GROUPWORK

Coursework Declaration and Feedback Form

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Course Code: | **ENG 5027** | Course Name: | **Digital Signal Processing** | | |
| Course Co- Ordinator: | **Scott Watson** | Mentor: |  | Group Number Lab Group Tutorial Group | **36** |
| Title of Assignment: | **Assignment 1 (Fourier Transform – FFT)** | | | | |
| Date of Submission: | **20-10-2024** | | | | |
| ***Declaration of Originality and Submission Information*** | | *I affirm that this submission is my own / the groups original work in accordance with the University of Glasgow Regulations and the School of Engineering*  *Requirements*  ***All students should sign this form*** | | | |
| Student Name: | **Basav Prasad (21961731)** | Student Name: | **Cem Cetinkaya (2961798)** | Student Name: | **Mingxuan Sun (3025864)** |
| Student Name: | **Syafiq Fathullah (2830739)** | Student Name: | **Kabir Sani Muhammad (2973267)** | Student Name: | **Abdulrazaq Muhammad Sani (2962438)** |
|  | | | | | |
| Feedback from Lecturer to Student – to be completed by Lecturer or Demonstrator | | | | | |
| Grade Awarded:  Feedback (as appropriate to the coursework which was assessed) | | | | | |
| Lecturer/Demonstrator | | Date returned to the Teaching Office | | | |

****

**Assignment 1 (Fourier Transform – FFT)**

October 20, 2024

A assignment report submitted in partial fulfilment of the requirements for the credit of course -

Digital Signal Processing

**Introduction**

This report presents the analysis and enhancement of an audio signal using Python. The main focus was to examine the signal in both the time and frequency domains, identify key features such as fundamental frequencies, harmonics, and noise bands, and enhance the audio by improving vocal quality. The script was implemented to make the audio perceptually more pleasant, with clearer and more vibrant vocal content.

**Objectives**

1. **Load and Plot the Original Audio Signal**: Visualize the audio signal in both the time and frequency domains with proper labeling and logarithmic axes for the frequency plot.
2. **Identify Key Features**: Identify fundamental frequencies, harmonics, and noise bands in the audio spectrum.
3. **Improve Voice Quality**: Enhance the quality of the voice by manipulating the frequency bands above 3 kHz.
4. **Enhance the Voice**: Apply an aural exciter to enhance the vocal clarity and richness.

**Methods**

**Task 1: Loading and Plotting the Original Audio**

The audio file, “orig.wav”, was loaded using the Python wave module and converted into a NumPy array. The signal was normalized to the range –1 to 1, and a time axis was generated for plotting.

* **Time Domain Plot**: The normalized audio signal was plotted over time to visually observe amplitude variations.

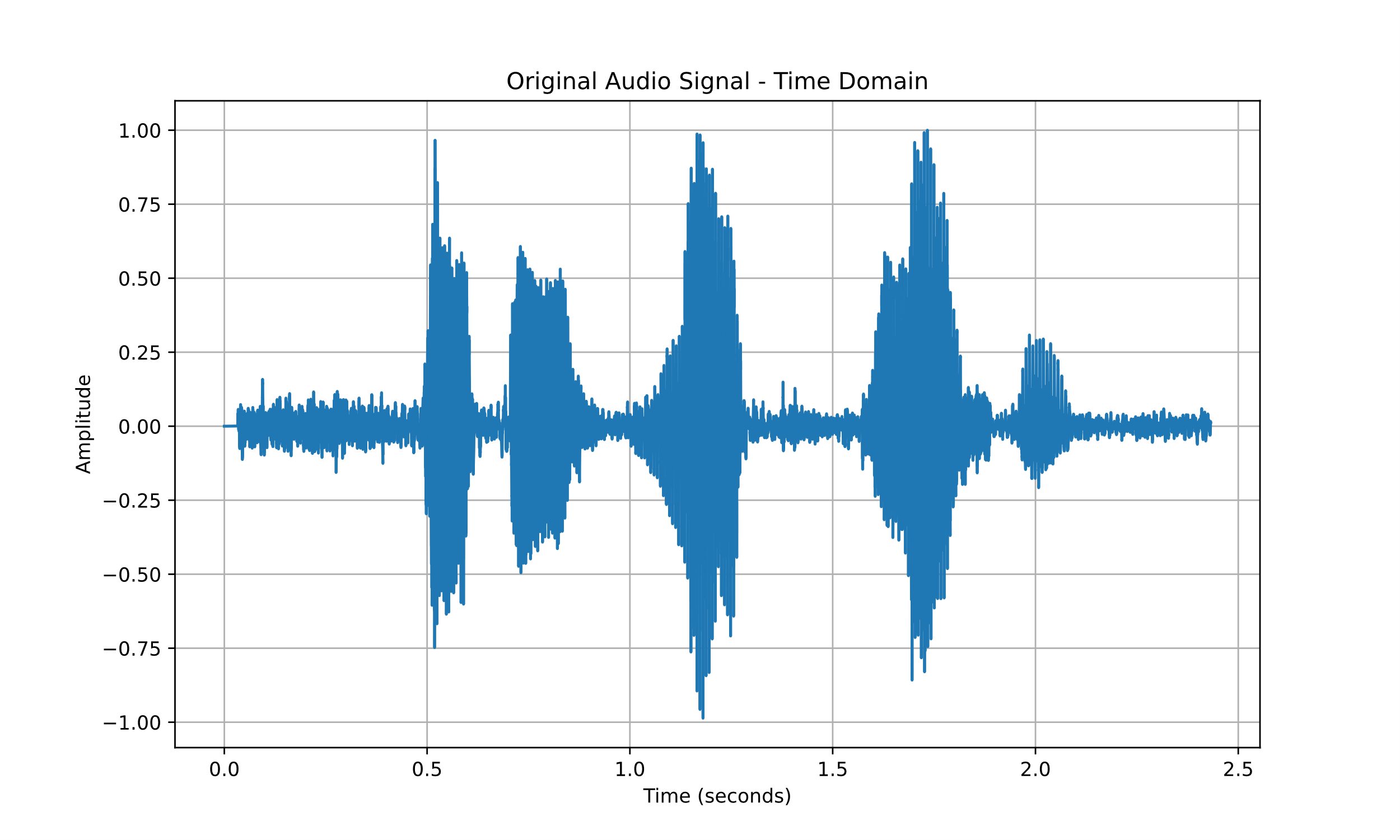


Figure - Time domain plot of orig.wav

* **Frequency Domain Plot**: A Fourier Transform was applied to convert the audio to the frequency domain, where the positive frequencies were extracted and plotted on a logarithmic scale (for both frequency and amplitude) to better visualize the signal spectrum.

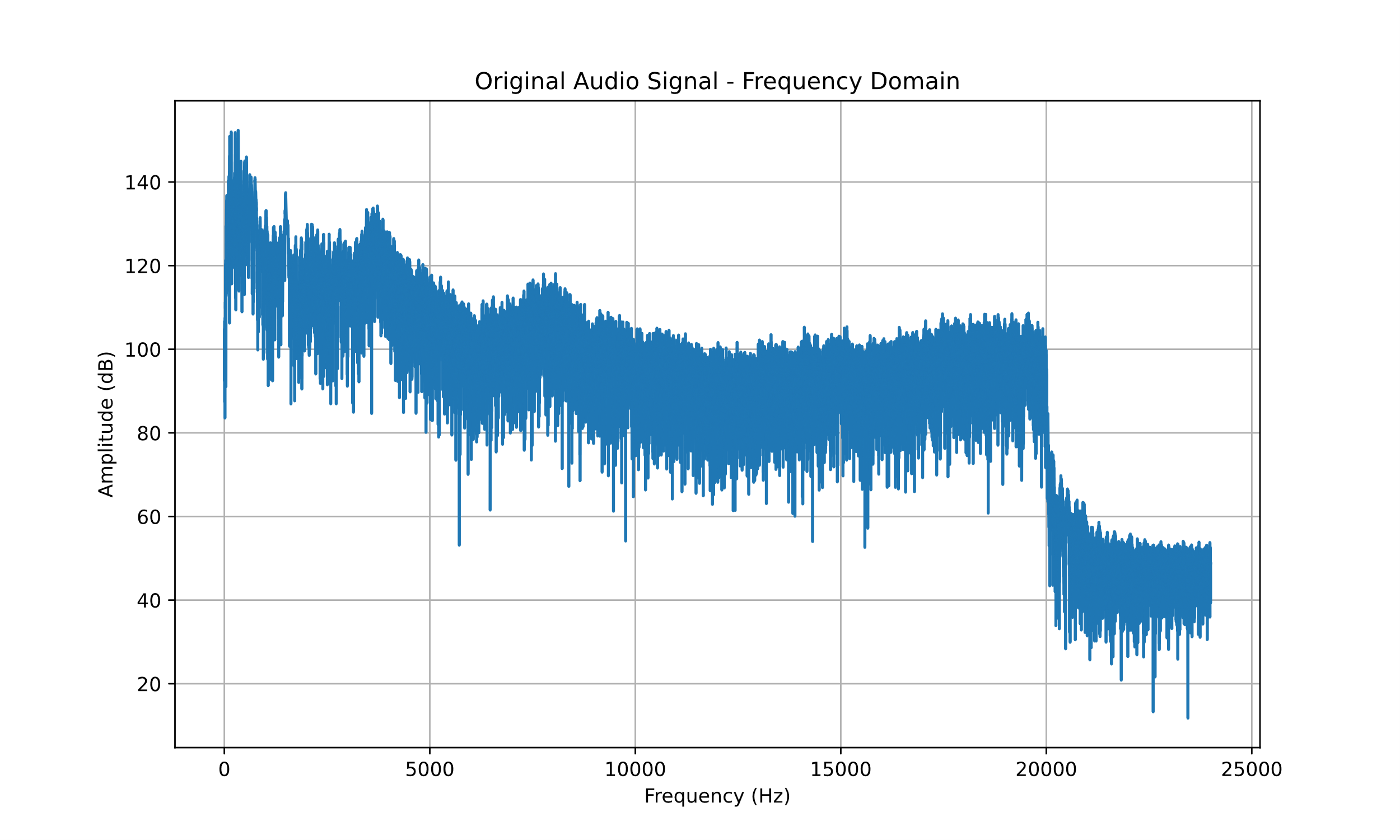


Figure - Frequency domain plot of orig.wav

**Task 2: Identifying Key Features**

* **Fundamental Frequency and Harmonics**: The fundamental frequency of the audio was identified at 168.174 Hz, and harmonic frequencies (multiples of the fundamental) were marked in the frequency domain plot. The harmonics were assumed to range from 2 to 5 times the fundamental frequency.
* **Noise Bands**: Noise bands were identified in the lower (below 85 Hz) and higher (above 8 kHz) frequency ranges, and these were visually annotated on the frequency domain plot. This provided context on how human-produced sounds are perceived and highlighted areas that may not contribute meaningfully to the vocal clarity.

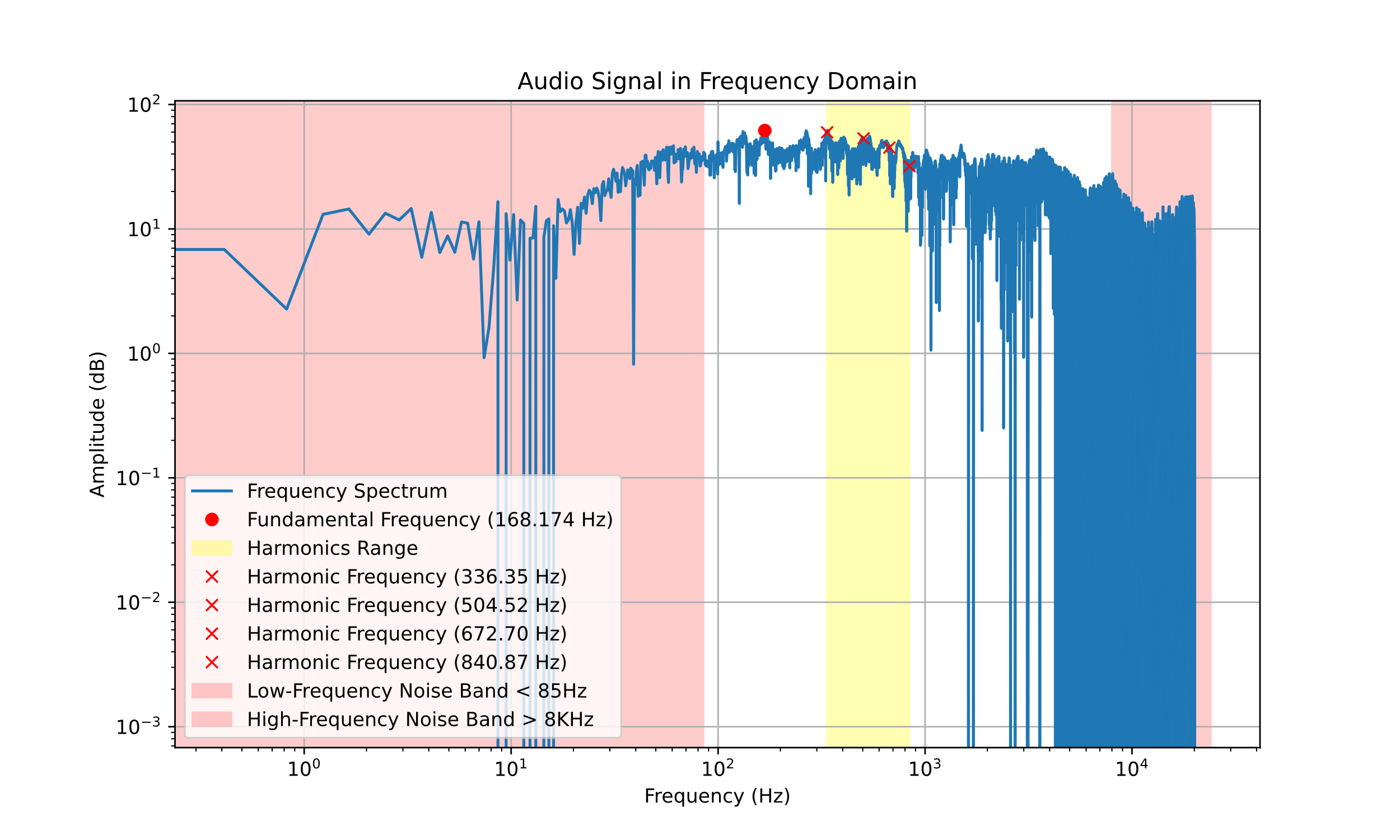


Fig.3

**Task 3: Improving Voice Quality**

To improve the quality of the voice, the frequency bands above 3 kHz were manipulated using the Fourier Transform. This process was aimed at enhancing the clarity and making the voice sound more pleasant and interesting. A custom smoother filter was applied between 3 kHz and 10 kHz to create a gradual transition, reducing sharp changes in this frequency range.

**Task 4: Enhancing the Voice with Aural Exciter**

* **Aural Exciter**: A non-linearity (hyperbolic tangent) was applied to frequencies between 3 kHz and 10 kHz to create the aural exciter effect. This enhanced the harmonic content in this frequency range, making the audio sound more vibrant and clearer.
* A scaling factor of 0.3 was used to ensure that only a small portion of the processed signal was added back to the original, preserving a natural sound while adding clarity.

**Saving and Plotting the Enhanced Audio**

The final enhanced audio was saved as a new file (“final\_enhanced\_audio.wav”) and plotted in both time and frequency domains to visually compare it with the original signal. The enhanced signal showed improved clarity, especially in the key vocal frequency ranges.

**Results**

* The **time domain plot** of the original audio showed the natural amplitude variations of the recorded voice.
* The **frequency domain plot** highlighted the fundamental frequency, harmonics, and noise bands, providing insights into the signal composition.
* After **manipulating the frequency bands above 3 kHz**, the audio showed enhanced clarity, making it sound more pleasant and distinct.
* The application of the **aural exciter** resulted in enhanced vocal brightness and richness, as evidenced by the final frequency domain plot, which showed a more pronounced harmonic structure in the key frequency ranges.

**Conclusion**

This project successfully analyzed and enhanced an audio signal by visualizing and modifying key components in both the time and frequency domains. The use of a smoother filter and aural exciter significantly improved the perceptual quality of the audio. The final enhanced audio exhibits better clarity, enriched harmonic content, and an overall more pleasant listening experience.

**Future Work**

* **Adaptive Enhancement**: Implement a more adaptive technique that can dynamically adjust the enhancement based on the characteristics of the audio.
* **Dynamic Scaling**: Experiment with dynamic scaling for the aural exciter to further optimize the balance between enhancement and naturalness.
* **Machine Learning Approaches**: Utilize machine learning models to automatically identify and enhance specific features in the audio signal.

**References**

* **Python Libraries**: numpy, matplotlib, wave, scipy.io.wavfile
* **Aural Exciter Concept**: [Aphex Aural Exciter](https://www.muzines.co.uk/articles/aphex-aural-exciter-type-b/2850)

**Methods**

**Task 1: Loading and Plotting the Original Audio**

The audio file, “Lab 1 test 2.wav”, was loaded using the Python wave module and converted into a NumPy array. The signal, originally in 16-bit format, was normalized to the range –1 to 1 by dividing each sample by , ensuring the values fell within the appropriate range for further processing.

* **Time Domain Plot**: The time domain plot was generated by creating a time axis from 0 to the total duration of the audio, calculated as the number of frames divided by the frame rate. The normalized audio signal was plotted against this time axis to visually observe amplitude variations.
* **Frequency Domain Plot**: A Fourier Transform (FFT) was applied to convert the audio signal from the time domain to the frequency domain. Only the positive frequencies were extracted and plotted. To visualize the wide range of frequency components effectively, both the frequency and amplitude axes were plotted on a logarithmic scale using plt.xscale('log') and plt.yscale('log'). This helped to emphasize the lower and higher frequency components that are important for audio analysis.

**Task 2: Identifying Key Features**

* **Fundamental Frequency and Harmonics**: The fundamental frequency of the audio was identified at 168.174 Hz. This frequency represents the lowest frequency of the voice signal, which is typically associated with the pitch of the spoken vowel. The fundamental frequency for male voices generally lies between 85 Hz and 255 Hz, while female voices are typically higher. Harmonic frequencies are integer multiples of the fundamental frequency, representing overtones that contribute to the richness of the sound. In this analysis, harmonic frequencies ranging from 2 to 5 times the fundamental frequency were marked on the frequency domain plot to observe their contributions.
* **Noise Bands**: Noise bands were identified in the lower (below 85 Hz) and higher (above 8 kHz) frequency ranges. Low-frequency noise often originates from background hum or vibrations, while high-frequency noise may be due to ambient sound or equipment interference. Identifying these bands was crucial to understand the components of the signal that do not meaningfully contribute to vocal clarity and might need to be reduced or filtered out.

**Task 3: Improving Voice Quality**

To improve the quality of the voice, the frequency bands above 3 kHz were manipulated using the Fourier Transform. Frequencies above 3 kHz are particularly important for the intelligibility and brightness of speech, as they capture many of the consonant sounds that make speech clearer. A custom smoother filter was applied between 3 kHz and 10 kHz. This filter was designed to create a gradual transition in this range, reducing abrupt changes that could lead to harsh sounds or unnatural artifacts. The filter mask was defined to gradually attenuate frequencies in this range, allowing smoother enhancement without introducing sharp artifacts.

**Task 4: Enhancing the Voice with Aural Exciter**

* **Aural Exciter**: To further enhance the vocal clarity and richness, an aural exciter effect was applied. This involved passing the frequency components between 3 kHz and 10 kHz through a non-linearity using the hyperbolic tangent (tanh) function. The non-linearity amplifies the harmonic content, making the voice sound more vibrant and adding a subtle brightness that enhances its presence. A scaling factor of 0.3 was used to ensure that only a controlled amount of the processed signal was added back to the original. This prevented the enhancement from sounding exaggerated or artificial, while still significantly improving the clarity.
* **Tuning for Optimal Results**: The choice of the 3 kHz to 10 kHz range was based on the understanding that this frequency band contains important harmonics and consonant sounds. Other approaches were also considered, such as using a broader range or applying non-linearity to the entire spectrum. However, these approaches often led to unnatural artifacts or excessive enhancement. By focusing on this specific band, the balance between clarity and naturalness was achieved effectively. The tuning of the scaling factor was also crucial; higher values led to a more pronounced effect but risked sounding artificial, while lower values did not produce enough enhancement. The value of 0.3 was found to be optimal after several iterations.

**Saving and Plotting the Enhanced Audio**

The final enhanced audio was saved as a new file (“final\_enhanced\_audio.wav”) and plotted in both time and frequency domains to visually compare it with the original signal. The time domain plot showed smoother transitions, while the frequency domain plot highlighted increased harmonic content in the key vocal frequency ranges, demonstrating the effectiveness of the enhancement.

**Results**

* The **time domain plot** of the original audio showed the natural amplitude variations of the recorded voice.
* The **frequency domain plot** highlighted the fundamental frequency, harmonics, and noise bands, providing insights into the signal composition.
* After **manipulating the frequency bands above 3 kHz**, the audio showed enhanced clarity, making it sound more pleasant and distinct.
* The application of the **aural exciter** resulted in enhanced vocal brightness and richness, as evidenced by the final frequency domain plot, which showed a more pronounced harmonic structure in the key frequency ranges.

**Conclusion**

This project successfully analyzed and enhanced an audio signal by visualizing and modifying key components in both the time and frequency domains. The use of a smoother filter and aural exciter significantly improved the perceptual quality of the audio. The final enhanced audio exhibits better clarity, enriched harmonic content, and an overall more pleasant listening experience.

**Future Work**

* **Adaptive Enhancement**: Implement a more adaptive technique that can dynamically adjust the enhancement based on the characteristics of the audio.
* **Dynamic Scaling**: Experiment with dynamic scaling for the aural exciter to further optimize the balance between enhancement and naturalness.
* **Machine Learning Approaches**: Utilize machine learning models to automatically identify and enhance specific features in the audio signal.

**References**

* **Python Libraries**: numpy, matplotlib, wave, scipy.io.wavfile
* **Aural Exciter Concept**: [Aphex Aural Exciter](https://www.muzines.co.uk/articles/aphex-aural-exciter-type-b/2850)