การวัดการตอบสนองความถี่โดยใช้เสียงมนุษย์

The measurement of frequency response by using voice as sound source

Mr. Kittitorn Himasuk

Mr. Kris Wannawong

Miss Watsaya Takkapaijit

Mr. Veerapat Pongyart

Submitted in partial fulfillment of the requirement for
the bachelor’s degree in engineering program
Department of Computer Engineering,
Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang
Academic year of 2018

Department of engineering

King Mongkut's Institute of Technology Ladkrabang

Thesis certificate

Thesis Title The measurement of frequency response by using voice as sound source

Student Mr. Kittitorn Himasuk, Mr. Kris Wannawong, Miss Watsaya Takkapaijit, Mr. Veerapat Pongyart

Student ID. 58010096, 58010150, 58011142, 58011181

Degree Bachelor of Engineering

Program Music Engineering and Multimedia

Thesis advisor Asst. Prof. Munhum Park

Thesis advisor	Signature
Asst. Prof. Munhum Park	Parkmunh

หัวข้อปริญญานิพนธ์ การวัดการตอบสนองความถี่โดยใช้เสียงมนุษย์

นักศึกษา นายกิตติธร หิมะสุข, นายคริษฐ์ วรรณวงศ์,

นางสาววัศยา ตรรกไพจิตร, นายวีรภัทร พงศ์ญาติ

รหัสประจำตัว 58010096, 58010150, 58011142, 58011181

ปริญญา วิศวกรรมศาสตร์บัณฑิต B.Eng.(Music E

สาขาวิชา วิศวกรรมดนตรี และสื่อประสม

พ.ศ. 2561

อาจารย์ที่ปรึกษาปริญญานิพนธ์ ผศ.ดร. มุนฮึม บัก

บทคัดย่อ

ในโครงงานนี้พวกเรามีจุดมุ่งหมายที่จะวัดค่าตอบสนองต่อย่านความถี่ต่างๆ (frequency responses) โดยใช้เสียงคนเป็นแหล่งกำเนิดเสียง การเตรียมการทำการทดลองนั้นพวกเราเริ่มจากการวัดอิมเพาซ์เรสปอนซ์ (impulse response) โดยใช้เสียงไซน์สวีพ (sine sweep) ที่เป็นเสียงมาตราฐานในการวัด ต่อมาพวกเราได้ทำการบันทึกเสียงนักร้องสี่คนโดยให้นักร้องทำการร้องออกมาใน สี่โครมาติกสเกล (chromatic scales) ซึ่งมีความแต่งต่างกันอยู่ที่ 25 เซน เพื่อทำการวิเคราะห์ค่า ซิกแนลทูน๊อยซ์เรโชว (signal-to-noise ratio) หลังจากที่ได้ผลของเสียงนักร้องทั้งสี่คนจึงนำเสียงที่ได้ทำการบันทึกมานั้นปล่อยเสียงผ่านลำโพงโดยมีไมโครโฟนสองตัว ซึ่งมีทั้งตำแหน่งใกล้และไกลเพื่อทำการบันทึกอีกครั้งและวิเคราะห์ผลลัพธ์ที่ออกมา เมื่อเราทำการวิเคราะห์เป็นที่เรียบร้อยแล้ว เราสามารถสรุปได้ว่าเราสามารถใช้เสียงของมนษย์เป็นแหล่งกำเนิดเสียงได้ถึงแม้ว่าเสียงของมนุษย์จะมีข้อจำกัดก็ตาม

Thesis	The measurement of frequency response by using voice as sound source
Student	Mr. Kittitorn Himasuk, Mr. Kris Wannawong, Miss Watsaya Takkapaijit, Mr. Veerapat Pongyart
Student ID.	58010096, 58010150, 58011142, 58011181
Degree	Bachelor of Engineering
Program	Music Engineering and Multimedia
Year	2018
Thesis Advisor	Asst. Prof. Munhum Park

Abstract

In the current project, we aimed to measure frequency responses (FR) by using voice as sound source. First, we measured the reference impulse response (IR) by using a sine sweep as the sound source. Then, we recorded the singing voices where four singers sang 4 chromatic scales differing by 25 cents, from which we could analyze the signal-to-noise ratio (SNR). Having found that the singing voice has sufficiently broad spectra, we played the recorded voices through a speaker while simultaneously recording the sound at 2 locations: one very close to the loudspeaker and the other at a distance. We analyzed the data and could conclude that human voice can be used as sound source although it is band-limited.

Acknowledgement

We would like to extend our gratitude to all those who enable us to complete this project. We would like to give a special thanks to all the singers who took part in the recording sessions.

Index

Page

บทคัดย่อ......I

Abstract......II

Acknowledgement......III

Index......IV

Figure index......VI

1 Introduction......1

1.1 Objectives......1

1.2 Literature reviews......1

1.3 Scope of Thesis......1

2 Theory......2

2.1 Impulse response and frequency response......2

2.1.1 Introduction......2

2.1.2 Measurement of impulse response......3

2.2 Spectrum of human voice......3

2.2.1 Human voice frequency range......3

2.2.2 Harmonics......3

2.2.3 Frequency & Pitch note......4

3 Methodology......6

3.1 Measuring impulse response using sine sweep......6

3.2 Recording voice from singers......7

3.3 Measuring impulse response using voice......10

4 Results & Discussions......11

4.1 Measuring Impulse response by sine sweep......11

4.1.1 Source mic and room mic......11

4.1.2 Frequency response......11

4.1.3 Impulse response......12

4.2 Recording voice from singer results......12

4.2.1 Spectrum of voice......12

4.2.1.1 Spectrum of voice Male 01......12

4.2.2 Maximum spectrum of voice compare to background noise......13

4.2.2.1 Spectrum of voice compare to noise floor example 13

4.2.3 Signal to noise ratio of 4 singers......14

4.3 Frequency response of all voice as sound source & sine sweep......14

4.4 Impulse response of all voice as sound source & sine sweep......15

5 Summary......16

Reference......17

Figure index

Figure Page

2.1 Block diagram of LTI system where given input is an impulse 2
2.2 Block diagram of LTI system given input is any signal 2
2.3 Measurement block diagram........ 3
2.4 Relation of pitch note & frequency........ 4
2.5 Relation of note & frequency........ 5
3.1 Measurement configuration........ 6
3.2 Python generated sine sweep signal........ 7
3.3 Configuration for the voice recording 8
3.4 Voice spectrum of first note........ 8
3.5 Voice spectrum of first & second note 9
3.6 All voice spectrum in 4 chromatic scale 9
3.7 Max. spectrum of voice and noise floor 10
3.8 Signal to noise ratio of one person........ 10
4.1 Source-mic signal........ 11
4.2 Room-mic signal........ 11
4.3 Frequency responses, H and H'........ 11
4.4 Impulse responses, h and h' 12
4.5 Spectrum of all note of Male01........ 13
4.6 Spectrum of voice and noise floor of Male01 13
4.7 Signal-to-noise ratio plotting procedure........ 14

4.8 Frequency responses using voice as sound source compared to those measured with sine sweep......14

4.9 Impulse response using voice as sound source & sine sweep......15

Introduction

Impulse response, h(t) is the output of a system when the given input is an impulse, δ(t). It is extremely useful to predict the output of a linear-time invariant system. In real life, it is very difficult for any input to perfectly be an impulse. So, rather than directly recording the impulse response, people analyze frequency band energies [1] first by taking Fourier transform. Then, we can take inverse Fourier transform to get impulse response. Nowadays, several sound sources are used in the measurement [2]. Also, there are various techniques for the impulse response measurement. In this project, we want to know if we can use human voice to measure the impulse response, and we will analyze the results in comparison with the sine sweep technique [4].

1.1 Objectives

- To measure frequency response and impulse response by using human voice.
- To compare the frequency responses of human voice with the frequency response of sine-sweep.

1.2 Literature reviews

- The basics of measuring impulse response and frequency response.
- Analysis of human singing voice.
- Data analysis on python.

1.3 Scope of Thesis

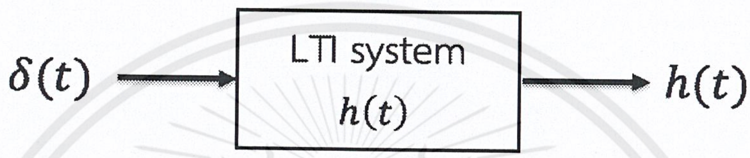
In order to compare the results, we measured singing voice of 4 amateur singers. Then, we analyzed the data to see frequency response and impulse response. Finally, we compared them with the frequency response and impulse response using sine-sweep.

Theory

2.1 Impulse response and frequency response

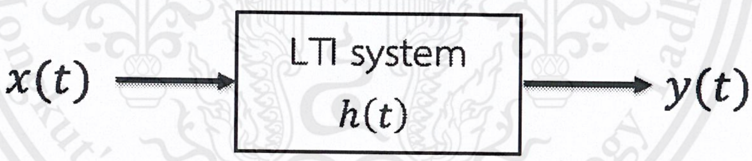
2.1.1 Introduction

Impulse response is the response of an LTI system, such as a filter, when the system's input is the unit impulse or $\delta(t)$. We call impulse response $h(t)$, which is the characteristic of the system (figure 2.1).



[Figure 2.1: Block diagram of LTI system where given input is an impulse]

Impulse response is extremely useful to predict the output of any LTI system. If we give the input signal, $x(t)$, to the system, we can predict the output, $y(t)$, by equation 2.1 where $y(t)$ is described by the convolution of $x(t)$ and $h(t)$.



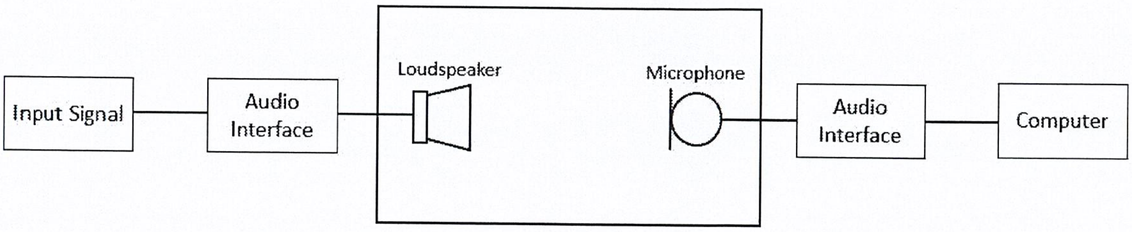
[Figure 2.2: Block diagram of LTI system given input is any signal]

$$y(t) = x(t) * h(t) \quad \text{[Equation 2.1]}$$

In reality, input signal cannot perfectly be an impulse. Therefore, people determine frequency response first. Frequency response, $H(j\omega)$, is the response of an LTI system in frequency domain, which contains magnitude and phase information. The relation between input, output and frequency response can be represented as in equation 2.2, where $Y(j\omega)$ is output signal, $X(j\omega)$ is input signal and $H(j\omega)$ is frequency response. When $H(j\omega)$ is calculated, we can simply take inverse Fourier transform to get impulse response, $h(t)$.

$$Y(j\omega) = X(j\omega)H(j\omega)$$ [Equation 2.2]

2.1.2 Measurement of impulse response



[Figure 2.3: Measurement block diagram]

According to figure 2.3, people normally use this measurement setup to measure impulse response. Sine sweep is used as a sound source. Then, amplifier in audio interface amplifies the signal, and RTA microphone is used to record the sound. After that, the measured data will be analyzed in computer.

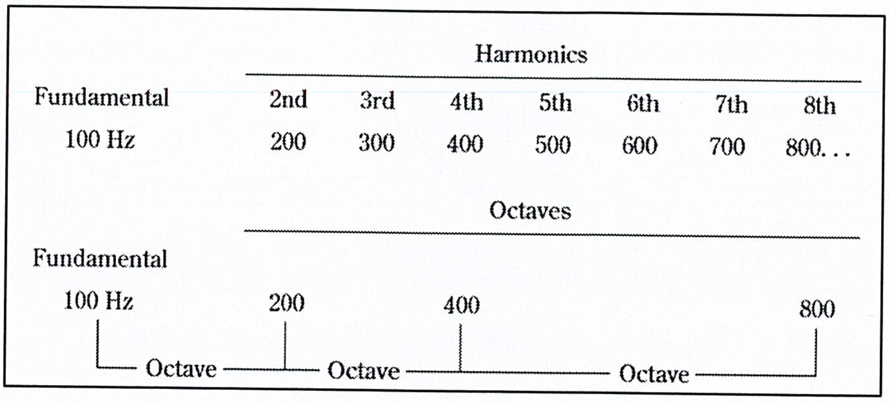
2.2 Spectrum of human voice

2.2.1 Human voice frequency range

The frequency range of average male voice is approximately between D2(73.42 Hz) - C6(1046.50) Hz, and for female voice is between E3(164.81) Hz - G6(1567.98) Hz[5]. So, the expected frequency response should be within these ranges and go above upper ranges because there should be some signal from 'harmonics', which will be explained in the next section.

2.2.2 Harmonics

Since human voice is complex wave, human voice is constructed by adding more than one frequency tone. The lowest frequency usually with the and highest amplitude tone is called fundamental. If the fundamental frequency of a sound is f, the harmonics of this sound will be $2f$, $3f$,$4f$ and so on. For example, in figure 2.4, if the fundamental frequency is 100 Hz, the harmonics will be 200 Hz, 300 Hz, 400 Hz and so on. The amplitude of harmonics is usually lower than the fundamental and may decrease when the frequency increases.



[Figure 2.4: Relation of pitch note & Frequency]

Reference: F. Alton Everest, 2001, Master Handbook of Acoustics, Fourth Edition. The McGraw Hill

2.2.3 Frequency & Pitch note

The frequency of note can be calculated, but first, you have to know the center frequency of note, which is 440 Hz. It matches the note for A4. Then you can use the equation below to find a frequency of the other notes.

$$f = 2^{\frac{n}{12}} \times 440\,Hz \qquad \text{[Equation 2.3]}$$

Where f is frequency of note.

n is number of semitones between the note and center frequency note. For example, if you want to find the frequency of F4 which is lower than A4 for 4 semitones, n must be equal to -4.

$$f_{F4} = 2^{\frac{-4}{12}} \times 440\,Hz$$

$$f_{F4} = 394.288\,Hz$$

If you want to find C5 which higher than A4 by 3 semitones, then n will be equal to 3.

$$f_{C5} = 2^{\frac{3}{12}} \times 440\,Hz$$

$$f_{C5} = 523.251\,Hz$$

You can see the relation of pitch note & frequency in figure 2.5.

NOTE TO FREQUENCY CHART

NOTE	HZ	NOTE	HZ	NOTE	HZ
A0	27.5000	D3	146.832	G5	783.991
A♯0 or Bb0	29.1352	D♯3 or Eb3	155.563	G♯5 or Ab5	830.609
B0	30.8677	E3	164.814	A5	880.000
C1	32.7032	F3	174.614	A♯5 or Bb5	932.328
C♯1 or Db1	34.6478	F♯3 or Gb3	184.997	B5	987.767
D1	36.7081	G3	195.998	C6	1046.50
D♯1 or Eb1	38.8909	G♯3 or Ab3	207.652	C♯6 or Db6	1108.73
E1	41.2034	A3	220.000	D6	1174.66
F1	43.6535	A♯3 or Bb3	233.082	D♯6 or Eb6	1244.51
F♯1 or Gb1	46.2493	B3	246.942	E6	1318.51
G1	48.9994	C4	261.626	F6	1396.91
G♯1 or Ab1	51.9131	C♯4 or Db4	277.183	F♯6 or Gb6	1479.98
A1	55.0000	D4	293.665	G6	1567.98
A♯1 or Bb1	58.2705	D♯4 or Eb4	311.127	G♯6 or Ab6	1661.22
B1	61.7354	E4	329.628	A6	1760.00
C2	65.4064	F4	349.228	A♯6 or Bb6	1864.66
C♯2 or Db2	69.2957	F♯4 or Gb4	369.994	B6	1975.53
D2	73.4162	G4	391.995	C7	2093.00
D♯2 or Eb2	77.7817	G♯4 or Ab4	415.305	C♯7 or Db7	2217.46
E2	82.4069	A4	440.000	D7	2349.32
F2	87.3071	A♯4 or Bb4	466.164	D♯7 or Eb7	2489.02
F♯2 or Gb2	92.4986	B4	493.883	E7	2637.02
G2	97.9989	C5	523.251	F7	2793.83
G♯2 or Ab2	103.826	C♯5 or Db5	554.365	F♯7 or Gb7	2959.96
A2	110.000	D5	587.330	G7	3135.96
A♯2 or Bb2	116.541	D♯5 or Eb5	622.254	G♯7 or Ab7	3322.44
B2	123.471	E5	659.255	A7	3520.00
C3	130.813	F5	698.456	A♯7 or Bb7	3729.31
C♯3 or Db3	138.591	F♯5 or Gb5	739.989	B7	3951.07
				C8	4186.01

[Figure 2.5 Relation of note & frequency] -- source

Reference: https://www.doctormix.com/blog/note-to-frequency-chart

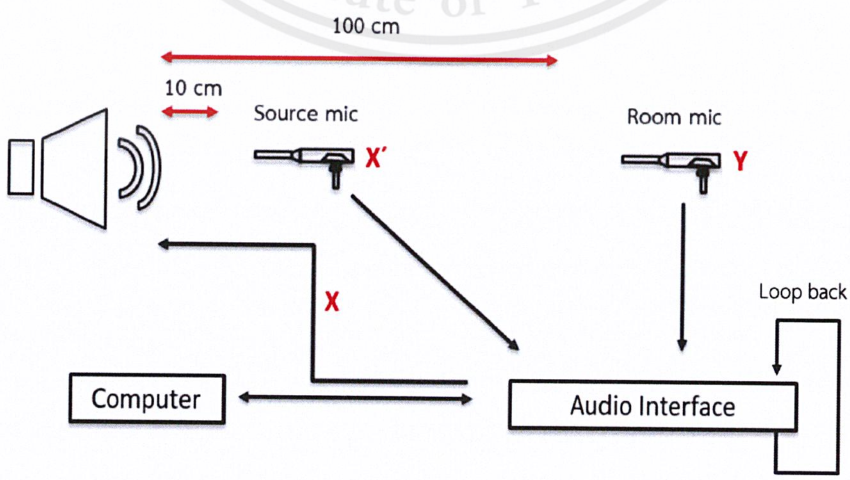
Methodology

To complete the goal of this project, we need to do 3 tasks to make sure that the impulse response measurement using human voice is possible: 1) Measuring impulse response using sine sweep, 2) recording voice for singers; and 3) measuring impulse response using human voice.

In the current project, we used the following devices for the measurement: Earthworks M30 microphone, RTA microphones, Shure MX153 wireless microphone, RME Fireface 802 audio interface and Vented-box speaker (custom-made).

3.1 Measuring impulse response using sine sweep

Since we do not know the actual signal of human voice, we needed to use microphone close to the source instead. However, we need to make sure that the signal from the microphone can be used. To compare between real input signal and signal from microphone close to the source, we measured impulse response by using sine sweep played back over the speaker (see figure 3.1). We setup 2 microphones. In this configuration, we call the first microphone 'source mic', which is a microphone 10 centimeters away from the speaker. The spectrum of the signal from the source mic is referred to as X'. The second microphone is 'room mic', which is a microphone 100 centimetres away from the speaker. The spectrum of the signal from the room mic is called Y. The sound source, of which the spectrum is X, is a sine sweep signal from 20 Hz to 20 kHz.



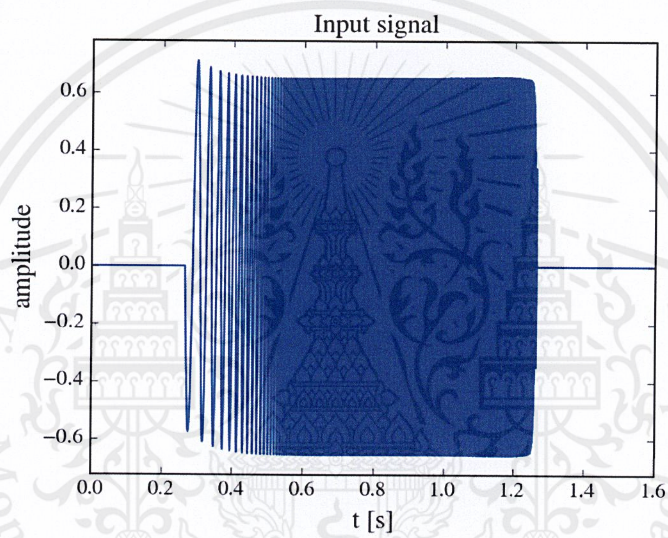
[Figure 3.1: Measurement configuration]

We used a dedicated Python script to generate the sine sweep (see figure 3.2) and to simultaneously record two channels. Then, we analyzed by plotting frequency response H and H', where H and H' are defined by the following equations.

$$H = \frac{Y}{X} \quad \text{[Equation 3.1]}$$

$$H' = \frac{Y}{X'} \quad \text{[Equation 3.2]}$$

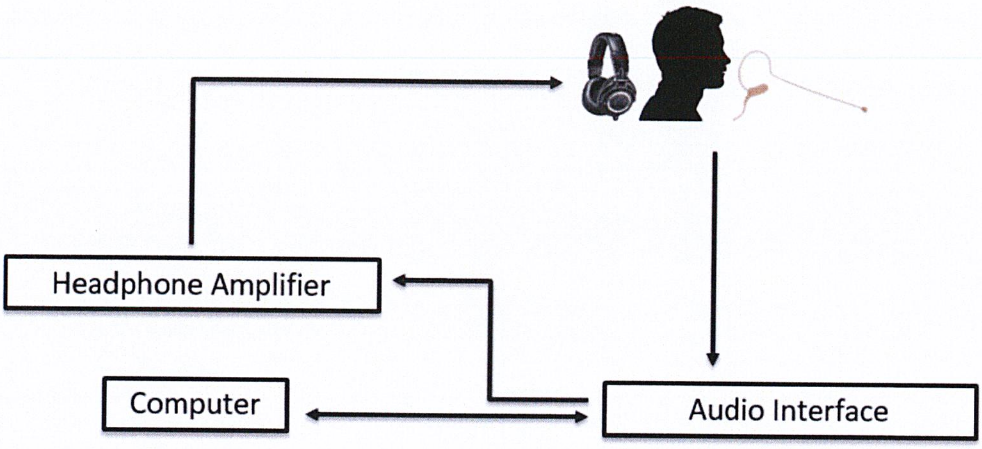
After we have H and H', we can use them to determine h and h' by the inverse Fourier transform, where h is an impulse response of H, and h' is an impulse response of H', respectively.



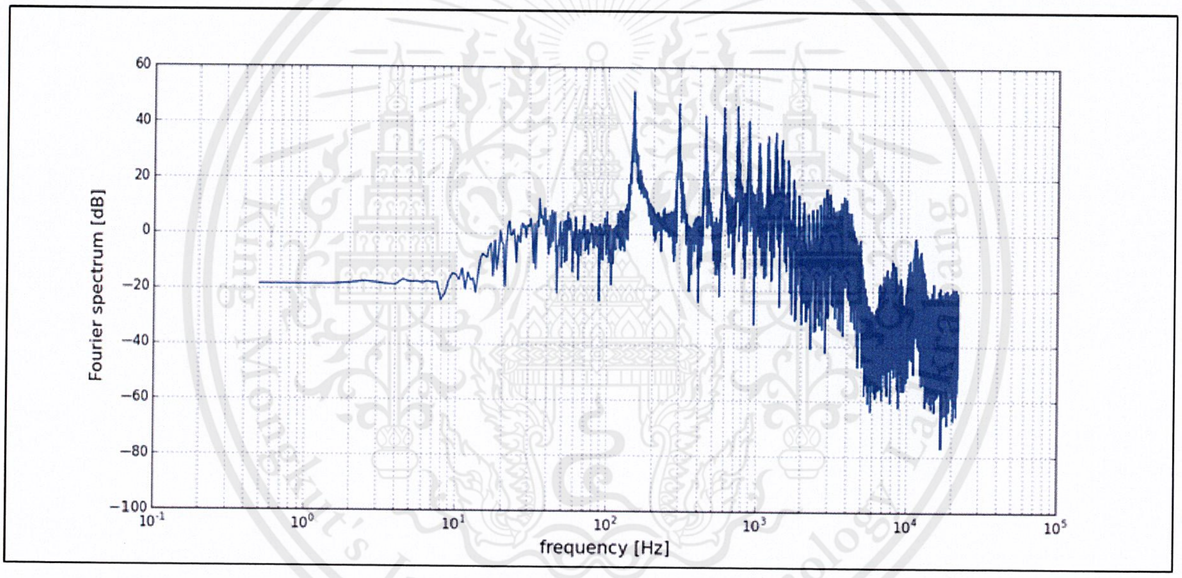
[Figure 3.2: Python generated sine sweep signal]

3.2 Recording voice from singers

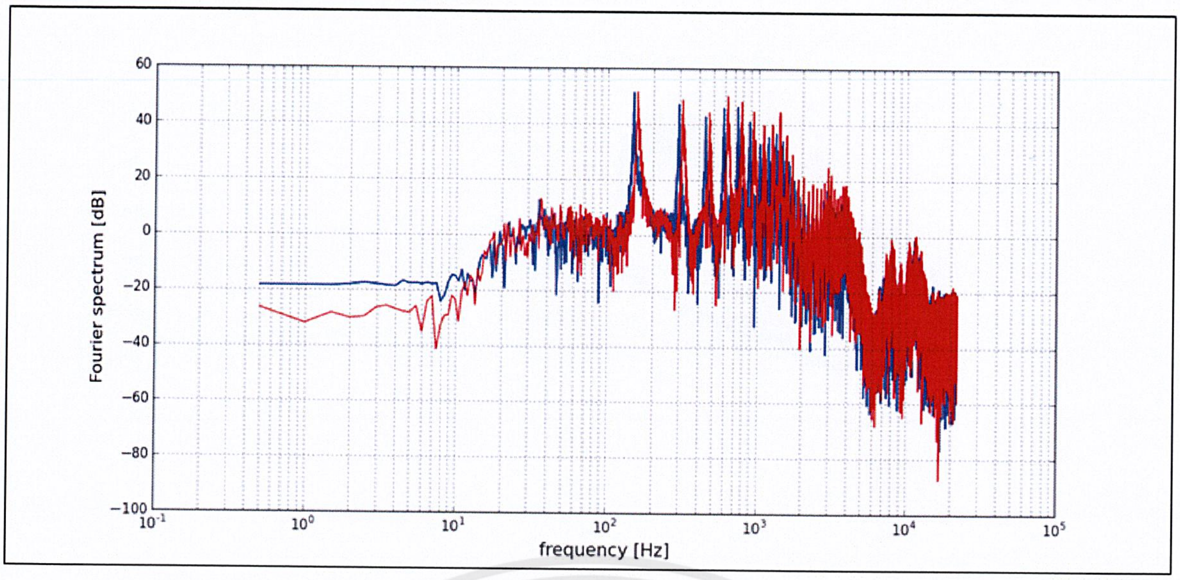
We invited 4 singers to record the voice (see figure 3.3). First, we had to find the lowest and highest notes that the singers could sing and let them sing the chromatic scale between their lowest and highest notes. Then, we increased the pitch by +25 cents, +50 cents and +75 cents. After we finished recording singer's voice, we plotted the spectra from the first note (see figure 3.4), second note (see figure 3.5) and so on. When we have the spectra of all notes (see figure 3.6), we analyzed and compared the maximum spectrum of all 4 scales to the averaged noise floor (see figure 3.7), and finally, we the noise floor from the maximum spectrum to get signal to noise ratio (see figure 3.8).



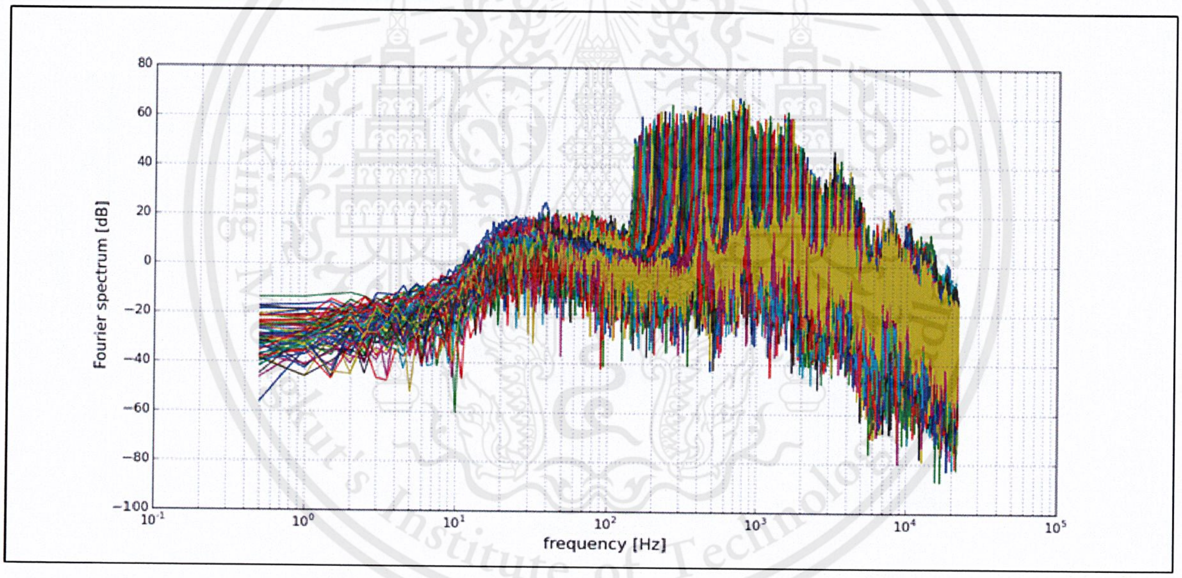
[Figure 3.3: Configuration for the voice recording]



[Figure 3.4: Voice spectrum of first note]



[Figure 3.5: Voice spectrum of first & second note]



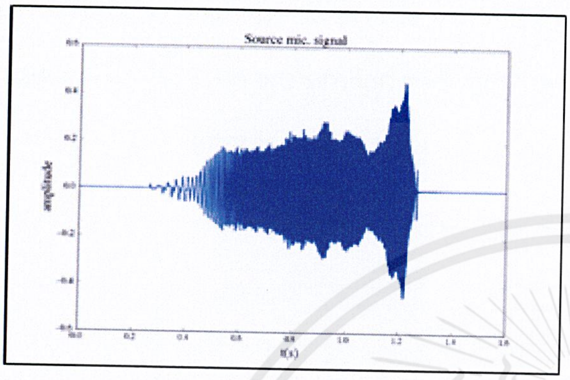
[Figure 3.6: All voice spectrum in 4 chromatic scale]

4

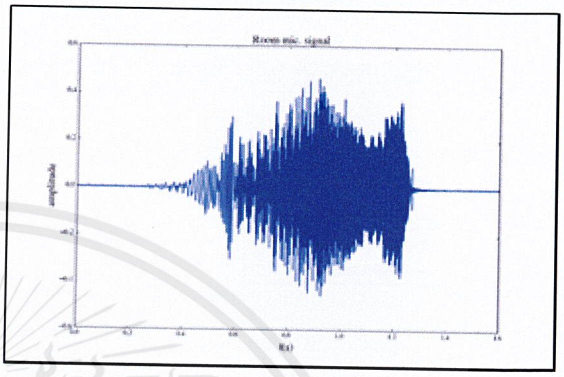
Results & Discussions

4.1 Measuring Impulse response by sine sweep

4.1.1 Source mic and room mic



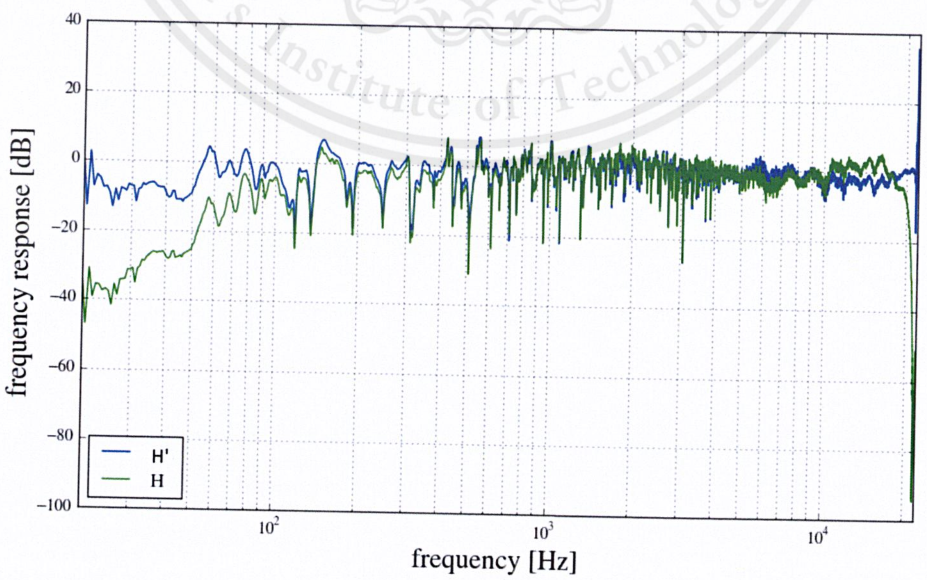
[Figure 4.1: Source-mic signal]



[Figure 4.2: Room-mic signal]

Figures 4.1 and 4.2 show the source-mic and room-mic signal using sine sweep as sound source. We can see that room-mic signal has lower amplitude and more noise than source-mic signal, because room mic was placed 100 cm away from speaker. Moreover, some reflections occurred in the room and some external noise. Furthermore, mid frequency of sine sweep tends to decrease as the distance increases.

4.1.2 Frequency response



[Figure 4.3: Frequency responses, H and H']

Summary

From the first part of the measurement results, the frequency response measured with sine sweep and that measured with the source mic signal were similar to each other in the frequency range from 100 Hz to 10,000 Hz. However, the response with sine sweep rolled off and differed from the response with source-mic signal at 20 Hz to 100 Hz and at the frequencies above 10 kHz because H contained one microphone response only in the numerator, which behaves as a band pass filter itself, whereas the microphone responses were cancelled out in H'.

It appears that human voice may be used as sound source with sufficient energy although band-limited. Comparing frequency response using sine sweep and human voice, it appears that human voice may be used as sound source to measure frequency response. However, the impulse responses of singers' voice appeared to be indecent since we did not apply any filters and thresholding to the frequency responses.

Reference

[1] J. S. Abel, N. J. Bryan, P. P. Huang, M. Kolar, and B. V. Pentcheva, "Estimating Room Impulse Responses from Recorded Balloon Pops," in *Audio Engineering Society Convention 129*, 2010.

[2] B. Kakonyi, R. Abdul Jaleel, and A. Novak, "Balloon Explosion," in *Audio Engineering Society Convention 144*, 2018.

[3 E. Shabalina, M. Kaiser, J. Ramuscak, and M. Vorländer, "Live Measurements of Subwoofers' Performance," *J. Audio Eng. Soc*, vol. 61, no. 3, pp. 138–142, 2013.

[4] G.-B. Stan, J.-J. Embrechts, and D. Archambeau, "Comparison of Different Impulse Response Measurement Techniques," *J. Audio Eng. Soc*, vol. 50, no. 4, pp. 249–262, 2002.

[5] Olga Banis, "What's My Voice Type? The Different Voice Types and How to Distinguish Them", September 2, 2012, <http://choirly.com/whats-my-voice-type/>