# Sonar Data Characterisation and Analysis

*Mika Levonen*


A thesis submitted for the degree of Doctor of Philosophy.
**The University of Edinburgh**.
July 2005

# Abstract

This thesis is concerned with sonar signal processing, in particular the statistical characterisation of sonar data. It proposes a number of signal processing methods that are suitable for sonar data. More specifically it deals with the signal processing of time series, from a sonar system, which were collected during a number of experiments in the Baltic Sea. The return signals in sonar can be viewed as a mixture of deterministic and stochastic parameters. Consequently, instead of trying to model the wave propagating environment itself this thesis will illustrate a variety of characteristic properties of the signals and propose suitable methods for solving the problems.

In Chapter 3 a statistical characterisation of sonar data is presented, different aspects of statistical properties of sonar data are addressed. The focus is on the level of stationarity of both active and passive sonar. For active sonar the issue of ping to ping stationarity, for both the actual ping and for the reverberation tail, are examined and found to be non-stationary. However, for some cases it is seen that there are consecutive pings that are stationary. This leads to the conclusion that it can be beneficial to use several pings from the same target. However it is necessary to remember that there is a large variation in the number of consecutive stationary pings. For passive sonar both ambient noise and tonals emitted from surface vessels are investigated. It is found that for all cases tested the sonar data is non-stationary. That is speaking of stationarity in the strict sense. There are however parts in the data that exhibit more stationary behaviour then the data in general.

The stationarity length of the ambient noise data is also examined, using data from a multi-sensor trial (almost 800 data files, see Section 2.5.1). A large proportion of the data set had a stationarity time of roughly 0.4 seconds, or slightly longer. This is also seen in data from the fibre glass boat trial (Section 2.5.6). The data seems to be stationary for about 0.4 seconds. Testing the passive data for symmetry and linearity show that the data is mostly linear, and symmetric.

The ambient noise is found to have more stationary properties at higher frequencies. This leads to the conclusion that it is more beneficial to operate sonar systems at higher frequencies, in environments that are equivalent to Baltic Sea.

In Chapter 4 a number of time-frequency methods are mentioned. Data analysis using the traditional Short Time Fourier Transform, and the Short Time Fractional Fourier Transform is presented. Performance comparisons are made between the two different transforms and it is shown that matching the transform to the nature of the data improves the quality of the time-frequency images.

In Chapter 5 the focus is on data conditioning methods. First a method using outlier rejection applied to bicoherence measures is discussed and the results of applying it to sonar data are presented. Then empirical mode decomposition analysis is used as a data conditioning tool for the two data sets and the results show that the EMD algorithm does not necessarily add to the traditional time/frequency analysis. Never the less, EMD can be a powerful tool for data where the instantaneous frequency varies slowly and the SNR is lower, such as examining signatures from ships.

# Declaration of originality

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the School of Electronics and Electrical Engineering at The University of Edinburgh, and at FOI, the Swedish Defence Research Agency.

Mika Levonen

# Acknowledgements

A thank you goes to my supervisors Stephen McLaughlin and Leif Persson. They have been a great support all the way through the process of writing this thesis, and during the three years of research.

Thank you parents, Margit and Markku Levonen, for putting me on this planet and knowing me far better than I do. Without you I would never even considered going to university in the first place.

Thanks to my employer, FOI, for sponsoring the project. And thanks to all colleagues at FOI for your support, bright ideas, and for the good company.

And then to the department of thanks that goes out to the people whose names have been left out to protect the innocent, however it is the authors belief that the ones mentioned will recognise them selves in their own right.

Cheers for the unforgettable guitar sessions, the weight lifting and the warm welcome to Edinburgh when I first arrived at the office on the fourth floor. I couldn't possibly forget how warm and cosy the crossfire between two distinguished gentlemen on opposite sides of the room made me feel. It was like being caught in the middle of a brass band.

A thanks to all the pg's (and everyone else) in the sas group for mostly banter, but every once in a while a flicker of enlightenment and wisdom, as well as the highly entertaining discussions, parties, pub sessions, and hill walks.

# Contents

# List of figures

# List of tables

# Acronyms and abbreviations

| | |
|---|---|
| AG | Array Gain |
| cdf | cumulative distribution function |
| CPA | Closest Point of Approach |
| d | detection index |
| DFT | Discrete Fourier Transform |
| dof | degrees of freedom |
| DI | Directivity Index |
| DT | Detection Threshold |
| EE | Echo Excess |
| EL | Echo Level |
| EMD | Empirical Mode decomposition |
| FFT | Fast Fourier Transform |
| FOI | The Swedish Defence Research Agency |
| FOM | Figure Of Merit |
| FrFT | Fractional Fourier Transform |
| FT | Fourier Transform |
| IMF | Intrinsic Mode Function |
| IT | Inner Triangle |
| NL | Noise Level |
| NML | Noise-Masking Level |
| OT | Outer Triangle |
| pdf | probability density function |
| PF | Performance Figure |
| RL | Reverberation Level |
| RML | Reverberation-Masking Level |
| SE | Signal Excess |
| SL | Source Level |
| SSP | Sound Speed Profile |
| SSS | Strict Sense Stationarity |

| | |
|---|---|
| STFT | Short Time Fourier Transform |
| SVP | Sound Velocity Profile |
| TFD | Time Frequency Distribution |
| TL | Transmission Loss |
| TS | Target Strength |
| VBR | Variable Bit Rate |
| WSS | Wide Sense Stationarity |

# Nomenclature

| | |
|---|---|
| $x(t)$ | Timeseries or signal |
| $\Phi_X(\xi)$ | The characteristic function for a random vector X |
| $M_X(\lambda)$ | The moment generating function for a random vector X |
| $K_X(\lambda)$ | The cumulant generating function for a random vector X |
| $\kappa_n(X) = \kappa_n(X_1, ..., X_n)$ | The $n^{th}$order cumulant of the vector $X = [X_1, ..., X_n]^T$ |
| $\mu_n(X) = \mu_n(X_1, ..., X_n)$ | The $n^{th}$order moment of the vector $X = [X_1, ..., X_n]^T$ |
| $\mu$ | The mean of $x(t)$ |
| $\sigma^2$ | The variance of $x(t)$ |
| $\gamma_1$ | The skewness of $x(t)$ |
| $\gamma_2$ | The kurtosis of $x(t)$ |
| $B(k,l)$ | The bispectrum at frequencies $(k,l)$ |
| $b^2(k,l)$ | The bicoherence at frequencies $(k,l)$ |
| $s^2(k,l)$ | The skewness function at frequencies $(k,l)$ |
| $P(t,\omega)$ | Time/frequency distribution |
| $P_{SP}(t,\omega)$ | The spectrogram |
| $W(t,\omega)$ | The Wigner distribution |
| $P_{CW}(t,\omega)$ | The Choi-Williams distribution |
| $D_f r, \phi$ | The Radon Wigner Transform of a signal $f(t)$ |
| $FrFT^\alpha f(x)$ | The Fractional Fourier Transform of $f(x)$ of order $\alpha$ |
| $STFrFT^\alpha f(x)$ | The Short Time FrFT of $f(x)$ of order $\alpha$ |

# Chapter 1
# Introduction

This thesis is concerned with sonar signal processing, in particular in the statistical characterisation of sonar data. More specifically it deals with the signal processing of time series which were collected during a number of experiments in the Baltic Sea.

Conventional sonar signal processing techniques are mainly limited to a variety of beam forming and detection methods. Sound propagation in shallow water is a difficult task due to the stochastic behaviour of the medium both temporally and spatially. This results in the statistics of the data being non-stationary, and the performance of traditional methods deteriorating. In the thesis a brief introduction is given to wave propagation in shallow waters, and the sonar equation is explained for both active and passive cases. The main goal of this thesis is to increase the understanding of underlying characteristics of the sonar data and thereby improving the possibilities of understanding the limitations of the achievable signal processing gain. And thereby also improving the reliability of estimates of such as, false alarm rates, the figure of merit of sonar systems, detection probabilities and so on. This is achieved by the following steps:

- Statistical analysis of sonar data to point out some weaknesses in the traditional assumptions of the data characteristics, such as the extent of stationarity.
- A comparison of time/frequency analysis methods is presented and the difference between the resulting images is examined visually as well as quantified using a number of suggested measures.
- Two methods of conditioning sonar data and analysis results, are applied and evaluated. The methods are the stepwise outlier rejection, and Empirical Mode Decomposition. Both are used in a novel fashion to improve on the image quality of lofargrams.

## 1.1 Motivation

The underwater environment presents a great challenge for any signal processing scheme. It imposes restrictions on the processing that are often more difficult than anything in electromagnetic signals in the air such as communications or radar. The sound propagation properties of water varies very much for different areas in the Baltic Sea. Even in an area no larger than the archipelago of Stockholm there are numerous different conditions to consider, such as different water depths, different Sea-floor characteristics, and the added effect of the weather conditions. All of which contributes to making it more difficult to predict the behaviour of the sound propagation. The aim for this thesis is to characterise sonar data for a few characteristic areas of the Stockholm archipelago with very shallow water.
The return signals in sonar systems can be viewed as a mixture of deterministic and stochastic parameters. Consequently, instead of trying to model the wave propagating environment itself this thesis will try to point out a few characteristic properties of the signals and propose suitable methods for solving the problems.

## 1.2 Thesis Organisation

The thesis is organised as follows:
Chapter 2 provides a basic background for sonar signal processing and gives an insight to the difficulties that can occur. The topic of underwater acoustics is discussed, including wave propagation, especially in shallow water, and the influence of environmental parameters on the wave propagation. Following this underwater acoustics overview, an explanation of the sonar equations for both the passive and active cases is given. Then the different experiments, conducted at the Swedish Defence Research Agency, that are analysed in this thesis are presented.
The properties of sonar data are investigated in Chapter 3. With the focus on the statistical properties of the sonar data. The issues of different degrees of stationarity, linearity, and symmetry are discussed for different types of sonar data. For active sonar data the stationarity of consecutive pings is examined for both the transmitted signal and the following reverberation tail (the returning echoes). For passive sonar data the stationarity time is addressed, for both ambient noise data and data where a target appears. All of the tests are applied to data from the Baltic Sea, so the data is considered to be from shallow waters at all times, the depth in the Baltic Sea being on average about 50 m and no more then a few hundred meters at most.

Chapter 4 is dedicated to a number of time/frequency analysis methods. Time frequency distributions have been used in sonar signal processing previously, however it is rarely implemented in other forms than the lofargram (Short Time Fourier Transform or spectrogram) so there is plenty of room for improvement, especially in the cases where one is interested in detecting impulsive sounds. Another issue addressed in chapter 4 is the quality of time/frequency displays. The question is how well a transforms performs compared to other transforms. An attempt is made to quantify the time/frequency image quality by using a number of different measures.
The characterisation of sonar data provides a direction for the signal processing schemes to consider. Due to the nature of the wave propagating medium for the sonar data it becomes clear that methods for non-stationary and non-linear signal processing have to be considered.

In cases where one finds that the data consists of different parts, i.e stationary (typically tones from machinery) and non-stationary parts (Noise in various form and shapes), there is an advantage to be gained if the non-stationary parts can be filtered out. Also in many applications the amount of data might not be adequate and therefore the possibilities of resampling are considered.

To address these issues two different approaches to data conditioning are presented in Chapter 5. Starting with a method of outlier rejection applied to lofargram images. Then followed by a recent method called EMD (Empirical Mode Decomposition). These methods are evaluated by applying them to one of the data sets that is also analysed in chapter 3.
The results and findings of this work are discussed in Chapter 6. This is then followed by a section on future research.

# Chapter 2
# Underwater Acoustics and sonar data from the Baltic Sea

To provide an insight to the nature of the time series that are produced using sonar systems a brief survey of sonar and underwater acoustics is provided. The intention of this survey is to give an overall picture for the different aspects of sonar and underwater acoustics [1], [2], and [3]. First the sonar equations are introduced as they provide a good overview of the interesting parameters in underwater acoustics. The variables in the sonar equations are then explained in more detail. The sonar equations are followed by an introduction to the physical events that drive wave propagation underwater.

## 2.1 Sonar

Sonar is an abbreviation for SOund Navigation and Ranging. The most common of any sonar system is the widely used echo sounder. Virtually every sea going vessel, starting from rather small boats to large ships, is equipped with at least one echo sounder. There are numerous uses for sonars, one is the aforementioned echo sounder used for bathymetry. Other widely spread uses are fishing sonars, sonars for seafloor mapping, for finding shipwrecks, for oil exploration, and the naval applications where sonars are mounted on surface ships or on submarines, or as towed arrays with apertures of hundreds of meters. Sonars are also found on torpedoes used in combination with electronics and guiding algorithms as homing devices. Other uses of sonars are sonobouys, deployed from surface ships or from the air (planes and helicopters). It is also fairly common to have helicopters carry dipping sonars in cases when one wants to change the location of the sonar system rapidly.

## 2.2 The sonar equations

The sonar equations describe the key parameters and can be used to assess sonar performance. One of the key application is to find the range of detection both for those systems trying to

detect something and the those that try to avoid being detected. The sonar equations can be presented in many different forms, for different uses. The form below is chosen due to the fact that it displays most of the included variables and allows a straight forward interpretation of the equations. That is the sonar has detected something if the Signal Excess ($SE$) is larger than zero, $SE < 0$.

For passive sonar the sonar equation for $SE$ is defined below as

$$SE = SL - TL + DI - NL - DT \tag{2.1}$$

where $SL$ is the source level of the sound emitting object, see Chapter 10 in [4], and Chapter 6 in [5]. $TL$ is the one way transmission loss, i.e.. the attenuation of the transmitted sound over the distance in question. $DI$ is the directivity index of the receiver array. $NL$ is the ambient noise level, and $DT$ is the detection threshold. All values for the sonar equation variables are given in dB.

For active sonar, the sonar equation has two main cases, the reverberation limited case and the noise limited case. Reverberation is a form of noise just as ambient noise is, but instead of being produced from other objects than the sonar itself, it is what is called signal induced noise. That means, that all undesired received echos are reverberation, see section 2.2.4. It can be compared to the interference phenomena in communications caused by multipath propagation. Reverberation can thought of as active noise. What distinguishes the two cases from each other is the use of the terms $RL$ (reverberation level) and $NL$. $RL$ and $NL$ are compared at the receiver and the one that has the larger value decides if it is the reverberation or noise limited case. Limited in this context means a limit in range, i.e. the distance over which a sonar is able to detect objects. The reverberation limited case is

$$SE = SL - 2TL + DI - RL + TS - DT \tag{2.2}$$

where $SL$ is the source level of the transducer in question. $DI$ is the the directivity index for both the receiver and transmitter arrays. $RL$ is the reverberation level. $TS$ is the target strength of the object the sonar is trying to detect. $TS$ is a measure of how well an object reflects acoustic waves. $DT$ is the detection threshold, and it varies with the type of transmitted signal and the signal processing the returned signal is subject to.

For the noise limited active case the equation looks almost the same with the exception of replacing $RL$ with $NL$.

$$SE = SL - 2TL + DI - NL + TS - DT \tag{2.3}$$

where all included variables are the same as above. All variables included in the equations above are described in more detail in the next sections.

### 2.2.1 Array Gain and Directivity Index

Under the assumption that the incoming sound wave is a unidirectional plane wave (coherent) and the ambient noise is incoherent then the AG (Array Gain) reduces to DI (Directivity Index). AG is the quotient between an omni directional sound field and the integral of the combined beam patterns of both receiver and transmitter, and is defined below as

$$AG = DI = 10log_{10}(\frac{\int_{4\pi} d\Omega}{\int_{4\pi} b_r(\theta,\phi)b_t(\theta,\phi)d\Omega}) \tag{2.4}$$

where $b_r(\theta,\phi)$, and $b_t(\theta,\phi)$, are the beam patterns for the receiver and transmitter, respectively. If the patterns have rotational symmetry and are non-directional in the $\phi$ plane then the equation above can be simplified and rewritten to the following

$$DI = 10log_{10}(\frac{2\pi}{\int_{-\pi}^{\pi} b_r(\theta)b_t(\theta)d\theta}) \tag{2.5}$$

as would be the case of a vertically non-directional (line array) active sonar with transmitter and receiver beams $b_r(\theta)$, and $b_t(\theta)$. In the passive sonar case the sound source (transmitter, usually a vessel of some sort) is thought of as omni-directional and the above is simplified to

$$DI = 10log_{10}(\frac{2\pi}{\int_{-\pi}^{\pi} b_r(\theta)d\theta}) \tag{2.6}$$

Now it is easy to see that if the receiver is omni directional then the DI is zero, since the integral $\int_{2\pi} d\theta$ is equal to $2\pi$.

The DI should be used with caution since the coherence of the signal is very likely to deteriorate over distance, and the noise might not be entirely isotropic. Then the AG will be over estimated using the above equations. But they do provide a fairly good approximation and certainly good

enough to go into the sonar equation. Tactically there are different implications in overestimating the array gain. The hunter is over confident in detecting a target whereas the hunted object is too pessimistic in the likelihood of being detected.

For the special case where the array is a linear one with $\lambda/2$ spacing of $n$ elements the AG is given as

$$AG = 10log_{10}(n) \tag{2.7}$$

This assumes a coherent plane wave in isotropic noise. But when the noise coherence increases the array gain becomes

$$AG = 10log_{10}(\frac{n}{1+(n-1)\rho}) \tag{2.8}$$

where $\rho$ is the noise coherence, and the array gain is clearly less than in Equation. 2.7 as the maximum in Equation. 2.8 for a fixed $n$ is achieved for $\rho$ equal to zero and the equations are equal to each other.

The array gain can be viewed as the best achievable performance for any beam forming algorithm, since the gain is described as a quotient between a single element (omni-directional) and the beamformed output from an array. Array processing and beamforming is a large research field in signal processing and there are numerous books, see [6], [7], [8], and articles written on the subject, see [9], [10],[11], [12]. However that falls outside the scope of this thesis since it will be assumed that the appropriate beamforming is already done to produce the desired time series or that the experimental setup does not allow any beamforming to be done (single hydrophones). Further reading on beamforming in a sonar context and other fields can be found in [13], [14], and [15] just to mention a few.

### 2.2.2 Detection Threshold

The detection threshold is probably the most interesting variable in the sonar equation, from a signal processing point of view. This is where the difference between signal processing schemes is demonstrated. When assessing sonar performance it is crucial to chose a DT that is appropriate for each case. Choosing a DT to low means the system is likely to have many false alarms and doing the opposite means that signals that should have been detected go by unnoticed. Generally speaking the detection threshold is defined as,

$$DT = 10log_{10}(\frac{S}{N}) \tag{2.9}$$

i.e. the signal to noise ratio, where $S$ is the signal level and $N$ is the noise level. If the detection threshold is the sought after variable, it can also be found from the sonar equation setting $SE$ = 0, and rearranging the variables to the following, $DT = SL - TL + DI - NL$, this being the passive case. Similarly, it is possible to derive an expression for the active case. When other terms in the sonar equation are the target then the DT must be estimated. Fortunately there are simple expressions for the detection threshold for the different cases of the sonar equation. The detection threshold is set differently with the level of knowledge about the received signal. It is common knowledge that the best detection performance is achieved using a matched filter. This implies complete knowledge of the received signal. Active sonar with a known and well defined pulse shape is a good example. In that case the detection threshold can be written as

$$DT = 10log_{10}(\frac{d}{2t}) \tag{2.10}$$

where $t$ is the pulse length (or observation time in the passive case), and $d$ is the detection index. This expression is valid for noise limited active sonar and for passive sonar when the signal shape is known. It can be seen that DT is decreased with longer pulses in the active case and by integration (observation) time in the passive case. The limiting factor for the pulse length in the active sonar case is reverberation, and it goes up with longer pulses. So at some point the reverberation becomes the limiting factor. However, with larger bandwidths the reverberation goes down, so there is a balance between RL and pulse length and bandwidth.

On the other side of the scale is the case without any prior knowledge of the received signal, for instance in the passive narrow-band case. However, there is an assumption that the signal consists of tonals such as the sound emitted from the drive or transmission of a ship. Then the detection threshold is set to

$$DT = 5log_{10}(\frac{dw}{t}) \tag{2.11}$$

where $d$ is the detection index, $w$ the receiver bandwidth, and $t$ the integration time.
Detection is another large field in signal processing, that falls slightly outside the scope of this thesis but it is an important part of sonar signal processing so it will be mentioned at times, and parallels are drawn to the most commonly used techniques for detection, see [16], [17], [18], [19], [20], and [21] just to mention a few.

### 2.2.3 Noise Level

NL (Noise Level) is a collective term for all noise sources in the sonar equation reverberation aside, so NL includes self noise and ambient noise, see [22], [23], and [24]. Ambient noise comes from different sources, shipping, from the sea it self i.e. waves, rain fall, and from marine life. Self noise is everything where the source of the noise can be related to the platform carrying the sonar in question. Starting from within there is electrical (thermal) noise, there is the platform itself (ship, submarine etc.), typically in the form of propulsion or auxiliary machinery. So depending on the frame of reference, part of one systems self noise is actually the signature that others want to detect. If the platform is mobile there is also flow noise in form of turbulent flow over the sonar. The ambient noise is often described with Knudssen spectra [25], [26], or Wenz curves, [27].
Most signal processing that is done to sonar data assumes that the ambient noise is stationary, it is also often assumed to be isotropic, and finally that it is Gaussian. In chapter 3 it will be shown that this is not necessarily true.

### 2.2.4 Reverberation Level

Reverberation is signal induced noise, see [28], and [29]. There are three different forms of reverberation, surface, volume, and bottom reverberation. In short all undesired echoes are reverberation. Surface reverberation is the back scattered wave from the sea surface, and it is dependent on the roughness of the sea. A rough sea will lead to higher levels of surface reverberation than a calm sea surface. This is due to the rough surface acting as several scatterers, scattering the signal rather than reflecting it. A calm, smooth sea surface will reflect rather than scatter the wave and provide multiple paths for the wave to travel, thus producing several echoes from the same objects. The total RL (Reverberation Level) can be larger or smaller depending on the rest of the contributing factors. A calm sea surface does not necessarily lower the total amount of reverberation. Volume reverberation is back scattered waves from the water volume and objects in the water volume. Bottom reverberation is the back scattered waves from the seafloor. Of all forms of reverberation the bottom reverberation is the most important and generally the largest contributing source of reverberation. It depends strongly on the acoustic properties of the seafloor. Acoustically soft materials, that is materials with a low acoustic impedance, give lesser echoes than materials with a high acoustic impedance, or acoustically hard materials. Also the scale of the bottom is important, if the bottom consists of large rocks

instead of non-consolidated sand the reverberation levels are inevitably higher. In other words, reverberation depends on how much of the sound waves are scattered back to the sonar. This is called the scattering strength of a bottom type, or volume or surface respectively. The scattering strength is the ratio of back scattered sound to incoming sound as in $S_{s,v} = 10log_{10}\frac{I_{scat}}{I_{inc}}$, where $S_{s,v}$ is the scattering strength for both surface (surface and seafloor) and volume. $I_{scat}$ and $I_{inc}$ are the scattered and incoming intensities respectively.

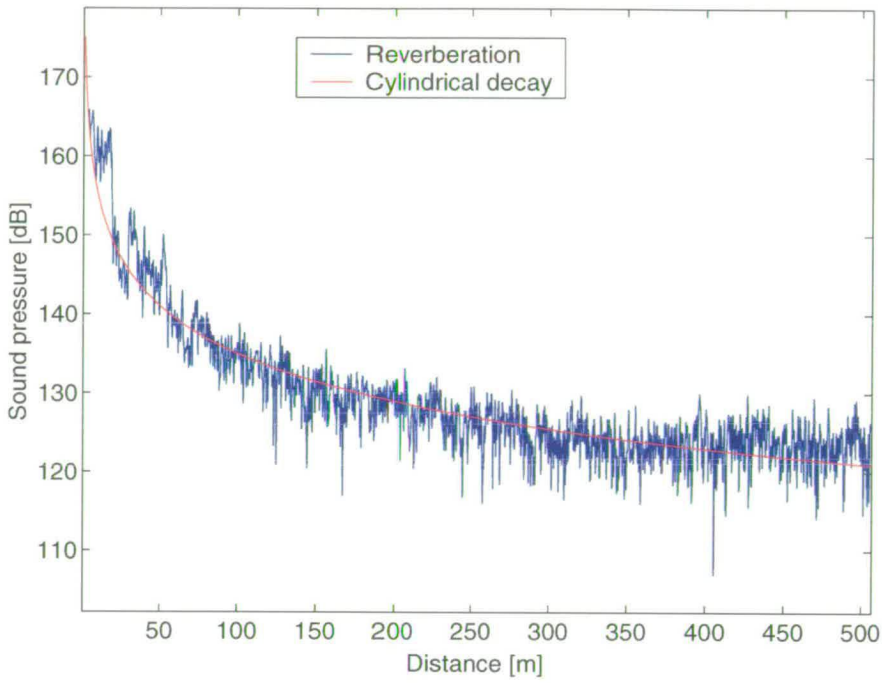


**Figure 2.1:** *An example of shallow water reverberation from a trial site outside of Stockholm.*

The RL term depends on the scattering strength and the surface or volume that is insonified, clearly the beam patterns of both the receiver and transmitter will play an important role in the amount of reverberation received. In the sonar equation the RL term is given as a function of RL versus time/distance, see Figure 2.1 for an example of shallow water reverberation from a trial site outside of Stockholm. The trial site is described in Sections 2.4.2 and 2.5. In the sonar equation the RL is then selected for the distance in question.

The statistics of the reverberation are largely unknown. This is due to the fact that each and every area will have different characteristics. The one thing that can be said is that, unlike ambient noise that in some cases can be thought of as isotropic, stationary and Gaussian, reverberation is not any of these. Due to the great variation of almost all parameters in the sea and

the fact that reverberation is a local phenomena driven by the signal transmitted the nature of reverberation changes over rather short periods of time. An attempt to shed some light on the statistics of reverberation is presented in chapter 3.

### 2.2.5 Source Level and signatures

The SL term is used in both the passive and active sonar equation. In the passive case it presents the sound unintentionally emitted by a vessel. The sound sources are typically the propeller, propeller shaft, drive and engine. The sound that is emitted from a vessel is called its signature. The signature is typically described by a number of tonals at certain frequencies, and a broad band frequency spectrum. The tonals from vessels are usually in the low end of the frequency spectrum, typically up to frequencies of about 1 kHz. The broad band spectrum description of a signature is often presented as a sparsely sampled power spectrum, in third note, octave, or decade frequency bands, see Figure 2.2 below. For a more thorough description of the sparsely sampled broadband signature see appendix A, Tab. A.1 on page 152.

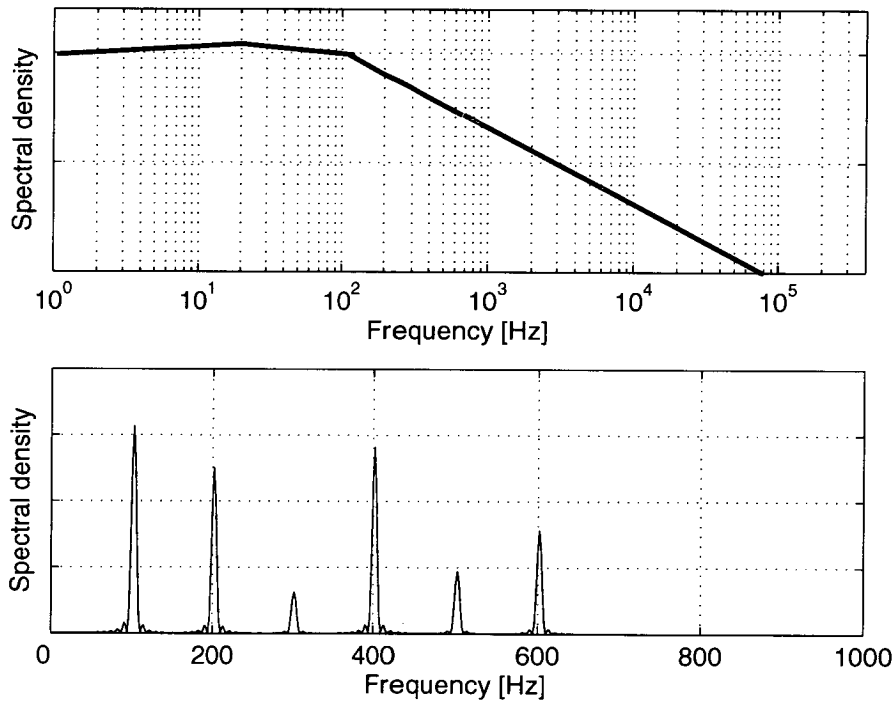


**Figure 2.2:** *A synthetic example of a broadband and a narrow band signature of a vessel.*

In the active case the source level is the sound pressure produced by a transducer (transmitter). The source level is mostly given as a sound pressure level compared to 1 $\mu$ Pascal at a distance

of one meter. For instance 210 dB (rel 1 $\mu Pa$ at 1m)

### 2.2.6 Transmission Loss

Transmission loss is the amount the emitted signal is attenuated from its source to its destination. The attenuation (loss of power) of a wave travelling through water, especially shallow water, has many factors affecting it. The water sound speed profile is one of the main factors as it bends the waves up or down. Then the bottom and surface conditions are also a big influence. Depending on the roughness of the sea a different amount of the waves are reflected and other parts just scattered. The same applies to the seafloor, with the added influence of the material in the bottom. All these wave attenuating and guiding factors are described in more detail in section 2.4.

## 2.3 Transducers and Receivers

Transducers and receivers are underwater loudspeakers and microphones and are also called hydrophones. They share similarities to the speakers and microphones used for aero acoustics but have one major difference. The media they operate in, and how that influences their design. In aero acoustics only a percentage as small as three percent of the power input is actually transfered in to the air as sound pressure. In water the figure is as high as 95 percent. As a consequence it might be considered that making a waterproof sub-woofer would be a good idea of how to produce underwater sound, Unfortunately, it is not as simple as that. Speakers in aero acoustics have a very long stroke (in comparison with their underwater counter part). Bearing in mind that water is about 1000 times more dense than air it becomes clear that this could not possibly work. Underwater transducers work with a much shorter stroke, and smaller membranes. A typical way of producing a speaker for aero acoustics is simply to use a coil to drive a piston inside the coil which in turn is attached to a speaker membrane. The way to build a underwater transducer is entirely different, but the designs are not worlds apart. The underwater speaker works in air, but not very efficiently, and since it is not designed for it, it doesn't last very long when pushed to its limits. The classical approach is building what is called a Tonpilz element, illustrated in Figure 2.3.

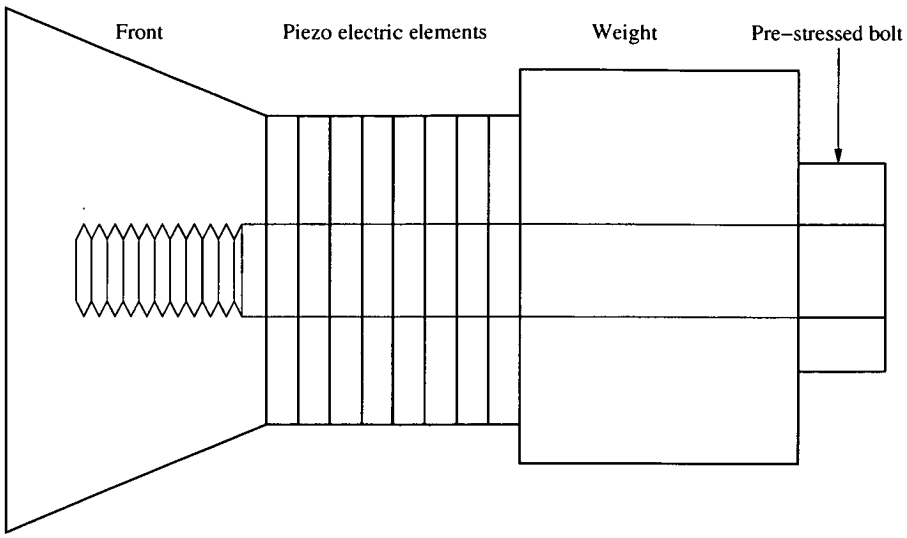


**Figure 2.3:** *Tonpilz element.*

The Tonpilz element is made of a stack of piezoelectric elements joined together with a bolt that goes through the stack. Piezoelectric materials are materials that can expand and shrink when a voltage is applied to the adjoint surfaces of the materials, see [30]. The bolt is fastened to a backing weight and to the front part of the element. The front of the element is attached to a membrane and that membrane is the part that is in contact with the water, also known as an acoustic window. The backing weight, the piston head, and ceramic mass together with the water load gives the resonant frequency of the transducer. The Tonpilz transducer is only one of many different ways of constructing a transducer, another is the flextentional transducer, see [2] p. 84-86. There are transducers, that use elecrostrictive materials. That is using the effect of applying an electric field across a dielectric material and that deforms it and this motion is used to produce sound waves. Magnetostrictive transducers use magnet deforming materials, that are usually attached to a diaphragm and that motion produces the sound waves, see [31], [32], and [33] for more detail on different kinds of materials and transducers. Recently there has been interest in fibre-optic transducers. The fibre-optics based transducers have the potential benefit of a very high sensitivity, see[34].

## 2.4 Oceanography

The Baltic Sea is in comparison with other seas a shallow one, its depth ranges from about 400 meters at the Landsort trench to the very shallow, fairly vast, coastal areas. The archipelago of Stockholm is of special interest of several reasons, it is a shallow water environment with fairly heavy shipping and highly varying bottom properties. It is also one of the waters where the Swedish navy operates. But first and fore most, the fact that wave propagation in shallow water is very hard to model and to predict makes for very interesting time series to analyse and to apply different signal processing schemes. The main influences on wave propagation are the water sound speed profile, sea floor properties, that is both the surface of the seafloor and the volume consisting of different materials, and the roughness of the sea or the sea state, i.e the surface. So how does all this affect the variables in the sonar equation? Transmission loss for one thing goes up if the two boundaries, bottom and surface, absorb and scatter the incoming waves rather than reflecting them. On the other hand if the surfaces do not reflect the propagated wave very well, the reverberation levels go down. With different sound speed profiles the wave propagation differs very much from time to time. The difference in TL can be in magnitudes of tens given the best and worst conditions.

### 2.4.1 Sound speed profiles and Wave propagation

We have already stated that the sound speed profile (SSP) affects the wave propagation in water, now we need to explain why and how, see [35] and [36]. First of all we look at the sound speed in water and what its dependencies are. The key parameters are the temperature of the water, the pressure (or depth), and salinity. The sound speed increases with increasing temperature, and increases with increasing pressure, it also increases with increasing salinity. The sound speed of the materials in the seafloor depend on the composition of the seafloor. Different materials have very different characteristics as far as their sound speed. The seafloor in the Stockholm archipelago varies quite a lot as can be seen in section 2.4.2. There are cases when the sound can travel faster in the sediments than it does in the water. Also at times there are bottom materials that have lower sound speeds than the water, which will give later arrivals of the emitted sound waves. And cases where the sound waves are trapped in the seafloor to emerge later at a greater distance. The sound speed in water is measured when ever ships are at sea to get an idea of the wave propagation properties. Whereas the seafloor properties are assumed to be more or less the same over greater areas and it would naturally be very difficult

to probe the sea floor at each instance the water sound speed is measured. On top of that it is not as great an influence as the sound speed so the approximations made are in most cases good enough. A typical example of how the water sound speed profile can change over just one day in the summer can be seen in Figure 2.4, and 2.5, this is not a very extreme example of changes in the SSP. The temperatures vary between roughly 8 and 13 degrees Celsius. During hot summer days the surface temperature can go up toward 24-25 degrees Celsius. That would make the profile even more rapidly changing, as in Figure 2.6.

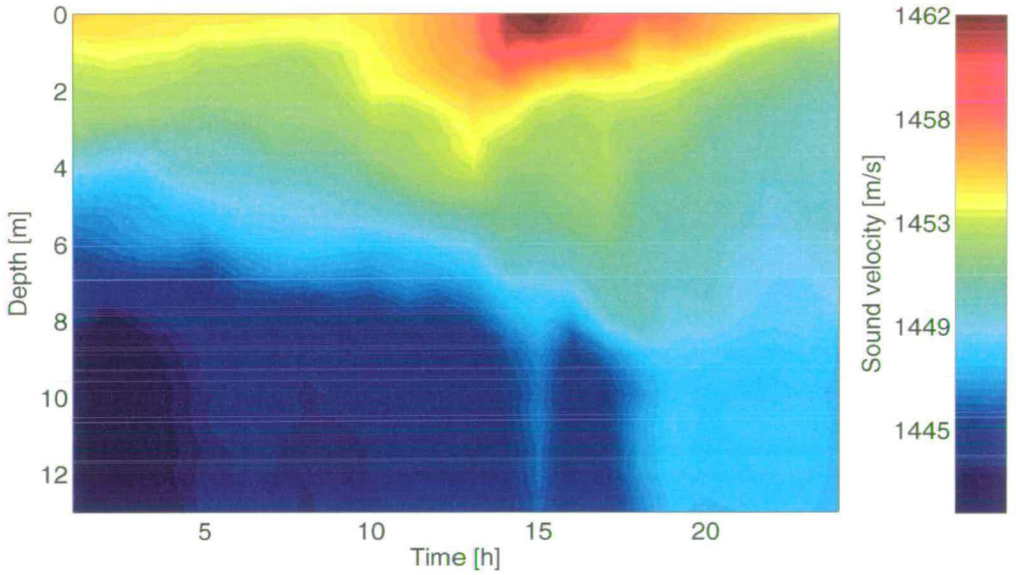


**Figure 2.4:** *Sound speed profiles from FOI's trial site in the archipelago of Stockholm during the first of June, 2004.*

Now that the variations in sound speed have been established, the influence of the SSP on the wave propagation can be addressed. Sea water is typically layered with cold water (4 degrees Celsius) on the bottom and warmer water on the surface. When ever there are rapid changes in the SSP this creates the effect of a wave passing between two different media, since water of different temperature (and therefore sound speed) have different density and impedance, see figures 2.7, and 2.8.

Waves passing from lower sound speed water to higher velocity water are refracted downward, and the opposite is true. So on a day when the SSP is upward refracting the propagated waves travel in rays that reflect from the surface and travel further away to yet again reflect from the surface. Or if the SSP is downward refracting the sound waves are directed to the sea floor and if the sea floor is acoustically soft the sound propagates down the sediments rather than reflecting

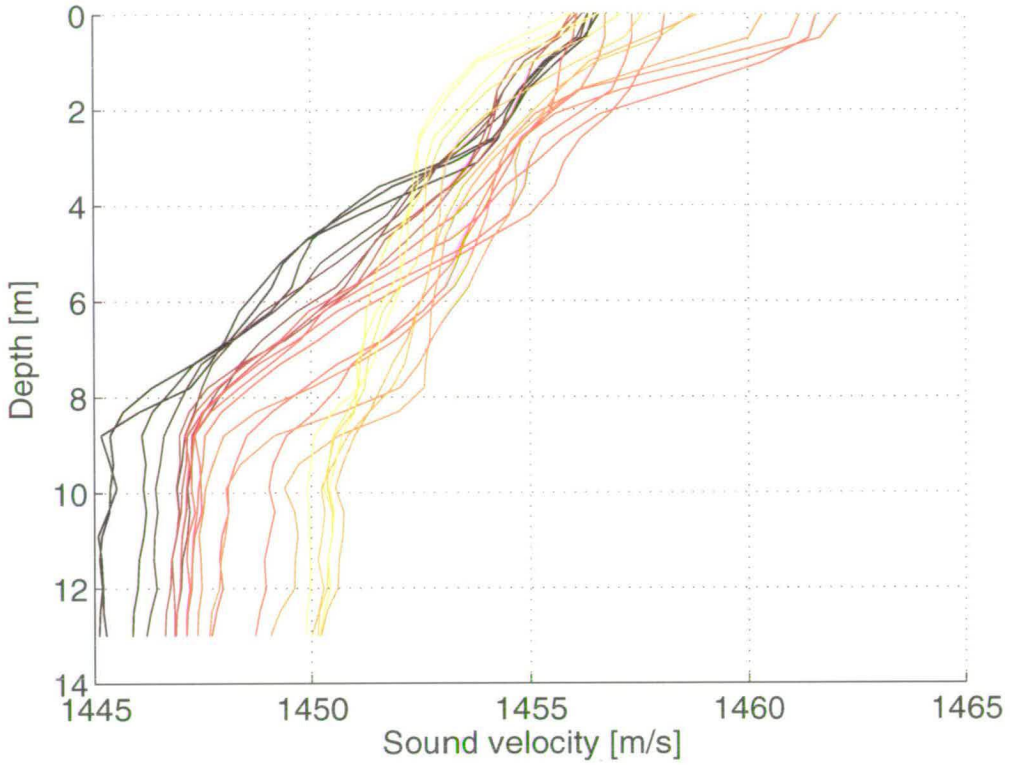


**Figure 2.5:** *Individual sound speed profiles, from a shallow bay south east of Stockholm, 1 June 2004. The lightest coloured one is from 1am, and the darkest one is from 12pm*

up in the water volume again. With an acoustically hard bottom the sound is reflected up again as can be seen in the middle display of Figure 2.8. Although the nature of wave propagation is to say the very least difficult it does create interesting effects. For instance, if a SSP has a shape that concentrates the waves, as in the bottom display in Figure 2.8, this creates a wave guide that can carry the sound waves hundreds of kilometres, this is known as a sound channel.

The panels in Figure 2.8 show something called ray diagrams, that display how wave fronts propagate under certain conditions. Ray diagrams follow from solving the wave equation using ray theory which is one of two common approaches of solving it, the other one being normal-mode theory. The wave equation is defined as follows,

$$\frac{\delta^2 p}{\delta t^2} = c^2\left(\frac{\delta^2 p}{\delta x^2} + \frac{\delta^2 p}{\delta y^2} + \frac{\delta^2 p}{\delta z^2}\right) \tag{2.12}$$

where $p$ is the pressure related to the coordinates $x, y, z$ at time $t$, and $c$ is the water sound speed, $c$ may vary with the $x, y, z$-coordinates and time as well. Solving the wave equation using normal-mode theory gives the complete solution to the pressure field. In normal-mode

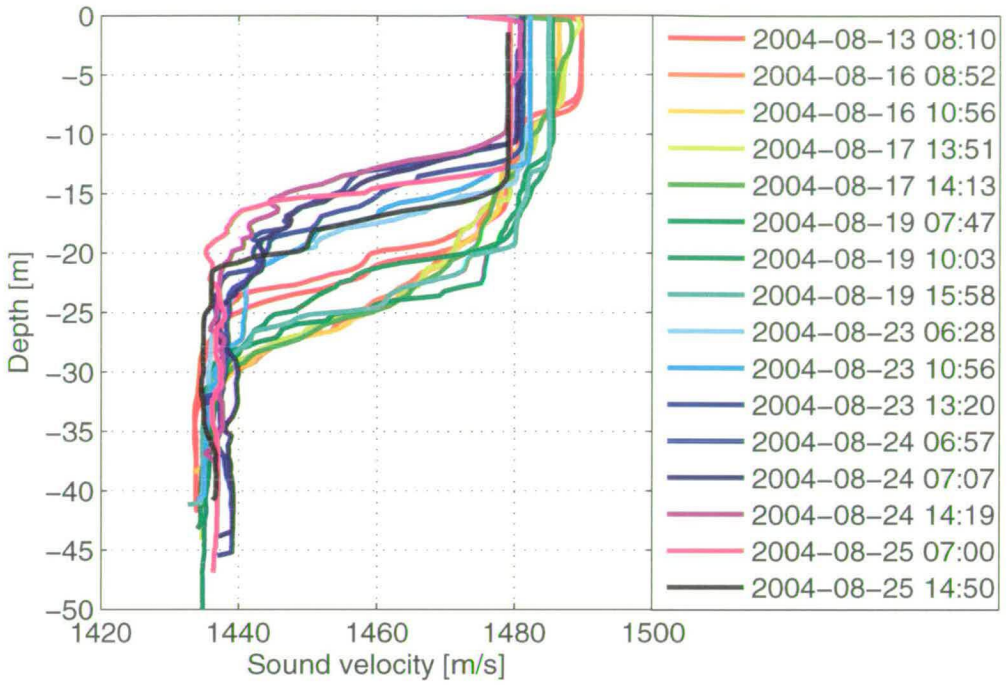


**Figure 2.6:** *Sound speed profiles from a sea trial conducted in August 2004. See section 2.5.1*

theory the propagation is described as a set of characteristic functions called normal modes, each of them being a solution to the wave equation. Then the normal modes are combined additively to produce the entire pressure field. The normal-mode solution of wave equation is valid for all frequencies but due to computational cost it is seldom used for anything but low frequencies, i.e few modes.

The other approach is, as previously stated, ray theory. The basis of ray theory is to postulate wavefronts, and the assumption that there are rays that these wavefronts follow. Along the wavefronts the pressure is considered to be constant. When approximating the SSP with a linear function as in Figure 2.8 a propagating wave can be shown to to follow a circle arc, see [1], [2], [3]. The circularity of the rays is used in computing the propagation of sound waves. The approach to ray tracing programs is to approximate the SSP with piecewise linear functions and to make use of Snell's law to find the direction that the rays will take, i.e how they are refracted when going from one layer of the media to another. Snell's law is defined below

$$\frac{cos\theta 1}{c1} = \frac{cos\theta 2}{c2} = \frac{cos\theta 3}{c3} = ... = constant \tag{2.13}$$

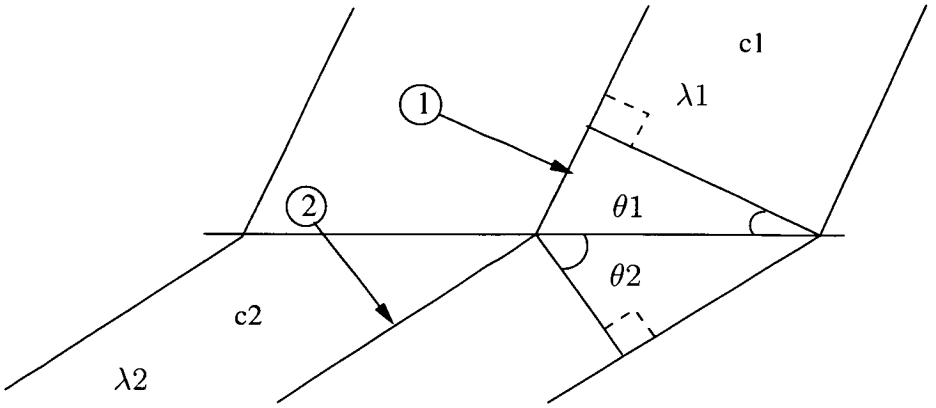


**Figure 2.7:** *Sound waves passing from a layer of water with sound speed c1 and wavelength λ1, to a layer of water with sound speed c2 and wavelength λ2, Where c1 > c2.*

Ray tracing is an appealing method for solving the wave equation as it also provides a graphical view of the propagated sound, but it comes with limitations. It is a high frequency approach to solve the wave propagation problem, it is only valid if the radius of the ray curve is larger than the wavelength, and that the rate if change in sound speed (with depth) is small, within a wave lengths depth change. Most wave propagation programs that solve the wave equation for the entire pressure field, use both ray and wave theory to solve the wave equation to keep the computational cost down and making efficient use of both methods within their valid and efficient regions.

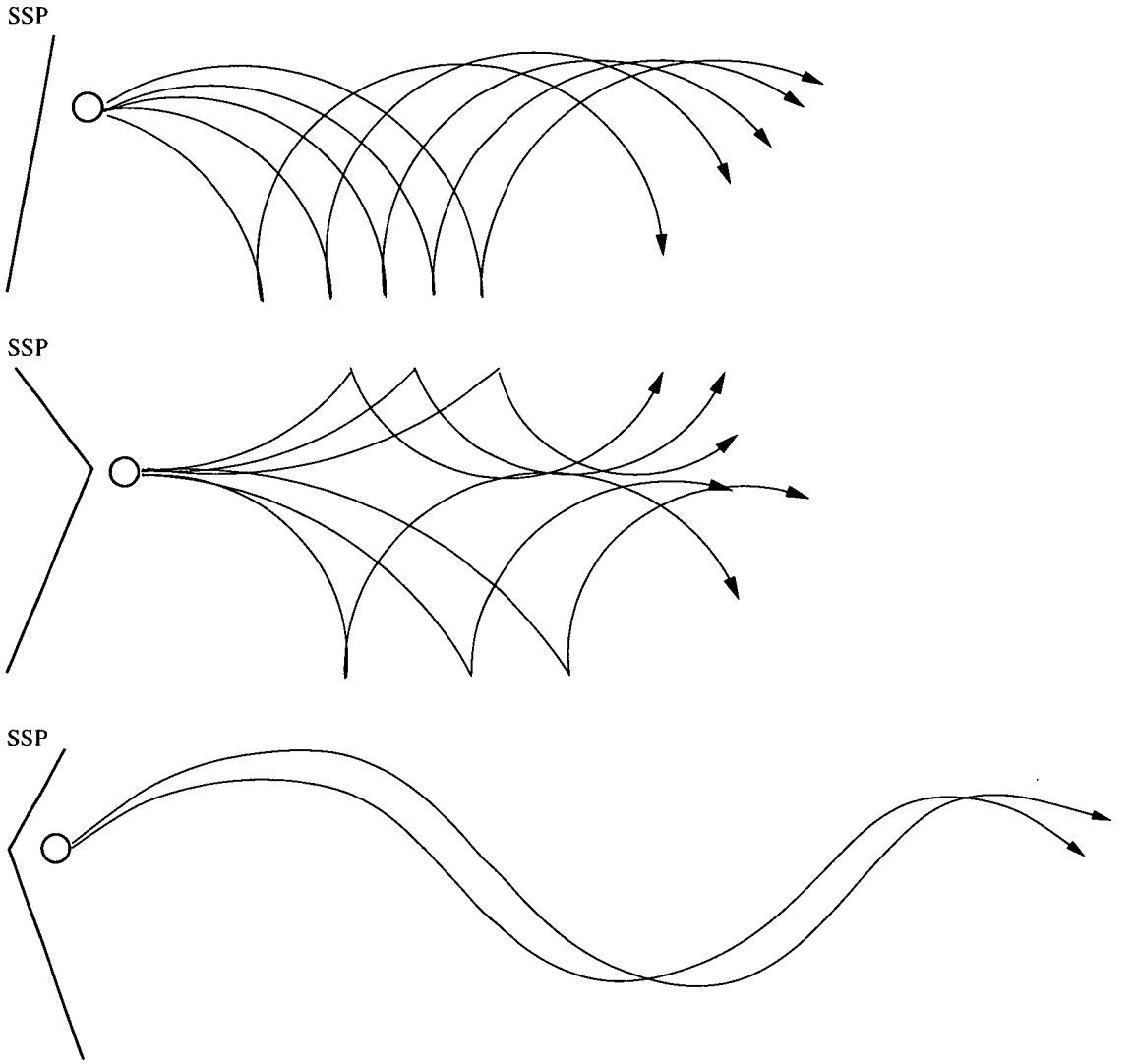


**Figure 2.8:** *Three examples of how a wave front can travel in water, with corresponding sound speed profiles to the left. The top panel shows a shallow water down refracting profile, the middle one a combination of both a down and upward refracting profile for shallow water, where the upper layer of the water has a positive gradient and refracts the waves upward and the bottom layer refracts the waves down to the sea floor. The bottom panel shows a SSP that creates a sound channel where the waves propagates in a wave guide made up of the two layers of water.*

### 2.4.2 The geology of the Stockholm Archipelago

The bedrock of the Stockholm Archipelago in the central Baltic Sea consists mainly of crystalline granites, gneisses and leptites formed more than 1800 Million years ago. This area is intersected by several tectonic lineaments and small fracture zones which contribute to the large extension of the archipelago. These structures were at least partly initiated in connection with the bedrock formation and have probably been reactivated several times since then. Even if this area is not tectonically active at present day some of these geological structures show evidence of energy release by small movements and by outflow of thermogenically altered gas [37], [38], and [39].

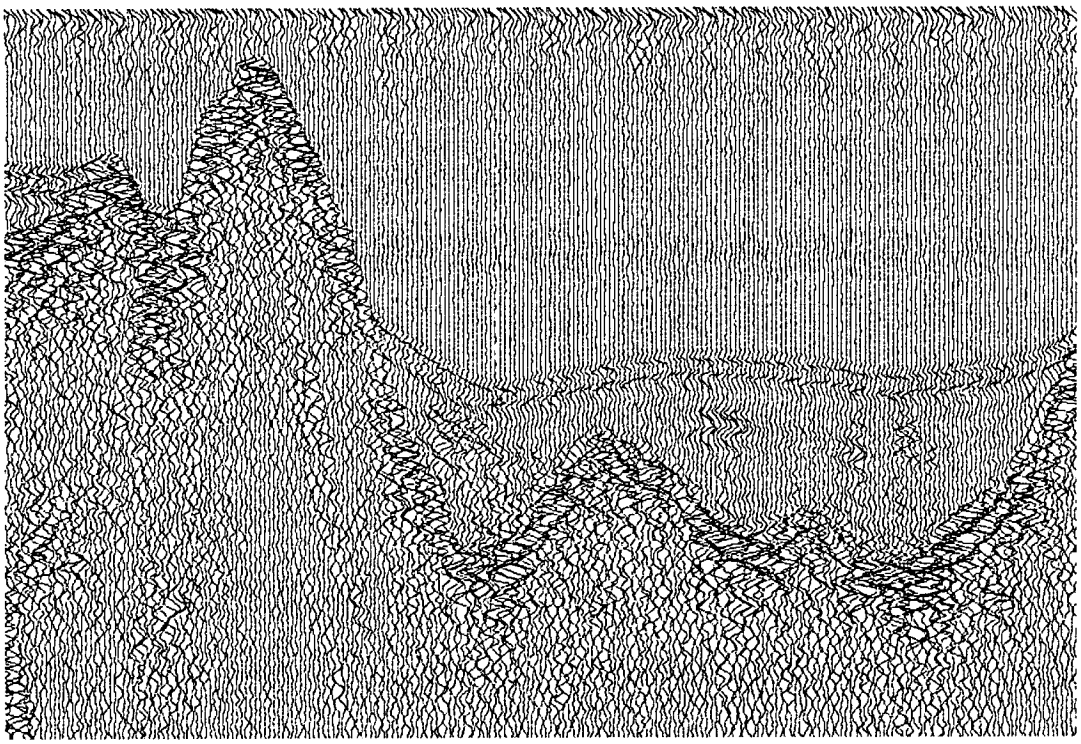

**Figure 2.9:** *A bottom profile from the archipelago of Stockholm. The maximum water depth is approximately 40m. The rocky outcrop to the left is at a depth of about 15m*

Mostly the bedrock is covered by quaternary sediments which may be generalised as follows; Covering the bedrock is a thin layer of till, typically some metres thick. On top of the till are glacial and post glacial clays deposited. Typically, the glacial clay is quite uniform in thickness being frequently a few metres thick. The glacial deposits are succeeded by post glacial clays and recent mud. The thickness of these sediments is highly variable due to local variations in

condition of accumulation and erosion. Locally the post glacial sediment amounts to several tens of metres in thickness. Due to the post-glacial process of uplift and erosion the upper sediments may locally be missing and pre-quaternary sediments exposed in the seafloor. See Figure 2.9

The image shown in Figure 2.9 is done stacking traces from a sea floor survey done by the Geological Survey of Sweden (SGU) for FOI during early spring 2000, [40], using an air gun (sleeve gun) as a sound source. Also see [41], and [42] for more information on seafloor characterisation. Compressed air is used to build up a high pressure and then releasing it through a nozzle underwater creating a bubble that collapses and thereby creating a wide band low frequency impulse. The research vessel Ocean Surveyor from SGU was used as the platform carrying a hydrophone array, and the sleeve gun. Ocean Surveyor travelled at 3m/s and fired one shot from the sleeve gun every second. The time series in the image are not processed in any way just placed adjacent to one and other. The figure shows the sea floor consisting of an upper softer layer with a faint echo and the much harder bedrock beneath. Also the typical rocky out crop can be seen to the left in the figure. The deepest part of the trial site as shown in the image is about 40m. The rocky outcrop is at a depth of about 15m.

## 2.5 Experiments and sonar data acquisition

At the Swedish Defence Research Agency, FOI, there is an ongoing experimental activity. The departments concerned with underwater acoustics and electromagnetics conduct a number of sea trials every year. These sea trials are usually designed to answer questions about the different parameters in the sonar equations. For instance the TL and RL was measured at a trial in 2002, described below. Aside from the sea trials FOI also has a tank for underwater acoustic and electromagnetic field measurements. The tank is primarily used to calibrate sensors and to test equipment prior to sea-trials. The experimental data that has been analysed in this thesis is described in the following sections and are summarised in Table 2.1.

### 2.5.1 The multi sensor trial

During August 2004 a trial was conducted in the Stockholm Archipelago with the objective of fusing data from different types of sensors. The equipment used during the trial was a 32 element acoustic array and a long base electromagnetic array including 8 pairs of elec-

| Sea Trial | Purpose | Data size | Frequency band |
|---|---|---|---|
| Multi Sensor Trial | Ambient Noise | 200GB | 0-12500Hz |
| BAROC Sea Trial | Reverbartion | 2GB | 0-30kHz |
| Utö Sea Trial | Reverberation | 2GB | 0-10kHz |
| The SAS Sea Trial | Synthetic Aperture | 600MB | 0-250kHz |
| The Ambient Noise experiment | Noise Level | 600MB | 0-22kHz |
| The Fibre Glass Boat Trial | Passive sonar | 1GB | 0-12500Hz |

**Table 2.1:** *The purpose, and amounts, of sea trial data in this thesis*

trodes. furthermore a tri-axial short base-line electrode system was used as a reference. The targets tracked during this trial was small surface ships and towed acoustic and electromagnetic sources. See Figure 2.10, and also see [43] for a more thorough explanation of the trial.

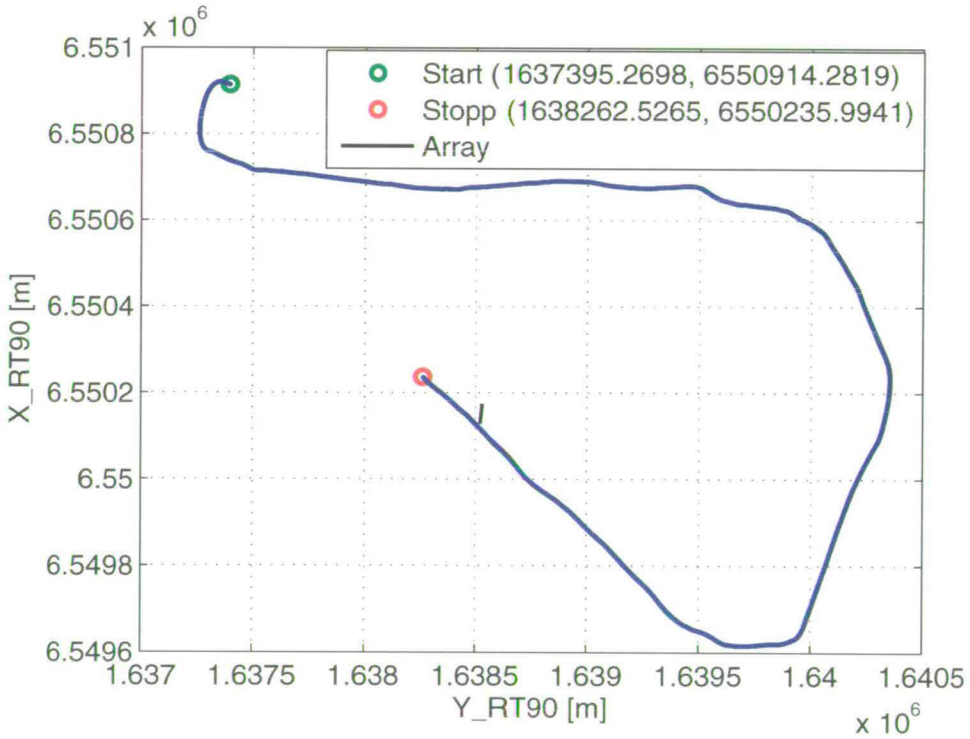


**Figure 2.10:** *The true position of HMS Ägir towing an acoustic source transmitting a sinusoid at 200Hz*

Also during this sea trial the ambient noise was measured during night times and at weekends, collecting 1 minute of data every fifteen minutes. This was done using the 32 channel array, and the data was sampled at 25 kHz.

| Frequencies and Bandwidths for the first pulse train | | | | |
|---|---|---|---|---|
| 350-450Hz | 600-700Hz | 700-800Hz | 800-900Hz | 900-1000Hz |
| 1000-1200Hz | 1200-1400Hz | 1400-1600Hz | 1600-1800Hz | 1800-2000Hz |
| 3000-3200Hz | 3200-3400Hz | 3400-3600Hz | 3600-3800Hz | 3800-4000Hz |
| 4000-4400Hz | 4400-4800Hz | 4800-5200Hz | 5200-5600Hz | 5600-6000Hz |
| Frequencies and Bandwidths for the second pulse train | | | | |
| CW 420Hz | 3000-5000Hz | 1000-2000Hz | CW 4000Hz | CW 1500Hz |

**Table 2.2:** *The two different pulse trains used during the BAROC sea trial.*

### 2.5.2 The BAROC sea trial

During September 2002 a sea trial was conducted in the vicinity of the Landsort trench in the northern part of the Baltic Sea. Trial was conducted as a joint project between FOI and FWG, and had three objectives. The measurements that were performed are Reverberation Level (RL) measurements, Transmission Loss (TL) measurements, and finally a number of tests done for the underwater communications group at FOI.

During the trial three vessels were used, WFS Planet (FWG), MzB Swedeneck (Naval Test Center WTD 71), and HMS URD (FOI).

The data of interest for this thesis is taken from the TL measurements. MzB Schwedeneck ran on given tracks to or from WFS Planet, towing three transmitters, see [44]. Every 80 seconds Schwedeneck transmitted a train of pulses consisting of 20 separated pulses (both time and frequency), where each pulse was a chirped signal starting and stopping at a given frequency as can be seen in Table 2.2. Each one of the pulses in this pulse train is one second long. These 20 chirp are followed by another set of pulses, three CW (continuous wave) pulses and another two chirps, the first CW pulse has a duration of one second. The chirps are two seconds long and the final two CW pulses are five seconds long. A schematic picture of the transmitted pulse train can be seen in Figures 2.11, and 2.12. On board WFS Planet two different receiving systems were used during the experiments. The one of interest in this case is the vertical array VAIII, nested array, designed for three frequencies, 1200, 2400, and 4800Hz. The acoustic section of the array is 40 metres long and holds 128 elements. The signals were sampled at 59900 Hz. Each file of data is 1.8GB in size, thus having roughly 4.5 hours of data in each file. In this thesis the results from only one of these data files are presented, since they are representative for the overall behaviour of the data from this particular trial.

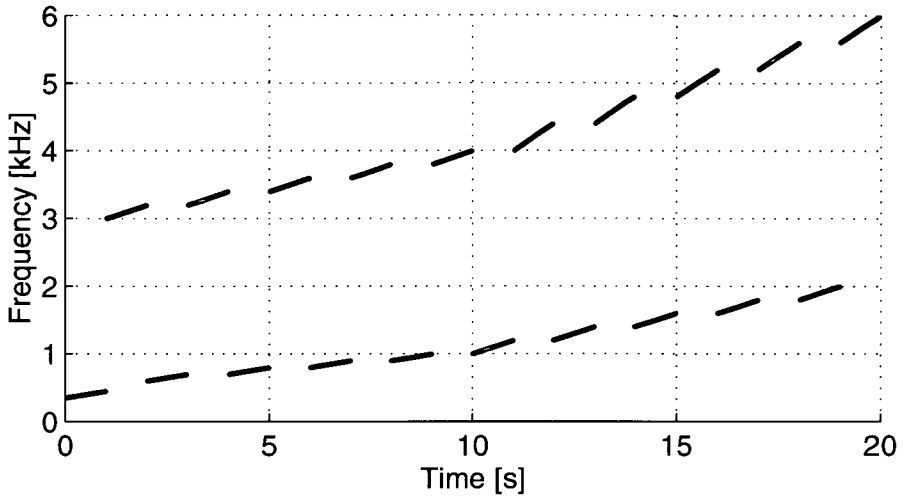


**Figure 2.11:** *The first part of the data set consists of 20 chirped pulses varying bandwidth with a fixed duration of one second. They were transmitted in 1 second intervals without time overlap.*

### 2.5.3 The Utö sea trial

The experiment was carried out in the archipelago of Stockholm during September 2001. The objective of this trial was to measure reverberation levels and transmission loss in a certain area. The equipment utilised during this trial was HMS URD, a ship of the Swedish navy for underwater acoustic applications, an array consisting of 32 elements, and an omni directional transmitter, see Figure 2.13. The data was sampled a rate of 20kHz. The results from the analysis of three data files from this trial are presented in this thesis. Each of the data files consists of 26 or 27 pings spaced 30 seconds apart. See [44] for a more thorough description of the trial.

### 2.5.4 The synthetic aperture sonar sea trial

In October 1999 another sea trial was conducted in a shallow bay the southern archipelago of Stockholm. The objective of the trial was to highlight possibilities of using synthetic aperture sonar (SAS). The pulse shape used during this trial was a 1ms, 90kHz centre frequency, 60kHz band width, linear chirp, see Figure 2.14. The trial equipment consisted of a remotely operated vehicle (ROV), on which a 32 element receiver array and a transmitter was mounted, see [45]. The data collected consists of recordings from the 32 element array sampled at 500kHz. Every record has 25 pings (pulses). In SAS both transmitter and receiver array are moved along a

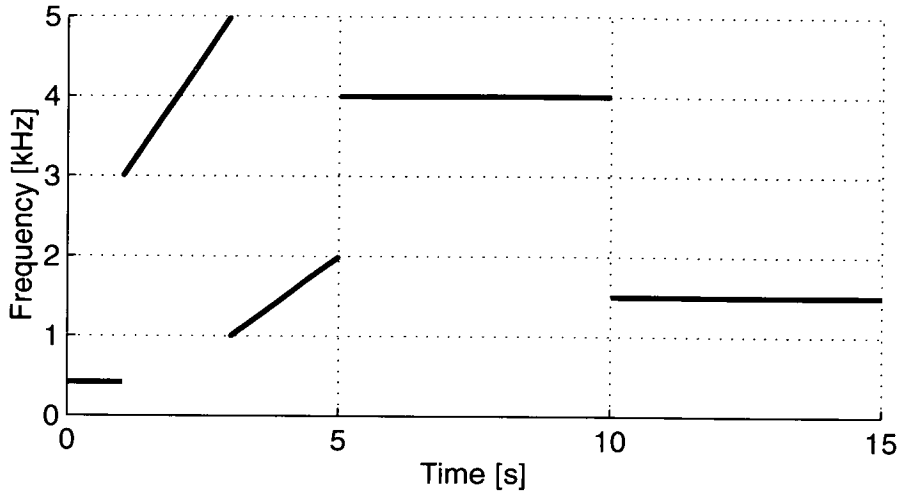


**Figure 2.12:** *The second part of the data set consists of 5 pulses. Starting with a CW pulse followed by two chirps, and another two CW pulses*

straight line to form an array of longer length than the physical length of the array, hence the name synthetic aperture sonar. The interesting part of this trial for the scope of this thesis is the pulse shape and the data it self, not the SAS application. The BAROC experiment has very low bandwidths (some are even CW pulses) and low chirp rates, and this set has, high chirp rates and bandwidths.

### 2.5.5 The ambient noise experiment

This data set consists of sonar, wind speed data, and sound speed recordings from, and nearby a shallow bay in the archipelago of Stockholm in the Baltic Sea. The sonar data was collected in 10 second sequences (sampled at 44kHz) on the hour, every hour during the period August 7th - 15th, 2001. Wind speed data is collected during the same period but in half an hour intervals, with maximum wind speeds for each measurement interval. The sound speed profiles were (are actually continuously being recorded) recorded with a bottom mounted array of thermistors, equidistantly spaced over the depth of 13m at the site, the recordings are done with 10 minute intervals around the clock.

### 2.5.6 The fibre glass boat trial

The sonar data was recorded in a shallow bay in the Baltic Sea with a water depth of about 30 m. The trials were performed using a 23-ft. fibreglass motorboat with a 4-cyl. 4-stroke diesel

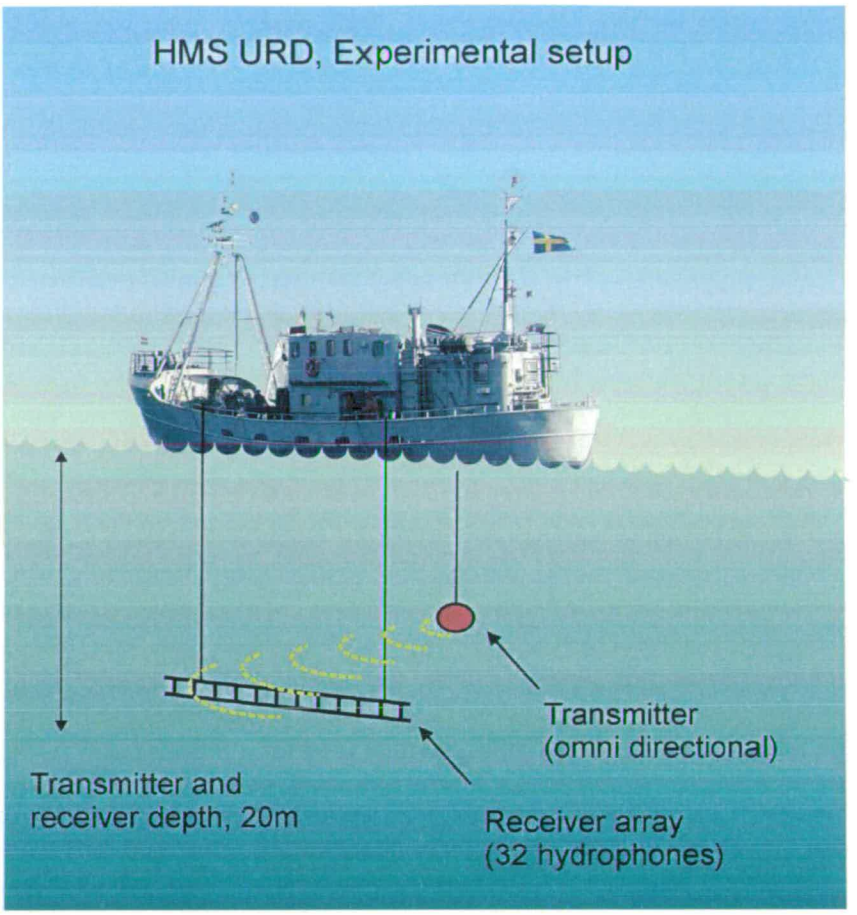


**Figure 2.13:** *Experimental setup for the Utö field trial in September 2001. The ship HMS URD belongs to the Swedish Navy. The underwater acoustic equipment consist of a 12 meter long array with 32 hydrophone, and an omni directional transmitter provided by CETUS AB*

engine as a sound source. The small boat started at a distance of about 600 m from the array of the bottom-mounted hydrophones. The boat was held on straight track runs with constant speed approaching and passing the array by a distance of about 100 m. The engine was held to run constantly at 2712 rpm during the recordings. The radiated underwater sound was recorded by four bottom-mounted hydrophones. The sampling rate was 25 kHz for all recording channels. For more details about the sea trial and recording procedure see, [46], [47], [48]. See Figure 2.15

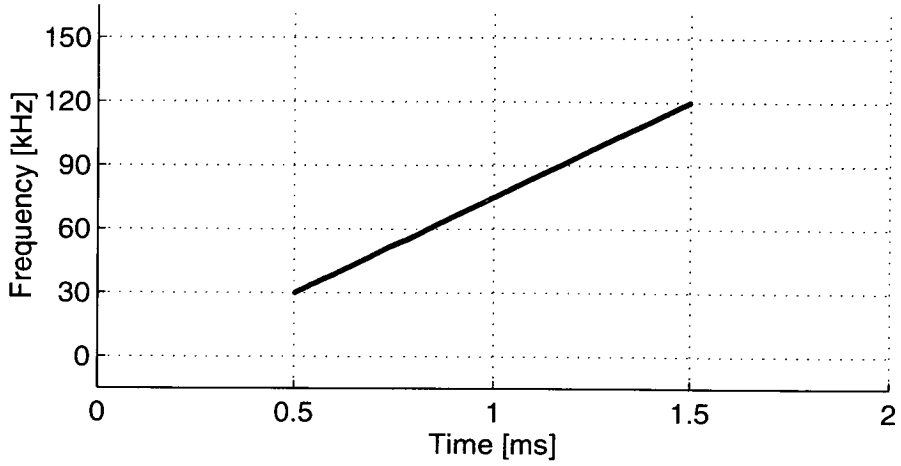


**Figure 2.14:** *The data in from the SAS experiment consists of a number broad band linear chirps, 1ms length, 90kHz centre frequency and 60 kHz bandwidth.*

## 2.6 Traditional Sonar Signal Processing, and result displays

In traditional sonar signal processing the first things that comes in to mind are lofargrams, demon and time bearing plots. An example of how a sonar signal might look like see Figure 2.16, where a time series from the fibre glass boat trial is displayed. Simply by looking at the time series on can see that the CPA (Closest Point of Approach) is at about 60 seconds in, the amplitude is significantly higher than anywhere else in the time series. The signal to noise ratio in this particular experiment is fairly good and the signature of the boat is detectable all the way through from the point where the boat starts to the end of on its roughly 800 meter long track. A lofargram is a spectrogram displayed with frequency along the horizontal axis and time running down the vertical axis. The lofargram is continuously updated with new data coming in from the top of the display. The lofargram is used for narrowband processing of signals and are the most commonly used tool among sonar operators, be it on board a submarine or a surface vessel, see Figures 2.17, 2.18, 2.19, and 2.20. It is an excellent tool to pick up the tonal sounds from machinery (engine, drive, propeller). The tonals leave vertical lines on the display that are easily detected by the trained eye of an operator. The lofargram display is made out of Fourier transformed segments of data, the length of each segment is decided by the operator and is called integration time among the sonar operators, or simply as FFT length for everyone else. To produce the lofargram the data segments are, as stated earlier placed on top of previous Fourier transformed segments. This is why it is such a powerful tool for narrowband analysis of the data, since the Fourier transform is a transform that decomposes the data into sinusoids, see Chapter 4 for a more through description of transforms and time frequency displays.

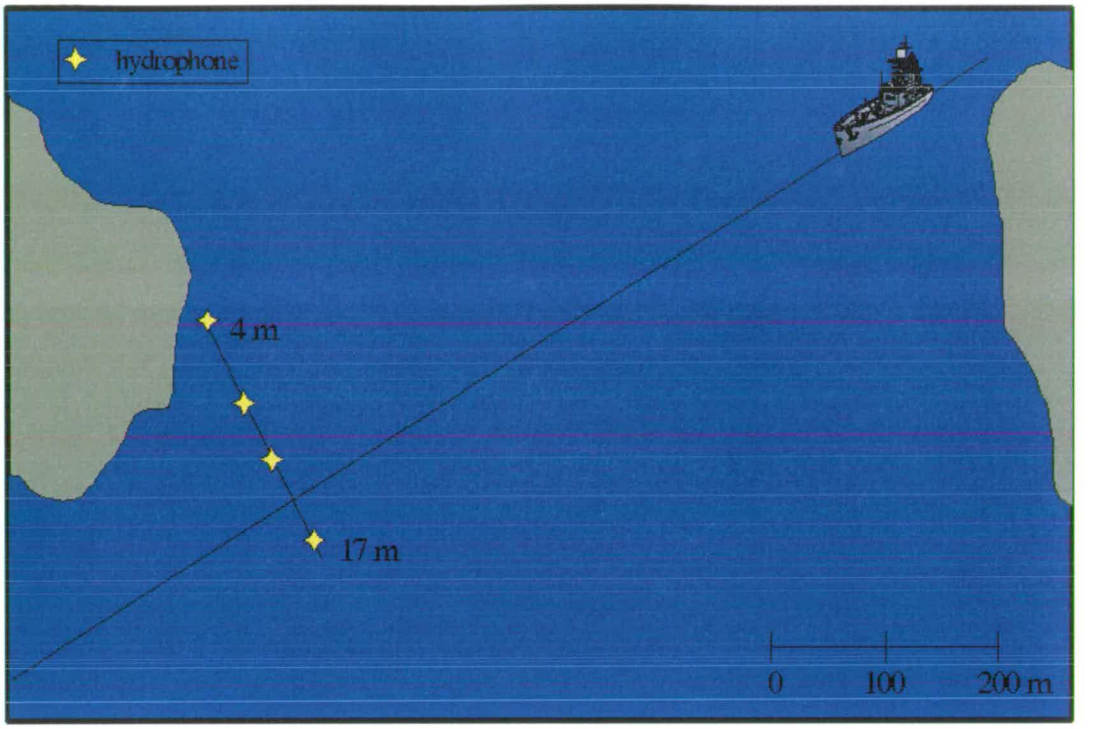


**Figure 2.15:** *The experimental setup for the fibreglass boat trial.*

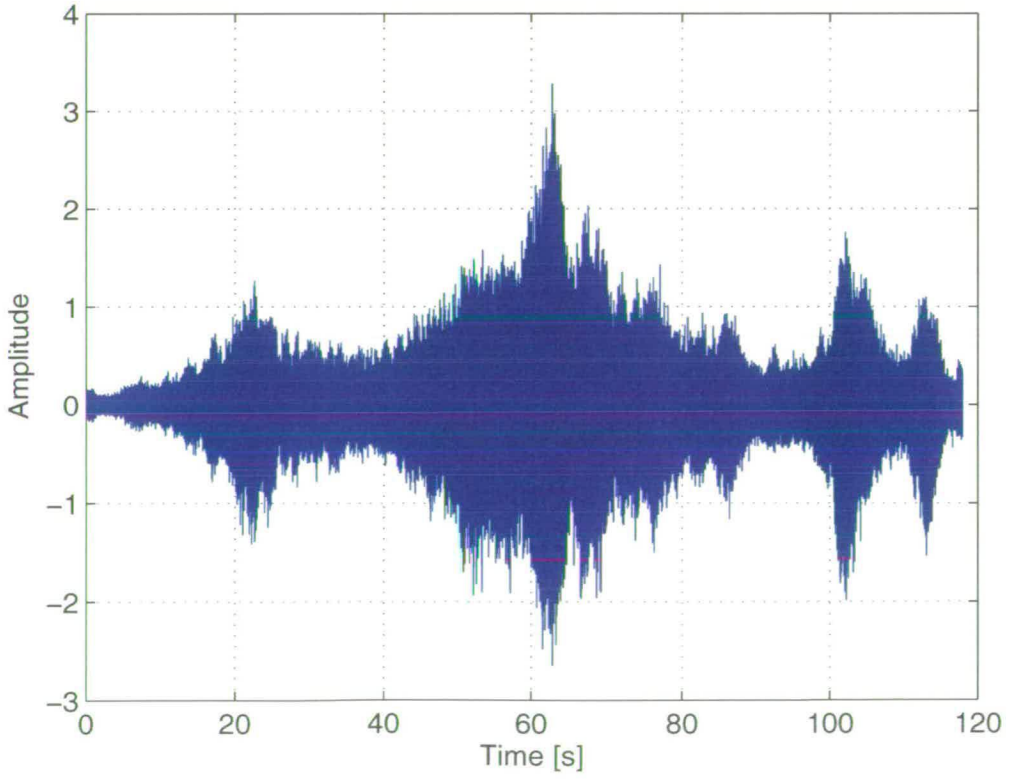


**Figure 2.16:** *A time series from the glass fibre boat experiment described in Section 2.5.6.*

A sonar operator has control over two parameters, the integration time and the frequency range of the display. In order to emphasise different parts of the spectrum, or lofargram for that matter, these two parameters are changed. It is easy to see what the effect of changing the FFT-length is. The display in Figure 2.17 is very blocky along the frequency axis and smooth along the time axis. This follows from the data having a sampling frequency of 2500 Hz and using a FFT-length of as little as 128 samples the frequency resolution is very low (2500/128 Hz). It is also easy to see the change in the following Figures 2.18, 2.19, and 2.20, with increasing FFT-length, the tonal lines appear clearer and the improvement is obvious to a FFT-length of 2048 samples. However, going to the other extreme with 16384 samples the lofargram does not appear any clearer in frequency, the tonals are not easier to detect and it is getting blocky along the time axis. In both Figures, 2.18, and 2.19 the tonal lines are clearly visible and so are the variations over time. The difference between them is as stated before, the time and frequency resolution. Which one of them is to be considered better is largely up to the observers preferences, and perhaps the ideal FFT-length would be in between 512 and 2048 (1024). Also see Chapter 3 where stationarity lengths are discussed, as this gives a hint of what FFT-length to use.

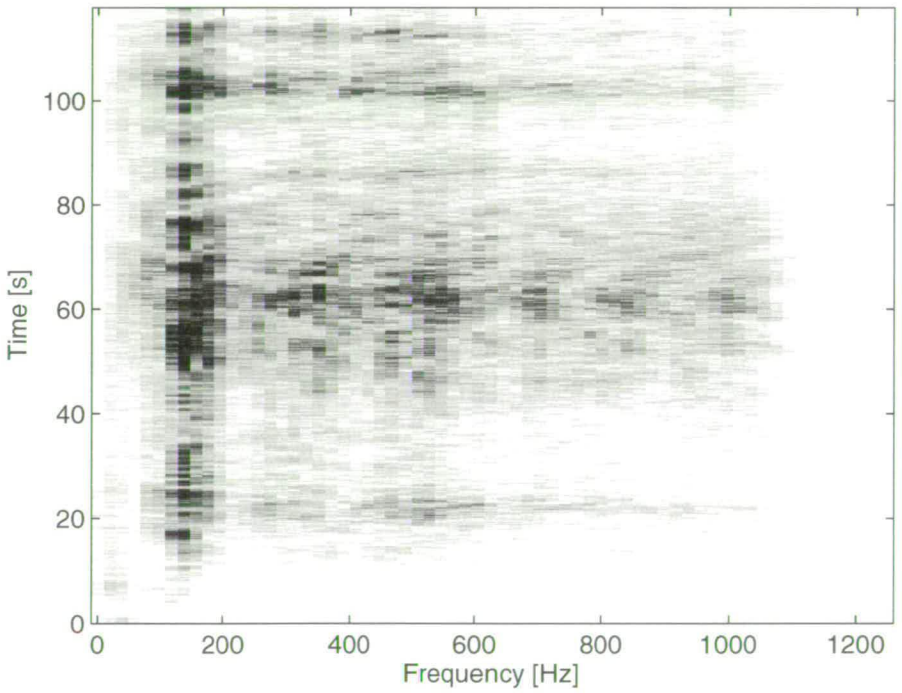


**Figure 2.17:** *A Lofargram display using a FFT length of 128 samples.*

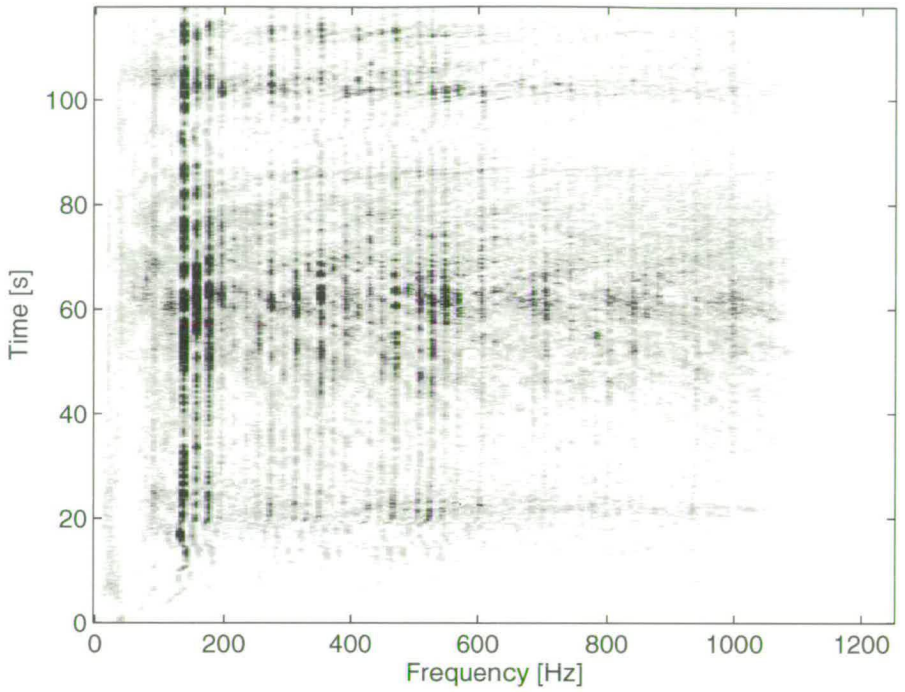


**Figure 2.18:** *A Lofargram display using a FFT length of 512 samples*

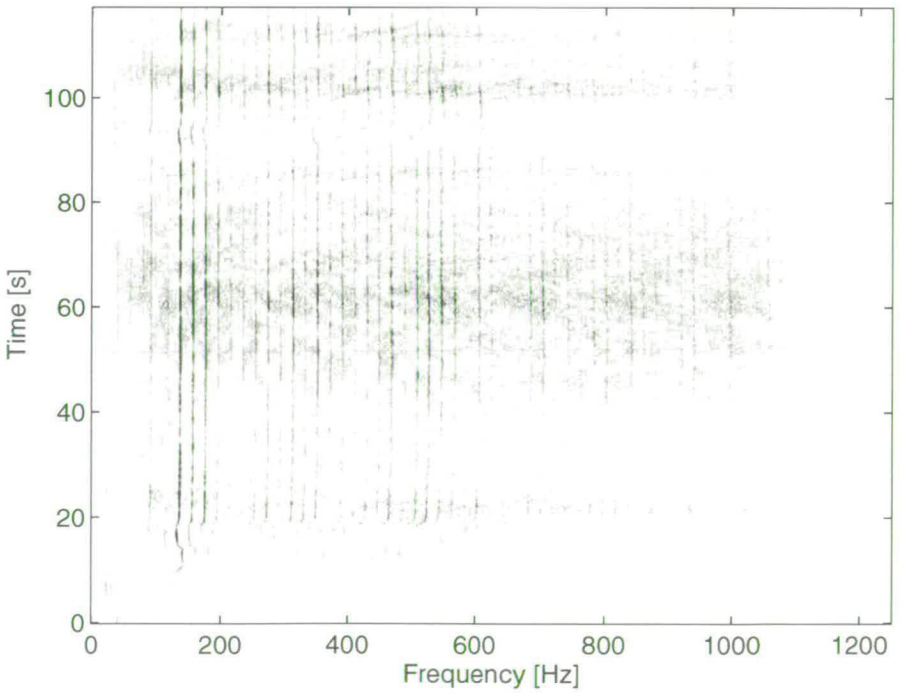


**Figure 2.19:** *A Lofargram display using a FFT length of 2048 samples*

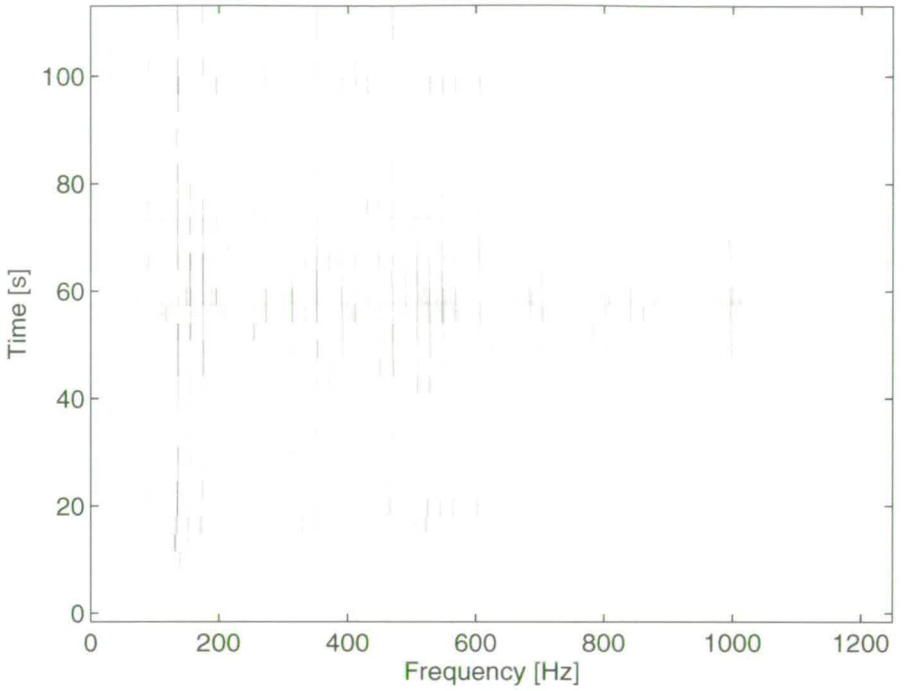


**Figure 2.20:** *Lofargram display using a FFT length of 16384 samples*

The lofargram is used to detect and to identify different vessels as all vessels have a different set of tonals when used, and that is what is called a signature as explained in Section 2.2.5. The lofargram analysis method is also called LOFAR and it stands for LOw Frequency Analysis and Ranging, as all tonals from machinery are inherently low frequency in their nature.

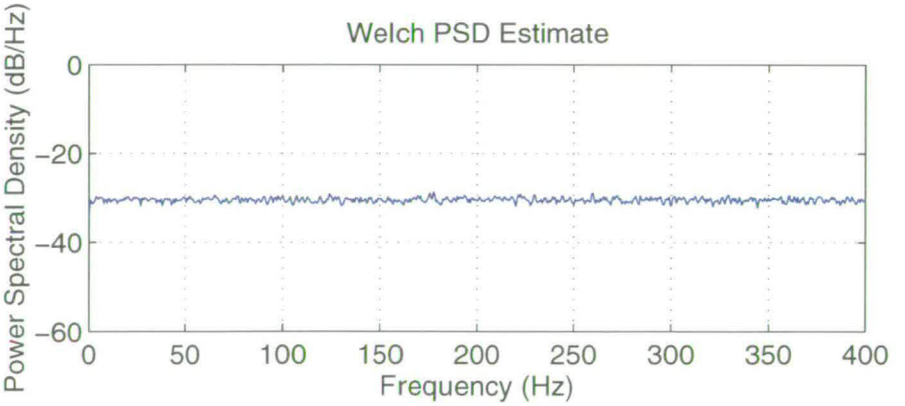


**Figure 2.21:** *A power spectrum estimate from synthetic data consisting of a sum of* $N(0,1$ *noise and modulated* $N(0,1)$ *noise.*

The demon display is another analysis method and it is an acronym for DEMOdulated Noise

[49], see Figure 2.22. The idea is that a narrow band signal carrying information about a target has modulated a broadband noise process. This is an example of cavitation noise, where collapsing bubbles are modulated by the propeller. The DEMON analysis that is presented in Figure 2.22 is done on synthetic data that consists of a sum of $N(0,1)$ noise and $N(0,1)$ noise modulated at 25Hz, $f(t)$. The amplitude for the modulated noise is half the amplitude for the non-modulated noise. The first step of the DEMON analysis is to segment the signal into $N$ segmants. Then each segment $f_i(t)$ is demodulated, ie

$$q_i(t) = f_i(t)e^{-j2\pi f_c/f_s t}, \tag{2.14}$$

where $f_c$ is the center frequency, $f_s$ the sampling frequency. Then the absolute value of each segment is obtained and the segments of demodulated signal forms a time/frequency surface, such as the one in Figure 2.22. In this specific case the frequency is known and the analysis can be done for that frequency band. In cases where the modualting frequency is unknown the analysis is done for an arbitrary number of modulating frequencies in order to cover the frequency band of interest.

Looking at Figure 2.21 one finds no discernible peaks in frequency. That is expected since there are no pure tonals in the data. However, by demodulating the noise and then doing the frequency analysis over again one finds a clear line in the demon plot and in the averaged demon plot in Figure 2.23, defined in Equation (2.15).

$$P_D = \sum_{i=0}^{N} |q_i(t)|, \tag{2.15}$$

where $q_i(t)$ is a demodulated segment of data and $N$ is the number of segments.

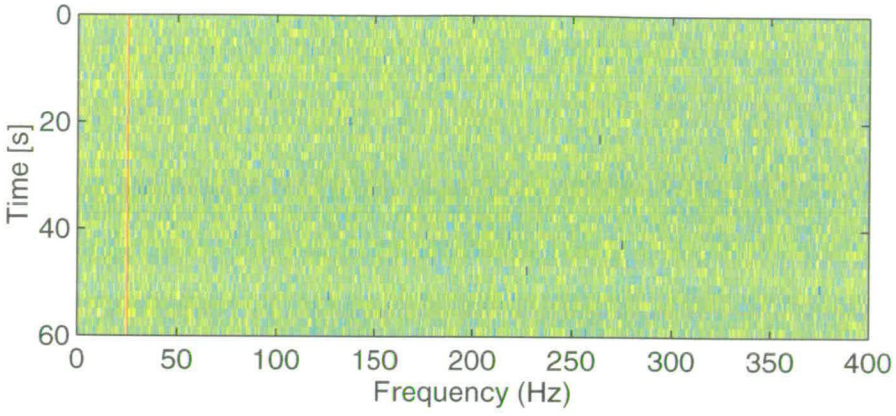


**Figure 2.22:** *DEMON (DEMOdulated noise) display of the synthetic data described above and in Figure 2.21*

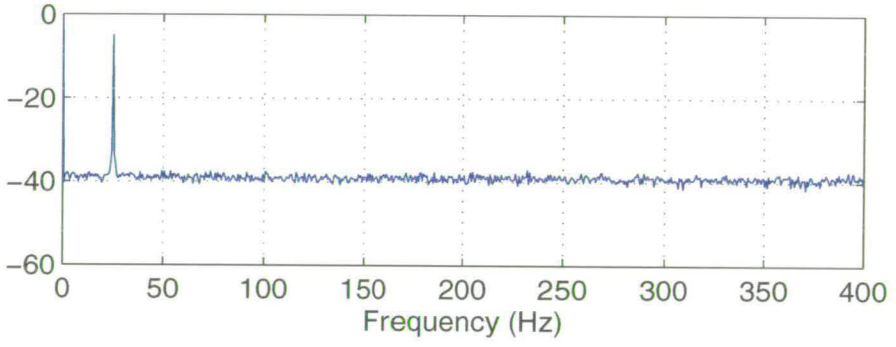


**Figure 2.23:** *An averaged DEMON display*

The above mentioned techniques are applied mainly to passive sonar, although they can be used on active sonar data as well. Especially different time/frequency displays can be useful for use with active data.

BTR (Bearing Time Record) plots are also commonly used when a target is identified then a tracker is set to it and the motion of the target can be followed on the time bearing display. The time bearing plot is exactly what it sounds like, an image of bearing versus time, see Figure 2.24.

Yet another type of bearing time display is the PPI plot (Planned Position Indicator) and since it is primarily used on active data the distance can be given as displayed in Figure 2.25.

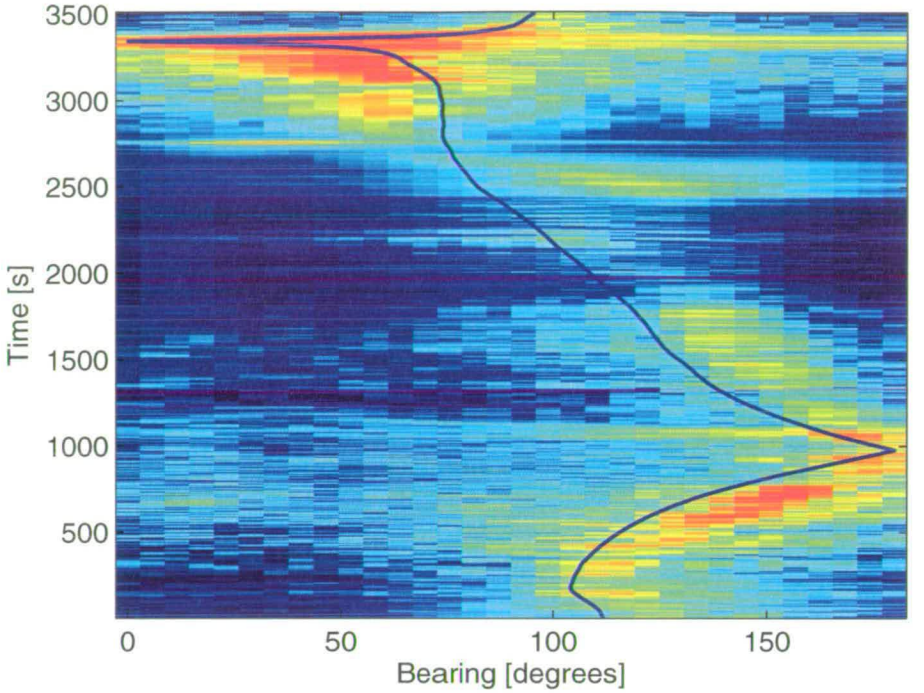


**Figure 2.24:** *BTR plot from the multi sensor trial during August 2004*

## 2.7 Summary

This brief introduction to underwater acoustics is from a rather practical point of view introducing the sonar equations that are used on a daily basis to estimate detection ranges and to design sonar systems. The surrounding media and its properties have been discussed briefly to provide a basic understanding of the difficulties of underwater acoustics and how that affects the signal processing techniques that are used. From a signal processing perspective, knowing the benefits and limitations of the methods, can justify a lower detection threshold and a higher probability of detection. There are after all significant advantages to gain under the assumption that without any knowledge about the received signal the detection threshold is set approximately 6 dB higher than in the case of exact knowledge about the emitted sound. So anything that can give a performance gain from the unknown signal case, up to or equal to knowing the signal properties and being able to use matched filtering, is of great value. That provides the direction for the rest of this thesis, signal processing and signal characteristics.

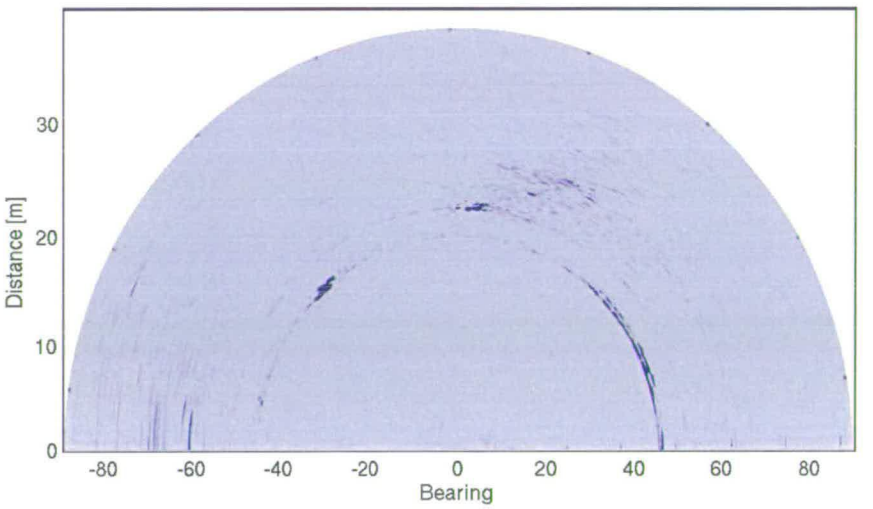


**Figure 2.25:** *PPI plot from the SAS trial*

# Chapter 3
# Statistical Characterisation of Sonar Data

In order to apply signal processing schemes correctly and obtain reliable results at the desired performance, it is necessary to understand the limitations and abilities of the methods in question. The key to this is the understanding of the underlying time series that is to be processed. Some assumptions about signals (or stochastic processes for that matter) are so widely and commonly used that they are almost considered as absolute truths. Most of the time signals are assumed to be stationary, if the opposite is true the results from the analysis are at best, worse than expected, and possibly erroneous and misleading. Another common assumption is that all noise is Additive White Gaussian Noise (AWGN), which of course makes for easier calculations but is not always true. In this thesis it is found that the stationarity characteristics of sonar data vary considerably from location to location over time. In order to assess these properties a number of tests are performed on the sonar data. The performance of these tests are also assessed using synthetic data. First however, this chapter begins with an introduction to certain aspects of mathematical statistics, stochastic processes, probabilities, and statistical signal processing. Then the tests and estimators for the relevant signal characteristics are defined and evaluated.

## 3.1 Distribution functions and probability density functions

The distribution function of a stochastic process $x$, is defined as follows

$$F(x) = \int_{-\infty}^{\infty} f(x)dx \tag{3.1}$$

where $f(x)$ is the pdf (probability density function). The distribution function is also called a CDF (Cumulative Distribution Function).

The concept of distribution functions and probability density functions are explained in great detail in all textbooks on mathematical statistics, but as a brief introduction to this section the example of throwing a die is examined. The throwing of a die has six possible outcomes, if $x$ is the outcome of the throwing of a die, then $x \in \{1, 2, 3, 4, 5, 6\}$. All six outcomes are equally probable. $P(x = X) = \frac{1}{6}$ and $\sum_{X=1}^{6} p(x = X) = 1$, which can be seen in Figure 3.1. Looking at the distribution function one finds the probability of $x \leq X$ is the sum of the outcomes. So for $P(x \leq 3) = 0.5$, which makes perfect sense, since it is intuitively easy to grasp that throwing 3 or less is just as likely as throwing 4 or more. Expanding that to the probability of $a \leq x \leq b$ it is easily found that by looking at the area under the pdf between a and b one finds the probability of $x$ being in that interval. For instance, $P(1 \leq x \leq 5) = \sum_{X=1}^{5} p(x = X) = \frac{5}{6}$.

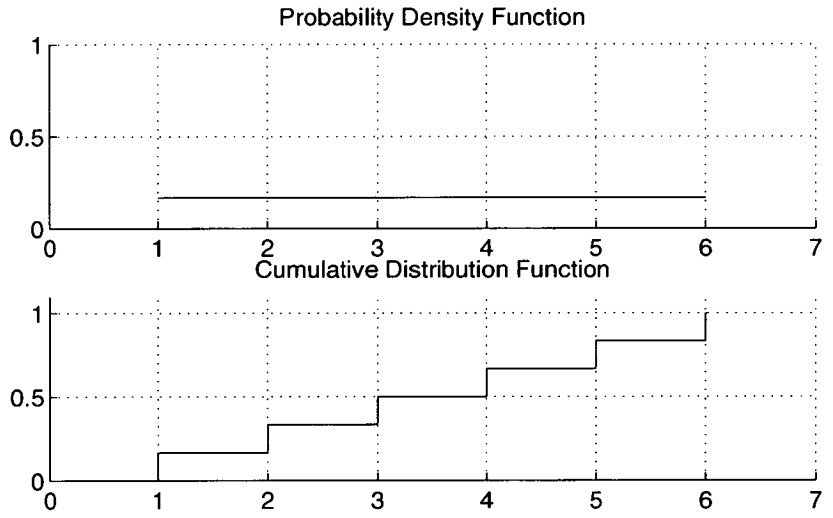


**Figure 3.1:** *The probability density function and the distribution function for throwing a die.*

A vector $X$ consisting of samples from a Gaussian process is produced. A Gaussian stochastic variable takes values according to the Gaussian pdf below

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-m}{\sigma})^2} \tag{3.2}$$

where $m, \sigma^2$ denotes mean and variance respectively. The mean and variance among other properties a pdf and a CDF can have will be addressed later in this chapter, in Section 3.6, and are defined in Section 3.2. Instead of the throwing a die case the variables in $X$ can take any

value that the Gaussian pdf allows with probability

$$P(a \leq x \leq b) = \int_a^b p(x)dx = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-m}{\sigma})^2} dx. \tag{3.3}$$

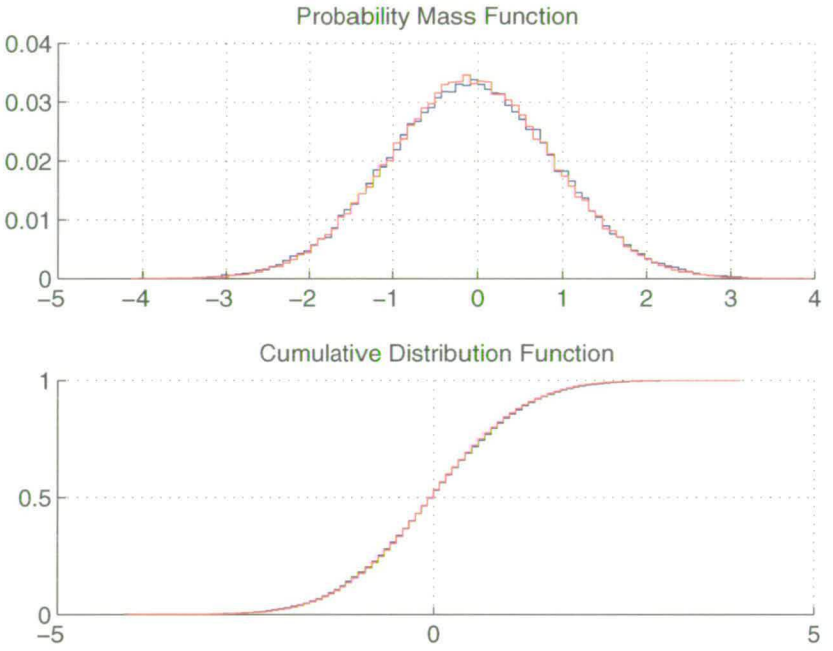


**Figure 3.2:** *The estimated probability density functions and cumulative distribution functions for two vectors of Gaussian ($N(0,1)$) variables.*

The CDF can be estimated by making a histogram of the vector $X$ and then cumulatively summing the number of outcomes in each bin. The CDF is produced using 100 bins and it is easy to see the quantisation on both the pdf and the CDF. With a longer vector $X$, more data samples, with the same pdf, and a larger number of bins the curves would appear smoother. In Figure 3.2 two CDFs and pdfs are illustrated. They are produced from two vectors of Gaussian variables, with a length of 1 million samples each. The vectors are of equal mean and variance, so in second order statistical terms, there is no statistically significant difference. In Figure 3.3 the pdfs are shown in a semi log plot, in order to enhance the difference between the two functions. It worth noting that in this way it is possible to see the difference in the tails of the distribution functions.

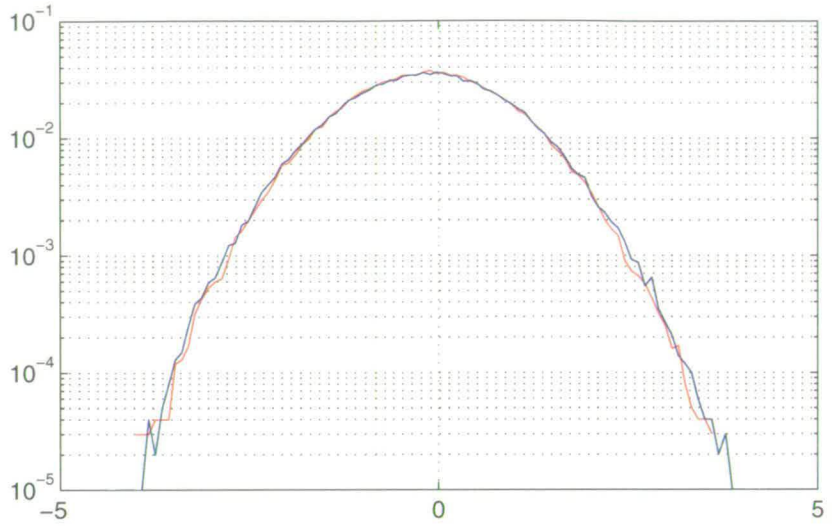


**Figure 3.3:** *A log scale plot (logarithmic y-axis and linear x-axis) of the estimated probability mass functions for the same two vectors of Gaussian ($N(0,1)$) variables as displayed in Figure 3.2.*

## 3.2 Moments and cumulants

The nth-order moment of a stochastic process with probability density function $f(x)$ is defined as

$$\mu_n = E\{x^n\} = \int_{-\infty}^{\infty} x^n f(x) dx \tag{3.4}$$

the first characteristic function for a distribution is defined as

$$\Phi(t) = E\{e^{jxt}\} = \int_{-\infty}^{\infty} e^{jxt} f(x) dx \tag{3.5}$$

and it is called the moment-generating function, [50], Chapter 5, [51] Chapters 10 and 15, and it is the mean value for $e^{jxt}$ where $x$ had the probability density function $f(x)$. Looking at the equation above one realises that the characteristic function $\Phi(t)$ is the inverse Fourier transform of the probability density function. The reciprocal of the equation above is the Fourier transform of $\Phi(t)$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-jxt} \Phi(t) dt. \tag{3.6}$$

In some literature, [51] the pdf $f(x)$ is termed a frequency function.

The relation between the first characteristic function and the moments is derived below. Since there is a one to one relationship between the characteristic function and the pdf this derivation also provides the relationship between the moments and the pdf, thus showing how the pdf can be described by the moments. Assuming that the $n$-th derivative of $\Phi(t)$ with respect to $t$ at $t = 0$ exists, it is equal to

$$\frac{d^n\Phi(t)}{dt^n}|_{t=0} = (j)^n \int_{-\infty}^{\infty} x^n f(x)dx = (j)^n \mu_n. \tag{3.7}$$

Then near $t = 0$ the MacLaurin series expansion of $\Phi$ yields

$$\Phi(t) = 1 + \sum_{n=1}^{k} \frac{\mu_n (jt)^n}{n!} \tag{3.8}$$

and to arrive at Equation (3.7) the following steps are made. Since the MacLaurin formula for a function $g(x)$ is given by

$$g(x) = \sum_{n=-\infty}^{\infty} \frac{g(x)^{(n)}(0)x^n}{n!} \tag{3.9}$$

where $g(x)^{(n)}$ denotes the $n$th derivative of $g(x)$. The value for $\Phi(0)$ and for the derivatives of $\Phi^{(n)}$ are given in Equation (3.10)

$$\begin{aligned}
\Phi(0) &= \int_{-\infty}^{\infty} f(x)dx = 1 \\
\Phi'(0) &= j \int_{-\infty}^{\infty} xf(x)dx = j\mu_1 \\
\Phi''(0) &= (j)^2 \int_{-\infty}^{\infty} x^2 f(x)dx = (j)^2 \mu_2 \\
&\vdots = \vdots \\
\Phi^{(n)}(0) &= (j)^n \int_{-\infty}^{\infty} x^n f(x)dx = (j)^n \mu_n
\end{aligned} \tag{3.10}$$

and then the coefficients of the MacLaurin series expansion of $\Phi(t)$ are

$$\Phi(t) = \Phi(0) + \frac{\Phi^{(')}(0)}{1}t + \frac{\Phi^{(")}(0)}{2}t^2 + \ldots + \frac{\Phi^{(k)}(0)}{k!}t^k + o(t^k) \tag{3.11}$$

and the remainder $o(t^k) \rightarrow 0$, according to

$$o\{t^k\} = \frac{\Phi(\Theta t) - \Phi(0)}{n!} \quad \rightarrow 0, \quad (0 < \Theta < 1) \tag{3.12}$$

since $\Phi(\Theta t) - \Phi(0)$ tends to zero with $t$, Chapter 12.1 in [51].

The following expression is attained from Equation (3.8) by substituting $(jt)$ with $s$

$$\Phi(s) = 1 + \sum_{n=1}^{k} \frac{\mu_n(s)^n}{n!} \tag{3.13}$$

From the equations above it is easy to see that, by differentiating the characteristic function at the origin ($s = 0$), the moments of order n are equal to the nth derivative of the characteristic function. Which yields $\Phi^{'} = \mu_1$, $\Phi^{"} = \mu_2, \ldots$, and $\Phi^{(n)} = \mu_n$. This is why $\Phi(s)$ is called the moment-generating function.

The second characteristic function is the logarithm of the first characteristic function

$$\Psi(t) = \ln(\Phi(t)) \tag{3.14}$$

and for $\Psi(t)$ there is a corresponding MacLaurin series expansion. Then by using

$$\ln(1+z) = \frac{z}{1} - \frac{z^2}{2} + \ldots \pm \frac{z^k}{k} + o(z^k) \tag{3.15}$$

and replacing $1 + z$ with $\Psi$, and rearranging the expression, thus arriving at

$$\Psi(t) = \ln(\Phi(t)) = \sum_{n=1}^{k} \frac{\kappa_n(jt)^n}{n!} \tag{3.16}$$

and again substituting $(jt)$ with $s$ the expression for $\Psi(t)$ is obtained.

$$\Psi(s) = \sum_{n=1}^{k} \frac{\kappa_n(s)^n}{n!}. \tag{3.17}$$

The coefficients $\kappa$ are called cumulants and $\Psi(s)$ is the cumulant-generating function. Just as

the derivatives of the first characteristic function are the moments, the derivatives of the second characteristic function are the cumulants. So $\Psi' = \kappa_1$, is the first order cumulant, $\Psi'' = \kappa_2$ is the second order cumulant, ..., and $\Psi^{(n)} = \kappa_n$ is the n:th order cumulant.

The derivation above is valid for a scalar stochastic variable $x$ but can be extended to the case where all the expressions above are given for a two-dimensional variable according to [51] Chapter 21.3. Thereafter according to [51] Chapters 11, and 22.4 the expressions can be given for a n-dimensional stochastic variable. The expressions for the general n-dimensional case are presented below but some of the detail in the derivation of the expressions is left out. For a n-dimensional variable or a vector of random variables $X = [X_1, X_2, \ldots, X_n]^T$ the nth-order moment is defined as

$$\mu_n([X_1, X_2, \ldots, X_n]) = E\{[X_1, X_2, \ldots, X_n]^n\} \tag{3.18}$$

Equivalent to the one-dimensional case there is a characteristic function for the vector $X$

$$\Phi_X(\xi) = E\{e^{j\xi^T X}\}. \tag{3.19}$$

It can be shown following the same steps as for the one-dimensional case that the characteristic function is related to the moments as follows

$$\mu_n([X_1, X_2, \ldots, X_n]) = (-j)^n \frac{\partial \Phi_X(\xi)}{\partial \xi_1 \ldots \xi_n}|_{\xi=0}. \tag{3.20}$$

Then substituting $j\xi$ with $\lambda$, and thus arriving at

$$M_X(\lambda) = E\{e^{\lambda^T X}\} \tag{3.21}$$

which is the moment generating function and using a MacLaurin expansion it is found that the nth-order moments of $X$ are the coefficients of $\prod_{i=0}^{n} \lambda_i$, and the moment generating function can be written as

$$\mu_n([X_1, X_2, \ldots, X_n]) = \frac{\partial M_X(\lambda)}{\partial \lambda_1 \ldots \lambda_n}|_{\lambda=0}. \tag{3.22}$$

The second characteristic function is as stated in Equation (3.14) is the logarithm of the first characteristic function, and is defined below for the vector $X$

$$\Psi_X(\xi) = \ln(\Phi_X(\xi)) = \ln(E\{e^{j\xi^T X}\}) \tag{3.23}$$

then substituting $j\xi$ with $\lambda$ the resulting function

$$K_X(\lambda) = E\{e^{\lambda^T X}\} \tag{3.24}$$

is the cumulant generating function. Then by following the same steps of MacLaurin series expansion the relation between the cumulant generating function and cumulants is found.

$$\kappa_n([X_1, X_2, \ldots, X_n]) = (-j)^n \frac{\partial \Psi_X(\xi)}{\partial \xi_1 \ldots \xi_n}|_{\xi=0} = \frac{\partial K_X(\lambda)}{\partial \lambda_1 \ldots \lambda_n}|_{\lambda=0} \tag{3.25}$$

where $\kappa_n(X) = \kappa_n([X_1, X_2, \ldots, X_n])$ denotes the nth-order cumulant for the random vector $X$. The moments and cumulants are obviously related to one and other, as it can be seen from the derivation of both. The explicit relationship between them can be seen below and the derivation of them can be found in [52], and the proof is found in [53]

$$\mu(X) = \sum_{P(n)\in\mathcal{P}(n)} \prod_{\mathbf{m}\in P(n)} \kappa(X_{\mathbf{m}}) \tag{3.26}$$

$$\kappa(X) = \sum_{P(n)\in\mathcal{P}(n)} (-1)^{r-1}(r-1)! \prod_{\mathbf{m}\in P(n)} \mu(X_{\mathbf{m}}) \tag{3.27}$$

where $\mathbf{n}$ is the set $\{1, 2, \ldots, n\}$, and $\mathbf{m}$ is a subset of $\mathbf{n}$, $X_{\mathbf{m}}$ is a vector of elements $X$ with indexes $\mathbf{m}$. $\mathcal{P}(n)$ is the set of all partitions of $\mathbf{n}$, and P(n) is a partition of $\mathbf{n}$. The expressions above in Equations (3.26) and (3.27) are usually written out for $n$ = 1,2,3, and 4

The mean of the vector $X$ is the first-order cumulant and moment

$$\kappa_1(X_1) = \mu_1(X_1) = E\{X_1\} \tag{3.28}$$

The second-order cumulant and moment, are also equal to the covariance of the vector $X$ as,

$$\kappa_2(X_1, X_2) = \mu_2(X_1, X_2) - \mu_1(X_1)\mu_1(X_2). \tag{3.29}$$

The third-order cumulant and moment relations are,

$$\begin{aligned}\kappa_3(X_1, X_2, X_3) = \; & \mu_3(X_1, X_2, X_3) - 3\mu_1(X_1)\mu_2(X_2, X_3) \\ & +2\mu_1(X_1)\mu_1(X_2)\mu_1(X_3) \quad . \end{aligned} \tag{3.30}$$

The fourth-order cumulant and moment relations are,

$$\begin{aligned}\kappa_4(X_1, X_2, X_3, X_4) = \; & \mu_4(X_1, X_2, X_3, X_4) - \mu_2(X_1, X_2)\mu_2(X_3, X_4) \\ & -\mu_2(X_1, X_3)\mu_2(X_2, X_4) - \mu_2(X_1, X_4)\mu_2(X_2, X_3) \\ & -\mu_1(X_1)\mu_3(X_2, X_3, X_4) - \mu_1(X_2)\mu_3(X_1, X_3, X_4) \\ & -\mu_1(X_3)\mu_2(X_1, X_2, X_4) - \mu_1(X_4)\mu_2(X_1, X_2, X_3) \\ & +2\mu_2(X_1, X_2)\mu_1(X_3)\mu_1(X_4) + 2\mu_2(X_1, X_3)\mu_1(X_2)\mu_1(X_4) \\ & +2\mu_2(X_1, X_4)\mu_1(X_2)\mu_1(X_4) + 2\mu_2(X_2, X_4)\mu_1(X_1)\mu_1(X_3) \\ & +2\mu_2(X_3, X_4)\mu_1(X_1)\mu_1(X_2) + 2\mu_2(X_2, X_3)\mu_1(X_1)\mu_1(X_4) \\ & -6\mu_1(X_1)\mu_1(X_2)\mu_1(X_3)\mu_1(X_4). \end{aligned} \tag{3.31}$$

Then for real valued, zero mean, moments and cumulants it simplifies to the following
The first-order cumulant and moment;

$$\kappa_1(X_1) = \mu_1(X_1) \tag{3.32}$$

The second-order cumulant and moment;

$$\kappa_2(X_1, X_2) = \mu_2(X_1, X_2) \tag{3.33}$$

The third-order cumulant and moment;

$$\kappa_3(X_1, X_2, X_3) = \mu_3(X_1, X_2, X_3) \tag{3.34}$$

The fourth-order cumulant and moment;

$$\kappa_4(X_1, X_2, X_3, X_4) = \mu_4(X_1, X_2, X_3, X_4) - \mu_2(X_1, X_2)\mu_2(X_3, X_4)$$
$$-\mu_2(X_1, X_3)\mu_2(X_2, X_4) - \mu_2(X_1, X_4)\mu_2(X_2, X_3) =$$
$$\mu_4 - 3\mu_2^2. \tag{3.35}$$

## 3.3 Properties of the characteristic function

It was found that the characteristic function is related to the pdf via the Fourier transform. This one to one relationship between them makes it possible to find the other if one is known or estimated. That also implies that the properties for the characteristic function are similar to those of the FT. The properties of distribution functions and probability density functions were described in Section 3.1 The characteristic function has the following properties:

$$\Phi(0) = 1$$

$$|\Phi(t)| \leq \int_{-\infty}^{\infty} f(x)d(x) = 1$$

which follows from $|e^{ixt}| = 1$ and the integral over a pdf is always equal to one.

$$\Phi(-t) = \overline{\Phi(t)}$$

where $\overline{a}$ denotes the complex conjugate of $a$.

## 3.4 Properties of Moments and cumulants

Although both the moments and cumulants can be used to describe a distribution function there are a few benefits using cumulants over using moments. First the properties that concern both are listed

**Both moments and cumulants are symmetric** in their arguments, i.e.

$$\kappa_n(\tau_1, ..., \tau_n) = \kappa_n(\tau_{i1}, ..., \tau_{in})$$

$$\mu_1(\tau_1, ..., \tau_n) = \mu_n(\tau_{i1}, ..., \tau_{in})$$

, where $(i_1, ..., i_n)$ is a permutation of $(1, ..., n)$.

**Both are additive in their arguments** , i.e.

$$\kappa_n(X_1 + Y_1, Z_2, .., Z_n) = \kappa_n(X_1, Z_2, .., Z_n) + \kappa_n(Y_1, Z_2, .., Z_n)$$

$$\mu_n(X_1 + Y_1, Z_2, .., Z_n) = \mu_n(X_1, Z_2, .., Z_n) + \mu_n(Y_1, Z_2, .., Z_n)$$

**For constant** $a_1, ..., a_n$

$$\kappa_n(a_1X_1, ..., a_nX_n) = \kappa_n(X_1, .., X_n) \prod_{i=1}^{n} a_i$$

$$\mu_n(a_1X_1, ..., a_nX_n) = \mu_n(X_1, .., X_n) \prod_{i=1}^{n} a_i$$

**For** $m$ **independent vector** valued random variables $X^r$, where $r = 1, 2, ..., m$ and $X^r$ has the components $X_1^r, ..., X_n^r$ then

$$\kappa_n(\sum_{r=1}^{m} X_r^1, ..., \sum_{r=1}^{m} X_r^n) = \sum_{r=1}^{m} \kappa_n(X_r^1, ..., X_r^n)$$

**For two independent vectors** $X_1$, and $X_2$, the cumulants $\kappa(X_1, X_2) = 0$, which is not true in general for moments, $\mu(X_1, X_2) \neq 0$.

**For a Gaussian process** the cumulants of higher than second order are zero, since the characteristic function for a jointly Gaussian distributed random variables $X_i, i = 1, ..., k$, with mean $\mu$ and the covariance matrix $R$ is given by

$$\phi(\xi) = e^{(j\xi\mu - \frac{1}{2}\xi^T(R)\xi)}.$$

The first and second order cumulants are the mean and covariance. However, the higher order moments are not zero in this case.

## 3.5 Estimation

Given the fact that for real life signals the pdf is very seldom known, the moments or cumulants have to be estimated in order to be able to describe the signal. In this section the estimation of the first four orders of central moments, mean, variance, skewness and kurtosis is the main scope. Since the moments and cumulants are closely related to frequency spectra of corresponding orders, the estimation procedures are discussed from that point of view, [54], [55], [56], [57]. Another reason for bringing up spectral estimation is that the practicalities for spectral estimates are the same as for the moments and cumulants.

In spectral estimation there are two main approaches to the estimation process, parametric and non-parametric. Parametric methods offer the potential for lower variance and higher resolution than the parametric ones but the suffer from the fact that it is difficult to chose the type of model AR, MA, or ARMA (Auto Regressive, Moving Average). Furthermore it is difficult to chose the model order. This is especially difficult for data from in an environment that changes over time. Non-parametric methods do not have these difficulties and once the underlying data is fairly well understood there are less need to make restrictive assumptions about the system.

Estimation theory is a field on it's own and this section is merely a brief overview, included here since the measures used in the statistical tests are estimated from real sonar data sets.

### 3.5.1 The Power Spectrum and Autocorrelation

In practice the power spectrum is estimated from a single realisation of a random process. The estimation can be done by the direct and indirect method. Either directly Fourier transforming the data vector and squaring the result arriving at the power spectrum or using the ACF for the vector of data and then Fourier transforming that to obtain the power spectrum (using the Wiener-Khintchine theorem). For practical purposes the direct method is convenient to use since it is computationally efficient. Going via the ACF requires extra steps and does not add anything to the quality of the estimate. Estimating the power spectrum using the direct method is described below.

The N-point, Discrete Fourier Transform (DFT) is defined as follows,

$$X^n(k) = \sum_{i=1}^{N} x(t)e^{-j2\pi kt/N} \tag{3.36}$$

and from the DFT follows the power spectrum for L segments of data.

$$P(k) = \frac{1}{L}\sum_{1}^{L} X^n(k)X^{n^*}(k). \tag{3.37}$$

A periodogram is estimated by windowing the data and then calculating the DFT on L data segments of length N. Each one of these segments is a part of the power spectrum above as,

$$P_i(k) = X^n(k)X^{n^*}(k). \tag{3.38}$$

So the power spectrum can also be defined as

$$P(k) = \frac{1}{L}\sum_{i=1}^{L} P_i(k). \tag{3.39}$$

which is also known as the averaged periodogram. There are several ways of estimating the power spectrum, the Bartlett method, the Welch method, the Blackman-Tukey method, just to mention a few. All of the above mentioned methods are different ways of averaging the periodogram, where the Bartlett method is the most basic one. The Welch method uses a modified periodogram and finally, the Blackman-Tukey uses smoothing of the periodogram.

If the objective is estimating the ACF then the process can be reversed. In some cases it makes sense to first obtain the power spectrum and then the ACF from that. Due to the efficiency of the FFT calculations it is usually better (faster) then the convolution process to obtain the ACF directly.

Then again if the objective is to get the second order cumulant of lag zero, the variance, then it is best obtained by the following,

$$\kappa_{2,0} = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{3.40}$$

which is the best unbiased variance estimate. It can however also be obtained from the power spectrum.

### 3.5.2 Higher order Spectra and Cumulants

Estimating the third order spectrum, the bispectrum can be done in a similar way as the power spectrum,

$$B(k,l) = \frac{1}{K}\sum_{i=1}^{K} X_i(k)X_i(l)X_i^*(k+l) \tag{3.41}$$

In this section two ways of normalising the bispectrum are presented. Adopting the terminology in [56] and [57], the normalised bispectral estimates are called skewness and squared bicoherence respectively, and are defined below as

$$s^2(k,l) = \frac{|B(k,l)|^2}{P(k)P(l)P(k+l)} \tag{3.42}$$

which is known as the skewness function [58]. Then the squared bicoherence, [59] is defined as

$$b^2(k,l) = \frac{|B(k,l)|^2}{C(k,l)P(k+l)} \tag{3.43}$$

where

$$C(k,l) = \frac{1}{L}\sum_{i=1}^{L} |X_i(k)X_i^*(l)|^2. \tag{3.44}$$

Higher than third order spectra can of course be defined in a similar fashion, for instance the fourth order spectrum is called a trispectrum. However, that falls outside the scope of this thesis since the intended use of the higher order spectra here is limited to third order spectra.

If the intention is to estimate the cumulants of the desired order the same approach may be adopted, segmenting and averaging. Thus arriving at robust estimates of the desired cumulant.

Special cases of the cumulants are, as mentioned earlier, the lag zero of the third and fourth order cumulants, namely the skewness and kurtosis. The skewness $\gamma_1$ can be estimated as follows

$$\kappa_{3,0} = \gamma_1 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^3/\sigma^3 \tag{3.45}$$

and the kurtosis $\gamma_2$ in a similar fashion as

$$\kappa_{4,0} = \gamma_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^4/\sigma^4 - 3. \tag{3.46}$$

## 3.6 The shape of a distribution function related to the mean, variance, skewness, and kurtosis

As mentioned in section 3.1 the pdf and CDF can be described by the mean, variance, and other higher order cumulants, the skewness and kurtosis. They are the first four orders of cumulants at lag zero. The mean of course not being a function of time lags but never the less the first order cumulant or moment and is included here for the sake of completeness. Starting with the mean of the variable $x$

$$E\{x(t)\} = m \tag{3.47}$$

where $m$ is an arbitrary constant.

Then the autocorrelation function (ACF) $R_x(\tau)$ is defined as,

$$E\{x(t)x(\tau - t)\} = f(\tau) \tag{3.48}$$

Which is equal to the covariance, in this special case. And at $\tau = 0$ it is the variance, $\sigma^2$.

For lag number zero, $\tau_1 = \tau_2 = 0$, the third-order moment and cumulant divided by the cube of the variance of the process $\sigma^3$, is known as the skewness of a process, and it is defined as follows

$$\gamma_1 = \frac{E\{(x-m)^3\}}{\sigma^3} = \kappa_{3,0} \tag{3.49}$$

The fourth order cumulant at lag zero, $\tau_1 = \tau_2 = \tau_3 = 0$, divided by $\sigma^4$, is known as the kurtosis of a process and is defined below as

$$\gamma_2 = \frac{E\{(x-m)^4\}}{\sigma^4} - 3 = \frac{\mu_4 - 3\mu_2^2}{\sigma^4} = \kappa_{4,0}. \tag{3.50}$$

The mean of a process gives the location of the centre of gravity of the pdf. Variance is a measure of dispersion. In other words it gives a measure of the spread or the variation from the mean.

The skewness is a measure of symmetry, so that any symmetrical probability density function has a skewness equal to zero, as is the case for a Gaussian distribution function. A distribution function with a positive skewness has a heavier tail on the right hand side of the centre of gravity, and a negative skewness indicates a heavy tail to the left, see Figures 3.4, and 3.5 for

examples of positively and negatively skewed pdfs.

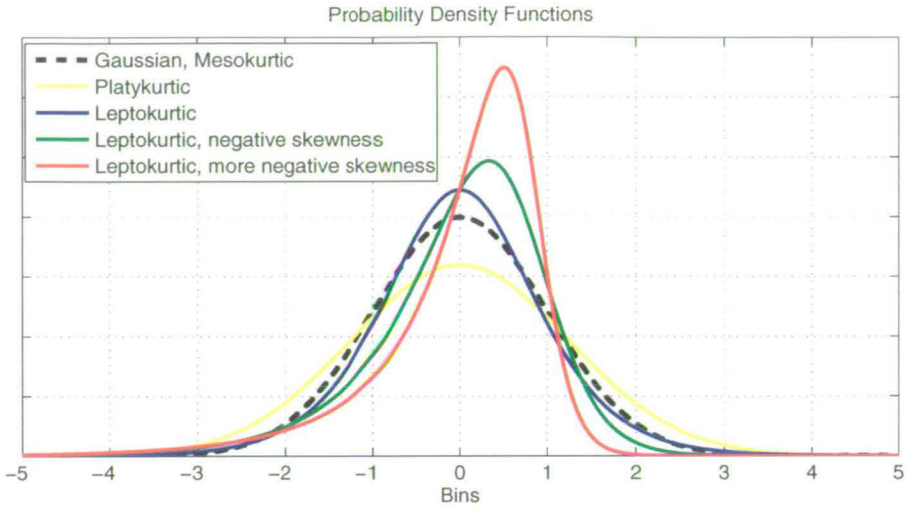


**Figure 3.4:** *An image of negatively skewed pdfs with both lower and higher kurtosis values (platy- and lepto-kurtic) compared to a Gaussian pdf.*

The kurtosis is a measure of the flattening, or pointedness, of the distribution function at its centre. The kurtosis is equal to zero for the Gaussian distribution function and functions that are more pointed than than the Gaussian distribution have a positive kurtosis and flatter distributions have a negative kurtosis value. When the distribution function is flatter than a Gaussian curve it is known as a platykurtic curve, if the curve is neither flatter nor more pointed than the Gaussian it is mesokurtic, and finally if the curve is more pointed than the Gaussian it is leptokurtic, see Figures 3.4, and 3.5 for examples of the different type of pdfs. Platy- and lepto-kurtic are also known as sub and super Gaussian curves. Kurtosis is also known as excess and is defined as above in [51] 15.8.

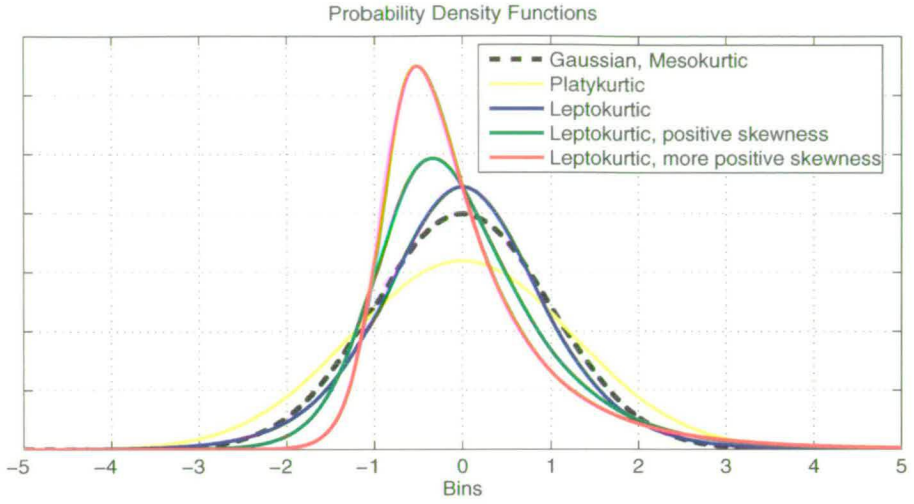


**Figure 3.5:** *An image of positively skewed pdfs with both lower and higher kurtosis values compared to a Gaussian pdf.*

## 3.7 Stationarity

Most signal processing techniques rely on the existence of stationarity in order to produce reliable and consistent estimates. Signals or stochastic processes can be categorised by their stationarity properties, going from stationary to the non-stationary, and other levels of stationarity as described below.

The level of stationary has implications on a number of factors in signal processing, but mainly on the level of reliability of any estimate. Estimates from highly non-stationary data can at times be so bad that the variance of the estimate is by far greater than the estimate itself.

One important parameter that is directly depending on stationarity times is integration time. It comes in as a parameter in most signal processing schemes since there is always a choise on how long data segments to use. In direction of arrival estimation where an important part is the estimation of a covariance matrix, it is crucial to use enough data to get a reliable estimate but it is on the other hand important to limit the amount of data. This is due the fact that the variance of the estimate will not get lower if the data is not stationary up to at least the lenght of the estimation window. This is also important in any estimation where a number of data segments are used to form an average of some sort. Again, if the data is too non-stationary the variance

of the estimate will be high and will not go down with a higher amount of data segments in the averaging process.

Time/frequency analysis has been one of the proposed methods of analysing non-stationary data and in that context it also ipmortant that the data is stationary at least up to the chosen data segment length.

Consider a stochastic process *x(t)* with two distinct sets of samples $x_{t_1}$ and $x_{t_1+\tau}$, where the second sample is merely displaced in time by a constant $\tau$. The two sets are characterised by their joint probability density functions,

$$p(x_{t_1}, x_{t_2}, x_{t_3}, ....) \tag{3.51}$$

and

$$p(x_{t_1+\tau}, x_{t_2+\tau}, x_{t_3+\tau}, ....). \tag{3.52}$$

These two pdfs may or may not be identical. If they are equal the process is stationary in the strict sense (or simply, stationary). However, there are a variety of different levels of stationarity and non-stationarity.

### 3.7.1 Non stationarity

When the two pdfs are not equal, as can be seen below

$$p(x_{t_1}, x_{t_2}, x_{t_3}, ....) \neq p(x_{t_1+\tau}, x_{t_2+\tau}, x_{t_3+\tau}, ....) \tag{3.53}$$

the process is said to be non-stationary.

### 3.7.2 Wide sense stationarity

Wide sense stationarity or weak stationarity is the least strict degree of stationarity and it is defined as follows. The first requirement is that the mean of the signal is constant, For discrete time series the mean is calculated as follows

$$\frac{1}{k}\sum_{k} x(k) = m. \tag{3.54}$$

This is rarely a limitation in reality since most of the time the mean of the signal is subtracted before any other analysis.

The second requirement is that the ACF (Auto Correlation Function) only depends on the time lag $\tau$, and not time itself, i.e. it is time independent.

The ACF for discrete time $R_x(k)$ is defined,

$$R_x(k) = \sum_k x(n)x(k-n) = f(k) \tag{3.55}$$

where $f(k)$ is a function of the time lag $k$. Stationarity in the strict sense implies wide sense stationarity, but the reverse is not true, except for those processes that are fully described by the first- and second-order moments, mean and variance, such as a Gaussian process. The higher than second-order moments of odd order for a Gaussian process are zero, and the even-order moments higher than second-order are described by the variance.

### 3.7.3 Nth-order stationarity

A more strict class of signals are the ones where the $n$th-order moment is time independent. The $n$th-order moment is defined as follows

$$E\{x(t)^n\} = \int_{-\infty}^{\infty} x(t)^n dF(x(t)) \tag{3.56}$$

where $F(X)$ is the distribution function for $x(t)$. For a process to be labelled $n$:th order stationary all the moments of the process up to order $n$ have to be time independent. The moments about the mean are defined as follows

$$\mu_n = \int_{-\infty}^{\infty} (x(t) - m)^n dF(x(t)) \tag{3.57}$$

which for a zero mean process reverts back to the previous form, [51] Chapter 15. The moments about the mean are also called central moments, from here on when talking about moments the central moment, $\mu_n$ is assumed, again there being no difference if the signal or process is zero mean.

### 3.7.4 Ergodicity

The definition for ergodicity is that the time average is equal to the ensemble average. Which means that the mean in time from one realisation of a stochastic process equals the mean over a number of realisations of the same process. This is what is known as mean ergodic. There's yet another class of ergodicity, variance ergodic processes. This in turn means that the time and ensemble variances are equal. Ergodicity implies stationarity in the strict sense, but stationarity does not imply ergodicity, i.e.

$$ergodicity => stationarity, \text{ but } stationarity \neq> ergodicity.$$

See Figure 3.6 for an overview of the stationarity properties of a signal, where it can be seen that the stricter forms of stationarity and ergodicity are subsets of all signals.

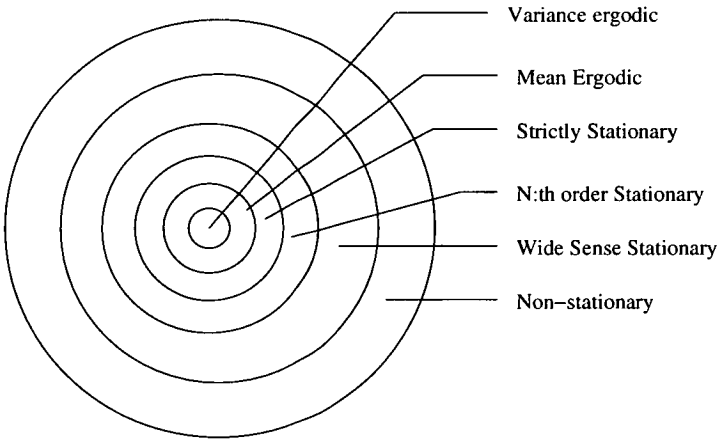


**Figure 3.6:** *Signal categories according to their stationarity properties*

## 3.8 Linearity Symmetry and Gaussianity

The principle of linearity is best described by considering a signal as an input to a system, or a filter for that matter. Then viewing the output of the system and deciding what the characteristics of the system are. The system in this case is the acoustic sound channel that the sonar signals travel through. Together with linearity properties in Table 3.1 the requirement for strict sense stationarity is repeated and also the requirements for causality and stability are mentioned for the sake of completing the table over system properties.

| System property | Input | Output |
|---|---|---|
| Linear | $f(t)$ | $g(t)$ |
| | $f_n(t)$ | $g_n(t)$ |
| | $a_n f_n(t)$ | $a_n g_n(t)$ |
| | $\sum_n a_n f_n(t)$ | $\sum_n a_n g_n(t)$ |
| Strictly Stationary | $f(t-\tau)$ | $g(t-\tau)$ |
| Stable | $\|f(t)\| < M$ | $\|g(t)\| < MI$ |
| Causal | $f(t) = 0$, for $t < t_1$ | $g(t) = 0$, for $t < t_1$ |

**Table 3.1:** *Overview of system properties and their inputs versus outputs.*

where, $n = 1, 2, 3...., a_n, M, I$ are constants. The principle of linearity is also known as the principle of superposition, a linear operator is synonymous with linear system, [60].

Alongside with linearity, symmetry and gaussianity are very common assumptions on signal properties. In [61], and [62] the concepts of symmetry and Gaussianity are discussed. Since the basic assumption is that the systems are linear and Gaussian there are tests done for the opposite case. Symmetry and Gaussianity are sometimes used interchangeably but it is not always a correct description of the signal properties. A gaussian pdf is symmetrical but symmetry does not always imply gaussianity, as it can be seen in Section 3.6 where some pdfs are symmetrical but more pointy than the Gaussian pdf.

## 3.9 Statistical test on sonar data

The tests described and applied in this section can all be described as hypothesis testing. Testing a hypothesis is inferring from a sample whether or not a statement about the population is true, such as sonar data is stationary or ambient noise is Gaussian. In hypothesis testing there are two hypothesis involved, the null hypothesis $H_0$ and the alternative hypothesis $H_1$. For the example above sonar data is stationary the null hypothesis is

$H_0$ : Sonar data is stationary

and the alternative hypothesis is

$$H_1: \quad \text{Sonar data is not stationary}$$

Once the hypothesis is postulated, a test statistic is chosen, and a decision rule is made. The decision rule is to find the possible values that the test statistic can take. Then the test statistic is calculated from a sample from the population, and a decision is made on accepting or rejecting the null hypothesis.

The statistical hypothesis test has a level of significance. That is, there is a certain probability that the test gives a faulty answer. The level of significance, also known as $\alpha$, is the maximum probability of rejecting a true null hypothesis. For a one sided test with the significance level $\alpha = 0.05$ there is a maximum of 5% probability to reject a true null hypothesis. There are two ways to make errors in hypothesis testing. The first being rejecting a true null hypothesis, the second is to accept a false null hypothesis. Rejecting a true null hypothesis is called a *type I error*, and accepting a false null hypothesis a *type II error*.
The critical region of a test statistic is also called a rejection region, and it is defined as the set of all points in the sample space where the null hypothesis is rejected. The set of all points in the sample space where the null hypothesis is accepted is called an acceptance region.
If the null hypothesis is false there is a possibility to accept it with a probability $\beta$. The Power of a test is the ability to reject a false null hypothesis, and it is denoted as $1 - \beta$.

| | Accept $H_0$ | Reject $H_0$ |
|---|---|---|
| $H_0$ is true | Correct decision probability = 1-$\alpha$ | Type 1 error probability = $\alpha$, level of significance |
| $H_0$ is false | Type II error probability = $\beta$ | Correct decision probability = 1 -$\beta$, Power |

Another part of hypothesis testing is the p-value. The p-values is the smallest significance level where the null hypothesis is rejected for a given observation, meaning that, this is the value that the computed test statistic is compared to.

### 3.9.1 Testing for Stationarity

This section discusses testing for different levels of stationarity, starting with stationarity in the strict sense then looking at nth-order stationarity and finally at wide sense stationarity. A number of tests are applied first to synthetic data to point out their strengths and weaknesses, then to real data that has been collected at FOI, over a number of years.

Data from four different experiments is analysed, first an active sonar data set from the BAROC experiment, described in Section 2.5.2, from the transmission loss part of the experiment. This is done to point out the difficulties that arise from moving the objects in the sea and therefore changing the sound channel. And to compare the pings as they are recorded trying to assess the possible benefits of stacking or coherently adding pings. The second set of data analysed is data from a reverberation measurement trial, described in Section 2.5.3. Data segments from between the pings are analysed to examine reverberation and not the pings themselves. This has been done using different lengths of the data, and from different parts of the reverberation tail. The third data set is from the fibre glass boat experiment, described in Section 2.5.6. This includes both background noise data and data when the boat is running. The fourth data set is for looking at back ground noise at different time of the day, described in Section 2.5.5.

### 3.9.2 Kolmogorov-Smirnov

Statistics of the Kolmogorov-Smirnov (K-S) type, Chapter 6 in [63], lend themselves very well to this kind of test. A test for strict sense stationarity, comparing two distribution functions and finding whether or not there is a significant difference between them. If they are approximately equal, according to the test statistic, then the process is stationary.

K-S type of statistics has been applied to sonar data before in [64], and [65] testing for stationarity and normality. However, the data sets are from different areas and there are many factors in underwater acoustics that a make significant difference to the wave propagation so the results from the analysis done here differ in some aspects to the findings in [64], and [65].

The K-S two sample test uses two estimated cumulative distribution functions to produce the test statistic. It is defined as the maximum absolute distance between the two unknown CDFs, and the formal definition follows

$$T_{KS} = sup_x |CDF_1(x) - CDF_2(x)|. \tag{3.58}$$

| Test statistic values for two Gaussian vectors with different variance | | | | | |
|---|---|---|---|---|---|
| $\sigma^2$ | 0.95 | 0.97 | 1.0 | 1.03 | 1.05 |
| Test statistic $T_{KS}$ | 0.0089 | 0.0071 | 0.0034 | 0.0057 | 0.0101 |
| Test statistic $T_{KS}$, analytical | 0.0073 | 0.0043 | 0 | 0.0041 | 0.0068 |

**Table 3.2:** *K-S test statistic values*

The assumptions made on the the underlying data samples are:

1. The samples are random samples
2. The two samples are mutually independent
3. The measurement scale is at least ordinal
4. For the test to be exact the variables are assumed to be continuous.

If the variables are discrete the test valid but becomes conservative, meaning that the risk of making a type I error is less than stated, [63].

Here the two sided K-S test has been applied with a significance level $\alpha$ of 0.01. That is reject the null hypothesis $H_0$ if $T$ exceeds the $1-\alpha$ quantile, Table A19 [63]. The approximation for large samples, $p = \frac{2.30}{\sqrt{n}}$ is used since the number of points, $n$ is larger than 40. The $p$-value is the threshold that the test statistic is compared to.

An example of how the test works is demonstrated by considering two vectors of Gaussian distributed data. Changing the variance in one of them until they are differing in a statistically significant way. In Table 3.2, are the values for the test statistic $T_{KS}$ calculated for two vectors of length $N = 100000$, which makes $p = 2.30/\sqrt{(100000)} = 0.0073$. One of the vectors has a variance equal to one and the variance of the other vector is varied according to Table 3.2 below

The last row of Table 3.2 shows the analytical test statistic value for Gaussian cdfs with the variance according to the top row in the table. This is only included to illustrate the principle of the test. In most cases the analytical pdf or cdfs are not available and the K-S test does not make any assumptions on the underlying distribution function, it operates on the estimated CDFs as previously stated.

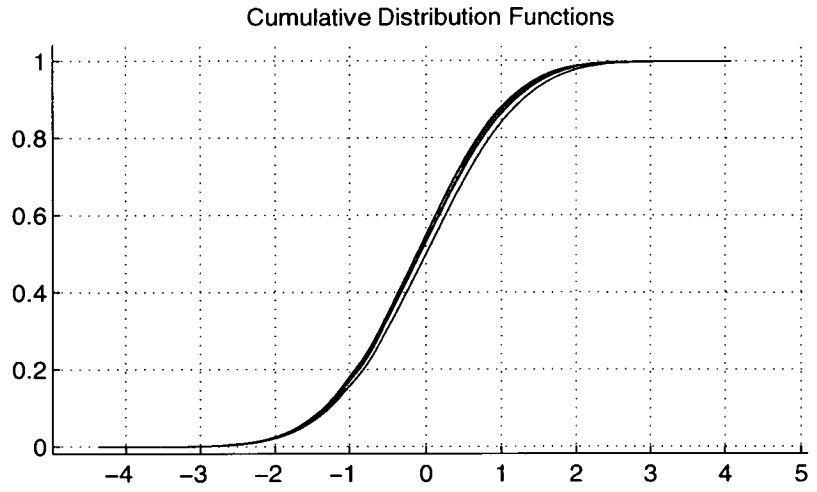


**Figure 3.7:** *The estimated cumulative distribution functions for five vectors of Gaussian variables, with* $\sigma^2 = \{0.95, 0.97, 1.0, 1.03, 1.05\}$

The five CDFs are presented in Figure 3.7. Only the CDFs with a variance of 0.95 and 1.05 are significantly different according to the test statistic $T_{KS}$. It should also be noted that even the vector with a variance equal to one is from another realisation of a Normal process so there is a difference between the vectors and the cumulative distribution functions too. But the difference is not statistically significant since the test statistic $T_{KS} < p - value$

### 3.9.3 Wide Sense Stationarity Quotient

A test for wide sense stationarity as suggested in [66] is the so called wide sense stationarity (WSS) quotient. It consists of two tests, the t-test, that tests for equality of means, $\hat{\mu}$ defined as

$$t = (\hat{\mu_i} - \hat{\mu_j})(\frac{N-1}{\hat{\sigma_i^2} + \hat{\sigma_j^2}})^{\frac{1}{2}} \tag{3.59}$$

where t has a Students t distribution with $\nu = 2N - 1$ degrees of freedom, N is the number of data points in each segment, and $i, j = 1, 2, \cdots, S$. The degrees of freedom is the number of parameters which may be independently varied.

The second part of the WSS test is the F-test, which is a test for equality of variances, $\hat{\sigma^2}$

$$F = \frac{\hat{\sigma_i^2}}{\hat{\sigma_j^2}} \tag{3.60}$$

where $F$ has an F-distribution with $\nu_1 = \nu_2 = N - 1$ degrees of freedom.

Then a matrix $P$ is defined, as follows

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0,S-1} \\ p_{10} & p_{11} & \cdots & p_{1,S-1} \\ \vdots & \vdots & \ddots & \vdots \\ p_{S-1,0} & p_{S-1,1} & \cdots & p_{S-1,S-1} \end{pmatrix}$$

Where $p_{ij}$ are the entries in P and are equal to 1 if and only if both test statistics are between their respective limits.

$$p_{ij} = \begin{cases} 1 & if \mid T \mid < t_{\nu,\frac{\alpha}{2}}, and \\ & F_{\nu_1,\nu_2,\frac{1-\alpha}{2}} \leq F \leq F_{\nu_1,\nu_2,\frac{\alpha}{2}} \\ 0 & otherwise \end{cases} \tag{3.61}$$

Then the WSS quotient can be defined as below,

$$WSSq = \frac{2}{S(S-1)} \sum_{i=1}^{S-1} \sum_{j=i+1}^{S} p_{ij} \tag{3.62}$$

where $S$ is the number of segments, and $p_{ij}$ are the elements in the matrix $P$. It should be noted that the summation is only over the elements in the upper triangle of the matrix since the values on the main diagonal are always ones and the lower triangle is the mirror image of the upper triangle.

The test has been slightly modified from its original form in [67], and [66] due to the fact that the nature of sonar data is different from the Variable Bit Rate (VBR) data examined in [67], and [66]. Sonar data does not consist of a number of frames and is treated as separate segments of any time series. So the summation over matrix elements goes to the number of segments instead of the quotient between the number of frames and number of samples in each segment, suggested in [67], and [66].

This test has been applied to all data sets with $\alpha = 0.05$. For wide sense stationary data the WSS quotient should be at least 0.90, since applying the two tailed test with $\alpha = 0.05$ one

should expect the test to reject as much as 10 % of the data even if it were stationary.

An example of how the WSS quotient works is shown in Figure 3.8 where the WSS quotient has been estimated from a synthetic data set consisting of the sum of three sinusoids with Gaussian (N(0,1)) distributed noise added. Which can be considered as a very simple model of sonar data. The sinusoids corresponding to the sound emitted from a vessels machinery and drive, and the Gaussian noise being the ambient noise. The quotient has been calculated for data lengths of 100 - 100000 samples. This synthetic data has a sampling frequency of 10kHz, thus making the length of the tested data from 0.01s to 10s. In Figure 3.8 it is clear to see that in order to get reliable results the amount of data should be no less than at least 10000 samples (1s). Since the WSS quotient is less then 0.9 for sample size equal to 100, and then drops even lower for a sample size of about 500. After a data length of about 1000 samples the quotient remains higher than the stated limit of WSS = 0.9 for stationary data. It should be noted that, given the frequency content of the data one should use segment lengths where at least one period of the slowest varying part of the signal is included. In the case of getting less than a period of a sinusoid into each segment the mean value will always differ from one segment to the next, and the test will consider the data non-stationary even if it is stationary.

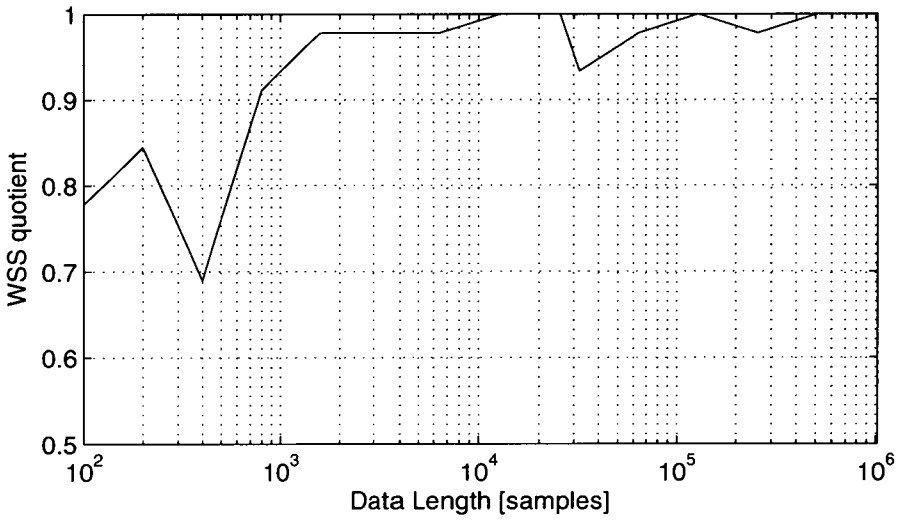


**Figure 3.8:** *The WSS quotient estimated for different sample sizes, from synthetic data.*

### 3.9.4 Symmetry and Linearity tests

A test for symmetry based on an estimate of the pdf can be done using the histogram. The test statistic is defined below as

$$T_{sym} = \frac{1}{N} \sum_{i=1}^{M/2} (H(i) - H(-i)) \tag{3.63}$$

where $H$ is the estimated histogram for a block of data with a length $N$ and $M$ is the number of bins in the histogram. The sum over the histogram divided with the number of data points $N$ is equal to one. The sum of the differences around the mean value of the normalised histogram should be zero for a symmetric data set. However, since the histogram is estimated from a limited data set some deviation from zero should be allowed. The data is considered symmetric if $|T_{sym}| \leq 0.05$, which means that the histogram is less than 5 percent skewed to either side.

Another symmetry test is proposed in [61], and is applied to underwater acoustic data in Chapter 1,[68]. and to seismic data in [62]. It is based on the normalised bispectrum, Equation (3.43), [56],[57], and consists of the sum over all frequency pairs inside the inner triangle (IT), see Figure 3.9.

$$S = 2 \sum_{f_1, f_2 \in IT} b^2(f_1, f_2) \tag{3.64}$$

The test statistic $S$ is $\chi^2$ distributed with dof equal to twice the number of grid-points in the IT. Testing for symmetry is now essentially the same as testing $S$ for consistency with a $\chi^2$-distribution. If the underlying data is symmetric in it's distribution, then $S \equiv 0$ under the null hypothesis. $H_0$ is rejected if $S$ differs significantly from 0.

A test for linearity is also proposed in [61], and it also based on the normalised bispectrum. The test statistic is $S_L$ is equal to the ordered values in $2|b(f_1, f_2)|^2$. Then the interquartile range $R$ of $S_L$ is estimated, $R$ is the difference between the first and third quartile, e.g. 50 per cent of the values around the mean. The hypothesis of linearity is rejected if $R$ is different from a theoretical value of the interquartile range of a $\chi^2(\lambda)$-distribution, where $\lambda$ is the estimated mean from the bicoherence $b^2(f_1, f_2), f_1, f_2 \in IT$. And where the bicoherence is a normalised bispectrum, much like the ordinary second order coherence function is a normalised power spectrum.

### 3.9.5 Stationarity testing based on the bispectrum

Following the reasoning from the previous section on the symmetry and linearity tests based on the bispectrum, tests for stationarity can also be defined on the same basis. Using the normailsed bispectrum, see [62] a test statistic is defined below as

$$z = \frac{\sum_{f_1, f_2 \in OT} |s(f_1, f_2)|^2 - 2}{2\sqrt{N^2/48}} \tag{3.65}$$

where $N^2/48$ is the number of estimated bifrequencies in the Outer Triangle (OT), see Figure 3.9, and $s(f_1, f_2)$ is defined in Equation (3.42). The total number of bifrequencies in the bispectral plane is $N^2$. This leads to the inner triangle having $N^2/16$ bifrequencies, since IT has the width $N/2$, and height $N/4$. Then the number of bifrequencies in both IT and OT is $N^2/12$, and from that it follows that OT has $N^2(\frac{1}{12} - \frac{1}{16}) = N^2/48$ as stated above.

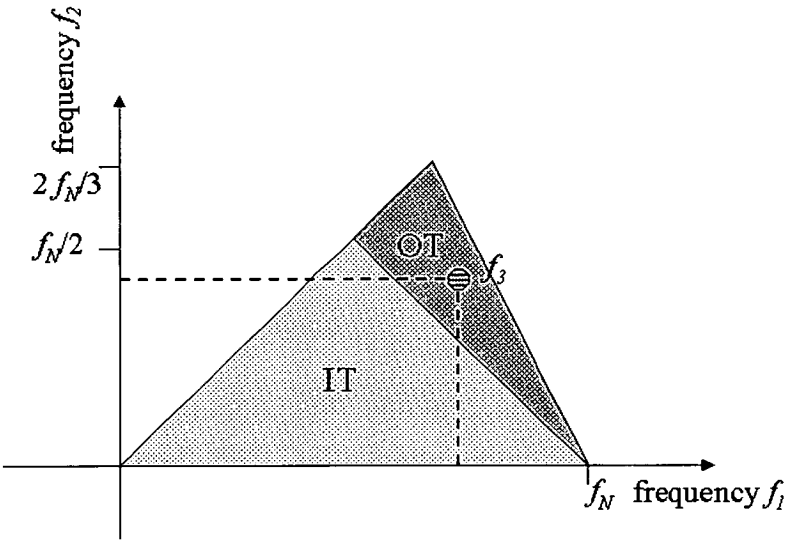


**Figure 3.9:** *The non-redundant principal domain of the bispectrum. In this thesis IT is used for symmetry and linearity tests and OT is used for a stationarity test. $f_N$ is the Nyquist frequency.*

### 3.9.6 Testing ping to ping stationarity

In active sonar it is of great interest to find out how similar consecutive pings are. Common techniques in sonar signal processing include stacking pings, which means averaging over consecutive pings. This is done to enhance the signal to noise ratio. Stacking pings works well under the assumption that the actual pings are stationary to some degree and that it is possible to average out the noise background over a number of pings. The idea is very similar to estimating power spectra by segmenting data and averaging over a spectrogram, and thereby improving the SNR of the spectral estimate, see Chapter 12, [69]. This technique can improve the probability of detection considerably, however unless the mentioned assumptions hold true there is very little to gain by performing the stacking, it can in fact make the situation worse if the pings are too different from each other and lead to the loss of information rather than a gain in the SNR. Under the assumption that the signals are coherent and the noise is non-coherent and white, the gain in SNR is proportional to the square root of the number of stacked pings, $(\sqrt{N})$.

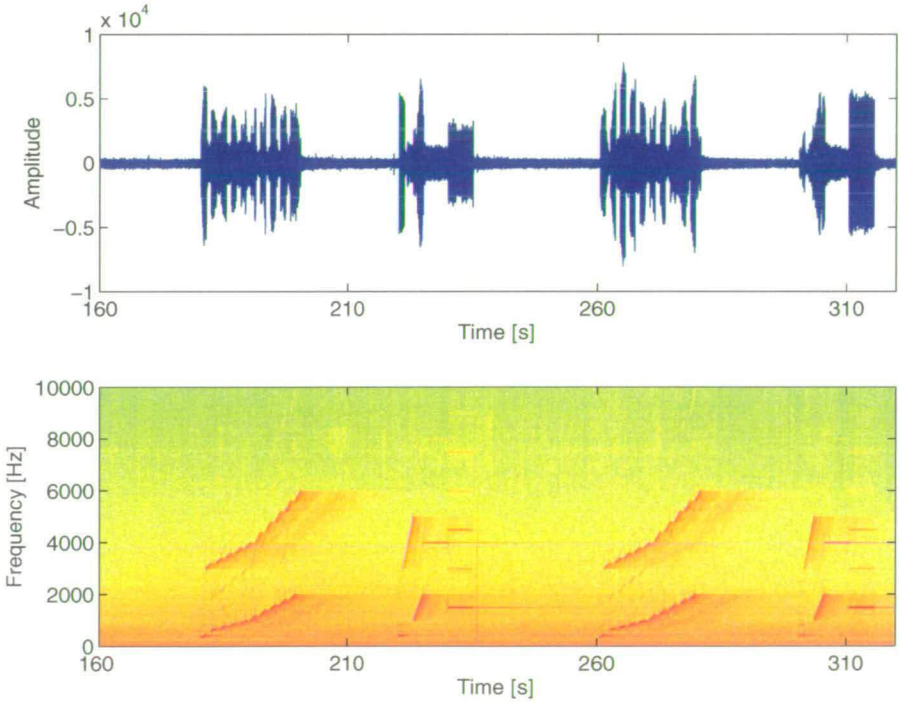


**Figure 3.10:** *Time series and spectrogram over the first four pings in the data file. Note that it is the 20 seconds long ping consisting of chirps that are analysed*

The first data set analysed is from the BAROC trial. The pings (here including the entire pulse train consisting of 20 chirps) from the TL measurement are compared to one another.

The comparison is made over the first 100 pings from the data file in question. Figure 3.10 illustrates the first four pings. The first ping is the 20 seconds pulse train consisting of 20 chirps, the second ping is a mixture of cw-pulses and chirps. These pings are repeated every 80 seconds. That means that data from pings within a time frame of 3.5 hours are compared to determine if the received pings remain the same and it is feasible to call them invariant, or if indeed the underwater-channel affects them enough to make the received pings into a non-stationary process. It is only the pulse train of 20 chirps that are analysed in this section.

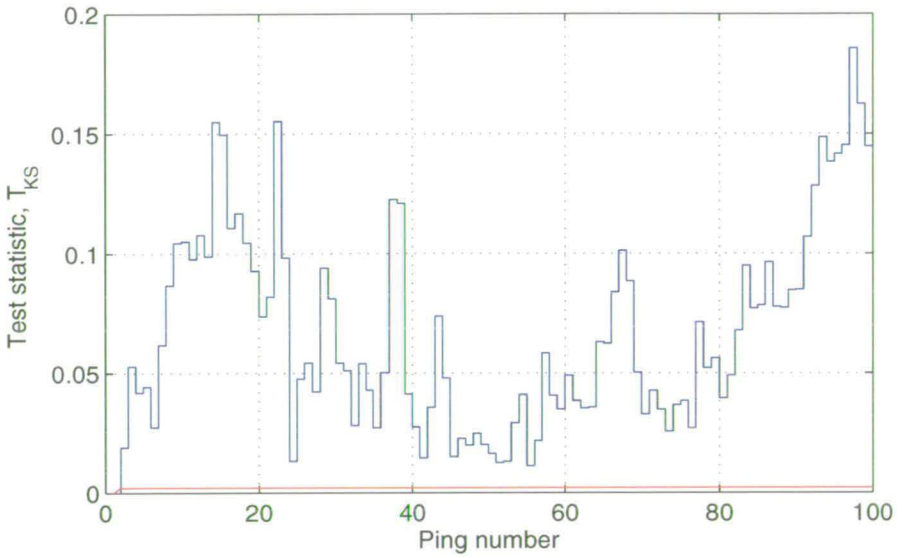


**Figure 3.11:** *The K-S test applied to BAROC data. Ping one is compared to all pings up to ping number 100*

The first test applied is the K-S two sample test. That is to determine if the pings are strictly stationary or not, see Figure 3.11, and Figure 3.12. Where ping number one is compared to the rest of the pings and then ping number 50 is compared to the rest up to ping number 100. Clearly the test statistic is above the $p$-value at all times so the data is not stationary. Reminding, that the $p$-value is the threshold which the test statistic is compared to. This is a result that is expected since there is 80 seconds between consecutive pings, and the vessel carrying the receiver is moving away from the transmitter. Thereby essentially creating a new channel for the sound to travel in. The bottom topography changes since the boat is moving, the SSP also changes from one point to another.

The K-S test is applied to all of the pings comparing all of them to one another and are at all times different enough for the K-S test to distinguish between them, see Figure 3.13 where the test statistic is plotted. The only place where the test is accepting the null hypothesis is along

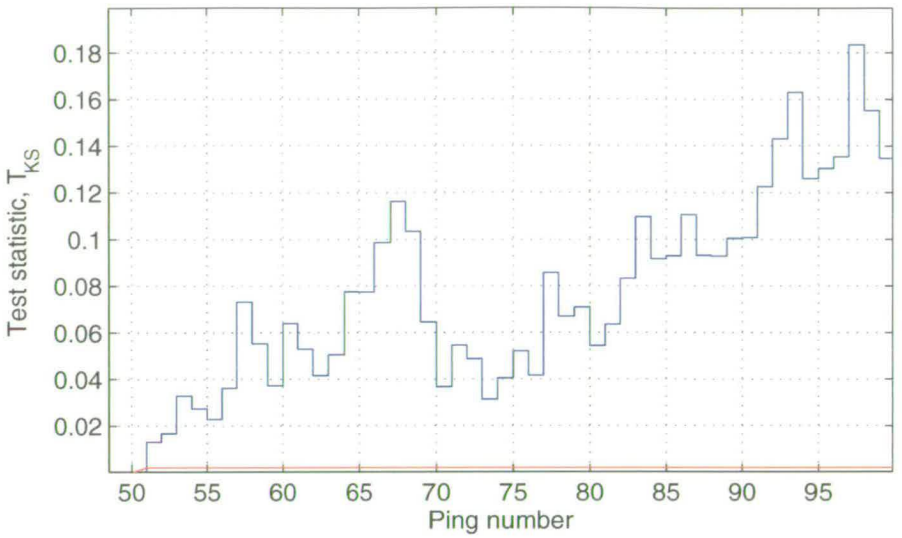


**Figure 3.12:** *The K-S test applied to BAROC data. Ping 50 compared to all pings from 51 up to ping number 100*

the diagonal, which is comparing a ping to itself. Otherwise the test statistic is larger than the $p$-value at all times.

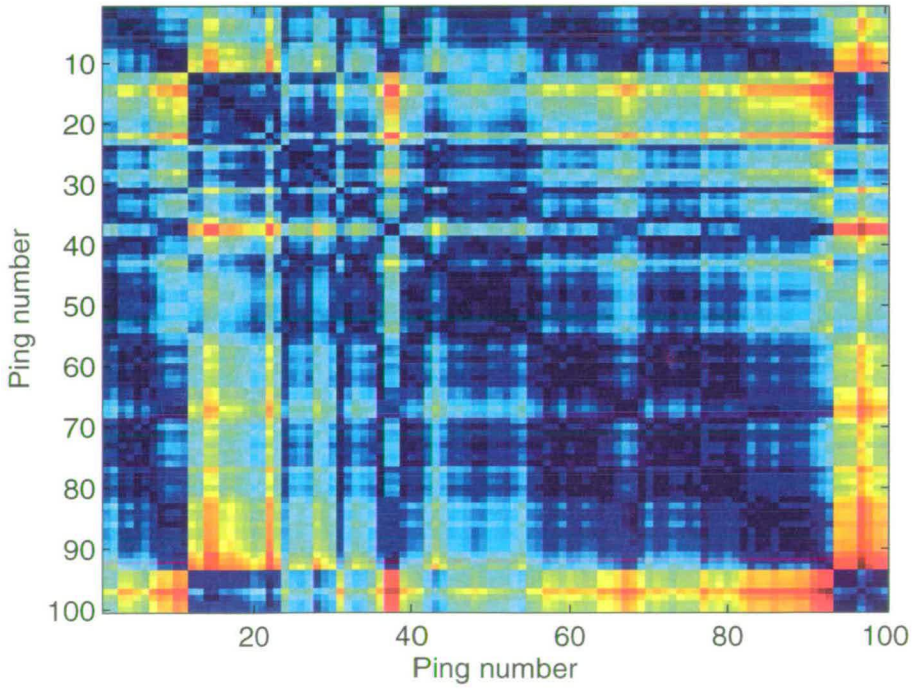


**Figure 3.13:** *The K-S test applied to BAROC data, including all pings up to number 100.*

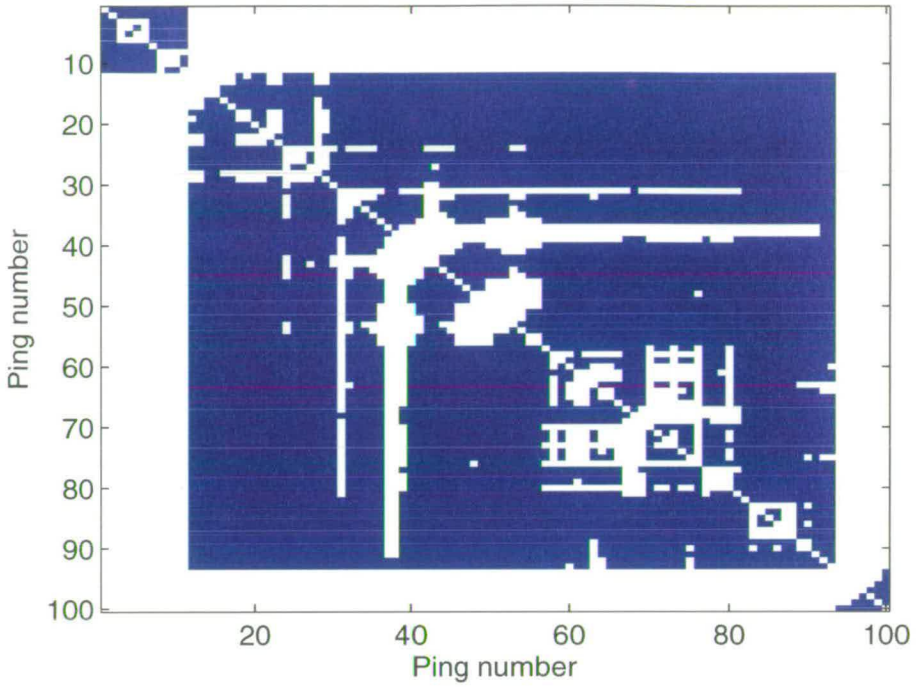


**Figure 3.14:** *The WSS quotient applied to BAROC data, including all pings up to number 100.*

The second test applied to the BAROC data is the WSS quotient, see Figure 3.14 where all pings are compared as in the KS test case. The WSS test has been done in two different ways, first alongside with the KS test, testing one ping against another, thus getting $S_{ij} = 1$ if all requirements are fulfilled in the test, and $S_{ij} = 0$ otherwise. Then creating a matrix $T_{WSS}$ with the $S_{ij}$ as elements. It can be seen in Figure 3.14 that the null hypothesis is not accepted at all times. The dark areas denotes the ones and the white areas are zeros. By summing the elements in $T_{WSS}$ the WSS quotient is obtained, and it has the value 0.58. The same value is obtained by reading all the data in and running the test by directly getting the WSS quotient. From this it can be concluded that the pings are not even wide sense stationary. However to some degree, pings following each other are similar enough, and the middle area in Figure 3.14 actually reaches wide sense stationarity. It should also be noted that the dark area correspond fairly well with the area of lower values of $T_{KS}$ in the middle of Figure 3.13.

This opens up another line of questioning, namely, how many pings in a row are wide sense stationary? As it can be seen in Figure 3.14 it depends on what the starting point is. Starting from the diagonal in Figure 3.14, one can trace a black line for up to 70 pings. Which means that the 70 following pings are wide sense stationary. This is however not the case in general,

| Ping number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr of WSS pings | 10 | 9 | 0 | 1 | 0 | 5 | 4 | 0 | 0 | 1 |
| Ping number | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nr of WSS pings | 0 | 0 | 0 | 0 | 78 | 76 | 75 | 0 | 70 | 69 |
| Ping number | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| Nr of WSS pings | 68 | 70 | 69 | 0 | 0 | 65 | 0 | 65 | 64 | 61 |
| Ping number | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| Nr of WSS pings | 0 | 0 | 8 | 7 | 6 | 4 | 3 | 0 | 0 | 0 |
| Ping number | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 |
| Nr of WSS pings | 0 | 0 | 50 | 48 | 7 | 3 | 2 | 0 | 0 | 0 |
| Ping number | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 |
| Nr of WSS pings | 0 | 0 | 40 | 39 | 0 | 37 | 12 | 0 | 10 | 2 |
| Ping number | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 | 69 | 70 |
| Nr of WSS pings | 0 | 0 | 6 | 0 | 4 | 3 | 1 | 0 | 0 | 2 |
| Ping number | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 | 80 |
| Nr of WSS pings | 1 | 0 | 2 | 1 | 0 | 17 | 0 | 0 | 0 | 13 |
| Ping number | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 |
| Nr of WSS pings | 12 | 11 | 0 | 2 | 1 | 0 | 0 | 5 | 0 | 3 |
| Ping number | 91 | 92 | 93 | 94 | 95 | 96 | 97 | 98 | 99 | 100 |
| Nr of WSS pings | 2 | 1 | 0 | 0 | 0 | 0 | 3 | 2 | 1 | 0 |

**Table 3.3:** *The number of consecutive wide sense stationary pings from the BAROC trial*

see Table 3.2 for the number of consecutively wide sense stationary pings. Again it should be noted that the test is carried out for the 100 first pings from the data file so the number of consecutively wide sense stationary pings does not mean anything for ping number 100, since there has not been a comparison done outside the 100 first pings. However, is should also be noted that there is not a single case where the number of consecutive WSS pings reach the limit of 100.

### 3.9.7 Testing Reverberation for stationarity

The reasons for testing the stationarity of reverberation are similar as for the actual ping, with the addition of also viewing reverberation as noise as well as the part of the ping that actually has the desired echoes in it. It is also of importance to find out what the actual statistics of the noise background in active sonar is.

The experimental setup for the Utö trial is an almost monostatic one, the tranmitter and receiver are mounted to different sides of a boat, and sometimes at different depths, but this difference in location is rather small compared to the distances the returning echoes are from. The monostatic setup of this sea trial makes it an ideal data set for a reverberation analysis. An example of how the data looks can be seen in Figure 3.15, where the first four pings from File 1 is displayed. The data has been beamformed with the main lobe perpendicular to the receiver array. That simply means that data from the 32 channels are added to each other. The transmitted pulse passes the receiving array and the signal is heavily distorted (clipped), making a statistical analysis of the transmitted ping difficult. However, the reverberation tail is free from this sort of distortion and it is suitable for further analysis, since the distortion is introduced from the receiving antenna elements and amplifiers and has nothing to do with the actual wave propagation. File 1 and File 3 have equal depths on both transmitter and receiver, 12m whereas for File 2, the transmitter is at 12m and the receiver is at 25m depth.

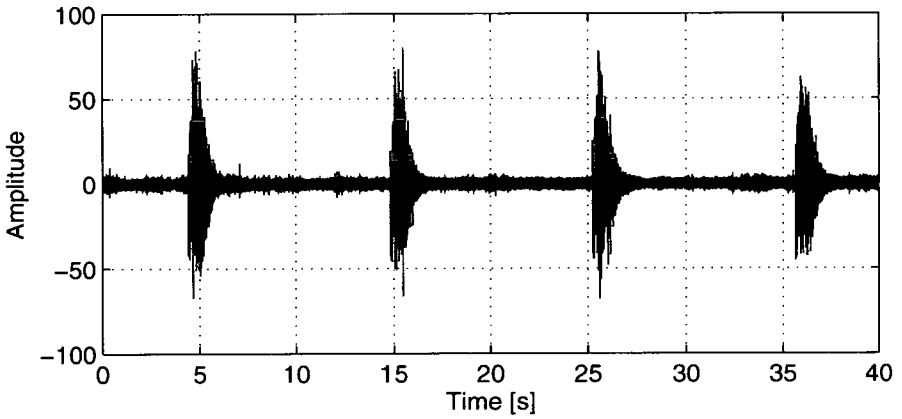


**Figure 3.15:** *The four first pings from File 1.*

This data is analysed with K-S test for strict sense stationarity and the result of that is displayed in Figure 3.16. The test is done comparing all pings to each other but only the results from comparing ping number one to all other pings in each file is presented here. This is partly due to the test statistic $T_{KS}$ exceeding the p-value and the null hypothesis is rejected, so the reverberation data is not stationary in a strict sense, and also due to the fact that it captures the behaviour of the rest of the test. The full test results can be seen in appendix C.

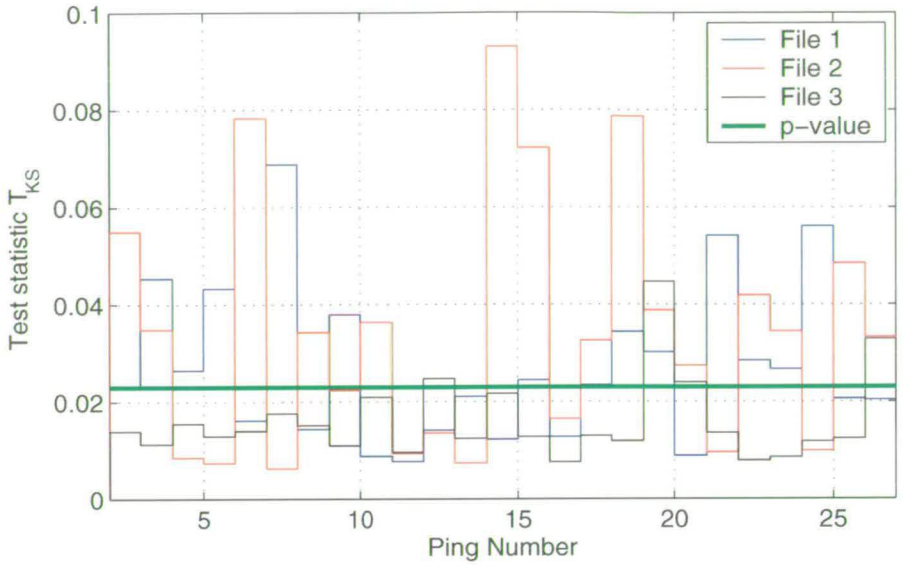


**Figure 3.16:** *The K-S test applied to Utö data. Ping one is compared to all pings up to ping number 27, for all three data files.*

The WSS quotient is calculated for the reverberation data, at about one and three seconds after the transmission of a ping. It is done for a data length of 2 seconds. For the first part of the test the WSS quotient reaches values of 0.49 File 1, 0.61 for File 2, and 0.47 for File 3. So the reverberation as such is not wide sense stationary either. But it is still interesting to find out if consecutive pings, or rather the reverberation following consecutive pings, are wide sense stationary to some degree. The results from the ping to ping wide sense stationarity analysis is presented in Tables, 3.4, to 3.9.Also see appendix C. Starting with the three that hold test results from one second after the ping transmission.

| Ping number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr of WSS pings | 0 | 4 | 0 | 1 | 0 | 0 | 20 | 0 | 9 | 0 |
| Ping number | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nr of WSS pings | 1 | 2 | 1 | 0 | 3 | 0 | 1 | 0 | 1 | 0 |
| Ping number | 21 | 22 | 23 | 24 | 25 | 26 | | | | |
| Nr of WSS pings | 6 | 1 | 0 | 3 | 0 | 1 | | | | |

**Table 3.4:** *The number of consecutive wide sense stationary pings from the Utö trial, file 1. Analysis starting one second after the ping was transmitted.*

The WSS quotient is also calculated at the very tail end of the reverberation and it reaches values of 0.45 File 1, 0.50 for File 2, and 0.41 for File 3. This is at a point when the reverberation level

| Ping number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr of WSS pings | 0 | 25 | 2 | 1 | 0 | 7 | 0 | 0 | 4 | 3 |
| Ping number | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nr of WSS pings | 0 | 1 | 0 | 7 | 2 | 0 | 0 | 3 | 2 | 1 |
| Ping number | 21 | 22 | 23 | 24 | 25 | 26 | | | | |
| Nr of WSS pings | 0 | 5 | 1 | 0 | 2 | 1 | | | | |

**Table 3.5:** *The number of consecutive wide sense stationary pings from the Utö trial, file 2. Analysis starting one second after the ping was transmitted.*

| Ping number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr of WSS pings | 0 | 16 | 2 | 0 | 1 | 2 | 3 | 2 | 1 | 0 |
| Ping number | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nr of WSS pings | 1 | 0 | 5 | 4 | 2 | 1 | 0 | 0 | 8 | 7 |
| Ping number | 21 | 22 | 23 | 24 | 25 | 26 | | | | |
| Nr of WSS pings | 6 | 3 | 2 | 1 | 0 | 1 | | | | |

**Table 3.6:** *The number of consecutive wide sense stationary pings from the Utö trial, file 3. Analysis starting one second after the ping was transmitted.*

is getting close to the ambient noise level. It is worth noting that the WSS quotient is lower for this part of the ping. The starting point is for this part of the analysis is at about 8 seconds in File 1, and at corresponding times after the ping in File 2 and 3.

| Ping number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr of WSS pings | 7 | 2 | 0 | 0 | 1 | 0 | 2 | 1 | 0 | 17 |
| Ping number | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nr of WSS pings | 0 | 1 | 0 | 10 | 9 | 2 | 1 | 0 | 1 | 0 |
| Ping number | 21 | 22 | 23 | 24 | 25 | 26 | | | | |
| Nr of WSS pings | 3 | 1 | 0 | 0 | 1 | 0 | | | | |

**Table 3.7:** *The number of consecutive wide sense stationary pings from the Utö trial, file 1. Analysis starting three seconds after the ping was transmitted.*

comparing the Tables 3.7 - 3.9 with 3.4 - 3.6 One finds that the number of consecutive pings that are wide sense stationary is lower. It can be said about both that the number of pings that are stationary after any given ping varies highly. It should be noted that toward the end of the Tables 3.4 to 3.9 there is not enough data to say anything about the number of stationary pings due to the number of available pings. However, it seems that in most of the cases it does not limit the number of pings that are stationary.

| Ping number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr of WSS pings | 0 | 25 | 2 | 1 | 0 | 5 | 4 | 2 | 1 | 0 |
| Ping number | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nr of WSS pings | 0 | 1 | 0 | 4 | 3 | 1 | 0 | 0 | 8 | 5 |
| Ping number | 21 | 22 | 23 | 24 | 25 | 26 | | | | |
| Nr of WSS pings | 4 | 3 | 2 | 0 | 0 | 1 | | | | |

**Table 3.8:** *The number of consecutive wide sense stationary pings from the Utö trial, file 2. Analysis starting three seconds after the ping was transmitted.*

| Ping number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Nr of WSS pings | 1 | 22 | 21 | 0 | 4 | 2 | 1 | 0 | 4 | 3 |
| Ping number | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| Nr of WSS pings | 0 | 1 | 0 | 10 | 2 | 1 | 0 | 6 | 0 | 4 |
| Ping number | 21 | 22 | 23 | 24 | 25 | 26 | | | | |
| Nr of WSS pings | 0 | 2 | 1 | 0 | 2 | 1 | | | | |

**Table 3.9:** *The number of consecutive wide sense stationary pings from the Utö trial, file 3. Analysis starting three seconds after the ping was transmitted*

### 3.9.8 Testing Ambient noise for stationarity

Tests on stationarity and Gaussianity have been done on ambient noise data but mostly on data from deep sea locations, [68] Chapter 1. In this section three different sets of data are analysed, one from relatively deep waters from the Baltic Sea and two from a shallow bay outside Stockholm. The deepest part of the Baltic Sea is no more than about 400 meters so all sonar data from the Baltic Sea is to be compared to littoral areas elsewhere.

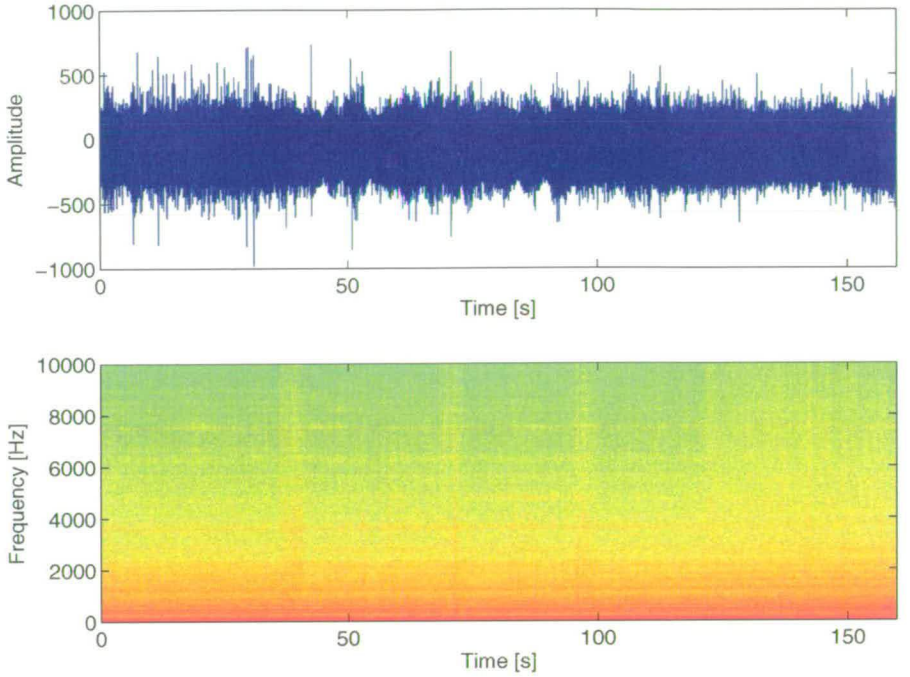


**Figure 3.17:** *Time series and spectrogram over the first 160 seconds of data from the BAROC data file.*

The ambient noise data from the BAROC trial is a part of the data set where there are no active transmission. The noise from this recording is dominated by shipping, since there was at least three ships in the vicinity at all times during the experiment, see Figure 3.17 for the time series and the spectrogram over the 160 first seconds of the BAROC data set. Although there are no very clear tonal lines in the spectrogram it is shipping that is the largest contributing factor in this data set. Both the KS and the WSS test have been applied to this data. The data was segmented into 100 segments of equal length, 1.6 seconds (with a sampling frequency of 59900Hz that makes 95840 samples). The analysis has been done on the data as it is and band pass filtered data to find if there possible is a dependency on frequency. It should be noted that the number of data points are the same even if the bandwidth of the data that is processed is different.

The KS test rejects the null hypothesis since there are times when the test statistic $T_{KS}$ exceeds the chosen quantile for all test cases. That is all the different frequency bands as well as the unfiltered data. However, it is interesting to follow how the sum over all test cases vary according

| Centre Frequency | 100 | 125 | 160 | 200 | 250 | 315 |
|---|---|---|---|---|---|---|
| Bandwidth | 25 | 28 | 35 | 49 | 58 | 70 |
| Centre Frequency | 400 | 500 | 630 | 800 | 1000 | 1250 |
| Bandwidth | 87 | 125 | 142 | 173 | 250 | 280 |
| Centre Frequency | 1600 | 2000 | 2500 | 3150 | 4000 | 5000 |
| Bandwidth | 350 | 490 | 575 | 705 | 870 | 1200 |
| Centre Frequency | 6300 | 8000 | 10000 | 12500 | 16000 | 20000 |
| Bandwidth | 1420 | 1730 | 2500 | 2800 | 3500 | 4900 |

**Table 3.10:** *Centre frequencies and bandwidths, in Hz, for the band pass filter applied to the BAROC ambient noise data prior to the statistical testing. Note that the centre frequency / bandwidth quotient is kept approximately constant through the filters.*

to

$$\frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} H_{ij} \tag{3.66}$$

where $H_{ij}$ are either 0 or 1 corresponding to $H_0$ and $H_1$, and $n$ is the side of the square matrix $H$. These results are presented in Figure 3.18.

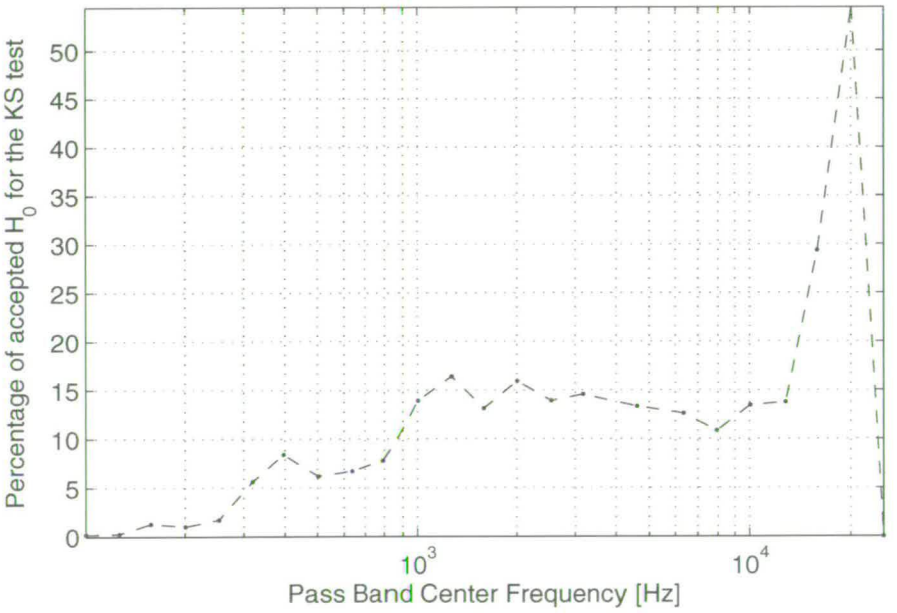


**Figure 3.18:** *The quotient of KS test outcomes that accepts the null hypothesis for BAROC data without any transmission. The first data point is not filtered, the others following are bandpass filtered according to Table 3.10, with exception that the last data point is from unfiltered data*

It is interesting to note that the factors that affect stationarity are lower in the spectrum and the percentage of accepted null hypothesis go up with frequency. This is most likely due to the low frequency sounds being emitted from the nearby ships and there being little acoustic energy higher up in frequency.

The results from WSS quotient for the BAROC ambient noise data, segmented as described above and filtered according to Table 3.10 are displayed in Figure 3.19. The last point in the result vector comes from data that has not been band pass filtered other than during the collection of the data. Contrary to the results using the KS test the low frequency parts seem to have a more constant mean and variance and the higher up in frequency one goes the WSS quotient drops lower. This points in the direction that there is information beyond second order statistics in the sonar data.

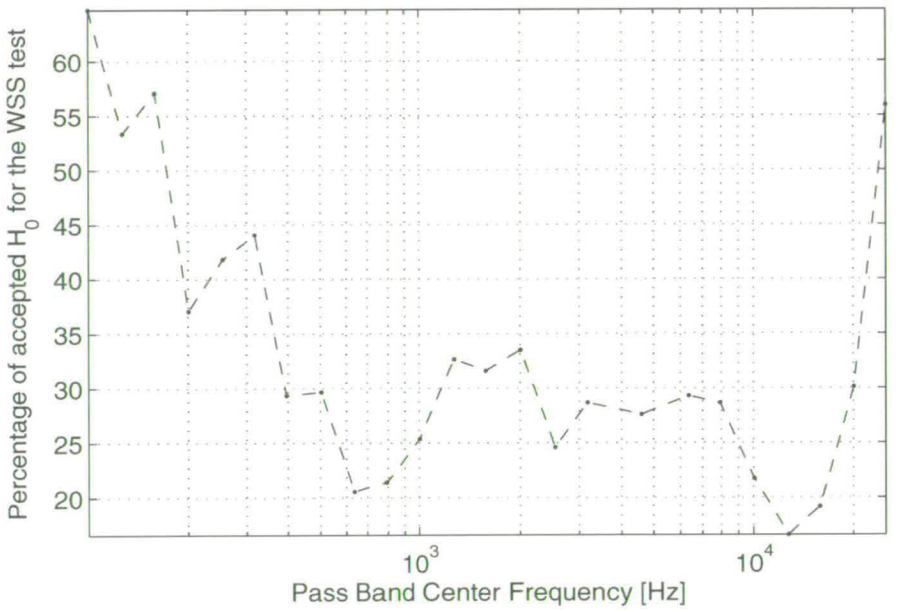


**Figure 3.19:** *The WSS quotient for BAROC data without any transmission. The last data point is not filtered, the others following are bandpass filtered according to Table 3.10.*

The influence of the data length to estimate the test statistic WSS is examined and the result can be seen in Figure 3.20, where the window length goes from 600 samples to 4.8 million samples (0.01s-80s). The highest value for the WSS quotient is reached at 2.3 million samples, and that equals 40 seconds of data. However, it is not feasible to use that many samples in each window, since the amount of data available is limited, and another limiting factor is the time it takes to make the estimation. A compromise that seems reasonable is to use data lengths of the order of

1.5s (100k samples).

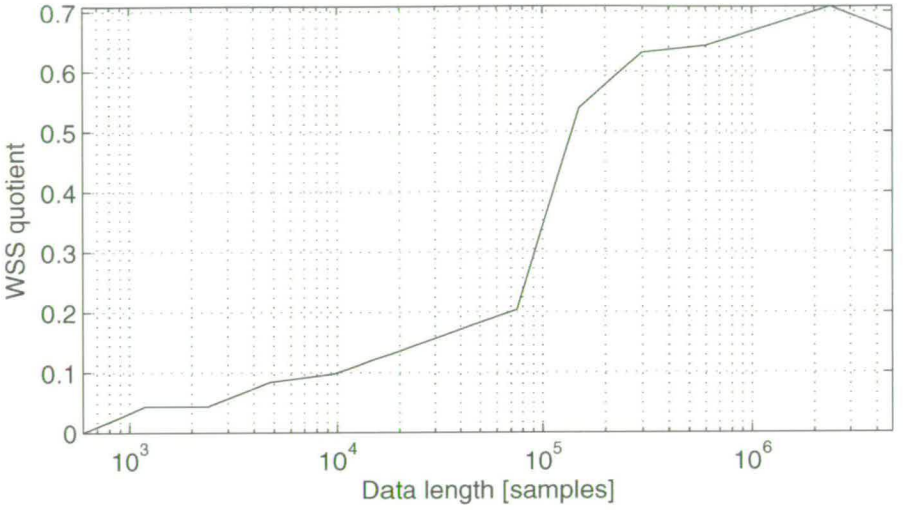


**Figure 3.20:** *The WSS quotient estimated for different lengths of data for the BAROC sea trial. Starting with length of 600 samples going up to 4.8 million samples. That corresponds to, 10ms to about 80s of data.*

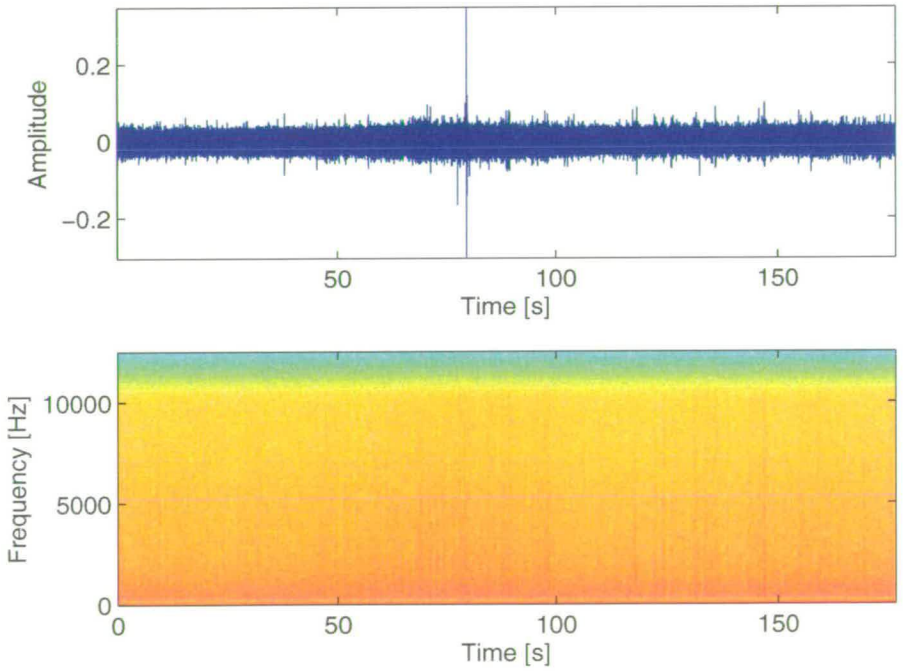


**Figure 3.21:** *The time series and the spectrogram of noise data from before the fibreglass boat trial.*

The second part of this ambient noise stationarity analysis is done on data collected during the fibreglass boat trial. During that trial ambient noise was collected both before and after the

actual main event, ie when the boat was in use and the straight line tracks were run, see Figure 3.21 for the time series and the spectrogram of the data from before the trial. The analysis results for the data from before the boat was in use is presented in Figures 3.23, and 3.25. First the KS test was applied to all of the ambient noise data comparing the data at the beginning and the end of the sea trial. And the null hypothesis is rejected since the test statistic ($T_{KS}$ =0.085) has a larger value than the p-value (0.0012). The two cumulative distribution functions are shown in Figure 3.22 and it is easy to see that there is a difference between them. The difference that is accepted by the KS test is however barely discernible to the eye when testing data sets of this length, 180s times 25kHz sampling frequency. So it is clear that the ambient noise from before and after the trial are statistically different. This is probably mainly due to differences in shipping, which basically is due to the different time of the day.

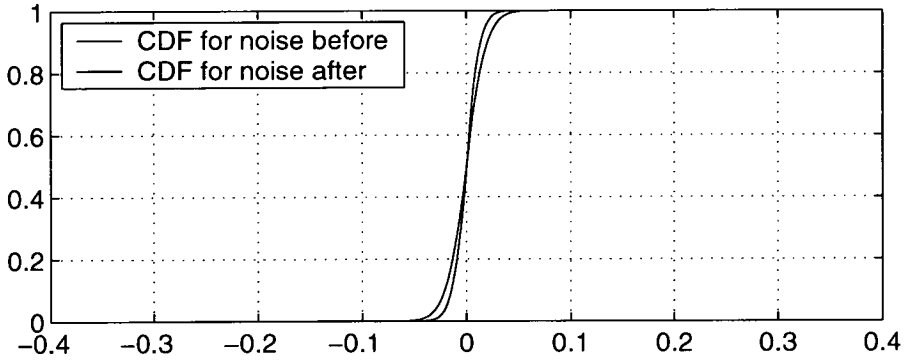


**Figure 3.22:** *The KS test applied to fibreglass boat trial ambient noise data, using all data points from the ambient noise recording before the boat trial and after the trial.*

Again it is found that the percentage of accepted null hypothesis go up with frequency, see Figure 3.23. All though the experimental setup differs quite significantly from the BAROC trial. There is no nearby shipping to speak of, and the only surface vessel in the area is not in use.

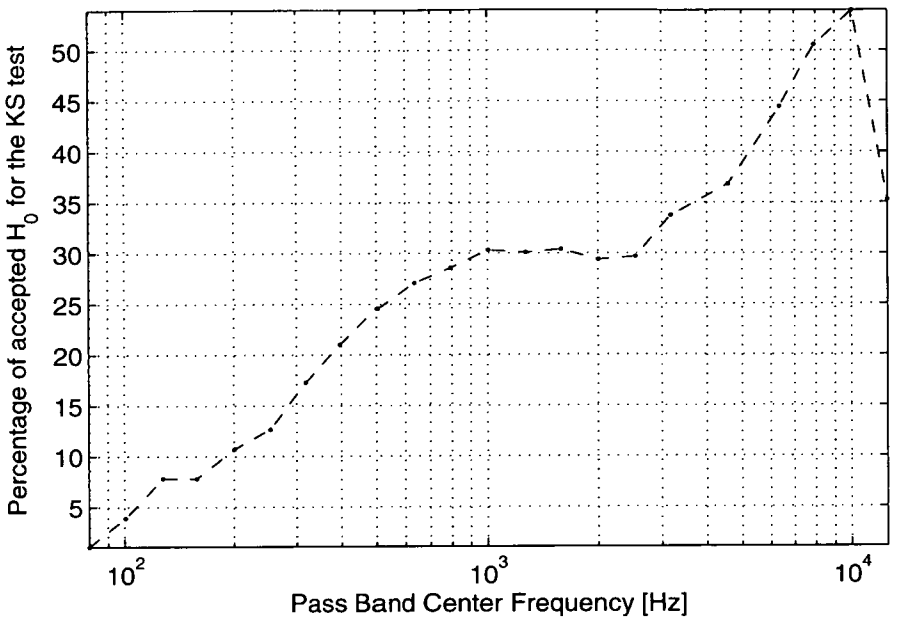


**Figure 3.23:** *The quotient of KS test outcomes that accepts the null hypothesis for the fibreglass boat trial ambient noise data. The first data point is not filtered, the following are bandpass filtered according to Table 3.10, with two exceptions. The highest centre frequency is 10kHz, and lower centre frequency is added, 71Hz. This is due to the lower sampling frequency 25kHz compared to 59,9kHz*

To find out how long the ambient noise stays stationary the KS test is also applied to different lengths of data from the fibreglass boat trial. The results from that analysis can be seen in Figure 3.24. The KS test is done to data lengths from 10ms to roughly 88 seconds (270 samples equals roughly 10ms of data, and 88s is roughly 2.1 million samples). There were four hydrophones in use during the trial and the KS test was applied to all of them. Following the curves in Figure 3.24 one finds that up to about 10k samples the test statistic $T_{KS}$ remains under the p-value on average. So the length of time that the ambient noise from that particular part of the Stockholm archipelago, under the circumstances of the trial, remain stationary is about 0.4 seconds.

The results from the WSS quotient test is quite different from the results of the BAROC data analysis, see Figure 3.25. Again this could be attributed to the different experimental setup.

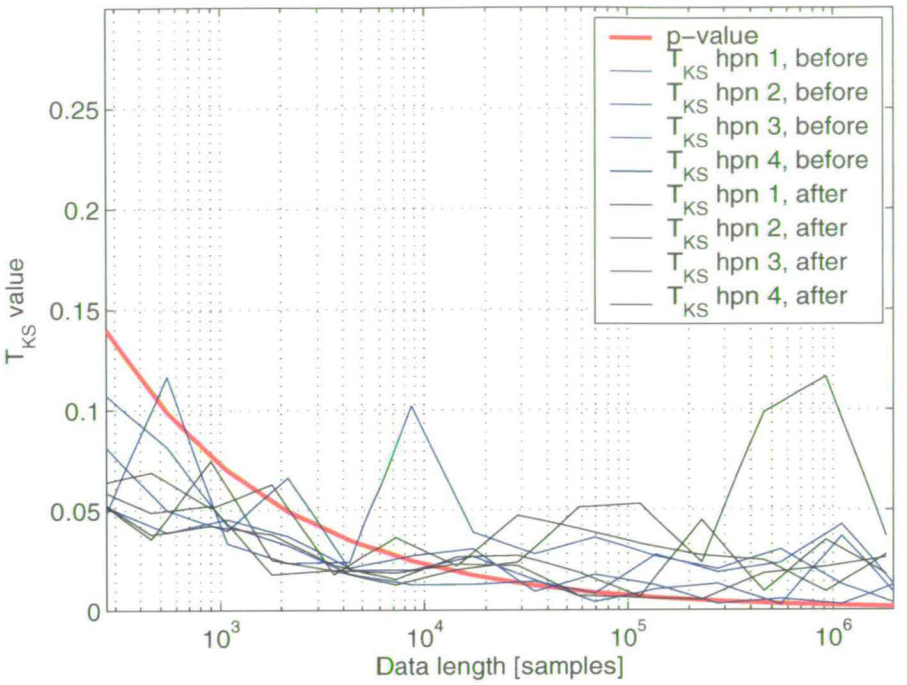


**Figure 3.24:** *The KS test applied to fibreglass boat trial ambient noise data, for all four available hydrophones, both before and after the boat trial. Using different data length to find how long the data stays stationary.*

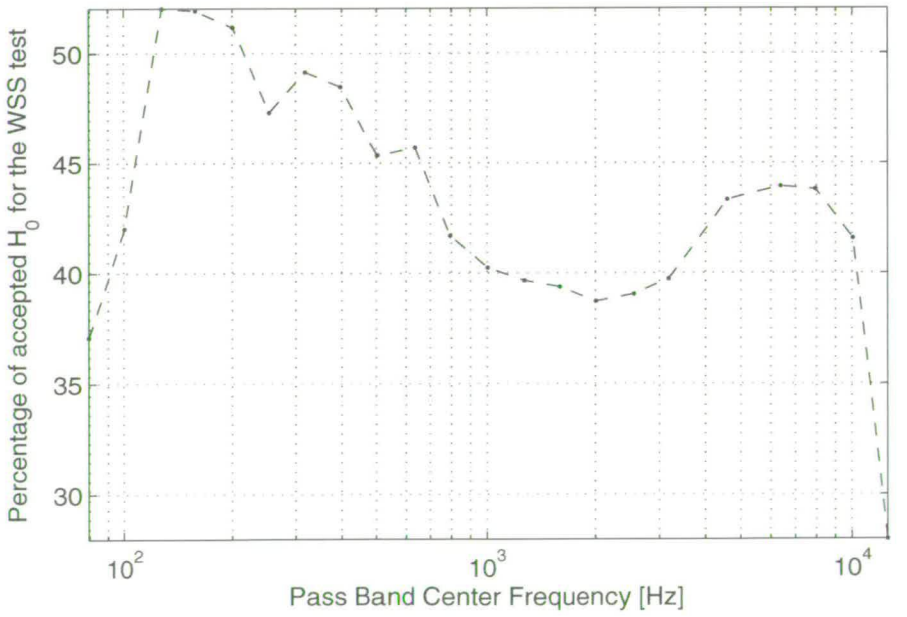


**Figure 3.25:** *The quotient of WSS quotient test outcomes that accepts the null hypothesis for the fibreglass boat trial ambient noise data. The first data point is not filtered, the others following are bandpass filtered according to Table 3.10, with two exceptions. The highest centre frequency is 10kHz, and lower centre frequency is added, 71Hz. This is due to the lower sampling frequency 25kHz compared to 59,9kHz*

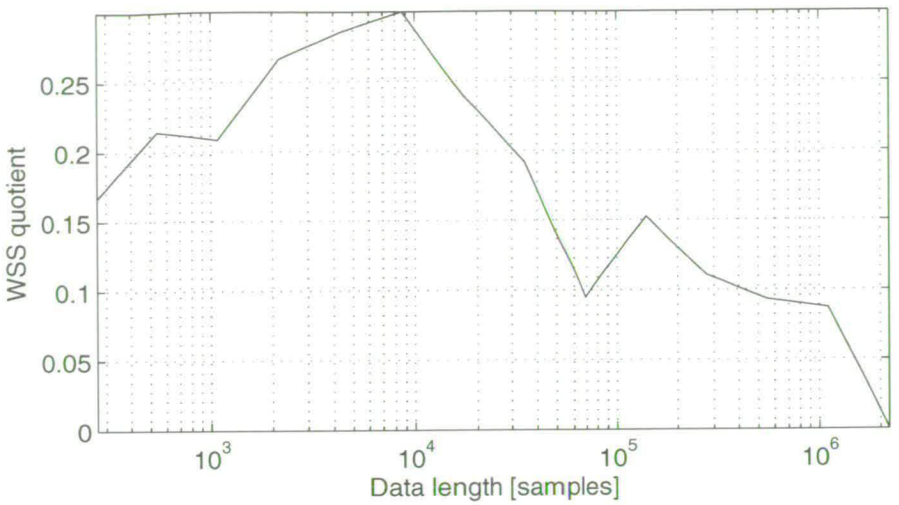


**Figure 3.26:** *The WSS quotient estimated for different lengths of data for the fibreglass boat experiment. Starting with length of 270 samples going up to 2.1 million samples, that equals 10ms to about 88s of data.*

From looking at the Figure 3.26 it is easy to see that the highest value of the WSS quotient is reached at a window length equal to 10k samples which is 0.4 seconds of data.

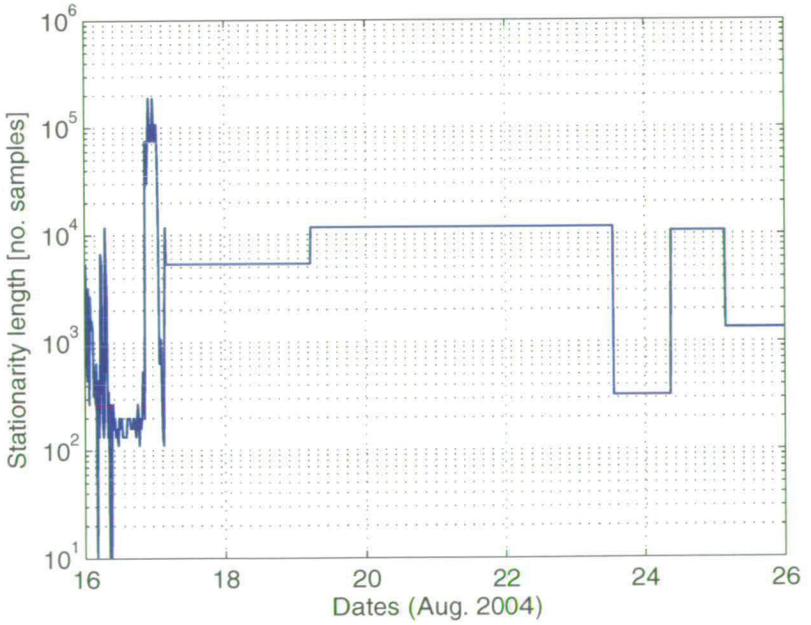


**Figure 3.27:** *The KS test applied to different lengths of data from the multi sensor trial in June 2004. Starting with lengths from 100 samples going up to 750,000 samples (that equals 4ms to about 30s of data).*

In Figure 3.27 the results from applying the KS test to ambient noise data from the multi sensor trial in June 2004, are displayed. The KS test was applied to 796 data files, each 60s of length. The data from each one minute file was rotated circularily, ten times to ensure that the results are not dependent on what part of the data was used. Then the test statistic was compared to the p-value, for each sample size respectively, to find the stationarity length of the data. As it can be seen there is a large proportion of the data that is staionary for roughly 10k samples, which equals 0.4 seconds. See appendix B where a table lists the number of stationary samples and two figures over the test statistic are shown.

### 3.9.9 Testing passive sonar data for stationarity

In this section the objective is to find out how much the sound emitted from surface vessels, which is the dominant contributor to the ambient noise, vary over time. More specifically, what kind of variation in the spectral level is found in a shallow water environment. Data from the fibreglass boat experiment is analysed by looking at the variation of the amplitude of tonals in a spectrogram. It is expected that the amplitude will vary due to the effect of multipath propagation. The lines in a spectrogram are essentially tones that arise from the machinery and drive of a vessel, See Figure 3.29. The tones being fairly narrow band in their nature, the multipath propagation gives either constructive or negative interference. At times the tones can be almost cancelled and at other times they might double in amplitude. Consider the third panel from the top in Figure 3.30, where the spectral level from the Fourth run from the fibreglass boat trial is displayed for the 137 Hz line, which happens to be the strongest. Also the power spectrum for the same run in Figure 3.28, where it is easy to see the spectral components that produce clear lines in a lofargram. The spectral level for the first seven of the peaks {78Hz, 90Hz, 137Hz, 156Hz, 176Hz, 181Hz, 195Hz} are displayed in Figure 3.30. The power spectrum is produced from a block of data that starts at 25 seconds into the data file up to 40 seconds, using 65336 samples, which gives a frequency resolution of about 0.4Hz. The relation of the first clear seven peaks in the power spectrum, to the boat are displayed in Table 3.11. In the spectrogram it is first and foremost the four lines {137Hz, 156Hz, 176Hz, 195Hz} that are multiples of the drive rotation frequency that are visible. The peaks in the power spectrum are mostly multiples of half the engine axis rotation frequency, 22.6Hz, and 19.5Hz which is the propeller shaft rotation frequency, corresponding to the engine rotation frequency divided by 2.31 (gear ratio of the primary drive). The frequencies 19.5 and 22.6Hz are not very clearly visible in the power spectrum, and they are even less visible in the spectrogram.

Multiples of these frequencies higher up in frequency are visible in both the power spectrum and the spectrogram.

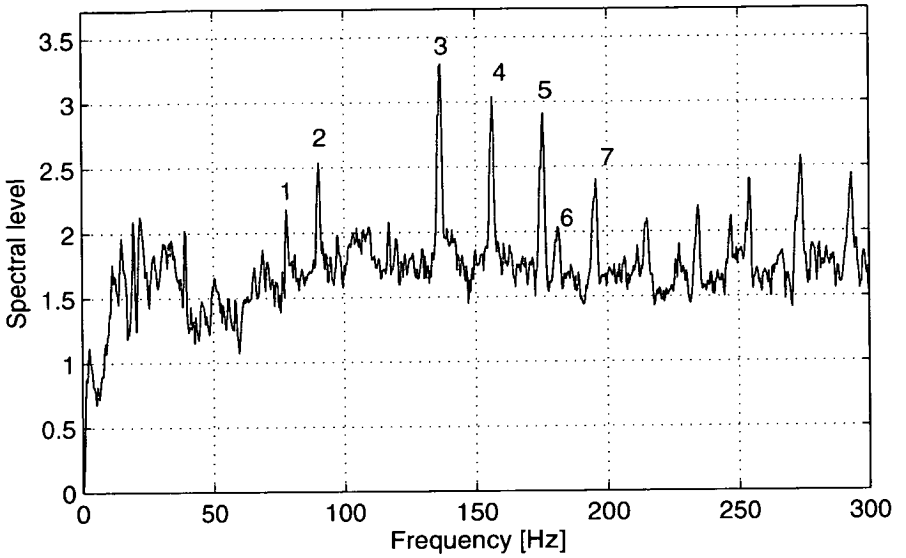


**Figure 3.28:** *A power spectrum from the fibreglass boat trial, run number 4, hydrophone 1.*

| Peak number | Frequency [Hz] | Relation to boat |
|---|---|---|
| 1 | 78 | Fourth harmonic of the drive |
| 2 | 90 | First harmonic of the engine rotation frequency |
| 3 | 137 | Seventh harmonic to the drive |
| 4 | 156 | Eighth harmonic to the drive |
| 5 | 176 | Ninth harmonic to the drive |
| 6 | 181 | Fourth harmonic to the engine rotation frequency |
| 7 | 195 | Tenth harmonic to the drive |

**Table 3.11:** *The seven first clear spectral peaks from run 4 and their relation to the boat.*

The first 20 seconds of data from this run was not used in the analysis of the data, even though it maybe displayed in the Figures. This is due to the boat not having reached a stable speed and that causes the frequencies of the emitted sound to vary significantly, as can be seen in Figure 3.29. Knowing the reason for these variations they are not an interesting part of the analysis. Furthermore it is clear to see and intuitive to understand that with the changes in the speed of the boat, and the engine and drive revelutions, there is nothing in the emitted sound that can be stationary.

The level of the first four clear spectral lines from the spectrogram can be seen in Figure 3.30,

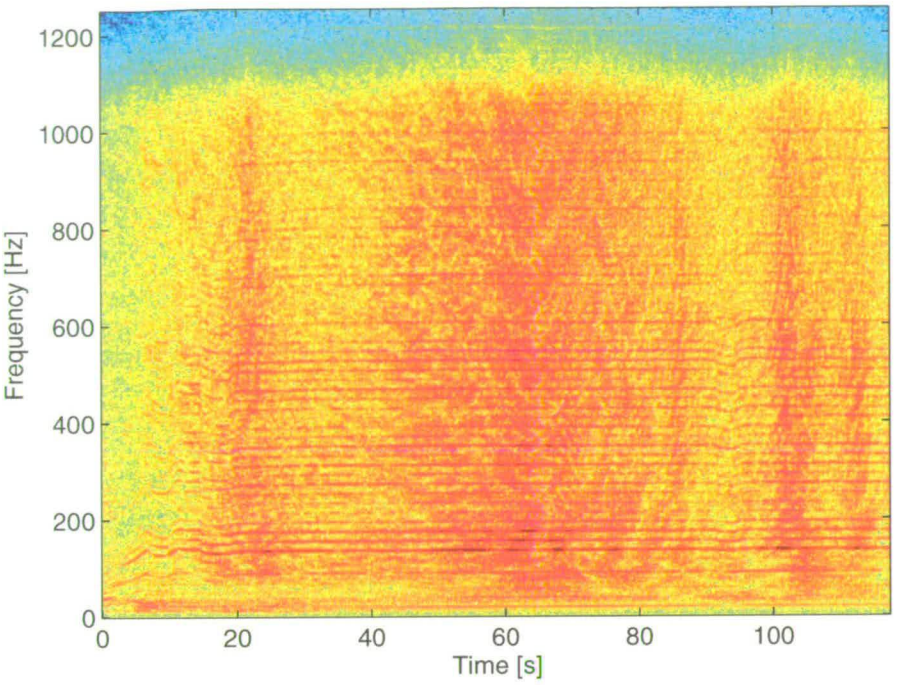


**Figure 3.29:** *Spectrogram for the fourth run with the fibreglass boat.*

they are the seventh to the tenth harmonics of the drive shaft rotation frequencies, numbered {3,4,5,7} in Figure 3.28, and Table 3.11. Also displayed in Figure 3.30 are the two first clear peaks from the power spectrum in Figure 3.28, numbered 1 and 2, corresponding to the fourth harmonic to the rotation frequency of the drive and the first harmonic to the engine rotation frequency. Another peak that is not visible in the spectrogram but that is quite clear in the power spectrum is the peak numbered 6, corresponding to the fourth harmonic of the engine rotation frequency. It is easy to see that the level varies considerably over the run. First form the start where the the level is almost at zero, mostly consisting of the back ground noise. Then at the Closest Point of Approach (CPA), which occurs at about 60 seconds from the start, where the levels are at their highest. Then falling down to almost background noise levels again as the boat has gone past CPA and the distance increases. In the time frame closest to CPA, around twenty seconds before and after, there is a region where the sound levels could have been almost stable since the distance is fairly constant between the boat and the hydrophone array. Only looking at spectrogram in Figure 3.29, this might be the conclusion since the tonal lines appear almost as strong and clear within this time frame. Never the less, it is clear to see that there are rather large variations in the levels. This is due to the multipath propagation of the emitted sound from the boat to the hydrophone.

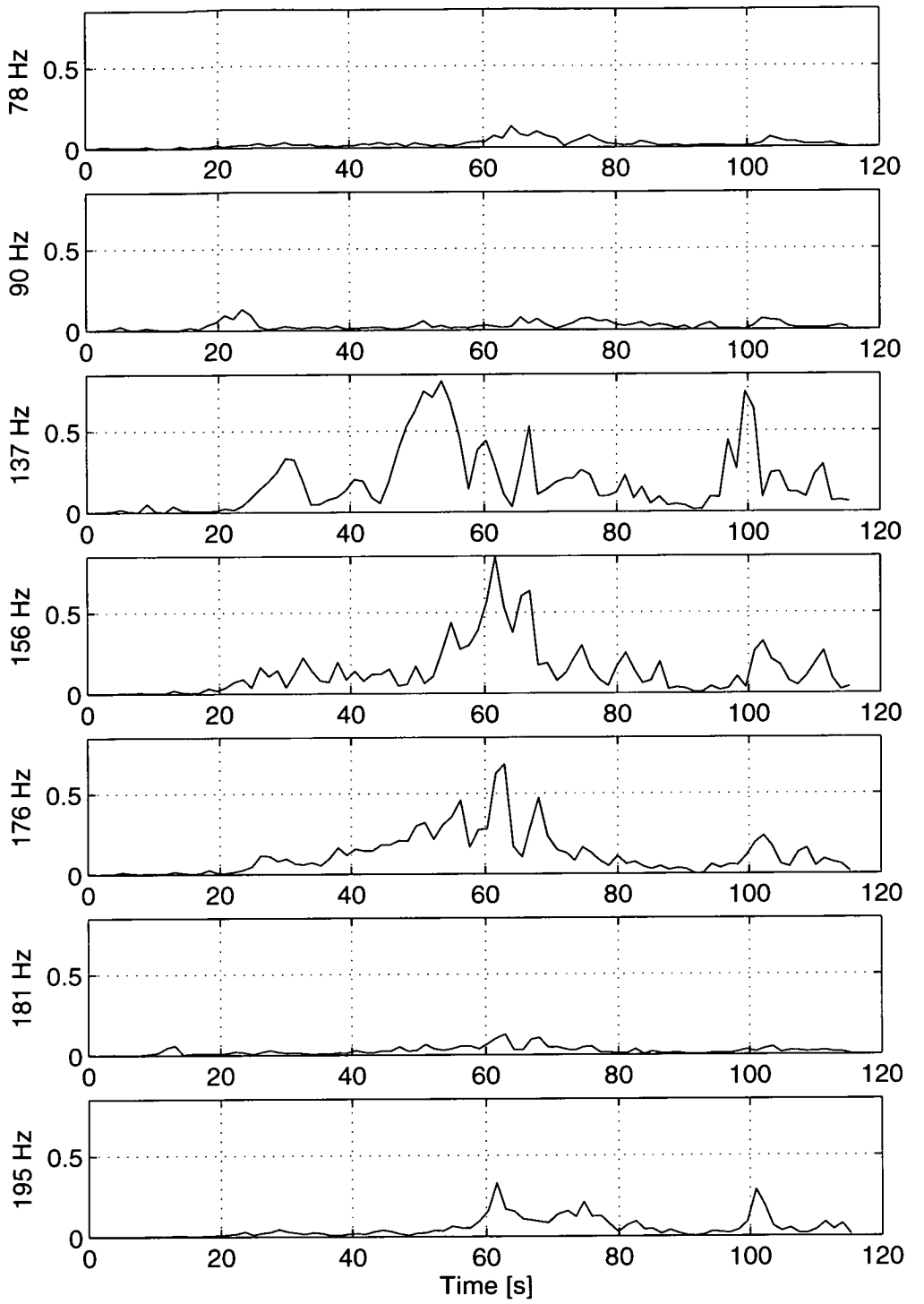


**Figure 3.30:** *Spectral levels for the first five peaks in the power spectrum displayed in Figure 3.28.*

In Figure 3.31 the results from a bispectrum based stationarity test are presented. The test is

applied to a downsampled version of the data since it is computationally demanding to estimate higher than second order spectra. The interesting part of the data is found below 1250 Hz in frequency so the test is done on the relevant part of the data. It is clear to see that the test statistic is over the threshold at all times and the null hypothesis of stationarity is rejected. Even if the KS test showed stationarity times up to 0.4 seconds it does not influence this test due to the fact that when estimating a bispectrum one needs to average over several time frames of data to ensure the significance of the estimates. In this case there are 25 windows, each of 512 samples in length, and that adds up to just above 5 seconds of data.

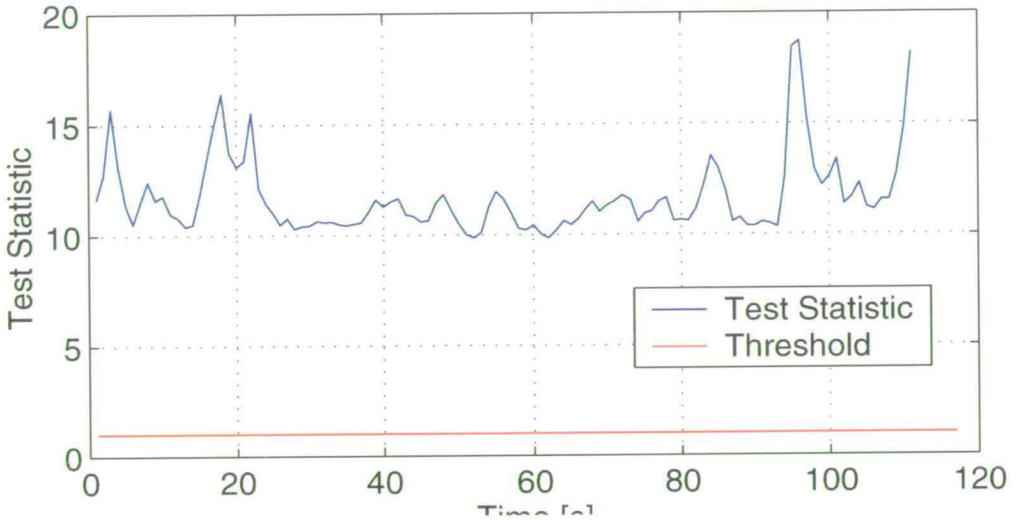


**Figure 3.31:** *A bicoherence based stationarity test applied to the data set.*

In Figure 3.32 the results from applying the KS test to the entire data file from run four from the fibreglass boat trial is presented. The test is applied to data lengths starting from 180 samples up to 1.5 million samples. In order to get data from all parts of the set the data was circularly shifted ten times using a tenth of the data length as a shift. Just like the KS test applied to the ambient noise data from the same trial the stationarity time seems to be about 10k samples, or roughly 0.4 seconds.

Yet another application of the KS test is looking at the spatial difference of this trial. This is done by comparing the different hydrophone signals since they were placed about 50 meters apart. Not surprisingly there is a significant difference between the distributions functions for the four separate hydrophones. The KS test rejects the null hypothesis and the data is spatially stationary either. The test statistic $T_{KS}$ is equal to $\{0.0295, 0.0888, 0.1427, 0.0905,$

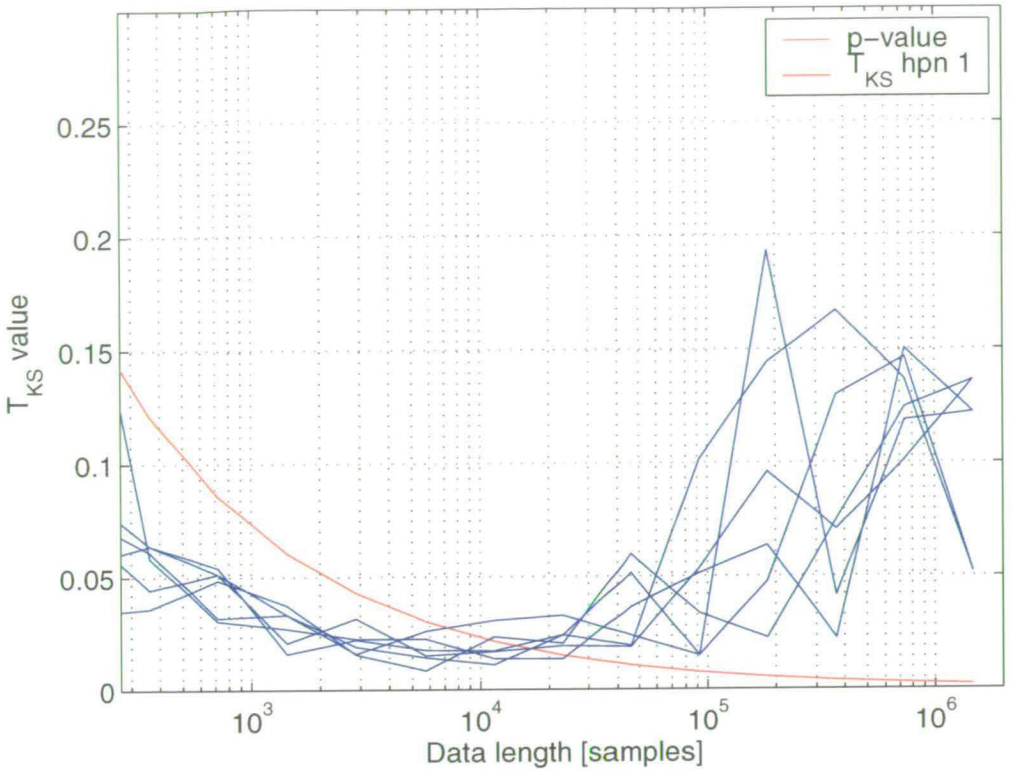


**Figure 3.32:** *The K-S test applied to run four from the fibreglass boat trial, on data lengths from 180 to 1.5 million samples. The data has been circularly shifted 10 times to get the different parts of the entire run into the test sequence for all data lengths.*

0.1210, 0.1517}, and that corresponds to comparing all the hydrophone signals to one and other according to, $\{s_1 - s_2, s_1 - s_3, s_1 - s_4, s_2 - s_3, s_2 - s_4, s_3 - s_4\}$. The p-value that $T_{KS}$ is compared to is equal to 0.0013, so $T_{KS}$ is at best a factor 20 too large for accepting $H_0$.

While knowing that the data is not spatially strictly stationary, this leaves a possibility that it is stationary to some degree. The WSS quotient is calculated for the data from the four hydrophones. That results in a WSS quotient equal to 0. This a bit surprising, but the explanation lies with the data set being small. There are only four signals to compare and that leads to a summation over only six elements in the $P$ array in Equation (3.62). Then by first removing the mean from the signal the WSS quotient rises to 0.5. And further by normalising the data so that the maximum value in each signal is equal the WSS quotient goes to 0.667. This leads to the conclusion that the data is slightly biased, and there are some low frequency disturbances in the data collection system (DC level).

However, trying the same approach to removing the mean and normalising the data do not

change the results for the KS test. With that in mind it is safe to say, with the basis of these test results that the data is not spatially stationary within the array length (200m), see Figure 2.15.

### 3.9.10 Testing for Symmetry and Linearity

The first symmetry test based on the histogram is applied to the fibre glass boat trial data. The legnth of data segments that the histograms are estimated from is equal to 10000 samples, which follows from the previous stationarity test on the data that suggests a stationarity time of about 0.4s. The histograms were estimated using 1000 bins. Examples of the histograms can be seen in Figure 3.33. It can be seen that there is a difference between the histograms which is also reflected in Figure 3.34 where the test statistic is displayed. The symmetry values differ from time to time.

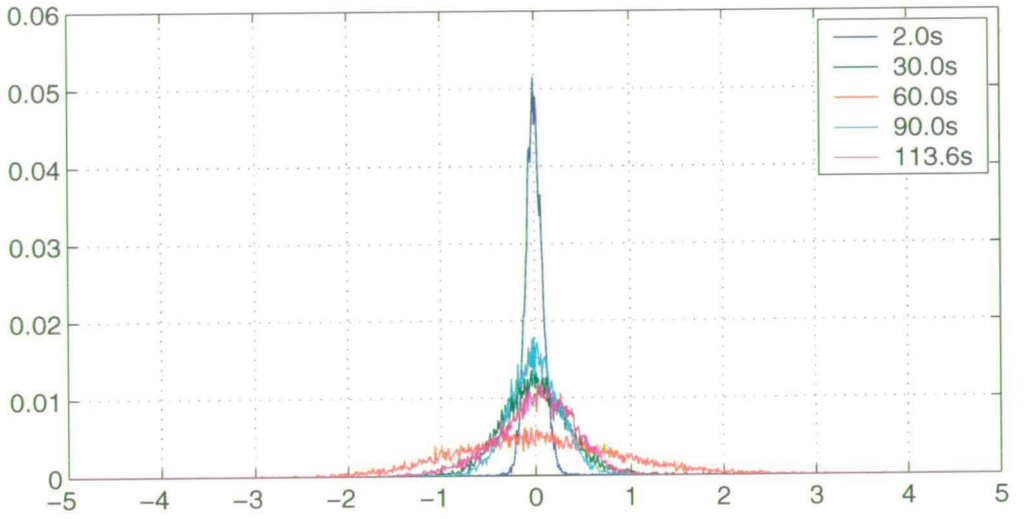


**Figure 3.33:** *Five Histograms from the fibre glass boat trial data set. The histograms are from (2.0s, 30.0s, 60.0s, 90.0s, and 113.6s) respectively.*

By studying Figure 3.34 it is clear that the test statistic lies within the limits on all occasions, except foir a short interval during the first 8 seconds and at one occasion at at about 12 seconds in and at the end of the run. So the data can be considered symmetric according to this test.

The next symmetry test is the test based on the skewness function, and it is applied to a down-sampled version of the same data. The reason for the down sampling is due to the computational cost of higher order measures and memory limitations. However, there is very little of interest above the 1250Hz mark, so the test still reflects the nature of the interesting part of the data set. Looking at Figure 3.35 one finds that the results from both tests give similar results. However,

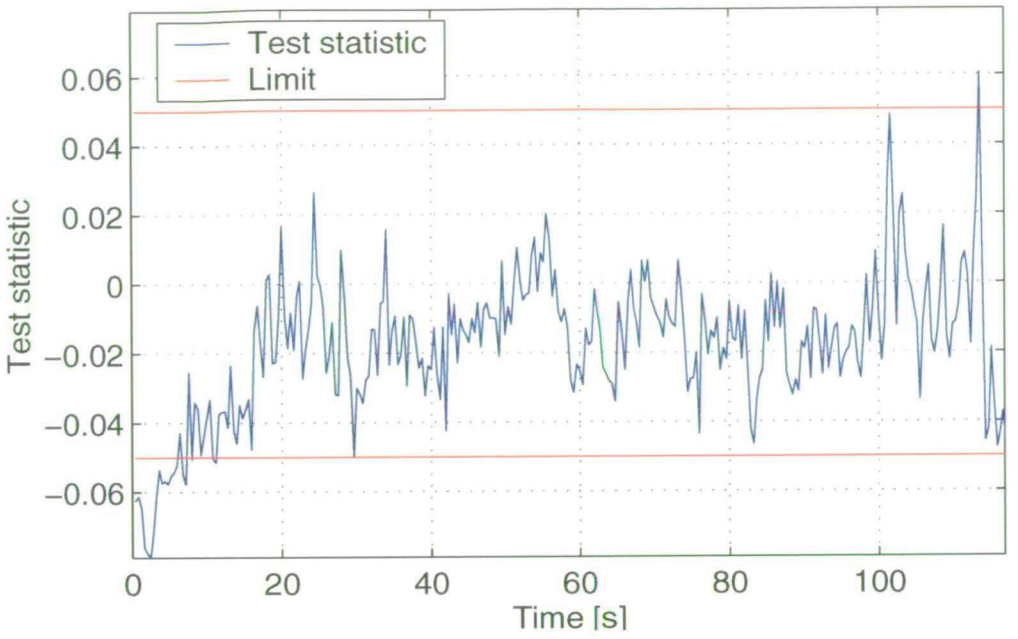


**Figure 3.34:** *The test statistic from the symmetry test according to Equation (3.63), applied to the fibre glass boat trial data.*

the regions at the start and at the end zones where the first symmetry test rejects the symmetry hypothesis are considered symmetric in the second test. The reason for that is likely the effect of the data being non-stationary, estimating the bispectrum involves a certain amount of averaging and the characteristics that render the data skewed are not captured in the smoothed out set. The skewness function is estimated using 25 windows, each of length 512 samples. That equals just over 5 seconds of downsampled data.

The result from the linearity test follows from the symmetry test, when the data is tested to be symmetrical it follows that the data is also linear. It should be noted that, since the stationarity tests reject the stationarity hypothesis and the linearity test is also based based on the normailsed bispectrum, this result should not be overiniterpreted, i.e, the data can still exhibit non-linear tendencies.

## 3.10 Conclusions

In this chapter different aspects of statistical properties of sonar data have been addressed. Mainly focusing on the level of stationarity on both active and passive sonar. It is important to bear in mind that all findings in this chapter are valid for shallow water environments with the

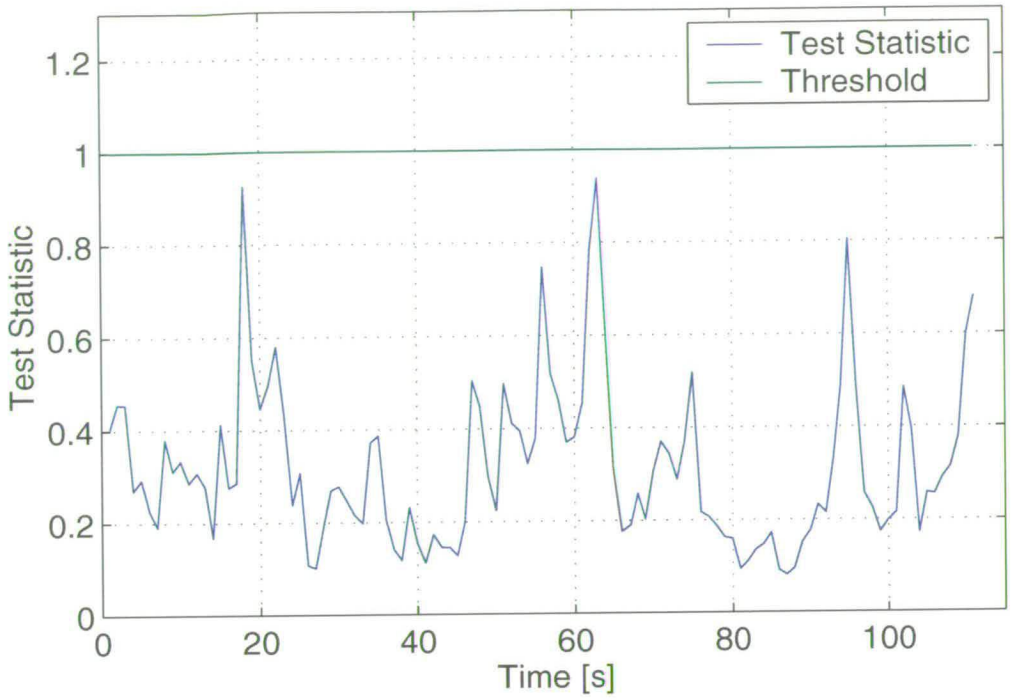


**Figure 3.35:** *The test statistic from the symmetry test based on the skewness function, applied to the fibre glass boat trial data.*

typical bottom composition of the Baltic Sea and the archipelago of Stockholm. It is however reasonable to assume that similar environments, given that the acoustic parameters also are of close to or equal to the ones in the Baltic Sea, would give similar results for the analysis done above. And the general conclusions drawn from the analysis could serve as guidelines for other shallow water areas, as for what kind of performance one might expect from applying signal processing schemes.

For active sonar the issues of ping to ping stationarity, for both the actual ping and for the reverberation tail, was examined and the data was found non-stationary. However, for some cases it is seen that there are consecutive pings that are stationary. This leads to the conclusion that it can be beneficial to use several pings from the same target. However it is necessary to remember that there is a great variation in the number of consecutive stationary pings.

For passive sonar both ambient noise and tonals emitted from surface vessels are investigated. It is found that for all cases tested the sonar data is non-stationary. That is speaking of stationarity in the strict sense. There are however parts in the data that exhibit more stationary behaviour then the data in general. From the fibre glass boat trial the data seems to be stationary for about

0.4 seconds. That is the pdfs for two consecutive segments maintain sufficiently equal for that period of time.

Testing the passive data for symmetry and linearity show that the data is mostly linear, and symmetrical.

The ambient noise is found to have more stationary properties at higher frequencies. This leads to the conclusion that it is more beneficial to operate sonar systems at higher frequencies. Again this is only valid for low salinity as in the Baltic Sea, where attenuation is not of great importance. However, in waters with higher salinity the sonar performance would degrade from using higher frequency. The stationarity length on ambient noise data was examined using data from the from the multisensor trial (almost 800 data files). A large proportion of the data set also had a stationarity time of roughly 0.4 seconds, or slightly longer.

# Chapter 4
# Time/Frequency Analysis of Sonar Data

Time/frequency analysis of signals is one way of addressing the issue of non-stationarity. Since the signal properties change over time there is a value in finding out when and where in both time and frequency these changes take place. In a sonar context this traditionally means using a lofargram, as described in Section 2.6.

The lofargram is a spectrogram with time running down on the y-axis and frequency on the x-axis. Traditionally time frequency analysis has been done on passive data, i.e. only listening with the sonar system, and no active transmission of sound. The lofargram is an excellent tool for finding narrowband tones from sources such as the machinery, drive, and propeller. These tones show up as lines in the lofargram and combined with an experienced sonar operator is indeed a powerful system of detecting and classifying objects in the water, both on the surface and in the volume, see the right panel in Figure 4.1. Also see Section 2.6 where a few examples are given on the lofargram of a typical sonar data recording.

The main reason for using time/frequency analysis instead of time series or frequency domain analysis is the need to find where different signal components are located in time.

The analysis of time-series data has traditionally been done by conventional time-domain techniques or in the frequency domain. Over the last ten years or more, interest in the use of Time-Frequency tools has increased. Popular techniques in the radar and sonar context have been the short time Fourier transform or the windowed power spectrum also known as the spectrogram. Later on other techniques have emerged that are more specific to the data at hand. Wigner Distributions [70] and others like it [71], have proved useful. Different wavelets has also been used [72], [73]. See Figure 4.2 for examples of chirps and chirplets, waves and wavelets.

In this chapter the spectrogram, and Short-Time Fractional Fourier Transform (STFrFT) [74], [75], [76] are used to analyse sonar data. The Wigner Distribution (WD), and the Radon Wigner Transform are briefly discussed since the WD is a well known and widely used time/frequency

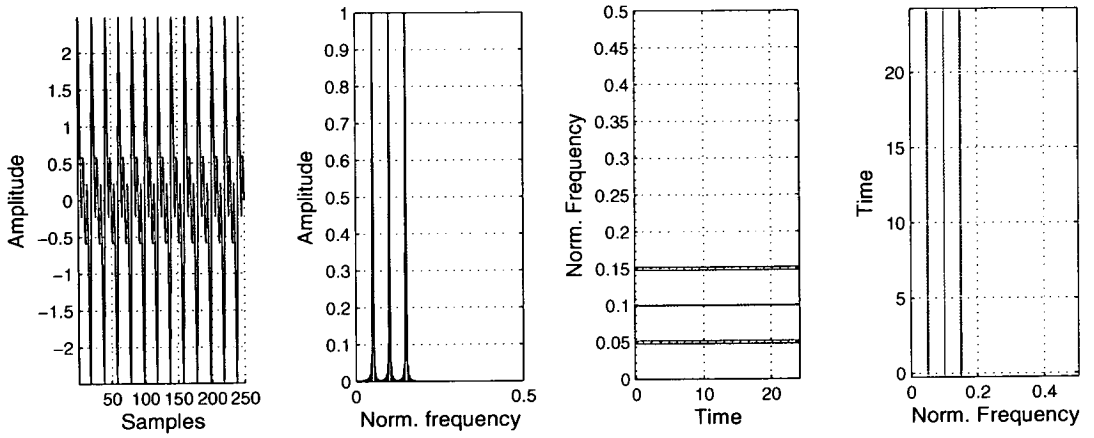


**Figure 4.1:** *The panels display, from the left: A sum of sinusoids, the power spectrum, a spectrogram, and the lofargram.*

method. The RWT is included due to it's relation to the FrFT.

One of the drawbacks with time/frequency methods like the WD is the cross terms that appear in the time/frequency images. These are discussed in more detail later on in this Chapter. However a major breakthrough was made with the Choi-Williams Distribution (CWD), [77], [78],[79], [80], which is one of the most common methods of coping with cross terms in the Wigner displays. It is discussed in Appendix D where it is applied to some synthetic chirps.

The nature of active sonar data makes it easier to use a transform well matched to the received signal since the output of the transmitter is known. However, the return signal is always distorted in a number of ways and so there is always room for improvement in trying to extract more information from the data. In [81] the chirplet transform is introduced and it also provides a useful framework for other time/frequency/scale transforms.

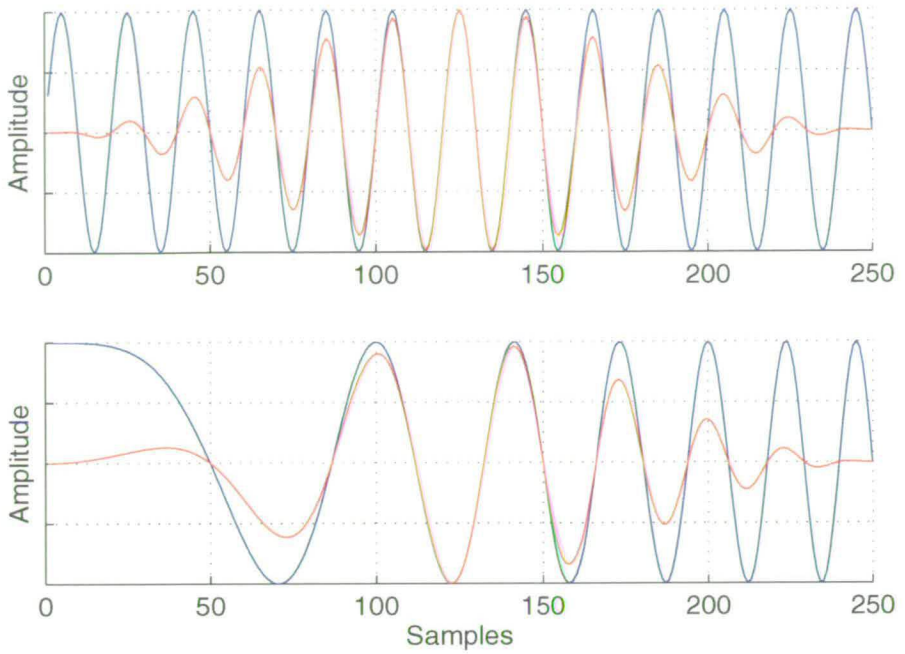


**Figure 4.2:** *The wave and wavelet, and the chirp and chirplet. Both wavelet and chirplet can be viewed as wavelet basis functions and are both localised in time as well as frequency.*

Since the main scope of this chapter is on methods of use in a sonar and sonar operator context, and the most widely used method is the lofargram, the comparison of time/frequency methods is done between the lofargram and the STFrFT. That is mainly due to the similar fashion of producing the surfaces. That is windowing the signal and placing the windowed segments adjacent to each other to form a surface.

Aside from applying the different transform methods an assessment is done to the quality of the produced image. Regardless of how good a transform seems there is often a question left unanswered, namely how much better one transform is compared to another, and in what respect. In order to answer a question like that a number of possible image quality measures, [82] have been explored.

## 4.1 Time/Frequency Theory

A time-frequency representation maps a one-dimensional signal $f(t)$ into a two-dimensional joint time-frequency distribution $P(t, \omega)$. The Time-Frequency Distributions represents the

energy distribution of the signal.

### 4.1.1 Joint Time-Frequency Energy Distributions

The time domain signal $f(t)$ is uniquely related with the signal $F(\omega)$ in the frequency domain by its Fourier transform.

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \tag{4.1}$$

The Fourier transform $F(\omega)$ is also called the spectrum of the signal $f(t)$, and it represents the frequency content in $f(t)$. The energy per time unit and frequency unit for the signal $f(t)$ is $E(t)$ and $E(\omega)$, respectively.

$$E(t) = |f(t)|^2 \tag{4.2}$$

$$E(\omega) = |F(\omega)|^2 \tag{4.3}$$

Then the total energy should be equally preserved in both time and frequency domains, which is expressed by Parceval's relation as

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |F(\omega)|^2 d\omega. \tag{4.4}$$

The joint time-frequency distribution $P(\omega, t)$ represents the energy distribution of the signal. The total energy of $P(\omega, t)$ should also be equal to the total energy of the signal.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(\omega, t) d\omega dt = \int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |F(\omega)|^2 d\omega \tag{4.5}$$

To get the time and frequency marginal of the TFD, the TFD is integrated over frequency and time, respectively,

$$P(t) = \int_{-\infty}^{\infty} P(\omega, t) d\omega \tag{4.6}$$

$$P(\omega) = \int_{-\infty}^{\infty} P(\omega, t)dt \tag{4.7}$$

The time and frequency marginals are both one dimensional distributions. The marginals should satisfy the conditions,

$$P(t) = |f(t)|^2 \tag{4.8}$$

$$P(\omega) = |F(\omega)|^2 \tag{4.9}$$

The conditions in Equation (4.8) and Equation (4.9) do not uniquely determine the time-frequency distribution. However, one can formulate an intuitive physical interpretation of the joint TFD, by defining the average and variance for the TFD. The average and variance in time gives an indication of where the signal energy is concentrated and the spread of the signal energy in time. The average in time is defined by,

$$\mu_t = E\{t\} = \int_{-\infty}^{\infty} tP(t)dt = \int_{-\infty}^{\infty} t|f(t)|^2 dt \tag{4.10}$$

The variance in time is given by,

$$\sigma_t^2 = \int_{-\infty}^{\infty} (t - \mu_t\})^2 P(t)dt = \int_{-\infty}^{\infty} (t - \mu_t)^2 |f(t)|^2 dt \tag{4.11}$$

The average and variance in frequency describes the location and the spread of the energy in frequency. The average in frequency is defined by,

$$\mu_\omega = E\{\omega\} = \int_{-\infty}^{\infty} \omega P(\omega)d\omega = \int_{-\infty}^{\infty} \omega |F(\omega)|^2 d\omega \tag{4.12}$$

The variance in frequency is given by,

$$\sigma_\omega^2 = \int_{-\infty}^{\infty} (\omega - \mu_\omega\})^2 P(\omega) d\omega = \int_{-\infty}^{\infty} (\omega - \mu_\omega)^2 |F(\omega)|^2 d\omega \tag{4.13}$$

### 4.1.2 The uncertainty principle

The uncertainty principle was articulated in the context of quantum mechanics, which has an inherently probabilistic concept. In transform theory the uncertainty principle states a fundamental relation between the signal and its transform. The relation can be expressed as,

$$\sigma_t \sigma_\omega \geq \frac{1}{2} \tag{4.14}$$

where $\sigma_t$ and $\sigma_\omega$ are the spread of the marginals (given by Equations (4.11) and (4.13)). This means that the signal and its transform cannot be arbitrarily narrow at the same time. The uncertainty principle only constrains the relation between the marginals and not the resolution in the TFD itself. Hence, the uncertainly principle and high resolution in time and frequency simultaneously in a TFD, do not contradict. This is due to the possibility of choosing different resolution in time and frequency in different parts of the TFD. However, in each part of the TFD the uncertainty principle always limits the resolution.

## 4.2 Time/Frequency Distributions

More recently other time/frequency and time/scale methods than the STFT have emerged. Of these newer methods the Wigner distribution is one of the first that springs to mind, also the wavelet transform is widely used. These methods offer better resolution in both time and frequency and provides a much better representation of broad band signals such as transients, be it noise or active transmissions. The nature of these transforms are better suited for that kind of analysis and one way of illustrating that is to look at one basis function of the wavelet transform, that is one of infinitely many. See Figure 4.2 where a generic wavelet and the chirplet are depicted. In cases where the nature of the data is even better understood and known as in active transmission of pulsed frequency sweeps (chirps) yet another method of analysing the sonar data presents itself, the Fractional Fourier Transform (FrFT). Never the less, the STFT is still a corner stone in time/frequency analysis and it is presented first of the methods mentioned

above.

### 4.2.1 The STFT and the Spectrogram

The short-time Fourier transform is the most widely used method for studying time-varying and non-stationary signals. The popularity and usefulness of the method lies in it's simplicity, and very likely the fact that the Fourier Transform is familiar in all disciplines of science.

The spectrogram can be used to study frequency properties at a particular time, or time properties at a particular frequency. But as previously stated arbitrarily high resolution in time and frequency simultaneously is not possible. When the time intervals become too short the information from the spectrum becomes meaningless, and shows no relation to the spectrum of the original signal. The reason is that short duration signals have large a bandwidth, and the spectra of such short duration signals have very little to do with the properties of the original signal. This is caused by the uncertainty principle applied to the short time intervals that we have created for the purpose of the analysis.

To study the properties of a signal at time $t$, the signal is multiplied with a window function $h(t)$ to produce the modified signal,

$$f_t(\tau) = f(\tau)h(\tau - t) \tag{4.15}$$

The modified signal is a function of two time variables, the fixed time $t$, and the running time, the timelag $\tau$. The window function is chosen to leave the signal as intact as possible around time $t$ without introducing artefacts to the transform, and to suppress the signal outside the window.

$$h(t) = \begin{cases} 1 & |t| < \tau/2 \\ 0 & |t| > \tau/2 \end{cases} \tag{4.16}$$

This gives the modified signal,

$$f_t(\tau) = \begin{cases} f(t) & |t| < \tau/2 \\ 0 & |t| > \tau/2 \end{cases} \tag{4.17}$$

Since the modified signal contains the original signal around time $t$, the Fourier transform will reflect the distribution around that time,

$$\begin{aligned} F_t(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_t(\tau) e^{-j\omega\tau} d\tau \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) h(\tau - t) e^{-j\omega\tau} d\tau \end{aligned} \tag{4.18}$$

The power spectral density at time $t$ is therefore,

$$\begin{aligned} P_{SP}(t,\omega) &= |F_t(\omega)|^2 \\ &= |\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\tau) h(\tau - t) e^{-j\omega\tau} d\tau|^2 \end{aligned} \tag{4.19}$$

For each different time $t$ a different spectrum is obtained. All of these spectra make up the time-frequency distribution and it is called the spectrogram. The spectrogram has the the following general properties:

The characteristic function of the spectrogram is given as,

$$\begin{aligned} M_{SP}(\Theta, t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F_t(\omega)|^2 e^{j(\omega\tau + \Theta t)} d\omega dt \\ &= A_f(\Theta, \tau) A_h(-\Theta, \tau) \end{aligned} \tag{4.20}$$

where $A_f(\Theta, \tau)$ is the ambiguity function of the signal, and is defined below

$$A_f(\Theta, \tau) = \int_{-\infty}^{\infty} f^*(t - \tau/2) f(t + \tau/2) e^{j\Theta t} dt, \tag{4.21}$$

and $A_h(\Theta, \tau)$ is the ambiguity function of the window defined as,

$$A_h(\Theta, \tau) = \int_{-\infty}^{\infty} h^*(t - \tau/2) h(t + \tau/2) e^{j\Theta t} dt. \tag{4.22}$$

The total energy is obtained by integrating the time-frequency distribution over time and frequency. However the total energy can also be obtained by evaluating the characteristic function

at zero. Then by using Equations (4.21) and (4.22),

$$\begin{aligned} E_{SP} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{SP}(t,\omega)dtd\omega = M_{SP}(0,0) \\ &= \int_{-\infty}^{\infty} |f(t)|^2 dt \times \int_{-\infty}^{\infty} |h(t)|^2 dt \end{aligned} \tag{4.23}$$

Which means that the total energy of the spectrogram can be obtained by multiplying the total energy of the signal with the total energy of the window. By choosing the energy of the window to be one, the total energy of the spectrogram will be equal to the total energy of the signal. Then the following is integrated over frequency to obtain the time marginal of the spectrogram,

$$\begin{aligned} P(t) &= \int_{-\infty}^{\infty} P_{SP}(t,\omega)d\omega = \int_{-\infty}^{\infty} |F_t(\omega)|d\omega \\ &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\tau)h(\tau-t)f^*(\tau')h^*(\tau'-t)\delta(\tau-\tau')e^{j\omega(\tau-\tau')}d\tau d\tau' d\omega \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\tau)h(\tau-t)f^*(\tau')h^*(\tau'-t)\delta(\tau-\tau')d\tau d\tau' \\ &= \int_{-\infty}^{\infty} |f(\tau)|^2|h(\tau-t)|^2 d\tau \end{aligned} \tag{4.24}$$

To obtain the frequency marginal a similar operation but over time is performed, which yields,

$$P(\omega) = \int_{-\infty}^{\infty} |F(\omega)|^2|H(\omega-\omega')|^2 d\omega' \tag{4.25}$$

It can be seen that the marginals of the spectrogram do not generally satisfy the marginal conditions, which are given by Equation (4.8) and Equation (4.9). The reason is that the spectrogram mixes the energy distributions of the signal and the window. Because the marginals are not satisfied, average of time and frequency will never be correct. However, if the energy of the window function is chosen set to one i.e. $\int_{-\infty}^{\infty} |H(\omega-\omega')|^2 d\omega' = 1$ then the marginals are satisfied.

In general arbitrarily high resolution in both time and frequency simultaneously is not possible. To get good time resolution a narrow window has got to be used in the time domain, vice versa for good frequency resolution, a narrow window in the frequency domain. But as is well known,

a narrow window in the time domain results in a broad window in the frequency domain, hence there is a trade-off between time and frequency resolution. The degree of trade-off depends on the window, signal, time, and frequency. The uncertainty principle (Equation (4.14)) quantifies these dependencies.

### 4.2.2 The Wigner distribution and the Radon-Wigner distribution

The Wigner distribution was the first distribution introduced that is qualitatively different from the spectrogram. The spectrogram is linear and the Wigner distribution is bilinear, see Chapter 8 [83]. The Wigner distribution is the foundation for many other bilinear distributions, such as Choi-Williams method. The Wigner distribution is defined in terms of the signal,

$$W(t,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f^*(t-\tau/2) f(t+\tau/2) e^{-j\omega\tau} d\tau \tag{4.26}$$

or its spectrum,

$$W(t,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F^*(\omega+\Theta/2) F(\omega-\Theta/2) e^{-j\Theta t} d\Theta \tag{4.27}$$

The Wigner distribution at a particular time is obtained by adding up pieces made up of the product of the signal at a past time with the signal at a future time, the time into the future being equal to the time into the past, see Equation (4.26) where the signal $f$ enters twice $t-\tau/2$ and $t+\tau/2$.

In order to understand the behaviour of the Wigner distribution at a given time. The distribution can be viewed as the past part of the signal folded over to the future part of it, as can be seen in the definition 4.26.

If there is an overlap in the signals, the properties in the overlap are also present in the Wigner distribution at the time. An example of this is shown in Figure 4.3. Where a signal consisting of:

- The top panel displays the Wigner distribution of a CW at 700 Hz of 0.02 seconds length, followed by another 0.02 seconds of zeros and then another CW at 1.2kHz, also of 0.02 seconds in length.

- The bottom panel displays the Wigner distribution of the same CW's as the top panel but the zeros are exchanged for Gaussian (N(0,1)) noise.

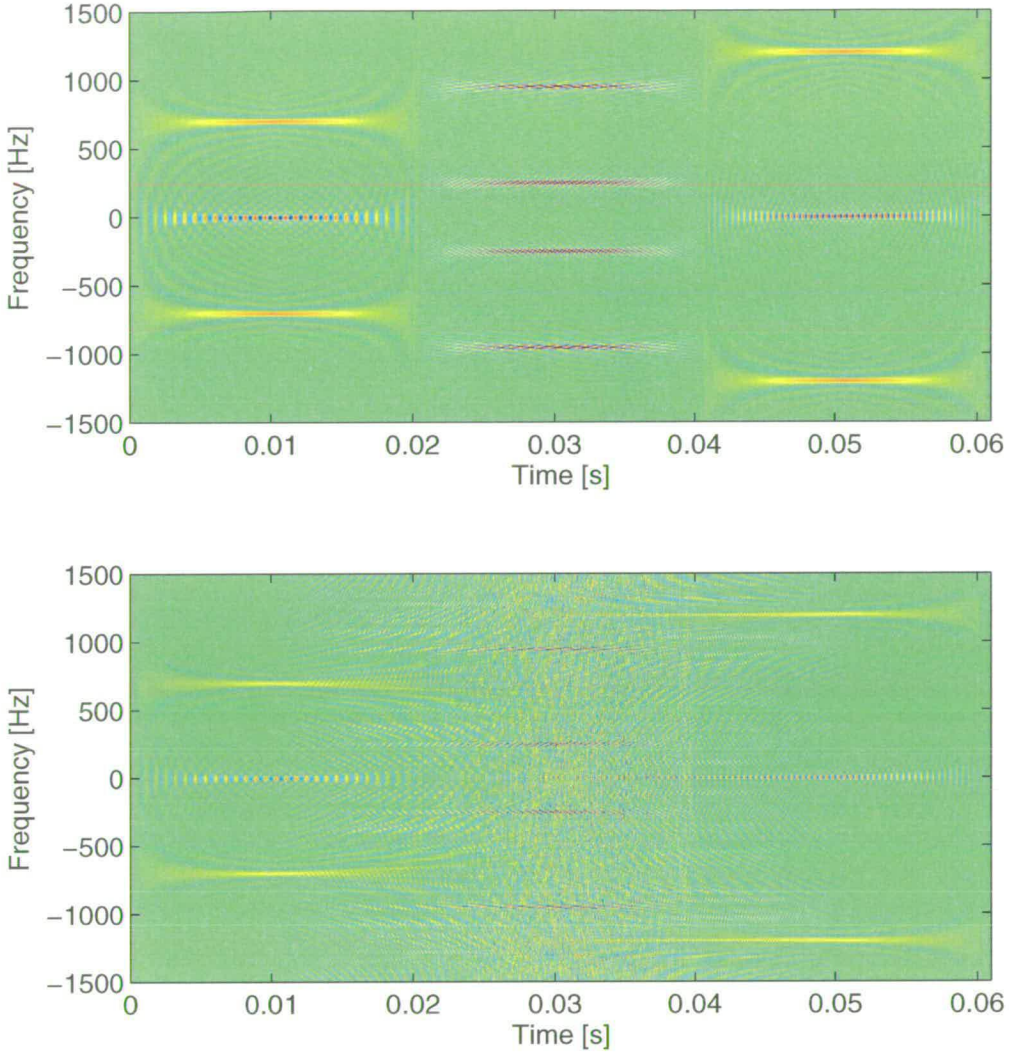


**Figure 4.3:** *The Wigner distribution for two CW signal separated in time, in the top panel the gap between the CW signals is zeros, whereas in the bottom the gap is Gaussian (N(0,1)) noise.*

The signals displayed in Figure 4.3 are finite duration signals with noise, or zeros present for only a short period of time, [0.02-0.04s]. The following applies regardless of what signal property might be present at any given time but it is easiest described in terms of noise. To determine if there is noise in the Wigner distribution at a time, one can mentally fold the signal at that time of interest. If there is noise in the overlap there will be noise in the Wigner distribution, even if

there is no noise in the signal itself at that time, see Figure 4.3 outside the interval [0.02-0.04s]. This yields that the Wigner distribution weights the far away times equally to the near times, which states that the Wigner distribution is highly nonlocal.

All properties for the time domain also hold for the frequency domain, because the Wigner transform is almost identical in form in both domains.

It has already been mentioned that the Wigner distribution is nonlocal. What will happen if we have a signal of infinite duration? From the discussion above it is clear that for an infinite duration signal, the Wigner distribution will be nonzero for all times.

Consider the signal already presented in the bottom panel Figure 4.3, it is a signal of finite duration. First the Wigner distribution is calculated for times before the signal starts. It is clear that folding the signal over from a time before the signal has started, does not result in any overlap. Neither will calculating the Wigner distribution for a time after the signal has ended.

$$W(t,\omega) = 0, \quad \text{for } t \text{ outside } \quad (t_1, t_2) \quad \text{if} \quad f(t) \quad \text{is zero outside} \quad (t_1, t_2) \tag{4.28}$$

As mentioned before the properties of the Wigner distribution are basically the same for the frequency domain as for the time domain. So for a band limited signal the Wigner distribution will be zero for frequencies outside the band, which gives

$$W(t,\omega) = 0, \quad \text{for } \omega \text{ outside } \quad (\omega_1, \omega_2) \quad \text{if} \quad F(\omega) \quad \text{is zero outside} \quad (\omega_1, \omega_2) \tag{4.29}$$

For the Wigner distribution the characteristic function is given by,

$$\begin{aligned} M_W(\Theta,\tau) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(t,\omega)e^{j(\Theta t+\tau\omega)}dtd\omega \\ &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j(\Theta t+\tau\omega)}f^*(t-\tau'/2)f(t+\tau'/2)e^{-j\omega\tau'}d\tau' d\tau d\omega \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j\Theta t}\delta(\tau-\tau')f^*(\tau-\tau'/2)f(\tau-\tau'/2)d\tau' dt' \\ &= \int_{-\infty}^{\infty} f^*(t-\tau/2)f(t+\tau/2)e^{j\Theta t}dt \\ &= A(\Theta,\tau) \end{aligned} \tag{4.30}$$

where $A(\Theta,\tau)$ is the symmetric ambiguity function. For the Wigner distribution the characteristic function equals the ambiguity function, see page 74 in [84]. In terms of the spectrum that

yields

$$M_W(\Theta,\tau) = \int_{-\infty}^{\infty} F^*(\omega+\Theta/2)F(\omega-\Theta/2)e^{j\tau\omega}d\omega \tag{4.31}$$

The Wigner distribution is real valued for all signals. This can be shown by looking at the complex conjugate of the Wigner distribution.

$$\begin{aligned} W^*(t,\omega) &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} f(t-\tau/2)f^*(t+\tau/2)e^{j\omega\tau}d\tau \\ &= -\tfrac{1}{2\pi}\int_{\infty}^{-\infty} f(t-\tau/2)f^*(t+\tau/2)e^{-j\omega\tau}d\tau \\ &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} f(t-\tau/2)f^*(t+\tau/2)e^{j\omega\tau}d\tau \\ &= W(t,\omega) \end{aligned} \tag{4.32}$$

Observe the change in the integration limits in the second row of Equation (4.38). The symmetry properties of the Wigner distribution are examined by substituting $-\omega$ for $\omega$. And that yields:

For a real signal, the same form is returned for $-\omega$ as for $\omega$. Real signals have symmetrical spectra which leads to the Wigner distribution being symmetrical in the frequency domain. Similarly, for a real spectra the time signal is symmetrical and the Wigner distribution is symmetrical in time.

$$\begin{gathered} W(t,\omega) = W(t,-\omega) \quad \text{for real valued signals.} \\ \Leftrightarrow \\ \text{Symmetrical spectra,} \quad F(\omega) = F(-\omega) \end{gathered} \tag{4.33}$$

$$\begin{gathered} W(t,\omega) = W(-t,\omega) \quad \text{for real valued spectra.} \\ \Leftrightarrow \\ \text{Time reversible signals,} \quad f(t) = f(-t) \end{gathered} \tag{4.34}$$

The Wigner distribution satisfies the marginals, starting with the time marginal,

$$\int_{-\infty}^{\infty} W(t,\omega)d\omega = |f(t)|^2 \tag{4.35}$$

and the frequency marginal,

$$\int_{-\infty}^{\infty} W(t,\omega)dt = |F(\omega)|^2 \tag{4.36}$$

Which can be verified by looking at the characteristic function

$$M_W(\Theta,0) = \int |f(t)|^2 e^{j\Theta t}dt \quad ; \quad M_W(0,\tau) = \int |F(\omega)|^2 e^{j\tau\omega}d\omega \tag{4.37}$$

That the Wigner distribution satisfies the marginals can also be calculated directly, starting with the time marginal, the distribution is integrated over frequency,

$$\begin{aligned} P(t) &= \int_{-\infty}^{\infty} W(t,\omega)d\omega \\ &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty}\tfrac{1}{2\pi}\int_{\infty}^{-\infty} f(t-\tau/2)f^*(t+\tau/2)e^{j\omega\tau}d\tau d\omega \\ &= \int_{-\infty}^{\infty} f(t-\tau/2)f^*(t+\tau/2)\delta(\tau)d\tau \\ &= |s(t)|^2 \end{aligned} \tag{4.38}$$

Then to obtain the frequency marginal of the Wigner distribution, the integration is done over time in a similar way, which yields,

$$P(\omega) = \int_{-\infty}^{\infty} W(t,\omega)dt = |F(\omega)|^2 \tag{4.39}$$

Equations (4.35), and (4.36) state that the Wigner distribution satisfies the marginal conditions. Since the marginals are satisfied, the total energy condition is automatically satisfied, that is

$$E = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W(t,\omega)dtd\omega = \int_{-\infty}^{\infty} |f(t)|^2 dt \tag{4.40}$$

An example of how the WD looks for two chirped signals can be seen in Figure 4.4. The signals are, starting with the one of slightly lower frequency, a chirp of approximately 0.04 seconds duration with a start frequency of 5.1kHz ending at 7kHz. The second chirp is one that starts at 7kHz and stops at 9.9kHz, and also of 0.04 seconds duration.

In Figure 4.5 the same two chirps are shown, with the difference that they are now separated in

time. It is noteworthy that the cross terms are visible where the signals are equal to zero. This is yet another example of the nonlocal behaviour of the WD described in the beginning of this section.

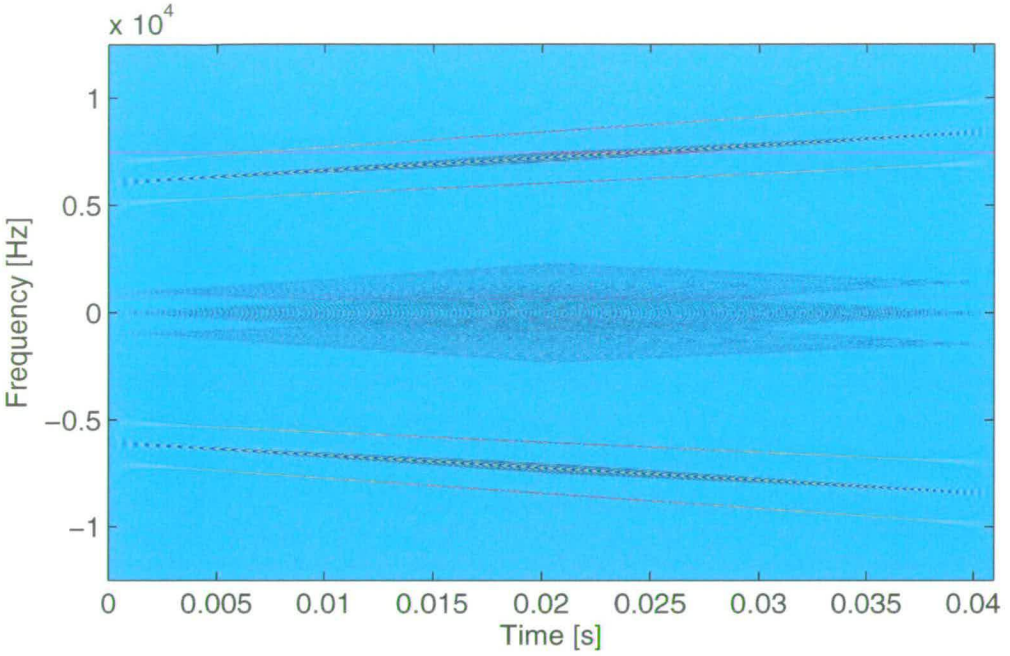


**Figure 4.4:** *The Wigner distribution for the sum of two chirps*

In [85], it is stated that the WD can be an optimal detector for linear frequency modulated pulses given that the noise is AWGN and that the WD is integrated along all possible rotational angles of the time/frequency plane. This is actually equivalent to performing a Radon Transform (RWT) on the signal, [86]. In [87] the theory on the discrete implentation of the RWT is presented, then it is followed by applications in [88].

In [89], [90] the RWT of a signal $f(t)$ is defined as

$$\begin{aligned} D_f(r,\phi) &= R_W[f(t)] = R[W_f(t,\omega)] \\ &= \int W_f(r\cos(\phi) - s\sin(\phi), r\sin(\phi) + s\cos(\phi))ds, \end{aligned} \tag{4.41}$$

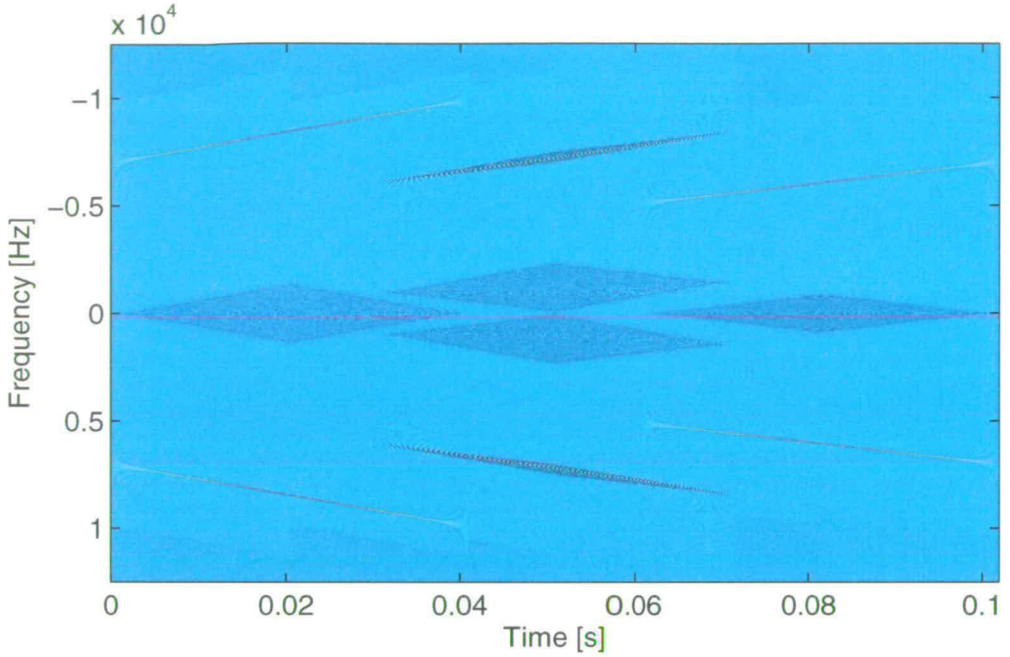


**Figure 4.5:** *The Wigner distribution for two chirps separated in time*

where

| | |
|---|---|
| r | the $t$-axis rotated an angle $\phi$ |
| s | the $\omega$-axis rotated an angle $\phi$ |
| $W_f$ | WD |
| $R[\,]$ | Radon Transform Operator |
| $R_W[\,]$ | Radon-Wigner Operator |
| $D_f(r,\phi)$ | RWT of a signal $f(t)$. |

An example of the RWT applied to synthetic data can be seen in Figure 4.6, where both the WD and the RWT of a linear chirp are displayed. The chirp has a bandwidth of 12kHz, a centerfrequency of 6.25kHz and a duration of 10ms.

Another example is shown in Figure 4.8 where the RWT and the WD of a chirp from the SAS trial data set are displayed. In Figure 4.7 one slice of the surface in Figure 4.8 is displayed.

This is at the optimum rotational angle $74^o$, i.e. the rotational angle where the maximum pulse compression is achieved.

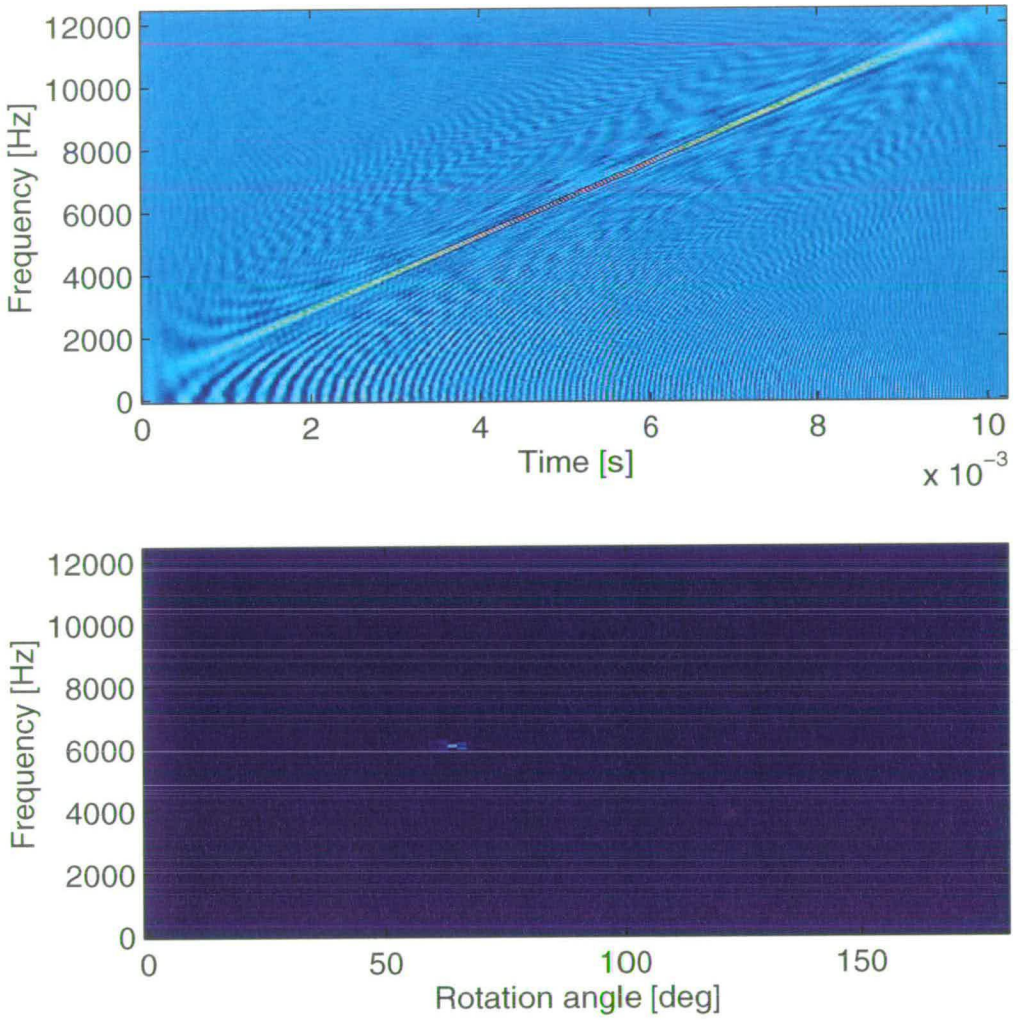


**Figure 4.6:** *The Wigner distribution and the Radon-Wigner transform of a linear chirp. The chirp has a bandwidth of 12kHz, a centerfrequency of 6.25kHz, and a duration of 10ms.*

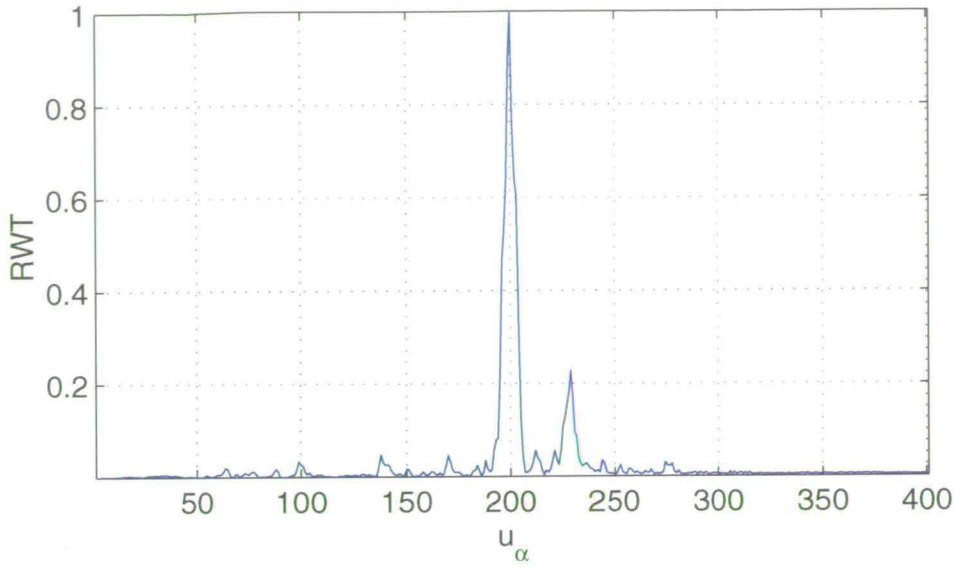


**Figure 4.7:** *Radon-Wigner transform of a linear chirp, at a rotational angle of 74 degrees. The chirp has a bandwidth of 12kHz, a centerfrequency of 6.25kHz, and a duration of 10ms The y-axis is the normalised amplitude of the RWT and the x-axis is the rotated frequency axis.*

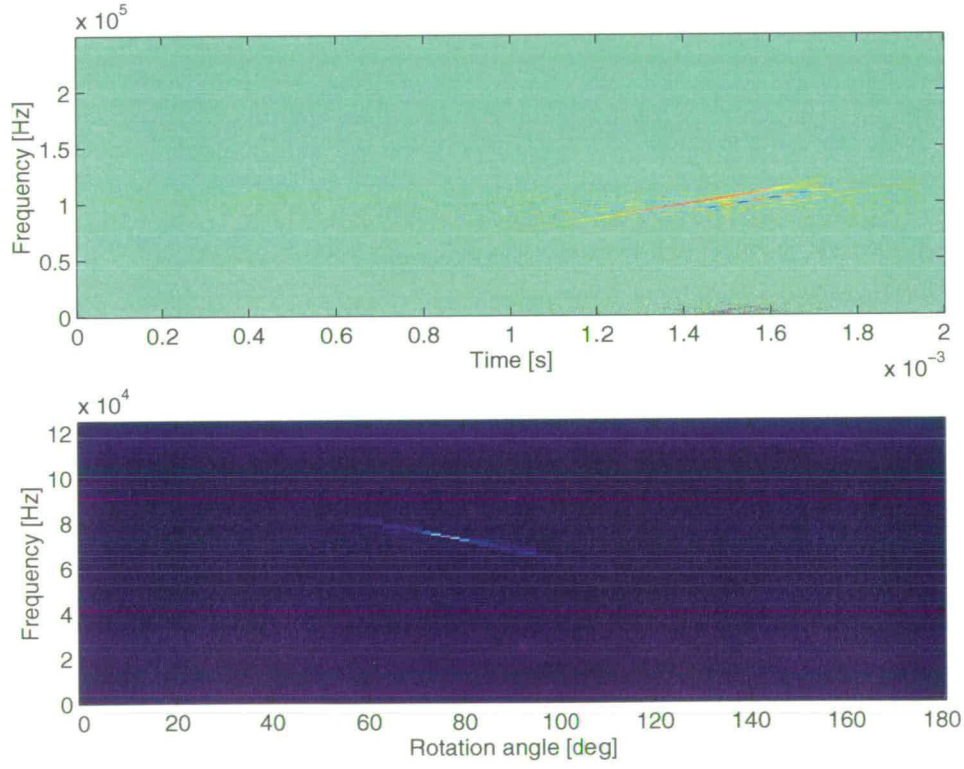


**Figure 4.8:** *The Wigner distribution and the Radon-Wigner transform of a linear chirp. The chirp has a bandwidth of 60kHz, a centerfrequency of 90kHz, and a duration of 1ms.*

### 4.2.3 The Short Time Fractional Fourier Transform

Before getting to the Short Time Fractional Fourier Transform (STFrFT) the Fractional Fourier Transform (FrFT) is introduced. The FrFT is a generalization of the Fourier Transform (FT) developed within the optics community which uses linear chirps as the basis functions, see [91] and [92].

The traditional FT decomposes signals by sinusoids, whereas the FrFT is a decomposition of signals by chirps. More formally the FrFT can be defined as follows, see [76]

$$FrFT^{\alpha}f(t) = \frac{e^{-j(\frac{1}{4}\pi\hat{\varphi}-\frac{1}{2}\varphi)}}{(2\pi|sin\varphi|)^{\frac{1}{2}}}e^{\frac{1}{2}jy^2\cot\varphi} \times \int_{-\infty}^{\infty} e^{-\frac{jty}{\sin\varphi}+\frac{1}{2}jt^2\cot\varphi}f(t)dt \tag{4.42}$$

where $\varphi = \frac{\pi\alpha}{2}$, $\hat{\varphi} = sgn(\varphi)$, $\alpha$ is the order of the transform. For $\alpha = \pm 1$ the FrFT reverts to the traditional FT, where t and y time and frequency respectively. The FrFT can defined in several different ways, see [84]. In the format the FrFT is presented above it can be viewed as four separate parts:

- A multiplication by a chirp in one domain
- A Fourier Transform
- A multiplication by a chirp in the transform domain
- A complex scaling.

The order of the transform is decided by the chirp properties, so that

$$\alpha = \frac{2}{\pi}\arctan(\frac{\delta f/\delta t}{2a}) \tag{4.43}$$

Where $a$ is the rate of change in the chirp, $\delta f$ is the frequency resolution equal to $f_s/N$, and $\delta t$ is the time resolution equal to $1/f_s$, $f_s$ is the sampling frequency and $N$ is the number of samples. The expression can be rewritten as

$$\alpha = \frac{2}{\pi} \arctan(\frac{fs^2/N}{2a}) \tag{4.44}$$

The rate of change in the chirp comes from the general expression for a linear chirp,

$$g(t) = e^{j\phi} = e^{j(at^2+bt+c)} \tag{4.45}$$

with the instantaneous frequency

$$\phi' = 2at + b \tag{4.46}$$

which is the derivative of the phase for a linear chirp.

In the case of active sonar data, the wave form is known and the transform order $\alpha$, can be given. However, if when the chirp properties are not known, $\alpha$ has to be optimized. Consider Figures 4.9,and 4.10 where the FrFT has been calculated for $\alpha$ ranging from -1 to 1. At $\alpha = \pm$ 0.997 one can see that the pulse compression is the highest and therefore this is the optimum transform order for this specific chirp.

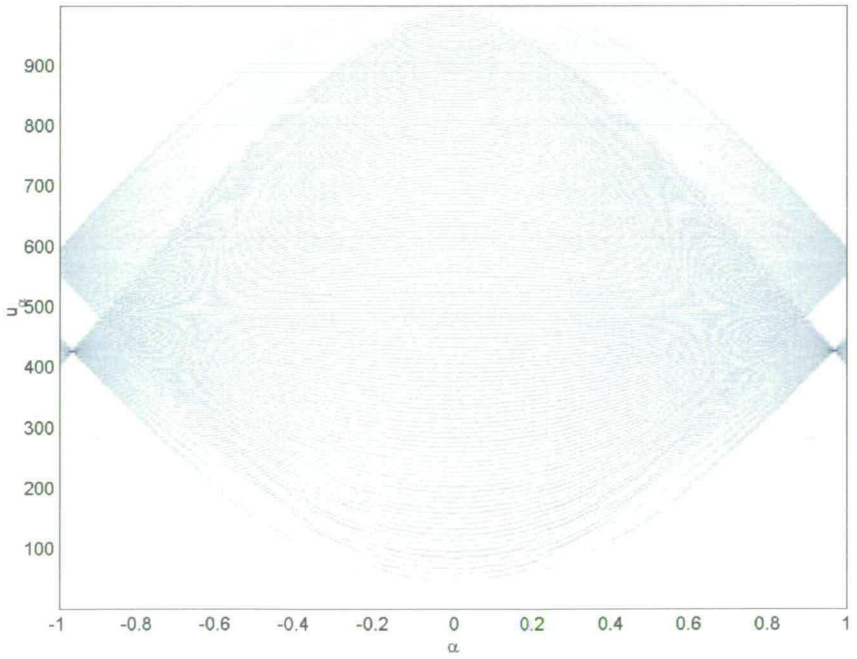


**Figure 4.9:** *Verifying the optimum choice of $\alpha$, values going from -1 to 1, highest level of pulse compression at $\alpha = 0.997$ and -0.997*

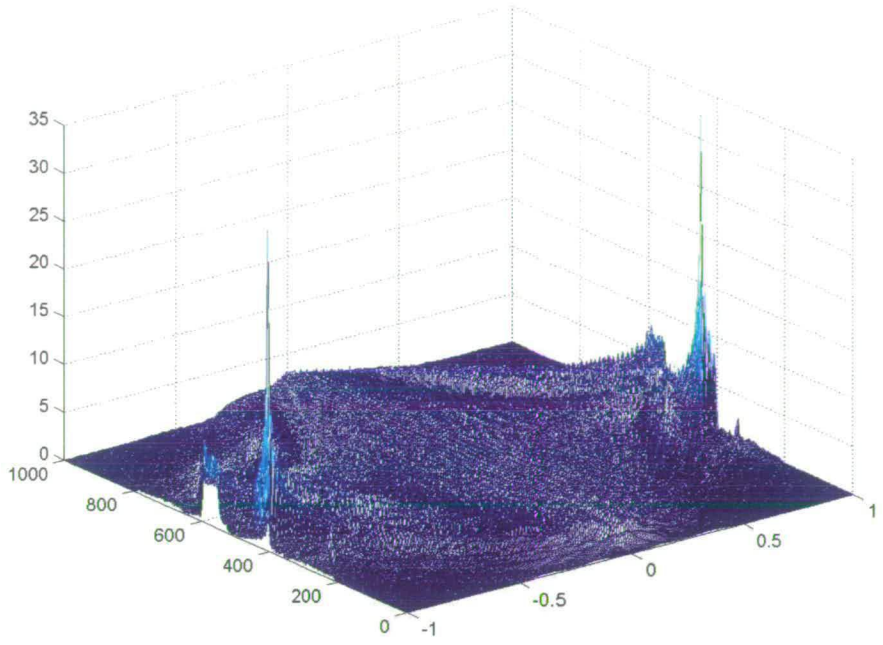


**Figure 4.10:** *The optimum choice of $\alpha$, same as in Figure 4.9 but a different view. Note the ridge that goes in an arch over the entire transform plane and has its peaks at the points $\alpha = 0.997$ and -0.997*

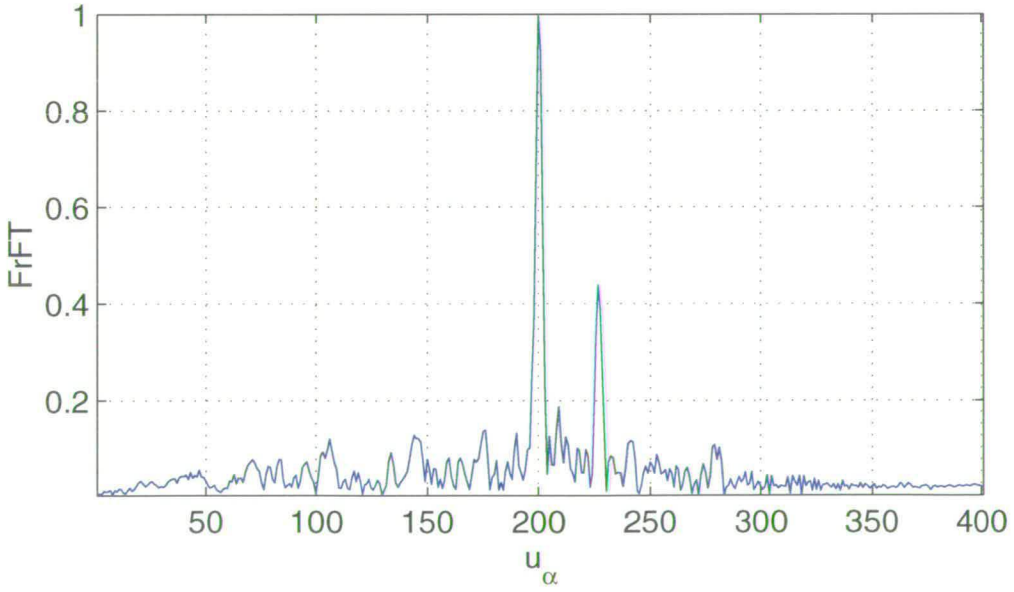


**Figure 4.11:** *The FrFT applied to a linear chirp. The chirp has a bandwidth of 12kHz, a centerfrequency of 6.25kHz, and a duration of 10ms. The y-axis is the normalised amplitude of the FrFT and the x-axis is the fractional frequency axis.*

In Figure 4.11 the FrFT of a linear chirp from the SAS sea trial is displayed. It is worth noting the similarity between the Figures 4.7 and 4.11, the RWT of a signal at the optimum rotational angle is equal to the FrFT of the signal. The differences in the images are due to the differences in the implementation of the FrFT, also see [93].

The FrFT, as with the FT, does not provide any information about time, only frequency. So there is no way of telling when an event including certain frequencies occur. The solution to this problem is the development of a Windowed FrFT [94] or the STFrFT, and it can be calculated in the same way as the Short Time Fourier Transform (STFT), or the spectrogram. The definiton for the spectrogram is shown in Section 4.2.1. Keeping the same approach in mind the STFrFT is defined as a windowed version of the FrFT, using a window function $h(t)$ to produce a modified signal $f_t(\tau)$

$$f_t(\tau) = f(\tau)h(\tau - t)$$

Just as before in Section 4.2.1 the modified signal is a function of two time variables, the fixed time $t$, and the running time, the timelag $\tau$. The window function is chosen to leave the signal as intact as possible around time $t$ without introducing artefacts to the transform, and to suppress the signal outside the window.

$$h(t) = \begin{cases} 1 & |t| < \tau/2 \\ 0 & |t| > \tau/2 \end{cases}$$

This gives the modified signal,

$$f_t(\tau) = \begin{cases} f(t) & |t| < \tau/2 \\ 0 & |t| > \tau/2 \end{cases}$$

Since the modified signal contains the original signal around time $t$, the FrFT will reflect the distribution around that time,

$$FrFT^{\alpha}f(t) = \frac{e^{-j(\frac{1}{4}\pi\hat{\varphi}-\frac{1}{2}\varphi)}}{(2\pi|sin\varphi|)^{\frac{1}{2}}} e^{\frac{1}{2}jy^2\cot\varphi} \times \int_{-\infty}^{\infty} e^{-\frac{jty}{\sin\varphi}+\frac{1}{2}jt^2\cot\varphi} f_t(\tau)d\tau \quad (4.47)$$

So that each $FrFT_t^{\alpha}f_t(\tau)$ makes up a slice in the time frequency display at time $t$. see Figure 4.12, for both the STFrFT and the spectrogram.

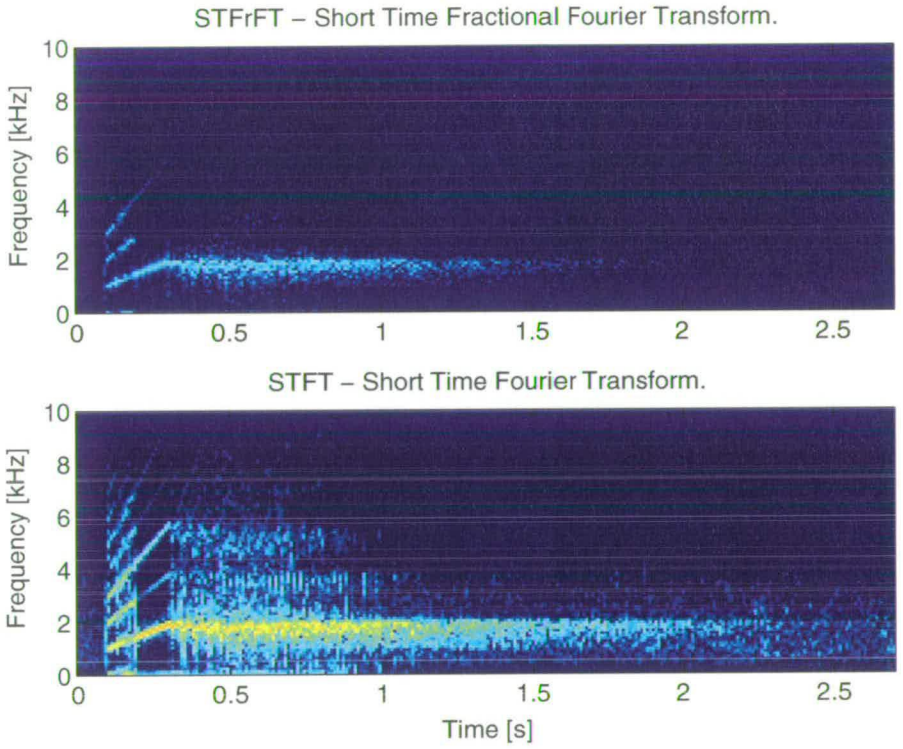


**Figure 4.12:** *STFrFT compared to the STFT of a sonar signal, a linear chirp with a bandwidth of 1kHz and a center frequency of 1.5kHz, and a duration of 200ms*

## 4.3 Assessing Image quality in the time/frequency displays

In [82] there are a number of image quality measures proposed. The problem that is addressed in this section is posed differently so they are not necessarily applicable as they are. The aim for the measures in [82] is quantifying image quality degradation due to compression of the images, whereas our aim is quantifying the advantage one transform method has over another.

A measure that aims at quantifying the focus of the image or rather how focused the structures

in the image are, is the $Q - value$, see Figure 4.13 for an explanation of the definition, Figures 4.14, and 4.15 for an example of how it works. The idea is, the better the transform behave, the more localized in time and frequency, the structures are. In this case the structures are mostly lines or curves in a typical time/frequency display, see the right panel in Figure 4.1, and Figure 4.15. The $Q - value$ is defined below as,

$$Q = 10 \log_{10} \frac{H}{W}, \tag{4.48}$$

where $H$ is the height of the peak over the noise floor, and $W$ is the width of the peak at the noise floor.

The $Q$ value springs from a number of different uses and different areas, physics, electronics and indeed signal processing. Within the field of optics, it is a quotient between energy contained in a cavity and the energy that dissipates from it. see [95]. That means that the $Q$ value is a quality measure of the cavity in question.

In spectral analysis it is ratio of the mean squared over the variance of the frequency estimate, see [69]. And it is defined as follows,

$$Q_A = \frac{E[P_{xx}(f)]^2}{var[P^{xx}(f)]}, \tag{4.49}$$

where $P(f)$ is the power spectrum. The $Q$ values is defined for each frequency, $f$.

In electronics the Q-value is a goodness measure for filters and such, where it is defined as a quotient between the center frequency and width of the frequency response of the filter at the -3dB level,

$$Q_B = \frac{f_0}{\delta f}, \tag{4.50}$$

where $f_0$ is the center frequency and $\delta f$ is the -3dB bandwidth, see [96].

The two curves in Figure 4.14 correspond to the amount of pulse compression a transform gives. The larger the pulse compression the larger the $Q - value$, so for the curves in the Figure we

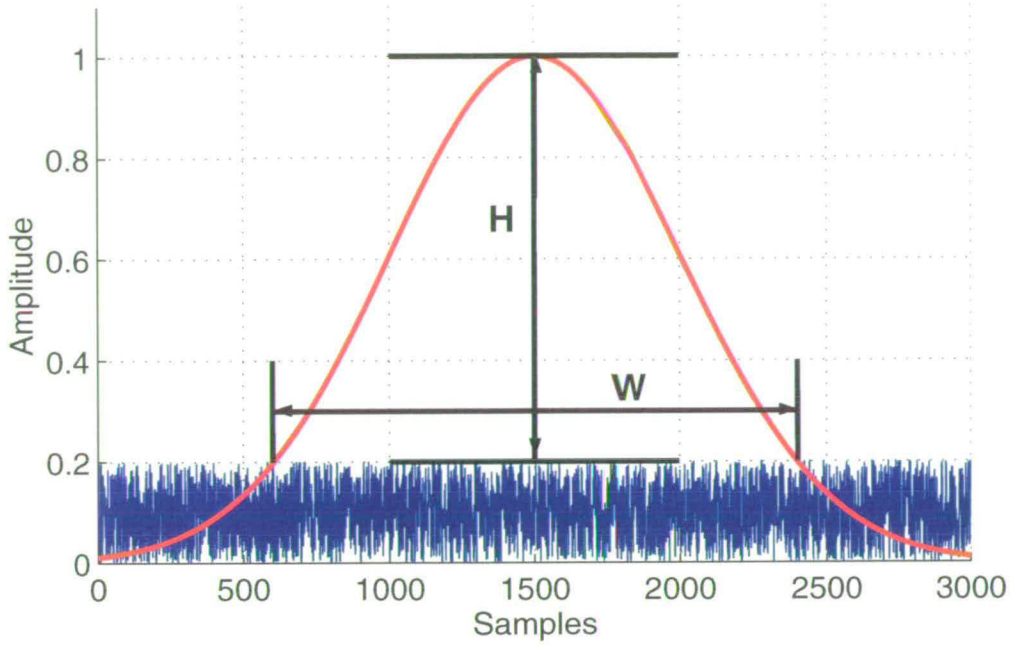


**Figure 4.13:** *The Q-value is defined as the quotient between the height and width of the peak over the noise floor.*

have $Q_{black} >> Q_{red}$

Note that neither the lines in Figure 4.14 or the two curves in Figure 4.15 are results from actual transforms. They are Gaussian probability density functions and are used to illustrate the the principle of more focused images and more localised transforms. The images in Figure 4.15 were produced taking the Gaussian curves in Figure 4.14 and moving their center of gravity along a sinusoid. Then these shifted Gaussians were placed next to each other to form an image that looks like a time/frequency plot. This was done because it is easy to have complete control over the appearence of a Gaussian, and it gives an idea of what kind of difference a more focused image can make.

Another measure is the average distance $AD$ as defined below. The expression has to be slightly rewritten, from the original in [82] where it was used as a measure of image dissimilarity between an original image and a compressed version, to its current form below

$$AD = \frac{\sum_i \sum_j \|R_{ij} - S_{ij}\|}{Nrows \times Ncols} \tag{4.51}$$

Where i = 1...Ncols, and j = 1...Nrows. The matrices are equally sized Ncols x Nrows. R

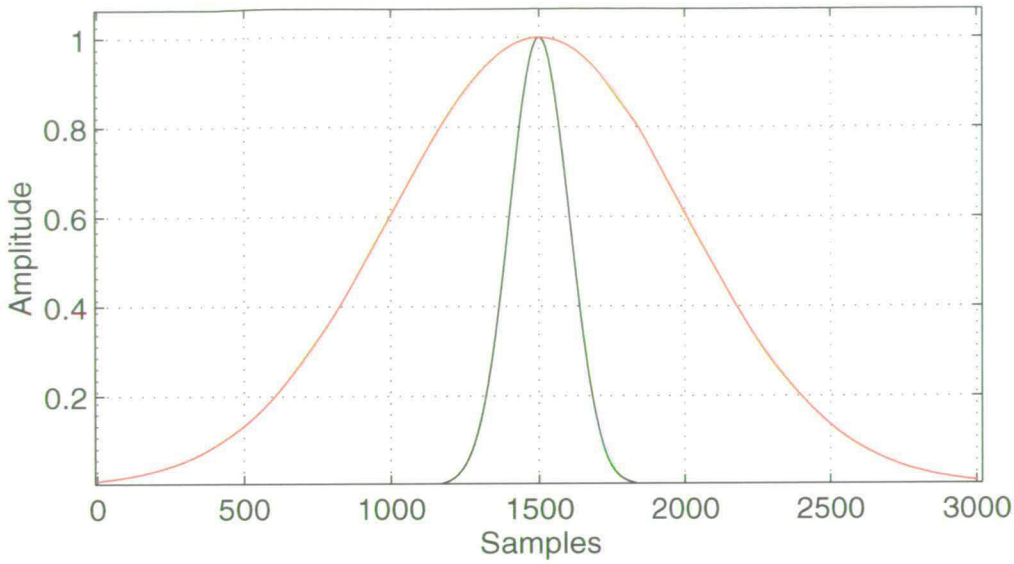


**Figure 4.14:** *Comparing the Q-values for two different transforms to give a measure of focus. For the curves in the figure we have $Q_{black} >> Q_{red}$, so the $Q - value$ is a lot higher for the black curve, ie. it is more focused.*

is the time frequency surface made with the transform of interest on real data and S is the same transform applied to synthetic data. The synthetic data is made out of synthetic chirps with arrival times and amplitudes matching the real chirps in the real data, with white noise added to avoid effects from the transforms being applied to zero vectors and also giving a fairer comparison between the transforms, see Figure 4.18.

A second measure that is applicable in this context is a measure based on the dynamics in the transform surface, a measure we choose to call $D$, for Dynamics.

$$D = 10 \log_{10} \frac{max(R)}{min(R)} \tag{4.52}$$

where R is the matrix containing the transform surface, be it performed on real or synthetic data, and $max(R)$, $min(R)$ refers to the largest and smallest value of $R$.

Yet another way of quantifying the quality of an time/frequency surface is to look at the condition number of the transform matrix. That is, the ratio of largest to smallest singular value of the matrix. The measure is defined as the quotient between the condition number for the real

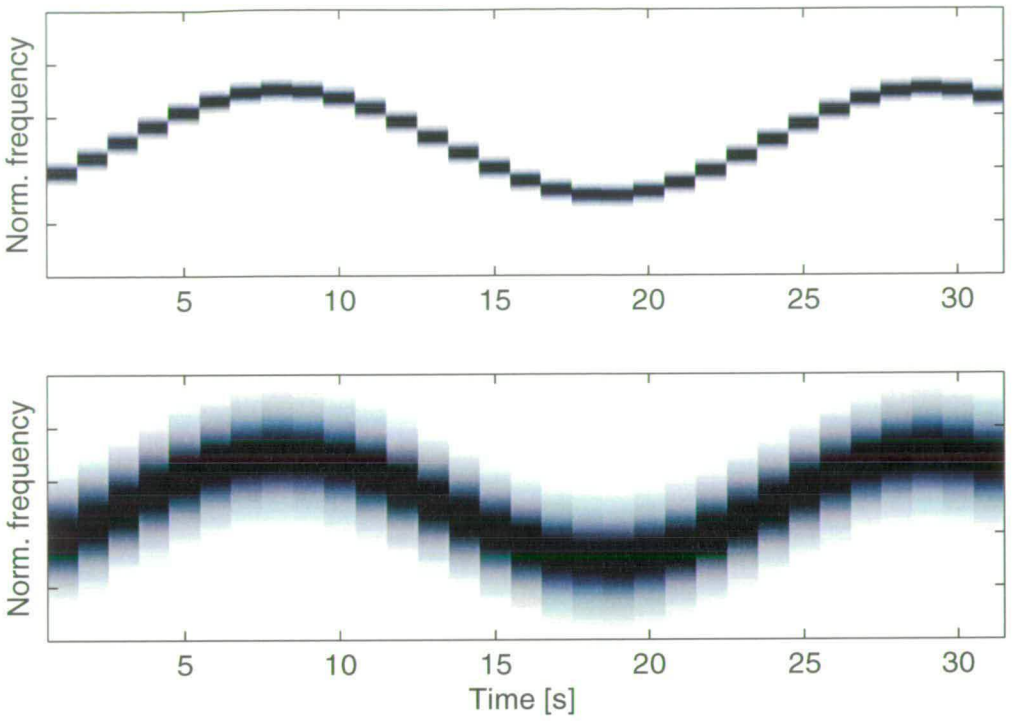


**Figure 4.15:** *Comparing the Q-values for two different transforms to give a measure of focus. The upper panel is made stacking displaced versions of the black curve in Figure 4.14 adjacent to one aother, and the lower one is made in the same way, using the red curve in Figure 4.14.*

and synthetic matrices, as below

$$Cq = 10Log_{10}\frac{cond(R)}{cond(S)} \tag{4.53}$$

where R and S are the real and synthetic transform matrices. Knowing the condition number of the matrix tells us something about the amount of information in the matrix, and also provides a metric that makes it possible to compare the real and synthetic matrices. The idea is that the better the transform behaves on real data the smaller the difference between the real and synthetic matrices is, and also a transform that produces an image with as few artefacts and as little content, still producing the desired features, as possible will have a lower condition number. If the data one is analysing is strictly tonals without much variations, the image matrix produced should be almost be a linear combination of the first column in the transform matrix, hence giving the lower condition number.

### 4.3.1 Applying the Image quality measures

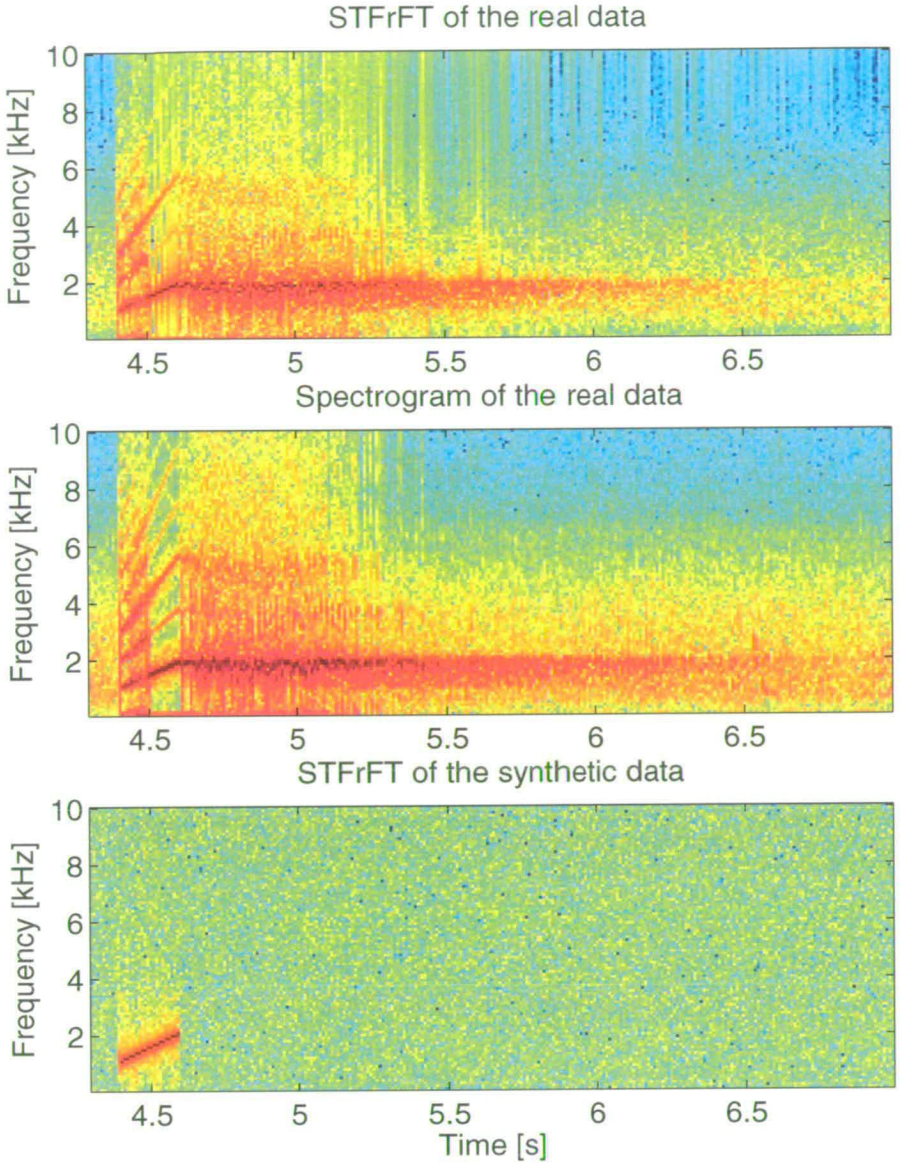


**Figure 4.16:** *STFrFT compared to the STFT of a sonar signal, a linear chirp with a bandwidth of 1kHz and a center frequency of 1.5kHz, and a duration of 200ms. The topmost panel displays the STFrFT, the middle one displays the STFT, both on real data. The bottom panel displays the STFrFT on the synthetic data.*

The data that is chosen for the time/frequency image quality assessment is two different sets. The first one is from the reverberation measurements done at Utö in 2001, see Section 2.5.3. The second data set is the SAS experiment from 1999, see Section 2.5.4. The reason for picking

these two data sets is that they represent two different chirps, the first one has a very low rate of change and a rather low bandwidth, and the second set consists of chirps with a high rate of change and large bandwidth.

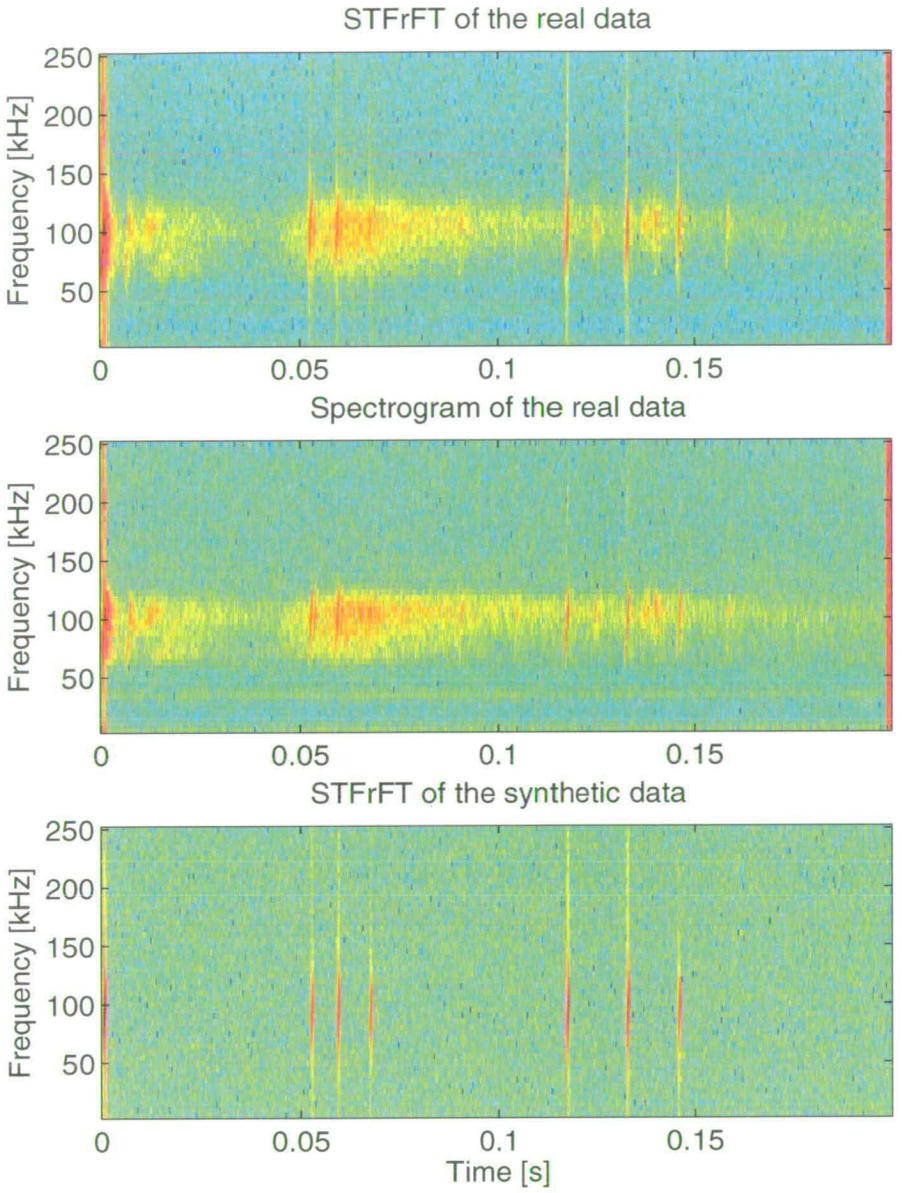


**Figure 4.17:** *STFrFT compared to the STFT of a sonar signal, a linear chirp with a bandwidth of 60kHz and a center frequency of 90kHz, and a duration of 1ms. The topmost panel displays the STFrFT, the middle one displays the STFT, both on real data. The bottom panel displays the STFrFT opn the synthetic data. The first of the chirps is the transmitted pulse and the other 6 are returns from objects in the water volume.*

In Figure 4.16 the analysis results of a linear chirp with 1kHz bandwidth, a center frequency of 1.5kHz, and a duration of 200ms is presented. Both the STFT and the STFrFT are calculated. For the sake of comparison the window lengths are the same. There is no overlap between consecutive frames. A Hanning window is applied when estimating the spectrogram. In Figure 4.17 the panels display the STFrFT and the spectrogram on the other linear chirp. It has a 60kHz bandwidth, 90kHz center frequency, and 1ms duration.

It can be seen that the STFrFT images have slightly fewer artifacts introduced by the transform. The reverberation after the chirp has been transmitted can be seen in the spectrogram and it is spread over the entire frequency band of the chirp, whereas it is more suppressed in the STFrFT and less spread out in frequency. In Figure 4.17 where the data with the higher chirp rate is presented the advantage of the STFrFT is easier seen. However, with longer analysis windows the fractional transform produces even better localized peaks, but by using longer windows the time resolution suffers as with any short time transform.

The real matrices $R$ for both the STFrFT and the spectrogram can be seen in the two top most panels in Figures 4.16 and 4.17.

All the image quality measures mentioned above, in Section 4.3, are applied to the data. And the results for the Average distance, Dynamics, and the Condition quotient are presented in Table 4.1 Out of these three measures the $Ad$ and $Cq$ is done as comparisons between synthetic and real data. An example of the real and synthetic timeseries can be seen in Figure 4.18.

| The chirp with a bandwidth of 1kHz | | |
|---|---|---|
| Method / Measure | Spectrogram | STFrFT |
| Average distance (Ad) | $3.8 \times 10^{-3}$ | $3.7 \times 10^{-3}$ |
| Condition Quotient (Cq) | 12.0 | 18.2 |
| Dynamics (D) | 65.3 | 51.8 |
| **The chirp with a bandwidth of 60kHz** | | |
| Method / Measure | Spectrogram | STFrFT |
| Average distance (Ad) | $1.5 \times 10^{-3}$ | $5.3 \times 10^{-4}$ |
| Condition Quotient (Cq) | 2.73 | 2.38 |
| Dynamics (D) | 56.3 | 50.4 |

**Table 4.1:** *The results from applying the Ad, D, and Cq measures to the Utö and SAS trial data.*

One can see that the average distance measure is smaller for both surfaces produced using the

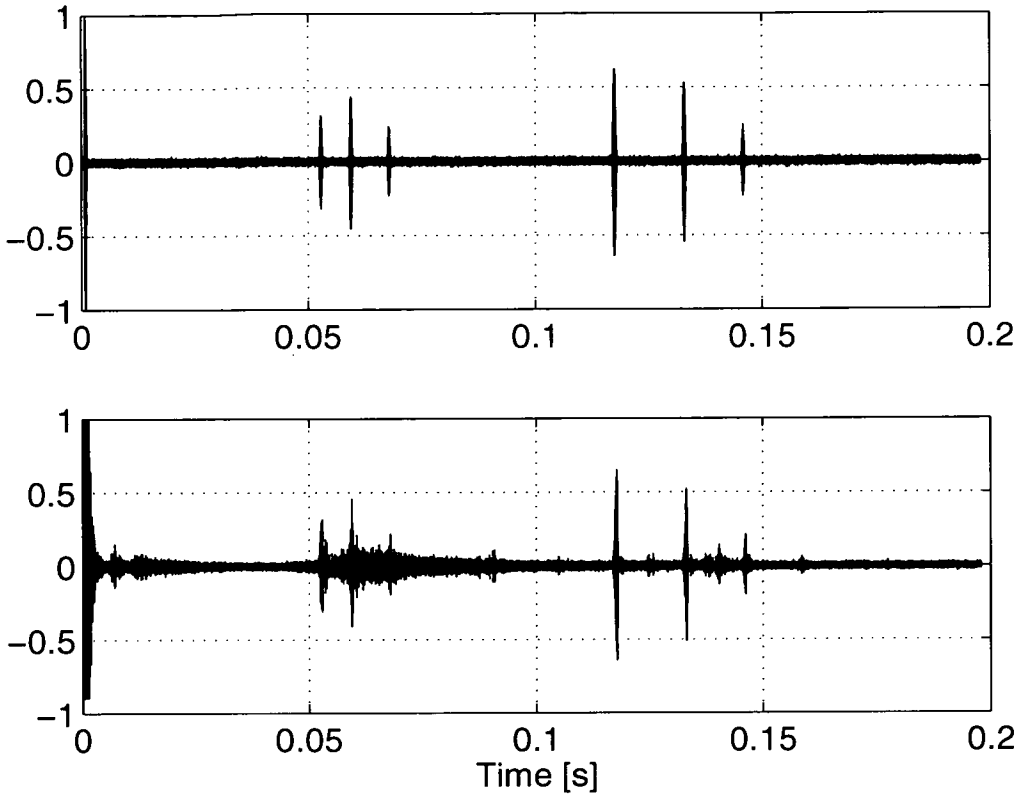


**Figure 4.18:** *The timeseries from the SAS experiment that the image quality measures are applied to. The synthetic time series is displayed in the top panel, and real time series in the bottom one.*

STFrFT, but the difference is smaller when the chirp rate is lower, as would be expected. The dynamics measure is greater for the images produced with the spectrogram.

The results from applying the condition number quotient are a bit inconclusive, for the chirp with the higher frequency the condition number is lower as it was expected since having a less cluttered image would produce a matrix with a smaller difference between largest and smallest eigenvalues. However in the case of the 1kHz chirp, although the image looks less cluttered the $Cq$ number is higher.

The Q-value has been applied to the higher chirp rate time frequency image. It was applied to the slices of Figure 4.17 where there is a clear return visible and the results can be seen in Figures 4.19, 4.20, 4.21, and 4.22 and Table 4.2.

In most cases the Q-value is larger for the STFrFT than for the STFT. However, there are cases when the STFT outperforms the STFrFT. In Figure 4.22 the Q-value is larger for the STFT.

| Time [s] | (Peak number) | $Q_2$-value, Spectrogram | $Q_1$-value, STFrFT |
|---|---|---|---|
| 0.0519 | (1) | 0.16 | 1.09 |
| 0.0588 | (2) | 0.13 | 1.59 |
| 0.0669 | (3) | 0.14 | 1.10 |
| 0.1163 | (4) | 0.11 | 0.28 |
| 0.1319 | (5) | 0.37 | 0.99 |
| 0.1450 | (6) | 0.12 | 1.20 |
| 0.1450 | (9) | 0.11 | 0.09 |

**Table 4.2:** *The results from applying the Q-value to the SAS trial data.*

This is due to the fact that, it is the STFrFT for the lower half of the STFrFT spectrum that is compared to the STFT. Where the FT is entirely symmetrical the FrFT is not so it is important to use the right part of the image one wants to make the comparison over.

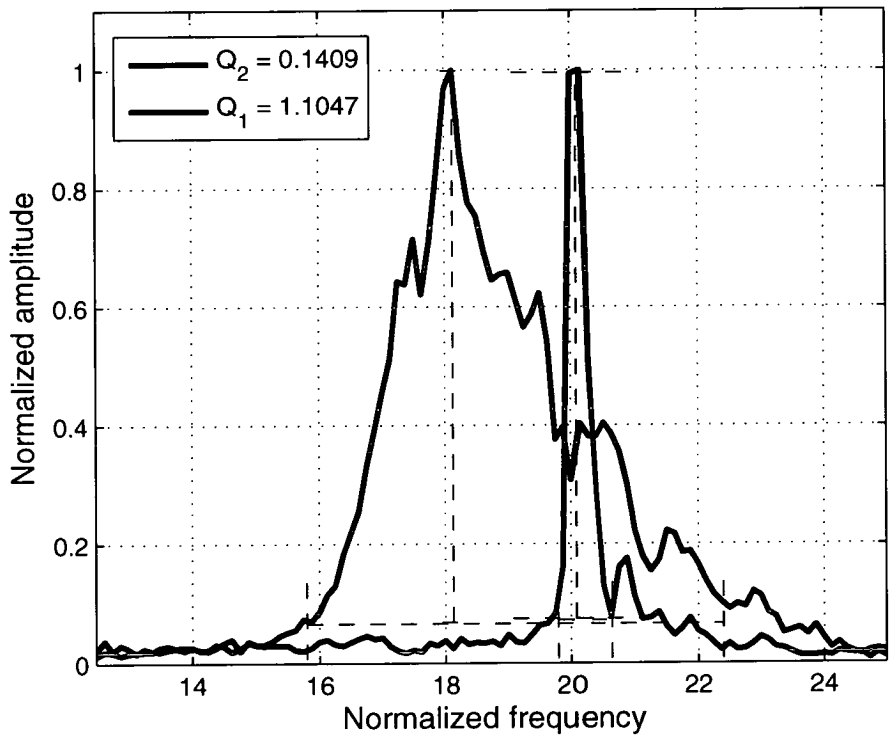


**Figure 4.19:** *The Q-value for peak number one.*

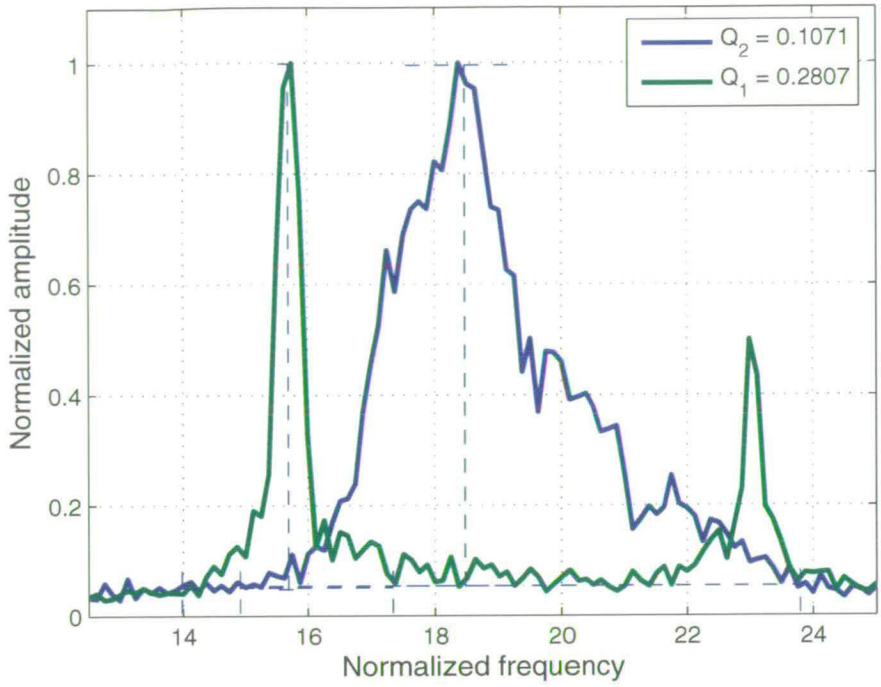


**Figure 4.20:** *The Q-value for peak number two.*

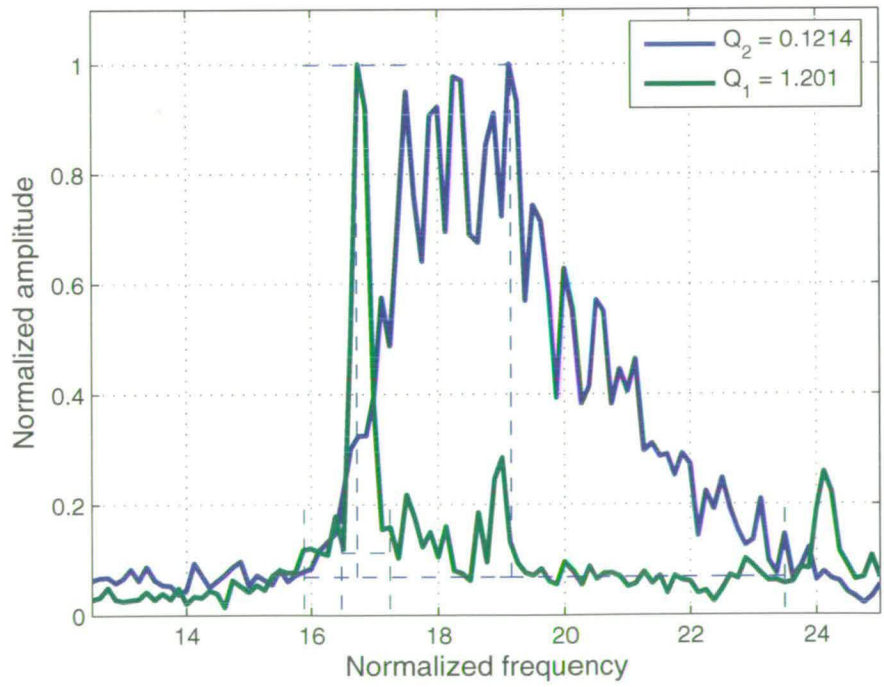


**Figure 4.21:** *The Q-value for peak number three.*

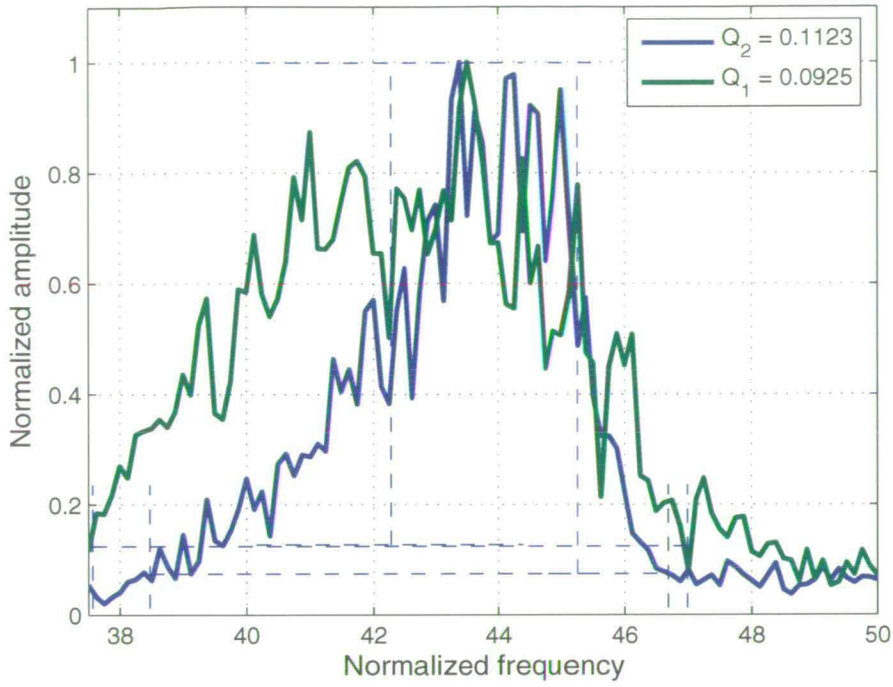


**Figure 4.22:** *The Q-value for peak number nine.*

## 4.4 Conclusions

Although mentioning a number of time/frequency analysis methods and pointing out their possible uses the time-frequency analysis in this chapter has been done using the traditional Short Time Fourier Transform, and the Short Time Fractional Fourier Transform. Performance comparisons are made between the two different transforms and it is shown that matching the transform to the nature of the data improves the quality of the time-frequency images. I.e. if the data consists of chirps then it is benefitial to decompose the timeseries using a chirp as a basis function rather then using sinusoids (the FrFT and FT respectively). It is important to note that when the chirp rate approaches zero (a sinusoid is a chirp with zero chirp rate) there is no advantages to use the STFrFT over the STFT to produce time/frequency images.

In the second section of this chapter a number of different metrics for image quality assessment are defined and evaluated using the time/frequency images produced with the aforementioned transform techniques.

The results show that the Fractional Fourier Transforms (FrFT) has advantages over the tradi-

tional short time Fourier transform, as they make use of the knowledge of the transmitted wave form.

Of all the implemented image quality measures ($Q - value, Ad, D, Cq$) the Q-value seems to be most consistent with expectation and proves reliable for the presented images. The other metrics have their use but should be applied with a bit of caution since the results found were inconclusive.

# Chapter 5
# Data Conditioning Methods applied to Sonar Data

Recorded sonar data is processed in many different ways to extract information from the signals. Trying to separate different sources of sound, or identifying vessels by examining the emitted tonals, is often done in the time-frequency domain, i.e. using the lofargram. Narrowband components appear as long-time duration lines and transients as broadband short-time duration pulses. The transients may be generated by the same source as the lines, from another near distance source, or be a part of the additional ambient noise. If the transients are considered as disturbances they are preferably suppressed in the Lofargram. Furthermore, conventional spectral estimation methods are based on the assumption of stationarity. If transients and narrowband components are mixed in the estimation window, distortion and masking of the lines may occur especially in the low SNR regimes. If suppression or removal of transients is possible, the source signal identification of spectral lines may be improved.

In this section two methods of improving sonar data are applied. The first one is a method of stepwise outlier rejection. The rejection of outliers is done in a bispectral domain, and incorporates different methods of replacing the rejected data.

A second method for improving the visibility of relevant spectral lines in a lofargram is also performed, instead of trying to get rid of data by regarding parts as outliers the signal is decomposed by using the method of empirical decomposition.

## 5.1 Outlier rejection

The stepwise outlier rejection method is used to detect and remove the data containing disturbances. This method of rejection of outliers has been used earlier for sonar data conditioning, see [59],[47], [97], and [48]. One disadvantage with the outlier rejection method is the requirement of relatively long data sets for significant estimates due to the properties in the bicoherence domain, see [98]. The procedure attempts to identify any quadratic phase couplings (QPC) for

testing of independence of sources, therefore, the amount of the data used in the estimation should be kept intact if possible. Instead of reducing the data set by the removal of the rejection method, the data points that are rejected are replaced.

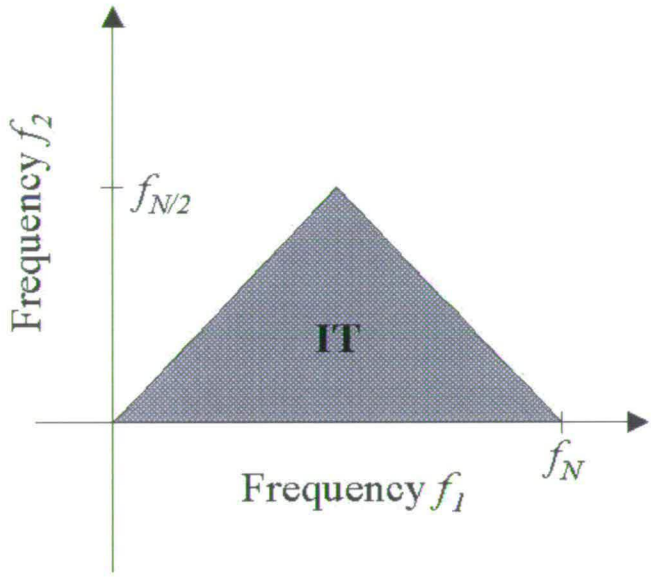


**Figure 5.1:** *The non-redundant part IT of the principal domain used for bispectral estimations. $f_N$ is the Nyquist frequency.*

The assumptions for the outlier rejection are; the estimates at each bifrequency are independent and if no transients are present in the estimation window the bispectral values follow a bivariate, normal distribution with unknown mean and variance, [99],[100].

The individual bispectrum $B(k,l)$ as defined in Equation (3.41) for each block $n = 1, 2, .., L$ is estimated, see [101]. Then the averaged bispectrum as defined in Equation (3.43) is estimated. The bispectral estimates are restricted to the non-redundant part of the inner triangle (IT),

$$f_1, f_2 \geq 0, f_1 \geq f_2, f_1 + f_2 \leq f_N$$

see [102] and Figure 5.1.

The difference for each frequency triplet from the global mean is compared with a statistical

table at the required significance level using a chi-squared test, [103]. If the value is larger than the threshold it is considered an outlier, and is rejected, removed from the bispectral estimate and replaced by a value depending on the method. The procedure is repeated until all values are considered as inliers to form the modified bicoherence used in the projection in Equation (3.43) to form a modified Lofargram, (see Section 2.6 for examples of Lofargrams). The different data replacement methods are summarised in Table 5.1. For estimation of passive underwater target features such as spectral lines in the Lofargram, it is important to use as significant values as possible. If the sonar data contain short-time disturbances, with unknown amplitude, arrival and duration time, the estimated spectra lines may be distorted. In this case it is difficult for a limited time and frequency resolution to achieve significant estimates. For reduction and removal of disturbances a stepwise outlier rejection algorithm is used, see [59].

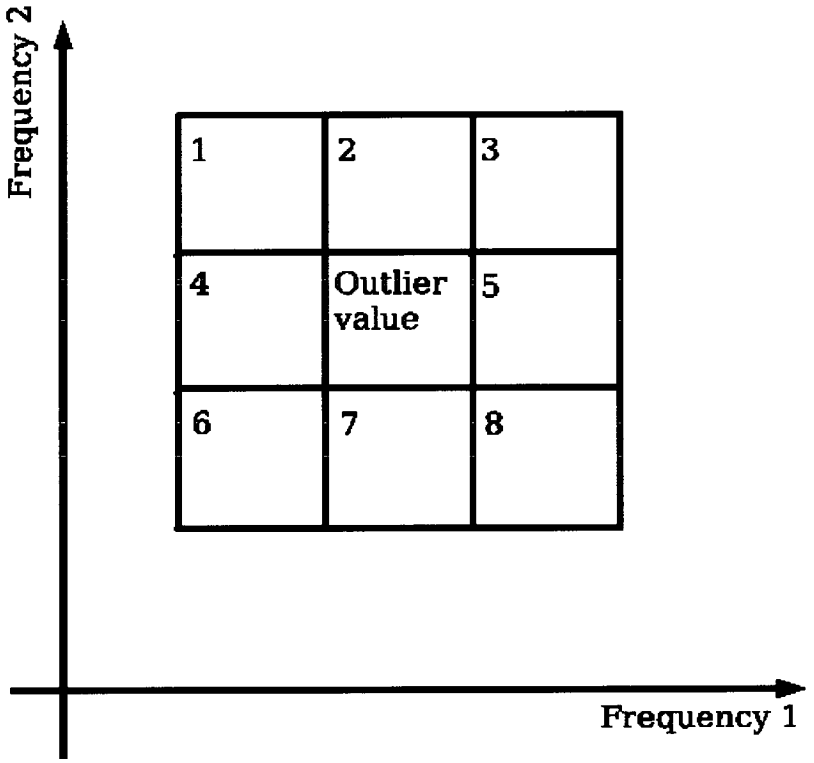


**Figure 5.2:** *The nearest neighbour (NN) of an outlier value in the bispectral domain used by method B, C, D and F.*

Method A is based on replacing each outlier value in the bispectrum domain by a zero. Instead of using zeros as replacement values, the other methods B to D and F use nearest neighbour

| Method | Replaced outlier value |
|---|---|
| A | Zero |
| B | Mean of nearest neighbour (NN) |
| C | Median of NN |
| D | Minimum of NN |
| E | Mean of a global slice |
| F | Mean of a NN slice |

**Table 5.1:** *Data replacement methods for replacing outlier values.*

(NN) values in the bispectrum domain as can be seen in Figure 5.2. Around the outlier value at a distance of at least one value from the bispectral manifolds, the number of nearest neighbour values are eight. For method B the mean of the eight values is estimated, for method C and D the median and the minimum value are estimated, respectively. Methods E and F use slices of the NN to replace the outlier value. The slices are produced by holding the value of frequency $f_1$ constant, i.e. producing a slice along the frequency $f_2$ axis. Three values in both directions, not counting the outlier value are used, by taking the mean of the six values. For method F the mean of all values along the slice is estimated.

### 5.1.1 Data analysis and results

The data set analysed in this section is the fibre glass boat trial as described in section 2.5.6. First the bicoherence for a segment 10-40 seconds of the data is shown in Figure 5.3. Then the same data is processed with the the outlier rejection method and the data considered as outliers is replaced using method D, see Table 5.1, and the result can be seen in Figure 5.4. It is clear to see that the number of visible points are reduced, and the ones remaining are the ones that are quadratically phase coupled. Amongst the coupled frequency pairs the most prominent ones are the ones that are related to the different rotating parts of the machinery, such as the engines crank shaft, primary drive, and the propeller.

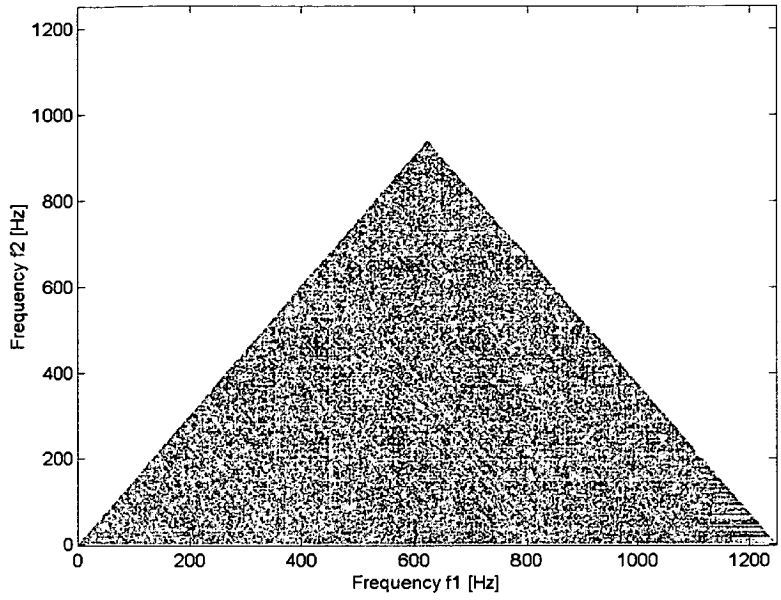


**Figure 5.3:** *Bicoherence estimated from the same data as in Figure 2.18.*

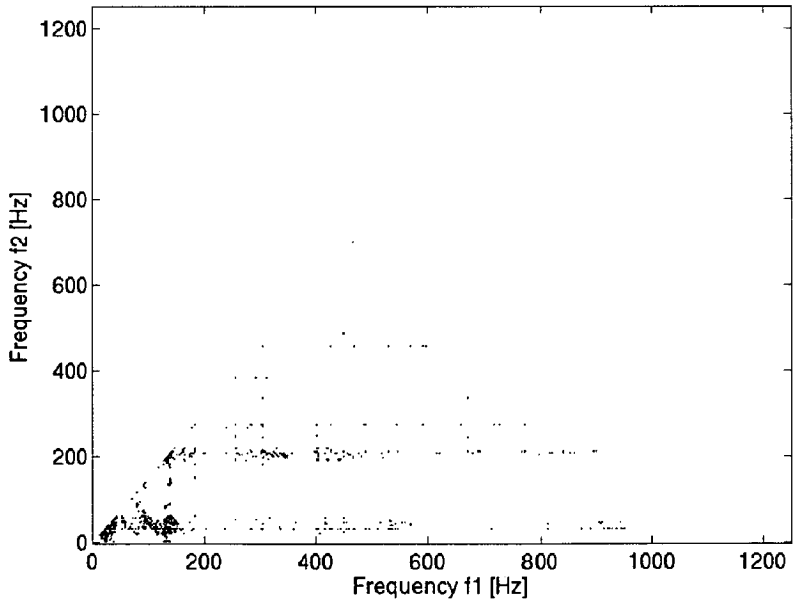


**Figure 5.4:** *Bicoherence with outlier rejection and data replacement method D applied for the same data as in Figure 5.3*

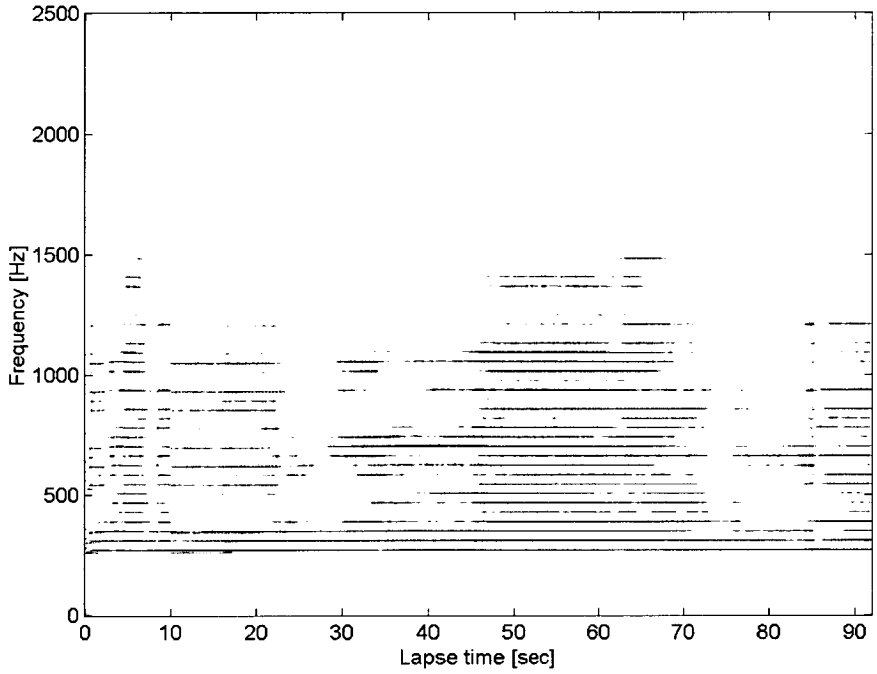


**Figure 5.5:** *Time-frequency display (Lofargram) by bicoherence projection with outlier rejection and method A.*

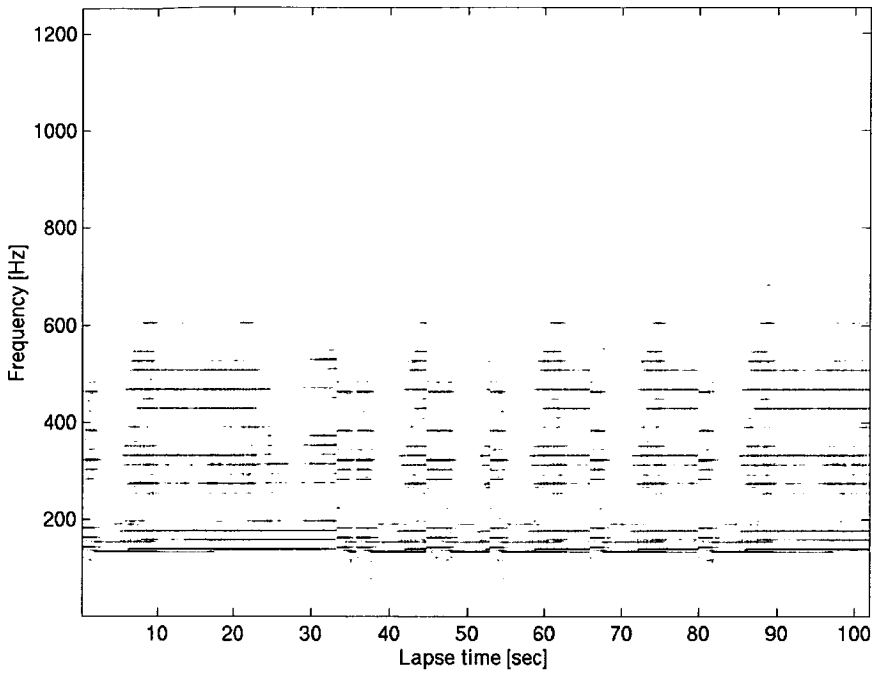


**Figure 5.6:** *Time-frequency display (Lofargram) by bicoherence projection with outlier rejection and method B.*

As can be seen in Section 2.5.6 the length of a complete recording is about 115 seconds in total. In the analysis the data was divided into time frames of 15 sec. Very little information about the source was found above 1 kHz so the data is lowpass filtered and down sampled to 2.5 kHz. The frames of data are segmented into 36 non-overlapping blocks of 1024 samples each. The estimation is performed in 0.2 sec steps of the time frame. The estimate is from the same data as in Section 2.6.

The Lofargram in Figures 5.5, 5.6, 5.7, and 5.8, display the final result with data modified by outlier rejection and data replacement methods A, B, D and E. The lofargrams are estimated by a projection of the bicoherence, so that the frequencies are averaged over $f_1$ and then as usual in lofargram analysis, the slices produced are placed adjacent to one and other to form a time/frequency display. It can be seen that the spectral lines are more stable for the Lofargram based on the modified data throughout the entire run. It is however not very easy to distinguish between the different data replacement methods. Although it seems that the methods B,D and E allow more variation in the spectral lines whereas method A appears to give a thin sharp line in the lofagram. Depending on the data at hand and the goal for the conditioning there is reason to use either methods A or B (or depending on the width of the peaks in the bicoherence surface

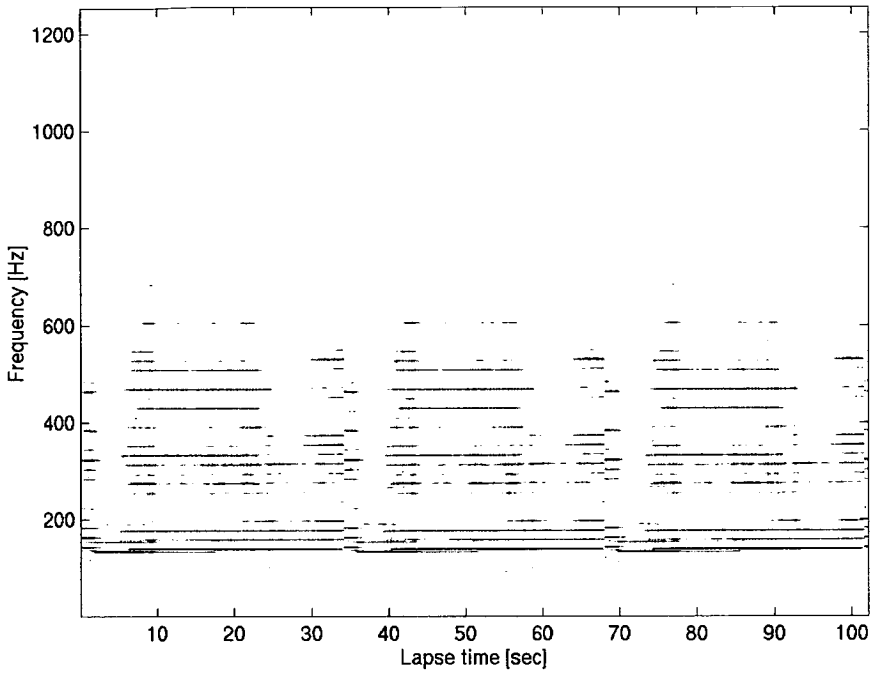


**Figure 5.7:** *Time-frequency display (Lofargram) by bicoherence projection with outlier rejection and method D.*

the methods C, D, and E might be good candidates).

Method A would be best suited when the target that appear as spectral lines in the lofargram holds a constant speed (i.e. constant engine and propeller revolutions). Conversely, method B would work better when it is known that the speed of the target might not be constant and the slight variations in the spectral lines are not to be considered as outliers. All of the methods of replacing data are done after removing the outliers, so the methods B-E (based on the value of the nearest neighbours) affect the distribution of the bicoherence surface the least.

## 5.2 Empirical Mode Decomposition

A new way of analysing non stationary data has recently been developed by [104], termed Empirical Mode Decomposition (EMD). It is a purely data driven analysis method that decomposes the signal into so called Intrinsic Mode Functions (IMF's). The introduction of EMD has aided the understanding of instantaneous frequency concept, and hence the well known Hilbert Transform is refered to as the HH-Transform (HHT) in many recent publications (the Hilbert

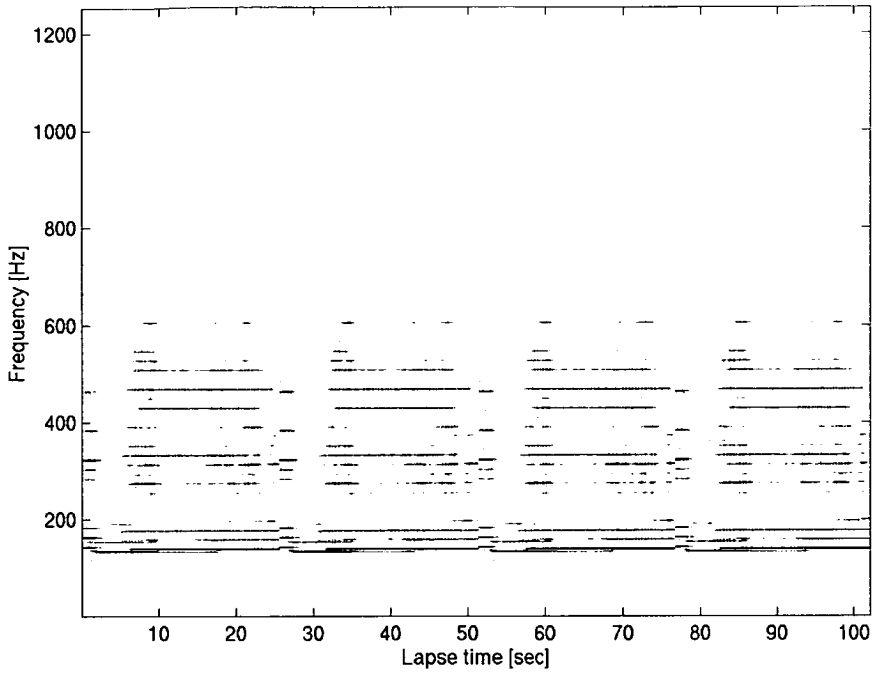


**Figure 5.8:** *Time-frequency display (Lofargram) by bicoherence projection with outlier rejection and method E.*

Transform a method for obtaining the analytical signal). Further work on the algorithm has been done in [105], and [106] to name a few. The algorithm has been applied to different fields such as Geophysics in [107], and another application for the EMD algorithm is naturally sonar data, which follows from the previous attempts of applying different time/frequency methods to analyse the data. In this section the EMD algorithm is explained and then applied to a sonar data set, see section 2.5.6. The data set has been analysed using the spectrogram and also been tested for various statistical properties, in chapter 3.

### 5.2.1 The EMD Algorithm

For a signal $x(t)$ the EMD algorithm can be summarised into five steps as follows

1. Identifying all extrema of $x(t)$.
2. Interpolation between the maxima and minima, and retrieving a function of the envelope $e_{min}(t)$ and $e_{max}(t)$.
3. Computing the average of the envelope function $m(t) = (e_{min}(t) + e_{max}(t))/2$

4. The extraction of the detail from the data $d(t) = x(t) - m(t)$

5. Iterate on the residual $m(t)$ until it is sufficiently small

The first step from the summary is rather self explanatory and is left as it is. The following steps might however, require further explanation. The second step of the summary is the definition of the minimum and maximum functions of the envelope, so that $(e_{min}(t)$ describes all the minima from the original time series and, $e_{max}(t)$ describes the maxima from the time-series. Then the third step is simply the mean of these two envelope functions. The fourth step of extracting the so called detail from the data is a matter of finding the most rapidly oscillating modes from the data and extracting it, thus creating the IMF's. Then the fifth and final step is to iterate the entire process until the residual is sufficiently small. Small enough in this case means leaving a monotonic function as a residue that no more IMF's can be extracted from. This is also known as the sifting process. The entire signal can be recovered from the IMF's and the residual such that

$$x(t) = \sum_{i=1}^{n} d_i + r_n \tag{5.1}$$

The decomposition of the signal into empirical modes is made under the following three assumptions:

1. The signal has at least two extrema, a minimum and a maximum

2. The Characteristic time scale of the signal is defined by the time between the extrema

3. If data is completely devoid of extrema but contain only inflection points then it can be differentiated to reveal the extrema, thus it needs to be n times differentiable.

The IMF's satisfy the following two conditions:

1. In the whole data set the number of extrema and zero crossings is equal or differ at most by one.

2. At any point, the mean value of the envelope defined by the local maxima and, the envelope defined by the local minima, is equal to zero.

### 5.2.2 Data analysis and results

The fibre glass boat trial data is analysed to see if the EMD algorithm can reveal more information from the data than the traditional time/frequency methods. This data set is rich in harmonics from the rotating parts of the boats engine and drive, and it is interesting to find out if the EMD algorithm can separate the different components of the signal and maybe relate the different tonals to each other, see Figure 5.9 for the Power spectrum of the fibre glass boat data up to 300 Hz, and Table 5.2 for the frequencies that can be seen as peaks. The fibre glass boat data also has some outliers in it according to the data conditioning done in the previous section. These outliers are usually spiky noise, and broad band in their nature. This leads to a degradation of the spectral lines in the spectrogram or the lofargram. If these noise components can be caught in one or several of the imfs then the lofargram can be improved by only using those imfs that include the harmonics.

The data analysis in this section falls into three parts:

- Producing Hilbert spectra of the data.
- Comparing the intrinsic mode functions and their spectra with the ordinary power spectrum of the signal.
- Conditiong lofagrams by using the intrinsic mode functions as an input to making the lofargram.

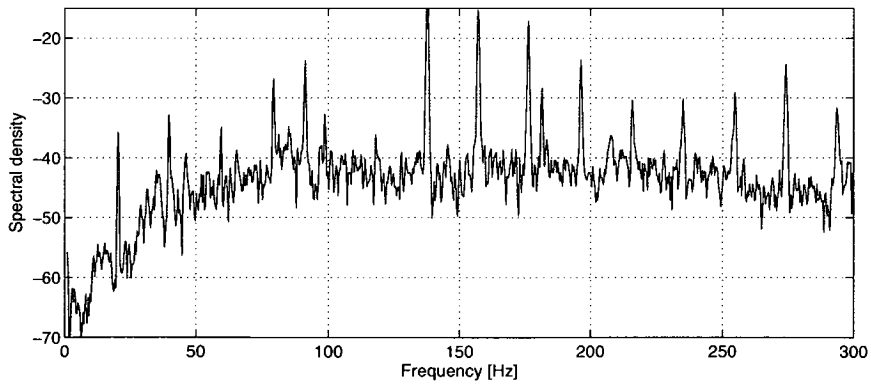


**Figure 5.9:** *The Power spectrum of the fibre glass boat trial data, for 65-80 seconds, including CPA. Also see Table 5.2 where the frequency peaks are noted.*

| The frequncy peaks from Figure 5.9 in Hz. | | | | |
|---|---|---|---|---|
| 19.5 | 64.4 | 117.1 | 180.8 | 234.2 |
| 39.5 | 78.2 | 136.6 | 195.3 | 253.7 |
| 45.4 | 90.4 | 156.1 | 206.9 | 273.1 |
| 58.7 | 97.7 | 175.8 | 214.7 | 296.2 |

**Table 5.2:** *The indentified frequency peaks in Figure 5.9.*

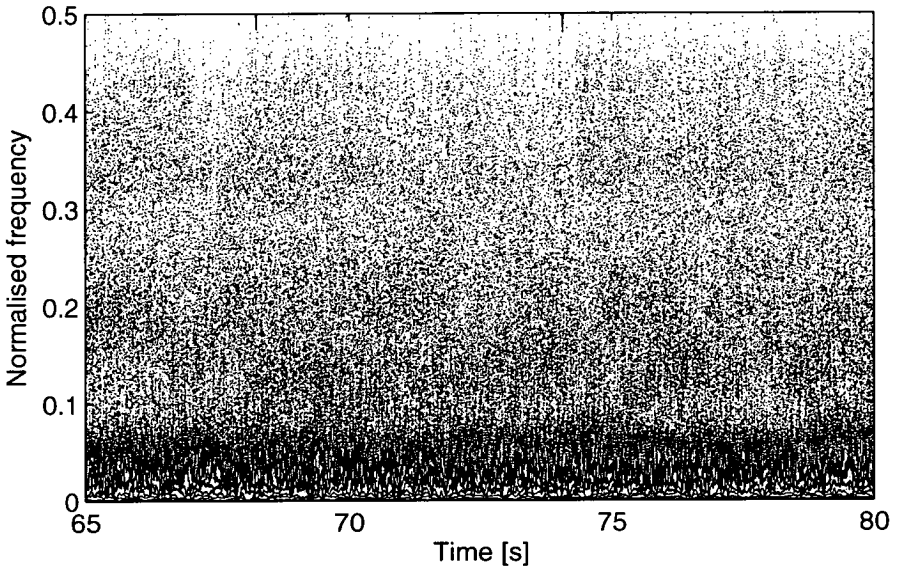


**Figure 5.10:** *The Hilbert spectrum of the fibre glass boat trial data for the part that includes CPA, 65-80 seconds*

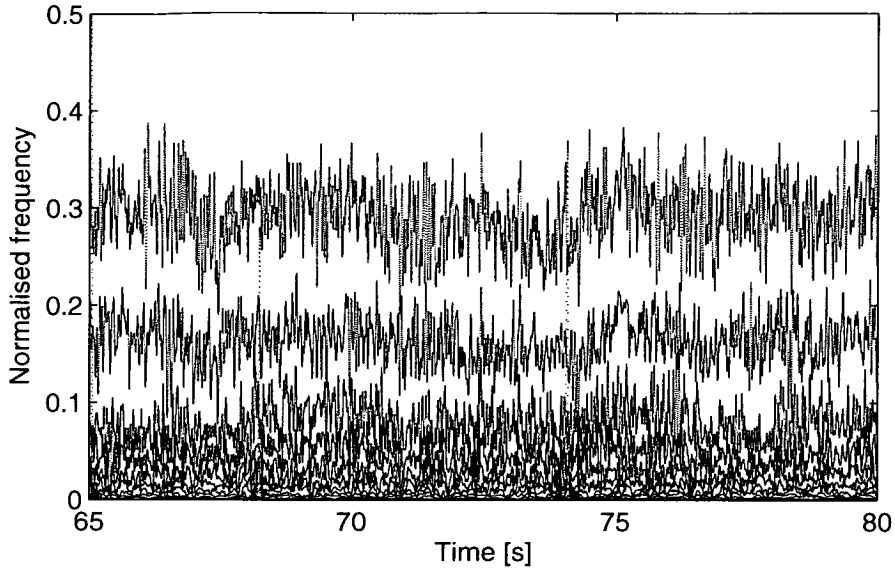


**Figure 5.11:** *The Hilbert spectrum of the fibre glass boat trial data for the part that includes CPA, 65-80 seconds. The image is downsampled with a factor of 50.*

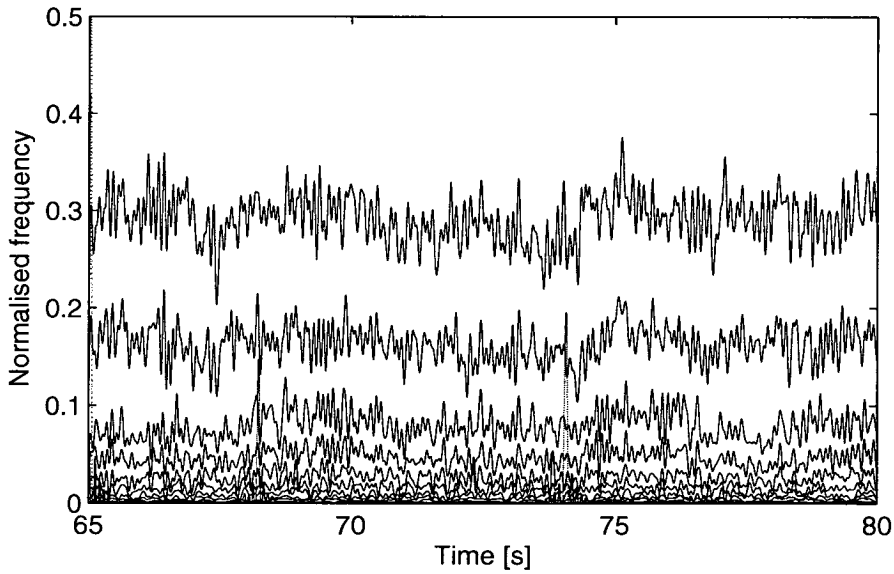


**Figure 5.12:** *The Hilbert spectrum of the fibre glass boat trial data for the part that includes CPA, 65-80 seconds. The image is downsampled with a factor of 100.*

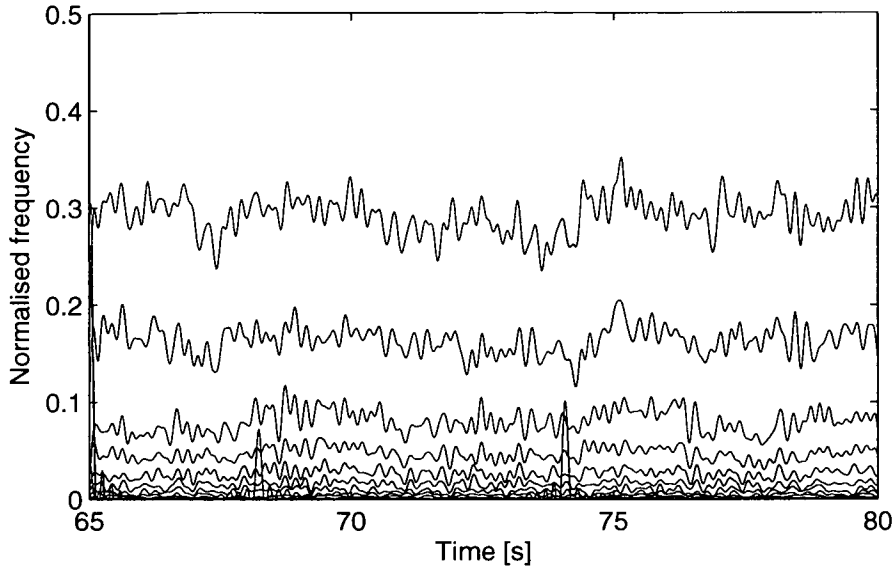


**Figure 5.13:** *The Hilbert spectrum of the fibre glass boat trial data for the part that includes CPA, 65-80 seconds. The image is downsampled with a factor of 200.*

Initially the EMD algorithm was used to produce a hilbert spectrum, see Figure 5.10, and it can be seen in the Figure 5.10 that it is difficult to to obtain any information from the hilbert spectrum. This is due to the fact that the instantaneous frequencies vary rapidly in their amplitude.

However, the data used to produce the images can be lowpass filtered so that the variations become less rapid. This can be achieved by first decimating the data and then resampling it again. For the image shown in Figure 5.11 the data was decimated by a factor of 50, and then resampled again to regain the same number of samples. Now the amplitudes of the instantaneous frequencies are becoming easier to follow. By doing this decmimation and resampling process the actual resolution of the instantaneous frequency has gone down by as much as the decimation factor implies.

In figures 5.12 and 5.13, the data shown was decimated and resampled using a factor of 100 and 200, respectively. In performing this decimation and resampling process the amplitudes of the instantaneous frequencies become a lot more stable and it is posible to distinguish different parts of the signal. It must however be kept in mind that, while performing the decimation and resampling and bringing out features in the data the results must be interpreted with utmost caution. Since the actual instantaneous frequencies as they are shown in Figure 5.10 are the ones that most resemble the physical process behind the produced image.

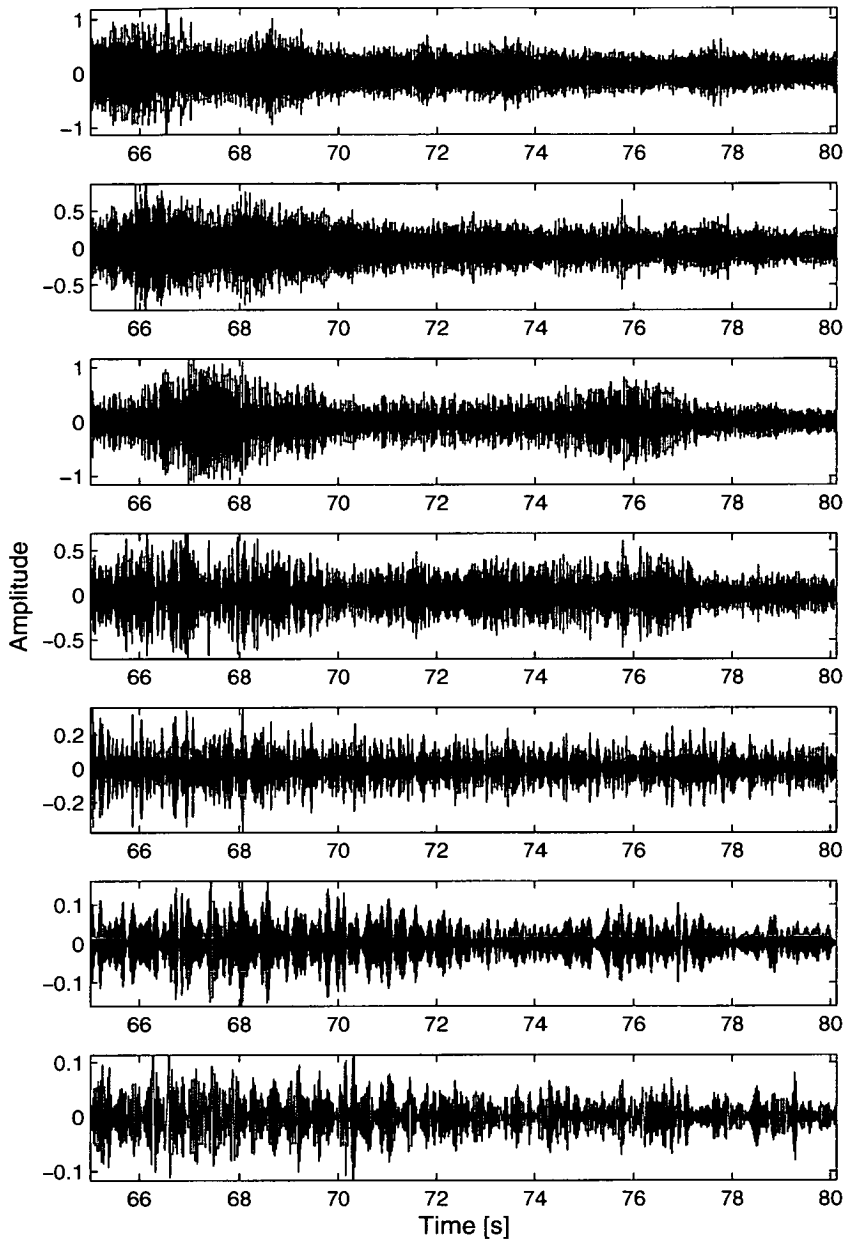


**Figure 5.14:** *The first seven intrinsic mode functions from the fibre glass boat trial data.*

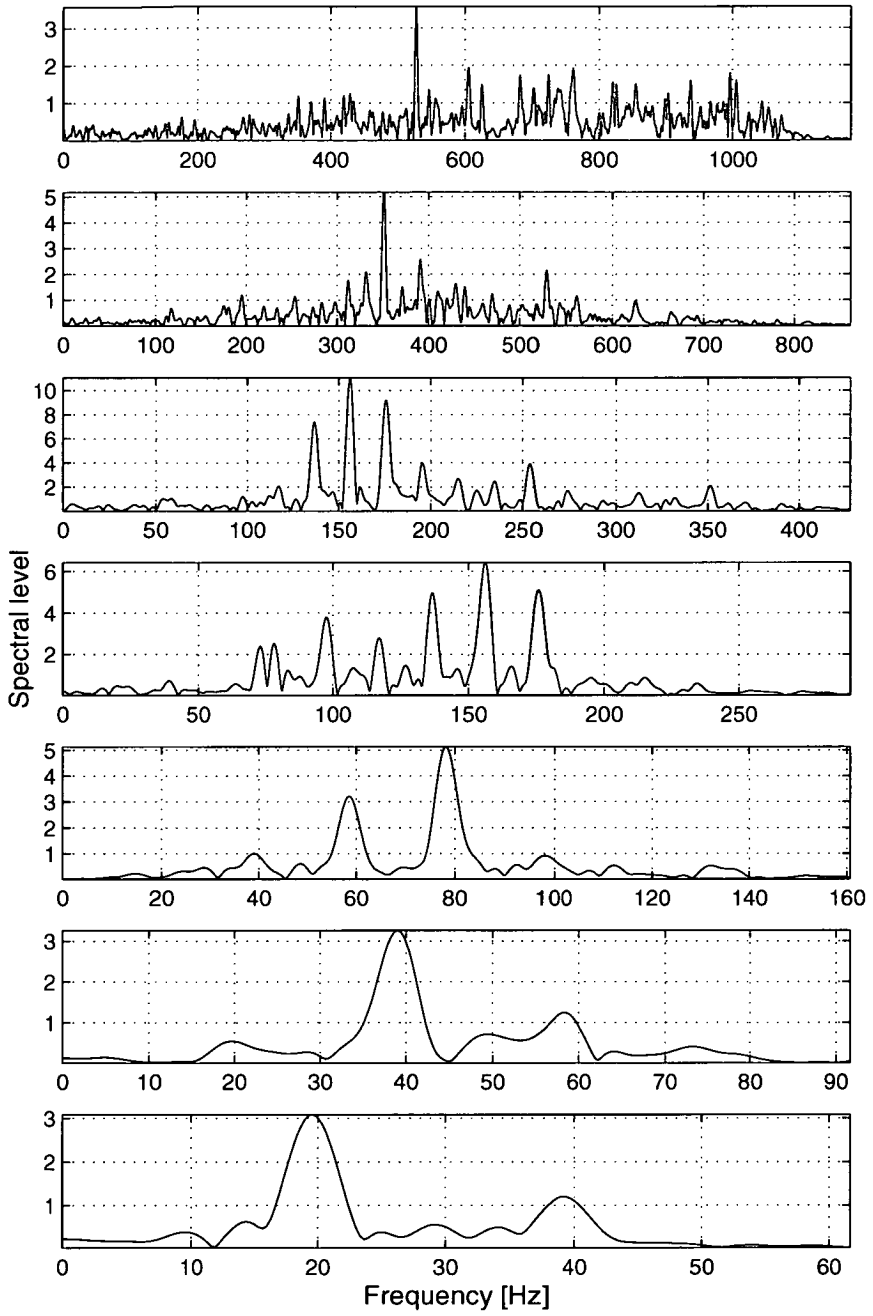


**Figure 5.15:** *The Fourier transform of the first seven intrinsical mode functions from the fibre glass boat trial data.*

Figure 5.14 displays the first seven intrinsic mode functions from the middle part of the fibre glass boat sea trial. Then the power spectrum was estimated for each of the imfs, see Figure 5.15. Some of the frequency peaks that are seen in the power spectrum (Figure 5.9) can also be found in the panels of Figure 5.15. The lowest panel displays the lowest frequencies and the top most panel the highest frequencies. The lowest panel displays the two peaks at 19.5Hz and 39Hz. The second panel from the bottom displays the 39Hz and 58.7Hz peaks. The peaks at 45.4Hz and 64.4Hz are not visually distinguishable in any of the panels. In the third panel from the bottom, the frequencies 58.7Hz and 78.2 Hz are visible. Moving up in frequency the are a number of peaks missing in the spectra from the imfs and some that are difficult to tell apart. The ones missing are 90.4Hz, and 206.9Hz. The frequencies that are difficult to tell apart are 175.8Hz and 180.8Hz, but they can be seen in the spectrum in the second panel from the top. Then the top most panel displays frequencies that are above the ones in the spectrum in Figure 5.9.

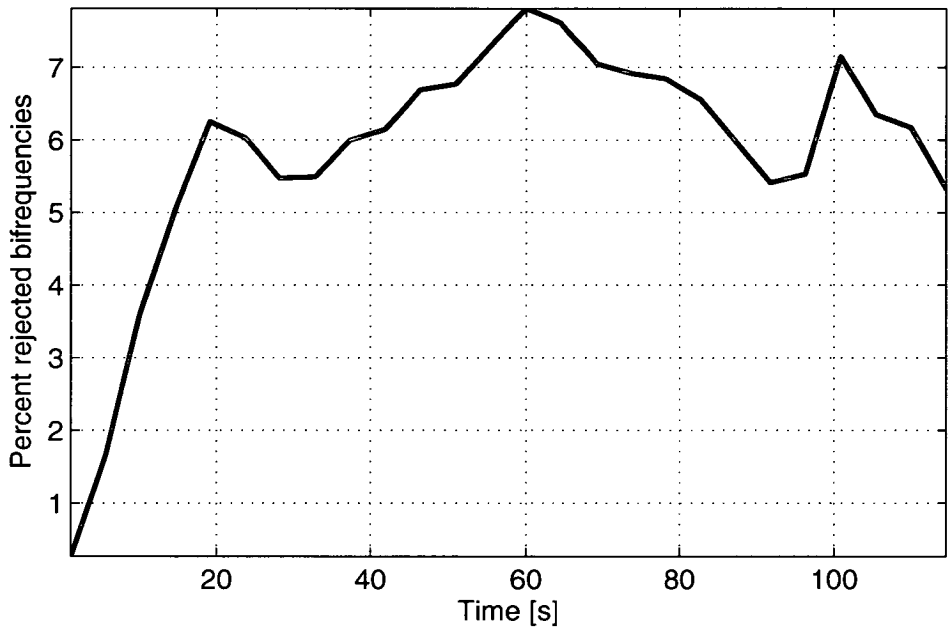


**Figure 5.16:** *The percentage of rejected data by the stepwise outlier algorithm.*

The third part and final part of the data analysis is done making use of the EMD algorithm to pick out the interesting imfs and then summing them to create a lofargram. The data used to

create the modified lofagrams is defined in the equation below

$$s(t) = \sum_{i=N_1}^{N_2} d_i, \tag{5.2}$$

where $d_i$ are the imfs and $N_1$, $N_2$ are the number of the included imfs. In the cases presented below $N_1 = 2$, and $N_2 = 7$, so that out of 14 imfs that makes up the entire data set, only 5 of them are used to create the improved lofargrams.

The number of rejected blocks from the stepwise outlier rejection is used as an input for the further EMD analysis of the data set. In the beginning of the boat run there are very few data points that are considered to be outliers whereas in the middle of the run at CPA there are more data points that are rejected, see Figure 5.16.

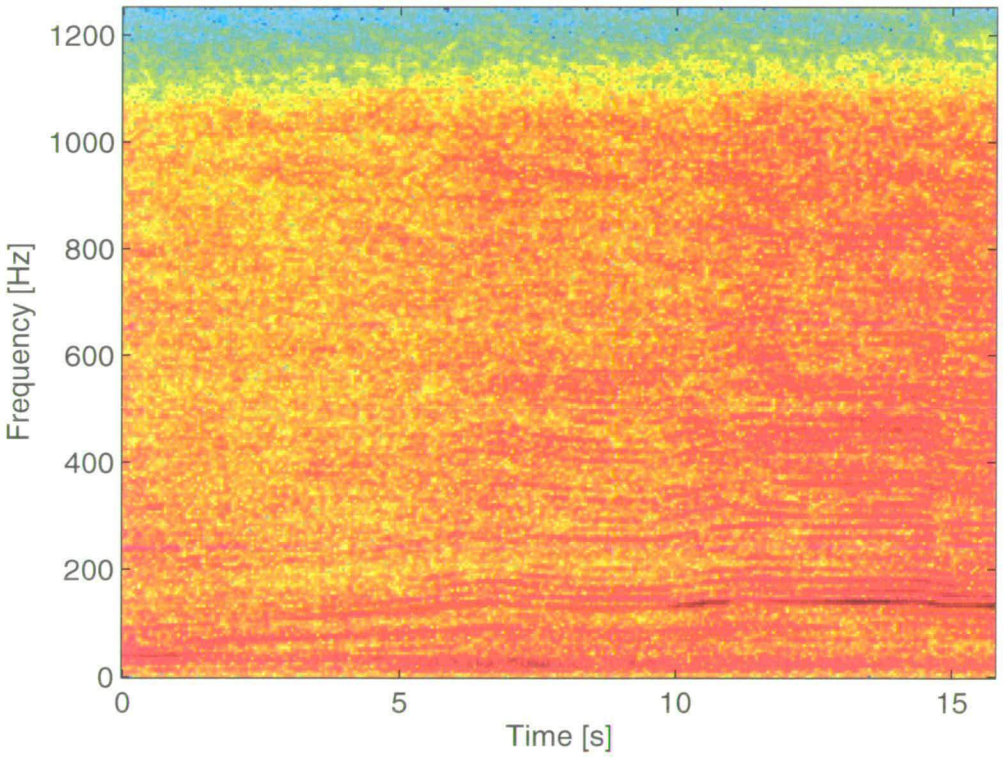


**Figure 5.17:** *A spectrogram of the first 15 seconds of data from the run.*

The two lofagrams from the beginning of the boat run are displayed in Figures 5.17 and 5.18. Where the lofagram in Figure 5.17 is produced from the raw data data set and the lofagram in

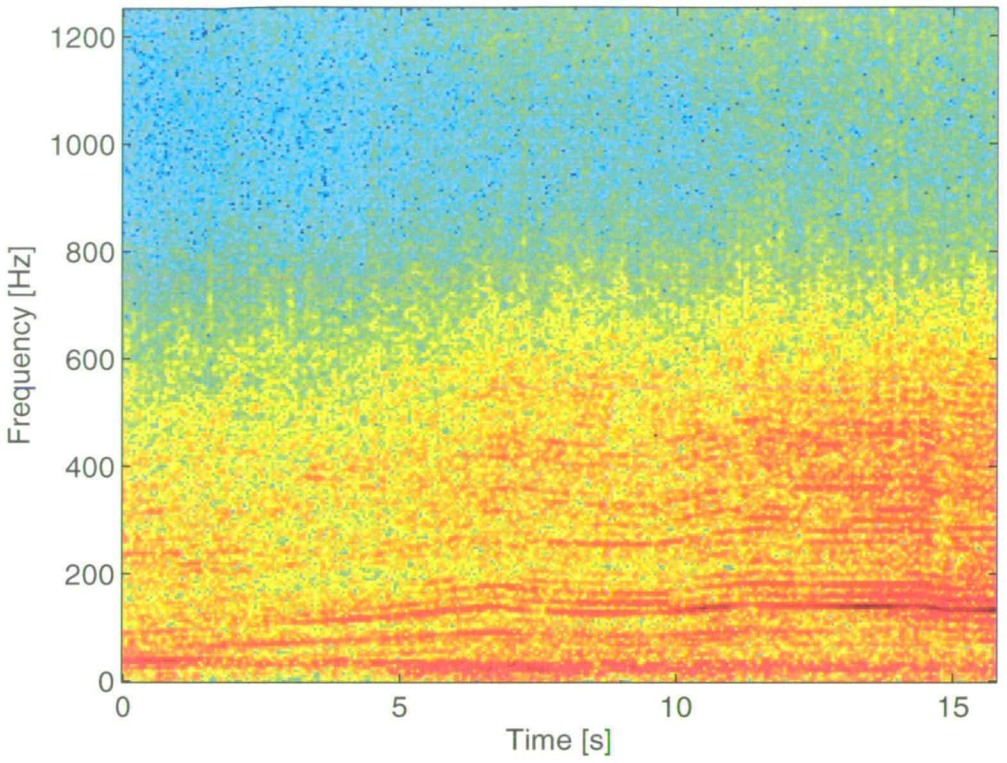


**Figure 5.18:** *A spectrogram of the sum of imfs from imf two to imf seven. From the first 15 secinds of data*

Figure 5.18 is from the reduced times series $s(t)$.

The two lofagrams from the middle of the boat run (including CPA) are displayed in Figures 5.19 and 5.20. Where the lofagram in Figure 5.19 is produced from the raw data data set and the lofagram in Figure 5.20 is from the reduced times series $s(t)$.

In both the lofagrams produced using the reduced data set $s(t)$ it is easier to see the lines made up of tonals in the data. The images are less cluttered, which is a result of removing the first high frequency imf and then by removing the latter imfs the low frequency content below 10Hz are also suppressed. This gives image of the remaining data a better dynamic and the wanted features are easier to see. It is however possible to see the lines in the lofagrams made out of the original data set too.

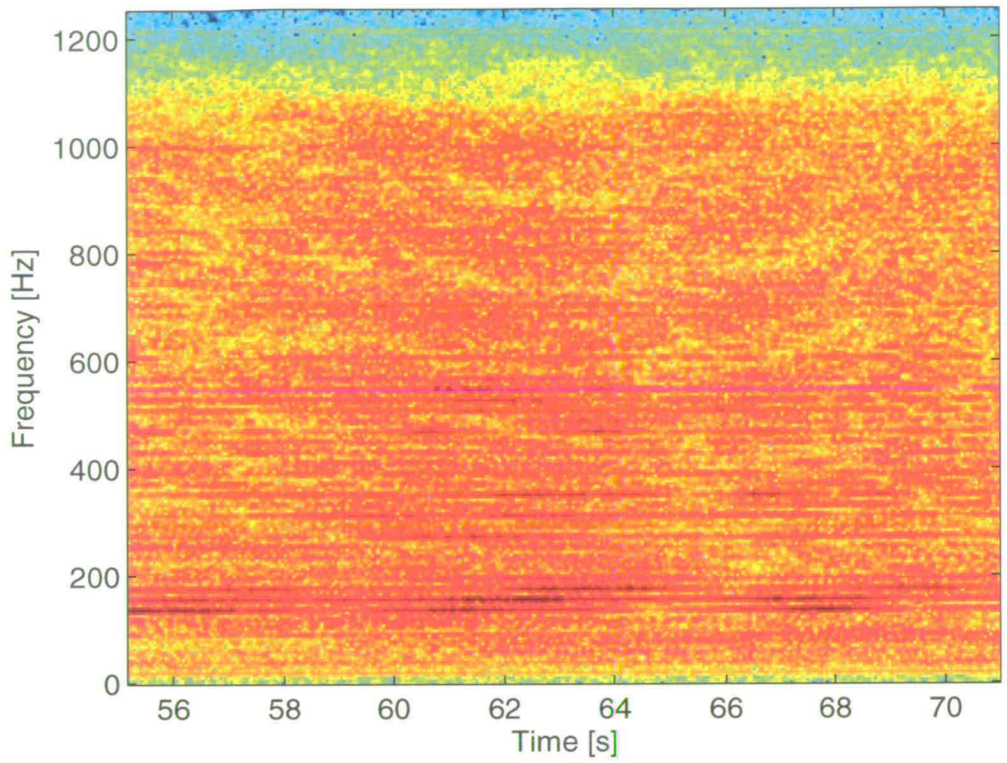


**Figure 5.19:** *A spectrogram of the middle part, 15 seconds of the run including CPA.*

## 5.3 Conclusions

Features in displays can be enhanced by different data conditioning schemes, in this chapter the application was on a Lofargram and the improvement of the spectral lines.

Both the stepwise outlier rejection method and the emd method of cleaning data improved the lofagrams.

The different methods of replacing data in the outlier rejection scheme had less impact on the lofargram than expected but it did show some difference. Depending on the objective of the data conditioning the different data replacement strategies gives slighlty different results results. When the frequencies are known to vary it is beneficial to use method B-E, and when the frequencies are more stable then replacement method A is prefered.

The Hilbert spectra on the IMFs from this data set does not really add to the understanding of the data other than that the instantaneous frequencies vary rapidly.

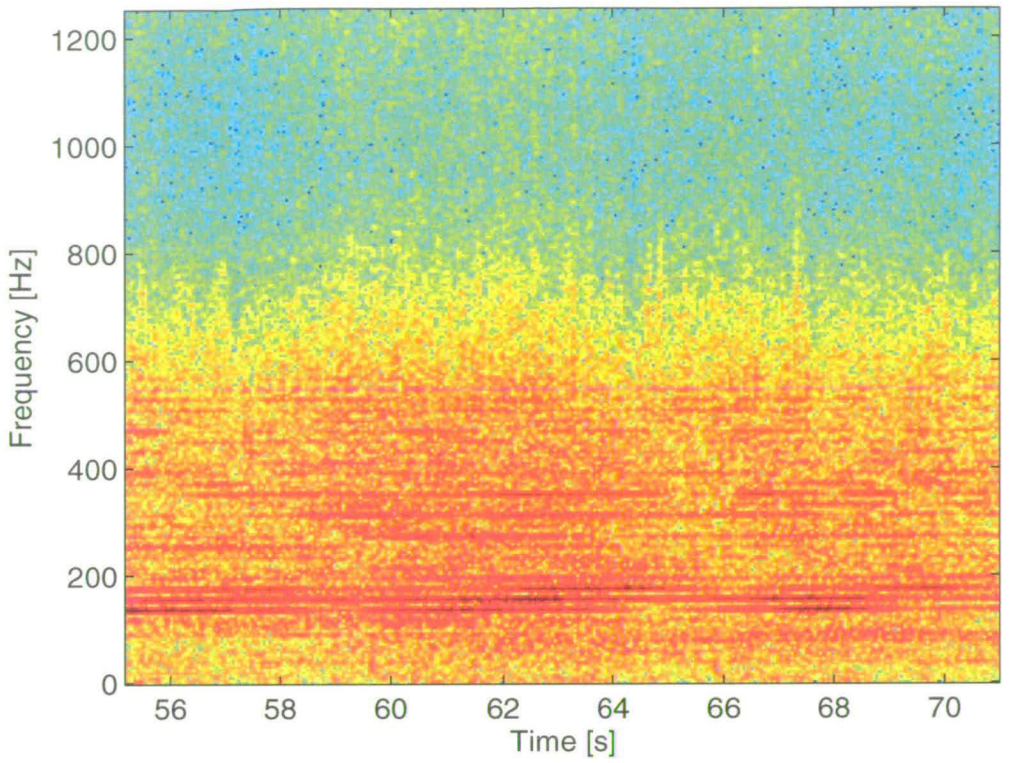


**Figure 5.20:** *A spectrogram of the sum of imfs from imf two to imf seven. From the middle part of the run, including CPA*

The power spectra made from the imfs show most of the frequency content in the original signal but some peaks are not easily distinguished. It should be noted that by using all the imfs and the residual all of the frequncies are accounted for.

As a data conditioning method the EMD shows promise and the lofagrams produced from a sum of the imfs of interest are clearly an improvement over the lofagram produced from the raw data.

# Chapter 6
# Conclusions and Suggestions for future Research

## 6.1 Introduction

The work in this thesis falls into two categories, a statistical characterisation of sonar data and a part that is aimed at time/frequency analysis of sonar data. The latter part is divided into two chapters. One chapter that is aimed at producing time/frequency images and assessing their quality, and a second chapter where the aim is to improve on the quality of time/frequency surfaces.

## 6.2 Achievements

In Chapter 3 on statistical characterisation of sonar data, different aspects of statistical properties of sonar data have been addressed. Mainly focusing on the level of stationarity on both active and passive sonar. It is important to bear in mind that all findings in this chapter are valid for shallow water environments with the typical bottom composition of the Baltic Sea and the archipelago of Stockholm. It is however reasonable to assume that equivalent environments, given that the acoustic parameters also are of close to or equal to the ones in the Baltic Sea, would produce similar results for the analysis reported in this thesis. The general conclusions drawn from the analysis could serve as guidelines for other shallow water areas, as for what kind of performance one might expect from applying signal processing schemes.

For active sonar the issues of ping to ping stationarity, for both the actual ping and for the reverberation tail, was examined and was found to be non-stationary. However, for some cases it is seen that there are consecutive pings that are stationary. This leads to the conclusion that it can be beneficial to use several pings from the same target. However it is necessary to remember that there is a large variation in the number of consecutive stationary pings.

For passive sonar both ambient noise and tonals emitted from surface vessels are investigated. It

is found that for all cases tested the sonar data is non-stationary. That is speaking of stationarity in the strict sense. There are however parts in the data that exhibit more stationary behaviour then the data in general. From the fibre glass boat trial the data seems to be stationary for about 0.4 seconds. That is the PDFs for two consecutive segments maintain sufficiently equal for that period of time.

Testing the passive data for symmetry and linearity show that the data is mostly linear, and symmetrical.

The ambient noise is found to have more stationary properties at higher frequencies. This leads to the conclusion that it is more beneficial to operate sonar systems at higher frequencies. This is only valid in shallow water environments equivalent to the Baltic Sea, where attenuation due to salinity is a lesser factor in the sound propagation. However, in waters with higher salinity the sonar performance would degrade from using higher frequency.

The stationarity length on ambient noise data was examined using data from the from the multi-sensor trial (almost 800 data files). A large proportion of the data set also had a stationarity time of roughly 0.4 seconds, or slightly longer.

In Chapter 4 a number of time-frequency methods were examined. The data analysis was done using the traditional Short Time Fourier Transform, and the Short Time Fractional Fourier Transform. Performance comparisons were made between the two different transforms and it is shown that matching the transform to the nature of the data improves the quality of the time-frequency images, i.e. if the data consists of chirps then it is beneficial to decompose the time series using a chirp as a basis function rather then using sinusoids (the FrFT and FT respectively). It is important to note that if the chirp rate tends towards zero there is no advantages to use the STFrFT over the STFT to produce time/frequency images.

In the second section of Chapter 4 a number of different metrics for image quality assessment are defined and evaluated using the time/frequency images produced with the aforementioned transform techniques.

Of all the implemented image quality measures ($Q - value, Ad, D, Cq$) the Q-value seems to be most consistent with expectation and proves reliable for the presented images. The other metrics have their use but should be applied with caution since the results were inconclusive.

Features in displays can be enhanced by different data conditioning schemes, as can be seen in Chapter 5. The application was to improve the spectral lines in a lofargram.

Both the stepwise outlier rejection method and the EMD method of cleaning data improved the lofargrams.

The different methods of replacing data in the outlier rejection scheme had less impact on the lofargram than expected but it did show some difference. Depending on the objective of the data conditioning the different data replacement strategies gives slightly different results. When the frequencies are known to vary it is beneficial to use method B-E, and when the frequencies are more stable then replacement method A is preferred.

The Hilbert spectra on the IMFs from this data set does not really add to the understanding of the data other than that the instantaneous frequencies vary highly.

The power spectra made from the IMFs show most of the frequency content in the original signal but some peaks are not easily distinguished. It should be noted that by using all the IMFs and the residual all of the frequencies are accounted for.

As a data conditioning method the EMD shows promise and the lofargrams produced from a sum of the IMFs of interest are clearly an improvement over the lofargram produced from the raw data.

## 6.3 Suggested Future Work

Concluding this thesis I suggest a number of possible additions to the work done in the main chapters.

The statistical evaluation of the sonar data in Chapter 3 can be extended with an analysis of a longer ambient noise measurement. This would be done to enable an analysis of the seasonal variations, mainly the difference in the sound speed profile and the influence of ice cover. It would also enable the analysis of the influence of different weather conditions, such as wind speed, wind direction, and wave height (sea state).

The time-frequency methods examined in Chapter 4 could be used from a detection point of view, just as using a matched filter and making the detection of signals in a time lag domain it

could be done in any other domain such as the FrFT produces. The performance of the FrFT could be evaluated in an intercept sonar situation (ping stealing, or torpedo detection).

The data conditioning could be extended to reconstructing data from the enhanced lofargram, or any other altered time/frequency surface. It would be possible to examine the EMD method from a more image processing point of view, since the images produced using the EMD were using significantly less data than the original.

And finally, all of the work done in this thesis is done on sonar data. Lately there has been an increasing interest in using other sensors for underwater applications, such as magnetic and electric. The background noise statistics for that kind of sensors in a shallow water environment is at some extent still unexamined.

# Appendix A
# Description of the Broad Band Signature of a vessel

The broad band signature of vessels are typically given as a simplified spectrum consisting of 56 third note band values.

More familiar to most people is probably the concept of an octave. An octave means the eighth note on a scale and its pitch is twice as high as the base note. For instance, if the frequency for the base note is 400 Hz then the octave to that is 800 Hz. So a spectrum in octave bands starting at 1Hz would have its second band centered around 2Hz, its third band at 4Hz and so on. It is very common to talk about filters and frequency drop off after the cut off frequency in terms of dB/octave or dB/decade. A decade is the as the name suggests a tone that is pitched ten times higher than the base tone. So a spectrum given in decade bands would if the first frequency is 1Hz have 10Hz as its second frequency and 100Hz as its third, and son on. The third note band notation is used because it produces as closer sampled spectrum, since the ratio between the frequency bands is 5;4 as can be seen in Tab A.1 below.

The spectrum is typically presented as a vector of 56 values, one for each frequency band. The spectral density for each band is retreived by reducing the value for the frequency band in question by $10log_{10}(w)$, where $w$ is the band width of the frequency band.

**Table A.1:** *Table over the Frequencies in the broad band description of a signature*

| Freq. band nr. | Center Freq. [Hz] | Lower limit [Hz] | Upper limit [Hz] | Band width [Hz] |
|---|---|---|---|---|
| 1 | 1.25 | 1.13 | 1.41 | 0.28 |
| 2 | 1.6 | 1.41 | 1.76 | 0.35 |
| 3 | 2 | 1.76 | 2.2 | 0.44 |
| 4 | 2.5 | 2.2 | 2.8 | 0.6 |
| 5 | 3.15 | 2.8 | 3.5 | 0.7 |
| 6 | 4 | 3.5 | 4.4 | 0.9 |

| | | | | |
|---|---|---|---|---|
| 7 | 5 | 4.4 | 5.7 | 1.3 |
| 8 | 6.3 | 5.7 | 7.1 | 1.4 |
| 9 | 8 | 7.1 | 8.8 | 1.7 |
| 10 | 10 | 8.8 | 11.3 | 2.5 |
| 11 | 12.5 | 11.3 | 14.1 | 2.8 |
| 12 | 16 | 14.1 | 17.6 | 3.5 |
| 13 | 20 | 17.6 | 22 | 4.4 |
| 14 | 25 | 22 | 28 | 6 |
| 15 | 31.5 | 28 | 35 | 7 |
| 16 | 40 | 35 | 44 | 9 |
| 17 | 50 | 44 | 57 | 13 |
| 18 | 63 | 57 | 71 | 14 |
| 19 | 80 | 71 | 88 | 17 |
| 20 | 100 | 88 | 113 | 25 |
| 21 | 125 | 113 | 141 | 28 |
| 22 | 160 | 141 | 176 | 35 |
| 23 | 200 | 176 | 225 | 49 |
| 24 | 250 | 225 | 283 | 58 |
| 25 | 315 | 283 | 353 | 70 |
| 26 | 400 | 353 | 440 | 87 |
| 27 | 500 | 440 | 565 | 125 |
| 28 | 630 | 565 | 707 | 142 |
| 29 | 800 | 707 | 880 | 173 |
| 30 | 1000 | 880 | 1130 | 250 |
| 31 | 1250 | 1130 | 1414 | 284 |
| 32 | 1600 | 1414 | 1760 | 346 |
| 33 | 2000 | 1760 | 2250 | 490 |
| 34 | 2500 | 2250 | 2825 | 575 |
| 35 | 3150 | 2825 | 3530 | 705 |
| 36 | 4000 | 3530 | 4400 | 870 |
| 37 | 5000 | 4400 | 5650 | 1250 |
| 38 | 6300 | 5650 | 7070 | 1420 |
| 39 | 8000 | 7070 | 8800 | 1730 |

| | | | | |
|---|---|---|---|---|
| 40 | 10000 | 8800 | 11300 | 2500 |
| 41 | 12500 | 11300 | 14100 | 2800 |
| 42 | 16000 | 14100 | 17600 | 3500 |
| 43 | 20000 | 17600 | 22500 | 4900 |
| 44 | 25000 | 22500 | 28300 | 5800 |
| 45 | 31500 | 28300 | 35300 | 7000 |
| 46 | 40000 | 35300 | 44000 | 8700 |
| 47 | 50000 | 44000 | 56500 | 12500 |
| 48 | 63000 | 56500 | 70700 | 14200 |
| 49 | 80000 | 70700 | 88000 | 17300 |
| 50 | 100000 | 88000 | 113000 | 25000 |
| 51 | 125000 | 113000 | 141400 | 28400 |
| 52 | 160000 | 141400 | 176000 | 34600 |
| 53 | 200000 | 176000 | 225000 | 49000 |
| 54 | 250000 | 225000 | 283000 | 58000 |
| 55 | 315000 | 283000 | 353000 | 70000 |
| 56 | 400000 | 353000 | 440000 | 87000 |

# Appendix B
# KS testing the ambient noise data from the multisensor trial

The results in this appendix are an extention to the analysis results displayed in Figure 3.27. Figure B.1 is from the early part of the trial (Aug 16 - Aug 17) where the stationarity times varied a lot but also provided the longest and of course the shortest stationarity times. Figure B.1 is an example of when the stationarity time was exceptionally long compared to other parts in this data set, at around 4-5 seconds. Figure B.3 is also from the first days of the sea trial, and is an example where the data is strictly non-stationary at all times. Figure B.3 is from the middle part of the trial (Aug 19 - Aug 23) where the stationarity time remained constant over the period, where the typical stationarity time was about 0.4-0.5 seconds.

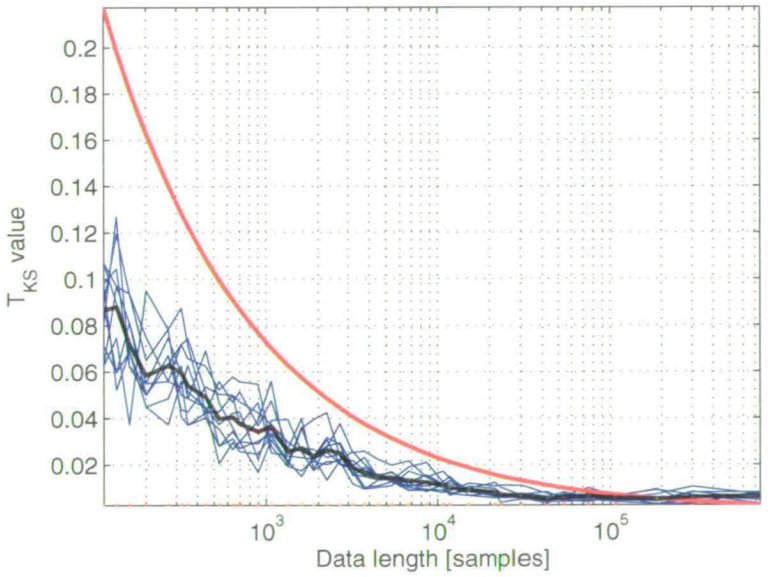


**Figure B.1:** *The KS test applied to different lengths of data from the multi sensor trial in June 2004. Example file 1, where the test statistic is less than the p-value for sample sizes up to 185,000 samples.*

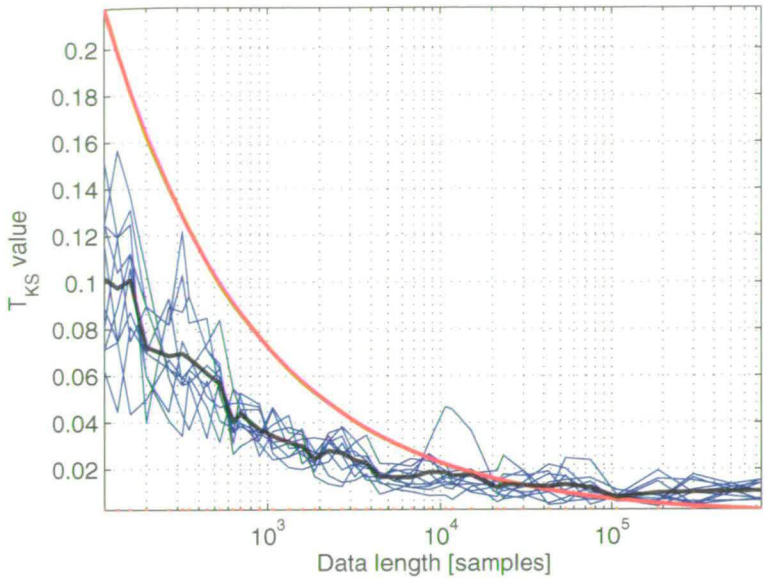


**Figure B.2:** *The KS test applied to different lengths of data from the multi sensor trial in June 2004. Example 2 where the test statistic is less than the p-value for all sample sizes of about 10,000 samples.*

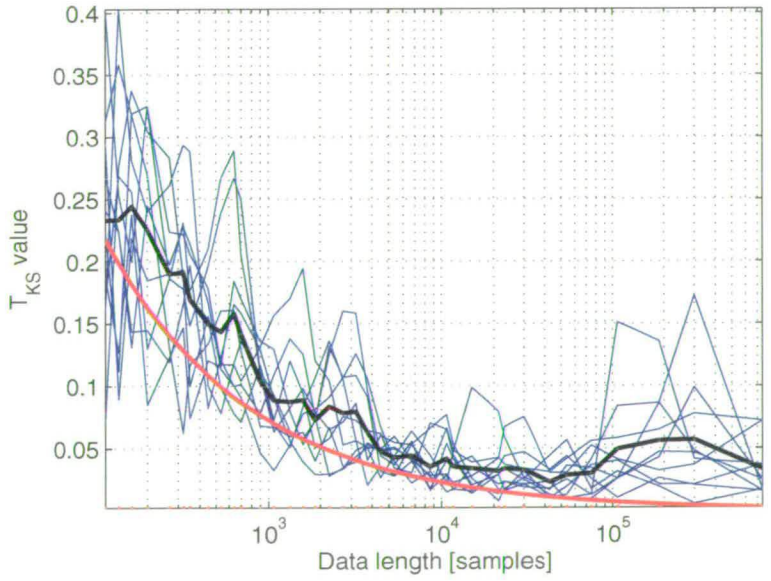


**Figure B.3:** *The KS test applied to different lengths of data from the multi sensor trial in June 2004. Example 3 where the test statistic is larger than the p-value for all sample sizes.*

Table B.1 shows the amount data used for the stationarity test. The x-axis of Figures B.1, B.2, and B.3 is a logarithmic scale where the smallest and largest value corresponds to the smallest and largest value in Table B.1 respectively.

| | | | |
|---|---|---|---|
| 112 | 700 | 4480 | 42880 |
| 134 | 896 | 5360 | 53600 |
| 160 | 1072 | 6700 | 75040 |
| 200 | 1340 | 8576 | 107200 |
| 268 | 1600 | 10720 | 187600 |
| 320 | 1876 | 11725 | 300160 |
| 350 | 2240 | 15008 | 750400 |
| 448 | 2680 | 21440 | |
| 536 | 3200 | 23450 | |
| 640 | 3752 | 30016 | |

**Table B.1:** *Number of samples used in the stationary testing of the ambient noise data collected during June 2004, where the sampling frequency was 25kHz.*

# Appendix C
# Reverberation stationarity analysis on the Utö data

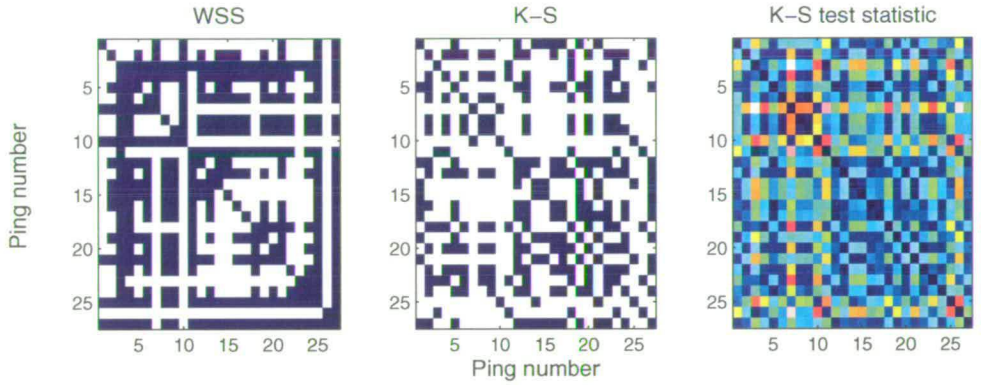


**Figure C.1:** *The WSS and the K-S test applied to data from the reverberation sea trial at Utö, File number 1. The panels show, from left to right, the WSS hypothesis, the K-S hypothesis, and the K-S test statistic.*

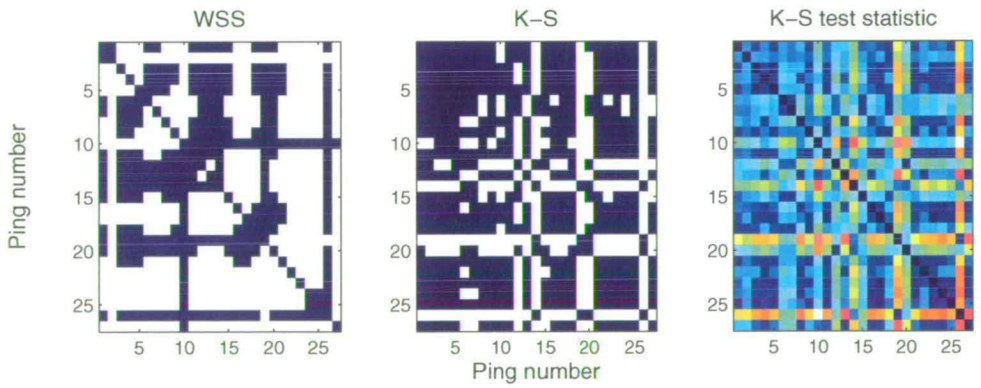


**Figure C.2:** *The WSS and the K-S test applied to data from the reverberation sea trial at Utö, File number 2.*

Figures C.1, C.2 and C.3 show the full analysis results from the three files from the reverberation sea trial at Utö 2001 that was analysed in Chapter 3. The tests are applied to the data starting 1 second after each ping has been transmitted.

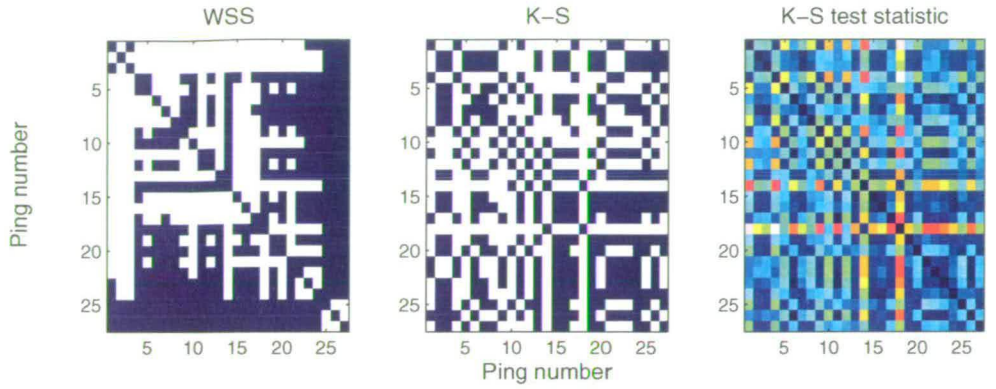


**Figure C.3:** *The WSS and the K-S test applied to data from the reverberation sea trial at Utö, File number 3.*

The K-S test rejects the null hypothesis for all files but as it can be seen in the right side panel in figures C.1, C.2 and C.3 the value of the test statistic varies and it is smaller in the regions around the main diagonal. The mean value of the test statistic shown in the figures is 0.02 and the threshold it is compared to has a value of 0.01. The colormap for the right side panel in the figures goes from blue to dark red, where blue values are low and the red ones are high. Some of the test cases (ping to ping) are stationary but the entire data set as such, is not.

The percentage of stationary pings is shown in Table C.1, for both the K-S and the WSS tests.

In the figures C.1, C.2 and C.3, in the left and middle panels, the null hypothesis is rejected where the figure is colored white, and conversely the figure is blue when the null hypothesis is accepted.

| Percentage of accepted $H_0$ | File 1 | File 2 | File 3 |
|---|---|---|---|
| WSS | 49 | 61 | 47 |
| SS | 54 | 30 | 46 |

**Table C.1:** *The percentage of accepted null hypothesis for the K-S (strictly stationary) and WSS (wide sense stationnary) tests for File 1 - File 3.*

# Appendix D
# The Choi-Williams distribution

Before moving on to other time/frequency distributions, the idea of a general class of transforms that can be presented in the form of an integral over a quadratic form of a signal $f(t)$ and a kernel $\phi(\Theta, \tau)$ is introduced, see Chapter 9 [83].

$$W_{CW} = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^*(u - \tau/2) f(t + \tau/2) \phi(\Theta, \tau) e^{-j(\Theta t + \omega\tau - \Theta u)} du d\tau d\omega \quad \text{(D.1)}$$

Using the above notation the Wigner distribution has a kernel $\phi(\Theta, \tau) = 1$, and the form in Equation (4.26) is obtained. The Spectrogram has the kernel

$$\phi(\Theta, \tau) = \int_{-\infty}^{\infty} h^*(u - \tau/2) e^{-j\Theta u} h(t + \tau/2) du \quad \text{(D.2)}$$

It becomes clear that with this formulation a time/frequency distribution with the desired characteristics can be obtained. The Choi-Williams distribution was devised to address the one big disadvantage with the Wigner distribution, the cross terms that are present for multi-component signals. The kernel function for the Choi-Williams,[77], distribution has the following form

$$\phi(\Theta, \tau) = e^{-\Theta^2 \tau^2 / \sigma} \quad \text{(D.3)}$$

Which is a product kernel, and $\sigma$ is a parameter that is varied to obtain different levels of cross term suppression. For large $\sigma$ the Choi-Williams distribution approaches the Wigner distribution. Moreover, both marginals are satisfied, which can be seen by looking at the following property, $\phi(0, \tau) = \phi(\Theta, 0) = 1$.

Then by substituting the kernel into the general class function the following is obtained

$$P_{CW}(t,\omega) = \frac{1}{4\pi^{3/2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\tau^2/\sigma}}) e^{-\frac{(u-\tau)^2}{4\tau^2/\sigma} - j\tau\omega} f^*(u-\tau/2) f(t+\tau/2) du d\tau \quad \text{(D.4)}$$

The problem with cross terms has been studied extensively and this usually involves violating the desired properties like the marginals. Choi and Williams introduced this method to reduce the cross terms without violating any of the desired properties of the distribution. An example of how the Choi-Williams distribution affects the cross terms is given below. Consider the sum of two sine waves,

$$f(t) = A_1 e^{j\omega_1 t} + A_2 e^{j\omega_2 t} \quad \text{(D.5)}$$

For the signal in Equation (D.5) the distribution can be calculated exactly from Equation (D.4), which yields,

$$P_{CW}(t,\omega) = A_1^2 \delta(\omega-\omega_1) + A_2^2 \delta(\omega-\omega_2) + 2A_1 A_2 cos((\omega_2-\omega_1))\eta(\omega,\omega_1,\omega_2,\sigma) \quad \text{(D.6)}$$

where

$$\eta(\omega,\omega_1,\omega_2,\sigma) = \sqrt{\frac{1}{4\pi(\omega_1-\omega_2)^2/\sigma}} exp\left[-\frac{(\omega-(\omega_1+\omega_2)/2)^2}{4(\omega_1-\omega_2)^2/\sigma}\right] \quad \text{(D.7)}$$

Seeing that the amplitudes $A_1$, and $A_2$ are followed by $\delta$ functions that are infinite at $\omega_1 = \omega_2$, the distribution will be infinitely peaked at $\omega_1 = \omega_2$, which is at the auto terms. The amplitude of the cross term depends on $\eta(\omega,\omega_1,\omega_2,\sigma)$. The limit of $\eta(\omega,\omega_1,\omega_2,\sigma)$, $\sigma \rightarrow \infty$, as shown below

$$\lim_{\sigma\rightarrow\infty} \eta(\omega,\omega_1,\omega_2,\sigma) = \delta(\omega-(\omega_1+\omega_2)/2) \quad \text{(D.8)}$$

gives back the Wigner distribution, and the infinite peak at $\omega = (\omega_1+\omega_2)/2$.

An example of how the Choi-Williams distribution works for a multicomponent signal can be seen in Figure D.1 where the distribution is estimated for $\sigma = \{10^5, 10^9\}$ respectively. It is clear to see the effect that the kernel has on the distribution. For $\sigma = 10^5$ the distribution is

clear of cross terms but the lines corresponding to the two chirps are also slightly smeared out in frequency, but the two chirps are well localised, so this choice of $\sigma$ is clearly better then the one displayed in the bottom panel of the figure. The signals that are analysed here are the same ones as in Section 4.2.2, and displayed in Figures 4.4.

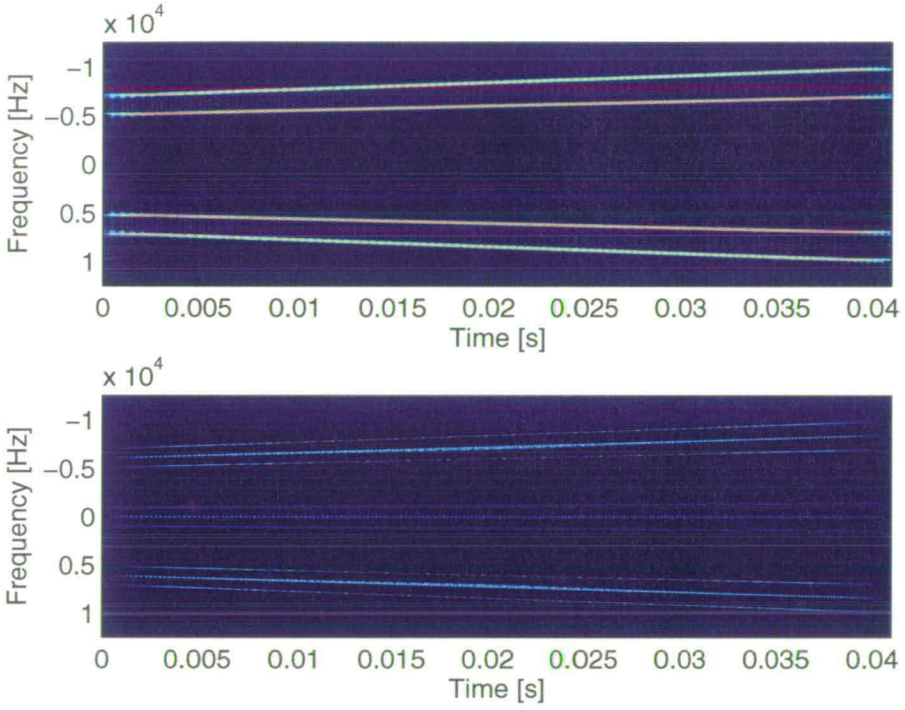


**Figure D.1:** *The Choi-Williams distribution for two chirps, using* $\sigma = \{10^5, 10^9\}$

For $\sigma = 10^9$, displayed in the bottom panel of Figure D.1, the distribution is approximately back to the Wigner distribution and the cross terms are back as can they were in Figure 4.4.

# List of Publications

- M Levonen, L Persson, and S McLaughlin "Stationarity Testing of Sonar Data Using the KS Test", Proceedings of 5th nordic signal processing symposium, Oct. 2002
- M Levonen and S McLaughlin "Analysing Sonar Data Using the Fractional Fourier Transform", Proceedings of 5th nordic signal processing symposium, Oct. 2002
- M Levonen and S McLaughlin "Fractional Fourier transform techniques applied to active sonar", Proceedings of OCEANS 2003. pp.1894 - 1899, Vol.4, Sept. 2003
- M Levonen and L Persson "Conditioning of sonar data", Proceedings of OCEANS 2003. pp.1900 - 1904 Vol.4, Sept. 2003
- M Levonen and L Persson "Temporal properties of ambient noise in the Baltic Sea", Proceedings of UDT 2003, June 24-26, Malmö
- M Levonen, L Persson, and S McLaughlin "Image quality assessment in time/frequency surfaces", Proceedings of MTS/IEEE TECHNO-OCEAN '04, pp.665 - 669 Vol.2, Nov. 2004

- L Persson and M Levonen "Statistical tests for sonar data conditioning", Proceedings of MTS/IEEE Oceans '02, pp. 2234 - 2238 vol.4, Oct 2002
- S McLaughlin, M Levonen, R Lennartsson, J Robinson, and L Persson "Comparison of quadratic phase coupling detectors on sonar data", Proceedings of Oceans '03. pp.1685 - 1690 vol 3. Sept 2003
- R Lennartsson, S McLaughlin, M Levonen, J Robinson, and L Persson "Energy separation algorithms applied to sonar data", Proceedings of MTS/IEEE TECHNO-OCEAN '04, pp.781 - 785 Vol.2, Nov 2004
- E Dalberg, R Lennartsson, M Levonen, and L Persson "Properties of acoustic ambient noise in the Baltic SEA", Proceedings of "Underwater Acoustic Measurements: Technologies & Results" Heraklion, Crete, Greece, 28th June - 1st July 2005

# References


[1] R. J. Urick, *Principles of underwater sound.* Los Altos, California: Peninsula Publishing, third ed., 1983.

[2] W. S. Burdic, *Underwater Acoustic System Analysis.* Prentice-Hall, 1990, 2nd edition.

[3] A. D. Waite, *Sonar for practising Engineers. 2:d edition.* Thomson Marconi Sonar Limited, 1998.

[4] Finn B. Jenssen, William A. Kuperman, Michael B. Porter, and Henrik Schmidt, *Computational Ocean Acoustics.* Springer-Verlag, 2000.

[5] Xavier Lurton, *An Introduction to Underwater Acoustics Principles and Applications.* Springer, 2002.

[6] Harry L. Van Trees, *Optimum Array Processing.* John Wiley & Sons, Inc., 2002.

[7] J.H. Justice, N.L. Owsley, J.L. Yen, and A.C. Kak, *Array Signal Processing.* Prentice-Hall, 1985.

[8] Prabhakar S. Naidu, *Sensor Array Signal Processing.* CRC Press Prentice-Hall, 2001.

[9] R.J. Kozick, B.M. sadler, "Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3520–3535, 2000.

[10] S.D. Gray, J.C. Preisig, and D. Brady, "Multiuser detection in a horizontal underwater acoustic channel using array observations," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 148–160, 1997.

[11] P. Stoica, M. Viberg, B. Ottersten, "Instrumental variable approach to array processing in spatially correlated noise fields," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 121–133, 1994.

[12] P. Chevalier, L. Albera, A. Ferreol, P. Comon, "On the virtual array concept for higher order array processing," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1254–1271, 2005.

[13] C.A. Speirs, D.T. Highes, J.L. Mather and A.R. Green, "Practical considerations of using adaptive beamforming with real sonar data," *Proceedings of Electronic Engineering in Oceanography, Seventh international Conference on Technology transfer from Research to Industry*, pp. 120–127, June 1997.

[14] K. Lo, "Adaptive array processing for wide-band active sonars," *IEEE Journal of Oceanic Engineering*, vol. 29, pp. 837–846, July 2004.

[15] Yao Kung, R.E. Hudson, C.W. Reed, Chen Daching and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1555–1567, Oct 1998.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* Upper Saddle River, New Jersey: Prentice Hall, 1993.

[17] M. Wazenski, D. Alexandrou, D. DeFatta, "Roc evaluation of adaptive beamforming in a simulated shallow water environment," *IEEE Journal of Oceanic Engineering*, vol. 21, no. 1, pp. 94–99, 1996.

[18] Y.T. Chan, G.H. Niezgoda,S.P. Morton, "Passive sonar detection and localization by matched velocity filtering," *IEEE Journal of Oceanic Engineering*, vol. 20, no. 3, pp. 197–189, 1995.

[19] L.L. Scharf and M.L. McCloud, "Blind adaptation of zero forcing projections and oblique pseudo-inverses for subspace detection and estimation when interference dominates noise," *IEEE Transactions on Signal Processing*, vol. 50, no. 12, pp. 2938–2946, 2002.

[20] S.C. White and N.C. Beaulieu, "On the application of the Cramer-Rao and detection theory bounds to mean square error of symbol timing recovery," *IEEE Transactions on Communications*, vol. 40, no. 10, pp. 1635–1643, 1992.

[21] Harry L. van Trees, "Applications of state-variable techniques in detection theory," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 653–669, 1970.

[22] N.O. Booth, R. Judd, and H. Buckner, "Measurement of vertical noise directionality with a mixed-polarity vertical array," *Journal of Oceanic Engineering*, vol. 28, no. 3, pp. 537–543, 2003.

[23] J. Foster, "Rain measurement on buoys using hydrophones," *Journal of Oceanic Engineering*, vol. 19, no. 1, pp. 23–29, 1994.

[24] H.C. Pumphrey and J.E. Ffowcs Williams, "Bubbles as sources of ambient noise," *Journal of Oceanic Engineering*, vol. 15, no. 4, pp. 268–274, 1990.

[25] V.O. Knudsen, R.S. Alford, and J.W. Emling, "Underwater ambient noise," *Journal of Marine Research*, vol. 7, pp. 410–429, 1948.

[26] H.W. Marsh, "Origin of the knudsen spectra," *Journal of Marine Research*, vol. 35, pp. 409–410, 1963.

[27] G.M. Wenz, "Acoustic ambient noise in the ocean," *Journal of the Acoustical Society of America*, vol. 34, pp. 1936–1936, 1962.

[28] D.W. Miklovic,M.E. Huster,and M. Moser-Breaker, "The spatial variability of surface reverberation under high and low wind speeds," *Journal of the Acoustical Society of America*, vol. 20, no. 4, pp. 337–339, 1995.

[29] G.A. Scanlon,R.H. Bourke, and J.H. Wilson, "Estimation of bottom scattering strength from measured and modeled mid-frequency sonar reverberation levels," *Journal of Oceanic Engineering*, vol. 21, no. 4, pp. 440–445, 1996.

[30] R. Ting, "A review on the development of piezoelectric composites for underwater acoustic transducer applications," *IEEE transactions on Instrumentation and Measurement*, vol. 41, pp. 64–67, Feb 1992.

[31] J.P. Voccio, C.H. Joshi and J.F. Lindberg, "Application of high-temperature superconducting wires to magnetostrictive transducers for underwater sonar," *IEEE Transactions on Magnetics*, vol. 30, pp. 1693–1698, Jul 1994.

[32] S. Pilgrim, "Electrostrictive ceramics for low-frequency active transducers," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 47, pp. 861–876, Jul 2000.

[33] G. Sessler, "What's new in electroacoustic transducers," *ASSP Magazine IEEE*, vol. 1, pp. 3–13, Oct 1984.

[34] A. Dandridge, "Fiber-optic sensors make waves in acoustics, control, and navigation," *Circuits and Devices Magazine, IEEE*, vol. 6, pp. 12–19, Nov 1990.

[35] P. Abbot, S. Celuzza,I. Dyer, B. Gomes, J. Fulford, and J. Lynch, "Effects of east china sea shallow-water environment on acoustic propagation," *Journal of Oceanic Engineering*, vol. 28, no. 2, pp. 192–211, 2003.

[36] Jr W Hodgkiss, "An oceanic reverberation model," *Journal of Oceanic Engineering*, vol. 9, no. 2, pp. 63–74, 1984.

[37] R. Arvidsson, *Seismodynamics of Sweden deduced from earthquake-source and seismic-wave characteristics*. PhD thesis, Uppsala University, 1991.

[38] T Floden and P Söderberg, "Shallow gas traps and gas migration models in crystalline bedrock areas offshore sweden," *Baltica*, 1994.

[39] T Floden and P Söderberg, "Pockmarks and related seabed structures in some areas of precambrian bedrock in sweden," *Geological Survey of Finland*, 1988.

[40] S. Ivansson, N. Jacobsen, M. Levonen, B. Nilsson, P. Morén and P. Söderberg, "Acoustic detection of buried objects," *Swedish Defence Research Agency, Div. of Systems Technology, FOI-R-0324-SE*, Dec 2001.

[41] B. Berntsen, S.A. Frivik and J.M. Hovem, "MAST III - ISACS Final Report 1997-1999," *NTNU*, 2000.

[42] B. Berntsen, O. Bergem, E. Pouliquen and N.G. Pace, "Preliminary studies of seafloor characterisation using BORIS simulations," *Saclantcen, Report*, Sept 1997.

[43] Lennart Crona, Eva Dalberg, Ron Lennartsson, Björn Lundqvist, Per Morén, Leif Perssson and Per Söderberg, "An experiment on passive multi-sensor underwater surveillance.," *FOI-R–1336–SE, ISSN 1650-194*, Aug 2004.

[44] Jörgen Pihl, Sven Ivansson, Peter Karlsson, Ron Lennartsson, Mika Levonen, Per Morén, Bernt Nilsson, Marie Olsson, Örjan Staaf, Gunnar Sundin and Per Söderberg, "Low frequency active sonar - reverberation and performance in the baltic," *FOI-R–0887–SE, ISSN 1650-1942*, June 2003.

[45] G. Shippey, J. Pihl and M. Jönsson, "Autopositioning for wideband synthetic aperture sonar using fast factored back projection," in *Computer-aided detection and computer-aided classification conference, Halifax, Canada*, Nov 2001.

[46] L. Persson, R.K. Lennartsson, J.W.C. Robinson and S. McLaughlin, "Quadratic phase coupling analysis of passive sonar data using biphase techniques," in *Proceedings of OCEANS 2000 MTS/IEEE Conference and Exhibition*, vol. 2, pp. 1053–1057, Sept 2000.

[47] R. K. Lennartsson, J.W.C. Robinson, L. Persson, M.J. Hinich and S. McLaughlin, "Passive sonar signature estimation using bispectral techniques," in *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing*, pp. 281–285, Aug 2000.

[48] L. Persson, R.K. Lennartsson and S. McLaughlin, "Sonar data conditioning by means of stepwise outlier rejection," in *Proceedings of OCEANS 2001 MTS/IEEE Conference and Exhibition*, vol. 3, pp. 1629–1635, Nov 2001.

[49] Johan G. Lourens and Johan A. du Preez, "Sonar ML Estimator for Ship Propeller Speed," *IEEE Journal of Oceanic Engineering*, vol. 23, pp. 448–453, Oct 1998.

[50] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York, NY: McGraw Hill, fourth ed., 2002.

[51] H. Cramér, *Mathematical Methods of Statistics*. Almqvist & Wiksells, 1945.

[52] D. Brillinger, "An introduction to polyspectra," *Annals of Mathematical Statistics*, vol. 36, pp. 1351–1374, 1965.

[53] B. Porat, *Digital Processing of Random Signals: Theory and Methods*. Englewood Cliffs, New Jersey: Prentice Hall, 1994.

[54] M. B. Priestly, *Spectral Analysis and Time Series*. London: Academic Press, 1981.

[55] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, New Jersey: Prentice Hall, 1993.

[56] J.W.A. Fackrell, S. McLaughlin and P.R. White, "Bicoherence estimation using the direct method: Part 1: Theoretical considerations," *Applied Signal Processing*, vol. 2, no. 3, pp. 155–168, 1995.

[57] J.W.A. Fackrell, S. McLaughlin and P.R. White, "Bicoherence estimation using the direct method: Part 2: Practical considerations," *Applied Signal Processing*, vol. 2, no. 4, pp. 186–199, 1995.

[58] M. J. Hinich, "Non-linear and non-Gaussian ocean noise," *Journal of the Acoustical Society of America*, vol. 82, no. 4, pp. 1386–1394, 1987.

[59] J. W. Fackrell and S. MacLaughlin, "Robust non-parametric bicoherence estimation by stepwise outlier rejection," *Electronics Letters*, vol. 36, no. 4, pp. 368–370, 2000.

[60] M. Båth, *Spectral analysis in Geophysics*. Elsevier Scientific Publishing Company, 1974.

[61] M. J. Hinich, "Testing for Gaussianity and linearity of stationary time series," *Journal of Times Series Analysis*, vol. 3, no. 3, pp. 169–176, 1982.

[62] L. Persson, "Statistical tests for regional seismic phase characterisations," *Journal of Seismology, Kluwer*, pp. 19–33, 2003.

[63] W. J. Conover, *Practical nonparametric statistics.* John Wiley & Sons, Inc., third ed., 1999.

[64] M.J. Wilmut, N.R. Chapman and S. Anderson, "Statistical properties of horizontal, vertical and omnidirectional underwater noise fields at low frequency," *Proceedings of OCEANS'93*, vol. 3, pp. 269–273, 1993.

[65] Lisa A. Pflug, Pamela Jackson, George E. Ioup and Juliette W. Ioup, "Variability in higher order statistics of measured shallow-water shipping noise," *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistic*, pp. 400–404, July 1997.

[66] S. Bates and S. McLaughlin, "Is VBR data non-stationary or self-similar," in *European Simulation Multi Conference*, (ESM), 1996.

[67] S. Bates, *Traffic Characterisation and Modelling for Call Admission Control Schemes on Asynchronous Transfer Mode Networks.* PhD thesis, The University of Edinburgh, 1997.

[68] Edward J. Wegman, Stuart C. Schwartz and John B. Thomas, ed., *Topics in Non-Gaussian Signal Processing.* New York: Springer-Verlag, 1988.

[69] J. Proakis and D. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications.* Prentice Hall, 1995.

[70] G. Gaunaurd and H. Strifors, "Signal analysis by means of time-frequency (Wigner-Type) distributions - applications to sonar and radar echoes.," *Proceedings of the IEEE, 84(9))*, pp. 1231–1248, Sept 1996.

[71] L. Cohen, "Time-frequency distributions-a review," *Proceedings of the IEEE*, vol. 77, pp. 941–981, Jul 1989.

[72] I. Daubechies, *Ten Lectures on Wavelets.* Philadelphia: SIAM, 1992.

[73] A. Cohen and J. Kovacevic, "Wavelets: The mathematical background," *Procedings of the IEEE*, vol. 84, no. 4, pp. 514–522, 1996.

[74] S. Pei and M. Yeh, "Improved discrete fractional Fourier transform," *Optics Letters, 22(14))*, pp. 1047–1049, July 1997.

[75] S. Pei and J. Ding, "Relations between fractional operations and time-frequency distributions and their applications," *IEEE Trans. on Signal Processing, 49(8))*, pp. 1047–1049, August 2001.

[76] C. C. Capus, *Time-Frequency methods based on the Fractional Fourier transform.* PhD thesis, Heriot-Watt University, 2002.

[77] I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 37, pp. 862–891, Nov 1989.

[78] D. Barry, "Fast calculation of the Choi-Williams time-frequency distribution," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 40, pp. 450–455, Feb 1992.

[79] D.L. Jones and T.W. Parks, "A resolution comparison of several time-frequency representations," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 40, pp. 413–420, Feb 1992.

[80] A.C.A. Figueiredo and M.F.F. Nave, "Improved time-frequency visualization of chirping mode signals in tokamak plasmas using the choi-williams distribution," *IEEE Transactions on Plasma Sciences*, vol. 33, pp. 468–469, Apr 2005.

[81] S. Mann and S. Haykin, "The chirplet transform: Physical considerations," *IEEE Trans on Signal Processing*, vol. 43, pp. 2745–2761, Nov 1995.

[82] A. Eskicioglu and P. Fisher, "Image quality measured and their performance," *IEEE Transactions on Communications*, vol. 43, pp. 2959–2965, December 1995.

[83] L. Cohen, *Time Frequency Analysis*. Upper Saddle River, New Jersey: Prentice Hall, Inc, 1995.

[84] Haldun M. Ozaktas, Zeev Zalevsky and M. Alper Kutay, *The Fractional Fourier Transform with Applications in Optics and Signal Processing*. Chichester, England: John Wiley & Sons, Ltd, 2001.

[85] S. Kay and G. F. Boudreaux-Bartels, "On the optimality of the Wigner distribution for detection," *IEEE transactions of International Conference on ICASSP*, vol. 10, pp. 1017–1020, April 1985.

[86] P. A. Rattey and A. G. Lindgren, "Sampling the 2-D Radon transform," *Signal Processing*, vol. 29, pp. 994–1002, October 1981.

[87] Daniel P.K. Lun, T.C Hsung and T.W Shen, "Orthogonal discrete periodic Radon transform. Part I: theory and realisation," *Signal Processing*, vol. 83, pp. 941–955, 2003.

[88] Daniel P.K. Lun, T.C Hsung and T.W Shen, "Orthogonal discrete periodic Radon transform. Part II: applications," *Signal Processing*, vol. 83, pp. 957–971, 2003.

[89] J. C. Wood and D. T. Barry, "Tomographic time-frequency analysis and its application toward time-varying filtering and adaptive kernel design for multicomponent linear-fm signals," *IEEE Transactions if Signal Processing*, vol. 42, pp. 2094–2104, Aug 1994.

[90] J. C. Wood and D. T. Barry, "Linear Signal Synthesis Using the Radon-Wigner Transform," *IEEE Transactions on Signal Processing*, vol. 42, pp. 2105–2111, Aug 1994.

[91] Haldun M. Ozaktas and David Mendlovic, "Fractional Fourier transforms and their optical implementation. I," *Journal of the Optical Soceity of America*, vol. 10, no. 9, pp. 1875–1881, 1993.

[92] Haldun M. Ozaktas and David Mendlovic, "Fractional Fourier transforms and their optical implementation. II," *Journal of the Optical Soceity of America*, vol. 10, no. 12, pp. 2522–2531, 1993.

[93] Adolf W. Lohmann, "Image Rotation, Wigner Rotation, and the Fractional Fourier transform," *Journal of the Optical Soceity of America*, vol. 10, no. 10, pp. 2181–2186, 1993.

[94] Ljubiša Stancović, Tatiana Alieva and Martin J. Bastiaans, "Time-frequency analysis based on the windowed fractional Fourier transform," *Signal Processing*, vol. 83, pp. 2459–2468, 2003.

[95] U. Ringstrom and L-E. Selin, *Vågrorelselara, Akustik, Optik.* Stockholm: Teknisk Hogskolelitteratur i Stockholm AB THS, 5 ed., 1989. In Swedish.

[96] M. H. Jones, *A practical introduction to electronic circuits.* Cambridge: CamBridge University Press, 1977.

[97] J.W.A. Fackrell and S. McLaughlin, "Robust nonparametric bicoherence estimation by stepwise outlier rejection," in *Electronics Letters*, vol. 36, pp. 368–370, Feb 2000.

[98] L. Persson, R.K. Lennartsson and S. McLaughlin, "Quadratic phase coupling analysis of passive sonar data using biphase technique," *Proc. IEEE/MTS Oceans, Providence, USA, 2000*, vol. 2, pp. 1053–1057, Sept 2000.

[99] S. Elgar, "Relationships involving third moments and bispectra of a harmonic process," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 1725–1726, Dec 1987.

[100] S. Elgar and R.T. Guza, "Statistics of bicoherence," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1667–1668, Oct 1988.

[101] D. R. Brillinger, *Time Series Data Analysis and Theory.* New York, NY: Holt, Rinehart and Winston, 1975.

[102] C. Nikias and J. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, pp. 10–37, July 1993.

[103] Vic Barnett and Toby Lewis, *Outliers in statistical data.* John Wiley & sons. Inc, third ed., 1994.

[104] N E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N-C. Yen, C.C. Tung and H.H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. London*, vol. 454, pp. 903–995, Jul 1998.

[105] P. Flandrin, G. Rilling and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters*, vol. 11, pp. 112–114, Feb 2004.

[106] G. Rilling, P. Flandrin and P. Goncalves, "On empirical mode decomposition and its algorithms," *Proceedings of IEEE EURASIP Workshop on Nonlinear Signal and Image Processing*, June 2003.

[107] P. Gloersen and N. Huang, "Comparison of inter annual intrinsic modes in hemispheric sea ice covers and other geophysical parameters," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 1062–1074, May 2003.