DEFENCE DÉFENSE

Linear and Quadratic Time-Frequency Representations

T. Thayaparan
Defence Research Establishment Ottawa

Defence R&D Canada

DEFENCE RESEARCH ESTABLISHMENT OTTAWA

TECHNICAL MEMORANDUM
DREO TM 2000-080
November 2000

National Defence Défense nationale

Canada

Abstract

This report is reviewing both linear and quadratic time-frequency representations. The linear representations discussed are Short-Time Fourier Transform and S-transform. The quadratic representation discussed is the Wigner distribution. We outline the motivations, interpretations, mathematical fundamentals, properties, and applications of these linear and quadratic time-frequency representations. We also compare these three different time-frequency analysis techniques and show that each technique has its strengths and drawbacks. Simulated data sets have been used for the comparison. The choice of the particular time-frequency representation depends upon the specific area of application and what we aim to achieve with a local frequency analysis. We show that time-frequency analysis methods should enable us to classify signals with a considerably greater reflection of the physical situation than can be achieved by the conventional Fourier Transform method alone.

Résumé

Ce rapport examine les représentations linéaire et quadratique de temps-fréquence. Les représentations linéaires qu'on examine sont la transformée ' et la transformée de Fourier à temps réduite. La représentation quadratique qu'on examine est la distribution de Wigner. Nous donnons les grandes lignes au sujet des motivations, interprétations, mathématique fondamentale, propriétés, et les applications de ces représentations linéaires et quadratiques de temps-fréquence. Nous comparons aussi ces trois techniques différentes d'analyses de temps-fréquence et montrons que chaque techniques a ses forces et ses problèmes. Des ensembles de données simulées ont été utilisés pour des comparaisons. Le choix de la représentation en temps-fréquence dépend du domaine d'application spécifique et ce que nous pouvons accomplir avec une analyse de fréquence locale. Nous montrons que les méthodes d'analyse temps-fréquence devraient nous permettre de classifier des signaux avec une interprétation plus grande de la situation physique que peut être accompli par la méthode conventionnelle seule.

Executive Summary

One of the central problems in High Frequency (HF) radar data is the analysis of a time series. The problem at hand is how to extract the information present in the data and use it to its full potential. In order to address this problem we turn to the field of signal analysis and data representations. The Fourier Transform is at the heart of a wide range of techniques that are used in HF radar data analysis and processing. Mapping the data into the temporal frequency domain is an effective way of recording the data such that their global characteristics can be assessed. However, the change of frequency content with time is one of the main features we observe in HF radar data. Because of this change of frequency content with time, radar signals belong to the class of non-stationary signals. The analysis of non-stationary signals requires technique that extend the notion of a global frequency spectrum to a local frequency description. The spectral energy density function that is obtained by means of a Fourier Transform, the so-called power spectrum, shows the frequencies that are present in our data, but does not reveal where changes in the frequency content occur. Consequently, for the interpretation of radar data in terms of a changing frequency content, we need a representation of our data as a function of both time and frequency. Only, quite recently, the joint time-frequency representation of signals has become a major area of research in signal processing.

The time-frequency representations characterize signals over a time-frequency plane. They thus combine time-domain and frequency-domain analyses to yield a potentially more revealing picture of the temporal localization of a signal's spectral components. They may also serve as a basis for signal detection, characterization, coding, and processing. A complete and comprehensive theory for joint time-frequency analysis does not yet exist. There is no unique time-frequency representation of a signal that satisfies all the properties of a physically correct joint time-frequency energy density function. However, discarding the requirement that all properties must be satisfied in one time-frequency representation, a class of joint time-frequency representations can be derived that serves as a model of a local power spectrum.

In order to employ the concept of a local power spectrum in HF radar signal analysis we need to decide which time-frequency representation is to be used. This choice cannot be made on the basis of a mathematical analysis alone. We should also take into account what we aim to achieve with a local frequency analysis. Our main area of application of the

time-frequency representation is the analysis of experimental HF radar data. A study of the properties of different time-frequency representations should provide us a guideline as to which representation is to be chosen in these applications.

This report is reviewing both linear and quadratic time-frequency representations. The linear representations discussed are Short-Time Fourier Transform and S-transform. The quadratic representation discussed is the Wigner distribution. We outline the motivations, interpretations, mathematical fundamentals, and properties of these linear and quadratic time-frequency representations. We also compare these three different time-frequency analysis techniques and show that each technique has its strengths and drawbacks. Simulated data sets have been used for the comparison. The comparison allows us to determine what would be the most effective methods under different conditions. The choice of the particular time-frequency representation depends upon the specific area of application and what we aim to achieve with a local frequency analysis. We show that time-frequency analysis methods should enable us to classify signals with a considerably greater reflection of the physical situation than can be achieved by the conventional Fourier Transform method alone.

Author, Thayananthan, Thayaparan, Linear and quadratic time-frequency representations, Defence Research Establishment Ottawa, DREO TM 2000-080, November 2000.

Sommaire

Un des problèmes centraux dans les données des radars à haute fréquence est l'analyses des séries temporelles. Le problème consiste à savoir comment extraire l'information présente dans les données et de l'utiliser à son plein potentiel. En vue d'adressé ce problème nous nous tournons vers le domaine des représentations et d'analyses des signaux. La transformée de Fourier est au cœur d'un large éventail de techniques qui sont utilisées dans l'analyse et le traitement des données des radars à hautes fréquences. Le mappage des données dans le domaine temporel de fréquence est une façon effective d'enregistrer les données tel que leurs caractéristiques globales peuvent être évaluer. Cependant, le changement des composantes spectrales avec le temps est une des caractéristiques que nous observons dans les données radar à haute fréquence. Ces signaux radars appartiennent à la classe des signaux non stationnaires à cause de ce changement de composantes spectrales avec le temps, L'analyses des signaux non stationnaires exige des techniques qui étende la notion de spectre de fréquence globale à une description locale. La courbe de densité spectrale d'énergie qui est obtenu au moyen de la transformée de Fourier, qui est aussi appelée le spectre de puissance, montre les fréquences qui sont présente dans nos données, mais ne montre pas où les changements de fréquences se produisent. En conséquence, pour l'interprétation des données radars en terme de changement de composantes spectrales, nous avons besoin d'une représentation de nos données en fonction du temps ainsi que des fréquences. La représentation temps-fréquence est récemment devenue un des domaines majeurs de recherche dans les traitements des signaux.

Les représentations temps-fréquence caractérisent les signaux dans un plan de fréquence et de temps. Ainsi ils combinent les analyses du domaine des fréquences et du temps produisant ainsi une image beaucoup plus claire de la localisation temporelle des composantes spectrales des signaux. Ils peuvent aussi servir comme une base pour la détection, la caractérisation, le codage, et le traitement. Il n'y a pas de représentation unique de signaux qui satisfassent toutes les propriétés représentation correcte d'une courbe de densité spectrale d'énergie temps-fréquence. Cependant, en mettant de cotée le fait que toute les propriétés doivent être satisfaites dans la représentation temps-fréquence, une classe de représentation temps-fréquence peut être dériver comme un modèle de spectre de puissance locale.

Avant d'employer le concept de puissance spectral locale dans l'analyse des signaux radars à haute fréquence, nous avons besoin de décider quelle représentation temps-fréquence sera

utiliser. Ce choix ne peut être fait seulement sur une base mathématique. Nous devons aussi tenir compte de ce que nous voulons accomplir avec l'analyse de fréquence locale. Notre application principale des représentations temps-fréquence est l'analyse des données expérimentales des radars à haute fréquence. Une étude des propriétés des différentes représentations temps-fréquence devrait nous fournir un guide a savoir quelle représentation doit être choisie pour cette application.

Ce rapport examine les représentations linéaire et quadratique de temps-fréquence. Les représentations linéaires qu'on examine sont la transformée S et la transformée de Fourier à temps réduite. La représentation quadratique qu'on examine est la distribution de Wigner. Nous donnons les grandes lignes au sujet des motivations, interprétations, mathématique fondamentale, propriétés, et les applications de ces représentations linéaires et quadratiques de temps-fréquence. Nous comparons aussi ces trois techniques différentes d'analyses de temps-fréquence et montrons que chaque technique a ses forces et ses problèmes. Des ensembles de données simulées ont été utilisés pour des comparaisons. Le choix de la représentation en temps-fréquence dépend du domaine d'application spécifique et ce que nous pouvons accomplir avec une analyse de fréquence locale. Nous montrons que les méthodes d'analyse temps-fréquence devraient nous permettre de classifier des signaux avec une interprétation plus grande de la situation physique que peut être accompli par la méthode conventionnelle seule.

Auther, Thayananthan, Thayaparan, Représentation linaire et quadratique en temps-fréquence, Center de Recherches pour la Défense, Ottawa, DREO TM 2000-080, Novembre 2000.

Contents

1 Introduction 1

2 Description of Signals in the Time and Frequency Domains 5

2.1 Introduction 5

2.2 Description of Signals in the Time Domain 6

2.2.1 Instantaneous Power 7

2.2.2 Total Energy 7

2.2.3 Time Averages, Mean Time, and Duration 8

2.3 Description of Signals in the Frequency Domain 9

2.3.1 Fourier Transform 9

2.3.2 Discrete Fourier Transform 10

2.3.3 Spectral Amplitude and Phase 11

2.3.4 Energy Spectrum Density 11

2.3.5 Mean Frequency, Bandwidth, and Frequency Averages 12

2.4 The Frequency and the Time Operators 13

2.4.1 The Frequency Operator 13

2.4.2 The Time Operator 16

2.4.3 Mean Frequency and Instantaneous Frequency 17

2.4.4 Mean Time and Group Delay 18

2.4.5 The Translation Operator 19

2.5 The Covariance of a Signal 20

2.5.1 The Covariance of a Spectrum 20

2.5.2 When is the Covariance Equal to Zero ? 21

2.6 Power Spectrum 22

2.7 Nonadditivity of Spectral Properties 23

3 Instantaneous Frequency and the Analytical Signal **24**

3.1 Introduction 24

3.2 The Analytical Signal 25

3.2.1 Energy of the Analytic Signal 27

3.2.2 Paradoxes Regarding the Analytic Signal 27

4 The Uncertainty Principle **30**

4.1 Introduction 30

4.1.1 A More General Uncertainty Principle 31

4.2 Proof of the Uncertainty Principle 32

4.3 The Uncertainty Principle for the Short-Time Fourier Transform 34

4.3.1 Time-Dependent and Window-Dependent Uncertainty Principle . . . 36

5 Time-Frequency Distributions: Fundamental Ideas **37**

5.1 introduction 37

5.2 Marginals 37

5.3 Total Energy 38

5.4 Characteristic Functions 38

5.4.1 One-Dimensional 38

5.4.2 Two-Dimensional 40

5.5 Global Averages 41

5.6 Time and Frequency Shift Invariance 41

5.7 Weak and Strong Finite Support 42

5.7.1 Distributions Concentrated in a Finite Region 43

5.8 Uncertainty Principle 43

6 The Short-Time Fourier Transform **45**

6.1 Introduction 45

6.2 The Short-Time Fourier Transform and Spectrogram 46

6.2.1 Characteristic Function 47

6.3 General Properties 48

6.3.1 Total Energy 48

6.3.2 Marginals 48

6.3.3 Averages of Time and Frequency Functions 49

6.3.4 Finite Support 50

6.3.5 Localization Trade-Off 50

6.3.6 One Window or Many ? 51

6.3.7 Entanglement and Symmetry Between Window and Signal 51

7 The Wigner Distribution **52**

7.1 Introduction 52

7.1.1 Time-Dependent Power Spectrum 52

7.2 The Wigner Distribution 54

7.2.1 Range of the Wigner Distribution 55

7.2.2 The Characteristic Function of the Wigner Distribution 56

7.3 General Properties of Wigner distribution 61

7.3.1 Time and Frequency Shift Invariance 61

7.3.2 Reality 62

7.3.3 Symmetry 62

7.3.4 Time and Frequency Marginals 63

7.3.5 Instantaneous Frequency and Group Delay 64

7.4 The Wigner Distribution of the Sum of Multiple Signals 64

7.5 Pseudo Wigner Distribution 67

7.6 Discrete Wigner distribution 68

7.7 Comparison of the Wigner Distribution with the Spectrogram 69

8 The S-Transform **71**

8.1 Introduction 71

8.2 Derivation of the S-Transform from the Short Time Fourier Transform . . . 71

8.3 Properties of the S-Transform 74

8.3.1 Inverse of the S-Transform and the Fourier Transform 74

8.3.2 Progressive Resolution of the Time-Frequency Domain 75

8.3.3 Frequency Resolution Depends on the Signal 76

8.3.4 The S Transform and Generalized Instantaneous Frequency 76

8.3.5 Linearity 80

8.3.6 The Generalized S-Transform 80

8.4 The Discrete S-Transform 81

8.5 Discrete Inverse S-Transform 83

8.6 Comparison of Time-Frequency Representations 84

9 Conclusion **90**

List of Figures

1 The instantaneous frequency for the signal $s(t) = A_1 e^{j10t} + A_2 e^{j20t}$. The spectrum consists of two frequencies, at $\omega = 10$ and $\omega = 20$. In (a) $A_1 = 0.2$ and $A_2 = 1$. The instantaneous frequency is continuous and ranges outside the bandwidth. In (b) $A_1 = -1.2$ and $A_2 = 1$. Although the signal is analytic the instantaneous frequency may become negative [after Cohen, 1995]. 29

2 The uncertainty relation for time and frequency marginals. A broad waveform in time gives a narrow frequency spectrum and vice versa. 31

3 A sample time series with a non-zero region. (b) The Wigner distribution of the time series in (a). The Wigner distribution has finite support. It is zero in the region for $t < t_1$ where t_1 is the start of the signal, and in the region for $t > t_2$ where t_2 is the end of the signal. 57

4 A sample time series with zero-filled region. (b) The Wigner distribution of the time series in (a). The Wigner distribution has compact support, but it is not identically equal to zero when the signal is zero. In fact, the zero region of the signal corresponds to the maximum peak in the Wigner distribution. . 58

5 A sample time series with three distinct regions. (b) The Wigner distribution of the time series in (a). The cross-terms in the Wigner distribution heavily contaminate the signal. and the high frequency burst is not detected. 59

6 A sample time series with a sinusoid in the first 100 points, and normally distributed noise in the last 28 points. The vertical black line denotes the boundary between the two regions. (b) The Wigner distribution of the time series in (a) without the noise region. (c) The Wigner distribution of the time series in (a). The noisy region is propagated into the sinusoidal part of the signal. 60

7 The Wigner distribution of the sum of two chirps illustrating the cross terms [after Qian and Chen, 1996]. 66

8 The bottom plot is a time waveform that contains four frequency tones. The right plot is the traditional power spectrum. The middle one is the joint time-frequency representation [after Qian and Chen, 1996]. 66

9 (a) One period of a signal of frequency $2/256$ and of a signal with frequency $= 3/256$. The difference in period is easily observed. There are some 42 points between trough of the $3/256$ signal, and that of the $2/256$ signal. (b) One period of a signal of frequency $33/256$ and of a signal with frequency $= 34/256$. The different period is very difficult to detect. Thus the frequency dependent resolution of the S-transform is explained (note the change in the x-axis. 77

10 (a) The resolution of time frequency space when the signal is viewed as a time series, having precise time resolution, and no frequency resolution. (b) The resolution of time frequency space when the signal is viewed in the S-Transform representation, having frequency dependent resolution. (c) The resolution of time frequency space when the signal is viewed as a spectrum, having no time resolution, and precise frequency resolution [after Stockwell, 1999]. 78

11 A 256 point sinusoidal time series with a frequency of $32/256$. It is non-zero for a region of 128 points. (b) A region of the amplitude spectrum of the time series in (a). (c) The same time series in (a) but now zero padded up to a length of 768. Note the change in the x-axis. (d) A region of the amplitude spectrum of the time series plotted in (c). The resolution has not changed, but the frequency sampling interval is smaller. (e) The time series in (a) but now exists (non-zero) for 384 points. (f) A region of the amplitude spectrum of the time series plotted in (e). Because the signal has lasted for a longer time interval, it is possible to measure the frequency with much higher resolution than in (d). 79

12 A chirp function with a linearly increasing frequency. (b) The instantaneous frequency of the time series in (a). (c) Amplitude of the S-transform for the time series in (a). (d) The generalized instantaneous frequency calculated from the S-transform for the time series in (a). (e) Amplitude of the STFT for the time series in (a). (f) The Winger distribution for the time series in (a). For such a simple and clean time series, all methods do a very good job in finding the frequency chirp. 85

13 Another chirp function with a linearly increasing frequency added to one with a decreasing frequency. (b) The instantaneous frequency of the time series in (a). (c) Amplitude of the S-transform for the time series in (a). (d) The generalized instantaneous frequency as calculated from the S-transform for the time series in (a). (e) Amplitude of the STFT for the time series in (a). (f) The Winger distribution for the time series in (a). 86

14 A sample time series with 3 distinct components. (b) The instantaneous frequency of the time series in (a). (c) Amplitude of the S-transform for the time series in (a). (d) Amplitude of the STFT for the time series in (a). (e) The Winger distribution for the time series in (a). 87

15 A complicated tiem series with a sinusoidally modulated frequency in the first 512 points, then two sets of crossing chirps in the next 512 points. All throughout the time series, high frequency bursts of short duration are added. (b) The instantaneous frequency of the time series in (a). This identifies the sinusoidal modulation in the first half well, but misses the bursts and the crossed chirps. (c) Amplitude of the S-transform for the time series in (a). This identifies all features. (d) Amplitude of the STFT for the time series in (a). This does a good job, however the high frequency bursts are under represented. (e) The Winger distribution for the time series in (a). The signal is poorly represented. 88

1 Introduction

Spectral analysis began in the second half of the 19th century, with the discovery that the line spectra of atoms and molecules are unique identifiers, and that the characteristics of the spectral lines, shape, and width, are indications of the environment. About the same time, the study of continuous spectra, particularly black-body radiation, became a forefront problem. The explanation of these spectra revolutionized chemistry, physics, and engineering, as they were the seeds for the discovery of modern physics and chemistry. The main effort of spectral analysis in the last 150 years has been the study of signals that do not change in time. However, in the last 50 years it has been realized that for many natural and man-made signals, their spectra to change with time.

One of the central problems in High Frequency (HF) radar data is the analysis of a time series. The problem at hand is how to extract the information present in the data and use it to its full potential. In order to address this problem we turn to the field of signal analysis and data representations. The Fourier Transform is at the heart of a wide range of techniques that are used in HF radar data analysis and processing. Mapping the data into the temporal frequency domain is an effective way of recording the data such that their global characteristics can be assessed. However, the change of frequency content with time is one of the main features we observe in HF radar data. Because of this change of frequency content with time, radar signals belong to the class of non-stationary signals. The analysis of non-stationary signals requires technique that extend the notion of a global frequency spectrum to a local frequency description. The spectral energy density function that is obtained by means of a Fourier Transform, the so-called power spectrum, shows the frequencies that are present in our data, but does not reveal where changes in the frequency content occur. Consequently, for the interpretation of radar data in terms of a changing frequency content, we need a representation of our data as a function of both time and frequency. Only, quite recently, the joint time-frequency representation of signals has become a major area of research in signal processing.

The development of the physical and mathematical ideas needed to explain and understand time-varying spectra has evolved into the field now called "time-frequency analysis". It has been an exciting development due to the introduction of many concepts and myriad of applications. The basic idea is to find a density in time and frequency that indicates which frequencies are present in the signal and how they change in time. This is in contrast to the Fourier transform that shows only the frequencies present in the signal, but does not reveal how frequencies change in time. The ultimate aim is to have a consistent theory of such densities and use them in the same manner as any other joint density. Such densities of time and frequency are called distributions or representations.

The time-frequency representations characterize signals over a time-frequency plane. They thus combine time-domain and frequency-domain analyses to yield a potentially more revealing picture of the temporal localization of a signal's spectral components. They may also serve as a basis for signal detection, characterization, coding, and processing. A complete and comprehensive theory for joint time-frequency analysis does not exist yet. There is no unique time-frequency representation of a signal that satisfies all the properties of a physically correct joint time-frequency energy density function. However, discarding the requirement that all properties must be satisfied in one time-frequency representation, a class of joint time-frequency representations can be derived that serves as a model of a local power spectrum.

In order to employ the concept of a local power spectrum in HF radar signal analysis we need to decide which time-frequency representation is to be used. This choice cannot be made on the basis of a mathematical analysis alone. We should also take into account what we aim to achieve with a local frequency analysis. Our main area of application of the time-frequency representation is the analysis of experimental HF radar data. A study of the properties of different time-frequency representations should provide us a guideline as to which representation is to be chosen in these applications.

This report is reviewing both linear and quadratic time-frequency representations. The linear representations discussed are Short-Time Fourier Transform (STFT) and S-transform. The quadratic representation discussed is Wigner distribution. We outline the motivations, interpretations, mathematical fundamentals, and properties of these linear and quadratic time-frequency representations. We also compare these three different time-frequency analysis techniques and show that each technique has its strengths and drawbacks. The simulated data sets have been used for the comparison. The comparison allows us to determine what would be the most effective methods under different conditions. The choice of the particular time-frequency representation depends upon the specific area of application and what we aim to achieve with a local frequency analysis. We show that time-frequency analysis methods should enable us to classify signals with a considerably greater reflection of the physical situation than can be achieved by the conventional Fourier Transform method alone.

The report is organized into nine sections. In Section 2, we review the time domain and frequency domain as defined by the Fourier Transform. We review in more detail some important relationships between the time and frequency representations, which are fundamental for the time-frequency analysis. In particular, we introduce the concept of instantaneous frequency, mean frequency, mean time, group delay and power spectrum.

Signals in nature are real. Nevertheless, it is often advantageous to define a complex signal that in some sense or other corresponds to the real signal. In Section 3 we describe the motivations for seeking a complex signal representation and its relation to the concept of instantaneous frequency.

Because the time and frequency representations are related via the Fourier Transform, the signal's time and frequency behaviors are not independent. For example, when a signal's time duration gets narrower, its frequency bandwidth must become wider. We cannot make the time duration and frequency bandwidth arbitrarily small simultaneously. This assertion is traditionally named uncertainty principle, which plays an important role in the joint time-frequency analysis. In Section 4 we give mathematical proof of the uncertainty principle.

In Section 5 we introduce the concept of time and frequency marginals, total energy, characteristic functions, global averages, time and frequency shift invariance, and weak and strong finite support, which are important for the time-frequency analysis.

It is the primary goal of this report to introduce the signal's joint time-frequency representations. Analogous to the classical Fourier analysis, we present in parallel the methods of linear and quadratic (or bilinear) joint time-frequency representations. Sections 6 and 8 are devoted to the Short-Time Fourier Transform and the S-transform, respectively. Section 7 is devoted to the Wigner distribution.

Section 6 discusses linear time-frequency representation, the Short-Time Fourier Transform. The STFT is considered in some detail, with emphasis placed on basic properties, running-window, time resolution versus frequency resolution, and the limitations and strengths of the STFT.

Section 7 considers quadratic time-frequency representation, the Wigner distribution. We review the basic properties, time-dependent power spectrum, and the pseudo Wigner distribution. We discuss the occurrence of quadratic cross-term interference that significantly obscures the applications of the Wigner distribution. We also compare the Wigner distribution with STFT.

Section 8 discusses linear time-frequency representation, the S-transform. We derive the S-transform from the STFT. In this section we review the general properties of S-transform, progressive resolution of the time-frequency domain, and the generalized S-transform. In this section we also present computer simulation results with the aim of comparing different time-frequency representations. Finally, general conclusions are summarized in Section 9.

2 Description of Signals in the Time and Frequency Domains

2.1 Introduction

This section provides a brief review of the fundamentals to time and frequency analysis. Signal analysis is the study and characterization of the basic properties of signals and was historically developed concurrently with the discovery of the fundamental signals in nature, such as the electric field, sound wave, and electric currents. A signal is generally a function of many variables. For example, the electric field varies in both space and time. Our main emphasis will be the time variation, although the ideas developed are easily extended to spatial and other variables, and we do so in the latter part of the report.

Among the number of infinite possible representations to study the signal, the most important are time and frequency, because they are closely related to our everyday life. The time variation of a signal is fundamental because time is fundamental. However, if we want to gain more understanding, it is often advantageous to study the signal in a different representation. This is done by expanding the signal in a complete set of functions, and from a mathematical point of view there are an infinite number of ways this can be done. What makes a particular representation important is that the characteristics of the signal are understood better in that representation because the representation is characterized by a physical quantity that is important in nature or for the situation at hand. Besides time, the most important representation is frequency.

Based on frequency behaviors, signals can further be grouped into two categories. First is the one whose frequency contents are not changed with time. It has been well known that the frequency behavior of this kind of signal can be well characterized by the conventional Fourier Transform (FT). Very often, people call this type of signal a stationary signal. The second type of signals are those frequency contents evolve with time. This kind of signal is usually called a non-stationary signal. The majority of signals encountered in the real world belong to this category. Because the conventional Fourier Transform does not tell how a signal's frequency contents change in time, the classical Fourier analysis is not adequate for many real signals.

In recent years, the time honoured technique of Fourier analysis has given way to more advanced representations known as joint time-frequency representations. In music, the changing tones are essential. A bird's chirp is a good example of frequency changing with time, as well as the voice of a person in conversation, where the harmonic tones carry all the information. While these examples are auditory, simply watching a sunset as the sky changes color is another case of changing frequencies. As with all these examples, it is change in frequency with time that contains the information, not necessarily the frequency itself. The goal of this report is to systematically discuss new representations that describe a signal's behavior in time and frequency domains simultaneously.

The discussion in this section mainly follows that of *Brigham* [1975], *Oppenheim and Schafer* [1975], *Papoulis* [1977], *Bracewell* [1978], *Bloomfield* [1976] *Portnoff* [1980], *Cohen* [1989, 1995], *Rioul and Vetterli* [1991], *Hlawatsch and Boudreaux-Bartels* [1992], *Strang and Nguyen* [1995], *Vetterli and Kovacevic* [1995], *Qian and Chen* [1996], *Hlawatsch* [1998], *Mertins* [1999] in addition to the references given. Although the materials presented in this section may not be completely new, it is certainly beneficial to go through them before reading the rest of the report. The concepts introduced in this section will be extensively used for the developments.

2.2 Description of Signals in the Time Domain

The term signals generally refers to a function of one or more independent variables, which contain information about the behavior or nature of some phenomenon. The common examples of the signals include electrical current, image, speech signals, pressure, and electromagnetic field, etc., which all produced by some time-varying processes. The simplest time-varying signal is the sinusoid. It is a solution to many of the fundamental equations, such as Maxwell equations, and is common in nature. We shall denote a signal by $s(t)$. It is characterized by a constant amplitude, a, and constant frequency, ω_0,

$$s(t) = a\ \cos \omega_0 t \tag{1}$$

The general signal can be written in the form

$$s(t) = a(t)\ \cos \theta(t) \tag{2}$$

where the amplitude, $a(t)$, and phase, $\theta(t)$, are now arbitrary functions of time. To emphasize that they generally change in time, the phrases amplitude modulation and phase modulation are often used, since the word modulation means change. It is also often advantageous to write a signal in complex form

$$s(t) = A(t)\, e^{i\varphi(t)} = s_r + i\, s_i \tag{3}$$

where s_r and s_i are real and imaginary parts of the signal. It is important to note that the phase and amplitude of the real signal are not generally the same as the phase and amplitude of the complex signal. We have emphasized this by using different symbols for the phases and amplitudes in equations (2) and (3).

2.2.1 Instantaneous Power

In the case of electromagnetic theory, the electric energy density is the absolute square of the electric field and similarly for the magnetic field. This was derived by Poynting using Maxwell's equations and is known as Poynting's theorem. In circuits, the energy density is proportional to the voltage squared. For a sound wave it is the pressure squared. Therefore, the energy or intensity of a signal is generally $\mid s(t) \mid^2$. That is, in a small interval of time, Δt, it takes $\mid s(t) \mid^2 \Delta t$ amount of energy to produce the signal at that time. Since $\mid s(t) \mid^2$ is the energy per unit time it may be appropriately called the energy density or the instantaneous power since power is the amount of work per unit time [*Cohen,* 1989, 1995]. Therefore

$$\begin{aligned} \mid s(t) \mid^2 &= \text{instantaneous power at time } t \\ \mid s(t) \mid^2 \Delta t &= \text{the fractional energy in the time interval } \Delta t \text{ at time } t \end{aligned}$$

2.2.2 Total Energy

If $\mid s(t) \mid^2$ is the energy per unit time, then the total energy is obtained by summing or integrating over all time [*Cohen,* 1989, 1995; *Mertins,* 1999]

$$E = \int \mid s(t) \mid^2 \, dt \tag{4}$$

For signals with finite energy we can take, without loss of generality, the total energy to be equal to one. For many signals the total energy is infinite. For example, a pure sine wave has infinite total energy, which is reasonable since to keep on producing it, work must be expended continually.

2.2.3 Time Averages, Mean Time, and Duration

In this section and the following sections, we review some basic connections between a signal's time and frequency representations. The most important relationship in terms of joint time-frequency analysis, however, is the relationship between a signal's duration and frequency bandwidth. The concepts introduced in this section play significant roles in joint time-frequency analysis.

The average time is defined as [*Young,* 1962; *Taylor,* 1982; *Cohen,* 1989, 1995; *Qian and Chen,* 1996; *Mertins* , 1999]:

$$< t >= \int t \mid s(t) \mid^2 \, dt \tag{5}$$

The reasons for defining an average are that it may give a gross characterization of the density and it may give an indication of where the density is concentrated.

A measure for the duration of a signal, denoted by T, can be found as the standard deviation σ_t around the average time [*Young,* 1962; *Taylor,* 1982; *Cohen,* 1989, 1995; *Qian and Chen,* 1996; *Mertins,* 1999]:

$$T^2 = \sigma_t^2 = \int (t- < t >)^2 \mid s(t) \mid^2 \, dt =< t^2 > - < t >^2 \tag{6}$$

where $< t^2 >$ is defined similarly to $< t >$. The standard deviation is an indication of the duration of the signal: in a time $2\sigma_t$ most of the signal will have gone by. If the standard deviation is small then most of the signal is concentrated around the mean time and it will go by quickly, which is an indication that we have a signal of short duration; similarly for long duration. It should be pointed out that there are signals for which the standard deviation is infinite, although they may be finite energy signals. That usually indicates that the signal is very long lasting. The average of any function of time, $g(t)$, is obtained by [*Cohen*, 1966, 1989, 1995]

$$< g(t) >= \int g(t) \mid s(t) \mid^2 dt \tag{7}$$

Note that for a complex signal, time averages depend only on the amplitude.

2.3 Description of Signals in the Frequency Domain

Although a given signal can be represented in many different ways, the most important are the time and frequency representations. The significance of the quantity time is easy to understand, because it is fundamental. The majority of signals encountered in our everyday life are directly related to time. The frequency representations, on the other hand, were not popular until the early 19th century when Fourier first proposed the harmonic trigonometric series. Since then, the frequency representation has become one of the most powerful and standard tools for studying signals. By using frequency representations, we could better understand many physical phenomenon and accomplish many things that cannot achieved based on time representations. The bridge between time and frequency is the $Fourier\ Transform$.

2.3.1 Fourier Transform

The signal can be represented as a sum of sinusoid frequency components [*Brigham*, 1974; *Bloomfield,* 1976; *Papoulis* 1977; *Bracewell,* 1978]

$$s(t) = \frac{1}{\sqrt{2\pi}} \int S(\omega)\, e^{j\omega t}\, d\omega \tag{8}$$

The waveform is made up of the addition (linear superposition) of the simple waveforms, $e^{j\omega t}$, each characterized by the frequency, ω, and contributing a relative amount indicated by the coefficient, $S(\omega)$. $S(\omega)$ is obtained from the signal by

$$S(\omega) = \frac{1}{\sqrt{2\pi}} \int s(t)\, e^{-j\omega t}\, dt \tag{9}$$

and is called the spectrum or the Fourier Transform. Since $S(\omega)$ and $s(t)$ are uniquely related we may think of the spectrum as the signal in the frequency domain or frequency space or frequency representation.

2.3.2 Discrete Fourier Transform

In dealing with a discrete time series of N points with a sampling interval of T, the Discrete Fourier Transform (DFT) is used (where $k = 0...N-1$ and $m = 0...N-1$) [*Papoulis,* 1977; *Hlawatsch,* 1998; *Mertins,* 1999]

$$S[\frac{n}{NT}] = \frac{1}{N} \sum_{k=0}^{N-1} s[kT]\, e^{\frac{-i2\pi nkT}{NT}} \tag{10}$$

and its inverse relationship is:

$$s[kT] = \sum_{n=0}^{N-1} S[\frac{n}{NT}]\, e^{\frac{i2\pi nkT}{NT}} \tag{11}$$

The Fourier Transform and its inverse establish a one-to-one relation between the time domain and frequency domain. The time domain and frequency domain constitute two alternative ways of looking at a signal. Although the Fourier Transform allows a passage from one domain to the other, it does not allow a combination of the two domains. In particular, most time information is easily accessible in the frequency domain. While the spectrum $S(\omega)$ shows the overall strength with which any frequency f is contained in the signal $s(t)$, it does not generally provide easy to interpret information about the *time localization* of spectral components. Strictly speaking, this information is contained in the space spectrum but often comes in a form that is not easily interpreted as is discussed in the following sections.

2.3.3 Spectral Amplitude and Phase

As with the signal, it is often advantageous to write the spectrum in terms of its amplitude and phase [*Papoulis,* 1977; *Bracewell,* 1978; *Cohen,* 1995; *Mertins,* 1998]

$$S(\omega) = B(\omega)\, e^{j\psi(\omega)} \tag{12}$$

We call $B(\omega)$ the spectral amplitude and $\psi(\omega)$ the spectral phase to differentiate them from the phase and amplitude of the signal.

2.3.4 Energy Spectrum Density

In analogy with the time waveform we can take $\mid S(\omega) \mid^2$ to be the energy spectrum density per unit frequency [*Cohen,* 1989, 1995]:

$$\begin{aligned} \mid S(\omega) \mid^2 &= \text{energy spectrum density at frequency } \omega \\ \mid S(\omega) \mid^2 \, \Delta\omega &= \text{the fractional energy in the frequency interval } \Delta\omega \text{ at frequency } \omega \end{aligned}$$

That $\mid S(\omega) \mid^2$ is the energy density can be seen by considering the simple case of one component, $s(t) = S(\omega_0)\, e^{j\omega_0 t}$, characterized by the frequency, ω_0. Since the signal energy is $\mid s(t) \mid^2$, then for this case the energy density is $\mid S(\omega_0) \mid^2$. Since all the energy is in one frequency, $\mid S(\omega_0) \mid^2$ must then be the energy for that frequency.

The total energy of the signal should be independent of the method used to calculate it. Hence, if the energy density per unit frequency is $\mid S(\omega) \mid^2$, the total energy should be the integral of $\mid S(\omega) \mid^2$ over all frequencies and should equal the total energy of the signal calculated directly from the time waveform [*Brigham,* 1974; *Bloomfield,* 1976; *Papoulis* 1977; *Bracewell,* 1978; *Cohen,* 1989, 1995; *Mertins,* 1998]

$$E = \int \mid s(t) \mid^2 \, dt = \int \mid S(\omega) \mid^2 \, d\omega \tag{13}$$

This identity is commonly called Parceval's or Rayleigh's theorem. To prove it consider [*Cohen,* 1995]

$$E = \int |\, s(t)\, |^2 \; dt \;\; = \;\; \frac{1}{2\pi} \int\int\int \; S^\star(\omega')\, S(\omega)\, e^{j(\omega-\omega')t}\, d\omega\, d\omega'\, dt \tag{14}$$

$$= \int\int \; S^\star(\omega')\, S(\omega)\, \delta(\omega-\omega')\, d\omega\, d\omega' \tag{15}$$

$$= \int |\, S(\omega)\, |^2 \; d\omega \tag{16}$$

where in going from equation (14) to (15) we have used

$$\frac{1}{2\pi} \int \; e^{j(\omega-\omega')t}\, dt = \delta(\omega-\omega') \tag{17}$$

The energy density spectrum tells us which frequencies existed during the total duration of the signal. It gives us no indication as to when these frequencies existed. The mathematical and physical ideas needed to understand and describe how the frequencies are changing in time is the subject of this report.

2.3.5 Mean Frequency, Bandwidth, and Frequency Averages

If $|\, S(\omega)\, |^2$ represents the energy density in frequency then we can use it to calculate averages, the motivation being the same as in the time domain, namely that it gives a rough idea of the main characteristics of the spectral density.

In a similar way as the average time, the average frequency, $<\omega>$, and its standard deviation, σ_ω, are given by [*Young,* 1962; *Taylor,* 1982; *Cohen,* 1989, 1995; *Qian and Chen,* 1996; *Mertins,* 1998]:

$$<\omega> \;\; = \;\; \int \omega \; |\, S(\omega)\, |^2 \; d\omega \tag{18}$$

$$B^2 = \sigma_\omega^2 \;=\; \int (\omega - <\omega>)^2 \mid S(\omega) \mid^2 \, d\omega \tag{19}$$

$$= \; <\omega^2> - <\omega>^2 \tag{20}$$

B is commonly known as the bandwidth. The average of any frequency function, $g(\omega)$, is

$$< g(\omega) >= \int g(\omega) \mid S(\omega) \mid^2 \, d\omega \tag{21}$$

2.4 The Frequency and the Time Operators

From the definition, the equation (18), shows that when we want to calculate the average frequency we first have to obtain the spectrum. But that is not so. There is an important method that avoids the calculation of the spectrum, simplifies the algebra immensely, and moreover will be central to our development in the later sections for deriving time-frequency representations.

2.4.1 The Frequency Operator

For convenience one defines the frequency operator [*Qian and Chen,* 1996; *Cohen,* 1989, 1995]:

$$\mathcal{W} = \frac{1}{j}\frac{d}{dt} \tag{22}$$

and it is understood that repeated use, denoted by $\mathcal{W}^n$, is to mean repeated differentiation,

$$\mathcal{W}^n \, s(t) = \left(\frac{1}{j}\right)^n \frac{d^n}{dt^n} \, s(t) \tag{23}$$

We are now in a position to state and prove the general result that the average of a frequency function can be calculated directly from the signal by way of [*Qian and Chen,* 1996; *Cohen,* 1989, 1995]

$$< g(\omega) > \; = \int g(\omega)\,|S(\omega)|^2\,d\omega \tag{24}$$

$$= \int s^{\star}(t)\,g(\mathcal{W})\,s(t)\,dt \tag{25}$$

$$= \int s^{\star}(t)\,g\left(\frac{1}{j}\frac{d}{dt}\right)\,s(t)\,dt \tag{26}$$

In words: Take the function $g(\omega)$ and replace the ordinary variable ω by the operator $\frac{1}{j}\frac{d}{dt}$; operate on the signal, multiply by the complex conjugate signal, and integrate. Before we prove this we must face a side issue and discuss the meaning of $g(\mathcal{W})$ for an arbitrary function. If g is ω^n, then the procedure is clear, as indicated by equation (23). If g is the sum of powers, then it is also clear. For a general function we first expand the function in a Taylor series and then substitute the operator $\mathcal{W}$ for ω [*Cohen,* 1995]. That is,

$$\text{if} \quad g(\omega) = \sum g_n\,\omega^n \quad \text{then} \quad g(\mathcal{W}) = \sum g_n\,\mathcal{W}^n \tag{27}$$

To prove the general result, equation (26), we first prove it for the case of the average frequency [*Cohen,* 1995]:

$$< \omega > = \int \omega|S(\omega)|^2\,d\omega \; = \; \frac{1}{2\pi}\int\!\!\int\!\!\int \omega\,s^{\star}(t)\,s(t')\,e^{j(t-t')\omega}\,d\omega\,dt'\,dt \tag{28}$$

$$= \frac{1}{2\pi j}\int\!\!\int\!\!\int s^{\star}(t)\,s(t')\,\frac{\partial}{\partial t}e^{j(t-t')\omega}\,d\omega\,dt'\,dt \tag{29}$$

$$= \frac{1}{j}\int\!\!\int s^{\star}(t)\,\frac{\partial}{\partial t}\delta(t-t')\,s(t')\,dt'\,dt \tag{30}$$

$$= \int s^{\star}(t)\,\frac{1}{j}\frac{d}{dt}\,s(t)\,dt \tag{31}$$

These steps can be repeated as often as necessary to prove it for $g = \omega^n$. Hence

$$< \omega^n > = \int s^{\star}(t)\left(\frac{1}{j}\frac{d}{dt}\right)^n s(t)\,dt = \int s^{\star}(t)\,\mathcal{W}^n\,s(t)\,dt \tag{32}$$

Having proved the general result for functions of the form $g = \omega^n$, we now prove it for an arbitrary function, $g(\omega)$, by expanding the function in a Taylor series [*Cohen,* 1995]

$$< g(\omega) > = \int g(\omega)\,|S(\omega)|^2\,d\omega \;=\; \int \sum g_n\,\omega^n\,|S(\omega)|^2\,d\omega \tag{33}$$

$$= \sum g_n \int s^\star(t)\,\mathcal{W}^n\,s(t)\,dt \tag{34}$$

$$= \int s^\star(t)\,g(\mathcal{W}^n)\,s(t)\,dt \tag{35}$$

Manipulation Rules.

The frequency operator is a Hermitian operator, which means that for any two signals, $s_1(t)$ and $s_2(t)$ [*Cohen,* 1989, 1995; *Hlawatsch,* 1998; *Mertins,* 1999],

$$\int s_1^\star(t)\,\mathcal{W}\,s_2(t)\,dt = \int s_2(t)\,\{\mathcal{W}\,s_1(t)\}^\star\,dt \tag{36}$$

This is ready proved by integrating by parts. Also, a real function, $g(\omega)$, of a Hermitian operator, $g(\mathcal{W})$, is also Hermitian. That is,

$$\int s_1^\star(t)\,g(\mathcal{W})\,s_2(t)\,dt = \int s_2(t)\,\{g(\mathcal{W})\,s_1(t)\}^\star\,dt \qquad [\text{if } g(\omega) \text{ is real}] \tag{37}$$

An important property of Hermitian operators is that their average value as defined by equation (35) must be real, so in the manipulation of averages we can simply discard the imaginary terms since we are assured that they add up to zero.

We now derive the second simplification for the average square of frequency [*Cohen,* 1995]. We have

$$< \omega^2 > = \int s^\star(t)\,\mathcal{W}^2\,s(t)\,dt \;=\; \int s^\star(t)\,\mathcal{W}\,\mathcal{W}\,s(t)\,dt \tag{38}$$

$$= \int \mathcal{W}\,s(t)\,\{\mathcal{W}\,s(t)\}^\star\,dt \tag{39}$$

$$= \int |\mathcal{W}\, s(t)|^2\, dt \tag{40}$$

This is an immense simplification since not only do we not have to find the spectrum, we also avoid a double differentiation. Therefore, the average frequency and average square of frequency is given by [*Cohen*, 1995]

$$< \omega > = \int \omega |S(\omega)|^2\, d\omega = \int s^\star(t)\, \frac{1}{j}\, \frac{d}{dt}\, s(t)\, dt \tag{41}$$

$$\begin{aligned} < \omega^2 > = \int \omega^2 |S(\omega)|^2\, d\omega &= \int s^\star(t) \left(\frac{1}{j}\frac{d}{dt}\right)^2 s(t)\, dt && (42) \\ &= -\int s^\star(t)\, \frac{d^2}{dt^2}\, s(t)\, dt && (43) \\ &= \int \left| \frac{d}{dt}\, s(t) \right|^2 dt && (44) \end{aligned}$$

The bandwidth is given by [*Cohen*, 1995]

$$\begin{aligned} \sigma_\omega^2 &= \int (\omega - < \omega >)^2\, |S(\omega)|^2\, d\omega && (45) \\ &= \int s^\star(t) \left(\frac{1}{j}\frac{d}{dt} - < \omega > \right)^2 s(t)\, dt && (46) \\ &= \int \left| \left(\frac{1}{j}\frac{d}{dt} - < \omega > \right) s(t) \right|^2 dt && (47) \end{aligned}$$

2.4.2 The Time Operator

In the above discussion we emphasized that we can avoid the necessity of calculating the spectrum for the calculation of averages of frequency functions. Similarly if we have a spectrum and want to calculate time averages, we can avoid the calculation of the signal. The time operator is defined by [*Qian and Chen,* 1996; *Cohen,* 1989, 1995]

$$\mathcal{T} = -\frac{1}{j}\frac{d}{d\omega} \tag{48}$$

and the same arguments and proofs as above lead to [*Cohen,* 1995]

$$< g(t) > = \int g(t)\,|s(t)|^2\,dt = \int S^{\star}(\omega)\,g(\mathcal{T})\,S(\omega)\,d\omega \tag{49}$$

In particular,

$$< t > = \int t\,|s(t)|^2\,dt = S^{\star}(\omega)\left(-\frac{1}{j}\frac{d}{d\omega}\right)S(\omega)\,d\omega \tag{50}$$

$$\begin{aligned} < t^2 > = \int t^2\,|s(t)|^2\,dt &= S^{\star}(\omega)\left(-\frac{1}{j}\right)^2\frac{d^2}{d\omega^2}S(\omega)\,d\omega && (51)\\ &= -\int S^{\star}(\omega)\frac{d^2}{d\omega^2}S(\omega)\,d\omega && (52)\\ &= \int\left|\frac{d}{d\omega}S(\omega)\right|^2 d\omega && (53) \end{aligned}$$

2.4.3 Mean Frequency and Instantaneous Frequency

Consider first

$$\begin{aligned} \mathcal{W}\,s(t) = \mathcal{W}\,A(t)\,e^{\varphi(t)} &= \frac{1}{j}\frac{d}{dt}A(t)\,e^{\varphi(t)} && (54)\\ &= \left(\varphi^{'}(t) - j\,\frac{A^{'}(t)}{A(t)}\right)s(t) && (55) \end{aligned}$$

Therefore, the mean frequency is [*Fink,* 1966; *Mandel,* 1974; *Cohen and Lee,* 1986; *Cohen,* 1995]

$$< \omega > = \int \omega |S(\omega)|^2 \, d\omega \;=\; \int s^{\star}(t) \, \frac{1}{j} \frac{d}{dt} \, s(t) \, dt \tag{56}$$

$$= \int \left(\varphi^{'}(t) - j \, \frac{A^{'}(t)}{A(t)} \right) \, A^2(t) \, dt \tag{57}$$

The second term is zero. This can be seen in two ways. First, since that term is purely imaginary part it must be zero for $< \omega >$ to be real. Alternatively, we note that the integrand of the second term is a perfect differential that integrates to zero [*Cohen,* 1995]. Hence

$$< \omega > = \int \varphi^{'}(t) \, |s(t)|^2 \, dt = \int \varphi^{'}(t) \, A^2(t) \, dt \tag{58}$$

This is an interesting and important result because it says that the average frequency may be obtained by integrating "something" with the density over all time. This something must be the instantaneous value of the quantity for which we are calculating the average. In this case the something is the derivative of the phase, which may be appropriately called the frequency at each time or the *instantaneous frequency* , $\omega_i(t)$ [*Cohen,* 1989, 1995; *Qian, Chen,* 1996; *Vakman,* 1996]

$$\omega_i(t) = \varphi^{'}(t) \tag{59}$$

Instantaneous frequency, as an empirical phenomenon, is experienced daily as changing colors, changing pitch, etc. Whether or not the derivative of the phase meets our intuitive concept of instantaneous frequency is a central issue and is addressed in subsequent sections.

2.4.4 Mean Time and Group Delay

The identical derivations can be used to write the mean time. In particular [*Cohen,* 1989, 1995; *Qian and Chen,* 1996; *Vakman,* 1996]

$$< t > = - \int \psi^{'}(\omega)\,|S(\omega)|^2\, d\omega \tag{60}$$

It says that if we average $-\psi^{'}(\omega)$ over all frequencies we will get the average time. Therefore we may consider $-\psi^{'}(\omega)$ to be the average time for a particular frequency. This is called the *group delay* and we shall use the following notation for it [*Rihaczek,* 1969; Papoulis, 1977; *Cohen,* 1995]

$$t_g(\omega) = -\psi^{'}(\omega) \tag{61}$$

2.4.5 The Translation Operator

Many results in signal analysis are easily derived by the use of the translation operator, $e^{j\tau\mathcal{W}}$, where τ is a constant. Its effect on a function of time is [*Cohen*, 1989, 1995; *Qian and Chen,* 1996; *Hlawatsch,* 1998; *Mertins,* 1999]

$$e^{j\tau\mathcal{W}}\, f(t) = f(t+\tau) \tag{62}$$

That is, the translation operator translates function by τ. Note that it is not Hermitian. To prove equation (62) consider [*Cohen,* 1995]

$$e^{j\tau\mathcal{W}}\, f(t) = \sum_{n=0}^{\infty} \frac{(j\tau)^n\, \mathcal{W}^n}{n!}\, f(t) = \sum_{n=0}^{\infty} \frac{\tau^n}{n!}\, \frac{d^n}{dt^n}\, f(t) \tag{63}$$

But this is precisely the Taylor expansion of $f(t+\tau)$ and hence equation (62) follows. Similarly, the operator $e^{-j\theta\mathcal{T}}$ translates frequency functions,

$$e^{j\theta\mathcal{T}}\, S(\omega) = S(\omega+\theta) \tag{64}$$

2.5 The Covariance of a Signal

If we want a measure of how time and instantaneous frequency are related, the calculation of the covariance or the correlation coefficient is desired. Consider the quantity [*Cartwright*, 1974; *Cohen*, 1989, 1995; *Mertins*, 1999]

$$< t\,\varphi'(t) >= \int t\,\varphi'(t)\,|s(t)|^2\,dt \tag{65}$$

which may be thought of as the average of time multiplied by the instantaneous frequency. If time and frequency have nothing to do each other then we would expect $< t\,\varphi'(t) >$ to equal $< t >< \varphi'(t) >=< t >< \omega >$. Therefore the excess of $< t\,\varphi'(t) >$ over $< t >< \omega >$ is a good measure of how time is correlated with instantaneous frequency. This is precisely what is called the covariance of a signal by [*Cartwright*, 1974]

$$\mathrm{Cov}_{t\omega} =< t\,\varphi'(t) > - < t >< \omega > \tag{66}$$

The correlation coefficient is the normalized covariance

$$r = \frac{\mathrm{Cov}_{t\omega}}{\sigma_t\,\sigma_\omega} \tag{67}$$

The reason for defining the correlation coefficient in the standard considerations is that it ranges from minus one to one and hence gives an absolute measure. That is not the case here, but nonetheless it does give a good indication of the relationship between the time and frequency.

2.5.1 The Covariance of a Spectrum

Suppose we consider in the frequency domain so that time is t_g and frequency is ω. It is reasonable to define the covariance by [*Cartwright*, 1974; *Cohen*, 1995]

$$\mathrm{Cov}_{t\omega} =< t_g\,\omega > - < t >< \omega > \tag{68}$$

with

$$< t_g \, \omega >= - < \omega \, \psi'(\omega) >= - \int \omega \, \psi'(\omega) \, |S(\omega)|^2 \, d\omega \tag{69}$$

For equations (66) and (68) to be identical we must have

$$\int t \, \varphi'(t) \, |s(t)|^2 \, dt = - \int \omega \, \psi'(\omega) \, |S(\omega)|^2 \, d\omega \tag{70}$$

It is an interesting identity because it connects the phases and amplitudes of the signal and spectrum.

2.5.2 When is the Covariance Equal to Zero ?

If the covariance is to be an indication of how instantaneous frequency and time are related, then when the instantaneous frequency does not change the covariance should be zero [*Cohen,* 1995]. That is indeed the case. Consider

$$s(t) = A(t) \, e^{j\omega_0 t} \tag{71}$$

where the amplitude modulation is arbitrary. Now

$$< t \, \varphi'(t) >= \int t \, \omega_0 \, |A(t)|^2 \, dt = \omega_0 \, < t > \tag{72}$$

But since $< \omega >= \omega_0$, we have

$$< \omega >< t >= \omega_0 \, < t > \tag{73}$$

and therefore the covariance and correlation coefficient are equal to zero. Similarly, if we have a spectrum of the form $S(\omega) = B(\omega) \, e^{j\omega \, t_0}$, then there is no correlation between time and frequency. In general,

$$\mathrm{Cov}_{t\omega} = 0 \quad ; \quad r = 0 \qquad \text{for} \quad s(t) = A(t)\, e^{j\omega_0 t} \tag{74}$$

or

$$S(\omega) = B(\omega)\, e^{j\omega\, t_0} \tag{75}$$

2.6 Power Spectrum

The square of the Fourier Transform $| S(\omega) |^2$ is called *power spectrum*, which indicates how the signal energy is distributed in the frequency domain. While the Fourier Transform $S(\omega)$ is a linear function of the analyzed signal, the power spectrum $| S(\omega) |^2$ is quadratic to the signal $s(t)$. The Fourier Transform $S(\omega)$ in general is complex, whereas the power spectrum $| S(\omega) |^2$ is always real. The Fourier Transform and the power spectrum are the two most important tools for frequency analysis [*Lampard,* 1954; *Schroeder and Atal,* 1962; *Papoulis,* 1977; *Cohen,* 1989, 1995; *Qian and Chen,* 1996; *Mertins,* 1999].

According to the Wiener-Khinchin theorem, the power spectrum can also be written as the Fourier Transform of the signal's auto-correlation function, i.e.,

$$| S(\omega) |^2 = \int R(\tau)\, e^{-j\omega\tau}\, d\tau \tag{76}$$

where the auto-correlation function $R(\tau)$ is computed by

$$R(\tau) = \int s^\star(t-\tau)\, s(t)\, dt \tag{77}$$

The representation (76) is useful, which leads to a feasible way of designing the joint time-frequency representations. For example, if we make $R(\tau)$ time dependent, such as $R(\tau, t)$, then the resulting Fourier Transform manifestly is the function of time and frequency, i.e.,

$$P(t, \omega) = \int R(t, \tau)\, e^{-j\omega\tau}\, d\tau \tag{78}$$

which links the power spectrum to time. Hence, we name $P(t, \omega)$ the time-dependent spectrum. Good examples include the STFT spectrogram as well as Wigner distribution. We shall elaborate on this subject in more detail in the following sections.

2.7 Nonadditivity of Spectral Properties

Many of the conceptual difficulties associated with time-frequency analysis are a reflection of the basic properties of signals and spectra. It must be emphasized the frequency content is not additive. Suppose we have a signal composed of two parts, the spectrum will be the sum of the corresponding spectrum of each part of the signal [*Papoulis,* 1977; *Cohen,* 1989; *Qian and Chen,* 1996; *Hlawatsch,* 1998]

$$s = s_1 + s_2 \quad ; \quad S = S_1 + S_2 \tag{79}$$

However, the energy density is not the sum of the energy densities of each part

$$\mid S \mid^2 \;=\; \mid S_1 + S_2 \mid^2 = \mid S_1 \mid^2 + \mid S_2 \mid^2 + 2\,\Re\{S_1^* S_2\} \tag{80}$$

$$\neq \; \mid S_1 \mid^2 + \mid S_2 \mid^2 \tag{81}$$

Thus the frequency content is not the sum of the frequency content of each signal. The physical reason is that when we add two signals, the waveforms may add and interfere in all sort of ways to give different weights to the original frequencies. Mathematically this is reflected by the fact that the energy density spectrum is the absolute square of the sum of the spectra. How the intensities change is taken into account by equation (80).

3 Instantaneous Frequency and the Analytical Signal

3.1 Introduction

The concept of "instantaneous" frequency has a long history in physics and astronomy. Historically the methodology and description of instantaneous frequency has not always been associated with time-frequency distributions or a time-varying spectrum. A comprehensive theory of joint time-frequency distributions would be able to encompass and clarify the concept of instantaneous frequency, so it is important to appreciate the work has been done along these lines. It was *Armstrong's* [1936] discovery that frequency modulation for radio transmission reduces noise significantly, which produced a concerted effort to understand and describe the mathematical and conceptual description of frequency modulation and instantaneous frequency. Early comprehensive works on the analysis of frequency modulation were those of *Carson and Fry* [1937] and *Van der Pol* [1946], who defined instantaneous frequency as the rate of change of phase of the signal. This definition implies that we have some procedure for forming a complex signal from a real one.

Signals in nature are real. Nevertheless, it is often advantageous to define a complex signal that in some sense or other corresponds to the real signal. In this section we describe the motivations for seeking a complex signal representation and its relation to the concept of instantaneous frequency. One of the motives for defining the complex signal is that it will allow us to define the phase, from which we can obtain the instantaneous frequency.

We seek a complex signal, $z(t)$, whose real part is the "real signal", $s_r(t)$, and whose imaginary part, $s_i(t)$, is our choice, chosen to achieve a sensible physical and mathematical description,

$$z(t) = s_r + j\, s_i \doteq A(t)\, e^{j\varphi(t)} \tag{82}$$

If we can fix the imaginary part we can then unambiguously define the amplitude and phase by [*Cohen,* 1995]

$$A(t) = \sqrt{s_r^2 + s_i^2} \quad ; \quad \varphi(t) = \arctan\left(\frac{s_r}{s_i}\right) \tag{83}$$

using equation (59), we get

$$\omega_i(t) = \varphi^{'}(t) = \frac{(s_i^{'}\, s_r - s_r^{'}\, s_i)}{A^2} \tag{84}$$

for the instantaneous frequency.

3.2 The Analytical Signal

A major step was made by *Gabor* [1946], who from the observation that both $\sin\omega t$ and $\cos\omega t$ transform into an exponential $e^{j\omega t}$ if we use only their positive spectrum, generalized to the arbitrary case with the prescription to "suppress the amplitudes belonging to negative frequencies and multiply the amplitudes of positive frequencies by 2". He noted that this procedure is equivalent to adding to the signal an imaginary part, which is the Hilbert transform of the signal. The positive frequencies are multiplied by 2 to preserve the total energy of the original signal (see Section (3.2.1)).

To see how Hilbert transform arises from the above prescription, suppose the real signal, $s(t)$, has the spectrum, $S(\omega)$, then the complex signal, $z(t)$, whose spectrum is composed of the positive frequencies of $S(\omega)$ only, is given by the inverse transform of $S(\omega)$, where the integration goes only over the positive frequencies [*Gabor,* 1946; *Bedrosian,* 1953; *Nuttall,* 1966; *Rihaczek,* 1966; *Cohen,* 1995; *Vakman,* 1996],

$$z(t) = 2\,\frac{1}{\sqrt{2\pi}} \int_0^{\infty} S(\omega)\, e^{j\omega t}\, dt \tag{85}$$

We now obtain the explicit form for $z(t)$ in terms of the real signal $s(t)$ [*Cohen,* 1995]. Since

$$S(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(t)\, e^{-j\omega t}\, dt \tag{86}$$

we have

$$z(t) \;=\; 2\,\frac{1}{\sqrt{2\pi}} \int_0^{\infty} \int s(t^{'})\, e^{-j\omega t^{'}}\, e^{j\omega t}\, dt^{'}\, d\omega \tag{87}$$

$$= \frac{1}{\pi} \int_0^{\infty} \int s(t')\, e^{j\omega(t-t')}\, dt'\, d\omega \tag{88}$$

and using

$$\int_0^{\infty} e^{j\omega x}\, d\omega = \pi\, \delta(x) + \frac{j}{x} \tag{89}$$

we obtain

$$z(t) = \frac{1}{\pi} \int_0^{\infty} s(t') \left[\pi\, \delta(t-t') + \frac{j}{t-t'} \right] dt' \tag{90}$$

yielding

$$\mathcal{A}[s] = z(t) = s(t) + \frac{j}{\pi} \int \frac{s(t')}{t-t'}\, dt' \tag{91}$$

We use the notation $A[s]$ to denote the analytic signal corresponding to the signal s. The reason for the name analytic is that these types of complex functions satisfy the Cauchy-Riemann conditions for differentiability and have been traditionally called analytic functions. The second part of equation (91) is the Hilbert transform of the signal and there are two conventions to denote the Hilbert transform of a function, $\hat{s}(t)$ and $H[s(t)]$ [*Bedrosian,* 1953; *Nuttall,* 1966; *Rihaczek,* 1966; *Cohen,* 1995; *Vakman,* 1996]

$$H[s(t)] = \hat{s}(t) = \frac{j}{\pi} \int \frac{s(t')}{t-t'}\, dt' \tag{92}$$

Therefore using equation (59) the Instantaneous Frequency (IF) is given by (59) [*Papoulis,* 1977; *Vakman,* 1996]:

$$IF = \frac{\partial \varphi(t)}{\partial t} = \frac{\partial}{\partial t} \arctan\left(\frac{\hat{s}(t)}{s(t)} \right) \tag{93}$$

The derivative of the phase of the analytic signal conforms to our expectations of instantaneous frequency for a wide variety of cases, particularly narrow-band signals. There has been considerable controversy over whether this represents the proper mathematical expression of instantaneous frequency, and a number of other definitions have been given [*Shekel,* 1953; *Gupta,* 1975]. In a series of seminal papers *Vakman* [1972, 1976, 1980] has addressed the concepts of instantaneous frequency and the analytic signal and has brought forth the fundamental issues regarding these subjects. The review article on the subject is given by *Vakman and Vainshtein* [1978]

3.2.1 Energy of the Analytic Signal

Because we have insisted that the real part of the complex signal be the original signal, normalization is not preserved. Recall that the spectrum of the original real signal satisfies $|S(\omega)| = |S(-\omega)|$ and therefore the energy of the original signal is [*Cohen,* 1995]

$$E_s = \int |S(\omega)|^2 \, d\omega = 2 \int_0^\infty |S(\omega)|^2 \, d\omega = \frac{1}{2} \int_0^\infty |2S(\omega)|^2 \, d\omega = \frac{1}{2} E_z \tag{94}$$

That is, the energy of the analytic signal is twice the energy of the original signal. In addition, the energy of the real part is equal to the energy of the imaginary part

$$E_s = E_{H[s]} \tag{95}$$

which can be seen by considering $|z(t)|^2 = |s(t) + j\, H[s]|^2$. When this is expanded the middle term is [*Cohen,* 1995]

$$\int \int \frac{s^\star(t)\, s(t') + s(t)\, s^\star(t')}{t - t'} \, dt' \, dt = 0 \tag{96}$$

since the integrand is a two dimensional odd function.

3.2.2 Paradoxes Regarding the Analytic Signal

There are five paradoxes or difficulties regarding the notion of instantaneous frequency if it is defined as the derivative of the phase of the analytic signal [*Cohen,* 1989, 1995]. First, instantaneous frequency may not be one of the frequencies in the spectrum. That is strange because if instantaneous frequency is an indication of the frequencies that exist at each time, how can it not exist when we do the final bookkeeping by way of the spectrum? Second, if we have a line spectrum consisting of only a few sharp frequencies, then the instantaneous frequency may be continuous and range over an infinite number of values. Third, although the spectrum of the analytic signal is zero for negative frequencies, the instantaneous frequency may be negative. Fourth, for a bandlimited signal the instantaneous frequency may go outside the band. All these points are illustrated by the following simple example.

Example: Instantaneous Frequency for the Sum of Two Sinusoids

Consider

$$\begin{aligned} s(t) &= s_1(t) + s_2(t) && (97) \\ &= A_1\, e^{j\omega_1 t} + A_1\, e^{j\omega_1 t} && (98) \\ &= A(t)\, e^{j\varphi(t)} && (99) \end{aligned}$$

where the amplitudes A_1 and A_2 are taken to be constants and ω_1 and ω_2 are positive. The spectrum of this signal consists of two delta functions at ω_1 and ω_2,

$$S(\omega) = A_1\, \delta(\omega - \omega_1) + A_2\, \delta(\omega - \omega_2) \tag{100}$$

Since we take ω_1 and ω_2 to be positive, the signal is analytic. Solving for the phase and amplitude,

$$\varphi(t) = \arctan\left(\frac{A_1 \sin\omega_1 t + A_2 \sin\omega_2 t}{A_1 \cos\omega_1 t + A_2 \cos\omega_2 t}\right) \tag{101}$$

$$A^2(t) = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\omega_2 - \omega_1)t \tag{102}$$

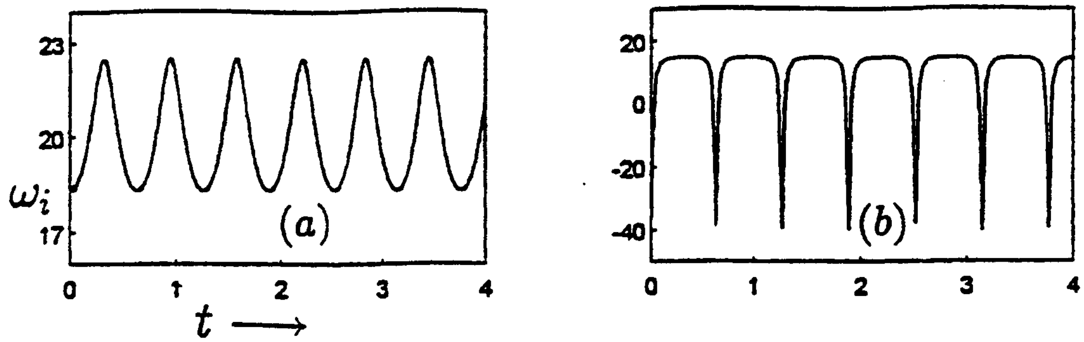

Figure 1: The instantaneous frequency for the signal $s(t) = A_1 e^{j10t} + A_2 e^{j20t}$. The spectrum consists of two frequencies, at $\omega = 10$ and $\omega = 20$. In (a) $A_1 = 0.2$ and $A_2 = 1$. The instantaneous frequency is continuous and ranges outside the bandwidth. In (b) $A_1 = -1.2$ and $A_2 = 1$. Although the signal is analytic the instantaneous frequency may become negative [after Cohen, 1995].

and taking the derivative of the phase we obtain

$$w_i = \varphi'(t) = \frac{1}{2}(\omega_2 + \omega_1) + \frac{1}{2}(\omega_2 - \omega_1)\frac{A_2^2 - A_1^2}{A^2(t)} \tag{103}$$

By taking different values of the amplitudes and frequency we can illustrate the points above. This is done in Figure 1.

One last paradox regarding the analytic signal needs to be discussed. If instantaneous frequency is an indication of the frequencies that exist at time t, one would presume that what the signal did a long time ago and is going to do in the future should be of no concern; only the present should count. However, to calculate the analytic signal at time t we have to know the signal for all time. This paradox has been analyzed by *Vakman* [1976]. A more detail discussion about the instantaneous frequency and the analytical signal can be found in *Vakman* [1976].

4 The Uncertainty Principle

4.1 Introduction

The uncertainty principle states that a signal cannot be arbitrarily concentrated with respect to both time and frequency [*Papoulis,* 1977; *De Bruijn,* 1967; *Folland and Sitaram,* 1997] The basic consequence of the Fourier Transform relation is that a narrow waveform yields a wide spectrum and a wide waveform yields a narrow spectrum and both the time waveform and frequency spectrum cannot be made arbitrarily small simultaneously. This observation is known as the uncertainty principle of signal analysis. The uncertainty principle is a fundamental statement regarding Fourier Transform pairs. A graphical illustration of the uncertainty relation between the time and frequency representations of a signal is shown in Figure 2.

The uncertainty principle expresses a fundamental relationship between the standard deviation of a function and the standard deviation of its Fourier Transform. The standard deviations of the time and frequency density functions, σ_t and σ_ω, are defined as the parameters that describes the broadness of the signal in time and frequency domains respectively. For convenience we repeat the definitions here:

$$T^2 = \sigma_t^2 = \int (t- < t >)^2 \mid s(t) \mid^2 \, dt \tag{104}$$

$$B^2 = \sigma_\omega^2 = \int (\omega- < \omega >)^2 \mid S(\omega) \mid^2 \, d\omega \tag{105}$$

With these two measures, T and B, for the broadness of a signal, the uncertainty relation can be expressed as [*Gabor*, 1946]

$$TB \geq \frac{1}{2} \tag{106}$$

We refer to the inequality equation (106) as Gabor's relation, in order to distinguish it from Heisenburg's uncertainty relation of quantum mechanics. Although the two relations are the same in their mathematical expression, they relate to different physical concepts. In quantum mechanics the uncertainty relation emerges in a probabilistic context, while in

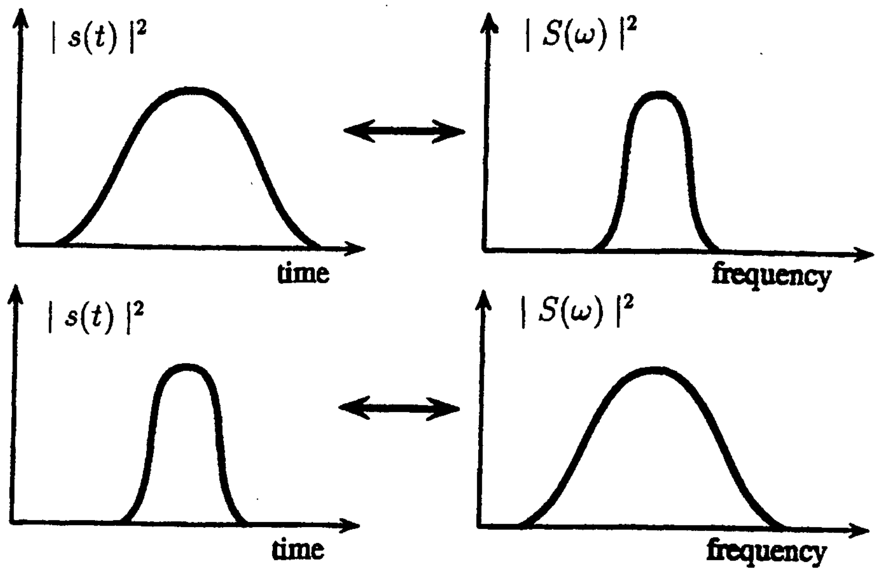

Figure 2: The uncertainty relation for time and frequency marginals. A broad waveform in time gives a narrow frequency spectrum and vice versa.

signal analysis it is an expression of the simple fact that one cannot make T and B arbitrarily small. An excellent discussion of the uncertainty principle in signal analysis and its relation to quantum mechanical concepts can be found in *Cohen* [1989].

4.1.1 A More General Uncertainty Principle

A stronger version of the uncertainty principle is [*Cohen,* 1995]

$$\sigma_t\,\sigma_\omega \geq \frac{1}{2}\sqrt{1 + 4\,\mathrm{Cov}_{t\omega}^2} \tag{107}$$

where $\mathrm{Cov}_{t\omega}$ is the covariance as defined by equation (66).

4.2 Proof of the Uncertainty Principle

First, let us note that no loss of generality occurs if we take signals that have zero mean time and zero mean frequency. The reason is that the standard deviation does not depend on the mean because it is defined as the broadness about the mean. If we have a signal s_{old}, then a new signal defined by [*Cohen,* 1995]

$$s_{\text{new}}(t) = e^{-j<\omega>(t+<t>)}\, s_{\text{old}}(t+ < t >) \tag{108}$$

has the same shape both in time and frequency as s_{old} except that it has been translated in time and frequency so that the means are zero. Conversely, if we have a signal $s_{\text{new}}(t)$ that has zero mean time and zero mean frequency and we want a signal of the same shape but with particular mean time and frequency, then

$$s_{\text{old}}(t) = e^{-j<\omega>t}\, s_{\text{new}}(t- < t >) \tag{109}$$

The bandwidth expressed in terms of the signal is as per equation (44):

$$\sigma_\omega^2 = \int \omega^2 \mid S(\omega) \mid^2 d\omega = \int \mid s^{'}(t) \mid^2 dt \tag{110}$$

The duration is

$$\sigma_t^2 = \int t^2 \mid s(t) \mid^2 dt \tag{111}$$

and therefore

$$\sigma_t^2 \sigma_\omega^2 = \int \mid t\, s(t) \mid^2 dt \times \int \mid s^{'}(t) \mid^2 dt \tag{112}$$

It should be noted that no other assumptions or ideas are used in equation (112). The fact that s and S are Fourier Transform pairs is reflected in equation (110).

Now, for any two functions (not only Fourier Transform pairs)

$$\int \mid f(x) \mid^2 \, dx \int \mid g(x) \mid^2 \, dx \geq \left| \int f^\star(x) \, g(x) \, dx \right|^2 \tag{113}$$

which is commonly known as the Schwarz inequality [*Cohen,* 1995]. Taking $f = t\,s$ and $g = s'$ gives

$$\sigma_t^2 \sigma_\omega^2 \geq \left| \int t\, s^\star(t)\, s'(t) dt \right|^2 \tag{114}$$

There are many proof of the inequality. A simple one is to note that for any two functions

$$\int \mid f(x) \mid^2 \, dx \int \mid g(x) \mid^2 \, dx - \left| \int f^\star(x) \, g(x) \, dx \right|^2 = \frac{1}{2} \int \int |f(x)\, g(y) - f(y)\, g(x)|^2 \, dx \, dy \tag{115}$$

which is readily verified by direct expansion of the right hand side. Since the right hand side is manifestly positive, we have equation (113).

The integrand, written in terms of amplitude and phase, is [*Cohen,* 1995]

$$t\, s^\star(t)\, s'(t) = t\, A'\, A + j\, t\, \varphi'\, A^2 \tag{116}$$

$$= \frac{1}{2} \frac{d}{dt} t\, A^2 - \frac{1}{2} A^2 + j\, t\, \varphi'(t) \tag{117}$$

The first term is a perfect differential and integrates to zero. The second term gives one half since we assume the signal is normalized and the third term gives j times the covariance of the signal. Hence [*Cohen,* 1995]

$$\sigma_t^2 \sigma_\omega^2 \geq \left| \int t\, s^\star(t)\, s'(t) dt \right|^2 = | -\frac{1}{2} + j\, \mathrm{Cov}_{t\omega} |^2 = \frac{1}{4} + \mathrm{Cov}_{t\omega}^2 \tag{118}$$

Therefore we have the uncertainty principles as given by equation (107). Since $\mathrm{Cov}_{t\omega}^2$ is always positive, it can, if we choose so, be dropped to obtain the more usual form, equation (106).

It is important to note that the proof depends on only four things: first, on $| s(t) |^2$ being the density in time; second, on taking $| S(\omega) |^2$ as the density in frequency; third that $s(t)$ and $S(\omega)$ are Fourier pairs; and fourth, on defining T and B as standard deviations of time and frequency.

It should also be noted from the preceding steps that the uncertainty principle is calculated *only* from the marginals (see Section 5.2). Hence any joint distribution that yields the *marginals* will give the uncertainty principle. It has nothing to do with correlations between the time and frequency or the measurement for small times and frequencies. What does it say is that the marginals are functionally dependent. But the fact that marginals are related does not imply correlation between the variables and has nothing to do with the existence or nonexistence of a joint distribution [*Cohen,* 1989; *Qian and Chen,* 1996; *Hlawatsch,* 1998; *Mertins,* 1999].

4.3 The Uncertainty Principle for the Short-Time Fourier Transform

There are many things one can do to signals to study them. However, if we do something to a signal that modifies it in some way, one should not confuse the uncertainty principle applied to the modified signal with the uncertainty principle as applied to the original signal [*Cohen,* 1995]. One of the methods used to estimate properties of a signal is to take only a small piece of the signal around the time of interest and study that piece while neglecting the rest of the signal. In particular, we can take the Fourier Transform of the small piece of the signal to estimate the frequencies at that time. If we make the time interval around the time t small, we will have a very high bandwidth. This statement applies to the modified signal, that is, to the short interval that we have artificially constructed for the purpose of analysis. The process of chopping up a signal for the purpose of analysis is called the Short-Time Fourier Transform procedure. Although we will be studying the short-Time Fourier Transform in Section 6, this is an appropriate place to consider the uncertainty principle for it.

From the original signal $s(t)$ one defines a short duration signal around the time of interest, t, by multiplying it by a window function that is peaked around the time, t, and falls off rapidly. This has the effect of emphasizing the signal at time, t, and suppressing it for times far away from that time. In particular, we define the normalized short duration signal at time, t, by [*Cohen,* 1995]

$$\eta_t(\tau) = \frac{s(\tau)\, h(\tau - t)}{\sqrt{\int \, |s(\tau)\, h(\tau - t)|^2 \, d\tau}} \tag{119}$$

where $h(t)$ is the window function, t is the fixed time for which we are interested, and τ is now the running time. This normalization ensures that

$$\int \, |\eta_t(\tau)|^2 \, d\tau = 1 \tag{120}$$

for any t. Now $\eta_t(\tau)$ as a function of the time τ is of short duration since presumably we have chosen a window function to make it so. The time, t, acts as a parameter. The Fourier Transform of the small piece of the signal, the modified signal, is

$$F_t(\omega) = \frac{1}{\sqrt{2\pi}} \int \, e^{-j\omega\tau} \, \eta_t(\tau) \, d\tau \tag{121}$$

$F_t(\omega)$ gives us an indication of the spectral content at the time t. For the modified signal we can define all the relevant quantities such as mean time, duration, and bandwidth in the standard way, but they will be time dependent. The mean time and duration for the modified signal are [*Cohen,* 1995]

$$< \tau >_t = \int \, \tau |\eta_t(\tau)|^2 \, d\tau = \frac{\int \, \tau \, |s(\tau)\, h(\tau - t)|^2 \, d\tau}{\int \, |s(\tau)\, h(\tau - t)|^2 \, d\tau} \tag{122}$$

$$T_t^2 = \int \, (\tau - < \tau >)^2 \, |\eta_t(\tau)|^2 \, d\tau = \frac{\int \, (\tau - < \tau >_t)^2 \, |s(\tau)\, h(\tau - t)|^2 \, d\tau}{\int \, |s(\tau)\, h(\tau - t)|^2 \, d\tau} \tag{123}$$

Similarly, the mean frequency and bandwidth for the modified signal are [*Cohen,* 1995]

$$< \omega >_t = \int \, \omega \, |F_t(\omega)|^2 \, d\omega = \int \, \eta_t^*(\tau) \, \frac{1}{j} \, \frac{d}{d\tau} \, \eta_t(\tau) \, d\tau \tag{124}$$

$$B_t^2 = \int \, (\omega - < \omega >_t)^2 \, |F_t(\omega)|^2 \, d\omega \tag{125}$$

4.3.1 Time-Dependent and Window-Dependent Uncertainty Principle

Since we have used a normalized signal to calculate the duration and bandwidth, we can immediately write that [*Cohen,* 1995]

$$B_t T_t \geq \frac{1}{2} \tag{126}$$

This is the uncertainty principle for the Short-Time Fourier Transform. It is a function of time, the signal, and the window. It should not be confused with the uncertainty principle applied to the signal. It is important to understand this uncertainty principle because it places limits on the technique of the Short-Time Fourier Transform procedure. However, it places no constraints on the original signal. It is not possible to have arbitrarily good time resolution simultaneously with good frequency resolution. A long time window gives poor time resolution but relatively good frequency resolution. A large bandwidth window gives poor frequency resolution but relatively good time resolution.

5 Time-Frequency Distributions: Fundamental Ideas

5.1 introduction

The basic objective of time-frequency analysis is to devise a function that will describe the energy density of a signal simultaneously in time and frequency, and that can be used and manipulated in the same manner as any density. We now begin our study of how to construct such distributions and in this section we describe the main ideas. The discussion in this section mainly follows that of *Moyal* [1949], *Papoulis* [1977], *Cohen* [1989, 1995], *Rioul and Vetterli* [1991], *Hlawatsch and Boudreaux-Bartels*, *Qian and Chen* [1996], *Hlawatsch* [1998], *Mertins* [1999], in addition to the references given. To crystallize our aim we recall that the instantaneous power or intensity in time is

$$\begin{aligned} |s(t)|^2 &= \text{instantaneous power or energy density at time } t\text{, or} \\ |s(t)|^2 \, \Delta t &= \text{the fractional energy in the time interval } \Delta t \text{ at time } t \end{aligned}$$

and the density in frequency, the energy density spectrum, is

$$\begin{aligned} |S(\omega)|^2 &= \text{energy spectrum density at } \omega\text{, or} \\ |S(\omega)|^2 \, \Delta\omega &= \text{the fractional energy in the frequency interval } \Delta\omega \text{ at frequency } \omega \end{aligned}$$

What we seek is a joint density, $P(t, \omega)$, so that

$$\begin{aligned} P(t,\omega) &= \text{the intensity at time } t \text{ and frequency } \omega\text{, or} \\ P(t,\omega)\, \Delta t \, \Delta\omega &= \text{the fractional energy in the time-frequency cell } \Delta t \, \Delta\omega \text{ at } t, \omega \end{aligned}$$

5.2 Marginals

Summing up the energy distribution for all frequencies at a particular time should give the instantaneous energy, and summing up over all times at a particular frequency should give the energy density spectrum. Therefore, ideally, a joint density in time and frequency should satisfy [*Papoulis,* 1977; *Cohen,* 1989, 1995; *Hlawatsch and Boudreaux-bartels,* 1992; *Qian and Chen,* 1996; *Mertins,* 1999]

$$\int P(t,\omega)\, d\omega = |s(t)|^2 \tag{127}$$

$$\int P(t,\omega)\, dt = |S(\omega)|^2 \tag{128}$$

which are called the time and frequency marginal conditions.

5.3 Total Energy

The total energy of the distribution should be the total energy of the signal [*Papoulis,* 1977; *Cohen,* 1989, 1995; *Hlawatsch and Boudreaux-Bartels,* 1992; *Qian and Chen,* 1996; *Mertins,* 1999]

$$E = \int\int P(t,\omega)\, d\omega\, dt = \int |s(t)|^2\, dt = \int |S(\omega)|^2\, d\omega \tag{129}$$

Note that if the joint density satisfies the marginals, it automatically satisfies the total energy requirement, but the converse is not true. It is possible that a joint density can satisfy the total energy requirement without satisfying the marginals. The spectrogram that we study in the next section is one such example. The total energy requirement is a weak one and that is why many distributions that do not satisfy it may nonetheless give a good representation of the time-frequency structure.

5.4 Characteristic Functions

5.4.1 One-Dimensional

The characteristic function is a powerful tool for the study and construction of densities. Suppose $P(x)$ is a one-dimensional density of the quantity x. The characteristic function is the Fourier Transform of the density [*Moyal,* 1949; *Cohen,* 1966, 1989, 1995]

$$M(\theta) = \int e^{j\theta x} P(x)\, dx =< e^{j\theta x} > \tag{130}$$

The characteristic function is the average of $e^{j\theta x}$, where θ is a parameter. By expanding the exponential we have

$$M(\theta) = \int e^{j\theta x} P(x)\, dx = \int \sum_{i=1}^{\infty} \frac{(j\theta x)^n}{n!} P(x)\, dx = \sum_{i=1}^{\infty} \frac{j^n \theta^n}{n!} < x^n > \tag{131}$$

which is a Taylor series in θ with coefficients $j^n < x^n >$. Since the coefficients of a Taylor are given by the nth derivative of the function evaluated at zero, we have [*Cohen,* 1995]

$$< x^n >= \frac{1}{j^n} \frac{\partial^n M(\theta)}{\partial \theta^n} |_{\theta=0} \tag{132}$$

The fact that moments can be calculated by differentiation rather than by integration is one the advantages of the characteristic function, since differentiation is always easier than integration. Of course, one has to find the characteristic function and that may be hard.

Fourier Transform pairs are uniquely related and hence the characteristic function determines the distribution [*Cohen,* 1989, 1995],

$$P(x) = \frac{1}{2\pi} \int M(\theta)\, e^{-j\theta x}\, d\theta \tag{133}$$

Some general properties that a function must posses if it is a proper characteristic function are easily obtained. By proper we mean a characteristic function that comes from a normalized positive density [*Cohen,* 1995]. First, taking $\theta = 0$ we see that

$$M(0) = \int P(x)\, dx = 1 \tag{134}$$

Taking the complex conjugate of equation (130) and using the fact that densities are real, we have

$$M^{\star}(\theta) = \int e^{-j\theta x} P^{\star}(x)\, dx = M(-\theta) \tag{135}$$

or

$$M^{\star}(-\theta) = M(\theta) \tag{136}$$

The absolute value of the characteristic function is always less than or equal to one [*Cohen,* 1995],

$$| M(\theta) | \leq 1 \tag{137}$$

This follows from

$$| M(\theta) | = \int e^{j\theta x} P(x)\, dx \leq \int | e^{j\theta x} | \, | P(x) | \; dx = \int P(x)\, dx = 1 \tag{138}$$

We know that the characteristic function at the origin is equal to one and therefore

$$| M(\theta) | \leq M(0) \tag{139}$$

5.4.2 Two-Dimensional

The two-dimensional characteristic function $M(\theta, \tau)$, is the average of $e^{j\theta x + j\tau y}$, [*Moyal,* 1949; *Cohen,* 1966, 1989, 1995]

$$M(\theta, \tau) = < e^{j\theta x + j\tau y} > = \int\int e^{j\theta x + j\tau y} P(x, y)\, dx\, dy \tag{140}$$

and the distribution function may be obtained from $M(\theta, \tau)$ by Fourier inversion,

$$P(x, y) = \frac{1}{4\pi^2} \int\int M(\theta, \tau)\, e^{-j\theta x - j\tau y}\, d\theta\, d\tau \tag{141}$$

Therefore the joint characteristic function of a time-frequency density is given by [*Cohen,* 1995]

$$M(\theta,\tau) =< e^{j\theta t+j\tau\omega} > = \int\int P(t,\omega)\, e^{j\theta t+j\tau\omega}\, dt\, d\omega \tag{142}$$

5.5 Global Averages

The average value of any function of time and frequency is to be calculated in the standard way [*Moyal, 1949; Cohen,* 1966, 1995, 1999]

$$< g(t,\omega) > = \int\int g(t,\omega)\, P(t,\omega)\, dt\, d\omega \tag{143}$$

If the marginals are satisfied then we are guaranteed that averages of the form

$$\begin{aligned} < g_1(t) + g_2(\omega) > &= \int\int \{g_1(t) + g_2(\omega)\}\, P(t,\omega)\, dt\, d\omega && (144) \\ &= \int g_1(t)\, |s(t)|^2\, dt + \int g_2(\omega)\, |S(\omega)|^2\, d\omega && (145) \end{aligned}$$

will be correctly calculated since the calculation requires only the satisfaction of the marginals.

5.6 Time and Frequency Shift Invariance

Suppose we have a signal $s(t)$ and another signal that is identical to it but translated in time by t_0. We want the distribution corresponding to each signal to be identical in form, but that the one corresponding to the time shifted signal be translated by t_0. That is [*Papoulis,* 1977; *Cohen,* 1989, 1995; *Qian and Chen,* 1996; *Mertins,* 1999]

$$\text{if} \quad s(t) \rightarrow s(t-t_0) \quad \text{then} \quad P(t,\omega) \rightarrow P(t-to,\omega) \tag{146}$$

Similarly, if we shift the spectrum by a constant frequency we expect the distribution to be shifted by that frequency,

$$\text{if} \quad S(\omega) \rightarrow S(\omega - \omega_0) \quad \text{then} \quad P(t,\omega) \rightarrow P(t, \omega - \omega_0) \tag{147}$$

Both of these cases can be handled together. If $s(t)$ is the signal, then a signal that is translated in time by t_0 and translated in frequency by ω_0 is given by $e^{j\omega_0 t}\, s(t - t_0)$. Accordingly, we expect the distribution to be shifted in time and frequency in the same way,

$$\text{if} \quad s(t) \rightarrow e^{j\omega_0 t}\, s(t - t_0) \quad \text{then} \quad P(t,\omega) \rightarrow P(t - t_0, \omega - \omega_0) \tag{148}$$

5.7 Weak and Strong Finite Support

Suppose a signal doesn't start until t_1. We want the joint distribution also to not start until t_1. Similarly, if the signal stops after time t_2 we expect the distribution to be zero after that time. If that is the case we say the distribution has weak finite time support. The reason for the word weak will be apparent shortly. Similarly, if the spectrum is zero outside a frequency band, then the distribution should also be zero outside the band. In such a case we say that the distribution has weak finite spectral support. We can express these requirements mathematically as [*Cohen,* 1989, 1995; *Qian and Chen,* 1996]

$$P(t,\omega) \;=\; 0 \quad \text{for } t \text{ outside } (t_1, t_2) \qquad \text{if } s(t) \text{ is zero outside } (t_1, t_2) \tag{149}$$

$$P(t,\omega) \;=\; 0 \quad \text{for } \omega \text{ outside } (\omega_1, \omega_2) \qquad \text{if } S(\omega) \text{ is zero outside } (\omega_1, \omega_2) \tag{150}$$

Now suppose we have a signal that stops for a half hour and then starts again. We would expect the distribution to be zero for that half hour. Similarly, if we have a gap in the spectrum, then we expect the distribution to be zero in that gap. If a distribution satisfies these requirements, namely that it is zero whenever the signal is zero or is zero whenever the spectrum is zero, then we say the distribution has strong finite support:

$$P(t,\omega) \;=\; 0 \qquad \text{if } s(t) \text{ is zero for a particular time} \tag{151}$$

$$P(t,\omega) \;=\; 0 \qquad \text{if } S(\omega) \text{ is zero for a particular frequency} \tag{152}$$

Strong finite support implies weak finite support, but not conversely.

5.7.1 Distributions Concentrated in a Finite Region

A signal cannot be both of finite duration and bandlimited in frequency [*Cohen,* 1995]. Therefore if a distribution satisfies the weak finite support property it cannot be limited to a finite region of the time-frequency plane. If it were, it would be both time and frequency limited, which is impossible. If it turns out that a distribution is limited in a finite region, then it does not satisfy the finite support properties and/or the marginals.

5.8 Uncertainty Principle

In Section 4 we emphasized that the uncertainty principle depends on only three statements. First and second are that the time and frequency standard deviations are calculated using $|s(t)|^2$ and $|S(\omega)|^2$ as the respective densities,

$$T^2 \;=\; \int (t- < t >)^2\, |s(t)|^2\, dt \tag{153}$$

$$B^2 \;=\; \int (\omega- < \omega >)^2\, |S(\omega)|^2\, d\omega \tag{154}$$

and the third is that $s(t)$ and $S(\omega)$ are Fourier Transform pairs. From a joint distribution the standard deviations are obtained by

$$\sigma_t^2 \;=\; \int\!\!\int (t- < t >)^2\, P(t,\omega)\, dt\, d\omega = \int (t- < t >)^2\, P(t)\, dt \tag{155}$$

$$\sigma_\omega^2 \;=\; \int\!\!\int (\omega- < \omega >)^2\, P(t,\omega)\, dt\, d\omega = \int (\omega- < \omega >)^2\, P(\omega)\, d\omega \tag{156}$$

When the standard deviations calculated using the joint distribution give the same answer as when calculated by equations (153)-(154), we will get the correct uncertainty principle. This will be the case when the marginals are correctly given,

$$P(t) = |s(t)|^2 \qquad \text{and} \qquad P(\omega) = |S(\omega)|^2 \qquad \text{for uncertainty principle} \tag{157}$$

Therefore, any joint distribution that yields the correct marginals will yield, and is totally consistent with, the uncertainty principle [*Cohen,* 1995].

6 The Short-Time Fourier Transform

6.1 Introduction

The Short-Time Fourier Transform (STFT) is the most widely used method for studying non-stationary signals [*Dziewonski et al,* 1969; *Oppenheim,* 1970, 1975; *Levshin,* 1972; *Kodera et al.,* 1976; *Allen and Rabiner,* 1977; *Altes,* 1980; *Portnoff,* 1980; *Cohen,* 1989, 1995; *Qian and Chen,* 1996; *Hlawatsch,* 1998; *Mertins,* 1999]. The concept behind it is simple and powerful. Suppose we listen to a piece of music that lasts an hour where in the beginning there are violins and at the end drums. If we Fourier analyze the whole hour, the energy spectrum will show peaks at the frequencies corresponding to the violins and drums. That will tell us that there were violins and drums but will not give us any indication of when the violins and drums were played. The most straightforward thing to do is to break up the hour into five minute segments and Fourier analyze each interval. Upon examining the spectrum of each segment we will see in which five minute intervals the violins and drums occurred. If we want to localize even better, we break up the hour into one minute segments or even smaller time intervals and Fourier analyze each segment. That is the basic idea of the Short-Time Fourier Transform: break up the signal into small time segments and Fourier analyze each time segment to ascertain the frequencies that existed in that segment. The totality of such spectra indicates how the spectrum is varying in time.

Can this process be continued to achieve finer and finer time localization? Can we make the time intervals as short as we want? The answer is no, because after a certain narrowing the answers we get for the spectrum become meaningless and show no relation to the spectrum of the original signal. The reason is that we have taken a perfectly good signal and broken it up into short duration signals. But short duration signals have inherently large bandwidths, and the spectra of such short duration signals have very little to do with the properties of the original signal. This should be attributed not to any fundamental limitation, but rather to a limitation of the technique which makes short duration signals for the purpose of estimating the spectrum. Sometimes this technique works well and sometimes it does not. It is not the uncertainty principle as applied to the signal that is the limiting factor; it is the uncertainty principle as applied to the small time intervals that we have created for the purpose of analysis. The distinction between the uncertainty principle for the small time intervals created for analysis and the uncertainty principle for the original signal should be clearly noted and the two should not be confused.

It should be always noted that in the Short-Time Fourier Transform the properties of the signal are scrambled with the properties of the window function, the window function being the means of chopping up the signal. Unscrambling is required for proper interpretation and estimation of the original signal.

The above difficulties notwithstanding, the Short-Time Fourier Transform method is ideal in many respects. It is well defined, based on reasonable physical principles, and for many signals and situations it gives an excellent time-frequency structure consistent with our intuition. However, for certain situations it may not be the best method available in the sense that it does not always give us the clearest possible picture of what is going on. Thus other methods have been developed, which are discussed in subsequent sections.

6.2 The Short-Time Fourier Transform and Spectrogram

To study the properties of the signal at time t, one emphasizes the signal at that time and suppresses the signal at other times. This is achieved by multiplying the signal by a window function, $h(t)$, centered at t, to produce a modified signal [*Cohen,* 1995],

$$s_t(\tau) = s(\tau)\, h(\tau - t) \tag{158}$$

The modified signal is a function of two times, the fixed time we are interested in, t, and the running time, τ. The window function is chosen to leave the signal more or less unaltered around the time t but to suppress the signal for times distant from the time of interest. That is,

$$s_t(\tau) \sim \left\{ \begin{array}{ll} s(\tau) & \text{for } \tau \text{ near } t \\ 0 & \text{for } \tau \text{ far away from } t \end{array} \right\} \tag{159}$$

The term "window" comes from the idea that we are seeking to look at only a small piece of the signal as when we look out of a real window and see only a relatively small portion of the scenery. In this case we want to see only a small portion.

Since the modified signal emphasizes the signal around the time t, the Fourier Transform will reflect the distribution of frequency around that time,

$$S_t(\omega) = \frac{1}{2\pi} \int e^{-j\omega\tau} \, s_t(\tau) \, d\tau \tag{160}$$

$$= \frac{1}{2\pi} \int e^{-j\omega\tau} \, s(\tau) \, h(\tau - t) \, d\tau \tag{161}$$

The energy density spectrum at time t is therefore [*Schroeder and Atal,* 1962; *Kodera et al.,* 1976; *Portnoff,* 1980; *Rabiner and Allen,* 1980; *Crochiere and Rabiner,* 1983; *Nawab and Quatieri,* 1988; *Cohen,* 1989, 1995]

$$P_{SP}(t,\omega) = |S_t(\omega)|^2 = \left| \frac{1}{\sqrt{2\pi}} \int e^{-j\omega\tau} \, s(\tau) \, h(\tau - t) \, d\tau \right|^2 \tag{162}$$

For each different time we get a different spectrum and the totality of these spectra is the time-frequency distribution, P_{SP}. It goes under many names, depending on the field; we shall use the most common phraseology, "spectrogram."

Since we are interested in analyzing the signal around the time t, we presumably have chosen a window function that is peaked around t. Hence the modified signal is short and its Fourier Transform, equation (161), is called the Short-Time Fourier Transform. However, it should be emphasized that often we will not be taking narrow windows - which is done when we want to estimate time properties for a particular frequency. When we want to estimate time properties for a given frequency we do not take short times but long ones, in which case the Short-Time Fourier Transform may be appropriately called the Long-Time Fourier Transform or the Short-Frequency Time Transform.

6.2.1 Characteristic Function

The characteristic function of the spectrogram is straight forwardly obtained [*Moyal,* 1949, *Cohen,* 1966, 1989, 1995],

$$M_{SP}(\theta,\tau) = \int\!\!\int |S_t(\omega)|^2 \, e^{j\theta t + j\tau\omega} \, dt \, d\omega \tag{163}$$

$$= A_s(\theta,\tau) \, A_h(-\theta,\tau) \tag{164}$$

where

$$A_s(\theta,\tau) = \int s^\star(t-\frac{1}{2}\tau)\, s(t+\frac{1}{2}\tau)\, e^{j\theta t}\, dt \tag{165}$$

is the ambiguity function of the signal, and A_h is the ambiguity function of the window defined in the identical manner, except that we use $h(t)$ instead of $s(t)$. Note that $A(-\theta,\tau) = A^\star(\theta,-\tau)$, a relation we will use later.

6.3 General Properties

6.3.1 Total Energy

The total energy is obtained by integrating over all time and frequency. However, we know that it is given by the characteristic function evaluated at zero (see Section 5.4). Using equations (164) and (165) we have

$$E_{SP} = \int\int P_{SP}(t,\omega)\, dt\, d\omega = M_{SP}(0,0) \tag{166}$$

$$= A_s(0,0)\, A_h(0,0) \tag{167}$$

$$= \int |s(t)|^2\, dt \times \int |h(t)|^2\, dt \tag{168}$$

Therefore, we see that if the energy of the window is taken to be one, then the energy of the spectrogram is equal to the energy of the signal.

6.3.2 Marginals

The time marginal is obtained by integrating over frequency [*Cohen,* 1995],

$$P(t) = \int |S_t(\omega)|^2\, d\omega \tag{169}$$

$$= \frac{1}{2\pi} \int s(\tau)\, h(\tau - t)\, s^\star(\tau')\, h^\star(\tau' - t)\, e^{-j\omega(\tau-\tau')}\, d\tau\, d\tau'\, d\omega \tag{170}$$

$$= \int s(\tau)\, h(\tau - t)\, s^\star(\tau')\, h^\star(\tau' - t)\, \delta(\tau - \tau')\, d\tau\, d\tau' \tag{171}$$

$$= \int |s(\tau)|^2\, |h(\tau - t)|^2\, d\tau \tag{172}$$

$$= \int A^2(\tau)\, A_h^2(\tau - t)\, d\tau \tag{173}$$

Similarly, the frequency marginal is

$$P(\omega) = \int B^2(\omega')\, B_H^2(\omega - \omega')\, d\omega' \tag{174}$$

As can be seen from these equations, the marginals of the spectrogram generally do not satisfy the correct marginals, namely $|s(t)|^2$ and $|S(\omega)|^2$,

$$P(t) \neq A^2(t) = |s(t)|^2 \tag{175}$$

$$P(\omega) \neq B^2(\omega) = |S(\omega)|^2 \tag{176}$$

The reason is that the spectrogram scrambles the energy distributions of the window with those of the signal. This introduces effects unrelated to the properties of the original signal.

Notice that the time marginal of the spectrogram depends only on the magnitude of the signal and window and not on their phases. Similarly, the frequency marginal depends only on the amplitudes of the Fourier Transforms.

6.3.3 Averages of Time and Frequency Functions

Since the marginals are not satisfied, averages of time and frequency functions will never be correctly given [*Cohen,* 1995],

$$< g_1(t) + g_2(t) > \; = \; \int\int \{g_1(t) + g_2(\omega)\} \, P_{SP}(t,\omega) \, d\omega \, dt \tag{177}$$

$$< g_1(t) + g_2(t) > \; \neq \; \int g_1(t) \, |s(t)|^2 \, dt + \int g_2(\omega) \, |S(\omega)|^2 \, d\omega \tag{178}$$

This is in contrast to other distributions we will be studying where these types of averages are always correctly given.

6.3.4 Finite Support

Recall from our discussion in Section 5.7 that for a finite duration signal we expect the distribution to be zero before the signal starts and after it ends. This property was called the finite support property. The spectrogram does not satisfy this property, because the modified signal as a function of t will not necessarily be zero since the window may pick up some of the signal. That is, even though $s(t)$ may be zero for a time t, $s(\tau)\, h(\tau - t)$ may not be zero for that time. This will always be the case for windows that are not time limited. But even if a window is time limited we will still have this effect for time values that are close to the beginning or end of the signal. Similar considerations apply to the frequency domain. Therefore the spectrogram does not possess the finite support property in either time or frequency.

6.3.5 Localization Trade-Off

If we want good time localization we have to pick a narrow window in the time domain, $h(t)$, and if we want good frequency localization we have to pick a narrow window, $H(\omega)$, in the frequency domain. But both $h(t)$ and $H(\omega)$ cannot be made arbitrarily narrow; hence there is an inherent trade-off between time and frequency localization in the spectrogram for a particular window. The degree of trade-off depends on the window, signal, time, and frequency. The uncertainty principle for the spectrogram quantifies these trade off dependencies, as we discussed in Section 5.8.

6.3.6 One Window or Many ?

We have just seen that one window, in general, cannot give good time and frequency localization. That should not cause any problem of principle as long as we look at the spectrogram as a tool at our disposal that has many options including the choice of window. There is no reason why we cannot change the window depending on what we want to study. That can sometimes be done effectively, but not always. Sometimes a compromise window does very well. One of the advantages of other distributions that we will be studying is that both time and frequency localization can be done concurrently.

6.3.7 Entanglement and Symmetry Between Window and Signal

The results obtained using the spectrogram generally do not give results regarding the signal solely, because the Short-Time Fourier Transform entangles the signal and window. Therefore we must be cautious in interpreting the results and we must attempt to disentangle the window. That is not always easy. In fact, because of the basic symmetry in the definition of the Short-Time Fourier Transform between the window and signal, we have to be careful that we are not using the signal to study the window.

7 The Wigner Distribution

7.1 Introduction

The representations that describe a signal's frequency behavior fall predominantly into two categories: linear representations such as the Fourier Transform and quadratic representations such as the power spectrum. Previously we described the linear joint time-frequency representations, the Short-Time Fourier Transform. In this section, we will introduce the counterpart to the power spectrum: the quadratic, or bilinear, joint time-frequency representation. Although dozens of bilinear joint time-frequency representations have been proposed over the last five decades, we shall discuss the Wigner distribution because it is simple and powerful.

In the following sections, we discuss the motivation and general properties of the Wigner distribution. What makes the Wigner distribution so unique are its descriptions of a signal's time-varying nature better than many other representations, such as STFT Spectrogram. Moreover, the Wigner distribution possesses many properties useful for signal analysis. The problems of the Wigner distribution have been so-called cross-term interference that severely limits the application of the Wigner distribution. The discussion in this section mainly follows that of *Papoulis,* 1977; *Claasen and Mecklenbrauker* [1980a,b, 1983], *Cohen* [1989, 1995] *Hlawatsch and Boudreaux-Bartels,* [1992], *Qian and Chen* [1996], *Hlawatsch* [1998], *Mertins* [1999], in addition to the references given.

7.1.1 Time-Dependent Power Spectrum

The square of the Fourier Transform is called the *power spectrum*, which characterizes the signal's energy distribution in the frequency domain. While the Fourier Transform is linear, the power spectrum (PS) is the quadratic function of frequencies. Accordingly, we also use the square of Short-Time Fourier Transform to describe the signal's energy distribution in joint time-frequency domain. According to the Wiener-Khinchin theorem, the power spectrum can also be considered as the Fourier Transform of the auto-correlation function $R(\tau)$ [*Lampard,* 1954; *Schroeder and Atal,* 1962; *Papoulis,* 1977; *Cohen,* 1989; *Qian and Chen,* 1996]

$$\mathrm{PS}(t,\omega) = \mid S(\omega) \mid^2 = \int R(\tau)\, \exp(-j\omega\tau)\, d\tau \tag{179}$$

where $R(\tau)$ is computed by

$$R(\tau) = \int s(t)\, s^*(t-\tau)\, d\tau \tag{180}$$

Equation (179) is not a function of time, which indicates how much energy is present in frequency over ω over the entire time period. But it does not show how the spectrum is distributed in time. Based on equation (179), there is no way to tell whether or not a signal's power spectrum changes over time. Therefore, the standard power spectrum is inadequate a depict signals whose frequency contents evolve with time.

By examining equation (179), we can see that one possibly way to depict a time dependent spectrum is to make the auto-correlation function time-dependent. The resultant Fourier Transform of the time-dependent auto-correlation function $R(t,\tau)$, with respect to variable τ, is then a function of time, i.e.,

$$P(t,\omega) = \int R(t,\tau)\, \exp(-j\omega\tau)\, d\tau \tag{181}$$

We name $P(t,\omega)$ a $time-dependent\ power\ spectrum$ [*Qian and Chen,* 1996].

Apparently, the choice of $R(t,\tau)$ is not arbitrary. For example, because $P(t,\omega)$ presumably describes the time-dependent spectrum, adding all instantaneous-time power spectrum $P(t_0,\omega)$ should yield the total power spectrum $|S(\omega)|^2$, i.e.,

$$\int P(t,\omega)\, dt = \mid S(\omega) \mid^2 \tag{182}$$

which is traditionally called the $frequency\ marginal\ condition$ [*Cohen,* 1989; *Qian and Chen,* 1996]. Conversely, the integration along the frequency axis should be equal to the instantaneous energy, i.e.,

$$\frac{1}{2\pi}\int P(t,\omega)\, d\omega = \mid s(t) \mid^2 \tag{183}$$

which is commonly known as the *time marginal condition* [*Cohen,* 1989; *Qian and Chen,* 1996]. If $P(t, \omega)$ represents signal energy distribution in the joint time-frequency domain, then we hope it is real valued, i.e.,

$$P(t, \omega) = P^{\star}(t, \omega) \tag{184}$$

From the conventional energy concept, we also wish that the time-dependent spectrum would be non-negative.

7.2 The Wigner Distribution

The Wigner distribution is the prototype of distributions that are qualitatively different from the spectrogram. The discovery of its strengths and shortcomings has been a major thrust in the development of the field. It has often been studied in contrast to the spectrogram. The Wigner distribution was originally developed for the area of quantum mechanics in 1932 [*Wigner,* 1932] and was introduced for signal analysis by a French scientist Ville 16 years later [*Ville,* 1948]. In the Wigner distribution, the time-dependent auto-correlation function is chosen to be

$$R(t, \tau) = s^{\star}(t - \frac{1}{2}\tau)\, s(t + \frac{1}{2}\tau) \tag{185}$$

Substituting the above time-dependent auto-correlation into equation (181) yields [*Papoulis,* 1977; *Claasen and Mecklenbrauker,* 1980; *Cohen,* 1989; *Hlawatsch and Boudreaux-Bartels,* 1992; *Qian and Chen,* 1996; *Hlawatsch,* 1998; *Mertins,* 1999]

$$\begin{aligned} W(t, \omega) &= \frac{1}{2\pi} \int s^{\star}(t - \frac{1}{2}\tau)\, s(t + \frac{1}{2}\tau)\, e^{-j\tau\omega}\, d\tau & (186) \\ &= \frac{1}{2\pi} \int S^{\star}(\omega + \frac{1}{2}\theta)\, S(\omega - \frac{1}{2}\theta)\, e^{-jt\theta}\, d\theta & (187) \end{aligned}$$

The equivalence of the two expressions is easily checked by writing the signal in terms of the spectrum and substituting into equation (186). The Wigner distribution is said to be bilinear in the signal because the signal enters twice in its calculation.

Notice that to obtain the Wigner distribution at a particular time we add up pieces made up of the product of the signal at a past time multiplied by the signal at a future time, the time into the past being equal to the time into the future. Therefore, to determine the properties of the Wigner distribution at a time t we mentally fold the left part of the signal over to the right to see if there is any overlap. If there is, then those properties will be present now, at time t. Everything we have said for the time domain holds for the frequency domain because the Wigner distribution is basically identical in form in both domains. Another important point is that the Wigner distribution weighs the far away times equally to the near times. Hence the Wigner distribution is highly non-local.

Equation (187) is usually called as the auto-Wigner distribution. Accordingly, the cross-Wigner distribution is defined as

$$W(t,\omega) = \frac{1}{2\pi} \int s(t+\frac{1}{2}\tau)\, g^{\star}(t-\frac{1}{2}\tau)\, e^{-j\tau\omega}\, d\tau \tag{188}$$

7.2.1 Range of the Wigner Distribution

The Wigner distribution satisfies the finite support properties in time and frequency [*Cohen,* 1987a,b; *Cohen,* 1989; *Hlawatsch,* 1984, 1988, 1998; *Qian and Chen,* 1996; *Mertins,* 1999]

$$W(t,\omega) = 0 \quad \text{for } t \text{ outside } (t_1,t_2) \qquad \text{if } s(t) \text{ is zero outside } (t_l,t_2) \tag{189}$$

$$W(t,\omega) = 0 \quad \text{for } \omega \text{ outside } (\omega_1,\omega_2) \qquad \text{if } S(\omega) \text{ is zero outside } (\omega_l,\omega_2) \tag{190}$$

The time and frequency support of the Wigner distribution are claimed as desirable properties. If the time series is non-zero in a certain range (t_1,t_2) and zero elsewhere, then the Wigner distribution is also non-zero in this range (t_1,t_2) and zero elsewhere. This is shown in Figure 3a and Figure 3b. The time support characteristic of the Wigner distribution seems to be a very attractive property, but may be a bit misleading. One should not infer that any zero-valued region in the time series has a corresponding zero-valued region in the Wigner distribution. This is true only if the zero-filled region extends to $\pm\infty$. Figure 4 illustrates that the cross-term inherent in the Wigner distribution contaminate the zero region if it has non-zero regions on either side. Here the time series has a large zero region in the middle. The dominant signal seen in the Wigner distribution is centered around time=64, in the middle of the zero region. This plot inspires the seeming paradox that the marginal property is satisfied (i.e., the integral over frequency of Wigner distribution gives the instantaneous power in the time series), and yet the Wigner distribution is non-zero in the region where the time series is zero (likewise for the integral over time and the non-zero frequency regions). Figure 5 shows the result of performing the Wigner distribution on a three component signal. In Figure 6, a time series with a signal in the first part and noise in the second part, illustrates the fact that a noisy region can contaminate the non-noisy region in the Wigner representation.

Other distributions such as the Cone-Kernal distribution [*Zhao et al.,* 1990], the Choi-Williams distribution *Choi and Williams,* 1989], as well as the smoothed Pseudo Wigner Distribution (see the following subsection) [*Hlawatsch and Boudreaux-Bartels,* 1992] reduce the effect of these cross-terms, but in doing so fail to satisfy the marginal properties of the Wigner distribution.

7.2.2 The Characteristic Function of the Wigner Distribution

We have [*Cohen,* 1989, 1995]

$$M(\theta,\tau) = \int\int e^{j\theta t + j\tau\omega}\, W(t,\omega)\, dt\, d\omega \tag{191}$$

$$= \frac{1}{2\pi}\int\int\int e^{j\theta t + j\tau\omega}\, s^\star(t-\frac{1}{2}\tau')\, s(t+\frac{1}{2}\tau')\, e^{-j\tau'\omega}\, d\tau'\, dt\, d\omega \tag{192}$$

$$= \int\int e^{j\theta t}\, \delta(\tau-\tau')\, s^\star(t-\frac{1}{2}\tau')\, s(t+\frac{1}{2}\tau')\, d\tau'\, dt \tag{193}$$

$$= \int s^\star(t-\frac{1}{2}\tau)\, s(t+\frac{1}{2}\tau)\, e^{j\theta t}\, dt \tag{194}$$

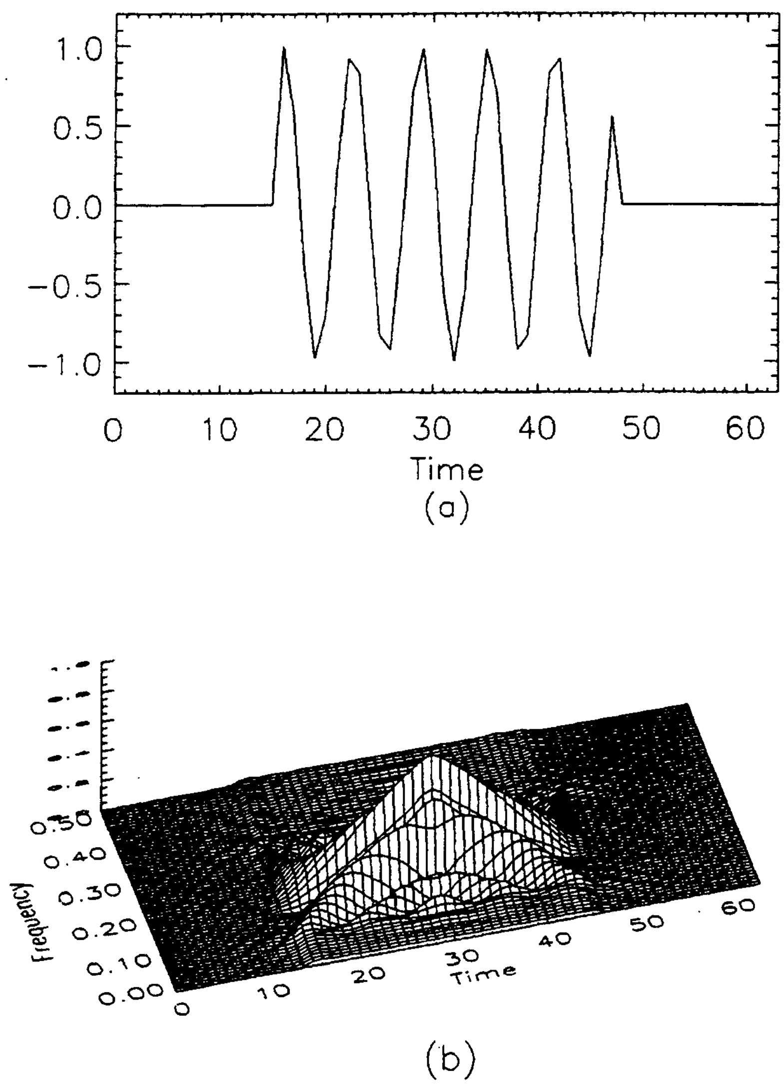

Figure 3: A sample time series with a non-zero region. (b) The Wigner distribution of the time series in (a). The Wigner distribution has finite support. It is zero in the region for $t < t_1$ where t_1 is the start of the signal, and in the region for $t > t_2$ where t_2 is the end of the signal.

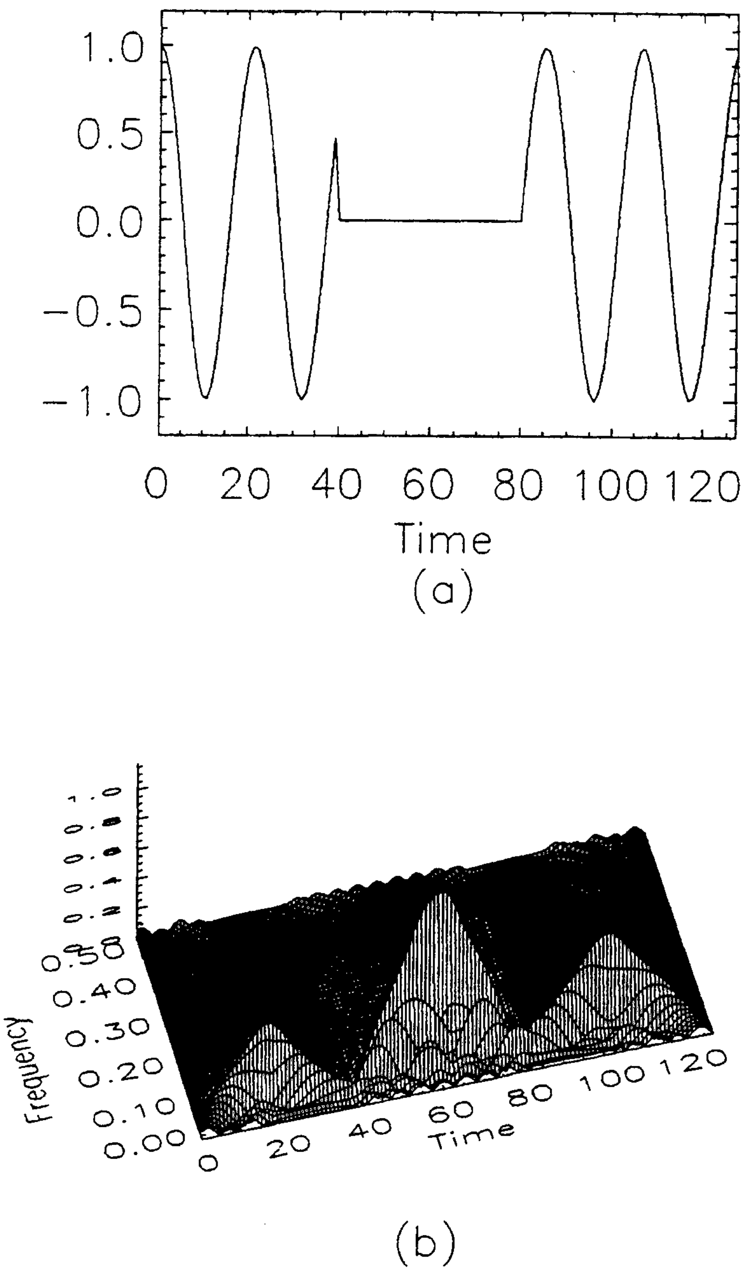

Figure 4: A sample time series with zero-filled region. (b) The Wigner distribution of the time series in (a). The Wigner distribution has compact support, but it is not identically equal to zero when the signal is zero. In fact, the zero region of the signal corresponds to the maximum peak in the Wigner distribution.

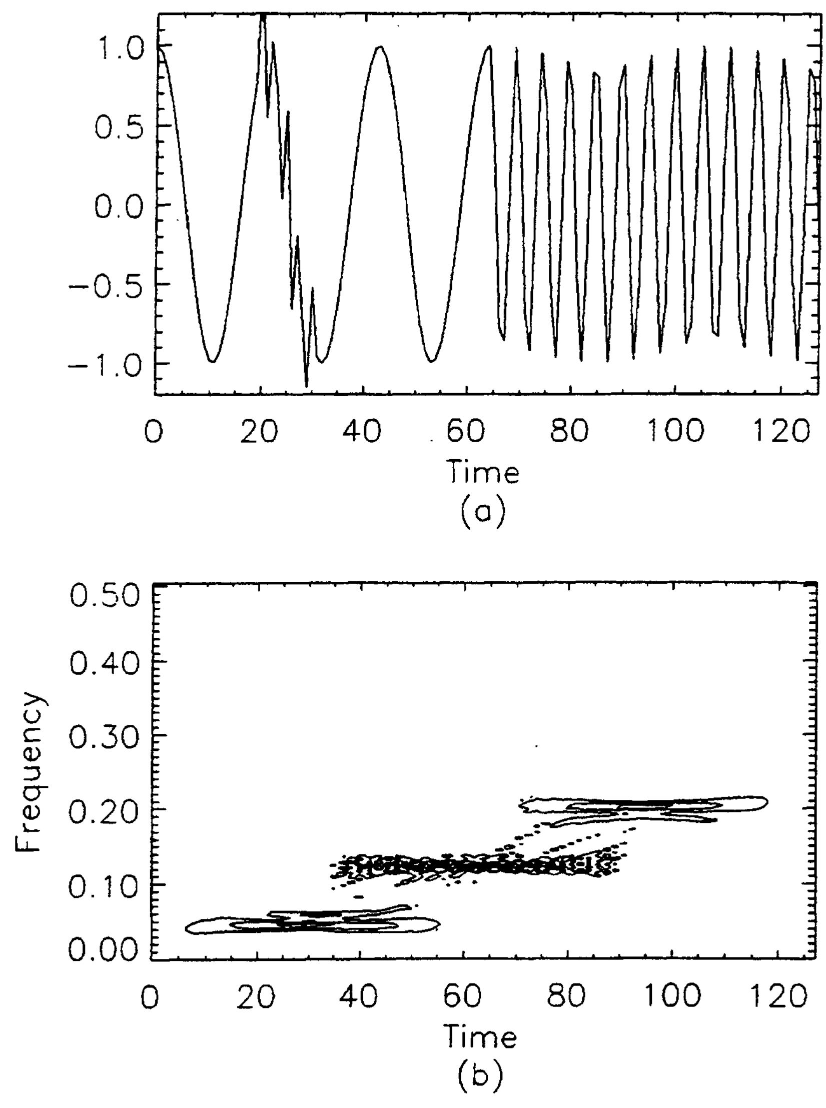

Figure 5: A sample time series with three distinct regions. (b) The Wigner distribution of the time series in (a). The cross-terms in the Wigner distribution heavily contaminate the signal. and the high frequency burst is not detected.

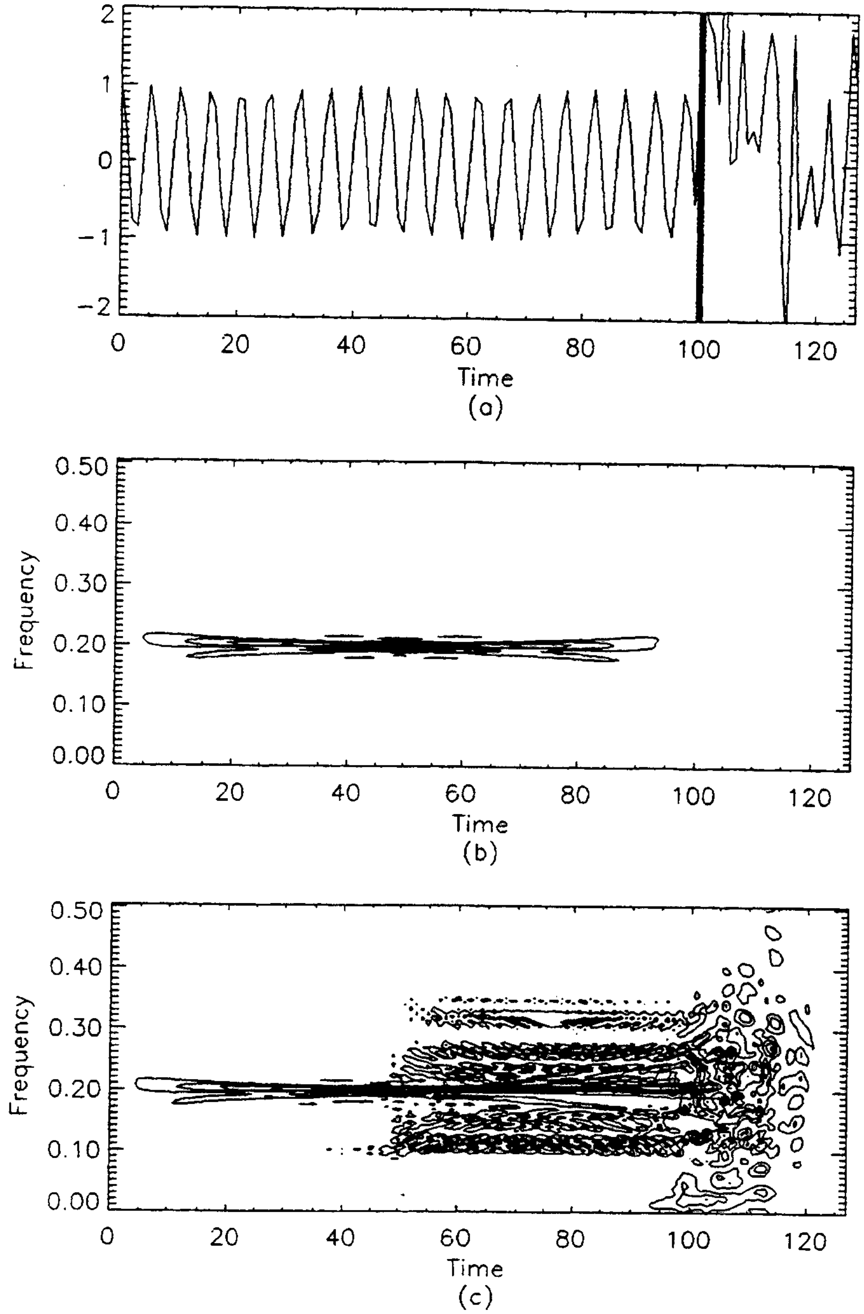

Figure 6: A sample time series with a sinusoid in the first 100 points, and normally distributed noise in the last 28 points. The vertical black line denotes the boundary between the two regions. (b) The Wigner distribution of the time series in (a) without the noise region. (c) The Wigner distribution of the time series in (a). The noisy region is propagated into the sinusoidal part of the signal.

$$= \quad A(\theta, \tau) \tag{195}$$

This function and variants of it have played a major role in signal analysis. This particular form is called the symmetric ambiguity function. It was first derived by *Ville* [1949] and *Moyal* [1949] and its relation to matched filters was developed by *Woodward* [1953]. We have previously discussed it in the calculation of the characteristic function of the spectrogram, equation (165). In terms of the spectrum the characteristic function is [*Cohen,* 1989, 1995]

$$M(\theta, \tau) = \int S^{\star}(\omega + \frac{1}{2}\theta)\, S(\omega - \frac{1}{2}\theta)\, e^{j\tau\omega}\, dt \tag{196}$$

7.3 General Properties of Wigner distribution

In the preceding section, we discussed the concept of time-dependent spectrum and the Wigner distribution. Compared to STFT, the Wigner distribution not only has a better resolution, but also does not suffer the window effects. We now discuss the basic properties of the Wigner distribution.

7.3.1 Time and Frequency Shift Invariance

If we time shift the signal by t_0 and/or shift the spectrum by ω_0, then the Wigner distribution is shifted accordingly [*Cohen,* 1989, 1995],

$$\text{if} \quad s(t) \rightarrow e^{j\omega_0 t}\, s(t - t_0) \quad \text{then} \quad W(t, \omega) \rightarrow W(t - t_0, \omega - \omega_0) \tag{197}$$

To see this we replace the signal by $e^{j\omega_0 t} s(t - t_0)$ in the Wigner distribution and call W_{sh} the shifted distribution,

$$W_{sh}(t, \omega) = \frac{1}{2\pi} \int e^{-j\omega_0 (t - \tau/2)} s^{\star}(t - t_0 - \frac{1}{2}\tau) \times e^{-j\omega_0 (t + \tau/2)} s(t - t_0 + \frac{1}{2}\tau)\, e^{-j\tau\omega} d\tau \tag{198}$$

$$= \frac{1}{2\pi} \int s^{\star}(t - t_0 - \frac{1}{2}\tau) s(t - t_0 + \frac{1}{2}\tau) e^{-j\tau(\omega - \omega_0)} d\tau \tag{199}$$

$$= W(t - t_0), \omega - \omega_0) \tag{200}$$

7.3.2 Reality

The Wigner distribution is always real, even if the signal is complex. This can be verified by considering the complex conjugate of $W(t, \omega)$ [*Cohen,* 1989, 1995],

$$W^{\star}(t, \omega) = \frac{1}{2\pi} \int s(t - \frac{1}{2}\tau) \, s^{\star}(t + \frac{1}{2}\tau) \, e^{j\tau\omega} \, d\tau \tag{201}$$

$$= -\frac{1}{2\pi} \int_{\infty}^{-\infty} s(t + \frac{1}{2}\tau) \, s^{\star}(t - \frac{1}{2}\tau) \, e^{-j\tau\omega} \, d\tau \tag{202}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} s(t + \frac{1}{2}\tau) \, s^{\star}(t - \frac{1}{2}\tau) \, e^{-j\tau\omega} \, d\tau \tag{203}$$

$$= W(t, \omega) \tag{204}$$

The fact that the Wigner distribution is real for any signal can also be seen from the characteristic function. Recall that $M^{\star}(-\theta, -\tau) = M(\theta, \tau)$ is the condition for a distribution to be real. But the characteristic function of the Wigner distribution is the ambiguity function, $A(\theta, \tau)$, equation (194), which does satisfy this property.

7.3.3 Symmetry

Substituting $-\omega$ for ω into the Wigner distribution we see that we obtain the identical form back if the signal is real [*Cohen,* 1989, 1995]. But real signals have symmetrical spectra. Therefore, for symmetric spectra the Wigner distribution is symmetrical in the frequency domain. Similarly for real spectra the time waveform is symmetrical and the Wigner distribution is symmetric in time. Therefore,

$$W(t,\omega) = W(t,-\omega) \text{for real signals} \equiv \text{symmetrical spectra, } S(\omega) = S(-\omega) \tag{205}$$

$$W(t,\omega) = W(-t,\omega) \text{for real spectra} \equiv \text{symmetrical signals, } s(t) = s(-t) \tag{206}$$

7.3.4 Time and Frequency Marginals

The Wigner distribution satisfies the time-frequency marginals

$$\int W(t,\omega)\, d\omega = |s(t)|^2 \tag{207}$$

$$\int W(t,\omega)\, dt = |S(\omega)|^2 \tag{208}$$

Both of these equations can be readily verified by examining $M(\theta, 0)$ and $M(0, \tau)$. By inspection of equation (194) and equation (196) we have

$$M(\theta,0) = \int |s(t)|^2 \, e^{i\theta t}\, dt \quad ; \quad M(0,\tau) = \int |S(\omega)|^2 \, e^{i\tau\omega}\, d\omega \tag{209}$$

But these are the characteristic functions of the marginals and hence the marginals are satisfied. To do it directly [*Cohen,* 1995],

$$\begin{aligned} P(t) = \int W(t,\omega)\, d\omega &= \frac{1}{2\pi} \int\int s^\star(t - \frac{1}{2}\tau)\, s(t + \frac{1}{2}\tau)\, e^{-j\tau\omega}\, d\tau\, d\omega & (210) \\ &= \int s^\star(t - \frac{1}{2}\tau)\, s(t + \frac{1}{2}\tau)\, \delta(\tau)\, d\tau & (211) \\ &= |s(t)|^2 & (212) \end{aligned}$$

and similarly for the marginal in frequency. Since the marginals are satisfied, the total energy condition is also automatically satisfied,

$$E = \int\int W(t,\omega)\, d\omega\, dt = \int |s(t)|^2\, d\tau = 1 \tag{213}$$

7.3.5 Instantaneous Frequency and Group Delay

Let $s(t) = A(t)\exp(j\varphi(t))$, where amplitude $A(t)$ and phase $\varphi(t)$ both are real-valued functions. Then

$$< \omega >_t = \frac{\int \omega\, W(t,\omega)\, d\omega}{\int W(t,\omega)\, d\omega} = \frac{1}{|\, A(t)\, |^2} \int \omega\, W(t,\omega)\, d\omega = \varphi'(t) \tag{214}$$

which says that, at time t, the mean instantaneous frequency of Wigner distribution is equal to the mean instantaneous frequency of the analyzed signal [*Cohen,* 1995; *Qian and Chen,* 1996].

Assume that the Fourier Transform of signal $s(t)$ is $S(\omega) = B(\omega)\exp(j\psi(\omega))$. Then the first derivative of the Phase $\phi'(\omega)$ is called the group delay. For Winger distribution, we have

$$< \omega >_s = \frac{\int t\, W(t,\omega)\, dt}{\int W(t,\omega)\, dt} = \frac{1}{|\, S(\omega)\, |^2} \int t\, W(t,\omega)\, dt = -\psi'(\omega) \tag{215}$$

which says that the conditional mean time of the Wigner distribution is equal to the group delay [*Cohen,* 1995; *Qian and Chen,* 1996]. These results are important because they are always true for any signal. Recall that for the spectrogram they were never correctly given.

7.4 The Wigner Distribution of the Sum of Multiple Signals

As discussed in the preceding sections, the Wigner distribution not only possesses many useful properties, but also has better resolution than the STFT spectrogram. Although the Wigner distribution has existed for a long time, its applications are very limited. One main deficiency of the Wigner distribution is the so-called cross-interference [*Papoulis,* 1977; *Hlawatsch,* 1984, 1998; *Cohen,* 1989, 1995]. Suppose we express a signal as the sum of two pieces,

$$s(t) = s_1(t) + s_2(t) \tag{216}$$

Substituting this into the definition, we have

$$W(t,\omega) = W_{11}(t,\omega) + W_{22}(t,\omega) + W_{12}(t,\omega) + W_{21}(t,\omega) \tag{217}$$

where

$$W_{12}(t,\omega) = \frac{1}{2\pi} \int s_1^\star(t - \frac{1}{2}\tau)\, s_2(t + \frac{1}{2}\tau)\, e^{-j\tau\omega}\, d\tau \tag{218}$$

This is called Wigner distribution. In terms of the spectrum it is

$$W_{12}(t,\omega) = \frac{1}{2\pi} \int S_1^\star(\omega + \frac{1}{2}\theta)\, S_2(\omega - \frac{1}{2}\theta)\, e^{-jt\theta}\, d\theta \tag{219}$$

The cross Wigner distribution is complex. However, $W_{12} = W_{21}^\star$, and therefore $W_{12}(t,\omega) + W_{21}(t,\omega)$ is real. Hence

$$W(t,\omega) = W_{11}(t,\omega) + W_{22}(t,\omega) + 2\,\Re\{W_{12}(t,\omega)\} \tag{220}$$

We see that the Wigner distribution of the sum of the two signals is not the sum of the Wigner distribution of each signal but has the additional term $2\Re\{W_{12}(t,\omega)\}$. The term is often called the interference term or the cross-term and it is often said to give rise to artifacts. Because the cross-term usually oscillates and its magnitude is twice as large as that of the auto-terms, it often obscures the useful time-dependent spectrum patterns.

Figures 7 and 8 give two examples to get a better idea about the cross-term interference. The Wigner distribution sometimes places values in the middle of the two signals both in time and in frequency. Sometimes these values are in places in the time-frequency plane at odds with what is expected. A typical case is illustrated in Figure 7.

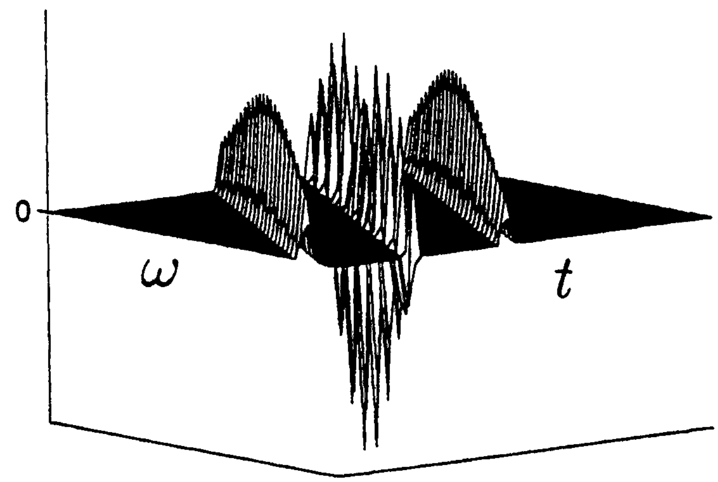

Figure 7: The Wigner distribution of the sum of two chirps illustrating the cross terms [after Qian and Chen, 1996].

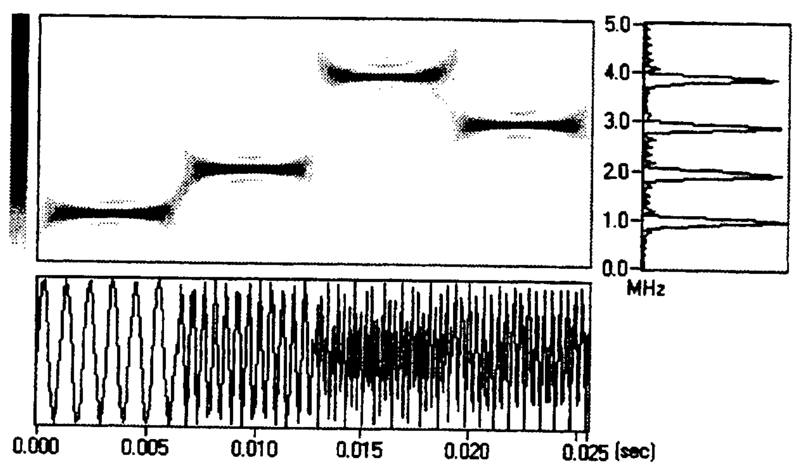

Figure 8: The bottom plot is a time waveform that contains four frequency tones. The right plot is the traditional power spectrum. The middle one is the joint time-frequency representation [after Qian and Chen, 1996].

From equation (220), each pair of auto-terms creates one cross-term. For N individual components, the total number of cross-terms is $N(N-1)/2$. In the simple case, such as in Figure 8, we can easily identify the cross-term interference. For real signals, the pattern of the cross-terms, which usually overlap with auto-terms could be more complicated and confusing. Consequently, the desired time-dependent spectrum could be deceiving and confusing.

Figure 8 illustrates the sum of four frequency tones. The bottom plot is time waveform. The right plot is the traditional power spectrum. The middle plot is the desired time-dependent spectrum. The conventional power spectrum indicates that there are four different frequency tones, but it is not clear when those different frequency tones occur. The time-dependent power spectrum not only shows four frequency tones, but also tells when they take place. How to reduce the cross-term interference without destroying the useful properties of the Wigner distribution has been very important to time-frequency analysis.

7.5 Pseudo Wigner Distribution

For a given time the Wigner distribution weighs equally all times of the future and past. Similarly, for a given frequency it weighs equally all frequencies below and above that frequency. There are two reasons for wanting to modify this basic property of the Wigner distribution. First, in practice we may not be able to integrate from minus to plus infinity and so one should study the effects of limiting the range. Second, in calculating the distribution for a time t, we may want to emphasize the properties near the time of interest compared to the far away times. To achieve this, note that for a given time the Wigner distribution is the Fourier Transform with respect to τ of the quantity $s^\star(t-\frac{1}{2}\tau)\, s(t+\frac{1}{2}\tau)$. The variable τ is called the lag variable. Therefore if we want to emphasize the signal around time t, we multiply this product by a function that is peaked around $\tau = 0$, $h(\tau)$ say, to define the pseudo Wigner distribution [*Claasen and Mecklenbrauker,* 1980b; *Cohen,* 1989, 1995; *Hlawatsch and Boudreaux-Bartels,* 1992; *Qian and Chen,* 1996; *Hlawatsch,* 1998]

$$W_{PS}(t,\omega) = \int h(\tau)\, s^\star(t-\frac{1}{2}\tau)\, s(t+\frac{1}{2}\tau)\, e^{-j\omega\tau}\, d\tau \tag{221}$$

The Wigner distribution is highly non-local and the effect of the windowing is to make it less so. One of the consequences of this is that the pseudo Wigner distribution suppresses, to some extent, the cross-terms for multicomponent signals. This is because we have made the Wigner distribution local. While windowing the lag does suppress the cross-terms, it also destroys many of the desirable properties of the Wigner distribution. For example the marginals and instantaneous frequency properties no longer hold.

7.6 Discrete Wigner distribution

The continuous time Wigner distribution introduced in the previous sections is of great value in analyzing and gaining insight into the properties of continuous time signals. Because the majority of signals that we deal with are discrete time signals, in the present section we shall address the subject of discrete Wigner distribution [*Claasen and Mecklenbrauker,* 1980a, 1983; *Cohen,* 1987a,b; *Cohen,* 1989; *Mertins,* 1999].

By letting $u = \tau/2$ in equation (186), the Wigner distribution becomes [*Qian and Chen,* 1996]

$$W(t,\omega) = 2 \int s^\star(t-u)\, s(t+u)\, e^{-2j\omega u}\, du \tag{222}$$

Assume the interval of the integration (222) is Δ. We have an approximation of the integral by

$$W(t,\omega) = 2\Delta \sum_n s^\star(t-n\Delta)\, s(t+n\Delta)\, e^{-2j\omega n\Delta} \tag{223}$$

If the signal $s(t)$ is sampled in every T second, $T = \Delta$, then we obtain the discrete time Wigner distribution as

$$W(mT,\omega) = 2T \sum_n s^\star((m-n)T)\, s((m+n)T)\, e^{-2j\omega nT} \tag{224}$$

where n and m are integer-vulued. Obviously,

$$W(mT, \omega + \frac{\pi}{T}) = W(mT,\omega) \tag{225}$$

Therefore, the period of $W(mT, \omega)$ is π/T rather than $2\pi/T$ required by sampling theory. The equation (225) implies that the highest frequency component in equation (225) must be less than or equal to $\pi/(2T)$. If the signal bandwidth is larger than $\pi/(2T)$, then aliasing will occur. In order to obtain an aliasing-free discrete time Wigner distribution, we have to double the sampling rate. The simplest way of doubling sampling is to apply the interpolation filter. If the original sample interval is T second, then the interval of the interpolated samples is $T/2$ second. Applying this result into equation (224), we have

$$W(mT/2, \omega) = 2T/2 \sum_n s^\star((m-n)T/2)\, s((m+n)T/2)\, e^{-2j\omega nT/2} \tag{226}$$

The period of (226) becomes $2\pi/T$, which is exactly what we anticipate. Let the normalized frequency $\theta = \omega T/2$. Without loss of generality, we further assume that $T = 2$, then equation (226) reduces to

$$W(m, \theta) = 2 \sum_{n=-\infty}^{\infty} s^\star(m-n)\, s(m+n)\, e^{-2j\theta n} \tag{227}$$

7.7 Comparison of the Wigner Distribution with the Spectrogram

It has often been said that one of the advantages of the Wigner distribution over the spectrogram is that we do not have to bother with choosing the window. This viewpoint misses the essence of the issue. The spectrogram is not one distribution, it is an infinite class of distributions and to say that an advantage is that one does not have to choose makes as much sense as saying one book is better than a library because we don't have to choose which book to read. Here is the point: The Wigner distribution in some respects is better than any spectrogram. It is not that we do not have to bother about choosing a window, it is that even if we bothered we wouldn't find one that produces a spectrogram that is better than the Wigner. In particular, the Wigner distribution gives a clear picture of the instantaneous frequency and group delay. In fact, the conditional averages are exactly the instantaneous frequency and group delay. This is always true for the Wigner distribution; it is never true for the spectrogram. We could search forever and never find a window that will produce a spectrogram that will give the instantaneous frequency and group delay, although sometimes a good approximation is achieved.

Another property of the Wigner distribution is that it satisfies the marginals and always gives the correct answers for averages of functions of frequency or time and always satisfies the uncertainty principle of the signal. On the other hand, the spectrogram never gives the correct answers for these averages and never satisfies the uncertainty principle of the signal.

The main deficiency of the Wigner distribution is the so-called cross-term interference. At any time instant, if there is more than one frequency tone, then the Wigner distribution may become messed up because of the presence of undesired terms. However, the cross-terms highly oscillate and are localized, which always occur in the midway of the pair of corresponding auto-terms. On the other hand, although spectrogram resolve the components at certain cases, it often cannot resolve the components effectively.

8 The S-Transform

8.1 Introduction

The S-Transform is a new transform that produces a time frequency representation of a time series. It is a generalization of the Fourier Transform to the case of non-stationary time series. It uniquely combines a frequency dependent resolution with simultaneously localizing the real and imaginary spectra. It was first published in *Stockwell et al.* [1996c], and since has seen several interesting applications [*Eramian,* 1996; *Chu,* 1996; *Osler and Chapman,* 1996; *Stockwell,* 1999; *Stockwell et al.*, 1996a; *Stockwell et al.*, 1996b; *Varanini et al.*, 1997; *Mansinha et al.*, 1997a; *Mansinha et al.*, 1997b; *Fritts,* 1998; *Eramian et al.*, 1999].

8.2 Derivation of the S-Transform from the Short Time Fourier Transform

What follows is the original derivation of the S-Transform that demonstrates the relationship between the S-Transform and the Short-Time Fourier Transform.

Recall that the spectrum $S(f)$ of a time series $s(t)$ is given by the standard Fourier analysis as [*Stockwell et al.*, 1996c; *Mansinha et al.*, 1997b]:

$$S(f) = \int_{-\infty}^{\infty} s(t)\, e^{-i2\pi ft}\, dt \tag{228}$$

and its inverse relationship is:

$$s(t) = \int_{-\infty}^{\infty} S(f)\, e^{i2\pi ft}\, dt \tag{229}$$

The spectrum $S(f)$ can be referred to as the "time-averaged spectrum".

If the time series $s(t)$ is windowed (or multiplied point by point with) a window function $g(t)$ then the resulting spectrum is

$$S_g(f) = \int_{-\infty}^{\infty} s(t)\, g(t)\, e^{-i2\pi f t}\, dt \tag{230}$$

The S-Transform can be found by first defining a particular window function, a normalized Gaussian

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{t^2}{2\sigma^2}} \tag{231}$$

and then by allowing it to be a function of translation τ and dilation (or window width) σ [*Stockwell et al.*, 1996c; *Mansinha et al.*, 1997b].

$$ST^*(\tau, f, \sigma) = \int_{-\infty}^{\infty} s(t)\, \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(t-\tau)^2}{2\sigma^2}}\, e^{-i2\pi f t}\, dt \tag{232}$$

which, with a particular value of σ, is similar in definition to the Short-Time Fourier transform (see Section 6). The Gaussian window is chosen because it is the most compact in time and frequency [*Janssen*, 1992]. In fact, this is a special case of the Multiresolution Fourier Transform [*Wilson et al.*, 1992]. Because this is a function of three independent variables, it is impractical as a tool for analysis. Simplification can be achieved by adding the constraint restricting the width of the window σ to be proportional to the inverse of the frequency (or proportional to the period) [*Stockwell et al.*, 1996c; *Mansinha et al.*, 1997b]

$$\sigma(f) = \frac{1}{\mid f \mid} \tag{233}$$

Thus one has the S-Transform [*Stockwell et al.*, 1996c; *Stockwell*, 1999; *Mansinha et al.*, 1997b]:

$$ST(\tau, f) = \frac{\mid f \mid}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(t)\, e^{-\frac{(t-\tau)^2 f^2}{2}}\, e^{-i2\pi f t}\, dt \tag{234}$$

The one dimensional function of the time variable τ and fixed parameter f_1 defined by $ST(\tau, f_1)$ is called a *voice* (as with Wavelet Transforms). The one dimensional function of the frequency variable f and fixed parameter τ_1 defined by $ST(\tau_1, f)$ is called a *local*, or a *local spectrum* [*Stockwell et al.*, 1996c; *Stockwell*, 1999].

One can see that the zero frequency voice of the S-Transform is identically equal to zero. This adds no information, therefore $ST(\tau, 0)$ is defined to be independent of time and equal to the average of the function $s(t)$ [*Stockwell et al.*, 1996c; *Stockwell,* 1999],

$$ST(\tau, 0) = \lim_{T \to \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} s(t)\, dt \tag{235}$$

The S-Transform can be written as a convolution of two functions over the variable t [*Stockwell et al.,* 1996c; *Stockwell,* 1999]

$$ST(\tau, f) = \int_{-\infty}^{\infty} p(t, f)\, g(\tau - t, f)\, dt \tag{236}$$

or

$$ST(\tau, f) = p(\tau, f) \star g(\tau, f) \tag{237}$$

where

$$p(\tau, f) = s(\tau)\, e^{-i2\pi f \tau} \tag{238}$$

and

$$g(\tau, f) = \frac{\mid f \mid}{\sqrt{2\pi}}\, e^{-\frac{\tau^2 f^2}{2}} \tag{239}$$

Let $B(\alpha, f)$ be the Fourier Transform (from τ to α) of the S-Transform $ST(\tau, f)$. By the convolution theorem [*Brigham,* 1975] the convolution in the τ (time) domain becomes a multiplication in the α (frequency) domain [*Stockwell et al.,* 1996c; *Stockwell,* 1999]:

$$B(\alpha, f) = P(\alpha, f)\, G(\alpha, f) \tag{240}$$

Likewise, $P(\alpha, f)$ and $G(\alpha, f)$ are the Fourier Transform of $p(\tau, f)$ and $g(\tau, f)$. Explicitly,

$$B(\alpha, f) = S(\alpha + f)\, e^{-\frac{2\pi^2\alpha^2}{f^2}} \tag{241}$$

where $S(\alpha + f)$ is the Fourier Transform of (238), and the exponential term is the Fourier transform of the Gaussian function (239). Thus the S-Transform is the inverse Fourier Transform (from α to τ) of the above equation (for $f \neq 0$) [*Stockwell et al.,* 1996c; *Stockwell,* 1999].

$$ST(\tau, f) = \int_{-\infty}^{\infty} S(\alpha + f)\, e^{-\frac{2\pi^2\alpha^2}{f^2}}\, e^{i2\pi\alpha\tau}\, d\alpha \tag{242}$$

The exponential function in equation (242) is the frequency dependent localizing window and is called the *voice Gaussian*. It plays the role of a low pass filter for each particular voice.

The S-Transform improves on the Short-Time Fourier Transform in that it has better resolution in phase space (i.e., a more narrow time window for higher frequencies), giving a fundamentally more sound time frequency representation [*Daubachie,* 1990].

8.3 Properties of the S-Transform

Because of the absolutely referenced phase information in the S-Transform, many advantageous characteristics of the S-Transform arise.

8.3.1 Inverse of the S-Transform and the Fourier Transform

If the S-Transform is indeed a representation of the local spectrum, one would expect that the simple operation of averaging the local spectra over time would give the Fourier Transform spectrum [*Stockwell et al.,* 1996c; *Stockwell,* 1999]. This is indeed the case with the S-Transform

$$\int_{-\infty}^{\infty} ST(\tau, f)\, d\tau = S(f) \tag{243}$$

where $S(f)$ is the Fourier Transform of $s(t)$. It follows that $s(t)$ is exactly recoverable from $ST(\tau, f)$

$$s(t) = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} ST(\tau, f)\, d\tau \right\} e^{i2\pi ft}\, df \tag{244}$$

This shows that the S-Transform is a generalization of the Fourier Transform to non-stationary time series.

8.3.2 Progressive Resolution of the Time-Frequency Domain

The sampled frequencies of a time series consisting of N points with a sampling interval of T are $f_n = n/(NT)$. Thus the periods sampled range from NT for the first harmonic (ignoring the DC level), $NT/2$ for the second harmonic, and so on as function of $1/n$. Because of this, one would expect that it is easier to tell (given one period of an oscillation) the difference between the second and third harmonic (where the difference in period is $\Delta T = (NT/3 - NT/2) = NT/6$) and the 33rd and 34th harmonic (where the difference in period is $\Delta T = (NT/33 - NT/34) = 0.00089 \star NT/2)$.

This is illustrated in Figure (9). In Figure 9a, "short-lived *low*-frequency signals" are shown (short-lived because they last for one period). Because the periods are quite different, it is very easy to discern the difference. Thus it is easy to distinguish short-lived low-frequency signals from each other.

In Figure 9b, the "short-lived *high*-frequency signals" are shown (again, short-lived because they last for only one period). It is very difficult to resolve these two signals, since their periods are quite similar (in fact the difference in period is smaller that the sampling interval of the time series). Thus it is difficult to distinguish short-lived high-frequency signals from each other.

This illustrates that, for short-lived oscillations, the high frequencies are difficult to resolve from each other, and hence high frequencies have poor frequency resolution. It also illustrates that, for short-lived oscillations, low frequencies are easy to resolve from each other, and hence low frequencies have good frequency resolution. Short-lived high frequency signals have good time resolution, compared to short-lived low frequency signals. This pattern (Figure 10) of (high frequency) = (good time resolution) and (poor frequency resolution) while (low frequency) = (poor time resolution) and (good frequency resolution), is characteristic of the S-transform. The relationship between temporal and frequency resolution is directly related to uncertainty principle.

8.3.3 Frequency Resolution Depends on the Signal

Increasing the resolution with which one can measure the frequency of an oscillating function depends on the length of time that signal exists within the time series. A common error is to state that the resolution depends on the length of time that one measures. This is illustrated in Figure 11. In part (a), a time series consists of a pure sinusoid of frequency $f = 0.125$. The non-zero region of the signal lasts for 128 points (there are 256 total points in the time series). Part (b) shows the amplitude spectrum in the region $f = 0.1$ to $f = 0.15$. One can see the width of the central lobe of the peak, which limits the resolution which one can resolve frequency components, and the familiar sinc function sidelobes. When one adds more points to the time series by zero padding (Figure 11c), the width of the central lobe does not change, it is merely sampled more finely (Figure 11d). It is only when the signal itself exists for a long time that extending the integration range (as in Figure 11e) of the Fourier Transform will actually increase the frequency resolution (Figure 11f). This fact comes into play with the S-Transform. It is suggested that the proper method to define resolution is to examine the time-frequency space defined by the S-transform. Only if S-transform examination shows that the signal lasts for several periods would it be appropriate to average over a longer time interval. Frequency resolution depends on the length of the signal, not the length of the measurement, therefore the frequency dependent resolution of the S-transform is appropriate.

8.3.4 The S Transform and Generalized Instantaneous Frequency

It can be shown that the S-Transform provides an extension of instantaneous frequency to broadband signals [*Bracewell,* 1978]. A particular voice of the S-Transform can be written as [*Stockwell et al.,* 1996a,c; *Stockwell,* 1999; *Mansinha et al.,* 1997b]

$$ST(\tau, f_o) = A(\tau, f_o)\, e^{i\Phi(\tau, f_o)} \tag{245}$$

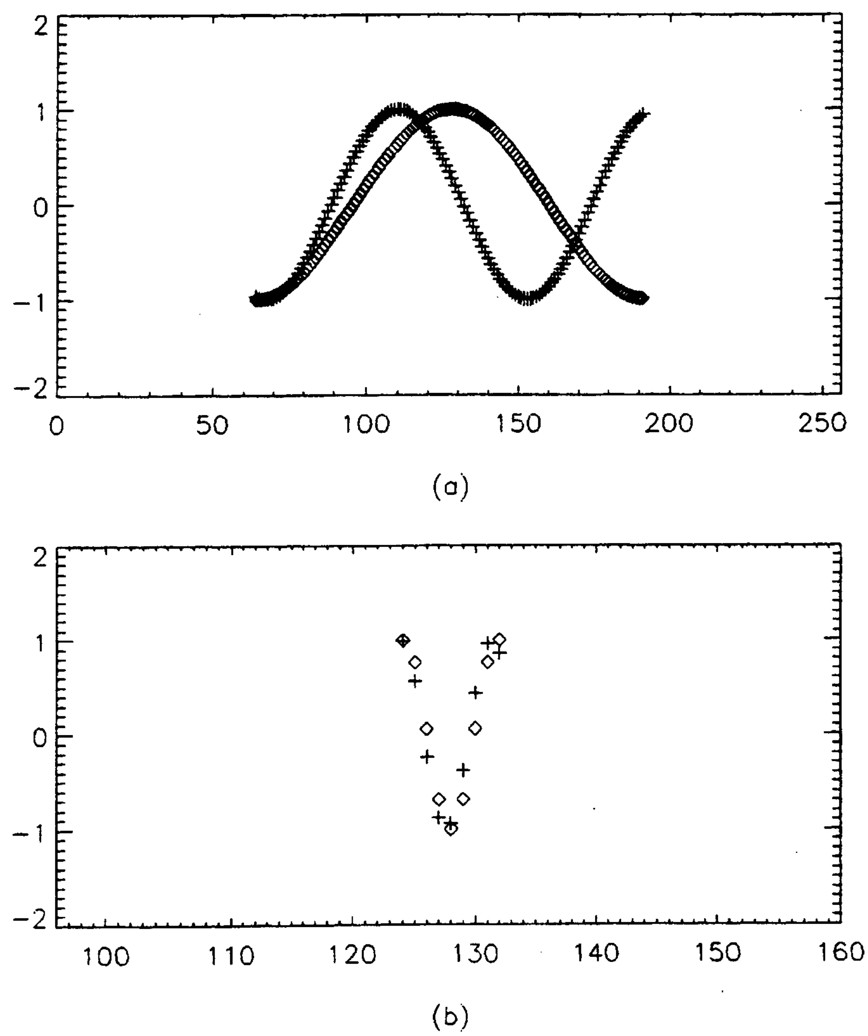

Figure 9: (a) One period of a signal of frequency 2/256 and of a signal with frequency = 3/256. The difference in period is easily observed. There are some 42 points between trough of the 3/256 signal, and that of the 2/256 signal. (b) One period of a signal of frequency 33/256 and of a signal with frequency = 34/256. The different period is very difficult to detect. Thus the frequency dependent resolution of the S-transform is explained (note the change in the x-axis.

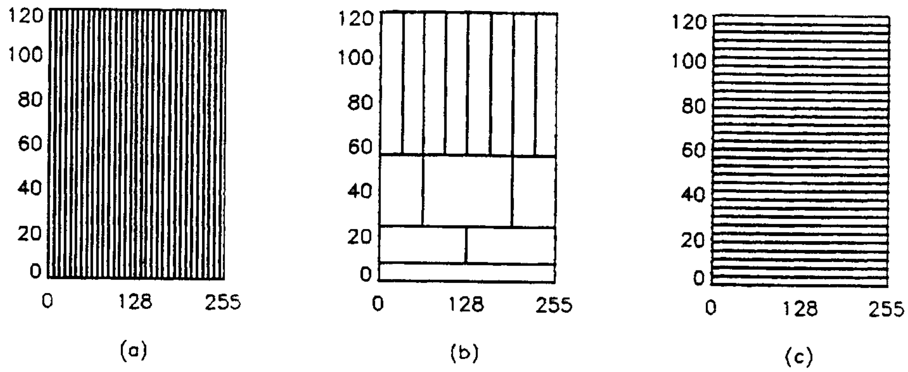

Figure 10: (a) The resolution of time frequency space when the signal is viewed as a time series, having precise time resolution, and no frequency resolution. (b) The resolution of time frequency space when the signal is viewed in the S-Transform representation, having frequency dependent resolution. (c) The resolution of time frequency space when the signal is viewed as a spectrum, having no time resolution, and precise frequency resolution [after Stockwell, 1999].

where

$$A(\tau, f_o) = \sqrt{\Re\{S(\tau, f_o)\}^2 + \Im\{S(\tau, f_o)\}^2} \tag{246}$$

and

$$\Phi(\tau, f_o) = \arctan\left(\frac{\Im\{S(\tau, f_o)\}}{\Re\{S(\tau, f_o)\}}\right) \tag{247}$$

Since a voice isolates one specific component, one may use the phase in equation (245) to determine the instantaneous frequency [*Bracewell,* 1978]

$$IF(\tau, f_o) = \frac{1}{2\pi}\frac{\partial}{\partial\tau}\{2\pi\tau f_o + \Phi(\tau, f_o)\} \tag{248}$$

The validity of equation (248) can be seen for the simple case of $s(t) = \cos(2\pi\omega t)$ where the phase function $\Phi(\tau, f) = 2\pi(\omega - f)\tau$.

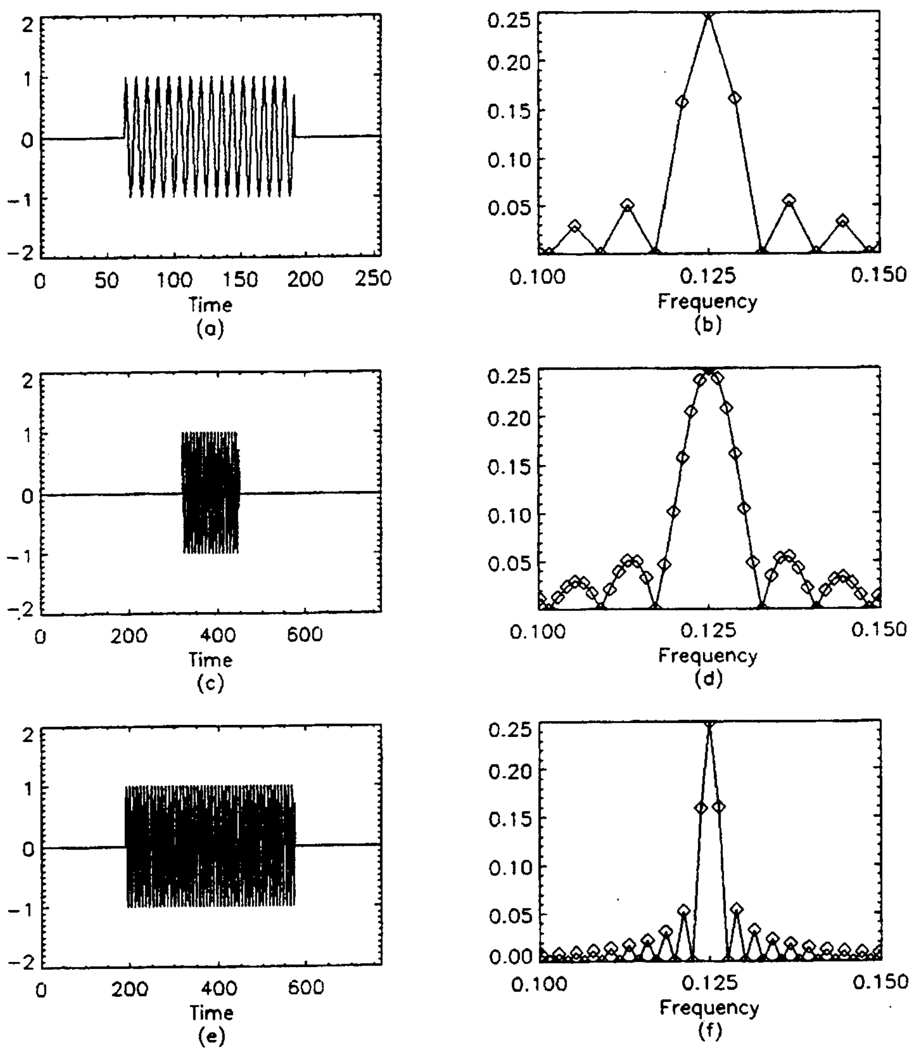

Figure 11: A 256 point sinusoidal time series with a frequency of 32/256. It is non-zero for a region of 128 points. (b) A region of the amplitude spectrum of the time series in (a). (c) The same time series in (a) but now zero padded up to a length of 768. Note the change in the x-axis. (d) A region of the amplitude spectrum of the time series plotted in (c). The resolution has not changed, but the frequency sampling interval is smaller. (e) The time series in (a) but now exists (non-zero) for 384 points. (f) A region of the amplitude spectrum of the time series plotted in (e). Because the signal has lasted for a longer time interval, it is possible to measure the frequency with much higher resolution than in (d).

8.3.5 Linearity

The S-Transform is a linear operation on the time series $s(t)$. This is important for the case of additive noise in which one can model the data as $data(t) = signal(t) + noise(t)$ and thus the operation of the S-Transform leads to [*Stockwell et al.*, 1996c]

$$ST\{data\} = ST\{signal\} + ST\{noise\} \tag{249}$$

This is an advantage over the bilinear class of time-frequency representations (TFRs) where one finds [*Cohen,* 1989]

$$TFR\{data\} = TFR\{signal\} + 2 \star TFR\{signal\} \star TFR\{noise\} + TFR\{noise\} \tag{250}$$

8.3.6 The Generalized S-Transform

Definition (234) was derived under the assumption that the width σ of the Gaussian modulation function is proportional to the inverse of frequency. Thus the Gaussian window function is [*Stockwell et al.*, 1996c; *Stockwell,* 1999; *Mansinha et al.*, 1997b]:

$$g(t, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \tag{251}$$

with

$$\sigma = \frac{k}{f} \tag{252}$$

Normally k is set to unity as in definition (234), providing approximately one modulated sine and cosine cycle. However k may be increased for increasing frequency resolution, with a corresponding loss of resolution in time. Expression (234) is then modified slightly:

$$ST(\tau, f) = \int_{-\infty}^{\infty} s(t) \frac{\left|\frac{f}{k}\right|}{\sqrt{2\pi}} e^{-\frac{(t-\tau)^2 f^2}{2k^2}} e^{-i2\pi f t} \, dt \tag{253}$$

The inverse S-transform is still found by equation (244).

8.4 The Discrete S-Transform

The discrete analog of equation (242) is used to calculate the discrete S-Transform by taking advantage of the efficiency of the FFT (Fast Fourier Transform) and the convolution theorem.

Let $s[kT]$, $k = 0, l, ..., N-1$ denote a discrete time series, corresponding to $s(t)$, with a time sampling interval of T. The discrete Fourier Transform is given by [*Brigham,* 1975]

$$S[\frac{n}{NT}] = \frac{1}{N} \sum_{k=0}^{N-1} s[kT]\, e^{-\frac{i2\pi nk}{N}} \tag{254}$$

where n = 0,1,...,$N-1$. In the discrete case, the S-Transform is the projection of the vector defined by the time series $s[kT]$ onto a spanning set of vectors. The spanning vectors are not orthogonal, and the elements of the S-Transform are not independent. Each basis vector (of the Fourier Transform) is divided into N localized vectors by an element-by-element product with the N shifted gaussians, such that the sum of these N localized vectors is the original basis vector.

Using the discrete analog of equation (242), the S-Transform of a discrete time series $s[kT]$ is given by (letting $f \to n/NT$ and $T \to jT$) [*Stockwell et al.,* 1996b,c; *Mansinha et al.,* 1997b]

$$ST[jT, \frac{n}{NT}] = \sum_{m=0}^{N-1} S[\frac{m+n}{NT}]\, e^{-\frac{2\pi^2 m^2}{n^2}}\, e^{\frac{i2\pi mj}{N}} \tag{255}$$

and for the $n = 0$ voice it is equal to the constant defined by [*Stockwell et al.,* 1996c]

$$ST[jT, 0] = \frac{1}{N} \sum_{m=0}^{N-1} s(\frac{m}{NT}) \tag{256}$$

where j, m and $n = 0, 1, ..., N-1$. Equation (360) puts the constant average of the time series into the zero frequency voice, thus assuring the inverse is exact for the general time series. The discrete S-Transform suffers the familiar problems from sampling and finite length, giving rise to implicit periodicity in the time and frequency domains. The convolution operations are implicitly 'wrap-around', giving rise to edge effects.

The calculation of the S-Transform is very efficient, using the convolution theorem both ways, each to the advantage, and utilizing the efficiency of the Fast Fourier Transform algorithm. An approximate count of the number of operations is given in the parentheses after each numbered section below (letting $n = n/NT$, $m = m/NT$, $k = kT$ and $j = jT$). Thus using equation (256)

1. Fourier transform the original time series $s[k]$, with N points and sample interval T, to give $S[m]$ using a FFT (Fast Fourier Transform) routine. This is only done once ($NlogN$ operations) [*Stockwell et al.*, 1996c; *Stockwell*, 1999].

2. Calculate the localizing Gaussian $G[n, m]$ for the required frequency n, (N assignment statements).

3. Shift the spectrum $S[m]$ to $S[(m+n)]$ for the frequency n. (One pointer addition) [1st use of the convolution theorem].

4. Multiply $S[(m+n)]$ by $G[n, m]$ to give $B[n, m]$. (N multiplications) [2nd use of the convolution theorem].

5. Inverse Fourier transform $B[n, m]$ m to j to give the row of $ST[n, j]$ corresponding to the frequency n. ($NlogN$ operations).

6. Repeat steps 3, 4 and 5 until all the rows of $ST[n, j]$ corresponding to all discrete frequencies n have been defined.

The computational efficiency of the Fast Fourier transform has been used whenever possible. The total number of operations is approximately $N(N + NlogN)$ operations. Considering that the S-Transform has N^2 points, the number of operations per computed transform point is the same as the standard FFT when the time series is multiplied by an apodizing function ($N + NlogN$ operations).

8.5 Discrete Inverse S-Transform

The discrete inverse of the S-Transform is performed through the intermediate step of computing the discrete Fourier Transform. Summing the S-matrix along the voices (rows) gives ($n \neq 0$) [*Stockwell et al.*, 1996c; *Stockwell*, 1999]

$$\sum_{j=0}^{N-1} ST[\frac{n}{NT}, jT] = \sum_{j=0}^{N-1} \sum_{m=0}^{N-1} S[\frac{m+n}{NT}]\, e^{-\frac{2\pi^2 m^2}{n^2}}\, e^{\frac{i2\pi mj}{N}} \tag{257}$$

Reordering the sequence of summation, we have

$$\sum_{j=0}^{N-1} ST[\frac{n}{NT}, jT] = \sum_{m=0}^{N-1} S[\frac{m+n}{NT}]\, e^{-\frac{2\pi^2 m^2}{n^2}} \sum_{j=0}^{N-1} e^{\frac{i2\pi mj}{N}} \tag{258}$$

By the orthogonal property, the sum over j is zero unless $m =$ zero in which case it is equal to N. Thus the average of the voices of $ST(n/NT, jT)$ is [*Stockwell et al.*, 1996c; *Stockwell*, 1999]

$$\sum_{j=0}^{N-1} ST[\frac{n}{NT}, jT] = \sum_{m=0}^{N-1} N\delta_{m,0}\, S[\frac{m+n}{NT}]\, e^{-\frac{2\pi^2 m^2}{n^2}} \tag{259}$$

That is,

$$\frac{1}{N} \sum_{j=0}^{N-1} ST[\frac{n}{NT}, jT] = S[\frac{n}{NT}] \tag{260}$$

Therefore the discrete inverse of the S-Transform is ($\forall n$):

$$s[kT] = \frac{1}{N} \sum_{n=0}^{N-1} \left\{ \sum_{j=0}^{N-1} ST[\frac{n}{NT}, jT] \right\} e^{\frac{i2\pi nk}{N}} \tag{261}$$

In the limit at $n = 0$, the width of the voice Gaussian decreases to zero. The zero frequency is the average of the time series and is constant. The value of $ST(n/NT, jT)$ for $n = 0$ is simply the average of $s(kT)$. Every value along the voice for $n = 0$ can be filled with this value. In the representation equation (255), the voice Gaussian function at $n = 0$ is replaced with the Kronecker delta function $\delta_{m,0}$. The S-Transform is exactly invertible.

8.6 Comparison of Time-Frequency Representations

Figure 12 shows a chirp signal. In part (b) the Instantaneous Frequency calculated using the Hilbert transform is shown. It accurately represents the linearly increasing frequency. There are some edge effects due to the implicit wrap around of the time series, in the Fourier representation used to calculate the Hilbert Transform. Part (c) shows the amplitude of the S-Transform of the time series in (a). It also accurately represents the chirp signal. The edges have been removed with a 5% Hanning window taper. Part (d) is the Short-Time Fourier Transform calculated using a Gaussian window (for comparison with the S-Transform) with a standard deviation of 8 points. Part (e) shows the Wigner Distribution of the chirp signal. All methods faithfully represent the chirp signal.

Figure 13 shows the combination of a chirp signal with an increasing frequency, with that of one with a decreasing frequency. It is interesting to see that, in the time domain, the nature of the signal is not at all obvious, as it has been with previous examples. In part (b) the Instantaneous Frequency calculated using the Hilbert transform is shown. It performs quite poorly with the multi-component signal. Part (c) shows the amplitude of the S-Transform of the time series in (a). It also accurately represents the crossing chirp signal. Part (d) is the Short-Time Fourier Transform calculated using a Gaussian window with a standard deviation of 8 points. It does a nice job of representing the signal. Part (e) shows the Wigner Distribution of the crossed chirp signal. While the chirps are detectable, the cross-terms greatly reduce the reliability of detecting them.

Figure 14 shows a sample time series composed of 3 distinct local frequency components. The first half has a low frequency, the second half has a high frequency, and in the first half a high frequency component is added. In part (b) the Instantaneous Frequency calculated using the Hilbert transform is shown. Again this performs poorly with the multi-component signal. In the region where two frequencies are present, the instantaneous frequency oscillates wildly. Part (c) shows the amplitude of the S-Transform of the time series in (a). It accurately represents all three components. The low frequency component has very good frequency resolution, while the high frequency transient signal has very good time resolution, and is reliably detected. Part (d) is the Short-Time Fourier Transform calculated using a

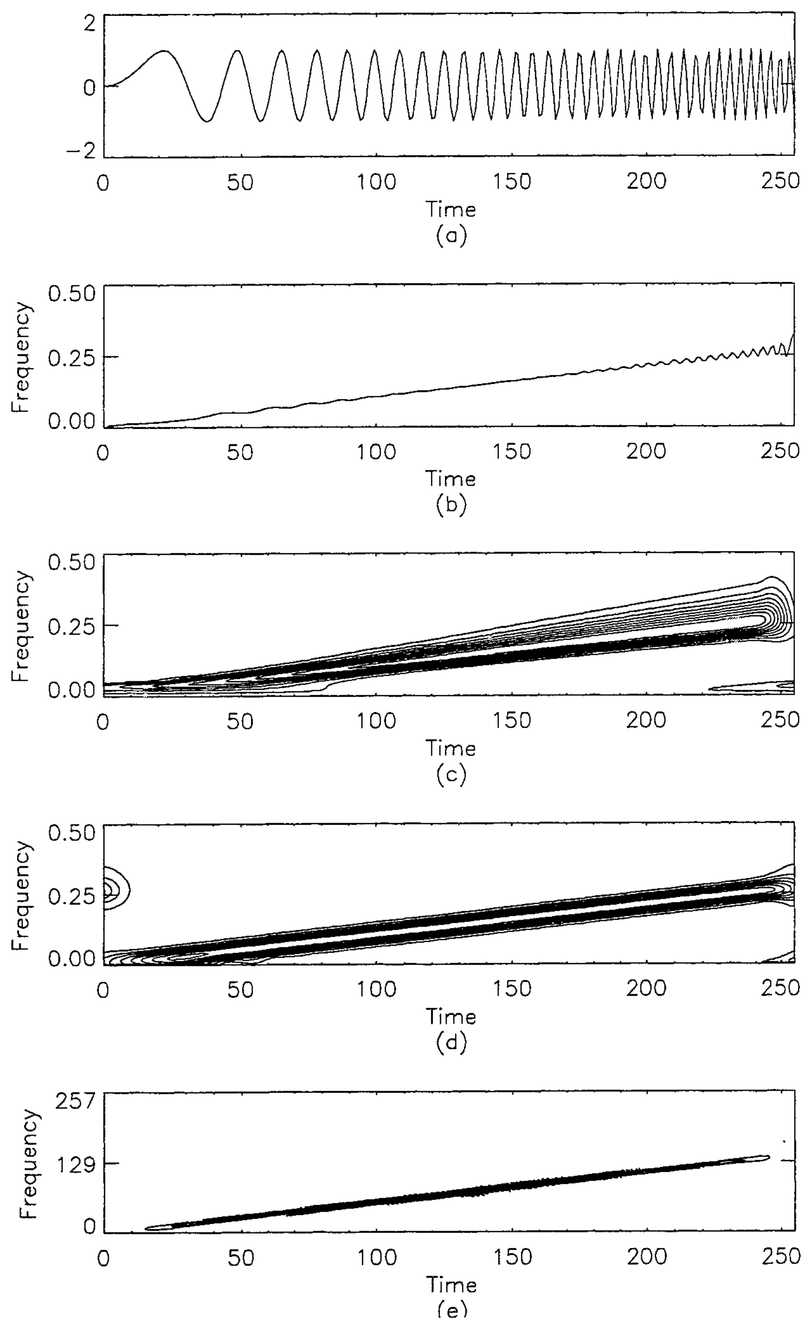

Figure 12: A chirp function with a linearly increasing frequency. (b) The instantaneous frequency of the time series in (a). (c) Amplitude of the S-transform for the time series in (a). (d) The generalized instantaneous frequency calculated from the S-transform for the time series in (a). (e) Amplitude of the STFT for the time series in (a). (f) The Winger distribution for the time series in (a). For such a simple and clean time series, all methods do a very good job in finding the frequency chirp.

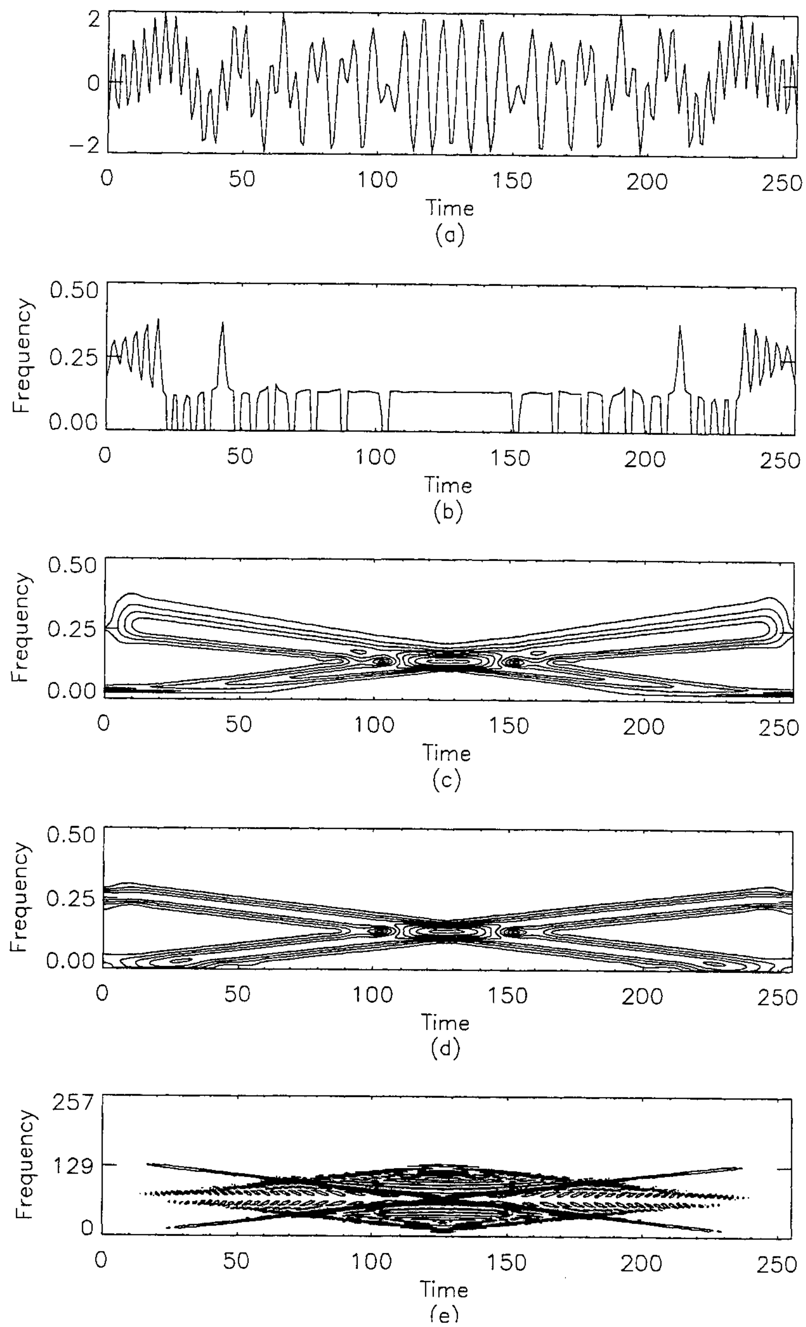

Figure 13: Another chirp function with a linearly increasing frequency added to one with a decreasing frequency. (b) The instantaneous frequency of the time series in (a). (c) Amplitude of the S-transform for the time series in (a). (d) The generalized instantaneous frequency as calculated from the S-transform for the time series in (a). (e) Amplitude of the STFT for the time series in (a). (f) The Winger distribution for the time series in (a).

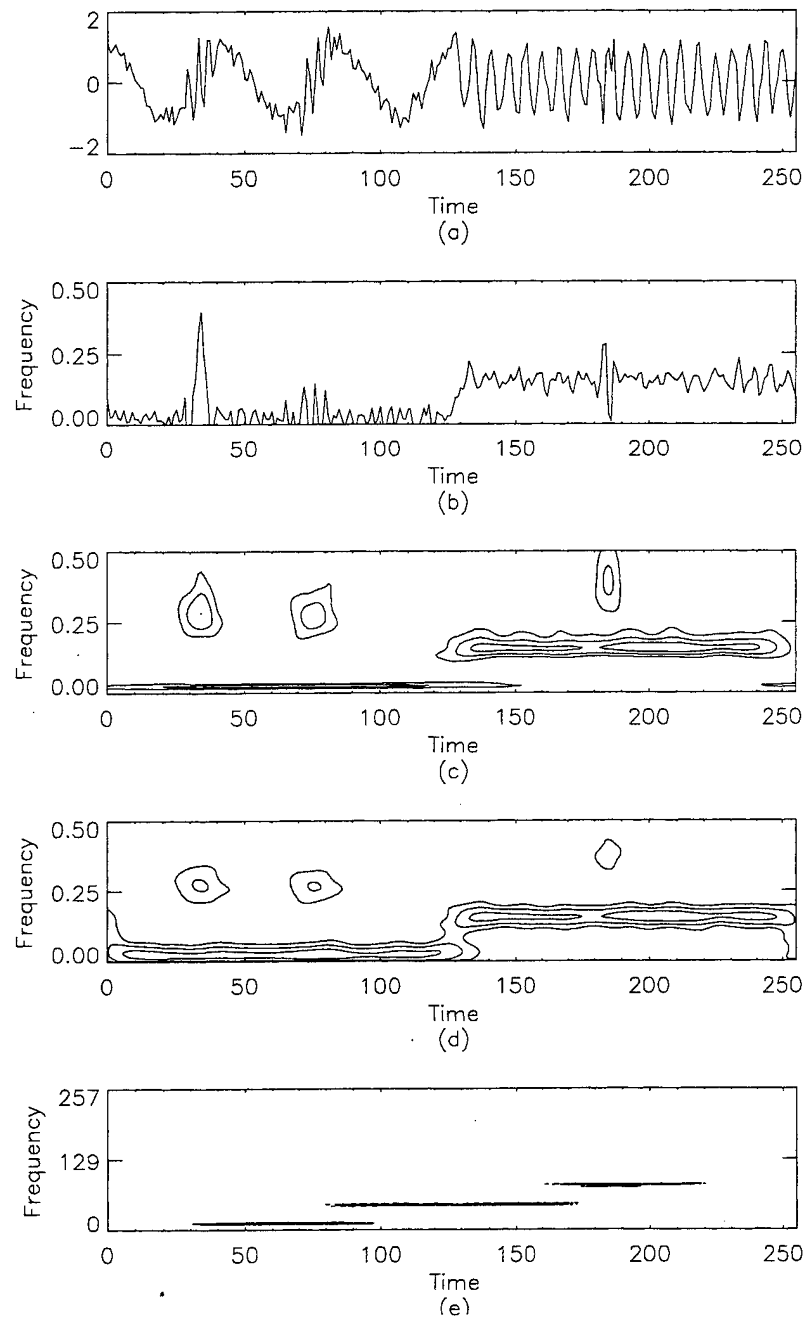

Figure 14: A sample time series with 3 distinct components. (b) The instantaneous frequency of the time series in (a). (c) Amplitude of the S-transform for the time series in (a). (d) Amplitude of the STFT for the time series in (a). (e) The Winger distribution for the time series in (a).

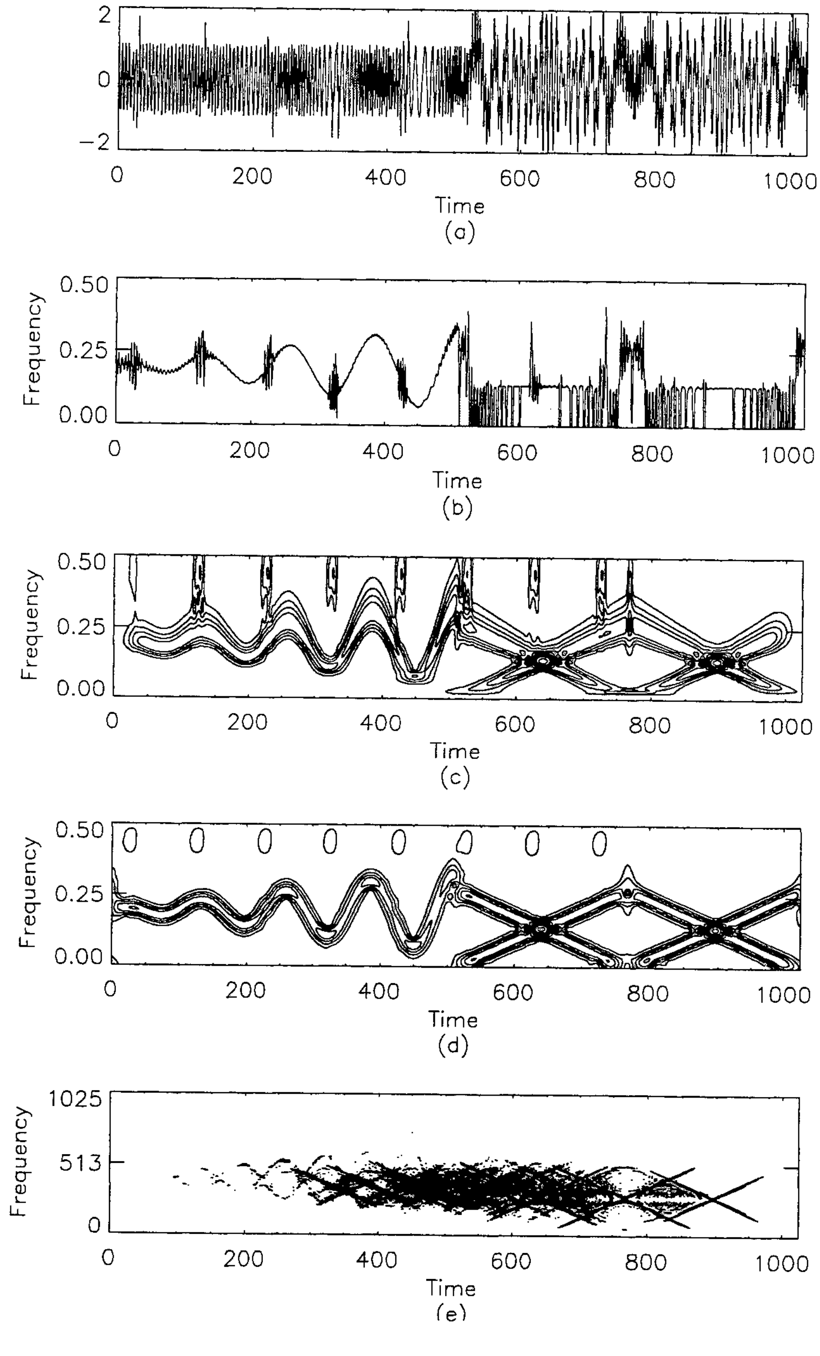

Figure 15: A complicated tiem series with a sinusoidally modulated frequency in the first 512 points, then two sets of crossing chirps in the next 512 points. All throughout the time series, high frequency bursts of short duration are added. (b) The instantaneous frequency of the time series in (a). This identifies the sinusoidal modulation in the first half well, but misses the bursts and the crossed chirps. (c) Amplitude of the S-transform for the time series in (a). This identifies all features. (d) Amplitude of the STFT for the time series in (a). This does a good job, however the high frequency bursts are under represented. (e) The Winger distribution for the time series in (a). The signal is poorly represented.

Gaussian window with a standard deviation of 8 points. This does detect both the low frequency signals, but the high frequency transient is washed out at this resolution due to the fact that the constant width window is averaging over a long period of time. The low frequency component is represented with poor frequency resolution. Part (e) shows the Wigner Distribution of the crossed chirp signal. While the two low frequencies are detectable, the cross-terms greatly reduce the reliability of detecting them. Also the high frequency transient is not detected.

Figure 15 shows a sample time series composed of 3 distinct local frequency components as in Figure 14 but now with a noise component of standard deviation 0.5 added. The S-Transform does detect the signal, but the other methods perform poorly. For simple signals the Wigner distribution does very well, but so does the instantaneous frequency. Because of the simplicity of the instantaneous frequency, it should perhaps be attempted first. For noisy signal, or ones with complicated time frequency structure, in order to be sure of the representation one should use the S-Transform.

In studying the various techniques, it is apparent that every one has its strengths and weaknesses. For locally narrowband signals the Instantaneous Frequency and the Wigner Distribution work very well. For most time series, the Short-Time Fourier Transform does give a reasonable representation of the time frequency structure. However, for the general non-stationary time series, the S-Transform does a better job of simultaneously localizing all the frequency components.

9 Conclusion

The time-frequency representations are powerful tools for the analysis and processing of non-stationary signals for which separate time-domain and frequency-domain analyses are not adequate. In this report, we outline the motivations, interpretations, mathematical fundamentals, and properties of linear and quadratic time-frequency representations.

This report is reviewing both linear and quadratic time-frequency representations. The linear representations discussed are Short-Time Fourier Transform and S-transform. The quadratic representation discussed is the Wigner distribution. We also compare these three different time-frequency analysis techniques and show that each technique has its strengths and drawbacks. Simulated data sets have been used for the comparison. For locally narrow band signals the Instantaneous Frequency and Wigner distribution methods work very well. For most time series, the Short-Time Fourier Transform does give a reasonable representation of the time-frequency structure. However, for the general non-stationary time series, the S-transform does a better job of simultaneously localizing all the frequency components.

Although we have attempted to provide a coherent framework of time-frequency representations, a truly unified framework is difficult to obtain because the large variety of exiting methods and approaches cause the field of time-frequency analysis to be somewhat disparate. It is clear that the choice of the particular time-frequency representation depends upon the specific area of application and what we aim to achieve with a local frequency analysis. We show that time-frequency analysis methods should enable us to classify signals with a considerably greater reflection of the physical situation than can be achieved by the conventional Fourier transform method alone.

Our main areas of application of the time-frequency representation are the analysis of experimental HF radar data and ISAR (Inverse Synthetic Aperture Radar) image data. We will examine and compare different time-frequency representations discussed in this report to detect accelerating target such as aircraft by a HF radar in the presence of clutter background.

Acknowledgments

I would like to thank Dr. B. Stockwell for his comments and suggestions for improvement of this work. I also wish to thank him for supplying me with the S-transform algorithms.

References

[1] Allen, J. B. and L. R. Rabiner, A unified approach to short-time Fourier analysis and synthesis, *Proc. IEEE.* 65, 1558-1564, 1977.

[2] Altes, Detection, estimation and classification with spectrograms, *J. Acoust. Soc. Am.*, 67, 1232-1246, 1980.

[3] Armstrong, E. H., A method of reducing disturbances in radio signaling by a method of frequency modulation, *Proc. IRE.*, 24, 669-740, 1936.

[4] Bedrosian, E., A product theorem for Hilbert transforms, *Proc. IEE.*, 51, 686-689, 1963.

[5] Bracewell, R. N., The Fourier Transform and its applications, McGraw-Hill, New York, 1978.

[6] Bloomfield, P., Fourier analysis of time series, *Wiley*, 1976.

[7] Brigham, E. O., The fast Fourier transform, *Prentice-Hall Inc.*, New Jersey, 1975.

[8] Carson, J. R. and T. C. Fry, Variable frequency electric circuit theory with application to the theory of frequency modulation, *Bell Sys. Tech. J.*, 16, 513-540, 1937.

[9] Cartwright, N., Correlations without joint distributions in quantum mechanics, *Found. Phys.*, 4, 127-136, 1974.

[10] Choi, H. I. and W. J. Williams, Improved time-frequency representation of multicomponent signals using exponential kernels, *IEEE transactions on Acoustics, Speech and Signal Processing*, 37, 862-871, 1989.

[11] Claasen, T. A. C. M. and W. F. G. Mecklenbrauker, The Wigner Distribution - a tool for time-frequency signal analysis; part-II: discrete time signals", *Phillips J. Res.*, 35, 276-300, 1980a.

[12] Claasen, T. A. C. M. and W. F. G. Mecklenbrauker, The Wigner distribution - a tool for time-frequency signal analysis; part III: relations with other time-frequency signal transformations, *Phillips J. Res.*, 35, 372-389, 1980b.

[13] Claasen, T. A. C. M. and W. F. G. Mecklenbrauker, The aliasing problem in discrete time Wigner distributions, *IEEE Trans. Acoust., Speech, Signal Processing*, 31, 1067-1072, 1983.

[14] Cohen, L., Generalized phase-space distribution functions, *J. Math. Phys.,* 7, 781-786, 1966.

[15] Cohen, L., On a fundamental property of the Wigner distribution, *IEEE Trans. Acoust., Speech, Signal Processing,* 35, 559-561, 1987a.

[16] Cohen, L., Winger distribution for finite duration or band-limited signals and limiting cases, *IEEE Trans. Acoust., Speech, Signal Processing,* 35, 796-806, 1987b.

[17] Cohen, L. and C. Lee, Instantaneous frequency, its standard deviation and multicomponent signals, *Proc. SPIE.,* 975, 186-208, 1988.

[18] Cohen, L., Time-frequency distribetions - a review, *Proc. IEEE,* 77, 7, July, 1989.

[19] Cohen, L., Time-frequency analysis, *Prentice Hall Signal Processing Series,* Prentice-Hall Inc., 7, New Jersey, 1995.

[20] Crochire, R. E. and L. R. Rabiner, Multi-rate digital signal processing, *Prentice-Hall Inc.,* New Jersey, 1983.

[21] Daubechies, I., The wavelet transform, time-frequency localization and signal analysis, *IEEE transactions on Information Theory,* 36 (5), 395-457, 1990.

[22] De Bruijn, N. G., Uncertainty principles in Fourier analysis, Inequalities, *Academic Press,* New York, 1967.

[23] Dziewonski, A., S. Block, and M. Landisman, A technique for the analysis of transient signals, *Bull. Seism. Soc. Am.,* 59, 427-444, 1969.

[24] Eramian, M., R. A. Schincariol, R. G. Stockwell, R. P. Lowe, and L. Mansinha, Review of applications of 1-d and 2-d S-transform, *Wavelet Applications IV*, 3078, 1996

[25] Eramian, M., R. A. Schincariol, L. Mansinha, and R. G. Stockwell, Generation of aquifer heterogeneity maps using two dimensional spectral texture segmentation technique, *Mathematical Geology,* 31 (3), 327-348, 1999.

[26] Fink, L. M., Relations between the spectrum and instantaneous frequency of a signal, *Problems Inform. Transm.,* 2, 11-21, 1966.

[27] Folland, G. B. and A. Sitaram, The uncertainty principle: a mathematical survey, *J. Fourier Anal. Appl.*, 3 (3), 207-238, 1997.

[28] Fritts, D. C., D. M. Riggin, B. B. Balsley, and R. G. Stockwell, Recent results with an MF radar at mcmurdo, antarctica: chracteristics and variability of motions near 12-hour period in the mesosphere, *Geophysical Research Letters,* 25 (3) 297-300, 1998.

[29] Gabor, D., Theory of communication, *J. IEE (London),* 93, 429-457, 1946.

[30] Gupta, M. S., Definitions of instantaneous frequency and frequency measurability, *Am. J. Phys.,* 43, 1087-1088, 1975.

[31] Hlawatsch, F. and G. F. Boudreaux-Bartels, Linear and quadratic time-frequency signal representations, IEEE Signal Processing Magazine, April, 1992.

[32] Hlawatsch, F., Interference terms in the Wigner distribution, *Digital Signal Processing, Elsevier,* Amsterdam, 1984.

[33] Hlawatsch, F., A note on Wigner distribution for finite duration or band-limited signals and limiting cases, *IEEE Trans. Acoust., Speech, Signal Processing,* 36, 927-929, 1988.

[34] Hlawatsch, F., Time-frequency analysis and synthesis of linear linear signal spaces, Kluwer Academic Publishers, Boston, 1998.

[35] Janssen, A., Optimality property of the Gaussian window spectrogram, *IEEE Transactions on Acoustics, Speech and Signal Processing,* ASSP-39 (1) 1991.

[36] Kodera, K., C. de Villedary, and R. Gendrin, A new method for the numerical analysis of non-stationary signals, *Phys. Earth and Planetary Interiors,* 12, 142-150, 1976.

[37] Lampard, D. G., Generalization of the Wiener-Knintchine theorem to nonstationary processes, *J. Appl. Phys.,* 25, 802, 1954.

[38] Levshin, A. L., V. F. Pisarenko, and G. A. Pogrebinsky, On a frequency-time analysis of oscillations, *Ann. Geophys.,* 28, 211-218, 1972

[39] Qian, S. and D. Chen, Joint time-frequency analysis, *Prentice-Hall Inc.,* New Jersey, 1996.

[40] Mandel, L., Interpretation of instantaneous frequency, *Amer. Jour. Phys.,* 42, 840-846, 1992.

[41] Mansinha, L., R. G. Stockwell, and R. P. Lowe, Pattern analysis with two dimensional spectral localization: application of two dimensional S-transforms, *Pyhsica A,* 239, 286-295, 1997a.

[42] Mansinha, L., R. G. Stockwell, R. P. Lowe, M. Eramian, and R. A. Schincariol, Local spectrum analysis of 1-d and 2-d data, *Physics of the Earth and Planetary Interiors,* 103, 239-336, 1997b.

[43] Mertins, A., Signal analysis, *John Wiley & Sons*, Toronto, 1999.

[44] Moyal, J. E., Quantum mechanics as a statistical theory, *Proc. Cambridge, Phil. Soc.*, 45, 99-124, 1949.

[45] Nawab, S. H. and T. Quatieri, Advanced topics in signal processing, *Prentice-Hall,* New Jersey, 1988. IEEE

[46] Nuttall, A. H., On the Quadrature approximation to the Hilbert transform of modulated signals, *Proc. IEE.,* 54, 1458, 1966.

[47] Oppenheim, A. V., Speech spectrograms using the fast Fourier transform, *IEEE Spectrum,* 7, 57-62, 1970.

[48] Oppenheim, A. V. and R. W. Schafer, Digital signal processing, *Prentice-Hall*, New Jersey, 1975.

[49] Papoulis, A., Signal analysis, *McGraw-Hill Book Company,* New York, 1977.

[50] Portnoff, M. R., Time-frequency representations of digital signals and systems based on short-time Fourier analysis, *IEEE Trans. Acoust. Speech, Signal Proc.,* 28, 55-69, 1980.

[51] Rabiner, L. R. and J. B. Allen, On the implementation of a short-time spectral analysis method for system identification, *IEEE Trans. Acoust. Speech, Signal Proc.,* 28, 69-79, 1980.

[52] Ricaczek, W., Hilbert transforms and complex representation of real signals, *Proc. IEE.,* 54, 434-435, 1966.

[53] Ricaczek, W., Principles of high-resolution radar, *McGraw-Hill,* New York, 1969.

[54] Rioul, O. and M. Vetterli, Wavelets and signal processing, *IEEE signal processing magazine,* 8, 14-38, 1991.

[55] Schroeder, M. R. and B. S. Atal, Generalized short-time power spectra and autocorrelation functions, *J. Acoust. Soc. Am.,* 34, 1679-1683, 1962.

[56] Shekel, J., Instantaneous frequency, *Proc. IRE.,* 41, 548, 1953.

[57] Strang, G. and T. Q. Nguyen, Wavelets and filter banks, *Wellesley Cambridge Press,* New York, 1995.

[58] Stockwell, R. G., R. P. Lowe, and L. Mansinha, Instantaneous wavevector analysis, *Wavelet Applications IV*, 3078, 349-358, 1996a.

[59] Stockwell, R. G., Localization of the complex spectrum: S-transform, *M.Sc. Thesis,* University of Western Ontario, Ontario, 1994.

[60] Stockwell, R. G., S-transform analysis of gravity wave activity, *Ph.D. Thesis,* University of Western Ontario, Ontario, 1999.

[61] Stockwell, R. G., R. P. Lowe, and L. Mansinha, Localized cross spectral analysis with phase corrected wavelets, *Wavelet Applications IV*, 2762, 557-564, 1996b.

[62] Stockwell, R. G., R. P. Lowe, and L. Mansinha, Localization of the complex spectrum: The S-transform, *IEEE Transactions on Signal Processing*, 44 (4) 998-1001, 1996c.

[63] Taylor, J. R., An introduction to error analysis, *Mill Valley,* California, U.S.A., 1982.

[64] Vakman, D. E., On the definition of concepts of amplitude, phase and instantaneous frequency of a signal, *Trans. in Radio Eng. Electron. Phys.,* 754-759, 1972.

[65] Vakman, D. E., Do we know what are the instantaneous frequency and instantaneous amplitude of a signal ? *Trans. in Radio Eng. Electron. Phys.,* 21, 95-100, 1976.

[66] Vakman, D. E., On the analytic signal, the teager-kaiser energy algorithm, and other methods for defining amplitude and frequency, *IEEE transaction on Signal Prosessing,* 44 (4) 791-797, 1996.

[67] Vakman, D. E. and L. A. Vainshtein, Amplitude, phase, frequency - fundamental concepts of oscillation theory, *Sov. Phys. Ush.,* 20, 1002-1016, 1978.

[68] Vakman, D. E., Measuring the frequency of an analytical signal, *Trans. in Radio Eng. Electron. Phys.,* 63-69, 1980.

[69] Van der Pol, B., The foundamental principles of frequency modulation, *J. IEE (London),* 93, 153-158, 1946.

[70] Varanini, M., G. D. Paolis, M. Emdin, A. Macerata, S. Pola, M. Cipriani, and C. Marchesi, Spectral analysis of cardiovascular time series by the S-transform, *Computers in Cardiology,* 24, 383-386, 1997.

[71] Vetterli, M. and J. Kovacevic, Wavelets and subband coding, *Prentice-Hall,* New Jersey, 1995.

[72] Ville, J. Thovrie et applications de la notion de signal analytique, *Cables et Transmission,* 2, 61-74, 1948.

[73] Wigner, E., On the quantum correction for thermodynamic equilibrium, *Phys. Rev.,* 40, 749-759, 1932.

[74] Wilson, R., A. D. Calway, and E. R. S. Pearson, A generalized wavelet transform for Fourier analysis: the multiresolution Fourier and its application to image and audio signal analysis, *IEEE Trans. Info. Theory,* 38, 674-690, 1992.

[75] Woodward, P. M., Probability and information theory with application to radar, *Pergamon Press,* London, 1953.

[76] Young, H. D., Statistical treatment of experimental data, *McGraw-Hill,* New York., 1962.

[77] Young, R. K., Wavelet theory and its applications, *Kluwer Academic Publishers,* Boston, 1993.

[78] Zhao, Y., L. E. Atlas, and R. J. Marks, The use of cone-shaped kernals for generalized time-frequency representations of nonstationary signals, *IEEE Transactions on Acoustics, Speech and Signal Processing,* 38, 1084-1091, 1990.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

   Defence Research Establishment Ottawa
   Department of National Defence
   Ottawa, Ontario, Canada K1A 0Z4

2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable)

   UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C or U) in parentheses after the title.)

   Linear and Quadratic Time-Frequency Representations (U)

4. AUTHORS (Last name, first name, middle initial)

   Thayaparan, Thayananthan

5. DATE OF PUBLICATION (month and year of publication of document)

   November 2000

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

   97

6b. NO. OF REFS (total cited in document)

   78

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

   DREO TECHNICAL MEMORANDUM

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

   Defence Research Establishment Ottawa
   Department of National Defence
   Ottawa, Ontario, Canada K1A 0Z4

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

   05AB11

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

   DREO TECHNICAL MEMORANDUM 2000-080

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

   (X) Unlimited distribution
   ( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
   ( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
   ( ) Distribution limited to government departments and agencies; further distribution only as approved
   ( ) Distribution limited to defence departments; further distribution only as approved
   ( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

(U) This report is reviewing both linear and quadratic time-frequency representations. The linear representations discussed are Short-Time Fourier Transform and S-transform. The quadratic representation discussed is Wigner distribution. We outline the motivations, interpretations, mathematical fundamentals, properties, and applications of these linear and quadratic time-frequency representations. We also compare these three different time-frequency analysis techniques and show that each technique has its strengths and drawbacks. The simulated data sets have been used for the comparison. The choice of the particular time-frequency representation depends upon the specific area of application and what we aim to achieve with a local frequency analysis. We show that time-frequency analysis methods should enable us to classify signals with a considerably greater interpretation of the physical situation than can be achieved by the conventional Fourier Transform method alone.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Time-Frequency Analysis
Short Time Fourier Transform
Wigner Distribution
S-Transform
Fourier Transform
Uncertainty Principle
Instantaneous Frequency
Group Delay
Analytical Signal
Hilbert Transform
Power Spectrum
Auto-Correlation Function
Characteristic Function
Marginals

Defence R&D Canada

is the national authority for providing
Science and Technology (S&T) leadership
in the advancement and maintenance
of Canada's defence capabilities.

R et D pour la défense Canada

est responsable, au niveau national, pour
les sciences et la technologie (S et T)
au service de l'avancement et du maintien des
capacités de défense du Canada.

DEFENCE DÉFENSE

www.drdc-rddc.dnd.ca