SATFD

Lab. 1

February 23, 2024

1 Introduction

Musical instruments are divided into wind, stringed and percussion instruments. In stringed instruments, the source of sound is a plucked string (e.g., a guitar) or a struck string (e.g., a piano). While the theory describing the vibrations that are induced in a rigid string is very complex, to a first approximation, these vibrations can be described by a linear wave equation.

$$\frac{\partial^2 y}{\partial t^2} = \nu^2 \frac{\partial^2 y}{\partial x^2} \tag{1}$$

The speed of wave propagation can be expressed as:

$$\nu = \sqrt{\frac{T_0}{\rho_0}} \tag{2}$$

where T_0 is the tension of the string in the equilibrium position, and ρ_0 is the linear density (mass per unit length) of the string.

For a given string, these values are constant, which means that, to the first approximation, the speed of sound in the string can be considered constant, independent of the magnitude of that sound. Consequently, such vibrations can be induced in the string, for which there will be wave nodes at both ends of the string, while there will be odd multiples of halves of the wavelength of the string. The wavelength is related to its frequency as follows:

$$\nu = \lambda f \tag{3}$$

If the resting string length is L (for a free string of an acoustic guitar is L = 620 mm, for an electric L = 624 mm), then we can expect excitation of waves of the the following lengths (k means a fraction of $\frac{1}{k}$ string length corresponding to half the wavelength):

$$\lambda_k = \frac{2L}{k} \tag{4}$$

Which is equivalent to the frequencies:

$$f_k = \frac{\nu}{2L} \cdot k \tag{5}$$

If the fundamental frequency for $k = 1$ is denoted as f_1, it can be noted that the string can produce successive multiples of of this frequency - the so-called higher harmonic frequencies (overtones).

2 Elements of acoustics: equal temperament

The basic interval used in acoustics and harmonic analysis is the octave. Two sounds differ by an octave if their frequencies are in a ratio of 1:2 to each other. If the fundamental tone is f_1, then the tone an octave lower will have a frequency of $\frac{f_1}{2}$, and the tone an octave higher will have a frequency of $2f_1$. The tone two octaves higher has a frequency of $4f_1$, and the tone two octaves lower has a frequency of $\frac{f_1}{4}$, and so on. An octave is divided into twelve semitones. The names of these semitones are:

$$C_m; C\#_m; D_m; D\#_m; E_m; F_m; F\#_m; G_m; G\#_m; A_m; A\#_m; B_m,$$

where m denotes the octave number. The frequencies of semitones in relation to C_m have historically been chosen differently, however, since the 19th century, the so-called equally tempered tuning has been universally introduced, based on the fundamental tone $A_4 = 440Hz$ (it's worth noting that in this system, successive notes are labeled with the letters A, B, C, D, E, F, G). The equally tempered system is a musical tuning that involves dividing the octave into twelve equal semitones. The ratio of frequencies between two consecutive notes in the equally tempered system is $\sqrt[12]{2}$ (Wikipedia). The fundamental tone is the tone A_4 (the sound A in the 4th octave) with a frequency $f_A = 440$ Hz. In music, there are many different tunings (Wikipedia). Octaves are numbered with natural numbers starting from 0. For example, the range of notes on a piano covers octaves from 0 (A_0) to 7 (A_7). The frequency of each note f_p can be expressed in the number of semitones p that separate it from the note A_4:

$$f_p = f_A \cdot 2^{\frac{p}{12}} \tag{6}$$

$$p = 12 \cdot (ln_2 f - ln_2 f_A) \tag{7}$$

It can be observed that the scale of notes perceived by humans as uniform is essentially a logarithmic scale. This occurs because the human ear does not perceive the absolute frequency difference Δf, but only the relative frequency ratio $\frac{\Delta f}{f}$. Consequently, the intervals between successive notes increase proportionally to the frequencies of those notes. From relation (6), it can be easily shown that the relative frequency difference between adjacent semitones is constant, independent of their frequencies, and equal to: $2^{\frac{1}{12}} - 1$.

An interesting observation is the positioning of successive harmonics of the fundamental tone on the equally tempered scale. Table (1) presents the order of the harmonic and the interval of this harmonic from the fundamental tone, expressed in the number of semitones, while calculated according to the formula $n_k = 12 \ln_2 k$. Additionally, the name of the interval used in music is provided. Where the interval is approximate, the name is given in parentheses.

Table 1: Harmonics of the fundamental tone on the equally tempered scale.

Order of harmonic k	Interval n_k in semitones	Frequency ratio	Interval name
1	0	1:1	unison
2	12	2:1	octave
3	19.020	$\sim 3:2$	(octave + fifth)
4	24	4:1	2 octaves
5	27.863	$\sim 7:3$	(2 octaves + third)
6	31.020	$\sim 5:2$	(2 octaves + fifth)
7	33.688	$\sim 14:5$	(2 octaves + seventh)
8	36	8:1	3 octaves

As on can see from Table (1), higher harmonics of the fundamental frequency correspond to tones that are distant from the fundamental tone by one or several octaves ($k = 1, 2, 4, 8$), or approximately by an octave or several octaves and a fifth ($k = 3, 6, 12$). Because the frequency deviation is negligible in the case of octaves and small in the case of fifths, these tones are called perfect consonances. They form the basis of major chords. As seen, subsequent groups of harmonics ($k = 5, 10, 15$) and ($k = 7, 14, 21$) are distant from the fundamental tone by a certain number of octaves and thirds or sevenths, respectively. These are called imperfect consonances. Ancient scholars noticed that tones whose frequencies are multiples sound harmonious (con-sonans - hence consonance). This phenomenon is described by the so-called natural scale, in which intervals are described by frequency ratios of the form $p : q$, where p and q are small natural numbers (Wikipedia). Unfortunately, constructing such a scale where all frequencies of tones are multiples of each other is impossible because then the differences between successive tones become too large. Therefore, in the 19th century, the equally tempered scale was proposed, which, although introduces small deviations for fifths and slightly larger ones for thirds and sevenths, maintains a constant frequency ratio between adjacent tones, equal to $2^{\frac{1}{12}}$.

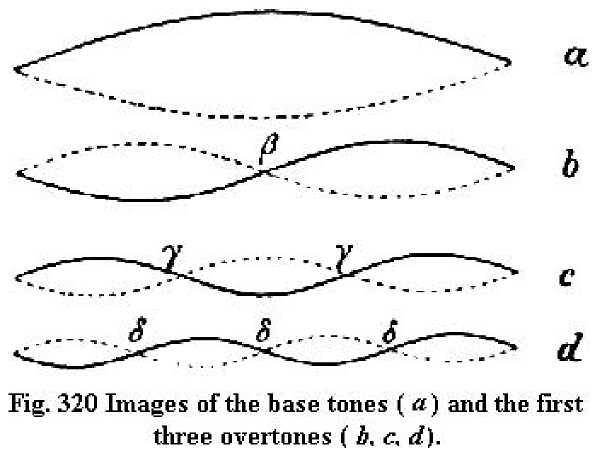

Figure 1: Stationary waves for strings. According to Rainer Radok, Acoustics

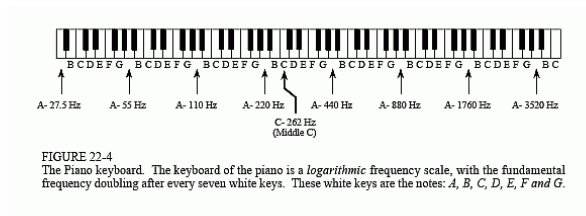

Figure 2: Division of sounds into octaves. According to "The Scientist and Engineer's Guide to Digital Signal Processing" by Steven W. Smith, Ph.D.

An example of a power spectrum including the fundamental tone and the first six higher harmonics is shown in Figure 3. The horizontal axis is on a linear scale to emphasize the fact that the even-tempered scale is logarithmic (frequency differences between successive tones depend on the frequency). The fundamental tone is C_4. As you can see, successive harmonics fall (approximately) on tones (intervals are given in parentheses): C_5 (octave), E_6 (third), G_6 (fifth), and $A\#_6$ (seventh). In practice, the tone is determined by the first 10–12 harmonics. In the intervals between harmonics, a spectrum of non-zero power is visible. Its sources are twofold. The first source is the higher frequencies (non-harmonics)—the so-called aliquots. In the first approximation of the theory, aliquots do not exist. They determine the timbre of the sound: for two different instruments, the power distribution of the aliquots will be different.

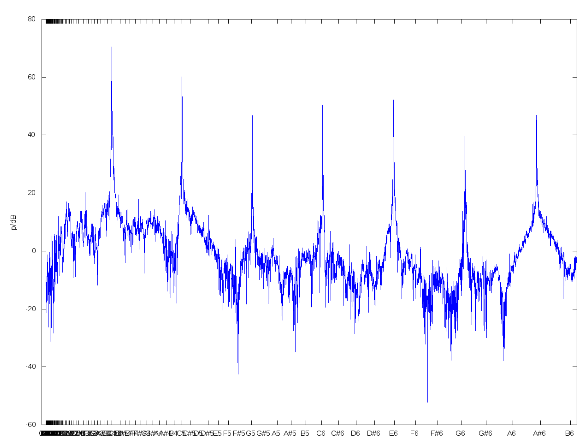

Figure 3: The power spectrum of the sound C_4 from a piano.

All these tones played simultaneously with the basic tone will sound nicely together, creating a consonance (as opposed to a dissonance, which does not sound nice). For example, the tonic triad in the C key consists of C, E and G tones, i.e., tones that are a third and a fifth away from the basic tone, respectively. You can see that these intervals are not random, but result from the location of the harmonic frequencies of the vibrating string. Consonances, or consonances, are the basis of harmony - the science of combining sounds.

3 Realization of the task

The suggested environment is MATLAB or Python.

Data:

- chord.wav - chord in WAV format,
- ex1.m - main script in Matlab format,
- tones.m - auxiliary script in Matlab format.

Suggested steps:

1. Load a wav file from disk with the wavread or load function (if there is no wavread).
2. Make a graph of the time series as a function of time.
3. Evaluate the stationarity of the signal (melody or chord).
4. Carry out a Fourier transform for the entire signal or for the relevant parts of the signal, e.g., with a Hamming window (or others suitably selected) and then a power spectrum.
5. Plot the power spectrum in units of dB/Hz, for the interval 16 Hz–4 kHz.
6. Then identify the chord or melody recorded in the file (i.e., list all the sounds included in the chord/melody).
7. Suggested method: truncate the power spectrum graph (but not the spectrum itself) vertically and horizontally, so that only the highest peaks on a double logarithmic scale remain.
8. Then use the data returned by the tones function as a source for tick positions and horizontal axis labels. You can also use the code of the notes function and write a method that takes the frequency and fundamental frequency and returns the closest tone from the equal-tempered scale and possible deviation from it in absolute (Hz) and relative units (either divided by the frequency itself or expressed in percent of interval).
9. Then create a method that identifies the peaks and then turns their central frequencies into semitone names. If you're working on a melody, you need to start with a method that identifies the moments at which successive tones begin - clearly visible on a graph as a function of time.

Notes:

1. Sources from the internet can be used provided that all sources are cited.
2. The instruction 'str = cell2str(cell)' is used to convert a cell object to a string.
3. Measured frequencies may differ slightly from theoretical frequencies because the instrument was not perfectly tuned.
4. The integrity of the researcher requires not listening to the melody file before identification ;-).
5. Each sound of the chord consists of a fundamental tone and a rich collection of overtones (higher tones), the presence of which is due to the fact that the Fourier spectrum of a single sound is not a spike.
6. Overtones have frequencies higher than the fundamental frequency (analysis should start from the lowest frequencies).

7. Among the overtones of a single sound, peaks of higher harmonics of the fundamental tone are extracted against the background of 1/f noise.

8. Higher tones of the chord may coincide with overtones of lower tones - this is the phenomenon of consonance - hence it sounds nice.

9. Sounds belonging to the chord should have higher intensities than the intensities within the overtones. But this is not necessarily a rule.

10. Tip. The chord is in the major key. You can find all major chords on the internet, determine their frequencies for each octave, and compare their theoretical spectrum with the obtained spectrum. For example, you can calculate the correlation of the theoretical spectrum with the experimental spectrum. In practice, it is better to describe the theoretical spectrum as Gaussian curves with a certain width.

11. There are 6 sounds in the chord, higher sounds may appear an octave or several octaves above the fundamental tone or any of the tones of the major triad.

Content and form of the report: electronic form (Word, TeX, HTML, pdf, ps, ...), the report should contain the name of the sample, power spectrum, description of the frequency identification method, and the identified chord: 6 tones or melody written with the names of the semitones. The report should include a complete set of Matlab scripts used or a session recording if someone worked interactively.