Lecture #7: Discrete-time Signals and Sampling

1. Introduction

Unlike continuous-time signals, discrete-time signals have defined values only at a fixed number of points in time; discrete-time signals typically arise when we *sample* a continuous-time signal. In this class, we will consider discrete-time signals whose defined values occur only at regularly spaced intervals, in other words, discrete-time signals with a constant *sampling frequency*. To convert a continuous-time signal $x_c(t)$ to a discrete-time signal $x[n]$, we can use the following equation:

$$x[n] = x_c(nT_s) = x(n/f_s) \tag{1}$$

where T_s denotes the *sampling period*, f_s denotes the *sampling frequency*, and,

$$f_s = \frac{1}{T_s}. \tag{2}$$

Note that in equation (1) we define the discrete-time signal $x[n]$ as a function of the dimensionless time index n, $n \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. Figure 1 plots the discrete-time sequences $x[n]$ of tow continuous-time signals $x(t)$, and a sampling frequency of $f_s = 5$ Hertz (samples/second).

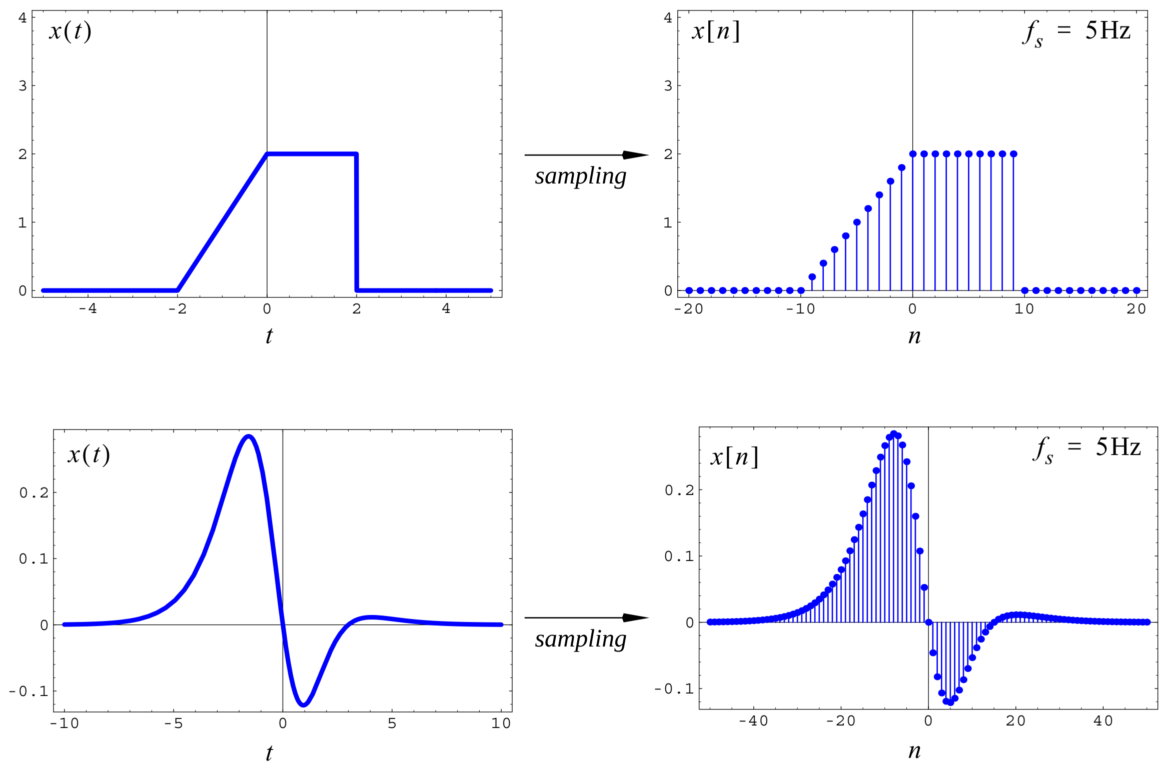

Figure 1

If we want to reconstruct a continuous-time signal from a discrete-time signal, knowledge of the sampling frequency is critically important, since without it, we do not know to which real time t each time index n corresponds. For example, music recorded on a CD is stored as a sequence of numbers. By international agreement, all music stored on CDs is sampled at 44.1 kHz; without this critical piece of information, the stored music could not be faithfully reconstructed in the continuous-time domain (on your stereo). Computer music files, such as mp3's, have their sampling frequency encoded in the file itself. If the encoded sampling frequency in the mp3 file is incorrect for some reason, the music that your computer would play would sound either too fast or too slow. To demonstrate this point, I took a short piece of music (from Kenny Roger's "The Gambler"), sampled at 32kHz, and played the original music during lecture, assuming that the sampling frequency was 24kHz (incorrect), 32kHz (correct) and 40kHz (incorrect). The piece of music played back with the incorrect sampling frequency of

24kHz sounds slowed down, while it sounds sped up with the incorrect sampling frequency of 40kHz. All three music files are posted on the web site in *wav* and *mp3* formats.

2. Sampling

A. How fast is enough?

An extremely important question that must be answered in the sampling process is "How fast is enough?" That is, how fast do we have to sample a continuous-time signal in order to be able to reconstruct it from just the discrete-time signal and knowledge of the sampling frequency? A truly remarkable theorem, the *Shannon Sampling Theorem*, answers this question:

> ***Shannon Sampling Theorem**: A continuous-time signal $x_c(t)$ with frequencies no higher than f_{max} can be reconstructed exactly from its samples $x[n] = x_c(nT_s)$, if the samples are taken at a sampling frequency $f_s > 2f_{max}$, that is, at a sampling frequency greater than $2f_{max}$. The frequency $2f_{max}$ is known as the Nyquist frequency.*

We have already seen that any signal can be converted into its frequency domain representation, which tells us what frequencies are contained within that signal. Now, we begin to see one reason why that frequency-domain representation is so important. It allows us to determine (or filter our signal) so that we know f_{max}; with that knowledge, we can confidently sample at a frequency greater than the Nyquist frequency and *lose no information in that sampling process*.

As an example, consider a 1Hz cosine wave:

$$x_c(t) = \cos(2\pi t) \tag{3}$$

For this signal, $f_{max} = 1$ Hz, so that the Nyquist frequency is 2Hz. In Figure 2, we plot the sequences $x[n]$ corresponding to eight different sampling frequencies: 100Hz, 10Hz, 5Hz, 3Hz, 2.1Hz, 1.9Hz, 1.4Hz and 1.1Hz. The Sampling Theorem tells us that we should be able to reconstruct $x_c(t)$ from just the samples for all sampling frequencies greater than 2Hz (top five plots in Figure 2). This seems surprising given the relative sparsity of data for sampling frequencies 5Hz, 3Hz and 2.1Hz. As you will remember, most of the class thought that the 100Hz sampling frequency would be sufficient for reconstructing the original waveform; some of the class thought the 10Hz sampling frequency would be sufficient; however, very few thought that any of the lower frequencies would be sufficient for perfect reconstruction of the original signal. Therefore, the Sampling Theorem is not only remarkably powerful, but also remarkably surprising.

B. Ideal reconstruction of discrete-time signal

Note that the Sampling Theorem does not tell us specifically how we should reconstruct the continuous-time signal from the discrete-time samples. Clearly, if we are to reconstruct the original sine wave in Figure 2 from the samples corresponding to the smaller sampling frequencies above 2Hz, we cannot simply use linear interpolation. Figure 3 shows, for example, two naive reconstructions of the cosine wave for the sampling frequency $f_s = 2.1$ Hz. The reconstruction on the left simply approximates the reconstructed function $x_r(t)$ as a sequence of pulses spaced T_s seconds apart. The reconstruction on the right linearly interpolates between consecutive samples; neither reconstruction could, in any sense of the phrase, be considered a "perfect reconstruction" of the original signal.

To achieve the ideal reconstruction, each sample should be the weighted and time-shifted peak of the sinc function,[1]

$$\text{sinc}(t) = \frac{\sin(t)}{t} \tag{4}$$

plotted in Figure 4 below. Specifically, the ideal reconstruction is given by,

1. We will see later why this is so.

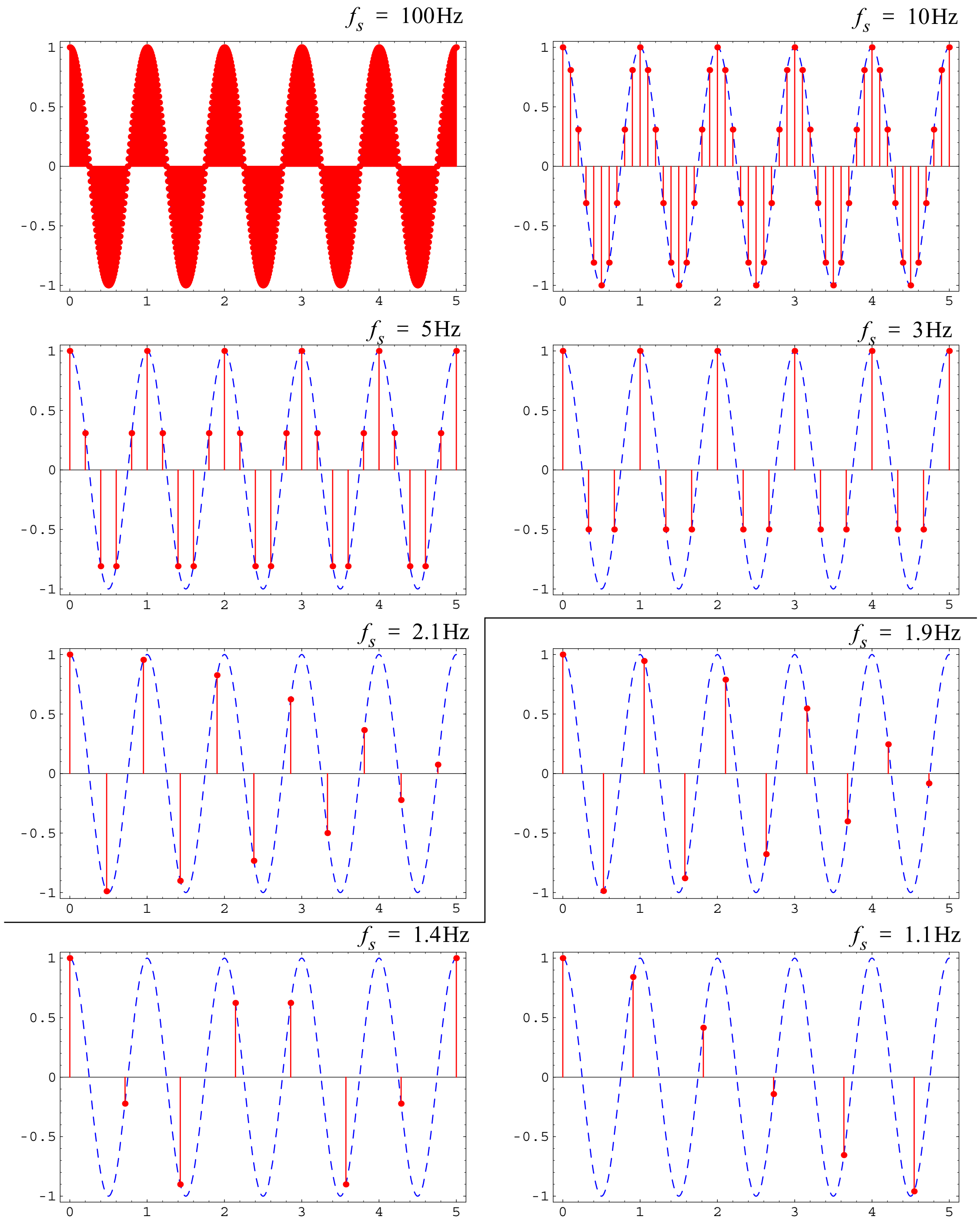

Figure 2: A 1Hz cosine wave and different sampling frequencies

$$x_r(t) = \sum_n x[n]\mathrm{sinc}[\pi f_s(t - n/f_s)]\,. \qquad (5)$$

where $x_r(t)$ refers to the reconstructed, continuous-time signal. You may wonder why the value of the $\mathrm{sinc}(t)$ function is 1 at $t = 0$, since both the numerator and denominator of equation (4) is zero at $t = 0$. When we have such an undefined ratio, we can use *L'Hospital's Rule* (from calculus) to determine the value at $t = 0$:

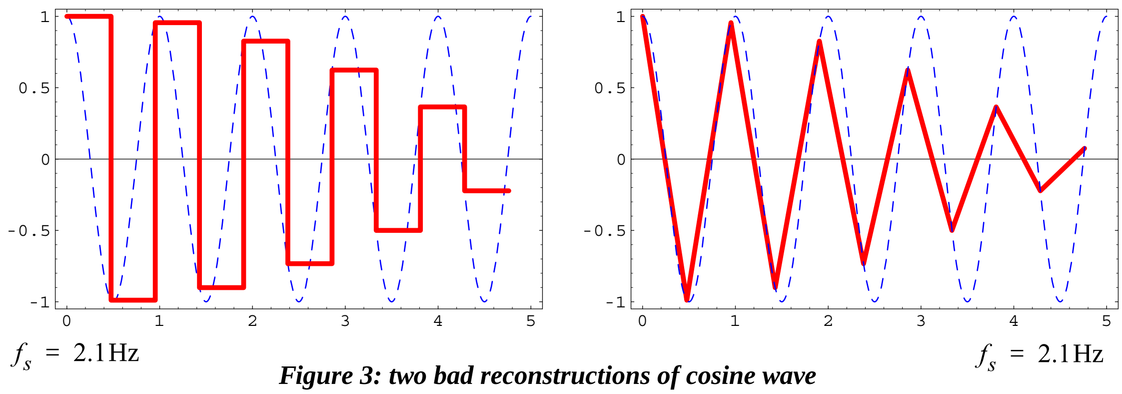

Figure 3: two bad reconstructions of cosine wave

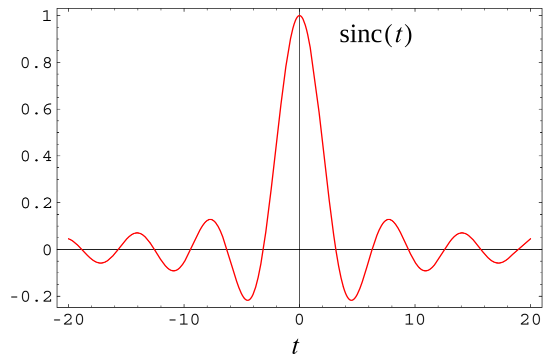

Figure 4

$$\lim_{t \to 0} \text{sinc}(t) = \lim_{t \to 0} \frac{\sin(t)}{t} = \lim_{t \to 0} \frac{\frac{d}{dt}\sin(t)}{\frac{d}{dt}(t)} = \lim_{t \to 0} \frac{\cos(t)}{1} = \lim_{t \to 0} \cos(t) = 1 . \quad (6)$$

Now, in Figure 5 below we plot this ideal reconstruction for sampling frequencies f_s = 5Hz, 3Hz, 2.1Hz, 1.9Hz, 1.4Hz and 1.1Hz, using equation (5) above. Note that for sampling frequencies greater than 2Hz (the Nyquist frequency for the 1Hz signal), the reconstruction is very good, while for sampling frequencies less than 2Hz, the reconstructed signal $x_r(t)$ does not approximate $x_c(t)$ in equation (3) very well.

At this point, the observant reader will notice that the reconstruction for $f_s = 2.1$ Hz, while good, is not perfect. The reason for this is that we are not actually reconstructing,

$$x_c(t) = \cos(2\pi t) \text{ [equation (3)]} \quad (7)$$

but instead are reconstructing the *time-limited* continuous-time signal below:

$$x_c(t) = \cos(2\pi t)[u(t) - u(t-5)] . \quad (8)$$

As we will see later, the continuous-time magnitude frequency rerpresentation $|X_c(f)|$ for (8) above is actually not frequency-limited such that $f_{max} = 1$ Hz, as illustrated in Figure 6, which plots $|X_c(f)|$ for the time-limited signal (8). Note that while the frequency content of the time-limited cosine waveform still has dominant peaks at ±1 Hz, time limiting the cosine causes the frequency spectrum to spread out over the entire frequency spectrum. Therefore, the 2.1Hz sampling frequency is strictly speaking not larger than $2f_{max}$, so that the reconstruction in Figure 5 is not perfect.

Furthermore, there is a practical problem with the reconstruction in equation (5); the value of the reconstructed signal $x_r(t)$ is dependent on all samples for all values of t. For example, suppose we wanted to

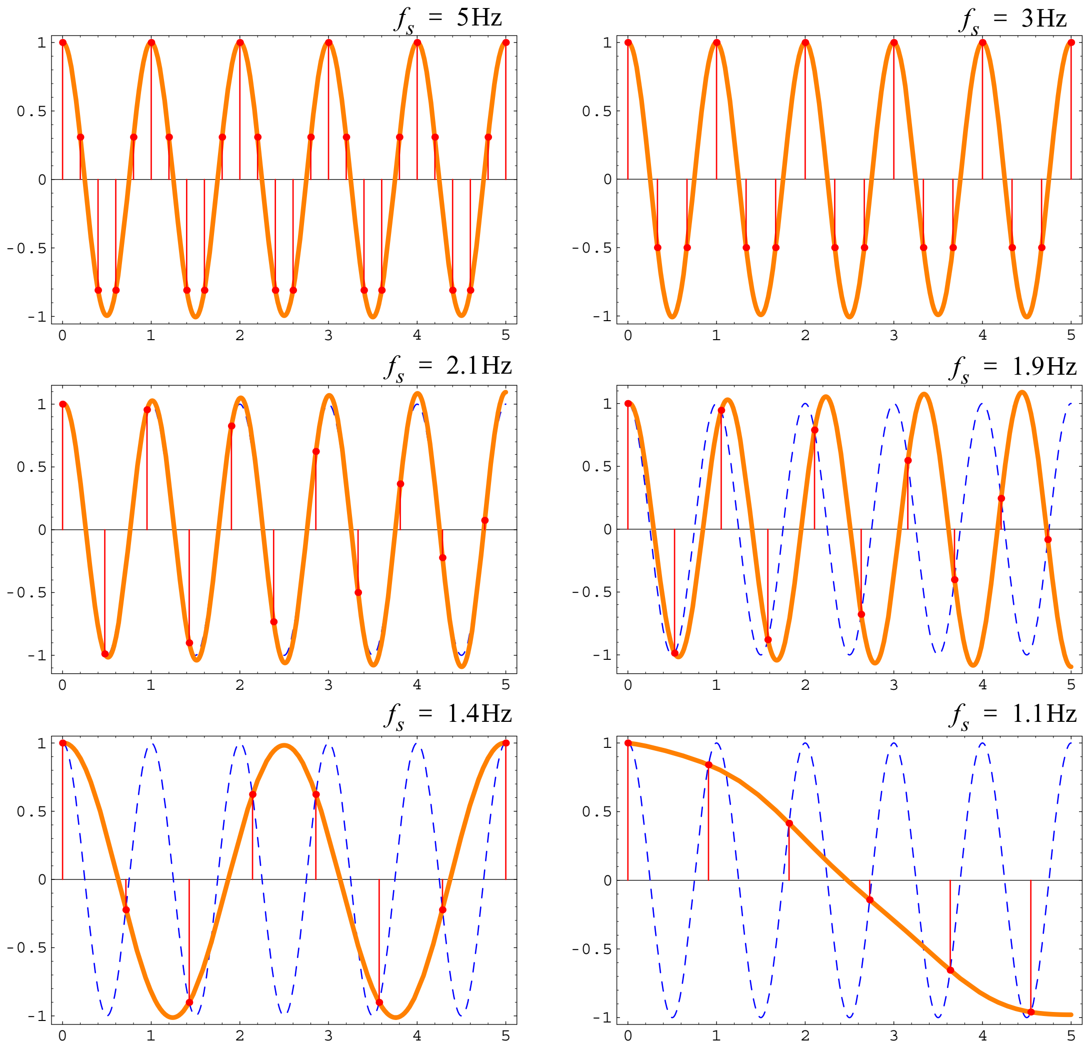

Figure 5: Ideal reconstructions from samples

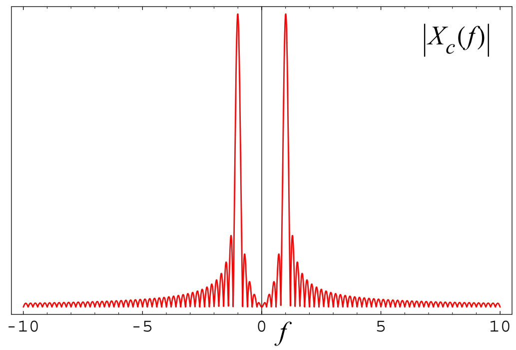

Figure 6: Time-limited frequency spectrum of 1Hz cosine signal

know the value of the original signal for $t = 2$; in the ideal reconstruction, $x_r(2)$ is dependent not only on samples close to $t = 2$, but all samples, no matter how distant (in time) from $t = 2$.

Therefore, as a practical matter, the ideal reconstruction is not realizable in real systems. We can compensate for this fact by *oversampling*; that is sampling a signal $x_c(t)$ at a frequency significantly larger than the Nyquist frequency. Let us again consider the example of CDs. The human ear cannot hear frequencies higher than approximately 20 kHz; a sampling frequency of 44.1 kHz is therefore more than $2f_{max}$, where f_{max} is the maximum frequency that we should be interested in for an audio system intended for human enjoyment.

C. Conclusion

Next time, we will continue our investigation of the sampling process to understand what is so critical about the Nyquist frequency. Not surprisingly, our development will rely on the transformation between the time and frequency domain. Finally, all the figures related to sampling were generated with the *Mathematica* notebook “intro_sampling.nb.”