Ollscoil Mhá Nuad
Maynooth University

SEMESTER 2
2023-2024

CS425
Audio & Speech Processing

Dr. C. Gurrin, Prof. R.J. Farrell, Mr. D. Balasubramaniam

Time allowed: 2 hours

Answer at least ***three*** questions

Your mark will be based on your best ***three*** answers

All questions carry equal marks

Instructions

	Yes	No	N/A
Formulae and Tables book allowed *(i.e. available on request)*	X		
Formulae and Tables book required *(i.e. distributed prior to exam commencing)*		X	
Statistics Tables and Formulae allowed *(i.e. available on request)*		X	
Statistics Tables and Formulae required *(i.e. distributed prior to exam commencing)*		X	
Dictionary allowed *(supplied by the student)*		X	
Non-programmable calculator allowed	X		
Students required to write in and return the exam question paper		X	

[25 marks]

1 (a) Explain Shannon's sampling theorem. Use a diagram to illustrate sampling a simple sinusoid with different sampling rate. Given that CD-quality audio employs a sampling rate of 44,100 Hz, discuss how this relates to the characteristics of human hearing. [6 marks]

(b) A signal contains frequency components up to 5 kHz. What is the minimum sampling rate required to accurately digitize this signal according to the Nyquist theorem? Describe what could happen if the signal is sampled at 4 kHz instead. [6 marks]

(c) Compare the number of quantization levels and signal-to-noise ratios between 8-bit and 16-bit systems. Given the expression, [6 marks]

$$SNR_{dB} = 6.02_{dB} * N_{bits} + 1.76_{dB}$$

calculate the difference in signal-to-noise ratio for these two systems. Explain how this relates to dynamic range and perceived volume after digital-to-analog conversion.

(d) Even if the Shannon sampling theorem is strictly observed, an anti-aliasing filter is typically used before sampling. Explain why this is necessary. [2 marks]

(e) Say whether the following statements are true or false: [5 marks]

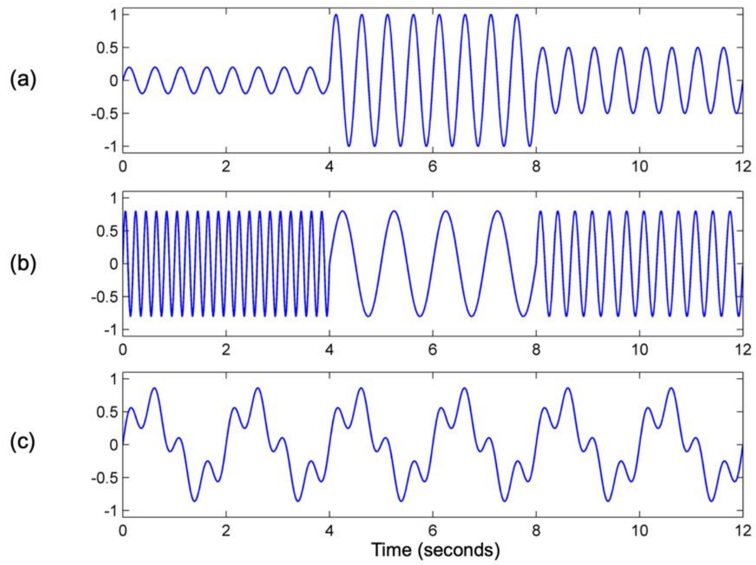
i. Sampling at exactly twice the highest frequency component of a signal always prevents aliasing.
ii. Nyquist frequency is half the sampling rate.
iii. A higher bit depth in digital audio reduces the dynamic range.
iv. Increase in sampling rate decreases the audio file size.
v. WAV is a lossless audio file format.

[25 marks]

2 (a) Given a sampling frequency of 8000Hz and an analog signal that contains 4 sinusoids at 1800Hz, 3300Hz, 4700Hz and 6210Hz, compute the frequencies at which the sinusoids components will appear in the sampled signal. [5 marks]

(b) [10 marks]

i. Sketch the magnitude Fourier transform of the following signals.
ii. Sketch the spectrogram of the following signals. Assume a window length that corresponds to a duration of about one second.



(c) What is the formula to calculate the magnitude and phase of a given complex frequency coefficient? [3 marks]
Find the magnitude and phase (in degree) for the following.
   i. 3+4i
   ii. 2+2i

(d) Describe the concepts of windowing, hopping, and padding in the context of the Short-Time Fourier Transform (STFT) and their importance in analyzing non-stationary signals. [7 marks]

[25 marks]

3 (a) Discuss the importance of using complex exponentials in the analysis of signals through the Fourier Transform. How do complex exponentials facilitate the representation of both amplitude and phase information? [5 marks]

(b) Provide the first five sine waves of the Fourier series to approximate a square wave signal. The fundamental being $\sin(2\pi x)$ [3 marks]

(c) Given the signal x[n]={1,3,5,-4,2} and the impulse response h[n]={1,−3,2}, compute the convolution y[n]=x[n]∗h[n] manually, showing each step of your calculation. [7 marks]

(d) Write a pseudocode for the Fast Fourier Transform algorithm, assuming the length of the input signal is a power of 2. Include brief comments to explain each step. [10 marks]

[25 marks]

4 (a) What is a Cepstrum and Explain MFCC and why is it suitable for speech. [8 marks]

(b) What is Pitch, Formants and Fricatives of a speech signal? explain them with examples. [6 marks]

(c) Describe the limitations of the linear speech model. [6 marks]

(d) Explain Hidden Markov Model and its use in speech processing [5 marks]