Joonas Tuovinen

Signal Processing in a Semi-Automatic Piano Tuning System

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.

Espoo 18.2.2019

Thesis supervisor and advisor:

Prof. Vesa Välimäki

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE MASTER'S THESIS

Author: Joonas Tuovinen

Title: Signal Processing in a Semi-Automatic Piano Tuning System

Date: 18.2.2019 Language: English Number of pages: 7+51

Department of Signal Processing and Acoustics

Professorship: Audio Signal Processing Code: ELEC3030

Supervisor and instructor: Prof. Vesa Välimäki

Tuning a piano is a complicated and time-consuming process, mostly done by professional tuners. To make the process faster and non-dependent on the skills of a professional tuner, a semi-automatic piano tuning system is developed. The system composes of a stepper motor, an aluminium frame, an Arduino Uno, a microphone and a computer. The stepper motor changes the tune of piano strings by turning pins connected to them, the aluminium frame holds the motor in place and the Arduino controls the motor. The microphone and the computer are used as a part of a closed loop control system to tune the strings automatically. The control system requires the current fundamental frequency as well as the optimal, "in tune", fundamental frequency. In this thesis, literature as well as algorithms on finding the current fundamental frequency are reviewed and a novel tuning process for calculating the optimal fundamental frequency is developed. The novel tuning process is compared against a tuning done by a professional tuner. The process is well suited for the tuning of a piano, as there is very little difference between the tuning done by the process and the tuner. Between keys A_0 and $G\#_5$ the difference between the tunings is 2.5 cents and above that, in a region where the pitch of a string is much harder to hear, the difference is 8.1 cents.

Keywords: acoustic signal processing, audio systems, automatic control, music, spectral analysis

AALTO-YLIOPISTO
SÄHKÖTEKNIIKAN KORKEAKOULU

DIPLOMITYÖN TIIVISTELMÄ

Tekijä: Joonas Tuovinen

Työn nimi: Puoliautomaattisen pianonviritysjärjestelmän signaalinkäsittely

Päivämäärä: 18.2.2019 Kieli: Englanti Sivumäärä: 7+51

Signaalinkäsittelyn ja akustiikan laitos

Professuuri: Audiosignaalikäsittely Koodi: ELEC3030

Valvoja ja ohjaaja: Prof. Vesa Välimäki

Pianon viritys on monimutkainen ja aikaa vievä prosessi, jonka pääosin tekevät ammattilaisvirittäjät. Pianon virittämisen helpottamiseksi sekä viritysajan lyhentämiseksi on kehitetty puoliautomaattinen pianonviritysjärjestelmä. Järjestelmä koostuu askelmoottorista, alumiinikehikosta, Arduino Uno mikroprosessorista, mikrofonista sekä tietokoneesta. Askelmoottori ja alumiinirunko mahdollistavat pianon kielien vireen säätämisen ja Arduino, mikrofoni sekä tietokone toimivat osina takaisinkytkettyä säädintä, jonka avulla kielet pystytään virittämään automaattisesti. Kielen virettä hallitaan säätämällä kielen perustaajuutta. Tässä työssä tutkitaan miten pianon kielen perustaajuuden saa selvitettyä sekä kehitetään uusi menetelmä, jonka avulla saadaan laskettua perustaajuus, jossa kieli on parhaassa mahdollisessa vireessä. Kun tätä menetelmää verrataan ammattilaisvirittäjän tekemään viritykseen, niiden välillä on eroa vain 2,5 senttiä koskettimien A_0 sekä $G\#_5$ välillä, sekä 8,1 senttiä koskettimien A_5 ja C_8 välillä. Ylemmillä koskettimilla oleva suurempi ero johtuu siitä, että kielien virettä on vaikeampi kuulla näillä koskettimilla.

Avainsanat: akustinen signaalinkäsittely, audiojärjestelmät, automaattinen ohjaus, musiikki, spektrianalyysi

Preface

This Master's work was carried out in the Department of Signal Processing and Acoustics at Aalto university with funding Nordic Sound and Music Computing Network (NordicSMC), project no.86892. The funding period was from June 2018 throughout November 2018.

I want to thank Professor Vesa Välimäki on his guidance on the topic. I would also like to thank Jamin Hu, a student at the Sibelius Academy, who made the central part of the tuning system, the tuning robot, and came up with the subject for the thesis. Finally thanks to all my wonderful collages at the Aalto Acoustics Lab who helped with my work and provided a warm and pleasant working environment.

Otaniemi, 18.2.2019

Joonas M. Tuovinen

Contents

Symbols and abbreviations

Symbols

B	inharmonicity coefficient
f_k	*kth* partial of a tone
f_0	fundamental frequency
E	Young's modulus
L	length
r	radius
T	tension

Abbreviations

dB	decibel
DFT	discrete Fourier transform
ET	equal temperament
ET12	equal temperament twelve-tone
FFT	fast Fourier transform
MAT	median-adjustive trajectories
PFD	partial frequency deviation
PID	proportional-integral-derivative

1 Introduction

Tuning a piano is known to be a complicated process, which takes a considerable amount of time and patience. Tuning all the 200 plus strings of the instrument is a daunting task, especially as doing it incorrectly may leave the instrument in even worse tune. Because of this, the instrument is not commonly tuned by its owner, but the job is left to hands of a professional tuner.

The scale of the piano is based on the equal temperament twelve-tone scale (ET12) which specifies the fundamental frequencies of each string in the piano. The difficulty of tuning a piano comes from the fact that mode frequencies of piano strings deviate from the harmonic series, in an effect called inharmonicity. Because of this, the fundamental frequencies of piano strings can not be tuned to follow this scale, as doing so would lead the instrument to sound out of tune [1]. Instead, tuners use the beating effect, which is an amplitude modulation produced by two frequencies close to each other, to tune the instrument [2].

To make the process of tuning a piano faster and non-dependent on the skills of a professional tuner, a semi-automatic piano tuning system is developed. The aim of the system is to tune a grand piano with the help of a non-professional tuner. The instrument is to be tuned with at least the same accuracy as a tuning done by a professional tuner, and in less time. The tuning system is semi-automatic as some parts of the tuning process such as pressing a key and muting the strings that are not tuned at the moment can not be done by the system. However, these tasks are so simple that they can be performed without any expertise on piano tuning.

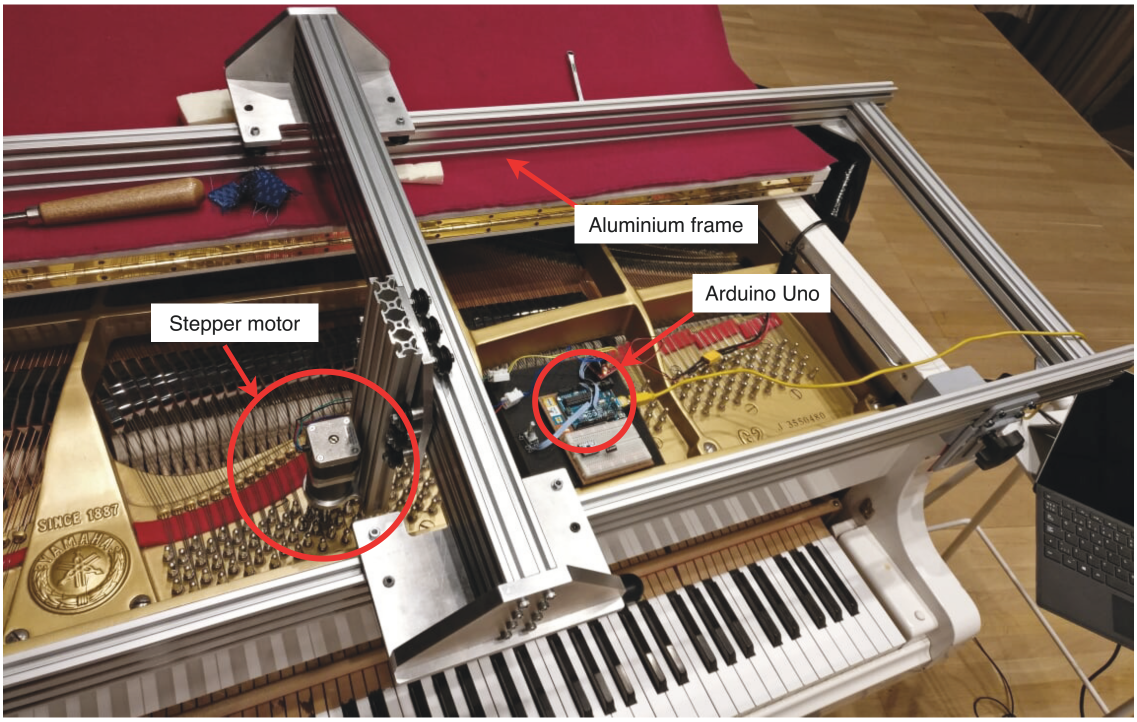
Piano tuners use a tuning lever to tune the strings. The lever is used to turn the pins , which change the tension (T) of the string, as the strings are wrapped around the pins. This change in tension changes the fundamental frequency (f_0) of the string, as T and f_0 are related by equation:

$$f_0 = \frac{\sqrt{\frac{T}{m/L}}}{2L}, \tag{1}$$

where m is the mass of the string and L is the length of the string.

The first step in making the piano tuning system was to create a structure which allows the automatic control of string tension. The proposed system, developed by Sibelius Academy student Jamin Hu, composes of a stepper motor, an aluminium frame and an Arduino Uno microprocessor. The stepper motor is capable of turning the pins of the piano with high precision (small angle) and has enough torque to turn even the high tension strings in the bass region of the piano. The aluminium frame has two functions: first, it keeps the body of the stepper motor in place, so a pin can be turned, and second, it allows the stepper motor to be placed on any of the pins of the grand piano. The Arduino is used to control the stepper motor. This part of the system, called the tuning robot, can be seen in Figure 1.

The next step is to automatically control the system. The control system needs to be able to determine the number of steps needed to get the strings of the piano in tune. A closed loop control system is used to do this. The general structure of a closed loop control system can be seen in Figure 2a. The system takes a reference



a

Figure 1: The tuning robot, composed out of a stepper motor, an aluminium frame and a Arduino.

value as its input, and the aim of the control system is to minimize the difference between the reference and a value measured from the output of the system. This is done with a controller, which controls the input of the process, based on the difference between the reference and the measured output.

Figure 2b shows a diagram of the piano tuning system. Here, the controller is the Arduino, which controls the number of steps that the stepper motor takes (process). The output of the system is measured as the f_0 of the string. To be able to measure this from the tone produced by a string, a microphone and a computer are added to the system. In addition to providing the measured output (current value of f_0), these two components are used to calculate the reference. The reference is the target, "in tune", f_0 of the string. The difference between the current and the target value of f_0 is used by the controller (Arduino) to calculate the appropriate number of steps needed to correct the difference in frequency.

The fundamental frequency of the tone is used for tuning as it, as well as the inharmonicity constant (which is discussed later) can be used to calculate the mode frequencies of piano tones. Piano tuners use beats, generated by mode frequencies of two tones, to tune the instrument, so fundamental frequency can be used to control the frequency of these beats.

In this thesis, existing methods and related literature on finding the fundamental frequency of a piano string are reviewed. This is done to find the first unknown in the input to the controller, the current value of f_0. To find the second unknown, the target value of f_0, literature about tuning a piano as well as existing methods

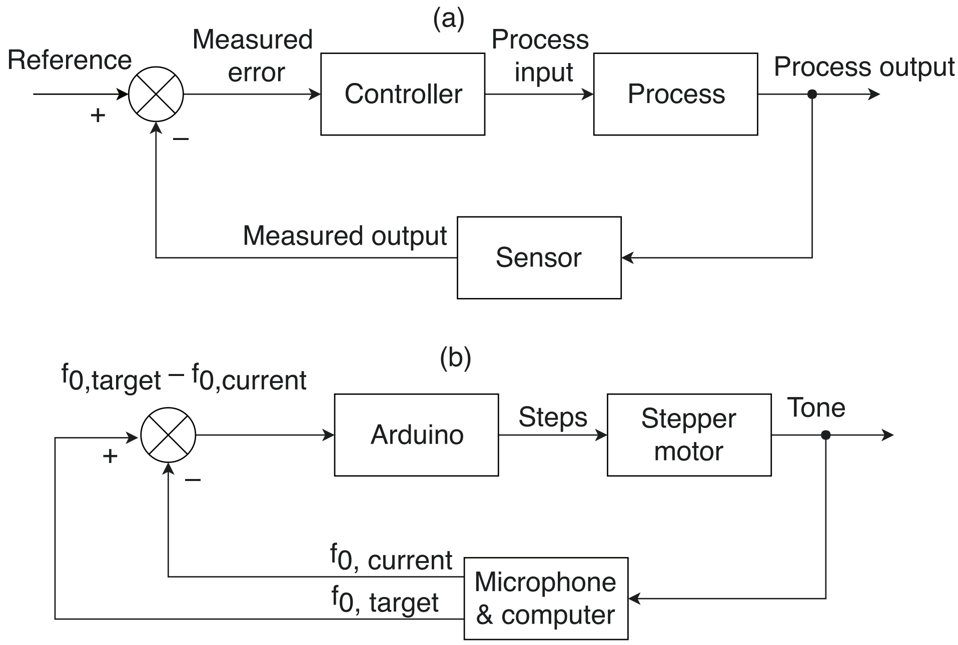


Figure 2: (a) general structure of a closed-loop control system (b) the structure of the closed-loop control system used in the piano tuning system.

for automatically calculating the target value of f_0 are reviewed. After this, a novel tuning process for finding the target value of f_0 is developed. The novel process finds the target frequency in a similar way to aural tuning (tuning by ear), as beating rates between currently tuned string and previously tuned strings are taken into consideration to find the best possible tuning.

After the current and the target value of f_0 are known, the controller can be used to automatically adjust the tension of a string to minimize the difference between the current value, and the target value of f_0. This is done using a proportional-integral-derivative (PID) control scheme. The PID controller can be used to automatically adjust the f_0 of the strings from the current one to the target value, thus tuning the strings of the piano. This part of the process is out of the scope of this thesis and will not be discussed further.

The thesis is structured as follows. In Section 2, literature about the physics of piano's sound production mechanism is reviewed in order to understand how the piano produces a tone and how the features of the various parts of the mechanism affect its tone and tuning. In Section 3, algorithms for finding the current value of f_0 are reviewed. In Section 4, literature related to the tuning of a piano are reviewed, in order to understand how a piano should be tuned, and tuning by ear, as well as mathematical simulations of tuning are reviewed. In Section 5, the novel tuning process is presented. In Section 6, the performance of the novel tuning process is evaluated. Finally, Section 7 concludes the thesis.

2 Physics of Piano

In this section, the physics of piano's sound production mechanism are reviewed. First, the general structure of the piano is discussed as well as how different parts interact with each other to get a clear picture of the instrument. After that the interactions affecting the tone of the piano, and thus the current fundamental frequency as well as its tuning, are reviewed.

2.1 General Structure

The piano has over 200 strings covering more than seven octaves (A_0 to C_8), which are played with the 88 keys of the keyboard. The keys are connected to the strings with the action, which transfers energy from the fingers of the player to the strings, by throwing a hammer against a string. The strike of a hammer causes a string to vibrate in frequencies determined by the features of the string, and this vibration is transferred to the soundboard through the bridge. The soundboard is a large, thin plate, which amplifies the vibrations of the strings and is the main source of radiated sound in a piano [3]. A diagram of a simplified version of a piano can be seen in Figure 3.

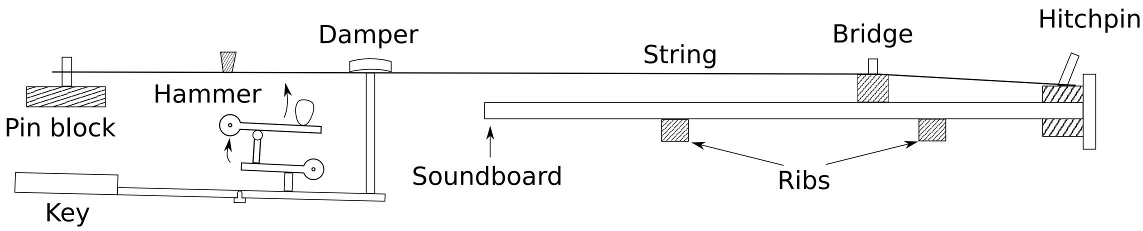


Figure 3: A simplified version of the piano adapted from [3].

Throughout the years, the piano has evolved to two distinct instruments: the modern grand piano and the upright piano. This thesis will solely focus on the modern grand piano as the tuning robot, used as a part of the tuning system is designed for it. A typical grand piano has 243 strings with the first eight keys of the bass end connected to a single string (hammer hits a single string), the next five keys connected to two strings and the rest connected to three strings. These strings vary in length from 2 m in the bass end to 5 cm in the treble end, and some of these strings are wrapped once or twice in wire. The strings are connected to two separate bridges: the bass bridge and the treble bridge. Figure 4 shows a drawing of a grand piano from above [4]. It can be seen how the strings on the bass bridge are on top of the strings in the treble bridge to maximize the length of the strings and to position the bass bridge closer to the middle of the soundboard.

It should be noted that the specification of a typical grand piano does not apply to all grand pianos. The number of strings connected to a key, the length of the strings and the number of double wound, single wound and unwound strings are design choices made by the manufacturer of the piano.

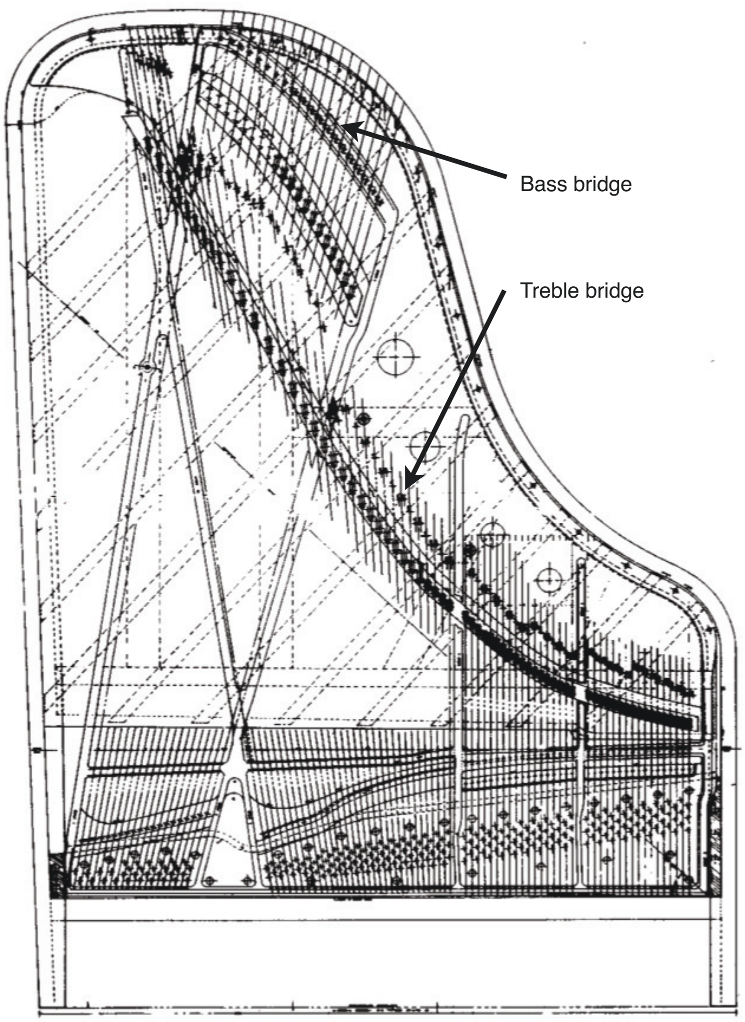


Figure 4: Drawing of a grand piano from above adopted from [4].

2.2 Strings and Inharmonicity

The strings of the piano vibrate in a series of modes. These modes have certain frequencies and amplitudes, which mainly determine the pitch of the string. The first seven modes of an ideal string can be seen in Figure 5.

An ideal string is a string with no stiffness and uniform linear density. The fundamental frequency of an ideal string can be calculated if the frequencies of its modes are known. This is because these mode frequencies are harmonics (integer multiples) of the fundamental and thus follow the equation:

$$f_k = f_0 k, \tag{2}$$

where k is the number the mode. This equation can be used to calculate f_0 if f_k and k are known.

However, Equation 2 can not be used to find the f_0 of piano strings, as they have stiffness. This stiffness acts as a restoring force, making the mode frequencies

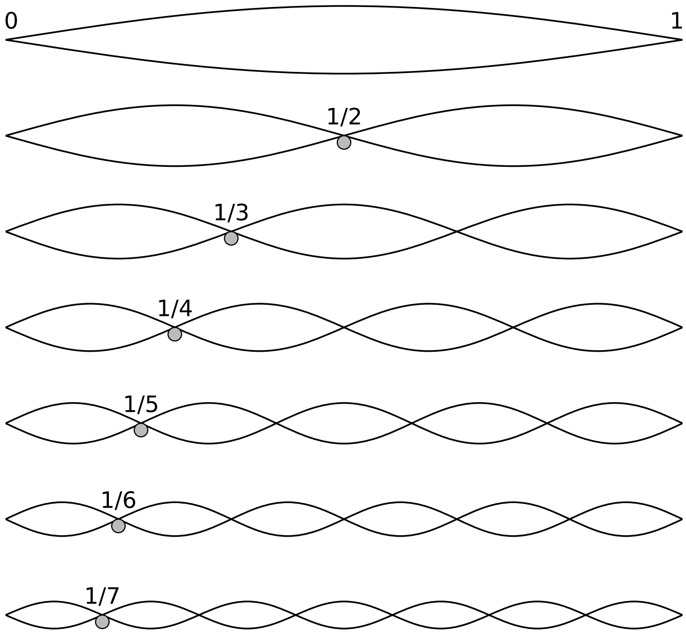


Figure 5: Modes of an ideal string. Figure adopted from [5].

slightly higher from their harmonic counterparts. This deviation from the harmonic series is called inharmonicity. The amount deviation is greater with higher modes, as they have more bends [3]. The mode frequencies of a piano string can be calculated from equation [6]:

$$f_k = kf_0\sqrt{1 + Bk^2} \tag{3}$$

where f_k is the frequency of the *kth* mode (also known as the *kth* partial), f_0 is the fundamental frequency, and B is the inharmonicity coefficient of the string. The value of inharmonicity coefficient of a solid string depends on the length, tension and radius of the string, according to the following equation [3]:

$$B = \frac{\pi^3 r^4 E}{8TL} \tag{4}$$

where r is the radius of the string, E is Young's modulus, T is tension, and L is the length of the string. The strings in the bass end are wrapped in wire to lower their fundamental frequency by increasing their linear mass. This increases their inharmonicity slightly, but not as much as adding the linear mass by increasing the radius of the solid string. Figure 6 shows how strings are wrapped in wire. The upper figure shows a single would string and the lower figure shows a double wound string.

Figure 7 shows the inharmonicity constants of piano strings obtained from the recordings of a Yamaha grand piano, as a function of the key number (1 is A_0 and 88 is C_8). It can be seen that as we go up the scale of the piano and the strings get shorter, the inharmonicity of the strings gets higher. At the bass, end the inharmonicity starts increasing again, because strings are wrapped in wire.

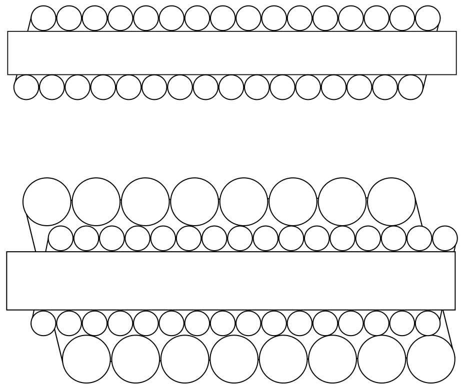


Figure 6: Graph of strings wrapped in wire.

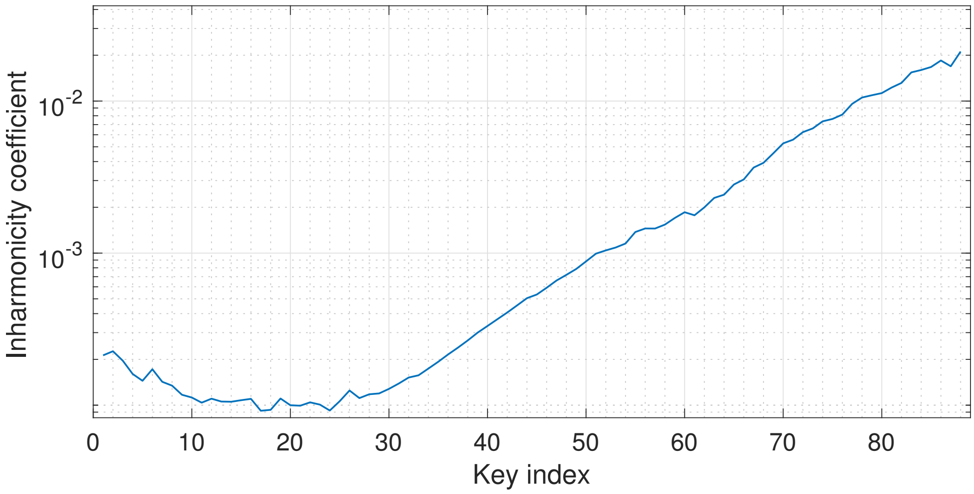


Figure 7: Inharmonicity coefficients of piano strings.

Figure 8 shows the effect of inharmonicity on the spectrum of a piano tone (A_4). The dotted lines show the harmonics of the fundamental frequency and the numbered peaks show the real peaks. It can be seen how the inharmonicity of the string affects the first few modes only slightly, but as the partial number grows greater, the difference between the harmonics and the real peaks get greater.

From Figure 8 varying magnitudes of string partials can also be seen. It can be seen how the first partial is the strongest one, dominating the sound of the instrument [7]. In the mid and high registers the first partial is usually the one with the greatest magnitude. In the bass register, one of the higher partials is often the strongest one.

The spectrum of a piano tone varies greatly between the bass, mid and treble registers. Figure 9 shows the partials of notes (a) D_1, (b) D_4 and (c) D_6. It can be seen how the higher partials of D_1 are stronger than the lowest ones and how the spectrum contains a high number of partials. Notes D_4 and D_6 on the other had contain a lower number of partials, with the first partial being the strongest one.

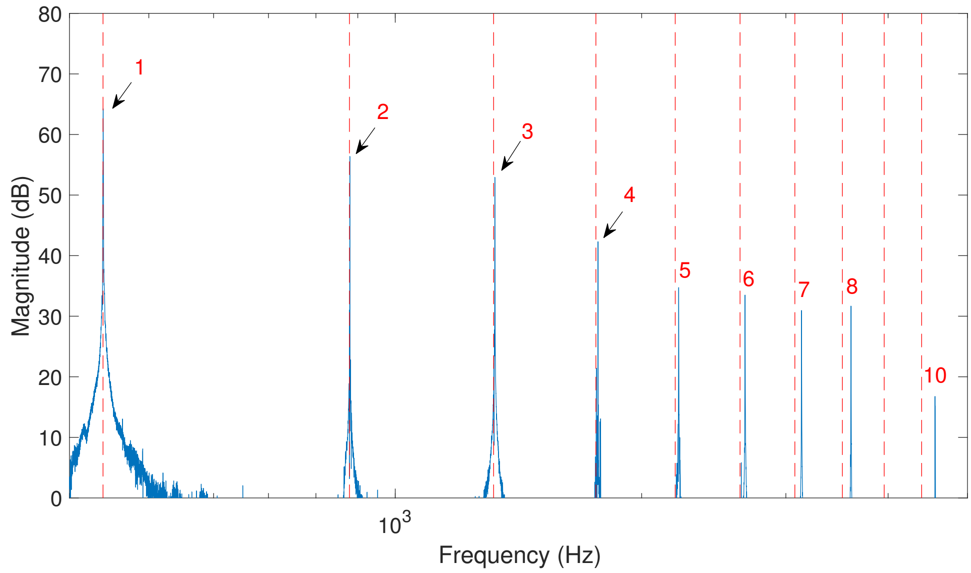


Figure 8: Figure shows the effect of inharmonicity in the spectrum of note A_4. Numbered peaks are spectral peaks belonging to string modes and dashed lines show the mode frequencies of a ideal string with harmonic mode frequencies.

2.3 Phantom Partials

In addition to the spectral components produced by string modes, other components can be found in the spectrum of a piano tone as well. These additional partials are called phantom partials. They appear at integer multiples of normal partials as well as at sums of two normal partials [8]. Phantom partials at integer multiples of normal partials are called "even" partials. Transverse vibrations of piano strings at a single frequency cause a longitudinal force at each end of the string at twice the transverse frequency [4]. This force produces the even phantom partials in the soundboard. Odd phantom partials are produced by nonlinearities in the structure of the piano or the string [8].

Figure 10 shows the spectrum of E_2. It can be seen how there are phantom partials close to the *13th*, *18th* and *19th* partial. The phantom partial close to the 13th partial has almost the same magnitude as the "real" partial produced by a string mode. These phantom partials make the task of finding the value of f_0 for a string much more difficult, as they can be confused with the one produced by string modes.

2.4 String Unisons

Most keys in the keyboard of the piano are connected to more than one strings. These sets of strings, known as string unisons, are used to amplify the movement of the soundboard, thus amplifying the amplitude of radiated sound. This is done because most strings do not have enough mass to achieve desired level of loudness

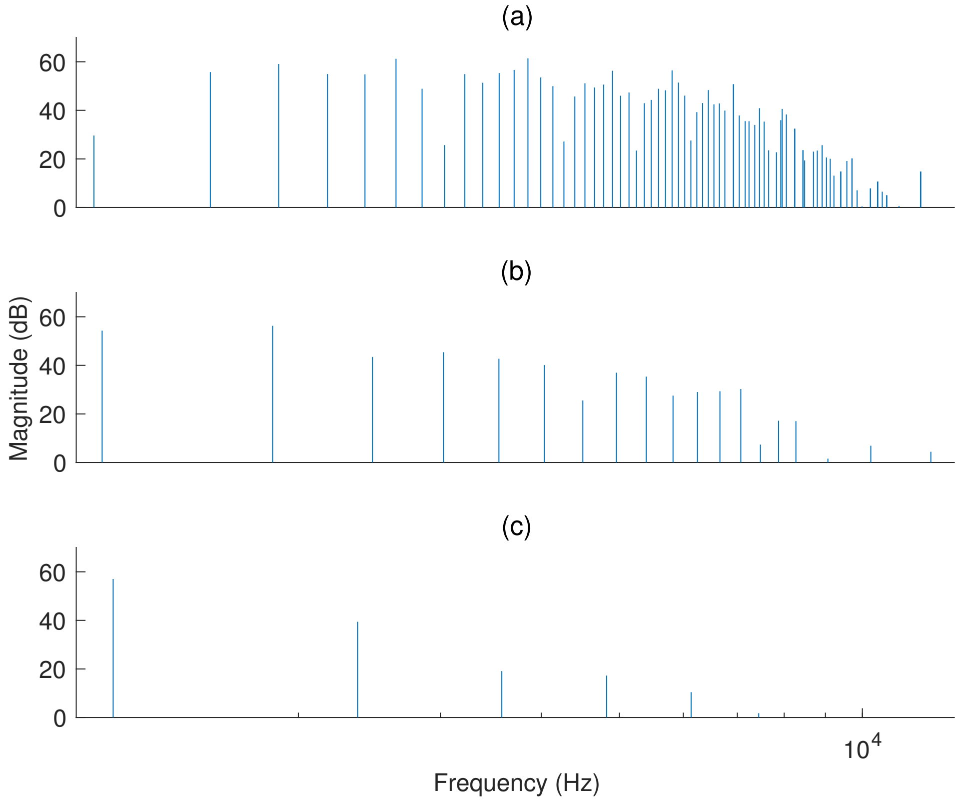


Figure 9: Magnitudes of partial frequencies of note (a) D_1, (b) D_4 and (c) D_6.

[7]. In the process of tuning the piano, a single string from each string unison is tuned by comparing it with other string unison and after that, all the other strings in the unison are tuned close to that string [2].

Unlike the name suggest, the strings in an unison are not tuned to the exact same tune. Instead the strings are tuned to slightly different frequencies so that the decay time of the tone is maximized. The decay time of piano tones is mainly determined by how rapidly the energy from the string is transferred to the soundboard trough the bridge. The mechanical impedance of the soundboard is much greater than the impedance of strings, so energy is transferred fairly slowly from the strings to the soundboard. Typical mechanical impedance ratio between the parallel vibrations of the strings and the soundboard is 200:1. The impedance mismatch is even greater between the parallel vibrations of the strings and the soundboard, as the soundboard is much stiffer in this direction [7].

The decay curves of piano tones usually have two stages: an initial fast decay and a slower final decay. This phenomenon is called double decay [9]. The effect is the most noticeable with string unisons. When strings vibrate, energy is transferred through the bridge to the soundboard causing the soundboard to vibrate with the same frequency as the string. If two strings vibrate with the same phase, there is constructive interference between the vibrations and energy is transferred rapidly

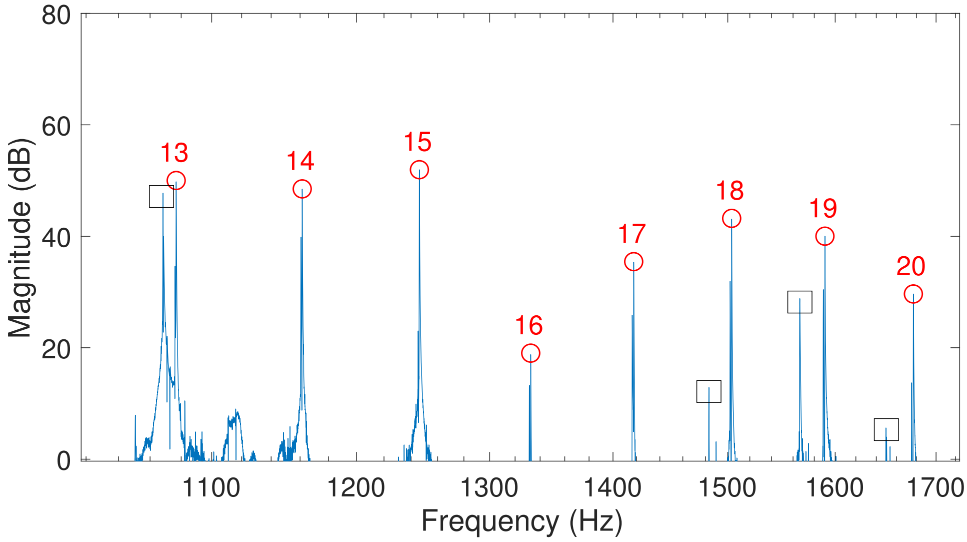


Figure 10: Part of the spectrum of E_2 showing inharmonic partials (∘) and phantom partials (□).

from the strings to the soundboard, making the amplitude of the radiated sound higher and the decay time quicker. If two strings have opposite phases, there is destructive interference between the strings. This means that less energy is transferred from the strings to the soundboard making the amplitude of radiated sound lower and the decay time longer. When two strings have slightly different frequencies, they have the same phase at first, so the amplitude of radiated sound is higher and the decay time is faster ("fast" decay phase). As they get out of phase, less energy is transferred to the soundboard, the amplitude of the radiated sound is lower and the decay time is longer ("aftersound") [9]. This means that by controlling the difference between string frequencies in an unison, the relation between "fast" decay and "aftersound" can be controlled.

Figure 11 shows the decay times of key C_4 with different tunings of the three strings of the unison. It can be seen that if the unison is tuned to the same frequency, the tone will decay quickly with no aftersound. It can also be seen, that if the strings are 2 cents out of tune, there will be multiple fast decay and aftersound phases, as the strings get in an out of tune with each other. With the 1.5 cents sharp tuning, the decay curve seems to be most consistent.

These results match quite well with the results of a study done by Roger E. Kirk [10] where unison groups of three strings where tuned with various differences in fundamental frequency. The participants were asked to tell which tuning sounds the best. The mean maximum difference between the string fundamental frequencies for expert listeners was 1.7 cents. The result was also consistent with the results obtained by measuring a concert grand pianos right after tuning, with the mean maximum difference between fundamental frequencies being in that case 1.6 cents.

It is also possible to witness the double decay phenomenon with a single string [12]. This is because the double decay effect has other contributing factors as well.

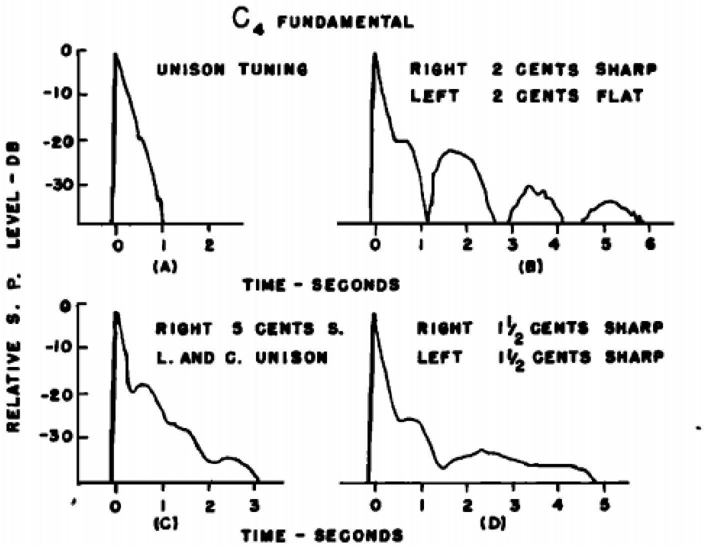


Figure 11: Decay times of three string unison with different tunings. The figure has been adopted from [11].

One of those factors is the double polarization of piano strings. Double polarization means that piano strings vibrate in vertical and horizontal directions due to the nonlinearity of the string [13]. Because the soundboard of a piano is much stiffer in horizontal direction, horizontal vibrations move the soundboard less, which causes radiated sound to have lower amplitude, but longer decay time.

3 Finding Fundamental Frequency

In this section, inharmonicity coefficient estimation algorithms are reviewed. These algorithms are designed to estimate the inharmonicity coefficient of a tone, and they produce an estimation for the fundamental frequency as a side product. These algorithms look at the spectrum of a tone to find its partial frequencies and use these values to estimate the B and f_0 based on the equation:

$$f_k = kf_0\sqrt{1 + Bk^2}. \tag{5}$$

There have been a several algorithms tackling the problem of automatic estimation of f_0 and B values [14, 15, 16, 17, 18]. In 1994, Galembo and Askenfelt found out that a rough estimation of the B of piano tones can be achieved with cepstrum analysis and harmonic product spectrum analysis [14]. Later, they attained a better estimate by using multiple band pass filters, with the centre frequencies situated at the frequencies calculated with Equation 3 [15]. The method is accurate, but suffers from long runtime [16] compared to more recent methods.

Partial frequencies deviation (PFD) [16] and Median-Adjustive Trajectories (MAT) [17] algorithms are more recently developed algorithms. They both are fast as well as accurate, and look like good candidates for the piano tuning system to use as its fundamental frequency estimator. The two algorithms tackle the challenge of finding the inharmonicity coefficient (and f_0) from two, completely different angles. For this reason, both algorithms are reviewed here and later, in Section 6 they are compared to find the better of the two.

Both of these algorithms are based on finding spectral peaks, so a short review of discrete Fourier transform (DFT), which is used to find the spectral components of the signal, is included. After this, PDF and MAT are reviewed.

3.1 Calculating the spectrum of a signal

According to Fourier's theorem, all sounds can be modelled as a sum of fixed sinusoids [19]. The Fourier transform is used to calculate the amplitude, frequency and phase of these sinusoids from the time-domain representation of a signal. For a discrete signal, such as a sequence of time-domain sampled values $x(n)$, the Discrete Fourier Transform (DFT) performs the transform between time and frequency domain. The DFT is expressed as [19, 20]:

$$X(m) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi nm/N}, \tag{6}$$

where $X(m)$ is a complex value, expressing the phase and amplitude of a discrete frequency value, N is the number of samples in the sequence and j is the imaginary unit. From this complex value, the magnitude of the each frequency component can be calculated with equation:

$$X_{mag}(m) = |X(m)| = \sqrt{X_{real}(m)^2 + X_{imag}(m)^2} \tag{7}$$

and the phase of each frequency component can be calculated with equation:

$$X_{phase}(m) = \tan^{-1}\Big(\frac{X_{imag}(m)}{X_{real}(m)}\Big). \tag{8}$$

The magnitude of DFT represents the maximum amplitude of the sine wave. It is often converted to decibel (dB) scale as it better represents the sensitivity of human hearing. This conversion can be made with equation:

$$X_{dB} = 20log_{10}(X_{mag}). \tag{9}$$

A more efficient way of calculating the DFT is using the fast Fourier transform (FFT) algorithm [21]. FFT removes redundant calculations from DFT, and thus makes the calculation much faster [20].

3.1.1 Using the FFT to Calculate Spectrum

The process of finding the amplitude and phase of each frequency component of a piano tone starts by taking a sequence of samples. This sampling is done with a specific sampling frequency f_S, which specifies how often a sample is taken from the continuous signal and a specific bit depth, which specifies the number of different values between 1 and -1 the sample can have depending on its amplitude.

The frequency range of human hearing is between 20 and 20000 Hz [22] and so the semi-automatic piano tuning system uses a commonly used sampling rate of 44.1kHz, which captures the full range of human hearing, as the Nyquist theorem states that the sampling rate must be at least twice as much as the bandwidth of the signal [22]. The system uses a bit depth of 16, which provides $2^16 = 65536$ discrete values between 1 and -1.

The number of samples N taken from a piano tone determines the lowest detectable frequency as well as the frequency resolution of DFT, as each frequency bin of DFT $X(m)$ is an integer multiple of f_S/N. This means that if the frequency range between 20Hz-20000Hz is desired to be examined, the frequency resolution of DFT must be at least 20, meaning $f_s/N = 20$. The frequency resolution of each sine wave, as in how well the frequency of a sine wave can be approximated from the DFT as well as how well can two sine waves be distinguished from one and other is a bit more complicated and will be discussed next.

The magnitude response (amplitude as a function of frequency) of a real sinusoid, time limited by a rectangular window N, can be approximated by the sinc function [20]:

$$|X(m)| = |\frac{A_0 N}{2}\frac{\sin[\pi(f-m)]}{\pi(f-m)}| \tag{10}$$

where A_0 is the amplitude of the sinusoid, f is its frequency and m is the bin index of the DFT. The centre of the main lobe of the sinc function is located at frequency f and each zero crossing of the function is located at a distance of f_S/N from that point. The magnitude response of a sinc function can be seen in Figure 12. It can be seen that the first zero crossings have distance of f_s/N from the centre of the main lobe and the following ones have a distance of f_s/N from each other.

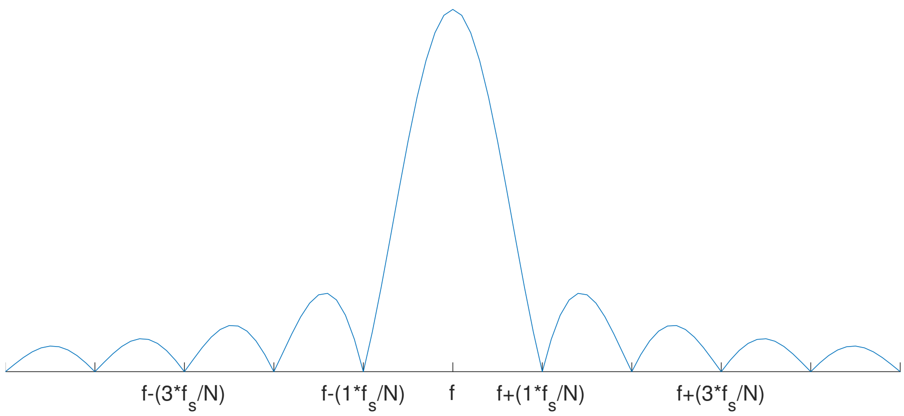


Figure 12: Magnitude response of a rectangularly windowed sinusoid

This means that if the frequency of the sinusoid in the input of the DFT is a multiple of the frequency resolution of DFT ($f = k\frac{f_s}{N}$, where k is an integer number), a single peak can be seen in the amplitude response of the signal as the centre of the sinc function will be located at frequency bin $k\frac{f_s}{N}$ and the zero crossings of the sinc functions will be located at other frequency bins. However, if f is not a integer multiple of the frequency resolution, the frequency bin closest to the main lobe is not in its centre and the magnitude of the side lobes of the sinc function show up in all the other frequency bins of the DFT in an phenomenon called leakage.

To minimize leakage, windowing functions can be used to reduce the magnitude of the side lobes. However, this is done with a cost in frequency resolution, as using a windowing function increases the size of the main lobe. The width of the main lobe and magnitude of the side lobes depend on the windowing function. Figure 13 shows the magnitude response of a sine wave with a frequency of 200 Hz using three different windowing functions. The rectangular window corresponds to using no window. It has a narrow main lobe and high magnitude side lobes. The Blackman window function has three times as wide of a main lobe and much lower amplitude side lobes than the rectangular window. The hamming window is between these two windows with a main lobe two times as wide as the rectangular window and side lobes with magnitude peaks between those of rectangular window and Blackman window.

Window functions can be applied to a signal by multiplying the input sequence of a DFT $x(n)$ with a sequence of window coefficients $w(n)$ calculated with a window function:

$$X_w(m) = \sum_{n=0}^{N-1} w(n)x(n)e^{-j2\pi nm/N} \tag{11}$$

The width of the main lobe of a windowing function can be reduced by using

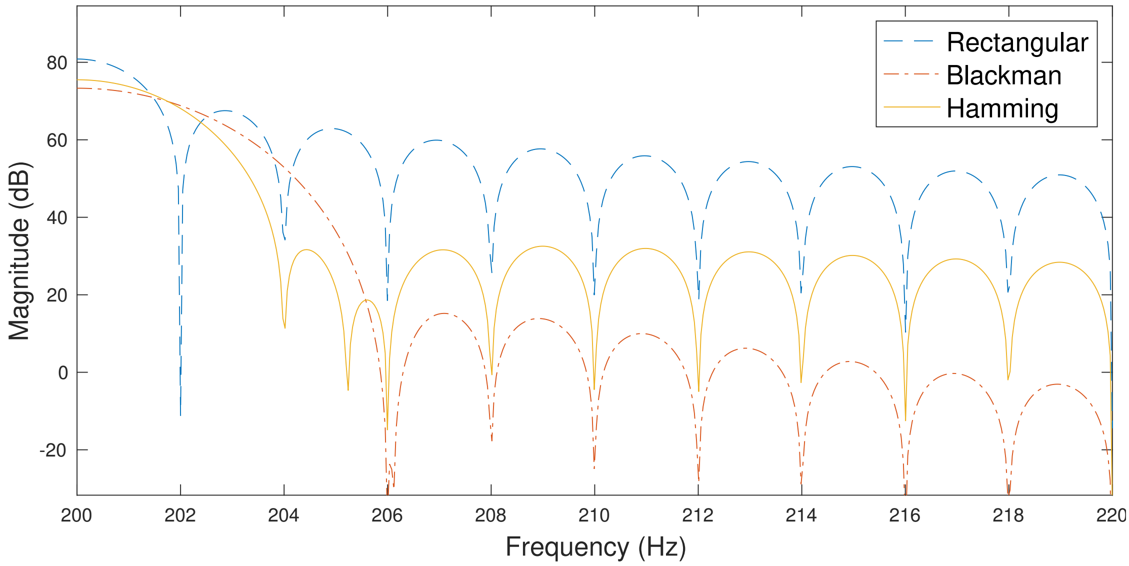


Figure 13: Magnitude response of sine wave with a frequency of 200 Hz using a rectangular window, Blackman window and a Hamming window

more samples to compute the DFT. This means that the size of the window, used for FFT should depend of the frequency of each piano tone. Lower tones tend to have phantom partials which are only a few Hertz from the real ones, and to distinguish these from each other, it is good to use a high frequency resolution. For example, if 44100 samples are used, a frequency resolution of 1 Hz can be obtained. For tones in the higher register, windows that are as long are not needed as phantom partials usually deviate from the real ones by a larger margin. This is good, because the decay times of higher notes are shorter, and the tones only last few tenths of a second.

To find the centre of the main lobe, zero padding the input signal can be used to get more points from the magnitude response of the windowing function. This can be seen in Figure 14 which includes 22050 point (0.5 seconds at 44.1 kHz) FFT of a sine wave with a frequency of 41 Hz without zero padding as well as with zero padding to make the length of the DFT reach 2^{20}, which corresponds to approximately 24 seconds at the sampling rate of 44.1 kHz.

Another, more efficient way to find the peak of the main lobe is to use parabolic interpolation [19]. Parabolic interpolation is used after FFT is performed and the method fits a parabola to three point to estimate the local maxima. It is usually best to use both zero padding and parabolic interpolation, first using zero padding within the time frame of the application as zero padding increases the process time of FFT and in real-time applications algorithms must be performed within a certain time. After that parabolic interpolation can be used to enhance the accuracy of the estimation.

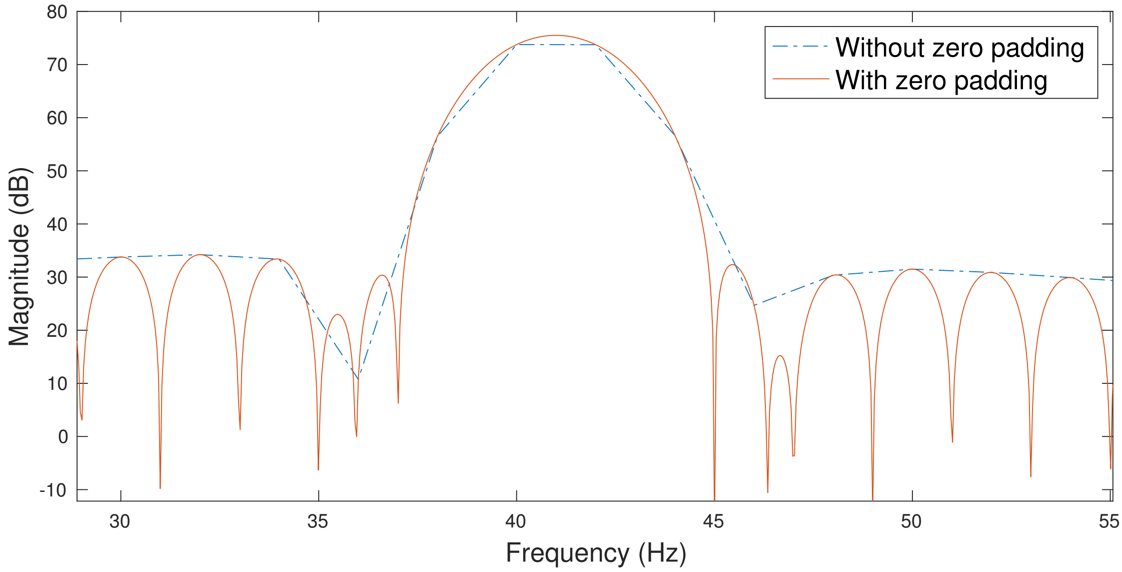


Figure 14: Effect of zero padding.

3.2 Partial Frequency Deviation

The Partial Frequency Deviation (PFD) algorithm [16] is based on minimizing the difference between estimated partial frequency values $\hat{f}_k$ and values of frequency peaks found in the spectrum of the signal which belong to those partials f_k. The algorithm iteratively enhances the accuracy of its estimations by changing the values of B and f_0 based on the trend of deviation between $\hat{f}_k$ and f_k.

The algorithm needs a sequence of samples from the estimated tone. The length of this sequence determines the frequency resolution of the estimation. A rough estimate of the first partial $\hat{f}_0$ as well as the inharmonicity coefficient $\hat{B}$ are needed as well. The estimation $\hat{B}$ does not need to be very accurate as small deviations in the estimate do not affect the final result very much [16]. The estimate for f_0 can be made by finding the first partial from the spectrum of the tone and calculating the fundamental frequency with that and the estimate for B from Equation 5.

In the first stage of the algorithm, the spectrum of the signal is obtained with the FFT algorithm. This is done with a 2^{16}-point FFT with zero padding and a Blackman window [16]. Next, the number of spectral peaks is reduced by dividing the spectrum into sub-bands, each with a width of $5f_1$, and selecting 10 of the highest spectral peaks in each band. This reduces the number of spectral peaks to $2K_{max}$, where K_{max} is the maximum number peaks used for the estimation. After this, an iteration loop starts refining the estimations of B and f_0. The iteration loop is first used to estimate the value of B and this estimation is then used to estimate the value of f_0.

The iteration loop includes three stages:

1. The amount of deviation between estimated values of partial frequencies $\hat{f}_k$ and the found spectral peaks is calculated. This is done by calculating estimates for partial frequencies $\hat{f}_k$, by using $\hat{f}_0$ and $\hat{B}$ in Equation 3. FFT used to compute

the spectra has limited frequency resolution depending on the number of points used to calculate it, so to avoid bias due to this, values of $\hat{f}_k$ are quantized using the resolution of the FFT. After this, spectral peaks are searched for in a closed interval $[\hat{f}_k - \Delta f, \hat{f}_k + \Delta f]$ near the estimated values and the frequency deviation between peak frequencies f_k and estimations $\hat{f}_k$ is calculated $D_k = \hat{f}_k - f_k$.

2. Trend of deviation is determined by comparing subsequent partials. This is done by taking their derivative. If there are more positive derivatives, the trend is positive and vice versa.

3. New estimation of B or f_0 is done depending on if it is used to make a B estimation or f_0 estimation.

   - B estimation is done by multiplying the previous estimate with 10^{δ}, where δ is an adaptive step size which depends on the trend of deviation. The value of δ is set to 1 at first. If the trend of deviation is positive the sign of δ is positive and if it is negative, the sign of δ is negative. If the trend goes from positive to negative or vice versa, δ is divided by two.
   - f_0 estimation is done by modifying the value of f_0 according to equation: $f_0 = f_0(1 \pm \mu)$, where μ is an adaptive step size. The value of μ is initialized to 0.005 and if the trend of deviation changes its sign, the value is divided by 2.

When the initial B estimation is done, the loop is terminated after 40 iterations or if $|\delta| < 10^{-4}$. The iteration loop of f_0 estimation is terminated after 100 iterations or if $|\mu| < 10^{-5}$. Figure 15 shows a block diagram of the process.

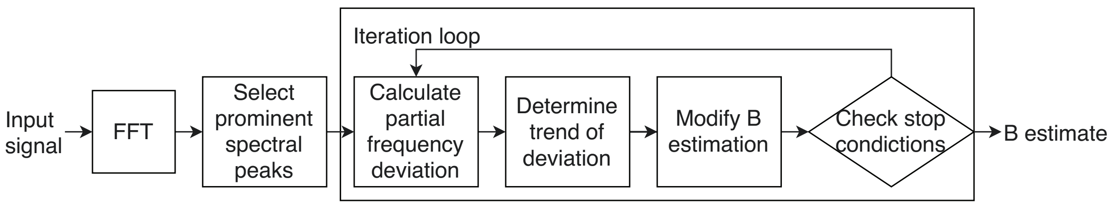


Figure 15: Block diagram of PFD process, adopted from [16].

The smoothness of the D_k curve provides an indication of the accuracy of the algorithm. Figure 16 shows the D_k curve of A_2 after 1 (a), 2 (b), 5 (c) and 10 (d) iterations. It can be seen how at the start, as the estimates are not accurate, the differences between the estimated values and found peaks are great, but as new estimations are made, the deviation curve gets smoother and the differences between estimated values and found peaks get smaller.

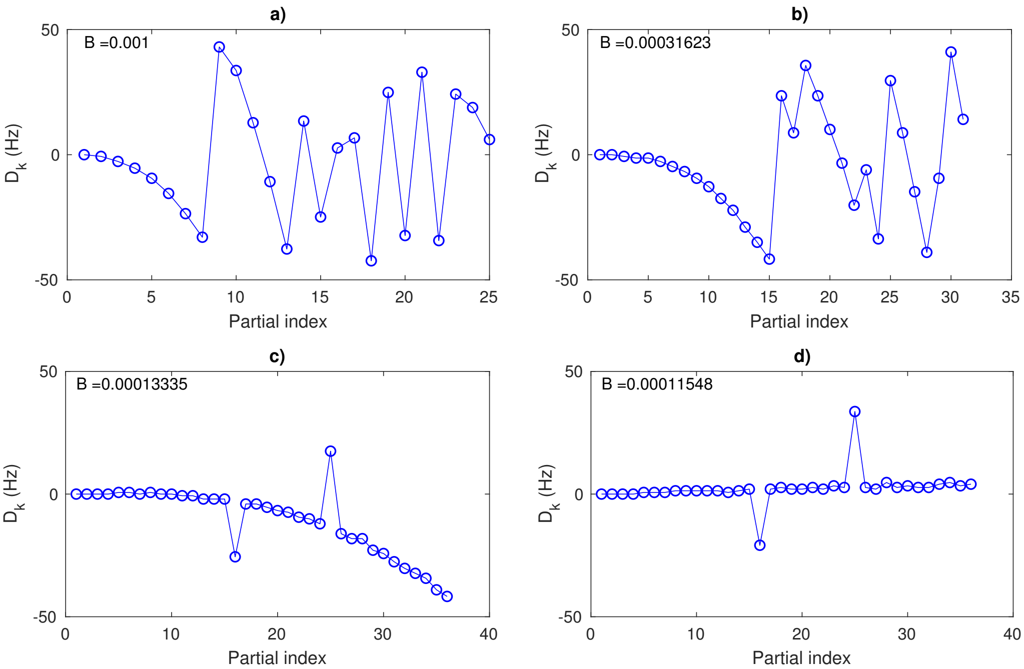


Figure 16: Block diagram of PFD process, adapted from [16].

3.3 Median-Adjustive Trajectories

The Median-Adjustive Trajectories algorithm calculates estimations for B and f_0 based on the frequencies of known partials, and finds new partials based on estimations made with these estimated values of B and f_0. The algorithm is based on the idea that if the frequencies of two partials are known, the value of B can be calculated purely based on them.

If the Equation 3 is solved in terms of fundamental frequency with partial number m:

$$f_0 = \frac{f_m}{k\sqrt{1+Bm^2}}, \tag{12}$$

and then 12 is substituted into 3:

$$f_k = kf_0\sqrt{1+Bk^2} = k\frac{f_m}{k\sqrt{1+Bm^2}}\sqrt{1+Bk^2} \tag{13}$$

The value of B can be calculated with the partial frequencies f_k and f_m:

$$B = \frac{(f_k\frac{m}{k})^2 - f_m^2}{k^2 f_m^2 - m^2(f_k\frac{m}{k})^2}. \tag{14}$$

This means that if the frequencies of first two partials, f_1 and f_2 are found from the spectrum of the tone, an estimation for the value of B can be made. This estimation, along with the frequencies of these partials can be then used in Equation

12 to make estimations for the value of f_0. With these estimations for B and f_0, the estimation for a new partial can then be made.

A block diagram of the process can be seen in Figure 17. The algorithm requires an estimation for the values of f_0 and B as it needs to find the first two partials before making its first estimation. First estimates $f_1 = 1f_0\sqrt{1 + B1^2}$ and $f_2 = 2f_0\sqrt{1 + B2^2}$ are used to find spectral peaks within a range $[\hat{f}_k - \Delta f, \hat{f}_k + \Delta]$. The original algorithm used estimates $f_1 = 1f_0$ and $f_2 = 2f_0$, but the new estimates were found to provide better accuracy.

After finding the spectral peaks belonging to the first two partials a new estimate of B is calculated with Equation 14. This estimation is stored into an array of B estimations and a median of the array is used to calculate estimations (one per partial) of f_0 with Equation 12. The f_0 estimations are stored in an array and the median of this array as well as the median of the B array is used to estimate location of the next partial f_3 with Equation 3. If the amplitude of the partial exceeds a set threshold, the process of calculating new B and f_0 estimates and finding a new frequency peak belonging to the next partial is repeated.

The number of B estimations produced by the algorithm can be calculated from equation:

$$E = \frac{K^2 - K}{2} \tag{15}$$

where K is the number of measured partials and the number of estimated f_0 values is K. After the algorithm has terminated due to a frequency peak lower than the set threshold, a median value of the estimations is taken as the output.

Figure 18 shows the spectrum and estimated values of f_k (dashed lines) before the algorithm starts searching for each peak. In Figure 18a initial estimates of $f_0 = 110$ Hz and $B = 10^{-3}$ are used to find the first two peaks (circles), in 18b these first two peak have been used to estimate new f_k values and the third peak is looked for with this estimate. Figures 18c and d show the algorithm looking for the fourth and fifth peak. The arrows point to the spectral peaks that the f_k estimations are trying to find. It can be seen how after the first estimations, the adjacent spectral peaks could be confused based on the estimations, but as higher peaks are used to estimate the f_0 and B values, the estimations for the higher partials get much better.

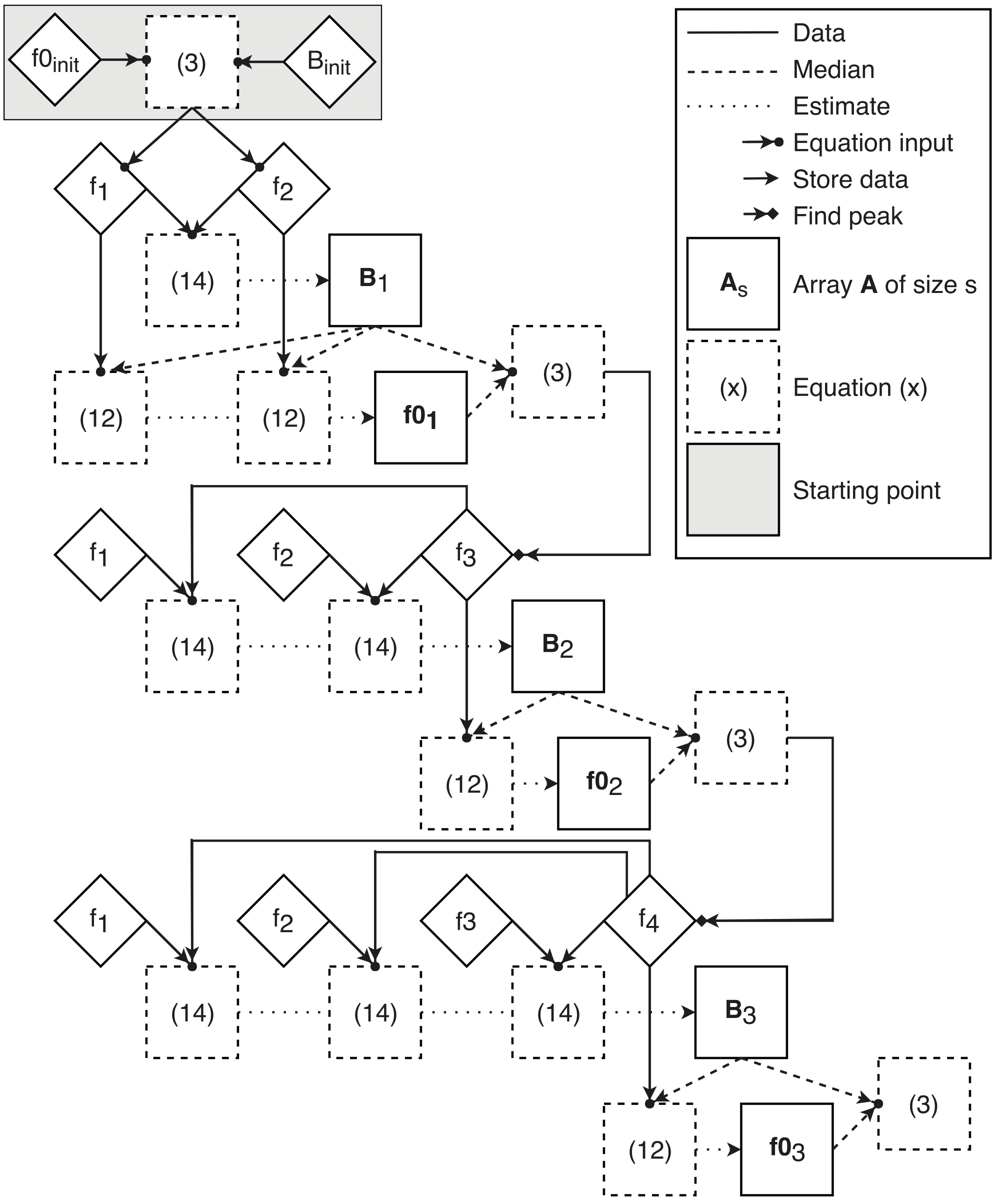


Figure 17: Block diagram of MAT algorithm, adapted from [17].

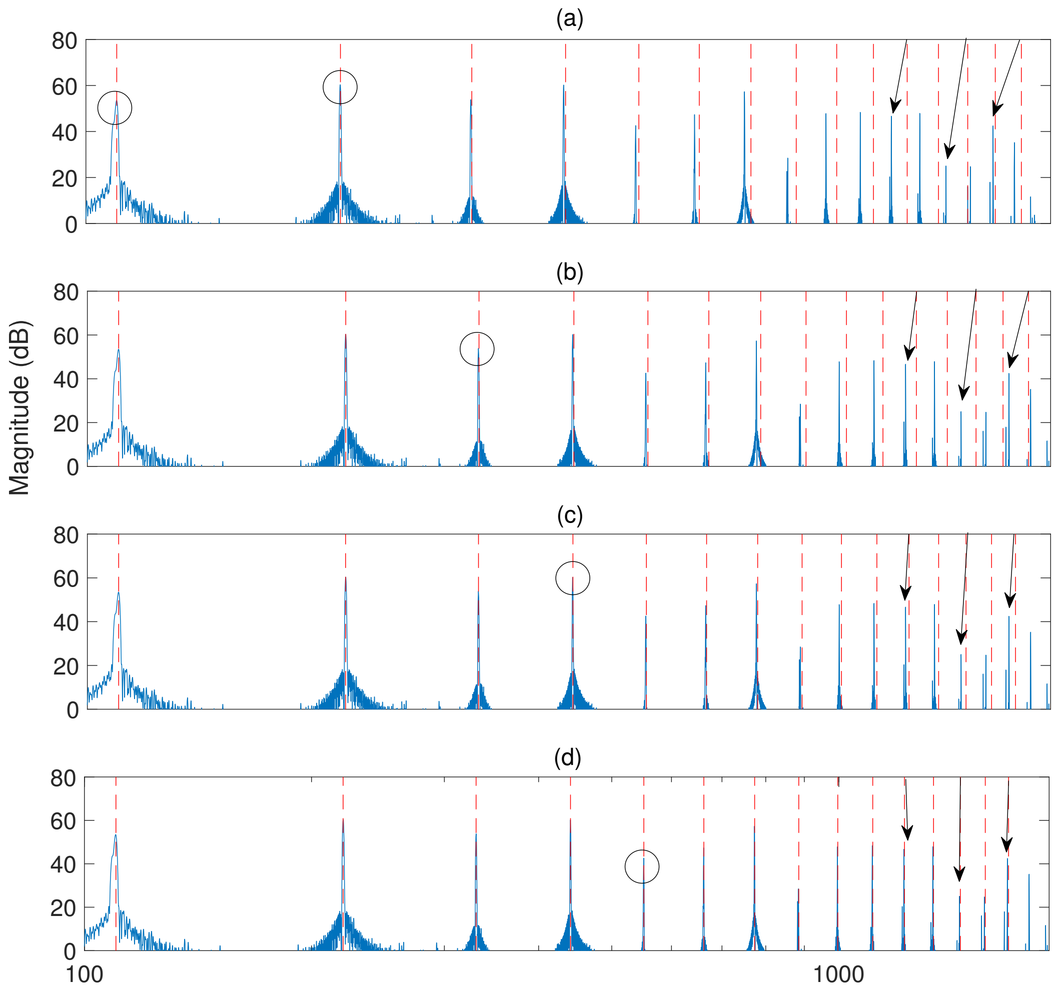


Figure 18: Frequency spectrum of A_2 during the process of the MAT algorithm. The circles show the spectral peaks found at that specific point of the algorithm, the dashed lines show estimated partials from the first to the fifteenth and the arrows show, which peaks are looked for with the estimations for the *11th*, *13th 15th* partial.

4 Piano Tuning Process

In this section, the tuning process of a piano is reviewed. Tuning a piano is generally done by a professional tuner with the help of a tuning fork, which is used to tune a reference note, a tuning lever, which is used to change the tension of the strings and some sort of material to dampen all except one string in each string unison.

The tuning is done aurally by counting the period time of beats, which are amplitude modulations formed by two frequencies close to each other. Beats can be used to tune the instrument as the scale of the piano specifies that tones produced by keys, which have a certain distance, or interval, from each other should have a specific beating rate. However, as the tone of a piano consists of multiple partials, these beats are formed between multiple sets of frequencies. Tuners tune the strings in a way, where none of these beats deviate too much from the beating rates defined by the scale.

This section includes a short literature review of the beats as well as loudness and masking, which are psychoacoustic phenomena affecting how well a beat is heard. After that the scale of the piano is discussed, and how beats are used to tune the instrument. After learning all needed concepts, the process of aural tuning and existing mathematical simulations of the tuning process are reviewed.

4.1 Beating

When two pure tones have frequencies close to each other, but not quite the same, their superposition will have a periodic change in its amplitude. This amplitude modulation is called beating, and the frequency of these modulations can be calculated from the equation [23]

$$f_B = f_2 - f_2 = \Delta f, \tag{16}$$

where f_1 and f_2 are the frequencies of the two tones. Figure 19 illustrates this effect. Figure 19a shows a sine wave with a frequency of 110 Hz, Figure 19b shows a sine wave with a frequency of 100 Hz and Figure 19c shows the superposition of the two waves. It can be seen that the period time of the amplitude modulation is 0.1 seconds which translates to a frequency of 10 Hz.

Equation 16 applies only until a certain point. As the two frequencies get further away from each other the frequency of beats get faster at first, until unpleasant roughness between the two frequencies emerges. From this roughness, two distinct tones can be heard after Δf exceeds the limit of frequency discrimination and after Δf surpasses the critical band, the roughness disappears and only two distinct frequencies can be heard [23].

The maximum amplitude of the beating effect is the sum of the two tones:

$$A_{max} = A_1 + A_2, \tag{17}$$

where A_1 is the amplitude of the first tone and A_2 is the amplitude of the other one. It can be seen that the maximum amplitude of the superposition in 19c is 2 as the amplitudes of the two waves are 1 ($A_{max} = 1 + 1 = 2$).

The minimum amplitude of the beating effect is the difference between the two tones:

$$A_{min} = |A_1 - A_2|, \tag{18}$$

and it can be seen that the minimum amplitude of the superposition is zero as there is no difference between the two amplitudes.

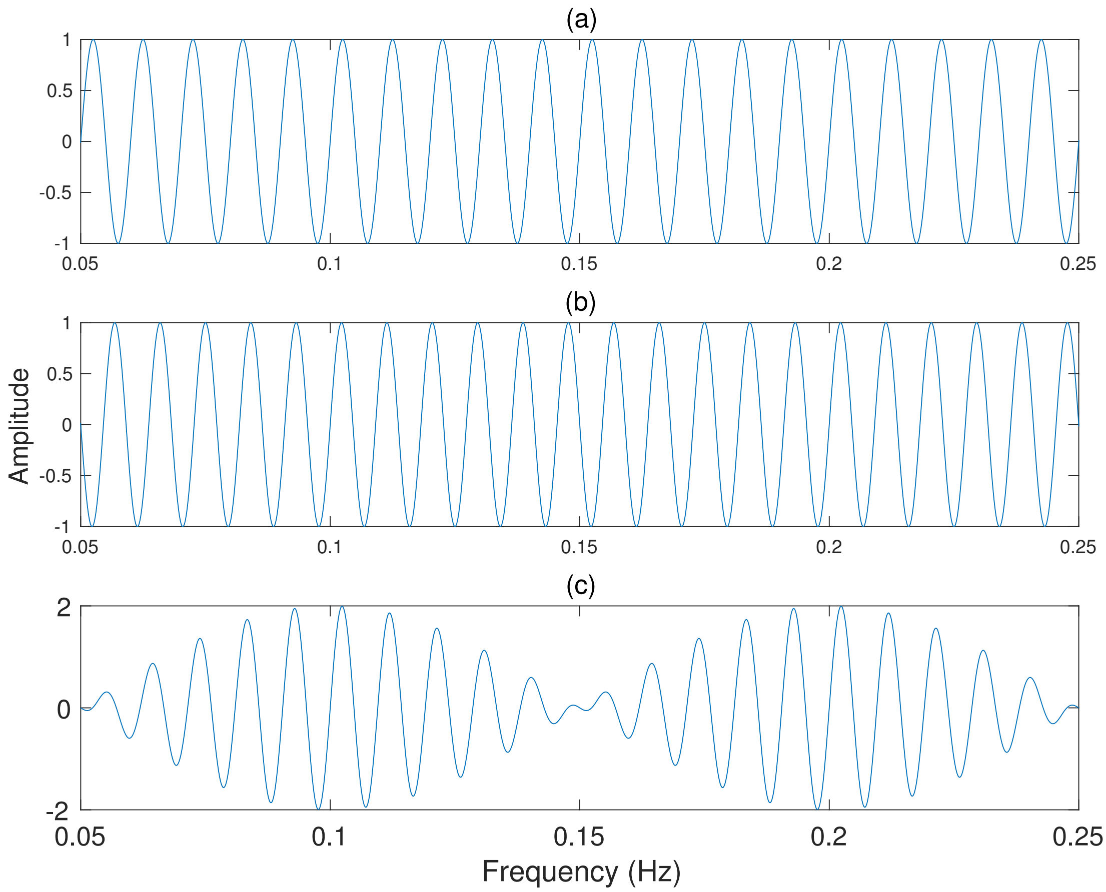


Figure 19: (a) sine wave with frequency of 110 Hz (b) sine wave with frequency 100 Hz (c) superposition of the two waves.

These concepts related to two pure tones can be expanded to complex tones, as all complex tones can be thought as a sum of sine waves. Each of the partials close to each other produce beats with differing levels and frequencies.

4.2 Loudness

To find out how well a listener can hear the beats produced by a set of partials, the perceived level of the sound waves have to be considered. Partials with frequencies between 20Hz and 20000 Hz and sound pressure level (SPL) between the threshold of hearing and the threshold of pain are perceived by a listener [22]. The frequency sensitivity of human hearing in that range can be approximated by using a weighting curve on the sound pressure levels of string modes to find the perceived level of sound or sound level [22]. Figure 20 shows weighting curves A, B, C and D. Most commonly used weighting curve is the A-weighting curve, which is commonly used in noise measurements [22].

When a microphone is used to record piano tones, there is no knowledge about the sound pressure level of the radiated sound. This means that there is no way to tell if a partials has a sound level above the threshold of human hearing. However, the weighting functions are not dependent on the SPL of the radiated sound so relative sound levels can be calculated.

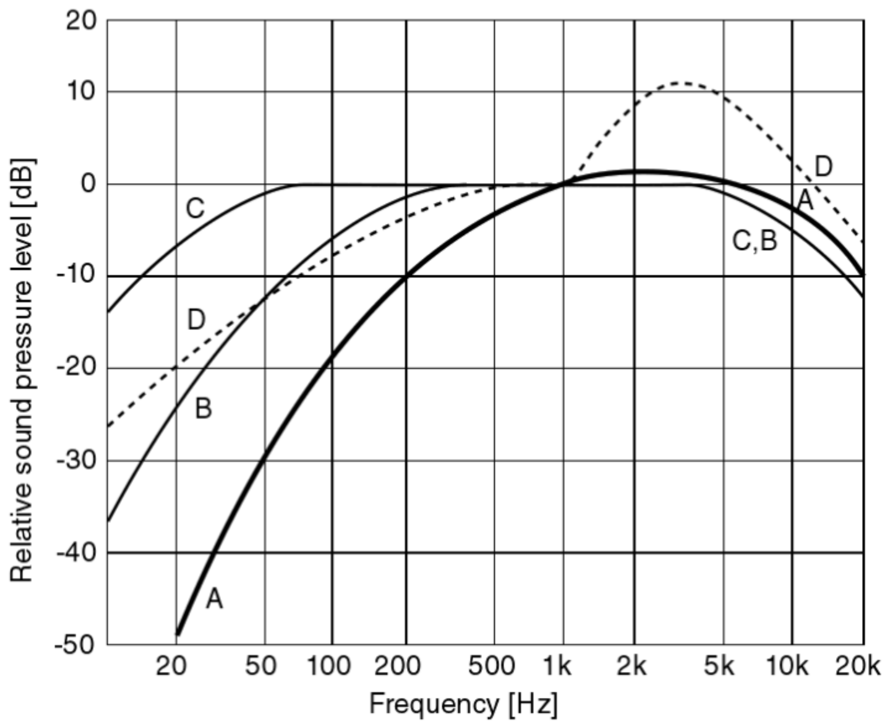


Figure 20: Weighting curves A, B, C and D [22].

4.3 Masking

As the partials of piano tones have varying sound levels, softer sounds may be unheard because of louder ones in a phenomenon called masking. The effect of masking can be estimated by calculating a masking threshold. If a tone is below the masking threshold of another tone, it will not be heard by a listener. Figure 21 shows the masking threshold of a probe tone in the presence of a masker tone. It can be seen that as the partials of the two tones are close to each other, the masking threshold is lower, because the beating effect lowers the threshold [22].

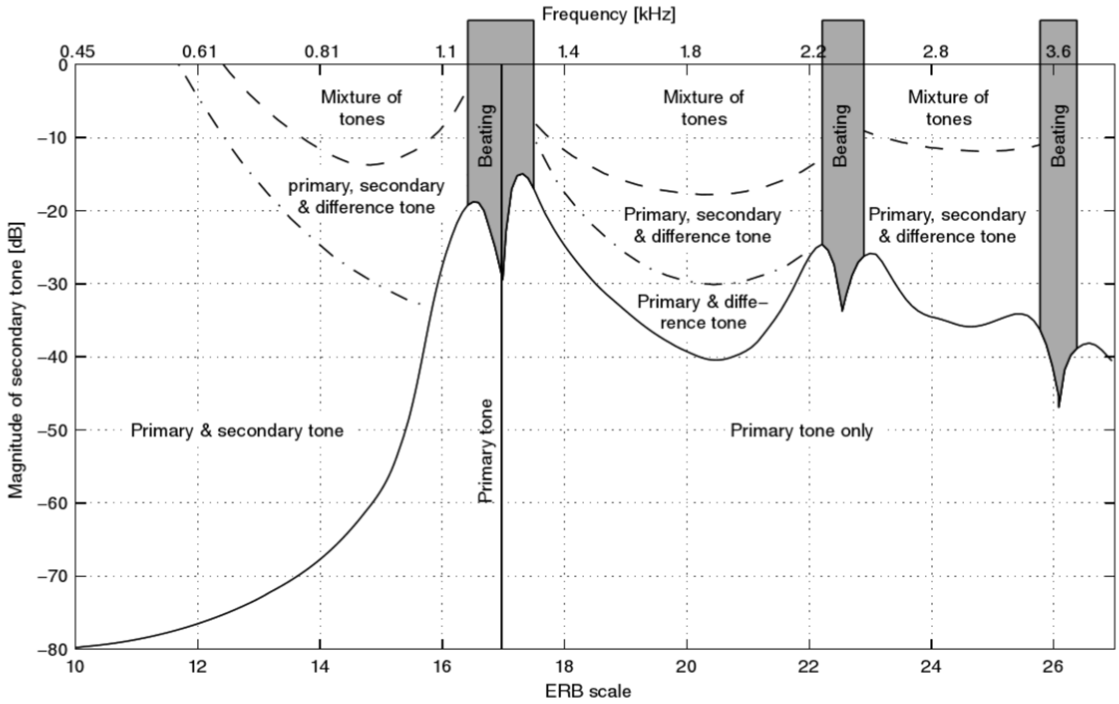


Figure 21: Masking threshold of a probe in the presence of a masker tone [22].

An estimation of the masking effect can be calculated by using a spreading function. A popular spreading function proposed by Schroeder is expressed by equation [24]:

$$10 \log_{10} SF(dz)/dB = 15.81 + 7.5(dz + 0.474) - 17.5\sqrt{1 + (dz + 0.474)^2} \qquad (19)$$

where dz is the difference between the two partials in Bark scale. Conversion between Hz and bark frequency scale can be calculated with the following equation [25]:

$$z(f) = 13 \arctan 0.00076 f + 3.5 \arctan (f/7500)^2 \qquad (20)$$

This spreading function is independent of the maskers SPL and thus can be used with recordings of piano tones. The spreading function in Equation 19 is shifted slightly lower depending on the tonality of the masker. As all the partials of a piano tone are tonal, the following variable is added to Equation 19 [26]:

$$\Delta_{TM}(z_{masker}) = -6.025 - 0.257 z_{masker} \qquad (21)$$

where z_{masker} is the frequency of the masker in bark scale.

Figure 22 shows the spreading functions for each partial of F-sharp in the fourth octave. It can be seen how partials 4, 6, 7, 8 and 9 are masked by the other partials.

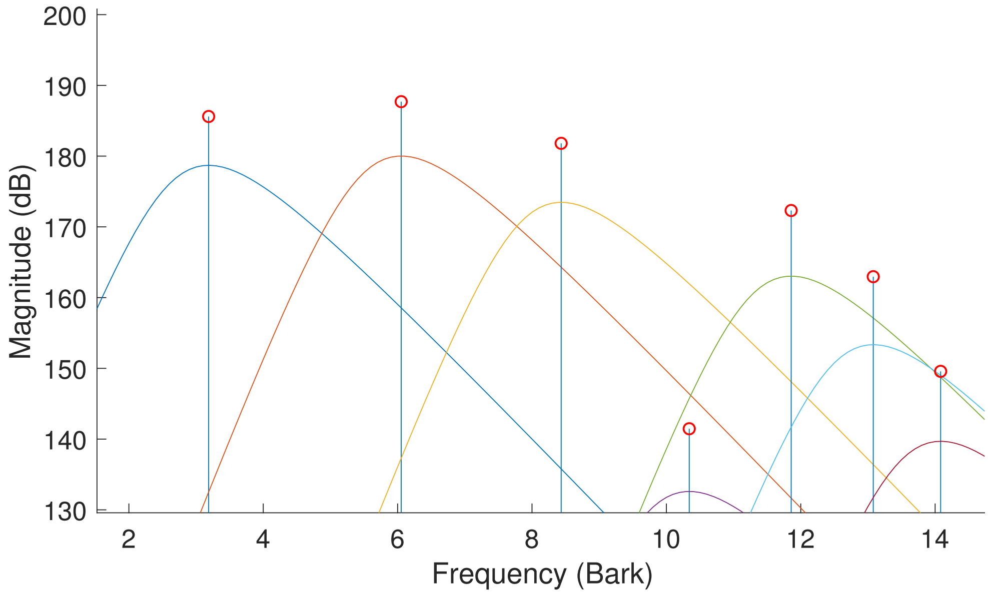


Figure 22: Spreading functions for the partials of a F-sharp in the fourth octave.

4.4 The Scale of the Piano

The tuning of a piano is based on the equal temperament scale, which makes all of the steps in the scale equal. This means that the ratio between the fundamental frequencies of subsequent tones in the scale should be the same [3]. More specifically, the scale is the equal temperament twelve-tone (ET12) scale which makes the ratio between an octave 2:1 and divides each octave in 12 steps.

The frequency ratio of 2:1 means that the fundamental frequencies of two tones constructing an octave should have a ratio of 2:1. As all steps in the scale should be equal, this leads to the 12 steps between each octave have a frequency ratio of $\sqrt[12]{2}$:1, as $12\ \sqrt[12]{2} = 2$.

This kind of spacing of steps in the scale leads to all other intervals deviating from their simple frequency ratios according to Table 1. The deviation in Table 1 is measured in cents, which is a logarithmic unit of measure. The deviation between two frequencies in cent can be calculated with equation:

$$\text{cents} = 1200 \log_2 (b/a), \tag{22}$$

where a positive value indicates that b is greater than a. The values in Table 1 tell that when two keys have a specified distance (interval) between them, and the frequency ratio of the interval is $k : l$, the deviation D between the fundamental frequency of the lower key $f_{0,lower}$ and the higher one $f_{0,higher}$ should follow equation:

$$D = 1200 \log_2 \left(\frac{l f_{0,higher}}{k f_{0,lower}} \right) \tag{23}$$

These deviations lead to specified beating rates between two harmonic tones as the partials of a harmonic tone follow equation:

$$f_k = kf_0, \tag{24}$$

which leads to all partials which are integer multiples of the frequency ratio specified by the interval to have the same distance. However, as piano tones are not harmonic and follow equation:

$$f_k = kf_0\sqrt{1 + Bk^2}, \tag{25}$$

tuning the instrument to fundamental frequencies specified by the ET12 scale does not lead to same beating rates between all partials.

Table 1: Partial closest to each other in an interval.

Interval	Frequency ratio	Deviation (cents)	Distance (semitones)
Octave	2:1	0	12
Perfect fifth	3:2	-1.96	7
Perfect fourth	4:3	+1.96	5
Major sixth	5:3	+15.64	9
Major third	5:4	+13.69	4

To achieve similar beating rates as an instrument with harmonic tones, piano tuners listen to beats and tune the instrument based on them [2]. Figure 23 shows the first two matching partials of the octave between Bb_7 and Bb_8 in a tuning done by a professional tuner. It can be seen how it is not possible to match all the partials, but a compromise between the two sets producing as little beats as possible is made.

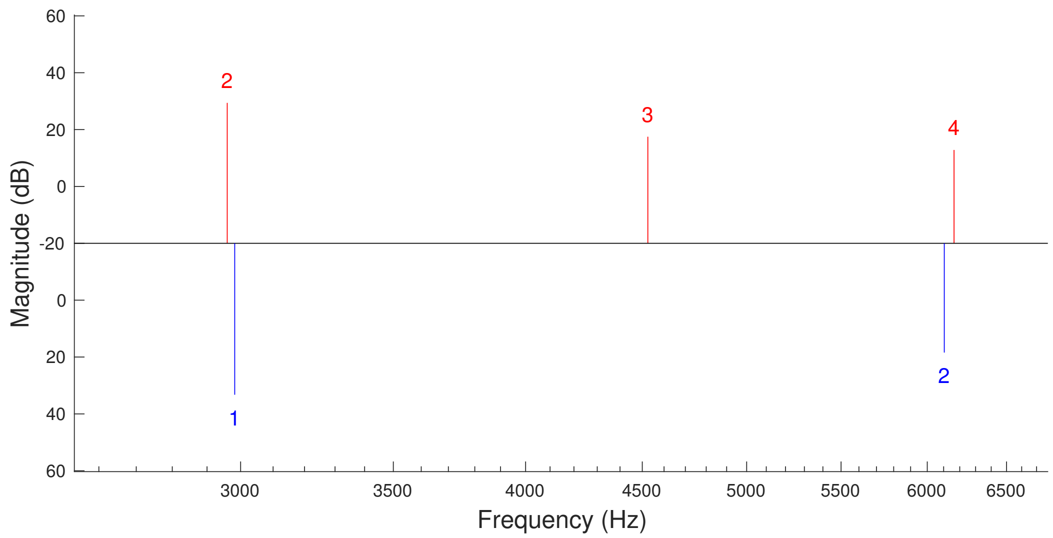


Figure 23: Partial frequencies of A_4 (higher) and A_5 (lower).

The process of aural tuning leads the scale of a piano to be "stretched". This is because higher partials of bass strings, with higher level of inharmonicity, are louder than lower ones, which means that the tuner listens to beats based on those partials, which leads to wider intervals. With treble strings the inharmonicity is very high already with low partials, so counting beats based on those lead to stretched intervals as well. Figure 24 shows tuning done by a professional tuner. The figure shows how much the tuning deviates from the ET12 scale (in cents). It can be seen that in the middle of the scale, the tuning is very close to ET12. This is because the tuner starts the tuning process there. In the bass end and in the treble end there is a considerable amount of deviation from the ET12 scale.

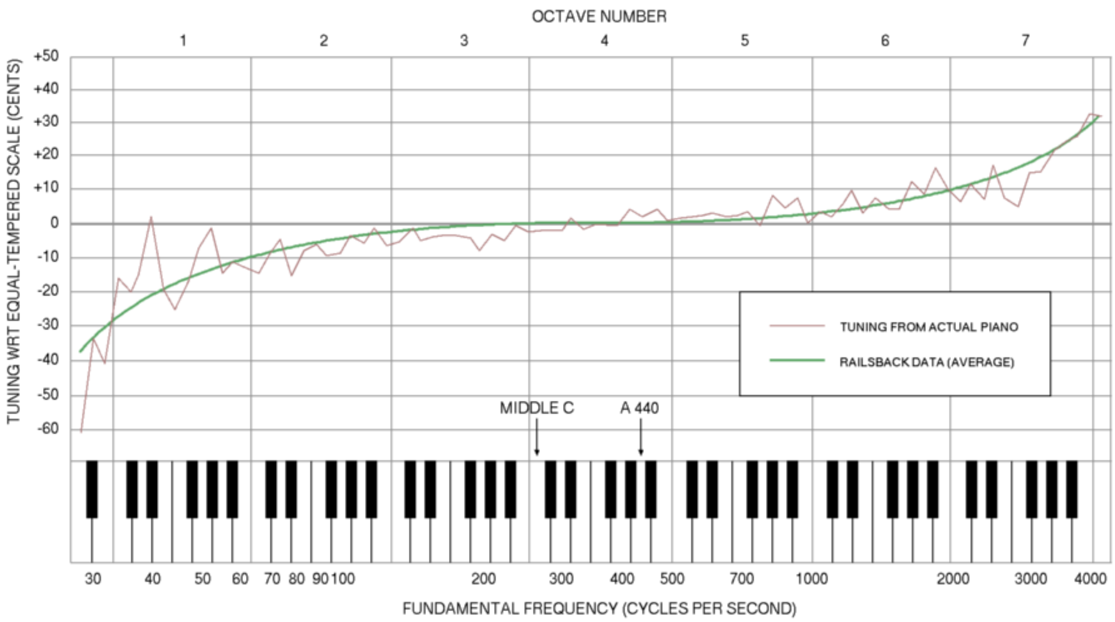


Figure 24: Tuning done by a professional tuner (deviation from theoretical ET12 scale). Data from [1], figure adopted from [27]

4.5 Aural Tuning

In the process of aural tuning, the deviations between the partials of an interval are used to tune the instrument. The order in which the keys of the instrument are tuned is usually quite similar between tuners, and follow these steps [2]:

1. Tuning the reference note: The first partial of a reference note is tuned to the same frequency as a tuning fork by minimizing beats between the two tones. The reference note is usually A_4 and the frequency of the principal mode (mode that is heard) of the tuning fork is usually 440 Hz.

2. Tuning the reference octave: A reference octave is tuned with the help of the reference note. The octave is usually from F_3 to F_4 and is tuned close to the

ET12. There are multiple ways to do this, depending on the tuning system, but the principle is usually the same.

Intervals are tuned according to the theoretical values between the partials of two tones by counting beats. For example if A_4 is tuned as a reference note, according to ET12 the partials of A_4 and A_3 should not have a beating effect, so A_3 is tuned by minimizing the beating effect between the two tones. After this, F_3 could be tuned by counting beats between F_3 and A_3. If the theoretical partial frequencies of both tones (according to ET12) are calculated, as is done in Table 2, it can be seen that the fifth partial of F_3 is close to the third partial of A_3 and the beating effect between them should have a frequency of $873.1Hz - 880Hz = -6.9Hz$. This means that there should be 6.9 periods of the beating effect (amplitude modulation) in one second. The rest of the reference octave would be then tuned by doing similar calculations and counting beats.

Table 2: Theoretical partial frequencies of F_3 and A_3.

Partial	F_3	A_3
f_1	174.6	220.0
f_2	349.2	440.0
f_3	523.8	880.0
f_4	698.5	1760.0
f_5	873.1	3520.0

3. Tuning the high middle and high treble: All the notes above the reference octave are tuned by controlling the beating effect between two tones. Notes are first tuned by minimizing beats between octaves and then fine tuning is done by listening that octaves, double octaves, tenths and seventeens all sound good and there is no unpleasant roughness. The intervals used to tune a piano may differ with varying tuning systems, but the principle is the same.

4. Tuning low middle and bass: The strings below the reference octave are tuned in a similar way to the ones above it. The only difference is that the upper partials of bass tones can be heard much more clearly as there are more of them in the audible range and their amplitudes are greater. Also larger intervals are usually used in the low end, as the beats between the partials of smaller intervals start to get harder to distinguish. Intervals of fifth, octave and thirteen are usually used in the higher end of the bass and seventeenth, double octave, double octave and a third as well as double octave and a fifth, triple octave and triple octave and a third may be used in the very bass end.

4.6 Mathematical Simulation of Tuning Process

There have been a few attempts of tuning the piano with a mathematical model. Lattard [28] proposed a method of using the first five partials of each tone to calculate best possible beating rates for several intervals. The method was used to tune the reference octave from F_3 to F_4 by first recording each note and finding the values of f_0 and B for each tone. After this, the values of f_0 were adjusted until each interval had beating rates that were close enough to the ones defined by ET12.

Another method proposed by Hinrichsen [29] minimizes the Shannon entropy between each tone in the scale. The entropy between two frequencies decreases as they overlap and is in minimum when they have the same frequency. The method gives a fairly good result but does not fit the requirements of the tuning system as each key of the piano has to be pre-recorded, which means that the tuning process would take more time then that of a professional tuner.

5 Novel Tuning Process

In this section, a novel tuning process for the grand piano is proposed. The main part of the process is tuning one string of each string unison to a target frequency, and the other strings in the unison will then be tuned to maximize the decay time by utilizing the double decay effect.

Similar to the process of aural tuning, the strings are tuned one by one, starting with a reference string (A_4) and continuing with the reference octave ($F_3 - F_4$), high middle, high treble, low middle and the bass. The fundamental frequency of each string is changed by changing its tension, as changing the tension of a string changes the fundamental frequency of a string according to equation

$$f_0 = \frac{\sqrt{\frac{T}{m/L}}}{2L} \tag{26}$$

where T is the tension of the string, m is the mass of the string and L is the length of the string. The fundamental frequency and inharmonicity coefficient of each string is stored after tuning and used to tune the following strings by calculating a target f_0 based on the beating rates of several intervals. As the first string (reference string) does not have any previously tuned strings to compare against, a reference frequency (close to 440 Hz) is used to tune it.

Frequency estimation algorithms need rough estimates for the values of f_0 and B in order to find better ones. These estimates are provided by the tuning process as it is always known which key is tuned and this provides the theoretical f_0 for each key. The estimate for the value of B is gotten from keys near the currently tuned one, which have already been tuned (and B has been estimated).

5.1 Finding the Target Fundamental Frequency

The target fundamental frequency for each string is found by optimizing the beating rates between the string that is currently tuned and all the strings that have been already tuned and are certain interval away from that string. The ET12 scale specifies beating rates for each interval and those beating rates stay constant when they are measured in cents. The beating rate in cents between two frequencies can be calculated with equation:

$$1200 \log_2 \left(\frac{f_{l,n+m}}{f_{k,n}} \right) = c \tag{27}$$

where c is the difference in cents, $f_{k,n}$ is the kth partial of the nth note and $f_{k,n+m}$ is the lth partial of the note which has a m semitone difference (interval) from n. Table 3 shows the values of c, k, l and m for intervals usually used for tuning a piano.

As the partials of piano tones can be calculated with equation:

$$f_k = k f_0 \sqrt{1 + Bk^2} \tag{28}$$

Table 3: Values of c, k, l and m for several intervals

Interval	**c**	**k**	**l**	**m**
Octave (up)	0	2	1	+12
Perfect fifth (up)	-1.96	3	2	+7
Perfect fourth (up)	+1.96	4	3	+5
Major sixth (up)	+15.64	5	3	+9
Major third (up)	+13.69	5	4	+4
Octave (down)	0	1	2	-12
Perfect fifth (down)	+1.96	2	3	-7
Perfect fourth (down)	-1.96	3	4	-5
Major sixth (down)	-15.64	3	5	-9
Major third (down)	-13.69	4	5	-4

Equation 27 can be written as:

$$1200 \log_2 \frac{l f_{0,n+m}\sqrt{1 + B_{n+m} l^2}}{k f_{0,n}\sqrt{1 + B_n k^2}} - c = 0. \tag{29}$$

An estimation for the value of $f_{0,n}$ can be made by solving it from Equation 29 with the assumption that the inharmonicity coefficient of the string does not change during the tuning. Other coefficients in the equation are known as $n + m$ is the index of a previously tuned string with known values of f_0 and B.

The assumption that B does not change during the tuning is fairly accurate, as the change in tension changes f_0 much more than B. This can be seen by using the specifications of a real piano string to measure the accuracy of this assumption. Table shows the specifications for one of the strings of key D_3 of the mini-sized Estonian-Minion piano with values of f_0 and B calculated with Equations 26 and 4.

If the tension of the string is lowered to 558 N, the f_0 of the string drops down approximately one semitone and B increases to $B_{new} = 1.5 \times 10^{-4}$. The effect of

Table 4: Specifications of the mini-sized Estonian-Minion piano [30]

Note	D_3
T (N)	626
L (mm)	964.1
r (mm)	0.5625
m/L (g/m)	7.8
E (MPa)	2100000
f_0 (Hz)	146.92
B	1.3×10^{-4}

Table 5: Difference between the first five partial frequencies calculated with differing levels of B (one semitone apart).

Partial	**Difference in cents**
1	0.0142
2	0.0569
3	0.1280
4	0.2273
5	0.3548

this increase in B can be examined by calculating the partials of the string with Equation 3 by using the value of f_0 in Table 4 and values of B before (from Table 4) and after (B_{new}) the change in tension. Table 5 shows the difference between the first five partial frequencies calculated with the two values of B. The increase of B rises the partials well under half a cent, so the approximation is fairly accurate with a change like this.

Solving $f_{0,n}$ from Equation 29 results in:

$$f_{0,n} = \frac{A}{C} \tag{30}$$

where

$$A = \frac{l f_{0,n+m}\sqrt{1 + B_{n+m}l^2}}{k\sqrt{1 + B_n k^2}} \tag{31}$$

and

$$C = 2^{c/1200}. \tag{32}$$

As the result of Equation 29 is zero for every interval, the sum of multiple intervals is zero as well:

$$\sum_{i=0}^{N-1}\left[1200\log_2\left(\frac{l_i f_{0,n+m_i}\sqrt{1 + B_{n+m_i}l_i^2}}{f_{0,n}k_i\sqrt{1 + B_n k_i^2}}\right) - 2^{c_i/1200}\right] = 0 \tag{33}$$

where N is the number of intervals and k_i, l_i, m_i and c_i are the values k, l, m and c for a specific interval. Solving $f_{0,n}$ from this equation results in:

$$f_{0,n} = \left(\frac{\prod_{i=0}^{N-1} A_i}{2^{C_{sum}/1200}}\right)^{1/N} \tag{34}$$

where

$$A_i = \frac{l_i f_{0,n+m_i}\sqrt{1 + B_{n+m_i}l^2}}{k_i\sqrt{1 + B_n k_i^2}} \tag{35}$$

and

$$C_{sum} = \sum_{i=1}^{N-1} c_i \tag{36}$$

5.2 Weights

When an interval between two piano tones produces beats, these beats are not only produced by the first matching partials of each interval according to the k and l values of Table 3, but also by higher partials that are integer multiples of k and l. All the partials producing beats in an interval can be taken into consideration by calculating weights that depend on the maximum amplitude and the amount of amplitude modulation that each pair of partials produce as well as by ignoring beats that fall under the masking threshold of other beats.

Weights can be added to Equation 34 to acquire the following equation:

$$f_{0,n} = \left(\frac{\prod_{i=0}^{N-1} {A_i}^{w_i}}{2^{C_{sum,w}/1200}} \right)^{1/\sum_{i=0}^{N-1} w_i}, \tag{37}$$

where:

$$C_{sum,w} = \sum_{i=1}^{N-1} c_i w_i \tag{38}$$

where w_i is the weights of a specific interval.

Weights are calculated with the following steps:

1. Find magnitudes of partial frequencies: The magnitudes of partial frequencies can be found after finding the estimates for f_0 and B with the PFD and then using a small window close to the estimated partial frequencies to find the magnitudes of spectral peaks. With the MAT algorithm, partial magnitudes can be stored as the algorithm finds partials.

2. Apply A-Weighting: A-Weighting is applied to estimate the frequency sensitivity of human hearing. A-Weighting is applied with equation:

$$A(f) = 20 \log_{10} (R_A(f)) + 0.17, \tag{39}$$

where

$$R_A(f) = \frac{12194^2 f^4}{(f^2 + 20.6^2)\sqrt{(f^2 - 107.7^2)(f^2 + 737.9^2)}(f^2 + 12194^2)} \tag{40}$$

3. Masking: The maximum amplitude of a beat, as well as the difference between two partials producing a beat are taken into consideration when determining which beats are not heard by the listener. If the maximum amplitude of a

beat is under the masking threshold of another beat, it will not be taken into consideration when calculating weights. The same is done if the magnitudes of partials producing beats deviate too much as in this case the louder partial will mask the softer one, and no beats will be heard.

4. Weights: After all partials under the masking threshold are taken out of consideration, the weights for each set of partials are calculated. The weights are calculated with the following equation:

$$w_i = \frac{M_i}{\sum_{n=0}^{N-1} M_n} \tag{41}$$

where w_i is the weights of ith matching partials, M_i is the maximum magnitude of the beats produced by the partials and N is the total number of matching partials over the masking threshold.

The weights are distributed in a way where the sum of the weights for each interval is one. This way the weight of each interval is the same. Figure 25 shows the weights of the octave between E_3 and E_4. It can be seen that the sum of these weights is one and that all beats under the threshold have a weight of zero.

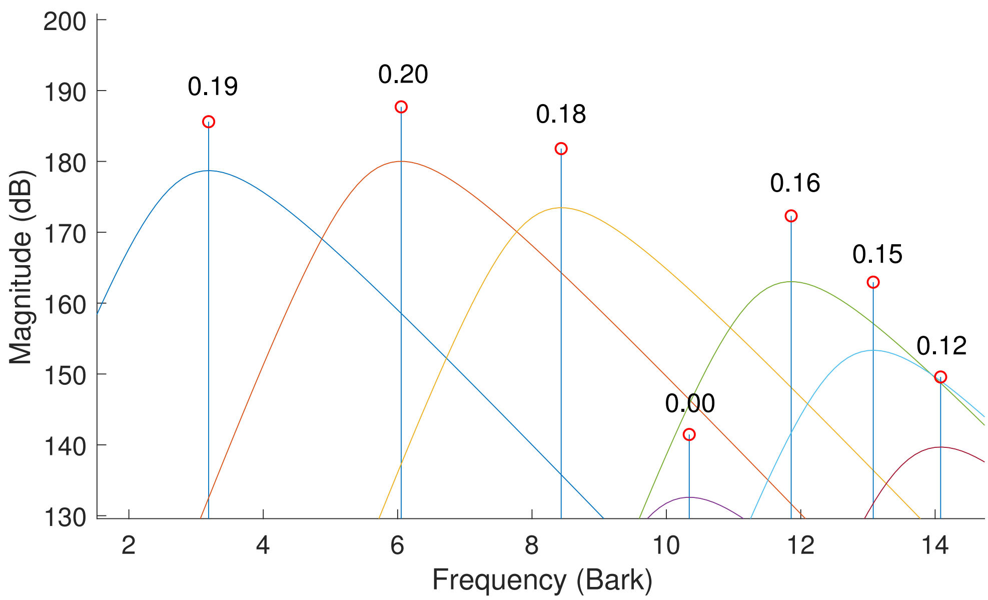


Figure 25: Weights of the beats between E_3 and E_4 (octave).

6 Results

In this section, the performance of the novel tuning process is evaluated. Before doing this, the inharmonicity coefficient algorithms are compared to find out which one of them, PFD or MAT, is better suitable to work as a part of the tuning process.

6.1 Partial Frequency Estimation Algorithms

The inharmonicity coefficient estimation algorithms were compared by looking at their computation times (runtime), as well as their accuracy. The accuracy of the algorithms was determined by comparing the partial frequencies calculated with Equation 3, using f_0 and B estimates gotten with the algorithms, to frequency peaks in the spectrum of the tone. Each spectrum was examined manually to ensure that phantom partials were not treated as partial frequencies. For tones with a high number of partials, the first 20 partials were used to calculate the error, because beyond that point, phantom partials started to be so prominent that the evaluation became much more difficult.

Runtimes of both algorithms for each key can be seen in Figure 26. It can be seen that the MAT algorithm is faster for almost every key of the keyboard. The runtimes of the MAT algorithm were between 0.04 and 0.33 seconds with an average runtime of 0.18 seconds. The runtimes of the PFD algorithm were between 0.18 and 0.58 seconds with an average of 0.27 seconds.

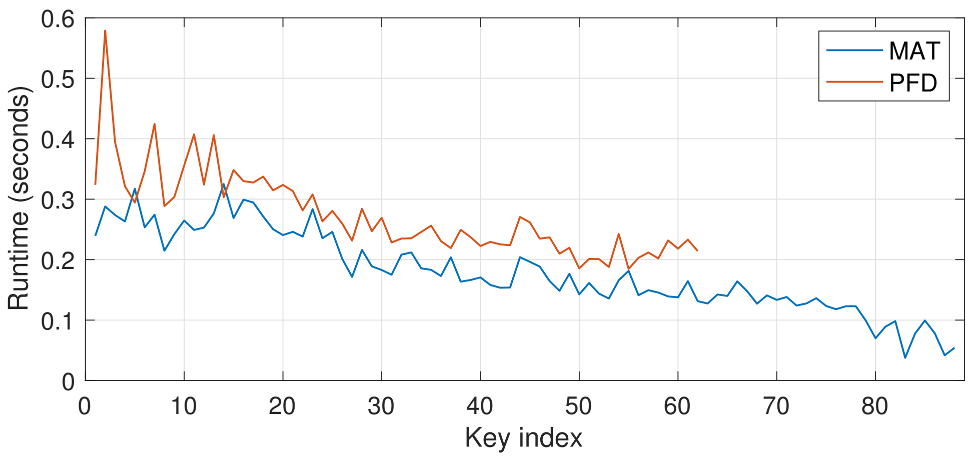


Figure 26: Runtime of PFD and MAT algorithms.

The accuracy of both algorithms can be seen in Figure 27. The accuracy is displayed as the difference between the estimated and real values of partial frequencies in cents. It can be seen that the PFD algorithm is missing the value for its accuracy beyond the 62nd key. This is because the algorithm could not make estimations for higher keys. Even though the algorithm could not perform estimation for the higher keys, it still could be considered to be used for the lower keys, if it gave more accurate estimations. However, this is not the case. The MAT algorithm outperforms the PFD algorithm with few exceptions. The RMS error for the MAT algorithm is between 0.001 and 12 cents with an average of 1.95 cents. The RMS error for the PFD algorithm is between 0.25 and 14.95 cents with an average of 3.29 cents.

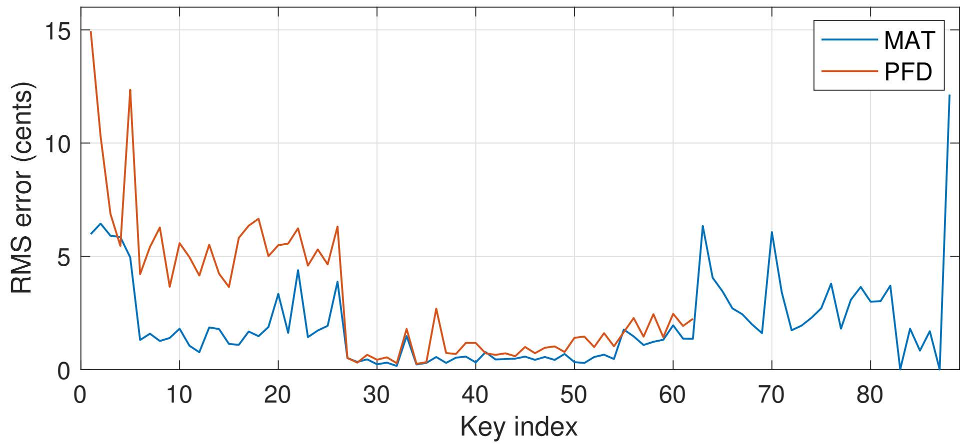


Figure 27: Error of PFD (orange) and MAT (blue) algorithms.

With superior runtimes as well as better accuracy, it is clear that the MAT algorithm should be used as a part of the novel tuning process. However, it should be noted that the algorithm does have a high error in both ends of the keyboard which could have a major effect on the performance of the tuning process.

6.2 Novel Tuning Process

Real piano tones were recorded from a Yamaha grand piano to evaluate the performance of the novel tuning process. The piano had been tuned on the previous morning by a professional tuner. The leftmost string from each of the 88 string unisons of the piano was recorded with a DPA 4061 microphone using a Focusrite Scarlett 2i4 and a MacBook Pro.

The tuning process was emulated by resampling the recordings to achieve the fundamental frequency calculated by the novel tuning system. After that, the frequencies of the first partials of recorded piano tones (tuning done by professional tuner) were compared against the first partials of resampled piano tones (tuning done by novel tuning system).

6.2.1 Weights

The novel tuning system is evaluated with and without weights. The model with no weights uses only the first matching partials of each interval and the ideal fundamental frequency is calculated with Equation 34. The model with weights uses Equation 37 to take into account higher partials of each interval.

6.2.2 Intervals

The note A_4 is used as the reference note and the reference frequency is set to be the same as the first partial of A_4 according to the tuning done by the professional tuner. This is done so that the two tunings can be compared against each other. After this, the reference octave is tuned using the intervals displayed in Figure 28. The crosses in the figure show the note that is being tuned and the circles show the reference notes. The intervals follow the "Defebaugh F-F" temperament, used by piano tuners [2] with the exception that in addition to intervals used in the original temperament, all other available intervals of thirds, fourths, fifths, sixths, sevenths and octaves are used.

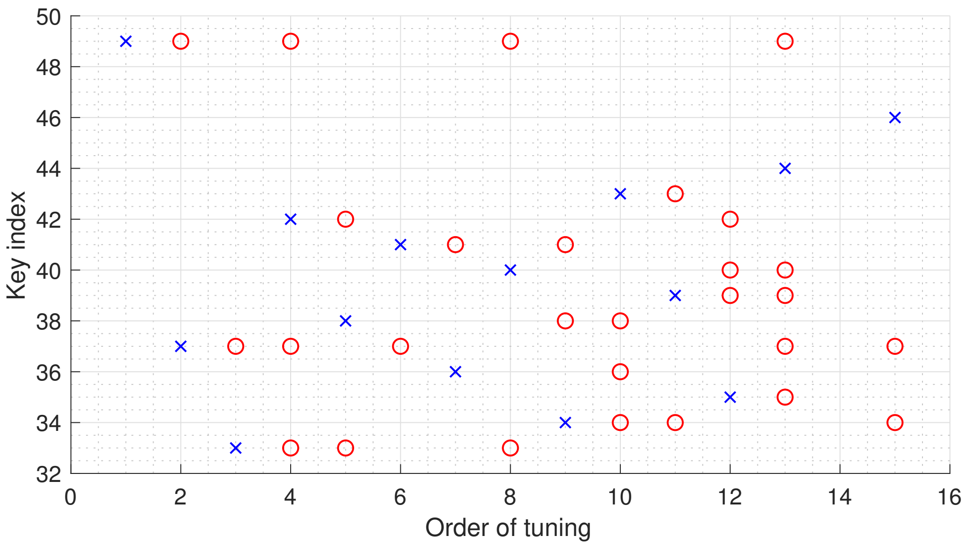


Figure 28: Intervals used for tuning in the reference octave. The crosses represent the key that is tuned and the circles represent the keys that are used as a reference.

Outside the reference octave three differing sets of intervals are used:

1. Octaves: The first set includes only octaves.
2. Octaves and double octaves: The second set includes octaves and double octaves. Double octaves are used when keys two octaves apart from each other have been tuned.
3. Intervals suggested by aural tuning system: For the third set, intervals suggested by an aural tuning system are used [2]. Keys above the reference octave, intervals shown in Figure 29 are used. For the first three keys (46-48) octaves and sixths are used and after that tenth (octave and a third), seventeenth (octave and a fifth) and double octave are used. For keys below the reference octave intervals shown in Figure 30 are used. First fifth, octave and tenth are used for keys 32 to 22 and after that seventeenth, double octave and double octave and a third is added to the set of intervals for the rest of the keys (22-1).

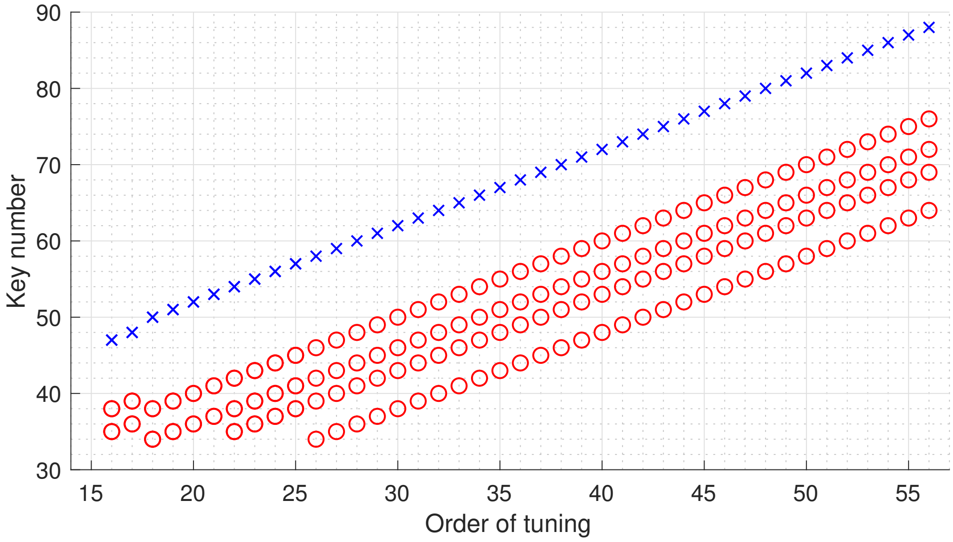


Figure 29: Intervals used to tune keys above the reference octave with the third set of intervals. The crosses represent the key that is tuned and the circles represent the keys that are used as a reference.

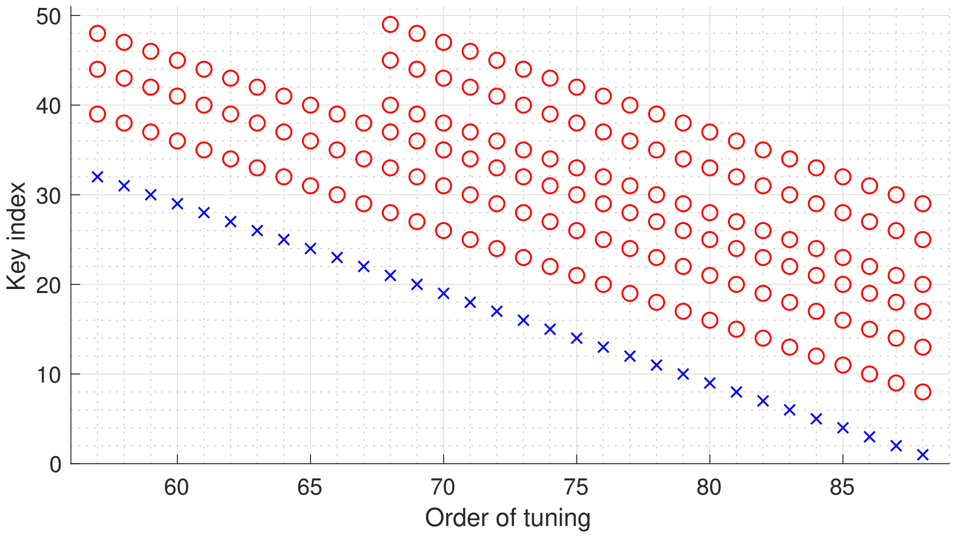


Figure 30: Intervals used to tune keys below the reference octave with in the third set of intervals. The crosses represent the key that is tuned and the circles represent the keys that are used as a reference.

6.3 Comparison

6.3.1 No Weights

All three sets of intervals were first evaluated without weights, using first matching partials of each interval. The reference octave is tuned in each set with the same intervals, shown in Figure 28. The tuning done with the novel process came very close to the tuner within the reference octave with a RMS error of 1.3 cents.

Tuning done with the first set of intervals can be seen in Figure 31. Using only an octave for the tuning, provided tuning close to the ET12 below the reference octave, with a RMS difference of 10.8 cents from the tuner. Above the reference octave novel tuning came closer to the tuner with a RMS difference of 6.1 cents. Difference between tuning done by the novel process and tuning can be seen in Table 6. Overall, the difference is 8.1 cents.

Table 6: Average deviation between a professional tuner and the novel process without weights using the first set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	10.8
Reference octave	1.3
$F\#_3$ to C_8	6.8
All	8.1

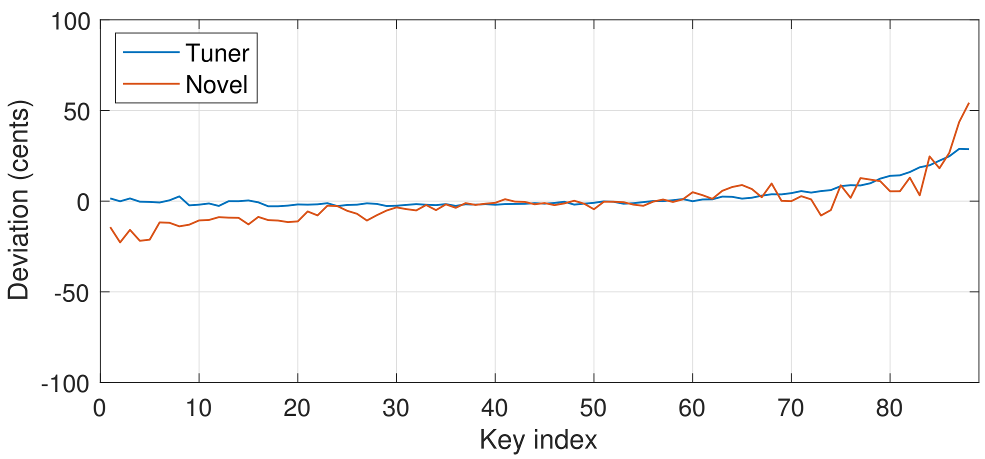


Figure 31: Tuning done without weights and the first set of intervals.

Tuning done with the second set of intervals is shown in Figure 32. In this case, the tuning below the reference octave is slightly better with a RMS difference of 9.3 cents, but tuning above the reference octave deviates more from the tuner with a RMS error of 7.3 cents. Difference between the tunings can be seen in Table 7. Overall, difference is 7.3 cents.

Figure 33 shows tuning done with the third set of intervals. Below the reference octave this set provides by far the closest tuning to the tuner, with a RMS difference

Table 7: Average deviation between a professional tuner and the novel process with no weights and the second set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	9.3
Reference octave	1.3
$F\#_3$ to C_8	7.3
All	7.4

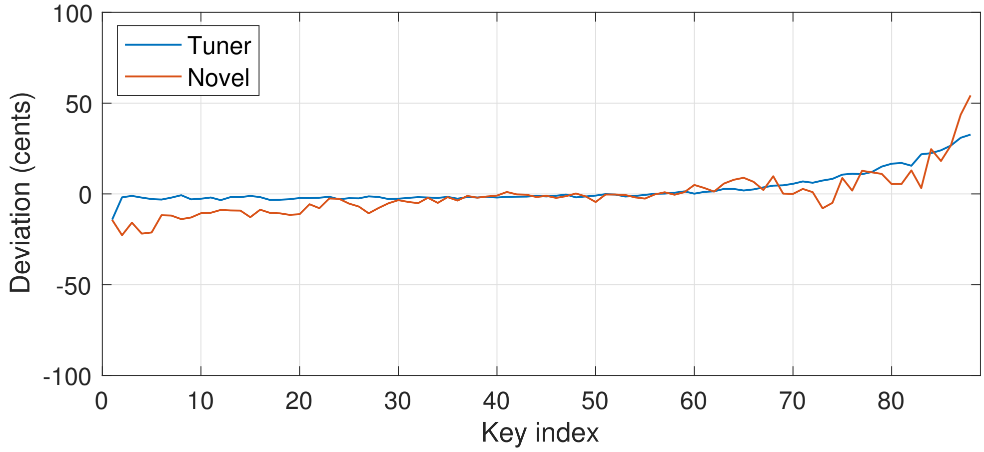


Figure 32: Tuning done with no weights and the second set of intervals.

of 3.2 cents. On the other hand, above the reference octave there is a great amount of deviation for the tuner with a RMS difference of 22.58 cents. Difference between the tunings can be seen in Table 8. The overall difference in the final set is 15.2 cents.

By combining the third set of intervals below the reference octave and the first set of intervals above the reference octave, tuning closest to the tuner is achieved. Figure 34 shows this tuning. Difference between his tuning and that of the tuner can be seen in Table 9. With these intervals, an overall difference of only 5.1 cents is attained.

Table 8: Average deviation between a professional tuner and the novel process with no weights and the third set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	3.22
Reference octave	1.3
$F\#_3$ to C_8	22.6
All	15.2

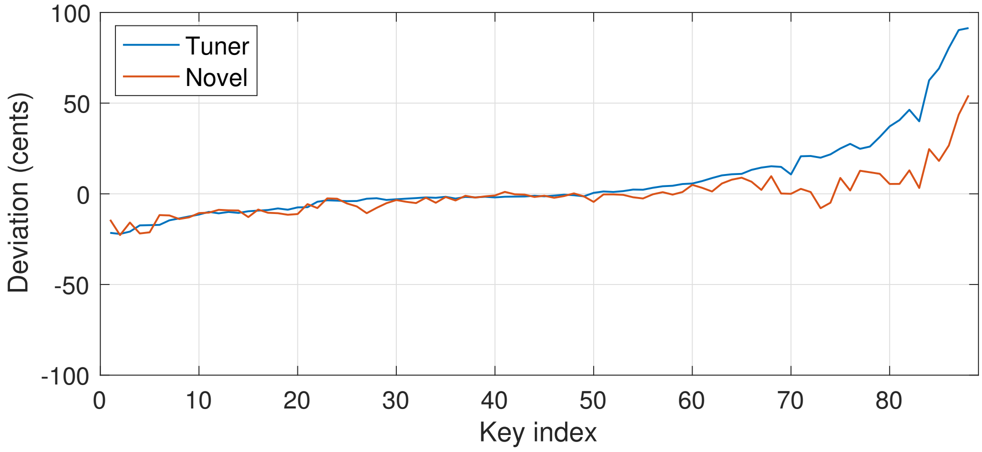


Figure 33: Tuning done with no weights and the third set of intervals.

Table 9: Average deviation between a professional tuner and the novel process with no weights, using the third set of intervals below the reference octave and the first set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	3.22
Reference octave	1.3
$F\#_3$ to C_8	6.8
All	5.1

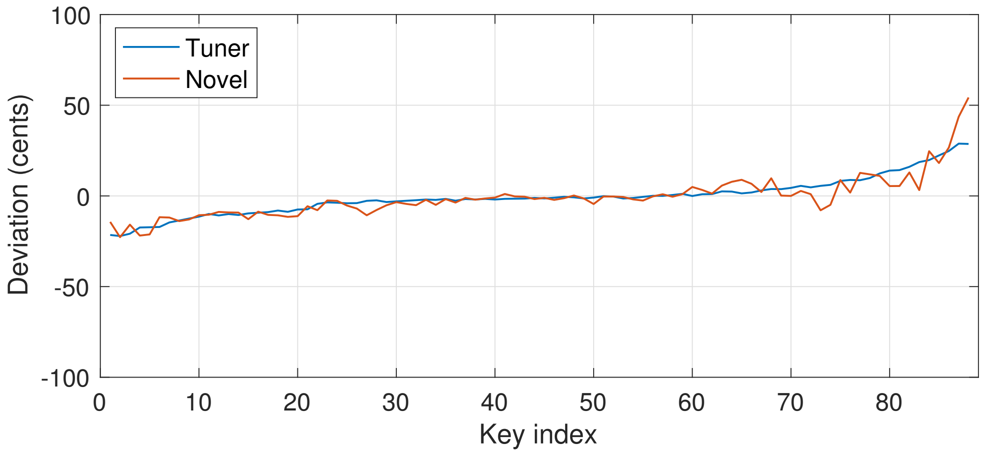


Figure 34: Tuning done with no weights, using the third set of intervals below the reference octave and the first set of intervals above the reference octave.

6.3.2 With Weights

Taking all the partials of piano tones into consideration when using the novel tuning process leads to a RMS error of more than 74 cents, which is unacceptable. Because of this only the first ten, fifteen and twenty partials were first considered with the first set of intervals and the partial count with the least deviation from the tuner was then chosen to be used for evaluation of the other interval sets.

Figure 35 shows tuning done with the first set of intervals using the first ten partials for weights. The reference octave is close to the tuner with a RMS difference of 1.4 cents, which is slightly higher than without weights. Below the reference octave, the tuning is between the ET12 and tuning done by the tuner with a RMS difference of 5.67 cents. Above the reference octave, the tuning has quite large difference with a RMS of 9.6 cents. Difference between the tunings can be seen in Table 10. Overall, difference between the tunings is 7.4 cents.

Table 10: Average deviation between a professional tuner and the novel process with weights using first 10 partials and the first set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	5.6
Reference octave	1.4
$F\#_3$ to C_8	9.6
All	7.4

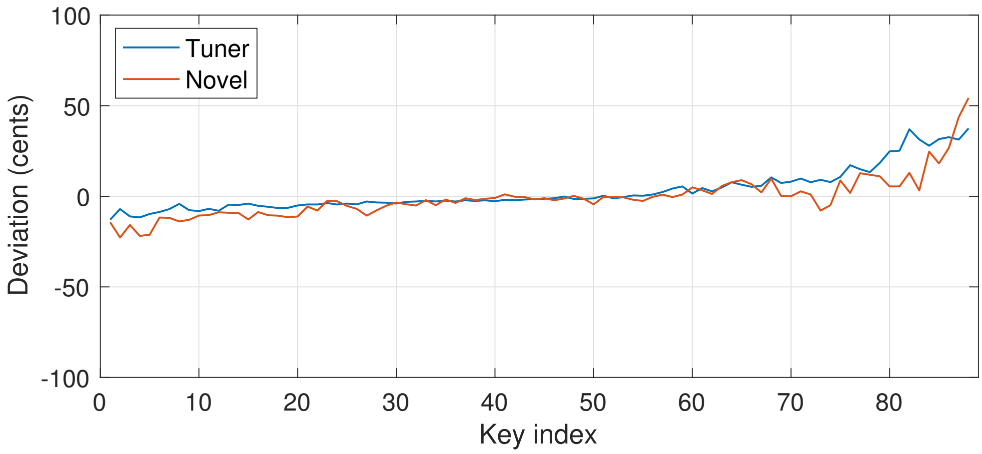


Figure 35: Tuning done with weights using the first ten partials and the first set of intervals.

Figure 36 shows tuning done with the first set of intervals and the first fifteen partials. The reference octave has the same difference as tuning done with 10 partials with a RMS of 1.4 cents. Below the reference octave, the tuning is much better with a RMS of 4.1 cents. Above the reference octave, the difference is also similar but a slightly smaller with a RMS of 9.4 cents. Difference between the tunings can be seen in Table 11. Overall, the difference between the tunings is 7.0 cents.

Figure 37 shows tuning done with the first set of intervals and the first twenty partials. The difference in the reference octave is 1.4 cents

Table 11: Average deviation between a professional tuner and the novel process with weights using the first 15 partials and the first set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	4.1
Reference octave	1.4
$F\#_3$ to C_8	9.4
All	7.0

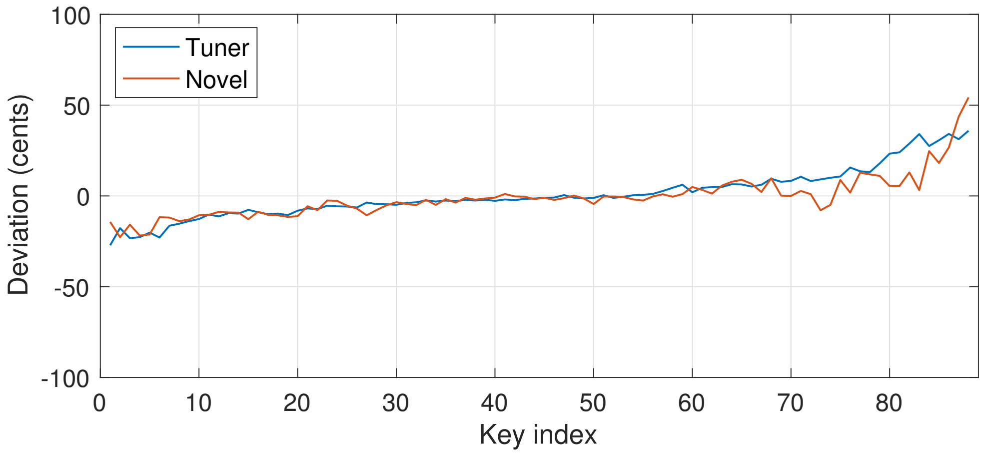


Figure 36: Tuning done with weights using the first fifteen partials and the first set of intervals.

It can be seen that the reference octave (deviation of 1.38 cents) and keys above the reference octave (deviation of 9.19 cents) are similar to the results obtained using the first fifteen partials, as the tones do not have much more than 15 partials. Below the reference octave, the tuning produces a much more stretched tuning with a difference of 9.30 cents. If more partials would be taken into consideration, the tuning would become even more stretched.

Using the first fifteen partials provided the best result for the first set of intervals. It can be seen that the choice of partial count affected only keys below the reference octave. This is because many of the higher keys do not have more than 10 partials. The other sets of intervals were also evaluated with several partial counts and as

Table 12: Average deviation between a professional tuner and the novel process with weights using first 20 partials and the first set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	9.3
Reference octave	1.4
$F\#_3$ to C_8	9.2
All	8.5

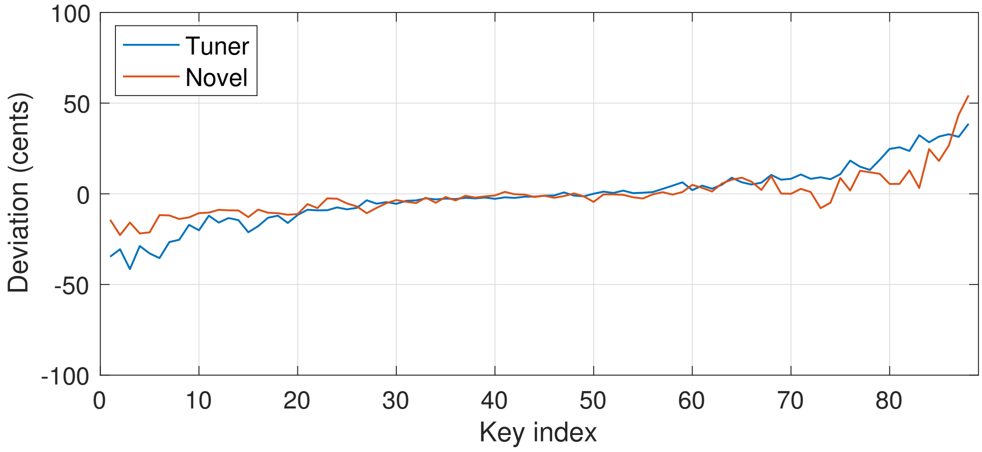


Figure 37: Tuning done with weights and the first set of intervals. First twenty partials used for weights.

with the first set of intervals, 15 partials provided tuning closest to the tuner.

Figure 38 shows tuning done with the second set of intervals using the first fifteen partials. Tuning in the reference octave is the same for the rest of the interval sets with a RMS difference of 1.4 cents. Below the reference octave, the tuning is very close to the tuning done by the tuner with a deviation of 3.4 cents. Above the reference octave, there is less deviation than with the first set of intervals, with a difference of 8.31 cents. The overall deviation is 6.12 cents. Table 13 shows the difference between the tunings.

Figure 38 shows tuning done with the third set of intervals and the first fifteen partials. The tuning is rather stretched with the third set as there is 7.4 cents of deviation below the reference octave and 26.6 above it. The overall deviation is 18.3 cents. Table 14 shows difference between the tunings.

With weights, the first fifteen partials of set 2 (Figure 38) provide the tuning closest to that of the tuner with a deviation of 6.1 cents which is much higher than the ideal tuning without weights (5.1 cents). The overall deviation of every tuning can be seen in Table 15. The best tuning both with and without weights is highlighted with bold text in the table.

Table 13: Average deviation between a professional tuner and the novel process with weights using the first 15 partials and the second set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	3.4
Reference octave	1.4
$F\#_3$ to C_8	8.3
All	6.1

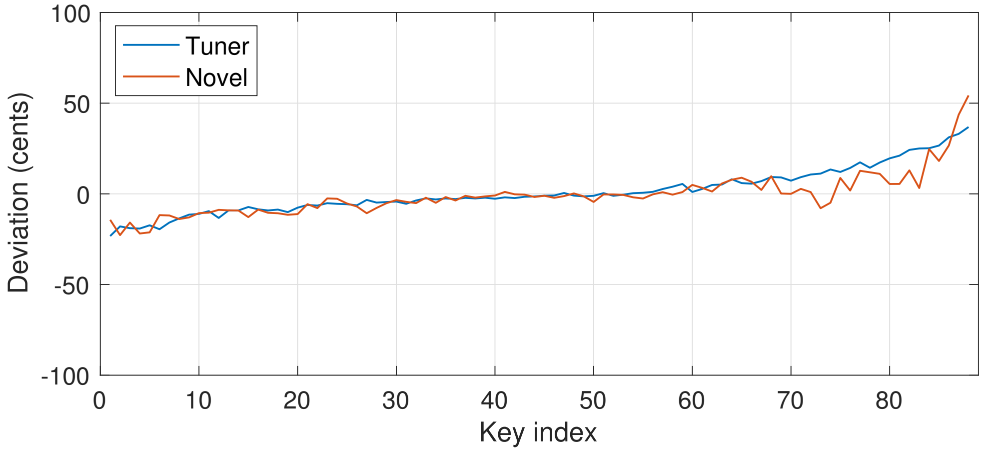


Figure 38: Tuning done with weights and the second set of intervals. The first fifteen partials used for weights.

Table 14: Average deviation between a professional tuner and the novel process with weights using first 15 partials and the third set of intervals.

Keys	**RMS deviation (cents)**
A_0 to E_3	7.3
Reference octave	1.4
$F\#_3$ to C_8	26.6
All	18.3

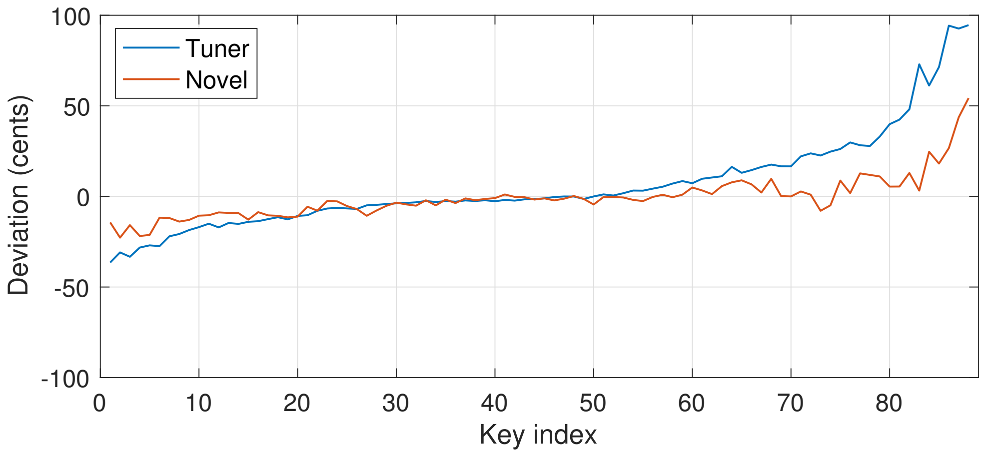


Figure 39: Tuning done with weights and the third set of intervals. The first fifteen partials used for weights.

Table 15: Average deviation between all tunings. Best tunings with and without weights are highlighted with bold text.

Tuning	**the RMS deviation (cents) of all keys**
Without weights	
Interval set 1	8.1
Interval set 2	7.4
Interval set 3	15.2
Interval set 3+1	**5.1**
With weights	
Interval set 1, 10 partials	7.4
Interval set 1, 15 partials	7.0
Interval set 1, 20 partials	8.5
Interval set 2, 15 partials	**6.1**
Interval set 3, 15 partials	18.3

7 Conclusion

In this thesis, a novel piano tuning process was developed. The process calculates optimal beating rates for several intervals, using the values of the fundamental frequency and the inharmonicity coefficient. To maximize the number of intervals that can be compared against, the process also determines the order in which the keys of the piano should be tuned in.

The process was evaluated by comparing it to a tuning conducted by an expert tuner. As the process can be done using several intervals and differing weights, comparisons were made with and without weights, using three sets of intervals. Using no weights with intervals in the following order:

- F_3 to F_4 (reference octave): "Defebaugh F-F"-temperament.
- $F\#_4$ to C_8: Match octaves.
- E_3 to $F\#_2$: Match the fifth, the octave and the 10th.
- F_2 to A_0: Match the fifth, the octave, the 10th, the 17th, the double octave and the double octave and a third.

provided the best result with an overall RMS deviation of 5.1 cents.

With weights, a deviation of 6.1 cents was achieved. To get this result, the number of partials taken into consideration had to be restricted to 15. The model used to calculate weights is overly simplified and this is probably the reason why all partials could not be taken into consideration. A more accurate model could possibly provide a tuning which is better than the tuning done with no weights.

The novel tuning process is well suited for the semi-automatic piano tuning system. Most of the deviation within the tuning is in the treble end of the instrument, where decay times of tones are short and the pitch of a tone is much harder to hear. It is possible that the tuning realized by the novel process is even better than that of the tuner, as the tuning accomplished by the tuner deviated above and below the ET12 in the treble end.

References

[1] D. W. Martin and W. Ward, "Subjective evaluation of musical scale temperament in pianos," *The Journal of the Acoustical Society of America*, vol. 33, no. 5, pp. 582–585, 1961.

[2] A. A. Reblitz, *Piano Servicing, Tuning, & Rebuilding: For the Professional, the Student, the Hobbyist*. Vestal Press, 1976.

[3] T. D. Rossing, F. R. Moore, and P. A. Wheeler, *The Science of Sound*. Addison Wesley, 1990.

[4] H. A. Conklin Jr, "Design and tone in the mechanoacoustic piano. part iii. piano strings and scale design," *The Journal of the Acoustical Society of America*, vol. 100, no. 3, pp. 1286–1298, 1996.

[5] "Harmonic partials on strings." https://en.wikipedia.org/wiki/Harmonic_series_(music)#/media/File:Harmonic_partials_on_strings.svg. Accessed: 2019-02-13.

[6] H. Fletcher, E. D. Blackham, and R. Stratton, "Quality of piano tones," *The Journal of the Acoustical Society of America*, vol. 34, no. 6, pp. 749–761, 1962.

[7] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*. Springer, 1991.

[8] H. A. Conklin Jr, "Generation of partials due to nonlinear mixing in a stringed instrument," *The Journal of the Acoustical Society of America*, vol. 105, no. 1, pp. 536–545, 1999.

[9] G. Weinreich, "Coupled piano strings," *The Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1474–1484, 1977.

[10] R. E. Kirk, "Tuning preferences for piano unison groups," *The Journal of the Acoustical Society of America*, vol. 31, no. 12, pp. 1644–1648, 1959.

[11] T. C. Hundley, H. Benioff, and D. W. Martin, "Factors contributing to the multiple rate of piano tone decay," *The Journal of the Acoustical Society of America*, vol. 64, no. 5, pp. 1303–1309, 1978.

[12] J. Burred, *The Acoustics of the Piano*. PhD thesis, Ph. D. thesis, Professional Conservatory of Music Arturo Soria, Madrid, Spain, 2004.

[13] N. Giordano and A. Korty, "Motion of a piano string: Longitudinal vibrations and the role of the bridge," *The Journal of the Acoustical Society of America*, vol. 100, no. 6, pp. 3899–3908, 1996.

[14] A. Galembo and A. Askenfelt, "Measuring inharmonicity through pitch extraction," *Journal STL-QPSR*, vol. 35, no. 1, pp. 135–144, 1994.

[15] A. Galembo and A. Askenfelt, "Signal representation and estimation of spectral parameters by inharmonic comb filters with application to the piano," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 2, pp. 197–203, 1999.

[16] J. Rauhala, H.-M. Lehtonen, and V. Välimäki, "Fast automatic inharmonicity estimation algorithm," *The Journal of the Acoustical Society of America*, vol. 121, no. 5, pp. EL184–EL189, 2007.

[17] M. Hodgkinson, J. Wang, J. Timoney, and V. Lazzarini, "Handling inharmonic series with median-adjustive trajectories," in *Proc. 12th Int. Conf. Digital Audio Effects (DAFx-09)*, (Como, Italy), pp. 1–7, Sept. 2009.

[18] J. Rauhala and V. Välimäki, "F0 estimation of inharmonic piano tones using partial frequencies deviation method.," in *Proc. Int. Computer Music Conf. (ICMC-07)*, (Copenhagen, Denmark), pp. 453–456, Aug. 2007.

[19] J. O. Smith *et al.*, *Spectral Audio Signal Processing*, vol. 1334027739. W3K, 2011.

[20] R. G. Lyons, *Understanding Digital Signal Processing (3rd Edition)*. Prentice Hall, nov 2010.

[21] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Mathematics of Computation*, vol. 19, no. 90, pp. 297–301, 1965.

[22] V. Pulkki and M. Karjalainen, *Communication Acoustics: an Introduction to Speech, Audio and Psychoacoustics*. John Wiley & Sons, 2015.

[23] J. G. Roederer, *The Physics and Psychophysics of Music: An Introduction*. Springer, 2001.

[24] M. R. Schroeder, B. S. Atal, and J. Hall, "Optimizing digital speech coders by exploiting masking properties of the human ear," *The Journal of the Acoustical Society of America*, vol. 66, no. 6, pp. 1647–1652, 1979.

[25] T. Painter and A. Spanias, "A review of algorithms for perceptual coding of digital audio signals," in *13th International Conference on Digital Signal Processing Proceedings*, vol. 1, pp. 179–208, IEEE, 1997.

[26] Y. Lin and W. H. Abdulla, "Audio watermarking techniques," in *Audio Watermark*, pp. 51–94, Springer, 2015.

[27] "Railsback stretch." `https://upload.wikimedia.org/wikipedia/commons/a/ae/Railsback2.png`. Accessed: 2019-02-13.

[28] J. Lattard, "Influence of inharmonicity on the tuning of a piano—measurements and mathematical simulation," *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 46–53, 1993.

[29] H. Hinrichsen, "Entropy-based tuning of musical instruments," *Revista brasileira de Ensino de Física*, vol. 34, no. 2, pp. 1–8, 2012.

[30] A. Stulov, "Physical modelling of the piano string scale," *Applied Acoustics*, vol. 69, no. 11, pp. 977–984, 2008.