

Time-Frequency Analysis of Time-Varying Signals and Non-Stationary Processes

An Introduction

Maria Sandsten

2020



LUND UNIVERSITY

Centre for Mathematical Sciences

CENTRUM SCIENTIARUM MATHEMATICARUM

Contents

1	Introduction	3
1.1	Spectral analysis history	3
1.2	A time-frequency motivation example	5
2	The spectrogram	9
2.1	Spectrum analysis	9
2.2	The uncertainty principle	10
2.3	STFT and spectrogram	12
2.4	Gabor expansion	14
2.5	Wavelet transform and scalogram	17
2.6	Other transforms	19
3	The Wigner distribution	21
3.1	Wigner distribution and Wigner spectrum	21
3.2	Properties of the Wigner distribution	23
3.3	Some special signals.	24
3.4	Time-frequency concentration	25
3.5	Cross-terms	27
3.6	Discrete Wigner distribution	29
4	The ambiguity function and other representations	35
4.1	The four time-frequency domains	35
4.2	Ambiguity function	39
4.3	Doppler-frequency distribution	44
5	Ambiguity kernels and the quadratic class	45
5.1	Ambiguity kernel	45
5.2	Properties of the ambiguity kernel	46
5.3	The Choi-Williams distribution	48
5.4	Separable kernels	52

5.5	The Rihaczek distribution	54
5.6	Kernel interpretation of the spectrogram	57
5.7	Multitaper time-frequency analysis	58
6	Optimal resolution of time-frequency spectra	61
6.1	Concentration measures	61
6.2	Instantaneous frequency	63
6.3	The reassignment technique	65
6.4	Scaled reassigned spectrogram	69
6.5	Other modern techniques for optimal resolution	72
7	Stochastic time-frequency analysis	75
7.1	Definitions of non-stationary processes	75
7.2	The mean square error optimal kernel	80

Chapter 1

Introduction

We usually differ between deterministic signals and stochastic realizations. The values of a deterministic signal is explicitly known at all points in time and if the signal has a periodicity, typically a sinusoidal signal, the analysis is usually made in the frequency domain. A stochastic realization is just one of many possible in a collection, generated from a stochastic process model. The stochastic process is characterized by certain statistical properties, mainly described by the covariance function and the spectral density. A stationary stochastic process has the same properties for all points in time and with a reasonable length of the realization, these properties can be reliably estimated. When the properties of the signal changes with time, e.g. the sinusoidal frequency is increasing or the covariance function and spectral density are time-dependent, the analysis tools should be able to follow this behaviour. This forms the basic need for **time-frequency analysis** which has been an important field of research for the last 30 years.

1.1 Spectral analysis history

This is a short story of the history behind frequency analysis in signal processing, which starts with the famous glass prism experiment in 1704 by Sir Isaac Newton (1642-1727) where the sunbeams resolved into the colors of the rainbow. The result was an image of the frequencies contained in the sunlight. The glass prism was actually the first spectrum analyser.

Hundred years later, in 1807, Jean Baptiste Joseph Fourier (1768-1830) studied the most basic problem regarding heat, a discontinuity in temperature when hot and cold objects were put together. He invented the mathematics where discontinuous functions can be expressed as a sum of continuous frequency functions, and although his ideas were seen as absurd from the established scientists at the time, including

Pierre Simon Laplace (1749-1827) and Joseph Louis Lagrange (1736-1813), they were eventually accepted and is what we today call the **Fourier expansion**.

Joseph von Fraunhofer (1787-1826) invented the spectroscope in 1815 and discovered what has become known as the Fraunhofer lines. Later, Robert Bunsen (1811-1899) studied the light from a burning rag soaked with salt solution, where the spectrum appeared as a bright yellow line, and further experiments showed that every material has its own unique spectrum, with different frequency contents. Based on this discovery, Bunsen and Gustav Kirchhoff (1824-1887) showed that light spectra can be used for recognition, detection and classification of substances.

The spectra of electrical signals can be obtained by using narrow bandpass-filters. The measured output from the filter is squared and averaged and the result is recognized as the average power at the frequencies of the filter. Using many narrow bandpass-filters, an image of the frequency content is obtained. This spectrum analyzer performance is equivalent to modulating (shifting) the signal in frequency and using only one lowpass-filter. The modulation procedure will move the power of the signal at a certain frequency down to the base-band, i.e., around zero frequency. This interpretation is called the heterodyne technique, and is the basic idea behind Fourier based frequency and spectrum analysis. For a broader and deeper view on frequency and spectral estimation, see [1].

A first step in modern spectrum analysis, in the sense of sampled discrete-time analysis, was made by Sir Arthur Schuster (1851-1934) already in 1898, where he found hidden periodicities in sunspot numbers using Fourier series, or what is well known today as the periodogram, [2]. This method has been frequently used for spectrum estimation ever since James Cooley (1926-2016) and John Tukey (1915-2000) developed the Fast Fourier Transform (FFT) algorithm in 1965, [3, 4, 5]. However, it should be noted that Carl Friedrich Gauss (1777-1855) invented the FFT-algorithm in an unpublished manuscript already in 1805, long before the existence of any computers. The FFT is the efficient way of calculating the discrete Fourier transform of a signal which explores the structure of the Fourier transform algorithm by minimizing the number of multiplications and summations. This made the Fourier transform to actually become a tool and not just a theoretical description, [6].

All the techniques for classical spectrum estimation rely on the fact that the signal characteristics, such as frequencies, amplitudes, covariances and spectral densities, do not change with time. And if all these properties change slowly enough with time, also the time-dependent frequency spectrum of the signal can be estimated, using short-time versions of these methods. However, if the signal has new components showing up or disappearing at certain points in time, and has fast varying frequency, other appropriate methods optimal for the application purpose are needed.

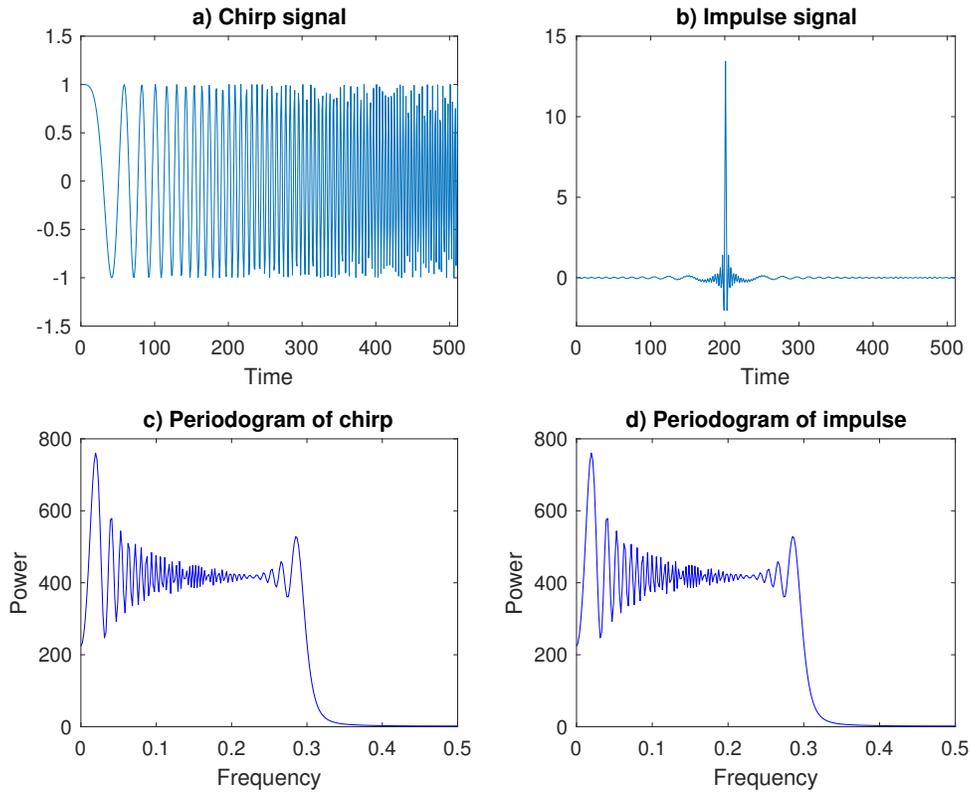


Figure 1.1: Two different time-varying signals a) and b) with identical periodograms, c) and d).

1.2 A time-frequency motivation example

With an illustrating example, the need for time-frequency analysis becomes obvious. Two different time-varying signals, a linear **chirp**, a sinusoidal signal with linearly increasing frequency, and an impulse could have identical spectral estimates, Figure 1.1. The two signals have the same magnitude function but different phase functions and conclusively the periodogram, the square of the magnitude, does not give a total picture as we see clearly in time that the two signals are very different. The **spectrogram** is a 3-dimensional representation that shows how the spectral density or the power of a signal vary with time. Other names that are found in different applications are **spectral waterfall** or **sonogram**.

There are several advantages of using the spectrogram, e.g., fast implementation

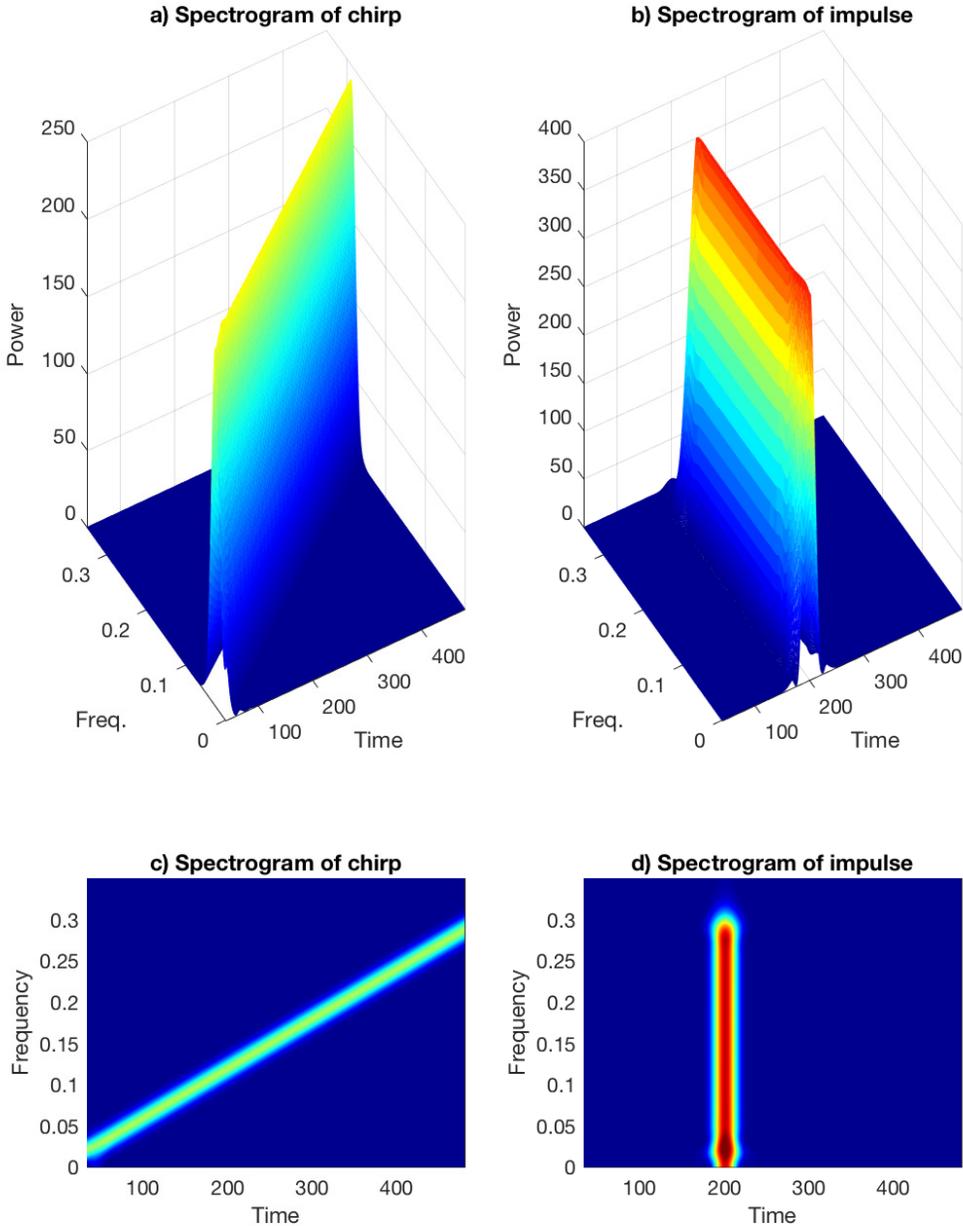


Figure 1.2: Spectrograms of the linear chirp signal and the impulse signal, (red color: high power, blue color: low power).

using the FFT, easy interpretation and the connection to the periodogram. The time domain data is divided in shorter sub-sequences, which often overlap, and for each short sequence, the calculation of the periodogram is made, giving frequency spectra for all sub-sequences. These frequency spectra are then ordered on a corresponding time-scale and form a three-dimensional picture, (time, frequency, power). The upper part of Figure 1.2 presents 3-dimensional views of the powers at different points of time and frequency of the chirp signal in Figure 1.2a and the impulse in Figure 1.2b, where low to high power are colored from blue to red. In the lower part of Figure 1.2, the views are right from above of the 3-dimensional figures and the power values low to high are only possible to view using the color-scale from blue to red. The changes of power with time and frequency are however easily interpretable and also here we clearly see that the chirp signal has linearly increasing frequency, Figure 1.2c and the impulse has all frequencies present at a single point of time, Figure 1.2d.

Chapter 2

The spectrogram

Dennis Gabor suggested a representation of a signal in two dimensions, time and frequency. In 1946 he defined certain elementary signals, one "quantum of information", that occupies the smallest possible area in the two-dimensional space, the "information diagram", [7]. Gabor had a large interest in holography and carried out basic experiments at that time called "wavefront reconstruction". The reconstruction is not needed to perform a short-time Fourier transform and a spectrogram, but for appropriate analysis and synthesis of non-stationary signals the Gabor expansion is essential. In 1971 Dennis Gabor received the Nobel Prize for his discovery of the principles underlying the science of holography.

2.1 Spectrum analysis

The Fourier transform of a continuous-time integrable signal $x(t)$, $-\infty < t < \infty$, is defined as

$$X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft} dt, \quad -\infty < f < \infty, \quad (2.1)$$

where the signal can be recovered by the inverse Fourier transform,

$$x(t) = \mathcal{F}^{-1}\{X(f)\} = \int_{-\infty}^{\infty} X(f)e^{i2\pi ft} df, \quad -\infty < t < \infty. \quad (2.2)$$

The absolute value of the Fourier transform gives us the **magnitude function**, $|X(f)|$ and the argument is the **phase function**, $\arg\{X(f)\}$. The **spectrum** is given from the squared magnitude function,

$$S_x(f) = |X(f)|^2, \quad -\infty < f < \infty. \quad (2.3)$$

Using the Wiener-Khintchine theorem, the **spectral density** of a zero-mean stationary stochastic process $x(t)$, $-\infty < t < \infty$, can be calculated as the Fourier transform of the **covariance function** $r_x(\tau)$,

$$S_x(f) = \mathcal{F}\{r_x(\tau)\} = \int_{-\infty}^{\infty} r_x(\tau) e^{-i2\pi f\tau} d\tau, \quad -\infty < f < \infty, \quad (2.4)$$

where $r_x(\tau)$ is defined as

$$r_x(\tau) = E[x(t - \tau)x^*(t)], \quad -\infty < \tau < \infty, \quad (2.5)$$

with $E[\]$ denoting expected value and $*$ complex conjugate. The covariance function $r_x(\tau)$ can be recovered by the inverse Fourier transform of the spectral density,

$$r_x(\tau) = \mathcal{F}^{-1}\{S_x(f)\} = \int_{-\infty}^{\infty} S_x(f) e^{i2\pi f\tau} df, \quad -\infty < \tau < \infty. \quad (2.6)$$

2.2 The uncertainty principle

A signal cannot be both time-limited and frequency-limited at the same time. This is intuitively understood with the simple example where an arbitrary signal $x(t)$ which is multiplied with a rectangular function. The signal will certainly be time-limited to the interval defined by the rectangular function, i.e., the windowed signal will have **compact support**. The Fourier transform, however, will be of infinite bandwidth as the multiplication in time is transferred to a convolution in frequency with an infinite length sinc-function, ($\text{sinc}(x) = \sin(\pi x)/(\pi x)$), which means that even if the Fourier transform $X(f)$ is limited, the resulting convolution will be of infinite length.

Another concentration measure which is often used is the **effective duration** defined as

$$T_e = \sqrt{\frac{\int_{-\infty}^{\infty} t^2 |x(t)|^2 dt}{\int_{-\infty}^{\infty} |x(t)|^2 dt}}, \quad (2.7)$$

and the corresponding **effective bandwidth**,

$$B_e = \sqrt{\frac{\int_{-\infty}^{\infty} f^2 |X(f)|^2 df}{\int_{-\infty}^{\infty} |X(f)|^2 df}}, \quad (2.8)$$

where the signal as well as the Fourier transform are assumed to be located symmetrically around $t = 0$ and $f = 0$. The **uncertainty principle** or the **bandwidth-duration product**, defined as $B_e \cdot T_e$, serves as a measure of the information concentration in the signal. It can be shown that the bandwidth-duration product always fulfills

$$B_e \cdot T_e \geq \frac{1}{4\pi}, \quad (2.9)$$

and that only the Gaussian signal reaches the equality, [7]. We compute the bandwidth-duration product of a Gaussian signal, defined as

$$x(t) = e^{-at^2}, \quad -\infty < t < \infty, \quad (2.10)$$

with the Fourier transform

$$X(f) = \int_{-\infty}^{\infty} e^{-at^2} e^{-i2\pi ft} dt = \sqrt{\frac{\pi}{a}} e^{-\frac{\pi^2 f^2}{a}}, \quad -\infty < f < \infty. \quad (2.11)$$

The integral is derived using the formula

$$\int_{-\infty}^{\infty} e^{-ax^2+bx} dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}}, \quad (2.12)$$

from which also the well known Gaussian integral, $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$, is recognized when $a = 1$ and $b = 0$.

The integrals included in the expression for the duration T_e in Eq. (2.7) are then easily calculated as

$$\begin{aligned} \int_{-\infty}^{\infty} |x(t)|^2 dt &= \int_{-\infty}^{\infty} e^{-2at^2} dt = \sqrt{\frac{\pi}{2a}}, \\ \int_{-\infty}^{\infty} t^2 |x(t)|^2 dt &= \int_{-\infty}^{\infty} t^2 e^{-2at^2} dt = \frac{1}{4a} \sqrt{\frac{\pi}{2a}}, \end{aligned}$$

using partial integration for the second one. Similarly, the integrals included in the expression of the bandwidth B_e in Eq. (2.8) are found as

$$\begin{aligned} \int_{-\infty}^{\infty} |X(f)|^2 df &= \int_{-\infty}^{\infty} \frac{\pi}{a} e^{-\frac{2\pi^2 f^2}{a}} df = \sqrt{\frac{\pi}{2a}}, \\ \int_{-\infty}^{\infty} f^2 |X(f)|^2 df &= \int_{-\infty}^{\infty} \frac{\pi}{a} f^2 e^{-\frac{2\pi^2 f^2}{a}} df = \frac{1}{4\pi} \sqrt{\frac{a}{2\pi}}. \end{aligned}$$

The resulting bandwidth-duration product is

$$B_e \cdot T_e = \sqrt{\frac{a}{4\pi^2}} \cdot \sqrt{\frac{1}{4a}} = \frac{1}{4\pi}, \quad (2.13)$$

which is the minimum value of Eq. (2.9).

2.3 STFT and spectrogram

A natural extension of the Fourier transform when the signals are time-varying or non-stationary is the **short-time Fourier transform (STFT)**, which is defined as

$$X(t, f) = \int_{-\infty}^{\infty} x(t_1) h^*(t_1 - t) e^{-i2\pi f t_1} dt_1, \quad -\infty < t, f < \infty, \quad (2.14)$$

where $h(t)$ is a window function centered at time t . The window function cuts the signal just close to the time t and the Fourier transform will be an estimate locally around this time instant. The usual way of calculating the STFT is to use a fixed positive even window, $h(t)$, of a certain shape, which is centered around zero and has power $\int_{-\infty}^{\infty} |h(t)|^2 dt = 1$. Similar to the ordinary Fourier transform and spectrum we can formulate the spectrogram as

$$S_x(t, f) = |X(t, f)|^2, \quad -\infty < t, f < \infty, \quad (2.15)$$

which is frequently used for analyzing time-varying and non-stationary signals. We illustrate the spectrogram with an example: In Figure 2.1a, a sequence of data, x_n , $n = 0, 1, 2, \dots, N - 1$, consisting of several short frequency components is shown. A musical interpretation of this is three tones of increasing tone height. The resulting spectrogram of the tones is presented in Figure 2.1b and shows a clear view of three frequency components and their locations in time and frequency. In practice, the measured signal is usually sampled with some sample distance T , i.e. $x_n = x(nT)$,

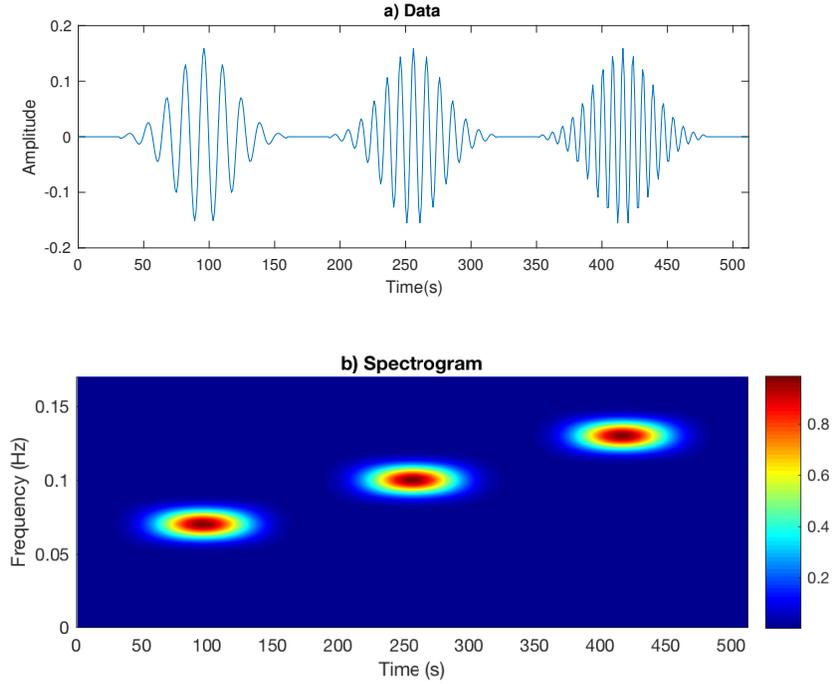


Figure 2.1: Spectrogram using a Hanning window.

related to the sample frequency F_s where $T = 1/F_s$. The discrete-time and discrete-frequency spectrogram is defined as

$$S_x(n, l) = \left| \sum_{n_1=0}^{N-1} x_{n_1} h^*(n_1 - n + M/2) e^{-i2\pi n_1 \frac{l}{L}} \right|^2, \quad (2.16)$$

where the window function $h(n)$ is of length M and energy normalized according to

$$h(n) = \frac{h_1(n)}{\sqrt{\sum_{n=0}^{M-1} h_1^2(n)}}, \quad n = 0 \dots M - 1. \quad (2.17)$$

The length (and shape) of the window function $h(n)$ is very important as it determines the resolution in time and frequency, as shown in Figure 2.2, where the real part of the three Gaussian component signal is shown in Figure 2.2a. Figures 2.2b, c and d show the spectrograms when $M = 32$, $M = 64$, and $M = 128$, are chosen as window lengths of the spectrogram Hanning window. With short window length the two components located at the same time interval are smeared together as the frequency resolution

decreases, Figure 2.2b. In Figure 2.2c, the window length is about the same as the component lengths, which is a rule of thumb in calculations using the spectrogram. The optimal choice of window length however depends on the relative locations and power levels of the components and is not usually easy to decide in practice. With long window length, Figure 2.2d, the frequency resolution becomes reasonable but now the two low-frequency components are smeared together due to low time resolution. Additionally, mainlobe width and sidelobe height related to the window shape, affect the resolution and possibility to detect weak amplitude signals, in exactly the same way as for the usual windowed periodogram, [8].

Similarly, the spectrogram calculation is a fast and efficient computation using the Fast Fourier transform (FFT) algorithm where the choice of FFT-length L gives computed spectrogram values for frequencies, $l = 0, \frac{F_s}{L}, \frac{2F_s}{L}, \dots$. The number of frequency values should always be $L = 2^I$, where I is some integer value for the best performance of the FFT but does not have to be chosen larger than N , i.e. the whole data length, which is a usual mistake. The only length to consider is the windowed sequence length M , so that a reasonable number of frequency values are considered and that L is larger than M . In Figures 2.1 and 2.2, $L = 1024$. An even more increased speed in calculation is easily achieved if the computations are not made for all values of n in Eq. (2.16), but just for, $n = 0, N_{step}, 2N_{step}, \dots$, where N_{step} can often be chosen as large as $M/8$ without any significant visual change of the spectrogram.

2.4 Gabor expansion

The short-time Fourier transform and the Gabor expansion was related by Martin Bastiaans in 1980, [9]. A **Gabor system** is defined by translations and modulations. For a signal $x(t)$, $-\infty < t < \infty$, the Gabor expansion is defined as

$$x(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{m,k} g(t - mT_0) e^{i2\pi k F_0 t}, \quad (2.18)$$

where T_0 and F_0 denote the time and frequency sampling steps. The Gabor function, also called Gabor atom, is defined by $g(t - mT_0) e^{i2\pi k F_0 t}$ and the coefficients $a_{m,k}$ are called the Gabor coefficients. Gabor chose the Gaussian signal

$$g(t) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\alpha}{2} t^2}, \quad (2.19)$$

as the elementary function or **synthesis window**, because it is optimally concentrated in the joint time-frequency domain in terms of the uncertainty principle. However, in

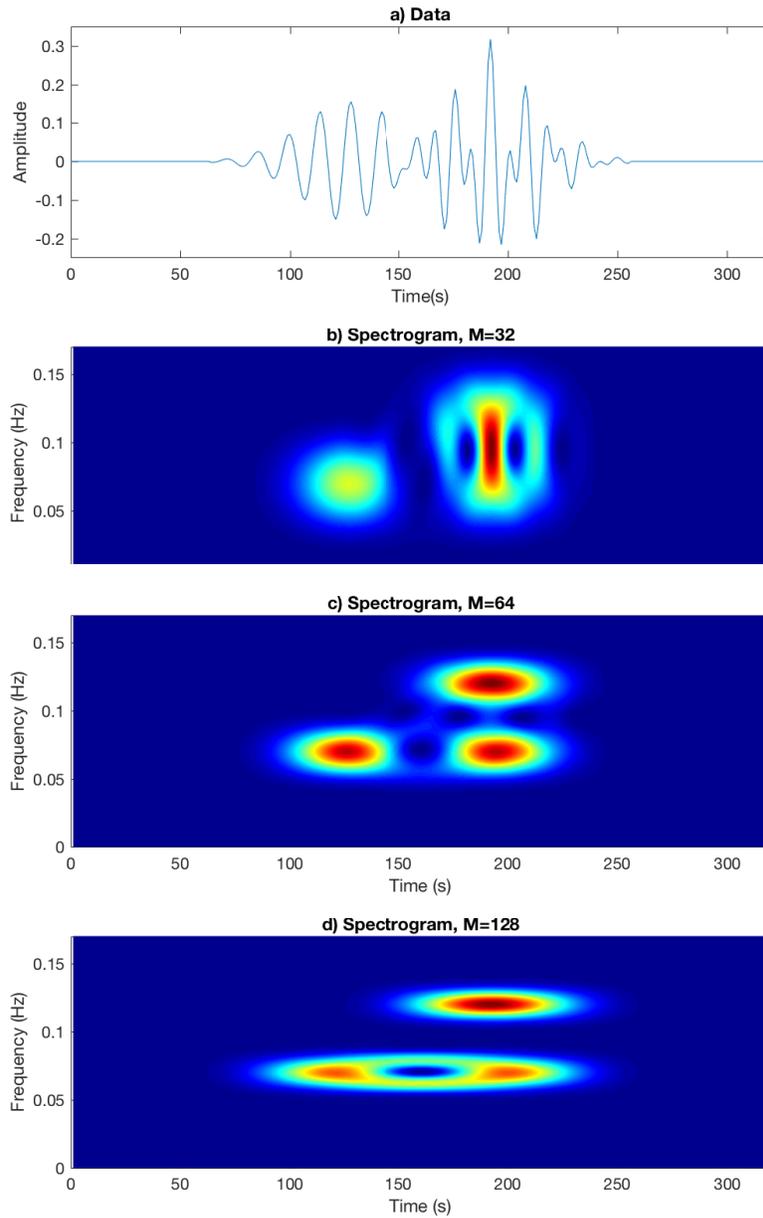


Figure 2.2: Spectrogram examples using a Hanning window with different window lengths; a) The real part of the signal consisting of three Gaussian windowed sinusoidal components located at (t_0, f_0) , $(130, 0.07)$, $(180, 0.07)$ and $(180, 0.12)$; b) $M = 32$; c) $M = 64$; d) $M = 128$.

practice, the Gaussian window cannot be used as it does not possess the property of compact support, and using the Gaussian window actually means using a truncated version of this window. The related sampled STFT, also known as the Gabor transform, is

$$X(mT_0, kF_0) = \int_{-\infty}^{\infty} x(t_1)w^*(t_1 - mT_0)e^{-i2\pi kF_0t_1} dt_1, \quad (2.20)$$

where $X(mT_0, kF_0) = a_{m,k}$, if the sampling distances T_0 and F_0 satisfy the **critical sampling** relation $F_0T_0 = 1$. In this case there is an uniquely defined **analysis window** $w(t)$ given from the synthesis window $g(t)$. This is seen if Eq. (2.20) is substituted into Eq. (2.18) as

$$x(t) = \int_{-\infty}^{\infty} x(t_1) \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} g(t - mT_0)w^*(t_1 - mT_0)e^{i2\pi kF_0(t-t_1)} dt_1, \quad (2.21)$$

which is true if

$$\sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} g(t - mT_0)w^*(t_1 - mT_0)e^{i2\pi kF_0(t-t_1)} = \delta(t - t_1). \quad (2.22)$$

The expression could be solved and the analysis window $w(t)$ follows uniquely from a given synthesis window, [9]. For discrete-time signals, calculation of the analysis window is made using the Zak transform, [10].

However, the analysis window does not necessary have attractive properties from a time-frequency analysis perspective, i.e., localization in time- and frequency. This window is not limited in frequency either as the sharp edges in time will correspond to a well spread function in frequency. If instead an oversampled lattice is used, i.e., $F_0T_0 < 1$, the analysis window is no longer unique, and a choice can be made from a time-frequency localization perspective. A mean square error solution gives the analysis window which is most similar to the synthesis window, [11], see Figure 2.3. The Gaussian synthesis window in Figure 2.3a generates at critical sampling the analysis window of Figure 2.3b which is not attractive for generating a time-frequency distribution, but might be the best for other applications, e.g., edge detection in image analysis. For increased oversampling, the analysis window becomes more and more similar to the synthesis window, Figure 2.3c and d.

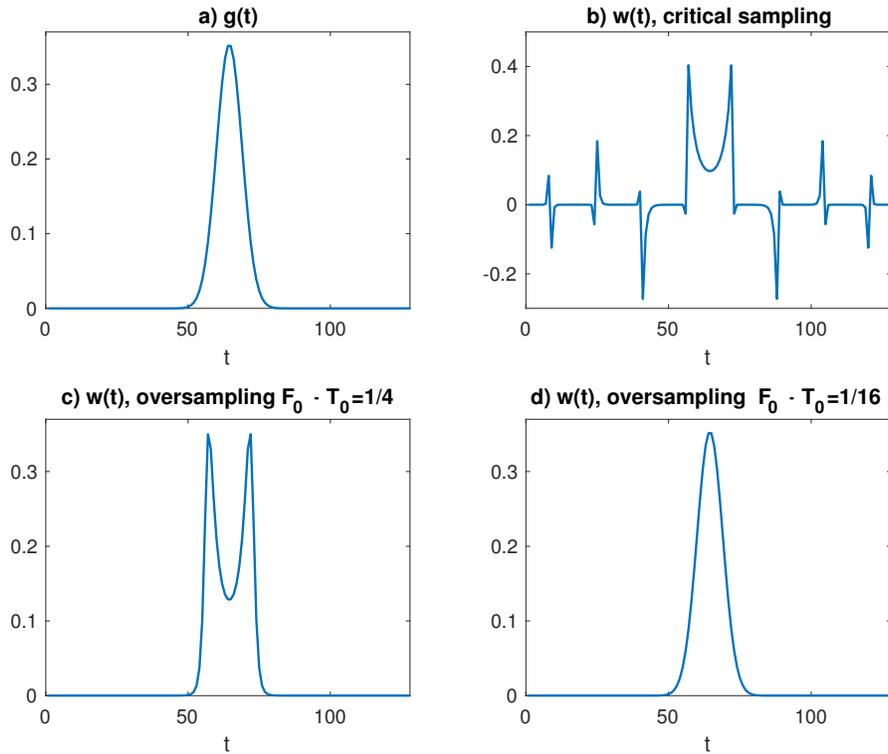


Figure 2.3: a) A Gaussian synthesis window $g(t)$. The analysis window $w(t)$ at; b) critical sampling; c) oversampling a factor 4; d) oversampling a factor 16.

2.5 Wavelet transform and scalogram

Wavelets analyse data at different scales or resolutions. The definition of the continuous wavelet transform (CWT) is

$$CWT(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t_1) h^*\left(\frac{t_1 - b}{a}\right) dt_1 \quad (2.23)$$

where, similarly to the STFT and the spectrogram, a **scalogram** is produced from the squared absolute value of the CWT.

The first and important issue, is to find the prototype function, the mother wavelet. The most trivial function is the Haar wavelet, which has compact support, is first mentioned in the thesis by Alfréd Haar in 1909, [12]. The first non-trivial wavelet was

invented by Yves Meyer. Actually he was waiting for a photocopier where a colleague was copying a paper on Morlet's wavelets (ondelettes), and started a conversation. The Meyer wavelets are continuously differentiable but do not have compact support, in contrast to the Haar wavelet. From the mid-1980s Meyer, together with Ingrid Daubechies and Ronald Coifman, made the earlier work of wavelet theory into a unified picture. Especially should the work by Daubechies be noticed, as she found a systematical method to construct the compact support, orthogonal wavelet. The famous Daubechies wavelets (DBW) are the most applied today for compression and reconstruction, [13, 14].

The discrete wavelet transform (DWT) can be computed with multiresolution analysis, a fast and efficient procedure initially applied to image compression, using the pyramid algorithm by Stéphane Mallat, [15]. He performed the work on relationships between quadrature mirror filters, pyramidal algorithms and orthonormal wavelet bases, which has been very valuable for the computation of orthogonal and inversable basis functions as the computation of the DWT is very efficient using a tree structure. The original signal passes through a low-pass and a high-pass filter, generating a smoothed approximation and a detail (noise). The wavelets are defined with $a = 2^l$ and $b = n2^l$ for integer vales of l and n . However, for a better view of the scalogram, without considering reconstruction, the so called sampled CWT is used, where l and n are not necessarily integers.

A sufficient condition for the reconstruction of any signal of finite energy is that the wavelet form an orthonormal basis and that a scaling function exists. The number of vanishing moments for with wavelet as well as the scaling function is important for compression purpose. The higher number of vanishing moments, the higher is the possible signal compression for that basis. The Meyer wavelet is not compactly supported, however there exists a good approximation leading to FIR-filters, which then allows for use of the DWT. The mexican hat as well as the Morlet wavelets are not orthogonal, nor do the corresponding scaling function exists. These two are accordingly not useful for the DWT, but are still appropriate for the analysis using the scalogram. The DBW family includes the Haar wavelet as the first and subsequent members of the family have increasing number of vanishing moments. Symmetry of wavelets are important for dephasing in, e.g., image processing. The DBW is not symmetrical. If symmetry is important, e.g. the symlet or the coiflet wavelets can be used.

The DWT is actually a subset of the wavelet packet transform, as the DWT is represented as a tree of low- and high-pass filters where only the lower resolution components are saved. Discrete wavelet packet analysis (DWPA) is similar but also the high-frequency components are saved. Meyer was the Abel Prize Laureate of 2017, see <https://www.abelprize.no/c69461/seksjon/vis.html?tid=69535> for his great contribu-

tions to the development of the wavelet theory. Here is also an opportunity to view special Abel lectures given by Meyer, Mallat, Daubechies and Emmanuel Jean Candés.

2.6 Other transforms

The fractional Fourier transform (FrFT) is a generalization of the Fourier transform and has been used extensively in optics but also in other application areas, [16, 17]. The FrFT is defined as

$$X_\alpha(u) = \int_{-\infty}^{\infty} x(t)K(\alpha, t, u)dt, \quad (2.24)$$

where the kernel $K(\alpha, t, u)$ is given by

$$K(\alpha, t, u) = \frac{\exp(i\alpha/2)}{\sqrt{i \sin(\alpha)}} \exp\left(i\pi \frac{(t^2 + u^2) \cos(\alpha) - 2ut}{\sin(\alpha)}\right). \quad (2.25)$$

Note that the FrFT is a generalization of the ordinary Fourier transform, where $\alpha = \pi/2$ and $\alpha = -\pi/2$.

The Stockwell transform, [18], (sometimes also referred to as the S-transform), is applied for localization of the complex valued spectrum. The original definition is

$$S_x(t, f) = \int_{-\infty}^{\infty} x(t_1) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(t-t_1)^2 f^2}{2}} e^{-i2\pi f t_1} dt_1, \quad (2.26)$$

where it can be seen that the Gaussian window function width is decreasing for higher frequency values which is similar to the wavelet transform. For the differences and similarities of the Stockwell transform and the wavelet transform arguments are given in [19, 20]. The interest to apply the Stockwell transform has increased in recent years, e.g., to applications such as detection of epileptic seizures from EEG, [21] and double-talk-detection for acoustic echo cancellation, [22]. A signal-adaptive method is found in [23], where the width of the Gaussian window function is estimated using some concentration criterion.

Higher-order generalizations of the STFT such as, e.g., the local polynomial Fourier transform (LPFT) are applied in many different areas to reduce noise interference. The LPFT relies on the assumption that the phase of an time-varying signal can be approximated by a polynomial. A review of the developments including different application areas is found in [24].

Chapter 3

The Wigner distribution

It could be expected that time-frequency basics is the spectrogram but we will discover that this is just one of many possible time-frequency representations derived from the **Wigner distribution**. The Wigner distribution, where the use of distribution should not be understood in the sense of statistics, was suggested in a famous publication from 1932 in the area of quantum mechanics by Eugene Wigner, [25]. Wigner also received the Nobel Prize in 1963, together with Maria Goeppert Mayer and Hans Jensen, for the discovery concerning the theory of the atomic nucleus and elementary particles. The suggested definition of the Wigner distribution, the one also used today, was actually chosen "because it seems to be the simplest". Jean-André Ville, [26] redefined the **analytic signal**, where a real-valued signal is converted into a complex-valued signal of non-negative frequency content. The Wigner distribution using the analytic signal, which is usually applied in practice today, is therefore often also called the **Wigner-Ville distribution**. The Wigner distribution has the best possible concentration but somewhere the cost of this optimal concentration needs to be paid and the problem shows up as **cross-terms**, that is large oscillating terms located in the middle between the actual signal components.

3.1 Wigner distribution and Wigner spectrum

For the deterministic signal $x(t)$, $-\infty < t < \infty$, the Wigner distribution is defined as

$$W_x(t, f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right)e^{-i2\pi f\tau} d\tau, \quad -\infty < t, f < \infty, \quad (3.1)$$

where we for a non-stationary stochastic process define the **instantaneous autocorrelation function**, (IAF),

$$r_x(t, \tau) = E[x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})], \quad (3.2)$$

and the corresponding **Wigner spectrum**

$$S_x(t, f) = \int_{-\infty}^{\infty} r_x(t, \tau) e^{-i2\pi f\tau} d\tau, \quad -\infty < t, f < \infty, \quad (3.3)$$

which is well in accordance with Eq. (3.1) also fulfills the basic properties of the Wigner distribution. Note the small but important difference between the two IAFs in Eq. (3.2) and Eq. (2.5).

The analytic signal

For real-valued signals and process realizations, the **Hilbert transform** is often used to transform the signal into the **analytic signal** $z(t)$, $-\infty < t < \infty$. The Hilbert transform of a real-valued signal $x(t)$, $-\infty < t < \infty$, is computed as

$$\mathcal{H}\{x(t)\} = \mathcal{F}^{-1}\{(-i \operatorname{sign}(f))\mathcal{F}\{x(t)\}\}, \quad (3.4)$$

giving the analytic signal

$$z(t) = x(t) + i\mathcal{H}\{x(t)\}. \quad (3.5)$$

The resulting spectrum of $z(t)$ is zero for negative values of f and of identical shape of the spectrum of $x(t)$ for $f \geq 0$. The calculation of the analytic signal can be made in the time domain as well, [27], but the interpretation in frequency domain is much easier. There are several advantages of using the corresponding analytic signal as input to the Wigner distribution, which is exemplified in Section 3.3. However, first we introduce some general properties of the Wigner distribution.

3.2 Properties of the Wigner distribution

The most important properties of the Wigner distribution are the following:

- The Wigner distribution is always real-valued even if the signal is complex-valued, i.e., $W_x^*(t, f) = W_x(t, f)$.
- For real-valued signals the frequency domain is symmetrical, i.e., $W_x(t, -f) = W_x(t, f)$, (compare with the definition of spectrum and spectrogram for real-valued signals).
- A shift of the signal in time or frequency, will cause the Wigner distribution to be shifted accordingly, it is **time-shift and frequency-shift invariant**, i.e., if $y(t) = x(t - t_0)$ then

$$W_y(t, f) = W_x(t - t_0, f), \quad (3.6)$$

and if $y(t) = x(t)e^{i2\pi f_0 t}$ then

$$W_y(t, f) = W_x(t, f - f_0). \quad (3.7)$$

- The Wigner distribution also satisfies the so called **time- and frequency marginals**, defined as

$$\int_{-\infty}^{\infty} W_x(t, f) df = |x(t)|^2, \quad \text{Time marginal}, \quad (3.8)$$

which is a energy conservation property of the time signal, where

$$\int_{-\infty}^{\infty} W_x(t, f) dt = |X(f)|^2, \quad \text{Frequency marginal}, \quad (3.9)$$

is related to the periodogram estimate of the total signal. If both these marginals are satisfied, the total energy condition is also automatically satisfied,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(t, f) dt df = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df = E_x, \quad (3.10)$$

where E_x is the total energy of the signal.

3.3 Some special signals.

To understand the properties of the Wigner distribution we introduce some special signals. We define a complex-valued sinusoidal signal of frequency f_0 , $x(t) = e^{i2\pi f_0 t}$, $-\infty < t < \infty$, and calculate the Wigner distribution,

$$\begin{aligned} W_x(t, f) &= \int_{-\infty}^{\infty} e^{i2\pi f_0(t+\frac{\tau}{2})} e^{-i2\pi f_0(t-\frac{\tau}{2})} e^{-i2\pi f\tau} d\tau \\ &= \int_{-\infty}^{\infty} e^{i2\pi(f_0\tau - f\tau)} d\tau = \delta(f - f_0). \end{aligned}$$

Similarly, for a complex-valued linear chirp signal $x(t) = e^{i2\pi\frac{\beta}{2}t^2 + i2\pi f_0 t}$, $-\infty < t < \infty$, we find

$$\begin{aligned} W_x(t, f) &= \int_{-\infty}^{\infty} e^{i2\pi(\frac{\beta}{2}(t+\frac{\tau}{2})^2 + f_0(t+\frac{\tau}{2}))} e^{-i2\pi(\frac{\beta}{2}(t-\frac{\tau}{2})^2 + f_0(t-\frac{\tau}{2}))} e^{-i2\pi f\tau} d\tau \\ &= \int_{-\infty}^{\infty} e^{i2\pi(\beta t + f_0 - f)\tau} d\tau \\ &= \delta(f - f_0 - \beta t). \end{aligned}$$

We should also mention that for $x(t) = \delta(t - t_0)$, the Wigner distribution is given as

$$W_x(t, f) = \delta(t - t_0). \quad (3.11)$$

The actual computation of the Wigner distribution in this case, involves the multiplication of two delta-functions, which should not be allowed, but we leave this theoretical dilemma to the serious mathematics. For the impulse, the mono-component complex-valued sinusoid and linear chirp signal, the Wigner distribution gives exactly the **instantaneous frequencies**, i.e. perfectly localized time-frequency representations. For these three signals, the time-frequency resolution of the Wigner distribution is unbeatable. Also for short components, such as a Gaussian windowed sinusoid, the Wigner distribution concentration is preferable.

An interesting interpretation is found for the Wigner distribution of a FrFT $X_\alpha(t)$. The relation is

$$W_{X_\alpha}(t, f) = W_x(t \cos(\alpha) - f \sin(\alpha), t \sin(\alpha) + f \cos(\alpha)), \quad (3.12)$$

which corresponds to a rotation with the angle α of the time-frequency plane, [28]. It has been applied to improve time-frequency concentration and filtering of non-stationary disturbances, [29]. However, the discrete-time implementation is not always easy as different invented implementations gives very different results.

3.4 Time-frequency concentration

We simplify and calculate the Wigner distribution of a Gaussian function of center locations zero frequency and zero time. It can easily be shown that the calculations are the same for all other possible time and frequency locations using the time-shift and frequency-shift properties. The signal is defined as

$$x(t) = \left(\frac{\beta}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\beta}{2}t^2}, \quad -\infty < t < \infty, \quad (3.13)$$

and the Wigner distribution is calculated as

$$\begin{aligned} W_x(t, f) &= \sqrt{\frac{\beta}{\pi}} \int_{-\infty}^{\infty} e^{-\frac{\beta}{2}(t+\frac{\tau}{2})^2} e^{-\frac{\beta}{2}(t-\frac{\tau}{2})^2} e^{-i2\pi f\tau} d\tau \\ &= \sqrt{\frac{\beta}{\pi}} e^{-\beta t^2} \int_{-\infty}^{\infty} e^{-\frac{\beta}{4}\tau^2 - i2\pi f\tau} d\tau \\ &= 2e^{-(\beta t^2 + \frac{4\pi^2 f^2}{\beta})}, \end{aligned} \quad (3.14)$$

using the relation in Eq. (2.12). The resulting Wigner distribution is depicted in Figure 3.1a. The maximum value is 2 for $t = f = 0$ and to measure the concentration, we decide to move a factor e downhill from the peak value, to the height $2e^{-1} \approx 0.736$, where we find the elliptic cross-section,

$$\beta t^2 + \frac{4\pi^2 f^2}{\beta} = 1, \quad (3.15)$$

with the area $A = \frac{1}{2}$, Figure 3.1b (dark red area).

For comparison, using a Gaussian window $h(t) = \left(\frac{\alpha}{\pi}\right)^{\frac{1}{4}} e^{-\frac{\alpha}{2}t^2}$, the resulting spectrogram can be calculated as

$$S_x(t, f) = \frac{2\sqrt{\alpha\beta}}{\alpha + \beta} e^{-\frac{\alpha\beta}{\alpha+\beta}t^2 - \frac{1}{\alpha+\beta}4\pi^2 f^2}. \quad (3.16)$$

Studying the spectrogram one can note that the area of the ellipse a factor e downhill from the peak value is

$$A = \frac{\alpha + \beta}{2\sqrt{\alpha\beta}}. \quad (3.17)$$

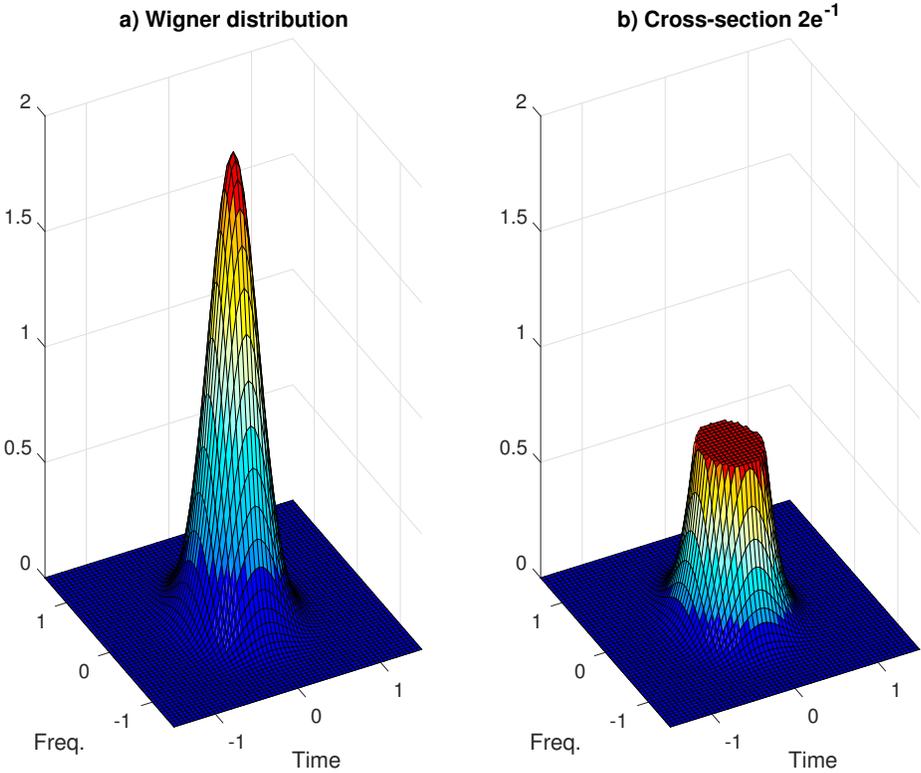


Figure 3.1: a) Wigner distribution of a Gaussian function; b) the cross-section area (dark red) at $2e^{-1}$.

Differentiation gives a minimum area $A = 1$ for $\alpha = \beta$. The conclusion of this small example is twofold, the best concentration of the spectrogram of a Gaussian signal is given if the window length is equal to the signal component. However, the Wigner distribution is much better and still gives half the time-frequency spread compared to the optimal spectrogram. So, why don't we always use the Wigner distribution in all calculations and why is there a huge field of research to find new well-resolved time-frequency distributions? We will look more closely into the major drawback of the Wigner distribution in the next section.

3.5 Cross-terms

The Wigner distribution **cross-terms**, are located in the middle between and can be twice as large as the different signal components. And it does not matter how far apart the different signal components are, cross-terms show up anyway and between all components of the signal as well as of the disturbance components. This makes the Wigner distribution useless for most signals that are not just toy signals.

The explanation of these phenomenon is found if we study the two-component signal $x(t) = x_1(t) + x_2(t)$ for which the Wigner distribution is

$$W_x(t, f) = W_{x_1}(t, f) + W_{x_2}(t, f) + 2\Re[W_{x_1, x_2}(t, f)], \quad (3.18)$$

where $W_{x_1}(t, f)$ and $W_{x_2}(t, f)$ are called **auto-terms**, and are the Wigner distributions of $x_1(t)$ and $x_2(t)$ respectively. The term

$$2\Re[W_{x_1, x_2}(t, f)] = 2\Re[\mathcal{F}\{x_1(t + \tau/2)x_2^*(t - \tau/2)\}]$$

is called cross-term. The cross-term will always be present, located midway between the two auto-terms and oscillating proportionally to the distance between the auto-terms. The direction of the oscillation will be orthogonal to the line connecting the auto-terms, see Figure 3.2, which shows the Wigner distribution of the signal displayed in Figure 2.1. The Wigner distribution is shown in a 3-dimensional plot as well as in a 2-dimensional color representation and we should note that we also have negative amplitude values present. With some knowledge, possibly from the view of Figure 2.1 we can locate the actual components, the auto-terms. For simplicity, they are called $C1$, $C2$ and $C3$ and they are marked in the lower figure. We find two of them as yellow-colored smaller components where the middle one, $C2$, is larger, marked with a red-yellow square-pattern. This pattern consists of the middle component and the

added corresponding oscillating cross-term from the outer components ($C1$ and $C3$). Then there is a cross-term found in the middle between $C1$ and $C2$ and one between $C2$ and $C3$. In conclusion, from these three signal components, the Wigner distribution will give the three components and additionally three cross-terms (one between each possible pair of components). All the cross-terms oscillates and it should be noted that they also adopt negative values (blue colour), which seems strange if we would like to interpret the Wigner distribution computation as a time-varying power estimate. This is the major draw-back of the Wigner distribution, and this example shows that the result could easily be misinterpreted.

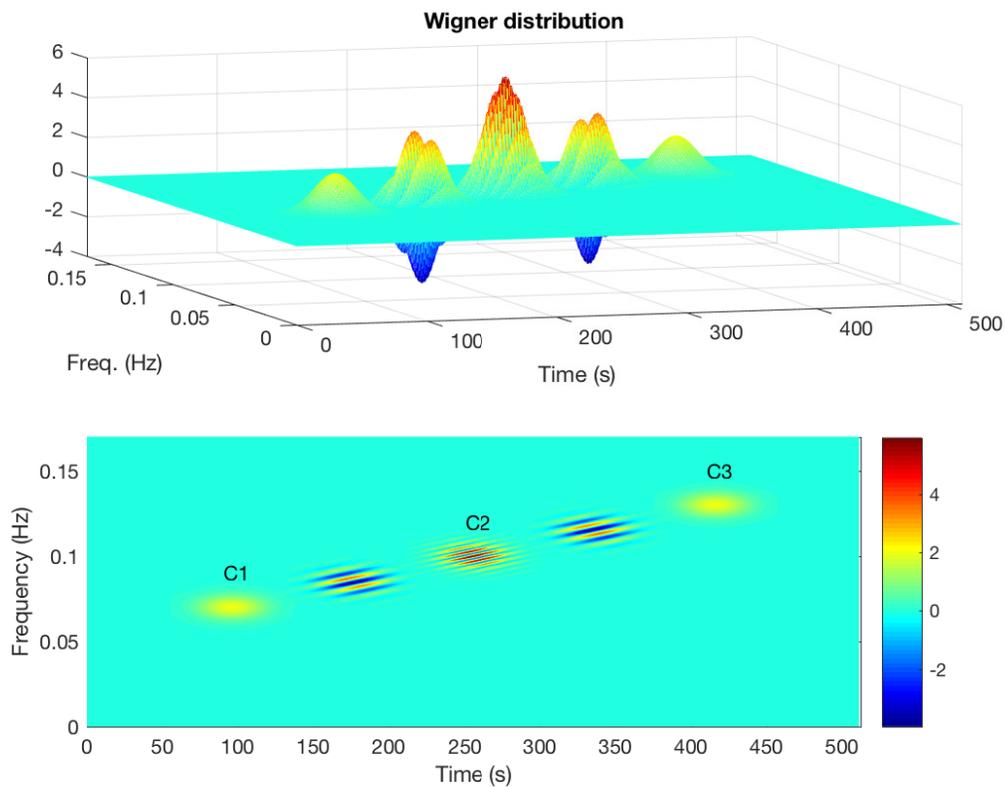


Figure 3.2: The Wigner distribution of the signal shown in Figure 2.1a.

The Wigner distribution does not have **strong time- and frequency support** as cross-terms arise in between signal components of a multi-component signal. Strong support means that whenever the signal is zero, then the distribution also should be zero. However, a **weak time- and frequency support** is satisfied, as the Wigner

distribution is zero before the signal starts and after the signal ends, i.e., zero for the outer limits of the whole signal.

Cross-terms of the spectrogram

Actually, the spectrogram also gives cross-terms but they show up as strong components mainly when the signal components are close to each other.

This is seen in Figure 3.3 where two Gaussian windowed components are located close in the time- and frequency plane. With $x_1(t)$ and $x_2(t)$, the spectrogram of the sum of the signals, $x(t) = x_1(t) + x_2(t)$ is

$$\begin{aligned} |X(t, f)|^2 &= |X_1(t, f) + X_2(t, f)|^2 \\ &= |X_1(t, f)|^2 + |X_2(t, f)|^2 + X_1(t, f)X_2^*(t, f) + X_1^*(t, f)X_2(t, f) \\ &= |X_1(t, f)|^2 + |X_2(t, f)|^2 + 2\Re[X_1(t, f)X_2^*(t, f)], \end{aligned}$$

depicted in Figure 3.3b. The two terms $|X_1(t, f)|^2$ and $|X_2(t, f)|^2$ are the spectrograms of the two components, respectively, where $2\Re[X_1(t, f)X_2^*(t, f)]$ could be referred to as a cross-term. However, it is only when $X_1(t, f)$ and $X_2^*(t, f)$ covers the same time- and frequency interval, the total density also involves the cross-term. This term is affected by different phases of the signal components as well as their magnitude functions and is most often called **leakage** but could also be defined as a cross-term. It is most prominent when the signal components are fairly close in the time-frequency plane, which is the main difference if we compare with the Wigner distribution, Figure 3.3c, where the cross-term always shows up independently of the distance between the components. For a thorough analysis and comparison, see [30].

When the signal stops for while and then starts again, i.e., if there is an interval in the signal that is zero, it does **not** imply that the Wigner distribution is zero in that time interval, as the cross-term between the components appear. The same applies to frequency intervals where the spectrum is zero, it does not imply that the Wigner distribution is zero in that frequency interval.

3.6 Discrete Wigner distribution

The discrete-time and discrete-frequency Wigner distribution is defined as

$$W_x[n, l] = 2 \sum_{m=-\min(n, N-1-n)}^{\min(n, N-1-n)} x_{n+m}x_{n-m}^* e^{-i2\pi m \frac{l}{L}}, \quad (3.19)$$

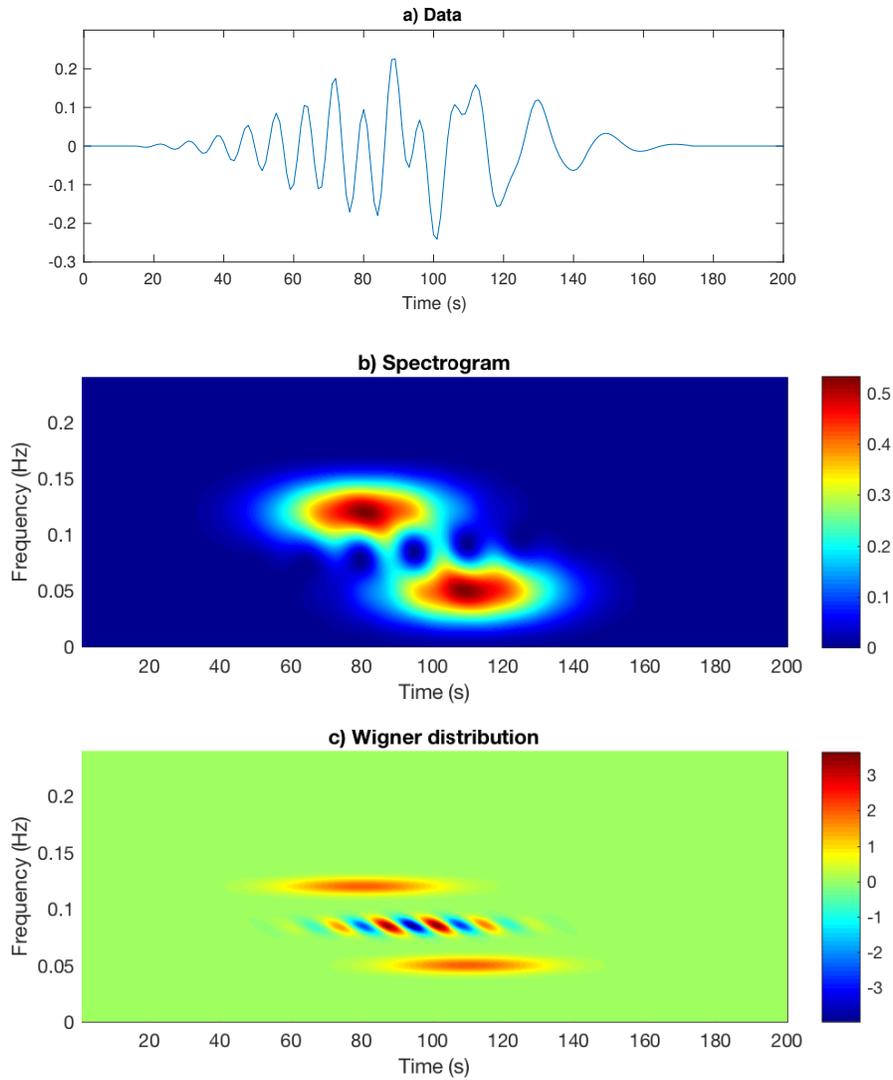


Figure 3.3: The spectrogram and the Wigner distribution of two closely spaced complex Gaussian windowed sinusoids; a) Data sequence (real part); b) Spectrogram with Hanning window of length $M=32$; c) Wigner distribution.

for a discrete-time signal x_n , $n = 0 \dots N - 1$, where L is the number of frequency values and $f = l/(2 \cdot L)$, [31]. The discrete Wigner distribution is not as reliable and predictable as the corresponding discrete spectrogram, which has caused a long-living debate on how the Wigner distribution should be computed for discrete-time signals, [32]. We illustrate the main issues with a simple example using a real-valued signal, in contrary to the figures previously shown. The spectrogram of a real-valued Gaussian windowed sinusoid, with frequency $f_0 = 0.3$, is shown in Figure 3.4a, where we clearly can identify the positive frequency component at $f = 0.3$ and also the negative frequency component at $f = -0.3$. We remember that we always get symmetrical spectra for real-valued signals, but usually we do not show the negative frequency axis. In Figure 3.4b the Wigner distribution of the same real-valued signal is visualized and the picture becomes more difficult to interpret. We see a component at $f_0 = 0.3$ and one at $f_0 = -0.3$ as expected and the expected cross-term between these two is placed at $f = 0$, in the middle between them. Cross-terms at $f = 0$ will always be present for all real-valued signals and is accordingly an issue when low-frequency parts of the signal are of certain interest.

More confusing is the two extra components found at $f \pm 0.2$. These components are in fact **aliasing**. We are used to expect aliasing when we sample a continuous-time signal with a sample frequency that is $f_s < 2f_{max}$ where f_{max} is the highest frequency of the signal. However, in the context of the Wigner distribution it is the calculation using the discrete Fourier transform in Eq. (3.19) that introduces the aliasing. The aliasing occurs for normalized signal frequencies above 0.25, i.e., for all discrete-time frequencies between 0.25 and 0.5, which in this example means that the frequency 0.3 is aliased around 0.25 and become 0.2.

A better understanding is given if we recall that aliasing is actually caused by periodic repetition of the frequency range. Usually, the periodicity is one but now the periodicity is 0.5, meaning that the picture including only the two components at $f = \pm 0.3$ and the cross-term at $f = 0$ is repeated and centered around $f = 0.5$ giving the cross-term located at 0.5 and the actual component located at $f = -0.3$ will now be found at $f = 0.2$ causing aliasing. Similarly if the picture is repeated around $f = -0.5$, where the $f = 0.3$ will give an aliased term at $f = -0.2$.

One elegant solution to all this problems, is to use the analytic signal in the calculation of the discrete Wigner distribution. We recall that the analytic signal does not include any negative frequencies, implying that the frequency component at $f = -0.3$ is no longer present and thereby the cross-term at $f = 0$ also disappears, Figure 3.5. The periodicity of 0.5 will cause the now empty frequency range of -0.5 to 0 to be repeated at 0 to 0.5 but as the component at frequency $f = -0.3$ does not exist it can not be repeated at $f = 0.2$, i.e., no resulting aliasing. This means that the signal

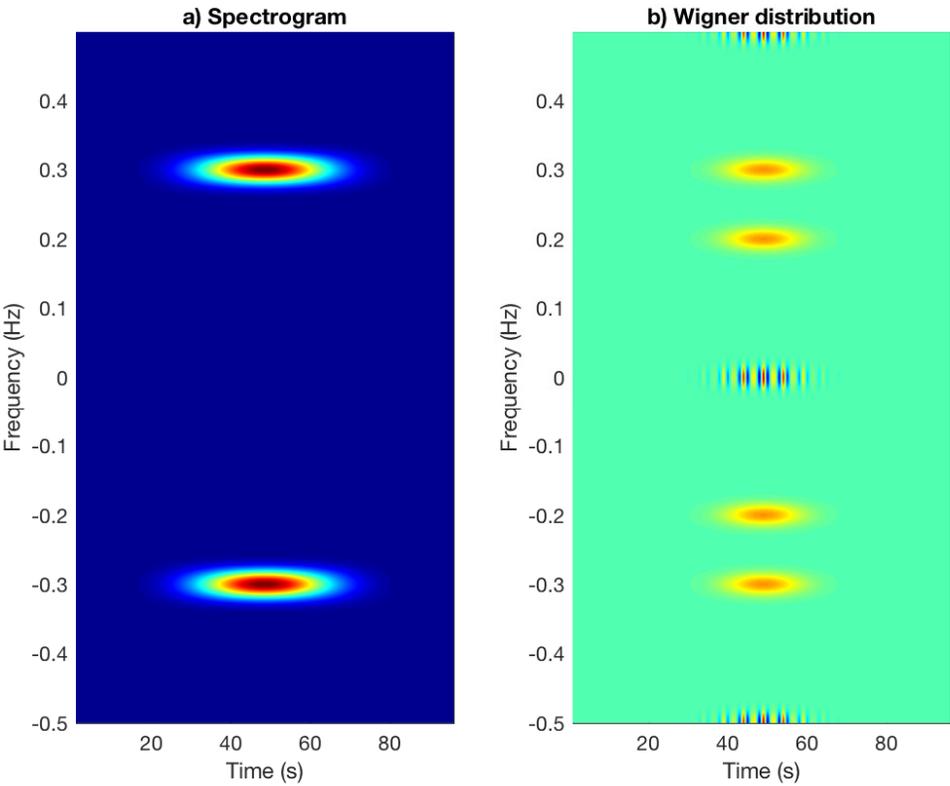


Figure 3.4: a) Spectrogram; b) Wigner distribution of a real-valued Gaussian windowed sinusoid, $f_0 = 0.3$.

frequency content of an analytic signal can now be up to $f = 0.5$ as we are used to handle for discrete-time signals. **Using the analytic signal in calculations avoids unnecessary cross-terms and expands the possible frequency range of the input signals.** Therefore, the analytic signal is always used in calculations of the Wigner distribution.

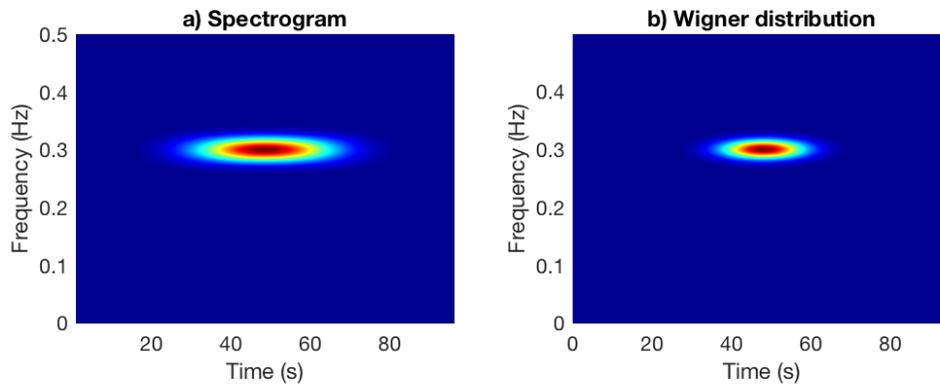


Figure 3.5: a) Spectrogram; b) Wigner distribution of the corresponding analytic signal.

Chapter 4

The ambiguity function and other representations

The **ambiguity** or **doppler-lag function** is the Fourier transform in both variables of the Wigner distribution. The ambiguity function has some nice properties, which are useful especially for cross-term reduction. However, if we make the Fourier transform in just one of the variables, two other functions are given, in total four domains.

The word "ambiguity" is a bit ambiguous as it then should stand for something that is not clearly defined. The name comes from the radar field and was introduced in [33], describing the equivalence between time-shifting and frequency-shifting for linear FM signals, which are used in radar. It was however first introduced by Jean-André Ville, [26], and by José Enrique Moyal, [34], who suggested the name **characteristic function**.

4.1 The four time-frequency domains

The four different domains for representation of a time-varying signal are presented, the time-lag domain in the variables (t, τ) , the time-frequency (Wigner) domain in (t, f) , the ambiguity (doppler-lag) domain in (ν, τ) and finally the doppler-frequency domain in (ν, f) . A schematic overview is given in Figure 4.1. Studying simple signals in the different domains give us information on the interpretation.

We can recall that a Gaussian signal without any oscillation, i.e., located at $f = 0$ and also centered at $t = 0$ shows up as a 2-dimensional Gaussian function, located at origin in all other domains. A shift of the Gaussian envelope signal in time to $t = 20$ and a oscillation in frequency of $f = 0.2$, gives a complex-valued signal visualized in the time-lag domain in Figure 4.2a (real part). The figure shows the time-shift location and the frequency as an oscillation in the direction of the lag-variable τ . In the ambiguity

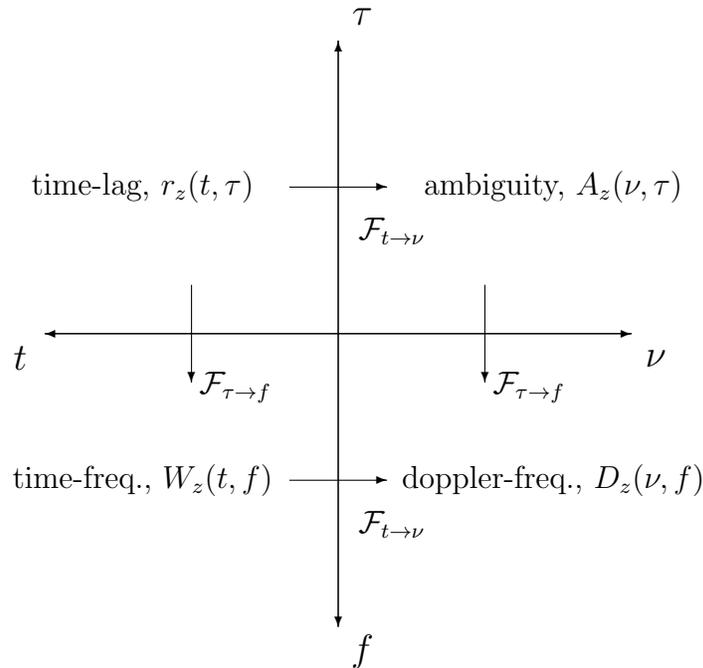


Figure 4.1: The four possible domains of time-frequency analysis.

domain, Figure 4.2b, (real part), the combination of time- and frequency shifts show up as interfering oscillations of the function located at origin both in τ and ν . The view of the ambiguity domain is not intuitive but the Fourier transforms from τ to f and ν to t into the time-frequency domain gives the interpretable Wigner distribution visualized in Figure 4.2c where the doppler-frequency domain in Figure 4.2d shows the frequency shift location but the time-location as an oscillation in the direction of the ν -variable.

For a multicomponent signal, a combination of a Gaussian envelope signal at center time $t = 20$ and center frequency $f = 0.3$, and one at center time $t = 40$ and center frequency $f = 0.1$, a more complex pattern is seen in all domains of Figure 4.3. In the time-frequency domain, Figure 4.3c, the two functions are found at their specific locations in time and frequency with the cross-term in between. A Fourier transform in the t -variable to the ν -variable will give us a Gaussian envelope with oscillations between negative and positive values. The cross-term also shows up in the other domains at various places. In the time-lag domain, Figure 4.3a, the auto-terms are located at $\tau = 0$ where the cross-term shows up both at $\tau = -20$ and $\tau = 20$. Similarly in the for doppler-frequency domain, Figure 4.3d, the auto-terms are located at $\nu = 0$ and the cross-term at $\nu = -0.2$ and $\nu = 0.2$. Viewing the ambiguity function,

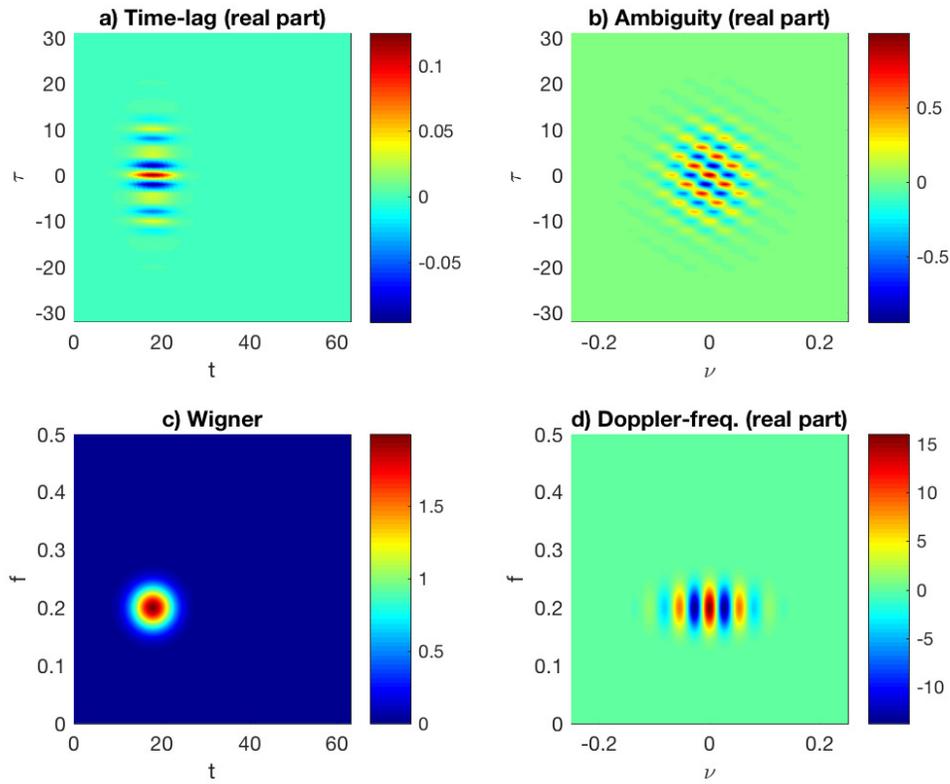


Figure 4.2: A Gaussian windowed signal, centered at $f = 0.2$ and $t = 20$ and the representation in all four domains; a) Time-lag domain (real part); b) Ambiguity domain (real part); c) Wigner domain; d) Doppler-frequency domain (real part).

Figure 4.3b, these properties can be used to differ auto-terms from cross-terms as the auto-terms always are located at $\tau = \nu = 0$ and the cross-terms at positive and negative values of τ and ν . This property is used to design other distributions with reduced cross-term contribution.

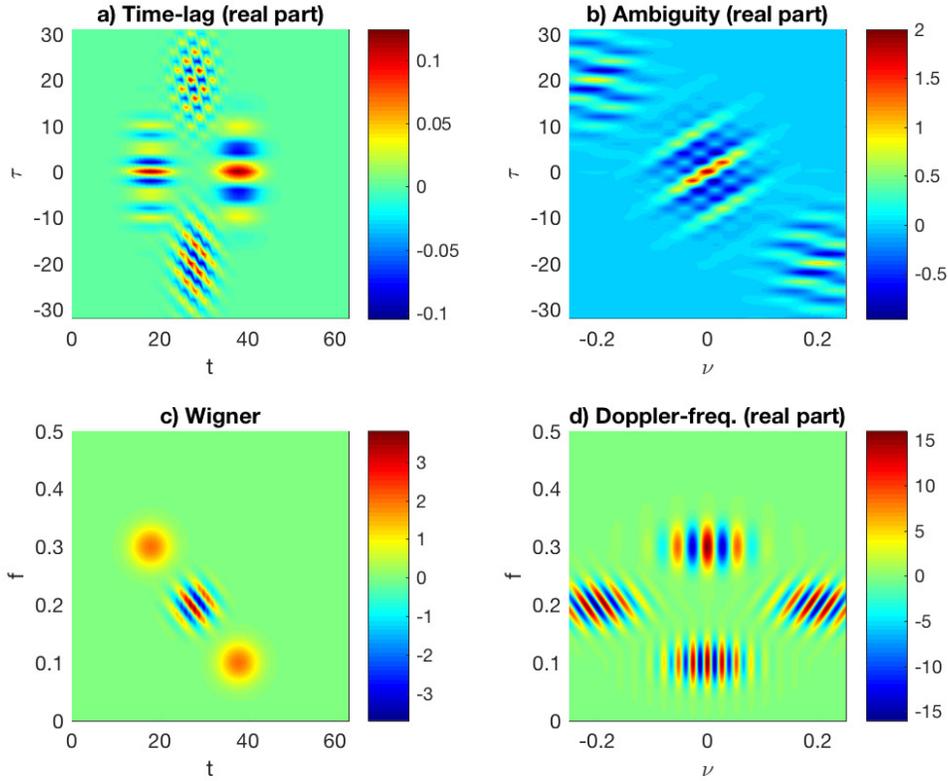


Figure 4.3: A multi-component with two Gaussian signals, one at $f = 0.3$, $t = 20$ and one at $f = 0.1$, $t = 40$ and the representation in all four domains; a) Time-lag domain (real part); b) Ambiguity domain (real part); c) Wigner domain; d) Doppler-frequency domain (real part).

4.2 Ambiguity function

The ambiguity function $A_z(\nu, \tau)$ is defined as

$$A_z(\nu, \tau) = \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-i2\pi\nu t} dt, \quad (4.1)$$

where usually the analytic signal $z(t)$ is used, although, without any restrictions, $z(t)$ could be replaced by $x(t)$. We note that the formulation is similar to the Wigner distribution, the difference is that the Fourier transform now is made in the t -variable instead of the τ -variable, giving the ambiguity function dependent of the two variables ν and τ . Similarly, the Fourier transform of the IAF in the variable t gives the **ambiguity spectrum**,

$$A_z^s(\nu, \tau) = \int_{-\infty}^{\infty} r_z(t, \tau) e^{-i2\pi\nu t} dt. \quad (4.2)$$

To learn how to use the ambiguity function for analysis, we compare the view of the Wigner distribution and the ambiguity function for some signal examples. Here we visualize the **absolute value** of the ambiguity function in the figures. In the first case presented in Figure 4.4a and b we see that for a Gaussian envelope signal centered at time $t_1 = 20$ and at frequency $f_1 = 0.3$, the ambiguity function will be located at $\tau = 0$ and $\nu = 0$. The frequency- and time-shift will show up as the oscillation frequency and direction of the oscillations but this information is not visualized in the absolute value. (Compare with the real-value visualized in Figure 4.2b. In Figure 4.4c, the Wigner distribution of the Gaussian envelope analytic signal located at $t_2 = 40$ with frequency $f_2 = 0.1$ is depicted and in Figure 4.4d the corresponding ambiguity function is located at the centre. It can easily be shown that the ambiguity function of a time- and frequency shifted mono-component signal will always relocate to $\tau = 0$ and $\nu = 0$ with identical absolute value for all time- and frequency shifts.

The great advantage of the ambiguity function shows up in the last example, Figure 4.4e and f, where the signal now consists of the sum of the two Gaussian components. The two signal components and the cross-term of the Wigner distribution are clearly visible in Figure 4.4e, where we find the cross-term located in the middle between the auto-terms, at time location $(t_1 + t_2)/2$ and frequency location $(f_1 + f_2)/2$ as discussed before. In the ambiguity function, Figure 4.4f, the signal components are summed at the centre, and the cross-term(s) show up located at doppler frequencies $\nu_1 = f_2 - f_1 = -0.2$ and $\nu_2 = f_1 - f_2 = 0.2$ and at lags $\tau_1 = t_2 - t_1 = 20$ and $\tau_2 = t_1 - t_2 = -20$. The cross-terms will always be located away from the centre, and also located further away if the time-frequency distance between the actual signal com-

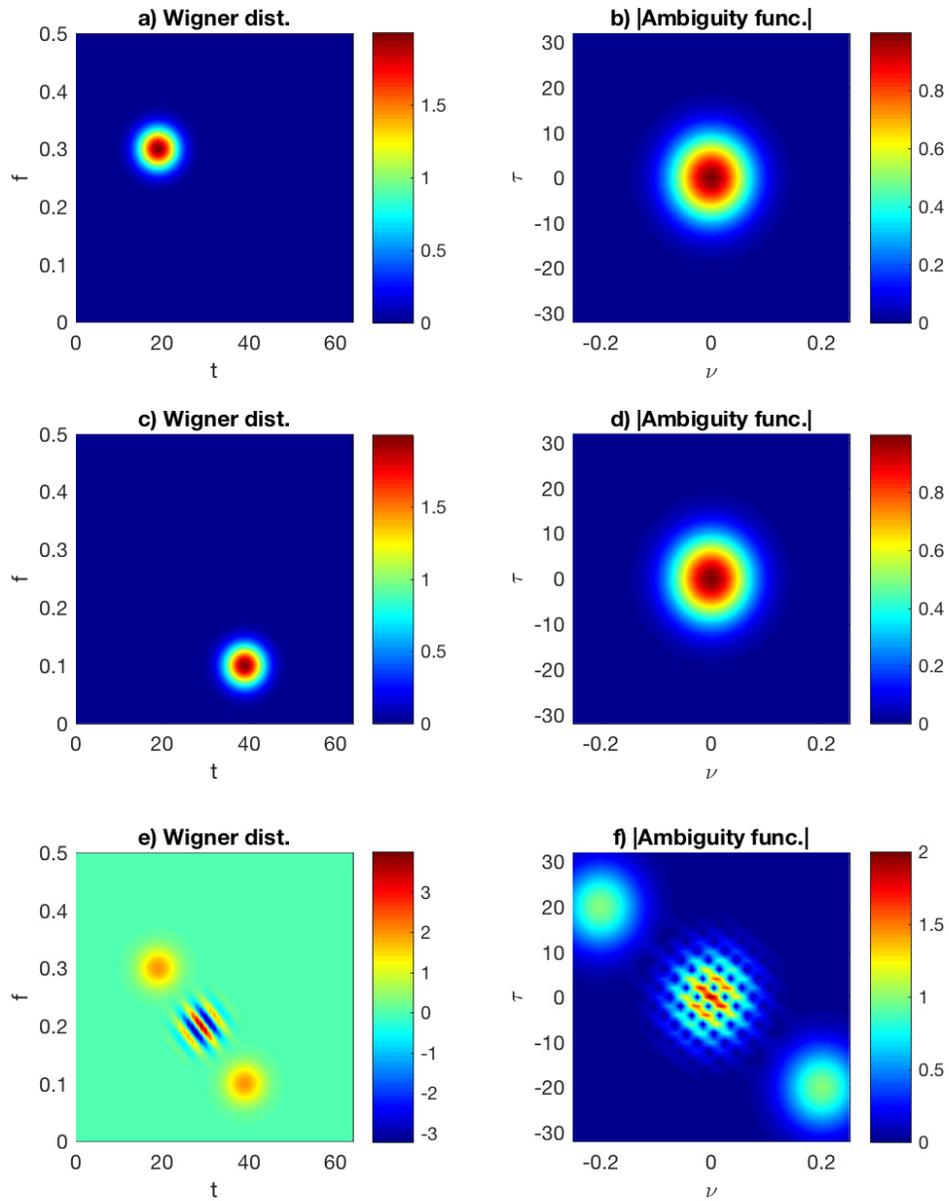


Figure 4.4: Different examples of Gaussian windowed complex-valued signals and their Wigner distributions and ambiguity functions (absolute value); a) and b) one Gaussian component with $t_1 = 20$ and $f_1 = 0.3$; c) and d) one Gaussian component with $t_2 = 40$ and $f_2 = 0.1$; e) and f) the sum of the two Gaussian components.

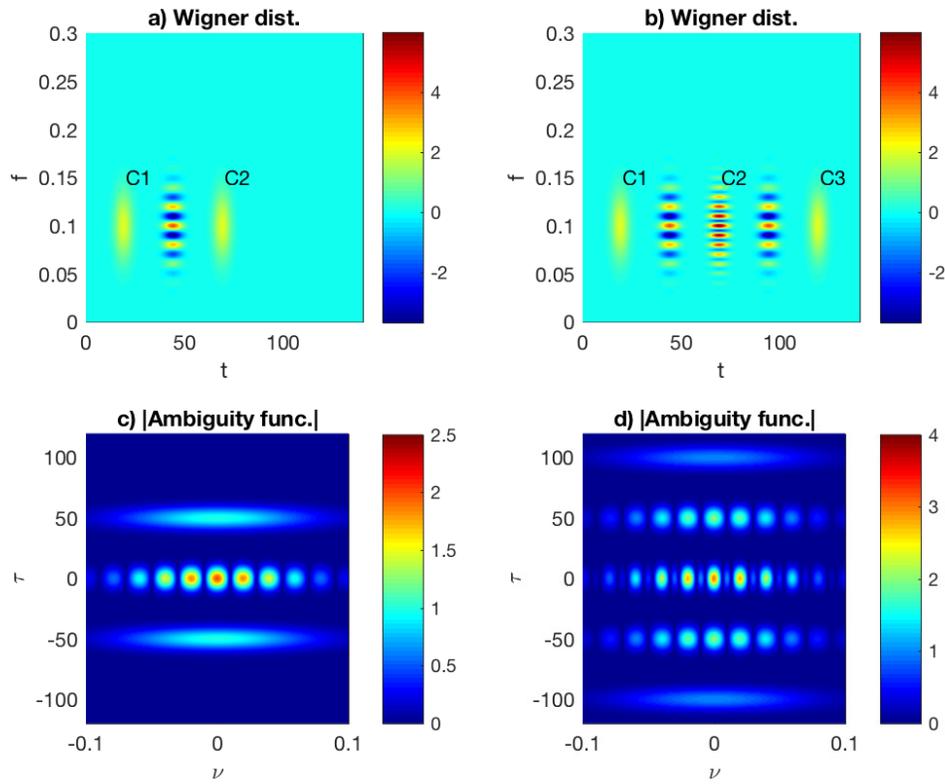


Figure 4.5: Comparison of the ambiguity function of one two-component and one three-component signal; a) and b) the Wigner distributions; c and d) the ambiguity functions (absolute values).

ponents is increased. A natural approach based on these findings is to keep the wanted auto-terms located at the centre and reduce the unwanted cross-term components located away from the centre of the ambiguity function by designing an appropriate function, a kernel, in the ambiguity domain.

Before proceeding to kernel design, we study another example where the signal consists of two Gaussian enveloped analytic components with $t_1 = 20$, $t_2 = 70$ and $f_1 = f_2 = 0.1$ ($C1$ and $C2$), see the Wigner distribution in Figure 4.5a with one cross-term in between the two auto-terms, which we compare with a three-component signal with an additional component located at $t_3 = 120$ and $f_3 = 0.1$, visualized in Figure 4.5b ($C1$, $C2$ and $C3$). If we study their corresponding ambiguity functions, Figures 4.5c and d respectively, we clearly see what we learned from the previous example, that the cross-terms of the distance $\Delta t = 50$ are located at $\tau = \pm 50$. For the second example, the cross-term between the outermost components $C1$ and $C3$ of Figure 4.5b is found at locations $\tau = \pm 100$ in Figure 4.5d. If we only study the middle parts of Figures 4.5c and d, the sum of the auto-terms, which are the ones we usually keep, we can note that also in the auto-term pattern located close to $\nu = \tau = 0$, the time-difference between $C1$ and $C2$ is visualized as small components with distance $\Delta\nu = 0.02 = 1/50$ in Figure 4.5c, a pattern that is repeated in Figure 4.5d. The time-difference 100 for the outermost components $C1$ and $C3$ in Figure 4.5b is only reflected as a weak pattern with components at $\Delta\nu = 0.01 = 1/100$.

This similarity of auto-term pattern, irrespectively of the number of components, has been used e.g. for clustering of bird song syllables in [35], where the ambiguity function is used for further extraction of features and has shown to be an excellent tool for clustering of complex bird song syllables. An example two bird song syllables, one with five components and one with six components, is visualized in Figures 4.6a and b. These two syllables should belong to the same class, irrespectively that they have five or six components. It is seen that the time-frequency distributions (spectrograms), Figures 4.6c and d, are more different than the ambiguity functions, Figures 4.6e and f.

Contrary to the similarity of the ambiguity domain auto-terms, the ambiguity domain cross-term patterns are clearly different and if the aim is to emphasize and actually classify small differences of the signals, the cross-term pattern can be used for classification, [36, 37].

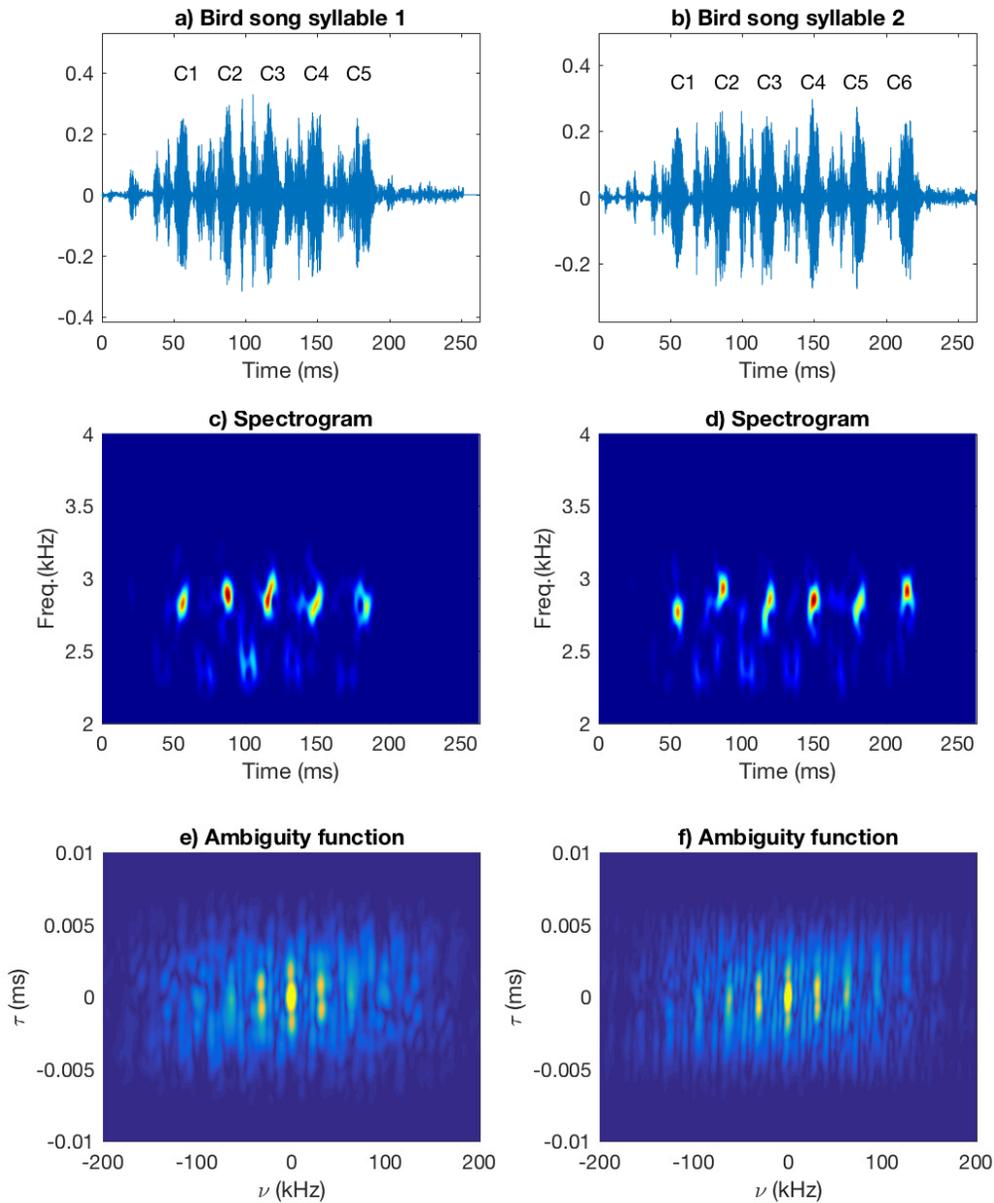


Figure 4.6: Comparison of the spectrogram and ambiguity function of; a) a 5-component bird syllable; b) a 6-component bird syllable; c) and d) the spectrograms; e and f) the ambiguity functions (absolute values).

4.3 Doppler-frequency distribution

The **doppler-frequency distribution** is found from

$$\begin{aligned} D_z(\nu, f) &= \int_{-\infty}^{\infty} A_z(\nu, \tau) e^{-i2\pi f\tau} d\tau \\ &= \int_{-\infty}^{\infty} W_z(t, f) e^{-i2\pi\nu t} dt, \end{aligned} \quad (4.3)$$

or from

$$D_z(\nu, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r_z(t, \tau) e^{-i2\pi(f\tau + \nu t)} dt d\tau. \quad (4.4)$$

The doppler-frequency distribution is also referred to as the **spectral autocorrelation function**. Using the change of variables $t + \tau/2 = t_1$ and $t - \tau/2 = t_2$ and reformulate Eq. (4.4) as

$$\begin{aligned} D_z(\nu, f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t + \frac{\tau}{2}) z^*(t - \frac{\tau}{2}) e^{-i2\pi(f\tau + \nu t)} dt d\tau, \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t_1) z^*(t_2) e^{-i2\pi(f(t_1 - t_2) + \nu(t_1 + t_2)/2)} dt_1 dt_2, \\ &= \int_{-\infty}^{\infty} z(t_1) e^{-i2\pi(f + \frac{\nu}{2})t_1} dt_1 \cdot \int_{-\infty}^{\infty} z^*(t_2) e^{i2\pi(f - \frac{\nu}{2})t_2} dt_2 \\ &= Z(f + \frac{\nu}{2}) Z^*(f - \frac{\nu}{2}), \end{aligned} \quad (4.5)$$

where $Z(f)$ is the Fourier transform of $z(t)$. We now see that the doppler-frequency distribution is the frequency dual of $r_z(t, \tau) = z(t + \tau/2) z^*(t - \tau/2)$.

Chapter 5

Ambiguity kernels and the quadratic class

A large number of time-frequency estimation methods can be found in research literature, almost all with the aim at reducing the Wigner distribution cross-terms. After the invention of the spectrogram and the Wigner distribution, a lot of other distributions and methods were suggested, such as the Rihaczek, Page and Levin distributions, [38, 39, 40]. In 1966, Leon Cohen suggested that most of these methods actually could be viewed in the same framework, a framework that is known as the **Cohen's class**, **quadratic class** or **bi-linear class**, [41]. The formulation by Cohen was initially restricted with constraints on the distributions to fulfill the marginals, where the quadratic class also includes methods which do not fulfill the marginals. On the other hand, Cohen's class also included signal-dependent kernels which cannot be members of the quadratic class. The aim is to design **time-frequency kernels** which are usually defined and optimized in the ambiguity domain.

5.1 Ambiguity kernel

A **filtered** ambiguity function is defined as the element-wise multiplication of the ambiguity function and an **ambiguity kernel**, $\phi(\nu, \tau)$,

$$A_z^Q(\nu, \tau) = A_z(\nu, \tau) \cdot \phi(\nu, \tau), \quad (5.1)$$

where $A_z(\nu, \tau)$ is defined in Eq. (4.1) and $A_z^Q(\nu, \tau)$ represents any member of the **quadratic class** of distributions. The corresponding **time-frequency kernel** is given

by,

$$\Phi(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\nu, \tau) e^{-i2\pi(f\tau - \nu t)} d\tau d\nu, \quad (5.2)$$

and the corresponding **smoothed** Wigner distribution is then found as the 2-dimensional convolution

$$W_z^Q(t, f) = W_z(t, f) * \Phi(t, f), \quad (5.3)$$

which can be expanded into

$$\begin{aligned} W_z^Q(t, f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_z(\nu, \tau) \phi(\nu, \tau) e^{-i2\pi(f\tau - \nu t)} d\tau d\nu, \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(u + \frac{\tau}{2}) z^*(u - \frac{\tau}{2}) \phi(\nu, \tau) e^{i2\pi(\nu t - f\tau - \nu u)} du d\tau d\nu, \end{aligned} \quad (5.4)$$

using Eq. (4.1). This form is the most recognized defining the quadratic class. We can note that the original Wigner distribution has the simple ambiguity kernel $\phi(\nu, \tau) = 1$ for all ν and τ , and the corresponding time-frequency (non)smoothing kernel $\Phi(t, f) = \delta(t)\delta(f)$.

5.2 Properties of the ambiguity kernel

To learn about the design properties of the ambiguity kernel we set $\tau = 0$ in Eq. (4.1) and get

$$A_z(\nu, 0) = \int_{-\infty}^{\infty} z(t) z^*(t) e^{-i2\pi\nu t} dt = \int_{-\infty}^{\infty} |z(t)|^2 e^{-i2\pi\nu t} dt, \quad (5.5)$$

which actually is the Fourier transform of the time marginal given in Eq. (3.8). Similarly, with $\nu = 0$ and using the Fourier transform of $z(t)$, $Z(f) = \int_{-\infty}^{\infty} z(t) e^{-i2\pi t f} dt$, we can reformulate Eq. (4.1) into

$$A_z(0, \tau) = \int_{-\infty}^{\infty} Z(f) Z^*(f) e^{i2\pi\tau f} df = \int_{-\infty}^{\infty} |Z(f)|^2 e^{i2\pi\tau f} df, \quad (5.6)$$

which is recognized as the inverse Fourier transform of the frequency marginal given in Eq. (3.9). The frequency marginal is equal to the usual spectral density and the inverse

Fourier transform is then the usual covariance function, i.e. the ambiguity function at the doppler axis $\nu = 0$ is the covariance function.

Conclusively, in order to preserve the time- and frequency marginals of the time-frequency distribution, the ambiguity kernel must fulfill

$$\phi(0, \tau) = \phi(\nu, 0) = 1, \quad (5.7)$$

meaning that the kernel must be one at both axes, independently of how it is defined for other τ and ν . We can also see that

$$\phi(0, 0) = 1, \quad (5.8)$$

will preserve the total energy E_x of the signal.

The Wigner distribution is always real-valued and for the smoothed time-frequency distribution to also become real-valued, the ambiguity kernel must fulfill the Hermitian property,

$$\phi(\nu, \tau) = \phi^*(-\nu, -\tau). \quad (5.9)$$

We leave the proof as an exercise for the reader!

Time-invariance and frequency-invariance properties are important and to find the restrictions of the ambiguity kernel we start with the expression in Eq. (5.4) using a time- and frequency-shifted signal $y(t) = z(t - t_0)e^{i2\pi f_0 t}$. We get

$$\begin{aligned} W_y^Q(t, f) &= \int \int \int y(u + \frac{\tau}{2})y^*(u - \frac{\tau}{2})\phi(\nu, \tau)e^{i2\pi(\nu t - f\tau - \nu u)} du d\tau d\nu, \\ &= \int \int \int z(u + \frac{\tau}{2} - t_0)e^{i2\pi f_0(u + \frac{\tau}{2})} z^*(u - \frac{\tau}{2} - t_0)e^{-i2\pi f_0(u - \frac{\tau}{2})} \dots \\ &\quad \dots \cdot \phi(\nu, \tau)e^{i2\pi(\nu t - f\tau - \nu u)} du d\tau d\nu, \\ &= \int \int \int z(u_1 + \frac{\tau}{2})z^*(u_1 - \frac{\tau}{2})\phi(\nu, \tau)e^{i2\pi(\nu(t-t_0) - (f-f_0)\tau - \nu u_1)} du_1 d\tau d\nu, \\ &= W_z^Q(t - t_0, f - f_0), \end{aligned} \quad (5.10)$$

where all integrals range from $-\infty$ to ∞ . The final step relies on that the kernel $\phi(\nu, \tau)$ is not a function of time nor of frequency and leads to the conclusion that the

quadratic distribution is time-shift invariant if the kernel is independent of time and frequency-shift invariant if the kernel is independent of frequency.

The Wigner distribution is known not to have strong time- and frequency support as cross-terms arise in between signal components of a multi-component signal. However, other members of the quadratic class are defined to have strong support. The restrictions on the ambiguity kernel for strong time support is

$$\int_{-\infty}^{\infty} \phi(\nu, \tau) e^{-i2\pi\nu t} d\nu = 0, \quad |\tau| \neq 2|t|, \quad (5.11)$$

and similarly strong frequency support implies that

$$\int_{-\infty}^{\infty} \phi(\nu, \tau) e^{-i2\pi\tau f} d\tau = 0, \quad |\nu| \neq 2|f|. \quad (5.12)$$

For proofs see [42].

5.3 The Choi-Williams distribution

The **Choi-Williams distribution**, [43], also called the exponential distribution, (ED), is perhaps the most applied distribution with ambiguity kernel defined as

$$\phi_{ED}(\nu, \tau) = e^{-\frac{\nu^2\tau^2}{\sigma}}, \quad (5.13)$$

where σ is a design parameter. The Choi-Williams kernel is symmetric in ν as well as τ so the property of Eq. (5.9) is fulfilled. It is also a **product kernel**, meaning that the dependency is in one dimension, $x = \nu\tau$, and thereby the design parameter set only includes one variable, σ . This is a great advantage, especially when optimizing the kernel for a certain performance of a given signal. The resulting smoothed time-frequency spectra from the Choi-Williams distribution is visualized in Figure 5.1 where three different parameter values of the Choi-Williams kernel are visualized for a two-component signal. In Figure 5.1a, $\sigma = 0.01$, producing a kernel that quickly approaches zero also for small absolute values of ν and τ .

Note that the kernel is one for all ν and τ at both of the axes as $\phi_{ED}(\nu, 0) = \phi_{ED}(0, \tau) = 1$ and accordingly the marginals are fulfilled. The resulting smoothed Wigner distribution in Eq. (5.3) is depicted in Figure 5.1b, where we see that the auto-terms are smeared. The reason is that the ambiguity kernel is too narrow, with resulting suppression of the auto-terms (located at $\tau = 0, \nu = 0$). Choosing $\sigma = 1$ results in the kernel presented in Figure 5.1c, and the resulting smoothed Wigner distribution in

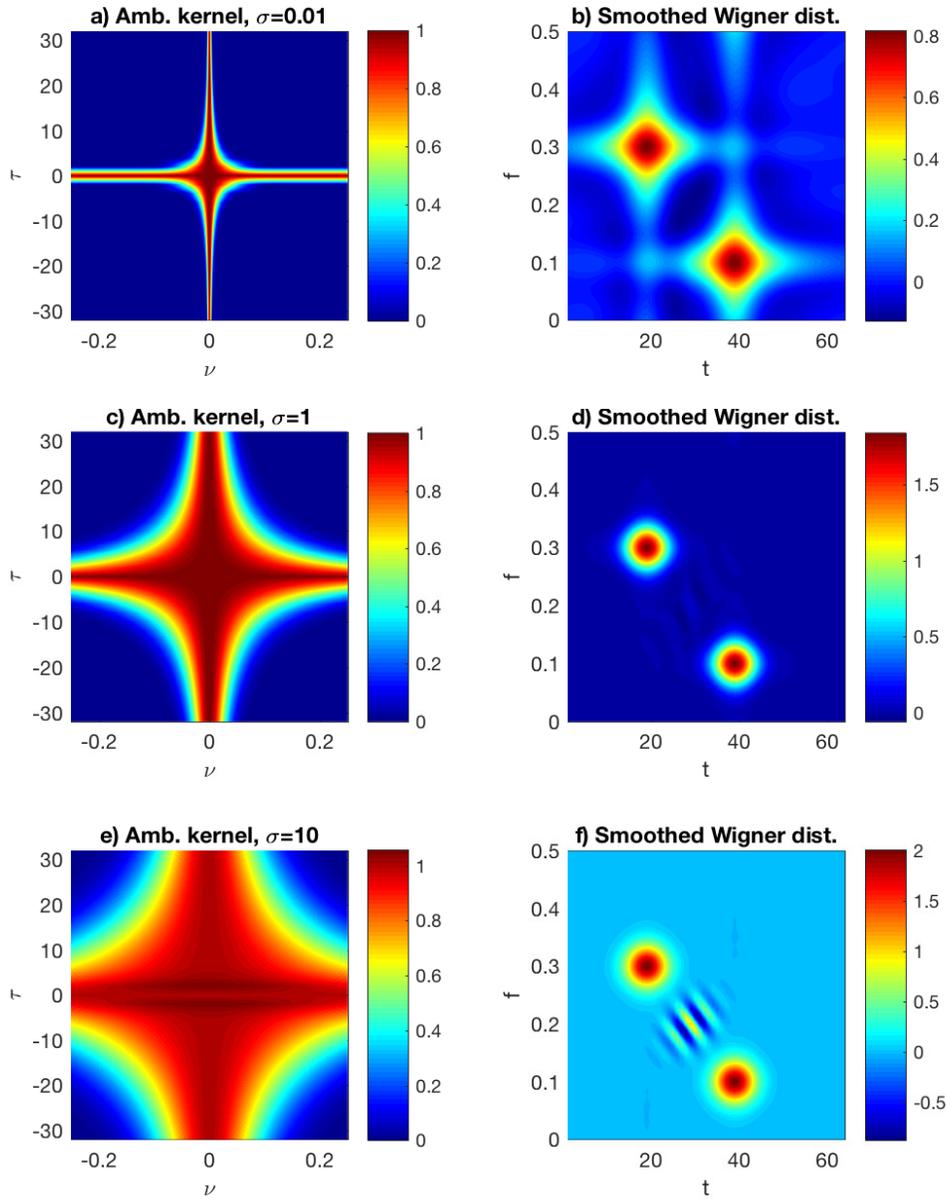


Figure 5.1: The ambiguity kernel and the corresponding Choi-Williams distribution next to each other for different choices of the parameter σ applied to a two-component Gaussian windowed complex-valued signal with $t_1 = 20$, $f_1 = 0.3$ and $t_2 = 40$, $f_2 = 0.1$; a) and b) $\sigma = 0.01$; c) and d) $\sigma = 1$; e) and f) $\sigma = 10$.

Figure 5.1d. The Gaussian shape and concentration of the auto-terms is now better preserved, where the cross-term still is reduced. The choice of σ is crucial and choosing a too large value, $\sigma = 10$ with the ambiguity kernel given as in Figure 5.1e will result in a smoothed Wigner distribution where the cross-term becomes more prominent, Figure 5.1f.

Other suppressions of cross-terms have been invented, e.g. the Born-Jordan distribution, also called the sinc-distribution, derived by Cohen, [41]. The ambiguity kernel for the Born-Jordan distribution is

$$\phi_{BJ}(\nu, \tau) = \text{sinc}(a\nu\tau) = \frac{\sin(\pi a\nu\tau)}{\pi a\nu\tau}, \quad (5.14)$$

which also fulfills the marginal properties and the Hermitian property. The properties of this kernel were not fully understood until in the 1990s, with the work of Jeong and Williams, [44]. Both these kernels have been widely applied in many different areas and are also sometimes referred to as reduced interference distributions (RID) kernels.

There is one major problem with kernels that satisfies the marginal properties, which we illustrate with another example. The Wigner distribution of a signal consisting of three Gaussian complex-valued components located at $t_1 = 20$, $f_1 = 0.1$, $t_2 = 40$, $f_2 = 0.1$ and $t_3 = 20$, $f_3 = 0.3$ is shown in Figure 5.2a.

The corresponding absolute value of the ambiguity function is depicted in Figure 5.2b, where the auto-terms all are located in the middle and the cross-terms are located away from the centre at different positions corresponding to the relative locations of the auto-terms. In Figure 5.2c and d the resulting smoothed Wigner distribution and filtered ambiguity function of the Choi-Williams distribution using $\sigma = 1$, i.e., the same kernel as depicted in Figure 5.1c. The kernel suppresses the cross-term between component F2 and F3 as this cross-term is localized at $\nu = 0.2$, $\tau = -20$ and $\nu = -0.2$, $\tau = 20$ in the ambiguity function, but the cross-term between F1 and F2, which will be located at the axis $\nu = 0$ at $\tau = \pm 20$ will not be fully suppressed as the kernel is one on the ν -axis. A similar result is given for the cross-term between F1 and F3 which will be located at the axis $\tau = 0$ at $\nu = \pm 0.2$. The cross-terms in the direction of the axes are suppressed but not so extensively as the cross-term between F2 and F3. This is a drawback of the kernels that fulfill the marginals. However, if energy preservation is not necessary in the representation other more efficient kernels can be designed.

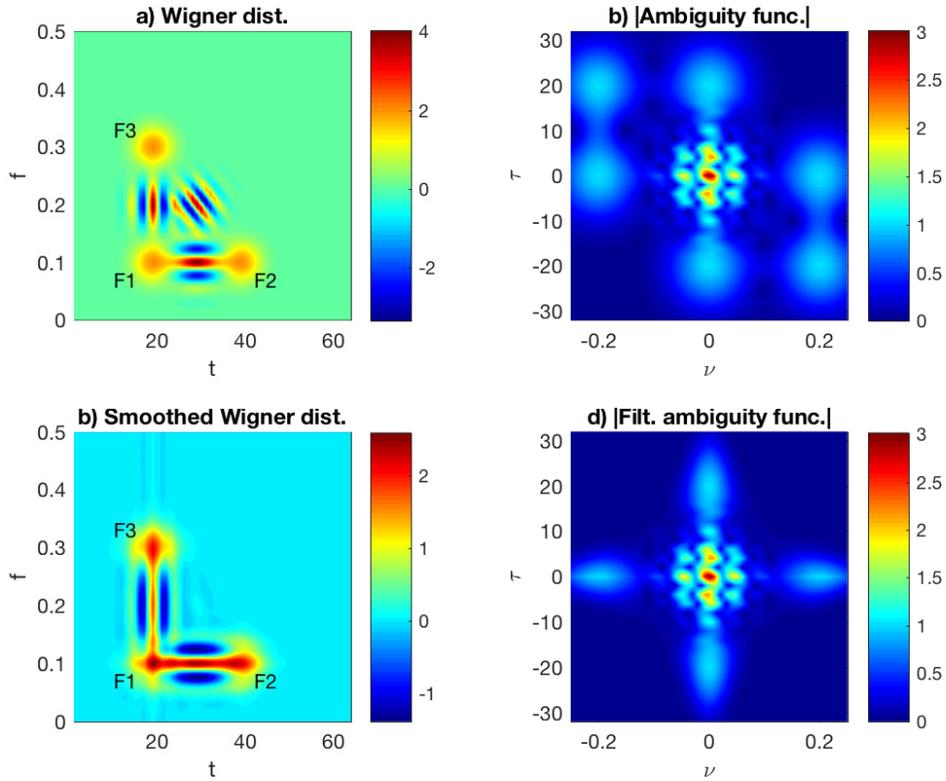


Figure 5.2: A signal of three Gaussian windowed complex-valued components; a) The Wigner distribution; b) The absolute value of the ambiguity function; c) The Choi-Williams distribution using $\sigma = 1$; d) The absolute value of the corresponding filtered ambiguity function.

5.4 Separable kernels

Another nice form of useful kernels are the **separable** kernels defined by

$$\phi(\nu, \tau) = G_1(\nu)g_2(\tau). \quad (5.15)$$

This form transfers easily to the time-frequency domain as $\Phi(t, f) = g_1(t)G_2(f)$ with $g_1(t) = \mathcal{F}^{-1}\{G_1(\nu)\}$ and $G_2(f) = \mathcal{F}\{g_2(\tau)\}$. The quadratic time-frequency formulation becomes

$$W_z^Q(t, f) = g_1(t) * W_z(t, f) * G_2(f), \quad (5.16)$$

as

$$A_z^Q(\nu, \tau) = G_1(\nu)A_z(\nu, \tau)g_2(\tau). \quad (5.17)$$

The separable kernel replaces the 2-D convolution of the quadratic time-frequency representation with two 1-D convolutions, which might be beneficial for some signals. Two special cases can be identified: if

$$G_1(\nu) = 1, \quad (5.18)$$

a **doppler-independent** kernel is found as $\phi(\nu, \tau) = g_2(\tau)$, and the resulting smoothed time-frequency distribution reduces to

$$W_z^Q(t, f) = W_z(t, f) * G_2(f), \quad (5.19)$$

which is a smoothing restricted only to the frequency direction. The doppler-independent kernel is also given the name **Pseudo-Wigner** or **windowed Wigner** distribution. The second case is when

$$g_2(\tau) = 1, \quad (5.20)$$

resulting in the **lag-independent** kernel, $\phi(\nu, \tau) = G_1(\nu)$ where the time-frequency formulation is a smoothing only in the variable t ,

$$W_z^Q(t, f) = g_1(t) * W_z(t, f). \quad (5.21)$$

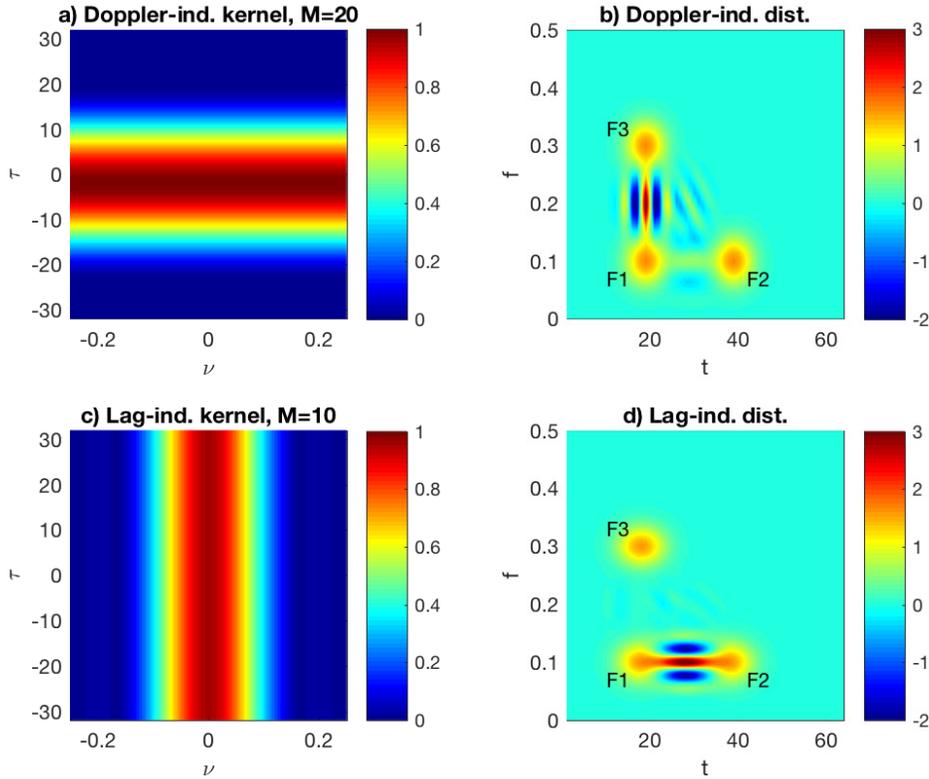


Figure 5.3: A signal of three Gaussian windowed complex-valued components; a) The doppler-independent kernel, $M = 20$; b) The doppler-independent distribution; c) The lag-independent kernel, $M=10$; d) The lag-independent distribution.

A comparison of the results of these two kernels for the previous example signal consisting of three Gaussian windowed components is presented in Figure 5.3. The doppler-independent kernel using a Hanning window of length $M = 20$ is shown in Figure 5.3a and the resulting frequency smoothed Wigner distribution in Figure 5.3b. It is clearly seen that the alternating cross-term between F1 and F2 as well as the one between F2 and F3 are reduced. An appropriate length of a window can be chosen related to the time-distance between the auto-term components as the resulting cross-term are located at the corresponding distance of τ . Compare the kernel in Figure 5.3a with the ambiguity function in see Figure 5.2b and note that the suppression will effect

the cross-terms located at $\tau = \pm 20$ but not the cross-terms on the τ -axis. Therefore the cross-term between F1 and F3 is almost unaffected.

In Figure 5.3c a lag-independent kernel using a Hanning window of length $M = 10$ is applied. The window length is no longer directly coupled to the view of the kernel as the window is applied in time, which is then Fourier transformed to ν . The resulting ambiguity kernel, Figure 5.3c, is interpretable anyhow. The frequency distance of the components F1 and F3 gives cross-terms located at the corresponding position in $\nu = \pm 0.2$, Figure 5.2b, and if the ambiguity kernel is more narrow in the variable ν , the cross-terms between F1 and F3 as well as F2 and F3, will be reduced. Similarly as for the previous example, cross-terms located on the ν -axis, i.e. the one between F1 and F2, will be preserved. The choice of Hanning windows applied here are just for illustration. Any other window, with certain properties regarding mainlobe width and sidelobe suppression, could be applied.

5.5 The Rihaczek distribution

The **Rihaczek distribution (RD)**, [38], also called the Kirkwood-Rihaczek distribution as it actually was derived much earlier in the context of quantum mechanics, [45], is derived and identified backwards from the total energy of a complex-valued deterministic signal according to

$$E = \int_{-\infty}^{\infty} |z(t)|^2 dt = \int_{-\infty}^{\infty} z(t)z^*(t) dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t)Z^*(f)e^{-i2\pi ft} df dt, \quad (5.22)$$

using the inverse Fourier transform of $z^*(t)$ in the last step. Similar to Eq. (3.10), where the total energy of the Wigner distribution is calculated, the RD is defined from the expression inside the double integral, i.e.,

$$W_z^{RD}(t, f) = z(t)Z^*(f)e^{-i2\pi ft}. \quad (5.23)$$

The RD distribution has both strong time support and strong frequency support. This is easily seen as the distribution is zero at the time intervals where $z(t)$ is zero and at the frequency intervals where $Z(f)$ is zero. We also note that the marginals are satisfied as

$$\int_{-\infty}^{\infty} W_z^{RD}(t, f) df = \int_{-\infty}^{\infty} z(t)Z^*(f)e^{-i2\pi ft} df = |z(t)|^2, \quad (5.24)$$

and

$$\int_{-\infty}^{\infty} W_z^{RD}(t, f) dt = Z^*(f) \int_{-\infty}^{\infty} z(t) e^{-i2\pi ft} dt = |Z(f)|^2. \quad (5.25)$$

The Fourier transform in two variables of Eq. (5.23) gives the ambiguity function from where the ambiguity kernel

$$\phi_{RD}(\nu, \tau) = e^{-i\pi\nu\tau}, \quad (5.26)$$

can be identified. We easily verify that the marginals are satisfied from the ambiguity kernel and we also note that the Rihaczek kernel is complex-valued, where the real and imaginary parts are shown in Figure 5.4a and b, respectively. Accordingly the Hermitian property is not fulfilled as both the real and the imaginary parts have positive symmetry, i.e. $\phi_{RD}(\nu, \tau) = \phi_{RD}(-\nu, -\tau)$, compare with Eq. (5.9). The absolute value $|\phi_{RD}(\nu, \tau)| = 1$ for all values of ν and τ and sometimes this type of kernels are referred to as **phase kernels**.

We return to the signal example presented in Figure 4.4, two complex-valued Gaussian functions, centered at times $t_1 = 20$ and $t_2 = 40$ with frequencies $f_1 = 0.3$ and $f_2 = 0.1$. In Figure 5.4c the absolute value of the RD is depicted resulting in four components located in a square. This confusing result is intuitively easy to understand if we study Eq. (5.23) where the signal $z(t)$ will be present just locally around time $t = 20$ and 40 and the Fourier transform $Z(f)$ will also be present locally around frequencies $f = 0.1$ and 0.3 . These two functions are combined to the two-dimensional time-frequency representation, which naturally then will have components appearing at combinations of exactly these time and frequency instants. However, to differ between auto-terms, the actual components, and the cross-terms is not an easy task from this view. We study the real-valued part of the distribution in Figure 5.4d, where we see that the auto-terms show up as expected at $t_1 = 20$, $f_1 = 0.3$ and $t_2 = 40$, $f_2 = 0.1$ where the two remaining oscillating components are the cross-terms.

The real-valued part of the RD is usually referred to as the **Levin distribution (LD)**, [46], was originally derived in a quantum-mechanical context and there called Margenau-Hill distribution, [40]. It is simply defined as

$$W_z^{LD}(t, f) = \Re[z(t)Z^*(f)e^{-i2\pi ft}], \quad (5.27)$$

and has the advantage of being real-valued, similar to the Wigner distribution. The Rihaczek and Levin distributions have their obvious drawbacks, such as e.g. the double

number of cross-terms, compared to the Wigner distribution. However, these distributions are intuitively nice from interpretation aspects, and are often applied using a short-time window, e.g. the windowed RD defined as

$$W_z^{wRD}(t, f) = z(t) [\mathcal{F}_{\tau \rightarrow f}\{z(\tau)w(\tau - t)\}]^* e^{-i2\pi ft}. \quad (5.28)$$

With an appropriate size of the window, the cross-terms of the RD can be reduced for signals where the components appear at different time instants.

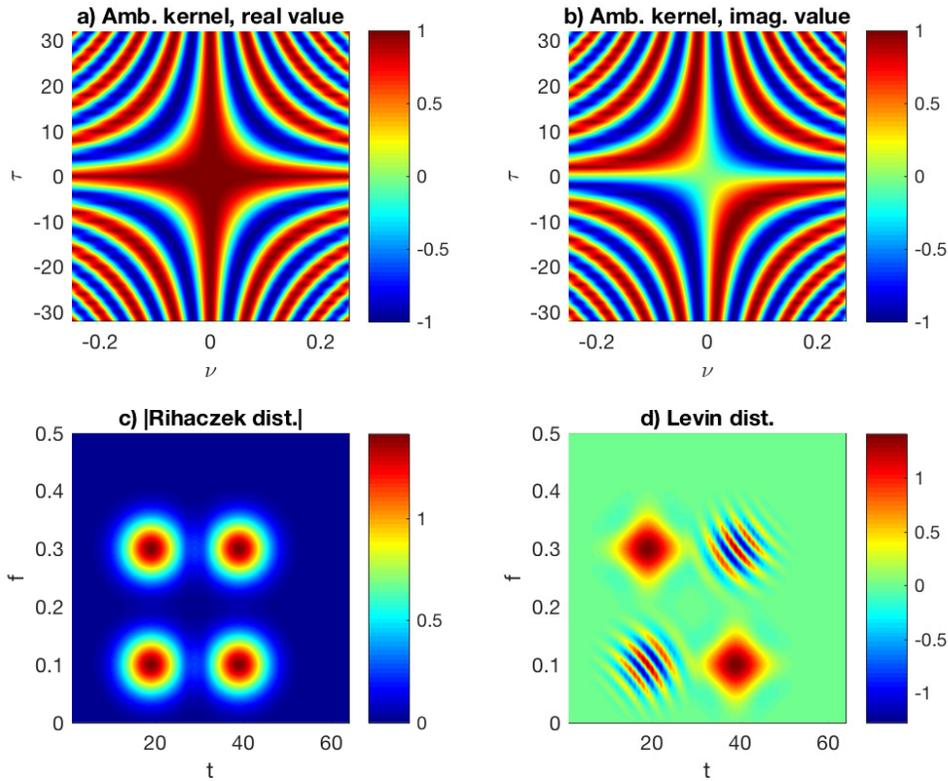


Figure 5.4: The ambiguity kernel of the Rihaczek distribution; a) real-valued part; b) imaginary part; c) The absolute value of the Rihaczek distribution for a two-component complex-valued Gaussian windowed signal with $t_1 = 20$, $f_1 = 0.3$ and $t_2 = 40$, $f_2 = 0.1$; d) The real-valued part of the Rihaczek distribution, also referred to as the Levin distribution.

5.6 Kernel interpretation of the spectrogram

The quadratic class as defined in Eq. (5.4) can also be reformulated as

$$W_z^Q(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(u + \frac{\tau}{2}) z^*(u - \frac{\tau}{2}) \rho(t - u, \tau) e^{-i2\pi f \tau} du d\tau, \quad (5.29)$$

where a **time-lag kernel** is defined as the inverse Fourier transform of the ambiguity kernel,

$$\rho(t, \tau) = \int_{-\infty}^{\infty} \phi(\nu, \tau) e^{i2\pi \nu t} d\nu. \quad (5.30)$$

It is then easily shown that the spectrogram also belongs to the quadratic class. The spectrogram, defined as in Eq. (2.15),

$$\begin{aligned} S_z(t, f) &= \left| \int_{-\infty}^{\infty} z(t_1) h^*(t_1 - t) e^{-i2\pi f t_1} dt_1 \right|^2, \\ &= \left(\int_{-\infty}^{\infty} z(t_1) h^*(t_1 - t) e^{-i2\pi f t_1} dt_1 \right) \left(\int_{-\infty}^{\infty} z^*(t_2) h(t_2 - t) e^{i2\pi f t_2} dt_2 \right), \end{aligned}$$

is reformulated using $t_1 = u + \frac{\tau}{2}$ and $t_2 = u - \frac{\tau}{2}$ to become

$$S_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(u + \frac{\tau}{2}) z^*(u - \frac{\tau}{2}) h^*(u + \frac{\tau}{2} - t) h(u - \frac{\tau}{2} - t) e^{-i2\pi f \tau} du d\tau, \quad (5.31)$$

where we compare with Eq. (5.29) and identify the time-lag kernel of the spectrogram as

$$\rho^h(t, \tau) = h^*(-t + \frac{\tau}{2}) h(-t - \frac{\tau}{2}). \quad (5.32)$$

The corresponding ambiguity kernel $\phi^h(\nu, \tau)$ is found as the Fourier transform of the time-lag kernel,

$$\phi^h(\nu, \tau) = \int_{-\infty}^{\infty} \rho^h(t, \tau) e^{-i2\pi \nu t} dt = \int_{-\infty}^{\infty} h^*(-t + \frac{\tau}{2}) h(-t - \frac{\tau}{2}) e^{-i2\pi \nu t} dt. \quad (5.33)$$

As expected, as the spectrogram is real-valued, the Hermitian property is fulfilled. The ambiguity kernel marginal properties are however difficult to reach using any reasonable window, which is seen e.g., from the case where $\tau = 0$, giving $\phi(\nu, 0) = \int |h(-t)|^2 e^{-i2\pi \nu t} dt$, which can only be equal to one if $h(t) = \delta(t)$, which obviously is not a useful window.

5.7 Multitaper time-frequency analysis

It has also been shown that the calculation of the two-dimensional convolution between the kernel and the Wigner distribution can be simplified using kernel decomposition and calculating a **multitaper** spectrogram, [27, 47]. This can be seen if we use the quadratic class definition based on the time-lag kernel, from Eq. (5.29), where we replace $u = (t_1 + t_2)/2$ and $\tau = t_1 - t_2$, resulting in

$$W_z^Q(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t_1)z^*(t_2)\rho\left(t - \frac{t_1 + t_2}{2}, t_1 - t_2\right)e^{-i2\pi f(t_1 - t_2)} dt_1 dt_2, \quad (5.34)$$

where the time-lag kernel is redefined into a rotated time-lag kernel,

$$\rho^{rot}(t_1, t_2) = \rho\left(\frac{t_1 + t_2}{2}, t_1 - t_2\right). \quad (5.35)$$

Solving

$$\int_{-\infty}^{\infty} \rho^{rot}(t_1, t_2)u(t_1)dt_1 = \lambda u(t_2), \quad (5.36)$$

results in eigenvalues λ_k and corresponding eigenfunctions $u_k(t)$. The kernel can be expressed as

$$\rho^{rot}(t_1, t_2) = \sum_{k=1}^{\infty} \lambda_k u_k^*(t_1)u_k(t_2), \quad (5.37)$$

if $\rho^{rot}(t_1, t_2) = (\rho^{rot}(t_2, t_1))^*$, and Eq. (5.34) is rewritten as a weighted sum of spectrograms according to

$$\begin{aligned} W_z^Q(t, f) &= \sum_{k=1}^{\infty} \lambda_k \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t_1)z^*(t_2)u_k^*(t_1 - t)u_k(t_2 - t)e^{-i2\pi ft_1}e^{i2\pi ft_2} dt_1 dt_2 \\ &= \sum_{k=1}^{\infty} \lambda_k \left| \int_{-\infty}^{\infty} z(t_1)u_k^*(t_1 - t)e^{-i2\pi ft_1} dt_1 \right|^2. \end{aligned} \quad (5.38)$$

Depending on the different λ_k , the number of spectrograms to be averaged are in practice just a few or an infinite number. With a small number of λ_k that differs significantly from zero, the multitaper spectrogram, defined as

$$S_z(t, f) = \sum_{k=1}^K \lambda_k \left| \int_{-\infty}^{\infty} z(t_1) u_k^*(t_1 - t) e^{-i2\pi f t_1} dt_1 \right|^2, \quad (5.39)$$

can be an efficient solution from implementation and optimization aspects. The number of windows K is usually in the range $2 - 10$. The best concentration in a time-interval and simultaneously in a specific frequency interval is given by the orthogonal prolate spheroidal wave functions, presented and investigated in a famous series of papers, [48, 49, 50, 51, 52]. In the final paper the discrete prolate spheroidal sequences (DPSS) are defined, nowadays often referred to as the **Slepian functions** or the Thomson multitapers, [53]. The Slepian functions are recognized to give orthonormal spectra for stationary white noise and also to be the most localized tapers in the frequency domain. They are the time-limited and band-limited eigenvectors of a rectangular region in the time-frequency domain.

An operator defining a circular region in the time-frequency domain, which is radially symmetric will have the **Hermite functions** as eigenfunctions, [13, 54]. In time-frequency analysis the Hermite functions have been shown to give the best time-frequency localization as well as orthonormality for white noise. Recently, many methods have been proposed based on the Hermite functions as multitapers, [55, 56, 57, 58, 59]. However, the Hermite functions and the Slepian functions are not orthonormal for spectra with sharp peaks, where the cross-correlation between sub-spectra give degraded performance, [60]. Approaches to approximate a kernel with a few multitaper spectrograms have been taken, e.g., decomposing the Choi-Williams kernel, [61, 62], least square optimization of the weights, [55] and an approach using a penalty function, [63]. In [56], time-frequency kernels and connected multitapers for statistically stable frequency marginals are developed.

Chapter 6

Optimal resolution of time-frequency spectra

Many time-frequency methods are devoted to suppression of cross-terms and to maintain the concentration of the auto-terms from the Wigner distribution. To measure auto-term concentration and possibly automatically determine parameters for a method, we need measurement criteria for the concentration. In this chapter, some often applied measures are discussed.

A well-known and large sub-area in time-frequency analysis is estimation of **instantaneous frequency** or instantaneous phase. The instantaneous frequency is the derivative of the phase, and is closely related to the definition of the analytic signal. For the analytic signal or any mono-component complex-valued signal, numerous of methods exist for extraction of the instantaneous frequency. The most modern of these are based on the **reassignment principle** and **synchrosqueezing**. The reassignment idea is to automatically reassign mass to the center of gravity and for linear chirps and constant sinusoids the true instantaneous frequency is given as a result. Recently, novel reassignment techniques are proposed estimation and detection also of short transient signals.

6.1 Concentration measures

The basic idea of a concentration measure can be found using this simple example from probability theory. We use N non-negative numbers, p_1, p_2, \dots, p_N where

$$p_1 + p_2 + \dots + p_N = 1. \tag{6.1}$$

A quadratic test function

$$\gamma = p_1^2 + p_2^2 + \dots + p_N^2, \quad (6.2)$$

will attend its minimum value when $p_1 = p_2 = \dots = p_N = 1/N$ (maximum spread) and the maximum value when only one $p_i = 1$ and all the others are zero (minimum spread). Incorporating a unity sum constraint in the test function gives

$$\gamma = \frac{p_1^2 + p_2^2 + \dots + p_N^2}{(p_1 + p_2 + \dots + p_N)^2}. \quad (6.3)$$

We can use this idea as a measure for the time-frequency concentration, e.g., as suggested in [64],

$$\gamma = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^4(t, f) dt df}{\left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^2(t, f) dt df\right)^2}, \quad (6.4)$$

where $W(t, f)$ is a chosen time-frequency distribution or spectrum. Other norms could of course be used, but this one is similar to the often used **kurtosis** in statistics. The measure in Eq. (6.4) is however not useful for signals including several components and a better measure is then to use a local function, also proposed in [64],

$$\gamma_{(t,f)} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q^2(t_1 - t, f_1 - f) W^4(t_1, f_1) dt_1 df_1}{\left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(t_1 - t, f_1 - f) W^2(t_1, f_1) dt_1 df_1\right)^2}, \quad (6.5)$$

where $Q(t, f)$ is a weighting function, typically the Gaussian function, that determines the region for the concentration measure.

Another more recent measure is based on the **Rényi entropy**, [65],

$$R_\alpha = \frac{1}{1 - \alpha} \log_2 \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^\alpha(t, f) dt df \right), \quad \alpha > 2. \quad (6.6)$$

where $1/(1 - \alpha)$ becomes negative and the resulting entropy becomes larger for less concentrated distributions and smaller for more highly concentrated distributions. Eq. (6.6) adopts the lowest value for the Wigner distribution of a single Gaussian component. A normalized Rényi entropy defined as

$$R_\alpha = \frac{1}{1 - \alpha} \log_2 \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{W(t, f)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t, f) dt df} \right)^\alpha dt df \right), \quad (6.7)$$

is suggested get similar entropy values independently of the power of the signal.

For odd values of α , R_α is asymptotically invariant to the time-frequency cross-terms, which actually means that the use of the Rényi entropy when optimizing auto-term concentration and cross-term suppression should be applied with care. But if only the concentration optimization is the issue, the value $\alpha = 3$ is often used, [66, 67, 68]. The Rényi entropy could also be applied for counting of the number of components in a multi-component signal, [69].

Using different normalizations, e.g. the volume normalized Rényi entropy, [70],

$$R_\alpha = \frac{1}{1 - \alpha} \log_2 \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{W(t, f)}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |W(t, f)| dt df} \right)^\alpha dt df \right), \quad (6.8)$$

the optimization of the suppression of cross-terms could be included in the measure, [71]. Other approaches for measuring the reduction of cross-terms is found in [72].

The Rényi entropy relates to the well-known Shannon entropy

$$H = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t, f) \log_2 W(t, f) dt df, \quad (6.9)$$

which could be recovered from Eq. (6.6) when $\alpha \rightarrow 1$, [65]. We should note that the Shannon entropy could not be used for a distribution including negative values, such as the Wigner distribution and many other distributions in the quadratic class, as the logarithm is applied inside the integral. In the definition of the Rényi entropy in Eq. (6.6) there is also an logarithm, but this is found outside the double integral, and most signals and smoothed distributions are positive in this context. However, examples of multi-component signals and α can be found where $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^\alpha(t, f) dt df < 0$, see [65].

6.2 Instantaneous frequency

A general definition of a single-component oscillating signal is

$$x(t) = A(t)e^{i\phi(t)}, \quad (6.10)$$

where $A(t)$ is defined as the **instantaneous amplitude** and $\phi(t)$ is the **instantaneous phase**. For a simple amplitude-modulated sinusoidal signal,

$$x(t) = A(t)e^{i2\pi f_0 t}, \quad (6.11)$$

the frequency for all t is easily found from the derivative of the phase as

$$f_0 = \frac{\phi'(t)}{2\pi}. \quad (6.12)$$

The more general case of a time-varying sinusoidal signal where $\phi(t) = 2\pi f(t)t$, we rely on that $\phi(t)$ is differentiable and can be evaluated at $t = t_1$ and $t = t_2$, where $t_2 > t_1$. Then there exists a time interval t_1 and t_2 such that

$$\phi(t_2) - \phi(t_1) = (t_2 - t_1)\phi'(t). \quad (6.13)$$

If p_I is the period of one oscillation of $x(t)$, then the **instantaneous frequency** is defined as $f_I = 1/p_I$. Evaluating for one period $\phi'(t)$ from Eq. (6.13) will be simplified to

$$\phi'(t) = \frac{\phi(t_2) - \phi(t_1)}{t_2 - t_1} = \frac{2\pi}{p_I} = 2\pi f_I, \quad (6.14)$$

as the period time relates $t_2 = t_1 + p_I$ and the resulting oscillation is $\phi(t_2) = \phi(t_1) + 2\pi$. The instantaneous frequency will, similarly to the constant-frequency example, then be

$$f_I = \frac{\phi'(t)}{2\pi}, \quad (6.15)$$

for the time $t_1 < t < t_2$.

An estimate is found as

$$\hat{f}_I(t) = \frac{\int_{-\infty}^{\infty} f \hat{S}(t, f) df}{\int_{-\infty}^{\infty} \hat{S}(t, f) df}, \quad (6.16)$$

where $S(t, f)$ could be the spectrogram or any other time-frequency distribution. The estimate is however sensitive to noise and a more straight forward and robust estimate for specifically the mono-component case is

$$\hat{f}_I(t) = \arg \max_f S(t, f). \quad (6.17)$$

Algorithms that uses demodulation of the estimated frequency to iteratively find the instantaneous phase exist, [73, 74], and post processing analysis is also applied nowadays, such as image analysis techniques, [75].

The IF of multi-component signals

One should however note that for multi-component signals, the definition of instantaneous frequency becomes difficult and usually these methods do not work properly, especially when the components are closely spaced or cross in the time-frequency domain. In the case of a real-valued signal, e.g.,

$$x(t) = A(t) \cos(2\pi f(t)t), \quad (6.18)$$

the instantaneous frequency will be zero as this signal has no phase in the sense defined above. The concept of instantaneous frequency is no longer valid and the methods that rely on the definition do not work properly. One example is the sum of two complex sinusoids,

$$x(t) = A_1 e^{i2\pi f_1 t} + A_2 e^{i2\pi f_2 t}, \quad (6.19)$$

where f_1 and f_2 are positive, i.e., an analytic signal. Depending on the values of the frequencies and the amplitudes very different instantaneous frequencies show up, see Figure 6.1a), where the parameters are $f_1 = 0.11$, $f_2 = 0.2$, $A_1 = 0.5$ and $A_2 = 1$ giving an instantaneous frequency varying between 0.17 and 0.29 which includes the frequency 0.2 but not at all 0.11. In Figure 6.1b), the amplitude of the first frequency is changed to $A_1 = -1.5$ and the resulting instantaneous frequency even becomes negative for certain time points.

6.3 The reassignment technique

The concept of reassignment was introduced in 1976, [76], but were not applied to a larger extent until it was reintroduced in [77]. For multi-component signals, the reassignment improves the readability as the cross-terms are reduced by a smoothing and the reassignment then squeezes the signal terms to be more localized. The method works well for long linear chirps and constant frequency signals where theoretically the instantaneous frequency is achieved.

The STFT of the signal $x(t)$ using the window $h(t)$ at time t is

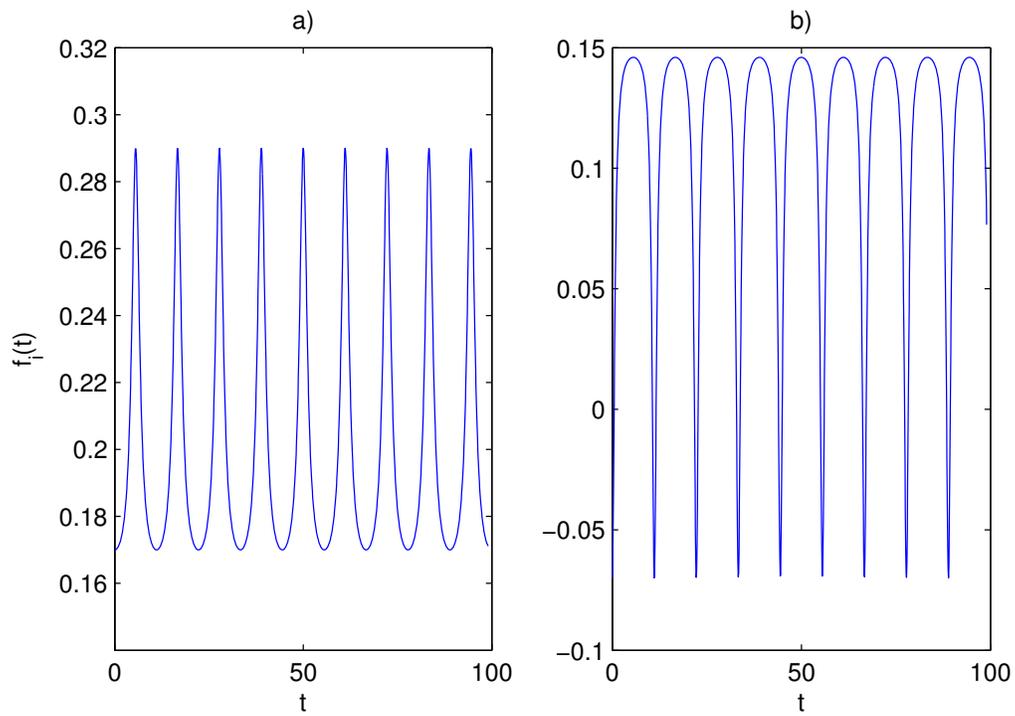


Figure 6.1: Example of instantaneous frequency of a two-component signal; a) $f_1 = 0.11$, $f_2 = 0.2$, $A_1 = 0.5$ and $A_2 = 1$; b) $f_1 = 0.11$, $f_2 = 0.2$, $A_1 = -1.5$ and $A_2 = 1$.

$$X^h(t, \omega) = \int_{-\infty}^{\infty} x(t_1)h^*(t_1 - t)e^{-i\omega t_1} dt_1, \quad (6.20)$$

where $\omega = 2\pi f$ and the corresponding spectrogram is found as

$$S_x^h(t, \omega) = |X^h(t, \omega)|^2. \quad (6.21)$$

The **reassigned spectrogram** is defined as

$$RS_x^h(t, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_x^h(t_1, \xi) \delta(t - \hat{t}_x(t_1, \omega_1), \omega - \hat{\omega}_x(t_1, \omega_1)) dt_1 d\omega_1, \quad (6.22)$$

where each value of the spectrogram is relocated to a corresponding \hat{t}_x and $\hat{\omega}_x$, where $\delta(t, \omega)$ is the two-dimensional Dirac impulse defined as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, \omega) \delta(t - t_0, \omega - \omega_0) dt d\omega = f(t_0, \omega_0). \quad (6.23)$$

The crucial part is the actual reassignment, which can be computed as

$$\hat{t}_x(t, \omega) = t + \Re\left(\frac{X^{th}(t, \omega)}{X^h(t, \omega)}\right), \quad (6.24)$$

$$\hat{\omega}_x(t, \omega) = \omega - \Im\left(\frac{X^{dh/dt}(t, \omega)}{X^h(t, \omega)}\right), \quad (6.25)$$

where $X^{th}(t, \omega)$ and $X^{dh/dt}(t, \omega)$ are STFTs of the signal $x(t)$ using the time multiplied window $t \cdot h(t)$ and the time derivative window $dh(t)/dt$ respectively, [77]. The spectrogram can be replaced by any other distribution belonging to the quadratic class, and the resulting reassignment process will lead to perfectly localized single component chirp signals, frequency tones and impulse signals, [77]. An example is given in Figure 6.2, with two linear Gaussian windowed chirp signals, one of increasing frequency and the other of decreasing frequency. The usual windowed spectrogram, of window length 32, gives the view in Figure 6.2a, and in Figure 6.2b the reassigned spectrogram is presented with a view close the instantaneous frequencies, except at the crossing of the two signals.

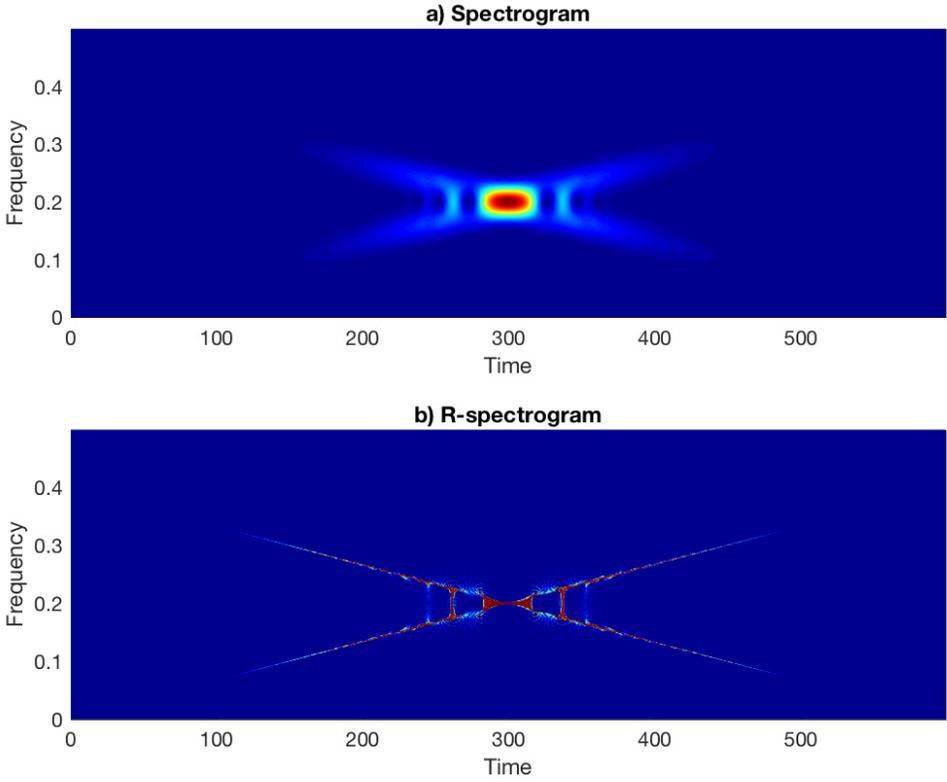


Figure 6.2: Two crossing Gaussian windowed linear chirp-signal; a) spectrogram; b) reassigned spectrogram.

6.4 Scaled reassigned spectrogram

For short transients, the reassignment does not give an optimal result but with use of a simple scaling of the relocation, optimal concentration of especially a Gaussian transient can be achieved, [78]. The Gaussian transient is defined as

$$x_0(t) = g(t - t_0)e^{i\omega_0 t}, \quad -\infty < t < \infty, \quad (6.26)$$

where

$$g(t) = e^{-\frac{t^2}{2\sigma^2}}, \quad -\infty < t < \infty. \quad (6.27)$$

The quadratic class of distributions obey time-frequency shift-invariance, implying that further analysis can be restricted to $x(t) = g(t)$, [27]. Using a unit energy Gaussian window with parameter λ , $h(t) = 1/(\pi^{1/4}\sqrt{\lambda})e^{-\frac{t^2}{2\lambda^2}}$, and introducing the scaling factors c_t and c_ω in Eqs. (6.24, 6.25), will lead to the following reassignment, [78],

$$\hat{t}_x(t, \omega) = t - c_t \left(\frac{\lambda^2}{\lambda^2 + \sigma^2} t \right), \quad (6.28)$$

$$\hat{\omega}_x(t, \omega) = \omega - c_\omega \left(\frac{\sigma^2}{\lambda^2 + \sigma^2} \omega \right). \quad (6.29)$$

For the matched window case, i.e., $\lambda = \sigma$, Eqs. (6.28, 6.29) simplify to

$$\hat{t}_x(t, \omega) = t - c_t \frac{t}{2}, \quad (6.30)$$

$$\hat{\omega}_x(t, \omega) = \omega - c_\omega \frac{\omega}{2}. \quad (6.31)$$

For perfect localization of the reassigned spectrogram, where the Gaussian function is centered at time and frequency zero, \hat{t} and $\hat{\omega}$ should accordingly be zero for all values of t and ω . This is achieved for $c_t = c_\omega = 2$ in Eqs. (6.30, 6.31), with the perfect localization given as

$$SRS_x^h(t, \omega) = \delta(t, \omega).$$

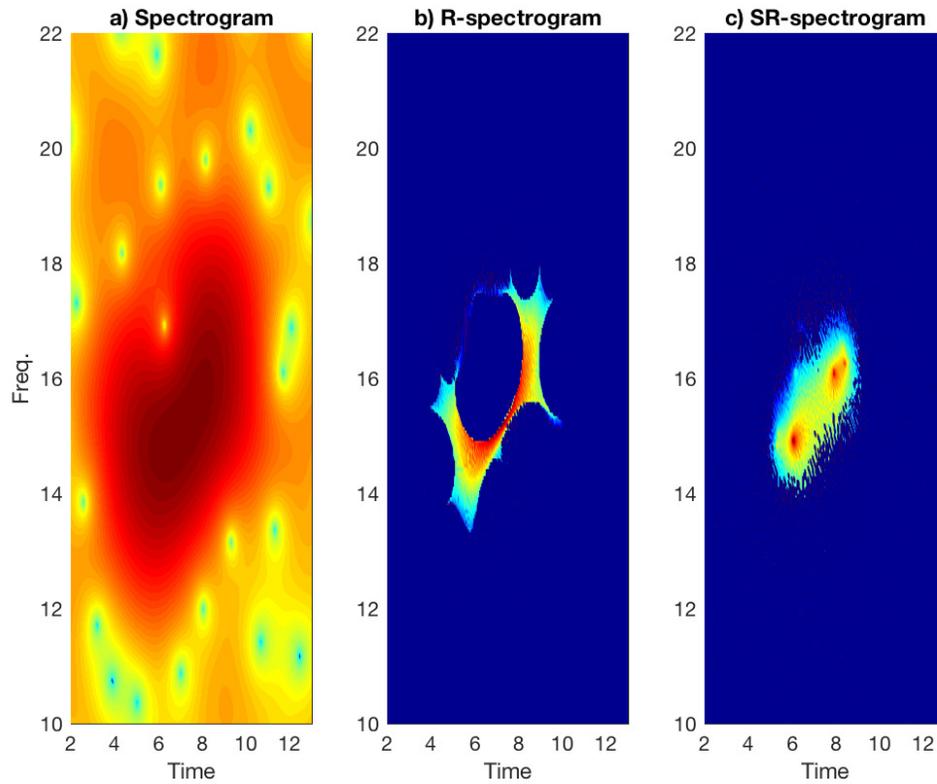


Figure 6.3: Time-frequency representations in dB-scale of different methods for the sum of two Gaussian functions centered at $t_1 = 6$, $\omega_1 = 15$ and $t_2 = 9$, $\omega_2 = 17$; a) spectrogram, b) reassigned spectrogram, c) scaled reassigned spectrogram for $\lambda = \sigma = 20$.

For the usual reassignment without scaling, $c_t = c_\omega = 1$,

$$\begin{aligned}\hat{t}_x(t, \omega) &= \frac{\sigma^2}{\lambda^2 + \sigma^2}t = \frac{t}{2} \\ \hat{\omega}_x(t, \omega) &= \frac{\lambda^2}{\lambda^2 + \sigma^2}\omega = \frac{\omega}{2}.\end{aligned}$$

with the last steps given for $\lambda = \sigma$. The corresponding reassigned spectrogram is found as, [79],

$$\begin{aligned}RS_x^h(t, \omega) &= \frac{2\sqrt{\pi}(\lambda^2 + \sigma^2)}{\lambda} e^{-\left(\left(\frac{\lambda^2 + \sigma^2}{\sigma^4}\right)t^2 + \left(\frac{(\lambda^2 + \sigma^2)\sigma^2}{\lambda^2}\right)\omega^2\right)} \\ &= 4\sqrt{\pi}\sigma e^{-2\left(\frac{t^2}{\sigma^2} + \sigma^2\omega^2\right)},\end{aligned}\tag{6.32}$$

with the last step given for $\lambda = \sigma$.

The performance is illustrated with a two-component signal of Gaussian transients defined as in Eq. (6.26) with $\sigma = 20$, $t_1 = 6$, $\omega_1 = 15$ and $t_2 = 9$, $\omega_2 = 17$. In Figure 6.3, the spectrogram, the usual reassignment and the scaled reassignment are presented for the matched window case $\lambda = 20$. The spectrogram of Figure 6.3a indicates that there are two components but the reassignment presented in Figure 6.3b shows a representation that could be interpreted as some other signal. For the scaled reassignment, the final representation of Figure 6.3c clearly shows the two components localized at their correct time-frequency positions.

We finally show an example in Figure 6.4, of a signal that seems to be a Gaussian windowed chirp-signal of increasing frequency but that actually is a sum of four equal length Gaussian functions and added white noise disturbance according to Figure 6.4a. The corresponding spectrogram is presented in Figure 6.4b, the result of the reassigned spectrogram is shown in Figure 6.4c and the scaled reassignment in Figure 6.4d. For the scaled reassignment, the masses are reallocated to the centers showing the appearance of four Gaussian functions, where for the usual reassignment the underlying signal appears to be a chirp signal.

We should note that in most cases the parameter σ is unknown and needs to be estimated for the matched window calculation and corresponding scaled reassignment. However, the scaled reassignment can be performed using either a set of candidate matched windows or candidate scaling factors. With use of the Rényi entropy for

measurement of the resulting concentration, an optimal reassignment can be found, [80, 81, 78]. The scaled reassignment has been applied to detection of dolphin echolocation signals, [82, 83].

6.5 Other modern techniques for optimal resolution

There is also an interest in discriminating between short transients and longer signals where the Levenberg-Marquardt (LM) reassignment, proposed in [84], is an adjustable method where the user can choose if a weak or strong localization should be made. For the linear chirp, the impulse and the constant frequency signal, the perfect localization is achieved already with the usual reassignment and the LM technique can only be used to make this concentration weaker.

Empirical mode decomposition (EMD) was introduced by Huang in [85] and decomposes multi-component non-stationary signals into intrinsic mode functions (IMF). To obtain each IMF, an iterative procedure called 'sifting' is used. The main idea is to recursively estimate and subtract the local mean until the resulting function becomes an IMF. A review of applications are found in [86]. The drawback of the EMD is that the resulting estimated IMFs are sensitive to disturbing noise. This phenomenon is known as the mode mixing problem where a solution is to use filterbank structures of time-varying filters, [87].

The synchrosqueezing transform (SST) is developed as an alternative to the EMD, [88] and reassigns energy in frequency at a certain time point. The SST operates on the STFT in contrary to reassignment that uses the spectrogram. Recently, many papers suggesting improved SST algorithms have been published. E.g., in [89] a nonlinear SST for weak signal detection is invented and the authors of [90] suggest a synchrosqueezing wavelet transform.

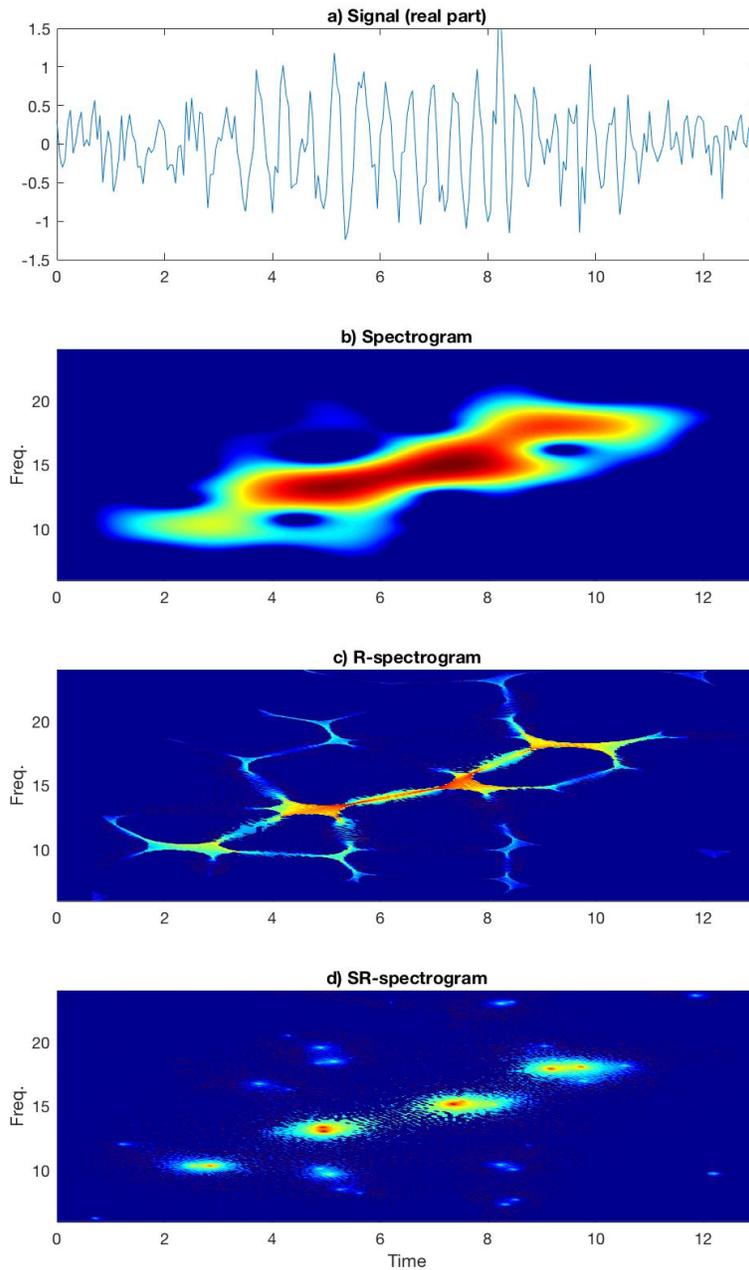


Figure 6.4: Example of time-frequency representations in dB-scale for four Gaussian functions ($\sigma_n = 20$) with increasing center frequencies disturbed by white noise, (SNR=12dB); a) the signal, b) spectrogram, c) reassigned spectrogram, d) scaled re-assigned spectrogram.

Chapter 7

Stochastic time-frequency analysis

The fundamental problem in **stochastic** time-frequency analysis is the estimation of a reliable time-frequency spectrum from a single realization of the process. For weakly stationary stochastic processes, the covariance function is assumed to be dependent only on the time difference, and with use of all available pairs of data of this time difference, a possibly reliable estimate of the covariance is given and accordingly an estimate of the spectral density. If no assumptions are made of the covariance or spectral properties, the resulting estimates are certainly to become unreliable. Therefore, different restrictions are made of the non-stationarity. Priestley, [91], suggested the oscillatory random process, a sinusoidal function of stochastic time-varying amplitude, giving the **evolutionary spectrum**. The **locally stationary process** by [92], restricts the covariance function to be a separable time-varying stationary covariance. However, many of the general quadratic class of time-frequency representation methods can be used for successful estimation, although they are actually developed for deterministic signals. For the class of so called **underspread processes**, the ambiguity spectrum is well concentrated around the origin, and it has been shown that many rules for time-invariant linear systems and stationary processes can be applied to these processes, [93].

7.1 Definitions of non-stationary processes

The spectral decomposition theorem or the Cramér harmonic decomposition, [94], for a zero-mean complex-valued weakly stationary process, $z(t)$, admits the Fourier-Stieltjes integral

$$z(t) = \int_{-\infty}^{\infty} e^{i2\pi ft} dZ(f), \quad (7.1)$$

where $dZ(f)$ is a complex-valued uncorrelated increment process. As the stationary process admits a frequency decomposition into uncorrelated random variables, we find $E[dZ(f_1)dZ^*(f_2)] = 0$, when $f_1 \neq f_2$, and $E[|dZ(f)|^2] = dF(f)$, where $dF(f)$ is related to the spectral density as $dF(f) = S_z(f)df$, [91]. We note that the decomposition in Eq. (7.1) has double orthogonality, i.e. with respect to the usual inner product,

$$\int_{-\infty}^{\infty} e^{i2\pi f_1 t} e^{-i2\pi f_2 t} dt = \delta(f_2 - f_1). \quad (7.2)$$

as well as with respect to the statistical inner product defined by the expectation

$$E[dZ(f_1)dZ^*(f_2)] = \delta(f_1 - f_2)S_z(f_1)df_2df_1. \quad (7.3)$$

The covariance function is found as

$$\begin{aligned} r_z(\tau) = E[z(t + \tau)z^*(t)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i2\pi f_1(t+\tau)} e^{-i2\pi f_2 t} E[dZ(f_1)dZ^*(f_2)] \\ &= \int_{-\infty}^{\infty} e^{i2\pi f \tau} dF(f), \\ &= \int_{-\infty}^{\infty} e^{i2\pi f \tau} S_z(f)df, \end{aligned} \quad (7.4)$$

where the above equations rely on the double orthogonality.

We continue to focus on the second-order properties and define a zero-mean complex-valued non-stationary stochastic process, $z(t)$, characterized by the covariance function,

$$\begin{aligned} r_z(t_1, t_2) = E[z(t_1)z^*(t_2)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i2\pi(f_1 t_1 - f_2 t_2)} E[dZ(f_1)dZ^*(f_2)], \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{i2\pi(f_1 t_1 - f_2 t_2)} S_z(f_1, f_2)df_1df_2, \end{aligned} \quad (7.5)$$

where $r_z(t_1, t_2)$ is related of the spectral distribution function $S_z(f_1, f_2)$, [95]. The corresponding process $z(t)$ is called **harmonizable**, [96, 97]. We note that the spectral increments $dZ(f)$ of $S_z(f_1, f_2)$ are not necessarily orthogonal as in the case for a stationary process. However, we could note that Eq. (7.5) simplifies into Eq. (7.4) for the double orthogonality of the stationary process.

If we want to maintain the orthogonality of the decomposition for the non-stationary process, the **Karhunen decomposition** can be applied, where the eigenfunctions $\psi(t, f)$ of the covariance kernel is used, [95]. The covariance has the form

$$\begin{aligned} r_z(t_1, t_2) = E[z(t_1)z^*(t_2)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(t_1, f_1)\psi^*(t_2, f_2)E[dZ(f_1)dZ^*(f_2)] \\ &= \int_{-\infty}^{\infty} \psi(t_1, f_2)\psi^*(t_2, f_2)S_z(f_2)df_2. \end{aligned} \quad (7.6)$$

The eigenfunctions are given from solving the eigenvalue problem given from the multiplication of $\psi(t_2, f_1)$ on both sides and integration over t_2 of Eq. (7.6),

$$\begin{aligned} \int_{-\infty}^{\infty} r_z(t_1, t_2)\psi(t_2, f_1)dt_2 &= \int_{-\infty}^{\infty} \psi(t_1, f_2) \left[\int_{-\infty}^{\infty} \psi(t_2, f_1)\psi^*(t_2, f_2)dt_2 \right] S_z(f_2)df_2 \\ &= \int_{-\infty}^{\infty} \psi(t_1, f_2)\delta(f_2 - f_1)S_z(f_2)df_2 \\ &= S_z(f_1)\psi(t_1, f_1), \end{aligned} \quad (7.7)$$

where the orthogonality relation

$$\int_{-\infty}^{\infty} \psi(t_2, f_1)\psi^*(t_2, f_2)dt_2 = \delta(f_2 - f_1), \quad (7.8)$$

is used. The decomposition of the variance leads to

$$r_z(t, t) = E[|z(t)|^2] = \int_{-\infty}^{\infty} |\psi(t, f_2)|^2 S_z(f_2)df_2, \quad (7.9)$$

which defines a time-dependent power spectrum as

$$S_z(t, f) = |\psi(t, f)|^2 S_z(f), \quad (7.10)$$

and the process $z(t)$ can be expanded as

$$z(t) = \int_{-\infty}^{\infty} \psi(t, f)dZ(f). \quad (7.11)$$

The advantage of the Karhunen representation is that the orthogonality is kept in time as well as frequency. However, the interpretation of the variable f as frequency is no longer obvious.

The evolutionary spectrum

The **oscillatory random process** [91], has basis function defined by

$$\psi(t, f) = A(t, f)e^{i2\pi ft}, \quad (7.12)$$

giving the **evolutionary spectrum**,

$$S_z(t, f) = |A(t, f)|^2 S_z(f). \quad (7.13)$$

These basis functions are not necessarily orthogonal and the oscillatory processes are also a rather restricted class, as a linear combination of oscillatory processes is not necessarily oscillatory. However, the spectral decomposition of a stationary process, Eq. (7.1), where the process is decomposed using the orthogonal basis functions $e^{i2\pi ft}$ is similar in the sense that an amplitude modulation of each complex exponential is applied for the non-stationary oscillatory process. As $A(t, f)$ is varying in time for each frequency, the decomposition is close to orthogonal when

$$\int_{-\infty}^{\infty} \psi(t, f_1)\psi^*(t, f_2)dt = \int_{-\infty}^{\infty} A(t, f_1)A^*(t, f_2)e^{i2\pi(f_1-f_2)t}dt \approx A^2\delta(f_1 - f_2), \quad (7.14)$$

i.e. $A(t, f_1)A^*(t, f_2) \approx A^2$.

Locally stationary processes

The **locally stationary process** (LSP), [92], has a covariance of the form

$$r_z(t_1, t_2) = q\left(\frac{t_1 + t_2}{2}\right) \cdot r(t_1 - t_2), \quad (7.15)$$

where $q(t)$ is a non-negative function and $r(t)$ positive definite. The LSP falls into the class of harmonizable processes and studying a formulation as a local symmetric covariance function, i.e.,

$$E\left[z\left(t + \frac{\tau}{2}\right)z^*\left(t - \frac{\tau}{2}\right)\right] = q(t) \cdot r(\tau), \quad (7.16)$$

we see that the LSP is a result of a modulation in time of a stationary covariance function. With $q(t) = e^{-t^2/2}$ and $r(\tau) = e^{-\frac{c}{4}\tau^2/2}$, we study some realizations in Figure 7.1 for different parameter values c . Figure 7.1a presents four realizations for $c = 1.1$,

which is a band limited process where most of the stochastic behavior shows up in the different amplitudes. In Figure 7.1b the process realizations for $c = 1.1$ are modulated with a cosine harmonic function of frequency $f_0 = 1$, which shows the ability to model any oscillating signal of random amplitude. For larger c , more broadband realizations are generated as seen in the four examples for $c = 10$ in Figure 7.1c.

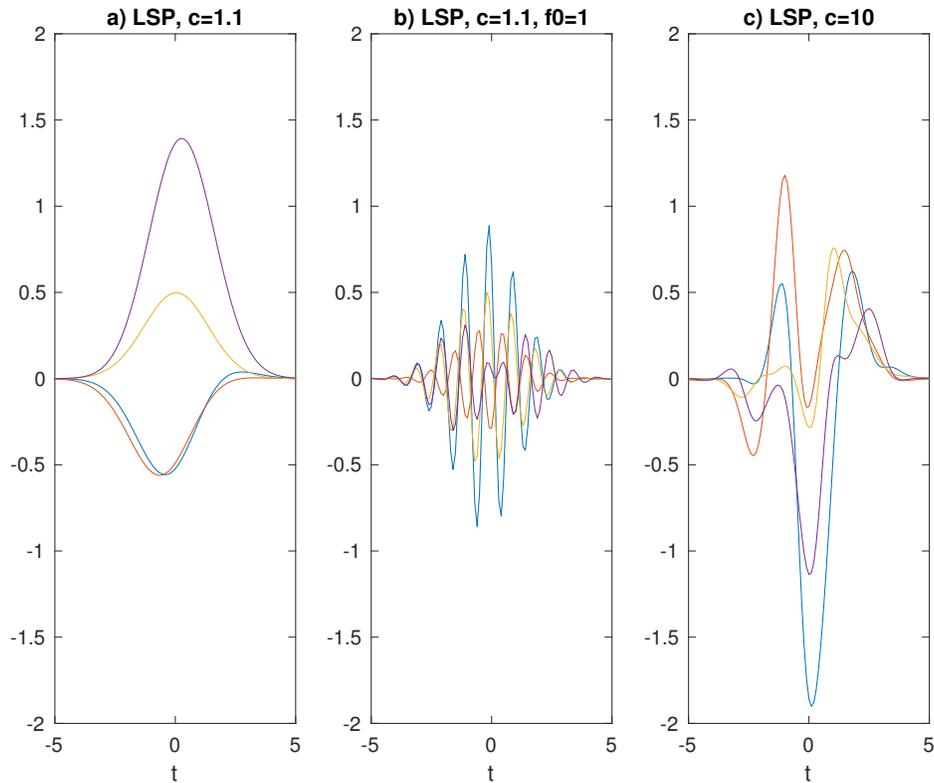


Figure 7.1: Examples of realizations (four in each case) of a locally stationary processes for different parameter values, c ; a) LSP for $c = 1.1$; LSP for $c = 1.1$ and modulation with $f_0 = 1$; c) LSP for $c = 10$.

Another similar definition is the **quasi-stationary process**, which apply a deterministic modulation directly on the stationary process, i.e.

$$z(t) = d(t)w(t), \tag{7.17}$$

where $w(t)$ is a complex-valued weakly stationary process. The local covariance function becomes

$$E[z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})] = d(t + \frac{\tau}{2})d(t - \frac{\tau}{2})r_w(\tau). \quad (7.18)$$

If the modulation is slow, i.e., the variation of $d(t)$ is small in comparison with the statistical memory of $w(t)$, then $d(t + \tau/2) \approx d(t) \approx d(t - \tau/2)$.

Another definition is connected to linear systems, where we can represent $z(t)$ as the output of a linear time-varying system \mathbf{H} whose input is stationary white noise, denoted $n(t)$, with power spectral density $S_n(f) = 1$,

$$z(t) = (\mathbf{H}n)(t) = \int_{-\infty}^{\infty} h(t, t')n(t')dt', \quad (7.19)$$

We note that if $z(t)$ is stationary, \mathbf{H} is time-invariant and the power spectral density of $z(t)$ is given by $S_z(f) = |H(f)|^2$, where $H(f) = \int_{-\infty}^{\infty} h(\tau)e^{-i2\pi f\tau}d\tau$. A well-defined framework has been proposed for the class of so called **underspread** processes, where the ambiguity spectrum is well concentrated around the origin, and it has been shown that many rules for time-invariant linear systems and stationary processes can be applied to these processes, [93]. The system \mathbf{H} is called an **innovations system** of the process $z(t)$ and the process is underspread if components that are sufficiently distant from each other in the time-frequency plane are effectively uncorrelated.

7.2 The mean square error optimal kernel

Optimal kernels for certain classes of non-stationary processes can be found, e.g. [98, 99, 100]. Sayeed and Jones [98] derived the optimal kernel in the **mean square error sense** for Gaussian harmonizable processes by minimizing the integrated expected squared error in the ambiguity domain,

$$J(\phi) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E[\hat{A}_z(\nu, \tau)\phi(\nu, \tau) - E[\hat{A}_z(\nu, \tau)]]^2 d\nu d\tau, \quad (7.20)$$

resulting in the optimal ambiguity kernel

$$\phi_{opt}(\nu, \tau) = \frac{E[\hat{A}_z(\nu, \tau)]^2}{E[\hat{A}_z^2(\nu, \tau)]}. \quad (7.21)$$

The optimal kernel for the LSP

This is a general formulation for all harmonizable processes and for the LSP, defined in Eq. (7.16), a separable and rank one covariance function, the ambiguity spectrum becomes

$$E[\hat{A}_z(\nu, \tau)] = A_z(\nu, \tau) = Q(\nu)r(\tau), \quad (7.22)$$

using $Q(\nu) = \int_{-\infty}^{\infty} q(t)e^{-i2\pi t\nu} dt$. The optimal kernel for Gaussian circularly symmetric LSP-processes was derived in [100] as,

$$\phi_{opt}(\nu, \tau) = \frac{|Q(\nu)|^2|r(\tau)|^2}{|Q(\nu)|^2|r(\tau)|^2 + (\int_{-\infty}^{\infty} |r(t)|^2 e^{-i2\pi t\nu} dt)(q * q)(\tau)}, \quad (7.23)$$

and the corresponding multitapers are shown to approximate Hermite functions. In [59], these multitapers and corresponding weights of Eq. (5.39) are studied and evaluated. Some examples of LSP-optimal windows and weights are seen in Figure 7.2. In Figure 7.2a and c the multitapers and weights corresponding to the optimal kernel when $c=1.1$ are shown. The use of the corresponding multitaper spectrogram will give the mean square error optimal estimate of processes such as those exemplified in Figure 7.1a but also for processes as exemplified in Figure 7.1b as the spectrogram is frequency-shift invariant. For the process in Figure 7.1c, the mean square error optimal multitaper spectrogram is given using the windows and weights presented in Figure 7.2b and d.

In a real data example where the electrical activity of the brain was measured, a 9 Hz flickering light was introduced at a certain time point, 4.7 s, and the light stimulation lasted 1 s. The Electroencephalogram (EEG) data was recorded and the resulting LSP multitaper spectrograms for $c = 1.1, 1.5$ and 4 are presented in Figure 7.3. The results of the Peak Matched Multiple Windows (PM MW) with number of multi tapers $K = 3$ and resolution $B = 4/64$, [101, 102], the Thomson multitapers, $K = 2, B = 4/64$, [53], and the Welch method with $K=2$ using a Hanning window and 50 % overlap are also presented for comparison. We use the windows lengths $N = 64$ for all methods. The time interval close to 4.7-5.7 s and the frequency interval around 9 Hz are studied. The introduced flickering of 9 Hz give a respond of 9 Hz, which is clearly seen in the spectrograms of the first row. We also see that a 10 Hz alpha-activity is started shortly before $t = 6$ s and ends just before $t = 7$ s.

A few contributions are found concerning the actual estimation of optimal kernels and multitapers from data, e.g., [103, 104, 105].

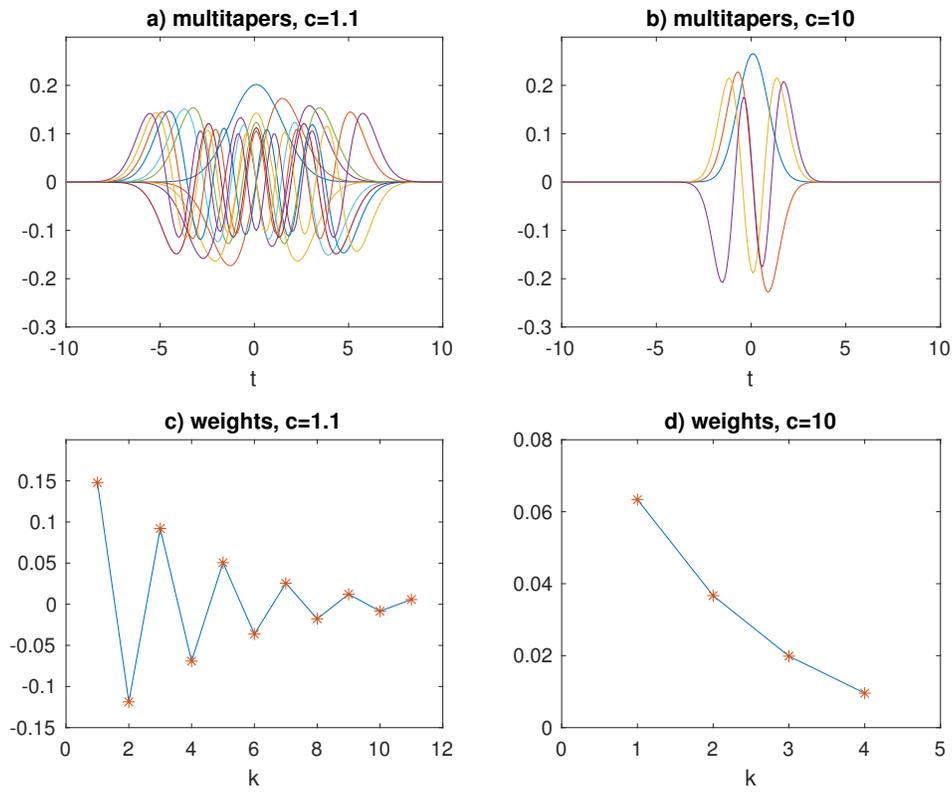


Figure 7.2: Examples of eigenvectors (multitapers) and eigenvalues (weights) of the LSP-optimal kernel; a) multitapers, $c=1.1$; b) multitapers, $c=10$; c) weights, $c=1.1$; d) weights, $c=10$.

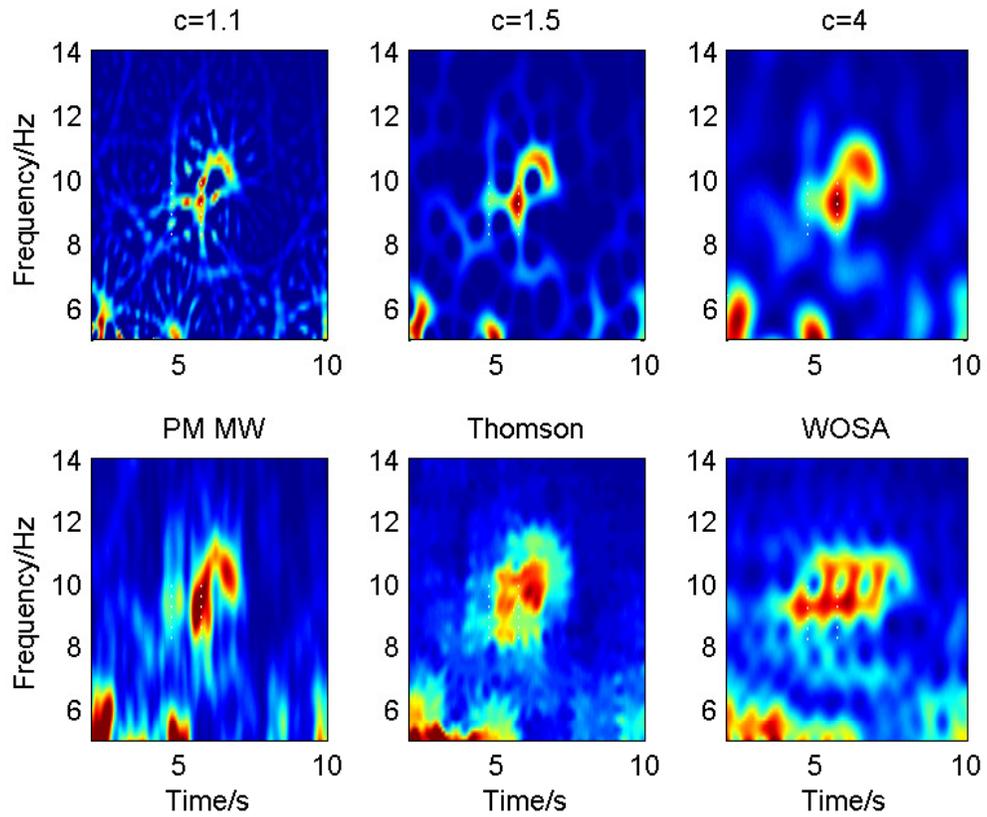


Figure 7.3: Examples of resulting multitaper spectrograms of EEG data.

Bibliography

- [1] E. A. Robinson. A historical perspective of spectrum estimation. *Proceedings of the IEEE*, 70(9):885–907, Sept 1982.
- [2] A. Schuster. On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena. *Terr. Magnet.*, 3:13–41, 1898.
- [3] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of Fourier series. *Math. Comput.*, 19:297–301, 1965.
- [4] J. W. Cooley, P. A. W. Lewis, and P. D. Welch. Historical notes on the fast Fourier transform. *IEEE Transactions Audio Electroacoust.*, AU-15:76–79, 1967.
- [5] J. W. Cooley, P. Lewis, and P. D. Welch. The fast Fourier transform and its applications. *IEEE Transactions Education*, E-12:27–34, March 1969.
- [6] D. R. Brillinger. John W. Tukey’s work on time series and spectrum analysis. *The Annals of Statistics*, 30(6):1595–1618, Dec 2002.
- [7] D. Gabor. Theory of communications. *Journal of the IEE*, 93:429–457, 1946.
- [8] G. Lindgren, H. Rootzén, and M. Sandsten. *Stationary stochastic processes for scientists and engineers*. CRC Press, 2014.
- [9] M. J. Bastiaans. Gabor’s expansion of a signal into Gaussian elementary signals. *Proceedings of the IEEE*, 68(4):538–539, April 1980.
- [10] A. J. E. M. Janssen. The Zak transform: a signal transform for sampled time-continuous signals. *Philips J. Res.*, 43:23–69, 1988.
- [11] J. Wexler and S. Raz. Discrete Gabor expansions. *Signal Processing*, 21(3):207–221, Nov 1990.

- [12] A. Graps. An introduction to wavelets. *IEEE Computational Science and Engineering*, 2(2):50–61, 1995.
- [13] I. Daubechies. Time-frequency localization operators: A geometric phase space approach. *IEEE Transactions on Information Theory*, 34(4):605–612, 1988.
- [14] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5):961–1005, 1990.
- [15] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.
- [16] H.M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural. Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms. *Journal of the optical society of America A - Optics image sciences and vision*, 11(2):547 – 559, 1994.
- [17] Z. Yudong, Y. Jian-Fei, P. Phillips, Y. Jie, W. Shuihua, Z. Zheng, and S. Ping. A comprehensive survey on fractional Fourier transform. *Fundamenta Informaticae*, 151(1-4):1 – 48, 2017.
- [18] R.G. Stockwell. Localization of the complex spectrum: the S-transform. *IEEE Transactions on Signal Processing*, 44(4):993, 1996.
- [19] P.C. Gibson, M.P. Lamoureux, and G.F. Margrave. Letter to the editor: Stockwell and wavelet transforms. *Journal of Fourier analysis and applications*, 12(6):713 – 721, 2006.
- [20] R.G. Stockwell. A basis for efficient representation of the s-transform. *Digital Signal Processing*, 17(1):371 – 393, 2007.
- [21] H. Kalbkhani and M. G. Shayesteh. Stockwell transform for epileptic seizure detection from eeg signals. *Biomedical Signal Processing and Control*, 38:108 – 118, 2017.
- [22] M. Hamidia and A. Amrouche. A new robust double-talk detector based on the Stockwell transform for acoustic echo cancellation. *Digital Signal Processing*, 60:99 – 112, 2017.
- [23] P. Soo-Chang and H. Shih-Gu. STFT with adaptive window width based on the chirp rate. *IEEE Transactions on Signal Processing*, 60(8):4065 – 4080, 2012.

- [24] L. Xiumei, B. Guoan, S. Stanković, and A.M. Zoubir. Local polynomial Fourier transform: A review on recent developments and applications. *Signal Processing*, 91(6):1370 – 1393, 2011.
- [25] E. P. Wigner. On the quantum correction for thermodynamic equilibrium. *Physics Review*, 40:749–759, June 1932.
- [26] J. Ville. Theorie et applications de la notion de signal analytique. *Cables et Transmissions*, 2A:61–74, 1948.
- [27] L. Cohen. *Time-Frequency Analysis*. Signal Processing Series. Prentice-Hall, Upper Saddle River, NJ, USA, 1995.
- [28] L.B. Almeida. The fractional Fourier transform and time-frequency representations. *IEEE Transactions on Signal Processing*, 42(11):3084–3091, 1994.
- [29] R. Tao, Y.-L. Li, and Y. Wang. Short-time fractional Fourier transform and its applications. *IEEE Transactions on Signal Processing*, 58(5):2568, 2010.
- [30] S. Kadambe and G. F. Boudreaux-Bartels. A comparison of the existence of cross terms in the Wigner distribution and the squared magnitude of the wavelet transform and the short time Fourier transform. *IEEE Transactions on Signal Processing*, 40(10):2498–2517, Oct. 1992.
- [31] T. A. C. M. Claasen and W. F. G. Mecklenbräuker. The Wigner distribution - part II: Discrete-time signals. *Philips J. Res.*, 35:276–350, 1980.
- [32] A. H. Costa and G. F. Boudreaux-Bartels. An overview of aliasing errors in discrete-time formulations of time-frequency representations. *IEEE Transactions on Signal Processing*, 47(5):1463–1474, 1999.
- [33] S. M. Sussman. Least-squares synthesis of radar ambiguity functions. *IRE Transactions Information theory*, 8:246–254, April 1962.
- [34] J. E. Moyal. Quantum mechanics as a statistical theory. *Proc. Camb. Phil. Soc.*, 45:99–124, 1949.
- [35] M. Große Ruse, D. Hasselquist, B. Hansson, M. Tarka, and M. Sandsten. Automated analysis of song structure in complex bird songs. *Animal Behaviour*, 112:39–51, 2016.
- [36] B. W. Gillespie and L. E. Atlas. Optimizing time-frequency kernels for classification. *IEEE Transactions on Signal Processing*, 49(3):485–496, 2001.

- [37] M. Sandsten and J. Brynolfsson. Classification of bird song syllables using Wigner-Ville ambiguity function cross-terms. In *EUSIPCO*, Kos island, Greece, 2017.
- [38] A. W. Rihaczek. Signal energy distribution in time and frequency. *IEEE Transactions Information Theory*, 14:369–374, May 1968.
- [39] C. H. Page. Instantaneous power spectra. *J. of Applied Physics*, 23:103–106, January 1952.
- [40] H. Margenau and R. N. Hill. Correlation between measurements in quantum theory. *Progress of Theoretical Physics*, 26:722–738, 1961.
- [41] L. Cohen. Generalized phase-space distribution functions. *Jour. Math. Phys.*, 7:781–786, 1966.
- [42] P. J. Loughlin, J. W. Pitton, and L. E. Atlas. Bilinear time-frequency representations: new insights and properties. *IEEE Transactions on Signal Processing*, 41(2):750–767, 1993.
- [43] H.-I. Choi and W. J. Williams. Improved time-frequency representation of multi-component signals using exponential kernels. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37:862–871, June 1989.
- [44] J. Jeong and W. Williams. Kernel design for reduced interference distributions. *IEEE Transactions on Signal Processing*, 40:402–412, 1992.
- [45] J. G. Kirkwood. Quantum statistics of almost classical ensembles. *Physics review*, 44:31–37, 1933.
- [46] M. J. Levin. Instantaneous spectra and ambiguity functions. *IEEE Transactions on Information theory*, 10:95–97, 1964.
- [47] G. S. Cunningham and W. J. Williams. Kernel decomposition of time-frequency distributions. *IEEE Transactions on Signal Processing*, 42:1425–1442, June 1994.
- [48] D. Slepian and H.O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - I. *The Bell System Technical Journal*, 40:43–63, 1961.
- [49] H.J. Landau and H.O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - II. *The Bell System Technical Journal*, 40:65–84, 1961.

- [50] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - III. The dimension of the space of essentially time- and band-limited signals. *The Bell System Technical Journal*, 41:1295–1336, 1962.
- [51] D. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty - IV. Extensions to many dimensions; generalized prolate spheroidal functions. *The Bell System Technical Journal*, 43:3009–3057, 1964.
- [52] D. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty - V: The discrete case. *The Bell System Technical Journal*, 57(5):1371–1430, May-June 1978.
- [53] D. J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, Sept 1982.
- [54] R. G. Shenoy and T. W. Parks. The Weyl correspondence and time-frequency analysis. *IEEE Transactions on Signal Processing*, 42(2):318–331, 1994.
- [55] F. Cakrak and P. J. Loughlin. Multiple window time-varying spectral analysis. *IEEE Transactions on Signal Processing*, 49(2):448–453, 2001.
- [56] S. Aviyente and W. J. Williams. Multitaper marginal time-frequency distributions. *Signal Processing*, 86:279–295, 2006.
- [57] J. Xiao and P. Flandrin. Multitaper time-frequency reassignment for nonstationary spectrum estimation and chirp enhancement. *IEEE Transactions on Signal Processing*, 55(6):2851–2860, 2007.
- [58] I. Orović, S. Stanković, T. Thayaparan, and L. J. Stanković. Multiwindow S-method for instantaneous frequency estimation and its application in radar signal analysis. *IET Signal Processing*, 4(4):363–370, 2010.
- [59] M. Hansson-Sandsten. Optimal estimation of the time-varying spectrum of a class of locally stationary processes using Hermite functions. *EURASIP Journal on Advances in Signal Processing*, 2011. Article ID 980805.
- [60] M. Hansson-Sandsten and J. Sandberg. Optimization of weighting factors for multiple window spectrogram of event related potentials. *EURASIP Journal on Advances in Signal Processing*, 2010. Article ID 391798.
- [61] S. Aviyente and W. J. Williams. Multitaper reduced interference distribution. In *Proceedings of the tenth IEEE Workshop on Statistical Signal and Array Processing*, pages 569–573. IEEE, 2000.

- [62] W. J. Williams and S. Aviyente. Spectrogram decompositions of time-frequency distributions. In *Proceedings of the ISSPA*, pages 587–590. IEEE, 2001.
- [63] M. Hansson-Sandsten. Multitaper Wigner and Choi-Williams distributions with predetermined doppler-lag bandwidth and sidelobe suppression. *Signal Processing*, 91(6):1457–1465, 2011.
- [64] D. L. Jones and T. W. Parks. A high-resolution data-adaptive time-frequency representation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38:2127–2135, 1990.
- [65] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel. Measuring time-frequency information content using Rényi entropies. *IEEE Transactions on Information Theory*, 47(4):1391–1409, May 2001.
- [66] S. Aviyente and W. J. Williams. Minimum entropy time-frequency distributions. *IEEE Signal Processing Letters*, 12:37–40, January 2005.
- [67] V. Sucic, N. Saulig, and B. Boashash. Analysis of local time-frequency entropy features for nonstationary signal components time supports detection. *Digital Signal Processing*, 34:56 – 66, 2014.
- [68] M. Hansson-Sandsten. Classification of bird song syllables using singular vectors of the multitaper spectrogram. In *EUSIPCO*, Nice, France, 2015.
- [69] V. Sucic, N. Saulig, and B. Boashash. Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency renyi entropy. *EURASIP Journal on Advances in Signal Processing*, page 125, 2011.
- [70] T.-H. Sang and W. J. Williams. Rényi information and signal-dependent optimal kernel design. In *Proceedings IEEE Int. Conf. Acoustics, Speech and Signal Processing*, volume 2, pages 997–1000. IEEE, 1995.
- [71] M. Sandsten, I. Reinhold, and J. Starkhammar. Automatic time-frequency analysis of echo-location signals using matched Gaussian multitaper spectrogram. In *INTERSPEECH 2017*, Stockholm, Sweden, 2017.
- [72] N. Saulig, I. Orović, and V. Sucic. Optimization of quadratic time-frequency distributions using the local Rényi entropy information. *Signal Processing*, 129:17 – 24, 2016.
- [73] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal. I. fundamentals. *Proceedings of the IEEE*, 80(4):520–538, 1992.

- [74] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal. II. algorithms and applications. *Proceedings of the IEEE*, 80(4):540–568, 1992.
- [75] N. A. Khan and M. Sandsten. Time-frequency image enhancement based on interference suppression in Wigner-Ville distribution. *Signal Processing*, 127:80–85, 2016.
- [76] 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.
- [77] F. Auger and P. Flandrin. Improving the readability of time-frequency and time-scale representations by the reassignment method. *IEEE Transactions on Signal Processing*, 43:1068–1089, May 1995.
- [78] M. Hansson-Sandsten and J. Brynolfsson. The scaled reassigned spectrogram with perfect localization for estimation of Gaussian functions. *IEEE Signal Processing Letters*, 22(1):100–104, 2015.
- [79] P. Flandrin. A note on reassigned Gabor spectrograms of Hermite functions. *J Fourier Analysis and Applications*, 19(2):285–295, 2013.
- [80] M. Sandsten, J. Brynolfsson, and I. Reinhold. The matched window reassignment. In *EUSIPCO*, Rome, Italy, 2018.
- [81] J. Brynolfsson and M. Sandsten. Parameter estimation of oscillating Gaussian functions using the scaled reassigned spectrogram. *Signal Processing*, 150:20–32, 2018.
- [82] I. Reinhold, M. Sandsten, and J. Starkhammar. Objective detection and time-frequency localization of components within transient signals. *J. Acoust. Soc. Am. (JASA)*, 143(4):2368–2378, 2018.
- [83] J. Starkhammar, I. Reinhold, P. W. Moore, D. S. Houser, and M. Sandsten. Detailed analysis of two detected overlaying transient components within the echolocation beam of a bottlenose dolphin (*tursiops truncatus*). *J. Acoust. Soc. Am. (JASA)*, 145(4), 2019.
- [84] F. Auger, E. Chassande-Mottin, and P. Flandrin. Making reassignment adjustable: The Levenberg-Marquardt approach. In *Proceedings of the ICASSP*, pages 3889–3892. IEEE, 2012.

- [85] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.
- [86] N. E. Huang and Z. Wu. A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics*, 46(2), 2008.
- [87] H. Li, Z. Li, and W. Mo. A time varying filter approach for empirical mode decomposition. *Signal Processing*, 138:146 – 158, 2017.
- [88] I. Daubechies, J. Lu, and H.-T. Wu. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Applied & Computational Harmonic Analysis*, 30(2):243 – 261, 2011.
- [89] S. Wang, X. Chen, Y. Wang, G. Cai, B. Ding, and X. Zhang. Nonlinear squeezing time-frequency transform for weak signal detection. *Signal Processing*, 113:195 – 210, 2015.
- [90] Q. Jiang and B. W. Suter. Instantaneous frequency estimation based on synchrosqueezing wavelet transform. *Signal Processing*, 138:167 – 181, 2017.
- [91] M. B. Priestley. Evolutionary spectra and non-stationary processes. *J. Roy. Statistics Soc. ser. B*, 27:204–229, 1965.
- [92] R. A. Silverman. Locally stationary random processes. *IRE Transactions on Information Theory*, 3:182–187, 1957.
- [93] G. Matz and F. Hlawatsch. Nonstationary spectral analysis based on time-frequency operator symbols and underspread approximations. *IEEE Transactions on Information Theory*, 52(3):1067–1086, March 2006.
- [94] H. Cramér. On the theory of stationary random processes. *Annals of Mathematics*, 41:215–230, Jan. 1940.
- [95] P. Flandrin. *Time-Frequency/Time-Scale Analysis*, volume 10 of *Wavelet analysis and its applications*. Academic Press, 1999.
- [96] W. Mark. Spectral analysis of the convolution and filtering of nonstationary stochastic processes. *Journal of Sound and Vibration*, 11(1):19–63, 1970.
- [97] W. Martin and P. Flandrin. Wigner-Ville spectral analysis of nonstationary processes. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(6):1461–1470, 1985.

- [98] A. M. Sayeed and D. L. Jones. Optimal kernels for nonstationary spectral estimation. *IEEE Transactions on Signal Processing*, 43(2):478–491, Feb. 1995.
- [99] I. Djurovic, V. Katkovnik, and L. Stanković. Median filter based realizations of the robust time-frequency distributions. *Signal Processing*, 81(8):1771–1776, 2001.
- [100] P. Wahlberg and M. Hansson. Kernels and multiple windows for estimation of the Wigner-Ville spectrum of Gaussian locally stationary processes. *IEEE Trans. on Signal Processing*, 55(1):73–87, January 2007.
- [101] M. Hansson and G. Salomonsson. A multiple window method for estimation of peaked spectra. *IEEE Trans. on Signal Processing*, 45(3):778–781, March 1997.
- [102] M. Hansson. Optimized weighted averaging of peak matched multiple window spectrum estimates. *IEEE Trans. on Signal Processing*, 47(4):1141–1146, April 1999.
- [103] R. Anderson and M. Sandsten. Inference for time-varying signals using locally stationary processes. *Journal of Computational and Applied Mathematics*, 347:24–35, 2019.
- [104] R. Anderson and M. Sandsten. Classification of EEG signals based on mean-square error optimal time-frequency features. In *EUSIPCO*, Rome, Italy, 2018.
- [105] R. Anderson, P. Jönsson, and M. Sandsten. Stochastic modeling and optimal time-frequency estimation of task-related HRV. *Applied Sciences*, 9(23), 2019.