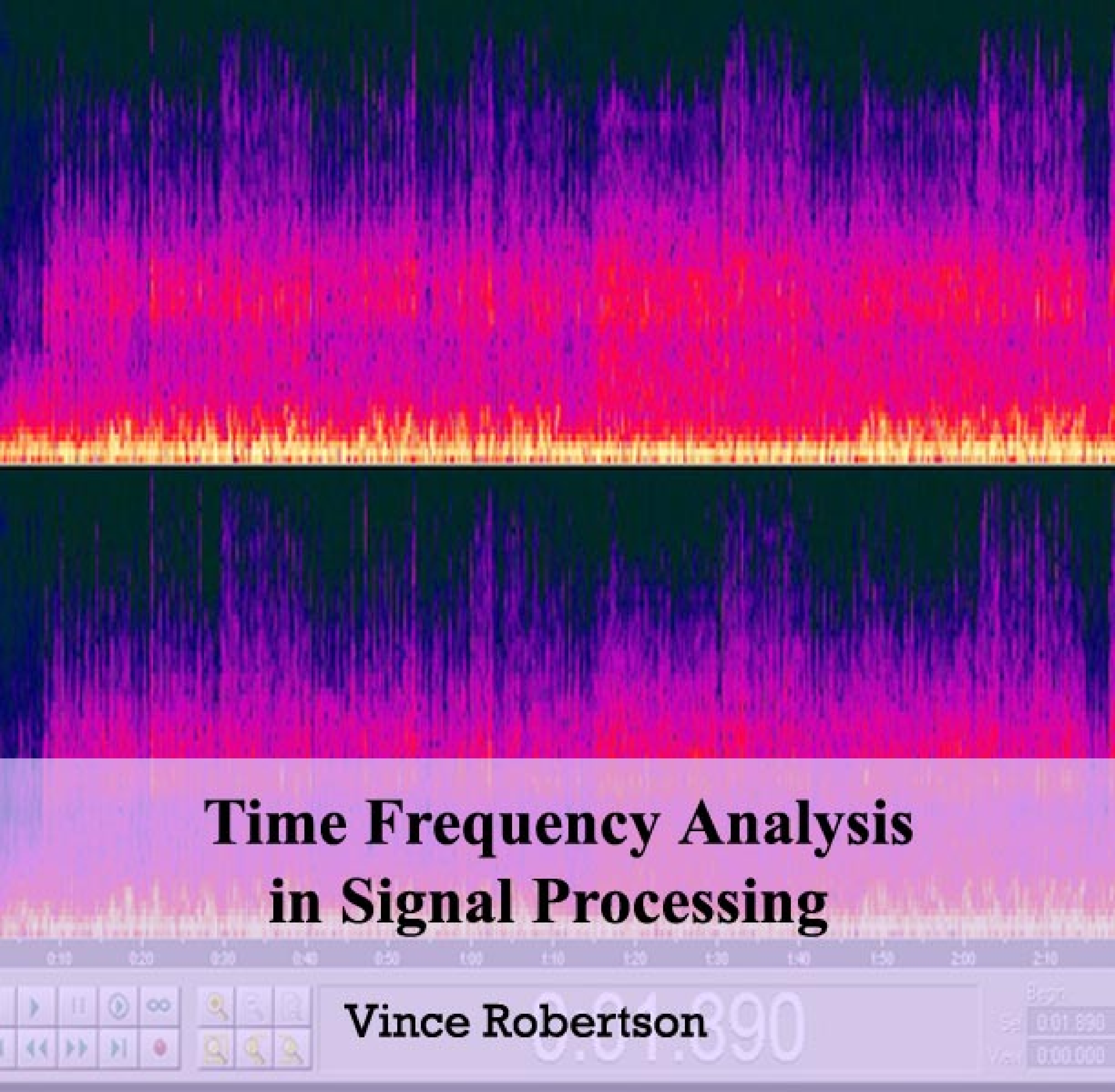


Time Frequency Analysis in Signal Processing

Vince Robertson

First Edition, 2012

ISBN 978-81-323-3110-0

Published by:
Research World
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter 1

Time Frequency Analysis

Need for a time–frequency approach

In signal processing, **time–frequency analysis** is a body of techniques and methods used for characterizing and manipulating signals whose statistics vary in time, such as transient signals.

It is a generalization and refinement of Fourier analysis, for the case when the signal frequency characteristics are varying with time. Since many signals of interest – such as speech, music, images, and medical signals – have changing frequency characteristics, time–frequency analysis has broad scope of applications.

Whereas the technique of the Fourier transform can be extended to obtain the frequency spectrum of any slowly growing locally integrable signal this requires a complete description of the signal's behavior over all time. Indeed, one can think of points in the (spectral) frequency domain as smearing together information from across the entire time domain. While mathematically elegant such a technique is not appropriate for analyzing a signal with indeterminate future behavior. For instance one must presuppose some degree of indeterminate future behavior in any telecommunications systems to achieve non-zero entropy (if you already know what the other guy will say you can't learn anything).

To harness the power of a frequency representation without the need of a complete characterization in the time domain, one first obtains a time–frequency distribution of the signal, which represents the signal in both the time and frequency domains simultaneously. In such a representation the frequency domain will only reflect the behavior of a temporally localized version of the signal. This enables one to talk sensibly about signals whose component frequencies vary in time.

For instance rather than using tempered distributions to globally transform the following function into the frequency domain one could instead use these methods to describe it as a signal with a time varying frequency.

$$x(t) = \begin{cases} \cos(\pi t); & t < 10 \\ \cos(3\pi t); & 10 \le t < 20 \\ \cos(2\pi t); & t > 20 \end{cases}$$

Once such a representation has been generated other techniques in time–frequency analysis may then be applied to the signal in order to extract information from the signal, to separate the signal from noise or interfering signals, etc.

Time–frequency distribution functions

Diversity of time–frequency formulations

There are several different ways to formulate a valid time–frequency distribution function, resulting in several well-known time–frequency distributions, such as:

- Short-time Fourier transform (including the Gabor transform),
- Wavelet transform,
- Bilinear time–frequency distribution function (Wigner distribution function),
- Modified Wigner distribution function, Gabor–Wigner distribution function, and so on.

More information about the history and the motivation of development of time–frequency distribution can be found in the entry Time–frequency representation.

Ideal TF distribution function

A time–frequency distribution function ideally has the following properties:

1. **High clarity** to make it easier to be analyzed and interpreted.
2. **No cross-term** to avoid confusing real components from artefacts or noise.
3. **A list of desirable mathematical properties** to ensure such methods benefit real-life application.
4. **Lower computational complexity** to ensure the time needed to represent and process a signal on a time–frequency plane allows real-time implementations.

Below is a brief comparison of some selected time–frequency distribution functions.

	Clarity	Cross-term	Good mathematical properties	Computational complexity
Gabor transform	Worst	No	Worst	Low

Wigner distribution function	Best	Yes	Best	High
Gabor-Wigner distribution function	Good	Almost eliminated	Good	High

To analyze the signals well, choosing an appropriate time–frequency distribution function is important. Which time–frequency distribution function should be used depends on the application being considered, as shown by reviewing a list of applications. The high clarity of the Wigner distribution function (WDF) obtained for some signals is due to the auto-correlation function inherent in its formulation; however, the latter also causes the cross-term problem. Therefore, if we want to analyze a single-term signal, using the WDF may be the best approach; if the signal is composed of multiple components, some other methods like the Gabor transform, Gabor-Wigner distribution or Modified B-Distribution functions may be better choices.

To illustrate this, we observe that by Fourier analysis, we can't recognize the two signals $x_1(t)$ and $x_2(t)$ below.

$$x_1(t) = \begin{cases} \cos(\pi t); & t < 10 \\ \cos(3\pi t); & 10 \le t < 20 \\ \cos(2\pi t); & t > 20 \end{cases}$$

$$x_2(t) = \begin{cases} \cos(\pi t); & t < 10 \\ \cos(2\pi t); & 10 \le t < 20 \\ \cos(3\pi t); & t > 20 \end{cases}$$

Thanks to the time–frequency analysis approach, we can still solve this problem of correctly identifying the two different signals.

Signal processing applications

The following applications need not only the time–frequency distribution functions but also some operations to the signal. The Linear canonical transform (LCT) is really helpful. By LCTs, the shape and location on the time–frequency plane of a signal can be in the arbitrary form that we want it to be. For example, the LCTs can shift the time–frequency distribution to any location, dilate it in the horizontal and vertical direction without changing its area on the plane, shear (or twist) it, and rotate it (Fractional Fourier transform). This powerful operation, LCT, make it more flexible to analyze and apply the time–frequency distributions. Here we list some applications of time–frequency analysis.

Instantaneous frequency estimation

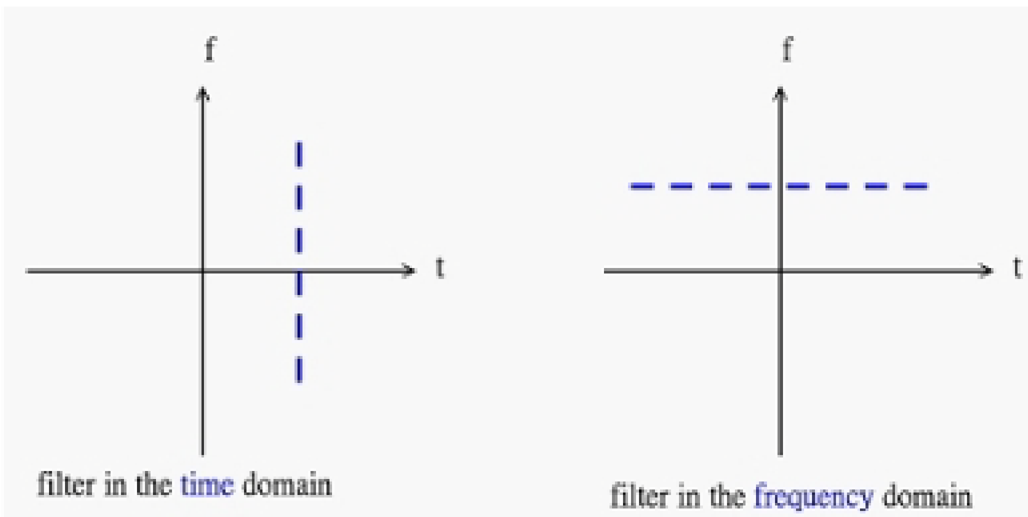
The definition of instantaneous frequency is the time rate of change of phase, or

$$\frac{1}{2\pi}\frac{d}{dt}\phi(t),$$

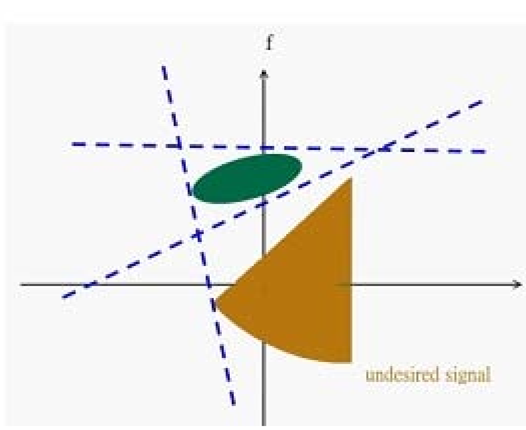
where φ(t) is the instantaneous phase of a signal. We can know the instantaneous frequency from the time–frequency plane directly if the image is clear enough. Because the high clarity is critical, we often use WDF to analyze it.

TF filtering and signal decomposition

The goal of filter design is to remove the undesired component of a signal. Conventionally, we can just filter in the time domain or in the frequency domain individually as shown below.



The filtering methods mentioned above can't work well for every signal which may overlap in the time domain or in the frequency domain. By using the time–frequency distribution function, we can filter in the Euclidian time–frequency domain or in the fractional domain by employing the fractional Fourier transform. An example is shown below.

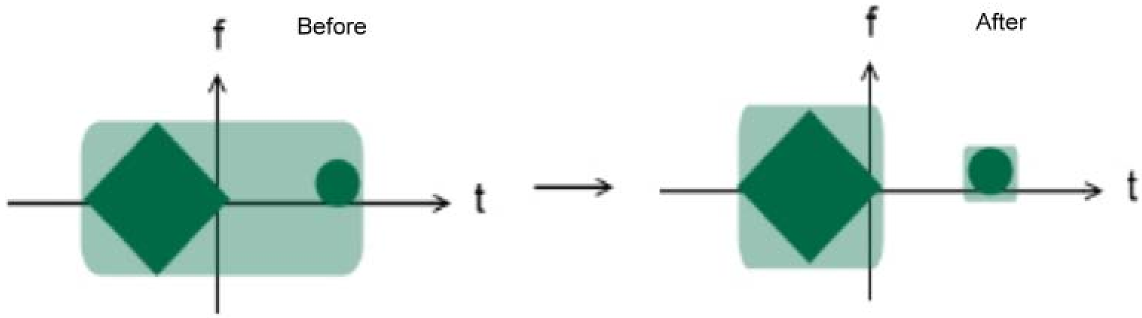


Filter design in time–frequency analysis always deals with signals composed of multiple components, so one cannot use WDF due to cross-term. The Gabor transform, Gabor-Wigner distribution function, or Cohen's class distribution function may be better choices.

The concept of signal decomposition relates to the need to separate one component from the others in a signal; this can be achieved through a filtering operation which require a filter design stage. Such filtering is traditionally done in the time domain or in the frequency domain; however, this may not be possible in the case of non-stationary signals that are multicomponent as such components could overlap in both the time domain and also in the frequency domain; as a consequence, the only possible way to achieve component separation and therefore a signal decomposition is to implement a time–frequency filter.

Sampling theory

By the Nyquist–Shannon sampling theorem, we can conclude that the minimum number of sampling points without aliasing is equivalent to the area of the time–frequency distribution of a signal. (This is actually just an approximation, because the TF area of any signal is infinite.) Below is an example before and after we combine the sampling theory with the time–frequency distribution:



It is noticeable that the number of sampling points decreases after we apply the time–frequency distribution.

When we use the WDF, there might be the cross-term problem (also called interference). On the other hand, using Gabor transform causes an improvement in the clarity and readability of the representation, therefore improving its interpretation and application to practical problems.

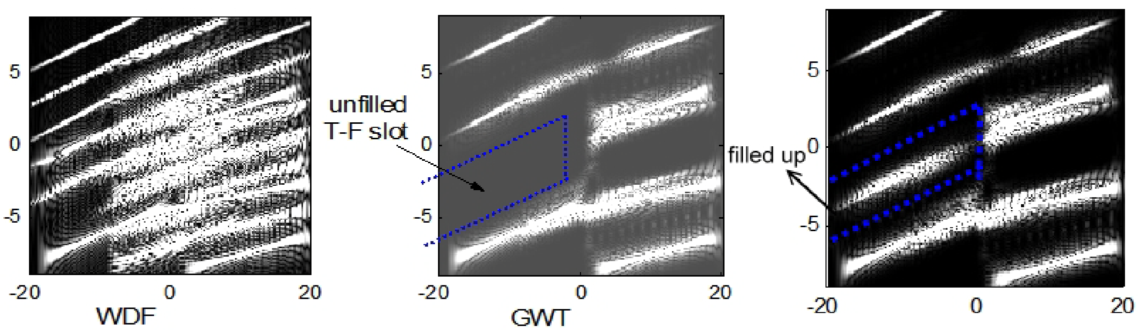
Consequently, when the signal we tend to sample is composed of single component, we use the WDF; however, if the signal consists of more than one component, using the Gabor transform, Gabor-Wigner distribution function, or other reduced interference TFDs may achieve better results.

The Balian–Low theorem formalizes this, and provides a bound on the minimum number of time-frequency samples needed.

Other applications

Modulation and multiplexing

Conventionally, the operation of modulation and multiplexing concentrates in time or in frequency, separately. By taking advantage of the time–frequency distribution, we can make it more efficient to modulate and multiplex. All we have to do is to fill up the time–frequency plane. We present an example as below.



As illustrated in the upper example, using the WDF is not smart since the serious cross-term problem make it difficult to multiplex and modulation.

Electromagnetic wave propagation

We can represent an electromagnetic wave in the form of a 2 by 1 matrix

$$\begin{bmatrix} x \\ y \end{bmatrix},$$

which is similar to the time–frequency plane. When electromagnetic wave propagates through free-space, the Fresnel diffraction occurs. We can operate with the 2 by 1 matrix

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

by LCT with parameter matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & \lambda z \\ 0 & 1 \end{bmatrix},$$

where z is the propagation distance and λ is the wavelength. When electromagnetic wave pass through a spherical lens or be reflected by a disk, the parameter matrix should be

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{-1}{\lambda f} & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{\lambda R} & 1 \end{bmatrix}$$

respectively, where f is the focal length of the lens and R is the radius of the disk. These corresponding results can be obtained from

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Optics, acoustics, and biomedicine

Light is a kind of electromagnetic wave, so we apply the time–frequency analysis to optics in the same way as to electromagnetic wave propagation. In the same way, a characteristic of acoustic signals is that, often, its frequency varies really severely with time. Because the acoustic signals usually contain a lot of data, it is suitable to use simpler TFDs such as the Gabor transform to analyze the acoustic signals due to the lower computational complexity. If speed is not an issue, then a detailed comparison with well defined criteria should be made before selecting a particular TFD. Another approach is to define a signal dependent TFD that is adapted to the data. In biomedicine, one can use time–frequency distribution to analyze the electromyography (EMG), Electroencephalography (EEG), Electrocardiogram (ECG) or otoacoustic emissions (OAEs).

Chapter 2

Bilinear Time–Frequency Distribution

Bilinear time-frequency distributions, or **quadratic time-frequency distributions**, arise in a sub-field field of signal analysis and signal processing called time-frequency signal processing, and, in the statistical analysis of time series data. Such methods are used where one needs to deal with a situation where the frequency composition of a signal may be changing over time; this sub-field used to be called time-frequency signal analysis, and is now more often called time-frequency signal processing due to the progress in using these methods to a wide range of signal processing problems.

Background

Methods for analysing time series, in both signal analysis and time series analysis, have been developed as essentially separate methodologies applicable to, and based in, either the time or the frequency domain. A mixed approach is required in time-frequency analysis techniques which are especially effective in analyzing non-stationary signals, whose frequency distribution and magnitude vary with time. Examples of these are acoustic signals. Classes of "quadratic time-frequency distributions" (or bilinear time-frequency distributions") are used for time-frequency signal analysis. This class is similar in formulation to Cohen's class distribution function that was used in 1966 in the context of quantum mechanics. This distribution function is mathematically similar to a generalized time-frequency representation which utilizes bilinear transformations. Compared with other time-frequency analysis techniques, such as short-time Fourier transform (STFT), the bilinear-transformation (or Quadratic Time-Frequency Distributions) may not have higher clarity for most practical signals, but it provides an alternative framework to investigate new definitions and new methods. While it does suffer from an inherent cross-term contamination when analyzing multi-component signals, by using a carefully chosen window functions, the interference can be significantly mitigated, at the expense of resolution.

Mathematical definition

The definition of the class of bilinear (or quadratic) time-frequency distributions is as follows:

$$C_x(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A_x(\eta,\tau)\Phi(\eta,\tau)\exp(-j2\pi(\eta t+\tau f))\,d\eta\,d\tau,$$

where $A_x(\eta, \tau)$ is the ambiguity function (AF) which will be further discussed later, and $\Phi(\eta, \tau)$ is the kernel function which is usually a low-pass function and is used to mask out the interference.

Ambiguity function

The class of bilinear (or quadratic) time-frequency distributions can be most easily understood in terms of the ambiguity function an explanation of which follows.

Consider the well known power spectral density $P_x(f)$ and the signal auto-correlation function $R_x(\tau)$ in the case of a stationary process. The relationship between these functions is as follows:

$$P_x(f) = \int_{-\infty}^{\infty} R_x(\tau)e^{-j2\pi f\tau}\,d\tau,$$
$$R_x(\tau) = E\left[x(t+\tau/2)x(t-\tau/2)\right].$$

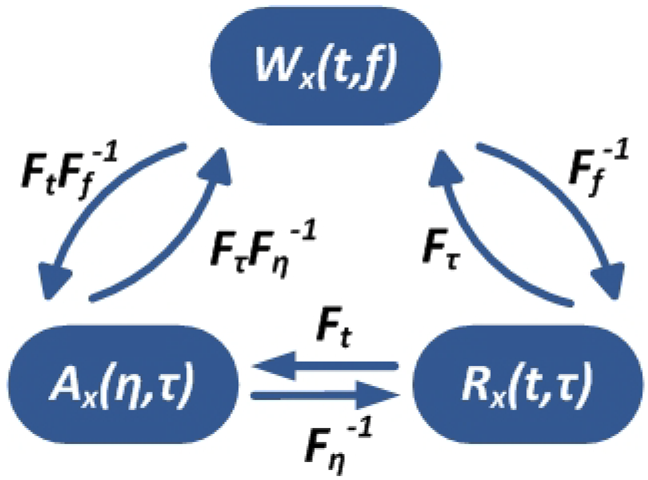
For a non-stationary signal $x(t)$, these relations can be generalized using a time-dependent power spectral density or equivalently the famous Wigner distribution function of $x(t)$ as follows:

$$W_x(t,f) = \int_{-\infty}^{\infty} R_x(t,\tau)e^{-j2\pi f\tau}\,d\tau,$$
$$R_x(t,\tau) = x(t+\tau/2) * x(t-\tau/2).$$

If the Fourier transform of the auto-correlation function is taken with respect to *t* instead of τ, we get the ambiguity function as follows:

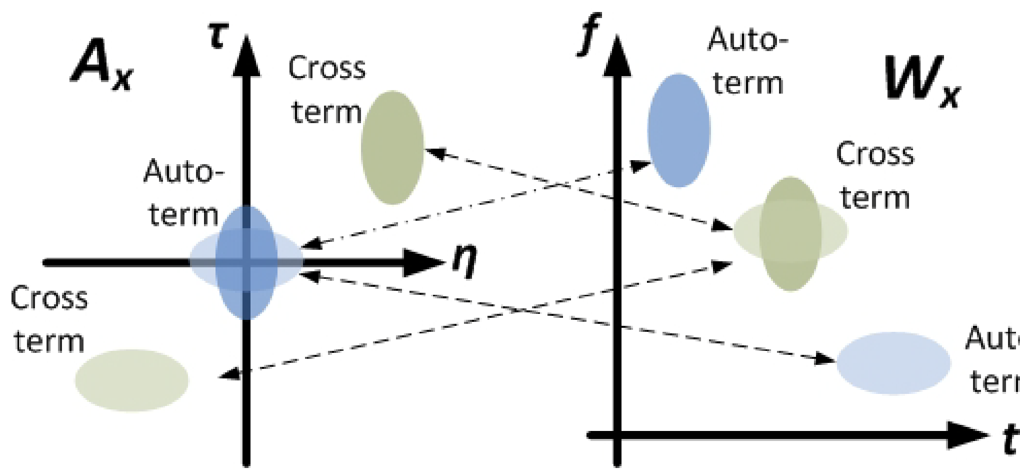
$$A_x(\eta,\tau) = \int_{-\infty}^{\infty} x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi t\eta}\,dt.$$

The relationship between the Wigner distribution function, the auto-correlation function and the ambiguity function can then be illustrated by the following figure.

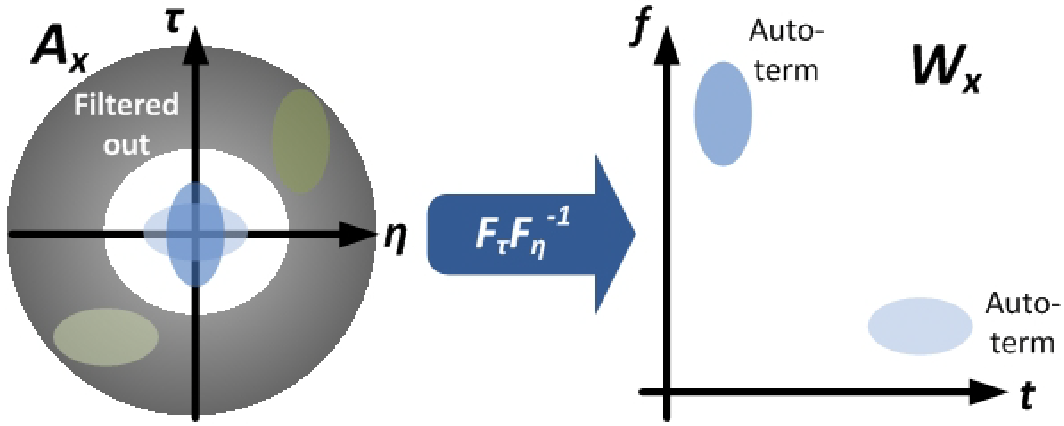


By comparing the definition of bilinear (or quadratic) time-frequency distributions with that of the Wigner distribution function, it is easily found that the latter is a special case of the former with $\Phi\left(\eta,\tau\right)=1$. Alternatively, bilinear (or quadratic) time-frequency distributions can be regarded as a masked version of the Wigner distribution function if a kernel function $\Phi\left(\eta,\tau\right)\neq 1$is chosen. A properly chosen kernel function can significantly reduce the undesirable cross-term of the Wigner distribution function.

What is the benefit of the additional kernel function? The following figure shows the distribution of the auto-term and the cross-term of a multi-component signal in both the ambiguity and the Wigner distribution function.



For multi-component signals in general, the distribution of its auto-term and cross-term within its Wigner distribution function is generally not predictable, and hence the cross-term cannot be removed easily. However, as shown in the figure, for the ambiguity function, the auto-term of the multi-component signal will inherently tend to close the origin in the η,τ plane, and the cross-term will tend to be away from the origin. With this property, the cross-term in can be filtered out effortlessly if a proper low-pass kernel function is applied in η,τ domain. The following is an example that demonstrates how the cross-term is filtered out.



Some time-frequency distributions

Wigner distribution function

Aforementioned, the Wigner distribution function is a member of the class of quadratic time-frequency distributions (QTFDs)with the kernel function $\Phi(\eta,\tau)=1$. The definition of Wigner distribution is as follows:

$$W_x(t,f)=\int_{-\infty}^{\infty} x(t+\tau/2)*x(t-\tau/2)e^{-j2\pi f\tau}\,d\tau.$$

Choi–Williams distribution function

The kernel of Choi–Williams distribution is defined as follows:

$$\Phi(\eta,\tau)=\exp\left[-\alpha(\eta\tau)^2\right],$$

where α is an adjustable parameter.

Cone-shape distribution function

The kernel of cone-shape distribution function is defined as follows:

$$\Phi(\eta,\tau)=\frac{\sin(\pi\eta\tau)}{\pi\eta\tau}\exp\left(-2\pi\alpha\tau^2\right),$$

where α is an adjustable parameter.

There are many more such QTFDs and a full list can be found in the next reference.

Chapter 3

Time-Frequency Analysis for Music Signal

Time-frequency analysis for music signal is one of the applications of time-frequency analysis. Musical sound can be more complicated than human vocal sound, occupying a wider band of frequency. Music signals are time-varying signals; while the classic Fourier transform is not sufficient to analyze them, time-frequency analysis is an efficient tool for such use. Time-frequency analysis is extended from the classic Fourier approach. Short-time Fourier transform (STFT), Gabor transform (GT) and Wigner distribution function (WDF) are famous time-frequency methods, useful for analyzing music signals such as notes played on a piano, a flute or a guitar.

Knowledge about music signal

Music is a type of sound that has some stable frequencies in a time period. Music can be produced by several methods. For example, the sound of a piano is produced by striking strings, and the sound of a violin is produced by bowing. All musical sounds have their fundamental frequency and overtones. Fundamental frequency is the lowest frequency in harmonic series. In a periodic signal, the fundamental frequency is the inverse of the period length. Overtones are integer multiples of the fundamental frequency.

Table. 1 the fundamental frequency and overtone

Frequency	**Order**		
$f = 440$ Hz	$N = 1$	Fundamental frequency	1st harmonic
f = 880 Hz	$N = 2$	1st overtone	2nd harmonic
$f = 1320$ Hz	$N = 3$	2nd overtone	3rd harmonic
$f = 1760$ Hz	$N = 4$	3rd overtone	4th harmonic

In musical theory, pitch represents the perceived fundamental frequency of a sound. However the actual fundamental frequency may differ from the perceived fundamental frequency because of overtones.

Short-time Fourier transform

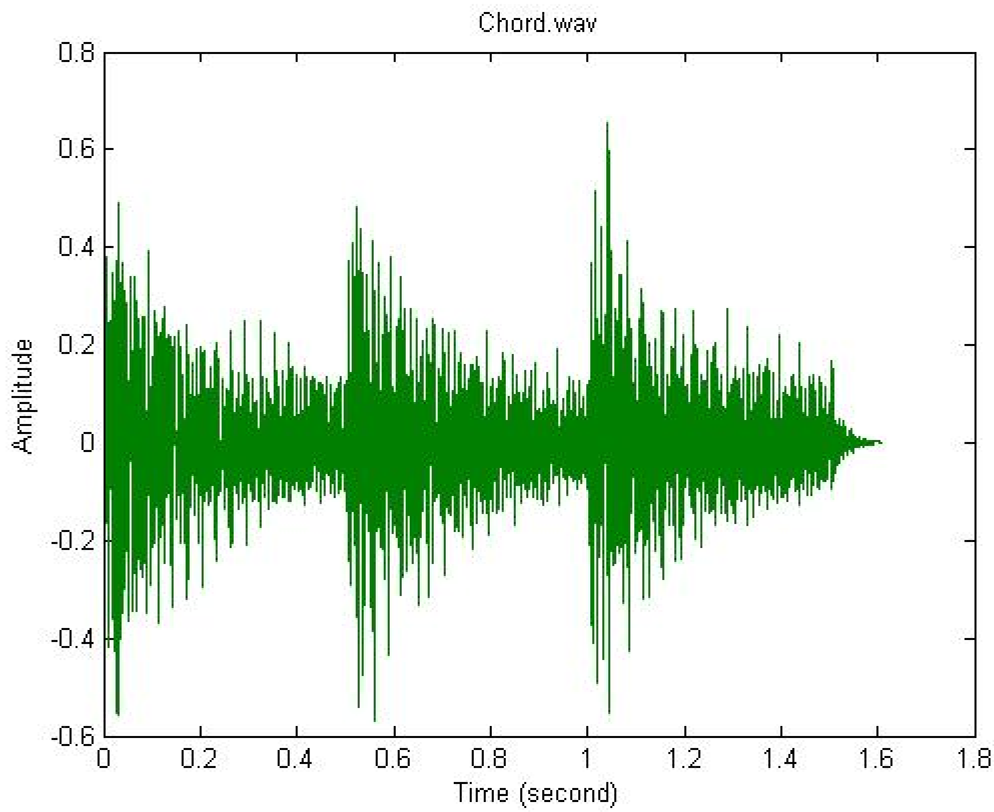


Fig.1 Waveform of the audio file "Chord.wav"

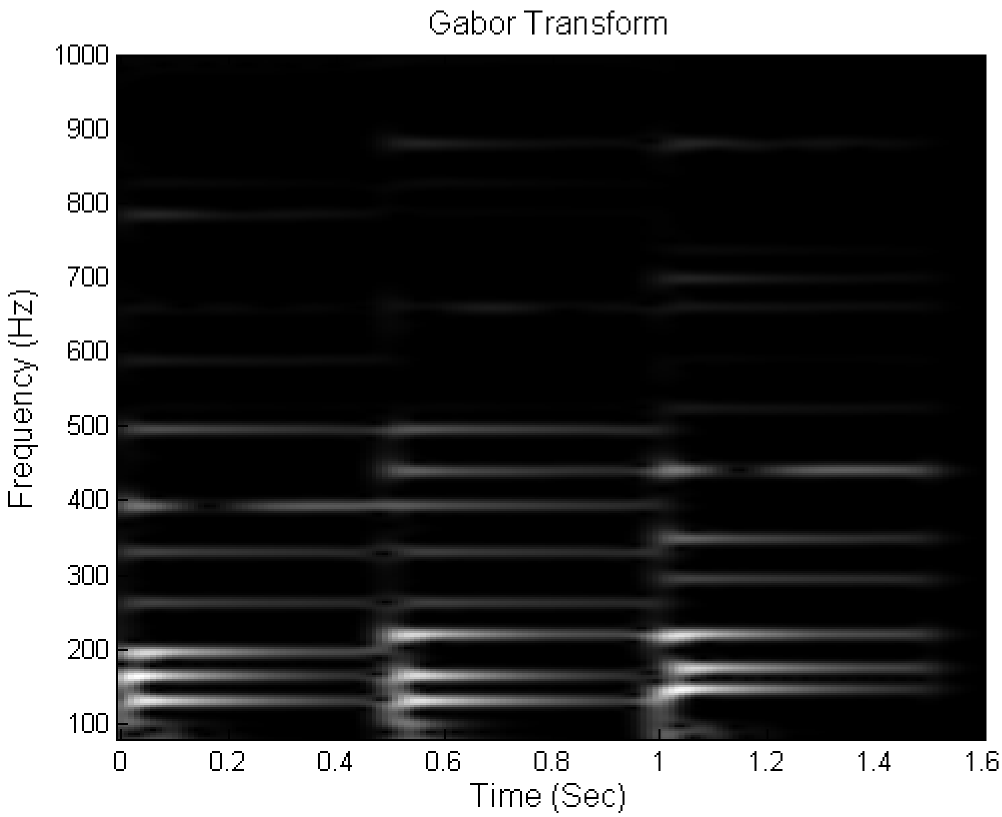


Fig.2 Gabor transform of "Chord.wav"

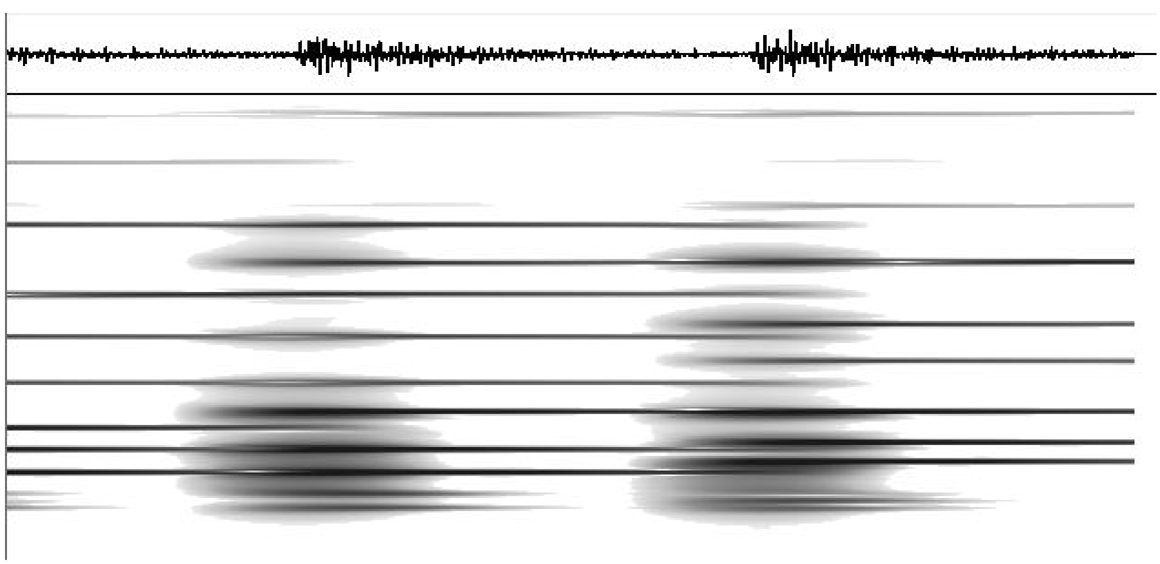


Fig. 3 Spectrogram of "Chord.wav"

Continuous STFT

Short-time Fourier transform is a basic type of time-frequency analysis. If there is a continue signal $x(t)$, we can compute the short-time Fourier transform by

$$\mathbf{STFT}\{x(t)\} \equiv X(t,f) = \int_{-\infty}^{\infty} x(\tau)w(t-\tau)e^{-j2\pi\tau}\,d\tau$$

,where $w(t)$ is a window function. When the $w(t)$ is a rectangular function, the transform is called Rec-STFT. When the $w(t)$ is a Gaussian function, the transform is called Gabor transform.

Discrete STFT

However, normally the musical signal we have is not a continuous signal. It is sampled in a sampling frequency. Therefore, we can't use the formula to compute the Rec-short-time Fourier transform. We change the original form to

$$X(n\,\Delta t, m\,\Delta f) = \sum_{p=n-Q}^{n+Q} x(p\,\Delta t)e^{-j2\pi pm\,\Delta t\,\Delta f}\,\Delta t$$

Let $t = n\,\Delta t$, $f = m\,\Delta f$, $\tau = p\,\Delta t$ and $B = Q\,\Delta t$. There are some constraints of discrete short-time Fourier transform:

- $\Delta t\,\Delta f = \frac{1}{N}$, where N is an integer.
- $N \geq 2Q + 1$
- $\Delta < \frac{1}{2f_{\mathbf{max}}}$, where f_{max} is the highest frequency in the signal.

STFT example

Fig.1 shows the waveform of a piano music audio file with 44100 Hz sampling frequency. And Fig.2 shows the result of short-time Fourior transform (we use Gabor transform here) of the audio file. We can see from the time-frequency plot, from $t = 0$ to 0.5 second, there is a chord with three notes, and the chord changed at $t = 0.5$, and then changed again at $t = 1$. The fundamental frequency of each note in each chord is show in the time-frequency plot.

Spectrogram

Figure 3 shows the spectrogram of the audio file shows in Figure 1. Spectrogram is the square of STFT, time-varying spectral representation. The spectrogram of a signal $s(t)$

can be estimated by computing the squared magnitude of the STFT of the signal $s(t)$, as shown below:

$$\mathbf{spectrogram}(t, f) = |\mathbf{STFT}(t, f)|^2$$

Although the spectrogram is profoundly useful, it still has one drawback. It displays frequencies on a uniform scale. However, musical scales are based on a logarithmic scale for frequencies. Therefore, we should describe the frequency in logarithmic scale related to human hearing.

Wigner distribution function

The Wigner distribution function can also be used to analyze music signal. The advantage of Wigner distribution function is the high clarity. However, it needs high calculation and has cross-term problem, so it's more suitable to analyze signal without more than one frequency at the same time.

Formula

The Wigner distribution function $W_x(t,f)$ is:

$$\mathbf{W}_x(t, f) = \int_{-\infty}^{\infty} x(t + \tau/2)x^*(t - \tau/2)e^{-j2\pi\tau f}\, d\tau,$$

where $x(t)$ is the signal, and $x^*(t)$ is the conjugate of the signal.

Chapter 4

Transformation between Distributions in Time-Frequency Analysis

Introduction

In the field of time-frequency analysis, the goal is to define signal formulations that are used for representing the signal in a joint time-frequency domain. There are several methods and transforms called "time-frequency distributions" (TFDs). The most useful and used methods form a class referred to as "quadratic" or bilinear time-frequency distributions. A core member of this class is the Wigner-Ville distribution (WVD), as all other TFDs can be written as a smoothed version of the WVD. Another popular member of this class is the spectrogram which is the square of the magnitude of the short-time Fourier transform (STFT). The spectrogram has the advantage of being positive and is easy to interpret, but has disadvantages like being irreversible which means that once the spectrogram of a signal is computed, the original signal can't be extracted from the spectrogram. The theory and methodology for defining a TFD that verifies certain desirable properties is given in the "Theory of Quadratic TFDs" . The scope is to outline some elements of the procedure to transform one distribution into another. The method used to transform a distribution is borrowed from quantum mechanics, even though the subject matter of the article is "signal processing". Noting that a signal can recovered from a particular distribution under certain conditions, given a certain TFD $\rho1(t,f)$ representing the signal in a joint time-frequency domain, another different TFD $\rho2(t,f)$ of the same signal can be obtained to calculate any other distribution, by simple smoothing or filtering; some of these relationships are shown below. A full treatment of the question can be given from a signal processing perspective.

General class

If we use the variable $\omega=2\pi f$, then, borrowing the notations used in the field of quantum mechanics, we can show that time-frequency representation, such as Wigner distribution function (WDF) and other bilinear time-frequency distributions, can be expressed as

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

where φ(θ,τ) is a two dimensional function called the kernel, which determines the distribution and its properties (for a signal processing terminology and treatment of this question, the reader is referred to the references already cited in the introduction).

For the kernel of the Wigner distribution function (WDF) is one. However, it is no particular significance should be attached to that since it is to write the general form so that the kernel of any distribution is one, in which case the kernel of the Wigner distribution function (WDF) would be something else.

Characteristic function formulation

The characteristic function is the double Fourier transform of the distribution. By inspection of Eq. (1), we can obtain that

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

where

$$\begin{aligned} M(\theta,\tau) &= \phi(\theta,\tau) \int s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) e^{j\theta u} \, du \\ &= \phi(\theta,\tau) A(\theta,\tau) \end{aligned} \quad (3)$$

and where $A(\theta,\tau)$ is the symmetrical ambiguity function. The characteristic function may be appropriately called the generalized ambiguity function.

Transformation between distributions

To obtain that relationship suppose that there are two distributions, C_1 and C_2, with corresponding kernels, φ_1 and φ_2. Their characteristic functions are

$$M_1(\phi,\tau) = \phi_1(\theta,\tau) \int s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) e^{j\theta u} \, du \quad (4)$$

$$M_2(\phi,\tau) = \phi_2(\theta,\tau) \int s^*(u - \frac{1}{2}\tau) s(u + \frac{1}{2}\tau) e^{j\theta u} \, du \quad (5)$$

Divide one equation by the other to obtain

$$M_1(\phi,\tau) = \frac{\phi_1(\theta,\tau)}{\phi_2(\theta,\tau)} M_2(\phi,\tau) \tag{6}$$

This is an important relationship because it connects the characteristic functions. For the division to be proper the kernel cannot to be zero in a finite region.

To obtain the relationship between the distributions take the double Fourier transform of both sides and use Eq. (2)

$$C_1(t,\omega) = \frac{1}{4\pi^2} \iint \frac{\phi_1(\theta,\tau)}{\phi_2(\theta,\tau)} M_2(\theta,\tau) e^{-j\theta t - j\tau\omega}\, d\theta\, d\tau \tag{7}$$

Now express M_2 in terms of C_2 to obtain

$$C_1(t,\omega) = \frac{1}{4\pi^2} \iiiint \frac{\phi_1(\theta,\tau)}{\phi_2(\theta,\tau)} C_2(t,\omega') e^{j\theta(t'-t)+j\tau(\omega'-\omega)}\, d\theta\, d\tau\, dt'\, d\omega' \tag{8}$$

This relationship can be written as

$$C_1(t,\omega) = \iint g_{12}(t'-t,\omega'-\omega) C_2(t,\omega')\, dt'\, d\omega' \tag{9}$$

with

$$g_{12}(t,\omega) = \frac{1}{4\pi^2} \iint \frac{\phi_1(\theta,\tau)}{\phi_2(\theta,\tau)} e^{j\theta t + j\tau\omega}\, d\theta\, d\tau \tag{10}$$

Relation of the spectrogram to other bilinear representations

Now we specialize to the case where one transform from an arbitrary representation to the spectrogram. In Eq. (9), both C_1 to be the spectrogram and C_2 to be arbitrary are set. In addition, to simplify notation, $\varphi_{SP} = \varphi_1$, $\varphi = \varphi_2$, and $g_{SP} = g_{12}$ are set and written as

$$C_{SP}(t,\omega) = \iint g_{SP}(t'-t,\omega'-\omega) C(t,\omega')\, dt'\, d\omega' \tag{11}$$

The kernel for the spectrogram with window, $h(t)$, is $A_h(-\theta,\tau)$ and therefore

$$g_{SP}(t,\omega) = \frac{1}{4\pi^2}\iint \frac{A_h(-\theta,\tau)}{\phi(\theta,\tau)} e^{j\theta t + j\tau\omega}\, d\theta\, d\tau$$
$$= \frac{1}{4\pi^2}\iiint \frac{1}{\phi(\theta,\tau)} h^*(u-\frac{1}{2}\tau)h(u+\frac{1}{2}\tau)e^{j\theta t + j\tau\omega - j\theta u}\, du\, d\tau\, d\theta$$
$$= \frac{1}{4\pi^2}\iiint h^*(u-\frac{1}{2}\tau)h(u+\frac{1}{2}\tau)\frac{\phi(\theta,\tau)}{\phi(\theta,\tau)\phi(-\theta,\tau)} e^{-j\theta t + j\tau\omega + j\theta u}\, du\, d\tau\, d\theta \quad (12)$$

If taking the kernels for which $\varphi(-\theta,\tau)\varphi(\theta,\tau) = 1$, $g_{SP}(t,\omega)$ is just the distribution of the window function, except that it is evaluated at $-\omega$. Therefore,

$g_{SP}(t,\omega) = C_h(t,-\omega)$ (13)

for kernels that satisfy $\varphi(-\theta,\tau)\varphi(\theta,\tau) = 1$

and

$$C_{SP}(t,\omega) = \iint C_s(t',\omega')C_h(t'-t,\omega'-\omega)\, dt'\, d\omega' \quad (14)$$

for kernels that satisfy $\varphi(-\theta,\tau)\varphi(\theta,\tau) = 1$

This was shown by Janssen. For the case where $\varphi(-\theta,\tau)\varphi(\theta,\tau)$ does not equal one, then

$$C_{SP}(t,\omega) = \iiiint G(t'',\omega'')C_s(t',\omega')C_h(t''+t'-t, -\omega''+\omega-\omega')\, dt'\, dt''\, d\omega\, d\omega'' \quad (15)$$

where

$$G(t,\omega) = \frac{1}{4\pi^2}\iint \frac{e^{-j\theta t - j\tau\omega}}{\phi(\theta,\tau)\phi(-\theta,\tau)}\, d\theta\, d\tau \quad (16)$$

Chapter 5

Analytic Signal

In mathematics and signal processing, the **analytic representation** of a real-valued function or signal facilitates many mathematical manipulations of the signal. The basic idea is that the negative frequency components of the Fourier transform (or spectrum) of a real-valued function are superfluous, due to the Hermitian symmetry of such a spectrum. These negative frequency components can be discarded with no loss of information, providing one is willing to deal with a complex-valued function instead. That makes certain attributes of the signal more accessible and facilitate the derivation of modulation and demodulation techniques, especially single-sideband. As long as the manipulated function has no negative frequency components (that is, it is still *analytic*), the conversion from complex back to real is just a matter of discarding the imaginary part. The analytic representation is a generalization of the phasor concept: while the phasor is restricted to time-invariant amplitude, phase, and frequency, the analytic signal allows for time-variable parameters.

Definition

If $x(t)$ is a real-valued signal with Fourier transform $X(f)$, and $\mathrm{u}(f)$ is the Heaviside step function, then the function**:**

$$\begin{aligned} X_{\mathrm{a}}(f) &\stackrel{\mathrm{def}}{=} \begin{cases} 2X(f), & \text{for } f > 0, \\ X(f), & \text{for } f = 0, \\ 0, & \text{for } f < 0, \end{cases} \\ &= X(f) \cdot 2\mathrm{u}(f) \end{aligned}$$

contains only the non-negative frequency components of $X(f)$. And the operation is reversible, due to the Hermitian property of $X(f)$**:**

$$X(f) = \begin{cases} \frac{1}{2}X_\mathrm{a}(f), & \text{for } f > 0, \\ X^*(|f|), & \text{for } f < 0. \end{cases} \qquad \text{complex conjugate}$$

The inverse Fourier transform of $X_\mathrm{a}(f)$ is the **analytic signal:**

$$\begin{aligned} x_\mathrm{a}(t) &= \underbrace{\mathcal{F}^{-1}\{X(f)\}}_{x(t)} * \underbrace{\mathcal{F}^{-1}\{2\mathrm{u}(f)\}}_{\delta(t)+j\cdot\frac{1}{\pi t}} \qquad \text{convolution} \\ &= x(t) + j\underbrace{\left[x(t) * \frac{1}{\pi t}\right]}_{\hat{x}(t)}, \end{aligned}$$

where $\hat{x}(t)$ is the Hilbert transform of $x(t)$, and j is the imaginary unit.

Example 1: $x(t) = \cos(\omega_0 t)$, for some real parameter $\omega_0 > 0$
$\hat{x}(t) = \cos(\omega_0 t - \frac{\pi}{2}) = \sin(\omega_0 t)$
$x_\mathrm{a}(t) = \cos(\omega_0 t) + j \cdot \sin(\omega_0 t) = e^{j\omega_0 t}$ (The 2nd equality is Euler's formula.)
This is a complex-valued signal with increasing phase (positive frequency).

It also follows from Euler's formula that $\cos(\omega_0 t) = \frac{1}{2}(e^{j\omega_0 t} + e^{-j\omega_0 t})$. So $x(t)$ comprises both positive **and** negative frequency components. $x_\mathrm{a}(t)$ is just the positive portion. When dealing with simple sinusoids or sums of sinusoids, we can deduce $x_\mathrm{a}(t)$ directly, without determining $\hat{x}(t)$ first.

Example 2: $x(t) = \cos(\omega t + \theta) = \frac{1}{2}(e^{j(\omega t+\theta)} + e^{-j(\omega t+\theta)})$

$$x_\mathrm{a}(t) = \begin{cases} e^{j|\omega|t} \cdot e^{j\theta}, & \text{if } \omega > 0, \\ e^{j|\omega|t} \cdot e^{-j\theta}, & \text{if } \omega < 0. \end{cases}$$

The removal of the negative frequency terms is also demonstrated in Example 3. We note that nothing prevents us from computing $x_\mathrm{a}(t)$ for a complex-valued $x(t)$. But it might not be a reversible representation, because the original spectrum is not symmetrical in general. So except for this example, the general discussion assumes real-valued $x(t)$.

Example 3: $x(t) = e^{-j\omega_0 t}$, for some real parameter $\omega_0 > 0$
$\hat{x}(t) = j \cdot e^{-j\omega_0 t}$
$x_\mathrm{a}(t) = e^{-j\omega_0 t} + j^2 \cdot e^{-j\omega_0 t} = 0$

Analytic signals are often shifted in frequency (down-converted) toward 0 Hz, which creates [non-symmetrical] negative frequency components. One motive is to allow lowpass filters with real coefficients to be used to limit the bandwidth of the signal. Another motive is to reduce the highest frequency, which reduces the minimum rate for alias-free sampling. A frequency shift does not undermine the mathematical tractability of the complex signal representation. So in that sense, the down-converted signal is still "analytic". However, restoring the real-valued representation is no longer a simple matter of just extracting the real component. Up-conversion is obviously required, and if the signal has been sampled (discrete-time), interpolation (upsampling) might also be necessary to avoid aliasing.

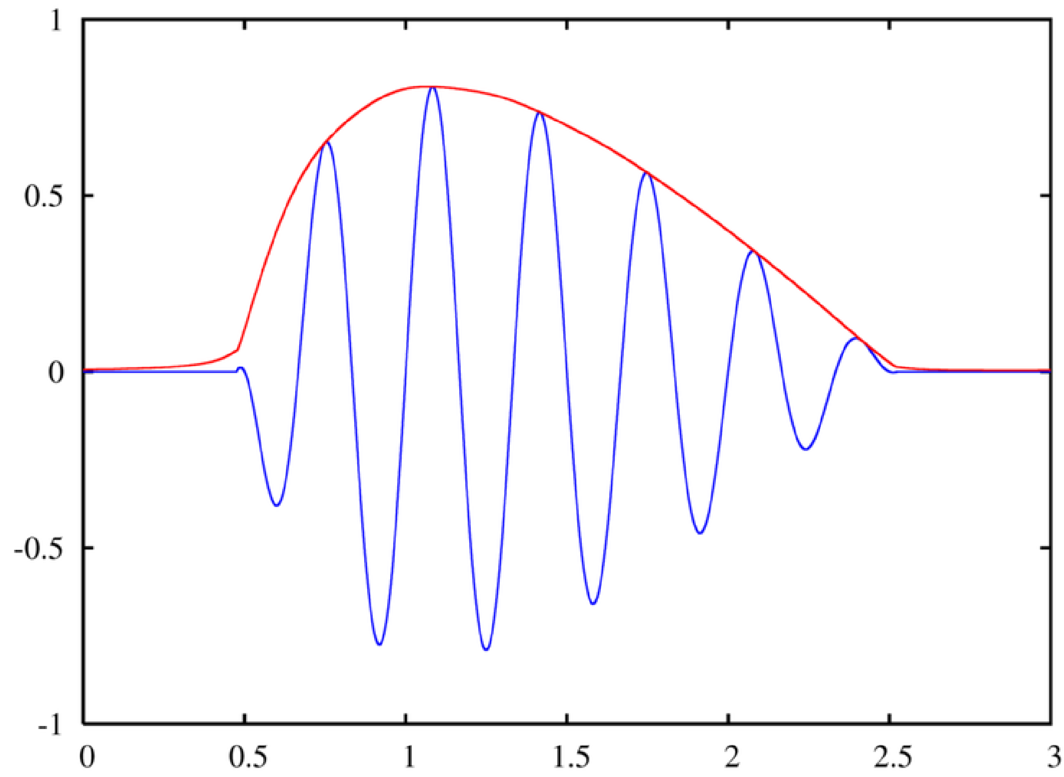
The complex conjugate of an analytic signal contains *only* negative frequency components. In that case also, there is no loss of information or reversibility by discarding the imaginary component. Obviously the real component of the complex conjugate is the same as the real component of the analytic signal. But in this case, its extraction restores the suppressed positive frequency components.

Another way to achieve a spectrum of negative frequencies is to frequency-shift the analytic signal sufficiently far in the negative direction. Extracting the real component again restores the positive frequencies. But in this case their order is reversed**...** the low-frequency component is now the high one. This can be used to demodulate a type of single sideband signal called *lower sideband* or *inverted sideband*.

Applications

The analytic signal can also be expressed in terms of complex polar coordinates, $x_{\mathrm{a}}(t) = A(t)e^{j\phi(t)}$, where**:**

$$A(t) = |x_{\mathrm{a}}(t)| = \sqrt{x^2(t) + \hat{x}^2(t)}$$
$$\phi(t) = \arg\{x_{\mathrm{a}}(t)\}\,.$$



A signal in blue and the magnitude of its analytic signal in red, showing the envelope effect

These functions are respectively called the amplitude envelope and instantaneous phase of the signal $x(t)$. In the accompanying diagram, the blue curve depicts $x(t)$ and the red curve depicts the corresponding $A(t)$.
The time derivative of the unwrapped instantaneous phase is called the instantaneous frequency:.

$$\omega(t) \stackrel{\mathrm{def}}{=} \phi'(t) = \frac{d}{dt}\phi(t).$$

The amplitude function, and the instantaneous phase and frequency are in some applications used to measure and detect local features of the signal. Another application of the analytic representation of a signal relates to demodulation of modulated signals. The polar coordinates conveniently separate the effects of amplitude modulation and phase (or frequency) modulation, and effectively demodulates certain kinds of signals.

The analytic signal can also be represented as:

$$x_\mathrm{a}(t) = \left[A(t)e^{j\phi(t)}e^{-j\omega_0 t}\right]\ e^{j\omega_0 t}$$

$$= \gamma(t) \cdot e^{j\omega_0 t},$$

where

$$\gamma(t) = A(t) \cdot e^{j(\phi(t) - \omega_0 t)}$$

is the signal's *complex envelope*. The complex envelope is not unique; on the contrary, it is determined by an arbitrary ω_0 assignment. This concept is often used when dealing with passband signals. If $x(t)$ is a modulated signal, ω_0 is usually assigned to be a carrier frequency. In other cases it is selected to be somewhere in the middle of the frequency band. Sometimes ω_0 is chosen to minimize

$$\int_0^{+\infty} (\omega - \omega_0)^2 |X_\mathrm{a}(\omega)|^2 \, d\omega.$$

Alternatively , ω_0 can be chosen to minimize the mean square error in linearly approximating the *unwrapped* instantaneous phase $\phi(t)$:

$$\int_{-\infty}^{+\infty} (\omega(t) - \omega_0)^2 |x_\mathrm{a}(t)|^2 \, dt$$

or another alternative (for some optimum θ):

$$\int_{-\infty}^{+\infty} \left(\phi(t) - (\omega_0 t + \theta)\right)^2 \, dt.$$

In the field of time-frequency signal processing, it was shown that the analytic signal was needed in the definition of the Wigner-Ville Distribution so that the method can have the desirable properties needed for practical applications. More details and other applications can be found in .

Extensions of the analytic signal to signals of multiple variables

The concept of analytic signal is well-defined for signals of a single variable which typically is time. For signals of two or more variables, an analytic signal can be defined in different ways, and two approaches are presented below.

Multi-dimensional analytic signal based on an ad-hoc direction

A straightforward generalization of the analytic signal can be done for a multi-dimensional signal once it is established what is meant by *negative frequencies* for this case. This can be done by introducing a normalized vector $\hat{n}$ in the Fourier domain and label any frequency vector *u* as negative if $u \cdot \hat{n} < 0$. The analytic signal is then produced by removing all negative frequencies and multiply the result by 2, in accordance to the procedure described for the case of one-variable signals. However, there is no particular direction for $\hat{n}$ which must be chosen unless there are some additional constraints. Therefore, the choice of $\hat{n}$ is ad-hoc, or application specific.

The monogenic signal

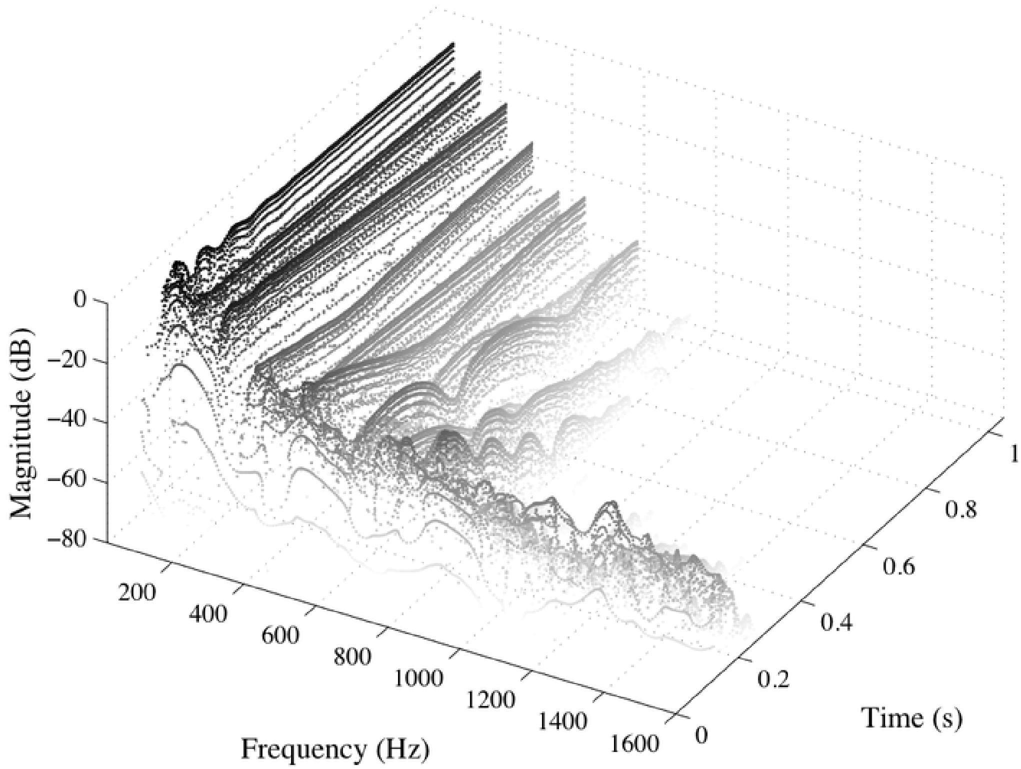
The real and imaginary parts of the analytic signal correspond to the two elements of the vector-valued monogenic signal, as it is defined for one-variable signals. However, the monogenic signal can be extended to arbitrary number of variables in a straightforward manner, producing an $(n + 1)$-dimensional vector-valued function for the case of n-variable signals.

Chapter 6

Reassignment Method

The **method of reassignment** is a technique for sharpening a time-frequency representation by mapping the data to time-frequency coordinates that are nearer to the true region of support of the analyzed signal. The method has been independently introduced by several parties under various names, including *method of reassignment*, *remapping*, *time-frequency reassignment*, and *modified moving-window method*. In the case of the spectrogram or the short-time Fourier transform, the method of reassignment sharpens blurry time-frequency data by relocating the data according to local estimates of instantaneous frequency and group delay. This mapping to reassigned time-frequency coordinates is very precise for signals that are separable in time and frequency with respect to the analysis window.

Introduction



Reassigned spectral surface for the onset of an acoustic bass tone having a sharp pluck and a fundamental frequency of approximately 73.4 Hz. Sharp spectral ridges representing the harmonics are evident, as is the abrupt onset of the tone. The spectrogram was computed using a 65.7 ms Kaiser window with a shaping parameter of 12.

Many signals of interest have a distribution of energy that varies in time and frequency. For example, any sound signal having a beginning or an end has an energy distribution that varies in time, and most sounds exhibit considerable variation in both time and frequency over their duration. Time-frequency representations are commonly used to analyze or characterize such signals. They map the one-dimensional time-domain signal into a two-dimensional function of time and frequency. A time-frequency representation describes the variation of spectral energy distribution over time, much as a musical score describes the variation of musical pitch over time.

In audio signal analysis, the spectrogram is the most commonly-used time-frequency representation, probably because it is well-understood, and immune to so-called "cross-terms" that sometimes make other time-frequency representations difficult to interpret. But the windowing operation required in spectrogram computation introduces an unsavory tradeoff between time resolution and frequency resolution, so spectrograms provide a time-frequency representation that is blurred in time, in frequency, or in both dimensions. The method of time-frequency reassignment is a technique for refocussing time-frequency data in a blurred representation like the spectrogram by mapping the data to time-frequency coordinates that are nearer to the true region of support of the analyzed signal.

The spectrogram as a time-frequency representation

One of the best-known time-frequency representations is the spectrogram, defined as the squared magnitude of the short-time Fourier transform. Though the short-time phase spectrum is known to contain important temporal information about the signal, this information is difficult to interpret, so typically, only the short-time magnitude spectrum is considered in short-time spectral analysis.

As a time-frequency representation, the spectrogram has relatively poor resolution. Time and frequency resolution are governed by the choice of analysis window and greater concentration in one domain is accompanied by greater smearing in the other.

A time-frequency representation having improved resolution, relative to the spectrogram, is the Wigner–Ville distribution, which may be interpreted as a short-time Fourier transform with a window function that is perfectly matched to the signal. The Wigner–Ville distribution is highly-concentrated in time and frequency, but it is also highly nonlinear and non-local. Consequently, this distribution is very sensitive to noise, and generates cross-components that often mask the components of interest, making it difficult to extract useful information concerning the distribution of energy in multi-component signals.

Cohen's class of bilinear time-frequency representations is a class of "smoothed" Wigner–Ville distributions, employing a smoothing kernel that can reduce sensitivity of the distribution to noise and suppresses cross-components, at the expense of smearing the distribution in time and frequency. This smearing causes the distribution to be non-zero in regions where the true Wigner–Ville distribution shows no energy.

The spectrogram is a member of Cohen's class. It is a smoothed Wigner–Ville distribution with the smoothing kernel equal to the Wigner–Ville distribution of the analysis window. The method of reassignment smoothes the Wigner–Ville distribution, but then refocuses the distribution back to the true regions of support of the signal components. The method has been shown to reduce time and frequency smearing of any member of Cohen's class In the case of the reassigned spectrogram, the short-time phase spectrum is used to correct the nominal time and frequency coordinates of the spectral data, and map it back nearer to the true regions of support of the analyzed signal.

The method of reassignment

Pioneering work on the method of reassignment was first published by Kodera, Gendrin, and de Villedary under the name of *Modified Moving Window Method* Their technique enhances the resolution in time and frequency of the classical Moving Window Method (equivalent to the spectrogram) by assigning to each data point a new time-frequency coordinate that better-reflects the distribution of energy in the analyzed signal.

In the classical moving window method, a time-domain signal, $x(t)$ is decomposed into a set of coefficients, $\varepsilon(t,\omega)$, based on a set of elementary signals, $h_\omega(t)$, defined

$$h_\omega(t) = h(t)e^{j\omega t}$$

where $h(t)$ is a (real-valued) lowpass kernel function, like the window function in the short-time Fourier transform. The coefficients in this decomposition are defined

$$\begin{aligned}\epsilon(t,\omega) &= \int x(\tau)h(t-\tau)e^{-j\omega[\tau-t]}d\tau \\ &= e^{j\omega t}\int x(\tau)h(t-\tau)e^{-j\omega\tau}d\tau \\ &= e^{j\omega t}X(t,\omega) \\ &= X_t(\omega) = M_t(\omega)e^{j\phi_\tau(\omega)}\end{aligned}$$

where $M_t(\omega)$ is the magnitude, and $\varphi_\tau(\omega)$ the phase, of $X_t(\omega)$, the Fourier transform of the signal $x(t)$ shifted in time by t and windowed by $h(t)$.

$x(t)$ can be reconstructed from the moving window coefficients by

$$\begin{aligned}x(t) &= \iint X_\tau(\omega)h^*_\omega(\tau-t)d\omega d\tau \\ &= \iint X_\tau(\omega)h(\tau-t)e^{-j\omega[\tau-t]}d\omega d\tau \\ &= \iint M_\tau(\omega)e^{j\phi_\tau(\omega)}h(\tau-t)e^{-j\omega[\tau-t]}d\omega d\tau \\ &= \iint M_\tau(\omega)h(\tau-t)e^{j[\phi_\tau(\omega)-\omega\tau+\omega t]}d\omega d\tau\end{aligned}$$

For signals having magnitude spectra, $M(t,\omega)$, whose time variation is slow relative to the phase variation, the maximum contribution to the reconstruction integral comes from the vicinity of the point t,ω satisfying the phase stationarity condition

$$\begin{aligned}\frac{\partial}{\partial\omega}\left[\phi_\tau(\omega)-\omega\tau+\omega t\right] &= 0 \\ \frac{\partial}{\partial\tau}\left[\phi_\tau(\omega)-\omega\tau+\omega t\right] &= 0\end{aligned}$$

or equivalently, around the point $\hat{t}, \hat{\omega}$ defined by

$$\hat{t}(\tau,\omega) = \tau - \frac{\partial \phi_\tau(\omega)}{\partial \omega} = -\frac{\partial \phi(\tau,\omega)}{\partial \omega}$$
$$\hat{\omega}(\tau,\omega) = \frac{\partial \phi_\tau(\omega)}{\partial \tau} = \omega + \frac{\partial \phi(\tau,\omega)}{\partial \tau}.$$

This phenomenon is known in such fields as optics as the principle of stationary phase, which states that for periodic or quasi-periodic signals, the variation of the Fourier phase spectrum not attributable to periodic oscillation is slow with respect to time in the vicinity of the frequency of oscillation, and in surrounding regions the variation is relatively rapid. Analogously, for impulsive signals, that are concentrated in time, the variation of the phase spectrum is slow with respect to frequency near the time of the impulse, and in surrounding regions the variation is relatively rapid.

In reconstruction, positive and negative contributions to the synthesized waveform cancel, due to destructive interference, in frequency regions of rapid phase variation. Only regions of slow phase variation (stationary phase) will contribute significantly to the reconstruction, and the maximum contribution (center of gravity) occurs at the point where the phase is changing most slowly with respect to time and frequency.

The time-frequency coordinates thus computed are equal to the local group delay, $\hat{t}_g(t,\omega)$, and local instantaneous frequency, $\hat{\omega}_i(t,\omega)$, and are computed from the phase of the short-time Fourier transform, which is normally ignored when constructing the spectrogram. These quantities are *local* in the sense that they are represent a windowed and filtered signal that is localized in time and frequency, and are not global properties of the signal under analysis.

The modified moving window method, or method of reassignment, changes (reassigns) the point of attribution of $\varepsilon(t,\omega)$ to this point of maximum contribution $\hat{t}(t,\omega), \hat{\omega}(t,\omega)$, rather than to the point t,ω at which it is computed. This point is sometimes called the *center of gravity* of the distribution, by way of analogy to a mass distribution. This analogy is a useful reminder that the attribution of spectral energy to the center of gravity of its distribution only makes sense when there is energy to attribute, so the method of reassignment has no meaning at points where the spectrogram is zero-valued.

Efficient computation of reassigned times and frequencies

In digital signal processing, it is most common to sample the time and frequency domains. The discrete Fourier transform is used to compute samples $X(k)$ of the Fourier transform from samples $x(n)$ of a time domain signal. The reassignment operations proposed by Kodera *et al.* cannot be applied directly to the discrete short-time Fourier transform data, because partial derivatives cannot be computed directly on data that is discrete in time and frequency, and it has been suggested that this difficulty has been the primary barrier to wider use of the method of reassignment.

It is possible to approximate the partial derivatives using finite differences. For example, the phase spectrum can be evaluated at two nearby times, and the partial derivative with respect to time be approximated as the difference between the two values divided by the time difference, as in

$$\begin{aligned} \frac{\partial \phi(t,\omega)}{\partial t} &\approx \frac{1}{\Delta t}\left[\phi(t+\tfrac{\Delta t}{2},\omega) - \phi(t-\tfrac{\Delta t}{2},\omega)\right] \\ \frac{\partial \phi(t,\omega)}{\partial \omega} &\approx \frac{1}{\Delta \omega}\left[\phi(t,\omega+\tfrac{\Delta \omega}{2}) - \phi(t,\omega-\tfrac{\Delta \omega}{2})\right] \end{aligned}$$

For sufficiently small values of Δt and $\Delta\omega$, and provided that the phase difference is appropriately "unwrapped", this finite-difference method yields good approximations to the partial derivatives of phase, because in regions of the spectrum in which the evolution of the phase is dominated by rotation due to sinusoidal oscillation of a single, nearby component, the phase is a linear function.

Independently of Kodera *et al.* , Nelson arrived at a similar method for improving the time-frequency precision of short-time spectral data from partial derivatives of the short-time phase spectrum. It is easily shown that Nelson's *cross spectral surfaces* compute an approximation of the derivatives that is equivalent to the finite differences method.

Auger and Flandrin showed that the method of reassignment, proposed in the context of the spectrogram by Kodera *et al.*, could be extended to any member of Cohen's class of time-frequency representations by generalizing the reassignment operations to

$$\begin{aligned} \hat{t}(t,\omega) &= t - \frac{\iint \tau \cdot W_x(t-\tau,\omega-\nu)\cdot\Phi(\tau,\nu)d\tau\, d\nu}{\iint W_x(t-\tau,\omega-\nu)\cdot\Phi(\tau,\nu)d\tau\, d\nu} \\ \hat{\omega}(t,\omega) &= \omega - \frac{\iint \nu \cdot W_x(t-\tau,\omega-\nu)\cdot\Phi(\tau,\nu)d\tau\, d\nu}{\iint W_x(t-\tau,\omega-\nu)\cdot\Phi(\tau,\nu)d\tau\, d\nu} \end{aligned}$$

where $W_x(t,\omega)$ is the Wigner–Ville distribution of $x(t)$, and $\Phi(t,\omega)$ is the kernel function that defines the distribution. They further described an efficient method for computing the times and frequencies for the reassigned spectrogram efficiently and accurately without explicitly computing the partial derivatives of phase.

In the case of the spectrogram, the reassignment operations can be computed by

$$\begin{aligned} \hat{t}(t,\omega) &= t - \Re\left\{\frac{X_{\mathcal{T}h}(t,\omega)\cdot X^*(t,\omega)}{|X(t,\omega)|^2}\right\} \\ \hat{\omega}(t,\omega) &= \omega + \Im\left\{\frac{X_{\mathcal{D}h}(t,\omega)\cdot X^*(t,\omega)}{|X(t,\omega)|^2}\right\} \end{aligned}$$

where $X(t,\omega)$ is the short-time Fourier transform computed using an analysis window $h(t)$, $X_{\mathcal{T}h}(t,\omega)$is the short-time Fourier transform computed using a time-weighted anlaysis window $h_T(t) = t \cdot h(t)$ and $X_{Dh}(t,\omega)$ is the short-time Fourier transform computed using a time-derivative analysis window $h_D(t) = \frac{d}{dt}h(t)$.

Using the auxiliary window functions $h_T(t)$ and $h_D(t)$, the reassignment operations can be computed at any time-frequency coordinate *t*,ω from an algebraic combination of three Fourier transforms evaluated at *t*,ω. Since these algorithms operate only on short-time spectral data evaluated at a single time and frequency, and do not explicitly compute any derivatives, the reassigned time-frequency coordinates $\hat{\omega}(t_n, \omega_k)$ and $\hat{t}(t_n, \omega_k)$ can be computed from three discrete short-time Fourier transforms evaluated at t_n,ω_k. This gives an efficient method of computing the reassigned discrete short-time Fourier transform provided only that the $|X(t,\omega)|^2$ is non-zero. This is not much of a restriction, since the reassignment operation itself implies that there is some energy to reassign, and has no meaning when the distribution is zero-valued.

Separability

The short-time Fourier transform can often be used to estimate the amplitudes and phases of the individual components in a *multi-component* signal, such as a quasi-harmonic musical instrument tone. Moreover, the time and frequency reassignment operations can be used to sharpen the representation by attributing the spectral energy reported by the short-time Fourier transform to the point that is the local center of gravity of the complex energy distribution.

For a signal consisting of a single component, the instantaneous frequency can be estimated from the partial derivatives of phase of any short-time Fourier transform channel that passes the component. If the signal is to be decomposed into many components,

$$x(t) = \sum_n A_n(t) e^{j\theta_n(t)}$$

and the instantaneous frequency of each component is defined as the derivative of its phase with respect to time, that is,

$$\omega_n(t) = \frac{d\theta_n(t)}{dt},$$

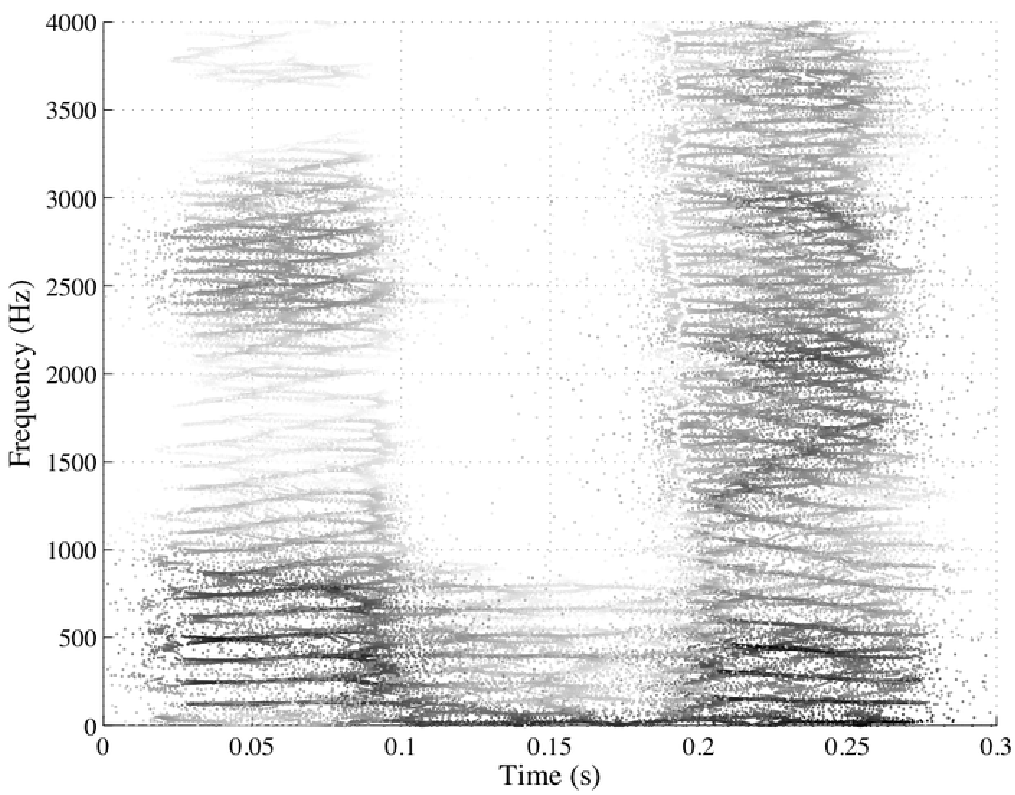
then the instantaneous frequency of each individual component can be computed from the phase of the response of a filter that passes that component, provided that no more than one component lies in the passband of the filter.

This is the property, in the frequency domain, that Nelson called *separability* and is required of all signals so analyzed. If this property is not met, then the desired multi-component decomposition cannot be achieved, because the parameters of individual components cannot be estimated from the short-time Fourier transform. In such cases, a different analysis window must be chosen so that the separability criterion is satisfied.

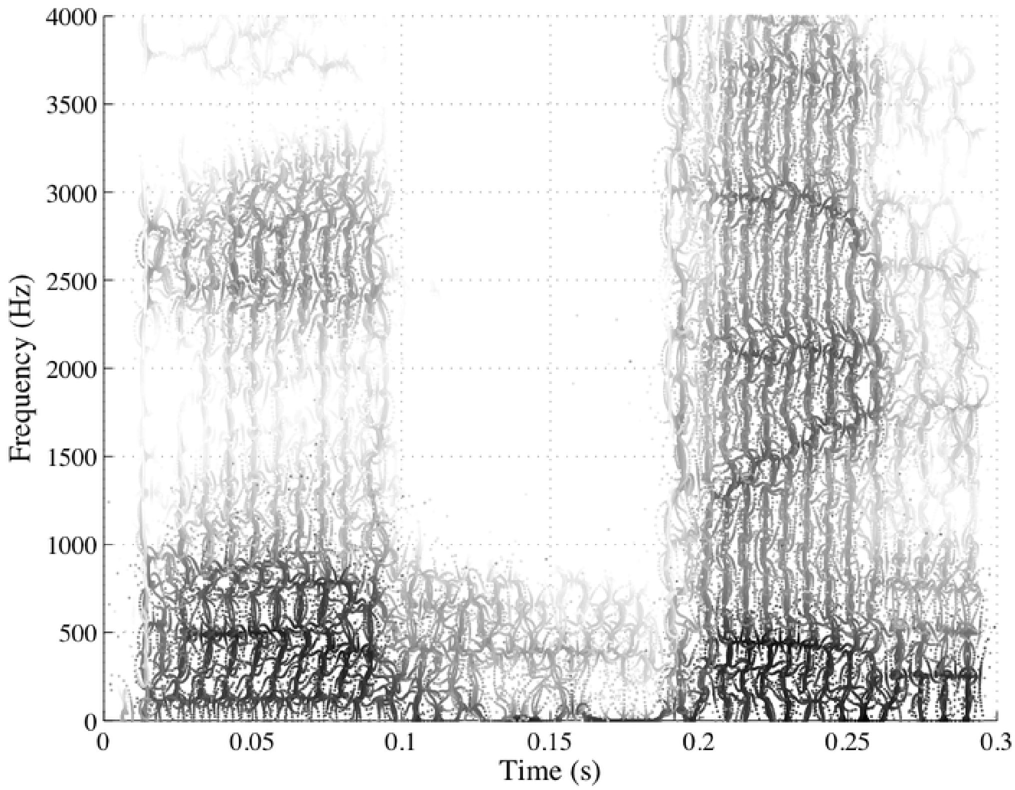
If the components of a signal are separable in frequency with respect to a particular short-time spectral analysis window, then the output of each short-time Fourier transform filter is a filtered version of, at most, a single dominant (having significant energy) component, and so the derivative, with respect to time, of the phase of the $X(t,\omega_0)$ is equal to the derivative with respect to time, of the phase of the dominant component at ω_0. Therefore, if a component, $x_n(t)$, having instantaneous frequency $\omega_n(t)$ is the dominant component in the vicinity of ω_0, then the instantaneous frequency of that component can be computed from the phase of the short-time Fourier transform evaluated at ω_0. That is,

$$\begin{aligned} \omega_n(t) &= \frac{\partial}{\partial t}\arg\{x_n(t)\} \\ &= \frac{\partial}{\partial t}\arg\{X(t,\omega_0)\} \end{aligned}$$

Thus, the partial derivative with respect to time of the phase of the short-time Fourier transform can be used to compute the instantaneous frequencies of the individual components in a multi-component signal, provided only that the components are separable in frequency by the chosen analysis window.



Long-window reassigned spectrogram of the word "open", computed using a 54.4 ms Kaiser window with a shaping parameter of 9, emphasizing harmonics.



Short-window reassigned spectrogram of the word "open", computed using a 13.6 ms Kaiser window with a shaping parameter of 9, emphasizing formants and glottal pulses.

Just as each bandpass filter in the short-time Fourier transform filterbank may pass at most a single complex exponential component, two temporal events must be sufficiently separated in time that they do not lie in the same windowed segment of the input signal. This is the property of separability in the time domain, and is equivalent to requiring that the time between two events be greater than the length of the impulse response of the short-time Fourier transform filters, the span of non-zero samples in $h(t)$.

Separability in time and in frequency is required of components to be resolved in a reassigned time-frequency representation. If the components in a decomposition are separable in a certain time-frequency representation, then the components can be resolved by that time-frequency representation, and using the method of reassignment, can be characterized with much greater precision than is possible using classical methods.

In general, there are an infinite number of equally-valid decompositions for a multi-component signal. The separability property must be considered in the context of the desired decomposition. For example, in the analysis of a speech signal, an analysis window that is long relative to the time between glottal pulses is sufficient to separate harmonics, but the individual glottal pulses will be smeared, because many pulses are covered by each window (that is, the individual pulses are not separable, in time, by the chosen analysis window). An analysis window that is much shorter than the time between glottal pulses may resolve the glottal pulses, because no window spans more than one pulse, but the harmonic frequencies are smeared together, because the main lobe of the analysis window spectrum is wider than the spacing between the harmonics (that is, the harmonics are not separable, in frequency, by the chosen analysis window).

Chapter 7

Short-Time Fourier Transform

The **short-time Fourier transform** (**STFT**), or alternatively **short-term Fourier transform**, is a Fourier-related transform used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time.

STFT

Continuous-time STFT

Simply, in the continuous-time case, the function to be transformed is multiplied by a window function which is nonzero for only a short period of time. The Fourier transform (a one-dimensional function) of the resulting signal is taken as the window is slid along the time axis, resulting in a two-dimensional representation of the signal. Mathematically, this is written as:

$$\mathbf{STFT}\left\{x(t)\right\} \equiv X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}\,dt$$

where $w(t)$ is the window function, commonly a Hann window or Gaussian "hill" centered around zero, and $x(t)$ is the signal to be transformed. $X(\tau,\omega)$ is essentially the Fourier Transform of $x(t)w(t-\tau)$, a complex function representing the phase and magnitude of the signal over time and frequency. Often phase unwrapping is employed along either or both the time axis, τ, and frequency axis, ω, to suppress any jump discontinuity of the phase result of the STFT. The time index τ is normally considered to be "*slow*" time and usually not expressed in as high resolution as time t.

Discrete-time STFT

In the discrete time case, the data to be transformed could be broken up into chunks or frames (which usually overlap each other, to reduce artifacts at the boundary). Each chunk is Fourier transformed, and the complex result is added to a matrix, which records magnitude and phase for each point in time and frequency. This can be expressed as:

$$\mathbf{STFT}\{x[n]\} \equiv X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}$$

likewise, with signal $x[n]$ and window $w[n]$. In this case, m is discrete and ω is continuous, but in most typical applications the STFT is performed on a computer using the Fast Fourier Transform, so both variables are discrete and quantized. Again, the discrete-time index m is normally considered to be "slow" time and usually not expressed in as high resolution as time n.

The magnitude squared of the STFT yields the spectrogram of the function:

$$\text{spectrogram}\{x(t)\} \equiv |X(\tau,\omega)|^2$$

Inverse STFT

The STFT is invertible, that is, the original signal can be recovered from the transform by the Inverse STFT.

Continuous-time STFT

Given the width and definition of the window function $w(t)$, we initially require the height of the window function to be scaled so that

$$\int_{-\infty}^{\infty} w(\tau)\,d\tau = 1.$$

It easily follows that

$$\int_{-\infty}^{\infty} w(t-\tau)\,d\tau = 1 \quad \forall\ t$$

and

$$x(t) = x(t)\int_{-\infty}^{\infty} w(t-\tau)\,d\tau = \int_{-\infty}^{\infty} x(t)w(t-\tau)\,d\tau.$$

The continuous Fourier Transform is

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\, dt.$$

Substituting *x*(*t*) from above:

$$X(\omega) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x(t)w(t-\tau)\, d\tau\right] e^{-j\omega t}\, dt$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t)w(t-\tau)\, e^{-j\omega t}\, d\tau\, dt.$$

Swapping order of integration:

$$X(\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t)w(t-\tau)\, e^{-j\omega t}\, dt\, d\tau$$
$$= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x(t)w(t-\tau)\, e^{-j\omega t}\, dt\right] d\tau$$
$$= \int_{-\infty}^{\infty} X(\tau,\omega)\, d\tau.$$

So the Fourier Transform can be seen as a sort of phase coherent sum of all of the STFTs of *x*(*t*). Since the inverse Fourier transform is

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)e^{+j\omega t}\, d\omega,$$

then *x*(*t*) can be recovered from $X(\tau,\omega)$ as

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(\tau,\omega)e^{+j\omega t}\, d\tau\, d\omega.$$

or

$$x(t) = \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} X(\tau,\omega)e^{+j\omega t}\, d\omega\right] d\tau.$$

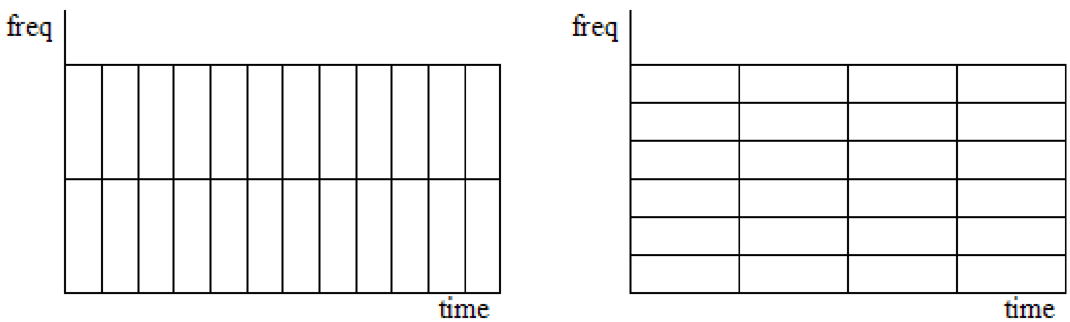
It can be seen, comparing to above that windowed "grain" or "wavelet" of *x*(*t*) is

$$x(t)w(t-\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\tau,\omega)e^{+j\omega t}\, d\omega.$$

the inverse Fourier transform of $X(\tau,\omega)$ for τ fixed.

Resolution issues

One of the downfalls of the STFT is that it has a fixed resolution. The width of the windowing function relates to how the signal is represented—it determines whether there is good frequency resolution (frequency components close together can be separated) or good time resolution (the time at which frequencies change). A wide window gives better frequency resolution but poor time resolution. A narrower window gives good time resolution but poor frequency resolution. These are called narrowband and wideband transforms, respectively.



Comparison of STFT resolution. Left has better time resolution, and right has better frequency resolution.

This is one of the reasons for the creation of the wavelet transform (or multiresolution analysis in general), which can give good time resolution for high-frequency events, and good frequency resolution for low-frequency events, which is the type of analysis best suited for many real signals.

This property is related to the Heisenberg uncertainty principle, but it is not a direct relationship. The product of the standard deviation in time and frequency is limited. The boundary of the uncertainty principle (best simultaneous resolution of both) is reached with a Gaussian window function, as the Gaussian minimizes the Fourier uncertainty principle.

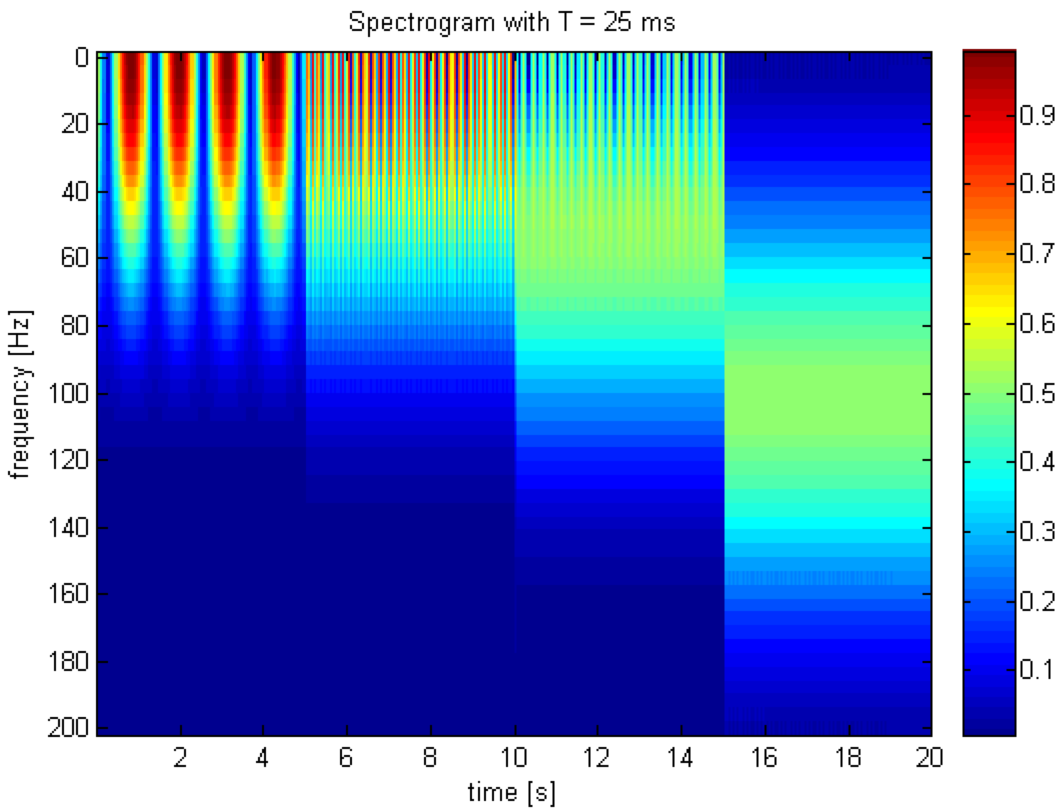
One can consider the STFT for varying window size as a two-dimensional domain (time and frequency), as illustrated in the example below, which can be calculated by varying the window size. However, this is no longer a strictly time–frequency representation.

Example

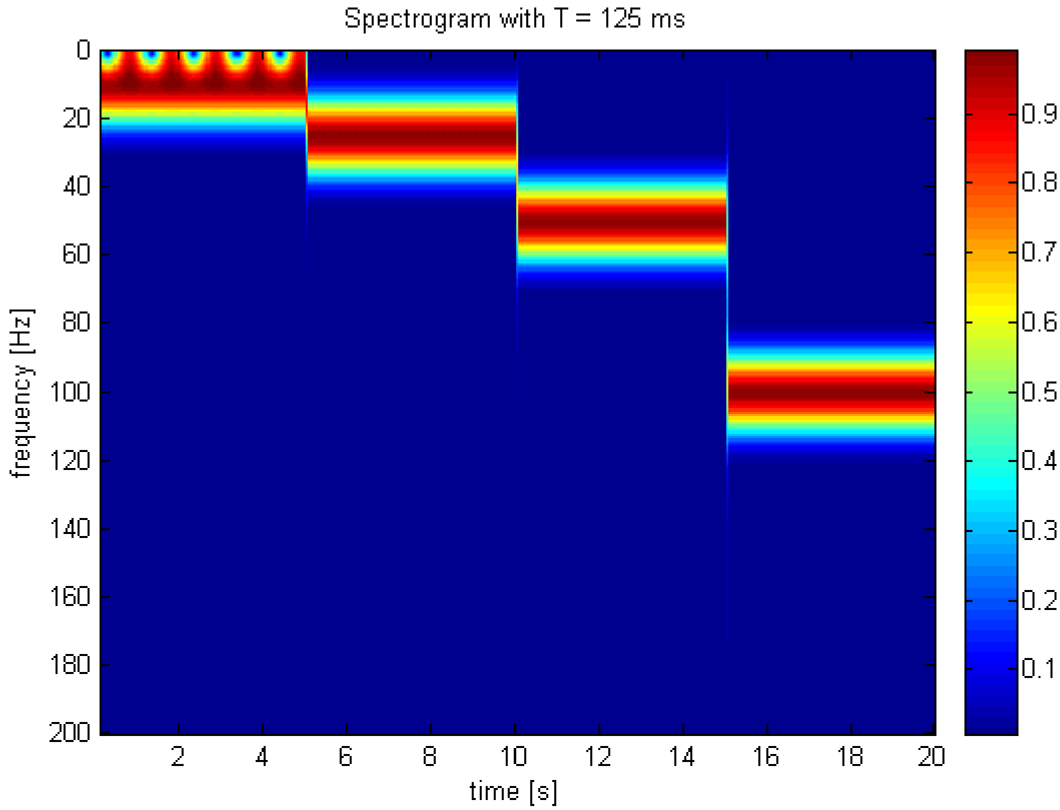
Using the following sample signal $x(t)$ that is composed of a set of four sinusoidal waveforms joined together in sequence. Each waveform is only composed of one of four frequencies (10, 25, 50, 100 Hz). The definition of $x(t)$ is:

$$x(t) = \begin{cases} \cos(2\pi 10t/\mathrm{s}) & 0\,\mathrm{s} \le t < 5\,\mathrm{s} \\ \cos(2\pi 25t/\mathrm{s}) & 5\,\mathrm{s} \le t < 10\,\mathrm{s} \\ \cos(2\pi 50t/\mathrm{s}) & 10\,\mathrm{s} \le t < 15\,\mathrm{s} \\ \cos(2\pi 100t/\mathrm{s}) & 15\,\mathrm{s} \le t < 20\,\mathrm{s} \end{cases}$$

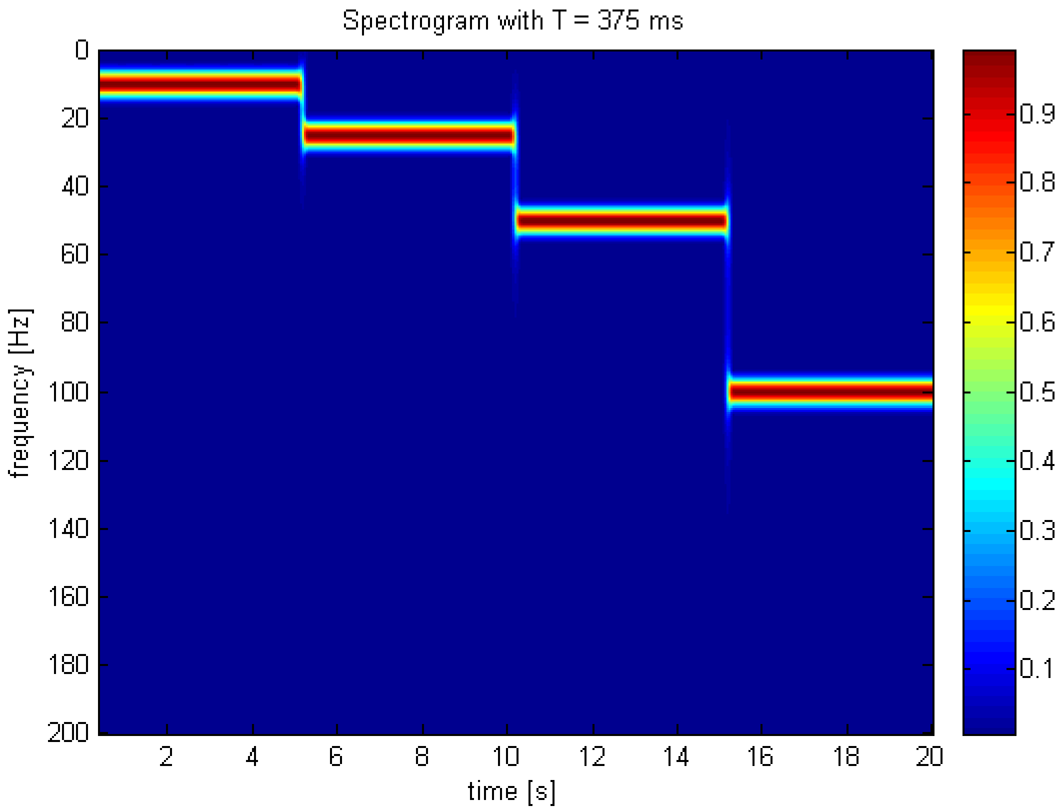
Then it is sampled at 400 Hz. The following spectrograms were produced:



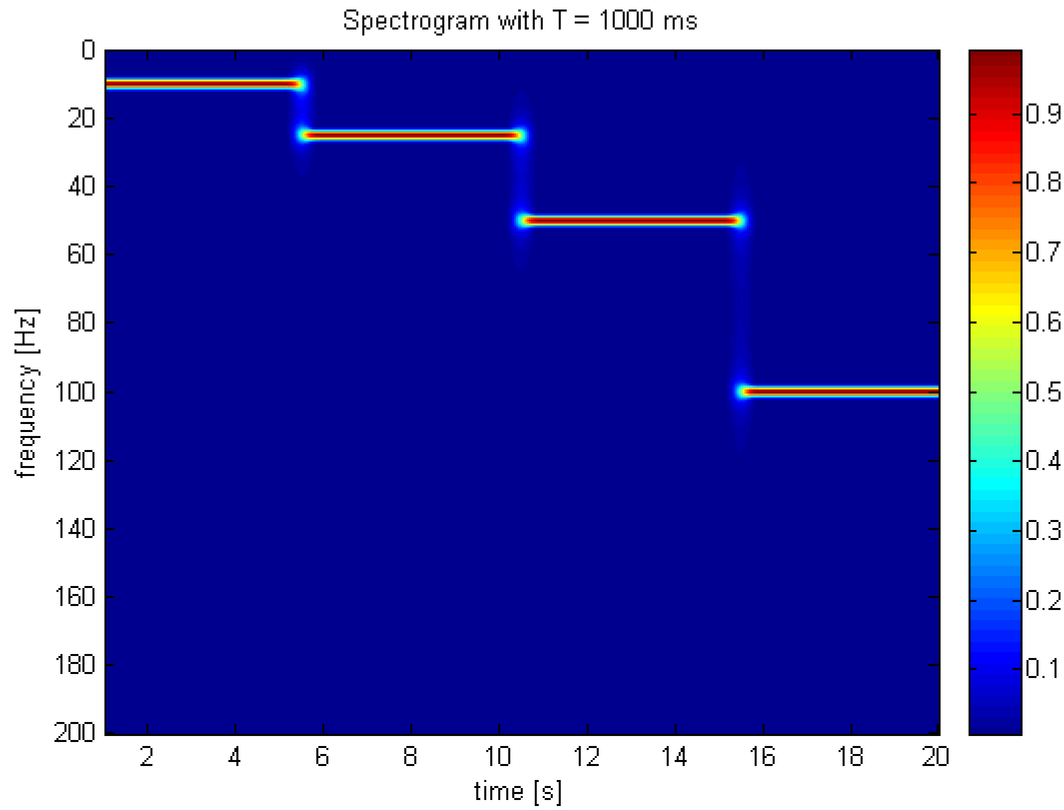
25 ms window



125 ms window



375 ms window



1000 ms window

The 25 ms window allows us to identify a precise time at which the signals change but the precise frequencies are difficult to identify. At the other end of the scale, the 1000 ms window allows the frequencies to be precisely seen but the time between frequency changes is blurred.

Explanation

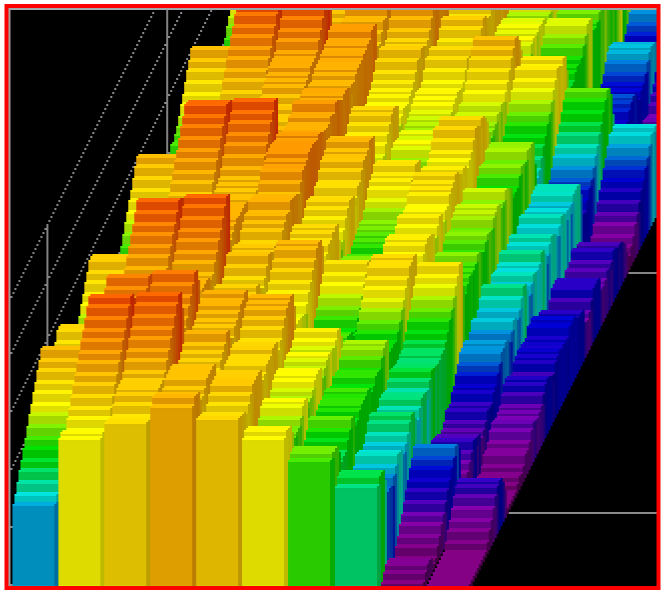
It can also be explained with reference to the sampling and Nyquist frequency.

Take a window of N samples from an arbitrary real-valued signal at sampling rate f_s . Taking the Fourier transform produces N complex coefficients. Of these coefficients only half are useful (the last $N/2$ being the complex conjugate of the first $N/2$ in reverse order, as this is a real valued signal).

These $N/2$ coefficients represent the frequencies 0 to $f_s/2$ (Nyquist) and two consecutive coefficients are spaced apart by f_s/N Hz.

To increase the frequency resolution of the window the frequency spacing of the coefficients needs to be reduced. There are only two variables, but decreasing f_s (and keeping N constant) will cause the window size to increase — since there are now fewer samples per unit time. The other alternative is to increase *N*, but this again causes the window size to increase. So any attempt to increase the frequency resolution causes a larger window size and therefore a reduction in time resolution—and vice-versa.

Application

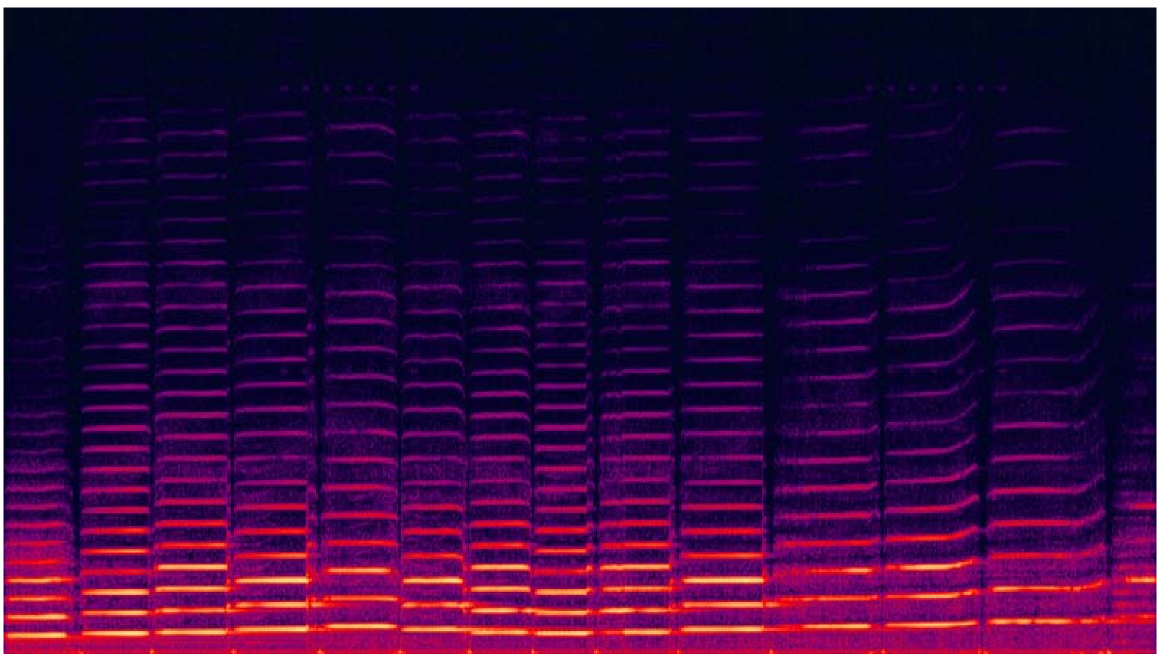


A STFT being used to analyze an audio signal across time.

STFTs as well as standard Fourier transforms and other tools are frequently used to analyze music. The spectrogram can, for example, show frequency on the horizontal axis, with the lowest frequencies at left, and the highest at the right. The height of each bar (augmented by color) represents the amplitude of the frequencies within that band. The depth dimension represents time, where each new bar was a separate distinct transform. Audio engineers use this kind of visual to gain information about an audio sample, for example, to locate the frequencies of specific noises (especially when used with greater frequency resolution) or to find frequencies which may be more or less resonant in the space where the signal was recorded. This information can be used for equalization or tuning other audio effects.

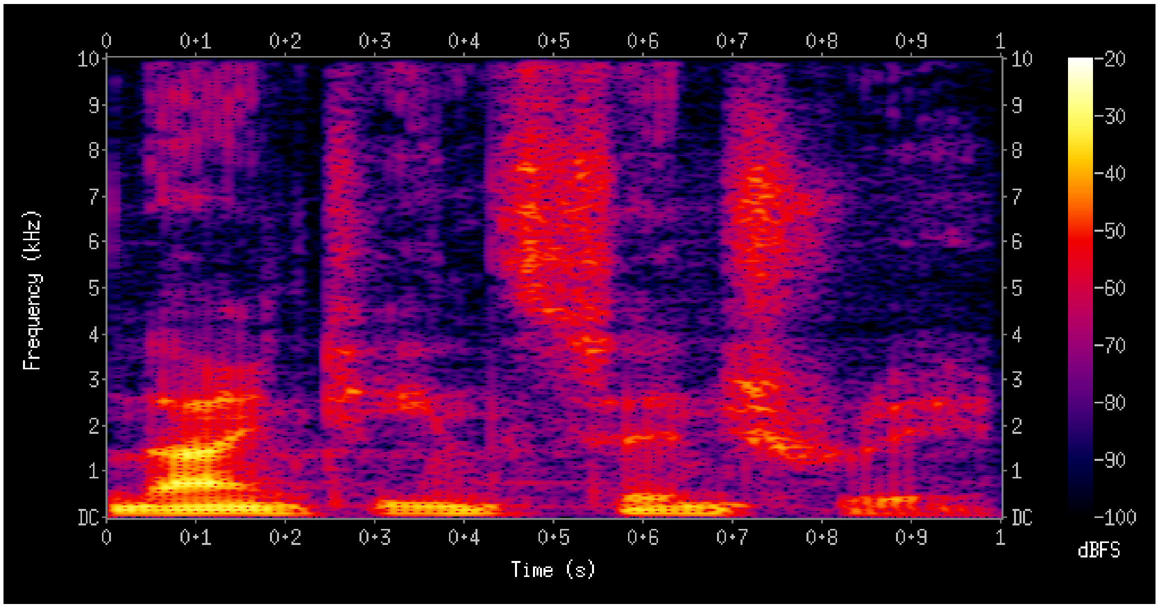
Chapter 8

Spectrogram



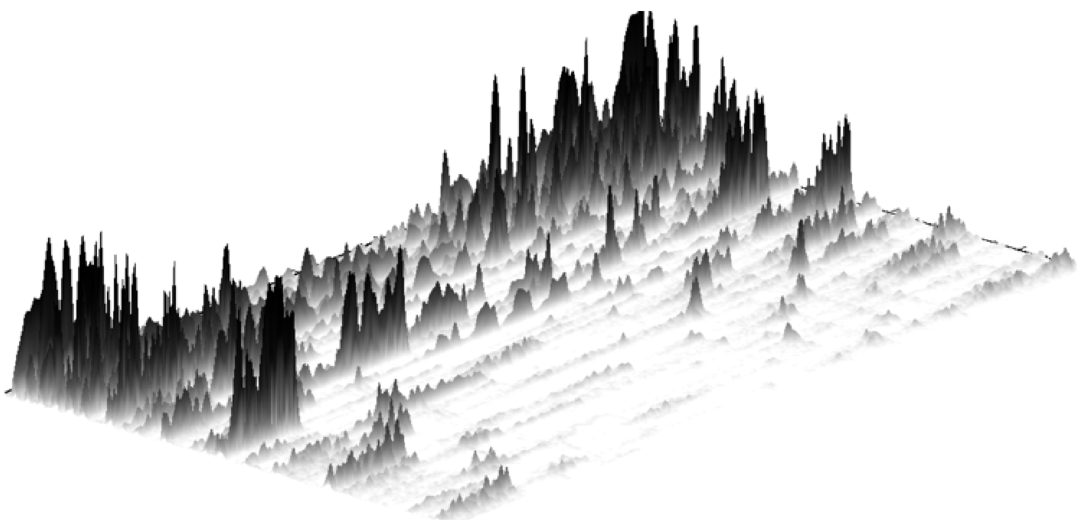
Spectrogram of violin playing.

A **spectrogram** is a time-varying spectral representation (forming an image) that shows how the spectral density of a signal varies with time. In the field of time–frequency signal processing, it is one of the most popular quadratic Time-Frequency Distribution that represents a signal in a joint time-frequency domain and that has the property of being positive. Also known as **spectral waterfalls**, **sonograms**, **voiceprints**, or **voicegrams**, spectrograms are used to identify phonetic sounds, to analyse the cries of animals; they were also used in many other fields including music, sonar/radar, speech processing, seismology, etc. The instrument that generates a spectrogram is called a **spectrograph** and is equivalent to a **sonograph**.



Spectrogram of a male voice saying 'nineteenth century'.

Format

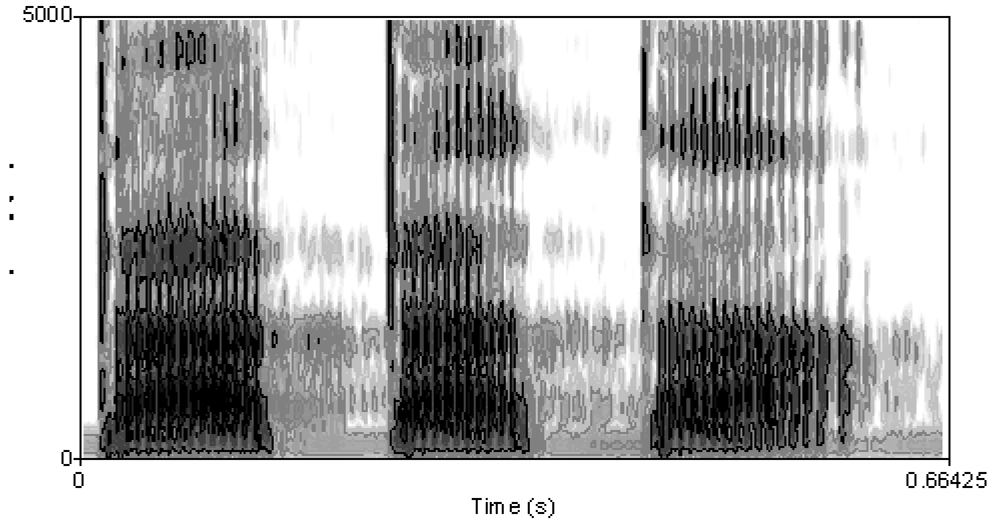


3D surface spectrogram of a part from a music piece.

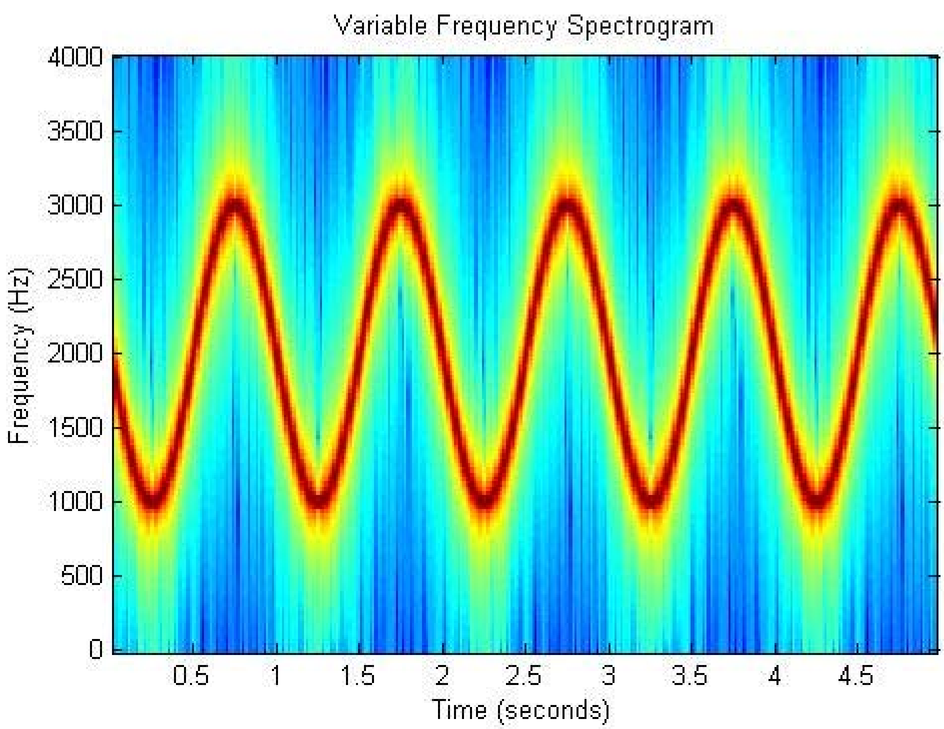
The most common format is a graph with two geometric dimensions: the horizontal axis represents time, the vertical axis is frequency; a third dimension indicating the amplitude of a particular frequency at a particular time is represented by the intensity or colour of each point in the image.

There are many variations of format: sometimes the vertical and horizontal axes are switched, so time runs up and down; sometimes the amplitude is represented as the height of a 3D surface instead of color or intensity. The frequency and amplitude axes can be either linear or logarithmic, depending on what the graph is being used for. Audio would usually be represented with a logarithmic amplitude axis (probably in decibels, or dB), and frequency would be linear to emphasize harmonic relationships, or logarithmic to emphasize musical, tonal relationships.

Generation



Spectrogram of a male voice saying 'tatata'.



Spectrogram of an FM signal. In this case the signal frequency is modulated with a sinusoidal frequency vs. time profile

Spectrograms are usually created in one of two ways: approximated as a filterbank that results from a series of bandpass filters (this was the only way before the advent of modern digital signal processing), or calculated from the time signal using the short-time Fourier transform (STFT). These two methods actually form two different quadratic Time-Frequency Distributions, but are equivalent under some conditions.

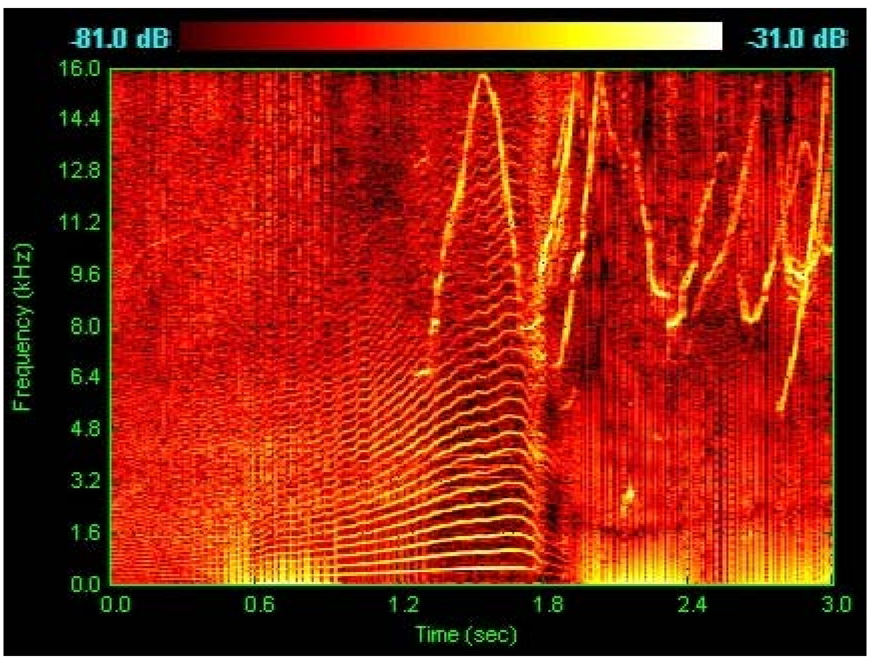
The bandpass filters method usually uses analog processing to divide the input signal into frequency bands; the magnitude of each filter's output controls a transducer that records the spectrogram as an image on paper.

Creating a spectrogram using the STFT is usually a digital process. Digitally sampled data, in the time domain, is broken up into chunks, which usually overlap, and Fourier transformed to calculate the magnitude of the frequency spectrum for each chunk. Each chunk then corresponds to a vertical line in the image; a measurement of magnitude versus frequency for a specific moment in time. The spectrums or time plots are then "laid side by side" to form the image or a three-dimensional surface.

The spectrogram of a signal s(t) can be estimated by computing the squared magnitude of the STFT of the signal s(t), as shown below:

$$\text{spectrogram}(t,\omega) = |\text{STFT}(t,\omega)|^2$$

Applications



Spectrogram of dolphin vocalizations; chirps, clicks and harmonizing are visible as inverted Vs, vertical lines and horizontal striations respectively

- Early analog spectrograms were applied to a wide range of areas including the study of bird calls, with current research continuing using modern digital equipment and applied to all animal sounds. Contemporary use of the digital spectrogram is especially useful for studying frequency modulation (FM) in animal calls. Specifically, the distinguishing characteristics of FM chirps, broadband clicks, and social harmonizing are most easily visualized with the spectrogram. A particularly interesting example for the use of the spectrogram is in analysis of the vocalizations of a pod of Dolphins.
- Spectrograms are useful in assisting in overcoming speech defects and in speech training for the portion of the population that is profoundly deaf
- The studies of phonetics and speech synthesis are often facilitated through the use of spectrograms.
- By reversing the process of producing a spectrogram, it is possible to create a signal whose spectrogram is an arbitrary image. This technique can be used to hide a picture in a piece of audio and has been employed by several electronic music artists.
- Some modern music is created using spectrograms as an intermediate medium; changing the intensity of different frequencies over time, or even creating new ones, by drawing them and then inverse transforming.
- Spectrograms can be used to analyze the results of passing a test signal through a signal processor such as a filter in order to check its performance.
- High definition spectrograms are used in the development of RF and microwave systems
- Spectrograms are now used to display S-parameters measured with vector network analyzers
- The US Geological Survey now provides real-time spectrogram displays from seismic stations

Limitations and resynthesis

From the formula above, it appears that a spectrogram contains no information about the exact phase of the signal that it represents. For this reason, it is not possible to reverse the process and generate a copy of the original signal from a spectrogram, though in situations where the exact initial phase is unimportant (of which audio may be one), it may be possible to generate a useful approximation of the original signal. The Analysis & Resynthesis Sound Spectrograph is an example of a computer program that attempts to do this. In fact, there is some phase information in the spectrogram, but it appears in another form, as time delay (or group delay) which is the dual of the Instantaneous Frequency; an experiment explaining and relating these two concepts is described in

Chapter 9

Fractional Fourier Transform

In mathematics, in the area of harmonic analysis, the **fractional Fourier transform** (**FRFT**) is a linear transformation generalizing the Fourier transform. It can be thought of as the Fourier transform to the *n*-th power where *n* need not be an integer — thus, it can transform a function to an *intermediate* domain between time and frequency. Its applications range from filter design and signal analysis to phase retrieval and pattern recognition.

The FRFT can be used to define fractional convolution, correlation, and other operations, and can also be further generalized into the linear canonical transformation (LCT). An early definition of the FRFT was given by Namias, but it was not widely recognized until it was independently reinvented around 1993 by several groups of researchers. Since then, there has been a surge of interest in extending Shannon's sampling theorem for signals which are bandlimited in Fractional Fourier domain.

A completely different meaning for "fractional Fourier transform" was introduced by Bailey and Swartztrauber as essentially another name for a z-transform, and in particular for the case that corresponds to a discrete Fourier transform shifted by a fractional amount in frequency space (multiplying the input by a linear chirp) and evaluating at a fractional set of frequency points (e.g. considering only a small portion of the spectrum). (Such transforms can be evaluated efficiently by Bluestein's FFT algorithm.) This terminology has fallen out of use in most of the technical literature, however, in preference to the FRFT.

Definition

If the continuous Fourier transform of a function $f(t)$ is denoted by $\mathcal{F}(f)$, then $\mathcal{F}^2(f) = \mathcal{F}(\mathcal{F}(f))$, and in general $\mathcal{F}^{(n+1)}(f) = \mathcal{F}(\mathcal{F}^n(f))$; similarly,

$\mathcal{F}^{-n}(F)$ denotes the n-th power of the inverse transform $\mathcal{F}^{-1}(F)$ of $F(\omega)$. The FRFT further extends this definition to handle non-integer powers $n = 2\alpha / \pi$ for any real α, denoted by $\mathcal{F}_\alpha(f)$ and having the properties:

$$\mathcal{F}_\alpha(f) = \mathcal{F}^{2\alpha/\pi}(f)$$

when $n = 2\alpha / \pi$ is an integer, and:

$$\mathcal{F}_{\alpha+\beta}(f) = \mathcal{F}_\alpha(\mathcal{F}_\beta(f)) = \mathcal{F}_\beta(\mathcal{F}_\alpha(f)).$$

More specifically, $\mathcal{F}_\alpha(f)$ is given by the equation:

$$\mathcal{F}_\alpha(f)(\omega) = \sqrt{\frac{1 - i\cot(\alpha)}{2\pi}} e^{i\cot(\alpha)\omega^2/2} \int_{-\infty}^{\infty} e^{-i\csc(\alpha)\omega t + i\cot(\alpha)t^2/2} f(t)\, dt.$$

Note that, for $\alpha = \pi / 2$, this becomes precisely the definition of the continuous Fourier transform, and for $\alpha = -\pi / 2$ it is the definition of the inverse continuous Fourier transform.

If α is an integer multiple of π, then the cotangent and cosecant functions above diverge. However, this can be handled by taking the limit, and leads to a Dirac delta function in the integrand. More easily, since $\mathcal{F}^2(f) = f(-t)$, $\mathcal{F}_\alpha(f)$ must be simply $f(t)$ or $f(-t)$ for α an even or odd multiple of π, respectively.

Related transforms

There also exist related fractional generalizations of similar transforms such as the discrete Fourier transform. The **discrete fractional Fourier transform** is defined in (Candan, Kutay & Ozaktas 2000) and (Ozaktas, Zalevsky & Kutay 2001, Chapter 6).

Wavelet–fractional Fourier transforms: uses other orthonormal bases for $L^2(\mathbb{R})$ instead of Hermite–Gaussian functions. The new orthonormal basis is derivated from multiresolution analysis and orthonormal wavelets.

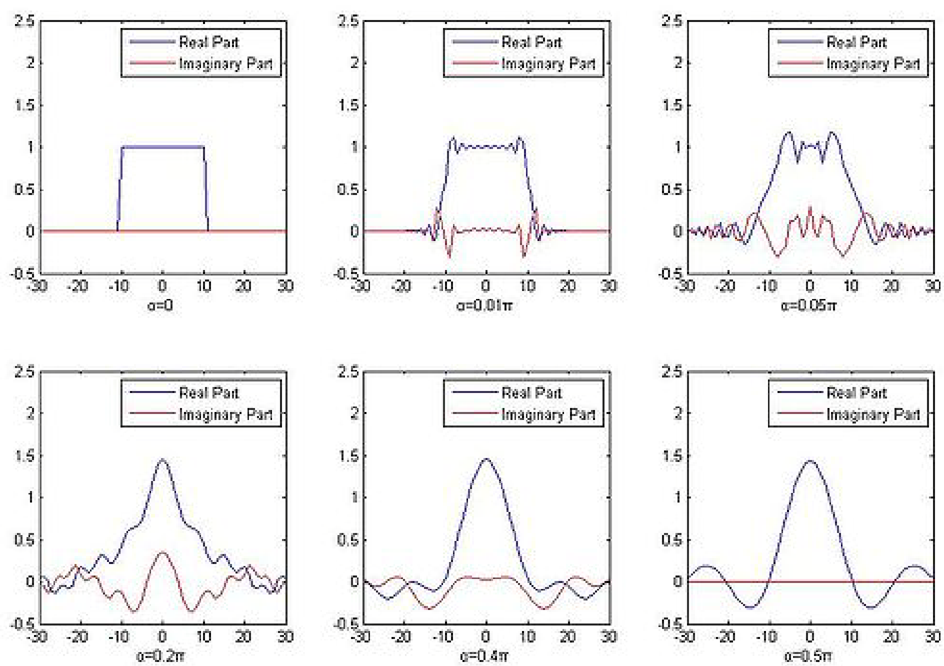
Generalization

The Fourier transform is essentially bosonic; it works because it is consistent with the superposition principle and related interference patterns. There is also a fermionic Fourier transform. These have been generalized into a supersymmetric FRFT, and a supersymmetric Radon transform. There is also a fractional Radon transform, a symplectic FRFT, and a symplectic wavelet transform. Because quantum circuits are based on unitary operations, they are useful for computing integral transforms as the latter are unitary operators on a function space. A quantum circuit has been designed which implements the FRFT.

Interpretation of the Fractional Fourier Transform

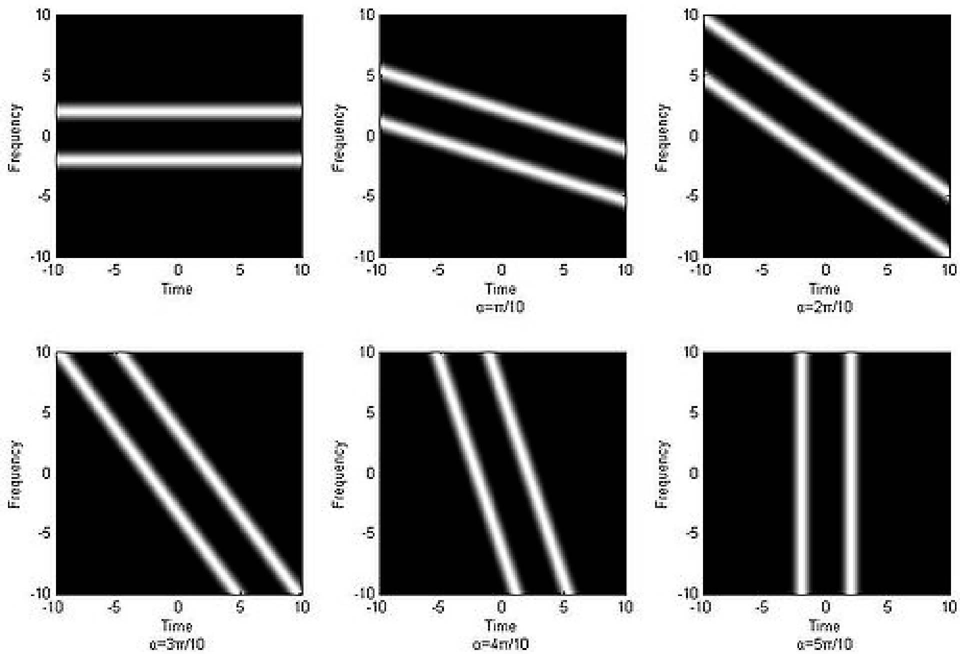
The usual interpretation of the Fourier transform is as a transformation of a time domain signal into a frequency domain signal. On the other hand, the interpretation of the inverse Fourier transform is as a transformation of a frequency domain signal into a time domain signal. Apparently, fractional Fourier transforms can transform a signal (either in the time domain or frequency domain) into the domain between time and frequency: it is a rotation in the time-frequency domain. This perspective is generalized by the linear canonical transformation, which generalizes the fractional Fourier transform and allows linear transforms of the time-frequency domain other than rotation.

Take the below figure as an example. If the signal in the time domain is rectangular (as below), it will become a sinc function in the frequency domain. But if we apply the fractional Fourier transform to the rectangular signal, the transformation output will be in the domain between time and frequency.



Fractional Fourier Transform

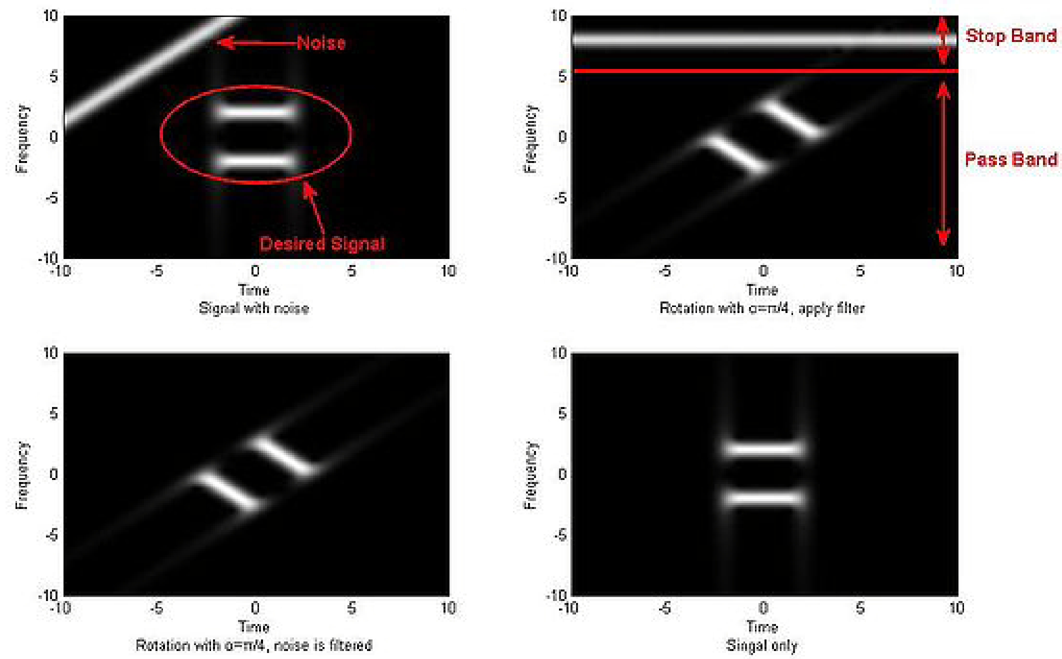
Actually, fractional Fourier transform is a rotation operation on the time frequency distribution. From the definition above, for $\alpha=0$, there will be no change after applying fractional Fourier transform, and for $\alpha=\pi/2$, fractional Fourier transform becomes a Fourier transform, which rotates the time frequency distribution with $\pi/2$. For other value of α, fractional Fourier transform rotates the time frequency distribution according to α. The following figure shows the results of the fractional Fourier transform with different values of α.



Time/Frequency Distribution of Fractional Fourier Transform.

Application

Fractional Fourier transform can be used in time frequency analysis and DSP. It is useful to filter noise, but with the condition that it does not overlap with the desired signal in the time frequency domain. Let's see the following example. We cannot apply a filter directly to eliminate the noise, but with the help of the fractional Fourier transform, we can rotate the signal (including the desired signal and noise) first. We then apply a specific filter which will allow only the desired signal to pass. Thus the noise will be removed completely. Then we use the fractional Fourier transform again to rotate the signal back and we can get the desired signal.



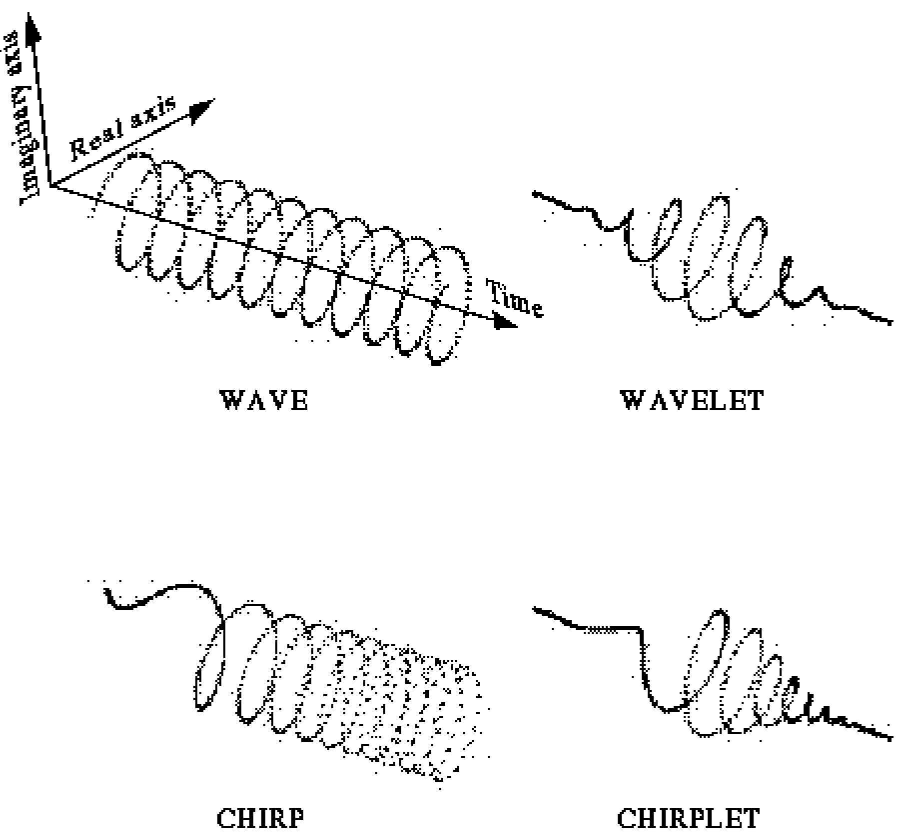
Fractional Fourier Transform in DSP.

Thus, using just truncation in the time domain, or equivalently low-pass filters in the frequency domain, one can cut out any convex set in time-frequency space; just using time domain or frequency domain methods without fractional Fourier transforms only allow cutting out rectangles parallel to the axes.

Chapter 10

Chirplet Transform & Instantaneous Phase

Chirplet Transform



Comparison of wave, wavelet, chirp, and chirplet

In signal processing, the **chirplet transform** is an inner product of an input signal with a family of analysis primitives called **chirplets**.

Similarity to other transforms

Much as in the wavelet transform, the chirplets are usually generated from (or can be expressed as being from) a single *mother chirplet* (analogous to the so-called *mother wavelet* of wavelet theory).

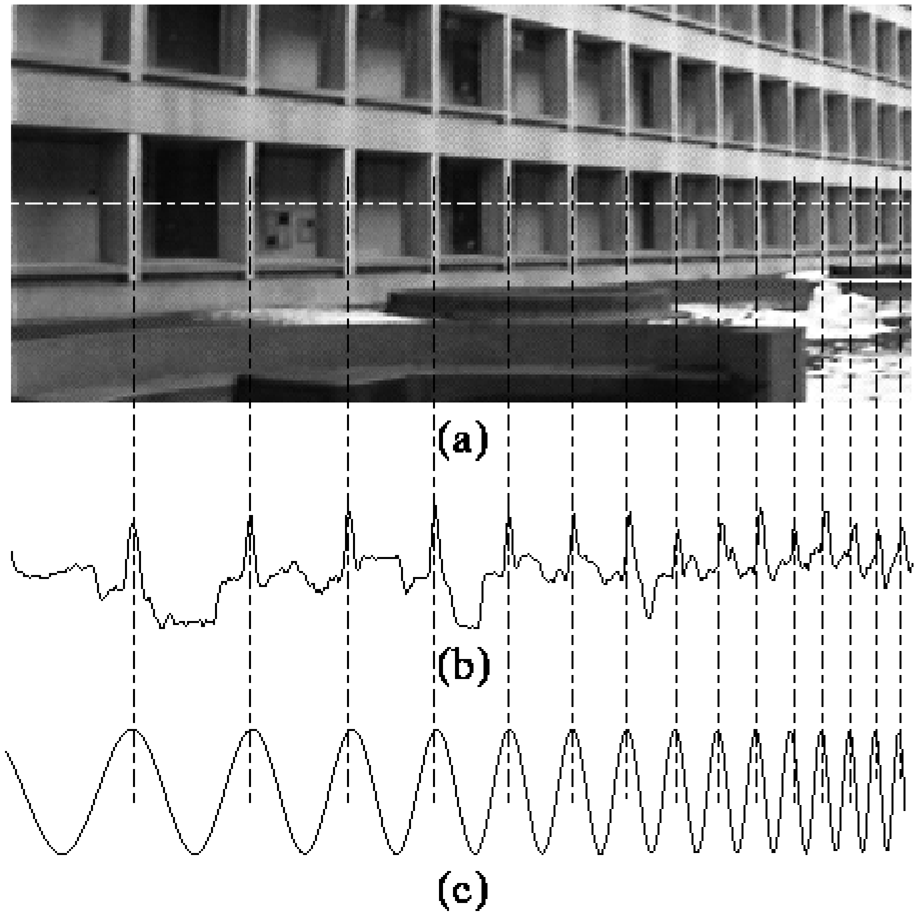
What is the chirplet and chirplet transform?

The term *chirplet transform* was coined by Steve Mann, as the title of the first published paper on chirplets. The term *chirplet* itself (apart from chirplet transform) was also used by Steve Mann, Domingo Mihovilovic, and Ronald Bracewell to describe a windowed portion of a chirp function. In Mann's words:

A wavelet is a piece of a wave, and a chirplet, similarly, is a piece of a chirp. More precisely, a chirplet is a windowed portion of a chirp function, where the window provides some time localization property. In terms of time–frequency space, chirplets exist as rotated, sheared, or other structures that move from the traditional parallelism with the time and frequency axes that are typical for waves (Fourier and short-time Fourier transforms) or wavelets.

The chirplet transform thus represents a rotated, sheared, or otherwise transformed tiling of the time–frequency plane. Although chirp signals have been known for many years in radar, pulse compression, and the like, the first published reference to the *chirplet transform* described specific signal representations based on families of functions related to one another by time–varying frequency modulation or frequency varying time modulation, in addition to time and frequency shifting, and scale changes. In that paper, the Gaussian chirplet transform was presented as one such example, together with a successful application to ice fragment detection in radar (improving target detection results over previous approaches). The term *chirplet* (but not the term *chirplet transform*) was also proposed for a similar transform, apparently independently, by Mihovilovic and Bracewell later that same year.

Applications



(a) In image processing, periodicity is often subject to a linear scaling. (b) In this image, repeating structures like the alternating dark space inside the windows, and light space of the white concrete, *chirp* (increase in frequency) towards the right. (c) The chirplet transform is able to represent this modulated variation compactly.

The chirplet transform is a useful signal analysis and representation framework that is widely used in

- Radar;
- Biomedical, the most commonly and widely used chirplet applications in biomedical being:
    - Heart sound processing;
    - EEG processing.
- Signal processing;
- Image processing;
- SETI@home uses chirp functions to compensate for Doppler drift;
- Chirplet Time Domain Reflectometry.

Taxonomy of chirplet transforms

There are two broad categories of chirplet transform:

- Fixed
- Adaptive

These categories may be further subdivided by:

- choice of chirp
- choice of window

In either the fixed or adaptive case, the chirplets may be:

- q-chirplets (quadratic chirplets) of the form $\exp[i2\pi(at^2 + bt + c)]$ or, in general, some kind of quadratically varying exponent, linear swept wave packet, or the like. These are sometimes called *linear FM chirplets* (linear frequency-modulated chirplets, since quadratic phase is linear frequency). Commonly used families of q-chirplets are metaplectomorphisms of one another (i.e., the energy distribution of any member of the family of q-chirplets can be generated from any other member by shear-in-time, shear-in-frequency, dilation, translation-in-time, and translation-in-frequency).
- w-chirplets, also known as *warblets*. A family of warblets are like the sound made by birds called warblers. Unwindowed warblets have a sinusoidally varying time–frequency distribution, or similar cyclostationary or periodically varying time–frequency plot. The sound of a police siren is an example, in which the pitch goes up and down periodically. Of course, the warblet is a "piece of" a warble (i.e., a windowed section of something that has a time–frequency periodicity).
- d-chirplets, also known as *Doppler chirplets*. These are analysis functions that mimic the Doppler shift of a passing tone (e.g., the sound you hear from a train whistle as it moves past).
- p-chirplets, in which the scale varies projectively. Whereas the wavelet transform is based on wavelets of the form $g(ax + b)$, the p-type chirplet transform is based on chirplets of the form $g((ax + b)/(cx + 1))$, where a is the scale, b is the translation, and c is the *chirpiness* (chirp-rate, as defined by the degree of perspective, or projection).

The choice of window is also another matter of decision. A Gaussian window is one possible choice, leading to a four parameter chirplet transform (for which time–shear and frequency–shear only give one degree of freedom that may thus be encapsulated as rotation angle—Radon transform of the Wigner distribution may, for example, be used, as may the fractional Fourier transform).

Another possible choice is the rectangular window, and discrete prolate spheroidal sequences (also called Multitaper#The_Slepian_sequences) may be used, by way of the *method of multiple mother chirplets*. This method gives a total chirplet transform as the sum of energies in various contributory chirplet transforms made from multiple windows, akin to the way in which DPSSs are used to get a perfect rectangular tiling of the time–frequency plane. Thus it is now possible to get perfect parallelogram tiling of the time–frequency plane, using the method of multiple mother chirplets.

Related work

The chirplet transform is a generalized representation that includes as special cases:

- The Fourier transform
- The short-time Fourier transform (STFT), also known as the spectrogram
- The Wigner-Ville distribution
- The wavelet transform
- Canonical conjugate variables
- Segal–Shale–Weil distribution

Josef Segman proposed the idea of incorporating scale into the Heisenberg group (position, momentum, phase, or equivalently any canonical conjugate variables taken together with phase, such as, for example, time, frequency, and phase). This gave rise to a four parameter space of time, frequency, phase, and scale. Segman introduced this idea of *phase scale*. (Personal communication with Mann, from Josef Segman, at Harvard University and at Massachusetts Institute of Technology). Further personal communication between Irving Segal (the principal behind the Segal, Shale Weil representation, known also as the metaplectic representation—a double covering of the symplectic group) and Mann led to additional insight into the chirplet transform, in particular, to the variation of the chirplet transform that is based on q-chirplets.

Instantaneous Phase

The notions of Instantaneous Phase and Instantaneous Frequency are important concepts in Signal Processing that occur in the context of the representation and analysis of time-varying signals . In signal processing, the **instantaneous phase** (or "local phase" or simply "phase") of a complex-valued function $x(t)$ is the real-valued function**:**

$$\phi(t) = \arg(x(t)).$$

And for a real-valued signal $s(t)$ it is determined from the signal's analytic representation, $s_{\mathrm{a}}(t)$**:**

$$\phi(t) = \arg(s_{\mathrm{a}}(t)).$$

When $\phi(t)$is constrained to an interval such as $(-\pi, \pi]$ or $[0, 2\pi)$, it is called the **wrapped** phase. Otherwise it is called **unwrapped**, which is a continuous function of argument t, assuming s_{a} is a continuous function of t. Unless otherwise indicated, the continuous form should be inferred.

Examples

Example 1: $s(t) = A \cdot \cos(\omega t + \theta)$, where A and ω are positive values.

$s_{\mathrm{a}}(t) = A \cdot e^{i(\omega t+\theta)}$

$\phi(t) = \omega t + \theta$

Example 2: $s(t) = A \cdot \sin(\omega t) = A \cdot \cos\left(\omega t - \frac{\pi}{2}\right)$

$s_{\mathrm{a}}(t) = A \cdot e^{i\left(\omega t - \frac{\pi}{2}\right)}$

$\phi(t) = \omega t - \frac{\pi}{2}$

For both of these sinusoidal examples, the local maxima of $s(t)$ correspond to

$$\phi(t) = N \cdot 2\pi$$

for integer values of N. Similarly, the local minima correspond to

$$\phi(t) = \pi + N \cdot 2\pi,$$

and the maximum rates of change correspond to

$$\phi(t) = \frac{\pi}{2} + N \cdot \pi.$$

For signals that are approximately sinusoidal, these properties can be used, e.g., in image processing and computer vision, to detect points that are close to edges or lines, and also to measure the position of these points with sub-pixel accuracy.

Instantaneous frequency

In general, the instantaneous angular frequency is defined as

$$\omega(t) = \phi'(t) = \frac{d}{dt}\phi(t)$$

and the **instantaneous frequency** (Hz) is.

$$f(t) = \frac{1}{2\pi}\phi'(t)$$
.

Conversely, the unwrapped phase can be represented in terms of an instantaneous frequency. When it is actually constructed/derived this way, this process is called **phase unwrapping:**

$$\begin{aligned}\phi(t) = 2\pi \int_{-\infty}^{t} f(\tau)\,d\tau &= 2\pi \int_{0}^{t} f(\tau)\,d\tau + 2\pi \int_{-\infty}^{0} f(t)\,dt \\ &= 2\pi \int_{0}^{t} f(\tau)\,d\tau + \phi(0)\end{aligned}$$

A similar definition can be used to describe phase as a function of frequency. In this case, the complex notation of phase generates values from +-180 degrees, but for unwrapped phase, an additional 360 degrees of phase must be added when the phase moves from +180 to -180 degrees. In this way, the phase will accumulate to any arbitrarily large value. Note from the definition that the phase function must exist for all frequency from f=0. In many cases, such as RF and Microwave measurements of cables or filters, the phase response is not measured continuously, but from some non-zero start frequency. In this case, phase unwrapping has ambiguities due to the lack of knowledge of phase response before the start frequency. In this case, it is common practice to assign the integral a value of zero for phase response before the start frequency. For devices such as cables, which have nearly constant phase vs frequency, this can give an error in the unwrapped phase response compared to the expected value as the phase before the start frequency is not included.

Complex representation

In some applications, such as averaging the values of phase at several moments of time, it may be useful to convert each value to a complex number, or vector representation**:**

$$\begin{aligned}e^{i\phi(t)} &= \frac{s_{\mathrm{a}}(t)}{|s_{\mathrm{a}}(t)|} \\ &= \cos(\phi(t)) + i \cdot \sin(\phi(t)).\end{aligned}$$ (Euler's formula)

This representation is similar to the wrapped phase representation in that it does not distinguish between multiples of 2π in the phase, but similar to the unwrapped phase representation since it is continuous. A vector-average phase can be obtained as the **arg** of the sum of the complex numbers. A full treatment of the question of instantaneous frequency estimation, computation and its application to real-life applications is given in and .

Chapter 11

Linear Canonical Transformation & Time–Frequency Representation

Linear Canonical Transformation

In Hamiltonian mechanics, the **linear canonical transformation** (**LCT**) is a family of integral transforms that generalizes many classical transforms. It has 4 parameters and 1 constraint, so it is a 3-dimensional family, and can be visualized as the action of the special linear group $SL_2(\mathbf{R})$ on the time–frequency plane (domain).

The LCT generalizes the Fourier, fractional Fourier, Laplace, Gauss–Weierstrass, Bargmann and the Fresnel transforms as particular cases. The name "linear canonical transformation" is from canonical transformation, a map that preserves the symplectic structure, as $SL_2(\mathbf{R})$ can also be interpreted as the symplectic group Sp_2, and thus LCTs are the linear maps of the time–frequency domain which preserve the symplectic form.

Definition

The LCT can be represented in several ways; most easily, it can be viewed as a 2×2 matrix with determinant 1, i.e., an element of the special linear group $SL_2(\mathbf{R})$. Taking a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad - bc = 1$, the corresponding integral transform is:

$$X_{(a,b,c,d)}(u) = \sqrt{-i} \cdot e^{i\pi \frac{d}{b} u^2} \int_{-\infty}^{\infty} e^{-i2\pi \frac{1}{b} ut} e^{i\pi \frac{a}{b} t^2} x(t)\, dt\,, \quad \text{when } b \neq 0,$$

$$X_{(a,0,c,d)}(u) = \sqrt{d} \cdot e^{i\pi c d u^2} x(du)\,, \quad \text{when } b = 0.$$

Special cases

Many classical transforms are special cases of the linear canonical transform:

- The Fourier transform corresponds to rotation by 90°, represented by the matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

- The fractional Fourier transform corresponds to rotation by an arbitrary angle; they are the elliptic elements of $SL_2(\mathbf{R})$, represented by the matrices:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

- The Fresnel transform corresponds to shearing, and are a family of parabolic elements, represented by the matrices:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & \lambda z \\ 0 & 1 \end{bmatrix}.$$

where z is distance and λ is wave length.

Composition

Composition of LCTs corresponds to multiplication of the corresponding matrices; this is also known as the "additivity property of the WDF".

In detail, if the LCT is denoted by $O_F^{(a,b,c,d)}$, i.e.

$$X_{(a,b,c,d)}(u) = O_F^{(a,b,c,d)}[x(t)]$$

then

$$O_F^{(a2,b2,c2,d2)}\left\{O_F^{(a1,b1,c1,d1)}[x(t)]\right\} = O_F^{(a3,b3,c3,d3)}[x(t)]\ ,$$

where

$$\begin{bmatrix} a3 & b3 \\ c3 & d3 \end{bmatrix} = \begin{bmatrix} a2 & b2 \\ c2 & d2 \end{bmatrix} \begin{bmatrix} a1 & b1 \\ c1 & d1 \end{bmatrix}.$$

In optics and quantum mechanics

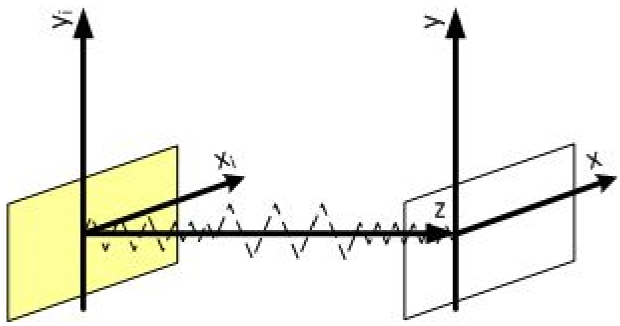
Paraxial optical systems implemented entirely with thin lenses and propagation through free space and/or graded index (GRIN) media, are quadratic phase systems (QPS); these were known before Moshinsky and Quesne (1974) called attention to their significance in connection with canonical transformations in quantum mechanics. The effect of any arbitrary QPS on an input wavefield can be described using the linear canonical transform, a particular case of which was developed by Segal (1963) and Bargmann (1961) in order to formalize Fok's (1928) boson calculus.

Applications

Canonical transforms provide a fine tool for the analysis of a class of differential equations. These include diffusion, the Schrödinger free particle, the linear potential (free-fall), and the attractive and repulsive oscillator equations. It also includes a few others such as the Fokker–Planck equation. Although this class is far from universal, the ease with which solutions and properties are found makes canonical transforms an attractive tool for problems such as these.

Wave propagation through air, a lens, and between satellite dishes are discussed here. All of the computations can be reduced to 2×2 matrix algebra. This is the spirit of LCT.

Electromagnetic wave propagation



Assuming the system looks like as depicted in the figure, the wave travels from plane x_i, y_i to the plane of x and y. The Fresnel transform is used to describe the electromagnetic wave propagation in air:

$$U_0(x,y) = -\frac{j}{\lambda}\frac{e^{jkz}}{z}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j\frac{k}{2z}[(x-x_i)^2+(y-y_i)^2]}U_i(x_i,y_i)\;dx_i\;dy_i,$$

with

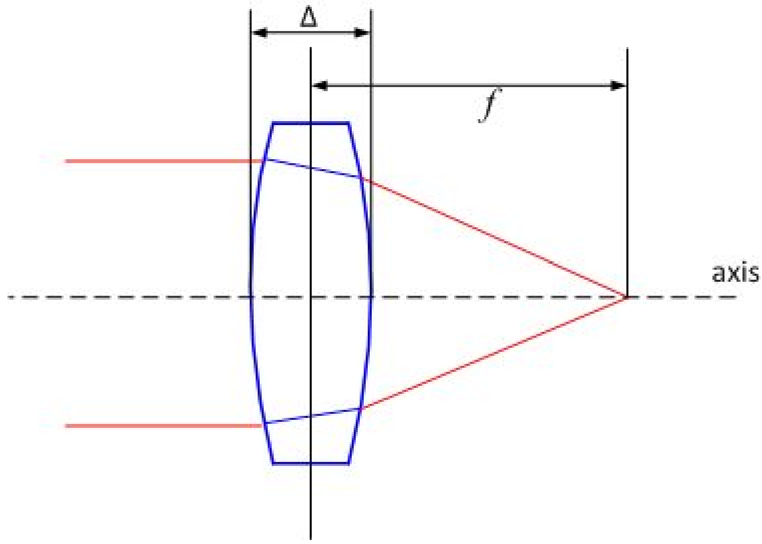
$k = 2\pi/\lambda$	: wave number;
λ	: wavelength;
z	: distance of propagation;
j	: imaginary unit.

This is equivalent to LCT (shearing), when

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & \lambda z \\ 0 & 1 \end{bmatrix}.$$

When the travel distance (z) is larger, the shearing effect is larger.

Spherical lens



With the lens as depicted in the figure, and the refractive index denoted as n, the result is:

$$U_0(x,y) = e^{jkn\Delta} e^{-j\frac{k}{2f}[x^2+y^2]} U_i(x,y)$$

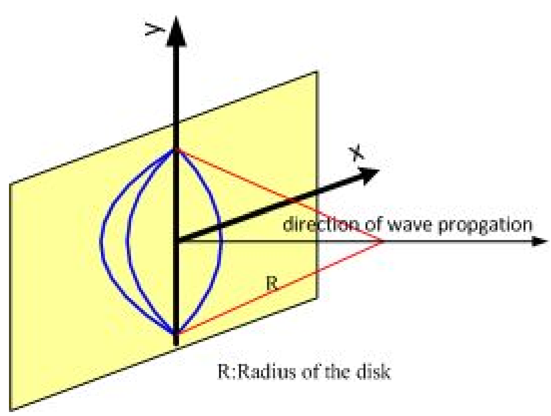
with f the focal length and Δ the thickness of the lens.

The distortion passing through the lens is similar to LCT, when

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{-1}{\lambda f} & 1 \end{bmatrix}.$$

This is also a shearing effect: when the focal length is smaller, the shearing effect is larger.

Satellite dish

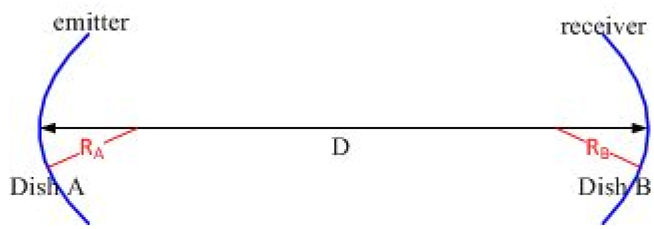


The satellite dish can be described as a LCT, with

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{-1}{\lambda R} & 1 \end{bmatrix}.$$

This is very similar to lens, except focal length is replaced by the radius of the dish. Therefore, if the radius is larger, the shearing effect is larger.

Example



The system considered is depicted in the figure to the right: two dishes – one being the emitter and the other one the receiver – and a signal travelling between them over a distance *D*. First, for dish A (emitter), the LCT matrix looks like this:

$$\begin{bmatrix} 1 & 0 \\ \frac{-1}{\lambda R_A} & 1 \end{bmatrix}.$$

Then, for dish B (receiver), the LCT matrix similarly becomes:

$$\begin{bmatrix} 1 & 0 \\ \frac{-1}{\lambda R_B} & 1 \end{bmatrix}.$$

Last, for the propagation of the signal in air, the LCT matrix is:

$$\begin{bmatrix} 1 & \lambda D \\ 0 & 1 \end{bmatrix}.$$

Putting all three components together, the LCT of the system is:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{-1}{\lambda R_B} & 1 \end{bmatrix} \begin{bmatrix} 1 & \lambda D \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{-1}{\lambda R_A} & 1 \end{bmatrix} = \begin{bmatrix} 1 - \frac{D}{R_A} & -\lambda D \\ \frac{1}{\lambda}(R_A^{-1} + R_B^{-1} - R_A^{-1} R_B^{-1} D) & 1 - \frac{D}{R_B} \end{bmatrix}.$$

Time–Frequency Representation

A **time–frequency representation** (**TFR**) is a view of a signal (taken to be a function of time) represented over both time ***and*** frequency. Time–frequency analysis means analysis into the time–frequency domain provided by a TFR. This is achieved by using a formulation often called "Time–Frequency Distribution", abbreviated as TFD.

TFRs are often complex-valued fields over time and frequency, where the modulus of the field represents "energy density" (the concentration of the root mean square over time and frequency) or amplitude, and the argument of the field represents phase.

Background and motivation

A signal, as a function of time, may be considered as a representation with perfect *time resolution*. In contrast, the magnitude of the Fourier transform (FT) of the signal may be considered as a representation with perfect *spectral resolution* but with no time information because the magnitude of the FT conveys frequency content but it fails to convey when, in time, different events occur in the signal.

TFRs provide a bridge between these two representations in that they provide *some* temporal information ***and*** *some* spectral information simultaneously. Thus, TFRs are useful for the representation and analysis of signals containing multiple time-varying frequencies.

Formulation of TFRs and TFDs

Quadratic forms

One form of TFR (or TFD) can be formulated by the multiplicative comparison of a signal with itself, expanded in different directions about each point in time. Such representations and formulations are known as quadratic TFRs or TFDs (QTFRs or QTFDs) because the representation is quadratic in the signal. This formulation was first described by Eugene Wigner in 1932 in the context of quantum mechanics and, later, reformulated as a general TFR by Ville in 1948 to form what is now known as the Wigner–Ville distribution, as it was shown in that Wigner's formula needed to use the anaytic signal defined in Ville's paper to be useful as a representation and for a practical analysis. Today, various QTFRs include but not limited to spectrogram (squared magnitude of short-time Fourier transform), scaleogram (squared magnitude of Wavelet transform) and the smoothed pseudo-Wigner distribution. In fact, a whole class of representations using bilinear time–frequency distributions fall in this category.

Although quadratic TFRs offer perfect temporal and spectral resolutions simultaneously, the quadratic nature of the transforms creates cross-terms. The following can be used to estimate which QTFRs contain cross terms.

Given a QTFR $E(t,f)$ defined on $\mathbb{R}^2$, define a constant $E_0 = \mathrm{Sup}|E(t,f)|,\ (t,f) \in \mathbb{R}^2$ and a set $\mathcal{C} = \{(t,f) \in \mathbb{R}^2 : |E(t,f)| > T,\ \forall T \in [0, E_0]\}$. The QTFR, $E(t,f)$ is cross-term free if $\mathcal{C}$ is a convex set.

Linear forms

The cross-terms caused by the bilinear structure of TFDs and TFRs may be useful in some applications such as classification as the cross-terms provide extra detail for the recognition algorithm. However, in some other applications, these cross-terms may plague certain quadratic TFRs and they would need to be reduced. One way to do this is obtained by comparing the signal with a different function. Such resulting representations are known as linear TFRs because the representation is linear in the signal.

The *windowed Fourier transform* (also known as the short-time Fourier transform) localises the signal by modulating it with a window function, before performing the Fourier transform to obtain the frequency content of the signal in the region of the window.

Wavelet transforms

Wavelet transforms, in particular the continuous wavelet transform, expand the signal in terms of wavelet functions which are localised in both time and frequency. Thus the wavelet transform of a signal may be represented in terms of both time and frequency.

The notions of time, frequency, and amplitude used to generate a TFR from a wavelet transform were originally developed intuitively. In 1992, Delprat et al. gave a quantitative derivation of these relationships, based upon a stationary phase approximation.

Linear canonical transformation

Linear canonical transformations are the linear transforms of the time–frequency representation that preserve the symplectic form. These include and generalize the Fourier transform, fractional Fourier transform, and others, thus providing a unified view of these transforms in terms of their action on the time–frequency domain.

Chapter 12

Wavelet

A **wavelet** is a wave-like oscillation with an amplitude that starts out at zero, increases, and then decreases back to zero. It can typically be visualized as a "brief oscillation" like one might see recorded by a seismograph or heart monitor. Generally, wavelets are purposefully crafted to have specific properties that make them useful for signal processing. Wavelets can be combined, using a "shift, multiply and sum" technique called convolution, with portions of an unknown signal to extract information from the unknown signal.

For example, a wavelet could be created to have a frequency of Middle C and a short duration of roughly a 32nd note. If this wavelet were to be convolved at periodic intervals with a signal created from the recording of a song, then the results of these convolutions would be useful for determining when the Middle C note was being played in the song. Mathematically, the wavelet will resonate if the unknown signal contains information of similar frequency - just as a tuning fork physically resonates with sound waves of its specific tuning frequency. This concept of resonance is at the core of many practical applications of wavelet theory.

As wavelets are a mathematical tool they can be used to extract information from many different kinds of data, including - but certainly not limited to - audio signals and images. Sets of wavelets are generally needed to analyze data fully. A set of "complementary" wavelets will deconstruct data without gaps or overlap so that the deconstruction process is mathematically reversible. Thus, sets of complementary wavelets are useful in wavelet based compression/decompression algorithms where it is desirable to recover the original information with minimal loss.

In formal terms, this representation is a wavelet series representation of a square-integrable function with respect to either a complete, orthonormal set of basis functions, or an overcomplete set or Frame of a vector space, for the Hilbert space of square integrable functions.

Name

The word *wavelet* has been used for decades in digital signal processing and exploration geophysics. The equivalent French word *ondelette* meaning "small wave" was used by Morlet and Grossmann in the early 1980s.

Wavelet theory

Wavelet theory is applicable to several subjects. All wavelet transforms may be considered forms of time-frequency representation for continuous-time (analog) signals and so are related to harmonic analysis. Almost all practically useful discrete wavelet transforms use discrete-time filterbanks. These filter banks are called the wavelet and scaling coefficients in wavelets nomenclature. These filterbanks may contain either finite impulse response (FIR) or infinite impulse response (IIR) filters. The wavelets forming a continuous wavelet transform (CWT) are subject to the uncertainty principle of Fourier analysis respective sampling theory: Given a signal with some event in it, one cannot assign simultaneously an exact time and frequency response scale to that event. The product of the uncertainties of time and frequency response scale has a lower bound. Thus, in the scaleogram of a continuous wavelet transform of this signal, such an event marks an entire region in the time-scale plane, instead of just one point. Also, discrete wavelet bases may be considered in the context of other forms of the uncertainty principle.

Wavelet transforms are broadly divided into three classes: continuous, discrete and multiresolution-based.

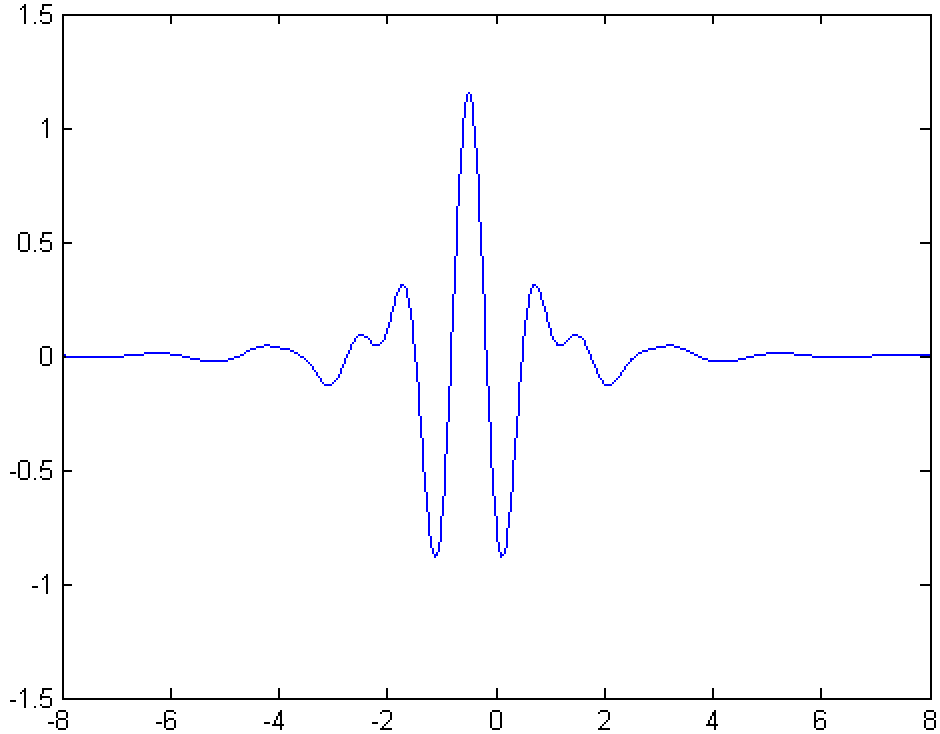
Continuous wavelet transforms (continuous shift and scale parameters)

In continuous wavelet transforms, a given signal of finite energy is projected on a continuous family of frequency bands (or similar subspaces of the L^p function space $L^2(\mathbb{R})$). For instance the signal may be represented on every frequency band of the form $[f,2f]$ for all positive frequencies $f>0$. Then, the original signal can be reconstructed by a suitable integration over all the resulting frequency components.

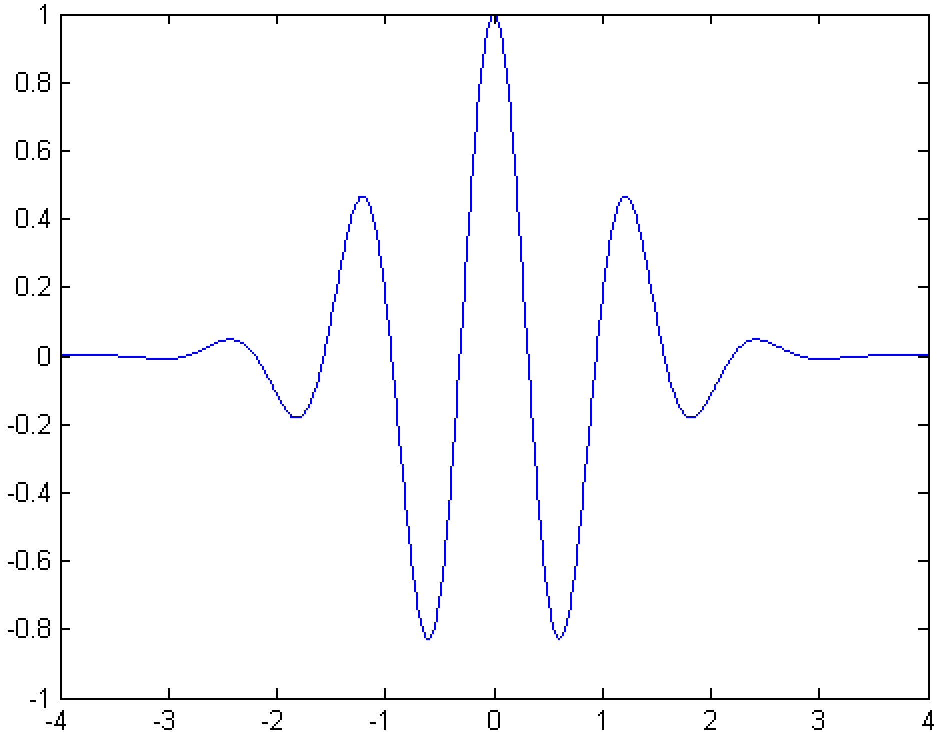
The frequency bands or subspaces (sub-bands) are scaled versions of a subspace at scale *1*. This subspace in turn is in most situations generated by the shifts of one generating function $\psi \in L^2(\mathbb{R})$, the *mother wavelet*. For the example of the scale one frequency band [1,2] this function is

$$\psi(t) = 2\,\mathrm{sinc}(2t) - \,\mathrm{sinc}(t) = \frac{\sin(2\pi t) - \sin(\pi t)}{\pi t}$$

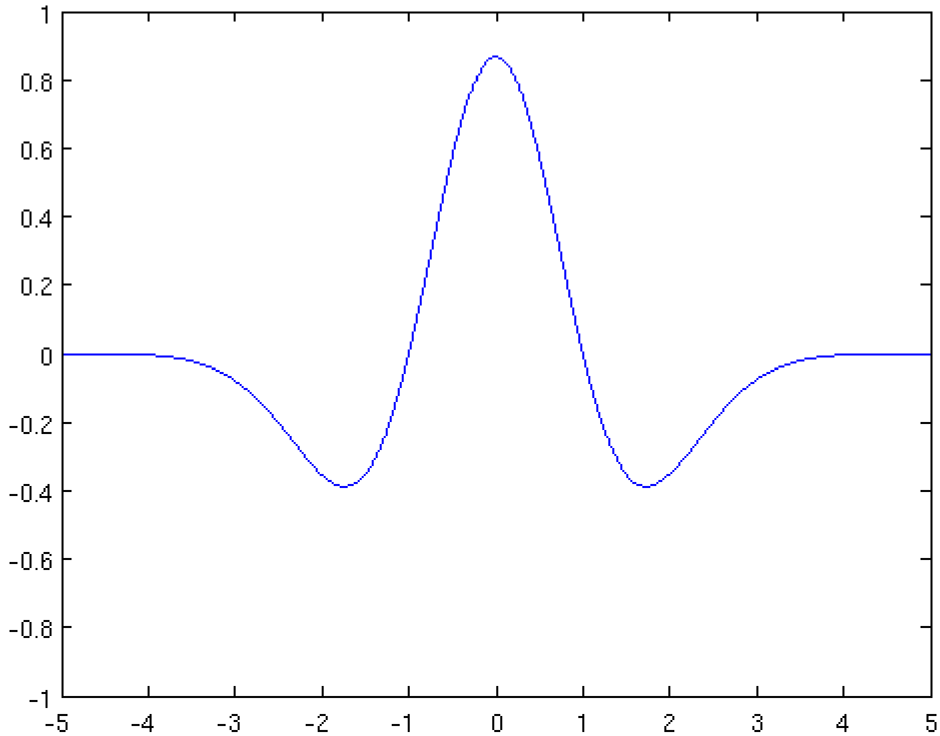
with the (normalized) sinc function. Other example mother wavelets are:



Meyer



Morlet



Mexican Hat

The subspace of scale *a* or frequency band $[1/a,\ 2/a]$ is generated by the functions (sometimes called *child wavelets*)

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right),$$

where *a* is positive and defines the scale and *b* is any real number and defines the shift. The pair *(a,b)* defines a point in the right halfplane $\mathbb{R}_+ \times \mathbb{R}$.

The projection of a function *x* onto the subspace of scale *a* then has the form

$$x_a(t) = \int_{\mathbb{R}} WT_\psi\{x\}(a,b)\cdot\psi_{a,b}(t)\,db$$

with *wavelet coefficients*

$$WT_\psi\{x\}(a,b) = \langle x, \psi_{a,b}\rangle = \int_{\mathbb{R}} x(t)\overline{\psi_{a,b}(t)}\,dt.$$

For the analysis of the signal x, one can assemble the wavelet coefficients into a scaleogram of the signal.

Discrete wavelet transforms (discrete shift and scale parameters)

It is computationally impossible to analyze a signal using all wavelet coefficients, so one may wonder if it is sufficient to pick a discrete subset of the upper halfplane to be able to reconstruct a signal from the corresponding wavelet coefficients. One such system is the affine system for some real parameters $a>1$, $b>0$. The corresponding discrete subset of the halfplane consists of all the points $(a^m, n\, a^m b)$ with integers $m, n \in \mathbb{Z}$. The corresponding *baby wavelets* are now given as

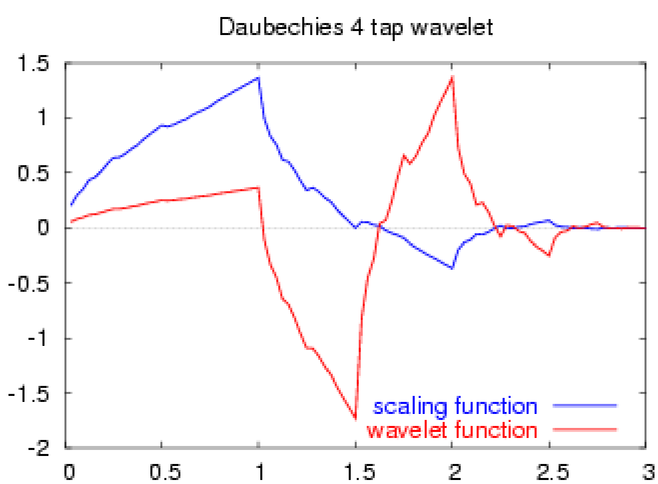
$$\psi_{m,n}(t) = a^{-m/2}\psi(a^{-m}t - nb).$$

A sufficient condition for the reconstruction of any signal x of finite energy by the formula

$$x(t) = \sum_{m\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} \langle x,\, \psi_{m,n}\rangle \cdot \psi_{m,n}(t)$$

is that the functions $\{\psi_{m,n} : m, n \in \mathbb{Z}\}$ form a tight frame of $L^2(\mathbb{R})$.

Multiresolution discrete wavelet transforms



D4 wavelet

In any discretised wavelet transform, there are only a finite number of wavelet coefficients for each bounded rectangular region in the upper halfplane. Still, each coefficient requires the evaluation of an integral. To avoid this numerical complexity, one needs one auxiliary function, the *father wavelet* $\phi \in L^2(\mathbb{R})$. Further, one has to restrict *a* to be an integer. A typical choice is *a=2* and *b=1*. The most famous pair of father and mother wavelets is the Daubechies 4 tap wavelet.

From the mother and father wavelets one constructs the subspaces

$$V_m = \operatorname{span}(\phi_{m,n} : n \in \mathbb{Z}), \text{ where } \varphi_{m,n}(t) = 2^{-m/2}\varphi(2^{-m}t - n)$$

and

$$W_m = \operatorname{span}(\psi_{m,n} : n \in \mathbb{Z}), \text{ where } \psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n).$$

From these one requires that the sequence

$$\{0\} \subset \cdots \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots \subset L^2(\mathbb{R})$$

forms a multiresolution analysis of $L^2(\mathbb{R})$ and that the subspaces $\ldots, W_1, W_0, W_{-1}, \ldots\ldots$ are the orthogonal "differences" of the above sequence, that is, W_m is the orthogonal complement of V_m inside the subspace V_{m-1}. In analogy to the sampling theorem one may conclude that the space V_m with sampling distance 2^m more or less covers the frequency baseband from *0* to 2^{-m-1}. As orthogonal complement, W_m roughly covers the band $[2^{-m-1}, 2^{-m}]$.

From those inclusions and orthogonality relations follows the existence of sequences $h = \{h_n\}_{n\in\mathbb{Z}}$ and $g = \{g_n\}_{n\in\mathbb{Z}}$ that satisfy the identities

$$h_n = \langle \phi_{0,0},\ \phi_{-1,n} \rangle \text{ and } \phi(t) = \sqrt{2}\sum_{n\in\mathbb{Z}} h_n \phi(2t - n)$$

and

$$g_n = \langle \psi_{0,0},\ \phi_{-1,n} \rangle \text{ and } \psi(t) = \sqrt{2}\sum_{n\in\mathbb{Z}} g_n \phi(2t - n).$$

The second identity of the first pair is a refinement equation for the father wavelet φ. Both pairs of identities form the basis for the algorithm of the fast wavelet transform.

Mother wavelet

For practical applications, and for efficiency reasons, one prefers continuously differentiable functions with compact support as mother (prototype) wavelet (functions). However, to satisfy analytical requirements (in the continuous WT) and in general for theoretical reasons, one chooses the wavelet functions from a subspace of the space

$L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. This is the space of measurable functions that are absolutely and square integrable:

$$\int_{-\infty}^{\infty} |\psi(t)|\, dt < \infty \text{ and } \int_{-\infty}^{\infty} |\psi(t)|^2\, dt < \infty.$$

Being in this space ensures that one can formulate the conditions of zero mean and square norm one:

$\int_{-\infty}^{\infty} \psi(t)\, dt = 0$ is the condition for zero mean, and
$\int_{-\infty}^{\infty} |\psi(t)|^2\, dt = 1$ is the condition for square norm one.

For ψ to be a wavelet for the continuous wavelet transform, the mother wavelet must satisfy an admissibility criterion (loosely speaking, a kind of half-differentiability) in order to get a stably invertible transform.

For the discrete wavelet transform, one needs at least the condition that the wavelet series is a representation of the identity in the space $L^2(\mathbb{R})$. Most constructions of discrete WT make use of the multiresolution analysis, which defines the wavelet by a scaling function. This scaling function itself is solution to a functional equation.

In most situations it is useful to restrict ψ to be a continuous function with a higher number *M* of vanishing moments, i.e. for all integer *m*<*M*

$$\int_{-\infty}^{\infty} t^m\, \psi(t)\, dt = 0.$$

The mother wavelet is scaled (or dilated) by a factor of *a* and translated (or shifted) by a factor of *b* to give (under Morlet's original formulation):

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right).$$

For the continuous WT, the pair *(a,b)* varies over the full half-plane $\mathbb{R}_+ \times \mathbb{R}$; for the discrete WT this pair varies over a discrete subset of it, which is also called *affine group*.

These functions are often incorrectly referred to as the basis functions of the (continuous) transform. In fact, as in the continuous Fourier transform, there is no basis in the continuous wavelet transform. Time-frequency interpretation uses a subtly different formulation (after Delprat).

Comparisons with Fourier transform (continuous-time)

The wavelet transform is often compared with the Fourier transform, in which signals are represented as a sum of sinusoids. The main difference is that wavelets are localized in both time and frequency whereas the standard Fourier transform is only localized in frequency. The Short-time Fourier transform (STFT) is more similar to the wavelet transform, in that it is also time and frequency localized, but there are issues with the frequency/time resolution trade-off. Wavelets often give a better signal representation using Multiresolution analysis, with balanced resolution at any time and frequency.

The discrete wavelet transform is also less computationally complex, taking $O(N)$ time as compared to $O(N \log N)$ for the fast Fourier transform. This computational advantage is not inherent to the transform, but reflects the choice of a logarithmic division of frequency, in contrast to the equally spaced frequency divisions of the FFT(Fast Fourier Transform). It is also important to note that this complexity only applies when the filter size has no relation to the signal size. A wavelet without compact support such as the Shannon wavelet would require $O(N^2)$. (For instance, a logarithmic Fourier Transform also exists with $O(N)$ complexity, but the original signal must be sampled logarithmically in time, which is only useful for certain types of signals.)

Definition of a wavelet

There are a number of ways of defining a wavelet (or a wavelet family).

Scaling filter

An orthogonal wavelet is entirely defined by the scaling filter - a low-pass finite impulse response (FIR) filter of length $2N$ and sum 1. In biorthogonal wavelets, separate decomposition and reconstruction filters are defined.

For analysis with orthogonal wavelets the high pass filter is calculated as the quadrature mirror filter of the low pass, and reconstruction filters are the time reverse of the decomposition filters.

Daubechies and Symlet wavelets can be defined by the scaling filter.

Scaling function

Wavelets are defined by the wavelet function $\psi(t)$ (i.e. the mother wavelet) and scaling function $\varphi(t)$ (also called father wavelet) in the time domain.

The wavelet function is in effect a band-pass filter and scaling it for each level halves its bandwidth. This creates the problem that in order to cover the entire spectrum, an infinite number of levels would be required. The scaling function filters the lowest level of the transform and ensures all the spectrum is covered.

For a wavelet with compact support, $\varphi(t)$ can be considered finite in length and is equivalent to the scaling filter g.

Meyer wavelets can be defined by scaling functions

Wavelet function

The wavelet only has a time domain representation as the wavelet function $\psi(t)$.

Applications of discrete wavelet transform

Generally, an approximation to DWT is used for data compression if signal is already sampled, and the CWT for signal analysis. Thus, DWT approximation is commonly used in engineering and computer science, and the CWT in scientific research.

Wavelet transforms are now being adopted for a vast number of applications, often replacing the conventional Fourier Transform. Many areas of physics have seen this paradigm shift, including molecular dynamics, ab initio calculations, astrophysics, density-matrix localisation, seismology, optics, turbulence and quantum mechanics. This change has also occurred in image processing, blood-pressure, heart-rate and ECG analyses, DNA analysis, protein analysis, climatology, general signal processing, speech recognition, computer graphics and multifractal analysis. In computer vision and image processing, the notion of scale-space representation and Gaussian derivative operators is regarded as a canonical multi-scale representation.

One use of wavelet approximation is in data compression. Like some other transforms, wavelet transforms can be used to transform data, then encode the transformed data, resulting in effective compression. For example, JPEG 2000 is an image compression standard that uses biorthogonal wavelets. This means that although the frame is overcomplete, it is a *tight frame*, and the same frame functions (except for conjugation in the case of complex wavelets) are used for both analysis and synthesis, i.e., in both the forward and inverse transform.

A related use is that of smoothing/denoising data based on wavelet coefficient thresholding, also called wavelet shrinkage. By adaptively thresholding the wavelet coefficients that correspond to undesired frequency components smoothing and/or denoising operations can be performed.

Wavelet transforms are also starting to be used for communication applications. Wavelet OFDM is the basic modulation scheme used in HD-PLC (a powerline communications technology developed by Panasonic), and in one of the optional modes included in the IEEE P1901 draft standard. The advantage of Wavelet OFDM over traditional FFT OFDM systems is that Wavelet can achieve deeper notches and that it does not require a Guard Interval (which usually represents significant overhead in FFT OFDM systems).

History

The development of wavelets can be linked to several separate trains of thought, starting with Haar's work in the early 20th century. Notable contributions to wavelet theory can be attributed to Zweig's discovery of the continuous wavelet transform in 1975 (originally called the cochlear transform and discovered while studying the reaction of the ear to sound), Pierre Goupillaud, Grossmann and Morlet's formulation of what is now known as the CWT (1982), Jan-Olov Strömberg's early work on discrete wavelets (1983), Daubechies' orthogonal wavelets with compact support (1988), Mallat's multiresolution framework (1989), Nathalie Delprat's time-frequency interpretation of the CWT (1991), Newland's Harmonic wavelet transform (1993) and many others since.

Timeline

- First wavelet (Haar wavelet) by Alfred Haar (1909)
- Since the 1970s: George Zweig, Jean Morlet, Alex Grossmann
- Since the 1980s: Yves Meyer, Stéphane Mallat, Ingrid Daubechies, Ronald Coifman, Victor Wickerhauser,

Wavelet transforms

A wavelet is a mathematical function used to divide a given function or continuous-time signal into different scale components. Usually one can assign a frequency range to each scale component. Each scale component can then be studied with a resolution that matches its scale. A wavelet transform is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as "daughter wavelets") of a finite-length or fast-decaying oscillating waveform (known as the "mother wavelet"). Wavelet transforms have advantages over traditional Fourier transforms for representing functions that have discontinuities and sharp peaks, and for accurately deconstructing and reconstructing finite, non-periodic and/or non-stationary signals.

Wavelet transforms are classified into discrete wavelet transforms (DWTs) and continuous wavelet transforms (CWTs). Note that both DWT and CWT are continuous-time (analog) transforms. They can be used to represent continuous-time (analog) signals. CWTs operate over every possible scale and translation whereas DWTs use a specific subset of scale and translation values or representation grid.

There are a large number of wavelet transforms each suitable for different applications.

Generalized transforms

There are a number of generalized transforms of which the wavelet transform is a special case. For example, Joseph Segman introduced scale into the Heisenberg group, giving rise to a continuous transform space that is a function of time, scale, and frequency. The CWT is a two-dimensional slice through the resulting 3d time-scale-frequency volume.

Another example of a generalized transform is the chirplet transform in which the CWT is also a two dimensional slice through the chirplet transform.

An important application area for generalized transforms involves systems in which high frequency resolution is crucial. For example, darkfield electron optical transforms intermediate between direct and reciprocal space have been widely used in the harmonic analysis of atom clustering, i.e. in the study of crystals and crystal defects. Now that transmission electron microscopes are capable of providing digital images with picometer-scale information on atomic periodicity in nanostructure of all sorts, the range of pattern recognition and strain/metrology applications for intermediate transforms with high frequency resolution (like brushlets and ridgelets) is growing rapidly.

Fractional wavelet transforms are based on scaling functions which, contrary to the "integer" wavelet transform, may have infinite support. Nonetheless, they allow non truncation of basis functions, exact treatment of boundaries, and perfect reconstruction.

List of wavelets

Discrete wavelets

- Beylkin (18)
- BNC wavelets
- Coiflet (6, 12, 18, 24, 30)
- Cohen-Daubechies-Feauveau wavelet (Sometimes referred to as CDF N/P or Daubechies biorthogonal wavelets)
- Daubechies wavelet (2, 4, 6, 8, 10, 12, 14, 16, 18, 20)
- Binomial-QMF (Also referred to as Daubechies wavelet)
- Haar wavelet
- Mathieu wavelet
- Legendre wavelet
- Villasenor wavelet
- Symlet

Continuous wavelets

Real-valued

- Beta wavelet
- Hermitian wavelet
- Hermitian hat wavelet
- Mexican hat wavelet
- Shannon wavelet

Complex-valued

- Complex mexican hat wavelet

- Morlet wavelet
- Shannon wavelet
- Modified Morlet wavelet

Chapter 13

Gabor–Wigner Transform & Wigner Distribution Function

Gabor–Wigner Transform

The **Gabor transform**, named after Dennis Gabor, and the **Wigner distribution** function, named after Eugene Wigner, are both tools for time-frequency analysis. Since the Gabor transform does not have high clarity, and the Wigner distribution function has a cross term problem, a 2007 study by S. C. Pei and J. J. Ding proposed a new combination of the two transforms that has high clarity and no cross term problem. Since the cross term does not appear in the Gabor transform, the time frequency distribution of the Gabor transform can be used as a filter to filter out the cross term in the output of the Wigner distribution function.

Mathematical definition

- **Gabor transform**

$$G_x(t,f) = \int_{-\infty}^{\infty} e^{-\pi(\tau-t)^2} e^{-j2\pi f\tau} x(\tau)\, d\tau$$

- **Wigner distribution function**

$$W_x(t,f) = \int_{-\infty}^{\infty} x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi\tau\, f}\, d\tau$$

- **Gabor–Wigner transform**

  There are many different combinations to define the Gabor–Wigner transform. Here four different definitions are given.

1. $D_x(t,f) = G_x(t,f) \times W_x(t,f)$
2. $D_x(t,f) = \min\left\{|G_x(t,f)|^2, |W_x(t,f)|\right\}$
3. $D_x(t,f) = W_x(t,f) \times \{|G_x(t,f)| > 0.25\}$
4. $D_x(t,f) = G_x^{2.6}(t,f)W_x^{0.7}(t,f)$

Performance of Gabor–Wigner transform

Here some examples are given to show the performance of four Gabor–Wigner transform comparing to Gabor transform and Wigner distribution function.

- $x(t) = \cos(8\pi t) + \cos(16\pi t)$
- $x(t) = e^{jt^3}$

The above examples illustrate that the Gabor–Wigner transform has less cross term and higher clarity than Gabor transform.

Wigner Distribution Function

The **Wigner distribution function (WDF)** was first proposed to account for quantum corrections to classical statistical mechanics in 1932 by Eugene Wigner, cf. Wigner quasi-probability distribution.

Given the shared algebraic structure between position-momentum and time-frequency pairs, it may also usefully serve as a transform in time-frequency analysis. Compared to a short-time Fourier transform, such as the Gabor transform, the Wigner distribution function can furnish higher clarity in some cases.

Mathematical definition

There are several different definitions for the Wigner distribution function. The definition given here is specific to time-frequency analysis. The Wigner distribution function $W_x(t,f)$ is

$$W_x(t,f) = \int_{-\infty}^{\infty} x(t+\tau/2)x^*(t-\tau/2)e^{-i2\pi\tau f}\,d\tau$$

where $i = \sqrt{-1}$ is the imaginary unit. The WDF is *essentially* the Fourier transform of the input signal's autocorrelation function — the Fourier spectrum of the product between the signal and its delayed, time reversed copy, as a function of the delay.

Time frequency analysis example

Here are some examples to illustrate how the WDF is used in time-frequency analysis.

Constant input signal

When the input signal is constant, its time-frequency distribution is a horizontal line on to the frequency axis. For example, if $x(t) = 1$, then

$$W_x(t,f) = \int_{-\infty}^{\infty} e^{-i2\pi\tau\, f}\, d\tau = \delta(f).$$

Sinusoidal input signal

When the input signal is a sinusoidal function, its time-frequency distribution is a horizontal line parallel to the frequency axis, at its sinusoidal frequency. For example, if $\boldsymbol{x(t) = e^{i2\pi ht}}$, then

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

Chirp input signal

When the input signal is a chirp function, the instantaneous frequency is a linear function. This means that the time frequency distribution should be a straight line. For example, if $x(t) = e^{i2\pi kt^2}$, then its instantaneous frequency is $\frac{1}{2\pi}\frac{d(2\pi kt^2)}{dt} = 2kt$, and by WDF

$$\begin{aligned}W_x(t,f) &= \int_{-\infty}^{\infty} e^{i2\pi k(t+\tau/2)^2} e^{-i2\pi k(t-\tau/2)^2} e^{-i2\pi\tau\, f}\, d\tau \\ &= \int_{-\infty}^{\infty} e^{i4\pi kt\tau} e^{-i2\pi\tau f}\, d\tau \\ &= \int_{-\infty}^{\infty} e^{-i2\pi\tau(f-2kt)}\, d\tau \\ &= \delta(f-2kt).\end{aligned}$$

Delta input signal

When the input signal is a delta function, since it is only non-zero at t=0 and contains infinite frequency components, its time-frequency distribution should be a vertical line across the origin. This means that the time frequency distribution of the delta function should also be a delta function. By WDF

$$\begin{aligned}W_x(t,f) &= \int_{-\infty}^{\infty} \delta(t+\tau/2)\delta(t-\tau/2) e^{-i2\pi\tau\, f}\, d\tau \\ &= 4\int_{-\infty}^{\infty} \delta(2t+\tau)\delta(2t-\tau) e^{-i2\pi\tau f}\, d\tau \\ &= 4\delta(4t)e^{i4\pi tf} \\ &= \delta(t)e^{i4\pi tf} \\ &= \delta(t).\end{aligned}$$

The Wigner distribution function is best suited for time-frequency analysis when the input signal's phase is 2nd order or lower. For those signals, WDF can exactly generate the time frequency distribution of the input signal.

Performance of Wigner distribution function

Here are some examples to show performance features of the Wigner distribution function preferable to the Gabor transform.

- $x(t) = \cos(2\pi t)$
- $x(t) = e^{i\pi t^2}$
- $x(t) = \begin{cases} 1 & |t| < 2 \\ 0 & \text{otherwise} \end{cases}$ rectangular function

Cross term property

The Wigner distribution function is not a linear transform. A cross term ("time beats") occurs when there is more than one component in the input signal, analogous in time to frequency beats. In the ancestral physics Wigner quasi-probability distribution, this term has important and useful physics consequences. The short-time Fourier transform does not have this feature. The following are some examples that show the cross term feature of the Wigner distribution function.

- $$x(t) = \begin{cases} \cos(2\pi t) & t \le -2 \\ \cos(4\pi t) & -2 < t \le 2 \\ \cos(3\pi t) & t > 2 \end{cases}$$

- $x(t) = e^{it^3}$

In order to reduce the cross term problem, many other transforms have been proposed, including the modified Wigner distribution function, the Gabor–Wigner transform, and Cohen’s class distribution.

Properties of the Wigner distribution function

The Wigner distribution function has several evident properties listed in the following table.

	Remarks	
1	Projection property	$\lvert x(t)\rvert^2 = \int_{-\infty}^{\infty} W_x(t,f)\,df \quad \lvert X(f)\rvert^2 = \int_{-\infty}^{\infty} W_x(t,f)\,dt$
2	Energy property	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_x(t,f)\,df dt = \int_{-\infty}^{\infty} \lvert x(t)\rvert^2\,dt = \int_{-\infty}^{\infty} \lvert X(f)\rvert^2\,df$
3	Recovery property	$\int_{-\infty}^{\infty} W_x(t/2,f)e^{i2\pi ft}\,df = x(t)x^*(0) \quad \int_{-\infty}^{\infty} W_x(t,f/2)e^{i2\pi ft}\,dt = X(f)X^*(0)$
4	Mean condition frequency and mean condition time	$X(f) = \lvert X(f)\rvert e^{i2\pi\psi(f)} \quad x(t) = \lvert x(t)\rvert e^{i2\pi\phi(t)}$ If $\phi'(t) = \lvert x(t)\rvert^{-2}\int_{-\infty}^{\infty} fW_x(t,f)\,df$ and $-\psi'(f) = \lvert X(f)\rvert^{-2}\int_{-\infty}^{\infty} tW_x(t,f)\,dt$
5	Moment properties	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} t^n W_x(t,f)\,dtdf = \int_{-\infty}^{\infty} t^n \lvert x(t)\rvert^2\,dt$ $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^n W_x(t,f)\,dtdf = \int_{-\infty}^{\infty} f^n \lvert X(f)\rvert^2\,df$
6	Real properties	$W_x^*(t,f) = W_x(t,f)$

7	Region properties	If $x(t) = 0\ for\ t > t_0$ then $W_x(t,f) = 0\ for\ t > t_0$ If $x(t) = 0\ for\ t < t_0$ then $W_x(t,f) = 0\ for\ t < t_0$
8	Multiplication theory	If $y(t) = x(t)h(t)$ then $W_y(t,f) = \int_{-\infty}^{\infty} W_x(t,\rho) W_h(t, f-\rho)\, d\rho$
9	Convolution theory	If $y(t) = \int_{-\infty}^{\infty} x(t-\tau)h(\tau)\, d\tau$ then $W_y(t,f) = \int_{-\infty}^{\infty} W_x(\rho, f) W_h(t-\rho, f)\, d\rho$
10	Correlation theory	If $y(t) = \int_{-\infty}^{\infty} x(t+\tau)h^*(\tau)\, d\tau$ then $W_y(t,\omega) = \int_{-\infty}^{\infty} W_x(\rho,\omega) W_h(-t+\rho,\omega)\, d\rho$
11	Time-shifting property	If $y(t) = x(t-t_0)$ then $W_y(t,f) = W_x(t-t_0, f)$
12	Modulation property	If $y(t) = e^{i2\pi f_0 t} x(t)$ then $W_y(t,f) = W_x(t, f-f_0)$

Chapter 14

Fourier Analysis

In mathematics, **Fourier analysis** is a subject area which grew from the study of Fourier series. The subject began with the study of the way general functions may be represented by sums of simpler trigonometric functions. Fourier analysis is named after Joseph Fourier, who showed that representing a function by a trigonometric series greatly simplifies the study of heat propagation.

Today, the subject of Fourier analysis encompasses a vast spectrum of mathematics. In the sciences and engineering, the process of decomposing a function into simpler pieces is often called Fourier analysis, while the operation of rebuilding the function from these pieces is known as **Fourier synthesis**. In mathematics, the term *Fourier analysis* often refers to the study of both operations.

The decomposition process itself is called a Fourier transform. The transform is often given a more specific name which depends upon the domain and other properties of the function being transformed. Moreover, the original concept of Fourier analysis has been extended over time to apply to more and more abstract and general situations, and the general field is often known as harmonic analysis. Each transform used for analysis has a corresponding inverse transform that can be used for synthesis.

Applications

Fourier analysis has many scientific applications — in physics, partial differential equations, number theory, combinatorics, signal processing, imaging, probability theory, statistics, option pricing, cryptography, numerical analysis, acoustics, oceanography, optics, diffraction, geometry, and other areas.

This wide applicability stems from many useful properties of the transforms:

- The transforms are linear operators and, with proper normalization, are unitary as well (a property known as Parseval's theorem or, more generally, as the Plancherel theorem, and most generally via Pontryagin duality)(Rudin 1990).
- The transforms are usually invertible.
- The exponential functions are eigenfunctions of differentiation, which means that this representation transforms linear differential equations with constant coefficients into ordinary algebraic ones (Evans 1998). Therefore, the behavior of a linear time-invariant system can be analyzed at each frequency independently.
- By the convolution theorem, Fourier transforms turn the complicated convolution operation into simple multiplication, which means that they provide an efficient way to compute convolution-based operations such as polynomial multiplication and multiplying large numbers (Knuth 1997).
- The discrete version of the Fourier transform (see below) can be evaluated quickly on computers using fast Fourier transform (FFT) algorithms. (Conte & de Boor 1980)

Fourier transformation is also useful as a compact representation of a signal. For example, JPEG compression uses a variant of the Fourier transformation (discrete cosine transform) of small square pieces of a digital image. The Fourier components of each square are rounded to lower arithmetic precision, and weak components are eliminated entirely, so that the remaining components can be stored very compactly. In image reconstruction, each Fourier-transformed image square is reassembled from the preserved approximate components, and then inverse-transformed to produce an approximation of the original image.

Applications in signal processing

When processing signals, such as audio, radio waves, light waves, seismic waves, and even images, Fourier analysis can isolate individual components of a compound waveform, concentrating them for easier detection and/or removal. A large family of signal processing techniques consist of Fourier-transforming a signal, manipulating the Fourier-transformed data in a simple way, and reversing the transformation.

Some examples include:

- Telephone dialing; the touch-tone signals for each telephone key, when pressed, are each a sum of two separate tones (frequencies). Fourier analysis can be used to separate (or *analyze*) the telephone signal, to reveal the two component tones and therefore which button was pressed.
- Removal of unwanted frequencies from an audio recording (used to eliminate hum from leakage of AC power into the signal, to eliminate the stereo subcarrier from FM radio recordings);
- Noise gating of audio recordings to remove quiet background noise by eliminating Fourier components that do not exceed a preset amplitude;
- Equalization of audio recordings with a series of bandpass filters;

- Digital radio reception with no superheterodyne circuit, as in a modern cell phone or radio scanner;
- Image processing to remove periodic or anisotropic artifacts such as jaggies from interlaced video, stripe artifacts from strip aerial photography, or wave patterns from radio frequency interference in a digital camera;
- Cross correlation of similar images for co-alignment;
- X-ray crystallography to reconstruct a crystal structure from its diffraction pattern;
- Fourier transform ion cyclotron resonance mass spectrometry to determine the mass of ions from the frequency of cyclotron motion in a magnetic field.
- Many other forms of spectroscopy also rely upon Fourier Transforms to determine the three-dimensional structure and/or identity of the sample being analyzed, including Infrared and Nuclear Magnetic Resonance spectroscopies.
- Generation of sound spectrograms used to analyze sounds.

Variants of Fourier analysis

Fourier analysis has different forms, some of which have different names. The more common variants are shown below. The different names usually reflect different properties of the function or data being analyzed. The resultant transforms can be seen as special cases or generalizations of each other.

(Continuous) Fourier transform

Most often, the unqualified term **Fourier transform** refers to the transform of functions of a continuous real argument, such as time (t). In this case the Fourier transform describes a function $f(t)$ in terms of basic complex exponentials of various frequencies. In terms of ordinary frequency, ν, the Fourier transform is given by the complex number**:**

$$F(\nu) = \int_{-\infty}^{\infty} f(t) \cdot e^{-i2\pi\nu t} dt.$$

Evaluating this quantity for all values of ν produces the *frequency-domain* function.

Fourier series

A Fourier series is a representation of a function in terms of a summation of a potentially infinite number of harmonically-related sinusoids or complex exponential functions with different amplitudes and phases. The amplitude and phase of a sinusoid can be combined into a single complex number, called a Fourier *coefficient*. The Fourier series is a periodic function. So it cannot represent any arbitrary function. It can represent either**:**

(a) a periodic function, or
(b) a function that is defined only over a finite-length interval (or compact support). Then the values produced by the Fourier series outside the finite interval are irrelevant.

The general form of a Fourier series is:

$$s(t) = \sum_{n=-\infty}^{\infty} S[n] \cdot e^{i2\pi \frac{n}{\tau} t},$$

where τ is the period (case a) or the interval length (case b), and the $S[n]$ sequence are the Fourier coefficients. When the coefficients are derived from a function, $f(t)$ as follows:

$$S[n] = \frac{1}{\tau} \int_{u}^{u+\tau} f(t) \cdot e^{-i2\pi \frac{n}{\tau} t}\, dt,$$

then, aside from possible convergence issues, $s(t)$ will equal $f(t)$ in the interval $[u,u+\tau]$. It follows that if $f(t)$ is τ-periodic (case a), $s(t)$ and $f(t)$ are equal everywhere.

The Fourier series is analogous to the inverse Fourier transform, in that it is the reconstruction of the original function that was transformed into the Fourier series coefficients.

Discrete-time Fourier transform (DTFT)

For functions of an integer index, the **discrete-time Fourier transform** (**DTFT**) provides a useful frequency-domain transform.

A useful "discrete-time" function can be obtained by sampling a "continuous-time" function, $s(t)$, which produces a sequence, $s(nT)$, for integer values of n and some time-interval T. If information is lost, then only an approximation to the original transform, $S(f)$, can be obtained by looking at one period of the periodic function:

$$S_T(f) = \sum_{k=-\infty}^{\infty} S\left(f - \frac{k}{T}\right) \equiv \sum_{n=-\infty}^{\infty} \underbrace{T \cdot s(nT)}_{s[n]} \cdot e^{-i2\pi f n T},$$

which is the DTFT. The identity above is a result of the Poisson summation formula. The DTFT is also equivalent to the Fourier transform of a "continuous" function that is constructed by using the $s[n]$ sequence to modulate a Dirac comb.

Applications of the DTFT are not limited to sampled functions. It can be applied to any discrete sequence.

Discrete Fourier transform (DFT)

When $s[n]$ is periodic, with period **N**, $S_T(f)$ is another Dirac comb function, modulated by the coefficients of a **Fourier series**. And the integral formula for the coefficients simplifies to:

$$S[k] = \sum_{n=0}^{N-1} s[n] \cdot e^{-i2\pi \frac{k}{N} n}$$ for all integer values of k.

This sequence is N-periodic, and so the entire sequence can be described by just N coefficients, known most often as the **DFT** and sometimes as the discrete Fourier series (**DFS**). The DFT also has an inverse transform that reproduces the periodic $s[n]$ sequence.

When $s[n]$ is not periodic, but its non-zero portion has finite duration (**N**), $S_T(f)$ is continuous and finite-valued. But a discrete subset of its values is sufficient to reconstruct/represent the (finite) portion of $s[n]$ that was analyzed, analogous to case b (above) of the Fourier series. And that subset is again the DFT.

- When **N** is larger than the non-zero portion of $s[n]$, known as zero-padding, the DFT computes more closely-spaced samples of one period of $S_T(f)$. That is frequently done to provide an interpolated view of the DTFT.
- The term DFT is ambiguous in the sense that it does not tell us whether the inverse transform is valid for all n or just for a sequence of length **N**; i.e. whether the original $s[n]$ sequence is periodic or finite. The term DFS, mentioned above, is sometimes used instead of DFT to convey that $s[n]$ is periodic.

The DFT can be computed using a fast Fourier transform (FFT) algorithm, which makes it a practical and important transformation on computers.

Fourier transforms on arbitrary locally compact abelian topological groups

The Fourier variants can also be generalized to Fourier transforms on arbitrary locally compact abelian topological groups, which are studied in harmonic analysis; there, the Fourier transform takes functions on a group to functions on the dual group. This treatment also allows a general formulation of the convolution theorem, which relates Fourier transforms and convolutions.

Time–frequency transforms

In signal processing terms, a function (of time) is a representation of a signal with perfect *time resolution,* but no frequency information, while the Fourier transform has perfect *frequency resolution,* but no time information.

As alternatives to the Fourier transform, in time–frequency analysis, one uses time–frequency transforms to represent signals in a form that has some time information and some frequency information – by the uncertainty principle, there is a trade-off between these. These can be generalizations of the Fourier transform, such as the short-time Fourier transform, the Gabor transform or fractional Fourier transform, or can use different functions to represent signals, as in wavelet transforms and chirplet transforms, with the wavelet analog of the (continuous) Fourier transform being the continuous wavelet transform.

History

A primitive form of harmonic series dates back to ancient Babylonian mathematics, where they were used to compute ephemerides (tables of astronomical positions).

In modern times, variants of the discrete Fourier transform were used by Alexis Clairaut in 1754 to compute an orbit, which has been described as the first formula for the DFT, and in 1759 by Joseph Louis Lagrange, in computing the coefficients of a trigonometric series for a vibrating string. Technically, Clairaut's work was a cosine-only series (a form of discrete cosine transform), while Lagrange's work was a sine-only series (a form of discrete sine transform); a true cosine+sine DFT was used by Gauss in 1805 for trigonometric interpolation of asteroid orbits. Euler and Lagrange both discretized the vibrating string problem, using what would today be called samples.

An early modern development toward Fourier analysis was the 1770 paper *Réflexions sur la résolution algébrique des équations* by Lagrange, which in the method of Lagrange resolvents used a complex Fourier decomposition to study the solution of a cubic: Lagrange transformed the roots x_1, x_2, x_3 into the resolvents:

$$\begin{aligned} r_1 &= x_1 + x_2 + x_3 \\ r_2 &= x_1 + \zeta x_2 + \zeta^2 x_3 \\ r_3 &= x_1 + \zeta^2 x_2 + \zeta x_3 \end{aligned}$$

where ζ is a cubic root of unity, which is the DFT of order 3.

A number of authors, notably Jean le Rond d'Alembert, , and Carl Friedrich Gauss used trigonometric series to study the heat equation, but the breakthrough development was the 1807 paper *Mémoire sur la propagation de la chaleur dans les corps solides* by Joseph Fourier, whose crucial insight was to model *all* functions by trigonometric series, introducing the Fourier series.

Historians are divided as to how much to credit Lagrange and others for the development of Fourier theory: Daniel Bernoulli and Leonhard Euler had introduced trigonometric representations of functions, and Lagrange had given the Fourier series solution to the wave equation, so Fourier's contribution was mainly the bold claim that an arbitrary function could be represented by a Fourier series.

The subsequent development of the field is known as harmonic analysis, and is also an early instance of representation theory.

The first fast Fourier transform (FFT) algorithm for the DFT was discovered around 1805 by Carl Friedrich Gauss when interpolating measurements of the orbit of the asteroids

Juno and Pallas, although that particular FFT algorithm is more often attributed to its modern rediscoverers Cooley and Tukey.

Interpretation in terms of time and frequency

In signal processing, the Fourier transform often takes a time series or a function of continuous time, and maps it into a frequency spectrum. That is, it takes a function from the time domain into the frequency domain; it is a decomposition of a function into sinusoids of different frequencies; in the case of a Fourier series or discrete Fourier transform, the sinusoids are harmonics of the fundamental frequency of the function being analyzed.

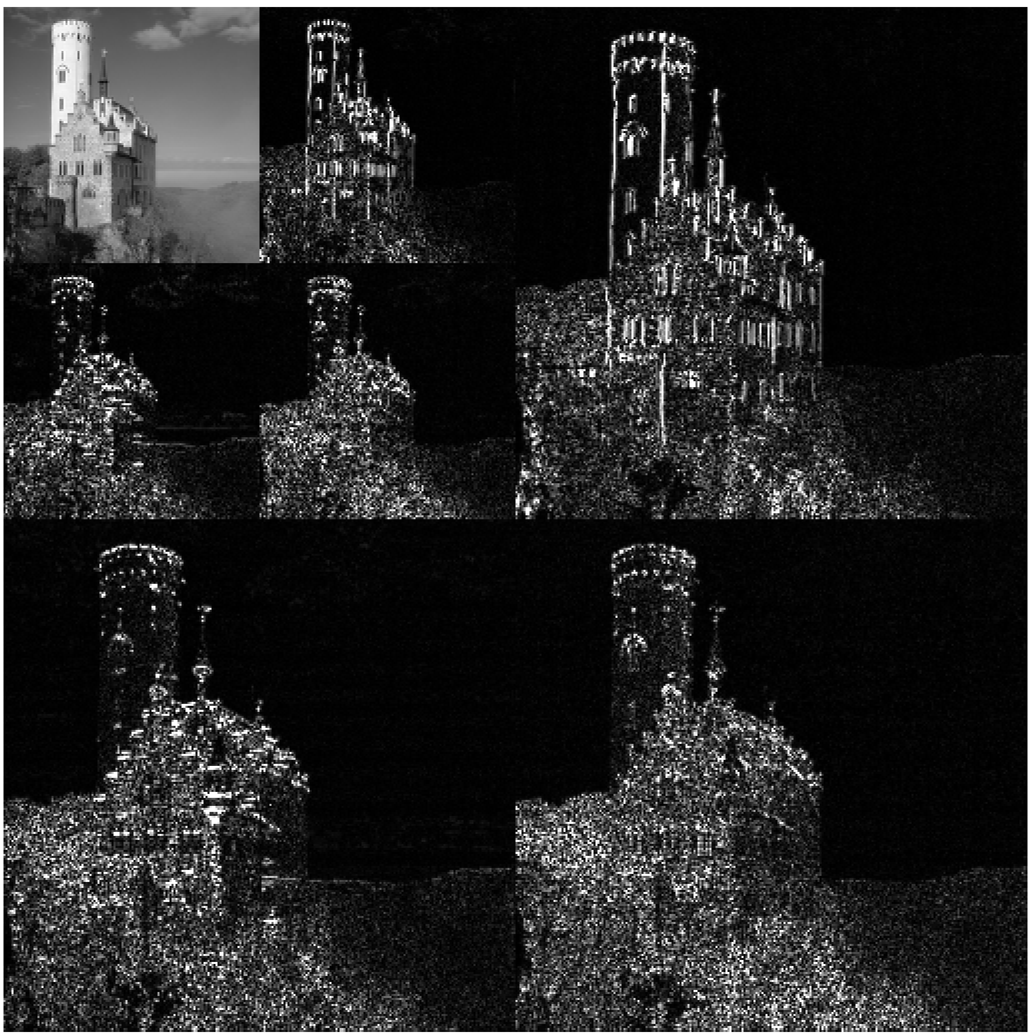
When the function f is a function of time and represents a physical signal, the transform has a standard interpretation as the frequency spectrum of the signal. The magnitude of the resulting complex-valued function F at frequency ω represents the amplitude of a frequency component whose initial phase is given by the phase of F.

Fourier transforms are not limited to functions of time, and temporal frequencies. They can equally be applied to analyze *spatial* frequencies, and indeed for nearly any function domain. This justifies their use in branches such diverse as image processing, heat conduction and automatic control.

Chapter 15

Wavelet Transform & Gabor Transform

Wavelet Transform



An example of the 2D discrete wavelet transform that is used in JPEG2000.

In mathematics, a **wavelet series** is a representation of a square-integrable (real- or complex-valued) function by a certain orthonormal series generated by a wavelet.

Formal definition

A function $\psi \in L^2(\mathbb{R})$ is called an **orthonormal wavelet** if it can be used to define a Hilbert basis, that is a complete orthonormal system, for the Hilbert space $L^2(\mathbb{R})$ of square integrable functions. The Hilbert basis is constructed as the family of functions $\{\psi_{jk} : j, k \in \mathbb{Z}\}$ by means of dyadic translations and dilations of ψ,

$$\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$$

for integers $j, k \in \mathbb{Z}$. This family is an orthonormal system if it is orthonormal under the inner product

$$\langle \psi_{jk}, \psi_{lm} \rangle = \delta_{jl}\delta_{km}$$

where δ_{jl} is the Kronecker delta and $\langle f, g \rangle$ is the standard inner product $\langle f, g \rangle = \int_{-\infty}^{\infty} \overline{f(x)} g(x) dx$ on $L^2(\mathbb{R})$. The requirement of completeness is that every function $f \in L^2(\mathbb{R})$ may be expanded in the basis as

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{jk}\psi_{jk}(x)$$

with convergence of the series understood to be convergence in norm. Such a representation of a function *f* is known as a **wavelet series**. This implies that an orthonormal wavelet is self-dual.

Wavelet transform

The **integral wavelet transform** is the integral transform defined as

$$\left[W_\psi f\right](a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x-b}{a}\right)} f(x) dx$$

The **wavelet coefficients** c_{jk} are then given by

$$c_{jk} = \left[W_\psi f\right](2^{-j}, k2^{-j})$$

Here, $a = 2^{-j}$ is called the **binary dilation** or **dyadic dilation**, and $b = k2^{-j}$ is the **binary** or **dyadic position**.

Wavelet compression

Wavelet compression is a form of data compression well suited for image compression (sometimes also video compression and audio compression). Notable implementations are JPEG 2000 for still images, and REDCODE, the BBC's Dirac, and Ogg Tarkin for video. The goal is to store image data in as little space as possible in a file. Wavelet compression can be either lossless or lossy.

Using a **wavelet transform**, the wavelet compression methods are adequate for representing transients, such as percussion sounds in audio, or high-frequency components in two-dimensional images, for example an image of stars on a night sky. This means that the transient elements of a data signal can be represented by a smaller amount of information than would be the case if some other transform, such as the more widespread discrete cosine transform, had been used.

Wavelet compression is not good for all kinds of data: transient signal characteristics mean good wavelet compression, while smooth, periodic signals are better compressed by other methods, particularly traditional harmonic compression (frequency domain, as by Fourier transforms and related). Data statistically indistinguishable from random noise is not compressible by any means.

Method

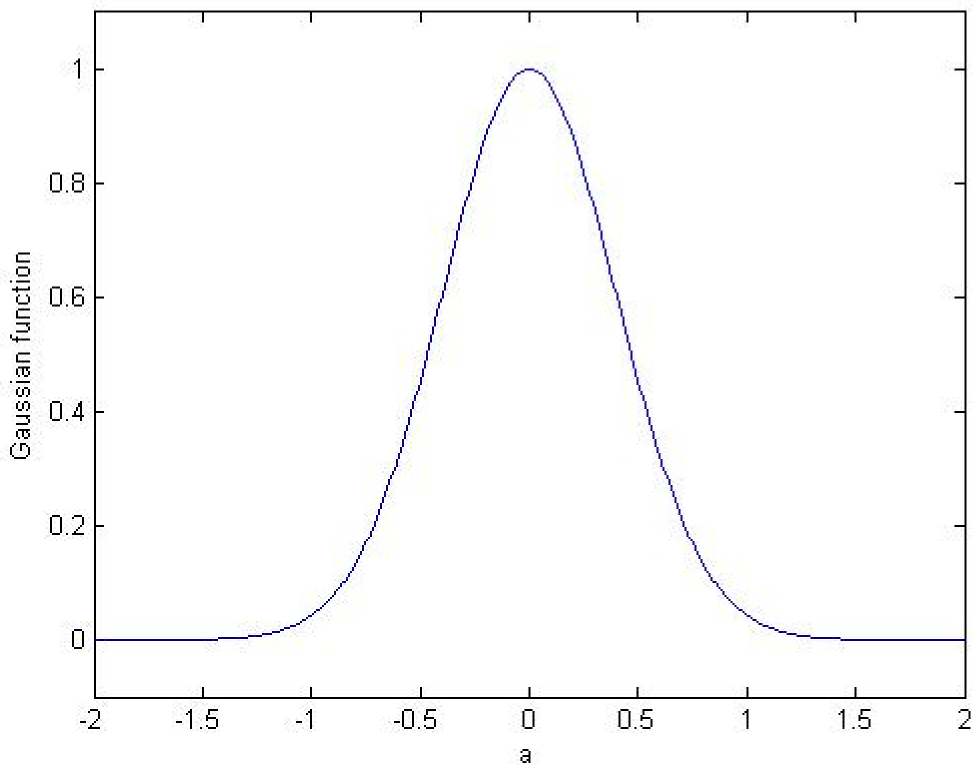
First a wavelet transform is applied. This produces as many coefficients as there are pixels in the image (i.e.: there is no compression yet since it is only a transform). These coefficients can then be compressed more easily because the information is statistically concentrated in just a few coefficients. This principle is called transform coding. After that, the coefficients are quantized and the quantized values are entropy encoded and/or run length encoded.

A few 1D and 2D applications of wavelet compression use a technique called "wavelet footprints".

Gabor Transform

The **Gabor transform**, named after Dennis Gabor, is a special case of the short-time Fourier transform. It is used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. The function to be transformed is first multiplied by a Gaussian function, which can be regarded as a window, and the resulting function is then transformed with a Fourier transform to derive the time-frequency analysis. The window function means that the signal near the time being analyzed will have higher weight. The Gabor transform of a signal x(t) is defined by this formula:

$$G_x(t,f) = \int_{-\infty}^{\infty} e^{-\pi(\tau-t)^2} e^{-j2\pi f\tau} x(\tau)\, d\tau$$



Magnitude of Gaussian function.

The Gaussian function has infinite range and it is impractical for implementation. But take a look at the distribution of Gaussian function.

$$\begin{cases} e^{-\pi a^2} \geq 0.00001; & |a| \leq 1.9143 \\ e^{-\pi a^2} < 0.00001; & |a| > 1.9143 \end{cases}$$

Gaussian function with $|a| > 1.9143$ can be regarded as 0 and also can be ignored. Thus the Gabor transform can be simplified as

$$G_x(t,f) = \int_{-1.9143}^{1.9143} e^{-\pi(\tau-t)^2} e^{-j2\pi f\tau} x(\tau)\, d\tau$$

This simplification makes the Gabor transform practical and realizable.

Inverse Gabor transform

The Gabor transform is invertible. The original signal can be recovered by the following equation

$$x(t) = \int_{-\infty}^{\infty} G_x(t,f) e^{j2\pi tf} \, df$$

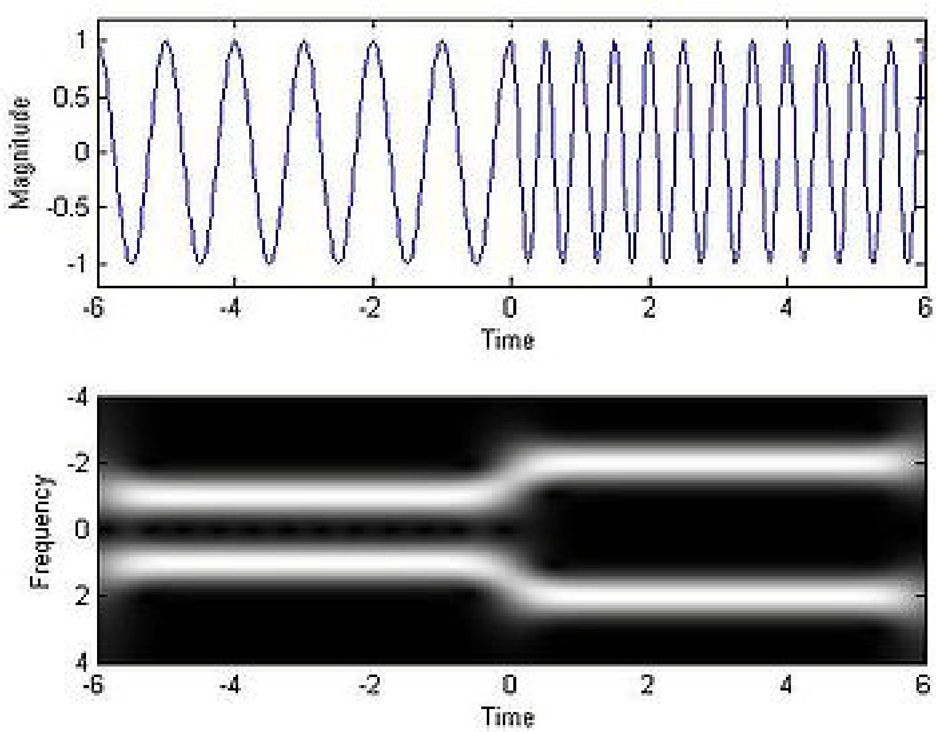
Properties of the Gabor transform

The Gabor transform has many properties like those of the Fourier transform. These properties are listed in the following tables.

	Signal	Gabor transform	Remarks
	$x(t)$	$G_x(t,f) = \int_{-\infty}^{\infty} e^{-\pi(\tau-t)^2} e^{-j2\pi f\tau} x(\tau)\, d\tau$	
1	$a \cdot x(t) + b \cdot y(t)$	$a \cdot G_x(t,f) + b \cdot G_y(t,f)$	Linearity property
2	$x(t-t_0)$	$G_x(t-t_0,f)e^{-j2\pi f t_0}$	Shifting property
3	$x(t)e^{j2\pi f_0 t}$	$G_x(t,f-f_0)$	Modulation property

		Remarks
1	$\int_{-\infty}^{\infty} \lvert G_x(t,f)\rvert^2 \, df = \int_{-\infty}^{\infty} e^{-2\pi(\tau-t)^2} \lvert x(\tau)\rvert^2 \, d\tau \approx \int_{u-1.9143}^{u+1.9143} e^{-2\pi(\tau-u)^2} \lvert x(\tau)\rvert^2 \, d\tau$	Power integration property
2	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_x(t,f) G_y^*(t,f) \, df \, dt = \int_{-\infty}^{\infty} x(\tau) y^*(\tau) \, d\tau$	Energy sum property
3	$\begin{cases} \int_{-\infty}^{\infty} \lvert G_x(t,f)\rvert^2 \, df < e^{-2\pi(t-t_0)^2} \int_{-\infty}^{\infty} \lvert G_x(t_0,f)\rvert^2 \, df; & \text{if} \\ \int_{-\infty}^{\infty} \lvert G_x(t,f)\rvert^2 \, dt < e^{-2\pi(f-f_0)^2} \int_{-\infty}^{\infty} \lvert G_x(t,f_0)\rvert^2 \, dt; & \text{if} \end{cases}$	Power decay property
4	$\int_{-\infty}^{\infty} G_x(t,f) e^{j2\pi ktf} \, df = e^{-\pi(k-1)^2 t^2} x(kt)$	Integration property
5	$\int_{-\infty}^{\infty} G_x(t,f) e^{j2\pi tf} \, df = x(t)$	Recovery property

Application and example



Time/frequency distribution.

The main application of the Gabor transform is used in time frequency analysis. Take the following equation as an example. The input signal has 1Hz frequency component when $t \le 0$ and has 2Hz frequency component when $t > 0$

$$x(t) = \begin{cases} \cos(2\pi t) & \text{for } t \le 0, \\ \cos(4\pi t) & \text{for } t > 0. \end{cases}$$

But if the total bandwidth available is 5Hz, other frequency bands except $x(t)$ are wasted. Through time frequency analysis by applying the Gabor transform, the available bandwidth can be known and those frequency bands can be used for other applications and bandwidth is saved. The right side picture show the input signal $x(t)$ and the output of the Gabor transform. As our expectation, the frequency distribution can be separate as two parts. One is $t \le 0$ and the other is $t > 0$. The white part is the frequency band occupied by $x(t)$ and the black part is not used.