Electronics and Electrical Communications Engineering Department

Faculty of Engineering

Cairo University

Submitted to: Dr. Mohsen Rashwan

Speech Processing Problems

DSP-1 Assignment 1 submitted for course ELC4011 “DSP-1 Applications“

4th Year

1st Semester - Academic Year 2025/2026

Prepared by:

NAME	SECTION	ID
Yousef Khaled Omar Mahmoud	4	9220984

Submission Date: 20 November 2025

1 Table of Contents

2 Table of Figures

3 List of tables

DSP-1 Assignment 1: Speech Processing Problems

Youssef Khaled Omar Mahmoud*1

*Electronics and Communication Department, Faculty of Engineering ,

Cairo University Giza, 12613, Egypt

yousef.mahmoud03@eng-st.cu.edu.eg

Abstract: ***Since***

Key words: ***).***

1 Introduction

.ss

A. Audio File Information

As shown in Table 1, the analysis in this report, particularly the practical implementation of windowing (Problem 1) and the context for LPC (Problem 2 & 3), is based on the following source audio file:

Table 1 Audio File Information[2]

Parameter	Value
File	..\Data\Test\C02n_1.wav
Sampling Frequency F_s	16000 Hz
Bits Per Sample (Bit Depth)	16 bits
Number of Channels	1
Total Samples	12800
Calculated Bit Rate	256000 bits/s

2 Analyze the Frequency Domain Characteristics of Windows

This section presents the time-domain definitions and frequency-domain analysis for the Rectangular, Hanning, and Hamming window functions, which are fundamental tools in digital signal processing for frame-based analysis.

A. Window Definitions

Given a window length N, the three specified window functions are defined in the time domain as follows:[1]

1) Rectangular Window W_{Rec} [1]

$$W_{\mathrm{Rec}}(n) = \begin{cases} 1 \ \ for\ 0 \le n \le N-1 \\ \quad 0\ Otherwise \end{cases}$$

2) Hanning Window W_{Han} [1]

$$W_{Han}(n) = 0.5 - 0.5\cos\left(\frac{2\pi n}{N}\right)\ \text{for}\ 0 \le n \le N-1$$

3) Hamming Window W_{Ham} [1]

$$W_{Ham}(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N}\right)\ \text{for}\ 0 \le n \le N-1$$

B. Time Domain Visualization

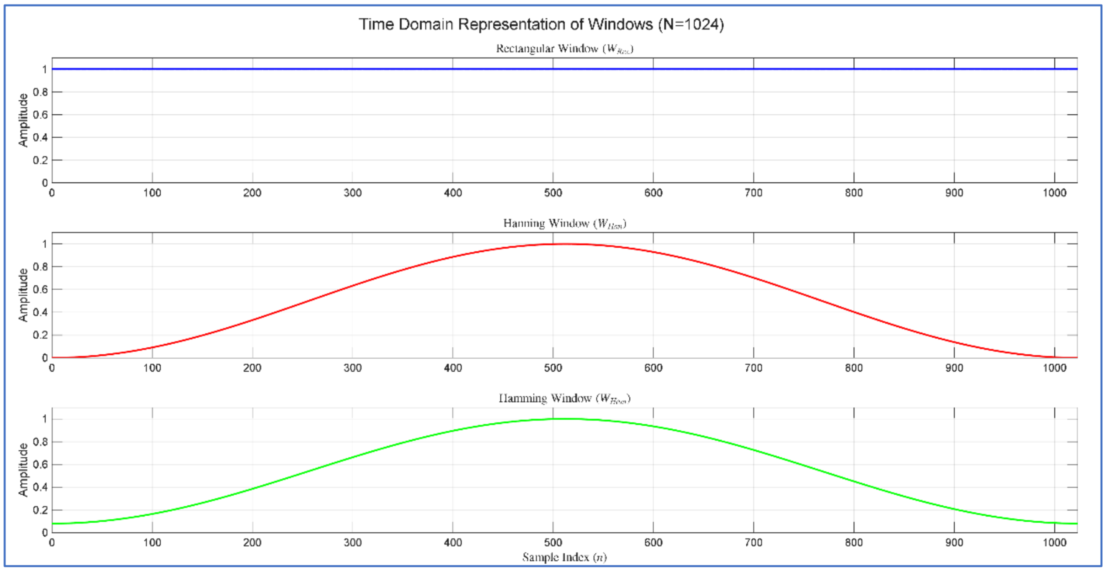

Figure 1 : Time Domain Representation of Rectangular, Hanning, and Hamming Windows (N=1024)[2]

As shown in Figure 1, the time domain representations of the three window functions for $N = 1024$ are visualized below. This plot demonstrates how the windows taper the input signal.

C. Frequency Domain Visualization

The frequency response, $W(f)$, is obtained by computing the Discrete-Time Fourier Transform (DTFT) of the time-domain window, typically approximated using the Fast Fourier Transform (FFT). The plots below show the log-magnitude spectrum, $20\log_{10}(|W(f)|)$, normalized such that the maximum peak is 0 dB.

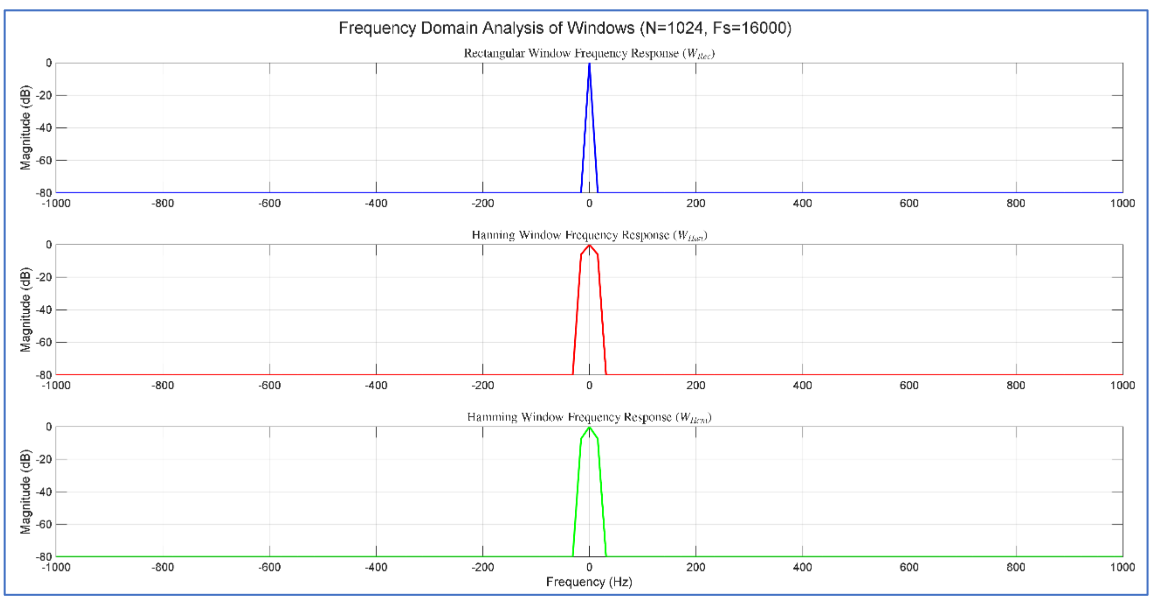

Figure 2 : Frequency Domain Analysis of Windows (Log Magnitude)[2]

As shown in Figure 2, the plots are centered around 0 Hz and zoomed to the range $[-\mathrm{F}_s/16,\ \ _s/16]$ H Z to clearly observe the main lobe and the first few side lobes. The magnitude floor is clipped at -80 dB for visualization clarity.

D. Discussion and Comments

As shown in Table 2, the primary function of a window is to mitigate **spectral leakage** by smoothly forcing the signal to zero at the frame boundaries. The analysis of the windows in the frequency domain reveals the critical trade-off between **frequency resolution** (main lobe width) and **spectral leakage** (side lobe attenuation).

Table 2 Comparison of Frequency Domain Characteristics for Rectangular, Hanning, and Hamming Windows

Characteristic	Rectangular (W_{Rec})	Hanning (W_{Han})	Hamming (W_{Ham})
Main Lobe Width	Narrowest (Highest Resolution)	Medium	Medium
Peak Side Lobe	Highest (≈ -13 dB)	Moderate (≈ -31 dB)	Lowest (≈ -41 dB)
Leakage/Attenuation	Worst Spectral Leakage	Good Attenuation	Best First Side Lobe Attenuation
Application	Analyzing transient, short bursts	General purpose, good compromise	Better for minimizing interference from strong nearby tones

Detailed Comments:

1) Rectangular Window (W_{Rec}):

- **Main Lobe:** Exhibits the narrowest main lobe, theoretically offering the best frequency resolution.
- **Side Lobes:** Its sudden truncation in the time domain results in the highest side lobes (only $\approx$ 13 dB down from the peak). This poor attenuation means energy from a strong frequency component "leaks" significantly into adjacent frequency bins, leading to severe spectral leakage.

2) Hanning Window (W_{Han}):

- **Tapering:** Provides a smooth taper to zero, which significantly reduces side lobe levels.
- **Side Lobes:** The peak side lobe is suppressed to about -31 dB This is a substantial improvement over the rectangular window.
- **Trade-off:** This suppression comes at the cost of a wider main lobe (approximately twice the width of the rectangular window), which slightly reduces frequency resolution.

3) Hamming Window (W_{Ham}):

- **Design:** The Hamming window is a modification of the Hanning window, designed specifically to minimize the height of the *first* side lobe.
- **Side Lobes:** Achieves the best first side lobe attenuation, suppressed to about -41 dB.
- **Trade-off:** While the first side lobe is the lowest, subsequent side lobes roll off more slowly than those of the Hanning window.

Conclusion: For applications like speech processing, where minimizing spectral leakage is crucial to accurately separating closely spaced harmonics and formants, the **Hanning** and **Hamming** windows are strongly preferred over the Rectangular window. The choice between Hanning and Hamming depends on whether a faster side lobe decay (Hanning) or the lowest possible first side lobe (Hamming) is desired.

3 Linear Predictive Coding Analysis based on Given Autocorrelation Data

E. Methodology: Solving the Yule-Walker Equations

Problem 2 requires the determination of the Linear Predictive Coding (LPC) coefficients, a_1 and a_2, and the minimum mean-squared prediction error, E_2, given the first three autocorrelation values of a signal, as discussed in the course material on **Speech Analysis** [3].

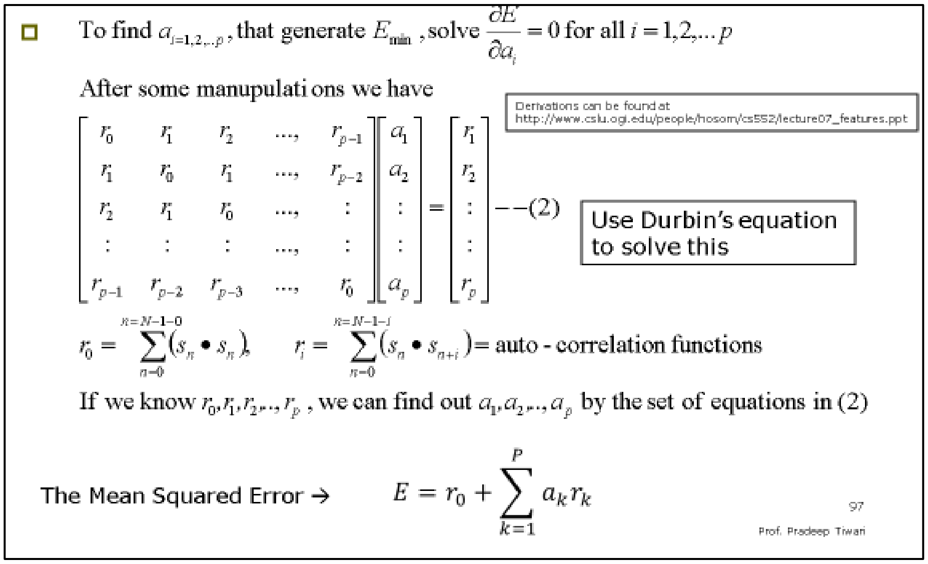

Figure 3 LPC Calculations [3]

As shown in Figure 3, The solution is found by solving the **Yule-Walker system of equations**, which relates the autocorrelation values **R** to the LPC coefficients **a** and the minimum mean-squared error E_p. For a prediction order p, the matrix equation is:

$$\mathbf{R_p a} = \mathbf{r}_\mathrm{p}$$

where:

- $\mathbf{R_p}$ is the $p \times p$ **Toeplitz matrix** containing autocorrelation values $R(|i-j|)$.
- $\mathbf{a} = [a_1, a_2, \dots, a_p]^T$ is the vector of LPC coef
- $\mathbf{r_p} = [R(1), R(2), \dots, R(p)]^T$ is the vector of shifted autocorrelation values.

The minimum mean-squared error E_p is calculated as:

$$E_p = R(0) + \sum_{k=1}^{p} a_k R(k)$$

F. Input Data and Results

The given autocorrelation data for a prediction order $p = 2$ is:

$$\mathbf{R} = [R(0), R(1), R(2)] = [1.0000, 0.7000, 0.4000]$$

Substituting these values into the $p = 2$ Yule-Walker matrix equation:

$$\begin{bmatrix} 1.0 & 0.7 \\ 0.7 & 1.0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.4 \end{bmatrix}$$

As shown in Table 3, Solving this linear system using the implementation (e.g., using matrix inversion or the Levinson-Durbin algorithm) yielded the following results:

Table 3 Problem 2 LPC Results [2]

Parameter	Value
Prediction Order p	2
LPC Coefficient a_1	0.8235
LPC Coefficient a_2	-0.1765
Minimum Mean-Squared Error E_2	1.5059

4 CONCLUSIONS

Due to the suc

BIOGRAPHY

Youssef Khaled Omar Mahmoud

Youssef Khaled is an undergraduate student in Electrical and Computer Engineering at the Faculty of Engineering, Cairo University, Egypt. He specializes in embedded systems, IoT, and intelligent control. Youssef has led research and development projects such as *AquaVision*, an AI-driven smart aquaculture system, and has hands-on experience with STM32, ESP32, C/C++, Python, and sensor–actuator integration. He has also contributed to robotics education and embedded system teams through IEEE CUSB and CUERT.

5 REFERENCES

[1] S. K. Mitra, "Digital Signal Processing: A Computer-Based Approach," 4th ed., McGraw-Hill, 2011.
[2] "DSP-Speech-Processing Assignment," GitHub Repository. [Online]. Available: https://github.com/youefkh05/DSP_Speech_Processing
[3] Dr. Mohsen Rashwan, "Lec3&4 Speech Analysis (DSP-1 Applications)," Cairo University, [Lecture Slides], 2025.

المقدمة باللغة العربية