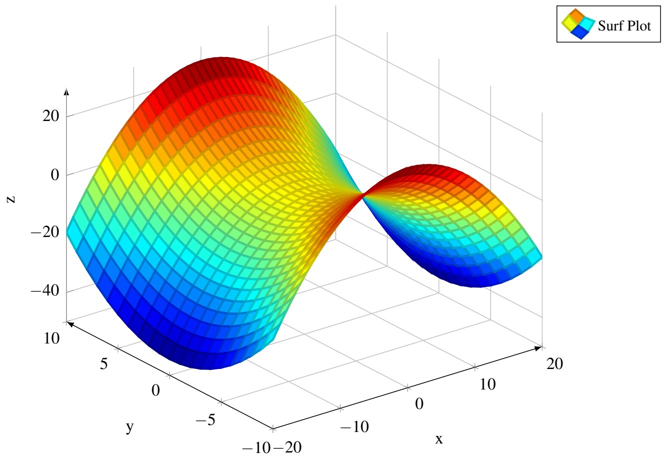
Worksheets for New Generation Guitar Tuner

September 2014

Group 714

Project worksheet



Aalborg University
Sound & Music Computing
Rendsburggade 14
DK-9000 Aalborg

AALBORG UNIVERSITY
STUDENT REPORT

AD:MT
Rendsburggade 14
DK-9000 Aalborg Ø
http://smc.aau.dk

Title:
Worksheet for 1st Semester M.Sc. Article

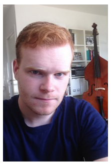
Theme:
Foundations of SMC

Project Period:
Autumn semester 2014

Project Group:
714

Authors:

Karl Heding



Jakob Aaen

Anders Christensen

Jacob Møller

Supervisor:
Mads Græsbøll Christensen

Number of Prints:
none

Number of Pages:
24

Date for Submission:
Dec 18, 2014

Abstract:

Empty.

Preface

Many electronic applications have signal processing as an important part of the system. The theme of this semester is .... The level is a 1st semester M.Sc project rated to 15 ECTS-points within AD:MT at Aalborg University. The target group of this report are supervisors and future developers of audio technologies such as students and other interested parties within the School of Information and Communication Technology at the The Faculty of Engineering and Science. The central aspects of this project are dealing with design considerations ... bla bla.

Reading Guide

The report is a worksheet for an article and therefore the structure is loose. There is one part called Worksheet.

The Analysis will be dealing with definition of the technology which will be developed and therefore it contains the necessary discussions for a Requirement Specification, which will result in guidelines and instructions for the Design and Implementation part. The Design and Implementation part contains development of the system which is implemented in the DSP and solutions to the issues discussed.

Appendices

Appendices are found after the main report **FiXme Fatal: insert reference** and on the attached CD. The appendices contains all source code, an extended analysis, measurement reports and work sheets. The CD also contains a digital version of the report and the web pages used. A complete list of the CD content can be found here **FiXme Fatal: insert ref**. When referred to the CD the following symbol is used

`path/to/stuff`

All figures, tables and equations are referred to by the number of the chapter they are used in, followed by a number indicating the number of figure, table or equation in the specific chapter. Hence each figure has a unique number, which is also printed at the bottom of the figure along with a caption. The same applies to tables and equations, the latter of which have no captions. Appendices are referred to by capital letters instead of chapter numbers.

Bibliography

At the very end of the main report, a Bibliography is listed which contains all sources of information used in the report. In the Bibliography books are indicated with author, title, publisher and year. Web pages are indicated with author and title. All information sources are referred to by the number which they feature in the list. This will look like this: [number]. If the sources reference is written just before a final dot, then it represents the source of information for the sentence in question. If the reference is placed after the dot it represents the source of information for the section in question.

Contents

Part I

Analysis

Chapter 1

Introduction

1.1 Motivation

Recently, new technologies for designing guitar tuners has been developed. These are called poly tuners. They have the ability to provide a the guitar player with a quick overview of the pitch of all strings, in order to speed up the process of tuning a guitar. Such technologies aims to estimate the pitch of several strings, simultaneously. Some limitations of such technologies are estimation of the extremely low fundamental frequencies and separation from background environmental noise**FiXme Fatal: Verify these assumptions** and it seems that the poly-tuner has difficulties if only one string is tabbed. Due to the limitations of the existing poly tuner, new methods for multi-pitch estimation is investigated. Therefore, this project aims to investigate the theory of multi-pitch estimation and consequently develop a software prototype application for improving some of these limitations **FiXme Fatal: Insert Table or a figure of overview of limitations**.

1.2 Initial Problem

Based on the motivation of this study an initial problem is formed in order to structure the further analysis.

How can multi-pitch estimation improve poly-tuner technology?

- Verify limitations of existing technologies.
- Investigate the physics of the guitar and how to model the guitar strings.
- Investigate methods for multi-pitch estimation.
- Set up a solution approach for improving the verified limitations.
- Investigate implementation methods for a software prototype.

FiXme Fatal: Elaborate on the items

Chapter 2

Preliminary Analysis

Before coming to theory which has to be used in Multi-Pitch Estimation, some fundamentals theory has to be described. First the signal models and covariance matrix model. Next section is is about maximum likelihood estimation, which also includes some basic theory.The last section is about physics for a guitar string

2.1 Signal Models and Covariance Matrix Model

We begin by forming the basic form of a signal for $n = 0, ..., N-1$

$$x_k(n) = \sum_{l=1}^{L_k} a_{k,l} e^{j\omega_k l n} + e_k(n) \tag{2.1}$$

The signal model $x(n)$ is calculated for every harmonic l up to L_k with the complex amplitude written as

$$a_{k,l} = A_{k,l} e^{j\phi_{k,l}} \tag{2.2}$$

Where $A_{k,l}$ is the amplitude for the signal k and the amplitude l and ϕ the phase. ω_k is the fundamental frequency for the k'th signal and $e_k(n)$ is the noise.

The unknown real parameters are put in a vector which we define with T denoting the transpose

$$\Theta_k = \left[\omega_k A_{k,l} \phi_{k,l} \cdots A_{k,L_k} \phi_{k,L_k}\right]^T \tag{2.3}$$

As seen, only the frequency for the k'th signal is needed as this is the fundamental frequency for all harmonics.

For cases with more sources, we write the total signal as

$$x_k(n) = \sum_{k=1}^{K} \sum_{l=1}^{L_k} a_{k,l} e^{j\omega_k l n} + e_k(n) \tag{2.4}$$

where we sum for the harmonics for each source, and take the sum of the sources to get the total signal.

As we later need to work with sub-vectors we define these for a signal with $M \leq N$ as

$$x(n) = \left[x(n) \cdots x(n+M-1)\right[^T \tag{2.5}$$

We define a Vandermonde structure like matrix with vectors $z(\omega) = \left[1 e^{j\omega} \cdots e^{j\omega(M-1}\right]^T$ as

$$Z(k) = \left[z(\omega_k) z(\omega_k 2) \cdots z(\omega_k L_k\right] \tag{2.6}$$

From the equation we split the matrix $Z_k(n)$ by

$$Z_k(n) = Z_k D_n \tag{2.7}$$

and can derive the diagonal matrix written with diag$(\cdot)$ for an easier read

$$D_n = diag\left(\left[e^{j\omega_k 1 n} \cdots e^{j\omega_k L_k n}\right]\right) \tag{2.8}$$

Giving us the possibility to divide into different parameters.

As we will use the covariance matrix of the sub-vectors we define it here with $E\{\cdot\}$ the expected and $(\cdot)^H$ the conjugate transpose as

$$R = E\left\{x(n)x^H(n)\right\} \tag{2.9}$$

This matrix is unknown in practice and we will therefore generally replace it with the sample covariance matrix

$$\widehat{R} = \frac{1}{N-M+1}\sum_{n=0}^{N-M} x(n)x^H(n) \tag{2.10}$$

For multiple sources we define the covariance matrix with Q as the covariance matrix for the noise $e(n)$ and P_k the covariance matrix for the amplitude as

$$R = \sum_{k=1}^{K} Z_k P_k Z_k^H + Q \tag{2.11}$$

$$Q = E\left\{e(n)e^H(n)\right\} \tag{2.12}$$

$$P_k = E\left\{a_k(n)a_k^H(n)\right\} \tag{2.13}$$

2.2 Multi-Pitch Estimation Section 2.2 (Maximum Likelihood Estimation)

The maximum likelihood method is used for finding parameters of a single periodic signal in Multi-Pitch estimation, but first a couple of basic terms and methods, which will be used, is needed to be shortly explained. The source for the basic is [3]

2.2.1 Basic Terms Explanation

The probability mass function, cumulative distribution, probability density function will be explained, the weighted least square and maximum likelihood estimators will be explained.

Probability Mass Function

If the random variables possible values are a sequence, they are said to be discrete. When looking on a discrete random variable X, the term to use it probability mass function, $p(a)$ of X.

$$p(a) = P\{X = a\} \tag{2.14}$$

The probability mass function has following values

$$p(x_i) > 0, \quad i = 1, 2... \tag{2.15}$$
$$p(x_i) = 0, \quad \text{other values of x} \tag{2.16}$$

The cumulative distribution function, CDF, is defined for any real number x as probability of the event $\{X \leq x\}$:

$$F(x) = P\{X \leq x\} \tag{2.17}$$

X must take on one of x_i values, which means:

$$\sum_{i=1}^{\infty} p(x_i) = 1 \tag{2.18}$$

The CDF of the probability mass function is a step-function:

$$p(a) = \sum_{all \leq a} p(x) \tag{2.19}$$

Probability Density Function

If the random variables are jointly continuously, or in other words the values are in a interval, the probability density function is used. X is an continuously random variable if a non-negative function $f(x)$ exist, defined from $x \in (-\infty, \infty)$ having property for any B real number:

$$P\{X \in B\} = \int_B f(x)dx \tag{2.20}$$

The function $f(x)$ is the probability density function of the random variable X. X must take on some value, it yields:

$$1 = P\{X \in (-\infty, \infty\} = \int_{-\infty}^{\infty} f(x)dx \tag{2.21}$$

The probability for a continuous random variable is zero for any particular value, in this case a:

$$0 = P\{X = a\} = \int_a^a f(x)dx \tag{2.22}$$

Relation between the cumulative distribution and the probability density is, the density is the derivative of the cdf.

Weighted Least Square

The regression model:

$$Y = \alpha + \beta x + e \tag{2.23}$$

has a variance, which changes according to its input level x_i. If variances are known, the α and β parameters should be estimated by using the weighted, w_i, least square method. The idea is to minimize the weighted sum of the squares. If the variance of Y_i:

$$\text{Var}(Y_i) = \frac{\sigma^2}{w_i} \tag{2.24}$$

then the estimators A and B should minimize:

$$\frac{1}{\sigma^2} \sum_i w_i (Y_i - A - Bx_i)^2 \tag{2.25}$$

Next step is to take the partial derivative to A then B and equal those to 0. The minimizers for A and B:

$$\sum_i w_i Y_i = A \sum_i w_i + B \sum_i w_i x_i \tag{2.26}$$

and

$$\sum_i w_i x_i Y_i = A \sum_i w_i x_i + B \sum_i w_i x_i^2 \tag{2.27}$$

This means two equations with two unknown parameters has to be solved.

Maximum Likelihood Estimators

Estimating a value of an unknown parameter is called an estimator. The normal estimating is to use the sample mean, e.g. 3 samples with values $x_1 = 3, x_2 = 5, x_3 = 1$, the estimated value will be 3.
Having some random variables $X_1..X_n$ given, with the joint distribution assumed to be given for the unknown parameter θ, the unknown θ can be estimated by using the random known variables. If the random variables are independent exponential random variables, they will have the same unknown θ. The joint density function of the random variables are then:

$$f(x_1, x_2, , , x_n) \tag{2.28}$$

The maximum likelihood estimator is obtained by letting $f(x_1, x_2, , , x_n|\theta)$ be the joint probability mass function of the random variables $X_1..X_n$ if they are discrete, and letting $f(x_1, x_2, , , x_n|\theta)$ joint probability density function when the random variables are jointly continuously.

A given signal is gaussian distributed with mean θ [2].fig. 2.1 shows that the input, $x[0]$ is dependeding on θ:

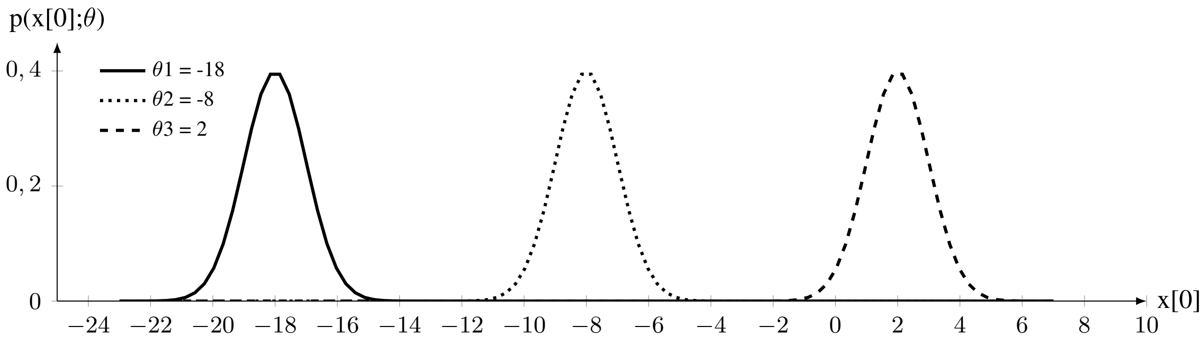


Figure 2.1: Dependence of PDF on unknown parameter

It can be seen that θ affacts the probability of $x[0]$. Since this it the case, it should be possible to infer what value θ should be according to the $x[0]$. If $x[0]$ is negative, the θ should be pick to be θ_3.

The likelihood is that $f(x_1, x_2, , , x_n)$ will be observed when θ is a true value. When it comes to the maximum likelihood estimate, $\hat{\theta}$, it is that value of θ which maximizes the likelihood function; $f(x_1, x_2, , , x_n|\theta)$. It should be noted that the log-likelihood and the likelihood have their maximum at the same θ.

2.2.2 Theory from Multi-Pitch Estimation section 2.2

From the Multi-Pitch Estimation [1], as mentioned in the beginning of the section, the maximum likelihood estimation is used to finding parameters of a single period signal. This method operate on a signal sub-vector constructed from the signal-pitch $\mathbf{x}_k(n)$:

$$\mathbf{x}_k(n) = [x_k(n)...x_k(n + M - 1)]^T \tag{2.29}$$

Often this kind of sub-vector can me modelled as a sum of the harmonics $\mathbf{L}_k$ in colored Gaussian noise $\mathbf{e}_k$ having covariance matrix $\mathbf{Q}_k$. This gives the following eqaution:

$$\mathbf{x}_k(n) = \mathbf{s}_k(n) + \mathbf{e}_k(n) = \mathbf{x}_k(n) = \mathbf{s}_k(n) + \mathbf{e}_k(n) \tag{2.30}$$

Where $\mathbf{a}_k = [A_{k,l}e^{j\phi_{k,l}}...A_{k,l}e^{j\phi_{k,l}}]$ which is a vector of complex amplitudes.
The Vandermonde matrix $\mathbf{Z}_k(n)$ is described earlier in **FiXme Fatal: which section?**. The total number of sub-vectors are $G = N - M + 1$ where N are observations and M are length of $\mathbf{x}_k(n)$.

The signal and noise parameter is called $\boldsymbol{\theta}_k$. In this parameter the fundamentel frequency ω_k and complex amplitude $A_{k,l}e^{j\phi_{k,l}}$. This means that it also contains the L_k order and the noise covariance $\mathbf{Q}_k$. By assuming

$\mathbf{Q}_k$ is invertibale , that $\mathbf{s}_k(n)$ is stationary and $\mathbf{e}_k$ is independent and identically distributed over n, the likelihood function of the observed signal set sub-vectors are:

$$p(\{\mathbf{x}_k(n)\};\boldsymbol{\theta}_k) = \prod_{n=0}^{G-1} p(\{\mathbf{x}_k(n)\};\boldsymbol{\theta}_k) \tag{2.31}$$

$$= \frac{1}{\pi^{MG}\det(\mathbf{Q}_k)^G} e^{-\sum_{n=0}^{G-1}\mathbf{e}_k^H(n)\mathbf{Q}_k^{-1}\mathbf{e}_k(n)} \tag{2.32}$$

By taking the logarithm the result is the log-likelihood function:

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}_k) = \sum_{n=0}^{G-1} \ln p(\{\mathbf{x}_k(n)\};\boldsymbol{\theta}_k) \tag{2.33}$$

$$= -GM\ln\pi - G\ln\det(\mathbf{Q}_k) - \sum_{n=0}^{G-1}\mathbf{e}_k^H(n)\mathbf{Q}_k^{-1}\mathbf{e}_k(n) \tag{2.34}$$

Finally the maximum likelihood estimates is then for the parameters $\boldsymbol{\theta}_k$:

$$\hat{\boldsymbol{\theta}} = \arg\max\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}_k) \tag{2.35}$$

It is found by estimateing the noise vector by:

$$\overset{\wedge}{\mathbf{e}}_k = \hat{\mathbf{x}}_k(n) - \hat{\mathbf{s}}_k(n) \tag{2.36}$$

Note that the deterministic part of $\mathbf{s}_k(n)$ is made by the parameters estimaters.

It should also be noted that the fundamental frequency is a nonlinear parameter and noise covariance matrix is unknown. The complex amplitude $\mathbf{a}_k$ are found by using weighted least-square:

$$\hat{\mathbf{a}} = \left(\sum_{n=0}^{G-1}\mathbf{Z}_k^H(n)\mathbf{Q}_k^{-1}\mathbf{Z}_k(n)\right)^{-1}\sum_{n=0}^{G-1}\mathbf{Z}_k^H(n)\mathbf{Q}_k^{-1}\mathbf{x}(n) \tag{2.37}$$

2.3 Guitar String Harmonics in Practice and Theory

In the following section the physics of the guitar strings will be analyzed, briefly. Firstly, this is done by analyzing the harmonics of 6 individual open strings of an electric guitar. Note that it is only done for one recording on six open strings on an electric guitar. All audio material and source code for figures in this section is available in the following directory. The software is very easy to alter to in order to analyse other recordings.

`analysis/code/string_analysis/`

In order to initialize the modeling of guitar strings, 6 recordings of guitar strings are analyzed. The spectrograms of their string response is used for evaluation of the relations between the harmonic frequencies and the pitch of a guitar string. These results will be compared to the inharmonicity model of [4], which relates the harmonic interval to the stiffness of a guitar string as in eq. (2.38).

$$f_n = nf_1[1 + (n^2 - 1)A], \tag{2.38}$$

for a solid wire without wrapping $A = \frac{\pi^3 r^4 E}{8TL^2}$,

- r is the radius of the string.
- E is Young's Modulus.

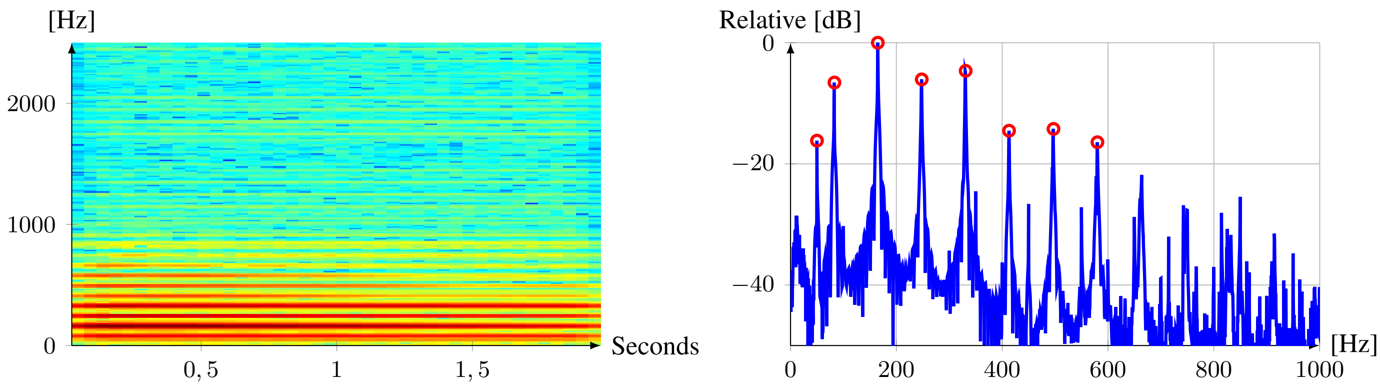


Figure 2.2: Frequency content of the lowest E-string with indicated peaks.

- T is the tension.
- L is the length.

The inharmonicity is smallest for long, thin wires under great tension. The preceding experiment is compared to the simpler form of a stiffness model of inharmonicity of strings from [1]. This will be called the Ξ model in the following. The Ξ model is shown in 2.39.

$$\Xi_{k,l} = \omega_k l \sqrt{1 + B_k l^2} \tag{2.39}$$

where

- k is the index of the fundamental frequency.
- ω is the fundamental frequency.
- l is the index of the harmonic frequency.
- $B << 1$ is an unknown stiffness parameter of the specific string.

2.3.1 Experiment

Six open strings is being recorded individually. They are all tuned referenced to $\mathtt{A} = 440\,\text{Hz}$. The harmonics is identified by locating the first eight peaks in the magnitude response of the frequency domain. Every figure that shows the frequency resopone in this section is produced from a two seconds file, sampled at $fs = 44.1\,\text{kHz}$. The spectrogram of fig. 2.2 depicts the frequencies of the lowest E-string. The harmonics below 800 Hz is extracted and the peak-indices can be located from a full windowed FFT, which is plotted as the magnitude response, normalized to the highest peak. At first glance, it seems that the harmonics are close to the integer multiple of the fundamentals. Therefore, all peaks are extracted and put in a table of harmonics table 2.1a. Be aware, that from the plots it is seen that they all suffer from 50 Hz hum noise, which is not shown in the table.

Table 2.1a shows that the harmonics are integer multiples of the fundamental w_k with a deviation $\mathbf{r} < 3.5\,\text{Hz}$. The deviation $\mathbf{r}$ is shown in table 2.1b. Furthermore, some of the harmonics are shared between the individual strings e.g. 440 Hz, 588 Hz and 988 Hz. The deviation is found from the distance to the expected integer multiple $\omega_{1,k}$. It is found from

$$\mathbf{r} = r_{i,k} = \omega_{1,k} i - \omega_{i,k}, \quad \mathtt{for} \quad i = \{1, \cdots, 7\}$$

, and k is the denoting the different sources. N.B. Due to the insufficient amount of tests, it is difficult to do a relevant conclusion from this standalone experiment. Still, it is implicated from the comparison in **??** of the eq. (2.39) to recording, that an this open guitar string has linear integer harmonics. The graphs show the recordings in red and the model is shown for seven different values of B. These six recordings are similar to the case of $B = 0$. It is obvious that the next step is to compare the Ξ model to strings that are shorter i.e. all strings closed to the 12^{th} fret.

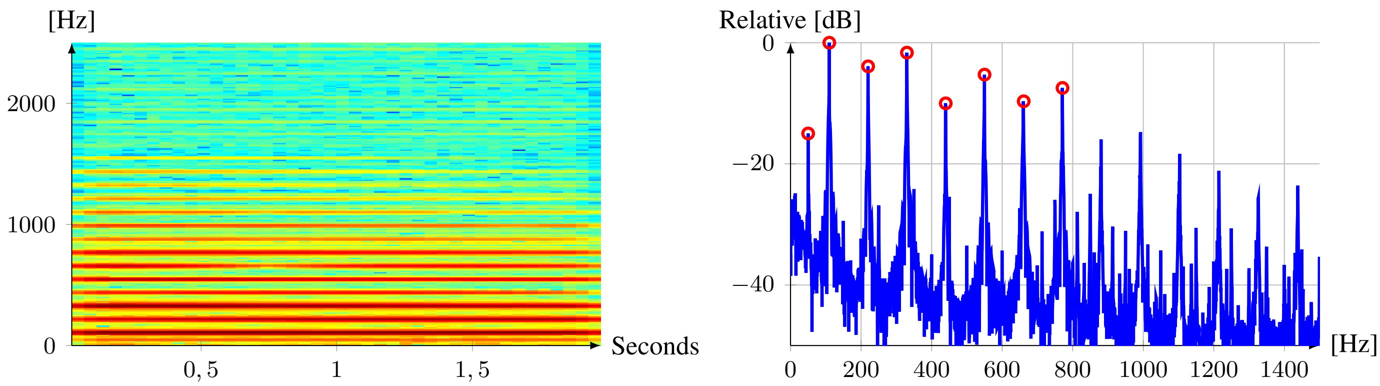


Figure 2.3: Frequency content of the A-string with indicated peaks.

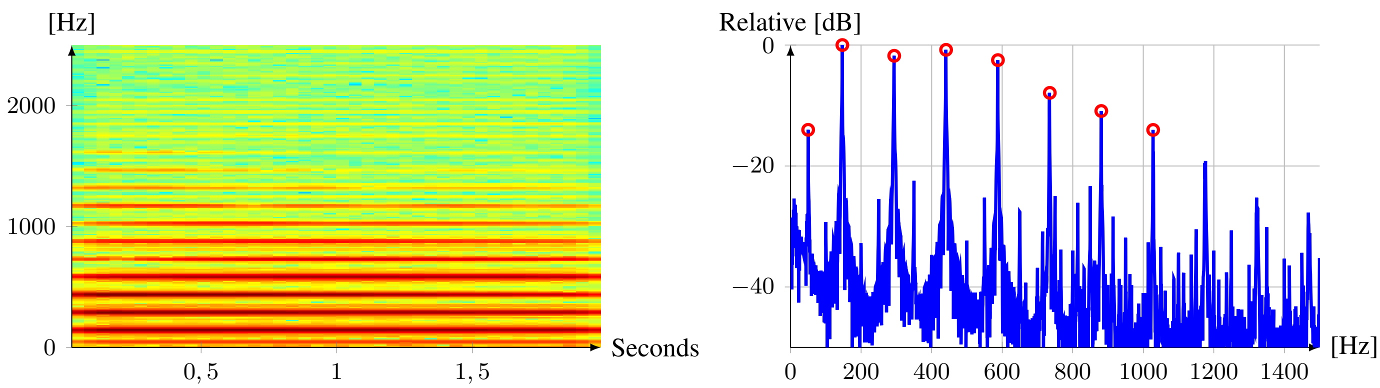


Figure 2.4: Frequency content of the D-string with indicated peaks.

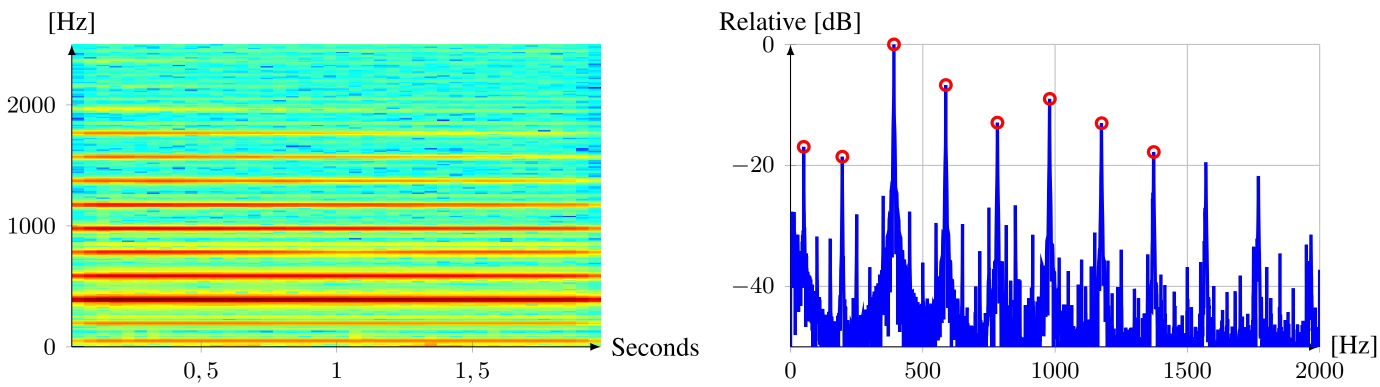


Figure 2.5: Frequency content of the G-string with indicated peaks.

The two models are compared in fig. 2.9, for $B = A$. The values of these are the same as in fig. 2.8. It can be seen that the two models are similar, only that the f_n model gives rise to higher harmonics along the l axis, which can be verified from eq. (2.38).

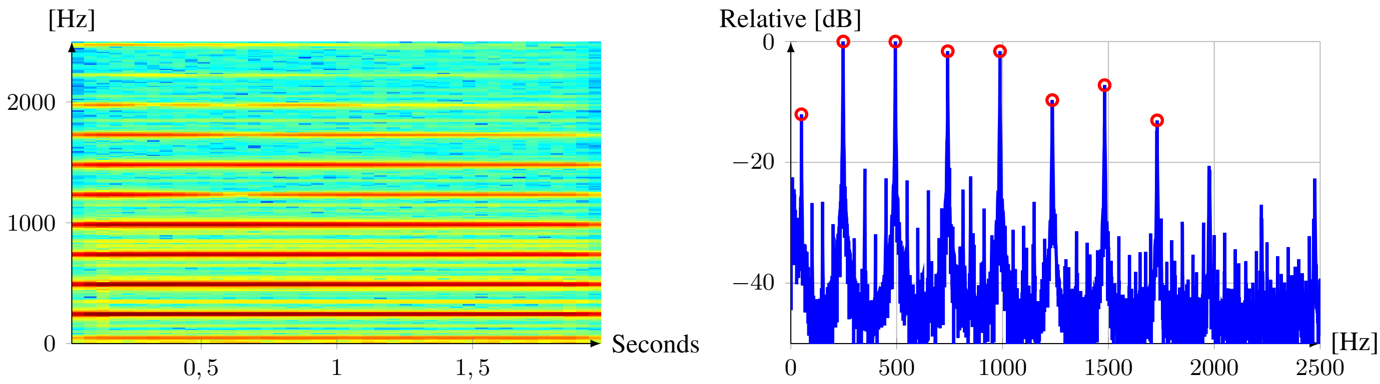


Figure 2.6: Frequency content of the H-string with indicated peaks.

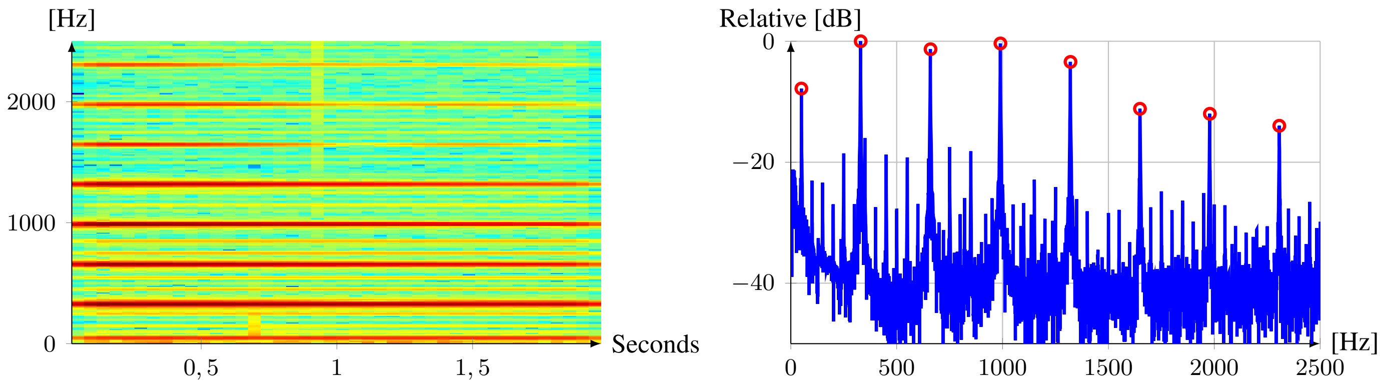


Figure 2.7: Frequency content of the highest E-string with indicated peaks.

1	E	A	D	G	H	E
1	83	110	147	196	248	330
2	165	220	294	392	494	660
3	248	330	441	587	471	990
4	331	440	588	782	988	1301
5	434	550	735	980	1235	1649
6	497	661	881	1176	1483	1979
7	580	771	1028	1373	1730	2307

(a) Measured Harmonic Frequencies

E	A	D	G	H	E
0	0	0	0	0	0
0.4993	0.4993	0.4993	0	0.9986	-0.4993
0.4993	0.9986	0.9986	0	1.9973	-1.4980
0.9986	1.4980	1.4980	0.4993	2.4966	-2.4966
0.9986	1.4980	1.9973	-0.9986	3.4952	-1.4980
0.4993	1.4980	2.4966	-1.4980	3.4952	-1.4980
0.0000	1.4980	2.9959	-2.4966	3.4952	-0.4993

(b) The distance to the integer multiples

Table 2.1: Harmonics frequencies of experiment.

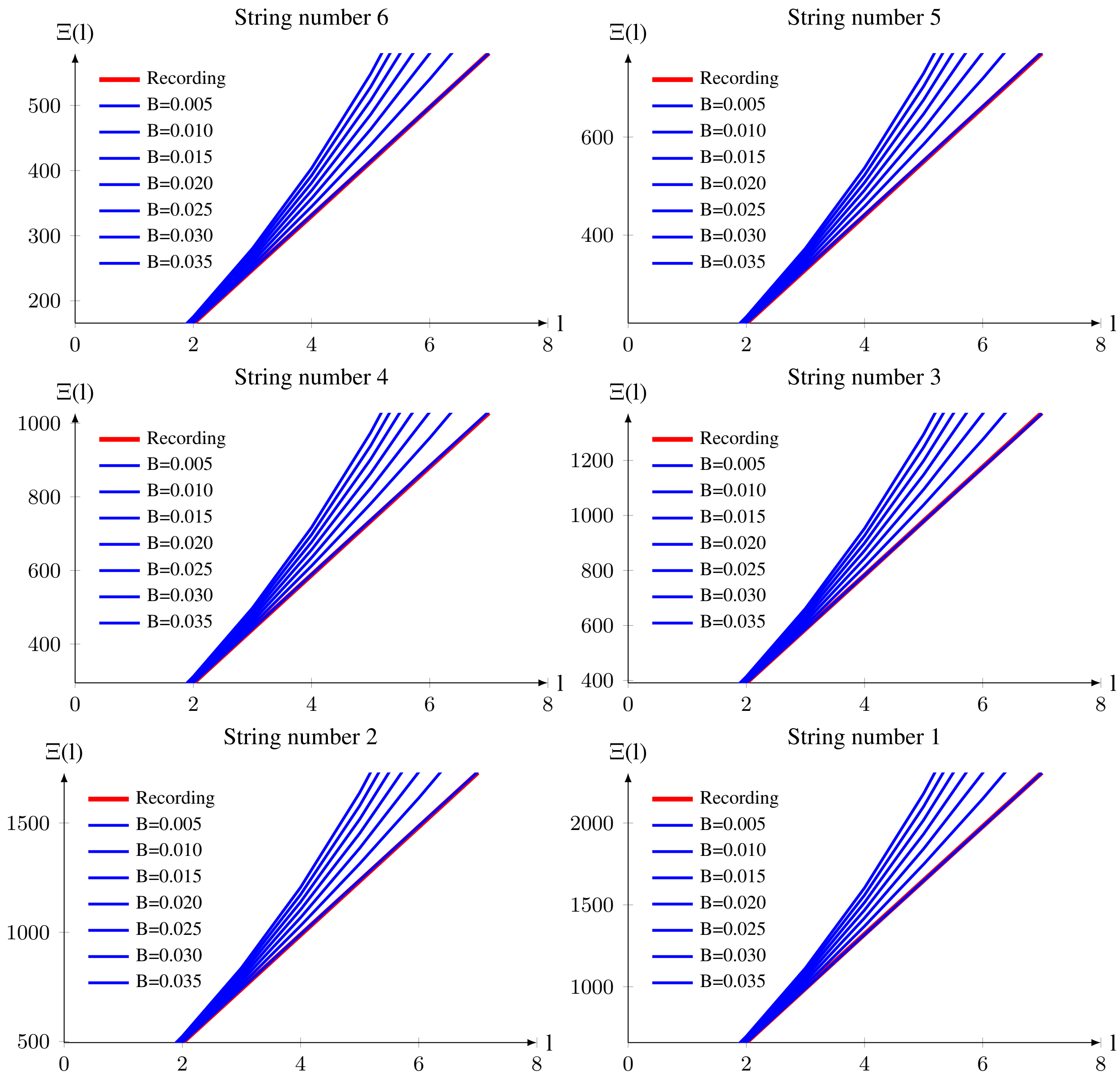


Figure 2.8: Ξ model compared to the recordings

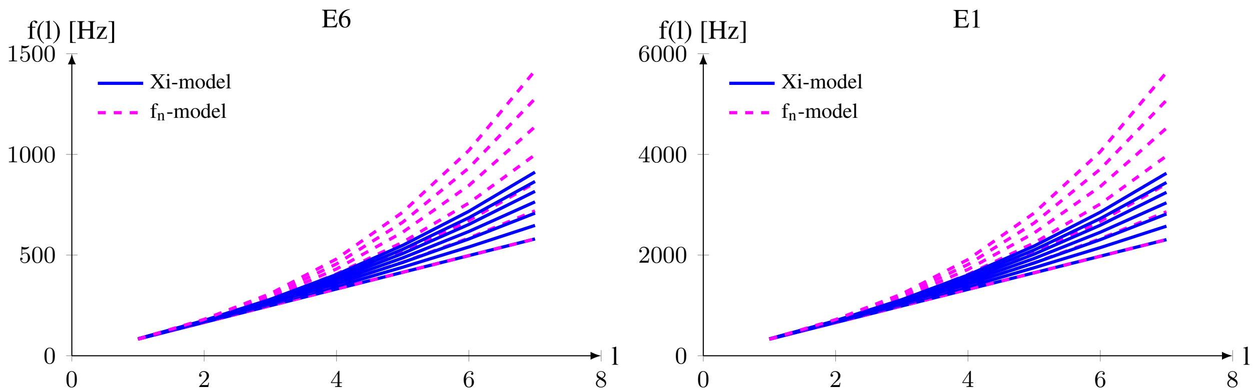


Figure 2.9: Ξ model compared to f_n model.

Chapter 3

Division of System into System Modules

Chapter 4

Interim Conclusion

Chapter 5

Requirements Specification

Chapter 6

Test Specification

Part II

Design

Part III

Implementation

Bibliography

[1] Mads Græsbøll Christensen and Andreas Jakobsson. *Multi-Pitch Estimation*. Morgan & Claypool, 2009.

[2] Steven M. Kay. *Fundamentals of Statistical Signal Processing*. Pearson Education, 2010.

[3] Sheldon M Ross. *Probability and Statistics for Engineers and Scientists*. unknown, 200x.

[4] F.Richard Moore Thomas D Rossing and Paul A. Wheeler. *The Science of Sound*. Addison Wesley, 2002.

Appendix CD

List of Corrections