A multi-resolution sinusoidal model

Part 1

You can download or listen to the sounds chosen here:

- Orchestra fragment
- One man Rumba

Part 2

- The first orchestra fragment is a polyphonic sound with percussion in low frequencies and some other higher pitch instruments. One can observe that we have some not very well defined harmonics for the attacks around 100 Hz and some better defined harmonics in the low frequency band around 2000 Hz. By observing the plots on Figure 1 we can see both that the original and synthesized signal in time domain are very similar. Also, using STFT we can see that in the time-frequency domain we have a good frequency resolution and the attacks are well defined. To do that 3 windows were used: a blackman window with size 4095 and FFT size 8192, another blackman with size 2047 and FFT size of 4096 and at last a hamming window with size 1023 and FFT size 2048.
- The second audio signal is a rumba fragment that is a polyphonic sound too with percussion in low frequencies and a fundamental frequency of 50 Hz. On the higher pitches lie some claps that reach even very high frequencies in the spectrogram around 20 kHz. Again we can see on Figure 2 a good reconstruction. To do that 3 windows were used: a blackman window with size 1001 and FFT size 2048 to ensure that we have a good time resolution and keep the attacks, another blackman with size 2047 and FFT size of 4096 and at last a hamming window with size 1023 and FFT size 2048.

Part 3

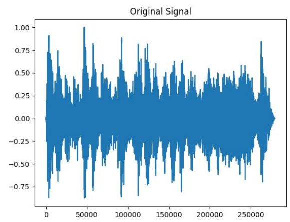
Multi-resolution analysis balances time and frequency resolution. It uses varying FFT sizes to provide high time resolution for capturing abrupt changes (attacks) and high frequency resolution for detailed spectral analysis. This approach enhances the analysis of complex signals while managing computational complexity by adjusting FFT sizes. In terms of computational complexity, while smaller FFT sizes are computationally less demanding, larger sizes provide detailed frequency information but at a cost of higher computational complexity. The process of sinusoidal model analysis can be extended depending on our needs. For HPS models, it helps in accurately separating harmonic content from -stochastic- noise by maintaining high frequency resolution for harmonic components and high time resolution for stochastic components. For HPR models we can isolate the harmonic elements and residual noise for a signal reconstruction that will be closer to the original one.

Part 4

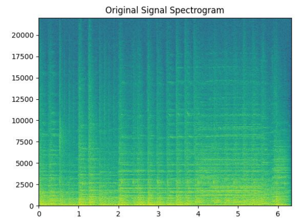
One potential challenge is that estimating the fundamental frequency f_0 can be complex. It is important to first analyze the lowest frequency band to accurately capture the lowest possible f_0 and then ensure that all multiples of this fundamental frequency are correctly identified as harmonics. n HPR models, distinguishing between harmonic and residual components requires robust F0 tracking. Variability in resolution can complicate the detection of harmonic structures and accurate f_0 extraction. For HPS models, distinguishing stochastic noise from harmonic components requires precise f_0 estimation and sinusoidal tracking. Multi-resolution analysis must balance between capturing fine details of harmonic components and separating stochastic noise. Thus we need some resolution management techniques to ensure accurate sinusoid tracking and F0 estimation.

Part 5

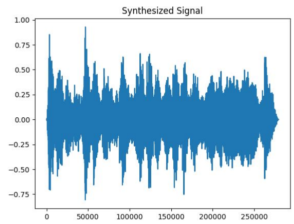
Some famous methods are adaptive windowing, more sophisticated frequency interpolation and harmonics peak interpolation techniques. Also hierarchical tracking methods that start with broader resolutions and step by step refine sinusoid tracking at higher resolutions, that allows for efficient tracking and f_0 estimation by having different resolution levels sequentially.



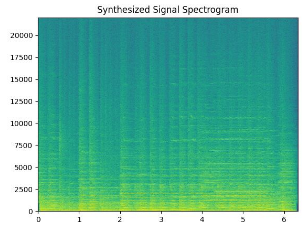
Original signal in time-domain



Original signal in time-frequency domain

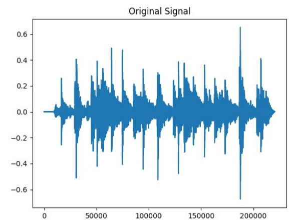


Synthesized signal in time-domain

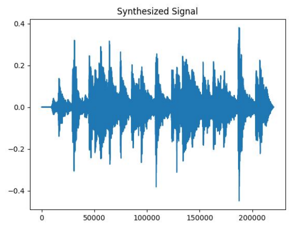


Synthesized signal in time-frequency domain

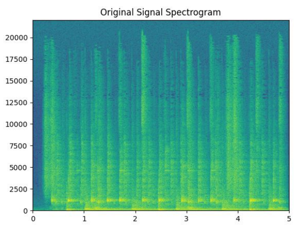
Figure 1: Orchestra fragment original signal vs sinusoidal model



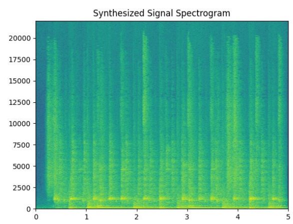
Original signal in time-domain



Original signal in time-frequency domain



Synthesized signal in time-domain



Synthesized signal in time-frequency domain

Figure 2: Orchestra fragment original signal vs sinusoidal model