1 Course Overview

"In his whole life man achieves nothing so great and so wonderful as what he achieved when he learned to talk."– Otto Jespersen

1.1 Course Syllabus

Please see syllabus handout. The syllabus for EE589 Digital Speech Processing is also available at:

http://www.ece.nmsu.edu/~pdeleon/Teaching/EE589/Syllabus.html

1.2 Course Outline

Please see outline handout. The outline for EE589 Digital Speech Processing is also available at:

http://www.ece.nmsu.edu/~pdeleon/Teaching/EE589/Outline.html

2 Lecture Outline

Reading: Chapters 1 and 2. Notes from your previous DSP class.

We will quickly review the following DSP topics

- Fourier Transforms
- Sampling and Reconstruction

It is expected that the student is familiar and comfortable with this material.

3 Fourier Transforms

3.1 Discrete-Time Fourier Transform (DTFT)

A discrete-time (DT) signal can have a representation using a basis function, $e^{j\omega n}$ where the independent variable, ω is *continuous* and represents frequency in units of radians per sample. The form is as follows

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)\, e^{j\omega n} d\omega \tag{1}$$

where

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n]\, e^{-j\omega n}. \tag{2}$$

In this representation, we have a continuous, complex-valued function, $X(\omega)$, known as the *spectrum*, which weights our basis function, $e^{j\omega n}$. The spectrum is periodic with a period of 2π rads/sample and for real-valued signals, the signal bandwidth is π rads/sample. The DT frequency variable, ω can be converted to units of Hertz (Hz) using

$$f = \frac{\omega}{2\pi} f_s \tag{3}$$

where f_s is the sample rate.

Example: Consider the rectangular pulse defined by (L odd)

$$x[n] = \begin{cases} 1, & |n| \leq (L-1)/2 \\ 0, & |n| > (L-1)/2. \end{cases} \tag{4}$$

The DTFT is computed as

Figure 1: Discrete-time rectangular pulse of length L.

$$\begin{aligned} X(\omega) &= \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega t} \\ &= \sum_{n=-(L-1)/2}^{(L-1)/2} e^{-j\omega n} \\ &= \frac{\sin(\omega L/2)}{\sin(\omega/2)}. \end{aligned} \tag{5}$$

Figure 2: DTFT of discrete-time rectangular pulse of length L.

3.2 Discrete Fourier Transform (DFT)

A finite-length DT signal can have a representation using a basis function, $e^{j\omega_k n}$ where the independent variable, ω_k is *discrete* and represents frequency in units of radians per sample. Consider the N-point DFT of a length L signal. The form is as follows

$$x[n] = \sum_{k=0}^{N-1} X[k] e^{j\omega_k n}, 0 \leq n \leq L-1 \tag{6}$$

where

$$X[k] = \sum_{n=0}^{L-1} x[n]\, e^{-j\omega_k n}, 0 \le k \le N-1 \tag{7}$$

with frequency points

$$\omega_k \equiv \frac{2\pi k}{N} \tag{8}$$

i.e., the DFT frequencies are uniformly spaced over the 2π signal bandwidth. For convenience, we refer to ω_k as the "kth DFT frequency."

In this representation, we have a discrete, complex-valued sequence, $X[k]$, known as the *spectrum*, which weights our basis function, $e^{j\omega_k n}$. The DT frequency variable, ω_k can be converted to units of Hertz (Hz) using

$$f_k = \frac{\omega_k}{2\pi} f_s \tag{9}$$

where f_s is the sample rate. Let the Nth root of unity (often called a twiddle factor) be denoted as

$$W_N \equiv e^{-j2\pi/N} \tag{10}$$

Our DFT equations can then also be written as

$$x[n] = \sum_{k=0}^{N-1} X[k]\, W_N^{-kn}, 0 \le n \le L-1 \tag{11}$$

where

$$X[k] = \sum_{n=0}^{L-1} x[n]\, W_N^{kn}, 0 \le k \le N-1. \tag{12}$$

Example: Consider the rectangular pulse defined by (L odd)

$$x[n] = \begin{cases} 1, & |n| \le (L-1)/2 \\ 0, & |n| > (L-1)/2. \end{cases} \tag{13}$$

The DFT is computed as

Figure 3: Discrete-time rectangular pulse of length L.

$$\begin{aligned} X[k] &= \sum_{n=-(L-1)/2}^{(L-1)/2} x[n]\, W_N^{kn} \\ &= \sum_{n=-(L-1)/2}^{(L-1)/2} W_N^{kn} \\ &= \frac{\sin(\pi k L/N)}{\sin(\pi k/N)}. \end{aligned} \tag{14}$$

Figure 4: DFT of discrete-time rectangular pulse of length L.

We can consolidate the above N DFT equations in (12) in matrix form

$$\begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[N-1] \end{bmatrix} = \begin{bmatrix} W_N^{(0)(0)} & W_N^{(0)(1)} & \cdots & W_N^{(0)(L-1)} \\ W_N^{(1)(0)} & W_N^{(1)(1)} & \cdots & W_N^{(1)(L-1)} \\ \vdots & \vdots & \ddots & \vdots \\ W_N^{(N-1)(0)} & W_N^{(N-1)(1)} & \cdots & W_N^{(N-1)(L-1)} \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[L-1] \end{bmatrix} \tag{15}$$

or

$$\begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[N-1] \end{bmatrix} = \begin{bmatrix} W_N^{0} & W_N^{0} & \cdots & W_N^{0} \\ W_N^{0} & W_N^{1} & \cdots & W_N^{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ W_N^{0} & W_N^{N-1} & \cdots & W_N^{(N-1)(L-1)} \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[L-1] \end{bmatrix} \tag{16}$$

or

$$\mathbf{X} = \mathbf{W}\mathbf{x}. \tag{17}$$

The N-point DFT can be thought of as a linear matrix (unitary) transformation (using $\mathbf{W}$, the DFT matrix) of the L-dimensional vector of time samples into an N-dimensional vector of frequency samples. Assuming $N = L$, i.e., $\mathbf{W}$ is a square matrix, then the inverse DFT (IDFT) is given by

$$\begin{aligned} \mathbf{x} &= \frac{1}{N}\mathbf{W}^{-1}\mathbf{X} \\ &= \frac{1}{N}\mathbf{W}^{H}\mathbf{X} \end{aligned} \tag{18}$$

where ^{H} denotes the Hermitian or complex conjugate transpose. The last line results from the (almost) unitary property of the DFT matrix.

Figure 5: Figure Fourier transforms as a change of basis

Figure 6: Perspectives in Fourier Analysis

3.3 Fast Fourier Transform (FFT)

The FFT is an efficient algorithm for computing the DFT assuming $N = L$ and N is a power of two. See any standard DSP text for details about this algorithm.

4 Sampling and Reconstruction

4.1 Introduction

The typical method of obtaining a DT signal from a CT signal is through periodic sampling. Here a sequence of samples, $x[n] = x_a(nT_s)$ is obtained from a CT signal $x_a(t)$ where T_s is the sampling period and its reciprocal f_s is the sampling frequency in samples/second.

In general, it is not possible to reconstruct $x_a(t)$ from the samples $x[n]$ since many CT signals can produce the same output sequence of samples. Under certain special situations, however, perfect reconstruction (no aliasing) of a CT signal from its samples *is possible* (ignoring quantization effects for now).

Theorem 1 *(Nyquist 1928) Suppose that $x_a(t)$ is sampled at a rate of $f_s = 1/T_s$ samples per second and suppose $x_a(t)$ is bandlimited to $f_{\max}$. Then $x_a(t)$ can be uniquely determined from its uniformly spaced samples, $x[n] = x_a(nT_s)$ if*

$$f_s \quad > \quad 2f_{\max}. \tag{19}$$

We call $f_{\max}$ the "Nyquist frequency" and $2f_{\max}$ the "Nyquist rate."

4.2 A Mathematical Foundation for Sampling

The basis for the sampling theorem is that sampling the signal, $x_a(t)$ at a rate f_s results in spectral duplicates (images) spaced by $2/T_s$ Hz. Sampling at the Nyquist rate or higher avoids aliasing (overlap and distortion of spectral duplicates), thus preserving the spectral integrity of the signal.

Figure 7: Spectrum of $x_a(t)$ and spectra of $x[n]$ for two different sample rates

We can develop the ideal sampling and reconstructions relations mathematically. Apply the "sampling function,"

$$p(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT_s) \tag{20}$$

to the CT signal to arrive at a "sampled signal"

$$\begin{aligned} x_p(t) &= x_a(t)p(t) \\ &= \sum_{k=-\infty}^{\infty} x_a(kT_s)\delta(t - kT_s). \end{aligned} \tag{21}$$

We let $x[n] = x_p(nT_s)$.

Figure 8: Sampling function and the "sampled signal"

The question to ask is "what does the spectrum of the sampled signal look like?" i.e., what is $X_p(\Omega)$? We have as the spectrum of the sampling function

$$P(\Omega) = \frac{2\pi}{T_s} \sum_{k=-\infty}^{\infty} \delta(\Omega - k\Omega_s) \tag{22}$$

where $\Omega_s = 2\pi f_s$. By computing the FT of (21) and making use of (22) and the window and sifting properties, we can show

$$\begin{aligned} X_p(\Omega) &= X_a(\Omega) * P(\Omega) \\ &= \frac{1}{T_s}\sum_{k=-\infty}^{\infty} X_a(\Omega - k\Omega_s). \end{aligned} \tag{23}$$

Figure 9: Spectrum of the sampled signal

When sampling at or above the Nyquist rate, the spectral images do not overlap and thus the signal spectrum is preserved. We can relate the spectrum of $x[n]$ or $X(e^{j\omega})$ to the spectrum of $x_a(t)$ or $X_a(\Omega)$ with

$$X(e^{j\omega}) = \frac{1}{T_s} X_a(\Omega T), \quad |\omega| \le \pi. \tag{24}$$

Note that $\Omega = \omega/T_s = \omega f_s$ (conversions are $\Omega = 2\pi f$ rads/s and $\omega = 2\pi f/f_s$ rads/sample). Sampling normalizes the time-axis from t in seconds to n in time instants spaced T_s seconds apart. Sampling also normalizes the signal bandwidth from $-f_s/2 \le \Omega \le f_s/2$ in Hz (cycles per second) to $-\pi \le \omega \le \pi$ in Hz (cycles per sample).

4.3 A Mathematical Foundation for Reconstruction

The basis for ideal reconstruction is to filter the central spectral image from the spectrum of the sampled signal. Since the spectral image is at baseband, this amounts to a lowpass filtering. In the time-domain, we can show that this filtering amounts to attaching the impulse response of the ideal lowpass filter, $h(t)$ to each sample (interpolation).

We first assume the DT signal $x[n]$ is "converted" to a CT signal $x_p(t)$ via

$$x_p(t) = x[n]p(t). \tag{25}$$

For the reconstructed signal, $x_r(t)$ we have

$$\begin{aligned} x_r(t) &= x_p(nT_s) * h(t) \\ &= \sum_{n=-\infty}^{\infty} x_p(nT_s)\frac{\sin[\pi(t-nT_s)/T_s]}{\pi(t-nT_s)/T_s}. \end{aligned} \tag{26}$$

Figure 10: Perfect reconstruction