Middle East Technical University
Electrical-Electronics Engineering Department

EE301 Term Project (2016-2017 Fall Semester)

Title: Analysis/synthesis and recognition of the sound signals of various musical instruments and human voice using Fourier Series Representation


Background:

Sound signals have a vitally important position in our lives. As a human being with auditory sensors, we are capable of hearing and listening to various kind of sounds. We not only hear or listen but also arrange different sound notes in a systematical way to compose songs, which we play with musical instruments. Furthermore, we also use our voice to sing a song. In this project, we will analyze the sound signals of various musical instruments and human voice in frequency domain. We will see that Fourier series representation has a core position in the analysis of such sound signals.

The Fourier series representation is defined for periodic signals which have a certain frequency. This frequency is also referred to as "fundamental frequency" as you have seen in the EE-301 course. As you have also learnt in EE-301, such periodic signals can be decomposed into components whose frequencies are integer multiples of the fundamental frequency. These frequencies are also referred to as "harmonics".

Musical instruments that have a string (which are also named as "string instruments") typically generate a periodic signal that consists of those "harmonics". For instance, when a certain note is played by picking a string of a guitar, in addition to the frequency of the note that is played (for example the "la" ['A' in English] note has a frequency of 440 Hz), the generated sound signal have components with frequencies at 2x440 Hz, 3x440 Hz, …, etc. This reason is related to the harmonic oscillations of the string of the guitar. Such oscillations are presented in Fig.1.

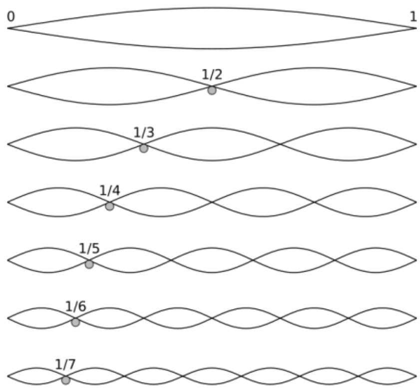

Fig.1: Fundamental and harmonic oscillations of a string

In Fig.1, the uppermost string oscillation has the frequency of the note to be played (for instance, it is 440 Hz for the “la” note). In addition to the oscillation in the fundamental frequency, there are other oscillations with a frequency of integer multiple of the fundamental frequency as can be observed in Fig.1. Such phenomena also apply to the string instrument violin. Even human voice have harmonics, since our vocal cords are also a sort of string. Therefore, the sound signals with such structure are periodic and they have a Fourier series representation. What makes the sound (or “timbre”) of the various musical instruments or the human voice different (even if they have the same fundamental frequency, or the note that is played by various instruments or sung by a human are the same) is the weights of the harmonic contents of the sounds they generate. The harmonic content weights correspond to the Fourier series coefficients. Remember that a periodic signal $s(t)$ with a certain fundamental frequency, w_0, can be decomposed as

$$s(t) = \sum_{k=-\infty}^{\infty} a_k e^{jkw_0 t}. \tag{1}$$

Even if the played note has a certain fundamental frequency, the values of a_k will differentiate the kind of the musical instrument or the voice of the human singing the note.

To start with, we will try to find the values of a_k for various musical instruments and a human voice.

Important Note: In this document, the frequency variable f ,whose units is Hz, is used instead of the frequency variable w, which is equal to $2\pi f$ and whose units is in radians/sec.

PHASE 1

Analysis of the sound signals

1) To represent a signal $s(t)$ as in (1), $s(t)$ must be an infinite duration signal. We will try to find a_k’s from a finite duration version of $s(t)$, namely $s_w(t)$. Find the continuous time Fourier transform (CTFT) of $s_w(t)$, namely $S_w(f)$, in terms of the Fourier transform of the infinite duration signal $s(t)$, which is denoted as $S(f)$. You may assume that $s_w(t)$ is equal to $s(t)$ times a rectangular window of length L. Without loss of generality, you may assume that the rectangular window is centered at $t = 0$.

2) Express $S_w(f)$ in terms of the Fourier Series coefficients of $s(t)$, namely a_k.

3) Explain how you can obtain a_k’s by sampling $S_w(f)$ by assuming that the fundamental frequency w_0 is an integer multiple of $2\pi/L$. What must be the frequency interval between the samples of $S_w(f)$?

4) Now, we will try to obtain $S_w(f)$ for a violin, flute and a human voice by using previously recorded .wav files. You can find those files in “Project, Part-1 Support Files” with the file names “violin_A4.wav”, “flute_A4.wav” and “voice_A4.wav”. In each .wav file, the “la” note is played with a violin, a flute or sung by a human. You can play those files to listen to their content.

   These .wav files contain samples of the continuous time sound signals, namely $s_w(nTs)$, T_s being the sampling period. Therefore, they are discrete time signals rather than being continuous. As you will learn later in the semester, if the sampling rate $F_s = 1/T_s$ is larger than the bandwidth[1] of $s_w(t)$, one can obtain $S_w(f)$ for $-\frac{F_s}{2} < f < +\frac{F_s}{2}$ based on $s_w(nTs)$. For the sound signals provided in the support files, the sampling rate is larger than the signal bandwidth. Therefore, we can find the CTFT, $S_w(f)$ through the samples $s_w(nT_s)$. Then we can obtain the Fourier series coefficients by inspecting $S_w(f)$ at some values of f. The concepts mentioned in this paragraph may not be clear to you. However, it will suffice that we can somehow find the samples of the CTFT of a signal in MATLAB and the Fourier series coefficients.

You are supposed to use the m-file "analyze_sound.m" in the support files. This m-file is provided as an incomplete m-file. You are supposed to write your own code segments in the m-file following the instructions specified in the m-file as comments. When you complete the m-file, it is supposed to be able to plot the samples of the CTFT for $-\frac{F_s}{2} < f < +\frac{F_s}{2}$, since F_s is larger than the signal bandwidth for an audio signal. The code segments related to CTFT computation are provided to you in the "analyze_sound.m" file.

Plot the time domain signals, $s_w(t)$, for the specified time intervals (1.8<t<1.84 for violin signal, 0.1<t<0.12 for flute and 0.17<t<0.21 seconds for voice signal). Are they periodic? Comment.

Plot also the magnitude and the phase of the samples of CTFT for $-\frac{F_s}{2} < f < +\frac{F_s}{2}$ of $s_w(t)$ for the three different audio signals (violin, flute and voice signals). Read the values of the peaks in the magnitude plot of CTFT at the harmonics of the fundamental frequency (which is 440 Hz for our case). Should these values be equal to a constant multiple of the magnitude of the Fourier series coefficients for the audio signals? Why? Explain the reason by considering your answer to Q3. Also note that the harmonics are not exactly at the multiples of 440 Hz. This is owing to the fact that there are some small tuning errors, thus the musical instrument or human voice does not exactly play or sing the note "la". For example, the first fundamental frequency of the voice signal is at about 463 Hz.

5) List the Fourier series coefficients for the 3 audio signals (violin, flute, voice) in a table. Normalize the values of the coefficients such that the one with the maximum magnitude has a magnitude of 1. You may list the first 10 coefficients (the one corresponding to the fundamental frequency and the higher harmonics) for the flute, 19 coefficients for the violin and 9 coefficients for the singer. Note that you are not required to find the Fourier series coefficient a_0 since the DC level for sound signals is equal to zero in general. You should provide the values of the Fourier coefficients in polar form, that is, you should provide the magnitude and the phase (in radians) of the Fourier series coefficients a_k.

6) Save the m-file that you have created with the file name "analyze_sound_Group_XX.m", where XX is your group number. For example, if your HW group is Group 5, you should save the m-file with the file name "analyze_sound_Group_05.m". Use "publish" command in MATLAB to publish your code in Microsoft Word format. Submit your m-file and the .doc file created by using "publish" command along with the other documents that you submit for Project Part-1 Submission.

[1] A windowed signal cannot be strictly band-limited. Yet, we can define an effective bandwidth which contains almost all the energy and this would be sufficient for the argument here.