Automatic Detection of Emergency Cars

Master Thesis

Group 1061

Aalborg University
Acoustics and Audio technology

Electronics and IT
Aalborg University
http://www.aau.dk

AALBORG UNIVERSITY

Title:
Automatic Detection of Emergency Cars

Theme:
Master Thesis

Project Period:
Spring Semester 2017

Project Group:
1061

Participant(s):
López Montoro, Jonás
Kongsbak Drastrup, Mads Johan

Supervisor(s):
Flemming Christensen
Dorte Hammershøi

Page Numbers: 99

Date of Completion:
June 7, 2017

Abstract:

Inattention is a major factor in traffic accidents taking place in a third of all accidents in Denmark 2010 [1]. This inattention to surrounding traffic and distractions, such as music or hands-free communication, within the car cabin can make drivers unaware of nearby emergency vehicles. This gives the incentive to develop proof of concept that an automatic detection of emergency vehicles can be created. This project studies the characteristics of the siren wail of an ambulance. In field measurements of this siren are done both in stationary and moving scenarios, in order to establish signal description of this siren. By knowing the unique characteristics of the siren wail as well as the acoustical properties of a signal in motion (The Doppler Effect), lays the foundation in order to develop three algorithms: Siren Detection, Direction Determination and Distance Determination. These three algorithms are used to detect any nearby ambulances, determine whether the ambulance is approaching or receding and finally estimating the distance to the incoming ambulance. These three algorithms are tested using the recorded measurements in various scenarios, both stationary and in motion. The results indicated that the siren detection works perfectly with all signal samples, the direction algorithms is flawed in terms of determining direction based on the properties of the Doppler effect and the distance determination still needs improvement in order to determine distance accurately.

Contents

Preface

This report was submitted to the Master Program in Acoustics & Audio Technology as a master thesis at the Acoustics, Department of Electronic Systems at Aalborg University. The report represents the work of group 1061 during the period: 1 February - 8 June 2017.

Relevant readers for this report will have theoretical knowledge of acoustics, corresponding to what is expected at 10th semester of the Acoustics & Audio Technology program of Aalborg University.

Finally the authors would like to thank the supervisors Dorte Hammershøi & Flemming Christensen for their constructive supervision during the writing of this report. Furthermore the authors would like to thank the helpful people of the Falck Station Aalborg, who provided information regarding their issued ambulances, siren and usage, as well as making measurements possible of an ambulance's siren.

Aalborg University, June 7, 2017

López Montoro, Jonás
<jlopez15@student.aau.dk>

Kongsbak Drastrup, Mads Johan
<mdrast11@student.aau.dk>

1 | Introduction

Emergency vehicles use sirens and lights to alert nearby cars of emergencies, where they are at haste of moving quickly about in traffic. Despite the use of lights and sirens of high levels, lots of road users can still be unaware of emergency vehicles in traffic. This is due to a number of reason; Luxury seem to be quite isolating regarding exterior noise, while smaller and cheaper cars seem to be generally more noisy [2]. Furthermore many modern cars come with gadgets like GPS navigation, Hi-Fi systems and hands-free communication with telephones. Being in either an isolated or noisy car cabin can make it difficult to hear exterior sounds, such as a siren - especially if the driver is listening to either radio, music or telephone calls, which is the case for 47 % of drivers [3]. An entirely different aspect is the human. Humans get inattentive when driving. According to the council of safe traffic in Denmark (Rådet for Sikkertrafik) inattention was a major factor in third of all accidents in 2010 [1]. Furthermore both according to an article by Warncke Claus [4], as well as a statement from one of the paramedics from Falck [5], drivers are often so distracted that they end up blocking ambulances thus delaying their response time. Therefore it is very helpful for the human drivers, to alert the driver in the presence of an emergency vehicles with the sirens on. By automatically turning off/lowering the volume significantly of the Hi-Fi system and hands-free communication it creates less distractions for the driver. Getting the attention of these drivers in traffic is critical for the emergency vehicles to get to the desired location as fast as possible, without being involved in any crashes, collisions or obstructions.

Furthermore the uprising interest of autonomous cars have emerged in the last couple of years and it is greatly anticipated to be a great part of the (near) future of driving [6], [7]. Google cars (Waymo) have driven 3 million miles autonomously as of May 2017 [8]. As well as many semi-autonomous solutions already available in high-end luxury cars such as Mercedes and Tesla [9], [10], where the advanced auto pilot is able to follow the car in front without any interaction of the driver. Just as with human drivers, computers running these autonomous cars need to be aware of the surroundings in traffic - that includes emergency vehicles - so they can act accordingly.

So in any case, whether it being autonomous or human drivers, there is an interest to alert drivers of nearby emergency vehicles. Thus introducing an automatic method for detection of emergency vehicles. Emergency vehicles use both lights and sound to alert nearby road users of its presence. When an emergency vehicle is not within line-of-sight, other road users might rely on the siren to know of the incoming emergency vehicle in the vicinity. By knowing the acoustic properties of the siren and the effects of it being emitted by a moving vehicle (The Doppler Effect), it might be possible to determine the relative position and direction, thus flagging and alerting the driver of the nearby emergency vehicle, so that they can act accordingly.

This leads to the following questions:

- How can an automatic detection of emergency vehicles be designed?
- How can such a system be implemented in real-life scenarios and what are the limitations of its application?

An automatic detection system of emergency cars can have many applications, but this project will focus on the context of awareness of surrounding emergency vehicles in traffic, whether it being human or non-human drivers.

Project Purpose

Ultimately the overall goal is proof of concept, that a system as described can automatically detect and alert the driver of nearby emergency vehicles, by using a signal description of the siren in order to create an automatic detection algorithm. Furthermore by knowing the properties of the Doppler Effect it should be possible to determine whether the emergency vehicle is approaching or receding the observer, and ultimately determine the relative position, sub-sequentially altering the driver.

2 | Problem Analysis

This chapter will explain the thought process of identifying the problems and considerations in order to achieve the project goal. This project will consider an ambulance as the emergency vehicle at which the proposed system will work on. The following sections will describe the ambulance, the siren of the ambulance and finally the a signal analysis of the siren. Furthermore looking into the theory behind the Doppler Effect and the propagation of the siren, it should yield the needed knowledge in order to begin implementing a system, i.e. MATLAB program, which satisfies the project purpose.

2.1 Implementation Strategy

This section will describe considerations regarding strategy of the User Interface. Furthermore this section will describe further considerations on how the system will work and perform in certain contexts.

2.1.1 User Interaction

To fulfil the project purpose, the siren of an ambulance must be detected and sub sequentially the driver must be alerted. In order to rule out false positives and falsely alerting the driver, the siren detecting algorithm must run across several segments in order to be sure it is indeed a true positive. This implies a balance between tolerances for errors and the timing. If the algorithm investigates enough segments, it will be sure to determine whether a siren is present or not within that time period, but the timing of presenting this information is essential and could be vital for both driver and the people inside the incoming ambulance. Presenting too early might yield a false positive, while presenting later, with greater certainty, can be "too late". Either way, by alerting at a time where the driver is already obstructing the ambulance or the driver has become aware of the ambulance before the system, the system would not work so a compromise has to be made.

This project will not deal with the actual implementation of the described system within cars, but rather a proof of concept will be studied and tested. However, considerations regarding the physical implementation are still valid. Preferably the user interaction would take place using a car's own User Interface (UI), in the sense that all other warnings occur within this interface. Once a siren is detected in the vicinity, the interface would prompt a message stating that a siren has been detected, the music and hands free communication would be lowered or maybe even muted, while eventually pinpointing the estimated source position of the siren.

Figure 2.1: Concept art of how user interaction could be integrated into the existing display in a car.

This type of implementation is not within the scope of this project, so this yields for another, more primitive solution for user interaction and alert the driver of incoming ambulances in the vicinity. The obvious choice is to have a simplified GUI on a PC running real-time MATLAB, since this is the setup for siren detection anyway. This would essentially show the same principle as shown in figure 2.1, part off not being an integrated implementation with the ability to control the Hi-Fi system within the car. Instead it can alert the driver by prompting with a red light and a message stating, that a siren is detected, sub sequentially determining the distance and direction to the source and presenting it on the screen.

1. Upon start-up the MATLAB program will initiate a start-up procedure. This will yield that the program will have to process several segments in order to describe the background noise, creating a noise buffer and is therefore unable to detect any sirens within the first hundreds milliseconds on start-up.

2. Once the initial start-up sequence is done, the program will continuously look for the characteristics of the siren wail within the noise buffer, which is established by the background noise.

3. Upon detection of a siren in the vicinity, the program will flash a red screen with the warning “Siren Detected!”.

4. The program will determine the direction and distance of the siren. This will be presented with an arrow or a number 0-360 (like a compass) stating the direction, and a

number of the estimated distance in meters.

5. Once the siren is out of the reach, thus no longer detected, the program will reset and continue as in step 2.

2.2 Microphone & Placement

Sound input is acquired using a single microphone. For a real life implementation, the microphone within the car cabin used for hands-free communication could be used. Or even the microphone from a smart-phone placed within the car could be used. While these are in principle sufficient to gather the input for the system, these are placed within the same insulated car cabin, where music or hands-free telephone calls can be present. With a sufficiently strong signal analysis and system to detect a siren, it should still be able to outperform the human hearing. The optimal solution is to have an external microphone placed on the car, which then negates the attenuation of the car cabin and any unwanted noise within the cabin, while the placing the microphone on the exterior would introduce other noises like wind noise. Ultimately the overall performance should be superior to the human hearing, i.e. being able to tell if a siren is present in the vicinity before the driver is, for this project to make sense.

2.3 The Ambulance & Siren

The current line of ambulances in Denmark are based on a Mercedes Sprinter 319 CDI, with a siren from Federal Signal mounted above the license plate, shifted to the left from the centre, in order to be aimed in-line-of-sight of the driver position of any cars in front of the ambulance [11]. See Appendix A for more pictures and measurements of the siren.

Figure 2.2: Front of the ambulance. The siren placed in the left hand side, just above the license plate.

Figure 2.3: The back and right side of the ambulance. The model number 319 CDI can be seen on the back of the ambulance.

2.3.1 Siren: Horn Loudspeaker

Due to the construction of the horn loudspeaker, the particle velocity is more in phase with the diaphragm which yields higher real acoustical impedance and thus higher efficiency, compared to the normal loudspeaker where the particle velocity goes sideways, making the acoustical impedance more reactive and ultimately yields a much lower electroacoustical efficiency. The horn construction used in the danish ambulances is not definitively known, other than the manufacturer is Federal Signal [11].

A horn loudspeaker is used for sirens as they have high electroacoustic efficiency and can achieve high directivity based on the horn construction. The horn loudspeaker has much higher electroacoustic efficiency than a normal moving diaphragm loudspeaker. Compared to the moving diaphragm loudspeaker, the horn loudspeaker has a small diaphragm and a low mass.

Constructing a horn loudspeaker with maximum efficiency at any given frequency, means that the effective mechanical reactance of the system is equal to the effective mechanical resistance - this means, that the throat mechanical resistance must be proportional to the frequency, to obtain the maximum efficiency. The mechanical resistance of the exponential horn construction is independent of frequency, but it has a lower cutoff frequency at which it transmits nothing and remains purely reactive. For the canonical horn construction the efficiency is higher at lower frequencies and the efficiency drops as the frequency increases [12], [13], [14].

The horn has the highest frequency efficiency at 1-1.2 kHz, based on the measurements done in Appendix A. This could indicate that the horn construction is indeed an exponential design, but without certainty, it cannot be proven. Inquires to Falck have been made to get further detailed information on the specific siren used, but so far the siren speaker is seen as a "black box" and is operated with a ICS2010 Integrated System Controller, with a BCT500CBDK control head from Federal Signal to control both siren and lights on the ambulance [15].

2.4 Signal Analysis

In this section the different signals of sirens are going to be analysed in order to be able to identify the main characteristics of them so the system can detect the sirens and extract all the important information from them such as the distance and direction of the sound source.

An ambulance has different siren modes although there are two different ones that are the most used: the wail and the yelp. The signal analysis is based on the measurements described in appendix A.

2.4.1 Wail

The wail is the most common used siren of all. It consists of a repetitive frequency sweep of approximately 5.8 seconds between the frequencies of 600Hz and 1200Hz. A spectrogram of the siren is shown in the figure 2.4. As well as the main signal, higher harmonics can be seen which will help to extract some information. The energy in the lower frequencies correspond to the background noise caused by the vehicles other than the ambulance. It is also described in section 2.4.4.

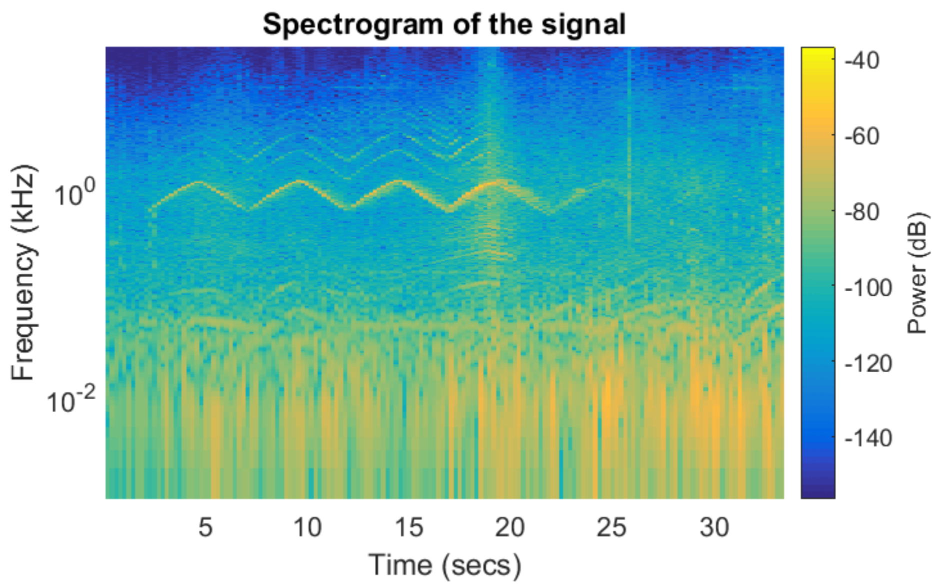

Figure 2.4: Spectrogram of the wail. The pattern of the frequency sweep is clearly visible.

In the figure below 2.5, the PSD of the siren can be seen. The main tone and its higher harmonics can be identified as well as the difference in power between them and the distance in frequency. This will be of extreme use to extract the information of distance and direction of the sound source.

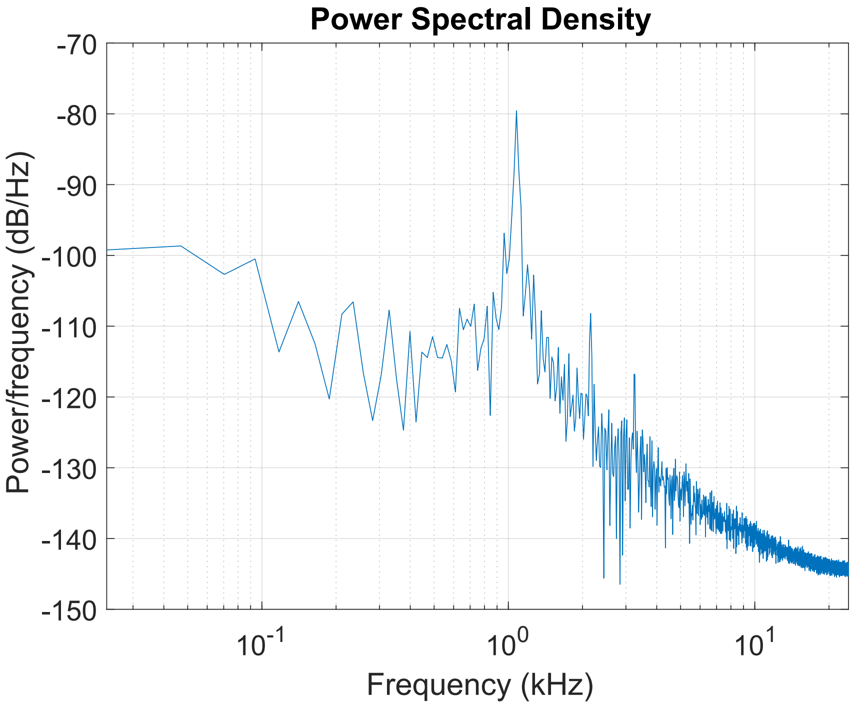

Figure 2.5: Power spectral density of 30ms of the wail.

The levels of the Wail siren are measured at 3 meters distance to the ambulance - from there the LC$_{peak}$ value is measured.

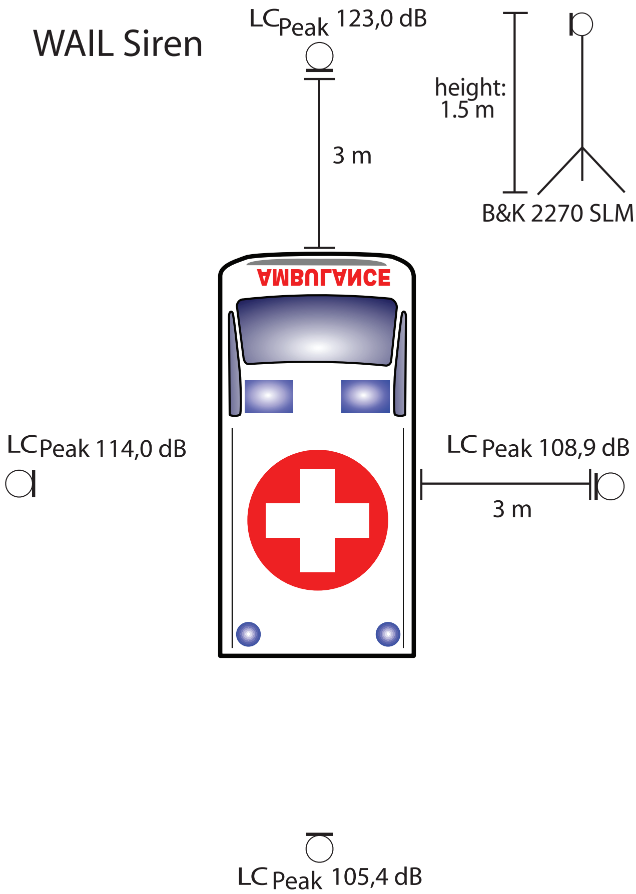

Figure 2.6: LC$_{peak}$ values of the wail siren measured at 3 meters distance to the siren from all sides.

2.4.2 Yelp

The Yelp is the second most used siren, specially when approaching an intersection. It consist of a much shorter frequency sweep of around 0,3 seconds between the frequencies of 600Hz and 1100Hz. In the figure 2.7 the spectrogram of the siren is shown where again the fundamental signal and its higher harmonics can be seen.

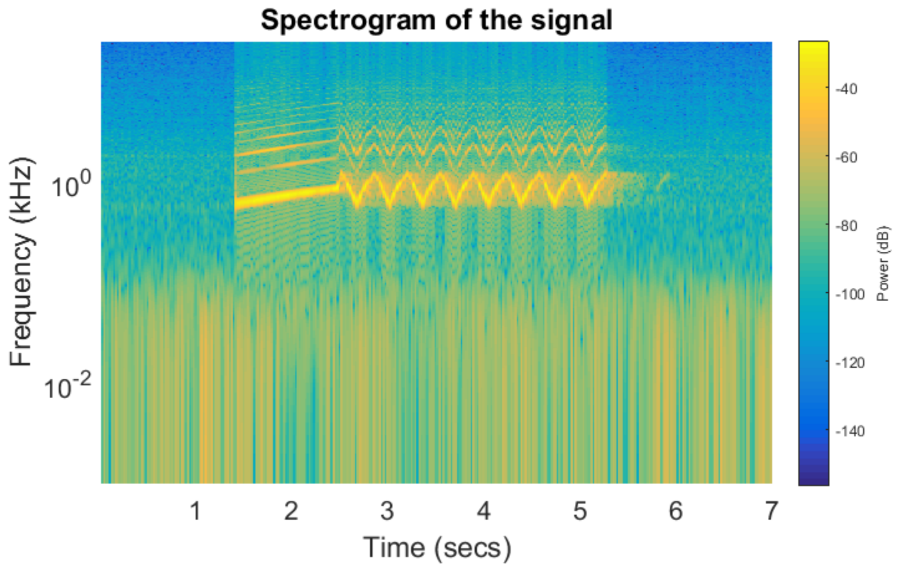

Figure 2.7: Spectrogram of the yelp. As the siren is turned on it "builds up" before doing the fast frequency sweep, which is clearly visible.

In the figure below 2.8 the power spectral density of the yelp siren can be seen. In this case the difference in level between the fundamental and the harmonics is lower than in the previous siren.

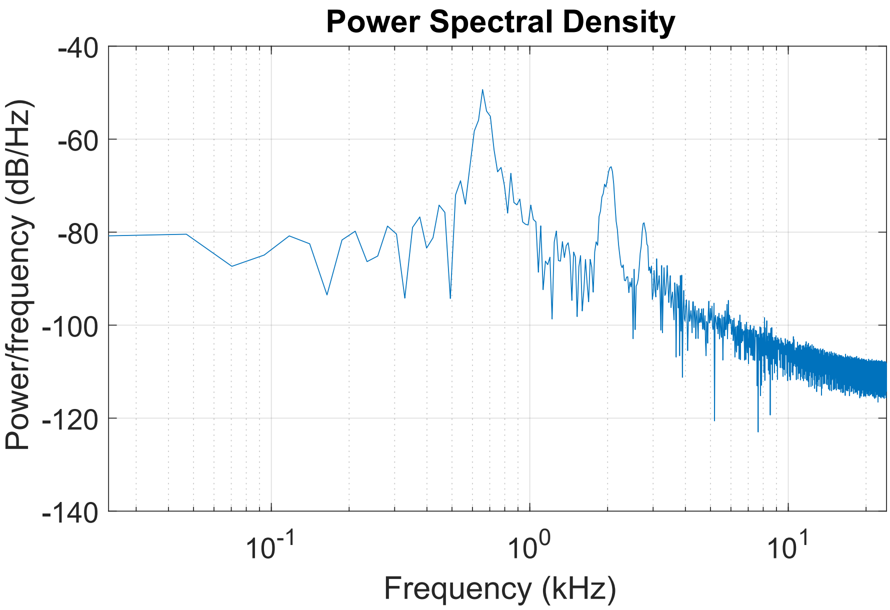

Figure 2.8: Power spectral density of 30ms of the yelp.

The levels of the Yelp siren are measured at 3 meters distance to the ambulance - from there the LC$_{peak}$ value is measured.

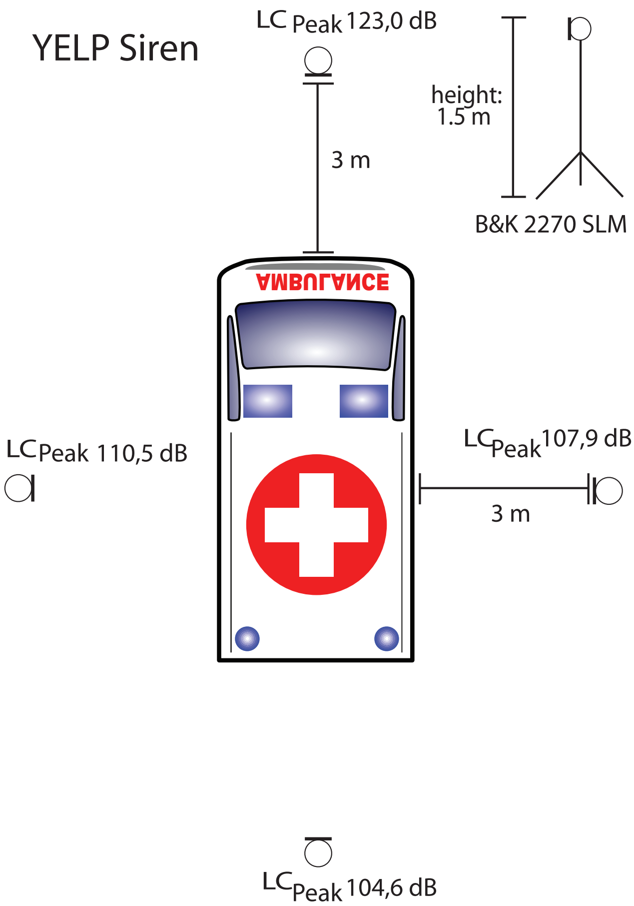

Figure 2.9: LC$_{peak}$ values of the yelp siren measured at 3 meters distance to the siren from all sides.

2.4.3 Directivity

The siren is placed just above the license plate, shifted off centre towards the driver position. Therefore a high directivity towards the left (driver's side) was assumed. The directivity is based on the eight levels measurements in the section above, including both siren types (wail/yelp).

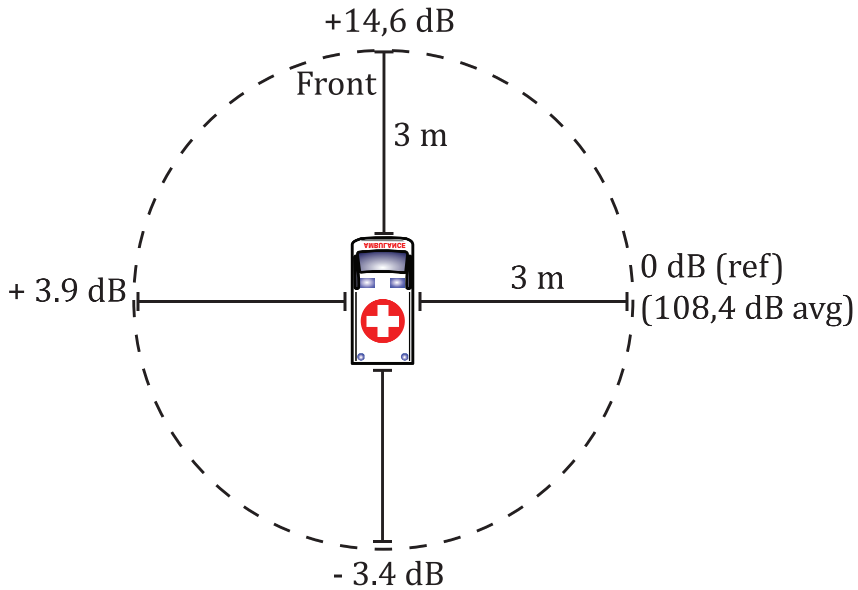

Figure 2.10: Because of the placement of the siren, there is a high directivity towards the left of the ambulance (driver's side).

2.4.4 Background noise

In order to identify all the characteristics of the measured signals, an analysis of the background noise also has to be done. In the figure below 2.11 a spectrogram of 35 seconds of background noise is shown. It can be seen that all the energy is concentrated in the lower frequencies which will not interfere with the emergency vehicles sirens, well below 600 Hz base tone of the siren.

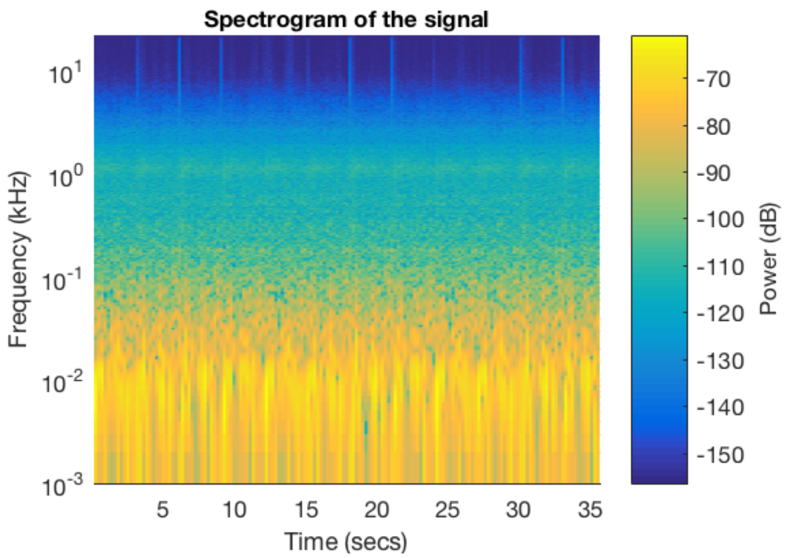

Figure 2.11: Spectrogram of the background noise. The noise mostly consists of low frequency noise at around 100 Hz and below.

2.5 Sound Propagation

This section will describe the sound propagation of the siren emitted by the ambulance.

2.5.1 Free Field Distance Approximation

As stated in the limitations in 2.9, this project will be limited to assuming siren detection in free-field conditions, since it is assumed that high diffusivity will impact performance severely. In a real world scenario the sound field will neither be true free- or diffuse field. Furthermore it can be assumed that there will be ground reflections, since the siren having wavelengths of up to 0.56 meter (from a base frequency of 600 Hz) which might introduce 1st order reflections since the placement of the siren is just above the license plate, ultimately yielding a sound field that is closer to half free field. In free field conditions the sound pressure decreases inverse proportionally to the distance to the sound source, thus giving a 6 dB decrease in sound pressure per doubling distance to the source, based on 2.1. While the horn construction of the siren will not yield omnidirectional sound propagation, due to the uncertainties and lack of documentation of the actual siren horn constructions, the high directivity of the horn will be ignored and it is assumed 6 dB decay per doubling distance.

$$20 \cdot log10(0.5) = -6.0 \tag{2.1}$$

Based on the measurements in Appendix A, the maximum (C-weighted) sound pressure level measured was 123 dB at 3 meters distance in front of the siren. Then it is possible to estimate the maximum distance of which it is possible to detect a siren, before the level of the siren drops below the background noise (traffic noise). Assuming average traffic noise of 60 dBA and idealised free field conditions:

$$\begin{aligned} r_2 &= r_1 * 10^{\frac{|L_1 - L_2|}{20}} \\ 4237 \text{ m} &= 3 \text{ m} * 10^{\frac{|123 \text{ dB} - 60 \text{ dB}|}{20}} \end{aligned} \tag{2.2}$$

Where:

- r_1 is the distance to the source. (Reference)
- L_1 is the SPL at r_1 (Reference)
- r_2 is the distance to the source from another point. (Maximum distance)
- L_2 is the SPL at the point r_2. (Background/Traffic noise level)

Thus a maximum range of 4237 m at which the sound will propagate to reach 60 dB in totally idealised conditions, before the siren is at the same level as the traffic noise thus no longer detectable. In real life conditions weather, wind, temperature, humidity and the surrounding environment will influence this range and a shorter range is to be expected [16].

2.5.2 Atmospheric Effects

Environmental factors and mechanisms that will affect the propagation of sound.

Air Absorption

Acoustic energy will be absorbed by the atmosphere because of molecular relaxation and viscosity effects. The greatest attenuation is found in the 10-30% humidity range and the attenuation is greater the higher the frequencies.

The effects of atmospheric conditions is under "normal" conditions largely insignificant, unless considering high frequencies over very long distances, which is not the case in this project. E.g. the attenuation due to air absorption of 1 kHz is ~0.4 dB per 100 meter while attenuation of 4 kHz is ~2.5 dB per 100 meter (for 20 °C at 70 % relative humidity) [17].

Wind and Temperature

The speed of sound depends on temperature - higher temperatures yield a higher speed of sound. The temperature in the atmosphere is not uniform, therefore local variations will occur. In normal weather conditions the temperatures are lower at higher altitudes, thus creating refraction or upward bents of the sound waves which could cause shadow zones in which sound does not penetrate. Furthermore scattering can occur which is the result of the sound waves being spread out in many other directions due to a local variation of air density or speed of sound.

Wind can greatly affect the sound propagation. When wind is blowing, there will be a wind (speed) gradient. The layer of air by the ground is stationary. Sound waves propagating upwind will bent upwards and viceversa. These wind gradients can have large effects on the otherwise predicted sound propagation considering only geometrical spreading atmospheric effects. Especially at longer distances (several hundred meters), this can be deemed problematic and will impact the overall range in which it is possible to detect the siren within this project [18].

2.5.3 Surface effects

Ground absorption

When sound is propagating close to the ground some attenuation will occur due to reflections. Since the siren is placed just above the license plate of the ambulance, reflections can occur due to the wavelength being 0.56 m (with a frequency of 600 Hz). A smooth, hard surface such as asphalt will yield very little absorption, compared e.g to grass. Much more important is the ground effect, where first order reflections can cause interference with the direct wave.

Barriers and trees

Propagation of sound through trees is mostly non problematic, unless it is the case of deep tree lines or actual forests. Solid barriers such as buildings are much more obscuring for the propagation of sound. These barriers are most effective at high frequencies since the low frequencies are easier diffracted along the edge of the barrier. The maximum attenuation of a barrier is around 40 dB due to the scattering of the atmosphere. Barriers can also hugely affect the temperature and wind gradients, which alone will affect the sound propagation a lot. The biggest concern regarding this project is considering streets in the city, where multiple barriers/buildings create multiple reflections from the source and ultimately yields a more diffuse sound field which makes it more difficult to detect the source position compared to free field [16], [19].

2.6 Doppler Effect

A sound source in motion will alter the wavelength and the received frequency of sound, although the emitted sound from the sound source is unchanged. When a sound source is in motion towards an observer, the (sound) waves are being pushed tighter together, thus resulting in the frequency going up and giving a higher pitch. The opposite occurs when the sound source moves away from an observer, resulting in a lower pitch. This phenomenon is called the Doppler effect [20].

$$f_{source} = \frac{c}{\lambda} \tag{2.3}$$

$$\begin{aligned} f_{approach} &= \frac{c}{\lambda_{approach}} = \frac{c}{c - v_s} f_{source} \\ f_{recede} &= \frac{c}{\lambda_{recede}} = \frac{c}{c + v_s} f_{source} \end{aligned} \tag{2.4}$$

Where:
f is the frequency of the source, the source approaching and receding an observation position respectably. c is the speed of sound. v_s is the relative velocity of the sound source in motion.

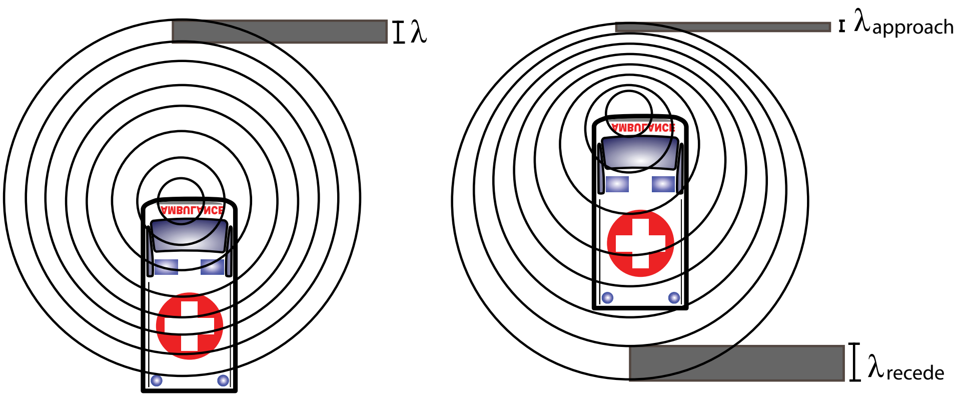

Figure 2.12: Illustration showing the Doppler Effect in practice of a moving ambulance compared to stationary.

2.6.1 Doppler Shift of the Harmonics

Based on the formula for the Doppler Effect in 2.4, it is evident that the speed of the source will affect the apparent frequency. Therefore different scenarios have been simulated in the tables below, showing the apparent frequency for a generic 1 kHz tone and a siren with a fundamental tone of 704 Hz, as well as the harmonics from the measurements done in Appendix A. The Doppler shift is calculated for relative speed of 10 m/s (36 km/h) and 20 m/s (72 km/h), showing a 3 and 6 percent difference compared to the apparent frequency without the Doppler Effect. This difference due to the Doppler Effect is constant regardless of frequency and increases by 3 percent with each increase of 10 m/s relative speed of the source.

So in essence the Doppler Effect will spread out the harmonics when the source is approaching an observer, and opposite it will cramp together the harmonics when the source is receding an observer. These properties can be used in order to determine the direction of the source relative to the observer.

Table 2.1: Example of Doppler Shift for a 1000 Hz tone approaching an observer

Approach	[Hz]	Shift 10 m/s [Hz]	Shift 20 m/s [Hz]	Diff 10 m/s [%]	Diff 20 m/s [%]
Fundamental	1000	1030	1062	2.96	6.01
2nd harmonic	2000	2060	2125	2.96	6.06
3rd harmonic	3000	3090	3187	2.96	6.04
4th harmonic	4000	4121	4250	2.98	6.06
5th harmonic	5000	5151	5312	2.98	6.05
6th harmonic	6000	6181	6375	2.97	6.06

Table 2.2: Example of Doppler Shift for a 1000 Hz tone receding an observer

Recede	[Hz]	Shift 10 m/s [Hz]	Shift 20 m/s [Hz]	Diff 10 m/s [%]	Diff 20 m/s [%]
Fundamental	1000	971	944	-2.94	-5.76
2nd harmonic	2000	1942	1889	-2.94	-5.71
3rd harmonic	3000	2914	2833	-2.91	-5.73
4th harmonic	4000	3885	3777	-2.92	-5.73
5th harmonic	5000	4857	4722	-2.90	-5.72
6th harmonic	6000	5828	5666	-2.91	-5.73

Table 2.3: Example of Doppler Shift for siren approaching an observer

Approach	[Hz]	Shift 10 m/s [Hz]	Shift 20 m/s [Hz]	Diff 10 m/s [%]	Diff 20 m/s [%]
Fundamental	704	725	748	2.94	6.06
2nd harmonic	1360	1401	1445	2.97	6.06
3rd harmonic	2067	2129	2196	2.96	6.05
4th harmonic	2770	2853	2943	2.95	6.06
5th harmonic	3476	3581	3693	2.98	6.05
6th harmonic	4152	4277	4411	2.97	6.05

Table 2.4: Example of Doppler Shift of a siren receding an observer

Recede	[Hz]	Shift 10 m/s [Hz]	Shift 20 m/s [Hz]	Diff 10 m/s [%]	Diff 20 m/s [%]
Fundamental	704	683	664	-3.03	-5.85
2nd harmonic	1360	1321	1284	-2.91	-5.75
3rd harmonic	2067	2007	1952	-2.95	-5.72
4th harmonic	2770	2690	2616	-2.93	-5.72
5th harmonic	3476	3367	3282	-3.19	-5.74
6th harmonic	4152	4033	3921	-2.91	-5.72

As it can be seen in the tables, e.g in tables 2.3 and 2.4, there is an upwards or downwards shift in the apparent frequency compared to emitted (stationary) frequencies. This difference is the effect of the Doppler Effect. This means that by measuring the apparent frequency, it can be determined whether the siren is approaching, receding or stationary relative to the observer. This information will be a key factor in determining the direction of the siren.

2.7 Fundamental/Harmonics positioning

In this section, the difficulties found in order to identify the fundamental frequency and its harmonics will be assessed. In the figure below 2.13 the spectrogram of the wail siren at 3 metres distance while the ambulance is not moving can be seen. It is easy to identify the position of the fundamental frequency and each of the higher harmonics.

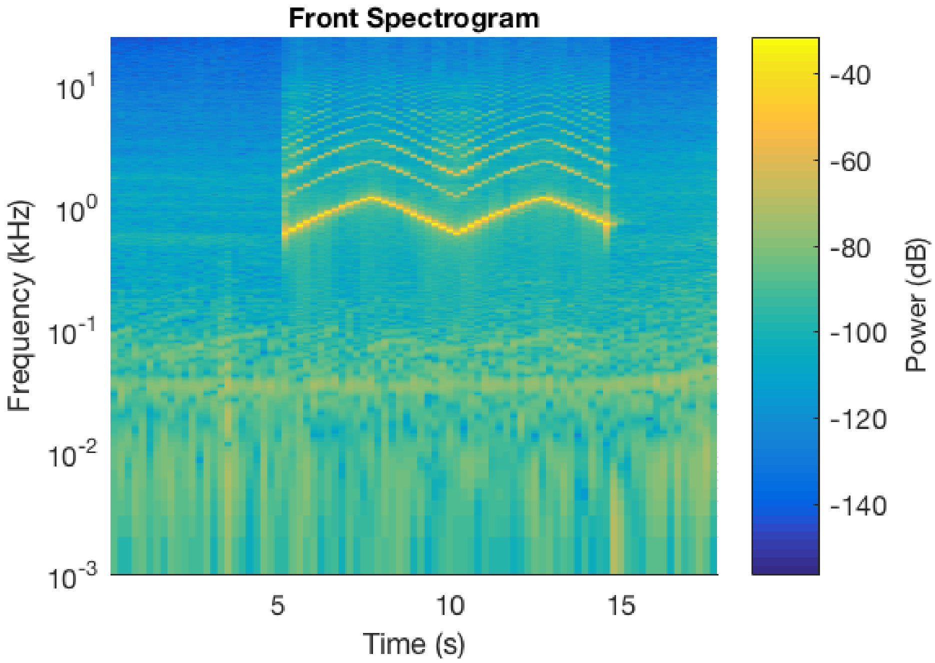

Figure 2.13: Spectrogram of wail siren while the ambulance is stationary.

Comparing the three figures below 2.14, 2.15, 2.16, it can be seen how the frequencies move up and down depending on the instant of the siren. Also, a really big difference between the magnitudes of the harmonics can be seen, even though the measurement conditions do not vary. This could make harder to detect the harmonics of the siren which would lead to a very difficult identification of the direction and speed of the emergency car. This difference in level may be caused because of reflections by nearby buildings.

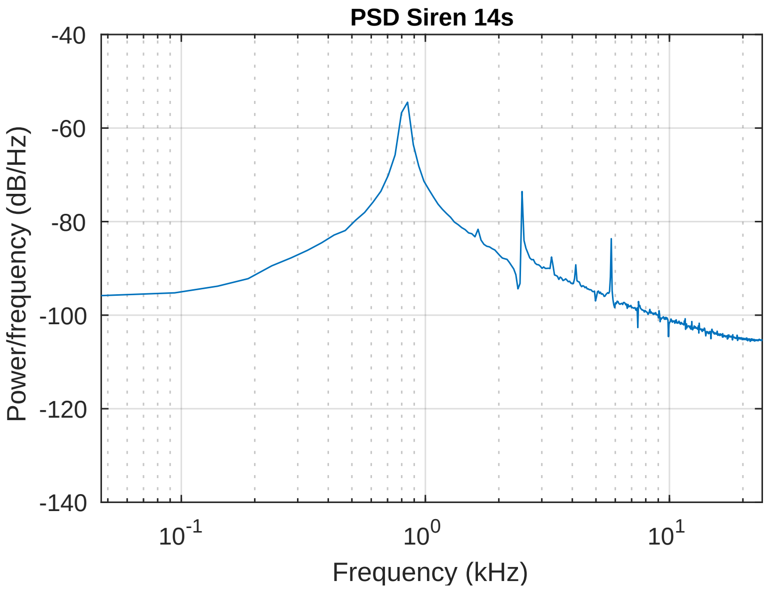

Figure 2.14: PSD of wail siren while the ambulance is stationary. Second 14. 1024 samples

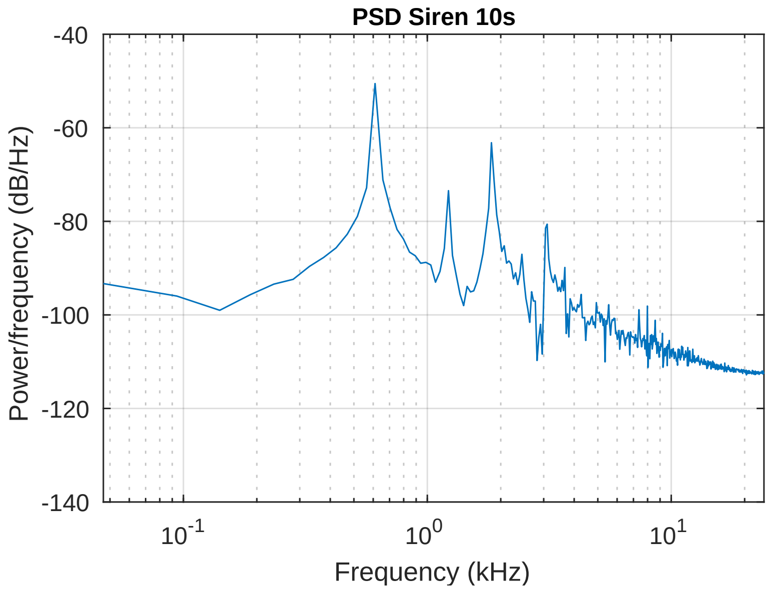

Figure 2.15: PSD of wail siren while the ambulance is stationary. Second 10. 1024 samples

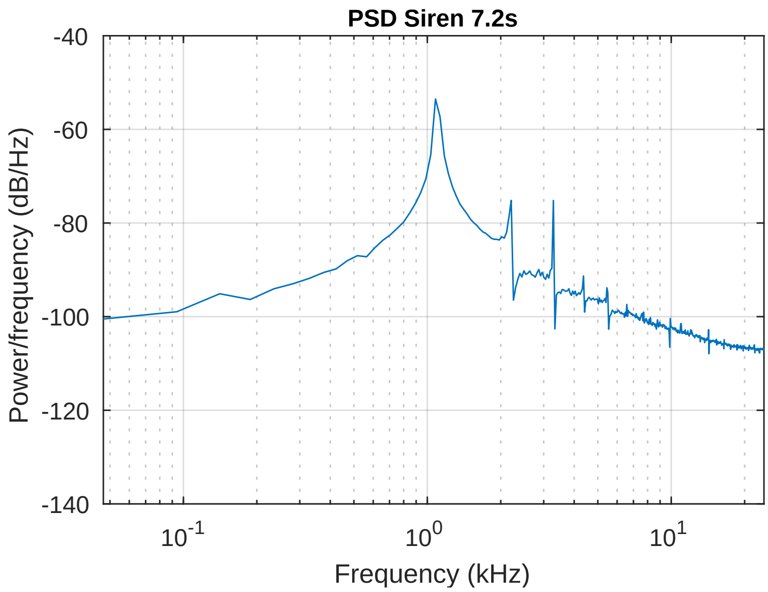

Figure 2.16: PSD of wail siren while the ambulance is stationary. Second 7.2. 1024 samples

The figures below 2.17 and 2.18, show a linear spectrogram of the siren while being stationary and moving respectively. It can be seen that the difference between fundamental and harmonics is constant for every instant of the siren.

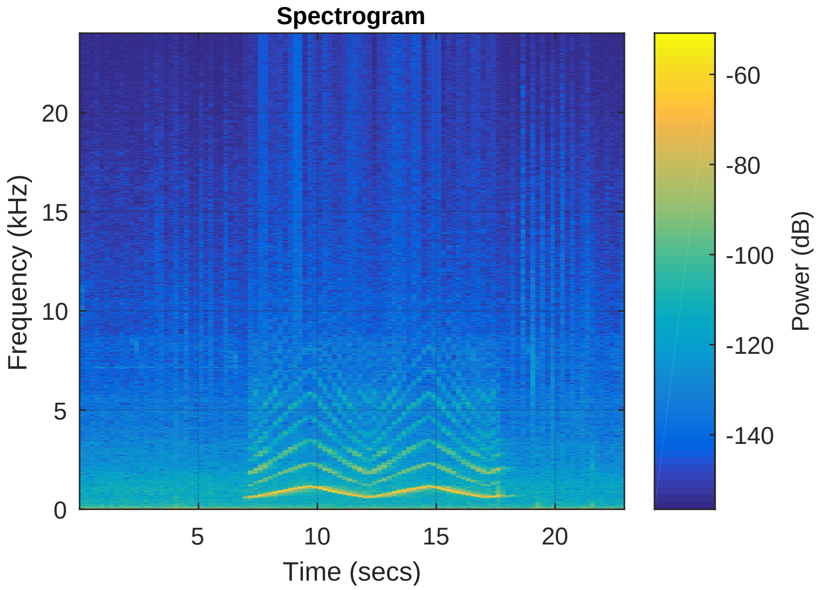

Figure 2.17: Linear spectrogram of wail siren while the ambulance is stationary.

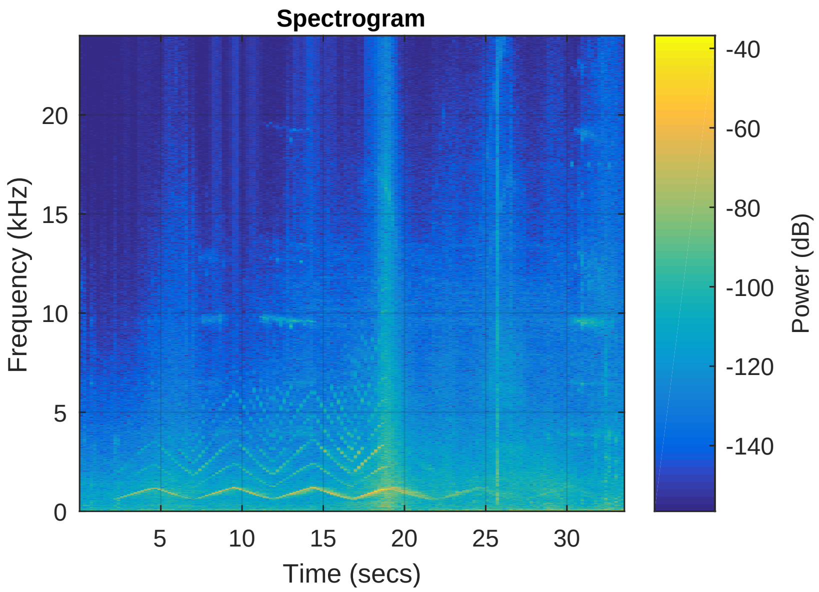

Figure 2.18: Linear spectrogram of wail siren while the ambulance is moving.

2.7.1 ARYule Method

In order to detect the signal fundamental frequency and its harmonics with more precision compared to the FFT, the ARYule method is used. This method computes the autoregressive all-pole model to fit the given data using the Yule-Walker equations [21]. Using only poles and no zeros to define the model allows to invest all the calculations to place where the signal peaks are, leaving the deeps out. However, this method does not detect properly the levels of the signal.

Comparing the figures below 2.19 and 2.20 it can be seen how the ARYule method only computes where the peaks are placed but giving a much more accurate result unlike the FFT method.

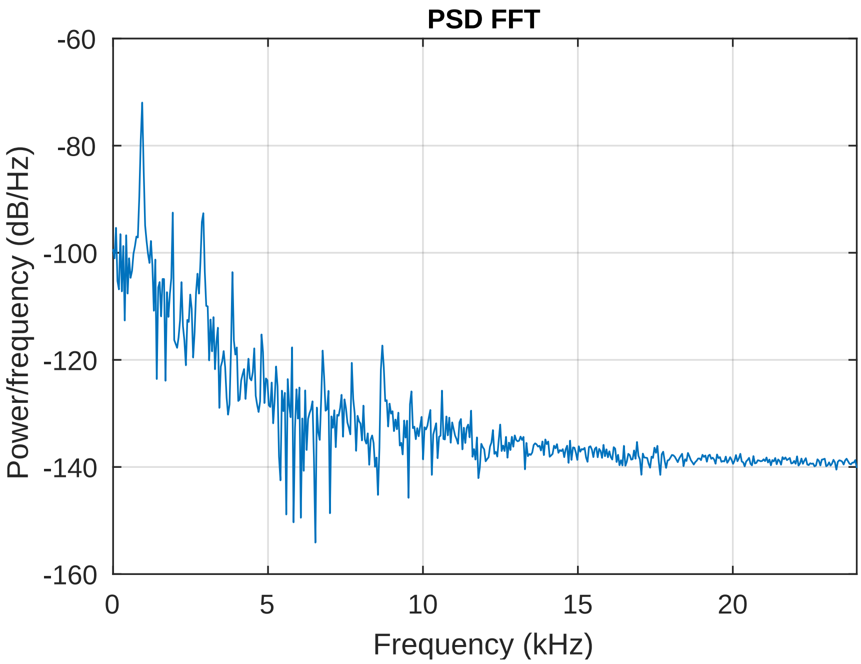

Figure 2.19: PSD of a segment using FFT method

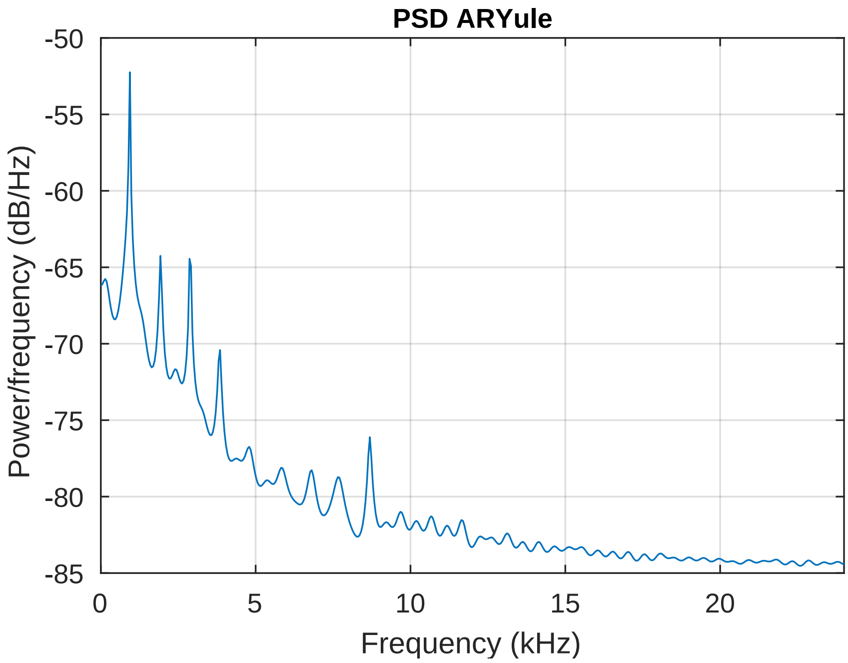

Figure 2.20: PSD of a segment using ARYule method

In the table 2.5 the frequencies detected by both method are shown. It can be seen that there is a difference between both of them.

Table 2.5: Frequencies detected by the FFT and ARYule method

	FFT	ARYule
Fundamental	939.3 Hz	937.5 Hz
2nd harmonic	1926 Hz	1922 Hz
3rd harmonic	2912 Hz	2859 Hz
4th harmonic	3851 Hz	3844 Hz

2.8 Scenarios

For this type of system to be successful in determining the siren of emergency vehicles, there are a lot of factors and different scenarios to consider, which can affect the success rate. The number of units, i.e sirens present, might very well be a limiting factor in the success rate. If there are multiple sirens from different origins, e.g. an ambulance coming from the front driving towards an observer, while another ambulance is coming from behind. This could prove to be problematic. The relative speed is also of essence of this project, since the Doppler Effect only occurs when there is a relative shift in speed from the ambulance compared to the observer. So if an observer is driving in a straight road with consistent speed, the ambulance would have to be driving faster in order for the apparent frequency of the siren to have an upwards shift, or slower but still approaching the car, to have a downwards shift of the emitted frequency. Using the properties of the Doppler Effect alone to determine the direction of an incoming ambulance can be proven to have limited usability in a real life application.

The environment has already been proved to cause possible problems in section 2.5.2. So atmospheric, meteorological and geographical factors all contribute to a decrease in performance of the system. The background and environment noise should not have a big impact, as the background noise will be considered in order to detect the siren, while a very diffuse sound field caused by multiple reflections from tightly packed buildings is almost certain to negate performance, as it might results in determining multiple sirens from multiple origin positions - just as with multiple ambulances present simultaneously.

In order to achieve the project goal of designing and implementing a prototype system for automatic detection of emergency vehicles, some limitations have to be established in order to be fulfilled within the project deadline.

2.9 Delimitations

- The emergency vehicle chosen for the initial system will be an ambulance. This is due to ambulances being deemed the most common of the emergency vehicles and the fact that lives are at stake, thus at most importance for the paramedics to get from A to B as fast and safe as possible. If the time frame of the project allows for it, other vehicles like police cars and fire trucks could be implemented in the system.
- Only the main siren type (Wail) will be consider for this project. The Wail is one of three siren types available with used siren and controller combination issued by Federal Signal in the danish ambulances. Stated by one of the paramedics is that the wail is almost exclusively used, while the yelp is only use occasionally. The third type is never used and not considered.
- The sound field will (most likely) highly influence the performance of directional pinpointing. A diffuse sound field, e.g. equivalent to being in the centre of a city, would yield lots of reflection and thus poorer performance compared to free field conditions. Thus the first priority is implementing the system in free field conditions. Diffuse field conditions will be done when free field conditions are deemed successful. A final solution should be able to handle both situations to be deemed robust and usable.
- Scenarios considering multiple sirens will be neglected in this project. The purpose is to have a reliable, robust determination of a single ambulance siren. While the system

built around this project might work with multiple sirens, it is not within the scope of this project and therefore there will not be focus on either having a robust system for multiple sirens or having a fail-safe system when multiple sirens are detected - instead this project is limited to scenarios with one ambulance siren present.

2.10 Summary

This section summarizes this chapter and outlines the analysis, investigation and concerns done in order to achieve the project goal as stated in 1. The project purpose is proof of concept that a system can be developed in order to automatically detect emergency vehicles and alert other drivers of their presence. The system should be able to detect the siren of an ambulance based on sound acquisition and using the properties of the signal to:

1. Determine whether a ambulance is present. (Siren Detection)
2. Determine the direction of the ambulance relative to the observer.
3. Determine the distance to the ambulance relative to the observer.

The siren detection will use the properties of the signal analysis done in section 2.4 and Appendix A. This project will only consider the wail siren tone, which is the main tone and is almost exclusively used. The wail consists of a repetitive frequency sweep of approximately 5.8 seconds between the frequencies of 600 and 1200 Hz. These properties will be used in order to establish an automatic detection algorithm of a siren with these acoustical properties.

The siren direction will use the level of the siren and properties of the Doppler Effect when the siren is in motion, which will result in a upwards or downwards frequency shift depending on if the source is approaching or receding the observer. The shift will be higher relative to reference/stationary as the velocity is higher of the source.

The distance to the siren is determined based on the free field approximation of sound propagation, which yields 6 dB per doubling distance. It is known that the assumption of the siren being a point source is incorrect due to the horn construction of the siren, and thus a high directivity meaning the sound propagation will not be completely omnidirectional/(hemispherical) with spherical sound waves. However, the sound decay might be in reality not a major importance, since it is a factor that can be changed or modified in future work.

These three objectives is what defines the overall project purpose. The next chapter will describe the considerations as well as the actual implementation of how these objectives are solved within MATLAB.

3 | Implementation

This chapter will describe the implementation within MATLAB. The accuracy of the methods will not be discussed in this chapter, all the performance analysis is done in section 4. Firstly pseudo codes will be presented, in order to let the reader get an general overview of how the system would operate. Sub sequentially each of the three different algorithms will be described. These algorithms are: Siren Detection algorithm 3.1, Siren Direction Detection algorithm 3.2 and Siren Distance Determination algorithm 3.3.

3.1 Siren Detection Algorithm

In this section an overview of the entire system is described in order to have a better understanding of the implemented algorithm. The general pseudocode is described in the listing below 3.1.

Listing 3.1: Pseudocode of the siren detection algorithm.

```
Noise buffer loading:
1 Sound Acquisition -> (1024 samples per segment)
2 Each segment is stored in a noise buffer of N segments
3 Average N segments in the noise buffer
4 Create Yulewalk filter from the buffer
5 Sound acquisition -> new segment
6 Apply designed filter
7 Find peaks (siren detection algorithm)
8 If no peaks are found
8.1 Segment is a possible noise
8.2 If determined as noise
8.2.1 Add segment to buffer (overwrite oldest input)
8.2.2 Average noise in buffer
8.2.3 Compute new filter
8.2.4 Return to point 5 (Sound Acquisition)
8.3 If determined as siren
```

```
8.3.1 Stored in siren array
8.3.2 Return to point 5 (Sound Acquisition)
9 If peaks are found
9.1 Segment is a possible siren
9.2 If determined as noise
9.2.1 Go to point 8.2.1
9.3 If determined as siren
9.3.1 Stored in siren array
9.3.2 Check direction of the source
9.3.3 Check distance of the source
9.9.4 Return to point 5 (Sound Acquisition)
```

For the siren detection algorithm the idea is to have a continuous sound acquisition with 1024 samples per segment. Each segment is stored in a N-segment noise buffer. An average of the N segments is calculated and used to create a Yulewalk filter - the yulewalk filter designs a recursive IIR digital filters using a least-squares fit to a specified frequency response [22]. Or in other words, the yulewalk filter is used to describe the noise buffer, that being the average background noise within the N-segments. The inverse of the calculated Yulewalk filter is applied to the input segment in order to have a "flat" response such that any peaks or irregularities will stand out as ”cups on a table”. Once the Yulewalk filter has been applied, an algorithm to find peaks will be run in order to determine if any peaks are "unwanted noise" (anything not related to a siren) or if in fact the peaks are related to a siren. If the segments are determined to not being a siren, then the current segment is added to the buffer, it overwrites the oldest input and again computes an average noise of the buffer, the Yulewalk filter is again calculated and updated to apply it to the following segments. If the segment is determined as siren, it is flagged and put under further investigation for several segments in order to be sure whether the flagged sound is indeed a siren. If the flagged sound is a siren, it will be analysed to determine the direction of the source, the distance relative to the observer and eventually return to the sound acquisition and repeat the process all over.

In the listing below 3.2, a more detailed pseudocode of how the siren is interpreted as noise or siren is described.

Listing 3.2: Pseudocode of the siren determination algorithm.

```
1 Find peaks
1.1 If no peaks have been found, it is a possible noise
1.1.1 If the previous 5 segments were noise or they were not
    possible sirens, the segment is determined as noise
1.1.2 Else, it is determined as siren
1.2 If peaks have been found, it is a possible siren
1.2.1 If the previous 10 segments were siren or they were not
    possible noise, the segment is determined as siren
1.2.2 Else, it is determined as noise
```

In order to detect peaks, a calibration has to be done to the system depending on the sensitivity of the transducer used to capture the sound. The system only uses one transducer so the calibration process only has to be done once. Once peaks have been or have been not found in the current segment, previous segments are taken into account to make the system more robust to noise. A number of five previous segments are used to determine if the current segment is noise although ten previous segments are used to determine if the current segment is noise. This difference is done in order to not give false detections.

In the figure below 3.1 the number of sirens detected is shown. Three different moments are seen: First segments where the siren is not present yet, siren detected when the siren is on and last segments where the siren has already faded away.

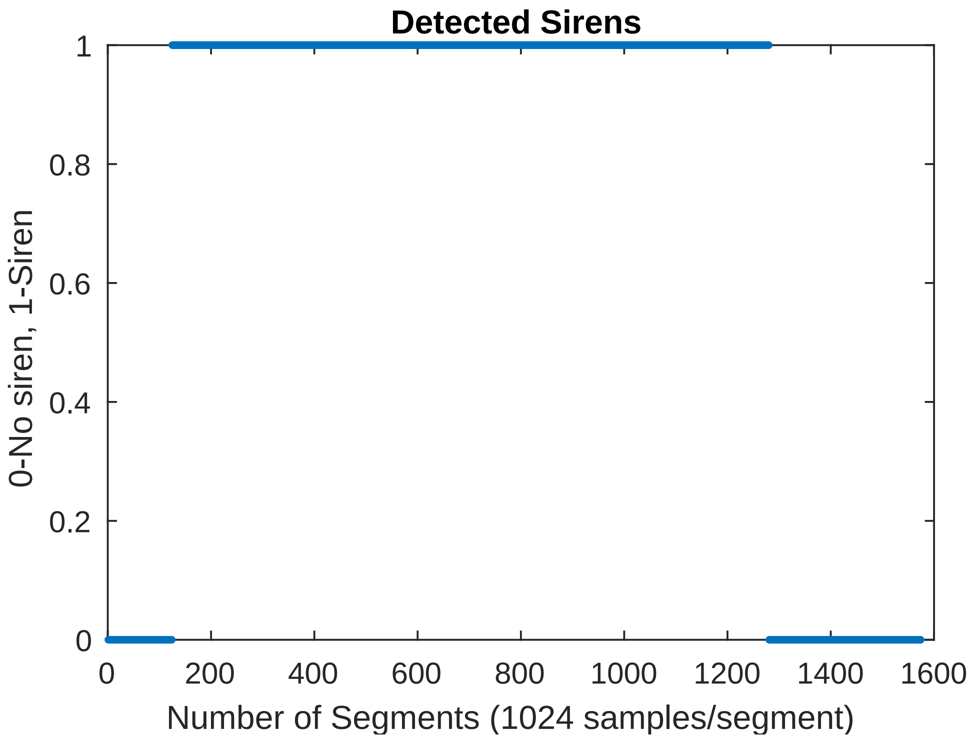

Figure 3.1: Detected Sirens.

3.1.1 Yulewalk Filter

The ideal situation to find peaks in a signal is to have a flat signal with only prominences of the desired peaks. This situation can be reached by filtering the signal with the inverse of its slope in order to make it flat, and the design of such filter can be done by using a Yulewalk filter. This filter has the advantage of just with the magnitude and frequency values of the desire frequency response it is able to achieve the wanted results.

In the figures below we can see the frequency response of the desired filter 3.2 and the final frequency response of the designed Yulewalk filter 3.3. The change in density of points perceived in the first figure is due to a larger number of points in the lower frequencies (frequencies where the siren is) is given to the Yulewalk filter than in the higher frequencies.

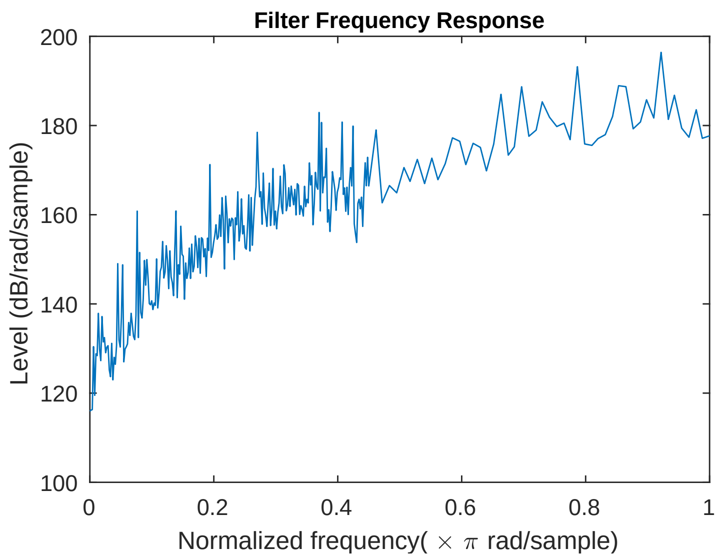

Figure 3.2: Desired Yulewalk filter Frequency Response.

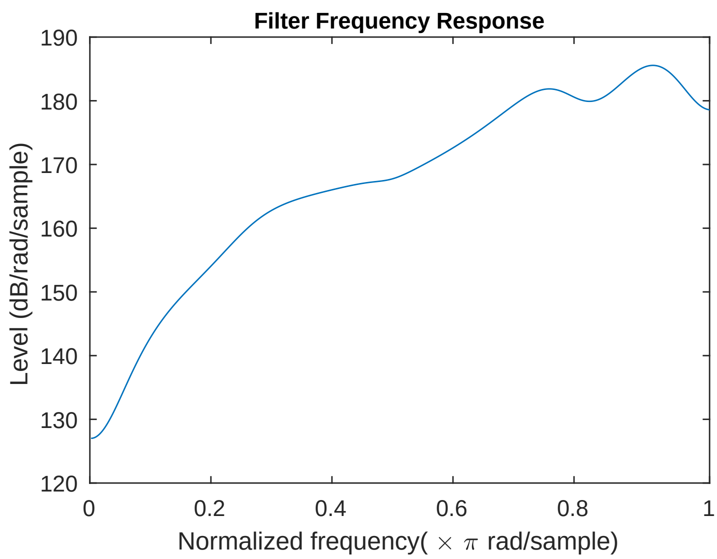

Figure 3.3: Final Yulewalk filter Frequency Response.

3.2 Direction Determination Algorithm

In this section, a description of the procedure followed in order to determine the direction of the sound source is described. Two different methods of detecting the direction of the sound source have been implemented. One using the level of the siren 3.2.1 and another one using the difference in frequency produced by the Doppler effect 3.2.2

3.2.1 Direction detection using level

In the listing below 3.3, a pseudocode of the sound source relative direction determination algorithm using the level of the signal is described. Note that the level of the siren varies depending on the moment of the period where it is observed.

Listing 3.3: Pseudocode of the siren relative direction determination algorithm using the level of the signal.

```
1 If at least a period of the siren has been analysed
1.1 If the maximum level of the current segment is greater
    than the mean of the maximum levels of the four segments
    in the previous period
1.1.1 The sound source is approaching the observer
1.2 If it is lower
1.2.1 The sound source is going away from the observer
1.3 Else, the direction is unknown
```

Once a siren has been detected, the maximum level of the current segment is obtained. Then, the level is compared with the mean values of the four segments located in the previous period of the siren. The levels cannot be compared between adjacent segments as they vary depending on the instant of the period of the siren, hence comparing them between same instants of different periods. If the current level is higher than the one of the previous period, the source is determine to be approaching the observer and viceversa.

The function below 3.4 corresponds to the one used to determine the direction of the ambulance.

Listing 3.4: Determination of the direction of the ambulance.

```
function pdir = getPDirection (i, Sirens, pSirens)
    pdir = -1;
    if i > 237 % At least a period of the siren
        if (~isempty(Sirens(i-234).siren) && ~isempty(Sirens(
            i-235).siren) && ...
                ~isempty(Sirens(i-236).siren) && ~isempty(
                    Sirens(i-237).siren))
```

```
                level = (Sirens(i-234).siren.Max + Sirens(i-235).
                    siren.Max + ...
                      Sirens(i-236).siren.Max + Sirens(i-237).siren
                          .Max)/4;
                if level < pSirens(i).siren.Max
                      pdir = 1;
                else
                      pdir = 0;
                end
            end
        end
end
```

In the figure below 3.4 can be seen the detected directions of the sirens. This algorithm needs some tuning as it gets some errors although it can be seen that there are three different stages: the first one when a period of a siren has not passed so the direction is unknown, the second one when the ambulance is approaching the observer, and the third when the ambulance is receding the observer.

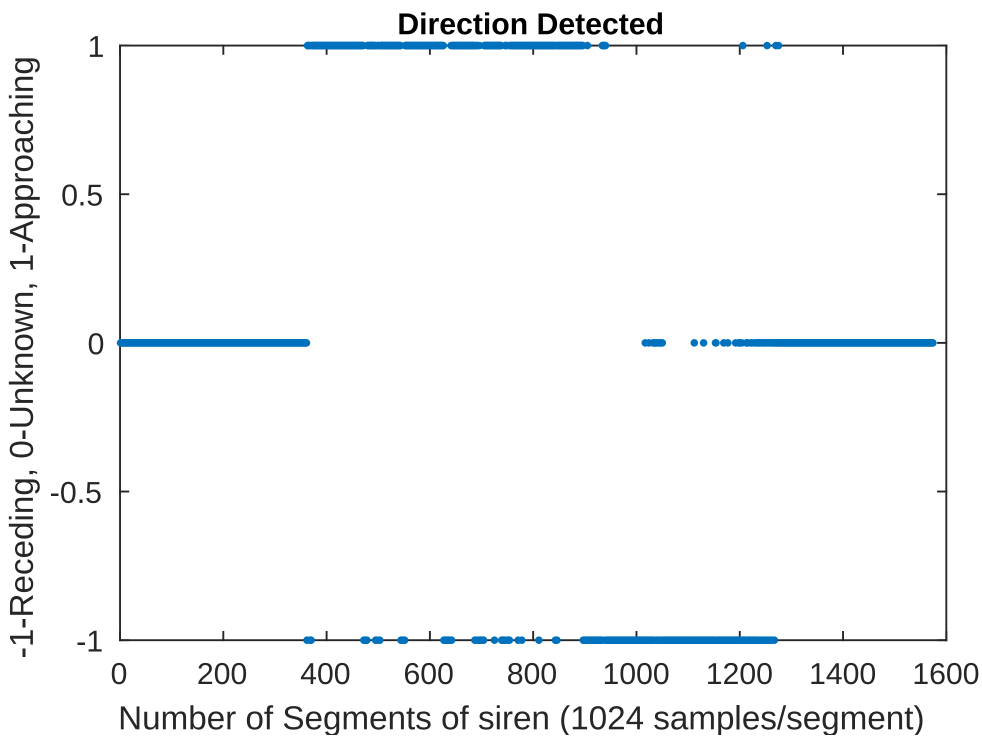

Figure 3.4: Direction of the detected sirens.

After adding several changes to the algorithm like taking into consideration previous direction detected the results improved considerably and can be seen in the figure below 3.5. The introduced changes can be seen in the function 3.5

Listing 3.5: Determination of the direction of the ambulance.

```
function direct = getDirection (i, pdirection, direction )
    if pdirection(i) == -1
        if direction(i-1) == 1
            direct = 1;
        elseif direction(i-1) == 0
            direct = 0;
        else
            direct = -1;
        end
    end

    if pdirection(i) == 0
        if direction(i-1)==0
            direct = 0;
        elseif direction(i-1)==-1
            direct = 0;
        elseif num1([pdirection(i-1) pdirection(i-2) pdirection(i-3) ...
                pdirection(i-4) pdirection(i-5) pdirection(i-6) ...
                pdirection(i-7) pdirection(i-8) pdirection(i-9) pdirection(i-10)]) >= 3
            direct = 1;
        else
            direct = 0;
        end
    end

    if pdirection(i) == 1
        if direction(i-1)==1
            direct = 1;
        elseif direction(i-1)==-1
            direct = 1;
        elseif num0([pdirection(i-1) pdirection(i-2) pdirection(i-3) ...
                pdirection(i-4) pdirection(i-5) pdirection(i-6) ...
                pdirection(i-7) pdirection(i-8) pdirection(i-9) pdirection(i-10)]) >= 3
```

```
                direct = 0;
            else
                direct = 1;
            end
        end
end
```

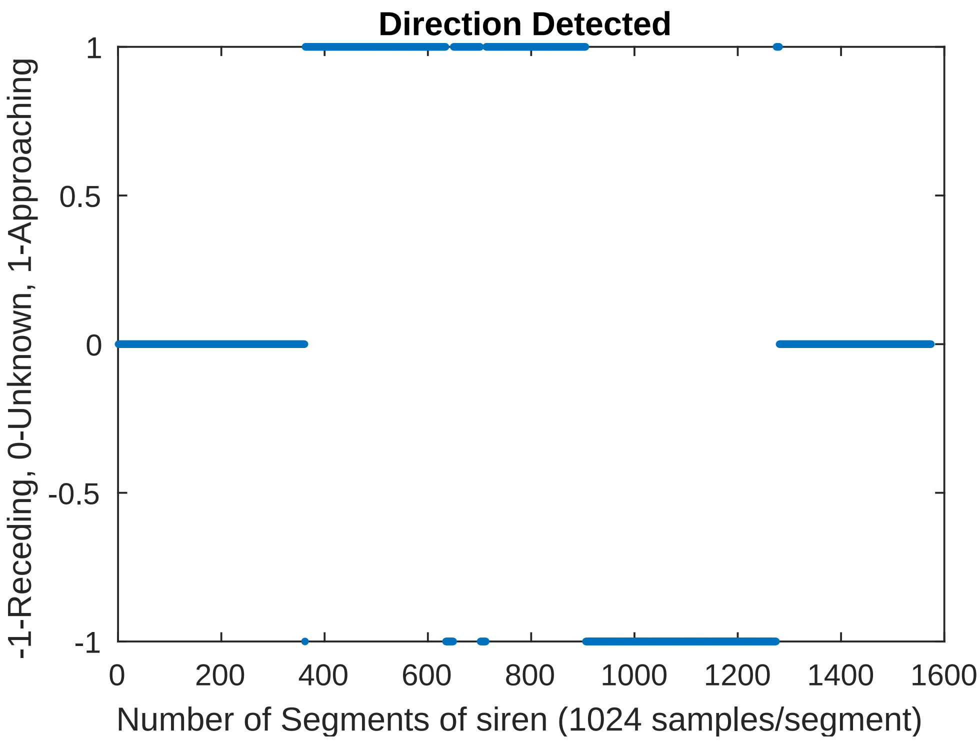

Figure 3.5: Direction of the detected sirens.

3.2.2 Direction detection using frequency

In this case, the direction detection is determined by the differences between fundamental and harmonics frequencies. It is shown in section 2.7 that said differences are constant for every instance of the siren. A pseudocode of this method can be seen in the listing 3.6. In order to acquire the exact frequencies, the method called ARYule is used. This method is described in section 2.7.1.

Listing 3.6: Pseudocode of the siren relative direction determination algorithm using frequency differences between fundamental and harmonics frequencies.

```
1 The difference in frequency between fundamental and
    harmonics for the current segment is obtained
2 The same is done for 20 segments before
3 The current value is compared with the previous ones
4 If it is higher than most of the previous ones
4.1 The source is approaching
4.2 Else, the source is receding
```

The results of using this method can be seen in the following figure 3.6.

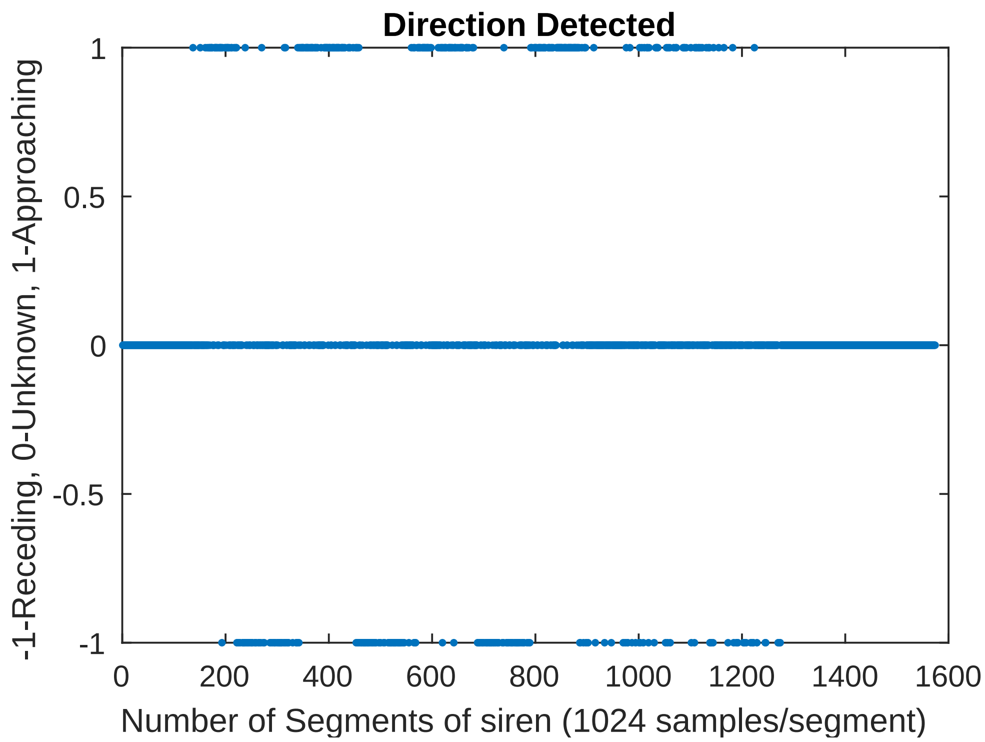

Figure 3.6: Direction of the detected sirens using the ARYule method. It can be seen as the result obtained varies exactly with the cycle of the siren, which yields an error stating the source as alternately approaching and receding.

From the figure above several conclusions can be extracted. The first one is that in a large number of segments, no two consecutive harmonics are found in order to obtain the difference between them so it can be compared with the previous segments, hence the large number of unknown results. And the second one is that the cycle of the siren seems to affect whether the algorithm detects if the ambulance is approaching or receding, which can be seen as the result obtained varies exactly with the cycle of the siren.

To assess the problem caused by the influence of the cycle of the siren another approach is taken. Once the difference has been acquire, the obtained value is then divided by the fundamental frequency in order to eliminate the influence of the cycle of the siren. In the listing below 3.7 the pseudocode of said approach is described.

Listing 3.7: Pseudocode of the siren relative direction determination algorithm using frequency differences between fundamental and harmonics frequencies.

```
1 The difference in frequency between fundamental and
    harmonics for the current segment is obtained
2 The obtained value is then divided by the fundamental
    frequency that would have had in case the source was
    stationary
3 If it is higher than 1
3.1 The source is approaching
3.2 Else, the source is receding
```

To obtain the difference in frequency between the fundamental and harmonics, first all the peaks in the segments are obtained, then it is determined which ones correspond to the fundamental and harmonics, later the differences between adjacent harmonics are computed, and lastly, the median value of all the obtained differences is calculated in order to make it more robust and consistent in case some differences cannot be obtained. In case that there are no differences detected, the direction is determined as unknown.

Once the difference have been obtained, it needs to be divided by the fundamental frequency which will yield to a value higher than 1 in case the source is approaching the observer and vice versa. To do so, first it is needed to compute the relationship between the fundamental frequency and difference obtained for a stationary source. This relationship is shown in the figure below 3.7.

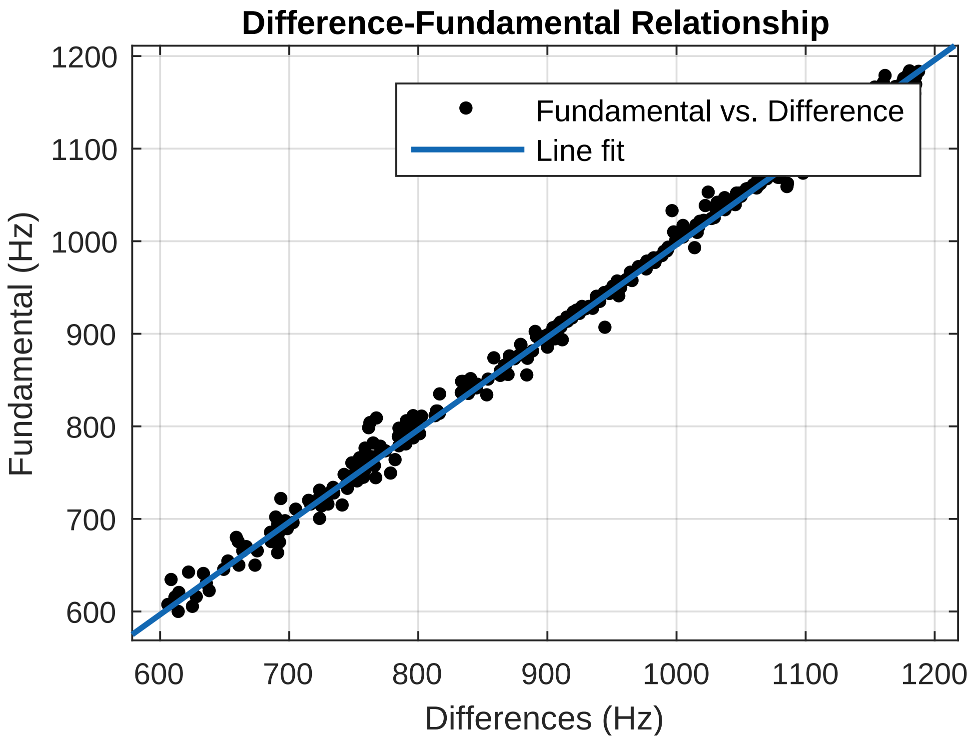

Figure 3.7: Relationship between the difference obtained and its fundamental for an stationary source. Fitting of the previous relationship. The relationship is used in order to divide obtained differences with the correct value, in order to assess the influence of the siren cycle.

A line is fitted into the points to later be able to divide the obtained difference by the correct value.

After having performed the previously described, the results are sorted into receding if lower than 1 and approaching if otherwise. The result obtained is shown in the figure 3.8 below. The influence of the siren cycle have been assessed, but the results are still not as expected. The source is still alternating between the approaching and receding which is clearly an error and does not correspond to reality.

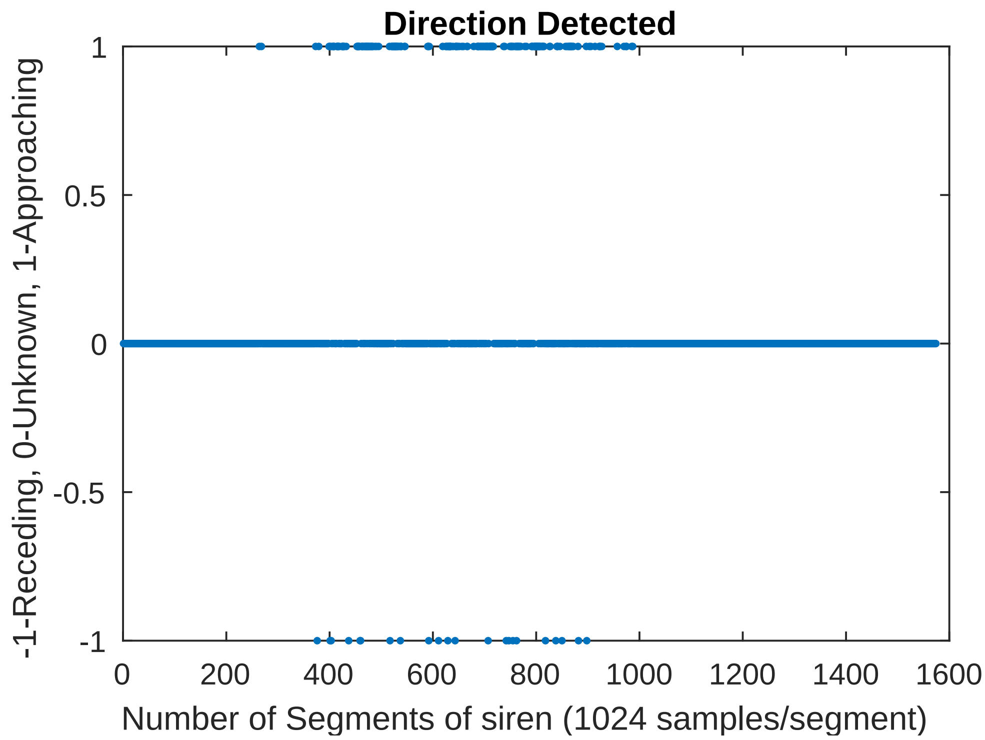

Figure 3.8: Direction of the detected sirens using the ARYule method. The influence of the siren cycle has been assessed, but the result is still not what is expected. The source is still alternating from approaching to receding, which is not the case in reality. The state is unknown if there is no difference detected.

3.3 Distance Determination Algorithm

In this section, a description of the procedure followed in order to determine the direction of the sound source is described. Two different methods have been implemented: one that takes into account all the obtained points, and another one that only uses the points where the horn's performance is optimal (between 1kHz and 1.2kHz).

Both methods have several things in common: They use the maximum level of the segment to determine the distance and they make use of an average between segments so they are more robust to noise.

In the listing below 3.8, a pseudocode of the sound source distance detection algorithm is described.

Listing 3.8: Pseudocode of the siren distance detection algorithm.

```
1 The maximum level of signal in the current segment is
   obtained
2 The direction of the source is taken into account (
   calculated before)
3 The distance of the source is obtained using one of both
   fitting depending on the direction
4 The calculated distance is averaged with previous segments
   to reduce noisy results
```

In the figure below 3.9 the maximum values of the detected sirens can be seen.

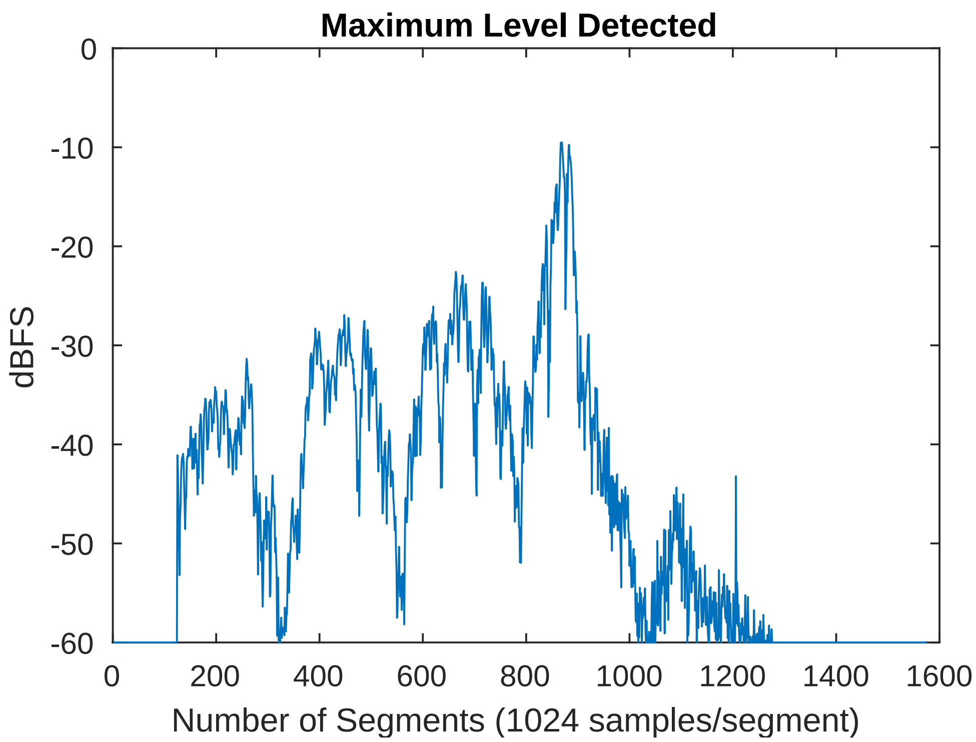

Figure 3.9: Maximum level of the detected sirens.

It can be observed that the level fluctuates depending on the moment of the cycle where it is measured. This fluctuation is caused by the difference in performance of the horn depending on the emitted frequency and makes it more difficult to determine the distance.

To determine the distance out of the level of the signal, two different curves have been fitted to the points obtained in the siren measurements A. Both fitted curves have a decay of 6 dB per doubling distance, however, one is used for when the source is approaching the observer and another one is used for when the source is receding. This is done in order to account for the directivity of the sound source 2.4.3. Both fitted curves can be seen in figures 3.10 and 3.11

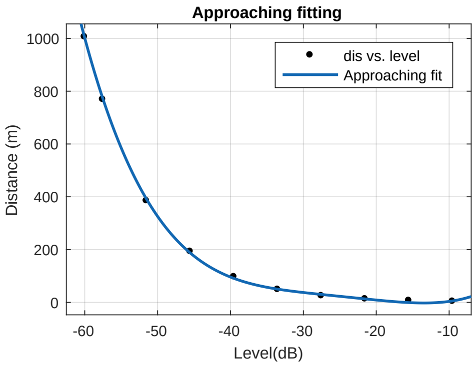

Figure 3.10: 4 degree polynomial fitting for approaching source.

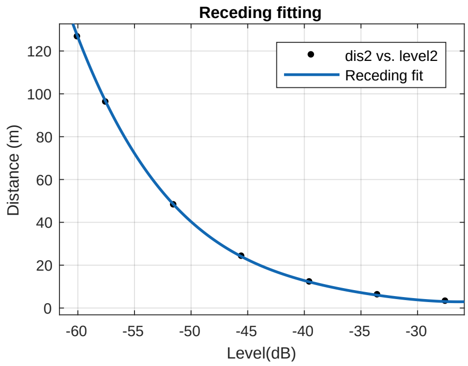

Figure 3.11: 4 degree polynomial fitting for receding source.

The function below 3.9 corresponds to the one used to determine the direction of the ambulance using only the points where the horn's performance is optimal . In order to use all the obtained points, only the condition where the points are discarded should be deleted.

Listing 3.9: Determination of the relative distance between the ambulance and the observer.

```
function distance = getDistance(fappr, frec, i, direction,
    pSirens)
    dire = direction(i);

    % Frequency where the max value is located
    peakFrec = pSirens(i).siren.LOCS(find(pSirens(i).siren.
        PKS==pSirens(i).siren.Max));

    % Range of frequencies where the horn is tuned to its
        optimal
    % performance
    if peakFrec > 1 && peakFrec < 1.2
        if dire == -1
            distance = NaN;
        end

        if dire == 1
            distance = fappr(pSirens(i).siren.Max);
```

```
            end

            if dire == 0
                distance = frec(pSirens(i).siren.Max);
            end
        else
            distance = NaN;
        end
end
```

In the figure below 3.12, the obtained distances without doing any averaging and for all the points are shown. Very erratic values can be seen and this is much improved first by averaging and then by only taking the horn's best performance range points.

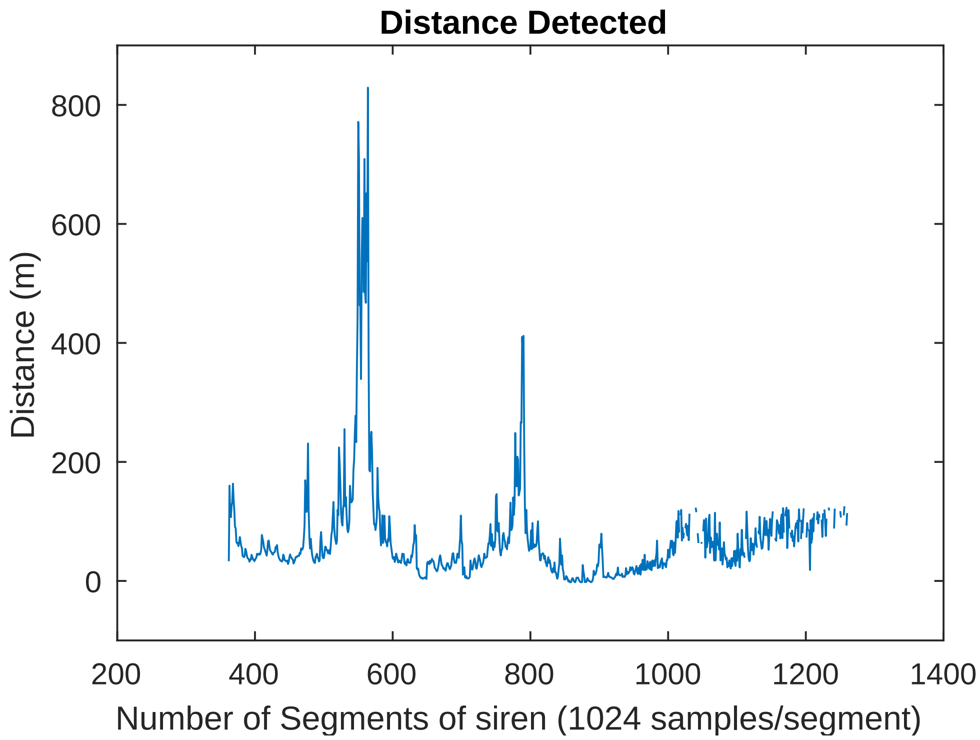

Figure 3.12: Distances detected for every segment.

In the figure below 3.13, the results obtained by using only the averaging are shown.

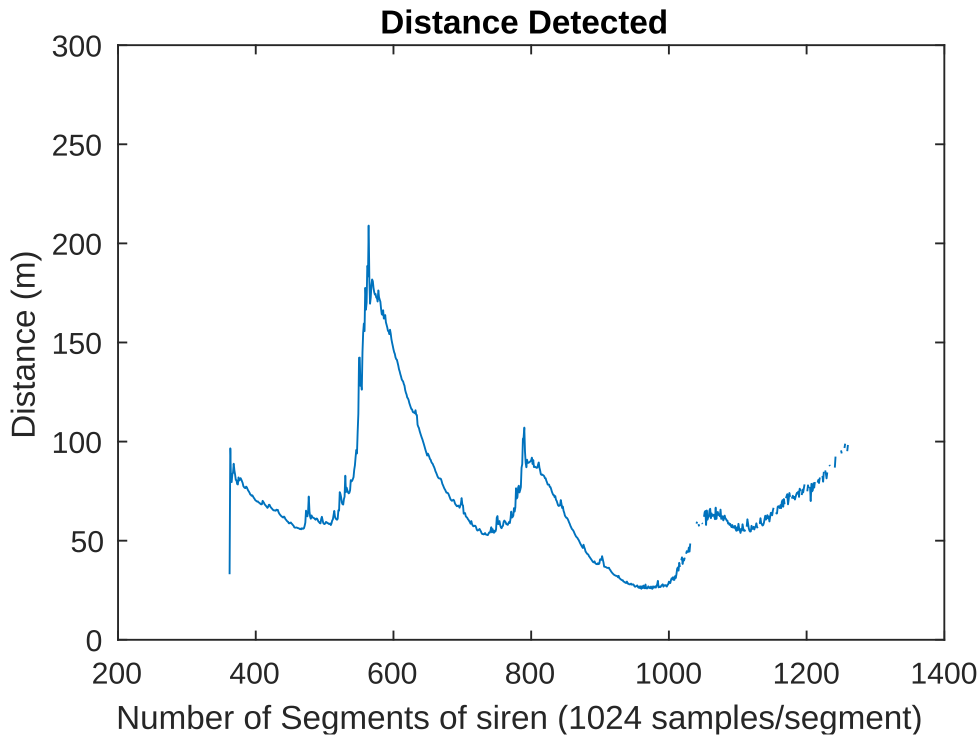

Figure 3.13: Distances detected with averaging.

In contrast with the results seen in figure 3.12, the far away distances that were wrong have disappeared and can be seen how the ambulance is approaching the observer until around the segment 950 and afterwards the distances start growing back as the ambulance recedes.

In the figure below 3.14, the results obtained by using only the optimal points plus the averaging are shown.

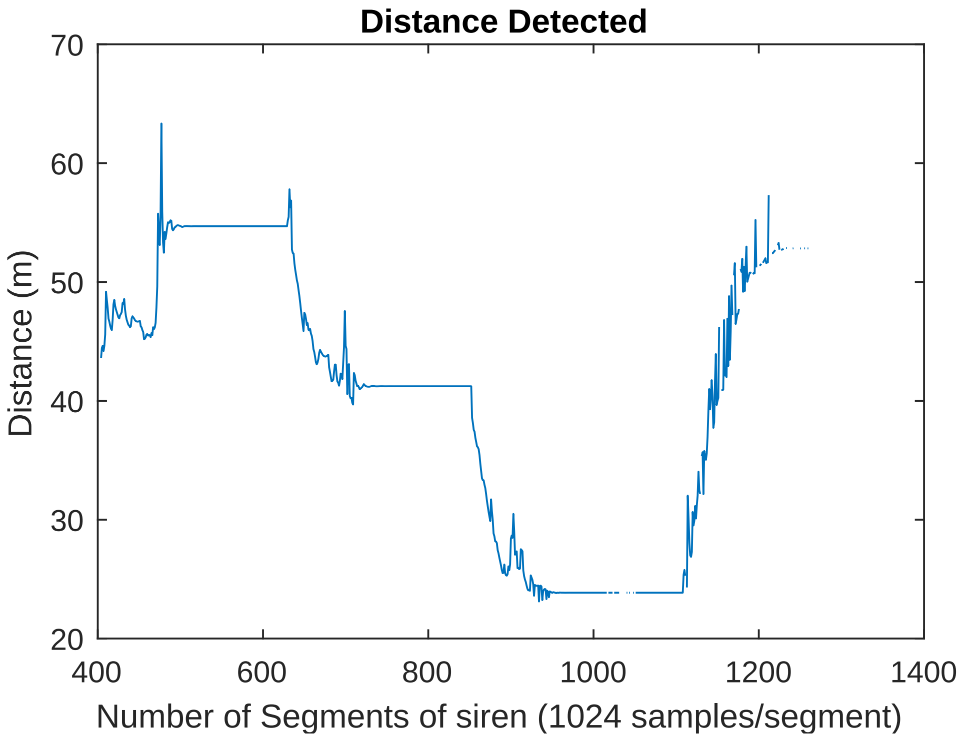

Figure 3.14: Distances detected with averaging.

In this case, the distances are much more consistent with the expected ones although there are some points where the distance is constant as we are taking the previous obtained value for the points out of the horn's optimal performance range.

3.4 General run of the code

In this section, a general run of the code will be shown where one of the recorded signal is used. The code is run segment by segment of 1024 samples/segment as it would be done in real-time. The code makes use of all the previous implemented algorithm showed in sections 3.1, 3.2 and 3.3.

Listing 3.10: The noise is loaded to the system.

```
%% Load file
filename = '/Users/Jonas/Desktop/svnP10/Siren recordings/Cuts
    /SLMCut.wav';
[sound, fs] = audioread(filename);
```

Listing 3.11: Global variables are created.

```
%% Global variables
currentSound = [];
noiseBuffer = [];
finalNoise = [];
sirenDetection = [];
bufferPos = 0;
buffersize = 5;

pSirens = struct([]);
Sirens = struct([]);
pNoises = struct([]);
Noises = struct([]);
filters = struct([]);
Pks = struct([]);
Locs = struct([]);
pdirection = [];
direction = [];
distance = [];
```

Listing 3.12: Window size and number of loops are calculated.

```
%% File windowing
window = 1024;
loop = round(length(sound)/window,0)-1;
```

Listing 3.13: First noise buffer is filled and averaged.

```
%% Noise buffer loading
noiseBuffer(:,1) = sound((1:window),1);

```

```
for n = 2:buffersize
    noiseBuffer(:,size(noiseBuffer,2)+1) = sound((((n-1)*
        window)+1:n*window),1);
end
finalNoise = sum(noiseBuffer,2)./buffersize;  % Average
    noises
```

In the figure below 3.15, the first averaged noise PSD is shown. It will be used to design a filter to compensate for the noise buried in the signal.

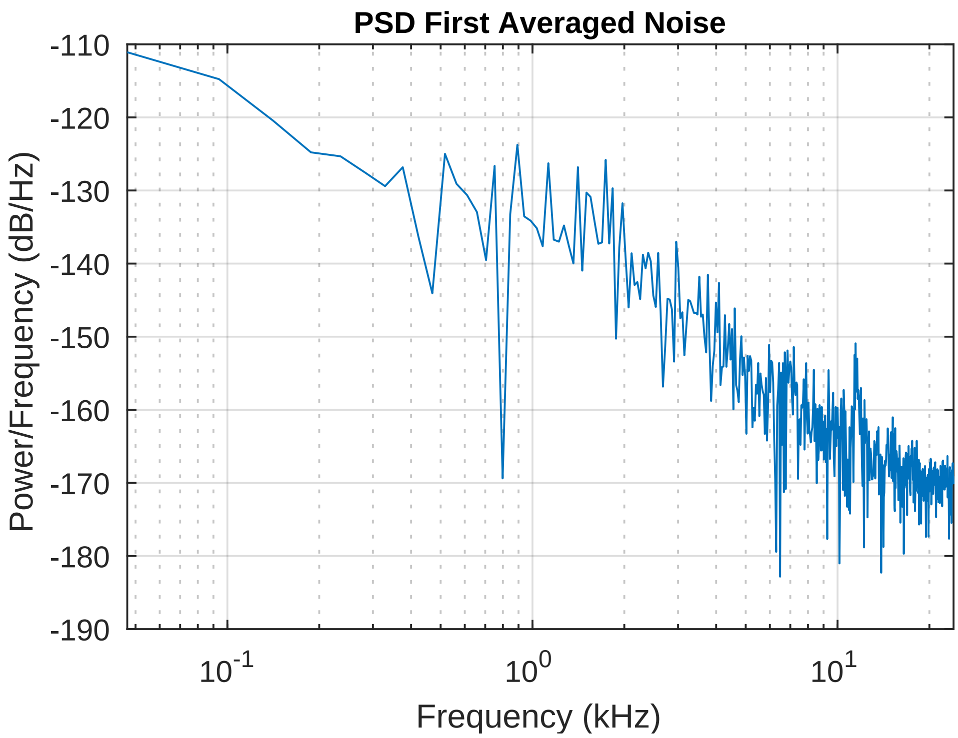

Figure 3.15: Power Spectral Density of the First Averaged Noise.

Listing 3.14: Yulewalk filter is computed.

```
%% Filter design

[Hz, filters(1).filter] = designFilter(finalNoise, Xaxe, fs);
```

The code below 3.15 corresponds to the function used to design the Yulewalk filter. It uses the magnitude values of the averaged noise's frequency response in order to design the filter. A greater number of the magnitudes in the frequencies below 10kHz are taken as it is the most important range to describe the signal. In figure 3.16 the desired filter can be seen and in figure 3.17 the final designed filter is shown.

Listing 3.15: Design of the Yulewalk filter.

```
function [Hz, filter] = designFilter( finalNoise, Xaxe, fs )
    N = 10;
    const=fs/2;
    % Power spectral density
    nfft = 2^nextpow2(length(finalNoise));
    Pfs = abs(fft(finalNoise,nfft)).^2/length(finalNoise)/fs;
    HpsdIS = dspdata.psd(Pfs(1:length(Pfs)/2),'Fs',fs);

    % Yulewalk filter desing
    data = 10*log10(HpsdIS.data);
    points = round(linspace(1, 230, 230),0);
    points2 = round(linspace(231, 512, 50),0);

    F = [(Xaxe(points)*1000)/const (Xaxe(points2)*1000)/const
        ];
    M = [-data(points)' -data(points2)'];

    filter = struct('F',F,'M',M);

    Hz = yulewalk(N,F,M);
end
```

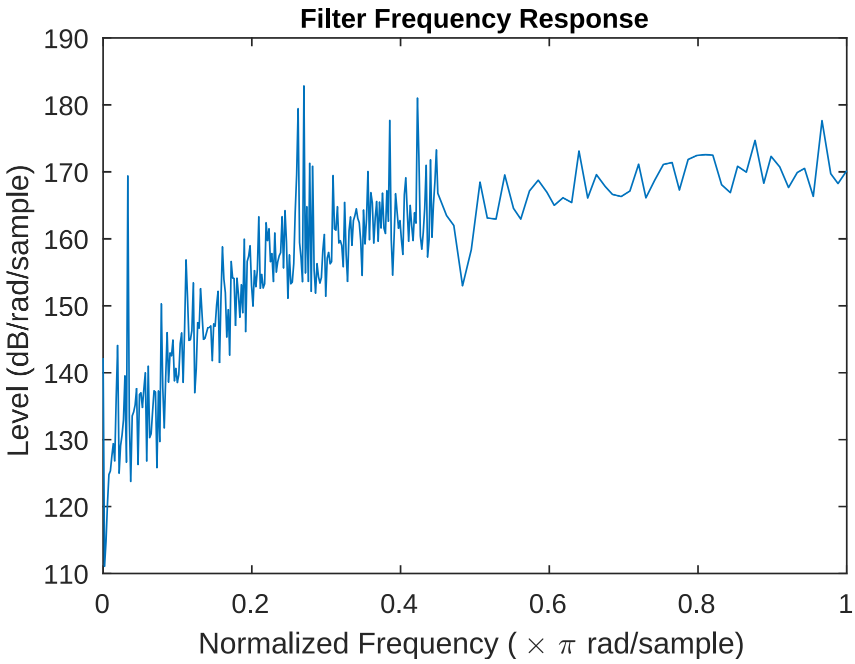

Figure 3.16: Desired Yulewalk filter Frequency Response.

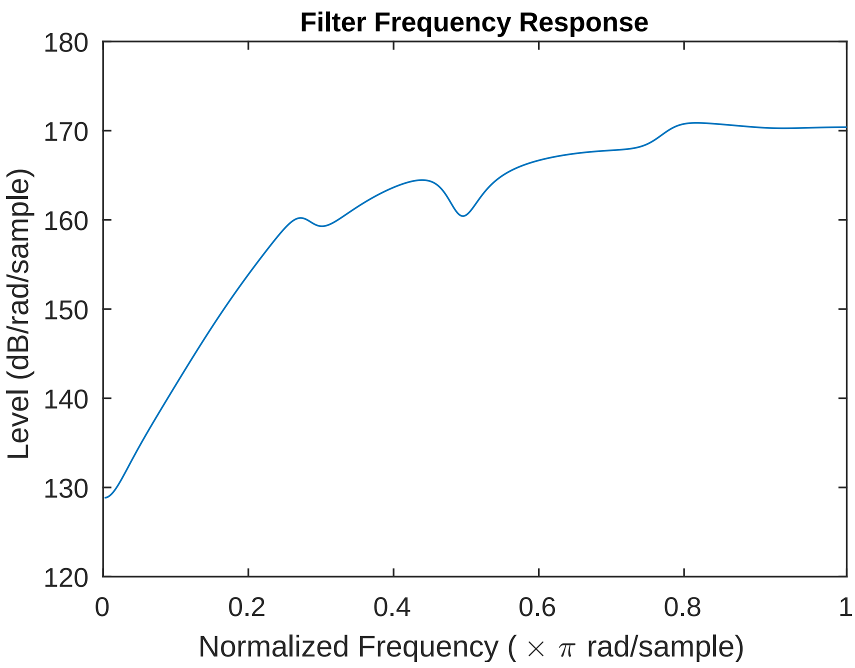

Figure 3.17: Final Yulewalk filter Frequency Response.

Listing 3.16: Loading of new segment and filtering.

```
for i = buffersize+1:loop                    % General loop
    %% Load sound
    currentSound(:,1) = sound((((i-1)*window)+1:i*window),1);

    %% Filtering
    filteredSound = filter(Hz,1,currentSound);
```

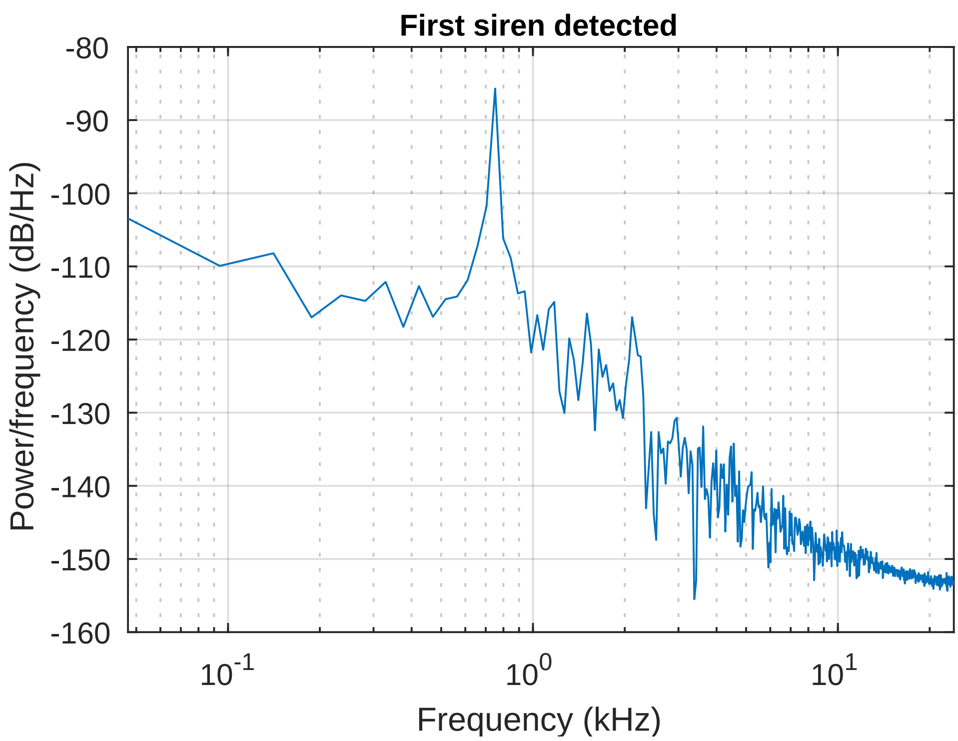

Figure 3.18: Power Spectral Density of the first detected siren.

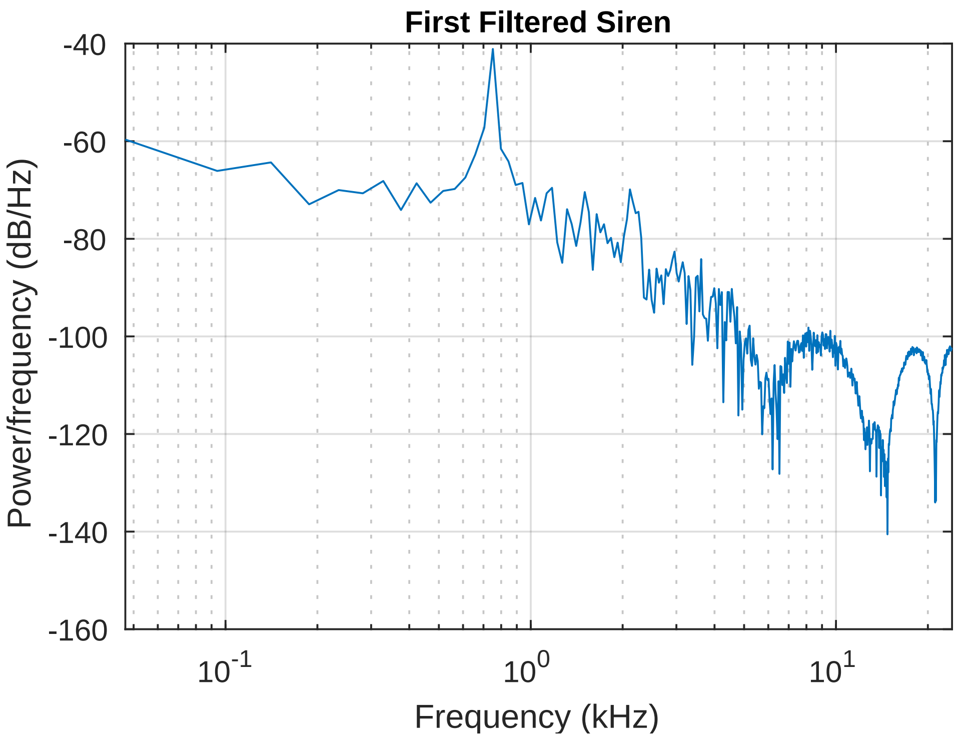

Figure 3.19: Power Spectral Density of the first filtered siren.

Both of the two following listings below are used to find peaks in the segment but the first one 3.17 is used to determine more precise levels which is useful for those algorithms that use level of the signal as the key parameter for them to work, and the second one 3.18 which detects smaller changes in frequency than the previous one which is used for the algorithms based on frequency shifts.

Listing 3.17: Peak finding using FFT.

```
%% Peak finding using FFT
Pfs = abs(fft(filteredSound,nfft)).^2/length(
    filteredSound)/fs;
HpsdFS = dspdata.psd(Pfs(1:length(Pfs)/2),'Fs',fs);

[PKS, LOCS] = findpeaks(10*log10(HpsdFS.data(10:215)),
    Xaxe(10:215),'MinPeakDistance',1,'MinPeakHeight',-60);
Pks(i).pks = struct('Segment',i,'pks',PKS);
Locs(i).locs = struct('Segment',i,'locs',LOCS);
```

Listing 3.18: Peak finding using ARYule.

```
%% Peak finding using ARYule
```

```
[d1,p1] = aryule(filteredSound,100);
pol = roots(d1);
orderpz = cplxpair(pol);
poles = poly(orderpz);
zeros = [1 0 0];
[H, W] = freqz(zeros,poles,[],fs);

[pks, locs] = findpeaks(abs(H),W);
Pks2(i).pks = struct('Segment',i,'pks',pks);
Locs2(i).locs = struct('Segment',i,'locs',locs);
```

Listing 3.19: Checking of any peaks found.

```
if isempty(PKS) % If no peaks are found
   pNoises(i).noise = struct('Segment',i,'Sound',
      currentSound(:,1),...
      'fSound',filteredSound,'PKS',-100,'LOCS',0.1,'Max'
         ,-60);

   pSirens(i).siren = [];
```

Listing 3.20: If determined as noise, include in buffer and new filter is computed.

```
    if isNoise (i, Noises, pSirens) % It is noise

        sirenDetection(i) = 0;
        direction(i) = 0;
        direction2(i) = 0;
        pdirection(i) = 0;
        distance(i) = NaN;
        Sirens(i).siren = [];
        Noises(i).noise = pNoises(i).noise;

        % Update buffer
        [noiseBuffer, bufferPos, finalNoise] =
           updateBuffer(bufferPos, noiseBuffer,
           currentSound, buffersize);

        %% Filter design

        [Hz, filters(i).filter] = designFilter(finalNoise
           , Xaxe, fs);
```

Listing 3.21: Else, it is siren and stored for future analysis.

```
    else % It is siren
```

```
            sirenDetection(i) = 1;
            direction(i) = direction(i-1);
            direction2(i) = direction2(i-1);
            pdirection(i) = 0;
            distance(i) = NaN;
            Sirens(i).siren = struct('Segment',i,'Sound',
               currentSound(:,1),...
            'fSound',filteredSound,'PKS',-100,'LOCS',0.1,'Max
               ',-60,'Direction',direction(i));
            Noises(i).noise = [];
        end
```

Listing 3.22: If peaks are found, add to possible sirens.

```
    else  % If you find peaks

        pNoises(i).noise = [];

        pSirens(i).siren = struct('Segment',i,'Sound',
           currentSound(:,1),...
            'fSound',filteredSound,'PKS',PKS, 'LOCS', LOCS, '
               Max', max(PKS));
```

Listing 3.23: If determined as siren, stored and direction and distance determination.

```
        if isSiren(i, Sirens, pNoises) % It is siren

            sirenDetection(i) = 1;

           % Direction of the siren

            pdirection(i) = getPDirection(i, Sirens, pSirens)
               ;

            direction(i) = getDirection(i, pdirection,
               direction);

            [nowFHPos, nowFHFrec, nowMulti, Diffs(i).diff] =
               getFandH (i, Pks2, Locs2 );
            direction2(i) = getDirection2(i, Diffs);

            [ normDif(end+1), div(end+1) ] = compDif( i,
               fitting2 , Diffs);

           % Distance of the source
```

```

            [distance(i), peakFrec(i)] = getDistance(fappr, frec, i, direction, distance, pSirens);

            f = fillmissing(distance,'previous');

            distance(i) = f(end);

            Sirens(i).siren = struct('Segment',i,'Sound', currentSound(:,1) ,...
            'fSound',filteredSound ,'PKS',PKS, 'LOCS', LOCS, 'Max', max(PKS),'Direction',direction(i));
            Noises(i).noise = [];
```

Listing 3.24: Else, it is noise and added to buffer and new filter is computed.

```
        else % It is noise
            sirenDetection(i) = 0;
            direction(i) = 0;
            direction2(i) = 0;
            pdirection(i) = 0;
            distance(i) = NaN;
            Sirens(i).siren = [];
            Noises(i).noise = pSirens(i).siren;

            % Update buffer
            [noiseBuffer, bufferPos, finalNoise] = updateBuffer(bufferPos, noiseBuffer, currentSound, buffersize);

            %% Filter design

            [Hz, filters(i).filter] = designFilter(finalNoise, Xaxe, fs);

        end
    end
end
```

In the figure below 3.20 the number of sirens detected is shown. It can be seen that the detection algorithm works as it detects when the ambulance is not present and when it is faded away.

The code below 3.25 corresponds to the function used to determine if the segment is siren.

Listing 3.25: Determination if segment is siren.

```
% If 10 segments before were siren or If 10 segments before
    weren't pNoise
% then it is siren
function siren = isSiren (i, Sirens , pNoises)
    siren = (~isempty(Sirens(i-1).siren) && ~isempty(Sirens(i
        -2).siren) && ...
        ~isempty(Sirens(i-3).siren) && ~isempty(Sirens(i-4).
            siren) && ...
        ~isempty(Sirens(i-5).siren) && ~isempty(Sirens(i-6).
            siren) && ...
        ~isempty(Sirens(i-7).siren) && ~isempty(Sirens(i-8).
            siren) && ...
        ~isempty(Sirens(i-9).siren) && ~isempty(Sirens(i-10).
            siren)) || ...
        (isempty(pNoises(i-1).noise) && isempty(pNoises(i-2).
            noise) && ...
        isempty(pNoises(i-3).noise) && isempty(pNoises(i-4).
            noise) && ...
        isempty(pNoises(i-5).noise) && isempty(pNoises(i-6).
            noise) && ...
        isempty(pNoises(i-7).noise) && isempty(pNoises(i-8).
            noise) && ...
        isempty(pNoises(i-9).noise) && isempty(pNoises(i-10).
            noise));
end
```

The code below 3.26 corresponds to the function used to determine if the segment is noise.

Listing 3.26: Determination if segment is noise.

```
% If 5 segments before were noise or If 5 segments before
    weren't pSiren
% then it is siren
function noise = isNoise (i, Noises, pSirens)
     noise = ((~isempty(Noises(i-1).noise) && ~isempty(Noises(
        i-2).noise) && ...
        ~isempty(Noises(i-3).noise) && ~isempty(Noises(i-4).
            noise) && ...
        ~isempty(Noises(i-5).noise)) || (isempty(pSirens(i-1).
            siren) && ...
        isempty(pSirens(i-2).siren) && isempty(pSirens(i-3).
            siren) && ...
```

```
            isempty(pSirens(i-4).siren) && isempty(pSirens(i-5).
                siren)));
end
```

The four following figures 3.20, 3.21, 3.22, 3.23 and 3.24 are the final outputs of the entire algorithm. The accuracy of the results is discussed in section 4.

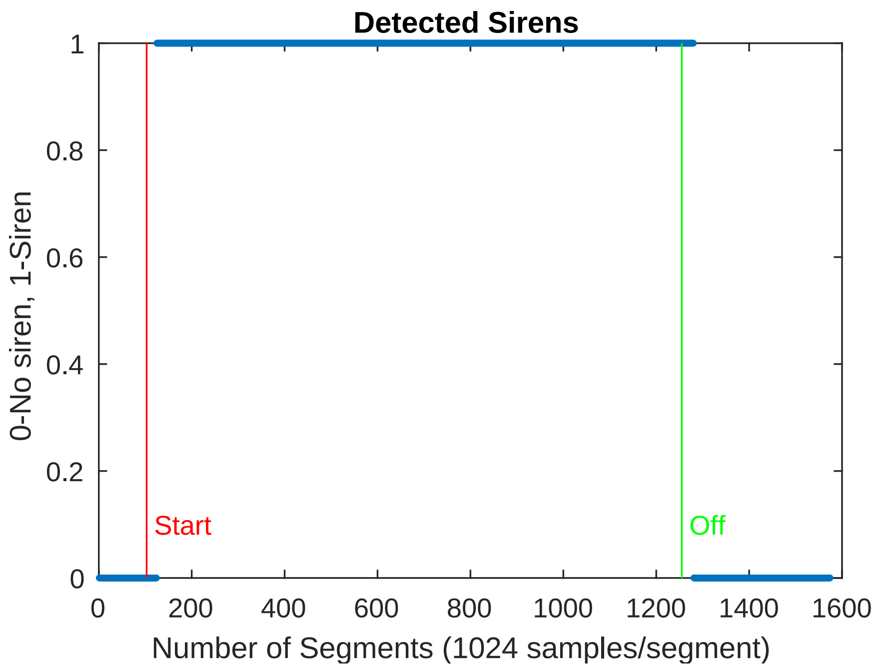

Figure 3.20: Detected Sirens.

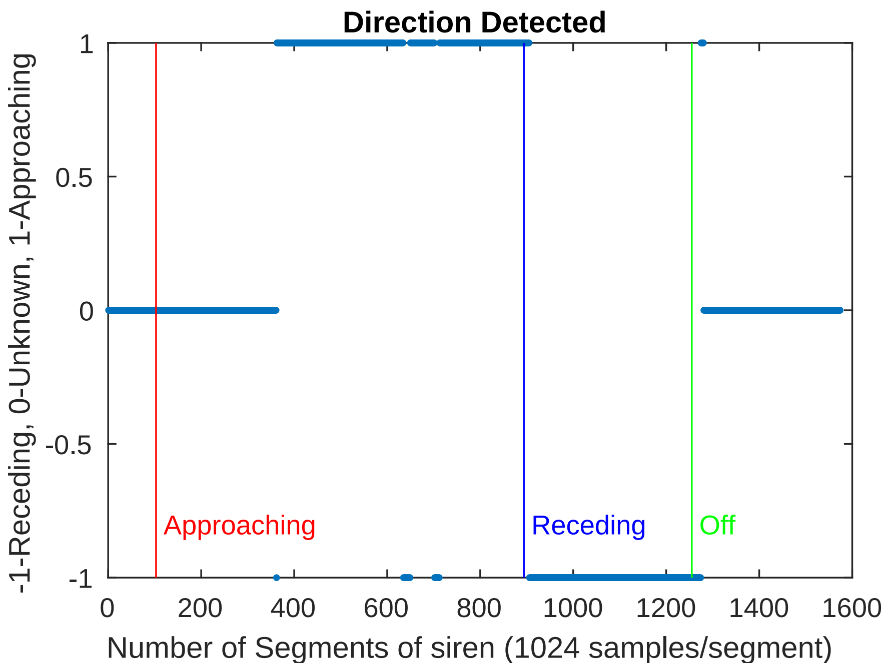

Figure 3.21: Sound source direction detection using level based algorithm.

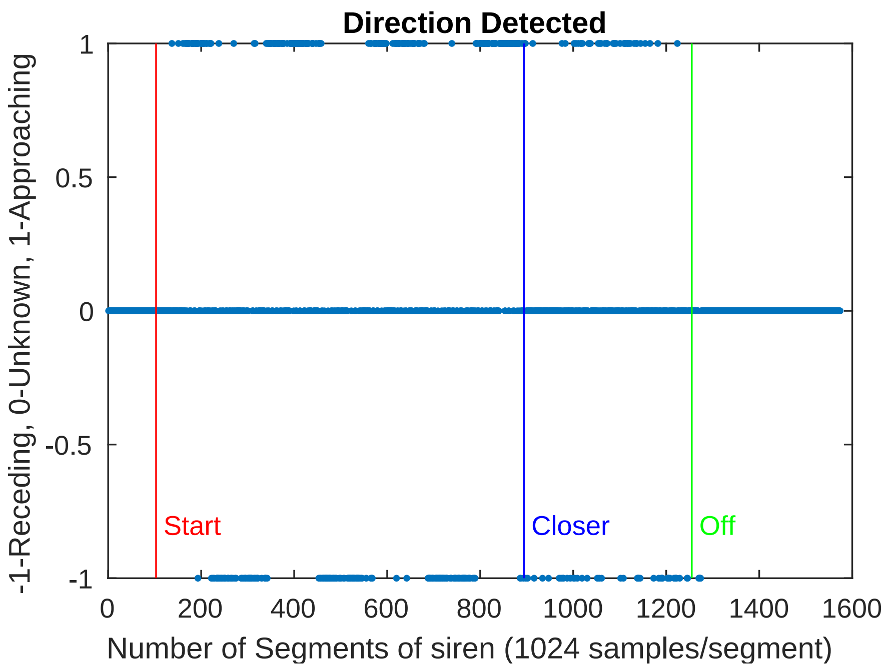

Figure 3.22: Sound source direction detection using frequency shift based algorithm.

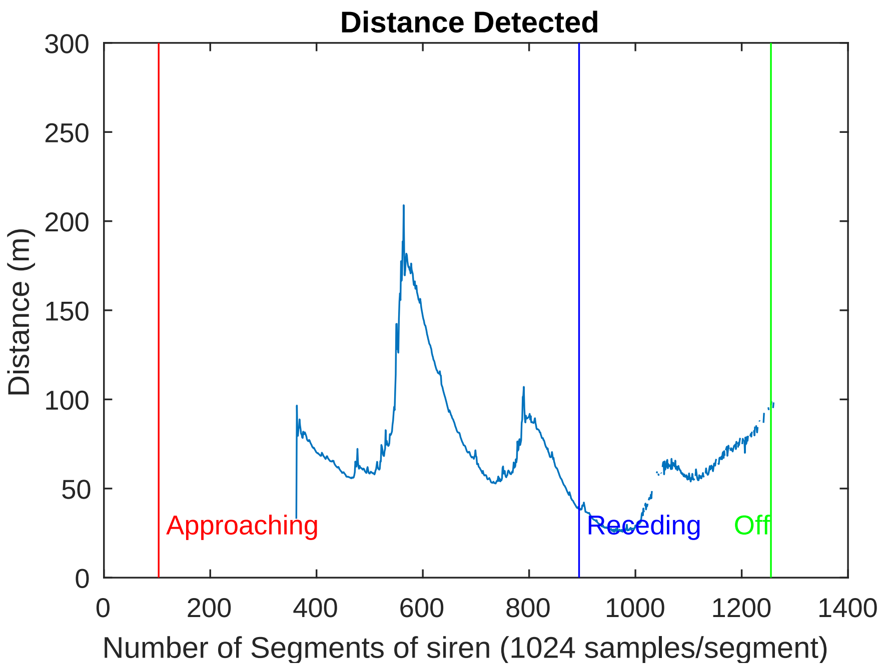

Figure 3.23: Sound source distance detection using all possible points.

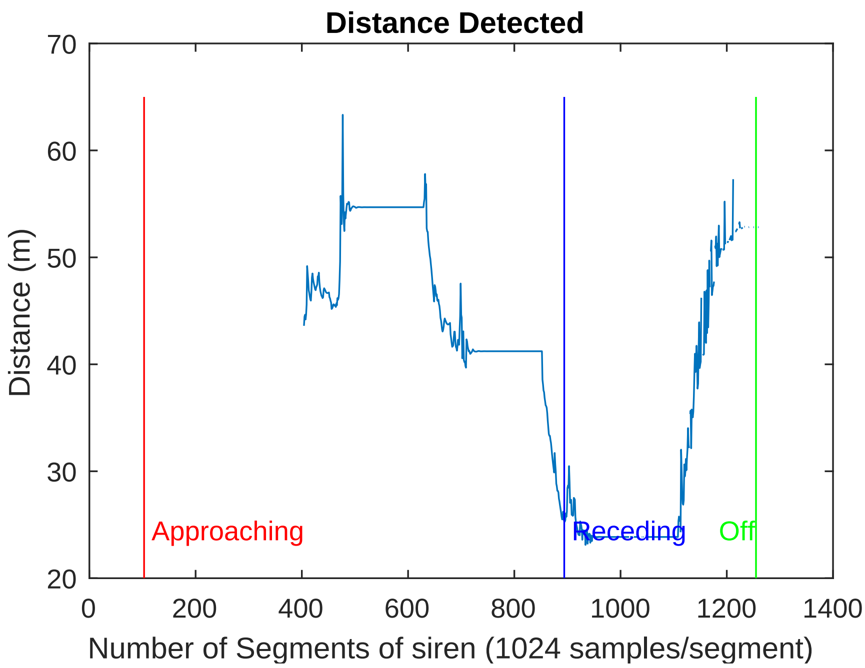

Figure 3.24: Sound source direction detection using horn's best performance range.

4 | Testing & Results

This chapter will describe the testing and results of the three algorithms: Siren Detection Algorithm, Direction Detection Algorithm and Distance Detection Algorithm developed within this project. The testing conditions consists of four different situations (Moving recording with a sound level meter, moving recording with a zoom recorder, background noise of the moving scenario and a stationary recording). Within these situations, these five different characteristics are tested: Siren detection 3.1, direction detection using level values 3.2.1, direction detection using frequency differences values 3.2.2, distance detection using all possible points and distance detection using only frequencies within horn's best performance range 3.3.

4.1 Test Signals

This section will describe each of the test signals. All the signals were prerecorded in different scenarios: two moving situations, one stationary and one background noise of the moving scenario.

A detailed description of how these samples are measured and each specific situation can be found in appendix A.

4.1.1 Sound Level Meter (In Motion)

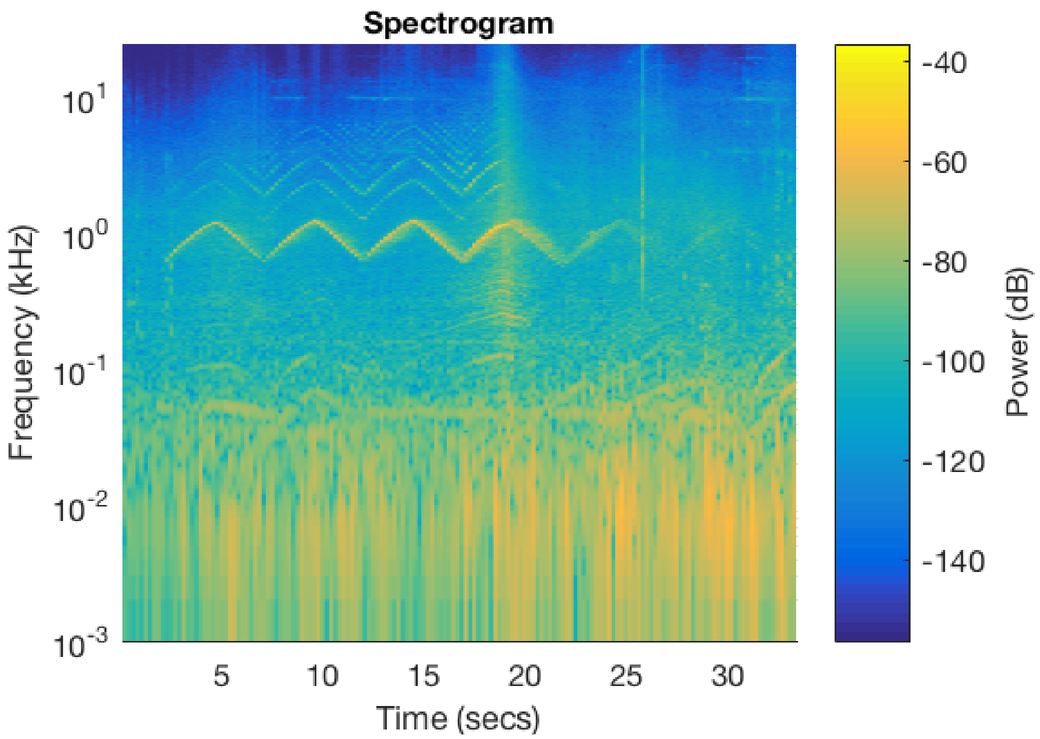

Figure 4.1: Spectrogram of the signal. In this sample the ambulance is approaching the measurement position recorded with a B&K 2270 Sound Level Meter with a measurement height of 1.5 meter. The ambulance exits the Falck Station approximately 200 meters away from the measurement position driving down Håndværkervej. In the recording the Doppler Effect is clearly audible.

The signal is recorded with a B&K 2270 Sound Level Meter with a 48000 kHz sampling frequency. As described in section 2.4, the wail consists of a approximately 6 second frequency sweep from 600-1200 Hz. In figure 4.1 a spectrogram of the sample is shown. At 2.2 seconds into the sample the siren wail from the ambulance appears. The ambulance passes the sound level meter at 19 seconds, and is receding from this point. The signal continues its presence until 32 seconds before it is gradually faded away and is no longer detected.

It is important to state that this signal sample is the signal which all the algorithms are build upon and is the signal which is mainly used within chapter 3. With this in mind, it can be assumed that the best performance will be seen with this sample. Furthermore the Doppler effect is clearly audible within this sample when listened to.

4.1.2 Zoom Recorder (In Motion)

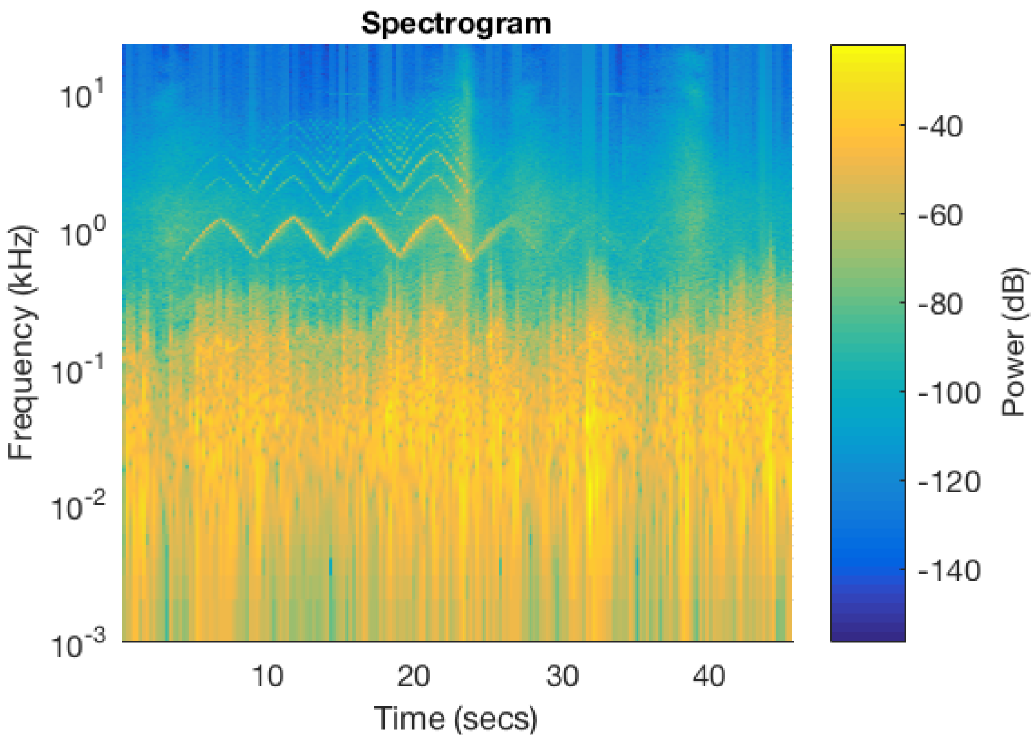

Figure 4.2: Spectrogram of the signal. This sample is equivalent to the scenario as in the Sound Level Meter (In Motion) in section 4.1.1, but with some important differences. The ambulance is still approaching the measurement device, but the device is a Zoom recorder, recording with a 44.1 kHz sampling frequency. Furthermore this recorder was placed 50 meters further away than the Sound Level Meter.

The signal is recorded with a Zoom H4 recorder with a sampling frequency of 44.1 kHz. This particular recording corresponds to the Zoom recorder placed 50 meters further away on Håndværkervej relative to the B&K 2270 Sound Level Meter. In figure 4.2 a spectogram of the signal is shown. The ambulance is approaching the Zoom recorder - 4.2 seconds into the recorded sample the siren wail appears. At 23.2 seconds the ambulance is passing the Zoom recorder. From this point the ambulance is receding the Zoom recorder and gradually fades away until 36 seconds where it is no longer detected.

The Doppler Effect is also clearly audible within the recording from the Zoom recorder.

4.1.3 Backside (Stationary)

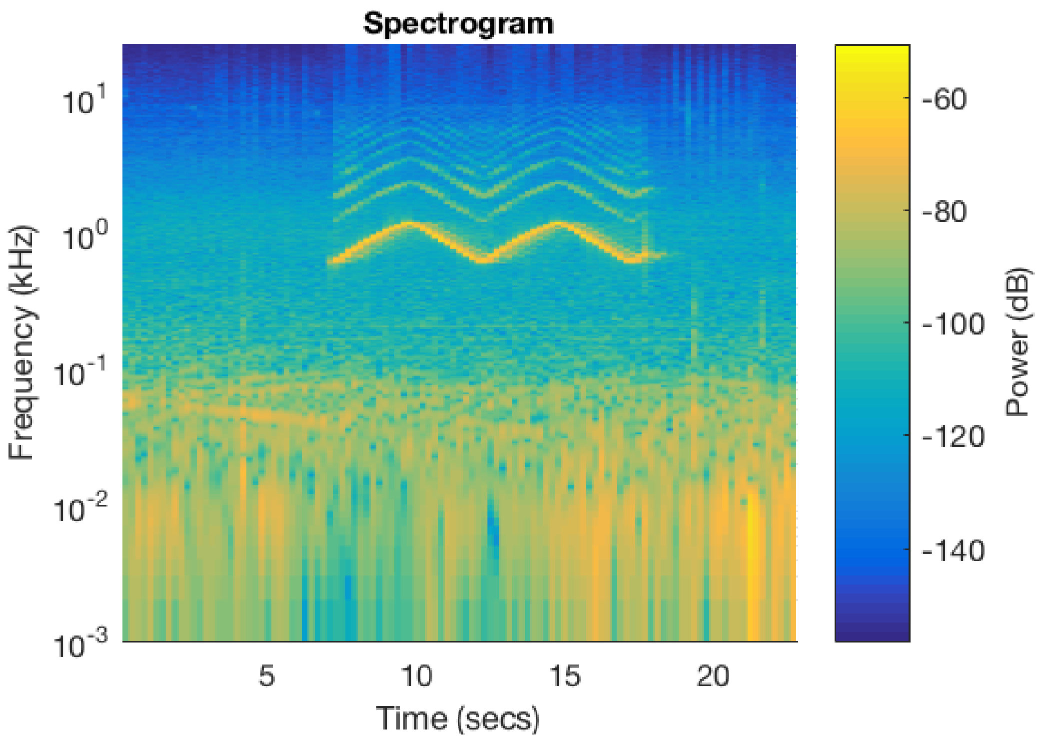

Figure 4.3: Spectrogram of the signal. In this sample the ambulance is stationary and is recorded from the backside at 3 meters distance to the ambulance with a measurement height of 1.5 meter. This sample is recorded with the B&K 2270 Sound Level Meter with sampling frequency of 48000 Hz.

The signal is recorded with a B&K 2270 Sound Level Meter with a 48000 kHz sampling frequency. In figure 4.3 a spectrogram of the signal can be seen. 7.2 seconds into the sample the siren wail is activated and appears in the spectrogram. Two full periods of the wail frequency sweep appears and is shut off at 18.2 seconds.

4.1.4 Background Noise

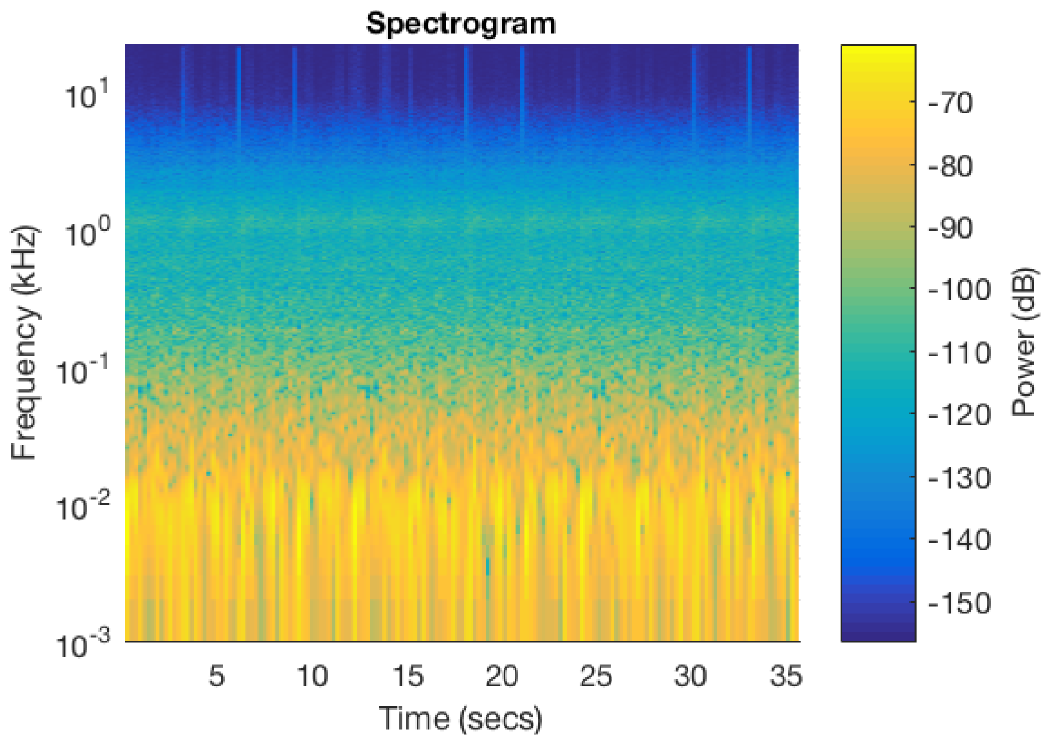

Figure 4.4: Spectrogram of the background signal. The background signal is an extraction of the Sound Level Meter (In Motion) sample. This sample is extracted from one part of this recording, in which the siren was not present. It shares the same measurement conditions and is also recorded with a 48000 Hz sampling frequency.

The background noise sample is an extraction of the Sound Level Meter sample, at the time where no siren was present. Therefore it features the exact same measurement conditions as the Sound Level Meter (In Motion) in section 4.1.1. In figure 4.4 a spectrogram of the signal can be seen. The duration of the sample is 35 seconds.

4.2 Siren Detection Algorithm

This section will test the Siren Detection Algorithm. The purpose for this algorithm is to reliably detect the wail siren of an ambulance within traffic noise. The siren detection is based on knowing signal characteristics specific to the Wail siren, which is a 5 second frequency sweep from 600-1200 Hz. Refer to section 3.1, where the siren detection algorithm is explained.

The four test signals are now tested with the siren detection algorithm. In figure 4.5 the results using the Sound Level Meter sample is shown. The performance for this signal is good, as the siren is instantly detected and keeps being detected throughout the sample, until the siren fades away into vicinity. This corresponds extremely well to the actual movement of the ambulance during recording of this sample. Mind that this is the sample which the algorithm is build for, so the best performance can be expected with this sample.

The Zoom sample also showed good detection of the siren, as seen in figure 4.6. The Zoom recorder is, as mentioned earlier, placed in the same conditions as Sound Level Meter. The results siren fits very well the scenario during recording, as the siren is approaching the siren is detected and keeps being detected as the ambulance is approaching, passes the Zoom recorder and recedes away into the vicinity. There is slight variation in the results at the end, where the detection algorithms shift to no siren, siren and back to no siren.

The stationary example, measured from the backside of the siren shows similarly good results as seen in figure 4.7. The detection matches when the siren is turned on and off.

Lastly the siren detection algorithm is tested for false positives. This is done using the recording sample of the background noise as seen in figure 4.8. In the 35 second duration of the signal there are no detection of a siren, which matches reality as the sample only consists of background noise of the traffic, i.e only cars and trucks passing by.

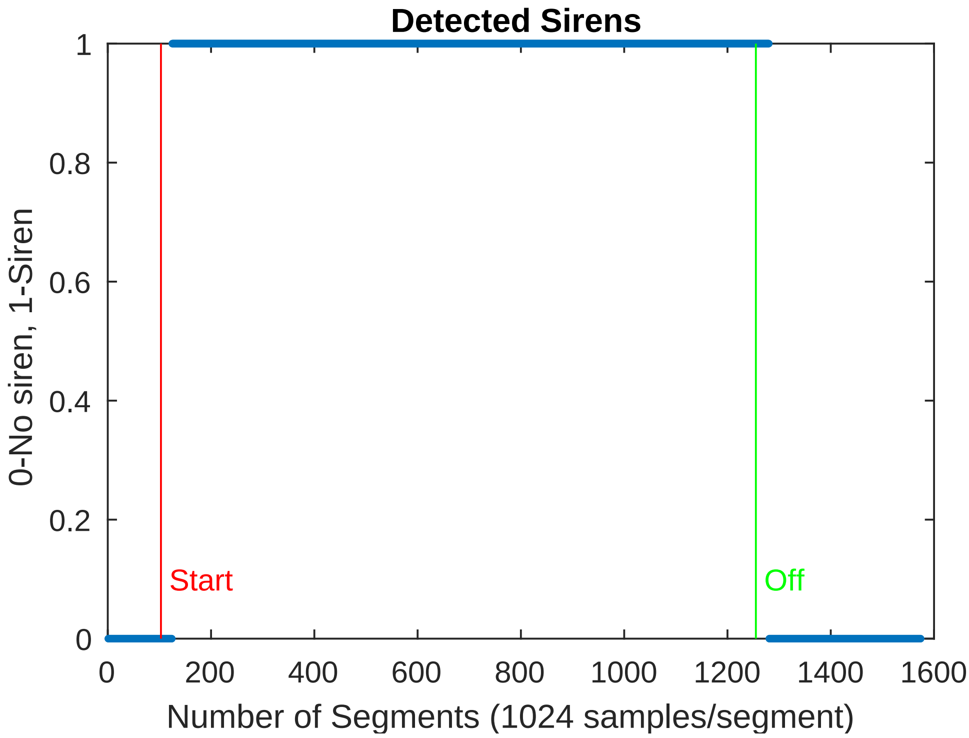

Figure 4.5: Shows the detection of the siren in the SLM sample, in which the siren is approaching.

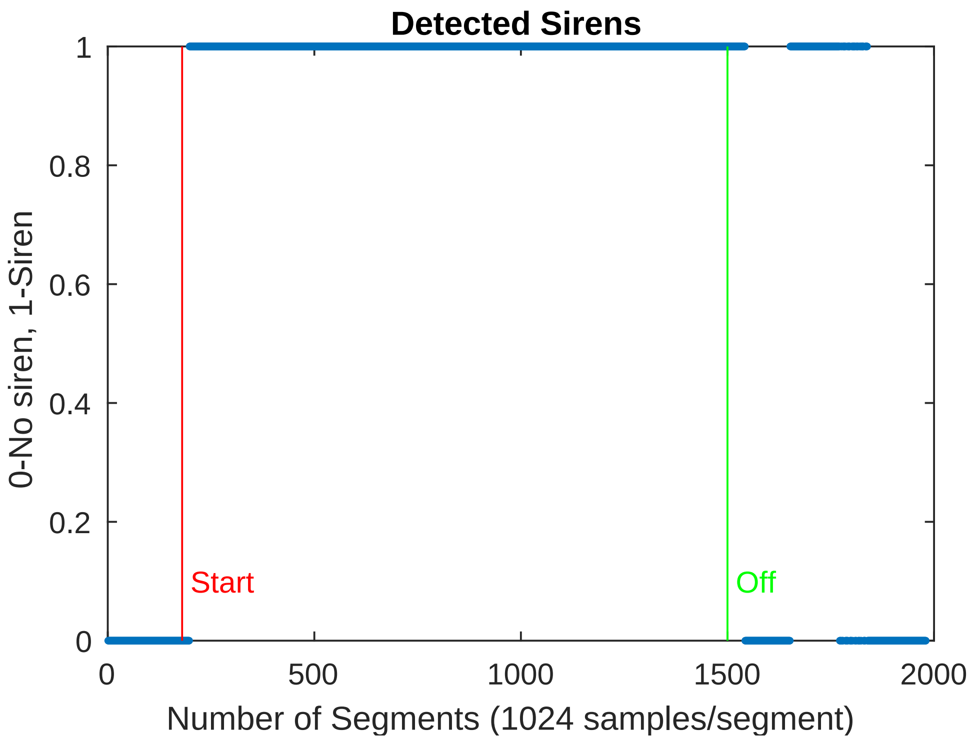

Figure 4.6: Shows the detection of the siren in the Zoom sample, in which the siren is approaching.

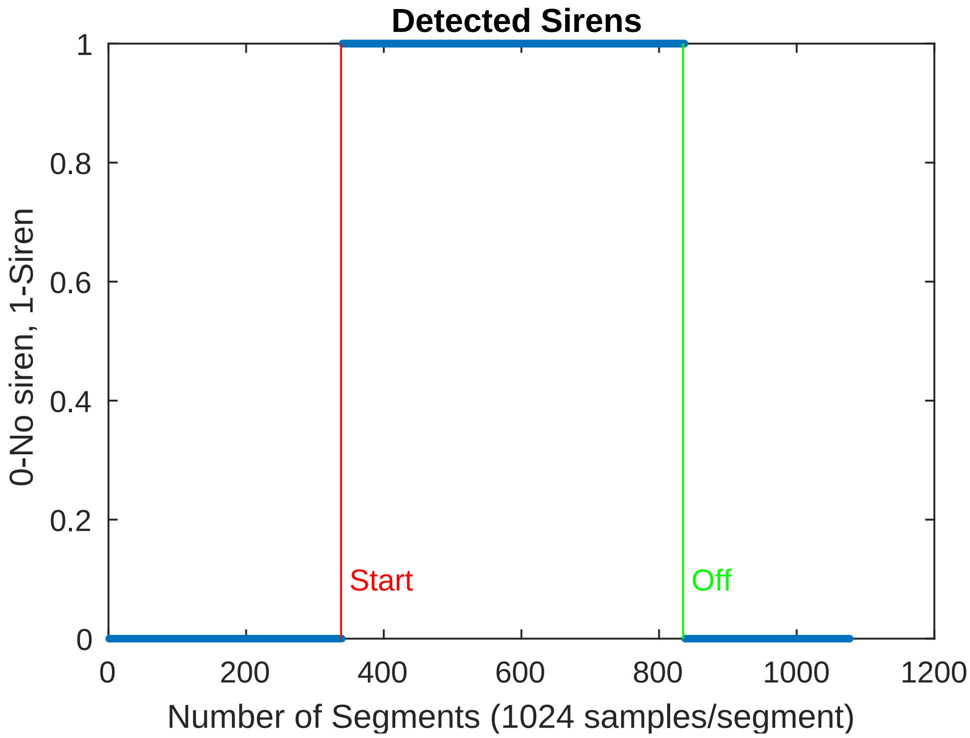

Figure 4.7: Shows the detection of the in the backside sample, in which the siren is stationary.

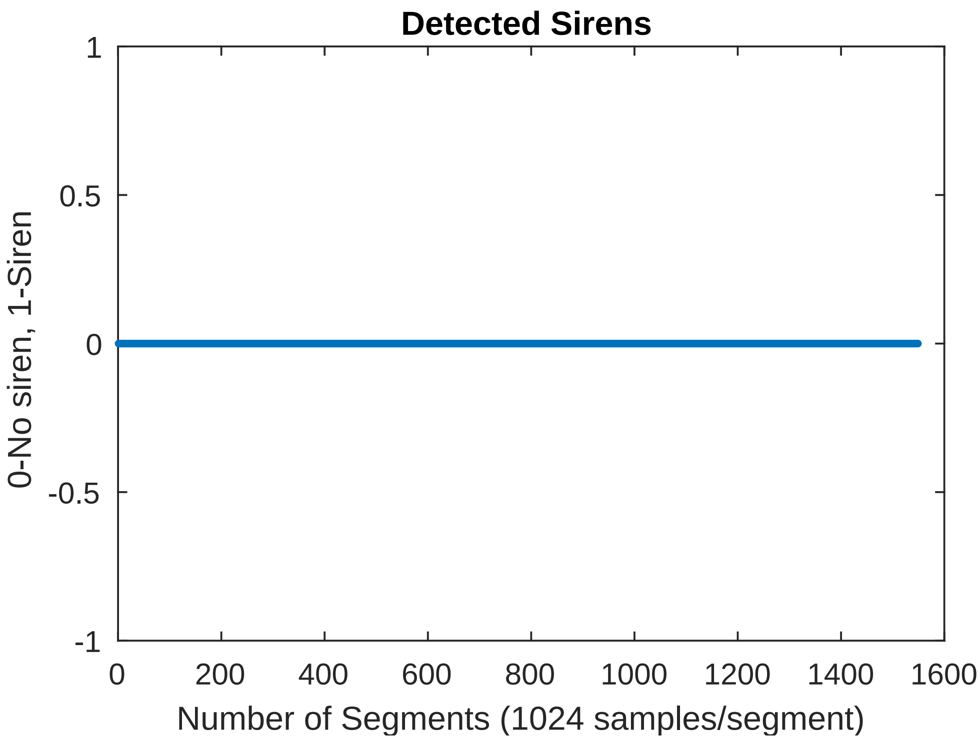

Figure 4.8: Background noise sample. The figures shows no siren detection within this 35 second sample.

4.3 Direction Detection Algorithm

This section will test the Direction Detection Algorithm. The purpose for this algorithm is to determine whether a detected siren corresponds to an ambulance in the vicinity approaching or receding the observer. The direction detection is done in two different ways: Using the level values as described in section 3.2.1 and using the frequency differences values as described in section 3.2.2, based on the shift that occurs due to the Doppler Effect.

The direction detection algorithm will be tested on three signals: Sound Level Meter sample, Zoom sample and the stationary sample. In figure 4.9 the results of the algorithm that uses the levels of the signal to determine the direction of the source are shown. It can be seen that there is a delay between the start of the approaching period and when the method actually detects that the source is approaching. This delay is caused by the algorithm as it has to wait for a whole period of the signal in order to be able to compare values with the previous period. The length of the delay is what was expected as it fits the length of a period of the siren. Other than the delay, the results follows exactly the movement of the source. As the siren is detected, it is determined to be approaching, then the source passes by and is from that point receding.

In figure 4.10 the results of the algorithm that uses the frequency differences between the fundamental and harmonics of the signal to determine the direction of the source are shown. The figure shows mostly the unknown state, because the algorithm is no able to detect any differences between the fundamental and harmonics. Earlier, as stated in section 3.2.2, there was a problem with the time signal variation of the siren wail having a greater impact on the frequency differences than the Doppler Effect, so this was assessed in order to account for cycle of the siren. This gave a marginally better result, but as it can be seen in the figure, the source is alternating between the three states, mainly being unknown when there are no detected differences. Once there is a difference detected, the direction alternates between approaching, receding and unknown. In reality the source is approaching, passes by and then recedes into the vicinity. The erratic behaviour can not be explained as the root cause is unknown.

Now comparing to the results with the Zoom recorder. As seen in figure 4.11, the results of the algorithm that uses the levels of the signal to determine the direction of the source are shown. It can be seen that there is a delay between the start of the approaching period and when the method actually detects that the source is approaching. This delay is again caused because the algorithm has to wait for a whole period of the signal in order to be able to compare values with the previous period. The length of the delay is again what was expected as it fits the length of a period of the siren. In this case there are some errors in the detection of the direction. Compared to the Sound Level Meter sample the performance is mostly the same, but with some errors in terms of consistency.

For the algorithm that uses the frequency differences between the fundamental and harmonics of the signal to determine the direction of the source are shown in figure 4.12. The results are in fact comparable with the result from the Sound Level Meter sample in figure 4.10. The same erratic pattern occurs, where the source is estimated alternately between approaching, receding and the unknown states.

Finally the algorithms are tested with the stationary signal. In figure 4.13 the results of the algorithm that uses the levels of the signal to determine the direction of the source are shown. In this case the direction is most of the time unknown although some other results are obtained due to very little differences in level.

In the figure 4.14, where the algorithm uses the frequency differences between the fundamental and harmonics of the signal to determine the direction of the source, again the direction detected is most of the time unknown with some other results although in this case they are caused by the cycle of the siren wail. Both results show irregularities, but should have been consistently stayed in the unknown state, as a stationary source should not yield any level differences nor Doppler effect which should indicate the siren is either moving towards or away.

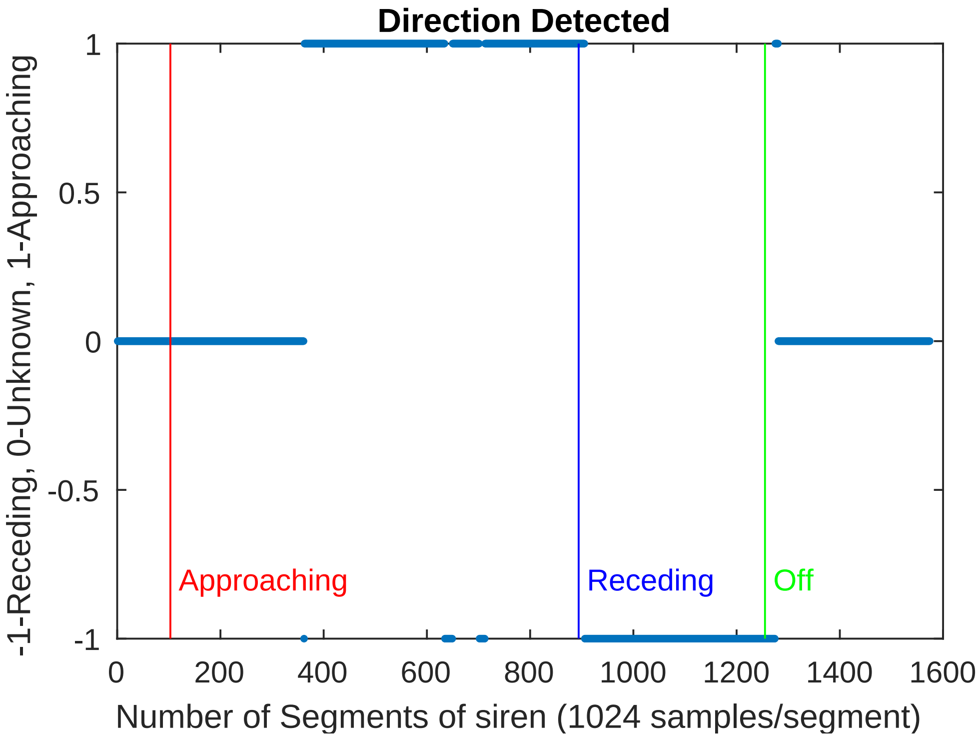

Figure 4.9: Shows the direction of the source using level values for the sound level meter sample. The figure shows the source being detected as approaching, passing and afterwards receding, which is what occured in reality.

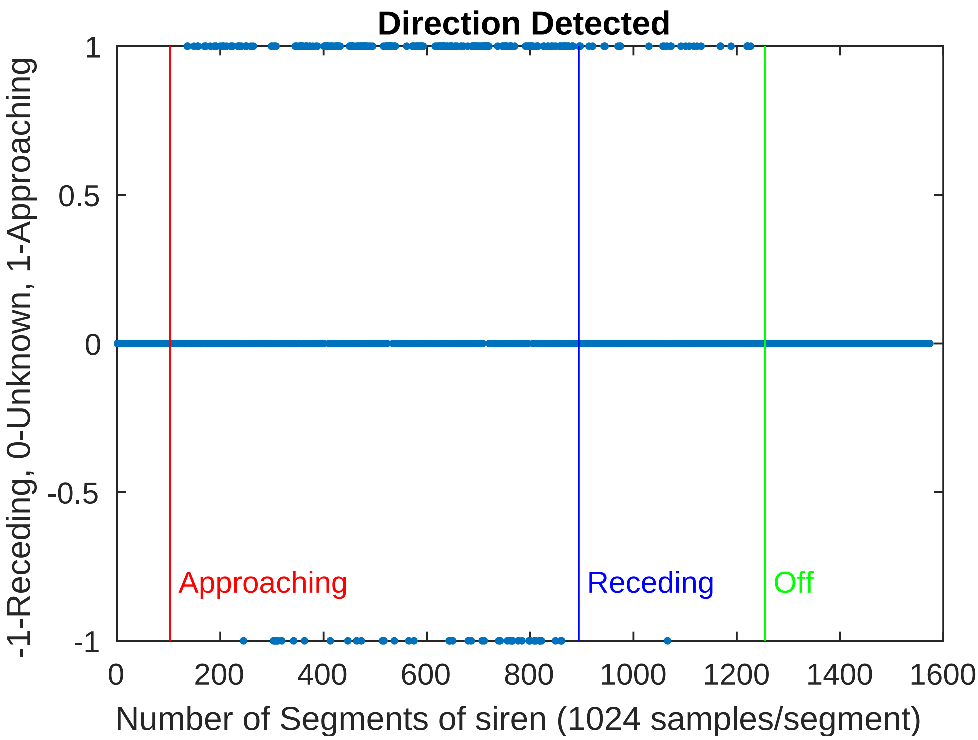

Figure 4.10: Shows the direction of the source using frequency differences values for the sound level meter sample. Using the frequency differences values yields a much poorer results, as the direction is mostly unknown and the approaching/receding/unknown states are erratic.

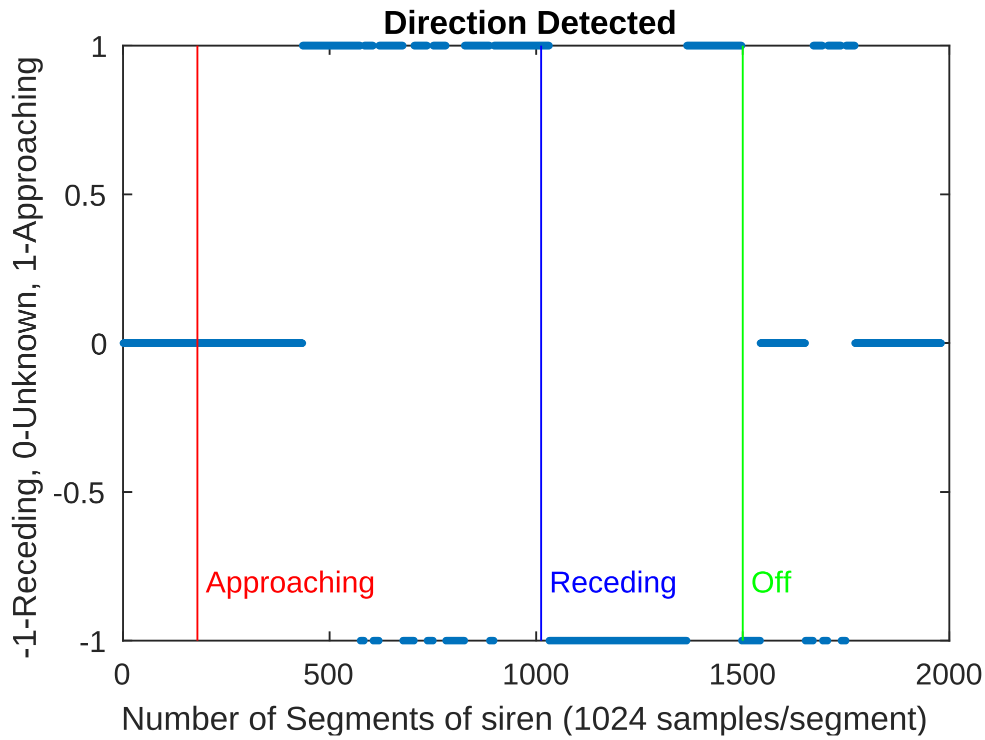

Figure 4.11: Shows the direction of the source using level values for the Zoom sample. It can be seen there is a delay before the direction is determined.

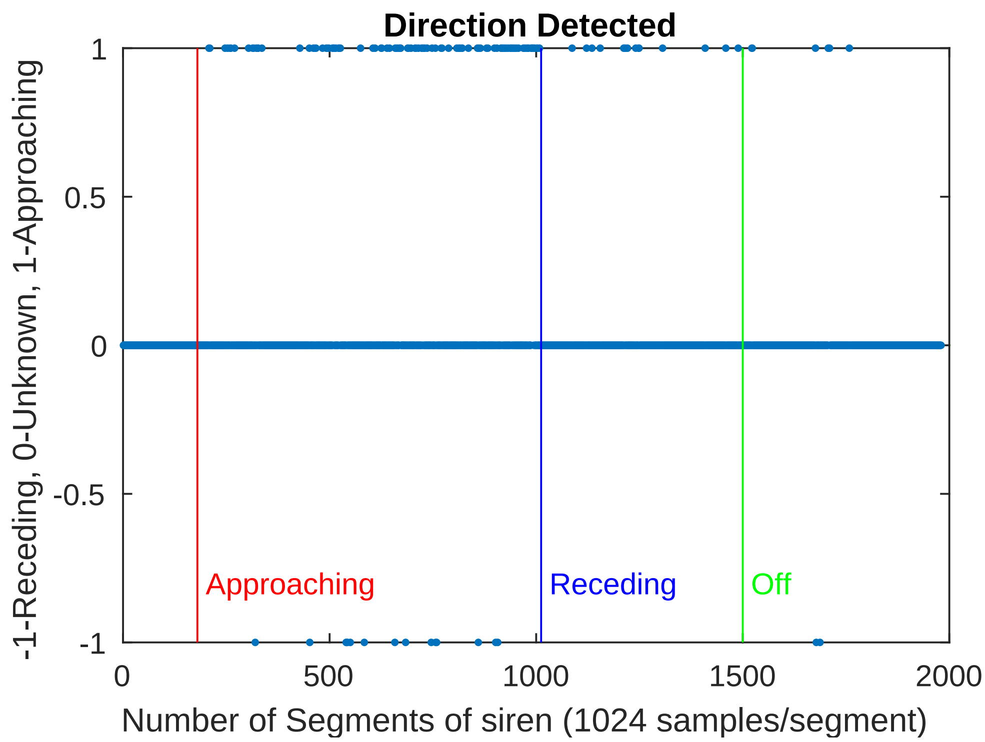

Figure 4.12: Shows the direction of the source using frequency differences values. The performance is similarly poor as with Sound Level Meter sample using the same method.

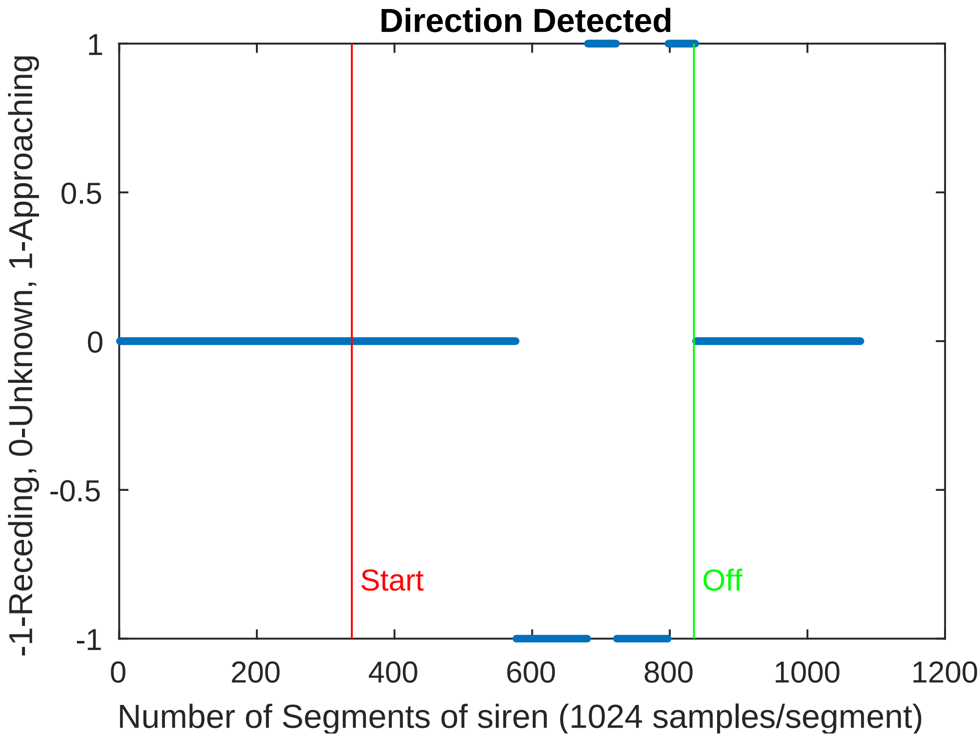

Figure 4.13: Shows the direction of the source using level values for the stationary sample. Since the siren is stationary the direction is mostly unknown, but some irregularities occur due small differences in the level of the signal.

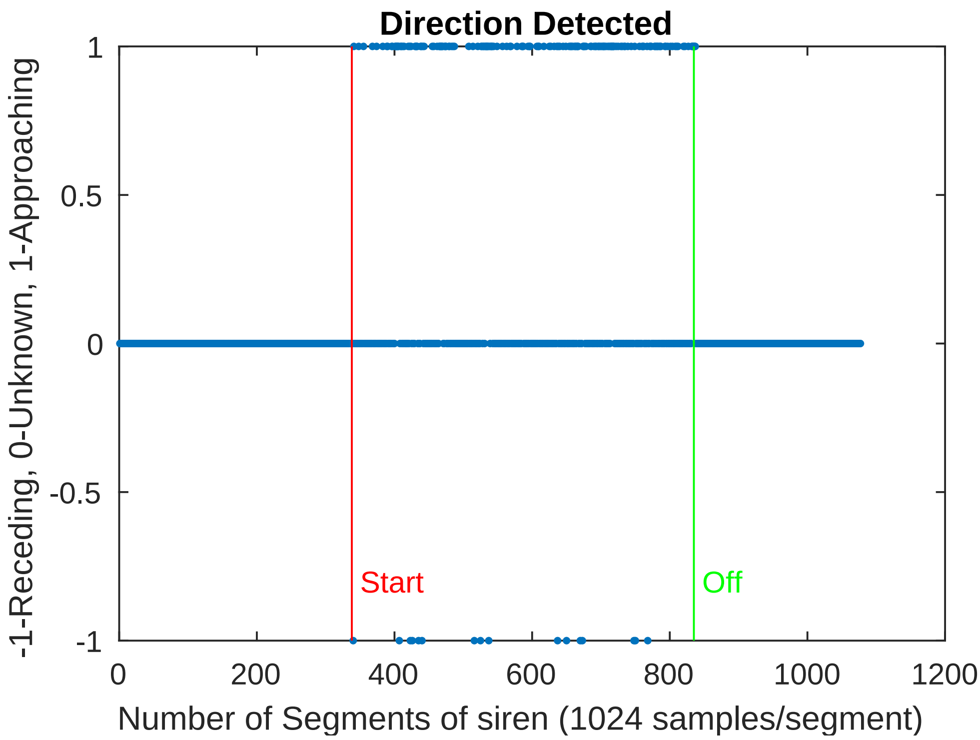

Figure 4.14: Shows the direction of the source using frequency differences values for the stationary sample. Since the siren is stationary the direction should be "unknown", as there should not occur any differences between the fundamental and the harmonics, since there is no Doppler Effect for the stationary sample.

4.4 Distance Detection Algorithm

This section will test the Distance Detection Algorithm. The purpose for this algorithm is to determine how far away the detected siren is. The distance detection is done in two ways: Using all possible points and using only frequencies within the siren horn's highest performance (1kHz-1.2kHz) as described in section 3.3.

The distance detection algorithm is tested using the same three signals: Sound Level Meter sample, Zoom sample and the stationary sample. In figure 4.15 the results of the algorithm that uses all the possible obtained points to determine the distance of the source are shown. The delay at the beginning is caused by the algorithm uses the previously detected directions to take into account the directivity of the source and since the direction detection has a delay, this method also contains the same delay. It can be seen that the distance decreases until the point where the source is at the closest point and then it starts increasing again. The shift between the closest distance and the closest point (blue marker) is due to an averaging in the algorithm which eliminates extreme values caused by a sudden decrease in the level detected. In figure 4.16 the results of the algorithm that uses only the frequencies within the horn's best performance range to determine the distance of the source are shown. The delay is again due to what was explained in the previous figure 4.15. In this case, when there are no values, the distance obtained corresponds to the last obtained one. This gives a less erratic result which correspond better to the source approaching and receding behaviour.

Comparing to the Zoom sample, the results of the algorithm that uses all the possible obtained points to determine the distance of the source are shown in figure 4.17. The delay at the beginning is caused because the algorithm uses the previously detected directions to take into account the directivity of the source and since the direction detection has a delay, this method also contains the same delay. It can be seen that the distance decreases until the point where the source is at the closest point and then it starts increasing again. The shift between the closest distance and the closest point (blue marker) is due to an averaging in the algorithm which eliminates extreme values caused by a sudden decrease in the level detected. Figure 4.18 shows the results of the algorithm that uses only the frequencies within the horn's best performance range to determine the distance of the source. The delay is again due to what was explained in the previous figure 4.17.

In the stationary example, shown in the figure 4.19 and 4.20 the results of both algorithms are quite poor since they should be at all times constant and at a distance of 3 meters. These poor results are caused partly because both algorithms use the previously obtained values of direction to determine the distance values and since they were already not correct, it is carried away to this method as well.

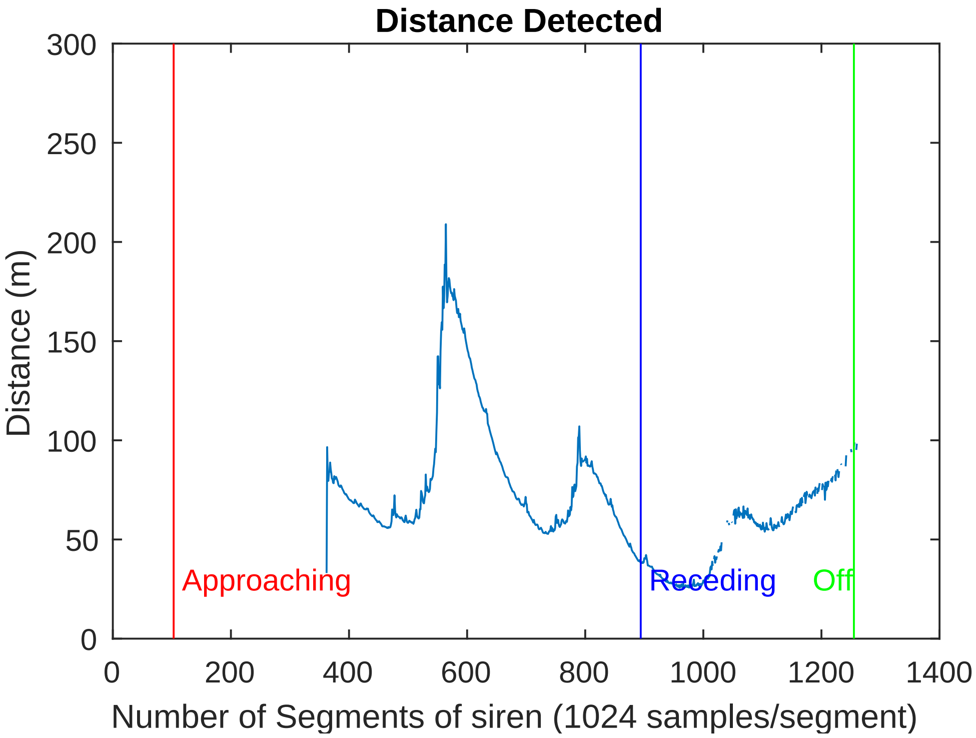

Figure 4.15: Shows the distance of the source using all possible points, using the Sound Level Meter sample.

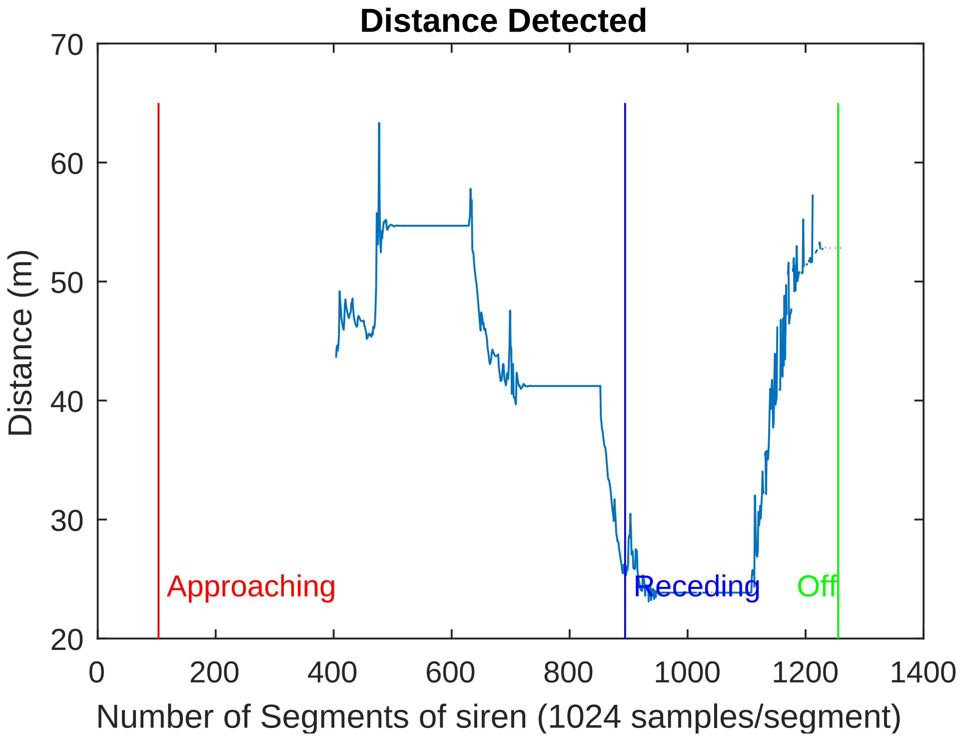

Figure 4.16: Shows the distance of the source using frequencies within horn's best performance range, using the Sound Level Meter sample.

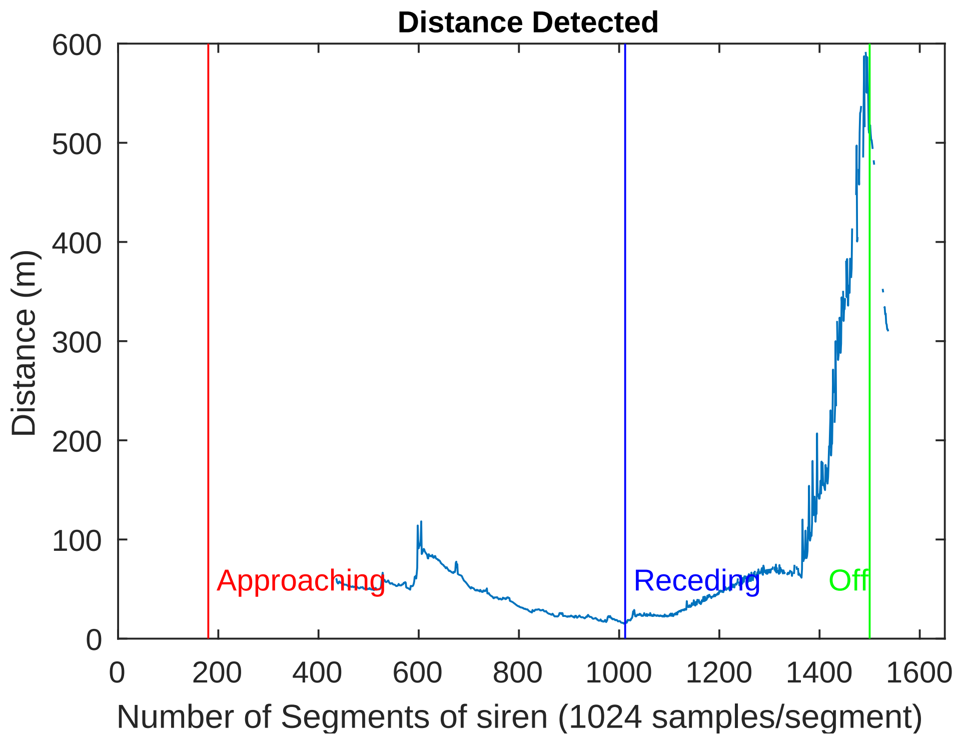

Figure 4.17: Shows the distance of the source using all possible points, using the Zoom sample.

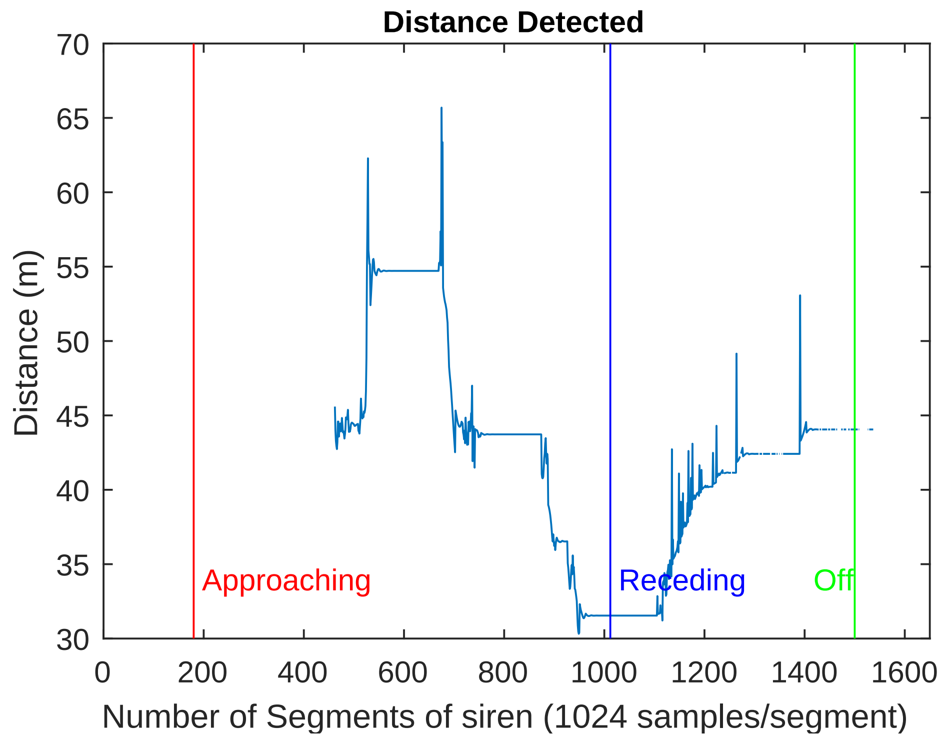

Figure 4.18: Shows the distance of the source using frequencies within horn's best performance range, using the Zoom sample.

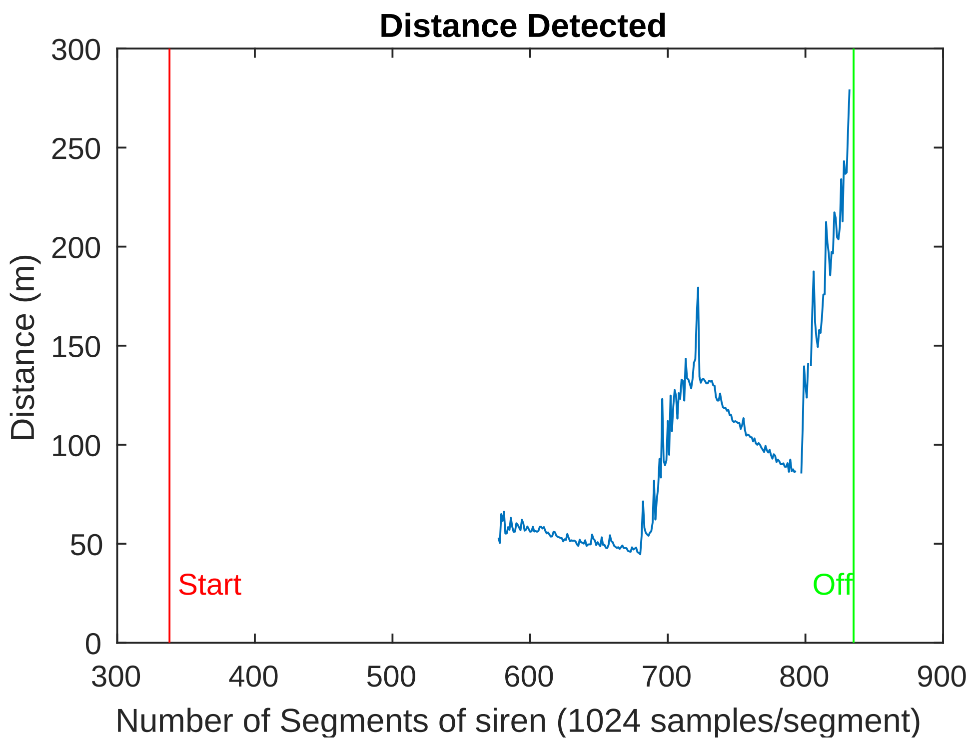

Figure 4.19: Shows the distance of the source using all possible points, using the stationary sample.

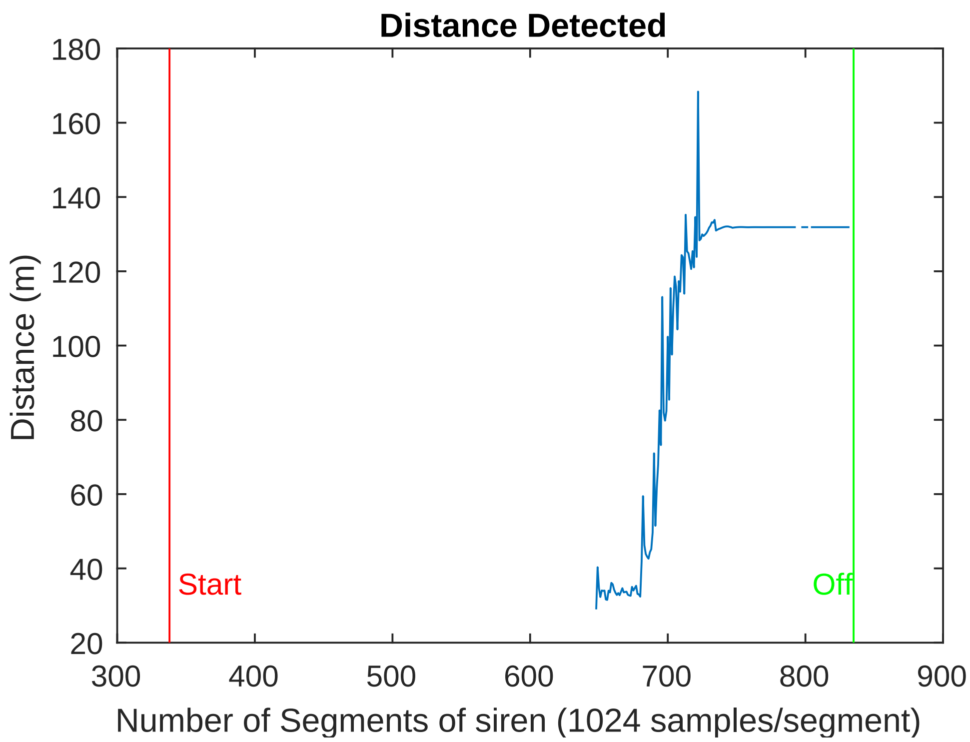

Figure 4.20: Shows the distance of the source using frequencies within horn's best performance range, using the stationary sample.

4.5 Results

For the siren detection algorithm shows very good results. While there are some variation in some of the samples, the algorithms appears to perform well in order to detect siren. For all samples with the siren present it detects the siren promptly and is relatively consistent in all cases, except the variation in the end of the sample for the Zoom recorder shown in figure 4.6. Both the moving and stationary samples measured with the Sound Level Meter showed great detection of the siren. Furthermore no false positives were detected within the background noise sample.

For the direction detection algorithm using the signal levels the results were still good although there were some minor errors. The direction algorithm using the level can be deemed acceptable, even the results vary depending on the tested signal. The performance using the algorithm determining direction based on the frequency differences is poor and erratic. The stationary sample also showed some irregularities regardless of the level or frequency difference method - in both cases the siren is in the "unknown" state, but it should be consistently unknown as the siren is stationary, so no level changes or the frequency differences should occur, since no Doppler Effect takes place. The reason for weird behaviour is not known, but there is clearly something wrong. The influence of time variation, or in other words the cycle of the siren wail has been addressed in section 3.2.2. Before, the siren wail would cause frequency differences much larger than the shifts due to the Doppler Effect. This was originally thought to be the cause of the erratic behaviour, but as it turns out it is not. Therefore the behaviour can not be explained for now and the algorithm using the frequency differences is unusable. The goal was to have a combined solution, and make use of properties of both the level as well as the frequency characteristics based on the differences occurring to the shift due to Doppler Effect.

Having the direction detection be based on the level and frequency differences would ideally create a robust solution, in order to detect whether incoming source is approaching or receding. Furthermore as of now, only the direction determination based on level is working. Being limited to the level of the siren is not ideal, as this does not yield a robust solution as well as creating specific demands into the microphone position. Being only reliant on level would totally reject the possibility of using a microphone from inside the car cabin, e.g. the hands-free communication or even the microphone of a phone, as the acoustical properties for the siren inside the car are incomparable with the measurements done within this project.

For the distance determination algorithm the results are quite erratic. There is quite a bit of variation in the results depending on the method. Using all possible points gives varying results throughout the whole signal, whereas the other method only considering frequencies within the horn's best performance range (1.1-1.2 kHz) shows less erratic results. In any case, it can be seen that the distance decreases when the source is approaching and increases when the source is receding the observer. For the stationary example the distance should be constant, but is showing some incorrect results. Furthermore the distance shown for the stationary is much further than reality, as it is 3 meter in reality and shows much longer distances due to the averaging across different instants to eliminate erratic behaviours. Note that the distance detection algorithm makes use of the direction detection algorithm so any mistakes in the latter one, makes the distance detection results to be wrong. Once the direction is able to be obtained robustly, the distance results should improve accordingly.

5 | Conclusion

The goal of this project was to achieve proof of concept, that a system can be developed which can automatically detect and alert a driver of nearby emergency vehicles by using the properties of the measured signal. The system should be able to detect whether the emergency vehicle in the vicinity is approaching or receding the observer, as well as estimating the distance.

In this project, a test library of sounds has been developed by carrying out in field measurements in different situations including measurements where the source was stationary and in movement. An algorithm able to detect the existence of an ambulance siren within a signal has been created. This is done by looking for the presence of a siren in terms of level and frequency and then discriminate whether it is noise or siren. The accuracy of this algorithm has been proved to give consistently good and robust results in all the tested situations.

Furthermore, two different methods to determine the direction of the source have been designed. One of them using the signal level, and the other one by analysing the frequency differences between the fundamental and harmonics of the signal, based on the properties of the Doppler effect. In the case of the first method, the results were most of the time accurate, meanwhile, the second method did not give correct results for any of the tested situations. The direction detection based on the frequency differences yields erratic results, for which the reason cannot be explained. Ideally, direction determination would rely both on the level and frequencies of the signal in order to create a robust algorithm. As the direction determination is only working based on level, it limits the implementation capabilities as any microphone placed within the inside the cabin of a car, e.g. using the microphone for hands-free communication or a phone, would yield vastly different acoustical properties of a measured siren.

Additionally, to obtain the relative distance between the source and the observer, two methods were built based on the level of the signal. One using the whole frequency range, and the other one making use only of the siren's horn optimal performance. Both methods give reasonable results although they could be improved.

The overall conclusion of this project is that a robust solution for detecting a siren of an emergency vehicle, determining direction and estimating the distance could be deemed possible. This project only partly accomplishes this goal, as it only satisfies the detection of sirens. The level alone is not a robust solution for determining the direction and the distance estimation is not showing any numbers close to reality. Further research is therefore necessary in order to correct the direction determination using frequencies differences and make the distance estimation viable.

6 | Further Research

This chapter will comment on the aspects that can be further researched. As it can be seen in the testing and results chapter, there is a problem with the current state of the direction determination based on the frequency differences of the fundamental and the harmonics. The problem is that the time variation of the siren wail yields a much stronger frequency difference and the difference values between the fundamental and harmonics is completely negated. For this to work properly, a stronger model which describes the periodicity of the siren wail has to be established.

Speed detection is also lacking and could be proven to be very useful in real world scenarios. The speed could be determined by using a combination of the frequency shifts by the Doppler effect and the signal detected levels.

An entirely different approach of dealing with this problem is to use machine learning in order to make a learning algorithm, which can learn the periodicity of the siren wail. Using machine learning it can improve the robustness of all the algorithms used within this project, as the machine learning algorithm will gradually learn more of the characteristics of the desired signal, i.e. the siren wail.

For further research it would be nice to implement a real-time implementation, as the currently the algorithms are only tested using pre-recorded signal samples. Using a real-time implementation the algorithms could be tested in a more realistic scenario. One example could be playing the test signals from an external sound player to see whether the real-time implementation would work. Furthermore testing the whole project in real life, in real time would be the ideal scenario in which to thoroughly test the algorithms for siren detection, direction determination and finally distance determination.

This project was chosen to consider on the siren wail of an ambulance, but in a final solution both the siren wail and yelp (the two siren types used by the ambulances), as well as the siren types used by fire trucks and police cars should be considered in order to be deemed complete and robust solution to work in practice.

Finally as mentioned in section 2.1, a final solution would be integrated into a car. This would mean choosing a microphone for sound acquisition either inside or outside the cabin to use as input for algorithms to perform siren detection, direction determination and distance - and speed - determination. Furthermore the driver needs to be alerted by incoming emergency vehicles in the vicinity. This could be done by showing an alert message in the car computer behind the steering wheel, as well as lowering the volume/muting the car sound system and hands-free communication.

Bibliography

[1] M. Møller, *Distraktion i forbindelse med bilkørsel*. DTU, 2010. Chap. 4.4.

[2] M. Schwoerer, "The art of noise." http://www.thetruthaboutcars.com/2009/02/the-art-ofnoise/, 2009. Comparison of car cabin noise at 130 km/h.

[3] R. F. S. Trafik, "Uopmærksomhed i trafikken." http://newision.dk/websites/sikkertrafik/, 2015.

[4] C. Warncke, "Uopmaerksomhed forsinker udrykning." https://transporttidende.com/artikler/sikkerhed-c22/uopmaerksomhed-forsinker-udrykning-p42032, 2015.

[5] Falck, "Paramedic at Falck Station Håndværkervej 27, 9000 Aalborg," 2017. Personal Interview.

[6] C. Insight, "44 Corporations Working On Autonomous Vehicles." https://www.cbinsights.com/blog/autonomous-driverless-vehicles-corporations-list/, 2017.

[7] L. Walford, "Autonomous cars: What does the future look like?." http://www.jsonline.com/story/sponsor-story/griffin-ford/2017/02/06/autonomous-cars-what-does-the-future-look-like/97544742/, 2017.

[8] Waymo, "Journey towards self-driving technology." https://waymo.com/journey/, 2017.

[9] Mercedes Benz, "Drive Pilot Technology." https://www.mercedes-benz.com/en/mercedes-benz/innovation/with-intelligent-drive-more-comfort-in-road-traffic/, 2017.

[10] Tesla Inc., "Auto Pilot." https://www.tesla.com/autopilot, 2017.

[11] Federal Signal Vama, “Siren Loudspeakers.” https://www.fedsigvama.com/en/productos/fuego-y-ambulancia/senalizacion-acustica/, 2017.

[12] Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, James V. Sanders, *Fundamentals of Acoustics*. John Wiley & Sons, Inc., 2000. pp. 414-416.

[13] Harry F. Olson, *Acoustical Engineering*. D. VAN NOSTRAND COMPANY, INC., 1957. Chap. 7: Horn Loudspeakers.

[14] Bjørn Kolbrek, *Horn Theory: An Introduction, part 1*. AudioXpress, 2008.

[15] Federal Signal Vama, “Integrated Control Systems.” http://fsvama.gr/pdf/FSV2015PC.pdf, 2017.

[16] J. S. Lamancusa, *Outdoor Sound Propagation*. Penn State University, 2009. Chap. 10.

[17] Cyril M. Harris, *Absorption of Sound in Air versus Humidity and Temperature*. Columbia University, 1966. pp. 149-159.

[18] U. Ingaard, *A Review of the Influence of Meteorological Conditions on Sound Propagation*. The Journal of the Acoustical Society of America 25, 1953. pp. 405-411.

[19] D. Aylor, *Noise Reduction by Vegetation and Ground*. The Journal of the Acoustical Society of America 51, 1972. pp. 197-205.

[20] Eric W. Weisstein, “The Doppler Effect.” http://scienceworld.wolfram.com/physics/DopplerEffect.html, 2007.

[21] MathWorks, “ARYule: Autoregressive all-pole model parameters — Yule-Walker method.” https://se.mathworks.com/help/signal/ref/aryule.html, 2017.

[22] MathWorks, “Yulewalk Filter: Recursive digital filter design.” https://se.mathworks.com/help/signal/ref/yulewalk.html, 2017.

A | Ambulance Siren Measurements

Introduction

For detecting sirens of emergency vehicles, it is of interest to measure and study the sirens emitted by these vehicles. Initial measurements will be limited to ambulances, but fire trucks and police cars are also of interest. The goal is to perform measurements of the siren of an ambulance, both in motion and stationary, to determine the frequency content of the siren, distinguishing the fundamental and harmonic frequencies for a signal analysis. By also measuring the ambulance in motion, the Doppler Effect will cause an apparent change in frequency by the motion of the source relative to a fixed measurement position. Documenting the frequency shift as the ambulance is moving towards and away from the fixed measurement position will be crucial for the automatic detection of emergency vehicles which is the overall project goal.

Procedure

First, a measurement of the siren with the ambulance being stationary will be performed from front, sides and back. Both sides should be measured to determine symmetry. Secondly, measurements of the ambulance in motion should be done. The ambulance should be driving towards a fixed measurement position, pass by and drive passed it, thus getting the Doppler Effect of the siren in practice. Optimally the ambulance will have consistent speed when approaching and passing by the measurement position in a straight direction, but most likely the measurement scenario will not be with this amount of control.
The measurements will be conducted on open street, by Falck (a departure site for ambulances). Street noise has to be considered and might influence measurements.

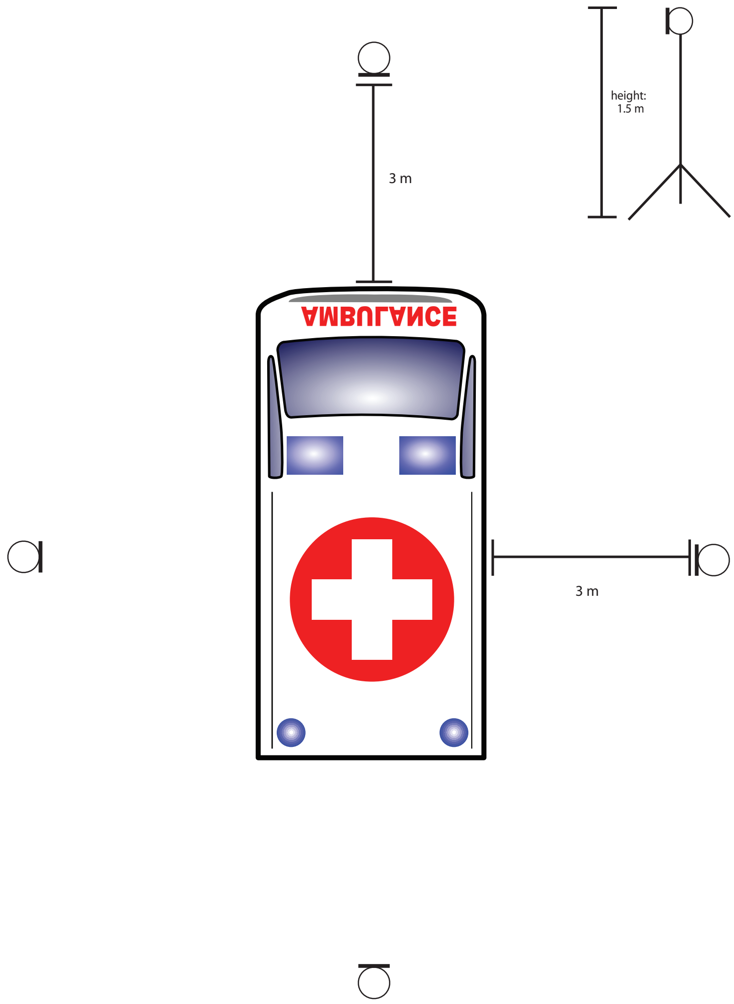

Figure A.1: Measurement points equidistant to ambulance.

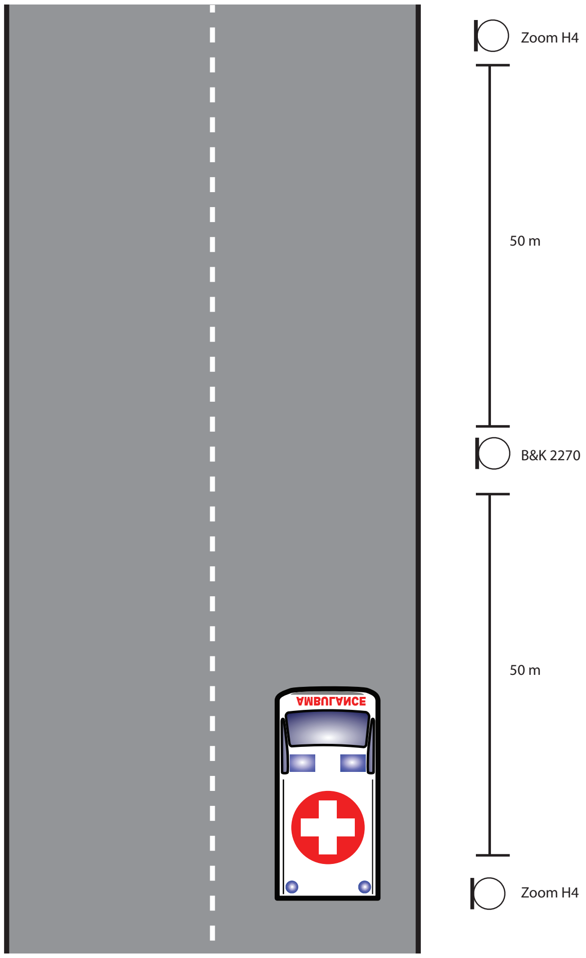

Figure A.2: One stationary measurement position along road or runaway at which the ambulance is driving. This measurement will record the Doppler Effect in effect, as well as showing how noise from road, other potential cars, etc. will affect the siren.

A background noise measurement at the same position but without the siren present will be performed to be able to subtract the background noise from the measurements.

The measurements will be conducted with a B&K Sound Level Meter, which will store the measurements on memory and will be extracted for post-processing in MATLAB.

Scenario

The measurement was conducted on Håndværkervej 27, 9000 Aalborg from 13.30-14.50 on 15th of February 2017.

- Temperature: 3° C.
- Relative humidity: 75 %.
- Wind speed: 2 m/s.
- Atmospheric pressure: 1033 mBar.

List of Equipment

- B&K 2270 Sound Level Meter
- 2 x Zoom H4 Recorders
- Leica laser distance meter (for short use)
- Leica laser Range Meter (for long use)
- Thermometer/Hygrometer
- Microphone stands

Measurement Setup 1: Stationary

In the first measurement setup an ambulance is placed in the parking lot of the Falck station. Note that the station building is place within 10 meters which could cause reflections in some instances. Measurements were conducted sequentially, starting from the right side of the ambulance and moving counter-clockwise around the ambulance. Each measurement position is 3 meters away from the ambulance.

The main siren mode "Wail" was measured first, while a second "Yelp" mode was measured secondly, thus resulting in a total of eight stationary measurements.

A.0.1 Wail

The wail is the most common used siren of all. It consists of a repetitive frequency sweep of approximately 5 seconds between the frequencies of 600Hz and 1200Hz. A spectrogram of the siren is shown in the figure A.3. As well as the main signal, higher harmonics can be seen which will help to extract some information. The energy in the lower frequencies correspond to the background noise caused by the vehicles other than the ambulance. It is also described in section 2.4.4

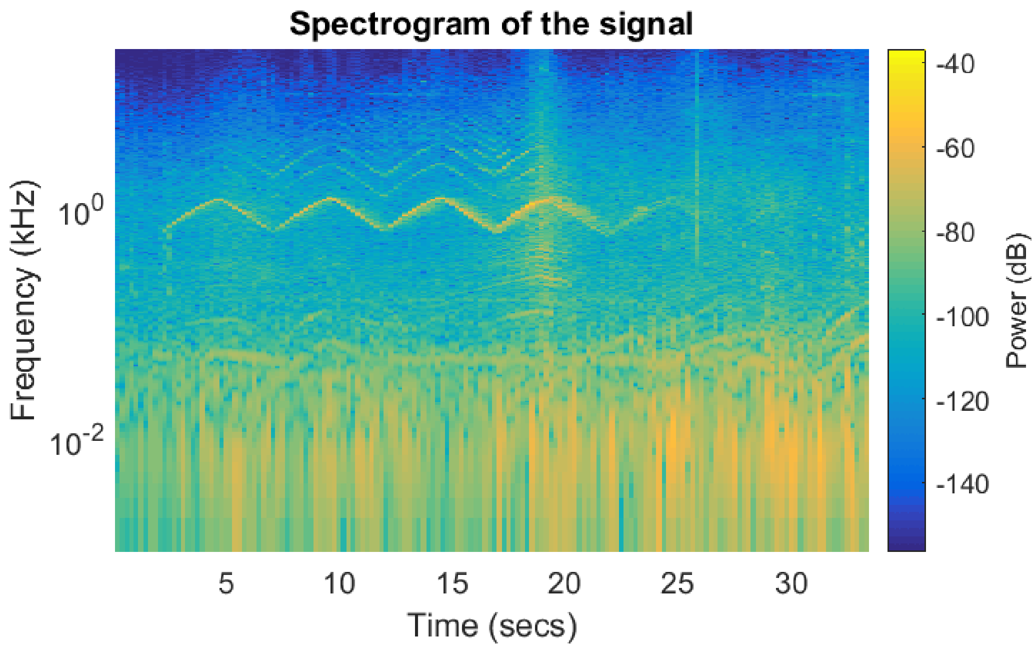

Figure A.3: Spectrogram of the wail.

In the figure below 2.5, the PSD of the siren can be seen. The main tone and its higher harmonics can be identified as well as the difference in power between them and the distance in frequency. This will be of extreme use to extract the information of distance, direction and speed of the sound source.

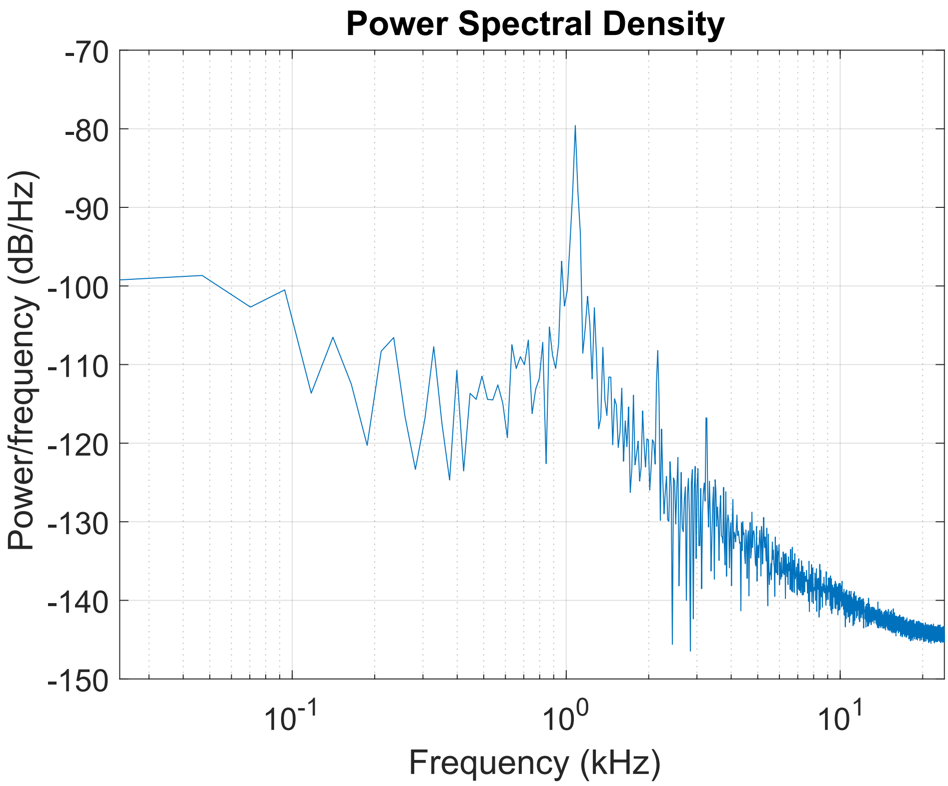

Figure A.4: Power spectral density of 30ms of the wail.

The levels of the Wail siren are measured at 3 meters distance to the ambulance - from there the LC$_{peak}$ value is measured.

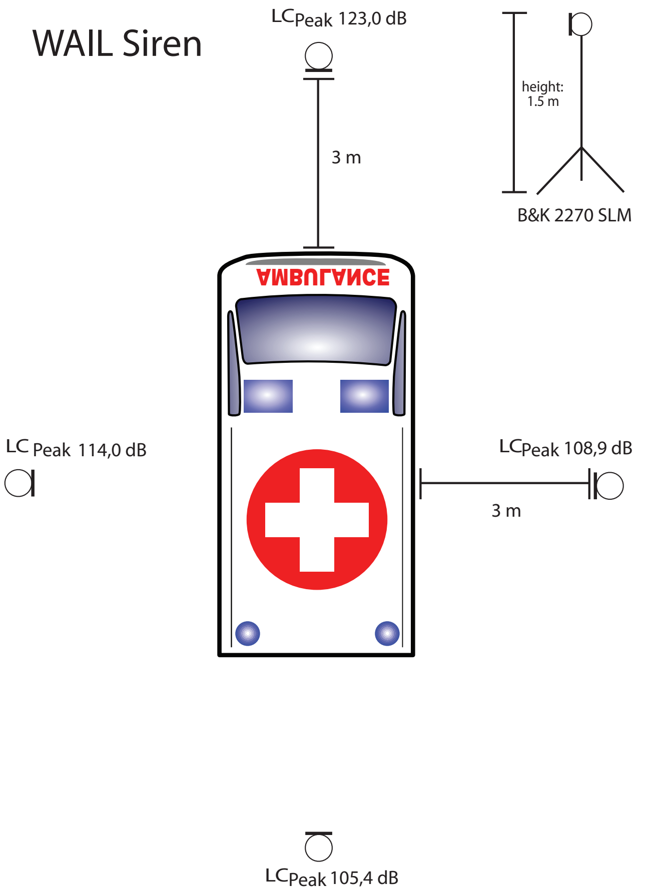

Figure A.5: LC$_{peak}$ values of the wail siren measured at 3 meters distance to the siren from all sides.

A.0.2 Yelp

The Yelp is the second most used siren, specially when approaching an intersection. It consist of a much shorter frequency sweep of around 0,3s between the frequencies of 600Hz and 1100Hz.In the figure 2.7 the spectrogram of the siren is shown where again the fundamental signal and its higher harmonics can be seen.

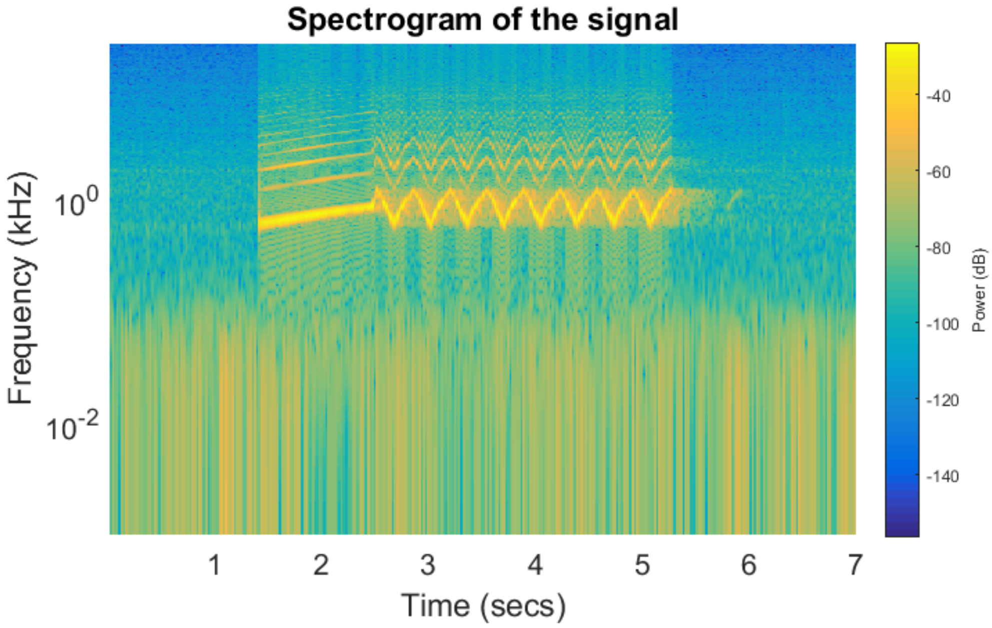

Figure A.6: Spectrogram of the yelp. As the siren is turned on it "builds up"

In the figure below 2.8 the power spectral density of the yelp siren can be seen. In this case the difference in level between the fundamental and the harmonics is lower than in the previous siren.

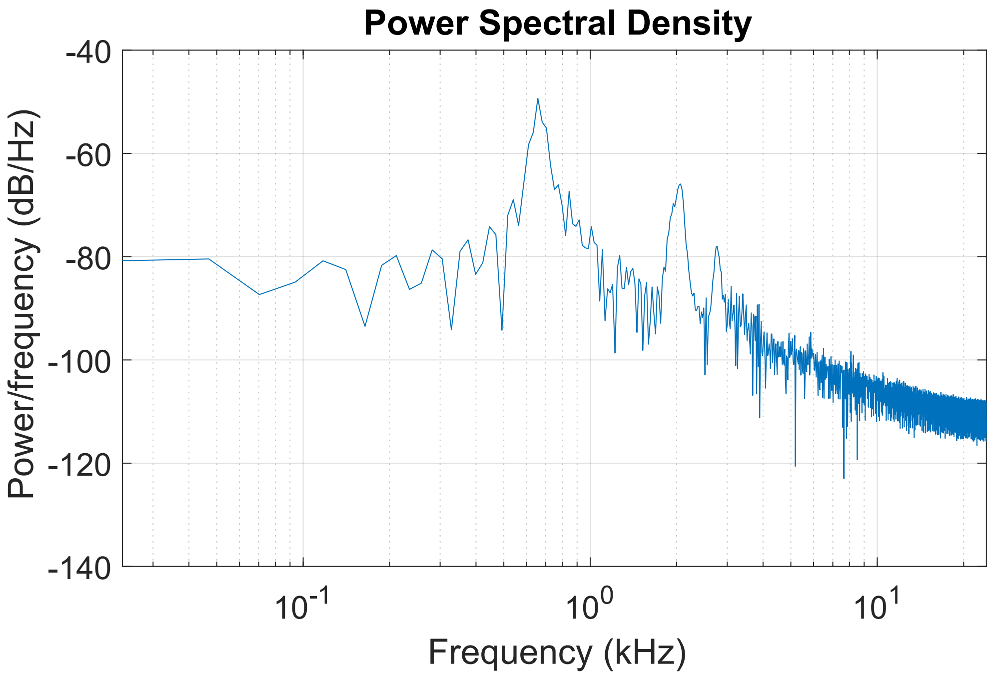

Figure A.7: Power spectral density of 30ms of the yelp.

The levels of the yelp siren are measured at 3 meters distance to the ambulance - from there the LC_{peak} value is measured.

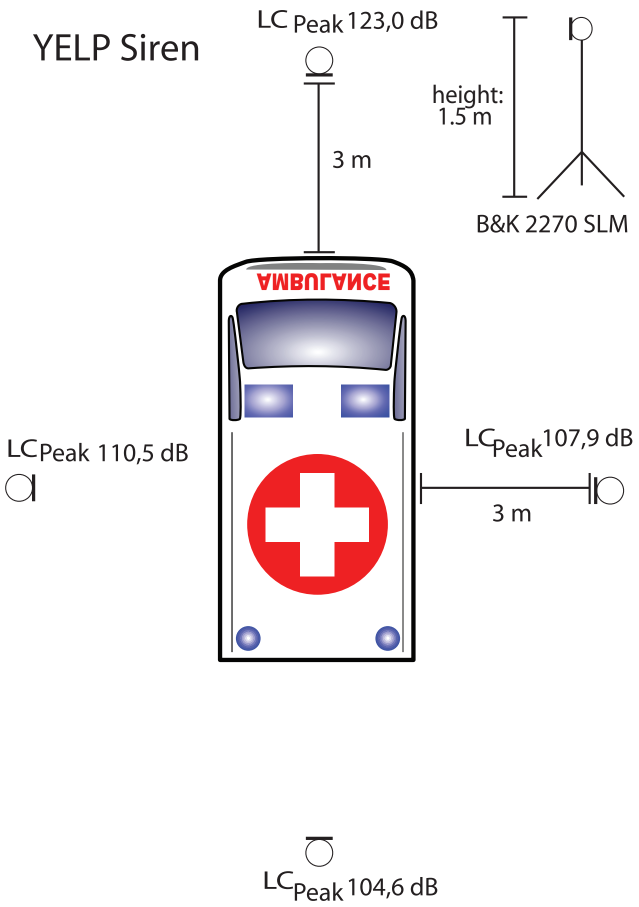

Figure A.8: LC_{peak} values of the yelp siren measured at 3 meters distance to the siren from all sides.

Measurement Setup 2: In Motion

Second measurement setup was placed along Håndværkervej to catch one of the ambulances driving with the sirens on. The goal here is to measure the siren in action, while driving. Unfortunately it was not possible to have position tracking on the siren, so the acceleration and velocity cannot be determined. Secondly to record the effects of the Doppler effect of the siren. The traffic noise/background noise measured 76 dB $LC_{eq_{10\text{min}}}$ on this busy street.

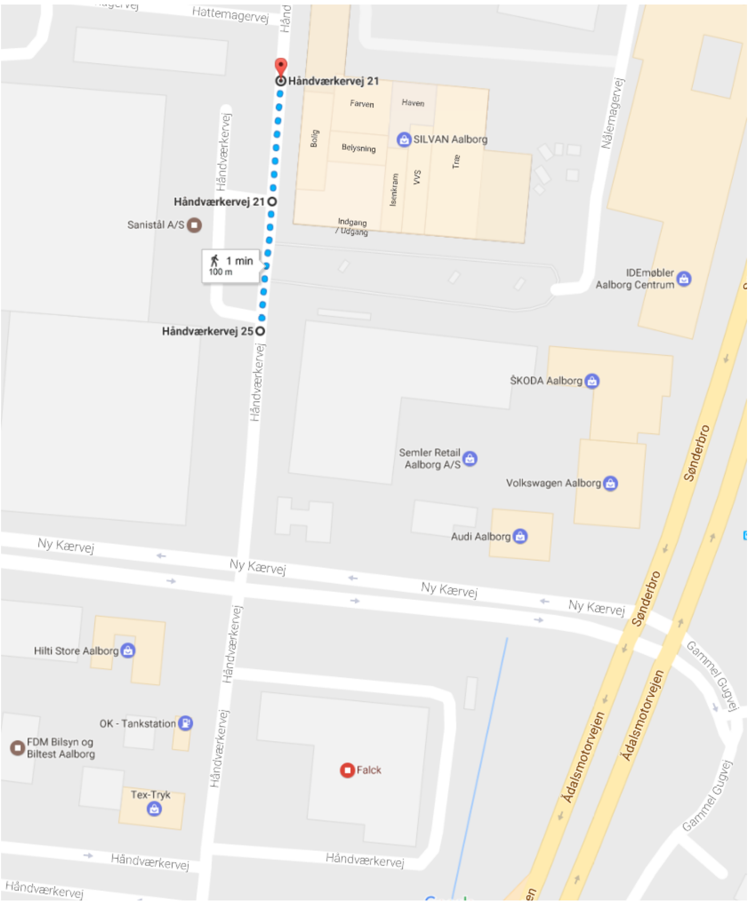

Figure A.9: Three measurement points along Håndværkervej with a 50 meter spacing. The measurement positions are positioned 200 meter from the exist from the Falck station to negate the initial acceleration from exiting the Falck station.

B | Siren Analysis of Fundamental & Harmonic Frequencies

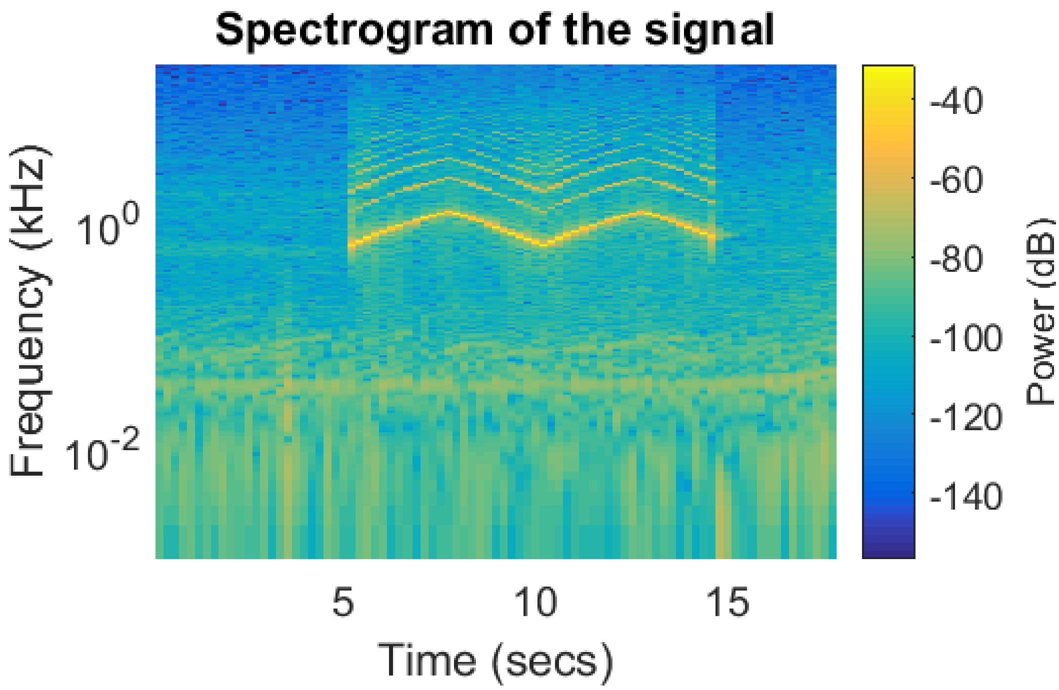

Figure B.1: Spectrogram of siren stationary at 3 meter distance, in front of the ambulance.

Table B.1: Fundamental and harmonics frequencies (stationary).

Front Stationary (reference)	Fundamental	2nd	3rd	4th	5th	6th
Start (5.5 sec)	704 Hz	1360 Hz	2067 Hz	2770 Hz	3476 Hz	4152 Hz
Peak (8.5 sec)	1032 Hz	2065 Hz	3097 Hz	4152 Hz	5185 Hz	6217 Hz
Dip (11 sec)	809 Hz	1618 Hz	2439 Hz	3248 Hz	4068 Hz	4842 Hz
End (14.5 sec)	727 Hz	1408 Hz	2135 Hz	3075 Hz	3613 Hz	4293 Hz

Table B.2: Fundamental and harmonics ratios (stationary).

Ratios	Fundamental-2nd	2nd-3rd	3rd-4th	4th-5th	5th-6th
Start (5.5 sec)	1.93	1.52	1.34	1.25	1.19
Peak (8.5 sec)	2.0	1.5	1.34	1.28	1.19
Dip (11 sec)	2.0	1.5	1.33	1.25	1.19
End (14.5 sec)	1.93	1.51	1.4	1.17	1.18

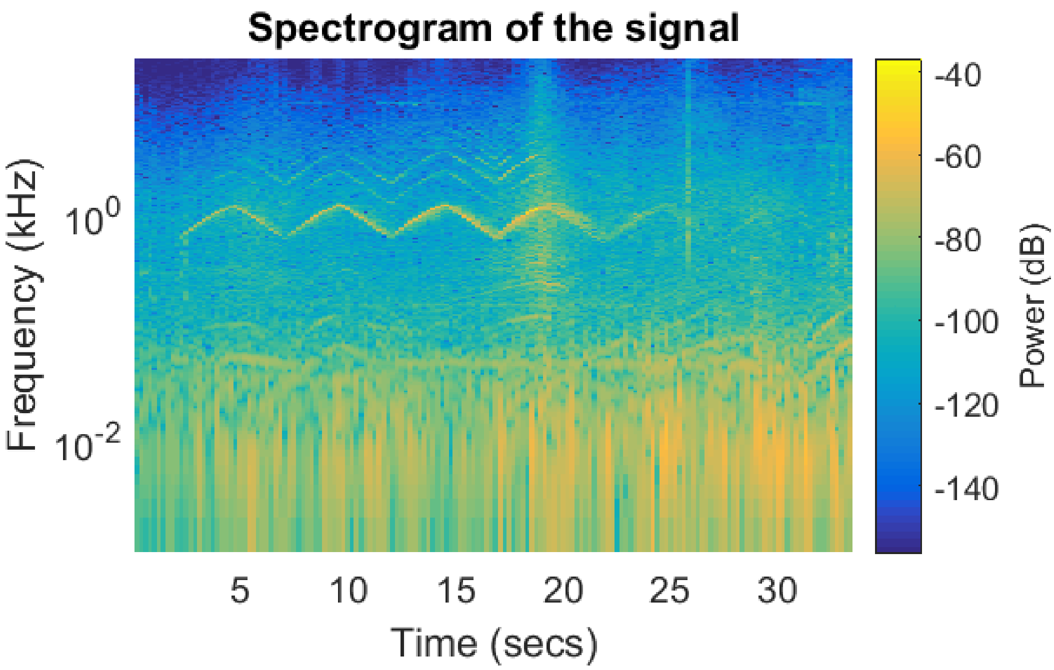

Figure B.2: Spectrogram of siren approaching Sound Level Meter.

Table B.3: Fundamental and harmonics frequencies (moving).

Moving Ambulance	Fundamental	2nd	3rd	4th	5th	6th
Start (2.2 sec)	633 Hz	1243 Hz	2065 Hz	3425 Hz	3871 Hz	4223 Hz
Before close (16.7 sec)	656 Hz	1267 Hz	1877 Hz	2557 Hz	3191 Hz	3824 Hz
Passing SLM (peak) (19 sec)	1126 Hz	2252 Hz	3355 Hz	4551 Hz	5701 Hz (noise)	7038 Hz (noise)
Leaving (22 sec)	656 Hz	1360 Hz	1876 Hz	2849 Hz	3424 Hz	3951 Hz

Table B.4: Fundamental and harmonics ratios (moving).

Ratios	Fundamental-2nd	2nd-3rd	3rd-4th	4th-5th	5th-6th
Start (2.2 sec)	1.96	1.66	1.66	1.13	1.09
Before close (16.7 sec)	1.93	1.48	1.24	1.19	-
Passing SLM (peak) (19 sec)	1.99	1.48	1.35	1.25	1.23
Leaving (22 sec)	2.07	1.38	1.2	1.15	-