University of Pennsylvania
Department of Electrical and System Engineering
Digital Audio Basics

ESE1500, Spring 2023 | Midterm | Wednesday, March 1

- Exam ends at 12:55PM; begin as instructed (target 12:05PM)
- Do not open exam until instructed to begin exam.
- Problems weighted as shown.
- Calculators allowed.
- Closed book = No text or notes allowed.
- Provided reference materials on next to last page.
- Show work for partial credit consideration.
- Unless otherwise noted, answers to two significant figures are sufficient.
- Sign Code of Academic Integrity statement (see last page for code).

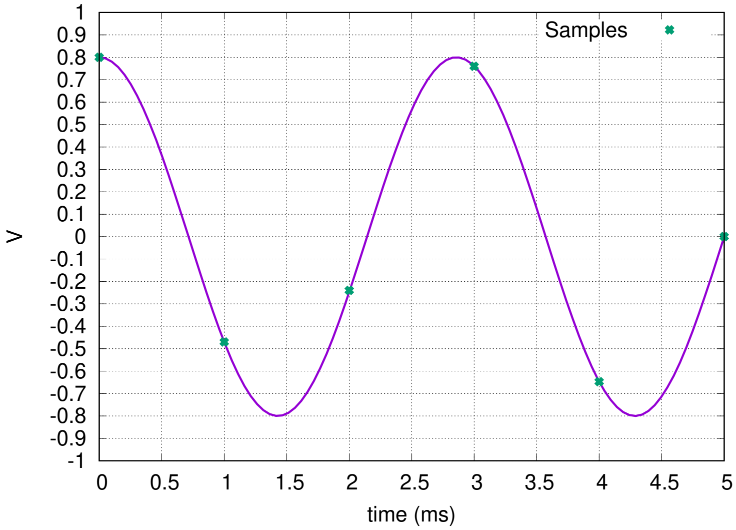
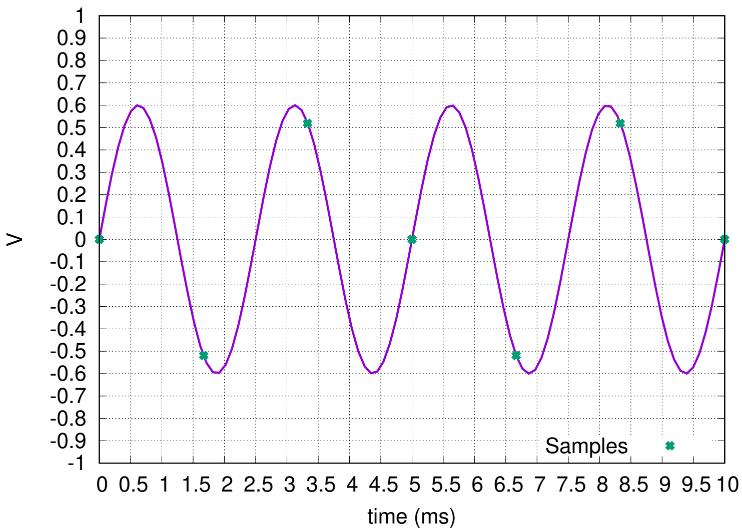
____

I certify that I have complied with the University of Pennsylvania's Code of Academic Integrity in completing this exam.

Name: Solution

1	2				3					4				5
	a	b	c	d	a	b	c	d	e	a	b	c	d	
16	4	4	10	4	2	4	5	5	4	5	5	5	7	20

1. Consider the following sampling cases, complete the table entries.

Waveform and Sample	Waveform Frequency (Hz)	Sample Frequency (Hz)	Inferred Frequency (Hz) from Samples	Properly Sampled?
	350	1000	350	Yes
	400	600	200	No

2. Assume most human speech occurs under 5.5 KHz, you can get reasonable recovery quantized to 256 amplitude levels, and your primary concern is to preserve and reconstruct human speech.

   (a) What is the lowest sampling frequency you can use?

   11 KHz – Nyquist sampling rate at twice the maximum frequency in the signal

   (b) Starting from a CD Quality PCM recording (44KHz sampling, 16b amplitudes):

      i. While still using a PCM scheme, by what factor can you compress the representation of speech? (CD-PCM/your-PCM)?

      factor of 8 – 4x from dropping 44KHz to 11KHz sampling rate, 2x from dropping 16b amplitude to 8b amplitude

      ii. What kind of compression is this?

      lossy – we are discarding both high frequencies and resolution that we cannot recover.

   (c) Assuming there are no aliasing artifacts in the original CD-quality PCM, how would you re-encode the sound to achieve your identified compression without introducing new aliasing artifacts?

   We need to remove the frequencies above 5.5 KHz to make sure we don't introduce aliasing. Simple keeping every fourth time sample would result in aliasing of signals between 5.5 KHz and 22 KHz. So, instead we must convert to the frequency domain using a Fourier Transform. In the frequency domain, we then drop all the frequencies above 5.5KHz and then convert back to the time domain (inverse Fourier Transform) with the 11 KHz sample rate, saving only 8b amplitudes for each time point.

   (d) Using a 25 ms window, this sample rate can identify 550 frequencies. Assuming there are on average 2 frequencies to encode per human cirtical band and you can use 8b amplitude per frequency, by what additional factor can you typically compress the speech representation (your-PCM/your-frequency-based)?

   11 KHz PCM encoding is $11,000 \times 8 = 88{,}000$ b/s

   20 critical bands (bands 21–24 are above 5.5 KHz)

   2 frequencies per band

   18b for (freq,amplitude pair) $= \log_2(550)+8$

   $20 \times 2 \times 18/0.025 = 28{,}800$ b/s

   Ratio $\frac{88000}{28800} = 3.055 \approx 3.1$

3. Consider the following quote from Benjamin Franklin:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
w	h	e	n		y	o	u	'	r	e		f	i	n	i	s	h	e	d		c	h	a

24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
n	g	i	n	g	,		y	o	u	'	r	e		f	i	n	i	s	h	e	d

This has 46 symbols from a set of 18 unique symbols.

symbol		'	,	a	c	d	e	f	g	h	i	n	o	r	s	u	w	y	**sum**
count	5	2	1	1	1	2	5	2	2	4	5	5	2	2	2	2	1	2	46

(a) How many bits to encode this using a uniform encoding where each symbol is encoded using the same number of bits?

$46 \times \lceil \log_2(18) \rceil = 46 \times 5 = 230$

(b) What property or properties of this quote make it amenable to lossless compression?

Symbols occur with different frequencies. Some symbols occur more often than others; we can give short encodings to the more frequently occuring symbols.

(c) What is the Shannon Entropy lower bound for encoding this entire quote?

$$\text{Lower Bound} = -\sum_i \log_2(P(c[i])) \qquad (1)$$

Hint: there are only 4 different counts, so 4 different P(c[i]) values to calculate.

$4 \times 5 \times \log_2(5/46) + 1 \times 4 \times \log_2(4/46) + 9 \times 2 \times \log_2(2/46) + 4 \times 1 \times \log_2(1/46) = 181.6 \approx 182$

(d) Consider the following set of variable-length binary encodings. Assign each symbol to an encoding to minimize the encoded length.

encode	100	011	010	001	000	1010	11111	11110	11101
symbol	i	space	n	e	h	'	u	f	r
encode	11100	11001	11000	10111	10110	110111	110110	110101	110100
symbol	o	y	d	s	g	c	w	,	a

(e) For the above assignment, how many bits are required to encode the quote?

$4 \times 5 \times 3 + 1 \times 4 \times 3 + 1 \times 2 \times 4 + 8 \times 2 \times 5 + 4 \times 1 \times 6 = 184$

(This page left mostly blank for pagination. Feel free to perform calculations here.)

4. Consider a collection of zoo sounds. There are tracks of sounds of 20 animals. Assume each animal makes sounds dominated by a different set of frequencies; different animals may share some frequencies. The sounds can be assembled to create a zoo soundscape by combining some subset of the 20 animal sound tracks. Assume the original samples are 44KHz-sampled audio, and you want to produce combined soundscapes that are also 44KHz-samples of audio. Assume frequency-encodings below operate on 25 ms time windows and only store the present frequencies in each time window.

   (a) How large is a PCM-encoded, composite (multiple-animal) soundscape compared to a PCM-encoded single animal track? (give ratio composite-soundscape/single-animal-track)

   1 – the PCM encoding uses all the time samples and is not impacted by the number of present frequencies.

   (b) Which is likely larger, a PCM-encoded, single-animal track or a frequency-encoded single-animal track and why?

   PCM-encoded single-animal track will likely be larger; the single-animal track will only need to store the present frequencies; the present frequencies for an animal are likely small compared to the entire frequency range, so will take much less space than the time samples.

   (c) What is the upper-bound on size of a frequency-encoded, 5-animal composite soundscape compared to the size of a frequency-encoded, single-animal track? (give ratio 5-animal-composite-soundscape/single-animal-track)

   5× – Worst case, we add all the frequencies together. If there are 5× as many frequencies to represent, it will take 5× the number of bits. Strictly 5× the size of the largest single-animal soundtrack; and this assumes that the number of frequencies in an animal soundtrack is generally small (less than 1/5th) the total number of frequencies this sampling can possibly represent.

   (d) Assuming you are producing MP3-quality audio and exploiting human psychoacoustics, what may make the frequency-encoded, multi-animal composite soundscape smaller than the upper-bound estimate?

   Masking – some of the frequencies between animals may be the same or may be close enough that one animal's sound masks another at a particular frequency. As a result, a psychoacoustic-encoded representation (including an MP3) will retain fewer than 5× the number of frequencies in the single animal tracks.

5. Given: $f(t) = 0.5\cos(2\pi \cdot 400t) + \sin(2\pi \cdot 1300t)$
give the first 5 time-sample values of $f(t)$ for a 12 KHz sample rate.

sample	value
0	0.50
1	1.1
2	1.4
3	1.3
4	0.74

(This page left mostly blank for pagination and workspace.)

Human auditory critical bands:

Band Number	Low	High
1	20	100
2	100	200
3	200	300
4	300	400
5	400	510
6	510	630
7	630	720
8	720	920
9	920	1080
10	1080	1370
11	1270	1480
12	1480	1720
13	1720	2000
14	2000	2320
15	2320	2700
16	2700	3150
17	3150	3700
18	3700	4400
19	4400	5300
20	5300	6400
21	6400	7700
22	7700	9500
23	9500	12000
24	12000	15500

Code of Academic Integrity

Since the University is an academic community, its fundamental purpose is the pursuit of knowledge. Essential to the success of this educational mission is a commitment to the principles of academic integrity. Every member of the University community is responsible for upholding the highest standards of honesty at all times. Students, as members of the community, are also responsible for adhering to the principles and spirit of the following Code of Academic Integrity.*

Academic Dishonesty Definitions

Activities that have the effect or intention of interfering with education, pursuit of knowledge, or fair evaluation of a student's performance are prohibited. Examples of such activities include but are not limited to the following definitions:

A. Cheating Using or attempting to use unauthorized assistance, material, or study aids in examinations or other academic work or preventing, or attempting to prevent, another from using authorized assistance, material, or study aids. Example: using a cheat sheet in a quiz or exam, altering a graded exam and resubmitting it for a better grade, etc.

B. Plagiarism Using the ideas, data, or language of another without specific or proper acknowledgment. Example: copying another person's paper, article, or computer work and submitting it for an assignment, cloning someone else's ideas without attribution, failing to use quotation marks where appropriate, etc.

C. Fabrication Submitting contrived or altered information in any academic exercise. Example: making up data for an experiment, fudging data, citing nonexistent articles, contriving sources, etc.

D. Multiple Submissions Multiple submissions: submitting, without prior permission, any work submitted to fulfill another academic requirement.

E. Misrepresentation of academic records Misrepresentation of academic records: misrepresenting or tampering with or attempting to tamper with any portion of a student's transcripts or academic record, either before or after coming to the University of Pennsylvania. Example: forging a change of grade slip, tampering with computer records, falsifying academic information on one's resume, etc.

F. Facilitating Academic Dishonesty Knowingly helping or attempting to help another violate any provision of the Code. Example: working together on a take-home exam, etc.

G. Unfair Advantage Attempting to gain unauthorized advantage over fellow students in an academic exercise. Example: gaining or providing unauthorized access to examination materials, obstructing or interfering with another student's efforts in an academic exercise, lying about a need for an extension for an exam or paper, continuing to write even when time is up during an exam, destroying or keeping library materials for one's own use., etc.

* If a student is unsure whether his action(s) constitute a violation of the Code of Academic Integrity, then it is that student's responsibility to consult with the instructor to clarify any ambiguities.