ELEC S347F Multimedia Technologies,
Spring Term, 2021
Tutorial 06 Solution

Question 1: Given that the threshold of hearing (TOH) is at 10^{-12} W/m^2 and the threshold of pain for hearing is at 1 W/m^2.

(a) Determine the threshold of pain in terms of decibel.

(b) Find the sound intensity for conversation of 60 dB.

(c) Suppose whisper produces a sound of 20 dB. How many times more intense is the sound of a conversation if it produces a sound of 60 dB?

(a) $10 \log_{10}(1/10^{-12}) = 10 \log_{10}(10^{12}) = 10(12) = 120$ dB.

(b) $60 \text{ dB} = 10 \log_{10}(P/10^{-12})$

$\Rightarrow 60 = 10 \log_{10}(P) - 10 \log_{10}(10^{-12})$

$\Rightarrow 60 = 10 \log_{10}(P) + 120$

$\Rightarrow 10 \log_{10}(P) = -60$

$\Rightarrow \log_{10}(P) = -6$

$\Rightarrow P = 10^{-6}$ W/m^2.

(c) 60 dB – 20 dB = 40 dB = 10^4 times more intense.

Alternative Solution:

60 dB => P = 10^{-6} W/m^2

20 dB => P = 10^{-10} W/m^2

$10^{-6} / 10^{-10} = 10^4$

Question 2: Given that the sampling rate for Super Audio Compact Disc (SACD) format is 64 times of a standard CD format.

(a) What is the sampling period of SACD format?

(b) What is the range of frequencies that can be represented in the recordings in SACD?

(a) Sampling frequency = 44.1 kHz x 64 = 2822.4 kHz = 2.8224 MHz
Sampling period = 1/2.8224 MHz = 0.3543 us (or 354.3 ns)

(b) According to Nyquist's sampling theorem, the max. freq. component = 2.822 MHz / 2 = 1.411 MHz. Therefore, the range of frequencies is 0 Hz to 1.411 MHz

Question 3: Given the psychoacoustic model that signal-to-mask ratios for bands 1, 2 and 3 are for signals above 80 dB in band 2 a masking of 30 dB in band 1 and 40 dB in band 3. The first 3 levels of the critical bands of the audio is listed below.

(d) Assume the 32 critical bands are divided into equal-width and the human audible frequencies are ranged from 20 Hz to 20,000 Hz. Find the range of frequencies of Band 1, 2 and 3 respectively.

(e) Determine which band of data can be removed by frequency masking.

Band	1	2	3
Level (dB)	20	90	55

(a) The width of each band (20000 – 20)/32 = 19980/32 = 624.375

Band 1: [20, 644.375), Band 2: [644.375, 1268.75), Band 3: [1268.75, 1893.125)

(b) For band 1: 20 dB < 30 dB so ignore it.

For band 3: 55 dB >= 40 dB so send it.

Question 4: Given the sequence of the Middle/Side channels of a MP3 audio as follows:

Middle:	35	10	54	38	27	37	27
Side:	1	0	2	2	-3	-3	-1

(a) Find the sequence of the left channel of the above sequence.

(b) Find the sequence of the right channel of the above sequence.

Middle: L + R, Side: L – R

L = (M + S)/2, R = (M – S)/2

(a) Left:	18	5	28	20	12	17	13
(b) Right:	17	5	26	18	15	20	14

Question 5: A typical audio compact disc (CD) can record 74 minutes of CD quality music.

(a) Determine the data rate of CD quality audio and telephony quality speech, respectively.

(b) Determine the no. of bytes of data that can be stored in the CD.

(c) Determine the no. of hours of telephony quality speech can be recorded on the same CD.

(d) Determine the no. of hours of MP3 music of typical data rate at 128 kbps that can be recorded on the same CD.

Given: CD quality audio: 44.1 kHz, 16 bit, Stereo

Telephony quality speech: 8 kHz, 8 bit, mono

(a) CD quality audio: 44.1x1000x16x2 = 1,411,200 bps (1.4112 Mbps)

Telephony quality speech: 8x1000x8x1 = 64,000 bps (64 kbps)

(b) 74x60x1,411,200 = 6,265,728,000 bits = 783,216,000 bytes (or 746.9 MB)

(c) 6,265,728,000/64,000 = 97902 seconds = 1631.7 minutes = 27.2 hours

Alternative solution:

74 mins x 1,411,200 / 64,000 = 1631.7 minutes = 27.2 hours

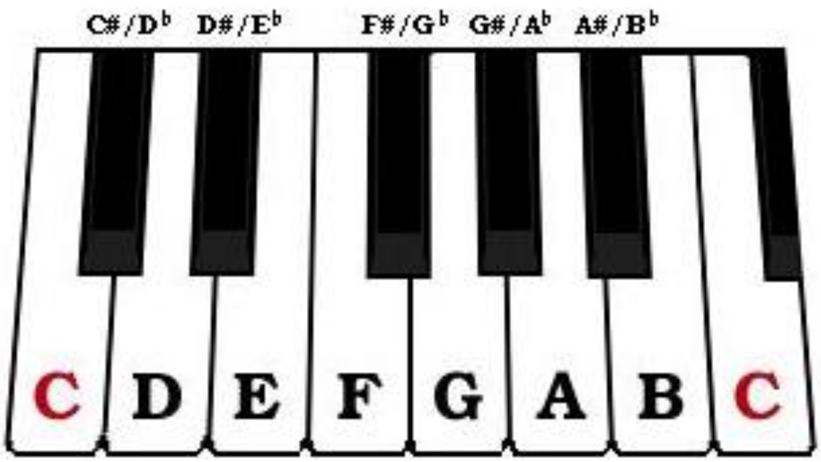
(d) 6,265,728,000 / 128,000 = 48,951 seconds = 815.85 minutes = 13.6 hours
Note: typical compression rate of MP3 is about 10:1 to 12:1.
74 mins x 11 = 814 mins. Roughly equals to the answer obtained in part (c).

Comment: both (b) and (c) use lossy method to compress the audio signals.
Telephony quality speech is done by a low-pass filter which filter out signal with frequency greater than 4 kHz.
MP3 music is done by removing the inaudible data and quantization further introduce information loss.

Question 6: The difference in frequency between consecutive keys on a piano is not linearly spaced, but logarithmically spaced.

(a) Given that the range of a piano is seven octaves plus three semitones. If the lowest note played by a piano is approximately 27.5 Hz, what is the highest frequency the piano can generate?
(b) Chords are a set of notes (usually three) that sounds good when they are played together. A major chord consists of three notes played together: a note X, note X + 4 semitones, and note X + 7 semitones. If note X is A3 (220 Hz), determine the name and calculate the frequency of the remaining two notes that will form a major chord for note X, respectively.
(c) Consider a note W of frequency 493.88 Hz. What is the name of note W?

[C#3 D#3 F#3 G#3 A#3] C#4 (black keys)
[Db3 Eb3 Gb3 Ab3 Bb3] Db4 (Alternative name)



[C3 D3 E3 F3 G3 A3 B3] C4 (white keys)

(a) 7 octaves plus 3 semitones = 7 x 12 + 3 = 87 semitones
The highest frequency = $27.5 \times 2^{87/12} = 4186$ Hz

(b) Y = X + 4 semitones = C#4

$220 \times 2^{4/12} = 277.18$

Z = X + 7 semitones = E4

$220 \times 2^{7/12} = 329.63$

(c) $220 \times 2^{k/12} = 493.88$

$\Rightarrow k = \log_2(493.88/220) \times 12$

$\Rightarrow k = 14$

There are 14 semitones in between A3 and W.

An octave consists of 12 semitones. 14 semitones = 1 Octave + 2 semitones

Therefore, W is B4.

Question 7: Given the MIDI message format (in binary) as follows:

Message	Status Byte	Data Byte 1	Data Byte 2
Note Off	1000 CCCC	0NNN NNNN	0VVV VVVV
Note On	1001 CCCC	0NNN NNNN	0VVV VVVV
Panning	1011 CCCC	0000 1010	0SSS SSSS
Change Program	1100 CCCC	0PPP PPPP	-
System Reset	1111 1111	-	-

CCCC: Channel Number NNNNNNN: Note Number

VVVVVVV: Velocity PPPPPPP: Program Number

SSSSSSS: Stereo Position (0: absolute left, 64_{10}: center, 127_{10}: absolute right)

The General MIDI instrument patch map (partial):

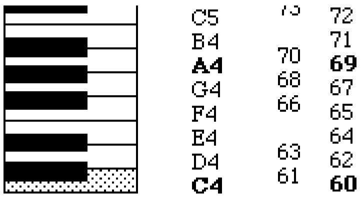
Program No.	Program	Program No.	Program
1	Acoustic Grand	5	Rhodes Piano
2	Bright Acoustic	6	Chorused Piano
3	Electric Grand	7	Harpsichord
4	Honky-Tonk	8	Clavinet

The MIDI velocity and dynamic mapping table:

Velocity (in dec)	Dynamic	Velocity (in dec)	Dynamic
8	***pppp***	64	***mf***
20	***ppp***	80	***f***
31	***pp***	96	***ff***
42	***p***	112	***fff***
53	***mp***	127	***ffff***

The MIDI Note Number table:

ELEC S347F Multimedia Technologies,
Spring Term, 2021
Tutorial 06 Solution



Consider a MIDI file contains the following corresponding actions:

1. Reset the MIDI device
2. Change the instrument on Channel 2 to Electric Grand Piano
3. Pan the stereo position of Channel 2 to absolute left
4. Play the F4 note using *ffff* loudness on Channel 2
5. Turn off the same note

(a) Determine the MIDI message (in hex) for each of the actions.

(b) Refer to the above MIDI encoding structure, what are the limitations of MIDI?

(a) Note: Channel and Program are 1-index.

Channel 2 is coded as 0001 (0000: channel 1, 0001: channel 2, .., 1111: channel 16)

Program 3 is coded as 000 0010 (000 0000: program 1, 000 0001: program 2, .., 111 1111: program 128)

1. System Reset: 1111 1111 (FF)
2. Change Prog.: 1100 0001 0000 0010 (C1 02)
3. Pan to Left: 1011 0001 0000 1010 0000 0000 (B1 0A 00)
4. Note On: 1001 0001 0100 0001 0111 1111 (91 41 7F)
5. Note Off: 1000 0001 0100 0001 0000 0000 (81 41 00)

(Note: velocity is typically 0 for note off message)

(b) Limited no. of channels (limited the no. of musical instruments played simultaneously)

Limited no. of program (limited the no. of musical instruments available)

Limited resolution in data values (e.g. limited the no. of velocity (volume) levels of instrument)