Lecture 7

Audio frequencies

Introduction

Audio is one of the most common application areas of DSP, and we are using audio DSP as the central theme of this course. When developing DSP algorithms it is important to understand the characteristics of the signal to be processed. Speech and music are the most common types of audio signals. Therefore this lecture is devoted to a brief discussion of audio frequencies.

As we go through this lecture it can be helpful to generate and listen to a sine wave of a particular frequency with the following Scilab commands

```
F = 440;
tmax = 4;
Fs = 44100;
t = [0:1/Fs:tmax];
sound(sin(2*%pi*F*t),Fs,16);
```

Here `Fs` is the sampling frequency and `F` is the sine wave frequency (both in Hz). The signal lasts `tmax` seconds.

Human hearing

The “standard” range of human hearing is from 20 Hz to 20 kHz [1]. Actually, the ability to hear different frequencies varies among individuals. In particular, as people age they tend to lose their ability to hear high-frequency sounds. Many animals can hear frequencies beyond the human range. Cats can hear frequencies above 75 kHz. The moth *galleria mellonella* is sensitive to frequencies up to 300 kHz [2]. The frequency ranges of a few musical instruments and the human voice are shown in Fig. 1.

“High-fidelity” recordings necessarily cover the entire range of (ideal) human hearing. Traditional analog telephone systems transmitted frequencies between about 300 Hz and 3.4 kHz which is sometimes called the “voice band.” AM radio signals cover roughly this same band. Not surprisingly, “talk radio” and news are popular on AM stations while music tends to be limited to FM stations (which transmit higher bandwidth signals, usually in stereo).

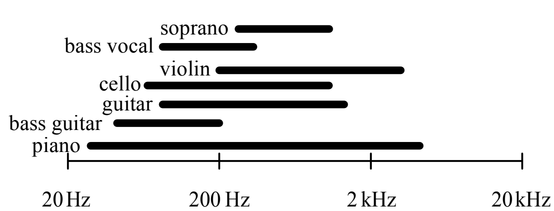

Fig. 1: Frequency ranges of various musical instruments and human voice. Frequencies are fundamental tones. Instruments also generate higher-frequency harmonics. Note logarithmic scale.

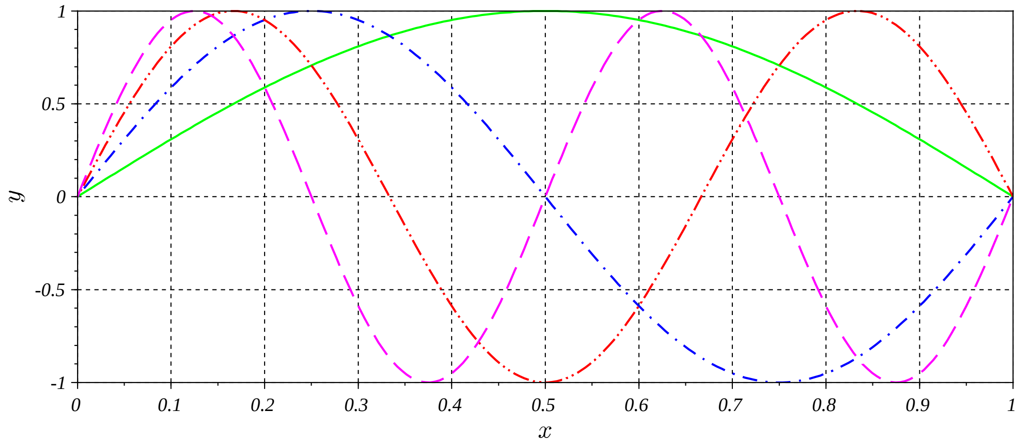

Fig. 2: A resonant structure can oscillate at frequencies which are integer multiples of a fundamental tone.

Resonant structures, speech and music

The human voice is the original musical instrument. Although we probably don't think of our everyday speech as music, speech and music have much in common since they are generated using similar principles.

Harmonics

Non-electronic musical instruments (including the human voice) typically generate notes using a *resonant structure* such as a string or air cavity. If L is the length of the structure and v is the speed of sound in that structure then the time it takes a sound wave to travel "round-trip" through the structure is

$$t_{rt} = \frac{2L}{c} \tag{1}$$

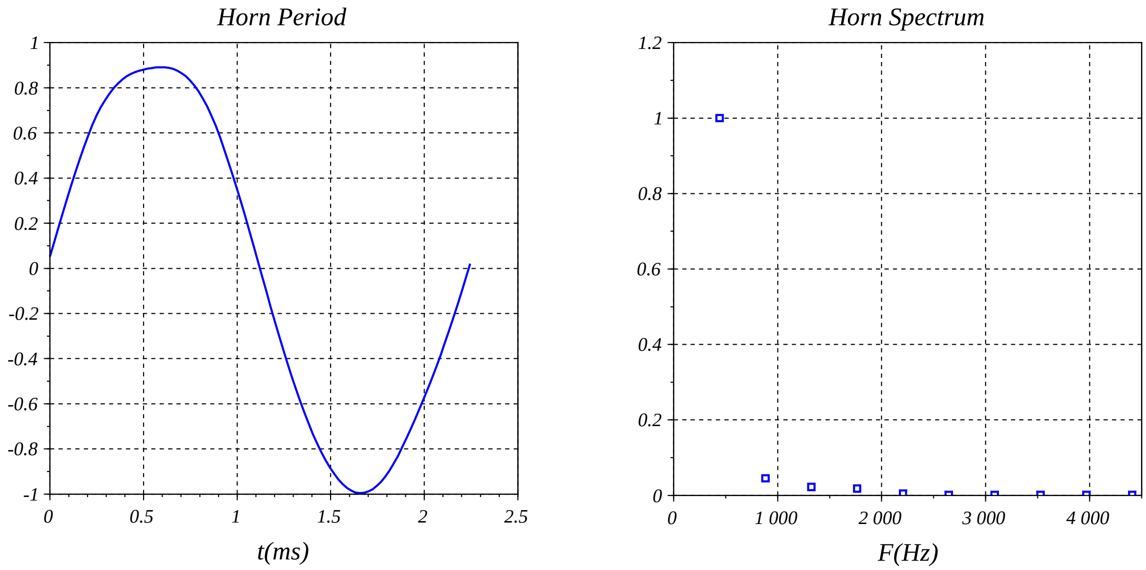

Fig. 3: (Left) one period of a horn ("French horn") playing 440-Hz note. (Right) spectral amplitudes.

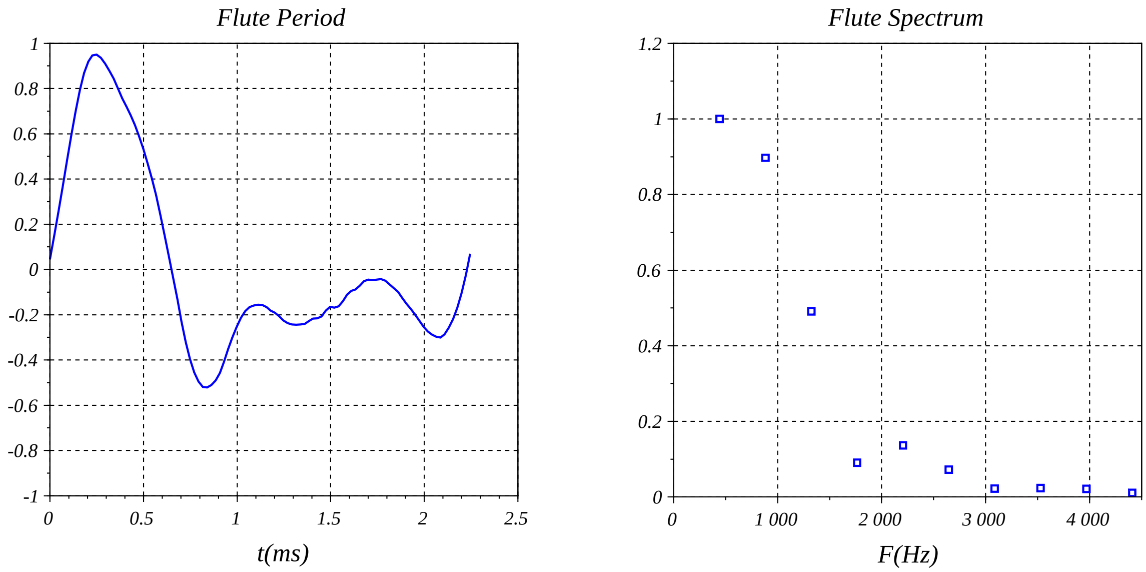

Fig. 4: (Left) one period of a flute playing 440-Hz note. (Right) spectral amplitudes.

The inverse of this time is the *fundamental frequency* of the structure. It is the frequency at which one oscillation corresponds to one round-trip through the structure. Integer multiples of the fundamental frequency

$$f_n = n\frac{v}{2L} \tag{2}$$

for $n=1,2,3,\ldots$ are its resonant frequencies or *harmonics* (Fig. 2). The structure can oscillate at one or more of these frequencies simultaneously. The first frequency f_1 is the fundamental while the others are called *overtones*. The overtone frequencies are integer multiples of the fundamental frequency.

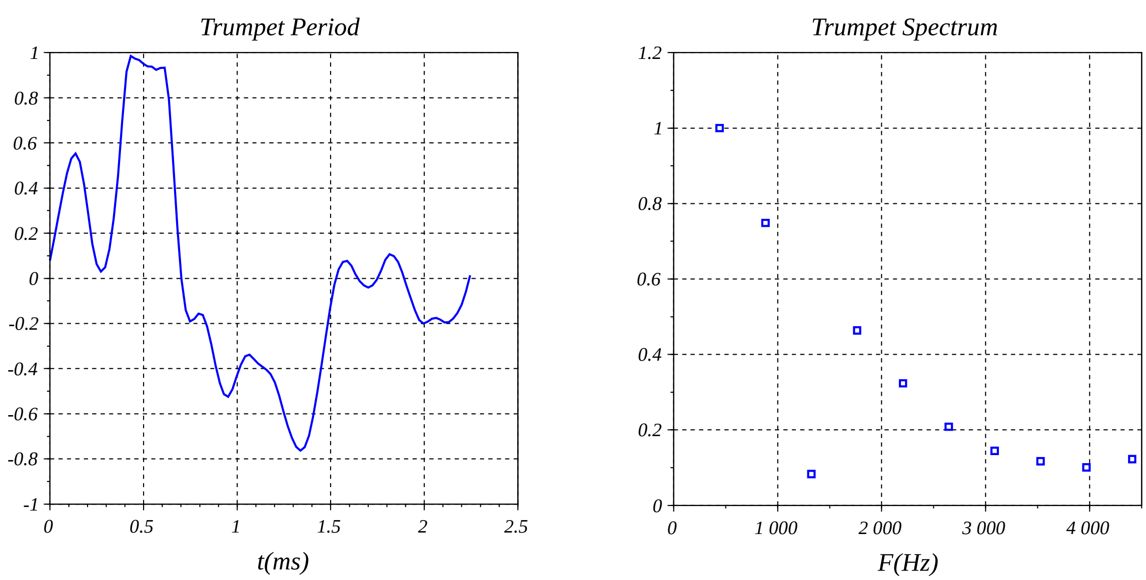

Fig. 5: (Left) one period of a trumpet playing 440-Hz note. (Right) spectral amplitudes.

A periodic signal with fundamental frequency F_1 can be represented by the Fourier series

$$s(t)=\sum_{n=1}^{\infty} A_n \cos(2\pi n F_1 t+\theta_n) \tag{3}$$

Figures 3, 4, and 5 show one period of, respectively, a horn, flute, and trumpet playing a note of fundamental frequency 440 Hz. Also shown are the corresponding spectra which are the harmonic amplitudes A_n at frequencies 440, 880, 1320, 1760, etc. Hz. The horn is very "pure" with almost all of the signal in the fundamental. With the flute and trumpet we see more signal in higher harmonics. This is perceived as the trumpet being "sharper" and the horn "smoother" with the flute in-between. To hear these sounds play the file `hornFluteTrumpet.mp3`.

Musical notes

Out of the continuous range of audible frequencies, music focuses on a discrete set that taken together have certain aesthetic properties. The most common standard for musical note frequencies is *twelve-tone equal temperament* based on a reference frequency of 440 Hz. In this system note frequencies are indexed by an integer. The n^{th} frequency is

$$f_n = 440\cdot\alpha^n \text{ Hz}$$

where

$$\alpha = 2^{\frac{1}{12}} = 1.059463\ldots$$

and n can take on negative, zero or positive values. Adjacent notes differ in frequency by about 6%. The Appendix provides a brief description of the logic behind this standard.

Negative indices are inconvenient for computer applications. The *Musical Instrument Digital Interface* (MIDI) specifies frequencies as

$$f_m = 440\cdot\alpha^{(m-69)} \text{ Hz}$$

where $0\le m\le 127$ indexes frequencies over the range

$$f_0 = 8.18 \text{ Hz}, f_{127} = 12{,}543.85 \text{ Hz}$$

On a standard 88-key piano (Fig. 6) the lowest note is $m=21$ and the highest note is $m=108$.

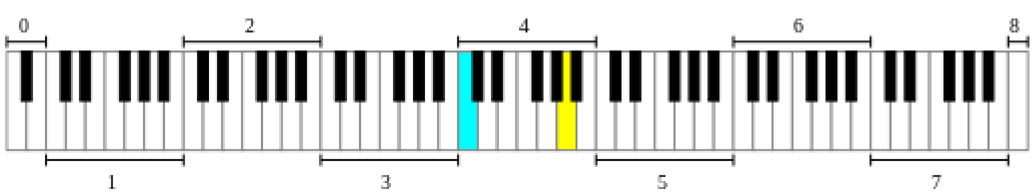

Fig. 6: 88-key piano keyboard. Numbers refer to standard octaves (see Table 1). Yellow key is A-440 (A4) $(n=0, m=69)$. *Cyan key is "middle C" (C4)* $(n=-9, m=60)$.

References

1. https://en.wikipedia.org/wiki/Hearing_range
2. https://www.nature.com/news/moth-smashes-ultrasound-hearing-records-1.12941
3. https://en.wikipedia.org/wiki/Scientific_pitch_notation
4. https://en.wikipedia.org/wiki/Equal_temperament

Appendix – list of musical notes, frequencies and MIDI numbers

	Octave						
Note	1	2	3	4	5	6	7
C	32.70 (24)	65.41 (36)	130.81 (48)	261.63 (60)	523.25 (72)	1046.50 (84)	2093.00 (96)
C#	34.65 (25)	69.30 (37)	138.59 (49)	277.18 (61)	554.37 (73)	1108.73 (85)	2217.46 (97)
D	36.71 (26)	73.42 (38)	146.83 (50)	293.66 (62)	587.33 (74)	1174.66 (86)	2349.32 (98)
D#	38.89 (27)	77.78 (39)	155.56 (51)	311.13 (63)	622.25 (75)	1244.51 (87)	2489.02 (99)
E	41.20 (28)	82.41 (40)	164.81 (52)	329.63 (64)	659.26 (76)	1318.51 (88)	2637.02 (100)
F	43.65 (29)	87.31 (41)	174.61 (53)	349.23 (65)	698.46 (77)	1396.91 (89)	2793.83 (101)
F#	46.25 (30)	92.50 (42)	185.00 (54)	369.99 (66)	739.99 (78)	1479.98 (90)	2959.96 (102)
G	49.00 (31)	98.00 (43)	196.00 (55)	392.00 (67)	783.99 (79)	1567.98 (91)	3135.96 (103)
G#	51.91 (32)	103.83 (44)	207.65 (56)	415.30 (68)	830.61 (80)	1661.22 (92)	3322.44 (104)
A	55.00 (33)	110.00 (45)	220.00 (57)	440.00 (69)	880.00 (81)	1760.00 (93)	3520.00 (105)
A#	58.27 (34)	116.54 (46)	233.08 (58)	466.16 (70)	932.33 (82)	1864.66 (94)	3729.31 (106)
B	61.74 (35)	123.47 (47)	246.94 (59)	493.88 (71)	987.77 (83)	1975.53 (95)	3951.07 (107)

Table 1: Frequency in Hz and MIDI index (m) of musical notes.

Appendix – chords, scales and equal temperament

Chords and scales

The development of music has been strongly influenced by the perception of *consonance vs. dissonance* – which sounds "go together" vs. which do not. While this is ultimately subjective, there is broad-enough consensus to allow a theory of music to be developed. Below we briefly describe concepts used in the European musical tradition.

The following Scilab code plays two notes, one at $F=261.63\,\text{Hz}$ (C4, "middle C") and a second at $F=261.63\cdot r\,\text{Hz}$ where *r* is the ratio of the frequencies. The two notes form a musical *interval*.

```
F = 261.63;
r = 3/2;
tmax = 4;
Fs = 44100;
t = [0:1/Fs:tmax];
sound(0.5*(sin(2*%pi*F*t)+sin(2*%pi*r*F*t)),Fs,16);
```

Certain intervals are perceived to be more consonant than others, and likewise for certain combinations of intervals which form musical *chords*. Halving or doubling the frequency of a note is perceived to generate a lower or higher version of the same note. The two notes are said to be in *unison*, and we say the second note is an *octave down* or an *octave up* from the original. The octave is the "most consonant" interval (Run the above code with `r = 2` and `r = 1/2` to hear this.) Since any note can be moved up or down an octave (r multiplied by 2 or ½), we can limit consideration to intervals between "notes in a single octave" $(1 \le r < 2)$ by multiplying any r value outside this interval by 2 or ½ as needed.

When an interval is the ratio of small integers the waveform has a short period, and most listeners perceive it to be consonant. Other intervals have a long period (infinite if r is irrational) and are perceived to be dissonant (Fig. 7). After the octave, the next most consonant interval is $r=3/2$ which is said to form a *perfect fifth* (for reasons we will see below). After the perfect fifth, the next most consonant interval is $r=4/3$, called a *perfect fourth*.

The "base note" $r=1$ is called the *tonic*, the $r=3/2$ note is called the *dominant*, and the $r=4/3$ note is called the *subdominant*. Lowering $r=4/3$ by an octave to $r=2/3$ we have

$$\begin{matrix} \frac{2}{3} & \frac{1}{1} & \frac{3}{2} \\ \text{subdominant} & \text{tonic} & \text{dominant} \end{matrix}$$

These notes and *chords* built upon them play a central role in (most) "tonal" music.

Following the pattern $\frac{3}{2}, \frac{4}{3}$ we come to the interval $\frac{5}{4}$, called a *major third*. Play the file `intervals.mp3` to hear the following sequence of intervals

$$\left[\frac{1}{1}, \frac{3}{2}\right] ; \left[\frac{1}{1}, \frac{4}{3}\right] ; \left[\frac{1}{1}, \frac{5}{4}\right]$$

To most people this sounds musically pleasing and seems to move toward a final "resolution."

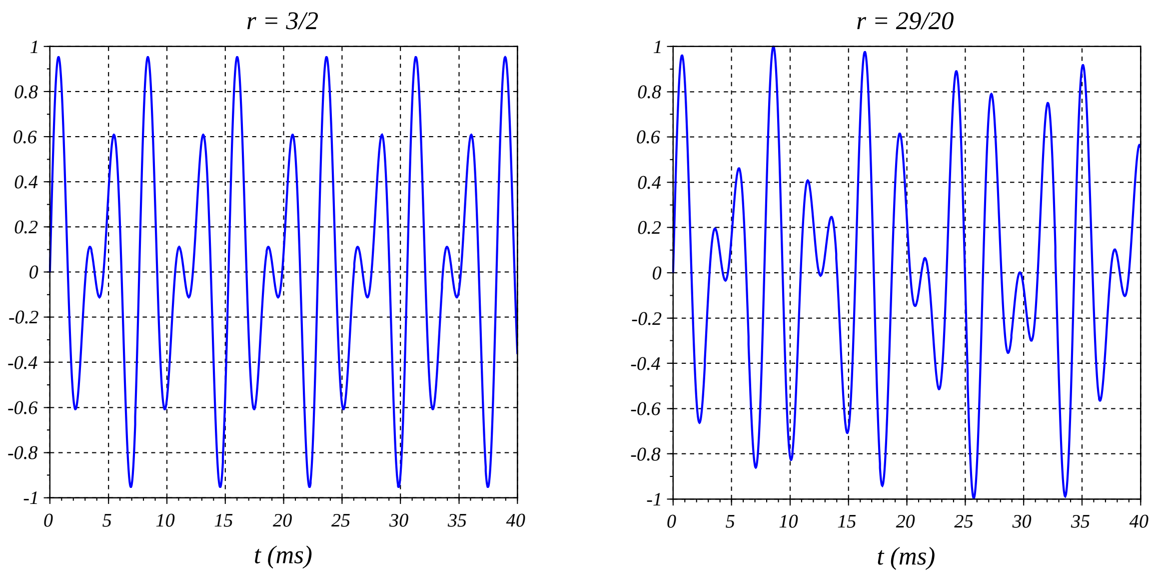

Fig. 7: Sum of sine waves of frequencies 261.63 Hz and 261.63·r Hz with (left) r = 3/2 and (right) r = 29/20. The first case leads to a signal with short period (2 cycles of the lower frequency) while the second case leads to a signal with long period (20 cycles of lower frequency). To most listeners the first interval sounds consonant ("in tune") while the second sounds dissonant ("out of tune").

Combining the major-third and perfect-fifth intervals produces a *major chord*

$$\left[\frac{1}{1},\frac{5}{4},\frac{3}{2}\right]$$

Building major chords on the subdominant, tonic, and dominant tones we have

$$\underset{\text{subdominant}}{\frac{2}{3}\left[\frac{1}{1},\frac{5}{4},\frac{3}{2}\right]}\quad\underset{\text{tonic}}{\left[\frac{1}{1},\frac{5}{4},\frac{3}{2}\right]}\quad\underset{\text{dominant}}{\frac{3}{2}\left[\frac{1}{1},\frac{5}{4},\frac{3}{2}\right]}\rightarrow\underset{\text{subdominant}}{\left[\frac{4}{3},\frac{5}{3},\frac{1}{1}\right]}\quad\underset{\text{tonic}}{\left[\frac{1}{1},\frac{5}{4},\frac{3}{2}\right]}\quad\underset{\text{dominant}}{\left[\frac{3}{2},\frac{15}{8},\frac{9}{8}\right]}$$

where at right we have moved all notes into our single octave $(1\le r< 2)$. Play `progression.mp3` to hear the *chord progression*: tonic, subdominant, dominant, tonic. (For the last tonic chord we move the tonic note 1/1 up an octave to 2/1 to give the progression a sense of "moving up.") Countless folk and popular songs are built from these three chords alone.

Arranging the notes of the tonic, subdominant and dominant chords in order we obtain the *C major scale*

$$\begin{array}{ccccccc}\frac{1}{1} & \frac{9}{8} & \frac{5}{4} & \frac{4}{3} & \frac{3}{2} & \frac{5}{3} & \frac{15}{8}\\ \text{C} & \text{D} & \text{E} & \text{F} & \text{G} & \text{A} & \text{B}\end{array}$$

Play `majorScaleJust.mp3` to hear these notes (we add a final C an octave up $(r=2/1)$ so the scale has a sense of "completion"). These correspond to the "white keys" on the piano (Fig. 8) and are denoted by the first seven letters of the alphabet, although not in alphabetical order: C, D, E, F, G, A, B. They are also known as: do, re, me, fa, so, la, ti. An entire keyboard repeats this scale over several octaves (Fig. 6). The interval $r=3/2$ is called a "fifth" because G is the fifth note of the C-major scale; likewise $r=4/3$ is called a "fourth," and $r=5/4$ is called a "third." This method of fixing scale tones as the ratios of small integers is called *just intonation*.

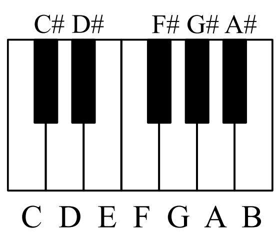

Fig. 8: One octave of a piano keyboard.

An obvious question is why the letters used to signify the "white-key" major scale notes start at C and not at A. This is the result of historical developments that go back to (at least) the 10th century. If we play these same tones starting with note A we have the *A-minor scale*

$$\begin{array}{ccccccc}\frac{5}{6} & \frac{15}{16} & \frac{1}{1} & \frac{9}{8} & \frac{5}{4} & \frac{4}{3} & \frac{3}{2}\\ \text{A} & \text{B} & \text{C} & \text{D} & \text{E} & \text{F} & \text{G}\end{array}$$

Here we've shifted notes A and B down an octave so that the scale moves up monotonically. Play `minorScaleJust.mp3` to hear these notes (we add a final A an octave up so the scale has a sense of "completion"). Major scales are typically perceived as being "happy" while minor scales are perceived as being "somber." The third interval (from A to C) defines a ratio

$$\frac{1/1}{5/6}=\frac{6}{5}$$

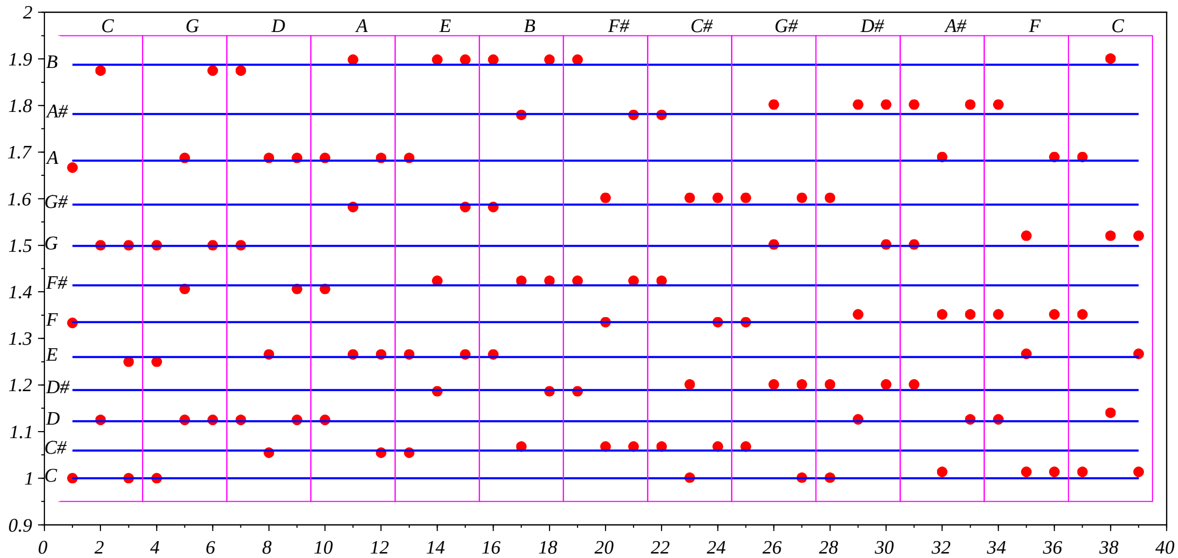

Fig. 9: Subdominant-dominant-tonic progressions following the circle of fifths with just intonation. Vertical axis is ratios of chord tones to C4. Dots are note values using just intonation. Horizontal lines are note values with equal temperament. With just intonation the process does not "close on itself." Instead, the C-major scale tones we start with at left end up slightly higher in frequency at right.

The old tonic is the new subdominant while the old dominant is the new tonic. We say that we have *modulated* from the key of C to the key of G.

The new dominant chord ("D major")

$$\left[\frac{9}{8}, \frac{45}{32}, \frac{27}{16}\right]$$

has a note (45/32) not found in the C-major scale. This falls between notes F and G:

$$\frac{4}{3}=1.33333\ldots \;,\; \frac{45}{32}=1.40625 \;,\; \frac{3}{2}=1.5$$

We call it "F#" and assign it to a black key on the piano (Fig. 8). Seemingly $27/16$ is also a new note. However, it is very close to the existing note "A"

$$\frac{27}{16}=1.6875\approx\frac{5}{3}=1.666\ldots$$

The ratio of this "new note" to A is only

$$\frac{27/16}{5/3}=\frac{81}{80}=1.0125$$

whereas the ratio of F# to F is

$$\frac{45/32}{4/3}=\frac{135}{128}=1.0546875$$

For instruments that can produce a continuous range of note frequencies, such as voice and fretless string instruments (violin, etc.), the existence of closely-spaced notes is not a problem.

But for instruments that use keys or frets to produce a discrete set of frequencies, such as piano, flute, or guitar, it is. Do we add a special key for 27/16, resulting in two keys that sound almost the same, or do we call 5/3 “close enough?”

We can continue to change keys, moving our progression up a perfect fifth each time (multiplying through by 3/2 and moving all notes into our single octave $0 \le r < 1$ by multiplying by 2 or ½ as needed). The resulting frequencies are shown as dots in Fig. 9. Play `circleFifthsJust.mp3` to hear these 39 chords. This sequence of changing keys, C-G-D-etc., is known as the *circle of fifths*.

After twelve of these modulations we *almost* return to the key of C. And, this procedure *almost* defines twelve distinct notes, the seven white keys and five black keys on the piano, and twelve associated “major keys.” But, in fact, each note appears at slightly different frequencies during this process. Compare the first three chords in the block labeled “C” to the last three chords. The original C-major chord

$$\left[1, \frac{5}{4}, \frac{3}{2}\right] = [1, 1.25, 1.5]$$

ends up as

$$\left[\frac{531441}{524288}, \frac{2657205}{2097152}, \frac{1594323}{1048576}\right] \approx [1.0136, 1.2671, 1.5205]$$

This is almost the same, but noticeably different when played in succession. So with just intonation the circle of fifths does not “close on itself,” and each of the twelve notes takes on slightly different frequencies in different musical keys.

A solution is to give up requiring intervals to be rational numbers and instead require all intervals to be an integer power of the constant

$$\alpha = 2^{\frac{1}{12}}$$

The twelve notes of a single octave are then α^n, $n = 0,1,2,3,4,5,6,7,8,9,10,11$. This system is called *twelve-tone equal temperament*. Since $\alpha^{-12} = 1/2$ and $\alpha^{12} = 2$, decreasing or increasing the exponent of α by 12 moves the note down or up an octave. The difference in the major scale frequency ratios between just intonation and equal temperament is quite small (Fig. 10) which is why this strategy “works.”

$$\begin{array}{ccccccccccccccc} 1 & \frac{9}{8} & \frac{5}{4} & \frac{4}{3} & \frac{3}{2} & \frac{5}{3} & \frac{15}{8} & = & 1 & 1.125 & 1.250 & 1.333 & 1.500 & 1.667 & 1.875 \\ \alpha^0 & \alpha^2 & \alpha^4 & \alpha^5 & \alpha^7 & \alpha^9 & \alpha^{11} & = & 1 & 1.122 & 1.260 & 1.335 & 1.498 & 1.682 & 1.888 \end{array}$$

Fig. 10: Major scale notes with (top) just intonation and (bottom) equal temperament.

The intervals of a major chord are now

$$[\alpha^0, \alpha^4, \alpha^7]$$

with α^7 corresponding to the perfect fifth interval. Moving all tones up a perfect fifth we have

$$\alpha^7[\alpha^0, \alpha^4, \alpha^7] = [\alpha^7, \alpha^{11}, \alpha^{14}] \rightarrow [\alpha^7, \alpha^{11}, \alpha^2]$$

We will always end up with the same twelve notes, so the circle of fifths now “closes” on itself.

Play `chordJustVsEqual.mp3` to hear the difference of a major chord in just intonation and equal temperament. There is a slight, but noticeable, difference.