Department of Informatics
University of Bergen
Norway

Master of Science Thesis

Recursive Time-Frequency Reassignment

Geir Kjetil Nilsen

2007

Acknowledgements

I would first of all like to thank Prof. Hans Munthe-Kaas for making it possible for me to work with signal processing and for sharing his wide knowledge on our frequent meetings at Matematisk Institutt. I am grateful for his support and for the many doors he has opened for me.

I would like to thank Øyvind Lunde Rørtveit for his ever-ready attitude to help and discuss problems. Also thanks to Christine Drengenes and Joakim Grahl-Knudsen for reading the thesis.

Finally, I would like to thank Bjørn Harald Fotland for his company through 5 years at Institutt for Informatikk, and my parents for supporting me in every thinkable way.

Geir Kjetil Nilsen
Bergen, 2007

Contents

Summary

Time-Frequency Analysis (TFA) is the name of a scientific field for which the agenda is to provide joint time-frequency representations (TFRs) of time-series. TFA has numerous applications, ranging from the obvious visualization of signals with both time and frequency as coordinates, to being the basis for more sophisticated algorithms such as the phase vocoder [11].

In this thesis, a fast algorithm for producing TFRs is proposed. The resulting TFRs have optional time-frequency resolutions up to optimality. The method exploits that implementations of recursive linear systems compared to non-recursive linear systems yields a substantial reduction in the computational cost. As these TFRs rise from recursive linear systems, they are here reffered to as Recursive TFRs (RTFRs).

A theory connecting the RTFRs to the Short-Time Fourier Transform (STFT) spectrograms will be presented. It is shown that the RTFRs are a special case of the STFT spectrograms via an special infinite window function. The theory reveals that the RTFRs suffer from the same uncertainty as the regular spectrograms. This invites for combining the RTFRs with a recent method called time-frequency reassignment. The extra computations needed in the reassignment method turns out to be available through simple modifications of the transfer function describing the recursive system. This has the implication that the Reassigned RTFRs (RRTFRs) can be computed with the same complexity as the RTFRs. It has the main implication of making time-frequency reassignment well suited for *real-time implementations*.

The theory connecting the RTFRs to the STFT spectrograms also expanded the literature resources used in this thesis. It was recently[1] discovered that the idea of using an infinite window function with the STFT is not new [2, 34, 35, 39, 40]. Some of these had also developed recursive implementations based on the same type of STFT window functions. Nevertheless, exploiting these recursive structures in use with the reassignment method is new. The main contribution of original work in this thesis is hence called recursive time-frequency reassignment.

The thesis is organized as follows: Chapter 1 gives the most basic definitions, while Chapter 2 presents an overview of the basic principles behind time-frequency analysis, focusing on the fundamental limitation called time-frequency

[1] September, 2007

uncertainty and the method known as time-frequency reassignment.

In Chapter 3 a summary of well-known TFRs is presented. Chapter 4 introduces the proposed recursive system for producing TFRs. The theory in Chapter 4 was developed independently of any other, but was slightly updated according to the new literature mentioned above. As the theory is presented using continuous mathematics, the end of Chapter 4 deals with important discretization issues done in a manner not found elsewhere.

In Chapter 5 the time-frequency reassignment method is presented from scratch, while in Chapter 6 the time-frequency reassignment method is brought into the recursive system described in Chapter 4.

Chapter 7 consists of three sections. The first section is regarding implementation issues, while the second section gives numerical results. The last section gives concluding remarks, and ideas for further work.

A journal paper based on the results in this thesis has been submitted to IEEE Transactions on Signal Processing. The paper is attached to this thesis, and can be found in the Appendix.

Chapter 1

Basic Definitions

This chapter summarizes the basic terminology required in the following chapters. It is, however, assumed that the reader is familiar with the most used concepts trough signal processing, system theory and/or differential calculus. For introductory material, see for instance [22, 24, 32, 45].

1.1 Bandwidth

The classical definition of the bandwidth, Δ_ω, of a filter, $h(t)$, is the second moment of its squared magnitude response

$$\Delta_\omega = \frac{\int \omega^2 |H(\omega)|^2 d\omega}{\int |H(\omega)|^2 d\omega}, \tag{1.1}$$

and is thus a measure of its extent in frequency.

1.2 Duration

The classical definition of the time duration, Δ_t, of a filter, $h(t)$, is the second moment of its magnitude squared

$$\Delta_t = \frac{\int t^2 |h(t)|^2 dt}{\int |h(t)|^2 dt}, \tag{1.2}$$

and is therefore a measure of its time extent.

1.3 Time-Frequency Uncertainty

The time-frequency uncertainty for a time-domain filter, $h(t)$, is given by the product

$$\Delta_t \Delta_\omega, \tag{1.3}$$

where Δ_t and Δ_ω represents $h(t)$'s duration and bandwidth defined in (1.1) and (1.2), respectively. Thus, the lower the uncertainty value, the higher accuracy. On the other hand, the term time-frequency resolution is defined as a quantity inversely proportional to the time-frequency uncertainty. Hence, the greater the resolution value, the higher accuracy.

1.4 Q Factor

The Q factor of a filter is a measure of the sharpness of its response. As Q increases, the range of frequencies compared with its centre frequency becomes more narrow. The Q factor is defined as

$$Q = \frac{\omega_0}{\Delta\omega}, \tag{1.4}$$

where ω_0 is the centre frequency of the magnitude response, and $\Delta\omega$ is the bandwidth.

1.5 Hilbert Transform

The Hilbert Transform, $\mathcal{H}\{x(t)\}$ of a signal $x(t)$ can be defined as

$$\mathcal{H}\{x(t)\} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}d\tau, \tag{1.5}$$

when the integral exists. There are subtle mathematical issues regarding the existence of the above integral. The interested reader can consult [20] for a excellent survey. The Hilbert Transform shows up in a number of situations; it might be used implicitly or explicitly in time-frequency analysis as it is related to the concept of analytic signals.

1.6 Analytic Signals and The Analytic Associate

Given a real, continuous time-domain signal, $x(t)$, its continuous complex analytic associate, $z(t)$, of the real variable t is defined

$$z(t) = x(t) + i\mathcal{H}\{x(t)\}. \tag{1.6}$$

Signals $z(t)$ that satisfy (1.6) are called analytic. The effect of taking the complex modulus of an analytic signal gives the envelope effect illustrated in figure (1.1). This has the effect of removing the oscillations from the waveform by encapsulating it. In general, analytic signals contains no negative frequency components. A complex coefficient filter can be such that the real and imaginary parts of its output form an approximate Hilbert Transform pair, i.e. the filter will produce complex analytic output from real input.

1.7 Instantaneous Frequency

The instantaneous (radian) frequency, $\omega_i(t)$, of a complex continuous analytic time-domain signal $z(t)$ of the real variable t is defined

$$\omega_i(t) = \frac{\partial}{\partial t}arg\{z(t)\} \tag{1.7}$$

where the $arg\{\}$ operator represents the complex argument. If the signal $z(t)$ in (1.7) is not analytic, it must first be replaced by its analytic associate defined in (1.6).

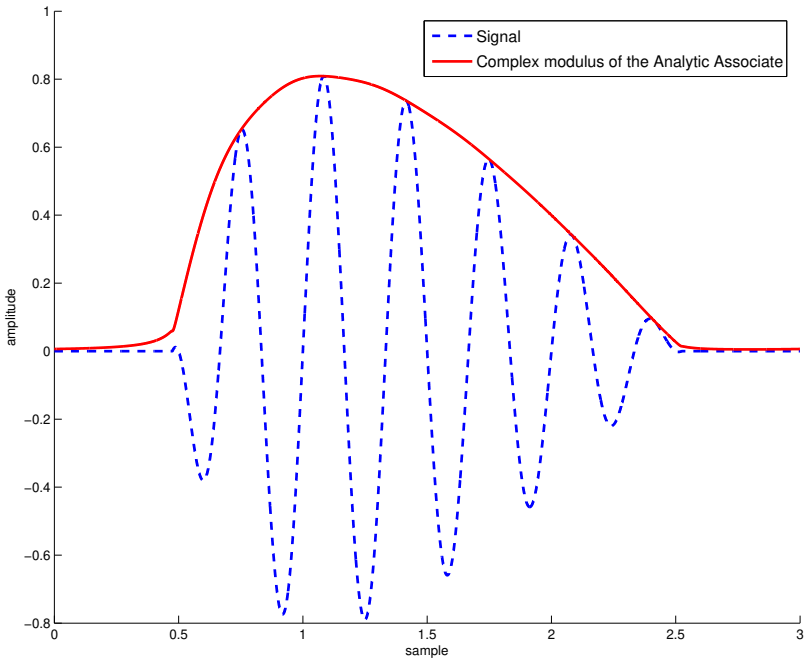


Figure 1.1: *The signal (stipulated blue), and the complex modulus of its analytic associate (red).*

At this point is important to emphasize that, by definition, instantaneous frequency is only physically meaningful when $z(t)$ just has one single frequency component present at any time instant. For instance, as shown by [15], if $z(t)$ consists of two frequency components with equal amplitudes, the instantaneous frequency will be the mean of the two frequencies. Worse, if the amplitudes are not equal, the instantaneous frequency will vary periodically with the difference frequency of the two. As the number of components increase, the behaviour turns more and more chaotic. This limitation is the reason why this function generally cannot be used as a meaningful TFR.

1.8 Group Delay

The group delay, $t_g(\omega)$, of a complex continuous analytic frequency-domain signal $Z(\omega)$ of the real variable ω is defined

$$t_g(\omega) = -\frac{\partial}{\partial\omega} arg\{Z(\omega)\} \tag{1.8}$$

If $z(t)$ is not analytic, it must as above be replaced by its analytic associate. This quantity is the dual of instantaneous frequency; it gives the time as a function of frequency. As with the instantaneous frequency, the group delay is also only physically meaningful for component signals.

1.9 The Laplace Transform

The Laplace Transform, $H(s) = \mathcal{L}\{h(t)\}$, of a continuous time-domain signal $h(t)$ of the real variable t is defined

$$H(s) = \mathcal{L}\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-st}dt. \tag{1.9}$$

with the complex variable $s = \sigma + i\omega$.

1.10 The z-Transform

The z-Transform, $H(z) = \mathcal{Z}\{h(n)\}$, of a discrete time-domain signal $h(n)$ of the real variable n is defined

$$H(z) = \mathcal{Z}\{h(n)\} = \sum_{n=0}^{\infty} h(n)z^{-n}. \tag{1.10}$$

The z-transform is the Laplace Transform's discrete counterpart. The complex variable z can therefore be seen to be $z = e^{sT}$, where T is the sampling period.

1.11 The Bilinear Transform

The Bilinear Transform is a mapping of the continuous, complex Laplace-domain variable, s, to the continuous complex z-domain variable z. The mapping can be derived in many ways, but they are all related to fact that they are finite approximations to $z = e^{sT}$.

The simplest approximation is by using the first order Taylor series expansion. This gives $s = (1 - z^{-1})/T$ which is the first order finite difference. This makes sense, since multiplication by s in the Laplace-domain corresponds to differentiation in the time-domain.

The best approximation of a function by a rational function of given order is known to be the Padé approximant. The first order Padé approximant gives the definition of the Bilinear Transform

$$s = \frac{2(1 - z^{-1})}{T(1 + z^{-1})}. \tag{1.11}$$

1.12 The Discrete Fourier Transform (DFT)

The DFT, $X(m)$, of a discrete signal $x(n)$ of the real variable n, consisting of N samples is defined

$$X(m) = \sum_{n=0}^{N-1} x(n)e^{\frac{-2\pi inm}{N}} \tag{1.12}$$

where $m = 0, 1, \ldots, N - 1$ is a real, discrete variable.

1.12.1 The Fast Fourier Transform (FFT)

The FFT is a fast algorithm implementing the DFT in (1.12) [7]. The algorithm exploits Danielson-Lanczos Lemma [41], and applies a Divide and Conquer strategy to yield a computational complexity of $O(N \log N)$. This opposed to $O(N^2)$ as direct implementation of (1.12) would require.

Chapter 2

An Overview of Time-Frequency Analysis (TFA) with Focus on Reassignment

”Visible Patterns of Sound”. The title was published in SCIENCE [25] in November 1945. It became clear that Bell Telephone Laboratories had held their cards strategically tight from the beginning of the war. It was time to bring true pictures of sound to public attention. The machine called the ”spectrograph” was an analog device capable of visualizing sound. Deaf could now use the telephone, and for those who already could speak, it could help them improve their speech. The possibilities were endless and potentially revolutionary. This very first approach of automatically producing images of sound was the beginning of TFA.

The legendary 1946 article, ”Theory of Communication” [13] by Gabor, was the first paper to put the workings of the spectrograph into strict mathematical terms [28]. His work could be seen as a digital approximation of the spectrograph device, and was a framework for representing one-dimensional signals in two dimensions, with time and frequency as coordinates.

Gabor’s work was influenced by a fundamental result of Heisenberg [18, 42]. This principle is important in quantum mechanics, and states that the exact velocity and position of a particle cannot simultaneously be known. Gabor extended this to time and frequency, and derived an analogous principle. For all measurable continuous functions, it is true that its time extent multiplied by its frequency extent always is greater or equal to a certain constant. Since this product therefore cannot be infinitesimal, there will always be a uncertainty in time or frequency. The term time-frequency uncertainty was therefore given to this product.

Gabor’s principle has the consequence that any physical signal must occupy at least a minimum area in the time-frequency plane. Gabor went on to show that for Gaussians this minimum is achieved exactly. This suggested that by expanding signals in Gaussian elementary components (named Gabor logons), one will have a time-frequency representation (TFR) optimally localized in both time and frequency.

However, Gabor did not characterize this representation as a set of short Fourier Transforms applied on a segmented signal. Much later the equivalence between Gabor’s expansion and the Short-Time Fourier Transform (STFT) using

a Gaussian window function was realized. The modulus squared of the STFT was established as the spectrogram, and became the most popular and efficient way of computing TFRs using the Fast Fourier Transform (FFT) algorithm. The STFT will be discussed in greater detail in Chapters 3, 4 and 5. However, the spectrogram in accordance with Gabor's result, will always be smeared in time if the time uncertainty is chosen too high. In the same way, it will be smeared in frequency if it is chosen too low. Since the frequency uncertainty is proportional to the reciprocal of the time uncertainty, the same smearing occurs the other way around.

Despite this fundamental difficulty, Ackroyd [1] used Rihaczek's [26] complex energy density theory and was able to show that the spectrogram can be written as a double convolution between the energy densities of the signal and the window. This conveyed insight into the internal structure of the spectrogram: the energy density of the analyzed signal can be seen smoothed by the energy density of the STFT window function.

In 1989 Cohen published a review [6] on TFA. His work could be seen as an unification of all the TFA theory stemming from many different branches in science. He found a general formula reducible to most of the TFRs known at that time. The class corresponding to all the members of this formula is today called Cohen's class. The spectrogram was shown to be just one of these members. It became evident that not only was the spectrogram a smoothed version of the complex energy density function of Rihaczek, but all members of Cohen's class were smoothed versions of the TFR called the Wigner-Ville Distribution.

A possible solution counteracting the smoothing of the spectrogram was proposed by Kodera et al. [21] in 1976. This was the invention of what is today known as time-frequency reassignment. Kodera et al. proposed a clever enhancement reversing the effect of the smoothing. The crucial step is to *reassign* each point in the spectrogram to new coordinates better fitting with the true energy density of the analyzed signal. It was shown that the reassigned coordinates were related to the definitions for the instantaneous frequency and group delay of the analyzed signal. These coordinates were also shown to be related to the partial derivatives of the STFT phase. However, Kodera et al. only used approximate finite differences to compute these partial derivatives.

Various researchers have continued the development of the method [3, 8, 12, 21, 23]. These developments apply particularly to the spectrogram, but also in general to other members of Cohen's class. Auger and Flandrin's [3, 8] generalization of the reassignment method lead to an important discovery. The reassignment coordinates for the spectrogram were found to be computed efficiently and exactly via only two extra STFTs. Nelson [23] independently derived the same results for the spectrogram, but calculated the reassignment coordinates by his approximative cross spectral method discussed in Section 5.6.2.

Although introduced in the 1970s, the reassignment method is still not widely used. The main reason is speculated to be implementation difficulties *before* Auger and Flandrin's generalization [8]. Another possible reason is the irreversibility, i.e. when the reassignment method is applied, one cannot get back the original signal because the phase information corresponding to the reassigned spectrogram is lost. Ongoing research by Fulop and Fitz [29] concerning

the phase corrections that must be made to attain reversibility, is promising. Relatively fresh papers on the subject of reassignment are still being published, widening its applications and users. For instance, Plante, as well as Nelson, [10, 23] used the method to analyze speech signals. Hainsworth [16, 17] for both review and analysis, together with an in-depth thesis regarding musical transcription. Fulop and Fitz [28], and Auger and Flandrin [8] also regarding implementation issues. This thesis further expands the applications of time-frequency reassignment. It will be seen that the proposed recursive time-frequency reassignment algorithm is especially well suited for *real-time implementations*.

Chapter 3

Linear and Bilinear Systems for Time-Frequency Representations (TFRs)

3.1 Introduction

The process of producing a two-dimensional TFR from a one-dimensional signal must involve some form of filtering. As will be demonstrated in this chapter, well-known TFRs can be modelled by either linear or bilinear systems.

3.2 Frequency Representations

A time-domain signal $x(t)$ can be transformed into the complex frequency-domain signal $X(\omega)$. The operation connecting the domains is called the Fourier Transform.

3.2.1 The Fourier Transform

The Fourier Transform of a signal $x(t)$ can be defined as

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt. \tag{3.1}$$

This well-known transform takes a time-domain signal completely into the dual frequency-domain. The frequency-domain representation makes it possible to study all frequencies present in the signal. The Fourier transform can be regarded as inner products between the analyzed signal $x(t)$ and all possible frequency shifted complex exponentials. The complex exponential in the Fourier transform is often called the Fourier basis. The inner products with the Fourier basis identify the amplitude and phase of each individual frequency component occurring in the transformed signal.

To be practical as a visual frequency representation, the Fourier transformed signal is usually studied in terms of its amplitude information. The power spectrum is defined as the complex modulus squared of (3.1)

$$|X(\omega)|^2. \tag{3.2}$$

As the inner products are linear operations, the Fourier transform can be regarded as a linear system. On the other hand, if one considers the power spectrum it must rather be regarded as a bilinear system because of the modulus squared operation.

3.2.2 The Constant Q Transform, Logarithmic Frequency Scale & Multiresolution

Frequency representations are frequently used to analyse musical signals. Since musical signals often are harmonic, it can be convenient to have a logarithmic frequency scale such that the perceived sounds better correspond with the visual representation. Within this context, as introduced by [5], the Constant Q transform is closely related to the Fourier transform. The spacing of the Fourier basis in the frequency-domain can be chosen logarithmically, and its extent can be made dependent of its centre frequency. Hence, the Q value can be chosen to be constant, thus the name constant Q. Visually, this implies a Fourier basis with narrow frequency-domain extent on low frequencies (i.e. wide time-domain extent), and wide frequency-domain extent on high frequencies (i.e. narrow time-domain extent). This setup is related to the human auditory system, and it will yield frequency dependent resolution in the frequency domain; that is, multiresolution. Multiresolutional analysis is especially well suited for musical signal, as sounds of low frequency often are close in frequency, while sounds of high frequency often are close in time.

For the same reasons as with the Fourier transform, the Constant Q transform can be regarded as a linear system, and the belonging power spectrum as a bilinear system.

3.3 TFRs

There is a wide range of material available on the topic of Fourier analysis (see for instance [19]), and in this brief introduction it is assumed that the reader is familiar with the basic concepts. From now the focus will be on creating a hybrid representation involving time and frequency simultaneously.

3.3.1 The Spectrogram

The Fourier basis can be combined with a window function $\mathcal{W}(t)$ that localizes its time-domain extent. This basically means that the infinite time-domain extent of the Fourier basis is reduced. Hence, the inner products now become time-dependent, and this is exactly the definition of convolutions. The result is usually called the Short-Time Fourier Transform (STFT). The STFT of the signal $x(t)$ can be defined

$$X(t,\omega) = e^{i\omega t} \int_{-\infty}^{\infty} x(\tau)\mathcal{W}(t-\tau)e^{-i\omega\tau}d\tau \tag{3.3}$$

The TFR called the spectrogram is defined as the modulus squared of (3.3)

$$|X(t,\omega)|^2. \tag{3.4}$$

As (3.3) can be regarded as linear convolutions between the input signal $x(t)$ and frequency shifted windows, the STFT is nothing else than a linear system.

The spectrogram, on the other hand, must be seen as a bilinear system for the same reasons as with the power spectrum.

The main limitation of the STFT spectrogram is its blurry appearance. Even for single component, steady state signals, the spectrogram is smoothed because of the fixed window function. This will be discussed in detail in Section 3.3.4, while a possible solution to the problem is introduced in Chapter 5.

The idea of combining the constant Q property with the spectrogram would yield a generalization first presented by [44] in 1978. The window function will be such that its bandwidth is dependent on the centre frequency ω. This is, in fact, the first step in the direction of the invention of Wavelet theory [36]. By introducing a ω-factor inside the argument of the window function, and recalling the scaling theorem of Fourier analysis, it becomes evident that this modified STFT can be made constant Q. In addition one might choose the frequencies ω logarithmically for the same reasons as pointed out earlier.

3.3.2 The Scalogram

The Continuous Wavelet Transform (CWT) of a signal $x(t)$ can be defined as

$$X(t,s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(\tau)\Psi\left(\frac{t-\tau}{s}\right) d\tau, \tag{3.5}$$

with $\Psi(t)$ being called the Wavelet basis. The fundamental idea with the CWT is that the frequency shifting operation which occurs in the STFT is replaced by a frequency scaling operation. Recall that the windowed Fourier basis in the STFT is of constant duration with an increasing number of oscillations as higher frequencies are analyzed. With the CWT the duration of the Wavelet basis decreases as higher frequencies are analyzed, but the number of oscillations remains the same. The factor in front of the integral is to ensure that the energy of the Wavelet, $\Psi(t)$, is normalized for all scales. The important difference is that scaling automatically yields constant Q analysis. The CWT is thus a generalization of the constant Q STFT.

This transform yields no TFR, but rather a Time-Scale Representation (TSR). The TSR called the scalogram is defined as the modulus squared of (3.5)

$$|X(t,s)|^2. \tag{3.6}$$

The CWT can be seen as linear convolutions between the input signal signal $x(t)$ and scaled Wavelets. The CWT therefore also forms a linear system. However, as with the power spectrum and the spectrogram, also the belonging scalogram must necessarily be seen as a bilinear system.

Like the spectrogram, also the scalogram suffers from blurring in the time-scale plane. This is caused by the same reason and will be explained in Section 3.5.

The applications and theory of Wavelets are enormous, and the CWT is only one of many related transforms. Nevertheless, the TSR produced by the CWT is the only one well suited for visual inspection. For a practical guide refer to [36], or [14] for a more extensive reference.

3.3.3 The Wigner-Ville Distribution (WVD)

The WVD[1] was first introduced by E. Wigner in 1932 in context of quantum mechanics [43], and later independently developed by J. Ville for signal analysis [30, 38]. The WVD of a analytic signal $x(t)$ can be defined as

$$X(t,\omega) = \int_{-\infty}^{\infty} x^{\star}(t - \frac{1}{2}\tau)x(t + \frac{1}{2}\tau)e^{-i\omega\tau}d\tau, \tag{3.7}$$

where $\star$ denotes complex conjugation. The WVD is different from both the (constant Q) STFT and the CWT. Instead of using a predefined window, such as $\mathcal{W}(t)$ in (3.3), or the predefined Wavelet, such as $\Psi(t)$ in (3.5), the WVD can be seen as taking the time-reversed version of the analyzed analytic signal $x(t)$ as the window [8]. The WVD therefore introduce bilinearity coincidentally with the convolutions, meaning that this system is inevitably bilinear.

Since the window now is dependent on the analyzed signal, the WVD does not suffer from blurring in the time-frequency plane. However, the bilinearity also has the major drawback of introducing non-negliable cross-terms when the analyzed signal contains more than one component. The cross-terms are spawned from the bilinearity like $2xy$ in $(x+y)^2$. It is easily verified that the cross-terms will be highly oscillatory. As discussed in the next section, the cross-terms can be reduced at the expense of re-introducing blurring!

3.4 Cohen's Class of TFRs

Cohen [6] showed that a large class of TFRs – now called Cohen's class – can be written in the general form

$$X(t,\omega) = \frac{1}{4\pi}\int\int\int x^{\star}(t - \frac{1}{2}\tau)x(t + \frac{1}{2}\tau)\Phi(\theta,\tau)e^{-i\theta t - i\tau\omega + i\theta u}d\theta d\tau du, \tag{3.8}$$

where $x(t)$ is the analytic input signal, and where the function $\Phi(\theta,\tau)$ is called the kernel function specified in the Doppler-lag domain. The expression in (3.8) reduces to the WVD in (3.7) if $\Phi(\theta,\tau) = 1$ is chosen. It can also be shown that with the kernel function

$$\Phi(\theta,\tau) = \int \mathcal{W}^{\star}(u - \frac{1}{2}\tau)\mathcal{W}(u + \frac{1}{2}\tau)e^{-i\theta u}du, \tag{3.9}$$

equation (3.8) precisely reduces to the spectrogram in (3.4).

It is evident that Cohen's general expression (3.8) also involves the same bilinearity on the input signal $x(t)$ as seen in the WVD. This must seemingly introduce cross-terms in all the members of Cohen's class. However, as the cross-terms are highly oscillatory, smoothing can serve to reduce them. This is equivalent to choosing the kernel function in (3.8) to be a function with low-pass characteristics in the time-frequency domain. Indeed, this is the case with the spectrogram kernel function in (3.9). The spectrogram can for these reasons be seen as a smoothed version of the WVD. That is, the cross-terms are traded

[1] Distribution and representation is used interchangly in the literature.

with a blurry representation. In fact, all members of Cohen's class can be seen as smoothed versions of the WVD.

The possibilities branching from this very general theory are numerous. For a comprehensive reference see [4].

3.5 The Affine Class

Rioul and Flandrin [27] extended Wavelet transforms to a general TSR class now called the affine class. The implication is that the scalogram can be seen as a affine smoothed version of the WVD. That is, a frequency dependent smoothing of the WVD. In other words, the scalogram belongs to the affine class, not to Cohen's class.

Chapter 4

A Recursive Linear System for TFRs

4.1 Introduction

The aim of introducing a different underlying system in context of TFRs is twofold. Firstly, the computational cost of the system's implementation should be low, and secondly, the implementation should be easy. By defining a recursive linear system with its rational transfer function, the filter order is reduced drastically, and so is the computational cost. Discretized recursive linear systems are easily implemented in terms of linear implicit recursions.

In the previous chapter it was argued that all the well-known TFA methods can be modelled by linear or bilinear systems in terms of linear convolutions. As a motivation, the STFT is again considered. The definition of the STFT is repeated here for convenience. It is defined

$$X(t,\omega) = e^{i\omega t}\int_{-\infty}^{\infty} x(\tau)\mathcal{W}(t-\tau)e^{-i\omega\tau}d\tau. \tag{4.1}$$

The impulse response of this system is easily found to be

$$h(t) = \mathcal{W}(t)e^{i\omega t}, \tag{4.2}$$

where $\mathcal{W}(t)$ is the window. The magnitude response of (4.2) has its centre located at ω provided that the window has low-pass characteristics. Moreover, the magnitude response is also zero for all negative frequencies, so it follows that (4.2) is analytic. The output of STFT system (4.1) will also be analytic, since it forms linear convolutions between $x(t)$ and (4.2).

The crucial point is now that by a clever choice of $\mathcal{W}(t)$, (4.1) can become a solution to an ordinary differential equation. The differential will act on the output's time dimension, and represent a recursion propagating over in the discrete case as a linear implicit recursion.

This convolutional view, however, makes the discretized spectrogram very redundant. In other words, the number of samples in the time-dimension will be equal to the number of samples in the analyzed signal. A more efficient method avoiding the redundancy is possible. Essentially, this alternative method stems from viewing (4.1) as a set of windowed Fourier transforms instead. Now one can introduce a window hop factor not necessarily equal to one, so that the output

will implicitly be downsampled by the same factor. Anyhow, the redundancy is, in fact, necessary for the time-frequency reassignment method presented in Chapter 5 to work. It is therefore appropriate to compare computational cost with the convolutional view.

4.2 A Second Order Recursive System

It turns out that the simplest recursive system existing, is the preferable choice for reasons that will be pointed out in Section (6.6.4). First, a second order recursive system of this kind is considered, then afterwards it is generalized to order k.

The second order recursive system can be expressed as

$$\frac{\partial^2}{\partial t^2}y(t,\omega_p) + 2(\sigma_p - i\omega_p)\frac{\partial}{\partial t}y(t,\omega_p) + (\sigma_p^2 - 2i\omega_p\sigma_p - \omega_p^2)y(t,\omega_p) = \sigma_p^2 x(t), \quad (4.3)$$

where $y(t,\omega_p)$ is the complex, continuous two-dimensional output signal of the real, continuous variables t and ω_p. The input signal, $x(t)$, is a real continuous one-dimensional signal of the real variable t. The recursion in the continuous system is represented by the differential acting on the output's time dimension. The transfer function of this system is in the Laplace-domain given by

$$H(s) = \frac{\sigma_p^2}{(s+p^\star)^2}, \quad (4.4)$$

where $\star$ denotes the complex conjugate, and the pole p defined as the complex constant

$$p = \sigma_p + i\omega_p. \quad (4.5)$$

Note that the constant in the nominator of (4.4) is present just to ensure that the magnitude response is normalized to unity at $\omega = \omega_p$. The impulse response corresponding to (4.3) is therefore

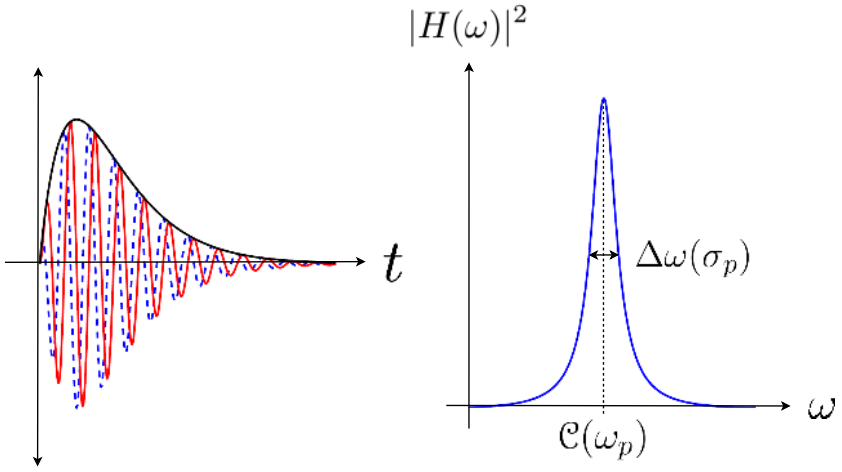
$$h(t) = \mathcal{L}^{-1}\{H(s)\} = \sigma_p^2 t e^{-p^\star t}u(t), \quad (4.6)$$

where $u(t)$ is the heaviside step function. Since p is a complex constant, (4.6) is a causal decaying complex exponential, and the decay rate is controlled by the real part of p, that is, σ_p. This constant is thus related to the duration of (4.6) – or equivalently – the bandwidth of the magnitude response of (4.4). It should be noticed that (4.6) can be written as

$$h(t) = \underbrace{\sigma_p^2 t e^{-\sigma_p t}u(t)}_{\text{window}}\, e^{i\omega_p t}. \quad (4.7)$$

The recursive system in (4.3) can therefore be seen as a special case of the STFT utilizing a causal, asymmetric and infinite window.

Two functions can now be parametrized. Using that σ_p is related to the bandwidth of (4.4), and that ω_p is the centre frequency of (4.4); let the bandwidth be the function $\Delta\omega(\sigma_p)$ of σ_p, and let the centre of the magnitude response be the function $\mathcal{C}(\omega_p)$ of ω_p. Figure (4.1) shows the shape of the magnitude response and the impulse response together with the relations to the two functions just defined. Note that the illustrated bandwidth should be taken conceptually.



(a) *Impulse Response* **(b)** *Magnitude Response*

Figure 4.1: *(a) shows the real, and imaginary (stipulated) parts of $h(t)$. The complex modulus (black) shows the window shape. (b) Magnitude Response for the same fixed σ_p and ω_p.*

4.3 Generalization

The second order transfer function (4.4) has two repeated poles, so its magnitude response is the product of two single one-pole transfer functions. By using k repeated poles, the magnitude response is generally raised to the kth power. The general kth order transfer function $H^k(s)$ is thus

$$H^k(s) = \frac{\sigma_p^k}{(s+p^\star)^k}. \tag{4.8}$$

This means that the corresponding impulse response of (4.8) consists of k repeated convolutions of first order impulse responses. As a consequence of the central limit theorem, repeated convolutions of a distribution converges to the Gauss distribution. This is especially useful because – in accordance with Gabor's result – the Gauss distribution has the optimal time-frequency resolution. Introducing the parameter k is therefore equivalent to introducing a optional time-frequency resolution up to optimality.

Thel ordinary differential equation describing the generalized recursive system is given by

$$\sum_{j=0}^{k} \binom{k}{j} (\sigma_p - i\omega_p)^j \frac{\partial^{k-j}}{\partial t^{k-j}} y(t,\omega_p) = \sigma_p^k x(t), \tag{4.9}$$

and the corresponding impulse response is given by

$$h^k(t) = \mathcal{L}^{-1}\{H^k(s)\} = \sigma_p^k \frac{t^{k-1}}{(k-1)!} e^{-p^\star t} u(t). \tag{4.10}$$

4.3.1 The Generalized Window Function

Considering the generalized kth order impulse response (4.10), the generalized kth order window function is found to be

$$\mathcal{W}^k(t) = \sigma_p^k \frac{t^{k-1}}{(k-1)!} e^{-\sigma_p t} u(t). \tag{4.11}$$

Even if the extent of (4.11) is infinite for positive t, it decays to zero since exponential growth always overtakes polynomial growth. It is thus stable for all orders k. However, convolutions with an infinite window must yield an infinite result. In the discrete case this is readily omitted by cutting the convolution after a finite number of time steps. The cutting is feasible since the significant contribution from the window's extent is concentrated around relatively small t.

4.4 General Solution of the kth Order Ordinary Differential Equation

The solution of the kth order differential equation in (4.9) is a system consisting of one linear convolution between $x(t)$ and (4.10) per ω_p. Equation (4.10) can be parametrized with ω_p so that the solution can be written

$$y^k(t, \omega_p) = \int_{-\infty}^{\infty} x(\tau) h^k(t-\tau, \omega_p) d\tau \tag{4.12}$$

$$= \int_{-\infty}^{\infty} x(\tau) \mathcal{W}^k(t-\tau) e^{i\omega_p(t-\tau)} d\tau \tag{4.13}$$

$$= e^{i\omega_p t} \int_{-\infty}^{\infty} x(\tau) \mathcal{W}^k(t-\tau) e^{-i\omega_p \tau} d\tau \tag{4.14}$$

The recursive system is thus equivalent with the STFT via the infinite window function in (4.11). The power of the recursive approach – seen from a computational point of view – comes from that (4.14) is equivalent to (4.9), and that the discretized version of (4.9) can be computed efficiently in terms of linear implicit recursions.

4.5 Time-Frequency Resolution

Applying the classical definitions for bandwidth (1.1) and duration (1.2) to (4.10) or (4.8) yields no time-frequency resolution at all. This is because the second moment of $|H^k(s)|^2|_{s=i\omega}$ does not converge. Appropriate definitions were proposed by [35]. The bandwidth is given as the duration of a rectangular window with height equal to $max\{|H^k(s)|^2|_{s=i\omega}\}$ and energy equal to the energy of $H(s)|_{s=i\omega}$

$$\Delta_\omega = \frac{1}{\max\{|H^k(s)|^2|_{s=i\omega}\}} \int_{-\infty}^{\infty} |H^k(s)|^2 ds|_{s=i\omega}, \tag{4.15}$$

and equivalently for Δ_t

$$\Delta_t = \frac{1}{\max\{|h^k(t)|^2\}} \int_{-\infty}^{\infty} |h^k(t)|^2 dt. \tag{4.16}$$

Since $\max\{|H^k(s)|^2|_{s=i\omega}\} = 1$, and as using Parseval's formula yields the integral in (4.15) to be written in terms of (4.16), the bandwidth can be written

$$\Delta_\omega = 2\pi \int_{-\infty}^{\infty} |h^k(t)|^2 dt = \frac{2\pi\sigma_p \left[2(k-1)\right]!}{\left[(k-1)!\right]^2 2^{2k-1}}. \tag{4.17}$$

By differentiating $|h^k(t)|^2$ and setting it equal to zero yields $\max\{|h^k(t)|^2\}$

$$\frac{\partial}{\partial t}|h^k(t)|^2 = 0 \Leftrightarrow t = \frac{k-1}{\sigma_p} \Rightarrow \max\{|h^k(t)|^2\} = \left[\frac{(k-1)^{k-1}\sigma_p e^{-(k-1)}}{(k-1)!}\right]^2. \tag{4.18}$$

The duration is finally found to be

$$\Delta_t = \frac{\left[2(k-1)\right]!}{\sigma_p \left(k-1\right)^{2(k-1)} e^{-2(k-1)} 2^{2k-1}}. \tag{4.19}$$

Using $k = 1$ yields the time-frequency uncertainty $\Delta_t\Delta\omega = \pi/2$, while using the optimum Gaussian window would yield $\Delta_t\Delta\omega = \pi$. As the number of repeated poles in (4.8) increases, the time-frequency resolution approaches the optimum. The convergence is fast enough to yield $\Delta_t\Delta_\omega = 3.077$ when $k = 5$.

4.5.1 Optimum Time-Frequency Resolution Balance

The bandwidth parameter σ_p is cancelled in $\Delta_t\Delta_\omega$ and has therefore nothing to do with the overall time-frequency uncertainty nor resolution. It has, however, the purpose of balancing the trade-off between resolution in time and resolution frequency. Setting $\Delta_t = \Delta_\omega$ and solving for σ_p yields

$$\sigma_p = \frac{(k-1)!}{\sqrt{2\pi}(k-1)^{(k-1)}e^{1-k}}. \tag{4.20}$$

Using this value for σ_p can therefore therefore be regarded as the optimal choice given that resolution is equally important in both domains.

4.6 The Output of the Recursive System is Analytic

The system in (4.9) clearly has analytic output, since it is a special case of the STFT. However, this can also be shown mathematically. Rewriting (4.12) slightly yields

$$\begin{aligned} y^k(t,\omega_p) &= \frac{\sigma_p^k}{(k-1)!}\int_{-\infty}^{\infty}(t-\tau)^{k-1}e^{-\sigma_p(t-\tau)}\cos\left[\omega_p(t-\tau)\right]x(\tau)d\tau \qquad (4.21)\\ &+ \frac{i\sigma_p^k}{(k-1)!}\int_{-\infty}^{\infty}(t-\tau)^{k-1}e^{-\sigma_p(t-\tau)}\sin\left[\omega_p(t-\tau)\right]x(\tau)d\tau \\ &= X_R^k(t,\omega_p) + iX_I^k(t,\omega_p), \qquad (4.22)\end{aligned}$$

and using that

$$\mathcal{H}\{\cos\left[\omega_p(t-\tau)\right]\} = \sin\left[\omega_p(t-\tau)\right], \tag{4.23}$$

equation (4.21) can be written

$$y^k(t,\omega_p) = \frac{\sigma_p^k}{(k-1)!}\int_{-\infty}^{\infty}(t-\tau)^{k-1}e^{-\sigma_p(t-\tau)}\cos\left[\omega_p(t-\tau)\right]x(\tau)d\tau \tag{4.24}$$

$$+\frac{i\sigma_p^k}{(k-1)!}\int_{-\infty}^{\infty}(t-\tau)^{k-1}e^{-\sigma_p(t-\tau)}\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\cos\left[\omega_p(t-\tau)\right]}{t-\tau'}d\tau' x(\tau)d\tau$$

$$=\frac{\sigma_p^k}{(k-1)!}\int_{-\infty}^{\infty}(t-\tau)^{k-1}e^{-\sigma_p(t-\tau)}\cos\left[\omega_p(t-\tau)\right]x(\tau)d\tau \tag{4.25}$$

$$+\frac{i}{\pi}\int_{-\infty}^{\infty}\frac{i\sigma_p^k}{(k-1)!}\int_{-\infty}^{\infty}\frac{(t-\tau)^{k-1}e^{-\sigma_p(t-\tau)}\cos\left[\omega_p(t-\tau)\right]x(\tau)}{t-\tau'}d\tau d\tau'$$

$$=X_R^k(t,\omega_p)+i\mathcal{H}\{X_R^k(t,\omega_p)\} \tag{4.26}$$

The output of the kth order recursive system defined by (4.9) is then, in accordance with definition (1.6), analytic.

4.7 Multiresolution & Frequency Scale

The system described above is easily made multiresolutional by ensuring that

$$Q=\frac{\mathcal{C}(\omega_p)}{\Delta\omega(\sigma_p)} \tag{4.27}$$

is constant. That is, by choosing $\Delta\omega(\sigma_p)$ proportional to $\mathcal{C}(\omega_p)$. To have a linear frequency scale on $y^k(t,\omega_p)$, one can set

$$\mathcal{C}(\omega_p)=\omega_p. \tag{4.28}$$

A logarithmic frequency scale can also be used. For instance, to synchronize with the well-tempered scale (12-sub-octaves) one may simply choose

$$\mathcal{C}(\omega_p)=f_0 2^{\omega_p/12} \tag{4.29}$$

with f_0 being the lowest frequency.

4.8 The TFR

The output from the system in (4.9), is as shown a two-dimensional complex analytic function of time and frequency. In analogy with the STFT spectrogram, one may take the complex modulus squared to compute the non-oscillatory envelope's energy

$$|y^k(t,\omega_p)|^2. \tag{4.30}$$

The envelope effect was illustrated in Chapter 2. The envelope is, in a sense, an anti-aliasing filter, as it in the discrete case yields a smooth surface well suited for visualization (or direct downsampling). The TFR (4.30) is hereby reffered to as the Recursive TFR (RTFR).

4.9 Non-linear Phase

It is well-known that recursive systems have non-linear phase responses. This is, however, not a serious problem, because the group delay of (4.8) evaluated on the centre frequency $\omega = \omega_p$ is equal for all ω_p. This is due to that a change of the centre frequency ω_p is just a translation along the imaginary axis in (4.8). Thus, there is no skewness in the time delay per horizontal "line" in the RTFRs.

The group delay is directly linked to the system order k, meaning that the higher the system order, the more delay is introduced. It will be seen that the recursive reassignment method presented in Chapter 6, will eliminate this time delay completely.

4.9.1 Time-Domain Signal Shape Invariance

The main shortcoming of the RTFRs is their lack of time-domain signal shape invariance. As a consequence of the non-linear phase, the group delay within the pass-band of the filters is non-linear. This will cause a distortion in the time-domain waveforms observable in the RTFRs. As the group delay will be maximal at $\omega = \omega_p$, it is expected that the time delay in the RTFRs will be slightly more prominent on the centre frequency for any component. Nevertheless, a consequence of the central limit theorem is that when the order k of the system increases, the group delay in the pass-band is linearized around ω_p. This diminish the time-domain shape invariance when the system order increases. If the TFA application, however, is especially sensitive for the preservation of time-domain signal shapes, one solution is to move to the bidirectional RTFRs outlined in the end of Chapter 6.

4.10 Discretization of the Recursive Linear System

There are several ways of discretizing the transfer function described in (4.8). Since it can be characterized as a bandpass filter without any undesired high frequency components, one can put a requirement on the sampling rate, and then take samples directly. This type of discretization is known as the impulse invariance method.

Substituting $t = nT$ in (4.10), where n is the time step, and $T = 1/f_s$ is the sampling period and f_s is the sampling frequency, yields

$$h^k(nT) = \sigma_p^k \frac{(nT)^{k-1}}{(k-1)!} e^{-pnT} u(nT). \tag{4.31}$$

By now, requiring $f_s > 2 \max\{\mathcal{C}(\omega_p)\}$, no aliasing, in accordance with Nyquist's sampling theorem, will occur.

Applying the z-transform on (4.31) gives

$$H^k(z) = \mathcal{Z}\{h^k(nT)\} = \sigma_p^k \sum_{n=0}^{\infty} \frac{(nT)^{k-1}}{(k-1)!} e^{-p^\star nT} z^{-n} \tag{4.32}$$

$$= \sigma_p^k \frac{T^{k-1}}{(k-1)!} \sum_{n=0}^{\infty} n^{k-1} e^{-p^\star nT} z^{-n} \tag{4.33}$$

$$= \sigma_p^k \frac{T^{k-1}}{(k-1)!} \overbrace{-z\frac{\partial}{\partial z}}^{k-1 \text{ times}} \frac{1}{1 - e^{-p^\star T} z^{-1}}, \tag{4.34}$$

With an extra scaling factor T for correct normalization, this can be written as

$$H^k(z) = \sigma_p^k T^k \frac{\sum_{j=1}^{k} A_{k-1,k-j} (e^{-p^\star T} z^{-1})^{j-1} + \delta(k-1)}{(k-1)!(1 - e^{-p^\star T} z^{-1})^k}, \tag{4.35}$$

where $\delta(\cdot)$ represents the Kronecker delta function, and with $A_{n,k}$ representing the Eulerian numbers given by

$$A_{k,n} = \sum_{j=0}^{n} (-1)^j \binom{k+1}{j} (n+1-j)^k. \tag{4.36}$$

Table (4.1) lists the digital filter coefficients for orders $k = 1 \ldots 5$.

4.11 Linear Implicit Recursions – the Difference Equation

Consider the case where $k = 2$. The second order transfer function $H^2(z)$ is thus

$$H^2(z) = \sigma_p^2 T^2 \frac{e^{-p^\star T} z^{-1}}{(1 - e^{-p^\star T} z^{-1})^2} = \frac{Y(z)}{X(z)}. \tag{4.37}$$

This corresponds to a difference equation in the time-domain

$$\begin{gathered} Y(z)(1 - 2e^{-p^\star T} z^{-1} + e^{-2p^\star T} z^{-2}) = \sigma_p^2 T^2 e^{-p^\star T} X(z) z^{-1} \\ \Updownarrow \\ y(n, \omega_p) = 2e^{-(\sigma_p - i\omega_p)T} y(n-1) - e^{-2(\sigma_p - i\omega_p)T} y(n-2) + \sigma_p^2 T^2 e^{-(\sigma_p - i\omega_p)T} x(n-1), \end{gathered} \tag{4.38}$$

with all initial conditions equal to zero. Equation (4.38) is now fully equivalent with the continuous system in (4.3). For more details regarding implementation issues, see Chapter 7.

4.12 Optimal Time-Frequency Resolution Balance in the Discrete Case

To find the optimal time-frequency resolution balance in the discrete case, the definitions of bandwidth and duration also need to be discretized. Sampling the

k	b_n, a_n
1	$b_0 = \sigma_p T,$ $a_0 = 1, a_1 = -e^{-p^\star T}$
2	$b_0 = 0, b_1 = \sigma_p^2 T^2 e^{-p^\star T},$ $a_0 = 1, a_1 = -2e^{-p^\star T}, a_2 = e^{-2p^\star T}$
3	$b_0 = 0, b_1 = \sigma_p^3 \frac{T^3}{2} e^{-p^\star T}, b_2 = \sigma_p^3 \frac{T^3}{2} e^{-2p^\star T},$ $a_0 = 1, a_1 = -3e^{-p^\star T}, a_2 = 3e^{-2p^\star T}, a_3 = -e^{-3p^\star T}$
4	$b_0 = 0, b_1 = \sigma_p^4 \frac{T^4}{6} e^{-p^\star T}, b_2 = \sigma_p^4 \frac{2T^4}{3} e^{-2p^\star T}, b_3 = \sigma_p^4 \frac{T^4}{6} e^{-3p^\star T},$ $a_0 = 1, a_1 = -4e^{-p^\star T}, a_2 = 6e^{-2p^\star T}, a_3 = -4e^{-3p^\star T}, a_4 = e^{-4p^\star T}$
5	$b_0 = 0, b_1 = \sigma_p^5 \frac{T^5}{24} e^{-p^\star T}, b_2 = \sigma_p^5 \frac{11T^5}{24} e^{-2p^\star T}, b_3 = \sigma_p^5 \frac{11T^5}{24} e^{-3p^\star T}, b_4 = \sigma_p^5 \frac{T^5}{24} e^{-4p^\star T},$ $a_0 = 1, a_1 = -5e^{-p^\star T}, a_2 = 10e^{-2p^\star T}, a_3 = -10e^{-3p^\star T},$ $a_4 = 5e^{-4p^\star T}, a_5 = -e^{-5p^\star T}$

Table 4.1: *A summary of the digital filter coefficients for $k = 1...5$.*

integrals in (4.16) and (4.15) with $t = nT$, the optimal σ_p in the discrete case is found to be

$$\sigma_p = \frac{(k-1)!}{\sqrt{2\pi T}(k-1)^{(k-1)} e^{1-k}}. \tag{4.39}$$

This is worth emphasizing, because in the classical STFT, the time-frequency resolution balance is only affectable by the chosen number of samples in the window. On the other hand, here the number of samples in the window will implicitly be infinite, so the time-frequency resolution balance can only be changed by σ_p.

Chapter 5

The Time-Frequency Reassignment Method

5.1 Motivation

As pointed out in Chapter 2, the fundamental limitation of TFRs is Heisenberg's (Gabor's) uncertainty principle [13]. If the bandwidth of (4.8) is chosen too wide, the duration of the impulse response (4.10) becomes too narrow. Vice versa, if the bandwidth is chosen too narrow, then the duration of the impulse response becomes too wide. This affects the uncertainty in both domains: higher accuracy in one will yield lower accuracy in the other. However, in the middle there must be a point where the bandwidth and the duration are equal, so the uncertainty in both domains are equally bad or equally good. The value for σ_p serving this property was given in equation (4.39). The uncertainty is, however, independent of σ_p, and it was shown that the uncertainty is not minimal for the recursive filters used. In accordance with Gabor's results it is necessary to use a Gaussian for minimal uncertainty. This is not possible within any recursive system. The reason is that the window in (4.11), or any other window stemming from a recursive system, cannot be true Gaussian because rational transfer functions never have perfectly symmetric impulse responses.

Because of the uncertainty principle, it is seemingly impossible for a TFR to have higher resolution than the one achieved by using the Gaussian window. This is, however, not the case. The crucial point is that the uncertainty principle also applies for the analyzed signal. This means that it is not necessary for the window to be infinitesimal in both domains in order to lower the uncertainty. If the bandwidth or duration adaptively depends on the analyzed signal's duration or bandwidth, the uncertainty will be zero. This resembles the WVD, and, it is the reason why the WVD provides maximal concentration in the time-frequency plane. However, the cross-term artefacts discussed in Chapter 3 makes even the WVD less attractive when the analyzed signal contains more than one component on a *global* level.

One related technique used to cancel the uncertainty is time-frequency reassignment. This method is also adaptive, but the adaption is not from using a window with variable bandwidth or duration. It is rather from exploiting that it is possible to calculate the offset positions of the signal segments captured by the windows. This information can be used to adjust the precision in the TFR such that the uncertainty again will be zero. However, if this was the end of the

story, the uncertainty principle would be false. The limitation of the reassignment method is simply that only one significant signal component can fall inside *one single window*. The attractiveness of the reassignment method comes from that this only pose a limitation on a *local* level.

5.2 Introduction

The train of thought which ends with the reassignment principle, starts with the Rihaczek Distribution introduced by [26]. The following will use the spectrogram as a starting point, this despite that the theory can be generalized to arbitrary TFRs by using the formulation of [3].

The Rihaczek Distribution [26] of an analytic signal $x(t)$ is defined

$$E_x(t,\omega) = x(t)X^\star(\omega)e^{-i\omega t}, \tag{5.1}$$

where $X(\omega)$ is the Fourier transform of $x(t)$, and $\star$ denotes complex conjugation. This TFR is today known to be a member of Cohen's class with the kernel function $\Phi(\theta,\tau) = e^{i\tau\theta/2}$. Since all the members of Cohen's class are smoothed versions of the WVD, it follows that the (5.1) also is a smoothed version of the WVD.

Just before the time when the reassignment method was invented, it was shown [1] that the spectrogram as defined in (3.4), can be written as

$$|X(t_0,\omega_0)|^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E_x(t,\omega)E_{\mathcal{W}}(t_0-t,\omega-\omega_0)dtd\omega, \tag{5.2}$$

where E_x and $E_{\mathcal{W}}$ are the Rihaczek Distributions of the input signal $x(t)$ and the window function $\mathcal{W}(t)$, respectively.

One single point's value in the spectrogram is therefore essentially found by evaluating the two-dimensional convolution between E_x and $E_{\mathcal{W}}$. Since the Rihaczek Distribution contains cross-terms and is complex valued, it is difficult to interpret. However, since the window has low-pass characteristics, it follows that (5.2) is a smoothing operation so that the cross-terms are traded with a blurry representation. This can be seen as the main reason why the spectrogram does not sharply reproduce the distribution of the original spectrum.

Equation (5.2) also reveals that one cannot know which part of the integrand that contributes mostly to the integration. This insight, can, as the next section shows, be exploited to define a new spectrogram where each point is moved to new coordinates better fitting with the energy density of the input signal.

5.3 The Reassignment

Kodera et al. [21] proposed a clever procedure reversing the effect of the smoothing in (5.2) without re-introducing cross-terms. Firstly, for each point (t_0,ω_0) in the spectrogram, define the set $\mathcal{N}\{(t_0,\omega_0)\}$ consisting of all the points in the close vicinity of (t_0,ω_0). Secondly, for each of these sets, define the maximum intensity point $(t_0,\omega_0)_{\max}$. Thirdly, and finally, for each of these sets, reassign and accumulate all the points onto the maximum intensity point $(t_0,\omega_0)_{\max}$. This will yield a sharpening effect, counteracting the smoothing from the windowing.

As each point (t_0, ω_0) in the spectrogram is the result of the two-dimensional convolution (5.2), it means that each point is a weighted sum of the Rihaczek Distribution of the input signal inside a region centred at (t_0, ω_0) with radius approximately equal to the time-frequency extent of the window. That is, the region defined by the window's bandwidth, Δ_ω, and duration, Δ_t. In other words, the time-frequency range of the neighbourhood function $\mathcal{N}\{\cdot\}$ must correspond to the time-frequency extent of the window. In Figure (5.1) this concept is illustrated using a Gabor logon. Figure (5.2) shows how the same logon is affected by the reassignment.

The process just described could simply be implemented by a brute force strategy, but it turns out that there is a far more sophisticated and efficient way of doing this.

To find where the contribution to the integral in (5.2) is locally maximal around the point (t_0, ω_0), one can start by considering the time and frequency variables separately. Note that the imaginary part of (5.2) is zero. In the local domain defined by the bandwidth of the window, the curve through the spectrogram in the direction of frequency, around the point (t_0, ω_0) can be written as

$$\eta(\omega, t_0, \omega_0) = \int_{-\infty}^{\infty} \mathrm{Re}\{E_x(t, \omega) E_{\mathcal{W}}(t_0 - t, \omega - \omega_0)\} dt. \tag{5.3}$$

By taking the normalized first moment of (5.3), one gets the point in frequency which can be seen as the mass center with respect to signal energy inside the local frequency domain of the window

$$\hat{\omega_i}(t_0, \omega_0) = \frac{\int \omega \eta(\omega, t_0, \omega_0) d\omega}{\int \eta(\omega, t_0, \omega_0) d\omega}. \tag{5.4}$$

Note that the local frequency-domain of the window corresponds to its bandwidth. This explains the quantity Δ_ω in Figure (5.1).

Similarly, in the local domain defined by the duration of the window, the curve through the spectrogram in the direction of time, around the point (t_0, ω_0) can be written as

$$\xi(t, t_0, \omega_0) = \int_{-\infty}^{\infty} \mathrm{Re}\{E_x(t, \omega) E_{\mathcal{W}}(t_0 - t, \omega - \omega_0), \} d\omega. \tag{5.5}$$

By taking the first moment,

$$\hat{t_g}(t_0, \omega_0) = \frac{\int t \xi(t, t_0, \omega_0) dt}{\int \xi(t, t_0, \omega_0) dt}, \tag{5.6}$$

one gets the point in time which can be seen as the mass center with respect to signal energy inside the local time-domain of the window. The local time-domain of the window corresponds to its duration. This quantity is also illustrated in Figure (5.2). Together, (5.4) and (5.6) can be used as new coordinates for each point in (5.2) such that the magnitude better fits with the input signal's true energy.

In other words, it makes sense to define a new reassigned spectrogram, $X_R(t_0, \omega_0)$, where each point (t_0, ω_0) is the sum of all points in the regular spectrogram with

new coordinates (t_0, ω_0). Mathematically, this can be written as

$$X_R(t_0,\omega_0) = \int\int |X(t',\omega')|^2 \delta\left[t_0 - \hat{t_g}(t',\omega')\right] \delta\left[\omega_0 - \hat{\omega_i}(t',\omega')\right] dt' d\omega', \quad (5.7)$$

where $\delta[\cdot]$ represents the Kronecker delta function.

5.4 The Reassignment Coordinates and their Relation to Phase

Kodera et al. [21] showed that by considering the output of the STFT (4.1) in polar coordinates,

$$X(t_0,\omega_0) = a(t_0,\omega_0)e^{i\phi(t_0,\omega_0)}, \quad (5.8)$$

the coordinates defined by (5.4) and (5.6) can be written in terms of the partial derivatives of the phase $\phi(t_0, \omega_0)$. This result is re-derived here for completeness. Note that in the following equations, the window and its Fourier transform is only distinguishable by their running variable t or ω.

The frequency reassignment coordinate in (5.4) can be found as follows

$$\begin{aligned}
\hat{\omega_i}(t_0,\omega_0) &= \mathrm{Re}\left[\frac{\int \omega \int E_x(t,\omega)E_{\mathcal{W}}(t_0-t,\omega-\omega_0)dtd\omega}{\int\int E_x(t,\omega)E_{\mathcal{W}}(t_0-t,\omega-\omega_0)dtd\omega}\right] && (5.9)\\
&= \mathrm{Re}\left[\frac{\int \omega \int E_x(t,\omega)E_{\mathcal{W}}(t_0-t,\omega-\omega_0)dtd\omega}{X(t_0,\omega_0)X^\star(t_0,\omega_0)}\right] && (5.10)\\
&= \mathrm{Re}\left[\frac{\int \omega e^{-i\omega t_0}X^\star(\omega)\mathcal{W}^\star(\omega-\omega_0)d\omega}{X^\star(t_0,\omega_0)}\right] && (5.11)\\
&= \mathrm{Re}\left[-\frac{1}{iX^\star(t_0,\omega_0)}\frac{\partial}{\partial\omega_0}X^\star(t_0,\omega_0)\right] && (5.12)\\
&= \mathrm{Re}\left[-\frac{1}{ia^\star(t_0,\omega_0)e^{-i\phi(t_0,\omega_0)}}\frac{\partial}{\partial\omega_0}a^\star(t_0,\omega_0)e^{-i\phi(t_0,\omega_0)}\right] && (5.13)\\
&= \frac{\partial}{\partial t_0}\phi(t_0,\omega_0). && (5.14)
\end{aligned}$$

The time reassignment coordinate in (5.6) can likewise be found

$$\begin{aligned}
\hat{t_g}(t_0,\omega_0) &= \mathrm{Re}\left[\frac{\int t \int E_x(t,\omega)E_{\mathcal{W}}(t_0-t,\omega-\omega_0)d\omega dt}{\int\int E_x(t,\omega)E_{\mathcal{W}}(t_0-t,\omega-\omega_0)d\omega dt}\right] && (5.15)\\
&= \mathrm{Re}\left[\frac{\int t \int E_x(t,\omega)E_{\mathcal{W}}(t_0-t,\omega-\omega_0)d\omega dt}{X(t_0,\omega_0)X^\star(t_0,\omega_0)}\right] && (5.16)\\
&= \mathrm{Re}\left[\frac{\int te^{i\omega_0(t_0-t)}x(t)\mathcal{W}(t_0-t)dt}{X(t_0,\omega_0)}\right] && (5.17)\\
&= \mathrm{Re}\left[t_0 - \frac{1}{iX(t_0,\omega_0)}\frac{\partial}{\partial t_0}X(t_0,\omega_0)\right] && (5.18)\\
&= \mathrm{Re}\left[t_0 - \frac{1}{ia(t_0,\omega_0)e^{i\phi(t_0,\omega_0)}}\frac{\partial}{\partial t_0}a(t_0,\omega_0)e^{i\phi(t_0,\omega_0)}\right] && (5.19)\\
&= t_0 - \frac{\partial}{\partial\omega_0}\phi(t_0,\omega_0). && (5.20)
\end{aligned}$$

The two equations (5.14) and (5.20) resembles the definitions of instantaneous frequency (1.7) and group delay (1.8), but since the phase now is a two-dimensional function of both time and frequency, and in that sense "local", [23] decided to name (5.14) and (5.20) channelized instantaneous frequency and local group delay, respectively.

In the introduction it was mentioned that also the reassignment method has its limitation. Recall that in the definition of instantaneous frequency (1.7), it is presumed that the signal only has one (frequency) component at any time instant. As a consequence of (5.14), the frequency reassignment is only meaningful with the same presumption. It is therefore expected that the frequency reassignment coordinates will behave chaotic if the (frequency) window is not able to separate close frequency components. The same statements are also true for the time reassignment, because (5.20) involves (1.8), and the latter has just the same limitation.

5.5 Theoretical Efficiency

The theoretical efficiency of the reassignment method was not addressed by its inventors [21]. Auger and Flandrin [3] much later showed that the method yields TFRs with perfect localization of chirps and impulses. For completeness, this is re-derived from the results in [3] below.

5.5.1 Perfect Localization of Impulses & Chirps

If the input signal is defined as an impulse at time instant t_1

$$x(t) = A\delta(t - t_1), \tag{5.21}$$

then the STFT defined by (4.1) is

$$\begin{aligned} X(t,\omega) &= e^{i\omega t}\int_{-\infty}^{\infty} x(\tau)\mathcal{W}(t-\tau)e^{-i\omega\tau}d\tau \\ &= Ae^{i\omega t}\int \delta(\tau - t_1)\mathcal{W}(\tau - t)e^{-i\omega\tau}d\tau \\ &= Ae^{i\omega t}\mathcal{W}(t_1 - t)e^{-i\omega t_1} \\ &= Ae^{-i\omega(t_1-t)}\mathcal{W}(t_1 - t). \end{aligned} \tag{5.22}$$

The spectrogram becomes

$$|X(t,\omega)|^2 = A^2\mathcal{W}^2(t_1 - t). \tag{5.23}$$

The spectrogram is clearly smoothed by the window, because if we perturb t with an amount smaller than half the duration of $\mathcal{W}(t)$, then there will still be a contribution. However, the frequency reassignment coordinate is

$$\frac{\partial}{\partial t}\arg\{X(t,\omega)\} = \frac{\partial}{\partial t} - \omega(t_1 - t) = \omega, \tag{5.24}$$

and the time reassignment coordinate is

$$t - \frac{\partial}{\partial \omega}\arg\{X(t,\omega)\} = t - \frac{\partial}{\partial \omega} - \omega(t_1 - t) = t_1. \tag{5.25}$$

The reassigned spectrogram is finally found to be found

$$X_R(t_0,\omega_0) = \int\int |X(t,\omega)|^2 \delta\left[t_0 - \hat{t_g}(t,\omega)\right] \delta\left[\omega_0 - \hat{\omega_i}(t,\omega)\right] dtd\omega \tag{5.26}$$

$$= A^2 \int\int \mathcal{W}^2(t_1 - t)\delta\left[t_0 - t_1\right] \delta\left[\omega_0 - \omega\right] dtd\omega \tag{5.27}$$

$$= A^2 \left[\int\int \mathcal{W}^2(t_1 - t)\delta\left[\omega_0 - \omega\right] dtd\omega\right] \delta\left[t_0 - t_1\right]. \tag{5.28}$$

The reassigned spectrogram is now independent of the window. A perturbation on t_0 yields no contribution because of the infinitesimal narrow delta function.

Likewise, if the input signal is defined as a chirp

$$x(t) = Ae^{i(\alpha/2t^2 + \omega_1 t)}, \tag{5.29}$$

then the reassigned STFT spectrogram is found to be

$$X_R(t_0,\omega_0) = 2\pi A^2 \left[\int\int |X(t,\omega)|^2 \delta\left[t_0 - t_g(\hat{t},\omega)\right] dtd\omega\right] \delta\left[\omega_0 - \omega_1 - \alpha t_0\right]. \tag{5.30}$$

5.6 Computing the Partial Derivatives of the Phase

Three main methods exist for computing the partial derivatives of the phase $\phi(t,\omega)$. This section introduce all three, and their shortcomings will be discussed.

5.6.1 The Finite Difference Method

The method used by the pioneer of reassignment, [21], was simply a first order finite difference approximation to the derivative operators in (5.14) and (5.20), that is, for (5.14),

$$\frac{\partial}{\partial t_0}\arg\{X(t_0,\omega_0)\} \approx \arg\{X(t_0+1,\omega_0)\} - \arg\{X(t_0,\omega_0)\} \tag{5.31}$$

$$\approx \phi(t_0+1,\omega_0) - \phi(t_0,\omega_0) \tag{5.32}$$

$$= \Delta_{t_0}\phi(t_0,\omega_0). \tag{5.33}$$

This method has two severe drawbacks. First, note that the phase must be continuous to be differentiable. When using inverse tangent routines to compute the argument function, one must expect 2π jump discontinuities. Hence, the phase, using this method, must be unwrapped by appropriately adding multiples of 2π to the principal value until the discontinuities are removed. This unwrapping is not a very reliable operation [23]. See for instance [37] for a constructed situation (often arising in practice) in which it fails. Second, finite differences are approximations, so, at least in theory, it can be done better. Finite differences are easily seen unreliable when the analyzed signal contains high frequencies relative to the sampling frequency. By taking the z-transform of (5.33) with respect

to t_0, one gets

$$\begin{aligned}\mathcal{Z}\{\Delta_{t_0}\phi(t_0,\omega_0)\} &= \mathcal{Z}\{\phi(t_0+1,\omega_0)-\phi(t_0,\omega_0)\} && (5.34)\\ &= \phi(z,\omega_0)z-\phi(z,\omega_0) && (5.35)\\ &= \phi(z,\omega_0)(z-1)=\phi(z,\omega_0)H_D(z), && (5.36)\\ & && (5.37)\end{aligned}$$

which means that the filter acting as the differentiator has the transfer function given by

$$H_D(z) = z - 1. \tag{5.38}$$

Taking the complex modulus when z is evaluated on the unit circle yields the magnitude response of (5.38)

$$\begin{aligned}H_D(\omega) &= e^{i\omega T}-1 && (5.39)\\ &= e^{\frac{i\omega T}{2}}(e^{\frac{i\omega T}{2}}-e^{\frac{-i\omega T}{2}}) && (5.40)\\ |H_D(\omega)| = |e^{\frac{i\omega T}{2}}||(e^{\frac{i\omega T}{2}}-e^{\frac{-i\omega T}{2}})| &\qquad = 2|\sin(\frac{\omega T}{2})|, \omega\in[-\pi,\pi] && (5.41)\end{aligned}$$

On the other hand, the ideal differentiator is found by taking the Fourier transform of the differentiation operator applied to $\phi(t,\omega_0)$

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{\partial}{\partial t}\phi(t,\omega_0)e^{-i\omega t}dt &= -i\omega\Phi(\omega,\omega_0) && (5.42)\\ &= \Phi(\omega,\omega_0)H_I(\omega), && (5.43)\end{aligned}$$

meaning that the ideal differentiator filter is given by

$$H_I(\omega) = -i\omega. \tag{5.44}$$

Taking the complex modulus yields its magnitude response

$$|H_I(\omega)| = \omega, \omega\in[-\pi,\pi] \tag{5.45}$$

Comparing (5.41) to (5.45) shows, as asserted, that the finite difference is unreliable on high frequencies relative to the sampling frequency. See Figure (5.3) for a plot of the two.

5.6.2 The Cross Spectral Method

The second method is the one invented and used by [23]. It is called the cross spectral method, and it provides the solution to the first drawback in the previous method. The method is best explained by the following few lines

$$\begin{aligned}&\mathrm{Im}\{\log(a(t_0+1,\omega_0)e^{i\phi(t_0+1,\omega_0)}[a(t_0,\omega_0)e^{i\phi(t_0,\omega_0)}]^*)\} && (5.46)\\ &= \mathrm{Im}\{\log(a(t_0+1,\omega_0)+\log(e^{i\phi(t_0+1,\omega_0)})+\log(a^*(t_0,\omega_0))-\log(e^{i\phi(t_0,\omega_0)})\} && (5.47)\\ &= \mathrm{Im}\{\log(e^{i\phi(t_0+1,\omega_0)})-\log(e^{i\phi(t,\omega_0)})\} && (5.48)\\ &= \phi(t_0+1,\omega_0)-\phi(t_0,\omega_0) && (5.49)\\ &\approx \hat{w}_i(t_0,\omega_0) && (5.50)\end{aligned}$$

where $\star$ as usual denotes complex conjugation. Hence, the cross spectral method avoids the need for an argument function, and thereby also the need for phase-unwrapping. On the other hand, from a computational point of view, it introduces a costly logarithm function.

5.6.3 The Method of Auger and Flandrin

The third method was introduced by [3] only 12 years ago. It is general enough to work for all members of Cohen's class, even though its more computationally convenient for certain of them. In the spectrogram case, it is basically seen as two extra STFTs with two slightly modified window functions. It was shown that the new frequency coordinate (i.e. channelized instantaneous frequency) can be computed as

$$\hat{\omega_i}(t_0, \omega_0) = \omega_0 + \mathrm{Im}\left[\frac{X_{\mathcal{D}\mathcal{W}}(t_0, \omega_0) X^{\star}_{\mathcal{W}}(t_0, \omega_0)}{|X_{\mathcal{W}}(t_0, \omega_0)|^2}\right], \tag{5.51}$$

where the subscripts is used to distinguish the window functions. $X_{\mathcal{D}\mathcal{W}}$ is the STFT using the differentiated window function $\frac{\partial}{\partial t}\mathcal{W}(t)$, and $X_{\mathcal{W}}$ is the STFT using the unaltered window function $\mathcal{W}(t)$. Similarly, the new time coordinate (i.e. local group delay) can be computed by using with the STFT a time ramped window function, $t \times \mathcal{W}(t)$, denoted by $X_{\mathcal{T}\mathcal{W}}$,

$$\hat{t_g}(t_0, \omega_0) = t_0 - \mathrm{Re}\left[\frac{X_{\mathcal{T}\mathcal{W}}(t_0, \omega_0) X^{\star}_{\mathcal{W}}(t_0, \omega_0)}{|X_{\mathcal{W}}(t_0, \omega_0)|^2}\right]. \tag{5.52}$$

This superior method eliminates both drawbacks from the first method. It is not an approximation, nor does it require phase-unwrapping. It is also computationally fast since it only require two extra STFTs.

The next chapter pulls the pieces from the previous chapters together, and brings reassignment into the recursive system. It will be seen that the reassignment coordinates belonging to the recursive system can be computed equally fast as the RTFRs, and is thus in analogy with the method of Auger and Flandrin.

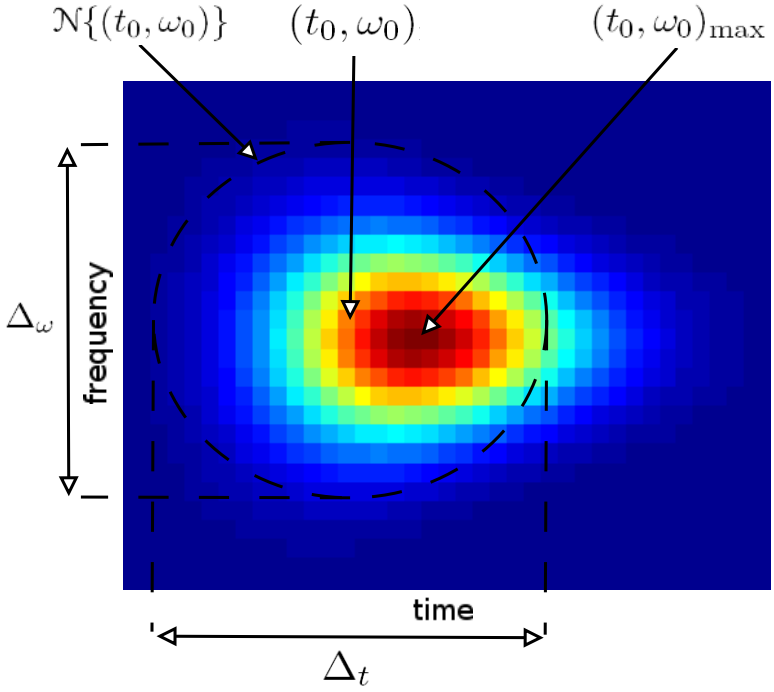


Figure 5.1: *Before reassignment is applied to the TFR: the stipulated circle's height represents the local frequency-domain of the window, i.e. its bandwidth Δ_ω. The width of the stipulated circle represents the local time-domain of the window, i.e. its duration Δ_t. The point $(t_0, \omega_0)_{max}$ is the maximum intensity point within the local domain described by Δ_ω and Δ_t. The point in the center of the window (t_0, ω_0) is to be reassigned onto the maximum intensity point $(t_0, \omega_0)_{max}$.*

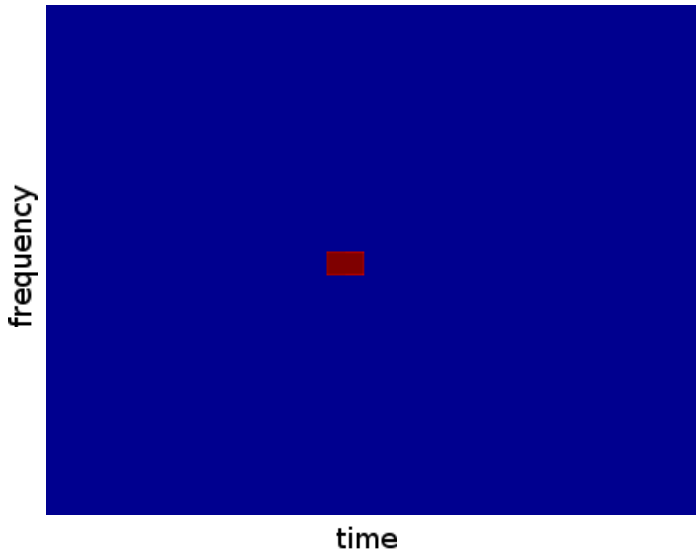


Figure 5.2: *After reassignment is applied to the TFR: all the points in the spectrogram have collapsed onto the logon's energy mass centre. Note that if there were two mass centres so close that one window could not separate them, the reassignment would not be meaningful. This situation can arise if there exists two or more maximum intensity points not being neighbours in the same local domain. Or equivalently, if two signal components are closer than Δ_t in time, or closer than Δ_ω in frequency.*

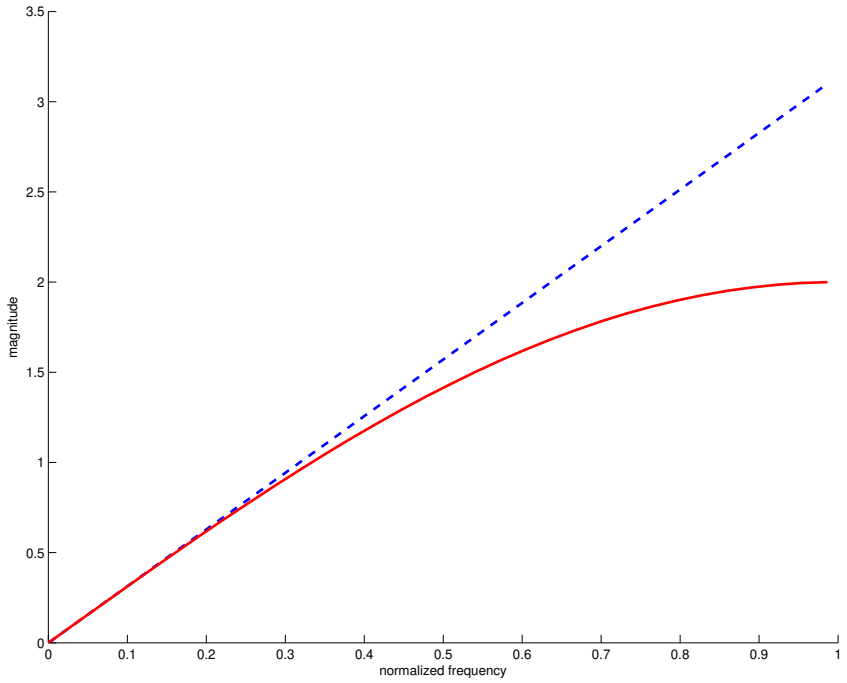


Figure 5.3: *The red curve represents the magnitude response of the finite difference differentiator $H_D(\omega)$, while the blue stipulated curve represents the ideal differentiator $H_I(\omega)$.*

Chapter 6

Recursive Time-Frequency Reassignment

6.1 Introduction

This chapter shows how the reassignment method, presented in Chapter 5, can be used within the recursive system presented in Chapter 4. It will be shown that the recursive system has corresponding recursive expressions yielding the time and frequency reassignment coordinates. As a consequence, the instantaneous frequency and group delay are found to be computed by linear implicit recursions as well. The new expressions involve three different transfer functions, where two of them are related to the other by simple operations.

6.2 Recursive Reassignment

The output from the recursive system defined by its transfer function in (4.8) can be written as in (4.12). It is repeated here for convenience, but notice that since the system output is analytic, it makes sense to consider the output in polar coordinates.

$$|y^k(t,\omega_p)|e^{i\phi^k(t,\omega_p)} = y^k(t,\omega_p) = \int_{-\infty}^{\infty} h^k(t-\tau,\omega_p)x(\tau)d\tau \tag{6.1}$$

The squared modulus of (6.1), in accordance with Rihazcek's theory [26], can be found by integrating over the energy densities of the analyzed signal $x(t)$ and the impulse response $h^k(t)$. The point is that the time and frequency reassignment coordinates also here will be the partial derivatives of the phase $\phi^k(t,\omega_p)$. This is easily verified by that (6.1) is nothing else than a special case of the STFT using the special window function given in (4.11).

6.3 The Partial Phase Derivatives in the Recursive System

The new frequency coordinate $\hat{\omega_i}(t, \omega_p)$ is thus to compute

$$\begin{aligned}
\hat{\omega_i}(t, \omega_p) &= \frac{\partial}{\partial t}\phi^k(t, \omega_p) && (6.2)\\
&= \frac{\partial}{\partial t}\arg\{y^k(t, \omega_p)\} && (6.3)\\
&= \frac{\partial}{\partial t}\mathrm{Im}\{\log\left(y^k(t, \omega_p)\right)\} && (6.4)\\
&= \mathrm{Im}\{\frac{\partial}{\partial t}\log\left(y^k(t, \omega_p)\right)\} && (6.5)\\
&= \mathrm{Im}\{\frac{\frac{\partial}{\partial t}y^k(t, \omega_p)}{y^k(t, \omega_p)}\}. && (6.6)
\end{aligned}$$

Similarly, the new time coordinate can be computed

$$\hat{t_g}(t, \omega_p) = t - \mathrm{Im}\{\frac{\frac{\partial}{\partial \omega_p}y^k(t, \omega_p)}{y^k(t, \omega_p)}\}. \quad (6.7)$$

Equation (6.6) shows that if the output of the recursive system can be differentiated with respect to time, then the new frequency coordinate can be found by taking the ratio between it and the unaltered output. Likewise, if the output is differentiated with respect to frequency, then the new time coordinate can essentially be found by an equivalent ratio. The next section shows that it is possible to find two new recursive systems implementing the derivative operators in equations (6.6) and (6.7).

6.3.1 The Transfer Function for the New Frequency Coordinate

Starting from (6.6), and substituting $y^k(t, \omega_p)$ by (6.1), one gets

$$\begin{aligned}
\hat{\omega_i}(t, \omega_p) &= \mathrm{Im}\{\frac{\frac{\partial}{\partial t}\int_{-\infty}^{\infty} h^k(t-\tau, \omega_p)x(\tau)d\tau}{y^k(t, \omega_p)}\}\\
&= \mathrm{Im}\{\frac{\int_{-\infty}^{\infty}\frac{\partial}{\partial t}h^k(t-\tau, \omega_p)x(\tau)d\tau}{y^k(t, \omega_p)}\}. && (6.8)
\end{aligned}$$

To take the derivative with respect to time of the impulse response, is in the Laplace domain equivalent to a multiplication by the Laplace variable s. Now one can define a new transfer function $H^k_{\mathcal{D}}(s)$ corresponding to the system in the numerator of (6.8)

$$H^k_{\mathcal{D}}(s) = \mathcal{L}\{\frac{\partial}{\partial t}h^k(t, \omega_p)\} = sH^k(s), \quad (6.9)$$

and call this the frequency reassignment filter. Note that $H^k(s)$ is the filter defined in (4.8). Equation (6.8) can finally be written

$$\hat{\omega_i}(t, \omega_p) = \mathrm{Im}\{\frac{y^k_{\mathcal{D}}(t, \omega_p)}{y^k(t, \omega_p)}\}, \quad (6.10)$$

where $y^k_{\mathcal{D}}(t, \omega_p)$ denotes the kth order frequency reassignment filter defined by (6.9) applied on the input signal $x(t)$.

6.3.2 The Transfer Function for the New Time Coordinate

Starting from (6.7), and substituting $y^k(t,\omega_p)$ by (6.1), one gets

$$\hat{t_g}(t,\omega_p) = t - \mathrm{Im}\{\frac{\frac{\partial}{\partial\omega_p}\int_{-\infty}^{\infty} h^k(t-\tau,\omega_p)x(\tau)d\tau}{y^k(t,\omega_p)}\} \tag{6.11}$$

$$= t - \mathrm{Im}\{\frac{\int_{-\infty}^{\infty} \frac{\partial}{\partial\omega_p} h^k(t-\tau,\omega_p)x(\tau)d\tau}{y^k(t,\omega_p)}\}. \tag{6.12}$$

There is no simple relation between the two domains in this case. For the differentiation of the impulse response with respect to frequency, it is necessary to formally take the derivative. As it happens, the impulse response of this system is easily differentiated to yield a simple relation anyway

$$\frac{\partial}{\partial\omega_p}h^k(t,\omega_p) = \frac{\partial}{\partial\omega_p}\sigma_p^k\frac{t^{k-1}}{(k-1)!}e^{-(\sigma_p-i\omega_p)t}u(t) = it\frac{t^{k-1}}{(k-1)!}e^{-p^\star t}u(t) = ith^k(t,\omega_p). \tag{6.13}$$

Multiplication by t in the time-domain is equivalent to taking the negative of the derivative with respect to s in the Laplace domain. One can therefore define a new transfer function $H_{\mathcal{J}}^k(s)$ corresponding to the system in the numerator of (6.12)

$$H_{\mathcal{J}}^k(s) = \mathcal{L}\{ith^k(t,\omega_p)\} = -i\frac{\partial}{\partial s}H^k(s), \tag{6.14}$$

and call this the time reassignment filter. The new filter coefficients are related through the differentiation operator applied on the original filters transfer function, provided that (6.13) holds. If so, (6.14) can in terms of the original transfer function be written

$$H_{\mathcal{J}}^k(s) = \frac{ik}{\sigma_p}H^{k+1}(s), \tag{6.15}$$

Equation (6.12) can finally be written

$$\hat{t_g}(t,\omega_p) = t - \mathrm{Im}\{\frac{y_{\mathcal{J}}^k(t,\omega_p)}{y^k(t,\omega_p)}\}, \tag{6.16}$$

where $y_{\mathcal{J}}^k(t,\omega_p)$ denotes the kth order time reassignment filter defined by (6.15) applied on the input signal $x(t)$.

6.4 Discretization of the Time and Frequency Reassignment Filters

This section shows how the two additional systems in the previous section can be discretized to yield digital filter coefficients. Figure (6.1) shows the magnitude of the original filter's impulse response (4.10). Figure (6.2) and (6.3) shows the magnitudes of the impulse responses of the time and frequency reassignment filters, respectively.

6.4.1 The Frequency Reassignment Filter

Starting with the kth order frequency reassignment filter

$$H_{\mathcal{D}}^k(s) = sH^k(s) = \frac{s\sigma_p^k}{(s+p)^k} \tag{6.17}$$

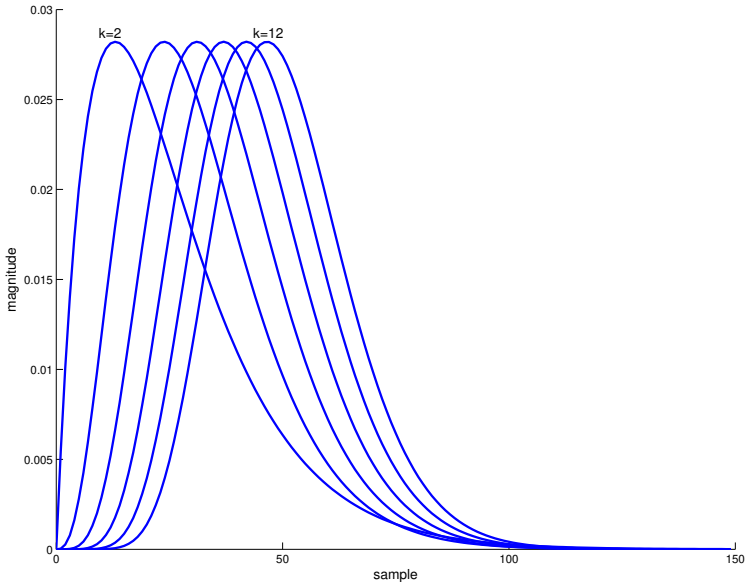


Figure 6.1: *The original filter's impulse response corresponding to (4.8) for increasing orders k.*

Its impulse response $h^k_{\mathcal{D}}(t)$ is

$$h^k_{\mathcal{D}}(t) = \mathcal{L}^{-1}\{H^k_{\mathcal{D}}(s)\} = \frac{\partial}{\partial t} h^k(t) = \frac{\partial}{\partial t} \frac{\sigma_p^k t^{k-1}}{(k-1)!} e^{-p^\star t} u(t) \tag{6.18}$$

$$= \frac{\sigma_p^k t^{k-2}}{(k-2)!} e^{-p^\star t} - p^\star \frac{\sigma_p^k t^{k-1}}{(k-1)!} e^{-p^\star t} \tag{6.19}$$

Sampling $h^k_{\mathcal{D}}(t)$ by $t = nT$ yields

$$h^k_{\mathcal{D}}(nT) = \frac{\sigma_p^k (nT)^{k-2}}{(k-2)!} e^{-p^\star nT} - p^\star \frac{\sigma_p^k (nT)^{k-1}}{(k-1)!} e^{-p^\star nT} \tag{6.20}$$

$$\tag{6.21}$$

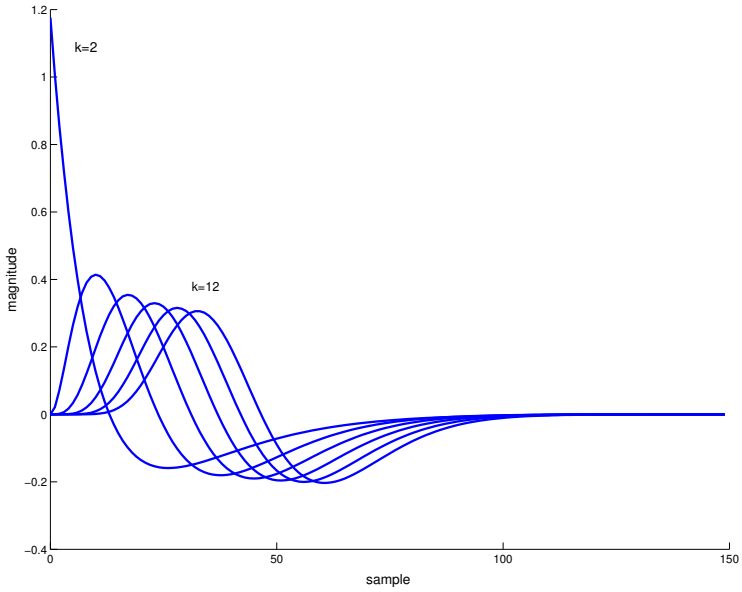


Figure 6.2: *The frequency reassignment filter's impulse response corresponding to (6.9) for increasing orders k.*

Taking the z-transform gives

$$H_{\mathcal{D}}^{k}(z) = \mathcal{Z}\{h_{\mathcal{D}}^{k}(nT)\} = \frac{\sigma_p^k T^{k-2}}{(k-2)!} \sum_{n=0}^{\infty} n^{k-2} e^{-p^\star nT} z^{-n} - \frac{\sigma_p^k p^\star T^{k-1}}{(k-1)!} \sum_{n=0}^{\infty} n^{k-1} e^{-p^\star nT} z^{-n} \tag{6.22}$$

$$= \frac{\sigma_p^k T^{k-2}}{(k-2)!} \overbrace{-z\frac{\partial}{\partial z}}^{k-2 \text{ times}} \frac{1}{1-e^{-p^\star T} z^{-1}} - \frac{\sigma_p^k p^\star T^{k-1}}{(k-1)!} \overbrace{-z\frac{\partial}{\partial z}}^{k-1 \text{ times}} \frac{1}{1-e^{-p^\star T} z^{-1}} \tag{6.23}$$

With an extra scaling factor T for correct normalization, (6.23) can be written in terms of the original transfer function (4.35)

$$H_{\mathcal{D}}^{k}(z) = \sigma_p H^{k-1}(z) - p^\star H^k(z) \tag{6.24}$$

Equation (6.24) reveals that the actual filtering process essentially can be done by taking the difference of the outputs of two systems of order k and $k-1$. However, Table (6.1) lists the digital filter coefficients for orders $k = 1 \ldots 4$ which can be used directly instead of via equation (6.24).

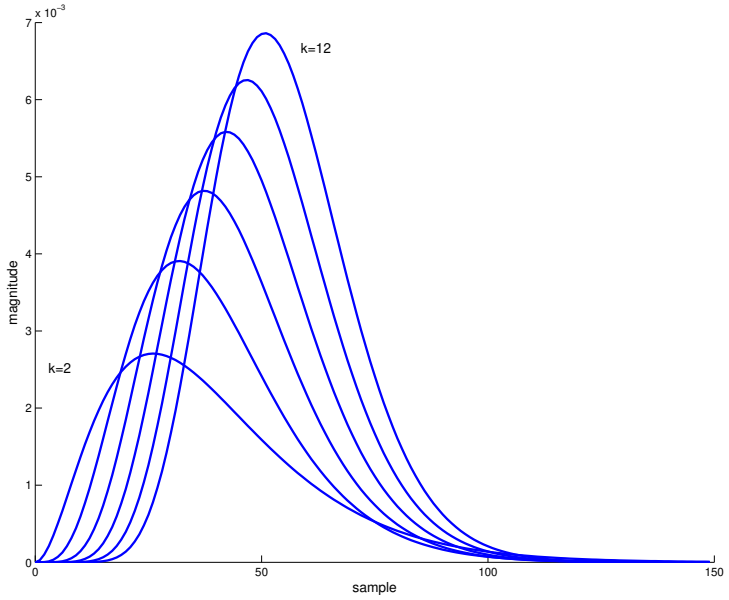


Figure 6.3: *The time reassignment filter's impulse response corresponding to (6.14) for increasing orders k.*

6.4.2 The Time Reassignment Filter

By following the same procedure as in the previous section, the time reassignment filter $H_{\mathcal{T}}^k(s)$ in (6.15) is easily digitized. The transfer function of the time reassignment filter is related to the original transfer function by

$$H_{\mathcal{T}}^k(z) = \frac{ik}{\sigma_p} H^{k+1}(z). \tag{6.25}$$

Thus, the time reassignment filter coefficients for a system of order k are found by using the coefficients from the original system of order $k+1$ except for a multiplication of the b_n coefficients with ik/σ_p.

Reaching the conclusion that recursive reassignment for a system of order k is related to three systems of increasing orders $k-1$, k and $k+1$.

6.4.3 The Discretized Expressions for the Reassignment Coordinates

When discretized, expressions (6.10) and (6.16) becomes

$$\hat{\omega_i}(n,\omega_p) = \mathrm{Im}\{\frac{y_{\mathcal{D}}^k(n,\omega_p)}{\Omega y^k(n,\omega_p)}\}, \tag{6.26}$$

and

$$\hat{t_g}(n,\omega_p) = n - \mathrm{Im}\{\frac{y_{\mathcal{T}}^k(n,\omega_p)}{T y^k(n,\omega_p)}\}, \tag{6.27}$$

where $n = 0, 1, \ldots$ now is the discrete time variable, Ω represents the sampling period i frequency, and T as usual is the sampling period in time.

k	b_n, a_n
1	$b_0 = -\sigma_p p^\star T$ $a_0 = 1, a_1 = -e^{-p^\star T}$
2	$b_0 = \sigma_p^2 T, b_1 = -\sigma_p^2 T(1 + p^\star T)e^{-p^\star T}$, $a_0 = 1, a_1 = -2e^{-p^\star T}, a_2 = e^{-2p^\star T}$
3	$b_0 = 0, b_1 = \sigma_p^3 \frac{T^2}{2}(2 - p^\star T)e^{-p^\star T}, b_2 = -\sigma_p^3 \frac{T^2}{2}(2 + p^\star T)e^{-2p^\star T}$, $a_0 = 1, a_1 = -3e^{-p^\star T}, a_2 = 3e^{-2p^\star T}, a_3 = -e^{-3p^\star T}$
4	$b_0 = 0, b_1 = \sigma_p^4 \frac{T^3}{6}(3 - p^\star T)e^{-p^\star T}, b_2 = -\sigma_p^4 \frac{2T^4}{3} p^\star e^{-2p^\star T}, b_3 = -\sigma_p^4 \frac{2T^3}{3}(3 + p^\star T)e^{-3p^\star T}$, $a_0 = 1, a_1 = -4e^{-p^\star T}, a_2 = 6e^{-2p^\star T}, a_3 = -4e^{-3p^\star T}, a_4 = e^{-4p^\star T}$

Table 6.1: *A summary of the frequency reassignment filter coefficients for $k = 1 \ldots 4$.*

6.5 Recursive Reassignment Eliminates the Time Delay in the RTFRs

The recursive reassignment method eliminates the group delay introduced by the recursive filters. The group delay of (4.35) is added to the group delay of the input signal when it is passed through the system. Since the partial phase derivative with respect to frequency is nothing else than the "local" group delay of the input signal, it follows from (6.12) that it is subtracted to yield zero time delay in the RTFRs.

6.6 Time-Domain Shape Invariance

As mentioned in section (4.9), the (R)RTFRs can be modified to preserve time-domain signal shape exactly. This section outlines a possible solution.

6.6.1 Bidirectional Filtering

The reason of the shape invariance is the non-linear phase of the recursive filters. The following analysis shows how the phase delay vanish to zero if a signal is filtered bidirectional [31]. Let $v(n)$ be the output signal, $x(n)$ the input signal, $h(n)$ the causal impulse response, and let $\star$ denote convolution

$$v(n) = (h \star x)(n)$$

now time-reversing $v(n)$ to get $v(-k)$ and convolving this with $h(n)$ a second time to get $w(n)$

$$w(n) = (h \star v(-k))(n)$$

and then again time-reversing the output $w(n)$ to get the final output $y(n)$

$$y(n) = w(-n) = (h \star v(-k))(-n) \tag{6.28}$$
$$= (h(-k) \star v)(n) \tag{6.29}$$
$$= (h(-k) \star (h \star x))(n) \tag{6.30}$$

Equation (6.30) shows that bidirectional filtering is the same as filtering with the impulse response, followed by a second filtering with the time-reversed impulse response.

By using the z-transform, the effect can be studied easily. The z-transform of a time-reversed signal $x(-n)$ is

$$\sum_{n=0}^{\infty} x(-n)z^{-n} = \sum_{m=0}^{\infty} x(m)z^{m} = \sum_{m=0}^{\infty} x(m)(z^{-1})^{-m} = X(z^{-1}) \tag{6.31}$$

Using the Convolution theorem twice, the z-transform of (6.30) is given by

$$Y(z) = H(z^{-1})H(z)X(z) \tag{6.32}$$

or equivalently, evaluated on the unit circle by setting $z = e^{i\omega}$,

$$Y(\omega) = H(e^{-i\omega})H(e^{i\omega})X(e^{i\omega}) = |H(e^{i\omega})|^2 X(e^{i\omega}). \tag{6.33}$$

As $|H(e^{i\omega})|$ is a real function, it follows that the phase response is zero. As a bonus, the magnitude response is squared, which in turn yields lower time-frequency uncertainty. The disadvantage is that the number of filter coefficients doubles, hence, double computational cost.

6.6.2 Complex Coefficient Filters

The filter in (4.8) is a complex coefficient filter (complex filter). Complex filters were the first time discussed in [33]. If the filter in (4.8) was a real coefficient filter (real filter), then the system could not produce the analytic output necessary for the reassignment coordinates in (5.6) and (5.4). If complex filters are executed bidirectional, then the phase information vanish as shown in (6.33). The output of the filter will therefore be real, so the problem with the reassignment coordinates is reintroduced. The solution lies in the fact that the input signal itself may first be made analytic, and then real filters can be executed bidirectional still to produce the necessary analytic output. This scheme will thus as well preserve the time-domain signal shape because of the zero phase.

Hence, the idea is to first use a Hilbert transformer on the input signal to produce its analytic associate, then modify the complex filters in (4.8) to be a real filters.

6.6.3 From Complex Filters to Real Filters

The complex filter described in (4.8), of first order (i.e. $k = 1$), may be converted to a real coefficient filter by introducing a conjugate pole p. To ensure that the magnitude at the center frequency ω_p is normalized for all ω_p, it is necessary to

introduce a zero as well. The resulting filter is often called the peak filter [24]. In the Laplace domain it can be written

$$H(s) = \frac{\alpha s}{s^2 + \alpha s + \Omega_0^2} = \frac{\alpha s}{(s+p)(s+p^\star)} \tag{6.34}$$

The corresponding digitized transfer function can easily be found via the Bilinear Transform defined in (1.11):

$$H(z) = H(s)|_{s=\frac{1-z^{-1}}{1+z^{-1}}} = \frac{\alpha s}{s^2 + \alpha s + \Omega_0^2}|_{s=\frac{1-z^{-1}}{1+z^{-1}}} \tag{6.35}$$

$$= \frac{\overbrace{1 - \frac{1}{1+\tan\left(\frac{\Delta\omega}{2}\right)}}^{b_0} + \overbrace{\left(\frac{1}{1+\tan\left(\frac{\Delta\omega}{2}\right)} - 1\right)}^{b_2} z^{-2}}{\underbrace{1}_{a_0} - \underbrace{\frac{2\cos\omega_p}{1+\tan\left(\frac{\Delta\omega}{2}\right)}}_{a_1} z^{-1} + \underbrace{\left(\frac{2}{1+\tan\left(\frac{\Delta\omega}{2}\right)} - 1\right)}_{a_2} z^{-2}}, \tag{6.36}$$

with centre frequency $\omega_p \in [0, \pi]$ and bandwidth $\Delta\omega \in [0, \pi]$.

6.6.4 Bidirectional Calculation of the Reassignment Coordinates

It was pointed out in Chapter 4 that the simplest recursive system there is, also turns out to be the preferable choice. The following arguments will motivate this assertion. When moving from the complex filter in (4.8), to the real filter (6.34), the imaginary part of the poles no longer equals the centre frequency of the magnitude response. In fact, the filter in (4.8) is the only filter having this desired property [31]. As a result, one can no longer take the derivative of the filters impulse response with respect to $\mathrm{Im}\{p^\star\} = \omega_p$ to find the time reassignment coordinate. For instance, the new centre frequency of the two-pole real filter (6.34) is easily seen to be $\Omega_0 = |p| = \sqrt{\sigma_p^2 + \omega_p^2}$. Its impulse response is, however, on the form

$$h(t) = \frac{\alpha}{2\omega_p}\left[pe^{-pt} - p^\star e^{-p^\star t}\right]. \tag{6.37}$$

It reveals that there is no longer a simple (derivative) relationship that can be used to find the same type of reassignment filters as for the filter in (4.8). Thus, one must move to an approximative method for computing the reassignment coordinates. For example by using the cross-spectral method discussed in (5.6.2).

The conclusion of this chapter is that the recursive filter in (4.8) is preffered over any other recursive filter because it yields simple reassignment expressions.

Chapter 7

Algorithms, Numerical Results, Concluding Remarks & Further Work

7.1 Introduction

All resulting theory in this thesis is implemented in MATLAB. The programs can be found on the authors website

`http://www.ii.uib.no/~geirkn/rrspec/`

This chapter starts with explaining the implementations of the algorithms producing both the RTFRs described in Chapter 4, and the RRTFRs described in Chapter 6. Then a few RTFRs and RRTFRs for various signals is shown. A couple of spectrograms are also shown for comparison. Finally comes concluding remarks and a discussion regarding ideas for further work.

7.2 Algorithms

7.2.1 Producing the RTFRs

One of the advantages of the RTFRs, is the easy implementation of the algorithm producing them. No FFTs (1.12.1) are needed, and the prominent part of the algorithm is the implementation of the standard difference equation describing the discrete recursive LTI system. In MATLAB, this filtering can be done by the built-in direct form II transposed implementation of the standard difference equation

$$\begin{aligned} y(n) = b_0x(n) + b_1x(n-1) + \ldots + b_kx(n-k) \\ - a_1y(n-1) - a_2y(n-2) - \ldots - a_ky(n-k) \end{aligned} \tag{7.1}$$

available through the **filter()** function. The coefficients $b_0, \ldots, b_k, a_1, \ldots, a_k$ can be found in table (4.1), or they can be calculated on the fly using the MATLAB function ***kcoef()*** available at the above URL.

The algorithm is thus to execute (7.1) for each frequency in the range of $\mathcal{C}(\omega_p)$. If the cardinality of the set of chosen frequencies is Ω, and the length of the signal to analyze is N, the filtering process produces a matrix, say, $\mathrm{RTFR}(n, \omega_p)$, of size $N \times \Omega$.

To have a linear or logarithmic frequency scale, one chooses $\mathcal{C}(\omega_p)$ according to (4.28) or (4.29), respectively. For constant Q analysis, the bandwidth given by (1.1) must be chosen such that (4.27) is constant for all ω_p in the range of $\mathcal{C}(\omega_p)$. The final step is to take the squared magnitude of $RTFR(n, \omega_p)$ according to (4.30).

It should also be noted that the RTFRs are especially well suited for real-time processing. This is due to that the difference equation in (7.1) can be computed using a cyclic buffer of only $k\Omega$ units. However, this is not possible if the MATLAB function **filter()** is used. The point is that if the difference equation (7.1) is implemented explicitly, the innermost loop can iterate over the frequencies. This means that it is not necessary to keep a history of more than k time samples in the outermost loop.

A MATLAB implementation of the RTFR is available online as *rspec.m* at the above URL.

Computational Cost

If N is the number of samples in the input signal $x(n)$ and k is the order of the system, then the cost is $\Omega \times 2k \times N$, that is, $O(\Omega kN)$.

This can be compared to the redundant classical spectrogram which cost is either $O(\Omega N \log N)$ or $O(N\Omega \log \Omega)$ depending on the implementation.

7.2.2 Producing the RRTFRs

The extra calculations needed to produce RRTFRs from the RTFRs is by implementation of two extra difference equations. These should be executed for each ω_p in $\mathcal{C}(\omega_p)$ as well. Having the output of these systems, equations (6.26) and (6.27) can be implemented directly.

The coefficients for the first difference equation is found in Table (6.1), given system order k. The result of this filtering process is applied with (6.26) to get an additional matrix, say $\mathrm{RCIF}(n, \omega_p)$. Coefficients for the second difference equation is found by using the relation in (6.25). This second filtering process together with (6.27) should produce yet another matrix, $\mathrm{RLGD}(n, \omega_p)$, say.

Note that as an alternative, one can filter the input signal with three systems of increasing order. Then equation (6.24), and (6.25) can be applied on the outputs for use within (6.26) and (6.27). This approach makes Table (6.1) superfluous.

The matrix representing the RRTFR, say, $\mathrm{RRTFR}(n, \omega_p)$, can now be produced by a mapping of all the points in $|\mathrm{RTFR}(n, \omega_p)|^2$ using $\mathrm{RCIF}(n, \omega_p)$ and $\mathrm{RLGD}(n, \omega_p)$ as new coordinates. Considering (5.7), the mapping is simply to move and accumulate all points (n, ω_p) in $|\mathrm{RTFR}(n, \omega_p)|^2$ onto new coordinates given by $[\mathrm{RLGD}(n, \omega_p), \mathrm{RCIF}(n, \omega_p)]$. This can be written mathematically as

$$RRTFR(n, \omega_p) = \sum_{n'} \sum_{\omega'_p} |RTFR(n', \omega'_p)|^2 \delta \left[n - \mathrm{RLGD}(n', \omega'_p) \right] \delta \left[\omega_p - \mathrm{RCIF}(n', \omega'_p) \right], \tag{7.2}$$

where $\delta[\cdot]$ represents the discrete Kronecker delta function.

A MATLAB implementation of the RRTFR is available online as *rrspec.m* at the above URL.

Computational Cost

The calculations of RCIF(n, ω_p) and RLGD(n, ω_p) thus have the same computational cost as the RTFR(n, ω_p). The implementation of (6.26) and (6.27) requires only constant time per sample. The additional mapping involves at a maximum $N \times \Omega$ iterations. On average less, because when a point is zero it is meaningless to reassign it. This gives a total of $7k \times \Omega \times N$, which also is $O(\Omega k N)$.

This can be compared to the classical reassigned spectrogram using the method of Auger and Flandrin which cost is either $O(\Omega N \log N)$ or $O(N\Omega \log \Omega)$.

7.3 Numerical Results

7.3.1 Synthetic Data

The synthetic signal used in this section is a superimposition of two sinusoids close in time, two sinusoids close in frequency, and one quadratic chirp.

Figure (7.1) shows the *ideal* time-frequency representation for this synthetic signal. Note that this is just an illustration made from the a priori known times and frequencies. In the following figures, the same signal is analyzed

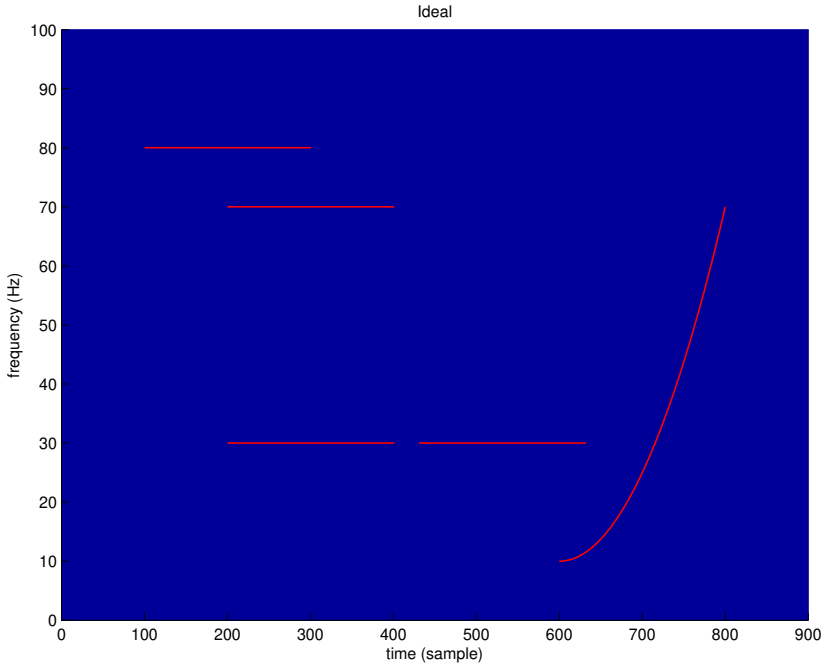


Figure 7.1: *Ideal TFR*

using the RTFR algorithm. Figure (7.3) shows the 1st order (k=1) RTFR, while Figures (7.4) through (7.6) results from using increasing orders k. The bandwidth parameter σ_p is chosen according to (4.39) for optimal time-frequency resolution balance.

The theoretical time-frequency resolution was in Chapter 4 shown to be a function of k approaching the optimum as k increases. It is evident from these figures that the time-frequency resolution increases with k. This is especially easy to see in the regions where the two sinusoids are close in time or close in frequency. The higher the time-frequency resolution, the better the windows in (4.11) can separate close components in time and frequency.

The next figures shows the reassigned versions using the RRTFR algorithm. Figure (7.7) shows the 1st order RRTFR, while figures (7.8) through (7.10) results from using increasing orders k. The bandwidth parameter σ_p is chosen according to (4.39) for optimal time-frequency resolution balance. It is seen that the 1st order RRTFRs are of little value. However, when the order increases, the quality becomes close to the ideal situation.

The next two Figures (7.11) and (7.12) shows how the 4th order RTFRs are affected by choosing the bandwidth parameter σ_p too low and too high, respectively. It is clear that the optimal time-frequency resolution balance provides the best trade-off given that accuracy in time and frequency are equally important. The next two Figures (7.13) and (7.14) shows how the 4th order reassigned versions are affected by the same choices of σ_p.

The reassignment works best if the time-frequency resolution balance is optimal. The presumption of only one significant component per window (discussed in Chapter 5) is most likely satisfied in both time and frequency when the uncertainties are simultaneously minimal. Comparing the ideal TFR in Figure (7.1) with the 4th order RRTFR (using optimal time-frequency resolution balance) in Figure (7.10) shows that the latter is close to the ideal situation.

The 4th order RRTFR in Figure (7.10) can be compared to the classical reassigned spectrogram using various finite windows. Figure (7.15) is the result of using the Hamming window of 1/4 of the analyzed signal's number of samples. Figures (7.16) and (7.17) results from using the Hanning window and the Blackman window (same length), respectively. The classical spectrograms are computed using the Time-Frequency Toolbox available at *http://tftb.nongnu.org/*.

The figures in this section have shown that for the synthetic signal there is a very little difference between the classical reassigned spectrograms and the 4th order RRTFRs. It is evident that the RRTFRs of 4th order is highly competitive with the classical reassigned spectrograms.

Time-Domain Shape Invariance

The problem of time-domain shape invariance discussed in Sections 4.9 and 6.6 can be illustrated by taking a Gabor logon as the input signal with the RTFR algorithm. Figure (7.2) shows the resulting RTFRs using increasing system orders. It is evident that the time-domain shape invariance decreases as the system order increases.

7.3.2 Real Data

Moving to real data yields difficulty when the spectrograms are to be compared. The main reason is that the ideal situation is not known. In Figures (7.18) and (7.19) a real signal recorded from the authors accordion is analyzed using the 4th

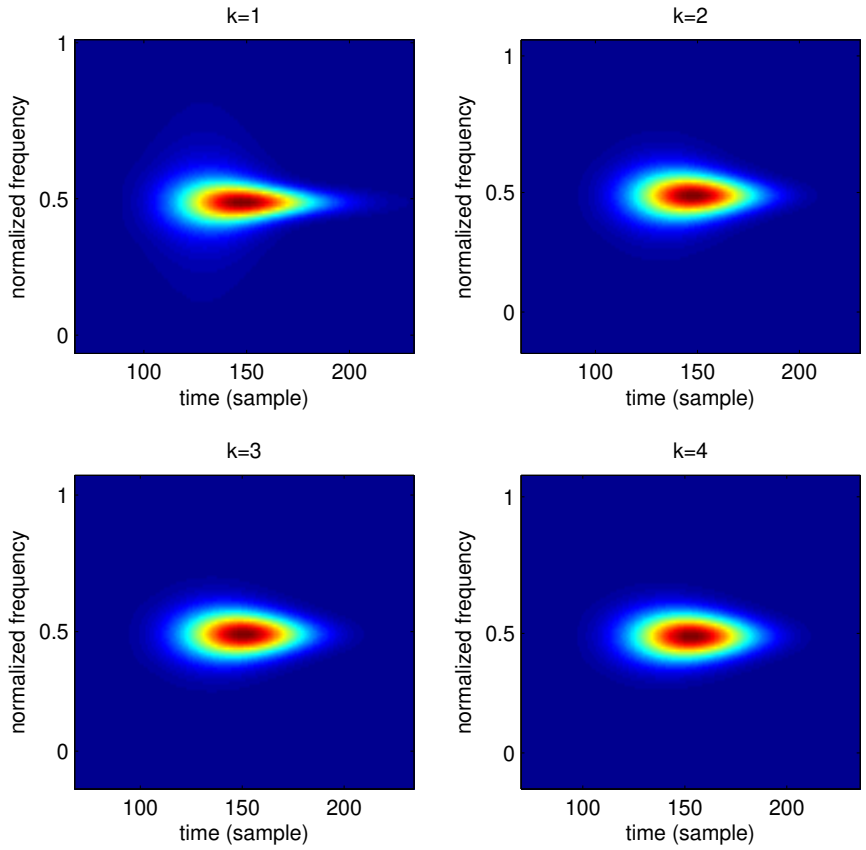


Figure 7.2: *Gabor logons are processed with the RTFR algorithm using increasing system orders: the time-domain shape invariance diminishes as the system order increases. This is seen by that the Gabor logon approaches its true symmetric shape when the system order increases.*

order RRTFR algorithm. For comparison, the classical reassigned spectrogram using the Hamming, Hanning and Blackman windows are also shown. The time-frequency resolution in the classical spectrograms are balanced by trial and error. It is evident from these figures that the RRTFRs of 4th order at least can compete with the classical reassigned spectrograms. More research is required to draw a theoretical conclusion. See Section 7.4.

7.4 Concluding Remarks & Further Work

Utilizing a recursive linear system for making TFRs has the benefit of requiring lower computational cost than of non-recursive linear systems. In addition these systems are easily implemented by implicit linear recursions. The recursive filter proposed has been chosen for several reasons. Firstly, it is a complex analytic filter. This means that the TFRs can be visualized in terms of their envelope. Secondly, the filter order is directly linked to the time-frequency

resolution. As the number of poles increases, the time-frequency resolution approaches the Gaussian optimum. Thirdly, the optimal time-frequency resolution balance is readily achieved by choosing the bandwidth/duration according to (4.39). Fourthly, the proposed filter also lead to simple time-frequency reassignment coordinate expressions. Three filters of increasing order has been shown to be sufficient to calculate the time and frequency reassignment coordinates. It has the main implication of making time-frequency reassignment well suited for real-time processing. As a side-effect, the proposed filter happens to form the basis for a special case of the STFT utilizing a causal, asymmetric, infinite window.

The fourth order (and higher orders) filters yields highly competitive TFRs compared to the classical reassigned spectrograms.

As the reassignment coordinates essentially are the instantaneous frequency and the group delay of the analyzed signal, a side-effect is that the results in Chapter 6 also can be regarded as new instantaneous frequency and group delay estimators. Hainsworth [17] used Cramer-Rao Bound methods to show that the group delay and instantaneous frequency estimators using Auger and Flandrin's [3] method introduce a bias when they are discretized. This was not predicted in their theory because it was derived in continuous mathematics. It leads to the contradictory claim that the discretized instantaneous frequency and group delay estimators are dependent on the window. Further research is required to find out what impact the infinite windows used in this thesis has on the proposed estimators.

The true optimum bandwidth depends adaptively on the analyzed signals instantaneous frequency, and the true optimum duration depends on the analyzed signals group delay. An idea is therefore to dynamically update the bandwidth parameter according to the instantaneous frequency which in turn is measurable through the frequency reassignment filter.

Another idea is to see if it is possible to calculate non-redundant reassigned spectrograms. In other words, such that the reassigned spectrograms are implicitly downsampled by a constant factor.

A related idea is to see if it is possible to develop an inverting transform such that the original signal can be recovered from the RTFRs (and/or RRTFRs). This should be possible by utilizing the same special window in the regular inverse STFT, but can inverse filters be found to yield a recursive inversion as well?

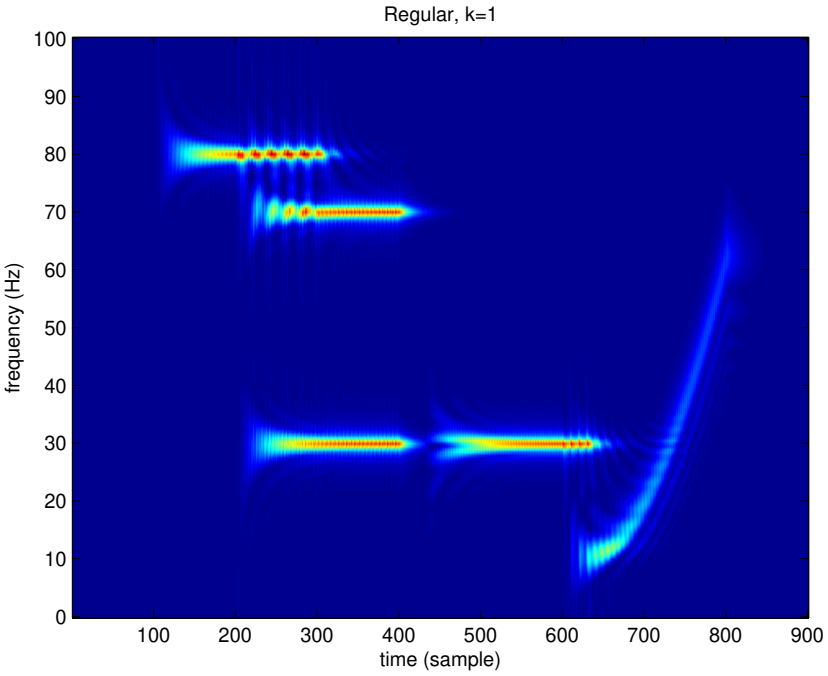


Figure 7.3: *1st order (k=1) RTFR*

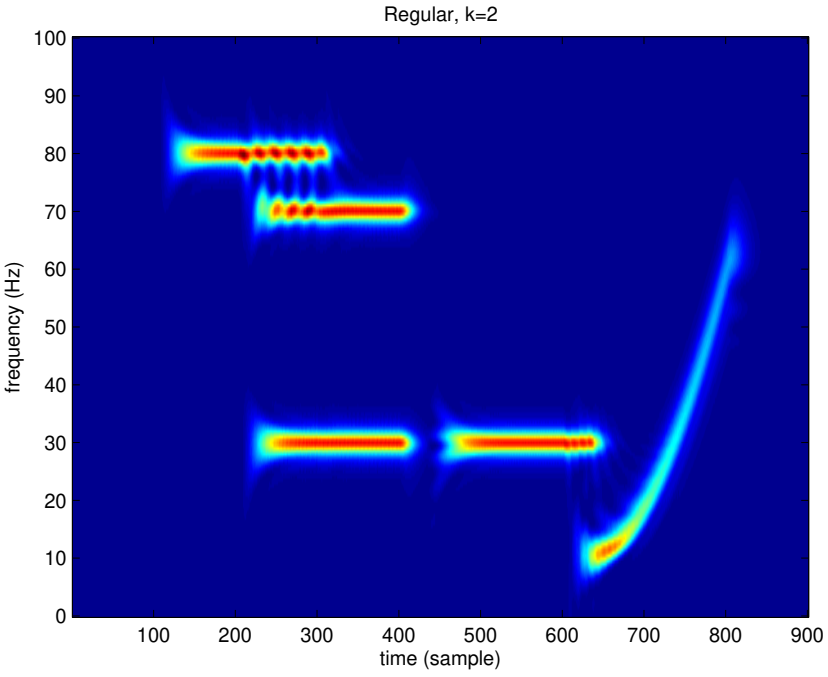


Figure 7.4: *2nd order (k=2) RTFR*

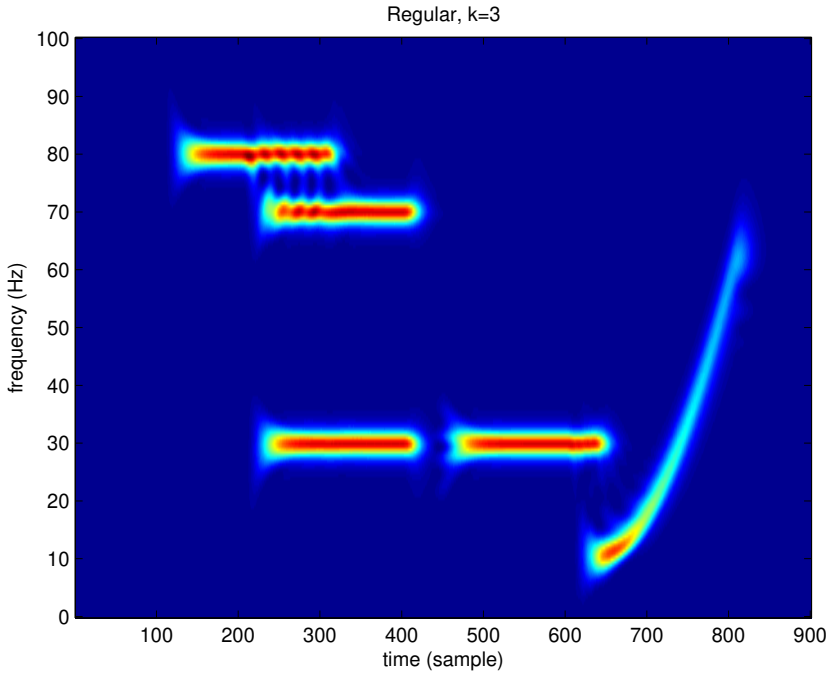


Figure 7.5: *3rd order (k=3) RTFR*

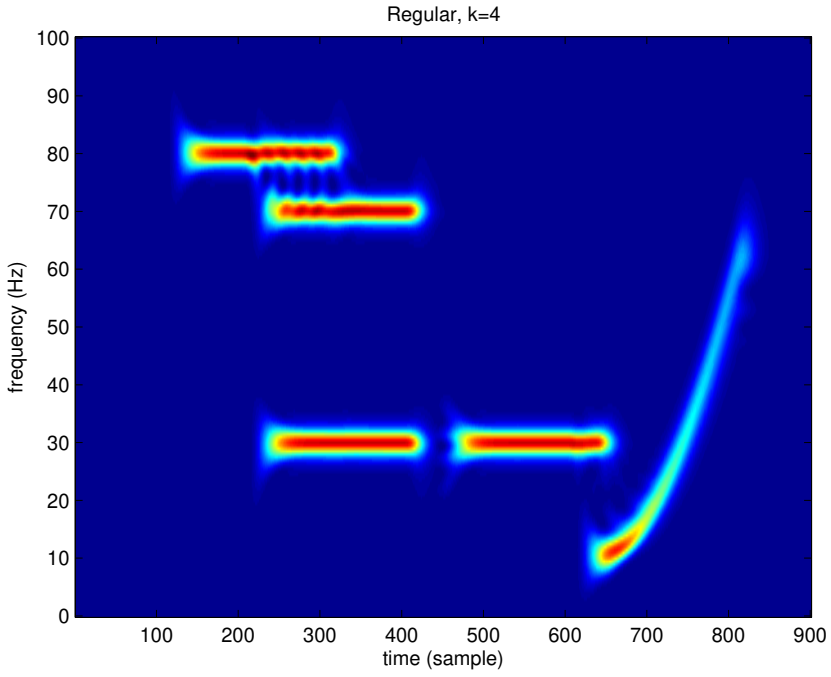


Figure 7.6: *4th order (k=4) RTFR*

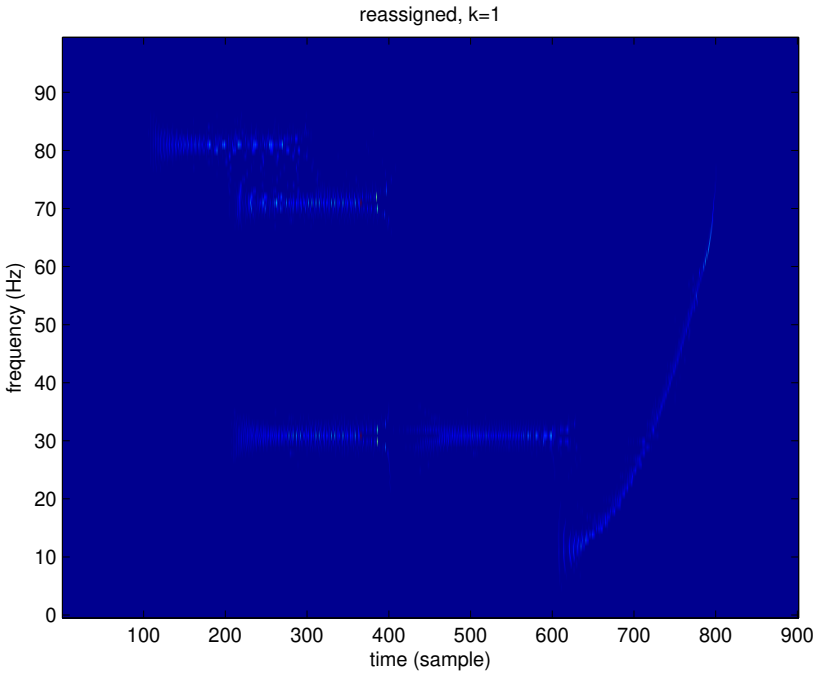


Figure 7.7: *1st order (k=1) RRTFR*

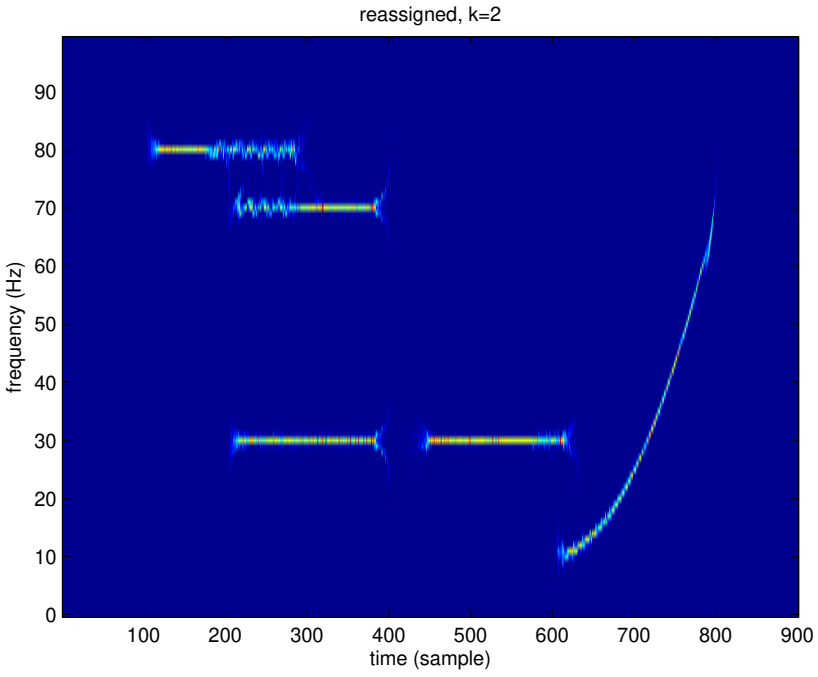


Figure 7.8: *2nd order (k=2) RRTFR*

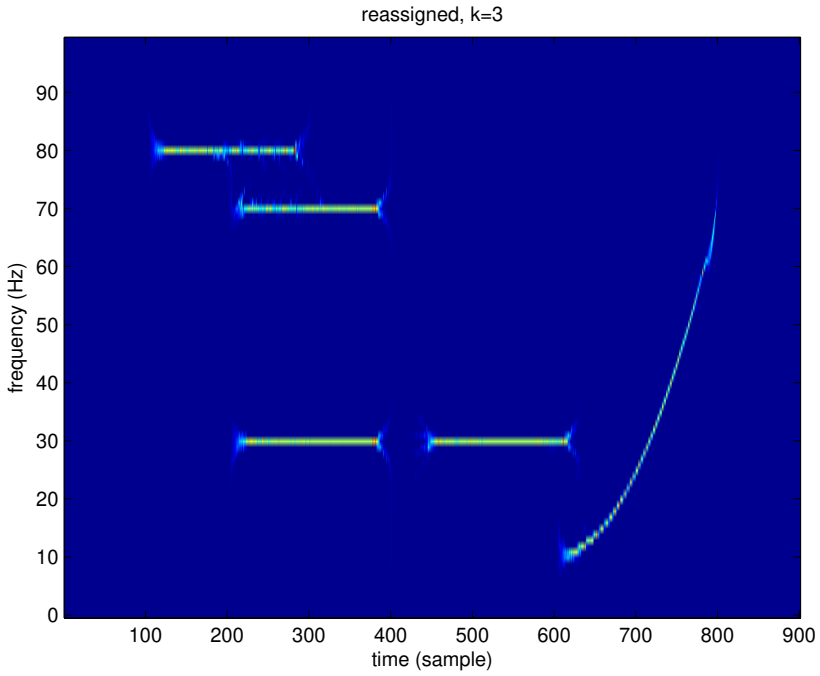


Figure 7.9: *3rd order (k=3) RRTFR*

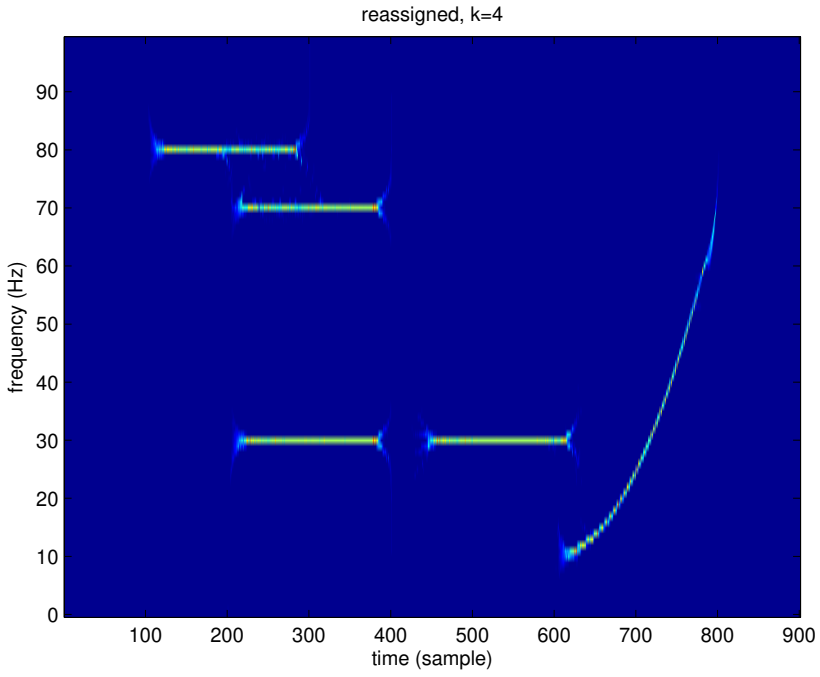


Figure 7.10: *4th order (k=4) RRTFR*

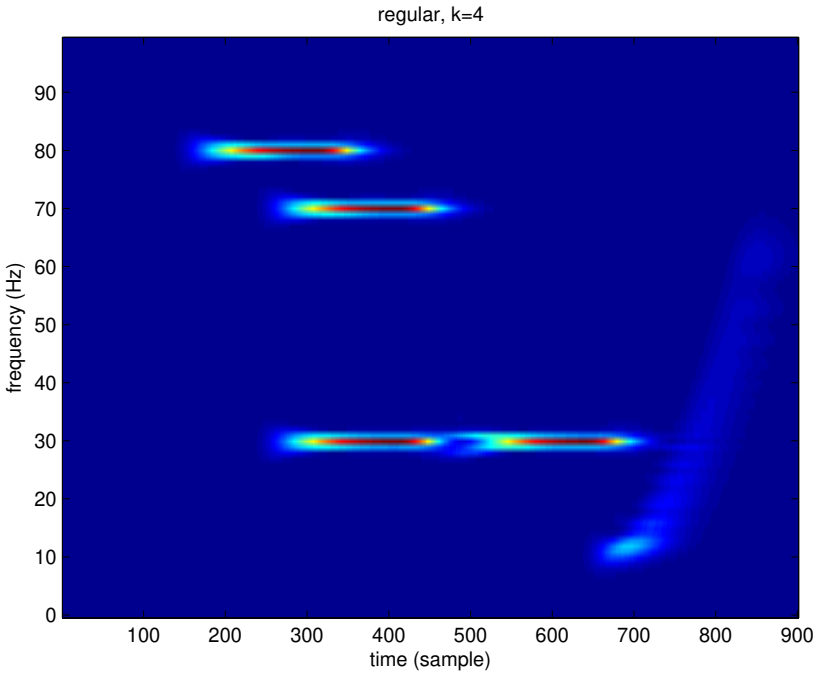


Figure 7.11: *4th order (k=4) RTFR with* ***too low*** σ_p

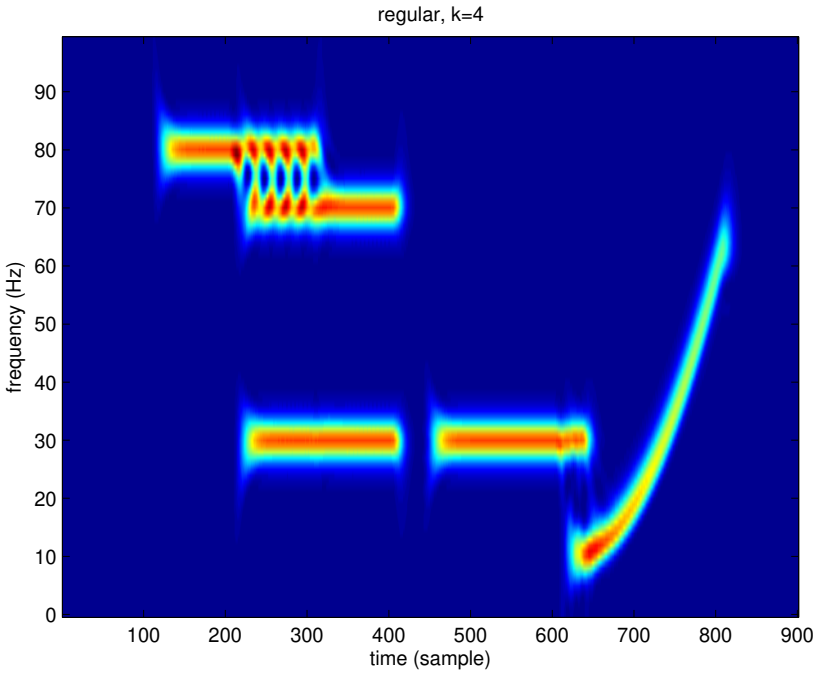


Figure 7.12: *4th order (k=4) RTFR with* ***too high*** σ_p

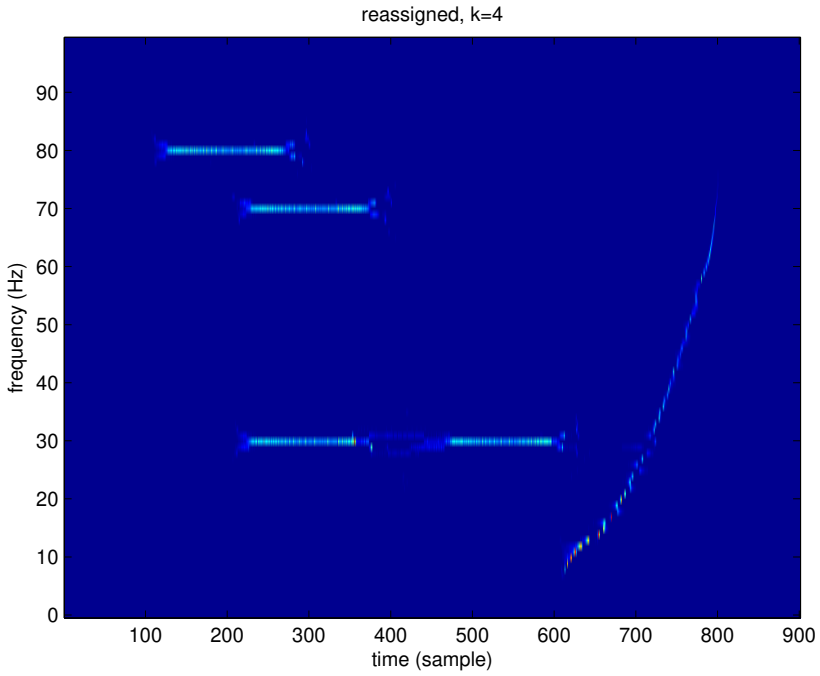


Figure 7.13: *4th order (k=4) RRTFR with* ***too low*** σ_p

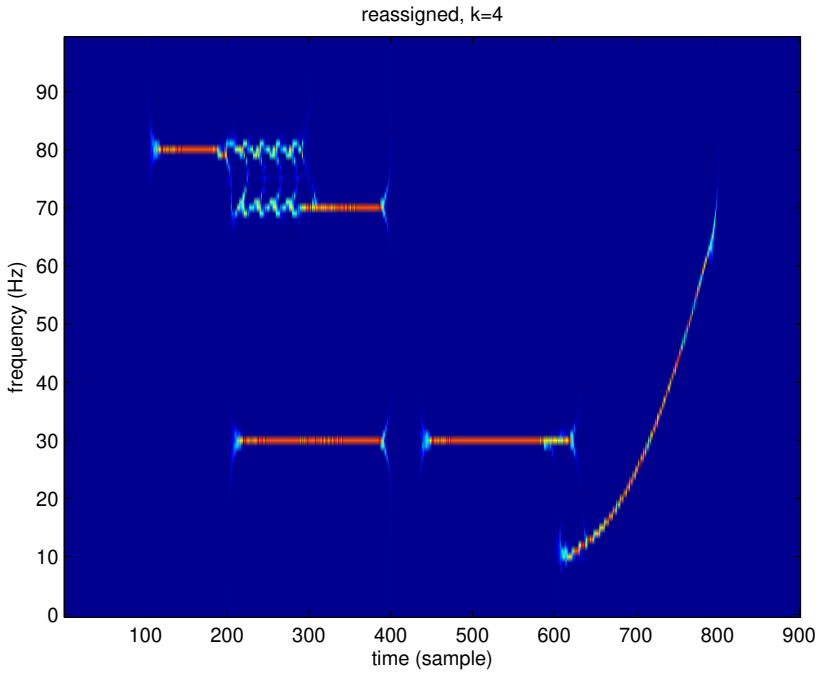


Figure 7.14: *4th order (k=4) RRTFR with* ***too high*** σ_p

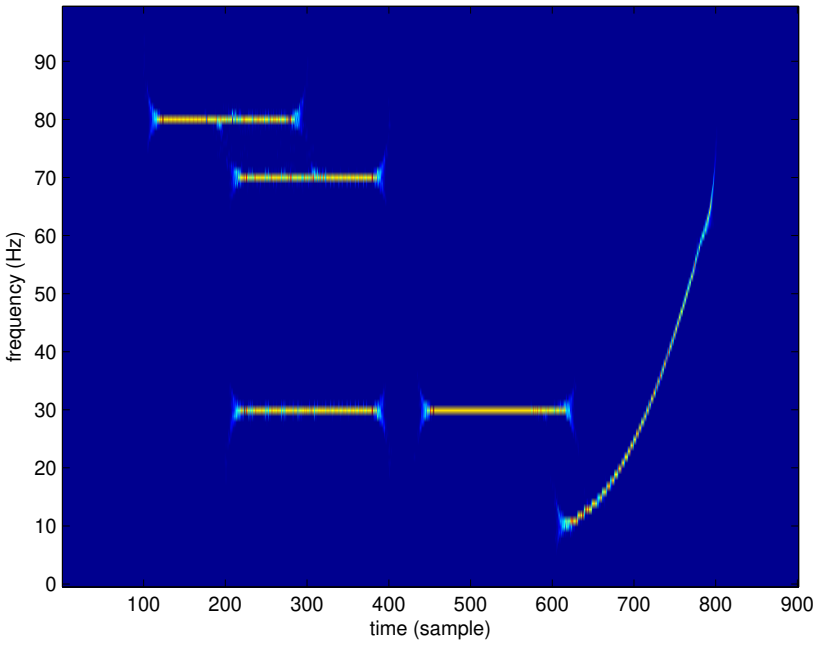


Figure 7.15: *Classical Reassigned Spectrogram using the Hamming Window*

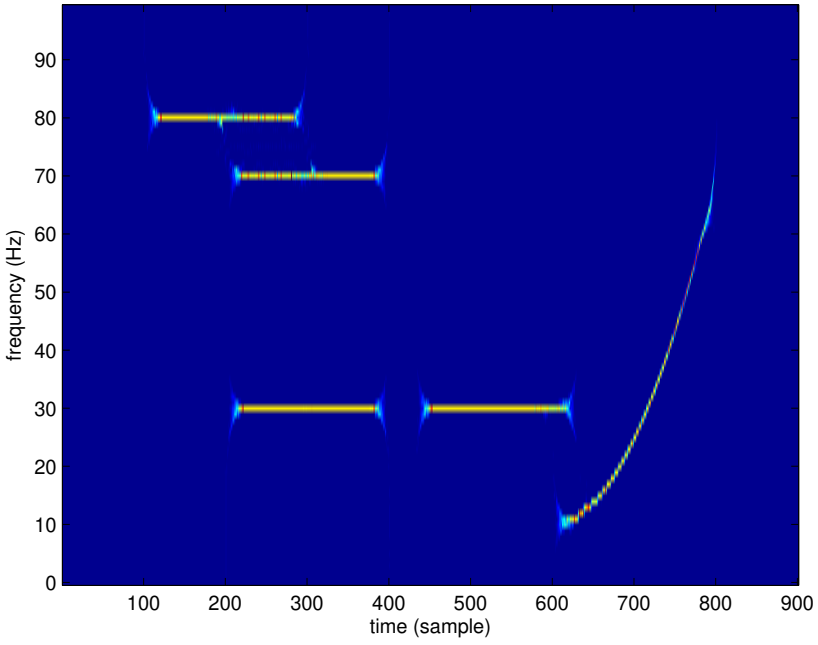


Figure 7.16: *Classical Reassigned Spectrogram using the Hanning Window*

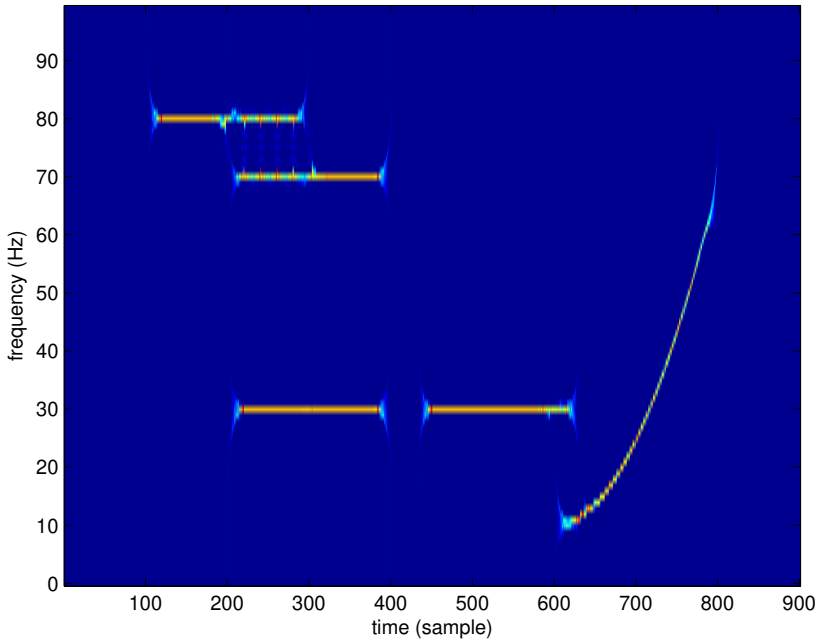


Figure 7.17: *Classical reassigned spectrogram using the Blackman Window*

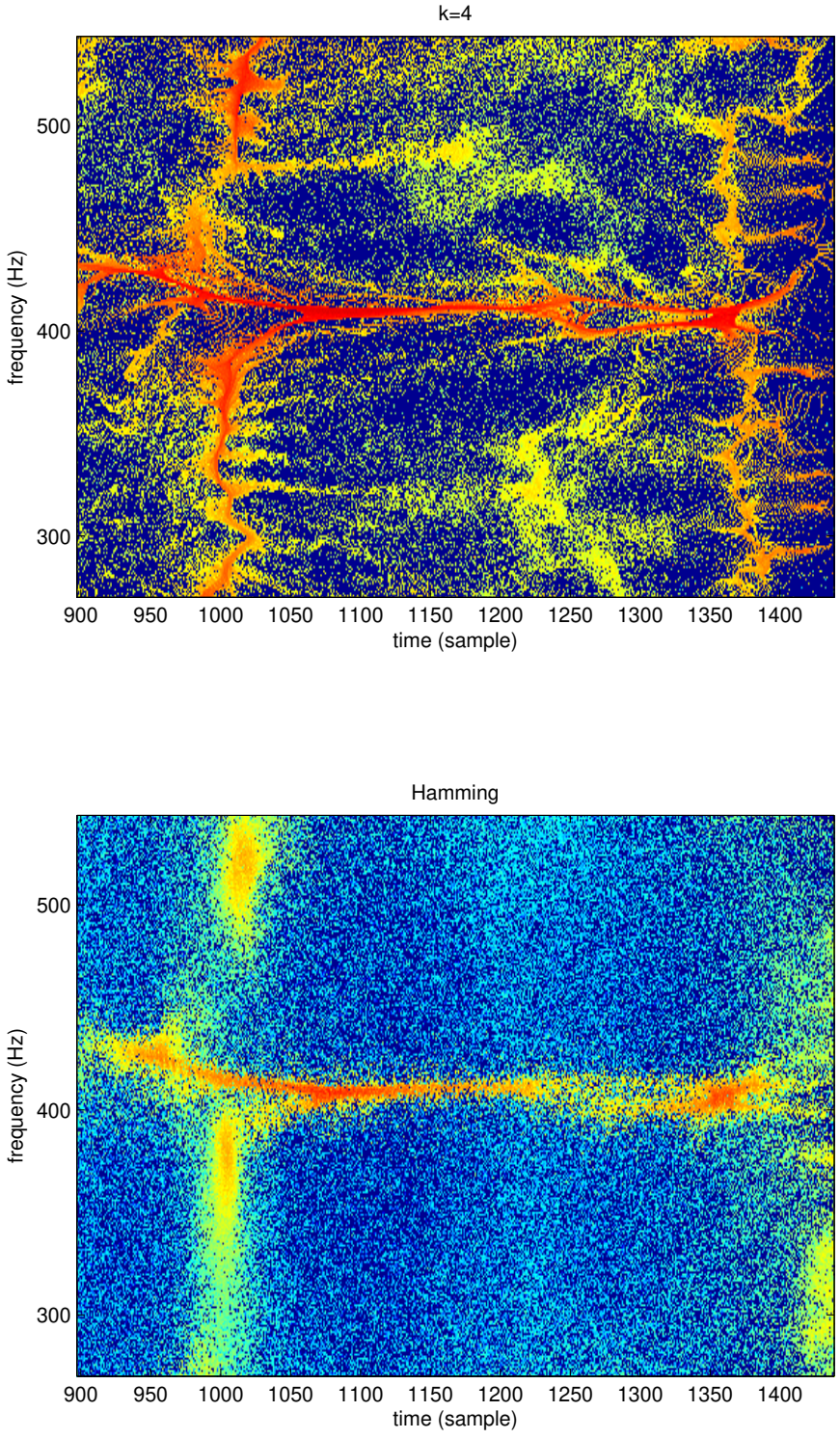


Figure 7.18: *4th order (k=4) RRTFR (top), classical reassigned spectrogram using the Hamming window(bottom).*

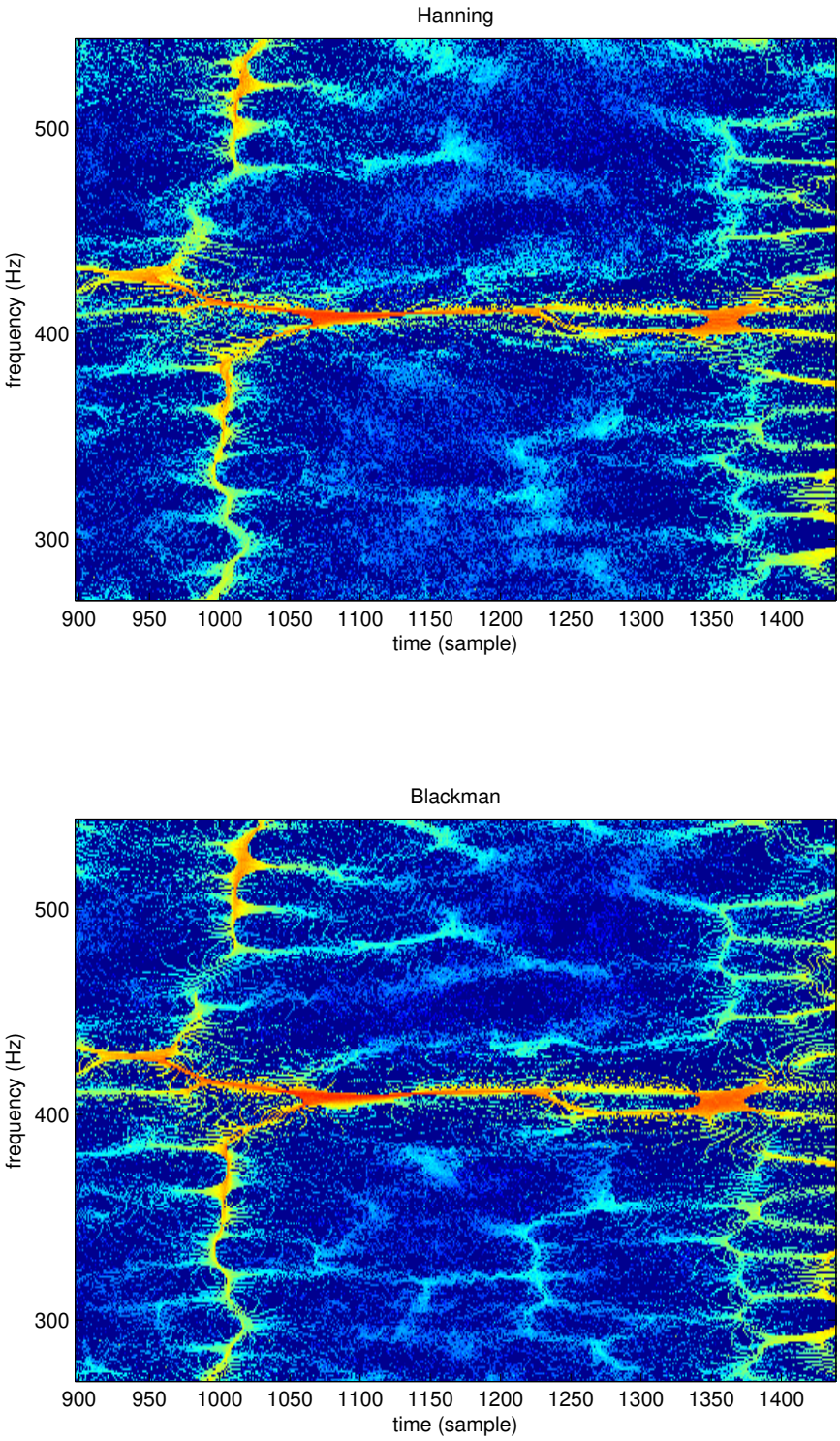


Figure 7.19: *Classical reassigned spectrogram using the Hanning window (top), classical reassigned spectrogram using the Blackman window (bottom).*

Appendix

Journal paper follows ...

Recursive Time-Frequency Reassignment

Geir Kjetil Nilsen

***Abstract*—A fast time-frequency reassignment algorithm for a special case of the Short-Time Fourier Transform (STFT) spectrogram is presented. The algorithm is based on introducing recursion into the linear system describing the regular STFT.**

***Index Terms*—STFT, spectrogram, linear implicit recursion, time-frequency analysis, reassignment.**

I. Introduction

The idea of calculating the STFT in terms of linear implicit recursions is not new [3], [4], [5], [7], [6]. Chen et al. [7] was the first to utilize a causal, infinite window function to compute the STFT recursively, although Friedlander et al. [1] several years earlier used the same window function with the Gabor transform for transient detection. Žnidar [4] developed a perfect reconstruction filter bank based on the same type of underlying filters, but did not at that point characterize this as a STFT with an infinite window. Tomažic and Žnidar [3], [5] together developed a fast recursive STFT algorithm based on the same concept. The main difference from [7] was that their algorithm has a general window hop of N rather than one. It was also shown that as the number of repeated poles on the window increases, the Gabor uncertainty approached the optimal Gaussian uncertainty. Amin and Feng [6] at the same time generalized the STFT with arbitrary infinite windows using cascading (IIR) filter structures. Unser [24] also employed these windows as wavelets.

The main issues concerning the STFT are computational efficiency, time-frequency resolution and stability of the inversion. Here the main concerns will be the first two. The Fourier transform of these infinite windows contains not only zeros, but also poles. The number of poles will be seen to offer a computational cost to time-frequency resolution tradeoff. Moreover, the implementation will be easy in terms of linear implicit recursions.

The method known as *time-frequency reassignment* has shown to yield TFRs with superior time-frequency resolution. The pioneers of the method, Kodera et al. [10], was influenced by the complex energy density theory of Rihaczek [12]. It was shown by Ackroyd [21] that the STFT spectrogram could be expressed as a double convolution over time and frequency. The integrand expression involved the energy densities corresponding to both the window and the signal. This conveyed an important insight into the structure of the spectrogram: the reason why the spectrogram does not sharply reproduce the distribution of the original spectrum is that the energy density of the analyzed signal is smoothed by the energy density of the window. This lead Kodera et al. [10] to propose an enhancement of the spectrogram reversing the smoothing from the windowing. The main idea was to *reassign* each point in the spectrogram with new coordinates better fitting with the energy density of the analyzed signal.

The result was seemingly a TFR beating Gabor's

G. K. Nilsen is at the Department of Computer Science, University of Bergen, Norway, e-mail: geirkn@ii.uib.no

uncertainty principle. Of course this contradicton had a natural explanation. It is best understood by noting that the new coordinates for the reassignment are related to the *instantaneous frequency* (IF) and *group delay* (GD) of the analyzed signal. The definition of IF is only physically meaningful when the signal contains only one frequency component at any given time instant. If the signal consists of two frequency components with equal amplitudes, the IF will be the mean of the two frequencies. Even worse, if the amplitudes are not equal, the IF will vary periodically with the difference frequency of the two. As the number of components increase, the behaviour turns more and more chaotic [20]. The reassignment method applied on the spectrogram will therefore fail if more than one significant signal component falls into one single window in the STFT. Thus, the reassignment method is based on a presumption not taken into account in the uncertainty principle. However, the attractiveness of the method comes from the fact that this is only a *local* limitation. The method have been used and discussed in several areas of research [11], [9], [17], [18]. The reason why it has not been widely used is speculated to be its irreversibility, i.e. when the method is applied, the resulting TFR cannot be used to recover the original signal. This is because the phase information does no longer correspond to the reassigned magnitude. However, recent research by Fulop and Fitz [16] concerning the phase corrections that must be made to attain inversion, is promising.

The reassignment coordinates were shown by Kodera et al. to be related to the partial derivates of the STFT phase. Kodera et al. did, nevertheless, only use approximative finite differences which also requires use of the unreliable operation known as phase-unwrapping [22]. Nelson [11] invented his approximative cross-spectral method making phase-unwrapping superfluous. However, Auger and Flandrin [2] generalized the reassignment to all the members of Cohen's class [15], and it was shown that the reassignment coordinates for the STFT spectrogram can be computed exactly by two extra STFTs using two modified windows.

In Section II the special case of the STFT yielding a recursive linear system is presented using a new formulation. Section III shows that the time-frequency reassignment coordinates can be computed by simple modifications of the transfer functions describing the recursive system. In Section IV, the computational cost is briefly discussed and a couple of classically and recursively computed spectrograms are compared.

II. The STFT Special Case

The STFT is the continuous version of the Gabor transform [8]. It can be regarded as a set of linear convolutions,

$$y(t,\omega) = e^{i\omega t}\int_{-\infty}^{\infty} x(\tau)\mathcal{W}(t-\tau)e^{-i\omega\tau}d\tau, \quad (1)$$

between the input signal $x(t)$ and the impulse responses given by

$$h(t,\omega) = \mathcal{W}(t)e^{i\omega t}, \quad (2)$$

where $\mathcal{W}(t)$ is the finite window. The spectrogram is defined as the complex modulus squared of (1).

Classical implementations of the STFT are hence based on the convolution theorem and FFTs. This approach, however, makes the spectrogram redundant. A faster method avoiding the redundancy is also possible. Essentially, this alternative method stems from viewing (1) as a set of windowed Fourier transforms instead. Now one can introduce a window hop factor not necessarily equal to one, so that the output will implicitly be downsampled by the same factor. Redundancy is, in fact, necessary for the reassignment method presented in

Section III to work. It is therefore appropriate to compare computational cost with the convolutional view.

The crucial point is that by a clever choice of $\mathcal{W}(t)$, (1) can become a solution to an ordinary differential equation. The differential will act on the time dimension, and represent a recursion propagating over in the discrete case as a linear implicit recursion.

If the window is defined

$$W^k(t) = \frac{\sigma_p^k t^{k-1}}{(k-1)!} e^{-\sigma_p t} u(t), \tag{3}$$

with k to be explained subsequently, then the impulse response (2) can be written

$$h^k(t, \omega_p) = \frac{\sigma_p^k t^{k-1}}{(k-1)!} e^{-\sigma_p t} e^{i\omega_p t} u(t) = \mathcal{W}^k(t) e^{i\omega_p t}. \tag{4}$$

The transfer function of (4) is given by

$$H^k(s) = \frac{Y^k(s)}{X^k(s)} = \frac{\sigma_p^k}{(s + p^\star)^k}, \tag{5}$$

where $\star$ denotes complex conjugation, and the pole p is a complex constant defined by

$$p = \sigma_p + i\omega_p. \tag{6}$$

It is evident that (5) is a rational with one pole raised to the kth power. This means that (4) consists of k repeated convolutions of first order impulse responses.

To apply the filter in (4), it is to be convoluted with the input signal $x(t)$ for each ω_p

$$y^k(t, \omega_p) = \int_{-\infty}^{\infty} x(\tau) h^k(t - \tau, \omega_p) d\tau. \tag{7}$$

The STFT in (1) is equivalent to (7) using the special window in (4), but (7) is also a solution to the kth order ordinary differential equation

$$\sum_{n=0}^{k} \binom{k}{n} (\sigma_p - i\omega_p)^n \frac{\partial^{k-n}}{\partial t^{k-n}} y(t, \omega_p) = \sigma_p^k x(t), \tag{8}$$

with all initial conditions equal to zero. The power of the recursive approach – seen from a computational point of view – is that (7) is equivalent to (8), and that the discretized version of (8) can be computed efficiently in terms of linear implicit recursions.

Reasonable definitions for the bandwidth and the duration of (5) and (4), respectively, were proposed by [3]. The classical definition of bandwidth cannot be applied, because the second moment of $|H(s)|^2|_{s=i\omega}$ does not converge. The bandwidth using this definition gives

$$\Delta_\omega = \frac{2\pi\sigma_p[2(k-1)]!}{[(k-1)!]^2 2^{2k-1}}, \tag{9}$$

and the duration is found to be

$$\Delta_t = \frac{[2(k-1)]!}{\sigma_p \, (k-1)^{2(k-1)} \, e^{-2(k-1)} 2^{2k-1}}. \tag{10}$$

The time-frequency resolution is inversely proportional to the uncertainty $\Delta_t \Delta_\omega$. Using $k = 1$ yields $\Delta_t \Delta\omega = \pi/2$, while using the optimum Gaussian window yields $\Delta_t \Delta\omega = \pi$. As the number of repeated poles (k) in (5) increases, the time-frequency resolution approaches the optimum. This is a direct consequence of the central limit theorem, readily understood through the repeated convolutions in (4). The convergence is fast enough to yield $\Delta_t \Delta_\omega = 3.077$ when $k = 5$.

The real part of the pole in (6) controls the decay rate of the exponential. This parameter can therefore be used to balance the tradeoff between time and frequency resolution. The *optimum time-frequency resolution balance* is found by setting $\Delta_\omega = \Delta_t$, and solving for σ_p. It is given by

$$\sigma_p = \frac{(k-1)!}{\sqrt{2\pi}(k-1)^{(k-1)} e^{1-k}}. \tag{11}$$

Note that the balancing parameter σ_p is cancelled in $\Delta_t \Delta_\omega$, so it has nothing to do with the time-frequency resolution (or uncertainty).

Quite simple, one can also get constant Q analysis by choosing σ_p such that ω_p is proportional to Δ_ω.

A. Discretizing the Recursive System

Using the *impulse invariant* discretization method [19], equation (4) can be sampled such that the sampling period T is sufficiently small in Nyquist sense. The digitized transfer function can be written

$$H^k(z) = \sigma_p^k T^k \frac{\sum_{j=1}^{k} A_{k-1,k-j}(e^{-pT}z^{-1})^{j-1} + \delta(k-1)}{(k-1)!(1-e^{-p^\star T}z^{-1})^k}, \quad (12)$$

with $A_{n,k}$ representing the Eulerian numbers given by

$$A_{k,n} = \sum_{j=0}^{n}(-1)^j \binom{k+1}{j}(n+1-j)^k. \quad (13)$$

Table (I) lists the coefficients for orders $k = 1$ to $k = 5$. Increasing k also means increasing the number of filter coefficients, so the parameter k gives a tradeoff between computational cost and time-frequency resolution. The implementation can now simply be done by using these coefficients with the standard difference equation[1]

$$y(n) = \sum_{j=0}^{k} b_j x(n-j) - \sum_{j=1}^{k} a_j y(n-j) \quad (14)$$

for each chosen ω_p.

In the classical STFT, the number of samples in the window is used to balance time-frequency resolution tradeoff. However, the window length in the special case is implicitly infinite, so the time-frequency resolution balance can only be influenced by the parameter σ_p. The *optimum time-frequency resolution balance* is in the discrete case found to be

$$\sigma_p = \frac{(k-1)!}{\sqrt{2\pi T}(k-1)^{(k-1)}e^{1-k}}. \quad (15)$$

k	b_n	a_n
1	$b_0 = \sigma_p T$	$a_0 = 1, a_1 = -e^{-p^\star T}$
2	$b_0 = 0, b_1 = \sigma_p^2 T^2 e^{-p^\star T}$	$a_0 = 1, a_1 = -2e^{-p^\star T}, a_2 = e^{-2p^\star T}$
3	$b_0 = 0, b_1 = \sigma_p^3 \frac{T^3}{2} e^{-p^\star T}$, $b_2 = \sigma_p^3 \frac{T^3}{2} e^{-2p^\star T}$	$a_0 = 1, a_1 = -3e^{-p^\star T}$, $a_2 = 3e^{-2p^\star T}, a_3 = -e^{-3p^\star T}$
4	$b_0 = 0, b_1 = \sigma_p^4 \frac{T^4}{6} e^{-p^\star T}$, $b_2 = \sigma_p^4 \frac{2T^4}{3} e^{-2p^\star T}$, $b_3 = \sigma_p^4 \frac{T^4}{6} e^{-3p^\star T}$	$a_0 = 1, a_1 = -4e^{-p^\star T}$, $a_2 = 6e^{-2p^\star T}, a_3 = -4e^{-3p^\star T}$, $a_4 = e^{-4p^\star T}$
5	$b_0 = 0, b_1 = \sigma_p^5 \frac{T^5}{24} e^{-p^\star T}$, $b_2 = \sigma_p^5 \frac{11T^5}{24} e^{-2p^\star T}$, $b_3 = \sigma_p^5 \frac{11T^5}{24} e^{-3p^\star T}$, $b_4 = \sigma_p^5 \frac{T^5}{24} e^{-4p^\star T}$	$a_0 = 1, a_1 = -5e^{-p^\star T}$, $a_2 = 10e^{-2p^\star T}, a_3 = -10e^{-3p^\star T}$, $a_4 = 5e^{-4p^\star T}, a_5 = -e^{-5p^\star T}$

TABLE I
SUMMARY OF FILTER COEFFICIENTS FOR ORDERS $k = 1...5$

III. RECURSIVE REASSIGNMENT

The reassignment method also fits well within a recursive system. In analogy with Auger and Flandrin's method [2], this section shows that if the recursive filter in (5) is used, the reassignment coordinates can be computed in terms of linear implicit recursions as well.

Kodera [10] showed that the coordinates for the reassignment are related to the partial derivatives of the phase of the STFT. However, the partial derivates also match the definitions of *IF* and *GD*. As IF and GD are only defined for *analytic* signals, it is necessary for the output of the STFT to be analytic. This is also the case, because the real and imaginary parts of the complex impulse response in (2) forms a Hilbert transform pair. Convolutions with (2) will have the same property, and therefore also be *analytic*.

The chosen recursive filter must therefore be a complex analytic filter. The impulse response in (4) is *analytic* because (5) does not form pairs of conjugate poles. Moreover, the filter in (5) is the only recursive filter with the property that the imaginary part of its pole also is its center frequency [23]. This means that the recursive filter in (5) leads to simple reassignment expressions, as well as being the basis for a special case of the STFT.

Since the recursive system in (5) can be interpreted as

[1]For instance available through the MATLAB function *filter()*.

the STFT, it makes sense to use the expressions found by Kodera et al. [10] involving the partial phase derivatives. The reassignment coordinates can be expressed as

$$\hat{\omega_i}(t,\omega_p) = \frac{\partial}{\partial t}\phi(t,\omega_p) = \text{Im}\left[\frac{\frac{\partial}{\partial t}y^k(t,\omega_p)}{y^k(t,\omega_p)}\right] \quad (16)$$

and

$$\hat{t_g}(t,\omega_p) = t - \frac{\partial}{\partial \omega_p}\phi(t,\omega_p) = t - \text{Im}\left[\frac{\frac{\partial}{\partial \omega_p}y^k(t,\omega_p)}{y^k(t,\omega_p)}\right], \quad (17)$$

where $y^k(t,\omega_p)$ represents the output from the kth order recursive system described by (5) or (7).

The differentiation of $y^k(t,\omega_p)$ with respect to time propagates over in the transfer function of the recursive system as a multiplication by s in the Laplace domain. The new frequency coordinate in (16) can therefore be found by applying the *frequency reassignment filter*

$$H^k_{\mathcal{D}}(s) = sH^k(s), \quad (18)$$

to the input signal $x(t)$, and then use the output in the denominator of (16). The differentiation of $y^k(t,\omega_p)$ with respect to frequency also yields a simple relation. The new time coordinate in (17) can be found by applying the *time reassignment filter*

$$H^k_{\mathcal{T}}(s) = -i\frac{\partial}{\partial s}H^k(s), \quad (19)$$

to the input signal $x(t)$, and then use the output in the denominator of (17). The kth order recursively reassigned spectrogram $y^k_R(t,\omega_p)$ can be expressed as

$$y^k_R(t,\omega) = \int\int |y^k(t',\omega')|^2 \delta\left[t - \hat{t_g}(t',\omega')\right] \cdot\delta\left[\omega - \hat{\omega_i}(t',\omega')\right] dt'd\omega', \quad (20)$$

where $\delta[\cdot]$ represents the Kronecker delta function.

A. Discretizing the Recursive Reassignment Filters

Using the same discretization method on (18) and (19) yields corresponding digital transfer functions. The relation between the *frequency reassignment filter* (18) and the original filter (5) is in the discrete case given by

$$H^k_{\mathcal{D}}(z) = \sigma_p H^{k-1}(z) - p^\star H^k(z), \quad (21)$$

while the relation between the *time reassignment filter* (19) and the original filter (5) is found to be

$$H^k_{\mathcal{T}}(z) = \frac{ik}{\sigma_p}H^{k+1}(z). \quad (22)$$

The discrete reassignment filters are all related to three systems of orders $k-1$, k and $k+1$. This means that Table (I) or equation (12) can be used to find all the needed filter coefficients, and then equations (21) and (22) can be applied on the outputs. When discretized, the reassignment coordinates in (16) and (17) becomes

$$\hat{\omega_i}(n,\omega_p) = \text{Im}\left[\frac{y^k_{\mathcal{D}}(n,\omega_p)}{\Omega y^k(n,\omega_p)}\right] \quad (23)$$

and

$$\hat{t_g}(n,\omega_p) = n - \text{Im}\left[\frac{y^k_{\mathcal{T}}(n,\omega_p)}{T y^k(n,\omega_p)}\right], \quad (24)$$

where Ω denotes the sampling period in frequency (i.e. the uniform difference between adjacent chosen frequencies ω_p), while $y^k_{\mathcal{D}}(n,\omega_p)$ and $y^k_{\mathcal{T}}(n,\omega_p)$ represents the outputs after (21) and (22) are applied, respectively.

It should also be noted that the complex divison in (23) and (24) can be avoided by multiplying both numerators and denominators by the complex conjugate of $y^k(n,\omega_p)$.

The discrete kth order recursively reassigned spectrogram $y^k_R(n,\omega_p)$ can now be expressed as

$$y^k_R(n,\omega_p) = \sum_{n'}\sum_{\omega'_p} |y(n',\omega'_p)|^2 \delta\left[n - \hat{t_g}(n',\omega'_p)\right] \cdot\delta\left[\omega_p - \hat{\omega_i}(n',\omega'_p)\right]. \quad (25)$$

Equation (25) simply means to map and accumulate each point in $|y(n,\omega_p)|^2$ onto the new coordinates defined by (23) and (24).

IV. Computational Cost and Numerical Results

The recursively reassigned STFT computed by the method presented in the previous section costs $O(\Omega k N)$ operations, where Ω is the cardinality of the set of chosen frequencies, the magnitude of the constant k controls the time-frequency resolution, and N is the number of samples in the analyzed signal. This can be compared to the classically reassigned STFT spectrogram using the method of Auger and Flandrin [2]. Its cost is either $O(\Omega N \log N)$ or $O(N\Omega \log \Omega)$.

The recursive time-frequency reassignment algorithm presented in this paper makes time-frequency reassignment well suited for real-time implementations. As seen by (14), a cyclic buffer of $k\Omega$ units is the only required storage space.

The recursive reassignment method is also seen to eliminate the GD introduced by the recursive filters. The GD of (12) is added to the GD of the input signal when it is passed through the system. Since the partial phase derivative w.r.t. frequency is nothing else than the "local" GD of the input signal, it follows from (17) that it is subtracted *to yield zero GD* in the TFRs.

A MATLAB implementation of the recursively reassigned spectrogram can be found online at `http://www.ii.uib.no/~geirkn/rrspec/`

The synthetic signal used in the following figures is a superimposition of two sinusoids close in time, two sinusoids close in frequency, and a quadratic chirp. Figure (1) shows the *ideal* time-frequency representation for this synthetic signal. Note that this is just an illustration made from the a priori known times and frequencies. Figure (2) shows the 4th order reassigned spectrogram. This can be compared to the classical reassigned spectrogram using the Hanning window in Figure (3).

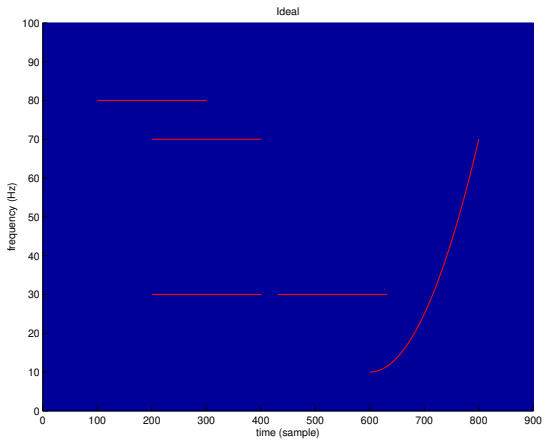


Fig. 1. The ideal TFR.

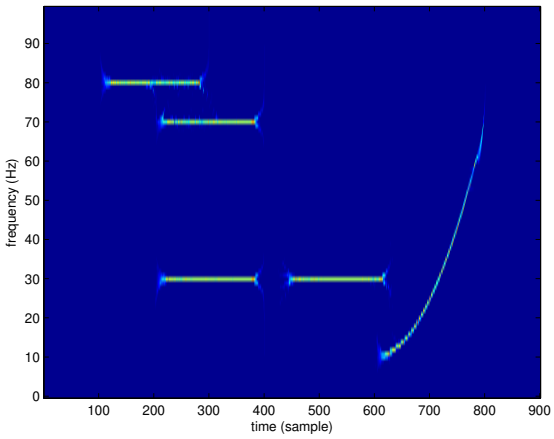


Fig. 2. The 4th order (k=4) reassigned spectrogram.

V. Conclusion

A fast algorithm for computing a special case of the reassigned STFT spectrogram has been presented. The corresponding STFT windows offer a tradeoff between computational cost and time-frequency resolution. The *optimum time-frequency resolution balance* is also found by an explicit formula. The proposed algorithm makes time-frequency reassignment well suited for real-

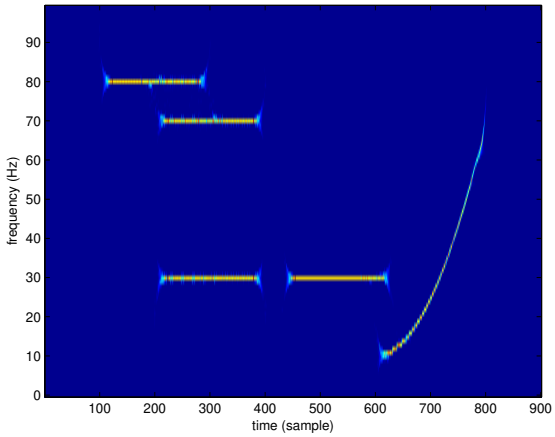


Fig. 3. Classical reassigned spectrogram using the Hanning window.

time implementations. The special windows have been pointed out by Chen et al. [7] to overcome the boundary effects associated with finite windows. As a consequence, undesired frequency-domain sidelobes will vanish. Hainsworth [18] used Cramer-Rao Bound methods to show that the GD and IF estimators using Auger and Flandrin's [2] method introduce a bias when they are discretized. This was not predicted in their theory because it was derived in continuous mathematics. It leads to the contradictory claim that the discretized IF and GD estimators are dependent on the window. Further research is required to find out what impact the special windows in this paper has on the proposed IF and GD estimators.

Geir K. Nilsen November 8, 2007

ACKNOWLEDGMENT

The author would like to thank Prof. Hans Munthe-Kaas for his help and support. Without him this paper had never been written. Also thanks to Øyvind Lunde Rørtveit for their ever-ready attitude to help and discuss problems.

REFERENCES

[1] B. Friedlander and B. Porat, "Detection of transient signals by the gabor representation," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, pp. 169–180, 1989.

[2] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Process.*, vol. 43, pp. 1068–1089, 1995.

[3] S. Tomažic and S. Žnidar, "A fast recursive stft algorithm," *Electrotechnical Conference, 1996. MELECON '96., 8th Mediterranean*, vol. 2, pp. 1025–1028, 1996.

[4] S. Žnidar, "Cosine-modulated perfect reconstruction iir filter bank with good time-frequency resolution," *Digital Signal Processing Proceedings, 1997. DSP 97., 1997 13th International Conference on*, vol. 2, pp. 1067–1070, 1997.

[5] S. Tomažic, "On short-time fourier transform with single-sided exponential window," *Signal Process.*, vol. 55, no. 2, pp. 141–148, 1996.

[6] M. G. Amin and K. Di Feng, "Short-Time Fourier Transforms Using Cascade Filter Structures," *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 42, pp. 631–641, 1995.

[7] T. W. S. W. Chen, N. Kehtarnavaz, "An efficient recursive algorithm for time-varying fourier transform," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2488–2490, 1993.

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

[9] K. F. Sean A. Fulop, "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications," *J. Acoust. Soc. Am.*, vol. 119 (1), pp. 360–371, 2005.

[10] R. G. K. Kodera, C. de Villedary, "A new method for the numerical analysis of non-stationary signals," *Phys. Earth Planet. Inter.*, vol. 12, pp. 142–150, 1976.

[11] D. J. Nelson, "Cross-spectral methods for processing speech," *J. Acoust. Soc. Am.*, vol. 110, pp. 2575–2592, 2001.

[12] A. W. Rihaczek, "Signal energy distribution in time and frequency," *IEEE Transactions on Information Theory*, vol. IT-14 NO.3, pp. 369–374, 1967.

[13] K. et al., "Analysis of time-varying signals with small bt values," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-26 (1), pp. 64–76, 1978.

[14] E. C.-M. F. Auger, P. Flandrin, "Time-frequency reassignment: From principles to algorithms," *Applications in Time-Frequency Signal Processing*, vol. Ch. 5, pp. 179–203, 2003.

[15] L. Cohen, "Time-frequency distributions: a review," *Proc. of IEEE*, pp. 941–979, July 1989.

[16] K. F. Sean A. Fulop, "A unified theory of time-frequency reassignment," *Unpublished Preprint submitted to Digital Signal Processing*, 2005.

[17] S. W. Hainsworth, "Analysis of musical audio for polyphonic transcription," 2001.

[18] ——, "Time frequency reassignment: A review and analysis," 2001.

[19] B. G. Lawrence R. Rabiner, *Theory and Application of DIGITAL SIGNAL PROCESSING*. Prentice-Hall, 1975.

[20] H. I. H. Suzuki, F. Ma and O. Yamazaki, "Instantaneous frequencies of signals obtained by the analytic signal method," *Acoust. Sci. & Tech.*, vol. 27, pp. 163–170, 2006.

[21] M. H. Ackroyd, "Short-time spectra and time-frequency energy distributions," *The Journal of the Acoustical Society of America*, vol. 50, pp. 1229–1231, 1971.

[22] J. M. Tribolet, "A new phase unwrapping algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-25 (2), 1977.

[23] J. O. Smith, *Introduction to Digital Filters, August 2006 Edition*, May 2006.

[24] M. Unser, "Fast Gabor-like windowed Fourier and continuous wavelet transforms," *IEEE Sig. Proc. Letters,*, vol. 1, May 1994

Bibliography

[1] Ackroyd, M. H.: "Short-time spectra and time-frequency energy distributions". *The Journal of the Acoustical Society of America*, **50**, pp. 1229–1231, 1971.

[2] Amin, M. G. and Feng, K.: "Short-time fourier transforms using cascade filter structures". *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, **42**, pp. 631–641, 1995.

[3] Auger, F. and Flandrin, P.: "Improving the readability of time-frequency and time-scale representations by the reassignment method". *IEEE Trans. Signal Process.*, **43**, pp. 1068–1089, 1995.

[4] Boashash, B.: *Time Frequency Signal Analysis and Processing: A comprehensive Reference*. Elsevier, 2003.

[5] Brown, J. C.: "Calculation of a constant q spectral transform". *J. Acoust. Soc. Am.*, **89**, pp. 425–434, 1991.

[6] Cohen, L.: "Time-frequency distributions: a review". *Proc. of IEEE*, pp. 941–979, July 1989.

[7] Cooley, J. W. and Tukey, J. W.: "An algorithm for the machine calculation of complex fourier series". *Math. Comput.*, **19**, pp. 297–301, 1965.

[8] F. Auger, P. Flandrin, E. C.-M.: "Time-frequency reassignment: From principles to algorithms". *Applications in Time-Frequency Signal Processing*, **Ch. 5**, pp. 179–203, 2003.

[9] F. Hlawatsch, G. F. B.-B.: "Linear and qudratic time-frequency signal representations". *IEEE SP Magazine*, pp. 21–67, 1992.

[10] F. Plante, G. Meyer, W. A. A.: "Speech signal analysis with reallocated spectrogram". *Proceedings of the IEEE-SP International Symposium on TimeFrequency and Time-Scale Analysis*, pp. 640–643, 1994.

[11] Flanagan, J. L. and Golden, R. M.: "Phase vocoder". *The Bell System Technical Journal*, pp. 1493–1509, November 1966. `http://www.ee.columbia.edu/~dpwe/e6820/papers/FlanG66.pdf`.

[12] Friedman, D. H.: "Instantaneous-frequency distribution vs. time: An interpretation of the phase structure of speech". *IEEE*, pp. 1121–1124, 1985.

[13] Gabor, D.: "Theory of communication". *J. IEE (London)*, **93(III)**, pp. 429–457, 1946.

[14] Gonzalez, R. C. and Woods, R. E.: *Digital Image Processing, 3rd Ed.*. Prentice Hall, 2006.

[15] H. Suzuki, F. Ma, H. I. and Yamazaki, O.: "Instantaneous frequencies of signals obtained by the analytic signal method". *Acoust. Sci. & Tech.*, **27**, pp. 163–170, 2006.

[16] Hainsworth, S. W.: "Analysis of musical audio for polyphonic transcription". 2001.

[17] Hainsworth, S. W.: "Time frequency reassignment: A review and analysis". 2001.

[18] Heisenberg, W.: "Uber den anschaulichen inhalt der quantentheoretischen kinematik und mechanik". *Zeitschrift fur Physik*, **43**, pp. 172–198, 1927.

[19] Hsu, H. P.: *Applied Fourier Analysis*. Harcourt Brace Jovanovich, 1984.

[20] Johansson, M.: "The hilbert transform". *Master Thesis*, ?

[21] K. Kodera, C. de Villedary, R. G.: "A new method for the numerical analysis of non-stationary signals". *Phys. Earth Planet. Inter.*, **12**, pp. 142–150, 1976.

[22] Lawrence R. Rabiner, B. G.: *Theory and Application of DIGITAL SIGNAL PROCESSING*. Prentice-Hall, 1975.

[23] Nelson, D. J.: "Cross-spectral methods for processing speech". *J. Acoust. Soc. Am.*, **110**, pp. 2575–2592, 2001.

[24] Orfanidis, S. J.: *Introduction to signal processing*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1995.

[25] Potter, R. K.: "Visible patterns of sound". *SCIENCE*, **102**, pp. 463–470, 1945.

[26] Rihaczek, A. W.: "Signal energy distribution in time and frequency". *IEEE Transactions on Information Theory*, **IT-14 NO.3**, pp. 369–374, 1967.

[27] Rioul, O. and Flandrin, P.: "Time-scale energy distributions: A general class extending wavelet transforms". *IEEE Trans. Signal Process.*, **40**, pp. 1746–1757, 1992.

[28] Sean A. Fulop, K. F.: "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications". *J. Acoust. Soc. Am.*, **119 (1)**, pp. 360–371, 2005.

[29] Sean A. Fulop, K. F.: "A unified theory of time-frequency reassignment". *Unpublished Preprint submitted to Digital Signal Processing*, 2005.

[30] Selin, I.: "Theory and applications of the notion of complex signal". **T-92**, 1958.

[31] Smith, J. O.: *Introduction to Digital Filters, August 2006 Edition*. May 2006.

[32] Smith, S. W.: *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Publishing, first edition, 1997.

[33] Thomas H. Crystal, L. E.: "The design and applications of digital filters with complex coefficients". 1968.

[34] Tomažic, S.: "On short-time fourier transform with single-sided exponential window". *Signal Process.*, **55**(2), pp. 141–148, 1996.

[35] Tomažic, S. and Žnidar, S.: "A fast recursive stft algorithm". *Electrotechnical Conference, 1996. MELECON '96., 8th Mediterranean*, **2**, pp. 1025–1028, 1996.

[36] Torrence, C. and Compo, G. P.: "A practical guide to wavelet analysis". *Bulletin of the American Meteorological Society*, **79**(1), pp. 61–78, 1998. `citeseer.ist.psu.edu/article/torrence98practical.html`.

[37] Tribolet, J. M.: "A new phase unwrapping algorithm". *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-25 (2)**, 1977.

[38] Ville, J.: "ThÃ©orie et applications de la notion de signal analytique." **2a**, pp. 61–74, 1948.

[39] Žnidar, S.: "Cosine-modulated perfect reconstruction iir filter bank with good time-frequency resolution". *Digital Signal Processing Proceedings, 1997. DSP 97., 1997 13th International Conference on*, **2**, pp. 1067–1070, 1997.

[40] W. Chen, N. Kehtarnavaz, T. W. S.: "An efficient recursive algorithm for time-varying fourier transform". *IEEE Transactions on Signal Processing*, **41**, pp. 2488–2490, 1993.

[41] Weisstein, E. W.: "Danielson-lanczos lemma". *Matworld - A Wolfram Web Resource. http://mathworld.wolfram.com/Danielson-LanczosLemma.html*.

[42] Wheeler, J. A. and Zurek, H.: "Quantum theory and measurement". *Princeton Univ. Press*, pp. 62–84, 1983.

[43] Wigner, E.: "On the quantum correction for thermodynamic equilibrium". *Physical Review*, **40**, pp. 749–759, 1932.

[44] Youngberg, J. E.: "Constant-q signal analysis and synthesis". *IEEE Trans. Signal Process.*, **78**, pp. 375–378, 1978.

[45] Ziemer, Tranter, F.: *Signals & Systems: Continuous and Discrete Fourth Edition*. Prentice Hall, 1998.