# Lecture 1: Reviews

## Reading: Chapter 1

- **Digital** signals are represented by discrete-time and discrete-valued data.
- **Analog** signals, on the other hand, are represented by continuous-time and continuous-valued data.

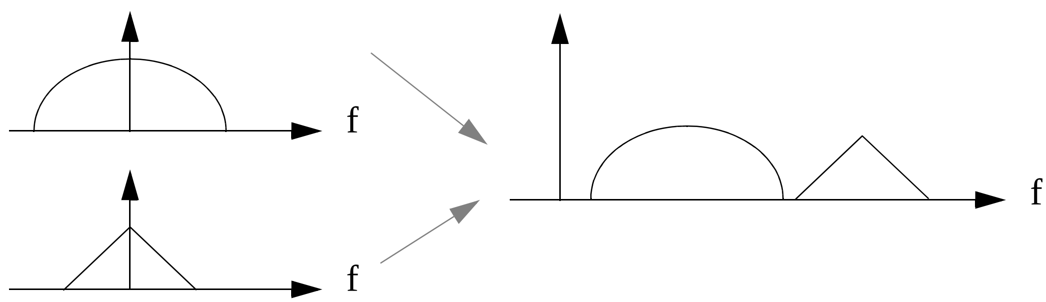
Q: Are there signals that are discrete-time but continuous-valued?

- Why do we emphasize filtering?

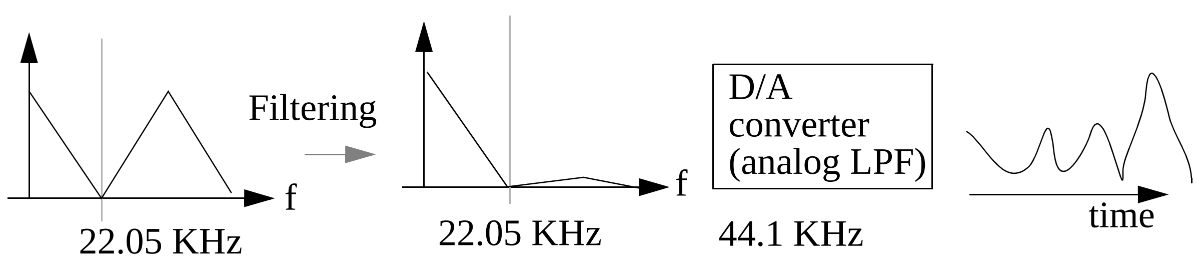
  One of the most useful operations in signal processing is to limit signal bandwidth, which implies filtering. Also, we often need to select certain signal properties, which again uses filtering.

- Applications:

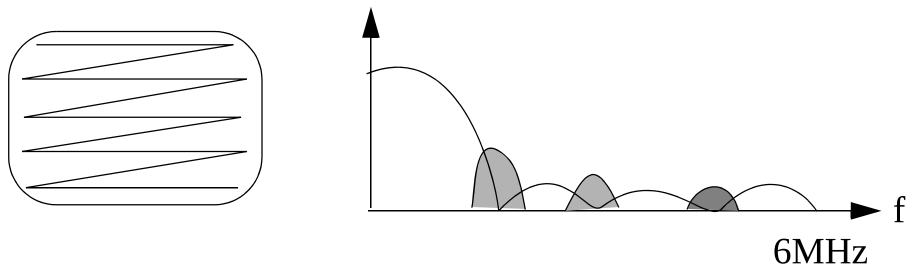
  1. FDM - Sharing of available frequency bandwidth. Examples are FM radio, paging systems, cellular phones, etc.



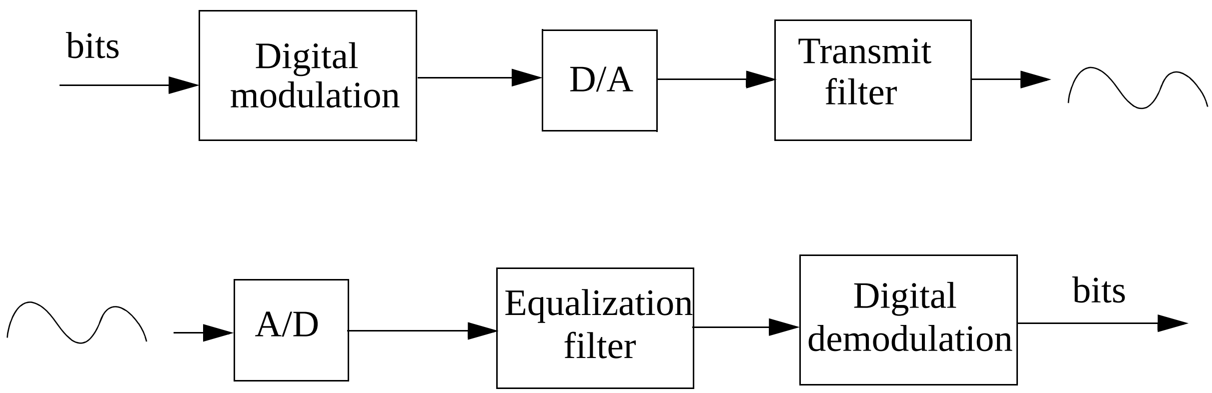

  2. CD Player - Low pass filter for over-sampling interpolation. Attenuate signal components above 22.05 KHz before D/A conversion.

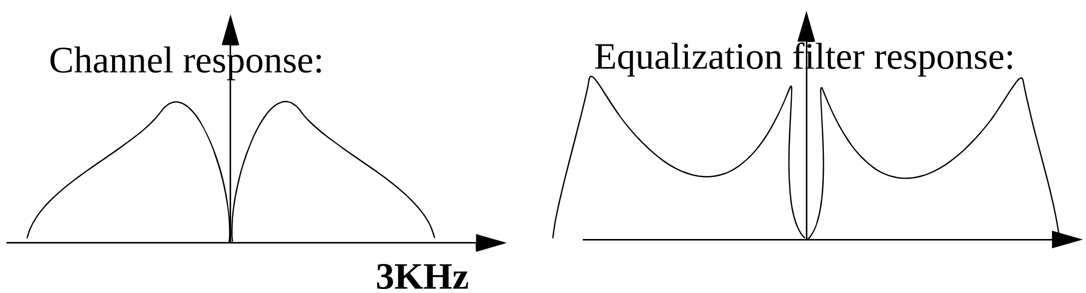


3. Color TV - Introduction of color components into luminance sprectrum without destroying the original black-and-white reception (an interesting engineering task).




4. Voice-band data modem - Transmission of digital data over telephone lines, an emphasis on filtering for two reasons: transmit filtering to limit signal bandwidth and equalization filter to compensate for channel distortion.



# Discrete-Time Signals and Systems

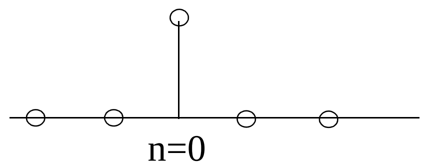
**Reading: Chapter 2.**

- Discrete-time signal is usually indexed by an integer $n$, representing the $n$-th sample in an infinitely long sequence.

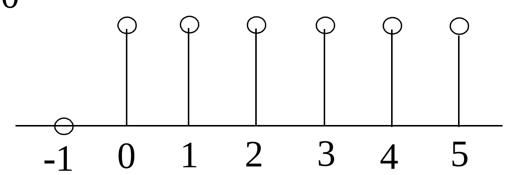
  General form: $x(n)$, $n$ is an integer ($x(n)$ is not defined for non-integer $n$).

  Examples:

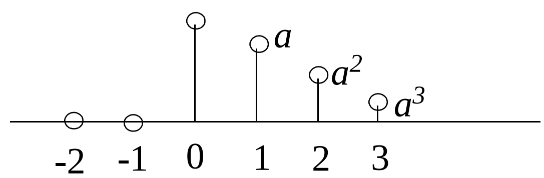
  1. $\delta(n)$ =

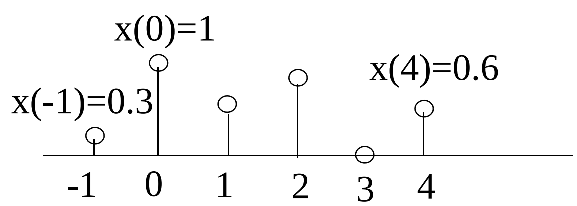
  2. $u(n) = \sum_{k=0}^{\infty} \delta(n-k)$

  3. $x(n) = a^n u(n)$

  4. $x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n-k)$

- **Fourier transform** and **inverse Fourier transform** are by far the most useful transformations in signal analysis. Their discrete-time version is defined as follows:

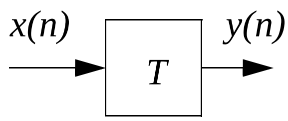
$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}$$

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

  Properties of the Fourier transform include linearity, the time-delay relationship $x(n-n_d) \leftrightarrow X(e^{j\omega})e^{-j\omega n_d}$ , and symmetry.

## Discrete-Time Systems

- Linearity:




If $y_1(n) = T\{x_1(n)\}$ and $y_2(n) = T\{x_2(n)\}$,
then $T\{\alpha x_1(n) + \beta x_2(n)\} = \alpha y_1(n) + \beta y_2(n)$
for all $\alpha$ and $\beta$ belonging to $R$.

Linearity ensures that the system follows the principle of superposition, so that each component of the input signal can be dealed with individually.

- Time-invariance:

If $y(n) = T\{x(n)\}$, then $T\{x(n-n_o)\} = y(n-n_o)$ for all integer $n_o$.

Time-invariance ensures that the same input signal always generates the same output signal, regardless of **when** the input signal is applied to the system.

- Linear time-invariant systems: By linearity, we can express the system output as follows.

$$y(n) = T\left\{ \sum_{k=-\infty}^{\infty} x(k)\delta(n-k) \right\} = \sum_{k=-\infty}^{\infty} x(k)T\{\delta(n-k)\}$$

Let's define $T\{\delta(n-k)\}$ to be the response of the system to $\delta(n-k)$, denoted as $h_k(n)$. By time-invariance, we can write $h_k(n)$ as $h_0(n-k)$, where $h_0(n)$ is the response to $\delta(n)$, defined as the **impulse response** of the system. Then the output equation can be rewritten as:

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h_k(n) = \sum_{k=-\infty}^{\infty} x(k)h_0(n-k)$$

The above equation defines a **convolution.** Note that the behavior of a linear time-invariant system can be solely characterized by its impulse response, while for a system that is non-linear or time-variant, its impulse response describes only part of the system's characteristics.

- From the definition of a **transfer function** (the Z-transform of the system's impulse response), we know that using the transfer function to describe a system makes sense only when the system is linear and time-invariant. When the system is either non-linear or time-variant, its transfer function doesn't carry too much information.
- If a system is linear time-invariant, with its frequency response denoted as $H(e^{j\omega})$, then we know that the frequency response of the system's output is merely $X(e^{j\omega})H(e^{j\omega})$, the product of the frequency response of the input signal and $H(e^{j\omega})$. This notation has a very important implication, that is, a linear time-invariant system is **not** capable of generating any frequency components that do not already exist in the input. Conversely, if a system is capable of generating frequency components not present in the input signal, then the system cannot possibly be linear time-invariant. Examples of such systems are frequency modulator, squarers, etc.

Q: Is the sampling operation a linear time-invariant system?

Q: How should we describe the behavior of a system that is not linear time-invariant?

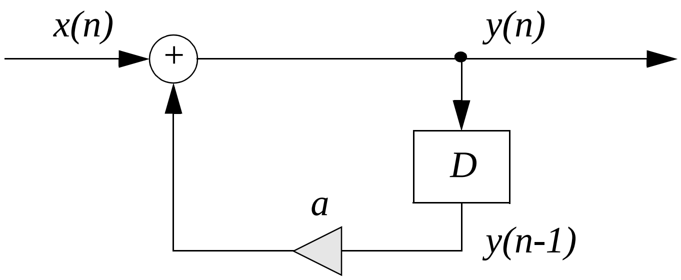
# Linear Constant-Coefficient Different Equations (LCCDE)

- LCCDE describes a system using the following equation:

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

LCCDE describes a subset of all possible linear time-invariant systems, if the system described by LCCDE can be shown to be linear. Take a look at the following example.

*Example:* The following system has an LCCDE of $y(n)-ay(n-1) = x(n)$.




If y(-1)=1, is this system linear? Is the system time-invariant?

Let's try one sample:

$y_1(0) = T\{x_1(0)\} = x_1(0) + a$.

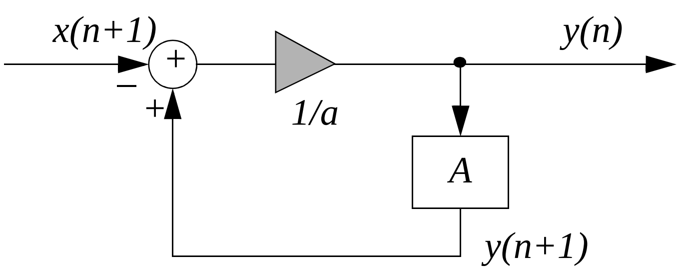
$y_2(0) = T\{x_2(0)\} = x_2(0) + a$.

$T\{x_1(0) + x_2(0)\} = x_1(0) + x_2(0) + a \ ?=? \ y_1(0) + y_2(0)$.

You can check on the property of time-invariance yourself.

Therefore for an LCCDE to be truly linear and time-invariant (LTI), we need the condition called "the initial-rest condition", which makes sure that all initial conditions are set to zero before the input is applied.

We have another problem. The LCCDE in the example above has more than one implementation. Take a look at the following diagram:



Although the system as drawn is not **causal**, nonetheless it has the same LCCDE as given in the example. Therefore to define a unique LTI system as described by a LCCDE, we have to make the following assumption: that the system is causal.

Q: Suppose a system is described as $y(n) = \sum_{m=0}^{M} b_m x(n-m)$ .

Without further assumptions, can you tell if this system is LTI? What is its impulse response?

# Z-Transform

### Reading: Chapter 3

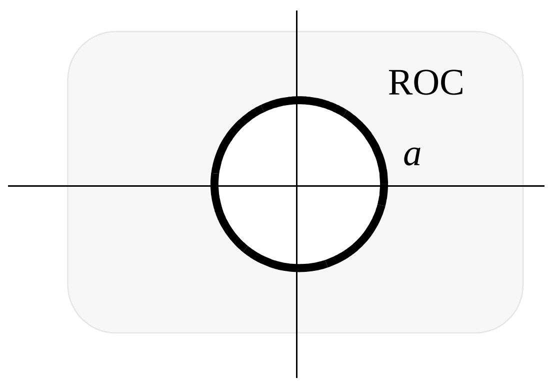
- Z-transform is a transformation between integer index *n* and a complex number *z*. It transforms a time-domain sequence to a function defined on a complex plane.

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} .$$

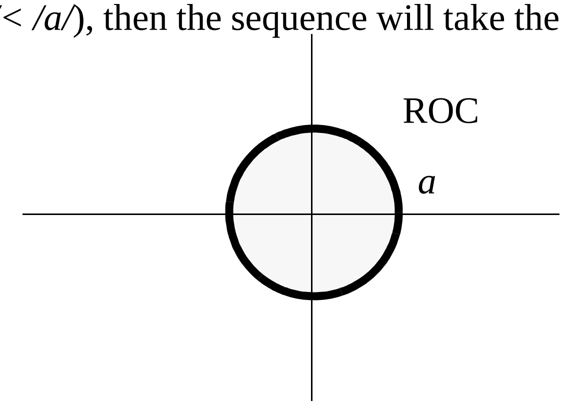
*Example:* Consider the sequence $x(n) = a^n u(n)$. Its Z-transform $X(z)$ is

$$\sum_{n=-\infty}^{\infty} x(n)z^{-n} = \sum_{n=0}^{\infty} a^n z^{-n} = \frac{1}{1-az^{-1}} .$$

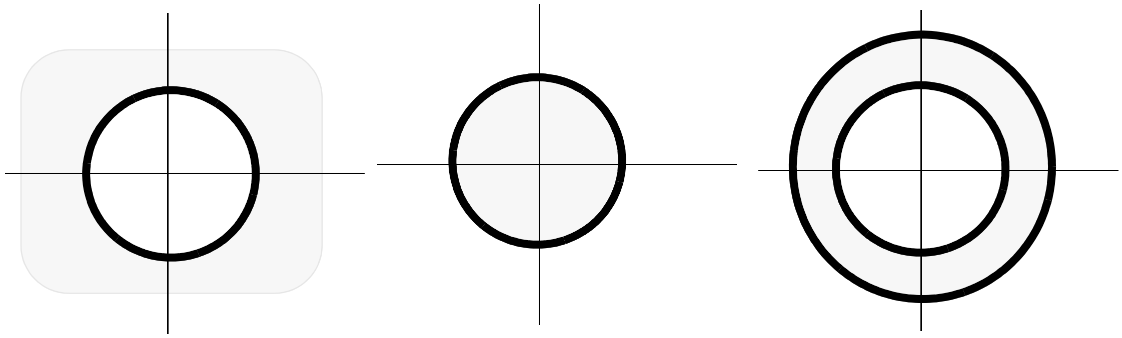
If $|az^{-1}| < 1$, then the infinite summation converges. By convergence we mean that $X(z_o)$ is finite if evaluated at a particular point $z_o$.The region of convergence for the above $X(z)$ is $|z| > |a|$.




If given the above Z-transform and we assume that it is an anti-causal sequence (/z/< /a/), then the sequence will take the form of $x(n) = - a^n u(-n-1)$.



There are only three legal shapes of ROCs (proved by complex variable theorems), as shown below.



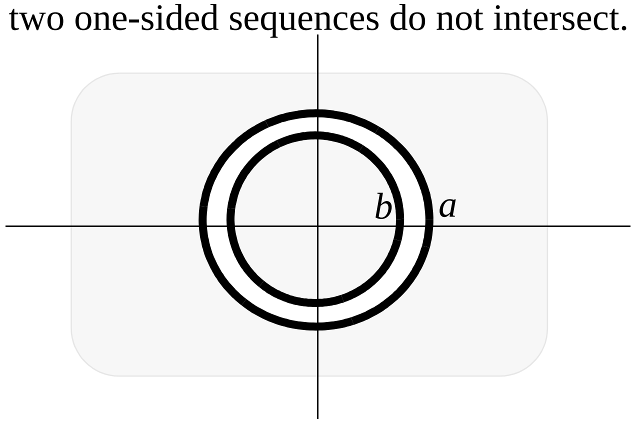
Right-sided sequences Left-sided sequences Two-sided sequences

There are infinitely many sequences whose Z-transforms do not have a ROC. It simply means that at no point on the complex Z-plane is the Z-transform evaluated to be finite (nothing to be alarmed about). Look at the following example.

*Example:* If $x(n) = a^n u(n) + b^n u(-n)$ and $|a| > |b|$, then $X(z)$ has no ROC.

$$X(z) = \sum_{n=0}^{\infty} a^n z^{-n} + \sum_{n=-\infty}^{0} b^n z^{-n}.$$

The above summation is infinite at all points on the Z-plane, because the two ROCs of the two one-sided sequences do not intersect.




Q: What is the ROC of a finite-sequence?

- FACT: A rational-polynomial $X(z)$ and its ROC uniquely defines a time-domain sequence. Without ROC, depending on the number of poles in the rational polynomial (poles determine the boundaries of ROCs), there may be many time-domain sequences corresponding to the same $X(z)$.

## Solution of LCCDE using Z-Transform

- Given a LCCDE with the initial-rest condition, we usually like to know its impulse response. Since LCCDE is defined as a recursive equation, using time-domain analysis may be complicated. One easy solution is to work in the transform domain. Given a LCCDE,

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

If we apply Z-transform to both sides of the equation, by linearity and the time-delay relationship, we get

$$\sum_{k=0}^{N} a_k Y(z) z^{-k} = \sum_{k=0}^{M} b_k X(z) z^{-k} \text{, which gives}$$

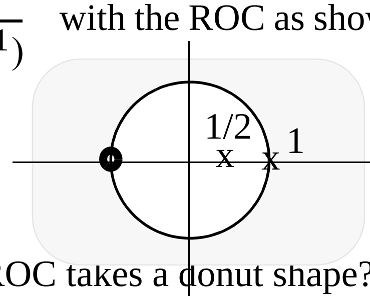
$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}} \ .$$

The above equation gives the Z-transform of the system's impulse response in a rational-polynomial form. To solve for the time-domain impulse response, we can apply inverse Z-transform, which usually is quite a pain to do. Since *H(z)* takes on this special form of a rational polynomial, we have a short-cut solution: the high-school math of partial fraction expansion!

*Example:* Suppose $H(z) = \frac{1 + 2z^{-1} + z^{-2}}{(1 - 0.5z^{-1})(1 - z^{-1})}$ with the ROC as shown.

$$H(z) = 2 - \frac{9}{1 - 0.5z^{-1}} + \frac{8}{1 - z^{-1}}$$

$$h(n) = 2\delta(n) - 9(0.5)^n u(n) + 8u(n) \quad .$$




Q: What is the corresponding time-sequence if the ROC takes a donut shape?

Q: What is the time-sequence if the ROC takes a cookie shape?

# Transform Analysis of LTI Systems

**Reading: Sec. 5.0 - 5.3.**

- **Frequency response** is the magnitude and phase response of a LTI system, obtained by evaluating the Z-transform on the unit circle on the complex Z-plane.

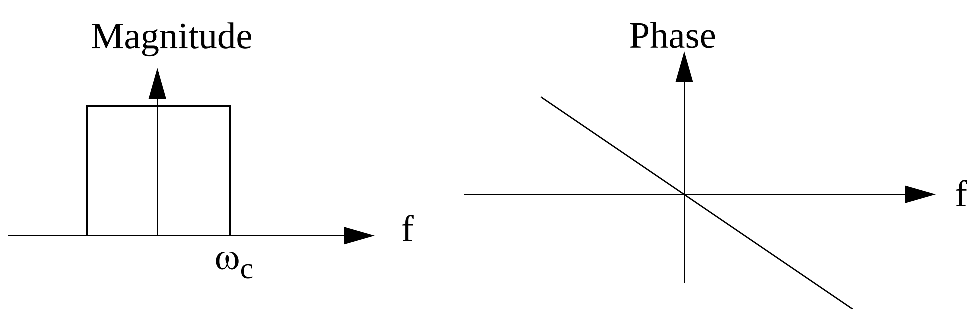
$$H(z)\big|_{z = e^{j\omega}} = \left|H(e^{j\omega})\right| \angle H(e^{j\omega}) .$$

The calculation of frequency response involves calculating the absolute value of $H(e^{j\omega})$ and its angle with respect to the real axis as a function of $\omega$. Since we are most interested in Z-transforms expressed in a rational-polynomial form (our digital filters), we have an easier way to calculate the magnitude and the phase using the geometric interpretation of poles and zeros, although we will most likely use Matlab to plot the response for us.

*Example:* An ideal low-pass filter has a frequency response of

$$H(e^{j\omega}) = e^{-j\omega n_d}, |\omega| < \omega_c \text{ and 0 otherwise.}$$

Its frequency response is




Its impulse response is $h(n) = \dfrac{\sin\omega_c(n-n_d)}{\pi(n-n_d)}$ .

- Phase distortion is usually measured by a quantity called group delay, which is defined as the negative of the derivative of the phase response with respect to frequency:

$$\tau(\omega) = -\frac{d}{d\omega}([\angle H(e^{j\omega})]) .$$

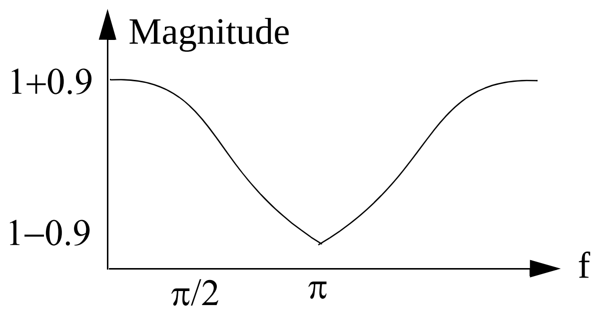
The deviation of the group delay away from a constant indicates the degree of nonlinearity of the phase, thus giving a measure of phase distortion.

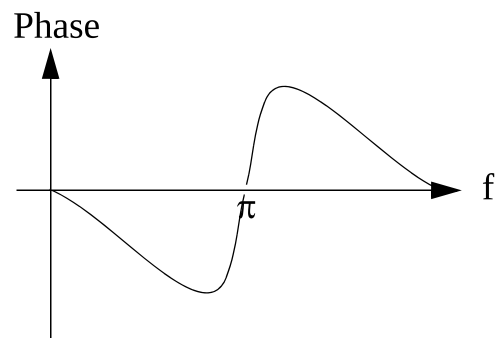
*Example:* Draw the frequency response of $H(z) = 1 - \gamma e^{j\theta} z^{-1}$ .

$$|H(e^{j\omega})| = \sqrt{1 + \gamma^2 - 2\gamma\cos(\omega - \theta)}$$

$$\angle H(e^{j\omega}) = \angle(1 - \gamma e^{j\theta} e^{-j\omega}) = \tan^{-1}\left[\frac{\gamma\sin(\omega - \theta)}{1 - \gamma\cos(\omega - \theta)}\right] .$$

Let's choose $\gamma$ to be 0.9 and $\theta$ to be $\pi$. Choosing a few points on the frequency axis from 0 to $\pi$ (magnitude response is even-symmetric around $\pi$), we can make the following sketch:




- Stability and causality: It is very easy to determine if a LTI system is stable or causal from the Z-transform of its impulse response. If the ROC contains the unit circle, then the system is stable (its frequency response exists). If the ROC includes infinity, then the system is causal (obvious from the fact that a right-sided impulse response has a ROC outside a disc, including infinity). Note that stability and causality are two totally unrelated properties of a system. A system can be causal and unstable, or stable but noncausal.

*Example:* The transfer function of a communication channel is described by $H(z) = \frac{-0.5 + z^{-1}}{1 - 0.9z^{-1}}$ , with a ROC of $|z|>0.9$. We want to design an inverse filter to compensate for the distortion introduced by the channel.

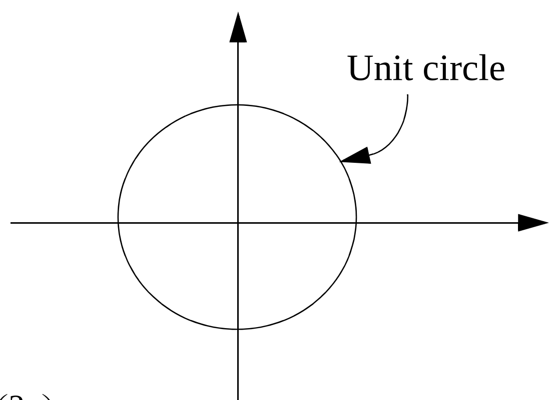
The inverse filter is $H(z)^{-1} = \frac{1 - 0.9z^{-1}}{-0.5 + z^{-1}} = \frac{-2 + 1.8z^{-1}}{1 - 2z^{-1}}$ .

Given the above transfer function, we have two choices in designing the inverse filter. One is to use a stable, but non-causal impulse response $h(n)=2^{n+1}u(-n-1)-1.8(2)^{n-1}u(-n)$. The other is to use a causal, but unstable impulse response $h(n)=-2^{n+1}u(n)+1.8(2)^{n-1}u(n-1)$. Well, what should we do?

## Geometric Interpretation of Poles and Zeros

- Poles in a Z-transform define the boundaries of ROCs, because by definition the Z-transform evaluated at a pole gives an infinite value. Poles at the origin, however, have no effects on the magnitude response and merely represent a delay in the time-domain sequence.
- Zeros on the unit circle are used to quickly bring the magnitude response to zero, for the design of band-selective filters. Zeros at the origin have no effects on the magnitude response and merely represent an advance in the time-domain sequence.
- IMPORTANT!!! When you try to draw a pole-zero plot, make sure that you convert the Z-transform from a function of $z^{-1}$ to a function of $z$, and do partial fraction expansion on $z$, not on $z^{-1}$.

  *Example:* Given a system with a transfer function $H(z) = 1-z^{-M}$, draw its pole-zero plot.




  Zeros at $z = e^{j\left(\frac{2\pi}{M}\right)i}$, $i = 0, 1, 2, ... , M-1$.

Q: Are there poles in the system?

Q: What is the impulse response of the system?

Q: Do causal FIR filters always have poles?

Q: What's the system output if the input is $x(n) = e^{-j\omega_o n}$, $\omega_o = 2\pi/M$? Can you show it by calculating $Y(z)=X(z)H(z)$?

Q: What can we tell of $|H(e^{j\omega})|$?

# Network Structures for Discrete-Time Systems

**Reading: Sec. 6.0 - 6.5.**

- For any LCCDE with the initial-rest condition and causality assumption, there still exist many ways of implementing it. Although the transfer function of these implementations are all the same, properties such as finite-precision effects depend on the exact implementation structure.
- The first structure is the canonical form structure, which minimizes the number of delay elements in implementing a LCCDE. This structure has very low tolerance to quantization error and therefore is not widely used in practice.
- The second structure is called the cascade form, which groups poles and zeros into complex-conjugate pairs. Each pair is called a biquad, or second-order section, which is the basic building block for cascaded filters. Since poles and zeros are grouped by complex-conjugate pairs, each section uses only real coefficients (given real coefficients in its LCCDE).

$$\frac{(z-z_1)(z-z_1^*)}{(z-p_1)(z-p_1^*)} = \frac{1-2Re[z_1]z^{-1}+|z_1|^2z^{-2}}{1-2Re[p_1]z^{-1}+|p_1|^2z^{-2}}$$

Cascade form is represented by $H(z) = b_0\frac{\prod(z-z_m)(z-z_m^*)}{\prod(z-p_k)(z-p_k^*)}$, where each pole and zero is combined with its complex conjugate.

- Parallel form is obtained by partial fraction expansion of the transfer function

  into second-order sections, which gives a summation of biquads represented by $H(z) = \sum\frac{e_{0k}+e_{1k}z^{-1}}{1-a_{1k}z^{-1}-a_{2k}z^{-2}}$ .
- There are other structures such as the lattice structure or normalized structure, which usually use more delay elements and computation hardware but guarantee stability at the present of quantization error.