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

An Introduction

Maria Sandsten

2016
6.4 Reassigned spectrogram ................................................. 62
6.5 Scaled reassigned spectrogram ........................................ 64

7 Stochastic time-frequency analysis ................................... 71
  7.1 Definitions of non-stationary processes .......................... 72
  7.2 The mean square error optimal kernel ............................ 76
Chapter 1

Introduction

Signals in nature vary greatly and sometimes it can be difficult to characterize them. We usually differ between deterministic and stochastic (random) signals where a deterministic signal is explicitly known and a stochastic process is one realization from a collection of signals which are characterized by different properties, such as expected value and variance. We also differ between a time-invariant and time-varying signal and between stationary and non-stationary processes. Time is a fundamental way of studying a signal but we can also study the signal in other representations, where one of the most important is frequency. Time-frequency analysis of time-varying signals and non-stationary processes, has been a field of research for a number of years.

1.1 The history of spectral analysis

The short description here just intend to give an orientation on the history behind spectral analysis in signal processing and for a more thorough description see, e.g., [1].

Sir Isaac Newton (1642-1727) performed an experiment in 1704, where he used a glass prism to resolve the sunbeams into the colors of the rainbow, a spectrum. The result was an image of the frequencies contained in the sunlight. Jean Baptiste Joseph Fourier (1768-1830) invented the mathematics to handle discontinuities in 1807 where the idea was that a discontinuous function can be expressed as a sum of continuous frequency functions. This idea was seen as absurd from the established scientists at that time, including Pierre Simon Laplace (1749-1827) and Joseph Louis Lagrange (1736-1813). The application behind the idea was the most basic problem regarding heat, a discontinuity in temperature when hot and cold objects were put together. However, the idea was eventually accepted and is what we today call the Fourier expansion.

Joseph von Fraunhofer (1787-1826) invented the spectroscope in 1815 for measurement of the index of refraction of glasses. He discovered what today has become known as the Fraunhofer lines. Based on this, an important discovery was made by Robert Bunsen (1811-1899) in the 1900th century. He studied the light from a burning rag soaked with
salt solution. The spectrum from the glass prism consisted of lines and especially a bright yellow line. Further experiments showed that every material has its own unique spectrum, with different frequency contents. The glass prism was actually the first spectrum analyser. Based on this discovery, Bunsen and Gustav Kirchhoff (1824-1887) discovered 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 one filter is squared and averaged and the result is the average power at the center-frequency of the filter. Another bandpass-filter with different center-frequency will give the average power at another frequency. With many filters of different center-frequencies, an image of the frequency content is obtained. This spectrum analyzer performance is equivalent to modulating (shifting) the signal in frequency and using one narrow-banded lowpass-filter. The modulation procedure will move the power of the signal at the modulation frequency to the base-band. Thereby only one filter, the lowpass-filter, is needed. The modulation frequency is different for every new spectrum value. This interpretation is called the heterodyne technique.

1.2 Spectrum analysis

A first step in modern spectrum analysis, in the sense of sampled discrete-time analysis, was made by Sir Arthur Schuster already in 1898, where he found hidden periodicities in sunspot numbers using Fourier series, or what is known today as the periodogram, [2]. This method has been frequently used for spectrum estimation ever since James Cooley and John Tukey 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 already in 1805, long before the existence of any computers. He did not publish the result though. The FFT is the efficient way of calculating the Fourier transform (series) 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. A recent and more thorough review of Cooley’s and Tukey’s work in time series analysis and spectrum estimation can be found in [6].

To find a reliable spectrum with optimal frequency resolution from a sequence of (short length) data has become a large research field. For time-invariant or stationary time-series, two major approaches can be found in the literature, the use of windowed periodograms, including lag-windowed covariance estimates, which represents the non-parametric approach or the classical spectral analysis and the parametric approach or the modern spectral analysis model which includes methods related to the autoregressive (AR) technique. Developments and combinations of these two approaches have during the last decades been utilized in a huge number of applications, [7]. A fairly recent, short, pedagogical overview is given in, [8, 9]. The Slepian functions of the Thomson
multitapers, [10] are well established in spectrum analysis today, and are recognized to give orthonormal spectra, (for stationary white noise) as well as to be the most localized tapers in the frequency domain.

High-resolution spectral estimation of sinusoids with well known algorithms such as MUSIC and others is also a very large research field and descriptions of some of these algorithms could be found in, e.g., [7]. These algorithms are outstanding in certain applications but in some cases the usual periodogram is applicable even for sinusoids, e.g., to detect or estimate a single sinusoid disturbed by noise, [11].

All these techniques for spectrum estimation rely on the fact that the properties, such as frequencies and amplitudes, of the signal change slowly with time. If this is true, also the time-dependent frequency spectrum of the signal can be estimated and visualized, using short-time versions of above methods. However, if the signal is really time-varying, e.g. with step-wise changes, or actually is a non-stationary process, other methods should be applied.

We give an example where the need for time-frequency analysis becomes obvious. We analyze two obviously different time-varying signals, see Figure 1.1a) and b). However, their periodograms will be exactly the same, see Figure 1.1c) and d), and we can be sure of that this is not the total picture. We see clearly from Figure 1.1a) and b) that they are different but in what sense? With use of time-frequency analysis, in the form of spectrograms, these differences become visible, Figure 1.2a) and b), where red color indicates high power and blue color is low power.

1.3 Time-frequency analysis

Eugene Wigner (1902-1995) 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 Wigner distribution, almost as well known as the spectrogram, was suggested in a famous paper from 1932, ”because it seemed to be the simplest”, [12]. The paper is actually in the area of quantum mechanics and not at all in signal analysis and the use of ”distribution” should not be understood in the sense of statistics. The drawback of the Wigner distribution is that it gives cross-terms, that is located in the middle between and can be twice as large as the different signal components. A large area of research has been developed with the ambition to reduce these cross-terms, where different time-frequency kernels are proposed.

The design of these kernels are usually made in the so called ambiguity domain, also known as the doppler-lag domain, where the ambiguity function is defined as the Fourier transform in both directions of the Wigner distribution. 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 [13], 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, [14], and by José Enrique Moyal, [15], suggesting the name **characteristic function**, where the concept of the Wigner distribution also was used for signal analysis for the first time. In his paper, [14], Ville also redefined the **analytic signal**, where a real-valued signal is converted into a complex-valued signal of non-negative frequency content. The Wigner distribution of the analytic signal, which is usually applied in practice today, is therefore called the **Wigner-Ville distribution**.

Dennis Gabor (1900-1979) also suggested a representation of a signal in two dimensions, time and frequency. In his paper from 1946, [16], Gabor defined certain elementary signals, one ”quantum of information”, that occupies the smallest possible area in the two-dimensional space, the ”information diagram”. He had a large interest in holography and in 1948 he carried out the basic experiments in holography, at that time called ”wavefront reconstruction”. In 1971 he received the Nobel Prize for his discovery of the principles underlying the science of holography. The Gabor expansion was related by Martin Bastiaans to the **short-time Fourier transform** and the spectrogram in 1980, [17].

Figure 1.1: Two different signals with exactly the same periodograms.
Figure 1.2: Spectrograms of the chirp and impulse signals, (Red color: high power, blue color: low power).

After the invention of the spectrogram and the Wigner-Ville distribution, a lot of other distributions and methods were suggested, such as the Rihaczek distribution, [18], the Page distribution, [19] and the Margenau-Hill or Levin distribution, [20]. In 1966 an important formulation was made by Leon Cohen, also in quantum mechanics, [21], which included these and an infinite number of other methods, known as the Cohen’s class. The formulation by Cohen was restricted with constraints of the doppler-lag kernels so the so called marginals should be satisfied. Some years later a quadratic class of methods have been defined, also including kernels which do not satisfy the marginals, but otherwise fulfill the properties of the Cohen’s class. We could however note, that Cohen’s class from the beginning also included signal-dependent kernels which can not be included in the quadratic class as the methods then no longer are a quadratic form of the signal. The general quadratic class, commonly referred to as the Cohen’s class, of time-frequency representation methods is the most applied today and a huge number of kernels can be found for different representations. The theories and methods of time-frequency analysis are usually defined in continuous-time and frequency and the the first discrete-time and
discrete-frequency (DT-DF) kernel was introduced by Claassen and Mecklenbräuker, [22], as late as 1980. The transformations in the DT-DF case are not always straightforward, [23, 24].
Chapter 2

Spectrogram and Wigner distribution

A spectrogram is a spectral representation that shows how the spectral density of a signal varies with time. Other names that are found in different applications are spectral waterfall or sonogram. The time domain data is divided in shorter sub-sequences, which usually overlaps, and for each sequence, the calculation of the squared magnitude of the Fourier transform 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, squared magnitude). The Wigner distribution, another famous spectral representation, comes from the area of quantum mechanics and has the best possible time- and frequency concentration of a signal. However, this time-frequency representation has severe drawbacks, such as cross-terms and resulting negative values in the time-frequency spectrum.

2.1 Fourier transform and spectrum analysis

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

\[
X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft}dt, \tag{2.1}
\]

where the signal \( x(t) \) can be recovered by the inverse Fourier transform,

\[
x(t) = \mathcal{F}^{-1}\{X(f)\} = \int_{-\infty}^{\infty} X(f)e^{i2\pi ft}df. \tag{2.2}
\]

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,
\[ S_x(f) = |X(f)|^2. \]  

(2.3)

We should also remember that using the Wiener-Khintchine theorem, the power spectrum of a zero-mean stationary stochastic process \( x(t) \), can be calculated as the Fourier transform of the covariance function \( r_x(\tau) \),

\[ S_x(f) = \mathcal{F}\{r_x(\tau)\} = \int_{-\infty}^{\infty} r_x(\tau)e^{-i2\pi ft}\,d\tau, \]

(2.4)

where \( r_x(\tau) \) is defined as

\[ r_x(\tau) = \mathcal{E}[x(t - \tau)x^*(t)], \]

(2.5)

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

\[ r_x(\tau) = \mathcal{F}^{-1}\{S_x(f)\} = \int_{-\infty}^{\infty} S_x(f)e^{i2\pi ft}\,d\tau. \]

(2.6)

### 2.2 The 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

\[ X(t, f) = \int_{-\infty}^{\infty} x(t_1)h^*(t_1 - t)e^{-i2\pi f t_1}\,dt_1, \]

(2.7)

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

\[ S_x(t, f) = |X(t, f)|^2, \]

(2.8)

which is used very frequently for analyzing time-varying and non-stationary signals. We can also extend the Wiener-Khintchine theorem, and define the power spectrum of a zero-mean non-stationary stochastic process \( x(t) \), to be calculated as the Fourier transform of the time-dependent covariance function \( r_x(t, \tau) \),
S_{xx}(t, f) = \mathcal{F}\{r_x(t, \tau)\} = \int_{-\infty}^{\infty} r_x(t, \tau) e^{-i2\pi f \tau} d\tau, \quad (2.9)

where \( r_x(t, \tau) \) is defined as

\[ r_x(t, \tau) = \mathcal{E}[x(t - \tau)x^*(t)]. \quad (2.10) \]

Using the spectrogram, the signal is divided into several shorter pieces, where a spectrum is estimated from each piece, which gives us information about where in time different frequencies occur. There are several advantages of using the spectrogram, e.g., fast implementation using the fast Fourier transform (FFT), easy interpretation and the clear connection to the periodogram.

We illustrate this with an example: In Figure 2.1a) a signal consisting of several frequency components is shown. A musical interpretation is three tones of increasing height. A common estimate of the spectrum is defined by the Fourier transform of a windowed sampled data vector, \( \{x(n), n = 0, 1, 2, \ldots N - 1\} \), \( (N = 512) \), where \( t = n/F_s \), \( (F_s \) is the sample frequency and in this example \( F_s = 1) \). The (modified) periodogram is defined as

\[ S_x(l) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n)h(n)e^{-i2\pi n \frac{l}{L}} \right|^2, \quad l = 0 \ldots L - 1 \quad (2.11) \]

where the frequency scale is \( f = l/L \cdot F_s \). The window function \( h(n) \) is normalized according to

\[ h(n) = \frac{h_1(n)}{\sqrt{\frac{1}{N} \sum_{n=0}^{N-1} h_1^2(n)}}, \quad (2.12) \]

where \( h_1(n) \) can be any window function, including \( h_1(n) = 1 \), (the rectangle window).

The periodogram (i.e. using a rectangle window), in Figure 2.1b) shows three frequency components but can not give any information on how these signals are located in time. Figure 2.1c) presents the modified periodogram (Hanning window), which is a very common analysis window, giving an estimated spectrum that seems to consist of just one strong frequency component and two weaker. These examples show how important it is to be careful using different spectral estimation techniques. The spectrogram is defined as

\[ S_x(n, l) = \frac{1}{M} \left| \sum_{n_1=0}^{N-1} x(n_1)h(n_1 - n + M/2)e^{-i2\pi n_1 \frac{l}{T}} \right|^2, \quad (2.13) \]
Figure 2.1: A three-component signals with increasing 'tone', frequency; a) data; b) periodogram; c) modified periodogram

for time-discrete signals and frequencies, where the window function is of length $M$ and normalized according to Eq. (2.12) with $N = M$. The resulting spectrogram of the three musical tones is presented in Figure 2.2 and shows a clear view of three frequency components and their locations in time and frequency.

The length (and shape) of the window function $h(t)$ is very important as it determines the resolution in time and frequency, as shown in Figure 2.3. The signal shown in Figure 2.3a) consists of two Gaussian windowed sinusoidal components located at the same time interval around $t = 100$, but with different frequencies, $f_0 = 0.07$ and $f_0 = 0.17$ and another component with frequency $f_0 = 0.07$ located around $t = 160$. The three components show up clearly in Figure 2.3c) where the window length of a Hanning window is $M = 32$. Figures 2.3b) and d) show the result when too short or too long window lengths are chosen. For a too short window length, Figure 2.3b), the two components located at the same time interval are smeared together as the frequency resolution becomes really bad for a window length of $M = 8$. With a long window, Figure 2.3d), the frequency
resolution is good but now the time resolution is so bad that the two components with the same frequency are smeared together.

The time- and frequency resolutions have to fulfill the so called uncertainty inequality, to which we will return to later. This gives a strong limitation on using the spectrogram for signals where the demand on time- and frequency resolution is very high. For now, we just note that usually a proper trade-off between time- and frequency resolution is found by choosing the length of the window about the same size as the time-invariance or stationarity of the individual signal components. The following simple calculation example show this: If the signal and the window function are

\[
x(t) = \left(\frac{\beta}{\pi}\right)^\frac{1}{4} e^{-\frac{\beta}{2}t^2}, \quad h(t) = \left(\frac{\alpha}{\pi}\right)^\frac{1}{4} e^{-\frac{\alpha}{2}t^2},
\]

respectively, then the resulting spectrogram can be calculated to be

\[
\text{Spectrogram, Hanning window, } M=128
\]
Figure 2.3: Spectrogram examples with different window lengths.

\[ S_x(t, f) = \frac{2\sqrt{\alpha \beta}}{\alpha + \beta} e^{-\frac{\alpha \beta}{\alpha + \beta} t^2 - \frac{1}{2\pi} \frac{1}{\alpha + \beta} \pi^2 f^2}. \] (2.15)

Studying the spectrogram one can note that the area of the ellipse \( e^{-1} \) down from the peak value is

\[ A = \frac{\alpha + \beta}{\sqrt{\alpha \beta} \pi}. \] (2.16)

Differentiation gives \( \alpha = \beta \) and a minimum area \( A = 2\pi \). The conclusion of this small example is that the a minimum time-frequency spread is given if the window length is equal to the signal component length for a signal with Gaussian envelope using a Gaussian function as window.

In certain applications where the lengths of the signal components are very different for different components adaptive windows can be used, [25, 26]. A time-variable Gaussian window is then defined as
\[ h_{t_1}(t) = \left(\frac{c_{t_1}}{\pi}\right) e^{-\frac{c_{t_1}}{2} t^2}, \quad (2.17) \]

for certain time-points defined by \( t_1 \). The adaptation of the parameter \( c_{t_1} \) of the window can be made in different ways, e.g., using locally maximized concentration

\[
\max_{c_{t_1}} \frac{\int |S_x(t_1, f_1)|^4 df_1}{\left( \int |S_x(t_1, f_1)|^2 df_1 \right)^2}, \quad (2.18)
\]

This way of maximizing the concentration is similar to maximizing other measures of sharpness, focus, peakiness or minimizing e.g., Rényi entropy.

### 2.3 The Wigner distribution and Wigner spectrum

The Wigner distribution defined for a deterministic time-varying signal is

\[
W_x(t, f) = \int_{-\infty}^{\infty} x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) e^{-i2\pi f \tau} d\tau, \quad (2.19)
\]

where we find

\[
r_x(t, \tau) = x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}), \quad (2.20)
\]

which for a non-stationary process, is actually the \textbf{instantaneous autocorrelation function}, (IAF), as

\[
r_x(t, \tau) = \mathcal{E}[x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2})], \quad (2.21)
\]

and we find an extension of the time-varying spectral density given as

\[
S_x(t, f) = \int_{-\infty}^{\infty} r_x(t, \tau) e^{-i2\pi f \tau} d\tau, \quad (2.22)
\]

which is well in accordance with Eq. (2.19). Note the difference between the two IAFs in Eq. (2.21) and Eq. (2.10). The formulation based on Eq. (2.21) and Eq. (2.22) is called the \textbf{Wigner spectrum} which also fulfills the basic properties of the Wigner distribution. As most of the research connected to the stochastic Wigner spectrum stems from the deterministic Wigner distribution, we start by introducing the properties of the Wigner distribution:
• The Wigner distribution is always real-valued even if the signal is complex-valued. For real-valued signals the frequency domain is symmetrical, $W_x(t, f) = W_x(t, -f)$, (compare with the definition of spectrum and spectrogram for real-valued signals).

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

\[
\int_{-\infty}^{\infty} W_x(t, f) df = |x(t)|^2, \quad \text{Time marginal} \tag{2.23}
\]

and

\[
\int_{-\infty}^{\infty} W_x(t, f) dt = |X(f)|^2. \quad \text{Frequency marginal} \tag{2.24}
\]

If the marginals are satisfied, the total energy condition is also automatically satisfied,

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(t, f) dt df = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df = E_x, \tag{2.25}
\]

where $E_x$ is the energy of the signal.

• If we shift the signal in time or frequency, the Wigner distribution will be shifted accordingly, it is **time-shift and frequency-shift invariant**, i.e., if $y(t) = x(t-t_0)$ then

\[
W_y(t, f) = W_x(t, f), \tag{2.26}
\]

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

\[
W_z(t, f) = W_x(t, f - f_0). \tag{2.27}
\]

• The Wigner distribution has always better resolution than the spectrogram for a so called mono-component signal, i.e., a single component signal. For the same signal as in the spectrogram window length optimization example, i.e.,

\[
x(t) = (\frac{\beta}{\pi})^{\frac{1}{4}} e^{-\frac{\beta}{4} t^2}, \tag{2.28}
\]
the Wigner distribution can be calculated as

$$W_x(t, f) = 2e^{-(\beta t^2 + \frac{1}{\beta} 4\pi^2 f^2)}.$$ (2.29)

Again, studying the area of this ellipse $e^{-1}$ down from the peak value gives $A = \pi$ which is half of the corresponding area of the spectrogram!

- For mono-component complex-valued sinusoids or linear frequency chirp signal that exists for all values of $t$ the Wigner distribution gives exactly the instantaneous frequency, i.e. a perfectly localized time-frequency representation. We find the Wigner distribution of $x(t) = e^{i2\pi f_0 t}$, $-\infty < t < \infty$, (complex-valued sinusoidal signal of frequency $f_0$), to be

$$W_x(t, f) = \int e^{i2\pi f_0 (t+\frac{\tau}{2})}e^{-i2\pi f_0 (t-\frac{\tau}{2})}e^{-i2\pi f\tau}d\tau$$

$$= \int e^{i2\pi (f_0 \tau - f)\tau}d\tau = \delta(f - f_0).$$

For the chirp signal $x(t) = e^{i2\pi \frac{\beta}{2} t^2 + i2\pi f_0 t}$, we compute the Wigner distribution as

$$W_x(t, f) = \int e^{i2\pi \left( \frac{\beta}{2}(t+\frac{\tau}{2})^2 + f_0 (t+\frac{\tau}{2}) \right)}e^{-i2\pi \left( \frac{\beta}{2}(t-\frac{\tau}{2})^2 + f_0 (t-\frac{\tau}{2}) \right)}e^{-i2\pi f\tau}d\tau$$

$$= \int e^{i2\pi (\beta t + f_0 - f)\tau}d\tau = \delta(f - f_0 - \beta t).$$

We should also mention that for $x(t) = \delta(t - t_0)$, the Wigner distribution is given as $W_x(t, f) = \delta(t - t_0)$. The 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 these and some other simple deterministic signals, the calculations of the Wigner distribution is possible to perform by hand, see also examples in the exercises. 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? The problem is well-known as cross-terms, which we will look more closely into in the next chapter.
Chapter 3

Cross-terms and the Wigner-Ville distribution

The Wigner distribution gives cross-terms, that is 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 disturbances. This makes the Wigner distribution useless for most signals that are not just toy signals. However, one important step on the way to make the Wigner distribution more useful is to calculate the Wigner distribution of the analytic signal, redefined by Ville, [14]. This form is usually applied in practice today, and is therefore called the Wigner-Ville distribution.

3.1 Cross-terms

The Wigner distribution has weak time support as it is limited by the duration of the signal,

\[ x(t) = 0 \text{ when } t < t_1 \text{ and } t > t_2, \]  
\[ W_x(t, f) = 0 \text{ for } t < t_1 \text{ and } t > t_2. \]  

(3.2)

where \( t_1 < t_2 \), then

Similarly it has weak frequency support and is thereby limited by the bandwidth of the signal,

\[ X(f) = 0 \text{ when } f < f_1 \text{ and } f > f_2, \]  

(3.3)
where \( f_1 < f_2 \), then

\[
W_x(t, f) = 0 \quad \text{for} \quad f < f_1 \quad \text{and} \quad f > f_2. \tag{3.4}
\]

However, 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. 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. The properties of strong time- and frequency supports are not fulfilled.

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

\[
W_x(t, f) = W_{x_1}(t, f) + W_{x_2}(t, f) + 2\Re[W_{x_1,x_2}(t, f)], \tag{3.5}
\]

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

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

is called cross-term. The cross-term will always be present and will be located midway between the auto-terms. The cross-term will oscillate proportionally to the distance between the auto-terms and the direction of the oscillations will be orthogonal to the line connecting the auto-terms, see Figure 3.1, which shows the Wigner distribution of the sequence displayed in Figure 2.1, (only the part with frequencies larger than zero). If we compare Figure 3.1 with Figure 2.1 we can locate the actual components, the auto-terms. For simplicity, they are called component \( F_1 \), \( F_2 \) and \( F_3 \). We find two of them as yellow components and the middle one, \( F_2 \), marked with a red-yellow square-pattern. This pattern consists of the middle component and the added corresponding cross-term from the outer components, (oscillating and located in the middle between \( F_1 \) and \( F_3 \)). Then there are a cross-term in the middle between \( F_1 \) and \( F_2 \) and one between \( F_2 \) and \( F_3 \). 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 a major draw-back of the Wigner distribution. This example, only including three components, shows that the result using the Wigner distribution could easily be misinterpreted.
3.2 The analytic signal and the Wigner-Ville distribution

A signal $z(t)$ is **analytic** if

$$Z(f) = 0 \text{ for } f < 0,$$

where $Z(f) = \mathcal{F}\{z(t)\}$, i.e. the analytic signal requires the spectrum to be identical to the spectrum for the real signal for positive frequencies and to be zero for negative frequencies. The definition

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

where $x(t)$ is replaced by the analytic signal $z(t)$, is called the **Wigner-Ville distribution**.
This is the most applied definition today.

We illustrate the advantages of the Wigner-Ville distribution with an example: Two real-valued sinusoids, with frequencies \( f_0 = 0.15 \) and \( f_0 = 0.3 \), windowed with time-localized Gaussian windows, around \( t = 80 \) and \( t = 200 \) respectively, are added together and form a two-component sequence \( (N = 256) \), see Figure 3.2. The spectrogram is shown in Figure 3.3a), where we clearly can identify the two components and their time-and frequency locations. As the signal is real-valued, we also get two components located at \( f = -0.15 \) and \( f = -0.3 \), although we usually not visualize the negative frequency part of the spectrum. In Figure 3.3b) the Wigner distribution is visualized, using the discrete Wigner distribution, defined as

\[
W_x[n, l] = 2 \sum_{m=0}^{\min(n,N-1-n)} x[n + m]x^*[n - m]e^{-i2\pi m \frac{l}{L}},
\]

for a discrete-time signal \( x[n] \), \( n = 0 \ldots N - 1 \), where \( L \) is the number of frequency values and \( f = l/(2 \cdot L) \). Now the picture becomes more difficult to interpret. The low-frequency component shows up at \( t = 80 \) and \( f = \pm 0.15 \) and the high-frequency component at \( t = 200 \) and \( f = \pm 0.3 \) as expected, (yellow-orange components), but we also get other contributions that actually look like components, one at \( t = 80, f = \pm 0.35 \) and the other at \( t = 200, f = \pm 0.2 \). These components can not be recognized as oscillating cross-terms, but could instead be referred to aliasing. This is a well known fact for the Discrete Fourier Transform (DFT) although that we are used to that the period is one, not 0.5 which is the period if the discrete Wigner distribution. The reason for the period of 0.5 is that we actually have down-sampled the signal a factor 2, \( \tau/2 \) is replaced with \( m \), when the discrete Wigner distribution is calculated.

A consequence of using the discrete Wigner distribution in this form, is that \( f_0 \) (here 0.3) above \( f = 0.25 \) will alias to a frequency \( 0.5 - f_0 \) (here 0.2) and will show up in the spectrum image as a component. The only indication of that this frequency component is not a actual component, is that there is not a cross-term midway between the frequency component at \( f = 0.3 \) and the one at \( f = 0.2 \), the actual cross-terms are found between \( f = 0.15 \) and \( f = 0.3 \). We also see the cross-terms midway between the frequency located at \( f = f_0 \) and \( f = -f_0 \), (at \( f = 0 \)) as well as the cross-terms midway between e.g., \( f = 0.15 \) and \( f = -0.3 \), (located at \( f = (0.15 + 0.3)/2 \approx 0.22 \). The picture of a two-component real-valued signal becomes very complex and the interpretation difficult.

One elegant solution is to use the analytic signal and calculate the Wigner-Ville distribution, i.e., \( x[n] \) in Eq. (??) is replaced by the analytic signal \( z[n] \), with resulting time-frequency spectrum shown in Figure 3.3c). The negative frequency components \( f = -0.15 \) and \( f = -0.3 \) disappear and the only cross-term is between the two components at positive frequencies. The spectrum repeats with period 0.5 but as we know from the beginning that no frequencies are located at negative frequencies (according to the definition of the analytic signal), we can just ignore the repeated picture at negative fre-
quencies. We also see that no aliasing occur for the frequency at $f = 0.3$, (as there are no negative frequencies that can repeat above 0.25). This means, that the signal frequency content of a analytic signal can be up to $f = 0.5$ as usual for discrete-time signals. There are other possible methods to avoid aliasing for signal frequencies higher than 0.25 and still compute the Wigner distribution for the real-valued signals, e.g., could the signal be up-sampled a factor two before the computation of the IAF, and thereby the aliasing is moved to the frequency 0.5 as in usual sampling.

![Figure 3.2: A two-component real-valued sequence.](image)

### 3.3 Cross-terms of the spectrogram

Actually, the spectrogram also give cross-terms but they show up as strong components mainly when the signal components are close to each other. This is seen in the following example: If we have two signals $x_1(t)$ and $x_2(t)$, the spectrum, the squared absolute value of the Fourier transform, of the resulting sum of the signals, $x(t) = x_1(t) + x_2(t)$ is

$$|X(f)|^2 = |\mathcal{F}\{x(t)\}|^2 = |X_1(f) + X_2(f)|^2$$

$$= |X_1(f)|^2 + |X_2(f)|^2 + X_1(f)X_2^*(f) + X_1^*(f)X_2(f)$$

$$= |X_1(f)|^2 + |X_2(f)|^2 + 2\Re[X_1(f)X_2^*(f)].$$

When the spectrum of $X_1(f)$ and $X_2(f)$ covers the same frequency interval the density of the sum involves not only the sum of the spectra for the two different signals but also a third term $2\Re[X_1(f)X_2^*(f)]$. This term is effected by different phases of the signal as well as the magnitude function. This term is called leakage but could also be defined as a cross-term. We extend these considerations to the spectrogram expressed by
\[ |X(t, f)|^2 = |\mathcal{F}\{x(t_1)h^*(t_1 - t)\}|^2 = |X_1(t, f) + X_2(t, f)|^2 = |X_1(t, f)|^2 + |X_2(t, f)|^2 + 2\Re[X_1(t, f)X_2^*(t, f)], \quad (3.9) \]

where the third term is included if the same time-frequency region is covered by \(X_1(t, f)\) and \(X_2(t, f)\). This is the main difference if we compare with the Wigner distribution. For the spectrogram, the cross-term component shows up if the signal components \(X_1(t, f)\) and \(X_2^*(t, f)\) overlap. This is seen in Figure 3.4, where the two complex-valued sinusoids are located at a large distance and a small distance. The cross-term from the Wigner distribution is clearly seen, placed midway between the two components. For the spectrogram there is no cross-term present when the frequency distance between the two components is large, but when the two components are more closely spaced the effect of the overlapped components becomes visible, Figure 3.5. For a thorough analysis and comparison, see [27].
Figure 3.4: The spectrogram and the Wigner distribution of two complex Gaussian windowed sinusoids.
Figure 3.5: The spectrogram and the Wigner distribution of two more closely spaced complex Gaussian windowed sinusoids.
Chapter 4

Ambiguity functions and ambiguity kernels

A huge number of methods with the aim to reduce cross-terms of the Wigner-Ville distribution can be found in the research literature. The aim is to design different time-frequency kernels which are usually defined and optimized in the so called ambiguity domain. The main reason is that auto-terms and cross-terms relocalize in a beneficial way.

4.1 Definition and properties

The ambiguity function is defined as

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

where usually the analytic signal \( z(t) \) is used. (Without any restrictions, \( z(t) \) could be replaced by \( x(t) \)). We note that the formulation is similar to the Wigner-Ville distribution, the difference is that the Fourier transform is made in the \( t \)-variable instead of the \( \tau \)-variable, giving the ambiguity function dependent of the two variables \( \nu \) and \( \tau \). The Fourier transform of the IAF in the variable \( t \) gives

\[ A_s(\nu, \tau) = \int_{-\infty}^{\infty} r_x(t, \tau)e^{-i2\pi\nu t}dt, \]  
(4.2)

which we call the ambiguity spectrum.

To be able to interpret the ambiguity function we compare the Wigner-Ville distribution and real-valued part of the ambiguity function for different signals. (The ambiguity function will almost always be complex-valued.) In the first case in Figure 4.1a) and b)
we see that a Gaussian function centered at time $t = 40$ with centre frequency $f = 0.05$ will be a Gaussian function in the time-frequency representation at the corresponding locations in time and frequency. The ambiguity function however, will be located at $\tau = 0$ and $\nu = 0$, where the frequency- and time-shift will show up as the oscillation frequency and direction on the oscillations respectively. In Figure 4.1c) and d), the Gaussian function with larger centre frequency $f = 0.2$ located at $t = 0$ also shows up in the centre of the ambiguity function with a different oscillation frequency and direction of the oscillation. A mono-component signal will always relocate to $\tau = 0$ and $\nu = 0$. However, the clear advantage of the ambiguity function shows up in the last example, Figure 4.1e) 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 4.1e), where we find the cross-term located in the middle between the auto-terms, at time location $(t_b + t_a)/2$ and frequency location $(f_b + f_a)/2$. In the ambiguity function, Figure 4.1f), the signal components add together at the centre, where the cross-term(s) now show up located at $\nu = f_b - f_a$ and $\tau = t_b - t_a$ as well as at $\nu = f_a - f_b$ and $\tau = t_a - t_b$.

As the signal (auto-term) components always will be located at the centre of the ambiguity function, independently of where they are located in the time-frequency plane, and the cross-terms always will be located away from the centre, a natural approach is to keep the terms located at the centre and reduce the components away from the centre of the ambiguity function.

### 4.2 Ambiguity kernels

A filtered ambiguity function is defined as the element-wise multiplication of the ambiguity function and the ambiguity kernel,

$$A_z^F(\nu, \tau) = A_z(\nu, \tau) \cdot \phi(\nu, \tau).$$

The corresponding time-frequency kernel is given by,

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

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

$$W_z^F(t, f) = W_z(t, f) * * \Phi(t, f),$$

We can note that the original Wigner-Ville distribution has the simple ambiguity domain
Figure 4.1: A time- and frequency shifted Gaussian function and the representation as Wigner-Ville distribution and ambiguity function (real part); a) and b), $f_1 = 0.1$, $t_1 = 20$; c) and d), $f_2 = 0.3$, $t_2 = 0$; e) f), The sum of the two Gaussian functions
Ambiguity functions and ambiguity kernels

kernel \( \phi(\nu, \tau) = 1 \) for all \( \nu \) and \( \tau \), and the corresponding time-frequency (non-)smoothing kernel is \( \Phi(t, f) = \delta(t)\delta(f) \).

To learn about the design properties of the ambiguity kernel we recall the time marginal, and if we put \( \tau = 0 \) in Eq.(4.1) we get

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

which is the Fourier transform of the time marginal. Similarly, using \( Z(f) = \int_{-\infty}^{\infty} z(t)e^{-i2\pi tf} dt \) we get

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

which is the inverse Fourier transform of the frequency marginal. The frequency marginal is the usual spectral density and the inverse Fourier transform of the spectral density, \( A_z(0, \tau) \), can then be interpreted as the usual covariance function. We can conclude from this, that in order to preserve the marginals of the time-frequency distribution, the ambiguity kernel must fulfill

\[
\phi(0, \tau) = \phi(\nu, 0) = 1.
\]

From this we also conclude that

\[
\phi(0, 0) = 1,
\]

to preserve the total energy of the signal. The Wigner distribution is real-valued and for the filtered ambiguity function to also become real-valued, the kernel must fulfill

\[
\phi(\nu, \tau) = \phi^*(-\nu, -\tau). \quad (Exercise!)
\]

The weak time support of a quadratic distribution is satisfied if

\[
\int_{-\infty}^{\infty} \phi(\nu, \tau)e^{-i2\pi\nu t} d\nu = 0 \quad |\tau| \leq 2|t|,
\]

and the weak frequency support if

\[
\int_{-\infty}^{\infty} \phi(\nu, \tau)e^{-i2\pi\tau f} d\tau = 0 \quad |\nu| \leq 2|f|.
\]
The Wigner distribution has weak time- and frequency support but not strong time- and frequency support as cross-terms arise in between signal components of a multi-component signal. Strong time support means that whenever the signal is zero, then the distribution also is zero, giving a stronger restriction on the ambiguity kernel,

$$\int_{-\infty}^{\infty} \phi(\nu, \tau)e^{-i2\pi\nu t}d\nu = 0 \quad |\tau| \neq 2|t|, \quad (4.13)$$

and similarly strong frequency support implies that

$$\int_{-\infty}^{\infty} \phi(\nu, \tau)e^{-i2\pi\nu f}d\tau = 0 \quad |\nu| \neq 2|f|. \quad (4.14)$$

For proofs see [28].

### 4.3 RID and product kernels

Methods to reduce the cross-terms, also sometimes referred to as interference terms, have been proposed and a number of useful kernels, that falls into the socalled Reduced Interference distributions (RID) class can be found. The maybe most applied RID is the Choi-Williams distribution, [29], also called the Exponential distribution (ED), with the ambiguity kernel defined as

$$\phi_{ED}(\nu, \tau) = e^{-\nu^2\sigma^2}, \quad (4.15)$$

where $\sigma$ is a design parameter. The Choi-Williams distribution also falls into the subclass of product kernels which have the advantage in optimization of being dependent of one variable only, i.e., $x = \nu\tau$. The Choi-Williams distribution also satisfy the marginals as

$$\phi_{ED}(0, \tau) = \phi_{ED}(\nu, 0) = 1. \quad (4.16)$$

The kernel is also Hermitian, i.e.,

$$\phi_{ED}^*(\nu, \tau) = \phi_{ED}(-\nu, -\tau), \quad (4.17)$$

ensuring realness of the time-frequency distribution.

A comparison of the Choi-Williams distribution with the Wigner-Ville distribution for two example signals is seen in Figure 4.2 where the Choi-Williams distribution gives a
reduction of the cross-terms by smearing out them, but still keeps most of the concentration of the auto-terms, when the signal components are located at different frequencies as well as different times. However, it is important to note that this nice view are not fully repeated for the two complex sinusoids located at the same frequency, see Figure 4.3, where the cross-terms still are present but comparing the size of them will show that the cross-terms of the Wigner-Ville distribution is twice the height of the auto-terms but the Choi-Williams distribution cross-terms are about half the height of the auto-terms, a reduction of four times. Still, the remaining amplitude of the cross-terms could cause misinterpretation of the distributions. A similar performance would be seen for two components located at the same time.

Figure 4.2: The Wigner distribution and Choi-Williams distribution of two complex sinusoids with Gaussian envelopes located at $t_1 = 64$, $f_1 = 0.1$ and $t_2 = 128$, $f_2 = 0.2$.

Other ways of dealing with the cross-terms have been applied, e.g. the Born-Jordan distribution derived by Cohen, [21], using a rule defined by Born and Jordan. The properties of this kernel were not fully understood until in the 1990s, with the work of Jeong and Williams, [30]. The ambiguity kernel for the Born-Jordan distribution, then also called is the Sinc-distribution, is
which also fits into the RID class and is a product kernel.

4.4 Separable kernels

Another nice form of simple kernels are the separable kernels defined by

\[ \phi(\nu, \tau) = G_1(\nu)g_2(\tau). \]  

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
\[ W_z^Q(t,f) = g_1(t) * W_z(t,f) * G_2(f), \]  

(4.20)

as

\[ A_z^Q(\nu,\tau) = G_1(\nu)A_z(\nu,\tau)g_2(\tau). \]  

(4.21)

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

\[ G_1(\nu) = 1, \]  

(4.22)

a doppler-independent kernel is given as \( \phi(\nu,\tau) = g_2(\tau) \), and the quadratic time-frequency distribution reduces to

\[ W_z^Q(t,f) = W_z(t,f) * G_2(f), \]  

(4.23)

which is a smoothing in the frequency domain meaning that only the weak time support is satisfied. The doppler-independent kernel is also given the name Pseudo-Wigner or windowed Wigner distribution. The other case is when

\[ g_2(\tau) = 1, \]  

(4.24)

giving the lag-independent kernel, \( \phi(\nu,\tau) = G_1(\nu) \) where the weak frequency support is satisfied and the time-frequency formulation gives only a smoothing in the variable \( t \),

\[ W_z^Q(t,f) = g_1(t) * W_z(t,f). \]  

(4.25)

A comparison for some example signals show the advantage of the frequency smoothing for the Doppler-independent kernel as the cross-terms for signals with different time location disappears with use of averaging (smoothing) in the frequency direction, as the cross-terms oscillates in the frequency direction, Figure 4.4a) and b). For the lag-independent kernel we see the advantage of the smoothing in the time-direction in Figure 4.4c), where the cross-terms disappears in the frequency direction.
Figure 4.4: The resulting time-frequency plots for three complex sinusoids with Gaussian envelopes, located at $t_1 = 64$, $f_1 = 0.1$ and $t_2 = 128$, $f_2 = 0.1$ and $t_3 = 128$, $f_3 = 0.2$; a) The Wigner distribution; b) Doppler-independent kernel; c) Lag-independent kernel.
Chapter 5

Quadratic distributions and Cohen’s class

In the 1940s to 60s, during and after the invention of the spectrogram and the Wigner-Ville distribution, a lot of other distributions and methods were invented, e.g., Rihaczek distribution, [18], Page distribution, [19] and Margenau-Hill or Levin distribution, [20]. In 1966 a formulation was made by Leon Cohen in quantum mechanics, [21], which included these and an infinite number of other methods as kernel functions. Many of the distributions satisfied the marginals, the instantaneous frequency condition and other properties and all of these methods are nowadays referred to as the Cohen’s class. Later a quadratic class was defined, also including kernels that not satisfy the marginals, which was a restriction of the original Cohen’s class. However, the quadratic class is nowadays often referred to as the Cohen’s class.

5.1 Quadratic distributions

Quadratic time-frequency distributions can always be formulated as the multiplication of the ambiguity function and the ambiguity kernel,

\[ A^Q_z(\nu, \tau) = A_z(\nu, \tau) \cdot \phi(\nu, \tau), \tag{5.1} \]

and the corresponding smoothed Wigner-Ville distribution is found as the 2-dimensional convolution,

\[ W^Q_z(t, f) = W_z(t, f) * \Phi(t, f), \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_z(\nu, \tau) \phi(\nu, \tau) e^{-i2\pi(f\tau - \nu t)} d\tau d\nu. \tag{5.2} \]
Using Eq. (4.1) we find

\[ W_Q^z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(u + \frac{\tau}{2})z^*(u - \frac{\tau}{2})\phi(\nu, \tau)e^{i2\pi(\nu t - f \tau - \nu u)}du d\tau d\nu, \quad (5.3) \]

which is the most recognized form defining the quadratic class. The time-invariance and frequency-invariance properties are important and to find the restrictions of the ambiguity kernel we study the quadratic distribution for a time- and frequency-shifted signal \( w(t) = z(t - t_0)e^{i2\pi f_0 t} \) giving

\[ W_Q^w(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(u + \frac{\tau}{2} - t_0)e^{i2\pi f_0(u + \frac{\tau}{2})}z^*(u - \frac{\tau}{2} - t_0)e^{-i2\pi f_0(u - \frac{\tau}{2})} \ldots \]

\[ \ldots \phi(\nu, \tau)e^{i2\pi(\nu t - f \tau - \nu u)}du d\tau d\nu, \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(u_1 + \frac{\tau}{2})z^*(u_1 - \frac{\tau}{2})\phi(\nu, \tau)e^{i2\pi(f_0 \tau + \nu t - f \tau - \nu u_1 - \nu t_0)}du_1 d\tau d\nu, \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(u_1 + \frac{\tau}{2})z^*(u_1 - \frac{\tau}{2})\phi(\nu, \tau)e^{i2\pi(\nu(t - t_0) - (f - f_0)\tau - \nu u_1)}du_1 d\tau d\nu, \]

\[ = W_Q^z(t - t_0, f - f_0), \quad (5.4) \]

where we have assumed that the kernel \( \phi(\nu, \tau) \) is not a function of time nor of frequency. This 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. However, note that the kernel still can be signal dependent.

### 5.2 Kernel interpretation of the spectrogram

The spectrogram actually also belong to the quadratic class which can be seen from the following derivation:

\[
S_z(t, f) = |\int_{-\infty}^{\infty} h^*(t - t_1)z(t_1)e^{-i2\pi ft_1}dt_1|^2,
\]

\[
= (\int_{-\infty}^{\infty} h^*(t - t_1)z(t_1)e^{-i2\pi ft_1}dt_1)(\int_{-\infty}^{\infty} h(t - t_2)z^*(t_2)e^{i2\pi ft_2}dt_2).
\]
Replacing $t_1 = t' + \frac{\tau}{2}$ and $t_2 = t' - \frac{\tau}{2}$,

$$S_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t' + \frac{\tau}{2})z^*(t' - \frac{\tau}{2})h(t - t' - \frac{\tau}{2})h^*(t - t' + \frac{\tau}{2})e^{-i2\pi f \tau}d\tau dt', \quad (5.5)$$

where we can identify the time-lag distribution $r_z(t, \tau) = z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})$ and a time-lag kernel

$$\rho^h(t, \tau) = h(t + \frac{\tau}{2})h^*(t - \frac{\tau}{2}), \quad (5.6)$$

will give

$$S_z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r_z(t', \tau)\rho^h(t - t', \tau)^*e^{-i2\pi f \tau}d\tau dt = W_z^h(t, f). \quad (5.7)$$

(The proof of the time-lag formulation of the quadratic time-frequency formulation is calculated in exercise 5.1.)

Another formulation of the spectrogram can be given by starting with a kernel $\phi_h(\nu, \tau)$ which is the ambiguity function of a window $h(t)$, i.e.,

$$\phi_h(\nu, \tau) = \int_{-\infty}^{\infty} h(t + \frac{\tau}{2})h^*(t - \frac{\tau}{2})e^{-i2\pi \nu t}dt. \quad (5.8)$$

Then we drop the integral limits for simplification and find,

$$W_z^h(t, f) = \int \int \int z(t_1 + \frac{\tau}{2})z^*(t_1 - \frac{\tau}{2})h(t_2 + \frac{\tau}{2})h^*(t_2 - \frac{\tau}{2})e^{-i2\pi (\nu t_1 + \nu t_2 - \nu t + f \tau)}dt_1 dt_2 d\tau d\nu,$$

which after integration over $\nu$ becomes

$$= \int \int \int z(t_1 + \frac{\tau}{2})z^*(t_1 - \frac{\tau}{2})h(t_2 + \frac{\tau}{2})h^*(t_2 - \frac{\tau}{2})\delta(t - t_1 - t_2)e^{-i2\pi f \tau} d\tau dt_1 dt_2.$$

Using $t_2 = t - t_1$ gives,

$$W_z^h(t, f) = \int \int z(t_1 + \frac{\tau}{2})z^*(t_1 - \frac{\tau}{2})h(t - t_1 + \frac{\tau}{2})h^*(t - t_1 - \frac{\tau}{2})e^{-i2\pi f \tau} d\tau dt_1,$$

and with $t_1 + \frac{\tau}{2} = t'$ and $t_1 - \frac{\tau}{2} = t''$, 39
\[ = \int \int z(t')z^*(t'')h(t - t'')h^*(t - t')e^{-i2\pi f(t' - t'')} dt' dt'', \]

which equals the spectrogram,

\[ S_z(t, f) = (\int_{-\infty}^{\infty} h^*(t - t')z(t')e^{-i2\pi ft'} dt')(\int_{-\infty}^{\infty} h(t - t'')z^*(t'')e^{i2\pi ft''} dt''), \]

\[ = |\int_{-\infty}^{\infty} h^*(t - t')z(t')e^{-i2\pi ft'} dt'|^2. \]

If we study the assumed the ambiguity kernel of the spectrogram above, we see that the marginals never could be fulfilled. Why?

### 5.3 Multitaper time-frequency analysis

We will also be able to find a connection between the spectrogram formulation and the Wigner-Ville distribution/spectrum using the concept of multitapers. It has been shown that the calculation of the two-dimensional convolution between the kernel and the Wigner spectrum estimate of a process realization can be simplified using kernel decomposition and calculating a multitaper spectrogram, \[31, 32\]. The time-lag estimation kernel is rotated and the corresponding eigenvectors and eigenvalues are calculated. The resulting smoothed time-frequency distribution is given as the weighted sum of the spectrograms of the data with the different eigenvectors as sliding windows and the eigenvalues as weights, \[33\]. This can be seen if we use the quadratic class definition based on the time-lag kernel,

\[ W_Q z(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})\rho(t - u, \tau)e^{-i2\pi ft} dud\tau. \] (5.9)

Replacing \( u = (t_1 + t_2)/2 \) and \( \tau = t_1 - t_2 \), we find

\[ W_Q^* z(t, f) = \int \int z(t_1)z^*(t_2)\rho(t - \frac{t_1 + t_2}{2}, t_1 - t_2)e^{-i2\pi f(t_1 - t_2)} dt_1 dt_2 \]

\[ = \int \int z(t_1)z^*(t_2)\rho_{\text{rot}}(t_1, t_2)e^{-i2\pi f(t_1 - t_2)} dt_1 dt_2 \] (5.10)

where

\[ \rho_{\text{rot}}(t_1, t_2) = \rho\left(\frac{t_1 + t_2}{2}, t_1 - t_2\right). \] (5.11)
If the kernel $\rho^{rot}(t_1, t_2)$ satisfies the Hermitian property

$$\rho^{rot}(t_1, t_2) = (\rho^{rot}(t_2, t_1))^*, \quad (5.12)$$

then solving the integral

$$\int \rho^{rot}(t_1, t_2)u(t_1)dt_1 = \lambda u(t_2), \quad (5.13)$$

results in eigenvalues $\lambda_k$ and eigenfunctions $u_k$ which form a complete set. The kernel can be expressed as

$$\rho^{rot}(t_1, t_2) = \sum_{k=1}^{\infty} \lambda_k u_k^*(t_1)u_k(t_2). \quad (5.14)$$

Using the eigenvalues and eigenvectors, Eq. (5.10) is rewritten as a weighted sum of spectrograms,

$$W_z^Q(t, f) = \sum_{k=1}^{\infty} \lambda_k \int \int z(t_1)z^*(t_2)e^{-i2\pi ft_1}e^{i2\pi ft_2}u_k^*(t-t_1)u_k(t-t_2)dt_1dt_2 \quad (5.15)$$

$$= \sum_{k=1}^{\infty} \lambda_k \int |z(t_1)e^{-i2\pi ft_1}u_k^*(t-t_1)dt_1|^2. \quad (5.16)$$

Depending on the different $\lambda_k$, the number of spectrograms that are averaged could be just a few or an infinite number. With just a few $\lambda_k$ that differs zero the multitaper spectrogram,

$$S_z(t, f) = \sum_{k=1}^{K} \lambda_k \int z(t_1)e^{-i2\pi ft_1}u_k^*(t-t_1)dt_1|^2. \quad (5.17)$$

is an efficient solution from implementation as well as optimization aspects. Approaches to approximate a kernel with a few multitaper spectrograms have been taken, e.g., decomposing the RID kernel, [34, 35], least square optimization of the weights, [36] and an approach using a penalty function, [37]. In [38], time-frequency kernels and connected multitapers for statistically stable frequency marginals are developed.

The Slepian functions (Thomson multitapers) are recognized to give orthonormal spectra for stationary white noise and also to be the most localized tapers in the frequency domain. In time-frequency analysis the Hermite functions have been shown to give the
best time-frequency localization and orthonormality in the time-frequency domain, [39]. Recently, many methods have been proposed, where the orthonormal multitapers are the Hermite functions, e.g., [36, 38, 40, 41, 42]. However, similarly as for the Slepian functions, for spectra with peaks, the cross-correlation between sub-spectra give degraded performance, [43].

5.4 Some historical important distributions

The Rihaczek distribution (RD), [18], is derived from the energy of a complex-valued deterministic signal over finite ranges of \( t \) and \( f \), and if these ranges become infinitesimal, a energy density function is obtained. The energy of a complex-valued signal is

\[
E = \int_{-\infty}^{\infty} |z(t)|^2 dt = \int_{-\infty}^{\infty} z(t)z^*(t)dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z(t)Z^*(f)e^{-i2\pi ft}df dt, \quad (5.18)
\]

where then for infinitesimal \( t \) and \( f \) we define

\[
R_z(t, f) = z(t)Z^*(f)e^{-i2\pi ft}, \quad (5.19)
\]

as the Rihaczek distribution. (Sometimes this distribution is also called the Kirkwood-Rihaczek distribution as it actually was derived earlier in the context of quantum mechanics, [44].) We note that the marginals are satisfied as \( \int_{-\infty}^{\infty} R_z(t, f)df = |z(t)|^2 \), (seen from Eq. (5.18)), and

\[
\int_{-\infty}^{\infty} R_z(t, f)dt = Z^*(f) \int_{-\infty}^{\infty} z(t)e^{-i2\pi ft}dt = |Z(f)|^2. \quad (5.20)
\]

Other important properties of the Rihaczek distribution are both strong time support and strong frequency support, which 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.

The Fourier transform in two variables of Eq. (5.19) gives the ambiguity function from where the ambiguity kernel

\[
\phi(\nu, \tau) = e^{-i\pi \nu \tau}, \quad (5.21)
\]

can be identified. From this we also verify that the marginals are satisfied as \( \phi(\nu, 0) = \phi(0, \tau) = 1 \).

Comparing the results of the Wigner-Ville and Rihaczek distributions for some signals show that for a sum of two complex-valued sinusoid signal, (\( f_1 = 0.05 \) and \( f_2 = 0.15 \)),
multiplied with a Gaussian functions and located at the same time-instant, Figure 5.1a), the Rihaczek distribution Figure 5.1c) (real value) and d) (imaginary value) does not generate any cross-terms as the Wigner-Ville distribution does, Figure 5.1b). The view of the components might be difficult to interpret as they seem to oscillate quite strange. However, in Figure 5.2, for the sum of two complex-valued sinusoids, \( f_1 = 0.05 \) and \( f_2 = 0.15 \), with Gaussian envelope located at different time intervals and with different frequencies, the Rihaczek distribution produces twice as many cross-terms compared to the Wigner-Ville distribution. Why the cross-terms appear as in Figure 5.2 is intuitively easy to understand if we study Eq. (5.19) where the signal \( z(t) \) will be present around time \(-30 \) and \( 30 \) and the Fourier transform \( Z(f) \) will be present around frequencies \( f_1 = 0.05 \) and \( f_2 = 0.15 \). These two functions are combined to the two-dimensional time-frequency representation, which naturally then will have components appering at all combinations of these time and frequency instants.

The real part of the Rihaczek distribution is also defined as the Levin distribution (LD), [45], i.e.,
Figure 5.2: The Wigner-Ville and Rihaczek distributions of two complex sinusoids with Gaussian envelope located at different time intervals and with different frequencies.
\[ L_z(t, f) = \Re[z(t)Z^*(f)e^{-i2\pi ft}]. \] (5.22)

It also satisfies the marginals and has strong time and frequency support. The advantage of the Levin distribution is that it is real-valued (similar to the Wigner-Ville distribution), even for complex-valued signals. It was also derived in quantum-mechanical context and is therefore sometimes called Margenau-Hill distribution, [20].

The Rihaczek and Levin distributions have their obvious drawbacks, but as they also are intuitively nice from interpretation aspects, they are often applied where the signal is windowed, e.g., a windowed Rihaczek distribution is defined as
\[ R^w_z(t, f) = z(t) \left[ F_{\tau \to f} \{ z(\tau)w(\tau - t) \} \right]^* e^{-i2\pi ft}. \] (5.23)

With an appropriate size of the window, the cross-terms of the Rihaczek distribution can be reduced.

Another suggestion of distribution is the running transform, \( Z_-(t, f) \), or the Page distribution, [19], which is the Fourier transform of the signal up to time \( t \).
\[ Z_-(f) = \int_{-\infty}^{t} z(t')e^{-i2\pi ft'} dt'. \] (5.24)

If the signal is zero after time \( t \), the marginal in frequency is \( |Z_-(f)|^2 \). If a distribution satisfy the marginal, then
\[ \int_{-\infty}^{t} P_-(t', f) dt' = |Z_-(f)|^2. \] (5.25)

Differentiating with respect to time gives
\[
P_-(t, f) = \frac{\partial}{\partial t} \left[ \int_{-\infty}^{t} z(t')e^{-i2\pi ft'} dt' \right]^2
= \frac{\partial}{\partial t} \left[ \left( \int_{-\infty}^{t} z(t')e^{-i2\pi ft'} dt' \right) \left( \int_{-\infty}^{t} z^*(t')e^{i2\pi ft'} dt' \right) \right]
= z(t)e^{-i2\pi ft}Z_-(f) + Z_-(f)z^*(t)e^{i2\pi ft}
= 2\Re \left[ z^*(t)Z_-(f)e^{i2\pi ft} \right],
\]
to be recognized to be somewhat similar to the Levin distribution. Comparing the results of the previous examples, we see the characteristics of the Page distribution in Figure 5.3, where one of the cross-terms from the Rihaczek distribution disappears but the other
remains. Can you see the reason why the first appearing cross-term in time is the one that disappears? Study the definition in Eq. (5.24) and the limits of the integral.

5.5 The four domains of time-frequency analysis

As noted we can also formulate the quadratic class as

\[
W_z^Q(t, f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r_z(u, \tau)\rho(t - u, \tau)e^{-i2\pi f\tau} dud\tau, \tag{5.26}
\]

where the **time-lag kernel** is defined as

\[
\rho(t, \tau) = \int_{-\infty}^{\infty} \phi(\nu, \tau)e^{i2\pi \nu t} d\nu, \tag{5.27}
\]

and the IAF as
\[ r_z(t, \tau) = z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2}). \]  

(5.28)

The **Doppler-frequency distribution** is found from the Fourier transform of the Wigner-Ville distribution in \( t \) or the Fourier transform of the ambiguity function in \( \tau \), i.e.,

\[
D_z(\nu, f) = \int_{-\infty}^{\infty} A_z(\nu, \tau)e^{-i2\pi f \tau} d\tau
\]

\[
= \int_{-\infty}^{\infty} W_z(t, f)e^{-i2\nu t} dt.
\]

(5.29)

The doppler-frequency distribution is the corresponding **spectral autocorrelation function** which is also calculated from the IAF as

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

(5.30)

Replacing \( t + \tau/2 = t_1 \) and \( t - \tau/2 = t_2 \) we reformulate,

\[
D_z(\nu, f) = \int_{-\infty}^{\infty} z(t_1)e^{-i2\pi(f+\frac{\nu}{2})t_1} dt_1 \cdot \int_{-\infty}^{\infty} z^*(t_2)e^{i2\pi(f-\frac{\nu}{2})t_2} dt_2
\]

\[
= Z(f + \frac{\nu}{2})Z^*(f - \frac{\nu}{2}),
\]

(5.31)

where \( Z(f) \) is the Fourier transform of \( z(t) \). We now see that the doppler-frequency distribution is the frequency dual of the IAF.

The four different domains for representation of a time-varying signal are presented, the instantaneous autocorrelation function (IAF) in the variables \( (t, \tau) \), the Wigner distribution in \( (t, f) \), the ambiguity function in \( (\nu, \tau) \) and finally the doppler-frequency distribution in \( (\nu, f) \). A schematic overview is given in Figure 5.4. Studying simple signals in the different domains give us information of the interpretation. A Gaussian envelope signal centered at \( t = 0 \) will show up as a Gaussian signal in all other domains too as the Fourier transform also has a Gaussian shape. It is important to remember which variable that is used in the Fourier transform and which variable that is fixed when transforming from one domain to another. When we move up, down, left or right between the figures, only one variable change for each step.

If we delay the Gaussian signal in time to \( t = 20 \), the IAF and Wigner distribution will shift the information by the movement of the Gaussian to \( t = 20 \), Figure 5.5. For the ambiguity function and doppler-frequency distribution the result is complex valued.
Figure 5.4: The four possible domains in time-frequency analysis.
and we choose to plot the real value of the functions. The Fourier transform from the 
$t$-variable to the $\nu$-variable will give us a Gaussian envelope with a phase corresponding 
to the time-shift of $t = 20$. The phase-shift shows up in the real value (and the imaginary 
value) and is seen in the plot as a function shifting between negative and positive values 
if studied in the $\nu$-variable. Increasing the time-shift will cause these these shifts to occur 
more frequently, i.e., a more rapid oscillation.

Figure 5.5: The Gaussian signal shifted to $t = 20$ and the representation in all four 
domains.

If we now replace the Gaussian signal with a complex-valued signal of frequency $f = 
0.2$ with a Gaussian envelope, movement in frequency show up in the Wigner and the 
doppler-frequency distribution as the functions now are located at $f = 0.2$. For the IAF 
and the ambiguity function however, we now get a function shifting from positive to 
negative values in the $\tau$-variable, Figure 5.6.

The combination of a shift in time of $t = 20$ with a shift in frequency $f = 0.2$, can 
clearly be seen in the Wigner distribution of Figure 5.7. For the ambiguity function, the
combination of time- and frequency shift show up as diagonal lines but the function is still located at zero both in $\tau$ and $\nu$. The IAF give time shift and in the doppler-frequency distribution the frequency shift of $f = 0.2$ shows up.

With a multicomponent signal, a combination of a Gaussian signal a time $t = 0$ and the complex signal with frequency $f = 0.2$, Gaussian envelope and located at $t = 20$, a more complex pattern can be seen in all domains of Figure 5.8. For the Wigner distribution, the two functions are seen at $t = 0, f = 0$ and $t = 20, f = 0.2$ with the cross-term in between. The cross-term also show up in the other domains at various places. In the IAF domain the auto-terms are located at $\tau = 0$ as before where the cross-terms show up at $\tau = -20$ and $\tau = 20$. The same thing is seen in the for doppler-frequency distribution where the auto-terms are located at $\nu = 0$ and the cross-terms at positive and negative $\nu$-values respectively. For the ambiguity function, these properties can be used to differ auto-terms from cross-terms as the auto-terms are located at $\tau = \nu = 0$.
Moving the complex sinusoids more apart in time will cause a corresponding shift in the IAF domain as well as the ambiguity function. For the doppler-frequency distribution however, this larger shift just shows up as the frequency of the shifts of the functions, the location of the auto-terms as well as the cross-terms remain the same. On the other hand, if the we increase the frequency shift but keep the time shift constant, the different terms of the IAF will remain at the same location but the doppler distribution location will change similarly to the ambiguity function. The auto-terms will still be located at the origin and the cross-terms somewhere outside the origin of the ambiguity function.

In the last example in Figure 5.9 we show the four domains in the case of a real cosine signal of frequency $f = 0.2$ with Gaussian envelope centered at $t = 0$. The auto-terms as well as the cross-terms show up as expected in the different domains. In the Wigner domain the auto-terms are located at $f = \pm 0.2$ and the cross-term at $f = 0$. In the
ambiguity function, the auto-term is at $\nu = \tau = 0$ and the cross-terms outside. For the IAF everything is located at $t = \tau = 0$ as $t = 0$ as all term are located at $t = 0$ and hence no time-difference between them exist and thereby the terms are located at $\tau = 0$ as well. The doppler-frequency distribution has the difference in frequency $f = \pm 0.2$, which shows up both in $f$ and in $\nu$. Note that the difference is doubled in the $\nu$-domain.

In [46, 47] features extracted from the time-frequency, ambiguity and doppler-frequency domains are used to classify the song syllables of male great reed warblers. An example of different representations are seen in Figure 5.10. Studying the time-frequency domain, a syllable and the same syllable modulated in frequency will not give comparable TF-representations, i.e., a classification based on this will indicate a difference. Similarly, time differences will show up and indicate a difference. In the ambiguity domain, differences in modulation of time and frequency will however give a similar appearance. Therefore, in different applications, different domains are beneficial.
Figure 5.9: A real-valued cosine signal with $f = 0.2$ and Gaussian envelope and the representation in all four domains.
Figure 5.10: Example of a syllable in a) the time domain, b) the time-frequency domain, c) the ambiguity domain, d) the doppler-frequency domain.
Chapter 6

Optimal time-frequency concentration

Many of the proposed methods in time-frequency analysis are devoted to suppression of cross-terms and to maintain the concentration of the autoterms. To measure concentration and possibly automatically determine parameters for optimal methods, measurement criteria for the concentration are needed. Such criteria were defined already in [16] and data-adaptive methods could be found e.g. in [25].

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

$$p_1 + p_2 + \ldots + p_N = 1.$$  \hfill (6.1)

A quadratic test function

$$\gamma = p_1^2 + p_2^2 + \ldots + p_N^2,$$  \hfill (6.2)

will attend its minimum value when $p_1 = p_2 = \ldots = p_N = 1/N$ (maximum spread) and the maximum value when only one $p_i = 1$ and all the others are zero (minimum spread).

Incorporating the unity sum constraint in the test function gives

$$\gamma = \frac{p_1^2 + p_2^2 + \ldots + p_N^2}{(p_1 + p_2 + \ldots + p_N)^2},$$  \hfill (6.3)

We can use this idea as a measure for the time-frequency concentration, e.g., as suggested in [25],

$$\gamma = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^4(t, f) dt df}{(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^2(t, f) dt df)^2},$$  \hfill (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 to use a local function, also proposed in [25],

\[
\gamma_{(t,f)} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q^2(t_1 - t, f_1 - f)W^4(t_1, f_1)dt_1df_1}{(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(t_1 - t, f_1 - f)W^2(t_1, f_1)dt_1df_1)^2},
\]

(6.5)

where \( Q(t, f) \) is a weighting function that determines the region for the concentration measure. One suggestion is to use a Gaussian function as weighting function, [25]. Another more recent measure is based on the Rényi entropy, [48],

\[
R_{\alpha} = \frac{1}{1-\alpha} \log_2(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^\alpha(t, f)dtdf), \quad \alpha > 2.
\]

(6.6)

The behavior depends on \( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^\alpha(t, f)dtdf \) for \( \alpha > 2 \) similarly to the above mentioned functions and the logarithm is a monotone function. However, as \( 1/(1 - \alpha) \) is negative for \( \alpha > 2 \), the entropy of Eq. (6.6) becomes larger for less concentrated distributions. The Rényi entropy could be interpreted as counting the number of components in a multi-component signal. For odd values of \( \alpha > 1 \), \( R_{\alpha} \) is asymptotically invariant to time-frequency cross-terms and does not count them. The \( R_{\alpha} \) attends the lowest value for the Wigner distribution of a single Gaussian windowed signal. It is also invariant to time- and frequency shifts of the signal.

The Rényi entropy can be related to the well-known Shannon entropy

\[
H = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(t, f) \log_2 W(t, f)dtdf,
\]

(6.7)

which could be recovered from the Rényi entropy, from the limit of \( \alpha \to 1 \), [48]. We should note that the Shannon entropy could not be used for a distribution including any negative values at any time-frequency value as the logarithm is applied inside the integral. For the Rényi entropy however, the logarithm is applied outside the integral, and many signals and especially smoothed distributions are positive in this context. However, examples of multi-component signals and \( \alpha \) can be found where \( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W^\alpha(t, f)dtdf < 0 \), see [48]. Often, the value of \( \alpha = 3 \) is applied. Many of the time-frequency distributions in the literature have been based either implicitly or explicitly on information measures, e.g., [49, 50].
6.1 Uncertainty principle

A signal cannot be both time-limited and frequency-limited at the same time. This is seen, e.g., in the example where a multiplication of any signal \( x(t) \) with a rectangle function certainly will time-limit the signal,

\[
x_T(t) = x(t) \cdot \text{rect}[t/T].
\]  

(6.8)

In the frequency domain, the result will be of infinite bandwidth as the multiplication is transferred to a convolution with an infinite sinc-function, \((\sin x/x)\).

\[
X_T(f) = X(f) \ast T\text{sinc}(fT).
\]  

(6.9)

Similarly, a bandwidth limitation by a rectangle function in the frequency domain will cause a convolution in the time domain with an infinite sinc-function.

We can measure the frequency bandwidth using the so called effective bandwidth \( B_e \),

\[
B_e^2 = \frac{\int_{-\infty}^{\infty} f^2 |X(f)|^2 df}{\int_{-\infty}^{\infty} |X(f)|^2 df},
\]  

(6.10)

where the denominator is the total energy of the signal and analogously the so called effective duration \( T_e \) is defined as

\[
T_e^2 = \frac{\int_{-\infty}^{\infty} t^2 |x(t)|^2 dt}{\int_{-\infty}^{\infty} |x(t)|^2 dt}.
\]  

(6.11)

In these definitions the signal as well as the spectrum are assumed to be located symmetrically around \( t = 0 \) and \( f = 0 \). These expressions can be compared with the variance of a stochastic variable, i.e., if \(|X(f)|^2\) is interpreted as the probability density function of the stochastic variable \( f \) the denominator would be 1 and the numerator would be the variance of \( f \). Using these measures we can define the bandwidth-duration product, \( B_e \cdot T_e \), which for a Gaussian signal is

\[
B_e \cdot T_e = \frac{1}{4\pi},
\]  

(6.12)

[16], and for all other signals

\[
B_e \cdot T_e > \frac{1}{4\pi}.
\]  

(6.13)
Another name for the bandwidth-duration product is the uncertainty principle which serves as a measure of the information richness in the signal. Gabor chose the Gaussian signal as his elementary function in the Gabor expansion, [16], as this function is the most concentrated both in time and frequency.

### 6.2 Gabor expansion

For a continuous signal \( x(t) \), the Gabor expansion is defined as

\[
x(t) = \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_{m,k} g(t - mT_0) e^{i2\pi kF_0 t},
\]

(6.14)

where \( T_0 \) and \( F_0 \) denote the time and frequency sampling steps. The coefficients \( a_{m,k} \) are called the Gabor coefficients. Gabor chose the Gaussian function

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

(6.15)

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 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

\[
X(mT_0, kF_0) = \int_{-\infty}^{\infty} x(t_1) w^*(t_1 - mT_0) e^{-i2\pi kF_0 t_1} dt_1,
\]

(6.16)

where \( X(mT_0, kF_0) = a_{m,k} \), if the sampling distances \( T_0 \) and \( F_0 \) satisfy the critical sampling relation \( F_0 T_0 = 1 \). In this case there is an uniquely defined analysis window \( w(t) \) given from the synthesis window \( g(t) \). This is seen if Eq. (6.16) is substituted into Eq. (6.14) as

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

(6.17)

which is true if

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

(6.18)
Using the Poisson-sum formula and the biorthogonality, see [17], the double summation is reduced to a single integration,

$$\int_{-\infty}^{\infty} g(t) w^* (t - mT_0) e^{i2\pi kF_0 t} = \delta_k \delta_m,$$

(6.19)

where $\delta_k$ is the Kronecker delta and the analysis window $w(t)$ follows uniquely from a given synthesis window. However, this window does not have attractive properties from a time-frequency analysis perspective, as it is not well located in time and frequency. If instead an oversampled lattice is used, i.e., $F_0 T_0 < 1$, the analysis window is no longer unique, and a choice can be made from a time-frequency analysis perspective.

For discrete-time signals, calculation of the analysis window is made using the Zak transform, [51] and a mean square error solution gives a analysis window which is most similar to the synthesis window, [52], see Figure 6.1. The Gaussian synthesis window in Figure 6.1a) generates at critical sampling the analysis window of Figure 6.1b) which is not attractive for generating a time-frequency distribution, but might be the best for other applications, e.g., image analysis. For higher oversampling Figure 6.1c) and d), the analysis window becomes more similar to the synthesis window.

The Gabor expansion is often used in applications where analysis as well as reconstruction of the signal is needed but also for estimation of e.g., linear time-variant (LTV) filters, [53].

### 6.3 Instantaneous frequency

A well-known sub-area in time-frequency analysis is estimation of **instantaneous frequency**. The instantaneous frequency is the derivative of the phase, and is closely related to the definition of the analytic signal as we from the complex-valued analytic signal can get a phase signal. Here, one of several approaches for the definition of the instantaneous frequency are presented, where we start by a simple amplitude-modulated signal,

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

(6.20)

where $A(t)$ is the **instantaneous amplitude**. The frequency of this signal for all $t$ is found as

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

(6.21)

We can now extend this constant-frequency example to a signal
Figure 6.1: 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.
where \( f(t) \) is time-variable and where \( \phi(t) \) is defined as the **instantaneous phase**. If \( \phi(t) \) is differentiable and is evaluated at \( t = t_1 \) and \( t = t_2 \), where \( t_2 > t_1 \), there exists a time interval between \( t_1 \) and \( t_2 \) such that

\[
\phi(t_2) - \phi(t_1) = (t_2 - t_1)\phi'(t).
\]  

(6.23)

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, i.e., \( t_2 = t_1 + p_I \), will result in that \( \phi(t_2) = \phi(t_1) + 2\pi \) and \( \phi'(t) \) from Eq. (6.23) will be simplified to

\[
\phi'(t) = 2\pi/p_I = 2\pi f_I.
\]  

(6.24)

The instantaneous frequency will then, similarly to the constant-frequency example, be

\[
f_I = \frac{\phi'(t)}{2\pi},
\]  

(6.25)

for the time \( t \) during that oscillation.

In the case of a real-valued signal, however, e.g.,

\[
x(t) = A(t)\cos(2\pi f(t)t),
\]  

(6.26)

the instantaneous frequency will be zero as this signal has no phase in the sense defined above. This will of course be the case for any real-valued signal but also for multi-component complex-valued signals, the definition of instantaneous frequency becomes difficult. The concept of instantaneous frequency is no longer valid and the methods that rely on the definition of instantaneous frequency do not work properly. One example is the sum of two complex sinusoids,

\[
x(t) = A_1e^{i2\pi f_1t} + A_2e^{i2\pi f_2t},
\]  

(6.27)

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 6.2a), 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 6.2b), 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.

![Graphs](image)

Figure 6.2: 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 $A_2 = 1$.

### 6.4 Reassigned spectrogram

The **reassignment principle** was introduced in 1976 by Kodera, de Villedary and Gendrin, [54], but were not applied to a larger extent until when Auger and Flandrin reintroduced the method, [55]. The idea of reassignment is to keep the localization of a single component by reassigning mass to the center of gravity where the cross-terms are reduced by smoothing of the Wigner distribution. 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 chirps and constant frequency signals.

Recently, there has been a renewed interest in the reassignment technique and the related synchrosqueezing, [56, 57, 58, 59]. Especially, there is an interest to discriminate between short pulses (transients) and stationary tones. The Levenberg-Marquardt (LM) reassignment, proposed in [58], 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.
For a Gaussian function (logon), the usual reassignment does not give perfect localization,
but with the LM technique, this is possible, i.e., to relocate all mass to the center of the
logon. Another approach is to use a simple scaling of the usual reassignment, [60]. The
method has advantages in the case where the length of the Gaussian function is known
or can be estimated. An example is presented where it is shown that the usual time-
frequency reassignment might hide the actual model of data in cases where the model
consists of several Gaussian components.

The short-time Fourier transform of the signal $x(t)$ using the window $h(t)$ at time $t$
and frequency $\omega = 2\pi f$ is

$$F^h_x(t, \omega) = \int x(s)h^*(s - t)e^{-i\omega s}ds,$$

where integrals are from $-\infty$ to $\infty$ when not specified and the corresponding spectrogram
is found as

$$S^h_x(t, \omega) = |F^h_x(t, \omega)|^2.$$  \hspace{1cm} (6.29)

The reassigned spectrogram, where the spectrogram values are relocated to the corre-
sponding $\hat{t}_x$ and $\hat{\omega}_x$, is defined as

$$RS^h_x(t, \omega) = \int \int S^h_x(s, \xi)\delta(t - \hat{t}_x(s, \xi), \omega - \hat{\omega}_x(s, \xi))dsd\xi,$$

where $s$ and $\xi$ are the integration variables in time and frequency, respectively, and where
$\delta(t, \omega)$ is the two-dimensional Dirac impulse defined as

$$\int \int f(t, \omega)\delta(t - t_0, \omega - \omega_0)dtd\omega = f(t_0, \omega_0).$$  \hspace{1cm} (6.31)

The reassignment can be computed as

$$\hat{t}_x(t, \omega) = t + \Re\left(\frac{F^h_x(t, \omega)}{F^h_x(t, \omega)}\right),$$ \hspace{1cm} (6.32)

$$\hat{\omega}_x(t, \omega) = \omega - \Im\left(\frac{F^h_x/dt(t, \omega)}{F^h_x(t, \omega)}\right),$$ \hspace{1cm} (6.33)

where $\Re$ and $\Im$ represents real and imaginary parts respectively, [56]. The spectrogram can
be replaced by any other distribution belonging to the Cohen class, and the resulting reassig-
nement process will lead to perfectly localized single component chirp signals, frequency
tones and impulse signals, [55]. An example is given in Figure 6.3. For multi-component
signals, the reassignment improves the readability as the cross-terms are reduced by the
smoothing of the specific distribution and the reassignment then squeezes the signals
terms.

\begin{align*}
\hat{t}_x(t, \omega) &= t + c_t \Re \left( \frac{F_x(t, \omega)}{F_x^b(t, \omega)} \right), \\
\hat{\omega}_x(t, \omega) &= \omega - c_\omega \Im \left( \frac{F_x(t, \omega)}{F_x^b(t, \omega)} \right).
\end{align*}

6.5 Scaled reassigned spectrogram

Introducing the scaling factors $c_t$ and $c_\omega$, the reassignment can be computed as

\begin{align*}
\hat{t}_x(t, \omega) &= t + c_t \Re \left( \frac{F_x(t, \omega)}{F_x^b(t, \omega)} \right), \\
\hat{\omega}_x(t, \omega) &= \omega - c_\omega \Im \left( \frac{F_x(t, \omega)}{F_x^b(t, \omega)} \right).
\end{align*}
A Gaussian windowed constant frequency signal

\[ x_0(t) = g(t - t_0)e^{i\omega t}, \quad (6.36) \]

where

\[ g(t) = e^{-\frac{t^2}{2\sigma^2}}, \quad -\infty < t < \infty, \quad (6.37) \]

with scaling parameter \( \sigma \) is often used to model a short non-stationary signal. The quadratic class of distributions obey time-frequency and shift-invariance, implying that further analysis can be restricted to \( x(t) = g(t) \), [33]. In most cases the parameter \( \sigma \) is unknown and needs to be estimated. Using a unit energy Gaussian window with scaling parameter \( \lambda \), \( h(t) = 1/(\pi^{1/4}\sqrt{\lambda})e^{-\frac{t^2}{2\lambda^2}} \), will lead to the following reassignment, [60],

\[
\begin{align*}
\hat{t}_x(t, \omega) &= t - c_t \left( \frac{\lambda^2}{\lambda^2 + \sigma^2} t \right), \\
\hat{\omega}_x(t, \omega) &= \omega - c_\omega \left( \frac{\sigma^2}{\lambda^2 + \sigma^2} \omega \right). \quad (6.38, 6.39)
\end{align*}
\]

For the so called matched window case, i.e., \( \lambda = \sigma \), Eqs. (6.38,6.39) simplify to

\[
\begin{align*}
\hat{t}_x(t, \omega) &= t - c_t \frac{t}{2}, \\
\hat{\omega}_x(t, \omega) &= \omega - c_\omega \frac{\omega}{2}. \quad (6.40, 6.41)
\end{align*}
\]

For perfect localization of the reassigned Gaussian function time-frequency spectrum, centered at time and frequency zero, \( \hat{t} \) and \( \hat{\omega} \) should accordingly be zero for all values of \( t \) and \( \omega \). If we now assume \( c_t = c_\omega = 2 \) in Eqs. (6.40,6.41), the perfect localization with

\[ ScRS^h_x(t, \omega) = \delta(t, \omega), \]

is found for the matched window case. For the usual reassignment without scaling, Eqs. (6.32,6.33),

\[
\begin{align*}
\hat{t}_x(t, \omega) &= \frac{\sigma^2}{\lambda^2 + \sigma^2} t = \frac{t}{2}, \\
\hat{\omega}_x(t, \omega) &= \frac{\lambda^2}{\lambda^2 + \sigma^2} \omega = \frac{\omega}{2}.
\end{align*}
\]
with the last steps given for $\lambda = \sigma$. The corresponding reassigned spectrogram is found as, \[56\],

$$RS_x(t, \omega) = \frac{2\sqrt{\pi}(\lambda^2 + \sigma^2)}{\lambda} e^{-\left(\frac{\lambda^2 \sigma^2 + \sigma^2}{\lambda^2 + \sigma^2}\right) t^2 + \left(\frac{\omega^2 \sigma^2 + \sigma^2}{\lambda^2 + \sigma^2}\right) \omega^2} = 4\sqrt{\pi} \sigma e^{-2\left(\frac{t^2}{\sigma^2} + \sigma^2 \omega^2\right)}, \quad (6.42)$$

for $\lambda = \sigma$.

The performance is illustrated with a two-component case $x(t) = x_1(t) + x_2(t)$ where $x_n(t)$ is defined as in Eq. (6.36) with the time- and frequency translations $(t_0, \omega_0)$ replaced with $(t_n, \omega_n)$ for $n = 1, 2$. In Fig. 6.4, the spectrogram, the reassignment and the scaled reassignment are presented for the matched window case $\lambda = \sigma = 20$, and $t_1 = 6, \omega_1 = 15$ and $t_2 = 9, \omega_2 = 17$.

![Figure 6.4: 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 $\lambda = \sigma = 20$.](image)

The spectrogram of Fig. 6.4a) indicates that there are two components but the reassignment presented in Fig. 6.4b) shows a representation that could be interpreted as
some other signal. For the scaled reassignment, the final representation of Fig. 6.4c) clearly shows the two components localized at the correct time-frequency positions.

In Fig. 6.5 the performances of the relocations are illustrated where Fig. 6.5a) shows the usual reassignment $R$ for $\lambda = \sigma_n = 20$ where the arrows start at $(t, \omega)$ and end at $(\hat{t}_x(t, \omega), \hat{\omega}_x(t, \omega))$. An underlying contour plot of the corresponding spectrogram is also shown. The reassignment is only performed for power values of the spectrogram larger than 1% of the maximum power of the spectrogram and thereby the reassignments of infinite size obtained for singular points of the spectrogram, are avoided. The performance is well in concordance with Fig. 6.4b), where most of the transferred mass is located in the shape of two combined irregular circles. In Fig. 6.5b), the performance of the scaled reassignment is illustrated where the transfers will be twice as large as for the usual reassignment and the total effect will be that most of the mass will be located at the centers of the two components $(t_1 = 6, \omega_1 = 15)$ and $(t_2 = 9, \omega_2 = 17)$ in concordance with Fig. 6.4c).

Figure 6.5: Illustration of the reassignment where the arrows start at $(t, \omega)$ and end at $(\hat{t}_x(t, \omega), \hat{\omega}_x(t, \omega))$ for $\lambda = \sigma = 20$; a) reassigned spectrogram, b) scaled reassigned spectrogram. The spectrogram is shown as a contour plot.

Another usual signal in simulation and modeling, is the (infinite) chirp-signal. The ReSpect as well as other methods are known to perform very well for these type of signals.
and they are often used as test-signals in performance analysis. We show an example in Fig. 6.6a), of a signal that seems to be a Gaussian windowed chirp-signal of linearly increasing frequency but that actually is a sum of four equal length Gaussian functions ($\sigma_n = 20$, $y(t) = \sum_{n=1}^{4} a_n x_n(t)$, defined by Eq. (6.36). The time- and frequency-shifts ($t_n$, $\omega_n$) are (3, 10), (5, 13), (7, 15), (9, 18) and the amplitudes $a_n$ are $-0.09 - i0.42$, $0.89 + i0.09$, $-0.14 + i0.90$, $0.57 + i0.50$ respectively. Circularly white Gaussian noise with standard variation 0.28 is added, (SNR=12 dB). The optimal window length is estimated (using Rényi entropy over the total area) to be $\lambda = 19$ for the ScRe-Spect and this window length is used for all methods. The corresponding spectrogram is presented in Fig. 6.6b), the result of ScRe-Spect is shown in Fig. 6.6c) and the Re-Spect in Fig. 6.6d). The colorscales are in dB, normalized by the maximum value and the FFT-length is 2048 where the number of time- and frequency samples in the figures are (261 X 294). For the ScReSpect, the masses are reallocated to the centers showing the appearance of four Gaussian functions. For the Re-Spect however, the underlying signal appears to be a linear chirp.
Figure 6.6: 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) scaled reassigned spectrogram, d) reassigned spectrogram. (In all cases an estimated value of $\lambda = 19$ is used as window length.)
Chapter 7

Stochastic time-frequency analysis

In the recent 10-15 years, the attempts to find optimal estimation algorithms for non-stationary processes, have increased, [61]. The term Wigner spectrum has been established where the ambiguity function is replaced by the ambiguity spectrum. The fundamental problem in stochastic time-frequency analysis is the estimation of a reliable time-frequency spectrum from a single realization of the process. Many kernels, optimized for different purposes, can be found, e.g., [62, 63, 38, 64]. A well-defined framework has also been proposed for the class of so called underspread processes, for which the ambiguity spectrum of the process is concentrated around the origin. For these processes it has been shown, at least approximately, that many rules for time-invariant linear systems and stationary processes applies, [65].

Time-frequency analysis of non-stationary and time-varying signals, in general, has been a field of research for a number of years. The general quadratic class of time-frequency representation methods can be used for this purpose, and huge number of known time-frequency kernels can be found for different representations. However, many of these are actually developed for deterministic signals (possibly with a disturbance) and used for a more exploratory description of multicomponent signals. In the recent 10-15 years, the attempts to find optimal estimation algorithms for non-stationary processes, have increased, [61]. A fundamental problem is the estimation of a time-frequency spectrum from a single realization of the process. The IAF for the non-stationary process is

$$r_z(t, \tau) = \mathcal{E}[z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})], \quad (7.1)$$

with the estimate from one realization as

$$\hat{r}_z(t, \tau) = z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2}). \quad (7.2)$$

The Wigner spectrum is defined as
\[
W_z(t, f) = \int_{-\infty}^{\infty} r_z(t, \tau) e^{-i2\pi f \tau} d\tau, \tag{7.3}
\]

with an estimate found as \[
\hat{W}_z(t, f) = \int_{-\infty}^{\infty} \hat{r}_z(t, \tau) e^{-i2\pi f \tau} d\tau.
\]

### 7.1 Definitions of non-stationary processes

The spectral decomposition theorem for a stationary process admits the Cramér harmonic decomposition

\[
z(t) = \int_{-\infty}^{\infty} e^{i2\pi f_1 t} dZ(f_1). \tag{7.4}
\]

This decomposition has double orthogonality, which means that

\[
\int_{-\infty}^{\infty} e^{i2\pi f_1 t} e^{-i2\pi f_2 t} dt = \delta(f_2 - f_1). \tag{7.5}
\]

and that

\[
\mathcal{E}[dZ(f_1)dZ^*(f_2)] = \delta(f_1 - f_2)S_z(f_1)df_1df_2. \tag{7.6}
\]

A zero-mean Gaussian non-stationary process is characterized by the covariance function

\[
p_z(t_1, t_2) = \mathcal{E}[z(t_1)z^*(t_2)], \tag{7.7}
\]

which is dual to the spectral distribution function \(P_z(f_1, f_2)\) by means of the Fourier-type relation,

\[
\mathcal{E}[z(t_1)z^*(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_z(f_1, f_2) e^{i2\pi(f_1t_1 - f_2t_2)} df_1df_2. \tag{7.8}
\]

The spectral increments of \(P_z(f_1, f_2)\) are not orthogonal as is the case for a stationary process, where the spectral distribution function reduces to \(P_z(f_1, f_2) = \delta(f_1 - f_2)S_z(f_1)\). However, for the existence of a decomposition, the requirement

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |P_z(f_1, f_2)| df_1df_2 < \infty, \tag{7.9}
\]

has to be fulfilled. The corresponding process \(z(t)\) is called harmonizable, [66, 67].
If we want to maintain the double orthogonality of the decomposition, the so-called Karhunen decomposition, can be applied, where the eigenfunctions of the covariance kernel is used, [61]. The covariance has the form

$$E[z(t_1)z^*(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(t_1, f_1)\psi^*(t_2, f_2)E[dZ(f_1)dZ^*(f_2)]$$

$$= \int_{-\infty}^{\infty} \psi(t_1, f_1)\psi^*(t_2, f_1)S_z(f_1)df_1.$$

Using the orthogonality relation

$$\int_{-\infty}^{\infty} \psi(t_1, f_1)\psi^*(t_2, f_2)dt_1 = \delta(f_2 - f_1), \quad (7.10)$$

we find the Karhunen representation as

$$\int_{-\infty}^{\infty} p_z(t_1, t_2)\psi(t_2, f_1)dt_2 = S_z(f_1)\psi(t_1, f_1). \quad (7.11)$$

The advantage of the Karhunen representations is the orthogonality in time as well as frequency. However, the interpretation of the variable $f_1$ as frequency is no longer obvious.

The decomposition of the autocovariance leads to

$$E[|z(t)|^2] = \int_{-\infty}^{\infty} |\psi(t_1, f_1)|^2S_z(f_1)df_1, \quad (7.12)$$

which defines a time-dependent power spectrum as

$$W_z(t_1, f_1) = |\psi(t_1, f_1)|^2S_z(f_1). \quad (7.13)$$

As the non-stationarity property is very general, different restrictions have been applied. One definition is the locally stationary process (LSP), [68], which has a covariance of the form

$$p_z(t_1, t_2) = q\left(\frac{t_1 + t_2}{2}\right) \cdot r(t_1 - t_2), \quad (7.14)$$

where $q(t)$ is a non-negative function and $r(t)$ positive definite. Studying a formulation as a local symmetric covariance function, i.e.,

$$E[z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})] = q(t) \cdot r(\tau), \quad (7.15)$$
we see that the LSP is a result of a modulation in time of a stationary covariance function. With \( q(\tau) = e^{-\frac{\tau^2}{2}} \) and \( r(\tau) = e^{-\frac{c\tau^2}{2}} \), we study some realizations in Figure 7.1 for different parameter values. When \( q(t) \) tends to be a constant function, the LSP becomes more similar to a stationary process.

![Figure 7.1: Examples of realizations of a locally stationary processes for different parameter values, \( c \).](image)

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

\[
z(t) = d(t)w(t),
\]

(7.16)

where \( w(t) \) is a complex-valued weakly stationary process. The local covariance function becomes
\[ \mathcal{E}[z(t + \frac{\tau}{2})z^*(t - \frac{\tau}{2})] = d(t + \frac{\tau}{2})d(t - \frac{\tau}{2})r_w(\tau). \] (7.17)

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

A third definition is connected to linear systems, where we can represent \( z(t) \) as the output of a linear time-varying system \( H \) whose input is stationary white noise, denoted \( n(t) \), with power spectral density \( S_n(f) = 1, \) \[69\],

\[ z(t) = (Hn)(t) = \int_{-\infty}^{\infty} h(t,t')n(t')dt', \] (7.18)

We note that if \( z(t) \) is stationary, \( H \) is time-invariant and the power spectral density of \( z(t) \) is given by \( S_z(f) = |H(f)|^2 \), where \( H(f) = \int_{-\infty}^{\infty} h(\tau)e^{-i2\pi f \tau}d\tau \). The system \( H \) 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. Many papers relating to time-frequency analysis and linear time-varying systems can be found, e.g., \[65\],

Priestley, \[70\], suggested the **oscillatory random process**, which is a rather restricted class of processes, as a linear combination of oscillatory processes is not necessarily oscillatory, where a basis function

\[ \psi(t,f) = A(t,f)e^{i2\pi ft}, \] (7.19)

is applied giving the process \( z(t) \) as

\[ z(t) = \int_{-\infty}^{\infty} \psi(t,f)dZ(f). \] (7.20)

It is worth noting that the basis functions are not necessarily orthogonal for the Priestley decomposition. Remembering the Cramér decomposition of a stationary process \( z(t) = \int_{-\infty}^{\infty} e^{i2\pi ft}dZ(f) \) where the process is decomposed using the orthogonal basis functions \( e^{i2\pi ft} \) we see that the Priestley decomposition is similar in the sense that an amplitude modulation of each complex exponential is applied for the non-stationary oscillatory process. If \( A(t,f) \) is slowly varying in time for each frequency the decomposition is close to orthogonal as

\[ \int_{-\infty}^{\infty} \psi(t,f_1)\psi^*(t,f_2)dt = \int_{-\infty}^{\infty} A(t,f_1)A^*(t,f_2)e^{i2\pi(f_1-f_2)t}dt \approx \delta(f_1 - f_2), \] (7.21)

when \( A(t,f_1)A^*(t,f_2) \approx 1 \). The so-called **evolutionary spectrum** is found as
\[ S_z(t, f) = |A(t, f)|^2 S_z(f). \] (7.22)

### 7.2 The mean square error optimal kernel

Sayeed and Jones [62] 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,

\[ J(\phi) = \int \int \mathcal{E}[\hat{A}_z(\nu, \tau)\phi(\nu, \tau) - \mathcal{E}[\hat{A}_z(\nu, \tau)]]^2 d\nu d\tau, \] (7.23)

resulting in the optimal ambiguity kernel

\[ \phi_{opt}(\nu, \tau) = \frac{\mathcal{E}[\hat{A}_z(\nu, \tau)]^2}{\mathcal{E}[\hat{A}_z^2(\nu, \tau)]}. \] (7.24)

Based on the fact that the ambiguity spectrum of a LSP is separable (rank one),

\[ \mathcal{E}[\hat{A}_z(\nu, \tau)] = A_z(\nu, \tau) = Q(\nu)r(\tau), \] (7.25)

where \( Q(\nu) = \int q(t)e^{-i2\pi t\nu}dt \), the optimal kernel for Gaussian circularly symmetric processes was derived in [63] as,

\[ \phi_{opt}(\nu, \tau) = \frac{|Q(\nu)|^2|r(\tau)|^2}{|Q(\nu)|^2|r(\tau)|^2 + (\int |r(t)|^2e^{-i2\pi \nu t}dt(q*q)(\tau))}. \] (7.26)

The multitapers of the LSP-optimal kernel are also shown to approximate Hermite functions and in [42], the these multitapers and corresponding weights (eigenvalues) are studied and evaluated for the multitaper spectrogram, Eq. (5.17). Some examples of LSP-optimal windows and weights are seen in Figure 7.2.

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 \( B = 4/64 \), the Thomson multitapers, \( K = 2, B = 4/64 \) 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.

Optimization of the weighting factors in the spectrogram sum has also been made using different criteria, e.g., minimum mean square error, [43], detection of instantaneous frequency, [36, 41] and minimum (Rényi) entropy, [50]. In all these cases, the multitaper spectrograms are based on the use of Hermite functions as windows. Very few contributions are found concerning the actual design or estimation of suitable multitapers for a specified non-stationary processes in a certain application, i.e., designing from the actual properties of the spectrum.
Figure 7.3: Examples of resulting multitaper spectrograms of EEG data.
Bibliography


82


