Math 761 Fast Fourier Transform Project

Restoration of Analog Audio Recordings

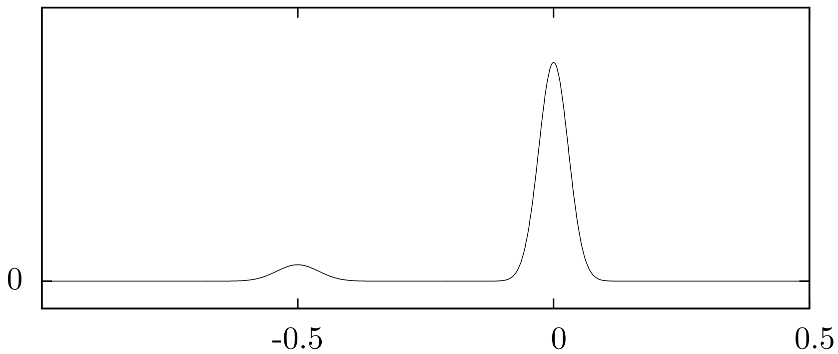
This project will explore use of the fast Fourier transform to restore sound recordings that have undergone quality loss from being stored on analog magnetic tapes. Three types of signal corruption will be addressed by our digital signal processing:

1. Loss of the high frequencies which result from tape azimuth and alignment errors, finite size of the tape head gap and diffusion in the magnetic media.
2. Print through or pre-echo resulting from transfer of magnetic oxide from one layer of the media to the next when the tape is stored tail-in for long periods of time.
3. Random noise which comes from the circuitry of the tape recorder, digitization equipment and external noise sources.

Part 1. Loss of high frequencies along with tape print through may be modeled by convolution of the original signal x with the function

$$k(t) = \frac{\alpha}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{t^2}{2\sigma^2}\Big) + \frac{\beta}{\sqrt{4\pi\sigma^2}} \exp\Big(-\frac{(t+0.5)^2}{4\sigma^2}\Big).$$

The Gaussian of height α represents the loss of high frequencies and the Gaussian of height β represents a 0.5 second pre-echo. Graphically, k looks like



Thus, $w = k * x$ represents the original signal x after having been degraded by loss of high frequencies and tape print through.

The file `signalw.dat` contains $N = 2^{18}$ decimal numbers between -1 and 1 representing about 6 seconds of audio sampled at 44.1 khz with seven digits of precision where $\alpha = 1.75$, $\beta = 0.5$ and $\sigma = 0.0004$. The convolution theorem allows us to solve for $\hat{x}_j$ from $\hat{w}_j = \hat{k}_j\hat{x}_j$ and then take the inverse transform of $\hat{w}_j/\hat{k}_j$ to obtain x. When forming the quotient $\hat{w}_j/\hat{k}_j$ numerically it is important not to divide by anything close to zero. Since the corrupted signal in `signalw.dat` is given with seven decimal digits of precision, then zero is about $\epsilon = 10^{-6}$. Therefore, we can approximate

$$\hat{x}_j \approx \begin{cases} \hat{w}_j/\hat{k}_j & \text{for } |\hat{k}_j| > \epsilon \\ 0 & \text{otherwise.} \end{cases}$$

Write a program which uses the fast Fourier transform to approximate x from w or download the program `prog1p1`. Compute the relative error

$$E(X,x)=\sqrt{\sum_{n=1}^{N}|X_n-x_n|^2\Big/\sum_{n=1}^{N}|x_n|^2}$$

where X is the approximation and x is the original signal given in `signalx.dat`.

Part 2. In this section we shall recover x when random noise is also present. Let η be a white noise source that is roughly 10 000 times larger than the seven-digit roundoff error already present in $k*x$. Thus, $y=k*x+\eta$ represents the original signal x after having been degraded by loss of high frequencies, tape print through and random noise.

The file `signaly.dat` contains about 6 seconds of the degraded audio signal y. Because of the noise, we need to relax our notion of what numbers are close to zero when solving for x. In particular, a larger value of ϵ will be necessary to recover x from y using the approximation

$$\hat{x}_j \approx \begin{cases} \hat{y}_j/\hat{k}_j & \text{for } |\hat{k}_j|>\epsilon \\ 0 & \text{otherwise.} \end{cases}$$

Modify the program in the previous part to compute the relative error in the recovered value of x when $\epsilon=10^{-6}$ and when $\epsilon=0.01$.

Part 3. In the 1940s Wiener and Kolmogorov derived an optimal filter to approximate x from y based on two assumptions:

1. The noise is statistically independent of the signal.
2. The noise has a constant power spectrum across the entire bandwidth of the signal.

Note that these two assumptions actually define white noise. In this case

$$\hat{x}_j \approx \frac{\hat{y}_j\,\overline{\hat{k}_j}}{|\hat{k}_j|^2+\epsilon^2}$$

will optimally recover x from the corrupted signal y. Here ϵ plays the same role as before but is now interpreted in terms of the power spectrum of the noise.

Write a program to recover x using the Wiener–Kolmogorov filter. Let X_ϵ be the approximation of x obtained for a particular choice of ϵ. Find

$$\min\left\{\, E(X_\epsilon,x) : \epsilon>0 \,\right\}$$

using whatever technique you prefer. What value of ϵ gives the smallest error and most accurate approximation? How much better is this optimal filter than the method used to recover x in the previous part of this project?