad (5.8.11)$$

In other words, a TF distribution $P_s(t,\omega)$ is a generalized-marginal distribution if and only if it satisfies (5.8.11). We now focus on the TF distributions $P_s(t,\omega)$ in (5.8.10), for a given angle α in (5.8.11). Let us see what the left hand side of (5.8.11) with angle α for $P_s(t,\omega)$ in (5.8.10) is.

$$\begin{aligned}&\int P_s(\tilde{t}\cos\alpha-\tilde{\omega}\sin\alpha,\tilde{t}\sin\alpha+\tilde{\omega}\cos\alpha)d\tilde{t}\\ &=\frac{1}{\cos\alpha}\int\int e^{-j\tilde{\omega}\tau(\sin\alpha\tan\alpha+\cos\alpha)-ju\tau\tan\alpha}\phi(-\tau\tan\alpha,\tau)s(u+\frac{\tau}{2})s^*(u-\frac{\tau}{2})d\tau du\\ &=\int e^{-j\tilde{\omega}\tau}\phi(-\tau\sin\alpha,\tau\cos\alpha)A_s(-\tau\sin\alpha,\tau\cos\alpha)d\tau,\end{aligned}$$

where A_s is the ambiguity function of s. It was proved (see, for example, [6]) that

$$A_s(-\tau\sin\alpha,\tau\cos\alpha)=A_{\mathbf{F}_{\alpha+\pi/2}s}(\tau,0).$$

Table 5.8.1: Generalized-Marginal Kernels

Name	Kernel $\phi(\theta,\tau)$	Generalized-Marginal $\tilde{\phi}(\theta,\tau)$ with angles α_k, $1 \le k \le N$.
Margenau-Hill	$\cos\frac{\theta\tau}{2}$	$\cos[0.5\prod_{k=1}^{N}(\theta\cos\alpha_k+\tau\sin\alpha_k)]$
Rihaczek	$e^{0.5j\theta\tau}$	$e^{0.5j\prod_{k=1}^{N}(\theta\cos\alpha_k+\tau\sin\alpha_k)}$
sinc	$\frac{\sin a\theta\tau}{a\theta\tau}$	$\frac{\sin[a\prod_{k=1}^{N}(\theta\cos\alpha_k+\tau\sin\alpha_k)]}{a\prod_{k=1}^{N}(\theta\cos\alpha_k+\tau\sin\alpha_k)}$
Page	$e^{0.5\theta\lvert\tau\rvert}$	$e^{0.5j\prod_{k_1=1}^{N_1}\lvert\theta\cos\alpha_{k_1}+\tau\sin\alpha_{k_1}\rvert\prod_{k_2=1}^{N_2}(\theta\cos\alpha_{k_2}+\tau\sin\alpha_{k_2})}$
Choi-Williams	$e^{-\frac{\theta^2\tau^2}{\sigma}}$	$e^{-\frac{1}{\sigma}\prod_{k=1}^{N}(\theta\cos\alpha_k+\tau\sin\alpha_k)^2}$

Therefore,

$$\int P_s(\tilde{t}\cos\alpha-\tilde{\omega}\sin\alpha,\tilde{t}\sin\alpha+\tilde{\omega}\cos\alpha)d\tilde{t}$$
$$=\int\int e^{j(u-\tilde{\omega})\tau}\phi(-\tau\sin\alpha,\tau\cos\alpha)|\mathbf{F}_{\alpha+\pi/2}s(u)|^2dud\tau.$$

Therefore, the generalized-marginal property holds if and only if

$$\int e^{j(u-\tilde{\omega})\tau}\phi(-\tau\sin\alpha,\tau\cos\alpha)d\tau=\delta(u-\tilde{\omega}),$$

i.e., $\phi(-\tau\sin\alpha,\tau\cos\alpha)=1$.

Although the above discussion is for one angle only, it is straightforward to generalize it to several angles α_k for $k=1,2,...,N$. Therefore, we have obtained the following result.

Theorem 5.8.1: *A time-frequency distribution $P_s(t,\omega)$ in (5.8.10) in Cohen's class with a kernel $\phi(\theta,\tau)$ is a generalized-marginal time-frequency distribution with angles α_k, $k=1,2,...,N$, as in (5.8.11) if and only if*

$$\phi(-\tau\sin\alpha_k,\tau\cos\alpha_k)=1,\ \textit{for all real}\ \tau,\ \textit{and}\ k=1,2,...,N, \tag{5.8.12}$$

in other words, ϕ is 1 on the lines perpendicular to the lines L_{α_k}, $k=1,2,...,N$, and passing through the origin.

With this result, one can easily modify the well-known kernels so that the corresponding modified TF distributions are generalized-marginals. We now list them in Table 5.8.1.

Theorem 5.8.1 also tells us that the Wigner-Ville distribution satisfies all marginal properties for all angles because $\phi(\theta, \tau) = 1$ for all real θ and τ. There is, however, a tradeoff between the number of generalized-marginals you want to impose and the freedom of choosing a kernel $\phi(\theta, \tau)$ in Cohen's class. As more generalized-marginals are required, there is less freedom in choosing kernels.

5.8.3 Summary and Conclusions

Fractional Fourier transforms can be used to introduce generalized marginal time-frequency distributions. Such distributions have been shown in the literature to be among the best time-frequency distributions in the sense of having the highest resolution for chirp-type signals.

Fractional Fourier transforms are further developed in Article 4.8.

References

[1] A. C. McBride and F. H. Kerr, "On Namias's fractional Fourier transforms," *IMA J. of Applied Mathematics*, vol. 39, no. 2, pp. 159–175, 1987.

[2] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Processing*, vol. 42, pp. 3084–3091, November 1994.

[3] A. W. Lohmann and B. H. Soffer, "Relationships between the Radon-Wigner and fractional Fourier transforms," *J. Optical Soc. of America A*, vol. 11, pp. 1798–1801, June 1994.

[4] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[5] X.-G. Xia, Y. Owechko, B. H. Soffer, and R. M. Matic, "On generalized-marginal time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 44, pp. 2882–2886, November 1996.

[6] H. L. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. III: "Radar-Sonar Signal Processing and Gaussian Signals in Noise". New York: Wiley, 1971. Reprinted Malabar, FL: Krieger, 1992. Reprinted New York: Wiley, 2001.

[7] S. Qian and D. Chen, *Joint Time-Frequency Analysis: Methods & Applications.* Upper Saddle River, NJ: Prentice-Hall, 1996.

Part III

Time-Frequency Methods

This Page Intentionally Left Blank

Chapter 6

Implementation and Realization of TFDs

Algorithms and computational issues are the keys to efficiently utilizing the properties of Time-Frequency Distributions (TFDs) for real life applications. This chapter presents procedures, techniques and methodologies for the efficient implementation of TFDs. The topic is covered in five articles with appropriate cross-referencing.

The discrete-time equivalent formulation of quadratic TFDs is defined for the purpose of digital computation (Article 6.1). An alternative method for realization of quadratic TFDs is to use the short-time Fourier transform (STFT) as a basis (6.2). The Gabor time-frequency representation may be expanded on a rectangular lattice, using the Fourier and Zak transforms for direct implementations (6.3). The computation of other quadratic TFDs can also be facilitated by using spectrogram decomposition (6.4). Finally, the computational procedure for implementing quadratic time-frequency methods directly is outlined, along with the necessary algorithms and MATLAB^{TM} code fragments (6.5).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

6.1 DISCRETE TIME-FREQUENCY DISTRIBUTIONS [0]

For the purposes of digital storage and processing, any real-life signal that is not discrete-time and time-limited must be made so by sampling and windowing. Moreover, if we wish to evaluate a continuous time-frequency distribution (TFD) numerically in a finite number of operations, we must be content with computing a sampled form of the TFD from a finite number of samples of the signal. For such reasons, we need to define discrete, time-limited equivalents of continuous TFDs. This article derives discrete forms of the Wigner-Ville distribution (WVD), the windowed WVD and the general quadratic TFD, and gives thirteen examples of discrete-time kernels. Thus it extends the material present in Chapters 2 and 3.

If the signal $z(t)$ is ideally sampled at times $t = n/f_s$, where n is an integer and f_s is the sampling rate, it becomes

$$z(t) \sum_{n=-\infty}^{\infty} \delta\big(t - \tfrac{n}{f_s}\big) = \sum_{n=-\infty}^{\infty} z\big(\tfrac{n}{f_s}\big)\, \delta\big(t - \tfrac{n}{f_s}\big) \tag{6.1.1}$$

where $\delta(\ldots)$ denotes the unit impulse function. We shall use a wide caret ($\widehat{\ }$) to denote a TFD that has been modified by sampling. In the lag (τ) domain, we shall consider only *ideal* sampling at $\tau = m/f_s$, where m is an integer. This in turn will draw attention to the discrete time values $t = n/f_s$, where n is an integer; but it will not be necessary to specify ideal sampling in the time domain.

6.1.1 The Discrete Wigner-Ville Distribution (DWVD)

The WVD of a continuous-time signal $z(t)$ has the form

$$W_z(t,f) = \underset{\tau \to f}{\mathcal{F}}\big\{z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\big\} = \int_{-\infty}^{\infty} z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\, e^{-j2\pi f\tau} d\tau. \tag{6.1.2}$$

Changing the variable of integration to $\theta = \tau/2$, we obtain the alternative definition

$$W_z(t,f) = 2\int_{-\infty}^{\infty} z(t+\theta)\, z^*(t-\theta)\, e^{-j4\pi f\theta} d\theta \tag{6.1.3}$$

which we shall find more convenient for the purpose of conversion to discrete form. The product $z(t+\theta)\, z^*(t-\theta)$ is the instantaneous autocorrelation function (IAF) in terms of t and θ. Let us consider this IAF as a function of θ and suppose that it is sampled.

Theorem 6.1.1: If $W_z(t,f)$ is modified by ideally sampling the IAF in θ at $\theta = \tau/2 = m/f_s$, where m is an integer and f_s is the sampling rate, and if the

[0]Authors: **Boualem Boashash** and **Gavin R. Putland**, Signal Processing Research Centre, Queensland University of Technology, GPO Box 2434, Brisbane, Q 4001, Australia (b.boashash@qut.edu.au, g.putland@qut.edu.au). Reviewers: G. Matz and LJ. Stanković.

modified TFD is denoted by $\widehat{W}_z(t,f)$, then

$$\widehat{W}_z\left(\tfrac{n}{f_s}, \tfrac{kf_s}{2N}\right) = 2 \sum_{|m|<N/2} z\left(\tfrac{n+m}{f_s}\right) z^*\left(\tfrac{n-m}{f_s}\right) e^{-j2\pi km/N} \tag{6.1.4}$$

where N is the duration (in samples) of $z(n/f_s)$; and the time support for $\widehat{W}_z$ is the same as for $z(n/f_s)$.

Proof/explanation: After sampling, the integrand in Eq. (6.1.3) becomes

$$z(t+\theta)\, z^*(t-\theta)\, e^{-j4\pi f\theta} \sum_{m=-\infty}^{\infty} \delta\left(\theta - \tfrac{m}{f_s}\right) \tag{6.1.5}$$

and the WVD becomes

$$\widehat{W}_z(t,f) = 2 \sum_{m=-\infty}^{\infty} z\left(t+\tfrac{m}{f_s}\right) z^*\left(t-\tfrac{m}{f_s}\right) e^{-j4\pi fm/f_s}. \tag{6.1.6}$$

Now let t be conveniently restricted to $t = n/f_s$ where n is an integer (this represents sampling of the signal in time, but the sampling is not ideal). Then, if $z(t)$ has finite duration, the time support of $z(n/f_s)$ can be written in the form $|n - n_0| < N/2$, where n_0 is real. Hence the time support of $\widehat{W}_z$ is

$$|n \pm m - n_0| < N/2 \tag{6.1.7}$$

where both signs must be satisfied. Because the sign can be chosen so as to increase the magnitude, we must have $|n - n_0| + |m| < N/2$, hence $|n - n_0| < N/2$ and $|m| < N/2$. So the time support for $\widehat{W}_z$ is the same as for $z(n/f_s)$, while the summation in Eq. (6.1.6) is restricted to $|m| < N/2$, giving a maximum of N terms.[1] The sampling in τ makes $\widehat{W}_z(t,f)$ periodic in f with period $f_s/2$, while the time-limiting gives a frequency resolution of N bins per period. So it is convenient to let

$$f = \frac{kf_s}{2N} \tag{6.1.8}$$

where k is an integer. With these restrictions, Eq. (6.1.6) reduces to Eq. (6.1.4). ∎

With a change of notation, Eq. (6.1.4) becomes

$$W_z[n,k] = 2 \sum_{|m|<N/2} z[n+m]\, z^*[n-m]\, e^{-j2\pi km/N}. \tag{6.1.9}$$

This $W_z[n,k]$ is called the **discrete WVD** or **DWVD**. If $z[n]$ has a duration not exceeding N samples, the DWVD is represented as an $N \times N$ real matrix. If the summand is extended periodically in m with period N (i.e. extended periodically in τ with period $2N/f_s$), we obtain

$$W_z[n,k] = 2 \underset{m \to k}{\mathrm{DFT}} \{z[n+m]\, z^*[n-m]\}\ ; \quad m \in \langle N \rangle \tag{6.1.10}$$

where $\langle N \rangle$ means any set of N consecutive integers.[2]

[1] N terms for odd N; $N-1$ terms for even N.

[2] For even N, the periodic extension is padded with a zero term.

It remains to find the minimum value of f_s that avoids aliasing in the frequency and Doppler domains. As usual, we define the instantaneous autocorrelation function as $K_z(t,\tau) = z(t+\frac{\tau}{2})\,z^*(t-\frac{\tau}{2})$, the spectrum as $Z(f) = \underset{t\to f}{\mathcal{F}}\{z(t)\}$, the WVD as $W_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\{K_z(t,\tau)\}$, and the ambiguity function as $A_z(\nu,\tau) = \underset{t\to\nu}{\mathcal{F}}\{K_z(t,\tau)\}$. Using these notations and the familiar properties of the FT, it can be shown that

$$W_z(t,f) = \left[2Z(2f)\,e^{j4\pi ft}\right] \underset{f}{*} \left[2Z^*(2f)\,e^{-j4\pi ft}\right] \tag{6.1.11}$$

$$A_z(\nu,\tau) = \left[Z(\nu)\,e^{j\pi\nu\tau}\right] \underset{\nu}{*} \left[Z^*(-\nu)\,e^{-j\pi\nu\tau}\right]. \tag{6.1.12}$$

If $Z(f)$ is zero outside the band $|f| < B/2$, then the spectrum on the right of Eq. (6.1.11), i.e. the WVD, is zero outside the band $|f| < B/2$, and the spectrum on the right of Eq. (6.1.12) is zero outside the band $|\nu| < B$. Aliasing will be avoided if the sampling rate in τ (namely $f_s/2$, since f_s is the sampling rate in θ) is at least B and the sampling rate in t (namely f_s) is at least $2B$; that is, aliasing will be avoided if $f_s \le 2B$. The minimum sampling rate in τ makes the WVD periodic in f with period B.

The sampling rate may be reduced in the case of an analytic signal. If $z(t)$ is the analytic associate of the real signal $s(t)$, whose spectrum $S(f)$ is zero outside the band $|f| < B/2$, then $Z(f)$ is zero outside the band $0 \le f < B/2$, so that the spectrum on the right of Eq. (6.1.11), i.e. the WVD, is zero outside the band $0 \le f < B/2$, and the spectrum on the right of Eq. (6.1.12) is zero outside the band $|\nu| < B/2$. Aliasing will be avoided if the sampling rate in τ is at least $B/2$ and the sampling rate in t is at least B; that is, aliasing will be avoided if $f_s \le B$. The minimum sampling rate in τ makes the WVD periodic in f with period $B/2$.

Now consider $W_z[n,k]$ as a matrix with k as the column index. If the sampling rate in t is B, there is no zero-padding in the frequency domain for an analytic signal, and all the columns of $W_z[n,k]$ are needed for the positive frequencies. If the sampling rate in t is $2B$, only half the columns of $W_z[n,k]$ are needed to represent the positive frequencies. The negative-frequency elements may be assumed to be zero if $z[n]$ is analytic, provided of course that the assumption of analyticity is not invalidated by short-segment effects, such as windowing of the signal in the n domain and/or windowing of the IAF in the m domain. The latter kind of windowing is discussed under the next heading.

6.1.2 The Windowed DWVD

Nonlinear IF laws and multiple signal components give rise to artifacts (interference terms, cross-terms) in the WVD [see Article 4.2]. The effect of nonlinear IF laws can be reduced, with a concomitant loss of frequency resolution, by windowing the IAF in the lag direction (i.e. in τ or θ) before taking the FT. For continuous time, if the window is

$$g(\tau) = \bar{g}(\tfrac{\tau}{2}) = \bar{g}(\theta)\ , \tag{6.1.13}$$

then the resulting TFD, denoted by $W_z^g(t,f)$, is

$$W_z^g(t,f) = \int_{-\infty}^{\infty} g(\tau)\, z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})\, e^{-j2\pi f\tau} d\tau \tag{6.1.14}$$

$$= 2\int_{-\infty}^{\infty} \bar{g}(\theta)\, z(t+\theta)\, z^*(t-\theta)\, e^{-j4\pi f\theta} d\theta \tag{6.1.15}$$

and is called the **windowed WVD** or **pseudo-WVD**. The effects of sampling and time-limiting on the windowed WVD are described by the following theorem.

Theorem 6.1.2: If $W_z^g(t,f)$ is modified by ideally sampling $g(\tau)$ in θ at $\theta = \tau/2 = m/f_s$, where m is an integer and f_s is the sampling rate, and if $g(\tau)$ is time-limited so that

$$g(\tau) = 0 \quad \text{for} \quad |\theta| = |\tfrac{\tau}{2}| \geq \tfrac{M}{2f_s} \tag{6.1.16}$$

where M is a positive integer, and if the modified TFD is denoted by $\widehat{W}_z^g(t,f)$, then

$$\widehat{W}_z^g\big(\tfrac{n}{f_s}, \tfrac{kf_s}{2M}\big) = 2\sum_{|m|<M/2} \bar{g}\big(\tfrac{m}{f_s}\big)\, z\big(\tfrac{n+m}{f_s}\big)\, z^*\big(\tfrac{n-m}{f_s}\big)\, e^{-j2\pi km/M} \tag{6.1.17}$$

and the time support for $\widehat{W}_z^g$ is the same as for $z(n/f_s)$.

Proof: Apart from the limits on τ, θ and m, which lead to the substitution $f = \frac{kf_s}{2M}$, the explanation is similar to that of Theorem 6.1.1. ∎

With a change of notation, Eq. (6.1.17) becomes

$$W_z^g[n,k] = 2\sum_{|m|<M/2} g[m]\, z[n+m]\, z^*[n-m]\, e^{-j2\pi km/M}. \tag{6.1.18}$$

This $W_z^g[n,k]$ is called the **windowed DVWD** or **pseudo-DWVD**. If $z[n]$ has a duration not exceeding N samples, the windowed DWVD is represented as an $N \times M$ matrix. If the summand is extended periodically in m with period M (i.e. extended periodically in τ with period $2M/f_s$), we obtain

$$W_z^g[n,k] = 2\,\underset{m\to k}{\mathrm{DFT}}\,\{g[m]\, z[n+m]\, z^*[n-m]\}\ ; \quad m \in \langle M \rangle\,. \tag{6.1.19}$$

6.1.3 The Discrete Quadratic TFD

In terms of continuous variables, the general quadratic TFD is

$$\rho_z(t,f) = \underset{\tau\to f}{\mathcal{F}}\big\{G(t,\tau) \underset{t}{*} [z(t+\tfrac{\tau}{2})\, z^*(t-\tfrac{\tau}{2})]\big\} \tag{6.1.20}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(u,\tau)\, z(t-u+\tfrac{\tau}{2})\, z^*(t-u-\tfrac{\tau}{2})\, du\, e^{-j2\pi f\tau} d\tau \tag{6.1.21}$$

where $G(t,\tau)$ is the time-lag kernel. If we define

$$\bar{G}(t,\theta) = G(t,\tau) \tag{6.1.22}$$

where $\theta = \tau/2$, we obtain the alternative definition

$$\rho_z(t,f) = 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{G}(u,\theta)\, z(t-u+\theta)\, z^*(t-u-\theta)\, e^{-j4\pi f\theta}\, du\, d\theta. \tag{6.1.23}$$

Theorem 6.1.3: If $\rho_z(t,f)$ is modified by ideally sampling $G(u,\tau)$ in θ at $\theta = \tau/2 = m/f_s$ and in u at $u = p/f_s$, where m and p are integers and f_s is the sampling rate, and if $G(u,\tau)$ is time-limited so that

$$G(u,\tau) = 0 \quad \text{for} \quad |\theta| = |\tfrac{\tau}{2}| \geq \tfrac{M}{2f_s} \text{ or } |u| \geq \tfrac{P}{2f_s} \tag{6.1.24}$$

where M and P are positive integers, and if the modified TFD is denoted by $\widehat{\rho}_z(t,f)$, then

$$\widehat{\rho}_z\big(\tfrac{n}{f_s}, \tfrac{kf_s}{2M}\big) = 2\sum_{|m|<\frac{M}{2}}\ \sum_{|p|<\frac{P}{2}} \bar{G}\big(\tfrac{p}{f_s}, \tfrac{m}{f_s}\big)\, z\big(\tfrac{n-p+m}{f_s}\big)\, z^*\big(\tfrac{n-p-m}{f_s}\big)\, e^{-j2\pi km/M}. \tag{6.1.25}$$

Proof/explanation: The sampled version of $G(u,\theta)$ is

$$\widehat{G}(u,\theta) = G(u,\theta)\sum_{p=-\infty}^{\infty}\delta\big(u-\tfrac{p}{f_s}\big)\sum_{m=-\infty}^{\infty}\delta\big(\theta-\tfrac{m}{f_s}\big) \tag{6.1.26}$$

$$= G(u,\theta)\sum_{p=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\delta_2\big(u-\tfrac{p}{f_s}\,,\,\theta-\tfrac{m}{f_s}\big) \tag{6.1.27}$$

where $\delta_2(u,\theta)$ is the two-dimensional unit impulse function. When $G(u,\theta)$ is replaced by $\widehat{G}(u,\theta)$, Eq. (6.1.23) becomes

$$\widehat{\rho}_z(t,f) = 2\sum_{m=-\infty}^{\infty}\sum_{p=-\infty}^{\infty} \bar{G}\big(\tfrac{p}{f_s}, \tfrac{m}{f_s}\big)\, z\big(t-\tfrac{p}{f_s}+\tfrac{m}{f_s}\big)\, z^*\big(t-\tfrac{p}{f_s}-\tfrac{m}{f_s}\big)\, e^{-j4\pi fm/f_s} \tag{6.1.28}$$

The time-limiting of $G(u,\tau)$ restricts the summation to $|m| < M/2$ and $|p| < P/2$. The sampling in τ makes $\rho_z(t,f)$ periodic in f with period $f_s/2$, while the time-limiting in τ (or m) gives a frequency resolution of M bins per period. So it is convenient to let

$$f = \frac{kf_s}{2M} \tag{6.1.29}$$

where k is an integer. With these restrictions, and with $t = n/f_s$, Eq. (6.1.28) reduces to Eq. (6.1.25). ∎

With a change of notation, Eq. (6.1.25) becomes

$$\rho_z[n,k] = 2\sum_{|m|<\frac{M}{2}}\ \sum_{|p|<\frac{P}{2}} G[p,m]\, z[n-p+m]\, z^*[n-p-m]\, e^{-j2\pi km/M} \tag{6.1.30}$$

$$= 2\sum_{|m|<\frac{M}{2}} G[n,m] \underset{n}{*} \big(z[n+m]\, z^*[n-m]\big)\, e^{-j2\pi km/M}. \tag{6.1.31}$$

This $\rho_z[n,k]$ is the generalized **discrete quadratic TFD**.

If the summand is extended in m with period M, Eq. (6.1.31) becomes

$$\rho_z[n,k] = 2 \underset{m\to k}{\mathrm{DFT}} \left\{ G[n,m] \underset{n}{*} (z[n+m]\, z^*[n-m]) \right\};\ m \in \langle M \rangle . \tag{6.1.32}$$

Because the time (n) support for $z[n+m]\, z^*[n-m]$ is the same as for $z[n]$, the time support for $\rho_z[n,k]$ is the same as for $G[n,m] \underset{n}{*} z[n]$. If the latter has a duration not exceeding N samples, then the non-zero elements of $\rho_z[n,k]$ may represented as an $N \times M$ matrix and the linear convolution in Eqs. (6.1.31) and (6.1.32) may be interpreted as modulo-N. If $G[n,m]$ is real and even in m, then the argument of the DFT is Hermitian in m, so that $\rho_z[n,k]$ is real.

Eq. (6.1.32) shows that the implementation of the discrete quadratic TFD involves construction of the discrete IAF, followed by convolution in n with the time-lag kernel, followed by discrete Fourier transformation. These steps may be simplified by taking advantage of symmetries, as explained in [1,2].

It remains to determine the effect of $G(t,\tau)$ on the required sampling rate. If $\underset{\tau\to f}{\mathcal{F}} \{G(t,\tau)\} = \gamma(t,f)$ and $\underset{t\to\nu}{\mathcal{F}} \{G(t,\tau)\} = g(\nu,\tau)$, we have the familiar results

$$\underset{\tau\to f}{\mathcal{F}} \left\{ g(t,\tau) \underset{t}{*} K_z(t,\tau) \right\} = \gamma(t,f) ** W_z(t,f) \tag{6.1.33}$$

$$\underset{t\to\nu}{\mathcal{F}} \left\{ g(t,\tau) \underset{t}{*} K_z(t,\tau) \right\} = g(\nu,\tau)\, A_z(\nu,\tau). \tag{6.1.34}$$

Comparing the above with Eqs. (6.1.11) and (6.1.12), we see that there is an additional spreading of the spectrum in f but not in ν. If $\gamma(t,f)$ is zero outside the band $|f| < B_G/2$, then the total bandwidth of the WVD is increased by B_G, so that the required sampling rate is increased by B_G in τ, or $2B_G$ in θ, and the sampling rate in t must be increased to match.

6.1.3.1 Special Cases

If $G[n,m] = \delta[n]\, g[m]$, then Eq. (6.1.32) reduces to Eq. (6.1.19), so that the discrete quadratic TFD reduces to the windowed DWVD. If, in addition, $g[m] = 1$ (that is, if $G[n,m] = \delta[n]$), then Eqs. (6.1.19) and (6.1.32) reduce to Eq. (6.1.10), so that the windowed DWVD and the discrete quadratic TFD reduce to the DWVD.

Two of the three theorems above concern the sampling of a window or kernel function. Theorem 6.1.1 seems to be an exception in that the entire IAF is sampled (which is possible only in theory, as one cannot compute a continuous IAF in practice). However, because the WVD may be considered as a windowed WVD with $g(\tau) = 1$, Theorem 6.1.1 can be restated in terms of sampling the lag window, like Theorem 6.1.2.

6.1.3.2 Doppler-Frequency Form

The Doppler-frequency form of the general quadratic TFD is

$$\rho_z(t,f) = \mathcal{F}^{-1}_{t\leftarrow\nu}\left\{\mathcal{G}(\nu,f) \underset{f}{*} [Z(f+\tfrac{\nu}{2})\, Z^*(f-\tfrac{\nu}{2})]\right\} \tag{6.1.35}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{G}(\eta,\nu)\, Z(f-\eta+\tfrac{\nu}{2})\, Z^*(f-\eta-\tfrac{\nu}{2})\, d\eta\, e^{j2\pi\nu t} d\nu \tag{6.1.36}$$

where $\mathcal{G}(\nu,f)$ is the Doppler-frequency kernel. If we define

$$\bar{\mathcal{G}}(\xi,f) = \mathcal{G}(\nu,f) \tag{6.1.37}$$

where $\xi = \nu/2$, we obtain the alternative definition

$$\rho_z(t,f) = 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{\mathcal{G}}(\xi,\eta)\, Z(f-\eta+\xi)\, Z^*(f-\eta-\xi)\, e^{j4\pi\xi t} d\eta\, d\xi. \tag{6.1.38}$$

As the time-lag definition of the quadratic TFD leads to Theorem 6.1.3, so the Doppler-frequency definition leads to the following result.

Theorem 6.1.4: If $\rho_z(t,f)$ is modified by ideally sampling $\mathcal{G}(\eta,\nu)$ in ξ at $\xi = \nu/2 = \frac{lf_s}{2N}$ and in η at $\eta = \frac{qf_s}{2N}$, where l and q are integers and N is a positive integer and f_s is a positive constant, and if $\mathcal{G}(\eta,\nu)$ is band-limited so that

$$\mathcal{G}(\eta,\nu) = 0 \quad \text{for} \quad |\xi| = |\tfrac{\nu}{2}| \geq \tfrac{Lf_s}{4N} \text{ or } |\eta| \geq \tfrac{Qf_s}{4N} \tag{6.1.39}$$

where L and Q are positive integers, and if the modified TFD is denoted by $\widehat{\rho}_z(t,f)$, then

$$\widehat{\rho}_z\left(\tfrac{n}{f_s}, \tfrac{kf_s}{2N}\right) = 2\sum_{|l|<\frac{L}{2}} \sum_{|q|<\frac{Q}{2}} \bar{\mathcal{G}}\left(\tfrac{lf_s}{2N}, \tfrac{qf_s}{2N}\right) Z\left(\tfrac{[k-q+l]f_s}{2N}\right) Z^*\left(\tfrac{[k-q-l]f_s}{2N}\right) e^{j2\pi ln/N}. \tag{6.1.40}$$

Proof: Parallel to the proof of Theorem 6.1.3 and the subsequent discussion. ∎

With a change of notation, Eq. (6.1.40) becomes

$$\rho_z[n,k] = 2\sum_{|l|<\frac{L}{2}} \sum_{|q|<\frac{Q}{2}} \mathcal{G}[l,q]\, Z[k-q+l]\, Z^*[k-q-l]\, e^{j2\pi ln/N} \tag{6.1.41}$$

$$= 2\sum_{|l|<\frac{L}{2}} \mathcal{G}[l,k] \underset{k}{*} (Z[k+l]\, Z^*[k-l])\, e^{j2\pi ln/N}. \tag{6.1.42}$$

If the summand is extended in l with period L, Eq. (6.1.42) becomes

$$\rho_z[n,k] = 2\ \underset{n\leftarrow l}{\mathrm{IDFT}}\left\{\mathcal{G}[l,k] \underset{k}{*} (Z[k+l]\, Z^*[k-l])\right\};\ l\in\langle L\rangle\,. \tag{6.1.43}$$

The frequency support for $\rho_z[n,k]$ is the same as for $G[l,k] \underset{k}{*} Z[k]$. If the latter support has a width not exceeding K frequency samples, then the non-zero elements of $\rho_z[n,k]$ may be represented as an $L\times K$ matrix and the linear convolution in Eqs. (6.1.42) and (6.1.43) may be interpreted as modulo-K. While the dimensions of the TFD matrix seem to differ from those in Theorem 6.1.3, the dimensions are upper bounds and may be matched, if desired, by zero-padding.

Table 6.1.1: Kernel requirements for selected properties ("Prop.") of quadratic TFDs, in time-lag and Doppler-frequency domains, for general continuous and discrete TFDs, and discrete TFDs with Doppler-independent (DI) and lag-independent (LI) kernels. Note that $\theta = \tau/2$ and $\xi = \nu/2$.

Prop.	KERNEL CONSTRAINTS			
	Continuous	Discrete	Discrete DI $G[n,m]=\delta[n]g_2[m]$ $\mathcal{G}[l,k]=G_2[k]$	Discrete LI $G[n,m]=g_1[n]$ $\mathcal{G}[l,k]=G_1[l]\delta[k]$
RE:	$G(t,\tau)$ $= G^*(t,-\tau)$.	$G[n,m]$ $= G^*[n,-m]$.	$g_2[m] = g_2^*[-m]$	$g_1[n]$ is real.
TM:	$G(t,0) = \delta(t)$	$G[n,0] = \delta[n]$	$g_2[0] = 1$	WVD only
FM:	$\mathcal{G}(0,f) = \delta(f)$	$\mathcal{G}[0,k] = \delta[k]$	WVD only	$G_1[0] = 1$
IF:	$G(t,0) = \delta(t)$; $\int f\mathcal{G}(\nu,f)\,df = 0$.	$G[n,0] = \delta[n]$; $\sum_k k\mathcal{G}[l,k] = 0$.	$g_2[0] = 1$; $\sum_k kG_2[k] = 0$.	WVD only
TD:	$\mathcal{G}(0,f) = \delta(f)$; $\int tG(t,\tau)\,dt = 0$.	$\mathcal{G}[0,k] = \delta[k]$; $\sum_n nG[n,m] = 0$.	WVD only	$G_1[0] = 1$; $\sum_n ng_1[n] = 0$.
TS:	$G(t,\tau) = 0$ if $\lvert\theta\rvert < \lvert t\rvert$.	$G[n,m] = 0$ if $\lvert m\rvert < \lvert n\rvert$.	Always	WVD only
FS:	$\mathcal{G}(\nu,f) = 0$ if $\lvert\xi\rvert < \lvert f\rvert$.	$\mathcal{G}[l,k] = 0$ if $\lvert l\rvert < \lvert k\rvert$.	WVD only	Always

6.1.4 Desirable Properties; Kernel Constraints

The desirable properties of continuous TFDs (defined in Section 3.1.1, p. 60 ff) are easily redefined for discrete TFDs. Some important examples are given below.

Realness (RE) says simply that the TFD is real.

The **marginal conditions**, which may be considered optional for signal-processing purposes, are the **time marginal (TM)**

$$\sum_k \rho_z[n,k] = |z[n]|^2 \tag{6.1.44}$$

and the **frequency marginal (FM)**

$$\sum_n \rho_z[n,k] = |Z[k]|^2. \tag{6.1.45}$$

The **IF property** says that the periodic first moment of the TFD w.r.t. frequency is the instantaneous frequency. Its dual, which seems to be regarded as less important, is the **time delay property (TD)**, and says that the periodic first moment of the TFD w.r.t. time is the time delay.

The **time support (TS)** property says that if $z[n] = 0$ everywhere except $n_1 \le n \le n_2$, then $\rho_z[n,k] = 0$ everywhere except $n_1 \le n \le n_2$. Similarly, the

Table 6.1.2: Kernels of selected TFDs in time-lag and Doppler-frequency domains. For the spectrogram and windowed Levin (w-Levin) distributions, the window $w[n]$ is assumed to be real and even. $W[k]$ denotes the sampled spectrum of the window $w(\tau)$. In the PROPERTY column, an exclamation (!) means that the property is always satisfied, while an asterisk (*) means that the property is satisfied subject to non-degenerate constraints on the window and/or parameter.

	KERNEL		PROPERTY						
Distribution	$G[n,m]$	$\mathcal{G}[l,k]$	RE	TM	FM	IF	TD	TS	FS
WVD	$\delta[n]$	$\delta[k]$	!	!	!	!	!	!	!
Levin	$\frac{1}{2}\delta[n+m]$ $+\frac{1}{2}\delta[n-m]$	$\frac{1}{2}\delta[k+l]$ $+\frac{1}{2}\delta[k-l]$	!	!	!	!	!	!	!
Born-Jordan	$\left[\frac{1}{4\lvert\alpha m\rvert}\,\mathrm{rect}\left(\frac{n}{4\alpha m}\right)\right]$ $**\,[\mathrm{sinc}\,n\,\mathrm{sinc}\,m]$	$\left[\frac{1}{4\lvert\alpha l\rvert}\,\mathrm{rect}\left(\frac{k}{4\alpha l}\right)\right]$ $**\,[\mathrm{sinc}\,k\,\mathrm{sinc}\,l]$	!	!	!	!	!	*	*
Modified B	$\frac{\cosh^{-2\beta}n}{\sum_n \cosh^{-2\beta}n}$		!		!		!		!
windowed WVD	$\delta[n]\,w[m]$	$W[k]$	*	*		*		!	
w-Levin	$\frac{1}{2}w[m]\,\delta[n+m]$ $+\frac{1}{2}w[m]\,\delta[n-m]$	$\frac{1}{2}W[k+l]$ $+\frac{1}{2}W[k-l]$	!	*		*		!	
Zhao-Atlas-Marks	$[w[m]\,\mathrm{rect}\left(\frac{an}{4m}\right)]$ $**\,[\mathrm{sinc}\,n\,\mathrm{sinc}\,m]$		*	*		*		*	
Rihaczek	$\delta[n-m]$	$\delta[k+l]$		!	!			!	!
w-Rihaczek	$w^*[-m]\,\delta[n-m]$	$W^*[k+l]$		*				!	
Page	$\delta[n-\lvert m\rvert]$	$\frac{1}{2}\delta[k+l]+\frac{1}{2}\delta[k-l]$ $+\frac{jl}{\pi(k^2-l^2)} * \mathrm{sinc}\,l$	!	!	!	!		!	
Choi-Williams	$\frac{\sqrt{\pi\sigma}}{2\lvert m\rvert}\exp\left(\frac{-\pi^2\sigma n^2}{4m^2}\right)$ $**\,[\mathrm{sinc}\,n\,\mathrm{sinc}\,m]$	$\frac{\sqrt{\pi\sigma}}{2\lvert l\rvert}\exp\left(\frac{-\pi^2\sigma k^2}{4l^2}\right)$ $**\,[\mathrm{sinc}\,k\,\mathrm{sinc}\,l]$	!	!	!	!	!		
B	$\left(\frac{\lvert 2m\rvert}{\cosh^2 n}\right)^{\beta} * \mathrm{sinc}\,m$		!						
spectrogram	$w[n+m]\,w[n-m]$	$W[k+l]\,W[k-l]$	!						

frequency support (FS) property says that if $Z[k]=0$ everywhere except $k_1 \le k \le k_2$, then $\rho_z[n,k]=0$ everywhere except $k_1 \le k \le k_2$.

The class of TFDs satisfying realness, the time and frequency marginals, the IF property and the time and frequency support properties is called class $\mathcal{P}$.

The properties listed above are selected from Table 3.3.1 on p. 75. For each property, Table 3.3.1 gives necessary and sufficient conditions on the kernel of the general continuous quadratic TFD. To obtain the corresponding conditions for discrete TFDs, we first express the conditions entirely in the time-lag and Doppler-frequency domains, using the scaled variables θ and ξ where convenient, obtaining the "Continuous" column of Table 6.1.1. Then we sample the kernels as specified

in Theorems 6.1.3 and 6.1.4, obtaining the "Discrete" column of Table 6.1.1. The remaining columns are obtained by specialization.

The sampling of the time-lag kernel will be free of aliasing if the kernel is first band-limited in t and θ to $\pm f_s/2$. This causes $\delta(t)$ to be discretized as $\delta[n]$. Similarly, the sampling of the Doppler-frequency kernel will be free of aliasing if the kernel is first time-limited in f and ξ to $\pm N/f_s$. This causes $\delta(f)$ to be discretized as $\delta[k]$.

6.1.5 Examples

By sampling the kernels of common continuous quadratic TFDs [see Table 3.3.2 on p. 76], we obtain the new Table 6.1.2, which lists the kernels for the discrete versions of those TFDs. The convolutions with sinc functions are performed *before* restricting the variable to integer values; this requires oversampling and is computationally inefficient. The convolutions in n and m arise from the band-limiting of the time-lag kernel prior to sampling. This band-limiting of the kernel does *not* affect the result of convolving the kernel with the IAF, because the IAF is assumed to be similarly band-limited. Similarly, the convolutions in k and l arise from the time-limiting of the Doppler-Frequency kernel prior to sampling. Where no convolution appears in the kernel, either a sinc function has been converted to a discrete delta function by the sampling, or a window function is assumed to provide sufficient filtering.

6.1.6 Summary and Conclusions

Ideal sampling of window or kernel functions leads to straightforward definitions of discrete-time forms of the WVD, the windowed WVD and other quadratic TFDs. Use of the analytic signal minimizes the required sampling rate.

Further theoretical details may be found in [3–5]. Some practical computational issues will be examined in Article 6.5.

References

[1] B. Boashash and P. J. Black, "An efficient real-time implementation of the Wigner-Ville distribution," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 35, pp. 1611–1618, November 1987.

[2] B. Boashash and A. Reilly, "Algorithms for time-frequency signal analysis," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 7, pp. 163–181, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[3] J. C. O'Neill and W. J. Williams, "Shift-covariant time-frequency distributions of discrete signals," *IEEE Trans. Signal Processing*, vol. 47, pp. 133–146, January 1999.

[4] M. S. Richman, T. W. Parks, and R. G. Shenoy, "Discrete-time, discrete-frequency time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 46, pp. 1517–1527, June 1998.

[5] A. H. Costa and G. F. Boudreax-Bartels, "An overview of aliasing errors in discrete-time formulations of time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 47, pp. 1463–1474, May 1999.

6.2 QUADRATIC AND HIGHER ORDER TIME-FREQUENCY ANALYSIS BASED ON THE STFT [0]

The oldest, simplest, and most commonly used tool for time-frequency (TF) analysis of a signal $x(t)$ is the spectrogram, defined as the squared magnitude of the short time Fourier transform (STFT) [1]. The STFT is

$$F_x(t,f) = \int_{-\infty}^{\infty} x(t+\tau)w(\tau)e^{-j2\pi f\tau}d\tau, \tag{6.2.1}$$

where $w(\tau)$ is a real-valued even lag window [cf. Section 2.3.1]. Implementations (hardware and software) of this transform are already widely present in practice. The STFT is linear and very simple for realization. However, it has some serious drawbacks. The most important one lies in its low concentration in the TF plane, when highly nonstationary signals are analyzed. In order to improve TF representation, various quadratic distributions have been introduced. The most important member of this class is the pseudo Wigner distribution (WD)

$$W_x(t,f) = \int_{-\infty}^{\infty} w(\frac{\tau}{2})w(-\frac{\tau}{2})x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau. \tag{6.2.2}$$

The WD itself has a drawback. Namely, in the case of multicomponent signals, $x(t) = \sum_{p=1}^{P} x_p(t)$, it produces emphatic cross-terms that can completely mask the auto-terms and make this distribution useless for analysis. This is why many other quadratic reduced interference distributions have been introduced (Choi-Williams, Zhao-Atlas-Marks, Born-Jordan, Butterworth, Zhang-Sato...) [1], [Article 6.4]. The cross-terms reduction in these distributions is based on a kind of the Wigner distribution smoothing, which inherently leads to the auto-terms degradation [2]. In contrast to these TF representations, which are focused on the preservation of marginal properties and the cross-terms reduction, the S-method (SM), which is the topic of this article, is derived with the primary goal to preserve the auto-terms quality as in the WD, while avoiding (reducing) the cross-terms. The software and hardware realization of this method is very efficient, since it is completely based on the STFT. The SM can, in a straightforward manner, be extended to the cross-terms free (reduced) realization of the higher order TF representations, time-scale representations, and multidimensional space/spatial-frequency representations.

6.2.1 STFT Based Realization of the Quadratic Representations

6.2.1.1 Basic S-Method Form

Relation between the STFT and the WD, [3],

$$W_x(t,f) = 2\int_{-\infty}^{\infty} F_x(t,f+\theta)F_x^*(t,f-\theta)d\theta \tag{6.2.3}$$

[0]Author: **LJubiša Stanković**, Elektrotehnicki fakultet, University of Montenegro, 81000 Podgorica, Montenegro (l.stankovic@ieee.org). Reviewers: J. F. Böhme, S. Carstens-Behrens and B. Ristic.

has led to the definition of a TF representation, referred to as the S-method (SM),

$$\mathrm{SM}_x(t,f) = 2\int_{-\infty}^{\infty} P(\theta)F_x(t,f+\theta)F_x^*(t,f-\theta)d\theta. \tag{6.2.4}$$

The special cases of the SM are two most important TF distributions: (1) For $P(\theta) = 1$ the WD follows, $\mathrm{SM}_x(t,f) = W_x(t,f)$, and (2) For $P(\theta) = \delta(\theta)/2$, the spectrogram $\mathrm{SM}_x(t,f) = |F_x(t,f)|^2 = S_x(t,f)$ is obtained. By changing the width of window $P(\theta)$, denoted by $2L_P$ ($P(\theta) = 0$ for $|\ \theta\ | > L_P$), between these two extreme cases we can get a gradual transition from the spectrogram to the WD. The best choice of L_P would be the value when $P(\theta)$ is wide enough to enable complete integration over the auto terms, but narrower than the distance between the auto-terms, in order to avoid the cross terms, Fig. 6.2.1. Then, *the SM produces the sum of the WDs of individual signal components, avoiding cross-terms.*

Proposition: Consider the signal $x(t) = \sum_{p=1}^{P} x_p(t)$, where $x_p(t)$ are monocomponent signals. Assume that the STFT of each component lies inside the region $D_p(t,f)$, $p = 1,2,...,P$. Denote the length of the p-th region along f, for a given t, by $2B_p(t)$, and its central frequency by $f_{0p}(t)$. The SM of $x(t)$ produces the sum of the WDs $W_{x_p}(t,f)$ of each signal component $x_p(t)$,

$$\mathrm{SM}_x(t,f) = \sum_{p=1}^{P} W_{x_p}(t,f), \tag{6.2.5}$$

if the regions $D_p(t,f)$, $p = 1,2,...,P$, do not overlap, $D_p(t,f) \cap D_q(t,f) = \emptyset$ for $p \neq q$ (meaning cross-terms free spectrogram), and if the width of the rectangular window $P(\theta)$, for a point (t,f), is defined by $L_P(t,f) = B_p(t) - |f - f_{0p}(t)|$ for $(t,f) \in D_p(t,f)$, and 0 elsewhere.

Proof: Consider a point (t,f) inside a region $D_p(t,f)$. The integration interval in (6.2.4), for the p-th signal component is symmetrical with respect to $\theta = 0$. It is defined by the smallest absolute value of θ for which $f+\theta$ or $f-\theta$ falls outside $D_p(t,f)$, i.e., $|f+\theta-f_{0p}(t)| \geq B_p(t)$ or $|f-\theta-f_{0p}(t)| \geq B_p(t)$. For $f > f_{0p}(t)$ and positive θ, the integration limit is reached first in $|f+\theta-f_{0p}(t)| \geq B_p(t)$ for $\theta = B_p(t) - (f - f_{0p}(t))$. For $f < f_{0p}(t)$ and positive θ, the limit is reached first in $|f-\theta-f_{0p}(t)| \geq B_p(t)$ for $\theta = B_p(t) - (f_{0p}(t) - f)$. Thus, having in mind the interval symmetry, an integration limit which produces the same value of integral (6.2.4) as the value of (6.2.3), over the region $D_p(t,f)$, is given by $L_P(t,f)$ in the Proposition. Therefore, for $(t,f) \in D_p(t,f)$ we have $\mathrm{SM}_x(t,f) = W_{x_p}(t,f)$. Since $L_P(t,f) = 0$ for $(t,f) \notin D_p(t,f)$, $p = 1,2,...,P$.

Note: Any window $P(\theta)$ with constant width $L_P \geq \max_{(t,f)}\{L_P(t,f)\}$ produces $\mathrm{SM}_x(t,f) = \sum_{p=1}^{P} W_{x_p}(t,f)$, if the regions $D_p(t,f)$, $p = 1,2,..,P$, are at least $2L_P$ apart along the frequency axis, i.e., $|f_{0p}(t) - f_{0q}(t)| > B_p(t) + B_q(t) + 2L_P$, for each p, q and t. This is the SM with constant window width (6.2.4). *If two components overlap for some time instants t, then the cross-term will appear, but only between these two components and for that time instants.*

The SM belongs to the general class of quadratic TF distributions, whose inner product form reads

$$\rho_x(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{G}(t_1,t_2)\left[x(t+t_1)e^{-j2\pi f t_1}\right]\left[x(t+t_2)e^{-j2\pi f t_2}\right]^* dt_1\, dt_2. \tag{6.2.6}$$

If the inner product kernel $\tilde{G}(t_1,t_2)$ is factorized in the Hankel form $\tilde{G}(t_1,t_2) = 2w(t_1)p(t_1+t_2)w(t_2)$, then by substituting its value into (6.2.6), with $P(-f) = \mathcal{F}_{t\to f}\{p(t)\}$, we get (6.2.4). Note that the Toeplitz factorization of the kernel $\tilde{G}(t_1,t_2) = 2w(t_1)p(t_1-t_2)w(t_2)$ results in the smoothed spectrogram. The

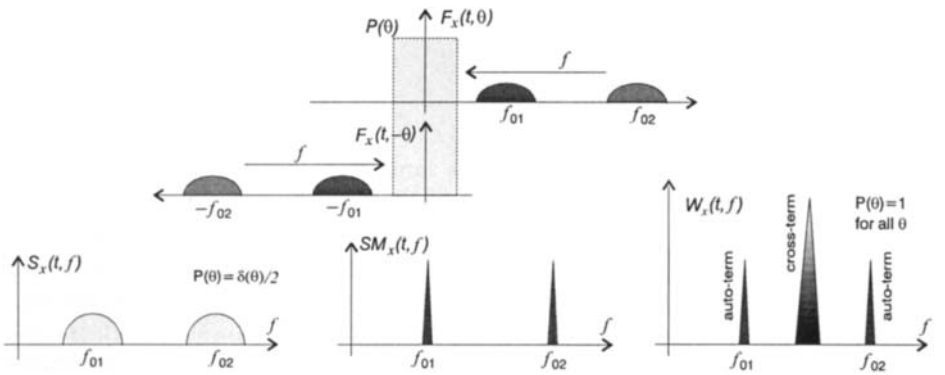


Fig. 6.2.1: Illustration of the SM calculation including two special cases: the WD and the spectrogram

smoothed spectrogram composes two STFTs in the same direction, resulting in the distribution spread, in contrast to the SM, where two STFTs are composed in counterdirection, resulting in the concentration improvement (Fig. 6.2.1; reference [4]; Article 9.1).

The SM kernel in Doppler-lag domain is given by $c(\nu,\tau) = P(\nu/2) *_\nu A_{ww}(\nu,\tau)$, where $A_{ww}(\nu,\tau)$ is the ambiguity function of $w(\tau)$, and $*_\nu$ denotes a convolution in ν. Generally, this kernel is not a separable function.

6.2.1.2 Other Forms of the S-Method in Quadratic Representations

Time direction form of the SM is

$$\mathrm{SM}_x(t,f) = 2\int_{-\infty}^{\infty} P(\tau)F_x(t+\tau,f)F_x^*(t-\tau,f)e^{-j4\pi f\tau}d\tau. \tag{6.2.7}$$

It results from the same analysis as (6.2.4), based on the frequency domain windowed WD, $W_x(t,f) = \int_{-\infty}^{\infty} W(\theta/2)W(-\theta/2)X(f+\theta/2)X^*(f-\theta/2)\exp(j2\pi t\theta)d\theta$.

Fractional domain form: The frequency and time direction forms of the SM can be generalized to any direction in the time-frequency plane. Consider the fractional FT of $x(t)$, denoted by $X^\alpha(u)$ [Articles 4.8, 5.8]. Its STFT is

$$F_x^\alpha(u,v) = \int_{-\infty}^{\infty} X^\alpha(u+\tau)h(\tau)\exp(-j2\pi v\tau)d\tau. \tag{6.2.8}$$

where $h(\tau)$ is the lag window. The SM in the fractional domain, is defined by

$$\mathrm{SM}_x^\alpha(u,v) = 2\int_{-\infty}^{\infty} P(\theta)F_x^\alpha(u,v+\theta)F_x^{\alpha *}(u,v-\theta)d\theta, \tag{6.2.9}$$

It can be easily realized based on the signal's fractional FT and (6.2.8).

Using the STFT rotational property, $F_x^{\alpha}(u,v)\exp(-j\pi uv) = F_x^0(t,f)\exp(-j\pi tf)$ with $u = t\cos\alpha + f\sin\alpha$, $v = -t\sin\alpha + f\cos\alpha$, [5], we can rewrite (6.2.9) as

$$\begin{aligned}\mathrm{SM}_x^{\alpha}(t,f) &= 2\int_{-\infty}^{\infty} P(\theta)F_x^0(t-\theta\sin\alpha, f+\theta\cos\alpha) \\ &\quad \times F_x^{0*}(t+\theta\sin\alpha, f-\theta\cos\alpha)e^{j4\pi f\theta\sin\alpha}d\theta, \end{aligned} \tag{6.2.10}$$

For $\alpha = 0$ it gives (6.2.4), while (6.2.7) follows for $\alpha = -\pi/2$. For the derivation of (6.2.10) the lag window $h(\tau)$ is formally assumed as $(W^{\alpha}(-\tau))^*$. Optimal direction for the fractional SM calculation can be obtained based on the fractional moments analysis in [Article 4.8]. It has been used in [5].

Affine SM form: Continuous wavelet transform (WT) is defined by $D_x(t,f) = \int_{-\infty}^{\infty} x(\tau)h^*((\tau-t)f/f_0)d\tau/\sqrt{|f_0/f|}$. As in [6] we used frequency instead of scale $a = f_0/f$. Consider $h(t)$ in the form $h(t) = w(t)\exp(j2\pi f_0 t)$ which provides a strong formal connection of the WT with the STFT. The pseudo affine Wigner distribution is defined by

$$W_x^a(t,f) = \int_{-\infty}^{\infty} w(\frac{\tau}{2f_0}f)w(-\frac{\tau}{2f_0}f)x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})e^{-j2\pi\tau f}d\tau. \tag{6.2.11}$$

The affine SM form reads:

$$\mathrm{SM}_x^a(t,f) = 2\int_{-\infty}^{\infty} P(\theta)D_x(t,f;f_0+\theta)D_x^*(t,f;f_0-\theta)d\theta, \tag{6.2.12}$$

where $D_x(t,f;f_0+\theta)$ is the WT calculated with $h(t) = w(t)\exp(j2\pi(f_0+\theta)t)$. If $P(\theta) = \delta(\theta)/2$, then $\mathrm{SM}_x^a(t,f)$ is equal to the scalogram of $x(t)$, while for $P(\theta) = 1$ it produces $W_x^a(t,f)$ defined by (6.2.11). This form of the SM has been extended to other time-scale representations in [6].

6.2.2 Discrete Realization of the Basic S-Method Form

Discrete SM, for a rectangular window $P(\theta)$, follows from (6.2.4)

$$\mathrm{SM}_x(n,k) = \frac{2}{T_w}[|F_x(n,k)|^2 + 2\mathrm{Re}\{\sum_{i=1}^{L_P} F_x(n,k+i)F_x^*(n,k-i)\}] \tag{6.2.13}$$

where: $F_x(n,k) = DFT_{i\to k}\{x(n+i)w(i)\Delta t\}$, Δt is the sampling interval, $T_w = N\Delta t$ is the width of $w(\tau)$, and $2L_P+1$ is the width of $P(\theta)$ in the discrete domain. For notation simplicity we will assume normalized $2/T_w = 1$. Recursive relation for the SM calculation is

$$\mathrm{SM}_x(n,k;L_P) = \mathrm{SM}_x(n,k;L_P-1) + 2\mathrm{Re}\{F_x(n,k+L_P)F_x^*(n,k-L_P)\} \tag{6.2.14}$$

where $\mathrm{SM}_x(n,k;0) = |F_x(n,k)|^2$, and $\mathrm{SM}_x(n,k;L_P)$ denotes $\mathrm{SM}_x(n,k)$ in (6.2.13) calculated with L_P. In this way we start from the spectrogram, and gradually make

the transition toward the WD. The calculation in (6.2.13) and (6.2.14) needn't be done for each point (n,k) separately. It can be performed for the whole matrix of the SM and the STFT. This can significantly save time in some matrix based calculation tools. In the SM calculation: (1) There is no need for analytic signal calculation since the cross-terms between negative and positive frequency components are removed in the same way as are the other cross-terms [7]. (2) If we take that $F_x(n,k) = 0$ outside the basic period, i.e., when $k < -N/2$ or $k > N/2 - 1$, then there is no aliasing when the STFT is alias-free (in this way we can calculate the alias-free WD by taking $L_P = N/2$ in (6.2.13)).

For the SM realization we have to implement the STFT first, based either on the FFT routines or recursive approaches suitable for hardware realizations [3,7]. After we get the STFT we have to "correct" the obtained values, according to (6.2.13), by adding few terms $2\mathrm{Re}\{F_x(n,k+i)F_x^*(n,k-i)\}$ to the SPEC values.

There are two possibilities to implement the summation in (6.2.13):
(1) With a signal independent L_P. Theoretically, in order to get the WD for each individual component, the length L_P should be such that $2L_P$ is equal to the width of the widest auto-term. This will guarantee cross-terms free distribution for all components which are at least $2L_P$ samples apart. For components and time instants where this condition is not satisfied, the cross-terms will appear, but still in a reduced form (see also [Article 7.3]).
(2) With a signal dependent $L_P = L_P(n,k)$ where the summation, for each point (n,k), lasts until the absolute square value of $F_x(n,k+i)$ or $F_x(n,k-i)$ is smaller than an assumed reference level R. If a zero value may be expected within a single auto-term, then the summation lasts until two subsequent zero values of $F_x(n,k+i)$ or $F_x(n,k-i)$ are detected. The reference level is defined as a few percent of the spectrogram's maximal value at a considered instant n, $R_n = \max_k\{S_x(n,k)\}/Q^2$, where Q is a constant. Index n is added to show that the reference level R is time dependent. Note that if $Q^2 \to \infty$, the WD will be obtained, while $Q^2 = 1$ results in the spectrogram. A choice of an appropriate value for design parameter Q^2 will be discussed in *Example 2.*

Example 1: Consider a real-valued multicomponent signal

$$x(t) = \cos(1200(t+0.1)^2) + e^{-36(t-1/3)^2}\cos(1200(t+1/2)^2) + e^{-36(t-2/3)^2}\cos(1200(t-1/3)^2) + \cos(960\pi t)$$

within the interval $[0,1]$, sampled at $\Delta t = 1/1024$. This sampling rate is very close to the Nyquist rate for this signal, that is $1/960$. The Hanning window of the width $T_w = 1/4$ is used. The spectrogram is shown in Fig. 6.2.2(a). Its "corrected" version (the SM), according to (6.2.13), with five terms, $L_P = 5$, is shown in Fig. 6.2.2(c). The auto-terms are concentrated almost as in the WD, Fig. 6.2.2(b). The Choi-Williams distribution (CWD), whose kernel reads $c(\nu,\tau) = \exp(-(\nu\tau)^2)$, is shown in Fig. 6.2.2(d). Normalized values $-\sqrt{\pi N/2} \le |2\pi\nu| \le \sqrt{\pi N/2}$, $-\sqrt{\pi N/2} \le |\tau| \le \sqrt{\pi N/2}$, and 128 samples within that interval, are used. If the analytic part of $x(t)$ were used, similar results would be obtained [see Fig. 9.1.1 on p. 379].

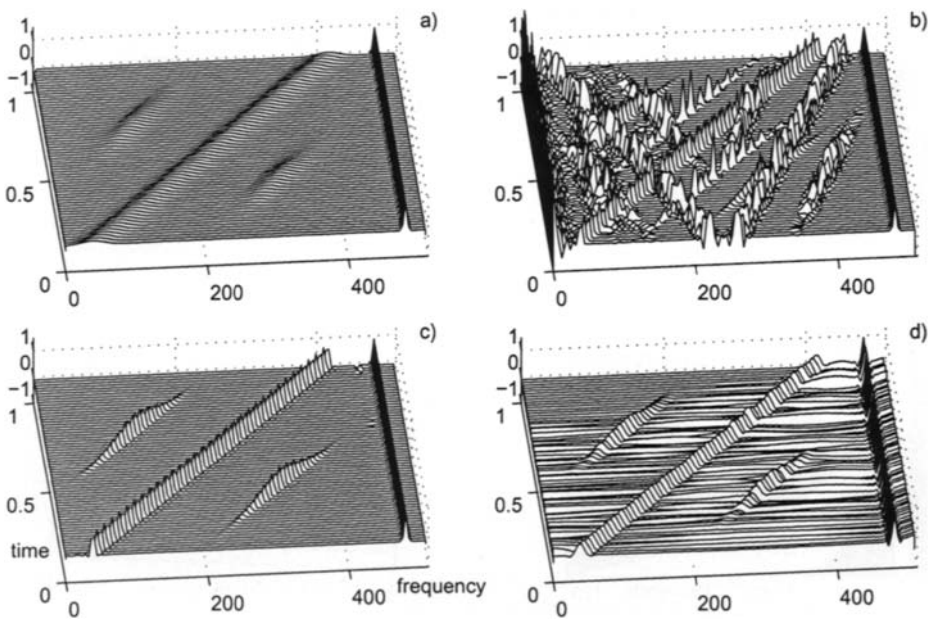


Fig. 6.2.2: Time-frequency representation of a real-valued multicomponent signal: (a) Spectrogram, (b) Pseudo Wigner distribution, (c) S-method with five "correcting terms ", $L_P = 5$, (d) Choi-Williams distribution, as the representative of reduced interference distributions.

Example 2: The adaptive SM realization will be illustrated on a three-component real signal, with a nonlinear FM component,

$$x(t) = e^{-t^2}\cos(25\pi t) + \cos(120t^3 + 45\pi t) + 1.5e^{-25t^2}\cos(40\pi t^2 + 150\pi t)$$

with the sampling interval $\Delta t = 1/256$. The signal is considered within the time interval $[-1, 1]$. The Hanning window of the width $T_w = 1$ is used. The spectrogram is presented in Fig. 6.2.3(a), while the SM with the constant $L_P = 3$ is shown in Fig. 6.2.3(b). The concentration improvement with respect to the case $L_P = 0$, Fig. 6.2.3(a), is evident. Further increase of L_P would improve concentration, but it would also cause that some cross-terms appear. Some small changes are already noticeable between the components with quadratic and constant IF. An improved concentration, without cross-terms, can be achieved by using the variable window width L_P. The regions $D_i(n,k)$, determining the summation limit $L_P(n,k)$ for each point (n,k), are obtained by imposing the reference level R_n corresponding to $Q^2 = 50$. They are defined as: $D_i(n,k) = 1$ when $S_x(n,k) \geq R_n = \max_k\{S_x(n,k)\}/Q^2$, and $D_i(n,k) = 0$ elsewhere, Fig. 6.2.3(c). White regions mean that the value of spectrogram is below 2% of its maximal value at that time instant n, meaning that the concentration improvement is not performed at these points. The signal dependent SM is given in Fig. 6.2.3(d). The method sensitivity, with respect to the

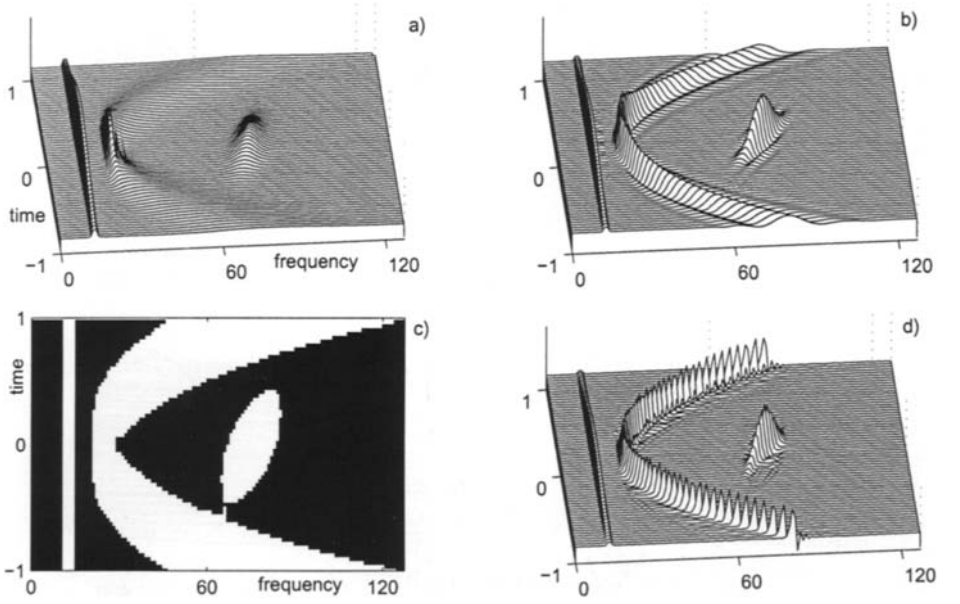


Fig. 6.2.3: Time-frequency analysis of a multicomponent signal: (a) Spectrogram, (b) The S-method with a constant window, with $L_P = 3$, (c) Regions of support for the S-method with a variable window width calculation, corresponding to $Q^2 = 50$, (d) The S-method with the variable window width calculated using regions in (c).

value of Q^2, is low.

6.2.3 STFT Based Realization of Higher Order Representations

In order to improve distribution concentration in the case of nonlinear FM signals, the higher order time-varying spectra have been defined (Wigner higher order spectra, Multitime Wigner distributions). For practical realizations the most interesting are the versions of these spectra reduced to the two-dimensional TF plane [8]. Here, we will present the L-Wigner distribution (LWD) and the fourth order polynomial Wigner-Ville distribution (PWVD).

6.2.3.1 The L-Wigner Distribution

The L-Wigner Distribution (LWD) is defined by [7,9]

$$LD_L(t,f) = \int_{-\infty}^{\infty} w_L(\tau) x^{*L}(t - \frac{\tau}{2L}) x^L(t + \frac{\tau}{2L}) e^{-j2\pi f\tau} d\tau. \qquad (6.2.15)$$

For $L = 1$ it reduces to the WD.

The LWD is a windowed slice of the multitime Wigner distributions:

$$W_x^K(t_1,...,t_k,f) = \int_{-\infty}^{\infty} x^*(\sum_{i=1}^{K} t_i + \frac{\tau}{K+1}) \prod_{i=1}^{L-1} x^*(-t_i + \frac{\tau}{K+1}) \prod_{i=L}^{K} x(t_i - \frac{\tau}{K+1}) e^{j2\pi f\tau} d\tau,$$

along the line $t_1 = t_2 = ... = t_{L-1} = -t$, $t_L = t_{L+1} = ... = t_K = t$, where the auto-terms in $W_x^K(t_1,...,t_k,f)$ are located, for $L = (K+1)/2$ [9].

Similarly, starting from the Wigner higher order spectra, dual to $W_x^K(t_1,t_2,...,t_k,f)$, we get a distribution dual to (6.2.15),

$$LW_L^f(t,f) = \int_{-\infty}^{\infty} W_L(\theta) X^{*L}(f + \frac{\theta}{2L}) X^L(f - \frac{\theta}{2L}) e^{-j2\pi\theta t} d\theta. \tag{6.2.16}$$

studied in details in [10]. Its realization is formally the same as for the time domain LWD.

For a frequency modulated signal $x(t) = \exp(j\phi(t))$, the LWD produces [7]

$$LW_L(t,f) = W_L(f - \phi'(t)/2\pi) *_f FT\{e^{(j(\phi^{(3)}(t+\tau_1)+\phi^{(3)}(t-\tau_2))/(48L^2))}\},$$

where τ_1, τ_2 are the values of τ within the lag window $w_L(\tau)$, and $W_L(f) = \mathcal{F}_{\tau\to f}\{w_L(\tau)\}$. For $L \to \infty$, the LWD tends to a distribution completely concentrated along the IF, i.e., $LW_L(t,f) \to W_L(f - \phi'(t)/2\pi)$.

The relationship between $LW_{2L}(t,f)$ and $LW_L(t,f)$ is of form (6.2.3),

$$LW_{2L}(t,f) = 2\int_{-\infty}^{\infty} LW_L(t,f+\theta) LW_L(t,f-\theta) d\theta$$

The realization of cross-terms and alias free version of the LWD may be efficiently done in the discrete domain, by using the SM form (6.2.13), as:

$$LW_{2L}(n,k) = LW_L^2(n,k) + 2\sum_{i=1}^{L_P} LW_L(n,k+i) LW_L(n,k-i), \tag{6.2.17}$$

with $LW_1(n,k) = W_x(n,k)$, and $W_x(n,k)$ calculated according to (6.2.13). Form (6.2.17) is very convenient for software and hardware realizations since the same blocks, connected in cascade, can provide a simple and efficient system for higher order TF analysis, based on the STFT in the initial step, and the signal sampled at the Nyquist rate. Numerical examples and illustrations of the LWD can be found in [7,9–11].

6.2.3.2 Polynomial Wigner-Ville Distribution

Modification of the presented method for the realization of the PWVD is straightforward. The fourth order PWVD can be written in a frequency scaled form [12]

$$PW_x(t,f) = \frac{1}{2.7}\int_{-\infty}^{\infty} x^2(t+\frac{\tau}{4}) x^{*2}(t-\frac{\tau}{4}) x^*(t+A\frac{\tau}{2}) x(t-A\frac{\tau}{2}) e^{-j\frac{2\pi f}{2.7}\tau} d\tau, \tag{6.2.18}$$

where $A = 0.85/1.35$ and $f' = f/2.7$. Note that $PW_x(t,f') = \frac{1}{2.7} LW_2(t,f') *_{f'} W_x^A(t,f')$, where $W_x^A(t,f') = FT\{x^*(t+A\frac{\tau}{2}) x(t-A\frac{\tau}{2})\}$ is the scaled and reversed

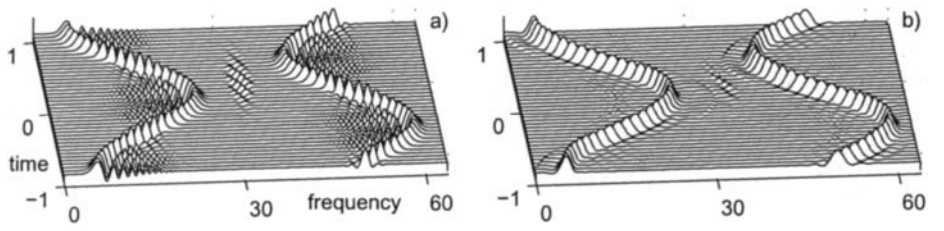


Fig. 6.2.4: Time-frequency representation of a real-valued multicomponent signal: (a) The SM (cross-terms and alias free version of the WD), (b) Polynomial Wigner-Ville distribution realized based on the STFT by using the SM and its order recursive form.

version of the WD. The cross-terms free realization of the WD and LWD is already presented. In the discrete implementation of the above relation, the only remaining problem is the evaluation of $W_x^A(t, f')$ on the discrete set of points on frequency axis, $f' = -k\Delta f'$. Since $W_x^A(t, f')$ is, by definition, a scaled and reversed version of $W_x(t, f')$, its values at $f' = -k\Delta f'$ are the values of $W_x(t, f')$ at $f' = k\Delta f'/A$. However, these points do not correspond to any sample location along the frequency axis grid. Thus, the interpolation has to be done (one way of doing it is in an appropriate zero padding of the signal). A discrete form of convolution (6.2.18), including rectangular window $P(\theta)$ and the above considerations, is

$$PW_x(n,k) = \sum_{i=-L_P}^{L_P} LW_2(n, k+i)\hat{W}_x(n, k+i/A) \tag{6.2.19}$$

where $2L_P + 1$ is the width of $P(\theta)$ in the discrete domain, while $\hat{W}_x(n, k+i/A)$ is the WD approximation. We can simply use $\hat{W}_x(n, k+i/A) = \mathrm{SM}_x(n, k+[i/A])$ where $[i/A]$ is the nearest integer to i/A, or use the linear interpolation of the SM values at two nearest integers. The terms in (6.2.19), when $k+i$ or $k+[i/A]$ is outside the basic period, are considered as being zero in order to avoid aliasing.

Example 3: Consider real-valued multicomponent signal

$$x(t) = \cos(20\sin(\pi t) + 30\pi t) + \sin(20\cos(\pi t) + 100\pi t)$$

within $-1 \leq t < 1$, with $\Delta t = 1/128$. In the realization, a Hanning window of the width $T_w = 2$ is used. Based on the STFT (using its positive frequencies), the cross-terms free WD is obtained from (6.2.13) with $L_P = 15$, and denoted by SM, Fig. 6.2.4(a). Then the LWD, with $L = 2$, is calculated according to (6.2.17). It is combined with the linearly interpolated SM value into the PWVD (6.2.19), shown in Fig. 6.2.4(b). For the precise implementation of $[i/A]$ the lag window has been zero-padded by a factor of 2.

6.2.4 Summary and Conclusions

The STFT based realization of quadratic TF representations, having the auto-terms close or the same to the ones in the WD, but without (or with reduced)

cross-terms, is presented. For this realization the S-method is used. The method is generalized, in an order recursive form, for the realization of higher order TF representations. Applications of the presented method on, for example, time-scale representations [6], and multidimensional space/spatial-frequency analysis [13], are straightforward. Hardware realization of the S-method is also simple and direct [7].

References

[1] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] L. Stanković, "Auto-term representation by the reduced interference distributions: A procedure for kernel design," *IEEE Trans. Signal Processing*, vol. 44, pp. 1557–1563, June 1996.

[3] L. Stanković, "A method for time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 225–229, January 1994.

[4] L. L. Scharf and B. Friedlander, "Toeplitz and Hankel kernels for estimating time-varying spectra of discrete-time random processes," *IEEE Trans. Signal Processing*, vol. 49, pp. 179–189, January 2001.

[5] L. Stanković, T. Alieva, and M. Bastiaans, "Fractional-Fourier-domain weighted Wigner distribution," in *Proc. Eleventh IEEE Workshop on Statistical Signal Processing*, pp. 321–324, Singapore, 6–8 August 2001.

[6] P. Gonçalvès and R. G. Baraniuk, "Pseudo affine Wigner distributions: Definition and kernel formulation," *IEEE Trans. Signal Processing*, vol. 46, pp. 1505–1517, June 1998.

[7] L. Stanković, "A method for improved distribution concentration in the time-frequency analysis of the multicomponent signals using the L-Wigner distribution," *IEEE Trans. Signal Processing*, vol. 43, pp. 1262–1268, May 1995.

[8] B. Ristic and B. Boashash, "Relationship between the polynomial and higher order Wigner-Ville distribution," *IEEE Signal Processing Letters*, vol. 2, pp. 227–229, December 1995.

[9] L. Stanković, "Multitime definition of the Wigner higher order distribution: L-Wigner distribution," *IEEE Signal Processing Letters*, vol. 1, pp. 106–109, July 1994.

[10] L. Stanković, "An analysis of Wigner higher order spectra of multicomponent signals," *Ann. Télécommunications*, vol. 49, pp. 132–136, March/April 1994.

[11] L. Stanković, "An analysis of some time-frequency and time-scale distributions," *Ann. Télécommunications*, vol. 49, pp. 505–517, September/October 1994.

[12] B. Boashash and B. Ristic, "Polynomial time-frequency distributions and time-varying higher order spectra: Application to the analysis of multicomponent FM signal and to the treatment of multiplicative noise," *Signal Processing*, vol. 67, pp. 1–23, May 1998.

[13] S. Stanković, L. Stanković, and Z. Uskoković, "On the local frequency, group shift, and cross-terms in some multidimensional time-frequency distributions: A method for multidimensional time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1719–1725, July 1995.

[14] L. Stanković and J. F. Böhme, "Time-frequency analysis of multiple resonances in combustion engine signals," *Signal Processing*, vol. 79, pp. 15–28, November 1999.

6.3 GABOR'S SIGNAL EXPANSION FOR A NON-ORTHOGONAL SAMPLING GEOMETRY[0]

6.3.1 Historical Perspective

In 1946 [1], Gabor suggested the representation of a time signal in a combined time-frequency domain (see Section 2.3.3); in particular he proposed to represent the signal as a superposition of shifted and modulated versions of a so-called elementary signal or synthesis window $g(t)$. Moreover, as a synthesis window $g(t)$ he chose a Gaussian signal, because such a signal has good localization in both the time domain and the frequency domain. The other choice that Gabor made was that his signal expansion was formulated on a rectangular lattice in the time-frequency domain $(mT, k\Omega)$, where the sampling distances T and Ω satisfied the relation $\Omega T = 2\pi$.

The coefficients in Gabor's signal expansion can be determined by using an analysis window $w(t)$. In the case of critical sampling, i.e., $\Omega T = 2\pi$, the analysis window $w(t)$ follows uniquely from the given synthesis window $g(t)$. However, such a unique analysis window appears to have some mathematically very unattractive properties. For this reason, the expansion should be formulated on a denser lattice, $\Omega T < 2\pi$. This makes the analysis window no longer unique and thus allows for finding an analysis window that is optimal in some way. We can, for instance, look for the analysis window that resembles best the synthesis window; a better resemblance can then be reached for a higher degree of oversampling.

A better resemblance can also be reached if we adapt the structure of the lattice to the form of the window as represented in the time-frequency domain. For the Gaussian window, for instance, the time-frequency representation has circular contour lines, and circles are better packed on a hexagonal lattice than on a rectangular lattice. Gabor's signal expansion on such a hexagonal, non-orthogonal lattice then leads to a better resemblance between the window functions $g(t)$ and $w(t)$ than the expansion on a rectangular, orthogonal lattice does.

6.3.2 Gabor's Signal Expansion on a Rectangular Lattice

We start with the usual Gabor expansion [1–5] on a rectangular time-frequency lattice, in which case a signal $\varphi(t)$ can be expressed as a linear combination of properly shifted and modulated versions $g_{mk}(t) = g(t - mT)\exp(jk\Omega t)$ of a synthesis window $g(t)$:

$$\varphi(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{mk}\, g_{mk}(t). \tag{6.3.1}$$

The time step T and the frequency step Ω satisfy the relationship $\Omega T \leq 2\pi$; note that the factor $2\pi/\Omega T$ represents the degree of oversampling, and that in his original

[0]Authors: **Martin J. Bastiaans** and **Arno J. van Leest**, Technische Universiteit Eindhoven, Faculteit Elektrotechniek, Postbus 513, 5600 MB Eindhoven, Netherlands (M.J.Bastiaans@tue.nl). Reviewers: Joel M. Morris and Shie Qian.

paper [1] Gabor considered the case of critical sampling, i.e. $\Omega T = 2\pi$. The expansion coefficients a_{mk} follow from sampling the windowed Fourier transform with analysis window $w(t)$, $\int_{-\infty}^{\infty} \varphi(t)\, w^*(t-\tau)\, \exp(-j\omega t)\, dt$, on the rectangular lattice ($\tau = mT, \omega = k\Omega$):

$$a_{mk} = \int_{-\infty}^{\infty} \varphi(t)\, w_{mk}^*(t)\, dt. \tag{6.3.2}$$

This relationship is known as the Gabor transform.

The synthesis window $g(t)$ and the analysis window $w(t)$ are related to each other in such a way that their shifted and modulated versions constitute two sets of functions that are biorthogonal:

$$\sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} g_{mk}(t_1)\, w_{mk}^*(t_2) = \delta(t_1 - t_2). \tag{6.3.3}$$

If the biorthogonality condition (6.3.3) is satisfied, the Gabor transform (6.3.2) and Gabor's signal expansion (6.3.1) form a transform pair in the following sense: if we start with an arbitrary signal $\varphi(t)$ and determine its expansion coefficients a_{mk} via the Gabor transform (6.3.2), the signal can be reconstructed via the Gabor expansion (6.3.1).

The biorthogonality relation (6.3.3) leads immediately to the equivalent but simpler expression

$$\frac{2\pi}{\Omega} \sum_{m=-\infty}^{\infty} g(t - mT)\, w^*\left(t - \left[mT + n\frac{2\pi}{\Omega}\right]\right) = \delta_n, \tag{6.3.4}$$

where δ_n is the Kronecker delta. In the case of critical sampling, i.e., $\Omega T = 2\pi$, the biorthogonality relation (6.3.4) reduces to

$$T \sum_{m=-\infty}^{\infty} g(t - mT)\, w^*(t - [m+n]T) = \delta_n \tag{6.3.5}$$

and the analysis window $w(t)$ follows uniquely from a given synthesis window $g(t)$, or vice versa. An elegant way to find the analysis window if the synthesis window is given, is presented in the next section.

6.3.3 Fourier Transform and Zak Transform

It is well known (see, for instance, [2–5]) that in the case of critical sampling, $\Omega T = 2\pi$, Gabor's signal expansion (6.3.1) and the Gabor transform (6.3.2) can be transformed into product form. We therefore need the Fourier transform $\bar{a}(t/T, \omega/\Omega)$ of the two-dimensional array of Gabor coefficients a_{mk}, defined by

$$\bar{a}(x, y) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{mk}\, \exp[-j2\pi(my - kx)], \tag{6.3.6}$$

and the Zak transforms $\tilde{\varphi}(xT, 2\pi y/T; T)$, $\tilde{g}(xT, 2\pi y/T; T)$, and $\tilde{w}(xT, 2\pi y/T; T)$ of the signal $\varphi(t)$ and the window functions $g(t)$ and $w(t)$, respectively, where the Zak transform $\tilde{f}(t, \omega; \tau)$ of a function $f(t)$ is defined as (see, for instance, [2,5])

$$\tilde{f}(t, \omega; \tau) = \sum_{n=-\infty}^{\infty} f(t + n\tau) \exp(-jn\tau\omega). \tag{6.3.7}$$

Note that the Fourier transform $\bar{a}(x, y)$ is periodic in x and y with period 1, and that the Zak transform $\tilde{f}(t, \omega; \tau)$ is periodic in ω with period $2\pi/\tau$ and quasi-periodic in t with period τ: $\tilde{f}(t + m\tau, \omega + 2\pi k/\tau; \tau) = \tilde{f}(t, \omega; \tau) \exp(jm\omega\tau)$.

Upon substituting from the Fourier transform (6.3.6) and the Zak transforms [cf. Eq. (6.3.7)] into Eqs. (6.3.1) and (6.3.2), it is not too difficult to show that Gabor's signal expansion (6.3.1) can be transformed into the product form

$$\tilde{\varphi}\left(xT, y\frac{2\pi}{T}; T\right) = \bar{a}(x, y)\, \tilde{g}\left(xT, y\frac{2\pi}{T}; T\right), \tag{6.3.8}$$

while the Gabor transform (6.3.2) can be transformed into the product form

$$\bar{a}(x, y) = T\, \tilde{\varphi}\left(xT, y\frac{2\pi}{T}; T\right) \tilde{w}^*\left(xT, y\frac{2\pi}{T}; T\right). \tag{6.3.9}$$

In particular the product form (6.3.9) is useful for determining Gabor's expansion coefficients. Since a Zak transform is merely a Fourier transform [cf. Eq. (6.3.7)], the expansion coefficients can be determined by Fourier transformations and multiplications; and if things are formulated for discrete-time signals, we can use the *fast* Fourier transform to formulate a fast algorithm for the Gabor transform [3,4].

The relationship between the Zak transforms of the analysis window $w(t)$ and the synthesis window $g(t)$ then follows from substituting from Eq. (6.3.9) into Eq. (6.3.8) and reads

$$T\, \tilde{g}\left(xT, y\frac{2\pi}{T}; T\right) \tilde{w}^*\left(xT, y\frac{2\pi}{T}; T\right) = 1. \tag{6.3.10}$$

From the latter relationship we conclude that (the Zak transform of) the analysis window $w(t)$ follows uniquely from (the Zak transform of) the given synthesis window $g(t)$. In general, however, the unique analysis window $w(t)$ has some very unattractive mathematical properties. We are therefore urged to consider Gabor's signal expansion on a denser lattice, in which case the analysis window is no longer unique. This enables us to choose an analysis window that is better suited to our purpose of determining Gabor's expansion coefficients.

6.3.4 Rational Oversampling

In the case of oversampling by a rational factor, $2\pi/\Omega T = p/q \geq 1$, with p and q relatively prime, positive integers, $p > q \geq 1$, Gabor's expansion (6.3.1) and the

Gabor transform (6.3.2) can be transformed into the sum-of-products forms [3,4], cf. Eqs. (6.3.8) and (6.3.9),

$$\varphi_s(x,y) = \frac{1}{p}\sum_{r=0}^{p-1} g_{sr}(x,y)\, a_r(x,y) \quad (s=0,1,\ \ldots,q-1) \tag{6.3.11}$$

$$a_r(x,y) = \frac{pT}{q}\sum_{s=0}^{q-1} w^*_{sr}(x,y)\,\varphi_s(x,y) \quad (r=0,1,\ \ldots,p-1), \tag{6.3.12}$$

respectively, where we have introduced the shorthand notations

$$\begin{aligned} a_r(x,y) &= \tilde{a}(x,y+r/p) \\ \varphi_s(x,y) &= \tilde{\varphi}((x+s)pT/q, 2\pi y/T; pT) \\ g_{sr}(x,y) &= \tilde{g}((x+s)pT/q, 2\pi(y+r/p)/T; T) \\ w_{sr}(x,y) &= \tilde{w}((x+s)pT/q, 2\pi(y+r/p)/T; T), \end{aligned}$$

with $0 \le x < 1$ and $s = 0,1,\ \ldots, q-1$ (and hence $0 \le (x+s)/q < 1$), and $0 \le y < 1/p$ and $r = 0,1,\ \ldots, p-1$ (and hence $0 \le y + r/p < 1$). The relationship between the Zak transforms of the analysis window $w(t)$ and the synthesis window $g(t)$ then follows from substituting from Eq. (6.3.12) into Eq. (6.3.11) and reads [cf. Eq. (6.3.10)]

$$\frac{T}{q}\sum_{r=0}^{p-1} g_{s_1 r}(x,y)\, w^*_{s_2 r}(x,y) = \delta_{s_1 - s_2}, \tag{6.3.13}$$

with $s_1, s_2 = 0,1,\ \ldots, q-1$. The latter relationship represents a set of q^2 equations for pq unknowns, which set is underdetermined since $p > q$, and we conclude that the analysis window does not follow uniquely from the synthesis window.

After combining the p functions $a_r(x,y)$ into a p-dimensional column vector $\boldsymbol{a}(x,y)$, the q functions $\varphi_s(x,y)$ into a q-dimensional column vector $\boldsymbol{\phi}(x,y)$, and the $q \times p$ functions $g_{sr}(x,y)$ and $w_{sr}(x,y)$ into the $q \times p$-dimensional matrices $\boldsymbol{G}(x,y)$ and $\boldsymbol{W}(x,y)$, respectively, the sum of products forms can be expressed as matrix-vector and matrix-matrix multiplications:

$$\boldsymbol{\phi}(x,y) = \frac{1}{p}\,\boldsymbol{G}(x,y)\,\boldsymbol{a}(x,y) \tag{6.3.14}$$

$$\boldsymbol{a}(x,y) = \frac{pT}{q}\,\boldsymbol{W}^*(x,y)\,\boldsymbol{\phi}(x,y) \tag{6.3.15}$$

$$\boldsymbol{I}_q = \frac{T}{q}\,\boldsymbol{G}(x,y)\,\boldsymbol{W}^*(x,y), \tag{6.3.16}$$

where $\boldsymbol{I}_q$ denotes the $q \times q$-dimensional identity matrix and where, as usual, the asterisk in connection with vectors and matrices denotes complex conjugation *and* transposition.

The latter relationship again represents q^2 equations for pq unknowns, and the $p \times q$ matrix $\boldsymbol{W}^*(x,y)$ cannot be found by a simple inversion of the $q \times p$ matrix $\boldsymbol{G}(x,y)$. An 'optimum' solution that is often used, is based on the generalized inverse and reads $\boldsymbol{W}^*_{opt}(x,y) = (q/T)\, \boldsymbol{G}^*(x,y)\, [\boldsymbol{G}(x,y)\, \boldsymbol{G}^*(x,y)]^{-1}$. This solution for $\boldsymbol{W}(x,y)$ is optimum in the sense that (i) it yields the analysis window $w(t)$ with the lowest L^2 norm, (ii) it yields the Gabor coefficients a_{mk} with the lowest L^2 norm, and (iii) it yields the analysis window that – in an L^2 sense, again – best resembles the synthesis window.

The 'optimum' solution gets better if the degree of oversampling p/q becomes higher. However, there is another way of finding a better solution, based on the structure of the lattice. If the lattice structure is adapted to the form of the window function as it is represented in the time-frequency domain, the 'optimum' solution will be better, even for a lower degree of oversampling. We will therefore consider the case of a non-orthogonal sampling geometry, but we will do that in such a way that we can relate this non-orthogonal sampling to orthogonal sampling. In that case we will still be able to use product forms of Gabor's expansion and the Gabor transform, and benefit from all the techniques that have been developed for them.

6.3.5 Non-Orthogonal Sampling

The rectangular (or orthogonal) lattice that we considered in the previous sections, where sampling occurred on the lattice points ($\tau = mT, \omega = k\Omega$), can be obtained by integer combinations of two orthogonal vectors $[T,0]^t$ and $[0,\Omega]^t$, see Fig. 6.3.1(a), which vectors constitute the lattice generator matrix

$$\begin{bmatrix} T & 0 \\ 0 & \Omega \end{bmatrix}.$$

We now consider a time-frequency lattice that is no longer orthogonal. Such a lattice is obtained by integer combinations of two linearly independent, but no longer orthogonal vectors, which we express in the forms $[aT, c\Omega]^t$ and $[bT, d\Omega]^t$, with a, b, c and d integers, and which constitute the lattice generator matrix

$$\begin{bmatrix} aT & bT \\ c\Omega & d\Omega \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & \Omega \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Without loss of generality, we may assume that the integers a and b have no common divisors, and that the same holds for the integers c and d; possible common divisors can be absorbed in T and Ω. Note that we only consider lattices that have samples on the time and frequency axes and that are therefore suitable for a discrete-time approach, as well.

The area of a cell (a parallelogram) in the time-frequency plane, spanned by the two vectors $[aT, c\Omega]^t$ and $[bT, d\Omega]^t$, is equal to the determinant of the lattice generator matrix, which determinant is equal to $\Omega T D$, with $D = |ad - bc|$. To be

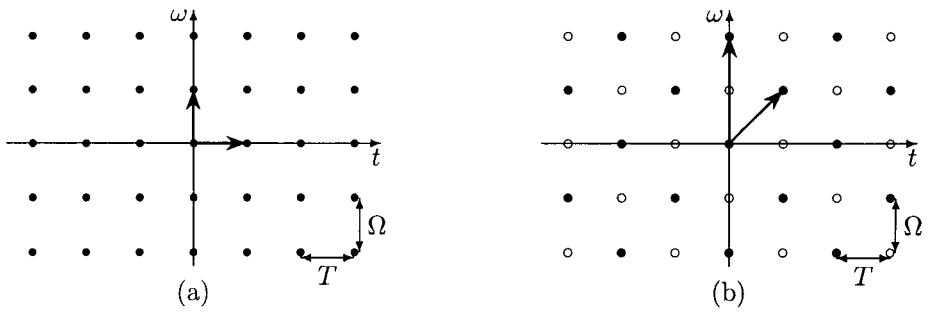


Fig. 6.3.1: (a) A rectangular lattice with lattice vectors $[T,0]^t$ and $[0,\Omega]^t$, and thus $R=0$ and $D=1$; and (b) a hexagonal lattice with lattice vectors $[T,\Omega]^t$ and $[0,2\Omega]^t$, and thus $R=1$ and $D=2$.

usable as a proper Gabor sampling lattice, this area should satisfy the condition $D \leq 2\pi/\Omega T$.

There are a lot of lattice generator matrices that generate the same lattice. We will use the one that is based on the Hermite normal form, unique for any lattice,

$$\begin{bmatrix} T & 0 \\ R\Omega & D\Omega \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & \Omega \end{bmatrix} \begin{bmatrix} 1 & 0 \\ R & D \end{bmatrix},$$

where R and D are relatively prime integers and $0 \leq |R| < D$. Sampling then occurs on the lattice points $(\tau = mT, \omega = [mR+nD]\Omega)$, and it is evident that these points of the non-orthogonal lattice form a subset of the points $(\tau = mT, \omega = k\Omega)$ of the orthogonal lattice. To be more specific: the non-orthogonal lattice is formed by those points of the rectangular (orthogonal) lattice for which $k - mR$ is an integer multiple of D. Note that the original rectangular lattice arises for $R = 0$ and $D = 1$, see Fig. 6.3.1(a), and that a hexagonal lattice occurs for $R = 1$ and $D = 2$, see Fig. 6.3.1(b).

6.3.6 Gabor's Signal Expansion on a Non-Orthogonal Lattice

If we define the two-dimensional array λ_{mk} as

$$\lambda_{mk} = \sum_{n=-\infty}^{\infty} \delta_{k-mR-nD}, \tag{6.3.17}$$

Gabor's signal expansion on a non-orthogonal lattice can be expressed as [cf. Eq. (6.3.1)]

$$\varphi(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \lambda_{mk}\, a_{mk}\, g_{mk}(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a'_{mk}\, g_{mk}(t), \tag{6.3.18}$$

while – with a different analysis window $w(t)$, though! – the expansion coefficients a_{mk} are still determined by the Gabor transform (6.3.2). Of course, since we only

need the limited array $a'_{mk} = \lambda_{mk}\, a_{mk}$ – which is, in fact, a properly sampled version of the full array a_{mk} – we need only calculate the coefficients a_{mk} for those values of m and k for which $k - mR$ is an integer multiple of D. We note that the Fourier transform $\bar{a}'(x,y)$ of the limited array a'_{mk} is related to the Fourier transform $\bar{a}(x,y)$ of the full array a_{mk} via the periodization relation

$$\bar{a}'(x,y) = \frac{1}{D} \sum_{n=0}^{D-1} \bar{a}\left(x - \frac{n}{D}, y - \frac{nR}{D}\right) \tag{6.3.19}$$

and thus

$$a'_r(x,y) = \frac{1}{D} \sum_{n=0}^{D-1} a_r\left(x - \frac{n}{D}, y - \frac{nR}{D}\right). \tag{6.3.20}$$

In the non-orthogonal case, the biorthogonality condition takes the form [cf. Eq. (6.3.3)]

$$\sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \lambda_{mk}\, g_{mk}(t_1)\, w^*_{mk}(t_2) = \delta(t_1 - t_2) \tag{6.3.21}$$

and leads to the equivalent but simpler expression [cf. Eq. (6.3.4)]

$$\frac{2\pi}{D\Omega} \sum_{m=-\infty}^{\infty} g(t - mT)\, w^*\left(t - \left[mT + n\frac{2\pi}{D\Omega}\right]\right) \exp\left(j2\pi m \frac{nR}{D}\right) = \delta_n. \tag{6.3.22}$$

Note that for $R = 0$ and $D = 1$, for which we have a rectangular lattice [see Fig. 6.3.1(a)], Eq. (6.3.22) reduces to Eq. (6.3.4), and that for $R = 1$ and $D = 2$, for which we have a hexagonal lattice [see Fig. 6.3.1(b)], Eq. (6.3.22) takes the form

$$\frac{\pi}{\Omega} \sum_{m=-\infty}^{\infty} g(t - mT)\, w^*\left(t - \left[mT + n\frac{\pi}{\Omega}\right]\right) (-1)^{mn} = \delta_n. \tag{6.3.23}$$

The biorthogonality condition expressed in terms of the Zak transforms of the window functions now takes the form, cf. Eq. (6.3.13),

$$\frac{T}{Dq} \sum_{r=0}^{p-1} g_{s_1,r}(x,y)\, w^*_{s_2,r}\left(x - \frac{n}{D}, y - \frac{nR}{D}\right) = \delta_n \delta_{s_1 - s_2}, \tag{6.3.24}$$

with $s_1, s_2 = 0, 1, \ldots, q-1$ and $n = 0, 1, \ldots, D-1$, and allows an easy determination of the analysis window $w(t)$ for a given synthesis window $g(t)$. For $R = 0$ and $D = 1$, for instance, relation (6.3.24) reduces to Eq. (6.3.13), while for $R = 1$, $D = 2$, $q = 1$, and p an even integer – which corresponds to the integer ($p/2$-times) oversampled hexagonal case – it reduces to

$$\frac{T}{2} \sum_{r=0}^{p-1} g_{0,r}(x,y)\, w^*_{0,r-np/2}(x,y)\, (-1)^{nr} = \delta_n \quad (n = 0, 1;\ p \text{ even}), \tag{6.3.25}$$

from which the Zak transform $\tilde{w}(t, \omega; T)$ and hence the window function $w(t)$ can easily be determined.

Since we have related Gabor's signal expansion on a non-orthogonal lattice to sampling on a denser but orthogonal lattice, followed by restriction to a sub-lattice that corresponds to the non-orthogonal lattice, we can still use all the techniques that are developed for rectangular lattices, in particular the technique of determining Gabor's expansion coefficients via the Zak transform, cf. Eq. (6.3.12).

6.3.7 Summary and Conclusions

Gabor's signal expansion and the Gabor transform on a rectangular lattice have been introduced, along with the Fourier transform of the array of expansion coefficients and the Zak transforms of the signal and the window functions. Based on these Fourier and Zak transforms, the sum-of-products forms for the Gabor expansion and the Gabor transform, which hold in the rationally oversampled case, have been derived.

We have then studied Gabor's signal expansion and the Gabor transform based on a non-orthogonal sampling geometry. We have done this by considering the non-orthogonal lattice as a sub-lattice of an orthogonal lattice. This procedure allows us to use all the formulas that hold for the orthogonal sampling geometry. In particular we can use the sum-of-products forms that hold in the case of a rationally oversampled rectangular lattice.

We finally note that if everything remains to be based on a rectangular sampling geometry (as in Article 4.9), it will be easier to extend the theory of the Gabor scheme to higher-dimensional signals; see, for instance, [6], where the multi-dimensional case is treated for continuous-time as well as discrete-time signals.

References

[1] D. Gabor, "Theory of communication," *J. IEE*, vol. 93(III), pp. 429–457, November 1946.

[2] M. J. Bastiaans, "Gabor's signal expansion and its relation to sampling of the sliding-window spectrum," in *Advanced Topics in Shannon Sampling and Interpolation Theory* (R. J. Marks II, ed.), pp. 1–35, New York: Springer, 1993.

[3] M. J. Bastiaans, "Gabor's expansion and the Zak transform for continuous-time and discrete-time signals: Critical sampling and rational oversampling," Research Report 95-E-295, Eindhoven University of Technology, Eindhoven, Netherlands, 1995.

[4] M. J. Bastiaans and M. C. W. Geilen, "On the discrete Gabor transform and the discrete Zak transform," *Signal Processing*, vol. 49, pp. 151–166, March 1996.

[5] H. G. Feichtinger and T. Strohmer, eds., *Gabor Analysis and Algorithms: Theory and Applications*. Berlin/Boston: Birkhäuser, 1998.

[6] A. J. van Leest, *Non-separable Gabor schemes: Their design and implementation*. PhD thesis, Eindhoven University of Technology, Eindhoven, Netherlands, 2001.

6.4 SPECTROGRAM DECOMPOSITIONS OF TIME-FREQUENCY DISTRIBUTIONS [0]

Cohen's class of bilinear time-frequency distributions (or the quadratic class; see p. 68n) has attracted much attention over the past decade or so. While such TFDs have yielded a proliferation of methods and are becoming increasingly popular in applications, their progress has been slowed, at least in part, by the computational burdens which are generally much greater than for spectrograms or wavelet-based methods. This article demonstrates methods for greatly reducing the computational burden, by decomposing the TFDs into linear weighted sums of spectrograms. The computational efficiency comes from economizing the number of STFTs that must be computed to form these spectrograms. Surprisingly, one can reduce the number of STFTs to a small number by using Haar windows for the STFT computation [1–3]. Finally, for certain kernels, special windows may be designed which can represent a TFD very well by using only three or four windows [3,4].

It should be noted that several attempts have been made to decompose *time-frequency distributions* themselves, in order to isolate specific TFD components or to synthesize TFDs using several components. While this is an interesting and valuable topic, it is beyond the scope of this article, which deals specifically with decomposition of the *kernel* of the TFD.

6.4.1 Decomposition Based Approaches

Cunningham and Williams have shown that a TFD of Cohen's class can be expressed as a weighted sum of spectrograms, where the spectrogram windows are an orthonormal set [5,6], and have given a rigorous base for further work. They were motivated by some ideas from White [7] and from Amin [8], who has continued his work in this area. The windows result from an eigensystem decomposition of the kernel. This approach has been used to approximate various TFDs [6]. A twist on this concept is to use windows which are shifted and scaled versions of each other [1–3]. In these studies, windows from the Haar basis set were used. It can be shown that any TFD can be represented by n+1 Haar windows, where $n = \log_2(N-1)$, and N is the number of sample points. Thus, a 257-point kernel can be represented by 9 Haar functions, one at each scale. It is important to note that the cross-spectrograms, i.e. spectrograms formed by STFTs using different windows, are used as well as the auto-spectrograms to form a basis for the TFD.

The Reduced Interference Distribution [9–11] or RID is a well-defined set of Cohen's-class TFDs [12] that can be constructed using a simple set of constraints, while retaining a number of very desirable mathematical properties. Faster computation of this class of TFDs was a particular motivation for the development of the spectrogram decomposition methods discussed in this article.

---

[0]Authors: **William J. Williams** and **Selin Aviyente**, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor MI 48109, USA (wjw@eecs.umich.edu, saviyent@eecs.umich.edu). Reviewers: B. Boashash and P. Flandrin.

6.4.2 Decomposition of Time-Frequency Kernels

For a given time-frequency kernel, the first step in the spectrogram decomposition algorithm is to decompose the kernel in terms of spectrogram windows. The approach suggested by Cunningham and Williams used an eigensystem type approach to decompose the 2-D kernel expressed in time-lag (n, m) as the outer product of a set of orthonormal basis functions. These basis functions become the windows of the spectrogram decomposition and the eigenvalues are the weightings of the spectrogram summations. Let $C(n, k)$ be a TFD with a real, bounded kernel of interest; then it can be expressed as

$$C(n,k) = \sum_{i=1}^{M} \lambda_i \left|\mathrm{STFT}_i(n,k)\right|^2 . \tag{6.4.1}$$

Denote the spectrogram windows as $w_i(n)$, and the signal as $x(n)$; then the STFTs can be represented as

$$\mathrm{STFT}_i(n,k) = \sum_{m=0}^{N-1} x(m-n) w_i(m)\, e^{-j2\pi mk/N} . \tag{6.4.2}$$

This is the Cunningham-Williams decomposition in a nutshell. The problem is that while the eigendecomposition gives an orthonormal set of windows, there is no orderly relationship between the windows as is the case for a Principal Component Analysis (PCA) or Karhunen-Loève Transform (KLT) for a set of signals. This is the "best basis" for the kernel in the same sense as is the KLT for representation of a set of signals. There is no redundancy or relationship between windows that allows more efficient computation. The principal eigenvector does the best job of representing the kernel in terms of an outer product with itself, the next eigenvector does the next best job when its contribution is compared to the remaining others, and so on.

Generally, the kernel can be adequately represented using a small number of the potential windows. However, a spectrogram must be computed for each window. In the present article we suggest methods which use specially designed windows, potentially increasing the speed and efficiency of representation.

6.4.3 Development of the Method

Time- and frequency-shift-covariant bilinear discrete TFRs are specified by a discrete kernel, and can be rewritten in the inner product form of

$$\begin{aligned}
&\mathrm{TFR}_x(n,\omega;\psi) \\
&= \sum_{n_1}\sum_{n_2} x(n+n_1)\, e^{-j\omega(n+n_1)} \psi\big(-\tfrac{n_1+n_2}{2}, n_1-n_2\big) \big[x(n+n_2) e^{-j\omega(n+n_2)}\big]^* \\
&= \big\langle \tilde{\psi}\, \overline{S}_{-n}\, \overline{M}_{-\omega}\, x\,,\, \overline{S}_{-n}\, \overline{M}_{-\omega}\, x \big\rangle
\end{aligned} \tag{6.4.3}$$

where $\overline{S}_{-n}$ and $\overline{M}_{-\omega}$ are, respectively, the time and frequency shift operators on ℓ_2, the space of finite-energy discrete-time signals, and $\tilde{\psi}$ is a bounded linear operator on ℓ_2 [5,6].

The spectral representation of $\tilde{\psi}$ may be used to express the TFR as a weighted sum of spectrograms or "projectograms". If the kernel is associated with a bounded, self-adjoint linear operator, then the kernel may be decomposed by an eigendecomposition such that one can represent the TFR as being composed of a finite series of spectrograms. The orthonormal windows forming the spectrograms are the eigenfunctions of the decomposition. The eigenvalues of the decomposition provide the weights for summing the set of spectrograms. The viewpoint may be taken that the projections of the signal on the eigenvectors of the kernel decomposition are then time and frequency shifted by the time and frequency shift operators, yielding, essentially, the STFT. The magnitude squared STFT is the "projectogram" or spectrogram associated with that particular window. A total of N windows are required to completely represent an $N \times N$ kernel.

The spectrogram decomposition can more generally be expressed as

$$\begin{aligned} &\mathrm{TFR}_x(n,\omega;\psi) \qquad (6.4.4)\\ &= \sum_{n_1} x(n+n_1)\, e^{-j\omega(n+n_1)} \sum_{n_1} x^*(n+n_2)\, e^{j\omega(n+n_2)} \sum_{l=1}^{N}\sum_{k=1}^{N} \lambda_{k,l}\, w_k(n_1)\, w_l(n_2). \end{aligned}$$

Here, w_k and w_l are the windows and $\lambda_{k,l}$ is the coefficient for each k,l. Then, $\sum_{n_1} w_k(n_1)x(n+n_1)e^{-j\omega(n+n_1)}$ is recognized as $\mathrm{STFT}_k(n,\omega)$ and the kernel is realized by the outer product $\sum_{l=1}^{N}\sum_{k=1}^{N}\lambda_{k,l}w_k(n_1)w_l(n_2)$.

By realizing that the spectrogram (cross and auto) is

$$\mathrm{SP}_{k,l}(n,\omega) = \mathrm{STFT}_k(n,\omega)\,\mathrm{STFT}_l^*(n,\omega) \qquad (6.4.5)$$

one may write Eq. (6.4.4) as

$$\mathrm{TFR}_x(n,\omega;\psi) = \sum_{l=1}^{N}\sum_{k=1}^{N}\lambda_{k,l}\mathrm{SP}_{k,l}(n,\omega). \qquad (6.4.6)$$

6.4.4 Wigner Example

The kernel matrix for an 8×8 discrete Wigner matrix is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \qquad (6.4.7)$$

Using the Haar basis vectors

$$\frac{1}{\sqrt{8}}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix} \tag{6.4.8}$$

one may obtain the projection of the kernel matrix on the outerproduct formed by all possible pairs of Haar basis vectors. The projection value is obtained by multiplying the kernel and the outerproduct matrices element by element and summing over all elements of the result. This provides the weighting $\lambda(k,l)$ for combining the spectrograms of Eq. (6.4.4). The array of weights is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{6.4.9}$$

Generally, good TFR representation is possible using only a fraction of the windows required for full representation. The spectrogram itself, of course, requires only one term, since it has only one window in its decomposition. In general, N^2 cross-spectrograms are required, but for many kernels only a few weights are nonzero and other weights are small enough to set to zero. However, even better results may obtained by using a special set of orthogonal windows, one may represent the TFR using many fewer cross-spectrograms.

6.4.5 Optimum Orthogonal Windows

The kernel for the binomial TFD [11] is shown here for illustration for $N = 9$:

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{8} & \frac{4}{16} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{3}{8} & \frac{6}{16} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \frac{3}{8} & \frac{4}{16} & 0 \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ 0 & \frac{4}{16} & \frac{3}{8} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{6}{16} & \frac{3}{8} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{4}{16} & \frac{1}{8} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{16} & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.4.10}$$

The method used by Cunningham and Williams would require the decomposition

$$\mathbf{V}\lambda = \mathbf{B}\mathbf{V} \tag{6.4.11}$$

where λ is a diagonal matrix of eigenvalues. The corresponding rows and columns of $\mathbf{V}$ form outer products which, weighted by the appropriate eigenvalues and summed, form the kernel

$$\mathbf{B} = \mathbf{V}\lambda\mathbf{V}'. \tag{6.4.12}$$

One would wish the eigenvalues to decline quickly, indicating that a truncated outerproduct reconstruction could suffice. Generally, it has been found that about 17 outerproducts are sufficient to well represent a 256×256 kernel matrix [5].

More efficiency can be gained taking advantages of symmetry. The upper right and lower left portions of $\mathbf{B}$ exhibit symmetry, so all of the kernel information is carried by the submatrix

$$\mathbf{C} = \begin{bmatrix} \frac{1}{16} & 0 & 0 & 0 & 0 \\ \frac{1}{8} & \frac{4}{16} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{8} & \frac{6}{16} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{8} & \frac{4}{16} & 0 \\ 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \end{bmatrix}. \tag{6.4.13}$$

Much greater efficiency can be gained by rotating matrix $\mathbf{C}$ to form

$$\mathbf{D} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{8} & \frac{4}{16} & 0 \\ \frac{1}{4} & \frac{3}{8} & \frac{6}{16} & 0 & 0 \\ \frac{1}{8} & \frac{4}{16} & 0 & 0 & 0 \\ \frac{1}{16} & 0 & 0 & 0 & 0 \end{bmatrix} \tag{6.4.14}$$

where the 1 in the center of the kernel matrix has been replaced by zero and a single companion matrix with the missing 1 (and zero elements otherwise) is created. These matrices can be augmented with additional zeros in their rows and columns to a size equal to the original kernel. Then, a third matrix can be formed from $\mathbf{C}$ by flipping $\mathbf{C}$ along the anti-diagonal. These three zero augmented matrices can be summed to form the original kernel. The eigenvectors of $\mathbf{D}$ can now be used as windows once the matrix is rotated back to its original position. However, the windows resulting from the eigenvectors now come in pairs such that the proper outerproducts are formed using a window and its time-reversed pair. Due to symmetry, only the upper right quadrant of $\mathbf{B}$ as represented by $\mathbf{C}$, plus the impulse window which provides the 1 in the center, is required to compute the STFTs. Details about the construction of TFDS from the STFTs are given elsewhere [1–4].

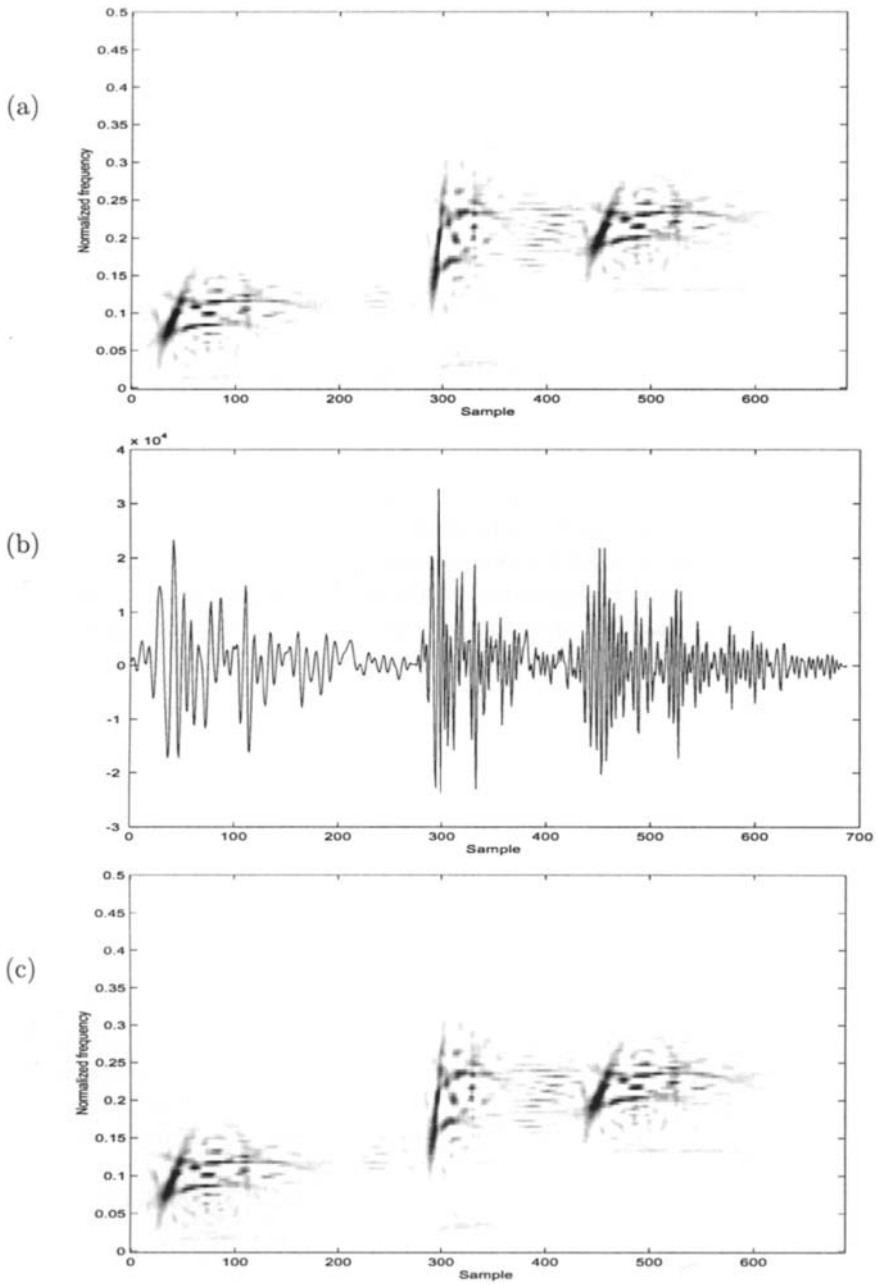


Fig. 6.4.1: Comparison of results for dolphin clicks: (a) True Binomial TFD, (b) Time series, and (c) Approximated Binomial TFD using a 4-spectrogram approximation.

6.4.6 Kernel Decomposition Results

Computation of TFDs by this method is efficient and fast. An example (Fig. 6.4.1) is provided for a 129 × 129 binomial kernel matrix. Only four distinct windows plus the impulse window are required to provide a very good representation of a complex signal. The kernel reconstructed from the truncated outer product series has only about 7% residual error. The signal consists of a dolphin click followed by a time-shifted and scaled version of the click (2:1 compressed in time and normalized for energy) followed by a time-shifted and frequency-shifted version of the click. This signal has been used elsewhere to illustrate the time-shift, scale and frequency shift covariance of RIDs. Nine STFTs must be computed for this approximation. One additional trivial computation is required for the impulse window. Clearly, the true and approximated TFD results are very similar. If one looks closely it appears that the cross-term activity is decreased in the approximate result. This is consistent with previous observations that noise and cross-term activity are abated in the approximations. One could compute each of the STFTs required in parallel. The formation of the cross-spectrograms from the STFTs and the summation of those results would require additional hardware or software, but these operations involve only multiplication of STFT points, multiplication of the resulting cross-spectrograms by the required eigenvalue coefficients and finally, summation of the results. In addition, a number of signal processing "tricks" may be employed to speed up the computations due to symmetry, the realness of the TFD result, etc.

Some TFDs that fall into the RID class may be approximated with a smaller number of windows. The Born-Jordan TFD requires only three distinct windows for a very nice representation. In fact, one may compute a legitimate RID using an impulse and one other distinct window [4]. However, this minimum window TFD may not have desirable representation properties in some cases. The windows derived by these methods have an interesting self-scaling property which accounts for the scale covariance being retained [13].

6.4.7 Summary and Conclusions

The liability of bilinear TFDs from Cohen's class in terms of computational burden can be overcome to a considerable degree by employing various means of spectrogram decomposition. In addition to fast computation, noise and cross-terms may be decreased in the approximated form. It is clear that such TFDs can now be computed very rapidly using parallel computation or dedicated hardware, bringing their use into a wider range of practical applications.

The next article considers the computation of quadratic TFDs discretized by the method of Article 6.1.

References

[1] W. J. Williams and J. C. O'Neill, "Decomposition of time-frequency distributions using scaled window spectrograms," in *Proc. SPIE: Advanced Signal Processing Algorithms,*

vol. 2563, pp. 44–58, Soc. of Photo-optical Instrumentation Engineers, San Diego, CA, 10–12 July 1995.

[2] W. J. Williams, T.-H. Sang, J. C. O'Neill, and E. J. Zalubas, "Wavelet windowed time-frequency distribution decompositions," in *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations VII*, vol. 3162, pp. 149–160, Soc. of Photo-optical Instrumentation Engineers, San Diego, July–August 1997.

[3] W. J. Williams, "Reduced interference time-frequency distributions: Scaled decompositions and interpretations," in *Wavelet Transforms and Time-Frequency Signal Analysis* (L. Debnath, ed.), ch. 12, Boston: Birkhäuser, 2001.

[4] W. J. Williams and S. Aviyente, "Optimum window time-frequency distribution decompositions," in *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers*, pp. 817–821, Pacific Grove, CA, 1–4 November 1998.

[5] G. S. Cunningham and W. J. Williams, "Kernel decomposition of time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, pp. 1425–1442, June 1994.

[6] G. S. Cunningham and W. J. Williams, "Fast implementations of generalized discrete time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, pp. 1496–1508, June 1994.

[7] L. B. White, "Transition kernels for bilinear time-frequency signal representations," *IEEE Trans. Signal Processing*, vol. 39, pp. 542–544, February 1991.

[8] M. G. Amin, "Spectral decomposition of time-frequency distribution kernels," *IEEE Trans. Signal Processing*, vol. 42, pp. 1156–1165, May 1994.

[9] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 37, pp. 862–871, June 1989.

[10] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, pp. 402–412, February 1992.

[11] W. J. Williams and J. Jeong, "Reduced interference time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 3, pp. 74–97, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[12] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[13] W. J. Williams and S. Aviyente, "Minimal-window time-frequency distributions," in *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, vol. 3807, pp. 446–457, Soc. of Photo-optical Instrumentation Engineers, Denver, CO, 19–21 July 1999.

6.5 COMPUTATION OF DISCRETE QUADRATIC TFDs[0]

6.5.1 General Computational Procedure

Article 6.1 deals with definitions and properties of the discrete WVD (DWVD) and other discrete quadratic TFDs. It shows that the general discrete quadratic TFD of an analytic signal $z[n]$ is

$$\rho_z[n,k] = 2 \sum_{|m|<\frac{M}{2}} \sum_{|p|<\frac{P}{2}} G[p,m]\, z[n-p+m]\, z^*[n-p-m]\, e^{-j2\pi km/M} \qquad (6.5.1)$$

$$= 2 \underset{m\to k}{\mathrm{DFT}} \left\{ G[n,m] \underset{n}{*} (z[n+m]\, z^*[n-m]) \right\};\ m \in \langle M \rangle \qquad (6.5.2)$$

where the support dimensions of the kernel do not exceed M samples in the lag (m) direction and P samples in the time (n) direction, and $\langle M \rangle$ means any set of M consecutive integers; cf. [1, p. 444]. So the general procedure for evaluating such a TFD is:

1. Formation of the instantaneous autocorrelation function (IAF) $K_z[n,m] = z[n+m]\, z^*[n-m]$;
2. Discrete convolution in n (time) with the smoothing function $G[n,m]$;
3. Discrete Fourier transformation mapping m (lag) to k (frequency).

For the DWVD, which has $G[n,m] = \delta[n]$, step 2 reduces to an identity transformation and may be omitted. The windowed DWVD has $G[n,m] = \delta[n]\, g[m]$, so that step 2 reduces to multiplication of the IAF by $g[m]$. Some other quadratic TFDs, however, have special forms leading to computational procedures which are *not* degenerate cases of the above, and which may be simpler or faster.

This article addresses some of the practical issues in computing quadratic TFDs of a real signal, examines various cases of the above procedure, and considers the spectrogram as one example of a special form leading to a simpler, faster algorithm.

6.5.2 Computation of the Analytic Signal

The usual definitions of quadratic TFDs, especially the WVD and the windowed WVD, assume an analytic signal in order to avoid interference terms between positive and negative frequencies. For computational purposes, an analytic signal also avoids the need for $2\times$ oversampling prior to computation of the IAF (see Section 6.1.1 and ref. [2]). So, given a real signal $s(t)$, we must first compute the analytic signal $z(t)$ associated with $s(t)$. The simplest method is the direct approach of filtering out the negative frequencies in the frequency domain. If a real signal $s[n]$ is given for $n = 0, 1, 2, \ldots, N-1$ and periodically extended with period N, where N is even (or is made even by zero-padding), the algorithm is:

[0]Authors: **Boualem Boashash** and **Gavin R. Putland**, Signal Processing Research Centre, Queensland University of Technology, GPO Box 2434, Brisbane, Q 4001, Australia (b.boashash@qut.edu.au, g.putland@qut.edu.au). Reviewers: S. L. Marple, A. Reilly and V. Sucic.

1. Compute $S[k] = \mathrm{DFT}\{s[n]\}$ for $k = 0, 1, \ldots, N-1$;

2. Compute
$$Z[k] = \begin{cases} S[k] & \text{for } k = 0, \frac{N}{2} \\ 2\,S[k] & \text{for } k = 1, 2, \ldots, \frac{N}{2}-1 \\ 0 & \text{otherwise ;} \end{cases} \tag{6.5.3}$$

3. Compute $z[n] = \mathrm{IDFT}\{Z[k]\}$, where $\mathrm{IDFT}\{\ldots\}$ denotes the inverse DFT.

The treatment of the Nyquist term ($k = N/2$) and the precise meaning of "analytic" for a discrete-time periodic signal are explained in [3]; these issues become significant if the signal has non-zero amplitude at the Nyquist frequency. Further details on implementation of TFDs, including computation of the analytic signal, are given in [4]. A time-domain algorithm for computing the analytic signal using FIR filters is described in [5].

6.5.3 Real-Time Computation of TFDs

The formula for the discrete quadratic TFD [Eq. (6.5.2)] involves the expression $z[n+m]$ where m is allowed to be positive, together with $z^*[n-m]$ where m is allowed to be negative. The same applies to the DWVD

$$W_z[n,k] = 2 \underset{m \to k}{\mathrm{DFT}} \{z[n+m]\, z^*[n-m]\} \; ; \quad m \in \langle N \rangle \tag{6.5.4}$$

and the windowed DWVD

$$W_z^g[n,k] = 2 \underset{m \to k}{\mathrm{DFT}} \{g[m]\, z[n+m]\, z^*[n-m]\} \; ; \quad m \in \langle M \rangle \tag{6.5.5}$$

(both of these expressions are derived in Article 6.1). Both cases involve **time-advanced** signals; for any value of n, the computation of the TFD involves signal values up to $z[n+\lambda]$, where λ is some positive integer. In real-time computation, we cannot compute the TFD for time n until we know the signal values up to $z[n+\lambda]$; thus λ is the minimum **latency** of the computation. In the case of the DWVD [Eq. (6.5.4)], the range of m for which the IAF can be non-zero is maximized when n is at the center of the time-support of the signal; so the latency reaches a peak of half the signal duration. For the windowed DWVD, the latency is limited to half the window duration. For the general discrete quadratic TFD, the latency is limited to half the sum of the dimensions of the G matrix. The latency of the analytic signal computation must be added to that of the TFD computation. In all cases a smaller value of M not only reduces latency but also produces shorter FFTs, hence shorter computational delays; but the cost is reduced frequency resolution.

Latency is one of two measures of merit for real-time computation of TFDs. The other measure is throughput, which depends on the efficiency of the numerical algorithms. Eq. (6.5.2) can be written

$$\rho_z[n,k] = 2 \underset{m \to k}{\mathrm{DFT}} \{R_z[n,m]\} \tag{6.5.6}$$

where $R_z[n,m] = G[n,m] \underset{n}{*} (z[n+m]\, z^*[n-m])$. Similarly,

$$\rho_z[n+1,k] = 2 \underset{m \to k}{\mathrm{DFT}} \{R_z[n+1,m]\}. \tag{6.5.7}$$

The above two equations represent successive time-slices of the TFD. Multiplying the second equation by j, adding the result to the first equation and using the linearity of the DFT, we obtain

$$\rho_z[n,k] + j\rho_z[n+1,k] = 2 \underset{m \to k}{\mathrm{DFT}} \{R_z[n,m] + jR_z[n+1,m]\}. \tag{6.5.8}$$

If the TFD is known *a priori* to be real, as it usually is, then Eq. (6.5.8) means that the successive time slices of the TFD are respectively the real and imaginary parts of the right-hand side, which involves only one FFT [6]. Thus the realness property can enhance efficiency by halving the required number of FFTs. It can also halve the storage requirement as it implies Hermitian symmetry in the smoothed IAF.

6.5.4 Computational Approximations for Discrete-Time Kernels

Table 6.5.1 reproduces the "$G[n,m]$" column of Table 6.1.2 (p. 240) and adds two special cases often found in the literature: $\mathrm{BJ}_{1/2}$ denotes the Born-Jordan distribution with $\alpha = 1/2$, while ZAM_2 denotes the Zhao-Atlas-Marks distribution with $a = 2$. Some entries in the "$G[n,m]$" column of Table 6.5.1 call for continuous convolution prior to sampling. At best, the evaluation of such a convolution in the time-lag domain requires oversampling. At worst, it requires the numerical evaluation of an improper integral arising from a singularity in $G(t,\tau)$. In either case, computational inefficiencies will arise unless the smoothing effect of the convolution can be approximated in some other way. The chosen approximations, shown in the right-hand column of Table 6.5.1, are explained below.

In the case of the B-distribution, the sole purpose of the convolution is to avoid aliasing. Without the convolution, and for typical values of the parameter β (e.g. $\beta = 0.01$), the time-lag kernel would be a continuous function with a narrow slot at $m = 0$ caused by the factor $|2m|^\beta$. This factor is approximately unity for small nonzero values of m. The convolution fills in the slot, so that the factor is approximately unity at $m = 0$ also. This effect can be approximated by replacing $|2m|$ with $[4m^2+1]^{1/2}$, as is done in Table 6.5.1.

In the case of ZAM distribution, for a suitable (unbounded) $w[m]$, the convolution also ensures that $G[n,0] = \delta[n]$, which in turn verifies the TM property. Without the convolution, we would have

$$G_{\mathrm{ZAM}}[n,m] = w[m]\,\mathrm{rect}\left(\tfrac{an}{4m}\right) = \begin{cases} w[m] & \text{if } |an| \le |2m| \\ 0 & \text{otherwise.} \end{cases} \tag{6.5.9}$$

This gives $G[n,0] = w[0]\,\delta[n]$, which verifies the TM property provided that $w[0] = 1$. Accordingly, Eq. (6.5.9) is used in Table 6.5.1, although other approximations are possible. For example, we could sacrifice the TM property in favor of

Table 6.5.1: Computational approximations for time-lag kernels of selected discrete quadratic TFDs. In the "Distribution" column, subscripts indicate parameter values while the prefix "w-" means "windowed" by the function $w[m]$. For the spectrogram and w-Levin distributions, the window w is assumed to be real and even. The "$G[n,m]$" column shows the exact kernels required for the avoidance of aliasing in the Doppler-frequency domain. If $G[n,m]$ cannot be computed as written, the "Approx." column shows the suggested computational approximation.

Distribution	$G[n,m]$	Approx.
WVD	$\delta[n]$	
Levin	$\frac{1}{2}\delta[n+m] + \frac{1}{2}\delta[n-m]$	
BJ	$\left[\frac{1}{\lvert 4\alpha m\rvert}\,\mathrm{rect}\left(\frac{n}{4\alpha m}\right)\right]$ $**$ $[\mathrm{sinc}\, n\, \mathrm{sinc}\, m]$	$\approx \begin{cases} \frac{1}{\lvert 4\alpha m\rvert+1} & \text{if } \lvert 2n\rvert \le \lvert 4\alpha m\rvert + 1 \\ 0 & \text{otherwise.} \end{cases}$
$\mathrm{BJ}_{1/2}$	$\left[\frac{1}{\lvert 2m\rvert}\,\mathrm{rect}\left(\frac{n}{2m}\right)\right]$ $**$ $[\mathrm{sinc}\, n\, \mathrm{sinc}\, m]$	$\approx \begin{cases} \frac{1}{\lvert 2m\rvert+1} & \text{if } \lvert n\rvert \le \lvert m\rvert \\ 0 & \text{otherwise.} \end{cases}$
Modified B	$\frac{\cosh^{-2\beta} n}{\sum_n \cosh^{-2\beta} n}$	
w-WVD	$\delta[n]\, w[m]$	
w-Levin	$\frac{1}{2} w[m]\,(\delta[n+m] + \delta[n-m])$	
ZAM	$\left[w[m]\,\mathrm{rect}\left(\frac{an}{4m}\right)\right]$ $**$ $[\mathrm{sinc}\, n\, \mathrm{sinc}\, m]$	$\approx \begin{cases} w[m] & \text{if } \lvert an\rvert \le \lvert 2m\rvert \\ 0 & \text{otherwise.} \end{cases}$
ZAM_2	$\left[w[m]\,\mathrm{rect}\left(\frac{n}{2m}\right)\right]$ $**$ $[\mathrm{sinc}\, n\, \mathrm{sinc}\, m]$	$\approx \begin{cases} w[m] & \text{if } \lvert n\rvert \le \lvert m\rvert \\ 0 & \text{otherwise.} \end{cases}$
Rihaczek	$\delta[n-m]$	
w-Rihaczek	$w^*[-m]\,\delta[n-m]$	
Page	$\delta[n - \lvert m\rvert]$	
CW	$\frac{\sqrt{\pi\sigma}}{\lvert 2m\rvert}\exp\left(\frac{-\pi^2\sigma n^2}{4m^2}\right)$ $**$ $[\mathrm{sinc}\, n\, \mathrm{sinc}\, m]$	$\approx \begin{cases} \delta[n] & \text{if } m = 0 \\ \sqrt{\frac{\pi\sigma}{4m^2+\pi\sigma}}\exp\left(\frac{-\pi^2\sigma n^2}{4m^2+\pi\sigma}\right) & \text{otherwise.} \end{cases}$
B	$\left[\frac{\lvert 2m\rvert}{\cosh^2 n}\right]^{\beta} * \mathrm{sinc}\, m$	$\approx \left[\frac{\sqrt{4m^2+1}}{\cosh^2 n}\right]^{\beta}$
spectrogram	$w[n+m]\, w[n-m]$	

some smoothing by using the approximation

$$G_{\mathrm{ZAM}}[n,m] \approx \tfrac{1}{2} w[m]\big[1 + \tanh(|4m| - |2an|)\big]\ , \qquad (6.5.10)$$

and we could salvage the TM property by using a separate definition for $m = 0$.

For the Born-Jordan (BJ) and Choi-Williams (CW) distributions, the convolutions are needed to remove singularities at $m = 0$ and ensure that $G[n,0] = \delta[n]$. For the BJ distribution, we can remove the singularity and approximate the spreading in the $[n,m]$ plane by replacing $|4\alpha m|$ with $|4\alpha m|+1$. The result is

$$G_{\mathrm{BJ}}[n,m] \approx \tfrac{1}{|4\alpha m|+1}\,\mathrm{rect}\big(\tfrac{n}{|4\alpha m|+1}\big)\ , \tag{6.5.11}$$

which is equivalent to the rule given in Table 6.5.1. For the CW distribution, a similar effect is obtained by replacing $|2m|$ with $[4m^2+\pi\sigma]^{1/2}$. This step, by itself, gives the kernel

$$G_{\mathrm{CW}}[n,m] \approx \sqrt{\tfrac{\pi\sigma}{4m^2+\pi\sigma}}\,\exp\big(\tfrac{-\pi^2\sigma n^2}{4m^2+\pi\sigma}\big). \tag{6.5.12}$$

For $n = m = 0$, this reduces to $G[0,0] = 1$, which is consistent with the requirement that $G[n,0] = \delta[n]$. However, for $m = 0$, Eq. (6.5.12) reduces to $G[n,0] = e^{-\pi n^2}$, which is only an approximation to $\delta[n]$. Accordingly, a two-part definition of the kernel is used in the "Approx." column of the table. With $n = 0$, the kernel as defined in the table reduces to

$$G_{\mathrm{CW}}[0,m] = \sqrt{\tfrac{\pi\sigma}{4m^2+\pi\sigma}} \tag{6.5.13}$$

which takes the value 1 at $m = 0$ and $1/\sqrt{2}$ at $m = \pm\sqrt{\pi\sigma/4}$. For realistic values of σ (e.g. $\sigma \geq 1$), this gives a reasonable degree of smoothing in the m direction.

An alternative approach to the problem of singularities, which is not pursued here, is to evaluate the kernels in the Doppler-lag $[l,m]$ domain. This is efficient if we intend to evaluate the time-convolution by the FFT method, which also uses the $[l,m]$ domain. But it is still an approximation (if the time-lag form of the kernel is taken as the definition) because the analytical formulae for continuous FTs of standard signals are only approximations when applied to the DFT.

6.5.5 Special Case: Direct Form of the Discrete Spectrogram

The short-time Fourier transform (STFT) of the continuous-time signal $x(t)$ with real window $w(t)$ is defined (in Section 2.3.1) as

$$F_x^w(t,f) \triangleq \underset{\tau\to f}{\mathcal{F}}\{x(\tau)\,w(\tau-t)\} = e^{-j2\pi ft}\underset{\tau\to f}{\mathcal{F}}\{x(\tau+t)\,w(\tau)\} \tag{6.5.14}$$

$$= e^{-j2\pi ft}\int_{-\infty}^{\infty} x(\tau+t)\,w(\tau)\,e^{-j2\pi f\tau}d\tau. \tag{6.5.15}$$

It is shown in Chapter 2 that the spectrogram $S_x^w(t,f)$, which is simply the squared magnitude of the STFT, can also be considered as a quadratic TFD with kernel $w(t+\frac{\tau}{2})\,w(t-\frac{\tau}{2})$. The discrete form of this kernel is $w[n+m]\,w[n-m]$. Hence the discrete spectrogram can be conveniently evaluated using the general procedure described in Section 6.5.1 above. But it is simpler and more efficient to discretize the continuous spectrogram directly.

Theorem 6.5.1: If the spectrogram S_x^w is modified by ideally sampling $w(\tau)$ at

$$\tau = m/f_s \tag{6.5.16}$$

where m is an integer and f_s is the sampling rate, and if

$$w(\tau) = 0 \quad \text{for} \quad |\tau| \geq \tfrac{M}{2f_s} \tag{6.5.17}$$

where M is a positive integer, and if the modified TFD is denoted by $\widehat{S}_x^w$, then

$$\widehat{S}_x^w\big(\tfrac{n}{f_s}, \tfrac{kf_s}{M}\big) = \left| \sum_{|m|<M/2} x\big(\tfrac{m+n}{f_s}\big)\, w\big(\tfrac{m}{f_s}\big)\, e^{-j2\pi km/M} \right|^2. \tag{6.5.18}$$

Proof/explanation: When $w(\tau)$ is sampled, the integrand in Eq. (6.5.15) becomes

$$x(\tau+t)\, w(\tau)\, e^{-j2\pi f\tau} \sum_{m=-\infty}^{\infty} \delta\big(\tau - \tfrac{m}{f_s}\big) \tag{6.5.19}$$

so that the STFT becomes

$$\widehat{F}_x^w(t,f) = e^{-j2\pi ft} \sum_{m=-\infty}^{\infty} x\big(\tfrac{m}{f_s}+t\big)\, w\big(\tfrac{m}{f_s}\big)\, e^{-j2\pi fm/f_s}. \tag{6.5.20}$$

By Eqs. (6.5.16) and (6.5.17), the summation is restricted to $|m| < M/2$, giving a maximum of M terms.[1] The sampling in τ makes $\widehat{F}_x^w(t,f)$ periodic in f with period f_s, while the time-limiting in τ gives a frequency resolution of M bins per period. So it is convenient to let

$$f = kf_s/M \tag{6.5.21}$$

where k is an integer. With these restrictions, Eq. (6.5.20) becomes

$$\widehat{F}_x^w\big(t, \tfrac{kf_s}{M}\big) = e^{-j2\pi kf_s t/M} \sum_{|m|<M/2} x\big(\tfrac{m}{f_s}+t\big)\, w\big(\tfrac{m}{f_s}\big)\, e^{-j2\pi km/M}. \tag{6.5.22}$$

Putting $t = n/f_s$ to match the quantization of τ, then taking the squared magnitude of the discrete STFT, we obtain Eq. (6.5.18). ∎

With a change of notation, Eq. (6.5.18) becomes

$$S_x^w[n,k] = \left| \sum_{|m|<M/2} x[m+n]\, w[m]\, e^{-j2\pi km/M} \right|^2. \tag{6.5.23}$$

This $S_x^w[n,k]$ is the **discrete spectrogram** of the discrete-time signal $x[n]$ with window $w[m]$. If the summand is extended periodically in m with period M (i.e. extended periodically in τ with period M/f_s), we obtain

$$S_x^w[n,k] = \left| \sum_{m\in\langle M\rangle} x[m+n]\, w[m]\, e^{-j2\pi km/M} \right|^2 \tag{6.5.24}$$

[1] M terms for odd M; $M-1$ terms for even M.

where $\langle M \rangle$ denotes any set of M consecutive integers.[2] This may be written

$$S_x^w[n,k] = \left| \underset{m \to k}{\mathrm{DFT}} \{x[m+n]\, w[m]\} \right|^2 \; ; \quad m \in \langle M \rangle \,. \tag{6.5.25}$$

The time support of $S_x^w[n,k]$ is that of $x[n] * w[n]$, corresponding to $x(t) * w(t)$. If this has a duration not exceeding N samples, then the non-zero elements of the discrete spectrogram may be represented by an $N \times M$ real matrix. Only half of the M columns are needed for the non-negative frequencies, which are sufficient if $x(t)$ is real.

Eq. (6.5.23) involves $x[n+m]$ where $|m| < M/2$ and M is the window length in samples. So, in real-time computations, the latency of the discrete spectrogram computed by this formula is half the window length.

6.5.6 Sample Code Fragments

In view of the current popularity of MATLAB™, we illustrate this Article with some code fragments from the experimental MATLAB function `tlkern.m`, which computed all of the TFDs plotted in Article 5.7. The input parameters of the function specify the kernel in terms of a time-dependent factor $g_1[n]$, a lag-dependent factor $g_2[m]$, and an "auxiliary factor" $g_3[n,m]$. The overall time-lag kernel $G[n,m]$ is then computed as

$$G[n,m] = g_2[m]\big(g_1[n] \underset{n}{*} g_3[n,m]\big) = \big(g_2[m]\, g_1[n]\big) \underset{n}{*} g_3[n,m]. \tag{6.5.26}$$

This scheme allows the computation of a wide variety of quadratic TFDs in under 320 lines of code, including exceptions for direct computation of the discrete spectrogram.

6.5.6.1 Example 1: MBD (General Algorithm)

For a separable kernel, the auxiliary factor would normally be omitted (i.e. taken as $\delta[n,m]$), while the time-dependent and lag-dependent factors would have input parameters specifying their types (e.g. Hamming or Hanning) and durations (in samples). Although the kernel of the modified B-distribution (MBD) is separable (see Article 5.7 and Table 6.5.1), its parameter β is *not* a duration. The MBD kernel is therefore specified using the factors

$$g_1[n] = \delta[n] \; ; \quad g_2[m] = 1 \; ; \quad g_3[n,m] = \frac{\cosh^{-2\beta} n}{\sum_n \cosh^{-2\beta} n} \,. \tag{6.5.27}$$

Notice that the auxiliary factor is the complete kernel.

To compute the MBD in Fig. 5.7.2(e) on p. 220, the function `tlkern` is called with the following significant parameters:

[2] For even M, the periodic extension is padded with a zero term.

```
s  = signal vector
N  = 128 = assumed period
tr = 1 = time resolution
tf = 'delta' = string specifying g1[n]
lf = '1' = string specifying g2[m]
af = 'mb' = string specifying form of g3[n,m]
ap = 0.2 = auxiliary parameter (β).
```

All internal computations, including IAF generation, are designed to be valid for periodic signals. Therefore, to compute the IAF of a *non*-periodic signal such as the one in Fig. 5.7.2(e), the assumed period `N` must be at least twice the signal length to avoid wrap-around effects. Because the time support of the IAF is identical to that of the signal, the same value of `N` is also sufficient to avoid wrap-around in the subsequent convolution with $g_1[n]$.

The output is the real matrix `tfd(1:Mpad+1,1:Nsel)`, whose dimensions `Mpad+1` and `Nsel` are assigned early by the statements

```
Mpad = 2^ceil(log(2*M)/log(2));  % lag-to-frequency FFT length
Ncut = min(N,length(s));         % duration of TF plot
Nsel = ceil(Ncut/tr);            % no. traces in TF plot
```

where `M` is the support length of the kernel in the lag direction; in this case `M` has been set to `length(s)` because of the constant "lag-dependent" factor.

Preliminaries: The analytic signal is computed by the frequency-domain method. If `N` is even, the Nyquist term has MATLAB index `N/2+1` and the amplitude at that frequency is left unchanged [3]. If `N` is odd, there is no Nyquist term. The following code handles both cases:

```
Noff = fix(N/2);
z = fft(real(s),N);              % s truncated or padded
z(2:N-Noff) = 2*z(2:N-Noff);     % positive frequencies
z(Noff+2:N) = 0;                 % negative frequencies
z = ifft(z);
```

For this kernel, the time-dependent factor `g1` and the lag-dependent factor `g2` are computed by the statements

```
g1(1:N) = 0;
...
g1(1) = 1;
...
g2(1:Mpad) = 1;
```

where "..." denotes one or more line(s) of control code, or code that is skipped in this case. The auxiliary factor `g3` (the whole kernel in this case) is computed by

```
Moff = fix(M/2);
...
g3(1:N,1:Mpad) = 0;
...
```

```
temp(1:N) = 0;
for n = -Noff:Noff
  temp(1+rem(N+n,N)) = (cosh(n))^(-2*ap);
end
temp = temp/sum(temp); % normalize
for m = -Moff:Moff
  g3(:,1+rem(Mpad+m,Mpad)) = temp.';
end
```

where ap denotes the auxiliary parameter (β), and the remainder (rem) function causes high array indices to represent negative values of time and lag.

Step 1—Formation of the IAF: The IAF matrix K(1:N,1:Mpad) is formed by

```
for n = 1:N
  for m = -Moff:Moff
  % K(n,m) = z(n+m)z^*(n-m), with corrected indices:
    K(n,1+rem(Mpad+m,Mpad)) = z(1+rem(2*N+n+m-1,N))*conj(z(1+rem(2*N+n-m-1,N)));
  end
end
```

where the "corrected" indices allow handling of periodic signals.

Step 2—Convolution in time: The assembly of the time-lag kernel and the convolution in time with the IAF are performed together. The smoothed IAF is

$$R_z[n,m] = K_z[n,m] \underset{n}{*} G[n,m] = K_z[n,m] \underset{n}{*} \big(g_2[m]\, g_1[n] \underset{n}{*} g_3[n,m]\big). \qquad (6.5.28)$$

The above convolutions may be taken as circular if the assumed period is sufficiently long, in which case

$$R_z[n,m] = \underset{n\leftarrow l}{\mathrm{IDFT}} \left\{ \underset{n\to l}{\mathrm{DFT}} \{K_z[n,m]\} \underset{n\to l}{\mathrm{DFT}} \{g_3[n,m]\} \underset{n\to l}{\mathrm{DFT}} \{g_1[n]\, g_2[m]\} \right\}. \qquad (6.5.29)$$

This is implemented by the following code, in which K(:,mcorr) is initially the m^{th} column of the IAF, but is overwritten by the m^{th} column of the *smoothed* IAF:

```
for m = -Moff:Moff
  mcorr = 1+rem(Mpad+m,Mpad);
  ...
  K(:,mcorr) = ifft(fft(K(:,mcorr)).*fft(g3(:,mcorr)).*fft(g1.'*g2(mcorr)));
  ...
end
```

(The factor $g_2[m]$ could be taken outside the IDFT, but this would not improve the efficiency of the code because g2(mcorr) is a scalar.) The FFT method of convolution is useful in *experimental* code because of its generality, but is not necessarily the most efficient method, especially if one of the convolved sequences is short.

Now we apply the time-resolution (tr):

```
for nsel = 1:Nsel
% nselth column of r is selected row of K:
  n = 1+tr*(nsel-1);
  r(:,nsel) = K(n,:).';
end
```

Step 3—DFT: The final DFT (lag to frequency) is computed by

```
r = fft(r);
```

which, for the sake of generality, does *not* take advantage of realness.

Final adjustments: The following code scales the TFD and repeats its first row (the zero-frequency row) so that the TFD spans a full cycle in the frequency domain:

```
tfd = [real(r);real(r(1,:))].*(Ncut/Nsel/Mpad);
```

The scaling ensures that the sum of the matrix elements is close to the signal energy regardless of the time resolution.

6.5.6.2 Example 2: Spectrogram (Special Case)

The spectrogram in Fig. 5.7.3(f) on p. 221 was computed by the same function `tlkern`. For the spectrogram, the parameters `s`, `N` and `tr` are the same as for the MBD, while `tf` is ignored. Other significant parameters are

```
lf = 'rect' = string specifying type of window
 M = 17 = window length (in samples)
af = 'sg' = string calling for spectrogram.
```

The output is `tfd(1:Mpad/2+1,1:Nsel)`, where the dimensions are assigned as for the MBD, except that the window duration `M` is read as an input parameter and *not* overwritten.

The analytic signal is computed as for the MBD, although this is not strictly necessary for the spectrogram.

The rectangular window is computed by

```
Moff = fix(M/2);
g2(1:Mpad) = 0;
...
for m = -Moff:Moff
  g2(1+rem(Mpad+m,Mpad)) = 1;
end
```

The matrix `K(1:N,1:Mpad)` normally represents the IAF, but for the spectrogram it is assigned differently:

```
for n = 1:N
  for m = -Moff:Moff
  % K(n,m) = z(n+m)g2(m), with corrected indices:
    K(n,1+rem(Mpad+m,Mpad)) = z(1+rem(2*N+n+m-1,N))*g2(1+rem(Mpad+m,Mpad));
  end
end
```

The code that applies the time resolution and performs the final DFT (lag to frequency) is the same as for the MBD. But the final adjustment is different:

```
tfd = (abs(r(1:Mpad/2+1,:))).^2.*(Ncut/Nsel/Mpad/sum(g2.^2));
```

The magnitude-squared operation alters the relationship between the window and the scaling of the TFD. Also note that the above step uses only half the columns of the Fourier-transformed r matrix, namely those corresponding to the non-negative frequencies. Efficiency could be further improved by exploiting the analytic signal to halve the sampling rate.

6.5.7 The *TFSA* package

The *Time-Frequency Signal Analysis* (*TFSA*) package is a set of functions developed over more than a decade at the Signal Processing Research Centre, Queensland University of Technology, for computing modulated signals, quadratic and polynomial TFDs, ambiguity functions, wavelet transforms and scalograms, and various estimates of instantaneous frequency. As this is a production package rather than an experimental package, computationally intensive functions are precompiled and optimized for efficiency, and an interactive user interface is added. The current version is distributed as a MATLAB toolbox, so that TFSA functions can be used with other computational and graphical functions of MATLAB. Further information is available at http://www.sprc.qut.edu.au/ or http://www.eese.bee.qut.edu.au/research/spr/ .

6.5.8 Summary and Conclusions

High-level programming languages with built-in FFT functions and matrix operations have made it possible to construct compact yet highly versatile functions for computing quadratic TFDs. Use of a common algorithm for all TFDs is convenient for the programmer. But, as illustrated by the direct form of the spectrogram, efficiency can sometimes be improved by using different algorithms in special cases.

References

[1] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[2] B. Boashash, "Note on the use of the Wigner distribution for time-frequency signal analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 36, pp. 1518–1521, September 1988.

[3] S. L. Marple Jr., "Computing the discrete-time "analytic" signal via FFT," *IEEE Trans. Signal Processing*, vol. 47, pp. 2600–2603, September 1999.

[4] B. Boashash and A. Reilly, "Algorithms for time-frequency signal analysis," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 7, pp. 163–181, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[5] A. Reilly, G. Frazer, and B. Boashash, "Analytic signal generation—Tips and traps," *IEEE Trans. Signal Processing*, vol. 42, pp. 3241–3245, November 1994.

[6] B. Boashash and P. J. Black, "An efficient real-time implementation of the Wigner-Ville distribution," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 35, pp. 1611–1618, November 1987.

Chapter 7

Measures, Performance Assessment and Enhancement

This chapter describes a number of time-frequency methods specifically developed for performance enhancement for a given application. The adopted performance measures are defined using objective criteria. The topic is covered in five articles.

Hyperbolic FM signals are well described by the Affine Wigner-Ville distribution, a method related to time-scale analysis and the wavelet transform (Article 7.1). A general procedure for enhancing the time-frequency resolution and readability of TFDs is the reassignment principle (7.2). Techniques for measuring the concentration of TFDs and for automatic optimization of their parameters may be based on entropy measures (7.3). Another approach defines a resolution performance measure using local measurements in the (t, f) domain, such as relative amplitudes of auto-terms and cross-terms (7.4). Finally, attempts to unify time-frequency, time-scale, filter banks, wavelets and the discrete-time Gabor transform using product functions and cascaded frames may assist in the selection of the best-performing method for a given application (7.5).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

7.1 TIME-FREQUENCY ANALYSIS BASED ON THE AFFINE GROUP.[0]

7.1.1 Scale Transformations and their Constructive Role in the Time-Frequency Analysis of Real Signals

The notion of time-frequency analysis is very familiar from the experience of hearing perception and its practical interest in signal theory seems a common sense topic. However, at this level of generality, there is no indication for giving an analytic content to the subject and, in fact, numerous techniques have been proposed (see Chapters 2 and 5). The introduction of the affine group allows to overcome this lack of universality by adding a consistency condition on which a common consent is possible.

At first it is necessary to recognize that signals describe the time evolution of a physical quantity (an acoustic pressure for example) and that the measurement of this quantity can be done only after a system of reference and a system of physical units have been chosen. In fact, depending on these choices, different descriptions (i.e. different signals) can be associated with the same physical phenomenon. Obviously the communication theory has to deal with the situation by interpreting equivalently the various options. We will show that the affine group is the right mathematical tool for managing this equivalence.

In signal theory there is only one reference variable which is the time. Hence the changes of reference system are naturally interpreted as clock changes involving changes of time origin and changes of time unit. The analytical description of any change is thus given by a transformation of the form:

$$t \longrightarrow t' = at + b, \tag{7.1.1}$$

where t and t' are the time variables and where a and b are real numbers with $a > 0$. The set of all such transformations constitutes the affine group.

These transformations will also affect physical units (those which are derived from the time unit) by multiplying each of them by some power of the dilation a. As a result a change of clock of the form (7.1.1) will induce on a given signal $s(t)$ a change of the form:

$$s(t) \longrightarrow s'(t) = a^r s(a^{-1}(t-b)), \tag{7.1.2}$$

where r is a real exponent depending on the physical nature of the signal under study. The exponent r will be called index of (dimensional) conformity.

The above remarks are only useful preliminaries before tackling the time-frequency problem. Actually time-frequency analysis is an alternative description in which signals are not characterized by their instantaneous values but by their form which is described in terms of modulations of amplitude and frequency. In practice, this description is summarized by a real function $P(t, f)$ which gives an image of

[0]Authors: **J. Bertrand**, LPTMC, University Paris VII, Case 7020, 75251 Paris Cedex 05, France (bertrand@ccr.jussieu.fr), and **P. Bertrand** (pibert@libertysurf.fr).

the spreading of the signal in the time-frequency half-plane ($f > 0$). The basic problem of time-frequency analysis is to make this approach effective by setting up the correspondence between the physical signal $s(t)$ and its time-frequency representation $P(t, f)$. It is well-known that the operation cannot be linear and the usual method consists in defining $P(t, f)$ as an Hermitian functional on a Hilbert space whose elements are in direct relation with the physical signals. In affine theory this classical approach is followed, the specific point being the introduction of a Hilbert space where an irreducible unitary representation of the affine group does exist.

Any real signal $s(t)$ is entirely characterized by the positive frequency part of its Fourier transform, i.e. by the function:

$$Z(f) \equiv Y(f) \int_{-\infty}^{\infty} e^{-j2\pi ft}\, s(t)\, dt, \tag{7.1.3}$$

where $Y(f)$ is the Heaviside step function.

In agreement with (7.1.2), a general affine transformation on signal $s(t)$ acts on $Z(f)$ according to:

$$Z(f) \longrightarrow Z'(f) \equiv a^{r+1}\, e^{-j2\pi bf}\, Z(af), \tag{7.1.4}$$

where r is the index of conformity of the physical signal. In contrast with (7.1.2), the transformation (7.1.4) corresponds to an irreducible representation of the affine group. Moreover, as can be verified, it conserves the norm defined by:

$$||Z||^2 \equiv \int_0^{\infty} |Z(f)|^2 f^{2r+1}\, df. \tag{7.1.5}$$

In the following, we will adopt this norm for which the representation (7.1.4) of the affine group is unitary.

In a change of reference system, the time-frequency representation of a signal will also be changed. This change will depend on the change in the variables (cf. (7.1.1)) and on the meaning of the function $P(t, f)$ as a quantitative representation of the spreading of the signal in the time-frequency half-plane. This leads to introduce the general transformation:

$$P(t, f) \longrightarrow P'(t, f) \equiv a^q P(a^{-1}(t - b), af), \tag{7.1.6}$$

where q is a real number which can be considered as an index of significance. The integral of $P(t, f)$ on the half-plane that is invariant by transformation (7.1.6) has the form:

$$\int f^q\, P(t, f)\, dt\, df. \tag{7.1.7}$$

For the common choice $q = 0$, a probabilistic comprehension of the spreading is possible. However, in special applications, the choice of other values for q can be pertinent.

The occurrence of the dimensional indices r and q is a new feature that has been introduced by the consideration of scaling transformations. It did not appear in the case of the Wigner-Ville representation which is based, in an essential way, on translations in both time and frequency through the so-called Weyl-Heisenberg group. In fact, it can be seen directly that the invariance of the scalar product associated with the norm (7.1.5) and of the integral (7.1.7) under frequency translations is only possible if $r = -1/2$ and $q = 0$ respectively.

7.1.2 Tomographic Derivation of the Affine Wigner Function

Tomographic methods are now of routine use in signal analysis. Their emergence in time-frequency analysis is related to the recognition that the general marginals of the well-known Wigner-Ville function along straight lines of arbitrary slope, are positive and easily interpretable in terms of the signal. An immediate result has been the construction of the Wigner-Ville function by relying only on the Heisenberg group which performs time and frequency translations [1,2]. Practical applications have arisen in the domain of quantum optics, where the method of tomography [3] allows to determine precisely the state of a system from experimental data [4]. More recently, there has been a renewed interest for the so-called fractional Fourier transform arising naturally in the expression for the Radon transform of the Wigner-Ville function (see Articles 4.8 and 5.8).

In fact, the relation between tomography and Wigner function is not accidental but is connected in an essential way to the underlying group. In signal analysis, the straight lines of a given direction in the time-frequency plane arise as families of curves invariant by a subgroup of the translations group. In the affine case, a tomography can be set up in an analogous way by considering all the subgroups which are acting as dilations around a fixed time. This will be done now.

The linear chirps play a major role in usual tomography since they are associated with straight lines in the time-frequency plane. They form families of signals which are characterized by their invariance, up to phase, under the subgroup of translations in a given direction. In the affine case, the same role is played by signals that are invariant (up to a phase) in dilations centered at the instant $t = \xi$. These are the hyperbolic chirps defined by:

$$\psi_\beta^\xi(f) = f^{-j2\pi\beta - r - 1} e^{-j2\pi\xi f}, \quad \beta \text{ real.} \tag{7.1.8}$$

It can be verified that such signals conserve the same form in the transformation of type (7.1.4) submitted to the constraint:

$$b = \xi(1 - a). \tag{7.1.9}$$

These transformations are just those of the subgroup of dilations of origin $t = \xi$.

A signal $Z(f)$, with index of conformity r, can be decomposed on the $\psi_\beta^\xi(f)$, ξ fixed, and the coefficients of the development are equal to:

$$\left(Z, \psi_\beta^\xi\right) = \int_0^\infty Z(f)\, f^{j2\pi\beta + r}\, e^{j2\pi\xi f}\, df, \tag{7.1.10}$$

where the scalar product is defined from (7.1.5). The expression (7.1.10) is a general Mellin transform whose inverse is easily obtained.

The group delays of the hyperbolic chirps (7.1.8) at frequency f are found to be

$$t = \xi + \beta/f, \quad \beta \text{ real.} \tag{7.1.11}$$

They correspond to curves in the time-frequency half plane (t, f), $f > 0$, that are invariant by dilation (7.1.9).

Let $P^q(t, f; Z)$ denote the time-frequency representation of signal $Z(f)$ that is under construction. The tomographic condition relates the integral of $P^q(t, f)$ along hyperbolas (7.1.11) to the coefficients (7.1.10) of signal $Z(f)$ on the basis of hyperbolic chirps. It takes the form:

$$\int_{-\infty}^{\infty} \int_{0}^{\infty} P^q(t, f; Z)\, \delta((t-\xi)f - \beta)\, f^q\, dt\, df = \left| \left(Z, \psi_\beta^\xi \right) \right|^2, \tag{7.1.12}$$

where each member is separately invariant by affine transformations. Considered for all real values of β and ξ, this relation has the form of a generalized Radon transform for $P^q(t, f)$. Its inversion gives the expression of the *Affine Wigner Function* [1,5]:

$$\begin{aligned} P^q(t, f; Z) \;=\; & f^{2r-q+2} \int_{-\infty}^{\infty} e^{j2\pi u f t} \\ & \times\; Z\left(\frac{f u e^{u/2}}{2\sinh(u/2)} \right) Z^*\left(\frac{f u e^{-u/2}}{2\sinh(u/2)} \right) \left(\frac{u}{2\sinh(u/2)} \right)^{2r+2} du. \end{aligned} \tag{7.1.13}$$

The integrand can be seen to depend only on the functions $\lambda(u)$ and $\lambda(-u)$ where:

$$\lambda(u) = \frac{u e^{u/2}}{2\sinh(u/2)}. \tag{7.1.14}$$

In formula (7.1.13), the index of significance q can be chosen according to the type of time-frequency representation needed, but r is necessarily equal to the index of conformity attached to $Z(f)$.

When an affine change is performed on signal $Z(f)$ according to (7.1.4), the function (7.1.13) is clearly transformed as in (7.1.6). Hence, the Affine Wigner Function verifies the condition:

$$P^q(t, f; Z') = a^q\, P^q(a^{-1}(t-b), af; Z), \tag{7.1.15}$$

where $Z'(f)$ is defined in (7.1.4). This property of invariance of the correspondence between a signal $Z(f)$ and its time-frequency representation $P^q(t, f)$ is essential for an analysis founded on the affine group. However, it does not characterize univocally the Affine Wigner Function. In fact, an infinite family of functions satisfying condition (7.1.15) can be determined: They form the affine analog of Cohen's class.[1] But the Affine Wigner Function (7.1.13) stands out as the only one

[1] That is, the quadratic class; see p. 68n.

with a tomographic construction based on the affine group alone. In this sense, it occupies the same position in the affine class as the usual Wigner-Ville function does in Cohen's class.

7.1.3 Discussion of the Result in terms of Broad-Band Corrections for the Usual Wigner Function

It is instructive to start using the Affine Wigner Function (7.1.13) for the analysis of narrow-band signals. In that case, the factor f^{2r-q+1} is approximately constant, with f replaced by the central frequency f_0 of the band. Moreover, the integrand in (7.1.13) is different from zero only when the arguments of Z and Z^* are both in the band. This requires for u to stay in the vicinity of the value $u = 0$. In fact, it can be observed that the extension of the domain of u will decrease along with the bandwidth of the signal. This allows to replace the function $\lambda(u)$ by its first order development about $u = 0$:

$$\lambda(u) \simeq 1 + u/2, \tag{7.1.16}$$

when analyzing narrow-band signals.
A change of variables from u to $v \equiv uf$ then leads to the approximate form:

$$P^q(t,f) \approx f_0^{2r-q+1} \int_{-\infty}^{\infty} e^{j2\pi vt} Z\left(f + \frac{v}{2}\right) Z^*\left(f - \frac{v}{2}\right) dv. \tag{7.1.17}$$

Thus for narrow-band signals, the expression of the Wigner-Ville function appears as a simplified expression of the Affine Wigner Function.

In the general case, it remains interesting to study the properties of the function (7.1.13) in relation with those of the usual Wigner-Ville function.

The total integral of $P^q(t,f)$ over the time-frequency space is now written in the invariant form:

$$\int_{-\infty}^{\infty} \int_{0}^{\infty} P^q(t,f;Z)\, f^q\, dt\, df = ||Z||^2, \tag{7.1.18}$$

where $||Z||$ is defined in (7.1.5).

Integrating the distribution on the time only, we get:

$$\int_{-\infty}^{\infty} P^q(t,f;Z)\, dt = f^{2r+1-q}\, |Z(f)|^2. \tag{7.1.19}$$

It can be observed that it is only for the choice $q = 2r+1$ that the familiar expression is exactly recovered.

Since $P^q(t,f)$ represents a spreading of the signal in the time-frequency plane, it can be used to compute the mean value of the epoch t for a fixed value of the frequency. The resulting expression, whatever the value of q, is the usual group delay of the signal:

$$\left(\int_{-\infty}^{\infty} tP^q(t,f;Z)\, dt\right) \left(\int_{-\infty}^{\infty} P^q(t,f;Z)\, dt\right)^{-1} = -(1/2\pi)\frac{d\theta(f)}{df}, \tag{7.1.20}$$

where the phase $\theta(f)$ of the analytic signal $Z(f)$ is defined by:

$$Z(f) = |Z(f)|\, e^{j\theta(f)}. \tag{7.1.21}$$

A domain where it is particularly important to be careful with the indices r and q is in the study of the effect of transducing filters on the time-frequency representations. Suppose a signal $Z(f)$, having an index of conformity equal to r, passes through a stationary linear device whose action changes its physical nature. The emerging signal $\widetilde{Z}(f)$ can be written as:

$$\widetilde{Z}(f) = T(f)\, Z(f), \tag{7.1.22}$$

where the function $T(f)$ characterizes the device. In general, $\widetilde{Z}(f)$ will transform under dilations with an index of conformity $\tilde{r}$ different from r. Hence, for consistency, the function $T(f)$ must be assigned an index α determined by relation (7.1.22) as:

$$\alpha = \tilde{r} - r - 1. \tag{7.1.23}$$

Notice that the identity filter corresponds here to $\alpha = -1$.

These operations have a counterpart in the domain of time-frequency functions. Let $P^q_{(r)}(t, f; Z)$, $P^\gamma_{(\alpha)}(t, f; T)$ and $P^{\tilde{q}}_{(\tilde{r})}(t, f; \widetilde{Z})$ be the representations corresponding to functions $Z(f)$, $T(f)$ and $\widetilde{Z}(f)$ respectively. The dependence of the representations on the indices of conformity of the signals has been shown explicitly, to avoid ambiguity. The time-frequency analog of relation (7.1.22) has the form of a convolution in time provided the indices are related according to:

$$\tilde{q} = q + \gamma + 1. \tag{7.1.24}$$

In practical situations, where $\tilde{q} = q$ and $\gamma = -1$, the relation is:

$$P^q_{(\tilde{r})}(t, f; \widetilde{Z}) = \int_{-\infty}^{\infty} P^{-1}_{(\alpha)}(t - t', f; T)\, P^q_{(r)}(t', f; Z)\, dt', \tag{7.1.25}$$

where the index α is given by (7.1.23).

Consider now the case of a transducer defined by:

$$T(f) = f^{-h},\ h \text{ real}, \tag{7.1.26}$$

with conformity index $\alpha = h - 1$. The only action of such a device is to replace the signal $Z(f)$ with index of conformity r by the signal $Z'(f) = f^{-h}\, Z(f)$ with index $\tilde{r} = r + h$. Relation (7.1.25) now becomes:

$$P^q_{(r+h)}(t, f; f^{-h} Z) \equiv P^q_{(r)}(t, f; Z), \tag{7.1.27}$$

an identity that can also be verified directly on the expression (7.1.13). Thus, once q has been chosen, computing $P^q(t, f)$ for a signal $Z(f)$ or for any of its transforms (as defined by (7.1.22) and (7.1.26)) yields the same result, provided care is taken to use the correct conformity indices. In the present context, the property leads to consider the device defined by (7.1.26) as a perfect transducer.

7.1.4 Hyperbolic Chirps and Affine Group Extension

The properties of marginalization, localization and extended invariance that are well-known for the Wigner-Ville function have direct analogs in the present case.

The marginal condition leading to a density in f is derived from relation (7.1.19). In addition, marginal densities in β are obtained for each value of ξ when integrating $P^q(t,f)$ along hyperbolas $(t-\xi)f = \beta$. From the tomographic condition (7.1.12) it results that these densities have the form:

$$\rho_\xi(\beta) = |(Z, \psi_\beta^\xi)|^2 \tag{7.1.28}$$

and verify:

$$\int_{-\infty}^{\infty} \rho_\xi(\beta)\, d\beta \;=\; ||Z||^2. \tag{7.1.29}$$

Localization in the time-frequency space arises for general hyperbolic chirps. For these signals, a direct computation gives:

$$P^q(t,f;\psi_\beta^\xi) = f^{-q}\,\delta((t-\xi)f - \beta). \tag{7.1.30}$$

The special case $\beta = 0$ corresponds to a localized signal $\psi_0^\xi(f) = f^{-r-1}e^{-j2\pi\xi f}$ attached to the instant $t = \xi$. The latter form can be obtained directly by requiring that after a clock change, labeled by (a,b) and acting as in (7.1.4), the signal keeps a localized form at the transformed instant $t' = a\xi + b$.

Another case of localization, that can be considered as a limit of the hyperbolic chirp behavior, concerns the pure frequency signal with index of conformity r:

$$Z_{f_0}(f) \equiv f^{-r}\,\delta(f - f_0) \tag{7.1.31}$$

which is represented by:

$$P^q(t,f;Z_{f_0}) = f^{1-q}\,\delta(f - f_0). \tag{7.1.32}$$

The property of invariance, under affine transformations, of the correspondence between a signal $Z(f)$ and its representation $P^q(t,f;Z)$ can be extended. To this end, we introduce the transformations performing a translation on the β-parameter of the hyperbolic chirps according to:

$$\psi_\beta^\xi(f) \longrightarrow \psi_{\beta+c}^\xi(f) = f^{-j2\pi(\beta+c)-r-1}\, e^{-j2\pi\xi f} \tag{7.1.33}$$

where c is a real number. These transformations act on an arbitrary signal as:

$$Z(f) \longrightarrow Z_c(f) = f^{-j2\pi c}\, Z(f). \tag{7.1.34}$$

Moreover, they combine with the affine transformations to form a three-parameter group G_0 which is is the largest group conserving the family of hyperbolas (7.1.11)

as a whole and the family of hyperbolic chirps. These general transformations, labeled by (a, b, c), act on a signal according to:

$$Z(f) \longrightarrow Z_{a,b,c}(f) = a^{r+1} e^{-j2\pi(bf + c\ln f)}\, Z(f), \tag{7.1.35}$$

and on its Affine Wigner Function as:

$$P^q(t, f; Z) \longrightarrow P^q(t, f; Z_{a,b,c}) = a^q P^q\left(a^{-1}(t - b - cf^{-1}), af; Z\right). \tag{7.1.36}$$

Thus the correspondence between the signal $Z(f)$ and the function $P^q(t, f; Z)$ is invariant not only by the affine group but by its extension G_0.

7.1.5 Unitarity Property and Some of its Consequences

The Affine Wigner Function (7.1.13) verifies the so-called unitarity (or Moyal) property:

$$\int_{-\infty}^{\infty}\int_{0}^{\infty} P^q(t, f; Z)\, P^q(t, f; Z')\, f^{2q}\, dt\, df = \left|(Z, Z')\right|^2, \tag{7.1.37}$$

where the scalar product (Z, Z'), issued from definition (7.1.5), depends on the index r of the signal.

A special case of relation (7.1.37) is obtained when $Z'(f)$ is the hyperbolic chirp $\psi_\beta^\xi(f)$ (cf. (7.1.8)) so that $P^q(t, f; Z')$ has the form (7.1.30). The result is just the tomographic condition which was introduced in (7.1.12).

A more general form of the unitarity property (see [6], formulas (III.15-17)) finds a direct application in the reconstruction of the signal from its affine Wigner function. In fact, it allows to write directly the formula:

$$Z(f_1)\, Z^*(f_2) = (f_1 f_2)^{-2r-1} \int f^{2r+2+q}\, P^q(t, f; Z)\, e^{j2\pi u f t} \tag{7.1.38}$$

$$\delta\left(f_1 - \frac{f u e^{-u/2}}{2\sinh(u/2)}\right) \delta\left(f_2 - \frac{f u e^{u/2}}{2\sinh(u/)2}\right) \left(\frac{u}{2\sinh(u/2)}\right)^{2r+2} du\, dt\, df,$$

where, as usual, the symbols δ hold for Dirac distributions.

The knowledge of $\rho(f_1, f_2) = Z(f_1)Z^*(f_2)$ allows to reconstruct the analytic signal up to a constant phase. Explicitly, we have:

$$|Z(f)| = \sqrt{\rho(f, f)}\,, \qquad e^{i(\theta_1 - \theta_2)} = \frac{\rho(f_1, f_2)}{\rho(f_1, f_1)\,\rho(f_2, f_2)}\,, \tag{7.1.39}$$

where the decomposition (7.1.21) of $Z(f)$ has been used. This result shows that the affine Wigner function (7.1.13) is a representation of the signal which does not discard any information but a constant phase.

The unitarity relation can also be used to define a regularized version $\widetilde{P}^q(t, f)$ of $P^q(t, f)$. To this end, a basic function $\Phi(f)$ is chosen and its representation

$P^q(t,f;\Phi)$ written down. The smoothed version $\widetilde{P}^q(t,f;Z)$ of the representation $P^q(t,f;Z)$ is then defined by the convolution on the affine group:

$$\widetilde{P}^q(t,f;Z) = f^{-q}\int_{-\infty}^{\infty}\int_{0}^{\infty} P^q(t',f';Z)\,P^q(f(t'-t),f'/f;\Phi)\,f'^{2q}\,dt'\,df', \quad (7.1.40)$$

where the kernel is the time-frequency representation of function Φ. A more practical form of the smoothed function $\widetilde{P}^q$ is obtained when taking into account the transformation law (7.1.15) of $P^q(t,f;\Phi)$ and the unitarity property (7.1.37):

$$\widetilde{P}(t,f;Z) = \left|(Z,\Phi_{(t,f)})\right|^2, \quad (7.1.41)$$

where $\Phi_{(t,f)}(f')$ is defined according to:

$$\Phi_{(t,f)}(f') = f^{-r-1}e^{-j2\pi f't}\,\Phi(f'/f). \quad (7.1.42)$$

The set of functions $\Phi_{(t,f)}(f')$ is recognized as a family of wavelets obtained from the mother wavelet $\Phi(f')$ by an affine transformation such that $a = f^{-1}$, $b = t$. Thus the right-hand side of relation (7.1.41) is the square modulus of the wavelet coefficient of $Z(f)$. Conversely, the above developments allow to interpret the square modulus of the wavelet coefficient as the result of a smoothing in the time-frequency half-plane. In particular, this provides a guide to assess the properties of a mother wavelet.

For example, the function $\Phi(f)$ may be chosen so that its representation $P^q(t,f;\Phi)$ has the best possible concentration in the time-frequency plane. An optimal choice for $\Phi(f)$ is found to be the Klauder wavelet [5]. It can be written using as parameters the mean frequency $f_0 =< f >$, the mean square deviation $\sigma_f \equiv< f^2 > -f_0^2$, and the relative bandwidth $\rho = \sigma_f/f_0$. The explicit expression of that wavelet is:

$$\Phi(f) \equiv K(\rho)\,f_0^{-r-1}\left(\frac{f}{f_0}\right)^{-r-1+\frac{1}{2\rho^2}}\exp\left(-\frac{1}{2\rho^2}\frac{f}{f_0}\right), \quad (7.1.43)$$

where $K(\rho)$ is a normalization constant.

The function (7.1.43) has a time-frequency representation approximately localized in the vicinity of point $t = 0$, $f = f_0$. By varying σ_f, it is possible to make it spread along one or the other of the time and frequency directions, without changing the point it is attached to.

7.1.6 Summary and Conclusions

The group of affine transformations on the time, or clock changes, has been introduced in an attempt to perform an analysis of real signals that is independent of the system of reference and of the system of units employed. The approach has proved operational, leading to an adapted time-frequency representation, the

Affine Wigner Function, that has many properties similar to those of the usual Wigner-Ville function.

The actions of time dilations on signals and time-frequency distributions are respectively characterized by two indexes r and q which are real numbers. The index of conformity$r$ depends on the physical origin of the signal and controls its behavior in a change of time units. The index of significance q, attached to the time-frequency representation, can be chosen according to the kind of description we require.

Both indexes r and q are special features coming from the introduction of dilations and allowing to express properly their effects. Those indexes are particularly important in the time-frequency analysis of problems involving transductions. They can be overlooked only in the limit of narrow relative bandwidth, in which case the usual Wigner-Ville representation appears as a universal approximation.

There are several ways to build a time-frequency analysis satisfying the above constraints of independence relatively to clock changes. They result in an affine analog of Cohen's class. However, what is called the Affine Wigner Function in the present work stands out as the unique time-frequency distribution obtained by a tomographic method fitted to the affine group. This distribution is unitary, gives a realistic description of the hyperbolic chirps and does not degrade the information contained in the original signal, except for a constant phase factor.

References

[1] J. Bertrand and P. Bertrand, "Représentations temps-fréquence des signaux," *Comptes rendus de l'Académie des Sciences, Paris*, vol. 299, Ser. 1, pp. 635–638, 1984.

[2] J. Bertrand and P. Bertrand, "A tomographic approach to Wigner's function," *Foundations of Physics*, vol. 17, pp. 397–405, 1987.

[3] K. Vogel and H. Risken, "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase," *Physical Review A*, vol. 40, pp. 2847–2849, 1989.

[4] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum," *Physical Review Letters*, vol. 70, pp. 1244–1247, 1993.

[5] J. Bertrand and P. Bertrand, "Affine time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 5, pp. 118–140, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[6] J. Bertrand and P. Bertrand, "Symbolic calculus on the time-frequency half-plane," *J. of Mathematical Physics*, vol. 39, pp. 4071–4090, August 1998.

7.2 TIME-FREQUENCY REASSIGNMENT [0]

Time-frequency and time-scale representations [1] aim to extract relevant information from a signal by representing it over a two-dimensional plane. These tools have been extensively studied in the past twenty years, resulting today in many useful analysis methods. Among them, the spectrogram and the smoothed versions of the Wigner-Ville distribution [1] are probably the most widely used, but their applicability is limited by localization trade-offs, which may be troublesome in some applications. For the spectrogram, a shorter analysis window yields a better time resolution and henceforth a poorer frequency resolution, as a consequence of the Gabor-Heisenberg inequality [1]. For the smoothed versions of the Wigner-Ville distribution, a larger smoothing kernel yields reduced cross-terms, but also a poorer localization of the signal components (see "Article 4.2").

These shortcomings must be overcome in order to obtain time-frequency "pictures" that can be both easily read by non-experts and easily included in a signal processing application. This is exactly what the *reassignment principle* has been devised for. Initially introduced in 1976 by Kodera, Gendrin and de Villedary [2], this idea first remained little known and rarely used. But recently, advances obtained during the eighties in time-frequency analysis have made its rebirth possible [3], which considerably extended its applicability, both conceptually and computationally.

7.2.1 Basic Principle

For a sake of simplicity, we will first present the basics of reassignment in the case of the spectrogram, which was the only case considered by Kodera *et al* [2]. Its application to other representations will be discussed afterwards. The spectrogram, which is the squared modulus of the short–time Fourier transform

$$S_x^h(t,f) = \left|F_x^h(t,f)\right|^2, \tag{7.2.1}$$

$$F_x^h(t,f) = \int x(u)\, h^*(t-u)\, e^{-i2\pi fu}\, du \tag{7.2.2}$$

can also be expressed as a two-dimensional smoothing of the Wigner–Ville distribution [1]

$$S_x^h(t,f) = \iint W_x(u,\nu)\, W_h(t-u, f-\nu)\, du\, d\nu. \tag{7.2.3}$$

In these expressions, t and f are respectively the time and frequency running variables, $x(t)$ is the analyzed signal, and $h(t)$ is the analyzing window. All integrals

[0]Authors: **François Auger**, GE44-IUT de Saint Nazaire, Bd de l'Université, BP 406, 44602 Saint Nazaire Cedex, France (auger@iutsn.univ-nantes.fr, auger@ge44.univ-nantes.fr); **P. Flandrin**, École Normale Supérieure de Lyon, Labo. de Physique, UMR 5672 CNRS, 46 allée d'Italie, 69364 Lyon Cedex 07, France (flandrin@ens-lyon.fr); **E. Chassande-Mottin**, Max Planck Institut für Gravitationphysik, Albert Einstein Institut, Am Mühlenberg, 1, D-14424 Golm, Germany (new email: ecm@obs-nice.fr). Reviewers: S. Barbarossa, R. Baraniuk, F. Hlawatsch and A. Papandreou-Suppappola.

have integration bounds running from $-\infty$ to $+\infty$. The latter expression shows explicitly that the value of the spectrogram at a given point (t, f) is a weighted sum of all the Wigner–Ville distribution values at the neighboring points $(t-u, f-\nu)$. The number $S_x^h(t, f)$ is therefore the sum of a whole energy distribution located around its geometrical center (t, f). Reasoning with a mechanical analogy, the situation is as if the total mass of an object was assigned to its geometrical center, an arbitrary point which except in the very specific case of an homogeneous distribution, has no reason to suit the actual distribution. A much more meaningful choice is to assign the total mass of an object — as well as the spectrogram value $S_x^h(t, f)$ — to the *center of gravity* of their respective distribution. This is exactly what the reassignment performs : at each time–frequency point (t, f) where a spectrogram value is computed, we also compute the coordinates $(\widehat{t}, \widehat{f})$ of the local centroid of the Wigner–Ville distribution W_x, as seen through the time–frequency window W_h centered at (t, f):

$$\widehat{t}_x(t,f) = \frac{1}{S_x^h(t,f)} \iint u\, W_x(u,\nu)\, W_h(t-u, f-\nu)\, du\, d\nu \tag{7.2.4}$$

$$\widehat{f}_x(t,f) = \frac{1}{S_x^h(t,f)} \iint \nu\, W_x(u,\nu)\, W_h(t-u, f-\nu)\, du\, d\nu. \tag{7.2.5}$$

Then, the spectrogram value $S_x^h(t, f)$ is moved from (t, f) to $(\widehat{t}, \widehat{f})$. This leads us to define the reassigned spectrogram as

$$\check{S}_x^h(t,f) = \iint S_x^h(u,\nu)\, \delta\left(t - \widehat{t}_x(u,\nu)\right)\, \delta\left(f - \widehat{f}_x(u,\nu)\right)\, du\, d\nu. \tag{7.2.6}$$

Originally, the reassignment operators $\widehat{t}$ and $\widehat{f}$ have been equivalently related to the phase of the STFT, an information which is generally discarded when computing the spectrogram:

$$\widehat{t}_x(t,f) = -\frac{1}{2\pi}\frac{\partial\varphi}{\partial f}(t,f) \tag{7.2.7}$$

$$\widehat{f}_x(t,f) = f + \frac{1}{2\pi}\frac{\partial\varphi}{\partial t}(t,f). \tag{7.2.8}$$

with $\varphi(t, f) = \arg F_x^h(t, f)$. These expressions may be interpreted respectively as the local group delay and the local instantaneous frequency of the signal observed inside the time-frequency domain imposed by the analysis window h. But it has been shown in [3] that a much more efficient implementation is possible thanks to a third expression involving two additional STFTs with particular analysis windows :

$$\widehat{t}_x(t,f) = t - \Re\left\{\frac{F_x^{th}(t,f)}{F_x^h(t,f)}\right\}, \tag{7.2.9}$$

$$\widehat{f}_x(t,f) = f + \Im\left\{\frac{F_x^{dh/dt}(t,f)}{2\pi F_x^h(t,f)}\right\} \tag{7.2.10}$$

As presented here, the reassignment principle can be used with a large number of distributions, beyond the spectrogram case. For example, if the WVD of the short-time window $h(t)$ in eq. (7.2.3) is replaced by an arbitrary (low-pass) kernel $\Pi(u,\nu)$, one recognizes the general form of the quadratic time–frequency energy distributions that are covariant under time and frequency shifts, referred to as the Cohen's class[1] [1]:

$$\rho_x^{\Pi}(t,f) = \iint W_x(u,\nu)\,\Pi(t-u,f-\nu)\,du\,d\nu. \tag{7.2.11}$$

The local centroids are then given by

$$\widehat{t}_x(t,f) = \frac{1}{\rho_x^{\Pi}(t,f)} \iint u\,W_x(u,\nu)\,\Pi(t-u,f-\nu)\,du\,d\nu \tag{7.2.12}$$

$$\widehat{f}_x(t,f) = \frac{1}{\rho_x^{\Pi}(t,f)} \iint \nu\,W_x(u,\nu)\,\Pi(t-u,f-\nu)\,du\,d\nu. \tag{7.2.13}$$

and the corresponding reassigned distribution becomes

$$\check{\rho}_x^{\Pi}(t,f) = \iint \rho_x^{\Pi}(u,\nu)\,\delta\left(t-\widehat{t}_x(u,\nu)\right)\,\delta\left(f-\widehat{f}_x(u,\nu)\right)\,du\,d\nu. \tag{7.2.14}$$

From a theoretical point of view, this reassigned representation is no longer bilinear, but it still remains an energy distribution covariant under time and frequency shifts. One of the most important properties of the reassignment principle is that the application of the reassignment process defined by eqs. (7.2.12), (7.2.13) and (7.2.14) to any distribution of the Cohen's class yields perfectly localized distributions for chirp signals, frequency tones and impulses, since the WVD does so, and since the centroid of a linear distribution necessary lies on the line. When applied to multicomponent signals, reassignment improves readability by overcoming — to a certain extent — the usual trade-off between cross-term level and localization: the underlying *smoothing* of the standard distribution guarantees some cross-term reduction, whereas reassignment acts as a *squeezing* that re-focuses the signal terms that had been spread out by smoothing (see Figure 7.2.1).

Among the examples of Cohen's class members studied in [3], the case of the smoothed pseudo Wigner–Ville distribution yields a very versatile signal analysis tool, with independently adjustable time and frequency smoothings:

$$SPWV_x^{g,h}(t,f) = \iint g(t-u)H(f-\nu)\,W_x(u,\nu)\,du\,d\nu \tag{7.2.15}$$

[1] That is, the quadratic class as defined on p. 68.

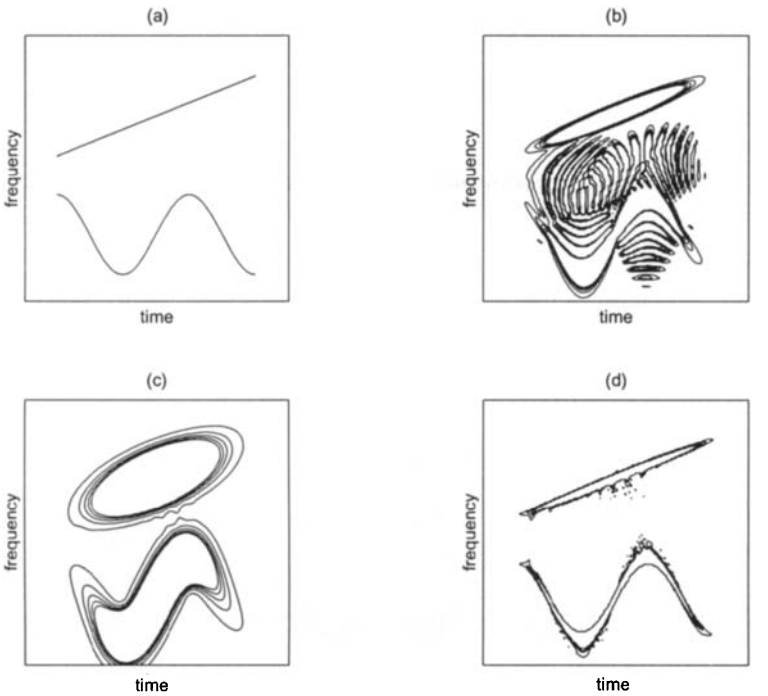


Fig. 7.2.1: Cross-term level and localization trade-off in the Cohen's class. We consider here the time-frequency representation of a signal composed of two different chirps whose instantaneous frequencies are detailed in (a). The kernel of the time-frequency distribution in the Cohen's class defines the degree of smoothing which is applied [1]: a weak smoothing favors a sharp localization at the expense of the presence of cross-terms (an example is the Wigner-Ville distribution displayed in a contour plot in (b)). Conversely, a stronger one leads to a lower cross-term level but also to a poorer localization (such as the spectrogram in (c)). Thanks to its smoothing–squeezing scheme, the reassignment method overcomes this trade-off as shown in (d) with the reassigned spectrogram.

Its reassigned version can be computed easily with two additional SPWDs:

$$\widehat{t}_x(t,f) \;=\; t - \frac{SPWV_x^{tg,h}(t,f)}{SPWV_x^{g,h}(t,f)}, \tag{7.2.16}$$

$$\widehat{f}_x(t,f) \;=\; f + i\,\frac{SPWV_x^{g,dh/dt}(t,f)}{2\pi\, SPWV_x^{g,h}(t,f)} \tag{7.2.17}$$

A different kind of generalization can be obtained when switching to time-scale energy distributions of the affine class [1], i.e., the quadratic distributions covariant under time shifts and dilations:

$$\Omega_x^{\Pi}(t,a) = \iint W_x(u,\nu)\,\Pi\left(\frac{t-u}{a}, a\nu\right) du\, d\nu \tag{7.2.18}$$

Within this framework, the reassignment operator in time is given directly by

$$\widehat{t}_x(t,a) = \frac{1}{\Omega_x^{\Pi}(t,a)} \iint u\, W_x(u,\nu)\, \Pi\left(\frac{t-u}{a}, a\nu\right) du\, d\nu, \tag{7.2.19}$$

whereas the reassignment operator in scale requires an intermediate step in the frequency domain

$$\widehat{a}_x(t,a) = \frac{f_0}{\widehat{f}_x(t,a)}, \quad \text{with} \quad f_0 = \iint f\, \Pi(t,f)\, dt\, df \tag{7.2.20}$$

$$\text{and} \quad \widehat{f}_x(t,a) = \frac{1}{\Omega_x^{\Pi}(t,a)} \iint \nu\, W_x(u,\nu)\, \Pi\left(\frac{t-u}{a}, a\nu\right) du\, d\nu$$

The most important case among this class is the scalogram (the squared modulus of the wavelet transform) [1], obtained by choosing for Π the WVD of the chosen wavelet. Simple and efficient expressions of the reassignment operators also exist in this case [3,4].

7.2.2 Variations and Related Approaches

7.2.2.1 Two Variations

The original purpose of the reassignment principle was the design of time-frequency distributions with increased readability. But some useful information on the signal structure can also be directly extracted from the reassignment operators, as shown by the following two extensions:

Signal/noise discrimination and supervised reassignment. When the analyzed signal includes broadband noise, the reassignment process yields peaked areas in noise-only regions, whereas rather smooth energy distributions are expected there. For such situations, an improved reassignment algorithm referred to as *supervised reassignment* [5] has been designed. This approach first attempts to discriminate between "signal+noise" and "noise only" regions in the time-frequency plane by means of a detector applied to the reassignment operators. Reassignment is then only performed for the points considered to belong to "signal+noise" regions.

Signal component extraction and differential reassignment. Many signal processing problems such as denoising and signal classification can be solved by a relevant tiling of the time-frequency plane, so as to isolate each signal "component" (although this concept is not clearly defined). For such applications, a new reassignment process called *differential reassignment* [6] has been considered. Whereas the original reassignment principle moves each value by one finite jump, differential reassignment considers each time-frequency point as the starting point of an elementary particle whose velocity field is deduced from the reassignment operators. The final points called *asymptotic reassignment points* are gathered and lead to a time-frequency map in which each cell indicates a signal component.

7.2.2.2 Related Approaches

Although original in many respects, the concept of reassignment is obviously connected with several approaches that have been proposed independently. We lack space to discuss these interactions precisely, but we cite:

- The instantaneous frequency density [7], which yields at each time sample a histogram of the frequency reassignment operator of the spectrogram.
- The extraction of ridges and skeletons out of the phase structure of the wavelet transform [8,9]. These ridges are made of the fixed points of the reassignment operators, either horizontally ($\widehat{a}_x(t,a) = a$) or vertically ($\widehat{t}_x(t,a) = t$).
- The synchrosqueezed plane [10], which also moves the scalogram values, but by a scale displacement only.

7.2.3 Summary and Conclusions

Reassignment can be viewed as the second step of a process whose goal is to build a readable time–frequency representation. It consists of:

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 which survived the smoothing.

As a result, this approach yields — without a drastic increase in computational complexity — enhanced contrast (when compared to smoothed distributions such as spectrograms) with a much reduced level of interferences (when compared to the Wigner-Ville distribution). This is especially true when the signal–noise ratio is not too low, and when the signal components are not "too close" to each other. Finally, MATLAB^{TM} implementations of the algorithms discussed here are included in a freeware available at http://crttsn.univ-nantes.fr/~auger/tftb.html .

References

[1] P. Flandrin, *Time-Frequency/Time-Scale Analysis*. San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[2] K. Kodera, C. de Villedary, and R. Gendrin, "A new method for the numerical analysis of nonstationary signals," *Physics of the Earth & Planetary Interiors*, vol. 12, pp. 142–150, 1976.

[3] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[4] P. Flandrin, E. Chassande-Mottin, and P. Abry, "Reassigned scalograms and their fast algorithms," in *Proc. SPIE: Wavelet Applications in Signal and Image Processing III*, vol. 2569, pp. 152–158, Soc. of Photo-optical Instrumentation Engineers, San Diego, CA, 12–14 July 1995.

[5] E. Chassande-Mottin, F. Auger, and P. Flandrin, "Supervised time-frequency reassignment," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 517–520, Paris, 18–21 June 1996.

[6] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin, "Differential reassignment," *IEEE Signal Processing Letters*, vol. 4, pp. 293–294, October 1997.

[7] D. Friedman, "Instantaneous frequency distribution vs. time: An interpretation of the phase structure of speech," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'85)*, pp. 1121–1124, Tampa, FL, 26–29 March 1985.

[8] R. Carmona, W. L. Hwang, and B. Torrésani, *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms with an Implementation in S.* San Diego: Academic Press, 1998.

[9] P. Guillemain and R. Kronland-Martinet, "Horizontal and vertical ridges associated to continuous wavelet transforms," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 63–66, Victoria, BC, 4–6 October 1992.

[10] S. Maes, "The synchrosqueezed representation yields a new reading of the wavelet transform," in *Proc. SPIE: Wavelet Applications II*, vol. 2491, pp. 532–559, Soc. of Photo-optical Instrumentation Engineers, Orlando, FL, 17–21 April 1995.

7.3 MEASURING TIME-FREQUENCY DISTRIBUTIONS CONCENTRATION [0]

Efficient measurement of time-frequency distributions (TFDs) concentration can provide a quantitative criterion for evaluation of various distributions performance. It can be used for adaptive and automatic parameter selection in time-frequency analysis, without supervision of a user. Measures for distribution concentration of monocomponent signals date back to [1, 2]. For more complex signals, some quantities from statistics and information theory were the inspiration for defining measures of the TFDs concentration [3,4]. They provided good quantitative measure of the auto-terms concentration. Various and efficient modifications are used in order to take into account the appearance of oscillatory cross-terms.

The application of concentration measures will be demonstrated on automatic determination of the "best window length" for the spectrogram or "the best number of terms" in the method that provides transition form the spectrogram toward the pseudo Wigner distribution (pseudo WD) [Article 6.2].

7.3.1 Concentration Measurement

The basic idea for measuring TFDs concentration can be explained on a simplified example motivated by the probability theory. Consider a set of N nonnegative numbers $p_1, p_2, ..., p_N \geq 0$, such that $p_1 + p_2 + ... + p_N = 1$. Form a simple test function $M(p_1, p_2, ..., p_N) = p_1^2 + p_2^2 + ... + p_N^2$. It is easy to conclude that $M(p_1, p_2, ..., p_N)$, under the constraint $p_1 + p_2 + ... + p_N = 1$, has the minimal value for $p_1 = p_2 = ... = p_N = 1/N$, i.e., for maximally spread values of $p_1, p_2, ..., p_N$. The highest value of $M(p_1, p_2, ..., p_N)$, under the same constraint, is achieved when only one p_i is different from zero, $p_i = \delta(i - i_0)$, where i_0 is an arbitrary integer $1 \leq i_0 \leq N$. This case corresponds to the maximally concentrated values of $p_1, p_2, ..., p_N$, at a single $p_{i_0} = 1$. Therefore, the function $M(p_1, p_2, ..., p_N)$ can be used as a measure of concentration of the set of numbers $p_1, p_2, ..., p_N$, under the unity sum constraint.[1] In general, the constraint can be included in the function itself by using the form $M(p_1, p_2, ..., p_N) = \left(p_1^2 + p_2^2 + ... + p_N^2\right) / \left(p_1 + p_2 + ... + p_N\right)^2$. For nonnegative $p_1, p_2, ..., p_N$ this function has the minimum for $p_1 = p_2 = ... = p_N$, and reaches its maximal value when only one p_i is different from zero.

In time-frequency analysis this idea has been used in order to measure the concentration. Several forms of the concentration measure, based on this fundamental idea, are introduced.

1. **Measure based on the ratio of norms:** For the WD of energy normalized signals, the relation $\sum_n \sum_k \rho_x^2(n,k) \equiv 1$ holds. Therefore, substitution

[0] Author: **LJubiša Stanković**, Elektrotehnicki fakultet, University of Montenegro, 81000 Podgorica, Montenegro (l.stankovic@ieee.org). Reviewers: W. J. Williams and P. Flandrin.

[1] In probability theory, the famous Shannon entropy $-\sum_i p_i \log(p_i)$ is commonly used for the same purpose. It produces the maximal value for the lowest concentration of probabilities p_i, $p_1 = p_2 = ... = p_N = 1/N$, and the minimal value for the highest concentration $p_i = \delta(i - i_0)$.

$p_i \to \rho_x^2(n,k)$ in the basic example, gives a function that can be used for measuring the concentration of the time-frequency representation $\rho_x(n,k)$:

$$M_{JP} = \left(\frac{L_4}{L_2}\right)^4 = \frac{\sum_n \sum_k \rho_x^4(n,k)}{(\sum_n \sum_k \rho_x^2(n,k))^2}. \tag{7.3.1}$$

This form is just the fourth power of the ratio of L_4 and L_2 norms of $\rho_x(n,k)$.[2] It has been introduced by Jones and Parks in [3]. They have used the magnitude of the signal's short-time Fourier transform as the time-frequency representation $\rho_x(n,k)$ in (7.3.1). High values of M_{JP} indicate that the representation $\rho_x(n,k)$ is highly concentrated, and vice versa. In general, any other ratio of norms L_p and L_q, $p > q > 1$, can also be used for measuring the concentration of $\rho_x(n,k)$ [3].

When there are two or more components (or regions in time-frequency plane of a single component) of approximately equal energies (importance), whose concentrations are very different, the norm based measures will favor the distribution with a "peaky" component, due to raising of distribution values to a high power. It means that if one component (region) is "extremely highly" concentrated, and all the others are "very poorly" concentrated, then the measure will not look for a trade-off, when all components are "well" concentrated. In order to deal with this kind of problems, common in time-frequency analysis, a concentration measure could be applied to smaller, local time-frequency regions [3]:

$$M_{JPL}(n,k) = \frac{\sum_n \sum_k Q^2(m-n,l-k)\rho_x^4(m,l)}{(\sum_n \sum_k Q(m-n,l-k)\rho_x^2(m,l))^2} \tag{7.3.2}$$

The localization weighting function $Q(n,k)$ determines the region where the concentration is measured. In [3] the Gaussian form of this function is used.

2. **Rényi entropy based measures:** The second class of TFD measures is defined in analogy with the Rényi entropy. It has been introduced in time-frequency analysis by Williams et al. [4,5], with a significant contribution of [6,7] in establishing the properties of this measure. The Rényi entropy, applied on the TFD $\rho_x(n,k)$, has the form

$$R_\alpha = \frac{1}{1-\alpha}\log_2\left(\sum_n \sum_k \rho_x^\alpha(n,k)\right) \tag{7.3.3}$$

with $\alpha > 2$ being recommended for the TFD measures [7]. For $\alpha = 2$ and the WD of energy normalized signals ($\sum_n \sum_k \rho_x^2(n,k) \equiv 1$), we have $R_2 = 0$ for all signals. Note that the logarithm is a monotone function. Thus, the behavior of R_α is determined by the argument $\sum_n \sum_k \rho_x^\alpha(n,k)$ behavior, as explained at the beginning of this section. In contrast to the measure (7.3.1), the entropy (7.3.3) has larger values for less concentrated distributions due to a negative coefficient

[2] In statistics, similar form (known as kurtosis) is used as a measure of the flatness or peakedness of a distribution. Kurtosis is zero for a Gaussian distribution. Values greater than zero mean that the distribution has more of a peak than a Gaussian distribution, while values less than zero mean flatter distributions.

$1/(1-\alpha)$ for $\alpha > 2$. This will be the case for all other measures which will be presented in the sequel.

It is interesting to note that the **Shannon entropy**

$$H = -\sum_n \sum_k [\rho_x(n,k) \log_2 \rho_x(n,k)]$$

could be recovered from the Rényi entropy, from the limit case $\alpha \to 1$, [7]. The Shannon entropy could not be used for general TFDs $\rho_x(n,k)$, which can assume negative values [7].

3. **Normalized Rényi entropy measures.** In order to avoid the problem which could be caused by the fact that the Rényi entropy based measure with $\alpha = 3$ ignore the presence of oscillatory cross-terms (when the auto-terms are well separated), some kind of normalization should be done. It can be done in various ways, leading to a variety of possible measure definitions [4].

Normalization with the distribution volume is performed as:

$$RV_3 = -\frac{1}{2} \log_2 \sum_n \sum_k \left[\rho_x(n,k) / \sum_n \sum_k |\rho_x(n,k)|\right]^3. \tag{7.3.4}$$

If the distribution contains oscillatory values, then summing their absolute values means that large cross-terms will decrease the measure RV_α. This is the expected behavior of a measure, since it will seek for a balance between the cross-terms suppression and auto-terms enhancement. The volume normalized form of measure has been used for adaptive kernel design in [4].

4. The basic idea for the measure that will be presented next comes from an obvious **classical definition of the time-limited signal duration**. If a signal $x(n)$ is time-limited to the interval $n \in [n_1, n_2 - 1]$, i.e., $x(n) \neq 0$ only for $n \in [n_1, n_2 - 1]$, then the duration of $x(n)$ is $d = n_2 - n_1$. It can be written as $d = \lim_{p\to\infty} \sum_n |x(n)|^{1/p}$. The same definition applied to a two-dimensional function $\rho_x(n,k) \neq 0$ only for $(n,k) \in D_x$, gives

$$N_D = \lim_{p\to\infty} \sum_n \sum_k |\rho_x(n,k)|^{1/p} \tag{7.3.5}$$

where N_D is the number of points within D_x. In reality, there is no a sharp edge between $\rho_x(n,k) \neq 0$ and $\rho_x(n,k) = 0$, so the value of (7.3.5) could, for very large p, be sensitive to small values of $\rho_x(n,k)$. The robustness may be achieved by using lower order forms, for example with $p = 2$. Therefore, the concentration can be measured with the function of the form

$$M_p^p = (\sum_n \sum_k |\rho_x(n,k)|^{1/p})^p, \tag{7.3.6}$$

with $\sum_n \sum_k \rho_x(n,k) = 1$, and $p > 1$.

After we have presented several possible forms for measuring the concentration of TFDs, we can summarize **a procedure for constructing a TFD measure**

based on one-dimensional classical signal analysis definitions, or definitions from either probability, quantum mechanics, or information theory:

(i) In the classical signal analysis definitions, consider the signal power $|x(t)|^2$ (spectral energy density $|X(f)|^2$) as the probability density function in time (frequency). This idea comes from quantum mechanics, where the absolute square of the wave function is the position's probability density function.

(ii) Assume that the TFD $\rho_x(t,f)$ can be treated as a joint two-dimensional probability density function.

(iii) According to these assumptions, reintroduce one-dimensional definition into joint two-dimensional time-frequency domain.

(iv) Additional modifications, interpretations, constraints, and normalizations are needed in order to get forms that can be used in time-frequency analysis. For example, several possible forms of the Rényi entropy measure in time-frequency domain have been proposed and used in various problems.

Example: Consider the classic **Leipnik entropy measure** [2], and **Zakai's entropy parameter** $\delta_t = -\int_{-\infty}^{\infty} |x(t)|^2 \ln |x(t)|^2 \, dt$ of signal $x(t)$ [2]. According to the procedure for constructing a time-frequency form, based on a classical signal processing relation, we get

$$\delta_t = -\int_{-\infty}^{\infty} |x(t)|^2 \ln |x(t)|^2 \, dt \quad \rightarrow \quad -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho_x(t,f) \ln \rho_x(t,f) dt df \tag{7.3.7}$$

This is exactly the well known **Shannon entropy**. It has already been discussed in [7] with respect to its (non)applicability in time-frequency problems. In a similar way, a logarithm of the general Zakai's signal duration (uncertainty)

$$Z_\alpha = \log_2 T_{2\alpha} = \frac{1}{1-\alpha} \log_2 \frac{\int_{-\infty}^{\infty} \left(|x(t)|^2\right)^\alpha dt}{\left(\int_{-\infty}^{\infty} |x(t)|^2 \, dt\right)^\alpha},$$

according to the proposed procedure, transforms into the Rényi entropy measure,

$$Z_\alpha \rightarrow \frac{1}{1-\alpha} \log_2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho_x^\alpha(t,f) dt df = R_\alpha$$

where $|x(t)|^2$ has been replaced by $\rho_x(t,f)$, and the unit signal energy is assumed.

Remark: In the probability theory all results are derived for the probability values p_i, assuming that $\sum_i p_i = 1$ and $p_i \geq 0$. The same assumptions are made in classical signal analysis for the signal power. Since a general TFD commonly does not satisfy both $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho_x(t,f) = 1$ and $\rho_x(t,f) \geq 0$, the obtained measures of TFD concentration may just *formally look like the original* entropies or classical signal analysis forms, while they can have different behavior and properties.[3]

[3]Quantum mechanics forms can also be used for the definition of highly concentrated signal representations. One of them is the "pseudo quantum" signal representation [8] in the form of $SD_x(t,\wp) = \int_{-\infty}^{\infty} x^{[L]}(t+\tau/(2L)) x^{*[L]}(t-\tau/(2L)) e^{-j\wp\tau} d\tau$, with $x^{[L]}(t) = A(t)\exp(jL\phi(t))$ for $x(t) = A(t)\exp(j\phi(t))$. For example, for $x(t) = A\exp(-at^2/2 + jbt^2/2 + jct)$ we get $SD_x(t,\wp) = A^2\exp(-at^2)\sqrt{4\pi/(a/L^2)}\exp(-(\wp - bt - c)^2/(a/L^2))$. For $a/L^2 \rightarrow 0$ it results in $SD(t,\wp) = 2\pi A^2 \exp(-at^2)\delta(\wp - bt - c)$, what is just an ideally concentrated distribution along the instantaneous frequency. For a large a, if L^2 is large enough so that $a/L^2 \rightarrow 0$, we get the distribution highly concentrated in a very small region around the point $(t,\wp) = (0,c)$.

7.3.2 Numerical Examples

Consider the spectrogram

$$S_x^w(n,k) = |F_x^w(n,k)|^2 / E$$

where $F_x^w(n,k) = DFT_{m\to k}\{w(m)x(n+m)\}$ is the short-time Fourier transform (STFT); E is the energy of the lag window $w(m)$. Among several spectrograms, calculated with different window lengths or forms, the best one according to the one of proposed concentration measures, denoted by $\mathcal{M}[\rho_x(n,k)]$, will be that which minimizes (or maximizes, depending on the used measure form):

$$w^+ = \arg\min_w \{\mathcal{M}[S_x^w(n,k)]\}. \tag{7.3.8}$$

Let us illustrate this by an example. Consider the signal

$$x(t) = \cos(50\cos(\pi t) + 10\pi t^2 + 70\pi t) + \cos(25\pi t^2 + 180\pi t) \tag{7.3.9}$$

sampled at $\Delta t = 1/256$, within $-1 \le t < 1$. The Hanning window $w(m)$ with different lengths is used in the spectrogram calculation. Here, we have used the measure (7.3.6) with $p = 2$, although for this signal all presented measures would produce similar results [9]. Note that the presented measures would significantly differ if, for example, the second component were pure sinusoid $\cos(180\pi t)$ instead of $\cos(25\pi t^2 + 180\pi t)$.

For wide lag windows, signal nonstationarity makes the spectrogram very spread in the time-frequency plane, having relatively large measure $\mathcal{M}[S_x^w(n,k)] = M_2^2$, Figs. 7.3.1(a), 7.3.1(b). For narrow lag windows its Fourier transform is very wide, causing spread distributions and large M_2^2, Figs. 7.3.1(d), 7.3.1(e). Obviously, between these two extreme situations there is a window that produces an acceptable trade-off between the signal nonstationarity and small window length effects. The measure M_2^2 is calculated for a set of spectrograms with $N = 32$ up to $N = 256$ window length, Fig. 7.3.1(f). The minimal measure value, meaning the best concentrated spectrogram according to this measure, is achieved for $N = 88$, Fig. 7.3.1(f). The spectrogram with $N = 88$ is shown in Fig. 7.3.1(c).

The same procedure will be used for determination of the optimal number of terms L, in a transition from the spectrogram to the pseudo WD, according to the recursive form of the S-method (SM) [Article 6.2]:

$$\mathrm{SM}_x(n,k;L) = \mathrm{SM}_x(n,k;L-1) + 2\Re\{F_x(n,k+L)F_x^*(n,k-L)\} \tag{7.3.10}$$

where $\mathrm{SM}_x(n,k;0) = |F_x^w(n,k)|^2$, and both $k+L$ and $k-L$ are within the basic frequency period. Note that $\mathrm{SM}(n,k;N/2)$ is equal to the pseudo WD. The optimal distribution $\mathrm{SM}_x^+(n,k;L)$, on our way from $L = 0$ (the spectrogram) toward $L = N/2$ (the pseudo WD), is the one calculated with L producing the minimal value of $\mathcal{M}[\mathrm{SM}_x(n,k;L)]$,

$$L^+ = \arg\min_L \{\mathcal{M}[\mathrm{SM}_x(n,k;L)]\}. \tag{7.3.11}$$

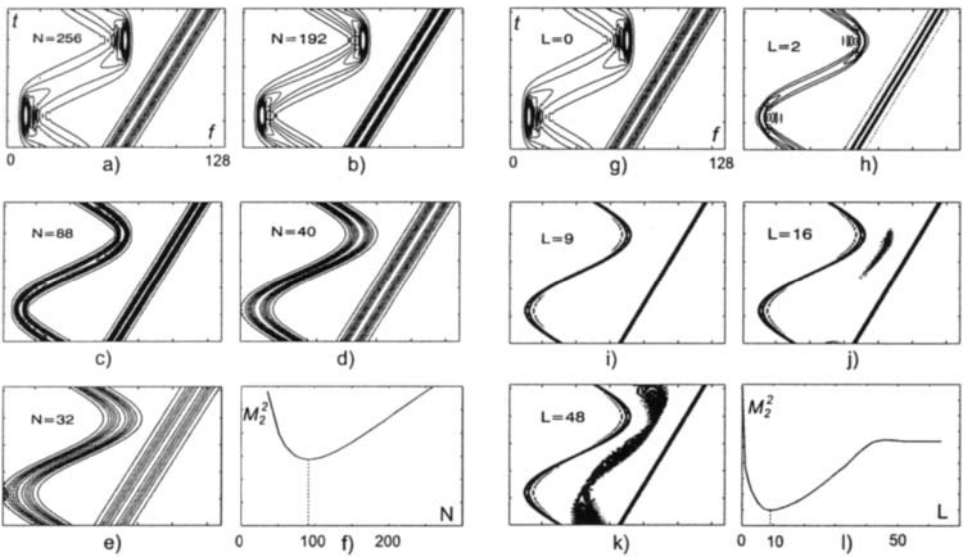


Fig. 7.3.1: (a)-(e) Spectrogram for various window lengths, and (f) its measure $\mathcal{M}\left[S_x^w(n,k)\right] = M_2^2$. The lowest M_2^2 is achieved for $N = 88$, being the best window length choice according to this measure. (g)-(k) The S-method for various values of parameter L, and (l) its measure $\mathcal{M}\left[\mathrm{SM}_x(n,k;L)\right] = M_2^2$. The lowest M_2^2 is obtained for $L = 9$.

Here, instead of $|\mathrm{SM}_x(n,k;L)|$, a nonnegative part of $\mathrm{SM}_x(n,k;L)$ will be used. Distributions $\mathrm{SM}_x(n,k;L)$ should be properly scaled in order to satisfy unbiased energy condition. The same signal is used for the illustration of the SM. Since this method is based on the WD, the best results will be achieved with a wide lag window in the STFT calculation, $N = 256$. The spectrogram ($L = 0$) is shown in Fig. 7.3.1(g). By increasing L the SM improves concentration of the spectrogram toward the pseudo WD quality, meaning lower measure $\{\mathcal{M}\left[\mathrm{SM}_x(n,k;L)\right]\} = M_2^2$, Fig. 7.3.1(h), 7.3.1(i). After L has reached the value equal to the distance between the auto-terms, cross-terms start to appear, increasing M_2^2, Figs. 7.3.1(j), 7.3.1(k). Minimal M_2^2 means a trade-off between the auto-terms concentration and the cross-terms appearances, Fig. 7.3.1(k). The SM with L corresponding to minimal M_2^2 is shown in Fig. 7.3.1(l).

The concentration measure is illustrated on time-frequency analysis of a pressure signal in the BMW engine with speed 2000 [rev/min], Fig. 7.3.2, [Article 15.2].

7.3.3 Parameter Optimization

Parameter optimization may be done by a straightforward computation of a distribution measure $\mathcal{M}[\rho_x(n,k)]$, for various parameter values. The best choice according to this criterion (optimal distribution with respect to this measure) is the

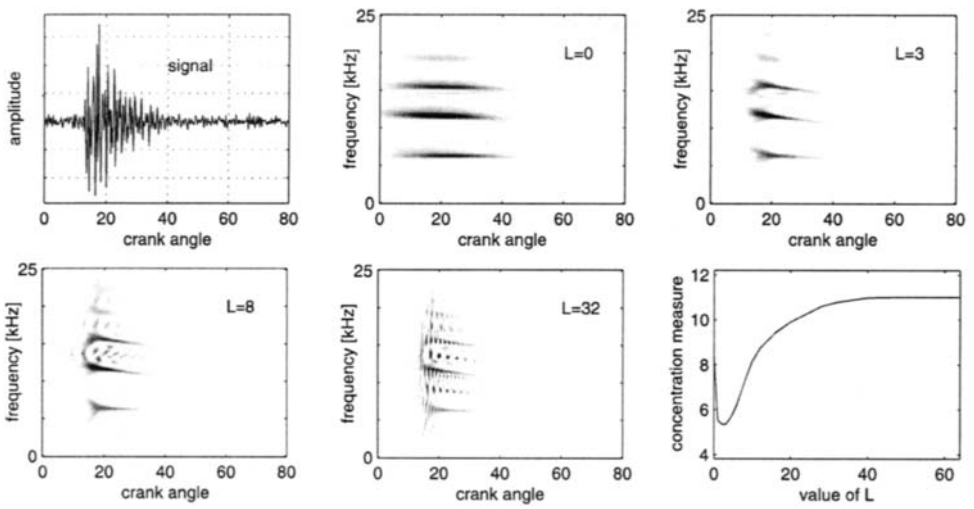


Fig. 7.3.2: Concentration measure illustration on time-frequency analysis of a car engine pressure signal. Signal, and its S-method based time-frequency representations are given. Time is rescaled into corresponding crank-angle. The best choice according to this measure was $L = 3$.

distribution which produces the minimal value of $\mathcal{M}[\rho_x(n,k)]$. In the cases when one has to consider a wide region of possible parameter values for the distribution calculation (like for example window lengths in spectrogram), this approach can be numerically inefficient. Then, some more sophisticated optimization procedures, like the one using the steepest descent approach described in [4], can be used. Its simplified version will be presented here [9].

The gradient of a measure $\mathcal{M}[\rho_x(n,k)]$, with respect to a distribution's generalized optimization parameter denoted by ξ, is

$$\frac{\partial \mathcal{M}[\rho_x(n,k)]}{\partial \xi} = \frac{\partial \mathcal{M}[\rho_x(n,k)]}{\partial \rho_x(n,k)} \frac{\partial \rho_x(n,k)}{\partial \xi}.$$

Iterations, starting from a very low concentrated distribution toward the maximally concentrated one, i.e., toward the measure minimum, can be done according to

$$\xi_{m+1} = \xi_m - \mu \partial \mathcal{M}[\rho_x(n,k)]/\partial \xi \tag{7.3.12}$$

where μ is the step, which should be chosen in the same way as the step in the other adaptive algorithms. The step should not be too small (since the convergence would be too slow), and not too large (to miss the minimum, or cause the divergence).

In discrete implementations, the gradient $\partial \mathcal{M}[\rho_x(n,k)]/\partial \xi$ can be approximated based on $\mathcal{M}[\rho_x(n,k;\xi_m)]$ calculated with ξ_m and its previous value ξ_{m-1}

$$\xi_{m+1} = \xi_m - \mu \frac{\mathcal{M}[\rho_x(n,k;\xi_m)] - \mathcal{M}[\rho_x(n,k;\xi_{m-1})]}{\xi_m - \xi_{m-1}}. \tag{7.3.13}$$

Example: The optimization procedure will be illustrated on the signal $x(t)$, its spectrogram, and the measure form Example 2. The optimal window length is obtained in few iterations by using (7.3.13), starting from the very narrow window. Values of $\xi_0 = N = 16$ and $\xi_1 = N = 20$ in the initial and first iteration, are assumed. The next value of $\xi_{m+1} \equiv N$ is calculated according to (7.3.13). During the iterations we get $\xi_m = 16, 20, 76$, and 90. The algorithm is stopped at $\xi_m = 90$, when $|\xi_{m+1} - \xi_m| < 2$, since even number of samples are used in the realization. Note that the obtained optimal value is within ± 2 of the value obtained by direct calculation. The value of parameter $\mu = 1/3$ has been used in all examples.

7.3.4 Summary and Conclusions

Measurement of time-frequency distributions concentration, with application to an automatic optimization of distribution parameters, is presented. It is based on the forms borrowed from the classical signal analysis, probability, or information theory, with appropriate interpretations and adjustments.

References

[1] D. Gabor, "Theory of communication," *J. IEE*, vol. 93(III), pp. 429–457, November 1946.

[2] D. E. Vakman, *Sophisticated Signals and the Uncertainty Principle in Radar.* New York: Springer, 1968. Translated by K. N. Trirogoff; edited by E. Jacobs.

[3] D. L. Jones and T. W. Parks, "A high-resolution data-adaptive time-frequency representation," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 2127–2135, December 1990.

[4] T.-H. Sang and W. J. Williams, "Rényi information and signal-dependent optimal kernel design," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'95)*, vol. 2, pp. 997–1000, Detroit, 9–12 May 1995.

[5] W. J. Williams, M. L. Brown, and A. O. Hero III, "Uncertainty, information and time-frequency distributions," in *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations II*, vol. 1566, pp. 144–156, Soc. of Photo-optical Instrumentation Engineers, San Diego, 24–26 July 1991.

[6] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Trans. Information Theory*, vol. 47, pp. 1391–1409, May 2001.

[7] P. Flandrin, R. G. Baraniuk, and O. Michel, "Time-frequency complexity and information," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, vol. 3, pp. 329–332, Adelaide, Australia, 19–22 April 1994.

[8] L. Stanković, "Highly concentrated time-frequency distributions: Pseudo-quantum signal representation," *IEEE Trans. Signal Processing*, vol. 45, pp. 543–551, March 1997.

[9] L. Stanković, "A measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, pp. 621–631, March 2001.

7.4 RESOLUTION PERFORMANCE ASSESSMENT FOR QUADRATIC TFDs[0]

7.4.1 Selecting and Comparing TFDs

Quadratic time-frequency distributions (TFDs) are effective tools for extracting information from a non-stationary signal, such as the number of components, their durations and bandwidths, components' relative amplitudes and instantaneous frequency (IF) laws (see Chapters 1 and 2). The performance of TFDs depends on the type of signal (see Chapter 3) [1,2]. For example, in the case of a monocomponent linear FM signal, the Wigner-Ville distribution is known to be optimal in the sense that it achieves the best energy concentration around the signal IF law (see Article 2.1 for more details) [1].

In applications involving multicomponent signals, choosing the right TFD to analyze the signals is an immediate critical task for the signal analyst. How best to make this assessment, using current knowledge, is the subject of this article.

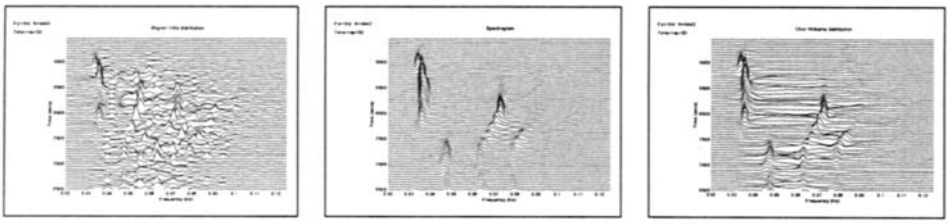
Let us, for example, consider a multicomponent whale signal, represented in the time-frequency domain using the Wigner-Ville distribution, the spectrogram, the Choi-Williams distribution, the Born-Jordan distribution, the Zhao-Atlas-Marks (ZAM) distribution, and the recently introduced B-distribution [3] (see Fig. 7.4.1).

To determine which of the TFDs in Fig. 7.4.1 "best" represents this whale signal (i.e. which one gives the best components' energy concentration and best interference terms suppression, and allows the best estimation of the components' IF laws) one could *visually* compare the six plots and choose the most appealing. The spectrogram and the B-distribution, being almost free from the cross-terms, seem to perform best.

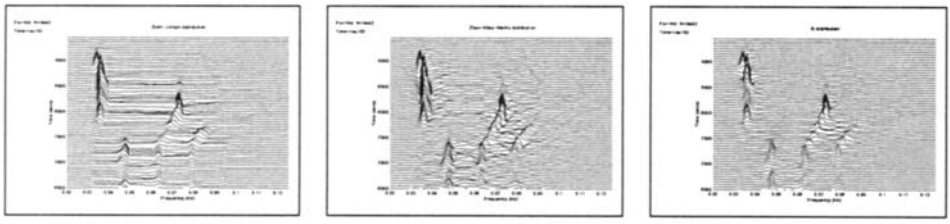
The performance comparison based on the visual inspection of the plots becomes more difficult and unreliable, however, when the signal components are closely-spaced in the time-frequency plane. To objectively compare the plots in Fig. 7.4.1 requires to use a quantitative performance measure for TFDs. There have been several attempts to define objective measures of "complexity" for TFDs (see Section 7.3.1). One of these measures, the Rényi entropy given in [4], has been used by several authors in preference to e.g. the bandwidth–duration product given in [1]. The performance measure described in this article, unlike the Rényi entropy, is a local measure of the TFD resolution performance, and is thus more suited to the selection problem illustrated by Fig. 7.4.1. This measure takes into account the characteristics of TFDs that influence their resolution, such as energy concentration, components separation, and interference terms minimization. Methodologies for choosing a TFD which best suits a given signal can then be developed by optimizing the resolution performance of considered TFDs and modifying their parameters to better match application-specific requirements.

---

[0]Authors: **Boualem Boashash** and **Victor Sucic**, Signal Processing Research Centre, Queensland University of Technology, GPO Box 2434, Brisbane, Q 4001, Australia (b.boashash@qut.edu.au, v.sucic@qut.edu.au). Reviewers: W. J. Williams and LJ. Stanković.



(a) Wigner–Ville (b) Spectrogram (Hanning, $L = 511$) (c) Choi-Williams ($\sigma = 10$)



(d) Born–Jordan (e) ZAM ($a = 2$) (f) B ($\beta = 0.01$)

Fig. 7.4.1: TFDs of a multicomponent whale signal.

7.4.2 Performance Criteria for TFDs

In the case of *monocomponent* FM signals, the best TFD is that which maximizes energy concentration about the signal instantaneous frequency. This is achieved by minimizing component sidelobe amplitude A_s relative to mainlobe amplitude A_m, and mainlobe bandwidth B relative to central frequency f (see Fig. 7.4.2).

The instantaneous concentration performance of a TFD may thus be quantified by the measure p expressed as:

$$p(t) = \left| \frac{A_s(t)}{A_m(t)} \right| \frac{B(t)}{f(t)} \tag{7.4.1}$$

A good performance is characterized by a small value of the measure p. For example, for the Wigner-Ville distribution of a linear FM signal with infinite duration, the bandwidth B and the sidelobe amplitude A_s are zero [1], and we obtain $p = 0$.

For *multicomponent* FM signals, the performance of a TFD can be *quantitatively* assessed in terms of:

- the energy concentration of the TFD about the respective instantaneous frequency of each component, as expressed by Eq. (7.4.1), and
- the components resolution, as measured by the frequency separation of the components' mainlobes, including the effect of cross-terms.

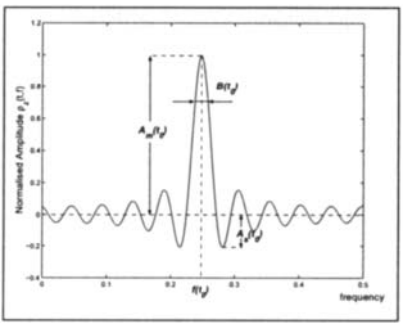


Fig. 7.4.2: Slice of a TFD of a monocomponent signal at time $t = t_0$. The dominant peak is the component, while the other peaks are the sidelobes. For clarity of presentation, we limit ourselves to measuring the mainlobe bandwidth at 0.71 of the component normalized amplitude A_m.

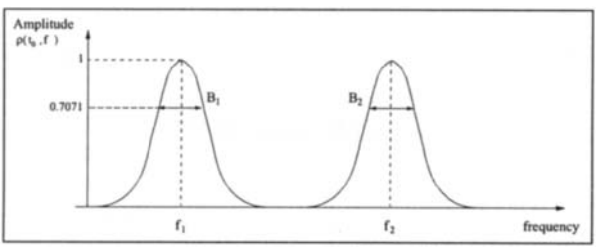


Fig. 7.4.3: Diagram illustrating the resolution of a two-component signal in the absence of cross-terms. The lobes are clearly distinguished from each other; the components are said to be resolved.

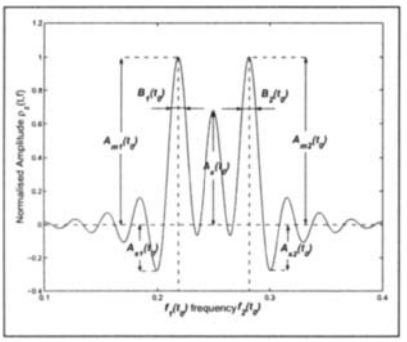
For stationary signals, the frequency resolution in a power spectral density estimate of a signal composed of two single tones, f_1 and f_2 (see Fig. 7.4.3), may be defined as the minimum difference $f_2 - f_1$ for which the following inequality holds:

$$f_1 + B_1/2 < f_2 - B_2/2\,, \qquad f_1 < f_2 \tag{7.4.2}$$

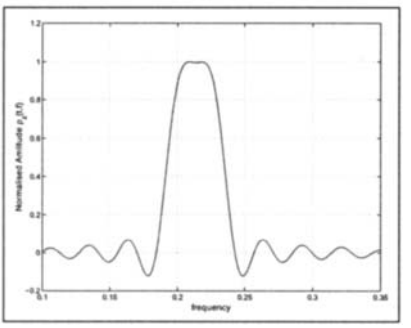
where B_1 and B_2 are the respective bandwidths of the first and the second sinusoid.

In the case of non-stationary signals, for a TFD $\rho_z(t, f)$ of a two-component signal, the above definition of resolution is valid for every time slice of a cross-term-free TFD, such as the spectrogram. However, for TFDs exhibiting cross-terms, we need to take into account the effect of cross-terms on resolution.

A slice of a typical quadratic TFD, with components clearly resolved, is shown in Fig. 7.4.4(a), where $B_1(t_0)$, $f_1(t_0)$, $A_{s_1}(t_0)$ and $A_{m_1}(t_0)$ represent respectively the instantaneous bandwidth, the IF, the sidelobe amplitude and the mainlobe amplitude of the first component at time $t = t_0$. Similarly, $B_2(t_0)$, $f_2(t_0)$, $A_{s_2}(t_0)$ and $A_{m_2}(t_0)$ represent the instantaneous bandwidth, the IF, the sidelobe amplitude and the mainlobe amplitude of the second component at the same time t_0. The amplitude $A_x(t_0)$ is that of the cross-term. An example of a quadratic TFD with non-resolved components is shown in Fig. 7.4.4(b).



(a) The two dominant peaks are the (resolved) signal components, the middle peak is the cross-term, and the other peaks are the sidelobes



(b) The two components and the cross-term have merged into a single peak; we say that the components are *not* resolved

Fig. 7.4.4: Slice of a TFD of a two-component signal taken at time $t = t_0$.

7.4.3 Resolution Performance Measure for TFDs

Eq. (7.4.2) and Fig. 7.4.4(a) suggest that the resolution performance of a TFD for a neighboring pair of components in a multicomponent signal may be defined by the minimum difference $D(t) = f_2(t) - f_1(t)$ for which we still have a positive separation measure $S(t)$ between the components' mainlobes centered about their respective IFs, $f_1(t)$ and $f_2(t)$. For best resolution performance of TFDs, $S(t)$ should be as close as possible to the true difference between the actual IFs.

The components' separation measure $S(t)$ is expressed as [5]:

$$S(t) = \left(f_2(t) - \frac{B_2(t)}{2}\right) - \left(f_1(t) + \frac{B_1(t)}{2}\right) \tag{7.4.3}$$

The resolution also depends on the following set of variables, all of which should be as small as possible:

(a) the normalized *instantaneous bandwidth* of the signal component $B_k(t)/f_k(t)$, $k = 1, 2$, which is accounted for in $S(t)$ (Eq. (7.4.3)),

(b) the ratio of the *sidelobe amplitude* $|A_{s_k}(t)|$ to the *mainlobe amplitude* $|A_{m_k}(t)|$, $k = 1, 2$, of the components, and

(c) the ratio of the *cross-term amplitude* $|A_x(t)|$ to the *mainlobe amplitudes* of the signal components $|A_{m_k}(t)|$, $k = 1, 2$.

It follows that the best TFD for multicomponent signals analysis is the one that concurrently *minimizes* the positive quantities (a), (b), (c), and *maximizes* $S(t)$.

Hence, by combining the above variables, expressions for a measure $P(t)$ of the resolution performance of a given TFD can be defined. Two have been proposed

in [5], among these a normalized performance measure expressed as:

$$P(t) = 1 - \frac{1}{3}\left\{\left|\frac{A_s(t)}{A_m(t)}\right| + \frac{1}{2}\left|\frac{A_x(t)}{A_m(t)}\right| + \left(1 - \frac{S(t)}{D(t)}\right)\right\} \tag{7.4.4}$$

where, for a pair of signal components, $A_m(t)$ and $A_s(t)$ are respectively the average amplitudes of the components' mainlobes and sidelobes, $A_x(t)$ is the cross-term amplitude, $S(t)$, defined by Eq. (7.4.3), is a measure of the components' separation in frequency, and $D(t) = f_2(t) - f_1(t)$ is the difference between the components' actual IFs. The algorithm presented in [6] describes how the parameters in Eq. (7.4.4) are measured in practice.

The measure $P(t)$ is close to 1 for well-performing TFDs and 0 for poorly-performing ones. Therefore, when designing a TFD we want to maximize $P(t)$ in order to reduce the cross-terms, while preserving the components' resolution.

In some applications involving real-life signals, we may need to better discriminate between different TFDs resolution performances in a given set of K TFDs. In this case, a suitable alternative to $P(t)$ that was proposed in [7] could be used. It is expressed as:

$$M_j(t) = 1 - \frac{1}{3}\left\{\frac{\left|\frac{A_{s_j}(t)}{A_{m_j}(t)}\right|}{\max\limits_{1\le k\le K}\left(\left|\frac{A_{s_k}(t)}{A_{m_k}(t)}\right|\right)} + \frac{\left|\frac{A_{x_j}(t)}{A_{m_j}(t)}\right|}{\max\limits_{1\le k\le K}\left(\left|\frac{A_{x_k}(t)}{A_{m_k}(t)}\right|\right)} + \frac{\frac{B_j(t)}{D_j(t)}}{\max\limits_{1\le k\le K}\left(\frac{B_k(t)}{D_k(t)}\right)}\right\} \tag{7.4.5}$$

where $M_j(t)$ $(1 \le j \le K)$ is the resolution performance measure of the j–th TFD, and B is the average instantaneous bandwidth of the components mainlobes. The measure $M(t)$ is used in Section 7.4.5 to compare the performances of quadratic TFDs of a real-life signal, as it discriminates better than the measure $P(t)$ for real-life signals [8].

7.4.4 Application to the Selection of the Optimal TFD for a Given Multicomponent Signal

A methodology for selecting the optimal TFD for resolving closely-spaced components in a multicomponent signal involves then the following steps:

1. Define a set of comparison criteria describing the information sought from TFDs (Section 7.4.2).

2. Objectively measure the resolution performance of TFDs based on these criteria (use the measure P defined by Eq. (7.4.4)).

3. Optimize each TFD to match the criteria as close as possible [5,6]: Select as the optimal TFD kernel parameter value the one which maximizes the overall performance measure P_{overall}, taken to be the mean of the instantaneous measures P in a time interval of interest.

Table 7.4.1: Optimization results for the TFDs of signal $s(t)$ defined by Eq. (7.4.6).

TFD	*Optimal value of the kernel parameter*	P_{overall}
Spectrogram	Bartlett window, length 31	0.86
Wigner-Ville	N/A in this case	0.62
Choi-Williams	$\sigma = 1$	0.82
Born-Jordan	N/A in this case	0.81
Zhao-Atlas-Marks (ZAM)	$a = 2$	0.67
Modified B	$\beta = 0.04$	0.88

4. Quantitatively compare TFDs and select the best one: An optimized TFD which has the largest value of P_{overall} is selected as *best* for representing the given signal in the joint time-frequency domain.

Example: We define the following two-component signal in noise:

$$\begin{aligned} s(t) &= s_1(t) + s_2(t) + n(t) \\ &= \cos\left(2\pi\left(0.1\,t + \tfrac{\alpha}{2}\,t^2\right)\right) + \cos\left(2\pi\left(0.2\,t + \tfrac{\alpha}{2}\,t^2\right)\right) + n(t) \end{aligned} \qquad (7.4.6)$$

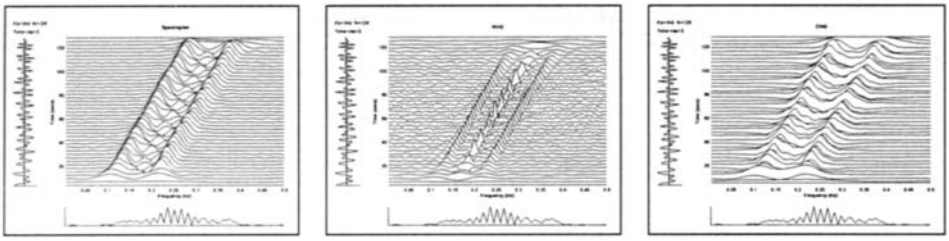
where $\alpha = 0.0016$ is the component bandwidth-duration ratio (duration $T = 128$), and $n(t)$ is additive white Gaussian noise with signal-to-noise ratio SNR = 10 dB. The sampling frequency is $f_s = 1$ Hz.

The signal $s(t)$ is analyzed in the time-frequency domain using the following TFDs: the spectrogram, the Wigner-Ville distribution, the Choi-Williams distribution, the Born-Jordan distribution, the Zhao-Atlas-Marks (ZAM) distribution, and the Modified B-distribution [9].

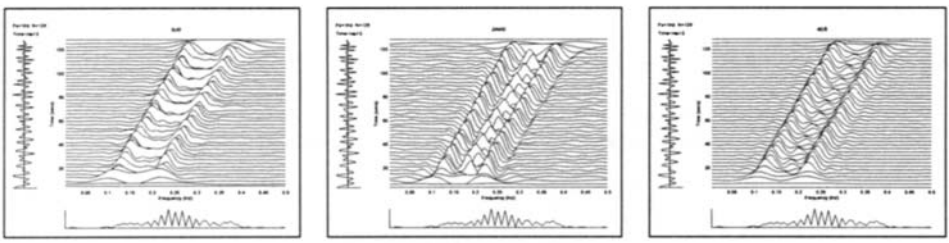
To find the optimal TFD for resolving the two components of $s(t)$, we first find the optimal values of the TFDs kernel parameters, as described in the above methodology. The Wigner-Ville distribution and the Born-Jordan distribution have no "smoothing" parameters, hence do not need optimizing. The optimized TFD with the largest P_{overall} among the considered TFDs is then selected as optimal for representing $s(t)$. Table 7.4.1 lists the results of the optimization process, and it shows that the signal optimal TFD is the Modified B-distribution with the parameter $\beta = 0.04$. All optimized TFDs are plotted in Fig. 7.4.5. From the signal optimal TFD important signal parameters can be measured (see Table 7.4.2). In addition, by optimizing components' concentration and resolution, more accurate components IF laws' estimates are obtained from the peaks of the optimal TFD's dominant ridges in the time-frequency plane [1] (see Fig. 7.4.6).

7.4.5 Use of the Performance Measure in Real-Life Situations

The methodology defined in this section enables to select a real-life signal best-performing TFD in an objective, automatic way. Its use should make time-frequency techniques more applicable in practice (e.g. machine condition monitoring described in Articles 15.2 and 15.6, or other applications presented in Part V).



(a) Spectrogram (Bartlett, $L = 31$) (b) Wigner–Ville (c) Choi-Williams ($\sigma = 1$)



(d) Born–Jordan (e) ZAM ($a = 2$) (f) Modified B ($\beta = 0.04$)

Fig. 7.4.5: Optimized TFDs of signal $s(t)$ defined by Eq. (7.4.6).

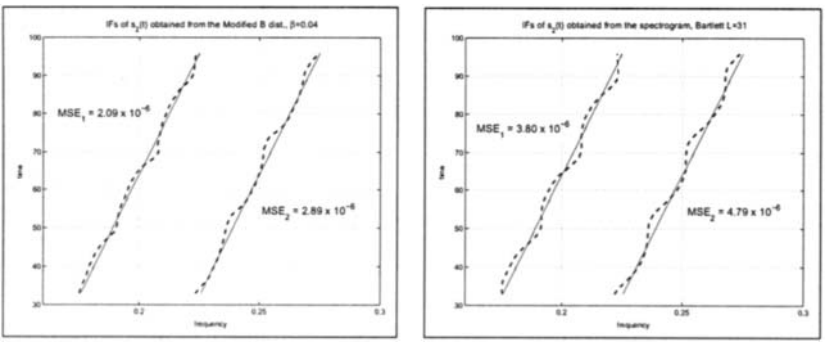
Table 7.4.2: Parameters of $s_1(t)$ and $s_2(t)$ (Eq. (7.4.6)) measured from the signal $s(t)$ optimal TFD (Modified B-distribution, $\beta = 0.04$). The values shown are the averages over $t \in [32, 96]$.

Parameter	*Component* $s_1(t)$	*Component* $s_2(t)$
Instantaneous bandwidth B	0.0194	0.0195
Mainlobe amplitude $\lvert A_m \rvert$	1.0002	0.9574
Sidelobe amplitude $\lvert A_s \rvert$	0.0900	0.0858
Cross-term amplitude $\lvert A_x \rvert$	0.1503	

The methodology consists of the following steps:

1. Represent the signal in the time-frequency domain with a quadratic TFD, i.e. a smoothed Wigner-Ville distribution (see Article 3.2). Following the approach described in Article 5.7, we smooth the WVD in both time t and lag τ with the Hanning window of length equal to a quarter of the signal duration. This time-frequency smoothing is intended to suppress the WVD inner and outer artifacts, while preserving components time-frequency features (see Article 4.2 for more details).

2. For the different time instants of the smoothed WVD, select the two closest dominant peaks in the frequency direction. To achieve the best resolution of



(a) Modified B ($\beta = 0.04$) (b) Spectrogram (Bartlett, $L = 31$)

Fig. 7.4.6: Comparison of the measured (dashed) and true (solid) IF laws of the component $s_1(t)$ (left) and $s_2(t)$ (right) of the signal $s(t)$ defined by Eq. (7.4.6). The mean-square-errors (MSEs) of the IF estimates obtained from the peaks of the signal optimized Modified B-distribution (best-performing TFD) are given in (a), and those obtained from the peaks of the signal optimized spectrogram (second best TFD) in (b).

the signal components, the best resolution of the two closest components at an observed time instant is sufficient [10]. Note that if a signal is monocomponent or no components exist at a particular time, this time instant is not considered.

3. For the selected pairs of components, optimize different TFDs using the resolution performance measure M defined by Eq. (7.4.5). The measure M is used over P since it is a better discriminator of real-life signals TFDs resolution performances [11]. The kernel parameter value, which from a set of different values considered, maximizes the overall performance measure (the mean of M over the observed times) is selected as the kernel parameter optimal value.

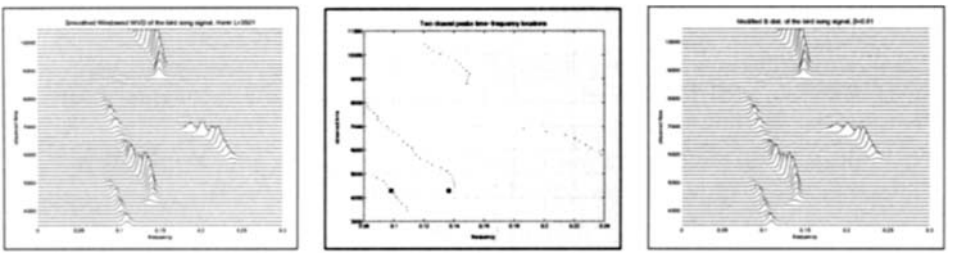
4. Calculate the measure M of the optimized TFDs for each of the selected pairs of signal components. The TFD which maximizes the average (over time) M is selected as the signal *best-performing* TFD among the considered TFDs.

Example: To illustrate how to use this methodology in practice, let us find the best-performing TFD for the Noisy Miner (*Manorina melanocephala*) song signal. The same TFDs we considered in the synthetic signal optimal TFD selection example will be considered in this real-life signal example.

We start by representing the signal in the time-frequency domain with the Wigner-Ville distribution smoothed in time and lag with the Hanning windows of length $L = 3501$ (Fig. 7.4.7(a)).

For each time instant of the smoothed WVD we then identify the pair of closest components. From Fig. 7.4.7(b) we can see that different components form such pairs at different times.

Next, as described in steps 3 and 4 of the above-defined methodology, the six considered TFDs are first optimized, after which their resolution performances are



(a) Smoothed WVD (Hanning, $L = 3501$)

(b) Pairs of closest components

(c) Modified B ($\beta = 0.91$)

Fig. 7.4.7: Optimization of the bird song signal TFDs. The signal Smoothed WVD is shown in (a), and the pairs of its closest components in (b), with the overall closest pair (at time $t = 4295$) marked by the squares. The signal optimized Modified B-distribution is shown in (c).

Table 7.4.3: Optimization and comparison results for the TFDs of the Noisy Miner song signal. The values of M_{overall} indicate that the spectrogram performs better than most traditional TFDs in this case. Only the Modified B-distribution performs better than the spectrogram and all others.

TFD	*Optimal value of the kernel parameter*	M_{overall}
Spectrogram	Bartlett window, length 3501	0.90
Wigner-Ville	N/A in this case	0.50
Choi-Williams	$\sigma = 0.004$	0.74
Born-Jordan	N/A in this case	0.65
Zhao-Atlas-Marks (ZAM)	$a = 2$	0.63
Modified B	$\beta = 0.91$	0.93

evaluated using the measure M. Table 7.4.3 shows the signal TFDs kernel parameters optimization and the TFDs resolution performance results. The Modified B-distribution for $\beta = 0.91$, plotted in Fig. 7.4.7(c), is found to have the largest value of M_{overall} (the mean of M over the time instants). Therefore, we select this TFD as best to represent the Noisy Miner song signal in the time-frequency plane.

7.4.6 Summary and Conclusions

This article defines a measure for assessing the resolution performance of quadratic TFDs in separating closely-spaced components in the time-frequency domain. The measure takes into account key attributes of TFDs, such as components' mainlobes and sidelobes, and cross-terms. The introduction of this measure allows to quantify the quality of TFDs instead of relying solely on visual inspection of plots. The resolution performance measure also allows for selecting the optimal TFD in a given practical application, and improving methodologies for designing high resolution quadratic TFDs, such as the Modified B-distribution.

References

[1] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[2] L. Stanković, "An analysis of some time-frequency and time-scale distributions," *Ann. Télécommunications*, vol. 49, pp. 505–517, September/October 1994.

[3] B. Barkat and B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis," *IEEE Trans. Signal Processing*, vol. 49, pp. 2232–2239, October 2001.

[4] T.-H. Sang and W. J. Williams, "Rényi information and signal-dependent optimal kernel design," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'95)*, vol. 2, pp. 997–1000, Detroit, 9–12 May 1995.

[5] B. Boashash and V. Sucic, "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 51, pp. 1253–1263, May 2003.

[6] V. Sucic and B. Boashash, "Parameter selection for optimising time-frequency distributions and measurements of time-frequency characteristics of non-stationary signals," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, vol. 6, pp. 3557–3560, Salt Lake City, UT, 7–11 May 2001.

[7] V. Sucic, B. Boashash, and K. Abed-Meraim, "A normalised performance measure for quadratic time-frequency distributions," in *Proc. Second IEEE Internat. Symp. on Signal Processing and Information Technology (ISSPIT'02)*, pp. 463–466, Marrakech, Morocco, 18–21 December 2002.

[8] V. Sucic and B. Boashash, "Selecting the optimal time-frequency distribution for real-life multicomponent signals under given constraints," in *Proc. Eleventh European Signal Processing Conf. (EUSIPCO-02)*, vol. 1, pp. 141–144, Toulouse, 3–6 September 2002.

[9] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 50, pp. 1866–1876, August 2002.

[10] V. Sucic and B. Boashash, "Optimisation algorithm for selecting quadratic time-frequency distributions: Performance results and calibration," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, vol. 1, pp. 331–334, Kuala Lumpur, 13–16 August 2001.

[11] V. Sucic and B. Boashash, "An approach for selecting a real-life signal best-performing time-frequency distribution," in *Proc. Seventh Internat. Symp. on Signal Processing and its Applications (ISSPA'03)*, vol. 1, pp. 100–104, Paris, 1–4 July 2003.

7.5 JOINT-DOMAIN REPRESENTATIONS VIA DISCRETE-DOMAIN FRAMES [0]

Representation of a signal in joint domains has been a very active area in signal processing. Prime examples of such representation are the short-time Fourier transform (STFT) and the discrete-time Gabor transform (DTGT) in the joint time-frequency domain, and filter banks and wavelets in the joint time-scale domain. The varieties of joint-domain transforms are unified in this article in the form of product function and cascaded frames.

7.5.1 Frames and Reconstruction Collections

A **frame** is a collection of sequences, which we refer to as frame elements, that generates a finite representation of a signal via the inner product of the signal with the frame elements. An overview of frame theory can be found in [1,2]. Mathematically, a collection of sequences $\{v_m\}$ in a Hilbert space $\mathbb{H}$, $m \in \mathbb{Z}$, the set of integers, forms a frame for $\mathbb{H}$, if there exist two numbers A and B, called the lower and upper frame bound, respectively, such that $0 < A \leq B < \infty$ and

$$A\|f\|^2 \leq \sum_m |\langle f, v_m \rangle|^2 \leq B\|f\|^2, \qquad \forall f \in \mathbb{H}. \tag{7.5.1}$$

Every $f \in \mathbb{H}$ can be reconstructed using a corresponding dual frame $\{u_m\}$ as $f = \sum_m \langle f, v_m \rangle u_m = \sum_m \langle f, u_m \rangle v_m$. The corresponding dual frame $\{u_m\}$ in $\mathbb{H}$ is defined by $u_m = S^{-1}v_m$, where S is the frame operator defined via $Sf = \sum_m \langle f, v_m \rangle v_m$.

A collection $\{u_m\}$ in $\mathbb{H}$ is defined to be a **reconstruction collection** [3] (RC) for $\{v_m\}$ if every $x \in \mathbb{H}$ can be written as $x = \sum_m \langle x, v_m \rangle u_m$. It is assumed that all the collections or sequences described in this article belong to the same Hilbert space of interest. An RC $\{u_m\}$ is related to a frame $\{v_m\}$ via the following theorem [3].

Theorem 7.5.1: *A collection $\{u_m\}$ is a reconstruction collection for a frame $\{v_m\}$ for $\mathbb{H}$ if and only if*

$$\sum_m v_m^*(l)u_m(k) = \delta(k-l). \tag{7.5.2}$$

The RC concept is similar to the pseudo-frame concept [4], where two Bessel sequences $\{u_m\}$ and $\{v_m\}$ are a pseudo frame and pseudodual frame pair for $\mathbb{H}$ if and only if $\langle f, g \rangle = \sum_m \langle f, v_m \rangle \langle u_m, g \rangle$, $\forall f, g \in \mathbb{H}$. The pseudo-frame theory, however, does not restrict the analysis collection to be a frame. By restricting the analysis collection to be a frame, the boundedness and continuity of the transform is ensured.

---

[0]Authors: **Joel M. Morris**, Computer Science and Electrical Engineering Department, University of Maryland Baltimore County, 1000 Hilltop Circle, Catonsville, MD 21250, USA (morris@umbc.edu), and **Sanjay M. Joshi**, Lucent Technologies, 1701 Harbor Bay Parkway, Alameda, CA 94502, USA (joshi@ieee.org). Reviewers: X.-G. Xia, R. Baraniuk, and M. J. Bastiaans.

The spaces of interest are the spaces of periodic and non-periodic square-summable sequences. The space $l^2(\mathbb{Z}/L)$ is the space of all sequences $f(k)$ periodic in k with integer period L. The inner product for $l^2(\mathbb{Z}/L)$ is defined as $\langle x, y\rangle = \sum_{k=0}^{L-1} x(k)y^*(k)$. The designation $\underline{L}$ represents the set $\{0, 1, 2, \ldots, L-1\}$. The space $l^2(\mathbb{Z})$ is the space of all square-summable sequences, with the inner product defined as $\langle x, y\rangle = \sum_{k=-\infty}^{\infty} x(k)y^*(k)$.

7.5.2 Product-Function Frames

Product-function frames (PFFs) generalize the windowed transforms, such as the discrete-time Gabor transforms, as seen in the following subsections.

7.5.2.1 PFF for Periodic Spaces

A PFF with two defining factor functions g_m and ϕ_n in $l^2(\mathbb{Z}/L)$ is defined via $\{g_m\phi_n\}$. The elements of this frame are $(g_m\phi_n)(k) = g_m(k)\phi_n(k)$, where $m \in \underline{M}$ and $n \in \underline{N}$. Since the span of $\{g_m\phi_n\}$ is $l^2(\mathbb{Z}/L)$, it is necessary that MN should be greater than or equal to its dimension L, *i.e.*, $MN \geq L$.

We choose the collection $\{\phi_n\}$ as an orthonormal basis for $l^2(\mathbb{Z}/N)$ and take $L = MN' = M'N$, where M' and N' are integers. This implies $l^2(\mathbb{Z}/N) \subset l^2(\mathbb{Z}/L)$. The following theorem [5] characterizes the frame condition for $\{g_m\phi_n\}$ in terms of matrices $\mathbf{G}(i)$, $i \in \underline{N}$, whose elements are given by

$$G_{qp}(i) = \sum_{m=0}^{M-1} g_m^*(i + qN)g_m(i + pN), \qquad p, q \in \underline{M'}.$$

Theorem 7.5.2: *A collection $\{g_m\phi_n\}$ is a frame for $l^2(\mathbb{Z}/L)$ iff $\mathbf{G}(i)$ are non-singular $\forall i \in \underline{N}$.*

Let an RC of a frame $\{g_m\phi_n\}$ be denoted by $\{\gamma_m\phi_n\}$, defined similar to $\{g_m\phi_n\}$. The collections $\{g_m\phi_n\}$ and $\{\gamma_m\phi_n\}$ are related via the following theorem [5].

Theorem 7.5.3: *A frame $\{\gamma_m\phi_n\}$ is a reconstruction collection for $\{g_m\phi_n\}$ iff $g_m(k)$ and $\gamma_m(k)$ satisfy*

$$\sum_{m=0}^{M-1} g_m(k + qN)\gamma_m^*(k + pN) = \delta(q - p), \qquad k \in \underline{N} \text{ and } q, p \in \underline{M'}. \tag{7.5.3}$$

Let $g_m(k) = g(k - mN')$ be a time-shifted version of a sequence $g(k)$, where N' is the time-shift parameter. Equation (7.5.3) then becomes

$$\sum_{m=0}^{M-1} g^*(k + qN - mN')\gamma(k + pN - mN') = \delta(q - p), \quad k \in \underline{N},\ q, p \in \underline{M'}. \tag{7.5.4}$$

This condition is the same as the one for the **Gabor transform** in [6] and in several other papers in different forms. Since ϕ_n can be chosen as any orthonormal

transform, the same pair of g and γ can be used for any windowed orthonormal transform. **Windowed transforms** are equivalent to oversampled (or critically-sampled) **modulated filterbanks** [7]. The relationship between the analysis and synthesis filters is in the time domain as given by Eq. (7.5.4).

Instead of using a single window for the Gabor transform, we can use a number of windows, resulting in a **multi-window Gabor transform** [8]. The equations are modified such that g_m is replaced by $g_{p,m}$, where p denotes the window number and m denotes the shift.

We now define a new family of transforms for periodic discrete-time sequences called **product transforms** [5]. Let $g_m(k) = g(k)\psi_m(k)$ and $\gamma_m(k) = \gamma(k)\psi_m(k)$, where $\{\psi_m\}$ is an orthonormal basis for $l^2(\mathbb{Z}/M)$. From Theorem 7.5.3, g and γ must satisfy

$$g^*(k+qN)\gamma(k+pN)\sum_{m=0}^{M-1}\psi_m^*(k+qN)\psi_m(k+pN) = \delta(q-p), \tag{7.5.5}$$

where $k \in \underline{N}$ and $p, q \in \underline{M'}$, to guarantee $\{g_m\phi_n\}$ and $\{\gamma_m\phi_n\}$ are reconstruction collections for each other. The dimension of the space spanned by $\{g\psi_m\phi_n\}$ is equal to the least common multiple LCM (M, N) [5].

This scheme has been utilized for fast computation of the DFT, *i.e.*, **FFT**, by using N and M-point DFTs as $\{\phi_n\}$ and $\{\psi_m\}$, where M and N are relatively-prime integers. The same algorithm can be used for any orthonormal transform of periodic sequences. The frame-theoretic structure derived above allows selecting different transforms for $\{\phi_n\}$ and $\{\psi_m\}$ to obtain a variety of transforms.

7.5.2.2 PFF for Non-Periodic Spaces

For $l^2(\mathbb{Z})$, the elements of a PFF $\{g_m\phi_n\}$ are $(g_m\phi_n)(k) = g_m(k)\phi_n(k)$, like the periodic case. The indices m and n are integers and $m \in \mathbb{Z}$ and $n \in \underline{N}$.

Following Daubechies' method [1], a sufficient condition for a product-function collection in $l^2(\mathbb{Z})$ to be a frame can be stated as follows [3].

Theorem 7.5.4: *Let $\{g_m\phi_n\}$ be a collection in $l^2(\mathbb{Z})$, where $m \in \mathbb{Z}$, $n \in \underline{N}$, and $\{\phi_n\}$ is an orthonormal basis for $l^2(\mathbb{Z}/N)$. If*

$$m(g) = \inf_{l\in\mathbb{Z}}\sum_m |g_m(l)|^2 > 0, \qquad M(g) = \sup_{l\in\mathbb{Z}}\sum_m |g_m(l)|^2 < \infty, \tag{7.5.6}$$

and

$$\sup_{s\in\mathbb{Z}}\left[(1+s^2)^{(1+\epsilon)/2}\beta(s)\right] = C_\epsilon < \infty, \tag{7.5.7}$$

for some $\epsilon > 0$, where

$$\beta(s) = \sup_{l\in\mathbb{Z}}\sum_m |g_m(l)g_m(l-s)|, \tag{7.5.8}$$

then $\{g_m\phi_n\}$ is a frame.

Examples of such collections are windowed transforms and a new scale-modulation transform discussed later in this article.

A necessary condition for a frame is as follows [3].

Corollary 7.5.5: *If a collection $\{g_m\phi_n\}$ defined in Theorem 7.5.4 is a frame for $l^2(\mathbb{Z})$, then Eq. (7.5.6) is true.*

For a PFF $\{g_m\phi_n\}$, let an RC be defined as $\{\gamma_m\phi_n\}$. Note that this is not the only solution possible. Substituting in the generalized reconstruction condition (Eq. 7.5.2), we get

$$\sum_m g_m^*(k - nN)\gamma_m(k) = \delta(n). \tag{7.5.9}$$

Equation 7.5.9, thus, is a necessary and sufficient condition for a collection $\{\gamma_m\phi_n\}$ to be an RC for a frame $\{g_m\phi_n\}$.

For a windowed transform, such as the discrete-time Gabor transform, $g_m(k)$ is defined as a shifted version of a mother window $g(k)$, *i.e.*, $g_m(k) = g(k - mM)$, where $M \geq 1$ is known as the shift-parameter. Assuming that $\gamma_m(k)$ also has the same structure (it is not necessary), the reconstruction condition (Eq. 7.5.9) becomes

$$\sum_m g^*(k - mM - nN)\gamma(k - mM) = \delta(n), \tag{7.5.10}$$

which is analogous to the biorthogonal-like condition [9] for defining pairs of biorthogonal-like sequences $\{g(k)\}$ and $\{\gamma(k)\}$ used for the discrete-time Gabor expansion in $l^2(\mathbb{Z})$. Collections $\{g_{mn}\}$ and $\{\gamma_{mn}\}$, where $g_{mn}(k) = g(k - nN)e^{j2\pi mk/M}$, $g(k) \in l^2(\mathbb{Z})$, and γ_{mn} defined similarly, are a pair of collections of biorthogonal-like sequences if

$$\left\langle g, \gamma^{(uv)} \right\rangle = \frac{N}{M}\delta(u)\delta(v), \tag{7.5.11}$$

where $\gamma^{(uv)}(k) = \gamma(k - uM)e^{j2\pi vk/N}$ [9]. Equation 7.5.11 is called the **biorthogonal-like condition**.

The product-function frame structure separates the computation of the window functions from the orthonormal basis used, thus reducing the complexity. The same windows g and γ, for instance, could be used with the discrete cosine transform (DCT) to obtained a windowed DCT. Thus, we have a generalized structure for windowed transforms.

If the frame element component g_m is replaced by a component g_{pq}, where $g_{pq}(k) = g_p(k - qT)$, the necessary and sufficient conditions for the **multi-Gabor expansions in** $l^2(\mathbb{Z})$ are obtained. In this frame, g_p are different windows, T is the time-shift parameter, and qT is the shift of each window. This structure, hence, extends the multi-window Gabor schemes described in [8] for the finite-dimensional space of periodic signals, $l^2(\mathbb{Z}/L)$, to the infinite-dimensional space $l^2(\mathbb{Z})$.

A new type of transform using a **scale-modulation** structure [3] is now derived from the PFFs. Consider $\mathbb{C}^L$, the space of length-L complex sequences, where for two sequences $x(k)$ and $y(k)$, $k \in \underline{L}$, in $\mathbb{C}^L$, the inner product is defined as

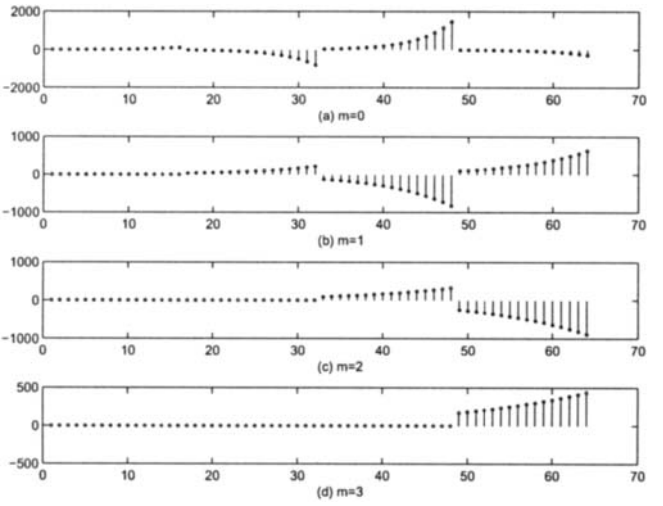


Fig. 7.5.1: Reconstruction collection windows γ_m for the scaled windows discussed in the text.

$\langle x, y \rangle = \sum_{k=0}^{L-1} x(k)y^*(k)$ and the norm is defined via $\|x\| = \sqrt{\langle x, x \rangle}$. It is a finite-dimensional subspace of $l^2(\mathbb{Z})$ with dimension L. Let $L = MN$, where M and N are integers. Define $g_m(k)$, $m = 0, 1, 2, \ldots, M-1$, to be scaled versions of a mother window, each of length $(m+1)N$ and norm 1. For example, an exponentially-decaying mother window generates scaled windows given by $g_m(k) = e^{-dk/((m+1)N)}$, where $0 \leq m \leq M-1$, $0 \leq k < (m+1)N$. Let the ϕ_n elements used with these windows be $\phi_n(k) = e^{j2\pi nk/N}$, where $0 \leq n, k \leq N-1$. Note that the ϕ_n elements are periodically extended to a length of MN. For such a frame with $d = 4$ and $N = 16$, the corresponding RC elements are shown in Fig. 7.5.1.

This new transform was used in [3] to analyze an exponentially decaying modulated signal, as found in magnetic resonance techniques. It was shown that the scale-modulation transform is likely to yield better signal estimation in the presence of noise.

The shift-modulation structure of the Gabor transforms has been in use for several years. The shift-scale structure of the wavelet transforms is a more recent development. The obvious third choice is the scale-modulation structure, which is provided by the scale-modulation transform.

7.5.3 Cascaded Frames

Although the product-function structure unifies the windowed transforms, it is not a good structure to represent the other popular joint time-frequency analysis scheme: wavelets. The discrete wavelet transform (DWT) and all of its variations can be unified under a second structure, cascaded frames. In simpler terms, a cascaded

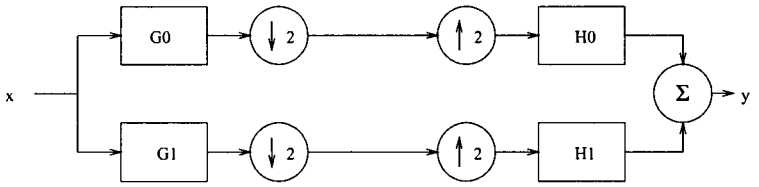


Fig. 7.5.2: Perfect reconstruction filterbank.

frame is a 'transform of transforms', although the discussion is presented in terms of the more general frames concept.

Suppose a frame $\{u_{0,m}\}$ maps a signal f in $l^2(\mathbb{Z}/L_0)$ to a vector of coefficients $\mathbf{c}_1 = [c_{1,0}, c_{1,1}, \ldots, c_{1,L_1}]$, $L_1 \geq L_0$. The coefficients themselves can be treated as a signal in $l^2(\mathbb{Z}/L_1)$. We can then use a frame $\{u_{1,n}\}$ for $l^2(\mathbb{Z}/L_1)$ to map the vector $\mathbf{c}_1$ to another vector $\mathbf{c}_2$ in some $l^2(\mathbb{Z}/L_2)$, $L_2 \geq L_1$. A signal in $l^2(\mathbb{Z}/L_0)$, thus, is effectively mapped to a coefficient vector $\mathbf{c}_2$ in $l^2(\mathbb{Z}/L_2)$. This can be viewed as a frame $\{u_p\}$, $p \in \underline{L_2}$, operating on a signal f in $l^2(\mathbb{Z}/L_0)$. We call this structure a cascaded frame. The following theorem proves that the cascaded structure corresponds to a frame [3].

Theorem 7.5.6: *Suppose* $\{u_{0,m}\}$, $m \in \underline{L_1}$, *is a frame for* $l^2(\mathbb{Z}/L_0)$, *and* $\{u_{1,n}\}$, $n \in \underline{L_2}$, *is a frame for* $l^2(\mathbb{Z}/L_1)$, *where* $L_0 \leq L_1 \leq L_2$. *Then there exists a frame* $\{u_p\}$, $p \in \underline{L_2}$, *for* $l^2(\mathbb{Z}/L_0)$ *obtained by cascading* $\{u_{0,m}\}$ *and* $\{u_{1,n}\}$.

When the number of elements in a frame for a finite-dimensional space is finite, frame theory is characterized in terms of matrix theory. Computation of an RC becomes an inversion problem of square matrices in the case of critically-sampled frames, and that of rectangular matrices in the case of oversampled frames.

The cascaded structure can be extended to $l^2(\mathbb{Z})$ via the following theorem [3].

Theorem 7.5.7: *If* $\{u_m\}$, $m \in \mathbb{Z}$, *and* $\{v_n\}$, $n \in \mathbb{Z}$, *are frames for* $l^2(\mathbb{Z})$, *then cascading* $\{u_m\}$ *and* $\{v_n\}$ *results in a frame for* $l^2(\mathbb{Z})$.

We now discuss some examples of cascaded frames in $l^2(\mathbb{Z})$. Consider a perfect reconstruction (PR) filterbank (FB) shown in Fig. 7.5.2. One of the analysis filters, say $G0$, is typically a low-pass filter and the other is a high-pass filter. The output of each filter, thus, gives the information about the signal in a particular time period in the corresponding frequency band. In **discrete wavelet transforms** (DWT), the output of $G0$ is further processed using the same filterbank, thus increasing the frequency resolution at the cost of reducing the time resolution for the transform coefficients. This cascading process is continued repeatedly to obtain the complete DWT. PR is guaranteed as long as the corresponding synthesis FB is present on the synthesis side.

In **wavelet packet transforms** (WPT), the output of any filter may be subjected to another layer of cascading filters. This gives much more flexibility in

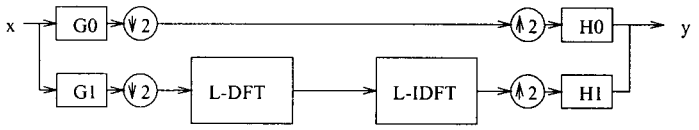


Fig. 7.5.3: Cascaded WPT and DFT.

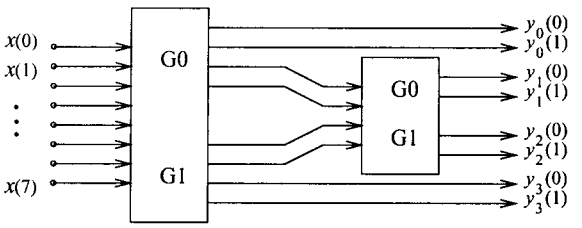


Fig. 7.5.4: An Extension of a WPT. $x(0)$ through $x(7)$ are input signal samples. y_0 through y_3 are the output signals.

choosing the time and frequency resolution for the transform coefficients. In **hybrid WPT**, the filters $G0$ and $G1$ may be different for each cascading layer. All these results can be readily extended to $l^2(\mathbb{Z}/L)$ [10].

We need not restrict ourselves to cascaded filterbanks only. We can replace some of the filterbanks by any frame. Even a frame for a periodic space $l^2(\mathbb{Z}/L)$, when used as a filterbank, provides a frame for $l^2(\mathbb{Z})$, since it is effectively a windowed transform, where the window is a rectangular window of length L moved by L samples for each shift [3]. Thus, we can combine the WPT with the DFT to obtain different TF plane partitions. Fig. 7.5.3 shows such a scheme.

Instead of cascaded filterbanks, we can use the DFT structure repeatedly to obtain various TF plane partitions, resulting in **block transform packets**.

New arrangements of the PRFBs, such as shown in Fig. 7.5.4, can be used to obtain more extensions of the WPT. The box labeled 'G0 G1' is an analysis FB like the one shown in Fig. 7.5.2. Similar to the wavelet transforms, all these frames can be easily extended to $l^2(\mathbb{Z}/L)$ [10].

7.5.4 Summary and Conclusions

In this article, we have seen how two frame structures, product-function frames and cascaded frames, generalize the popular transform techniques for joint-domain representation: discrete-time Gabor transforms for periodic and non-periodic spaces, and the discrete-time wavelet transforms. The necessary and sufficient conditions for the elements to form a frame were stated. It was demonstrated that the popular discrete joint-domain representation techniques, namely, the discrete-time Gabor and wavelet transforms, are special cases of these structures. Several new transforms, namely, product-transform, scale-modulation transform, and various extensions of

the discrete wavelet transforms, were derived as special cases of these structures. These two structures result in a variety of discrete transforms, giving a user more freedom in choosing a transform better suited for a particular type of signal. To conclude, we believe that this work paves the way for many new interesting discrete-domain representation schemes in signal processing.

References

[1] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Information Theory*, vol. 36, pp. 961–1005, September 1990.

[2] C. E. Heil and D. F. Walnut, "Continuous and discrete wavelet transforms," *SIAM Review*, vol. 31, pp. 628–666, December 1989.

[3] S. M. Joshi, *Joint-Domain Representations Using Discrete Domain Frames.* PhD thesis, Computer Science and Electrical Engineering Department, University of Maryland, Baltimore County, Catonsville, MD 21250, July 2000.

[4] S. Li and D. M. Healy, Jr., "A parametric class of discrete Gabor expansions," *IEEE Trans. Signal Processing*, vol. 44, pp. 201–211, February 1996.

[5] S. M. Joshi and J. M. Morris, "Some results on product-function frames," *Signal Processing*, vol. 80, pp. 737–740, April 2000.

[6] M. R. Portnoff, "Time-frequency representation of digital signals and systems based on short-time Fourier analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 28, pp. 55–69, February 1980.

[7] H. Bölcskei and F. Hlawatsch, "Oversampled modulated filter banks," in *Gabor Analysis and Algorithms: Theory and Applications* (H. G. Feichtinger and T. Strohmer, eds.), ch. 9, pp. 295–322, Berlin/Boston: Birkhäuser, 1998.

[8] M. Zibulski and Y. Y. Zeevi, "Discrete multiwindow Gabor-type transforms," *IEEE Trans. Signal Processing*, vol. 45, pp. 1428–1442, June 1997.

[9] J. M. Morris and Y. Lu, "Generalized Gabor expansions of discrete-time signals in $l^2(\mathbb{Z})$ via biorthogonal-like sequences," *IEEE Trans. Signal Processing*, vol. 44, pp. 1378–1391, June 1996.

[10] G. Strang and T. Q. Nguyen, *Wavelets and Filter Banks.* Wellesley, MA: Wellesley-Cambridge Press, 1996.

Chapter 8

Multi-Sensor and Time-Space Processing

This chapter presents time-frequency methods suitable for multi-sensor and time-space processing. The topic is covered in five articles with appropriate cross-referencing to other relevant chapters.

In underwater acoustics and telecommunications, separation of signal mixtures is traditionally based on methods such as Independent Component Analysis (ICA) or Blind Source Separation (BSS). These can be formulated using TFDs for dealing with the case when the signals are non-stationary (Article 8.1). Multi-sensor data can be processed with TFDs for channel estimation and equalization. In Blind Source Separation (BSS) and Direction of Arrival (DOA) estimation problems, the time-frequency approach to array signal processing leads to improved spatial resolution and source separation performances. Methods include time-frequency MUSIC, AD-MUSIC and TFD-based BSS (8.2). In sensor array processing, for source localization, TFDs provide a good framework for hypothesis testing as they possess additional degrees of freedom provided by the t and f parameters (8.3). TFD-based array detection is formulated using the Weyl correspondence. The TFD-based structure allows the optimal detector to be implemented naturally and efficiently (8.4). In the underdetermined case, the time-frequency formulations, methodologies and algorithms for BSS are implemented using two different approaches, namely vector clustering and component extraction. Algorithms and MATLAB^{TM} code for time-frequency BSS using the above-mentioned methods are provided (8.5).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

8.1 BLIND SOURCE SEPARATION USING TIME-FREQUENCY DISTRIBUTIONS [0]

Blind source separation (BSS) is a fundamental problem in signal processing that is sometimes known under different names: blind array processing, signal copy, independent component analysis, waveform preserving estimation, etc. In all these instances, the underlying model is that of n 'statistically' independent signals whose m (possibly noisy) mixtures are observed. Neither the structure of the mixtures nor the source signals are known to the receivers. In this environment, we want to identify the mixtures (blind identification problem) and decouple the mixtures (blind source decoupling).

BSS can have many applications in areas involving processing of multi-sensor signals. Examples of these applications include: Source localization and tracking by radar and sonar devices; speaker separation (cocktail party problem); multiuser detection in communication systems; medical signal processing (e.g., separation of EEG or ECG signals); industrial problems such as fault detection; extraction of meaningful features from data, etc.

Research in this area has been very active over the last two decades. Surprisingly, this seemingly impossible problem has elegant solutions that depend on the nature of the mixtures and the nature of the source statistical information.

Most approaches to blind source separation assume (explicitly or not) that each source signal is a sequence of independently and identically distributed (i.i.d.) variables [1]. In this context, the blind source separation is possible only if at most one of the sources has a Gaussian distribution. In contrast, if the source sequences are *not* i.i.d., we can blindly separate the sources even for Gaussian processes. Several authors have considered the case where each source sequence is a temporally correlated stationary process [2], in which case blind source separation is possible if the source signals have different spectra. Other contributors [3,4], have addressed the case where the second 'i' of 'i.i.d.' is failing, that is, the non stationary case.

Our focus is the exploitation of signal non stationarity for blind source separation. In this case, one can use time frequency analysis to separate and recover the incoming signals. The underlying problem can be posed as a signal synthesis from the time frequency (t-f) plane with the incorporation of the spatial diversity provided by the multisensor. Moreover, the effects of spreading the noise power while localizing the source energy in the t-f domain amounts to increasing the signal to noise ratio (SNR) and hence improved performance. This article reviews the BSS techniques that exploit the joint eigenstructure of a combined set of time frequency distribution matrices. The first part is devoted to the BSS problem of instantaneous mixtures while the second part considers the general case of blind separation of convolutive mixtures.

---

[0]Authors: **Karim Abed-Meraim** (Sig. & Image Proc. Dept., Telecom Paris, France, abed@tsi.enst.fr), **Adel Belouchrani** (EE Dept., Ecole Nationale Polytechnique, Algiers, Algeria, belouchrani@hotmail.com), and **Rahim Leyman** (Center for Wireless Communication, National University of Singapore, larahim@cwc.nus.edu.sg). Reviewers: B. Barkat and A. Cickochi.

8.1.1 Separation of Instantaneous Mixtures

8.1.1.1 Data Model

Consider m sensors receiving an instantaneous linear mixture of signals emitted from $n \leq m$ sources. The $m \times 1$ vector $\mathbf{x}(t)$ denotes the output of the sensors at time instant t which may be corrupted by additive noise $\mathbf{n}(t)$. Hence, the linear data model may be given by:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t), \tag{8.1.1}$$

where the $m \times n$ matrix $\mathbf{A}$ is called the 'mixing matrix'. The n source signals are collected in a $n \times 1$ vector denoted $\mathbf{s}(t)$ which is referred to as the source signal vector. The sources are assumed to have different structures and localization properties in the time-frequency domain. The mixing matrix $\mathbf{A}$ is full column rank but is otherwise unknown. In contrast to traditional parametric methods, no specific structure of the mixture matrix is assumed.

8.1.1.2 Spatial Time-Frequency Distributions

The discrete-time form of a quadratic-class time-frequency distribution (TFD), for a signal $x(t)$, may be expressed (cf. [5] and Article 6.1) as

$$D_{xx}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi(m,l) x(t+m+l) x^*(t+m-l) e^{-j4\pi fl} \tag{8.1.2}$$

where t and f represent the time index and the frequency index, respectively. The kernel $\phi(m,l)$ characterizes the distribution and is a function of both the time and lag variables. The cross-TFD of two signals $x_1(t)$ and $x_2(t)$ is defined by

$$D_{x_1x_2}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi(m,l) x_1(t+m+l) x_2^*(t+m-l) e^{-j4\pi fl} \tag{8.1.3}$$

Expressions (8.1.2) and (8.1.3) are now used to define the following data *spatial time-frequency distribution* (STFD) *matrix*,

$$\mathbf{D_{xx}}(t,f) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \phi(m,l) \mathbf{x}(t+m+l) \mathbf{x}^H(t+m-l) e^{-j4\pi fl} \tag{8.1.4}$$

where $[\mathbf{D_{xx}}(t,f)]_{ij} = D_{x_ix_j}(t,f)$, for $i,j = 1,\cdots,m$. $\mathbf{x}^H$ denotes the conjugate transpose of $\mathbf{x}$.

Under the linear data model of equation (8.1.1) and assuming noise-free environment, the STFD matrix takes the following structure:

$$\mathbf{D_{xx}}(t,f) = \mathbf{A}\mathbf{D_{ss}}(t,f)\mathbf{A}^H \tag{8.1.5}$$

where $\mathbf{D_{ss}}(t,f)$ is the source TFD matrix whose entries are the auto- and cross-TFDs of the sources, respectively, defined as $D_{s_is_i}(t,f)$ and $D_{s_is_j}(t,f)$ for $i \neq j$.

8.1.1.3 BSS based on STFD

Let us now briefly review the principle of blind source separation based on the STFD [3,6].

Whitening: Let $\mathbf{W}$ denotes a $n \times m$ matrix, such that $(\mathbf{WA})(\mathbf{WA})^H = \mathbf{I}$, i.e. $\mathbf{U} = \mathbf{WA}$ is a $n \times n$ unitary matrix (this matrix is referred to as a whitening matrix, since it whitens the signal part of the observations). Pre- and Post-multiplying the STFD matrices $\mathbf{D_{xx}}(t,f)$ by $\mathbf{W}$ leads to the *whitened STFD-matrices*, defined as:

$$\underline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{W}\mathbf{D_{xx}}(t,f)\mathbf{W}^H \tag{8.1.6}$$

From the definition of $\mathbf{W}$ and Eq. (8.1.5), we can express $\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)$ as

$$\underline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{U}\mathbf{D_{ss}}(t,f)\mathbf{U}^H \tag{8.1.7}$$

Note that the whitening matrix can be computed in different ways. It can be obtained, for example, as an inverse square root of the observation autocorrelation matrix [3] or else computed from the STFD matrices as shown in [7].

Source separation using joint diagonalization (JD): By selecting auto-term t-f points, the whitened STFD matrices will have the following structure,

$$\underline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{U}\mathbf{D_{ss}}(t,f)\mathbf{U}^H \tag{8.1.8}$$

where $\mathbf{D_{ss}}(t,f)$ is diagonal.[1] This expression shows that any whitened STFD-matrix is diagonal in the basis of the columns of the matrix $\mathbf{U}$ (the eigenvalues of $\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)$ being the diagonal entries of $\mathbf{D_{ss}}(t,f)$).

If, for a point (t,f) of the time frequency domain, the diagonal elements of $\mathbf{D_{ss}}(t,f)$ are all distinct, the missing unitary matrix $\mathbf{U}$ may be uniquely (up to permutation and scaling ambiguity[2]) retrieved by computing the eigendecomposition of $\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)$. Indeterminacy occurs in the case of degenerate eigenvalues, i.e., when $D_{s_i s_i}(t,f) = D_{s_j s_j}(t,f)$, $i \neq j$. It does not seem possible to a priori choose the t-f point such that the diagonal entries of $\mathbf{D_{ss}}(t,f)$ are all distinct. Moreover, when some eigenvalues of $\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)$ come to degeneracy, the robustness of determining $\mathbf{U}$ from eigendecomposition of a single whitened STFD matrix is seriously impaired. The situation is more favorable when considering joint diagonalization of a combined set $\{\underline{\mathbf{D}}_{\mathbf{xx}}(t_i,f_i)|i = 1,\cdots,p\}$ of p (source auto-term) STFD matrices. This amounts to incorporating several t-f points in the source separation problem which reduces the likelihood of having degenerate eigenvalues. It is noteworthy that two source signals with identical t-f signatures cannot be separated even with the inclusion of all information in the t-f plane.

The joint diagonalization (JD) [2] of a set $\{\mathbf{M}_k|k = 1..p\}$ of p matrices is defined as the maximization of the JD criterion:

[1] Since the off-diagonal elements of $\mathbf{D_{ss}}(t,f)$ are cross-terms, the source TFD matrix is quasi-diagonal for each t-f point that corresponds to a true power concentration, i.e. a source auto-term.

[2] The BSS problem has inherent ambiguity concerning the order and amplitudes of the sources.

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^{p} \sum_{i=1}^{n} |\mathbf{v}_i^H \mathbf{M}_k \mathbf{v}_i|^2 \tag{8.1.9}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_n]$. An efficient joint approximate diagonalization algorithm exists in [2] and it is a generalization of the Jacobi technique [8] for the exact diagonalization of a single normal matrix.

Source separation using joint anti-diagonalization (JAD): By selecting cross-term t-f points, the whitened STFD matrices will have the following structure:

$$\underline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{U}\mathbf{D}_{\mathbf{ss}}(t,f)\mathbf{U}^H \tag{8.1.10}$$

where $\mathbf{D}_{\mathbf{ss}}(t,f)$ is anti-diagonal.[3] The missing unitary matrix $\mathbf{U}$ is retrieved by Joint Anti-Diagonalization (JAD) of a combined set $\{\underline{\mathbf{D}}_{\mathbf{xx}}(t_i,f_i)|i=1,\cdots,q\}$ of q source cross-term STFD matrices [6].

The joint anti-diagonalization is explained by first noting that the problem of anti-diagonalization of a single $n \times n$ matrix $\mathbf{N}$ is equivalent[4] to maximizing

$$C(\mathbf{N},\mathbf{V}) \stackrel{\text{def}}{=} -\sum_{i=1}^{n} |\mathbf{v}_i^H \mathbf{N} \mathbf{v}_i|^2 \tag{8.1.11}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_n]$. Hence, JAD of a set $\{\mathbf{N}_k|k = 1..q\}$ of q $n \times n$ matrices is defined as the maximization of the JAD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^{q} C(\mathbf{N}_k,\mathbf{V}) = -\sum_{k=1}^{q} \sum_{i=1}^{n} |\mathbf{v}_i^H \mathbf{N}_k \mathbf{v}_i|^2 \tag{8.1.12}$$

under the same unitary constraint.

More generally, we can combine joint diagonalization and joint anti-diagonalization of two sets $\{\mathbf{M}_k|k = 1..p\}$ and $\{\mathbf{N}_k|k = 1..q\}$ of $n \times n$ matrices by maximizing the JD/JAD criterion:

$$C(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \left(\sum_{k=1}^{p} |\mathbf{v}_i^H \mathbf{M}_k \mathbf{v}_i|^2 - \sum_{k=1}^{q} |\mathbf{v}_i^H \mathbf{N}_k \mathbf{v}_i|^2 \right) \tag{8.1.13}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_n]$. The combined JD/JAD criterion can be applied to a combined set of p (source auto-term) STFD matrices and q (source cross-term) STFD matrices to estimate the unitary matrix $\mathbf{U}$.

Remarks: (1) The success of the JD or JAD of STFD matrices in determining the unitary matrix $\mathbf{U}$ depends strongly on the correct selection of the auto-term and cross-term points [6]. Therefore, it is crucial to have a selection procedure that

[3]Since the diagonal elements of $\mathbf{D}_{\mathbf{ss}}(t,f)$ are auto-terms, the source TFD matrix is quasi anti-diagonal (i.e. its diagonal entries are close to zero) for each t-f point corresponding to a cross-term.

[4]This is due to the fact that the Frobenius norm of a matrix is constant under unitary transform, i.e. norm($\mathbf{N}$) =norm($\mathbf{V}^H\mathbf{N}\mathbf{V}$).

is able to distinguish between auto-term and cross-term points based only on the STFD matrices of the observations. In [6], a selection approach that exploits the anti-diagonal structure of the source cross-term STFD matrices has been proposed. More precisely, for a source cross-term STFD matrix, we have

$$\begin{aligned} \text{Trace}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)) &= \text{Trace}(\mathbf{U}\mathbf{D}_{\mathbf{ss}}(t,f)\mathbf{U}^H) \\ &= \text{Trace}(\mathbf{D}_{\mathbf{ss}}(t,f)) \approx 0. \end{aligned}$$

Based on this observation, the following testing procedure applies:

$$\begin{aligned} &\text{if } \frac{\text{Trace}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))}{\text{norm}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))} < \epsilon \longrightarrow \text{ decide that } (t,f) \text{ is a cross-term} \\ &\text{if } \frac{\text{Trace}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))}{\text{norm}(\underline{\mathbf{D}}_{\mathbf{xx}}(t,f))} > \epsilon \longrightarrow \text{ decide that } (t,f) \text{ is an auto-term} \end{aligned}$$

where ϵ is a 'small' positive real scalar.

(2) In practice, the source cross-term STFD matrices will not be purely anti-diagonal. This is because some auto-terms, through their side lobes or main lobes, will intrude over the cross-term regions. The cross-terms will be however the dominant components. This situation is similar to the earlier work on joint diagonalization of STFD matrices selecting auto-term points [3], where the source auto-term STFD matrices are not purely diagonal because of cross-term intrusion. This impairment is mitigated by the joint approximation property of the JD/JAD algorithm and by its robustness.

(3) Other classes of TFDs and techniques can also be used in BSS. In [9] a cumulant-based 4th-order Wigner distribution or Wigner trispectrum was used for source separation. In [10] blind separation of *more* sources than sensors (underdetermined BSS problem) has been solved using a TF domain orthogonality concept. Implementation details and the corresponding MATLAB^{TM} code of the above algorithm are presented in Article 8.4. Sample code for computation of TFDs is given in Article 6.5.

8.1.2 Separation of Convolutive Mixtures

8.1.2.1 Data Model

Consider now a convolutive multiple input multiple output (MIMO) linear time invariant model given by:

$$x_i(t) = \sum_{j=1}^{n} \sum_{l=0}^{L} a_{ij}(l) s_j(t-l) \quad \text{for } i = 1, \cdots, m \tag{8.1.14}$$

where $s_j(t)$, $j = 1, \cdots, n$, are the n source signals (model inputs), $x_i(t)$, $i = 1, \cdots, m$, are the $m > n$ sensor signals (model outputs), a_{ij} is the transfer function between the j-th source and the i-th sensor with an overall extent of $(L+1)$ taps. As

before, the sources are assumed to have different structures and localization properties in the time-frequency domain and the channel matrix $\mathbf{A}$ defined in (8.1.16) is full column rank.

In matrix form, we have

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \tag{8.1.15}$$

where

$$\begin{aligned}\mathbf{s}(t) &= [s_1(t), \cdots, s_1(t-(L+L')+1), \cdots, s_n(t-(L+L')+1)]^T \\ \mathbf{x}(t) &= [x_1(t), \cdots, x_1(t-L'+1), \cdots, x_m(t-L'+1)]^T\end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \cdots & \mathbf{A}_{mn} \end{bmatrix} \tag{8.1.16}$$

with

$$\mathbf{A}_{ij} = \begin{bmatrix} a_{ij}(0) & \cdots & a_{ij}(L) & \cdots & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & \cdots & a_{ij}(0) & \cdots & a_{ij}(L) \end{bmatrix}. \tag{8.1.17}$$

Note that $\mathbf{A}$ is a $[mL' \times n(L+L')]$ matrix and $\mathbf{A}_{ij}$ are $[L' \times (L+L')]$ matrices. L' is chosen such that $mL' \geq n(L+L')$.

We retrieve here the same formalism as in the instantaneous mixture case. The data STFD matrices still have the same expression as in (8.1.5). But the source auto-term (respectively, cross-term) matrices $\mathbf{D_{ss}}(t,f)$ are no longer diagonal (respectively, anti-diagonal), but block-diagonal[5] (resp., block anti-diagonal) where each diagonal block is of size $(L+L') \times (L+L')$. It is this block-diagonal or block anti-diagonal structure that we propose to exploit, in the next subsection, to achieve BSS.

8.1.2.2 BSS using STFD Matrices

In this subsection we generalize the BSS method developed earlier to the case of convolutive mixtures.

Whitening: The first step of our procedure consists of whitening the data vector $\mathbf{x}(t)$. This is achieved by applying to $\mathbf{x}(t)$ a *whitening matrix* $\mathbf{W}$, *i.e.* a $[n(L'+L) \times mL']$ matrix verifying:

$$\mathbf{W} \lim_{T\to\infty} \left(\frac{1}{T} \sum_{t=1}^{T} \mathbf{x}(t)\mathbf{x}(t)^H \right) \mathbf{W}^H = \mathbf{W}\mathbf{R}_x\mathbf{W}^H = (\mathbf{W}\mathbf{A}\mathbf{R}_s^{\frac{1}{2}})(\mathbf{W}\mathbf{A}\mathbf{R}_s^{\frac{1}{2}})^H = \mathbf{I} \tag{8.1.18}$$

[5]The block diagonal structure comes from the fact that the cross-terms between $s_i(t)$ and $s_i(t-d)$ are not zero and depend on the local correlation structure of the signal.

where $\mathbf{R}_x$ and $\mathbf{R}_s$ denote the autocorrelation matrices[6] of $\mathbf{x}(t)$ and $\mathbf{s}(t)$, respectively. Equation (8.1.18) shows that if $\mathbf{W}$ is a whitening matrix, then

$$\mathbf{U} = \mathbf{W}\mathbf{A}\mathbf{R}_s^{\frac{1}{2}} \tag{8.1.19}$$

is a $[n(L'+L) \times n(L'+L)]$ unitary matrix where $\mathbf{R}_s^{\frac{1}{2}}$ (Hermitian square root matrix of $\mathbf{R}_s$) is block diagonal. The whitening matrix $\mathbf{W}$ can be determined from the eigendecomposition of the data autocorrelation $\mathbf{R}_x$ as in [3].

Separation using joint block diagonalization: Consider now the whitened STFD matrices $\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)$ defined in (8.1.6). By (8.1.15) and (8.1.19), we obtain the key relation:

$$\underline{\mathbf{D}}_{\mathbf{xx}}(t,f) = \mathbf{U}\mathbf{R}_s^{-\frac{1}{2}}\mathbf{D}_{\mathbf{ss}}(t,f)\mathbf{R}_s^{-\frac{1}{2}}\mathbf{U}^H = \mathbf{U}\mathbf{D}(t,f)\mathbf{U}^H \tag{8.1.20}$$

where we have set $\mathbf{D}(t,f) = \mathbf{R}_s^{-\frac{1}{2}}\mathbf{D}_{\mathbf{ss}}(t,f)\mathbf{R}_s^{-\frac{1}{2}}$.

Since the matrix $\mathbf{U}$ is unitary and $\mathbf{D}(\mathbf{t},\mathbf{f})$ is block diagonal, the latter just means that any whitened STFD matrix is block diagonal in the basis of the column vectors of matrix $\mathbf{U}$. The unitary matrix can be retrieved by computing the block diagonalization of some matrix $\underline{\mathbf{D}}_{\mathbf{xx}}(t,f)$. But to reduce the likelihood of indeterminacy and increase the robustness of determining $\mathbf{U}$, we consider the joint block diagonalization of a set $\{\underline{\mathbf{D}}_{\mathbf{xx}}(t_i,f_i);\ \ i=1,\cdots,p\}$ of p whitened STFD matrices.[7]

This joint block-diagonalization (JBD) is achieved by the maximization under unitary transform of the following criterion,

$$C(\mathbf{U}) \stackrel{\text{def}}{=} \sum_{k=1}^{p}\sum_{l=1}^{n}\sum_{i,j=(L'+L)(l-1)+1}^{(L'+L)l} |\mathbf{u}_i^*\underline{\mathbf{D}}_{\mathbf{xx}}(t_k,f_k)\mathbf{u}_j|^2 \tag{8.1.21}$$

over the set of unitary matrices $\mathbf{U} = [\mathbf{u}_1,\cdots,\mathbf{u}_{n(L'+L)}]$. Note that an efficient Jacobi-like algorithm for joint block diagonalization algorithm exists in [11,12].

Once the unitary matrix $\mathbf{U}$ is determined (up to a block diagonal unitary matrix $\mathbf{D}$ coming from the inherent indeterminacy of the JBD problem [13]), the recovered signals are obtained up to a filter by

$$\hat{\mathbf{s}}(n) = \mathbf{U}^H\mathbf{W}\mathbf{x}(n) \tag{8.1.22}$$

According to (8.1.15) and (8.1.19), the recovered signals verify,

$$\hat{\mathbf{s}}(n) = \mathbf{D}\mathbf{R}_s^{-\frac{1}{2}}\mathbf{s}(n) \tag{8.1.23}$$

where, we recall that, the matrix $\mathbf{R}_s^{-\frac{1}{2}}$ is block diagonal and $\mathbf{D}$ is a block diagonal unitary matrix.

---

[6] In practice, $\mathbf{R}_x$ and $\mathbf{R}_s$ are replaced by their time-averaged estimates, e.g., $\mathbf{R}_x = (\sum_{t=1}^{T}\mathbf{x}(t)\mathbf{x}(t)^H)/T$.

[7] A similar procedure can be used with joint block anti-diagonalization of source cross-term STFD matrices.

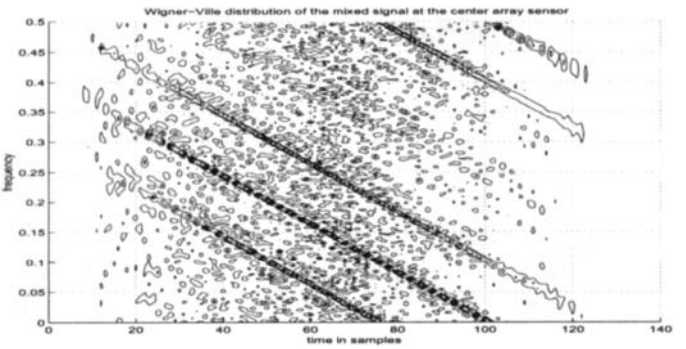


Fig. 8.1.1: WVD of two mixed signals at 0 dB SNR.

Remarks: (1) In practice, only n signals among the $n(L'+L)$ recovered ones are selected. This is done by choosing the signals which lead to the smallest cross-terms coefficients. Note that this information is a byproduct of the joint block diagonalization procedure and hence this selection needs no additional computations.

(2) Here we have considered source separation up to a filter, instead of the full MIMO deconvolution procedure. Note that if needed a SIMO (Single Input Multi Output) deconvolution/equalization [14] can be applied to the separated sources.

8.1.3 Illustrative Examples

The two following examples illustrate the effectiveness of the time frequency approach in achieving blind separation of non-stationary sources.

8.1.3.1 Separation of Instantaneous Mixtures

In this experiment, we consider two chirp signals ($n = 2$), depicted by

$$s_1(t) = \exp(-j0.004\pi t^2) \;\; ; \;\; s_2(t) = \exp(-j0.004\pi t^2 - j\pi 0.4t) \;,$$

impinging on an array of $m = 5$ sensors at 30 and 60 degrees. White Gaussian noise was added, leading to an SNR of 0 dB. The Wigner-Ville distribution (WVD) of the mixture at the middle sensor is depicted in Figure 8.1.1.

From Figure 8.1.1, we selected eight arbitrary t-f points, among which one was a cross-term. By applying the JD/JAD algorithm, we obtain the results displayed in Figure 8.1.2 with a signal rejection level estimate of -26 dB.

8.1.3.2 Separation of Convolutive Mixtures

The parameter settings of this example are as follows:

- $n = 2$ (speech signals sampled at 8kHz), $m = 3$, $L = 1$ and $L' = 2$.

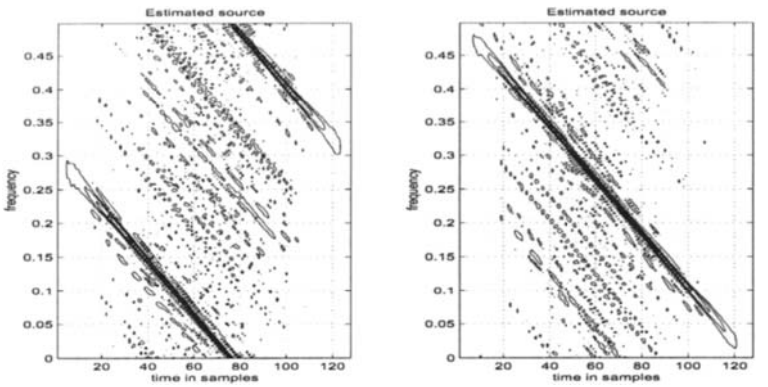


Fig. 8.1.2: WVDs of the two chirps using JD/JAD with seven auto-terms and one cross-term.

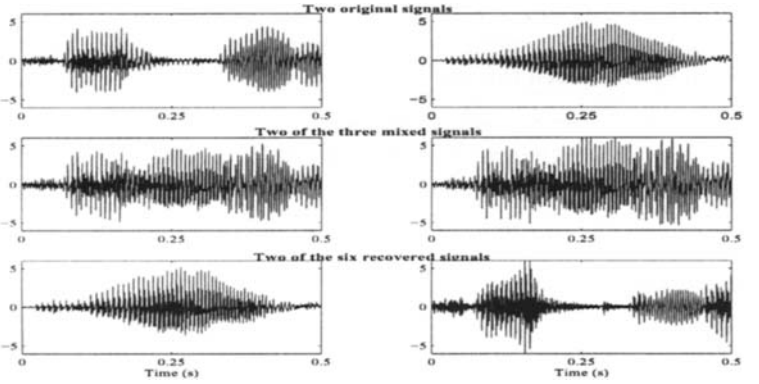


Fig. 8.1.3: Speech signal separation.

- The transfer function matrix of the simulated multi-channel is given by:

$$\mathbf{A}(z)\begin{bmatrix} 1 & 0.85+0.1z^{-1} \\ 0.7+0.4z^{-1} & 0.25+z^{-1} \\ 1+0.5z^{-1} & 0.7+0.85z^{-1} \end{bmatrix}.$$

Figure 8.1.3 shows the original speech signal, their convolutive mixture and the recovered speech signals by the time frequency separation algorithm.

8.1.4 Summary and Conclusions

The problem of blind separation of linear mixtures of non-stationary source signals based on time frequency distributions has been investigated. Both instantaneous and convolutive mixture cases have been considered. For both cases, solutions based on the use of the joint matrix structure (i.e., diagonal, anti-diagonal, block diagonal,

or block anti-diagonal) of a combined set of spatial time frequency distribution matrices selected in both the auto-term and cross-term regions, have been presented.

References

[1] J.-F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 86, pp. 2009–2025, October 1998. Special issue on blind identification and estimation.

[2] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "Blind source separation using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434–444, February 1997.

[3] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, pp. 2888–2898, November 1998.

[4] L. Parra and C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Trans. on Speech & Audio Processing*, vol. 8, pp. 320–327, May 2000.

[5] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[6] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. Zoubir, "Joint anti-diagonalization for blind source separation," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, Salt Lake City, UT, 7–11 May 2001.

[7] Y. Zhang and M. G. Amin, "Blind separation of sources based on their time-frequency signatures," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, vol. 5, pp. 3132–3135, Istanbul, 5–9 June 2000.

[8] G. H. Golub and C. F. Van Loan, *Matrix computations.* Baltimore, MD: Johns Hopkins University Press, 1989.

[9] A. R. Leyman, Z. M. Kamran, and K. Abed-Meraim, "Higher-order time frequency-based blind source separation technique," *IEEE Signal Processing Letters*, vol. 7, pp. 193–196, July 2000.

[10] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, and B. Boashash, "Separating more sources than sensors using time-frequency distributions," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, vol. 2, pp. 583–586, Kuala Lumpur, 13–16 August 2001.

[11] A. Belouchrani, M. G. Amin, and K. Abed-Meraim, "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices," *IEEE Signal Processing Letters*, vol. 4, pp. 266–268, September 1997.

[12] A. Belouchrani, K. Abed-Meraim, and Y. Hua, "Jacobi-like algorithms for joint block diagonalization: Application to source localization," in *Proc. Sixth IEEE Internat. Workshop on Intelligent Signal Processing and Communication Systems (ISPACS'98)*, vol. 1, pp. 133–137, Melbourne, Australia, 5–6 November 1998.

[13] H. Bousbia-Saleh, A. Belouchrani, and K. Abed-Meraim, "Jacobi-like algorithm for blind signal separation of convolutive mixtures," *Electronics Letters*, vol. 37, no. 16, pp. 1049–1050, 2001.

[14] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, pp. 1310–1322, August 1997.

8.2 SPATIAL TIME-FREQUENCY DISTRIBUTIONS AND THEIR APPLICATIONS [0]

8.2.1 Spatial Time-Frequency Distributions

The evaluation of quadratic time-frequency distributions (TFDs) of nonstationary signals impinging on a multi-sensor receiver yields spatial time-frequency distributions (STFDs), which permit the application of eigenstructure subspace techniques to solving a large class of channel estimation and equalization, blind source separation (BSS), and high resolution direction-of-arrival (DOA) estimation problems [1–3]. STFD based techniques are appropriate to handle sources of nonstationary waveforms that are highly localized in the time-frequency (t-f) domain. In the area of BSS, the use of the STFDs allows the separation of sources with identical spectral shape, but with different t-f localization properties, i.e., different t-f signatures. For both source separation and DOA estimation problems, spreading the noise power while localizing the source energy in the t-f domain amounts to increasing the robustness of eigenstructure signal and noise subspace estimation methods with respect to channel and receiver noise. This in turn leads to an improvement of spatial resolution and source separation performance.

The quadratic class of STFD matrix of a signal vector $\mathbf{x}(t)$ is defined as

$$\mathbf{D_{xx}}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\nu,\tau)\,\mathbf{x}(u+\tfrac{\tau}{2})\,\mathbf{x}^H(u-\tfrac{\tau}{2})\,e^{j2\pi(\nu u-\nu t-f\tau)}d\tau\,du\,d\nu, \tag{8.2.1}$$

where $g(\nu,\tau)$ is the kernel function (see Chapter 3).

In narrowband array processing, when n signals arrive at an m-element array (see Fig. 8.2.1), the linear data model

$$\mathbf{x}(t) = \mathbf{y}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{d}(t) + \mathbf{n}(t) \tag{8.2.2}$$

is commonly assumed, where $\mathbf{x}(t)$ is the $m \times 1$ data vector received at the array, $\mathbf{d}(t)$ is the $n \times 1$ source data vector, the $m \times n$ spatial matrix $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_n]$ represents the mixing matrix, $\mathbf{a}_i$ is the steering vector of ith signal, and $\mathbf{n}(t)$ is an additive noise vector whose elements are modeled as stationary, spatially and temporally white, zero-mean complex random processes, independent of the source signals.

Under the uncorrelated signal and noise assumption and the zero-mean noise property, the expectation of the crossterm TFD matrices between the signal and noise vectors is zero, i.e., $E[\mathbf{D_{yn}}(t,f)] = E[\mathbf{D_{ny}}(t,f)] = \mathbf{0}$, and it follows

$$E[\mathbf{D_{xx}}(t,f)] = \mathbf{D_{yy}}(t,f) + E[\mathbf{D_{nn}}(t,f)] = \mathbf{A}\mathbf{D_{dd}}(t,f)\mathbf{A}^H + \sigma^2\mathbf{I}, \tag{8.2.3}$$

where σ^2 is the noise power, and $\mathbf{I}$ is the identity matrix. Eq. (8.2.3) is similar to that which has been commonly used in array processing based on second-order statistics,

[0]Authors: **Moeness G. Amin** and **Yimin Zhang**, Department of Electrical and Computer Engineering, Villanova University, Villanova, PA 19085, USA (moeness@ece.villanova.edu, yimin@ieee.org). Reviewers: A. Gershman and K. Abed-Meraim.

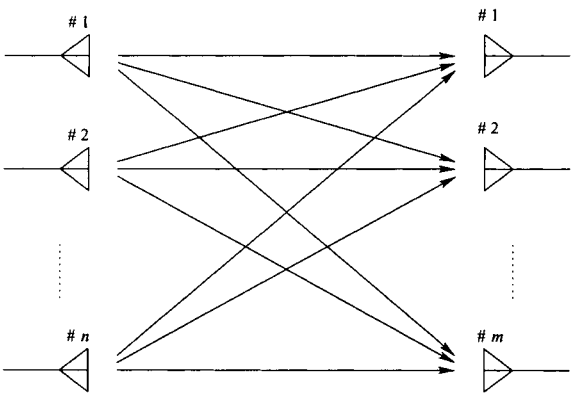


Fig. 8.2.1: m-element array with n signal arrivals.

relating the signal correlation matrix to the data spatial correlation matrix [1]. This implies that key problems in various applications of array processing, specifically those dealing with nonstationary signal environments, can be approached using quadratic transformations. If $\mathbf{D_{dd}}(t,f)$ is a full-rank matrix, the two subspaces spanned by the principal eigenvectors of $\mathbf{D_{xx}}(t,f)$ and the columns of $\mathbf{A}$ become identical. In this case, direction finding techniques based on eigenstructures can be applied. If $\mathbf{D_{dd}}(t,f)$ is diagonal, i.e., the signal cross-TFDs at the t-f point (t,f) are zeros, then both the mixing matrix and the signal waveforms can be recovered using BSS methods.

8.2.2 Fundamental Properties

There are five key advantages of array signal processing using STFD. In order to properly explain these advantages, we use the diagram in Fig. 8.2.2. We consider two sources A and B incident on a multi-sensor array. Source A occupies the t-f region R_a, whereas source B occupies the t-f region R_b. The t-f signatures of the two sources overlap, but each source still has a t-f region that is not intruded over by the other source.

(1) Eq. (8.2.3) can be easily derived for any arbitrary joint-variables. Time and frequency are indeed the two most commonly used and physically understood parameters. However, by replacing the STFDs by spatial arbitrary joint-variable distributions, one can relate the sensor joint-variable distributions to the sources joint-variable distributions through the same mixing matrix $\mathbf{A}$. As shown in the Examples section, there are situations where it is preferable to consider other domains such as the ambiguity lag-Doppler domain, where the locations of the signals and their cross-terms are guided by properties and mechanisms different than those associated with the t-f domain (see Article 3.2).

(2) Eq. (8.2.3) is valid for all t-f points. It is well known that direction finding techniques require $\mathbf{D_{dd}}(t,f)$ to be full rank, preferably diagonal. On the other

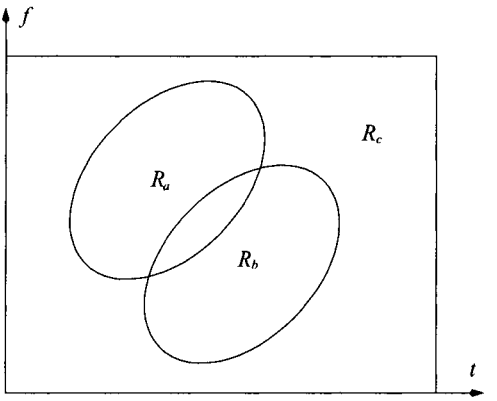


Fig. 8.2.2: Signals with different time-frequency signature.

hand, BSS techniques demand the diagonal structure of the same matrix without degenerate eigenvalues. These properties along with high signal-to-noise ratio (SNR) requirements may be difficult to achieve using a single t-f point. Two different methods can be used for integrating several t-f points into Eq. (8.2.3). One method is based on a simple averaging performed over the signatures of the sources of interest, whereas the second method is based on incorporating several desired t-f points into joint diagonalization or joint block-diagonalization schemes.

(3) The TFD of the white noise is distributed all over the t-f domain, whereas the TFDs of the source waveforms are likely to be confined to much smaller regions. Referring to Fig. 8.2.2, the noise is spread over both R_a and R_b as well as the complement region R_c. If the t-f points (t, f) used in either the averaging or joint diagonalization approaches belong to the noise only region R_c, then no information of the incident waveforms is used and, as such, no reasonable source localization and signal separation outcomes can be obtained. On the other hand, if all points (t, f) in Fig. 8.2.2 are used, and the employed TFD satisfies the marginal constraints, then it can be easily shown that only the signal average power is considered. As a result, the problem simplifies to the second-order covariance based matrix approach, traditionally used in high resolution DOA estimation. This is an important feature, as it casts the conventional techniques as special cases of the array signal processing framework based on t-f analysis. Finally, if we confine the (t, f) points to R_a and R_b, then only the noise part in these regions is included. The result of leaving out the points (t, f) that are not part of the t-f signatures of the signal arrivals is enhancing the input SNR, which is utilized by the source localization and signal separation techniques.

(4) By only selecting t-f points that belong to the t-f signature of one source, then this source will be the only one considered by Eq. (8.2.3). This selection, in essence, is equivalent to implicitly performing spatial filtering and removing other sources from consideration. It is important to note, however, that such removal does

not come at the expense of reduction of the number of degrees-of-freedom (DOFs), as it is the case in beamspace processing, but the problem remains a sensor space processing with the original number of DOFs kept intact. This property represents a key contribution of TFDs to the direction finding and DOA estimation areas. An antenna array can be used to localize a number of sources equal or even greater than its number of sensors. The fundamental condition is that there must be t-f regions over which the respective t-f signatures of the sources do not overlap. Referring to Fig. 8.2.2 and considering the case of two sensors, if all t-f points incorporated in direction finding belong to region R_a and not R_b, then the signal subspace defined by Eq. (8.2.3) is one-dimensional. Thus, by excluding source B, a one-dimensional noise subspace is established. This allows us to proceed with high resolution techniques for localization of source A. In a general scenario, one can localize one source at a time or a set of selected sources, depending on the array size, overlapping and distinct t-f regions, and the dimension of the noise subspace necessary to achieve the required resolution performance. The same concepts and advantages of t-f point selection discussed above for direction finding can be applied to BSS problems.

(5) The *a priori* knowledge of some temporal characteristics or the nature of time-varying frequency contents of the sources of interest may permit us to directly select the t-f regions used in Eq. (8.2.3). For instance, it is known that, in the ambiguity domain, all fixed frequency sinusoidal signals map to the time-lag axis. By only incorporating the points on this axis, we have, in fact, opted to separate and localize all narrowband signals in broadband communications platforms.

8.2.3 Examples

In this Section, we present simulation examples to demonstrate the fundamental offerings discussed in the previous Section. Time-frequency MUSIC (t-f MUSIC), ambiguity-domain MUSIC (AD-MUSIC), and the BSS based on STFDs are three different techniques chosen for the demonstration. The algorithms involved in the implementation of the techniques are given in Tables 8.2.1, 8.2.2 and 8.2.3 [1,2,4].

Example I [4]. Consider the scenario of a four-element equi-spaced linear array spaced by half a wavelength, where one chirp signal and two sinusoidal signals are received. The data record has 128 samples. All three signals have the same SNR of 20 dB. The DOAs of the chirp signal and the two sinusoidal signals are 15°, 10°, and 0°, respectively. While the ambiguity function of the chirp signal sweeps the ambiguity domain with contribution at the origin, the exact autoterm ambiguity function of the narrowband arrivals $s_1(t)$ and $s_2(t)$ is zero for non-zero frequency-lags and may have non-zero values only along the vertical axis $\nu = 0$.

In this simulation example, we selected 24 points on the time-lag axis, excluding the origin, and as such emphasizing the narrowband components. Fig. 8.2.3 shows the ambiguity function where the two vertical lines away from the origin represent the crossterms between the sinusoidal components. Fig. 8.2.4 shows the two estimated spatial spectra for three independent trials, one corresponds to the

Table 8.2.1: Time-Frequency MUSIC.

STEP I	Form K matrices $\mathbf{D_{xx}}(t_i, f_i)$ for the selected (t_i, f_i) points, $i = 1, \cdots, K$.
STEP II	The eigenvectors of $E[\mathbf{D_{xx}}(t,f)]$ corresponding to the $m-n$ smallest eigenvalues,$\mathbf{e}_1, \cdots, \mathbf{e}_{m-n}$, are obtained by joint block-diagonalization, or the eigen-decomposition of averaged matrix $\frac{1}{K}\sum_{i=1}^{K}\mathbf{D_{xx}}(t_i, f_i)$.
STEP III	Estimate the number of signals from the eigenvalues, and estimate the DOAs from the peaks of the t-f MUSIC spectra $f(\theta) = \left\|\hat{\mathbf{E}}_n^H \mathbf{a}(\theta)\right\|^{-2}$, where $\hat{\mathbf{E}}_n = [\mathbf{e}_1, \cdots, \mathbf{e}_{m-n}]$, and $\mathbf{a}(\theta)$ is the steering vector corresponding to DOA θ.

Table 8.2.2: Ambiguity-Domain MUSIC

Ambiguity-Domain MUSIC follows the same procedure as time-frequency MUSIC by using $\mathbf{D_{xx}}(\nu_i, \tau_i)$ instead of $\mathbf{D_{xx}}(t_i, f_i)$, $i = 1, \cdots, K$.

conventional method and the other corresponds to the AD-MUSIC. There are two dominant eigenvalues for the case of the AD-MUSIC, since we have not deliberately considered the chirp signal through our careful selection of the ambiguity-domain points. It is clear that the AD-MUSIC resolves the two sinusoidal signals, while the conventional MUSIC could not separate the three signals.

Example II [5]. Consider a uniform linear array of eight sensors separated by half a wavelength. Two chirp signals emitted from two sources positioned at $(\theta_1, \theta_2) = (-10°, 10°)$, respectively. The data record has 1024 samples. The start and end frequencies of the chirp signal of the source at θ_1 are $f_{s1} = 0$ and $f_{e1} = 0.5$, while the corresponding two frequencies for the signal of the other source at θ_2 are $f_{s2} = 0.5$ and $f_{e2} = 0$, respectively.

Fig. 8.2.5 displays the standard deviations of the DOA estimation $\hat{\theta}_1$ versus SNR. The curves in this figure show the theoretical and experimental results of the

Table 8.2.3: Blind Source Separation Based on STFDs

STEP I	Estimate the auto-correlation matrix $\hat{\mathbf{R}}_{\mathbf{xx}}$ from T data samples. Denote by $\lambda_a, \cdots, \lambda_n$ the n largest eigenvalues and the corresponding eigenvectors of $\hat{\mathbf{R}}_{\mathbf{xx}}$.
STEP II	An estimate $\hat{\sigma}^2$ of the noise variance is the average of the $m-n$ smallest eigenvalues of $\hat{\mathbf{R}}_{\mathbf{xx}}$. The whitening matrix is formed as $\hat{\mathbf{W}} = \left[(\lambda_1 - \hat{\sigma}^2)^{-\frac{1}{2}}\mathbf{h}_1, \cdots, (\lambda_n - \hat{\sigma}^2)^{-\frac{1}{2}}\mathbf{h}_n\right]^H$.
STEP III	Form K matrices by computing the STFD of whitened vector $\mathbf{z}(t) = \hat{\mathbf{W}}\mathbf{x}(t)$ for a fixed set of (t_i, f_i) points, $i = 1, \cdots, K$, corresponding to signal autoterms.
STEP IV	A unitary matrix $\hat{\mathbf{U}}$ is then obtained as joint diagonalizer of the set $\mathbf{D}_{\mathbf{zz}}(t_i, f_i), i = 1, \cdots, K$.
STEP V	The source signals are estimated as $\hat{\mathbf{s}}(t) = \hat{\mathbf{U}}^H\hat{\mathbf{W}}\mathbf{x}(t)$, and the mixing matrix $\mathbf{A}$ is estimated as $\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#}\hat{\mathbf{U}}$.

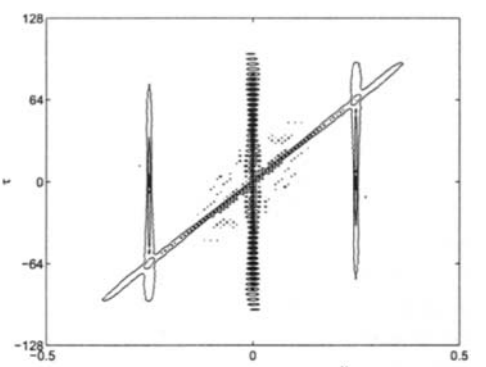


Fig. 8.2.3: The ambiguity functions of the chirp signal and two sinusoidal signals.

conventional MUSIC and t-f MUSIC. Pseudo Wigner-Ville distribution with window length $L = 33$ and 129 are considered. The Cramer-Rao Bound (CRB) is also shown in Fig. 8.2.5. Both signals are selected when performing t-f MUSIC. Simulation

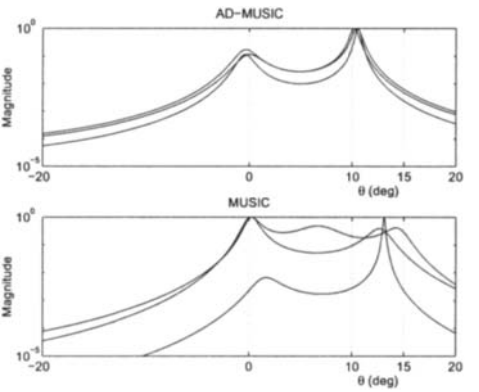


Fig. 8.2.4: The estimated spatial spectra of AD-MUSIC and conventional MUSIC.

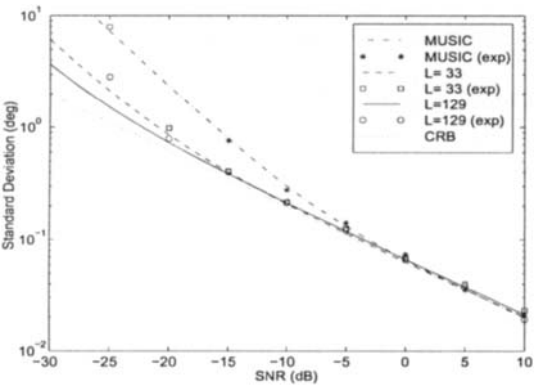


Fig. 8.2.5: The standard deviations of DOA estimation $\hat{\theta}_1$ vs. SNR.

results are averaged over 100 independent trials of Monte Carlo experiments. The advantages of t-f MUSIC in low SNR cases are evident from this figure. Fig. 8.2.6 shows estimated spatial spectra at SNR=-20 dB based on t-f MUSIC ($L = 129$) and the conventional MUSIC. The t-f MUSIC spectral peaks are clearly resolved.

Example III [1]. In Fig. 8.2.7, we show an example of the application of STFDs to the BSS problem. A three-element equi-spaced linear array is considered where the interelement spacing is half a wavelength. Two chirp signals arrive at $-10°$ and $10°$, respectively. The number of data samples used to compute the STFD is 128. The number of t-f points employed in the joint diagonalization is p=128, with equal number of points on each signature. Fig. 8.2.7(b) shows the Choi-Williams distributions of two linear mixtures of the original chirp signals depicted in Fig. 8.2.7(a), corresponding to the data at the first and the second sensors. Using the STFDs, we are able to recover the original signals from their observed mixture, as shown in Fig. 8.2.7(c).

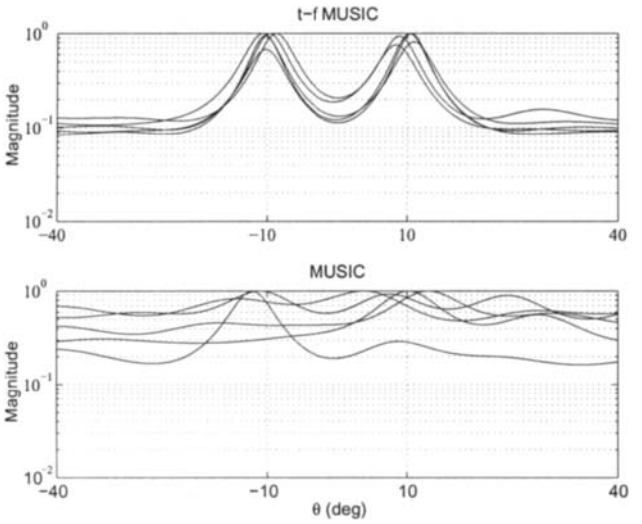
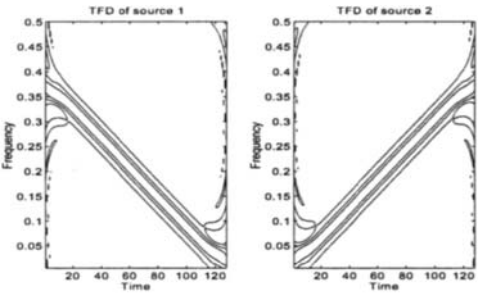


Fig. 8.2.6: The estimated spatial spectra of t-f MUSIC and conventional MUSIC.

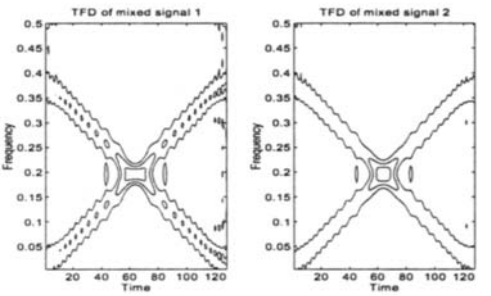
8.2.4 Crossterm Issues in STFD

There are two sources of crossterms. The first type are the crossterms that are the results of the interactions between the components of the same source signal. The other type of crossterms are those generated from the interactions between two signal components belonging to two different sources. These crossterms are associated with cross-TFDs of the source signals and, at any given t-f point, they constitute the off-diagonal entries of the source TFD matrices $\mathbf{D_{dd}}(t, f)$ defined in (8.2.3). Although the off-diagonal elements do not necessarily destroy the full-rank matrix property necessary for direction finding application [6], they violate the basic assumption in the problem of source separation regarding the diagonal structure of the source TFD matrix. We must therefore select the t-f points that belong to autoterm regions where crossterm contributions are at minimum, e.g., by using *a priori* information of the source signals.

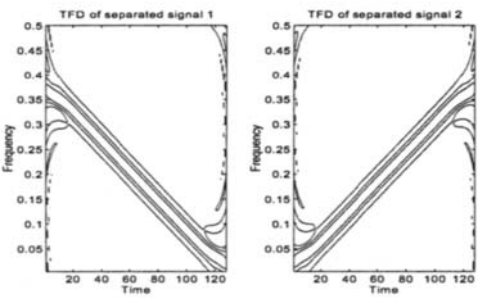
The method of spatial averaging of the STFD introduced in [7] does not reduce the crossterms as in the case with reduced interference distribution kernels, but rather move them from their locations on the off-diagonal matrix entries to be part of the matrix diagonal elements. The other parts of the matrix diagonal elements represent the contribution of the autoterms at the same point. Therefore, not only we are able to set the off-diagonal elements of the source TFD matrix to zeros, but also we can improve performance by selecting the t-f points of peak values, irrespective of whether these points belong to autoterm or crossterm regions.



(a) TFDs of the source signals



(b) TFDs of the mixed signals



(c) TFDs of the separated signals

Fig. 8.2.7: Blind source separation based on STFDs.

8.2.5 Summary and Conclusions

The spatial time-frequency distribution (STFD) is an important tool for temporal and spatial separations of sources emitting nonstationary signals. It is a discriminatory tool that allows a consideration of only a subset of source signals impinging on a multi-sensor receiver. This property enhances signal parameter estimation and permits direction finding and signal separation to be applied to a number of sources that is equal or even exceeds the number of sensors.

All material presented in this article is based on the model (8.2.2). One important change in the direction of the research in the time-frequency array signal processing area was given in [8], where the strict model of (8.2.2) was relaxed and a direction finding technique employing a STFD-based wideband root-MUSIC was proposed. Another research direction is the utilization and integration of crossterms into STFDs. It has recently been shown [9] that source separation can be performed based on both autoterms and crossterms through joint diagonalization and joint anti-diagonalization schemes of STFD matrices. Article 8.4 presents algorithms and sample code for these methods.

References

[1] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, pp. 2888–2898, November 1998.

[2] A. Belouchrani and M. G. Amin, "Time-frequency MUSIC: A new array signal processing method based on time-frequency signal representation," *IEEE Signal Processing Letters*, vol. 6, pp. 109–110, May 1999.

[3] Y. Zhang, W. Mu, and M. G. Amin, "Time-frequency maximum likelihood methods for direction finding," *J. Franklin Institute*, vol. 337, pp. 483–497, July 2000.

[4] M. G. Amin, A. Belouchrani, and Y. Zhang, "The spatial ambiguity function and its applications," *IEEE Signal Processing Letters*, vol. 7, pp. 138–140, June 2000.

[5] Y. Zhang, W. Mu, and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Trans. Signal Processing*, vol. 49, pp. 747–759, April 2001.

[6] M. G. Amin and Y. Zhang, "Direction finding based on spatial time-frequency distribution matrices," *Digital Signal Processing: A Review Journal*, vol. 10, pp. 325–339, October 2000.

[7] Y. Zhang and M. G. Amin, "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays," *IEEE Trans. Signal Processing*, vol. 48, pp. 2892–2902, October 2000.

[8] A. B. Gershman and M. G. Amin, "Wideband direction-of-arrival estimation of multiple chirp signals using spatial time-frequency distributions," *IEEE Signal Processing Letters*, vol. 7, pp. 152–155, June 2000.

[9] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. Zoubir, "Joint anti-diagonalization for blind source separation," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, Salt Lake City, UT, 7–11 May 2001.

8.3 QUADRATIC DETECTION IN ARRAYS USING TFDs[0]

The detection of signals in noise is a classical hypothesis-testing problem. The use of a sensor array can considerably enhance signal detection by providing a large gain in the SNR and allowing for signal source localization. The long-lasting interest devoted to the field of sensor array processing can be traced to the large number of applications where data is collected in both space and time. Some important applications include radar/sonar signal processing; ultrasonic, optical, and tomographic imaging; earth geophysical exploration such as crust mapping and oil exploration; and more recently space-time processing for wireless communications. We assume the very popular uniform linear array (ULA) geometry and a single signal source arriving from the far-field at some angle θ with respect to the array. The goal is to detect the presence of this signal source from observations contaminated by additive noise at the sensors.

In many situations, the channel may induce unknown modifications to the time and frequency parameters of the transmitted signal. Such situations include the well-known delay-Doppler situation in radar/sonar detection problems. Time-frequency distributions (TFDs), which jointly describe the time and frequency characteristics of the signal, are powerful tools for designing the optimal detector in such situations. It has been shown that the optimal quadratic detector for the detection of nonstationary Gaussian signals (i.e. signals whose statistics vary rapidly but are Gaussian at each fixed time instance) with unknown time and frequency offsets in the presence of noise can be implemented naturally within the quadratic class of TFDs [1]. Here we show how this idea can be extended to detection using an array of sensors, even those which exhibit imperfect spatial coherence. Imperfect coherence (i.e. a loss in the spatial correlation for a fixed time instance) is often attributed to signal scattering in multipath channels. Specifically, we consider the problem of detecting arbitrary nonstationary second-order signals with unknown time and frequency offsets arriving in a linear array with an unknown angle of arrival. We explicitly show how the optimal detector for such a problem can be implemented naturally and efficiently in the time-frequency domain.

8.3.1 The Detection Problem

Consider the following composite hypothesis-testing problem in continuous time:

$$\begin{aligned} H_0 : x(t) &= n(t) \\ H_1 : x(t) &= s(t-\tau)e^{j2\pi\nu t} + n(t) \end{aligned} \tag{8.3.1}$$

where $t \in T$, the time interval of observation, x is the observed signal, n is zero-mean complex white Gaussian noise with variance σ^2, and s is a zero-mean complex arbitrary second-order signal with correlation function $R_s(t_1, t_2)$. The parameters

[0]Authors: **Anil M. Rao** and **Douglas L. Jones**, Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign (anil-rao@dsp.csl.uiuc.edu, jones@dsp.csl.uiuc.edu). Reviewers: Akbar M. Sayeed and Graeme S. Jones.

(τ, ν) represent time and frequency offset parameters that are assumed to be unknown. These parameters arise in the classic radar delay/Doppler scenario where the delay is due to the range of the target and the Doppler is due to the velocity of the target. In statistical hypothesis testing, for each observation, x, a real-valued test statistic, $L(x)$, is compared to a threshold to decide in favor of H_0 or H_1; that is, to decide whether the signal is present or not. TFDs provide a natural detection framework for such hypothesis-testing problems for two main reasons: first, detecting a second-order signal (such as a Gaussian signal) in the presence of Gaussian noise involves a quadratic function of the observations [2], and bilinear TFDs are quadratic in the observations; second, TFDs possess additional degrees of freedom provided by the time and frequency parameters.

8.3.2 Quadratic Detection in a Linear Array

In the linear array configuration, the signal comes in to the array of M sensors with spacing d at angle θ, where θ is assumed to be unknown. We will denote the signal at the ith sensor by $x_i(t)$, $i = 1, \ldots, M$. Due to the linear array configuration, the signal at the ith sensor is a delayed version of the signal at the first sensor, and the value of the delay depends on the unknown angle of arrival, θ. That is, when the signal is present $x_i(t) = s(t - (i-1)D) + n_i(t)$, where $D = \frac{d}{c}\sin(\theta)$ and c is the velocity of propagation in the medium. A concern arises when considering the use of a very large array in order to achieve high array gain; the signal received at widely separated sensors may have reduced coherence due to the complexity in the propagation of the signal from the source to spatially separated receivers (see [3–6] and references therein). Loss of spatial coherence can often be attributed to complex signal scattering arising in multipath propagation [7]. In particular, the larger the angular spread of the multipath the more rapid the loss in spatial coherence (the smaller the coherence length). Since we are only considering the second-order statistics of the signal, the model for partial coherence used will be given in terms of the correlation function. An exponential power-law model has been suggested [3,8] whereby the cross-correlation function between the ith and jth sensors will be scaled by the coefficient $c_{ij} = e^{-\frac{|i-j|}{L}}$, where L is a dimensionless characteristic correlation length. We may arrange the decorrelation coefficients in matrix form as $\mathbf{C} = \{c_{ij}\}$. It will be convenient to deal with the aligned sensor outputs; that is, let $y_i(t) = x_i(t + (i-1)D)$. It will also be convenient to arrange the aligned sensor outputs in vector form as $\mathbf{Y}^\theta = [y_1(t)\ y_2(t)\ \ldots\ y_M(t)]^T$ where the superscript θ denotes the dependence of aligning the sensor signals on the unknown angle of arrival.

The optimal test statistic based on the deflection[1] criterion and using a gen-

[1]Deflection-optimal detectors can be interpreted as "maximum signal-to-noise ratio (SNR)" detectors because deflection is a measure of SNR. The deflection is defined as

$$q = \frac{E[L(x)|H_1] - E[L(x)|H_0]}{Var[L(x)|H_0]}.$$

eralized likelihood ratio test (GLRT) technique to deal with the unknown time, frequency, and angle parameters is given by

$$L_{opt} = \max_{(\tau,\nu,\theta)} \langle \left(\mathbf{C} \otimes \mathbf{R}_s^{(\tau,\nu)} \right) \mathbf{Y}^\theta, \mathbf{Y}^\theta \rangle, \tag{8.3.2}$$

where $\otimes$ represents the Kronecker product, $\mathbf{R}_s$ denotes the linear operator defined by the corresponding correlation function R_s as

$$(\mathbf{R}_s x)(t) = \int R_s(t,\tau) x(\tau) d\tau \ , \tag{8.3.3}$$

and $R_s^{(\tau,\nu)}(t_1,t_2) = R_s(t_1-\tau, t_2-\tau) e^{j2\pi\nu t_1} e^{-j2\pi\nu t_2}$ denotes the correlation function corresponding the signal source with the unknown time and frequency offsets. Upon expanding the inner product in (8.3.2) we obtain

$$L_{opt} = \max_{(\tau,\alpha,\theta)} \sum_{i=1}^{M} \sum_{j=1}^{M} c_{ij} \langle \mathbf{R}_s^{(\tau,\nu)} y_i^\theta(t), y_j^\theta(t) \rangle. \tag{8.3.4}$$

8.3.3 TFD Based Array Detection

The connection to TFDs is made through the use of the Weyl correspondence which relates inner products, positive definite linear operators, and the Wigner distribution. Using the fact that the Weyl correspondence involves a covariance to time, frequency, and scale offsets, using the methods in [1] it can be verified that the test statistic in (8.3.4) may conveniently be expressed in terms of TFDs, allowing for a natural and efficient implementation of the optimal detector. The optimal test statistic may be expressed as

$$L_{opt} = \max_{(t,f,\theta)} \sum_{i=1}^{M} \sum_{j=1}^{M} c_{ij} p^\theta_{y_i y_j}(t,f; g = WS_{R_s}), \tag{8.3.5}$$

where g is the Doppler-lag kernel filter and WS_{R_s} is the Weyl-symbol of R_s defined by

$$WS_{R_s}(u,v) = \int R_s(u+\tau/2, u-\tau/2) e^{-j2\pi v\tau} d\tau \ . \tag{8.3.6}$$

We use the superscript θ here to denote the fact that the TFD must be formed for each hypothesized angle of arrival. Observe that in (8.3.5) we must form the sum of all weighted cross-TFDs; we will refer to this quantity as a matrix TFD. Figure 8.3.1 illustrates the detector structure for the partially coherent case. Because the detector involves forming TFDs of signals that are aligned to examine different spatial directions, we may think of this detection structure in terms of time-frequency-space or time-scale-space. Since we have not assumed any spatial statistical characteristics of the signal, the kernel is the same regardless of the angle of arrival being analyzed.

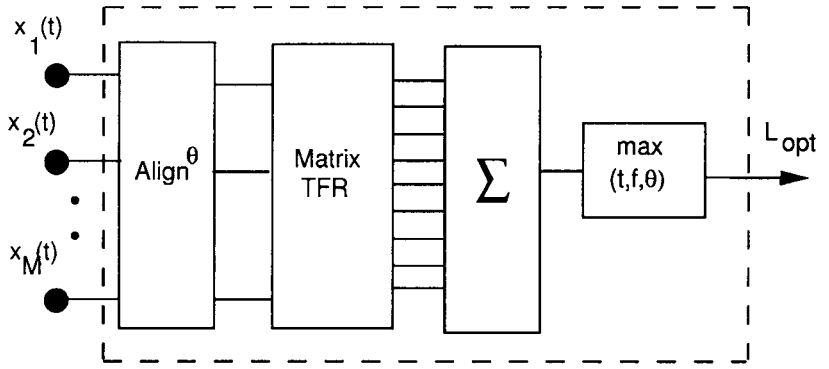


Fig. 8.3.1: TFD-Based Optimal Quadratic Array Processor.

If the array environment is perfectly coherent, then $c_{ij} = 1 \ \forall i, j$. Using fundamental properties of bilinear TFDs, it can be verified that the optimal test statistic will involve first summing the sensor observations and then applying the TFD (with the kernel as before) for each hypothesized angle of arrival and choosing the maximum value. Hence the matrix processing of Figure 8.3.1 is replaced with simple summation. If the array environment is noncoherent, then $\mathbf{C} = \mathbf{I}$ and it can be verified that the optimal test statistic will include first taking the TFD of each sensor observation (again, with kernel as before), and then summing the resulting TFDs for each hypothesized angle of arrival and choosing the maximum value.

8.3.4 Summary and Conclusions

We demonstrated that time-frequency based detectors are naturally suited to quadratic detection in an array environment. By using a GLRT approach, the deflection-optimal test statistic was cast in the form of TFDs. The TFD-based structure allows the optimal detector to be implemented naturally and efficiently by exploiting the many degrees of freedom available. In the general case of a partially coherent environment, the test statistic included a weighted sum of all cross-TFDs of the aligned sensor outputs for each value of hypothesized angle of arrival. Completely coherent and noncoherent cases were shown to be special cases of the partially coherent model. In the coherent case, the optimal test statistic simplified to include a single auto-TFD of the sum of the aligned sensor outputs for each hypothesized angle of arrival. In the noncoherent case, the optimal test statistic simplified to include the sum of auto-TFDs of the aligned sensor outputs for each hypothesized angle of arrival.

References

[1] A. M. Sayeed and D. J. Jones, "Optimal detection using bilinear time-frequency and time-scale representations," *IEEE Trans. Signal Processing*, vol. 43, pp. 2872–2883, December 1995.

[2] H. V. Poor, *An Introduction to Signal Detection and Estimation.* New York: Springer, 1988.

[3] D. R. Morgan and T. M. Smith, "Coherence effects on the detection performance of quadratic array processors with application to large-array matched-field beamforming," *J. Acoustical Soc. of America*, vol. 87, pp. 737–747, February 1990.

[4] A. M. Rao and D. L. Jones, "A denoising approach to multisensor signal estimation," *IEEE Trans. Signal Processing*, vol. 48, pp. 1225–1234, May 2000.

[5] A. M. Rao and D. L. Jones, "Nonstationary array signal detection using time-frequency and time-scale representations," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98)*, vol. 4, pp. 1989-1992, Seattle, 12–15 May 1998.

[6] A. M. Rao and D. L. Jones, "Efficient structures for quadratic time-frequency and time-scale array processors," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 397–400, Pittsburgh, PA, 6–9 October 1998.

[7] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Communications*, vol. 5, pp. 10–22, February 1998.

[8] H. Cox, "Line array performance when the signal coherence is spatially dependent," *J. Acoustical Soc. of America*, vol. 54, pp. 1743–1746, July 1973.

8.4 IMPLEMENTATION OF STFDs-BASED SOURCE SEPARATION ALGORITHMS [0]

Spatial time-frequency distributions (STFDs) are a generalization of the time-frequency distributions to a vector of multi-sensor signals.

Under a linear model, the STFDs and the commonly known covariance matrix exhibit the same eigenstructure. In applications involving multi-sensor signals, the aforementioned structure is often exploited to estimate some signal parameters through subspace-based techniques.

Algorithms based on the covariance matrix assume stationary signals. When the frequency content of the measured signals is time-varying, approaches of this class can still be applied; however, they do not use the signal time-frequency information. Algorithms based on STFDs properly use this time-frequency information to significantly improve performance. This improvement comes essentially from the fact that the effects of spreading the noise power while localizing the source energy in the (t, f) domain amounts to increasing the signal-to-noise ratio (SNR).

STFDs-based algorithms exploit the time-frequency representation of the signals together with the spatial diversity provided by the multiple sensors.

The concept of the STFD was introduced for the first time in 1996. It was used successfully in solving the problem of the blind separation of non-stationary signals [1–3]. This concept was then applied to solve the problem of direction of arrival (DOA) estimation [4]. Since then, other works have been conducted in this area using the new concept of STFDs [5–9].

This article examines the implementation of STFDs-based source separation algorithms. Some theoretical aspects of these algorithms are presented in Articles 8.1 and 8.2. The MATLAB^{TM} codes of the aforementioned algorithms are provided in the next sections.

8.4.1 The Spatial TFD (STFD)

Given an analytic signal vector $\mathbf{z}(t)$, the spatial instantaneous autocorrelation function (SIAF), which is the generalization of the instantaneous autocorrelation function (IAF) defined in Chapter 2, is given by

$$\mathbf{K}_{\mathbf{z},\mathbf{z}}(t,\tau) = \mathbf{z}(t+\tfrac{\tau}{2})\,\mathbf{z}^*(t-\tfrac{\tau}{2}) \tag{8.4.1}$$

Define also the smoothed SIAF as

$$\mathbf{R}_{\mathbf{z},\mathbf{z}}(t,\tau) = G(t,\tau) \underset{t}{*} \mathbf{K}_{\mathbf{z},\mathbf{z}}(t,\tau), \tag{8.4.2}$$

where $G(t,\tau)$ is the time-lag kernel. The time convolution operator $*_t$ is applied to each entry of the matrix $\mathbf{K}_{\mathbf{z},\mathbf{z}}(t,\tau)$. The class of quadratic Spatial TFDs (STFDs)

[0]Authors: **Adel Belouchrani**, EE Dept., Ecole Nationale Polytechnique, Algiers, Algeria (belouchrani@hotmail.com). Reviewers: K. Abed-Meraim and H. Bousbia-Salah.

are then defined as

$$\mathbf{D}_{\mathbf{z},\mathbf{z}}(t,f) = \underset{\tau \to f}{\mathcal{F}}\{\mathbf{R}_{\mathbf{z},\mathbf{z}}(t,\tau)\}, \tag{8.4.3}$$

where the Fourier transform $\mathcal{F}$ is applied to each entry of the matrix $\mathbf{R}_{\mathbf{z},\mathbf{z}}(t,\tau)$.

The discrete-time definition equivalent to Eqs. (8.4.3) and (8.4.2), which leads to the simple implementation of a STFD, is

$$\mathbf{D}_{\mathbf{z},\mathbf{z}}(n,k) = \underset{m \to k}{\mathcal{DF}}\{G(n,m) \underset{n}{*} \mathbf{K}_{\mathbf{z},\mathbf{z}}(n,m)\}, \tag{8.4.4}$$

which can be written

$$\mathbf{D}_{\mathbf{z},\mathbf{z}}(n,k) = \sum_{m=-M}^{M} \sum_{p=-M}^{M} G(n-p,m)\mathbf{z}(p+m)\mathbf{z}^*(p-m)e^{-j4\pi\frac{mk}{N}}, \tag{8.4.5}$$

where the discrete Fourier transform $\mathcal{DF}$ and the discrete time convolution operator $*_n$ are applied to each entry of the matrix $G(n,m) *_n \mathbf{K}_{\mathbf{z},\mathbf{z}}(n,m)$ and matrix $\mathbf{K}_{\mathbf{z},\mathbf{z}}(n,m)$, respectively. $N = 2M+1$ is the signal length. Note that the STFD of a signal vector is a matrix whose diagonal entries are the classical auto-TFDs of the vector components and the off-diagonal entries are the cross-TFDs.

A more general definition of the STFD can be given as

$$\mathbf{D}_{\mathbf{z},\mathbf{z}}(n,k) = \sum_{m=-M}^{M} \sum_{p=-M}^{M} \mathbf{G}(n-p,m) \odot \mathbf{z}(p+m)\mathbf{z}^*(p-m)e^{-j4\pi\frac{mk}{N}}, \tag{8.4.6}$$

where $\odot$ denotes the Hadamard product, and $[\mathbf{G}(n,m)]_{ij} = G_{ij}(n,m)$ is the time-lag kernel associated with the pair of the sensor signals $z_i(n)$ and $z_j(n)$.

The following MATLAB code calculates half of the STFD of a vector signal $\mathbf{z}(n)$ (the other half is obtained by Hermitian symmetry):

```
%                              MATLAB code 1
% Initialization
 D = zeros(fft_length,sample_size,sensor_number,sensor_number);
% STFD  computation
 for   k       = 1 :   K; % K: sensor number
   for l       = k :   K;
    D(:,:,k,l) = Cross-TFD(z(k,:),z(l,:),window_length,fft_length);
   end
 end
```

where the function

```
Cross-TFD(z(k,:),z(l,:),window_length,fft_length)
```

calculates the cross-TFD of the signals $z_k(n)$ and $z_l(n)$. [For the implementation of such functions, see Chapter 6.]

8.4.1.1 Structure under Linear Model

Consider the following linear model of the multi-sensor signal $\mathbf{z}(n)$,

$$\mathbf{z}(n) = \mathbf{A}\mathbf{s}(n), \tag{8.4.7}$$

where $\mathbf{A}$ is a $K \times L$ matrix ($K \geq L$) and $\mathbf{s}(n)$ is a $L \times 1$ vector which is referred to as the source signal vector.

Under this linear model the STFDs take the following structure:

$$\mathbf{D}_{\mathbf{zz}}(n,k) = \mathbf{A}\mathbf{D}_{\mathbf{ss}}(n,k)\mathbf{A}^H \tag{8.4.8}$$

where $\mathbf{D}_{\mathbf{ss}}(n,k)$ is the source STFD of vector $\mathbf{s}(n)$ whose entries are the auto- and cross-TFDs of the source signals.

The auto STFD denoted by $\mathbf{D}^a_{\mathbf{zz}}(n,k)$ is the STFD $\mathbf{D}_{\mathbf{zz}}(n,k)$ evaluated at auto-term (t,f) points only. Correspondingly, the cross STFD $\mathbf{D}^c_{\mathbf{zz}}(n,k)$ is the STFD $\mathbf{D}_{\mathbf{zz}}(n,k)$ evaluated at cross-term (t,f) points.

Note that the diagonal (off-diagonal) elements of $\mathbf{D}_{\mathbf{ss}}(n,k)$ are auto-terms (cross-terms). Thus, the auto (cross) STFD $\mathbf{D}^a_{\mathbf{ss}}(n,k)$ ($\mathbf{D}^c_{\mathbf{ss}}(n,k)$) is diagonal (off-diagonal) for each (t,f) point that corresponds to a source auto-term (cross-term), provided the window effect is neglected.

8.4.2 STFDs-Based Source Separation

The multi-sensor signal $\mathbf{z}(n)$ is assumed to be non-stationary and to satisfy the linear model (8.4.7).

The problem of source separation consists of identifying the matrix $\mathbf{A}$ and/or recovering the source signals $\mathbf{s}(n)$ up to a fixed permutation and some complex factors (see Article 8.1 for more details).

Denote by $\mathbf{W}$ the $L \times K$ whitening matrix, such that

$$(\mathbf{WA})(\mathbf{WA})^H = \mathbf{U}\mathbf{U}^H = \mathbf{I}. \tag{8.4.9}$$

Pre- and post-multiplying the STFD $\mathbf{D}_{\mathbf{zz}}(n,k)$ by $\mathbf{W}$ leads to the whitened STFD, defined as

$$\underline{\mathbf{D}}_{\mathbf{zz}}(n,k) = \mathbf{W}\mathbf{D}_{\mathbf{zz}}(n,k)\mathbf{W}^H = \mathbf{U}\mathbf{D}_{\mathbf{ss}}(n,k)\mathbf{U}^H \tag{8.4.10}$$

where the second equality stems from the definition of $\mathbf{W}$ and Eq. (8.4.8).

By selecting auto-term (t,f) points, the whitened auto STFD will have the following structure:

$$\underline{\mathbf{D}}^a_{\mathbf{zz}}(n,k) = \mathbf{U}\mathbf{D}^a_{\mathbf{ss}}(n,k)\mathbf{U}^H \tag{8.4.11}$$

where $\mathbf{D}^a_{\mathbf{ss}}(n,k)$ is diagonal. The missing unitary matrix $\mathbf{U}$ is retrieved (up to permutation and phase shifts) by Joint Diagonalization (JD) of a combined set $\{\underline{\mathbf{D}}^a_{\mathbf{zz}}(n_i,k_i)|i = 1,\cdots,P\}$ of P auto STFDs. The incorporation of several auto-term (t,f) points in the JD reduces the likelihood of having degenerate eigenvalues and increases robustness to a possible additive noise.

The selection of cross-term (t, f) points leads the following whitened cross STFD:

$$\underline{\mathbf{D}}^{c}_{\mathbf{zz}}(n,k) = \mathbf{U}\mathbf{D}^{c}_{\mathbf{ss}}(n,k)\mathbf{U}^{H} \tag{8.4.12}$$

where $\mathbf{D}^{c}_{\mathbf{ss}}(n,k)$ is off-diagonal. The unitary matrix $\mathbf{U}$ is found up to permutation and phase shifts by Joint Off-Diagonalization (JOD) of a combined set $\{\underline{\mathbf{D}}^{c}_{\mathbf{zz}}(n_i,k_i)|i=1,\cdots,Q\}$ of Q auto STFDs.

The unitary matrix $\mathbf{U}$ can also be find up to permutation and phase shifts by a combined JD/JOD of the two sets $\{\underline{\mathbf{D}}^{a}_{\mathbf{zz}}(n_i,k_i)|i=1,\cdots,P\}$ and $\{\underline{\mathbf{D}}^{c}_{\mathbf{zz}}(n_i,k_i)|i=1,\cdots,Q\}$. Once the unitary matrix $\mathbf{U}$ is obtained from either the JD or the JOD or the combined JD/JOD, an estimate of the mixing matrix $\mathbf{A}$ can be computed by the product $\mathbf{W}^{\#}\mathbf{U}$, where $\#$ denotes the pseudo-inverse operator. An estimate of the source signals $\mathbf{s}(n)$ is then obtained by the product $\mathbf{A}^{\#}\mathbf{z}(n)$.

8.4.3 Implementation of the Whitening

The implementation of the whitening goes as follows:

- Estimate the sample covariance $\mathbf{R}$ from T data samples.
- Denote by $\lambda_1,\ldots,\lambda_L$ the L largest eigenvalues, and by $\mathbf{h}_1,\ldots,\mathbf{h}_L$ the corresponding eigenvectors of $\mathbf{R}$.
- Under the white noise assumption, an estimate of the noise variance σ^2 is the average of the $K-L$ smallest eigenvalues of $\mathbf{R}$.
- The whitening matrix is computed by

$$\mathbf{W} = [(\lambda_1-\sigma^2)^{-\frac{1}{2}}\mathbf{h}_1,\ldots,(\lambda_L-\sigma^2)^{-\frac{1}{2}}\mathbf{h}_L]^T \tag{8.4.13}$$

- The whitened signals are then computed by $\underline{\mathbf{z}}(n) = \mathbf{W}\mathbf{z}(n)$.

Note that in the last step, the dimension of the problem is reduced from K to L. The above procedure is valid in the case of more sensors than sources, i.e. $K > L$. When $K = L$, no noise is assumed and the whitening matrix is computed as the matrix square root of the sample covariance matrix $\mathbf{R}$.

The following MATLAB code gives the implementation of the whitening :

```
%                        MATLAB code 2
% Computation of the covariance matrix
 R          = z*z'/sample_size;
 if L  <  K,                  % Assumes white noise
                              % K: sensor number, L: source number
% Compute the eigen decomposition of R
  [U,d]     = eig(R);
  [power,k]= sort(diag(real(d)));
  sigma     = mean(power(1:K-L));
  wl        = ones(L,1)./sqrt(power(K-L+1:K)-sigma);
% Computation of the whitening matrix
  W         = diag(wl)*U(1:K,k(K-L+1:K))';
```

```
 else                          % Assumes no noise
  W        = inv(sqrtm(R));
 end;
% Whitening the signals
 z         = W*z;
```

8.4.4 Selection of Auto-Terms and Cross-Terms

The selection procedure of the auto-terms and cross-terms exploits the off-diagonal structure of the source cross STFD matrices. More precisely, for a source cross STFD, we have

$$\text{Trace}(\underline{\mathbf{D}}^{c}_{\mathbf{zz}}(n,k)) = \text{Trace}(\mathbf{U}\mathbf{D}^{c}_{\mathbf{ss}}(n,k)\mathbf{U}^{H}) = \text{Trace}(\mathbf{D}^{c}_{\mathbf{ss}}(n,k)) \approx 0. \qquad (8.4.14)$$

Based on this observation, the following testing procedure applies:

if $\text{Trace}(\underline{\mathbf{D}}_{\mathbf{zz}}(n,k))/\text{norm}(\underline{\mathbf{D}}_{\mathbf{zz}}(n,k)) < \epsilon$ Then decide that (n,k) is a cross-term
if $\text{Trace}(\underline{\mathbf{D}}_{\mathbf{zz}}(n,k))/\text{norm}(\underline{\mathbf{D}}_{\mathbf{zz}}(n,k)) > \epsilon$ Then decide that (n,k) is an auto-term

where ϵ is a 'small' positive real scalar.

The following MATLAB code implements the selection procedure:

```
%                          MATLAB code 3
% Initialization
M   = [];  % will contain the auto STFDs
N   = [];  % will contain the cross STFDs
mp  = 2*L; % number of selected frequencies at a given time sample
           % where L is the source number
esp = 0.1; % threshold
% To select only (t-f) points with significant energy
[Dsort,Ii] = sort(D);     % D is the STFD computed at MATLAB code 1
for tp=1:skip:sample_size % skip is some chosen step
  for k = 1:mp
   fp   = Ii(fft_length-k,tp);
% Selection of the STFDs associated to the (t-f) point tp,fp
   Ds = squeeze(D(fp,tp,:,:));
% Form the other half of Ds by Hermitian symmetry
   Ds = Ds+ (Ds-diag(diag(Ds)))';
   if (trace(Ds)/norm(Ds))> esp,  % Selection criterion
    M = [M Ds];                   % Selection of auto STFDs
   else
    N = [N Ds];                   % Selection of cross STFDs
   end;
 end;
end;
% Note that the matrices M and N are the concatenation matrices of
% the auto-STFD matrices and cross-STFD matrices, respectively.
```

8.4.5 Implementation of JD and JOD

The joint diagonalization (JD) [10] of a set $\{\mathbf{M}_k | k = 1, \cdots, P\}$ of P matrices of dimensions $L \times L$ is defined as the maximization of the JD criterion

$$C_{\text{JD}}(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{k=1}^{P} \sum_{i=1}^{L} |\mathbf{v}_i^* \mathbf{M}_k \mathbf{v}_i|^2 \tag{8.4.15}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_L]$.

The joint Off-diagonalization (JOD) [8] of a set $\{\mathbf{N}_k | k = 1, \cdots, Q\}$ of Q matrices of dimensions $L \times L$ is defined as the maximization of the JOD criterion

$$C_{\text{JOD}}(\mathbf{V}) \stackrel{\text{def}}{=} -\sum_{k=1}^{P} \sum_{i=1}^{L} |\mathbf{v}_i^* \mathbf{N}_k \mathbf{v}_i|^2 \tag{8.4.16}$$

over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_L]$.

The combined joint diagonalization and joint off-diagonalization [8] of two sets $\{\mathbf{M}_k | k = 1, \cdots, P\}$ and $\{\mathbf{N}_k | k = 1, \cdots, Q\}$ of $L \times L$ matrices is defined as the maximization, over the set of unitary matrices $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_L]$, of the JD/JOD criterion

$$C_{\text{JD/JOD}}(\mathbf{V}) \stackrel{\text{def}}{=} \sum_{i=1}^{L} \left(\sum_{k=1}^{P} |\mathbf{v}_i^* \mathbf{M}_k \mathbf{v}_i|^2 - \sum_{k=1}^{Q} |\mathbf{v}_i^* \mathbf{N}_k \mathbf{v}_i|^2 \right). \tag{8.4.17}$$

Maximization of Eq. (8.4.17) is performed by successive Givens rotations as follows: A unitary matrix $\mathbf{V} = [v_{ij}]$ is sought such that the criterion in Eq. (8.4.17) is maximized w.r.t. the matrices $\mathbf{M}_k = [m_{ij}^k]$, $k = 1, \cdots, P$, and $\mathbf{N}_k = [n_{ij}^k]$, $k = 1, \cdots, Q$. At each Givens rotation step, a pivot pair r, s is chosen. With δ_{ij} denoting the Kronecker delta operator, v_{ij} is set to δ_{ij}, except for the elements $v_{rr} = v_{ss} = \cos(\theta)$, $v_{rs} = e^{j\phi} \sin(\theta)$ and $v_{sr} = -e^{-j\phi} \sin(\theta)$. One can show [10] that optimization of Eq. (8.4.17) is equivalent to the maximization of

$$\mathcal{Q} = \mathbf{v}^T \Re e(\mathbf{G}) \mathbf{v} \tag{8.4.18}$$

where

$$\mathbf{G} = \textstyle\sum_{k=1}^{P} \mathbf{g}_{1,k} \mathbf{g}_{1,k}^T - \sum_{k=1}^{Q} \mathbf{g}_{2,k} \mathbf{g}_{2,k}^T \tag{8.4.19}$$

$$\mathbf{v}^T = [\cos 2\theta, -\sin 2\theta \cos \phi, -\sin 2\theta \sin \phi] \tag{8.4.20}$$

$$\mathbf{g}_{1,k}^T = [m_{rr}^k - m_{ss}^k, m_{rs}^k + m_{sr}^k, j(m_{sr}^k - m_{rs}^k)] \tag{8.4.21}$$

$$\mathbf{g}_{2,k}^T = [n_{rr}^k - n_{ss}^k, n_{rs}^k + n_{sr}^k, j(n_{sr}^k - n_{rs}^k)]. \tag{8.4.22}$$

The next step is to recognize that the particular parameterization of $\mathbf{v}$ is equivalent to the condition $\mathbf{v}^T \mathbf{v} = 1$. Maximization of a quadratic form under the unit norm constraint of its argument is obtained by taking $\mathbf{v}$ to be the eigenvector of $\Re e(\mathbf{G})$ associated with the largest eigenvalue.

The reader may check that if we set $P = 1$ and $Q = 0$,[1] the above boils down to the standard Jacobi diagonalization procedure [11]. Also note that the main computational cost in this kind of technique is the update under Givens rotations of the various matrices involved in the diagonalization. This makes the computational cost of the JD/JOD procedure similar to $P+Q$ times the diagonalization of a single matrix.

The following MATLAB code gives the implementation of the JD/JOD:

```
%                    MATLAB code 4
% Initialization
 L         = min(size(M)); % M: auto STFDs given at MATLAB code 3
 nm1       = max(size(M));
 nm2       = max(size(N)); % N: cross STFDs given at MATLAB code 3
 V         = eye(L);       % L is the source number
 threshold = 1/sqrt(sample_size)/100;
 more      = 1;
 while more, more=0;
 for p=1:L-1,
  for q=p+1:L,
% Givens rotations
  g1=[M(p,p:L:nm1)-M(q,q:L:nm1);M(p,q:L:nm1)+M(q,p:L:nm1);
      i*(M(q,p:L:nm1)-M(p,q:L:nm1))];
  g2=[M(p,p:L:nm2)-M(q,q:L:nm2);M(p,q:L:nm2)+M(q,p:L:nm2);
      i*(M(q,p:L:nm2)-M(p,q:L:nm2)) ];
  [vcp,d] = eig(real(g1*g1'-g2*g2')); [la,Ki]=sort(diag(d));
   angles=vcp(:,Ki(1));angles=sign(angles(1))*angles;
   c=sqrt(0.5+angles(1)/2);
   sr=0.5*(angles(2)-j*angles(3))/c; sc=conj(sr);
   yes = abs(sr) > threshold;
   more=more | yes ;
   if yes,                          % Update of the M, N and V matrices
    colp1=M(:,p:L:nm1);colq1=M(:,q:L:nm1);
    M(:,p:L:nm1)=c*colp1+sr*colq1;M(:,q:L:nm1)=c*colq1-sc*colp1;
    rowp1=M(p,:);rowq1=M(q,:);
    M(p,:)=c*rowp1+sc*rowq1;M(q,:)=c*rowq1-sr*rowp1;
    colp2=N(:,p:L:nm2);colq2=N(:,q:L:nm2);
    N(:,p:L:nm2)=c*colp2+sr*colq2;N(:,q:L:nm2)=c*colq2-sc*colp2;
    rowp2=N(p,:);rowq2=N(q,:);
    N(p,:)=c*rowp2+sc*rowq2;N(q,:)=c*rowq2-sr*rowp2;
    temp=V(:,p);
    V(:,p)=c*V(:,p)+sr*V(:,q);V(:,q)=c*V(:,q)-sc*temp;
   end;     % if
  end;      % q loop
 end;       % p loop
end;        % while
```

8.4.6 Summary and Conclusions

The detailed implementation of non-stationary source separation algorithms based on Spatial TFDs has been presented together with the associated MATLAB codes.

[1] That is, only matrix $\mathbf{M}_1$ is incorporated in the criterion of Eq. (8.4.17).

The implementation details included whitening, selection of auto-term and cross-term regions and the combined Joint Diagonalization and Joint Off-Diagonalization.

References

[1] A. Belouchrani and M. G. Amin, "New approach for blind source separation using time frequency distributions," in *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures and Implementations VI*, vol. 2846, pp. 193–203, Soc. of Photo-optical Instrumentation Engineers, Denver, CO, 4–9 August 1996.

[2] A. Belouchrani and M. G. Amin, "Blind source separation using time-frequency distributions: Algorithm and asymptotic performance," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. 5, pp. 3469–3472, Munich, 21–24 April 1997.

[3] A. Belouchrani and M. G. Amin, "Blind source separation based on time-frequency signal representation," *IEEE Trans. Signal Processing*, vol. 46, pp. 2888–2898, November 1998.

[4] A. Belouchrani and M. G. Amin, "Time-frequency MUSIC: A new array signal processing method based on time-frequency signal representation," *IEEE Signal Processing Letters*, vol. 6, pp. 109–110, May 1999.

[5] A. S. Kayhan and M. G. Amin, "Spatial evolutionary spectrum for DOA estimation and blind signal separation," *IEEE Trans. Signal Processing*, vol. 48, pp. 791–798, March 2000.

[6] A. R. Leyman, Z. M. Kamran, and K. Abed-Meraim, "Higher-order time frequency-based blind source separation technique," *IEEE Signal Processing Letters*, vol. 7, pp. 193–196, July 2000.

[7] Y. Zhang, W. Mu, and M. G. Amin, "Subspace analysis of spatial time-frequency distribution matrices," *IEEE Trans. Signal Processing*, vol. 49, pp. 747–759, April 2001.

[8] A. Belouchrani, K. Abed-Meraim, M. G. Amin, and A. Zoubir, "Joint anti-diagonalization for blind source separation," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, Salt Lake City, UT, 7–11 May 2001.

[9] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, and B. Boashash, "Separating more sources than sensors using time-frequency distributions," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, vol. 2, pp. 583–586, Kuala Lumpur, 13–16 August 2001.

[10] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "Blind source separation using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434–444, February 1997.

[11] G. H. Golub and C. F. Van Loan, *Matrix computations*. Baltimore, MD: Johns Hopkins University Press, 1989.

8.5 UNDERDETERMINED BLIND SOURCE SEPARATION FOR FM-LIKE SIGNALS [0]

In many real-life engineering applications, including sonar and telecommunications, the signal under consideration may consist of a mixture of a number of signals. The user may not be interested in the whole mixture signal (also called a multi-component signal), but rather in a particular component of it. For instance, in telecommunications the received signal may be a mixture of several source signals (multiple access interference) but the user may wish to recover only one or some individual source signals. The various component signals may be non-stationary (i.e., their spectral contents may vary with time); and since time-frequency (TF) analysis has proved to be a powerful tool in the analysis of such signals, we use this technique to extract the desired source signal. In particular, we use the time-frequency distribution (TFD) of the mixture signal in order to separate and recover any signal component (see Article 8.1 for a detailed justification).

In this article, we focus on the important problem of underdetermined blind source separation (UBSS) where the number of sources is larger than the number of sensors. The sources under consideration are non-stationary frequency-modulated (FM) signals encountered in radar and certain wireless communication systems. For instance, linear FM chirp signaling is considered for frequency-hopped CDMA (Code Division Multiple Access) systems [1–3], in which context UBSS coincides with the blind multiuser detection problem.

We present here two classes of methods where UBSS is achieved using source spatial signatures in conjunction with their TF signatures. The two methods are based on a vector clustering technique and a mono-component extraction technique, respectively.

8.5.1 Data Model and Assumptions

Assume that an n-dimensional vector $\mathbf{s}(t) = [s_1(t), \ldots, s_n(t)]^T \in \mathbb{C}^{(n\times 1)}$ corresponds to n non-stationary complex source signals $s_i(t)$, $i = 1, \ldots, n$. The source signals are transmitted through a medium so that an array of m sensors picks up a set of mixed signals represented by an m-dimensional vector $\mathbf{x}(t) = [x_1(t), \ldots, x_m(t)]^T \in \mathbb{C}^{(m\times 1)}$. For an instantaneous linear mixture medium, the observed signals can, then, be modeled as:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \boldsymbol{\eta}(t), \tag{8.5.1}$$

where $\mathbf{A} \in \mathbb{C}^{(m\times n)}$ is the mixing matrix and $\boldsymbol{\eta}(t) = [\eta_1(t), \eta_2(t), \ldots, \eta_m(t)]^T \in \mathbb{C}^{(m\times 1)}$ is the observation noise vector. For the UBSS problem, i.e. for $n > m$, the mixing matrix $\mathbf{A}$ is not (left) invertible. However, the column vectors of matrix

[0]Authors: **Karim Abed-Meraim** and **Linh-Trung Nguyen**, Sig. & Image Proc. Dept., Telecom Paris, France (abed@tsi.enst.fr, trung@tsi.enst.fr), and **Adel Belouchrani**, EE Dept., Ecole Nationale Polytechnique, Algiers, Algeria (belouchrani@hotmail.com). Reviewers: B. Boashash, V. Chandran and A. Beghdadi.

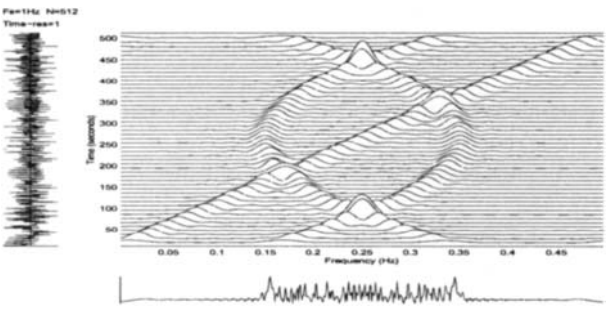


Fig. 8.5.1: A time-frequency distribution of a multicomponent signal.

$\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n]$ are assumed to be *pairwise linearly independent*, i.e., for any $i, j \in 1, 2, \ldots, n$ and $i \neq j$, $\mathbf{a}_i$ and $\mathbf{a}_j$ are linearly independent.

Here, the sources are assumed to be multicomponent FM signals. By a multicomponent signal, we mean a signal whose TF representation presents multiple ridges in the TF plane. Analytically, the kth source may be defined as,

$$s_k(t) = \sum_{l=1}^{M_k} s_{k,l}(t) \tag{8.5.2}$$

where each component $s_{k,l}(t)$, of the form

$$s_{k,l}(t) = a_{k,l}(t)\, e^{j\phi_{k,l}(t)}, \tag{8.5.3}$$

is assumed to have only one ridge in the TF plane. An example of a multicomponent signal, consisting of three components, is displayed in Fig. 8.5.1.

8.5.2 Separation using Vector Clustering

In this approach, the sources are assumed to have different structures and localization properties in the TF domain. More precisely, we assume the sources to be orthogonal in the TF domain (Fig. 8.5.2) in the sense that their TF supports[1] are disjoint.

The above assumption can be applied to any TFD. It is clear that the TF orthogonality is too restrictive and will almost never be satisfied exactly in practice. Fortunately, only approximate orthogonality, referred to as *quasi-orthogonality*, is needed in practice to achieve source separation [4]. Note, that two FM signals with different FM laws satisfy the quasi-orthogonality assumption.

Under the above assumption, one can notice that two auto-term spatial time frequency distribution (STFD) matrices (see Article 8.1 for a thorough definition)

[1] A TF support of a given signal $s(t)$ is defined by $\{(t, f) | D_{ss}(t, f) \neq 0\}$ where $D_{ss}(t, f)$ represents the TFD of $s(t)$.

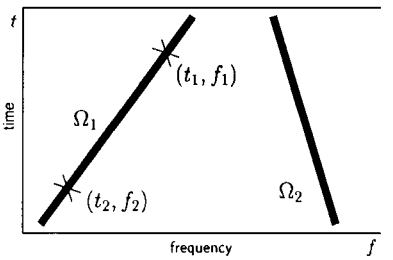


Fig. 8.5.2: TF orthogonal sources; the TF supports of two sources are disjoint.

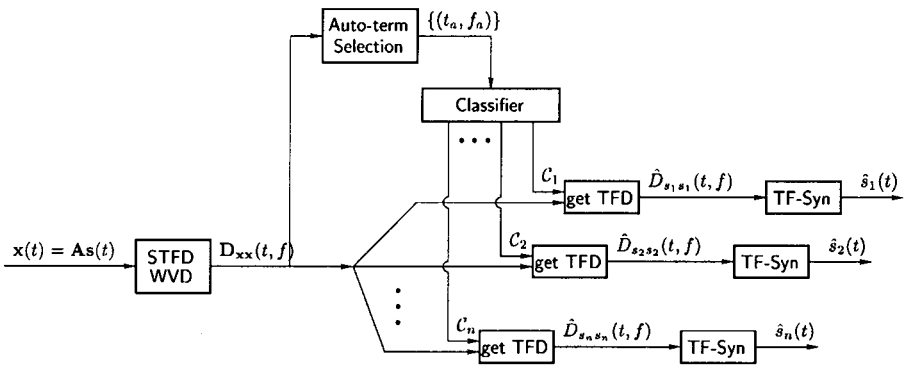


Fig. 8.5.3: Diagram of the UBSS algorithm.

of the observation $\mathbf{D_{xx}}(t_1, f_1)$ and $\mathbf{D_{xx}}(t_2, f_2)$ corresponding to the same source $s_i(t)$ are such that:

$$\begin{aligned} \mathbf{D_{xx}}(t_1, f_1) &= D_{s_i s_i}(t_1, f_1)\mathbf{a}_i\mathbf{a}_i^H, \\ \mathbf{D_{xx}}(t_2, f_2) &= D_{s_i s_i}(t_2, f_2)\mathbf{a}_i\mathbf{a}_i^H. \end{aligned} \tag{8.5.4}$$

Eq. (8.5.4) indicates that $\mathbf{D_{xx}}(t_1, f_1)$ and $\mathbf{D_{xx}}(t_2, f_2)$ have the same principal eigenvector $\mathbf{a}_i$. The idea of the proposed algorithm is to group together the auto-term points associated with the same principal eigenvector representing a particular source signal. The TFD of this source $(D_{s_i s_i}(t, f))$ is obtained as the principal eigenvalues of the STFD matrices at the auto-term points. The proposed algorithm is shown diagrammatically in Fig. 8.5.3. It consists of the following four steps:

(1) STFD computation and noise thresholding: Given a finite set of observation vectors, the STFD matrices of the observation $\mathbf{D_{xx}}(t, f)$ can be estimated using time-lag domain discrete implementation as shown in Articles 6.1 and 8.4. These STFD matrices are next processed to extract the source signals. In order to reduce the computational complexity by processing only "significant" STFD matrices, a noise thresholding is applied to the signal TFD. More precisely, a threshold ϵ_1 (typically, $\epsilon_1 = 0.05$ of the point with maximum energy) is used to keep only the

points $\{(t_s, f_s)\}$ with sufficient energy:

$$\text{Keep } (t_s, f_s) \text{ if } \|\mathbf{D_{xx}}(t_s, f_s)\| > \epsilon_1. \tag{8.5.5}$$

(2) Auto-term selection: The second algorithm step consists of separating the auto-term points from cross-term points using an appropriate testing criterion. For that, we exploit the sources' TF orthogonality. Under this assumption, each auto-term STFD matrix is of rank one, or at least has one "large" eigenvalue compared to its other eigenvalues. Therefore, one can use rank selection criteria, such as MDL (minimum description length) or AIC (Akaike information criterion), to select auto-term points as those corresponding to STFD matrices of selected rank equal to one. For simplicity, we use the following criterion:

$$\text{if } \left| \frac{\lambda_{\max}\{\mathbf{D_{xx}}(t,f)\}}{\text{norm}\{\mathbf{D_{xx}}(t,f)\}} - 1 \right| > \epsilon_2 \longrightarrow \text{decide that } (t,f) \text{ is a cross-term point}$$

where ϵ_2 is a small positive scalar (typically, $\epsilon_2 = 10^{-4}$), and $\lambda_{\max}\{\mathbf{D_{xx}}(t,f)\}$ represents the largest eigenvalue of $\mathbf{D_{xx}}(t,f)$.

(3) Clustering and source TFD estimation: Once the auto-term points have been selected, a clustering step based on the spatial signatures of the sources is performed. This clustering is based on the observation that two STFD matrices corresponding to the same source signal have the same principal eigenvectors. Moreover, the corresponding principal eigenvalues are given by the desired source TFD. This leads to the core step of our source separation method that uses vector clustering. This is implemented by the following four operations:

(a) For each auto-term point, (t_a, f_a), compute the main eigenvector, $\mathbf{a}(t_a, f_a)$, and its corresponding eigenvalue, $\lambda(t_a, f_a)$, of $\mathbf{D_{xx}}(t_a, f_a)$.

(b) As the vectors $\{\mathbf{a}(t_a, f_a)\}$ are estimated up to a random phase $e^{j\phi}$, $\phi \in [0, 2\pi)$, we force them to have, without loss of generality, their first entries real and positive.

(c) These vectors are then clustered into different classes $\{\mathcal{C}_i\}$. Mathematically, $\mathbf{a}(t_i, f_i)$ and $\mathbf{a}(t_j, f_j)$ belong to the same class if:

$$d(\mathbf{a}(t_i, f_i), \mathbf{a}(t_j, f_j)) < \epsilon_3 \tag{8.5.6}$$

where ϵ_3 is a properly chosen positive scalar and d is a distance measure (different strategies for choosing the threshold ϵ_3 and the distance d or even the clustering method can be found in [5]). As an example, in the simulated experiment, we used the angle between the two vectors as a distance measure:

$$d(\mathbf{a}_i, \mathbf{a}_j) = \arccos(\tilde{\mathbf{a}}_i^T \tilde{\mathbf{a}}_j) \tag{8.5.7}$$

where $\tilde{\mathbf{a}} = [\text{Re}(\mathbf{a})^T, \text{Im}(\mathbf{a})^T]^T / \|\mathbf{a}\|$.

(d) Set the number of sources equal to the number of "significant" classes and, for each source s_i (i.e. each class $\mathcal{C}_i$), estimate its TFD as:

$$\hat{D}_{s_i s_i}(t,f) = \begin{cases} \lambda(t_a, f_a), & \text{if } (t,f) = (t_a, f_a) \in \mathcal{C}_i \\ 0, & \text{otherwise.} \end{cases} \tag{8.5.8}$$

(4) Source signal synthesis: Use an adequate source synthesis procedure to estimate the source signals, $s_i(t)$ $(i = 1, \ldots, n)$, from their respective TFD estimates $\hat{D}_{s_i s_i}$. TF synthesis algorithms can be found in Chapter 11 and ref. [6].

An example in Fig. 8.5.4 illustrates the performance of the proposed algorithm.

8.5.3 Separation using Monocomponent Extraction

To achieve UBSS, we introduce here a four-step procedure consisting of:

1. Computation and spatial averaging of the observed signal TFDs;
2. Component extraction to separate all signal (mono) components;
3. Component clustering to group together components belonging to the same multicomponent source signal;
4. TF signal synthesis to recover the original source waveforms.

To have a "clean" TFD (i.e. one that can reveal the features of the signal clearly, without any "ghost" components), we use a newly developed high resolution quadratic TFD called the B-distribution (see Articles 2.7, 3.3 and 5.7). In addition, we use a spatial averaging that mitigates further the sources cross-terms by a factor depending on their spatial signatures angle (see Article 8.2). More precisely, we compute the averaged TFD (on which line detection is applied) as:

$$D(t,f) = \text{Trace}(\mathbf{D_{xx}}(t,f)) = \sum_{l=1}^{m} D_{x_l x_l}(t,f). \tag{8.5.9}$$

For component clustering, we use the observation that the STFD matrices at two auto-term points corresponding to the same source signal have the same principal eigenvector. Therefore, the proposed component clustering procedure consists of grouping together components associated with the same spatial direction representing a particular source signal. This spatial direction is estimated as the averaged value over all component points of the principal eigenvectors of the corresponding STFD matrices. More precisely, for each extracted component C, one estimates the corresponding spatial direction as:

$$\mathbf{a}_C = \frac{1}{N_{\mathcal{I}_C}} \sum_{i \in \mathcal{I}_C} \mathbf{a}(t_i, f_i) \tag{8.5.10}$$

where $\mathcal{I}_C$ denotes the set of points of component C, $N_{\mathcal{I}_C}$ denotes the number of points in $\mathcal{I}_C$ and $\mathbf{a}(t_i, f_i)$ is the estimated principal eigenvector of the ith component point STFD matrix $\mathbf{D_{xx}}(t_i, f_i)$. These vectors are then clustered into different classes using the clustering algorithm in Section 8.5.2.

For the component extraction, two different techniques, presented next, can be used. The first one is based on a "road extraction" procedure initially proposed for road detection in satellite image processing [7] while the second uses a "peak detection and tracking" procedure [8].

8.5.3.1 A "Road Network Tracking" Approach

The component-extraction method is divided into three main steps [9]: (i) preprocessing: because of the particularity of the TFD image, a preprocessing is needed before applying the component extraction procedure; (ii) line detection giving local binary detection of the potential linear structures (segments); and (iii) global optimization giving a set of labeled components. Due to space limitation, we only give a brief review of the principle of the method. Additional information and mathematical details can be found in [7] and references therein.

(1) Preprocessing: First the TF image is transformed to a real positive-valued image by forcing to zero all negative values[2] of the TFD and by using a gray scale in the range from 1 to 256. Also, line detectors are usually limited to a line width of 5 pixels. If the researched components do not respect this limit (which is usually the case for a TF image), an image subsampling by block-averaging is applied to reduce the pixel size. Despite the blurring effect, this filter presents the advantage of reducing the noise in the TF image. Moreover, as the TF image is unisotropic (i.e., it contains horizontal lines as can be observed in Fig. 8.5.1), this image downsampling [see Fig. 8.5.4(f)] removes this particular feature of the TF image.

(2) Line detection (Local optimization): A line detector is applied at each pixel of the image. We used a detector proposed in [7] for radar image processing. For a given direction, its response is based on the ratio of the means computed on both sides of the suspected line and the mean of the line itself. Height directions are studied and the best response is kept. The resulting image is then binarized using a simple thresholding. If statistics on the image are available (noise distribution, additive or multiplicative noise, etc.), a statistical study of the line detector performance can be made to choose the more adapted threshold (for instance the threshold corresponding to a fixed false alarm rate in homogeneous areas).

(3) Road detection (global optimization): This step is a global step introducing constraints on the shape of the linear features to extract connected components and to suppress the false alarms [7]. It works on segments extracted on the thresholded line response image by thinning and linearization. The previously detected segments are connected depending on proximity and alignment constraints (specially

[2]Negative values correspond mainly to undesired cross-terms or noise.

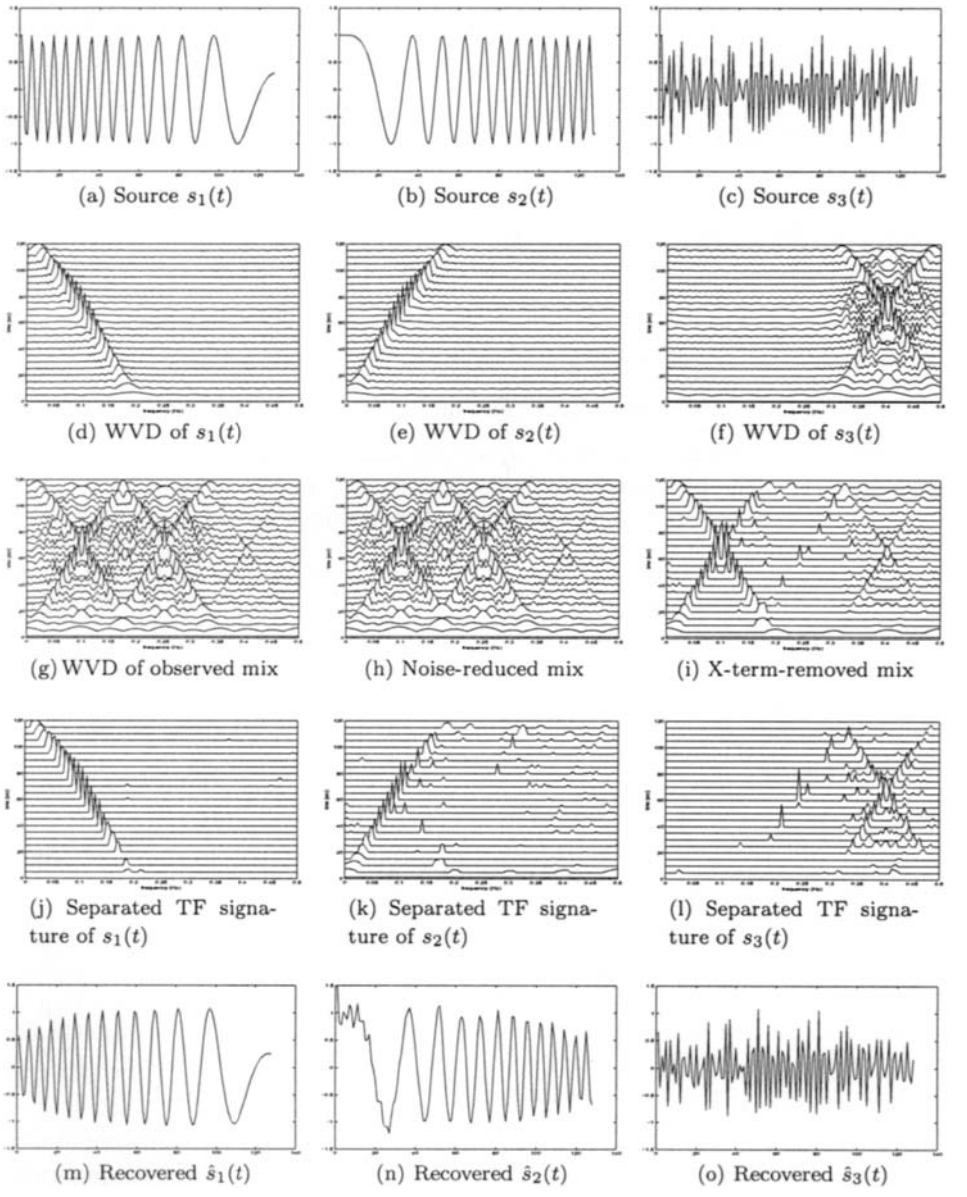


Fig. 8.5.4: Testing the "vector clustering" UBSS algorithm for $m = 2$ mixtures of two monocomponent and one multicomponent LFM signals denoted by $s_1(t)$, $s_2(t)$ and $s_3(t)$ (a–c). The recovered source signals (m–o) indicated the success of the UBSS algorithm. Source $s_3(t)$ was not falsely separated into two monocomponent sources.

on the line curvature) to form coherent components. Small isolated segments are suppressed. The algorithm depends on the following thresholds: the maximum gap between two segments to connect them, the allowed angular difference between the two segments, and the minimum size of a component. The result is a labeled image of components.

Fig. 8.5.5 shows an example illustrating the separation quality that can be obtained by the above technique.

8.5.3.2 A "Peak Detection and Tracking" Approach

We present here a component separation algorithm [8] relying on the additional assumption that all components of the signal exist at almost all time instants. Fig. 8.5.6, Table 8.5.1 and Table 8.5.2 illustrate this algorithm. Fig. 8.5.6 provides the algorithm flowchart, Table 8.5.1 describes the estimation technique of the number of components and Table 8.5.2 details the component separation technique.

(1) Noise thresholding: The first step of the algorithm consists in noise thresholding to remove the undesired "low" energy peaks in the TF domain.[3] We set to zero the TFD values smaller than a properly chosen threshold ϵ. Using MATLAB^{TM} notation, this operation can be written as:

$$D_{\text{th}}(t,f) = D(t,f).*(D(t,f) > \epsilon). \tag{8.5.11}$$

(2) Estimation of the number of components: In general, for a noiseless and cross-terms free TFD, the number of components at a given time instant t can be estimated as the number of peaks of the TFD slice $D(t,f)$. Here, we propose a simple technique to estimate the number of components in the case where all components exist simultaneously at almost all time instants. In this situation, we can efficiently evaluate the number of components (see Table 8.5.1) as the maximum argument of the histogram of the number of peaks computed for each time instant t in the range $[1,2,\ldots,t_{\max}]$ (where $t_{\max} \times f_{\max}$ is the dimension of the TFD matrix).

(3) Component separation procedure: The proposed algorithm assumes that (i) all components exist at all time instants in the TF plane and (ii) any component intersection is a crossing point. Under these two assumptions, we note that if, at a time instant t, two or more components are crossing, then the number of peaks (at this particular slice $D(t,f)$) is smaller than the total number of components d. The details of the proposed separation technique is outlined in Table 8.5.2.

To validate the proposed algorithm, we reconsider the same multicomponent signal analyzed in Fig. 8.5.1. This signal consists of a mixture of two quadratic frequency modulated (FM) components and a linear FM component. The mixture signal is added to a zero-mean white Gaussian noise, with SNR equal to 0 dB.

[3] This noise thresholding is justified by the fact that the noise energy is spread over the whole TF domain while the components' energies are well localized around their respective IFs leading to high energy peaks for the latter (assuming no cross-terms).

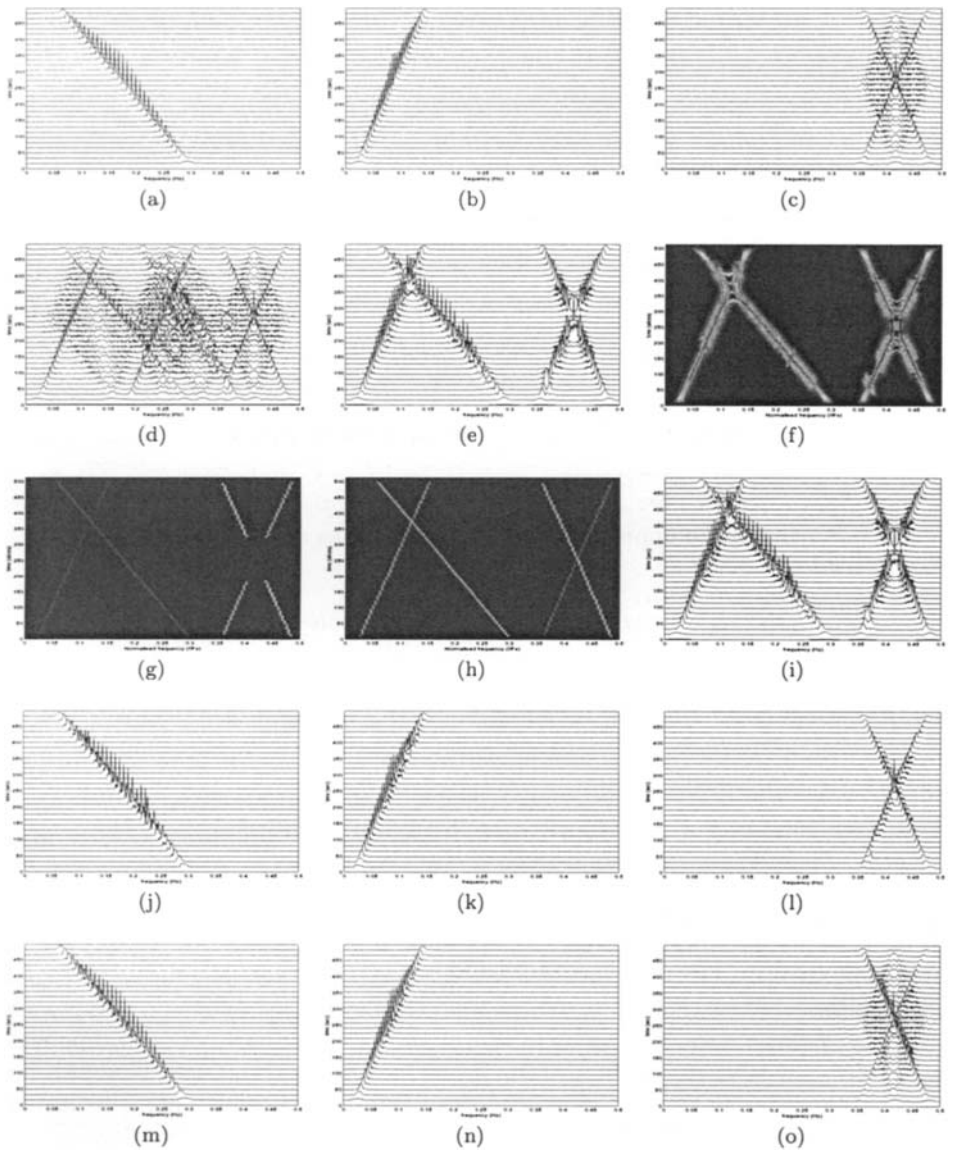


Fig. 8.5.5: (a–c) WVD of $s_1(t)$, $s_2(t)$, $s_3(t)$; (d,e) spatial-averaged TFD of the mixture outputs using WVD and MWVD; (f) STFD mixture converted to image; (g–h) extraction of source components using "road network tracking"; (i) auto-term points of known components; (j–l) TFD estimates of the sources; (m–o) TFD of estimated sources after TF synthesis: $m = 2$ sensors and $\mathrm{SNR} = 10\,\mathrm{dB}$.

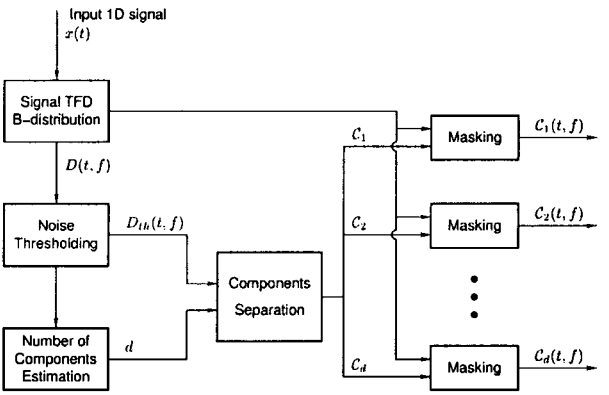


Fig. 8.5.6: Flowchart of the "component extraction" based algorithm.

Table 8.5.1: Estimation of the number of components.

1. For $t = 1, \ldots, t_{\max}$

   number_components(t) = number of peaks of $D(t,:)$

   end

2. Compute the histogram H of number_components.

3. Evaluate the number of components as $d = \arg\max H$

The B-distribution of the noisy signal as well as the components resulting from the separation algorithm are displayed in Fig. 8.5.7. Another relevant component separation algorithm can be found in [10].

8.5.4 Summary and Conclusions

This article presents TF-based methods for the underdetermined blind separation of FM-like signals. The first class of methods is based on a vector clustering approach while the second is based on a component extraction approach. Two component extraction techniques are introduced using a "road network tracking" algorithm (initially developed for road detection in satellite image processing) and a dedicated "peak detection and tracking" algorithm, respectively.

Simulation examples are provided to illustrate the performances of considered blind separation methods. Note that the UBSS method based on vector clustering is more general in the sense that it can be applied to separate a larger class of (not necessarily FM) signals as long as they satisfy the TF orthogonality property. In fact, it can be seen as a member of the "sparse decomposition" based UBSS methods [11] which essentially use the sparse representation of the source signals in a

Table 8.5.2: Component separation procedure for the "peak detection and tracking" algorithm.

1. Assign an index to each of the d components in an orderly manner.
2. For each time instant t (starting from $t = 1$), find the component frequencies as the peak positions of the TFD slice $D(t, f)$.
3. Assign a peak to a particular component based on the smallest distance to the peaks of the previous slice $D(t-1, f)$ (IFs continuous functions of time).
4. If at a time instant t a crossing point exists (i.e., number of peaks smaller than the number of components), identify the crossing components using the smallest distance criterion by comparing the distances of the actual peaks to those of the previous slice.
5. Permute the indices of the corresponding crossing components.

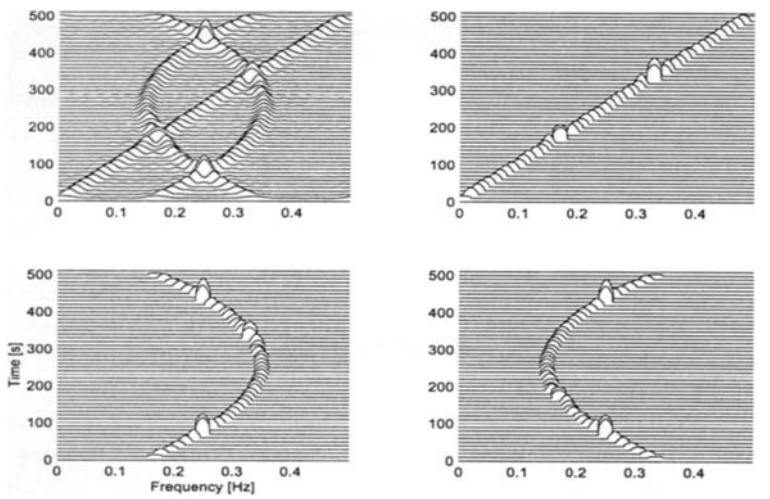


Fig. 8.5.7: The B-distribution of the original signal (top left) as well as the extracted components using the "peak detection and tracking" algorithm.

given transformed domain (in our case, it is the TF domain). The "vector clustering" based method is however less performant than the second class of methods using component extraction when considering only FM-type signals. Several component extraction techniques exist in the literature [8–10, 12] among which two have been presented in this article. The best promising technique (i.e. the one providing the best separation quality) is the "image processing" based one using a "road detection"

algorithm. In terms of computational cost, it is however much more expensive than the "peak detection and tracking" technique.

References

[1] S. El-Khamy, S. E. Shaaban, and E. A. Thabet, "Frequency-hopped multi-user chirp modulation (FH/M-CM) for multipath fading channels," in *Proc. Sixteenth National Radio Science Conference (NRSC'99)*, pp. C6/1–8, Ain Shams Univ., Cairo, 23–25 February 1999.

[2] C. Gupta and A. Papandreou-Suppappola, "Wireless CDMA communications using time-varying signals," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, vol. 1, pp. 242–245, Kuala Lumpur, 13–16 August 2001.

[3] A. Persson, T. Ottosson, and E. Strom, "Time-frequency localized CDMA for downlink multi-carrier systems," in *Proc. Seventh Internat. Symp. on Spread Spectrum Techniques and Applications (ISSSTA-02)*, vol. 1, pp. 118–122, Prague, 2–5 September 2002.

[4] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, and B. Boashash, "Separating more sources than sensors using time-frequency distributions," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, vol. 2, pp. 583–586, Kuala Lumpur, 13–16 August 2001.

[5] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Kluwer, 1991.

[6] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 34, pp. 442–451, June 1986.

[7] F. Tupin, H. Maître, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: Application to road network extraction," *IEEE Trans. Geoscience & Remote Sensing*, vol. 36, pp. 434–453, March 1998.

[8] K. Abed-Meraim and B. Barkat, "Blind source separation using the time-frequency distribution of the mixture signal," in *Proc. Second IEEE Internat. Symp. on Signal Processing and Information Technology (ISSPIT'02)*, vol. 2, pp. 663–666, Marrakech, Morocco, 18–21 December 2002.

[9] K. Abed-Meraim, N. Linh-Trung, V. Sucic, F. Tupin, and B. Boashash, "An image processing approach for underdetermined blind separation of nonstationary sources," in *Proc. Third Internat. Symp. on Image & Signal Processing and Analysis (ISPA-03)*, Rome, 18–20 September 2003.

[10] B. Barkat and K. Abed-Meraim, "A blind components separation procedure for FM signal analysis," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'02)*, vol. 2, pp. 1425–1428, Orlando, FL, 13–17 May 2002.

[11] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Processing*, vol. 81, pp. 2353–2362, November 2001.

[12] F. S. Cohen, S. Kadambe, and G. F. Boudreaux-Bartels, "Tracking of unknown nonstationary chirp signals using unsupervised clustering in the Wigner distribution space," *IEEE Trans. Signal Processing*, vol. 41, pp. 3085–3101, November 1993.

Part IV

Statistical Techniques

This Page Intentionally Left Blank

Chapter 9

Random Processes and Noise Analysis

Time-Frequency Distributions (TFDs) have been studied initially for deterministic signals. Noise is an ever permanent feature in signals and so there is a need to extend the formulation of TFDs to the case of random signals. This chapter describes time-frequency methodologies developed for dealing with randomness in non-stationary signals and systems. The topic is covered in five articles with appropriate cross-referencing.

Noise analysis for TFDs is presented with derivations of mean and variance of TFDs (Article 9.1). Both cases of additive and multiplicative noise, including white and colored noise, are treated. Time-varying random processes and systems can be represented with dispersive time-frequency characteristics using the Weyl symbol (9.2). This approach allows the adaptation of analysis tools to systems or processes that have specific time-frequency (TF) characteristics. TFDs such as the spectrogram and WVD can be designed that are robust to impulse noise (9.3). Time-varying power spectra can be defined based on generalizations of the Wigner-Ville spectrum and evolutionary spectrum. These are particularly suitable for underspread non-stationary processes (i.e. processes with small time-frequency correlations) (9.4). Time-varying random channels are also described using a time-frequency approach (9.5).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

9.1 ANALYSIS OF NOISE IN TIME-FREQUENCY DISTRIBUTIONS [0]

This article presents noise analysis for commonly used time-frequency distributions such as those presented in Chapters 2 and 3. The Wigner distribution, as a basic time-frequency representation, is studied first. The bias and variance in the case of complex white noise are derived. The analysis of noise is extended to other quadratic distributions, and to different types of additive and multiplicative noise, including: stationary white noise, nonstationary white noise, and colored stationary noise. Exact expressions for the mean value and the variance of quadratic distributions for each point in the time-frequency plane are given.

9.1.1 Wigner Distribution

The pseudo Wigner distribution (WD) of a discrete-time noisy signal $x(n) = s(n) + \epsilon(n)$ is defined by[1]:

$$W_x(n,f) = \sum\nolimits_m w(m)w(-m)x(n+m)x^*(n-m)e^{-j4\pi fm}. \tag{9.1.1}$$

where $w(m)$ is a real-valued lag window, such that $w(0) = 1$.

Consider first the case when $s(n)$ is deterministic and the noise $\epsilon(n)$ is a white, Gaussian, complex, stationary, zero-mean process, with independent real and imaginary parts having equal variances. Its autocorrelation function is $R_{\epsilon\epsilon}(m) = \sigma_\epsilon^2\delta(m)$. The WD mean for the noisy signal $x(n)$ is

$$\begin{aligned} E\{W_x(n,f)\} &= \sum\nolimits_m w(m)w(-m)s(n+m)s^*(n-m)e^{-j4\pi fm} \\ &\quad + \sum\nolimits_m w(m)w(-m)R_{\epsilon\epsilon}(2m)e^{-j4\pi fm} \\ &= 2\int_{-1/4}^{1/4} W_s(n,f-\alpha)F_w(2\alpha)d\alpha + \sigma_\epsilon^2, \end{aligned} \tag{9.1.2}$$

where $F_w(f) = \mathcal{F}_{m\to f}[w(m)w(-m)]$ is the Fourier transform (FT) of the product $w(m)w(-m)$, and $W_s(n,f)$ is the original WD of $s(n)$, without a lag window.

The lag window $w(m)$ causes the **WD bias**. The second term on the right-hand side in (9.1.2) is constant, so one can assume that it does not distort the WD. Expanding $W_s(n,f-\alpha)$ into a Taylor series, around f, we get

$$2\int_{-1/4}^{1/4} W_s(n,f-\alpha)F_w(2\alpha)d\alpha \cong W_s(n,f) + \frac{1}{8}\frac{\partial^2 W_s(n,f)}{\partial f^2}m_2 + ... \tag{9.1.3}$$

---

[0] Author: **LJubiša Stanković**, Elektrotehnicki fakultet, University of Montenegro, 81000 Podgorica, Montenegro (l.stankovic@ieee.org). Reviewers: M. G. Amin, and M. J. Bastiaans.

[1] Notation $\sum_m$, without limits, will be used for $\sum_{m=-\infty}^{\infty}$. The constant factor of 2 is omitted in the WD definition, and in other TFD definitions.

Thus, the bias can be approximated by

$$\text{bias}\,(n,f) \cong \frac{1}{8}\frac{\partial^2 W_s(n,f)}{\partial f^2} m_2 = \frac{1}{8} b(n,f) m_2,$$

where $m_2 = \int_{-1/2}^{1/2} f^2 F_w(f) df$. For the regions where the WD variations in the frequency direction are small, the bias is small, and vice versa.

The WD estimator **variance**, at a given point (n,f), is defined by:

$$\sigma^2_{WD}(n,f) = E\{W_x(n,f)W_x^*(n,f)\} - E\{W_x(n,f)\}E\{W_x^*(n,f)\}. \tag{9.1.4}$$

For signals $x(n) = s(n) + \epsilon(n)$ it results in

$$\begin{aligned}\sigma^2_{WD}(n,f) \;=\; & \sum_{m_1}\sum_{m_2} w(m_1)w(-m_1)w(m_2)w(-m_2)e^{-j4\pi f(m_1-m_2)} \\ & \times[s(n+m_1)s^*(n+m_2)R_{\epsilon\epsilon}(n-m_2,n-m_1) \\ & +s^*(n-m_1)s(n-m_2)R_{\epsilon\epsilon}(n+m_1,n+m_2) \\ & +s(n+m_1)s(n-m_2)R^*_{\epsilon\epsilon^*}(n-m_1,n+m_2) \\ & +s^*(n-m_1)s^*(n+m_2)R_{\epsilon\epsilon^*}(n+m_1,n-m_2) \\ & +R_{\epsilon\epsilon}(n+m_1,n+m_2)R_{\epsilon\epsilon}(n-m_2,n-m_1) \\ & +R_{\epsilon\epsilon^*}(n+m_1,n-m_2)R^*_{\epsilon\epsilon^*}(n-m_1,n+m_2)].\end{aligned} \tag{9.1.5}$$

The fourth-order moment of noise is reduced to the correlation functions by using the relation $E\{z_1z_2z_3z_4\} = E\{z_1z_2\}E\{z_3z_4\} + E\{z_1z_3\}E\{z_2z_4\} + E\{z_1z_4\}E\{z_2z_3\}$, which holds for Gaussian zero-mean random variables z_i, $i = 1,2,3,4$. For the considered complex noise $R_{\epsilon\epsilon}(n,m) = \sigma^2_\epsilon\delta(n-m)$ and $R_{\epsilon\epsilon^*}(n,m) = 0$. The variance of the WD estimator reduces to

$$\sigma^2_{WD}(n,f) = \sigma^2_\epsilon \sum_m w^2(m)w^2(-m)\left[2\,|s(n+m)|^2 + \sigma^2_\epsilon\right].$$

It is frequency independent. For constant modulus signals, $s(n) = a\exp[j\phi(n)]$, the variance is constant $\sigma^2_{WD}(n,f) = \sigma^2_\epsilon E_w(2a^2+\sigma^2_\epsilon)$, where $E_w = \sum_m[w(m)w(-m)]^2$ is the energy of $w(m)w(-m)$ window. A finite energy lag window is sufficient to make the variance of $W_x(n,f)$ finite.

The optimal lag window width can be obtained by minimizing the error $e^2 = \text{bias}^2(n,f) + \sigma^2_{WD}(n,f)$. For example, for constant modulus signals, and the Hanning window $w(m)w(-m)$ of the width N, when $E_w = 3N/8$ and $m_2 = 1/(2N^2)$, we get:

$$e^2 \cong \frac{1}{256N^4} b^2(n,f) + \frac{3N}{8}\sigma^2_\epsilon(2a^2+\sigma^2_\epsilon).$$

From $\partial e^2/\partial N = 0$ the approximation of optimal window width follows:

$$N_{opt}(n,f) \cong \sqrt[5]{\frac{b^2(n,f)}{24\sigma^2_\epsilon(2a^2+\sigma^2_\epsilon)}}.$$

An approach to the calculation of the estimate of $N_{opt}(n,f)$, without using the value of $b^2(n,f)$, is presented in [1], [Article 10.2]. Other statistical properties of the Wigner distribution are studied in [2].

9.1.2 Noise in Quadratic Time-Frequency Distributions

A discrete-time form of the Cohen class[2] of distributions of noise $\epsilon(n)$ is defined by:

$$\rho_\epsilon(n,f;G) = \sum_l \sum_m G(m,l)\epsilon(n+m+l)\epsilon^*(n+m-l)e^{-j4\pi fl}, \tag{9.1.6}$$

where $G(m,l)$ is the kernel in the time-lag domain.

Its **mean value**, for a general nonstationary noise, is

$$E\{\rho_\epsilon(n,f;G)\} = \sum_l \sum_m G(m,l)R_{\epsilon\epsilon}(n+m+l,n+m-l)e^{-j4\pi fl},$$

where $R_{\epsilon\epsilon}(m,n)$ is the noise autocorrelation function. For special cases of noise the values of $E\{\rho_\epsilon(n,f;G)\}$ follow.

(1) Stationary white noise, $R_{\epsilon\epsilon}(m,n) = \sigma_\epsilon^2\delta(m-n)$,

$$E\{\rho_\epsilon(n,f;G)\} = \sigma_\epsilon^2 g(0,0).$$

(2) Nonstationary white noise, $R_{\epsilon\epsilon}(m,n) = I(n)\delta(m-n)$, $I(n) \geq 0$,

$$E\{\rho_\epsilon(n,f;G)\} = \sum_m G(m,0)I(n+m).$$

(3) Stationary colored noise, $R_{\epsilon\epsilon}(m,n) = R_{\epsilon\epsilon}(m-n)$,

$$E\{\rho_\epsilon(n,f;G)\} = \int_{-1/2}^{1/2} \mathcal{G}(0,2(f-\alpha))S_{\epsilon\epsilon}(\alpha)d\alpha,$$

where $S_{\epsilon\epsilon}(f) = \mathcal{F}_{m\to f}[R_{\epsilon\epsilon}(m)]$ is the noise power spectrum density, and the kernel forms in time-lag, Doppler-lag, and Doppler-frequency domains are denoted by:

$$\sum_m G(m,l)e^{-j2\pi\nu m} = g(\nu,l) = \int_{-1/2}^{1/2} \mathcal{G}(\nu,f)e^{j2\pi fl}df. \tag{9.1.7}$$

The **variance** of $\rho_\epsilon(n,f;G)$, is defined by

$$\sigma_{\epsilon\epsilon}^2(n,f) = E\{\rho_\epsilon(n,f;G)\rho_\epsilon^*(n,f;G)\} - E\{\rho_\epsilon(n,f;G)\}E\{\rho_\epsilon^*(n,f;G)\}.$$

For Gaussian noise, as in (9.1.4)-(9.1.5), we get:

$$\begin{aligned}\sigma_{\epsilon\epsilon}^2(n,f) = &\sum_{l_1}\sum_{l_2}\sum_{m_1}\sum_{m_2} G(m_1,l_1)G^*(m_2,l_2) \\ &\times[R_{\epsilon\epsilon}(n+m_1+l_1,n+m_2+l_2)R_{\epsilon\epsilon}^*(n+m_1-l_1,n+m_2-l_2) \\ &+R_{\epsilon\epsilon^*}(n+m_1+l_1,n+m_2-l_2)R_{\epsilon^*\epsilon}(n+m_1-l_1,n+m_2+l_2)]e^{-j4\pi f(l_1-l_2)}.\end{aligned} \tag{9.1.8}$$

Form of $\sigma_{\epsilon\epsilon}^2(n,f)$ for the specific noises will be presented next.

[2] That is, the quadratic class; see p. 68n.

Complex stationary and nonstationary white noise

For nonstationary complex white noise, with independent real and imaginary part of equal variance, $R_{\epsilon\epsilon}(m,n) = I(n)\delta(m-n)$, $R_{\epsilon\epsilon^*}(n,m) = 0$, we get

$$\sigma^2_{\epsilon\epsilon}(n,f) = \sum_l \sum_m |G(m,l)|^2 I(n+m+l)I^*(n+m-l) = \rho_I(n,0;|G|^2). \quad (9.1.9)$$

For stationary white noise, $I(n) = \sigma^2_\epsilon$, the variance is proportional to the kernel energy,

$$\sigma^2_{\epsilon\epsilon}(n,f) = \sigma^4_\epsilon \sum_l \sum_m |G(m,l)|^2 . \quad (9.1.10)$$

Colored stationary noise

For complex colored stationary noise, the variance (9.1.8) can be written as

$$\sigma^2_{\epsilon\epsilon}(n,f) = \sum_{l_1} \sum_{m_1} G(m_1,l_1) \Big\{ \sum_{m_2} \sum_{l_2} G^*(m_2,l_2) \times [R_{\epsilon\epsilon}(m_1-m_2+l_1-l_2) R^*_{\epsilon\epsilon}(m_1-m_2-(l_1-l_2))] e^{-j4\pi f(l_1-l_2)} \Big\},$$

or

$$\sigma^2_{\epsilon\epsilon}(n,f) = \sum_l \sum_m G(m,l) \left\{ G(m,l) *_l *_m [R^*_{\epsilon\epsilon}(m+l) R_{\epsilon\epsilon}(m-l) e^{j4\pi fl}] \right\}^*,$$

where "$*_l *_m$" denotes a two-dimensional convolution in l, m. Consider the product of $G(m,l)$ and $Y^*(m,l) = \left\{ G(m,l) *_l *_m [R^*_{\epsilon\epsilon}(m+l) R_{\epsilon\epsilon}(m-l) e^{j4\pi fl}] \right\}^*$ in the last expression. Two-dimensional FTs of these terms are $\mathcal{G}(\nu,\xi)$ and $y(\nu,\xi) = \mathcal{G}(\nu,\xi) S_{\epsilon\epsilon}(f-(\xi-\nu)/2) S^*_{\epsilon\epsilon}(f-(\xi+\nu)/2)/2$. According to the Parseval's theorem we get:

$$\begin{aligned} \sigma^2_{\epsilon\epsilon}(n,f) &= \frac{1}{2} \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |\mathcal{G}(\nu,\xi)|^2 S^*_{\epsilon\epsilon}(f - \frac{\xi}{2} + \frac{\nu}{2}) S_{\epsilon\epsilon}(f - \frac{\xi}{2} - \frac{\nu}{2}) d\nu d\xi \\ &= \rho_{S_{\epsilon\epsilon}}(0,f;|\mathcal{G}|^2), \end{aligned} \quad (9.1.11)$$

for $|f-(\xi-\nu)/2| < 1/2$ and $|f-(\xi+\nu)/2| < 1/2$. The transforms in (9.1.11) are periodic in ν and ξ with period 1. It means that we should take into account all ν and ξ when $|f-[(\xi+k_1)-(\nu+k_2)]/2| < 1/2$ and $|f-[(\xi+k_1)+(\nu+k_2)]/2| < 1/2$, where k_1 and k_2 are integers.

Note that the FT of a colored stationary noise is a white nonstationary noise, with autocorrelation in the frequency domain

$$R_{\Xi\Xi}(f_1,f_2) = \sum_m \sum_n E\{\epsilon(m)\epsilon^*(n)\} e^{(-j2\pi f_1 m + j2\pi f_2 n)} = S_{\epsilon\epsilon}(f_2)\delta_p(f_1-f_2),$$

where $\delta_p(f)$ is a periodic delta function with period 1. Thus, (9.1.11) is just a form dual to (9.1.9).

Analytic noise

In the numerical implementation of quadratic distributions, an analytic part of the signal is commonly used, rather than the signal itself. The analytic part of noise can be written as $\epsilon_a(n) = \epsilon(n) + j\epsilon_h(n)$, where $\epsilon_h(n) = \mathcal{H}[\epsilon(n)]$ is the Hilbert transform of $\epsilon(n)$. Spectral power density of $\epsilon_a(n)$, within the basic period $|f| < 1/2$, for a white noise $\epsilon(n)$, is $S_{\epsilon_a\epsilon_a}(f) = 2\sigma_\epsilon^2 U(f)$, where $U(f)$ is the unit step function. The variance follows from (9.1.11) in the form

$$\sigma_{\epsilon\epsilon}^2(n,f) = 2\sigma_\epsilon^4 \int_{-1/2}^{1/2} \int_{-d(f,\xi)}^{d(f,\xi)} |\mathcal{G}(\nu,\xi)|^2 \, d\xi d\nu \quad \text{for} \quad |2f| \le \frac{1}{2}, \tag{9.1.12}$$

where the integration limits are defined by $d(f,\xi) = |\arcsin(\sin(\pi(2f-\xi)))|/\pi$ (for details see [3]).

The kernel $\mathcal{G}(\nu,\xi)$ is mainly concentrated at and around the (ν,ξ) origin and $\xi = 0$ axis. Having this in mind, as well as the fact that $|\mathcal{G}(\nu,\xi)|^2$ is always positive, we may easily conclude that the minimal value of $\sigma_{\epsilon\epsilon}^2(n,f)$ is for $f = 0$. The maximal value will be obtained for $|f| = 1/4$. It is very close to [3]:

$$\max\{\sigma_{\epsilon\epsilon}^2(n,f)\} \cong 2\sigma_\epsilon^4 \int_{-1/2}^{1/2} \int_{-1/2}^{1/2} |\mathcal{G}(\nu,\xi)|^2 \, d\xi d\nu = 2\sigma_\epsilon^4 \sum_l \sum_m |G(m,l)|^2 .$$

Real noise

Now consider a real stationary white Gaussian noise $\epsilon(n)$ with variance σ_ϵ^2. In this case, variance (9.1.8) contains all terms. It can be written as:

$$\sigma_{\epsilon\epsilon}^2(n,f) = \sigma_\epsilon^4 \sum_l \sum_m [|G(m,l)|^2 + G(m,l)G^*(m,-l)e^{-j8\pi fl}]. \tag{9.1.13}$$

For distributions whose kernel is symmetric with respect to l, $G(m,l) = G(m,-l)$ holds. The FT is therefore applied to the positive and even function $|G(m,l)|^2$. The transform's maximal value is reached at $f = 0$, and $|f| = 1/4$. Accordingly:

$$\max\left\{\sigma_{\epsilon\epsilon}^2(n,f)\right\} = 2\sigma_\epsilon^4 \sum_l \sum_m |G(m,l)|^2 . \tag{9.1.14}$$

The crucial parameter in all previous cases is the kernel energy $\sum_l \sum_m |G(m,l)|^2$. Its minimization is thoroughly studied in [4]. It has been concluded that, out of all the quadratic distributions satisfying the marginal and time-support conditions, the Born-Jordan distribution is optimal with respect to this parameter.

9.1.3 Noisy Signals

Analysis of deterministic signals $s(n)$ corrupted by noise, $x(n) = s(n) + \epsilon(n)$, is highly signal dependent. It can be easily shown [3,4], that the distribution variance $\sigma_\rho^2(n,f)$ consists of two components:

$$\sigma_\rho^2(n,f) = \sigma_{\epsilon\epsilon}^2(n,f) + \sigma_{s\epsilon}^2(n,f). \tag{9.1.15}$$

The first variance component, and the distribution mean value, have already been studied in detail.[3] For the analysis of the second, signal dependent, component $\sigma^2_{s\epsilon}(n,f)$ we will use the inner product form of the Cohen class of distributions:

$$\rho_x(n,f;\tilde{G}) = \sum_l \sum_m \tilde{G}(m,l)\left[x(n+m)e^{-j2\pi fm}\right]\left[x(n+l)e^{-j2\pi fl}\right]^*, \quad (9.1.16)$$

where $\tilde{G}(m,l) = G((m+l)/2,(m-l)/2)$. Calculation of $\tilde{G}(m,l)$ is described in the next section. For a real and symmetric $G(m,l)$, and complex noise, we get

$$\begin{aligned}\sigma^2_{s\epsilon}(n,f) \;=\; & 2\sum_{l_1}\sum_{m_1}\sum_{l_2}\sum_{m_2}\tilde{G}(m_1,l_1)\tilde{G}^*(m_2,l_2)s(n+m_1)s^*(n+m_2)\\ & \times R_{\epsilon\epsilon}(n+l_2,n+l_1)e^{-j2\pi f(m_1-l_1-m_2+l_2)},\end{aligned}$$

what can be written as

$$\sigma^2_{s\epsilon}(n,f) = 2\sum_{m_1}\sum_{m_2}\tilde{\Phi}(m_1,m_2)[s(n+m_1)e^{-j2\pi fm_1}][s(n+m_2)e^{-j2\pi fm_2}]^*, \quad (9.1.17)$$

where the new kernel $\tilde{\Phi}(m_1,m_2)$ reads

$$\tilde{\Phi}(m_1,m_2) = \sum_{l_1}\sum_{l_2}\tilde{G}(m_1,l_1)\tilde{G}^*(m_2,l_2)e^{-j2\pi f(l_2-l_1)}R_{\epsilon\epsilon}(n+l_2,n+l_1). \quad (9.1.18)$$

The signal dependent part of the variance $\sigma^2_\rho(n,f)$ is a quadratic distribution of the signal, with the new kernel $\tilde{\Phi}(m_1,m_2)$, i.e., $\sigma^2_{s\epsilon}(n,f) = 2\rho_s(n,f;\tilde{\Phi})$.

Special case 1: **White stationary complex noise**, when $R_{\epsilon\epsilon}(n+l_1,n+l_2) = \sigma^2_\epsilon\delta(l_1-l_2)$, produces

$$\tilde{\Phi}(m_1,m_2) = \sigma^2_\epsilon\sum_l \tilde{G}(m_1,l)\tilde{G}^*(m_2,l). \quad (9.1.19)$$

For time-frequency kernels we assumed realness and symmetry throughout the article, i.e., $\tilde{G}^*(m_2,l) = \tilde{G}(l,m_2)$. Thus, for finite limits (9.1.19) is a matrix multiplication form, $\tilde{\mathbf{\Phi}} = \sigma^2_\epsilon\tilde{\mathbf{G}}\cdot\tilde{\mathbf{G}}^* = \sigma^2_\epsilon\tilde{\mathbf{G}}^2$. Boldface letters, without arguments, will be used to denote a matrix. For example $\tilde{\mathbf{G}}$ is a matrix with elements $\tilde{G}(m,l)$. Thus,

$$\sigma^2_{s\epsilon}(n,f) = 2\rho_s(n,f;\sigma^2_\epsilon\tilde{\mathbf{G}}^2). \quad (9.1.20)$$

Note: Any two distributions with kernels $\tilde{G}_1(m,l) = \tilde{G}_2(m,-l)$ have the same variance, since

$$\sum_l \tilde{G}_1(m_1,l)\tilde{G}_1^*(m_2,l) = \sum_l \tilde{G}_1(m_1,-l)\tilde{G}_1^*(m_2,-l) = \sum_l \tilde{G}_2(m_1,l)\tilde{G}_2^*(m_2,l).$$

Corollary: A distribution with real and symmetric product kernel $g(\nu\tau)$ and the distribution with its dual kernel $g_d(\nu\tau) = \mathcal{F}_{\alpha\to\nu,\beta\to\tau}[g(\alpha\beta)]$ have the same variance.

[3] An analysis of the bias, i.e., kernel influence on the form of $\rho_s(n,f;G)$ may be found in [1].

Proof: Consider all coordinates in the analog domain. The time-lag domain forms of $g(\nu\tau)$, $G(t,\tau) = \mathcal{F}_{\nu\to t}[g(\nu\tau)]$, and $g_d(\nu\tau)$, $G_d(\nu\tau) = \mathcal{F}_{\nu\to t}[g_d(\nu\tau)]$ are related by $G(t,\tau) = G_d(\tau,t)$. In the rotated domain this relation produces $\tilde{G}(t_1,t_2) = \tilde{G}_d(t_1,-t_2)$, what ends the proof, according to the previous note.

Example: The WD has the kernel $g(\nu\tau) = 1$, $\tilde{G}(m,l) = \delta(m+l)$. According to the Corollary, the WD has the same variance as its dual kernel counterpart, with $g(\nu\tau) = \delta(\nu,\tau)$, $\tilde{G}(m,l) = \delta(m-l)$. This dual kernel corresponds to the signal energy $\sum_m |x(n+m)|^2$ (see (9.1.16)). Thus, the WD and the signal energy have the same variance. The same holds for the smoothed spectrogram, and the S-method [1], [Article 6.2], whose kernels are $\tilde{G}(m,l) = w(m)p(m+l)w(l)$, and $\tilde{G}(m,l) = w(m)p(m-l)w(l)$, respectively. Their variance is the same.

Eigenvalue decomposition: Assume that both the summation limits and values of $\tilde{G}(m,l)$ are finite. It is true when the kernel $G(m,l)$ is calculated from the well defined kernel in a finite Doppler-lag domain, $G(m,l) = \mathcal{F}_{\nu\to m}[g(\nu,l)]$, using a finite number of samples. The signal dependent part of the variance $\sigma^2_{s\epsilon}(n,f)$ can be calculated, like other distributions from the Cohen class, by using eigenvalue decomposition of matrix $\tilde{\mathbf{G}}$, [5,6]. The distribution of non-noisy signal (9.1.16) is

$$\rho_s(n,f) = \sum_{i=-N/2}^{N/2-1} \lambda_i S_s(n,f;q_i) = \rho_s(n,f;\lambda,q), \tag{9.1.21}$$

where λ_i and $q_i(m)$ are eigenvalues and eigenvectors of the matrix $\tilde{\mathbf{G}}$, respectively, and $S_s(n,f;q_i) = \left|\sum_{i=-N/2}^{N/2-1} s(n+m)q_i(m)e^{-j2\pi fm}\right|^2$ is the spectrogram of signal $s(n)$ calculated by using $q_i(m)$ as a lag window. Since $\tilde{\mathbf{\Phi}} = \sigma_\epsilon^2\tilde{\mathbf{G}}^2$, its eigenvalues and eigenvectors are $\sigma_\epsilon^2|\lambda_i|^2$ and $q_i(m)$, respectively. Thus, according to (9.1.20)

$$\sigma^2_{s\epsilon}(n,f) = 2\sigma_\epsilon^2 \sum_{i=-N/2}^{N/2-1} |\lambda_i|^2 S_s(n,f;q_i) = 2\sigma_\epsilon^2 \rho_s(n,f;|\lambda|^2,q). \tag{9.1.22}$$

Relation between the original kernel and variance $\sigma^2_{s\epsilon}(n,f)$ kernel: According to (9.1.21), we can conclude that the original kernel in the Doppler-lag domain can be decomposed into $g(\nu,l) = \sum_{i=-N/2}^{N/2-1} \lambda_i a_i(\nu,l)$, where $a_i(\nu,l)$ are ambiguity functions of the eigenvectors $q_i(m)$. The kernel of $\rho_s(n,f;|\lambda|^2,q)$, in (9.1.22), is $g_\sigma(\nu,l) = \sum_{i=-N/2}^{N/2-1} |\lambda_i|^2 a_i(\nu,l)$. A detailed analysis of distributions, with respect to their eigenvalue properties, is presented in [6], [Article 6.4]. In the sense of that analysis, the signal dependent variance is just "an energetic map of the time-frequency distribution" of the original signal.

The mean value of variance (9.1.17) is:

$$\overline{\sigma^2_{s\epsilon}(n,f)} = \int_{-1/2}^{1/2} \sigma^2_{s\epsilon}(n,f)df = 2\sigma_\epsilon^2 \sum\nolimits_m \tilde{\Phi}(m,m)\,|s(n+m)|^2\,. \tag{9.1.23}$$

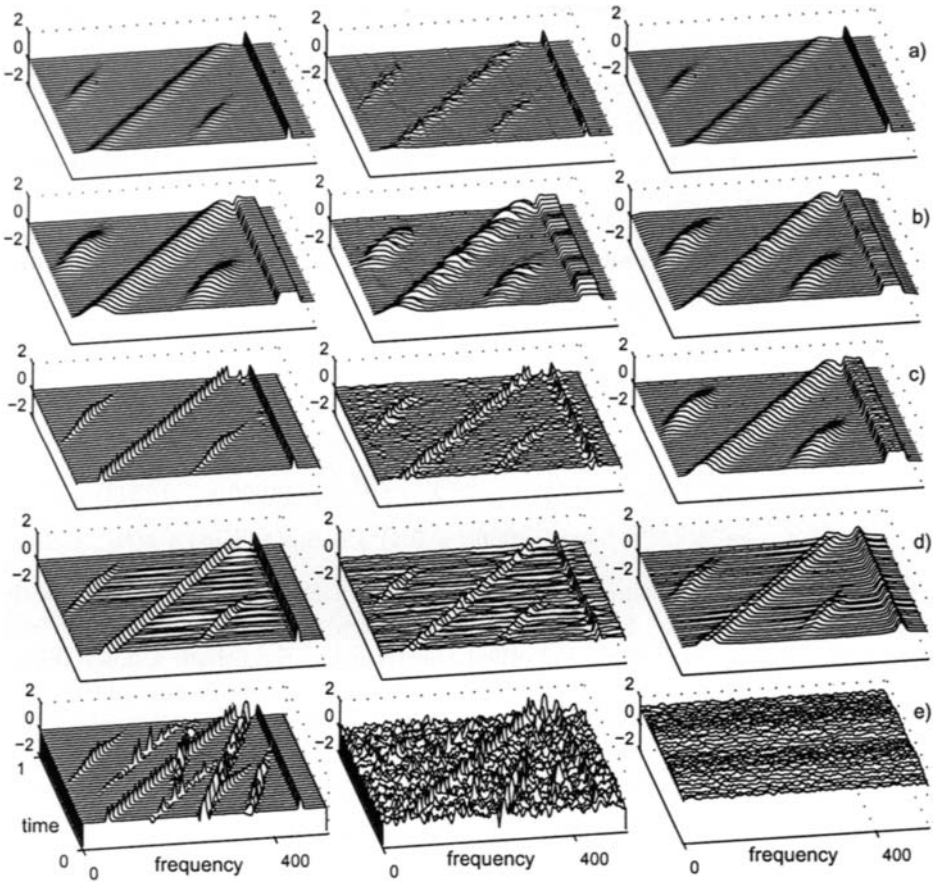


Fig. 9.1.1: Time-frequency representations of a non-noisy signal (First column); One realization of time-frequency representations of the signal corrupted by a white stationary complex noise (Second column); Variances of the distributions, obtained numerically by averaging over 1000 realizations (Third column): (a) Spectrogram, (b) Smoothed spectrogram, (c) S-method, (d) Choi-Williams distribution, (e) Pseudo Wigner distribution.

For frequency modulated signals $s(n) = a\exp[j\phi(n)]$ it is a constant proportional to the kernel energy [4].

Special case 2: For **nonstationary white complex noise**, (9.1.18) results in:

$$\tilde{\Phi}(m_1, m_2) = \sum_{i=-N/2}^{N/2-1} I(n+l)\tilde{G}(m_1, l)\tilde{G}^*(m_2, l), \tag{9.1.24}$$

or $\tilde{\mathbf{\Phi}} = \tilde{\mathbf{G}}\mathbf{I}_n\tilde{\mathbf{G}}$, where $\mathbf{I}_n$ is a diagonal matrix, with the elements $I(n+l)$. For the

quasi-stationary case, $I(n+l_1)\delta(l_1-l_2) \cong I(n)\delta(l_1-l_2)$, we have $\mu_i = I(n)\,|\lambda_i|^2$, with all other parameters as in (9.1.22).

Special case 3: In the case of **colored stationary complex noise,** relations dual to those in Special case 2, hold (like (9.1.9) and (9.1.11)).

Special case 4: Let $x(n) = s(n)(1+\mu(n))$, where $\mu(n)$ is a **multiplicative noise.** We can write $x(n) = s(n) + s(n)\mu(n) = s(n) + \epsilon(n)$, where $\epsilon(n) = s(n)\mu(n)$ is an additive noise. Thus, the case of this kind of multiplicative noise can be analyzed in the same way as the additive noise. For example, if the noise $\mu(n)$ is a nonstationary white complex one with $R_{\mu\mu}(m,n) = I_\mu(n)\delta(n-m)$, then $R_{\epsilon\epsilon}(m,n) = I_\epsilon(n)\delta(n-m)$, where $I_\epsilon(n) = |s(n)|^2 I_\mu(n)$.

9.1.4 Numerical Example

Consider the signal

$$\begin{aligned} x(t) &= \exp(j1100(t+0.1)^2) + e^{-25(t-0.25)^2}\exp(j1000(t+0.75)^2) \\ &\quad + e^{-25(t-0.67)^2}\exp(j1000(t-0.4)^2) + \exp(j2850t) + \epsilon(t), \end{aligned}$$

within the interval $[0,1]$, sampled at $\Delta t = 1/1024$. A Hanning lag window of the width $T_w = 1/4$ is used. Stationary white complex noise with variance $\sigma_\epsilon^2 = 2$ is assumed. The spectrogram, smoothed spectrogram, S-method, Choi-Williams distribution (CWD), and the WD, of signal without noise are presented in the first column of Fig. 9.1.1, respectively. For the CWD, the kernel $g(\nu,\tau) = \exp(-(\nu\tau)^2)$ is used, with normalized coordinates $-\sqrt{\pi N/2} \le |2\pi\nu| < \sqrt{\pi N/2}$, $-\sqrt{\pi N/2} \le |\tau| < \sqrt{\pi N/2}$, and 128 samples within the intervals. Elements of the matrix $\tilde{\mathbf{G}}$ were calculated as, [6]

$$\tilde{G}(m,l) = \sum_{p=-N/2}^{N/2} g(p\Delta\nu,(m-l)\Delta\tau)\exp(-j2\pi(m+l)p/(2N))\Delta\nu. \quad (9.1.25)$$

The normalized eigenvalues of the matrix $\tilde{\mathbf{\Phi}}$ were $\lambda_i = \{1,-0.87,0.69,-0.58,0.41,-0.30,...\}$ and $\mu_i = |\lambda_i|^2 = \{1,0.76,0.47,\ 0.33,0.17,0.09,...\}$. In the spectrogram and smoothed spectrogram the whole signal dependent part of variance is "located" just on the signal components, while in the WD it is "spread" over the entire time-frequency plane. Variance behavior in other two distributions is between these two extreme cases. As it has been shown, the variances in the smoothed spectrogram and the S-method are the same [Fig. 9.1.1(b) and (c)].

9.1.5 Summary and Conclusions

The variance values for a white nonstationary complex noise, with $R_{\epsilon\epsilon}(m,n) = I(n)\delta(m-n)$, $I(n) \ge 0$, for some distributions, are summarized next.
-Pseudo Wigner distribution $W_s(n,f;w)$, with $\tilde{G}(m,l) = w(m)\delta(m+l)w(l)$:

$$\sigma^2_{WD}(n,f) = \sigma^2_{\epsilon\epsilon}(n,f) + \sigma^2_{s\epsilon}(n,f) = W_I(n,0;w^2) + 2W_{I,|s|^2}(n,0;w^2), \quad (9.1.26)$$

where $W_{I,|s|^2}$ denotes the cross Wigner distribution for $I(n)$ and $|s(n)|^2$.
-Spectrogram $S_s(n,f;w)$, with $\tilde{G}(m,l) = w(m)w(l)$:

$$\sigma^2_{SPEC}(n,f) = S_I(n,0;w^2) + 2F_I(n,0;w^2)S_s(n,f;w). \tag{9.1.27}$$

The STFT of $I(n)$, calculated using the window $w^2(m)$, is denoted by $F_I(n,f;w^2)$.
-A general quadratic distribution, with kernel $\tilde{G}(m,l) = G((m+l)/2,(m-l)/2)$, in (9.1.6) or (9.1.16), and $\tilde{\mathbf{G}}$ being a matrix with elements $\tilde{G}(m,l)$:

$$\sigma^2_\rho(n,f) = \rho_I(n,0;|G|^2) + 2\rho_s(n,f;\tilde{\mathbf{G}}\mathbf{I}_n\tilde{\mathbf{G}}). \tag{9.1.28}$$

The first two formulae are special cases of (9.1.28). Expressions for stationary white noise follow with $I(n) = \sigma^2_\epsilon$. Dual expressions hold for a colored stationary noise. Further details can be found in [7].

References

[1] L. Stanković and V. Katkovnik, "The Wigner distribution of noisy signals with adaptive time-frequency varying window," *IEEE Trans. Signal Processing*, vol. 47, pp. 1099–1108, April 1999.

[2] P. Duvaut and D. Declerq, "Statistical properties of the pseudo-Wigner-Ville representation of normal random processes," *Signal Processing*, vol. 75, pp. 93–98, 5 January 1999.

[3] L. Stanković and V. Ivanović, "Further results on the minimum variance time-frequency distributions kernels," *IEEE Trans. Signal Processing*, vol. 45, pp. 1650–1655, June 1997.

[4] M. G. Amin, "Minimum-variance time-frequency distribution kernels for signals in additive noise," *IEEE Trans. Signal Processing*, vol. 44, pp. 2352–2356, September 1996.

[5] M. G. Amin, "Spectral decomposition of time-frequency distribution kernels," *IEEE Trans. Signal Processing*, vol. 42, pp. 1156–1165, May 1994.

[6] G. S. Cunningham and W. J. Williams, "Kernel decomposition of time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, pp. 1425–1442, June 1994.

[7] L. Stanković, "Analysis of noise in time-frequency distributions," *IEEE Signal Processing Letters*, vol. 9, pp. 286–289, September 2002.

[8] 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.), pp. 211–267, Amsterdam: Elsevier, 1997.

[9] S. B. Hearon and M. G. Amin, "Minimum-variance time-frequency distribution kernels," *IEEE Trans. Signal Processing*, vol. 43, pp. 1258–1262, May 1995.

[10] L. Stanković and S. Stanković, "On the Wigner distribution of the discrete-time noisy signals with application to the study of quantization effects," *IEEE Trans. Signal Processing*, vol. 42, pp. 1863–1867, July 1994.

9.2 STATISTICAL PROCESSING OF DISPERSIVE SYSTEMS AND SIGNALS[0]

9.2.1 Processing Tools For Time-Varying Systems and Signals

Time-frequency representations (TFRs) such as the Wigner distribution (WD) are powerful tools for analyzing deterministic signals whose spectral characteristics vary with time. In many applications, one can observe linear time-varying (LTV) systems (e.g. wireless communication or sonar channels) and nonstationary random processes (e.g. noisy signals with random parameters). Due to their time fluctuations, both of these can exhibit time-frequency (TF) characteristics that could be embedded in the model design of a system or in the autocorrelation function of a random process. Thus, it would be ideal to design transforms or TFRs that would extract useful information from these characteristics [1–5]. The transforms could preserve some important system or random process changes. Specifically, signals propagating over LTV systems may be shifted by a constant amount in frequency or shifted in a non-linear time-dependent manner due to a *dispersive* change in the signal's instantaneous frequency (IF). Note that dispersive implies that the IF change or shift varies non-linearly with time. Information of a constant frequency change, for example, is very useful in improving the Doppler diversity of a communications channel (see [6] and Article 13.2). Furthermore, information of an IF change on the input signal to a dispersive system can be used to improve system performance.

An appropriate analysis tool for LTV systems and random processes is a class of TFRs called *TF symbols* [1–5] which are 2-D functions that depend on an integral operator $\mathcal{L}$ defined as $(\mathcal{L}x)(t) = \int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}(t,\tau)\, x(\tau)\, d\tau$. The kernel $\Upsilon_{\mathcal{L}}(t,\tau)$ of the operator can be considered as the time-varying impulse response of an LTV system or the autocorrelation function of a nonstationary random process.

Narrowband Weyl Symbol The narrowband *Weyl symbol* (WS) of a linear operator $\mathcal{L}$ on $L_2(\mathbb{R})$, defined as

$$\mathrm{WS}_{\mathcal{L}}(t,f) = \int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}\left(t+\frac{\tau}{2}, t-\frac{\tau}{2}\right) e^{-j2\pi\tau f}\, d\tau\,, \tag{9.2.1}$$

is an important tool for analyzing LTV systems and nonstationary random processes characterized by constant TF shifts and scale changes [1, 2]. It can be interpreted as a time-varying transfer function of an LTV system (see [1] and Articles 4.7 & 12.4) or as a time-varying spectrum of a nonstationary random process (Article 9.4). When $\mathcal{L}$ is the autocorrelation operator $\mathcal{R}_x$ of a random process $x(t)$ whose kernel

[0]Authors: **A. Papandreou-Suppappola**, Telecommunications Research Center, Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-7206 USA (papandreou@asu.edu), **B.-G. Iem**, Electronics Department, Kangnung National University, Kangwondo, 210-702 South Korea (iembg@yahoo.com), and **G. Faye Boudreaux-Bartels**, Department of Electrical and Computer Engineering, University of Rhode Island, Kingston, RI 02881 USA (boud@ele.uri.edu). Reviewers: A. M. Sayeed and G. Matz.

Table 9.2.1: Summary of some commonly used operators. Depending on the sign of t, $\mathrm{sgn}(t) = \pm 1$.

Operators	Operator Definitions
$(\mathcal{S}_\tau\, x)(t)$	$x(t-\tau)$
$(\widetilde{\mathcal{S}}_\zeta^{(\xi)}\, x)(t)$	$(\mathcal{W}_\xi^{-1}\, \mathcal{S}_{t_r\zeta}\, \mathcal{W}_\xi\, x)(t)$
$(\mathcal{M}_\nu\, x)(t)$	$x(t)\, e^{j2\pi\nu t}$
$(\widetilde{\mathcal{M}}_\beta^{(\xi)}\, x)(t)$	$x(t)\, e^{j2\pi\beta\, \xi(\frac{t}{t_r})} = (\mathcal{W}_\xi^{-1}\, \mathcal{M}_{\beta/t_r}\, \mathcal{W}_\xi\, x)(t)$
$(\mathcal{H}_\beta\, x)(t)$	$x(t)\, e^{j2\pi\beta \ln(\frac{t}{t_r})}$
$(\mathcal{P}_\beta^\kappa\, x)(t)$	$x(t)\, e^{j2\pi\beta\, \mathrm{sgn}(t)\, \lvert\frac{t}{t_r}\rvert^{\kappa-1}}$
$(\mathcal{E}_\beta\, x)(t)$	$x(t)\, e^{j2\pi\beta e^{t/t_r}}$
$(\mathcal{C}_a\, x)(t)$	$\frac{1}{\sqrt{\lvert a\rvert}}\, x(\frac{t}{a})$
$(\widetilde{\mathcal{C}}_a^{(\xi)}\, x)(t)$	$(\mathcal{W}_\xi^{-1}\, \mathcal{C}_a\, \mathcal{W}_\xi\, x)(t)$
$(\mathcal{W}_\xi\, x)(t)$	$x\left(t_r \xi^{-1}\left(\frac{t}{t_r}\right)\right) \left\lvert t_r\, \upsilon\left(t_r \xi^{-1}\left(\frac{t}{t_r}\right)\right)\right\rvert^{-1/2}$

is the autocorrelation function $\Upsilon_{\mathcal{R}_x}(t,\tau) = E\{x(t)\, x^*(\tau)\}$, the WS of $\mathcal{R}_x$ is the expected value of the WD, i.e. $\mathrm{WS}_{\mathcal{R}_x}(t,f) = E\{\mathrm{WD}_x(t,f)\}$ (called the Wigner-Ville spectrum in [7] and Section 2.2.1). Here, $E\{\cdot\}$ denotes statistical expectation and the WD is defined as $\mathrm{WD}_x(t,f) = \int_{-\infty}^{\infty} x(t+\frac{\tau}{2})\, x^*(t-\frac{\tau}{2})\, e^{-j2\pi\tau f}\, d\tau$. When $\mathcal{L}$ is an innovations system, the squared magnitude of (9.2.1) is known as the Weyl spectrum in [8]. Note that the WS has been used to provide a definition of a TF concentration measure [2], and is useful in TF detection applications [7,8].

The WS satisfies many desirable properties. Specifically, it preserves constant TF shifts, and scale changes (see Table 9.2.1) on LTV systems and nonstationary random processes. For example, the WS of the autocorrelation operator $\mathcal{R}_x$ of a nonstationary random process $x(t)$ satisfies [1]

$$y(t) = (\mathcal{S}_\tau\, x)(t) = x(t-\tau) \;\Rightarrow\; \mathrm{WS}_{\mathcal{R}_y}(t,f) = \mathrm{WS}_{\mathcal{R}_x}(t-\tau, f) \quad (9.2.2)$$

$$y(t) = (\mathcal{M}_\nu\, x)(t) = x(t)\, e^{j2\pi\nu t} \;\Rightarrow\; \mathrm{WS}_{\mathcal{R}_y}(t,f) = \mathrm{WS}_{\mathcal{R}_x}(t, f-\nu) \quad (9.2.3)$$

$$y(t) = (\mathcal{C}_a\, x)(t) = \frac{1}{\sqrt{\lvert a\rvert}}\, x\left(\frac{t}{a}\right) \;\Rightarrow\; \mathrm{WS}_{\mathcal{R}_y}(t,f) = \mathrm{WS}_{\mathcal{R}_x}\left(\frac{t}{a}, af\right) \quad (9.2.4)$$

where $\mathcal{S}_\tau$, $\mathcal{M}_\nu$ and $\mathcal{C}_a$ are the constant time-shift, constant frequency-shift and scale operators, respectively. The WS also satisfies the unitarity property given by

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{WS}_{\mathcal{L}}(t,f)\, \mathrm{WS}^*_{\mathcal{N}}(t,f)\, dt\, df = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}(t,\tau)\, \Upsilon^*_{\mathcal{N}}(t,\tau)\, dt\, d\tau \quad (9.2.5)$$

where $\Upsilon_{\mathcal{L}}(t,\tau)$ and $\Upsilon_{\mathcal{N}}(t,\tau)$ are the kernels of the operators $\mathcal{L}$ and $\mathcal{N}$, respectively. This is an important property for preserving energy or norms. Since it is unitary,

the WS is *associated* with the unitary WD and preserves the quadratic form

$$\int_{-\infty}^{\infty} (\mathcal{L}x)(t)\, x^*(t)\, dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{WS}_{\mathcal{L}}(t,f)\, \mathrm{WD}_x(t,f)\, dt\, df\,. \tag{9.2.6}$$

Here, the correlation of the WS of a system $\mathcal{L}$ with the WD of the system's input $x(t)$ results in the correlation of the system's output $(\mathcal{L}x)(t)$ with $x(t)$ [1,2].

Narrowband Spreading Function The spreading function (SF) of a linear operator $\mathcal{L}$ is defined as the 2-D Fourier transform of the WS,

$$\begin{aligned}\mathrm{SF}_{\mathcal{L}}(\tau,\nu) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{WS}_{\mathcal{L}}(t,f)\, e^{-j2\pi(t\nu - f\tau)}\, dt\, df \\ &= \int_{-\infty}^{\infty} \Upsilon_{\mathcal{L}}\left(t+\frac{\tau}{2}, t-\frac{\tau}{2}\right) e^{-j2\pi t\nu}\, dt\,. \end{aligned} \tag{9.2.7}$$

Using the SF, the output of an LTV system $\mathcal{L}$ can be interpreted as a superposition of TF shifted versions of the input signal each weighted by the SF [2]

$$(\mathcal{L}x)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{SF}_{\mathcal{L}}(\tau,\nu)\, e^{-j\pi\tau\nu}\, x(t-\tau)\, e^{j2\pi\nu t}\, d\tau\, d\nu\,. \tag{9.2.8}$$

The system output interpretation in (9.2.8) is comparable to that of a linear time-invariant (LTI) system, and it can be used to formulate TF receiver structures. For example, it can provide critical information on the output signal of a time and frequency selective wireless communications channel. The SF is also used in the classification of operators as underspread or overspread [1].

Wideband Weyl Symbol The TF version of the affine WS , introduced in [2], is called P_0-Weyl symbol (P_0WS) in [3,4], and for $\lambda(\alpha) = \frac{\alpha/2}{\sinh(\alpha/2)}$, it is defined as

$$\mathrm{P}_0\mathrm{WS}_{\mathcal{B}}(t,f) = f\int_{-\infty}^{\infty} \Gamma_{\mathcal{B}}\left(f\lambda(\alpha)e^{\frac{\alpha}{2}}, f\lambda(\alpha)e^{-\frac{\alpha}{2}}\right)\lambda(\alpha)\, e^{j2\pi f t\alpha}\, d\alpha\,,\quad f>0\,,$$

where $(\mathcal{B}X)(f) = \int_0^{\infty} \Gamma_{\mathcal{B}}(f,\nu)\, X(\nu)\, d\nu$, and $\Gamma_{\mathcal{B}}(f,\nu)$ is the kernel of the frequency domain operator $\mathcal{B}$ on $L_2(\mathbb{R}^+)$. The P_0WS is a unitary symbol that is associated with the unitary Bertrand P_0-distribution [9]. The wideband SF (WSF) is the 2-D modified Fourier transform of $\mathrm{P}_0\mathrm{WS}_{\mathcal{B}}(t,f)$ [2], and they are both important for analyzing random processes and LTV systems characterized by constant or hyperbolic time shifts and scale changes [2–4].

9.2.2 Dispersive Time-Frequency Symbols

In nature, there are systems and random processes characterized by dispersive time or frequency shifts (see [9] and Article 5.6). For example, the ocean is a system that can cause echoes with power dispersive characteristics from acoustic waves reflected from immersed spherical shells [10]. The WS is not well-suited to analyze

such systems and processes as it is better matched to constant TF shifts [1, 3, 4]. Instead, as will be shown next, modified versions of the narrowband WS and SF were developed for analyzing random processes and LTV systems characterized by dispersive (such as power) IF shifts [3]. Note that extensions of the wideband P_0WS for arbitrary dispersive analysis is provided in [5].

Dispersive Weyl Symbol Let the linear operator $\mathcal{Z}$, defined in the time domain on $L_2([p, q])$, be $(\mathcal{Z}x)(t) = \int_p^q \Upsilon_{\mathcal{Z}}(t,\tau)\, x(\tau)\, d\tau$, with kernel $\Upsilon_{\mathcal{Z}}(t,\tau)$. The *dispersive Weyl symbol*, $\mathrm{DWS}_{\mathcal{Z}}^{(\xi)}(t,f)$, of $\mathcal{Z}$ is defined in [3] as

$$\mathrm{DWS}_{\mathcal{Z}}^{(\xi)}(t,f) = \int_c^d \Upsilon_{\mathcal{Z}}\Big(l(t,\zeta),\, l(t,-\zeta)\Big)\, e^{-j2\pi \frac{f\zeta}{\upsilon(t)}}\, \frac{d\zeta}{|\upsilon(l(t,\zeta))\,\upsilon(l(t,-\zeta))|^{\frac{1}{2}}} \tag{9.2.9}$$

$$= \mathrm{WS}_{\mathcal{W}_\xi\, \mathcal{Z}\, \mathcal{W}_\xi^{-1}}\left(t_r\, \xi\left(\frac{t}{t_r}\right),\, \frac{f}{t_r\, \upsilon(t)}\right) \tag{9.2.10}$$

where $l(t,\zeta) = t_r\, \xi^{-1}(\xi(\frac{t}{t_r}) + \frac{\zeta}{2})$, $\xi^{-1}(\xi(b)) = b$, $\upsilon(t) = \frac{d}{dt}\xi(\frac{t}{t_r})$, and the integration range $[c, d]$ in (9.2.9) and $[p, q]$ above depend on the range and domain, respectively, of $\xi(\cdot)$. Note that $t_r > 0$ is a reference time point that is needed for unit precision (unless otherwise stated, it could be taken as $t_r = 1$). Thus, the DWS is obtained by warping the operator $\mathcal{Z}$ and the WS in (9.2.1) using a unitary transformation ([3, 9, 11] and Article 5.6) based on the one-to-one warping function $\xi(b)$. With $(\mathcal{W}_\xi^{-1}\, \mathcal{W}_\xi\, x)(t) = x(t)$, the unitary warping operator $\mathcal{W}_\xi$ in (9.2.10) is given by

$$(\mathcal{W}_\xi\, x)(t) = x\left(t_r\, \xi^{-1}\left(\frac{t}{t_r}\right)\right)\, \left|t_r\, \upsilon\left(t_r\, \xi^{-1}\left(\frac{t}{t_r}\right)\right)\right|^{-1/2}.$$

The DWS preserves dispersive IF shifts on a random process $x(t)$, i.e.

$$y(t) = x(t)\, e^{j2\pi\beta\, \xi(\frac{t}{t_r})} \quad \Rightarrow \quad \mathrm{DWS}_{\mathcal{R}_y}^{(\xi)}(t,f) = \mathrm{DWS}_{\mathcal{R}_x}^{(\xi)}(t, f - \beta\, \upsilon(t)), \tag{9.2.11}$$

where[1] $y(t) = (\widetilde{\mathcal{M}}_\beta^{(\xi)}\, x)(t) = x(t)\, e^{j2\pi\beta\xi(\frac{t}{t_r})}$, $\widetilde{\mathcal{M}}_\beta^{(\xi)}$ is the IF shift operator, and $\beta\, \upsilon(t)$ is the change in IF. This follows as the WS preserves constant frequency shifts $(\mathcal{M}_\nu\, x)(t)$ in (9.2.3), and the warping in (9.2.10) yields $\mathcal{W}_\xi^{-1}\, \mathcal{M}_{\beta/t_r}\, \mathcal{W}_\xi = \widetilde{\mathcal{M}}_\beta^{(\xi)}$. Because of this important property, the DWS is potentially useful in analyzing random processes or LTV systems with characteristics that may be constant (when $\xi(b) = b$) or dispersive (when $\xi(b)$ is non-linear). The DWS also preserves warped time shifts $(\widetilde{\mathcal{S}}_\zeta^{(\xi)}\, x)(t) = (\mathcal{W}_\xi^{-1}\, \mathcal{S}_{t_r\zeta}\, \mathcal{W}_\xi\, x)(t)$ (cf. (9.2.2)) and warped scale changes $(\widetilde{\mathcal{C}}_a^{(\xi)}\, x)(t) = (\mathcal{W}_\xi^{-1}\, \mathcal{C}_a\, \mathcal{W}_\xi\, x)(t)$ (cf. (9.2.4)) as defined in [3].

The importance of the DWS is further emphasized when used to analyze random non-linear frequency modulated (FM) signals $x(t) = \mathbf{a}\, \varrho(t)\, e^{j2\pi\beta\phi(\frac{t}{t_r})}$ with phase function $\phi(b)$ and random amplitude $\mathbf{a}$ (see [9] and Article 5.6). When the phase

[1]The tilde above operators indicates the warped versions of the operators in (9.2.2)–(9.2.4).

function of the non-linear FM signal is matched to the warping function $\xi(b)$ used in the DWS, i.e. when $\phi(b) = \xi(b)$ and $\varrho(t) = |\frac{d}{dt}\phi(\frac{t}{t_r})|^{1/2} = |\upsilon(t)|^{1/2}$, then the DWS of the autocorrelation operator of the process produces an ideally localized representation,

$$\mathrm{DWS}_{\mathcal{R}_x}^{(\xi)}(t,f) = E\{|\mathsf{a}|^2\}\,\delta\left(f - \beta\,\upsilon(t)\right)\,,$$

provided that the range of $\xi(b)$ is $\mathbb{R}$. The DWS is localized along the curve $f = \beta\,\upsilon(t)$, corresponding to the IF of the non-linear FM signal.

Dispersive Spreading Function The dispersive spreading function (DSF) is obtained as the SF in (9.2.7) of the warped operator [3] $\mathcal{W}_\xi\,\mathcal{Z}\,\mathcal{W}_\xi^{-1}$,

$$\mathrm{DSF}_{\mathcal{Z}}^{(\xi)}(\zeta,\beta) = \mathrm{SF}_{\mathcal{W}_\xi\,\mathcal{Z}\,\mathcal{W}_\xi^{-1}}(t_r\,\zeta,\,\beta/t_r)\,. \tag{9.2.12}$$

For an LTV system $\mathcal{Z}$, the DSF provides an interpretation of the operator output as a weighted superposition of warped TF shifted versions of the input signal $x(t)$,

$$(\mathcal{Z}x)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{DSF}_{\mathcal{Z}}^{(\xi)}(\zeta,\beta)\,e^{-j\pi\zeta\beta}\,(\widetilde{\mathcal{M}}_\beta^{(\xi)}\,\widetilde{\mathcal{S}}_\zeta^{(\xi)}\,x)(t)\,d\zeta\,d\beta\,. \tag{9.2.13}$$

Depending on the warping function $\xi(b)$, the formulation in (9.2.13) simplifies to a specific interpretation on the operator output. For example, when $\xi(b) = \xi_{\ln}(b) = \ln b$, then $(\widetilde{\mathcal{M}}_\beta^{(\xi_{\ln})}\,\widetilde{\mathcal{S}}_\zeta^{(\xi_{\ln})}\,x)(t) = e^{-\frac{\zeta}{2}}\,x(te^{-\zeta})\,e^{j2\pi\beta\,\ln(\frac{t}{t_r})}$, and (9.2.13) describes the operator $\mathcal{Z}$ as a weighted superposition of hyperbolic IF shifts and scale changes (by e^ζ) on the input $x(t)$ [3].

Unitarity and Quadratic Form Properties of the DWS The DWS satisfies the unitarity property in (9.2.5) since [3]

$$\int_{-\infty}^{\infty}\int_{p}^{q} \mathrm{DWS}_{\mathcal{Z}}^{(\xi)}(t,f)\,\mathrm{DWS}_{\mathcal{X}}^{(\xi)\,*}(t,f)\,dt\,df = \int_{p}^{q}\int_{p}^{q} \Upsilon_{\mathcal{Z}}(t,\tau)\,\Upsilon_{\mathcal{X}}^{*}(t,\tau)\,dt\,d\tau. \tag{9.2.14}$$

The DWS is *associated* with $\mathrm{DWD}_x^{(\xi)}(t,f) = \mathrm{WD}_{\mathcal{W}_\xi x}\left(t_r\,\xi(\frac{t}{t_r}),\,\frac{f}{t_r\,\upsilon(t)}\right)$ which is the IF shift covariant version of the WD (see [9] and Article 5.6). Using the DWS, a general expression of the quadratic form in (9.2.6) is

$$\int_{p}^{q} (\mathcal{Z}x)(t)\,x^*(t)dt = \int_{-\infty}^{\infty}\int_{p}^{q} \mathrm{DWS}_{\mathcal{Z}}^{(\xi)}(t,f)\,\mathrm{DWD}_x^{(\xi)}(t,f)\,dt\,df \tag{9.2.15}$$

with potential detection applications for non-linear TF processes [3].

9.2.3 Special Cases of Dispersive Time-Frequency Symbols

Depending on the warping function $\xi(b)$, the DWS may simplify to TF symbols that are matched to linear or non-linear (dispersive) TF structures. Note that the trivial case of $\xi(b) = b$ simplifies the DWS to the WS. Some other examples are described below.

Hyperbolic WS and SF The hyperbolic Weyl symbol (HWS) is obtained as the warped version of the WS when $\xi(b) = \xi_{\ln}(b) = \ln b$ in the DWS formulation in (9.2.10) and (9.2.11). That is, for a linear time domain operator $\mathcal{J}$ on $L_2(\mathbb{R}^+)$, $\mathrm{HWS}_{\mathcal{J}}(t,f) = \mathrm{DWS}_{\mathcal{J}}^{(\xi_{\ln})}(t,f) = \mathrm{WS}_{\mathcal{W}_{\xi_{\ln}}\mathcal{J}\mathcal{W}_{\xi_{\ln}}^{-1}}\left(t_r \ln\left(\frac{t}{t_r}\right), \frac{tf}{t_r}\right)$, $t > 0$. This yields

$$\mathrm{HWS}_{\mathcal{J}}(t,f) = t\int_{-\infty}^{\infty} \Upsilon_{\mathcal{J}}\left(te^{\zeta/2}, te^{-\zeta/2}\right) e^{-j2\pi tf\zeta}\, d\zeta \qquad (9.2.16)$$

where $\Upsilon_{\mathcal{J}}(t,\tau)$ is the kernel of $\mathcal{J}$ [3]. The HWS is significant for processing systems and nonstationary signals with hyperbolic TF characteristics. Specifically, it preserves *hyperbolic* IF changes on a random process $x(t)$

$$y(t) = x(t)\, e^{j2\pi\beta \ln\left(\frac{t}{t_r}\right)} \quad \Rightarrow \quad \mathrm{HWS}_{\mathcal{R}_y}(t,f) = \mathrm{HWS}_{\mathcal{R}_x}\left(t, f - \frac{\beta}{t}\right),$$

where $y(t) = (\widetilde{\mathcal{M}}_{\beta}^{(\xi_{\ln})} x)(t) = (\mathcal{H}_{\beta}\, x)(t)$ is the hyperbolic IF shift operator (obtained when $\xi(b) = \ln b$ in (9.2.11)). The HWS also satisfies the scale covariance property in (9.2.4) since $\widetilde{\mathcal{S}}_{\zeta}^{(\xi_{\ln})} = \mathcal{C}_{e^{\zeta}}$ [3]. The HWS satisfies the unitarity property in (9.2.14) with $\xi(b) = \ln b$ and $[p,q] = [0,\infty)$. It is *associated* with the dual form of the Altes Q-distribution ([9] and Article 5.6).

As an example of the HWS, if the output of a system $\mathcal{J}$ is the scale convolution of an input signal $x(t)$ and some function $g(t)$, then the HWS in (9.2.16) is the Mellin transform of $g(t)$, i.e. $\mathrm{HWS}_{\mathcal{J}}(t,f) = \int_0^{\infty} g(\tau)\, \frac{1}{\sqrt{\tau}}\, e^{-j2\pi tf \ln\left(\frac{\tau}{t_r}\right)}\, d\tau$. This is intuitive as the Mellin is a natural transform for scale operations. For comparison, the WS in (9.2.1), $\mathrm{WS}_{\mathcal{J}}(t,f) = \int_0^{\infty} \frac{\sqrt{t_r}}{t-\tau/2}\, g\left(t_r \frac{t+\tau/2}{t-\tau/2}\right) e^{-j2\pi\tau f}\, d\tau$, of the *same* operator $\mathcal{J}$ is difficult to interpret.

The hyperbolic spreading function (HSF) is obtained from (9.2.12) when $\xi(b) = \xi_{\ln}(b) = \ln b$ as $\mathrm{HSF}_{\mathcal{J}}(\zeta,\beta) = \mathrm{DSF}_{\mathcal{J}}^{(\xi_{\ln})}(\zeta,\beta) = \mathrm{SF}_{\mathcal{W}_{\xi_{\ln}}\mathcal{J}\mathcal{W}_{\xi_{\ln}}^{-1}}(t_r\zeta, \beta/t_r)$ yielding [3]

$$\mathrm{HSF}_{\mathcal{J}}(\zeta,\beta) = \int_0^{\infty} \Upsilon_{\mathcal{J}}\left(te^{\zeta/2}, te^{-\zeta/2}\right) e^{-j2\pi\beta \ln\left(\frac{t}{t_r}\right)} dt\,.$$

It is related to the HWS using a modified Fourier transform and a Mellin transform as $\mathrm{HSF}_{\mathcal{J}}(\zeta,\beta) = \int_{-\infty}^{\infty}\int_0^{\infty} \mathrm{HWS}_{\mathcal{J}}(t,f) e^{j2\pi\zeta tf}\, e^{-j2\pi\beta \ln\left(\frac{t}{t_r}\right)}\, dt\, df$ [3]. The HSF provides an alternative interpretation of the operator output as a weighted superposition of hyperbolic IF shifted and scale changed versions of the input signal where the weight is the HSF, i.e.,

$$(\mathcal{J}x)(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{HSF}_{\mathcal{J}}(\zeta,\beta)\, e^{-j\pi\zeta\beta}\, (\mathcal{H}_{\beta}\, \mathcal{C}_{e^{\zeta}}\, x)(t)\, d\zeta\, d\beta, \quad t > 0.$$

It is useful for analyzing systems characterized by hyperbolic IF shifts and scale changes. The hyperbolic version of the quadratic form in (9.2.15) with $\xi(b) = \ln b$ is useful in detector formulations when signals in noise have hyperbolic TF characteristics [3].

Power WS and SF When $\xi(b) = \xi_\kappa(b) = \mathrm{sgn}(b)|b|^\kappa$ and $\upsilon_\kappa(t) = \frac{d}{dt}\xi_\kappa(\frac{t}{t_r})$, the DWS in (9.2.9) and the DSF in (9.2.12) simplify, respectively, to the κth power WS (PWS) and the κth power SF (PSF) for $\mathcal{L} \in L_2(\mathbb{R})$

$$\mathrm{PWS}_{\mathcal{L}}^{(\kappa)}(t,f) = \mathrm{DWS}_{\mathcal{L}}^{(\xi_\kappa)}(t,f) = \mathrm{WS}_{\mathcal{W}_{\xi_\kappa}\mathcal{L}\mathcal{W}_{\xi_\kappa}^{-1}}\Big(t_r\xi_\kappa\Big(\frac{t}{t_r}\Big),\, \frac{f}{t_r\upsilon_\kappa(t)}\Big),$$

$$\mathrm{PSF}_{\mathcal{L}}^{(\kappa)}(\zeta,\beta) = \mathrm{DSF}_{\mathcal{L}}^{(\xi_\kappa)}(\zeta,\beta) = \mathrm{SF}_{\mathcal{W}_{\xi_\kappa}\mathcal{L}\mathcal{W}_{\xi_\kappa}^{-1}}(t_r\zeta,\, \beta/t_r)$$

where $(\mathcal{W}_{\xi_\kappa}x)(t) = x\Big(t_r\xi_\kappa^{-1}(\frac{t}{t_r})\Big)\, |t_r\upsilon_\kappa(t_r\xi_\kappa^{-1}(\frac{t}{t_r}))|^{-1/2}$.

The PWS preserves power IF shifts on a random process $x(t)$, i.e.

$$y(t) = x(t)\, e^{j2\pi\beta\,\xi_\kappa(\frac{t}{t_r})} \quad \Rightarrow \quad \mathrm{PWS}_{\mathcal{R}_y}^{(\kappa)}(t,f) = \mathrm{PWS}_{\mathcal{R}_x}^{(\kappa)}\Big(t, f - \beta\,\upsilon_\kappa(t)\Big)$$

where $y(t) = (\widetilde{\mathcal{M}}_\beta^{(\xi_\kappa)}\, x)(t) = (\mathcal{P}_\beta^\kappa\, x)(t)$ is the power IF shift operator (the operator in (9.2.11) when $\xi(b) = \xi_\kappa(b)$). The PWS also preserves scale changes since warped scale covariance simplifies to the scale covariance in (9.2.4), i.e. $\widetilde{\mathcal{C}}_a^{(\xi_\kappa)} = \mathcal{C}_{\xi_{1/\kappa}(a)}$. The corresponding operator output can be interpreted as a weighted superposition of power IF shifted and power warped time-shifted versions of the input signal. Also, the PWS is unitary as it satisfies (9.2.14) and it is *associated* with the power WD ([9] and Articles 5.6 & 15.3) in the quadratic form in (9.2.15) when $\xi(b) = \xi_\kappa(b)$ [3].

Exponential WS and SF For a linear operator $\mathcal{L}$ defined on $L_2(\mathbb{R})$, the exponential WS, $\mathrm{EWS}_{\mathcal{L}}(t,\, f) = \mathrm{WS}_{\mathcal{W}_{\xi_e}\mathcal{L}\mathcal{W}_{\xi_e}^{-1}}(t_re^{t/t_r},\, fe^{-t/t_r})$, and the exponential SF, $\mathrm{ESF}_{\mathcal{L}}(\zeta,\, \beta) = \mathrm{SF}_{\mathcal{W}_{\xi_e}\mathcal{L}\mathcal{W}_{\xi_e}^{-1}}(t_r\zeta, \beta/t_r)$, are given as exponentially warped versions of the narrowband WS and SF, respectively. Here, $(\mathcal{W}_{\xi_e}x)(t) = x(t_r\ln(\frac{t}{t_r}))\,\sqrt{t_r/t}$, $t > 0$. The EWS and the ESF are obtained from the DWS in (9.2.9) and the DSF in (9.2.12), respectively, when $\xi(b) = e^b$. The EWS preserves exponential IF shifts on a random process $x(t)$,

$$y(t) = x(t)\, e^{j2\pi\beta\, e^{t/t_r}} \quad \Rightarrow \quad \mathrm{EWS}_{\mathcal{R}_y}(t,f) = \mathrm{EWS}_{\mathcal{R}_x}\Big(t, f - \beta\, e^{t/t_r}/t_r\Big)$$

where $y(t) = (\mathcal{D}_\beta^{(\xi_e)}\, x)(t) = (\mathcal{E}_\beta\, x)(t)$. Also, the EWS preserves constant time shifts in (9.2.2) since $\widetilde{\mathcal{C}}_a^{(\xi_e)} = \mathcal{S}_{t_r\ln a}$, is unitary, and satisfies the quadratic form in (9.2.15) with *association* with the exponential WD (see [9] and Article 5.6).

9.2.4 Analysis Application Examples

When a process or system has distinct TF characteristics in a particular application, it is important to choose an adequate analysis tool. As shown next, a dispersive WS produces an ideally localized spectrum of a process when they both have similar dispersive TF characteristics. Fig. 9.2.1 demonstrates the advantage of the HWS over the WS when used to analyze signals with hyperbolic TF characteristics. The

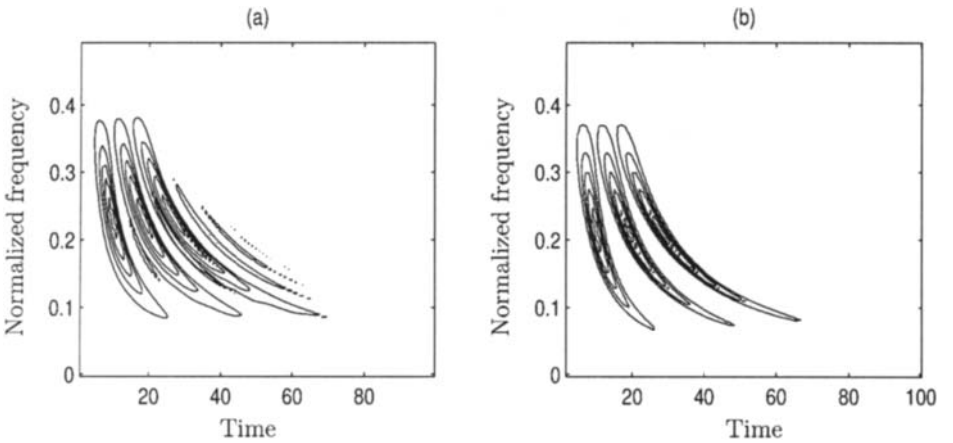


Fig. 9.2.1: (a) Narrowband Weyl symbol, and (b) hyperbolic Weyl symbol of a random process consisting of signal components with random amplitude and hyperbolic instantaneous frequency.

deterministic signal components are weighted by random, zero mean amplitudes, i.e. $x(t) = \left[\sum_{m=1}^{3} \mathbf{a}_m s_m(t)\right] w(t)$ where $w(t)$ is a shifted Hanning window, and $\mathbf{a}_m$ are uncorrelated random variables with zero means and constant variances $\sigma^2_{\mathbf{a}_m}$. Also, $s_m(t) = \frac{1}{\sqrt{t}} e^{j4\pi m \ln(\frac{t}{t_r})}$, $t > 0$, is a deterministic hyperbolic FM signal in time, with hyperbolic IF (see [9] and Article 5.6). For the ideal case without windowing, the theoretical HWS of the process $\tilde{x}(t) = \sum_{m=1}^{3} \mathbf{a}_m s_m(t)$ is

$$\mathrm{HWS}_{\mathcal{R}_{\tilde{x}}}(t,f) = \sum_{m=1}^{3} E\{|\mathbf{a}_m|^2\}\, \mathrm{HWS}_{\mathcal{R}_{s_m}}(t,f) = \frac{1}{t} \sum_{m=1}^{3} \sigma^2_{\mathbf{a}_m}\, \delta\left(f - \frac{2m}{t}\right), \quad t > 0\,.$$

It consists of three Dirac delta functions centered along the hyperbolae $f = 2m/t$, corresponding to the hyperbolic IF of $s_m(t)$, $m = 1, 2, 3$. Both the WS and the HWS produce no outer cross terms between neighboring hyperbolae since the random amplitudes $\mathbf{a}_m$ are uncorrelated. However, the WS in Fig. 9.2.1(a) has inner interference terms inside each hyperbola since it does not match the hyperbolic TF characteristics of the process. The HWS in Fig. 9.2.1(b) is an ideally concentrated TFR without any interference terms as it is matched to the process.

Fig. 9.2.2 demonstrates the advantage of the $\kappa = 3$ power WS over the narrowband WS when the analysis random process is a power FM signal $x(t) = \mathbf{a}\, \upsilon_\kappa(t)\, e^{j2\pi\beta\xi_\kappa(\frac{t}{t_r})}$ (also with $\kappa = 3$) [see [9] and Articles 5.6 & 15.3]. The power WS in Fig. 9.2.2(b) results in an ideal TFR for this process since, unlike the WS in Fig. 9.2.2(a), it matches the signal's power TF characteristics without any inner interference.

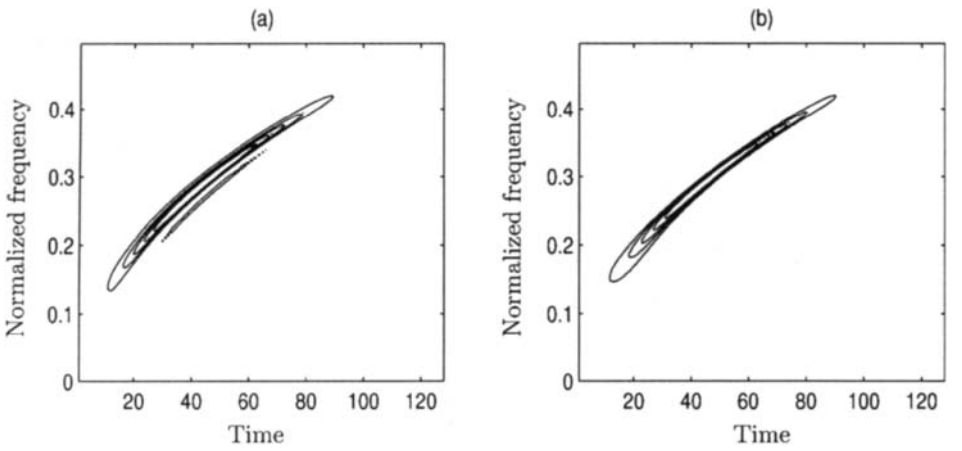


Fig. 9.2.2: (a) Narrowband Weyl symbol, and (b) power Weyl symbol of a random process consisting of a signal component with random amplitude and power instantaneous frequency.

9.2.5 Summary and Conclusions

This article discussed the use of TF symbols for the statistical signal processing of linear time-varying systems and nonstationary random processes. These symbols form the time-varying counterparts of transfer functions for systems and of power spectral densities for random processes. The Weyl symbol has been successfully used for narrowband systems with constant TF structures whereas the P_0-Weyl symbol has been used for wideband systems with either constant or hyperbolic TF characteristics. For systems with non-linear dispersive structures, new symbols were presented for appropriate matched processing. For example, such non-linear TF symbols would be suitable for analyzing tetherless underwater communication systems in dispersive mediums.

References

[1] W. Kozek, "On the transfer function calculus for underspread LTV channels," *IEEE Trans. Signal Processing*, vol. 45, pp. 219–223, January 1997.

[2] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 318–331, February 1994.

[3] B. G. Iem, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels, "New concepts in narrowband and wideband Weyl correspondence time-frequency techniques," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98)*, vol. 3, pp. 1573–1576, Seattle, 12–15 May 1998.

[4] B. G. Iem, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels, "Classes of smoothed Weyl symbols," *IEEE Signal Processing Letters*, vol. 7, pp. 186–188, July 2000.

[5] B. G. Iem, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels, "A wideband time-frequency Weyl symbol and its generalization," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 29–32, Pittsburgh, PA, 6–9 October 1998.

[6] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Communications*, vol. 47, pp. 123–132, January 1999.

[7] P. Flandrin, "A time-frequency formulation of optimum detection," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 36, pp. 1377–1384, September 1988.

[8] G. Matz and F. Hlawatsch, "Time-frequency formulation and design of optimal detectors," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 213–216, Paris, 18–21 June 1996.

[9] A. Papandreou-Suppappola, R. L. Murray, B. G. Iem, and G. F. Boudreaux-Bartels, "Group delay shift covariant quadratic time-frequency representations," *IEEE Trans. Signal Processing*, vol. 49, pp. 2549–2564, November 2001.

[10] 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. Grossmann, and P. Tchamitchian, eds.), pp. 147–153, Springer, 1989. Proc. of the Internat. Conf., Marseille, 14–18 December, 1987. 315pp.

[11] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: A new twist on signal processing," *IEEE Trans. Signal Processing*, vol. 43, pp. 2269–2282, October 1995.

9.3 ROBUST TIME-FREQUENCY DISTRIBUTIONS [0]

Study of the additive Gaussian noise influence on time-frequency (TF) distributions is an important issue (see Article 9.1). However, in many practical applications, especially in communications, signals are disturbed by a kind of impulse noise. These noises are commonly modeled by heavy-tailed (long-tailed) probability density functions (pdfs) [1]. It is well known that the conventional TF distributions are quite sensitive to this kind of noise, which is able to destroy sensitive signal information. The minimax Huber M-estimates [2] can be applied in order to design the periodogram and TF distributions, robust with respect to the impulse noise. For nonstationary signals, the robust TF distributions are developed as an extension of the robust M-estimation approach.

9.3.1 Robust Spectrogram

The standard short-time Fourier transform (STFT) at a given point (t, f), introduced in Section 2.3.1, can be alternatively defined as a solution of the following optimization problem:

$$F_z(t,f) = \arg\left\{\min_m I(t,f,m)\right\}, \tag{9.3.1}$$

$$I(t,f,m) = \sum_{n=-N/2}^{N/2-1} w(n\Delta t)\mathbf{F}(e(t,f,n)). \tag{9.3.2}$$

Here, the loss function is given as $\mathbf{F}(e) = |e|^2$, $w(n\Delta t)$ is a window function and Δt is a sampling interval. The error function has the form:

$$e(t,f,n) = z(t+n\Delta t)e^{-j2\pi fn\Delta t} - m, \tag{9.3.3}$$

where m is a complex-valued optimization parameter in (9.3.1). The error function can be considered as a residuum expressing the "similarity" between the signal and a given harmonic $\exp(j2\pi fn\Delta t)$.

The solution of (9.3.1) easily follows from

$$\frac{\partial I(t,f,m)}{\partial m^*} = 0 \tag{9.3.4}$$

in the form of the well-known standard STFT (cf. Section 6.5.5):

$$F_z(t,f) = \frac{1}{a_w}\sum_{n=-N/2}^{N/2-1} w(n\Delta t)z(t+n\Delta t)e^{-j2\pi fn\Delta t}, \tag{9.3.5}$$

[0]Authors: **Vladimir Katkovnik**, Department of Mechatronics, Kwangju Institute of Science and Technology (katkov@kjist.ac.kr); **Igor Djurović**, Department of Mechanical and System Engineering, Kyoto institute of Technology, Kyoto 606-8585, Japan (igordj@server1.cis.cg.ac.yu), on leave from the University of Montenegro; **LJubiša Stanković**, Elektrotehnicki fakultet, University of Montenegro, 81000 Podgorica, Montenegro (l.stankovic@ieee.org). Reviewer: S. Stanković.

where

$$a_w = \sum_{n=-N/2}^{N/2-1} w(n\Delta t). \tag{9.3.6}$$

The corresponding spectrogram is defined by

$$S_z(t,f) = |F_z(t,f)|^2. \tag{9.3.7}$$

The maximum likelihood (ML) approach can be used for selection of the appropriate loss function $\mathbf{F}(e)$ if the pdf $p(e)$ of the noise is known. The ML approach suggests the loss function $\mathbf{F}(e) \sim -\log p(e)$. For example, the loss function $\mathbf{F}(e) = |e|^2$ gives the standard STFT, as the ML estimate of spectra for signals corrupted with the Gaussian noise, $p(e) \sim \exp(-|e|^2)$. The standard STFT produces poor results for signals corrupted by impulse noise. Additionally, in many cases the ML estimates are quite sensitive to deviations from the parametric model and the hypothetical distribution. Even a slight deviation from the hypothesis can result in a strong degradation of the ML estimate. The minimax robust approach has been developed in statistics as an alternative to the conventional ML in order to decrease the ML estimates sensitivity, and to improve the efficiency in an environment with the heavy-tailed pdfs. The loss function

$$\mathbf{F}(e) = |e| = \sqrt{\Re^2\{e\} + \Im^2\{e\}} \tag{9.3.8}$$

is recommended by the robust estimation theory for a wide class of heavy-tailed pdfs. It is worth noting that the loss function

$$\mathbf{F}(e) = |\Re\{e\}| + |\Im\{e\}| \tag{9.3.9}$$

is the ML selection for the Laplacian distribution of independent real and imaginary parts of the complex valued noise.

Nonquadratic loss functions in (9.3.1) can improve filtering properties for impulse noises. Namely, in [3,4] it is proved that there is a natural link between the problem of spectra resistance to the impulse noise and the minimax Huber's estimation theory. It has been shown that the loss function derived in this theory could be applied to the design of a new class of robust spectra, inheriting properties of strong resistance to impulse noises.

In particular, the robust M-STFT has been derived by using the absolute error loss function $\mathbf{F}(e) = |e|$ in (9.3.1)-(9.3.4) [3]. It is a solution of the nonlinear equation:

$$F_z(t,f) = \frac{1}{a_w(t,f)} \sum_{n=-N/2}^{N/2-1} d(t,f,n) z(t+n\Delta t) e^{-j2\pi f n\Delta t}, \tag{9.3.10}$$

where:

$$d(t,f,n) = \frac{w(n\Delta t)}{|z(t+n\Delta t)e^{-j2\pi f n\Delta t} - F_z(t,f)|}, \tag{9.3.11}$$

and

$$a_w(t,f) = \sum_{n=-N/2}^{N/2-1} d(t,f,n). \tag{9.3.12}$$

If real and imaginary parts of the additive noise are independent, the statistically optimal robust estimation theory requires replacement of (9.3.1) with [4]:

$$F_z(t,f) = \arg\left\{\min_m I_1(t,f,m)\right\}, \tag{9.3.13}$$

$$I_1(t,f,m) = \sum_{n=-N/2}^{N/2-1} w(n\Delta t)[\mathbf{F}(\Re\{e_1\}) + \mathbf{F}(\Im\{e_1\})] \tag{9.3.14}$$

where e_1 is an error function of the form

$$e_1(t,f,n) = z(t+n\Delta t) - me^{j2\pi fn\Delta t}. \tag{9.3.15}$$

For $\mathbf{F}(e) = |e|$, the robust STFT (9.3.13) can be presented as a solution of (9.3.10), where $d(t,f,n)$ is given by:

$$d(t,f,n) = w(n\Delta t)\frac{|\Re\{e_1(t,f,n)\}| + |\Im\{e_1(t,f,n)\}|}{|\Re\{e_1(t,f,n)\}|^2 + |\Im\{e_1(t,f,n)\}|^2}. \tag{9.3.16}$$

The robust spectrogram defined in the form

$$S_z(t,f) = I_1(t,f,0) - I_1(t,f,F_z(t,f)) \tag{9.3.17}$$

is called the residual spectrogram, in order to distinguish it from the amplitude spectrogram (9.3.7). For the quadratic loss function $\mathbf{F}(e)$ the residual spectrogram (9.3.17) coincides with the standard amplitude spectrogram (9.3.7). In [4] it has been shown that, in a heavy-tailed noise environment, the residual robust spectrogram performs better than its amplitude counterpart.

The accuracy analysis of the robust spectrograms, as well as a discussion on further details on the minimax approach, can be found in [3,4].

9.3.2 Realization of the Robust STFT

9.3.2.1 Iterative Procedure

The expression (9.3.10) includes $F_z(t,f)$ on the right hand side. Therefore, to get the robust STFT we have to solve a nonlinear equation of the form $x = f(x)$. Here, we will use the fixed point iterative algorithm $x_i = f(x_{i-1})$, with the stopping rule $|x_i - x_{i-1}|/|x_i| < \eta$, where η defines the solution precision. This procedure, applied to (9.3.10), can be summarized as follows.

Step (0): Calculate the standard STFT (9.3.5): $F_z^{(0)}(t,f) = F_z(t,f)$, and $i = 0$.

Step (i): Set $i = i + 1$. Calculate $d^{(i)}(t, f, n)$ for $F_z^{(i-1)}(t, f)$ determined from (9.3.11) or (9.3.16). Calculate $F_z^{(i)}(t, f)$ as:

$$F_z^{(i)}(t,f) = \frac{1}{\sum_{n=-N/2}^{N/2-1} d^{(i)}(t,f,n)} \sum_{n=-N/2}^{N/2-1} d^{(i)}(t,f,n) z(t+n\Delta t) e^{-j2\pi f n \Delta t}. \quad (9.3.18)$$

Step (ii): If the relative absolute difference between two iterations is smaller than η:

$$\frac{|F_z^{(i)}(t,f) - F_z^{(i-1)}(t,f)|}{|F_z^{(i)}(t,f)|} \leq \eta, \quad (9.3.19)$$

then the robust STFT is obtained as $F_z(t, f) = F_z^{(i)}(t, f)$.

9.3.2.2 Vector Filter Approach

Note that the standard STFT (9.3.5) can be treated as an estimate of the mean, calculated over a set of complex-valued observations:

$$\mathbf{E}^{(t,f)} = \{z(t + n\Delta t)e^{-j2\pi f n \Delta t} : n \in [-N/2, N/2)\}. \quad (9.3.20)$$

If we restrict possible values of m in (9.3.1) to the set $\mathbf{E}^{(t,f)}$, the vector filter concept [5–7] can be applied to get a simple approximation of the robust estimate of the STFT. Here, the coordinates of vector-valued variable are real and imaginary parts of $z(t+n\Delta t)e^{-j2\pi f n \Delta t}$. The vector estimate of the STFT is defined as $F_z(t, f) = m$, where $m \in \mathbf{E}^{(t,f)}$, and for all $k \in [-N/2, N/2)$ the following inequality holds:

$$\sum_{n=-N/2}^{N/2-1} \mathbf{F}(|m - z(t + n\Delta t)e^{-j2\pi f n \Delta t}|) \leq$$
$$\sum_{n=-N/2}^{N/2-1} \mathbf{F}(|z(t + k\Delta t)e^{-j2\pi f k \Delta t} - z(t + n\Delta t)e^{-j2\pi f n \Delta t}|). \quad (9.3.21)$$

For $\mathbf{F}(e) = |e|$ this estimate is called the vector median.

The marginal median can be used for independent estimation of real and imaginary parts of $F_z(t, f)$. It results in

$$\begin{aligned} \Re\{F_z(t,f)\} &= \text{median}\{\Re\{z(t + n\Delta t)e^{-j2\pi f n \Delta t}\} : n \in [-N/2, N/2)\}, \\ \Im\{F_z(t,f)\} &= \text{median}\{\Im\{z(t + n\Delta t)e^{-j2\pi f n \Delta t}\} : n \in [-N/2, N/2)\} \end{aligned} \quad (9.3.22)$$

The separate estimation of the real and imaginary parts of $F_z(t, f)$ assumes independence of the real and imaginary parts of $z(t + n\Delta t)e^{-j2\pi f n \Delta t}$, what in general does not hold here. However, in numerous experiments the accuracy of the median estimates (9.3.21) and (9.3.22) is of the same order. A simplicity of calculation is the advantage of these median estimates over the iterative procedures.

9.3.3 Robust Wigner Distribution

The standard (windowed) Wigner distribution (WD) of a discrete-time signal is defined as

$$W_z(t,f) = \frac{1}{a_w} \sum_{n=-N/2}^{N/2} w(n\Delta t) z(t+n\Delta t) z^*(t-n\Delta t) e^{-j4\pi f n\Delta t}, \tag{9.3.23}$$

with the normalization factor

$$a_w = \sum_{n=-N/2}^{N/2} w(n\Delta t) \tag{9.3.24}$$

(cf. Article 6.1). It can be interpreted as a solution of the problem

$$W_z(t,f) = \arg\left\{ \min_m J(t,f,m) \right\}, \tag{9.3.25}$$

$$J(t,f,m) = \sum_{n=-N/2}^{N/2} w(n\Delta t) \mathbf{F}(|z(t+n\Delta t) z^*(t-n\Delta t) e^{-j4\pi f n\Delta t} - m|), \tag{9.3.26}$$

where $\mathbf{F}(e) = |e|^2$. For the loss function $\mathbf{F}(e) = |e|$, solution of (9.3.25)-(9.3.26) is a WD robust to the impulse noise. It can be obtained as a solution of the nonlinear equation [8]

$$W_z(t,f) = \frac{1}{a_{we}(t,f)} \sum_{n=-N/2}^{N/2} d(t,f,n) z(t+n\Delta t) z^*(t-n\Delta t) e^{-j4\pi f n\Delta t}, \tag{9.3.27}$$

with

$$\begin{aligned} d(t,f,n) &= \frac{w(n\Delta t)}{|z(t+n\Delta t) z^*(t-n\Delta t) e^{-j4\pi f n\Delta t} - W_z(t,f)|}, \\ a_{we}(t,f) &= \sum_{n=-N/2}^{N/2} d(t,f,n). \end{aligned} \tag{9.3.28}$$

An iterative procedure similar to the one described for the robust STFT can be used to find $W_z(t,f)$ from (9.3.27)-(9.3.28).

9.3.3.1 Properties of the Robust WD

(1) The robust WD is real-valued for real and symmetric window function:

$$W_z^*(t,f) = \frac{1}{a_{we}^*(t,f)} \sum_{n=-N/2}^{N/2} \frac{w^*(n\Delta t) z^*(t+n\Delta t) z(t-n\Delta t) e^{j4\pi f n\Delta t}}{|z^*(t+n\Delta t) z(t-n\Delta t) e^{j4\pi f n\Delta t} - W_z^*(t,f)|} =$$

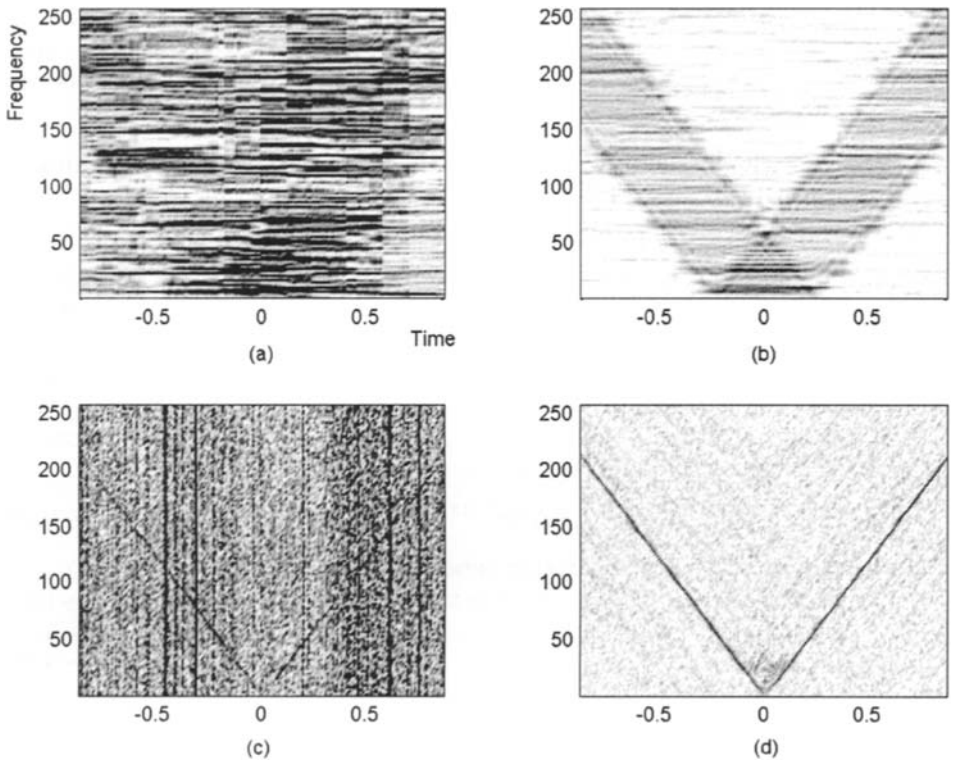


Fig. 9.3.1: TF representations of signal corrupted by impulse noise: (a) Standard spectrogram; (b) Robust spectrogram; (c) Standard WD; (d) Robust WD.

$$= \frac{1}{a_{we}(t,f)} \sum_{n=-N/2}^{N/2} \frac{w^*(-n\Delta t)z(t+n\Delta t)z^*(t-n\Delta t)e^{-j4\pi fn\Delta t}}{|z(t+n\Delta t)z^*(t-n\Delta t)e^{-j4\pi fn\Delta t} - W_z^*(t,f)|} = W_z(t,f). \tag{9.3.29}$$

(2) The robust WD is TF invariant. For signal $y(t) = z(t-t_0)e^{j2\pi f_0 t}$, we get $W_y(t,f) = W_z(t-t_0, f-f_0)$.

(3) For linear FM signals $z(t) = \exp(jat^2/2 + jbt)$, when $w(n\Delta t)$ is very wide window, the WD is an almost ideally concentrated TF distribution.

9.3.3.2 Median WD

For rectangular window, the standard WD can be treated as an estimate of the mean, calculated over a set of complex-valued observations

$$\mathbf{G} = \{z(t+n\Delta t)z^*(t-n\Delta t)e^{-j4\pi fn\Delta t} : n \in [-N/2, N/2]\}, \tag{9.3.30}$$

i.e.,

$$W_z(t,f) = \frac{1}{N+1} \sum_{n=-N/2}^{N/2} z(t+n\Delta t)z^*(t-n\Delta t)e^{-j4\pi fn\Delta t}. \tag{9.3.31}$$

From (9.3.29) follows that the robust WD is real-valued, thus the minimization of $J(t,f,m)$ can be done with respect to the real part of $z(t+n\Delta t)z^*(t-n\Delta t)e^{-j4\pi fn\Delta t}$ only. A form of the robust WD, the median WD, can be introduced as:

$$W_z(t,f) = \text{median}\{\Re\{z(t+n\Delta t)z^*(t-n\Delta t)e^{-j4\pi fn\Delta t}\}: n \in [-N/2, N/2]\}. \tag{9.3.32}$$

Generally, it can be shown that any robust TF distribution, obtained by using the Hermitian local auto-correlation function (LAF), $R_z(t,n\Delta t) = R_z^*(t,-n\Delta t)$ in the minimization, is real-valued. In the WD case this condition is satisfied, since $R_z(t,n\Delta t) = z(t+n\Delta t)z^*(t-n\Delta t)$. For a general quadratic distribution from the Cohen class with a Hermitian LAF, the proposed robust version reads

$$\rho_z(t,f) = \text{median}\{\Re\{R_z(t,n\Delta t)e^{-j4\pi fn\Delta t}\}: n \in [-N/2, N/2]\}, \tag{9.3.33}$$

where $R_z(t,n\Delta t)$ includes the kernel in time-lag domain.

Note, that for an input Gaussian noise the resulting noise in the WD has both Gaussian and impulse component, due to the WD's quadratic nature. Thus, as it is shown in [9], robust WD forms can improve performance of the standard WD, even in a high Gaussian input noise environment.

9.3.4 Example

Consider the nonstationary FM signal:

$$z(t) = \exp(j204.8\pi t|t|), \tag{9.3.34}$$

corrupted with a high amount of the heavy-tailed noise:

$$\varepsilon(t) = 0.5(\varepsilon_1^3(t) + j\varepsilon_2^3(t)), \tag{9.3.35}$$

where $\varepsilon_i(t)$, $i = 1,2$ are mutually independent Gaussian white noises $\mathcal{N}(0,1)$. We consider the interval $t \in [-7/8, 7/8]$ with a sampling rate $\Delta t = 1/512$ for spectrograms, and $\Delta t = 1/1024$ for WDs. The rectangular window width is $N = 256$ in all cases. The standard spectrogram and the WD (Figs. 9.3.1(a),(c)) are calculated according to (9.3.5) and (9.3.23). The robust spectrogram (Fig. 9.3.1(b)) is calculated by using iterative procedure (9.3.18)-(9.3.19). In this case, similar results would be produced by residual spectrogram (9.3.13)-(9.3.17), vector median (9.3.21), and marginal median (9.3.22). The robust WD (Fig. 9.3.1(d)) is calculated by using expression (9.3.32) for the considered TF point. It can be concluded from Fig. 9.3.1 that the robust spectrogram and the robust WD filter the heavy-tailed noise significantly better than the standard spectrogram and the standard WD. Note that the standard and the robust WD exhibit higher TF resolution in comparison with the corresponding spectrograms.

9.3.5 Summary and Conclusions

The TF distributions are defined within the Huber robust statistics framework. The loss function $\mathbf{F}(e) = |e|$ gives distributions robust to the impulse noise influence. They can be realized by using: the iterative procedures, the vector median, or the marginal median approach. All calculation procedures produce accuracy of the same order of magnitude.

Articles 6.1 and 6.5 further discuss the definition and computation of discrete-time TFDs.

References

[1] S. A. Kassam, *Signal detection in non-Gaussian noise.* Springer, 1988.

[2] P. J. Huber, *Robust statistics.* Wiley, 1981.

[3] V. Katkovnik, "Robust M-periodogram," *IEEE Trans. Signal Processing*, vol. 46, pp. 3104–3109, November 1998.

[4] V. Katkovnik, "Robust M-estimates of the frequency and amplitude of a complex-valued harmonic," *Signal Processing*, vol. 77, pp. 71–84, August 1999.

[5] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filters," *Proc. IEEE*, vol. 78, pp. 678–689, April 1990.

[6] I. Djurović, V. Katkovnik, and L. Stanković, "Median filter based realizations of the robust time-frequency distributions," *Signal Processing*, vol. 81, pp. 1771–1776, August 2001.

[7] I. Pitas and A. N. Venetsanopoulos, *Nonlinear digital filters: Principles and applications.* Kluwer, 1990.

[8] I. Djurović and L. Stanković, "Robust Wigner distribution with application to the instantaneous frequency estimation," *IEEE Trans. Signal Processing*, vol. 49, pp. 2985–2993, December 2001.

[9] I. Djurović, L. Stanković, and J. F. Böhme, "Estimates of the Wigner distribution in Gaussian noise environment," *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, vol. 56, no. 5, pp. 337–340, 2002.

9.4 TIME-VARYING POWER SPECTRA OF NONSTATIONARY RANDOM PROCESSES [0]

9.4.1 Nonstationary Random Processes

The second-order statistics of a (generally nonstationary) random process[1] $x(t)$ are characterized by the correlation function $r_x(t,t') = \mathrm{E}\{x(t)\,x^*(t')\}$ (with $\mathrm{E}\{\cdot\}$ denoting expectation). In the special case of a (wide-sense) *stationary* random process, the correlation function is of the form $r_x(t,t') = \tilde{r}_x(t-t')$ and the Fourier transform of $\tilde{r}_x(\tau)$,

$$P_x(f) = \int_{-\infty}^{\infty} \tilde{r}_x(\tau)\, e^{-j2\pi f\tau}\, d\tau \;\geq\; 0\,, \tag{9.4.1}$$

is known as the *power spectral density* (PSD) [1]. The PSD describes the distribution of the process' mean power over frequency f and is extremely useful in statistical signal processing. The time-frequency dual of stationary processes is given by *white* processes with correlation functions of the form $r_x(t,t') = q_x(t)\,\delta(t-t')$. Here, the *mean instantaneous intensity* $q_x(t) \geq 0$ is the time-frequency dual of the PSD.

In many applications, the random signals under analysis are nonstationary and thus do not possess a PSD. Various extensions of the PSD to the nonstationary case have been proposed, such as the *generalized Wigner-Ville spectrum* [2–7] and the *generalized evolutionary spectrum* [7,8]. In this article, we will briefly discuss these "time-varying power spectra" and show that they yield satisfactory descriptions for the important class of *underspread* nonstationary processes.

9.4.2 The Generalized Wigner-Ville Spectrum

The *generalized Wigner-Ville spectrum* (GWVS) of a nonstationary process $x(t)$ is defined as [2–7]

$$\overline{W}_x^{(\alpha)}(t,f) \;\triangleq\; \int_{-\infty}^{\infty} r_x^{(\alpha)}(t,\tau)\, e^{-j2\pi f\tau} d\tau$$

with

$$r_x^{(\alpha)}(t,\tau) \;\triangleq\; r_x\Big(t + \Big(\frac{1}{2} - \alpha\Big)\tau,\, t - \Big(\frac{1}{2} + \alpha\Big)\tau\Big)\,, \tag{9.4.2}$$

where α is a real-valued parameter. The GWVS equals the generalized Weyl symbol (see Article 4.7) of the correlation operator $\mathbf{R}_x$ (the linear operator whose kernel is the correlation function $r_x(t,t') = \mathrm{E}\{x(t)\,x^*(t')\}$) and, under mild assumptions, it equals the expectation of the generalized Wigner distribution [5] of $x(t)$. For $\alpha = 0$,

---

[0] Authors: **G. Matz** and **F. Hlawatsch**, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Gusshausstrasse 25/389, A-1040 Vienna, Austria (email: g.matz@ieee.org, fhlawats@pop.tuwien.ac.at, web: http://www.nt.tuwien.ac.at/dspgroup/time.html). Reviewers: M. Amin and A. Papandreou-Suppappola.

[1] In what follows, all random processes are assumed to be real or circular complex as well as zero-mean.

the GWVS becomes the ordinary *Wigner-Ville spectrum*, and for $\alpha = 1/2$ it reduces to the *Rihaczek spectrum* [2–7, 9]:

$$\overline{W}_x^{(0)}(t,f) = \int_{-\infty}^{\infty} r_x\Big(t+\frac{\tau}{2}, t-\frac{\tau}{2}\Big)\, e^{-j2\pi f\tau}\, d\tau\,,$$

$$\overline{W}_x^{(1/2)}(t,f) = \int_{-\infty}^{\infty} r_x(t,t-\tau)\, e^{-j2\pi f\tau}\, d\tau\,.$$

The GWVS $\overline{W}_x^{(\alpha)}(t,f)$ is a complete characterization of the second-order statistics of $x(t)$ since the correlation function $r_x(t,t')$ can be recovered from it. Integration of the GWVS gives the marginal properties

$$\int_{-\infty}^{\infty} \overline{W}_x^{(\alpha)}(t,f)\, dt = \mathrm{E}\{|X(f)|^2\}\,, \qquad \int_{-\infty}^{\infty} \overline{W}_x^{(\alpha)}(t,f)\, df = \mathrm{E}\{|x(t)|^2\}\,,$$

provided that the expectations on the right-hand sides exist. In this sense, the GWVS can be considered as a time-frequency (TF) distribution of the mean energy of $x(t)$. However, in general the GWVS is not real-valued; for $\alpha = 0$, $\overline{W}_x^{(0)}(t,f)$ is real-valued though possibly not everywhere nonnegative. For further interesting properties of the GWVS, see [2–7, 9].

We next discuss the GWVS of three fundamental types of processes.

- The GWVS of a stationary process with correlation function $r_x(t,t') = \tilde{r}_x(t-t')$ reduces to the PSD $P_x(f)$ for all t, i.e., $\overline{W}_x^{(\alpha)}(t,f) \equiv P_x(f)$.
- The GWVS of a (generally nonstationary) white process with correlation function $r_x(t,t') = q_x(t)\delta(t-t')$ reduces to the mean instantaneous intensity $q_x(t)$ for all f, i.e., $\overline{W}_x^{(\alpha)}(t,f) \equiv q_x(t)$.
- The GWVS of a stationary white process with correlation function $r_x(t,t') = \eta\,\delta(t-t')$ is given by $\overline{W}_x^{(\alpha)}(t,f) \equiv \eta$ (i.e., constant mean energy distribution over the entire TF plane).

These results show that the GWVS is consistent with the PSD of stationary processes and the mean instantaneous intensity of white processes.

The GWVS will be further considered in Section 9.4.6. Before that, we consider an alternative definition of a "time-varying power spectrum" in the next section.

9.4.3 The Generalized Evolutionary Spectrum

The PSD of a stationary random process $x(t)$ can alternatively be defined using an innovations system representation. Here, $x(t)$ is viewed as the output of a linear, time-invariant system $\mathbf{H}$ with impulse response $h(\tau)$ (the *innovations system*) that is driven by stationary white noise $n(t)$ with PSD $P_n(f) \equiv 1$, i.e., $x(t) = (\mathbf{H}n)(t) = \int_{-\infty}^{\infty} h(\tau)\, n(t-\tau)\, d\tau$. The PSD of $x(t)$ can then be written as

$$P_x(f) = |H(f)|^2, \tag{9.4.3}$$

where $H(f) = \int_{-\infty}^{\infty} h(\tau)\, e^{-j2\pi f\tau}\, d\tau$ is the transfer function (frequency response) of $\mathbf{H}$.

A similar innovations system representation is also possible in the nonstationary case. The innovations system $\mathbf{H}$ of a nonstationary random process $x(t)$ is a linear, *time-varying* system defined by $\mathbf{H}\mathbf{H}^{+} = \mathbf{R}_x$ (here, the superscript $^{+}$ denotes the adjoint [10]). Note that $\mathbf{H}$ is not uniquely defined; indeed, all innovations systems can be written as $\mathbf{H} = \mathbf{H}_p\mathbf{U}$ where $\mathbf{H}_p$ is the *positive (semi-) definite* [10] innovations system (which is unique) and $\mathbf{U}$ is a linear operator satisfying $\mathbf{U}\mathbf{U}^{+} = \mathbf{I}$ [8].

In analogy to the PSD expression in (9.4.3), the *generalized evolutionary spectrum* (GES) of a nonstationary process $x(t)$ is now defined as [7,8]

$$G_x^{(\alpha)}(t,f) \triangleq |L_{\mathbf{H}}^{(\alpha)}(t,f)|^2. \tag{9.4.4}$$

Here, $L_{\mathbf{H}}^{(\alpha)}(t,f)$ is the generalized Weyl symbol (see Article 4.7) of an innovations system $\mathbf{H}$ of $x(t)$, i.e.,

$$L_{\mathbf{H}}^{(\alpha)}(t,f) \triangleq \int_{-\infty}^{\infty} h\Big(t + \Big(\frac{1}{2} - \alpha\Big)\tau, t - \Big(\frac{1}{2} + \alpha\Big)\tau\Big)\, e^{-j2\pi f\tau}\, d\tau \tag{9.4.5}$$

where $h(t,t')$ is the kernel of $\mathbf{H}$. Note that the nonuniqueness of $\mathbf{H}$ implies a corresponding nonuniqueness of the GES. For $\alpha = 1/2$, $\alpha = -1/2$, and $\alpha = 0$, the GES reduces to the ordinary *evolutionary spectrum*[2] [11,12], the *transitory evolutionary spectrum* [8,13], and the *Weyl spectrum* [8], respectively.

In contrast to the GWVS, the GES is a nonnegative real-valued function. However, it is not a complete second-order description of $x(t)$ since in general the correlation function $r_x(t,t')$ cannot be recovered from it. For $\alpha = \pm 1/2$ and normal innovations system (i.e., $\mathbf{H}$ satisfies $\mathbf{H}\mathbf{H}^{+} = \mathbf{H}^{+}\mathbf{H}$ [10]; note, in particular, that $\mathbf{H}_p$ is always normal), the GES satisfies the marginal properties, i.e.,

$$\int_{-\infty}^{\infty} G_x^{(\pm 1/2)}(t,f)\, dt = \mathrm{E}\{|X(f)|^2\}, \qquad \int_{-\infty}^{\infty} G_x^{(\pm 1/2)}(t,f)\, df = \mathrm{E}\{|x(t)|^2\}.$$

Other properties of the GES are discussed in [7,8].

Next, we consider the GES of our three fundamental types of processes, assuming that the positive (semi-) definite innovations system $\mathbf{H}_p$ is used in the GES definition (9.4.4).

- For a stationary process with PSD $P_x(f)$, $\mathbf{H}_p$ is time-invariant with frequency response $H_p(f) = \sqrt{P_x(f)}$. Here, the GES reduces to the PSD $P_x(f)$ for all t, i.e., $G_x^{(\alpha)}(t,f) \equiv P_x(f)$.
- For a (generally nonstationary) white process with mean instantaneous intensity $q_x(t)$, $\mathbf{H}_p$ is "frequency-invariant" with kernel $h_p(t,t') = \sqrt{q_x(t)}\,\delta(t-t')$. The GES here reduces to $q_x(t)$ for all f, i.e., $G_x^{(\alpha)}(t,f) \equiv q_x(t)$.

[2] We note that Priestley's original definition of the evolutionary spectrum was based on a conceptually different approach using "oscillatory processes" [11,12].

- For a stationary and white process with correlation function $r_x(t,t') = \eta\,\delta(t-t')$, we have $\mathbf{H}_p = \sqrt{\eta}\,\mathbf{I}$ with $\mathbf{I}$ the identity operator. Thus, the GES is given by $G_x^{(\alpha)}(t,f) \equiv \eta$.

In Section 9.4.5, we shall consider conditions allowing the interpretation of the GWVS and GES as a "time-varying power spectrum." The formulation of these conditions will be based on a further TF representation of nonstationary processes, to be discussed next.

9.4.4 The Generalized Expected Ambiguity Function

The *generalized expected ambiguity function* (GEAF) is defined as [6–8]

$$\bar{A}_x^{(\alpha)}(\nu,\tau) \triangleq \int_{-\infty}^{\infty} r_x^{(\alpha)}(t,\tau)\, e^{-j2\pi\nu t}\, dt\,,$$

with $r_x^{(\alpha)}(t,\tau)$ as in (9.4.2). The interpretation of the GEAF is quite different from that of a "time-varying power spectrum:" For a given frequency lag ν and a given time lag τ, the GEAF $\bar{A}_x^{(\alpha)}(\nu,\tau)$ quantifies the statistical correlations of all process components separated in frequency by ν and in time by τ [6]. Hence, the extension of $\bar{A}_x^{(\alpha)}(\nu,\tau)$ about the origin of the (ν,τ) plane indicates the amount of "TF correlations" of $x(t)$. In particular, if $\bar{A}_x^{(\alpha)}(\nu,\tau)$ extends far in the ν direction, this indicates that $x(t)$ has a large spectral correlation width (i.e., $x(t)$ is highly nonstationary), and if $\bar{A}_x^{(\alpha)}(\nu,\tau)$ extends far in the τ direction, this indicates that $x(t)$ has a large temporal correlation width.

The GEAF equals the generalized spreading function (see Article 4.7) of the correlation operator $\mathbf{R}_x$. Like the GWVS, the GEAF is a complete second-order statistic. GEAFs with different α values differ merely by a phase factor, i.e.,

$$\bar{A}_x^{(\alpha_2)}(\nu,\tau) = \bar{A}_x^{(\alpha_1)}(\nu,\tau)\, e^{j2\pi(\alpha_1-\alpha_2)\nu\tau}\,.$$

Therefore, the GEAF magnitude is independent of α, $\left|\bar{A}_x^{(\alpha_1)}(\nu,\tau)\right| = \left|\bar{A}_x^{(\alpha_2)}(\nu,\tau)\right|$, and we may thus simply write $|\bar{A}_x(\tau,\nu)|$. GWVS and GEAF are related by a 2-D Fourier transform,

$$\overline{W}_x^{(\alpha)}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{A}_x^{(\alpha)}(\nu,\tau)\, e^{-j2\pi(\tau f-\nu t)}\, d\nu\, d\tau\,; \tag{9.4.6}$$

this extends the Wiener-Khintchine relation (9.4.1) to the nonstationary case.

Again, it is instructive to consider our three process types (see Fig. 9.4.1; this figure should be compared to that in Section 4.7.3):

- The GEAF of a stationary process $x(t)$ with correlation function $r_x(t,t') = \tilde{r}_x(t-t')$ is given by $\bar{A}_x^{(\alpha)}(\nu,\tau) = \delta(\nu)\,\tilde{r}_x(\tau)$ (i.e., only temporal correlations that are characterized by $\tilde{r}_x(\tau)$).

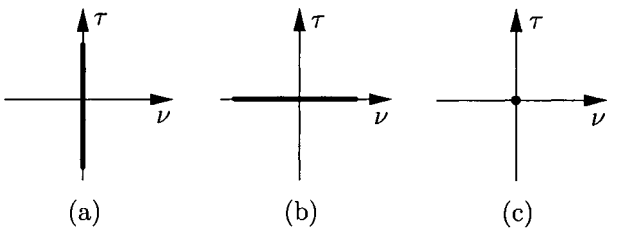


Fig. 9.4.1: Schematic representation of the GEAF magnitude of some (classes of) random processes: (a) stationary process, (b) white process, (c) stationary white process.

- The GEAF of a (generally nonstationary) white process $x(t)$ with correlation function $r_x(t,t') = q_x(t)\,\delta(t-t')$ is given by $\bar{A}_x^{(\alpha)}(\nu,\tau) = Q_x(\nu)\,\delta(\tau)$, where $Q_x(\nu)$ is the Fourier transform of the mean instantaneous intensity $q_x(t)$ (i.e., only spectral correlations that are characterized by $Q_x(\nu)$).

- The GEAF of a stationary and white process $x(t)$ with correlation function $r_x(t,t') = \eta\,\delta(t-t')$ is given by $\bar{A}_x^{(\alpha)}(\nu,\tau) = \eta\,\delta(\nu)\delta(\tau)$ (i.e., neither temporal nor spectral correlations).

9.4.5 Underspread Processes

A nonstationary random process is said to be *underspread* if its GEAF is well concentrated about the origin of the (ν,τ) plane, thus implying a small "TF correlation width." In contrast, a process with large TF correlation width is termed *overspread*. We will see in Section 9.4.6 that the GWVS and GES of an underspread process are approximately equivalent and can be interpreted as "time-varying power spectra."

There are two alternative mathematical definitions of underspread processes [6,7]. The first one [6] assumes that the GEAF $\bar{A}_x^{(\alpha)}(\nu,\tau)$ is supported in a compact region $\mathcal{G}_x$ about the origin of the (ν,τ) plane, i.e., $\left|\bar{A}_x(\nu,\tau)\right| = 0$ for $(\nu,\tau) \notin \mathcal{G}_x$. Let $\nu_x \triangleq \max_{(\nu,\tau)\in\mathcal{G}_x} |\nu|$ and $\tau_x \triangleq \max_{(\nu,\tau)\in\mathcal{G}_x} |\tau|$ denote the maximum frequency lag and time lag, respectively, for which the process $x(t)$ features TF correlations. The *TF correlation spread* of $x(t)$ is defined as $\sigma_x \triangleq 4\nu_x\tau_x$, which is the area of the rectangle $[-\nu_x,\nu_x] \times [-\tau_x,\tau_x]$ enclosing $\mathcal{G}_x$. The process $x(t)$ is considered underspread if $\sigma_x \ll 1$ [6].

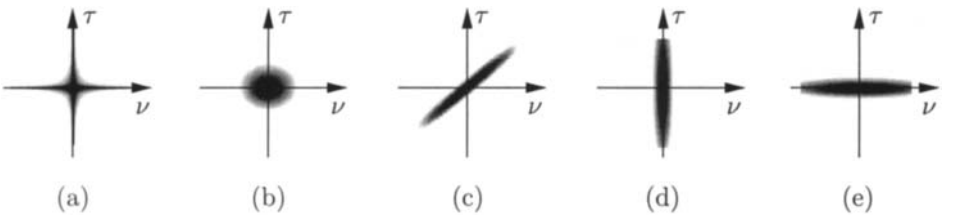


Fig. 9.4.2: Schematic representation of the GEAF magnitude of various types of nonstationary processes: (a) underspread process with small $m_x^{(1,1)}$; (b) underspread process with small $m_x^{(1,0)} m_x^{(0,1)}$; (c) "chirpy" underspread process [7]; (d) quasi-stationary process (small $m_x^{(0,1)}$); (e) quasi-white process (small $m_x^{(1,0)}$).

An alternative description of the GEAF's extension that avoids the assumption of compact GEAF support uses the normalized *weighted GEAF integrals*[3] [7]

$$m_x^{(\phi)} \triangleq \frac{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi(\nu,\tau)\,|\bar{A}_x(\nu,\tau)|\,d\nu\,d\tau}{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\bar{A}_x(\nu,\tau)|\,d\nu\,d\tau}\,.$$

Here, $\phi(\nu,\tau)$ is a nonnegative weighting function satisfying $\phi(\nu,\tau) \geq \phi(0,0) = 0$ and penalizing GEAF contributions located away from the origin. Important special cases are the *GEAF moments* $m_x^{(k,l)} \triangleq m_x^{(\phi_{k,l})}$ obtained with the weighting functions $\phi_{k,l}(\nu,\tau) = |\nu|^l|\tau|^k$ with $k,l \in \mathbb{N}_0$. A random process $x(t)$ can now be considered underspread if suitable weighted GEAF integrals or moments are "small." Processes that are underspread in the compact-support sense considered previously are easily shown to be a special case of this extended, more flexible definition of underspread processes.

Examples of underspread processes are illustrated in Fig. 9.4.2 (this figure should be compared to that in Section 4.7.4). We caution that the concept of underspread processes is not equivalent to that of quasi-stationary processes: indeed, a quasi-stationary process may be overspread if its temporal correlation width is very large. Finally, note that according to the Fourier transform relation (9.4.6), the GWVS of an underspread process is a *smooth* function.

9.4.6 Time-Varying Spectral Analysis of Underspread Processes

For *underspread* nonstationary processes, the GWVS and GES can be interpreted as "time-varying power spectra" that generalize the PSD of stationary processes and the mean instantaneous intensity of white processes. Indeed, small weighted GEAF integrals $m_x^{(\phi)}$ (or small moments $m_x^{(k,l)}$ or a small TF correlation spread σ_x) ensure the validity of the approximations described in what follows [6–8].

[3]Further definitions of weighted GEAF integrals and moments can be found in [7].

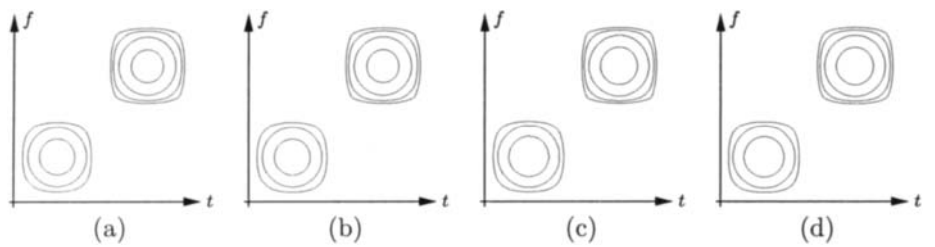


Fig. 9.4.3: GWVS and GES of an underspread process $x(t)$: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t,f)$, (b) real part of Rihaczek spectrum $\overline{W}_x^{(1/2)}(t,f)$, (c) Weyl spectrum $G_x^{(0)}(t,f)$, (d) evolutionary spectrum $G_x^{(1/2)}(t,f)$. In this and subsequent simulations, signal duration is 256 samples and normalized frequency ranges from $-1/4$ to $1/4$.

Approximate equivalence. In general, the GWVS and GES of a given process $x(t)$ may yield very different results which, moreover, may strongly depend on the parameter α used. However, for an underspread process $x(t)$, all these results are approximately equal, i.e.,

$$\begin{aligned}\overline{W}_x^{(\alpha_1)}(t,f) &\approx \overline{W}_x^{(\alpha_2)}(t,f)\,,\\ G_x^{(\alpha_1)}(t,f) &\approx G_x^{(\alpha_2)}(t,f)\,,\\ \overline{W}_x^{(\alpha_1)}(t,f) &\approx G_x^{(\alpha_2)}(t,f)\,.\end{aligned}$$

Indeed, it can be shown [7] that the approximation error $\overline{W}_x^{(\alpha_1)}(t,f) - \overline{W}_x^{(\alpha_2)}(t,f)$ is upper bounded as

$$\left|\overline{W}_x^{(\alpha_1)}(t,f) - \overline{W}_x^{(\alpha_2)}(t,f)\right| \;\leq\; 2\pi\,|\alpha_1-\alpha_2|\,\|\bar{A}_x\|_1\,m_x^{(1,1)}, \tag{9.4.7}$$

with $\|\bar{A}_x\|_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\bar{A}_x(\nu,\tau)|\,d\nu\,d\tau$. Thus, for an underspread system where $m_x^{(1,1)}$ is small, $\overline{W}_x^{(\alpha_1)}(t,f)$ and $\overline{W}_x^{(\alpha_2)}(t,f)$ will be approximately equal as long as $|\alpha_1 - \alpha_2|$ is not too large. Similar bounds can be developed for the approximation errors $G_x^{(\alpha_1)}(t,f) - G_x^{(\alpha_2)}(t,f)$ and $\overline{W}_x^{(\alpha_1)}(t,f) - G_x^{(\alpha_2)}(t,f)$ [7].

We can conclude from these results that for an underspread process, the choice of a specific spectrum is not critical. An example is shown in Fig. 9.4.3. For this example, the maximum normalized differences between the spectra shown are all around 0.03 (e.g. $\max_{t,f}\left|\overline{W}_x^{(0)}(t,f) - G_x^{(0)}(t,f)\right| / \max_{t,f}\left|\overline{W}_x^{(0)}(t,f)\right| = 0.029$). A counterexample involving an *overspread* process is shown in Fig. 9.4.4. Here, the results obtained with the various spectra are seen to be dramatically different, and indeed the maximum normalized differences range from 1 to 8.5 (e.g. $\max_{t,f}\left|\overline{W}_x^{(1/2)}(t,f) - G_x^{(1/2)}(t,f)\right| / \max_{t,f}\left|\overline{W}_x^{(1/2)}(t,f)\right| = 2.13$). It can be seen that all spectra contain oscillating components (so-called *statistical cross-terms*) which are indicative of TF correlations [7]. Such statistical cross-terms are reduced

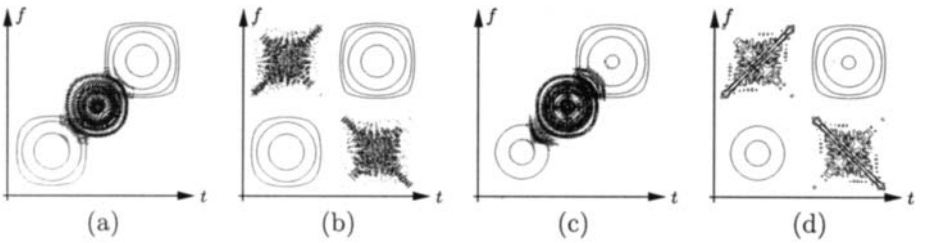


Fig. 9.4.4: GWVS and GES of an overspread process $x(t)$: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t,f)$, (b) real part of Rihaczek spectrum $\overline{W}_x^{(1/2)}(t,f)$, (c) Weyl spectrum $G_x^{(0)}(t,f)$, (d) evolutionary spectrum $G_x^{(1/2)}(t,f)$.

in extensions of the GWVS and GES that contain a TF smoothing [2–5, 7]. For underspread processes, on the other hand, a TF smoothing does not cause a big difference.

Approximate real-valuedness and positivity of the GWVS. The PSD of stationary processes and the mean instantaneous intensity of white processes are real-valued and nonnegative. This is also true for the GES of arbitrary processes. In contrast, the GWVS is real-valued only for $\alpha = 0$ and generally not everywhere nonnegative. In the case of underspread processes, however, it can be shown [7] that the imaginary part of the GWVS is approximately zero and the real part of the GWVS is approximately nonnegative, i.e.,

$$\Im\{\overline{W}_x^{(\alpha)}(t,f)\} \approx 0\,, \qquad \mathbf{R}e\{\overline{W}_x^{(\alpha)}(t,f)\} \gtrsim 0\,.$$

Upper bounds on the associated approximation errors (similar to (9.4.7)) can again be provided [7].

As an example, we reconsider the underspread process from Fig. 9.4.3. The normalized maximum of the imaginary part of the Rihaczek spectrum (the real part is shown in Fig. 9.4.3(b)) is $\max_{t,f} \left|\Im\{\overline{W}_x^{(1/2)}(t,f)\}\right| / \max_{t,f} \left|\overline{W}_x^{(1/2)}(t,f)\right| = 0.024$ and the normalized maximum of the negative real part is $\max_{t,f} \{ -\mathbf{R}e\{\overline{W}_x^{(1/2)}(t,f)\}\} / \max_{t,f} \left|\overline{W}_x^{(1/2)}(t,f)\right| = 0.006$.

Approximate input-output relations. If a stationary process $x(t)$ with PSD $P_x(f)$ is passed through a time-invariant linear system with impulse response $k(\tau)$ and transfer function $K(f)$, the output $y(t) = (x * k)(t)$ is also stationary and its PSD equals $P_y(f) = |K(f)|^2 P_x(f)$. Similarly, the response $y(t) = w(t)x(t)$ of a linear frequency-invariant system (see Article 4.7) to a white process $x(t)$ with mean instantaneous intensity $q_x(t)$ is again white with $q_y(t) = |w(t)|^2 q_x(t)$. A similar input-output relation does not exist for a general nonstationary process $x(t)$ that is passed through a general time-varying linear system $\mathbf{K}$. However, for an underspread process that is passed through an underspread system (i.e., a time-varying linear system introducing only small TF shifts, see Article 4.7), one can show the

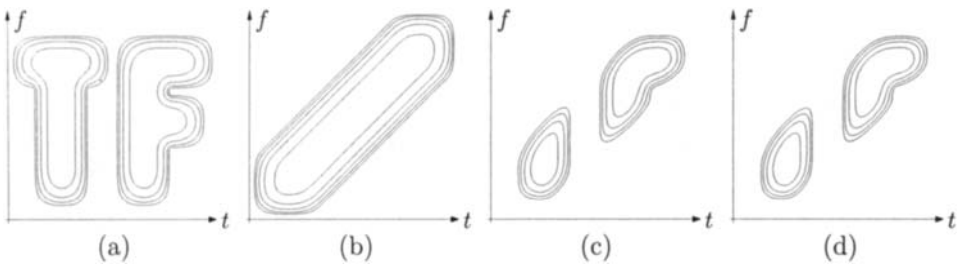


Fig. 9.4.5: Approximate input-output relation for the GWVS: (a) Wigner-Ville spectrum $\overline{W}_x^{(0)}(t,f)$ of input process $x(t)$, (b) Weyl symbol $L_{\mathbf{K}}^{(0)}(t,f)$ of LTV system $\mathbf{K}$, (c) Wigner-Ville spectrum $\overline{W}_y^{(0)}(t,f)$ of filtered process $y(t) = (\mathbf{K}x)(t)$, (d) approximation $|L_{\mathbf{K}}^{(0)}(t,f)|^2\, \overline{W}_x^{(0)}(t,f)$.

following approximate input-output relations of the GWVS and GES:

$$\overline{W}_y^{(\alpha)}(t,f) \approx \left|L_{\mathbf{K}}^{(\alpha)}(t,f)\right|^2 \overline{W}_x^{(\alpha)}(t,f)\,,$$
$$G_y^{(\alpha)}(t,f) \approx \left|L_{\mathbf{K}}^{(\alpha)}(t,f)\right|^2 G_x^{(\alpha)}(t,f)\,,$$

with $y(t) = (\mathbf{K}x)(t)$. Note that the generalized Weyl symbol $L_{\mathbf{K}}^{(\alpha)}(t,f)$ of $\mathbf{K}$ (see (9.4.5)) takes the place of the transfer function $K(f)$ or $w(t)$. An example for the Wigner-Ville spectrum (GWVS with $\alpha = 0$) is shown in Fig. 9.4.5. In this example, the normalized maximum approximation error is $\max_{t,f}\left\{\left|\overline{W}_y^{(0)}(t,f) - \left|L_{\mathbf{K}}^{(0)}(t,f)\right|^2 \overline{W}_x^{(0)}(t,f)\right|\right\} / \max_{t,f}\left|\overline{W}_y^{(0)}(t,f)\right| = 0.017$.

Discussion. The above approximations (more can be found in [6–8]) corroborate the interpretation of the GWVS and GES of underspread processes as *time-varying power spectra*. A mathematical underpinning of these approximations is provided by explicit upper bounds on the associated approximation errors [7]; these bounds involve the GEAF parameters $m_x^{(\phi)}$, $m_x^{(k,l)}$ or σ_x defined in Section 9.4.5. In the underspread case, these GEAF parameters are small and thus the approximations are guaranteed to be good. On the other hand, we caution that the approximations are *not* valid for overspread processes (cf. Fig. 9.4.4).

9.4.7 Summary and Conclusions

In this article, we have shown that for the practically important class of underspread processes (i.e., processes with small time-frequency correlations), the generalized Wigner-Ville spectrum and generalized evolutionary spectrum can be interpreted in a meaningful way as time-varying power spectra. Indeed, for underspread processes the generalized Wigner-Ville spectrum and the generalized evolutionary spectrum (approximately) satisfy desirable properties that any reasonable definition of a time-varying power spectrum would be expected to satisfy. We note that applications of the generalized Wigner-Ville spectrum in statistical signal processing are considered in Articles 12.1 and 12.4.

References

[1] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 3rd ed., 1991.

[2] 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.), ch. 9, pp. 208–232, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[3] P. Flandrin, "Time-dependent spectra for nonstationary stochastic processes," in *Time and Frequency Representation of Signals and Systems* (G. Longo and B. Picinbono, eds.), pp. 69–124, Vienna: Springer, 1989.

[4] 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.), pp. 211–267, Amsterdam: Elsevier, 1997.

[5] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[6] W. Kozek, F. Hlawatsch, H. Kirchauer, and U. Trautwein, "Correlative time-frequency analysis and classification of nonstationary random processes," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 417–420, Philadelphia, PA, 25–28 October 1994.

[7] G. Matz and F. Hlawatsch, "Time-varying spectra for underspread and overspread nonstationary processes," in *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers*, pp. 282–286, Pacific Grove, CA, 1–4 November 1998.

[8] G. Matz, F. Hlawatsch, and W. Kozek, "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes," *IEEE Trans. Signal Processing*, vol. 45, pp. 1520–1534, June 1997.

[9] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 33, pp. 1461–1470, December 1985.

[10] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science.* New York: Springer, 2nd ed., 1982.

[11] M. B. Priestly, "Evolutionary spectra and non-stationary processes," *J. Royal Statistical Soc. (Series B)*, vol. 27, no. 2, pp. 204–237, 1965.

[12] M. B. Priestley, *Spectral Analysis and Time Series—Part II.* London: Academic Press, 1981.

[13] C. S. Detka and A. El-Jaroudi, "The transitory evolutionary spectrum," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, vol. 4, pp. 289–292, Adelaide, Australia, 19–22 April 1994.

9.5 TIME-FREQUENCY CHARACTERIZATION OF RANDOM TIME-VARYING CHANNELS [0]

9.5.1 Time-Varying Channels

In many practical communication systems, the channel is modeled as linear but *time-varying* and *random*. Examples are the mobile radio, ionospheric, tropospheric, and underwater acoustic channels [1–7]. In this article, we will discuss time-frequency (TF) descriptions of both the channel (see also Article 4.7 for more details) and its second-order statistics.

The input-output relation of a linear, time-varying, random channel $\mathbf{H}$ is

$$r(t) = (\mathbf{H}s)(t) = \int_{-\infty}^{\infty} h(t,t')\, s(t')\, dt' = \int_{-\infty}^{\infty} \tilde{h}(t,\tau)\, s(t-\tau)\, d\tau\,, \qquad (9.5.1)$$

where $s(t)$ is the transmit signal, $r(t)$ is the received signal, $h(t,t')$ is the (random) kernel of $\mathbf{H}$, and $\tilde{h}(t,\tau) = h(t,t-\tau)$ is the (random) impulse response of $\mathbf{H}$. Two major physical phenomena underlying practical channels are *multipath propagation* and *Doppler spreading*. Multipath propagation (i.e., several different propagation paths from the transmitter to the receiver via various scattering objects) causes the received signal to consist of several delayed versions of the transmit signal. Doppler spreading is due to the movement of transmitter and/or receiver and/or scatterers; for a narrowband transmit signal $s(t)$, it causes the multipath signals to be frequency-shifted. The received signal $r(t)$ thus consists of several TF shifted (i.e., delayed and modulated) versions of the transmit signal $s(t)$ [1–4],

$$r(t) = \sum_{k=1}^{N} a_k\, s(t-\tau_k)\, e^{j2\pi\nu_k t}\,.$$

Here, N is the number of scatterers and τ_k, ν_k, and a_k are respectively the (random) delay, Doppler frequency, and reflectivity of the kth scatterer. The above relation can be extended to a continuum of scatterers (corresponding to a continuum of delays τ and Doppler frequency shifts ν) as

$$r(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)\, s_{\nu,\tau}^{(\alpha)}(t)\, d\tau\, d\nu\,, \qquad (9.5.2)$$

with $s_{\nu,\tau}^{(\alpha)}(t) = s(t-\tau)\, e^{j2\pi\nu t}\, e^{j2\pi\nu\tau(\alpha-1/2)}$, where α is a real-valued parameter[1] that is arbitrary but assumed fixed and $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)$ denotes the *generalized (delay-*

[0]Authors: **G. Matz** and **F. Hlawatsch**, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Gusshausstrasse 25/389, A-1040 Vienna, Austria (email: g.matz@ieee.org, fhlawats@pop.tuwien.ac.at, web: http://www.nt.tuwien.ac.at/dspgroup/time.html). Reviewers: P. Flandrin and A. Sayeed.

[1]The parameter α in $s_{\nu,\tau}^{(\alpha)}(t)$ corresponds to the infinitely many ways of defining a joint TF shift by combining time shifts and frequency shifts. In particular, the case $\alpha = 1/2$ corresponds to first shifting in time and then shifting in frequency, whereas $\alpha = -1/2$ corresponds to first shifting in frequency and then in time.

Doppler) spreading function (GSF) of the channel [1] (see also Article 4.7). The GSF is defined as

$$S_{\mathbf{H}}^{(\alpha)}(\nu,\tau) \triangleq \int_{-\infty}^{\infty} h\Big(t+\Big(\frac{1}{2}-\alpha\Big)\tau, t-\Big(\frac{1}{2}+\alpha\Big)\tau\Big)\, e^{-j2\pi\nu t}dt\,. \tag{9.5.3}$$

It can be shown that the input-output relation (9.5.2) is mathematically equivalent to (9.5.1).

In what follows, we will also use the *generalized Weyl symbol* (GWS)

$$L_{\mathbf{H}}^{(\alpha)}(t,f) \triangleq \int_{-\infty}^{\infty} h\Big(t+\Big(\frac{1}{2}-\alpha\Big)\tau, t-\Big(\frac{1}{2}+\alpha\Big)\tau\Big)\, e^{-j2\pi f\tau}d\tau \tag{9.5.4}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)\, e^{j2\pi(\nu t-\tau f)}\, d\tau\, d\nu\,. \tag{9.5.5}$$

The GWS can be interpreted (with certain precautions, see Article 4.7) as a "TF transfer function" of $\mathbf{H}$.

In the mobile communications literature, the parameter α is usually chosen as $1/2$. In this case, (9.5.3) and (9.5.4) become

$$S_{\mathbf{H}}^{(1/2)}(\nu,\tau) = \int_{-\infty}^{\infty} \tilde{h}(t,\tau)\, e^{-j2\pi\nu t}\, dt\,, \qquad L_{\mathbf{H}}^{(1/2)}(t,f) = \int_{-\infty}^{\infty} \tilde{h}(t,\tau)\, e^{-j2\pi f\tau}\, d\tau\,.$$

9.5.2 WSSUS Channels

Since the channel $\mathbf{H}$ is random, its GSF $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)$, GWS $L_{\mathbf{H}}^{(\alpha)}(t,f)$, and impulse response $\tilde{h}(t,\tau)$ are 2-D random functions (random processes). Hereafter, these random processes will be assumed zero-mean. The second-order statistics of $\mathbf{H}$ are characterized by the 4-D correlation functions $\mathrm{E}\{S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)\, S_{\mathbf{H}}^{(\alpha)*}(\nu',\tau')\}$, $\mathrm{E}\{L_{\mathbf{H}}^{(\alpha)}(t,f)\, L_{\mathbf{H}}^{(\alpha)*}(t',f')\}$, and $\mathrm{E}\{\tilde{h}(t,\tau)\,\tilde{h}^*(t',\tau')\}$, which are all mathematically equivalent.

Definition and description of WSSUS channels. An important simplification results from the assumption of *wide-sense stationary uncorrelated scattering (WSSUS)* [1,2,4,6,7]. For WSSUS channels, by definition, the reflectivities of scatterers corresponding to paths with different delay or Doppler are uncorrelated. In terms of the GSF $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)$, this means

$$\mathrm{E}\{S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)\, S_{\mathbf{H}}^{(\alpha)*}(\nu',\tau')\} = C_{\mathbf{H}}(\nu,\tau)\,\delta(\nu-\nu')\,\delta(\tau-\tau')\,, \tag{9.5.6}$$

i.e., $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)$ is a *wide-sense white* random process. The mean intensity function of this white random process, $C_{\mathbf{H}}(\nu,\tau) \geq 0$, is known as the *scattering function* [1–8].

Together with the Fourier transform relation (9.5.5), the WSSUS relation (9.5.6) implies that

$$\mathrm{E}\{L_{\mathbf{H}}^{(\alpha)}(t,f)\, L_{\mathbf{H}}^{(\alpha)*}(t',f')\} = R_{\mathbf{H}}(t-t', f-f')\,, \tag{9.5.7}$$

with the *TF correlation function* [1–7]

$$R_{\mathbf{H}}(\Delta t, \Delta f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu, \tau)\, e^{j2\pi(\nu\Delta t - \tau\Delta f)}\, d\tau\, d\nu\,. \tag{9.5.8}$$

The TF correlation function satisfies $\left|R_{\mathbf{H}}(\Delta t, \Delta f)\right| \leq R_{\mathbf{H}}(0,0)$ and $R^*_{\mathbf{H}}(-\Delta t, -\Delta f) = R_{\mathbf{H}}(\Delta t, \Delta f)$. Eq. (9.5.7) shows that the GWS $L^{(\alpha)}_{\mathbf{H}}(t,f)$ of a WSSUS channel is a 2-D *wide-sense stationary* process. According to (9.5.8), the scattering function $C_{\mathbf{H}}(\nu,\tau)$ is the Fourier transform of the correlation function $R_{\mathbf{H}}(\Delta t, \Delta f)$ of $L^{(\alpha)}_{\mathbf{H}}(t,f)$. Thus, $C_{\mathbf{H}}(\nu,\tau)$ can be interpreted as the *power spectral density* of $L^{(\alpha)}_{\mathbf{H}}(t,f)$. The *path loss* [3] is defined as

$$\rho^2_{\mathbf{H}} \triangleq R_{\mathbf{H}}(0,0) = \mathrm{E}\{\left|L^{(\alpha)}_{\mathbf{H}}(t,f)\right|^2\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu,\tau)\, d\tau\, d\nu\,.$$

Finally, in terms of the impulse response $\tilde{h}(t,\tau)$, the WSSUS property is expressed as

$$\mathrm{E}\{\tilde{h}(t,\tau)\, \tilde{h}^*(t',\tau')\} = r_{\mathbf{H}}(t-t',\tau)\, \delta(\tau-\tau')\,,$$

where $r_{\mathbf{H}}(\Delta t, \tau)$ is related by Fourier transforms to $C_{\mathbf{H}}(\nu,\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f)$. Thus, $\tilde{h}(t,\tau)$ is wide-sense stationary with respect to time t and uncorrelated for different delays τ. Note that this stationarity with respect to t refers to the second-order *statistics* of the channel and does not imply that the channel's realizations are time-invariant systems (cf. the first of the examples given below).

The 2-D functions $C_{\mathbf{H}}(\nu,\tau)$, $R_{\mathbf{H}}(\Delta t, \Delta f)$, and $r_{\mathbf{H}}(\Delta t, \tau)$ are *mathematically equivalent* descriptions of the second-order statistics of a WSSUS channel. They are related by Fourier transforms and do not dependent on α.

The composition (series connection) $\mathbf{H}_2\mathbf{H}_1$ of two statistically independent WSSUS channels $\mathbf{H}_1$ and $\mathbf{H}_2$ can be shown to be again a WSSUS channel. Its scattering function and TF correlation function are given by

$$\begin{aligned} C_{\mathbf{H}_2\mathbf{H}_1}(\nu,\tau) &= \left(C_{\mathbf{H}_2} ** C_{\mathbf{H}_1}\right)(\nu,\tau) \\ R_{\mathbf{H}_2\mathbf{H}_1}(\Delta t, \Delta f) &= R_{\mathbf{H}_2}(\Delta t, \Delta f)\, R_{\mathbf{H}_1}(\Delta t, \Delta f)\,, \end{aligned}$$

where $**$ denotes 2-D convolution. Note that $\mathbf{H}_2\mathbf{H}_1$ and $\mathbf{H}_1\mathbf{H}_2$ have the same second-order statistics.

From the 2-D functions $C_{\mathbf{H}}(\nu,\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f)$, several 1-D channel descriptions can be derived. In particular, the *delay power profile* and *Doppler power profile* are respectively defined as [3]

$$P_{\mathbf{H}}(\tau) \triangleq \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu,\tau)\, d\nu\,, \qquad Q_{\mathbf{H}}(\nu) \triangleq \int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu,\tau)\, d\tau\,.$$

Their Fourier transforms,

$$\begin{aligned} p_{\mathbf{H}}(\Delta f) &\triangleq \int_{-\infty}^{\infty} P_{\mathbf{H}}(\tau)\, e^{-j2\pi\tau\Delta f}\, d\tau = R_{\mathbf{H}}(0, \Delta f)\,, \\ q_{\mathbf{H}}(\Delta t) &\triangleq \int_{-\infty}^{\infty} Q_{\mathbf{H}}(\nu)\, e^{j2\pi\nu\Delta t}\, d\nu = R_{\mathbf{H}}(\Delta t, 0)\,, \end{aligned}$$

are known as *time correlation function* and *frequency correlation function*, respectively. Often, for the sake of simplicity, a separable model is assumed for the scattering function and the TF correlation function, i.e., $C_{\mathbf{H}}(\nu,\tau) = \frac{1}{\rho_{\mathbf{H}}^2} Q_{\mathbf{H}}(\nu)\, P_{\mathbf{H}}(\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f) = \frac{1}{\rho_{\mathbf{H}}^2} q_{\mathbf{H}}(\Delta t)\, p_{\mathbf{H}}(\Delta f)$.

Examples of WSSUS channels. In the following, we briefly present some special cases and important examples of WSSUS channels.

- *Time-invariant WSSUS channel.* The impulse response of a time-invariant channel has the form $\tilde{h}(t,\tau) = g(\tau)$. The WSSUS property here implies $\mathrm{E}\{g(\tau)\, g^*(\tau')\} = P_{\mathbf{H}}(\tau)\, \delta(\tau - \tau')$. It follows that $C_{\mathbf{H}}(\nu,\tau) = \delta(\nu)\, P_{\mathbf{H}}(\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f) = p_{\mathbf{H}}(\Delta f)$.

- *Frequency-invariant WSSUS channel.* Next, we consider a "frequency-invariant" channel with impulse response $\tilde{h}(t,\tau) = w(t)\, \delta(\tau)$, i.e., the input signal is simply multiplied by $w(t)$. With $W(\nu)$ denoting the Fourier transform of $w(t)$, the WSSUS property here implies $\mathrm{E}\{W(\nu)\, W^*(\nu')\} = Q_{\mathbf{H}}(\nu)\, \delta(\nu - \nu')$. It follows that $C_{\mathbf{H}}(\nu,\tau) = Q_{\mathbf{H}}(\nu)\, \delta(\tau)$ and $R_{\mathbf{H}}(\Delta t, \Delta f) = q_{\mathbf{H}}(\Delta t)$.

- *Random TF shift.* The GSF of a channel effecting a random frequency shift by ν_0 and a random time shift by τ_0 is given by $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau) = \delta(\nu - \nu_0)\, \delta(\tau - \tau_0)$. It can be shown that this channel is WSSUS with scattering function $C_{\mathbf{H}}(\nu,\tau) = f_{\nu_0,\tau_0}(\nu,\tau)$, where $f_{\nu_0,\tau_0}(\nu,\tau)$ is the joint probability density function of (ν_0,τ_0) [9]. Furthermore, $R_{\mathbf{H}}(\Delta t, \Delta f) = \Psi_{\nu_0,\tau_0}(\Delta t, -\Delta f)$, with $\Psi_{\nu_0,\tau_0}(\Delta t, \Delta f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\nu_0,\tau_0}(\nu,\tau)\, e^{j2\pi(\nu\Delta t + \tau\Delta f)}\, d\nu\, d\tau$ being the characteristic function of (ν_0,τ_0).

- *Typical mobile radio channel.* A channel model popular in the mobile radio literature [3] uses a separable scattering function $C_{\mathbf{H}}(\nu,\tau) = \frac{1}{\rho_{\mathbf{H}}^2} Q_{\mathbf{H}}(\nu) P_{\mathbf{H}}(\tau)$ with an exponential delay power profile

$$P_{\mathbf{H}}(\tau) = \begin{cases} \frac{\rho_{\mathbf{H}}^2}{\tau_0}\, e^{-\tau/\tau_0}, & \tau \geq 0, \\ 0, & \tau < 0, \end{cases} \tag{9.5.9}$$

  and a *Jakes* Doppler power profile

$$Q_{\mathbf{H}}(\nu) = \begin{cases} \dfrac{\rho_{\mathbf{H}}^2}{\pi\sqrt{\nu_{\max}^2 - \nu^2}}, & |\nu| \leq \nu_{\max}, \\ 0, & |\nu| > \nu_{\max}. \end{cases} \tag{9.5.10}$$

  The associated TF correlation function is $R_{\mathbf{H}}(\Delta t, \Delta f) = \frac{1}{\rho_{\mathbf{H}}^2} q_{\mathbf{H}}(\Delta t)\, p_{\mathbf{H}}(\Delta f)$, with

$$p_{\mathbf{H}}(\Delta f) = \frac{\rho_{\mathbf{H}}^2}{1 + j2\pi\tau_0\Delta f}, \qquad q_{\mathbf{H}}(\Delta t) = \rho_{\mathbf{H}}^2\, J_0(2\pi\nu_{\max}\Delta t),$$

  where $J_0(\cdot)$ denotes the zero-order Bessel function of the first kind.

Statistical input-output relations for WSSUS channels. The scattering function and TF correlation function are useful for formulating input-output relations that show how the second-order statistics of the channel output signal $r(t)$ depend on the second-order statistics of the input signal $s(t)$. Let $s(t)$ be a nonstationary random process that is statistically independent of the random channel $\mathbf{H}$. The second-order statistics of a nonstationary random process $x(t)$ with correlation operator $\mathbf{R}_x$ can be described by the *generalized Wigner-Ville spectrum* $\overline{W}_x^{(\alpha)}(t,f) \triangleq L_{\mathbf{R}_x}^{(\alpha)}(t,f)$ or, alternatively, by the *generalized expected ambiguity function* $\bar{A}_x^{(\alpha)}(\nu,\tau) \triangleq S_{\mathbf{R}_x}^{(\alpha)}(\nu,\tau)$ which is the 2-D Fourier transform of $\overline{W}_x^{(\alpha)}(t,f)$ (see Article 9.4). It can then be shown that

$$\overline{W}_r^{(\alpha)}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C_{\mathbf{H}}(\nu,\tau)\,\overline{W}_s^{(\alpha)}(t-\tau,f-\nu)\,d\tau\,d\nu\,, \tag{9.5.11}$$

$$\bar{A}_r^{(\alpha)}(\nu,\tau) = R_{\mathbf{H}}(\tau,\nu)\,\bar{A}_s^{(\alpha)}(\nu,\tau)\,. \tag{9.5.12}$$

For a (wide-sense) *stationary* transmit signal $s(t)$ with power spectral density $P_s(f)$ and autocorrelation $r_s(\tau)$, the received signal $r(t)$ is stationary as well and (9.5.11) and (9.5.12) reduce to

$$P_r(f) = \int_{-\infty}^{\infty} Q_{\mathbf{H}}(\nu)\,P_s(f-\nu)\,d\nu\,, \qquad r_r(\tau) = q_{\mathbf{H}}(\tau)\,r_s(\tau)\,.$$

Furthermore, $\mathrm{E}\{|r(t)|^2\} = \rho_{\mathbf{H}}^2\,\mathrm{E}\{|s(t)|^2\}$. Dual results are obtained for a nonstationary white transmit signal $s(t)$.

Finally, if $s(t)$ is *cyclostationary* with period T, cyclic correlation function $r_s^{(k)}(\tau)$, and cyclic spectral density $P_s^{(k)}(f)$ [10], the received signal $r(t)$ is cyclostationary with the same period T and we have

$$P_r^{(k)}(f) = \int_{-\infty}^{\infty} Q_{\mathbf{H}}^{(k)}(\nu)\,P_s^{(k)}(f-\nu)\,d\nu\,, \qquad r_r^{(k)}(\tau) = q_{\mathbf{H}}^{(k)}(\tau)\,r_s^{(k)}(\tau)\,,$$

where $q_{\mathbf{H}}^{(k)}(\Delta t) = R_{\mathbf{H}}(\Delta t, k/T)$ and $Q_{\mathbf{H}}^{(k)}(\nu) = \int_{-\infty}^{\infty} q_{\mathbf{H}}^{(k)}(\Delta t)\,e^{-j2\pi\nu\Delta t}\,d\Delta t$.

9.5.3 Underspread WSSUS Channels

A fundamental classification of WSSUS channels is into *underspread* and *overspread* channels [2,4,7]. As we will show in this section, underspread WSSUS channels have some interesting properties. We note that the underspread property for WSSUS random channels is analogous to the underspread property for deterministic time-varying systems that was considered in Article 4.7.

Definition of underspread channels. A WSSUS channel is *underspread* [4, 7] if its scattering function is highly concentrated about the origin.[2] The underspread property is practically relevant as most mobile radio channels are underspread.

[2]For simplicity, we assume that the scattering function is centered about $\tau = 0$, which means that an overall delay $\tau_0 > 0$ has been split off from the channel.

A simple method for quantifying the concentration of the scattering function $C_{\mathbf{H}}(\nu,\tau)$ is based on the assumption that the support of $C_{\mathbf{H}}(\nu,\tau)$ is contained within a rectangle $[-\nu_{\max},\nu_{\max}]\times[-\tau_{\max},\tau_{\max}]$ about the origin of the (ν,τ) plane. (This implies that, with probability one, the GSF $S_{\mathbf{H}}^{(\alpha)}(\nu,\tau)$ is supported within this rectangle as well.) The channel's *delay-Doppler spread* is then defined as the area of this rectangle, $\sigma_{\mathbf{H}} \triangleq 4\nu_{\max}\tau_{\max}$, and the channel is said to be underspread if $\sigma_{\mathbf{H}} \leq 1$ and overspread if $\sigma_{\mathbf{H}} > 1$ [4,7].

An alternative characterization of scattering function concentration that avoids the assumption of compact support uses normalized *weighted integrals* of the form

$$\overline{m}_{\mathbf{H}}^{(\phi)} \triangleq \frac{1}{\rho_{\mathbf{H}}^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\nu,\tau)\,C_{\mathbf{H}}(\nu,\tau)\,d\nu\,d\tau = \frac{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\nu,\tau)\,C_{\mathbf{H}}(\nu,\tau)\,d\nu\,d\tau}{\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}C_{\mathbf{H}}(\nu,\tau)\,d\nu\,d\tau},$$

where $\phi(\nu,\tau) \geq 0$ is a weighting function that satisfies $\phi(\nu,\tau) \geq \phi(0,0) = 0$ and penalizes scattering function components lying away from the origin. Special cases are the *moments* $\overline{m}_{\mathbf{H}}^{(k,l)} \triangleq \overline{m}_{\mathbf{H}}^{(\phi_{k,l})}$ obtained with the weighting functions $\phi_{k,l}(\nu,\tau) = |\nu|^l|\tau|^k$ with $k,l \in \mathbb{N}_0$. Of particular importance are the *delay spread* $\tau_{\mathbf{H}}$ and *Doppler spread* $\nu_{\mathbf{H}}$ that are obtained with $k=2$, $l=0$ and $k=0$, $l=2$, respectively:

$$\tau_{\mathbf{H}}^2 \triangleq \overline{m}_{\mathbf{H}}^{(2,0)} = \frac{1}{\rho_{\mathbf{H}}^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\tau^2\,C_{\mathbf{H}}(\nu,\tau)\,d\nu\,d\tau\,, \tag{9.5.13}$$

$$\nu_{\mathbf{H}}^2 \triangleq \overline{m}_{\mathbf{H}}^{(0,2)} = \frac{1}{\rho_{\mathbf{H}}^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\nu^2\,C_{\mathbf{H}}(\nu,\tau)\,d\nu\,d\tau\,. \tag{9.5.14}$$

Within this framework, a WSSUS channel is called underspread if specific weighted integrals and moments of the scattering function are small.

Approximate eigenfunctions and eigenvalues of underspread channels. It is known [2, 4, 7] that signals with good time and/or frequency concentration can pass an underspread WSSUS channel almost undistorted, i.e., merely multiplied by a random complex factor. We will analyze this effect using the approach in [11]. We note that similar results in a deterministic context are reported in Articles 4.7 and 13.3.

A normalized transmit signal $s(t)$ that remains undistorted, i.e., $(\mathbf{H}s)(t) = \lambda\,s(t)$, is an eigenfunction of the system $\mathbf{H}$; the associated eigenvalue is given by $\lambda = \langle \mathbf{H}s, s\rangle$. Since $\mathbf{H}$ is random, the relation $(\mathbf{H}s)(t) = \langle \mathbf{H}s, s\rangle\,s(t)$ is more appropriately formulated in the mean-square sense, i.e.,

$$\mathrm{E}\{\|\mathbf{H}s - \langle \mathbf{H}s, s\rangle\,s\|^2\} = 0\,.$$

The eigenfunctions of a WSSUS channel $\mathbf{H}$ are random and generally do not possess a specific structure. However, in the *underspread* case, TF translates of a function

$g(t)$ with good TF concentration are approximate eigenfunctions. Specifically, consider the TF translates

$$g_{t_0,f_0}(t) \;=\; g(t-t_0)\, e^{j2\pi f_0 t}\,,$$

where $g(t)$ is a normalized function that is well concentrated about the origin of the TF plane. One can show

$$\mathrm{E}\Big\{\big\|\mathbf{H}g_{t_0,f_0} - \langle \mathbf{H}g_{t_0,f_0}, g_{t_0,f_0}\rangle\, g_{t_0,f_0}\big\|^2\Big\} \;\leq\; \rho_{\mathbf{H}}^2\, \overline{m}_{\mathbf{H}}^{(\phi)}\,, \tag{9.5.15}$$

with $\phi(\nu,\tau) = 1 - |A_g^{(\alpha)}(\nu,\tau)|^2$ where $A_g^{(\alpha)}(\nu,\tau) = \int_{-\infty}^{\infty} g\big(t+(\frac{1}{2}-\alpha)\tau\big)\, g^*\big(t-(\frac{1}{2}+\alpha)\tau\big)\, e^{-j2\pi\nu t}\, dt$ denotes the generalized ambiguity function of $g(t)$. Therefore, if the channel is underspread, i.e., the channel's scattering function is concentrated about the origin (where $|A_g^{(\alpha)}(\nu,\tau)|^2 \approx |A_g^{(\alpha)}(0,0)|^2 = 1$ and thus $\phi(\nu,\tau) \approx 0$), the weighted integral $\overline{m}_{\mathbf{H}}^{(\phi)}$ will be small and one has the approximation (valid in the mean-square sense)

$$(\mathbf{H}g_{t_0,f_0})(t) \;\approx\; \langle \mathbf{H}g_{t_0,f_0}, g_{t_0,f_0}\rangle\, g_{t_0,f_0}(t)\,. \tag{9.5.16}$$

This shows that $g_{t_0,f_0}(t)$ is an approximate eigenfunction of $\mathbf{H}$. Furthermore, it can be shown that

$$\mathrm{E}\Big\{\big|\langle \mathbf{H}g_{t_0,f_0}, g_{t_0,f_0}\rangle - L_{\mathbf{H}}^{(\alpha)}(t_0,f_0)\big|^2\Big\} \;\leq\; \rho_{\mathbf{H}}^2\, \overline{m}_{\mathbf{H}}^{(\phi')}\,, \tag{9.5.17}$$

with $\phi'(\nu,\tau) = |1 - A_g^{(\alpha)}(\nu,\tau)|^2$. Thus, under the same conditions as before, we have

$$\langle \mathbf{H}g_{t_0,f_0}, g_{t_0,f_0}\rangle \;\approx\; L_{\mathbf{H}}^{(\alpha)}(t_0,f_0) \tag{9.5.18}$$

(again valid in the mean-square sense), which shows that the approximate eigenvalue $\langle \mathbf{H}g_{t_0,f_0}, g_{t_0,f_0}\rangle$ is approximately equal to the GWS at the TF point (t_0,f_0).

In contrast to the exact eigenfunctions of $\mathbf{H}$, the approximate eigenfunctions $g_{t_0,f_0}(t)$ are TF translates of a single prototype function $g(t)$ and thus highly structured; they do not depend on the specific channel realization and their parameters t_0, f_0 have an immediate physical interpretation.

To illustrate the above eigenfunction/eigenvalue approximations, we simulated the transmission of a signal $g_{t_0,f_0}(t)$, with $g(t)$ a Hanning window of duration $T_g = 128\,\mu$s, over a WSSUS channel. The channel's scattering function was $C_{\mathbf{H}}(\nu,\tau) = \frac{1}{\rho_{\mathbf{H}}^2}\, Q_{\mathbf{H}}(\nu)\, P_{\mathbf{H}}(\tau)$ with exponential $P_{\mathbf{H}}(\tau)$ (Eq. (9.5.9) with $\tau_0 = 1\,\mu$s) and Jakes-type $Q_{\mathbf{H}}(\nu)$ (Eq. (9.5.10) with $\nu_{\max} = 305\,$Hz). Fig. 9.5.1 illustrates the approximations (9.5.16) and (9.5.18) for a single channel realization. It is seen that the received signal $(\mathbf{H}g_{t_0,f_0})(t)$ and the approximation $L_{\mathbf{H}}^{(1/2)}(t_0,f_0)\, g_{t_0,f_0}(t)$ are practically identical. Furthermore, we used 500 realizations of $\mathbf{H}$ to estimate the normalized mean-square error $\mathrm{E}\big\{\big\|\mathbf{H}g_{t_0,f_0} - \langle \mathbf{H}g_{t_0,f_0}, g_{t_0,f_0}\rangle\, g_{t_0,f_0}\big\|^2\big\}/\rho_{\mathbf{H}}^2$ (see (9.5.15)). The result, $9\cdot 10^{-4}$, confirms the validity of the eigenfunction/eigenvalue approximation. The associated upper bound $\overline{m}_{\mathbf{H}}^{(\phi)}$ (see (9.5.15)) was calculated as $5\cdot 10^{-3}$. Finally,

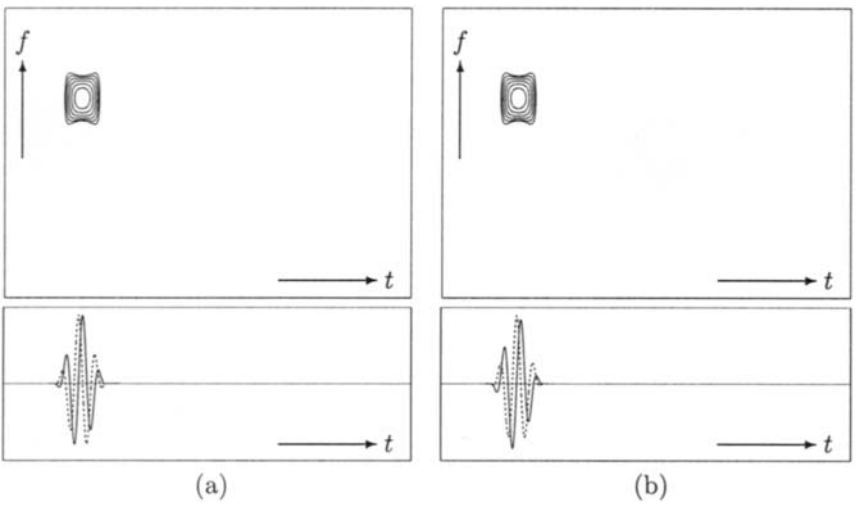


Fig. 9.5.1: Eigenfunction/eigenvalue approximation for an underspread WSSUS channel: (a) Wigner distribution [12] (top) and real and imaginary parts (bottom) of received signal $r(t) = (\mathbf{H}g_{t_0,f_0})(t)$, (b) Wigner distribution (top) and real and imaginary parts (bottom) of approximation $L_{\mathbf{H}}^{(1/2)}(t_0,f_0)\,g_{t_0,f_0}(t)$.

the normalized mean-square error $\mathrm{E}\{|\langle \mathbf{H}g_{t_0,f_0}, g_{t_0,f_0}\rangle - L_{\mathbf{H}}^{(1/2)}(t_0,f_0)|^2\}/\rho_{\mathbf{H}}^2$ (see (9.5.17)) was estimated as $5\cdot 10^{-6}$ and the associated upper bound $\overline{m}_{\mathbf{H}}^{(\phi')}$ was calculated as $2\cdot 10^{-3}$.

Sampling approximation for underspread channels. Next, we consider 2-D sampling of the channel's transfer function (GWS) $L_{\mathbf{H}}^{(\alpha)}(t,f)$. This is important for simplified channel representations that are used e.g. in the context of *orthogonal frequency division multiplexing (OFDM)* modulation [13,14].

Consider the representation of a WSSUS channel $\mathbf{H}$ by the samples $L_{\mathbf{H}}^{(\alpha)}(kT,lF)$ of its GWS taken on the uniform sampling grid (kT,lF). The reconstructed (interpolated) GWS is given by

$$\hat{L}_{\mathbf{H}}^{(\alpha)}(t,f) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} L_{\mathbf{H}}^{(\alpha)}(kT,lF)\,\mathrm{sinc}\Big(\frac{\pi}{T}(t-kT)\Big)\,\mathrm{sinc}\Big(\frac{\pi}{F}(f-lF)\Big),$$

with $\mathrm{sinc}(x) = \sin(x)/x$. For WSSUS channels with scattering function $C_{\mathbf{H}}(\nu,\tau)$ compactly supported within a rectangular area $[-\nu_{\max},\nu_{\max}]\times[-\tau_{\max},\tau_{\max}]$ and for sampling grid constants satisfying $T \le 1/(2\nu_{\max})$ and $F \le 1/(2\tau_{\max})$, the above reconstruction can be shown to be exact in the sense that $\mathrm{E}\{|\hat{L}_{\mathbf{H}}^{(\alpha)}(t,f) - L_{\mathbf{H}}^{(\alpha)}(t,f)|^2\} = 0$. Note that a smaller channel spread $\sigma_{\mathbf{H}} = 4\nu_{\max}\tau_{\max}$ allows for a coarser sampling grid and thus for a more parsimonious channel representation.

If the above conditions are not satisfied, the reconstructed GWS $\hat{L}_{\mathbf{H}}^{(\alpha)}(t,f)$ will

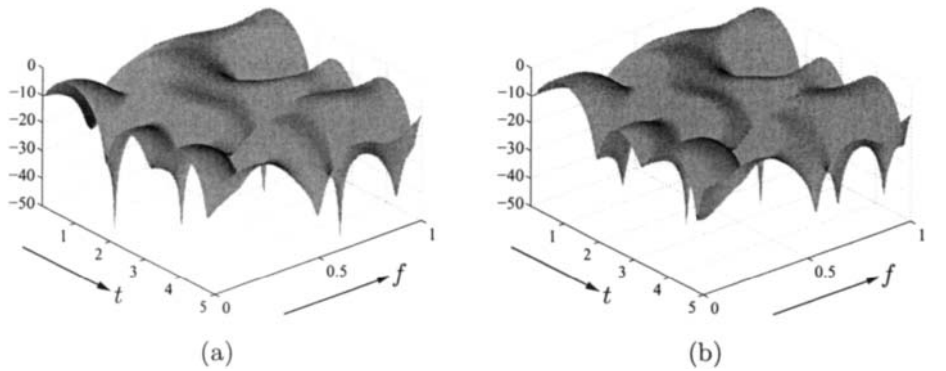


Fig. 9.5.2: Sampling approximation for an underspread WSSUS channel: (a) $|L_{\mathbf{H}}^{(1/2)}(t,f)|^2$ and (b) $|\hat{L}_{\mathbf{H}}^{(1/2)}(t,f)|^2$ (in dB; time in ms, frequency in MHz).

contain errors due to aliasing. However, it can be shown that these errors are bounded as

$$\mathrm{E}\{|\hat{L}_{\mathbf{H}}^{(\alpha)}(t,f) - L_{\mathbf{H}}^{(\alpha)}(t,f)|^2\} \leq 2\,\rho_{\mathbf{H}}^2\,(\tau_{\mathbf{H}}^2 F^2 + \nu_{\mathbf{H}}^2 T^2)\,, \tag{9.5.19}$$

where $\tau_{\mathbf{H}}$ and $\nu_{\mathbf{H}}$ are the delay spread and Doppler spread as defined in (9.5.13), (9.5.14). Thus, for WSSUS channels with small $\tau_{\mathbf{H}}$ and/or small $\nu_{\mathbf{H}}$, i.e., for *underspread* channels, a sampling of the transfer function will result in negligible errors provided that the sampling periods T and F are chosen appropriately. Specifically, the upper error bound in (9.5.19) is minimized when $T/F = \tau_{\mathbf{H}}/\nu_{\mathbf{H}}$.

For the WSSUS channel with exponential/Jakes scattering function that was considered further above, and for sampling periods $T = 138\,\mu$s, $F = 136.72\,$kHz, the normalized mean-square error $\mathrm{E}\{|\hat{L}_{\mathbf{H}}^{(1/2)}(t,f) - L_{\mathbf{H}}^{(1/2)}(t,f)|^2\}/\rho_{\mathbf{H}}^2$ was estimated from 500 channel realizations as $6.4 \cdot 10^{-3}$, and the upper bound $2\,(\tau_{\mathbf{H}}^2 F^2 + \nu_{\mathbf{H}}^2 T^2)$ was calculated as $3.2 \cdot 10^{-2}$. Fig. 9.5.2 shows the squared magnitude of the true channel transfer function $L_{\mathbf{H}}^{(1/2)}(t,f)$ and of the reconstruction $\hat{L}_{\mathbf{H}}^{(1/2)}(t,f)$ for a specific channel realization.

9.5.4 Summary and Conclusions

In this article, we have considered time-frequency characterizations of (the second-order statistics of) random linear time-varying channels satisfying the assumption of wide-sense stationary uncorrelated scattering (WSSUS). We have shown that the practically important class of *underspread* WSSUS channels allows some interesting approximations. In particular, underspread WSSUS channels possess approximate eigenfunctions with time-frequency shift structure (which suggests the use of OFDM), and they can be discretized by means of a time-frequency sampling. Related considerations and results can be found in Articles 13.2 and 13.3.

References

[1] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Communication Systems*, vol. 11, pp. 360–393, December 1963.

[2] R. S. Kennedy, *Fading dispersive communication channels.* New York: Wiley, 1969.

[3] J. D. Parsons, *The Mobile Radio Propagation Channel.* London: Pentech Press, 1992.

[4] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 3rd ed., 1995.

[5] T. S. Rappaport, *Wireless Communications: Principles & Practice.* Upper Saddle River, NJ: Prentice-Hall, 1996.

[6] K. A. Sostrand, "Mathematics of the time-varying channel," in *Proc. NATO Advanced Study Inst. on Signal Processing with Emphasis on Underwater Acoustics*, vol. 2, pp. 25.1–25.20, 1968.

[7] H. L. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. III: "Radar-Sonar Signal Processing and Gaussian Signals in Noise". New York: Wiley, 1971. Reprinted Malabar, FL: Krieger, 1992. Reprinted New York: Wiley, 2001.

[8] H. Artés, G. Matz, and F. Hlawatsch, "Unbiased scattering function estimation during data transmission," in *Proc. IEEE Vehicular Technology Conf. (VTC'99-Fall)*, pp. 1535–1539, Amsterdam, 19–22 September 1999.

[9] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[10] W. A. Gardner, ed., *Cyclostationarity in Communications and Signal Processing.* Piscataway, NJ: IEEE Press, 1995.

[11] W. Kozek and A. F. Molisch, "On the eigenstructure of underspread WSSUS channels," in *Proc. IEEE-SP Workshop on Signal Processing Advances in Wireless Communications (SPAWC'97)*, pp. 325–328, Paris, 16–18 April 1997.

[12] W. Mecklenbräuker and F. Hlawatsch, eds., *The Wigner Distribution —Theory and Applications in Signal Processing.* Amsterdam: Elsevier, 1997.

[13] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communications Magazine*, vol. 28, pp. 5–14, May 1990.

[14] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," *IEEE J. on Selected Areas in Communications*, vol. 16, pp. 1579–1589, October 1998.

This Page Intentionally Left Blank

Chapter 10

Instantaneous Frequency Estimation and Localization

A critical feature of a non-stationary signal is provided by its instantaneous frequency (IF), which accounts for the signal spectral variations as a function of time. This chapter presents methods and algorithms for the localization and estimation of the signal IF. The topic is covered in five articles with appropriate internal cross-referencing to this and other chapters.

The first conventional approach for IF estimation used the spectrogram. To account for its window-dependent resolution, improvements were made by introducing iterative methodologies on the estimate provided by the first moment of the spectrogram (Article 10.1). Another approach uses an adaptive algorithm for IF estimation using the peak of suitable TFDs with adaptive window length (10.2). This method was extended to the case of multicomponent signals using the modified B-distribution (10.3). When the signals considered have polynomial FM characteristics, both the peak of the polynomial WVD and higher-order ambiguity functions can be used (10.4). In the special case when the signals are subject to random amplitude modulation (or multiplicative noise), IF estimation procedures are described using the peak of the WVD for linear FM signals, and the peak of the PWVD for non-linear FM signals (10.5).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

10.1 ITERATIVE INSTANTANEOUS FREQUENCY ESTIMATION FOR RANDOM SIGNALS [0]

10.1.1 IF Estimation: Introduction and Background

The instantaneous frequency (IF) is a basic parameter which may be used to describe the nonstationarity in a process (see Article 1.3). It is used in many areas such as seismic, radar, sonar, communications and biomedical applications [1–3].

The concept of IF is often introduced in the context of frequency modulation in communication theory. The IF of a signal at time t_0 could be viewed as the frequency of the sinusoid which locally fits the signal at that time. Using the definition of Gabor and later Ville [4], the IF, $f(t)$, of a signal $x(t)$ can be defined in terms of the derivative of the phase of its analytic signal $z(t)$; i.e.,

$$f(t) = \frac{1}{2\pi}\frac{d}{dt}\arg(z(t)). \tag{10.1.1}$$

Another interpretation of the IF comes from the time-frequency distribution (TFD) point of view, where the IF of a signal at time t is defined as the weighted average of the frequencies which exist in the signal at time t [5]; i.e.,

$$\hat{f}(t) = \frac{\int_{-\infty}^{\infty} f\hat{P}(t,f)\,df}{\int_{-\infty}^{\infty} \hat{P}(t,f)\,df} \tag{10.1.2}$$

where $\hat{P}(t,f)$ is the TFD estimate.

The IF has many applications in time–frequency signal analysis. For example a matched spectrogram, which uses chirp windows, is shown to perform better than a regular spectrogram for chirp-like signals when the chirp rate of the window is matched to the IF of the signal [6]. The IF can also be utilized to obtain distributions mainly concentrated along the IF for monocomponent signals. For example, [7] shows how to obtain distributions infinitely concentrated along the IF for pure tones and for amplitude- and frequency-modulated signals, and considers kernel functions for bilinear distributions with these properties.

There are many ways to estimate the IF of a signal. The current IF estimation algorithms can be grouped as phase differencing methods, signal modeling methods (e.g., short time AR modeling, time–varying AR modeling), phase modeling methods, and time–frequency–representation methods. A good review of these methods can be found in [8].

One important time–frequency–representation method is the iterative cross Wigner-Ville IF estimation technique [8], in which, at each iteration, we construct a constant-amplitude FM signal whose IF matches the previous IF estimate, then take the cross Wigner-Ville distribution between the constructed signal and the

[0]Authors: **Amro El-Jaroudi** and **Mustafa K. Emresoy**, 348 Benedum Engineering Hall, Dept. of Electrical Engineering, University of Pittsburgh, Pittsburgh, PA, 15261, USA (amro@ee.pitt.edu, emresoy@siglab.ee.pitt.edu). Reviewers: M. Mesbah and G. R. Putland.

signal under analysis, and then use the cross-WVD to produce a new IF estimate. This method is shown to perform very well at low signal-to-noise ratios (cf. Subsection 3.1.2.3, especially at item 2).

In the next two Sections, we present an iterative algorithm to estimate the IF and matched spectrogram of a nonstationary sinusoidal signal, and then analyze the convergence of the algorithm. Our algorithm is similar to the cross Wigner-Ville algorithm in that we use a time-frequency representation iteratively to obtain an estimate of the IF. The iterative IF estimation method can be used to improve the performance of many time-frequency methods [9, 10].

10.1.2 Iterative Algorithm for IF Estimation

The basic idea behind the iterative IF estimation is to use a time–frequency estimator (in this case, the spectrogram) to obtain an IF estimate using (10.1.2) above, then use this IF estimate to recalculate the spectrogram. This process is repeated until convergence is reached. The principle at work here is that improving the IF estimate makes the matched spectrogram estimate better and vice versa. Consequently each iteration will lead to improvements in both quantities.

We will assume that the signal under analysis is a complex signal of the form

$$x(t) = A(t)\, e^{j\phi(t)} \tag{10.1.3}$$

where $A(t)$ is the time-varying amplitude and $\phi(t)$ is the phase of the signal.

In the iterative algorithm given below, the instantaneous frequency and the spectrogram as a TFD are estimated by using the whole frequency axis. Then the signal is "demodulated" by subtracting the phase estimate (obtained by integrating the IF estimate) from the phase of the signal. This demodulation shifts the spectrum of the nonstationary signal around zero frequency. The demodulated signal is then reanalyzed. Consequently we do not restrict the definition of the IF to use only positive frequencies, since after the first iteration the signal under analysis has non-redundant information over positive and negative frequencies. But when the algorithm is applied to real signals, we use the analytic signal to start the iterations, ensuring that the first IF estimate is not trivially zero. Note that although we use the analytic signal at first, the demodulated signal used in the subsequent iteration is not analytic. The outputs of the algorithm are the estimated IF and the matched spectrogram of the signal.

Algorithm:

1. Calculate the spectrogram $\hat{P}(t, f)$ of the signal (or of the analytic component if the signal is real).

2. Estimate $\hat{f}_i(t)$ by using the first moment of the spectrogram at each time and find the phase estimate

$$\hat{f}_i(t) = \frac{\int_{-\infty}^{\infty} f \hat{P}_i(t, f)\, df}{\int_{-\infty}^{\infty} \hat{P}_i(t, f)\, df} \tag{10.1.4}$$

$$\hat{\phi}_i(t) = 2\pi \int_{-\infty}^{t} \hat{f}_i(\tau)\, d\tau. \tag{10.1.5}$$

3. Demodulate the signal along the estimated IF

$$\tilde{x}_i(t) = A(t)\, e^{-j(\phi(t)-\hat{\phi}_i(t))} = A(t)\, e^{-j\tilde{\phi}_i(t)}. \tag{10.1.6}$$

4. Take the spectrogram $\tilde{P}_i(t,f)$ of $\tilde{x}(t)$ and compensate for the demodulation to obtain a new matched spectrogram estimate

$$\hat{P}_{i+1}(t,f) = \tilde{P}_i(t, f - \hat{f}_i(t)). \tag{10.1.7}$$

5. Go to step 2.

The algorithm is stopped once the IF estimate has converged. The convergence can be checked by comparing the difference between consecutive iterations with stopping threshold.

While the algorithm seems very simple, we show below that its convergence can be guaranteed given certain conditions on the analysis signal.

10.1.3 Convergence of the Estimation Algorithm

In order to examine the convergence of the iterative algorithm, we derive an expression for the IF estimate after each iteration. The IF estimate of the complex signal in (10.1.3) obtained by using the first moment of a spectrogram with a real analysis window, $h(t)$, is given by [5]

$$\hat{f}(t) = \frac{1}{2\pi} \sum_{k=0}^{\infty} \frac{M_k(t)}{k!}\, \phi^{(k+1)}(t) \tag{10.1.8}$$

where

$$M_n(t) = \frac{\int A^2(t+\tau)\, h^2(\tau)\, \tau^n d\tau}{\int A^2(t+\tau)\, h^2(\tau)\, d\tau} \tag{10.1.9}$$

and $\phi^{(n+1)}(t)$ is the $(n+1)^{\text{th}}$ derivative of the phase of the signal, with

$$\frac{M_0(t)}{0!} = 1. \tag{10.1.10}$$

Our hypothesis is that the general expression for the IF estimate at the end of the n^{th} iteration is

$$\hat{f}_n(t) = f(t) + (-1)^{(n+1)} \frac{1}{2\pi} \sum_{i_n=1}^{\infty} \cdots \sum_{i_1=1}^{\infty} \frac{M_{i_n}(t)}{i_n!} \Big\{ \cdots \Big\{ \frac{M_{i_1}(t)}{i_1!}\, \phi^{(i_1+1)}(t) \Big\}^{(i_2)} \cdots \Big\}^{(i_n)} \tag{10.1.11}$$

where $f(t) = \frac{1}{2\pi}\phi'(t)$.

We prove the hypothesis using the method of induction. To show that the equation (10.1.11) is true for the first iteration, we remove the summations i_2 to i_n and substitute $n = 1$ in equation (10.1.11), obtaining

$$\hat{f}_1(t) = f(t) + \frac{1}{2\pi}\sum_{i_1=1}^{\infty}\frac{M_{i_1}(t)}{i_1!}\,\phi^{(i_1+1)}(t). \tag{10.1.12}$$

This is the same as equation (10.1.8); i.e., equation (10.1.11) is true for $n = 1$.

For the inductive step, we assume that equation (10.1.11) is true for iteration $(n-1)$, so that

$$\hat{f}_{n-1}(t) = f(t) + (-1)^n\frac{1}{2\pi}\sum_{i_{n-1}=1}^{\infty}\cdots\sum_{i_1=1}^{\infty}\frac{M_{i_{n-1}}(t)}{i_{n-1}!}\Big\{\cdots\Big\{\frac{M_{i_1}(t)}{i_1!}\,\phi^{(i_1+1)}(t)\Big\}^{(i_2)}\cdots\Big\}^{(i_{n-1})}, \tag{10.1.13}$$

and show that it is also true for the n^{th} iteration. Following the algorithm given in the previous section, the phase error $\tilde{\phi}(t)$ at the n^{th} iteration is given by

$$\tilde{\phi}_n(t) = \phi(t) - 2\pi\int_{-\infty}^{t}\hat{f}_{n-1}(\tau)\,d\tau. \tag{10.1.14}$$

By combining the previous two equations, we obtain

$$\tilde{\phi}_n(t) = (-1)^{(n+1)}\sum_{i_{n-1}=1}^{\infty}\cdots\sum_{i_1=1}^{\infty}\int_{-\infty}^{t}\frac{M_{i_{n-1}}(\tau)}{i_{n-1}!}\Big\{\cdots\Big\{\frac{M_{i_1}(\tau)}{i_1!}\phi^{(i_1+1)}(\tau)\Big\}^{(i_2)}\cdots\Big\}^{(i_{n-1})}d\tau. \tag{10.1.15}$$

The IF estimate at the end of n^{th} iteration can be found by adding the IF estimate from iteration $n-1$ and the IF estimate of the signal $\tilde{x}(t)$ whose phase is given by (10.1.15); that is,

$$\hat{f}_n(t) = \hat{f}_{n-1}(t) + \frac{1}{2\pi}\sum_{i_n=0}^{\infty}\frac{M_{i_n}(t)}{i_n!}\,\tilde{\phi}_n^{(i_n+1)}(t). \tag{10.1.16}$$

Substituting for $\hat{f}_{n-1}(t)$ and $\tilde{\phi}_n(t)$ in the above equation, we find that the IF estimate at the end of the n^{th} iteration is

$$\begin{aligned}\hat{f}_n(t) = f(t) &+ (-1)^n\frac{1}{2\pi}\sum_{i_{n-1}=1}^{\infty}\cdots\sum_{i_1=1}^{\infty}\frac{M_{i_{n-1}}(t)}{i_{n-1}!}\Big\{\cdots\Big\{\frac{M_{i_1}(t)}{i_1!}\phi^{(i_1+1)}(t)\Big\}^{(i_2)}\cdots\Big\}^{(i_{n-1})}\\ &+ (-1)^{(n+1)}\sum_{i_n=0}^{\infty}\cdots\sum_{i_1=1}^{\infty}\frac{M_{i_n}(t)}{i_n!}\int_{-\infty}^{t}\Big\{\cdots\Big\{\frac{M_{i_1}(\tau)}{i_1!}\phi^{(i_1+1)}(\tau)\Big\}^{(i_2)}\cdots\Big\}^{(i_n+1)}d\tau\end{aligned}$$

which simplifies to equation (10.1.11), as claimed. This completes the proof.

In equation (10.1.11), the first term represents the actual IF whereas the second term represents the error in the IF estimate at the end of the n^{th} iteration. In general the error term is very difficult to analyze, being affected by both the amplitude and phase variations in the signal. Assuming, however, that the spectrogram window is Gaussian, and that the phase function of the signal has bounded derivatives of all orders, and that the signal is dominantly FM modulated (i.e. has constant or slowly varying amplitude), we may show that the iteration in (10.1.11) converges to the true IF of the signal. Let the spectrogram window be

$$h(t) = (\tfrac{\alpha}{\pi})^{1/4}\, e^{-\alpha t^2/2}, \tag{10.1.17}$$

and let the amplitude $A(t)$ be approximated by the constant A within the window, so that

$$x(t) = A\, e^{j\phi(t)}. \tag{10.1.18}$$

Making these two substitutions in (10.1.9), we find

$$M_n = \frac{(1+(-1)^n)\,\gamma(\frac{n+1}{2})}{2\sqrt{\pi}\,\alpha^{n/2}} \tag{10.1.19}$$

$$= \begin{cases} 0 & \text{if } n \text{ is odd} \\ \prod_{k=0}^{n/2} \frac{(n+(2k-1))\sqrt{\pi}}{\alpha^{n/2}} & \text{if } n \text{ is even.} \end{cases} \tag{10.1.20}$$

Then, from (10.1.11), the general expression for the IF estimate is

$$\hat{f}_n(t) = f(t) + (-1)^n \frac{1}{2\pi} \sum_{i_1,i_2,\ldots,i_n=2}^{\infty} \frac{M_{i_1} M_{i_2} \ldots M_{i_n}}{i_1!\, i_2! \ldots i_n!} \phi^{(i_1+i_2+\cdots+i_n+1)}(t). \tag{10.1.21}$$

If we define

$$c_m = \sum_{i_1+i_2+\cdots+i_n=m} \frac{M_{i_1} M_{i_2} \ldots M_{i_n}}{i_1!\, i_2! \ldots i_n!}, \tag{10.1.22}$$

then (10.1.21) can be written as

$$\hat{f}_n(t) = f(t) + \frac{(-1)^n}{2\pi} \sum_{m=2}^{\infty} c_m \phi^{(m+1)}(t). \tag{10.1.23}$$

It is important to note that $\{c_m\}$ are independent of the analysis signal and are mainly a function of the window parameter α. It is easy to show that after each iteration, two coefficients of the coefficient series $\{c_m\}$ become zero and overall the coefficients decay rapidly to zero. It is clear that, as $n \to \infty$, the coefficients go to zero. If we assume that the phase of the signal is continuous and has bounded derivatives, then the sum in equation (10.1.23) also goes to zero which implies $\hat{f}_n(t) \to f(t)$.

Convergence properties:

- Equation (10.1.20) implies that the convergence rate of the algorithm is proportional to $\alpha^{n/2}$. That is, for small α (long window) the convergence is slow whereas for large α (short window) the convergence is very fast.
- The convergence rate is also affected by the rate of change of the IF. For signals whose IF's have nonzero higher order derivatives, the convergence is slower. But for signals with polynomial phases, the algorithm theoretically converges with an iteration number equal to half of the degree of the polynomial. This can be seen by examining equation (10.1.21). At iteration n the first $2n$ coefficients in equation (10.1.23) become zero.
- If the phase of the signal has discontinuities at some points t_i, the derivatives of $\phi(t)$ go to $\pm\infty$. Then we cannot expect the algorithm to converge exactly to the $f(t)$. If $\phi(t)$ has bounded derivatives for all orders and for all t, then the algorithm converges to $f(t)$.

10.1.4 Summary and Conclusions

In this article, we presented an iterative algorithm to estimate the instantaneous frequency (IF) and matched spectrogram of nonstationary sinusoidal signals. The matched spectrogram obtained by this method is concentrated along the IF for monocomponent signals. The convergence analysis and the properties of the IF estimation algorithm are presented.

References

[1] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals," *Proc. IEEE*, vol. 80, pp. 520–538, April 1992.

[2] B. Ferguson, "A ground based narrow-band passive acoustic technique for estimating the altitude and speed of a propeller driven aircraft," *J. Acoustical Soc. of America*, vol. 92, pp. 1403–1407, September 1992.

[3] M. T. Taner, F. Koehler, and R. E. Sheriff, "Complex seismic trace analysis," *Geophysics*, vol. 44, pp. 1041–1063, June 1979.

[4] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmissions*, vol. 2A, no. 1, pp. 61–74, 1948. In French. English translation: I. Selin, *Theory and applications of the notion of complex signal*, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).

[5] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[6] D. L. Jones and T. W. Parks, "A high-resolution data-adaptive time-frequency representation," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 2127–2135, December 1990.

[7] L. Cohen, "Distributions concentrated along the instantaneous frequency," in *Proc. SPIE: Advanced Signal-Processing Algorithms, Architectures, and Implementations*,

vol. 1348, pp. 149–157, Soc. of Photo-optical Instrumentation Engineers, San Diego, 10–12 July 1990.

[8] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 2: Algorithms and applications," *Proc. IEEE*, vol. 80, pp. 540–568, April 1992.

[9] M. K. Emresoy and A. El-Jaroudi, "Iterative instantaneous frequency estimation and adaptive matched spectrogram," *Signal Processing*, vol. 64, pp. 157–65, January 1998.

[10] D. L. Jones and R. G. Baraniuk, "A simple scheme for adapting time-frequency representations," *IEEE Trans. Signal Processing*, vol. 42, pp. 3530–3535, December 1994.

10.2 ADAPTIVE INSTANTANEOUS FREQUENCY ESTIMATION USING TFDs[0]

Instantaneous frequency (IF) estimators based on maxima of time-frequency representations have variance and bias which are highly dependent on the lag window width. The optimal window width may be determined by minimizing the estimation mean squared error (MSE), provided that some signal and noise parameters are explicitly known. However, these parameters are not available in advance. This is especially true for the IF derivatives which determine the estimation bias. In this article, an adaptive algorithm for the lag window width determination, based on the confidence intervals intersection, will be presented [1–5]. This algorithm does not require knowledge of the estimation bias value. The theory and algorithm presented here are not limited to the IF estimation and time-frequency analysis. They may be applied to a parameter value selection in various problems.

10.2.1 Optimal Window Width

Consider a noisy signal:

$$x(n\Delta t) = s(n\Delta t) + \epsilon(n\Delta t), \quad s(t) = a\exp(j\phi(t)), \tag{10.2.1}$$

with $s(n\Delta t)$ being a signal and $\epsilon(n\Delta t)$ being a white complex-valued Gaussian noise with mutually independent real and imaginary parts of equal variances $\sigma_\epsilon^2/2$. Sampling interval is denoted by Δt. Consider the problem of the IF, $f_i(t) = \phi'(t)/2\pi$, estimation from the discrete-time observations $x(n\Delta t)$, based on maxima of a time-frequency distribution $\rho_x(t,f)$,

$$\hat{f}(t) = \arg\{\max_f \rho_x(t,f)\}. \tag{10.2.2}$$

Let $\triangle\hat{f}(t) = f_i(t) - \hat{f}(t)$ be the estimation error. The MSE, $E\{(\triangle\hat{f}(t))^2\}$, is used for the accuracy characterization at a given time instant t. Asymptotically, the MSE for commonly used time-frequency representations (e.g. the spectrogram, the pseudo Wigner distribution (WD), and its higher order versions) can be expressed in the following form [2–6] [Articles 10.3, 10.4]

$$E\left\{(\triangle\hat{f}(t))^2\right\} = \frac{V}{h^m} + B(t)h^n, \tag{10.2.3}$$

where h is a lag window $w_h(t)$ width, such that $w_h(t) = 0$ for $|t| > h/2$. It is related to the number of samples N by $h = N\Delta t$. The variance and the bias of estimate, for a given h, are

$$\sigma^2(h) = V/h^m, \quad \text{bias}\,(t,h) = \sqrt{B(t)h^n}. \tag{10.2.4}$$

[0]Author: **LJubiša Stanković**, Elektrotehnicki fakultet, University of Montenegro, 81000 Podgorica, Montenegro (l.stankovic@ieee.org). Reviewers: B. Barkat, I. Djurović and V. Ivanović.

The expression for $B(t)$ is a function of the IF derivatives.

For example, for the WD with a rectangular lag window we have [3]

$$E\left\{(\triangle\hat{f}(t))^2\right\} = \frac{6\sigma_\epsilon^2\Delta t}{(2\pi a)^2}\frac{1}{h^3} + \left(\frac{\phi^{(3)}(t)}{80\pi}\right)^2 h^4, \tag{10.2.5}$$

corresponding to $m = 3$ and $n = 4$ in (10.2.3). Values of m and n for some other distributions are indicated in Table 10.2.1, according to the results from [2–5].

The MSE in (10.2.3) has a minimum with respect to h. This minimum occurs for the optimal value of h given by

$$h_{opt}(t) = [mV/(nB(t))]^{1/(m+n)}. \tag{10.2.6}$$

Note that this relation is not useful in practice, because its right hand-side contains $B(t)$ which depends on derivatives of the unknown IF.

10.2.2 Adaptive Algorithm

Here, we present an adaptive method which can produce an estimate of $h_{opt}(t)$ without having to know the value of $B(t)$. For the optimal window width, according to (10.2.3), holds

$$\frac{\partial E\left\{(\triangle\hat{f}(t))^2\right\}}{\partial h} = -m\frac{V}{h^{m+1}} + nB(t)h^{n-1} = 0_{|h=h_{opt}}. \tag{10.2.7}$$

Multiplying (10.2.7) by h, we get the relationship between the bias and standard deviation, (10.2.4), for $h = h_{opt}$,

$$\text{bias}\,(t, h_{opt}) = \sqrt{\frac{m}{n}}\sigma(h_{opt}). \tag{10.2.8}$$

It will be assumed, without loss of generality, that the bias is positive. The IF estimate $\hat{f}_h(t)$ (obtained from (10.2.2) by using the lag window of width h) is a random variable distributed around the true IF $f_i(t)$ with the bias $\text{bias}\,(t, h)$ and the standard deviation $\sigma(h)$. Thus, we may write the relation:

$$\left|f_i(t) - \left(\hat{f}_h(t) - \text{bias}\,(t, h)\right)\right| \leq \kappa\sigma(h), \tag{10.2.9}$$

where the inequality holds with probability $P(\kappa)$ depending on parameter κ.[1] We will assume that κ is such that $P(\kappa) \to 1$.

[1] If we assume, for example, that the random variable $\hat{f}_h(t)$ is Gaussian, with the mean value $M = f_i(t) + \text{bias}\,(t, h)$ and the standard deviation $\sigma(h)$, then the probability that $\hat{f}_h(t)$ takes a value within the interval $[M - \kappa\sigma(h), M + \kappa\sigma(h)]$ is $P(\kappa) = 0.95$ for $\kappa = 2$, and $P(\kappa) = 0.997$ for $\kappa = 3$.

Let us introduce a set of discrete dyadic window-width values, $h \in H$,

$$H = \{h_s \mid h_s = 2h_{s-1}, s = 1, 2, ..., J\}. \tag{10.2.10}$$

Define the confidence intervals $D_s = [L_s, U_s]$ of the IF estimates, with the following upper and lower bounds

$$L_s = \hat{f}_{h_s}(t) - (\kappa + \Delta\kappa)\,\sigma(h_s), \quad U_s = \hat{f}_{h_s}(t) + (\kappa + \Delta\kappa)\,\sigma(h_s), \tag{10.2.11}$$

where $\hat{f}_{h_s}(t)$ is an estimate of the IF, for the window width $h = h_s$, and $\sigma(h_s)$ is its standard deviation. Assume that a window width denoted by $h_{s^+} \in H$ is of h_{opt} order, $h_{s^+} \sim h_{opt}$. Since h_{opt} does not correspond to any h_s from the set H, for the analysis that follows we can write $h_{s^+} = 2^p h_{opt}$, where p is a constant close to 0. According to (10.2.10) all other windows can be written as a function of h_{s^+} as

$$h_s = h_{s^+} 2^{(s-s^+)} = h_{opt} 2^{s-s^+ + p}, \quad (s - s^+) = ..., -2, -1, 0, 1, 2, ... \tag{10.2.12}$$

With this notation, having in mind (10.2.8), the standard deviation and the bias from (10.2.4) can be expressed by

$$\begin{aligned} \sigma(h_s) &= \sqrt{V/h_s^m} = \sigma(h_{opt}) 2^{-(s-s^+ + p)m/2}, \\ \text{bias}\,(t, h_s) &= \sqrt{B(t) h_s^n} = \sqrt{m/n}\;\sigma(h_{opt}) 2^{(s-s^+ + p)n/2}. \end{aligned} \tag{10.2.13}$$

For small window widths h_s, when $s \ll s^+$, the bias of $\hat{f}_{h_s}(t)$ is negligible, thus $f_i(t) \in D_s$ (with probability $P(\kappa + \Delta\kappa) \to 1$). Then, obviously, $D_{s-1} \cap D_s \neq \emptyset$, since at least the true IF, $f_i(t)$, belongs to both confidence intervals. For $s \gg s^+$ the variance is small, but the bias is large. It is clear that for $\text{bias}\,(t, h_s) \neq 0$ there exists such a large s that $D_s \cap D_{s+1} = \emptyset$ for a finite $\kappa + \Delta\kappa$.

The **idea behind the algorithm** is that $\Delta\kappa$ in D_s can be found in such a way that the largest s, for which the sequence of the pairs of the confidence intervals D_{s-1} and D_s has at least a point in common, is $s = s^+$. Such a value of $\Delta\kappa$ exists because the bias and the variance are monotonically increasing and decreasing functions of h, respectively, (10.2.13). As soon as this value of $\Delta\kappa$ is found, an intersection of the confidence intervals D_{s-1} and D_s,

$$\left|\hat{f}_{h_{s-1}}(t) - \hat{f}_{h_s}(t)\right| \leq (\kappa + \Delta\kappa)[\sigma(h_{s-1}) + \sigma(h_s)], \tag{10.2.14}$$

works as an indicator of the event $s = s^+$, i.e., the event $h_s = h_{s^+} \sim h_{opt}$. The value of h_{s^+} is the last h_s when (10.2.14) is still satisfied.

10.2.2.1 Parameters in the Adaptive Algorithm

There are three possible approaches to choosing algorithm parameters κ, $\Delta\kappa$, and p. Their performance do not differ significantly.

(1) When our knowledge about the variance and bias behavior, given by (10.2.3), is not quite reliable, an approximative approach for κ, $\Delta\kappa$, and p determination

Table 10.2.1: Parameters in the adaptive algorithm for various m, n, κ: $m = 3, n = 4$ for the spectrogram, Wigner and L-Wigner distribution based IF estimators; $m = 3, n = 8$ for the fourth order polynomial Wigner-Ville distribution, and local polynomial distribution based IF estimators; $m = 1, n = 4$ for the Wigner distribution as a spectrum estimator.

m	1	1	3	3	3	3	3	3
n	4	4	4	4	4	8	8	8
κ	2	3	2	3	5	2	3	5
$\Delta\kappa$	0.86	1.29	0.39	0.58	0.97	0.09	0.14	0.23
p	0.99	1.22	0.34	0.51	0.72	-0.13	-0.03	0.11
p_1	1.18	1.41	0.59	0.76	0.97	0.19	0.30	0.43

can be used. Then, we can assume a value of $\kappa \cong 2.5$, such that $P(\kappa) \cong 0.99$ for Gaussian distribution of estimation error. The value of $\Delta\kappa$ should take into account the bias for the expected optimal window width (10.2.8). It is common to assume that, for the optimal value of h, the bias and variance are of the same order, resulting in $\Delta\kappa \cong 1$. Then we can expect that the obtained value h_{s+} is close to h_{opt}, thus $p \cong 0$, and all parameters for **the key algorithm equation (10.2.14)** are defined. This simple heuristic form has been successfully used in [2] and [3], and *it is highly recommended for most of practical applications.* Estimation of the standard deviation $\sigma(h_s)$ will be discussed within the Numerical example.

(2) When the knowledge about the variance and bias behavior is reliable, i.e., when (10.2.3) accurately describes estimation error, then we can calculate all algorithm parameters. According to the algorithm basic idea, only three confidence intervals, D_{s+-1}, D_{s+}, and D_{s++1}, should be considered. The confidence intervals D_{s+-1} and D_{s+} **should have,** while D_{s+} and D_{s++1} **should not have,** at least one point in common. Assuming that relation (10.2.9) holds, and that the bias is positive, this condition means that the minimal possible value of upper D_{s+-1} bound, (10.2.11), denoted by $\min\{U_{s+-1}\}$, is always greater than or equal to the maximal possible value of the lower D_{s+} bound, denoted by $\max\{L_{s+}\}$, i.e., $\min\{U_{s+-1}\} \geq \max\{L_{s+}\}$. The condition that D_{s+} and D_{s++1} do not intersect is given by $\max\{U_{s+}\} < \min\{L_{s++1}\}$. According to (10.2.9) and (10.2.11) the above analysis results in

$$\begin{aligned} \text{bias}\,(h_{s+-1}) + \Delta\kappa\sigma(h_{s+-1}) &\geq \text{bias}\,(h_{s+}) - \Delta\kappa\sigma(h_{s+}), \\ \text{bias}\,(h_{s+}) + (2\kappa + \Delta\kappa)\sigma(h_{s+}) &< \text{bias}\,(h_{s++1}) - (2\kappa + \Delta\kappa)\sigma(h_{s++1}). \end{aligned} \quad (10.2.15)$$

Since the inequalities are written for the worst case, we can calculate the algorithm parameters by using the corresponding equalities. With (10.2.13) we get

$$\begin{aligned} \Delta\kappa &= 2\kappa/[2^{(m+n)/2} - 1], \\ 2^p &= \left[\Delta\kappa\sqrt{n/m}\left(2^{m/2}+1\right)/\left(1-2^{-n/2}\right)\right]^{2/(m+n)}. \end{aligned} \quad (10.2.16)$$

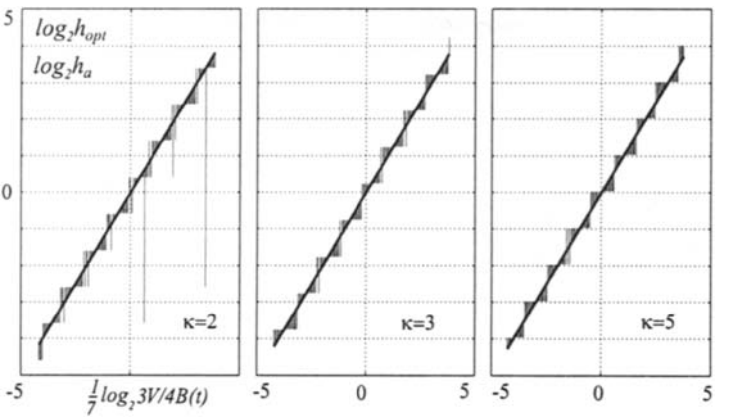


Fig. 10.2.1: Optimal window width (straight thick line), and adaptive window widths (end of the vertical lines, starting from the optimal window width line) for $m = 3$, $n = 4$, $V = 1$. The variance to bias ratio $V/B(t)$ is logarithmically varied. The adaptive width $h_a = h_{s+}/2^{p_1}$ is obtained from h_{s+}, according to (10.2.14), after correction for the corresponding values of p_1 given in Table 10.2.1.

Values of the parameters $\Delta\kappa$ and p for various distributions, i.e., for various values of m and n, are given in Table 10.2.1.

For further, and very **fine tuning of the algorithm parameters**, one may want that the adaptive window is unbiased in logarithmic, instead of in linear scale (due to definition (10.2.10)). The estimation bias and variance are exponential functions with respect to m and n, (10.2.13). Thus the confidence interval limits vary as $2^{(s-s^+)(m+n)/2}$. The mean value for this exponential function, for two successive confidence intervals, for example $(s - s^+) = 0$ and $(s - s^+) = 1$, is $(1 + 2^{(m+n)/2})/2$. It is shifted with respect to the geometrical mean $\sqrt{2^{(m+n)/2}}$ of these two intervals, by approximately $\Delta p \cong \left[\log_2\left((1 + 2^{(m+n)/2})/2\right)\right] \frac{2}{m+n} - \frac{1}{2}$, resulting in the total logarithmic shift $p_1 = p + \Delta p$, presented in Table 10.2.1. Therefore the adaptive window width (as an estimate of the optimal window width) should be $\hat{h}_{opt} = h_{s+}/2^{p_1}$.

Note that the set H of window widths h is a priori assumed. Therefore, as long as we can calculate p_1 , we can use it in the following ways: a) To calculate distribution with the new window width $h_a = h_{s+}/2^{p_1}$ as the best estimate of h_{opt}, b) To remain within the assumed set of $h_s \in H$, and to decide only whether to correct the obtained h_{s+} or not. For example, if $|p_1| \leq 1/2$ the correction is smaller than the window discretization step. Thus, we can use $h_a = h_{s+}$. For $1/2 < p_1 \leq 3/2$ it is better to use $h_a = h_{s+}/2 = h_{s+-1}$, as the adaptive window width value. Fortunately, the loss of accuracy for the adaptive widths h_a, as far as they are of h_{opt} order, is not significant since the MSE varies slowly around its stationary point. Thus, in numerical implementations we can use only the lag windows from the given set H.

10.2.2.2 Illustration

We have simulated the IF estimates as a random variable

$$\hat{f}_h(t) = \mathbf{a}\sqrt{V/h^m} + \sqrt{B(t)h^n} + f_i(t), \tag{10.2.17}$$

having the MSE given by (10.2.3), where $\mathbf{a} = \mathcal{N}(0,1)$ is a Gaussian (zero-mean, unity-variance) random variable, $m = 3$, $n = 4$, and $V = 1$. For the true IF value

$f_i(t)$, at a given t, any constant can be assumed. The bias parameter $B(t)$ in $\hat{f}_h(t)$ logarithmically varies within $\frac{1}{7}\log_2(mV/nB(t)) \in [-4,3]$, with step 0.05.

-For each value of parameter $B(t)$ we have calculated optimal window width according to (10.2.6), and plotted $\log_2(h_{opt})$ as a thick line in Fig. 10.2.1.

-The value of $\hat{f}_h(t)$ was simulated for each $B(t)$ and $h_s \in H$. The assumed set of possible window widths was $H = \{1/16, 1/8, 1/4, 1/2, 1, 2, 4, 8, 16, 32\}$, and $\kappa = 2$. The key algorithm relation (10.2.14) was tested each time, with the known standard deviation $\sigma(h_s) = \sqrt{V/h_s^m}$. The largest value of h_s when the key equation (10.2.14) was still satisfied was denoted by h_{s+}. Value $\Delta\kappa = 0.39$, corresponding to $m = 3$, $n = 4$, $\kappa = 2$, was used (Table 10.2.1). The adaptive values $h_a = h_{s+}/2^{p_1}$, $p_1 = 0.59$ (Table 10.2.1), produced in this way, are connected with the optimal window line, by thin vertical lines in Fig. 10.2.1.

-The same simulation is repeated with $\kappa = 3$ and $\kappa = 5$.

-We can conclude that the presented algorithm almost always chooses the width h_s from H which is the nearest to the optimal one. However, for relatively small $\kappa = 2$ there are few complete misses of the optimal window width, since (10.2.9) is satisfied only with probability $P(2) = 0.95$. For $\kappa = 2$, two successive confidence intervals do not intersect when the bias is small, producing false result, with probability of $2(0.05)^2 \sim 10^{-2}$ order.

(3) The third approach for the parameter $(\kappa + \Delta\kappa)$ estimation is based on the statistical nature of confidence intervals, and a posteriori check of the fitting quality [2]. This approach is beyond the scope of this article.

10.2.3 Numerical Example

In the example we assumed a signal of (10.2.1) form, with the given IF,

$$f_i(n\Delta t) = 128\arctan(250(n\Delta t - 0.5))/\pi + 128,$$

and the phase $\phi(n\Delta t) = 2\pi\Delta t\sum_{m=0}^{n} f_i(m\Delta t)$. The signal amplitude was $a = 1$, and $20\log(a/\sigma_\epsilon) = 10[dB]$ $(a/\sigma_\epsilon = 3.16)$. Considered time interval was $0 \le n\Delta t \le 1$, with $\Delta t = 1/1024$. The IF is estimated by using the discrete WD with a rectangular lag-window, $W_x^h(t,f) = DFT_{n\to f}[w_h(n\Delta t)x(t+n\Delta t)x^*(t-n\Delta t)]$, calculated with the standard FFT routines.

The algorithm is implemented as follows:

(1) A set H of window widths h_s, corresponding to the following number of signal samples $\mathbf{N} = \{4, 8, 16, 32, 64, 128, 256, 512\}$, is assumed. In order to have the same number of frequency samples, as well as to reduce the quantization error, all windows are zero-padded up to the maximal window width.

(2) For a given time instant $t = n\Delta t$, the WDs are calculated starting from the smallest toward the wider window widths h_s.

(3) The IF is estimated using equation (10.2.2) and $W_x^{h_s}(t,f)$.

(4) The confidence intervals intersection, (10.2.14), is checked for the estimated IF, $\hat{f}_{h_s}(t)$, and $\sigma(h_s) = \sqrt{3\sigma_\epsilon^2\Delta t/(2\pi^2 a^2 h_s^3)}$ with, for example, $\kappa + \Delta\kappa = 6$, when $p_1 \cong 1$, and $P(\kappa) \to 1$ (see Table 10.2.1, and the Comment that follows).

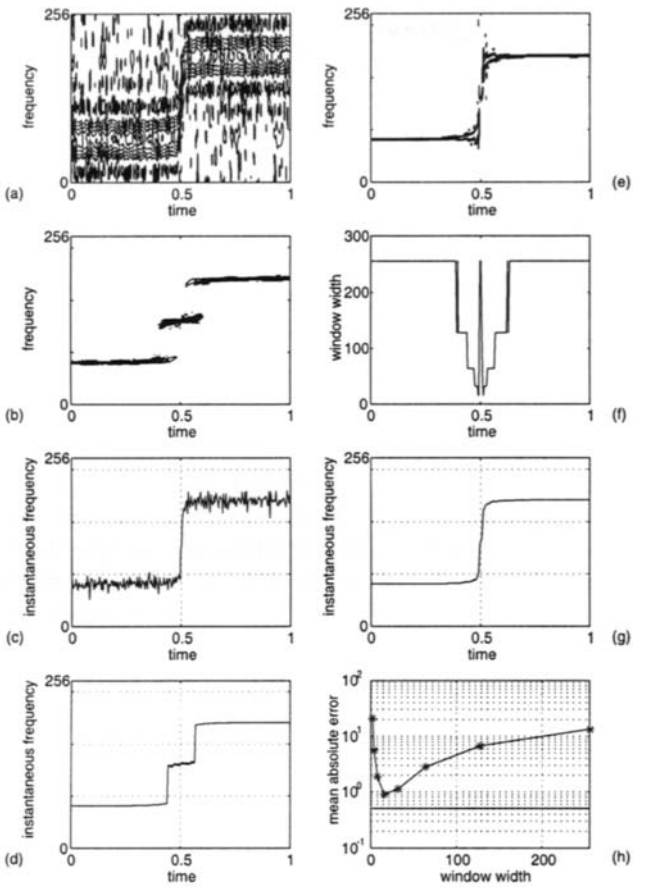


Fig. 10.2.2: Time-frequency analysis of a noisy signal: (a) Wigner distribution with $N = 16$; (b) Wigner distribution with $N = 256$; (c) Estimated instantaneous frequency using the Wigner distribution with $N = 8$; (d) Estimated instantaneous frequency using the Wigner distribution with $N = 256$; (e) Wigner distribution with adaptive window width; (f) Adaptive window width as a function of time; (g) Estimated instantaneous frequency using the Wigner distribution with the adaptive window width; (h) Absolute mean error as a function of the window width; the line represents the mean absolute error value for the adaptive window width.

(5) The adaptive window width $h_a = h_{s+}/2$ is obtained from the last $h_s = h_{s+}$ when (10.2.14) is still satisfied. Back to (2).

Comment: Estimation of the signal and noise parameters a and σ_ϵ^2 can be done by using $|\hat{a}|^2 + \hat{\sigma}_\epsilon^2 = \frac{1}{N}\sum_{n=1}^{N} |x(n\Delta t)|^2$. The variance is estimated by $\hat{\sigma}_\epsilon^2 = \hat{\sigma}_{\epsilon r}^2 + \hat{\sigma}_{\epsilon i}^2$, where

$$\hat{\sigma}_{\epsilon r,i} = \{ \text{median} \, (|x_{r,i}(n\Delta t) - x_{r,i}((n-1)\Delta t)| : n = 2, .., N)\}/(0.6745\sqrt{2}), \qquad (10.2.18)$$

with $x_r(n\Delta t)$ and $x_i(n\Delta t)$ being the real and imaginary part of $x(n\Delta t)$. It is assumed that N is large, and Δt is small [2–5]. For this estimation we oversampled the signal by factor of four.

The WDs with constant window widths $N_s = 16$ and $N_s = 256$ are presented

in Fig. 10.2.2(**a**), and Fig. 10.2.2(b), respectively. The IF estimates using the WDs with constant window widths $N_s = 8$, and $N_s = 256$ are given in Fig. 10.2.2(c) and Fig. 10.2.2(d). Fig. 10.2.2(e) shows the WD with adaptive window width. Values of the adaptive window width, determined by the algorithm, are presented in Fig. 10.2.2(f). We can see that when the IF variations are small the algorithm uses the widest window in order to reduce the variance. Around the point $n\Delta t = 0.5$, where the IF variations are fast, the windows with smaller widths are used. The IF estimate with adaptive window width is presented in Fig. 10.2.2(g). Mean absolute error, normalized to the discretization step, is shown in Fig. 10.2.2(h) for each considered window width. The line represents value of the mean absolute error for the adaptive window width.

10.2.4 Summary and Conclusions

An algorithm that can produce an accurate estimate of the optimal window width, without using the bias value, is presented. The IF estimates obtained by using this algorithm and the WD have lower error than by using the best constant-window width, which also is not known in advance. Additional examples, including distributions with adaptive order, the WD as a spectrum estimator, algorithm application to the sensor array signal tracking, as well as other realization details can be found in [2–5, 7].

References

[1] A. Goldenshluger and A. Nemirovski, "Adaptive denoising of signals satisfying differential inequalities," *IEEE Trans. Information Theory*, vol. 43, pp. 873–889, May 1997.

[2] V. Katkovnik and L. Stanković, "Periodogram with varying and data-driven window length," *Signal Processing*, vol. 67, pp. 345–358, 30 June 1998.

[3] V. Katkovnik and L. Stanković, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Trans. Signal Processing*, vol. 46, pp. 2315–2325, September 1998.

[4] L. Stanković and V. Katkovnik, "Instantaneous frequency estimation using the higher order L-Wigner distributions with the data driven order and window length," *IEEE Trans. Information Theory*, vol. 46, pp. 302–311, January 2000.

[5] L. Stanković and V. Katkovnik, "The Wigner distribution of noisy signals with adaptive time-frequency varying window," *IEEE Trans. Signal Processing*, vol. 47, pp. 1099–1108, April 1999.

[6] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals," *Proc. IEEE*, vol. 80, pp. 520–538, April 1992.

[7] A. B. Gershman, L. Stanković, and V. Katkovnik, "Sensor array signal tracking using a data-driven window approach," *Signal Processing*, vol. 80, pp. 2507–2515, December 2000.

10.3 IF ESTIMATION FOR MULTICOMPONENT SIGNALS[0]

10.3.1 Time-Frequency Peak IF Estimation

There is a wide range of applications where we encounter signals comprised of M components with different IF laws $f_m(t)$ and different envelopes $a_m(t)$, in additive noise. It is often desired from such an observed signal, to determine the number of components M, the IF law of each component and the corresponding envelope $a_m(t)$. This can be achieved by representing the observed signal $z(t)$ in a time-frequency (t-f) domain and use time-frequency filtering methods to recover the individual components [1]. Another approach involves extending algorithms for IF estimation of monocomponent FM signals to the case of multicomponent signals and design an algorithm that simultaneously tracks the various IF components of the observed signal [2,3]. Both approaches require the use of time-frequency distributions (TFDs) with very specific properties such as high time-frequency localization of the IF components and high reduction of cross-terms interferences.

The basic concept of instantaneous frequency is described in Pt. 1 of reference [4] and in Chapter 1 of this book. Methods of IF estimation are reported in Pt. 2 of [4] and in Chapter 10. Essential results are reproduced below, for greater clarity.

10.3.1.1 Spectrogram Peak IF estimation

Various approaches for IF estimation of monocomponent signals exist [4]. Most of these algorithms are suited to a particular class of signals, and both fixed and adaptive algorithms have been proposed. Our aim here is to approach the problem from a general viewpoint in order to define a general IF methodology that would be suitable for the largest class of signals found in practical applications. To illustrate this approach, we thus consider from the outset multicomponent signals in additive noise, which can be expressed as follows:

$$z(t) = \sum_{m=1}^{M} z_m(t) + \epsilon(t) = \sum_{m=1}^{M} a_m(t) e^{j\phi_m(t)} + \epsilon(t) \tag{10.3.1}$$

where the amplitudes $a_m(t)$ are the component amplitudes, $\phi_m(t)$ are the component phases, and $\epsilon(t)$ is a complex-valued white Gaussian noise process of independent and identically distributed (i.i.d.) real and imaginary parts with total variance σ_ϵ^2. The individual IF laws for each component are given by [1]:

$$f_m(t) = \frac{1}{2\pi} \frac{d\phi_m(t)}{dt} \quad ; \; m = 1, ..., M. \tag{10.3.2}$$

[0]Authors: **Zahir M. Hussain** and **Boualem Boashash**, Signal Processing Research Centre, Queensland University of Technology, 2 George Street, Brisbane, Queensland 4000, Australia (zahir.hussain@rmit.edu.au, b.boashash@qut.edu.au; Z. M. Hussain is now at the School of Electrical and Computer Systems Engineering, RMIT, GPO Box 2476V, Melbourne, Victoria 3001, Australia). Reviewers: LJ. Stanković and V. Katkovnik.

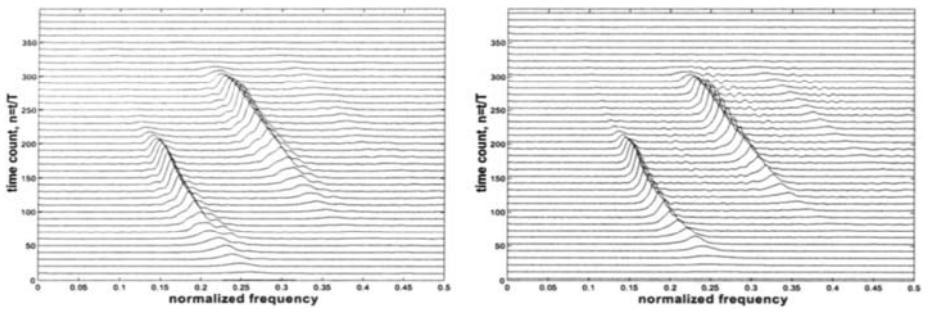


Fig. 10.3.1: Left: The spectrogram of a bat signal using a small analysis window. Right: The modified B-distribution of the same signal with parameter $\beta = 0.05$. Total signal length is $N = 400$ and sampling interval $\Delta t = 1$. The spectrogram cannot show the weakest component.

A conventional approach to represent and analyze such signals for IF estimation is to take the spectrogram of $z(t)$ and search for the peaks in the t-f domain (see Article 10.1). Curves formed by a continuum of these peaks describe the IF laws of the individual components of the observed signal $z(t)$, as illustrated in Fig. 10.3.1 using a bat signal. Analytically, this can be expressed as follows:

$$\hat{f}_m(t) = \arg[\max_f \rho_m(t, f)]; \quad 0 \leq f \leq f_s/2 \tag{10.3.3}$$

where $\rho_m(t, f)$ is the m^{th} peak of the spectrogram.

This spectrogram based approach has several advantages: it is easy to understand, easy to use, and there are no cross-terms producing unwanted interferences. A major disadvantage though is that the time-frequency resolution of the spectrogram for closely spaced components is often poor, especially if one of the components is much weaker, as illustrated in Fig. 10.3.2 for a two-component linear FM signal with one component weaker than the other. Fig. 10.3.2 also compares the performance of the spectrogram with the modified B-distribution (MBD) that was designed specifically for multicomponent IF estimation, as discussed later.

10.3.1.2 Peaks of WVD, PWVD, and RIDs

To improve upon the resolution of the spectrogram, various TFDs were proposed for IF estimation, one of the most important being the Wigner-Ville distribution (WVD). IF estimation using the peak of the Wigner-Ville distribution (WVD) is optimal for linear FM signals with high to moderate signal-to-noise ratios (SNRs) [4], but its performance degrades significantly at low SNRs, and in this case the cross WVD (XWVD) peak can be used as an IF estimator [5]. For polynomial FM signals it was shown that the polynomial WVD (PWVD) gives the best performance, especially at high SNRs (see [6] and Article 10.4). However, both WVD and PWVD suffer from cross-terms when used to analyze multicomponent signals. These cross-terms generate artifacts that obscure the (t-f) representation of the signal, leading to

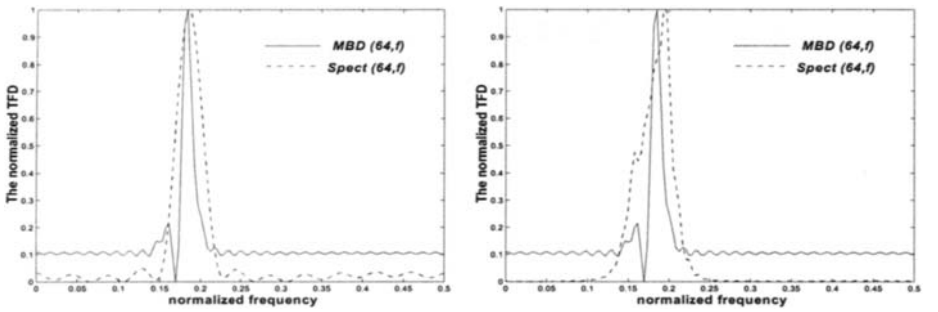


Fig. 10.3.2: Performance comparison between the spectrogram $\text{Spect}(t, f)$ and the modified B-distribution (MBD) for $\beta = 0.06$ using a two-component noise-free linear FM signal at the sampling instant $n = t/\Delta t = 64$. Total signal length is $N = 128$ and the sampling interval is $\Delta t = 1$. The right component is five times larger in amplitude than the left component. Left: Spectrogram with small analysis window length ($\Delta = 23$). Right: Spectrogram with large analysis window length ($\Delta = 83$). In both cases the spectrogram fails to resolve the two components. In addition, time resolution is bad for a large window length.

the development of reduced interference distributions (RIDs) to remedy the problem [7]. Straightforward IF estimation using the peak of RIDs give an IF estimate that is biased from the true IF law, and this bias is different for different RIDs. Although reduced, cross-terms still exist and can obscure weak components, hence the need to define special purpose RIDs with efficient cross-terms reduction, high time-frequency resolution and minimum bias from the true IF laws, such as the MBD [3,8].

10.3.2 Properties of IF Estimates Based on Quadratic TFDs

10.3.2.1 IF Estimates and Window Length

We consider an analytic signal $z(t)$ of the form $z(t) = ae^{j\phi(t)} + \epsilon(t)$ where the amplitude a is constant, and $\epsilon(t)$ is a complex-valued white Gaussian noise with independent identically distributed (i.i.d.) real and imaginary parts with total variance σ_ϵ^2. The IF of $z(t)$ is given by eq. (10.3.2), and it is assumed to be an arbitrary, smooth and differentiable function of time with bounded derivatives of all orders. The general equation for quadratic time-frequency representation of the signal $z(t)$ is given by [1]

$$\rho_z(t, f) = \underset{\tau \to f}{\mathcal{F}}[G(t, \tau) \underset{(t)}{*} K_z(t, \tau)]$$

where $G(t, \tau)$ is the time-lag kernel, $K_z(t, \tau) = z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})$ is the signal kernel or the instantaneous autocorrelation function (IAF), and $\underset{(t)}{*}$ denotes time convolution. For smoothing and localization on the IAF we apply a window function $w_h(\tau) = \frac{\Delta t}{h} w(\frac{\tau}{2h})$ on the instantaneous autocorrelation $K_z(t, \tau)$, where $w(t)$ is a real-valued symmetric window with unity length, i.e., $w(t) = 0$ for $\mid t \mid > \frac{1}{2}$; hence the window length is h.

The TFD is now dependent on the window length h as follows:

$$\rho_{z,h}(t,f) = \underset{\tau \to f}{\mathcal{F}}[w_h(\tau)G(t,\tau) \underset{(t)}{*} K_z(t,\tau)]. \tag{10.3.4}$$

If $\rho_{z,h}(t,f)$ is discretized over time, lag, and frequency then we have

$$\rho_{z,h}(n,k) = \sum_{l=-N_s}^{N_s-1} \sum_{m=-N_s}^{N_s-1} w_h(m\Delta t)K_z(l\Delta t, 2m\Delta t)G(n\Delta t - l\Delta t, 2m\Delta t)e^{-j2\pi\frac{km}{2N_s}} \tag{10.3.5}$$

where $2N_s$ is the number of samples and Δt is the sampling interval.

The IF estimate is a solution of the following optimization

$$\hat{f}_h(t) = \arg[\max_f \rho_{z,h}(t,f)] \quad ; \; 0 \le f \le f_s/2 \tag{10.3.6}$$

where $f_s = 1/\Delta t$ is the sampling frequency.

10.3.2.2 Bias and Variance of the IF Estimate

By extending the results in [9], the estimation bias and variance are found to be [3]

$$E[\Delta\hat{f}_h(t)] = \frac{L_h(t)}{2F_h}, \quad \text{var}(\Delta\hat{f}_h(t)) = \frac{\sigma_\epsilon^2}{2 \mid a \mid^2}[1 + \frac{\sigma_\epsilon^2}{2 \mid a \mid^2}]\frac{E_h}{F_h^2} \tag{10.3.7}$$

where

$$\begin{aligned}
\Delta\hat{f}_h(t) = \tfrac{1}{2\pi}\phi'(t) - \hat{f}(t) \; ; \quad F_h &= \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w_h(m\Delta t)(2\pi m\Delta t)^2 G(u, 2m\Delta t)du \\
L_h(t) &= \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w_h(m\Delta t)\Delta\phi(u, m\Delta t)(2\pi m\Delta t)G(t-u, 2m\Delta t)du \\
E_h &= \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} w_h(m\Delta t)^2(2\pi m\Delta t)^2 G(u, 2m\Delta t)du
\end{aligned} \tag{10.3.8}$$

where $\Delta\phi(t,\tau) = \phi(t+\tau/2) - \phi(t-\tau/2) - \tau\phi^{'}(t)$.

Equations (10.3.7) and (10.3.8) indicate that the bias and the variance of the estimate depend on the lag window length h for any kernel $G(t,\tau)$. To see how the bias and the variance vary with h, asymptotic analysis as $\Delta t \to 0$ is necessary for the chosen TFD.

10.3.2.3 TFD Properties Needed for Multicomponent IF Estimation

The results above indicate that a general method for IF estimation of multicomponent FM signals in additive Gaussian noise that is based on quadratic time-frequency distributions requires the quadratic TFDs to satisfy the following conditions [3]:

(1) $\rho(t, f)$ should have a high time-frequency resolution while suppressing cross-terms efficiently so as to give a robust IF estimate for mono- and multicomponent FM signals.
(2) $\rho(t, f)$ should enable amplitude estimation for the individual components of the signal, as the amplitude is necessary for evaluating the variance of the IF estimate for each component [2, 3, 9], and to allow for the reconstruction of the individual components of the signal.
(3) the choice of the lag window length should lead to a bias-variance tradeoff (see eqs. (10.3.13) and (10.3.14)).

Although some TFDs, like Choi-Williams distribution (CWD) and the spectrogram, can satisfy some of these conditions, they do not meet the second requirement, i.e. allowing direct amplitude estimation. The design of TFDs which satisfies all of these required properties is considered next.

10.3.3 Design of Quadratic TFDs for Multicomponent IF Estimation

10.3.3.1 Desirable Time-Lag Kernel

A TFD referred to as the B-distribution (BD) was proposed and shown to be superior to other fixed-kernel TFDs in terms of cross-terms reduction and resolution enhancement [10]. As it does not allow direct component amplitudes estimation [2], as per the second condition on TFDs required for multicomponent IF estimation listed above, the BD kernel was modified as [3]

$$G(t, \tau) = \mathrm{G}_\beta(t) = k_\beta / \cosh^{2\beta}(t) \tag{10.3.9}$$

where β is a real positive number and $k_\beta = \Gamma(2\beta)/(2^{2\beta-1}\ \Gamma^2(\beta))$, Γ stands for the gamma function. This modified B-distribution MBD(t, f) is also referred to as the hyperbolic T-distribution (HTD) in [8].

10.3.3.2 Relevant Properties of the Modified B-Distribution (MBD)

Most of the desirable properties of time-frequency distributions relevant to IF estimation (as explained in [1] and [4]) are satisfied by the MBD kernel. In particular, realness, time-shift and frequency shift invariance, frequency marginal and group delay, and the frequency support properties are satisfied. The time support property is not strictly satisfied, but it is approximately true [3]. The three required conditions listed in Subsection 10.3.2.3 are discussed in detail below.

(1) Reduced interference and resolution: This property is satisfied by MBD. For example, consider the sum of two complex sinusoidal signals $z(t) = z_1(t) + z_2(t) = a_1 e^{j(2\pi f_1 t + \theta_1)} + a_2 e^{j(2\pi f_2 t + \theta_2)}$ where a_1, a_2, θ_1 and θ_2 are constants. The TFD of the signal $z(t)$ is obtained as [3]

$$\mathrm{MBD}(t, f) = a_1^2\ \delta(f - f_1) + a_2^2\ \delta(f - f_2) + 2a_1 a_2 \gamma_\beta(t)\ \delta[f - (f_1 + f_2)/2] \tag{10.3.10}$$

where $\gamma_\beta(t) = |\Gamma(\beta + j\pi(f_1 - f_2))|^2 \cos\Big(2\pi(f_1 - f_2)t + \theta_1 - \theta_2\Big)/\Gamma^2(\beta)$. The cross-terms are oscillatory in time and depend on the frequency separation between signal components. If f_1 and f_2 are well separated then the term $|\Gamma(\beta + j\pi(f_1 - f_2))|^2$ can be substantially reduced, while $\Gamma^2(\beta)$ can be made high if β is small. When f_1 and f_2 are not well separated, the MBD still performs better than most quadratic TFDs (see Article 7.4).

(2) Direct amplitude and IF estimation: The MBD allows direct IF estimation by peak localization, i.e., at any time instant t, it has an absolute maximum at $f = \frac{1}{2\pi}\frac{d\phi(t)}{dt}$ for linear FM signals . For non-linear FM signals this estimate is biased, but this bias can be accounted for in the adaptive IF estimation, as presented next.

For an FM signal of the form $z(t) = a\, e^{j\phi(t)}$, the MBD is approximated by [3]

$$\mathrm{MBD}(t,f) \approx\mid a \mid^2 \int_{-\infty}^{\infty} \mathrm{G}_\beta(t-u)\, \delta\,[\frac{1}{2\pi}\phi^{'}(u) - f]du = |a|^2 \mathrm{G}_\beta(t - \psi(f))\psi^{'}(f) \tag{10.3.11}$$

where ψ is the inverse of $\frac{1}{2\pi}\phi^{'}$, i.e., $\frac{1}{2\pi}\phi^{'}(\psi(f)) = f$. Assuming that $\psi^{'}(f)$ is not a highly peaked function of f and knowing that $\mathrm{G}_\beta(t - \psi(f))$ is peaked at $t = \psi(f)$, the absolute maximum of $\mathrm{MBD}(t,f)$ for any time t would be at $\psi(f) = t$, or $f = \frac{1}{2\pi}\phi^{'}(t)$, which is the IF of the FM signal $z(t)$. For non-linear FM signals, the energy peak of the MBD is actually biased from the IF because of the extra term $\sum_{k=3(k\text{ odd})}^{\infty} \frac{\tau^{k-1}}{k!2^{k-1}}\phi^{(k)}(u)$. The major contribution in this term is due to $\phi^{(3)}(u)$. Therefore at the instants of rapid change in the IF law the bias is not negligible and eq. (10.3.11) would not be an accurate approximation to the MBD unless suitable windowing in the lag direction is used.

For linear FM signals we have $\phi^{(k)}(t) = 0$ for $k \geq 3$. Assuming that $z(t) = ae^{j2\pi(f_o t + \frac{\beta_o}{2}t^2)}$, where f_o and β_o are constants, we have

$$\mathrm{MBD}(t,f) = \frac{1}{\beta_o}|a|^2 \mathrm{G}_\beta\Big(t - \frac{1}{\beta_o}(f - f_o)\Big) \tag{10.3.12}$$

which has an absolute maximum at $f = f_o + \beta_o t$, the IF of the linear FM signal $z(t)$. As $\beta_o \to 0$, i.e., $z(t)$ approaches a sinusoid, we have $\mathrm{MBD}(t,f) \to |a|^2\delta(f - f_o)$, in accordance with eq. (10.3.10).

As for amplitude estimation, eqs. (10.3.11) and (10.3.12) indicate that the MBD can support amplitude estimation [3].

(3) Asymptotic Formulas Using MBD: The asymptotic formulas for the variance and the bias as $\Delta t \to 0$ using a rectangular lag window are given by [3]:

$$\mathrm{var}(\Delta\hat{f}_h(t)) = \frac{3\sigma_\epsilon^2}{2\pi^2 \mid a \mid^2}[1 + \frac{\sigma_\epsilon^2}{2|a|^2}]\frac{\Delta t}{h^3} \tag{10.3.13}$$

and

$$E(\Delta\hat{f}_h(t)) = \frac{h^2}{80}\int_{-\infty}^{\infty}\frac{\lambda(u)du}{\cosh^{2\beta}(t-u)}\ ;\quad E(\Delta\hat{f}_h(t)) \leq \frac{M_2}{40}h^2 \tag{10.3.14}$$

where $\lambda(t) = f^{(2)}(t+\tau_1) + f^{(2)}(t-\tau_1)$, $f^{(2)}(t)$ is the second derivative of the IF, and $\sup_t |f^{(2)}(t)| \leq M_2$. For small h, the optimal window length that minimizes the mean squared error is obtained by extending the result in [6] as:

$$\mathrm{h_{opt}}(t) = \Big[\frac{1800\sigma_\epsilon^2\Delta t(1+\frac{\sigma_\epsilon^2}{2|a|^2})}{\pi^2|a|^2(f^{(2)}(t) \underset{(t)}{*} 1/\cosh^{2\beta}(t))^2}\Big]^{\frac{1}{7}}. \tag{10.3.15}$$

Thus, the optimal window length depends on the second derivative of the instantaneous frequency $f^{(2)}(t)$, which is time and signal dependent. Eqs. (10.3.13) and (10.3.14) indicate that the variance and bias of the IF estimate using MBD have the same rates of change with respect to the window length h as those using WVD [9].

10.3.3.3 Examples of Quadratic TFDs Suitable for Multicomponent IF Estimation

TFDs with time-only (or lag-independent) kernels constitute a subclass of the quadratic class of TFDs. These TFDs share the important properties of cross-terms suppression, high-resolution, and supporting amplitude estimation, making them well suited for multicomponent IF estimation. The modified B-distribution (a. k. a. hyperbolic T-distribution) was defined earlier in this section. Another example is the exponential T-distribution, which is defined in terms of its time-lag kernel as [8]

$$G(t,\tau) = \mathrm{G}_\eta(t) = \sqrt{\eta/\pi}\exp(-\eta t^2)$$

where η is a real parameter and $\sqrt{\eta/\pi}$ is a normalization factor. The resulting TFD used for multicomponent IF estimation is then given by eq. (10.3.4).

10.3.4 An Adaptive Algorithm for Multicomponent IF Estimation

Eq. (10.3.15) shows that the optimal window length using the MBD is a function of time and depends on the second derivative of the IF law $f^{(2)}(t)$; it decreases when the IF law $f(t)$ has a high variation. Hence a time-varying window length is needed to optimize the estimation. The Stanković-Katkovnik adaptive algorithm developed in [9] for monocomponent FM signals can be used since the IF estimation variance is a continuously decreasing function of h while its bias is continuously increasing, as shown in eqs. (10.3.13) and (10.3.14); see also Article 10.2. These conditions are necessary for bias-variance tradeoff such that the algorithm converges at the optimum window length that resolves this tradeoff. It is shown in [9] that, if h is

small enough then the IF estimate $\hat{f}_h(t)$ is inside the confidence interval D defined as follows

$$D = [\hat{f}_h(t) - 2\kappa\sqrt{\text{var}(\Delta\hat{f}_h(t))}, \hat{f}_h(t) + 2\kappa\sqrt{\text{var}(\Delta\hat{f}_h(t))}] \tag{10.3.16}$$

with Gaussian probability $P(\kappa)$, κ being a parameter (usually 2); while for large h, $\hat{f}_h(t)$ is outside D. Hence, if we consider an increasing sequence of window lengths $\{h_r | r = 1 : N\}$ (N being the number of samples) and calculate the MBD (and hence $\hat{f}_{h_r}(t)$) for each h_r, then all $\{D_r\}$ have at least one point in common (which is $\hat{f}_{h_r}(t)$) if h_r is sufficiently small. The first h_r for which D_{r-1} and D_r have no point in common is considered optimal as it decides the bias-variance tradeoff.

The estimates for the amplitude of the signal a and the variance of noise σ_ϵ^2 [used in eq. (10.3.13) and implicitly in eq. (10.3.16)] were given in [9] as:

$$\widehat{a}^2 + \widehat{\sigma}_\epsilon^2 = \frac{1}{N}\sum_{n=1}^{N} \mid z(n\Delta t) \mid^2 \;\; ; \;\; \widehat{\sigma}_\epsilon^2 = \frac{1}{2N}\sum_{n=2}^{N} \mid z(n\Delta t) - z((n-1)\Delta t) \mid^2 \tag{10.3.17}$$

where N is the number of samples. For further details of this adaptive algorithm see [9].

For a multicomponent analytic signal of the form stated in eqs. (10.3.1) and (10.3.2), with $\{a_m\}$ constant, we can use the extension of the monocomponent IF estimation algorithm in [9] for multicomponent signals as described in [2,3]. This algorithm tracks component maxima in the time-frequency plane and requires a threshold $\mathcal{T}_\rho(t)$ so as to ignore the local maxima caused by the cross-terms and windowing. In fact, $\mathcal{T}_\rho(t)$ is application and distribution dependent.

The algorithm requires the knowledge of the confidence intervals $D_{r,m}$ for each component, where r refers to the window length (h_r) and m refers to the signal component. The calculation of $D_{r,m}$ depends on the estimation of the individual amplitudes a_m of the components. Using the MBD, the actual amplitudes $|\hat{a}_m|$ can be estimated as shown in [3]. Using $|\hat{a}_m|^2$ and $\hat{\sigma}^2$ to calculate $\text{var}(\Delta\hat{f}_h(t))$ [given by eq. (10.3.13) for MBD(t,f)], we can define the confidence intervals $\{D_{r,m}\}$ for all components as in [2,3]. The IF $f_m(t)$ is contained in at least one of the confidence intervals $\{D_{r,m}\}$ if h_r is sufficiently small, and the optimal window length is the first h_r (from the increasing sequence $\{h_r | r = 1 : N\}$) for which $D_{r-1,m}$ and $D_{r,m}$ have no point in common.

Example: We consider a three-component FM signal $z(n\Delta t)$ with amplitudes $a_1 = 0.5$, $a_2 = 1$, and $a_3 = 1.5$ and non-linear IF laws: $f_1 = 47 + 2.5\sinh^{-1}(20(n\Delta t - 0.2))$, $f_2 = 30 + 2.5\text{sgn}(40(n\Delta t - 0.6))$, and $f_3 = 10 + 2\sin(10(n\Delta t - 0.7))$, with $SNR = 15$ dB, $\beta = 0.1$, $\kappa = 2$, $0 \leq n\Delta t \leq 1$, and $\Delta t = 1/128$. Fig. 10.3.3 shows the result of the tracking adaptive algorithm for IF estimation of $z(n\Delta t)$ using the peaks of the MBD and the spectrogram.

Fig. 10.3.4 shows the conventional peak IF estimation for the same signal using MBD and the spectrogram. Both TFDs fail to give a robust IF estimation at the

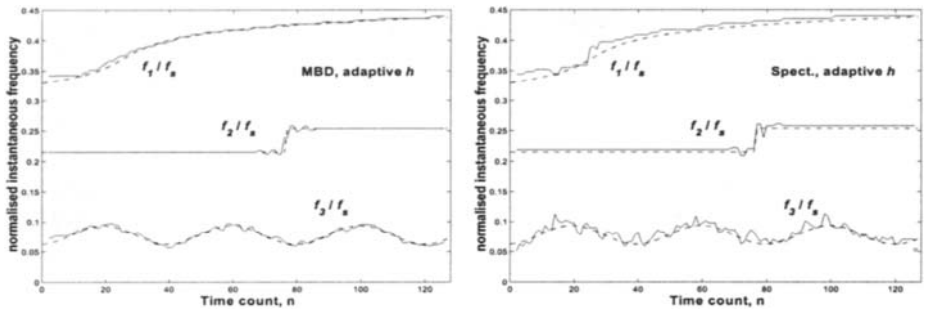


Fig. 10.3.3: Left: Adaptive IF estimation using the peak of the MBD for a three-component FM signal with total length $N = 128$, $SNR = 15$ dB, and $\Delta t = 1/128$. Dashed lines represent the true IF laws. Right: Adaptive IF estimation using the peak of the spectrogram for the same signal, assuming that component amplitudes are known.

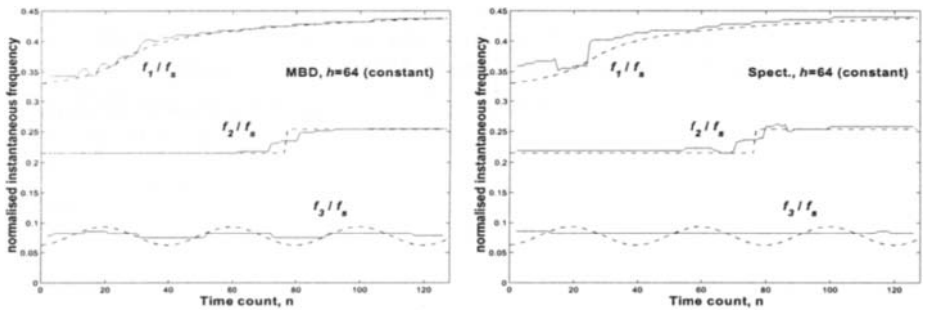


Fig. 10.3.4: Left: IF estimation for the three-component FM signal as in Fig. 10.3.3 using the conventional (constant window) peak estimation. Left: MBD peak IF estimation. Right: Spectrogram peak IF estimation. (In addition to failure of IF estimation at instants of rapid frequency change, the spectrogram has poor tracking performance. Note also that both methods cannot track the continuously varying frequency of the third component.)

instants of rapid frequency change. In addition, the spectrogram has poor time-frequency resolution, both in adaptive and constant-window IF estimation.

10.3.5 Summary and Conclusions

Concurrent IF estimation of the separate components of a multicomponent FM signal using TFD peaks location requires conditions on the selection of a suitable quadratic TFD. Required properties are: (1) high time-frequency resolution while suppressing cross-terms efficiently, (2) the TFD to enable direct amplitude estimation for the individual components, (3) the variance of the IF estimate using the TFD should be a continuously decreasing function of the lag window length while the bias is continuously increasing. Quadratic time-frequency distributions that satisfy these conditions were presented and discussed. A constant-window tracking

algorithm may not give a robust IF estimate if the IF changes rapidly with time due to the effect of the higher-order derivatives of the IF law. Hence an adaptive algorithm is used for robust multicomponent IF estimation.

References

[1] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[2] Z. M. Hussain and B. Boashash, "Adaptive instantaneous frequency estimation of multicomponent FM signals," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, vol. II, pp. 657–660, Istanbul, 5–9 June 2000.

[3] Z. M. Hussain and B. Boashash, "Multi-component IF estimation," in *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pp. 559–563, Pocono Manor, PA, 14–16 August 2000.

[4] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals; Part 2: Algorithms and applications," *Proc. IEEE*, vol. 80, pp. 519–568, April 1992. With introductions by Michael J. Riezenman.

[5] B. Boashash and P. O'Shea, "Use of the cross Wigner-Ville distribution for estimation of instantaneous frequency," *IEEE Trans. Signal Processing*, vol. 41, pp. 1439–1445, March 1993.

[6] 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 Processing*, vol. 47, pp. 2480–2490, September 1999.

[7] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. 40, pp. 402–412, February 1992.

[8] Z. M. Hussain and B. Boashash, "Design of time-frequency distributions for amplitude and IF estimation of multicomponent signals," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA '01)*, vol. 1, pp. 339–342, Kuala Lumpur, 13–16 August 2001.

[9] V. Katkovnik and L. Stanković, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Trans. Signal Processing*, vol. 46, pp. 2315–2325, September 1998.

[10] B. Barkat and B. Boashash, "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis," *IEEE Trans. Signal Processing*, vol. 49, pp. 2232–2239, October 2001.

10.4 ANALYSIS OF POLYNOMIAL FM SIGNALS IN ADDITIVE NOISE [0]

In this article, we focus our study on the analysis of polynomial frequency modulated signals (a.k.a. polynomial phase signals), corrupted by additive Gaussian noise. A noiseless complex polynomial frequency modulated (FM) signal, $z(t)$, can be expressed as:

$$z(t) = A \cdot \exp(j\phi(t)) = A \cdot \exp\left\{ j \sum_{k=0}^{P} a_k t^k \right\} \qquad t = [0, T]. \tag{10.4.1}$$

Here A is the signal amplitude, $\phi(t)$ is the signal phase, P is the polynomial phase order, the $\{a_k\}$ are arbitrary parameters and T is the signal duration. Polynomial FM signals are found in many important applications such as radar, sonar and telecommunications. The FM law of the signal corresponds to the instantaneous frequency (IF) trajectory, defined by:

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} = \frac{1}{2\pi} \sum_{k=1}^{P} k a_k t^{k-1} \qquad t = [0, T]. \tag{10.4.2}$$

In most applications, polynomial FM signals are immersed in noise. In additive noise, the signal model can be re-written as:

$$y(t) = z(t) + w(t) = A \cdot \exp\left\{ j \sum_{k=0}^{P} a_k t^k \right\} + w(t) \qquad t = [0, T], \tag{10.4.3}$$

where $w(t)$ is complex, Gaussian noise. It is often desired to estimate the IF of the class of signals defined by Equation (10.4.3). In the following text, we describe a method which can perform this task in an accurate way.

10.4.1 The Polynomial Wigner-Ville Distributions

10.4.1.1 IF Estimation

If the order of the polynomial phase in Equation (10.4.1) is equal to 1, then the signal, $z(t)$, is a complex sinusoidal (i.e. linear phase) signal. For this special case, the IF is a constant independent of the time variable. It can easily be estimated by using classical spectrum analysis methods. For instance, the frequency at which the Fourier transform of the data has its peak can be used as an estimate of the IF. The Fourier transform proves to be a useful tool in this case because it concentrates the signal energy in frequency (i.e. about the IF), while dispersing the noise energy

[0]Authors: **P. O'Shea**, School of Electrical and Electronic Systems Engineering, QUT, Brisbane, Australia, 4001 (pj.oshea@qut.edu.au), and **B. Barkat**, Nanyang Technological University, School of Electrical & Electronic Engineering, Block S2, Nanyang Avenue, Singapore 63978 (ebarkat@ntu.edu.sg).

over a wide band. In the Fourier (frequency) domain, the signal energy tends to be quite peaked, while the noise contribution tends to be broad and relatively low. The Fourier transform peak actually provides an *optimal* estimator for the frequency of a complex sinusoidal data sequence, assuming the noise is white and Gaussian [1].

If the order of the polynomial phase in Equation (10.4.1) is larger than 1, then the signal, $z(t)$, is a non-linear phase signal. For this general case the IF is not constant, but varies with time. The Fourier transform in this case is far less effective for IF estimation; the signal energy is spread over a continuum of frequencies and the signal-to-noise ratio (SNR) in the spectrum is reduced compared with the complex sinusoidal case. It would be convenient if an operator could be applied which could transform the non-linear phase signal into a linear phase one with frequency equal to the IF. Then the Fourier transform could be applied to estimate the IF.

One technique that can transform a polynomial FM signal into a sinusoid for a given time instant was proposed in [2] and is described in Articles 5.4 and 5.5. The general form for this transform is

$$K_z^P(t,\tau) = \prod_{i=1}^{I} z(t+c_i\tau)^{k_i} \cdot z^*(t-c_i\tau)^{k_i} \qquad t=[0,T], \tag{10.4.4}$$

where the c_i are the coefficients of the transformation, K_z^P, and I is the total number of distinct c_i values. The order of the transformation is $q = \sum_{i=1}^{I} 2k_i$. A general procedure to obtain the c_i, k_i and q for a fixed polynomial phase order, P, is outlined in [3]. For example, for $P = 2$, we find $q = 2$ and $c_1 = 0.5$. For $P = 3$ or $P = 4$ we find $q = 6$, $c_1 = 0.62$, $c_2 = 0.75$ and $c_3 = -0.87$ [3].

By applying the Fourier transform to the "kernel", K_z^P, of the signal we can determine the IF estimate. That is, the peak of the function,

$$W_z^P(t,f) = \int_{-\infty}^{\infty} K_z^P(t,\tau) \cdot e^{-j2\pi f\tau} d\tau, \tag{10.4.5}$$

yields the IF of the signal, $z(t)$. The function $W_z^P(t,f)$ is known as the polynomial Wigner-Ville distribution (PWVD) [2]. This terminology has arisen because $W_z^P(t,f)$ may be considered to be a generalization of the Wigner-Ville distribution (WVD), which has long been known to be effective for IF estimation of quadratic phase signals.

Figure 10.4.1 displays the PWVD of a quadratic ($P = 3$) FM signal. The plot on the left is the PWVD of a noiseless signal, while the plot on the right is the PWVD of the same signal immersed in Gaussian noise. In the latter case, the IF law of the original signal is still apparent, despite the degradation in the time-frequency distribution due to the noise. That is, the PWVD is useful for estimating the IF of *noisy* signals.

The PWVD is very good for analyzing polynomial phase signals at high signal-to-noise ratio (SNR) [4], but is very poor for low SNR signals. In order to achieve accurate IF estimation at low SNR values, one can use an iterative procedure [5].

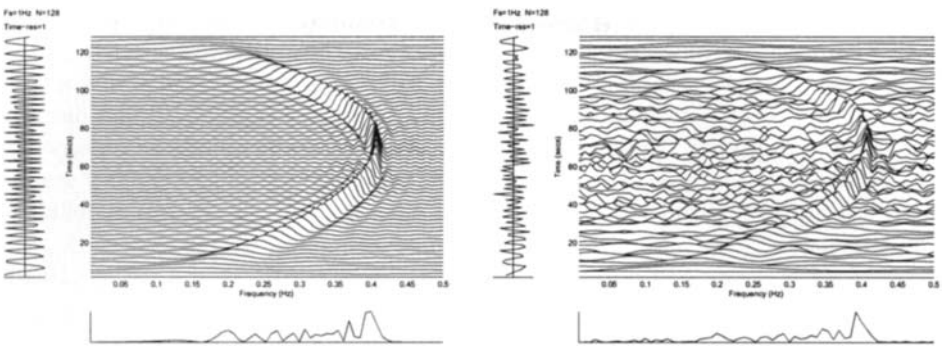


Fig. 10.4.1: PWVDs of a noiseless (left) and noisy (right) quadratic FM signal.

The first step involves forming a preliminary estimate of the IF law, $\hat{f}_i(t)$, by using a method which has a low SNR threshold but possibly limited accuracy (e.g. the peak of the short-time Fourier transform). $\hat{f}_i(t)$ is then used to reconstruct an estimate, $\hat{y}(t)$, of the noisy observation, $y(t)$. $\hat{y}(t)$ is used together with $y(t)$ to form a Cross Polynomial Wigner-Ville Distribution (XPWVD):

$$XW_y^P(t,f) = \int_{-\infty}^{\infty} XK_y^P(t,\tau) \cdot e^{-j2\pi f\tau} d\tau, \tag{10.4.6}$$

where

$$XK_y^P(t,\tau) = y(t+c_1\tau)^{k_1} \hat{y}^*(t-c_1\tau)^{k_1} \prod_{i=2}^{I} \hat{y}(t+c_i\tau)^{k_i} \cdot \hat{y}^*(t-c_i\tau)^{k_i} \;\; ; \;\; t = [0,T]. \tag{10.4.7}$$

XPWVD peak extraction is then used to form a new estimate of the IF law, which in turn is used to form an updated estimate of $y(t)$.

The process of (i) XPWVD formation, (ii) IF law estimation and (iii) re-estimation of $y(t)$ is repeated iteratively until there is an acceptably small difference between successive signal estimates. The SNR threshold of this method is typically very similar to that of the preliminary IF estimation method, but the accuracy tends to be much greater. More details about the technique can be found in [5].

10.4.1.2 Discrete-Time Implementation

In practice it is often necessary to operate on discrete-time signals. In this article $z(n)$, $y(n)$ and $K_y^P(n,m)$ signify the discrete-time signals obtained by sampling $z(t)$, $y(t)$ and $K_y^P(t,\tau)$ respectively with a sampling rate of 1. The number of samples available from both $z(n)$ and $y(n)$ is denoted $N(\approx T)$. The discrete-time version of the PWVD is:

$$W_y^P(n,f) = \underset{m \to f}{\text{DTFT}} \; K_y^P(n,m) \tag{10.4.8}$$

where the DTFT operation specifies discrete-time Fourier transformation from the m to f variables. It should be noted that to form the kernel term, $K_y^P(n,m)$, in the above expression, one must obtain samples of $y(t)$ at non-integer values of time. This typically means that in practice one must use interpolation of the discrete-time observation, $y(n)$. To limit the number of samples which must be obtained by interpolation it is convenient to introduce some 'lag scaling' and 'frequency scaling' operations into the definition for the discrete-time PWVD. The resulting definition then becomes:

$$W_y^P(n,f) = \underset{m \to \frac{f}{c_{i_{\max}}}}{\text{DTFT}} \; K_{y_s}^P(n,m) \tag{10.4.9}$$

where

$$K_{y_s}^P(t,\tau) = \prod_{i=1}^{I} y\left(t + \frac{c_i}{c_{i_{\max}}}\tau\right)^{k_i} \cdot y^*\left(t - \frac{c_i}{c_{i_{\max}}}\tau\right)^{k_i}, \tag{10.4.10}$$

and $K_{y_s}^P(n,m)$ is obtained from $K_{y_s}^P(t,\tau)$ by sampling at a rate of 1. $c_{i_{\max}}$ is the c_i value with the largest magnitude. The discrete-time PWVD defined above can be used for obtaining low variance IF estimates from discrete-time measurements. A derivation of the variance of such estimates is given in the Appendix.

10.4.1.3 Polynomial Phase Parameter Estimation

The PWVD can also be used to estimate the a_k parameters themselves, rather than simply the IF. It will be assumed for this estimation task that discrete-time measurements are available and that the noise on these measurements is white and Gaussian. Basically, the estimation procedure involves (a) forming P PWVD slices, (b) determining P IF estimates from these slices, (c) fitting a polynomial IF law to these P IF estimates, (d) deducing the a_k parameters from this polynomial IF law, and (e) performing a spectral zoom type technique to refine the parameter estimates. The procedure is specified mathematically below:

1. For the observation, $y(n)$, form the set of PWVD slices,

$$\mathbf{W_y^P}(f) = [W_y^P(n_1,f)\; W_y^P(n_2,f)\; \ldots W_y^P(n_P,f)]^T, \tag{10.4.11}$$

   where $n_1, n_2, \ldots, n_P$ are different, well-spaced discrete-time positions. There is some flexibility in the selection of $\{n_1,\ n_2, \ldots, n_P\}$. Having them equally spaced between $0.25N$ and $0.75N$ is typically a good choice. i.e. $n_i = 0.25N + 0.5N(i-1)/(P-1),\ \ i = 1, \ldots, P.$

2. Estimate a vector of IF estimates according to:

$$\hat{\mathbf{f}}_\mathbf{i} = \mathbf{arg}\ \ \underset{f}{\mathbf{max}}\ \{\mathbf{W_y^P}(f)\}, \tag{10.4.12}$$

   where **arg max** here denotes the vector argument of a vector (row by row) maximization.

3. Determine the vector of initial coarse estimates for the polynomial phase parameters estimates. This vector is denoted $\hat{\mathbf{a}}_\mathbf{i} = [\hat{a}_{i1} \;\; \hat{a}_{i2} \;\; \hat{a}_{i3} \;\; \ldots \;\; \hat{a}_{iP}]^T$, and is obtained by evaluating the matrix equation:

$$\hat{\mathbf{a}}_\mathbf{i} = 2\pi \mathbf{X}^{-1} \hat{\mathbf{f}}_\mathbf{i} \tag{10.4.13}$$

where

$$\mathbf{X} = \begin{bmatrix} 1 & 2n_1 & 3n_1^2 & \ldots & Pn_1^{P-1} \\ 1 & 2n_2 & 3n_2^2 & \ldots & Pn_2^{P-1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 2n_P & 3n_P^2 & \ldots & Pn_P^{P-1} \end{bmatrix} \tag{10.4.14}$$

4. Dechirp the observation by the estimated polynomial phase law:

$$y_d(n) = y(n) \exp(-j(\hat{a}_{i1} n + \hat{a}_{i2} n^2 + \ldots + \hat{a}_{ip} n^p)). \tag{10.4.15}$$

(The dechirped observation will now have spectral content which is highly localized around DC).

5. Filter $y_d(n)$ with an ideal low-pass filter whose bandwidth is B. Decimate the result by a factor, $1/B$, to obtain $y_o(n)$.

6. Unwrap the phase of $y_o(n)$. The vector of unwrapped phase values, $\mathbf{V} = [V(0) \; V(1) \; \ldots \; V(NB-1)]^T$, is also modeled as a (noisy) polynomial phase signal with phase parameters, $\mathbf{a_d}=[a_{d0} \; a_{d1} \; \ldots \; a_{dP}]^T$. One can then estimate $\mathbf{a_d}$ with the following linear regression:

$$\hat{\mathbf{a}_\mathbf{d}} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\mathbf{V}, \tag{10.4.16}$$

where

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 1 & 1 & \ldots & 1^P \\ 1 & 2 & \ldots & 2^P \\ \ldots & \ldots & \ldots & \ldots \\ 1 & NB-1 & \ldots & (NB-1)^P \end{bmatrix}, \tag{10.4.17}$$

7. Use the results from Step 6 to refine the parameter estimates from Step 3:

$$\hat{\mathbf{a}_\mathbf{f}} = \hat{\mathbf{a}}_\mathbf{i} + \hat{\mathbf{a}_\mathbf{d}}.\mathbf{B} \tag{10.4.18}$$

where $\mathbf{B} = [1 \; B \; B^2 \ldots B^P]^T$, and $\hat{\mathbf{a}}_\mathbf{d}.\mathbf{B}$ denotes element by element multiplication of $\hat{\mathbf{a}}_\mathbf{d}$ and $\mathbf{B}$.

Note that at least $P+1$ samples of $y_o(n)$ are required for the regression in Step 6 to be properly defined. In practice, it is recommended that a large number of samples be obtained. Note also that if $y(n)$ is white then $y_o(n)$ will also be white,

and being a filtered version of $y(n)$, will have a higher SNR (by a factor of $1/B$). The unwrapped phase of $y_o(n)$ will also be very close to white, provided that the SNR of $y_o(n)$ is above about 9dB [6]. The overall algorithm yields estimates which are asymptotically optimal above threshold. The threshold can be predicted with the formulae in [7].

10.4.2 Higher Order Ambiguity Functions

An alternative approach for analyzing polynomial phase signals has been provided by Peleg, Porat and Friedlander [8], [9]. The first step in the method is to apply a transform to convert an arbitrary P^{th} order polynomial phase signal into a linear phase one whose frequency is proportional to the a_P parameter. This can be done with "Higher order ambiguity functions (HAFs)". After the transformation in the first step, it is possible to estimate the a_P parameter with the Fourier transform. The second step in the method is to reduce the original signal to a $(P-1)^{th}$ order polynomial phase signal by "de-chirping" the highest order phase term out of the signal. A similar process can then be used to estimate the a_{P-1} parameter, and subsequently all lower order parameters. The formal algorithmic specification for the phase parameter estimators in [8] is:

1. Let $p = P$ and let $y^{(p)}(n) = y(n),\ 1 \le n \le N$.
2. Set $\tau_p = \frac{N}{p}$. Calculate the a_p estimate using the formula:

$$\begin{aligned}\hat{a}_p &= \frac{1}{p!\,\tau_p^{p-1}} \arg\max_{\omega} \left\{ \left| \sum_{n=(M-1)\tau_p+1}^{N} \Pi_{q=0}^{p-1} \left[y^{\$q}(n - q\tau_p)^{\binom{p-1}{q}} \right] e^{-j\omega n} \right| \right\} \\ &= \frac{1}{p!\,\tau_p^{p-1}} \arg\max_{\omega} \left\{ |\mathrm{HAF}_y(\omega, \tau_p)| \right\}, \end{aligned} \tag{10.4.19}$$

   where $y^{\$q}(n) = \begin{cases} y(n) \text{ if } q \text{ is even} \\ y^*(n) \text{ if } q \text{ is odd} \end{cases}$, * signifies complex conjugation, and $\mathrm{HAF}_y(\omega, \tau_p)$ denotes the p^{th} order "higher order ambiguity function".
3. Set $y^{(p-1)}(n) = y^p(n)\exp(-j\hat{a}_p n^p)$. Then set $p = p - 1$.
4. If $p \ge 1$, go back to Step2. Else proceed.
5. $\hat{a}_0 = \arg\{y^0(n)\},\ \hat{A} = |y^0(n)|$.

The parameter estimates obtained with the above algorithm are good but not optimal [9]. The estimates can be refined (to the point of optimality) with either a Newton algorithm or with a procedure similar to the one outlined in Steps 4-7 of Section 10.4.1.3.

Once the $a_0, a_1, \ldots, a_P$ parameters have been estimated, (10.4.1) can be used to obtain a reconstruction of the entire signal phase (and using (10.4.2), the IF).

Like the PWVD approach this method performs very well at high SNR, but suffers from threshold effects at low SNR. As with the PWVDs, an iterative method can be used to improve performance at low SNR, the details being provided in [10].

10.4.3 A Comparison of the Polynomial Wigner-Ville Distributions and the Higher Order Ambiguity Functions

10.4.3.1 Single Component Analysis

For analyzing a single polynomial phase signal the computational burdens of the PWVD and HAF methods are similar. The asymptotic variances of the a_k parameter estimates using the PWVD method are also similar. In fact, if the estimate refinement scheme outlined in Steps 4-7 of Section 10.4.1.3 is used for both methods, then the two techniques are both optimal above threshold. The SNR thresholds for the HAF and PWVD techniques are similar for lower order phase laws (i.e. third order and lower), while for higher order laws the PWVDs tend to have lower thresholds. Formulae for the threshold values are given in [8] and [7].

10.4.3.2 Multiple Component Analysis

Both the PWVDs and the HAFs are non-linear transforms and hence produce "cross-terms" when operating on a sum of two or more polynomial phase signals. These cross-terms complicate the task of analyzing such "multi-component" signals. The PWVDs produce cross-terms which are often spectrally concentrated, while the HAFs give rise to cross-terms which are spectrally dispersed. The spectrally dispersed cross-terms prove to be more conducive to analysis, because they appear more like white noise.

In analyzing multiple components with the HAFs one can proceed by first estimating the parameters of the highest amplitude component, and then removing the estimate of this component from the signal [11]. Then the next highest amplitude component can be analyzed. The process can be repeated until all components have been accounted for. Pre-requisites for the effective analysis of multi-component signals include (1) higher required SNR thresholds than for single component analysis, (2) good separation of the various components in the HAFs, and (3) phase polynomial orders of the constituent signals which are not too high. Requirement (2) is necessary because the level of dispersion of the cross-terms is related to the extent of the separation of the components in the HAFs. Requirement (3) is necessary because the energy of the cross-terms relative to the auto-terms increases with increasing HAF order.

10.4.4 Appendix: Asymptotic Mean-Square Error of a PWVD-Based IF Estimate

It is assumed that the IF of a polynomial phase signal is to be estimated, using the peak of the discrete-time PWVD. It is further assumed that the signal is embedded in complex, white, Gaussian noise with power, σ^2, and that the estimate of the IF

is required at time, $n = 0$. This appendix derives the asymptotic mean-square error (a.m.s.e.) of the IF estimate.

The PWVD based IF estimate, evaluated at $n = 0$, is defined by:

$$\hat{f}_0 = \arg\max_f \{W_y^P(n, f)|_{n=0}\} = \arg\max_f \{W_y^P(f)\}, \tag{10.4.20}$$

where

$$W_y^P(f) = \underset{m \to \frac{f}{c_{i_{\max}}}}{\text{DTFT}} \; K_{y_s}^P(n = 0, m). \tag{10.4.21}$$

The PWVD of a noisy polynomial phase signal, evaluated at $n = 0$, is:

$$W_y^P(f) = \sum_{m=-(N-1)/2}^{(N-1)/2} \prod_{i=1}^{I} \left[y(mc_i/c_{i_{\max}}) \cdot y^*\left(-mc_i/c_{i_{\max}}\right)\right]^{k_i} e^{-j2\pi f m/c_{i_{\max}}}. \tag{10.4.22}$$

In the absence of noise $W_y^P(f)$ has a global maxima at $f = f_0$, where f_0 is the value of the IF at $n = 0$. The perturbation of $W_y^P(f)$ which occurs when noise is added to the observation is given by:

$$\delta W_y^P(f) = \sum_{m=-(N-1)/2}^{(N-1)/2} zw(m) e^{-j2\pi f m/c_{i_{\max}}}, \tag{10.4.23}$$

where

$$zw(m) = \prod_{i=1}^{I} \left[y(mc_i/c_{i_{\max}})\, y^*\left(-mc_i/c_{i_{\max}}\right)\right]^{k_i} - \left[z(mc_i/c_{i_{\max}})\, z^*\left(-mc_i/c_{i_{\max}}\right)\right]^{k_i}. \tag{10.4.24}$$

With this perturbation, the peak of $W_y^P(f)$ shifts from $f = f_0$ to $f = f_0 + \delta f$. This appendix derives the a.m.s.e. of δf (i.e. the a.m.s.e. of the IF estimate). To achieve this, a formula for the asymptotic mean-square fluctuations of the maximum of a real valued random function is used [12]. This formula, re-expressed for the scenario in this section, is:

$$E\left\{\delta f^2\right\} = E\left\{\left[\frac{\partial \delta W_y^P(f_0)}{\partial f}\right]^2\right\}\left[\frac{\partial^2 W_y^P(f_0)}{\partial f^2}\right]^{-2} \tag{10.4.25}$$

where $E\{.\}$ denotes the expected value.

Under the simplifying assumption that all samples of the kernel are uncorrelated, relevant terms on the right hand side of (10.4.25) are found to be:

$$\frac{\partial^2 W_y^P(f_0)}{\partial f^2} \approx \frac{-4\pi^2 A^q N^3}{12 c_{i_{\max}}^2}, \tag{10.4.26}$$

$$\frac{\partial \delta W_y^P(f_0)}{\partial f} \approx -j2\pi \sum_{m=-N/2}^{N/2} m \cdot zw(m) e^{-j2\pi f m / c_{i_{\max}}}, \tag{10.4.27}$$

$$E\left\{\left[\frac{\partial \delta W_y^P(f_0)}{\partial f}\right]^2\right\} \approx \frac{4\pi^2 A^{2q} N^3 V_{\text{kern}}}{12}, \tag{10.4.28}$$

where V_{kern}, the "noise to signal ratio" in the kernel, is [4], [13]:

$$V_{\text{kern}} = \left[\left(\sum_{i=0}^{k_1}\binom{k_1}{i}^2 i! \frac{\sigma^{2i}}{A^{2i}}\right)^2 \text{X} \cdots \text{X} \left(\sum_{i=0}^{k_I}\binom{k_I}{i}^2 i! \frac{\sigma^{2i}}{A^{2i}}\right)^2\right] - 1. \tag{10.4.29}$$

(10.4.26) and (10.4.28) can be substituted into (10.4.25) to obtain:

$$E\left\{\delta f^2\right\} \approx \frac{12 \cdot c_{i_{\max}}^2 \cdot V_{\text{kern}}}{4\pi^2 \cdot N^3}. \tag{10.4.30}$$

10.4.5 Summary and Conclusions

In this article the problem of analyzing polynomial FM signals in additive noise has been addressed. Two different approaches have been described, the polynomial Wigner-Ville distributions and the higher order ambiguity functions. Both methods perform well at high SNR, and with some modifications, can be made to work well even at low SNRs. A brief comparison of both methods has been made.

References

[1] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Information Theory*, vol. 20, pp. 591–598, September 1974.

[2] B. Boashash and P. J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 216–220, January 1994.

[3] B. Barkat and B. Boashash, "Design of higher order polynomial Wigner-Ville distributions," *IEEE Trans. Signal Processing*, vol. 47, pp. 2608–2611, September 1999.

[4] 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 Processing*, vol. 47, pp. 2480–2490, September 1999.

[5] B. Ristic and B. Boashash, "Instantaneous frequency estimation of quadratic and cubic polynomial FM signals using the cross polynomial Wigner-Ville distribution," *IEEE Trans. Signal Processing*, vol. 44, pp. 1549–1553, 1996.

[6] S. A. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Information Theory*, vol. 31, pp. 832–835, November 1985.

[7] D. C. Reid, A. M. Zoubir, and B. Boashash, "Aircraft flight parameter estimation based on passive acoustic techniques using the polynomial Wigner-Ville distribution," *J. Acoustical Soc. of America*, vol. 102, pp. 207–23, July 1997.

[8] S. Peleg and B. Porat, "Estimation and classification of polynomial-phase signals," *IEEE Trans. Information Theory*, vol. 37, pp. 422–430, March 1991.

[9] B. Porat and B. Friedlander, "Asymptotic statistical analysis of the high-order ambiguity function for parameter estimation of polynomial phase signal," *IEEE Trans. Information Theory*, vol. 42, pp. 995–1001, May 1996.

[10] P. O'Shea, "An iterative algorithm for estimating the parameters of polynomial phase signals," in *Proc. Fourth Internat. Symp. on Signal Processing and its Applications (ISSPA '96)*, vol. 2, pp. 730–731, Gold Coast, Australia, 25–30 August 1996.

[11] S. Peleg and B. Friedlander, "Multicomponent signal analysis using the polynomial-phase transform," *IEEE Trans. Aerospace & Electronic Systems*, vol. 32, pp. 378–386, January 1996.

[12] S. Peleg and B. Porat, "Linear FM signal parameter estimation from discrete-time observations," *IEEE Trans. Aerospace & Electronic Systems*, vol. 27, pp. 607–616, July 1991.

[13] G. Reina and B. Porat, "Comparative performance analysis of two algorithms for instantaneous frequency estimation," in *Proc. Eighth IEEE Workshop on Statistical Signal and Array Processing (SSAP-96)*, pp. 448–451, Corfu, Greece, 24–26 June 1996.

10.5 IF ESTIMATION OF FM SIGNALS IN MULTIPLICATIVE NOISE [0]

10.5.1 Random Amplitude Modulation

Most IF estimation techniques, such as those presented in the previous articles of this chapter, assume that the signal of interest has a constant amplitude. While this is a valid assumption in a wide range of scenarios, there are several important applications in which this assumption does not hold. Indeed, in many situations the signal may be subjected to a *random* amplitude modulation which behaves as multiplicative noise. Examples include fading in wireless communications [1], fluctuating targets in radar [2], and structural vibration of a spacecraft during launch and atmospheric turbulence [3]. In this article, we focus on non-parametric methods. In particular, we show that the Wigner-Ville distribution (defined in Section 2.1.4) is able to display the IF of a signal affected by multiplicative noise, and that this representation is optimal in the sense of maximum energy concentration for a linear FM signal. For higher-order polynomial FM signals, the use of the polynomial Wigner-Ville distribution (PWVD), presented in Article 5.4, is shown to give optimal representations. Statistical performance of each case will be presented here.

10.5.2 Linear FM Signal

In this section, we study the case of a linear FM signal and assume that the multiplicative noise is a *real-valued* process.

10.5.2.1 Optimality of the Wigner-Ville Spectrum

First we show that the Wigner-Ville spectrum (WVS) is optimal, in the sense of IF localization, for the time-frequency analysis of linear FM signals affected by multiplicative noise. Consider the signal $y(t)$ given by

$$y(t) = a(t) \cdot z(t) \tag{10.5.1}$$

where $a(t)$ is a non-zero-mean real-valued stationary noise and $z(t)$ is a deterministic FM signal given by $z(t) = \exp\{j\phi(t)\}$. For a linear FM signal, $\phi(t)$ is a second-order polynomial. Using the expectation operator notation, the autocorrelation of the signal above can be expressed as

$$\begin{aligned} K_y(t,\tau) &= \mathrm{E}\left[y(t-\tfrac{\tau}{2})\, y^*(t+\tfrac{\tau}{2})\right] \\ &= \mathrm{E}\left[a(t-\tfrac{\tau}{2})\, a(t+\tfrac{\tau}{2})\right] \cdot \left\{z(t-\tfrac{\tau}{2})\, z^*(t+\tfrac{\tau}{2})\right\} \\ &= \mathcal{R}_a(\tau) K_z(t,\tau). \end{aligned} \tag{10.5.2}$$

[0] Authors: **B. Barkat**, Nanyang Technological University, School of Electrical & Electronic Engineering, Block S2, Nanyang Avenue, Singapore 639798 (ebarkat@ntu.edu.sg), and **B. Boashash**, Queensland University of Technology, Signal Processing Research Centre, GPO Box 2434, Brisbane, Q 4001, Australia (b.boashash@qut.edu.au). Reviewers: Dr. F. Sattar and S. Gulam Razul.

The WVS of $y(t)$, which is defined as the Fourier transform of $\mathcal{K}_y(t,\tau)$ [4], can be expressed as

$$\begin{aligned}\mathcal{W}_y(t,f) &= \mathcal{F}_{\tau\to f}\left\{\mathcal{K}_y(t,\tau)\right\} && (10.5.3)\\ &= \mathcal{F}_{\tau\to f}\left\{z(t-\tfrac{\tau}{2})\,z^*(t+\tfrac{\tau}{2})\right\} *_f \mathcal{F}_{\tau\to f}\left\{\mathrm{E}\left[a(t-\tfrac{\tau}{2})\,a(t+\tfrac{\tau}{2})\right]\right\} && (10.5.4)\\ &= W_z(t,f) *_f S_a(f) && (10.5.5)\end{aligned}$$

where $*_f$ is the convolution operation in the frequency space.

If we express the non-zero-mean random process $a(t)$ as $a(t) = \mu_a + a_0(t)$, where μ_a is a constant mean of $a(t)$ and $a_0(t)$ is a zero-mean noise with autocorrelation $R_{a_0}(\tau)$, we can re-write the WVS of $y(t)$ as

$$\mathcal{W}_y(t,f) = \mu_a^2 W_z(t,f) + S_{a_0}(f) *_f W_z(t,f). \tag{10.5.6}$$

For the case of a linear FM signal, the Wigner-Ville distribution (WVD) is given in Article 2.1 by [4]

$$W_z(t,f) = \delta(f - f_i(t)), \tag{10.5.7}$$

where $f_i(t)$ is the signal IF and δ is the Dirac delta function. In this case, we obtain

$$\mathcal{W}_y(t,f) = \mu_a^2\delta(f - f_i(t)) + S_{a_0}(f - f_i(t)). \tag{10.5.8}$$

Note that Eq. (10.5.8) exhibits the presence of a spectral line at the frequency $f_i(t)$ for all time instants. This means that, theoretically, the WVS always localizes the IF of a linear FM signal. This makes it a powerful tool in the analysis of linear FM signals affected by multiplicative noise. Also note that when $\mu_a = 0$, the WVS will not exhibit a peak at the signal IF, indicating a breakdown of the WVS to analyze the noisy signal.

As an illustration, consider a unit-modulus linear FM signal sampled at 1 Hz, whose frequency range lies between 0.1 Hz and 0.4 Hz. The signal length (in samples) is chosen as $N = 511$. This signal is multiplied by a real-valued i.i.d. Gaussian noise with a standard deviation $\sigma_a = 1$ and a mean equal to 0 and 1, respectively. Fig. 10.5.1 displays the WVS (one realization) of the noisy signal for both values of the mean. As expected, the WVS for the zero-mean case cannot reveal the signal IF; however, it can do so for the other case. In this last case, the peak of the WVS can be used to estimate the IF of the signal. In what follows, we will evaluate the statistical performance of such an estimator.

10.5.2.2 Statistical Performance Evaluation

Here, for a more complete study, we consider the presence of additive noise as well as the multiplicative noise. The objective is to derive the asymptotic variance of the IF estimator, based on the peak of the WVS, for this case.

Let the discrete-time version of the noisy signal be

$$y(n) = a(n)\,e^{j\phi(n)} + w(n)\ ,\quad n = 0,\ldots,N-1. \tag{10.5.9}$$

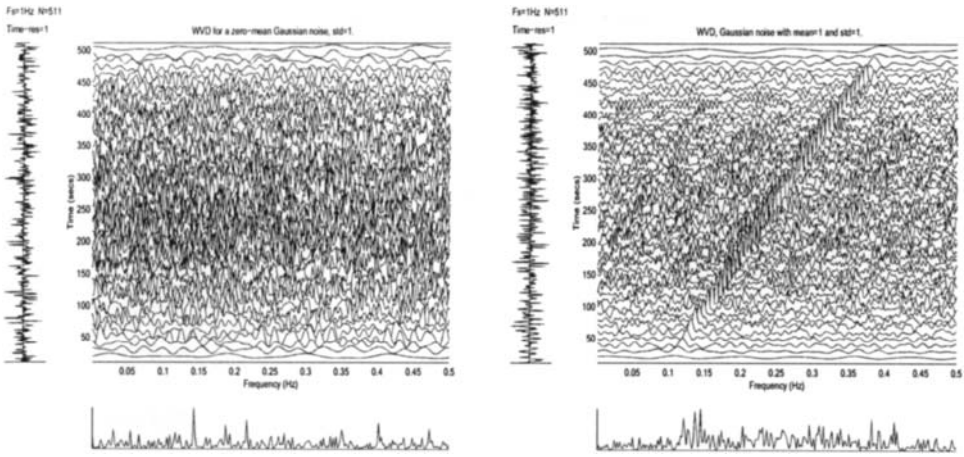


Fig. 10.5.1: The Wigner-Ville spectrum of a linear FM signal affected by a real-valued Gaussian multiplicative noise. The noise variance is 1 (both plots) and its mean is 0 (left plot) and 1 (right plot).

The process $a(n)$ is considered to be a real-valued stationary Gaussian noise with mean and variance given by μ_a and σ_a^2, respectively. The complex zero-mean additive process $w(n)$ is assumed to be stationary, white, circular and Gaussian with variance equal to σ_w^2. In addition, both noises are assumed to be independent.

The WVS used to estimate the signal IF is defined, in the discrete-time domain, using the expression given in Article 6.1, as [4]

$$\mathcal{W}_z(n,f) = \mathrm{E}\left[2\sum_{m=-M}^{M} y(n+m)\cdot y^*(n-m)\, e^{-j4\pi fm}\right]. \tag{10.5.10}$$

Using straightforward derivations, we can show that, for increasing window length $(2M+1)$, the WVS converges in probability to $\mu_a^2\delta(f-f_i(t))$ [5]. We can also show that the IF estimator asymptotic variance is approximately equal to [5]

$$\mathbf{Var}(\hat{f}_i(n)) = \frac{3}{(2\pi)^2\mathcal{S}_w(2M+1)^3}\left(2+\tfrac{2}{\mathcal{S}_a}+\tfrac{1}{\mathcal{S}_w}\right) \tag{10.5.11}$$

where $\mathcal{S}_a = \mu_a^2/\sigma_a^2$ and $\mathcal{S}_w = \mu_a^2/\sigma_w^2$. Note that:

(i) When $\mu_a = 0$, the variance goes to infinity indicating that the WVS based estimator breaks down. This result confirms the analysis presented earlier.

(ii) When $\mu_a = A$ where A is a constant and $\sigma_a = 0$, i.e., the signal under consideration is just a constant amplitude linear FM signal embedded in noise, the variance expression can be-rewritten as

$$\mathbf{Var}(\hat{f}_i(n)) = \frac{3\sigma_w^2[2A^2+\sigma_w^2]}{(2\pi)^2A^4(2M+1)^3}.$$

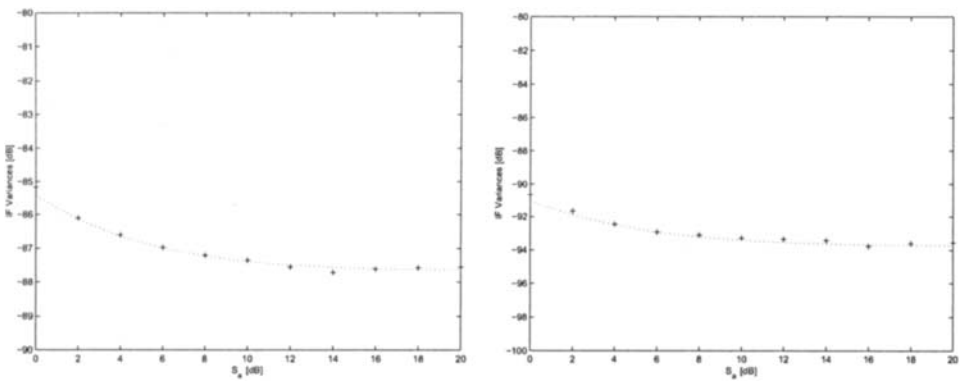


Fig. 10.5.2: Theoretical (dashed curve) and estimated ('+') variances of a linear FM signal corrupted by multiplicative and additive noise. Left plot is for $\mathcal{S}_w = 0\,\text{dB}$ and right plot is for $\mathcal{S}_w = 5\,\text{dB}$.

If, in addition, we assume high signal-to-noise ratio (SNR) (i.e., $A^2 \gg \sigma_w^2$) the asymptotic variance expression reduces to

$$\mathbf{Var}(\hat{f}_i(n)) = \frac{6\sigma_w^2}{(2\pi)^2(A^2)(2M+1)^3}$$

which is similar to the result obtained in [6].

The above theoretical results were confirmed by Monte-Carlo simulations. Specifically, we estimate the IF estimator variance using 5000 realizations of the signal given by (10.5.9). In Fig. 10.5.2, we display the theoretical (dashed curve) and the estimated ('+') variances plotted against $\mathcal{S}_a$ for $\mathcal{S}_w = 0$ and 5 dB, respectively.

If the signal under consideration is not a linear FM but a higher-order polynomial FM signal, the WVS becomes inappropriate because it introduces some artifacts which might hide the real features of the signal and its peak based IF estimator is biased for such signals [6]. In this situation, a different tool is needed for the analysis. This is the topic of the next section.

10.5.3 Polynomial FM Signals

Compared to the previous section, the extension here is two-fold: (i) the signal considered is assumed to be a polynomial FM signal of arbitrary order and (ii) the multiplicative noise is no longer limited to a real-valued process but is assumed to be a non-zero complex circular Gaussian process. Based on this, the noisy signal $y(t)$ is now written as

$$y(t) = a(t) \cdot z(t) + w(t) \tag{10.5.12}$$

where the stationary processes $a(t)$ and $w(t)$ are both assumed circular, complex, Gaussian and independent with means and variances given by (μ_a, σ_a^2) and $(0, \sigma_w^2)$,

respectively. The noiseless polynomial FM signal $z(t)$ is given by

$$z(t) = e^{j\phi(t)} = \exp\left\{ j \sum_{i=0}^{P} a_i t^i \right\} \tag{10.5.13}$$

where the a_i are real coefficients and P is the order of the polynomial phase. Note that in the derivation below we do not require the knowledge of the coefficients a_i; we only assume a polynomial FM signal.

The IF of the signal $z(t)$ is given in Article 1.3 by

$$f_i(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt} = \sum_{i=1}^{P} i a_i t^{i-1}$$

and our primary objective here is to estimate $f_i(t)$ from the noisy signal $y(t)$. For that purpose, we use the polynomial Wigner-Ville distribution (PWVD) defined in Article 5.4 as

$$W_z^{(q)}(t,f) = \int_{-\infty}^{\infty} \prod_{i=1}^{q/2} z(t + c_i\tau) z^*(t + c_{-i}\tau) e^{-j2\pi f\tau} d\tau \tag{10.5.14}$$

$$= \int_{-\infty}^{\infty} K_z^{(q)}(t,\tau) \cdot e^{-j2\pi f\tau} d\tau \tag{10.5.15}$$

where q is an even integer which indicates the order of non-linearity of the PWVD. The coefficients c_i and c_{-i} $(i = 1, 2, \ldots, q/2)$ are calculated so that the PWVD is real and equal to

$$W_z^{(q)}(n,f) = \delta(f - f_i(t)),$$

for signals given by Eq. (10.5.13). Note that the realness of the PWVD implies that $c_i = -c_{-i}$. Also note that the WVD is a member of the PWVDs class with parameters $q = 2$ and $c_1 = -c_{-1} = 0.5$. Full details of the design procedure may be found in [7] and Articles 5.4 and 5.5.

The choice of the PWVD stems from the fact that it yields a continuum of delta functions around the IF for a given polynomial FM signal. This property implies that the peak of the PWVD can be used as an IF estimator for polynomial FM in a noisy environment. In [8], the statistical performance of this estimator was evaluated for noisy signals described by Eq. (10.5.12). It shows that this estimator is unbiased and its asymptotic variance is approximately equal to [8]

$$\mathbf{Var}(\hat{f}_i(n)) = \frac{6\,(\sigma_a^2 + \sigma_w^2) \sum_{i=1}^{n_1} k_i^2}{(2\pi)^2 |\mu_a|^2 (2M+1)^3}. \tag{10.5.16}$$

In the above expression, n_1 represents the number of the different coefficients c_i in the PWVD kernel, while k_i (for $i = 1, \ldots, n_1$) represents the multiplicity of each of these coefficients c_i, and $(2M + 1)$ is the window length considered in the PWVD discrete-time implementation. Note that:

(i) When $\mu_a = 0$, the variance goes to infinity indicating that the PWVD based estimator breaks down.

(ii) When $\mu_a = A$ where A is a constant and $\sigma_a = 0$, i.e., the signal under consideration is just a constant amplitude polynomial FM signal embedded in complex Gaussian noise, the variance expression can be-rewritten as

$$\mathbf{Var}(\hat{f}_i) = \frac{6\,\sigma_w^2 \sum_{i=1}^{n_1} k_i^2}{(2\pi)^2 A^2 (2M+1)^3}.$$

The above expression is exactly similar to the result obtained in [6], which treats *constant* amplitude polynomial FM signals only.

10.5.3.1 Monte-Carlo Simulations

To confirm the validity of the above theoretical results, we consider the IF estimation of a quadratic FM signal at the middle of the signal interval. The peak of the sixth-order PWVD, whose signal kernel is given by [7]

$$\begin{aligned} K_y^{(6)}(t,\tau) &= [y(t+0.62\tau)\, y^*(t-0.62\tau)] \cdot [y(t+0.75\tau)\, y^*(t-0.75\tau)] \\ &\times [y(t-0.87\tau)\, y^*(t+0.87\tau)], \end{aligned}$$

is used here as the IF estimator. The noisy signal $y(t)$ is generated as suggested by Eq. (10.5.12). For this example, we choose the sampling period equal to $T = 1$, the signal length N equal to 129, the window length equal to $(2M + 1 = N = 129)$ and the noise variances to be equal (i.e., $\sigma_a^2 = \sigma_w^2$). In the simulations, the overall signal-to-noise ratio, defined as

$$\mathrm{SNR}_{w_1} = 10 \log_{10}(|\mu_a|^2/(\sigma_a^2 + \sigma_w^2)),$$

is varied in a 1 dB step from 0 to 15 dB. Monte-Carlo simulations for 1000 realizations are run for each value of SNR_{w_1}. The results of two different experiments, one performed for $|\mu_a| = 0.01$ and the other performed for $|\mu_a| = 1$, are displayed in Fig. 10.5.3 (left plot). We observe that, above a certain threshold, the estimated variances represented by '+' (for $|\mu_a| = 0.01$) and 'o' (for $|\mu_a| = 1$) are in total agreement with the theoretical ones given by Eq. (10.5.16) and represented by the continuous lines (superimposed).

Simulations run under the same noise conditions for other polynomial FM signals using the appropriate PWVD order also confirm the theoretical results presented above. One such case is when $y(t)$ is a linear FM signal and the PWVD considered is the second-order PWVD. The results of this experiment are displayed in the right plot of Fig. 10.5.3.

10.5.4 Time-Varying Higher-Order Spectra

Time-Varying Higher-Order Spectra (TV-HOS) based on the polynomial Wigner-Ville distribution are defined as the expected value of the PWVD [9], namely

$$W_z^{(q)}(t,f) = \mathrm{E}\left[W_z^{(q)}(t,f)\right] \qquad (10.5.17)$$

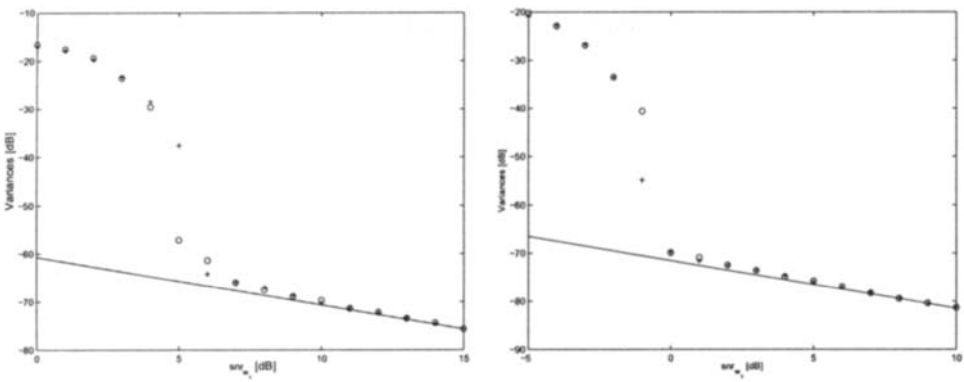


Fig. 10.5.3: Results of the experiments for a **quadratic** (left plot) and a **linear** (right plot) FM signal corrupted by complex Gaussian multiplicative and additive noise processes. The continuous lines (superimposed) represent the theoretical variances while '+' and 'o' correspond to the estimated variances for $|\mu_a| = 0.01$ and $|\mu_a| = 1$, respectively.

where $W_z^{(q)}(t,f)$ is the q^{th}-order PWVD defined by (10.5.14). Interchanging the expectation operator with the integration in the PWVD, one obtains

$$\mathcal{W}_z^{(q)}(t,f) = \int_{-\infty}^{\infty} \mathrm{E}\left[\prod_{i=1}^{q/2} z(t+c_i\tau)z^*(t+c_{-i}\tau)\right] e^{-j2\pi f\tau}d\tau \quad (10.5.18)$$

$$= \int_{-\infty}^{\infty} \mathcal{K}_z^{(q)}(t,\tau)\cdot e^{-j2\pi f\tau}d\tau \quad (10.5.19)$$

where $\mathcal{K}_z^{(q)}(t,\tau)$ represents a slice of a time-varying q^{th}-order moment function [9]. If the quantity $\mathcal{K}_z^{(q)}(t,\tau)$ is absolutely integrable, then, $\mathcal{W}_z^{(q)}(t,f)$ can be interpreted as a form of the time-varying higher-order moment spectrum. In [9], the authors showed that TV-HOS combine the advantages of classical time-frequency analysis with the benefits of higher-order spectra. To avoid the problem of non-superposition of the higher-order moments, the authors in [9] used higher-order cumulants instead.

It is important to note that since non-stationary random signals are non-ergodic, the ensemble averaging above cannot be replaced by time averaging. In this situation, local ergodicity has to be assumed.

Readers interested in TV-HOS are referred to [9] for more details, including examples of the efficacy of TV-HOS in the analysis of random FM signals affected by multiplicative noise. (See also Section 14.5.4 and the references in [9].)

10.5.5 Summary and Conclusions

The Wigner-Ville spectrum (WVS) and polynomial Wigner-Ville distributions (PWVDs) are considered for the analysis of polynomial FM signals corrupted by multiplicative and additive noise. In the noisy linear FM case, the WVS is shown

to give optimal IF localization. Accordingly, the peak of the WVS is proposed as an IF estimator. A statistical performance test shows that this estimator is very accurate even at low signal-to-noise ratio values. For the case of the noisy higher-order polynomial FM signal, the peak of the PWVD is shown to be a very consistent and accurate IF estimator.

References

[1] R. S. Kennedy, *Fading dispersive communication channels.* New York: Wiley, 1969.

[2] H. L. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. III: "Radar-Sonar Signal Processing and Gaussian Signals in Noise". New York: Wiley, 1971. Reprinted Malabar, FL: Krieger, 1992. Reprinted New York: Wiley, 2001.

[3] A. G. Piersol, "Power spectra measurements for space vibration data," *J. Spacecraft and Rockets*, vol. 4, p. 1613, December 1967.

[4] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[5] M. R. Morelande, B. Barkat, and A. M. Zoubir, "Statistical performance comparison of a parametric and a non-parametric method for IF estimation of random amplitude linear FM signals in additive noise," in *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pp. 262–266, Pocono Manor, PA, 14–16 August 2000.

[6] 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 Processing*, vol. 47, pp. 2480–2490, September 1999.

[7] B. Barkat and B. Boashash, "Design of higher order polynomial Wigner-Ville distributions," *IEEE Trans. Signal Processing*, vol. 47, pp. 2608–2611, September 1999.

[8] B. Barkat, "Instantaneous frequency estimation of nonlinear frequency-modulated signals in the presence of multiplicative and additive noise," *IEEE Trans. Signal Processing*, vol. 49, pp. 2214–2222, October 2001.

[9] B. Boashash and B. Ristic, "Polynomial time-frequency distributions and time-varying higher order spectra: Application to the analysis of multicomponent FM signal and to the treatment of multiplicative noise," *Signal Processing*, vol. 67, pp. 1–23, May 1998.

Chapter 11

Time-Frequency Synthesis and Filtering

To model and predict accurately the effects of linear systems on non-stationary signals in applications such as signal cleansing and enhancement, we need the capability to design time-varying linear systems with precise time-frequency specifications. This topic is covered in four articles with appropriate cross-referencing to other chapters.

The design of time-varying filters is useful in applications where it is desired to separate, suppress or reduce undesirable non-stationary signal components. This can be achieved with a number of methods such as the STFT and Gabor transform (Article 11.1). In particular, the use of the Gabor expansion for time-varying filtering is illustrated on an application to monitoring machine vibrations (Article 11.2). Another illustration of the procedure for designing a time-varying filter is provided in the context of an application involving hands-free telephone speech signals (11.3). Another important application of time-varying filtering, namely signal enhancement, is described using an iterative algorithm based on time-frequency peak filtering (11.4).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

11.1 LINEAR TIME-FREQUENCY FILTERS [0]

11.1.1 Time-Frequency Design of Linear, Time-Varying Filters

Linear, time-varying (LTV) filters are useful in many applications, especially for weighting, suppressing, or separating nonstationary signal components. The input-output relation of an LTV filter $\mathbf{H}$ with kernel (impulse response) $h(t,t')$ reads

$$y(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} h(t,t')\,x(t')\,dt'. \tag{11.1.1}$$

The nonstationary nature of input signal $x(t)$, output signal $y(t)$, and LTV filter $\mathbf{H}$ suggests the use of time-frequency (TF) representations for analyzing, designing, and/or implementing LTV filters.

There are two fundamentally different approaches to a TF design of LTV filters, namely, the "explicit" and "implicit" design philosophies [1,2]. Both are based on a prescribed TF weight function $M(t,f)$ that provides a TF specification of the desired filtering characteristic.

- **Explicit design:** The impulse response $h(t,t')$ of the LTV filter $\mathbf{H}$ is calculated (designed) such that a TF representation of $\mathbf{H}$ is equal to or best approximates the TF weight function $M(t,f)$. In this article, the TF representation of $\mathbf{H}$ will be chosen as the *generalized Weyl symbol* (see Article 4.7). An alternative explicit design of LTV filters using the *Wigner distribution of an LTV system* has been considered in [3]. The filtering itself is performed in the time domain according to (11.1.1).

- **Implicit design:** The LTV filter $\mathbf{H}$ is designed *implicitly* during the filtering, which is a three-step analysis-weighting-synthesis procedure. First (analysis step), a *linear* TF representation—such as the short-time Fourier transform—of the input signal $x(t)$ is calculated. Second (weighting step), this TF representation is multiplied by the TF weight function $M(t,f)$. Third (synthesis step), the output signal $y(t)$ is calculated in a linear manner from the TF function obtained in Step 2. Since all processing steps are linear, the overall procedure amounts to an LTV filter.

In this article, we will consider explicit TF filter designs based on the generalized Weyl symbol [1,2,4,5] and implicit TF filter designs based on the short-time Fourier transform [1,2,6–10] and the Gabor transform [2,11,12]. In particular, we will show that the resulting filters tend to perform similarly if the TF weight function $M(t,f)$ is sufficiently smooth.

---

[0]Authors: **F. Hlawatsch** and **G. Matz**, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Gusshausstrasse 25/389, A-1040 Vienna, Austria (email: fhlawats@pop.tuwien.ac.at, g.matz@ieee.org, web: http://www.nt.tuwien.ac.at/dspgroup/time.html). Reviewers: M. Amin and D. L. Jones.

11.1.2 Explicit Design—The Generalized Weyl Filter

The *generalized Weyl symbol* (GWS) of an LTV system $\mathbf{H}$ is defined as

$$L_{\mathbf{H}}^{(\alpha)}(t,f) \triangleq \int_{-\infty}^{\infty} h\Big(t+\Big(\frac{1}{2}-\alpha\Big)\tau, t-\Big(\frac{1}{2}+\alpha\Big)\tau\Big)\, e^{-j2\pi f\tau} d\tau\,, \tag{11.1.2}$$

where α is a real-valued parameter. The special cases $\alpha = 0$ and $\alpha = 1/2$ give the *Weyl symbol* and *Zadeh's time-varying transfer function*, respectively. For *underspread* LTV systems (i.e., LTV systems that produce only moderate TF displacements), the GWS can be interpreted as a TF transfer function describing the TF weighting produced by the system (see Article 4.7). Hence, a conceptually simple TF design of an LTV filter $\mathbf{H}$ from a prescribed TF weight function $M(t,f)$ is based on setting the filter's GWS equal to $M(t,f)$ [1,2,5],

$$L_{\mathbf{H}_{\mathrm{GWF}}}^{(\alpha)}(t,f) \equiv M(t,f)\,. \tag{11.1.3}$$

The impulse response of the LTV filter $\mathbf{H}_{\mathrm{GWF}}$ thus defined is obtained via the inverse of (11.1.2), i.e.,

$$h_{\mathrm{GWF}}(t,t') = \int_{-\infty}^{\infty} M\Big(\Big(\frac{1}{2}+\alpha\Big)t+\Big(\frac{1}{2}-\alpha\Big)t', f\Big)\, e^{j2\pi f(t-t')}\, df\,. \tag{11.1.4}$$

The filter $\mathbf{H}_{\mathrm{GWF}}$ in (11.1.4) is termed *generalized Weyl filter* [1]; it depends on the choice of the GWS parameter α used in (11.1.3). In particular, the choices $\alpha = 0$ and $\alpha = 1/2$ lead to the *Weyl filter* and *Zadeh filter*, respectively. For $\alpha = 0$, a real-valued weight function $M(t,f)$ will result in a self-adjoint [13] Weyl filter.

Dependence on α. The dependence of $\mathbf{H}_{\mathrm{GWF}}$ on α effectively disappears in the case of a *smooth* TF weight function $M(t,f)$ (yielding an *underspread* LTV system $\mathbf{H}_{\mathrm{GWF}}$ as discussed in Article 4.7; also note that smoothness of $M(t,f)$ is incompatible with a sharp TF cutoff). Let $\mathbf{H}_{\mathrm{GWF}}^{(1)}$ and $\mathbf{H}_{\mathrm{GWF}}^{(2)}$ be two generalized Weyl filters designed according to (11.1.4) with GWS parameter α_1 and α_2, respectively. Then, one can show that the difference $(\mathbf{H}_{\mathrm{GWF}}^{(1)}x)(t) - (\mathbf{H}_{\mathrm{GWF}}^{(2)}x)(t)$ of the output signals of these filters satisfies

$$\frac{\big\|\mathbf{H}_{\mathrm{GWF}}^{(1)}x - \mathbf{H}_{\mathrm{GWF}}^{(2)}x\big\|_2}{\|x\|_2} \le \epsilon_1 \triangleq |\alpha_1-\alpha_2|\left\|\frac{\partial^2 M}{\partial t\,\partial f}\right\|_2,$$

where $\|\cdot\|_2$ denotes the L_2 norm. Thus, it is seen that the generalized Weyl filter design is almost independent of α if ϵ_1 is small, i.e., if $M(t,f)$ is a smooth function.

TF projection filter. Formally, (11.1.3) can be viewed as the solution to the (unconstrained) minimization problem $\mathbf{H}_{\mathrm{GWF}} = \arg\min_{\mathbf{H}} \big\|M - L_{\mathbf{H}}^{(\alpha)}\big\|_2$. Solving this minimization problem under the side constraint that $\mathbf{H}$ be an orthogonal projection operator[1] yields the *TF projection filter* $\mathbf{H}_{\mathrm{P}}$ introduced in [4]. More specifically, let $u_k(t)$ and λ_k denote the eigenfunctions and eigenvalues, respectively, of

[1] An orthogonal projection operator is characterized by being self-adjoint ($\mathbf{H}^+ = \mathbf{H}$) and idempotent ($\mathbf{HH} = \mathbf{H}$) [13].

$(\mathbf{H}_{\mathrm{GWF}} + \mathbf{H}_{\mathrm{GWF}}^{+})/2$, where $\mathbf{H}_{\mathrm{GWF}}^{+}$ denotes the adjoint of $\mathbf{H}_{\mathrm{GWF}}$ [13]. Then, the impulse response of $\mathbf{H}_{\mathrm{P}}$ can be shown [4] to equal

$$h_{\mathrm{P}}(t,t') = \sum_{k\in\mathcal{I}} u_k(t)\, u_k^*(t')\,,$$

where $\mathcal{I}$ is the set of indices k for which $\lambda_k > 1/2$.

The TF projection filter is only able to pass or suppress signal components, with no other weights possible. It is advantageous in some situations since it is capable of realizing very sharp TF cutoff characteristics. However, compared to $\mathbf{H}_{\mathrm{GWF}}$, the calculation of $\mathbf{H}_{\mathrm{P}}$ requires the additional solution of an eigenproblem. An efficient online implementation of the TF projection filter is proposed in [14].

11.1.3 Implicit Design I—The STFT Filter

An *STFT filter* consists of the following three steps [1,2,6,7,9,10]:

- *Analysis*: Calculation of the short-time Fourier transform (STFT) [9,10] of the input signal $x(t)$,

  $$F_x^{\gamma}(t,f) = \int_{-\infty}^{\infty} x(t')\, \gamma_{t,f}^*(t')\, dt'\,,$$

  where $\gamma_{t,f}(t') = \gamma(t'-t)\, e^{j2\pi f t'}$ with $\gamma(t)$ being an analysis window (see Section 2.3.1).

- *Weighting*: Multiplication of the STFT by the TF weight function $M(t,f)$, i.e., calculation of $M(t,f)\, F_x^{\gamma}(t,f)$.

- *Synthesis*: The output signal $y(t)$ is obtained via an inverse STFT [9,10],

  $$y(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[M(t',f')\, F_x^{\gamma}(t',f')\right] g_{t',f'}(t)\, dt'df'\,.$$

  Here, $g_{t,f}(t') = g(t'-t)\, e^{j2\pi f t'}$ where $g(t)$ is a synthesis window that is usually assumed to satisfy $\int_{-\infty}^{\infty} g(t)\, \gamma^*(t)\, dt = 1$ (this guarantees perfect reconstruction for $M(t,f) \equiv 1$).

These steps implement an LTV filter—hereafter denoted $\mathbf{H}_{\gamma,g}$—that depends on the TF weight function $M(t,f)$ and the windows $\gamma(t)$ and $g(t)$.

Multiwindow STFT filter. An extension of the STFT filter $\mathbf{H}_{\gamma,g}$ is the *multiwindow STFT filter* [2,8]

$$\mathbf{H}_N \triangleq \sum_{i=1}^{N} \eta_i\, \mathbf{H}_{\gamma^{(i)},g^{(i)}}\,, \qquad \text{with} \quad \sum_{i=1}^{N} \eta_i = 1\,,\ \ \eta_i \in \mathbb{R}\,.$$

This is a linear combination of N STFT filters $\mathbf{H}_{\gamma^{(i)},g^{(i)}}$ with the *same* TF weight function $M(t,f)$ but different analysis windows $\gamma^{(i)}(t)$ and different synthesis windows $g^{(i)}(t)$. Note that the STFT filter $\mathbf{H}_{\gamma,g}$ is a special case with $N=1$. Using a larger number N of STFT filters yields increased flexibility of design at the expense

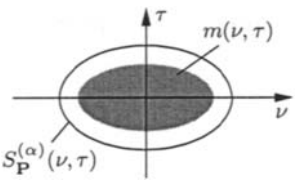


Fig. 11.1.1: Illustration of the case where the effective support of $S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)$ covers the effective support of $m(\nu,\tau)$.

of increased computational complexity. The impulse response of the multiwindow STFT filter $\mathbf{H}_N$ can be calculated as

$$h_N(t,\tilde{t}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} M(t',f')\,p(t-t',\tilde{t}-t')\,e^{j2\pi f'(t-\tilde{t})}\,dt'\,df',$$

with $p(t,t') = \sum_{i=1}^{N} \eta_i\, g^{(i)}(t)\,\gamma^{(i)*}(t')$. Furthermore, the GWS of $\mathbf{H}_N$ is obtained as

$$L_{\mathbf{H}_N}^{(\alpha)}(t,f) = M(t,f) ** L_{\mathbf{P}}^{(\alpha)}(t,f)\,, \tag{11.1.5}$$

where $**$ denotes two-dimensional convolution and $\mathbf{P}$ is the LTV system with impulse response $p(t,t')$.

Comparison with generalized Weyl filter. Comparing (11.1.5) with (11.1.3), we see that the multiwindow STFT filter $\mathbf{H}_N$ using TF weight function $M(t,f)$ is equivalent to a generalized Weyl filter using the modified TF weight function $\widetilde{M}(t,f) = M(t,f) ** L_{\mathbf{P}}^{(\alpha)}(t,f)$. For nonnegative coefficients η_i, $\widetilde{M}(t,f)$ will be a smoothed version of $M(t,f)$. However, for $\eta_i = (-1)^i$ and $N \to \infty$, it is possible to have $L_{\mathbf{H}_N}^{(\alpha)}(t,f) \to M(t,f)$ and thus $\mathbf{H}_N \to \mathbf{H}_{\mathrm{GWF}}$, i.e., the multiwindow STFT filter approaches the generalized Weyl filter using the TF weight function $M(t,f)$. It can be shown that the difference $(\mathbf{H}_N x)(t) - (\mathbf{H}_{\mathrm{GWF}} x)(t)$ of the output signals of $\mathbf{H}_N$ and $\mathbf{H}_{\mathrm{GWF}}$ (both based on the same TF weight function $M(t,f)$) satisfies

$$\frac{\|\mathbf{H}_N x - \mathbf{H}_{\mathrm{GWF}} x\|_2}{\|x\|_2} \le \epsilon_2 \triangleq \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |m(\nu,\tau)|^2\, \left|1 - S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)\right|^2 d\nu\, d\tau\right]^{1/2}. \tag{11.1.6}$$

Here, $m(\nu,\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} M(t,f)\,e^{-j2\pi(\nu t - \tau f)}\,dt\,df$ and $S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)$ is the generalized spreading function of $\mathbf{P}$ (see Article 4.7) The constant ϵ_2 is related to the operator $\mathbf{P}$ that characterizes the effect of the windows $\gamma^{(i)}(t)$, $g^{(i)}(t)$. in particular, ϵ_2 will be small if the effective support of $S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)$ covers the effective support of $m(\nu,\tau)$, so that $\left|1 - S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)\right|^2 \approx 0$ on the support of $m(\nu,\tau)$ (see Fig. 11.1.1). This is favored by a *smooth* TF weight function $M(t,f)$. Here, $m(\nu,\tau)$ is well concentrated about the origin and thus its effective support can easily be covered by $S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)$, even using a small N. Hence, for a smooth TF weight function $M(t,f)$, the generalized Weyl filter can easily be approximated by the (multiwindow) STFT filter. This will be verified experimentally in Section 11.1.6.

11.1.4 Implicit Design II—The Gabor Filter

The Gabor transform is the STFT evaluated on a TF lattice (nT, kF) with $n, k \in \mathbb{Z}$ [12]. A *Gabor filter* (see [2,11] and Article 11.2) consists of the following steps:

- *Analysis*: Calculation of the Gabor coefficients of the input signal $x(t)$ [12],

$$c_{n,k} = \int_{-\infty}^{\infty} x(t)\, \gamma_{n,k}^*(t)\, dt\,,$$

  where $\gamma_{n,k}(t) = \gamma(t - nT)\, e^{j2\pi kFt}$ with $\gamma(t)$ being a suitable analysis window.

- *Weighting*: Multiplication of the Gabor coefficients by the weights $M_{n,k} = M(nT, kF)$, i.e., calculation of $M_{n,k}\, c_{n,k}$.

- *Synthesis*: The output signal $y(t)$ is obtained via Gabor synthesis [12],

$$y(t) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} M_{n,k}\, c_{n,k}\, g_{n,k}(t)\,,$$

  where $g_{n,k}(t) = g(t - nT)\, e^{j2\pi kFt}$ with $g(t)$ being a suitable synthesis window.

This scheme implements an LTV filter that will be denoted $\widetilde{\mathbf{H}}_{\gamma,g}$. The windows $\gamma(t)$ and $g(t)$ are usually assumed to satisfy the perfect-reconstruction (biorthogonality) condition $\int_{-\infty}^{\infty} g(t)\, \gamma^*\left(t - \frac{n}{F}\right) e^{-j2\pi kt/T}\, dt = \delta_n\, \delta_k$, which presupposes critical sampling $(TF = 1)$ or oversampling $(TF < 1)$.

Multiwindow Gabor filter. An extension of the Gabor filter $\widetilde{\mathbf{H}}_{\gamma,g}$ is the *multiwindow Gabor filter* [2]

$$\widetilde{\mathbf{H}}_N \triangleq \sum_{i=1}^{N} \eta_i\, \widetilde{\mathbf{H}}_{\gamma^{(i)},g^{(i)}}\,, \qquad \text{with} \quad \sum_{i=1}^{N} \eta_i = 1\,,\ \ \eta_i \in \mathbb{R}\,,$$

i.e., a linear combination of Gabor filters $\widetilde{\mathbf{H}}_{\gamma^{(i)},g^{(i)}}$ with the same TF weights $M_{n,k}$ but different analysis windows $\gamma^{(i)}(t)$ and different synthesis windows $g^{(i)}(t)$. The Gabor filter $\widetilde{\mathbf{H}}_{\gamma,g}$ is reobtained with $N = 1$. Using a larger number N of Gabor filters allows to reduce the TF sampling density TF (cf. [15]). The impulse response of the multiwindow Gabor filter $\widetilde{\mathbf{H}}_N$ is given by

$$\tilde{h}_N(t, t') = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} M_{n,k}\, p(t - nT, t' - nT)\, e^{j2\pi kF(t-t')}\,,$$

with $p(t, t') = \sum_{i=1}^{N} \eta_i\, g^{(i)}(t) \gamma^{(i)*}(t')$, and the GWS of $\widetilde{\mathbf{H}}_N$ is

$$L_{\widetilde{\mathbf{H}}_N}^{(\alpha)}(t, f) = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} M_{n,k}\, L_{\mathbf{P}}^{(\alpha)}(t - nT, f - kF)\,,$$

where $\mathbf{P}$ is the LTV system with impulse response $p(t, t')$.

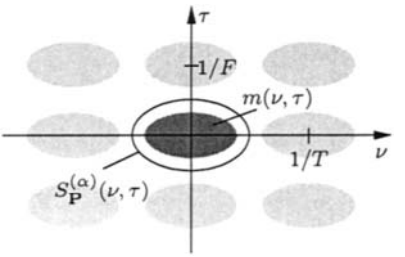


Fig. 11.1.2: Illustration of windowing and aliasing effects involved in the Gabor filter design. The dark gray ellipse indicates the effective support of $m(\nu,\tau)$. The light gray ellipses indicate the effective support of $m\big(\nu-\frac{n}{F},\tau-\frac{k}{T}\big)$ for $(n,k)\neq(0,0)$.

Comparison with generalized Weyl filter. We next analyze how close the multiwindow Gabor filter $\widetilde{\mathbf{H}}_N$ (using TF weights $M_{n,k}=M(nT,kF)$) is to the generalized Weyl filter $\mathbf{H}_{\mathrm{GWF}}$ (using TF weight function $M(t,f)$). One can show that

$$\frac{\big\|\widetilde{\mathbf{H}}_N x-\mathbf{H}_{\mathrm{GWF}}x\big\|_2}{\|x\|_2}\leq\epsilon_2+\epsilon_3\,,$$

where ϵ_2 was given in (11.1.6) and ϵ_3 is defined as

$$\epsilon_3\triangleq\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\big|S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)\big|^2\left|\sum_{\substack{n\neq0\\k\neq0}}m\Big(\nu-\frac{n}{F},\tau-\frac{k}{T}\Big)\right|^2d\nu\,d\tau\right]^{1/2}.$$

As in the case of the STFT filter, the term ϵ_2 is related to the operator $\mathbf{P}$ that describes the effect of the windows $\gamma^{(i)}(t)$, $g^{(i)}(t)$. If $M(t,f)$ is smooth so that $m(\nu,\tau)$ is well concentrated about the origin, a suitable choice of $\mathbf{P}$ allows to cover the effective support of $m(\nu,\tau)$ by the effective support of $S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)$, which results in a small value of ϵ_2 (cf. our discussion in Section 11.1.3). The additional term ϵ_3 is mainly due to potential aliasing errors which are caused by the sampling $M_{n,k}=M(nT,kF)$ that distinguishes the Gabor filter from the STFT filter. For ϵ_3 to be small, it is necessary that the effective support of $S_{\mathbf{P}}^{(\alpha)}(\nu,\tau)$ does not overlap with the periodic repetitions $m\big(\nu-\frac{n}{F},\tau-\frac{k}{T}\big)$ of $m(\nu,\tau)$ (see Fig. 11.1.2). This can be achieved if (i) $m(\nu,\tau)$ is well concentrated about the origin and thus $m\big(\nu-\frac{n}{F},\tau-\frac{k}{T}\big)$ is well localized about $\big(\frac{n}{F},\frac{k}{T}\big)$ and (ii) T and F are small enough to ensure that the periodic repetitions $m\big(\nu-\frac{n}{F},\tau-\frac{k}{T}\big)$ are sufficiently separated. For $m(\nu,\tau)$ well concentrated, the latter condition can be met even for $TF>1$. Hence, we conclude that for a smooth $M(t,f)$, the generalized Weyl filter can be accurately approximated by the (multiwindow) Gabor filter.

11.1.5 The Discrete-Time Case

While our discussion of TF filters has so far been placed in a continuous-time framework, practical implementation of these filters calls for a discrete-time formulation.

The input-output relation of a *discrete-time* LTV system $\mathbf{H}$ reads

$$y[n] = (\mathbf{H}x)[n] = \sum_{n'=-\infty}^{\infty} h[n,n']\,x[n']\,,$$

with $h[n,n']$ being the impulse response of $\mathbf{H}$. The GWS with arbitrary α is not easily reformulated in a discrete-time setting. However, for $\alpha = 0$ (Weyl symbol) and $\alpha = 1/2$ (Zadeh's time-varying transfer function), which are the main cases of practical interest, discrete-time formulations are given by

$$L_{\mathbf{H}}^{(0)}(n,\theta) = 2\sum_{m=-\infty}^{\infty} h[n+m,n-m]\,e^{-j4\pi\theta m}\,,$$

$$L_{\mathbf{H}}^{(1/2)}(n,\theta) = \sum_{m=-\infty}^{\infty} h[n,n-m]\,e^{-j2\pi\theta m}\,.$$

Here, θ denotes normalized frequency. We note that in order for $L_{\mathbf{H}}^{(0)}(n,\theta)$ to be in one-to-one correspondence to $h[n,n']$, $\mathbf{H}$ has to be a *halfband system*, i.e., an LTV system that accepts input signal components only within a specified halfband (e.g., $\theta \in [-1/4, 1/4)$) and maps them to a halfband output signal [2].

The TF system representations $L_{\mathbf{H}}^{(0)}(n,\theta)$ and $L_{\mathbf{H}}^{(1/2)}(n,\theta)$ can be used to design discrete-time LTV filters from a TF weight function $M(n,\theta)$ via an explicit filter design (cf. Section 11.1.2) [2]. The *discrete-time Zadeh filter* ($\alpha = 1/2$) is defined by setting $L_{\mathbf{H}}^{(1/2)}(n,\theta) = M(n,\theta)$; its impulse response is obtained as

$$h[n,n'] = \int_{-1/2}^{1/2} M(n,\theta)\,e^{j2\pi\theta(n-n')}\,d\theta\,.$$

In a similar manner, the *discrete-time Weyl filter* ($\alpha = 0$) is obtained as

$$h[n+m,n-m] = \int_{-1/4}^{1/4} M(n,\theta)\,e^{j2\pi\theta m}\,d\theta\,. \tag{11.1.7}$$

Since $L_{\mathbf{H}}^{(0)}(n,\theta)$ is meaningful only for halfband systems, $M(n,\theta)$ here is specified on the halfband $[-1/4, 1/4)$. According to (11.1.7), the impulse response $h[n_1,n_2]$ of the discrete-time Weyl filter is obtained only for n_1+n_2 even (since $n_1+n_2 = n+m+n-m = 2n$). If we assume $\mathbf{H}$ to be a halfband system, $h[n_1,n_2]$ is completely specified by these samples. The missing samples (for n_1+n_2 odd) could be obtained by interpolation; however, this is not necessary since the entire filtering can be performed using only the even-indexed samples [2]. We note that in some cases (especially for "chirpy" TF weight functions), the Weyl filter design results in better filtering performance than the Zadeh filter design [2].

Discrete-time versions of the implicit filter design methods from Sections 11.1.3 and 11.1.4 can be obtained in a straightforward manner; see [2,6,9,10] for discrete-time STFT filters and [2,11] for discrete-time Gabor filters.

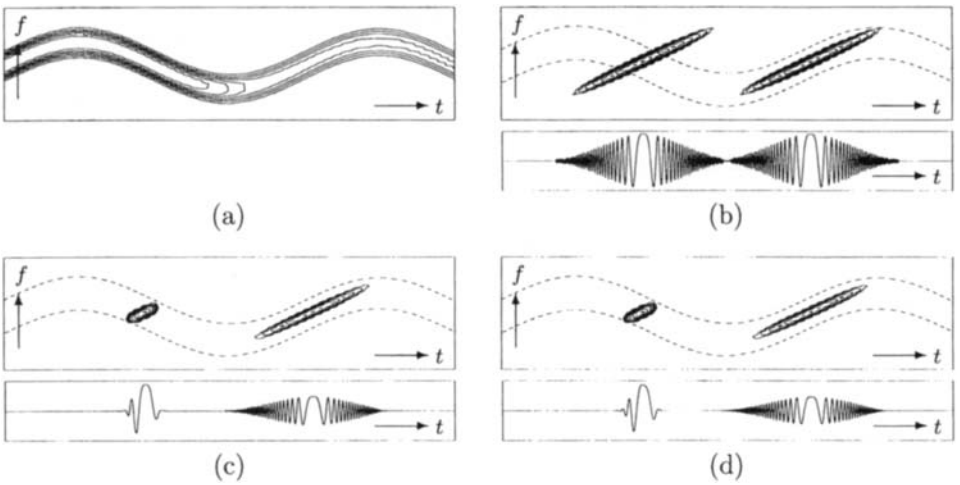


Fig. 11.1.3: Comparison of explicit and implicit TF filter designs: (a) Specified TF weight function $M(t,f)$, (b)–(d) spectrogram (top) and real part (bottom) of (b) input signal $x(t)$, (c) output signal $y(t)$ obtained with Zadeh filter, and (d) output signal $y(t)$ obtained with STFT filter. The dashed lines in the spectrograms indicate the TF pass region. The time duration is 2048 samples; the (normalized) frequency interval shown is $[-1/2, 1/2)$.

11.1.6 Simulation Results

Our first simulation example, shown in Fig. 11.1.3, compares the performance of the Zadeh filter $\mathbf{H}_{\mathrm{GWF}}$ (generalized Weyl filter with $\alpha = 1/2$) and the STFT filter $\mathbf{H}_{\gamma,g}$. The TF weight function $M(t,f)$ (see Fig. 11.1.3(a)) models a bandpass filter with sinusoidally time-varying center frequency and time-varying gain. The gain is 1 in the first (earlier) half and 1/2 in the second (later) half, with a roll-off in between. The two filters were applied to an input signal $x(t)$ consisting of two chirps (see Fig. 11.1.3(b)). The resulting output signals (shown in Figs. 11.1.3(c),(d)) are seen to conform to the specified TF weighting. Furthermore, they are effectively identical (we obtained $\|\mathbf{H}_{\gamma,g}x - \mathbf{H}_{\mathrm{GWF}}x\|_2 / \|\mathbf{H}_{\mathrm{GWF}}x\|_2 = 0.047$), which is due to the smoothness of $M(t,f)$ and confirms our approximation in (11.1.6).

The application of a multiwindow Gabor filter $\widetilde{\mathbf{H}}_N$ to speech enhancement (denoising) is considered in Fig. 11.1.4. The speech signal $s(t)$ and its noisy version $x(t) = s(t) + n(t)$ (where $n(t)$ is white noise with an SNR of 0 dB) are shown in Fig. 11.1.4(a),(b). The multiwindow Gabor filter has $N = 5$ branches and lattice parameters $T = 5.8\,\mathrm{ms}$, $F = 172.25\,\mathrm{Hz}$. The analysis/synthesis windows $\gamma^{(i)}(t) = g^{(i)}(t)$ and the branch weights η_i were chosen as discussed in [2, Section 4.6.5]. The weights $M_{n,k}$ were computed from the multiwindow Gabor coefficients $c^{(i)}_{n,k} = \int_{-\infty}^{\infty} x(t)\,\gamma^{(i)*}_{n,k}(t)\,dt$ in a signal-adaptive, online manner that does not require knowledge about the clean speech signal or its statistics [2, Section 4.7.3]. The filter output is shown in Fig. 11.1.4(c); the SNR improvement is 4.92 dB.

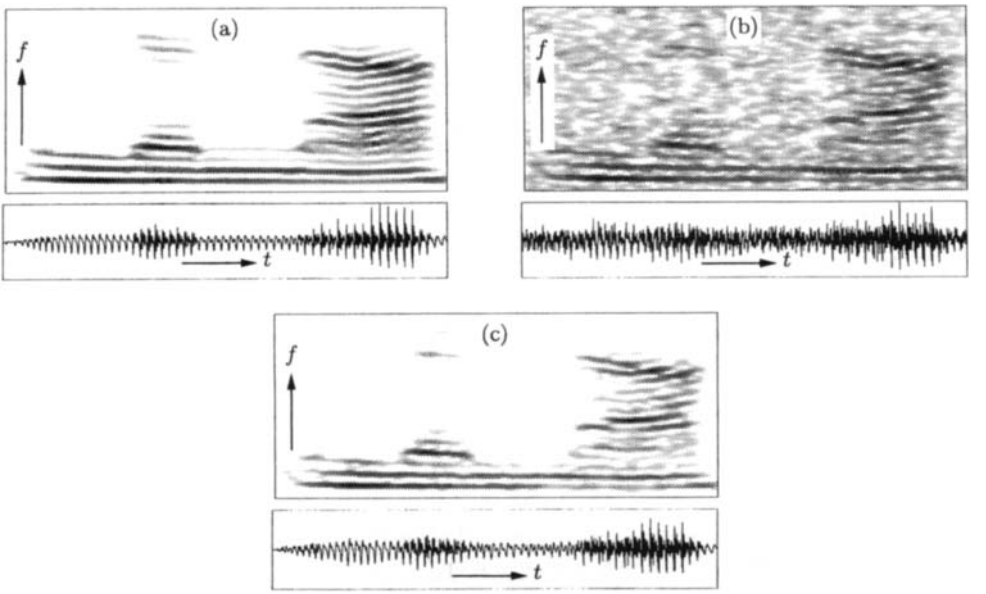


Fig. 11.1.4: Speech enhancement using a multiwindow Gabor filter $\widetilde{\mathbf{H}}_N$. The figure shows the smoothed pseudo-Wigner distribution (top) and the time-domain signal (bottom) of (a) the clean speech, (b) the noisy speech (input of $\widetilde{\mathbf{H}}_N$), and (c) the enhanced speech (output of $\widetilde{\mathbf{H}}_N$). The time duration is 4096 samples; the (normalized) frequency interval shown is $[0, 1/2)$.

11.1.7 Summary and Conclusions

We have discussed "explicit" and "implicit" time-frequency (TF) designs of linear, time-varying filters. These design methods are useful for filtering nonstationary signals if the filter characteristic can be specified in the TF domain via a TF weight function. All filters discussed (except the TF projection filter) tend to perform similarly if the TF weight function is sufficiently smooth. In the opposite case, however, different designs may result in filters that perform very differently [2].

We finally note that the application of TF filtering to nonstationary signal estimation and detection is considered in Article 12.4. Other TF approaches to time-varying filtering are described in Articles 11.2–11.4.

References

[1] W. Kozek and F. Hlawatsch, "A comparative study of linear and nonlinear time-frequency filters," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 163–166, Victoria, BC, 4–6 October 1992.

[2] G. Matz and F. Hlawatsch, "Linear time-frequency filters: On-line algorithms and applications," in *Applications in Time-Frequency Signal Processing* (A. Papandreou-Suppappola, ed.), ch. 6, pp. 205–271, Boca Raton, FL: CRC Press, 2002.

[3] 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.), ch. 9, pp. 235–287, Boston: Birkhäuser, 2001.

[4] F. Hlawatsch, *Time-Frequency Analysis and Synthesis of Linear Signal Spaces: Time-Frequency Filters, Signal Detection and Estimation, and Range-Doppler Estimation.* Boston: Kluwer, 1998.

[5] W. Kozek, "Time-frequency signal processing based on the Wigner-Weyl framework," *Signal Processing*, vol. 29, pp. 77–92, October 1992.

[6] 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 Processing*, vol. 15, pp. 193–201, September 1988.

[7] I. Daubechies, "Time-frequency localization operators: A geometric phase space approach," *IEEE Trans. Information Theory*, vol. 34, pp. 605–612, July 1988.

[8] W. Kozek, H. G. Feichtinger, and J. Scharinger, "Matched multiwindow methods for the estimation and filtering of nonstationary processes," in *Proc. IEEE Internat. Symp. on Circuits and Systems (ISCAS 96)*, vol. 2, pp. 509–512, Atlanta, GA, 12–15 May 1996.

[9] 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.), ch. 6, pp. 289–337, Englewood Cliffs, NJ: Prentice-Hall, 1988.

[10] M. R. Portnoff, "Time-frequency representation of digital signals and systems based on short-time Fourier analysis," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 28, pp. 55–69, February 1980.

[11] S. Farkash and S. Raz, "Linear systems in Gabor time-frequency space," *IEEE Trans. Signal Processing*, vol. 42, pp. 611–617, March 1994.

[12] H. G. Feichtinger and T. Strohmer, eds., *Gabor Analysis and Algorithms: Theory and Applications.* Berlin/Boston: Birkhäuser, 1998.

[13] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science.* New York: Springer, 2nd ed., 1982.

[14] G. Matz and F. Hlawatsch, "Time-frequency projection filters: Online implementation, subspace tracking, and application to interference excision," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'02)*, pp. 1213–1216, Orlando, FL, 13–17 May 2002.

[15] 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.), ch. 12, pp. 381–407, Berlin/Boston: Birkhäuser, 1998.

11.2 TIME-VARYING FILTER VIA GABOR EXPANSION[0]

11.2.1 Filtering a Six-Cylinder Engine Sound

One of the most important applications of the Gabor expansion is in time-varying filtering. It is well understood that the Fourier transform can be employed to effectively study a signal's frequency behavior as long as the signal's frequency content does not evolve during the period of observation. However, this is not the case when the signal's frequency changes rapidly.

Fig. 11.2.1 illustrates a sound waveform recorded during the run-up of a six-cylinder engine. Intuitively, the sound will be created not only by the engine rotation, but also by the other parts that vibrate due to the engine rotation. The sound waveform plotted in the bottom is indeed a combination of all kinds of vibrations caused by engine rotation. Moreover, the vibration frequencies are multiples of the fundamental frequency—the engine rotation speed. When the engine speed is constant, the classical Fourier transform can well isolate the vibrations. When the engine runs up/down, the fundamental frequency and its multiples increase/decrease with time. The corresponding frequency bandwidths will become wide and overlap each other, as shown in the left plot. Consequently, the Fourier transform based power spectrum will no longer be able to distinguish different individual vibrations. In the automobile industry and some other industries, such time-varying harmonics are named as orders to distinguish them from the time-invariant harmonics. To evaluate a vibration, engineers use adaptive filters, such as Kalman filter [1], to extract the time waveform corresponding to a particular order. Based on the time waveform, engineers can then further compute other information, such as phase and amplitude. Recently it has been discovered that the order tracking process can also be effectively achieved by the Gabor expansion (see Section 2.3.3 on the connection between the Gabor transform and the STFT). The middle plot of Fig. 11.2.1 shows the magnitudes of the Gabor coefficients for the sound signal. While neither the time waveform nor the Fourier transform based power spectrum gives us an idea about the structure of the engine sound, orders have distinct signatures in the joint time-frequency plot. Taking this advantage, we can select the desired Gabor coefficients and perform the Gabor expansion to obtain the corresponding time waveform. Such processing can be thought of as time-varying filtering.

11.2.2 Discrete Gabor Expansion

Before investigating the nature of the Gabor expansion based time-varying filter, first we shall briefly review the discrete Gabor expansion. For a discrete time signal s[k], its Gabor expansion is defined as [2]

$$s[k] = \sum_{m} \sum_{n=0}^{N-1} c_{m,n}\, h[k - mT]\, e^{j2\pi nk/N} \tag{11.2.1}$$

[0]Author: **Shie Qian**, National Instruments Corp., Austin, TX 78759, USA (shie.qian@ni.com). Reviewers: X.-G. Xia and F. Munk.

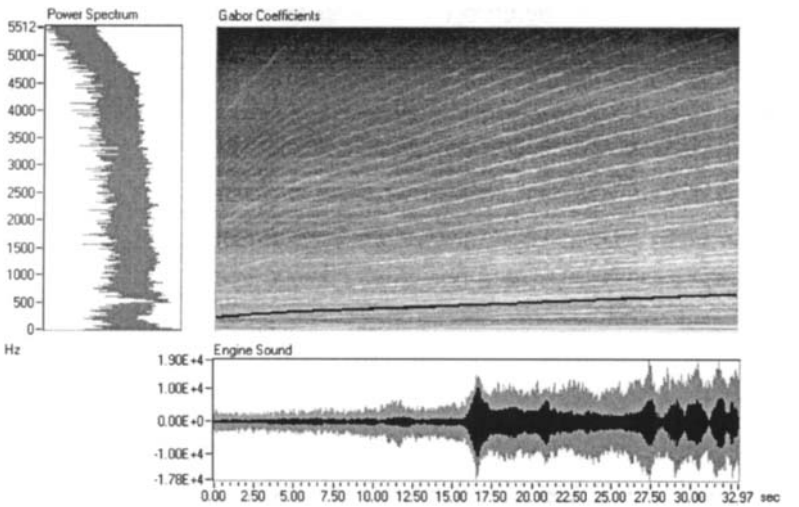


Fig. 11.2.1: Gabor expansion based Time-varying Filter for Order Tracking. The bottom plot illustrates the original engine run-up sound (light color) vs. the extracted sixth-order (dark color) time waveforms. The dark line in the Gabor coefficients plot marks the Gabor coefficients corresponding to the sixth order.

where the Gabor coefficients, $c_{m,n}$, are computed by the regular STFT (short-time Fourier transform), i.e.,

$$c_{m,n} = \sum_m s[k]\,\gamma^*[k - mT]\,e^{-j2\pi nk/N}. \tag{11.2.2}$$

The parameters T and N denote the time sampling interval (or decimation) and the number of frequency bins (or bands). The ratio of N and T determines the sampling rate. When this ratio is equal to one, we have critical sampling; in this case there is no redundancy for the resulting Gabor coefficients. When the sampling rate is greater than one, we have oversampling. For stable reconstruction, the ratio of N and T must be greater than or equal to one.

Note that the window functions $h[k]$ and $\gamma[k]$ are exchangeable. In other words, either of them can be used as the analysis or the synthesis function. They satisfy the so-called dual relation. The central issue of the Gabor expansion is how to compute the dual function for a given function (either $h[k]$ or $\gamma[k]$). Over the years, many schemes of computing the dual function have been proposed. Each method has its pros and cons. A special feature of the approach presented here is that the resulting dual window function always has the same length as the given function.

Without loss of generality, let us assume that the given window function is $h[k]$ with L points. Moreover, $h[k]$ has unit energy. Then, the corresponding dual

function $\gamma[k]$ can be solved by T independent linear equations, i.e.,

$$A_l \vec{\gamma}_l = \vec{u}_l \ , \qquad l = 0, 1, ..., T-1 \tag{11.2.3}$$

where the elements of the matrix A_k, and vectors $\vec{\gamma}_l$ and $\vec{u}_l$ are defined as

$$\begin{aligned} a_l[q,p] &= \tilde{a}[l + pT + qN] \\ \gamma_l[p] &= \gamma[l + pT] \\ u_l[q] &= (N^{-1}, 0, 0, ...)^T \end{aligned}$$

where $0 \le p < L/T$. The periodic auxiliary function $\tilde{a}[k]$ is defined as

$$\tilde{a}[k + i(2L - N)] = \begin{cases} h[k] & 0 \le k < L, \\ 0 & L \le k < 2L - N, \end{cases} \qquad 0 \le q < \frac{2L}{N} - 1. \tag{11.2.4}$$

If the window length is equal to the signal length, the periodic auxiliary function $\tilde{a}[k]$ is simply

$$\tilde{a}[k + iL] = h[k] \ , \quad 0 \le q < \frac{L}{N}.$$

Note that the solution of (11.2.3) is not unique for oversampling. To ensure that the analysis and synthesis window functions are both concentrated in the joint time-frequency domain, we require that the dual function is optimally similar, in the sense of LMSE (least mean square error), to the given window function, i.e.,

$$\Gamma = \min_{A\gamma = u} \left\| \frac{\vec{\gamma}}{|\vec{\gamma}|} - \vec{h} \right\|.$$

When the error is small, i.e. when $\vec{\gamma} \approx \vec{h}$, eq. (11.2.2) becomes

$$c_{m,n} = \sum_m s[k]\, h^*[k - mT]\, e^{-j2\pi nk/N}. \tag{11.2.5}$$

Then (11.2.5) and (11.2.1) form an *orthogonal-like Gabor transform pair*. In this case, the Gabor coefficients $c_{m,n}$ are exactly the signal's projection on the synthesis window function $h[k]$.

Since at critical sampling the pair of dual functions cannot be simultaneously concentrated in both the time and frequency domains, usually we always employ the oversampling scheme. Consequently, the resulting Gabor transformation is redundant. In this case, the Gabor coefficients will be the sub-space of two-dimensional functions. In other words, for an arbitrary two-dimensional function, there may be no corresponding time waveform. For example, assume that we have a modified two-dimensional function

$$\hat{c}_{m,n} = \phi_{m,n} c_{m,n}$$

where $\phi_{m,n}$ denotes a binary mask function, equal to either zero or one. Apply the Gabor expansion (11.2.1) to obtain

$$\hat{s}[k] = \sum_m \sum_{n=0}^{N-1} \hat{c}_{m,n}\, h[k - mT]\, e^{j2\pi nk/N}. \tag{11.2.6}$$

Then we find that

$$\sum_m \hat{s}[k]\,\gamma^*[k-mT]\,e^{-j2\pi nk/N} \neq \hat{c}_{m,n}.$$

The Gabor coefficients of the reconstructed time waveform $\tilde{s}[k]$ will not be inside the masked area determined by the mask function $\phi_{m,n}$.

11.2.3 Time-Varying Filtering

To overcome this problem, we will introduce an iteration approach as follows. That is, for a set of two-dimensional Gabor coefficients, first determine a binary mask matrix. Apply the mask to the two-dimensional Gabor coefficients to reserve desirable and remove unwanted coefficients. Then, compute the Gabor expansion. Once the time waveform is obtained, compute the new Gabor coefficients. Repeat this process until the time waveforms converge. For the sake of an efficient representation, let us rewrite the pair of the Gabor expansion (11.2.1) and (11.2.2) in matrix form, i.e.,

$$C = G\vec{s}$$
$$\vec{s} = H^T G\vec{s}$$

where H and $\mathbf{G}$ denote the analysis and synthesis matrices, respectively. Note that for oversampling

$$H^T G = I, \quad GH^T \neq I.$$

Then, the iterative process can be described as

$$\begin{aligned}
\hat{C}^1 &= \mathbf{\Phi} C \\
\vec{s}^1 &= H^T \hat{C}^1 \\
C^2 &= G\vec{s}^1 = GH^T\mathbf{\Phi}C \\
\hat{C}^2 &= \mathbf{\Phi}C^2 \\
\vec{s}^2 &= H^T\hat{C}^2 \\
C^3 &= G\vec{s}^2 = GH^T\mathbf{\Phi}GH^T\mathbf{\Phi}C = (GH^T\mathbf{\Phi})^2 C \\
&\cdots \\
C^k &= (GH^T\mathbf{\Phi})^{k-1}C
\end{aligned}$$

It can be shown [3] that *if and only if*

$$\sum_{i=0}^{\frac{L}{N}-1} \gamma^*[iN+k]h[iN+k+mT] = \sum_{i=0}^{\frac{L}{N}-1} h^*[iN+k]\gamma[iN+k+mT] \qquad (11.2.7)$$

for $0 \leq k < N$ and $0 \leq m < M$, where M is equal to the number time sampling points and L denotes the window length, *then*

1. C^k and $\vec{s}^k$ converge.

2. $C^k = \mathbf{\Phi} C^k, k \to \infty$. That is, the support of C^k in the time-frequency domain is inside the masked area.

Two trivial cases for (11.2.7) are

1. the critical sampling, $N = T$. Note that in this case the analysis and synthesis windows cannot both be localized in the joint time-frequency domain.

2. $\gamma[k] = h[k]$. In this case the Gabor coefficients $\mathbf{C}^2$, after the first iteration, are closest in the LMSE sense to the masked Gabor coefficients $\mathbf{\Phi C}$ (that is, desirable Gabor coefficients).

The second case usually implies heavy oversampling (or huge redundancy) and is therefore impractical due to the computation speed and memory consumption. Usually, we pursue the orthogonal-like representation introduced early. It has been found that for those commonly used window functions, such as Gaussian and Hanning windows, the difference between the analysis and synthesis windows would be negligible when the oversampling rate is four.

11.2.4 Numerical Simulation

Fig. 11.2.1 illustrates the application of the Gabor expansion based time-varying filter to order tracking. The dark waveform in the bottom plot depicts the sixth order extracted by the Gabor expansion based time-varying filter. In this example, $h[k]$ is the Hanning window. The oversampling rate was selected high enough (four) so that the dual function $\gamma[k]$ has a form that is almost identical to $h[k]$. It has been found that after a few iterations, the difference (LMSE) between $\bar{s}^k$ and $\bar{s}^{k-1}$ reduces to 10^4. The resulting time waveform is found to be almost identical to that computed by other methods.

11.2.5 Summary and Conclusions

In this article, the basic concept and a real-world application of the Gabor expansion based time-varying filter are introduced. Compared with other time-varying filter schemes, the method presented here is much simpler and thereby has great potential in wide range of applications.

References

[1] H. Vold and J. Leuridan, "High resolution order tracking at extreme slew rates using Kalman tracking filters," in *Proc. Noise & Vibration Conf. & Exposition*, Soc. of Automotive Engineers, Traverse City, MI, 17–20 May 1993. SAE paper no. 931288.

[2] S. Qian and D. Chen, *Joint Time-Frequency Analysis: Methods & Applications*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[3] X.-G. Xia and S. Qian, "Convergence of an iterative time-variant filtering based on discrete Gabor transform," *IEEE Trans. Signal Processing*, vol. 47, pp. 2894–2899, October 1999.

11.3 TIME-FREQUENCY FILTERING OF SPEECH SIGNALS IN HANDS-FREE TELEPHONE SYSTEMS [0]

Time-varying filtering of noisy speech signals is a very attractive challenge, with the main question: what does the most appropriate time-varying filter scheme look like? Speech signals are of highly nonstationary and multicomponent nature. If we deal with filtering of noisy speech signals, as they occur in hands-free telephone systems, then the desired scheme would provide a signal-to-noise ratio (SNR) greater than approximately 12dB. At the same time it should be suited for real-time implementation, with time delay less than 39ms for mobile telephony, and 2ms for circuit-switched telephony. The second requirement can cause additional difficulties and restrictions on finding an appropriate time-varying filter procedure.

The most commonly used approach in the filtering of speech signals is the so called quasi-stationary approach, where it is assumed that the signal is stationary in the time interval T, with T between 20 ms and 40ms being often used [1,2]. In this interval of time, classical speech enhancement schemes such as those given in Table 11.3.1 are used [3,4].

The noise is reduced by applying frequency-dependent suppression factors according to the various filtering rules given in Table 11.3.1. From the aspect of time-varying filtering we can say that this technique is quasi time-varying filtering. Thus, we have a sliding window of duration T along the signal where the filtering is performed after every T or after every $T/2$ (the second case is used in an overlap-add scheme in order to avoid block effects).

From the point of view of time-frequency analysis, having in mind the high nonstationarity of speech signals, we can conclude that the quasi-stationary approach of filtering is approximate in nature, and that it will more or less satisfy subjective perception requirements. If we want to achieve more accurate and more objective filtering of speech signals, time-varying filtering needs to be applied. Since a unique definition of time-frequency spectra does not exist (as is shown by the variety of TFDs derived in Chapter 2), several approaches to time-varying filtering have been proposed. We will use the one based on the Wigner distribution. It uses the Wigner spectrum, where the statistically independent cross-terms in Wigner distribution are averaged out. However, in order to calculate the Wigner spectrum it is necessary to have many different realizations of the same random process at a given instant. Obviously, in the case of real-time applications, the processing has to be based on a single noisy speech realization. It is the reason for using an approximation, in the sense that the Wigner spectrum is replaced by a cross-terms free (reduced) time-frequency distribution. According to the additional criterion of realization simplicity, special attention will be devoted to the filtering based on the spectrogram and distributions whose realization is directly related to the spectrogram. The use of other reduced interference time-frequency distributions in the

[0]Author: **Srdjan Stanković**, Elektrotehnicki fakultet, University of Montenegro, 81000 Podgorica, Montenegro (srdjan@cg.ac.yu). Reviewers: J. Tilp and LJ. Stanković.

Table 11.3.1: Filter transfer function for different algorithms, where $S_{xx}^2(\omega)$ and $S_{\epsilon\epsilon}^2(\omega)$ are the power spectra of the noisy signal and noise respectively, and λ is an overestimation factor.

Algorithm	*Wiener*	*Spectral subtraction*	*Maximum likelihood*	*Magnitude subtraction*
Filter tran. fun.	$1-\frac{S_{\epsilon\epsilon}^2(\omega)}{S_{xx}^2(\omega)}$	$\sqrt{1-\frac{\lambda S_{\epsilon\epsilon}^2(\omega)}{S_{xx}^2(\omega)}}$	$\frac{1}{2}[1+\sqrt{1-\frac{S_{\epsilon\epsilon}^2(\omega)}{S_{xx}^2(\omega)}}]$	$1-\sqrt{\frac{S_{\epsilon\epsilon}^2(\omega)}{S_{xx}^2(\omega)}}$

filtering, in place of the spectrogram, is straightforward.

11.3.1 Time-Variant Filtering of Speech Signals

By analogy with the filtering of stationary signals, nonstationary time-varying filtering of a noisy signal can be defined by [5–7]:

$$(Hx)(t)=\int_{-\infty}^{\infty} h(t+\frac{\tau}{2},t-\frac{\tau}{2})x(t+\tau)d\tau. \tag{11.3.1}$$

The signal

$$x(t)=s(t)+\epsilon(t)$$

is a noisy one with the desired signal $s(t)$ and the noise $\epsilon(t)$. Impulse response of the time-varying filter is $h(t+\frac{\tau}{2},t-\frac{\tau}{2})$. The optimal transfer function

$$L_H(t,f)=\int_{-\infty}^{\infty} h(t+\frac{\tau}{2},t-\frac{\tau}{2})x(t+\tau)e^{-j2\pi f\tau}d\tau$$

is defined by the relation [8–11]:

$$\overline{W}_{sx}(t,f)=L_H(t,f)\overline{W}_{xx}(t,f), \tag{11.3.2}$$

where

$$\overline{W}_{xx}(t,f)=E\{W_{xx}(t,f)\}=\int_{-\infty}^{\infty} E\{x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})\}e^{-j2\pi f\tau}d\tau \tag{11.3.3}$$

is the mean value of the Wigner distribution $W_{xx}(t,f)$ of signal $x(t)$ (i.e. the Wigner spectrum of signal $x(t)$ [12]). We can conclude that (11.3.2) is of the same form as the Wiener filter for the stationary case.

If the signal and noise are not correlated, we have:

$$L_H(t,f)=\frac{\overline{W}_{ss}(t,f)}{\overline{W}_{ss}(t,f)+\overline{W}_{\epsilon\epsilon}(t,f)}. \tag{11.3.4}$$

Consider now relation (11.3.4). Obviously, the mean value $E\{W_{ss}(t,f)\} = \overline{W}_{ss}(t,f)$ will eliminate uncorrelated cross-terms in the Wigner distribution, since

$$E\{s_i(t+\frac{\tau}{2})s_j^*(t-\frac{\tau}{2})\} = 0 \quad for \quad i \neq j,$$

as long as components $s_i(t)$ and $s_j(t)$ are not correlated [12]. However, if we have to perform filtering on the base of a single realization, the Wigner distribution should be used instead of the Wigner spectrum in (11.3.4). For filtering of multicomponent signals equation (11.3.4) is useless because emphatic cross-terms will appear. The problem of cross-terms will be partially overcome if we modify the definition (11.3.4) so that we apply some of the cross-terms reduced distributions $\rho(t,f)$ instead of the Wigner distribution. In this case we have:

$$L_H(t,f) = \frac{\rho_{ss}(t,f)}{\rho_{ss}(t,f) + \rho_{\epsilon\epsilon}(t,f)}. \tag{11.3.5}$$

It is clear that definition (11.3.5) is an approximation of (11.3.4) with $\rho(t,f)$ approximating the Wigner spectrum.

In order to obtain a more efficient filter, for numerical implementation, the previous definitions can be slightly modified by using their pseudo form:

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t+\frac{\tau}{2}, t-\frac{\tau}{2})w(\tau)x(t+\tau)d\tau. \tag{11.3.6}$$

Here, a lag window $w(\tau)$ is introduced. It can be shown that, for frequency modulated signals, $w(\tau)$ does not influence the output signal $(Hx)(t)$ if $w(0) = 1$ [6]. By using Parseval's theorem, (11.3.6) can be written in the form:

$$(Hx)(t) = \int_{-\infty}^{\infty} L_H(t,f)F_x(t,f)df \tag{11.3.7}$$

where

$$F_x(t,f) = \int_{-\infty}^{\infty} x(t+\tau)w(\tau)e^{-j2\pi f\tau}d\tau$$

is the short-time Fourier transform of the signal $x(t)$.

The choice of $\rho(t,f)$, in (11.3.5) will play a crucial role in the time-varying filter scheme. Obviously, for an efficient time-varying filtering, it is desired that the chosen $\rho(t,f)$ satisfies three main conditions:

(1) satisfactory noise reduction,

(2) appropriateness for real-time realization,

(3) its auto-terms are close to those in the Wigner spectrum.

The simplest and most commonly used $\rho(t,f)$, for which the real-time application is very well studied, is the spectrogram (see Section 2.3.1), which is the squared modulus of the short-time Fourier transform:

$$S_x(t,f) = |F_x(t,f)|^2 = \left|\int_{-\infty}^{\infty} x(t+\tau)w(\tau)e^{-j2\pi f\tau}d\tau\right|^2$$

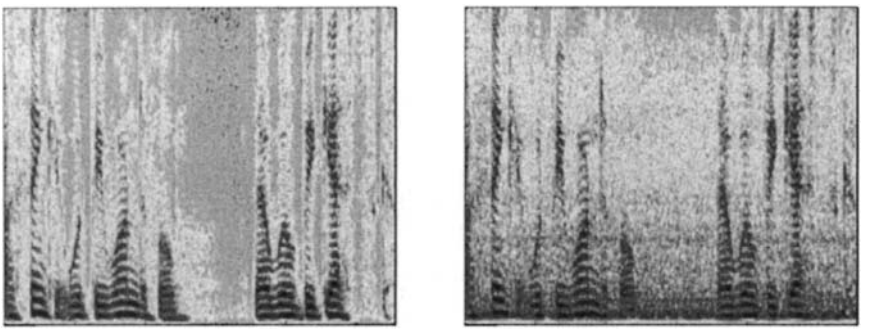


Fig. 11.3.1: Spectrogram of: (a) the clean speech signal, (b) the noisy signal filtered by a high-pass filter.

The main problem of using the short-time Fourier transform (and the spectrogram) is in determination of the window width $w(t)$. A narrow window produces better time resolution, while a wider window gives better frequency resolution. The window should be chosen by a compromise of these two opposite requirements.

Having in mind that a speech signal is approximately stationary within the interval T between 20ms and 40ms, for a sampling rate of $f_s = 8$kHz, we conclude that we can use a lag window width of $N = 256$ samples, corresponding to $T = 32$ms. In order to achieve a more accurate calculation of integral (11.3.7), zero padding up to 1024 samples will be used.

Consider now the spectrogram-based filtering of noisy speech signals, recorded in a car cruising along the highway. Estimations of the spectrogram of noise are performed in only one time instant during a speech pause. This assumption is made in order to have the worst filtering situation as in a real case. Since the noisy signal contains significant noise components in the low frequency range (below 98Hz) where, in our application, no speech components exist, the signal is prefiltered by using a high-pass filter with cut-off frequency 98 Hz. In this realization we will apply the time-varying Wiener filter definition (11.3.5), with the spectrogram instead of $\rho(t, f)$, and the time-varying version of the spectral subtraction definition:

$$L_{\mathrm{HW}}(t, f) = 1 - \frac{S_\epsilon(t, f)}{S_x(t, f)}. \tag{11.3.8}$$

$$L_{\mathrm{HSS}}(t, f) = \sqrt{1 - \lambda \frac{S_\epsilon(t, f)}{S_x(t, f)}} \tag{11.3.9}$$

In (11.3.9) λ is an overestimation factor applied in order to give some correction of the errors caused by the assumption that the noise is stationary in the interval between two pause estimations. The value $\lambda = 4$ is used. Modifications of the equations (11.3.8) & (11.3.9) are used after introducing a spectral floor [2]:

$$L_{\mathrm{HW}}(t, f) = \max\{L_{\mathrm{HW}}(t, f), \beta\} \tag{11.3.10}$$

and

$$L_{\mathrm{HSS}}(t,f) = \max\{L_{\mathrm{HSS}}(t,f), \beta\}. \tag{11.3.11}$$

In our examples the spectral floors are set to $\beta = 0.12$ and $\beta = 0.08$ (in (11.3.10), and (11.3.11) respectively).

Note that by increasing λ, better noise reducing is obtained, but the distortion of signal becomes significant. By increasing β more noise remains in signal, but speech distortion is audible. Thus, these two factors are chosen by compromise.

The time-frequency representations of a clean signal and a noisy signal are shown in Fig. 11.3.1(a) and (b), respectively. Time-frequency representation of denoised signal, filtered by using the time-varying Wiener filtering, and the time-varying spectral subtraction filtering, based on the spectrogram, are shown in Fig. 11.3.2(a) and (b). It is obvious that the noise suppression is better when the time-varying spectral subtraction filter definition is used, because overestimation factor λ provides better estimation of the spectrogram of noise.

Now, there is the question whether it is possible to use some other time-frequency distributions, in order to further improve the filtering results. The answer is yes. Namely, we can use reduced interference distributions which belong to the general Cohen class[1] of distributions [13]:

$$\rho_{xx}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\tau,\nu)x(u+\frac{\tau}{2})x^*(u-\frac{\tau}{2})e^{-j2\pi\nu t}e^{-j2\pi f\tau}e^{j2\pi\nu u}dud\theta d\tau \tag{11.3.12}$$

where the kernel $g(\tau,\nu)$ specifies the distribution. The most commonly used distributions include the Choi-Williams distribution, Zao-Atlas-Marks distribution, Born-Jordan distribution, Zhang-Sato distribution, S-method, etc.

When we use the reduced interference distributions, it is important to know that, in the case of a noisy signal, the distance between two auto-terms during voiced segments of speech is approximately equal to the value of the fundamental frequency in the case of a signal without noise. In the noisy case, we also have harmonically shaped components of noise, which can occur between the auto-terms of speech, causing additional cross-terms and errors in filtering [14].

A very simple and flexible implementation can be obtained by using the S-method (SM) [15], [Article 6.2], whose realization is straightforwardly based on the short-time Fourier transform. The result of this fact is that time-varying filtering based on the SM is a simple extension of the spectrogram-based filtering. Additionally, the SM of multicomponent signals:

$$x(t) = \sum_{i=1}^{N} x_i(t), \tag{11.3.13}$$

can assume the form $\mathrm{SM}_{xx}(t,f) \cong \sum_{i=1}^{N} W_{x_i x_i}(t,f)$, being a desired approximation of the Wigner distribution auto-terms.

[1] That is, the quadratic class; see p. 68n.

The SM is defined in the form:

$$\mathrm{SM}_{xx}(t,f) = 2\int_{-\infty}^{\infty} P(\theta)F_x(t,f+\theta)F_x^*(t,f-\theta)d\theta. \tag{11.3.14}$$

where $P(\theta)$ is a rectangular window in the frequency domain.

Discretization of the SM (11.3.14), taking a rectangular window for $P(l)$, produces:

$$\begin{aligned}\mathrm{SM}_{xx}(n,k) &= \sum_{l=-L}^{L} F_x(n,k+l)F_x^*(n,k-l) \\ &= |F_x(n,k)|^2 + 2\,\mathrm{Re}\{\sum_{l=1}^{L} F_x(n,k+l)F_x^*(n,k-l)\},\end{aligned} \tag{11.3.15}$$

From the previous equation we see that the SM realization is based on the spectrogram. Thus, filtering based on the SM will be a straightforward extension of the previously considered filter schemes:

$$L_{\mathrm{HW}}(t,f) = \max\left\{1 - \frac{\mathrm{SM}_\epsilon(t,f)}{\mathrm{SM}_x(t,f)}, \beta\right\} \tag{11.3.16}$$

and

$$L_{\mathrm{HSS}}(t,f) = \max\left\{\sqrt{1 - \lambda\frac{\mathrm{SM}_\epsilon(t,f)}{\mathrm{SM}_x(t,f)}}, \beta\right\}. \tag{11.3.17}$$

In our experiments we have used the SM with $L = 3$, and spectral floors $\beta = 0.12$ and $\beta = 0.08$, respectively [16].

The denoised signals by using the time-varying Wiener filtering and the time-varying spectral subtraction filtering, based on the SM, are shown in Fig. 11.3.2(c) and (d). By comparing the results with the ones produced by using the spectrogram based filtering, the improvements are obvious. It is important to note that the SM has a form very suitable for simple hardware realization. This property is attractive for on-line applications.

11.3.2 Summary and Conclusions

Time-varying filtering of speech signals disturbed by car noise is presented. On the base of the time-varying Wiener filter form, the time-varying spectral subtraction form of filtering is introduced. The filtering is performed on the base of the spectrogram and the S-method. The proposed filter schemes are efficient and suitable for hardware realization.

References

[1] C. Breining, P. Dreiseitel, E. Haensler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp, "Acoustic echo control: An application of very-high-order adaptive filters," *IEEE Signal Processing Magazine*, vol. 16, pp. 42–69, July 1999.

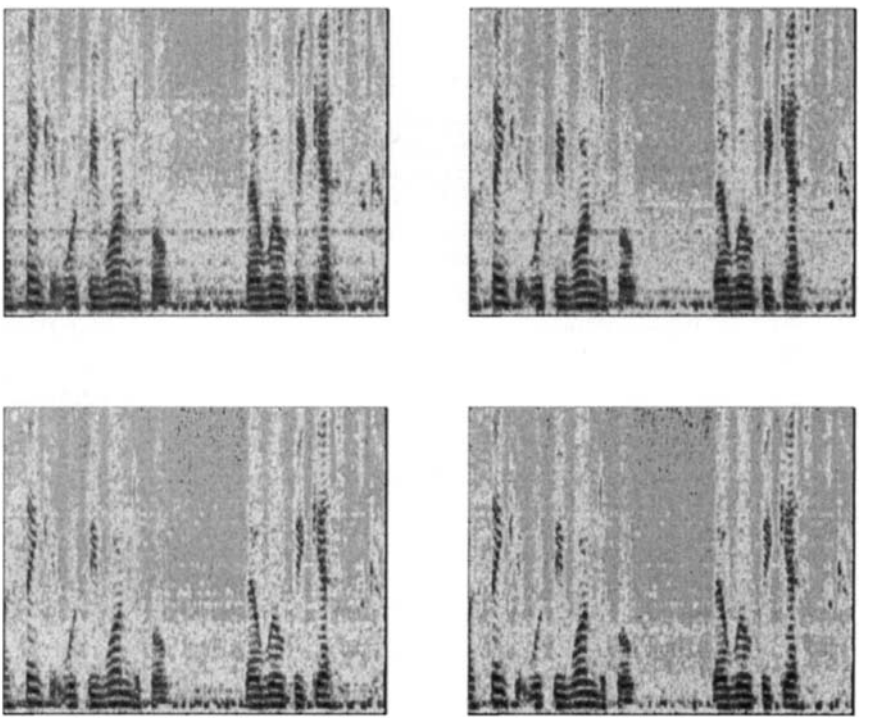


Fig. 11.3.2: Denoised signal obtained by filtering based on: (a) time-varying Wiener filter using the spectrogram, (b) time-varying spectral subtraction using the spectrogram, (c) time-varying Wiener filter using the SM, (d) time-varying spectral subtraction using the SM.

[2] S. V. Vaseghi, *Advanced Signal Processing and Digital Noise Reduction.* Wiley and Teubner, 1996.

[3] J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proc. IEEE*, vol. 67, pp. 1586–1604, December 1979.

[4] J. Yang, "Frequency domain noise suppression approaches in mobile telephone systems," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'93)*, vol. 2, pp. 363–366, Minneapolis, 27–30 April 1993.

[5] G. Matz, F. Hlawatsch, and W. Kozek, "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes," *IEEE Trans. Signal Processing*, vol. 45, pp. 1520–1534, June 1997.

[6] L. Stanković, S. Stanković, and I. Djurović, "Space/spatial frequency based filtering," *IEEE Trans. Signal Processing*, vol. 48, pp. 2343–2352, August 2000.

[7] H. L. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. I: "Detection, Estimation, and Linear Modulation Theory". New York: Wiley, 1968. Reprinted 2001.

[8] H. Kirchauer, F. Hlawatsch, and W. Kozek, "Time-frequency formulation and design of nonstationary Wiener filters," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'95)*, pp. 1549–1552, Detroit, 9–12 May 1995.

[9] W. Kozek, "Time-frequency signal processing based on the Wigner-Weyl framework," *Signal Processing*, vol. 29, pp. 77–92, October 1992.

[10] A. Papoulis, *Signal analysis*. New York: McGraw-Hill, 1977.

[11] L. Stanković, "On the time-frequency analysis based filtering," *Ann. Télécommunications*, vol. 55, pp. 216–225, May/June 2000.

[12] 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.), pp. 211–267, Amsterdam: Elsevier, 1997.

[13] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[14] S. Stanković, "About time-variant filtering of speech signals with time-frequency distributions for hands-free telephone systems," *Signal Processing*, vol. 80, pp. 1777–1785, September 2000.

[15] L. Stanković, "A method for time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 225–229, January 1994.

[16] S. Stanković and J. Tilp, "Time-varying filtering of speech signals using linear prediction," *Electronics Letters*, vol. 36, pp. 763–764, April 2000.

11.4 SIGNAL ENHANCEMENT BY TIME-FREQUENCY PEAK FILTERING [0]

11.4.1 Signal Enhancement and Filtering

Time-frequency peak filtering (TFPF) may be regarded as an unconventional alternative to the filtering methods described earlier in this chapter [1]. The signals considered are assumed to be sums of arbitrary numbers of band-limited non-stationary components in additive noise. For high SNR situations many signal processing algorithms work well but most perform poorly when SNR decreases below a given threshold [2]. In this case, signal enhancement algorithms are required to improve the SNR by reducing the distorting effects of noise. To this effect, both adaptive and fixed methods have been developed in the case of non-stationary signals in noise. Adaptive techniques are generally superior in performance to fixed methods, but they perform poorly in certain conditions, such as filtering of a non-stationary signal whose spectral content changes rapidly with time. For example, filters designed using a *least-mean-square* (LMS) approach may not adapt quickly enough to track the rapidly changing signal due to the delayed convergence of the algorithm. Further, adaptive methods require that the structure of the filter (such as the number of the taps) and an estimate of SNR be imposed for optimal performance. This is often not possible as assuming a model may lead to suboptimal results and even to erroneous conclusions about the signal. This suggests the need for a more general filtering method when the SNR is low and the underlying signal statistics vary rapidly with time.

The TFPF method is based on encoding the noisy signal as the IF of a unit amplitude frequency modulated (FM) analytic signal. The instantaneous frequency (IF) of the analytic signal is then estimated using standard time-frequency peak detection methods [2] to obtain an estimate of the underlying deterministic signal. For some signals, TFPF using a windowed WVD results in a significant enhancement of signals for SNR as low as -9 dB.

11.4.2 Time-Frequency Peak Filtering

11.4.2.1 Background and Definitions

Let us consider signals expressed as follows:

$$s(t) = x(t) + n(t) = \sum_{k=1}^{p} x_k(t) + n(t) \tag{11.4.1}$$

where $n(t)$ is an additive white Gaussian noise (WGN) and $x_k(t)$ are band-limited non-stationary deterministic components that may have overlapping frequency spectra. It is desired to recover the signal $x(t)$ given the observation of $s(t)$.

[0] Authors: **Boualem Boashash** and **Mostefa Mesbah**, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Australia (b.boashash@qut.edu.au, m.mesbah@qut.edu.au). Reviewers: M. G. Amin and D. L. Jones.

The IF of an analytic signal, $z(t) = a(t)e^{j2\pi\phi(t)}$, is defined in Chapter 1 and reference [2]:

$$f_z(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} \tag{11.4.2}$$

where $\phi(t)$ is the instantaneous phase and $a(t)$ is the instantaneous amplitude of the analytic signal $z(t)$ which can be expressed as [3]:

$$z(t) = a(t)e^{j2\pi\int_{-\infty}^{t} f_z(\lambda)d\lambda} \tag{11.4.3}$$

Among the existing techniques for IF estimation [2], we adopt the method that estimates the IF by taking the peak of the signal's TFD for its simplicity of implementation. The WVD is a natural first choice as a TFD for peak filtering given that the other quadratic TFDs are simply smoothed versions of the WVD [3]. The WVD of the analytic signal $z(t)$ is defined (in Section 2.1.4) as:

$$W_z(t,f) = \int_{-\infty}^{\infty} z(t+\tau/2)z^*(t-\tau/2)e^{-j2\pi f\tau}d\tau. \tag{11.4.4}$$

For monocomponent FM signals, the WVD will produce a time-frequency representation of the signal exhibiting significant energy concentration around the signal's IF. When the signal's IF is linear, delta functions will appear at the positions of the IF providing a perfect signal IF estimate. The IF estimate is found by maximizing the WVD over frequency [2]; that is

$$\hat{f}_z(t) = \underset{f}{\text{argmax}}[W_z(t,f)] \tag{11.4.5}$$

The IF estimate based on the peak of the WVD is unbiased and has variance approaching the Cramer-Rao lower bound for signals with linear IF laws in *additive* white zero-mean noise with moderate to high SNR [2]. However, as the order of the polynomial IF increases, the delta functions will be replaced by less peaky functions. The peak of these functions will lie away from the true IF, resulting in IF estimates which are biased. To remedy this, a windowed WVD is used, such that the signal IF is as close to linear as possible across the entire window length.

11.4.2.2 Basic Principle

Time-frequency peak filtering consists of a two step procedure whereby the signal to be filtered is first encoded as the IF of a unit amplitude FM modulated analytic signal. Then, the IF is estimated by taking the peak of a time-frequency distribution (TFD) to recover the filtered signal. This may be summarized as follows:

Step 1. Encode the noisy signal $s(t)$ via FM modulation as: $z_s(t) = e^{j2\pi\mu\int_0^t s(\lambda)d\lambda}$ where μ is a scaling parameter analogous to the FM modulation index.

Step 2. Estimate the peak of the WVD of the analytic signal $z_s(t)$:

$$\hat{x}(t) = \hat{f}_z(t) = \underset{f}{\text{argmax}}[W_{z_s}(t,f)]/\mu$$

11.4.2.3 Properties

The properties of the encoding and IF estimation steps of TFPF are derived for the case of WVD only. The use of other TFDs will lead to slightly different properties.

Property 1: The encoding step converts the additive noise $n(t)$ to multiplicative noise $z_n(t)$ that modulates the signal component $z_x(t)$; that is

$$z(t) = e^{j2\pi\mu\int_0^t s(\lambda)d\lambda} = z_x(t)z_n(t) \tag{11.4.6}$$

where the encoded noise and deterministic signal components are given by

$$z_x(t) = e^{j2\pi\mu\int_0^t x(\lambda)d\lambda} \qquad \text{and} \qquad z_n(t) = e^{j2\pi\mu\int_0^t n(\lambda)d\lambda} \tag{11.4.7}$$

Proof: Equation 11.4.6 is obtained by a direct substitution of 11.4.1 into 11.4.3.

Property 2: The Wigner-Ville Spectrum of the signal $z_s(t)$ is given by :

$$\text{WVS}_{z_s}(t,f) = E[W_{z_s}(t,f)] = \text{WVS}_{z_n}(t,f) \underset{f}{*} W_{z_x}(t,f) \tag{11.4.8}$$

where $E[\cdot]$ is the expectation operator, $\text{WVS}_{z_n}(t,f) = E[W_{z_n}(t,f)]$, and $\underset{f}{*}$ represents the convolution operation in the frequency domain.

Proof: This property follows from *property 1* and the direct application of the expectation operator to the WVD of the encoded signal $z_s(t)$.

This latter is given by:

$$E[W_{z_s}(t,f)] = \int_{-\infty}^{\infty} R_{z_n}(t,\tau)K_{z_x}(t,\tau)e^{-j2\pi f\tau}d\tau \tag{11.4.9}$$

where the time-dependent autocorrelation function of $z_n(t)$ is

$$R_{z_n}(t,\tau) = E[z_n(t+\tau/2)z_n^*(t-\tau/2)] = E[e^{j2\pi\mu\int_{t-\tau/2}^{t+\tau/2} n(\lambda)d\lambda}] \tag{11.4.10}$$

and the time-dependent bilinear product function of $z_x(t)$ is

$$K_{z_x}(t,\tau) = z_x(t+\tau/2)z_x^*(t-\tau/2) = e^{j2\pi\mu\int_{t-\tau/2}^{t+\tau/2} x(\lambda)d\lambda} \tag{11.4.11}$$

Equation (11.4.8) is then obtained by using the fact that the Fourier transform of a product in time is equivalent to the convolution in frequency.

Equation (11.4.8) shows that additive noise smears the encoded signal WVD, $W_{z_x}(t,f)$, through convolution. Therefore, the bias of TFPF is dependent on the encoded signal $z_x(t)$ as well as the shape of the encoded noise spectrum $\mathrm{WVS}_{z_n}(t,f)$. By restricting the shape of the encoded noise spectrum, a class of noise can be defined which does not introduce bias to the IF estimation. An example from this class is the WGN as will be seen next.

Property 3: The time dependent autocorrelation function of the encoded noise, $R_{z_n}(t,\tau;\mu)$, is equal to the characteristic function of $q(t,\tau) = 2\pi \int_{t-\tau/2}^{t+\tau/2} n(\lambda)d\lambda$; that is

$$R_{z_n}(t,\tau,\mu) = E[e^{j2\mu q(t,\tau)}] = \Phi_q(t,\tau,\mu) \tag{11.4.12}$$

where $\Phi_q(t,\tau,\mu)$ is the characteristic function of $q(t,\tau)$ defined by [4]:

$$\Phi_q(t,\tau,\mu) = E[e^{j\mu q(t,\tau)}] = \exp\left(\sum_{i=1}^{\infty} \frac{k_{qi}(t,\tau)(j\mu)^i}{i!}\right) \tag{11.4.13}$$

and $k_{qi}(t,\tau)$ is the ith cumulant of $q(t,\tau)$.

Proof: Equation 11.4.12 is obtained by forming the autocorrelation function of $z_n(t)$ in equation 11.4.7 and using the above definition of $q(t,\tau)$.

11.4.3 Accurate TFPF

Equation (11.4.8) suggests that in general, a bias in IF estimation is introduced by the time-frequency distribution of $z_x(t)$ (deterministic bias) and/or the noise (stochastic bias). In the case where the encoded signal $s(t)$ is composed of a deterministic signal $x(t)$ that is linear in time and embedded in stationary WGN $n(t)$, TFPF gives an unbiased estimate of the signal $x(t)$.

Proof: Consider the signal $s(t)$, given in (11.4.1), to be filtered using TFPF. For the case where $n(t)$ is stationary WGN, the ith cumulant of $n(t)$ is such that $k_{ni} = 0$ for $i \geq 3$ and $q(t,\tau)$ is Gaussian with $k_{qi} = 0$ for $i \geq 3$. Furthermore, if the noise is a zero-mean independent process, i.e. $R_n(\tau) = k_{n2}\delta(\tau)$, then [5, page 369]

$$k_{q1}(t,\tau) = 0 \qquad \text{and} \qquad k_{q2}(t,\tau) = 4\pi^2|\tau|k_{n2} \tag{11.4.14}$$

The characteristic function given in (11.4.13) becomes

$$\Phi_q(t,\tau,\mu) = e^{-2\pi^2\mu^2|\tau|k_{n2}} \tag{11.4.15}$$

Taking the Fourier transform of this expression gives

$$E[W_n(f,t)] = \frac{4\pi^2 k_{n2}\mu^2}{(2\pi^2 k_{n2}\mu^2)^2 + (2\pi f)^2} \tag{11.4.16}$$

This shows that the encoded noise spectrum is low-pass with a maximum at the frequency $0Hz$. Hence, WGN will not introduce any bias to the estimate of the IF. By replacing this last expression in (11.4.8) we obtain:

$$W_{z_s}(t,f) = W_{z_x}(t,f) *_f \frac{4\pi^2 k_{n2}\mu^2}{(2\pi^2 k_{n2}\mu^2)^2 + (2\pi f)^2} \tag{11.4.17}$$

This expression shows that the bias in the IF using the peak of $W_{z_s}(f,t)$ could only come from $W_{z_x}(t,f)$. For the case where the signal $x(t)$ is linear in time; that is $x(t) = \alpha t + C$, where α and C are constants, equation 11.4.17 becomes

$$W_{z_s}(t,f) = \delta(f - x(t)) *_f \frac{4\pi^2 k_{n2}\mu^2}{(2\pi^2 k_{n2}\mu^2)^2 + (2\pi f)^2} \tag{11.4.18}$$

$$= \frac{4\pi^2 k_{n2}\mu^2}{(2\pi^2 k_{n2}\mu^2)^2 + (2\pi f - 2\pi x(t))^2} \tag{11.4.19}$$

The delta function ensures that the peak of this function occurs at $x(t)$. Therefore if the signal $x(t)$ is linear in time and embedded in stationary WGN, TFPF gives an unbiased estimate.

Equation 11.4.17 shows that in the general case where the signal $x(t)$ is a non-linear function of time, the WVD-based TFPF is biased, requiring an appropriate windowing of the data. The window is chosen such that the signal within this window behaves almost linearly [6] (see Section 11.4.4). In the special case where the signal $x(t)$ is a finite-order polynomial in time, the deterministic bias can be completely eliminated if the WVD is replaced by the polynomial WVD (PWVD) of an appropriate order since the PWVD exhibits delta functions along the IF law for polynomial FM signals [7].

11.4.4 Discrete-Time Algorithm for TFPF

The implementation of TFPF using the windowed WVD requires both signal scaling before encoding to prevent aliasing, and the selection of the window length for reduced bias. These two aspects are discussed next.

11.4.4.1 Signal Scaling

FM modulation of un-scaled discrete time signals can lead to aliasing which produces discontinuities in the estimated IF at the frequency boundaries of the time-frequency plane. This is avoided by amplitude scaling of the noisy signal before frequency encoding. Without loss of generality and unless otherwise specified we assume that the signal $s(t)$ is sampled at a normalized sampling frequency of $1Hz$.

The scaled signal, $s_c(m)$, is obtained by using the following transformation.

$$s_c(m) = S[s(m)] = (a-b)\frac{s(m) - \min[s(m)]}{\max[s(m)] - \min[s(m)]} + b \tag{11.4.20}$$

where $S[.]$ is the scaling operator and the parameters a and b, which satisfy the constraint $.5 \geq a = \max[s_c(m)] > b = \min[s_c(m)] \geq 0$, are chosen to provide suitable frequency limits on the encoded signal. The operators max[.] and min[.] are the maximizing and minimizing functions respectively. The estimate of the desired signal, $\hat{x}(m)$, is recovered by an inverse scaling operation; that is

$$\hat{x}(m) = S^{-1}[\hat{x}_c(m)] = \frac{(\hat{x}_c(m) - b)(\max[s(m)] - \min[s(m)])}{a - b} + \min[s(m)] \quad (11.4.21)$$

where $\hat{x}_c(m)$ is the scaled signal obtained using TFPF on $s_c(m)$.

11.4.4.2 Reduced-Bias Window Length Selection

The bias-variance tradeoff is a key in the practical implementation of TFPF with the windowed WVD. Bias reduction requires a small window length to minimize the non-optimal nature of the WVD for higher than quadratic phase signals. On the other hand, variance reduction is achieved by increasing the window length to provide the local estimate with more information. To reduce the variance of the estimate while maintaining bias performance it becomes necessary to increase the sampling rate. Thus there is a tradeoff between estimator bias and sampling frequency which results in a relationship between window length and bias, for a given sampling frequency. Results relating to TFPF window lengths are derived in [6]. The basic results for window length τ_w, sampling frequency f_s and maximum value of IF, f_p are given below. For the case of signal estimation:

$$\tau_w \leq \frac{0.634 f_s}{\pi f_p} \quad (11.4.22)$$

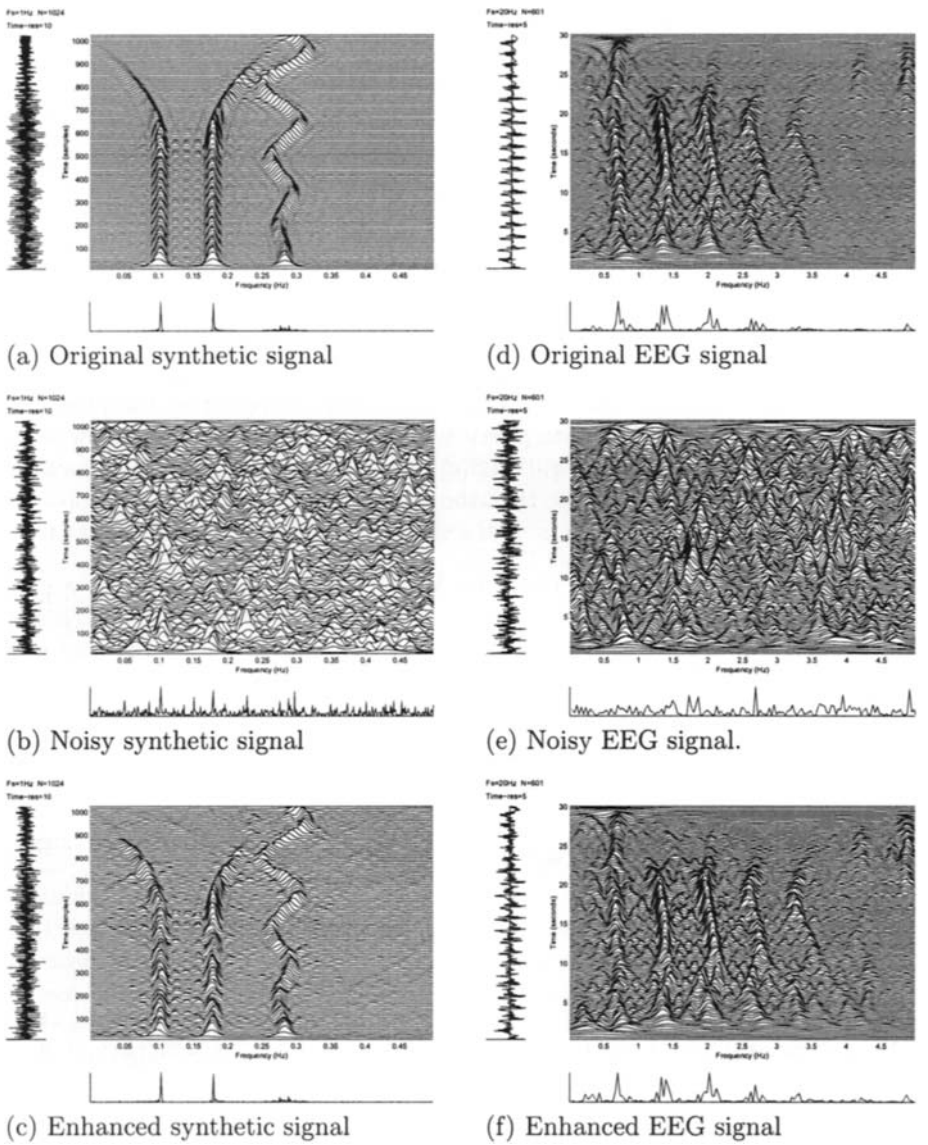
This equation gives the maximum window length as a function of maximum signal IF and sampling frequency. In a typical application a specified window length is required for a given SNR. The signal is sampled at a high enough rate to satisfy (11.4.22).

11.4.4.3 The Iterative TFPF Algorithm

In the first application, TFPF may not remove as much additive noise as desired. If this situation occurs, reapplication of the procedure to the filtered signal is recommended. This leads to a 3 step iterative process:

1. Scale and encode noisy signal.
2. Apply TFPF to yield a signal estimate $\hat{x}(t)$.
3. If $\hat{x}(t)$ contains substantial noise, go to step 1; else terminate the process.

Simulations demonstrate the convergence of the repeated scheme to a stable estimate $\hat{x}(t)$ [1].



(a) Original synthetic signal

(d) Original EEG signal

(b) Noisy synthetic signal

(e) Noisy EEG signal.

(c) Enhanced synthetic signal

(f) Enhanced EEG signal

Fig. 11.4.1: B-distributions of a synthetic signal (left column) and an EEG signal (right column), showing the original signal (top row), and the noisy signal before enhancement (middle row) and after enhancement (bottom row).

11.4.5 Examples and Results

Example 1 (A Multicomponent signal in WGN) Let us consider a multicomponent signal, $x(m)$, expressed by:

$$x(m) = \begin{cases} 0.85\sin(0.055m + 3.75 \times 10^{-4}\sin(0.000625m)m) + \sin(0.035m) \\ +\sin(0.020m), & 0 \le m < L/2 \\ 0.85\sin(0.055m + 3.75 \times 10^{-4}\sin(0.000625m)m) \\ +2(1 - m/L)\sin(0.035m + 2.75 \times 10^{-11}(m - L/2)^3) \\ +2(1 - m/L)\sin(0.020m - 2.75 \times 10^{-11}(m - L/2)^3), & L/2 \le m < L \end{cases}$$

where the data length L is taken as 32768 data points. For a time-frequency illustration of this signal, the B-distribution (BD) with smoothing parameter $\beta = 0.01$ is shown in Fig. 11.4.5(a).[1] White Gaussian noise was added to the above signal giving an SNR of -9 dB; the BD of the noisy signal is shown in Fig. 11.4.5(b). The windowed WVD peak filter was then implemented to recover $x(m)$ from the noisy signal. A window length of 15 data points was chosen to satisfy the window length constraints given in (11.4.22). Fig. 11.4.5(c) shows the clean recovery of the signal after three TFPF iterations. Note that the WVD is used as the vehicle for signal recovery while the B-distribution is used only for presentation of the results.

Example 2 (Newborn EEG data in WGN): Fig. 11.4.5(d) shows a time-frequency representation of a real newborn EEG signal using the B-distribution with $\beta = 0.01$. WGN is then added to the signal at $SNR = -9$ dB. The noisy signal in Fig. 11.4.5(e) shows that the time-frequency patterns of the EEG signal are not clearly visible. Using a window length of 20 data points, four iterations of the TFPF were used to recover a cleaner signal. The filtered signal in Fig. 11.4.5(f) demonstrates the efficiency of TFPF.

11.4.6 Summary and Conclusions

TFPF is a tool for signal enhancement, applicable to a large class of signals if the windowed WVD TFPF is used for reduced bias. This class includes those signals which may be represented as a sum of band-limited non-stationary processes in additive WGN. Testing on simulated and real data indicates that the method significantly enhances signals of this class by filtering out most of the additive noise. Further details of the time-frequency peak filtering method are provided in [8].

References

[1] M. J. Arnold, M. Roessgen, and B. Boashash, "Filtering real signals through frequency modulation and peak detection in the time-frequency plane," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, vol. 3, pp. 345–348, Adelaide, Australia, 19–22 April 1994.

[1] See p. 53 for a definition of the BD and its parameter.

[2] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals; Part 2: Algorithms and applications," *Proc. IEEE*, vol. 80, pp. 519–568, April 1992. With introductions by Michael J. Riezenman.

[3] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Analysis and Array Processing* (S. Haykin, ed.), vol. 1, ch. 9, pp. 418–517, Englewood Cliffs, NJ: Prentice-Hall, 1991.

[4] A. Papoulis, "Random modulation: A review," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 31, pp. 96–105, February 1983.

[5] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill, 3rd ed., 1991.

[6] M. Mesbah and B. Boashash, "Reduced bias time-frequency peak filtering," in *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, vol. 1, pp. 327–330, Kuala Lumpur, 13–16 August 2001.

[7] B. Boashash and P. J. O'Shea, "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 216–220, January 1994.

[8] B. Boashash and M. Mesbah, "Signal enhancement by time-frequency peak filtering," *IEEE Trans. Signal Processing*, vol. 51, November 2003.

This Page Intentionally Left Blank

Chapter 12

Detection, Classification and Estimation

Methods for the detection, estimation and classification of non-stationary signals can be enhanced by utilizing the time-frequency characteristics of such signals. Such time-frequency formulations are described in this chapter. The topic is covered in four articles with appropriate internal cross-referencing to this and other chapters.

The structure of TF methods is suitable for designing and implementing optimal detectors. Several approaches exist, such as decomposition of TFDs in sets of spectrograms (Article 12.1). For both analysis and classification, a successful TF methodology requires matching of TFDs with the structure of a signal. This can be achieved by a matching pursuit algorithm using time-frequency atoms adapted to the analyzed signals (12.2). We can perform system identification by exciting linear systems with a linear FM signal and relating TFDs of the input and output using time-frequency filtering techniques (12.3). Methods for time-frequency signal estimation and detection can be carried out using time-varying Wiener filters (12.4).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

12.1 OPTIMAL TIME-FREQUENCY DETECTORS [0]

12.1.1 Time-Frequency Detection

The most predominant application of time-frequency representations (TFRs) is *exploratory signal analysis*: using TFRs to expose time-varying spectral characteristics of signals. TFRs are widely used in this mode in applications ranging from speech analysis to biomedical signal analysis to geophysical exploration. However, the structure of TFRs can also be fruitfully exploited for designing and implementing optimal processors for detecting a variety of nonstationary signals in noisy environments. Perhaps the simplest and best-known examples are the short-time Fourier transform (STFT) and the cross-ambiguity function in radar/sonar processing [1].

The goal of this article is to highlight the key ideas behind the role of TFRs in optimal detection. Essentially, TFRs are useful for detecting nonstationary stochastic signals that exhibit certain unknown parameters, called *nuisance* parameters, such as unknown time or frequency shifts. The statistical characteristics of the signal are matched via proper choice of the TFR kernel and the nuisance parameters are dealt with by exploiting the *covariance* of TFRs to time and frequency shifts. The structure of optimal TFR detectors also suggests a design approach for suboptimal, low-complexity TFR detectors that can potentially deliver competitive performance in scenarios in which optimal detectors are prohibitively complex. The notion of time-frequency detectors also naturally generalizes to a certain class of quadratic signal representations that extend the concept of time and frequency.

The next section reviews relevant TFR concepts. Section 12.1.3 presents a class of detection scenarios for which TFRs are optimal and provides a characterization of the TFR detectors. Section 12.1.4 briefly discusses some extensions of the detection framework to generalized TFRs as well as suboptimal detectors. Finally, Section 12.1.5 provides some concluding and bibliographic remarks.

12.1.2 Time-Frequency Representations

The most widely used *linear* TFR is the STFT that correlates the signal of interest, $r(t)$, with TF shifted copies of a prototype (window) signal $g(t)$

$$F_r^g(t,f) = \int r(\tau) g^*(\tau - t) e^{-j2\pi f\tau} d\tau. \tag{12.1.1}$$

One of the most important *quadratic* TFRs is the real-valued Wigner distribution (WD) that essentially correlates the signal with time-frequency shifted copies of a time-reversed version of itself:

$$W_r(t,f) = \int r(t+\tau/2) r^*(t-\tau/2) e^{-j2\pi f\tau} d\tau. \tag{12.1.2}$$

[0]Author: **Akbar M. Sayeed**, Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706, USA (akbar@engr.wisc.edu). Reviewers: P. Flandrin and A. Papandreou-Suppappola.

A whole class of quadratic TFRs, sometimes called Cohen's class [2], can be generated from the WD via time-frequency filtering with a two-dimensional kernel

$$\rho_r(t,f;\gamma) = \int\int W_r(u,v)\gamma(u-t,v-f)dudv \tag{12.1.3}$$

where the entire class is characterized by different choices of the kernel $\gamma(t,f)$. A characterizing property of TFRs in Cohen's class is that they are *covariant* to time-frequency shifts in the signal:

$$r(t) \longrightarrow r(t-t_o)e^{j2\pi f_o t} \Longrightarrow \rho_r(t,f) \longrightarrow \rho_r(t-t_o, f-f_o). \tag{12.1.4}$$

The magnitude squared STFT, often called the spectrogram, is a member of Cohen's class with the kernel given by the WD of the window function: $|F_r^g(t,f)|^2 = \rho_r(t,f;\gamma = W_g)$.

An important relation in the context of detection is *Weyl correspondence* that relates a quadratic form to the WD

$$\langle \boldsymbol{L}x, x\rangle = \int\int x^*(t_1)L(t_1,t_2)x(t_2)dt_1dt_2 = \int\int W_x(t,f)\mathrm{WS}_L(t,f)dtdf \tag{12.1.5}$$

where $\boldsymbol{L}$ a linear operator defined by the kernel $L(t_1,t_2)$ and $\mathrm{WS}_L(t,f)$ denotes the Weyl symbol of the operator defined as [3]

$$\mathrm{WS}_L(t,f) = \int L(t+\tau/2, t-\tau/2)e^{-j2\pi f\tau}d\tau. \tag{12.1.6}$$

If $L(t_1,t_2) = R_s(t_1,t_2) = \mathcal{E}[s(t_1)s^*(t_2)]$, the correlation function of a random signal $s(t)$, then the Weyl symbol of $R_s(t_1,t_2)$ is also the Wigner-Ville spectrum of $s(t)$ defined as [2]

$$\mathrm{WVS}_s(t,f) = \mathcal{E}[W_s(t,f)] = \int R_s(t+\tau/2, t-\tau/2)e^{-j2\pi ft}d\tau = \mathrm{WS}_{R_s}(t,f). \tag{12.1.7}$$

The Wigner-Ville spectrum is a generalization of the notion of power spectral density to nonstationary processes and is an equivalent characterization of the second-order statistics of $s(t)$.

The kernel $\gamma(t,f)$ can be interpreted as the WS of an operator $\boldsymbol{L}_\gamma$ by inverting (12.1.6)

$$L_\gamma(t_1,t_2) = \int \gamma\left(\frac{t_1+t_2}{2}, f\right)e^{j2\pi f(t_1-t_2)}df. \tag{12.1.8}$$

If $\gamma(t,f)$ generates a real-valued TFR then $\boldsymbol{L}_\gamma$ is a Hermitian operator and under certain conditions (e.g, if it is a Hilbert-Schmidt operator) it admits an eigendecomposition $L_\gamma(t_1,t_2) = \sum_k \lambda_k u_k(t_1)u_k^*(t_2)$ where $\{\lambda_k\}$ are the real-valued eigenvalues and $\{u_k(t)\}$ are the corresponding orthonormal eigenfunctions. Using this eigendecomposition, the kernel can be expressed as a weighted sum of

WDs, $\gamma(t,f) = \sum_k \lambda_k W_{u_k}(t,f)$, and the corresponding TFR can be expressed as a weighted sum of spectrograms

$$\begin{aligned}\rho_r(t,f;\gamma) &= \sum_k \lambda_k \int\int W_r(u,v) W_{u_k}(u-t, v-f) du dv \\ &= \sum_k \lambda_k |F_r^{u_k}(t,f)|^2. \end{aligned} \tag{12.1.9}$$

This decomposition in terms of spectrograms (discussed in Article 6.4) provides an intimate connection between linear and quadratic TFRs and is very useful for efficient implementation of TFR-based detectors.

12.1.3 Time-Frequency Detection Framework

We motivate TFR-based detection by considering the classical binary hypothesis testing problem encountered in radar. Let $v(t)$ denote the complex baseband transmitted signal of duration T_o and bandwidth B_o. The received signal $r(t)$ is processed over an observation interval $[T_i, T_f]$ to detect the presence of a target. The two hypothesis on $r(t)$ are

$$\begin{aligned} H_0 &: r(t) = \epsilon(t), \quad t \in [T_i, T_f] \\ H_1 &: r(t) = a s(t; t_o, f_o) + \epsilon(t), \quad t \in [T_i, T_f] \ . \end{aligned} \tag{12.1.10}$$

Under the null hypothesis, H_0, the signal contains only the zero-mean complex additive white Gaussian noise (AWGN), $\epsilon(t)$, of variance σ_ϵ^2. Under the active hypothesis, H_1, a TF shifted version of the transmitted signal, $s(t;t_o,f_o) = v(t-t_o)e^{j2\pi f_o t}$, is received in the presence of noise. In (12.1.10), a denotes an unknown complex gain parameter. The unknown time and frequency parameters, (t_o, f_o), represent the delay and Doppler shift of the received signal relative to the transmitted signal and correspond to the unknown distance and velocity of the target. Let $\mathcal{S} = [T_{min}, T_{max}] \times [-F_{max}, F_{max}]$ denote the possible range of values for the nuisance parameters (t_o, f_o). The required observation interval is $[T_i, T_f] = [T_{min}, T_o + T_{max}]$ in this case.

For any *given* value of $(t_o, f_o) \in \mathcal{S}$, the optimal decision statistic, derived from the likelihood ratio [1,4], is the noncoherent matched filter statistic

$$D^{(t_o,f_o)} = |\langle r, s(t_o,f_o)\rangle|^2 = \left| \int_{T_i}^{T_f} r(t) v^*(t-t_o) e^{-j2\pi f_o t} dt \right|^2 = |F_r^v(t_o,f_o)|^2, \tag{12.1.11}$$

which correlates the received signal with the reference signal $s(t;t_o,f_o)$ and can be computed via a spectrogram with $v(t)$ as the window. The unknown nature of the nuisance parameters is usually accounted for via a generalized likelihood ratio test (GLRT) corresponding to the decision statistic

$$D(r) = \max_{(t,f)\in\mathcal{S}} D^{(t,f)}(r) = \max_{(t,f)\in\mathcal{S}} |F_r^v(t,f)|^2 \tag{12.1.12}$$

which represents the peak of the spectrogram over $\mathcal{S}$. The statistic $D(r)$ is compared to a threshold to declare the presence (H_1) or absence (H_0) of a target. Thus, the optimal detector correlates the received signal with all possible time-frequency shifted versions of the transmitted signal and picks the largest correlation as the decision statistic. The location of the peak is in fact a maximum likelihood (ML) estimate of the unknown parameters: $(\hat{t}_o, \hat{f}_o)_{ML} = \arg\max_{(t,f)\in\mathcal{S}} D^{(t,f)}$. The GLRT detector (12.1.12) is the statistic (12.1.11) computed at $(\hat{t}_o, \hat{f}_o)_{ML}$. The detector performance is governed by the shape of the magnitude squared of the auto-ambiguity function of the transmitted signal

$$A_v(\tau,\nu) = \int v(t+\tau/2)v^*(t-\tau/2)e^{-j2\pi\nu t}dt \tag{12.1.13}$$

near the origin [1]. Note that $|A_v(\tau,\nu)|^2 = |F_v^v(\tau,\nu)|^2 = \int\int W_v(u,v)W_v(u-\tau, v-\nu)dudv$. Ideally, $A_v(\tau,\nu)$ should approximate a two-dimensional Dirac delta function for perfect resolution in delay and Doppler. Waveforms with large time-bandwidth products, such as those derived from pseudo-random codes used in spread-spectrum communications [5], can be used to approximate this "thumbtack" function.

The simple detector (12.1.12) is based on a rank-1 quadratic[1] TFR. Higher-rank quadratic TFRs can realize detectors for a much richer class of nonstationary stochastic signals. The next subsection describes this quadratic TFR detection framework. In turn, in many cases *low-rank* approximations of TFR detectors, implemented via a bank of a *few* spectrograms, suffice.

12.1.3.1 Quadratic Time-Frequency Detectors

Quadratic TFRs can be exploited for detection scenarios of the form (12.1.10) where the signal $s(t;t_o,f_o)$ is a time-frequency shifted version of some underlying *stochastic* signal whose Wigner-Ville spectrum is relatively localized in time and frequency. This represents a fairly rich class of detection scenarios for modeling the inherent randomness in the signal as well as measurement uncertainties that manifest themselves as unknown time-frequency shifts. For example, radar applications in which the transmitted signal $v(t)$ encounters a randomly time-varying scatterer or channel may be modeled this way [1]. Similarly, signals involved in machine fault diagnostics may exhibit random characteristics along with timing jitters and frequency offsets due to physical mechanisms. The essential idea behind quadratic TFR detectors is to capture the signal statistics via the kernel $\gamma(t,f)$ and the nuisance parameters via the time-frequency covariance property (12.1.4).

Suppose that the signal $s(t;t_o,f_o)$ in (12.1.10) is a time-frequency shifted version of a zero-mean Gaussian signal $w(t)$; that is, $s(t;t_o,f_o) = w(t-t_o)e^{j2\pi f_o t}$ and $R_{s(t_o,f_o)}(t_1,t_2) = R_w(t_1-t_o,t_2-t_o)e^{j2\pi f_o(t_1-t_2)}$. In this case, for any *given* (t_o,f_o),

[1]The rank of a TFR is the number of significant nonzero eigenvalues of $\gamma(t,f)$ in (12.1.9).

the optimum decision statistic is [4,6,7]

$$D^{(t_o,f_o)}(r) = \left\langle \widehat{\boldsymbol{R}}_{s(t_o,f_o)} r, r \right\rangle = \left\langle W_r, \mathrm{WS}_{\widehat{R}_{s(t_o,f_o)}} \right\rangle, \tag{12.1.14}$$

where $\widehat{\boldsymbol{R}}_{s(t_o,f_o)} = (\boldsymbol{R}_{s(t_o,f_o)} + \sigma_\epsilon^2 \boldsymbol{I})^{-1} \boldsymbol{R}_{s(t_o,f_o)}$ is an operator defined by the signal correlation function and the second equality follows from (12.1.5). The simpler choice $\widehat{\boldsymbol{R}}_{s(t_o,f_o)} = \boldsymbol{R}_{s(t_o,f_o)}$ in (12.1.14) yields the *locally optimum*[2] detector [4].

The unknown nuisance parameters are again accounted for by exploiting the covariance property (12.1.4) which implies that $\mathrm{WS}_{R_{s(t_o,f_o)}}(t,f) = \mathrm{WS}_{R_w}(t-t_o, f-f_o)$ and $\mathrm{WS}_{\widehat{R}_{s(t_o,f_o)}}(t,f) = \mathrm{WS}_{\widehat{R}_w}(t-t_o, f-f_o)$, where $\widehat{\boldsymbol{R}}_w = (\boldsymbol{R}_w + \sigma_\epsilon^2 \boldsymbol{I})^{-1} \boldsymbol{R}_w$ [7].[3] Consequently, the quadratic decision statistic in (12.1.14) can be expressed as

$$\begin{aligned} D^{(t_o,f_o)} &= \int\int W_r(u,v) \mathrm{WS}_{\widehat{R}_w}(u-t_o, v-f_o) du dv \\ &= \rho_r\left(t_o, f_o; \gamma = \mathrm{WS}_{\widehat{R}_w}\right). \end{aligned} \tag{12.1.15}$$

Similarly, the locally optimal detector corresponds to $\widehat{R}_w = R_w$ in (12.1.15). Thus, the decision statistic corresponding to different values of (t_o, f_o) can be computed via a TFR with kernel $\mathrm{WS}_{\widehat{R}_w}$. The overall GLRT detector for unknown (t_o, f_o) is given by

$$\begin{aligned} D(r) &= \max_{(t,f)\in\mathcal{S}} \rho_r(t,f; \mathrm{WS}_{\widehat{R}_w}) \\ &= \max_{(t,f)\in\mathcal{S}} \int\int W_r(u,v) \mathrm{WS}_{\widehat{R}_w}(u-t, v-f) du dv \end{aligned} \tag{12.1.16}$$

where $\mathcal{S} = [T_{min}, T_{max}] \times [-F_{max}, F_{max}]$. If $\mathrm{WVS}_w(t,f) = \mathcal{E}[W_w(t,f)]$ is supported on $(t,f) \in [0,T_w] \times [-B_w, B_w]$, the required observation interval is $[T_i, T_f] = [T_{min}, T_{max} + T_w]$ and $W_r(t,f)$ in (12.1.16) is computed over the range $[T_{min}, T_{max} + T_w] \times [-B_w - F_{max}, B_w + F_{max}]$.

The detector (12.1.16) has the intuitively appealing interpretation of *nonstationary spectral correlation*: the WD of the observed waveform is correlated with all possible time-frequency shifted versions of $\mathrm{WS}_{\widehat{R}_w}$ which is the WVS of $w(t)$ in the case of the locally optimum detector and the WVS of a modified version of $w(t)$ in the case of the optimal detector. This interpretation of optimum TFR detection is illustrated in Fig. 12.1.1. The location of the maximum in (12.1.16) is the ML estimate of (t_o, f_o). Estimator performance is related to the shape of $F(\tau,\nu) = \int\int \mathrm{WVS}_w(u,v) \mathrm{WVS}_w(u-\tau, v-\nu) du dv$ near the origin.

---

[2]The first term in the expansion of the likelihood ratio as a function of signal-to-noise ratio (SNR) [4]. This decision statistic is useful in low-SNR scenarios.

[3]We note that both $\widehat{\boldsymbol{R}}_w$ and $\boldsymbol{R}_w$ have the same eigenfunctions; the eigenvalues of $\widehat{\boldsymbol{R}}_w$ are obtained from those of $\boldsymbol{R}_w$ by adding σ_ϵ^2 to each.

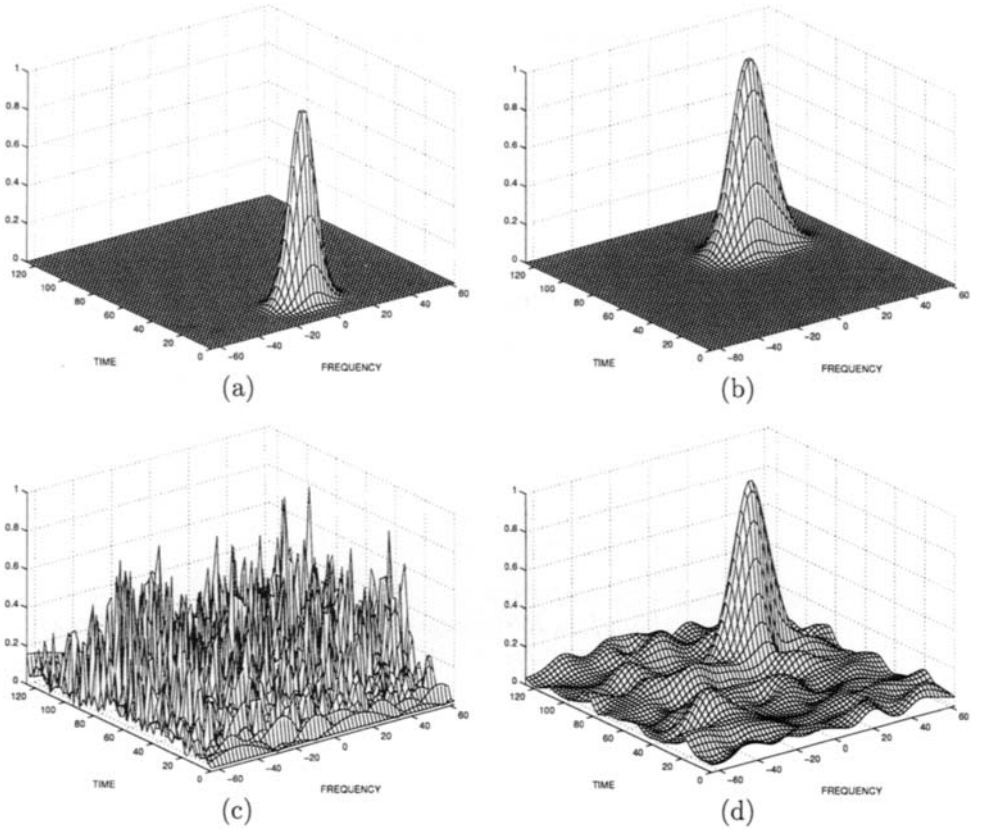


Fig. 12.1.1: Illustration of nonstationary spectral correlation underlying TFR detection. (a) The Wigner-Ville spectrum, $\mathrm{WVS}_w(t,f)$, of the underlying stochastic signal $w(t)$. (b) The *idealized* TFR, $\rho_r(t,f;\gamma = \mathrm{WVS}_w)$, in the absence of noise; that is, $r(t) = s(t;t_o,f_o)$. (c) The WD of the noisy observed signal $r(t)$. (d) The TFR, $\rho_r(t,f;\gamma = \mathrm{WVS}_w)$, of the noisy signal. The TFR in (d) is computed by smoothing the WD in (c) with the WVS in (a) and consists of the idealized TFR in (b) and additive noisy components. The peak of the TFR in (d) represents the decision statistic that is compared to a threshold to make a decision. The location of the peak represents ML estimates of the unknown delay and Doppler shift of the target.

The correlation function $\widehat{R}_w(t_1,t_2)$ is related to $R_w(t_1,t_2)$ by

$$\widehat{R}_w(t_1,t_2) = \sum_k \frac{\lambda_k}{\lambda_k + \sigma_\epsilon^2} u_k(t_1) u_k^*(t_2)$$

where $\{\lambda_k\}$ and $\{u_k\}$ denote eigenvalues and eigenfunctions, respectively, of $R_w(t_1,t_2)$. Analogous to (12.1.9), the decision statistic (12.1.15) can be expressed as a weighted sum of spectrograms

$$D^{(t_o,f_o)}(r) = \sum_k \widehat{\lambda}_k \left| F_r^{u_k}(t_o,f_o) \right|^2 \tag{12.1.17}$$

where $\widehat{\lambda}_k = \frac{\lambda_k}{\lambda_k+\sigma_\epsilon^2}$ in the case of optimum detection and $\widehat{\lambda}_k = \lambda_k$ in the case of locally optimum detection. The importance of the above decomposition stems from the fact that many signals encountered in practice exhibit *few* dominant eigenvalues. This implies that the quadratic TFR detector (12.1.16) can be efficiently implemented via a bank of a few spectrograms.

12.1.3.2 An Illustrative Example

We now illustrate the TFR detection framework in the context of an important application in which the transmitted signal $v(t)$ travels over a *randomly time-varying* channel. Such a channel model is appropriate in many applications, including radar [1], wireless communications [5], and underwater acoustical communication [1]. In this case, the stochastic signal $s(t;t_o,f_o)$ is related to $v(t)$ as

$$\begin{aligned} s(t;t_o,f_o) &= \int_0^{T_m} h(t,\tau)v(t-\tau-t_o)e^{j2\pi f_o t}d\tau \\ &= \int_0^{T_m}\int_{-B_d}^{B_d} H(\nu,\tau)v(t-\tau-t_o)e^{j2\pi(\nu+f_o)t}d\nu d\tau \end{aligned} \quad (12.1.18)$$

where $h(t,\tau)$ is the time-varying channel impulse response and the equivalent representation, $H(\nu,\tau) = \int h(t,\tau)e^{-j2\pi\nu t}dt$, is called the *delay-Doppler spreading function* [5]. In (12.1.18), the received signal is modeled as a linear combination (with stochastic weights) of a continuum of time-frequency shifted copies of $v(t)$. The parameters T_m and B_d are the *multipath* and *Doppler* spreads of the channel, respectively. The parameters (t_o,f_o) represent the *global* delay and Doppler shift encountered during propagation, and T_m and B_d represent the *local* time-frequency spreading around (t_o,f_o) produced by scattering. Each $(\nu,\tau) \in [-B_d,B_d]\times[0,T_m]$ represents a particular infinitesimal scatterer with stochastic gain $H(\nu,\tau)$. Under the wide-sense stationary uncorrelated scattering (WSSUS) model, $\{H(\nu,\tau)\}$ are modeled as a collection of uncorrelated Gaussian random variables [5].

The underlying Gaussian signal $w(t)$ in this case is characterized by

$$R_w(t_1,t_2) = \int_0^{T_m}\int_{-B_d}^{B_d} M(\nu,\tau)v(t_1-\tau)v^*(t_2-\tau)e^{j2\pi\nu(t_1-t_2)}d\nu d\tau. \quad (12.1.19)$$

where $M(\nu,\tau) = \mathcal{E}[|H(\nu,\tau)|^2]$ is the channel *scattering function* which quantifies the relative power contributed by each scatterer. Note that $\mathrm{WVS}_w(t,f)$ is supported on $[0,T_o+T_m]\times[-B_o-B_d,B_o+B_d]$. We focus on the locally optimum detector for simplicity which uses $\widehat{R}_w = R_w$ in (12.1.16). Using the definition of the Weyl symbol in (12.1.6), it can be shown that the detection statistic (12.1.15) takes the

form

$$
\begin{aligned}
D^{(t_o,f_o)}(r) &= \rho_r(t_o, f_o; \gamma = \mathrm{WVS}_w) \\
&= \int\int W_r(u,v)\mathrm{WVS}_w(u-t_o, v-f_o)dudv \qquad (12.1.20) \\
&= \int_0^{T_m}\int_{-B_d}^{B_d} M(\nu,\tau)\left[\int\int W_r(u,v)W_v(u-t_o-\tau, v-f_o-\nu)dudv\right] d\nu d\tau \\
&= \int_{-0}^{T_m}\int_{-B_d}^{B_d} M(\nu,\tau)\left|F_r^v(t_o+\tau, f_o+\nu)\right|^2 d\nu d\tau. \qquad (12.1.21)
\end{aligned}
$$

Thus, the test for any *given* (t_o, f_o) can be computed by smoothing $W_r(t,f)$ in the neighborhood of (t_o, f_o) with $\mathrm{WVS}_w(t,f)$ or by smoothing $|F_r^v(t,f)|^2$ in the neighborhood of (t_o, f_o) with $M(\nu,\tau)$. The overall GLRT detector (12.1.16) can be realized by: (1) computing the $W_r(t,f)$ for $(t,f) \in [T_{min}, T_{max}+T_o+T_m]\times[-F_{max}-B_o-B_d, F_{max}+B_o+B_d]$, or by computing $|F_r^v(t,f)|^2$ for $(t,f) \in [T_{min}, T_{max}+T_m]\times[-F_{max}-B_d, F_{max}+B_d]$, (2) computing $D^{(t,f)}$ over $\mathcal{S}$ by smoothing $W_r(t,f)$ with $\mathrm{WVS}_w(t,f)$ as in (12.1.20) or by smoothing the spectrogram with $M(\nu,\tau)$ as in (12.1.21), and (3) comparing the maximum of $\{D^{(t,f)} : (t,f) \in \mathcal{S}\}$ to a threshold.

While the detector (12.1.21) is in a fairly simple form, the low-complexity implementation in terms of the eigenfunctions of $R_w(t_1,t_2)$ takes an equally simple form. Due to the finite duration and bandwidth of $v(t)$, the decision statistic (12.1.21) admits the following approximate sampled representation

$$
D^{(t_o,f_o)}(r) \approx \sum_{l=0}^{\lceil T_m B_o\rceil}\sum_{m=-\lceil B_d T_o\rceil}^{\lceil B_d T_o\rceil} \widehat{M}(m/T_o, l/B_o)\left|F_r^v(t_o+l/B_o, f_o+m/T_o)\right|^2 \quad (12.1.22)
$$

where $\widehat{M}(\nu,\tau)$ represents a smooth version of $M(\nu,\tau)$ and the number of terms in the summation represents the number of delays and Doppler shifts that are *resolvable* at the receiver, which is also the number of dominant eigenvalues of $R_w(t_1,t_2)$. Note that the number of dominant eigenvalues is approximately $2T_mB_dT_oB_o$ which is typically relatively small since most practical channels are underspread ($2T_mB_d \ll 1$) [5].

12.1.4 Extensions

We now briefly discuss two extensions of the quadratic TFR detection framework described above. First, if the nuisance parameters are modeled as *random* with known probability density function, $p(t,f)$, the locally[4] optimum Bayesian test

[4]The optimum Bayesian detector essentially corresponds to replacing $\rho_r(t,f;\mathrm{WS}_{R_w})$ with $e^{-\rho_r(t,f;\mathrm{WS}_{\hat{R}_w})}$ in (12.1.23).

statistic can be realized in the time-frequency domain as

$$D_{\text{Bayes}}(r) = \int_{\mathcal{S}} \rho_r(t,f;\text{WVS}_w)p(t,f)dtdf$$
$$= \int\int W_r(t,f)\widehat{\text{WVS}}_w(t,f)dtdf. \qquad (12.1.23)$$

where $\widehat{\text{WVS}}_w(t,f) = \iint p(t,f)\text{WVS}_w(u-t,v-f)dudv$. The Bayesian detector computes a test statistic based on our prior expectation of the likelihood of different values of the nuisance parameters. The time-frequency formulation of a quadratic detector in (12.1.14) has an interesting interpretation in view of the form of the Bayesian detector (12.1.23) and the *quasi-energy density* interpretation of the WD: the quadratic detector (12.1.14) weights the WD of the noisy received signal with our prior expectation of distribution of the signal energy in the time-frequency plane. This fact can be also be exploited in the design of rank-1 detectors when the underlying reference signal can only be inferred from noisy measurements [8]. In the rank-1 case, $|\langle r,s\rangle|^2 = \langle W_r, W_s\rangle$, and if $W_s(t,f)$ is mostly concentrate along a curve, for example $W_s(t,f) = \delta(f-\nu(t))$, it can be estimated more accurately (as opposed to $s(t)$ directly) due to higher SNR along the curve in the time-frequency plane. A matched filter for the underlying signal can then be inferred via the instantaneous frequency or the group delay defined by the curve. Such weighted averaging of the TFR can also be exploited (in a suboptimal manner) to design detectors that suppress unwanted signals whose TF support does not completely overlap the support of the desired signal.

Another direction of generalization is going beyond TFRs to joint signal representations in terms of variables other than time and frequency — time-scale representations (TSRs), for example [9]. The GLRT detection framework described above is best suited to representations that possess some covariance property analogous to (12.1.4). For example, TSRs are covariant to time shifts and scale changes. In general, such covariance properties are imposed via a parameterized family of unity operators [10]. On the one hand, each family of unity operators defines a class of signal representations and on the other hand the parameters of the unitary operators provide a model for nuisance parameters in GLRT detection scenarios. In essence, each family of unity operators defines a one-to-one correspondence between a particular class of detection scenarios and a particular class of joint signal representations that serve as optimal detectors for such scenarios.

12.1.5 Summary and Conclusions

Time-frequency formulation of optimum quadratic detection, as in (12.1.14), along with a discussion of linear detectors was first developed by Flandrin in [6]; the basic ideas emphasizing the usefulness of a time-frequency approach to detection had been introduced earlier by Kumar and Carroll (1984) and Kay and Boudreaux-Bartels (1985) (see the references in [6]). Some elements of Bayesian detection were also introduced in [6]. Altes also provided a time-frequency formulation of optimum

detection and classification in terms of spectrograms in [11]. The GLRT formulation emphasized in this article, exploiting the degrees of freedom in a TFR (kernel and covariance properties), was developed in [7, 10]. The same general framework has been recently extended to design data-driven detectors and classifiers that infer the optimal processor directly from data [12]. While these formulations emphasize *optimum* detection, a potentially fruitful research direction, which is relatively unexplored, is the (suboptimal) combination of exploratory data analysis with decision theory: exploiting the degrees of freedom in the TFR to isolate essential signal features that can serve as inputs to pattern classifiers [13].

References

[1] H. L. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. III: "Radar-Sonar Signal Processing and Gaussian Signals in Noise". New York: Wiley, 1971. Reprinted Malabar, FL: Krieger, 1992. Reprinted New York: Wiley, 2001.

[2] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[3] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 318–331, February 1994.

[4] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer, 1988.

[5] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 3rd ed., 1995.

[6] P. Flandrin, "A time-frequency formulation of optimum detection," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 36, pp. 1377–1384, September 1988.

[7] A. M. Sayeed and D. J. Jones, "Optimal detection using bilinear time-frequency and time-scale representations," *IEEE Trans. Signal Processing*, vol. 43, pp. 2872–2883, December 1995.

[8] E. Chassande-Mottin and P. Flandrin, "On the time-frequency detection of chirps," *Applied & Computational Harmonic Analysis*, vol. 6, pp. 252–281, March 1999.

[9] O. Rioul and P. Flandrin, "Time-scale energy distributions: A general class extending wavelet transforms," *IEEE Trans. Signal Processing*, vol. 40, pp. 1746–1757, July 1992.

[10] A. M. Sayeed and D. L. Jones, "Optimum quadratic detection and estimation using generalized joint signal representations," *IEEE Trans. Signal Processing*, vol. 44, pp. 3031–3043, December 1996.

[11] R. A. Altes, "Detection, estimation and classification with spectrograms," *J. Acoustical Soc. of America*, vol. 67, pp. 1232–1246, April 1980.

[12] C. Richard and R. Lengellé, "Data-driven design and complexity control of time-frequency detectors," *Signal Processing*, vol. 77, pp. 37–48, August 1999.

[13] M. Davy, C. Doncarli, and G. F. Boudreaux-Bartels, "Improved optimization of time-frequency-based signal classifiers," *IEEE Signal Processing Letters*, vol. 8, pp. 52–57, February 2001.

12.2 TIME-FREQUENCY SIGNAL ANALYSIS AND CLASSIFICATION USING MATCHING PURSUITS [0]

12.2.1 Signal Time-Frequency Structures

Although time-frequency representations (TFRs) are the appropriate tools for nonstationary signal processing [1–3], it is important to match a TFR with the time-frequency (TF) structure of a signal for successful TF analysis and classification (see Section 1.1.5). In general, TFRs are ideally matched to one or two TF structures based on the properties they satisfy. However, some signals have multiple components with distinctively *different* instantaneous frequency (IF). For example, in a sonar acoustics application, a received signal (e.g. a linear frequency-modulated (FM) signal) may be interfered by whistles from a group of dolphins (e.g. with hyperbolic IF). The Wigner distribution (WD) can accurately represent signals with constant or linear IF whereas the Altes Q-distribution can accurately represent signals with hyperbolic (dispersive) IF. However, neither one of the two quadratic TFRs will ideally analyze all the components of a signal consisting of both linear and hyperbolic IF as in the aforementioned example. The TF analysis and classification of multicomponent signals is further complicated by the presence of cross terms when quadratic TFRs are used (see [3] and Article 4.2).

Some TFRs that can analyze signals with multiple TF structures include the spectrogram [3], reassigned TFRs (see [4] and Article 7.2), and various adaptive TFRs [5]. The IF of various signals was also estimated using TFRs [6]. Although they work well in many applications, such TFRs do not automatically yield the IF of a signal in closed form, and do not always provide a well-concentrated representation without cross terms for analyzing signals with nonlinear TF structures. Thus, it is advantageous for an adaptive TFR to exactly match and classify signal components as many natural or synthetic signals have different linear and nonlinear IF.

12.2.2 Matching Pursuits for Analysis and Classification

The matching pursuit iterative algorithm of Mallat and Zhang decomposes a signal into a linear expansion of waveforms selected from a redundant and complete dictionary [7]. It uses successive approximations of the signal with orthogonal projections on dictionary elements. The dictionary consists of a basic Gaussian atom that is TF shifted and scaled. A TFR (called the *modified Wigner distribution* in [7]) is obtained as a weighted sum of the WD of each selected element. This TFR is free of cross terms, and preserves signal energy, TF shifts, and scale changes on the analysis signal. In order to analyze linear FM signals more efficiently with fewer waveforms, rotated Gaussian atoms were included in the dictionary in [8]. On the other hand, a wave-based dictionary consisting of wavefronts, resonances, and linear

[0]Authors: **Antonia Papandreou-Suppappola**, Telecommunications Research Center, Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-7206 USA (papandreou@asu.edu), and **Seth B. Suppappola**, Acoustic Technologies, Inc., 1620 South Stapley Dr., Mesa, AZ 85204 USA (seth.suppappola@acoustictech.com). Reviewers: F. Auger and X. Xia.

FM signals was used to process scattering data in [9].

Although the matching pursuit algorithm in [7] works well for many signals, it uses many Gaussian atoms to represent a signal component with nonlinear TF characteristics. In addition, the modified WD is not well-concentrated along the nonlinear IF of some signal components. In order to be able to exactly match linear or nonlinear FM signals, a matching pursuit is used with a dictionary of waveforms that includes complex sinusoids with linear or nonlinear phase functions such as logarithmic and power functions. The aim of the adaptive (modified) matching pursuit (AMMP) is to be able to analyze and correctly classify multicomponent signals where each component has a specific (and often different) monotonic phase function that can be written in closed form [10]. The advantage of using a dictionary that is matched to the analysis data is that only a small number of elements are used to decompose the signal, and the algorithm gives fast and parsimonious results. At each iteration, the algorithm *adaptively* chooses the best dictionary element according to some condition, identifies its TF structure, and computes a specific TFR of the element. This TFR is adaptively chosen so that it provides a well-concentrated representation (without interference terms) of the selected element. The resulting TFR is formed as a weighted linear superposition of these matched TFRs. In this respect, it is similar to the TFR obtained in [11]. The computation of the TFR is only necessary for signal analysis. The algorithm can also be used to identify and classify each signal component. This will be useful, for example, in the classification of multiple received signals in a detection application. Thus, as will be shown, the AMMP decreases the number of algorithm iterations, improves the TF concentration of different multiple nonlinear FMs, and correctly classifies the IF of each signal component.

Adaptive matching pursuit and TFR The AMMP is based exclusively on the original algorithm in [7], but it has some major differences [10]. First, the AMMP uses more than one type of basic atom in the dictionary. Particularly, the dictionary consists of a large class of different basic atoms each of which has the form of a nonlinear FM signal [Article 5.6]

$$g(t;\,\xi,\lambda) = \sqrt{|v(t)|}\, e^{j2\pi\,\lambda\,\xi(\frac{t}{t_r})} \tag{12.2.1}$$

which is uniquely specified by its FM rate λ and its monotonic phase function $\xi(b)$. Note that $v(t) = \frac{d}{dt}\xi(\frac{t}{t_r})$ is the IF of the signal in (12.2.1), and $t_r > 0$ is a reference time. The dictionary may consist of only one type of FM signal with fixed $\xi(b)$ in (12.2.1) or a linear combination of them including sinusoids with $\xi(b) = b$, linear FM signals with $\xi(b) = \mathrm{sgn}(b)|b|^2$ (where $\mathrm{sgn}(b)$ is -1 for $b < 0$ and 1 for $b > 0$), hyperbolic FM signals with $\xi(b) = \ln b$, power FM signals with $\xi(b) = \mathrm{sgn}(b)|b|^\kappa$, and exponential FM signals with $\xi(b) = e^b$. The dictionary is formed by transforming the FM signal in (12.2.1) as

$$\begin{aligned} g(t;\,\xi,\lambda,\underline{\theta}) &= \Big(\mathcal{S}_\tau\, \mathcal{C}_a\, \mathcal{G}_c^{(\xi)}\, g(\xi,\lambda)\Big)(t) = \sqrt{|a|}\, g\Big(a(t-\tau);\,\xi,\lambda\Big)\, e^{j2\pi\, c\,\xi(a(\frac{t-\tau}{t_r}))} \\ &= \sqrt{|a\, v(a(t-\tau))|}\, e^{j2\pi\,(c+\lambda)\,\xi(a(\frac{t-\tau}{t_r}))}, \end{aligned} \tag{12.2.2}$$

with the parameter vector $\underline{\theta} = [c, a, \tau] \in (\Theta = \mathbb{R}^3)$. The dictionary is normally formed with a linear (i.e., grid) spacing of the parameters in $\underline{\theta}$, although this is not a requirement. The unitary operators $\mathcal{G}_c^{(\xi)}$, $\mathcal{C}_a$, and $\mathcal{S}_\tau$ result in an IF shift c, scale change a, and constant time shift τ, respectively, of the FM signal. Specifically, the operators transform a signal $x(t)$ as $(\mathcal{G}_c^{(\xi)}x)(t) = x(t)\,e^{j2\pi c\,\xi(\frac{t}{t_r})}$ (where $\xi(b)$ is the same function as in (12.2.1)), $(\mathcal{C}_a x)(t) = \sqrt{|a|}\,x(a\,t)$, and $(\mathcal{S}_\tau x)(t) = x(t-\tau)$. In (12.2.2), the transformation $\mathcal{G}_c^{(\xi)}$ results in a constant shift (from λ to $c+\lambda$) of the FM rate of the nonlinear FM signal instead of the constant frequency shift used in [7]. Hence, λ is not a parameter of the optimization, and is instead used to specify a base FM rate in (12.2.1). Shifts in the FM rate are affected by varying the c parameter. This is performed in this way since the signals considered may be wideband as well as dispersive, thus a shift of the IF is a better matched transformation (that covers the entire TF plane) than a constant frequency shift. With appropriate normalization, the energy of $g(t;\xi,\lambda,\underline{\theta})$ is restricted to be unity for every $\underline{\theta}$ in order to ensure energy preservation when $\xi(b)$ is fixed [7,10]. The iterative procedure of the AMMP first projects the analysis signal $x(t) = (R_0\,x)(t)$ onto each element of the dictionary, and selects[1] $g(t;\xi_0,\lambda,\underline{\theta}_0)$ based on the condition[2] $|\langle x,\, g(\xi_0,\lambda,\underline{\theta}_0)\rangle| \geq |\langle x,\, g(\xi,\lambda,\underline{\theta})\rangle|$, $\forall\,\underline{\theta} \in \Theta$ and for all possible phase functions $\xi(b)$ of the dictionary elements. This ensures that the signal component with the highest energy is separated first by being matched to the element $g(t;\xi_0,\lambda,\underline{\theta}_0)$ that yields the largest correlation with the signal. This results in

$$x(t) = \beta_0\, g(t;\xi_0,\lambda,\underline{\theta}_0) + (R_1 x)(t) \tag{12.2.3}$$

with the expansion coefficient $\beta_0 = \langle x,\, g(\xi_0,\lambda,\underline{\theta}_0)\rangle$. The function $\xi_0(b)$ is the phase function of the first selected element. For example, if this element is a hyperbolic FM signal, then $\xi_0(b) = \ln b$.

The second major difference of the AMMP from the algorithm in [7] is that the algorithm does not compute the WD of selected elements to form the modified WD. Instead, it adaptively uses the information that the first selected waveform has phase function $\xi_0(b)$ to compute its IF shift covariant *dispersively warped Wigner distribution* (DWD) [Article 5.6]. The DWD is a warped version of the WD, $W_y(t,f) = \int_{-\infty}^{\infty} y(t+\tau/2)y^*(t-\tau/2)e^{-j2\pi f\tau}d\tau$, with the warping based on a monotonic and (possibly) nonlinear parameter function $\zeta(b)$. In particular,

$$W_x^{(\zeta)}(t,f) = W_y\left(t_r\zeta(\frac{t}{t_r}),\, \frac{f}{t_r\mu(t)}\right) \tag{12.2.4}$$

where $\mu(t) = \frac{d}{dt}\zeta(\frac{t}{t_r})$, and $y(t) = (\mathcal{W}_\zeta x)(t) = t_r|\mu(t_r\zeta^{-1}(\frac{t}{t_r}))|^{-1/2}x(t_r\zeta^{-1}(\frac{t}{t_r}))$ [Article 5.6] is the warped signal. Note that a specific DWD is obtained simply by fixing its parameter function $\zeta(b)$. By matching $\zeta(b)$ in (12.2.4) to be equal to the

[1]Note that a subscript n in the parameters R_n, $\xi_n(b)$, $\underline{\theta}_n$, τ_n, and c_n, and a superscript n in a TFR $T^n(t,f)$ indicate the algorithm parameters at the $(n+1)$th iteration.

[2]The inner product is defined as $\langle x,\, g\rangle = \int_{-\infty}^{\infty} x(t)\,g^*(t)dt$.

phase function $\xi_0(b)$ in (12.2.3) (i.e. if $\zeta(b) = \xi_0(b)$), an *adaptive representation for multiple structures* (ARMUS) TFR, at this first iteration, is simply given as

$$T_x^0(t,f) = |\beta_0|^2 \, W^{(\xi_0)}_{g(\xi_0,\underline{\theta}_0)}(t,f).$$

At the second iteration, the residual function $(R_1 x)(t)$ is obtained by solving (12.2.3), and it is decomposed in a similar manner as the signal $x(t)$. At the $(n+1)$th iteration, the condition

$$|\langle R_n x,\, g(\xi_n, \underline{\theta}_n)\rangle| \;\geq\; |\langle R_n x,\, g(\xi, \underline{\theta})\rangle|, \quad \forall\, \underline{\theta} \in \Theta \tag{12.2.5}$$

is used to decompose the nth residual function $(R_n\, x)(t) = \beta_n\, g(t;\, \xi_n, \underline{\theta}_n) + (R_{(n+1)}\, x)(t)$ where

$$\beta_n = \langle\, R_n x,\, g(\xi_n, \underline{\theta}_n)\,\rangle \tag{12.2.6}$$

is the expansion coefficient. The DWD of $(R_n x)(t)$ is also obtained adaptively to match the TF structure of the nth residual function by letting $\zeta(b) = \xi_n(b)$ in (12.2.4).

After a total of N iterations, the matching pursuit algorithm results in the signal decomposition

$$x(t) = \sum_{n=0}^{N-1} \beta_n\, g(t;\, \xi_n, \lambda, \underline{\theta}_n) + (R_N\, x)(t)\,. \tag{12.2.7}$$

As the dictionary is complete [10], any signal can be represented as in (12.2.7) with $N = \infty$ (yielding $(R_N\, x)(t) = 0$) [7]. In actuality, when the signal components match the TF structure of the dictionary elements, the algorithm converges quickly. A maximum number of iterations and an acceptable small residue energy compared to the data energy are used as stopping criteria [7]. The resulting ARMUS of the signal at the Nth iteration is the weighted sum of the corresponding DWD of each selected dictionary element

$$\mathrm{ARMUS}_x(t,f) = T_x^{N-1}(t,f) = \sum_{n=0}^{N-1} |\beta_n|^2 \, W^{(\xi_n)}_{g(\xi_n,\underline{\theta}_n)}(t,f)\,, \tag{12.2.8}$$

with the weights $|\beta_n|^2$ defined in (12.2.6). Note that the same DWD (with fixed $\xi(b)$) may be used in (12.2.8) for components with the same TF structure but different FM rate.

It is important to note that the success of the algorithm depends highly on the choice of dictionary elements. As a result, pre-processing the data might be necessary in order to avoid a poor algorithm performance due to a mismatch between the analysis data and the dictionary elements used. Although this analysis technique follows directly from the matching pursuit in [7] with some simple modifications, it is very powerful once the dictionary is matched to the data not only for analysis but also for classification.

Properties of the algorithm and the TFR An important property of the matching pursuit in (12.2.7) is its covariance to certain signal changes. Consider the decomposed signal $x(t) = \sum_{n=0}^{\infty} \beta_n \, g(t; \xi, \lambda, \underline{\theta}_n)$ in (12.2.7) with $N = \infty$, and with identical TF structure dictionary elements (i.e. let $\xi_n(b) = \xi(b)$, $\forall n$). If the FM rate of a nonlinear FM signal $x(t)$ is shifted by a constant amount to form $y(t) = (\mathcal{G}_u^{(\xi)} x)(t) = x(t)\, e^{j2\pi u \, \xi(\frac{t}{t_r})}$, then its matching pursuit is simply given as $y(t) = \sum_{n=0}^{\infty} \beta_n \, g(t; \xi, \lambda, \hat{\underline{\theta}}_n)$. Note that the expansion coefficients β_n are not affected by this signal change. The parameter vector changes to $\hat{\underline{\theta}}_n = [c_n + u, a_n, \tau_n]$ indicating that the time shifts τ_n and the scale changes a_n remain the same, whereas the dictionary elements undergo a constant shift in their FM rate from $(\lambda + c_n)$ to $(\lambda + c_n + u)$. Note that if $\xi(b)$ is a power or a logarithmic function, then we can show that the corresponding AMMP is also covariant to scale changes [10].

The ARMUS TFR in (12.2.8) also satisfies various properties that are desirable in many applications. By simply combining the DWDs of each selected dictionary element, no cross terms are introduced in the TFR. Also, the TFR preserves the underlying TF structure of each analysis signal component, and it provides a highly concentrated representation of each component as it does not apply any smoothing. Specifically, the DWD with parameter $\xi(b)$ of a nonlinear FM signal with phase function $\xi(b)$ results in the highly concentrated representation $\mathrm{DWD}_{g(\xi,\lambda)}(t, f; \xi) = |\upsilon(t)| \, \delta(f - \lambda \, \upsilon(t))$ [Article 5.6] with $\upsilon(t) = \frac{d}{dt}\xi(\frac{t}{t_r})$. If a particular application uses signal components with only one type of TF structure, then we should form our dictionary using the corresponding nonlinear FM signal with matched IF. In such cases, the ARMUS satisfies other desirable signal properties such as the preservation of signal energy, and changes in the analysis signal's FM rate [10]. If the dictionary elements are either hyperbolic or power FM signals, then the TFR can also be shown to preserve scale changes.

Implementation issues Since the algorithm parameters vary in order to form the dictionary elements for the matching pursuit, the computation is intensive. However, if the data is pre-processed, one can form a dictionary with elements which approximately span the data in TF structure. Thus, the algorithm iterates more rapidly. Additional speedup is possible if the matched DWD of each dictionary element is computed ahead of time. Since the last operation on the basic atom in (12.2.2) is time shifting, the inner products in the AMMP condition in (12.2.5) can be computed as a cross-correlation instead of introducing another layer of dictionary elements over all possible time shifts. This increases the computational speed since correlations can be implemented using the fast Fourier transform (FFT). Also, the memory consumption by the dictionary is significantly reduced since additional dictionary elements are not needed for every time shift. Moreover, since the elements do not change, and the residual data is constant during a given AMMP algorithm iteration, additional speedup could be achieved by pre-computing and storing the FFTs of these sequences.

If the signal components are well-separated in time, the algorithm can simply find the time support and phase function of each selected element, and then use the

information to analyze the actual data (instead of the selected waveforms) with its matched DWD. This will greatly reduce computation as only a few DWDs need to be obtained. If classification is needed without analysis, then there is no need to compute the TFR. The AMMP will provide the IF and the FM rate of each component. It can also provide other relevant parameters depending on how adapted the dictionary is to a particular application.

12.2.3 Simulation Example

A synthetic multicomponent signal is used to demonstrate the significance of the AMMP method. The signal is 512 points long (although it need not be a power of two) and consists of *thirteen* components: four windowed hyperbolic FM signals, five windowed linear FM signals, and four Gaussian waveforms, all with varying parameters. Their "ideal" TF representation shown in Fig. 12.2.1(a) is obtained by plotting the IF of each of the hyperbolic and linear FM signals (hyperbolae and lines with positive slope) and the characteristic TF curve of each Gaussian component (ellipses). The WD in Fig. 12.2.1(b) suffers from cross terms and makes it difficult to identify the true TF structure of each component. On the other hand, the spectrogram in Fig. 12.2.1(c) suffers from loss of resolution due to smoothing that complicates the identification of the exact number of signal terms. Furthermore, the spectrogram cannot provide parameter values for any signal component.

The dictionary was formed using hyperbolic FM signals, linear FM signals, and Gaussian waveforms. Note that the Gaussian elements are not in the form of (12.2.1). However, they were included in the dictionary for a fair comparison with the matching pursuit in [7], and due to the *a priori* knowledge of their presence in the data. The decomposition approximates the data very well after only *twenty* iterations as demonstrated by overlaying the signal with its expansion. The ARMUS TFR in Fig. 12.2.1(d) provides a highly concentrated representation for all thirteen components without outer cross terms or loss of resolution. This is because it adaptively computes the Altes Q-distribution for selected elements with hyperbolic TF characteristics, and the WD for selected elements with linear or constant TF characteristics. Note that the mild spreading of the signal components as well as some inner interference terms are due to the fact that the data was windowed for processing. As the algorithm iterated twenty times, the ARMUS in Fig. 12.2.1(d) is a TFR of all twenty elements that the algorithm selected. However, only the thirteen elements matching the actual signal components are visible. This is because the energy term of the signal residues was very small for seven of the twenty iterations. If TFR analysis is not required, the ARMUS TFR computation step is removed from the algorithm, and the results are used only for classification. In this example, the components were classified in three signal types (hyperbolic FM, linear FM, and Gaussian). Furthermore, the values of the FM rate of each hyperbolic and linear waveform as well as the variance of each Gaussian component were directly obtained and found to be very similar to the parameters of the actual data.

For further comparison, the matching pursuit with only Gaussian dictionary

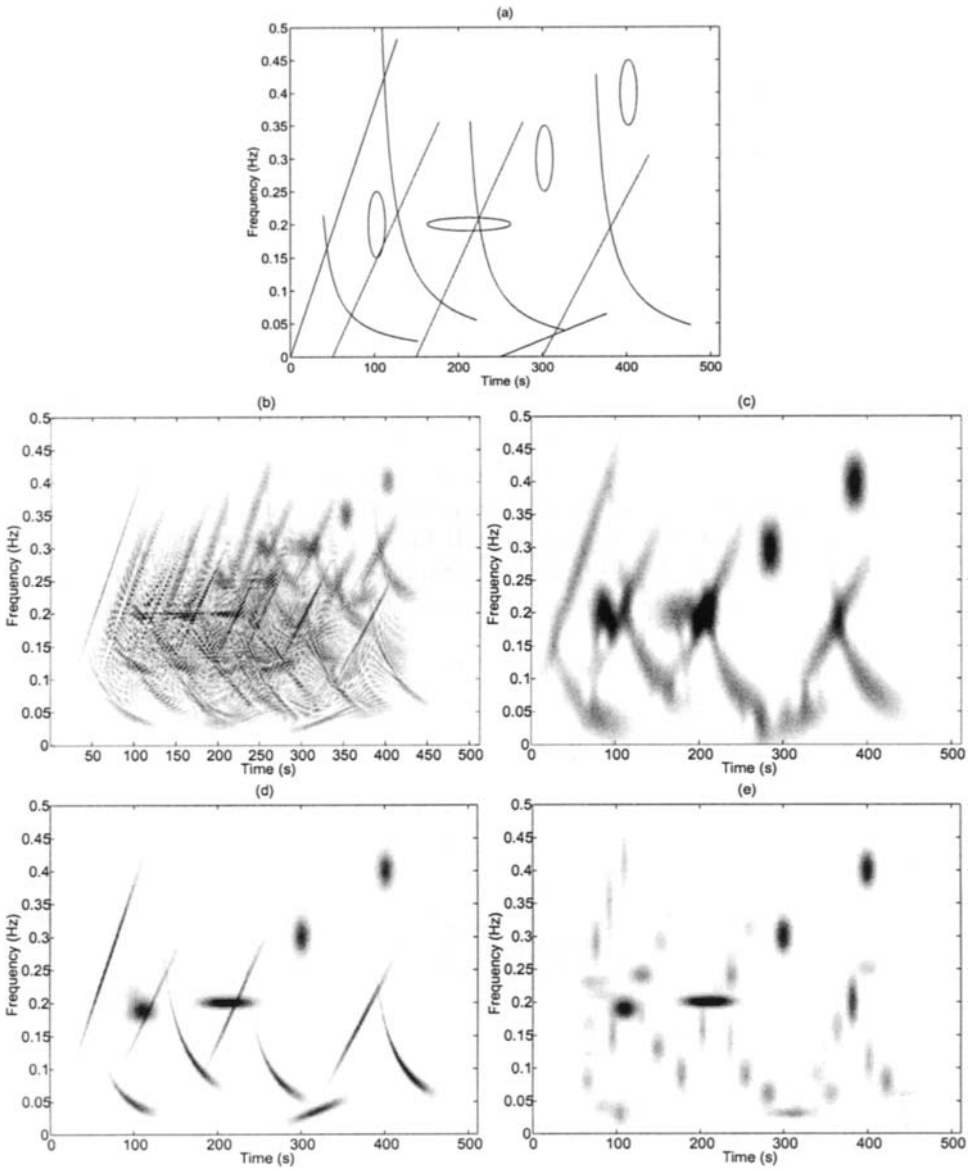


Fig. 12.2.1: (a) A linear combination of the ideal TF structure of each component of a signal consisting of four windowed hyperbolic FM signals, five windowed linear FM signals, and four windowed Gaussian signals. Note that the TF structure of a Gaussian waveform is represented as an ellipse delineating the approximate time and frequency support of the signal component. The signal is analyzed using (b) the Wigner distribution, (c) the spectrogram, (d) the ARMUS TFR, and (e) the modified WD in [7].

elements was used to decompose the signal, and then analyze it using the modified WD [7] as shown in Fig. 12.2.1(e). After 275 iterations, the TFR does not provide a concentrated representation that can easily identify the TF structure of the linear and hyperbolic FM components even though no cross terms are present. Also, the algorithm does not provide a closed form estimate of the IF and FM rate of the signal components for classification. On the other hand, as it is well-matched to Gaussians, the four Gaussian signal components are highly concentrated.

For relative comparison purposes, the speeds of the various methods can be contrasted as performed on a Pentium III based PC running MATLAB^{TM} under Linux. Since the Wigner distribution and the spectrogram use no form of matching pursuit iteration, they are considered to provide essentially instant results requiring only a few seconds to generate a plot. The ARMUS representation required approximately five minutes to compute, whereas the modified Wigner distribution consumed approximately forty-five minutes. The slowness of the modified Wigner distribution is primarily due to the high number of terms (i.e., iterations) required to approximate the analysis signal since Gaussian dictionary elements are not well matched to all of the analysis signal components. In fact, it takes several Gaussians to represent a single hyperbolic FM signal. The ARMUS representation, however, also contains hyperbolic and linear FM dictionary elements and is consequently better matched to the analysis signal.

A real data example using the AMMP method for classification can be found in [12].

12.2.4 Summary and Conclusions

In this article, we have presented a method of analyzing complex multicomponent time-frequency signal structures without the usual trade-off of TF resolution versus cross terms. The iterative approach is based on the matching pursuit of [7] but extended to include non-Gaussian signal types. Additionally, we generate a TFR which is free of cross terms by summing TFRs for the decomposed signal components. The TFR for each individual component is chosen to be ideally suited for the particular signal type extracted by the algorithm. Finally, since the TFR not only represents the individual components, but also parameterizes them, this algorithm lends itself nicely as a feature extraction tool for signal classification purposes.

References

[1] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[3] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[4] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[5] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Trans. Signal Processing*, vol. 41, pp. 1589–1602, April 1993.

[6] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals; Part 2: Algorithms and applications," *Proc. IEEE*, vol. 80, pp. 519–568, April 1992. With introductions by Michael J. Riezenman.

[7] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397–3415, December 1993.

[8] A. Bultan, "A four-parameter atomic decomposition of chirplets," *IEEE Trans. Signal Processing*, vol. 47, pp. 731–745, March 1999.

[9] M. R. McClure and L. Carin, "Matching pursuits with a wave-based dictionary," *IEEE Trans. Signal Processing*, vol. 45, pp. 2912–2927, December 1997.

[10] A. Papandreou-Suppappola and S. B. Suppappola, "Adaptive time-frequency representations for multiple structures," in *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pp. 579–583, Pocono Manor, PA, 14–16 August 2000.

[11] S. Qian and D. Chen, "Decomposition of the Wigner distribution and time-frequency distribution series," *IEEE Trans. Signal Processing*, vol. 42, pp. 2836–2842, October 1994.

[12] S. Pon Varma, A. Papandreou-Suppappola, and S. B. Suppappola, "Matching pursuit classification for time-varying acoustic emissions," in *Proc. 35th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, 4–7 November 2001. Paper TA2-3.

12.3 SYSTEM IDENTIFICATION USING TIME-FREQUENCY FILTERING[0]

12.3.1 Problem Description

A discrete time system identification problem can be stated as follows:

$$y[n] = \sum_k q[n-k]x[k] + \epsilon[n], \tag{12.3.1}$$

where $x[n]$ is a transmitted signal, $q[n]$ is the impulse response of a linear time invariant (LTI) system, $\epsilon[n]$ is an additive noise, and $y[n]$ is the received signal. The problem is to identify the LTI system transfer function $Q(\omega)$, i.e., the Fourier transform, of $q[n]$ given the input and the output signals $x[n]$ and $y[n]$.

The conventional method for solving the above problem is the least-squares solution method that is equal to the cross-spectral method in stationary cases, i.e., the system transfer function $Q(\omega)$ can be estimated by (see e.g. [1])

$$Q(\omega) = \frac{S_{xy}(\omega)}{S_{xx}(\omega)}, \tag{12.3.2}$$

where $S_{xy}(\omega)$ is the cross-spectrum of $x[n]$ and $y[n]$, and $S_{xx}(\omega)$ is the auto-spectrum of $x[n]$. When the additive noise $\epsilon[n]$ in (12.3.1) is a zero-mean Gaussian process and statistically independent of the input signal $x[n]$, the estimate in (12.3.2) is asymptotically unbiased but the performance is limited by the noise variance, or the signal-to-noise ratio (SNR). When this SNR is low, the performance of the estimate in (12.3.2) is poor as we will also see later. Since the auto-spectrum of the input signal $x[n]$ is in the denominator in the estimate (12.3.2), the input signal is, in general, chosen as a pseudo-random signal with flat spectrum. With this kind of input signals, noise reduction techniques before system identification do not apply.

In the following, we introduce a different technique [2] for the system identification problem. The main idea is as follows. Instead of pseudo-random signal $x[n]$, chirp type signals are transmitted as training signals, which have wideband characteristics in the frequency domain but are concentrated in the joint time-frequency (TF) domain. The TF concentration property usually holds after passing through an LTI system (this will be seen later). Since a joint TF distribution usually spreads noise and localizes signals, in particular chirps, the receiver may use a TF analysis technique to map the received signal $y[n]$ from the time domain into the joint TF domain. In this way, the SNR can be significantly increased in the joint TF domain [3]. Furthermore, TF filtering can be used in the TF plane to reduce the noise and the SNR in the time domain can be increased and therefore the system identification after denoising can improved.

[0]Author: **Xiang-Gen Xia**, Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA (xxia@ee.udel.edu). Reviewers: Paulo M. Oliveira and Shie Qian.

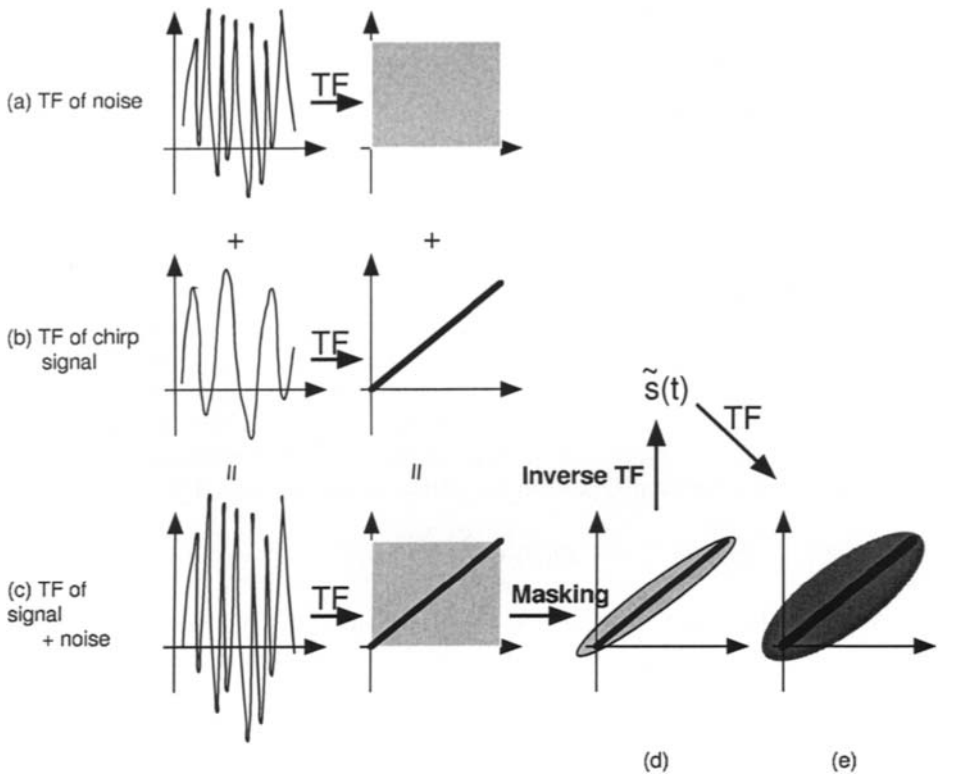


Fig. 12.3.1: TF filtering.

12.3.2 Time-Frequency Filtering

TF filtering (described in Chapter 11) includes three major components, namely a TF analysis or mapping, with a TF tool such as a TF distribution, from the time domain to the joint TF domain, a masking or filtering in the joint TF plane, and a TF synthesis from the joint TF plane back to the time domain, as illustrated in Fig. 12.3.1 for a chirp signal.

Differently from what happens with the Fourier transform in the Fourier transform based filtering, the TF mapping in a TF filtering may not be onto. In other words, not every signal $S[k,l]$ defined in the joint TF plane corresponds to a time domain signal $s[n]$ such that the TF mapping of the time domain signal $s[n]$ is exactly equal to $S[k,l]$. This causes problems in filtering in the TF domain, since the filtered signal in the TF domain may not correspond to any time domain signal as shown in Fig. 12.3.1 (a) and (e). An intuitive solution for this problem is to take the least-squares error (LSE) solution in the time domain, (see, for example, [4]). The LSE, however, usually does not have a desired TF characteristics in the TF domain. When a signal is very long, the computational load for the LSE solution

is significantly high because of the inverse matrix computation. Based on these observations, an iterative algorithm based on the discrete Gabor transform (DGT) was proposed in [5]; see Article 11.2. Conditions on the convergence, properties of the limit signals, and the relationship between the LSE solutions and solutions from the iterative algorithms were obtained in [6], where a significant improvement over the LSE solution was also shown. In the remainder of this article, we adopt the DGT as the TF analysis.

12.3.2.1 Discrete Gabor Transform

Let a signal $s[n]$, a synthesis window function $h[n]$ and an analysis window function $\gamma[n]$ be all periodic with same period L. Then,

$$s[k] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} C_{m,n} h_{m,n}[k], \tag{12.3.3}$$

$$C_{m,n} = \sum_{k=0}^{L-1} s[k] \gamma^*_{m,n}[k], \tag{12.3.4}$$

$$h_{m,n}[k] = h[k - m\Delta M] W_L^{n\Delta N k}, \tag{12.3.5}$$

$$\gamma_{m,n}[k] = \gamma[k - m\Delta M] W_L^{n\Delta N k}, \tag{12.3.6}$$

and $W_L = \exp(j2\pi/L)$, $j = \sqrt{-1}$. where ΔM and ΔN are the time and the frequency sampling interval lengths, and M and N are the numbers of sampling points in the time and the frequency domains, respectively, $M \cdot \Delta M = N \cdot \Delta N = L$, $MN \geq L$ (or $\Delta M \Delta N \leq L$). The coefficients $C_{m,n}$ are called the **discrete Gabor transform** (DGT) of the signal $s[k]$ and the representation (12.3.3) is called the **inverse discrete Gabor transform** (IDGT) of the coefficients $C_{m,n}$. Let H and G be the following $\Delta M \Delta N$ by L and MN by L matrices, respectively: the element at the $(mN + n)$th row and the kth column in H is

$$h[k - m\Delta M] W_L^{n\Delta N k}, \ 0 \leq m \leq M-1, 0 \leq n \leq N-1, 0 \leq k \leq L-1,$$

and the element at the kth row and the $(mN + n)$th column in G is

$$\gamma^*_{m,n}[k] = \gamma^*[k - m\Delta M] W_L^{-n\Delta N k}, \ 0 \leq m \leq M-1, \ 0 \leq n \leq N-1, \ 0 \leq k \leq L-1.$$

Then, the DGT and IDGT can be rewritten as

$$\mathbf{C} = G\mathbf{s} \quad \text{and} \quad \mathbf{s} = H\mathbf{C}, \tag{12.3.7}$$

where the following condition is needed

$$HG = I, \tag{12.3.8}$$

where I is the $L \times L$ identity matrix. (For more details, see [5] and Articles 4.9, 6.3 and 11.2.)

12.3.2.2 Iterative TF Filtering Algorithm

Let D denote the mask in the TF plane, i.e., an MN by MN diagonal matrix with diagonal elements either 0 or 1. Let $\mathbf{s}$ be a signal with length L in the time domain. The first step in the time-variant filtering is to mask the DGT of $\mathbf{s}$: $\mathbf{C}_1 = DG\mathbf{s}$, where D masks a desired domain in the TF plane. Since the DGT G is a redundant transformation, the IDGT of $\mathbf{C}_1$, $H\mathbf{C}_1$, may not fall in the mask. In other words, in general,

$$GH\mathbf{C}_1 \neq DGH\mathbf{C}_1, \tag{12.3.9}$$

which is illustrated in Fig. 12.3.1(e). An intuitive method to reduce the difference between the right and the left hand sides of (12.3.9) is to mask the right hand side of (12.3.9) again and repeat the procedure, which leads to the following iterative algorithm:

$$\begin{aligned} \mathbf{s}_0 &= \mathbf{s}, && (12.3.10)\\ \mathbf{C}_{l+1} &= DG\mathbf{s}_l, && (12.3.11)\\ \mathbf{s}_{l+1} &= H\mathbf{C}_{l+1}, && (12.3.12)\\ & l = 0, 1, 2, ... \end{aligned}$$

For more details, see Article 11.2 and reference [6].

12.3.3 Denoising for Received Signals through a Noisy Channel

We first describe some parameters. The training signal $x[n]$ is

$$x[n] = \cos\left(2\pi \left[\frac{n+15}{150}\right]^4\right), \quad n = 0, 1, ..., 499. \tag{12.3.13}$$

The synthesis and analysis window functions $h[n]$ and $\gamma[n]$ are the orthogonal-like functions from [5,6] and Article 11.2. Fig. 12.3.2(a) shows the DGT of $x[n]$.

We use 20 tap LTI systems in our numerical examples. The system model is

$$y[n] = \sum_{k=0}^{N-1} q[k]x[n-k] + \epsilon[n], \tag{12.3.14}$$

where $N = 20$, $\epsilon[n]$ is additive white Gaussian noise, independent of $x[n]$, and

$$s[n] = \sum_{k=0}^{N-1} q[k]x[n-k], \tag{12.3.15}$$

is considered as the signal. The original SNR for the received signal is calculated by

$$10\log_{10}\left(\frac{\sum_{n=0}^{499}|s[n]|^2}{\sum_{n=0}^{499}|\epsilon[n]|^2}\right).$$

In the following, we randomly generate the channel $q[n]$. As an example, a random channel, the noiseless signal $s[n]$, the received time waveform $y[n]$ with SNR= -4.5dB, and their Fourier spectrum are shown in Fig. 12.3.2(b). The DGT of the received signal $y[n]$ with -4.5dB SNR is shown in Fig. 12.3.2(c). In Fig. 12.3.2(c), one is still able to see the chirp pattern in the joint time and frequency plane although that is impossible in the time or the frequency domain alone in Fig. 12.3.2(b).

12.3.3.1 Mask D Design

The pattern in the DGT domain of the above signal $s[n]$ in (12.3.15) is similar to the one for the signal $x[n]$ in Fig. 12.3.2(a). This is not only true for this particular example, but has also been the case for our numerous examples. The reason can be found in the following analytic argument.

Assume the chirp signal $x[n] = \exp(jcn^r)$ for some constants $r \geq 2$ and $c \neq 0$. Then,

$$s[n] = \sum_k q[k]x[n-k] = x[n] \sum_k q[k] \exp(jc \sum_{l=0}^{r-1} c_l n^l k^{r-l}),$$

which is dominated by the original chirp $x[n]$ for finite tap LTI systems $q[k]$. It is because that the highest chirp order of $s[n]$, r, and the corresponding chirp rate are the same as those of $x[n]$, while the chirp order for the above multiplier of $x[n]$ in $s[n]$

$$\sum_k h[k] \exp(jc \sum_{l=0}^{r-1} c_l n^l k^{r-l})$$

is only $r-1$. As a special case, when $r = 2$, $s[n] = x[n]\tilde{Q}(2cn)$, where $\tilde{Q}(\omega)$ is the Fourier transform of the signal $q[n]x[n]$: $\tilde{Q}(\omega) = \sum_k q[k]x[k] \exp(-j2cnk)$. When the channel $q[n]$ has only a finite tap, the function $\tilde{Q}(\omega)$ is usually a smooth signal.

Since the transmitted signal $x[n]$ is known to both transmitter and the receiver, by the above property its pattern in the DGT domain may help in designing a mask in the DGT domain for filtering noise, which motivates the following design algorithm.

A Mask Design Procedure:

Step 1. Implement the DGT, obtaining the coefficients $C_{m,n}$ of the transmitted signal $x[k]$.

Step 2. Threshold the DGT coefficients $C_{m,n}$ and have a mask D_x from $C_{m,n}$:

$$D_x(m,n) = \begin{cases} 1, & \text{if } |C_{m,n}| > t_0, \\ 0, & \text{otherwise,} \end{cases}$$

where t_0 is a predesigned positive number that is called a *thresholding constant.*

Step 3. Implement Steps 1-2 for the received signal $y[k]$ and design a mask D_y with thresholding constant t_1 from the DGT coefficients of $y[n]$ with another predesigned constant $t_1 > 0$.

Step 4. The final mask is the product of D_x and D_y: $D = D_x D_y$.

Since the DGT of the signal $x[n]$ usually dominates the DGT of the signal $s[n]$, the pattern in the DGT domain of the signal $s[n]$ is usually in a close neighborhood of the pattern in the DGT domain of $x[n]$. Therefore, the mask D_x is usually designed so that it covers a relatively large area, i.e., the thresholding constant t_0 in Step 2 is usually chosen not too large. Since the received signal $y[n]$ is from a noisy channel, the resolution of its DGT pattern may be reduced and therefore the thresholding constant t_1 in Step 3 is usually chosen not too small. Otherwise, the mask D_y will cover too much area you do not want. Let us see an example. The mask D_x from $x[n]$, the mask D_y from $y[n]$, their product $D = D_x D_y$, and the mask D_s from the true signal $s[n]$ are shown in Fig. 12.3.2(d), respectively. The SNR in this case is SNR=-1.4dB. The thresholding constants in Steps 1-3 are $t_0 = 0.12 \cdot \max(DGT(y))$ and $t_1 = 0.15 \cdot \max(DGT(y))$. It should be pointed out that the above mask design procedure may be improved by using more sophisticated designs. Possible improvements are: (i) to find the optimal thresholding constants t_0 and t_1 by training a large number of signals and systems; (ii) to use more sophisticated statistical detection method in the DGT domain for the received signal $y[n]$ instead of a simple thresholding in Step 3; (iii) to smooth the mask $D = D_x D_y$ since the true mask D_s is usually smooth due to the nature of a chirp signal but D_y from the noisy signal $y[n]$ may not be smooth. Some morphological operations, such as dilation, may be used for smoothing the mask D.

12.3.3.2 Denoising Experiments

We now implement the denoising using the iterative TF filtering algorithm with three masking techniques: using the mask $D = D_x$ from the transmitted signal; using the mask $D = D_y D_x$ as designed by Steps 1-4; using the true mask $D = D_s$. We run 100 tests in terms of different LTI systems $q[n]$ (randomly generated) and different additive white Gaussian noises $\epsilon[n]$ for each masking method and take their mean SNR. Nine iterative steps are used in the iterative algorithm (12.3.10)-(12.3.12). Fig. 12.3.3(a) shows the curves of the mean SNR versus iterative steps for the three masking methods.

12.3.4 System Identification

In this section, we first use the iterative time-variant filter (12.3.10)-(12.3.12) developed in the previous sections to reduce the additive white Gaussian noise $\epsilon[n]$ from the received signal $y[n]$. In the iterative time-variant filter, for calculation simplicity we choose the first masking method, i.e., the mask $D = D_x$, throughout this section. With this mask, two iterations are used in the TF filtering. We then implement the conventional system identification method, as shown in Fig. 12.3.3(b).

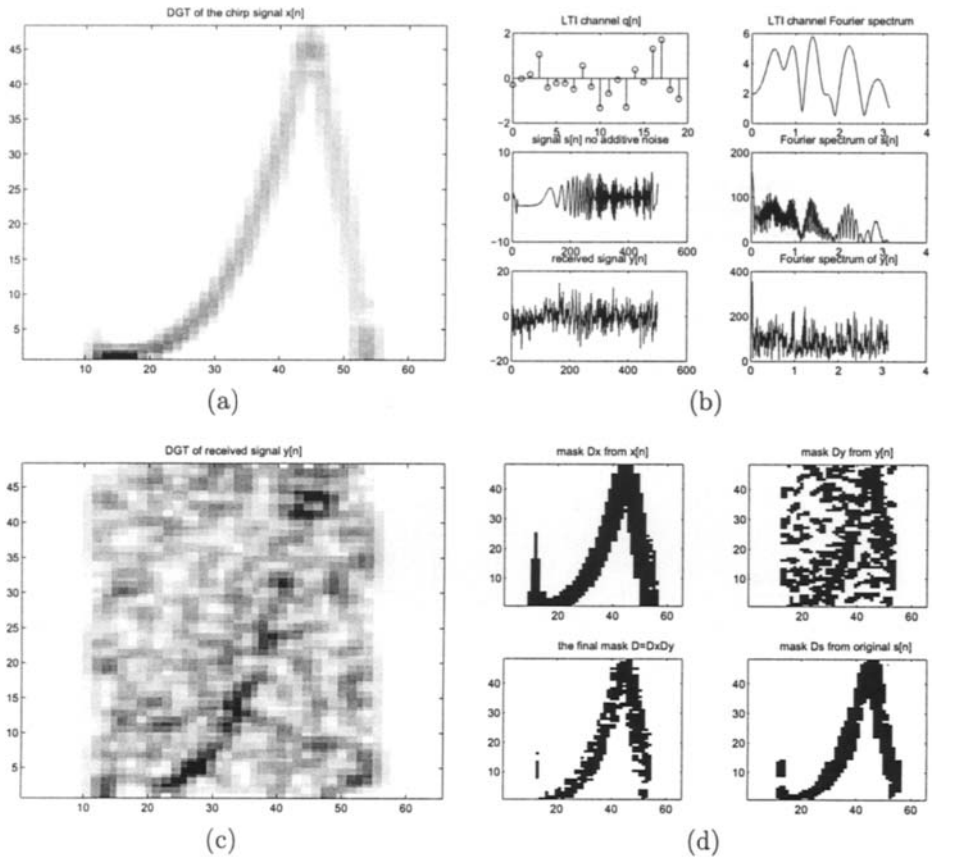


Fig. 12.3.2: (a) The discrete Gabor transform of signal $x[n]$. (b) An example of LTI channel $q[n]$, signal $s[n]$ and received signal $y[n]$, and their Fourier spectrum, where the SNR=−4.5dB for the additive white Gaussian noise. (c) The discrete Gabor transform of a received signal $y[n]$ with SNR=−4.5dB. (d) Example of masks D_x from $x[n]$, D_y from $y[n]$, the final mask $D = D_x D_y$ and the true mask D_s from $s[n]$.

The conventional system identification method used here is the cross-spectral method:

$$Q_{new}(\omega) = \frac{S_{\tilde{y}x}(\omega)}{S_{xx}(\omega)}, \tag{12.3.16}$$

where $x[n]$ is the chirp signal defined in (12.3.13) and $\tilde{y}$ is the signal after the denoising. It is compared with the conventional method without denoising, i.e.,

$$Q_{old_1}(\omega) = \frac{S_{yx}(\omega)}{S_{xx}(\omega)}, \tag{12.3.17}$$

where $x[n]$ is also the chirp signal. Since the system identification performance usually depends on the signal $x[n]$ you transmit, one might say that it is not fair to

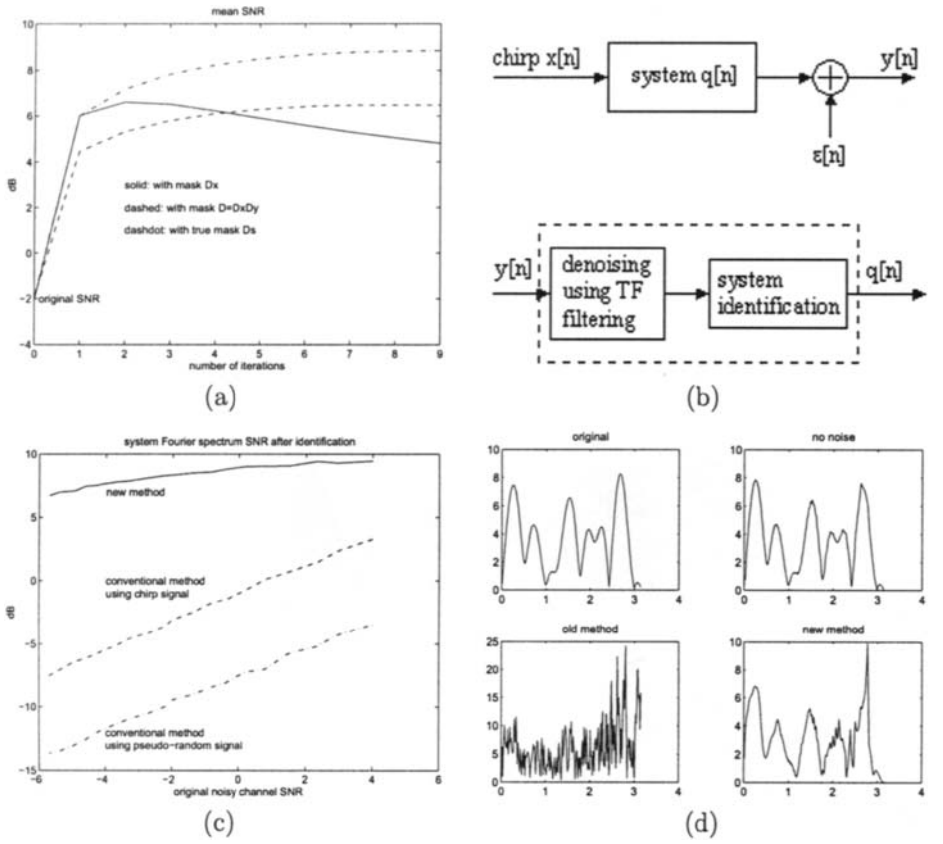


Fig. 12.3.3: (a) The mean SNR curves of the iterative time-variant filtering with the following masks: $D = D_x$, $D = D_x D_y$, and $D = D_s$. (b) New system identification method. (c) Comparison of system identification methods: the conventional method using chirp signals; the conventional method using pseudo-random signals; new method using chirp signals and time-variant filtering. (d) System identification examples: original spectrum $|Q(\omega)|$; identified spectrum without additive noise using the chirp signal; the conventional method with additive noise of SNR=-0.4dB; the new method with additive noise of SNR=-0.4dB.

compare them using the chirp signal that is preferred here for denoising purposes, but might not be preferred for other methods. For this reason, we also compare our new method with the conventional method using pseudo-random sequences:

$$Q_{old_2}(\omega) = \frac{S_{y\tilde{x}}(\omega)}{S_{\tilde{x}\tilde{x}}(\omega)}, \tag{12.3.18}$$

where $\tilde{x}[n]$ is a pseudo-random sequence.

Fig. 12.3.3(c) shows their performances, where 200 tests are used for the mean SNR curves for the system spectrum versus the original SNR. Our new method performs much better than others. Surprisingly, even for the conventional cross

spectral method, the chirp signal in (12.3.13) outperforms pseudo-random signals by approximately 6dB. In Fig. 12.3.3(d), some identification examples are shown, where the original SNR is -0.4dB. As a remark, all system identification calculations used in this article are based on the MATLAB^{TM} Signal Processing Toolbox.

12.3.5 Summary and Conclusions

System identification can be performed by transmitting chirp signals and applying iterative time-frequency filters based on a discrete Gabor transform/expansion. The filters are implemented by projecting the signal back and forth between the time domain and the joint time-frequency domain. This system identification method has better performance than the conventional methods at low signal-to-noise ratios.

References

[1] T. Södertröm and P. Stoica, *System Identification.* Englewood Cliffs. NJ: Prentice-Hall, 1989.

[2] X.-G. Xia, "System identification using chirp signals and time-variant filters in the joint time-frequency domain," *IEEE Trans. Signal Processing*, vol. 45, pp. 2072–2084, August 1997.

[3] X.-G. Xia and V. C. Chen, "A quantitative SNR analysis for the pseudo Wigner-Ville distribution," *IEEE Trans. Signal Processing*, vol. 47, pp. 2891–2894, October 1999.

[4] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 34, pp. 442–451, June 1986.

[5] S. Qian and D. Chen, *Joint Time-Frequency Analysis: Methods & Applications.* Upper Saddle River, NJ: Prentice-Hall, 1996.

[6] X.-G. Xia and S. Qian, "Convergence of an iterative time-variant filtering based on discrete Gabor transform," *IEEE Trans. Signal Processing*, vol. 47, pp. 2894–2899, October 1999.

12.4 TIME-FREQUENCY METHODS FOR SIGNAL ESTIMATION AND DETECTION [0]

This article uses the generalized Wigner-Ville spectrum (GWVS) and the generalized Weyl symbol (GWS) to develop time-frequency (TF) techniques for the estimation and detection of underspread nonstationary processes. These TF techniques extend optimal signal estimators (Wiener filters) and optimal signal detectors for the stationary case to underspread nonstationary processes. They are conceptually simple and intuitively appealing as well as computationally efficient and stable.

We will first review some fundamentals (for more details see Articles 4.7 and 9.4). The GWVS of a nonstationary random process $x(t)$ with correlation function $r_x(t,t') = \mathrm{E}\{x(t)\,x^*(t')\}$ is defined as

$$\overline{W}_x^{(\alpha)}(t,f) \triangleq \int_{-\infty}^{\infty} r_x\Big(t + \Big(\frac{1}{2}-\alpha\Big)\tau, t - \Big(\frac{1}{2}+\alpha\Big)\tau\Big)\, e^{-j2\pi f\tau} d\tau\,, \qquad (12.4.1)$$

where α is a real-valued parameter. The GWS of a linear, time-varying (LTV) system $\mathbf{H}$ with kernel (impulse response) $h(t,t')$ is defined as

$$L_{\mathbf{H}}^{(\alpha)}(t,f) \triangleq \int_{-\infty}^{\infty} h\Big(t + \Big(\frac{1}{2}-\alpha\Big)\tau, t - \Big(\frac{1}{2}+\alpha\Big)\tau\Big)\, e^{-j2\pi f\tau} d\tau\,.$$

In the case of a stationary process, the GWVS reduces to the conventional power spectral density. Similarly, for a linear, time-invariant (LTI) system, the GWS reduces to the conventional transfer function (frequency response).

A nonstationary process $x(t)$ is called *underspread* if components of $x(t)$ that are sufficiently separated in the TF plane are effectively uncorrelated (see Article 9.4). Two random processes $x(t)$, $y(t)$ are called *jointly underspread* if they satisfy similar underspread conditions [1]. An LTV system $\mathbf{H}$ is called underspread if it produces only small TF displacements of the input signal (see Article 4.7).

12.4.1 Nonstationary Signal Estimation

The enhancement or estimation of signals corrupted by noise or other interference is important in many applications. Here, we consider estimation of a nonstationary, zero-mean random signal $s(t)$ from a noise-corrupted observed signal $x(t) = s(t) + n(t)$ by means of an LTV system $\mathbf{H}$. The signal estimate is thus given by

$$\hat{s}(t) = (\mathbf{H}x)(t) = \int_{-\infty}^{\infty} h(t,t')\,x(t')\,dt'. \qquad (12.4.2)$$

The additive noise $n(t)$ is nonstationary and assumed uncorrelated with $s(t)$.

[0] Authors: **F. Hlawatsch** and **G. Matz**, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, Gusshausstrasse 25/389, A-1040 Vienna, Austria (fhlawats@pop.tuwien.ac.at, g.matz@ieee.org). Reviewers: P. Flandrin and S. Carstens-Behrens.

The time-varying Wiener filter. The LTV system $\mathbf{H}$ that minimizes the mean-square error (MSE) $\mathrm{E}\{\|\hat{s}-s\|^2\}$ is the *time-varying Wiener filter* [2,3]

$$\mathbf{H}_{\mathrm{W}} = \mathbf{R}_s(\mathbf{R}_s+\mathbf{R}_n)^{-1}. \tag{12.4.3}$$

Here, $\mathbf{R}_s$ and $\mathbf{R}_n$ denote the correlation operators[1] of signal and noise, respectively. For stationary random processes, $\mathbf{H}_{\mathrm{W}}$ is an LTI system whose frequency response is given by [2,3]

$$H_{\mathrm{W}}(f) = \frac{P_s(f)}{P_s(f)+P_n(f)}, \tag{12.4.4}$$

where $P_s(f)$ and $P_n(f)$ denote the power spectral density of signal and noise, respectively. This frequency-domain expression involves merely a product and a reciprocal of functions (instead of a product and an inverse of operators as in (12.4.3)) and thus allows a simple design and interpretation of time-invariant Wiener filters.

TF formulation of time-varying Wiener filters. We may ask whether a simple formulation similar to (12.4.4) can be obtained for the time-varying Wiener filter $\mathbf{H}_{\mathrm{W}}$ by replacing $H_{\mathrm{W}}(f)$ with the GWS $L^{(\alpha)}_{\mathbf{H}_{\mathrm{W}}}(t,f)$ and $P_s(f)$, $P_n(f)$ with the GWVS $\overline{W}_s^{(\alpha)}(t,f)$, $\overline{W}_n^{(\alpha)}(t,f)$. Indeed, for *jointly underspread* processes $s(t)$ and $n(t)$, it can be shown [1] that the time-varying Wiener filter $\mathbf{H}_{\mathrm{W}}$ can be written as the sum of two components: (i) an overspread (i.e., not underspread) system component that has negligible effect on the system's performance (MSE) and thus can be disregarded, and (ii) an underspread system component, hereafter denoted as $\mathbf{H}^u_{\mathrm{W}}$, that allows the approximate TF formulation

$$L^{(\alpha)}_{\mathbf{H}^u_{\mathrm{W}}}(t,f) \approx \frac{\overline{W}_s^{(\alpha)}(t,f)}{\overline{W}_s^{(\alpha)}(t,f)+\overline{W}_n^{(\alpha)}(t,f)}. \tag{12.4.5}$$

This TF formulation extends (12.4.4) to the underspread nonstationary case and allows a simple and intuitively appealing TF interpretation of the time-varying Wiener filter (see Fig. 12.4.1). Let $\mathcal{R}_s$ and $\mathcal{R}_n$ denote the effective support regions of $\overline{W}_s^{(\alpha)}(t,f)$ and $\overline{W}_n^{(\alpha)}(t,f)$, respectively. In the "signal only" TF region $\mathcal{R}_s\backslash\mathcal{R}_n$, (12.4.5) gives $L^{(\alpha)}_{\mathbf{H}^u_{\mathrm{W}}}(t,f)\approx 1$. Thus, $\mathbf{H}^u_{\mathrm{W}}$ passes all "noise-free" components of $x(t)$ without attenuation or distortion. In the "noise only" TF region $\mathcal{R}_n\backslash\mathcal{R}_s$, (12.4.5) gives $L^{(\alpha)}_{\mathbf{H}^u_{\mathrm{W}}}(t,f)\approx 0$, i.e., $\mathbf{H}^u_{\mathrm{W}}$ suppresses all components of $x(t)$ located in TF regions where there is no signal. Finally, in the "signal plus noise" TF region $\mathcal{R}_s\cap\mathcal{R}_n$, $|L^{(\alpha)}_{\mathbf{H}^u_{\mathrm{W}}}(t,f)|$ assumes values approximately between 0 and 1. Here, $\mathbf{H}^u_{\mathrm{W}}$ performs an attenuation that depends on the signal-to-noise ratio at the respective TF point.

[1] The correlation operator $\mathbf{R}_x$ of a nonstationary random process $x(t)$ is the positive (semi-) definite linear operator whose kernel equals the correlation function $r_x(t,t')=\mathrm{E}\{x(t)\,x^*(t')\}$.

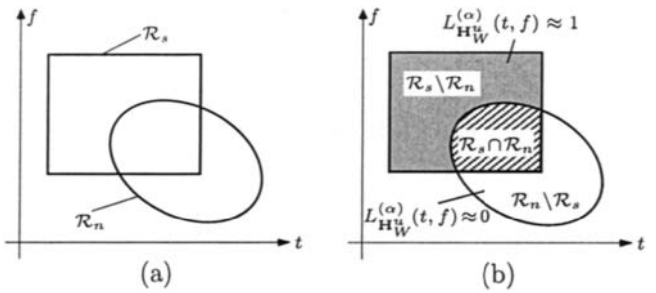


Fig. 12.4.1: TF interpretation of the time-varying Wiener filter $\mathbf{H}_W$ for jointly underspread signal and noise processes: (a) Effective TF support regions of signal and noise, (b) TF pass, stop, and transition regions of the time-varying Wiener filter.

TF design of time-varying Wiener filters. The TF formulation (12.4.5) suggests a simple *TF design* of nonstationary signal estimators. Let us define the "TF pseudo-Wiener filter" $\widetilde{\mathbf{H}}_W$ by setting its GWS equal to the right-hand side of (12.4.5) [1]:

$$L^{(\alpha)}_{\widetilde{\mathbf{H}}_W}(t,f) \triangleq \frac{\overline{W}_s^{(\alpha)}(t,f)}{\overline{W}_s^{(\alpha)}(t,f) + \overline{W}_n^{(\alpha)}(t,f)} . \tag{12.4.6}$$

For jointly underspread processes $s(t)$, $n(t)$ where (12.4.5) is a good approximation, the TF pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ will closely approximate (the underspread part of) the optimal Wiener filter $\mathbf{H}_W$; furthermore, it can be shown that $\widetilde{\mathbf{H}}_W$ will then be nearly independent of the value of α used in (12.4.6). For processes $s(t)$, $n(t)$ that are *not* jointly underspread, however, $\widetilde{\mathbf{H}}_W$ must be expected to perform poorly.

Whereas the TF pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ is *designed* in the TF domain, the signal estimate $\hat{s}(t)$ can be calculated in the time domain according to (12.4.2). The impulse response of $\widetilde{\mathbf{H}}_W$ is obtained from $L^{(\alpha)}_{\widetilde{\mathbf{H}}_W}(t,f)$ as (cf. Article 4.7)

$$\tilde{h}_W(t,t') = \int_{-\infty}^{\infty} L^{(\alpha)}_{\widetilde{\mathbf{H}}_W}\left(\left(\frac{1}{2}+\alpha\right)t + \left(\frac{1}{2}-\alpha\right)t', f\right) e^{j2\pi f(t-t')}\, df . \tag{12.4.7}$$

An efficient implementation of the TF pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ that is based on the multiwindow short-time Fourier transform is discussed in [1, 4].

Compared to the Wiener filter $\mathbf{H}_W$, the TF pseudo-Wiener filter $\widetilde{\mathbf{H}}_W$ possesses two practical advantages. First, the prior knowledge required for calculating $\widetilde{\mathbf{H}}_W$ is given by the GWVS $\overline{W}_s^{(\alpha)}(t,f)$ and $\overline{W}_n^{(\alpha)}(t,f)$ that are more intuitive and easier to handle than the correlation operators $\mathbf{R}_s$ and $\mathbf{R}_n$. Second, the TF design (12.4.6) is less computationally intensive and more stable than (12.4.3) since it requires pointwise (scalar) divisions of functions instead of operator inversions.

Robust TF Wiener filters. The performance of the filters $\mathbf{H}_W$ and $\widetilde{\mathbf{H}}_W$ is sensitive to deviations of the second-order statistics (correlations or GWVS) from the nominal

statistics for which these filters were designed. This motivates the use of *minimax robust Wiener filters* that optimize the worst-case performance (maximum MSE) within specified uncertainty classes of second-order statistics [5,6].

Consider a partition of the TF plane into K mutually disjoint TF regions $\mathcal{R}_i$, $i = 1, ..., K$. Extending the stationary case definition in [7], we define so-called *p-point uncertainty classes* $\mathcal{S}$ and $\mathcal{N}$ as the sets of all nonnegative TF functions (not necessarily valid GWVS) $\widetilde{W}_s(t,f)$ and $\widetilde{W}_n(t,f)$ that have prescribed energies s_i and n_i, respectively, within $\mathcal{R}_i$, i.e., $\iint_{\mathcal{R}_i} \widetilde{W}_s(t,f)\,dt\,df = s_i$ and $\iint_{\mathcal{R}_i} \widetilde{W}_n(t,f)\,dt\,df = n_i$ for $i = 1, ..., K$. For these uncertainty classes, the GWS of the *minimax robust TF Wiener filter* $\mathbf{H}_{\mathrm{R}}$ is given by [5,6]

$$L^{(\alpha)}_{\mathbf{H}_{\mathrm{R}}}(t,f) = \sum_{i=1}^{K} \frac{s_i}{s_i + n_i}\, I_{\mathcal{R}_i}(t,f)\,, \tag{12.4.8}$$

where $I_{\mathcal{R}_i}(t,f)$ is the indicator function of $\mathcal{R}_i$ (i.e., $I_{\mathcal{R}_i}(t,f)$ is 1 for (t,f) inside $\mathcal{R}_i$ and 0 outside $\mathcal{R}_i$). Note that $L^{(\alpha)}_{\mathbf{H}_{\mathrm{R}}}(t,f)$ is piecewise constant, expressing constant TF weighting within $\mathcal{R}_i$. The performance of $\mathbf{H}_{\mathrm{R}}$ is approximately independent of the actual second-order statistics as long as they are within $\mathcal{S}$, $\mathcal{N}$ [5,6]. Signal-adaptive, online implementations of robust time-varying Wiener filters using local cosine bases have been proposed in [6,8].

Simulation results. Fig. 12.4.2(a),(b) shows the Wigner-Ville spectra (GWVS with $\alpha = 0$) of jointly underspread signal and noise processes. The Weyl symbols (GWS with $\alpha = 0$) of the corresponding Wiener filter $\mathbf{H}_{\mathrm{W}}$, of its underspread part $\mathbf{H}^u_{\mathrm{W}}$, and of the TF pseudo-Wiener filter $\widetilde{\mathbf{H}}_{\mathrm{W}}$ are shown in Fig. 12.4.2(c)–(e). It is verified that the Weyl symbol of $\widetilde{\mathbf{H}}_{\mathrm{W}}$ approximates that of $\mathbf{H}^u_{\mathrm{W}}$. The mean SNR improvement achieved by the TF pseudo-Wiener filter $\widetilde{\mathbf{H}}_{\mathrm{W}}$ was obtained as 6.11 dB; this is nearly as good as that of the Wiener filter $\mathbf{H}_{\mathrm{W}}$ (6.14 dB).

To illustrate the performance of the robust TF Wiener filter $\mathbf{H}_{\mathrm{R}}$, we used $K = 4$ rectangular TF regions $\mathcal{R}_i$ to define p-point uncertainty classes $\mathcal{S}$ and $\mathcal{N}$ as described above. The regional energies s_i and n_i were obtained by integrating the nominal Wigner-Ville spectra in Fig. 12.4.2(a),(b) over the TF regions $\mathcal{R}_i$. The Weyl symbol of the robust TF Wiener filter $\mathbf{H}_{\mathrm{R}}$ in (12.4.8) is shown in Fig. 12.4.2(f). Fig. 12.4.3 compares the nominal and worst-case performance of the Wiener filter $\mathbf{H}_{\mathrm{W}}$ (designed for the nominal Wigner-Ville spectra in Fig. 12.4.2(a),(b)) with the performance of the robust TF Wiener filter $\mathbf{H}_{\mathrm{R}}$. It is seen that $\mathbf{H}_{\mathrm{R}}$ achieves a substantial performance improvement over $\mathbf{H}_{\mathrm{W}}$ at worst-case operating conditions with only a slight performance loss at nominal operating conditions.

12.4.2 Nonstationary Signal Detection

Next, we consider the discrimination of two nonstationary, zero-mean, Gaussian random signals $x_0(t)$ and $x_1(t)$. The hypotheses are

$$H_0: \; x(t) = x_0(t) \qquad \text{vs.} \qquad H_1: \; x(t) = x_1(t)\,.$$

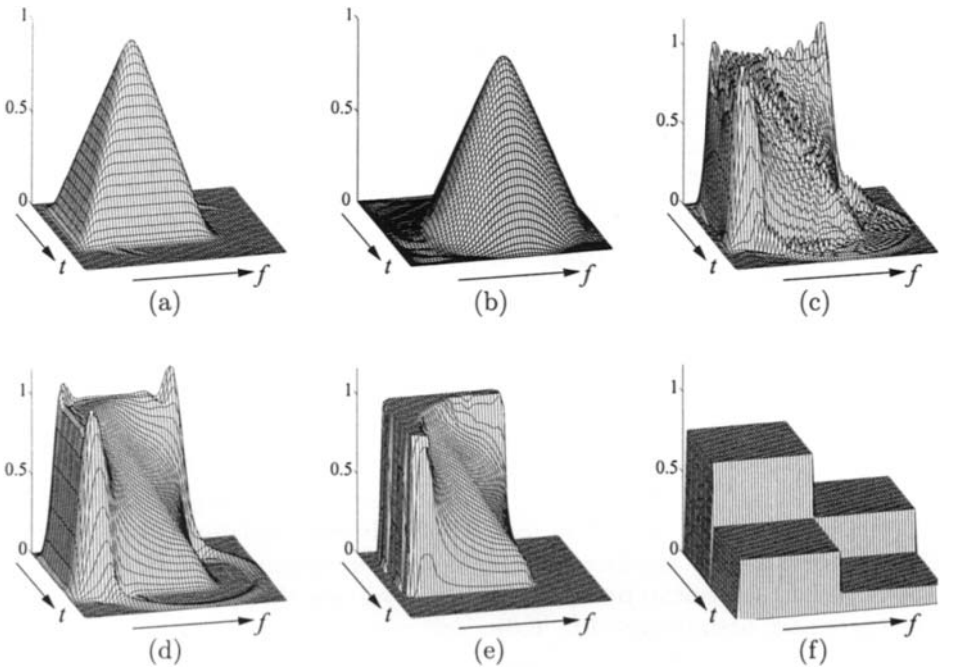


Fig. 12.4.2: TF representations of signal and noise statistics and of various Wiener-type filters: (a) Wigner-Ville spectrum of $s(t)$, (b) Wigner-Ville spectrum of $n(t)$, (c) Weyl symbol of Wiener filter $\mathbf{H}_{\mathrm{W}}$, (d) Weyl symbol of underspread part $\mathbf{H}_{\mathrm{W}}^{u}$ of $\mathbf{H}_{\mathrm{W}}$, (e) Weyl symbol of TF pseudo-Wiener filter $\widetilde{\mathbf{H}}_{\mathrm{W}}$, (f) Weyl symbol of robust TF Wiener filter $\mathbf{H}_{\mathrm{R}}$. The time duration is 128 samples; normalized frequency ranges from $-1/4$ to $1/4$.

The optimal detector. The optimal likelihood ratio detector [2, 3] calculates a quadratic form of the observed signal $x(t)$,

$$\Lambda(x) = \langle \mathbf{H}_{\mathrm{L}}\, x, x\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_{\mathrm{L}}(t,t')\, x(t')\, x^{*}(t)\, dt\, dt', \tag{12.4.9}$$

with the operator (LTV system) $\mathbf{H}_{\mathrm{L}}$ given by

$$\mathbf{H}_{\mathrm{L}} = \mathbf{R}_{x_0}^{-1} - \mathbf{R}_{x_1}^{-1} = \mathbf{R}_{x_0}^{-1}\big(\mathbf{R}_{x_1} - \mathbf{R}_{x_0}\big)\mathbf{R}_{x_1}^{-1}. \tag{12.4.10}$$

The test statistic $\Lambda(x)$ is then compared to a threshold to decide whether H_0 or H_1 is in force. For stationary processes, $\Lambda(x)$ can be expressed in terms of the Fourier transform $X(f)$ of $x(t)$ and the power spectral densities of $x_0(t)$ and $x_1(t)$ as

$$\Lambda(x) = \int_{-\infty}^{\infty} |X(f)|^2\, H_{\mathrm{L}}(f)\, df, \qquad \text{with } H_{\mathrm{L}}(f) = \frac{P_{x_1}(f) - P_{x_0}(f)}{P_{x_0}(f)\, P_{x_1}(f)}. \tag{12.4.11}$$

This frequency-domain expression involves simple products and reciprocals of functions (instead of operator products and inverses as in (12.4.10)) and thus allows a

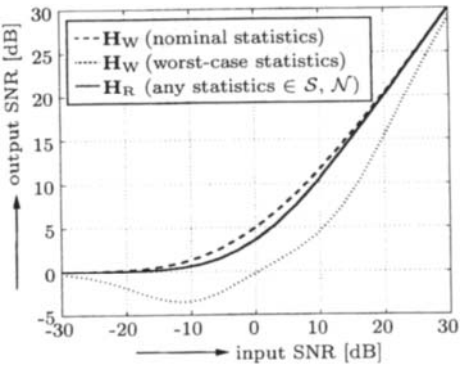


Fig. 12.4.3: Comparison of the performance (output SNR vs. input SNR) of the ordinary Wiener filter $\mathbf{H}_W$ and the robust TF Wiener filter $\mathbf{H}_R$ for various operating conditions.

simple interpretation and design of likelihood ratio detectors in the stationary case.

TF formulation of nonstationary detectors. It is known [9] that the quadratic test statistic in (12.4.9) can be rewritten as

$$\Lambda(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x^{(\alpha)}(t,f)\, L_{\mathbf{H}_L}^{(\alpha)*}(t,f)\; dt\,df\,, \tag{12.4.12}$$

where

$$W_x^{(\alpha)}(t,f) = \int_{-\infty}^{\infty} x\Big(t+\Big(\frac{1}{2}-\alpha\Big)\tau\Big)\, x^*\Big(t-\Big(\frac{1}{2}+\alpha\Big)\tau\Big)\, e^{-j2\pi f\tau} d\tau$$

is the generalized Wigner distribution [10] of the observed signal $x(t)$. Thus, $\Lambda(x)$ can be interpreted as a weighted integral of $W_x^{(\alpha)}(t,f)$, with the TF weight function being the conjugate of the GWS of the operator $\mathbf{H}_L$.

In analogy to Section 12.4.1, a simplified approximate TF formulation of $\Lambda(x)$ exists for *jointly underspread* processes $x_0(t)$, $x_1(t)$. Here, the operator $\mathbf{H}_L$ can be written as the sum of an overspread component whose effect is negligible and an underspread component, denoted $\mathbf{H}_L^u$, whose GWS can be approximated as [11]

$$L_{\mathbf{H}_L^u}^{(\alpha)}(t,f) \approx \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)}\,. \tag{12.4.13}$$

Substituting this approximation of $L_{\mathbf{H}_L^u}^{(\alpha)}(t,f)$ for $L_{\mathbf{H}_L}^{(\alpha)}(t,f)$ in (12.4.12), we obtain the following approximate TF formulation of our test statistic,

$$\Lambda(x) \approx \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x^{(\alpha)}(t,f) \left[\frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\,\overline{W}_{x_1}^{(\alpha)}(t,f)}\right]^* dt\,df. \tag{12.4.14}$$

This TF formulation extends (12.4.11) to the underspread nonstationary case and allows an intuitively appealing TF interpretation that is analogous to the one given in Section 12.4.1 in the context of the approximation (12.4.5).

TF design of nonstationary detectors. The TF formulation (12.4.14) suggests a simple *TF design* of nonstationary detectors. In analogy to (12.4.12), we define the test statistic

$$\tilde{\Lambda}(x) \triangleq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x^{(\alpha)}(t,f)\, L_{\widetilde{\mathbf{H}}_{\mathrm{L}}}^{(\alpha)*}(t,f)\, dt\, df\,,$$

where the operator (LTV system) $\widetilde{\mathbf{H}}_{\mathrm{L}}$ is defined by setting its GWS equal to the right-hand side of (12.4.13) [11]:

$$L_{\widetilde{\mathbf{H}}_{\mathrm{L}}}^{(\alpha)}(t,f) \triangleq \frac{\overline{W}_{x_1}^{(\alpha)}(t,f) - \overline{W}_{x_0}^{(\alpha)}(t,f)}{\overline{W}_{x_0}^{(\alpha)}(t,f)\, \overline{W}_{x_1}^{(\alpha)}(t,f)}\,.$$

For jointly underspread processes $x_0(t)$, $x_1(t)$ where (12.4.13) is a good approximation, $\widetilde{\mathbf{H}}_{\mathrm{L}}$ will closely approximate (the underspread part of) $\mathbf{H}_{\mathrm{L}}$, and thus the performance of the TF designed detector $\tilde{\Lambda}(x)$ will be similar to that of the optimal likelihood ratio detector $\Lambda(x)$. For processes $x_0(t)$, $x_1(t)$ that are *not* jointly underspread, however, $\tilde{\Lambda}(x)$ must be expected to perform poorly.

Whereas the detector $\tilde{\Lambda}(x)$ is designed in the TF domain, it can be implemented in the time domain in complete analogy to (12.4.9). The impulse response $\tilde{h}_{\mathrm{L}}(t,t')$ of $\widetilde{\mathbf{H}}_{\mathrm{L}}$ can be obtained from $L_{\widetilde{\mathbf{H}}_{\mathrm{L}}}^{(\alpha)}(t,f)$ by an inverse Weyl transformation (cf. (12.4.7)). An efficient implementation of the TF detector $\tilde{\Lambda}(x)$ that uses the multiwindow short-time Fourier transform is discussed in [12].

Compared to the likelihood ratio detector $\Lambda(x)$, the TF designed detector $\tilde{\Lambda}(x)$ is practically advantageous because the statistical *a priori* knowledge required for its design is formulated in the intuitively accessible TF domain, and because its design is less computationally intensive and more stable since operator inversions are replaced by pointwise divisions of functions. These advantages are analogous to the advantages of the TF pseudo-Wiener filter discussed in Section 12.4.1. Minimax robust detectors that are analogous to the minimax robust Wiener filter in Section 12.4.1 are reported in [13].

Simulation results. We first consider $x_0(t) = n(t)$ and $x_1(t) = s(t) + n(t)$, where signal $s(t)$ and noise $n(t)$ are jointly underspread, uncorrelated, zero-mean, Gaussian processes with Wigner-Ville spectra as shown in Fig. 12.4.4(a), (b). From Fig. 12.4.4(c), (d), we verify that the Weyl symbols of the optimal operator $\mathbf{H}_{\mathrm{L}}$ and the TF designed operator $\widetilde{\mathbf{H}}_{\mathrm{L}}$ are effectively identical. Also, Fig. 12.4.4(e), (f) shows that the performance of the TF designed detector $\tilde{\Lambda}(x)$ closely approximates that of the likelihood ratio detector $\Lambda(x)$.

Our next example concerns the detection of knocking combustions in car engines (see Article 15.2 and refs. [12,14,15] for background and details). Here, $x_0(t)$ is the nonknocking signal and $x_1(t)$ is the knocking signal. Estimates of the correlations of $x_0(t)$ and $x_1(t)$ were computed from a set of labeled training data,[2] and estimates

[2]We are grateful to S. Carstens-Behrens, M. Wagner, and J. F. Böhme for illuminating discussions and for providing us with the labeled car engine data.

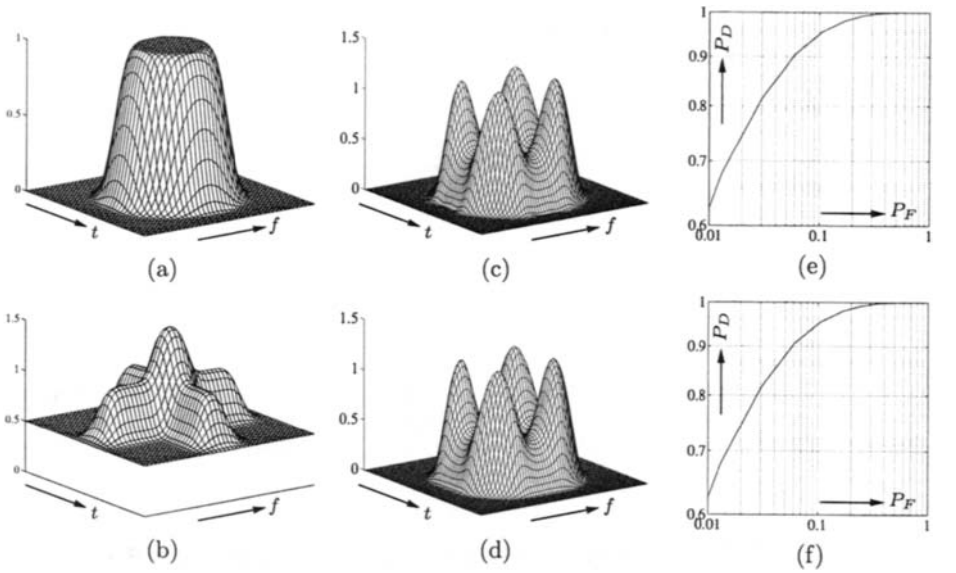


Fig. 12.4.4: Comparison of likelihood ratio detector $\Lambda(x)$ and TF designed detector $\tilde{\Lambda}(x)$: (a) Wigner-Ville spectrum of $s(t)$, (b) Wigner-Ville spectrum of $n(t)$, (c) Weyl symbol of $\mathbf{H}_{\mathsf{L}}$, (d) Weyl symbol of $\tilde{\mathbf{H}}_{\mathsf{L}}$, (e) receiver operator characteristics (ROC) [2] of $\Lambda(x)$, (f) ROC of $\tilde{\Lambda}(x)$. The ROCs were obtained by Monte Carlo simulation. The time duration is 128 samples; normalized frequency ranges from $-1/4$ to $1/4$.

of the Wigner-Ville spectra (shown in Fig. 12.4.5(a),(b)) were derived according to (12.4.1). The likelihood ratio detector $\Lambda(x)$ and the TF designed detector $\tilde{\Lambda}(x)$ were constructed using these estimated statistics, and the performance of these detectors was analyzed by applying them to a different set of labeled data. It can be seen from Fig. 12.4.5(c) that the TF designed detector performs significantly better than the theoretically optimal likelihood ratio detector. This is due to numerical problems that occurred in the design of the likelihood ratio detector. Specifically, the estimated correlation matrices[3] were poorly conditioned. Despite the use of pseudo-inverses, the inversion of these matrices (which is required for the design of the likelihood ratio detector) could not be stabilized sufficiently. In contrast, the design of the TF detector merely involves a pointwise division of the estimated Wigner-Ville spectra. This is much less affected by numerical problems since divisions by near-to-zero values can easily be stabilized by means of a thresholding.

12.4.3 Summary and Conclusions

The generalized Wigner-Ville spectrum (GWVS) provides a natural extension of the power spectral density to underspread, nonstationary random processes. Similarly,

[3]In the discrete-time case, correlation operators are replaced by correlation matrices.

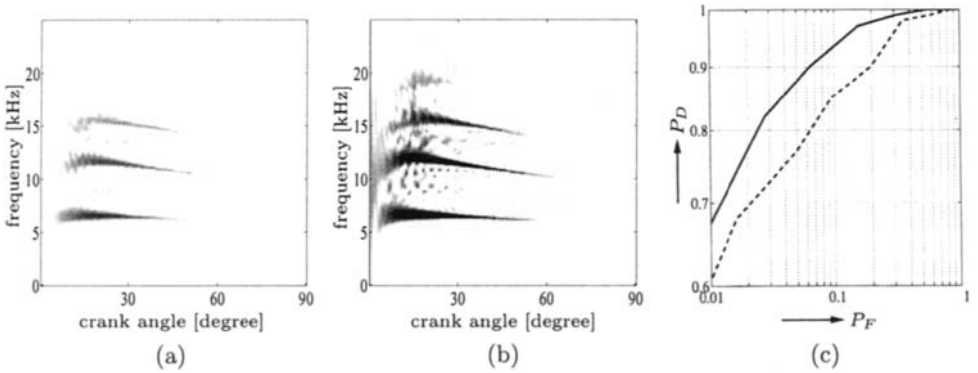


Fig. 12.4.5: Detection of knocking combustions: (a) Estimated Wigner-Ville spectrum of nonknocking combustion process $x_0(t)$, (b) estimated Wigner-Ville spectrum of knocking combustion process $x_1(t)$ (crank angle is proportional to time; signal length is 186 samples), (c) ROCs of likelihood ratio detector $\Lambda(x)$ (dashed line) and TF designed detector $\tilde{\Lambda}(x)$ (solid line).

the generalized Weyl symbol (GWS) provides a natural extension of the transfer function (frequency response) to underspread, time-varying linear systems. In this article, we have considered the application of the GWVS and GWS to the estimation and detection of underspread, nonstationary random processes. Using the GWVS and GWS, it was possible to extend classical stationary estimators and detectors to the nonstationary case in an intuitive manner. We note that the general approach—replacing the power spectral density with the GWVS and the transfer function with the GWS—is applicable to other problems of statistical signal processing as well, as long as the nonstationary processes involved are (jointly) underspread.

Further time-frequency methods for nonstationary signal estimation and detection are discussed in Articles 8.3, 9.2, 12.1, and 15.2 as well as in refs. [15–18].

References

[1] 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 Processing*, vol. 48, pp. 1417–1432, May 2000.

[2] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer, 1988.

[3] H. L. L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. I: "Detection, Estimation, and Linear Modulation Theory". New York: Wiley, 1968. Reprinted 2001.

[4] W. Kozek, H. G. Feichtinger, and J. Scharinger, "Matched multiwindow methods for the estimation and filtering of nonstationary processes," in *Proc. IEEE Internat. Symp. on Circuits and Systems (ISCAS 96)*, vol. 2, pp. 509–512, Atlanta, GA, 12–15 May 1996.

[5] G. Matz and F. Hlawatsch, "Minimax robust time-frequency filters for nonstationary signal estimation," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'99)*, pp. 1333–1336, Phoenix, AZ, 15–19 March 1999.

[6] G. Matz and F. Hlawatsch, "Minimax robust nonstationary signal estimation based on a p-point uncertainty model," *J. Franklin Institute*, vol. 337, pp. 403–419, July 2000.

[7] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, March 1985.

[8] G. Matz, F. Hlawatsch, and A. Raidl, "Signal-adaptive robust time-varying Wiener filters: Best subspace selection and statistical analysis," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, pp. 3945–3948, Salt Lake City, UT, 7–11 May 2001.

[9] P. Flandrin, "A time-frequency formulation of optimum detection," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 36, pp. 1377–1384, September 1988.

[10] 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.), pp. 59–133, Amsterdam: Elsevier, 1997.

[11] G. Matz and F. Hlawatsch, "Time-frequency formulation and design of optimal detectors," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 213–216, Paris, 18–21 June 1996.

[12] G. Matz and F. Hlawatsch, "Time-frequency methods for signal detection with application to the detection of knock in car engines," in *Proc. Ninth IEEE Workshop on Statistical Signal and Array Processing (SSAP-98)*, pp. 196–199, Portland, OR, 14–16 September 1998.

[13] G. Matz and A. Raidl, "Robust detection of nonstationary random signals belonging to p-point uncertainty classes," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'03)*, pp. 641–644, Hong Kong, scheduled 6–10 April 2003.

[14] S. Carstens-Behrens, M. Wagner, and J. F. Böhme, "Detection of multiple resonances in noise," *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, vol. 52, no. 5, pp. 285–292, 1998.

[15] G. Matz and F. Hlawatsch, "Time-frequency subspace detectors and application to knock detection," *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, vol. 53, pp. 379–385, December 1999.

[16] H. A. Khan and L. F. Chaparro, "Nonstationary Wiener filtering based on evolutionary spectral theory," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. 5, pp. 3677–3680, Munich, 21–24 April 1997.

[17] A. M. Sayeed and D. J. Jones, "Optimal detection using bilinear time-frequency and time-scale representations," *IEEE Trans. Signal Processing*, vol. 43, pp. 2872–2883, December 1995.

[18] J. A. Sills and E. W. Kamen, "Time-varying matched filters," *Circuits, Systems, & Signal Processing*, vol. 15, no. 5, pp. 609–630, 1996.

This Page Intentionally Left Blank

Part V

Engineering Applications

This Page Intentionally Left Blank

Chapter 13

Time-Frequency Methods in Communications

The wide range of potential applications of time-frequency methods made them an important tool in most fields of science and engineering. A large number of approaches exist, depending on the application considered. Key time-frequency methodologies are presented on specific applications in this Part V of the book and illustrated using selected examples. Telecommunications is one of the key industries where time-frequency methods are already playing an important role. The topic is represented by four articles selected for this chapter, complemented by articles in other chapters such as Articles 8.5 and 11.3.

Due to possible hostile jamming, broadband communication platforms use spread spectrum technology where interference protection is achieved by interference excision. By distributing the signature of received data over a time-frequency region, it is possible to attenuate strong interferences (Article 13.1). Linear dispersion in wireless communication channels distorts the transmitted signal in both time and frequency. This is accounted for by a time-frequency scattering function. In CDMA systems, fading and multi-access interference can be dealt with using time-frequency processing. A time-frequency RAKE receiver is described which accounts for both spectral and temporal channel variations (13.2). Eigenfunctions of linear systems can be modeled by signals with a time-frequency distribution well localized in the time-frequency plane. The knowledge of the eigenfunctions of time-varying transfer functions enables us to optimize the transmission strategy and take advantage of the channel dispersive properties (13.3). Detection and parameter estimation of chirps in communication systems may be implemented using the fractional FT (13.4).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

13.1 TIME-FREQUENCY INTERFERENCE MITIGATION IN SPREAD SPECTRUM COMMUNICATION SYSTEMS[0]

13.1.1 Spread-Spectrum Systems and Interference

In modern broadband communication systems, mitigation of correlated interference is an important aspect of system performance. Benign jamming sources exist ubiquitously in the transmission channel of multiple access systems like cellular telephony and wireless networks. Hostile jamming is certainly a significant issue in military communication systems. For these and many other reasons, broadband communication platforms employ spread spectrum (SS) technology, in which [1]

(1) The signal occupies a bandwidth much in excess of the minimum bandwidth necessary to send the information.

(2) Spreading is accomplished by means of spreading signal, which is often called a code signal that is independent of the data.

(3) At the receiver, despreading for the recovering the original data is accomplished by the correlation of the received spread signal with a synchronized replica of the spreading signal.

A commonly employed SS technique called direct sequence (DS) is to superimpose a pseudorandom noise (PN) sequence on the data bits. This effectively widens the signal bandwidth by a factor proportional to the ratio of the rate of the PN sequence divided by the data rate. The advantage of this spreading is that the wider bandwidth essentially marginalizes narrowband interference sources so they have a smaller effect on overall system performance.

At the receiver, the cross correlation with the replica of the PN sequence transfers the information signal back to its original bandwidth while reducing the level of the narrowband interference by spreading it across the bandwidth occupied by the PN sequence. Since the availability of the code at the receiver enables despreading and recovery of data while spreading and suppression of interference, any level of interference rejection can be achieved by using sufficient processing gain. This, however, may entail increasing the bandwidth of the transmitted signal beyond the limits of the available frequency spectrum. Therefore, signal processing techniques have been used in conjunction with the DS spread spectrum receiver to augment the processing gain, permitting greater interference protection without an increase in bandwidth [2].

Typically, interference excision in DSSS systems is performed in the following domains [3].

Frequency Domain: the FFT of the data over the information bit is weighted by appropriate values and then transformed back to the time domain. This is an effective method for stationary narrowband interference. Sidelobes may present a problem in removing the interference without losing some of the signal energy.

Time Domain: this includes clipping or gating the high energy regions. It also

[0]Authors: **Moeness G. Amin**, Villanova University, Villanova, PA, and **Alan R. Lindsey**, Air Force Research Laboratory, Rome, NY. Reviewers: S. Batalama and D. L. Jones.

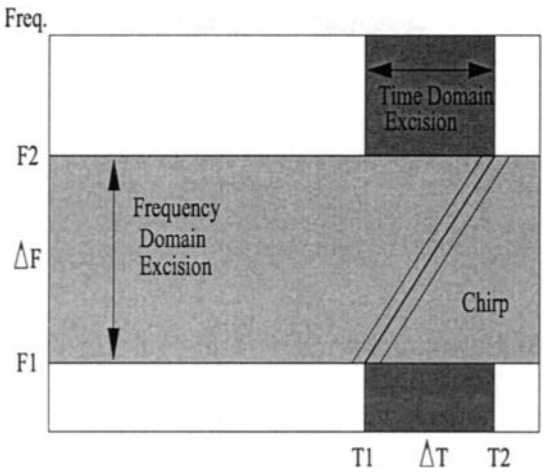


Fig. 13.1.1: Excision methods for nonstationary signals.

includes Wiener filtering, adaptive linear predictors, and smoothers. Tracking is highly dependent on the signal-to-noise ratio (SNR) and often fails under rapidly time-varying interference.

Time and Frequency Domains: a transversal filter is designed from the spectral information of the data. Spectral estimation methods combined with open-loop adaptive filtering have been shown to suffer from the same drawbacks as frequency-domain techniques.

Wavelet/Gabor domain: the discrete wavelet transform (DWT) or the Gabor transform is applied to the data, and the coefficients of high energy are removed prior to the inverse transform. The DWT is appropriate for cases of pulse jamming or interference with burst characteristics. The Gabor transform is an effective excision tool only when the interference is consistent with the corresponding tiling of the time-frequency $((t, f))$ plane. The same is true for the wavelet transform.

None of the above methods is capable of effectively incorporating the time-varying nature of the interference frequency characteristics. These methods all suffer from their lack of intelligence about the other domain(s) and therefore are limited in their results and their applicability. To illustrate, Fig. 13.1.1 shows that most frequency-and time-domain excisions, in essence, respectively, remove all desired signal information over the frequency band ΔF and time duration ΔT . As such, in the case of time-varying interferences, frequency-domain methods ignore the fact that only few frequency bins are contaminated by the jammer at a given time. Dually, time domain excision techniques, do not account for the cases where only few time samples are contaminated by the jammer for a given frequency. Applying either method will eliminate the interference but at the cost of unnecessarily reducing the desired signal energy.

The above example clearly demonstrates that nonstationary interferers, which

have model parameters that change with time, are particularly troublesome due to the inability of single-domain mitigation algorithms to adequately ameliorate their effects. In this challenging situation, and others like it, joint time-frequency techniques can provide significant performance gains, since the instantaneous frequency (IF), the instantaneous bandwidth, and the energy measurement, in addition to myriad other parameters, are available. The objective is then to estimate the time-frequency signature of the received data using time-frequency distributions (TFDs), attenuating the received signal in those *time-frequency regions* that contain strong interference. This is depicted by the region in between the fine lines in Fig. 13.1.1.

13.1.2 Typical Signal Model

The signal model $r(t)$ arriving at a receiver is in the form

$$r(t) = s(t) + n(t) + j(t) \tag{13.1.1}$$

which is composed of the desired spectrally spread signal $s(t)$, the composite additive jamming signal $j(t)$ which may be made up of several different jammers from various sources, and the obligatory uncorrelated thermal noise, $n(t)$, commonly assumed to be white with a Gaussian distribution.

13.1.3 A Time-Frequency Distribution Perspective

A straightforward application of time-frequency methods in the underlying problem is the design of adaptive notch filters based on the instantaneous frequency of the jamming signal. Amin pioneered this approach in [3], using the assumption that $n(t)$ and $s(t)$ are spread across the (t, f) plane and without features. In this model, the correlated features in $j(t)$ rise conspicuously in the (t, f) distribution performed at the input of the receiver, and subsequent instantaneous frequency information allows for the design of a strong notch filter that effectively excises only the portion of the (t, f) spectrum contaminated by the jammer. The process is illustrated in Fig. 13.1.2. The notch filter developed from the TFD can be very short having been shown to be very effective with as few as three or five coefficients.

Wang and Amin [4] considered the performance analysis of this system using a general class of multiple-zero FIR excision filters showing the dependence of the bit error rate (BER) on the filter order and its group delay. The effect of inaccuracies in the jammer IF on receiver performance was also considered, as a function of the filter notch bandwidth. Closed form approximations for signal to interference plus noise ratio (SINR) at the receiver are given for the various cases. The general form of the receiver SINR is given by

$$\text{SINR} = \frac{E^2[D_s]}{\text{Var}(D_s) + \text{Var}(D_j) + \text{Var}(D_n)} \tag{13.1.2}$$

where the three quantities in the denominator represent the variances of the decision variable, due to the SS signal, the jammer, and the noise respectively. $E[\cdot]$ is the

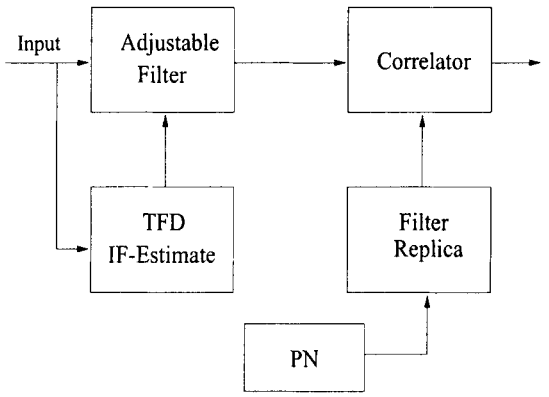


Fig. 13.1.2: TFD Excision System of Amin *et al.*

expectation operator, and with typical independent and uncorrelated characteristics of jammers and noise, it only leads to non-zero values when applied to the desired signal . If the jammer is fully excised then $\mathrm{Var}(D_j) = 0$. It is clear from the above equation that one of the main goals of the excision process should be reducing the self-noise represented by the term $\mathrm{Var}(D_s)$.

One of the drawbacks to the notch filter approach in [3] is the infinite notch depth due to the placement of the filter zeros. The effect is a "self-noise" inflicted on the received signal by the action of the filter on the PN sequence underlying the spread information signal. This problem led to the design of an open-loop filter with adjustable notch depth based on the jammer energy. The notch depth is determined by a variable embedded in the filter coefficients chosen as the solution to an optimization problem which maximizes receiver SINR cost function given in (13.1.2). The TFD is necessary for this work, even for single component signals, because simple IF estimators do not provide energy information. Amin, Wang, and Lindsey accomplished this work in [5], incorporating a "depth factor" into the analysis and redeveloping all the SINR calculations. The result was significant improvement in SINR, especially at mid-range jammer-to-signal ratios (JSRs), typically around 0 to 20 dB.

Barbarossa and Scaglione [6] proposed a two-step procedure based on dechirping techniques commonly applied in radar algorithms (see Article 13.3). In the first step the time varying jammer is converted to a fixed frequency sinusoid eliminated by time invariant filters. The process is reversed In the second step and the jammer-free signal is multiplied by the jammer TF signature to restore the DSSS signal and noise characteristics which have been strongly impacted in the first phase. Comparison of this technique with time-varying excision filters is yet to be conducted in terms of computational complexity and robustness to IF estimation errors.

Synthesis of the interfering signal from information available at the receiver followed by direct subtraction [7] is an approach put forth by Lach, Amin, and Lindsey.

In many situations, it is possible to make assumptions about certain key parameters which describe fully the nature and composition of a jammer, and then utilize signal processing to extract these parameters from the received signal. Once the parameters are determined, a replica of the jammer can be derived and subtracted from the incoming signal to produce an essentially jammer free channel. However, for this process to work, a jammer of constant modulus and polynomial phase is required. These parameters are extracted at the receiver via time-frequency distribution where the optimally matched signal in a least squares sense is constructed. The jamming signal is projected to the constant modulus and then phase-matched. The result, especially in high SNR environments, is a signal matching the jammer in amplitude, frequency profile and phase. The last step of generating the difference signal, which is theoretically also the desired signal after the subtraction of the jammer, is straightforward.

To overcome the drawbacks of the least squares synthesis methods Amin, Ramineni and Lindsey [8] proposed a projection filter approach in which the nonstationary interference subspace is constructed from its TF signature. Since the signal space at the receiver is not specifically mandated, it can be rotated such that a single jammer becomes one of the basis functions. In this way, the jammer subspace is one dimensional and its orthogonal subspace is jammer-free. A projection of the received signal onto the orthogonal subspace accomplishes interference excision with a minimal message degradation. The projection filtering methods compare favorably over the previous notch filtering systems.

13.1.4 Example

At this point, in order to further illustrate these excision methods, the work in [8] will be detailed since it includes comparisons between the two most prominent techniques based on time-frequency distributions currently being studied - notch filtering and projection filtering. The signal model is, as expected, given by (13.1.1), and the major theme of the work is to annihilate interference via projection of the received signal onto a "jammer-free" subspace generated from the estimated jammer characteristics. Fig. 13.1.3, reprinted from [8], clearly illustrates the trade-offs between projection and notch filtering based on the JSR. In the legend, the variable a represents the adaptation parameter for the notch filtering scheme and N represents the block size, in samples, for a 128 sample bit duration in the projection method. Thus, N=128 means no block processing and N=2 corresponds to 64 blocks per bit being processed for projection. Since the projection and non-adaptive notch filter techniques are assumed to completely annihilate the jammer, their performance is decoupled from the jammer power, and therefore correctly indicate constant SINR across the graph. The dashed line representing the notch filter with a=0 is really indicating no filtering at all, since the adaptation parameter controls the depth of the notch.

It is evident from Fig. 13.1.3 that without adaptation a crossover point occurs around 2 dB where filtering with an infinitely deep notch is advantageous. Thus

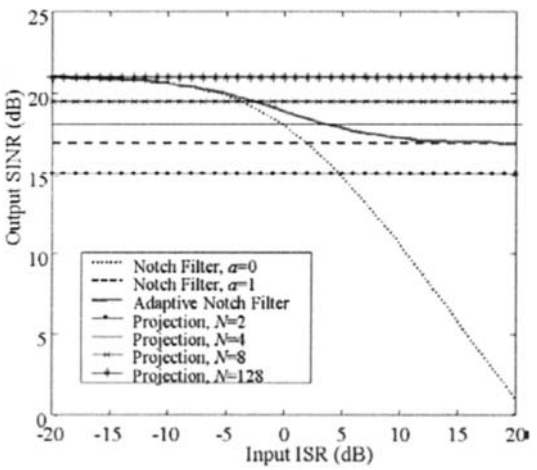


Fig. 13.1.3: Comparison between projection and notch filtering excision methods.

when jammer power exceeds this point, presumably a user would flip a switch to turn on the excision subsystem. However, with adaptation this process happens automatically, while giving superior performance in the midrange. For the projection technique, the block size determines receiver performance conspicuously (ceteris paribus). Most important to note, however, is the superior performance of projection over all methods when the block size is equal to the bit duration, i.e. no block processing. It is feasible that computational complexity may warrant a trade-off between SINR and block size, in which case a hybrid implementation may be of benefit - one that automatically switches between adaptive notch filtering and projection depending on the desired SINR. In any case, this example illustrates the parameters involved in the design of modern excision filters for nonstationary jammers.

13.1.5 Summary and Conclusions

The prime objective of interference suppression in broadband signal platforms is to cancel the interference with minimum distortions of the desired signal. Time-frequency signal representation provides the mechanism to achieve that objective for a large class of nonstationary interference signals. The signal localization in the time-frequency domain allows signal processing, acting on information of the instantaneous frequency and bandwidth, to play an effective role in enhancing the receiver performance and improving the bit error rates over existing techniques that deal with only the time-domain or the frequency-domain.

We have presented the problem of interference excision from a quadratic time-frequency perspective. Joint time-frequency and time-scale linear transforms have also been successfully applied to suppress nonstationary jammers. Further citations in this area can be found in the references of [2] and [9].

References

[1] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications* (3 vols.). Rockville, MD: Computer Science Press, 1985.

[2] M. G. Amin and A. N. Akansu, "Time-frequency for interference excision in spread-spectrum communications," in "Highlights of Signal Processing for Communications", *IEEE Signal Processing Magazine* (G. B. Giannakis, ed.), vol. 16(2), IEEE, March 1999.

[3] M. G. Amin, "Interference mitigation in spread-spectrum communication systems using time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 45, pp. 90–102, January 1997.

[4] C. Wang and M. G. Amin, "Performance analysis of instantaneous frequency based interference excision techniques in spread spectrum communications," *IEEE Trans. Signal Processing*, vol. 46, pp. 70–83, January 1998.

[5] M. G. Amin, C. Wang, and A. R. Lindsey, "Optimum interference excision in spread-spectrum communications using open-loop adaptive filters," *IEEE Trans. Signal Processing*, vol. 47, pp. 1966–1976, July 1999.

[6] S. Barbarossa and A. Scaglione, "Adaptive time-varying cancellations of wideband interferences in spread-spectrum communications based on time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 47, pp. 957–965, April 1999.

[7] S. Lach, M. G. Amin, and A. R. Lindsey, "Broadband nonstationary interference excision in spread-spectrum communications using time-frequency synthesis techniques," *IEEE J. on Selected Areas in Communications*, vol. 17, pp. 704–714, April 1999.

[8] R. S. Ramineni, M. G. Amin, and A. R. Lindsey, "Performance analysis of subspace projection techniques for interference excision in DSSS communications," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, vol. 5, pp. 2825–2828, Istanbul, 5–9 June 2000.

[9] G. J. Saulnier, M. J. Medley, and P. K. Das, "Wavelets and filter banks in spread spectrum communication systems," in *Subband and Wavelet Transforms: Design and Applications* (A. N. Akansu and M. J. T. Smith, eds.), ch. 10, pp. 309–346, Norwell, MA: Kluwer, 1996.

13.2 COMMUNICATION OVER LINEAR DISPERSIVE CHANNELS: A TIME-FREQUENCY PERSPECTIVE[0]

13.2.1 Linear Dispersive Channels

Linear dispersive channels are encountered in many applications, including mobile wireless communications and underwater acoustical communications [1]. In general, such channels disperse the transmitted signal in both time and frequency. Without proper system design, the effects of dispersion, such as signal fading, can significantly limit system performance. However, appropriate signaling and reception can significantly mitigate the effects of dispersion. Accurate modeling of channel characteristics is critical in this regard.

The effect of the channel on the transmitted signal intimately depends on the bandwidth-duration product (TBP) of the signaling waveforms. Narrowband signaling schemes with TBP ≈ 1 typically suffer from significant intersymbol interference (ISI) requiring sequence (Viterbi) decoding at the receiver. Linear equalizers constitute a sub-optimal low-complexity solution for narrowband systems. In contrast, wideband signaling schemes with TBP $\gg 1$ can significantly mitigate ISI and side-step the requirement for equalization. In fact, wideband systems can actually exploit the dispersion effects for improved performance. Time-frequency representations and methods provide useful insights in the analysis and design of such wideband communication systems. The goal of this article is to highlight the key aspects of this time-frequency perspective on wideband communication over linear dispersive channels. We develop the concepts in the context of spread-spectrum code division multiple access (CDMA) systems. However, the general ideas hold in other applications as well, including underwater acoustical communications.

The next section describes a canonical time-frequency model for wideband dispersive channels that lays the foundation of the time-frequency perspective. The model is used in Section 13.2.3 to discuss various key aspects of wireless CDMA communication system design.

13.2.2 Time-Frequency Model for Dispersive Channels

Let $x(t)$ denote the transmitted complex baseband signal. The received signal after passing through a linear dispersive channel can be generally expressed as [1, 2]

$$r(t) = s(t) + \epsilon(t) = \int_0^{T_m} h(t,\tau)x(t-\tau)d\tau + \epsilon(t), \qquad (13.2.1)$$

where $h(t,\tau)$ denotes the time-varying channel impulse response, $s(t)$ is the complex baseband signal at the output of the channel, and $\epsilon(t)$ is additive white Gaussian noise (AWGN) with power spectral density σ_ϵ^2. The maximum delay produced

[0]Author: **Akbar M. Sayeed**, Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706, USA (akbar@engr.wisc.edu). Reviewers: S. Barbarossa and M. G. Amin.

by the channel, T_m, is called the *multipath spread* of the channel. The following equivalent channel representation is particularly relevant from a time-frequency perspective

$$s(t) = \int_0^{T_m} \int_{-B_d}^{B_d} H(\nu, \tau) x(t-\tau) e^{j2\pi\nu t} d\nu d\tau, \tag{13.2.2}$$

where $H(\nu, \tau) = \int h(t, \tau) e^{-j2\pi\nu t} dt$ is the time-frequency *spreading function* of the channel [1, 2]. The variable τ corresponds to temporal (multipath) spreading produced by the channel and the variable ν corresponds to the spectral (Doppler) spreading produced by the channel. The maximum Doppler frequency produced by the channel, B_d, is called the *Doppler spread* of the channel. The representation (13.2.2) states that the output signal $s(t)$ is a linear combination of time- and frequency-shifted versions of the transmitted signal $x(t)$.

Canonical Time-Frequency Channel Model

The channel model (13.2.2) in terms of a continuum of scatterers is difficult to incorporate in receiver design. The finite duration T and (essentially) finite one-sided bandwidth B of the signaling waveform $x(t)$ can be exploited to derive a canonical time-frequency channel model that greatly facilitates system design. The canonical model asserts the following channel representation [2–4]

$$s(t) \approx \sum_{l=0}^{L} \sum_{m=-M}^{M} h_{ml} x_{ml}(t) = \sum_{l=0}^{L} \sum_{m=-M}^{M} \widetilde{H}\left(\frac{m}{T}, \frac{l}{B}\right) x\left(t - \frac{l}{B}\right) e^{j\frac{2\pi m t}{T}} \tag{13.2.3}$$

where $L = \lceil T_m B \rceil$, $M = \lceil T B_d \rceil$, and $\{h_{ml}\}$ are samples of a smoothed version, $\widetilde{H}(\nu, \tau)$, of the spreading function. The model (13.2.3) is a canonical uniform time-frequency (multipath-Doppler) sampling of the channel induced by the finite duration and bandwidth of $x(t)$. The sampling resolution in is $\Delta\tau = 1/B$ in time (multipath) and $\Delta\nu = 1/T$ in frequency (Doppler).

The canonical representation (13.2.3) states that the output signal $s(t)$ belongs to a subspace spanned by the basis functions

$$x_{ml}(t) = x\left(t - \frac{l}{B}\right) e^{j\frac{2\pi m t}{T}}, \quad l = 0, \cdots, L, \; m = -M, \cdots, M \tag{13.2.4}$$

that are discretely time-frequency shifted versions of the transmitted signal. The dimension of this (active) subspace is $N_a = (L+1)(2M+1) \approx 2TBT_mB_d + T_mB + 2TB_d + 1$ which is proportional to the TBP (TB) of the signaling waveform as well as the channel spread factor T_mB_d. Typical channels encountered in practice are *underspread*: $T_mB_d \ll 1$. While (13.2.4) provides a *fixed* basis for representing the received signal, an alternative signal representation in terms of approximate eigenfunctions of underspread channels is developed in Article 13.3 in this volume. However, the eigenfunctions are not fixed and need to be computed for each channel realization. Underspread channels imply that the output signal subspace is one-dimensional ($N_a = 1$) for narrowband ($TB = 1$) systems, whereas it is $N_a > 1$

dimensional for wideband ($TB \gg 1$) systems. As we will see, this increase in dimensionality due to wideband signaling plays an important role in improved communication over dispersive channels. Note that once the signaling waveform $x(t)$ and the channel spreads T_m and B_d are known at the receiver, all information about the channel is captured *linearly* via $\{h_{ml}\}$ in (13.2.3).

Statistical Channel Characterization

Communication system performance depends on statistical channel characteristics. A widely used model is the wide-sense stationary uncorrelated scattering (WSSUS) model in which $\{H(\nu,\tau)\}$ are modeled as uncorrelated Gaussian random variables [1,2]: $\mathcal{E}[H(\nu,\tau)H^*(\nu',\tau')] = \Psi(\nu,\tau)\delta(\tau-\tau')\delta(\nu-\nu')$ where $\Psi(\nu,\tau)$ is the time-frequency *scattering function* which quantifies the relative channel power at different delays and Doppler shifts. The support of $\Psi(\nu,\tau)$ is limited to the channel spreads: $[0,T_m] \times [-B_d,B_d]$. Under the WSSUS assumption, it can be shown that the channel coefficients $\{h_{ml}\}$ in the canonical time-frequency model (13.2.3) are approximately uncorrelated if $\Psi(\nu,\tau)$ is sufficiently smooth. For simplicity, we focus on Rayleigh fading WSSUS channels in which $\{h_{ml}\}$ are zero mean, uncorrelated Gaussian random variables. A more detailed characterization of randomly time-varying channels is developed in Article 9.5 in this volume.

13.2.3 Communication over Dispersive Channels

Signals with duration T and bandwidth B form a vector space of dimension $N_o \approx TB$. Spread-spectrum waveforms used in CDMA system take the form $q(t) = \sum_{n=0}^{N-1} c[n]v(t-nT_c)$ where $c[n]$ is the length-N spreading code associated with the waveform $q(t)$, T_c is the chip duration, and $N = T/T_c$ is the processing gain. For CDMA signals, $B \approx 1/T_c$ and $N \approx TB \approx N_o$. Given a particular spread-spectrum waveform $q(t)$, a complete basis for the signal space can be generated via distinct waveforms of the form $\{q_{ml}(t)\}$ as defined in (13.2.4). The two most significant factors affecting CDMA system performance are signal *fading* and *multiaccess interference* (MAI). Fading is due to the destructive combination of various time-frequency shifted signal copies and manifests itself large fluctuations in received signal power. MAI is caused by the multiple users simultaneously communicating over the channel and can drown the signal of the desired user.

We first discuss the role of time-frequency processing in single-user CDMA systems, highlighting the concept of diversity to combat fading. We then discuss the issue of interference suppression in multiuser systems via the notion of certain time-frequency subspaces. For simplicity, we focus on coherent receivers with binary phase shift keying (BPSK). We also make the assumption that $T_m \ll T$; that is, the ISI is negligible. Thus, symbol by symbol detection suffices.

13.2.3.1 Time-Frequency RAKE Receiver

Consider a single symbol transmission in a single-user system; that is, $x(t) = bq(t)$, where $q(t)$ is the wideband signaling waveform and $b \in \{-1,1\}$ is the transmitted

bit. The channel model (13.2.3) dictates the following front-end time-frequency correlation at the receiver

$$z_{ml} = \langle r, q_{ml} \rangle = \int_0^T r(t) q_{ml}^*(t) dt = b \sum_{l'=0}^{L} \sum_{m'=-M}^{M} \langle q_{m'l'}, q_{ml} \rangle h_{m'l'} + \langle \epsilon, q_{ml} \rangle. \tag{13.2.5}$$

The correlator outputs $\{z_{ml}\}$ form the N_a-dimensional representation of the noisy received signal with respect to the basis $\{q_{ml}(t)\}$. They are also the samples of the narrowband *cross-ambiguity function*, $A_{r,q}(\nu, \tau)$, between the received waveform $r(t)$ and the signaling waveform $q(t)$; that is, $z_{ml} = A_{r,q}\left(\frac{m}{T}, \frac{l}{B}\right)$ where $A_{r,q}(\nu, \tau) = \int r(t) q^*(t-\tau) e^{-j2\pi\nu t} dt$ [5]. The correlator outputs can be expressed in vector form as

$$\boldsymbol{z} = b\boldsymbol{Q}\boldsymbol{h} + \boldsymbol{w} \tag{13.2.6}$$

where $\boldsymbol{h}$ is the vector of channel coefficients, $\boldsymbol{w} \sim \mathcal{N}(0, \sigma_\epsilon^2 \boldsymbol{Q})$, and $\boldsymbol{Q}$ is an $N_a \times N_a$ matrix of correlation between the different basis waveforms consisting of the entries

$$\langle q_{m'l'}, q_{ml} \rangle = e^{j\frac{2\pi(m-m')l'}{TB}} A_{q,q}\left(\frac{m-m'}{T}, \frac{l-l'}{B}\right). \tag{13.2.7}$$

For typical spread-spectrum CDMA waveforms, and with chip-rate multipath sampling ($B = 1/T_c$ in (13.2.3)), the basis $\{q_{ml}(t)\}$ can be approximately orthogonal resulting in $\boldsymbol{Q} \approx \boldsymbol{I}$; that is, $\langle q_{m'l'}, q_{ml} \rangle \approx \delta_{l-l'}\delta_{m-m'}$. This implies that the *sampled* auto-ambiguity function of $q(t)$ in (13.2.7) is approximately a "thumbtack" function with peak at the origin.

Given estimates of the channel coefficients $\{\widehat{h}_{ml}\}$, the time-frequency correlator outputs $\{z_{ml}\}$ are coherently combined at the receiver to yield the final bit estimate

$$\begin{aligned} \hat{b} &= \operatorname{sign}\left\{ \operatorname{real}\left[\sum_{l=0}^{L} \sum_{m=-M}^{M} \widehat{h}_{ml}^* z_{ml} \right] \right\} \\ &= \operatorname{sign}\left\{ \operatorname{real}\left[\widehat{\boldsymbol{h}}^H \boldsymbol{z} \right] \right\} = \operatorname{sign}\left\{ \operatorname{real}\left[b\widehat{\boldsymbol{h}}^H \boldsymbol{Q}\boldsymbol{h} + \widehat{\boldsymbol{h}}^H \boldsymbol{w} \right] \right\} \end{aligned} \tag{13.2.8}$$

where the superscript H denotes complex conjugate transpose. The receiver (13.2.8) is a time-frequency generalization of the RAKE receiver used in conventional CDMA systems [4]. The time-frequency RAKE receiver performs joint multipath-Doppler combining to account for spectral and temporal channel variations within a symbol in contrast to the multipath combining in conventional RAKE receiver. Conditioned on the channel coefficients $\boldsymbol{h}$, and with perfect estimates ($\widehat{\boldsymbol{h}} = \boldsymbol{h}$), the bit-error-probability (BEP) of the receiver is given by $P_{e,\text{fading}}(\boldsymbol{h}) = Q(\text{SNR}(\boldsymbol{h}))$ where $\text{SNR}(\boldsymbol{h}) = 2\boldsymbol{h}^H \boldsymbol{Q}\boldsymbol{h}/\sigma_\epsilon^2$ and $Q(x) = \frac{1}{2\pi}\int_0^x e^{-t^2/2} dt$. The unconditioned BEP is given by averaging over the statistics of the channel coefficients: $P_{e,\text{fading}} = \mathcal{E}[P_{e,\text{fading}}(\boldsymbol{h})] = \mathcal{E}[Q(\text{SNR}(\boldsymbol{h}))]$.

13.2.3.2 Combating Fading with Time-Frequency Processing

The performance of the receiver (13.2.8) operating over an AWGN channel with the same *average* SNR as the fading channel is given by $P_{e,\text{AWGN}} = Q(\mathcal{E}[\text{SNR}(\boldsymbol{h})]) = Q(2\ \text{trace}(\boldsymbol{Q}\boldsymbol{R}_h)/\sigma_\epsilon^2)$ where trace(.) denotes the trace of a matrix and $\boldsymbol{R}_h = \mathcal{E}[\boldsymbol{h}\boldsymbol{h}^H]$ is the channel correlation matrix. We note that $P_{e,\text{fading}} \geq P_{e,\text{AWGN}}$ and the loss in performance under fading due to the fluctuations in $\text{SNR}(\boldsymbol{h})$ is quite significant. For example, achieving a BEP of 10^{-3} over a Rayleigh fading channel requires about 18 dB additional SNR compared to an AWGN channel with the same average SNR [1].

Diversity signaling and reception is a general technique for combating the effect of fading. The basic idea is to transmit the signal over multiple *independent* fading channels while keeping the total transmitted power constant. In the case of coherent reception, as the number of diversity channels increases, the performance of the diversity reception over a fading channel approaches that of an AWGN channel at the same average SNR [1].

The remarkable advantage of wideband signaling over WSSUS fading channels comes from the fact that the different delayed and Doppler shifted copies of the signaling waveform in (13.2.3) serve as N_a *virtual* diversity channels [3,4]. In essence, the receiver is able to resolve N_a weakly correlated time-frequency shifted copies ($\{q_{ml}(t)\}$) of the transmitted waveform that are linearly independent and serve as virtual diversity channels carrying the same information bit. Using the Karhunen-Loève expansion, the SNR can be written as $\text{SNR}(\boldsymbol{h}) = \sum_{n=1}^{N_a} \lambda_n \chi_n^2$ where $\{\chi_n^2\}$ are independent chi-squared random variables each with 2 degrees of freedom and $\{\lambda_n\}$ are the eigenvalues of $2\boldsymbol{Q}\boldsymbol{R}_h/\sigma_\epsilon^2$. Keeping the transmitted power constant, the law of large numbers dictates that $\text{SNR}(\boldsymbol{h}) \overset{N_a \to \infty}{\longrightarrow} \mathcal{E}[\text{SNR}(\boldsymbol{h})] = \text{trace}(2\boldsymbol{R}_h\boldsymbol{Q}/\sigma_\epsilon^2)$ and thus $P_{e,\text{fading}} \overset{N_a \to \infty}{\longrightarrow} P_{e,\text{AWGN}}$. Thus, as the number of time-frequency signal components increases in (13.2.3), the coherent receiver (13.2.8) completely combats the loss in performance due to fading. For given channel spread factor, $T_m B_d$, the larger the TBP of the signaling waveform, the higher the level of diversity. This advantage of wideband signaling is exploited in CDMA wireless communication systems via the use of direct-sequence spread-spectrum waveforms as well as in underwater acoustical communications via the use of linearly frequency-modulated waveforms. In particular, Doppler diversity, first introduced in [4], can be exploited in practice via time-selective signaling which increases the effective symbol duration [6].

13.2.3.3 Interference Suppression via Time-Frequency Subspaces

In CDMA systems, interference stems from multiple users simultaneously using the channel. In underwater acoustics, reverberation is the major source of interference. In either case, the time-frequency channel model (13.2.3) can be fruitfully exploited for effective interference suppression. The key idea is the notion of time-frequency subspaces associated with the desired signal and interference [7, 8]. We illustrate the concept in the context of CDMA systems.

Consider a CDMA communication system with K users. The k^{th} user is assigned a signaling waveform $q_k(t)$. According to the model (13.2.3), each user is associated

with a canonical basis $\{q_{k,ml}(t)\}$. For simplicity, we assume that all users are transmitting synchronously. The signal at any receiver can be represented as

$$r(t) \approx \sum_{k=1}^{K} s_k(t) + n(t) = \sum_{k=1}^{K} b_k \sum_{l=0}^{L_k} \sum_{m=-M_k}^{M_k} h_{k,ml} q_{k,ml}(t) + n(t). \tag{13.2.9}$$

Suppose that we are interested in decoding the signal of the first user. We consider *decentralized* detection in which only the code of the desired user is known at the receiver [8]. In this case, bit detection is accomplished by projecting onto the waveforms $\{q_{1,ml}\}$ associated with the first user. For a discussion of centralized detection which uses information about all users, we refer the reader to [7].

Recall that a subset of $\{q_{1,ml}\}$ provide a complete basis for the signal space. The key idea from the viewpoint of interference suppression is that of *active* and *inactive* subspaces, $\mathcal{S}_a$ and $\mathcal{S}_{ia}$, respectively. The active subspace, $\mathcal{S}_a$, is defined as the span of the basis elements that *lie within the the multipath-Doppler channel spread* and form the diversity channels: $\mathcal{S}_a = \text{span}\{q_{1,ml} : l = 0, 1, \cdots, L_1, \quad m = -M_1, \cdots, M_1\}$. The inactive subspace, $\mathcal{S}_{ia}$, is spanned by elements that *lie outside the channel spread*: $\mathcal{S}_{ia} = \text{span}\{q_{1,ml} : l \notin \{0, 1, \cdots, L_1\}, \quad m \notin \{-M_1, \cdots, M_1\}\}$.

The N_a-dimensional vector of correlator outputs corresponding to $\mathcal{S}_a$ can be expressed as

$$\boldsymbol{z}_a = b_1 \boldsymbol{Q}_{11} \boldsymbol{h}_1 + \sum_{k=2}^{K} b_k \boldsymbol{Q}_{1k} \boldsymbol{h}_k + \boldsymbol{w}_a = b_1 \boldsymbol{s}_1 + \boldsymbol{i}_a + \boldsymbol{w}_a \tag{13.2.10}$$

where the first terms represents the signal of interest (as in the single-user case), the second term represents the interference corrupting the active subspace, and the third term represents background noise. The matrix $\boldsymbol{Q}_{1k}$ denotes the correlation between the *active* basis waveforms of the first and the k^{th} user. Let $N_{ia} \leq N_o - N_a$ denote the dimension of the inactive subspace. The vector of correlator outputs corresponding to $\mathcal{S}_{ia}$ is given by

$$\boldsymbol{z}_{ia} = \sum_{k=2}^{K} b_k \widetilde{\boldsymbol{Q}}_{1k} \boldsymbol{h}_k + \boldsymbol{w}_{ia} = \boldsymbol{i}_{ia} + \boldsymbol{w}_{ia} \tag{13.2.11}$$

where $\widetilde{\boldsymbol{Q}}_{1k}$ represents the correlation between the *inactive* basis waveforms of the first user and the *active* basis waveforms of the k^{th} user. Note that $\boldsymbol{z}_{ia}$ does not contain the signal of interest — it only contains interference ($\boldsymbol{i}_{ia}$) and noise ($\boldsymbol{w}_{ia}$). By the very nature of spread-spectrum signals, $\boldsymbol{i}_a$ is correlated with $\boldsymbol{i}_{ia}$ and the basic idea is to use $\boldsymbol{i}_{ia}$ to suppress $\boldsymbol{i}_a$ corrupting the active subspace. After interference suppression, the "cleaned-up" version of $\boldsymbol{z}_a$ can be coherently combined as in the single-user case (see (13.2.8)) to decode the bits of the desired user. The dimension of $\mathcal{S}_{ia}$ controls the receiver complexity and the level of MAI suppression.

A linear receiver may be used to accomplish both interference suppression and diversity exploitation in the multiuser case. The bit estimate for b_1 is given by

$$\hat{b}_1 = \text{sign}\{\text{real}[\boldsymbol{g}^H \boldsymbol{z}]\} = \text{sign}\{\text{real}[\boldsymbol{g}_a^H \boldsymbol{z}_a + \boldsymbol{g}_{ia}^H \boldsymbol{z}_{ia}]\} \tag{13.2.12}$$

where the $N_a + N_{ia}$ dimensional filter $\boldsymbol{g}$ processes the active and inactive correlator outputs. If $\boldsymbol{g}$ is chosen to minimize the mean-squared-error (MSE) between $\widehat{b}_1$ and b_1, the linear filter admits an intuitively appealing decomposition: $\boldsymbol{g}_a^H \boldsymbol{z}_a + \boldsymbol{g}_{ia}^H \boldsymbol{z}_a = \boldsymbol{g}_a^H [\boldsymbol{z}_a - \boldsymbol{C}^H \boldsymbol{z}_{ia}]$. In this decomposition, the matrix filter $\boldsymbol{C}$ forms a minimum MSE (MMSE) estimate of $\boldsymbol{z}_a$ from $\boldsymbol{z}_{ia}$, thereby suppressing the interference corrupting the active subspace. The filter $\boldsymbol{g}_a$ then further suppresses any residual interference and exploits multipath-Doppler diversity to combat fading. The processing of N_a dimensional active correlator outputs, $\boldsymbol{z}_a$, is necessary for maximal diversity exploitation. The inactive subspace serves the sole purpose of interference suppression. It's dimension can be progressively increased to suppress any desired number of interfering users up to $N_o - N_a$. Thus, the notion of active/inactive time-frequency subspaces provides a natural framework for interference suppression and controlling the complexity of the multiuser time-frequency receiver.

13.2.3.4 An Illustrative Example

Recall that the key parameters are the signal space dimension, $N_o \approx TB$, which is proportional to the spreading gain $N = T/T_c$, and the products $T_m B$ and TB_d that control the level of multipath and Doppler diversity, respectively. We consider a system with a spreading gain of $N = 64$ and a multipath spread of $T_m = 2T_c$.

Consider first a single-user system to illustrate the effects of multipath-Doppler diversity. Suppose chip-rate sampling ($B = 1/T_c$) at the receiver so that $L = 2$, and $0 < TB_d < 0.5$ so that $M = 1$. This results in an $N_a = 9$ dimensional active subspace. Assume uniform power in all the multipath channel components. Let $\rho \in (0, 1/3)$ denote the ratio of the power in each of the $m = \pm 1$ Doppler components relative to the total power in the three ($m = -1, 0, 1$) components. As $TB_d \to 0.5$ (faster fading), $\rho \to 1/3$ to yield maximum Doppler diversity.[1] Figure 13.2.1(a) shows the BEP of the coherence time-frequency RAKE receiver (13.2.8) as a function of SNR for different values of ρ. It is evident that time-frequency RAKE receiver yields significant gains over RAKE as ρ increases. We note that while typical values of TB_d are relatively small in practice,[2] simple signaling schemes with longer symbol duration may be used to increase TB_d to achieve significant Doppler diversity [4].

Now consider a multiuser system to illustrate the role of active/inactive subspaces in interference suppression. We consider a slow-fading system ($TB_d = 0$) employing 4-fold oversampling ($B = 4/T_c$), resulting in a signal space dimension of $N_o = 4N = 256$ and $N_a = 9$. There are are 27 interfering users at the same power as the desired user. The signal-to-interference-and-noise-ratio (SINR) at the output of the desired user's receiver is an appropriate performance metric in the multiuser case. Figure 13.2.1(b) shows the SINR of the MMSE receiver as a func-

[1] For a uniform Doppler power spectrum, the values $\rho = 0.0125, 0.05, 0.25$ are achieved at $TB_d \approx 0.2, 0.4, 0.6$, respectively [4].

[2] For example, $TB_d \approx 0.013$, at a carrier frequency of 1.8 GHz, data rate of 10 kHz, and maximum speed of 50 mph.

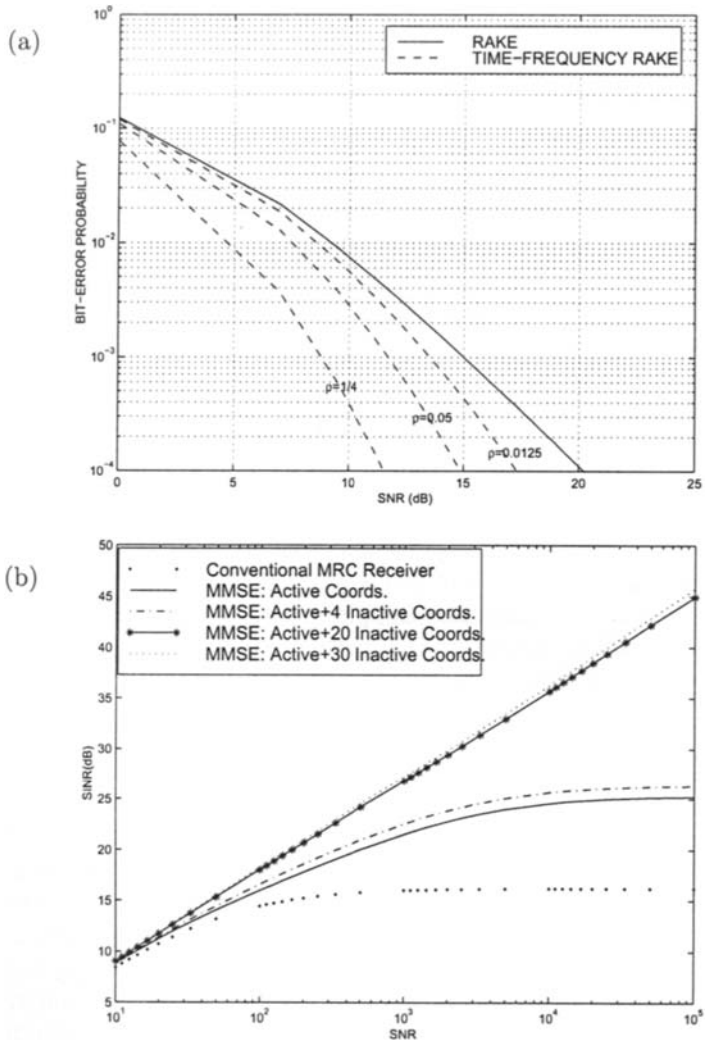


Fig. 13.2.1: (a) Joint multipath-Doppler diversity gains achieved by a single-user coherent time-frequency RAKE receiver as a function of ρ. The performance of conventional RAKE corresponds to $\rho = TB_d = 0$. (b) The monotonic increase in the output SINR of a multiuser time-frequency RAKE receiver with increasing dimension of the inactive time-frequency subspace. For $N_{ia} < 19$, the SINR saturates, whereas for $N_{ia} \geq 19$, interference is completely suppressed so that output SINR increases linearly with input SNR as in a single-user system.

tion of the transmitted SNR of the desired user for different values for the inactive subspace dimension, N_{ia}. As evident from Figure 13.2.1(b), the receiver goes from an interference-limited regime to a noise-limited regime as N_{ia} increases, with the

optimal cut-off value around $N_{ia} = 28 - N_a = 19$. Thus, the dimension of the inactive subspace can be progressively increased to achieve a desired level of interference suppression with commensurate receiver complexity.

13.2.4 Summary and Conclusions

The time-frequency channel model (13.2.3) at the heart of the framework presented in this article was first developed by Bello in his classic paper [2] on WSSUS channels. The monograph by Kennedy [3] exploits the model to address several key questions relating to reliable communication over dispersive channels, including capacity and error exponents. Results relating to the use of the model in wireless communications are more recent and are necessarily incomplete due to ongoing investigations. As evident from our brief discussion on diversity and interference suppression, the insights offered by the time-frequency perspective are very useful in efficient transceiver design. While not discussed in this article, similar ideas can be fruitfully exploited in the design of efficient transmission schemes for reaping the capacity of linear dispersive channels, including the design of orthogonal frequency division multiplexing (OFDM) schemes that are strong candidates for high rate wireless communication. Furthermore, the same basic principles apply in underwater acoustical communications as well.

References

[1] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 3rd ed., 1995.

[2] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Communication Systems*, vol. 11, pp. 360–393, December 1963.

[3] R. S. Kennedy, *Fading dispersive communication channels*. New York: Wiley, 1969.

[4] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Communications*, vol. 47, pp. 123–132, January 1999.

[5] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[6] S. Bhashyam, A. M. Sayeed, and B. Aazhang, "Time-selective signaling and reception for communication over multipath fading channels," *IEEE Trans. Communications*, vol. 48, pp. 83–94, January 2000.

[7] A. M. Sayeed, A. Sendonaris, and B. Aazhang, "Multiuser detection in fast fading multipath environments," *IEEE J. on Selected Areas in Communications*, vol. 16, pp. 1691–1701, December 1998.

[8] T. A. Kadous and A. M. Sayeed, "Decentralized multiuser detection for time-varying multipath channels," *IEEE Trans. Communications*, vol. 48, pp. 1840–1852, November 2000.

13.3 EIGENFUNCTIONS OF UNDERSPREAD LINEAR COMMUNICATION SYSTEMS [0]

The knowledge of the eigenfunctions of a linear system is a fundamental issue both from the theoretical as well as from the applications point of view. Nonetheless, no analytic solution is available for the eigenfunctions of a general linear system. There are two important classes of contributions suggesting analytic expressions for the eigenfunctions of slowly-varying operators: [1] and the references therein, where it was proved that the eigenfunctions of underspread operators can be approximated by signals whose time-frequency distribution (TFD) is well localized in the time-frequency plane, and [2] where a strict relationship between the instantaneous frequency of the channel eigenfunctions and the contour lines of the Wigner Transform of the operator kernel (or Weyl symbol) was derived for Hermitian slowly-varying operators. (See also Article 4.7.) In this article, following an approach similar to [2], we will show that the eigenfunctions can be found exactly for systems whose spread function is concentrated along a straight line and they can be found in approximate sense for those systems whose spread function is maximally concentrated in regions of the Doppler-delay plane whose area is smaller than one.

13.3.1 Eigenfunctions of Time-Varying Systems

The input/output relationship of a continuous-time (CT) linear system is [3]:

$$y(t) = \int_{-\infty}^{\infty} h(t,\tau)x(t-\tau)d\tau \tag{13.3.1}$$

where $h(t,\tau)$ is the system impulse response. Although throughout this section we will use the terminology commonly adopted in the transit of signals through time-varying channels, it is worth pointing out that the mathematical formulation is much more general. For example, (13.3.1) can be used to describe the propagation of waves through non homogeneous media and in such a case the independent variables t and τ are spatial coordinates. Following the same notation introduced by Bello [3], any linear time-varying (LTV) channel can be fully characterized by its impulse response $h(t,\tau)$, or equivalently by the *delay-Doppler spread function* (or simply spread function) $S(\nu,\tau) := \int_{-\infty}^{\infty} h(t,\tau)e^{-j2\pi\nu t}dt$, or by the *time-varying transfer function* $H(t,f) := \int_{-\infty}^{\infty} h(t,\tau)e^{-j2\pi f\tau}d\tau$.

Since the kernels of LTV systems in general are not self-adjoint, it is not possible to define the eigenfunctions of a linear system, but we can introduce the so called *left and right singular functions* (in the following we will use the term eigenfunction only for simplicity, meaning generically the left and right singular functions). In

[0]Author: **Sergio Barbarossa**, INFOCOM Department, University of Rome "La Sapienza", Via Eudossiana 18, 00184 Rome, Italy (sergio@infocom.uniroma1.it). Reviewers: G. Matz and A. M. Sayeed.

fact, if the system impulse response is square-integrable, i.e.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|h(t,\tau)|^2\,dt\,d\tau<\infty, \tag{13.3.2}$$

then there exists a countable set of *singular values* λ_i and two classes of orthonormal functions $v_i(t)$ and $u_i(t)$, named right and left singular functions, such that the following system of integral equations holds true

$$\lambda_i u_i(t) = \int_{-\infty}^{\infty} h(t,t-\tau)v_i(\tau)d\tau, \tag{13.3.3}$$

$$\lambda_i v_i(\tau) = \int_{-\infty}^{\infty} h^*(t,t-\tau)u_i(t)dt. \tag{13.3.4}$$

Inserting (13.3.3) in (13.3.4), we have

$$\lambda_i^2 v_i(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^*(t,t-\tau)h(t,\theta)v_i(t-\theta)d\theta dt. \tag{13.3.5}$$

so that $v_i(\tau)$ is the eigenfunction of the system whose kernel is

$$\tilde{h}(\tau,\theta) := \int_{-\infty}^{\infty} h^*(t,t-\tau)h(t,t-\theta)dt. \tag{13.3.6}$$

In practice, there are at least two quite common situations where $h(t,\tau)$ is not square-integrable: (i) linear time-invariant (LTI) channels, where $h(t,\tau)$ is constant along t [1]; and (ii) multipath channels with specular reflections, where $h(t,\tau)$ contains Dirac pulses. However, to avoid unnecessary complications with different notations as a function of the integrability assumption, in the following we will keep assuming (13.3.2), considering the aforementioned exceptions as limiting cases, as in [4, sec. 8].

13.3.2 Systems with Spread Function Confined to a Straight Line

If the spread function is confined to a line, i.e.

$$S(\nu,\tau) = g(\tau)\delta(\nu - f_0 - \mu\tau), \tag{13.3.8}$$

the singular functions are *chirp* signals, i.e.

$$v_i(t) = e^{j\pi\mu t^2}e^{j2\pi f_i t} \tag{13.3.9}$$

$$u_i(t) = e^{j\pi\mu t^2}e^{j2\pi f_i t}e^{j2\pi f_0 t} = v_i(t)e^{j2\pi f_0 t}. \tag{13.3.10}$$

[1] The LTI case as well as a large class of time-varying systems exhibiting some sort of stationarity can be dealt with by requiring the following integrability condition

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}dt\int_{-\infty}^{\infty}|h(t,\tau)|^2d\tau<\infty, \tag{13.3.7}$$

instead of (13.3.2).

In fact, the impulse response corresponding to (13.3.8) is

$$h(t,\tau) = g(\tau)e^{j2\pi\mu\tau t}e^{j2\pi f_0 t} \tag{13.3.11}$$

and, substituting (13.3.11) and (13.3.9) in (13.3.3) we get

$$\lambda_i u_i(t) = e^{j2\pi f_0 t}e^{j2\pi f_i t}e^{j\pi\mu t^2}G_\mu(f_i) = G_\mu(f_i)e^{j2\pi f_0 t}v_i(t), \tag{13.3.12}$$

where $G_\mu(f)$ is the Fourier transform (FT) of $g_\mu(t) := g(t)e^{j\pi\mu t^2}$. We can verify that (13.3.12) is satisfied if $u_i(t)$ is given by (13.3.10) and $\lambda_i = G_\mu(f_i)$. It is also straightforward to check that the two classes of functions $v_i(t)$ and $u_i(t)$ are orthogonal. Interestingly, the contour lines of $|H(t,f)|$ coincide with the instantaneous frequency of the eigenfunctions. In fact, the transfer function associated to (13.3.8) is $H(t,f) = G(f-\mu t)e^{j2\pi f_0 t}$, where $G(f)$ denotes the FT of $g(\tau)$, so that $|H(t,f)|$ is constant along lines of equation $f = \mu t + f_i$, which coincides with the instantaneous frequency of the right singular functions. Furthermore, if $f_0 = 0$, i.e. $S(\nu,\tau)$ is maximally concentrated along a line passing through the origin, the left and right singular functions are simply proportional to each other and we can talk of eigenfunctions and eigenvalues. Finally, it is worth noticing that if the spread function is mainly concentrated inside a rectangle of area $B_{max}\tau_{max} \ll 1$, thus $\mu\tau_{max}^2 \ll 1$ and $|H(t, f_i + \mu t)| = |G(f_i)| \approx |G_\mu(f_i)|$, so that the modulus of the i-th eigenvalue coincides approximately with the absolute value of the channel transfer function evaluated over the curve given by the eigenfunctions' instantaneous frequency.

In the following, we will show how these results can be generalized, albeit in approximate sense, to the more challenging case where the spread function is not confined to a straight line. But, before considering the more general case, it is worthwhile to remark that the model (13.3.8) encompasses three examples of systems commonly encountered in the applications, namely (i) time-invariant systems, where $S(\nu,\tau) = g(\tau)\delta(\nu)$, which corresponds to $\mu = 0$ and thus to having, as well known, sinusoidal eigenfunctions; (ii) multiplicative systems, where $S(\nu,\tau) = C(\nu)\delta(\tau)$, which corresponds to $\mu = \infty$ and thus to Dirac pulses as eigenfunctions; (iii) communication channels affected by *two-ray* multipath propagation, each ray having its own delay and Doppler frequency shift, i.e.

$$S(\nu,\tau) = \sum_{q=0}^{1} h_q\delta(\nu-\nu_q)\delta(\tau-\tau_q) \quad \text{or} \quad h(t,\tau) = \sum_{q=0}^{1} h_q e^{j2\pi f_q t}\delta(\tau-\tau_q). \tag{13.3.13}$$

In such a case, the eigenfunctions are chirp signals having different initial frequencies, but all with the same sweep rate $\mu = (f_1 - f_0)/(\tau_1 - \tau_0)$, which depends on the channel delay and Doppler parameters.

13.3.3 Analytic Models for Eigenfunctions of Underspread Channels

We extend now the analysis to systems whose spread function has a support, in the delay-Doppler domain, with small, but, differently from the previous case, non-null

area. Interestingly, this case encompasses all current digital communication systems. The aim of the ensuing analysis is to show that if $S(\nu,\tau)$ is mainly concentrated around the origin of the Doppler-delay plane, along one of the two axes but not along both, the main result derived above can be generalized, even though only in approximate sense.

First of all, proceeding as in [1], we define the absolute moments of $S(\nu,\tau)$ as

$$m_S^{(k,l)} := \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\nu|^k|\tau|^l|S(\nu,\tau)|d\nu d\tau}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|S(\nu,\tau)|d\nu d\tau}. \tag{13.3.14}$$

We say that a system is *underspread* if all the products $m_S^{(i,j)}m_S^{(k,l)}$ of order $i+j+k+l \geq 2$, where the indices are such that there is at least the product of a non null moment along τ times a non null moment along ν, are "small". This definition is not rigorous, but its meaning will be clarified within the proof of the main statement of this section. Since the partial derivatives of $H(t,f)$ can be upper bounded as follows

$$\left|\frac{\partial H^{k+l}(t,f)}{\partial t^k \partial f^l}\right| \leq (2\pi)^{k+l} m_S^{(k,l)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|S(\nu,\tau)|d\nu d\tau, \tag{13.3.15}$$

if $S(\nu,\tau)$ has small moments, $H(t,f)$ must be a *smooth function* in at least one direction.

In the following we show that, if the system is underspread, the singular function associated to the i-th singular value can be approximated by the following analytic function

$$v_i(t) := \sum_{m=1}^{K_i(t)} v_{i,m}(t) := \sum_{m=1}^{K_i(t)} a_{i,m}(t)e^{j\phi_{i,m}(t)}, \tag{13.3.16}$$

where (i) the instantaneous phase $\phi_{i,m}(t)$ is such that the corresponding instantaneous frequency $f_{i,m}(t) := \dot{\phi}_{i,m}(t)/2\pi$ of $v_{i,m}(t)$ is one of the real solutions of

$$|H(t, f_{i,m}(t))|^2 = \lambda_i^2, \quad \forall m; \tag{13.3.17}$$

(ii) the amplitude $a_{i,m}(t)$ is approximately constant and different from zero only within the time interval where $|H(t, f_{i,m}(t))|^2 = \lambda_i^2$ admits a real solution, and its value is such that $v_i(t)$ has unit norm; $K_i(t)$ is the number of solutions of (13.3.17), for each λ_i and t.

The existence of a real solution for $f_{i,m}(t)$ of (13.3.17) implies that the singular values λ_i must be bounded in the following interval: $\min_{t,f}|H(t,f)| \leq \lambda_i \leq \max_{t,f}|H(t,f)|$. Between these two boundaries, not all values of λ_i are possible: The only admissible values are the ones that allow the eigenfunctions to be orthonormal and respect Heisenberg's uncertainty principle, similarly to the area rule suggested in [2].

From (13.3.17) we notice that the instantaneous frequencies of the system eigenfunctions coincide with the contour lines of $|H(t,f)|$. Typically, the contour plots

are closed curves and then $K_i(t)$ is usually an even integer. In general, we have verified numerically that if there are more closed curves corresponding to the same eigenvalue λ_i, the multiplicity of the eigenvalue is equal to the number of closed curves corresponding to λ_i, with each closed curve giving rise to one eigenfunction.

We show now under which approximations, the function $v_i(\tau)$, as given in (13.3.16), is a solution of (13.3.5). Exploiting the system linearity, we compute the output $y_{i,m}(t)$ corresponding to each m-th component $v_{i,m}(t)$ in (13.3.16) and then we exploit the superposition principle to derive the output corresponding to $v_i(t)$. In our proof, we assume that the support of $h(t,\tau)$ along τ is small.[2] As a consequence, for each value of τ, the product $h^*(t,t-\tau)h(t,\theta)$ in (13.3.5) assumes significant values only for small values of both $t-\tau$ and θ. We can thus expand $v_{i,m}(t-\theta)$ in (13.3.5), around τ and keep only the lower order components

$$v_{i,m}(t-\theta) \approx a_{i,m}(\tau)e^{j\phi_{i,m}(\tau)}e^{j\dot{\phi}_{i,m}(\tau)(t-\theta-\tau)}, \tag{13.3.18}$$

having used a first order approximation for $\phi_{i,m}(t-\theta)$ and a zero-th order approximation for $a_{i,m}(t-\theta)$. Substituting (13.3.18) into (13.3.5) and invoking the principle of stationary phase [5] to derive an approximate analytic expression of the integral, we get the m-th output term

$$\begin{aligned} y_{i,m}(\tau) &\approx a_{i,m}(\tau)e^{j\phi_{i,m}(\tau)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^*(t,t-\tau)h(t,\theta)e^{-j\dot{\phi}_{i,m}(\tau)\theta}e^{j\dot{\phi}_{i,m}(\tau)(t-\tau)}d\theta dt \\ &= a_{i,m}(\tau)e^{j\phi_{i,m}(\tau)} \int_{-\infty}^{\infty} h^*(t,t-\tau)H(t,f_{i,m}(\tau))e^{j\dot{\phi}_{i,m}(\tau)(t-\tau)}dt. \end{aligned} \tag{13.3.19}$$

After a few algebraic manipulations involving the Taylor series expansion of both $h(t,t-\tau)$ and $H(t,f)$ about τ in their first argument and summing over m, we get

$$y(\tau) \approx \lambda_i^2 v_i(\tau) + \sum_{m=1}^{K_i} \underbrace{\sum_{k,l=0}^{\infty}}_{k,l\neq 0} \frac{H_{t,f}^{(k,0)}(\tau, f_{i,m}(\tau))H_{t,f}^{(l,k+l)^*}(\tau, f_{i,m}(\tau))}{(-j2\pi)^{k+l}k!\ l!}a_{i,m}(\tau)e^{j\phi_{i,m}(\tau)}, \tag{13.3.20}$$

where $H_{t,f}^{(k,l)}(t,f) := \partial^{k+l}H(t,f)/\partial t^k \partial f^l$. This equation shows that $v_i(\tau)$, as given in (13.3.16), is (approximately) the eigenfunction associated to the eigenvalue λ_i^2 if the perturbation, given by the second term of the right-hand side of (13.3.20), is small with respect to $\lambda_i^2 v_i(\tau)$. From (13.3.20), we notice that the perturbation is equal to the sum of complex functions given by the product of the partial derivatives of the system transfer function, evaluated along the curve where the modulus of the transfer function is constant. Furthermore, each term in the perturbation contains at least the first order derivative with respect to both time and frequency. Therefore,

[2] If this assumption is not true, to respect our main assumption about the concentration of the spread function, the spread of $S(\nu,\tau)$ along ν must be very small. In such a case, using duality arguments, we can derive equivalent results working with the spectrum of the eigenfunctions.

the perturbation is small with respect to the first term in (13.3.20) if the transfer function is smooth in at least one direction, i.e. time or frequency. Hence, the analytic model (13.3.16) is valid only for underspread systems, i.e. systems whose transfer function has small partial derivatives, by virtue of (13.3.15), in at least one direction. Furthermore, since the energy of the first term is λ_i^4, the approximation error is smaller for the highest eigenvalues.

Since so many approximations have been used to justify the analytic model (13.3.16), it is necessary to check the validity of such approximations. Given the crucial role played by the instantaneous frequency in the definition of the system eigenfunction and the interplay of time and frequency, the analysis of the time-frequency distribution (TFD) of the system eigenfunctions plays a fundamental role as a validation tool. Since the validation is necessarily numerical, we start deriving the equivalent discrete-time (DT) system corresponding to the continuous-time (CT) relationship (13.3.1). Specifically, we consider the system obtained by windowing $h(t,\tau)$ in time and in frequency. Assuming that the input signal $x(t)$ has a spectrum confined within the bandwidth $[-1/2T_s, 1/2T_s]$, we can express $x(t)$ as

$$x(t) = \sum_{k=-\infty}^{\infty} x[k]\mathrm{sinc}(\pi(\mathrm{t} - \mathrm{kT_s})/\mathrm{T_s}), \tag{13.3.21}$$

where $x[k] := x(kT_s)$ and$1/T_s$ is the sampling rate. Sampling the continuous time system output $y(t)$ at the same rate $1/T_s$,[3] we get the equivalent discrete-time I/O relationship

$$y[n] := y(nT_s) = \sum_{k=-\infty}^{\infty} h[n, n-k]x[k], \tag{13.3.22}$$

where $h[n,k]$ denotes the equivalent DT impulse responses, defined as

$$h[n, n-k] := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{sinc}(\pi(\mathrm{nT_s} - \theta)/\mathrm{T_s})\mathrm{sinc}(\pi(\theta - \tau - \mathrm{kT_s})/\mathrm{T_s})\mathrm{h}(\theta,\tau)\mathrm{d}\tau\mathrm{d}\theta. \tag{13.3.23}$$

Equation (13.3.22) is the DT counterpart of (13.3.1). Assuming that $h[n,k]$ has finite support over k, i.e. the channel is FIR of order L, and parsing the input sequence into consecutive blocks of size R, the discrete-time model leads directly to the matrix I/O relationship $\boldsymbol{y}(n) = \boldsymbol{H}(n)\boldsymbol{x}(n)$, where $\boldsymbol{H}(n)$ is the $P \times R$ channel matrix, with $P = R + L$, relative to the n-th transmitted block, whose (i,j) entry is $\{\boldsymbol{H}(n)\}_{i,j} = h[nP + i, i - j]$, whereas $\boldsymbol{x}(n) := (x[nR], \ldots, x[R + R - 1])^T$ and $\boldsymbol{y}(n) := (y[nP], \ldots, y[nP + P - 1])^T$ are the input and output blocks.

The discrete time counterpart of (13.3.3) and (13.3.4) is the singular value decomposition (SVD) of the channel matrix $\boldsymbol{H}(n)$, i.e. $\boldsymbol{H}(n) = \boldsymbol{U}(n)\boldsymbol{\Lambda}(n)\boldsymbol{V}^H(n)$, that allows us to write

[3]We assume that $1/T_s$ is large enough to respect the Nyquist principle for the system output $y(t)$; this means that, if we take into account the bandwidth increase due to the transit through a time-varying system, $1/T_s$ is strictly larger than the bandwidth of $x(t)$.

$$\boldsymbol{U}(n)\Lambda(n) = \boldsymbol{H}(n)\boldsymbol{V}(n), \quad or \quad \boldsymbol{H}^H(n)\boldsymbol{U}(n) = \boldsymbol{V}(n)\boldsymbol{\Lambda}(n), \tag{13.3.24}$$

where the columns of $\boldsymbol{U}(n)$ and $\boldsymbol{V}(n)$ are the left and right channel singular vectors associated to the singular values contained in the diagonal matrix $\boldsymbol{\Lambda}(n)$.

To check the validity of model (13.3.16), we proceed through the following steps. Given the impulse response $h(t,\tau)$ of the CT system, (i) we build the channel matrix $\boldsymbol{H}(n)$ of the equivalent DT system; (ii) we compute the SVD of $\boldsymbol{H}(n)$; (iii) we compute the TFD of the right and left singular vectors associated to the generic singular value λ_i; and (iv) we compare the energy distribution of these TFD with the contour plot of $|H(t,f)|$ corresponding to level λ_i. We used as a basic tool to analyze the signals in the time-frequency domain the Smoothed Pseudo-Wigner-Ville Distribution (SPWVD) with reassignment, presented in [6] and Article 7.2, for its property of having low cross terms without degrading the resolution. We considered as a test system a communication channel affected by multipath propagation, thus described by the CT impulse response

$$h(t,\tau) = \sum_{q=0}^{Q-1} h_q e^{j2\pi f_q t}\delta(\tau - \tau_q),$$

where each path is characterized by the triplet of amplitude h_q, delay τ_q and Doppler shift f_q. We generated the amplitudes h_q as independent identically distributed (iid) complex Gaussian random variables with zero mean and unit variance (the Rayleigh fading model), and the variables τ_q and f_q as iid random variables with uniform distribution within the intervals $[0, \Delta\tau]$ and $[-\Delta f/2, \Delta f/2]$, respectively. An example, relative to a multipath channel composed of $Q = 12$ paths, with $\Delta\tau = 4T_s$ and $\Delta f = 4/NT_s$, $N = 128$, is reported in Fig. 13.3.1 where we show: (a) the mesh plot of $|H(t,f)|$, (b) two contour plots of $|H(t,f)|$ corresponding to the levels λ_{16} (dashed line) and λ_{32} (solid line); (c) the contour plot of the SPWVD of $\boldsymbol{v}_{16}$; (d) the contour plot of the SPWVD of $\boldsymbol{v}_{32}$.

It is worth noticing how, in spite of the rather peculiar structure of the contour plots of $|H(t,f)|$, the SPWVDs of the two singular functions are strongly concentrated along curves coinciding with the contour lines of $|H(t,f)|$ corresponding to the associated singular values, as predicted by the theory.

It is also interesting to observe the *bubble*-like structure of the two SPWVDs. Indeed this behavior is quite common, because in general the contour lines of the time-varying transfer function are typically closed curves.

Before concluding this section, it is also important to provide some physical insight to justify the rather peculiar behavior of the channel eigenfunctions. Indeed, the bubble-like structure is perfectly functional to guaranteeing two of the fundamental properties of the eigenfunctions, namely orthogonality and system modes excitation. In fact, by construction, (13.3.16) and (13.3.17) insure that the instantaneous frequency curves of singular functions associated to distinct eigenvalues do not intersect. Therefore, if the WVDs of the eigenfunctions associated to distinct eigenvalues are well concentrated along their instantaneous frequency curve (i.e. if their amplitude modulation is negligible), the scalar product of their WVDs is null

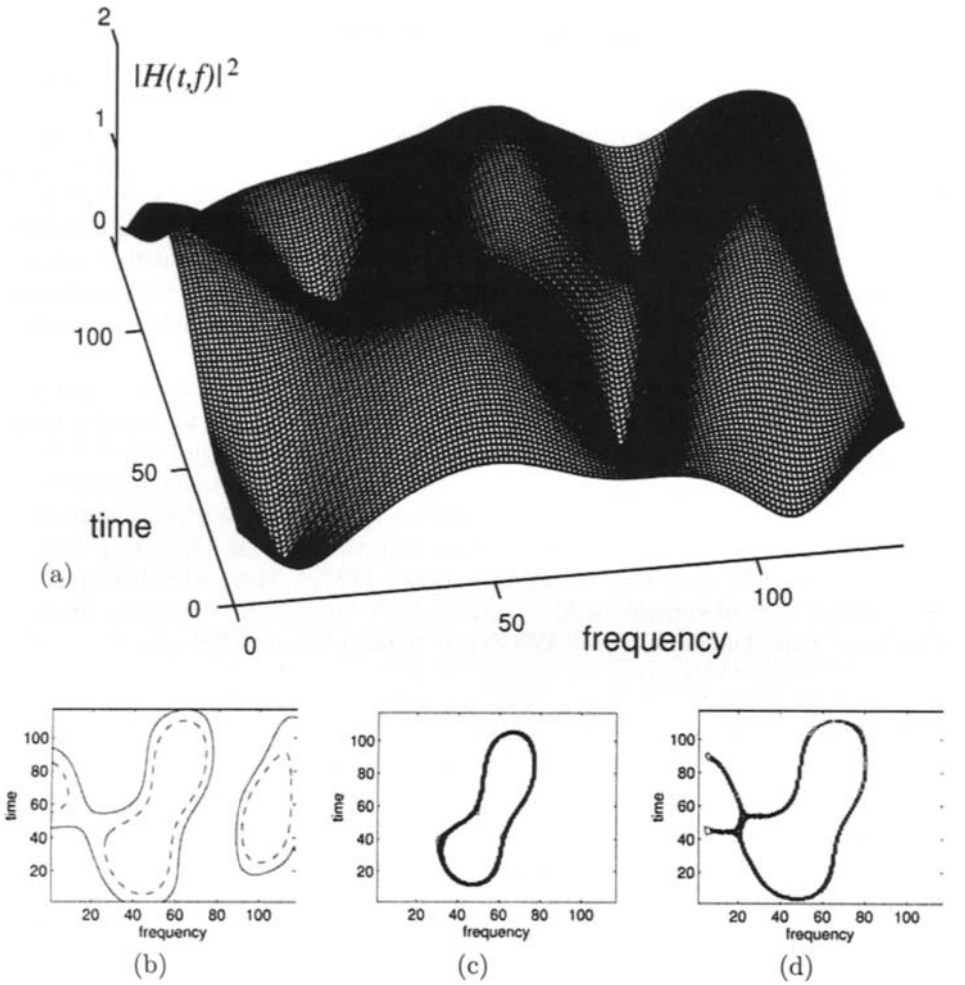


Fig. 13.3.1: Comparison between contour lines of $|H(t,f)|$ and TFDs of channel singular vectors: (a) 3D plot of $|H(t,f)|$; (b) contour lines of $|H(t,f)|$ corresponding to levels λ_{16} (dashed line) and λ_{32} (solid line); (c) SPWVD of $\boldsymbol{v}_{16}$; (d) SPWVD of $\boldsymbol{v}_{32}$.

and thus, by virtue of Moyal's formula, the eigenfunctions are orthogonal, as required. Considering now the modes of the system, we know that the unit energy input signal that maximizes the output energy is the right singular function associated to the highest singular value. Now, if we combine this basic property with the model given in (13.3.16) and (13.3.17), we can conclude that, not surprisingly, the input signal which maximizes the output energy is the signal whose energy is concentrated in the time-frequency region where the channel time-varying transfer function is maximum.

13.3.4 Optimal Waveforms for Digital Communications through LTV Channels

Let us consider one of the most interesting applications of the theory described above, i.e. the transmission of information symbols $s[k]$ through an LTV channel. In Article 13.2, for example, it is shown how to convert the channel dispersiveness, possibly in both time and frequency domains, into a useful source of diversity to be exploited to enhance the SNR at the receiver. Here we show that if the transmitter is able to predict the channel time-varying transfer function, at least within the next time slot where one is going to transmit, it is possible to optimize the transmission strategy and take full advantage of the diversity offered by the channel dispersiveness (see e.g. [7] for more details).

Considering a channel with approximately finite impulse response of order L, we can parse the input sequence in consecutive blocks of K symbols and insert null guard intervals of length L between successive blocks to avoid inter-block interference. If the symbol rate is $1/T_s$, the time necessary to transmit each block is KT_s. For each i-th block, we must consider the channel $h_i(t,\tau)$ obtained by windowing $h(t,\tau)$ in time, in order to retain only the interval $[iKT_s,(i+1)KT_s]$, and in frequency, keeping only the band $[-1/2T_s, 1/2T_s]$. The optimal strategy for transmitting a set of symbols $s_i[k] := s[iK+k]$, $k=0,\ldots,K-1$, in the presence of additive white Gaussian noise (AWGN), is to send the signal [4, sec. 8]

$$x_i(t) = \sum_{k=0}^{K-1} c_{i,k} s_i[k] v_{i,k}(t) \tag{13.3.25}$$

where $v_{i,k}(t)$ is the right singular function associated to the k-th eigenvalue of the channel response $h_i(t,\tau)$ in the i-th transmit interval and $c_{i,k}$ are coefficients used to allocate the available power among the transmitted symbols according to some optimization criterion [7]. Using (13.3.3), the received signal is thus

$$y_i(t) = \int_{-\infty}^{\infty} h_i(t,\tau) x_i(t-\tau) d\tau + w(t) = \sum_k c_{i,k} \lambda_{i,k} s_i[k] u_{i,k}(t) + w(t), \tag{13.3.26}$$

where $u_{i,k}(t)$ is the left singular function associated to the k-th singular value of $h_i(t,\tau)$ and $w(t)$ is AWGN. Hence, by exploiting the orthonormality of the functions $u_{i,k}(t)$, the transmitted symbols can be estimated by simply taking the scalar products of $y(t)$ with the left singular functions, i.e.

$$\hat{s}_i[m] = \frac{1}{\lambda_{i,m} c_{i,m}} \int_{-\infty}^{\infty} y(t) u_{i,m}^*(t) dt = s[m] + w_i[m], \tag{13.3.27}$$

where the noise samples sequence $w_i[m] := \int_{-\infty}^{\infty} w(t) u_{i,m}^*(t) dt$ constitutes a sequence of iid Gaussian random variables. In this way, the initial LTV channel, possibly dispersive in both time and frequency domains, has been converted into a set of parallel independent non-dispersive subchannels, with no intersymbol interference, and the symbol-by-symbol decision is also the maximum likelihood detector.

Most current transmission schemes turn out to be simple examples of the general framework illustrated above. For example, in communications through flat fading multiplicative channels, whose eigenfunctions are Dirac pulses, the optimal strategy is time division multiplexing. By duality, the optimal strategy for linear time-invariant channels is orthogonal frequency division multiplexing (OFDM). Interestingly, in the most general case (of underspread channels), the optimal strategy consists in sending symbols through channel-dependent *bubble-carriers.*

13.3.5 Summary and Conclusions

The analytic model for the eigenfunctions of underspread linear operators shown in this article, although approximate, shows that the energy of the system eigenfunctions is mainly concentrated along curves coinciding with level curves of the system transfer function. This property, for whose validation the analysis of the system eigenfunctions' TFD plays a fundamental role, gives a general framework for interpreting some current data transmission schemes and, most importantly, gives a new perspective on the optimal waveforms for transmissions over time-varying channels.

Pictorially speaking, if we draw a parallelism between time-frequency representations and musical scores, we may say that the eigenfunctions of underspread systems give rise to a polyphonic texture which reduces to monophonic lines only in the simple case of systems whose spread function is concentrated on a straight line. In the most general case, we have a polyphony of ascending and descending melodic lines which run in order to create bubbles whose shape is dictated by the contour lines of the system transfer function.

References

[1] 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. of Mathematical Physics*, vol. 39, pp. 4041–4070, August 1998. Special Issue on Wavelet and Time-Frequency Analysis.

[2] L. Sirovich and B. W. Knight, "On the eigentheory of operators which exhibit a slow variation," *Quarterly of Applied Mathematics*, vol. 38, pp. 469–488, 1980.

[3] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Communication Systems*, vol. 11, pp. 360–393, December 1963.

[4] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.

[5] A. Papoulis, *Signal analysis.* New York: McGraw-Hill, 1977.

[6] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[7] S. Barbarossa and A. Scaglione, "Time-varying fading channels," in *Signal Processing Advances in Wireless and Mobile Communications* (G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, eds.), vol. 2: "Trends in Single- and Multi-User Systems", ch. 1, Upper Saddle River, NJ: Prentice-Hall, 2001.

13.4 FRACTIONAL AUTOCORRELATION FOR DETECTION IN COMMUNICATIONS [0]

Linear frequency modulated (FM) signals, also called chirps, are commonly used in communications systems. For example, they are employed in modulation schemes that allow better resilience against interferences. Due to their frequent occurrence in communications, detection of chirp signals and estimation of their parameters have been of great interest. In this article, a method based on fractional autocorrelation is proposed for detection and parameter estimation of chirps. To that end, after briefly introducing the fractional Fourier transform (FRFT), formulations of fractional convolution and correlation are presented. Then, a detection statistic based on fractional autocorrelation is proposed and its performance is demonstrated through simulations.

13.4.1 Fractional Fourier Transform

The FRFT is a linear, energy preserving signal transformation that generalizes the conventional Fourier transform (FT) via an angle parameter ϕ [1,2]. For $\phi = \frac{\pi}{2}$, the FRFT simplifies to the conventional FT. The conventional FT of a time domain signal $s(t)$ can be interpreted as a counterclockwise rotation by $\frac{\pi}{2}$ radians in the time-frequency plane. The FRFT generalizes this rotation property by defining rotations associated with angles other than $\frac{\pi}{2}$.

The FRFT of a time domain signal $s(t)$ is defined [1,2] as

$$(\mathbb{F}^{\phi}s)(r) = S^{\phi}(r)$$
$$= \begin{cases} \sqrt{1-j\cot\phi}\, e^{j\pi r^2 \cot\phi} \int s(t)\, e^{j\pi t^2 \cot\phi}\, e^{-j2\pi tr \csc\phi} dt, & \phi \neq n\pi, \\ s(r), & \phi = 2n\pi, \\ s(-r), & \phi = (2n+1)\pi, \end{cases} \tag{13.4.1}$$

where $\mathbb{F}^{\phi}$ is the FRFT operator associated with angle ϕ. $S^{\phi}(r)$ denotes the fractional Fourier transformed signal and n is an integer. For $\phi = \frac{\pi}{2}$, the conventional FT is obtained. Similarly, for $\phi = 0$, one obtains the time domain representation $s(t)$ [1]. Furthermore, for $\phi = \pi$, the FRFT reduces to a time reversal operation [1],

$$(\mathbb{F}^{\pi}s)(t) = s(-t). \tag{13.4.2}$$

One fundamental property of the FRFT relates it to the WD. The WD of the time domain signal $s(t)$ is defined (in Section 2.1.3) as

$$W_s(t,f) = \int s(t+\frac{\tau}{2})\, s^*(t-\frac{\tau}{2})\, e^{-j2\pi f\tau} d\tau. \tag{13.4.3}$$

[0]Authors: **Olcay Akay**, Department of Electrical and Electronics Engineering, Dokuz Eylül University, Izmir, Turkey (olcay.akay@eee.deu.edu.tr), and **G. Faye Boudreaux-Bartels**, Department of Electrical Engineering, University of Rhode Island, Kingston, RI 02881, USA (boud@ele.uri.edu). Reviewers: M. J. Bastiaans and X.-G. Xia.

The relationship between the FRFT and the WD can be stated [1,2] as

$$W_{S^\phi}(t,f) = W_s(t\cos\phi - f\sin\phi, t\sin\phi + f\cos\phi) \tag{13.4.4}$$

where $W_{S^\phi}(t,f)$ denotes the WD of the FRFT signal, $S^\phi(r)$, whereas $W_s(t,f)$ is the WD of the original time domain signal, $s(t)$. According to (13.4.4), the WD of the fractional Fourier transformed signal $S^\phi(r)$ can be obtained by rotating the WD of the original signal $s(t)$ in the clockwise direction by ϕ.

13.4.2 Fractional Convolution and Correlation

Definitions of fractional convolution and correlation operations are obtained using the unitary fractional-shift operator $\mathbf{R}_\rho^\phi$, which is defined [3] as

$$(\mathbf{R}_\rho^\phi s)(t) = s(t-\rho\cos\phi)\, e^{-\jmath 2\pi\frac{\rho^2}{2}\cos\phi\sin\phi + \jmath 2\pi t\rho\sin\phi} \tag{13.4.5}$$

where ρ is the fractional lag variable. For $\phi = 0$, the unitary fractional-shift operator, $\mathbf{R}_\rho^\phi$, simplifies to the well-known unitary time-shift operator, $(\mathbf{R}_\tau^0 s)(t) = (\mathbf{T}_\tau s)(t) = s(t-\tau)$, with τ denoting the time lag. Analogously, as the special case for $\phi = \frac{\pi}{2}$, the unitary frequency-shift operator is obtained, $(\mathbf{R}_\nu^{\frac{\pi}{2}} s)(t) = (\mathbf{F}_\nu s)(t) = e^{\jmath 2\pi\nu t} s(t)$, with ν denoting the frequency lag.

Fractional convolution of two signals $s(t)$ and $h(t)$ is calculated [4] by evaluating the inner product[1] of the first signal $s(t)$ with the axis-reversed, complex-conjugated and fractionally-shifted version of the second signal $h(t)$,

$$\begin{aligned}(s *_\phi h)(r) &= \langle s, \mathbf{R}_r^\phi \mathbb{F}^\pi h^*\rangle = \langle s, \mathbf{R}_r^\phi \tilde{h}\rangle \\ &= e^{\jmath 2\pi\frac{r^2}{2}\cos\phi\sin\phi}\int s(\beta)\, h(r\cos\phi - \beta)\, e^{-\jmath 2\pi\beta r\sin\phi} d\beta. \end{aligned} \tag{13.4.6}$$

In (13.4.6), $*_\phi$ denotes fractional convolution associated with angle ϕ, $\tilde{h}(t) = h^*(-t)$, and $\mathbb{F}^\pi$ is the axis-reversal operator in (13.4.2). For $\phi = 0$, fractional convolution simplifies to linear time invariant (LTI) convolution given as

$$(s *_0 h)(t) = \langle s, \mathbf{T}_t \mathbb{F}^\pi h^*\rangle = \langle s, \mathbf{T}_t \tilde{h}\rangle = \int s(\beta)\, h(t-\beta)\, d\beta. \tag{13.4.7}$$

Similarly, for $\phi = \frac{\pi}{2}$, fractional convolution in (13.4.6) reduces to linear frequency invariant (LFrI) convolution which is obtained [4] by

$$\begin{aligned}(s *_{\frac{\pi}{2}} h)(f) &= \langle s, \mathbf{F}_f \mathbb{F}^\pi h^*\rangle = \langle s, \mathbf{F}_f \tilde{h}\rangle \\ &= \int s(\beta)\, h(-\beta) e^{-\jmath 2\pi f\beta} d\beta = \int S^{\frac{\pi}{2}}(\beta)\, H^{\frac{\pi}{2}}[-(f-\beta)] d\beta \end{aligned} \tag{13.4.8}$$

where $S^{\frac{\pi}{2}}(f)$ and $H^{\frac{\pi}{2}}(f)$ denote the conventional FTs of $s(t)$ and $h(t)$, respectively.

[1] The inner product of two functions $g(t)$ and $h(t)$ is defined as $\langle g, h\rangle = \int g(t)\, h^*(t)\, dt$.

Table 13.4.1: Definitions of fractional convolution and correlation at angle ϕ.

Convolution	$(s \star_\phi h)(r) = e^{\jmath 2\pi \frac{r^2}{2} \cos\phi \sin\phi} \int s(\beta)\, h(r\cos\phi - \beta)\, e^{-\jmath 2\pi \beta r \sin\phi} d\beta$
Cross-corr.	$(s \star_\phi h)(\rho) = e^{\jmath 2\pi \frac{\rho^2}{2} \cos\phi \sin\phi} \int s(\beta)\, h^*(\beta - \rho\cos\phi)\, e^{-\jmath 2\pi \beta \rho \sin\phi} d\beta$
Autocorr.	$(s \star_\phi s)(\rho) = e^{\jmath 2\pi \frac{\rho^2}{2} \cos\phi \sin\phi} \int s(\beta)\, s^*(\beta - \rho\cos\phi)\, e^{-\jmath 2\pi \beta \rho \sin\phi} d\beta$

Table 13.4.2: Linear time invariant (LTI) and frequency invariant (LFrI) convolution and correlation.

	LTI	LFrI
Convolution	$(s \star_0 h)(t) = \int s(\beta)\, h(t-\beta) d\beta$	$(s \star_{\frac{\pi}{2}} h)(f) = \int S^{\frac{\pi}{2}}(\beta)\, H^{\frac{\pi}{2}}[-(f-\beta)] d\beta$
Cross-corr.	$(s \star_0 h)(\tau) = \int s(\beta)\, h^*(\beta-\tau)\, d\beta$	$(s \star_{\frac{\pi}{2}} h)(\nu) = \int S^{\frac{\pi}{2}}(\beta) \left[H^{\frac{\pi}{2}}(\beta-\nu)\right]^* d\beta$
Autocorr.	$(s \star_0 s)(\tau) = \int s(\beta)\, s^*(\beta-\tau)\, d\beta$	$(s \star_{\frac{\pi}{2}} s)(\nu) = \int S^{\frac{\pi}{2}}(\beta) \left[S^{\frac{\pi}{2}}(\beta-\nu)\right]^* d\beta$

Analogously, fractional cross-correlation, denoted by $\star_\phi$, of $s(t)$ and $h(t)$ is calculated [4] using the fractional-shift operator, $\mathbf{R}_\rho^\phi$, within an inner product,

$$(s \star_\phi h)(\rho) = \langle s, \mathbf{R}_\rho^\phi h \rangle = e^{\jmath 2\pi \frac{\rho^2}{2} \cos\phi \sin\phi} \int s(\beta)\, h^*(\beta - \rho\cos\phi)\, e^{-\jmath 2\pi \beta \rho \sin\phi} d\beta. \tag{13.4.9}$$

Fractional autocorrelation at angle ϕ is similarly defined [4] by replacing the second function $h(t)$ in (13.4.9) with the function $s(t)$,

$$(s \star_\phi s)(\rho) = e^{\jmath 2\pi \frac{\rho^2}{2} \cos\phi \sin\phi} \int s(\beta)\, s^*(\beta - \rho\cos\phi)\, e^{-\jmath 2\pi \beta \rho \sin\phi} d\beta. \tag{13.4.10}$$

For the special cases of $\phi = 0$ and $\phi = \frac{\pi}{2}$, fractional correlation operations in (13.4.9) and (13.4.10) reduce to the corresponding LTI and LFrI correlation operations, respectively. Table 13.4.1 displays the definitions of fractional convolution and correlation operations. Definitions of LTI and LFrI convolution and correlation operations are listed in Table 13.4.2.

13.4.2.1 Alternative Formulations of Fractional Convolution and Correlation

In this subsection, alternative formulations for fractional convolution and correlation are presented. Having alternative equivalent formulations allows one to gain additional insight about these fractional operations. Furthermore, these alternative formulations suggest efficient ways for computer implementation.

The first alternative formulation of fractional convolution is given [4] as

$$(s *_\phi h)(r) = \int S^\phi(\beta)\, H^{-\phi}(r-\beta)\, d\beta = (S^\phi *_0 H^{-\phi})(r) \tag{13.4.11}$$

where $S^\phi(r) = (\mathbb{F}^\phi s)(r)$ and $H^{-\phi}(r) = (\mathbb{F}^{-\phi} h)(r)$. One can see from (13.4.11) that fractional convolution at angle ϕ can equivalently be calculated by a conventional LTI convolution of $S^\phi(r)$ and $H^{-\phi}(r)$. Another alternative formulation can be obtained by applying the convolution property of the conventional FT [4] to the right hand side of (13.4.11) and is given by

$$(s *_\phi h)(r) = \left(\mathbb{F}^{-\frac{\pi}{2}} \left\{ S^{\frac{\pi}{2}+\phi}(u) H^{\frac{\pi}{2}-\phi}(u) \right\}\right)(r). \tag{13.4.12}$$

According to this formulation, fractional convolution associated with angle ϕ can be computed by multiplying the FRFT at angle $\frac{\pi}{2}+\phi$ of the first signal, $S^{\frac{\pi}{2}+\phi}(u)$, with the FRFT at angle $\frac{\pi}{2}-\phi$ of the second signal, $H^{\frac{\pi}{2}-\phi}(u)$, and then taking a conventional inverse FT. This formulation suggests a way for efficient computer implementation of fractional convolution using the fast Fourier transform (FFT) along with the fast FRFT algorithm [5].

An alternative formula for fractional autocorrelation in (13.4.10) is given as [4]

$$(s \star_\phi s)(\rho) = \int S^\phi(\beta)[S^\phi(\beta-\rho)]^* d\beta = (S^\phi \star_0 S^\phi)(\rho). \tag{13.4.13}$$

Thus, as can be seen from (13.4.13), fractional autocorrelation can also be calculated via the LTI autocorrelation of $S^\phi(r)$. By applying the autocorrelation property of the conventional FT to the right hand side of (13.4.13), another alternative formulation of fractional autocorrelation is obtained as

$$(s \star_\phi s)(\rho) = \left(\mathbb{F}^{-\frac{\pi}{2}} \left\{ \left| S^{\frac{\pi}{2}+\phi}(u) \right|^2 \right\}\right)(\rho). \tag{13.4.14}$$

In (13.4.14), one FRFT with angle $\phi + \frac{\pi}{2}$ and one conventional inverse FT are used to compute fractional autocorrelation. This formulation suggests a fast discrete-time approximation of fractional autocorrelation. The discrete FRFT algorithm proposed in [5] has a computational load of $\mathcal{O}(N \log N)$ for a discrete-time signal of length N. Thus, utilizing the discrete FRFT and FFT algorithms together, a discrete-time approximation of fractional autocorrelation via (13.4.14) can be computed efficiently with a computational load of $\mathcal{O}(2N \log N + N)$ [4]. If one wants to calculate fractional autocorrelation for M different angles, then the total computational load is $\mathcal{O}[M(2N \log N + N)]$.

By computing the conventional FT of both sides of (13.4.14), one obtains

$$\left(\mathbb{F}^{\frac{\pi}{2}} \{(s \star_\phi s)(\rho)\}\right)(u) = |S^{\frac{\pi}{2}+\phi}(u)|^2. \tag{13.4.15}$$

Eq. (13.4.15) can be thought as the fractional autocorrelation theorem. For $\phi = 0$, it simplifies to the autocorrelation theorem of the conventional FT [4],

$$\left(\mathbb{F}^{\frac{\pi}{2}} \{(s \star_0 s)(\tau)\}\right)(f) = |S^{\frac{\pi}{2}}(f)|^2. \tag{13.4.16}$$

Table 13.4.3: Alternative formulations of the fractional convolution and correlation operations associated with the fractional domain at angle ϕ. The cross-AF, $AF_{s,h}(\tau,\nu) = \int s(t+\frac{\tau}{2})\,h^*(t-\frac{\tau}{2})\,e^{-\jmath 2\pi\nu t}dt$, simplifies to the auto AF in (13.4.17) when $h(t) = s(t)$. Also, $\tilde{h}(t) = h^*(-t)$.

Fractional	alternative I	alternative II	alternative III
Convolution $(s *_\phi h)(r) =$	$\int S^\phi(\beta)\,H^{-\phi}(r-\beta)\,d\beta$	$\left(\mathbb{F}^{-\frac{\pi}{2}}\left\{S^{\frac{\pi}{2}+\phi}(u)H^{\frac{\pi}{2}-\phi}(u)\right\}\right)(r)$	$AF_{s,\tilde{h}}(r\cos\phi, r\sin\phi)$
Cross-corr. $(s \star_\phi h)(\rho) =$	$\int S^\phi(\beta)[H^\phi(\beta-\rho)]^*d\beta$	$\left(\mathbb{F}^{-\frac{\pi}{2}}\left\{\left[S^{\frac{\pi}{2}+\phi}(u)\right]\left[H^{\frac{\pi}{2}+\phi}(u)\right]^*\right\}\right)(\rho)$	$AF_{s,h}(\rho\cos\phi, \rho\sin\phi)$
Autocorr. $(s \star_\phi s)(\rho) =$	$\int S^\phi(\beta)[S^\phi(\beta-\rho)]^*d\beta$	$\left(\mathbb{F}^{-\frac{\pi}{2}}\left\{\left\|S^{\frac{\pi}{2}+\phi}(u)\right\|^2\right\}\right)(\rho)$	$AF_s(\rho\cos\phi, \rho\sin\phi)$

The alternative formulations of fractional cross-correlation are similarly obtained by replacing the second $s(t)$ in (13.4.13) and (13.4.14) with $h(t)$. All alternative formulations of fractional convolution and correlation operations reported in this subsection are listed in the second and third columns of Table 13.4.3.

13.4.3 Fractional Autocorrelation and the Ambiguity Function

One of the fundamental properties of fractional autocorrelation is its relation to the ambiguity function (AF). The AF of $s(t)$ can be defined [6] as

$$AF_s(\tau,\nu) = \langle \mathbf{F}_{-\frac{\nu}{2}}\mathbf{T}_{-\frac{\tau}{2}}s, \mathbf{F}_{\frac{\nu}{2}}\mathbf{T}_{\frac{\tau}{2}}s\rangle = \int s(t+\frac{\tau}{2})\,s^*(t-\frac{\tau}{2})\,e^{-\jmath 2\pi\nu t}dt \quad (13.4.17)$$

where $(\mathbf{T}_\tau s)(t) = s(t-\tau)$ is the unitary time-shift and $(\mathbf{F}_\nu s)(t) = e^{\jmath 2\pi\nu t}s(t)$ is the unitary frequency-shift operator. A well-known property of the AF is that one can recover the LTI and LFrI autocorrelations in the fourth row of Table 13.4.2 by taking horizontal ($\phi = 0$) and vertical ($\phi = \frac{\pi}{2}$) cross sections of the AF, respectively [6];

$$(s \star_0 s)(\tau) = AF_s(\tau,0) \quad \text{and} \quad (s \star_{\frac{\pi}{2}} s)(\nu) = AF_s(0,\nu). \quad (13.4.18)$$

The relationship between the AF and fractional autocorrelation is given [4] as

$$(s \star_\phi s)(\rho) = AF_s(\rho\cos\phi, \rho\sin\phi). \quad (13.4.19)$$

This relationship generalizes the properties given in (13.4.18) for time ($\phi = 0$) and frequency ($\phi = \frac{\pi}{2}$) domain autocorrelations into other orientations of the AF. One can see from (13.4.19) that fractional autocorrelation at angle ϕ can also be recovered by taking a radial slice of the AF at angle ϕ in the ambiguity plane. Similarly, fractional convolution and fractional cross-correlation can be related to radial slices of the cross-AF as can be seen in the fourth column of Table 13.4.3.

13.4.4 Detection and Chirp Rate Parameter Estimation of Chirps

In this section a detection statistic is proposed for detection and chirp rate parameter estimation of chirp signals.

A continuous-time chirp can be formulated as $e^{\jmath 2\pi(\hat{\zeta}+\frac{\hat{m}}{2}t)t}$, where $\hat{m}$ and $\hat{\zeta}$ are the unknown chirp rate and initial frequency parameters, respectively. The AF of all chirps has a linear region of support that passes through the origin of the ambiguity plane [6, 7]. Since fractional autocorrelation corresponds to radial slices of the AF as in (13.4.19), a detection statistic based on fractional autocorrelation can be used for detection of chirps. By expressing the integral of the AF magnitude along the radial line with angle ϕ, the following detection statistic is derived [4],

$$L(m) = \int |(s \star_{\arctan(m)} s)(\rho)| \, d\rho. \tag{13.4.20}$$

Eq. (13.4.20) states that integration of the modulus of fractional autocorrelation with angle $\phi = \arctan(m)$ of the received signal $s(t)$ can be used for detection of chirp signals [4]. If the detection statistic, $L(m)$, exceeds a threshold for a certain chirp rate value, $\hat{m}$, then the detection statistic determines that a linear FM signal, with that particular chirp rate, is present in the received signal. The fast fractional autocorrelation algorithm via (13.4.14) does not require the calculation of the AF, and hence, is computationally efficient. In fact, if the number of chirp rates M, for which the detection statistic is calculated, is much less than the number of signal samples N, ($M << N$), then, computational savings provided by (13.4.20) over any detection statistic which requires calculation of the AF is considerable.

13.4.4.1 Simulation Examples

In this subsection, we test the detection statistic in (13.4.20) for detection and chirp rate parameter estimation of monocomponent and multicomponent chirp signals.

Example 1: In this simulation example, a discrete-time monocomponent chirp signal with chirp rate $\hat{m} = 0.35$ is corrupted with complex additive noise, $w[k]$, generated by two independent, zero-mean, Gaussian random processes of equal variance;

$$s[k] = e^{\jmath \hat{m} \frac{\pi}{1024} k^2} + w[k], \qquad k = 1, 2 \ldots, 512. \tag{13.4.21}$$

The normalized[2] detection statistic is calculated using (13.4.20) for different chirp rate values $m_l = 0.1 + (\frac{0.4}{200})l$, $l = 0, \ldots, 199$. The experiment was realized for the noise-free case, and for additive complex noise cases of -6 dB and -9 dB signal-to-noise ratio (SNR) values. Fig. 13.4.1 shows how the normalized detection statistic behaves in different levels of noise. It can be seen in Fig. 13.4.1 that the algorithm is able to detect the monocomponent chirp signal with the correct chirp rate parameter value, $\hat{m} = 0.35$, even in significant levels of noise.

[2]We normalize the detection statistic by the area under the received signal magnitude, i.e. $\frac{L(m)}{\sum_{k=1}^{N} |s[k]|}$, so that the detection threshold level is less affected by the noise level.

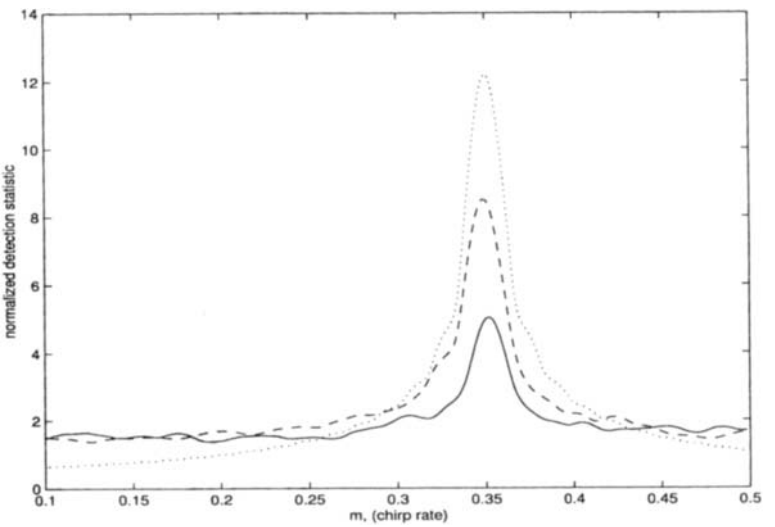


Fig. 13.4.1: Normalized (See Footnote 2) detection statistic. Dotted line: noise-free. Dashed line: SNR $= -6$ dB. Solid line: SNR $= -9$ dB.

Example 2: In this example, a discrete-time, multicomponent linear FM signal with initial frequency and chirp rate parameter values $\hat{\zeta}_0 = 6/1024$, $\hat{\zeta}_1 = 3/1024$, $\hat{\zeta}_2 = 2/1024$, $\hat{\zeta}_3 = 1/1024$, $\hat{m}_0 = 0.10$, $\hat{m}_1 = 0.20$, $\hat{m}_2 = 0.22$ and $\hat{m}_3 = 0.45$ is corrupted with complex additive noise, $w[k]$, generated by two independent, zero-mean, Gaussian random processes of equal variance;

$$s[k] = \sum_{i=0}^{3} e^{\jmath[2\pi\hat{\zeta}_i + \hat{m}_i \frac{\pi}{4096}k]k} + w[k], \qquad k = 1, 2 \ldots, 2048. \tag{13.4.22}$$

The normalized detection statistic is calculated via (13.4.20) with respect to different chirp rates $m_l = (\frac{0.5}{200})l$, $l = 0, \ldots, 199$. The experiment was done for the noise-free case, and for additive complex noise cases of -5 dB and -10 dB SNR values. As can be seen from Fig. 13.4.2, all four chirps are detected with their correct chirp rates. Two closely placed chirps with $\hat{m}_1 = 0.20$ and $\hat{m}_2 = 0.22$ are also resolved well.

13.4.5 Summary and Conclusions

In this article, a computationally efficient method based on fractional autocorrelation is proposed for detection and chirp rate parameter estimation of chirp signals that are utilized frequently in various applications of communications. Formulations of the fractional convolution and correlation operations associated with the fractional domains of the time-frequency plane are also presented. Those fractional domains are defined by the recently developed fractional Fourier transform (FRFT). The effectiveness of the proposed method is illustrated with simulation examples.

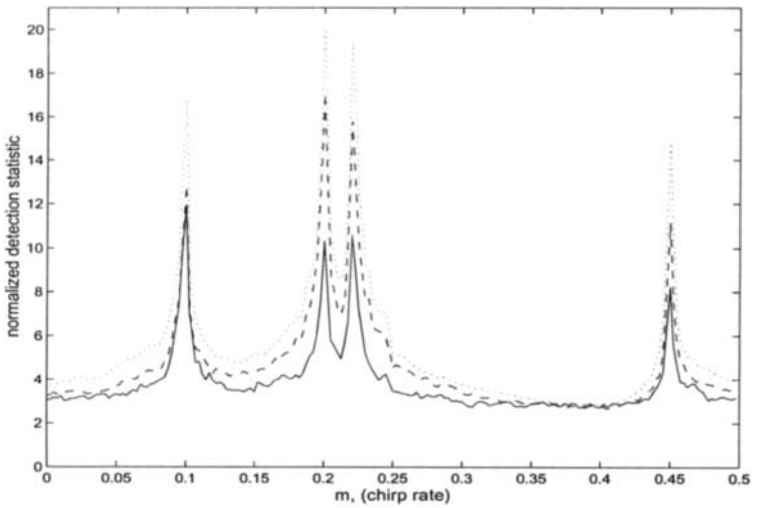


Fig. 13.4.2: Normalized (See Footnote 2) detection statistic. Dotted line: noise-free. Dashed line: SNR $= -5$ dB. Solid line: SNR $= -10$ dB.

Alternative presentations of the fractional Fourier transform are given in Articles 4.8 and 5.8.

References

[1] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Processing*, vol. 42, pp. 3084–3091, November 1994.

[2] H. M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural, "Convolution, filtering, and multiplexing in fractional Fourier domains and their relationship to chirp and wavelet transforms," *J. Optical Soc. of America A*, vol. 11, pp. 547–559, February 1994.

[3] O. Akay and G. F. Boudreaux-Bartels, "Unitary and Hermitian fractional operators and their relation to the fractional Fourier transform," *IEEE Signal Processing Letters*, vol. 5, pp. 312–314, December 1998.

[4] O. Akay and G. F. Boudreaux-Bartels, "Fractional convolution and correlation via operator methods and an application to detection of linear FM signals," *IEEE Trans. Signal Processing*, vol. 49, pp. 979–993, May 2001.

[5] H. M. Ozaktas, O. Arıkan, M. A. Kutay, and G. Bozdağı, "Digital computation of the fractional Fourier transform," *IEEE Trans. Signal Processing*, vol. 44, pp. 2141–2150, September 1996.

[6] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[7] B. Ristic and B. Boashash, "Kernel design for time-frequency signal analysis using the Radon transform," *IEEE Trans. Signal Processing*, vol. 41, pp. 1996–2008, May 1993.

This Page Intentionally Left Blank

Chapter 14

Time-Frequency Methods in Radar, Sonar & Acoustics

The fields of radar and sonar are traditionally key application areas and testing grounds for advances in signal processing. Time-frequency methodologies have made significant inroads already in these fields; their usefulness is demonstrated in five articles with appropriate internal cross-referencing to this and other chapters.

A baseband Doppler radar return from a helicopter target is an example of a persistent non-stationary signal. A linear time-frequency representation provides a high resolution suitable for preserving the full dynamic range of such complicated signals (Article 14.1). The synthetic aperture principle allows the combination of range resolution, achieved by the use of linear FM signals, with cross-range. For long observation intervals, the phase cannot be assumed to be a linear function of time; then time-frequency based imaging can obtain improvements in focus of the synthetic-aperture image (14.2). When a propeller-driven aircraft or a helicopter passes overhead, it produces a Doppler effect which allows the estimation of flight parameters by using IF estimation and the interference patterns formed in TFDs (14.3). To track a theater ballistic missiles launch, the WVD can be used effectively. Its peak provides a direct estimate of the instantaneous Doppler law giving the accelerating target dynamics (14.4). In sonar, there is a clear rationale for using time-frequency processing of returns to provide useful information about targets such as ships (14.5).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

14.1 SPECIAL TIME-FREQUENCY ANALYSIS OF HELICOPTER DOPPLER RADAR DATA [0]

14.1.1 Dynamic Range Considerations in TF Analysis

The classical technique for characterizing the time evolution behavior of sampled nonstationary signals has been the short-time Fourier transform (STFT), a time-frequency analysis technique that uses *linear* operations on sampled data. With momentary nonstationary signals of short duration, one often wishes to improve the time-vs-frequency resolving detail over that achievable with the STFT. Time-frequency distributions (TFDs) that use *quadratic* operations on sampled data, such as the Wigner-Ville distribution (WVD), are often able to achieve this improvement. For finite duration data records, this improvement in practice can be shown graphically to be approximately a factor of two in the time or frequency sharpening of the TFD response. However, this improvement is achieved in the presence of significant cross-term artifacts and with the sacrifice of detectable dynamic range (DNR). The cross terms, which are generated by multi-component signals (more than one signal present at an instant of time), additive noise, and analog-to-digital converter quantization effects, often obscure relevant signal components, particularly much weaker signal components (see Article 4.2). Numerous filtering and smoothing approaches have been developed to mitigate these cross-term artifacts, but this is always achieved at the expense of degrading the time-frequency sharpness of quadratic TFDs. Thus, quadratic TFDs tend to work best in mono-component signal situations.

The DNR between strongest and weakest signal components in practice can be quite significant. For example, modern radar systems, sonar systems, and medical Doppler ultrasound cardiovascular imaging systems can have DNRs that exceed 70 dB, 50 dB, and 100 dB, respectively. Signals from these systems almost always have multiple signal components present. Thus, quadratic TFD techniques are not good candidates for analyzing signals from such systems due to obscuring cross-term artifacts and dynamic range degradation. As this article will illustrate, it is still possible to achieve at least a factor of two improvement in the TFD response sharpness, without incurring the degradation caused by cross-term artifacts, by exploiting a pair of special two-dimensional (2-D) data arrays: the *windowed data function* (WDF) and the *complex WDF transform* (CWT). Application of variants of high-resolution 2-D spectral analysis techniques to the CWT will generate the desired high resolution TFD capability sought in quadratic TFDs, but without the cross-term artifacts.

14.1.2 Classical Linear and Quadratic TFDs

The baseline technique for comparing all time-frequency analysis methods is the STFT linear TFD. If $x(t)$ is the signal to be analyzed, define the short-time win-

[0]Author: **S. Lawrence Marple Jr.**, ORINCON Corporation, 9363 Towne Centre Drive, San Diego, California 92121 (lmarple@orincon.com). Reviewers: Joel M. Morris and Emin Tagluk.

dowed signal

$$x_h(t,\tau) = x(\tau)h^*(\tau - t), \tag{14.1.1}$$

in which $h(\tau)$ is the analysis window centered about $\tau = 0$, * denotes complex conjugation, and t represents the center time of the analysis interval on $x(t)$. The Fourier transform of the short-time windowed signal is therefore

$$F_x^h(t,f) = \underset{\tau \to f}{\mathcal{F}} \{x_h(t,\tau)\} = \int_{-\infty}^{\infty} x_h(t,\tau) \exp(-j2\pi f\tau)\, d\tau \quad . \tag{14.1.2}$$

which is the classical STFT. The localized STFT spectrum, or *spectrogram*, is simply the magnitude of the STFT

$$S_x^h(t,f) = \left|F_x^h(t,f)\right|^2 \tag{14.1.3}$$

which is then plotted to form the 2-D TFD gram. Trading off time-frequency concentration vs the resolution is achieved by the selection of the analysis window shape and duration.

The classical quadratic-based TFD method is the Wigner-Ville distribution $W_x(t,f)$ [1]. Most alternatives to the WVD are typically derivatives in which weighting or smoothing kernels are added to the functional definition of the quadratic TFD to better handle finite duration signal records or to suppress cross-term artifacts (see Chapter 3 and Article 5.7). As explained in Section 3.2.1, the WVD is linked to three companion quadratic functions: the *instantaneous autocorrelation function* (IAF) $K_x(t,\tau)$, the *spectral correlation function* (SCF) $k_x(\nu,f)$, and the *complex ambiguity function* (CAF) $A_x(\nu,\tau)$. The WVD can alternatively be computed as (1) the one-dimensional forward transform of the IAF, (2) the one-dimensional inverse transform of the SCF, or (3) the double transform of the CAF as follows

$$W_x(t,f) = \underset{\tau \to f}{\mathcal{F}} \{K_x(t,\tau)\} = \int_{-\infty}^{\infty} x(t+\tau/2)\, x^*(t-\tau/2)\, e^{-j2\pi f\tau} d\tau \tag{14.1.4}$$

$$= \underset{\nu \to t}{\mathcal{F}} \{k_x(\nu,f)\} = \int_{-\infty}^{\infty} X(f+\nu/2)\, X^*(f-\nu/2)\, e^{+j2\pi\nu t} d\nu \tag{14.1.5}$$

$$= \underset{\nu \to t}{\mathcal{F}} \underset{\tau \to f}{\mathcal{F}} \{A_x(\nu,\tau)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_x(\nu,\tau)\, e^{j2\pi[f\tau - \nu t]} d\nu\, d\tau. \tag{14.1.6}$$

in which $X(f)$ is the Fourier transform of $x(t)$, t is the "localized" center time, τ is the relative separation (shift) time displacement, f is the "localized" center frequency, and ν is the relative separation (shift) frequency displacement. More details on these relationships may be found in Part I. Fig. 14.1.1 depicts the key mathematical definitions and relationships among these four quadratic functions. Fig. 14.1.1 also graphically depicts the support regions of the quadratic functions for a finite-duration signal of N samples $x(n\Delta T) = x[n]$ at ΔT second intervals. Note that the temporal and spectral correlation functions have diamond-shaped support regions with time-shift and frequency-shift axes, respectively, that are twice the

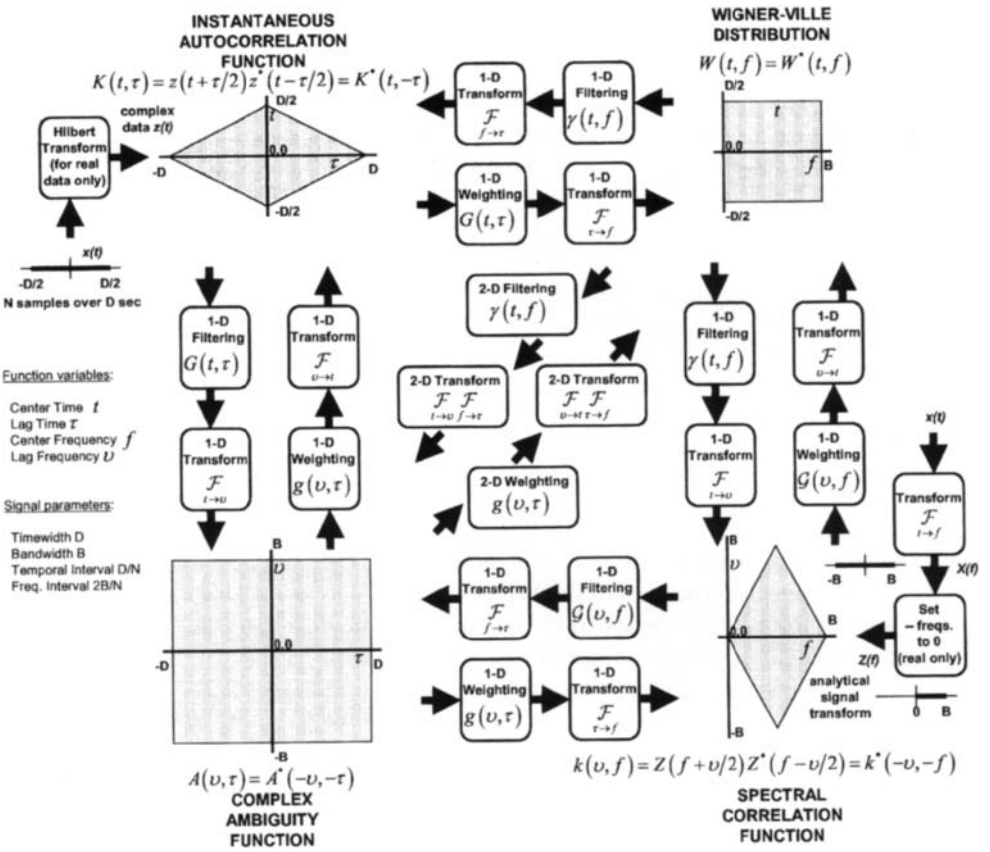


Fig. 14.1.1: Flow diagram and key relationships and support domains among the four quadratic representations: Wigner-Ville distribution, complex ambiguity function, instantaneous autocorrelation function, and spectral correlation function.

original signal time-width and its transform band-width, due to the correlation processes used to form the IAF and SCF. Having twice the temporal or spectral extents will yield, after transformation, WVD responses that are twice as sharp relative to the original signal's time extent or bandlimited frequency extent. Note that the temporal and frequency resolutions of the WVD are greatest at $t = 0$ and $f = 0$ and decrease as one advances to the outer time and frequency edges of the IAF or the SCF diamond-shaped regions of support.

In order to mitigate the deleterious effects of cross terms, a number of optional smoothing and filtering operations, marked in Fig. 14.1.1 as kernels $g, G, \gamma, \mathcal{G}$, have been developed to suppress the cross-term effects (see Articles 5.2 and 6.1 for more details). However, suppressing the cross terms will also degrade the time-frequency concentration; therefore, all kernel designs involve a tradeoff between minimizing

cross-term artifacts and maximizing time-frequency resolution.

14.1.3 Alternative High-Resolution Linear TFD

Let us revisit the 2-D STFT and develop a new perspective for creating the 2-D STFT spectrogram from a 1-D data record. Based on this new perspective, we will create an alternative 2-D matrix array using linear transform operations on data samples from which an opportunity to apply high-resolution 2-D spectral analysis procedures is generated. If one assumes a finite-width analysis window of M samples for a case of finite-duration data of N samples, and further assumes the analysis window center time is stepped through the data one sample at a time (multiple sample steps are also possible but not discussed here), then the operations to be performed on the data will start with the data matrix

$$\mathbf{x}(t,\tau) = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 & x[1] \\ \vdots & \iddots & & \iddots & x[1] & x[2] \\ 0 & & 0 & \iddots & & \vdots \\ \vdots & \iddots & x[1] & & & x[(M-1)/2] \\ 0 & \iddots & & & & \vdots \\ x[1] & & & & & x[N] \\ \vdots & & & & \iddots & 0 \\ x[N-(M-1)/2] & & & x[N] & \iddots & \vdots \\ \vdots & & \iddots & 0 & & 0 \\ x[N-1] & x[N] & \iddots & & \iddots & \vdots \\ x[N] & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \tag{14.1.7}$$

in which the analysis center time $t = m\Delta T$ corresponds to the row index m and the data samples within an analysis window interval are all the elements along a selected row, indexed by their column time $\tau = n\Delta T$. Note that the above data matrix has $N + M - 1$ rows corresponding to the center times ranging from $t = -(M+1)/2$ to $N + (M+1)/2$, which requires zero fill when the analysis window runs off the ends of the available data. Other ranges along the vertical axis are also possible to define such that zero fill is not required.

Noting the analysis window values as $h[1]$ to $h[M]$, the resultant *windowed data function* (WDF) is obtained element-by-element as

$$x_h(t,\tau) = x(\tau)h^*(\tau - t) \quad . \tag{14.1.8}$$

to form the WDF matrix $H(t,\tau)$ shown in Fig. 14.1.2. The STFT is then formed by taking the one-dimensional Fourier transform of each row of the WDF

$$F_x^h(t,f) = \underset{\tau \to f}{\mathcal{F}} \{H(t,\tau)\} \tag{14.1.9}$$

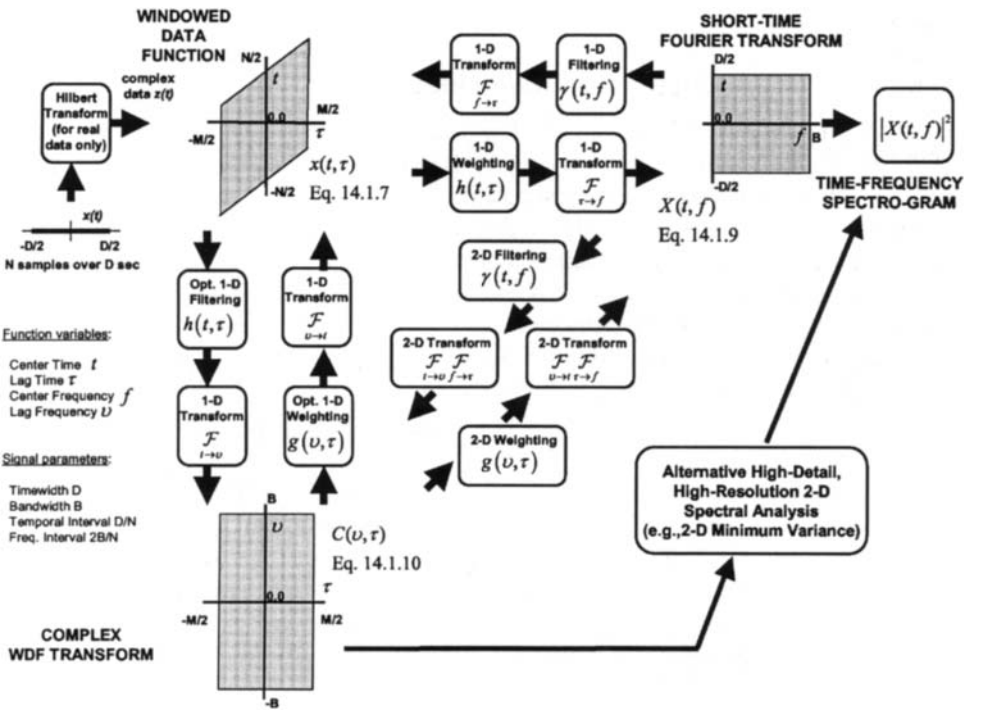


Fig. 14.1.2: Flow diagram of key relationships among windowed data function (WDF), complex WDF transform (CWT), short-time Fourier transform (STFT), and short-time Fourier transform spectrogram. Insertion of alternative 2-D spectral technique possibilities is indicated.

as indicated in Fig. 14.1.2.

Although counterintuitive, one could alternatively take the Fourier transform of each column (rather than each row) of the WDF to produce the *complex WDF transform* (CWT)

$$C(\nu,\tau) = \underset{t\to\nu}{\mathcal{F}}\{H(t,\tau)\} \tag{14.1.10}$$

as shown in Fig. 14.1.2. It may seem computationally disadvantageous to form the CWT, but the motivation is found in the similarity of the CWT to the CAF of Fig. 14.1.1. A double transform of the CWT yields the STFT as shown in Fig. 14.1.2, just as a double transform of the CAF yields the WVD in Fig. 14.1.1. By computing the CWT, we now open the possibility of creating a TFD gram via an alternative high-definition 2-D spectral analysis technique in lieu of the squared-magnitude STFT spectrogram. Because no cross terms are created in the formation of the CWT, there will be no cross-term artifacts introduced into the 2-D spectral analysis results.

Examples of high-definition 2-D spectral estimation procedures that may be

applied here include the 2-D autoregressive [2] and 2-D minimum variance techniques [3]. The 2-D data set that is used as the input to these methods is the CWT array. The final spectral analysis formula for generating the time-vs-frequency analysis gram will be a modification of the normal 2-D spectral analysis formula in order to account for the time-vs-frequency units of the CWT (rather than time-vs-time). For example, the 2-D autoregressive spectral formula is modified to have the structure

$$S_{\mathrm{TFAR}}(t,f) = \frac{1}{\left|\sum_m \sum_n a_{\nu,\tau}[m,n] \exp(-j2\pi[-m\Delta F t + n\Delta T f])\right|^2} \qquad (14.1.11)$$

in which the two-dimensional autoregressive parameters $a_{\nu,\tau}[m,n]$ are estimated from the CWT array elements, ΔT is the time-increment along the CWT rows, ΔF is the frequency-increment along the CWT columns, t represents the center time of the analysis evaluation, and f represents the center frequency of the analysis evaluation. Note the opposite signs of the time and frequency variables in the above equation, which contrast with the usual 2-D formulae that have the same sign. The modified 2-D spectral procedures are more appropriately termed the 2-D *time-vs-frequency autoregressive* (TFAR) and the 2-D *time-vs-frequency minimum variance* (TFMV) methods for TFD gram creation from the CWT array. Fast computational algorithms for 2-D autoregressive parameter estimation may be found in Marple [2,4] and for the 2-D minimum variance spectral function in Jakobsson et al. [3].

14.1.4 Application to Simulated and Actual Data

To illustrate the benefit of the new 2-D high-resolution approach to time-frequency analysis, consider the various TFDs shown in Fig. 14.1.3 for a simulated complex-valued waveform (has in-phase I and quadrature-phase Q components) of 256 samples consisting of two criss-crossing chirp (linear frequency modulated) signals and a sinusoidal FM modulation. The traditional stationary processing approach would be to form the assumed stationary correlation estimate using the entire 256-point data record, followed by Fourier transformation of this estimated correlation to create the long-term Fourier transform (LTFT) spectrum shown at the bottom of Fig. 14.1.3(a) (note that a logarithmic scale in dB is used to capture the full dynamic range of the spectral response). The I and Q temporal plots of the signal are drawn on the left and right sides of Fig. 14.1.3(a), respectively. Note that the LTFT spectral shapes of the criss-crossing chirps and the sinusoidal FM component are essentially the same, and thus the two signals are almost indistinguishable if the LTFT alone is used for signal classification.The three signals are distinguishable when the short-term Fourier transform spectrogram is used, as illustrated in Fig. 14.1.3(a) using a color-coded linear plot and in Fig. 14.1.3(c) using a log plot of the top 50 dB. The nonstationary frequency-varying character of the three signal components is now apparent. The Wigner-Ville TFD estimate is plotted with linear units in Fig. 14.1.3(b) and with log units in Fig. 14.1.3(d). The promise of the WVD to increase sharpness in the time-vs-frequency localization is apparent, but this has

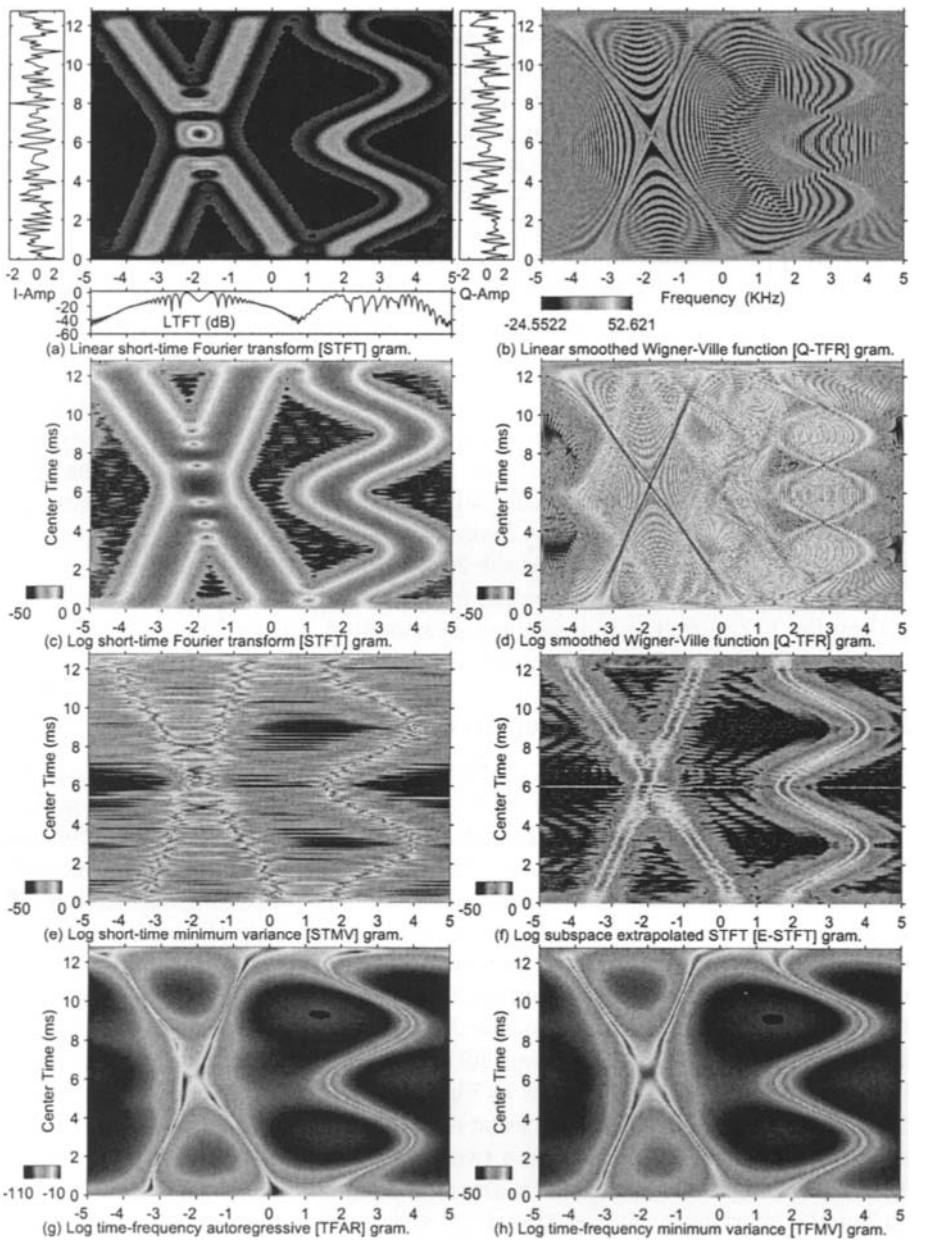


Fig. 14.1.3: Comparison of various time-frequency analysis techniques to extract the 2-D localized nonstationary features of a signal with three components: two criss-crossing chirps and one FM modulated waveform.

been achieved at the price of introducing significant cross-term artifacts, especially when emphasized by the logarithmic gram of Fig. 14.1.3(d). Four alternatives to STFT and WVD characterizations of the nonstationary behavior of the three signals are illustrated in Figs. 14.1.3(e-h). These include the TFAR and TFMV. These last two alternatives produce results as sharp as those obtained with WVD without the cross-term artifacts.

Consider the three TFD grams of Figs. 14.1.4–14.1.6 produced from a 330 ms record of baseband Doppler radar returns from a helicopter target. This is a case of a persistent nonstationary signal, in contrast with the more typical momentary nonstationary signals that characterize much of the time-frequency analysis literature. We shall use a sliding analysis window in this case, computing a single line of the TFD with each increment of the sliding analysis window. The helicopter target illuminated by the X-band CW radar was a two-engine Eurocopter Deutschland BO-105, as depicted in the inset of Fig. 14.1.4. The radar signal consisted of samples of the complex-valued (I/Q) baseband (carrier frequency demodulated down to 0 Hz) waveform; the temporal waveform samples are plotted as vertical displays on the left and right gram sides in Figs. 14.1.4–14.1.6. The helicopter is in motion, so the radar signal will have Doppler components due to the forward motion of the fuselage, the main rotor and tail rotor rotations, and multipath components between the fuselage and rotors. A LTFT spectral estimate of the entire 330 ms data record produced the spectrum plotted beneath each TFD gram in Figs. 14.1.4–14.1.6. Note that there is at least 70 dB of dynamic range between strongest and weakest signal components in this radar signal.

Fig. 14.1.4 is the result of a classical STFT analysis of the data record which has been adjusted to use sliding 128-sample Hamming analysis windows to best capture the most rapidly-varying Doppler components (main rotor) of the radar signal. Note that positive frequencies represent Doppler components moving toward the radar and negative frequencies represent Doppler components moving away from the radar in this baseband signal. Also note the color coding assignments of the logarithmic dynamic range that is is displayed; approximately the top 70 dB of each TFR gram is displayed. The dominant signal component is the +2.8 kHz line that represents the Doppler component off the helicopter fuselage due to the constant velocity motion of the helicopter toward the radar. Other features that can be seen in the STFT gram of Fig. 14.1.4 include the sinusoidal patterns bounded between -12 kHz and +18 kHz that correspond to the Doppler pattern of the four main rotor blades and the periodic broadband horizontal "flashes" corresponding to the periodic alignment of the main or tail rotors to maximally reflect the radar beam. Spectral lines at ±500 Hz also appear due to a 500 Hz modulated time tone that was added to the data tape.

Using the same sliding 128-sample analysis window increments as the STFT, the WVD gram shown in Fig. 14.1.5 was produced. A Choi-Williams smoothing filter was employed using a exponential parameter factor of 0.3 to mitigate the cross-term artifacts. Although there is an improved sharpness relative to the STFT spectrogram, the cross-term mitigation is insufficient as these terms are still much

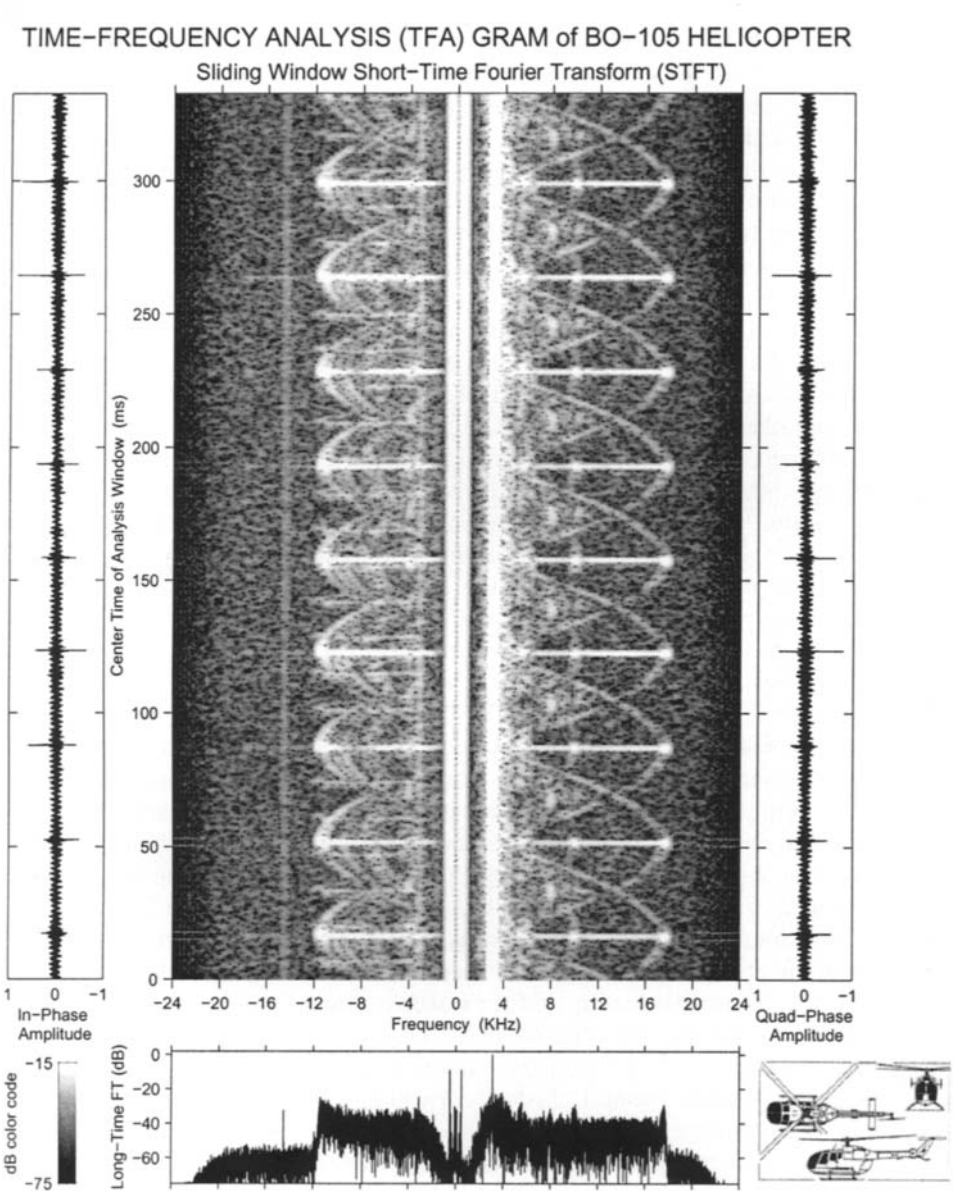


Fig. 14.1.4: Classical short-time Fourier transform (STFT) time-frequency analysis of 330 ms of complex (I/Q) baseband Doppler radar signal of a helicopter target in motion. Profile of BO-105 helicopter target is shown in lower right inset.

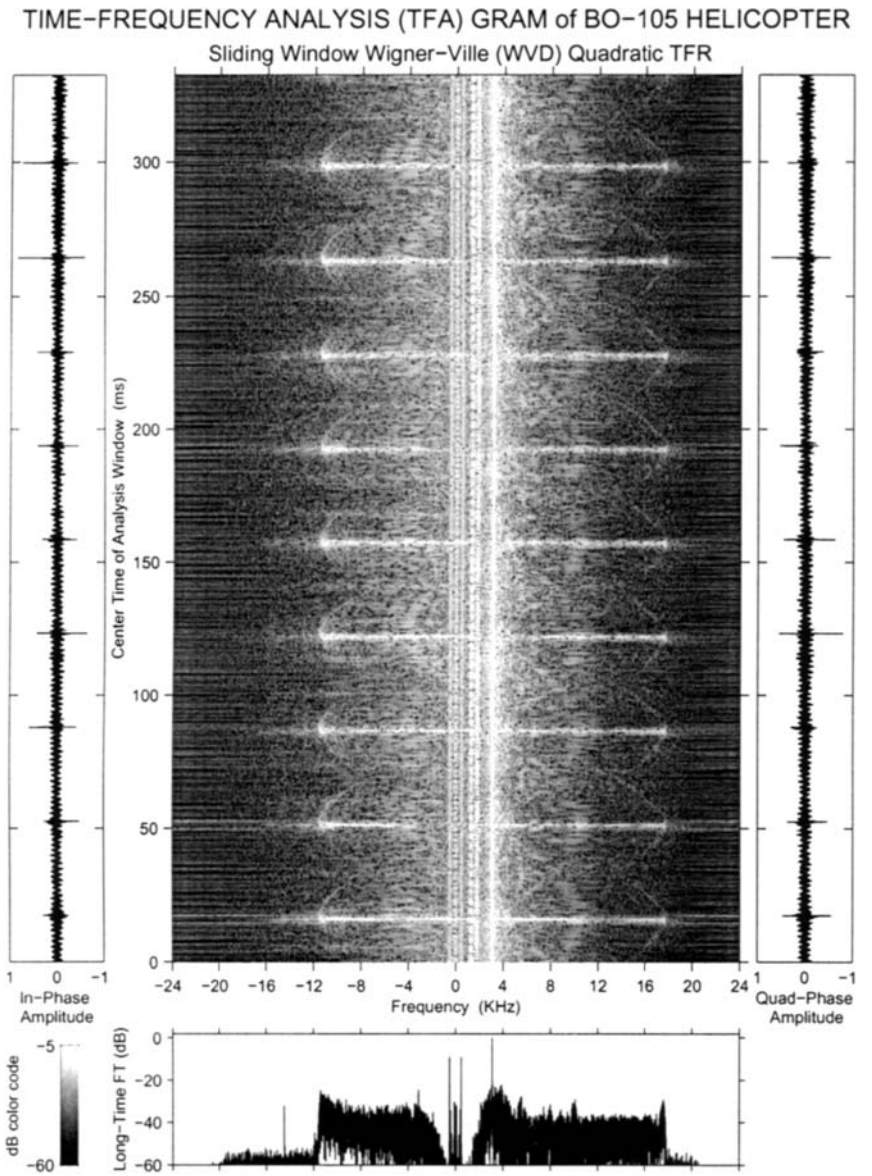


Fig. 14.1.5: Time-vs-frequency gram of Doppler radar signal using the Wigner-Ville distribution with sliding window and complex ambiguity function weighting using a Choi-Williams kernel.

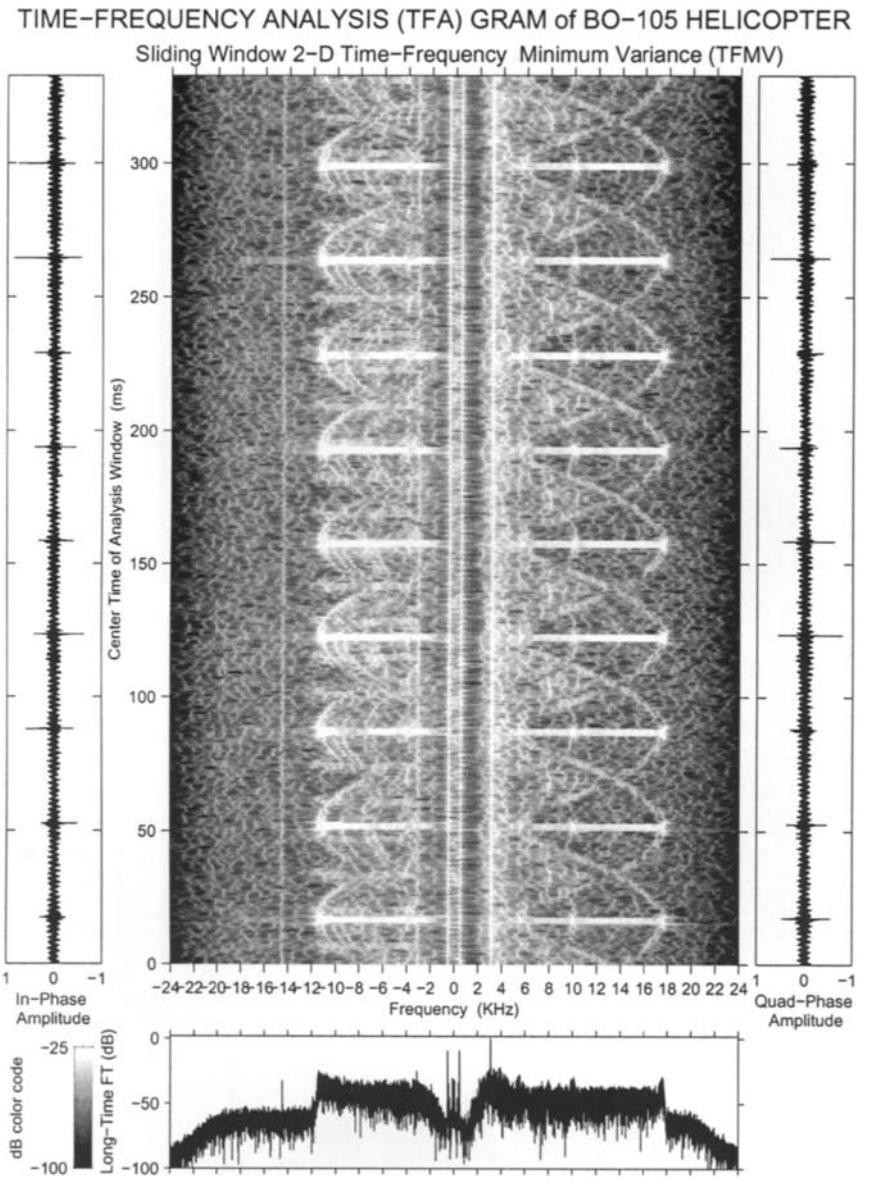


Fig. 14.1.6: Time-vs-frequency gram of Doppler radar signal using the two-dimensional minimum variance method with sliding window.

stronger than the weaker Doppler components that we are attempting to detect and to track. Fig. 14.1.6 is the TFD gram resulting from use of the TFMV technique.

14.1.5 Summary and Conclusions

The alternative high-resolution methods presented in this article clearly are capable of preserving the full dynamic range of features in complicated signals being analyzed while achieving sharp time-frequency responses comparable to those found in the WVD gram. These attributes are achieved without the artificial introduction of cross-term artifacts typical of some of the quadratic TFD methods.

References

[1] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner Distribution—A tool for time-frequency signal analysis," *Philips J. of Research*, vol. 35, pp. 217–250 (Part 1), 276–300 (Part 2) & 372–389 (Part 3), 1980.

[2] S. L. Marple Jr., "Two-dimensional lattice linear prediction parameter estimation method and fast algorithm," *IEEE Signal Processing Letters*, vol. 7, pp. 164–168, June 2000.

[3] A. Jakobsson, S. L. Marple Jr., and P. Stoica, "Computationally efficient two-dimensional Capon spectrum analysis," *IEEE Trans. Signal Processing*, vol. 48, pp. 2651–2661, September 2000.

[4] S. L. Marple Jr., *Digital Spectral Analysis with Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

14.2 TIME-FREQUENCY MOTION COMPENSATION ALGORITHMS FOR ISAR IMAGING [0]

The synthetic aperture principle is now a well established concept and has given rise to several applications in high resolution microwave imaging. Among the most impressive applications is the imaging of Venus' surface provided by the synthetic aperture radar (SAR) aboard the spacecraft *Magellan* (see e.g. `http://www.jpl.nasa.gov/magellan`). Venus' surface could not be observed by optical telescopes because of the thick layers of clouds covering the planet. It was thus necessary to use lower frequency waves to penetrate the clouds, but this would have come at the expense of resolution. The radar aboard the *Magellan* orbiter made it possible to send and receive electromagnetic waves through Venus' atmosphere and still maintain good resolution: the range resolution was achieved by transmitting a wideband impulsive linear FM (chirp) signal, whereas the cross-range resolution was obtained by exploiting the synthetic aperture principle.

The so-called **synthetic aperture** is formed by gathering the echoes collected by the radar platform during the interval in which the radar antenna beam insists on a specific site, and processing them coherently. The instantaneous phases of the echoes are compensated as to mimic the presence of a real antenna gathering the same echoes. If the phase compensation is perfect, we can consider the compensated signal as gathered by a synthetic aperture whose along-track length is equal to the distance traveled by the radar during the time interval within which each point on the ground is illuminated by the radar antenna. Clearly the relative motion between radar and observed scene, or target,[1] is fundamental to synthesizing an aperture wider than the real aperture.

The synthesis of the wide aperture requires an accurate knowledge of the temporal evolution of the phase shift induced by the relative radar/target motion. In formulas, if the radar sends an unmodulated sinusoid of frequency f_0, the echo from a point-like target whose distance from the radar is $r(t)$ assumes the form

$$y(t) = a(t)\, e^{j2\pi f_0 (t-2r(t)/c)} = a(t)\, e^{j2\pi f_0 t}\, e^{-j4\pi r(t)/\lambda}, \qquad (14.2.1)$$

where c is the speed of light, $\lambda = c/f_0$ is the transmission wavelength, and $a(t)$ is the amplitude modulation.

In the imaging of stationary scenes observed by a radar carried on a spacecraft or an aircraft, the variation of $r(t)$ is entirely due to the radar and an initial coarse estimation of $r(t)$ comes from the navigational system of the vehicle carrying the radar. However, in applications where the observed target is also moving, as in the so called Inverse SAR (ISAR), the distance $r(t)$ is not known at the radar site. Furthermore, also in cases where the target is not moving, the accuracy provided by the

[0] Author: **Sergio Barbarossa**, INFOCOM Department, University of Rome "La Sapienza", Via Eudossiana 18, 00184 Rome, Italy (sergio@infocom.uniroma1.it). Reviewers: M. Mesbah and G. R. Putland.

[1] In this article, the "target" means the object of the radar imaging.

navigational instruments may be insufficient, especially for high resolution imaging. In all these cases, it is necessary to estimate the phase history $\phi(t) := 4\pi r(t)/\lambda$ directly from the received data [1]. This requires the presence of prominent scatterers on the target, whose radar echoes are sufficiently stronger than the background to allow a reliable estimate. The situation is complicated, however, when more dominant scatterers occupy the same range cell, as is likely in many practical circumstances. In fact, if no *a priori* knowledge of the relative motion is available, it is safer not to use a high range resolution at the beginning, simply to avoid the range migration problem. But the more we decrease the range resolution, the higher is the probability of observing more dominant scatterers in the same range cell. In these cases, as proposed in [2], an iterative procedure can be followed, based on the following steps: (i) start using low range resolution data, for example by smoothing along the range direction, to avoid the range resolution problem; (ii) estimate the instantaneous phase of the dominant scatterer; (iii) estimate the relative motion law $r(t)$ of the dominant scatterer from the instantaneous phase; (iv) use the estimated motion law to compensate for the range migration occurring in the high range resolution data (hence the dominant scatterer automatically becomes the image center); (v) analyze the full range resolution data after compensation of the range migration.

Clearly step (i) increases the probability of having more dominant scatterers in the same range cell and, if the observed signal is really multi-component, the analysis becomes more complicated because we must separate the components corresponding to each echo before estimating their instantaneous phases. The separation is simple if the components have linear phase. In such a case, in fact, an FFT-based approach may be sufficient. However, especially when long observation intervals are used to obtain high cross-range resolution, the phase cannot be assumed to be a linear function of time. Time-frequency representations of the observed signals thus provide an important analysis tool in such cases. In this article we illustrate a method for separating the signal components and estimating their instantaneous phases using the so called Smoothed Pseudo Wigner-Ville Distribution with Reassignment (RSPWVD), described in Article 7.2 and originally proposed in [3], followed by a parametric estimation method. Intuitively speaking, the time-frequency analysis is used to get an initial idea about the kind of received signals. Based on this preliminary analysis, we can assume a parametric model and then use parametric estimation methods to improve the estimation accuracy. One possible method for combining TFDs and parametric modeling for retrieving the parameters of multi-component signals, with instantaneous frequencies (IFs) modeled as sinusoids or hyperbolas, was suggested in [4]; but clearly many alternatives can be followed to optimize the performance of the estimation method.

14.2.1 Echo from a Rotating Rigid Body

We assume that the target is a rigid body in the far field of the radar antenna and is characterized by a certain number of dominant scatterers. For a transmission

frequency $f_o = c/\lambda$, the echo from the k-th scatterer is

$$A_k\, e^{j2\pi f_o(t-2|\boldsymbol{r}_o-\boldsymbol{r}'_k(t)|/c)}, \qquad k=0,\ldots,K-1 \tag{14.2.2}$$

where $\boldsymbol{r}_o$ is the vector indicating the radar position, $\boldsymbol{r}'_k(t)$ indicates the k-th scatterer and K is the number of scatterers. Under the far field hypothesis, the echo can be approximated as

$$A_k\, e^{j2\pi f_o t}\, e^{-j4\pi R_o/\lambda}\, e^{j4\pi \hat{\boldsymbol{r}}_o.\boldsymbol{r}'_k(t)/\lambda} \tag{14.2.3}$$

where $R_o = |\boldsymbol{r}_o|$. The motion of a rigid body can always be expressed as the composition of the translation of one of its points plus the rotation of the body around that point. Every imaging or classification procedure must apply some kind of motion compensation and, in general, the translational motion is compensated first. This operation is performed by multiplying the radar echo by a reference signal matched to the echo from one dominant scatterer, and resampling the data in range to remove any range migration of the scatterer assumed as a reference. Taking as a reference the echo from the 0-th scatterer in (14.2.3) (setting $k=0$ and $A_0 = 1$), the signals after compensation take the form

$$A_k\, e^{j4\pi \hat{\boldsymbol{r}}_o.\boldsymbol{q}_k(t)/\lambda}, \qquad k=1,\ldots,K-1, \tag{14.2.4}$$

where $\boldsymbol{q}_k(t) := \boldsymbol{r}'_k(t) - \boldsymbol{r}'_0(t)$. Under the rigid-body constraint, the vectors $\boldsymbol{q}_k(t)$ can only rotate and the rotation matrix is the same for all points belonging to the target. The differential equation characterizing the rotation of the generic vector $\boldsymbol{q}_k(t)$ is

$$\frac{d\boldsymbol{q}_k(t)}{dt} = \boldsymbol{\omega}(t) \times \boldsymbol{q}_k(t)\,, \tag{14.2.5}$$

where $\boldsymbol{\omega}(t)$ is the vector containing the instantaneous pitch, roll and yaw pulsations $(\omega_p(t), \omega_r(t), \omega_y(t))$, and $\times$ denotes the vector (cross) product. Assuming a constant pulsation, i.e. $\boldsymbol{\omega}(t) = \boldsymbol{\omega}$, and indicating by Ω its modulus, i.e. $\Omega = \sqrt{\omega_p^2 + \omega_r^2 + \omega_y^2}$, and given an initial position $\boldsymbol{q}_k(0) = \boldsymbol{q}_k$ of the vector at time $t_0 = 0$, the solution of equation (14.2.5) is

$$\boldsymbol{q}_k(t) = \boldsymbol{a}_k + \boldsymbol{b}_k \cos(\Omega t) + \frac{\boldsymbol{c}_k}{\Omega} \sin(\Omega t), \tag{14.2.6}$$

where the vectors $\boldsymbol{a}_k$, $\boldsymbol{b}_k$, $\boldsymbol{c}_k$ are

$$\boldsymbol{a}_k = \frac{(\boldsymbol{\omega}.\boldsymbol{q}_k)}{\Omega^2}\boldsymbol{\omega}\,, \quad \boldsymbol{b}_k = \boldsymbol{q}_k - \frac{(\boldsymbol{\omega}.\boldsymbol{q}_k)}{\Omega^2}\boldsymbol{\omega}\,, \quad \boldsymbol{c}_k = \boldsymbol{\omega} \times \boldsymbol{q}_k. \tag{14.2.7}$$

Hence the echo from the generic k-th scatterer is

$$A_k\, e^{j4\pi \hat{\boldsymbol{r}}_o.\boldsymbol{q}_k(t)/\lambda} = A_k\, e^{j4\pi(m_k \cos(\Omega t+\phi_k)+\alpha_k)/\lambda}, \tag{14.2.8}$$

where

$$m_k = \sqrt{(\hat{\boldsymbol{r}}_o.\boldsymbol{b}_k)^2 + \frac{(\hat{\boldsymbol{r}}_o.\boldsymbol{c}_k)^2}{\Omega^2}}\,, \quad \phi_k = \arctan\left(\frac{\hat{\boldsymbol{r}}_o.\boldsymbol{c}_k}{\Omega \hat{\boldsymbol{r}}_o.\boldsymbol{b}_k}\right)\,, \quad \alpha_k = \hat{\boldsymbol{r}}_o.\boldsymbol{a}_k. \tag{14.2.9}$$

It is important to notice that the instantaneous phase of each echo contains a constant term plus a sinusoidal contribution having the *same frequency* for all the scatterers, but different amplitudes and initial phases. This is a consequence of the rigid-body constraint.

14.2.2 Signal Analysis based on Time-Frequency Representations

The general model for the frequency modulation induced by the relative radar-target motion can always be decomposed into the sum of a slow component, well approximated by a low-order polynomial, plus a possible fast component having a sinusoidal behavior. The echo can then be modeled as

$$x(t) = Ae^{j\sum_{m=0}^{M} a_m t^m} e^{j\beta\cos(\Omega t+\Phi)} + w(t), \tag{14.2.10}$$

where $w(t)$ is additive noise. The slow component is mainly due to the translation; the fast component depends on the rotation. If the sinusoidal component has a period ($T = 2\pi/\Omega$) much longer than the duration (T_0) of the observation interval, it can also be approximated as a low-order polynomial, so that the overall signal can be modeled as a polynomial-phase signal (PPS). Conversely, if T is much less than T_0, it is better to estimate the parameters of the sinusoidal component without any polynomial modeling.

As the velocity of variation is not known *a priori*, it has to be estimated from the data using, at least initially, a non-parametric approach. We used, as a preliminary tool, the Reassigned Smoothed Pseudo Wigner-Ville Distribution (RSPWVD) for its good localization and low cross-terms (see Article 7.2).

As an example, in Fig. 14.2.1 we report the RSPWVD of the echo from a rotating object with two dominant scatterers having the same backscattering coefficients. The two components are clearly visible and show the same period.

We can use the RSPWVD to extract an initial estimate of the period Ω and then exploit the parametric modeling, as shown in the following section. One possibility is to extract the peaks of the TFD, as in [5] and Articles 10.3 and 10.4, and then estimate the modulating frequency of the IF.

14.2.3 Parametric Estimation of Instantaneous Phases

With the aid of the echo modeling illustrated in Section 14.2.1, we can improve the performance of the estimation method over that obtainable from the simple use of the RSPWVD. In principle, considering signals expressed as in (14.2.8), we could smooth the WVD by integrating it over all possible sinusoids in the time-frequency plane, as a function of the sinusoidal parameters, using for example a generalized Hough Transform as in [4]. However this operation would be quite troublesome from the computational point of view. In fact such an operation is almost equivalent to computing the square modulus of the scalar product of the received signal and the signal model (14.2.8), with $A_k = 1$. Hence the detection and estimation of FM signals satisfying (14.2.8) can be carried out by searching for the peaks of the

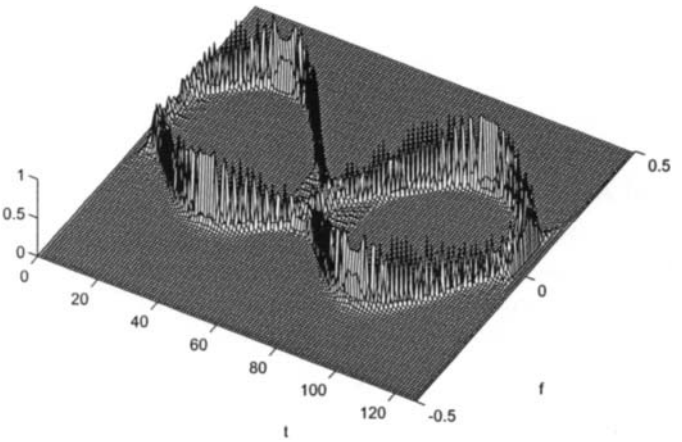


Fig. 14.2.1: Reassigned smoothed pseudo-Wigner-Ville distribution of the sum of two echoes from a rotating object.

function

$$P(m,\Omega,\phi) = \left| \sum_{n=0}^{N-1} x(n)\, e^{-j4\pi m \cos(\Omega n+\phi)/\lambda} \right|^2 , \tag{14.2.11}$$

where N is the number of samples: if a peak exceeds a suitable threshold, a sinusoidal FM signal is detected and its parameters are estimated as the coordinates of the peak. Of course the straightforward application of (14.2.11) is also quite troublesome from the computational point of view. Observing, however, that all signal components have the same pulsation Ω, we can estimate Ω from the RSPWVD and then use the estimated value $\hat{\Omega}$ in (14.2.11) to compute the $2D$ function

$$P(m,\phi) = \left| \sum_{n=0}^{N-1} x(n) e^{-j4\pi m \cos(\hat{\Omega} n+\phi)/\lambda} \right|^2 . \tag{14.2.12}$$

An example is shown in Fig. 14.2.2 for the case where three sinusoidal FM signals occupy the same range cell. Fig. 14.2.2 shows the function $P(m,\phi)$ given in (14.2.12). We can clearly observe the presence of three peaks. Comparing $P(m,\phi)$ with a suitable threshold, we detect the presence of dominant scatterers and estimate, at the same time, their modulation parameters m_k and ϕ_k as the coordinates of the peaks which exceed the threshold. A further simplification of the proposed procedure consists in analyzing time intervals smaller than the rotation period. In these sub-intervals, the instantaneous phases can be approximated by polynomials (the first terms of their Taylor series expansions). In this case we can use specific algorithms devised for the detection and parameter estimation of multicomponent polynomial phase signals embedded in noise, based on the so called high order ambiguity function [6].

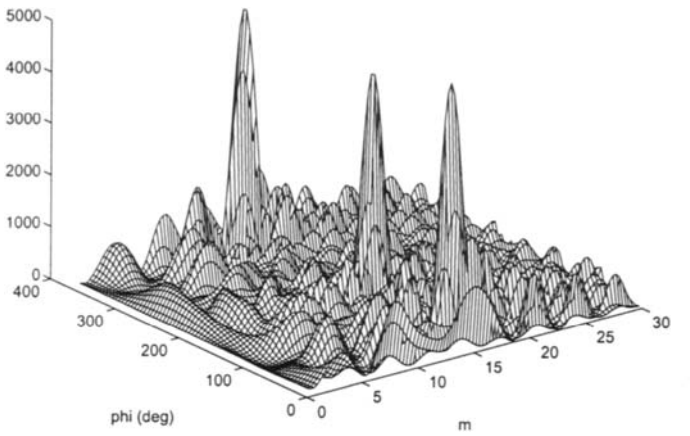


Fig. 14.2.2: $P(m, \phi)$ of a three-component sinusoidal FM signal.

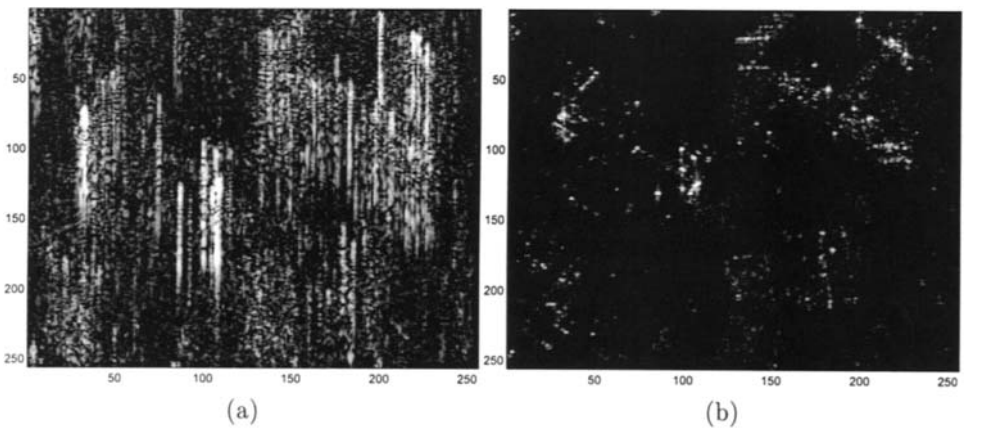


Fig. 14.2.3: Example of imaging: (a) Image defocused with a third-order polynomial instantaneous phase; (b) Focused image.

An example of application to real SAR images is given in Fig. 14.2.3, where we show a SAR image defocused by a third-order polynomial (left side) and the image focused by estimating the polynomial parameters and then compensating the instantaneous phase (right side).

The time-frequency analysis coupled with the parametric method is especially important when the illuminated spot contains moving targets. In such a case, the conventional SAR processing would provide a focused image of the stationary background and a defocused image of the moving object. Indeed, the only way to discriminate moving targets echoes from the background is to analyze their instantaneous frequencies: a different motion law must give rise to a different frequency modulation. Some examples of applications of this idea were shown in [7].

14.2.4 Summary and Conclusions

Estimating the frequency modulations of the radar echoes is a fundamental step in the formation of high-resolution images. The tool is especially important in the presence of targets moving on a stationary background, where the only possibility for discriminating moving from fixed targets involves the analysis of their instantaneous frequencies. As the observed signal often comes from multiple scatterers, the availability of TFDs with good localization and low cross-terms is particularly important. The other key idea is that, given a time-frequency representation of the signal of interest, we may derive a parametric model of the signal and then use parametric estimation techniques to improve the estimation accuracy.

References

[1] W. G. Carrara, R. S. Goodman, and R. M. Majewski, *Spotlight Synthetic Aperture Radar*. Norwood, MA: Artech House, October 1995.

[2] S. Barbarossa, "Detection and imaging of moving objects with synthetic aperture radar—Part 1: Optimal detection and parameter estimation theory," *IEE Proc., Part F: Radar & Signal Processing*, vol. 139, pp. 79–88, February 1992.

[3] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[4] S. Barbarossa and O. Lemoine, "Analysis of nonlinear FM signals by pattern recognition of their time-frequency representation," *IEEE Signal Processing Letters*, vol. 3, pp. 112–115, April 1996.

[5] 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 Processing*, vol. 47, pp. 2480–2490, September 1999.

[6] S. Barbarossa, A. Scaglione, and G. B. Giannakis, "Product high-order ambiguity function for multicomponent polynomial-phase signal modeling," *IEEE Trans. Signal Processing*, vol. 46, pp. 691–708, March 1998.

[7] S. Barbarossa and A. Scaglione, "Autofocusing of SAR images based on the product high-order ambiguity function," *IEE Proc.: Radar, Sonar & Navigation*, vol. 145, pp. 269–273, October 1998.

14.3 FLIGHT PARAMETER ESTIMATION USING DOPPLER AND LLOYD'S MIRROR EFFECTS[0]

14.3.1 Acoustical Doppler Effect

In 1842, Doppler predicted the change in the frequency of a wave observed at a receiver whenever the source or receiver is moving relative to the other or to the wave propagation medium. The effect was first verified for sound waves by Buys-Ballot in 1845 from experiments conducted on a moving train. Nowadays, more striking examples of the acoustical Doppler effect can be found in the time-frequency distributions (TFDs) from acoustic sensors, whether in air or under water, when a propeller-driven aircraft or a helicopter passes overhead. The dominant feature in the acoustic spectrum of a turboprop aircraft is the spectral line corresponding to the propeller blade-passage frequency. In the case of a helicopter, the dominant line corresponds to the main (or tail) rotor blade-passage frequency. This line of constant frequency represents the signal emitted by the source. During an aircraft transit, the instantaneous frequency (IF) of this signal, when received by a stationary sensor, is observed to change with time due to the acoustical Doppler effect.

14.3.1.1 Time-Frequency Model: Microphone in Air

Consider an acoustic source moving along a straight line at a constant subsonic speed v so that at time τ_c it is at the closest point of approach (CPA) to a stationary acoustic sensor at a separation distance (or slant range) of R_c. Suppose that the source emits an acoustic tone of constant frequency f_0 and that the isospeed of sound propagation in air is c_a. The IF of the signal received by the sensor at time t is given by [1]

$$f_a(t) = \alpha + \beta p(t; \tau_c, s) \tag{14.3.1}$$

where

$$\alpha = f_0 c_a^2/(c_a^2 - v^2) \tag{14.3.2}$$

$$\beta = -f_0 c_a v/(c_a^2 - v^2) \tag{14.3.3}$$

$$s = \frac{R_c\sqrt{c_a^2 - v^2}}{vc_a} \tag{14.3.4}$$

$$p(t; \tau_c, s) = \frac{t - \tau_c}{\sqrt{s^2 + (t - \tau_c)^2}}. \tag{14.3.5}$$

Examples of time-frequency (t, f_a) curves computed by (14.3.1–14.3.5), with NLS-optimized parameters, are shown as solid lines in Fig. 14.3.1.

---

[0]Authors: **Brian G. Ferguson** and **Kam W. Lo**, Defence Science and Technology Organisation, P.O. Box 44, Pyrmont, NSW 2009, Australia (Brian.Ferguson@dsto.defence.gov.au, Kam.Lo@dsto.defence.gov.au). Reviewers: G. J. Frazer and M. Mesbah.

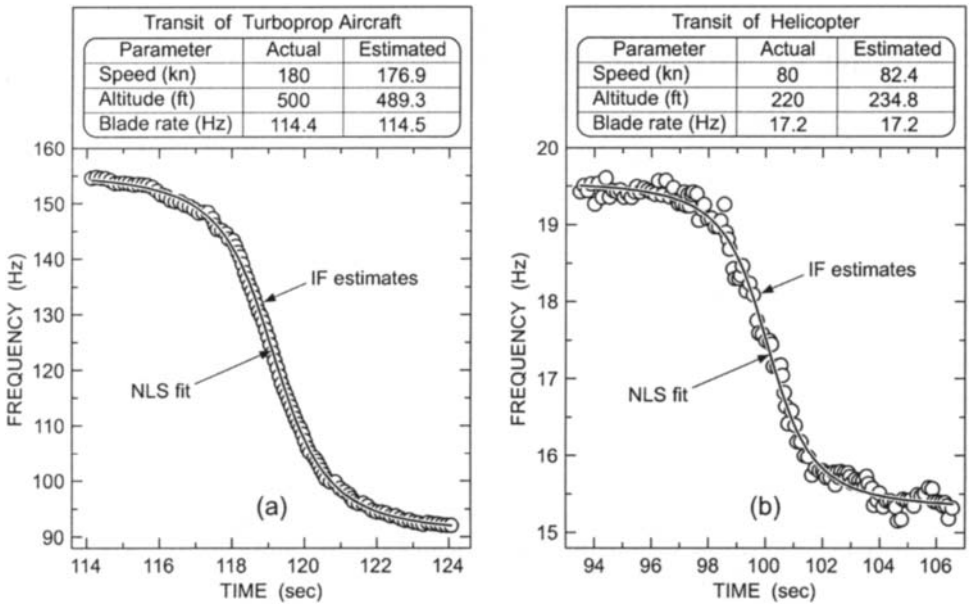


Fig. 14.3.1: (a) Variation with time of the instantaneous frequency estimates (circles) of the signal received by a microphone during the transit of a turboprop aircraft, and the nonlinear least squares (NLS) fit (solid line) to the observations. The actual and estimated values of the source parameters are tabled at the top of the diagram. (b) Similar to (a), but for the transit of a helicopter.

14.3.1.2 Time-Frequency Model: Hydrophone in Water

Now consider an airborne source traveling with a constant subsonic speed v at a constant altitude h in a constant direction so that its flight path passes directly over an acoustic sensor located at a depth d below the sea surface. Due to the propagation delay, the signal emitted by the source at time τ (the *source* time) arrives at the sensor at a later time t (the *sensor* time), which is given by [2,3]

$$t = \tau + \frac{l_a(\tau)}{c_a} + \frac{l_w(\tau)}{c_w} = \tau + \frac{h}{c_a \cos\phi_a(\tau)} + \frac{d}{c_w\sqrt{1-(c_w/c_a)^2\sin^2\phi_a(\tau)}} \quad (14.3.6)$$

where c_a and c_w are the respective isospeeds for sound propagation in air and water, $\phi_a(\tau)$ is the angle of incidence (in air) at source time τ, and $l_a(\tau)$ and $l_w(\tau)$ are the respective travel distances in air and water. Assume the source emits an acoustic tone with a constant frequency f_0. The IF of the signal received by the sensor at time t is given by [2,3]

$$f_w(t) = f_0\frac{d\tau}{dt} = \frac{f_0}{1+v\sin\phi_w(\tau)/c_w} = \frac{f_0}{1+v\sin\phi_a(\tau)/c_a} \quad (14.3.7)$$

where $\phi_w(\tau)$ is the angle of refraction (in water) at source time τ. Note that (14.3.6) and (14.3.7) constitute a set of parametric equations in $\phi_a(\tau)$. Given the

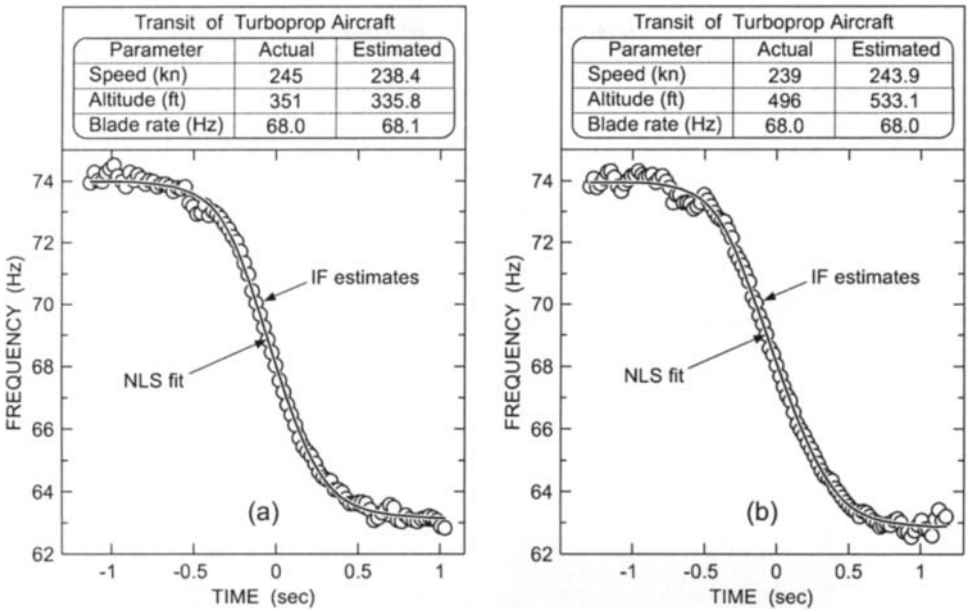


Fig. 14.3.2: (a) Variation with time of the instantaneous frequency estimates (circles) of the signal received by a hydrophone during the transit of a turboprop aircraft, and the nonlinear least squares fit (solid line) to the observations. The actual and source parameters are tabled at the top of the diagram. (b) Similar to (a), but for another transit of the turboprop aircraft.

source parameter set $\{f_0, v, h, \tau_c\}$ and sensor depth d, there is no explicit analytical expression for the IF f_w as a function of the sensor time t. However, a theoretical curve of f_w as a function of t can be obtained indirectly by first specifying the angle of incidence $\phi_a(\tau)$ and then calculating the corresponding values for f_w and t. Only acoustic ray paths for which the angle of incidence is less than the critical angle of incidence ϕ_c will have (some) acoustic energy transmitted across the air-sea interface into the underwater medium; i.e., $|\phi_a(\tau)|$ must be less than the critical angle of incidence ϕ_c, which is equal to $\sin^{-1}(c_a/c_w)$. Ray paths with angles of incidence exceeding the critical angle have all their acoustic energy reflected from the sea surface. Examples of time-frequency (t, f_w) curves computed by (14.3.6) and (14.3.7), with NLS-optimized parameters, are shown as solid lines in Fig. 14.3.2.

14.3.2 Acoustical Lloyd's Mirror Effect

When a broadband acoustic source (like a jet aircraft) is in motion, the TFD of the output of a sensor located above the ground displays a pattern of interference fringes known as the acoustical Lloyd's mirror effect. Fig. 14.3.3(a) shows a simulated TFD of the acoustic energy at the sensor output during the transit of an airborne source of broadband sound in an ideal environment. The resultant sound field at the sensor is the sum of the direct and ground-reflected sound fields. The reflected path is

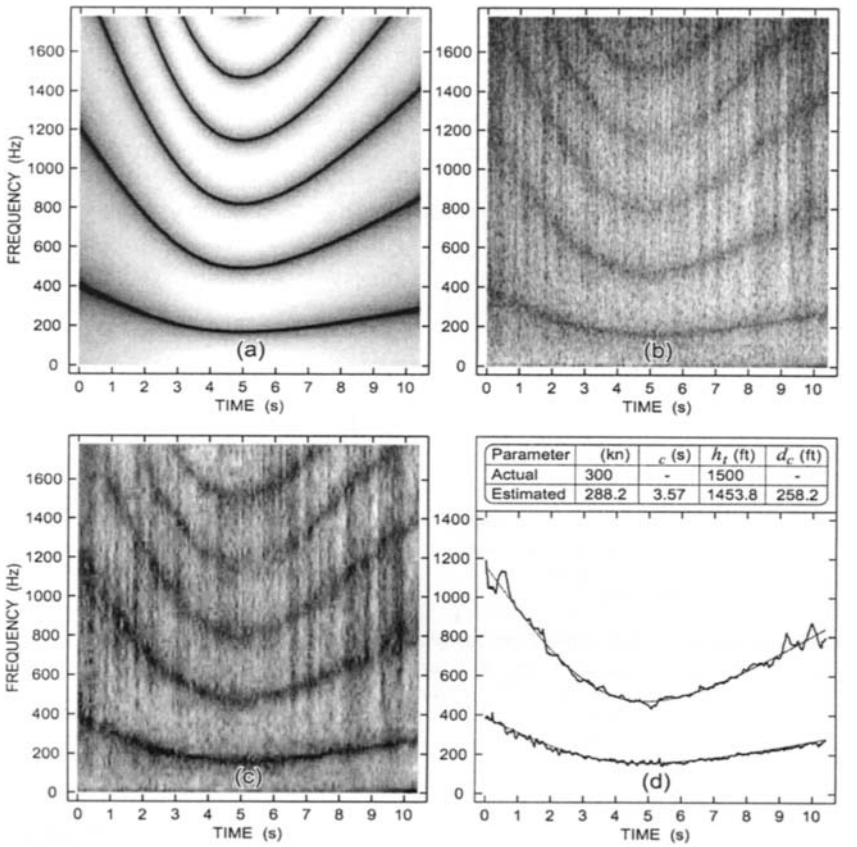


Fig. 14.3.3: (a) Simulated time-frequency distribution of a microphone's output for the transit of an airborne broadband acoustic source in an ideal environment. (b) Spectrogram of real data recorded from a microphone's output during a jet aircraft transit. (c) Image enhancement of (b) through pre-processing. (d) Predicted paths (thinner lines) of 1st and 2nd destructive-interference curves using NLS method superimposed on observed paths (thicker lines) from (c).

longer than the direct path between the source and the sensor, and the path length difference results in a phase difference between the two fields. At certain frequencies, the phase difference is of the correct amount to cause destructive interference (or cancellation) of the two sound fields.

14.3.2.1 Destructive-Interference Frequency Model

Consider an airborne source traveling in a straight line with a constant subsonic speed v at a constant altitude h_t over a hard ground. The acoustic sensor is located at a height h_r above the ground. The source is at the CPA to the sensor at time τ_c, with the ground range at CPA being d_c. Suppose the isospeed of sound propagation in air is c_a. The temporal variation of the nth order destructive-interference

frequency as a function of the four parameters $\{v_r, v_t, \tau_c, \gamma\}$ is modeled by [4]

$$f_n(t) = \frac{2n-1}{4} \frac{c_r^2}{c_r^2 - v_r^2} \left[\sqrt{\gamma^2(c_r^2 - v_r^2) + c_r^2 v_t^2 (t - \tau_c)^2} - v_r v_t (t - \tau_c) \right] \quad (14.3.8)$$

where $v_r = v/h_r$, $v_t = v/h_t$, $\gamma = \sqrt{1 + (d_c/h_t)^2}$, and $c_r = c_a/h_r$. The four parameters $\{v_r, v_t, \tau_c, \gamma\}$ are uniquely related to the four flight parameters $\{v, \tau_c, h_t, d_c\}$. Examples of destructive-interference frequency curves $\{f_n(t) : n = 1, 2\}$ computed by (14.3.8), with NLS-optimized parameters, are shown as thinner lines in Fig. 14.3.3(d).

14.3.3 Time-Frequency Signal Analysis

Various techniques for estimating the IF of a non-stationary signal have been considered in the literature (see [5] and Chapter 10). A common approach is to locate the peak of a TFD of the signal. In this article, an IF estimate corresponds to the propeller blade rate of a turboprop aircraft (or main rotor blade rate of a helicopter) as observed at an acoustic sensor. Shown in Figs. 14.3.1(a) and 14.3.1(b) are two examples of using the peaks of the spectrogram (short-time Fourier transform) to estimate the temporal variation of the IF (denoted by circles) of the signal received by a microphone located just above the ground during the respective transits of a turboprop aircraft and a helicopter. Shown in Figs. 14.3.2(a) and 14.3.2(b) are two examples of using the peaks of the Wigner-Ville distribution [Section 2.1.4] to estimate the temporal variation of the IF (denoted by circles) of the signal received by a hydrophone located under water during the respective transits of a turboprop aircraft.

Similarly, the temporal variation of the destructive-interference frequency of the signal received by an acoustic sensor located above the ground during the transit of an airborne source of broadband sound can be extracted by locating the minima of a TFD of the sensor's output. Shown in Fig. 14.3.3(d) is an example of using the minima of the spectrogram to estimate the temporal variation of the first and second destructive-interference frequencies (denoted by thicker lines) of the signal received by a microphone located above the ground during the transit of a jet aircraft.

14.3.4 Source Parameter Estimation: An Inverse Time-Frequency Problem

The inverse time-frequency problem is to estimate the source or flight parameters from the variation with time of the received signal's IF or destructive-interference frequency. The problem is solved by a nonlinear least-squares (NLS) method.

14.3.4.1 Narrowband Source in Level Flight with Constant Velocity: Microphone in Air

The source parameters $\{f_0, v, \tau_c, R_c\}$, or equivalently $\{\alpha, \beta, \tau_c, s\}$, are estimated by minimizing the sum of the squared deviations of the noisy IF estimates from their

predicted values [1]. Specifically, the NLS estimates of $\{\alpha, \beta, \tau_c, s\}$ are given by

$$\{\hat{\alpha}, \hat{\beta}, \hat{\tau}_c, \hat{s}\} = \arg\left\{ \min_{\{\alpha' \beta', \tau_c', s'\}} \sum_{k=1}^{K} \left[\alpha' + \beta' p(t_k; \tau_c', s') - g(t_k)\right]^2 \right\} \tag{14.3.9}$$

where $g(t_k)$ is the IF estimate at sensor time $t = t_k$ and K is the number of IF estimates. The four-dimensional minimization in (14.3.9) can be reduced to a two-dimensional maximization [1]:

$$\{\hat{\tau}_c, \hat{s}\} = \arg\left\{ \max_{\{\tau_c', s'\}} \frac{|\sum_{k=1}^{K} [g(t_k) - \bar{g}]\, p(t_k)|^2}{\sum_{k=1}^{K} [p(t_k) - \bar{p}]^2} \right\} \tag{14.3.10}$$

$$\hat{\beta} = \frac{\sum_{k=1}^{K} [g(t_k) - \bar{g}]\, \hat{p}(t_k)}{\sum_{k=1}^{K} [\hat{p}(t_k) - \bar{\hat{p}}]^2} \tag{14.3.11}$$

$$\hat{\alpha} = \bar{g} - \hat{\beta}\bar{\hat{p}} \tag{14.3.12}$$

where $\bar{g} = \frac{1}{K}\sum_k g(t_k)$, $p(t_k) = p(t_k; \tau_c', s')$, $\bar{p} = \frac{1}{K}\sum_k p(t_k)$, $\hat{p}(t_k) = p(t_k; \hat{\tau}_c, \hat{s})$, and $\bar{\hat{p}} = \frac{1}{K}\sum_k \hat{p}(t_k)$. Solving (14.3.2) and (14.3.3) using the estimated values for α and β gives the estimates of the source speed v and source frequency f_0 as

$$\hat{v} = -(\hat{\beta}/\hat{\alpha}) c_a \tag{14.3.13}$$

$$\hat{f}_0 = \hat{\alpha}(1 - \hat{v}^2/c_a^2). \tag{14.3.14}$$

From (14.3.4), the estimate of the CPA slant range R_c is given by

$$\hat{R}_c = \hat{s}\hat{v}c_a/\sqrt{c_a^2 - \hat{v}^2}. \tag{14.3.15}$$

The maximization in (14.3.10) is performed using the quasi-Newton method where the initial estimates of τ_c and s are given by the method described in [1]. The results of applying the source parameter estimation method to experimental data (represented by the circles) are shown at the top of Figs. 14.3.1(a) and 14.3.1(b). The estimates closely match the actual values of the aircraft's speed, altitude, and propeller or main rotor blade rate.

14.3.4.2 Narrowband Source in Level Flight with Constant Velocity: Hydrophone in Water

Define the cost function

$$P_f(f_0', v', h', \tau_c') = \sum_k \left[g(t_k) - f_0' z_k(v', h', \tau_c')\right]^2 \tag{14.3.16}$$

where $g(t_k)$ is the IF estimate at sensor time t_k, $\{f_0', v', h', \tau_c'\}$ are the hypothesized source parameters, and $f_0' z_k(v', h', \tau_c') = f_w(t_k)$ is the modeled frequency at sensor time t_k, which is computed using the approximate method found in [3]. Estimates

of the source parameters, $\{\hat{f}_0, \hat{v}, \hat{h}, \hat{\tau}_c\}$, are obtained by finding the set of values of $\{f_0', v', h', \tau_c'\}$ that minimizes P_f. It can be shown that

$$\{\hat{v}, \hat{h}, \hat{\tau}_c\} = \arg\left\{ \max_{\{v', h', \tau_c'\}} \frac{\left[\sum_k g(t_k)\, z_k(v', h', \tau_c')\right]^2}{\sum_k z_k^2(v', h', \tau_c')}, 0 < v' < c_a, h' > 0 \right\} \quad (14.3.17)$$

$$\hat{f}_0 = \frac{\sum_k g(t_k)\, z_k(\hat{v}, \hat{h}, \hat{\tau}_c)}{\sum_k z_k^2(\hat{v}, \hat{h}, \hat{\tau}_c)}. \quad (14.3.18)$$

Numerical procedures for maximizing (14.3.17) use standard (constrained) optimization methods. The initial estimates for $\{v, h, \tau_c\}$ are calculated using the procedure detailed in [3]. The results of applying the source parameter estimation method to experimental data (represented by the circles) are shown at the top of Figs. 14.3.2(a) and 14.3.2(b). The estimates closely match the actual values of the aircraft's speed, altitude and propeller blade rate.

14.3.4.3 Broadband Source in Level Flight with Constant Velocity: Microphone in Air

In a TFD of the sensor output, a destructive-interference curve is a sequence of connected points on the time-frequency plane with (locally) minimum amplitudes. The positions of these points, corresponding to the temporal variation of a particular order destructive-interference frequency, define the *path* (or shape) of the destructive-interference curve and their (logarithmic) magnitudes define the *intensity* of the curve. If the TFD is treated as an image, destructive-interference curves can be identified as dark fringes in the image – see Fig. 14.3.3(a). However, in practice, noise and background variations obscure the exact paths of these curves. Fig. 14.3.3(b) shows the spectrogram of a set of real data recorded from the output of a microphone during the transit of a jet aircraft. It is thus necessary that the time-frequency image be pre-processed to enhance the appearance of the destructive-interference curves before extracting the flight parameters from the image. The image shown in Fig. 14.3.3(b) is first normalized to remove background variations, and then wavelet de-noising is applied to the normalized image to suppress noise [6]. The result is an enhanced image [Fig. 14.3.3(c)] with the noise suppressed, yet the destructive-interference curves are not degraded in appearance.

Define the parameter vector $\boldsymbol{z} = [v_r, v_t, \tau_c, \gamma]^T$. A cost function that measures the difference between the observed and predicted paths of the destructive-interference curves is given by

$$P(\boldsymbol{z}) = \sum_{n=1}^{N} \sum_{k=1}^{K_n} \left[g_n(t_{nk}) - f_n(t_{nk}, \boldsymbol{z})\right]^2 \quad (14.3.19)$$

where $g_n(t)$ is the observed value of the nth order destructive-interference frequency at time t, which can be obtained by selecting the correct (local) minimum in the enhanced image, and $f_n(t, \boldsymbol{z})$ is the corresponding predicted value using (14.3.8). Minimizing $P(\boldsymbol{z})$ produces the NLS estimate $\hat{\boldsymbol{z}} = [\hat{v}_r, \hat{v}_t, \hat{\tau}_c, \hat{\gamma}]^T$ of $\boldsymbol{z}$. The speed, altitude

and CPA ground range of the source are then estimated as $\hat{v} = h_r \hat{v}_r$, $\hat{h}_t = \hat{v}/\hat{v}_t$, and $\hat{d}_c = |\hat{h}_t \sqrt{\hat{\gamma}^2 - 1}|$. The cost function is minimized using the Gauss-Newton method and the required initial estimate of $\boldsymbol{z}$ is obtained by the procedure detailed in [6].

The raw time-frequency image [Fig. 14.3.3(b)] is too noisy for flight parameter estimation. Indeed, direct application of the NLS method to this raw image produces very poor flight parameter estimates, necessitating the use of the enhanced image [Fig. 14.3.3(c)]. Fig. 14.3.3(d) shows the paths (thicker lines) of the first and second destructive-interference curves extracted from the enhanced image, which are then used in the NLS method to estimate the flight parameters. Only the first and second curves are used for flight parameter estimation, i.e., $n \leq 2$ in (19), because the higher order ($n > 2$) destructive-interference curves are too noisy. The estimated and actual values of the flight parameters are shown at the top of Fig. 14.3.3(d). The estimates of the speed and altitude are in good agreement with the actual values.

14.3.5 Summary and Conclusions

The acoustical Doppler effect enables the flight parameters and blade-passage frequencies of turboprop and rotary-wing aircraft to be estimated from the time-frequency signal analysis of acoustic sensor data. Also, during the transit of a broadband acoustic source, the acoustical Lloyd's mirror effect manifests itself as an interference pattern in the time-frequency distribution of the output of an acoustic sensor positioned above the ground. The flight parameters of jet aircraft can be estimated from the destructive interference fringes.

References

[1] B. G. Ferguson and B. G. Quinn, "Application of the short-time Fourier transform and the Wigner-Ville distribution to the acoustic localization of aircraft," *J. Acoustical Soc. of America*, vol. 96, pp. 821–827, 1994.

[2] B. G. Ferguson, "Time-frequency signal analysis of hydrophone data," *IEEE J. of Oceanic Engineering*, vol. 21, pp. 537–544, October 1996.

[3] B. G. Ferguson and K. W. Lo, "Transiting aircraft parameter estimation using underwater acoustic sensor data," *IEEE J. of Oceanic Engineering*, vol. 24, pp. 424–435, October 1999.

[4] K. W. Lo and B. G. Ferguson, "Passive estimation of aircraft motion parameters using destructive interference between direct and ground-reflected sound waves," in *Proc. Information Decision & Control 99*, pp. 171–176, IEEE, Adelaide, Australia, February 1999.

[5] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 2: Algorithms and applications," *Proc. IEEE*, vol. 80, pp. 540–568, April 1992.

[6] K. W. Lo, S. W. Perry, and B. G. Ferguson, "Aircraft flight parameter estimation using acoustical Lloyd's mirror effect," *IEEE Trans. Aerospace & Electronic Systems*, vol. 38, pp. 137–151, January 2002.

14.4 WIGNER-VILLE ANALYSIS OF HIGH FREQUENCY RADAR MEASUREMENTS OF A SURROGATE THEATER BALLISTIC MISSILE[0]

14.4.1 Experiment Description

In the autumn of 1997 four surrogate theater ballistic missiles (TBM) were launched from a temporary launch site in northern West Australia with the objective of testing a variety of TBM launch detection sensors. The missiles comprised two stages and used Terrier and improved Orion motors together with a passive payload designed so the flight vehicle had the same in-flight length as a representative TBM. The Terrier first stage burnt for approximately 5s following launch and the Orion second stage ignited at approximately 18s following launch and burnt for approximately 24s.

One of the sensors was an Australian Defence Science and Technology Organisation high frequency (HF) line-of-sight radar. The radar was positioned several tens of kilometers from the launch site and operated at a carrier frequency of either (approximately) 8MHz or 25MHz. The data reported here is from measurements made at 25MHz and is from the fourth surrogate TBM launch. The radar operated bistatically with the radar transmitter and receiver spaced approximately 20km apart to avoid the transmitter signal overloading the receiver via a direct line-of-sight or ground wave propagation path. This meant that there were no occluded regions in the range coverage. The transmitter floodlit the region of interest. The receiver used a "L" shaped array of 32 elements, with 16 elements on each arm of the "L" and each arm approximately 500m long. A 32 channel receiving system with analog de-ramping and digital output fed data to a real-time radar signal processor. Multiple directive beams covering the region of interest were formed using digital beamforming and both target elevation and azimuth could be determined. A full description of the radar is given in [1].

14.4.2 Signal Description

The radar used a linear frequency modulated continuous wave (LFMCW) waveform with a sweep (or waveform) repetition frequency (WRF) of 50Hz. A set of coherent measurements were collected, each of 256 sweeps or 5.12s duration. For hardware limitation reasons there was a short inter-dwell gap of approximately 12 sweeps ($12 \cdot \frac{1}{50}$s) where no data was recorded between each coherent measurement interval. The radar signal processor pulse compressed each sweep using stretch processing then formed 20 digital beams (10 for each arm of the "L" array). Doppler analysis of the 256 sweeps in each successive coherent processing interval was performed for each range cell in each beam direction.

The selection of WRF (50Hz) meant that Doppler measurements of what was a

[0]Author: **G. J. Frazer**, Defence Science and Technology Organisation (DSTO), PO Box 1500, Salisbury, SA 5108, Australia. Reviewers: B. Ristic and M. Turley (DSTO).

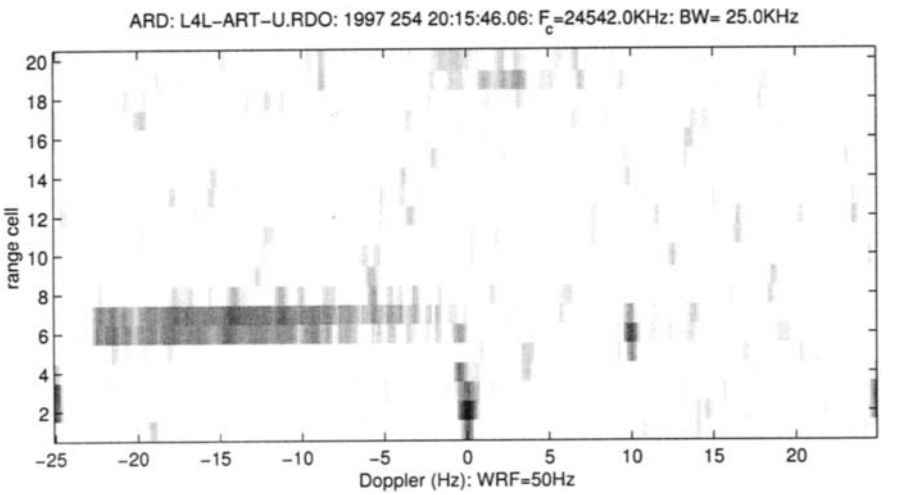


Fig. 14.4.1: Range-Doppler map showing the accelerating target smeared in Doppler (from -22Hz to 0Hz) in range cells 6 and 7. The direct wave from the transmitter and ground clutter is visible surrounding 0Hz Doppler and centered in range cell 2. The coasting spent first stage of the two stage TBM can be seen at range cell 6 with 10Hz Doppler.

very high velocity target were ambiguous for most of the flight. The long (with respect to target kinematics) coherent integration time (CIT) increased radar sensitivity although the target acceleration decreased the coherent processing gain achieved and limited the accuracy of velocity measurements. The radar used existing software for range, beamforming and Doppler processing which assumed constant velocity targets and rather than modify the software to support accelerating targets it was considered that some form of post event acceleration analysis could be used to mitigate the accelerating target mismatch loss.

Received radar data was displayed in real-time during the experiment and also recorded on tape for subsequent analysis. The data for off-line analysis was range processed (sweep compressed) and beamformed but not Doppler processed. It was organized into a sequence of dwells of data where each dwell contained two sets of 10 formed beams (one set for each arm of the "L" array) with a group of 20 range cells of interest and the complex signal corresponding to each of the 256 sweeps in a CIT for each range and beam. Sequences of dwells were collected into a single file.

A typical range-Doppler map seen by the operator is shown in Figure 14.4.1. It corresponds to one beam in the direction of the target and 20 range cells stacked vertically. For each range cell the Doppler spectrum has been determined from the appropriate range samples of each of the 256 pulse compressed sweeps, with one spectrum per range cell. Although not shown, the operator actually sees 20 such figures per dwell where each one corresponds to a different beam direction.

This particular example covers the time interval from launch time plus 16s to launch time plus 21s (T+16s : T+21s). The range-Doppler map shows the acceler-

ating target smeared in Doppler (from -22Hz to 0Hz) in range cells 6 and 7. The target velocity is such that this is an ambiguous Doppler measurement. The coasting spent first stage of the two stage TBM can be seen at range cell 6 with 10Hz Doppler. The direct wave from the transmitter as well as ground clutter is visible surrounding 0Hz Doppler and centered in range cell 2. An injected calibration signal can be seen in range cells 2 and 3 at -25Hz and +25Hz. Receding targets appear to incorrectly have positive Doppler however this is a frequency inversion artifact of the hardware design of the radar.

It is clearly difficult to determine the true time-varying velocity of the target since it changes significantly during the radar CIT. We are interested in determining the instantaneous Doppler law and hence the time-varying velocity of the target throughout this and all other dwells which contain the target during the powered segment of flight.

14.4.3 Signal Model

Let $z_{k,r,b,a,d}$ be the complex sample output from the radar signal processor prior to Doppler processing corresponding to the k^{th} sweep of the r^{th} range cell and b^{th} beam for the a^{th} arm of the "L" shaped receive array, for the d^{th} dwell. In the data $k \in [0,\ldots,255]$, $r \in [1,\ldots,20]$, $b \in [1,\ldots,10]$, $a \in [1,2]$ and $d \in [1,\ldots,10]$. Now let the complex signal corresponding to the coherent radar return from all sweeps for the r^{th} range cell, b^{th} beam and a^{th} arm in the d^{th} dwell be $z_{r,b,d,a}(k)$. This is the signal used for conventional Doppler processing. An example of the results of conventional Doppler processing applied to a set of range cells is shown in Figure 14.4.1.

Consider now a signal model for the radar return from a particular range and beam cell in a dwell which contained the target and where now for clarity we drop the notational dependence on r, b, d, a. Preliminary analysis using a spectrogram time-frequency distribution [2] suggested that the following discrete-time signal model would be at least piece-wise appropriate

$$z(t) = Ae^{j2\pi[f_0 t + \frac{\beta}{2}t^2]} + c(t) + n(t) \tag{14.4.1}$$

for $\{t : 0, \frac{1}{WRF}, \ldots, \frac{K}{WRF}\}$ and where A is the complex amplitude, f_0, β are the linear FM parameters, $c(t)$ represents clutter, and $n(t)$ represents noise. $T_{CIT} = \frac{K+1}{WRF}$ is the coherent integration time (CIT) of the radar. In the case of a bistatic HF line-of-sight radar, $c(t)$ includes contributions such as; the direct signal from the transmitter, range sidelobes from the direct signal, additional targets, say from a booster stage in a multi-stage rocket and meteor and ionospheric scatter. In general, both the clutter and noise are unknown, although we assume the relative energy is such that

$$\frac{|A|^2 T_{CIT}}{\left[\sum_0^{T_{CIT}} |c(t)|^2 + \sum_0^{T_{CIT}} |n(t)|^2\right]} \gg 1 \tag{14.4.2}$$

and thereby consider $z(t)$ as deterministic with unknown parameters, unknown

clutter and background noise, and with high signal to clutter plus noise energy ratio.

The objective is to determine the instantaneous Doppler law as part of the task of establishing the accelerating target dynamics. At a given radar carrier frequency the instantaneous Doppler as parameterized by f_0 and β can be converted to instantaneous target velocity. In some cases the instantaneous Doppler law of the target is not linear, or is piece-wise linear only, nevertheless, we adopt the model in (14.4.1) and if necessary apply our analysis over appropriate subsets of $\{t : 0 \leq t < T_{CIT}\}$.

14.4.4 Instantaneous Doppler Estimation

There are many methods which may have been applied to determine the instantaneous Doppler law (or f_0 and β) of the TBM target from the measured signal $z(t)$. Some are based on various time-frequency distributions [3–5]. Others are based on some estimation procedure for the parameters f_0 and β or the Doppler law, e.g. a maximum-likelihood criteria estimator for f_0 and β, or a least-squares criteria polynomial-phase-law estimator for the instantaneous Doppler law.

We have chosen to use a procedure based on the Wigner-Ville distribution (WVD) of the signal $z(t)$. We have done this for several reasons. Firstly, the clutter term $c(t)$ in (14.4.1) is unknown and at least early in the flight will include components due to the Terrier booster and range sidelobes from the direct signal from the transmitter. A simple phase difference based estimator we tried failed in these cases and more disturbingly provided no diagnostic as to the source of the error. On the other hand the WVD provided a clear visualization of the time-frequency law of the signal and since $c(t)$ was not too complicated the WVD display did not become overwhelmed by cross-terms in the WVD. Secondly, the data is in batch form (a sequence of dwells) with comparatively few samples in each dwell (256) which simplifies selection of the analysis window size for the WVD. There are small temporal gaps between dwells so that each dwell must be examined in isolation. Finally, our analysis is off-line with no requirement to be fully automatic. It was quite permissible to manually inspect and interpret the WVD of any signal of interest. Full details on the WVD are given in [2] and the references therein, and in Chapters 2 and 3.

14.4.5 Results

Standard analysis of the measured data from launch four consisted of extracting sequences of peaks corresponding to the target from beam-range-Doppler maps for the sequence of dwells covering the time interval of power flight. Each dwell comprised $(256 + 12) \cdot \frac{1}{WRF} = 5.36s$ elapsed time and a $256 \cdot \frac{1}{WRF} = 5.12s$ CIT. A single beam example of one of these maps from the dwell collected during second stage ignition is shown in Figure 14.4.1, the detail of which was discussed previously. The extracted Doppler measurements for the full period of powered flight, one per dwell, are shown in Figure 14.4.2. In this figure we have manually corrected for the

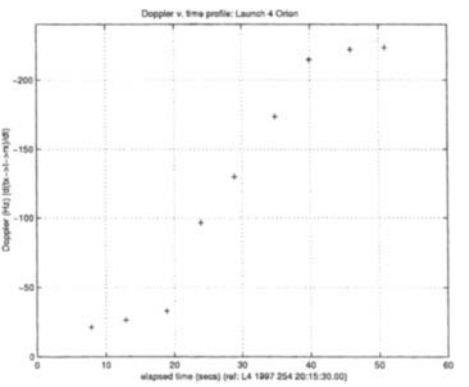


Fig. 14.4.2: The result of target Doppler law extraction based on conventional processing. One target Doppler measurement is determined for each dwell (5.12s).

discrete ambiguously measured Doppler caused by the selection of such a low WRF with respect to the target velocity.

Although not presented, similar results were obtained for target azimuth, elevation, range, and signal and noise energy. Although of no consequence here, determining azimuth, elevation and range required compensation for beam coning for the beams formed on each arm of the receive array and for range-Doppler coupling induced by our selection of waveform.

It is clear from Figure 14.4.2 that the temporal resolution of our Doppler measurements is poor. In fact, we have only 9 measurements for the period of powered flight, and any attempt to predict the full target trajectory following motor burnout, including impact location, will certainly be less accurate than desired. Many of the peaks are very approximate since the target accelerated during the CIT and therefore had no single velocity or Doppler.

Based on the approximate signal model in (14.4.1) and known localization properties of the WVD we have applied a procedure based on the WVD to improve upon the results derived from our standard analysis. Our objective was: to determine the target instantaneous Doppler law, to assist with distinguishing between accelerating and transient targets or scatterers, and to determine the instantaneous received target energy law and the processing loss caused by the target acceleration.

14.4.5.1 Instantaneous Doppler Law

The instantaneous Doppler law has been estimated from the peak of the WVD shown in Figure 14.4.3. Dwells prior to and later than this measurement interval have also been analyzed using the same procedure. The sequence of Doppler estimates, approximately one per sweep, is shown in Figure 14.4.4. There is an estimate every 20ms excluding the inter-dwell gap intervals and some missing estimates at the beginning and end of each CIT. It covers the period of powered flight and is an interval of approximately 10 CITs or some 50s. Contrasting Figures 14.4.2 and 14.4.4

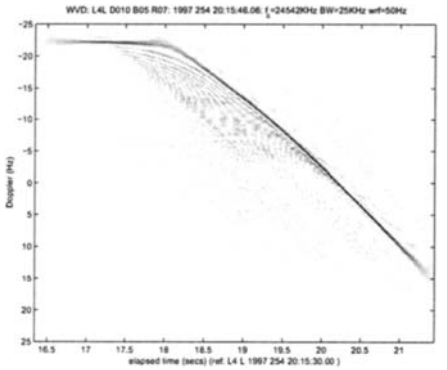


Fig. 14.4.3: WVD of the signal containing the accelerating target from range cell 7 in Figure 14.4.1. Second stage ignition occurred at 18s.

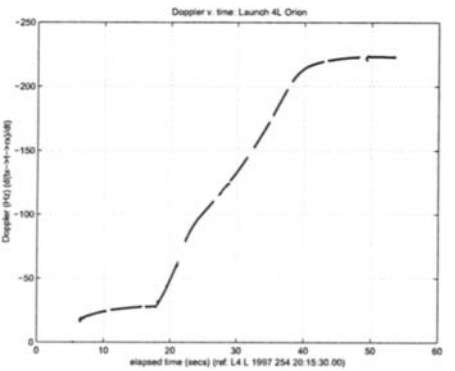


Fig. 14.4.4: Doppler v. elapsed time. Computed using the WVD. The gaps are due to missing sweeps during the radar inter-dwell gap and some lost instantaneous Doppler detections at either end of individual radar CIT. The plot is a sequence of point measurements, one per radar sweep, and not a continuous line, as it appears at this scale.

shows the benefit of determining accurate instantaneous Doppler law estimates as compared with conventional Doppler processing. The WVD based procedure has a temporal resolution of approximately the sweep duration, as compared with the CIT for conventional Doppler processing (20ms v. 5.12s). Note that the target second stage motor ignited at approximately 18s and that an accurate time of ignition can be determined from the instantaneous Doppler law.

Cross-terms are visible in the WVD shown in Figure 14.4.3 and are generated between the segment of the signal prior to second stage motor ignition and the segment which follows second stage ignition. The presence of the cross-terms did not impede the estimation of the instantaneous Doppler law.

Further work not reported here showed that smoothing of the instantaneous

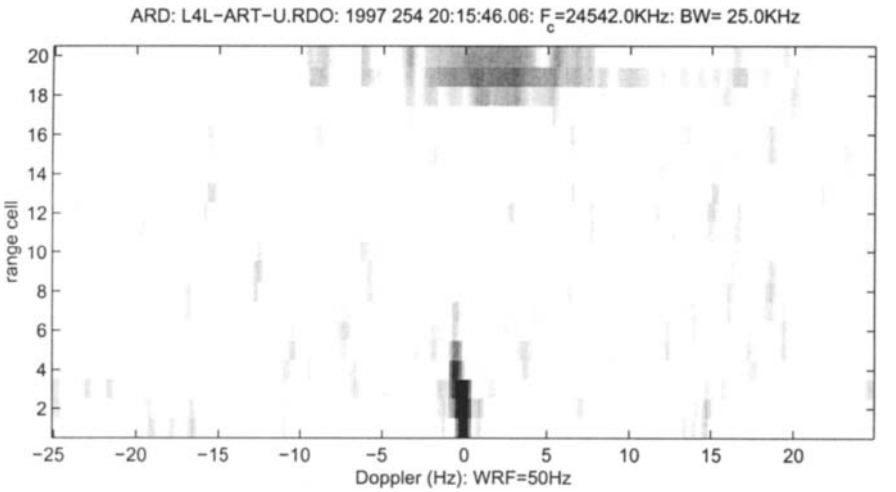


Fig. 14.4.5: Range-Doppler map showing the transient meteor scatterer in range cell 19. It is difficult to distinguish between this smeared feature and the smeared accelerating target shown in Figure 14.4.1.

Doppler law estimates using polynomial models reduces estimate variance and allows Doppler law prediction into the inter-dwell intervals. The corresponding polynomial velocity law was analytically integrated to produce range law estimates which were consistent with the directly measured range measurements.

14.4.5.2 Accelerating Target v. Transient

Figures 14.4.1 and 14.4.5 show the range–Doppler (RD) maps generated for two separate beam steer directions measured during the same 5.12s dwell. In Figure 14.4.1 the accelerating target is visible as a large smear in Doppler at range cell 7. In Figure 14.4.5 a transient meteor scatterer is visible at range cell 19, which is also smeared in the Doppler domain. We seek improved discrimination between the accelerating and the transient scatterers. Figures 14.4.3 and 14.4.6 show the WVD computed from the signal corresponding to the mentioned range cells. The instantaneous Doppler law of the accelerating target is visible and so is the transient behavior of the meteor scatterer. The WVD provides a display whose features readily allows the two types of scatterers to be distinguished.

14.4.5.3 Instantaneous Energy Law

Knowledge of the instantaneous Doppler law can also be used to construct a demodulation reference signal, $s(t)$. This signal has unity amplitude and instantaneous frequency law which is the conjugate of the estimated instantaneous Doppler law of the target, i.e. the instantaneous frequency law of $z(t)$. $s(t)$ can be used to

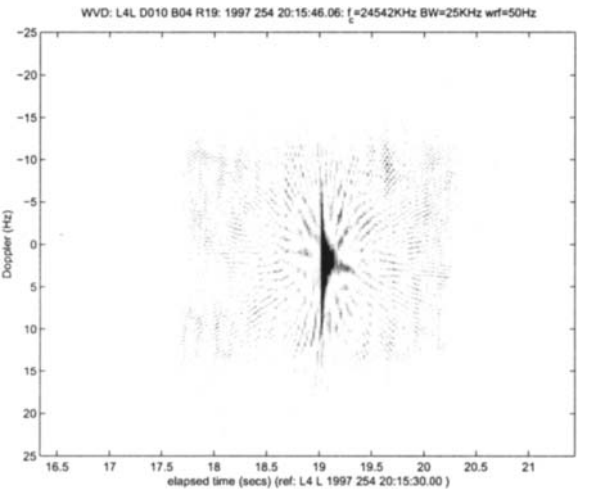


Fig. 14.4.6: WVD of the signal containing the transient meteor scatterer from range cell 19 in Figure 14.4.5.

demodulate $z(t)$ giving the approximately constant Doppler frequency signal $z'(t)$.

$$z'(t) = z(t) \cdot s(t) \tag{14.4.3}$$

The instantaneous energy of the demodulated signal is

$$\hat{p}(t) = \mathbf{G}|z'(t)|^2 \tag{14.4.4}$$

where $\mathbf{G}$ is some zero-phase smoothing operator. The instantaneous energy is shown in Figure 14.4.7, which shows three different levels of local smoothing, (i.e. different $\mathbf{G}$). Regardless of the smoothing operator selected there is a significant reduction in the target return energy immediately following second stage ignition between 18s and 18.5s. The cause of this effect is unresolved.

14.4.5.4 Processing Loss due to Target Acceleration

The processing loss due to target acceleration compared with a comparable target of constant velocity can be determined. One contrasts standard Doppler processing applied to the signal $z(t)$ and to the demodulated version $z'(t)$. It can be seen from Figure 14.4.8 that the processing loss is approximately 10dB for this particular CIT. This is the worst case since the dwell considered includes the period immediately before and immediately after second stage ignition and target acceleration was greatest in this measurement interval.

14.4.6 Summary and Conclusions

The Wigner-Ville distribution has been applied to HF line-of-sight radar measurements of a surrogate TBM launch. A procedure based on the WVD has been used

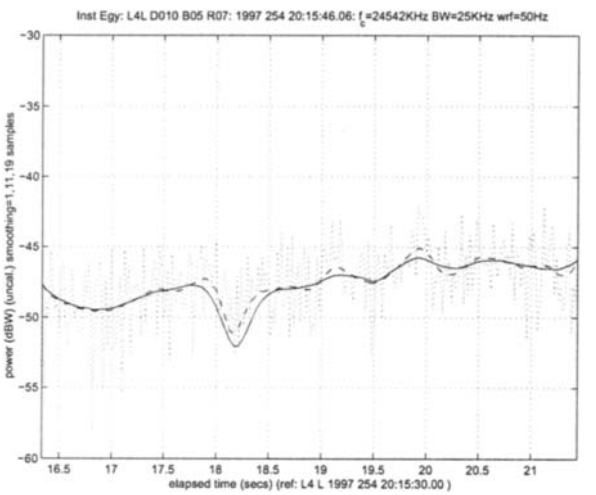


Fig. 14.4.7: Instantaneous energy law v. elapsed time for the interval 16.5s to 21.5s after launch. The three curves correspond to 1 (··), 11 (- -) and 19 (-) sample zero phase moving average smoothing.

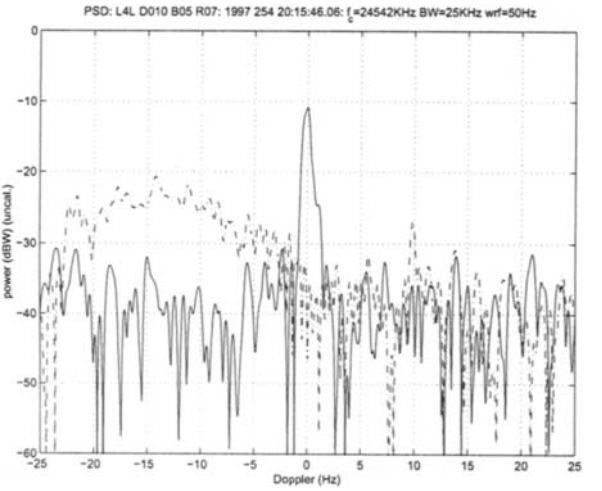


Fig. 14.4.8: The Doppler spectrum computed over the full CIT for the original signal (—·) and for the demodulated signal (-). The processing loss caused by assuming a constant velocity target is approximately 10dB in this case.

to estimate the Doppler law of the target at a temporal resolution of approximately 20ms compared with standard processing which had a temporal resolution of 5.12s and which contained error due to the target accelerating during the radar CIT. Examination of the display of the WVD has assisted with distinguishing between the accelerating target and interfering transient meteor scatterers. It has also al-

lowed determination of the instantaneous energy law of the target, identifying a still unexplained reduction in received signal level immediately following stage two ignition. Finally, analysis based on the procedure using the WVD has provided an estimate of the processing loss when standard Doppler processing is applied for this particular accelerating target.

References

[1] G. J. Frazer, "Project DUNDEE results: High frequency line-of-sight radar measurements and analysis," Tech. Rep. DSTO-TR-0734, Defence Science and Technology Organisation, 1998. AR 010-659.

[2] L. Cohen, "Time-frequency distributions—A review," *Proc. IEEE*, vol. 77, pp. 941–981, July 1989. Invited paper.

[3] B. Boashash, ed., *Time-Frequency Signal Analysis: Methods and Applications*. Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[4] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals," *Proc. IEEE*, vol. 80, pp. 520–538, April 1992.

[5] I. Djurović and L. Stanković, "Influence of high noise on the instantaneous frequency estimation using quadratic time-frequency distributions," *IEEE Signal Processing Letters*, vol. 7, pp. 317–319, November 2000.

[6] B. Boashash, G. Jones, and P. J. O'Shea, "Instantaneous frequency of signals: Concepts, estimation techniques and applications," in *Proc. SPIE: Advanced Algorithms and Architectures for Signal Processing IV*, vol. 1152, pp. 382–400, Soc. of Photo-optical Instrumentation Engineers, San Diego, 8–10 August 1989.

14.5 TIME-FREQUENCY METHODS IN SONAR[0]

14.5.1 Principles of Sonar

Sonar (for "Sound Navigation and Ranging", analogous to "Radar" for "Radio Detection and Ranging") was originally used for finding the distances and velocities of targets. It has evolved to systems that scan and are capable of localizing, identifying and classifying targets. Many other applications use sound under water, e.g. tracking and locating shoals of fish, identifying and tracking whales, sea-bottom surveying for mining, estimating depth or turbulence or temperature variations of the ocean, tracking aircraft flying above the ocean, vision systems for underwater robotic vehicles, etc. These tasks require the processing of a received signal to extract the relevant information. Some of these systems are active, that is, they transmit a sound signal and receive the reflected or backscattered signal. Others are passive and only receive sound emanating from other sources. The transmitters and receivers of sound in water are called hydrophones. Often arrays of hydrophones are employed to provide narrow beamwidths and increased sensitivity.

In **active sonar**, the transmitted signal is of known shape. The problem is to determine the delay in receiving the reflected or scattered signal. This delay is proportional to the target distance (i.e. range) and inversely proportional to the velocity of sound in water, which is about 1500 m/s. The wavelength of a 15 kHz component is therefore about 10 cm. In order to obtain better resolution, ultrasound at frequencies above 20 kHz is frequently employed in active sonar. The signal travels through water on its way to the target and back and is modified by the channel response. The received signal could be a superposition of scattered signals from several points on the target. It can also be a superposition of reflections from other objects in the environment (referred to as multipath). In addition, there may be noise.

An active sonar signal can be modeled using linear systems theory [1]. In the general case, both the object and the channel are linear and time-varying. The channel impulse response $c(\vec{r},t)$ and the insonified object's impulse response $o(\vec{r},t)$ are functions of space and time. The received signal $s_{\rm rec}(r;t)$ is related to the transmitted signal $s_{\rm tr}(t)$ through superposition integrals, as follows:

$$s_{\rm rec}(\vec{r};t) = \int_{-\infty}^{\infty} c(\vec{r};t,\tau)\, v(\vec{r};t,\tau)\, d\tau \;+\; n(t) \tag{14.5.1}$$

where $\vec{r}=[x,y,z]$ is the spatial position vector, t is the time, $n(t)$ is the noise, and

$$v(\vec{r};t) = \int_{-\infty}^{\infty} u(\vec{r};t,\tau)\, o(\vec{r};t,\tau)\, d\tau \tag{14.5.2}$$

where

$$u(\vec{r};t) = \int_{-\infty}^{\infty} c(\vec{r};t,\tau)\, s_{\rm tr}(\tau)\, d\tau. \tag{14.5.3}$$

[0]Author: **V. Chandran**, Queensland University of Technology, Brisbane, Australia (v.chandran@qut.edu.au). Reviewers: P. O'Shea, K. Abed-Meraim, E. Palmer and B. Boashash.

The spatial and temporal coherence of the ocean medium determine the largest aperture that can be used for image formation and the longest waveform duration that can be used in sonar systems. It is usually assumed that the coherence time of the channel is much greater than the time taken by the signal to travel to the target and back. Then the phase shift remains nearly constant over one pulse period. It is also assumed that the signal bandwidth is much smaller than the coherence bandwidth of the channel, which makes the channel wideband and frequency non-selective.

If the channel impulse response does not change over the duration of the signal, the system can be considered to be linear and time-invariant, and the superposition integrals are replaced by convolution integrals.

$$s_{\text{rec}}(t) = \left[\left(s_{\text{tr}}(t) \underset{t}{*} c(t)\right) \underset{t}{*} o(t)\right] \underset{t}{*} c(t) + n(t) \tag{14.5.4}$$

where $\underset{t}{*}$ represents convolution in the time domain. The channel impulse response consists of delayed impulses corresponding to various paths of propagation, such as echoes from the target and the sea-bottom. A typical channel impulse response that takes into account multiple reflections (but ignores point scatterers) is the discrete summation of impulses

$$c(t) = \sum_{i=1}^{M} c_i \, \delta(t - \tau_i) \tag{14.5.5}$$

where the c_i represent the attenuation suffered by each reflection term. Usually, in active sonar, the return from the target of interest is the strongest. The ideal condition is when there is only the one reflection, i.e. when $M = 1$. In practice, there are not only multiple reflections from other objects in the environment (such as fish or debris) but also backscattering from the ocean floor. If the lags for which the returned signal is analyzed are restricted to those corresponding to reasonable expected distances of the target, bottom backscatter and reflections from cluttering objects outside this "ballpark" can be ignored. The problem then essentially becomes one of detecting and locating the principal echo in additive noise that may be nonstationary owing to clutter.

Time-frequency analysis plays an important role in analyzing the received signal and estimating the range of the echo from the target because

- the transmitted signal is often frequency-modulated;
- there is a Doppler shift if the target is moving;
- time-delay is the parameter of most interest;
- noise level is high and signals are highly non-stationary, so that classical methods such as correlation or matched filters do not perform well, and
- time-frequency displays are physically meaningful for human observers to interpret.

In **passive sonar** (or acoustic) systems, there is no transmitted pulse; the sound emanates from the target itself. However, there may still be multiple paths by

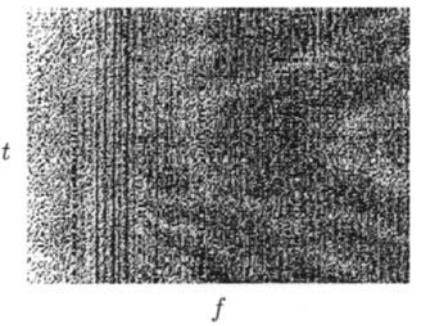


Fig. 14.5.1: Spectrogram of ship noise. Parabolic patterns are visible towards the right of the plot. The time (t) varies from 0 to 12 minutes and the frequency (f) from 0 to 280 Hz. (From ref. [2].)

which the sound travels to any hydrophone. In fact, the interference between the direct path and the bottom-returned path can be used to advantage in determining the distance and velocity of the source, and is particularly useful in detecting and tracking ships off the coast; it leaves parabolic patterns in the time-frequency plane as shown in [2]. If the channel is modeled as

$$c(t) = \delta(t) + c_1 \delta(t - \tau)\,, \tag{14.5.6}$$

then the magnitude-squared frequency response is $|C(\omega)|^2 = 1 + c_1^2 + 2c_1 \cos(\omega\tau)$. If h is the bottom depth, $h-k$ the hydrophone depth, and d the distance from the hydrophone to the ship along the surface, then it can be inferred from the geometry [2] that the direct signal travels a distance $\left(d^2 + (h-k)^2\right)^{1/2}$ and the bottom-reflected signal travels a distance $\left(d^2 + (h+k)^2\right)^{-1/2}$. The time delay between them is $\frac{1}{c}\left[\left(d^2 + (h+k)^2\right)^{-1/2} - \left(d^2 + (h-k)^2\right)^{1/2}\right]$, where c is the velocity of sound in water. Because d is often far greater than $h-k$ or $h+k$, the square roots can be approximated by first-order Taylor expansions to show that $\tau \approx \frac{2hk}{dc}$. Since the ship is in motion, τ will vary with time. If a is the minimum range of the ship and v its velocity of passage, and if the time origin is at the moment of closest approach, then $d^2 = a^2 + (vt)^2$. Again, when vt is much larger than a, this becomes $d \approx vt + \frac{a^2}{2vt}$, and the first interference maximum occurs at the angular frequency

$$\omega_{\max} = \frac{2\pi}{\tau} \approx \frac{\pi\left[vt + a^2/(2vt)\right]}{hk}. \tag{14.5.7}$$

Therefore, a time-frequency display will show parabolic patterns; see Fig. 14.5.1. Even if there is only flow noise from the vessel, these patterns can show up in long-time history averaged time-frequency representations. From these patterns, the speed of the vessel and its range can be estimated.

The passive sonar problem could be looked upon as one of blind deconvolution where the emitted signal parameters and the channel parameters need to be

simultaneously estimated (see also Article 8.1). Time-frequency methods are able to overcome the ill-conditioning typical of single sensor deterministic deconvolution techniques. Passive sonar can also be viewed as a transient signal detection problem in nonstationary noise when the signal-to-noise ratio is poor.

14.5.2 Classical Methods used in Sonar

14.5.2.1 Matched Filter

The optimal detector for a signal of known shape, $s(t)$, in additive Gaussian noise is the matched filter. If the received signal is $s_{\text{rec}}(t)$ and there is no Doppler shift, the matched filter can be formulated as a correlation, i.e.

$$\eta(\tau) = \int_{-\infty}^{\infty} s_{\text{rec}}(t)\, s^*(t-\tau)\, dt. \tag{14.5.8}$$

The greatest correlation between the delayed transmitted signal and the received signal will occur at the value of τ which corresponds to the delay of the principal echo. Hence the ideal matched filter output exhibits a peak at this delay. The matched filter output is thresholded to make a detection decision.

With increasing noise, dispersion in the medium and uncertainty in the transmitted signal shape, it becomes more and more difficult to obtain a discernible peak in the output at the correct location. When there is relative velocity between the source and receiver and a corresponding Doppler frequency ψ, the matched filter can be used for range-Doppler processing as

$$\eta(\psi, \tau) = \int_{-\infty}^{\infty} s_{\text{rec}}(t)\, s^*(t-\tau)\, e^{-j2\pi\psi t}\, dt \tag{14.5.9}$$

where the delay τ corresponds to the range parameter. The above version of the matched filter is also known as the Woodward ambiguity function.

14.5.2.2 Hypothesis Tests

Hypotheses tests, based on energy in sliding window signal segments, spectral density correlations and other statistics, are employed to decide whether a segment around a particular location resulted from signal (echo of the transmitted pulse in active sonar or signature of some acoustic source of interest in passive sonar) plus noise, or only noise. The alternative hypotheses for the detection of an unknown signal $s(t)$ in additive noise $n(t)$ are

$$H_0 : s_{\text{rec}}(t) = n(t) \tag{14.5.10}$$

$$H_1 : s_{\text{rec}}(t) = s(t) + n(t). \tag{14.5.11}$$

A test (or decision) statistic η is computed from the received signal $s_{\text{rec}}(t)$ or one of its transformed representations. If the value of the statistic exceeds a certain

threshold, the hypothesis is accepted. A measure of goodness of the test statistic (η) is the Signal-to-Noise (SNR) ratio [3]

$$\text{SNR} = \frac{|E(\eta/H_1) - E(\eta/H_0)|}{\sqrt{\frac{1}{2}\left[\text{var}(\eta/H_1) + \text{var}(\eta/H_0)\right]}} \tag{14.5.12}$$

where $E(...)$ denotes the expected value and var(...) the variance. This SNR can be used to compare test statistics. If the energy of the signal $s(t)$ is A, and if the noise is additive, Gaussian and white with power spectral density N_0, then the SNR of the matched filter output as a decision statistic is given by $\sqrt{A/N_0}$.

For classification of the signal into one of several known categories rather than detection of one category, the transient event would need to be segmented and a set of features extracted from the segment. Instead of thresholding a likelihood, a comparison of likelihoods can be used (e.g. the maximum-likelihood Bayes classifier).

Time-frequency methods are also used in hypothesis testing frameworks using a test statistic, and in classification frameworks using a set of features extracted from the time-frequency representation of the signal. The two-dimensional nature of the representation offers some potential for the selection of discriminating features that are also robust to noise.

14.5.3 Time-Frequency Approach to Sonar

14.5.3.1 Quadratic and Adaptive TFDs in Sonar

For sonar signal detection with the Wigner-Ville distribution (WVD), energy in the time-frequency plane is used as a test statistic. The WVD (see Article 2.1) is expressed as

$$W(t,f) = \int_{-\infty}^{\infty} z_{\text{rec}}\left(t - \tau/2\right) z^*_{\text{rec}}(t + \tau/2)\, e^{-j2\pi f\tau} d\tau \tag{14.5.13}$$

where $z_{\text{rec}}(t)$ is the analytic associate of $s_{\text{rec}}(t)$. The noise performance is given by [3]

$$\text{SNR} = \sqrt{\frac{A}{N_0}}\,\frac{1}{\sqrt{1 + N_0/A}}. \tag{14.5.14}$$

Thus the time-frequency representation is known to be sub-optimal and its performance degrades in high noise, due to its cross-terms. However, it provides the potential for time-varying filtering, and the effect of noise on the test statistic can be reduced by windowing the distribution. It is thus possible to obtain adaptive filtering prior to statistic computation or feature selection.

The cross WVD can be applied when the reference signal waveform is known (see Subsection 3.1.2.3):

$$W(t,f) = \int_{-\infty}^{\infty} z_{\text{rec}}(t - \tau/2)\, z^*(t + \tau/2)\, e^{-j2\pi f\tau} d\tau \tag{14.5.15}$$

where $z(t)$ is the analytic associate of $s(t)$. It is equivalent to the matched filter approach in this case, but with the added advantage of a two-dimensional time-frequency plane for adaptation and filtering.

If the transient signal is monocomponent and of known instantaneous frequency (IF), a two-dimensional window in the time-frequency plane that preserves points in a neighborhood of the IF, while filtering out all others, can be used to provide a test statistic with improved SNR. If the transient signal is multicomponent, one can compute and add the cross WVDs of the received signal with the separate components of the transient signal, thereby eliminating the effect of cross-terms while retaining auto-terms in the time-frequency plane. The IF laws in this case would need to have been identified and are assumed known [3].

Other quadratic TFDs such as the **cone-shaped kernel** time-frequency representation (CK-TFR) or ZAM distribution, given in Article 3.3, significantly attenuate the cross-terms and hence may be better suited for time-frequency displays of active sonar returns [4].

The **adaptive optimal kernel** (**AOK**), described in Article 5.3, uses a radially Gaussian signal-dependent kernel. The shape of the kernel is changed to optimally smooth the TFD as discussed in Chapter 3 and Article 5.3. The optimal kernel is one that maximizes a quality measure for the time-frequency representation subject to constraints that force the kernel to suppress cross-terms and to satisfy marginal distributions.

Performance comparisons of the AOK-based time-frequency representation with the STFT, Wigner and ZAM kernels on real multicomponent active-sonar signals are presented in [5].

Fig. 14.5.2 shows various time-frequency representations of a man-made underwater acoustic signal comprising three chirps in noise. The signal itself in plotted in the top box. The other boxes, labeled STFT, WIGNER (for WVD), ZAM and AOK, contain four time-frequency representation plots where the horizontal axis is time and the vertical axis is frequency. The chirps show up in the AOK plot as three line-like features decreasing in frequency with time. They overlap in time but not significantly in the time-frequency plane of the AOK TFR. Hence the chirps can be detected and delays estimated from this representation. Notice that the Wigner and ZAM distributions suffer more severely from cross-term artifacts while the resolution of the STFT is too poor to bring out the relevant features.

14.5.3.2 Gabor Expansion and Wavelets in Sonar

A **Cross-term Deleted Wigner representation** (**CDWR**), based on Gabor expansions (see Article 11.2), is applied in [6] to sonar detection and classification, on data collected by the US Navy. The test data include hyperbolic FM chirp signals and continuous-wave signals for two events. The decision statistic is a cross-correlation between the cross-CDWR of the received and replica signals and the auto-CDWR of the replica signal. This method achieves better detection accuracy than the square-law detector used by the US Navy. It is tailored to helicopter-

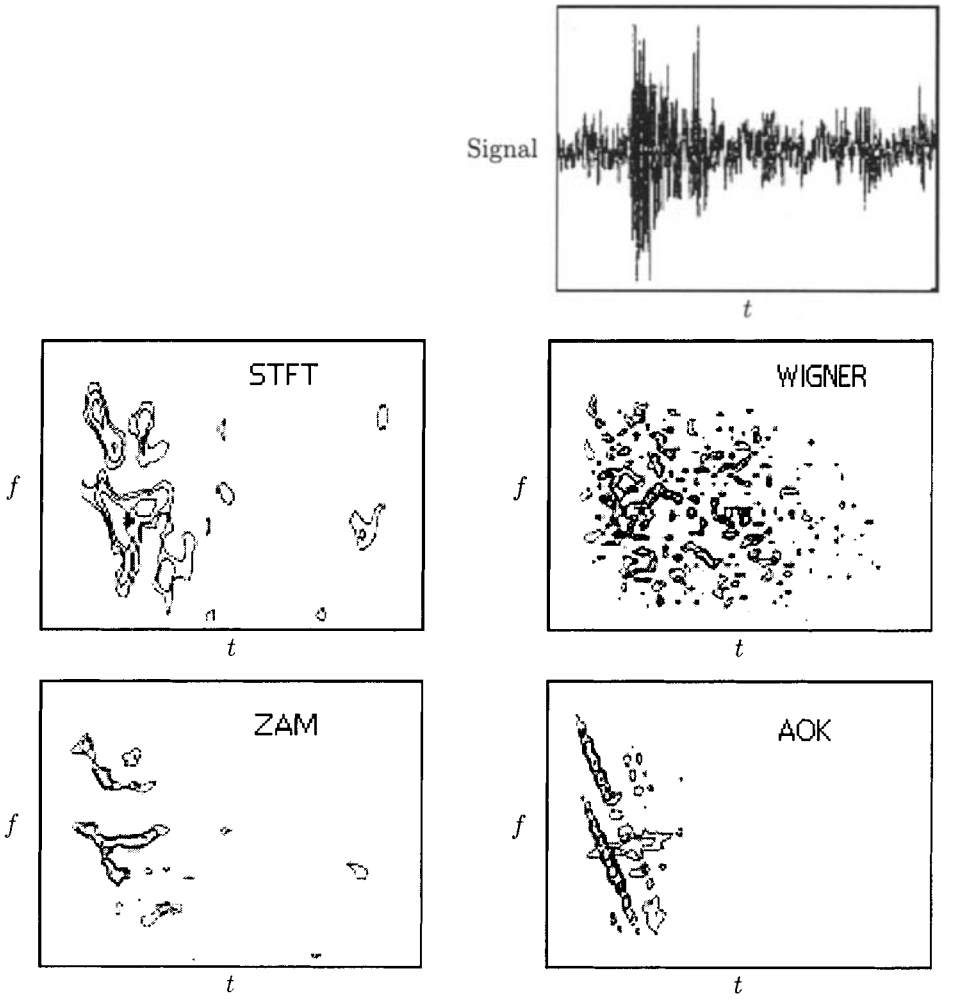


Fig. 14.5.2: Time trace (top) and various TFRs (labeled) of a sonar signal comprising three chirps. The AOK representation resolves features (nearly linear with negative slopes, showing frequencies decreasing with time) corresponding to the three chirps. The frequency sweep would typically be about 10 kHz to 20 kHz and the transients would last tens of milliseconds. (Adapted from [5].)

deployed arrays of hydrophones, which are often close to the target but have wide beamwidths and consequently poor signal-to-noise ratios.

The **wavelet transform** and its squared magnitude (the scalogram) are compared with the Wigner-Ville distribution for automatic classification of active sonar data in [7]. Classification accuracy of up to 92 percent is achieved at -4 dB SNR on synthetic data of acoustic scattering from high-aspect-ratio solids (length-to-width ratios of 5 and 10). The wavelet transform often outperforms the Wigner-Ville

distribution partly because it does not suffer from cross-terms. The scalogram, however, does suffer from such effects because it is a second-order representation.

Wavelet-based features and artificial neural networks based on the Bienenstock, Cooper and Munro (BCM) theory are used to classify underwater mammal sounds in [8]. Accuracy is shown to be significantly greater than for the use of the same number of Fourier energy and frequency features. The improvement is largely owing to the wavelets' non-uniform tiling of the time-frequency plane, and owing to BCM feature extraction that performs unsupervised nonlinear dimension reduction and supervised discriminant pursuit.

14.5.4 Prony and Higher-Order Spectral Methods in Sonar

The methods discussed below are "indirect" because although they provide time-frequency representations (primarily by means of sliding window analysis) they are adapted from techniques first proposed for spectral analysis, higher-order spectral analysis or other signal processing techniques. They are not members of the quadratic class of time-frequency representations either, but are sufficiently important to warrant description in this article.

Prony model method: The modified Prony method of detection of short-duration transient signals in noise is based on complex exponential signal modeling. It employs forward and backward predictions using high prediction orders, and eliminates noise zeroes from the forward and backward characteristic polynomials by using the fact that zeroes of white, stationary noise do not exhibit conjugate reciprocal relationships in the z-plane as do true signal zeroes. Singular value decomposition is used to provide further noise enhancement. After these steps, an enhanced short-duration signal spectrum for the analysis window is computed. A sliding window ensures that a time-frequency representation is obtained from which features can be extracted for detection and/or classification. The performance of this technique is compared with the STFT and the Wigner distribution on underwater acoustical data (tuna and sperm whale clicks) in [9]. The data was sampled at 23 kHz and the transient events lasted 10 to 30 milliseconds. Significantly improved features were extracted by the Prony technique.

Higher-Order Spectral methods: Gaussian noise immunity, translation invariance and other useful properties of higher order spectra are also used in obtaining robust time-frequency representations and in the feature-extraction stage after a representation. Higher-order spectra are Fourier representations of cumulants or moments of a stationary random process. They are functions of more than one frequency. The bispectrum is a function of two frequencies and is the Fourier transform of the third-order cumulant, which is a function of two lag variables. Unlike the power spectrum (which is the second order spectrum), higher-order spectra retain phase information. They are zero for Gaussian processes and can be used to detect non-Gaussian processes and identify nonlinear systems. For transient detection, as required in sonar, the stationarity assumption needs to be relaxed. Further, the

noise is often non-stationary as well and can contribute to non-zero higher-order spectral contributions. Time-varying versions of higher-order spectra include the third-order Wigner distribution called the Wigner Bispectrum [10], and a class of generalized Wigner-Ville distributions based on polynomial phase laws [11], which is applied to underwater acoustic signals.

A sonar image processing technique based on higher-order spectral feature extraction, presented in [12], is also indirectly related to time-frequency representation. This method applies to sonar scans that result in two-dimensional images with azimuth and range as independent variables. Sonar0, Sonar1 and Sonar3 sea-mine databases from the US Naval Surface Warfare Center, Coastal Systems Station, Florida (reference in [12]) are used. The three databases together contain 122 sea-mines in the training set and 123 in the testing set. The processing stages comprise adaptive Wiener filtering, sliding window matched filtering and adaptive thresholding. These steps account for non-stationary background noise. In this method, the thresholding after matched filtering is used for removing outlying noise rather than for detection. The resulting image is then used to extract features—the principal ones being bispectral and trispectral features (references in [12]) that are designed to be invariant to scaling and translation. Features that are robust to scaling make the task of the classifier easier. A statistical multi-stage classifier is used. The method achieves close to 90% accuracy with about 10% false alarms for detection of sea-mines. Since higher-order spectra are evaluated over overlapping windows, this method is an indirect "space-frequency" representation approach applied to 2D input. It is capable of being trained to achieve robustness to intra-class variations and some channel variations.

14.5.5 Dispersion and Angle Frequency Representation

Different propagating modes through the ocean (viewed as a waveguide) can exhibit different dispersion (group velocity versus frequency) curves. Different frequencies travel with different velocities within the same mode and between different modes of propagation. As a result, it is even possible for a linear FM up-ramp pulse to be converted into a down-ramp pulse of a different frequency law and increased duration, in extreme cases. In [13] a study of this effect is presented along with a method for deconvolution of the channel response. Such deconvolution is sometimes necessary before feature extraction to reduce classification or parameter estimation errors.

Dispersion can also be beneficial in classifying ocean bottoms. Bottom backscattering is a function of the type of surface (such as sand, pebble, clay, rock) and of the frequency of sound and incident angle. Wideband sonar (typically 20 kHz to 140 kHz [14]) is used to characterize the frequency dependence of bottom backscatter and estimate the backscattering coefficient or the impulse response of the bottom. Conventional methods are not adequate for this. The directivity of the transducer and propagation attenuation cannot be ignored and must be deconvolved from the echo response. An angle frequency representation is then obtained which exhibits

features similar to time-frequency representations that can be used for classification. When the scale of the roughness is much smaller than the depth, it affects only the phase and not appreciably the magnitude of the response at any frequency. However, there is an incident angle dependence which can be used for classification. A sand bottom shows a decrease in the backscattering coefficient as the angle increases, whereas, for surfaces with greater roughness such as the pebble bottom, the backscattering coefficient is nearly independent of the angle.

14.5.6 Summary and Conclusions

Time-frequency analysis may be applied with considerable success to active and passive underwater sonar data. Although multipath interference is usually a problem in radar and sonar applications, it is interesting to note that one of the first applications of time-frequency analysis to underwater sonar used it beneficially for detecting and tracking ships. Reduced-interference quadratic time-frequency representations allow time-frequency characteristics of transient signals to be displayed and interpreted without difficulty even at low signal-to-noise ratios and even for multicomponent signals. Filtering in the time-frequency plane allows for robust feature selection and decision statistics of better quality. Adaptations of other techniques such as Prony's method of spectral analysis, wavelet transforms, scale-space analysis and higher-order spectral analysis also allow time-frequency type representation and feature extraction for underwater acoustic signal processing.

Automated classification of underwater sources is still challenging. Multi-class problems with more than a few classes are yet to be solved satisfactorily, as the signatures often show significant intra-class variation due to changing channel conditions—multipath, dispersion, etc.—while the classification problem is coupled with the estimation of range, velocity and possibly other parameters.

Two-dimensional and three-dimensional scans make the classification task easier through the exploitation of spatial correlations of the returned signal or the match filtered output. However, they introduce additional difficulties with real-time processing. The concept of first expanding a signal in dimensionality to an information-rich space, and then selectively reducing the dimensionality through projections, slices or segmentation, is a powerful one and is being exploited in many pattern recognition problems. Time-frequency representations are one example of such an approach. They can reveal the nature of the evolution of correlations in a transient signal and provide information on the appropriate subspaces to extract features from. Practical exploitation of this in an automated detection, classification and tracking system requires its interfacing with other powerful concepts from other areas such as feature extraction, feature selection, adaptive information processing, classifier selection and classifier fusion.

References

[1] J. S. Jaffe, G. Chandran, and E. Reuss, "High frequency acoustic imaging in the ocean," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'95)*,

vol. 5, pp. 2793–2796, Detroit, 9–12 May 1995.

[2] J. G. Lourens, "Passive sonar detection of ships with spectro-grams," in *Proc. 3rd South African IEEE Conference on Theoretical and Practical Work in Communications and Signal Processing*, pp. 147–151, Johannesburg, June 1990.

[3] B. Boashash and P. J. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1829–1841, November 1990.

[4] W. L. J. Fox, J. C. Luby, J. W. Pitton, P. J. Loughlin, and L. E. Atlas, "Sonar and radar range-Doppler processing using a cone-shaped kernel time-frequency representation," in *Proc. 24th Asilomar Conf. on Signals, Systems, and Computers*, vol. 2, pp. 1079–1083, Pacific Grove, CA, 5–7 November 1990.

[5] R. Baraniuk, D. Jones, T. Brotherton, and S. L. Marple Jr., "Applications of adaptive time-frequency representations to underwater acoustic signal processing," in *Proc. 25th Asilomar Conf. on Signals, Systems, and Computers*, vol. 2, pp. 1109–1113, Pacific Grove, CA, 4–6 November 1991.

[6] S. Kadambe and T. Adali, "Application of cross-term deleted Wigner representation (CDWR) for sonar target detection/classification," in *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers*, vol. 1, pp. 822–826, Pacific Grove, CA, 1–4 November 1998.

[7] F. Lari and A. Zakhor, "Automatic classification of active sonar data using time-frequency transforms," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 21–24, Victoria, BC, 4-6 October 1992.

[8] 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 Processing*, vol. 46, pp. 1202–1207, May 1998.

[9] S. L. Marple Jr. and T. Brotherton, "Detection and classification of short duration underwater acoustic signals by Prony's method," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'91)*, vol. 2, pp. 1309–1312, Toronto, 14–17 May 1991.

[10] N. L. Gerr, "Introducing a third-order Wigner distribution," *Proc. IEEE*, vol. 76, pp. 290–292, March 1988.

[11] B. Boashash and G. Frazer, "Time-varying higher-order spectra, generalised Wigner-Ville distribution and the analysis of underwater acoustic data," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'92)*, vol. 5, pp. 193–196, San Francisco, 23–26 March 1992.

[12] V. Chandran, S. Elgar, and A. Nguyen, "Detection of mines in acoustic images using higher order spectral features," *IEEE J. of Oceanic Engineering*, vol. 27, pp. 610–618, July 2002.

[13] Z.-H. Michalopoulou, "Underwater transient signal processing: Marine mammal identification, localization, and source signal deconvolution," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, vol. 1, pp. 503–506, Munich, 21–24 April 1997.

[14] N. Ma, D. Vray, P. Delachartre, and G. Gimenez, "Sea-bottom backscattering modeling with a wideband constant beamwidth sonar at normal incidence," in *Proc. IEEE Ultrasonics Symposium*, vol. 2, pp. 1077–1080, Seattle, 7–10 November 1995.

This Page Intentionally Left Blank

Chapter 15

Time-Frequency Diagnosis and Monitoring

Time-frequency applications are now so widespread that they cannot be comprehensively covered in one volume. For this reason, this chapter attempts to further illustrate the time-frequency approach by selecting a few key generic applications of diagnosis and monitoring. The topic is represented by six articles covering a wide range of diverse applications.

Electrical power quality is often severely affected by transient disturbances. It is necessary to detect and assess their effect on voltage and current stability. This is achieved by a time-localized frequency analysis where the instantaneous frequency (IF) allows us to assess disturbance propagation (Article 15.1). In the automotive industry, the treatment and prevention of knock is a major problem for internal combustion engines as knock may lead to engine damage. The Wigner-Ville distribution is used to optimize the position for placement of knock sensors (15.2). Some applications involve signals that have dispersive group delays governed by a power law, such as dispersive propagation of a shock wave in a steel beam, and cetacean mammal whistles. A power class of TFDs suitable for such applications is formulated and a methodology described (15.3). In other applications of image processing, image quality may be assessed using a WVD based measure correlated with subjective human evaluations. A new SNR based on the WVD is shown to outperform the conventional SNR measure (15.4). In an application involving neonatal care, monitoring and diagnosing newborns for seizures is possible using a time-frequency approach exploiting characteristic patterns in the time-frequency plane. These patterns are used as a basis for a time-frequency matched-filter automatic detection (15.5). Machine condition monitoring is a task crucial to the competitiveness of a wide range of industries. Detecting and diagnosing faults in machines is possible using time-frequency approaches such as the WVD, wavelets and wavelet packets (15.6).

Time Frequency Signal Analysis and Processing
Boualem Boashash (Ed.)

15.1 TIME-FREQUENCY ANALYSIS OF ELECTRIC POWER DISTURBANCES [0]

With the advent of various power electronic devices in modern power systems, power quality is becoming an important issue for both customers and service providers. Ideally the voltage and current signals in power systems are supposed to be constant in amplitude and frequency. However, the voltage and current frequently suffer distortions in amplitude, frequency and phase due to various sources of disturbances [1]. The effects of periodic distortion or harmonics have been treated with Fourier series and are characterized by a total harmonic distortion (THD) index [1]. Recently transient disturbances, which occur within less than one cycle, are of great interest since they also degrade power quality. Specifically, the detection, estimation, classification and assessment of transient disturbance signals have become an important aspect in power quality analysis. To overcome the inability of Fourier analysis to provide time-localized frequency information for the transient disturbance signals, advanced signal processing techniques such as the wavelet transform [2] and time-frequency analysis [3] have been introduced to provide time-localized frequency analysis of the disturbances.

In this article, we focus on applications of time-frequency analysis to transient power quality events. A power system is mainly divided into distribution and transmission levels. We will provide one example from the distribution level for power quality assessment and the other from the transmission level for the propagation of disturbance waves.

15.1.1 Time-Frequency Analysis: Reduced Interference Distribution

As explained in Chapter 3, various types of time-frequency distributions may be expressed in terms of the following quadratic (or bilinear) distribution function with a kernel $g(\nu,\tau)$ [4]:

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j2\pi\nu(u-t)} g(\nu,\tau)\, z(u+\tfrac{\tau}{2})\, z^*(u-\tfrac{\tau}{2})\, e^{-j2\pi f\tau}\, d\nu\, du\, d\tau. \quad (15.1.1)$$

This $\rho_z(t,f)$ is the time-frequency distribution of $z(t)$. The arguments regarding the kernel selection and suppression of interference effects are also applicable to the case of the power quality analysis. As the disturbance signal is characterized by the presence of multiple frequency components over a short time, interference is also problematic and a high resolution time-frequency distribution is required. Among the various types of time-frequency distributions, the reduced interference distribution (RID) [4] has been shown to exhibit the most suitable properties for the analysis of power quality disturbance events (see [3] and Article 5.2).

[0]Authors: **Edward J. Powers**, **YongJune Shin** and **William M. Grady**, University of Texas at Austin, TX (ejpowers@mail.utexas.edu, june@ece.utexas.edu, grady@ece.utexas.edu). Reviewers: W. J. Williams and G. J. Frazer.

The advantage of the RID can be found in the motivation and philosophy of the RID. RID is a more general definition of the time-frequency distribution kernel that satisfies the following criteria: realness, time/frequency shift invariance, time/frequency marginal properties, instantaneous frequency (IF) / group delay (GD) availability, and time/frequency support. Various definitions of the RID kernel are available as a two dimensional low pass filter and the requirements listed above. Among the various types of the RID kernels, we employ the binomial distribution kernel provided in [4].

In the beginning of transient power quality signal analysis, the wavelet transform has been mainly utilized [2], because wavelet analysis provides time-localized properties. However, the time-frequency distribution, especially the RID, is a potentially more useful distribution [5] because it allows one not only to visualize the transient signal information but also to extract transient parameters useful in assessing the severity of various transient power quality events.

15.1.2 Power Quality Assessment via Time-Frequency Analysis

In this section, we present the application of time-frequency analysis to the assessment of power quality [3]. A capacitor switching disturbance waveform with its corresponding RID is provided in Fig. 15.1.1. The objective of capacitor switching in power systems is to correct the power factor and/or mitigate the effects of harmonics associated with nonlinearities. However, the switching capacitor event also generates undesirable disturbance waveforms as shown at the top of Fig. 15.1.1. At the bottom of Fig. 15.1.1 the corresponding RID is provided. The time-varying frequency content of the capacitor switching disturbance is clearly seen. Note that frequencies up to 3.5kHz appear to be associated with this particular event. This voltage signal was recorded in the field and is provided through the courtesy of EPRI (Electric Power Research Institute).

Besides the visualization of the time-varying spectral characteristics of the disturbance provided by the binomial RID, one can characterize the frequency variation of the disturbance in terms of instantaneous frequency.

For a time-frequency distribution $\rho_z(t,f)$ satisfying the IF property (see Section 3.1.1), the instantaneous frequency (IF) may be expressed as

$$f_i(t) = \frac{\int_{-\infty}^{\infty} f \cdot \rho_z(t,f)\, df}{\int_{-\infty}^{\infty} \rho_z(t,f)\, df}. \tag{15.1.2}$$

The instantaneous frequency is a normalized first order frequency moment for the time-frequency distribution and corresponds to the "average" frequency for given time t, where each frequency is weighted by the relative energy associated with that frequency at time t. Note that not all types of time-frequency distributions provide a reasonable estimate of the instantaneous frequency. For a reasonable estimate of the instantaneous frequency, a kernel must meet the requirements that $g(\nu, \tau = 0) = 1$ and that $\frac{\partial g(\nu,\tau)}{\partial \tau}|_{\tau=0} = 0$, which are satisfied by the binomial RID kernel used in this article (see Tables 3.3.1 and 6.1.1).

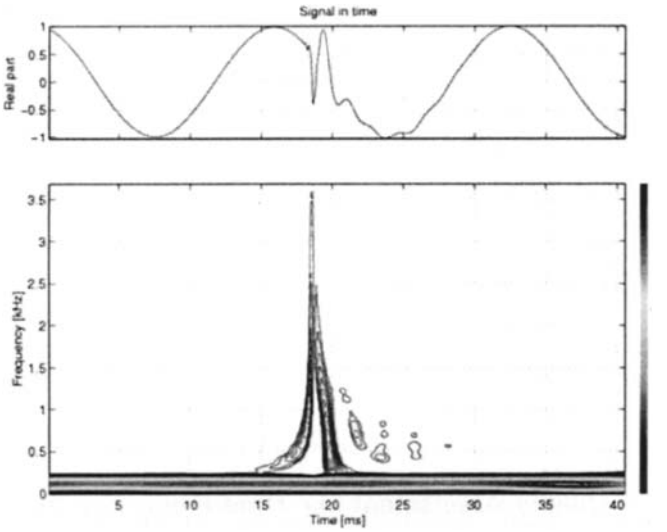


Fig. 15.1.1: Capacitor switching disturbance time series (top) and its reduced interference distribution (bottom). From [3], © 1999 IEEE.

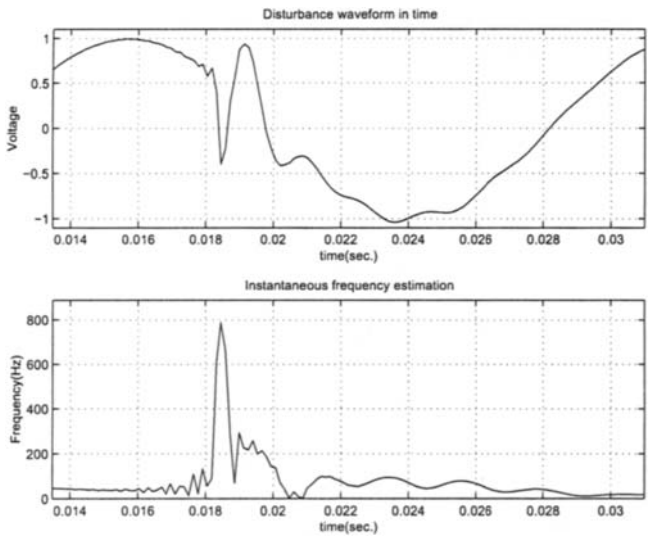


Fig. 15.1.2: Capacitor switching disturbance and corresponding instantaneous frequency based on the RID. From [3], © 1999 IEEE.

As a result, the instantaneous frequency of the disturbance provides a quantitative assessment of the transient frequency distortion. The same capacitor switching disturbance waveform (zoomed for the disturbance portion) used in Fig. 15.1.1, and its corresponding instantaneous frequency, are plotted in Fig. 15.1.2. The peak of

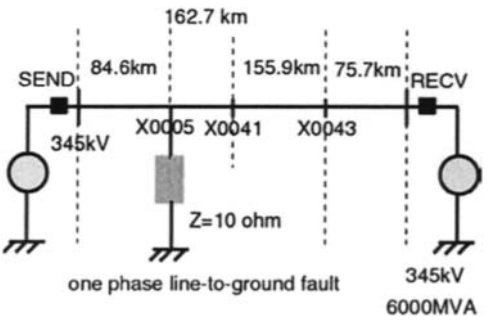


Fig. 15.1.3: EMTP simulation circuit configuration. From [6], © 2000 IEEE.

the instantaneous frequency is estimated to be around 800 Hz. During the disturbance (between times 16 ms and 28 ms), the disturbance is reflected in the change of the instantaneous frequency. After the transient disturbance the instantaneous frequency returns to 60 Hz.

15.1.3 Application of Instantaneous Frequency for Disturbance Propagation

The time of arrival of disturbance signals on high voltage transmission lines is of great interest for relay and fault localization. Traditional fault localization in a transmission line network is based on a fault-study using voltage and current measurements. The traditional methodology is subject to inaccurate results, because the calculation depends on the rough assumption of the fault impedance and the type of fault. Recently, power system monitoring systems employ GPS (Global Positioning Systems) receivers to provide time synchronized data. GPS synchronized data enables one to solve the fault location problem based on time-of-arrival of the disturbance waveforms. The propagation properties of high voltage transmission lines have been carefully treated and shown to be dispersive [6]. To treat the time synchronized disturbance data, an accurate estimation of the arrival time is critical. In this section, an application example is provided to show how the instantaneous frequency can be utilized for the arrival time estimation.

In Fig. 15.1.3 a simulation circuit diagram is provided. For a long transmission line (345kV), there occurs a typical line-to-ground fault which is 84.6 km away from 'SEND' and 394.3 km away from 'RECV' as indicated in Fig. 15.1.3. For this transmission line configuration, EMTP (Electro-Magnetic Transient Program) simulates the voltage and current disturbances. The corresponding voltage waveforms at individual buses (X0005, SEND, X0041, RECV) are provided in Fig. 15.1.4.

As the transmission line is characterized by frequency-dependent attenuation and dispersion, different frequencies suffer different amounts of attenuation and also propagate with different phase and group velocities; consequently, the wave-

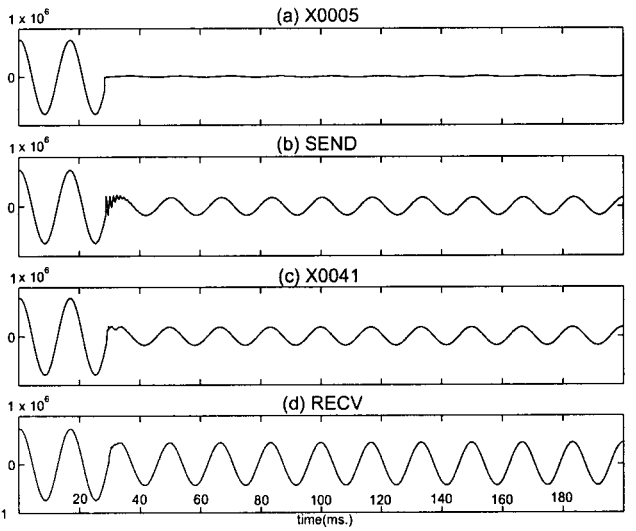


Fig. 15.1.4: Disturbance voltage waveforms recorded at individual buses. From [6], © 2000 IEEE.

forms observed at different buses appear "distorted" or "dispersed" from the original waveform. Therefore, it is difficult to assign time-of-arrival for "distorted" signals. There are perhaps many ways to determine time-of-arrival; however, in this article we focus on one, namely instantaneous frequency.

The corresponding zero sequence mode disturbance voltage is provided in Fig. 15.1.5. The zero sequence mode is a summation of the individual three-phase waveforms and is ideally zero for a balanced three-phase system. Thus it is very sensitive to a fault on any of the three phases as shown in Fig. 15.1.5. The reduced interference distribution has been calculated for the disturbance waveforms in zero sequence mode in order to generate the instantaneous frequency of the zero sequence disturbance signals. The instantaneous frequency, its peak value and time of arrival of the disturbance at various observation points are provided in Fig. 15.1.6. Note that the time axis in Fig. 15.1.6 is zoomed to within a 20-40 ms interval as indicated in Fig. 15.1.5. The time of arrival t^{arrival} has been assigned as follows:

$$t^{\text{arrival}} = \arg\{\max_t[f_i(t)]\}. \tag{15.1.3}$$

As the frequency bandwidth of the disturbance is broad since the disturbance is transient, the assignment of the arrival time via the peak instantaneous frequency is a reasonable approximation.

The arrival times and peak values of the instantaneous frequency are presented in Table 15.1.1. To convert times to distance we utilize the results of the analysis presented in reference [6], where it was shown that for a range of peak frequencies appearing in Table 15.1.1, the corresponding zero-sequence group velocity is $v_g = (2.6 \pm 0.2) \times 10^8$ m/sec. The corresponding estimates of distance are compared to the

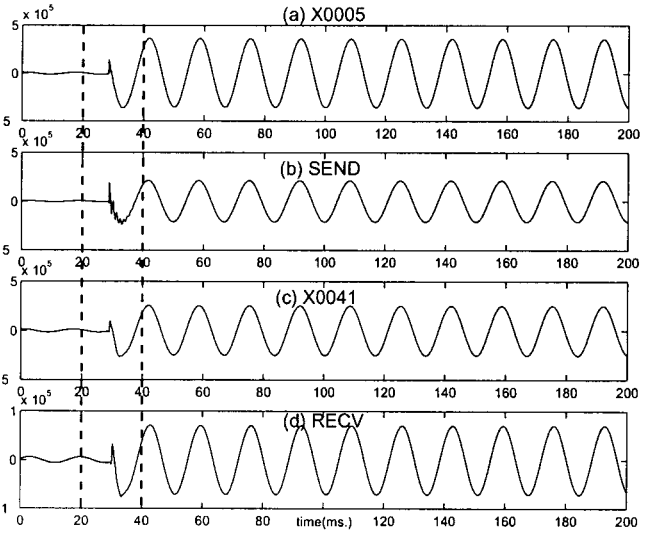


Fig. 15.1.5: Zero sequence disturbance voltage waveforms recorded at individual buses. From [6], © 2000 IEEE.

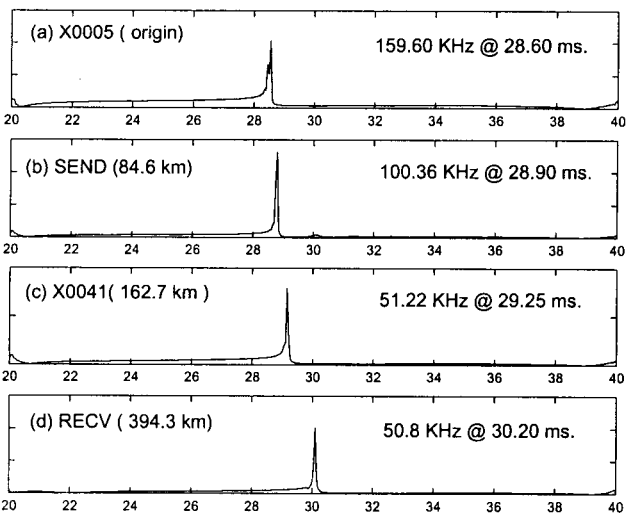


Fig. 15.1.6: Instantaneous frequency estimation of the disturbance voltage waveforms in zero sequence. From [6], © 2000 IEEE.

true distance in Table 15.1.1. Note the range of estimated distance agrees quite well with the known true distance. Note, also, that the peak instantaneous frequency is lower for the larger propagation distance. This is due to the fact that higher frequencies associated with the disturbance suffer greater attenuation than lower

Table 15.1.1: Summary of the zero sequence disturbance via RID.

Bus name (unit)	Arrival time (ms)	Peak IF (KHz)	True distance (Km)	Estimated distance (Km)
X0005	28.60	159.60	0.0	N/A
SEND	28.90	100.36	84.6	$72 \sim 84$
X0041	29.25	51.22	162.7	$156 \sim 182$
RECV	30.20	50.80	394.3	$384 \sim 448$

frequencies. Ongoing work in voltage-only distance localization involves refinement of the instantaneous frequency approach and consideration of the use of group delay.

15.1.4 Summary and Conclusions

In this article, we provided two examples of time-frequency analysis applied to power system disturbances. RID-based time-frequency analysis successfully provides a useful characterization of power system disturbance signals. In this article, we have also demonstrated that the concept of instantaneous frequency is quite useful in providing a simplified picture of the time-frequency features of a transient disturbance and in determining the arrival time of a disturbance observed in the zero-sequence mode of a three-phase system.

References

[1] J. Arrillaga, M. H. J. Bollen, and N. R. Watson, "Power quality following deregulation," *Proc. IEEE*, vol. 88, pp. 246–261, February 2000.

[2] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Trans. Power Delivery*, vol. 11, pp. 924–930, April 1996.

[3] Y. Shin, A. C. Parsons, E. J. Powers, and W. M. Grady, "Time-frequency analysis of power system disturbance signals for power quality," in *Proc. IEEE Power Engineering Soc. Summer Meeting*, vol. 1, pp. 402–407, Edmonton, AL (Canada), 18–22 July 1999.

[4] W. J. Williams and J. Jeong, "Reduced interference time-frequency distributions," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 3, pp. 74–97, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[5] L. Cohen, "Wavelet moments and time-frequency analysis," in *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, vol. 3807, pp. 434–445, Soc. of Photo-optical Instrumentation Engineers, Denver, CO, 19–21 July 1999.

[6] Y. Shin, E. J. Powers, W. M. Grady, and S. C. Bhatt, "Effects of dispersion on disturbance propagation on high voltage transmission lines," in *Proc. IEEE Power Engineering Soc. Summer Meeting*, vol. 2, pp. 851–854, Seattle, WA (USA), 16–20 July 2000.

15.2 COMBUSTION DIAGNOSIS BY TF ANALYSIS OF CAR ENGINE SIGNALS[0]

15.2.1 Knocking Combustions

A permanent challenge for car manufacturers is to increase efficiency, reduce pollution, and prolong life of internal combustion engines. A restriction to these aims is the occurrence of knock. Knock is an undesired spontaneous auto-ignition of the end gas causing a sharp increase of pressure and temperature [1]. Generally, rare knock has no effect to engine performance but frequent or very strong knock can damage the engine. Knock excites combustion chamber acoustic resonances that can be measured by special pressure sensors. But they are too expensive for use in serial vehicles. Nowadays, acceleration sensors mounted on the engine housing measure structure-borne sound as a distorted version of pressure to detect knock. Time-frequency analysis can help to understand the nature of pressure and structure-borne sound signals and to improve knock detection. The following applications for combustion diagnosis make use of the high frequency part of the signals which means frequencies above 3 kHz. Therefore, all signals were high pass filtered. Fig. 15.2.1 shows two pressure signals and a structure-borne sound signal of a BMW engine recorded simultaneously. The sensor of pressure 1 is mounted in the spark plug, the sensor of pressure 2 in the cylinder head.

15.2.2 Signal Models

Considering homogeneous gas distribution, the frequencies of the resonances depend on the speed of sound and of the combustion chamber geometry. The latter can be approximated by an ideal cylinder [2], or more accurately using finite element methods (FEM) [3]. Fig. 15.2.2 shows the instantaneous frequencies of a BMW engine estimated by FEM simulations [4]. But they do neither yield information about time instant when knock occurs nor the resonance amplitudes. The time instant of excitation is random, usually between 5° to 15° crank angle after top dead center of the piston. The amplitudes are damped due to heat and friction losses and to increasing combustion chamber volume. Previous considerations motivate to model the undisturbed pressure signal in time domain by [5]

$$X(t) = \sum_{p=1}^{P} A_p e^{-d_p(t-t_0)} \cos\left(2\pi \int_{t_0}^{t} f_p(\alpha(\tau))d\tau + \Phi_p\right) u(t-t_0). \qquad (15.2.1)$$

The pressure signal is a superposition of P resonances, index p refers to the pth component. A_p and Φ_p are random initial amplitude and phase, $e^{-d_p t}$ describes the damping of the oscillation, t_0 a suitable time instant after excitation, $u(t)$ the step

[0]Authors: **Johann F. Böhme**, Department of Electrical Engineering and Information Science, Ruhr-Universität Bochum, 44780 Bochum, Germany (boehme@sth.ruhr-uni-bochum.de) and **Sönke Carstens-Behrens**, Robert Bosch GmbH, 70049 Stuttgart, Germany (scb@ieee.org, scb@sth.ruhr-uni-bochum.de). Reviewers: LJ. Stanković and G. Matz.

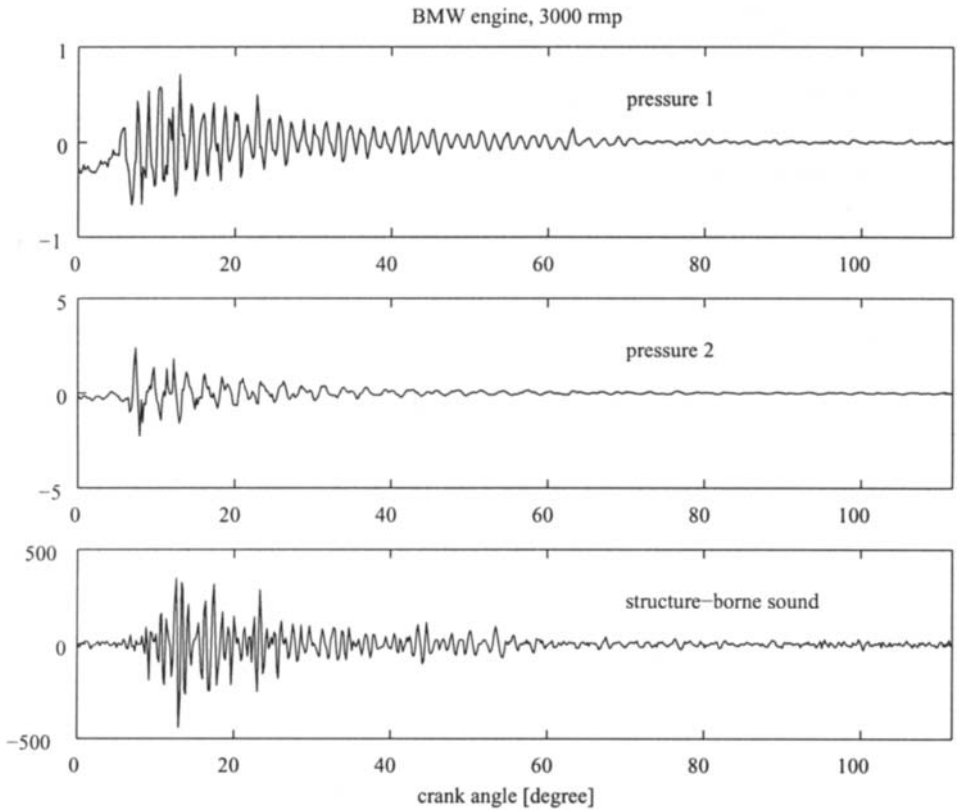


Fig. 15.2.1: Example of time signals of a knocking combustion.

function, and $f_p(\alpha(t))$ the instantaneous frequency at crank angle $\alpha(t)$. The engine housing transfer function is modeled as a linear, time-varying function $h(z(t), \tau)$, whereas time variance stems only from piston position $z(t)$ which is a function of crank angle. Merging pressure model and engine housing transfer function yields the model of the undisturbed structure-borne sound signal

$$\begin{aligned} Y(t) &= \int_{-\infty}^{\infty} h(z(t), \tau)\, X(t-\tau)\, d\tau \\ &\approx \sum_{p=1}^{P} A_p e^{-d_p(t-t_0)} \mathrm{Re}\left\{ H_{\mathrm{p}}(z(t), 2\pi f_p(\alpha(t)))\, e^{j2\pi \int_{t_0}^{t} f_p(\alpha(\tau))d\tau + \Phi_p} \right\} \end{aligned} \quad (15.2.2)$$

where

$$H_p(z(t), \omega) = \int_{-\infty}^{\infty} h(z(t), \tau)\, e^{d_p \tau} e^{-j\omega\tau}\, d\tau. \quad (15.2.3)$$

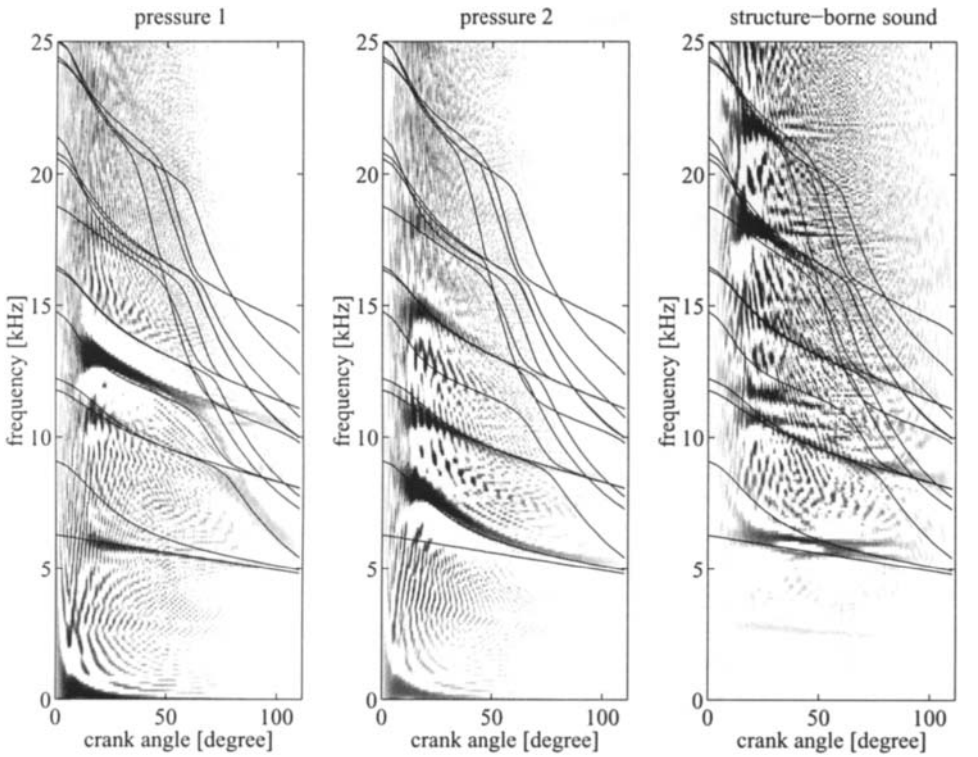


Fig. 15.2.2: Frequencies estimated by finite elements (solid lines) and estimated WVS of pressure and structure-borne sound signals of a BMW engine at 3000 rpm based on 96 realizations measured simultaneously.

The approximation is valid if the instantaneous frequencies $f_p(\alpha(t))$ do not change fast. Thus, the structure-borne sound signal is a complex amplitude modulated version of the pressure signal.

15.2.3 Signal Analysis using Wigner-Ville Spectrum

Pressure and structure-borne sound signals consist of a superposition of amplitude and frequency modulated components. Therefore, identification of the components only in time domain or frequency domain is difficult: a suitable time-frequency representation is required. The Wigner-Ville spectrum (WVS) yields a high time-frequency resolution. For a stochastic process $S(t)$ the WVS is defined by

$$W_S(t,f) = \int_{-\infty}^{\infty} r_S\left(t+\frac{1}{2}\tau, t-\frac{1}{2}\tau\right) e^{-j2\pi f\tau} d\tau \qquad (15.2.4)$$

where $r_S(t,t') = \mathrm{E}\{S(t)S^*(t')\}$ is the correlation function of $S(t)$ [see Article 9.4]. In case of constant speed and load pressure and structure-borne sound signals are

cyclo-stationary with low stochastic dependency from combustion to combustion [5]. Exploiting this fact, the correlation function can be estimated by

$$\hat{r}_S(t,t') = \frac{1}{L}\sum_{l=1}^{L} s_l(t)s_l(t') \tag{15.2.5}$$

where $s_l(t)$ is observed pressure or structure-borne sound of the lth combustion, $l = 1,\ldots,L$, with $s_l(t) = 0$ if t is not in the observation interval $[0,T]$ and $t = 0$ corresponds to top dead center for each combustion cycle. Substituting the estimated correlation function in Eq. (15.2.4) yields the estimated WVS of pressure or structure-borne sound, respectively,

$$\hat{W}_S(t,f) = \int_{-\infty}^{\infty} \hat{r}_S\left(t+\frac{1}{2}\tau, t-\frac{1}{2}\tau\right) e^{-j2\pi f\tau} d\tau \tag{15.2.6}$$

This WVS estimate has three advantages: it has a high time-frequency resolution, cross-terms are reduced effectively, and the algorithm is fast compared to other approaches that first estimate the WVS of each combustion and then take the mean over all combustions.

Model verification Fig. 15.2.2 shows the estimated WVS of the signals of two differently located pressure sensors and a structure-borne sound signal of a BMW engine. In order to stress weak signal components, signal amplitudes are scaled logarithmically, so we have to be careful in interpreting gray scales as amplitudes. In addition, Fig. 15.2.2 shows the instantaneous frequencies estimated by FEM simulations. Obviously, the pressure sensors observe different resonances due to their positions and pressure nodes and anti-nodes locations. The frequencies of the visible pressure resonances coincide well with the frequencies estimated by FEM simulations. The structure-borne sound signal shows similar behavior but there are some strange components at approximately 6.5 kHz and between 10 and 12 kHz. They may be a consequence of the engine housing transfer function, additional noise, or of insufficiently reduced cross-terms of the estimated WVS. As mentioned before, we have to be careful in interpreting the gray scale amplitudes. Nevertheless, we can recognize for pressure and structure-borne sound signals that knock occurs around 10° crank angle. Furthermore, signal power (and amplitude) is maximal at excitation of knock and decreases with increasing crank angle. In general, the results of the estimated WVS justify the decomposition of pressure and structure-borne sound signals into frequency and amplitude modulated components.

Knock sensor position optimization The number of resonances being observed by a sensor depends strongly on its position. This holds for pressure as well as for structure-borne sound. Therefore, it is very important to find a good acceleration sensor position. This can be done at test bed engines by optical methods like laser holography or by assessing the position of several acceleration sensors with statistical analysis [6]. A very simple, intuitive, and fast approach is to use the estimated

WVS. Since there are fast algorithms for its implementation, it can be calculated at the test bed while the engine is running. The operator compares the WVS of the pressure signal with those of the acceleration sensor signals. Using pressure as reference, structure-borne sound signals that contain the same resonances are more suitable for knock detection than signals with less or different resonances. As spin-off, observing the sensor signals by WVS is very useful for verifying measurement setup, a very important task at test bed measurements.

Non-equidistant sampling Knock detection usually bases on estimated resonance energies of structure-borne sound signals. As previous considerations have shown, the frequencies vary with crank angle. Optimum estimation of resonance energies can therefore be achieved by time-varying filtering. Alternatively, optimized non-equidistant sampling in time domain leads to resonances with approximately constant frequencies [7]. Using Wigner-Ville analysis, parameters can be found to optimally compensate the frequency modulation in a crank angle interval with significant resonance energy. Then, ordinary band pass filtering is sufficient to estimate the resonance energies well.

15.2.4 Signal Analysis using S-Method

The estimated WVS is a useful tool for analyzing a large number of combustions at once. But if the number of combustions is not large enough, cross-terms disturb the estimated spectrum. In many cases, we are interested in analyzing single combustions. There are a couple of good WVS estimators reducing or avoiding cross-terms; see e.g. [8, 9]. Among these, the S-method has been proven to be an appropriate tool to represent single knock signals in time-frequency domain [10]. Using the short-time Fourier transform of signal $s(t)$

$$F_s^w(t,\omega) = \int_{-\infty}^{\infty} s(t+\tau)\, w(\tau)\, e^{-j\omega\tau} d\tau \tag{15.2.7}$$

where $w(\tau)$ is a lag window, the S-Method can by defined as

$$SM(t,\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} P(\theta)\, F_s^w(t,\omega+\theta)\, F_s^{w*}(t,\omega-\theta)\, d\theta \tag{15.2.8}$$

with an appropriate window $P(\theta)$. The width of $P(\theta)$ affects the behavior of the S-Method with regard to reduction of cross-terms significantly. For a detailed discussion of the S-Method see Article 6.2. Exploiting Fast Fourier Transform algorithms, the S-Method can be implemented very efficiently [11], so that real-time processing is possible on a fast DSP.

Observation of single combustions Fig. 15.2.3 shows the estimated WVS and the S-Method of a single knocking combustion. The estimated WVS seems to contain mainly one component. There are some more but it is very difficult to distinguish between auto-terms and cross-terms. The time-frequency concentration of the S-Method is worse but it is easy to identify at least four components. Therefore, the

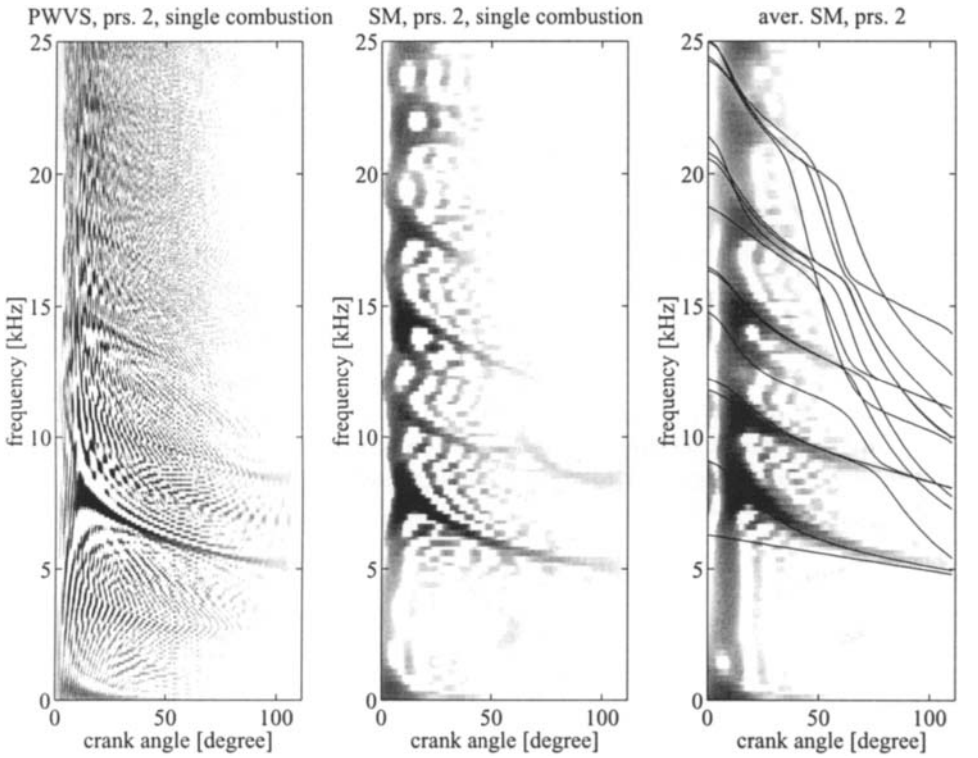


Fig. 15.2.3: Left: estimated WVS of a single knocking combustions of a BMW engine at 3000 rpm observed by pressure sensor 2; middle: S-Method of the same signal; right: S-Method averaged over 96 realizations.

S-Method is much more suitable to observe single combustions than the estimated WVS. Normally, the WVS is estimated by averaging over a large number of realizations. Analogue, the S-Method of a large number of single knocking combustions was averaged in Fig. 15.2.3 (right). Comparing this representation to the according WVS in Fig. 15.2.2 (middle), we recognize the worse resolution of the averaged S-Method. But the amplitude and frequency modulation is observable. In contrast to the WVS, the averaged S-Method is able to resolve the component that starts at 15 kHz at 0° crank angle which is the strongest component in pressure 1; see Fig. 15.2.2 (left). In the WVS this component is covered by cross-terms. Apparently, the number of realizations was too small to reduce cross-terms sufficiently.

Power and energy estimation As demonstrated before, the S-Method can be used to investigate single combustions. As shown in [10], it is possible to track the resonances and to estimate the power $A_p^2(t)$ of component p at each time instant t

by integrating along the frequency axis in the component's band $B_p(t)$:

$$\hat{A}_p^2(t) = \frac{1}{2\pi} \int_{B_p(t)} SM(t,\omega)\, d\omega. \tag{15.2.9}$$

In the simplest case, the band width $B_p(t)$ is constant. But considering knock signals, the band width decreases with crank angle; see Fig. 15.2.2. So the signal-to-noise ratio can be improved by adapting the band width. In Article 6.2 an adaption algorithm is proposed. Finally, integration of $\hat{A}_p^2(t)$ along the observation interval $[0, T]$ yields the estimated resonance energy $\hat{E}_p$ of component p

$$\hat{E}_p = \int_0^T \hat{A}_p^2(t)\, dt. \tag{15.2.10}$$

Knock detection The main application of estimating resonance energies of internal combustion engines is the detection of knock. A knock detector in today's cars based on structural-borne sound usually estimates the signal energy in a wide band, e.g. 5 to 12 kHz. The energy is compared to a threshold which yields the knock detection. This is a simple, fast and successfully applied method. But it can be improved by taking into consideration that signal energy is the sum of the energies of single components. Since the engine housing transfer function may attenuate or amplify the components differently, we get a wrong idea of the energy distribution inside the cylinder which is used as reference at test bed engines. Among other time-frequency detectors like those presented in Article 12.4 or [12], [13], the S-method is a useful tool for estimating the resonance energies of single components and is therefore advantageous for detecting knock; see e.g. [14].

15.2.5 Summary and Conclusions

On the basis of the WVS and the S-method, it was demonstrated that time-frequency representations are suitable and very useful tools for the task of combustion diagnosis of internal combustion engines. They allow an intuitive insight in relevant signal parameters and properties like amplitude and frequency modulation. Therefore, they simplify and improve investigations during the development and application of new engines. Furthermore, they help to improve signal processing for higher engine efficiency.

References

[1] J. B. Heywood, *Internal Combustion Engine Fundamentals.* McGraw-Hill, 1988.

[2] R. Hickling, D. A. Feldmaier, F. H. K. Chen, and J. S. Morel, "Cavity resonances in engine combustion chambers and some applications," *J. Acoustical Soc. of America*, vol. 73, pp. 1170–1178, April 1983.

[3] D. Scholl, C. Davis, S. Russ, and T. Barash, "The volume acoustic modes of spark-ignited internal combustion chambers," in *SI Engine Combustion*, no. 1315 in SAE Special Publications, pp. 1–8, Soc. of Automotive Engineers, 1998. Proc. SAE Internat. Congress & Exposition, Detroit, MI, 23–26 February 1998. SAE paper no. 980893.

[4] S. Carstens-Behrens and A. Podvoiski, "Estimation of combustion chamber resonances for improved knock detection in spark ignition engines," in *Proc. 18th CAD-FEM Users' Meeting, Internationale FEM-Technologietage*, Friedrichshafen, Germany, 20–22 September 2000.

[5] J. F. Böhme and D. König, "Statistical processing of car engine signals for combustion diagnosis," in *Proc. Seventh IEEE Workshop on Statistical Signal and Array Processing (SSAP-94)*, pp. 369–374, Quebec City, 26–29 June 1994.

[6] A. M. Zoubir, "Bootstrap multiple tests: An application to optimum sensor location for knock detection," *Applied Signal Processing*, vol. 1, pp. 120–130, 1994.

[7] D. König, "Application of time-frequency analysis for optimum non-equidistant sampling of automotive signals captured at knock," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'96)*, vol. 5, pp. 2746–2749, Atlanta, GA, 7–10 May 1996.

[8] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 37, pp. 862–871, June 1989.

[9] Y. Zhao, L. E. Atlas, and R. J. Marks II, "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1084–1091, July 1990.

[10] L. Stanković and J. F. Böhme, "Time-frequency analysis of multiple resonances in combustion engine signals," *Signal Processing*, vol. 79, pp. 15–28, November 1999.

[11] L. Stanković, "A method for time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 225–229, January 1994.

[12] G. Matz and F. Hlawatsch, "Time-frequency subspace detectors and application to knock detection," *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, vol. 53, pp. 379–385, December 1999.

[13] B. Samimy and G. Rizzoni, "Mechanical signature analysis using time-frequency signal processing: Application to internal combustion engine knock detection," *Proc. IEEE*, vol. 84, pp. 1130–1343, September 1996.

[14] S. Carstens-Behrens, M. Wagner, and J. F. Böhme, "Detection of multiple resonances in noise," *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, vol. 52, no. 5, pp. 285–292, 1998.

15.3 POWER CLASS TIME-FREQUENCY REPRESENTATIONS AND THEIR APPLICATIONS[0]

15.3.1 Power Class Quadratic Time-Frequency Representations

Various classes of quadratic time-frequency representations (QTFRs) are best suited for analyzing signals with certain types of time-frequency (TF) geometries. For example, when a signal has constant TF characteristics, Cohen's-class QTFRs (with signal-independent kernels; i.e. the quadratic class as defined on p. 68) [1,2] are most appropriate. The aforementioned QTFR classification is based on the grouping together of all QTFRs that satisfy the same two signal transformation covariance properties (see [3, 4] and Articles 4.3 and 5.6). Specifically, Cohen's class [1, 2] consists of QTFRs that are covariant to constant (nondispersive) time shifts and frequency shifts of the signal whereas the affine class (see [2, 5, 6] and Article 7.1) consists of QTFRs that are covariant to scale changes (dilations) and constant time shifts. Furthermore, the hyperbolic class [7,8] consists of QTFRs that are covariant to scale changes and hyperbolic dispersive time shifts and are best suited to analyze signals with hyperbolic (nonlinear) group delay. When the analysis signal has a group delay that is a power function of frequency, the aforementioned QTFRs do not provide an adequate representation as they do not match power TF characteristics. Thus, we designed *power class* QTFRs to successfully localize signals along their power law group delay functions [9, 10].

The importance of power QTFRs is pronounced by the fact that many applications involve signals with dispersive group delays governed by a power law that corresponds to some power parameter κ. Examples of such signals include the dispersive propagation of a shock wave in a steel beam ($\kappa = 1/2$), trans-ionospheric chirps measured by satellites ($\kappa = -1$), acoustical waves reflected from a spherical shell immersed in water, various cetacean mammal whistles, and signal solutions of the diffusion equation ($\kappa = 1/2$) (e.g., waves propagating along uniform distributed RC transmission lines). Power laws can also be used to roughly approximate other, more complex, group delays. References for these applications can be found in [9].

Localized signal analysis application. The type of signals found in the applications mentioned above constitute the family of *power impulses* that best typifies the power TF geometry. Power impulses are defined in the frequency domain as

$$I_c^{(\kappa)}(f) \triangleq \sqrt{|\tau_\kappa(f)|}\, e^{-j2\pi c\, \Lambda_\kappa\left(\frac{f}{f_r}\right)} = \sqrt{(|\kappa|/f_r)\,|f/f_r|^{\kappa-1}}\, e^{-j2\pi c\, \mathrm{sgn}(f)|f/f_r|^\kappa} \quad (15.3.1)$$

with monotonic phase spectrum $\Lambda_\kappa(f/f_r) = \mathrm{sgn}(f)\,|f/f_r|^\kappa$ and power group delay

[0]Authors: **Antonia Papandreou-Suppappola**, Telecommunications Research Center, Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-7206 USA (papandreou@asu.edu), **Franz Hlawatsch**, Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology, A-1040 Vienna, Austria (fhlawats@pop.tuwien.ac.at), and **G. Faye Boudreaux-Bartels**, Department of Electrical and Computer Engineering, University of Rhode Island, Kingston, RI 02881 USA (boud@ele.uri.edu). Reviewers: J. Bertrand and J. P. Ovarlez.

$\tau_g(f) = c\,\tau_\kappa(f) = c\,\frac{|\kappa|}{f_r}\,|f/f_r|^{\kappa-1} = c\,\frac{d}{df}\Lambda_\kappa(f/f_r)$ with $f \in \Re$. Here, $\mathrm{sgn}(f)$ provides the sign (± 1) of the frequency variable f, and $f_r > 0$ is a fixed reference frequency. For successful analysis, an ideal QTFR T must be localized along the group delay $\tau_g(f)$ of the power impulse in Equation (15.3.1). In particular,

$$T_{I_c^{(\kappa)}}(t,f) = |\tau_\kappa(f)|\;\delta\big(t - c\,\tau_\kappa(f)\big) \tag{15.3.2}$$

is very desirable in many applications where information about the signal analyzed could be obtained from the localized curve $t = c\,\tau_\kappa(f)$ in the TF plane. For example, the localization could be used in applications such as signal classification or estimation of the parameter c in (15.3.1). As we will show next, some power class QTFRs ideally provide the localized representation in (15.3.2) for analyzing power impulses as well as other signals with power group delay.

The power law TF structure can also be observed in applications where a system can cause a time shift that varies dispersively in frequency to an input signal with Fourier transform $X(f)$ yielding the output signal $Y(f) = e^{-j2\pi c\Lambda_\kappa(f/f_r)}\,X(f)$. Thus, power QTFRs could be used successfully in applications where a signal with constant group delay passes through a system with power dispersive TF characteristics that transforms the signal's constant group delay to a power group delay [9]. For example, the ocean is a medium (system) that could cause power dispersive changes to an underwater communications information message and its echoes. These changes could be accounted for at the receiver when matching processing tools, like power QTFRs, are used for detection.

The power QTFR classes. Following the covariance-based classification method, for an analysis signal $x(t)$ with Fourier spectrum $X(f)$, we define all κth power class QTFRs, $T_X^{(\kappa)}(t,f)$, to satisfy two specific covariance properties [9, 10]. The first property is covariance to scale changes on $x(t)$, i.e.,

$$T^{(\kappa)}_{\mathcal{C}_a X}(t,f) \;=\; T_X^{(\kappa)}(at, f/a)\,, \tag{15.3.3}$$

where the scaling operator $\mathcal{C}_a$ is defined as $(\mathcal{C}_a X)(f) = X(f/a)\,/\sqrt{|a|}$. The second property is covariance to power time shifts on $x(t)$ that correspond to frequency-dependent shifts, $\tau_\kappa(f)$, in the signal's group delay . Specifically,

$$T^{(\kappa)}_{\mathcal{D}_c^{(\kappa)} X}(t,f) \;=\; T_X^{(\kappa)}\big(t - c\,\tau_\kappa(f), f\big) \;=\; T_X^{(\kappa)}\left(t - c\,(\kappa/f_r)|f/f_r|^{\kappa-1},\; f\right) \tag{15.3.4}$$

for $f \in \Re$. The effect of the power time shift operator $\mathcal{D}_c^{(\kappa)}$ is given by [9]

$$(\mathcal{D}_c^{(\kappa)} X)(f) = e^{-j2\pi c\,\Lambda_\kappa(f/f_r)}\,X(f) = e^{-j2\pi c\,\mathrm{sgn}(f)|f/f_r|^\kappa}\,X(f)\,. \tag{15.3.5}$$

Here, $c \in \Re$ and $\kappa \in \Re$ ($\kappa \neq 0$) is the power parameter associated with each power class. The κth power function $\Lambda_\kappa(b) = \mathrm{sgn}(b)\,|b|^\kappa$, $b \in \Re$, corresponds[1] to a

[1]The definition $\Lambda_\kappa(b) = \mathrm{sgn}(b)\,|b|^\kappa$ versus $\Lambda_\kappa(b) = b^\kappa$ extends the power function to $b < 0$ so that $\Lambda_\kappa(b)$ is an odd, strictly monotonic function constituting a one-to-one mapping from $\Re$ to $\Re$.

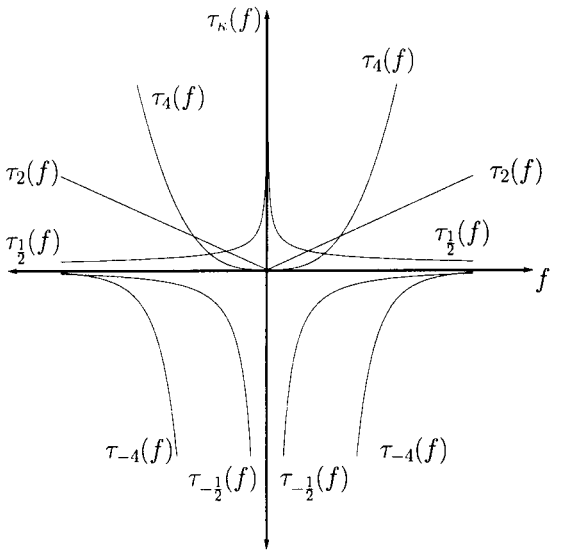


Fig. 15.3.1: The power group delay $\tau_\kappa(f)$ for various choices of the power parameter κ.

transformation of the phase spectrum of the signal as shown in Equation (15.3.5). The frequency-dependent time shift $\tau_\kappa(f)$ in (15.3.4) corresponds to the derivative of the power function, i.e., $\tau_\kappa(f) = \frac{d}{df}\Lambda_\kappa(f/f_r) = \frac{\kappa}{f_r}|f/f_r|^{\kappa-1}$. Fig. 15.3.1 depicts $\tau_\kappa(f)$ in the TF plane for various choices of the power parameter κ.

The importance of the κth power class QTFRs is directly linked to the two covariances in (15.3.3) and (15.3.4). The power time shift covariance in (15.3.4) is useful in analyzing signals passing through dispersive systems with power law group delay or signals localized along power law curves in the TF plane. On the other hand, the scale covariance in (15.3.3) is important for multiresolution analysis.

Power class formulation. It was shown in [9] that any QTFR of the κth power class can be expressed as

$$T_X^{(\kappa)}(t,f) = \frac{1}{|f|}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Gamma_T(f_1/f, f_2/f) \cdot e^{j2\pi(tf/\kappa)[\Lambda_\kappa(f_1/f)-\Lambda_\kappa(f_2/f)]}\, X(f_1)\, X^*(f_2)\, df_1\, df_2\,, \qquad (15.3.6)$$

where the two-dimensional (2-D) kernel $\Gamma_T(b_1, b_2)$ uniquely characterizes the QTFR. Specific choices of $\Gamma_T(b_1, b_2)$ define specific QTFRs $T^{(\kappa)}$ in the κth power class. Also note that a different power class is obtained by varying κ in (15.3.6). When $\kappa = 1$ we obtain the affine class [see Article 7.1], which is an important special case of the power classes corresponding to the constant (nondispersive) time shift $\tau_1(f) \equiv 1/f_r$.

The κth power class QTFR in (15.3.6) can also be obtained via a unitary warping operation (see Articles 4.5 & 5.6 and [9–11]). Specifically, if $T_X^{(A)}(t,f)$ is a QTFR

of the affine class, then the corresponding κth power class QTFR, $T_X^{(\kappa)}(t,f)$, can be obtained by warping the affine class QTFR, $T_X^{(A)}(t,f)$, according to [9]

$$\begin{aligned} T_X^{(\kappa)}(t,f) &= T^{(A)}_{\mathcal{U}_\kappa X}\left(\frac{t}{f_r\tau_\kappa(f)}, f_r\Lambda_\kappa(f/f_r)\right) \\ &= T^{(A)}_{\mathcal{U}_\kappa X}\left(\frac{t}{\kappa|f/f_r|^{\kappa-1}}, f_r\mathrm{sgn}(f)|f/f_r|^\kappa\right). \end{aligned} \tag{15.3.7}$$

Here, the unitary[2] frequency axis warping operator [9,10] $\mathcal{U}_\kappa$ is given by

$$(\mathcal{U}_\kappa X)(f) = \frac{X\big(f_r\Lambda_\kappa^{-1}(f/f_r)\big)}{\sqrt{f_r\,|\tau_\kappa(f_r\Lambda_\kappa^{-1}(f/f_r))|}} = \frac{1}{\sqrt{|\kappa|}|f/f_r|^{\frac{\kappa-1}{2\kappa}}} X\Big(f_r\mathrm{sgn}(f)|f/f_r|^{\frac{1}{\kappa}}\Big)$$

where the inverse function $\Lambda_\kappa^{-1}(b)$ satisfies $\Lambda_\kappa^{-1}(\Lambda_\kappa(b)) = \Lambda_\kappa(\Lambda_\kappa^{-1}(b)) = b$. The QTFRs of the affine class are defined as [2]

$$T_X^{(A)}(t,f) = \frac{1}{|f|}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Gamma_T(f_1/f, f_2/f)\, e^{j2\pi t(f_1-f_2)}\, X(f_1)\, X^*(f_2)\, df_1 df_2\,,$$

where $\Gamma_T(b_1,b_2)$ is a 2-D kernel characterizing the affine class QTFR (cf. (15.3.6)). Note that $T_X^{(A)}(t,f) = T_X^{(\kappa)}(t,f)\big|_{\kappa=1}$. The unitary warping relation in (15.3.7) preserves certain desirable characteristics of the affine class while transforming other ones to match the dispersive nature of the signals to be analyzed by power class QTFRs. For example, whereas both classes preserve scale changes of the signal, only the affine class preserves constant (nondispersive) time shifts. On the other hand, the warping in (15.3.7) transforms constant time shifts to power dispersive time shifts in the power class, and thus the constant time shift covariance of the affine class is transformed into the power time shift covariance of the power classes. The warping also provides an efficient method for computing power class QTFRs when algorithms for computing affine class QTFRs are available [10].

Class members. Specific QTFRs of the power classes satisfy various desirable properties in addition to the covariance properties in (15.3.3)–(15.3.4) satisfied by *all* members of the power classes. Some power class QTFRs of particular importance include the *power Wigner distribution*, the *powergram*, the *smoothed pseudo power Wigner distribution* [9], and the *Bertrand P_κ-distributions* (see Article 7.1 and [6,9]). All these QTFRs have counterparts in the affine class by virtue of the power warping relation in (15.3.7). For example, the power Wigner distribution,

$$\begin{aligned} W_X^{(\kappa)}(t,f) &= \left|\frac{f}{\kappa}\right| \int_{-\infty}^{\infty} X\Big(f\Lambda_\kappa^{-1}\Big(1+\frac{\beta}{2}\Big)\Big)\, X^*\Big(f\Lambda_\kappa^{-1}\Big(1-\frac{\beta}{2}\Big)\Big)\, e^{j2\pi\frac{tf}{\kappa}\beta}\, \frac{d\beta}{|1-\frac{\beta^2}{4}|^{\frac{\kappa-1}{2\kappa}}} \\ &= W_{\mathcal{U}_\kappa X}\Big(t/(f_r\tau_\kappa(f)),\, f_r\,\Lambda_\kappa(f/f_r)\Big) \end{aligned}$$

[2]Unitarity of the operator $\mathcal{U}_\kappa$ implies that $\mathcal{U}_\kappa$ preserves inner products. Specifically, the operator satisfies $\int_{-\infty}^{\infty}(\mathcal{U}_\kappa X)(f)\,(\mathcal{U}_\kappa X)^*(f)\,df = \int_{-\infty}^{\infty} X(f)\,X^*(f)\,df$.

(cf. (15.3.7)), is the power warped version of the well-known Wigner distribution (WD) , $W_X(t,f) = \int_{-\infty}^{\infty} X(f+\frac{\nu}{2})X^*(f-\frac{\nu}{2})\,e^{j2\pi t\nu}\,d\nu$ [1–3]. As such, it satisfies many properties such as a specific set of marginal properties and the perfect TF localization property for power impulses in (15.3.2). Just like the WD, the power WD contains oscillatory and partly negative cross terms when multicomponent signals are analyzed (see Article 4.2 and [10]). In order to suppress cross terms, a specific type of smoothing can be applied that is matched to the power TF geometry. The powergram and the smoothed pseudo power WD apply such a smoothing to the power WD, at the expense of the loss of some properties (such as the marginal properties) and the loss of TF resolution. The Bertrand P_κ-distributions (see Article 7.1 and [6]) are also perfectly localized for power impulses; moreover, they are the only power class QTFRs that preserve constant time shifts in addition to power dispersive time shifts. Power class members and their properties are discussed in detail in [9]. Next, we present examples with both synthetic and real data.

15.3.2 Power Class Applications

Synthetic data analysis example. The discrete implementation of power QTFRs (outlined in [9, 10]) was applied to analyze a two-component signal consisting of two power impulses with power parameter $\kappa_{\text{signal}} = 3$. For computational purposes, the impulses are windowed in the frequency domain. Figs. 15.3.2(a) and 15.3.2(b) show the results obtained with the power WD and a smoothed pseudo power WD with a very short smoothing window. Both QTFRs have power parameter $\kappa = 3$, matched to the power impulse parameter κ_{signal}. The power WD in Fig. 15.3.2(a) has very good TF concentration but large cross terms [10] which are effectively suppressed in the smoothed pseudo power WD in Fig. 15.3.2(b) with hardly any loss of TF concentration. Also shown (in Figs. 15.3.2(c) and 15.3.2(d)) are the results obtained with the WD and an affine-smoothed pseudo WD, both members of the affine class [2] (i.e., both QTFRs have power parameter $\kappa = 1 \neq \kappa_{\text{signal}}$). The WD in Fig. 15.3.2(c) is not matched to the power impulses, displaying complicated cross terms. The affine-smoothed pseudo WD in Fig. 15.3.2(d) does not suppress all the cross terms and has a larger loss of TF concentration than does the smoothed pseudo power WD in Fig. 15.3.2(b). Although all QTFRs in Fig. 15.3.2 are scale covariant, the results of the two power QTFRs with $\kappa = 3$ in Figs. 15.3.2(a) and 15.3.2(b) are better than those of the two affine QTFRs with $\kappa = 1$ in Figs. 15.3.2(c) and 15.3.2(d) because the former two are optimally matched to the $\kappa_{\text{signal}} = 3$ power law group delays of the power impulse signal components.

In order to further demonstrate the effect of mismatch in the signal parameter κ_{signal} and the QTFR power parameter κ, Figs. 15.3.2(e) and 15.3.2(f) show the results obtained when analyzing the above signal using the power WD and a smoothed pseudo power WD with QTFR power parameter $\kappa = 4$. Note that in Figs. 15.3.2(e) and 15.3.2(f) the power parameter of the power class QTFRs, $\kappa = 4$, is different from that of the signal, $\kappa_{\text{signal}} = 3$. The smoothed pseudo power WD in Fig. 15.3.2(f) has better cross term suppression and better TF concentration along

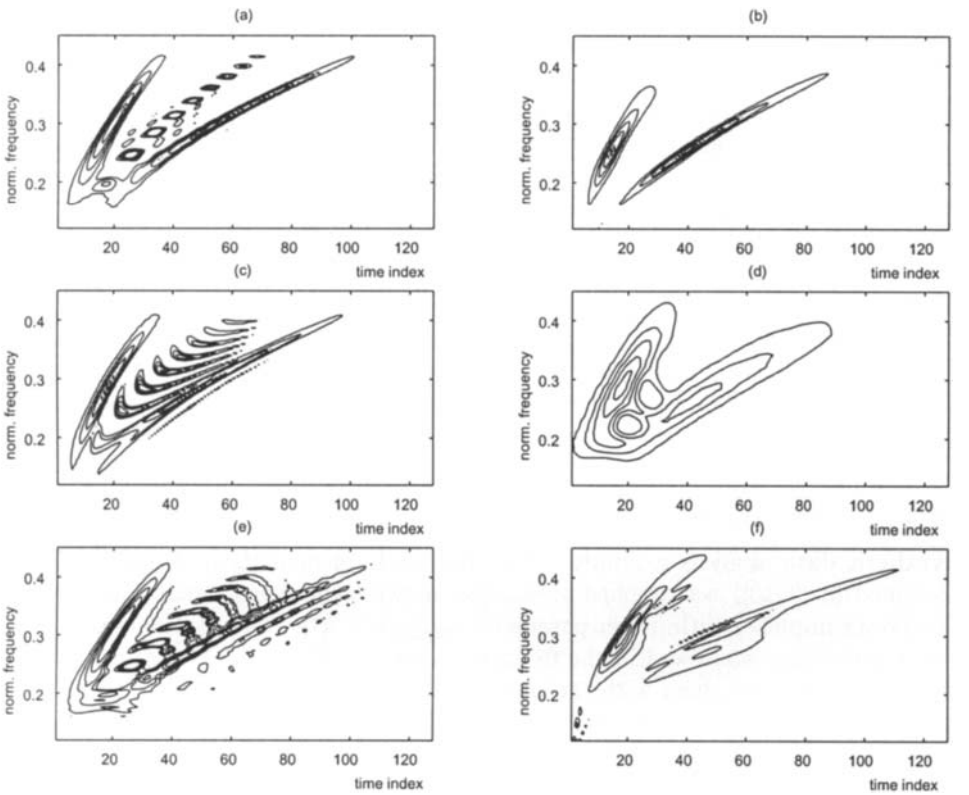


Fig. 15.3.2: Power class analysis of a two-component analytic signal consisting of the sum of two windowed power impulses with signal power parameter $\kappa_{\text{signal}} = 3$. (a) Power WD with $\kappa = 3$, (b) smoothed pseudo power WD with $\kappa = 3$, (c) WD ($\kappa = 1$), (d) affine-smoothed pseudo WD ($\kappa = 1$), (e) power WD with $\kappa = 4$, and (f) smoothed pseudo power WD with $\kappa = 4$.

the true group delay than the affine-smoothed pseudo WD in Fig. 15.3.2(d) since the power parameter mismatch in Fig. 15.3.2(f) is smaller than in Fig. 15.3.2(d).

Real data analysis example. Next, we demonstrate the use of power class QTFRs for analyzing real data with dispersive TF structure. Fig. 15.3.3 shows two power class QTFRs with $\kappa = 0.35$ and two affine ($\kappa = 1$) QTFRs of the measured impulse response of a steel beam with rectangular cross section[3] [12]. The impulse response was obtained by lightly tapping one end of the steel beam in the direction orthogonal to the flat side of the beam. Bending waves travel along the beam until they are reflected at the free end. They return to the point of impact, are reflected again, etc., thereby producing a series of echoes with increasing dispersion. The QTFRs in

[3]The data was obtained by J. Woodhouse in an experiment at Cambridge University. We are grateful to D. Newland and J. Woodhouse for making this data accessible to us.

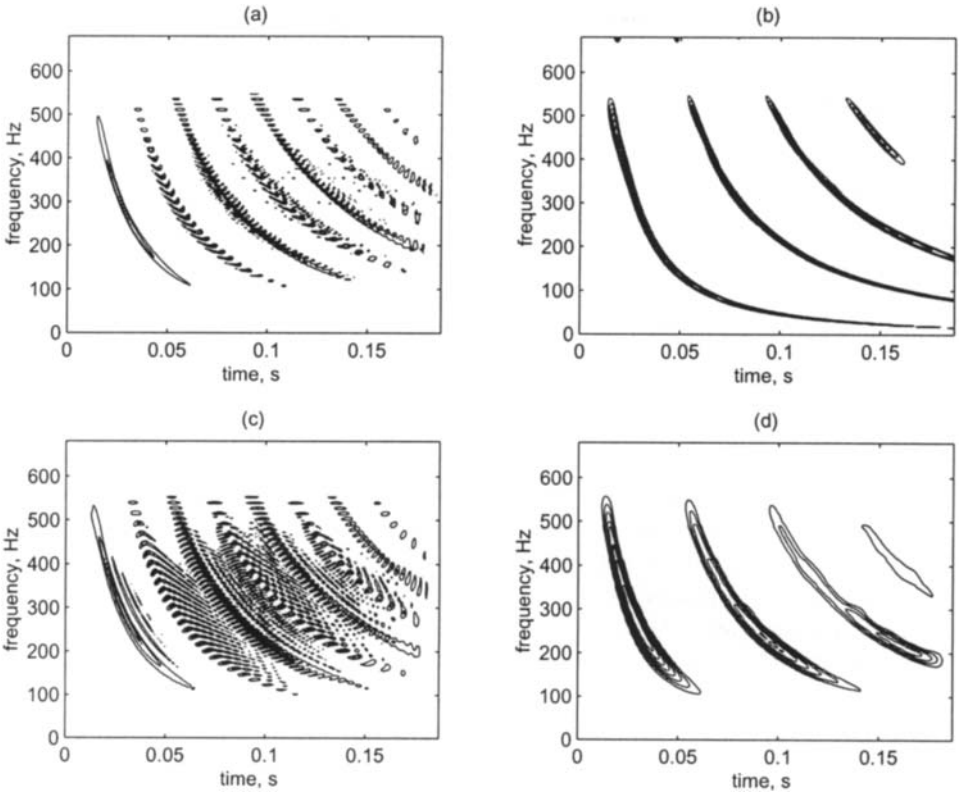


Fig. 15.3.3: Power class analysis of a bandpass-filtered segment of the measured impulse response of a steel beam (sampling freq. 4,096 Hz). (a) Power WD with $\kappa = 0.35$, (b) smoothed pseudo power WD with $\kappa = 0.35$, (c) WD ($\kappa = 1$), and (d) affine-smoothed pseudo WD ($\kappa = 1$).

Fig. 15.3.3 display a bandpass-filtered segment of the measured impulse response. As can be seen, the smoothed pseudo power WD with $\kappa = 0.35$ in Fig. 15.3.3(b) shows better resolution and/or cross term suppression than the other three QTFRs depicted. The specific value of $\kappa = 0.35$ was chosen empirically to match the TF curvature of the primary reflection.

15.3.3 Summary and Conclusions

This article presented QTFR classes specifically matched to signals and systems with power law group delay characteristics. These *power* QTFRs preserve scale changes and power law frequency-dependent time shifts of the signal under analysis. Thus, these QTFRs are potentially useful in applications where a propagation medium causes power dispersive time shifts as was demonstrated using a real data example. The implementation of power QTFRs can be based on a warping transformation that relates the κth power class with the affine class. Successful application of power

class QTFRs presupposes sufficient *a priori* knowledge about the signal to aid in choosing the appropriate power parameter κ.

References

[1] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] P. Flandrin, *Time-Frequency/Time-Scale Analysis.* San Diego: Academic Press, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[3] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[4] 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 Processing: A Review Journal*, vol. 8, no. 1, pp. 3–48, 1998.

[5] O. Rioul and P. Flandrin, "Time-scale energy distributions: A general class extending wavelet transforms," *IEEE Trans. Signal Processing*, vol. 40, pp. 1746–1757, July 1992.

[6] J. Bertrand and P. Bertrand, "A class of affine Wigner functions with extended covariance properties," *J. of Mathematical Physics*, vol. 33, pp. 2515–2527, July 1992.

[7] 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 Processing*, vol. 41, pp. 3425–3444, December 1993. Special Issue on Wavelets and Signal Processing.

[8] F. Hlawatsch, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels, "The hyperbolic class of quadratic time-frequency representations—Part II: Subclasses, intersection with the affine and power classes, regularity, and unitarity," *IEEE Trans. Signal Processing*, vol. 45, pp. 303–315, February 1997.

[9] 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 Processing*, vol. 47, pp. 3067–3083, November 1999.

[10] A. Papandreou-Suppappola, F. Hlawatsch, and G. F. Boudreaux-Bartels, "Power class time-frequency representations: Interference geometry, smoothing, and implementation," in *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pp. 193–196, Paris, 18–21 June 1996.

[11] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: A new twist on signal processing," *IEEE Trans. Signal Processing*, vol. 43, pp. 2269–2282, October 1995.

[12] D. E. Newland, "Time-frequency and time-scale analysis by harmonic wavelets," in *Signal Analysis and Prediction* (A. Prochazka, J. Uhlir, P. J. W. Rayner, and N. G. Kingsbury, eds.), ch. 1, Boston, MA: Birkhäuser, 1998.

15.4 IMAGE DISTORTION ANALYSIS USING THE WIGNER-VILLE DISTRIBUTION [0]

This article deals with the 2D Wigner-Ville distribution (WVD) in the context of image analysis applications. The properties that motivate the use of 2D WVD in image analysis are reviewed. The important issue of choosing the analytic image is emphasized, and practical implementation aspects are discussed.

The usefulness of Wigner-Ville distribution for image analysis is demonstrated by an application to image dissimilarities measurement. The WVD-based measure is correlated with subjective human evaluation, which is the premise towards an image quality assessor developed on this principle.

15.4.1 Image Quality and Joint Spatial/Spatial-Frequency Representations

The Wigner-Ville distribution (see Section 2.1.4) has been proved to be a powerful tool for analyzing the time-frequency distribution of nonstationary signals. WVDs were successfully applied in capturing essential nonstationary image structures [1–3]. In this context, the properties of joint spatial/spatial-frequency representations of images led to other applications of WVD to image processing [3,4].

This article aims to present the WVD from the point of view of image analysis, that is, to emphasis the desired properties and implementation considerations, and to present a specific application, namely analyzing and tracking image distortions for computing an image quality measure. While the WVD is widely used in applications involving 1D signals, the extension to multidimensional signals, in particular to 2D images, has not reached a similar development.

With the increasing use of digital video compression and transmission systems image quality assessment has become a crucial issue. In the last decade, there have been proposed numerous methods for image distortion evaluation inspired from the findings on Human Visual System (HVS) mechanisms [5]. In the vision research community, it is generally acknowledged that the early visual processing stages involve the creation of a joint spatial/spatial-frequency representation [6]. This motivates the use of the Wigner-Ville distribution as a tool for analyzing the effects induced by applying a distortion to a given image. The simple Wigner-based distortion measure introduced in this article does not take into account the masking effect. This factor needs to be introduced in future research. The qualitative correlation of the WVD-based measure with subjective human evaluation is illustrated through experimental results. This measure could be used for image quality assessment, or as a criterion for image coder optimization, or for bilinear problems encountered in computer vision, or for image segmentation.

---

[0]Authors: **A. Beghdadi**, L2TI-Institute Galilée, Université Paris 13, FR-93430 Villetaneuse, France (beghdadi@l2ti.univ-paris13.fr), and **R. Iordache**, Signal Processing Laboratory, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland (riordache@noos.fr).
Reviewers: B. Boashash, H. Wechsler and K. Abed-Meraim.

15.4.2 Continuous 2D Wigner-Ville Distribution

The 2D Wigner-Ville distribution (WVD) of a 2D image $f(x,y)$ is defined as [4]:

$$W_f(x,y,u,v) = \int_{\mathbb{R}^2} f\left(x+\tfrac{\alpha}{2}, y+\tfrac{\beta}{2}\right) f^*\left(x-\tfrac{\alpha}{2}, y-\tfrac{\beta}{2}\right) e^{-j2\pi(\alpha u+\beta v)}\, d\alpha\, d\beta \quad (15.4.1)$$

where x and y are the spatial coordinates, u and v are the spatial frequencies, and the asterisk denotes complex conjugation.

Historically, the function introduced in (15.4.1) is called the "Wigner distribution" of f, while the "Wigner distribution" of an analytic image associated to f is referred to as the "Wigner-Ville distribution" of f (see Chapter 2). For simplicity, in this article only "Wigner-Ville distribution" appellation is used, as defined in (15.4.1).

To the 2D image $f(x,y)$ corresponds a Wigner-Ville distribution $W_f(x,y,u,v)$, that provides a spatial/spatial-frequency representation of the image. To any point (x,y) of the image is assigned a 2D spatial-frequency spectrum. Applying Parseval's formula, the WVD can be written in terms of Fourier transform of $f(x,y)$, $F(u,v)$, as

$$W_f(x,y,u,v) = \int_{\mathbb{R}^2} F\left(u+\tfrac{\eta}{2}, v+\tfrac{\xi}{2}\right) F^*\left(u-\tfrac{\eta}{2}, v-\tfrac{\xi}{2}\right) e^{j2\pi(\eta x+\xi y)}\, d\eta\, d\xi\,. \quad (15.4.2)$$

The image can be reconstructed up to a sign ambiguity from its WVD:

$$f(x,y)f^*(0,0) = \int_{\mathbb{R}^2} W_f\left(\frac{x}{2}, \frac{y}{2}, u, v\right) e^{j2\pi(\alpha u+\beta v)} dx\, dy\,. \quad (15.4.3)$$

Among the properties of 2D Wigner-Ville distribution, the most important for image processing applications is that it is always a real-valued function and, at the same time, contains the phase information. The 2D Wigner-Ville distribution has many interesting properties related to translation, modulation, scaling, and convolution. For an in-depth description the reader is referred to [4]. In the sequel, the properties relevant for image analysis tasks are reviewed.

The Wigner-Ville distribution localizes the spatial/spatial-frequency structures of $f(x,y)$. For instance, if the signal energy is concentrated in space around (x_0,y_0) and in frequency around (u_0,v_0), then the energy of $W_f(x,y,u,v)$ is centered at (x_0,y_0,u_0,v_0) and has the same spatial and frequency spread as $f(x,y)$. Actually, it can be shown that the spatial-frequency support of $W_f(x,y,u,v)$, is included in the support interval of $F(u,v)$, for all (x,y), and that the spatial support of $W_f(x,y,u,v)$ is included in the support interval of $f(x,y)$ for all (u,v).

The local power at a fixed position (x_0,y_0) is equal to the marginal integral of the WVD over the spatial-frequency domain at the considered spatial position, while the spectral energy density at the spectral point (u_0,v_0) is equal to the marginal

integral of the WVD over the spatial domain at the considered spectral point:

$$|f(x_0, y_0)|^2 = \int_{\mathbb{R}^2} W_f(x_0, y_0, u, v)\, du dv \tag{15.4.4}$$

$$|F(u_0, v_0)|^2 = \int_{\mathbb{R}^2} W_f(x, y, u_0, v_0)\, dx\, dy\,. \tag{15.4.5}$$

As the result of these properties, the Wigner-Ville distribution is often thought as the image energy distribution in the joint spatial/spatial-frequency domain, which strongly encourages the use of 2D WVD in image analysis applications where the spatial/spatial-frequency features of images are of interest.

The main problem of the WVD is the so-called cross-term interference (see Article 4.2 for the one-dimensional case). Due to its bilinearity, the Wigner-Ville distribution of the sum of two images f_1 and f_2 introduces an interference term, usually regarded as undesirable artifacts in image analysis applications:

$$W_{f_1+f_2}(x, y, u, v) = W_{f_1}(x, y, u, v) + W_{f_2}(x, y, u, v) + 2\text{Real}\left[W_{f_1,f_2}(x, y, u, v)\right] \tag{15.4.6}$$

where the interference term W_{f_1,f_2} is the cross-WVD of f_1 and f_2:

$$W_{f_1,f_2}(x, y, u, v) = \int_{\mathbb{R}^2} f_1\left(x + \tfrac{\alpha}{2}, y + \tfrac{\beta}{2}\right) f_2^*\left(x - \tfrac{\alpha}{2}, y - \tfrac{\beta}{2}\right) e^{-j2\pi(\alpha u + \beta v)} d\alpha\, d\beta. \tag{15.4.7}$$

The interference terms are highly oscillating and their magnitude can be two times larger than that of the auto-terms. The artifacts introduced by the interference can be classified into two categories [2]: (a) interference terms due to the interaction between the conjugate symmetric spectral components of real images, and (b) interference terms due to the interaction between any two spectral components of the image (real or complex) at different spatial frequencies. Obviously, the WVD of any real image suffers from the first type of artifacts, due to its conjugate symmetric spectrum. Moreover, if the real image is multicomponent, which is usually the case in most of the applications, both kinds of artifacts are present, making the spatial/spatial-frequency representation difficult to interpret. The following subsection presents a concept that may be used to eliminate the first type of interferences and to reduce the second.

15.4.2.1 Analytic Image

If the spectrum of an image contains only positive (or only negative) frequency components, there is obviously no interference between the positive and negative frequencies in its WVD. The analytic image has such a spectral property, therefore a cleaner spatial/spatial-frequency representation of a real image, $f(x, y)$, may be obtained by computing the WVD of an analytic image, $z(x, y)$ associated to it.

The analytic image is the extension of 1D analytic signal to 2D. Contrary to the 1D analytic signal, the analytic image is not unique. For a reliable spatial/spatial-frequency representation of the real image, the analytic image should be chosen so that:
(a) the useful information from the 2D WVD of the real signal are found in the 2D WVD of the analytic image, and
(b) the 2D WVD of the analytic image minimize the interference effect.

As in 1D case where the analytic image is formed by suppressing the negative frequencies, the 2D analytic image may be constructed by setting to zero one half-plane of the spatial-frequency domain. By introducing a reference direction $(\cos\theta, \sin\theta)$, a pair of frequency (u, v) is called positive if $u\cos\theta + v\sin\theta > 0$, and is called negative if $u\cos\theta + v\sin\theta < 0$. According to this convention, the analytic signal is straightforward defined in the frequency domain.

The Fourier transform, $Z^{(\theta)}(u, v)$, of the analytic image with respect to direction $(\cos\theta, \sin\theta)$, $z^{(\theta)}(x, y)$, corresponding to the real image f is [7]:

$$Z^{(\theta)}(u, v) = F(u, v)\,(1 + \text{sign}(u\cos\theta + v\sin\theta))\,. \tag{15.4.8}$$

The choice of θ depends on the knowledge on the spectral characteristics of the real image. Unfortunately, $W_{z^{(\theta)}}$ contains interference cross-terms between the spectral components of the two quadrants in the positive frequency half-plane.

To eliminate these interference cross-terms a second definition of the analytic image seems appropriate, as a complex signal with single-quadrant spectrum [8]. To avoid any loss of information, two analytic images are associated to the real image.

The Fourier transforms $Z_1(u, v)$ and $Z_2(u, v)$, of the analytic images with the single-quadrant spectra, $z_1(x, y)$ and $z_2(x, y)$, corresponding to a real image f are defined as:

$$Z_1(u, v) = F(u, v)\,(1 + \text{sign}(u))\,(1 + \text{sign}(v)) \tag{15.4.9}$$

$$Z_2(u, v) = F(u, v)\,(1 + \text{sign}(u))\,(1 - \text{sign}(v))\,. \tag{15.4.10}$$

Using z_1 and z_2, the interference between frequencies from different quadrants is eliminated for any spectral distribution of the real signal. W_{z_1} reflects the spatial/spatial-frequency distribution of f for $uv > 0$, and W_{z_2} for $uv < 0$.

15.4.2.2 Continuous 2D Windowed Wigner-Ville Distribution

In practical applications, the images are of finite support. Therefore it is appropriate to apply Wigner analysis to a windowed version of the infinite support images.

As in the 1D case [9], the 2D windowed Wigner-Ville distribution is constructed as follows: To compute the windowed WVD at a spatial location (x_0, y_0), a windowed version of f is considered, where the window $w(x, y)$ is centered on (x_0, y_0):

$$f_w(x, y; x_0, y_0) = f(x, y)w(x - x_0, y - y_0). \tag{15.4.11}$$

Then the windowed WVD at (x_0, y_0) is defined as the WVD of f_w at $x = x_0,\ y = y_0$:

$$W_f^{(w)}(x,y,u,v) = \int_{\mathbb{R}^2} w\left(\tfrac{\alpha}{2},\tfrac{\beta}{2}\right) w^*\left(-\tfrac{\alpha}{2},-\tfrac{\beta}{2}\right) \times f\left(x+\tfrac{\alpha}{2},y+\tfrac{\beta}{2}\right) f^*\left(x-\tfrac{\alpha}{2},y-\tfrac{\beta}{2}\right) e^{-j2\pi(\alpha u+\beta v)}\, d\alpha\, d\beta\,. \tag{15.4.12}$$

The 2D windowed WVD can be shown to represent the convolution with respect to the spatial frequencies of W_f and W_w:

$$W_f^{(w)}(x,y,u,v) = \int_{\mathbb{R}^2} W_f(x,y,\xi,\eta) W_w(x,y,u-\xi,v-\eta) d\xi d\eta\,. \tag{15.4.13}$$

The effect of the windowing is to smear the WVD representation in the frequency plane only, so that the frequency resolution is decreased but the spatial resolution is unchanged.

15.4.3 Discrete 2D Wigner-Ville Distribution

In this article the discrete version of (15.4.1) is not discussed, as it does not provide any relevant information from the image analysis viewpoint. In this respect, it is more interesting to take into account the discrete version of the windowed WVD (15.4.13). For the sake of simplicity, square shaped images and analysis windows are considered.

Let $f(m,n)$ be the discrete image obtained by sampling $f(x,y)$, adopting the convention that the sampling period is normalized to unity in both directions. The following notation is made:

$$K(m,n,r,s) = w(r,s)w^*(-r,-s)f(m+r,n+s)f^*(m-r,n-s)\,. \tag{15.4.14}$$

The discrete space equivalent of the windowed Wigner-Ville distribution (15.4.13) is given by:

$$W_f^{(w)}(m,n,u,v) = 4 \sum_{r=-L}^{L} \sum_{s=-L}^{L} K(m,n,r,s) e^{-j4\pi(ru+sv)} \tag{15.4.15}$$

where $w(r,s) = 0$, for $|r|, |s| > L \in \mathbb{Z}^+$.

The properties of the discrete space WVD are similar to the continuous WVD, except for the periodicity in the frequency variables, which is one-half the sampling frequency in each direction:

$$W_{f_w}(m,n,u,v) = W_{f_w}\left(m,n,u+\frac{k}{2},v+\frac{\ell}{2}\right),\ \forall m,n,k,\ell \in \mathbb{Z}\,. \tag{15.4.16}$$

Therefore, if $f(x,y)$ is a real image, it should be sampled at twice the Nyquist rate to avoid aliasing effects in $W_{f_w}(m,n,u,v)$.

There are three practical solutions to avoid aliasing, when the original continuous image is not available:
(a) to filter out the frequency regions that cause aliasing (lowpass filtering),
(b) to up-sample the image by a factor of 2 in both directions and then appropriately band-limit the image using an anti-aliasing filter [1], and
(c) to compute the Wigner-Ville distribution of one or more analytic images associated to the real image, that eliminate or at least reduce the aliasing [2, 10].

To have a discrete spatial-frequency representation of the Wigner-Ville distribution, the frequency variables should be sampled. The 2D discrete windowed Wigner-Ville distribution of an image $f(n,m)$, $(n,m) \in \mathbb{Z}^2$ is defined as:

$$W_f^{(w)}(m,n,u_p,v_q) = 4 \sum_{r=-L}^{L} \sum_{s=-L}^{L} K(m,n,r,s) W_4^{rp+sq} \qquad (15.4.17)$$

where $N = (2L+2)$, $W_4 = e^{-j4\pi/N}$, and the normalized spatial-frequency pair is $(u_p, v_q) = (p/N, q/N)$.

By making a periodic extension of the kernel $K(m,n,r,s)$, for fixed (m,n), (15.4.17) can be transformed to match the standard form of a 2D DFT, except that the twiddle factor is W_4 instead of W_2 (see [9] for additional details for 1D case; the 2D construction is a direct extension). Thus standard FFT algorithms can be used to calculate the discrete $W_f^{(w)}$. The additional power of two represents a scaling along the frequency axes, and can be neglected in the calculations.

As the real-scene images have rich frequency content, the interference cross-terms may mask the useful components contribution. Therefore a commonly used method to reduce the interference in image analysis applications is to smooth the 2D discrete windowed WVD in the spatial domain using a smoothing window $h(m,n)$. The price to pay is the spatial resolution reduction. The result is the so-called 2D discrete pseudo-Wigner distribution (PWD) [1], which is obtained from the windowed WVD:

$$PW_f(m,n,u_p,v_q) = \sum_{k=-M}^{M} \sum_{\ell=-M}^{M} h(k,\ell) W_f^{(w)}(m+k,n+\ell,u_p,v_q). \quad (15.4.18)$$

In the case of a symmetric frequency window, i.e. $w(r,s) = w(-r,-s)$, the PWD can be defined by:

$$PW_f(m,n,u_p,v_q) = 4 \sum_{r=-L}^{L} \sum_{s=-L}^{L} |w(r,s)|^2 W_4^{rp+sq} \qquad (15.4.19)$$

$$\times \sum_{k=-M}^{M} \sum_{\ell=-M}^{M} h(k,\ell) f(m+k+r, n+\ell+s) f^*(m+k-r, n+\ell-s).$$

A very important aspect to take into account when using PWD is the choice of the two windows, $w(r,s)$ and $h(k,\ell)$. The size of the first window, $w(r,s)$, is dictated

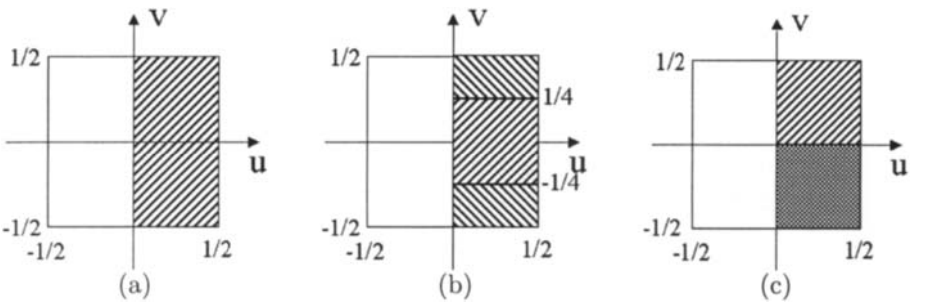


Fig. 15.4.1: Examples of discrete analytic images. The spatial frequency is normalized with respect to the sampling frequency. If the case, the frequency supports of different images are marked by different motives. (a) Analytic image used in [2], which is the discrete version of (15.4.8) for $\theta = 0$. (b) The discrete analytic image pair proposed in [10]. (c) Single-quadrant discrete analytic image pair, which are the discrete version of (15.4.9) and (15.4.10).

by the resolution required in the spatial-frequency domain. The spectral shape of the window should be an approximation of the delta function that optimizes the compromise between the central lobe width and the side lobes height. A window that complies with these demands is the 2D extension of Kaiser window, which was used in [1].

The role of the second window, $h(k, \ell)$, is to allow spatial averaging. Its size determines the degree of smoothing. The larger the size is, lower the spatial resolution becomes. The common choice for this window is the rectangular window.

15.4.3.1 Choice of Analytic Image in Computing Discrete Wigner-Ville Distributions

In the discrete case, there is an additional specific requirement when choosing the analytic image: the elimination of the aliasing effect. Taking into account that all the information of the real image must by preserved in the analytic image, only one analytic image cannot fulfill both requirements. Therefore, either one analytic image is used and some aliasing is allowed, or more analytic images are employed which satisfy two restrictions: (a) the real image can be perfectly reconstructed from the analytic images, and (b) each analytic image is alias free with respect to WVD.

The discrete version of the half-plane analytic image (15.4.8) was used for texture segmentation in [2]. While preserving all the information of the real image, it fails to produce a substantial reduction of the aliasing, as it does not prevent frequency fold-over along one direction (Fig. 15.4.1(a)).

To avoid aliasing, a solution is to use two analytic images, obtained by splitting the region of support of the half-plane analytic image into two equal area subregions. This is the procedure presented in [10], where one analytic image contains the low positive frequencies and the other the high positive frequencies (Fig. 15.4.1(b)).

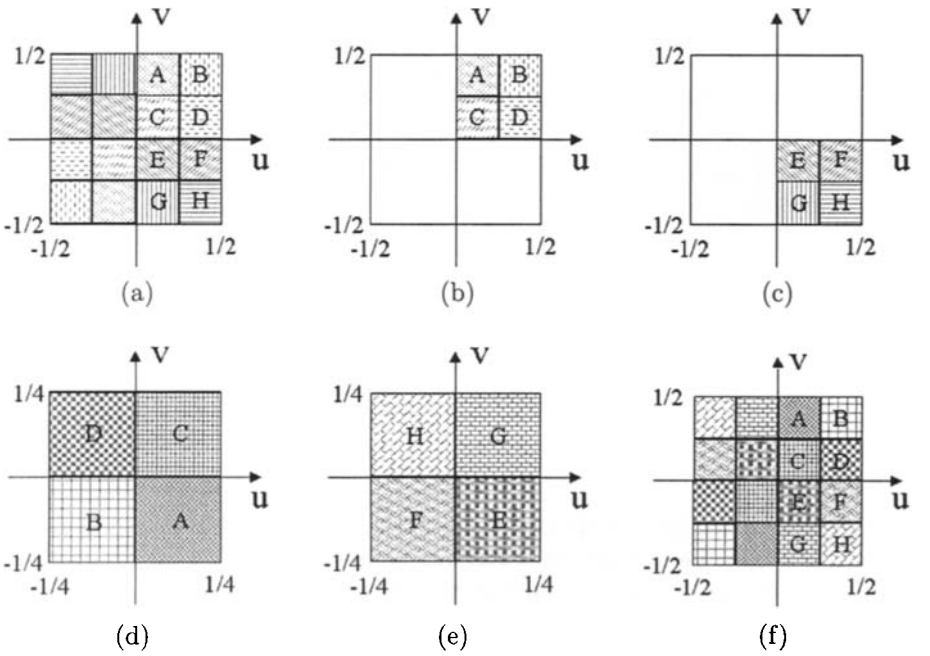


Fig. 15.4.2: Full-domain WVD computation using a single-quadrant analytic image pair. (a) Spectrum of the real image. (b) Spectrum of the upper-right quadrant analytic image (see (15.4.9)). (c) Spectrum of the lower-right quadrant analytic image (see (15.4.10)). (d) Spatial-frequency support of WVD of (b). (e) Spatial-frequency support of WVD of (c). (f) Spatial-frequency support of the full-domain WVD obtained from (d) and (e).

Although this method requires the computation of two WVD, no aliasing artifacts appear and the WVD of the analytic images can be combined to produce a spatial/spatial-frequency representation of the real image having the same frequency resolution and support as the original real image. This approach was successfully applied in texture analysis and segmentation in [3].

In general, the energy of natural images is concentrated at the low frequencies, so most probable the low frequency analytic image contains most of the energy of the real image, and, therefore, its WVD would exhibit strong cross-term interferences. A more equilibrated energy sharing is done when discrete single-quadrant analytic images are used (Fig. 15.4.1(c)). They are the discrete version of (15.4.9) and (15.4.10). As the positive frequency domain is split along v-axis, the inter-quadrant cross-terms are eliminated.

As in [10], a full-domain PWD of the real image $f(m,n)$, $FPW_f(m,n,u_p,v_q)$, can be constructed from $PW_{z_1}(m,n,u_p,v_q)$ and $PW_{z_2}(m,n,u_p,v_q)$. In the spatial-frequency domain, the full-domain PWD is, by definition, of periodicity 1 and symmetric with respect to the origin, as the WVD of a real image. It is completely

specified by:

$$FPW_f(m,n,u_p,v_q) = \begin{cases} PW_{z_1}(m,n,u_p,v_q), & 0 \le u_p < \frac{1}{2},\ 0 < v_q < \frac{1}{2} \\ PW_{z_2}(m,n,u_p,v_q), & 0 \le u_p < \frac{1}{2},\ 0 > v_q \ge -\frac{1}{2} \end{cases} \tag{15.4.20}$$

$$FPW_f(m,n,u_p,0) = PW_{z_1}(m,n,u_p,0) + PW_{z_2}(m,n,u_p,0), \quad 0 \le u_p < \frac{1}{2} \tag{15.4.21}$$

$$FPW_f(m,n,u_p,v_q) = FPW_f(m,n,-u_p,-v_q), \quad 0 > u_p, v_q \ge -\frac{1}{2} \tag{15.4.22}$$

$$FPW_f(m,n,u_p+k,v_q+l) = FPW_f(m,n,u_p,v_q), \quad \forall k,l,p,q \in \mathbb{Z}. \tag{15.4.23}$$

Fig. 15.4.2 illustrates the construction of the full-domain PWD from the PWD of the single-quadrant analytic images. The same shading indicates identical regions. Letters are used to follow the mapping of frequency regions of the real image; for instance, the region labeled A in (f) represents the mapping of the region A in the real image spectrum (a) on the spatial-frequency domain of the full-domain PWD.

In conclusion, the full-domain PWD provides information un-altered by aliasing artifacts on the spatial/spatial-frequency distribution of the real image over the entire frequency spectrum. A potential drawback of these approaches is that the additional sharp filtering boundaries may introduce ringing effects.

15.4.4 An Image Dissimilarity Measure based on the 2D Wigner-Ville Distribution

Structured distortions affecting an image, which are more annoying than the unstructured distortions, are usually highly concentrated in the spatial/spatial-frequency domain. Between two distortions with the same energy, i.e. same signal-to-noise ratio (SNR), the more disturbing is the one having a peaked energy distribution in spatial/spatial-frequency plane.

In terms of the effect on the WVD, the noise added to an image influences not only the coefficients in the positions where the noise has non-zero WVD coefficients, but also induces cross-interference terms (see (15.4.6)). Stronger the noise WVD coefficients are, more important the differences between the noisy image WVD and original image WVD become.

WVD-based SNR of a distorted version $g(m,n)$ of the original discrete image $f(m,n)$ is defined as:

$$\mathrm{SNR_W} = 10\log_{10}\frac{\sum_m\sum_n \max_{p,q}\left(|FPW_f(m,n,u_p,v_q)|\right)}{\sum_m\sum_n \max_{p,q}\left(|FPW_f(m,n,u_p,v_q) - FPW_g(m,n,u_p,v_q)|\right)}. \tag{15.4.24}$$

For each position (m,n) the highest energy WVD component is retained, as if the contribution of the other components are masked by it. Of course, the masking mechanisms are much more complex, but this coarse approximation leads to results which are more correlated to the HVS perception than SNR.

The use of maximum difference power spectrum as a nonlinearity transformation is motivated and inspired by some findings on nonlinearities in the HVS. Similar transformations have been successfully used to model intra-cortical inhibition in the primary visual cortex in an HVS-based method for texture discrimination [11].

Let η_1 and η_2 be two degradations having the same energy. The first, η_1 is additive white Gaussian noise, and the second, η_2 is an interference pattern. While the energy of the noise is evenly spread in the spatial/spatial-frequency plane, the energy of the structured degradation is concentrated in the frequency band of the interference. Thus the WVD of η_2 contains terms which have absolute values larger than any term of WVD of η_1, as the two degradations have the same energy. These peak terms induce larger local differences between WVD of $g_2 = f + \eta_2$ and WVD of f, which are captured by 'max' operation in the denominator of (15.4.24) and lead to a smaller $\mathrm{SNR_W}$ for g_2.

15.4.4.1 Results and Discussion

To show the interest of the proposed image distortion measure as compared to the signal-to-noise ratio (SNR), three types of degradation are considered: white noise, interference pattern, JPEG coding (Fig. 15.4.3). They induce nearly the same SNR, whereas visual comparison clearly reveals different perceptual quality.

The WVD-based distortion measure is proven to correlate with subjective quality evaluation done by five non-expert observers, which prefer the white noise distorted image to the interference perturbed image and to the JPEG coded image. The reason is that for random degradation the noise has the same effect in the entire spatial-frequency plane. Therefore, the maximum spectral difference at almost any spatial position is lower than the just noticeable perceptual difference. On the other hand, when the distortion is localized (as interference patterns or distortion induced by JPEG coding), the maximum spectral difference corresponding to an important proportion of the pixels has a significant value, much larger than the just noticeable perceptual difference.

15.4.5 Summary and Conclusions

This article considers the 2D WVD in the framework of image analysis. The advantages and drawbacks of this spatial/spatial-frequency analysis tool are recalled in the light of some pioneer and recent works in this field.

The usefulness of the WVD in image analysis is demonstrated by considering a particular application, namely distortion analysis. In this respect, a new image distortion measure is defined. It is calculated using the spatial/spatial-frequency representation of images obtained using the 2D WVD. The efficiency of this measure is validated through experiments and informal visual quality assessment tests. It is shown that this measure represents a promising tool for objective measure of image quality, although the masking mechanisms are neglected. To improve the reliability and the performance of the proposed method, a refinement to include a masking model is imperatively needed.

(a) f (b) g_1 (SNR = 23.70, SNR_W = 21.70)

(c) g_2 (SNR = 23.74, SNR_W = 17.66) (d) g_3 (SNR = 23.70, SNR_W = 14.07)

Fig. 15.4.3: Distorted versions of 256 × 256 pixel Parrot image, f: g_1 is obtained by adding white Gaussian noise on f; g_2 is a JPEG reconstruction of f, with a quality factor of 88; g_3 is the result of imposing a grid-like interference over f. The SNR and SNR_W values are given in dB.

It can be concluded that, taking into consideration some basic, well-established knowledge on the HVS (the joint spatial/spatial-frequency representation, and non-linear inhibition models), one can develop a simple image distortion measure correlated with the perceptual evaluation.

Further details on the use of the Wigner-Ville distribution for image analysis and processing can be found in [12–14].

References

[1] T. R. Reed and H. Wechsler, "Segmentation of textured images and Gestalt organization using spatial/spatial-frequency representations," *IEEE Trans. Pattern Analysis &*

Machine Intelligence, vol. 12, pp. 1–12, January 1990.

[2] Y. M. Zhu, R. Goutte, and M. Amiel, "On the use of two-dimensional Wigner-Ville distribution for texture segmentation," *Signal Processing*, vol. 30, pp. 329–353, February 1993.

[3] G. Cristóbal and J. Hormigo, "Texture segmentation through eigen-analysis of the Pseudo-Wigner distribution," *Pattern Recognition Letters*, vol. 20, pp. 337–345, 1999.

[4] L. Jacobson and H. Wechsler, "Joint spatial/spatial-frequency representation," *Signal Processing*, vol. 14, pp. 37–68, 1988.

[5] C. J. van den Branden Lambrecht, ed., *Signal Processing*, vol. 70, no. 3. EURASIP/Elsevier, 30 November 1998. Special issue on image quality assessment.

[6] M. A. García-Pérez and V. Sierra-Vázquez, "Visual processing in the joint spatial/spatial-frequency domain," in *Vision Models for Target Detection and Recognition: In Memory of Arthur Menendez* (E. Peli, ed.), ch. 2, pp. 16–62, World Scientific Publishing, 1995.

[7] T. Bülow and G. Sommer, "A novel approach to the 2D analytic signal," in *Proc. 8th Internat. Conf. on Computer Analysis of Images & Patterns (CAIP'99)*, Ljubljana, Slovenia, 1–3 September 1999 (F. Solina and A. Leonardis, eds.), no. 1689 in LECTURE NOTES IN COMPUTER SCIENCE, pp. 25–32, Springer, 1999.

[8] S. L. Hahn, "Multidimensional complex signals with single-orthant spectra," *Proc. IEEE*, vol. 80, pp. 1287–1300, August 1992.

[9] B. Boashash and P. J. Black, "An efficient real-time implementation of the Wigner-Ville distribution," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 35, pp. 1611–1618, November 1987.

[10] J. Hormigo and G. Cristóbal, "High resolution spectral analysis of images using the pseudo-Wigner distribution," *IEEE Trans. Signal Processing*, vol. 46, pp. 1757–1763, June 1998.

[11] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *J. Optical Soc. of America A*, vol. 7, pp. 923–932, 1990.

[12] H. O. Bartelt, K. H. Brenner, and A. W. Lohmann, "The Wigner distribution function and its optical production," *Optics Communications*, vol. 32, pp. 32–38, January 1980.

[13] G. Cristóbal, J. Bescós, and J. Santamaría, "Image analysis through the Wigner distribution function," *Applied Optics*, vol. 28, pp. 262–271, January 1989.

[14] H. Suzuki and F. Kobayashi, "A method of two-dimensional spectral analysis using the Wigner distribution," *Electronics & Communications in Japan, Part III: Fundamental Electronic Science*, vol. 75, pp. 1006–1013, January 1992.

15.5 TIME-FREQUENCY DETECTION OF EEG ABNORMALITIES [0]

This article presents an example of time-frequency methodology used for the detection of seizures in recorded EEG signals. The techniques used are adapted to the case of newborn EEGs, which exhibit some well defined features in the time-frequency domain that allow an efficient discrimination between abnormal EEGs and background.

15.5.1 EEG Abnormalities and Time-Frequency Processing

Neonatal seizures are usually the first signs of neurological abnormalities and can lead to permanent brain damage or even fatalities if not detected at the early stages. There are a number of disturbances underlying the seizure rather than a single identifiable cause making the identification process difficult. The problem in newborn is harder than in adults because the more obvious clinical symptoms such as muscle spasms, sporadic eye movements and drooling are often difficult to detect [1]. For this reason, Electroencephalogram (EEG) is practically the only tool available in seizure detection and characterization in newborns. Three major approaches have been used to detect seizures in newborns based on the assumption that the EEG signals are stationary or at least locally stationary [1,2]. However, a detailed examination of these signals shows that EEG signals exhibit significant non-stationary and multi-component features [see Fig. 15.5.1(a)]. making these three methods essentially invalid and at best only an approximation. This explains the relatively poor performance of these methods [2]. The non-stationarity and multicomponent nature of the EEG signal suggested the use of time-frequency (TF) signal processing to analyze and characterize the different newborn EEG patterns for developing a time-frequency seizure detection and classification [1,3].

15.5.2 EEG Seizures in Newborns

A seizure is defined to occur when there is an excessive synchronous discharge of neurons within the central nervous system. Its manifestation in the EEG, known as electrographic seizure, consists of a paroxysmal events which are trains of rhythmic repetitive sharp waves that emerge more or less abruptly from the ongoing background activities and have a distinct beginning and end. They may start with low voltages that increase usually as the discharge progresses. They often contain subharmonics and may have polyphasic contours or be sinusoidal. These discharges pattern can be divided into four categories: focal spike and sharp waves (> 2Hz), local low frequency discharges (around 1Hz), focal rhythmic discharge (0.5 Hz - 15 Hz), and multifocal patterns (EEG discharge originating from two or more loci) [1].

---

[0]Authors: **Boualem Boashash** and **Mostefa Mesbah**, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Australia (b.boashash@qut.edu.au, m.mesbah@qut.edu.au), and **Paul Colditz**, Perinatal Research Centre, Royal Women's Hospital, Brisbane, Australia (p.colditz@uq.edu.au). Reviewers: V. Chandran and A. Belouchrani.

The duration of rhythmic discharges is highly variable, from as short as 1 second to as long as 30 minutes. This fact contributed to the disagreement between the researchers about what constitutes a seizure. In order to consider an EEG discharge as a seizure, some researchers require that it must last at least 10 seconds, others require a minimum of 20 seconds, while a third group does not specify a time limit.

Seizure patterns are occasionally corrupted by artifacts and some abnormal background patterns such as burst suppression (BS). The most noticeable artifacts are the ones caused by the heartbeat (ECG), the eye movement (EOG) and head and body movements (EMG) [4].

15.5.3 Data Acquisition

Electrical signals produced in the brain can be monitored in a non-invasive manner by measuring variations in potential on the scalp. This EEG measurement is achieved by strategically placing several 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 created by measuring the voltage differences between neighboring electrodes. Five channels of EEG have been recorded in each session using the 10-20 International System of Electrode Placement. The EEG data has been recorded using a sampling frequency of 256 Hz. For artifact detection, three auxiliary signals representing electro-oculogram (EOG), electrocardiogram (ECG), and respiration are also recorded. Data used has been collected at the Royal Women's Hospital Perinatal Intensive Care Unit in Brisbane, Australia. The EEG signals containing seizures were obtained from two different newborn babies that have been clinically identified to have seizures. The gestational ages of the babies were 35 weeks and 40 weeks and 3 days. The recording lasted 137 minutes and 23 minutes respectively.

15.5.4 Selection of a Time-Frequency Distribution

The following characteristics were found to be typical of neonatal EEG signals [1]: non-stationary, occasionally multicomponent, low frequency signals in the range 0 to 5 Hz. These factors must be considered when selecting an optimal time-frequency distribution (TFD), as each TFD is more suited to representing signals with particular characteristics (see Chapter 3).

Since neonatal EEG signals are non-stationary and occasionally multicomponent, a desirable time-frequency distribution should have a good spectral resolution and reduced cross-terms. The performance and characteristics of several TFDs were compared to find an optimal representation of real neonatal EEG data in the TF domain. The scope of this comparison study has encompassed seven TFDs [1]. Each TFD has been applied to epochs of real neonatal EEGs for various data window lengths and individual TFD parameter values. The performances were compared visually and using an objective quantitative measure criterion (see Article 7.4). Based on this criterion, the B-distribution (BD) with the smoothing parameter β equals to 0.01 has been selected as the most suitable representation of

the EEG signals in the TF domain.

The B-distribution is defined in terms of its time-lag kernel (see chapters 2 and 3) and may be expressed as

$$\rho_z(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{|\tau|}{\cosh^2(t)}\right)^{\beta} z(t+\tau/2)z^*(t-\tau/2)e^{-j2\pi f\tau}\,du\,d\tau\,. \quad (15.5.1)$$

The parameter β $(0 < \beta \leq 1)$ controls the sharpness of the cut-off of the two-dimensional filter in the ambiguity domain. Hence, the EEG signals were represented in time frequency using the B-distribution with a smoothing parameter of 0.01, a window length of 127 samples, and a time resolution of 5 samples. The data has been resampled to 20 Hz for better representation of low frequency regions. The time-frequency analysis was performed using the commercial *TFSA* 5.2 MATLAB^{TM} toolbox (http://www.sprc.qut.edu.au//tfsa/~index.html).

15.5.5 EEG Pattern Analysis

The visual analysis of the time-frequency EEG data led to divide the time-frequency EEG patterns into two classes: seizure and background. The seizure patterns can be characterized in the time-frequency domain by a main ridge (component) as either a linear FM law or a piecewise linear FM while the background patterns exhibit a low frequency burst activities or irregular activities with no clearly defined patterns [1]. These observations correlate well with clinical information related to EEGs [5]. Representative TF representations of each of the subclasses are detailed below.

15.5.6 Analysis of Time-Frequency Seizure Patterns

15.5.6.1 Linear FM (LFM) Patterns

The EEG seizures analyzed in the TF domain that can be approximated by linear FMs with either fixed or time-varying amplitudes can be classified into the following sub-classes:

LFM Patterns with a Quasi-Constant Frequency: Fig. 15.5.1(b) shows a seizure that has a linear FM behavior with an almost constant frequency. The amplitude of the time-frequency seizure pattern increases at the onset and decreases toward the end. A major advantage of the TF representation is that we can easily distinguish the seizure from other phenomena such as burst activities as long as they occupy different TF regions. These unwanted signals can be removed from the EEG signal using a well designed TF filter without affecting much the seizure signal.

LFM Patterns with a Decreasing Frequency: Fig. 15.5.1(c) of this class differs from the one above by the fact that its frequency decreases with time [5] . By looking at the TF behavior of the seizure, we can easily deduce the precise non-stationary character of the seizure. The classical detection methods based on the stationarity assumption will most likely miss these patterns.

15.5.6.2 Piecewise LFM Patterns

Most of the patterns analyzed so far can be approximated to a good degree of accuracy by piecewise linear FM as shown in Fig. 15.5.1(d). These types of seizures usually comprises the different stages of the seizure [4].

15.5.6.3 EEG Background Patterns

By background, we mean any signal that is not classified as seizure. Two distinct patterns have been noticed: Burst of activity and an irregular activity with no clear pattern.

Burst of Activity: Fig. 15.5.1(e) is an example of this class characterized by a burst of activity. These are a short period signals with a high energy lasting for few seconds and usually occurring at frequencies below 4 Hz. These features are characteristic of *burst suppression.* Burst suppression is defined as burst of high voltage activity lasting 1-10 seconds and composed of various patterns (delta (0 - 4 Hz) and theta (4 - 8 Hz) with superimposed and intermixed spikes, sharp waves, and faster activity) followed by a marked background attenuation [5].

Activity Lacking a Specific Pattern: Fig. 15.5.1(f) is an example of an EEG epoch lacking a well-defined and consistent pattern. These type of activities are not constrained within the low frequency bands characterizing the EEG seizure.

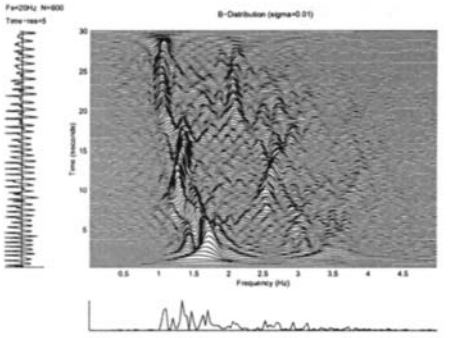
This time-frequency analysis indicates that a linear or piecewise linear instantaneous frequency (IF), obtained by taking the peak of the main component of a TFD, can be used as a critical feature of EEG seizure characteristics. These findings suggested to propose a TF-based seizure detector. This detector, called TF matched detector, performs a two dimensional correlation between the EEG signal and a reference template selected as a model in TF domain of the EEG seizure.

15.5.7 Time-Frequency Matched Detector

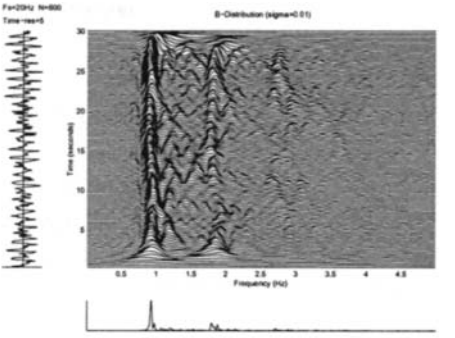
The matched filter is the simplest approach for constructing detectors and classifiers. It essentially reduces to a correlator receiver whose output is compared to a threshold. The threshold is chosen such that the probability of a false alarm is kept constant. The correlator receiver is implemented in time domain as a one-dimensional correlation between the received noisy signal $x(t)$ and a reference signal $s(t)$ or using the corresponding spectral representations. To extend this detector to handle nonstationary signals, the one-dimensional correlation is replaced by a two-dimensional correlation involving the TFD $\rho(t,f)$ of $x(t)$ and $s(t)$. The resulting test statistic is given by:

$$T(x) = \iint \rho_x(t,f)\rho_s(t,f)\,dt\,df \tag{15.5.2}$$

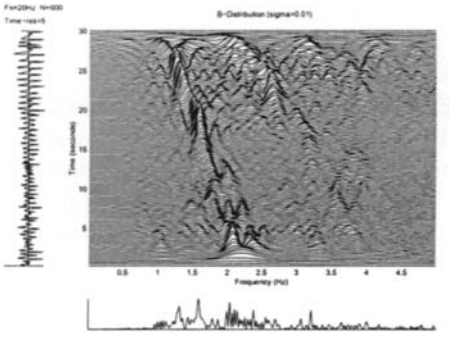
This type of detector has been implemented using different quadratic time-frequency distributions such as the spectrogram [6], the Wigner-Ville, and cross Wigner-Ville



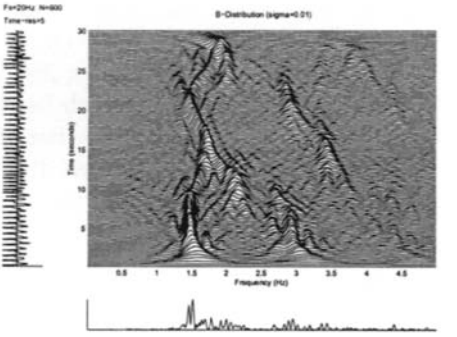
(a) Newborn seizure.



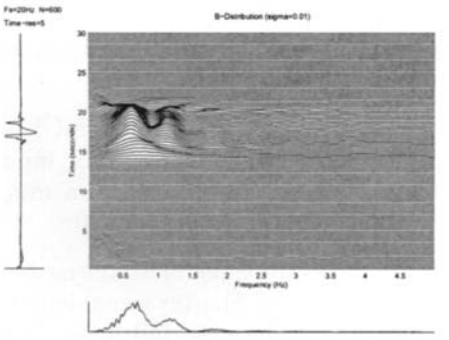
(b) Seizure exhibiting LFM behavior with constant frequency.



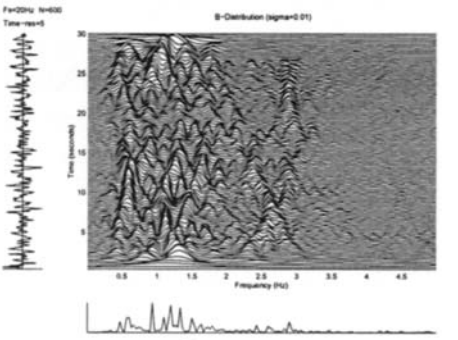
(c) Seizure exhibiting LFM behavior with decreasing frequency.



(d) Seizure exhibiting piecewise LFM behavior.



(e) Burst of activity.



(f) EEG lacking a specific pattern.

Fig. 15.5.1: B-distributions of EEG signals.

distributions [7] and the auto- and cross-ambiguity functions [8]. Using Moyal's formula, we get:

$$\iint \rho_x(t,f)\rho_s(t,f)\,dt\,df = \left|\int x(t)s^*(t)dt\right|^2 \tag{15.5.3}$$

This equality is only valid when the kernel filter is unimodular; that is its absolute value is equal to one all over the ambiguity domain. This is the case, for example, for the Wigner-Ville distribution and Rihaczek distribution [9]. Equation (15.5.3) is an alternative interpretation of the correlator receiver in terms of a correlation of the TFDs. Even though the B-distribution does not verify exactly Eq. (15.5.3), it has been used as the basis for the TF matched detector because of its superiority over the other TFDs in representing EEG signals as discussed in Section 15.5.4 (see also Articles 5.7 and 7.4).

For the case of a deterministic signal in additive noise (even white Gaussian noise) the TF-based correlator is suboptimal due to the nonlinearity of the quadratic TFDs which 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. Section 15.5.6 indicated that the EEG seizure could be characterized by a linear or a piecewise linear FM. To construct a TF-based matched detector, a representative TFD of a linear or piecewise linear FM, $\rho_{ref}(t,f)$, is selected to serve as template (reference). The correlator statistic $T(x)$ used is the two dimensional cross-correlation between the EEG signal TFD and the reference signal TFD, i.e.:

$$\iint \rho_{ref}(t,f)\rho_z^*(t,f)\,dt\,df \tag{15.5.4}$$

where z is the analytic signal corresponding to the EEG signal under consideration.

15.5.7.1 Implementation of the Time-Frequency Matched Detector

The implementation of the TF matched detector and a description of its main components are described below. More details may be found in [10].

Preprocessing: This stage includes artifact (such as ECG, EOG, and EMG) removal, noise filtering, and resampling the signal to comply with detector input specifications. A low pass filter along with an artifact removal algorithm using adaptive signal processing techniques were implemented for this purpose [4].

Signal Restructuring: EEG is segmented into an array of signals of fixed length (2 minutes) 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 minutes duration, each block is stored as a row of the newly formed array of signals. A protocol of 50% overlap of each block was adopted.

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.

Cross-correlation: The cross-correlation between the TF array of the EEG signal and the template (mask) is obtained using the two-dimensional cross-correlation function given by Eq. (15.5.4). The most crucial process is the choice of the template, in this case (see Section 15.5.6) the TFD of a linear FM or a piecewise linear FM. The time duration of the FM signal is set to 20 seconds as discussed below. To find the optimum slopes of the FM signal IF, that is the ones that corresponds to the best detection rate, a testing stage is necessary [10]. A similar testing stage is also required to select an optimum threshold that realizes a good compromise between the rate of good detections and the rate of false alarms.

Amplitude and Length Criteria: Ideally there will be one peak value in the output of the cross-correlation array, with its output amplitude determining the presence or absence of seizure. This proved to be unreliable, and it was decided to perform a search of sequential series of values over the amplitude threshold defined earlier. This proved to be successful, and a minimum ridge length of 20 seconds over the amplitude threshold was classified as a seizure. The 20-second length adopted is larger than the minimum 10-second length of EEG seizure adopted by many neurologists [5].

Map Seizure Decision to Real Time Location: This stage simply ties all of 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.

In order to validate and calibrate the detection algorithm, simulated data generated by the EEG model [2]. The model generates an EEG like signal characterized in time frequency by a linear IF with a random slope in the range of [-0.07 0]. These values were reported in [5]. The B-distribution was used to generate the reference template and the TFD of the simulated EEG. The signal used in the construction of the reference template is a linear FM. The average detection obtained was 99.1% while the false alarm rate was 0.4%. These results confirm the validity of the methodology since the template is well adapted to the EEG model.

15.5.8 Summary and Conclusions

The patterns obtained by a TF analysis of newborn EEG seizure signals show a linear FM or piecewise linear FM characteristic. This suggests a method of seizure detection and classification in the TF domain. A TF detector is proposed that involves cross-correlating the TFD of the EEG signal with a template. The design of the template takes into account the TF characteristics of the EEG seizure extracted in the TF domain. The performance of this time-frequency detector was tested on synthetic signals, corresponding to one specific type of seizure pattern (LFM). At

the time of publication, the methodology was being extended to deal with LFM patterns of varying slopes, and with piecewise linear FM patterns. The procedure will then allow classification within the selected sub-classes.

Another time-frequency approach to newborn EEG seizure detection is described in [11].

References

[1] B. Boashash and M. Mesbah, "A time-frequency approach for newborn seizure detection," *IEEE Engineering in Medicine & Biology Magazine*, vol. 20, pp. 54–64, September/October 2001.

[2] M. Mesbah and B. Boashash, "Performance comparison of seizure detection methods using EEG of newborns for implementation of a DSP subsystem," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'02)*, Orlando, FL, 13–17 May 2002. Paper no. 1932.

[3] B. Boashash, M. Mesbah, and P. Colditz, "Newborn EEG seizure pattern characterisation using time-frequency analysis," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, vol. 2, pp. 1041–1044, Salt Lake City, UT, 7–11 May 2001.

[4] P. Celka, B. Boashash, and P. Colditz, "Preprocessing and time-frequency analysis of newborn EEG seizures," *IEEE Engineering in Medicine & Biology Magazine*, vol. 20, pp. 30–39, September/October 2001.

[5] 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.), pp. 803–875, Baltimore, MD: Williams & Wilkins, 3rd ed., 1993.

[6] R. A. Altes, "Detection, estimation and classification with spectrograms," *J. Acoustical Soc. of America*, vol. 67, pp. 1232–1246, April 1980.

[7] B. Boashash and P. J. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1829–1841, November 1990.

[8] S. Pon Varma, A. Papandreou-Suppappola, and S. B. Suppappola, "Detecting faults in structures using time-frequency techniques," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, vol. 6, pp. 3593–3596, Salt Lake City, UT, 7–11 May 2001.

[9] P. Flandrin, "A time-frequency formulation of optimum detection," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 36, pp. 1377–1384, September 1988.

[10] B. Boashash and M. Mesbah, "Time-frequency methodology for newborn EEG seizure detection," in *Applications in Time-Frequency Signal Processing* (A. Papandreou-Suppappola, ed.), ch. 9, Boca Raton, FL: CRC Press, 2002.

[11] H. Hassanpour, M. Mesbah, and B. Boashash, "Comparative performance of time-frequency based newborn EEG seizure detection using spike signatures," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'03)*, vol. 2, pp. 389–392, Hong Kong, scheduled 6–10 April 2003.

15.6 TIME-FREQUENCY BASED MACHINE CONDITION MONITORING AND FAULT DIAGNOSIS [0]

15.6.1 Machine Condition Monitoring and Fault Diagnosis

Machine condition monitoring is the process of checking a machine for abnormal symptoms. Fault diagnosis, on the other hand, means deciding the nature and the cause of the fault by examining the symptoms [1]. The article aims at providing a methodology for potential users interested in implementing techniques pertaining to the area of machine condition monitoring using time-frequency analysis (TFA). It also provides three examples and some relevant references. Although this article focuses on one-dimensional time-domain signals, its methodology can be extended to images and image sequences.

15.6.1.1 Machine Condition Monitoring

In modern manufacturing, the quest for automation and flexibility has resulted in machines performing extremely complex processes. The performance of such processes highly depends on the trouble-free operation of all the components. When a fault occurs, it is critical to detect it, isolate the causes, and take appropriate maintenance action at an early stage. This helps prevent faults from developing into an eventual major machine failure and interrupting the production cycle. Consequently a number of techniques have been developed which monitor certain parameters within the machinery allowing its condition to be determined. These monitoring techniques have become known as machine condition monitoring.

The predictive maintenance through condition monitoring and diagnosis can significantly improve product quality, improve worker safety, and reduce the costs of maintenance. This is achieved by (1) allowing the early detection of potentially catastrophic faults which could be expensive to repair, and (2) allowing the implementation of condition based maintenance rather than periodic or failure based maintenance. In these cases, significant savings can be made by delaying scheduled maintenance until it is more convenient or necessary.

An efficient condition monitoring technique is capable of providing warning and predicting the faults at early stages by obtaining information about the machine in the form of primary data. Through signal processing (SP), the critical information from these data is captured and correlated to the condition of the machine. Effectiveness depends on matching the SP algorithms to the characteristics of the monitored signals.

Two types of condition monitoring and diagnosis systems are widely used: off-line and on-line. In an off-line (periodic) monitoring system, the monitored signal is measured at pre-selected time intervals. This approach is routinely used for fault diagnosis and trend analysis. In an on-line (permanent) monitoring system the

[0]Authors: **M. Mesbah, B. Boashash** and **J. Mathew**, Queensland University of Technology, CIEAM and SPRC, GPO Box 2434, Brisbane, Q 4001, Australia (m.mesbah@qut.edu.au, b.boashash@qut.edu.au, j.mathew@qut.edu.au). Reviewers: D. R. Iskander and V. Sucic.

signal is continuously measured and compared with a reference level. This type of system is intended to protect machines and/or operators by providing a warning about a possible malfunction of the machine and/or an imminent shutdown to prevent catastrophic failure.

Traditionally, human operators, using a combination of sight and sound, have performed machine condition monitoring. Recently, automatic techniques have been proposed to replace human operators. Some of these techniques rely on direct measurements while the majority depend on indirect measurements. Direct methods use sensing techniques that directly measure the extent of the deterioration, such as tool wear, in a machine. Indirect methods may rely on sensing different machine parameters such as forces, acoustic emission, temperature, vibration, current, voltage, torque, strain, and images of the tools in question. In techniques based on indirect measurement, features indicative of condition are extracted from these monitored signals and correlated to give a measure of the extent, the nature, and the location of the fault [2].

15.6.1.2 The Four Stages of Condition Monitoring and Diagnosis

In general, machine condition monitoring, as a pattern recognition problem, consists of four stages: data acquisition, feature extraction, feature selection, and decision-making. Data are acquired using transducers and normally recorded in either analog or digital form on magnetic tape or computer disk. (In simple systems it may be possible to perform the analysis in real-time).

A critical step of condition monitoring and diagnosis is feature extraction. It is generally not practical to automatically determine the machine condition using the collected raw signals and therefore some transformation or processing is required. This transformation usually involves as a first step mapping the original data from time-domain to another domain, such as the frequency or time-frequency domains, where the differences between the normal and abnormal behaviors are much more apparent. In this new domain, features that best describe the characteristics of the process condition are extracted. Feature extraction techniques include statistical methods, power spectral methods, and time-frequency methods as detailed in Section 15.6.2. In some cases, where the dimension of the feature space (or the number of features) is high, the dimension can be further reduced by retaining only the most valuable features and eliminating those that give little or no additional information. This dimension reduction processes is called feature selection.

The decision-making or classifier stage can be viewed as a process that automatically correlates the feature set, obtained from the previous stage, to the machine conditions [3]. It is usually done through supervised learning, where the operator instructs the computer of the possible patterns in the feature sets and relates them to the machine conditions. Sometimes it is difficult to generate data that reflects all uncertainties and differences within one class or group of faults in an experiment setting. In this case, an unsupervised learning strategy is used. Unsupervised learning is a task in which the number of classes is not known before classification and there

are no labeled training features available. The classifier itself should be capable of exploring the extracted features and deciding about the number of classes. Typical automatic or computer decision-making methods include pattern recognition, fuzzy logic, decision trees, and artificial neural networks. Ideally, there exists a one-to-one correlation between feature sets and machine conditions.

An alternative approach is to monitor the features and spot trends in them and thus predict failure. The decision to replace the faulty part is often taken when the feature crosses a given threshold [1].

15.6.1.3 Classical Signal Analysis Methods for Feature Extraction

Classical methods used for feature extraction can be classified into time domain and frequency domain.

Time Domain Methods: Probably the simplest approach proposed for fault detection in the time domain is through the measurement of the energy (mean square value) of the monitored signal. The method relies on the fact that as the machine's condition deteriorates, the vibration energy is expected to increase. Another approach is to use statistical parameters for fault detection. By treating the monitored signal as random variable, higher-order statistical moments, cumulants, and measures such as crest factor can also be used as features. Nonlinear signal based techniques have also been used for condition monitoring and fault diagnosis. In [4], for example, the correlation dimension was extracted from raw time-series acceleration data (collected from a rolling-element bearing) and used as a feature for detecting faults. Other methods such as level crossing, bandpass filtering, shock pulse, and autoregressive modeling are used (see for example [5]).

Frequency Domain Methods: The basic principle of spectral analysis is based on the fact that the spectrum of the monitored signal changes when faults occur. The nature and extent of the change depends of the nature of the fault and the machine being monitored. The condition of the machine is estimated through monitoring the change in the spectrum or a number of discriminating features extracted from the spectrum of the measured signal. These features are usually chosen as some specific frequency components that depend on the type of machine and the nature of the fault. They are compared to references established when the machine was known to work properly under similar conditions, and an appropriate decision is taken when the feature vector deviates from the reference by more than a predetermined threshold. In [6], the different changes in the vibration spectrum of rotating machines are surveyed and linked to different types of faults. Also, in [7, ch. 11] the most frequent failure modes are identified for the different machine-train components such as drives, steam turbines, gearboxes, and generators. For each component, a number of specific vibration frequencies are monitored for the diagnosis of incipient problems. These frequency-domain features, depending on the component and the nature of the failure, may include defect frequencies, the fundamental and harmonics of the rotational speed, the line frequency, the slip frequency, and

the tooth-mesh frequencies and the sidebands that surround them. Higher-order spectra such as the bispectrum and trispectrum are also used as a basis for condition monitoring. In [8], the bispectrum is used to analyze the acceleration signal obtained from a stamping process and to extract features related to defective parts.

15.6.1.4 Nonstationary Signals in Machines

Limitations of Classical Methods: Traditional time-domain and spectral analysis techniques have several shortcomings. For example, the Fourier transform is unable to accurately analyze and represent a signal that has non-periodic components such as a transient signal, as it is based on the assumption that the signal to be transformed is periodic. Another deficiency of the traditional spectral analysis is its inability to provide any information about the time dependency of the frequency content of non-stationary signals (see Article 1.1 for more details).

Motor current, for example, is well known to be a nonstationary signal whose properties vary with respect to the time-varying normal operating conditions of the motor, particularly with load. Also, for the case of rotating machines, the presence of certain frequency components within the spectrum has been shown to be an indication of a fault condition. However, since some of these frequencies depend on the rotational speed, it is not possible using spectral analysis to determine these frequencies when the bearing runs at variable rotational speed. Recent works have stressed the importance of machine monitoring during the transient states—such as start-up, shutdown, and acceleration periods—because some machine failures happen during these types of transition periods. Transient signals can be a good source of information about machine condition that is not available during steady states. Fourier transform based methods are known to be inadequate in representing this type of signals since the transient event can hardly be approximated by sines and cosines. For these reasons, Fourier transform based methods are unsuitable for machine monitoring in the above-mentioned circumstances [9].

The Need for Time-Frequency Methods: To overcome the shortcomings of the traditional spectral analysis techniques, nonstationary signal analysis approaches have been introduced. The most frequently used methods in the area of machine condition monitoring and diagnosis are quadratic time-frequency distributions (TFDs) and time-scale analysis (mainly the wavelet transforms (WT)). These methods represent the signals in a wider time-frequency space that allows easier and more precise discrimination between fault and normal machine conditions. Using time-frequency techniques, such as the Wigner-Ville distribution (WVD), a framework was developed that provided robust detection and classification schemes for helicopter gearbox faults [10]. This was achieved by showing that different faults produced different patterns in the time-frequency plane. The WVD-based patterns of vibration and acoustic signals were also used to detect faults in a number of machines and machine components such as engines [11] and gearboxes [12]. Other time-frequency distributions such as higher-order Wigner-Ville moment distributions [13] and reduced-interference time-frequency distributions (RIDs) [14] are used

for machine monitoring and diagnosis. Most of these methods, however, are visual-based detection/classification techniques which are meant to show the effectiveness of the respective TFDs for early detection of faults. The other methods are used as automatic feature extractors in an overall classification process. Some of the features extracted are amplitude values of the contour plots [12] and singular values of the TFD [14].

Due to their ability to represent nonstationary signals in general, and to detect and localize transient events in particular, wavelet transforms (both continuous and discrete) have been readily adopted in machine condition monitoring and diagnosis. They were used in detecting a large number of faults in different machines or machine components such as turning and drilling machines [1], gears or gear trains [15], and bearings [16]. As in the case of the TFDs, some of the proposed methods are used as feature extractors whose output is fed to a detector/classifier [1].

15.6.2 Time-Frequency Analysis Methods

Articles in Chapters 1 to 5 present detailed background on different time-frequency methods. The two most widely used time-frequency methodologies are the quadratic time-frequency distributions and the wavelet transforms. These two classes of representations are related through the STFT and the Gabor transform (see Articles 2.3 and 2.7). TFDs are suitable for large BT signals (see Chapter 1) while WTs give best results when used with low BT and transient signals.

15.6.2.1 Quadratic Time-Frequency Distributions

For nonstationary signals, the Wiener-Khintchine theorem indicates that the time-varying power spectral density, $S_x(t,f)$, of a real random signal $x(t)$ is related to the time-varying autocorrelation function, $R_x(t,\tau)$, by a Fourier transform relation; that is

$$S_x(t,f) = E\{W_x(t,f)\} = \int_{-\infty}^{\infty} R_x(t,\tau)\, e^{-j2\pi f\tau}\, d\tau. \tag{15.6.1}$$

The expression $S_x(t,f)$ given by Eq. (15.6.1) is the Wigner-Ville spectrum (WVS), which is the expectation value of the Wigner distribution (WD) $W_x(t,f)$. For practical reasons, $x(t)$ is replaced by its analytic associate $z(t)$ (see Sections 1.2.2, 1.2.3 and 2.1.4). It was shown that an estimate of $S_z(t,f)$ can be obtained from the quadratic class of TFDs [17], which was expressed in Section 3.2.2 as

$$\rho_z(t,f) = W_z(t,f) \underset{t}{*}\underset{f}{*} \gamma(t,f) \tag{15.6.2}$$

where $\gamma(t,f)$ is a two-dimensional kernel window which is application dependent, $W_z(t,f)$ is the WVD, and $\underset{tf}{**}$ indicates a double convolution. The example in Section 15.6.3.1 illustrates an application of the WVD to machine condition monitoring.

The kernel window, $\gamma(t,f)$, characterizes a particular time-frequency distribution and is generally chosen so as to obtain the best possible time-frequency resolution [14] (see Article 3.3 for more details).

15.6.2.2 Wavelet Transforms

Wavelet transforms are the localized equivalent of the Fourier transform. They provide a powerful tool for representing local features of a signal.

A finite-energy signal $x(t)$ can be represented by its Fourier transform $X(f)$:

$$x(t) = \int_{-\infty}^{\infty} X(f)\, e^{j2\pi ft}\, df \quad \text{where} \quad X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, dt. \tag{15.6.3}$$

Thus, the FT decomposes the time-domain signal into linear combinations of harmonics $e^{j2\pi ft}$. The wavelet transform (WT) is defined in the similar manner except that the harmonics are replaced by a series of wavelet basis functions given by [18]

$$\psi_{\tau s}(t) = \frac{1}{\sqrt{s}}\, \psi\left(\frac{t-\tau}{s}\right) \tag{15.6.4}$$

where τ and s are the translation and dilation (scale) parameters respectively. The function $\psi(...)$ is the transformation function called the mother wavelet. Using wavelet bases, the time-domain signal can be represented as

$$x(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{0}^{\infty} \Psi_x^{\psi}(\tau, s)\, \frac{1}{\sqrt{s}}\, \psi\left(\frac{t-\tau}{s}\right) \frac{ds}{s^2}\, d\tau \tag{15.6.5}$$

where

$$\Psi_x^{\psi}(\tau, s) = \int_{-\infty}^{\infty} x(t)\, \frac{1}{\sqrt{s}}\, \psi^*\left(\frac{t-\tau}{s}\right) dt \tag{15.6.6}$$

c_ψ is a constant that depends on the wavelet used and $\Psi_x^{\psi}(\tau, s)$ is the continuous wavelet transform of the signal $x(t)$. A number of mother wavelets have been proposed, such as the Mexican hat wavelet and the Morlet wavelet [18].

The discrete version of the WT is called discrete wavelet transform (DWT). It is realized by first discretizing the parameter scale s on a logarithmic grade. The time parameter is then discretized with respect to the scale parameter; that is a different sampling rate is used for every scale. In other words, the sampling is done on a dyadic sampling grid. With this sampling, a signal $x(t)$ can be decomposed into orthogonal basis functions (scaled and shifted versions of the mother wavelet ψ); that is [18]

$$x(t) = c_\psi \sum_l \sum_k a_{lk}\, s_0^{-l/2} \psi(s_0^{-l/2} t - k\tau_0) \tag{15.6.7}$$

where

$$a_{lk} = \int_{-\infty}^{\infty} x(t)\, s_0^{-l/2} \psi(s_0^{-l/2} t - k\tau_0)\, dt \tag{15.6.8}$$

with τ_0 and s_0 being positive constants usually taken as 1 and 2 respectively. The integer l describes the different levels of wavelets, and k covers the number of wavelets in each level.

The wavelet transform allows localization in both the time domain via translations of the mother wavelet, and in the scale (frequency) domain via dilations. The wavelet is irregular in shape and compactly supported, thus making it an ideal tool for analyzing signals of a transient nature. Irregularity of the wavelet basis lends it to analysis of signals with discontinuities or sharp changes, while the compactly supported nature of wavelets enables temporal localization of a signal's features.

The dilation function of the discrete wavelet transform can be represented as a tree of low- and high-pass filters, with each step transforming the low-pass filter. The original signal is successively decomposed into components of lower resolution, while the high-frequency components are not analyzed any further.

In contrast with the regular DWT, discrete wavelet packet analysis (DWPA) can significantly increase the versatility and power of the DWT. Unlike the DWT, DWPA utilizes both the low frequency components (approximations), and the high-frequency components (details). From this family of bases, a method for choosing the optimum scheme for a particular signal can be developed [18]. The two examples in Sections 15.6.3.2 and 15.6.3.3 illustrate the applications of DWT and DWPA to machine condition monitoring.

15.6.3 Examples of Condition Monitoring Using TFA

To illustrate how both time-frequency distributions and wavelet transforms are used in condition monitoring and diagnosis, we summarize three methods selected from the literature. These examples illustrate the time-frequency methodology adopted in this increasingly important area of engineering.

15.6.3.1 Gearbox Fault Detection Using Wigner-Ville Distribution

We consider the detection of a broken tooth in a spur gear using the WVD as the basis for feature extraction and pattern recognition techniques for classification [12].

Data Acquisition: The system considered is composed of a 24-tooth input gear driven by an electric motor and meshing with 16 teeth of a pinion whose rotational frequency is 37.5 Hz. The applied load was 70% of the maximum load. The study simulated five fault types, each involving the partial or total removal of one or more teeth. In particular, the faults were the removal of 25, 50, 75, and 100 percent of the face-width at a given radius, plus the same defect with 100% advancement on two pinion teeth. The acceleration vibration signal obtained from the above-mentioned system was low-pass filtered and sampled at a rate of 6.4 kHz.

Feature Extraction: The vibration signal is synchronously averaged in order to remove any periodic events not exactly synchronous with the gear of interest and to reduce the effects of noise and vibration sources other than that of the gear. In an industrial environment, where the problem of noise may become critical, efficient time-frequency based signal-cleansing techniques such as time-frequency peak filtering [19] (see also Article 11.4) may be required. The averaged signal is then transformed to the time-frequency domain using the pseudo-WVD (discrete

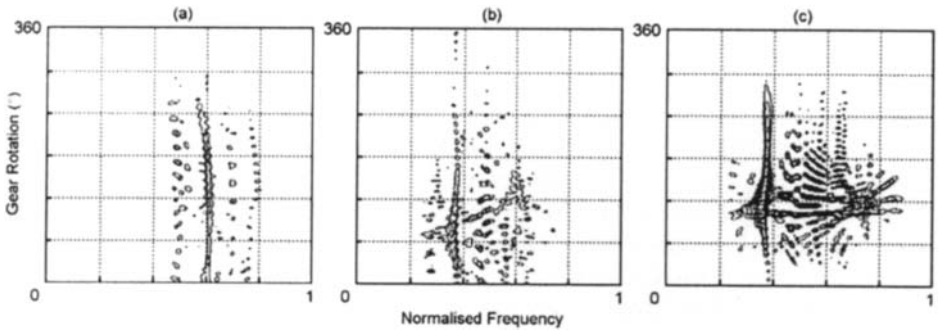


Fig. 15.6.1: Weighted WVD of the residual signal: (a) normal condition of the spur gear; (b) one broken tooth with 50% fault advancement; (c) one broken tooth with 100% fault advancement.

WVD) with a Hamming window. The negative values of the WVD are set to zero and the resulting distribution is normalized. The results are displayed in the form of contour plots. To enhance the sidebands around the meshing frequencies, the residual signal is obtained by removing the meshing harmonics using a band-stop filter. The extracted features are the amplitude values of the contour plots (see Fig. 15.6.1).

Feature Selection: To reduce the dimension of the feature vector, a selected number of WVD cross-sections at and around a chosen meshing frequency are selected.

Decision Making: Two classification approaches are considered: statistical and neural pattern recognition. In the first approach, to assign the feature vector from the last stage to one of the K classes considered, the Mahalanobis distance was chosen as the similarity measure. This measure is given by

$$d = \left[\sum_{k=1}^{K} (\vec{y} - \vec{\mu_k})^T \, \Sigma_k^{-1} (\vec{y} - \vec{\mu_k}) \right]^{1/2} \tag{15.6.9}$$

where $\vec{y}$ is the feature vector and $\vec{\mu_k}$ and Σ_k are the mean vector and covariance matrix representing the k^{th} class. The study considered only two classes; namely normal (no fault) and abnormal (fault) and used only one template representing the normal condition. In the second classification approach a neural network was trained in a supervised mode using the back-propagation algorithm [12].

15.6.3.2 Fault Diagnosis of Rotating Machinery Using Wavelet Transforms

In this example, the problem is to detect faults in a model drive-line consisting of various interconnected rotating parts that include a vehicle gearbox, two bearing housings, and an electric motor. All these parts are connected by flexible couplings and loaded by a disk brake as seen in Fig. 15.6.2 [20].

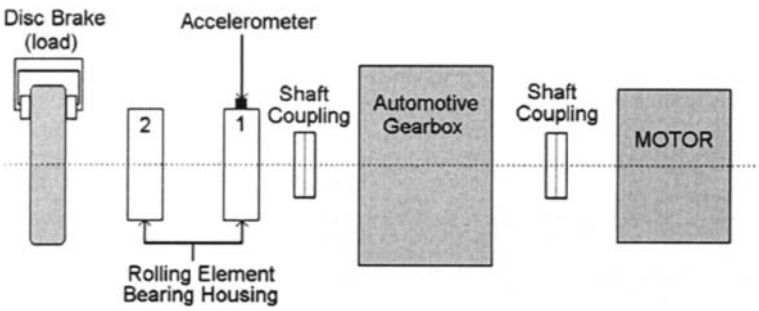
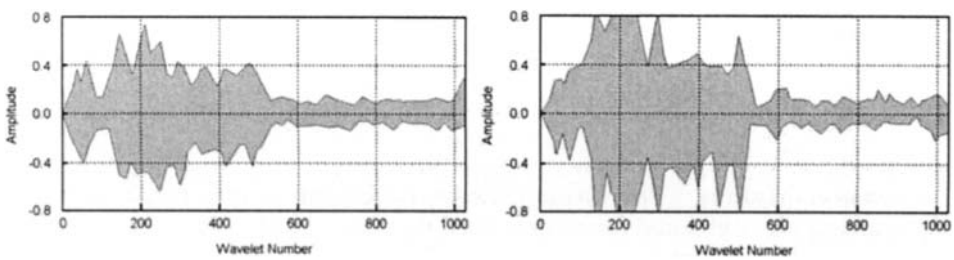


Fig. 15.6.2: Schematic presentation of the model drive-line.



(a) Envelope of the WT of a vibration signal representing a normal condition.

(b) Envelope of the WT of a vibration signal representing a faulty gear.

Fig. 15.6.3: WT of normal and faulty condition of the vibration signal.

Data Acquisition: Of the five gears (four forward and one reverse) only the reverse gear pinion is used in the experiment. On the gear pinion, two types of localized faults were simulated: a small "blip" of 2mm diameter on one tooth, and a triangular fracture on the face of one tooth. On the bearing housing, one fault was simulated by introducing a 1mm fracture across the inner race (Fig. 15.6.2). This gave six combinations of conditions for the pinion and housing, five of which represented fault conditions. An accelerometer was used to obtain the vibration signal from the bearing housing.

Feature Extraction: The vibration signals were transformed to the time-scale domain using the Daubechies 4th-order wavelet (D4) (see Fig. 15.6.3). After the transformation of the whole signal into the wavelet domain, a threshold value was chosen. This value was selected to be above the dominant component of the reference (normal) signal. The 10 most dominant amplitudes of the signals above the threshold value were selected to represent half of the feature vector. The other half consists of the 10 corresponding wavelet numbers (indicating both time and scale). As the number of features (20) is not large, no feature selection was needed.

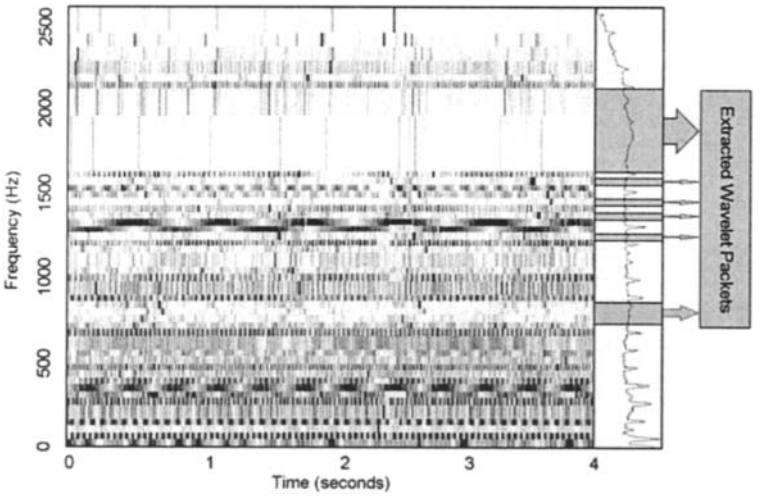


Fig. 15.6.4: DWPA representation of the vibration signal showing wavelet packets selected by ANFIS.

Decision Making: The classification was achieved using a two-layer neural network with sigmoid nodal function trained in a supervised mode using the back-propagation algorithm [20].

15.6.3.3 Extraction of Bearing Fault Transients Using DWPA

This example exploits the multiple band-pass filtering capability of the DWPA for the extraction of rolling-element bearing fault-related components. An algorithm is trained to recognize three types of localized faults; namely inner race, rolling element, and outer race faults [21].

Data Acquisition: The vibration signals are obtained from a rolling-element bearing test rig with a rolling-element fault and an operating speed of 60 rpm.

Feature Extraction: The extraction of high-frequency transients due to bearing impact resonance is achieved via best-basis DWPA representation using the Daubechies wavelet of order 20 and an adaptive network-based fuzzy inference system (ANFIS). ANFIS is a transformational model of integration where the final fuzzy inference system is optimized via artificial neural network training. Before the neuro-fuzzy network is trained (using wavelet packets extracted from vibration signals), suitable input parameters to train the network are selected. These parameters are kurtosis (a measure of spikiness) and the spectrum peak ratio (an indication of the presence of localized defects). The network is then trained using wavelet packets characterizing the above-mentioned types of faults. Fig. 15.6.4 illustrates how this method facilitates the extraction of bearing-fault-related components from a signal while rejecting the unwanted harmonics. The wavelet packets identified by ANFIS as containing bearing fault-related features are indicated on the figure [21].

15.6.4 Summary and Conclusions

Time-frequency analysis methods are applicable to the area of machine condition monitoring and diagnosis. They are capable of efficiently and unambiguously characterizing a large number of faults. TFA methods are used for detection, classification, and monitoring the progression of the faults and wear with time. This enables prediction and prevention of catastrophic failures. Time-frequency analysis techniques, in the form of either TFD or WT, are used as both visual indicators of the presence of faults and as a feature extractor in a fully automated pattern recognition process.

Articles 11.2 and 15.2 of this book describe two other time-frequency approaches to machine condition monitoring.

References

[1] R. Du, "Engineering monitoring and diagnosis using wavelet transforms," in *Computer-Aided Design, Engineering, and Manufacturing: Systems Techniques and Applications—Vol. I: Systems Techniques and Computational Methods* (C. T. Leondes, ed.), ch. 8, CRC Press, 2000.

[2] D. E. Dimla Snr., "Sensor signals for tool-wear monitoring in metal cutting operations—A review of methods," *Internat. J. Machine Tools & Manufacture*, vol. 40, pp. 1073–1098, June 2000.

[3] S. R. Kulkarni, G. Lugosi, and S. S. Venkatesh, "Learning pattern classification—A survey," *IEEE Trans. Information Theory*, vol. 44, pp. 2178–2206, October 1998.

[4] D. Logan and J. Mathew, "Using the correlation dimension for vibration fault diagnosis of rolling element bearings—I: Basic concepts," *J. Mechanical Systems and Signal Processing*, vol. 10, pp. 241–250, May 1996.

[5] D. C. Baillie and J. Mathew, "A comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings," *J. Mechanical Systems and Signal Processing*, vol. 10, pp. 1–17, January 1996.

[6] J. Courrech and R. L. Eshleman, "Condition monitoring of machinery," in *Harris' Shock and Vibration Handbook* (C. M. Harris and A. G. Piersol, eds.), ch. 16, New York: McGraw-Hill, 2002.

[7] R. K. Mobley, *Vibration Fundamentals*. Boston: Newnes, 1999.

[8] G. C. Zhang, M. Ge, H. Tong, Y. Xu, and R. Du, "Bispectral analysis for on-line monitoring of stamping operation," *Engineering Applications of Artificial Intelligence*, vol. 15, pp. 97–104, February 2002.

[9] R. Burnett, J. F. Watson, and S. Elder, "The application of modern signal processing techniques for use in rotor fault detection and location within three-phase induction motors," *Signal Processing*, vol. 49, pp. 57–70, February 1996.

[10] B. D. Forrester, "Time-frequency analysis in machine fault detection," in *Time-Frequency Signal Analysis: Methods and Applications* (B. Boashash, ed.), ch. 18, pp. 406–423, Melbourne/N.Y.: Longman-Cheshire/Wiley, 1992.

[11] B. Boashash and P. J. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques," *IEEE Trans. Acoustics, Speech, & Signal Processing*, vol. 38, pp. 1829–1841, November 1990.

[12] W. J. Staszewski, K. Worden, and G. R. Tomlinson, "Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition," *J. Mechanical Systems and Signal Processing*, vol. 11, pp. 673–692, September 1997.

[13] S. K. Lee and P. R. White, "Higher-order time-frequency analysis and its application to fault detection in rotating machinery," *J. Mechanical Systems and Signal Processing*, vol. 11, pp. 637–650, July 1997.

[14] S. Gu, J. Ni, and J. Yuan, "Non-stationary signal analysis and transient machining process condition monitoring," *Internat. J. Machine Tools & Manufacture*, vol. 42, pp. 41–51, January 2002.

[15] H. Zheng, Z. Li, and X. Chen, "Gear fault diagnosis based on continuous wavelet transform," *J. Mechanical Systems and Signal Processing*, vol. 16, pp. 447–457, March–May 2002.

[16] S. Prabhakar, A. S. Sekhar, and A. R. Mohanty, "Detection and monitoring of cracks in a rotor-bearing system using wavelet transforms," *J. Mechanical Systems and Signal Processing*, vol. 15, pp. 447–450, March 2001.

[17] B. Boashash and M. Mesbah, "A time-frequency approach for newborn seizure detection," *IEEE Engineering in Medicine & Biology Magazine*, vol. 20, pp. 54–64, September/October 2001.

[18] S. G. Mallat, *A Wavelet Tour of Signal Processing*. San Diego / London: Academic Press, 2nd ed., 1999.

[19] B. Boashash and M. Mesbah, "Signal enhancement by time-frequency peak filtering," *IEEE Trans. Signal Processing*, vol. 51, November 2003.

[20] B. A. Paya, I. I. Esat, and M. N. M. Badi, "Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor," *J. Mechanical Systems and Signal Processing*, vol. 11, pp. 751–765, September 1997.

[21] J. Altmann and J. Mathew, "DWPA multiple band-pass filtering for extraction of bearing fault transients from a strong continuous signal," in *Proc. Asia-Pacific Vibration Conference*, vol. 2, pp. 772–776, Hangzhou, China, October–November 2001.

Chapter 16

Other Applications

There are a number of applications that may be of interest to the reader but that could not be included in the chapters for obvious space reasons. A list of references to selected applications is provided below to further assist the reader of this book. They represent the areas of biomedical engineering [1–7], speech [8–10], radar [11–13], telecommunications [14, 15], plasma science [16], image [17, 18] and tomography [19]. Further references can be obtained by the reader using available computer searches on the relevant databases (e.g. Elsevier, IEEE, etc.).

References

[1] A. Monti, C. Medigue, and L. Mangin, "Instantaneous parameter estimation in cardiovascular time series by harmonic and time-frequency analysis," *IEEE Trans. Biomedical Engineering*, vol. 49, pp. 1547–1556, December 2002.

[2] H. G. van Steenis, W. L. J. Martens, and J. H. M. Tulen, "Time-frequency parameters of heart-rate variability," *IEEE Engineering in Medicine & Biology Magazine*, vol. 21, pp. 46–58, July-August 2002.

[3] R. Zou, W. A. Cupples, K. R. Yip, N. H. Holstein-Rathlou, and K. Chon, "Time-varying properties of renal autoregulatory mechanisms," *IEEE Trans. Biomedical Engineering*, vol. 49, pp. 1112–1120, October 2002.

[4] Y. Hu, K. D. Luk, W. W. Lu, and J. C. Leong, "Comparison of time-frequency analysis techniques in intraoperative somatosensory evoked potential (SEP) monitoring," *Computers in Biology & Medicine*, vol. 32, pp. 13–23, January 2002.

[5] P. Bonato, P. Boissy, U. D. Croce, and S. H. Roy, "Changes in the surface EMG signal and the biomechanics of motion during a repetitive lifting task," *IEEE Trans. Neural System & Rehabilitation Engineering*, vol. 10, pp. 38–47, March 2002.

[6] L. Rankine and M. Mesbah, "Significant atom determination of basis pursuit decomposition," in *Proc. Seventh Internat. Symp. on Signal Processing and its Applications (ISSPA'03)*, pp. 577–580, Paris, 1–4 July 2003.

[7] P. Zarjam, M. Mesbah, and B. Boashash, "Detection of newborn EEG seizure using optimal features based on discrete wavelet transform," in *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'03)*, vol. 2, pp. 265–268, Hong Kong, scheduled 6–10 April 2003.

[8] I. Magrin-Chagnolleau, G. Durou, and F. Bimbot, "Application of time-frequency principal component analysis to text-independent speaker identification," *IEEE Trans. on Speech & Audio Processing*, vol. 10, pp. 371–378, September 2002.

[9] S. Ghaemmaghami, M. Deriche, and B. Boashash, "Hierarchical approach to formant detection and tracking through instantaneous frequency estimation," *IEE Electronics Letters*, vol. 33, pp. 17–18, January 1997.

[10] M. D. Riley, *Speech Time-Frequency Representations.* Kluwer, 1989.

[11] X.-G. Xia, G. Wang, and V. Chen, "Quantitative SNR analysis for ISAR imaging using joint time-frequency analysis—Short time Fourier transform," *IEEE Trans. Aerospace & Electronic Systems*, vol. 38, pp. 649–659, April 2002.

[12] O. P. Kenny and B. Boashash, "Time-frequency analysis of backscattered signals from diffuse radar targets," *IEE Proc., Part F: Radar & Signal Processing*, vol. 140, pp. 198–208, June 1993.

[13] V. C. Chen and H. Ling, *Time-Frequency Transforms for Radar Imaging and Signal Analysis.* Scitech Publishing, 2002.

[14] A. Z. Sha'ameri, B. Boashash, and I. Ismail, "Design of signal dependent kernel functions for digital modulated signals," in *Proc. Fourth Internat. Symp. on Signal Processing and its Applications (ISSPA '96)*, vol. 2, pp. 527–528, Gold Coast, Australia, 25–30 August 1996.

[15] G. Azemi, B. Senadji, and B. Boashash, "A novel estimator for the velocity of a mobile base station in a micro-cellular system," in *Proc. IEEE Internat. Symp. on Circuits and Systems (ISCAS 03)*, vol. 2, pp. 212–215, Bangkok, 25–28 May 2003.

[16] A. C. A. Figueiredo and J. P. S. Bizarro, "Time-frequency images of magnetohydrodynamic phenomena in tokamak plasmas using a discrete-time Wigner distribution," *IEEE Trans. Plasma Science*, vol. 30, pp. 54–55, February 2002.

[17] I. Djurović and S. Stanković, "Estimation of time-varying velocities of moving objects by time-frequency representations," *IEEE Trans. Image Processing*, vol. 12, pp. 550–562, May 2003.

[18] S. L. Hahn and K. M. Snopek, "Double-dimensional distributions: Another approach to "quartic" distributions," *IEEE Trans. Signal Processing*, vol. 50, pp. 2987–2997, December 2002.

[19] C. Richard, "Time-frequency-based detection using discrete-time discrete-frequency Wigner distributions," *IEEE Trans. Signal Processing*, vol. 50, pp. 2170–2176, September 2002.

Time-Frequency Bibliography

Reference numbers in the text of this book refer to the local "References" list at the end of the current Chapter (in Part I and Chapter 16) or the current Article (elsewhere). The following "time-frequency" bibliography[0] is an alphabetical listing of all publicly-available sources that appear in the local "References" lists, excluding application-specific references not pertaining to the time-frequency field, and adding some other important references which, because of space and scope restrictions, are not cited in the text. It is intended as a snapshot, dated 2003, of the most important references in TFSAP and its applications.

If the same source appears in two or more local "References" lists, the reference numbers generally differ, but the substantive details have been verified and harmonized.[1]

[1] K. Abed-Meraim and B. Barkat. "Blind source separation using the time-frequency distribution of the mixture signal". In *Proc. Second IEEE Internat. Symp. on Signal Processing and Information Technology (ISSPIT'02)*, volume 2, pages 663–666, Marrakech, Morocco, 18–21 December 2002.

[2] K. Abed-Meraim, N. Linh-Trung, V. Sucic, F. Tupin, and B. Boashash. "An image processing approach for underdetermined blind separation of nonstationary sources". In *Proc. Third Internat. Symp. on Image & Signal Processing and Analysis (ISPA-03)*, Rome, 18–20 September 2003.

[3] S. S. Abeysekera and B. Boashash. "Methods of signal classification using the images produced by the Wigner-Ville distribution". *Pattern Recognition Letters*, 12(11):717–729, November 1991.

[4] R. Adelino and F. da Silva. "Atomic decomposition with evolutionary pursuit". *Digital Signal Processing: A Review Journal*, 13(2):317–337, April 2003.

[5] O. Akay and G. F. Boudreaux-Bartels. "Unitary and Hermitian fractional operators and their relation to the fractional Fourier transform". *IEEE Signal Processing Letters*, 5(12):312–314, December 1998.

[6] O. Akay and G. F. Boudreaux-Bartels. "Fractional convolution and correlation via operator methods and an application to detection of linear FM signals". *IEEE Trans. Signal Processing*, 49(5):979–993, May 2001.

[7] L. B. Almeida. "The fractional Fourier transform and time-frequency representations". *IEEE Trans. Signal Processing*, 42(11):3084–3091, November 1994.

[8] R. A. Altes. "Detection, estimation and classification with spectrograms". *J. Acoustical Soc. of America*, 67(4):1232–1246, April 1980.

[0] Authors: **Gavin R. Putland** and **Boualem Boashash**, Signal Processing Research Centre, Queensland University of Technology, Brisbane, Australia.

[1] A consolidated bibliographic database was compiled from the references supplied by contributing authors. Local reference lists were regenerated from that database using the `bibunits` package (v. 2.0 by Thorsten Hansen).

[9] R. A. Altes. "Wideband, proportional-bandwidth Wigner-Ville analysis". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 38(6):1005–1012, June 1990.

[10] R. A. Altes and E. L. Titlebaum. "Bat signals as optimally Doppler tolerant waveforms". *J. Acoustical Soc. of America*, 48:1014–1020, October 1970.

[11] J. Altmann and J. Mathew. "DWPA multiple band-pass filtering for extraction of bearing fault transients from a strong continuous signal". In *Proc. Asia-Pacific Vibration Conference*, volume 2, pages 772–776, Hangzhou, China, October–November 2001.

[12] M. G. Amin. "Time-frequency spectrum analysis and estimation for non-stationary random processes". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 9, pages 208–232. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[13] M. G. Amin. "Spectral decomposition of time-frequency distribution kernels". *IEEE Trans. Signal Processing*, 42(5):1156–1165, May 1994.

[14] M. G. Amin. "Minimum-variance time-frequency distribution kernels for signals in additive noise". *IEEE Trans. Signal Processing*, 44(9):2352–2356, September 1996.

[15] M. G. Amin. "Interference mitigation in spread-spectrum communication systems using time-frequency distributions". *IEEE Trans. Signal Processing*, 45(1):90–102, January 1997.

[16] M. G. Amin and A. N. Akansu. "Time-frequency for interference excision in spread-spectrum communications". In G. B. Giannakis, editor, "Highlights of Signal Processing for Communications", *IEEE Signal Processing Magazine*, volume 16(2). IEEE, March 1999.

[17] M. G. Amin, A. Belouchrani, and Y. Zhang. "The spatial ambiguity function and its applications". *IEEE Signal Processing Letters*, 7(6):138–140, June 2000.

[18] M. G. Amin, C. Wang, and A. R. Lindsey. "Optimum interference excision in spread-spectrum communications using open-loop adaptive filters". *IEEE Trans. Signal Processing*, 47(7):1966–1976, July 1999.

[19] M. G. Amin and Y. Zhang. "Direction finding based on spatial time-frequency distribution matrices". *Digital Signal Processing: A Review Journal*, 10(4):325–339, October 2000.

[20] J. C. Andrieux, R. Feix, G. Mourgues, P. Bertrand, B. Izrar, and V. T. Nguyen. "Optimum smoothing of the Wigner-Ville distribution". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 35(6):764–769, June 1987.

[21] L. Angrisani and M. D'Arco. "A measurement method based on a modified version of the chirplet transform for instantaneous frequency estimation". *IEEE Trans. Instrumentation & Measurement*, 51(4):704–711, August 2002.

[22] M. J. Arnold, M. Roessgen, and B. Boashash. "Filtering real signals through frequency modulation and peak detection in the time-frequency plane". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, volume 3, pages 345–348, Adelaide, Australia, 19–22 April 1994.

[23] H. Artés, G. Matz, and F. Hlawatsch. "Unbiased scattering function estimation during data transmission". In *Proc. IEEE Vehicular Technology Conf. (VTC'99-Fall)*, pages 1535–1539, Amsterdam, 19–22 September 1999.

[24] F. Auger and P. Flandrin. "Improving the readability of time-frequency and time-scale representations by the reassignment method". *IEEE Trans. Signal Processing*, 43(5):1068–1089, May 1995.

[25] G. Azemi, B. Senadji, and B. Boashash. "A novel estimator for the velocity of a mobile base station in a micro-cellular system". In *Proc. IEEE Internat. Symp. on Circuits and Systems (ISCAS 03)*, volume 2, pages 212–215, Bangkok, 25–28 May 2003.

[26] J. S. Bach (1685-1750). *St Matthew Passion*, nos. 2–3 (ed. A. Van Ryckeghem). *GMD Music Archive*, http://www.gmd.de/Misc/Music/scores/Introduction.html, 2000. Continued by *Werner Icking Music Archive*, http://icking-music-archive.sunsite.dk/ .

[27] R. Baraniuk, D. Jones, T. Brotherton, and S. L. Marple Jr. "Applications of adaptive time-frequency representations to underwater acoustic signal processing". In *Proc. 25th Asilomar Conf. on Signals, Systems, and Computers*, volume 2, pages 1109–1113, Pacific Grove, CA, 4–6 November 1991.

[28] R. G. Baraniuk. "Covariant time-frequency representations through unitary equivalence". *IEEE Signal Processing Letters*, 3(3):79–81, March 1996.

[29] R. G. Baraniuk. "Beyond time-frequency analysis: Energy densities in one and many dimensions". *IEEE Trans. Signal Processing*, 46(9):2305–2314, September 1998.

[30] R. G. Baraniuk. "Joint distributions of arbitrary variables made easy". *J. of Multidimensional Systems & Signal Processing*, 9(4):341–348, October 1998. Special issue on time-frequency analysis.

[31] R. G. Baraniuk and L. Cohen. "On joint distributions of arbitrary variables". *IEEE Signal Processing Letters*, 2(1):10–12, January 1995.

[32] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel. "Measuring time-frequency information content using the Rényi entropies". *IEEE Trans. Information Theory*, 47(4):1391–1409, May 2001.

[33] R. G. Baraniuk and D. L. Jones. "Signal-dependent time-frequency analysis using a radially Gaussian kernel". *Signal Processing*, 32(3):263–284, June 1993.

[34] R. G. Baraniuk and D. L. Jones. "A signal-dependent time-frequency representation: Optimal kernel design". *IEEE Trans. Signal Processing*, 41(4):1589–1602, April 1993.

[35] R. G. Baraniuk and D. L. Jones. "A signal-dependent time-frequency representation: Fast algorithm for optimal kernel design". *IEEE Trans. Signal Processing*, 42(1):134–146, January 1994.

[36] R. G. Baraniuk and D. L. Jones. "Unitary equivalence: A new twist on signal processing". *IEEE Trans. Signal Processing*, 43(10):2269–2282, October 1995.

[37] S. Barbarossa. "Detection and imaging of moving objects with synthetic aperture radar—Part 1: Optimal detection and parameter estimation theory". *IEE Proc., Part F: Radar & Signal Processing*, 139(1):79–88, February 1992.

[38] S. Barbarossa and O. Lemoine. "Analysis of nonlinear FM signals by pattern recognition of their time-frequency representation". *IEEE Signal Processing Letters*, 3(4):112–115, April 1996.

[39] S. Barbarossa and A. Scaglione. "Autofocusing of SAR images based on the product high-order ambiguity function". *IEE Proc.: Radar, Sonar & Navigation*, 145(5):269–273, October 1998.

[40] S. Barbarossa and A. Scaglione. "Adaptive time-varying cancellations of wideband interferences in spread-spectrum communications based on time-frequency distributions". *IEEE Trans. Signal Processing*, 47(4):957–965, April 1999.

[41] S. Barbarossa and A. Scaglione. "On the capacity of linear time-varying channels". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'99)*, pages 2627–2630, Phoenix, AZ, 15–19 March 1999.

[42] S. Barbarossa and A. Scaglione. "Time-varying fading channels". In G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, editors, *Signal Processing Advances in Wireless and Mobile Communications*, volume 2: "Trends in Single- and Multi-User Systems", chapter 1. Prentice-Hall, Upper Saddle River, NJ, 2001.

[43] S. Barbarossa, A. Scaglione, and G. B. Giannakis. "Product high-order ambiguity function for multicomponent polynomial-phase signal modeling". *IEEE Trans. Signal Processing*, 46(3):691–708, March 1998.

[44] B. Barkat. "Instantaneous frequency estimation of nonlinear frequency-modulated signals in the presence of multiplicative and additive noise". *IEEE Trans. Signal Processing*, 49(10):2214–2222, October 2001.

[45] B. Barkat and K. Abed-Meraim. "A blind components separation procedure for FM signal analysis". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'02)*, volume 2, pages 1425–1428, Orlando, FL, 13–17 May 2002.

[46] B. Barkat and B. Boashash. "Higher order PWVD and Legendre based time-frequency distribution". In *Proc. Sixth IEEE Internat. Workshop on Intelligent Signal Processing and Communication Systems (ISPACS'98)*, volume 2, pages 532–536, Melbourne, Australia, 5–6 November 1998.

[47] B. Barkat and B. Boashash. "Design of higher order polynomial Wigner-Ville distributions". *IEEE Trans. Signal Processing*, 47(9):2608–2611, September 1999.

[48] 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 Processing*, 47(9):2480–2490, September 1999.

[49] B. Barkat and B. Boashash. "A high-resolution quadratic time-frequency distribution for multicomponent signals analysis". *IEEE Trans. Signal Processing*, 49(10):2232–2239, October 2001.

[50] H. O. Bartelt, K. H. Brenner, and A. W. Lohmann. "The Wigner distribution function and its optical production". *Optics Communications*, 32(1):32–38, January 1980.

[51] M. J. Bastiaans. "Gabor's signal expansion and its relation to sampling of the sliding-window spectrum". In R. J. Marks II, editor, *Advanced Topics in Shannon Sampling and Interpolation Theory*, pages 1–35. Springer, New York, 1993.

[52] M. J. Bastiaans. "Gabor's expansion and the Zak transform for continuous-time and discrete-time signals: Critical sampling and rational oversampling". Research Report 95-E-295, Eindhoven University of Technology, Eindhoven, Netherlands, 1995.

[53] M. J. Bastiaans, T. Alieva, and L. Stanković. "On rotated time-frequency kernels". *IEEE Signal Processing Letters*, 9(11):378–381, November 2002.

[54] M. J. Bastiaans and M. C. W. Geilen. "On the discrete Gabor transform and the discrete Zak transform". *Signal Processing*, 49(3):151–166, March 1996.

[55] E. Bedrosian. "A product theorem for Hilbert transforms". *Proc. IEEE*, 51:686–689, 1963.

[56] P. A. Bello. "Characterization of randomly time-variant linear channels". *IEEE Trans. Communication Systems*, 11(4):360–393, December 1963.

[57] A. Belouchrani and M. G. Amin. "New approach for blind source separation using time frequency distributions". In *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures and Implementations VI*, volume 2846, pages 193–203. Soc. of Photo-optical Instrumentation Engineers, Denver, CO, 4–9 August 1996.

[58] A. Belouchrani and M. G. Amin. "Blind source separation using time-frequency distributions: Algorithm and asymptotic performance". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, volume 5, pages 3469–3472, Munich, 21–24 April 1997.

[59] A. Belouchrani and M. G. Amin. "Blind source separation based on time-frequency signal representation". *IEEE Trans. Signal Processing*, 46(11):2888–2898, November 1998.

[60] A. Belouchrani and M. G. Amin. "Time-frequency MUSIC: A new array signal processing method based on time-frequency signal representation". *IEEE Signal Processing Letters*, 6(5):109–110, May 1999.

[61] A. Belouchrani, M. G. Amin, and K. Abed-Meraim. "Direction finding in correlated noise fields based on joint block-diagonalization of spatio-temporal correlation matrices". *IEEE Signal Processing Letters*, 4(9):266–268, September 1997.

[62] M. Benidir. "Characterization of polynomial functions and application to time-frequency analysis". *IEEE Trans. Signal Processing*, 45(5):1351–1354, May 1997.

[63] M. Benidir and A. Ouldali. "Polynomial phase signal analysis based on the polynomial derivatives decompositions". *IEEE Trans. Signal Processing*, 47(7):1954–1965, July 1999.

[64] J. Bertrand and P. Bertrand. "Représentations temps-fréquence des signaux". *Comptes rendus de l'Académie des Sciences, Paris*, 299, Ser. 1:635–638, 1984.

[65] J. Bertrand and P. Bertrand. "A tomographic approach to Wigner's function". *Foundations of Physics*, 17:397–405, 1987.

[66] J. Bertrand and P. Bertrand. "Affine time-frequency distributions". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 5, pages 118–140. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[67] J. Bertrand and P. Bertrand. "A class of affine Wigner functions with extended covariance properties". *J. of Mathematical Physics*, 33(7):2515–2527, July 1992.

[68] J. Bertrand and P. Bertrand. "Symbolic calculus on the time-frequency half-plane". *J. of Mathematical Physics*, 39(8):4071–4090, August 1998.

[69] J. Bertrand, P. Bertrand, and J. P. Ovarlez. "The Mellin transform". In A. D. Poularikas, editor, *The Transforms and Applications Handbook*, chapter 11, pages 829–885. CRC Press, Boca Raton, FL, 1st edition, 1996.

[70] S. Bhashyam, A. M. Sayeed, and B. Aazhang. "Time-selective signaling and reception for communication over multipath fading channels". *IEEE Trans. Communications*, 48:83–94, January 2000.

[71] B. Boashash. "Wigner analysis of time-varying signals—Its application in seismic prospecting". In *Proc. European Signal Processing Conf. (EUSIPCO-83)*, pages 703–706, Nürnberg, September 1983.

[72] B. Boashash. "Note on the use of the Wigner distribution for time-frequency signal analysis". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 36(9):1518–1521, September 1988.

[73] B. Boashash. "Time-frequency signal analysis". In S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, volume 1, chapter 9, pages 418–517. Prentice-Hall, Englewood Cliffs, NJ, 1991.

[74] B. Boashash. "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals". *Proc. IEEE*, 80(4):520–538, April 1992.

[75] B. Boashash. "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals; Part 2: Algorithms and applications". *Proc. IEEE*, 80(4):519–568, April 1992. With introductions by Michael J. Riezenman.

[76] B. Boashash. "Estimating and interpreting the instantaneous frequency of a signal—Part 2: Algorithms and applications". *Proc. IEEE*, 80(4):540–568, April 1992.

[77] B. Boashash, editor. *Time-Frequency Signal Analysis: Methods and Applications.* Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[78] B. Boashash. "Time-Frequency Signal Analysis: Past, present and future trends". In C. T. Leondes, editor, *Digital Control and Signal Processing Systems and Techniques*, number 78 in CONTROL AND DYNAMIC SYSTEMS: ADVANCES IN THEORY AND APPLICATIONS, chapter 1, pages 1–69. Academic Press, San Diego, 1996.

[79] B. Boashash, editor. *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference.* Elsevier, Amsterdam, 2003.

[80] B. Boashash and P. J. Black. "An efficient real-time implementation of the Wigner-Ville distribution". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 35(11):1611–1618, November 1987.

[81] B. Boashash and B. Escudie. "Wigner-Ville analysis of asymptotic signals and applications". *Signal Processing*, 8(3):315–327, June 1985.

[82] B. Boashash and G. Frazer. "Time-varying higher-order spectra, generalised Wigner-Ville distribution and the analysis of underwater acoustic data". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'92)*, volume 5, pages 193–196, San Francisco, 23–26 March 1992.

[83] B. Boashash, G. Jones, and P. J. O'Shea. "Instantaneous frequency of signals: Concepts, estimation techniques and applications". In *Proc. SPIE: Advanced Algorithms and Architectures for Signal Processing IV*, volume 1152, pages 382–400. Soc. of Photo-optical Instrumentation Engineers, San Diego, 8–10 August 1989.

[84] B. Boashash and M. Mesbah. "A time-frequency approach for newborn seizure detection". *IEEE Engineering in Medicine & Biology Magazine*, 20(5):54–64, September/October 2001.

[85] B. Boashash and M. Mesbah. "Time-frequency methodology for newborn EEG seizure detection". In A. Papandreou-Suppappola, editor, *Applications in Time-Frequency Signal Processing*, chapter 9. CRC Press, Boca Raton, FL, 2002.

[86] B. Boashash and M. Mesbah. "Signal enhancement by time-frequency peak filtering". *IEEE Trans. Signal Processing*, 51(11), November 2003.

[87] B. Boashash, M. Mesbah, and P. Colditz. "Newborn EEG seizure pattern characterisation using time-frequency analysis". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, volume 2, pages 1041–1044, Salt Lake City, UT, 7–11 May 2001.

[88] B. Boashash and P. O'Shea. "Use of the cross Wigner-Ville distribution for estimation of instantaneous frequency". *IEEE Trans. Signal Processing*, 41(3):1439–1445, March 1993.

[89] B. Boashash and P. J. O'Shea. "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis techniques". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 38(11):1829–1841, November 1990.

[90] B. Boashash and P. J. O'Shea. "Polynomial Wigner-Ville distributions and their relationship to time-varying higher order spectra". *IEEE Trans. Signal Processing*, 42(1):216–220, January 1994.

[91] B. Boashash, P. J. O'Shea, and M. J. Arnold. "Algorithms for instantaneous frequency estimation: A comparative study". In *Proc. SPIE: Advanced Signal-Processing Algorithms, Architectures, and Implementations*, volume 1348, pages 126–148. Soc. of Photo-optical Instrumentation Engineers, San Diego, 10–12 July 1990.

[92] B. Boashash and A. Reilly. "Algorithms for time-frequency signal analysis". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 7, pages 163–181. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[93] B. Boashash and B. Ristic. "Analysis of FM signals affected by Gaussian AM using the reduced Wigner-Ville trispectrum". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'93)*, volume IV, pages 408–411, Minneapolis, 27–30 April 1993.

[94] B. Boashash and B. Ristic. "A time-frequency perspective of higher-order spectra as a tool for non-stationary signal analysis". In B. Boashash, E. J. Powers, and A. M. Zoubir, editors, *Higher-Order Statistical Signal Processing*, chapter 4, pages 111–149. Longman/Wiley, Melbourne/N.Y., 1995.

[95] B. Boashash and B. Ristic. "Polynomial time-frequency distributions and time-varying higher order spectra: Application to the analysis of multicomponent FM signal and to the treatment of multiplicative noise". *Signal Processing*, 67(1):1–23, May 1998.

[96] B. Boashash and V. Sucic. "Resolution measure criteria for the objective assessment of the performance of quadratic time-frequency distributions". *IEEE Trans. Signal Processing*, 51(5):1253–1263, May 2003.

[97] B. Boashash, L. B. White, and J. Imberger. "Wigner-Ville analysis of non-stationary random signals (with application to turbulent microstructure signals)". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'86)*, volume 4, pages 2323–2326, Tokyo, 7–11 April 1986.

[98] B. Boashash and H. J. Whitehouse. "High resolution Wigner-Ville analysis". In *Eleventh GRETSI Symp. on Signal Processing and its Applications*, pages 205–208, Nice, France, 1–5 June 1987.

[99] H. Bölcskei and F. Hlawatsch. "Oversampled modulated filter banks". In H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, chapter 9, pages 295–322. Birkhäuser, Berlin/Boston, 1998.

[100] P. J. Boles and B. Boashash. "Application of the cross-Wigner-Ville distribution to seismic data processing". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 20, pages 445–466. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[101] P. Bonato, P. Boissy, U. D. Croce, and S. H. Roy. "Changes in the surface EMG signal and the biomechanics of motion during a repetitive lifting task". *IEEE Trans. Neural System & Rehabilitation Engineering*, 10(1):38–47, March 2002.

[102] M. Born and P. Jordan. "Zur quantenmechanik". *Zeitschrift für Physik.*, 34:858–888, 1925.

[103] G. F. Boudreaux-Bartels. *Time-frequency signal processing algorithms: Analysis and synthesis using Wigner distribution.* PhD thesis, Rice University, 1983.

[104] G. F. Boudreaux-Bartels. "Mixed time-frequency signal transformations". In A. D. Poularikas, editor, *The Transforms and Applications Handbook*, chapter 12. CRC Press, Boca Raton, FL, 1st edition, 1996.

[105] G. F. Boudreaux-Bartels and T. W. Parks. "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 34(3):442–451, June 1986.

[106] 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 Processing*, 15:193–201, September 1988.

[107] T. Bülow and G. Sommer. "A novel approach to the 2D analytic signal". In F. Solina and A. Leonardis, editors, *Proc. 8th Internat. Conf. on Computer Analysis of Images & Patterns (CAIP'99)*, Ljubljana, Slovenia, 1–3 September 1999, number 1689 in LECTURE NOTES IN COMPUTER SCIENCE, pages 25–32. Springer, 1999.

[108] A. Bultan. "A four-parameter atomic decomposition of chirplets". *IEEE Trans. Signal Processing*, 47(3):731–745, March 1999.

[109] R. Burnett, J. F. Watson, and S. Elder. "The application of modern signal processing techniques for use in rotor fault detection and location within three-phase induction motors". *Signal Processing*, 49(1):57–70, February 1996.

[110] A. B. Carlson. *Communication Systems.* McGraw-Hill, Tokyo, 2nd edition, 1975.

[111] R. Carmona, W. L. Hwang, and B. Torrésani. *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms with an Implementation in S.* Academic Press, San Diego, 1998.

[112] W. G. Carrara, R. S. Goodman, and R. M. Majewski. *Spotlight Synthetic Aperture Radar.* Artech House, Norwood, MA, October 1995.

[113] S. Carstens-Behrens, M. Wagner, and J. F. Böhme. "Detection of multiple resonances in noise". *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, 52(5):285–292, 1998.

[114] P. Celka, B. Boashash, and P. Colditz. "Preprocessing and time-frequency analysis of newborn EEG seizures". *IEEE Engineering in Medicine & Biology Magazine*, 20(5):30–39, September/October 2001.

[115] V. Chandran, S. Elgar, and A. Nguyen. "Detection of mines in acoustic images using higher order spectral features". *IEEE J. of Oceanic Engineering*, 27(3):610–618, July 2002.

[116] E. Chassande-Mottin, F. Auger, and P. Flandrin. "Supervised time-frequency reassignment". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 517–520, Paris, 18–21 June 1996.

[117] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin. "Differential reassignment". *IEEE Signal Processing Letters*, 4(10):293–294, October 1997.

[118] E. Chassande-Mottin and P. Flandrin. "On the time-frequency detection of chirps". *Applied & Computational Harmonic Analysis*, 6(2):252–281, March 1999.

[119] V. C. Chen and H. Ling. *Time-Frequency Transforms for Radar Imaging and Signal Analysis*. Scitech Publishing, 2002.

[120] H.-I. Choi and W. J. Williams. "Improved time-frequency representation of multicomponent signals using exponential kernels". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 37(6):862–871, June 1989.

[121] T. A. C. M. Claasen and W. F. G. Mecklenbräuker. "The Wigner Distribution—A tool for time-frequency signal analysis". *Philips J. of Research*, 35:217–250 (Part 1), 276–300 (Part 2) & 372–389 (Part 3), 1980.

[122] T. A. C. M. Claasen and W. F. G. Mecklenbräuker. "The Wigner Distribution—A tool for time-frequency signal analysis; Part 1: Continuous-time signals". *Philips J. of Research*, 35(3):217–250, 1980.

[123] T. A. C. M. Claasen and W. F. G. Mecklenbräuker. "The Wigner Distribution—A tool for time-frequency signal analysis; Part 3: Relations with other time-frequency signal transformations". *Philips J. of Research*, 35(6):372–389, 1980.

[124] M. Coates and W. J. Fitzgerald. "Time-frequency signal decomposition using energy mixture models". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, volume 2, pages 633–636, Istanbul, 5–9 June 2000.

[125] F. S. Cohen, S. Kadambe, and G. F. Boudreaux-Bartels. "Tracking of unknown nonstationary chirp signals using unsupervised clustering in the Wigner distribution space". *IEEE Trans. Signal Processing*, 41(11):3085–3101, November 1993.

[126] L. Cohen. "Generalized phase-space distribution functions". *J. of Mathematical Physics*, 7(5):781–786, May 1966.

[127] L. Cohen. "Time-frequency distributions—A review". *Proc. IEEE*, 77(7):941–981, July 1989. Invited paper.

[128] L. Cohen. "Distributions concentrated along the instantaneous frequency". In *Proc. SPIE: Advanced Signal-Processing Algorithms, Architectures, and Implementations*, volume 1348, pages 149–157. Soc. of Photo-optical Instrumentation Engineers, San Diego, 10–12 July 1990.

[129] L. Cohen. "Introduction: A primer on time-frequency analysis". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 1, pages 3–42. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[130] L. Cohen. "The scale representation". *IEEE Trans. Signal Processing*, 41(12):3275–3292, December 1993.

[131] L. Cohen. *Time-Frequency Analysis.* Prentice-Hall, Englewood Cliffs, NJ, 1995.

[132] L. Cohen. "Wavelet moments and time-frequency analysis". In *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, volume 3807, pages 434–445. Soc. of Photo-optical Instrumentation Engineers, Denver, CO, 19–21 July 1999.

[133] A. H. Costa and G. F. Boudreax-Bartels. "An overview of aliasing errors in discrete-time formulations of time-frequency distributions". *IEEE Trans. Signal Processing*, 47(5):1463–1474, May 1999.

[134] H. Cox. "Line array performance when the signal coherence is spatially dependent". *J. Acoustical Soc. of America*, 54:1743–1746, July 1973.

[135] G. Cristóbal, J. Bescós, and J. Santamaría. "Image analysis through the Wigner distribution function". *Applied Optics*, 28(2):262–271, January 1989.

[136] G. Cristóbal and J. Hormigo. "Texture segmentation through eigen-analysis of the Pseudo-Wigner distribution". *Pattern Recognition Letters*, 20:337–345, 1999.

[137] G. S. Cunningham and W. J. Williams. "Fast implementations of generalized discrete time-frequency distributions". *IEEE Trans. Signal Processing*, 42(6):1496–1508, June 1994.

[138] G. S. Cunningham and W. J. Williams. "Kernel decomposition of time-frequency distributions". *IEEE Trans. Signal Processing*, 42(6):1425–1442, June 1994.

[139] R. N. Czerwinski and D. L. Jones. "Adaptive cone-kernel time-frequency analysis". *IEEE Trans. Signal Processing*, 43(7):1715–1719, July 1995.

[140] I. Daubechies. "Time-frequency localization operators: A geometric phase space approach". *IEEE Trans. Information Theory*, 34:605–612, July 1988.

[141] I. Daubechies. "The wavelet transform, time-frequency localization and signal analysis". *IEEE Trans. Information Theory*, 36(5):961–1005, September 1990.

[142] I. Daubechies. "The wavelet transform: A method for time-frequency localization". In S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, volume 1, chapter 8, pages 366–417. Prentice-Hall, Englewood Cliffs, NJ, 1991.

[143] M. Davy, C. Doncarli, and G. F. Boudreaux-Bartels. "Improved optimization of time-frequency-based signal classifiers". *IEEE Signal Processing Letters*, 8(2):52–57, February 2001.

[144] M. Davy and A. Doucet. "Copulas: A new insight into positive time-frequency distributions". *IEEE Signal Processing Letters*, 10(7):215–218, July 2003.

[145] N. G. de Bruijn. "A theory of generalized functions, with applications to Wigner distribution and Weyl correspondence". *Nieuw Archief voor Wiskunde (3)*, 21:205–280, 1973.

[146] S. R. Deans. "Radon and Abel transforms". In A. D. Poularikas, editor, *The Transforms and Applications Handbook*, chapter 8. CRC Press, Boca Raton, FL, 1st edition, 1996.

[147] C. S. Detka and A. El-Jaroudi. "The transitory evolutionary spectrum". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, volume 4, pages 289–292, Adelaide, Australia, 19–22 April 1994.

[148] I. Djurović, V. Katkovnik, and L. Stanković. "Median filter based realizations of the robust time-frequency distributions". *Signal Processing*, 81(8):1771–1776, August 2001.

[149] I. Djurović and L. Stanković. "Influence of high noise on the instantaneous frequency estimation using quadratic time-frequency distributions". *IEEE Signal Processing Letters*, 7(11):317–319, November 2000.

[150] I. Djurović and L. Stanković. "Robust Wigner distribution with application to the instantaneous frequency estimation". *IEEE Trans. Signal Processing*, 49(12):2985–2993, December 2001.

[151] I. Djurović, L. Stanković, and J. F. Böhme. "Estimates of the Wigner distribution in Gaussian noise environment". *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, 56(5):337–340, 2002.

[152] I. Djurović and S. Stanković. "Estimation of time-varying velocities of moving objects by time-frequency representations". *IEEE Trans. Image Processing*, 12(5):550–562, May 2003.

[153] R. Du. "Engineering monitoring and diagnosis using wavelet transforms". In C. T. Leondes, editor, *Computer-Aided Design, Engineering, and Manufacturing: Systems Techniques and Applications—Vol. I: Systems Techniques and Computational Methods*, chapter 8. CRC Press, 2000.

[154] P. Duvaut and D. Declerq. "Statistical properties of the pseudo-Wigner-Ville representation of normal random processes". *Signal Processing*, 75(1):93–98, 5 January 1999.

[155] G. Eichmann and N. M. Marinovich. "Scale-invariant Wigner distribution". In *Proc. SPIE: Analog Optical Processing and Computing*, volume 519, pages 18–25. Soc. of Photo-optical Instrumentation Engineers, Cambridge, MA, 25–26 October 1984.

[156] S. El-Khamy, S. E. Shaaban, and E. A. Thabet. "Frequency-hopped multi-user chirp modulation (FH/M-CM) for multipath fading channels". In *Proc. Sixteenth National Radio Science Conference (NRSC'99)*, pages C6/1–8, Ain Shams Univ., Cairo, 23–25 February 1999.

[157] M. K. Emresoy and A. El-Jaroudi. "Iterative instantaneous frequency estimation and adaptive matched spectrogram". *Signal Processing*, 64(2):157–65, January 1998.

[158] S. Farkash and S. Raz. "Linear systems in Gabor time-frequency space". *IEEE Trans. Signal Processing*, 42(3):611–617, March 1994.

[159] H. G. Feichtinger and T. Strohmer, editors. *Gabor Analysis and Algorithms: Theory and Applications*. Birkhäuser, Berlin/Boston, 1998.

[160] H. G. Feichtinger and T. Strohmer. *Advances in Gabor Analysis*. Birkhaüser, 2002.

[161] B. Ferguson. "A ground based narrow-band passive acoustic technique for estimating the altitude and speed of a propeller driven aircraft". *J. Acoustical Soc. of America*, 92(3):1403–1407, September 1992.

[162] B. G. Ferguson. "Time-frequency signal analysis of hydrophone data". *IEEE J. of Oceanic Engineering*, 21(4):537–544, October 1996.

[163] B. G. Ferguson and K. W. Lo. "Transiting aircraft parameter estimation using underwater acoustic sensor data". *IEEE J. of Oceanic Engineering*, 24(4):424–435, October 1999.

[164] B. G. Ferguson and B. G. Quinn. "Application of the short-time Fourier transform and the Wigner-Ville distribution to the acoustic localization of aircraft". *J. Acoustical Soc. of America*, 96:821–827, 1994.

[165] A. C. A. Figueiredo and J. P. S. Bizarro. "Time-frequency images of magnetohydrodynamic phenomena in tokamak plasmas using a discrete-time Wigner distribution". *IEEE Trans. Plasma Science*, 30(1):54–55, February 2002.

[166] P. Flandrin. "Some features of time-frequency representations of multicomponent signals". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'84)*, volume 3, pages 41B.4.1–41B.4.4, San Diego, 19–21 March 1984.

[167] P. Flandrin. "A time-frequency formulation of optimum detection". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 36(9):1377–1384, September 1988.

[168] P. Flandrin. "Time-dependent spectra for nonstationary stochastic processes". In G. Longo and B. Picinbono, editors, *Time and Frequency Representation of Signals and Systems*, pages 69–124. Springer, Vienna, 1989.

[169] P. Flandrin. *Time-Frequency/Time-Scale Analysis.* Academic Press, San Diego, 1999. Original French edition: *Temps-fréquence* (Paris: Hermès, 1993).

[170] P. Flandrin, R. G. Baraniuk, and O. Michel. "Time-frequency complexity and information". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, volume 3, pages 329–332, Adelaide, Australia, 19–22 April 1994.

[171] P. Flandrin, E. Chassande-Mottin, and P. Abry. "Reassigned scalograms and their fast algorithms". In *Proc. SPIE: Wavelet Applications in Signal and Image Processing III*, volume 2569, pages 152–158. Soc. of Photo-optical Instrumentation Engineers, San Diego, CA, 12–14 July 1995.

[172] P. Flandrin and P. Gonçalvès. "Geometry of affine time-frequency distributions". *Applied & Computational Harmonic Analysis*, 3:10–39, January 1996.

[173] P. Flandrin and W. Martin. "The Wigner-Ville spectrum of nonstationary random signals". In W. Mecklenbräuker and F. Hlawatsch, editors, *The Wigner Distribution — Theory and Applications in Signal Processing*, pages 211–267. Elsevier, Amsterdam, 1997.

[174] G. B. Folland. *Harmonic Analysis in Phase Space.* Number 122 in ANNALS OF MATHEMATICS STUDIES. Princeton University Press, Princeton, NJ, 1989.

[175] B. D. Forrester. "Time-frequency analysis in machine fault detection". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 18, pages 406–423. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[176] W. L. J. Fox, J. C. Luby, J. W. Pitton, P. J. Loughlin, and L. E. Atlas. "Sonar and radar range-Doppler processing using a cone-shaped kernel time-frequency representation". In *Proc. 24th Asilomar Conf. on Signals, Systems, and Computers*, volume 2, pages 1079–1083, Pacific Grove, CA, 5–7 November 1990.

[177] G. J. Frazer and B. Boashash. "Multiple window spectrogram and time-frequency distributions". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, volume IV, pages 293–296, Adelaide, Australia, 19–22 April 1994.

[178] D. Friedman. "Instantaneous frequency distribution vs. time: An interpretation of the phase structure of speech". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'85)*, pages 1121–1124, Tampa, FL, 26–29 March 1985.

[179] D. Gabor. "Theory of communication". *J. IEE*, 93(III)(26):429–457, November 1946.

[180] R. G. Gallager. *Information Theory and Reliable Communication.* Wiley, New York, 1968.

[181] M. A. García-Pérez and V. Sierra-Vázquez. "Visual processing in the joint spatial/spatial-frequency domain". In E. Peli, editor, *Vision Models for Target Detection and Recognition: In Memory of Arthur Menendez*, chapter 2, pages 16–62. World Scientific Publishing, 1995.

[182] W. A. Gardner, editor. *Cyclostationarity in Communications and Signal Processing.* IEEE Press, Piscataway, NJ, 1995.

[183] N. L. Gerr. "Introducing a third-order Wigner distribution". *Proc. IEEE*, 76(3):290–292, March 1988.

[184] A. B. Gershman and M. G. Amin. "Wideband direction-of-arrival estimation of multiple chirp signals using spatial time-frequency distributions". *IEEE Signal Processing Letters*, 7(6):152–155, June 2000.

[185] A. B. Gershman, L. Stanković, and V. Katkovnik. "Sensor array signal tracking using a data-driven window approach". *Signal Processing*, 80(12):2507–2515, December 2000.

[186] A. Gersho and R. M. Gray. *Vector Quantization and Signal Compression.* Kluwer, 1991.

[187] S. Ghaemmaghami, M. Deriche, and B. Boashash. "Hierarchical approach to formant detection and tracking through instantaneous frequency estimation". *IEE Electronics Letters*, 33(1):17–18, January 1997.

[188] S. C. Glinski. "Diphone speech synthesis based on a pitch-adaptive short-time Fourier transform". Master's thesis, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, 1981.

[189] G. H. Golub and C. F. Van Loan. *Matrix computations.* Johns Hopkins University Press, Baltimore, MD, 1989.

[190] P. Gonçalvès and R. G. Baraniuk. "Pseudo affine Wigner distributions: Definition and kernel formulation". *IEEE Trans. Signal Processing*, 46(6):1505–1517, June 1998.

[191] E. Grall-Maes and P. Beauseroy. "Mutual information-based feature extraction on the time-frequency plane". *IEEE Trans. Signal Processing*, 50(4):779–790, April 2002.

[192] S. Gu, J. Ni, and J. Yuan. "Non-stationary signal analysis and transient machining process condition monitoring". *Internat. J. Machine Tools & Manufacture*, 42(1):41–51, January 2002.

[193] P. Guillemain and R. Kronland-Martinet. "Horizontal and vertical ridges associated to continuous wavelet transforms". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 63–66, Victoria, BC, 4–6 October 1992.

[194] C. Gupta and A. Papandreou-Suppappola. "Wireless CDMA communications using time-varying signals". In *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA '01)*, volume 1, pages 242–245, Kuala Lumpur, 13–16 August 2001.

[195] S. L. Hahn. "Multidimensional complex signals with single-orthant spectra". *Proc. IEEE*, 80(8):1287–1300, August 1992.

[196] S. L. Hahn and K. M. Snopek. "Double-dimensional distributions: Another approach to "quartic" distributions". *IEEE Trans. Signal Processing*, 50(12):2987–2997, December 2002.

[197] A. Hanssen and L. L. Scharf. "A theory of polyspectra for nonstationary stochastic processes". *IEEE Trans. Signal Processing*, 51(5):1243–1252, May 2003.

[198] H. Hassanpour, M. Mesbah, and B. Boashash. "Comparative performance of time-frequency based newborn EEG seizure detection using spike signatures". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'03)*, volume 2, pages 389–392, Hong Kong, scheduled 6–10 April 2003.

[199] S. B. Hearon and M. G. Amin. "Minimum-variance time-frequency distribution kernels". *IEEE Trans. Signal Processing*, 43(5):1258–1262, May 1995.

[200] C. E. Heil and D. F. Walnut. "Continuous and discrete wavelet transforms". *SIAM Review*, 31(4):628–666, December 1989.

[201] C. W. Helstrom. "An expansion of a signal in Gaussian elementary signals". *IEEE Trans. Information Theory*, 12:81–82, January 1966.

[202] C. W. Helstrom. *The Statistical Theory of Signal Detection.* Oxford University Press, New York, 1968.

[203] F. B. Hildebrand. *Advanced Calculus for Engineers.* Prentice-Hall, New York, 1949.

[204] J. Hilgevoord and J. Uffink. "The mathematical expression of the uncertainty principle". In A. van der Merwe, G. Tarozzi, and F. Selleri, editors, *Microphysical Reality and Quantum Formalism*, number 25–26 in FUNDAMENTAL THEORIES OF PHYSICS, pages 91–114. Kluwer, Dordrecht, 1988. Proc. of the Conf. at Urbino, Italy, 25 Sep. to 3 Oct., 1985.

[205] F. Hlawatsch. "Interference terms in the Wigner distribution". In *Proc. Internat. Conf. on Digital Signal Processing*, pages 363–367, Florence, Italy, 5–8 September 1984.

[206] 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.

[207] F. Hlawatsch and H. Bölcskei. "Unified theory of displacement-covariant time-frequency analysis". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 524–527, Philadelphia, PA, 25–28 October 1994.

[208] F. Hlawatsch and G. F. Boudreaux-Bartels. "Linear and quadratic time-frequency signal representations". *IEEE Signal Processing Magazine*, 9(2):21–67, April 1992.

[209] F. Hlawatsch and P. Flandrin. "The interference structure of the Wigner distribution and related time-frequency signal representations". In W. Mecklenbräuker and F. Hlawatsch, editors, *The Wigner Distribution —Theory and Applications in Signal Processing*, pages 59–133. Elsevier, Amsterdam, 1997.

[210] F. Hlawatsch and G. Matz. "Quadratic time-frequency analysis of linear time-varying systems". In L. Debnath, editor, *Wavelet Transforms and Time-Frequency Signal Analysis*, chapter 9, pages 235–287. Birkhäuser, Boston, 2001.

[211] 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 Processing*, 48(5):1417–1432, May 2000.

[212] F. Hlawatsch, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels. "The hyperbolic class of quadratic time-frequency representations—Part II: Subclasses, intersection with the affine and power classes, regularity, and unitarity". *IEEE Trans. Signal Processing*, 45(2):303–315, February 1997.

[213] 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 Processing*, 47(11):3067–3083, November 1999.

[214] F. Hlawatsch, G. Tauböck, and T. Twaroch. "Covariant time-frequency analysis". In L. Debnath, editor, *Wavelets and Signal Processing*. Birkhäuser, Boston, 2003.

[215] F. Hlawatsch and T. Twaroch. "Covariant (α, β), time-frequency, and (a, b) representations". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 437–440, Paris, 18–21 June 1996.

[216] F. Hlawatsch and R. L. Urbanke. "Bilinear time-frequency representations of signals: The shift-scale invariant class". *IEEE Trans. Signal Processing*, 42:357–366, 1994.

[217] J. Hormigo and G. Cristóbal. "High resolution spectral analysis of images using the pseudo-Wigner distribution". *IEEE Trans. Signal Processing*, 46(6):1757–1763, June 1998.

[218] C. Hory, N. Martin, and A. Chehikian. "Spectrogram segmentation by means of statistical features for non-stationary signal interpretation". *IEEE Trans. Signal Processing*, 50(12):2915–2925, December 2002.

[219] Y. Hu, K. D. Luk, W. W. Lu, and J. C. Leong. "Comparison of time-frequency analysis techniques in intraoperative somatosensory evoked potential (SEP) monitoring". *Computers in Biology & Medicine*, 32(1):13–23, January 2002.

[220] P. J. Huber. *Robust statistics*. Wiley, 1981.

[221] Z. M. Hussain and B. Boashash. "Adaptive instantaneous frequency estimation of multicomponent FM signals". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, volume II, pages 657–660, Istanbul, 5–9 June 2000.

[222] Z. M. Hussain and B. Boashash. "Multi-component IF estimation". In *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pages 559–563, Pocono Manor, PA, 14–16 August 2000.

[223] Z. M. Hussain and B. Boashash. "Design of time-frequency distributions for amplitude and IF estimation of multicomponent signals". In *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, volume 1, pages 339–342, Kuala Lumpur, 13–16 August 2001.

[224] Z. M. Hussain and B. Boashash. "Adaptive instantaneous frequency estimation of multicomponent FM signals using quadratic time-frequency distributions". *IEEE Trans. Signal Processing*, 50(8):1866–1876, August 2002.

[225] 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 Processing*, 46(5):1202–1207, May 1998.

[226] B. G. Iem, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels. "New concepts in narrowband and wideband Weyl correspondence time-frequency techniques". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98)*, volume 3, pages 1573–1576, Seattle, 12–15 May 1998.

[227] B. G. Iem, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels. "A wideband time-frequency Weyl symbol and its generalization". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 29–32, Pittsburgh, PA, 6–9 October 1998.

[228] B. G. Iem, A. Papandreou-Suppappola, and G. F. Boudreaux-Bartels. "Classes of smoothed Weyl symbols". *IEEE Signal Processing Letters*, 7(7):186–188, July 2000.

[229] J. Imberger and B. Boashash. "Application of the Wigner-Ville distribution to temperature gradient microstructure: A new technique to study small-scale variations". *J. of Physical Oceanography*, 16:1997–2012, December 1986.

[230] L. Jacobson and H. Wechsler. "Joint spatial/spatial-frequency representation". *Signal Processing*, 14:37–68, 1988.

[231] J. S. Jaffe, G. Chandran, and E. Reuss. "High frequency acoustic imaging in the ocean". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'95)*, volume 5, pages 2793–2796, Detroit, 9–12 May 1995.

[232] A. Jakobsson, S. L. Marple Jr., and P. Stoica. "Computationally efficient two-dimensional Capon spectrum analysis". *IEEE Trans. Signal Processing*, 48(9):2651–2661, September 2000.

[233] J. Jeong and W. J. Williams. "Alias-free generalised discrete-time time-frequency distributions". *IEEE Trans. Signal Processing*, 40(11):2757–2765, November 1992.

[234] J. Jeong and W. J. Williams. "Kernel design for reduced interference distributions". *IEEE Trans. Signal Processing*, 40(2):402–412, February 1992.

[235] D. L. Jones and R. G. Baraniuk. "A simple scheme for adapting time-frequency representations". *IEEE Trans. Signal Processing*, 42(12):3530–3535, December 1994.

[236] D. L. Jones and R. G. Baraniuk. "Adaptive optimal-kernel time-frequency representation". *IEEE Trans. Signal Processing*, 43(10):2361–2371, October 1995.

[237] D. L. Jones and T. W. Parks. "A high-resolution data-adaptive time-frequency representation". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 38(12):2127–2135, December 1990.

[238] D. L. Jones and T. W. Parks. "A resolution comparison of several time-frequency representations". *IEEE Trans. Signal Processing*, 40(2):413–420, February 1992.

[239] G. Jones and B. Boashash. "Generalized instantaneous parameters and window matching in the time-frequency plane". *IEEE Trans. Signal Processing*, 45(5):1264–1275, May 1997.

[240] S. M. Joshi and J. M. Morris. "Some results on product-function frames". *Signal Processing*, 80(4):737–740, April 2000.

[241] S. Kadambe and T. Adali. "Application of cross-term deleted Wigner representation (CDWR) for sonar target detection/classification". In *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers*, volume 1, pages 822–826, Pacific Grove, CA, 1–4 November 1998.

[242] T. A. Kadous and A. M. Sayeed. "Decentralized multiuser detection for time-varying multipath channels". *IEEE Trans. Communications*, 48:1840–1852, November 2000.

[243] M. H. Kahaei, A. M. Zoubir, B. Boashash, and M. Deriche. "Tracking behaviour of lattice filters for linear and quadratic FM signals". In Wysocki et al. [506], pages 207–214.

[244] S. A. Kassam. *Signal detection in non-Gaussian noise*. Springer, 1988.

[245] S. A. Kassam and H. V. Poor. "Robust techniques for signal processing: A survey". *Proc. IEEE*, 73(3):433–481, March 1985.

[246] V. Katkovnik. "Robust M-periodogram". *IEEE Trans. Signal Processing*, 46(11):3104–3109, November 1998.

[247] V. Katkovnik. "Robust M-estimates of the frequency and amplitude of a complex-valued harmonic". *Signal Processing*, 77(1):71–84, August 1999.

[248] V. Katkovnik and L. Stanković. "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length". *IEEE Trans. Signal Processing*, 46(9):2315–2325, September 1998.

[249] V. Katkovnik and L. Stanković. "Periodogram with varying and data-driven window length". *Signal Processing*, 67(3):345–358, 30 June 1998.

[250] J. Kay and R. Lerner. *Lectures in Communications Theory*. McGraw-Hill, 1961.

[251] A. S. Kayhan and M. G. Amin. "Spatial evolutionary spectrum for DOA estimation and blind signal separation". *IEEE Trans. Signal Processing*, 48(3):791–798, March 2000.

[252] E. J. Kelly and R. P. Wishner. "Matched filter theory for high-velocity, accelerating targets". *IEEE Trans. Military Electronics*, 9:56–59, 1965.

[253] R. S. Kennedy. *Fading dispersive communication channels*. Wiley, New York, 1969.

[254] O. P. Kenny and B. Boashash. "Time-frequency analysis of backscattered signals from diffuse radar targets". *IEE Proc., Part F: Radar & Signal Processing*, 140(3):198–208, June 1993.

[255] H. A. Khan and L. F. Chaparro. "Nonstationary Wiener filtering based on evolutionary spectral theory". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, volume 5, pages 3677–3680, Munich, 21–24 April 1997.

[256] H. Kirchauer, F. Hlawatsch, and W. Kozek. "Time-frequency formulation and design of nonstationary Wiener filters". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'95)*, pages 1549–1552, Detroit, 9–12 May 1995.

[257] J. G. Kirkwood. "Quantum statistics of almost classical ensembles". *Physics Review*, 44:31–37, 1933.

[258] L. Knockaert. "A class of positive isentropic time-frequency distributions". *IEEE Signal Processing Letters*, 9(1):22–25, January 2002.

[259] K. Kodera, C. de Villedary, and R. Gendrin. "A new method for the numerical analysis of nonstationary signals". *Physics of the Earth & Planetary Interiors*, 12:142–150, 1976.

[260] W. Koenig, H. K. Dunn, and L. Y. Lacy. "The sound spectrograph". *J. Acoustical Soc. of America*, 18(1):19–49, 1946.

[261] D. König. "Application of time-frequency analysis for optimum non-equidistant sampling of automotive signals captured at knock". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'96)*, volume 5, pages 2746–2749, Atlanta, GA, 7–10 May 1996.

[262] P. J. Kootsookos, B. C. Lovell, and B. Boashash. "A unified approach to the STFT, TFDs, and instantaneous frequency". *IEEE Trans. Signal Processing*, 40(8):1971–82, August 1992.

[263] W. Kozek. "On the generalized Weyl correspondence and its application to time-frequency analysis of linear time-varying systems". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 167–170, Victoria, BC, 4–6 October 1992.

[264] W. Kozek. "Time-frequency signal processing based on the Wigner-Weyl framework". *Signal Processing*, 29(1):77–92, October 1992.

[265] W. Kozek. "On the transfer function calculus for underspread LTV channels". *IEEE Trans. Signal Processing*, 45(1):219–223, January 1997.

[266] W. Kozek, H. G. Feichtinger, and J. Scharinger. "Matched multiwindow methods for the estimation and filtering of nonstationary processes". In *Proc. IEEE Internat. Symp. on Circuits and Systems (ISCAS 96)*, volume 2, pages 509–512, Atlanta, GA, 12–15 May 1996.

[267] W. Kozek and F. Hlawatsch. "A comparative study of linear and nonlinear time-frequency filters". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 163–166, Victoria, BC, 4–6 October 1992.

[268] W. Kozek, F. Hlawatsch, H. Kirchauer, and U. Trautwein. "Correlative time-frequency analysis and classification of nonstationary random processes". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 417–420, Philadelphia, PA, 25–28 October 1994.

[269] W. Kozek and A. F. Molisch. "On the eigenstructure of underspread WSSUS channels". In *Proc. IEEE-SP Workshop on Signal Processing Advances in Wireless Communications (SPAWC'97)*, pages 325–328, Paris, 16–18 April 1997.

[270] W. Kozek and A. F. Molisch. "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels". *IEEE J. on Selected Areas in Communications*, 16:1579–1589, October 1998.

[271] S. Krishnamachari and W. J. Williams. "Adaptive kernel design in the generalized marginals domain for time-frequency analysis". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'94)*, volume 3, pages 341–344, Adelaide, Australia, 19–22 April 1994.

[272] S. R. Kulkarni, G. Lugosi, and S. S. Venkatesh. "Learning pattern classification—A survey". *IEEE Trans. Information Theory*, 44(6):2178–2206, October 1998.

[273] R. Kumaresan and S. Verma. "On estimating the parameters of chirp signals using rank reduction techniques". In *Proc. 21st Asilomar Conf. on Signals, Systems, and Computers*, pages 555–558, Pacific Grove, CA, 2–4 November 1987.

[274] S. Lach, M. G. Amin, and A. R. Lindsey. "Broadband nonstationary interference excision in spread-spectrum communications using time-frequency synthesis techniques". *IEEE J. on Selected Areas in Communications*, 17(4):704–714, April 1999.

[275] F. Lari and A. Zakhor. "Automatic classification of active sonar data using time-frequency transforms". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 21–24, Victoria, BC, 4-6 October 1992.

[276] S. K. Lee and P. R. White. "Higher-order time-frequency analysis and its application to fault detection in rotating machinery". *J. Mechanical Systems and Signal Processing*, 11(4):637–650, July 1997.

[277] M. J. Levin. "Instantaneous spectra and ambiguity functions". *IEEE Trans. Information Theory*, 10:95–97, January 1964.

[278] A. R. Leyman, Z. M. Kamran, and K. Abed-Meraim. "Higher-order time frequency-based blind source separation technique". *IEEE Signal Processing Letters*, 7(7):193–196, July 2000.

[279] S. Li and D. M. Healy, Jr. "A parametric class of discrete Gabor expansions". *IEEE Trans. Signal Processing*, 44(2):201–211, February 1996.

[280] E. H. Lieb. "Integral bounds for radar ambiguity functions and Wigner distributions". *J. of Mathematical Physics*, 31(3):594–599, March 1990.

[281] J. S. Lim and A. V. Oppenheim. "Enhancement and bandwidth compression of noisy speech". *Proc. IEEE*, 67(12):1586–1604, December 1979.

[282] N. Linh-Trung, A. Belouchrani, K. Abed-Meraim, and B. Boashash. "Separating more sources than sensors using time-frequency distributions". In *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA '01)*, volume 2, pages 583–586, Kuala Lumpur, 13–16 August 2001.

[283] K. W. Lo and B. G. Ferguson. "Passive estimation of aircraft motion parameters using destructive interference between direct and ground-reflected sound waves". In *Proc. Information Decision & Control 99*, pages 171–176. IEEE, Adelaide, Australia, February 1999.

[284] K. W. Lo, S. W. Perry, and B. G. Ferguson. "Aircraft flight parameter estimation using acoustical Lloyd's mirror effect". *IEEE Trans. Aerospace & Electronic Systems*, 38(1):137–151, January 2002.

[285] A. W. Lohmann, D. Mendlovic, and Z. Zalevsky. "Fractional transformations in optics". In E. Wolf, editor, *Progress in Optics—Vol. 38*, chapter 4, pages 263–342. Elsevier, Amsterdam, 1998.

[286] A. W. Lohmann and B. H. Soffer. "Relationships between the Radon-Wigner and fractional Fourier transforms". *J. Optical Soc. of America A*, 11(6):1798–1801, June 1994.

[287] C. T. Lombroso. "Neonatal EEG polygraphy in normal and abnormal newborns". In E. Niedermeyer and F. H. Lopes da Silva, editors, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, pages 803–875. Williams & Wilkins, Baltimore, MD, 3rd edition, 1993.

[288] P. Loughlin and B. Tacer. "On the amplitude and frequency-modulation decomposition of signals". *J. Acoustical Soc. of America*, 100:1594–1601, September 1996.

[289] P. J. Loughlin, J. W. Pitton, and L. E. Atlas. "Bilinear time-frequency representations: New insights and properties". *IEEE Trans. Signal Processing*, 41:750–767, 1993.

[290] J. G. Lourens. "Passive sonar detection of ships with spectro-grams". In *Proc. 3rd South African IEEE Conference on Theoretical and Practical Work in Communications and Signal Processing*, pages 147–151, Johannesburg, June 1990.

[291] B. Lovell, R. C. Williamson, and B. Boashash. "The relationship between instantaneous frequency and time-frequency representations". *IEEE Trans. Signal Processing*, 41(3):1458–1461, March 1993.

[292] N. Ma, D. Vray, P. Delachartre, and G. Gimenez. "Sea-bottom backscattering modeling with a wideband constant beamwidth sonar at normal incidence". In *Proc. IEEE Ultrasonics Symposium*, volume 2, pages 1077–1080, Seattle, 7–10 November 1995.

[293] S. Maes. "The synchrosqueezed representation yields a new reading of the wavelet transform". In *Proc. SPIE: Wavelet Applications II*, volume 2491, pages 532–559. Soc. of Photo-optical Instrumentation Engineers, Orlando, FL, 17–21 April 1995.

[294] I. Magrin-Chagnolleau, G. Durou, and F. Bimbot. "Application of time-frequency principal component analysis to text-independent speaker identification". *IEEE Trans. on Speech & Audio Processing*, 10(6):371–378, September 2002.

[295] S. G. Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, San Diego, 1st edition, 1998.

[296] S. G. Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, San Diego / London, 2nd edition, 1999.

[297] S. G. Mallat and Z. Zhang. "Matching pursuits with time-frequency dictionaries". *IEEE Trans. Signal Processing*, 41(12):3397–3415, December 1993.

[298] H. Margenau and R. N. Hill. "Correlation between measurements in quantum theory". *Progress of Theoretical Physics*, 26:722–738, 1961.

[299] S. L. Marple Jr. *Digital Spectral Analysis with Applications.* Prentice-Hall, Englewood Cliffs, NJ, 1987.

[300] S. L. Marple Jr. "Computing the discrete-time "analytic" signal via FFT". *IEEE Trans. Signal Processing*, 47(9):2600–2603, September 1999.

[301] S. L. Marple Jr. "Two-dimensional lattice linear prediction parameter estimation method and fast algorithm". *IEEE Signal Processing Letters*, 7(6):164–168, June 2000.

[302] S. L. Marple Jr. and T. Brotherton. "Detection and classification of short duration underwater acoustic signals by Prony's method". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'91)*, volume 2, pages 1309–1312, Toronto, 14–17 May 1991.

[303] W. Martin. "Time-frequency analysis of random signals". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'82)*, volume 3, pages 1325–1328, Paris, 3–5 May 1982.

[304] W. Martin and P. Flandrin. "Wigner-Ville spectral analysis of nonstationary processes". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 33(6):1461–1470, December 1985.

[305] G. Matz and F. Hlawatsch. "Time-frequency formulation and design of optimal detectors". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 213–216, Paris, 18–21 June 1996.

[306] G. Matz and F. Hlawatsch. "Time-frequency methods for signal detection with application to the detection of knock in car engines". In *Proc. Ninth IEEE Workshop on Statistical Signal and Array Processing (SSAP-98)*, pages 196–199, Portland, OR, 14–16 September 1998.

[307] 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. of Mathematical Physics*, 39(8):4041–4070, August 1998. Special Issue on Wavelet and Time-Frequency Analysis.

[308] G. Matz and F. Hlawatsch. "Time-varying spectra for underspread and overspread nonstationary processes". In *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers*, pages 282–286, Pacific Grove, CA, 1–4 November 1998.

[309] G. Matz and F. Hlawatsch. "Minimax robust time-frequency filters for nonstationary signal estimation". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'99)*, pages 1333–1336, Phoenix, AZ, 15–19 March 1999.

[310] G. Matz and F. Hlawatsch. "Time-frequency subspace detectors and application to knock detection". *Archiv für Elektronik und Übertragungstechnik (Internat. J. of Electronics & Communications)*, 53(6):379–385, December 1999.

[311] G. Matz and F. Hlawatsch. "Minimax robust nonstationary signal estimation based on a p-point uncertainty model". *J. Franklin Institute*, 337(4):403–419, July 2000.

[312] G. Matz and F. Hlawatsch. "Linear time-frequency filters: On-line algorithms and applications". In A. Papandreou-Suppappola, editor, *Applications in Time-Frequency Signal Processing*, chapter 6, pages 205–271. CRC Press, Boca Raton, FL, 2002.

[313] G. Matz and F. Hlawatsch. "Time-frequency projection filters: Online implementation, subspace tracking, and application to interference excision". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'02)*, pages 1213–1216, Orlando, FL, 13–17 May 2002.

[314] G. Matz, F. Hlawatsch, and W. Kozek. "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes". *IEEE Trans. Signal Processing*, 45(6):1520–1534, June 1997.

[315] G. Matz, F. Hlawatsch, and A. Raidl. "Signal-adaptive robust time-varying Wiener filters: Best subspace selection and statistical analysis". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, pages 3945–3948, Salt Lake City, UT, 7–11 May 2001.

[316] G. Matz and A. Raidl. "Robust detection of nonstationary random signals belonging to p-point uncertainty classes". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'03)*, pages 641–644, Hong Kong, scheduled 6–10 April 2003.

[317] A. C. McBride and F. H. Kerr. "On Namias's fractional Fourier transforms". *IMA J. of Applied Mathematics*, 39(2):159–175, 1987.

[318] M. R. McClure and L. Carin. "Matching pursuits with a wave-based dictionary". *IEEE Trans. Signal Processing*, 45(12):2912–2927, December 1997.

[319] W. Mecklenbräuker and F. Hlawatsch, editors. *The Wigner Distribution —Theory and Applications in Signal Processing*. Elsevier, Amsterdam, 1997.

[320] M. Mesbah and B. Boashash. "Reduced bias time-frequency peak filtering". In *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, volume 1, pages 327–330, Kuala Lumpur, 13–16 August 2001.

[321] M. Mesbah and B. Boashash. "Performance comparison of seizure detection methods using EEG of newborns for implementation of a DSP subsystem". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'02)*, Orlando, FL, 13–17 May 2002. Paper no. 1932.

[322] Y. Meyer. *Wavelets: Algorithms and applications*. Soc. for Industrial and Applied Mathematics, Philadelphia, PA, 1993. Translated and revised by Robert D. Ryan. Original French title: *Ondelettes et algorithems concurrents*.

[323] Z.-H. Michalopoulou. "Underwater transient signal processing: Marine mammal identification, localization, and source signal deconvolution". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, volume 1, pages 503–506, Munich, 21–24 April 1997.

[324] R. K. Mobley. *Vibration Fundamentals*. Newnes, Boston, 1999.

[325] A. Monti, C. Medigue, and L. Mangin. "Instantaneous parameter estimation in cardiovascular time series by harmonic and time-frequency analysis". *IEEE Trans. Biomedical Engineering*, 49(12):1547–1556, December 2002.

[326] M. R. Morelande, B. Barkat, and A. M. Zoubir. "Statistical performance comparison of a parametric and a non-parametric method for IF estimation of random amplitude linear FM signals in additive noise". In *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pages 262–266, Pocono Manor, PA, 14–16 August 2000.

[327] D. R. Morgan and T. M. Smith. "Coherence effects on the detection performance of quadratic array processors with application to large-array matched-field beamforming". *J. Acoustical Soc. of America*, 87(2):737–747, February 1990.

[328] J. M. Morris and Y. Lu. "Generalized Gabor expansions of discrete-time signals in $l^2(\mathbb{Z})$ via biorthogonal-like sequences". *IEEE Trans. Signal Processing*, 44(6):1378–1391, June 1996.

[329] S. H. Nawab and T. F. Quatieri. "Short-time Fourier transform". In J. S. Lim and A. V. Oppenheim, editors, *Advanced Topics in Signal Processing*, chapter 6, pages 289–337. Prentice-Hall, Englewood Cliffs, NJ, 1988.

[330] A. W. Naylor and G. R. Sell. *Linear Operator Theory in Engineering and Science*. Springer, New York, 2nd edition, 1982.

[331] D. E. Newland. "Time-frequency and time-scale analysis by harmonic wavelets". In A. Prochazka, J. Uhlir, P. J. W. Rayner, and N. G. Kingsbury, editors, *Signal Analysis and Prediction*, chapter 1. Birkhäuser, Boston, MA, 1998.

[332] R. M. Nickel, T.-H. Sang, and W. J. Williams. "A new signal adaptive approach to positive time-frequency distributions with suppressed interference terms". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98)*, volume 3, pages 1777–1780, Seattle, 12–15 May 1998.

[333] A. H. Nuttall. "On the quadrature approximation to the Hilbert transform of modulated signals". *Proc. IEEE*, 54:1458–1459, 1966.

[334] A. H. Nuttall. "Efficient evaluation of polynomials and exponentials of polynomials for equispaced arguments". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 35(10):1486–1487, October 1987.

[335] S. C. Olhede and A. T. Walden. "Generalized Morse wavelets". *IEEE Trans. Signal Processing*, 50(11):2661–2670, November 2002.

[336] P. M. Oliveira and V. Barroso. "Uncertainty in the time-frequency plane". In *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pages 607–611, Pocono Manor, PA, 14–16 August 2000.

[337] J. C. O'Neill and W. J. Williams. "Shift-covariant time-frequency distributions of discrete signals". *IEEE Trans. Signal Processing*, 47(1):133–146, January 1999.

[338] P. O'Shea. "An iterative algorithm for estimating the parameters of polynomial phase signals". In *Proc. Fourth Internat. Symp. on Signal Processing and its Applications (ISSPA'96)*, volume 2, pages 730–731, Gold Coast, Australia, 25–30 August 1996.

[339] A. Ouldali and M. Benidir. "Statistical analysis of polynomial phase signals affected by multiplicative and additive noise". *Signal Processing*, 78(1):19–42, October 1999.

[340] H. M. Ozaktas, O. Arıkan, M. A. Kutay, and G. Bozdağı. "Digital computation of the fractional Fourier transform". *IEEE Trans. Signal Processing*, 44(9):2141–2150, September 1996.

[341] H. M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural. "Convolution, filtering, and multiplexing in fractional Fourier domains and their relationship to chirp and wavelet transforms". *J. Optical Soc. of America A*, 11:547–559, February 1994.

[342] H. M. Ozaktas, M. A. Kutay, and D. Mendlovic. "Introduction to the fractional Fourier transform and its applications". In P. W. Hawkes, B. Kazan, and T. Mulvey, editors, *Advances in Imaging & Electron Physics—Vol. 106*, chapter 4, pages 239–291. Academic Press, San Diego, February 1999.

[343] C. H. Page. "Instantaneous power spectra". *J. of Applied Physics*, 23(1):103–106, January 1952.

[344] 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 Processing*, 41(12):3425–3444, December 1993. Special Issue on Wavelets and Signal Processing.

[345] A. Papandreou-Suppappola. "Generalized time-shift covariant quadratic time-frequency representations with arbitrary group delays". In *Proc. 29th Asilomar Conf. on Signals, Systems, and Computers*, pages 553–557, Pacific Grove, CA, October–November 1995.

[346] A. Papandreou-Suppappola, editor. *Applications in Time-Frequency Signal Processing*. CRC Press, Boca Raton, FL, 2002.

[347] A. Papandreou-Suppappola, F. Hlawatsch, and G. F. Boudreaux-Bartels. "Power class time-frequency representations: Interference geometry, smoothing, and implementation". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 193–196, Paris, 18–21 June 1996.

[348] 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 Processing: A Review Journal*, 8(1):3–48, 1998.

[349] A. Papandreou-Suppappola, R. L. Murray, B. G. Iem, and G. F. Boudreaux-Bartels. "Group delay shift covariant quadratic time-frequency representations". *IEEE Trans. Signal Processing*, 49(11):2549–2564, November 2001.

[350] A. Papandreou-Suppappola and S. B. Suppappola. "Adaptive time-frequency representations for multiple structures". In *Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (SSAP-2000)*, pages 579–583, Pocono Manor, PA, 14–16 August 2000.

[351] A. Papandreou-Suppappola and S. B. Suppappola. "Analysis and classification of time-varying signals with multiple time-frequency structures". *IEEE Signal Processing Letters*, 9(3):92–95, March 2002.

[352] A. Papoulis. *Signal analysis*. McGraw-Hill, New York, 1977.

[353] A. Papoulis. "Random modulation: A review". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 31(1):96–105, February 1983.

[354] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York, 3rd edition, 1991.

[355] T. W. Parks and R. G. Shenoy. "Time-frequency concentrated basis functions". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'90)*, pages 2459–2462, Albuquerque, NM, 3–6 April 1990.

[356] L. Parra and C. Spence. "Convolutive blind separation of non-stationary sources". *IEEE Trans. on Speech & Audio Processing*, 8(3):320–327, May 2000.

[357] J. D. Parsons. *The Mobile Radio Propagation Channel.* Pentech Press, London, 1992.

[358] B. A. Paya, I. I. Esat, and M. N. M. Badi. "Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor". *J. Mechanical Systems and Signal Processing*, 11(5):751–765, September 1997.

[359] S. Peleg and B. Friedlander. "Discrete polynomial-phase transform". *IEEE Trans. Signal Processing*, 43(8):1901–1914, August 1995.

[360] S. Peleg and B. Friedlander. "Multicomponent signal analysis using the polynomial-phase transform". *IEEE Trans. Aerospace & Electronic Systems*, 32(1):378–386, January 1996.

[361] S. Peleg and B. Porat. "Estimation and classification of polynomial-phase signals". *IEEE Trans. Information Theory*, 37(2):422–430, March 1991.

[362] S. Peleg and B. Porat. "Linear FM signal parameter estimation from discrete-time observations". *IEEE Trans. Aerospace & Electronic Systems*, 27(4):607–616, July 1991.

[363] A. Persson, T. Ottosson, and E. Strom. "Time-frequency localized CDMA for downlink multi-carrier systems". In *Proc. Seventh Internat. Symp. on Spread Spectrum Techniques and Applications (ISSSTA-02)*, volume 1, pages 118–122, Prague, 2–5 September 2002.

[364] B. Picinbono. "On Instantaneous Amplitude and Phase of Signals". *IEEE Trans. Signal Processing*, 45(3):552–560, March 1997.

[365] B. Picinbono and W. Martin. "Représentation des signaux par amplitude et phase instantanées". *Ann. Télécommunications*, 38:179–190, 1983.

[366] A. G. Piersol. "Power spectra measurements for space vibration data". *J. Spacecraft and Rockets*, 4:1613, December 1967.

[367] S. Pon Varma, A. Papandreou-Suppappola, and S. B. Suppappola. "Detecting faults in structures using time-frequency techniques". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, volume 6, pages 3593–3596, Salt Lake City, UT, 7–11 May 2001.

[368] S. Pon Varma, A. Papandreou-Suppappola, and S. B. Suppappola. "Matching pursuit classification for time-varying acoustic emissions". In *Proc. 35th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, 4–7 November 2001. Paper TA2-3.

[369] H. V. Poor. *An Introduction to Signal Detection and Estimation.* Springer, New York, 1988.

[370] B. Porat and B. Friedlander. "Asymptotic statistical analysis of the high-order ambiguity function for parameter estimation of polynomial phase signal". *IEEE Trans. Information Theory*, 42(3):995–1001, May 1996.

[371] M. R. Portnoff. "Time-frequency representation of digital signals and systems based on short-time Fourier analysis". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 28(1):55–69, February 1980.

[372] S. Prabhakar, A. S. Sekhar, and A. R. Mohanty. "Detection and monitoring of cracks in a rotor-bearing system using wavelet transforms". *J. Mechanical Systems and Signal Processing*, 15(2):447–450, March 2001.

[373] R. Price and E. M. Hofstetter. "Bounds on the volume and height distributions of the ambiguity function". *IEEE Trans. Information Theory*, 11:207–214, 1965.

[374] M. B. Priestley. *Spectral Analysis and Time Series—Part II.* Academic Press, London, 1981.

[375] M. B. Priestly. "Evolutionary spectra and non-stationary processes". *J. Royal Statistical Soc. (Series B)*, 27(2):204–237, 1965.

[376] J. G. Proakis. *Digital Communications.* McGraw-Hill, New York, 3rd edition, 1995.

[377] G. R. Putland and B. Boashash. "Can a signal be both monocomponent and multicomponent?". In *Third Australasian Workshop on Signal Processing Applications (WoSPA 2000)*, Brisbane, Australia, 14–15 December 2000. Paper no. 32.

[378] S. Qian. *Introduction to Time-frequency and Wavelet Transforms.* Prentice-Hall, Englewood Cliffs, NJ, 2002.

[379] S. Qian and D. Chen. "Decomposition of the Wigner distribution and time-frequency distribution series". *IEEE Trans. Signal Processing*, 42(10):2836–2842, October 1994.

[380] S. Qian and D. Chen. *Joint Time-Frequency Analysis: Methods & Applications.* Prentice-Hall, Upper Saddle River, NJ, 1996.

[381] S. Qian and D. Chen. "Joint time-frequency analysis". *IEEE Signal Processing Magazine*, 16(2):52–65, March 1999.

[382] S. Qian, Y. Rao, and D. Chen. "A fast Gabor spectrogram". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, volume 2, pages 653–656, Istanbul, 5–9 June 2000.

[383] R. S. Ramineni, M. G. Amin, and A. R. Lindsey. "Performance analysis of subspace projection techniques for interference excision in DSSS communications". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, volume 5, pages 2825–2828, Istanbul, 5–9 June 2000.

[384] L. Rankine and M. Mesbah. "Significant atom determination of basis pursuit decomposition". In *Proc. Seventh Internat. Symp. on Signal Processing and its Applications (ISSPA '03)*, pages 577–580, Paris, 1–4 July 2003.

[385] A. M. Rao and D. L. Jones. "Efficient structures for quadratic time-frequency and time-scale array processors". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 397–400, Pittsburgh, PA, 6–9 October 1998.

[386] A. M. Rao and D. L. Jones. "Nonstationary array signal detection using time-frequency and time-scale representations". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98)*, volume 4, pages 1989–1992, Seattle, 12–15 May 1998.

[387] A. M. Rao and D. L. Jones. "A denoising approach to multisensor signal estimation". *IEEE Trans. Signal Processing*, 48(5):1225–1234, May 2000.

[388] T. S. Rappaport. *Wireless Communications: Principles & Practice.* Prentice-Hall, Upper Saddle River, NJ, 1996.

[389] T. R. Reed and H. Wechsler. "Segmentation of textured images and Gestalt organization using spatial/spatial-frequency representations". *IEEE Trans. Pattern Analysis & Machine Intelligence*, 12(1):1–12, January 1990.

[390] D. C. Reid, A. M. Zoubir, and B. Boashash. "Aircraft flight parameter estimation based on passive acoustic techniques using the polynomial Wigner-Ville distribution". *J. Acoustical Soc. of America*, 102(1):207–23, July 1997.

[391] A. Reilly, G. Frazer, and B. Boashash. "Analytic signal generation—Tips and traps". *IEEE Trans. Signal Processing*, 42(11):3241–3245, November 1994.

[392] G. Reina and B. Porat. "Comparative performance analysis of two algorithms for instantaneous frequency estimation". In *Proc. Eighth IEEE Workshop on Statistical Signal and Array Processing (SSAP-96)*, pages 448–451, Corfu, Greece, 24–26 June 1996.

[393] C. Richard. "Time-frequency-based detection using discrete-time discrete-frequency Wigner distributions". *IEEE Trans. Signal Processing*, 50(9):2170–2176, September 2002.

[394] C. Richard and R. Lengellé. "Data-driven design and complexity control of time-frequency detectors". *Signal Processing*, 77(1):37–48, August 1999.

[395] M. S. Richman, T. W. Parks, and R. G. Shenoy. "Discrete-time, discrete-frequency time-frequency analysis". *IEEE Trans. Signal Processing*, 46(6):1517–1527, June 1998.

[396] D. C. Rife and R. R. Boorstyn. "Single-tone parameter estimation from discrete-time observations". *IEEE Trans. Information Theory*, 20(5):591–598, September 1974.

[397] A. W. Rihaczek. "Signal energy distribution in time and frequency". *IEEE Trans. Information Theory*, 14(3):369–374, May 1968.

[398] A. W. Rihaczek. *Principles of high-resolution radar.* McGraw-Hill, New York, 1969. Reprinted Los Altos, CA: Peninsula Publishing, 1985.

[399] M. D. Riley. *Speech Time-Frequency Representations.* Kluwer, 1989.

[400] O. Rioul and P. Flandrin. "Time-scale energy distributions: A general class extending wavelet transforms". *IEEE Trans. Signal Processing*, 40(7):1746–1757, July 1992.

[401] B. Ristic and B. Boashash. "Kernel design for time-frequency signal analysis using the Radon transform". *IEEE Trans. Signal Processing*, 41(5):1996–2008, May 1993.

[402] B. Ristic and B. Boashash. "Relationship between the polynomial and higher order Wigner-Ville distribution". *IEEE Signal Processing Letters*, 2(12):227–229, December 1995.

[403] B. Ristic and B. Boashash. "Instantaneous frequency estimation of quadratic and cubic polynomial FM signals using the cross polynomial Wigner-Ville distribution". *IEEE Trans. Signal Processing*, 44:1549–1553, 1996.

[404] G. Roberts, A. M. Zoubir, and B. Boashash. "Non-stationary, narrowband Gaussian signal discrimination in time-frequency space". In Wysocki et al. [506], chapter 18, pages 159–166.

[405] T. D. Rossing. *The Science of Sound.* Addison-Wesley, Reading, MA, 2nd edition, 1990.

[406] W. Rudin. *Real and complex analysis.* McGraw-Hill, New York, 1987.

[407] B. Samimy and G. Rizzoni. "Mechanical signature analysis using time-frequency signal processing: Application to internal combustion engine knock detection". *Proc. IEEE*, 84(9):1130–1343, September 1996.

[408] T.-H. Sang and W. J. Williams. "Rényi information and signal-dependent optimal kernel design". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'95)*, volume 2, pages 997–1000, Detroit, 9–12 May 1995.

[409] N. Sang-Won and E. J. Powers. "Volterra series representation of time-frequency distributions". *IEEE Trans. Signal Processing*, 51(6):1532–1537, July 2003.

[410] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann. "Power quality assessment via wavelet transform analysis". *IEEE Trans. Power Delivery*, 11(2):924–930, April 1996.

[411] G. J. Saulnier, M. J. Medley, and P. K. Das. "Wavelets and filter banks in spread spectrum communication systems". In A. N. Akansu and M. J. T. Smith, editors, *Subband and Wavelet Transforms: Design and Applications*, chapter 10, pages 309–346. Kluwer, Norwell, MA, 1996.

[412] A. M. Sayeed and B. Aazhang. "Joint multipath-Doppler diversity in mobile wireless communications". *IEEE Trans. Communications*, 47:123–132, January 1999.

[413] A. M. Sayeed and D. J. Jones. "Optimal detection using bilinear time-frequency and time-scale representations". *IEEE Trans. Signal Processing*, 43(12):2872–2883, December 1995.

[414] A. M. Sayeed and D. L. Jones. "A canonical covariance-based method for generalized joint signal representations". *IEEE Signal Processing Letters*, 3(4):121–123, April 1996.

[415] A. M. Sayeed and D. L. Jones. "Integral transforms covariant to unitary operators and their implications for joint signal representations". *IEEE Trans. Signal Processing*, 44(6):1365–1377, June 1996.

[416] A. M. Sayeed and D. L. Jones. "Optimum quadratic detection and estimation using generalized joint signal representations". *IEEE Trans. Signal Processing*, 44(12):3031–3043, December 1996.

[417] A. M. Sayeed, A. Sendonaris, and B. Aazhang. "Multiuser detection in fast fading multipath environments". *IEEE J. on Selected Areas in Communications*, 16:1691–1701, December 1998.

[418] L. L. Scharf and B. Friedlander. "Toeplitz and Hankel kernels for estimating time-varying spectra of discrete-time random processes". *IEEE Trans. Signal Processing*, 49(1):179–189, January 2001.

[419] M. O. Scully and L. Cohen. "Quasi-probability distributions for arbitrary operators". In Y. S. Kim and W. W. Zachary, editors, *The Physics of Phase Space: Nonlinear Dynamics and Chaos, Geometric Quantization, and Wigner Functions*, number 278 in LECTURE NOTES IN PHYSICS, pages 253–263. Springer, Berlin, 1987. Proc. First Internat. Conf. on the Physics of Phase Space, University of Maryland, College Park, MD, 20–23 May 1986.

[420] J. P. Sessarego, J. Sageloli, P. Flandrin, and M. Zakharia. "Time-frequency Wigner-Ville analysis of echoes scattered by a spherical shell". In J.-M. Combes, A. Grossmann, and P. Tchamitchian, editors, *Wavelets: Time-frequency Methods and Phase Space*, pages 147–153. Springer, 1989. Proc. of the Internat. Conf., Marseille, 14–18 December, 1987. 315pp.

[421] A. Z. Sha'ameri, B. Boashash, and I. Ismail. "Design of signal dependent kernel functions for digital modulated signals". In *Proc. Fourth Internat. Symp. on Signal Processing and its Applications (ISSPA'96)*, volume 2, pages 527–528, Gold Coast, Australia, 25–30 August 1996.

[422] R. G. Shenoy and T. W. Parks. "The Weyl correspondence and time-frequency analysis". *IEEE Trans. Signal Processing*, 42(2):318–331, February 1994.

[423] Y. Shin, A. C. Parsons, E. J. Powers, and W. M. Grady. "Time-frequency analysis of power system disturbance signals for power quality". In *Proc. IEEE Power Engineering Soc. Summer Meeting*, volume 1, pages 402–407, Edmonton, AL (Canada), 18–22 July 1999.

[424] Y. Shin, E. J. Powers, W. M. Grady, and S. C. Bhatt. "Effects of dispersion on disturbance propagation on high voltage transmission lines". In *Proc. IEEE Power Engineering Soc. Summer Meeting*, volume 2, pages 851–854, Seattle, WA (USA), 16–20 July 2000.

[425] W. M. Siebert. "Studies of Woodward's uncertainty function". *Quarterly Progress Report* (MIT Electronics Research Lab, Cambridge, MA), pages 90–94, 1958.

[426] J. A. Sills and E. W. Kamen. "Time-varying matched filters". *Circuits, Systems, & Signal Processing*, 15(5):609–630, 1996.

[427] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt. *Spread Spectrum Communications* (3 vols.). Computer Science Press, Rockville, MD, 1985.

[428] L. Sirovich and B. W. Knight. "On the eigentheory of operators which exhibit a slow variation". *Quarterly of Applied Mathematics*, 38:469–488, 1980.

[429] D. Slepian. "On bandwidth". *Proc. IEEE*, 64(3):292–300, March 1976.

[430] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum". *Physical Review Letters*, 70:1244–1247, 1993.

[431] T. Södertröm and P. Stoica. *System Identification.* Prentice-Hall, Englewood Cliffs. NJ, 1989.

[432] K. A. Sostrand. "Mathematics of the time-varying channel". In *Proc. NATO Advanced Study Inst. on Signal Processing with Emphasis on Underwater Acoustics*, volume 2, pages 25.1–25.20, 1968.

[433] L. Stanković. "An analysis of some time-frequency and time-scale distributions". *Ann. Télécommunications*, 49(9/10):505–517, September/October 1994.

[434] L. Stanković. "An analysis of Wigner higher order spectra of multicomponent signals". *Ann. Télécommunications*, 49(3/4):132–136, March/April 1994.

[435] L. Stanković. "A method for time-frequency analysis". *IEEE Trans. Signal Processing*, 42(1):225–229, January 1994.

[436] L. Stanković. "Multitime definition of the Wigner higher order distribution: L-Wigner distribution". *IEEE Signal Processing Letters*, 1(7):106–109, July 1994.

[437] L. Stanković. "A method for improved distribution concentration in the time-frequency analysis of the multicomponent signals using the L-Wigner distribution". *IEEE Trans. Signal Processing*, 43(5):1262–1268, May 1995.

[438] L. Stanković. "Auto-term representation by the reduced interference distributions: A procedure for kernel design". *IEEE Trans. Signal Processing*, 44(6):1557–1563, June 1996.

[439] L. Stanković. "Highly concentrated time-frequency distributions: Pseudo-quantum signal representation". *IEEE Trans. Signal Processing*, 45(3):543–551, March 1997.

[440] L. Stanković. "On the time-frequency analysis based filtering". *Ann. Télécommunications*, 55(5/6):216–225, May/June 2000.

[441] L. Stanković. "A measure of some time-frequency distributions concentration". *Signal Processing*, 81(3):621–631, March 2001.

[442] L. Stanković. "Analysis of noise in time-frequency distributions". *IEEE Signal Processing Letters*, 9(9):286–289, September 2002.

[443] L. Stanković. "Time-frequency distributions with complex argument". *IEEE Trans. Signal Processing*, 50(3):475–486, March 2002.

[444] L. Stanković, T. Alieva, and M. Bastiaans. "Fractional-Fourier-domain weighted Wigner distribution". In *Proc. Eleventh IEEE Workshop on Statistical Signal Processing*, pages 321–324, Singapore, 6–8 August 2001.

[445] L. Stanković and J. F. Böhme. "Time-frequency analysis of multiple resonances in combustion engine signals". *Signal Processing*, 79(1):15–28, November 1999.

[446] L. Stanković and V. Ivanović. "Further results on the minimum variance time-frequency distributions kernels". *IEEE Trans. Signal Processing*, 45(6):1650–1655, June 1997.

[447] L. Stanković and V. Katkovnik. "The Wigner distribution of noisy signals with adaptive time-frequency varying window". *IEEE Trans. Signal Processing*, 47(4):1099–1108, April 1999.

[448] L. Stanković and V. Katkovnik. "Instantaneous frequency estimation using the higher order L-Wigner distributions with the data driven order and window length". *IEEE Trans. Information Theory*, 46(1):302–311, January 2000.

[449] L. Stanković and S. Stanković. "On the Wigner distribution of the discrete-time noisy signals with application to the study of quantization effects". *IEEE Trans. Signal Processing*, 42(7):1863–1867, July 1994.

[450] L. Stanković and S. Stanković. "An analysis of instantaneous frequency representation using time-frequency distributions—Generalized Wigner distribution". *IEEE Trans. Signal Processing*, 43(2):549–552, February 1995.

[451] L. Stanković, S. Stanković, and I. Djurović. "Architecture for realization of the cross-terms free polynomial Wigner-Ville distribution". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'97)*, volume III, pages 2053–2056, Munich, 21–24 April 1997.

[452] L. Stanković, S. Stanković, and I. Djurović. "Space/spatial frequency based filtering". *IEEE Trans. Signal Processing*, 48(8):2343–2352, August 2000.

[453] S. Stanković. "About time-variant filtering of speech signals with time-frequency distributions for hands-free telephone systems". *Signal Processing*, 80(9):1777–1785, September 2000.

[454] S. Stanković and L. Stanković. "Approach to the polynomial Wigner distributions". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 153–156, Paris, 18–21 June 1996.

[455] S. Stanković and L. Stanković. "Introducing time-frequency distribution with a 'complex-time' argument". *Electronics Letters*, 32(14):1265–1267, July 1996.

[456] S. Stanković, L. Stanković, and Z. Uskoković. "On the local frequency, group shift, and cross-terms in some multidimensional time-frequency distributions: A method for multidimensional time-frequency analysis". *IEEE Trans. Signal Processing*, 43(7):1719–1725, July 1995.

[457] S. Stanković and J. Tilp. "Time-varying filtering of speech signals using linear prediction". *Electronics Letters*, 36(8):763–764, April 2000.

[458] W. J. Staszewski, K. Worden, and G. R. Tomlinson. "Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition". *J. Mechanical Systems and Signal Processing*, 11(5):673–692, September 1997.

[459] N. Stevenson, E. Palmer, J. Smeathers, and B. Boashash. "The *BT* product as a signal dependent sample size estimate in hypothesis testing: An application to linear/nonlinear discrimination in bandwidth limited systems". In *Proc. Seventh Internat. Symp. on Signal Processing and its Applications (ISSPA'03)*, pages 551–554, Paris, 1–4 July 2003.

[460] L. R. O. Storey. "An investigation of whistling atmospherics". *Phil. Trans. Roy. Soc.*, A246:113–141, 1953.

[461] G. Strang and T. Q. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, MA, 1996.

[462] V. Sucic and B. Boashash. "Optimisation algorithm for selecting quadratic time-frequency distributions: Performance results and calibration". In *Proc. Sixth Internat. Symp. on Signal Processing and its Applications (ISSPA'01)*, volume 1, pages 331–334, Kuala Lumpur, 13–16 August 2001.

[463] V. Sucic and B. Boashash. "Parameter selection for optimising time-frequency distributions and measurements of time-frequency characteristics of non-stationary signals". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'01)*, volume 6, pages 3557–3560, Salt Lake City, UT, 7–11 May 2001.

[464] V. Sucic and B. Boashash. "Selecting the optimal time-frequency distribution for real-life multicomponent signals under given constraints". In *Proc. Eleventh European Signal Processing Conf. (EUSIPCO-02)*, volume 1, pages 141–144, Toulouse, 3–6 September 2002.

[465] V. Sucic and B. Boashash. "An approach for selecting a real-life signal best-performing time-frequency distribution". In *Proc. Seventh Internat. Symp. on Signal Processing and its Applications (ISSPA'03)*, volume 1, pages 100–104, Paris, 1–4 July 2003.

[466] V. Sucic, B. Boashash, and K. Abed-Meraim. "A normalised performance measure for quadratic time-frequency distributions". In *Proc. Second IEEE Internat. Symp. on Signal Processing and Information Technology (ISSPIT'02)*, pages 463–466, Marrakech, Morocco, 18–21 December 2002.

[467] S. M. Sussman. "Least-squares synthesis of radar ambiguity functions". *IRE Trans. Information Theory*, 8:246–254, April 1962.

[468] H. Suzuki and F. Kobayashi. "A method of two-dimensional spectral analysis using the Wigner distribution". *Electronics & Communications in Japan, Part III: Fundamental Electronic Science*, 75(1):1006–1013, January 1992.

[469] H. H. Szu. "Two-dimensional optical processing of one-dimensional acoustic data". *Optical Engineering*, 21(5):804–813, September–October 1982.

[470] M. T. Taner, F. Koehler, and R. E. Sheriff. "Complex seismic trace analysis". *Geophysics*, 44(6):1041–1063, June 1979.

[471] S. Thangavelu. *Lectures on Hermite and Laguerre Expansions*. Princeton University Press, Princeton, NJ, 1993.

[472] S. A. Tretter. "Estimating the frequency of a noisy sinusoid by linear regression". *IEEE Trans. Information Theory*, 31(6):832–835, November 1985.

[473] F. Tupin, H. Maître, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky. "Detection of linear features in SAR images: Application to road network extraction". *IEEE Trans. Geoscience & Remote Sensing*, 36(2):434–453, March 1998.

[474] D. Vakman. "On the analytic signal, the Teager-Kaiser energy algorithm, and other methods for defining amplitude and frequency". *IEEE Trans. Signal Processing*, 44(4):791–797, April 1996.

[475] D. E. Vakman. *Sophisticated Signals and the Uncertainty Principle in Radar*. Springer, New York, 1968. Translated by K. N. Trirogoff; edited by E. Jacobs.

[476] A. J. van Leest. *Non-separable Gabor schemes: Their design and implementation*. PhD thesis, Eindhoven University of Technology, Eindhoven, Netherlands, 2001.

[477] H. G. van Steenis, W. L. J. Martens, and J. H. M. Tulen. "Time-frequency parameters of heart-rate variability". *IEEE Engineering in Medicine & Biology Magazine*, 21(4):46–58, July-August 2002.

[478] H. L. L. Van Trees. *Detection, Estimation, and Modulation Theory*, volume I: "Detection, Estimation, and Linear Modulation Theory". Wiley, New York, 1968. Reprinted 2001.

[479] H. L. L. Van Trees. *Detection, Estimation, and Modulation Theory*, volume III: "Radar-Sonar Signal Processing and Gaussian Signals in Noise". Wiley, New York, 1971. Reprinted Malabar, FL: Krieger, 1992. Reprinted New York: Wiley, 2001.

[480] S. V. Vaseghi. *Advanced Signal Processing and Digital Noise Reduction*. Wiley and Teubner, 1996.

[481] E. F. Velez and H. Garudadri. "Speech analysis based on smoothed Wigner-Ville distribution". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 15, pages 351–374. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[482] B. V. K. Vijaya Kumar and C. W. Carroll. "Performance of Wigner distribution function based detection methods". *Optical Engineering*, 23(6):732–737, November-December 1984.

[483] J. Ville. "Théorie et applications de la notion de signal analytique". *Cables et Transmissions*, 2A(1):61–74, 1948. In French. English translation: I. Selin, *Theory and applications of the notion of complex signal*, Rand Corporation Report T-92 (Santa Monica, CA, August 1958).

[484] K. Vogel and H. Risken. "Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase". *Physical Review A*, 40:2847–2849, 1989.

[485] H. Vold and J. Leuridan. "High resolution order tracking at extreme slew rates using Kalman tracking filters". In *Proc. Noise & Vibration Conf. & Exposition.* Soc. of Automotive Engineers, Traverse City, MI, 17–20 May 1993. SAE paper no. 931288.

[486] 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.

[487] W. Wang; and D. H. Johnson. "Computing linear transforms of symbolic signals". *IEEE Trans. Signal Processing*, 50(3):628–634, March 2002.

[488] J. Wexler and S. Raz. "Discrete Gabor expansions". *Signal Processing*, 21(3):207–221, November 1990.

[489] A. D. Whalen. *Detection of Signals in Noise.* Academic Press, 1971.

[490] L. B. White. "Transition kernels for bilinear time-frequency signal representations". *IEEE Trans. Signal Processing*, 39(2):542–544, February 1991.

[491] L. B. White and B. Boashash. "Cross spectral analysis of nonstationary processes". *IEEE Trans. Information Theory*, 36(4):830–835, July 1990.

[492] H. J. Whitehouse, B. Boashash, and J. M. Speiser. "High-resolution processing techniques for temporal and spatial signals". In M. Bouvet and G. Bienvenu, editors, *High-resolution methods in underwater acoustics*, chapter 4, pages 127–176. Springer, Berlin, 1991.

[493] E. P. Wigner. "On the quantum correction for thermodynamic equilibrium". *Physics Review*, 40:749–759, June 1932.

[494] W. J. Williams. "Reduced interference distributions: Biological applications and interpretations". *Proc. IEEE*, 84(9):1264–1280, September 1996.

[495] W. J. Williams. "Recent advances in time-frequency representations: Some theoretical foundation". In M. Akay, editor, *Time Frequency and Wavelets in Biomedical Signal Processing*, chapter 1. IEEE/Wiley, New York, 1998.

[496] W. J. Williams. "Reduced interference time-frequency distributions: Scaled decompositions and interpretations". In L. Debnath, editor, *Wavelet Transforms and Time-Frequency Signal Analysis*, chapter 12. Birkhäuser, Boston, 2001.

[497] W. J. Williams and S. Aviyente. "Optimum window time-frequency distribution decompositions". In *Proc. 32nd Asilomar Conf. on Signals, Systems, and Computers*, pages 817–821, Pacific Grove, CA, 1–4 November 1998.

[498] W. J. Williams and S. Aviyente. "Minimal-window time-frequency distributions". In *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, volume 3807, pages 446–457. Soc. of Photo-optical Instrumentation Engineers, Denver, CO, 19–21 July 1999.

[499] W. J. Williams, M. L. Brown, and A. O. Hero III. "Uncertainty, information and time-frequency distributions". In *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations II*, volume 1566, pages 144–156. Soc. of Photo-optical Instrumentation Engineers, San Diego, 24–26 July 1991.

[500] W. J. Williams and J. Jeong. "Reduced interference time-frequency distributions". In B. Boashash, editor, *Time-Frequency Signal Analysis: Methods and Applications*, chapter 3, pages 74–97. Longman-Cheshire/Wiley, Melbourne/N.Y., 1992.

[501] W. J. Williams and J. C. O'Neill. "Decomposition of time-frequency distributions using scaled window spectrograms". In *Proc. SPIE: Advanced Signal Processing Algorithms*, volume 2563, pages 44–58. Soc. of Photo-optical Instrumentation Engineers, San Diego, CA, 10–12 July 1995.

[502] W. J. Williams and T.-H. Sang. "Adaptive RID kernels which minimize time-frequency uncertainty". In *Proc. IEEE-SP Internat. Symp. on Time-Frequency & Time-Scale Analysis*, pages 96–99, Philadelphia, PA, 25–28 October 1994.

[503] W. J. Williams, T.-H. Sang, J. C. O'Neill, and E. J. Zalubas. "Wavelet windowed time-frequency distribution decompositions". In *Proc. SPIE: Advanced Signal Processing Algorithms, Architectures, and Implementations VII*, volume 3162, pages 149–160. Soc. of Photo-optical Instrumentation Engineers, San Diego, July–August 1997.

[504] J. C. Wood and D. T. Barry. "Tomographic time-frequency analysis and its application toward time-varying filtering and adaptive kernel design for multicomponent linear FM signals". *IEEE Trans. Signal Processing*, 42:2094–2104, 1994.

[505] P. M. Woodward. *Probability and Information Theory with Applications to Radar.* Pergamon, London, 1953.

[506] T. Wysocki, H. Razavi, and B. Honary, editors. *Digital Signal Processing for Communication Systems.* Number 403 in KLUWER INTERNAT. SERIES IN ENGINEERING AND COMPUTER SCIENCE. Kluwer, Boston, MA, 1997.

[507] X.-G. Xia. "System identification using chirp signals and time-variant filters in the joint time-frequency domain". *IEEE Trans. Signal Processing*, 45(8):2072–2084, August 1997.

[508] X.-G. Xia and V. C. Chen. "A quantitative SNR analysis for the pseudo Wigner-Ville distribution". *IEEE Trans. Signal Processing*, 47(10):2891–2894, October 1999.

[509] X.-G. Xia, Y. Owechko, B. H. Soffer, and R. M. Matic. "On generalized-marginal time-frequency distributions". *IEEE Trans. Signal Processing*, 44(11):2882–2886, November 1996.

[510] X.-G. Xia and S. Qian. "Convergence of an iterative time-variant filtering based on discrete Gabor transform". *IEEE Trans. Signal Processing*, 47(10):2894–2899, October 1999.

[511] X.-G. Xia, G. Wang, and V. Chen. "Quantitative SNR analysis for ISAR imaging using joint time-frequency analysis—Short time Fourier transform". *IEEE Trans. Aerospace & Electronic Systems*, 38(3):649–659, April 2002.

[512] J. Yang. "Frequency domain noise suppression approaches in mobile telephone systems". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP'93)*, volume 2, pages 363–366, Minneapolis, 27–30 April 1993.

[513] L. A. Zadeh. "Frequency analysis of variable networks". *Proc. IRE*, 3A-8:291–299, March 1950.

[514] P. Zarjam, M. Mesbah, and B. Boashash. "Detection of newborn EEG seizure using optimal features based on discrete wavelet transform". In *Proc. IEEE Internat. Conf. on*

Acoustics, Speech and Signal Processing (ICASSP'03), volume 2, pages 265–268, Hong Kong, scheduled 6–10 April 2003.

[515] A. I. Zayed. "A class of fractional integral transforms: a generalization of the fractional Fourier transform". *IEEE Trans. Signal Processing*, 50(3):619–627, March 2002.

[516] Y. Y. Zeevi, M. Zibulski, and M. Porat. "Multi-window Gabor schemes in signal and image representations". In H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, chapter 12, pages 381–407. Birkhäuser, Berlin/Boston, 1998.

[517] F. Zhang, G. Bi, and Y. Q. Chen. "Tomography time-frequency transform". *IEEE Trans. Signal Processing*, 50(6):1289–1297, June 2002.

[518] Y. Zhang and M. G. Amin. "Blind separation of sources based on their time-frequency signatures". In *Proc. IEEE Internat. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2000)*, volume 5, pages 3132–3135, Istanbul, 5–9 June 2000.

[519] Y. Zhang and M. G. Amin. "Spatial averaging of time-frequency distributions for signal recovery in uniform linear arrays". *IEEE Trans. Signal Processing*, 48(10):2892–2902, October 2000.

[520] Y. Zhang, W. Mu, and M. G. Amin. "Time-frequency maximum likelihood methods for direction finding". *J. Franklin Institute*, 337(4):483–497, July 2000.

[521] Y. Zhang, W. Mu, and M. G. Amin. "Subspace analysis of spatial time-frequency distribution matrices". *IEEE Trans. Signal Processing*, 49(4):747–759, April 2001.

[522] Y. Zhao, L. E. Atlas, and R. J. Marks II. "The use of cone-shaped kernels for generalized time-frequency representations of non-stationary signals". *IEEE Trans. Acoustics, Speech, & Signal Processing*, 38(7):1084–1091, July 1990.

[523] H. Zheng, Z. Li, and X. Chen. "Gear fault diagnosis based on continuous wavelet transform". *J. Mechanical Systems and Signal Processing*, 16(2–3):447–457, March–May 2002.

[524] Y. M. Zhu, R. Goutte, and M. Amiel. "On the use of two-dimensional Wigner-Ville distribution for texture segmentation". *Signal Processing*, 30(3):329–353, February 1993.

[525] M. Zibulski and Y. Y. Zeevi. "Discrete multiwindow Gabor-type transforms". *IEEE Trans. Signal Processing*, 45(6):1428–1442, June 1997.

[526] R. Zou, W. A. Cupples, K. R. Yip, N. H. Holstein-Rathlou, and K. Chon. "Time-varying properties of renal autoregulatory mechanisms". *IEEE Trans. Biomedical Engineering*, 49(10):1112–1120, October 2002.

Time-Frequency Index

This Page Intentionally Left Blank