Lecture 23: Aliasing in Frequency: the Sampling Theorem

Mark Hasegawa-Johnson

ECE 401: Signal and Image Analysis, Fall 2020

Outline

Two-sided spectrum

The **spectrum** of $x(t)$ is the set of frequencies, and their associated phasors,

$$\text{Spectrum}\left(x(t)\right) = \{(f_{-N}, a_{-N}), \ldots, (f_0, a_0), \ldots, (f_N, a_N)\}$$

such that

$$x(t) = \sum_{k=-N}^{N} a_k e^{j2\pi f_k t}$$

Fourier's theorem

One reason the spectrum is useful is that **any** periodic signal can be written as a sum of cosines. Fourier's theorem says that any $x(t)$ that is periodic, i.e.,

$$x(t+T_0) = x(t)$$

can be written as

$$x(t) = \sum_{k=-\infty}^{\infty} X_k e^{j2\pi k F_0 t}$$

which is a special case of the spectrum for periodic signals: $f_k = kF_0$, and $a_k = X_k$, and

$$F_0 = \frac{1}{T_0}$$

Fourier Series

- **Analysis** (finding the spectrum, given the waveform):

$$X_k = \frac{1}{T_0}\int_0^{T_0} x(t)e^{-j2\pi kt/T_0}dt$$

- **Synthesis** (finding the waveform, given the spectrum):

$$x(t) = \sum_{k=-\infty}^{\infty} X_k e^{j2\pi kt/T_0}$$

Outline

How to sample a continuous-time signal

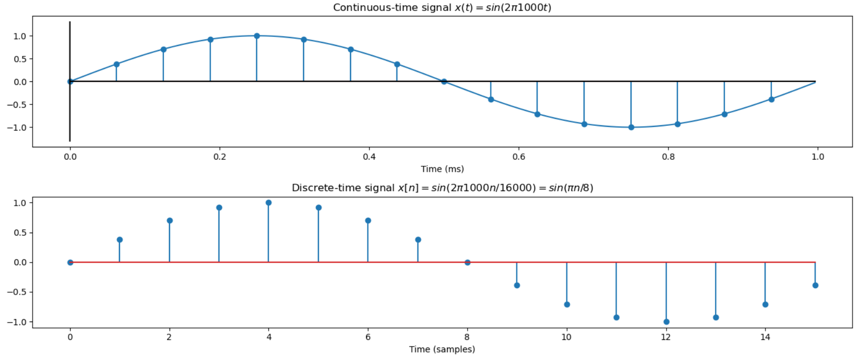
Suppose you have some continuous-time signal, $x(t)$, and you'd like to sample it, in order to store the sample values in a computer. The samples are collected once every $T_s = \frac{1}{F_s}$ seconds:

$$x[n] = x(t = nT_s)$$

Example: a 1kHz sine wave

For example, suppose $x(t) = \sin(2\pi 1000 t)$. By sampling at $F_s = 16000$ samples/second, we get

$$x[n] = \sin\left(2\pi 1000 \frac{n}{16000}\right) = \sin(\pi n/8)$$



Outline

Can every sine wave be reconstructed from its samples?

The question immediately arises: can every sine wave be reconstructed from its samples?
The answer, unfortunately, is "no."

Can every sine wave be reconstructed from its samples?

For example, two signals $x_1(t)$ and $x_2(t)$, at 10kHz and 6kHz respectively:

$$x_1(t) = \cos(2\pi 10000t), \quad x_2(t) = \cos(2\pi 6000t)$$

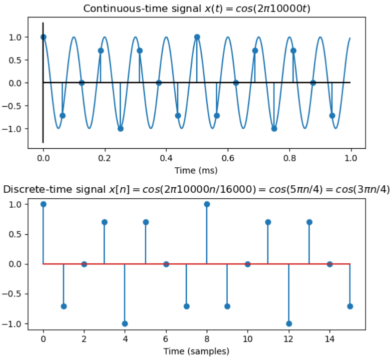
Let's sample them at $F_s = 16,000$ samples/second:

$$x_1[n] = \cos\left(2\pi 10000 \frac{n}{16000}\right), \quad x_2[n] = \cos\left(2\pi 6000 \frac{n}{16000}\right)$$

Simplifying a bit, we discover that $x_1[n] = x_2[n]$. We say that the 10kHz tone has been "aliased" to 6kHz:

$$x_1[n] = \cos\left(\frac{5\pi n}{4}\right) = \cos\left(\frac{3\pi n}{4}\right)$$
$$x_2[n] = \cos\left(\frac{3\pi n}{4}\right) = \cos\left(\frac{5\pi n}{4}\right)$$

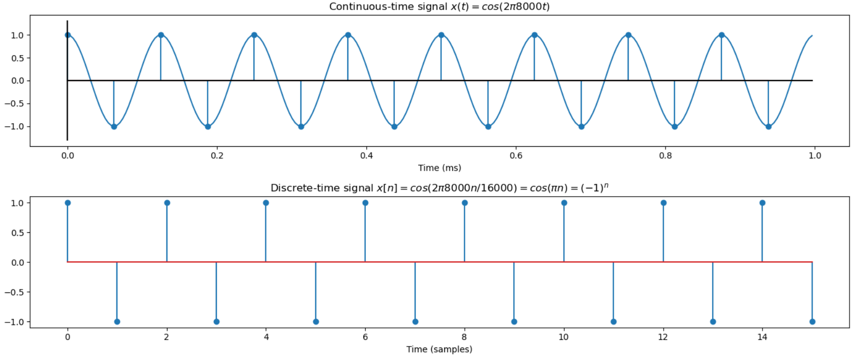
Can every sine wave be reconstructed from its samples?



What is the highest frequency that can be reconstructed?

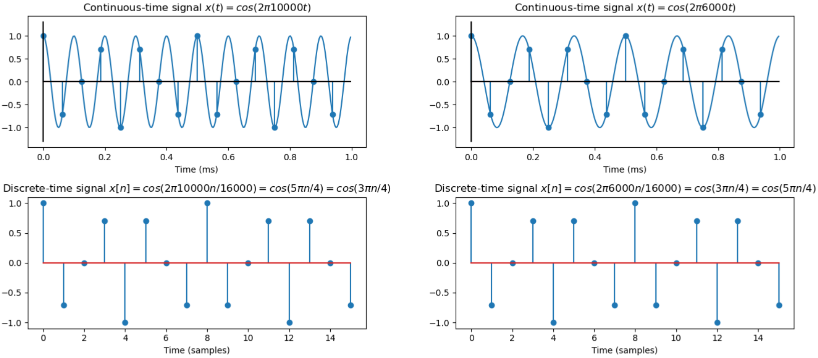
The highest frequency whose cosine can be exactly reconstructed from its samples is called the "Nyquist frequency," $F_N = F_S/2$. If $x(t) = \cos(2\pi F_N t)$, then

$$x[n] = \cos\left(2\pi F_N \frac{n}{F_S}\right) = \cos(\pi n) = (-1)^n$$



Sampling above Nyquist ⇒ Aliasing to a frequency below Nyquist

If you try to sample a signal whose frequency is above Nyquist (like the one shown on the left), then it gets **aliased** to a frequency below Nyquist (like the one shown on the right).



Outline

General characterization of continuous-time signals

Let's assume that $x(t)$ is periodic with some period T_0, therefore it has a Fourier series:

$$x(t) = \sum_{k=-\infty}^{\infty} X_k e^{j2\pi kt/T_0} = \sum_{k=0}^{\infty} |X_k| \cos\left(\frac{2\pi kt}{T_0} + \angle X_k\right)$$

Eliminate the aliased tones

We already know that $e^{j2\pi kt/T_0}$ will be aliased if $|k|/T_0 > F_N$. So let's assume that the signal is **band-limited:** it contains no frequency components with frequencies larger than $F_S/2$.

That means that the only X_k with nonzero energy are the ones in the range $-N/2 \leq k \leq N/2$, where $N = F_S T_0$.

$$x(t) = \sum_{k=-N/2}^{N/2} X_k e^{j2\pi kt/T_0} = \sum_{k=0}^{N/2} |X_k| \cos\left(\frac{2\pi kt}{T_0} + \angle X_k\right)$$

Sample that signal!

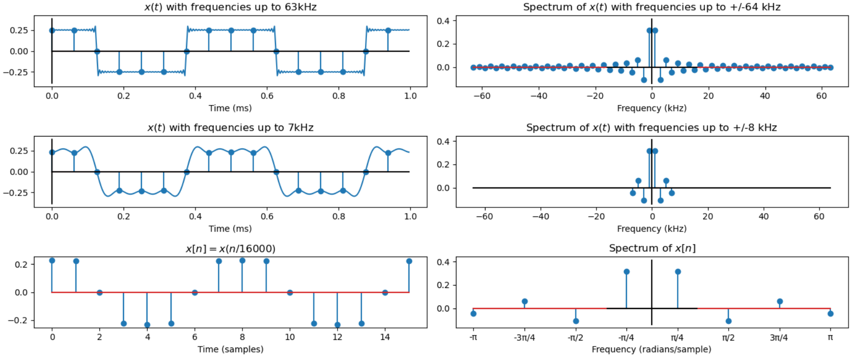
Now let's sample that signal, at sampling frequency F_S:

$$x[n] = \sum_{k=-N/2}^{N/2} X_k e^{j2\pi kn/F_S T_0} = \sum_{k=0}^{N/2} |X_k| \cos\left(\frac{2\pi kn}{N} + \angle X_k\right)$$

So the highest digital frequency, when $k = F_S T_0/2$, is $\omega_k = \pi$. The lowest is $\omega_0 = 0$.

$$x[n] = \sum_{\omega_k=-\pi}^{\pi} X_k e^{j\omega_k n} = \sum_{\omega_k=0}^{\pi} |X_k| \cos\left(\omega_k n + \angle X_k\right)$$

Spectrum of a sampled periodic signal



The sampling theorem

As long as $-\pi \leq \omega_k \leq \pi$, we can recreate the continuous-time signal by just regenerating a continuous-time signal with the corresponding frequency:

$$f_k \left[\frac{\text{cycles}}{\text{second}}\right] = \frac{\omega_k \left[\frac{\text{radians}}{\text{sample}}\right] \times F_S \left[\frac{\text{samples}}{\text{second}}\right]}{2\pi \left[\frac{\text{radians}}{\text{cycle}}\right]}$$

$$x[n] = \cos(\omega_k n + \theta_k) \quad \rightarrow \quad x(t) = \cos(2\pi f_k t + \theta_k)$$

The sampling theorem

A continuous-time signal $x(t)$ with frequencies no higher than f_{max} can be reconstructed exactly from its samples $x[n] = x(nT_S)$ if the samples are taken at a rate $F_s = 1/T_s$ that is greater than $2f_{max}$.

Outline

How can we get $x(t)$ back again?

We've already seen one method of getting $x(t)$ back again: we can find all of the cosine components, and re-create the corresponding cosines in continuous time.

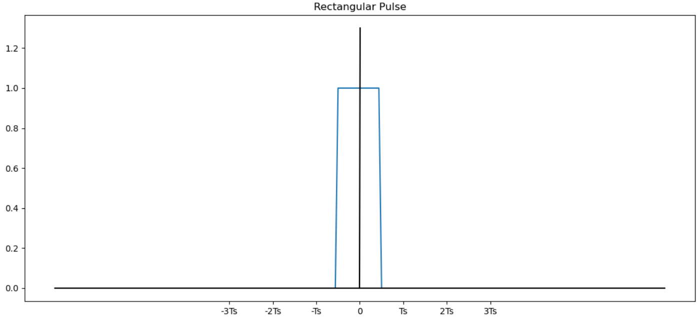
There is an easier way. It involves multiplying each of the samples, $x[n]$, by a short-time pulse, $p(t)$, as follows:

$$y(t) = \sum_{n=-\infty}^{\infty} y[n]p(t - nT_s)$$

Rectangular pulses

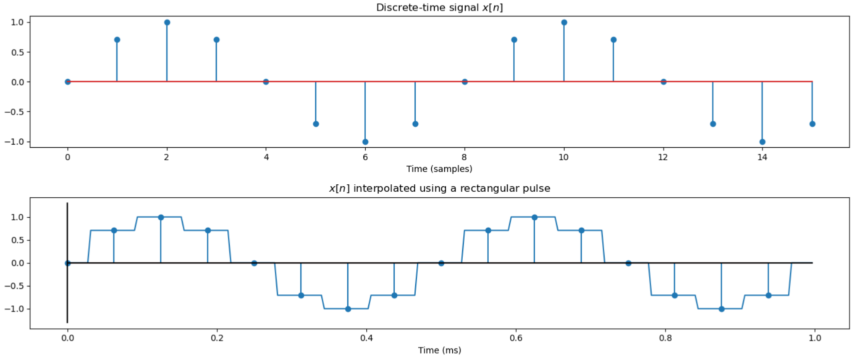
For example, suppose that the pulse is just a rectangle,

$$p(t) = \begin{cases} 1 & -\frac{T_S}{2} \le t < \frac{T_S}{2} \\ 0 & \text{otherwise} \end{cases}$$



Rectangular pulses = Piece-wise constant interpolation

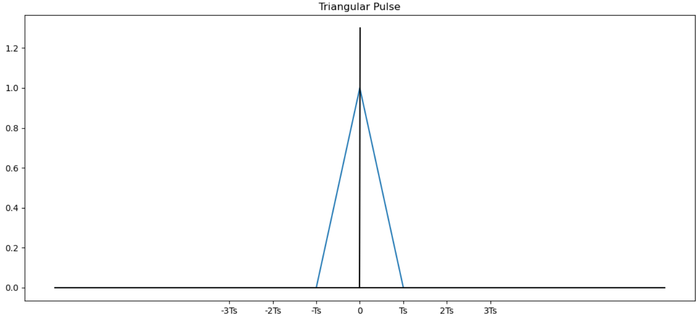
The result is a piece-wise constant interpolation of the digital signal:



Triangular pulses

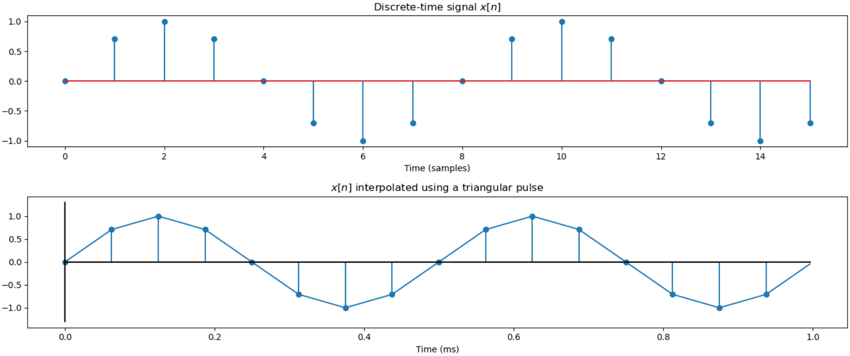
The rectangular pulse has the disadvantage that $y(t)$ is discontinuous. We can eliminate the discontinuities by using a triangular pulse:

$$p(t) = \begin{cases} 1 - \frac{|t|}{T_S} & -T_S \leq t < T_S \\ 0 & \text{otherwise} \end{cases}$$



Triangular pulses = Piece-wise linear interpolation

The result is a piece-wise linear interpolation of the digital signal:



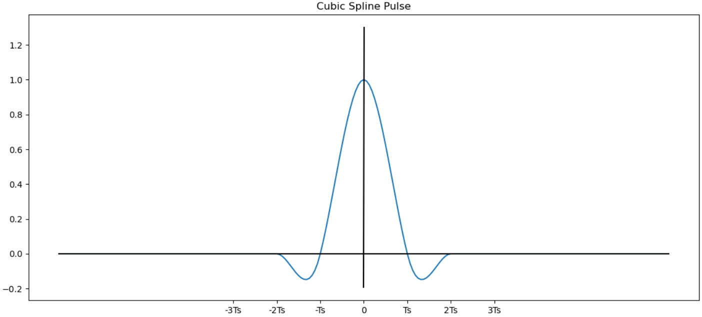
Cubic spline pulses

The triangular pulse has the disadvantage that, although $y(t)$ is continuous, its first derivative is discontinuous. We can eliminate discontinuities in the first derivative by using a cubic-spline pulse:

$$p(t) = \begin{cases} 1 - 2\left(\frac{|t|}{T_S}\right)^2 + \left(\frac{|t|}{T_s}\right)^3 & -T_S \le t < T_S \\ -\left(2 - \frac{|t|}{T_S}\right)^2 + \left(2 - \frac{|t|}{T_S}\right)^3 & T_S \le |t| < 2T_S \\ 0 & \text{otherwise} \end{cases}$$

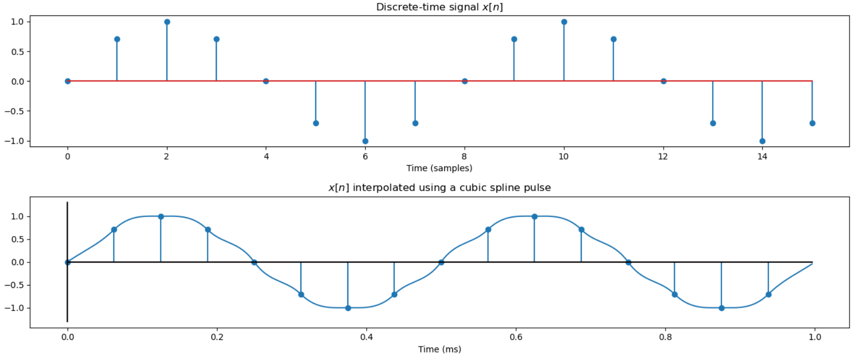
Cubic spline pulses

The triangular pulse has the disadvantage that, although $y(t)$ is continuous, its first derivative is discontinuous. We can eliminate discontinuities in the first derivative by using a cubic-spline pulse:



Cubic spline pulses = Piece-wise cubic interpolation

The result is a piece-wise cubic interpolation of the digital signal:



Sinc pulses

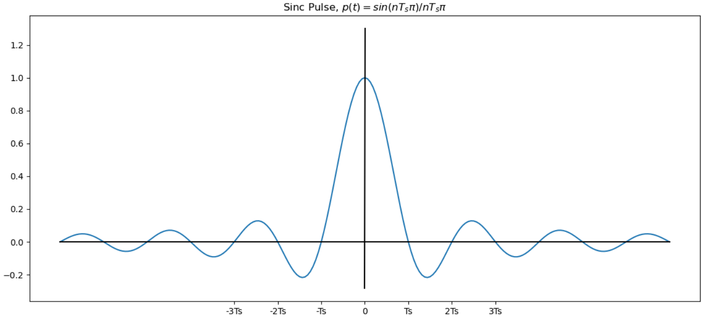
The cubic spline has no discontinuities, and no slope discontinuities, but it still has discontinuities in its second derivative and all higher derivatives. Can we fix those?

The answer: yes! The pulse we need is the inverse transform of an ideal lowpass filter, the sinc.

Sinc pulses

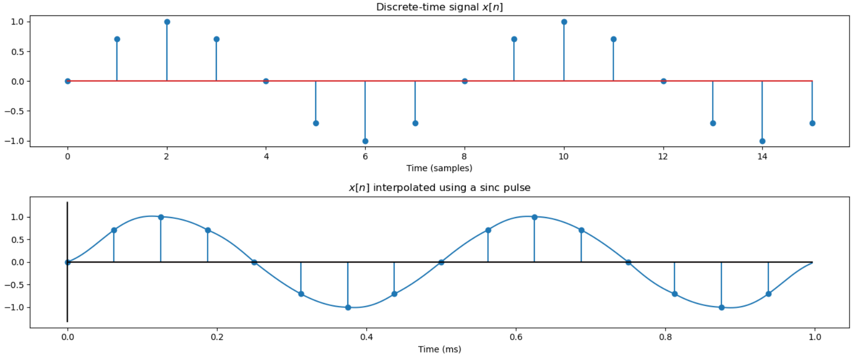
We can reconstruct a signal that has no discontinuities in any of its derivatives by using an ideal sinc pulse:

$$p(t) = \frac{\sin(\pi t/T_S)}{\pi t/T_S}$$



Sinc pulse = ideal bandlimited interpolation

The result is an ideal bandlimited interpolation:



Outline

Summary

A continuous-time signal $x(t)$ with frequencies no higher than f_{max} can be reconstructed exactly from its samples $x[n] = x(nT_S)$ if the samples are taken at a rate $F_s = 1/T_s$ that is greater than $2f_{max}$.

Ideal band-limited reconstruction is achieved using sinc pulses:

$$y(t) = \sum_{n=-\infty}^{\infty} y[n]p(t - nT_s), \quad p(t) = \frac{\sin(\pi t/T_S)}{\pi t/T_S}$$