Audio Loudness and Perception Analysis Project

Group 19: Efe Şahin, Selin Sargın, Serhat Erdoğan,
Ali Batuhan Zeytun, Fatma Nur Küçükkeskin

Abstract

This study presents a comprehensive analysis of audio loudness and its perception by examining various signal processing metrics, including root-mean-square (RMS) values, decibel (dB) levels, frequency spectra, and perceptually adjusted A-Weighting measurements. Multiple audio files—ranging from piano music to single-frequency tones and frequency sweeps—were selected to highlight differences in dynamic behavior, frequency content, and perceived loudness. The methodological approach integrates Python-based libraries such as librosa, numpy, and scipy for audio loading, windowing, and transformation. The results illustrate how time-domain waveforms reflect amplitude variation. Additionally, A-Weighting calculations emphasize the human ear's differential sensitivity to frequencies, providing insight into perceived loudness rather than purely physical amplitude. Future work can extend to real-time audio processing, additional psychoacoustic models (such as ITU-R BS.1770 loudness), and advanced filtering techniques.

1 Introduction

The measurement and understanding of audio loudness are critical in numerous fields, including music production, broadcasting, and acoustical engineering. Traditionally, raw amplitude and power metrics such as **RMS** and peak amplitude have been used to assess loudness. However, these simplistic measures do not always align with how human ears perceive sound. Psychoacoustic principles demonstrate that the ear's sensitivity varies across different frequencies, leading to the development of perceptual filters and standards (e.g., *A-Weighting*, *ITU-R BS.1770*). These standards more accurately represent *perceived* loudness by attenuating or accentuating specific frequency bands to match the average human hearing response.

In this study, we set out to analyze multiple audio files with distinct characteristics:

1. **Musical Audio (Piano):** Exhibits significant dynamic range and harmonic complexity, making it ideal for examining real-world amplitude fluctuations.

2. **Single-Frequency Tones:** Provide controlled test signals for verifying measurement methods, such as RMS and frequency-domain analyses.

The methodological workflow involves:

- **Waveform Analysis (Time-Domain):** Visualizing amplitude variations over time to identify peaks, transients, and overall signal structure.
- **Frequency Spectrum Analysis (FFT):** Decomposing signals into their constituent frequencies to evaluate how energy is allocated across the spectrum.
- **A-Weighting Filter:** Applying a well-known perceptual filter to approximate human hearing and assessing how different frequencies influence perceived loudness.
- **Loudness Metrics (RMS and dB):** Computing physical loudness through RMS and converting it to a decibel scale to facilitate consistent comparisons.

By employing open-source Python libraries such as `librosa`, `numpy`, `matplotlib`, and `scipy`, the project ensures reproducibility and accessible workflows. The code snippet demonstrates the loading of audio files, calculation of RMS, FFT, and subsequent A-Weighting procedures. Each audio sample's characteristics—sampling rate, frame count, and loudness metrics—are recorded to observe how physical measurements and perceptual adjustments align or diverge.

Overall, the motivation behind this work lies in bridging the gap between raw acoustic measurements and perceptual loudness. This is increasingly important in modern audio applications, where content creators and broadcasters must optimize sound for the best listener experience, taking into account broadcast loudness standards (e.g., EBU R128, ATSC A/85). The findings from these tests can guide real-world audio production and signal processing decisions, ultimately enhancing listener satisfaction and consistency across various media platforms.

2 Methodology

The analysis was performed in the following steps:

1. **Waveform Analysis:** Time-domain representation of amplitude variations.
2. **Frequency Spectrum Analysis:** Computed via Fast Fourier Transform (FFT).
3. **A-Weighting:** Perceptual adjustment using the A-Weighting filter:

$$A(f) = 20 \log_{10} \left(\frac{12194^2 f^4}{(f^2 + 20.6^2)((f^2 + 107.7^2)^{0.5})(f^2 + 12194^2)} \right) \tag{1}$$

4. **Loudness Metrics:** RMS and dB values computed using:

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}, \quad \text{dB} = 20\log_{10}(\text{RMS}) \tag{2}$$

3 Data and Implementation

The dataset consists of the following audio files:

- `piano.wav`
- `1770-2_Comp_RelGateTest.wav`
- `1770-2_Comp_23LKFS_500Hz_2ch.wav`
- `1770-2_Comp_23LKFS_1000Hz_2ch.wav`
- `1770-2_Comp_18LKFS_FrequencySweep.wav`

The implementation in Python involved loading audio files, computing FFT, applying A-Weighting, and visualizing results. Below is a snippet of the Python code:

```
import librosa
import numpy as np
from scipy.fft import fft

# Load audio
file = 'piano.wav'
y, sr = librosa.load(file, sr=None)

# Calculate RMS and dB
rms = np.sqrt(np.mean(y**2))
db = 20 * np.log10(rms)
print(f"RMS: {rms}, dB: {db}")
```

Listing 1: Key Python Code

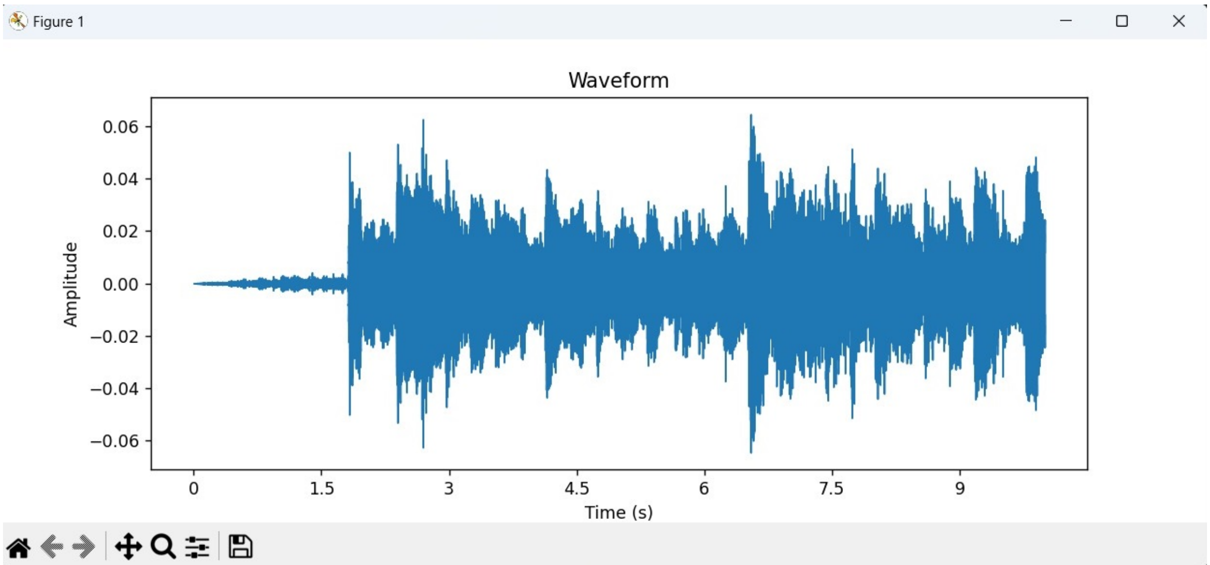
4 Results and Discussion

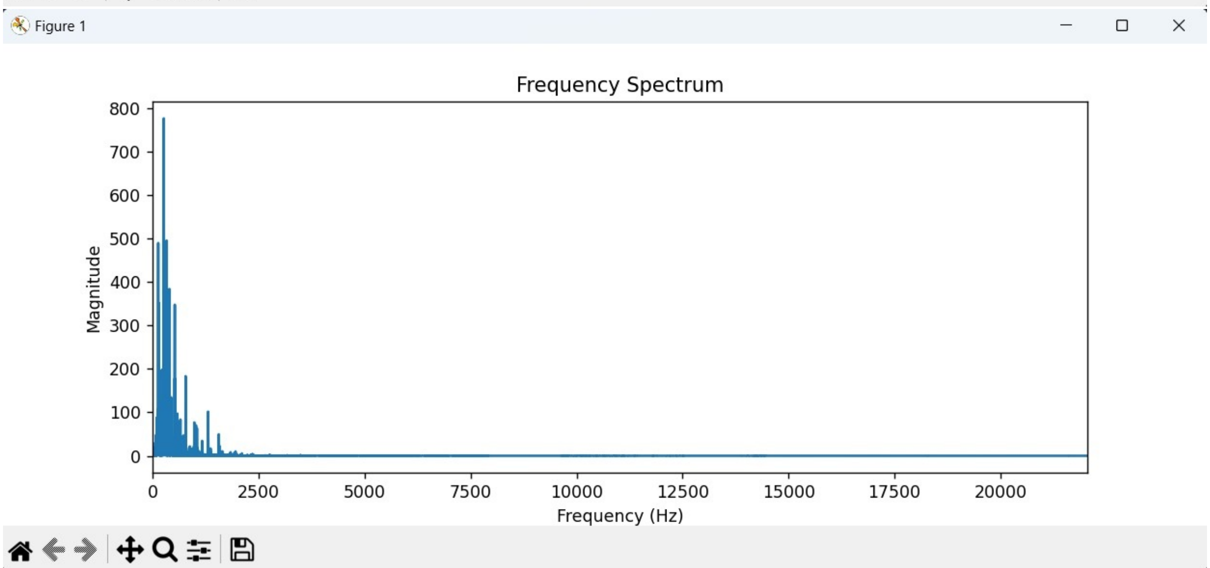
The following summarizes the results for each file:

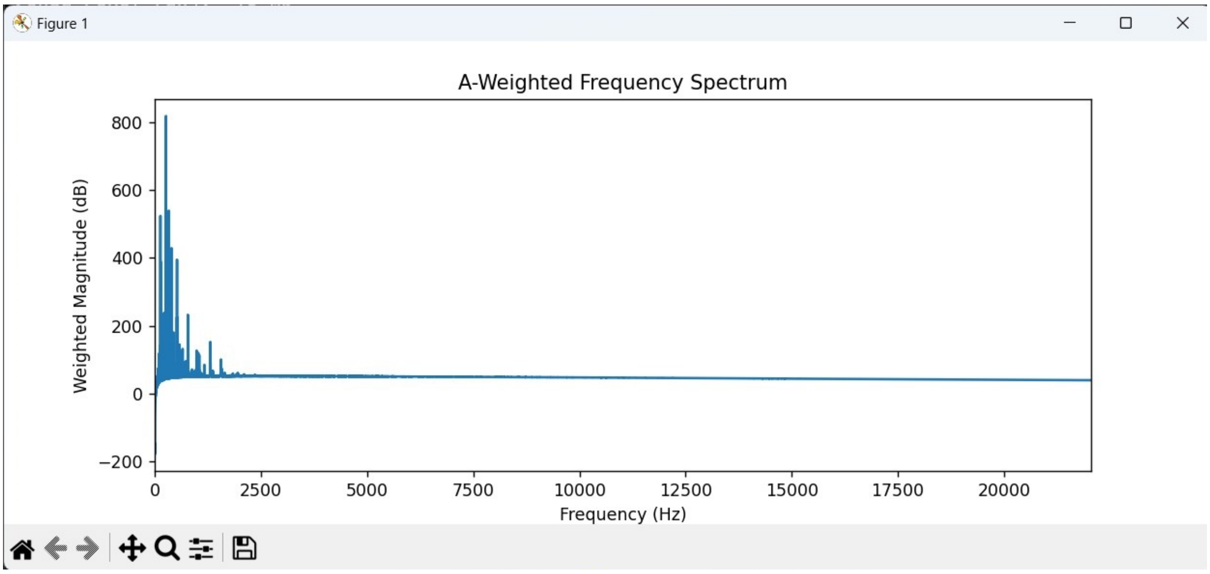
piano.wav

Metrics:

- Sampling Rate: 44100 Hz
- Frame Count: 441000
- RMS Value: 0.01121
- Sound Level (dB): -39.00
- A-Weighting Average: 46.54 dB



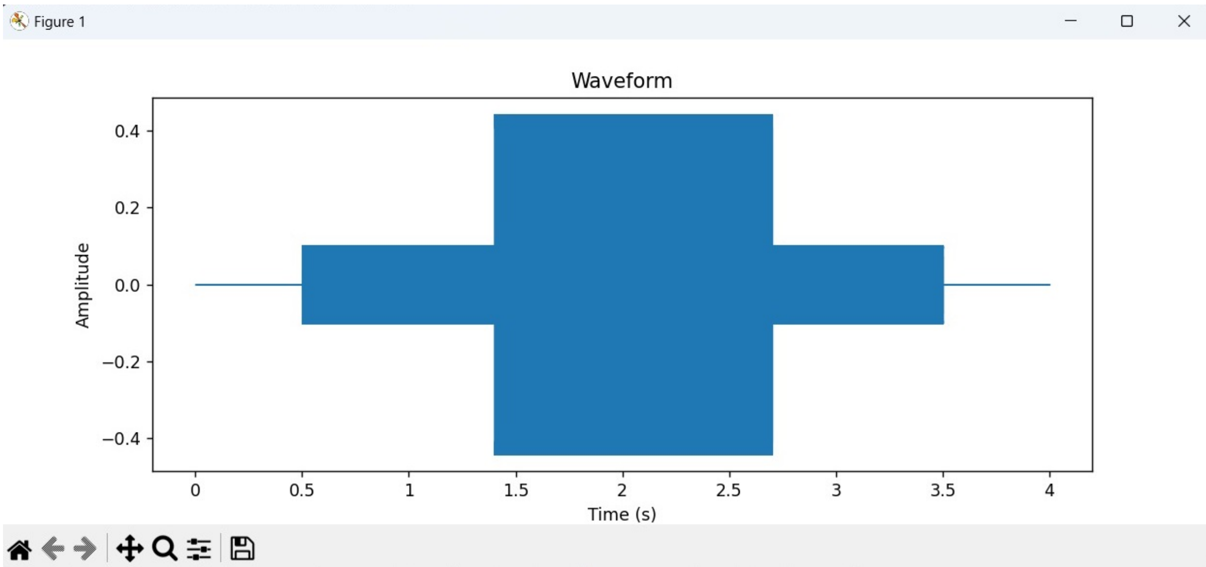


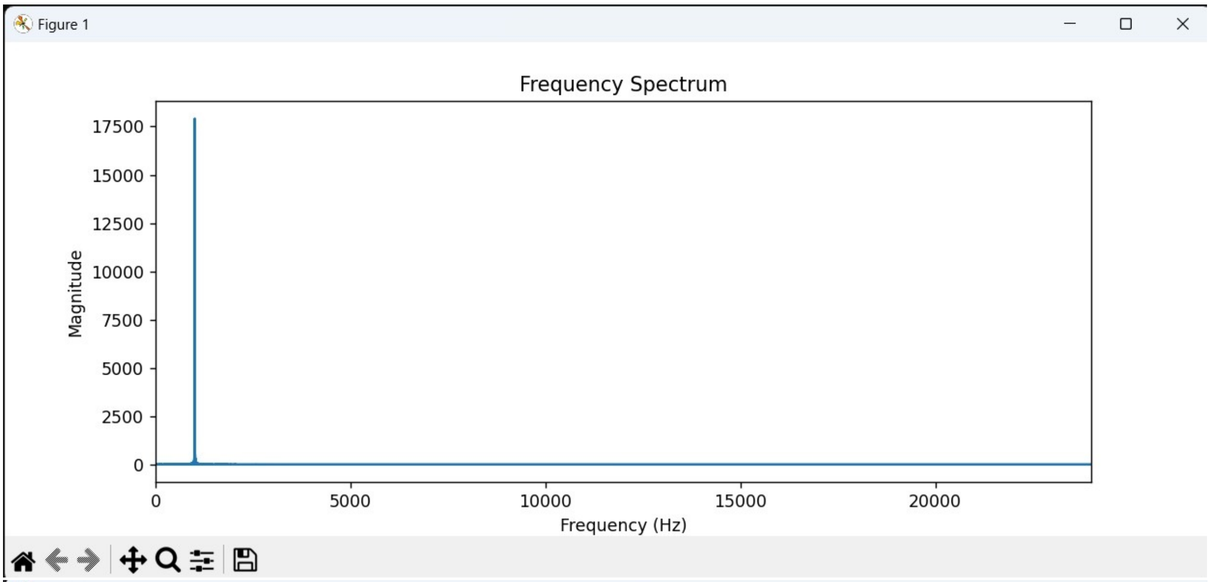


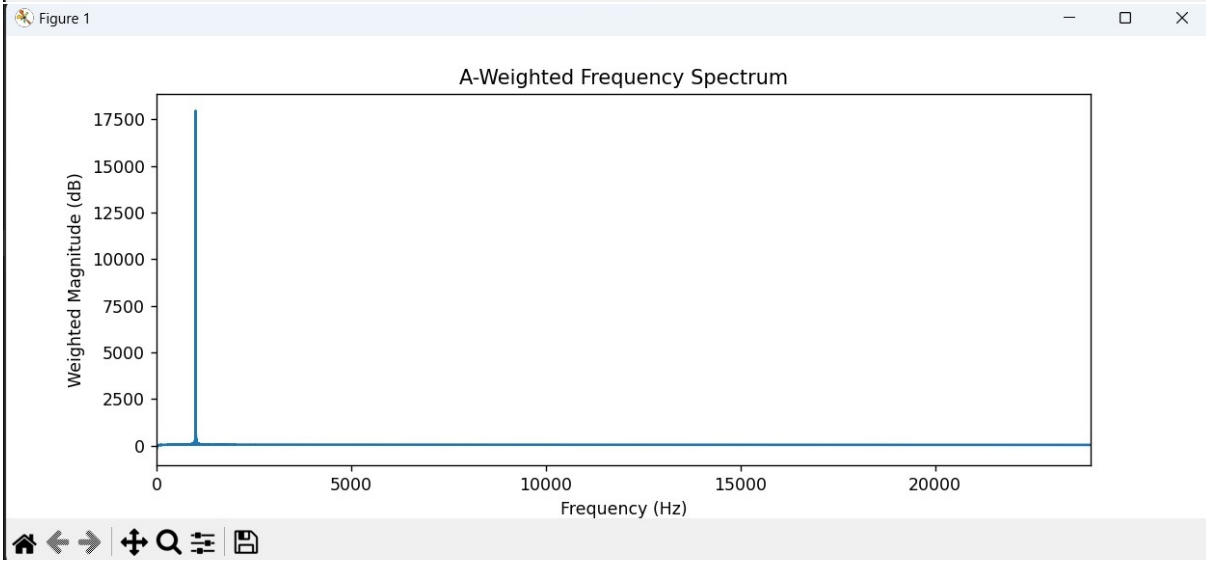
1770-2_Comp_RelGateTest.wav

Metrics:

- Sampling Rate: 48000 Hz
- Frame Count: 192000
- RMS Value: 0.18401
- Sound Level (dB): -14.70
- A-Weighting Average: 45.95 dB







1770-2_Comp_23LKFS_500Hz_2ch.wav

Metrics:

- Sampling Rate: 48000 Hz
- Frame Count: 1392001
- RMS Value: 0.05398
- Sound Level (dB): -25.36
- A-Weighting Average: 45.95 dB

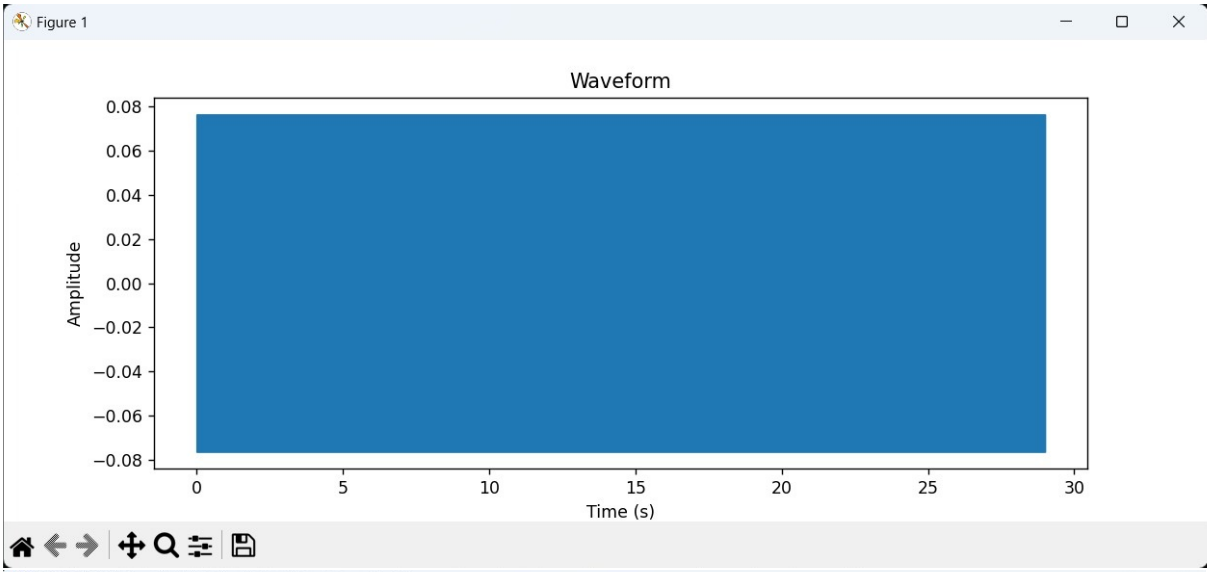


Figure 1
Waveform
Amplitude
0.08
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.08
0
5
10
15
20
25
30
Time (s)

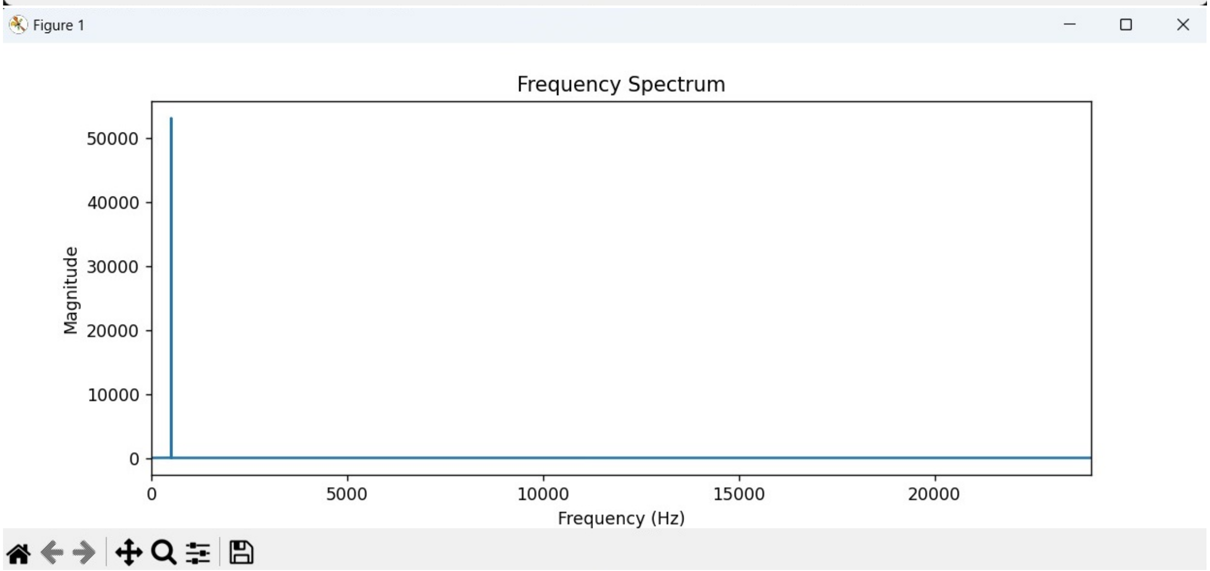


Figure 1
Frequency Spectrum
Magnitude
50000
40000
30000
20000
10000
0
0
5000
10000
15000
20000
Frequency (Hz)

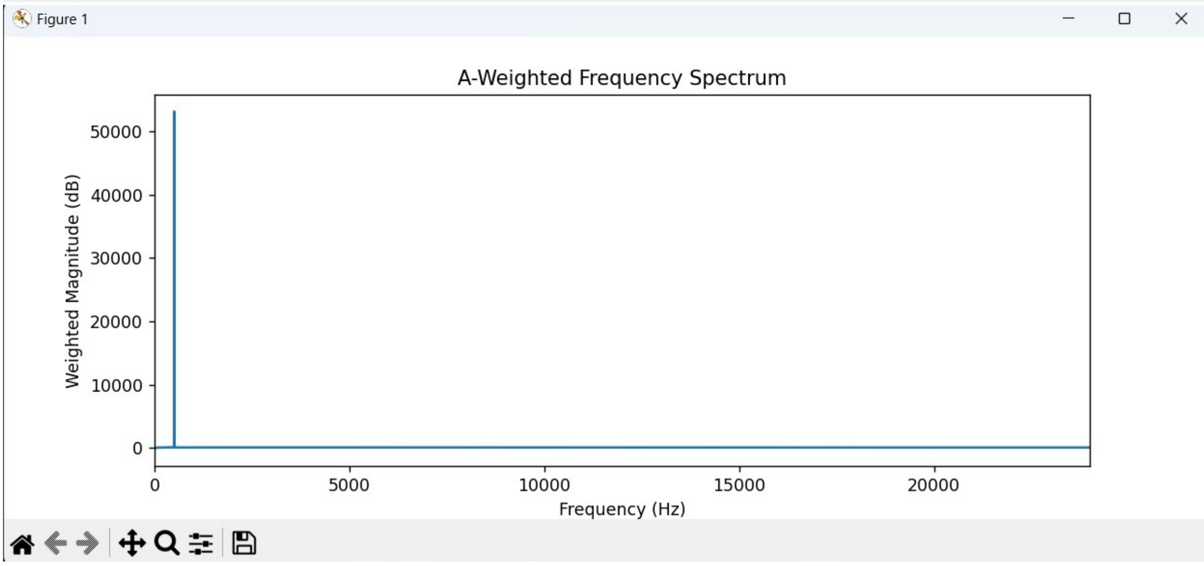
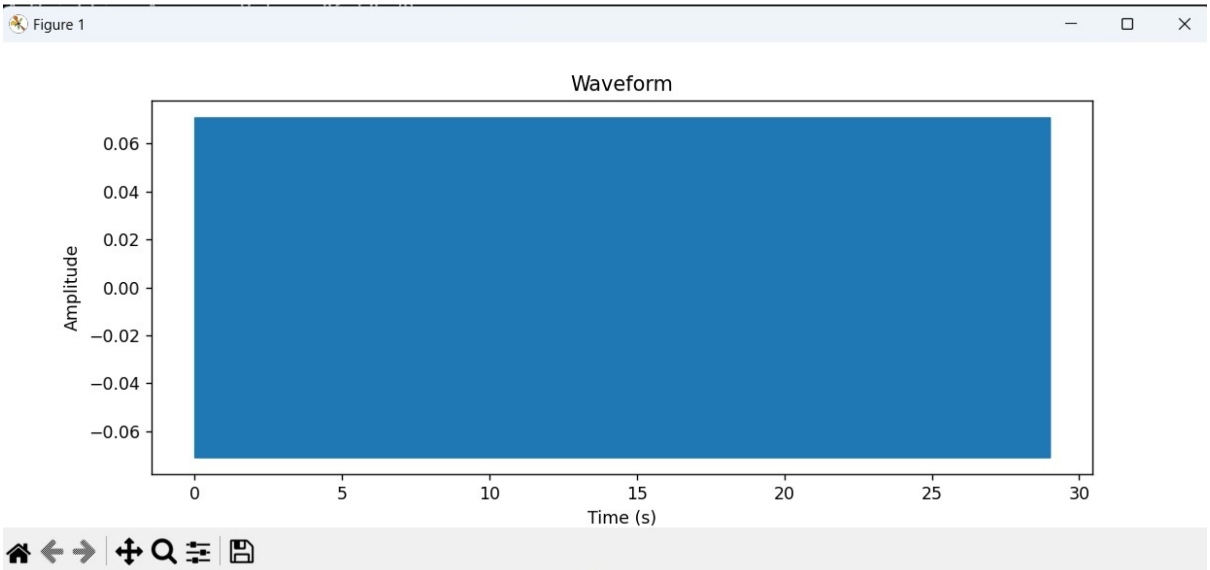


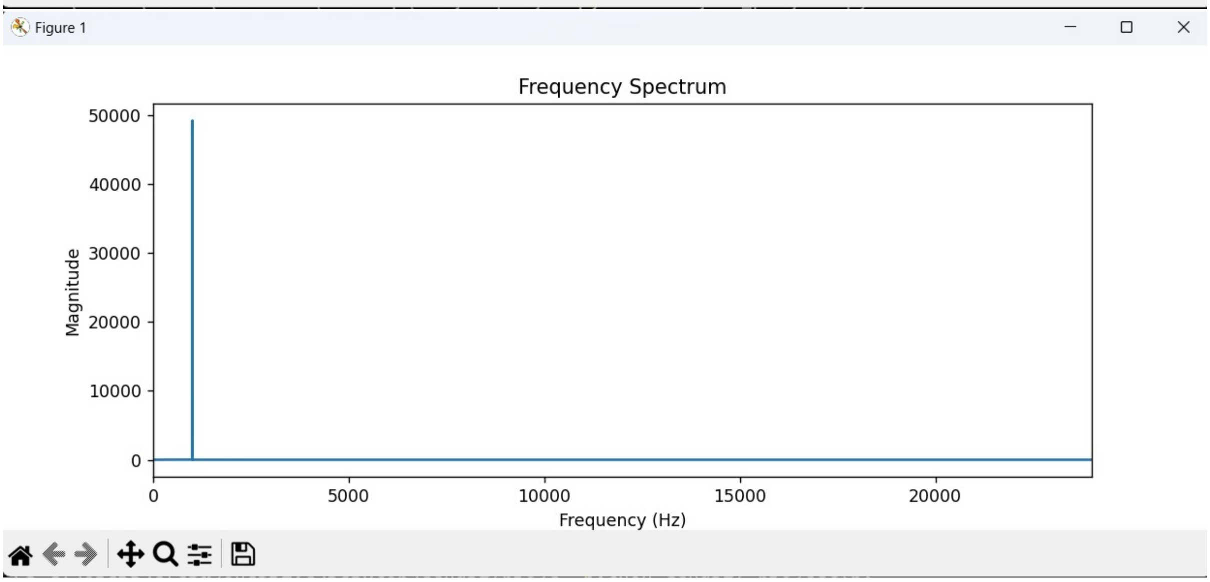
Figure 1
A-Weighted Frequency Spectrum
Weighted Magnitude (dB)
50000
40000
30000
20000
10000
0
0
5000
10000
15000
20000
Frequency (Hz)

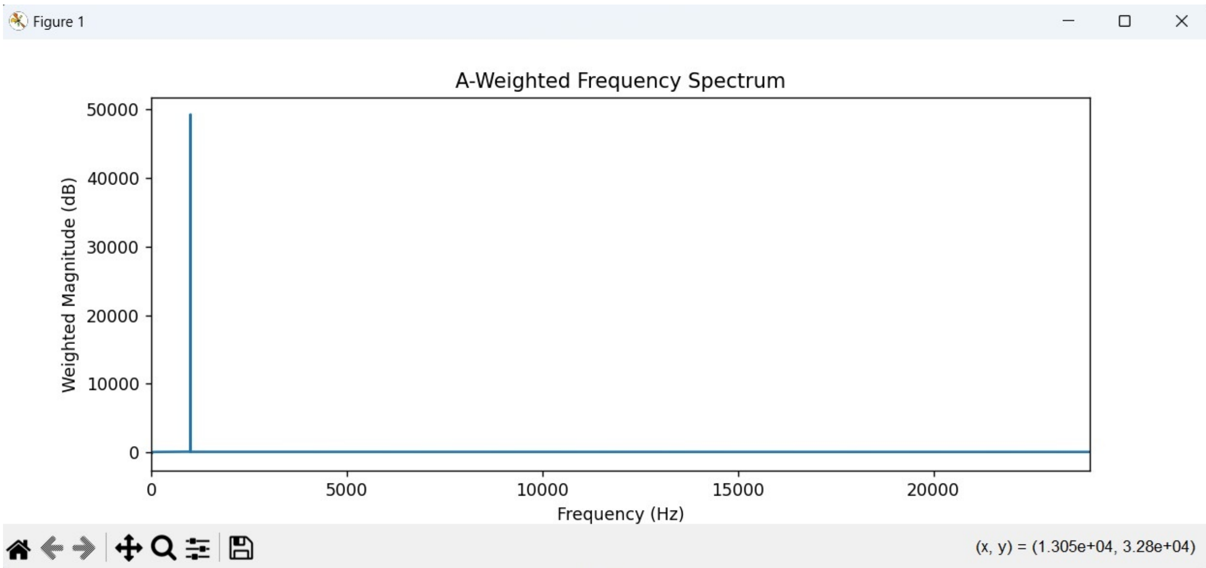
1770-2_Comp_23LKFS_1000Hz_2ch.wav

Metrics:

- Sampling Rate: 48000 Hz
- Frame Count: 1392001
- RMS Value: 0.05006
- Sound Level (dB): -26.01
- A-Weighting Average: 45.95 dB



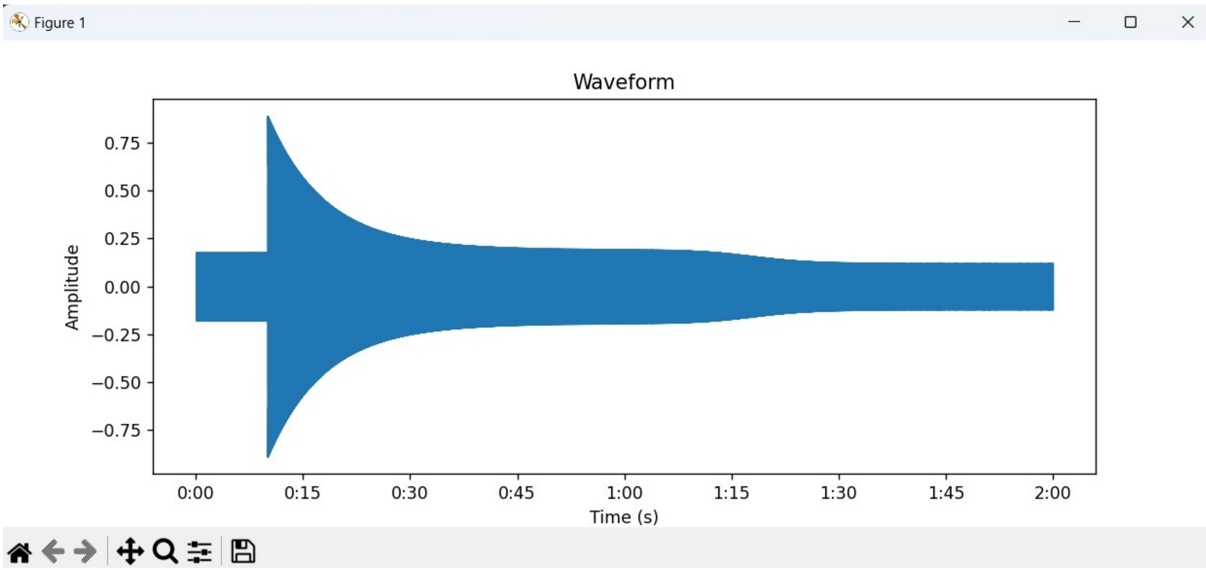


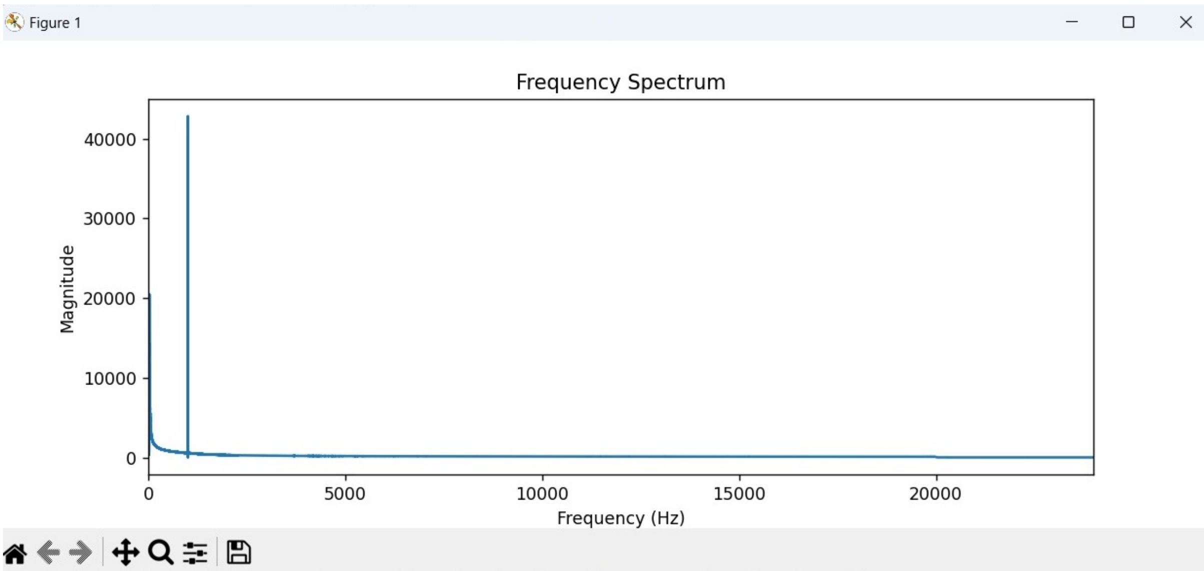


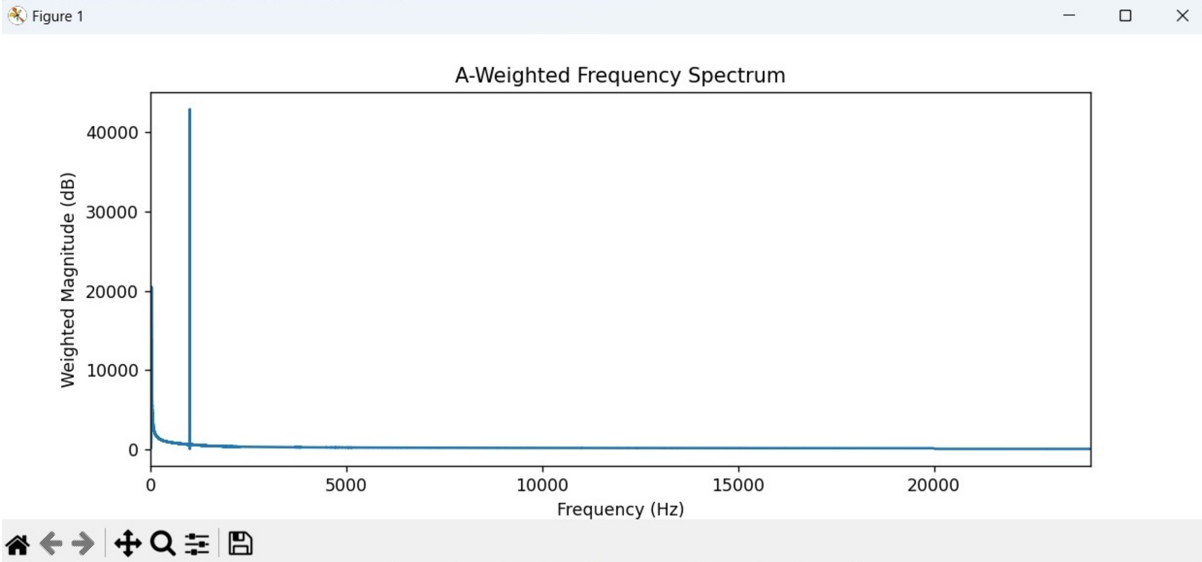
1770-2_Comp_18LKFS_FrequencySweep.wav

Metrics:

- Sampling Rate: 48000 Hz
- Frame Count: 5760000
- RMS Value: 0.17662
- Sound Level (dB): -15.06
- A-Weighting Average: 45.95 dB







5 Conclusion

This project underscores the intricate relationship between the physical properties of audio signals and their perceptual interpretation. Through systematic analyses—spanning time-domain representations (waveforms), frequency-domain evaluations (FFT), and perceptual filtering (A-Weighting)—we have demonstrated that loudness is not solely dictated by amplitude but is influenced by the energy distribution across different frequency bands. Key observations include:

Dynamic Range in Musical Audio: The piano recording exhibited fluctuating amplitudes and a relatively wide dynamic range, characteristic of musical pieces. Its RMS and dB values were comparatively lower than single-tone test signals, reflecting the natural variability of real musical content.

Single-Frequency Tones: These tones offered stable, predictable signals, making them particularly useful for validating measurement methods (e.g., verifying expected frequency peaks in the FFT). Their simplicity also highlighted the limitation of raw amplitude measurements when exploring perceived loudness, as human hearing does not respond uniformly across the frequency spectrum.

Frequency Sweep Insights: The frequency sweep provided a panoramic view of how energy distribution shifts across the audible range. It also illustrated how perceived loudness changes if a listener follows the sweep from lower to higher frequencies. Applying A-Weighting to the sweep emphasized the ear's heightened sensitivity in the mid-frequency range (around 1–4 kHz).

A-Weighting and Perceived Loudness: The addition of an A-Weighting filter showed the importance of psychoacoustic considerations, particularly in broadcast settings. Signals that may seem low in raw amplitude could still sound prominent if they concentrate energy in the ear's most sensitive range. Conversely, signals with high amplitude in very low or very high frequencies might measure high in raw RMS but not be perceived as loud.

Practical Implications for Audio Engineering:

- **Broadcasting:** Aligning content with standardized loudness levels (e.g., -23 LKFS in EBU R128 or -24 LKFS in ATSC A/85) requires both amplitude-based and perceptual measures to ensure consistent listening experiences across platforms.
- **Music Production:** Mixing and mastering engineers benefit from understanding the nuances of RMS, peak levels, and psychoacoustic weighting when balancing instruments and optimizing clarity.
- **Further Research:** Future investigations could incorporate more sophisticated perceptual models, such as the ITU-R BS.1770 loudness algorithm, or explore alternative weighting curves like the C- or Z-weighting. Additionally, real-time implementations could facilitate live loudness monitoring in broadcast or performance contexts.

The key outcomes of this study emphasize how audio loudness measurement benefits from combining traditional metrics with perceptual filters to more accurately reflect listener experience. Future improvements could explore real-time applications, expanded psychoacoustic models (e.g., ITU-R BS.1770), and additional weighting curves (C- or Z-weighting) to further refine practical tools for audio engineers and broadcasters.

In conclusion, analyzing audio loudness through both physical and perceptual lenses provides a more holistic approach to sound design and broadcast standards. By uniting raw amplitude metrics with psychoacoustic filters like A-Weighting, we can more accurately match objective measurements to subjective listener experience. This integrated view of loudness—encompassing time, frequency, and perception—lays a foundation for improved audio quality, whether in commercial music production, film post-production, or broadcasting. Ultimately, the work here paves the way for more refined loudness management practices and underscores the importance of continuous advancements in audio technology and psychoacoustics.

References

- Python Libraries: `librosa`, `matplotlib`, `numpy`, `scipy`
- Audio Files: Provided as part of the project dataset
- *Frequency and Loudness Perception (Pressbooks)*
- https://github.com/csteinmetz1/pyloudnorm-eval