Sound delay/Sound phased array[a)]

D. Lotnyk[1, b)]
Physical Department, Cornell University.

(Dated: 7 August 2018)

In the current project, the violin wav audio file was analyzed. Using Finite Impulse Response (FIR) filter with selected parameters is recommended to smooth an original data. The Fast Fourier Transform (FFT) was used to obtain the frequency spectra of original data, smoothed data, the sample and mix of sample + shifted data.

The huge hysteresis was observed between ascending and descending time delay of sample/shifted sample data – 0.04 sec vs 0.012 sec. The idea of using this effect in a sound phased array translators was proposed.

Keywords: Sound, wav-files, FFT, FIR, Sound Phased Array Antenna

I. SIGNAL ANALYSIS

In this work, I used the 'Violin-I.wav' sound file for analysis (see Fig. 1). First, I focused on the appropriate smoothing of the original signal. The standard filters such as 'median' or 'moving' corrupts the frequency spectrum and can not be used. Obvious choice was the FFT filter and, lately, I found out the FIR (finite impulse response) filter which was among the standard filters of signal processing in Matlab. The FIR is basically use the FFT algorithms as well. It is simply computed the DFT (Discrete Fourier Transform) of an initial filter design that you have using the FFT algorithm (if you don't have an initial estimate you can start with h[n]=delta[n]). In the Fourier domain or FFT domain you correct the frequency response according to your desired specs and compute the inverse FFT. In time-domain you retain only N of the coefficients (force the other coefficients to zero). Compute the FFT once again. Correct the frequency response according to specs. As one can see at Fig. 1, the filtered signal is more symmetric according to Ox-axis.

The frequency spectra of original and FIR-filtered data are shown in Fig. 2.

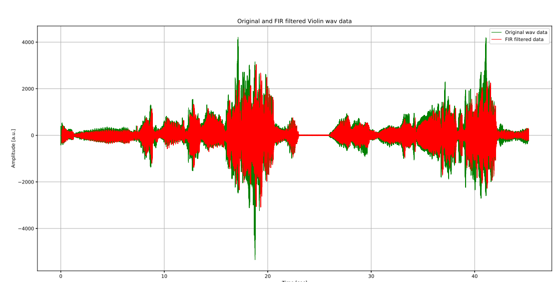

FIG. 1. Wave form of the original sound data (green) and FIR filtered (red).

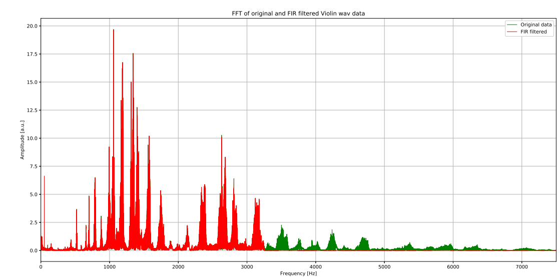

FIG. 2. The FFT of the original sound data (green) and FIR filtered (red).

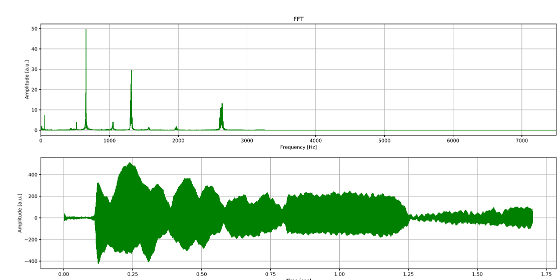

FIG. 3. The FFT of the sample data (top) and the sample itself (bottom).

The frequency spectrum after FIR filtering (red) is not different from the original one except for the cut-off frequency ≈ 3500 Hz.

II. SAMPLE SHIFTING

As a training sample (see Fig. 3), The beginning 1.7 sec part was chosen in order to define how two sample should timely separated to hear the difference.

Two audio files were created using the sample data. In the first one, the time delay increases from zero to $2000/48100 \approx 0.042$ sec in $200/48100 \approx 0.0042$ sec per step (48100 is the sampling frequency). As a result, the

a)I would like to thank prof. Graeme Bailey for mentoring this work.
b)Electronic mail: dl629@cornell.edu

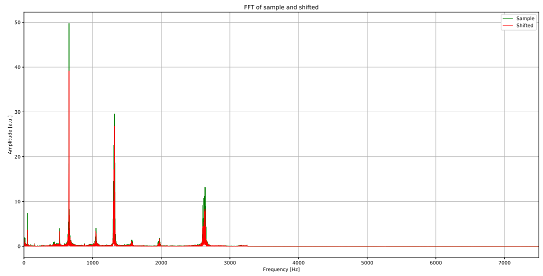

FIG. 4. The FFT of the original sound data (green) and FIR filtered (red).

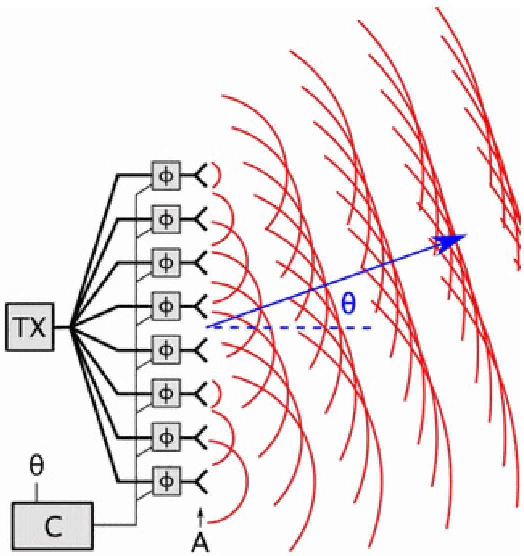

FIG. 5. The scheme of phased array antenna.

distinguished difference between two samples was heard only at 8-9 pulse which is ≈ 0.04 sec delay. The most impressive finding for me was the second audio file which was generated starting from $3000/48100 \approx 0.062$ sec and decrease in $300/48100 \approx 0.0042$ sec per step. In this case, the distinctive difference between samples was heard at 7-8 sample with the time delay only ≈ 0.012 sec!!!

Such a big difference was an unexpected result for me. The important thing is the frequency spectrum is not change (the peaks position) as one can see in Fig. 4. In that sense, it is hard to identify perfectly the non-simultaneous play. At least, FFT analysis is not appropriate to define it. Another unsuccessful at temp was to define a lag using 1D cross-correlation which is natural way to determine a lag between two signals.

III. CONCLUSIONS

In the current project, the violin wav audio file was analyzed. Using FIR filter with selected parameters is recommended to smooth an original data. The fast fourier transform was used to obtain the frequency spectra of original data, smoothed data, the sample and mix of sample + shifted data.

The huge hysteresis was observed between ascending and descending time delay of sample/shifted sample data – 0.04 sec vs 0.012 sec. The difference should be the pitch dependent. There was an attempt to identify the criteria of synchronize playing the same sample using FFT analysis and 1D cross correlation function. However, no clear method is proposed. Probably, there is another method/procedure to identify it.

One of the proposal which comes to mind since there is no such a easy way to identify/hear the difference of playing the same note is to use analog of the phased array antenna/emitter in radiolocation. Scheme is presented in Fig. 5.

The idea is to control the lag to control the beam array of the signal to accent or to confuse the listener. Such method could work at big concerts, in studios or in the game production