The Sampling Theorem and Normalized Time/Frequency

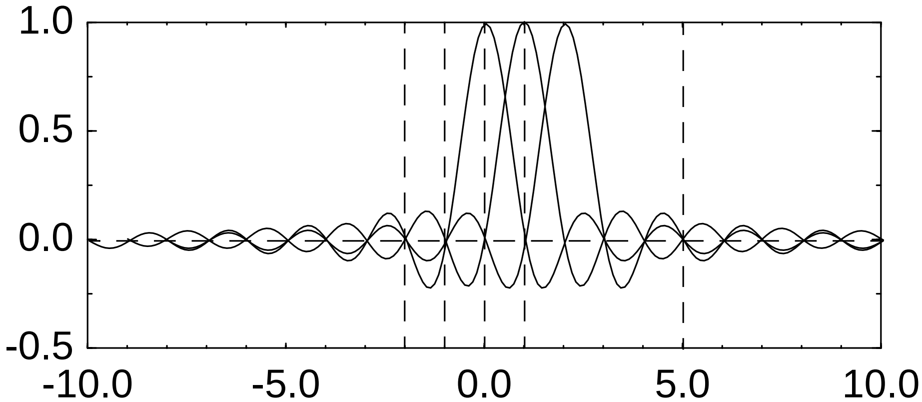
If the highest frequency contained in an analog signal, $x_a(t)$, is $F_{max} = B$ and the signal is sampled at a rate $F_s > 2F_{max} = 2B$, then $x_a(t)$ can be EXACTLY recovered from its sample values using the interpolation function:

$$g(t) = \frac{\sin(2\pi Bt)}{2\pi Bt} .$$

$x_a(t)$ may be expressed as:

$$x_a(t) = \sum_{n=-\infty}^{\infty} x_a\left(\frac{n}{F_s}\right) g\left(t - \frac{n}{F_s}\right)$$

where $x_a\left(\frac{n}{F_s}\right) = x_a(nT) = x(n)$.



Given a continuous signal:

$$x(t) = A\cos(2\pi ft + \theta) ,$$

A discrete-time sinusoid may be expressed as:

$$x(n) = A\cos\left(2\pi f\left(\frac{n}{f_s}\right) + \theta\right),$$

which, after regrouping, gives:

$$x(n) = A\cos(\omega n + \theta) ,$$

where $\omega = 2\pi\left(\frac{f}{f_s}\right)$, and is called normalized radian frequency and n represents normalized time.

Singularity Functions

Some Elementary Discrete-Time Signals:

(1) unit sample signal:

$$\delta(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

Using this function, we can write a mathematical expression for a sampled signal as:

$$x(n) = \sum_{k=-\infty}^{\infty} \left(x(t)\Big|_{t = \frac{k}{f_s}} \right) \delta(n-k)$$

Also, note that typical properties of continuous linear systems hold, such as the sifting property:

$$x(n) \otimes \delta(n-l) = \sum_{m=-\infty}^{\infty} x(m)\delta(l-m) = x(l)$$

(2) unit step signal:

$$u(n) = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

Using this function, we can write a mathematical expression for a finite duration segment of a signal, or window, of the signal:

$$w(n) = u(n) - u(N-n)$$
$$x_w(n) = x(n)w(n)$$

What is the impact of this on the spectrum of $x_w(n)$?

Energy and Power

Energy:

$$E \equiv \sum_{n=-\infty}^{\infty} |x(n)|^2$$

Average Power:

$$P = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |x(n)|^2$$

Finite Energy:

$$E_N = \sum_{n=-N}^{N} |x(n)|^2$$

$$E = \lim_{N \to \infty} E_N$$

$$P = \lim_{N \to \infty} \frac{1}{2N+1} E_N$$

Comments:

(1) If a signal’s energy is finite, $P = 0$.

(2) If a signal’s energy is infinite, its power may or may not be zero.

(3) RMS value is the square root of the power.

Examples:

(1) The average power of a sinewave is $\frac{A^2}{2}$.

(2) What does the following compute?

$$E(n) = E(n-1) + \alpha x^2(n)$$

Transforms

The z-transform of a discrete-time signal is defined as:

$$X(z) \equiv \sum_{n=-\infty}^{\infty} x(n) z^{-n} \qquad x(n) = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz$$

The Fourier transform of $x(n)$ can be computed from the z-transform as:

$$X(\omega) = X(z)\big|_{z = e^{j\omega}} = \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}$$

The Fourier transform may be viewed as the z-transform evaluated around the unit circle.

The Discrete Fourier Transform (DFT) is defined as a sampled version of the Fourier shown above:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}, \qquad k = 0, 1, 2, \ldots, N-1$$

The inverse DFT is given by:

$$x(n) = \sum_{n=0}^{N-1} X(k) e^{j2\pi kn/N}, \qquad n = 0, 1, 2, \ldots, N-1$$

The Fast Fourier Transform (FFT) is simply an efficient computation of the DFT.

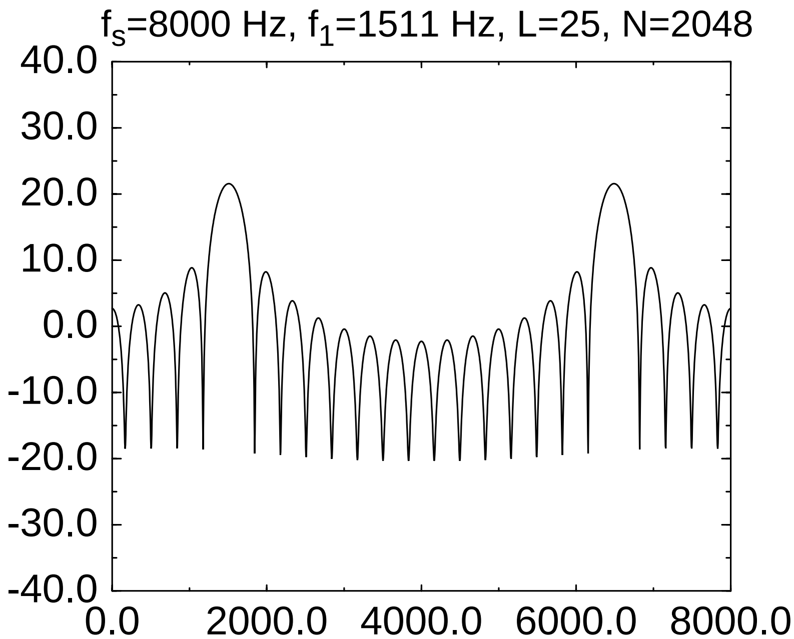
Note that these are not the only transforms used in speech processing (wavelets, Wigner distributions, fractals, etc.).

Time-Domain Windowing

Let $\{x(n)\}$ denote a finite duration segment of a signal:

$$\hat{x}(n) = x(n)w(n)$$

This introduces frequency domain aliasing (the so-called picket fence effect):



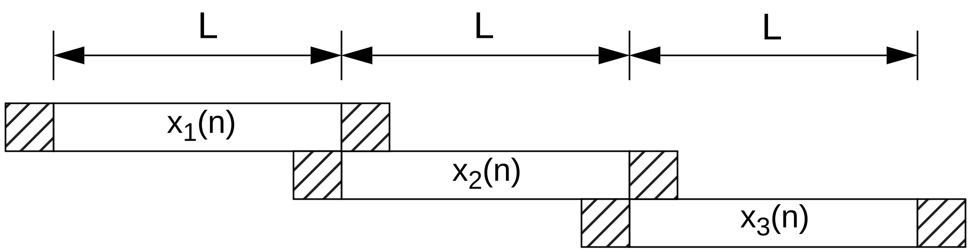
Popular Windows

Generalized Hanning: $$w_H(k) = w(k)\left[\alpha + (1-\alpha)\cos\left(\frac{2\pi}{N}k\right)\right] \qquad 0 < \alpha < 1$$

$$\alpha = 0.54, \qquad \textit{Hamming window}$$
$$\alpha = 0.50, \qquad \textit{Hanning window}$$

Frame-Based Analysis With Overlap

Consider the problem of performing a piecewise linear analysis of a signal:



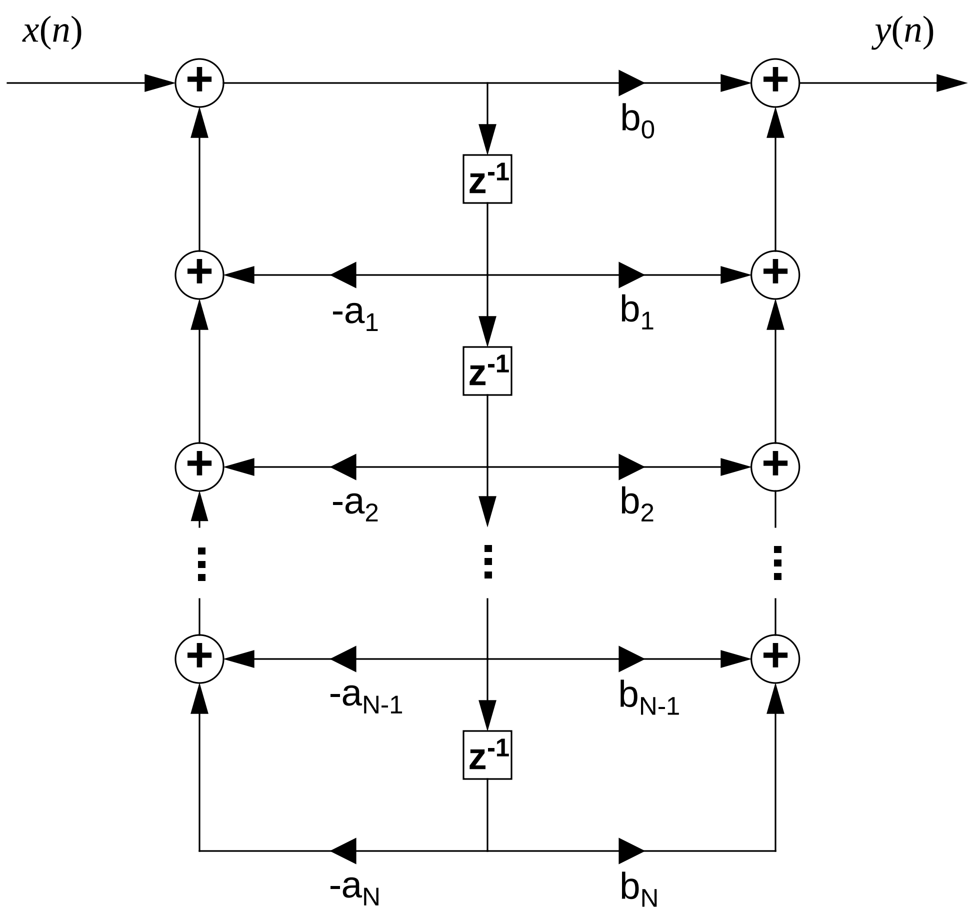
Difference Equations, Filters, and Signal Flow Graphs

A linear time-invariant system can be characterized by a constant-coefficient difference equations:

$$y(n) = -\sum_{k=1}^{N} a_k y(n-k) + \sum_{k=0}^{M} b_k x(n-k)$$

Is this system linear if the coefficients are time-varying?

Such systems can be implemented as signal flow graphs:



Minimum Phase, Maximum Phase, MIxed Phase, and Speech Perception

An FIR filter composed of all zeros that are inside the unit circle is minimum phase. There are many realizations of a system with a given magnitude response; one is a minimum phase realization, one is a maximum-phase realization, others are in-between. Any non-minimum phase pole-zero system can be decomposed into:

$$H(z) = H_{min}(z)H_{ap}(z)$$

It can be shown that of all the possible realizations of $|H(\omega)|$, the minimum-phase version is the most compact in time:

Define:

$$E(n) = \sum_{k=0}^{n} |h(k)|^2$$

Then, $E_{min}(n) \geq E(n)$ for all n and all possible realizations of $|H(\omega)|$.

Why is minimum phase such an important concept in speech processing?

We prefer systems that are invertible:

$$H(z)H^{-1}(z) = 1$$

We would like both systems to be stable. The inverse of a non-minimum phase system is not stable.

We end with a very simple question:

Is phase important in speech processing?