



Lab P-6: Synthesis of Sinusoidal Signals—A Music Illusion

Pre-Lab: Read the Pre-Lab and do all the exercises in the Pre-Lab section *prior to attending lab*.

Verification: The Exercise section of each lab should be completed **during your assigned Lab time** and the steps marked *Instructor Verification* signed off **during the lab time**. When you have completed a step that requires verification, demonstrate the result to your instructor and answer any questions about it. Turn in the completed verification sheet before you leave the lab.

Lab Report: The project requires an extensive programming effort and should be documented with a complete lab report. A good report should include the following items: a cover sheet, commented MATLAB code, explanations of your approach, conclusions and any additional tweaks that you implemented for the synthesis. Since the project must be evaluated by listening to the quality of the synthesized sounds, the criteria for judging a good result are given at the end of this lab description.

1 Introduction

This lab includes a project on sound synthesis with sinusoids. The sound synthesis will be done with sinusoidal waveforms of the form

$$x(t) = \sum_k A_k \cos(\omega_k t + \phi_k) \quad (1)$$

where the amplitudes can be manipulated to produce a musical illusion. The challenge of the lab is to adjust the amplitudes to improve the subjective quality for listening.

2 Pre-Lab

In this lab, periodic waveforms and music signals are created with the intention of playing them out through a speaker. Therefore, it is necessary to account for the fact that a conversion is needed from the digital samples, which are numbers stored in the computer memory to the actual voltage waveform that will be amplified and heard through the speakers. In particular, the Nyquist rate condition must be obeyed—samples must be created at a rate that is twice the highest frequency in the synthesized signal (see Chapter 4).

2.1 Piano Keyboard

Section 4 of this lab consists of synthesizing notes in one octave of a musical scale.¹ Since these signals require sinusoidal tones to represent piano notes, a quick introduction to the layout of the piano keyboard is needed. On a piano, the keyboard is divided into octaves—the notes in one octave being twice the frequency of the notes in the next lower octave. The white keys in each octave are named *A* through *G*. In order to define the frequencies of all the keys, one key must be designated as the reference. Usually, the reference note is the *A* above middle-*C*, called *A-440* (or A_4) because its frequency is 440 Hz. (In this lab, we are using the number 40 to represent middle *C*. This is somewhat arbitrary; for instance, the

¹If you have little or no experience reading music, don't be intimidated. Only a little music knowledge is needed to carry out this lab. On the other hand, the experience of working in an application area where you must quickly acquire new knowledge is a valuable one. Many real-world engineering problems have this flavor, especially in signal processing which has such a broad applicability in diverse areas such as geophysics, medicine, radar, speech, etc.

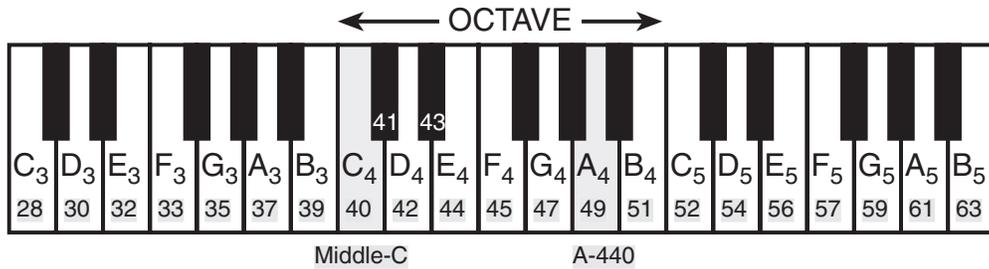


Figure 1: Layout of a piano keyboard. Key numbers are shaded. The notation C_4 means the C-key in the fourth octave.

Musical Instrument Digital Interface (MIDI) standard represents middle C with the number 60). Each octave contains 12 notes (5 black keys and 7 white) and the ratio between the frequencies of the notes is constant between successive notes. As a result, this ratio must be $2^{1/12}$. Since middle C is 9 keys below A-440, its frequency is approximately 261 Hz. Consult the text for even more details.

Musical notation shows which notes are to be played and their relative timing (half, quarter, or eighth). Figure 2 shows how the keys on the piano correspond to notes drawn in musical notation. The white keys are labeled as A , B , C , D , E , F , and G ; but the black keys are denoted with “sharps” or “flats.” A sharp such as A^\sharp is one key number larger than A ; a flat is one key lower, e.g., A^\flat_4 (A-flat) is key number 48.

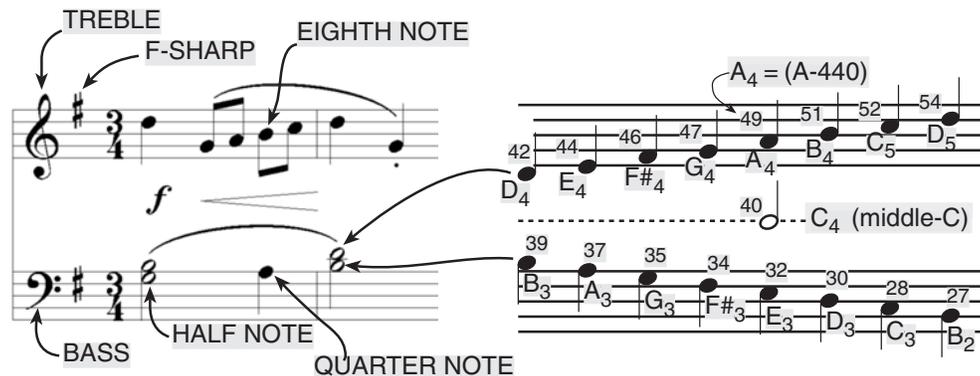


Figure 2: Musical notation is a time-frequency diagram where vertical position indicates which note is to be played. Notice that the shape of the note defines it as a half, quarter or eighth note, which in turn defines the duration of the sound.

Another interesting relationship is the ratio of fifths and fourths as used in a chord. Strictly speaking the fifth note should be 1.5 times the frequency of the base note. For middle-C the fifth is G, but the frequency of G is 391.99 Hz which is not exactly 1.5 times 261.63. It is very close, but the slight detuning introduced by the ratio $2^{1/12}$ gives a better sound to the piano overall. This innovation in tuning is called “equally-tempered” or “well-tempered” and was introduced in Germany in the 1760’s and made famous by J. S. Bach in the “Well Tempered Clavier.”

Thus, you can use the ratio $2^{1/12}$ to calculate the frequency of notes anywhere on the piano keyboard. For example, the E-flat above middle-C (black key number 43) is 6 keys below A-440, so its frequency should be $f_{43} = 440 \times 2^{-6/12} = 440/\sqrt{2} \approx 311$ Hertz.



2.2 Gaussian Forms

The Gaussian form is used in many different fields, but most often in probability where it is called the Gaussian distribution.² In this lab, we are going to use the Gaussian to control the amplitude weighting of sinusoids (called the envelope of the signal). The mathematical form of the Gaussian is

$$g(v) = \alpha e^{-(v-\mu)^2/2\sigma^2}$$

where v is the independent variable. Three parameters (α , μ , σ) are needed to specify the Gaussian. The plot of a Gaussian is the well-known “bell curve” where μ determines the peak location and σ the width of the bell-shaped peak. The plots in Fig. 3 show the general shape for three values of σ . Notice that the *width*

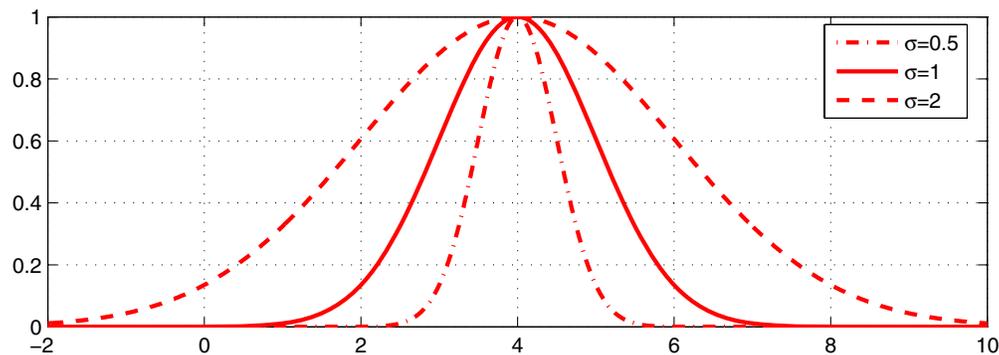


Figure 3: Plot of a Gaussian with parameters $\alpha = 1$, $\mu = 4$, and $\sigma = \frac{1}{2}$, 1, 2. When $\sigma = 2$ the 0.607 level extends from $v = 2$ to $v = 6$ giving a width of 4, i.e., the width is 2σ .

of the bell when measured at the $e^{-1/2} = 0.607$ level is equal to 2σ , which confirms the role of σ as the width-control parameter. In our application the amplitude is scaled by `soundsc` when a sound is played, so we choose $\alpha = 1$. However, for a probability distribution α would be chosen so that the total area under the Gaussian is equal to one, which leads to the value of α being $\alpha = 1/(\sigma\sqrt{2\pi})$.

Write a few lines of MATLAB code to make a plot of $10e^{-(v-1)^2/2(3)^2}$ over the range $-10 \leq v \leq 10$.

²Here are links for more info: https://en.wikipedia.org/wiki/Normal_distribution and <http://mathworld.wolfram.com/NormalDistribution.html>.



3 Warm-up

In the warm-up you will generate a Gaussian and then use its values to weight the amplitude of sinusoids to play a musical note.

3.1 Gaussian Weighting

When used with music synthesis, the Gaussian form must be a function of the log of frequency. Thus we need to define the following *frequency-weighting* function:

$$W(f) = g(\log_2(f)) = e^{-(\log_2(f) - \log_2(f_c))^2 / 2\sigma^2}$$

where f_c is the *center frequency* of the Gaussian where the peak is located. The parameter σ , which is usually called the *variance*, determines the width (2σ) of the Gaussian along the log-frequency axis.

- Write a short M-file that synthesizes a Gaussian given three inputs: center frequency, variance, and a vector of frequencies where the Gaussian is to be evaluated.
- The MATLAB code `ff=2.^(5:1/12:10)` generates frequencies that are uniformly spaced on a logarithmic axis.
- In order to demonstrate that your M-file works, make a plot of a Gaussian that is centered at 440 Hz and has a width equal to one octave. Make the plot over the following frequency range of six octaves: $55 \leq f \leq 1760$ Hz. Use `semilogx` to make the horizontal axis logarithmic, i.e., $\log_2(f)$.
Note: Since the independent variable is $\log_2(f)$, the variance controls width along the $\log_2(f)$ axis. Thus a variance of one corresponds to $\log_2(f/f_c) = 1$ which is a factor of two change in frequency.
- Make two more plots of the Gaussian, but make these plots versus f . The first plot should use MATLAB's `plot` command. Explain why the Gaussian appears distorted. Then use MATLAB's `semilogx` function to make the same plot. Explain why the Gaussian once again has its expected bell shape.

Instructor Verification (separate page)

3.2 Synthesize Octaves with Gaussian Weighting

Create a sound that is the sum of five notes for keys spaced by octaves and amplitude weighted by the Gaussian in the previous part. In a previous lab you completed the `key2note.m` function which synthesizes the correct sinusoidal signal for a particular key number, so you can use that function to do this task.

- Synthesize a signal that is 2 s long, sampled at $f_s = 8000$ Hz. Create five sinusoids and add them together so that they are played simultaneously. The five notes should be middle-C (also called C_4), C_5 , C_6 , C_3 and C_2 . The amplitudes of the five notes should be different; each note's amplitude should be obtained by sampling a Gaussian at the note's frequency. The Gaussian should be centered at $f = 440$ Hz, and have a width of one octave.
- Show a spectrogram for your synthesized sound, and identify all five components. Use a long window length in the spectrogram to get good frequency resolution. Check that there are peaks at the correct frequencies, and also examine the spectrogram image to see that the peaks have different amplitudes.

Instructor Verification (separate page)



4 Lab: A Musical Illusion

The objective in this lab project is to reproduce a musical illusion called the *Shepard Scale* in which the tones played seem to be continually rising forever, yet seem to stay within one octave. If you are familiar with M. C. Escher's drawing of a perpetual staircase³ called "Ascending and Descending," then this audio illusion is analogous. The illusion is referred to as "circularity in pitch judgement" in an article on the web site of the Acoustical Society of America (<http://asa.aip.org/demo27.html>). You can also find a demonstration audio file at the same ASA web site.

The basis of the Shepard scale is to create each "note" as the sum of many individual sinusoids, all separated by octaves. The use of several octaves allows the weighting to make the illusion. For example, the note A-440 might use sinusoids at the frequencies of 220 Hz, 880 Hz, 110 Hz, 1760 Hz, 55 Hz, 3520 Hz. By cleverly controlling the amplitudes of each sinusoid, the illusion is created by playing the amplitude-weighted notes in a scale (or within an octave) over and over. Specifically, the amplitude of the sinusoids would be largest at 440 Hz, a little smaller at 220 and 880 Hz, even smaller at 110 and 1760 Hz, and continually smaller for frequencies farther away from 440 Hz, i.e., the *amplitudes are frequency dependent*.

If we followed the guide of the ASA web site, we would use a cosine amplitude weighting. However, we use a Gaussian instead, because we can get nearly the same amplitude dependence, and we can vary the σ parameter to control the width of the Gaussian. For the Shepard scale, the peak of the Gaussian should be centered somewhere between 260 and 500 Hz, and σ should be fixed to get a fixed Gaussian shape from which the amplitude weights for all notes are obtained. Since the notes on a piano are spaced on a log scale, the Gaussian used for amplitude weighting must also be a function of $\log_2(f)$, instead of f .

The following steps should help you produce a working illusion:

- First of all, forget about the amplitude weighting. Synthesize a C-major scale, starting at middle-C, in which each note in the scale is accompanied by eight other notes separated by octaves, extending four octaves below and four octaves above. For example, to synthesize A-440, you would create the sum of nine sinusoids at the frequencies 27.5, 55, 110, 220, 440, 880, 1760, 3520 and 7040 Hz. Use a sampling frequency of 22050 Hz, so that you will have f_s greater than twice the highest frequency.
- Playing once through a scale requires that you play seven different notes, each note being the sum of nine octave-spaced keys. However, we need to play the scale over and over, so modify your code to play the C-Major scale five times (35 total notes). Adjust the length of the notes and introduce silence between the notes so that you can easily identify each one.
- Now introduce amplitude weighting using a Gaussian form centered somewhere between 260 and 500 Hz. The trick here is that the weighting must be a function of $\log_2(f)$, not f . Make a plot of the weighting function that you are going to use. Make the plot versus $\log_2(f)$ for the case when $\sigma = 2$.
- You should now be able to produce the illusion by using the Gaussian weight function to generate the amplitudes of the nine octave-related sinusoids that are added together for each note. Experiment with the σ parameter to get a good sounding illusion.
- Make a spectrogram of playing the scale three times and explain the illusion from the features that you can identify in the spectrogram. A spectrogram made by `plotspec` displays a linear amplitude scale, and that should make it easier to explain the illusion.
- Variations:* Play every note within the octave; all twelve of them and repeat. Does the illusion sound better in this case?
- Listening:* Prepare your MATLAB code so that it will be easy to demonstrate the illusion to your instructor at the beginning of the next lab.

³The idea of the perpetual staircase is attributed to Roger Penrose and his father Lionel (1958).



Lab: Shepard Scale

Instructor Verification Sheet

For each verification, be prepared to explain your answer and respond to other related questions that your lab instructor might ask. Turn this page in at the end of your lab period.

Name: _____

Date of Lab: _____

Part 3.1 Complete and demonstrate the function that synthesizes a Gaussian form versus $\log_2(f)$.

Verified: _____

Date/Time: _____

Part 3.2 Complete and demonstrate the synthesise of five Gaussian weighted sinusoids. Show a spectrogram and identify the frequencies and amplitudes of the five spectral components.

Verified: _____

Date/Time: _____

Sound Evaluation Criteria

Does the file play the correct notes?

All Notes _____ Most _____ Missing octaves _____ Wrong Amplitude Weighting _____

Overall Impression: _____

Good: Correct illusion; proper use of Gaussian envelope; all notes synthesized and in sync.

OK: Basic sinusoidal synthesis, but the illusion is not obvious

Poor: Synthesis does not work.; incorrect illusion.