Digital Signal Processing:

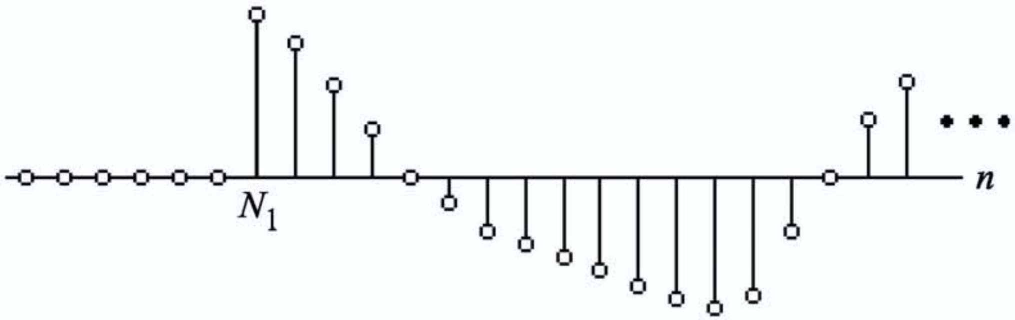
Mathematical and algorithmic manipulation of **discretized and quantized** or **naturally digital** signals in order to extract the most relevant and pertinent information that is carried by the signal.

- What is a signal?
- What is a system?
- What is processing?

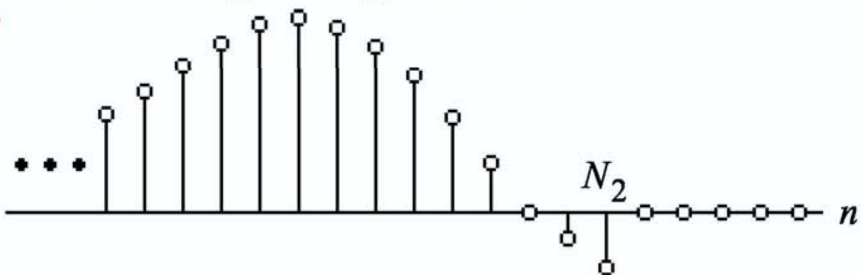
Discrete Signals

- A length-N sequence is often referred to as an **N-point sequence**
- The length of a finite-length sequence can be increased by ***zero-padding***, i.e., by **appending** it **with zeros**

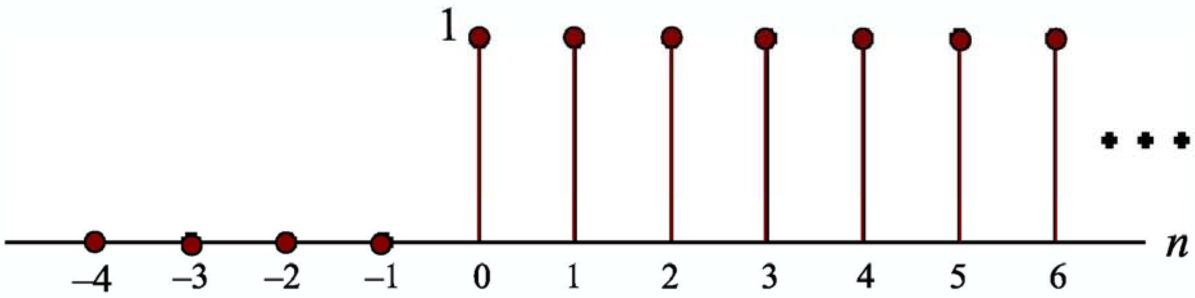
- A ***right-sided sequence*** $x[n]$ has zero-valued samples for $n<N_1$. If $N_1>0$, a right-sided sequence is called a ***causal sequence***



- A ***left sequence*** $x[n]$ has zero-valued samples for $n>N_2$. If $N_2<0$, a left-sided sequence is called an ***anti-causal sequence***



Unit Step (sequence)



$$u[n] = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases} = \sum_{i=0}^{\infty} \delta[n-i]$$

$$\delta[n] = u[n] - u[n-1]$$

PERIODICITY

- Sinusoidal sequence $A\cos(\omega_o n + \phi)$ and complex exponential sequence $Be^{j\omega_o n}$ are ***periodic sequences*** of period N, if $\omega_0 N = 2\pi r$, where N and r are positive integers
- Smallest value of N satisfying $\boxed{\omega_o N = 2\pi r}$ is the **fundamental period** of the sequence
- Any sequence that does not satisfy this condition is ***aperiodic***
- To verify the above fact, consider

$$x_1[n] = \cos(\omega_o n + \phi) \qquad x_2[n] = \cos(\omega_o (n+N) + \phi)$$

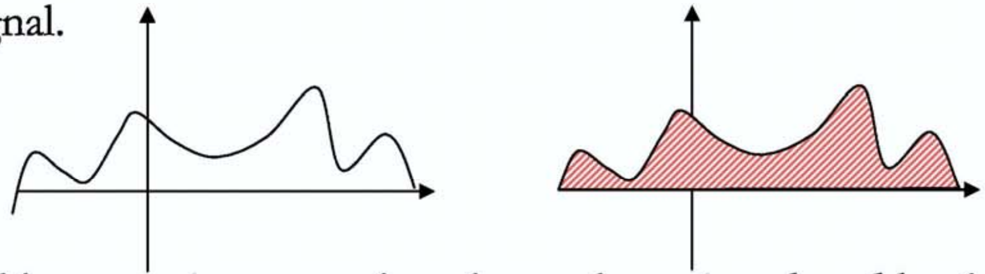
$$x_2[n] = \cos(\omega_o n + \phi)\cos\omega_o N - \sin(\omega_o n + \phi)\sin\omega_o N$$

$$= \cos(\omega_o n + \phi) = x_1[n] \quad \textbf{iff} \quad \sin\omega_o N = 0 \ \text{ and } \ \cos\omega_o N = 1$$

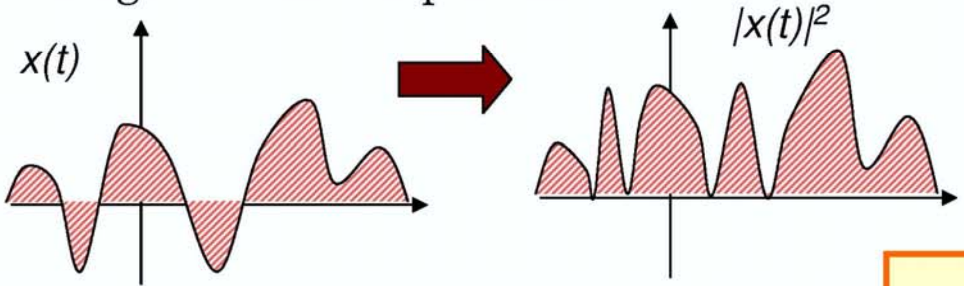
➪ These two conditions are met if and only if $\omega_o N = 2\pi r$ or $\dfrac{2\pi}{\omega_o} = \dfrac{N}{r}$

Energy & Power in Signals

- It is often useful to define the "size" or "strength" of a signal. That is, we would like to be able to use a single number that represents the average strength of the signal. How would we do that?
  - A reasonable answer would be to use the area under the curve. The larger the area, the stronger the signal.



  - But if the signal has negative areas, than the total area is reduced by the negative parts. Yet, a negative signal is not necessarily a weaker signal. In fact, -110 V will jolt you as much as +110 V will. So, we need another approach. Calculating the area under the "square of the absolute value of the signal" solves this problem



  - This area is defined as the energy of the (continuous time) signal.

$$E_x = \sum_{n=-\infty}^{\infty} \left(x[n] \right)^2$$

Causality

- A system is said to be ***causal***, if the output at time n_0 does not depend on the inputs that come after n_0.

- In other words, in a causal system, the n_0^{th} output sample $y[n_0]$ depends only on input samples $x[n]$ for $n \le n_0$ and does not depend on input samples for $n>n_0$.
- Here are some examples: Which systems are causal?

$$y[n] = \alpha_1 x[n] + \alpha_2 x[n-1] + \alpha_3 x[n-2] + \alpha_4 x[n-3]$$

$$y[n] = b_0 x[n] + b_1 x[n-1] + b_2 x[n-2] + a_1 y[n-1] + a_2 y[n-2]$$

$$y[n] = y[n-1] + x[n]$$

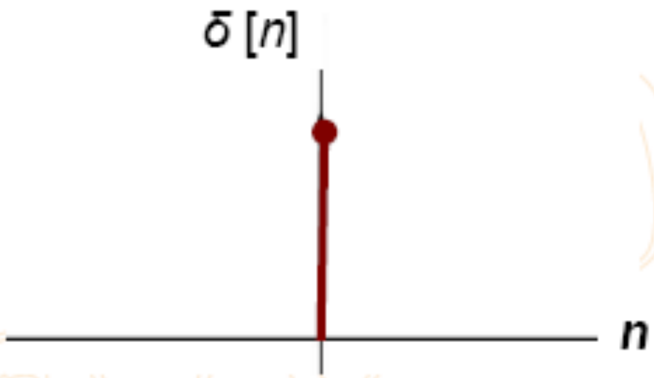
$$y[n] = x_u[n] + \frac{1}{2}\left(x_u[n-1] + x_u[n+1]\right)$$

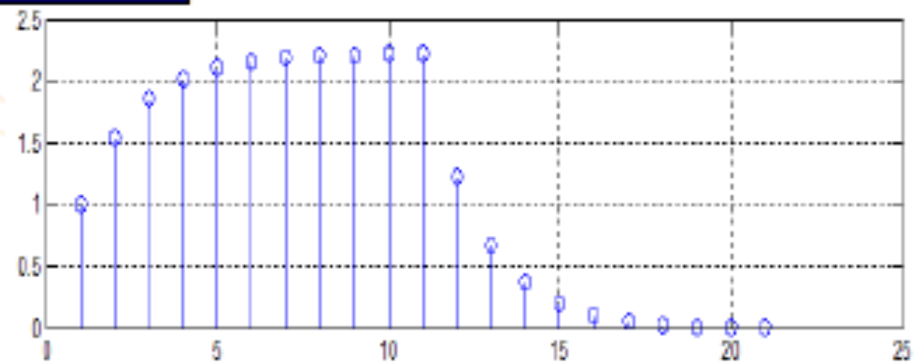
$$y[n] = x_u[n] + \frac{1}{3}\left(x_u[n-1] + x_u[n+2]\right) + \frac{2}{3}\left(x_u[n-2] + x_u[n+1]\right)$$

IMPULSE RESPONSE

➲ The response of a discrete system to a unit impulse sequence $\boldsymbol{\delta[n]}$ is called the ***impulse response*** of the system, and it is typically denoted by $\boldsymbol{h[n]}$

$$x[n]=\delta[n] \longrightarrow \boxed{\text{Discrete System}} \longrightarrow y[n]=h[n]$$





$$h[n]=h[n]*\delta[n]=\sum_{m=-\infty}^{\infty} h[m]\,.\,\delta[n-m]$$

FINITE IMPULSE RESPONSE SYSTEMS

- If the impulse response h[n] of a system is of finite length, that system is referred to as a ***finite impulse response (FIR)*** system

$$h[n]=0 \text{ for } n<N_1 \text{ and } n>N_2,\ \ N_1<N_2$$

  - The output of such a system can then be computed as a finite convolution sum

$$y[n]=\sum_{k=N_1}^{N_2} h[k]x[n-k]$$

  - E.g., h[n]=[1 2 0 -1] is a FIR system (filter)

  - FIR systems are also called ***nonrecursive systems*** (for reasons that will later become obvious), where the output can be computed from the current and past input values only – without requiring the values of ***previous outputs***.

INFINITE IMPULSE RESPONSE SYSTEMS

- If the impulse response is of infinite length, then the system is referred to as an ***infinite impulse response (IIR)*** system. These systems cannot be characterized by the convolution sum due to infinite sum.
  - Instead, they are typically characterized by constant coefficient linear difference equations (CCLDEs), as we will see later.
  - Recall accumulator and note that it can have an alternate – and more compact representation that makes the current output a function of previous inputs and outputs

$$y[n] = \sum_{\ell=-\infty}^{n} x[\ell] \quad \Longrightarrow \quad y[n] = y[n-1] + x[n]$$

  - The impulse response of this system (which is of infinite length), cannot be represented with a finite convolution sum. Note that, since the current output depends on the previous outputs, this is also called a ***recursive system***

Constant Coefficient Linear Difference Equations

- All discrete systems can also be represented using ***constant coefficient, linear difference equations,*** of the form

$$\boxed{y[n]+a_1 y[n-1]+a_2 y[n-2]+\cdots a_N y[n-N]=b_0 x[n]+b_1 x[n-1]+\cdots+b_M x[n-M]}$$

Outputs y[n] Inputs x[n]

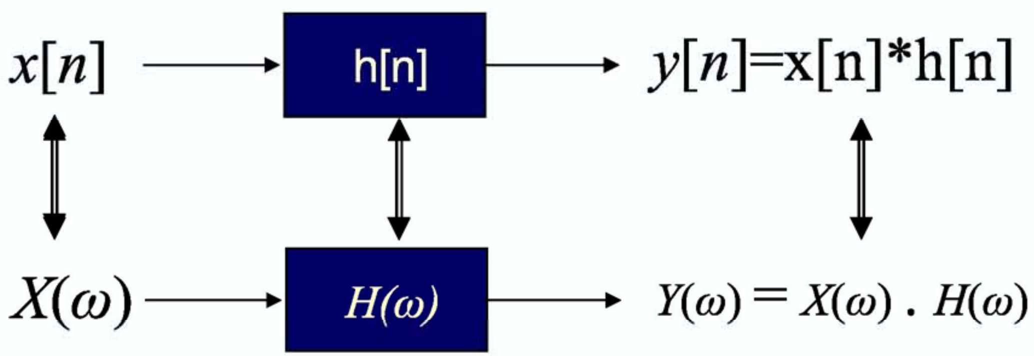
$$\boxed{\sum_{i=0}^{N} a_i y[n-k]=\sum_{i=0}^{M} b_i x[n-k]}$$

Constant coefficients

- Constant coefficients a_i and b_i are called ***filter coefficients***
- Integers M and N represent the maximum ***delay*** in the input and output, respectively. The larger of the two numbers is known as the ***order of the filter***.
- <u>Any</u> LTI system can be represented as two finite sum of products!

Important Theorems

- There are several important theorems related to DTFT.
- Theorem 1:
  - If x[n] is input to an LTI system with an impulse response of h[n], then the DTFT of the output is the product of $X(\omega)$ and $H(\omega)$



FREQUENCY RESPONSE

- Theorem 2:
  - If the input to an LTI system with an impulse response of h[n] is a complex exponential $e^{j\omega_0 n}$, then the output is the ***SAME*** complex exponential whose magnitude and phase are given by $|H(\omega)|$ and $\angle H(\omega)$, evaluated at $\omega = \omega_0$.

$$e^{j\omega_0 n} \longrightarrow \boxed{h[n]} \longrightarrow$$

$$y[n] = \sum_{k=-\infty}^{\infty} h[k] e^{j\omega_0 (n-k)}$$

$$= \underbrace{\left(\sum_{k=-\infty}^{\infty} h[k] e^{j\omega_0 k} \right)}_{\mathbf{H(\omega_0)}} e^{j\omega_0 n}$$

$$\boxed{y[n] = H(\omega_0) e^{j\omega_0 n}}$$

If the system input is a complex exponential at a specific frequency ω_0, then the system output is the same exponential, at the same frequency ω_0 but weighted by a complex amplitude that is a function of the input frequency. This complex amplitude, $H(\omega_0)$, is the DTFT of system impulse function h[n], evaluated at ω_0, and it is called the *frequency response* of the system.

Periodicity of DTFT

- Theorem 3:
  - The DTFT of a discrete sequence is periodic with the period 2π, that is

$$\boxed{X(\omega) = X(\omega + 2\pi k) \quad \text{for any integer } k}$$

- The periodicity of DTFT can be easily verified from the definition:

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

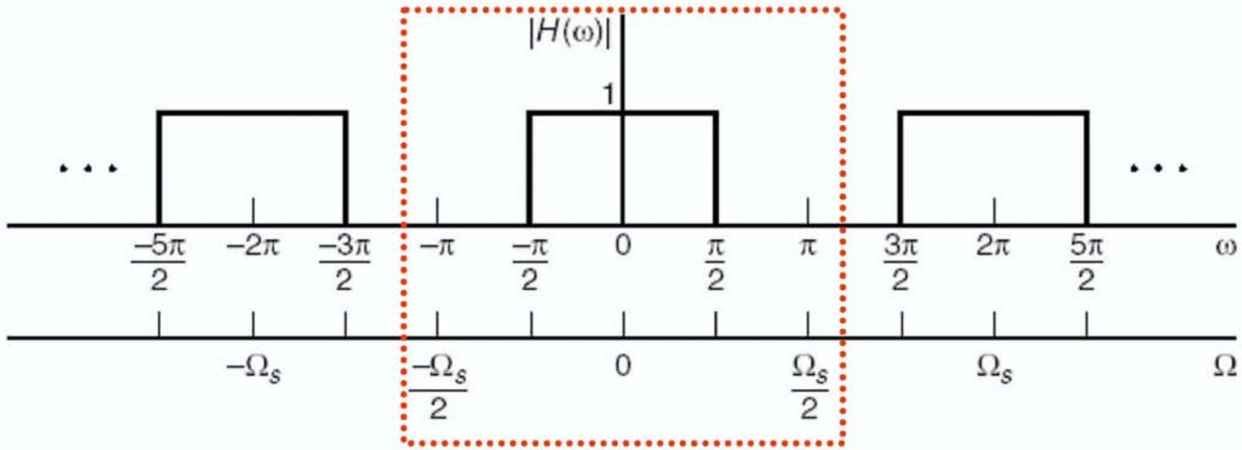
Why…

$$X(\omega + 2\pi k) = \sum_{n=-\infty}^{\infty} x[n] e^{-j(\omega + 2\pi k)n} = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} e^{-j2\pi kn}$$

$$= \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} = X(\omega) \quad \forall k$$

Implications of the Periodicity Property

$$\boxed{H(\omega) = H(\omega + 2\pi)}$$



↬ ***Theorem 4 (You-will-flunk-if-you-do-not-understand-this-fact** **theorem):***
The discrete frequency 2π rad. corresponds to the sampling frequency Ω_s used to sample the original continuous signal x(t) to obtain x[n].

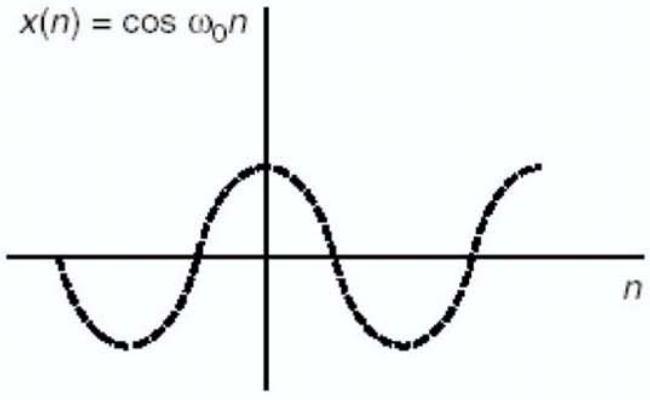
↬Proof: $x(t) = A\sin(\Omega t - \theta) \Rightarrow x(nT_s) = A\sin(\Omega T_s n - \theta)$

↬$\omega = \Omega T_s$ ➔For $\Omega = \Omega_s$, we have $\omega = \Omega_s T_s = 2\pi f_s T_s = 2\pi$

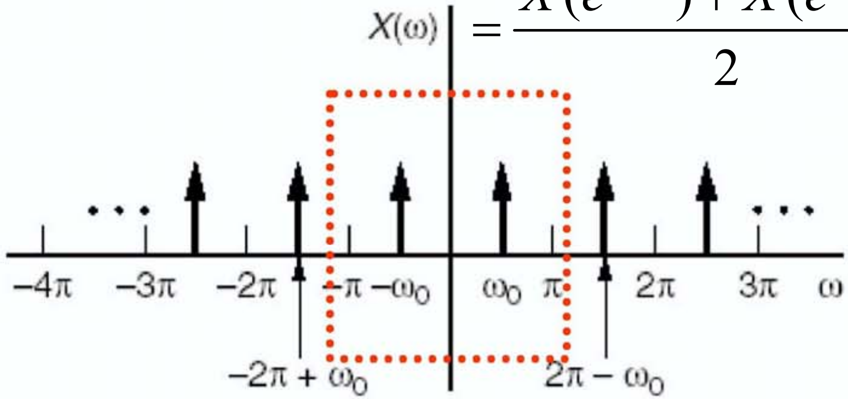
Important DTFT Pairs
The sinusoid at ω_0

➲ By far the most often used DTFT pair (it is less complicated then it looks):

$$x[n] = e^{j\omega_0 n} \quad ; \quad X(\omega) = \sum_{k=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi k)$$



(a)



$$X(\omega) = \frac{X(e^{j\omega_0 n}) + X(e^{-j\omega_0 n})}{2}$$

(b)

$$x[n] = e^{j\omega_0 n} \overset{\mathfrak{I}}{\Leftrightarrow} 2\pi \sum_{m=-\infty}^{\infty} \delta(\omega - \omega_0 \pm 2\pi m)$$

The above expression can also be obtained from the DTFT of the complex exponential through the Euler's formula.

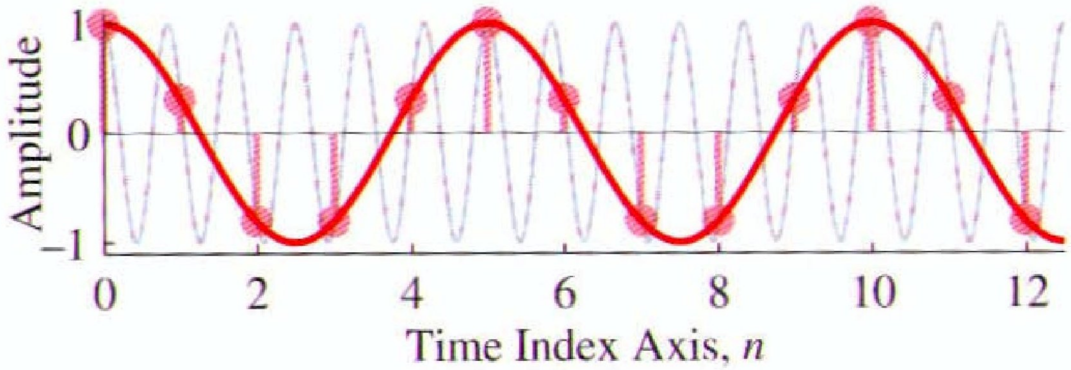
Other Important Properties of DTFT

- We will study the following properties of the DTFT:
  - Linearity ➔ DTFT is a linear operator
  - Time reversal ➔ $x[-n] \leftrightarrow X(-\omega)$
  - Time shift ➔ $x[n-n_0] \leftrightarrow X(\omega)e^{-j\omega n_0}$
  - Frequency shift ➔ $x[n]\, e^{j\omega_0 n} \leftrightarrow X(\omega- \omega_0)$
  - Convolution in time ➔ $x[n]*y[n] \leftrightarrow X(\omega).Y(\omega)$
  - Convolution in frequency
  - Differentiation in frequency ➔ $nx[n] \leftrightarrow j\,(dX(\omega)/d\omega)$
  - Parseval's theorem ➔ Conservation of energy in time and frequency domains
  - Symmetry properties

$$\boxed{x[n] \overset{\mathfrak{I}}{\Leftrightarrow} X(\omega)} \qquad \boxed{y[n] \overset{\mathfrak{I}}{\Leftrightarrow} Y(\omega)}$$

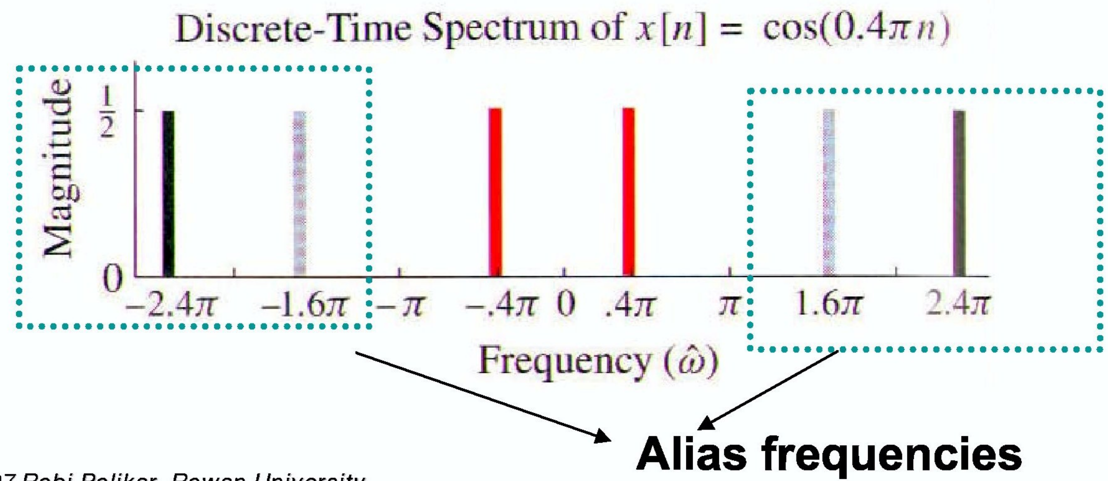
ALIASING

- Here is another example:
  - The signals in the following plot shows two sinusoids: $x_1[n]=\cos(0.4\pi n)$ and $x_2[n]=\cos(2.4\pi n)$. Note that these two signals are distinct, as the second one clearly has a higher frequency.
  - However, when sampled at, say integer values of n , they have the same values, that is $x_1[n]=x_2[n]$ for all integer $\boldsymbol{n}$. These two signals are aliases of each other. More specifically, in the DSP jargon, we say that the frequencies $\omega_1= 0.4\pi$ and $\omega_2=2.4\pi$ are aliases of each other.
  - This is why all signals and systems – when represented in frequency domain – are normalized to a 2π interval



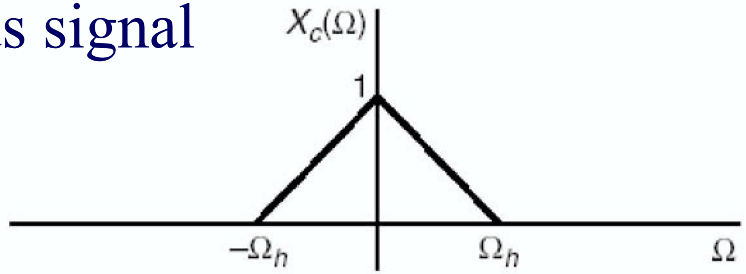
ALIASING

- In general, 2π multiples added or subtracted to a sinusoid gives aliases of the same signal.
  - The one at the lowest frequency is called the ***principal alias***, whereas those at the negative frequencies are called ***folded aliases***.
  - In summary, the frequencies at $\omega_0+2\pi k$ and $2\pi k-\omega_0$ for any integer *k*, are aliases of each other.
  - We can further show that for folded aliases, the algebraic sign of the phase angle is opposite that of the principal alias

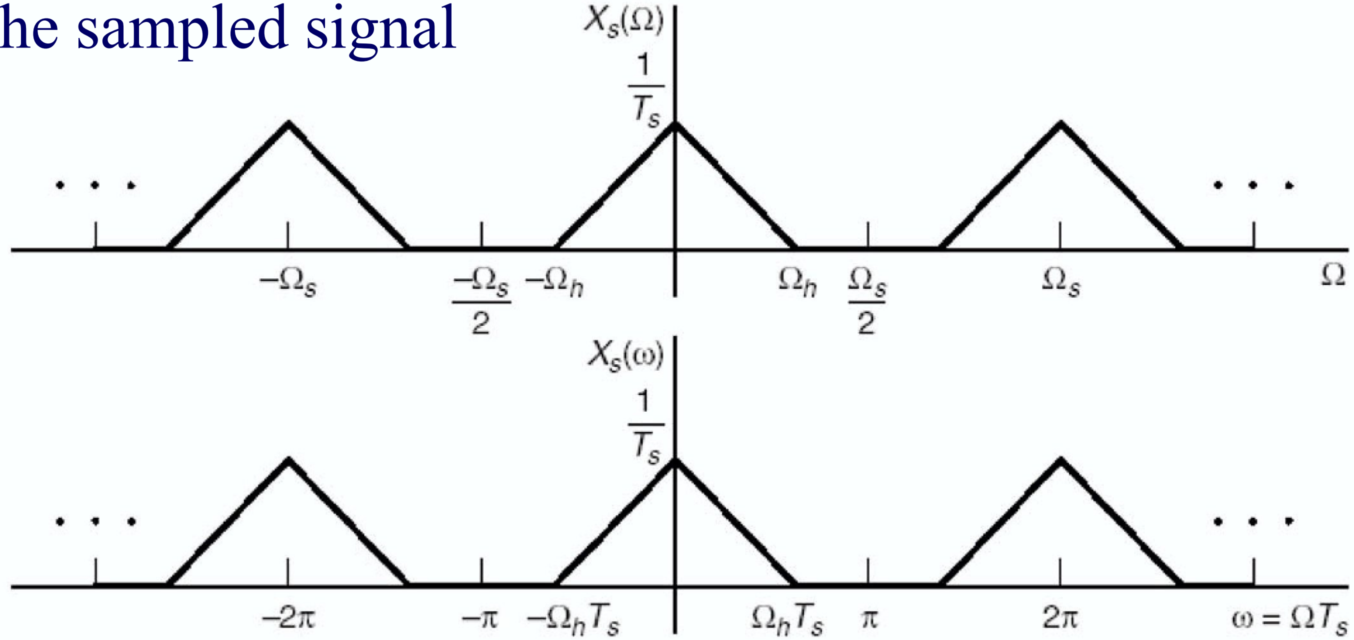


Effect of Sampling in the Frequency Domain

spectrum of the continous signal

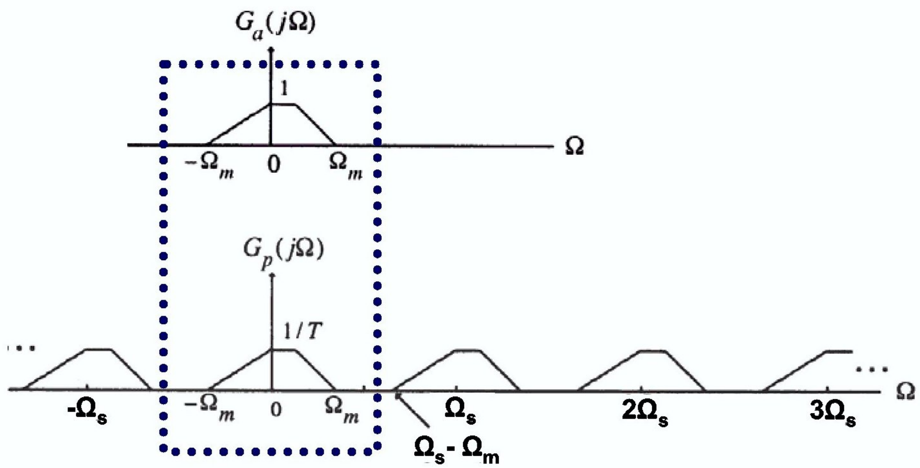


spectrum of the sampled signal



NYQUIST RATE

- Note that the key requirement for the $G_a(\Omega)$ recovered from $G_p(\Omega)$ is that $G_p(\Omega)$ should consist of non-overlapping replicas of $G_a(\Omega)$.



- Under what conditions would this be satisfied…?

?

If $\Omega_s \geq 2\Omega_m$, $g_a(t)$ can be recovered exactly from $g_p(t)$ by passing it through an ideal lowpass filter $H_r(\Omega)$ with a gain T_s and a cutoff frequency Ω_c greater than Ω_m and less than $\Omega_s - \Omega_m$. For simplicity, a half-band ideal filter is typically used in exercises.

DFT Analysis

- **Definition** - The simplest relation between a length-N sequence $x[n]$, defined for $0 \le n \le N-1$, and its DTFT $X(\omega)$ is obtained by uniformly sampling $X(\omega)$ on the ω-axis $0 \le \omega \le 2\pi$ at $\omega_k = 2\pi k/N$, $0 \le k \le N-1$

- From the definition of the DTFT we thus have

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-2\pi k n i/N} = \sum_{n=0}^{N-1} x[n] e^{-\omega_k n i}$$

DFT analysis equation

DFT Analysis

- Note the following:
  - ***k*** replaces ***ω*** as the frequency variable in the discrete frequency domain
  - ***X*[*k*]** is also (usually) a length-*N* sequence in the frequency domain, just like the signal x[n] is a length-*N* sequence in the time domain.
  - X[k] can be made to be longer than *N* points (as we will later see)
  - The sequence ***X*[*k*]** is called the ***discrete Fourier transform (DFT)*** of the sequence ***x*[*n*]**
- Using the notation $W_N = e^{-j2\pi/N}$ the DFT is usually expressed as:

$$X[k] = \sum_{n=0}^{N-1} x[n]\, W_N^{kn} \quad 0 \leqslant k \leqslant N-1$$

Inverse DFT (DFT Synthesis)

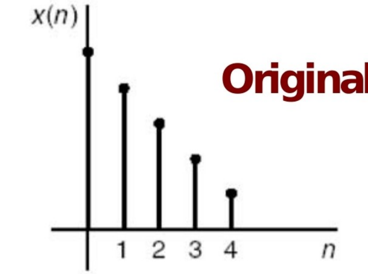
- The inverse discrete Fourier transform, also known as the synthesis equation, is given as

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]\, e^{j\frac{2\pi}{N}nk}$$
$$= \frac{1}{N}\sum_{k=0}^{N-1} X[k] W_N^{-kn}, \quad 0 \le n \le N-1$$

DFT synthesis equation

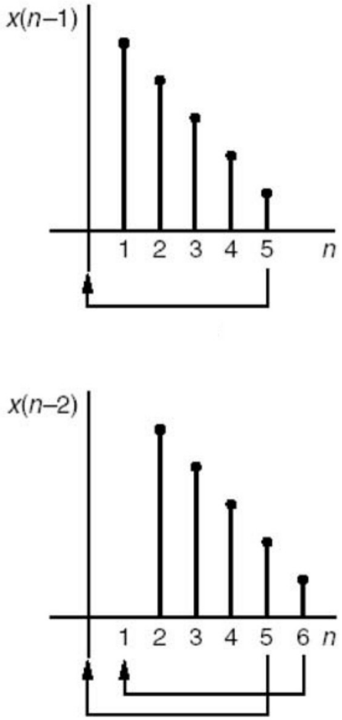
- To verify the above expression we multiply both sides of the above equation by $W_N^{\ell n} = e^{j\frac{2\pi}{N}n\ell}$ and sum the result from $n = 0$ to $n = N-1$

CIRCULAR SHIFT

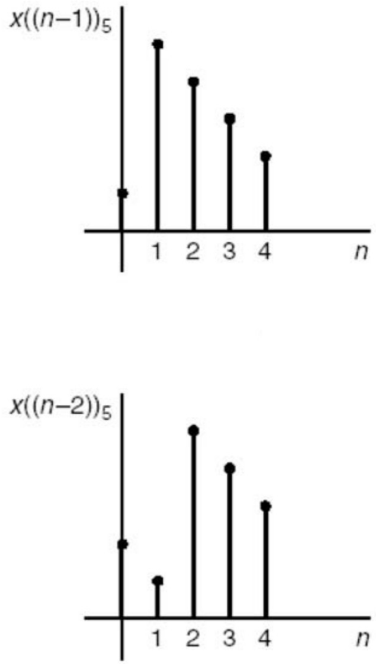


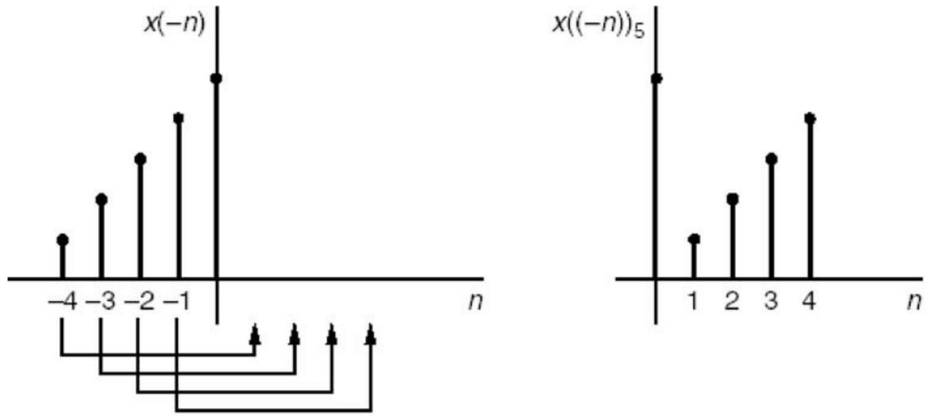
Original sequence

Linear shift



Circular shift





If the original sequence is defined in the time interval N_1 to N_2, then the circular shift can mathematically be represented as follows:

$$
\begin{aligned}
x_c[n] &= x[(n-L)]_N \\
&= x[(n-L) \bmod N] \\
&= x[(n-L+rN),\ \text{such that } N_1 \le n-L \le N_2]
\end{aligned}
$$

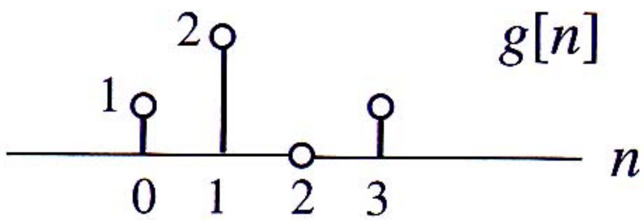
The circularly shifted sequence is obtained by finding an integer $\mathbf{r}$ such that $n\text{-}L\text{+}rN$ remains in the same domain as the original sequence.

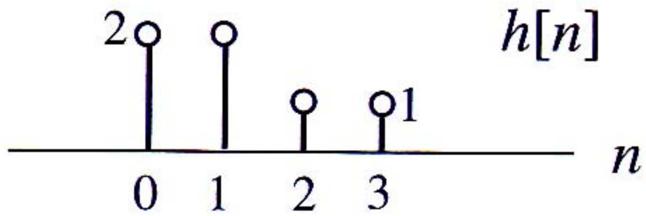
Circular Convolution

Linear convolution: $$y_L[n] = \sum_{m=0}^{N-1} g[m]h[n-m]$$

Circular convolution: $$y_C[n] = \sum_{m=0}^{N-1} g[m]h[< n-m >_N]$$

consider the following functions for a 4 point circular convolution





Circular Convolution

Evaluating $y_C[n] = \sum_{m=0}^{N-1} g[m]h[< n - m >_N]$ gives:

$$y_C[0] = g[0]h[0] + g[1]h[3] + g[2]h[2] + g[3]h[1]$$
$$y_C[1] = g[0]h[1] + g[1]h[0] + g[2]h[3] + g[3]h[2]$$
$$y_C[2] = g[0]h[2] + g[1]h[1] + g[2]h[0] + g[3]h[3]$$
$$y_C[3] = g[0]h[3] + g[1]h[2] + g[2]h[1] + g[3]h[0]$$

so we see that the 4 terms involve multiplying $g[n]$ with reversed and circularly shifted versions of $h[n]$ on the interval $n = 0 - 3$

Convolution Matrix

➲ The circular convolution can also be easily computed using the following **N-point** convolution matrix:

$$\boxed{y[n] = x[n] \circledast_N h[n]}$$

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \\ \vdots \\ y[N-1] \end{bmatrix} = \begin{bmatrix} h[0] & h[N-1] & h[N-2] & \cdots & h[1] \\ h[1] & h[0] & h[N-1] & \cdots & h[2] \\ h[2] & h[1] & h[0] & \cdots & h[3] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h[N-1] & h[N-2] & h[N-3] & \cdots & h[0] \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ x[2] \\ \vdots \\ x[N-1] \end{bmatrix}$$

The z-Transform
(The last one, really!)

- A generalization of the DTFT leads to the z-transform that may exist for many signals for which the DTFT does not.
  - DTFT is in fact a special case of the z-transform
    - …just like the Fourier transform is a special case of _______________(?)
- Furthermore, the use of the z-transform allows simple algebraic expressions to be used which greatly simplifies frequency domain analysis.
- Digital filters are designed, expressed, applied and represented in terms of the z-transform
- For a given sequence x[n], its z-transform X(z) is defined as

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

where z lies in the complex space, that is, $z = a + jb = re^{j\omega}$

Z ←→ DTFT

- From the definition of the z-variable

$$X(z)=\sum_{n=-\infty}^{\infty} x[n]z^{-n}=\sum_{n=-\infty}^{\infty} x[n](re^{j\omega})^{-n}=\sum_{n=-\infty}^{\infty} x[n]r^{-n}e^{-j\omega n}$$

  - It follows that the DTFT is indeed a special case of the z-transform, specifically, z-transform reduces to DTFT for the special case of $r=1$, that is, $|z|=1$, provided that the latter exists.

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

  - The contour $|z|=1$ is a circle in the z-plane of unit radius ➔ ***the unit circle***
  - Hence, the DTFT is really the z-transform evaluated on the unit circle
  - Just like the DTFT, z-transform too has its own, albeit less restrictive, convergence requirements, specifically, the infinite series $\sum_{n=-\infty}^{\infty} x[n]z^{-n}$ must converge.

- For a given sequence, the set R of values of z for which its z-transform converges is called the ***region of convergence (ROC).***

Convergence of the Z-Transform

- From our discussion with the DTFT, we know that the infinite series

$$X(z) = X(re^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] r^{-n} e^{-j\omega n}$$

converges if $x[n]r^{-n}$ is absolutely summable, that is, if

$$\boxed{\sum_{n=-\infty}^{\infty} \left| x[n] r^{-n} \right| < \infty}$$

- The area where this is satisfied defines the ROC, which in general is an annular region of the z-plane (since "z" is a complex number, constant z-values describe a circle in the z-plane)

$$R^- < |z| < R^+ \quad \text{where } 0 \le R^- < R^+ \le \infty$$

EXAMPLES

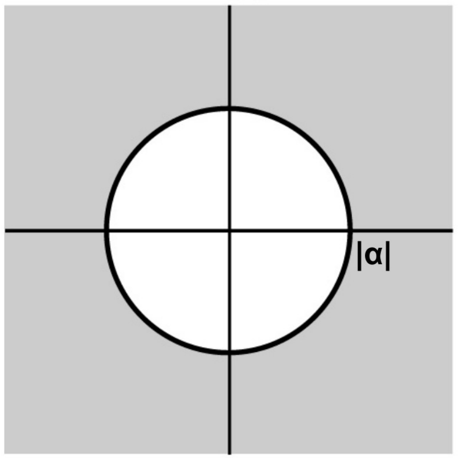
➲ Determine the z-transform and the corresponding ROC of the causal sequence $x[n]=\alpha^n u[n]$

$$X(z)=\frac{1}{1-\alpha z^{-1}}$$

for $\left|\alpha z^{-1}\right|<1$

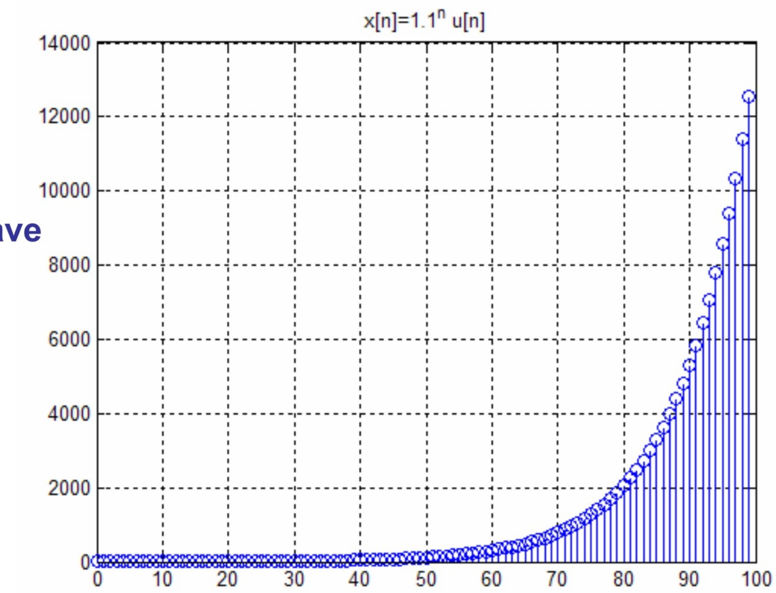
This power (geometric) series converges to

➔ ROC is the annular region $|z| > |a|$



•Note that this sequence does not have a DTFT if $|\alpha|>1$, however, it does have a z-transform!

•This is a right-sided sequence, which has an ROC that is outside of a circular area!



Examples

➲ The z-transform of the unit step sequence u[n] can be obtained from

$$X(z) = \frac{1}{1-\alpha z^{-1}}, \quad \text{for } \left|\alpha z^{-1}\right| < 1$$

by setting α=1 ➔ $$U(z) = \frac{1}{1-z^{-1}} = \frac{z}{z-1}, \quad \text{for } \left|z^{-1}\right| < 1$$

ROC is the annular regions |z|>1. Note that this sequence also does not have a DTFT!

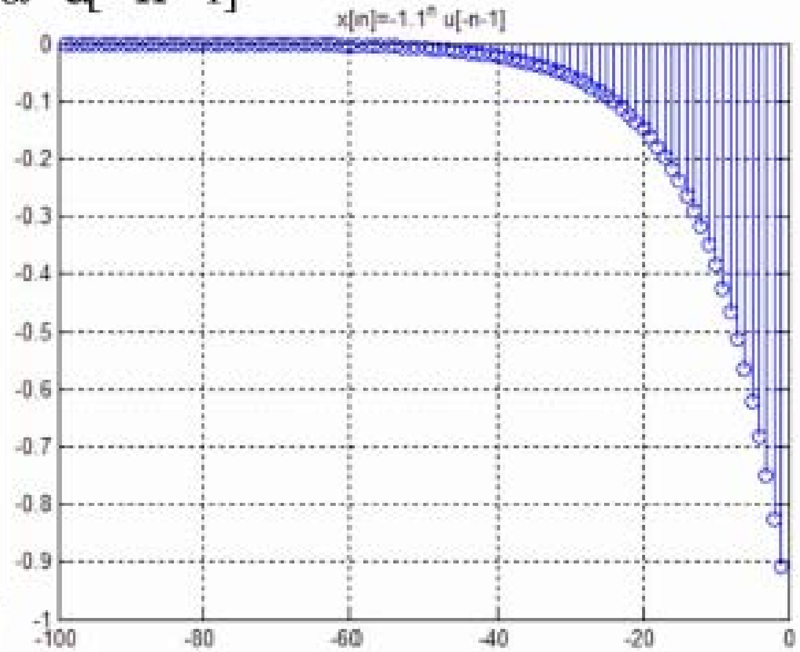
EXAMPLES

- Now consider the anti-causal sequence $y[n] = -\alpha^n u[-n-1]$

$$X(z) = -\sum_{n=-\infty}^{-1} \alpha^n z^{-n} = -\sum_{m=1}^{\infty} \alpha^{-m} z^m$$

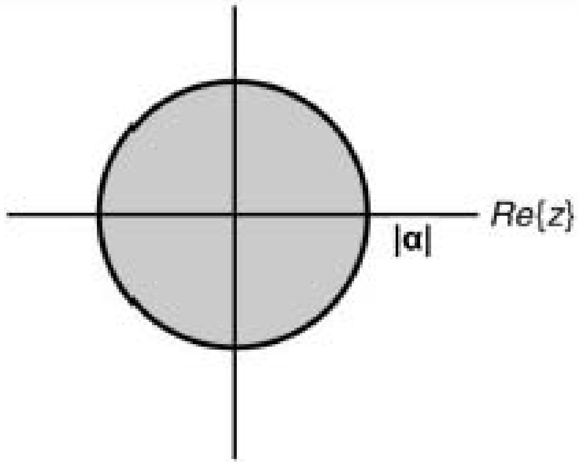
$$= -\alpha^{-1} z \sum_{n=0}^{\infty} \alpha^{-n} z^n = \frac{1}{1-\alpha z^{-1}}$$

for $\quad \left|\alpha^{-1} z\right| < 1$



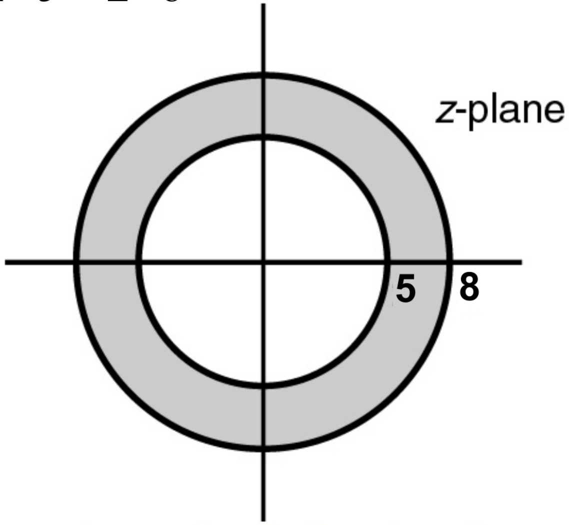
ROC is the annular region $|z| < |\alpha|$

- The z-transforms of the two sequences $\alpha^n u[n]$ and $-\alpha^n u[-n-1]$ are identical even though the two parent sequences are different
- Only way a unique sequence can be associated with a z-transform is by specifying its ROC
- This is a left-sided sequence, which has an ROC that is inside of a circular area!



EXAMPLE

- Consider $x[n]=5^n u[n]-8^n u[-n-1]$ ➔ $X(z)=\frac{z}{z-5}+\frac{z}{z-8}$
  - Corresponding ROCs are $|z|>5$ and $|z|<8$
  - Therefore the ROC for this signal is the annular region $5<|\mathbf{z}|<8$



  - Note that if the signal was $x[n]=8^n u[n]-5^n u[-n-1]$ the ROC would be empty! That is, the z-transform of this sequence does not exist!
  - **Now recall that DTFT was z-transform evaluated on the unit circle, that is for $z=e^{j\omega}$. Therefore, DTFT of a sequence exists (that is the series converges), if and only if the ROC includes the unit circle!**
  - The DTFT for the above example clearly does not exist, since the ROC does not include the unit circle!
  - Though, we must add that the existence of DTFT is not a guarantee for the existence of the z-transform.

Rational Z-Transforms

- A rational z-transform can be alternately written in factored form as

$$H(z) = \frac{b_0 \prod_{\ell=1}^{M} (1 - \zeta_\ell z^{-1})}{a_0 \prod_{\ell=1}^{N} (1 - p_\ell z^{-1})} = z^{(N-M)} \frac{p_0 \prod_{\ell=1}^{M} (z - \zeta_\ell)}{d_0 \prod_{\ell=1}^{N} (z - p_\ell)}$$

- At a root $z=\zeta_\ell$ of the numerator polynomial $H(\zeta_\ell)=0$, and as a result, these values of z are known as the ***zeros*** of $H(z)$
- At a root $z=p_\ell$ of the denominator polynomial $H(p_\ell) \rightarrow \infty$, and as a result, these values of z are known as the ***poles*** of $H(z)$
  - Note that H(z) has *M* finite zeros and *N* finite poles
  - If N > M there are additional *N-M* zeros at z = 0 (the origin in the z-plane)
  - If N < M there are additional *M-N* poles at z = 0
- Why is this important?
  - **As we will see later, a digital filter is designed by placing appropriate number of zeros at the frequencies (z-values) to be suppressed, and poles at the frequencies to be amplified!**

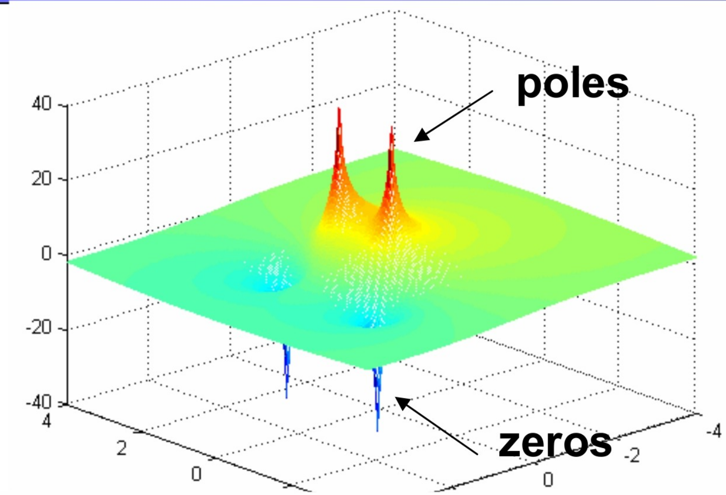
Some sense of Physical Interpretation of this Math Crap!

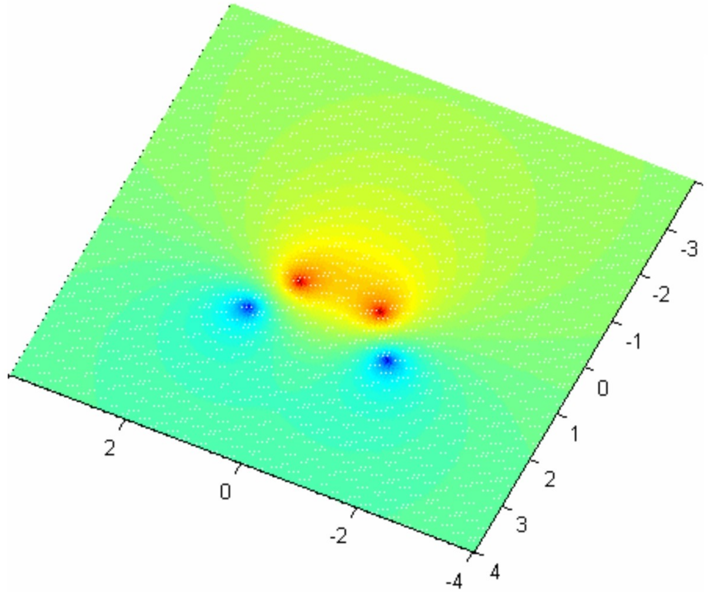
What does this look like???

$$G(z) = \frac{1 - 2.4\,z^{-1} + 2.88\,z^{-2}}{1 - 0.8\,z^{-1} + 0.64\,z^{-2}}$$

```
clear
close all
N=256;
rez=linspace(-4, 4, N);
imz=linspace(-4,4,N);
%create a uniform z-plane
for n=1:N
     z(n,:)=ones(1,N)*rez(n)+j*ones(1,N).*imz(1:N);
end
%Compute the H function on the z-plane
for n=1:N
   for m=1:N
      Hz(n,m)=H_fun(z(n,m));
   end
end
%Logarithmic mesh plot of the  H function
mesh(rez, imz, 20*log10(abs(Hz)))

======================================
function Hz=H_fun(z);
%Compute the transfer function
Hz=(1-2.4*z^(-1)+2.88*z^(-2))/(1-0.8*z^(-1)+0.64*z^(-2));
```





Stability & ROC In terms of Zeros & Poles

- Recall that for a ***system*** to be causal, its impulse response must satisfy h[n]=0, n<0, that is for a causal system, the impulse response is ***right sided.*** Based on this, and our previous observations, we can make the following important conclusions:
  - **The ROC of a causal system extends outside of the outermost pole circle**
  - **The ROC of an anticausal system (whose h[n] is purely left-sided) lies inside of the innermost pole circle**
  - **The ROC of a noncausal system (whose h[n] two-sided) is bounded by two different pole circles**
  - **Now, for an LTI system to be stable it must be absolutely summable, or in other words, it must have a DTFT. But for a system to have a DTFT, its ROC must include the unit circle. ➔ <u>An LTI system is stable, if and only if the ROC of its transfer function H(z) includes the unit circle!</u>**
  - **Furthermore, a causal system's ROC lies outside of a pole circle. If that system is also stable, its ROC must include unit circle➔ Then <u>a causal system is stable, if and only if, all poles are inside the unit circle!</u> Similarly, an anticausal system is stable, if and only if its poles lie outside the unit circle.**
    - **An FIR filter is always stable, why?**

Partial Fractional Expansion

Example 1: $H(z) = \dfrac{1}{(1-\frac{3}{4}z^{-1}+\frac{1}{8}z^{-2})} = \dfrac{1}{(1-\frac{1}{4}z^{-1})(1-\frac{1}{2}z^{-1})}$ $\quad |z| > 1/2$

this can be written as: $H(z) = \dfrac{A_1}{(1-\frac{1}{4}z^{-1})} + \dfrac{A_2}{(1-\frac{1}{2}z^{-1})}$

with: $A_1 = \left.\dfrac{1-\frac{1}{4}z^{-1}}{(1-\frac{1}{4}z^{-1})(1-\frac{1}{2}z^{-1})}\right|_{z=\frac{1}{4}} = -1; \qquad A_2 = \left.\dfrac{1-\frac{1}{2}z^{-1}}{(1-\frac{1}{4}z^{-1})(1-\frac{1}{2}z^{-1})}\right|_{z=\frac{1}{2}} = 2$

so: $H(z) = \dfrac{-1}{(1-\frac{1}{4}z^{-1})} + \dfrac{2}{(1-\frac{1}{2}z^{-1})}$

$h[n]$ is causal

$$h[n] = 2\left(\tfrac{1}{2}\right)^n \mu[n] - \left(\tfrac{1}{4}\right)^n \mu[n]$$

Partial Fractional Expansion

if $M \geq N$ a polynomial must be added of order $M - N$:

$$H(z) = \sum_{k=0}^{M-N} B_k z^{-k} + \sum_{i=0}^{N} \frac{A_i}{1 - d_i z^{-1}}$$

values of B_k are obtained by long division of the numerator of $H(z)$ by the denominator.

in the case of multiple poles of order s at z = d_i while all other poles are simple another term is required:

$$H(z) = \sum_{k=0}^{M-N} B_k z^{-k} + \sum_{k=0, k \neq i}^{N} \frac{A_k}{1 - d_k z^{-1}} + \sum_{m=1}^{s} \frac{C_m}{\left(1 - d_i z^{-1}\right)^m}$$

Partial Fraction Expansion

- Re-express the rational z-transform as a partial fraction expansion of simpler terms, whose inverse z-transforms are known.
  - Slightly different procedure depending on whether the system has simple poles or multiple poles
  - A rational H(z) can be expressed as

$$H(z)=\frac{Y(z)}{X(z)}=\frac{P(z)}{D(z)}=\frac{b_0+b_1 z^{-1}+b_2 z^{-2}+\cdots+b_M z^{-M}}{1+a_1 z^{-1}+a_2 z^{-2}+\cdots+a_N z^{-N}}=\frac{\sum_{i=0}^{M} b_i z^{-i}}{\sum_{i=0}^{N} a_i z^{-i}}$$

  - If $M \geq N$ then H(z) can be re-expressed through long division

$$\boxed{H(z)=\sum_{\ell=0}^{M-N}\eta_\ell z^{-\ell}+\frac{P_1(z)}{D(z)}}$$

$$H(z)=\frac{2+0.8z^{-1}+0.5z^{-2}+0.3z^{-3}}{1+0.8z^{-1}+0.2z^{-2}} \Rightarrow H(z)=-3.5+1.5z^{-1}+\frac{5.5+2.1z^{-1}}{1+0.8z^{-1}+0.2z^{-2}}$$

where the degree of $P_1(z)$ is less than N. The rational fraction $P_1(z)/D(z)$ is then called a ***proper fraction*** or ***proper polynomial*** .

Partial Fractional Expansion

Example 2:

$$H(z) = \frac{2z^2 + 4z + 2}{2z^2 - 3z + 1} = \frac{1 + 2z^{-1} + z^{-2}}{1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2}} = \frac{(1 + z^{-1})^2}{(1 - \frac{1}{2}z^{-1})(1 - z^{-1})}$$

As the numerator is the same order as the denominator we need to write *H(z)* as:

$$H(z) = B_0 + \frac{A_1}{(1 - \frac{1}{2}z^{-1})} + \frac{A_2}{(1 - z^{-1})} \qquad |z| > 1$$

we first need to determine B_o and a residual expression with a lower order numerator by dividing the denominator into the numerator

$$B_0 = + \frac{1 - 2z^{-1} + z^{-2}}{(1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2})} = 2 + \left[\frac{5z^{-1} - 1}{(1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2})} \right]$$

Partial Fractional Expansion

the next step is to further solve the second term: $\left[\dfrac{5z^{-1}-1}{(1-\frac{3}{2}z^{-1}+\frac{1}{2}z^{-2})}\right]$

and solve for A_1 and A_2:

$$A_1 = 2+\frac{5z^{-1}-1}{(1-\frac{1}{2}z^{-1})(1-z^{-1})}(1-\tfrac{1}{2}z^{-1})\Big|_{z=\frac{1}{2}} = -9$$

$$A_2 = 2+\frac{5z^{-1}-1}{(1-\frac{1}{2}z^{-1})(1-z^{-1})}(1-z^{-1})\Big|_{z=1} = +8$$

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

$h[n]$ is causal → $h[n] = 2\delta[n]-9\left(\frac{1}{2}\right)^n\mu[n]+8\mu[n]$

Frequency Response

- Recall that the output of an LTI system in the frequency domain is: $\boxed{H(\omega)=Y(\omega)/X(\omega)}$
  - where H(ω) is called the ***frequency response*** of the system, which relates the input and the output of an LTI system in the frequency domain. The frequency response can also be represented in terms of CCLDE coefficients:

$$\boxed{H(\omega)=\frac{\sum_{k=0}^{M} b_k e^{-j\omega k}}{\sum_{k=0}^{N} a_k e^{-j\omega k}}=\frac{b_0+b_1e^{-j\omega}+b_2e^{-j\omega 2}+\cdots+b_Me^{-j\omega M}}{a_0+a_1e^{-j\omega}+a_2e^{-j\omega 2}+\cdots+a_Ne^{-j\omega N}}}$$

- A generalization of the freq. response is the transfer func., computed in the z-domain
  - The function H(z), which is the z-transform of the impulse response h[n] of the LTI system, is called the ***transfer function*** or the ***system function*** $\boxed{H(z)=Y(z)/X(z)}$
  - Using the CCLDE coefficients

$$\boxed{H(z)=\frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}}=z^{(N-M)}\frac{\sum_{k=0}^{M} b_k z^{M-k}}{\sum_{k=0}^{N} a_k z^{N-k}}=\frac{b_0}{a_0}\cdot\frac{\prod_{k=1}^{M}(1-\zeta_k z^{-1})}{\prod_{k=1}^{N}(1-p_k z^{-1})}=\frac{b_0}{a_0}z^{(N-M)}\frac{\prod_{k=1}^{M}(z-\zeta_k)}{\prod_{k=1}^{N}(z-p_k)}}$$

CCLDE coefficients

Zeros & poles

Zero & pole factors

Frequency Response ←→ Transfer Function

- If the ROC of the transfer function $H(z)$ includes the unit circle, then the frequency response $H(\omega)$ of the LTI digital filter can be obtained simply as follows: $\boxed{H(e^{j\omega}) = H(\omega) = H(z)\big|_{z=e^{j\omega}}}$

- Assuming that the DTFT exists, starting with the factored z-transform, we can write the frequency response of a typical LTI system as

$$H(z) = \frac{b_0}{a_0} z^{(N-M)} \frac{\prod_{k=1}^{M}(z-\zeta_k)}{\prod_{k=1}^{N}(z-p_k)} \Rightarrow H(e^{j\omega}) = \frac{b_0}{a_0} e^{j\omega(N-M)} \frac{\prod_{k=1}^{M}(e^{j\omega}-\zeta_k)}{\prod_{k=1}^{N}(e^{j\omega}-p_k)}$$

From which we can obtain the magnitude and phase response:

$$\left|H(e^{j\omega})\right| = \left|\frac{b_0}{a_0}\right| \frac{\prod_{k=1}^{M}\left|e^{j\omega}-\zeta_k\right|}{\prod_{k=1}^{N}\left|e^{j\omega}-p_k\right|}$$

$$\arg H(e^{j\omega}) = \arg(b_0/a_0) + \omega(N-M) + \sum_{k=1}^{M}\arg(e^{j\omega}-\zeta_k) - \sum_{k=1}^{N}\arg(e^{j\omega}-p_k)$$

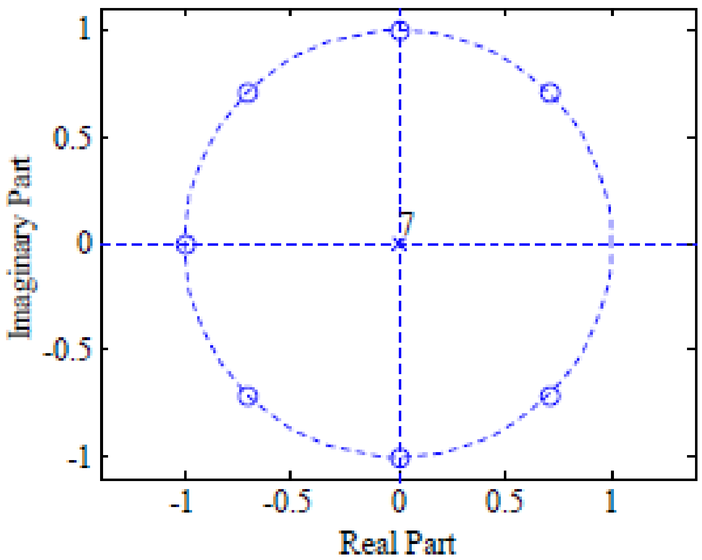
An Example

- Consider the M-point moving-average FIR filter with an impulse response

$$h[n]=\begin{cases}1/M, & 0\le n\le M-1\\ 0, & \text{otherwise}\end{cases}$$

$$H(z)=\frac{1}{M}\sum_{n=0}^{M-1} z^{-n}=\frac{1-z^{-M}}{M\left(1-z^{-1}\right)}=\frac{z^{M}-1}{M[z^{M-1}(z-1)]}$$

- Observe the following
  - The transfer function has M zeros on the unit circle at* $z=e^{j2\pi k/M}$,$0\le k\le M-1$
  - There are M-1 poles at $z=0$ and a single pole at $z=1$
  - The pole at $z=1$ exactly cancels the zero at $z=1$
  - The ROC is the entire z-plane except $z=0$



***To see this, try**
zplane(roots([1,0,0…,0, -1]))

Zero Phase Filters

- One way to avoid any phase distortion is to make sure the frequency response of the filter does not delay any of the spectral components. Such a transfer function is said to have a ***zero – phase*** characteristic.
- A zero – phase transfer function has no phase component, that is, the spectrum is purely real (no imaginary component) and non-negative
- However, it is NOT possible to design a causal digital filter with a zero phase.

LINEAR PHASE

- Note that a zero-phase filter cannot be implemented for real-time applications. Why?

- For a causal transfer function with a nonzero phase response, the phase distortion can be avoided by ensuring that the transfer function has (preferably) a unity magnitude and a ***linear-phase*** characteristic in the frequency band of interest

$$H(\omega) = e^{-j\alpha\omega} \quad \Rightarrow \quad |H(\omega)| = 1 \qquad \angle H(\omega) = \theta(\omega) = -\alpha\omega$$

  - Note that this phase characteristic is linear for all ω in $[0\ 2\pi]$.
  - Recall that the phase delay at any given frequency ω_0 was
  - If we have linear phase, that is, $\theta(\omega) = -\alpha\omega$, then the total delay at any frequency ω_0 is $\tau_0 = -\theta(\omega_0)/\ \omega_0 = -\alpha\omega_0/\omega_0 = \alpha$
  - Note that this is identical to the group delay $d\theta(\omega)/d\omega$ evaluated at ω_0

$$\boxed{\tau_p(\omega_o) = -\frac{\theta(\omega_o)}{\omega_o}}$$

$$\boxed{\tau_g(\omega_0) = -\left.\frac{d\theta(\omega)}{d\omega}\right|_{\omega=\omega_0}}$$

Linear Phase Filters

- It is typically impossible to design a linear phase IIR filter, however, designing FIR filters with precise linear phase is very easy:
- Consider a causal FIR filter of length M+1 (order M)

$$H(z) = \sum_{n=0}^{N} h[n]\, z^{-n} = h[0] + h[1]z^{-1} + h[2]z^{-2} + \cdots + h[M]z^{-M}$$

  - This transfer function has linear phase, if its impulse response h[n] is either ***symmetric***

$$\boxed{h[n] = h[M-n], \quad 0 \le n \le M}$$

or ***anti-symmetric***

$$\boxed{h[n] = -h[M-n], \quad 0 \le n \le M}$$

LINEAR PHASE FILTERS

This linear phase filter description can be generalised into a formalism for four type of FIR filters:

Type 1: symmetric sequence of odd length

Type 2: symmetric sequence of even length

Type 3: anti-symmetric sequence of odd length

Type 4: anti-symmetric sequence of even length

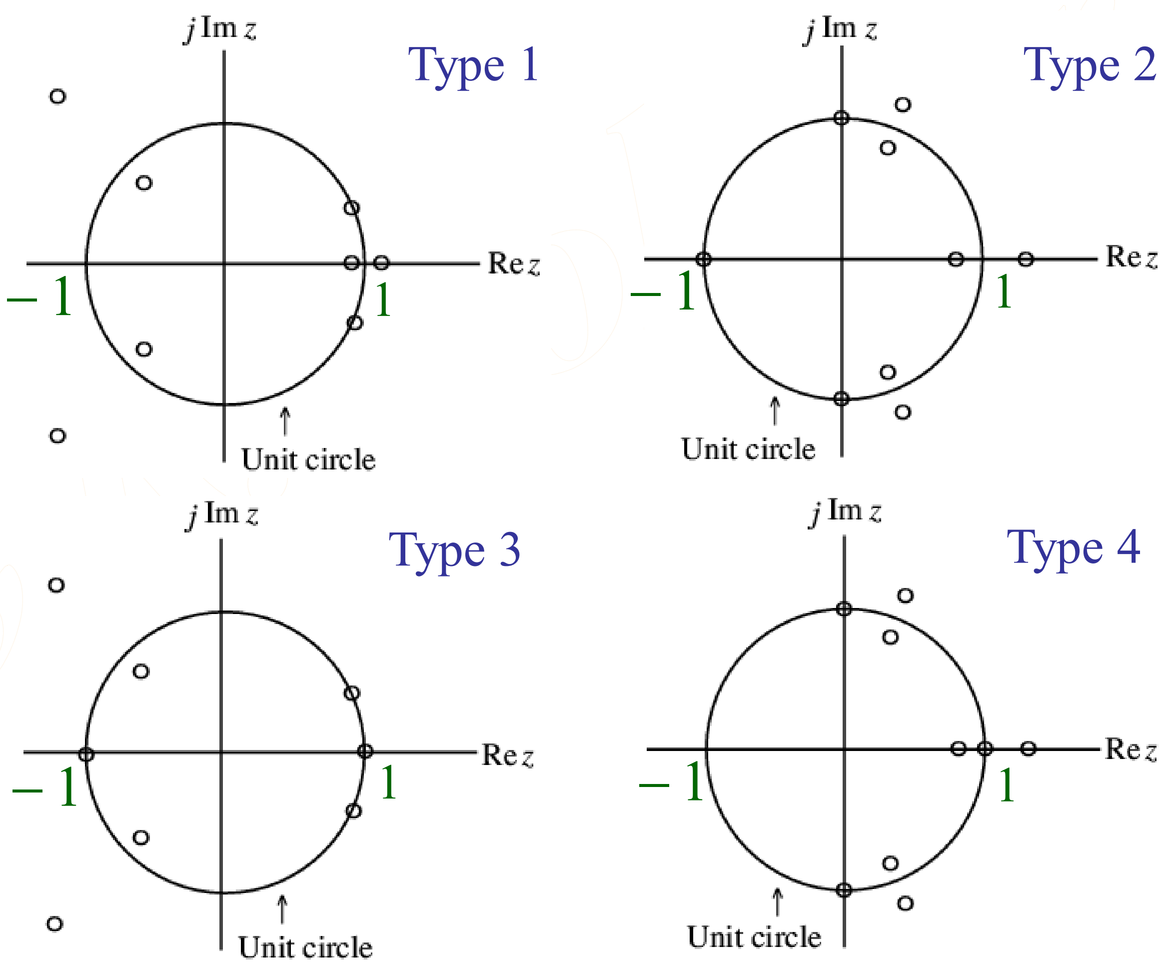
Linear Phase Filter Zero Locations

Type 1 FIR filter: Either an even number or no zeros at z = 1 and z=-1

Type 2 FIR filter: Either an even number or no zeros at z = 1, and an odd number of zeros at z=-1

Type 3 FIR filter: An odd number of zeros at z = 1 and and z=-1

Type 4 FIR filter: An odd number of zeros at z = 1, and either an even number or no zeros at z=-1



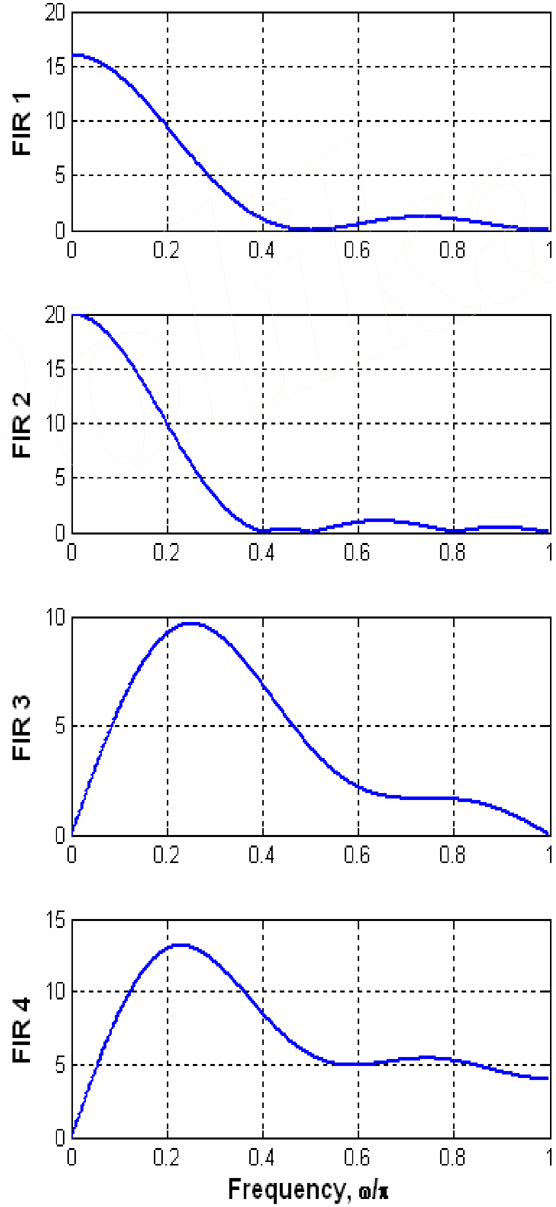
The presence of zeros at z=±1 leads to some limitations on the use of these linear-phase transfer functions for designing frequency-selective filters

A Type 2 FIR filter cannot be used to design a highpass filter since it always has a zero at z=-1

A Type 3 FIR filter has zeros at both z = 1 and z=-1, and hence cannot be used to design either a lowpass or a highpass or a bandstop filter

A Type 4 FIR filter is not appropriate to design a lowpass filter due to the presence of a zero at z = 1

Type 1 FIR filter has no such restrictions and can be used to design almost any type of filter



IIR LPF Filters

- A first-order causal lowpass IIR digital filter has a transfer function given by

$$\boxed{H_{LP}(z) = \frac{1-\alpha}{2}\left(\frac{1+z^{-1}}{1-\alpha z^{-1}}\right) = \frac{1-\alpha}{2}\left(\frac{z+1}{z-\alpha}\right)}$$

where $|\alpha| < 1$ for stability

Normalization term that ensures that the gain at $\omega = 0$ is 0 dB (magnitude of 1)

- The above transfer function has a zero at $z=-1$ i.e., at $\omega = \pi$ which is in the stopband
- $H_{LP}(z)$ has a real pole at $z = \alpha$
- As ω increases from 0 to π, the magnitude of the zero vector decreases from a value of 1 to 0, whereas, for a positive value of α, the magnitude of the pole vector increases from a value of $1-\alpha$ to $1+\alpha$
- The maximum value of the magnitude function is 1 at $\omega = 0$, and the minimum value is 0 at $\omega = \pi$

$$|H_{LP}(e^{j0})| = 1, \qquad |H_{LP}(e^{j\pi})| = 0 \Rightarrow \text{LPF}$$

Higher order IIR Filters

- Again, note that any of these filters can be designed to be of higher order, which typically provides sharper and narrower transition bands
- Simply cascade a basic filter structure as many times necessary to achieve the higher order filter.
- For example, to cascade K-first order LPFs

$$H_{LP}(z) = \frac{1-\alpha}{2}\left(\frac{1+z^{-1}}{1-\alpha z^{-1}}\right) \quad \Rightarrow \quad G_{LP}(z) = \left(\frac{1-\alpha}{2}\cdot\frac{1+z^{-1}}{1-\alpha z^{-1}}\right)^{K}$$

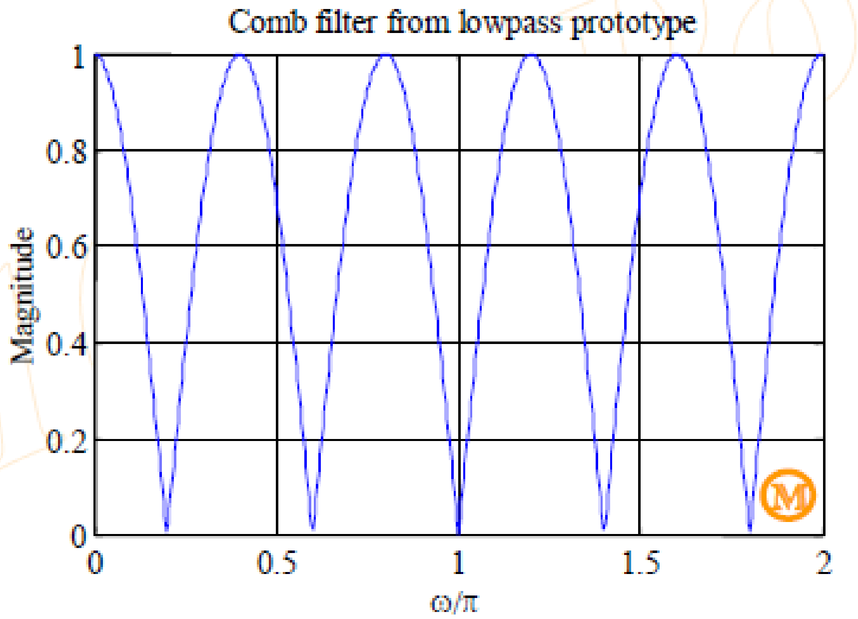
  - It can be shown that

$$\alpha = \frac{1+(1-C)\cos\omega_c - \sin\omega_c\sqrt{2C-C^2}}{1-C+\cos\omega_c} \qquad C = 2^{(K-1)/K}$$

Comb Filters

➲ Starting from a lowpass transfer function:

$$H(z) = \frac{1}{2}(1+z^{-1}) \Rightarrow H_{comb}(z) = H(z^L) = \frac{1}{2}(1+z^{-L})$$



(L=5) L *notches* are at ω=(2k+1)π/L and L *peaks* are at ω=2πk/L

$$0 \leq \omega < 2\pi$$

IIR Filter Structures

- The causal IIR digital filters are characterized by a real rational transfer function of z^{-1}, or equivalently, by a constant coefficient difference equation.
- From the difference equation representation, it can be seen that the realization of the causal IIR digital filters requires some form of feedback. Furthermore,
  - An N^{th} order IIR digital transfer function is characterized by $2N+1$ unique (***a*** and ***b***) coefficients, and in general, requires $2N+1$ multipliers and $2N$ two-input adders for implementation

$$\sum_{k=0}^{N-1} a_0\, y[n-k] = \sum_{l=0}^{M-1} b_0\, x[n-l]$$

$$y[n] = b_0 x[n] + b_1 x[n-1] + \cdots + b_{M-1} x[n-M+1] - a_1 y[n-1] - \cdots - a_{N-1} y[n-N+1]$$

- Given the filter CCLDE, we can implement it directly using the multiplier coefficients. This is called **Direct Form I** implementation.

IIR Filter Structures

- Consider a 3rd order example:

$$H(z) = \frac{P(z)}{D(z)} = \frac{p_0 + p_1 z^{-1} + p_2 z^{-2} + p_3 z^{-3}}{1 + d_1 z^{-1} + d_2 z^{-2} + d_3 z^{-3}}$$

  - Which can be split into two systems of H_1 (numerator) and H_2 (denominator):

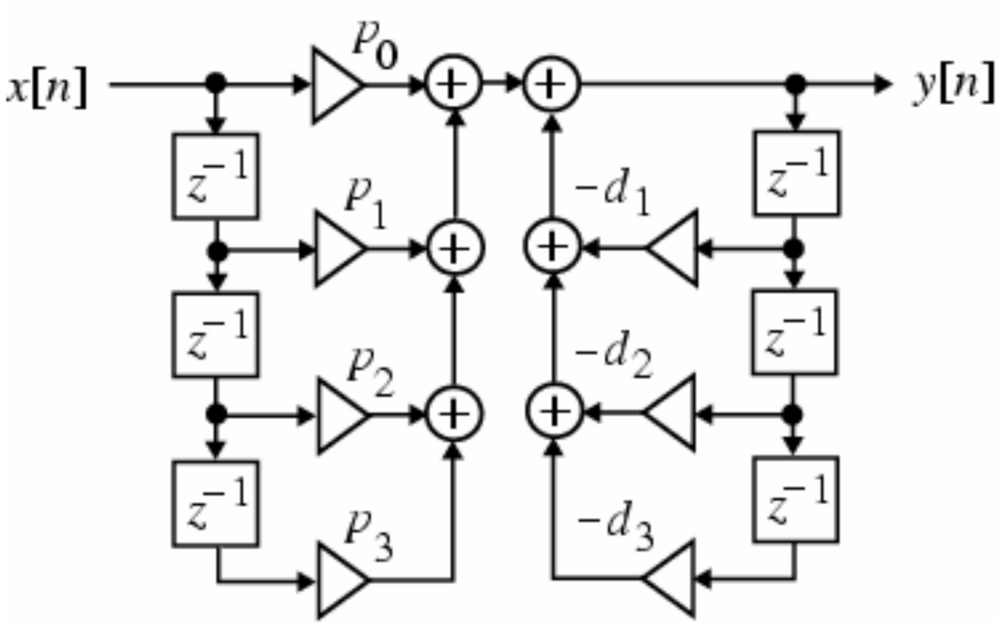
$$X(z) \longrightarrow \boxed{H_1(z)} \xrightarrow{W(z)} \boxed{H_2(z)} \longrightarrow Y(z)$$

$$\boxed{H_1(z) = \frac{W(z)}{X(z)} = P(z) = p_0 + p_1 z^{-1} + p_2 z^{-2} + p_3 z^{-3}}$$

$$\boxed{H_2(z) = \frac{Y(z)}{W(z)} = \frac{1}{D(z)} = \frac{1}{1 + d_1 z^{-1} + d_2 z^{-2} + d_3 z^{-3}}}$$

IIR Filter Structures: Direct Form I

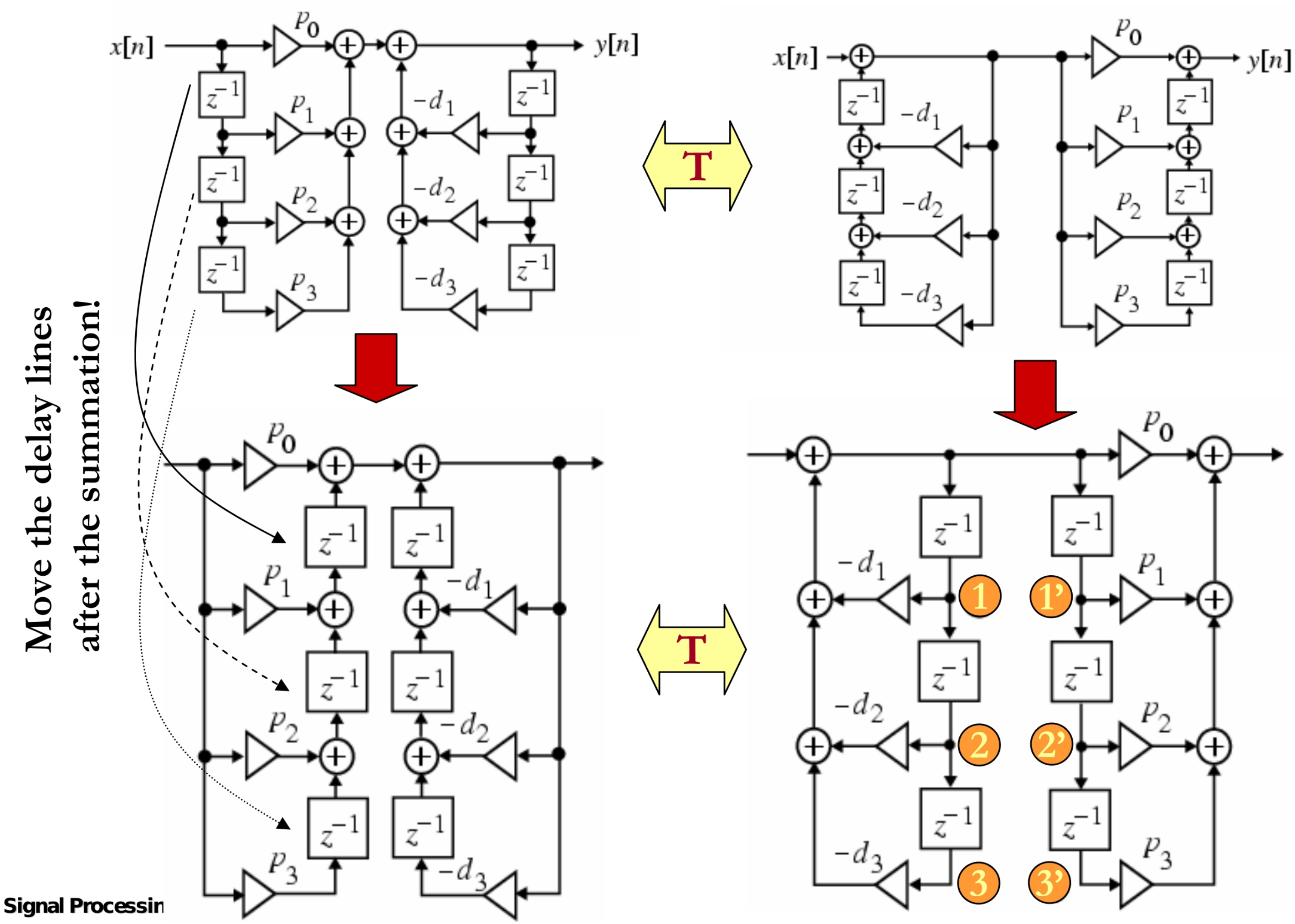
- A cascade of the two then gives us the overall H(z), whose implementation is known as **Direct Form I** implementation



  - Note that this structure is ***noncanonic*** since it employs 6 delays to realize a 3rd-order transfer function

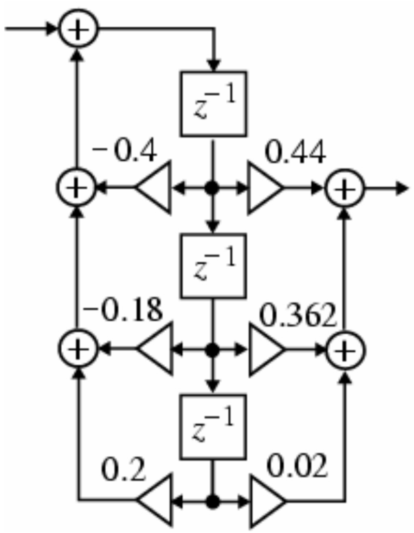
Other Non-Canonic Implementations

➲ Notice that we can also implement these structures as follows:

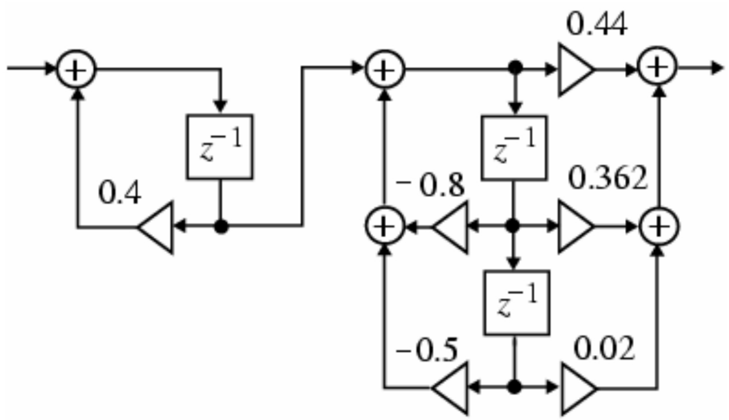


An Example

$$H(z)=\frac{0.44z^{-1}+0.362z^{-2}+0.02z^{-3}}{1+0.4z^{-1}+0.18z^{-2}-0.2z^{-3}}=\left(\frac{0.44+0.362z^{-1}+0.02z^{-2}}{1+0.8z^{-1}+0.5z^{-2}}\right)\left(\frac{z^{-1}}{1-0.4z^{-1}}\right)$$



Direct form II

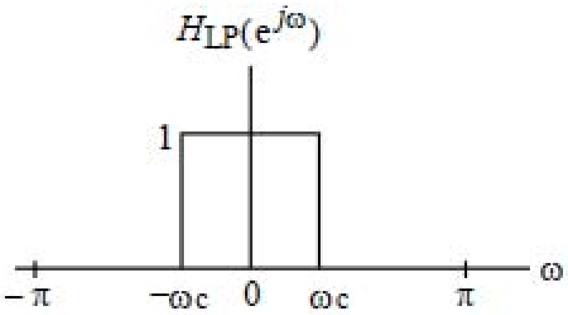


Cascade form

FIR or IIR???

- Several advantages to both FIR and IIR type filters.
- Advantages of FIR filters (disadvantages of IIR filters):
  - Can be designed with exact linear phase,
  - Filter structure always stable with quantized coefficients
  - The filter startup transients have finite duration
- Disadvantages of FIR filters (advantage of IIR filters)
  - Order of an FIR filter is usually much higher than the order of an equivalent IIR filter meeting the same specifications → higher computational complexity
  - In fact, the ratio of orders of a typical IIR filter to that of an FIR filter is in the order of tens.
- The nonlinear phase of an IIR filter can be minimized using an appropriate allpass filter, however, by that time the computational advantage of IIR is lost.
- However, in most applications that does not require real-time operation, phase is not an issue. Why…?

FIR (LOWPASS) FILTER DESIGN



Unrealizable!

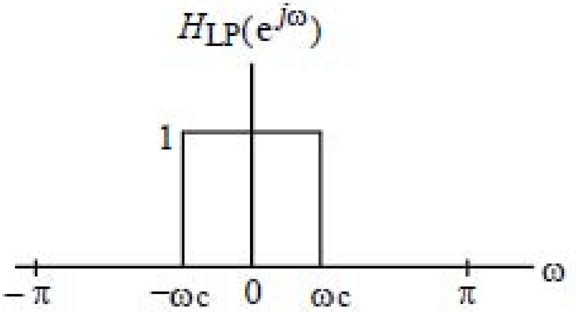
$$h_{LP[n]} = \frac{\sin(\omega_c n)}{\pi n}, \quad -\infty < n < \infty$$

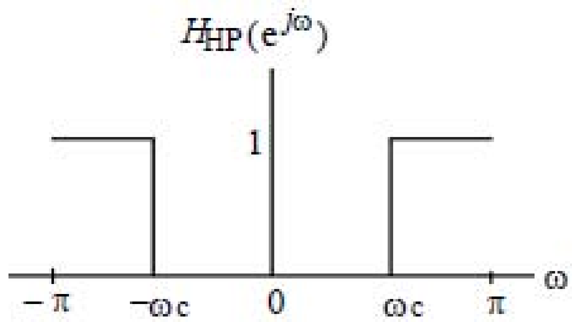
Realizable!

WINDOWING: zero coefficients outside $-M/2 \leq n \leq M/2$ and a shift to the right yields finite series with length $M+1$

$$h_{LP}[n] = \begin{cases} \dfrac{\sin(\omega_c (n - M/2))}{\pi(n - M/2)}, & 0 \leq n \leq M, n \neq \dfrac{M}{2} \\ \dfrac{\omega_c}{\pi}, & n = \dfrac{M}{2} \end{cases}$$

FIR Highpass Design

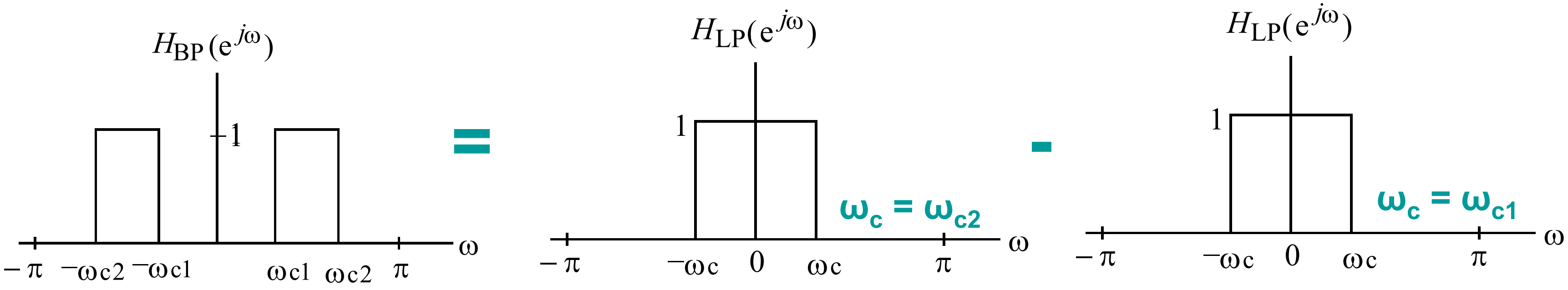




$$H_{HP}(\omega) = 1 - H_{LP}(\omega) \quad \Longleftrightarrow \quad h_{HP}[n] = \delta[n] - h_{LP}[n]$$

$$h_{HP}[n] = \begin{cases} -\dfrac{\sin(\omega_c n)}{\pi n}, & |n| > 0 \\ 1\text{-}\ \dfrac{\omega_c}{\pi}, & n = 0 \end{cases}$$

FIR BPF / BSF DESIGN



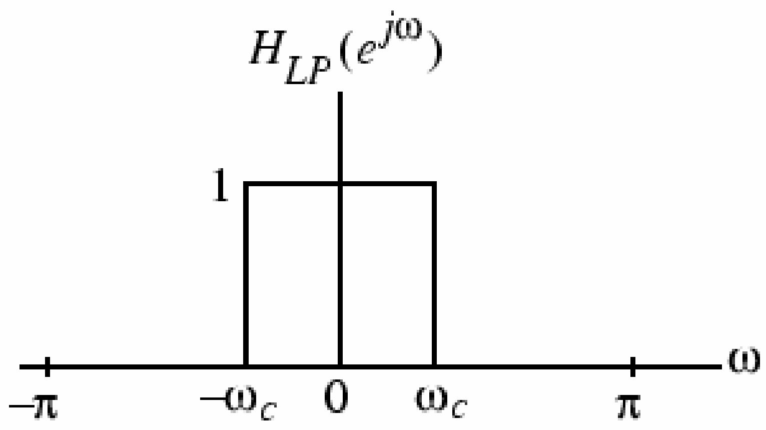
$$H_{BP}(\omega) = H_{LP}(\omega)\Big|_{\omega_c=\omega_{c2}} - H_{LP}(\omega)\Big|_{\omega_c=\omega_{c1}} \Longleftrightarrow h_{BP}[n] = h_{LP}[n]\Big|_{\omega_c=\omega_{c2}} - h_{LP}[n]\Big|_{\omega_c=\omega_{c1}}$$

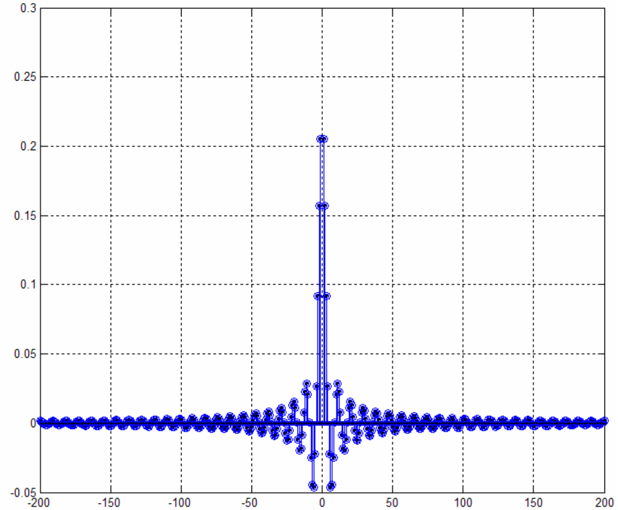
$$h_{BP[n]} = \begin{cases} \dfrac{\sin\left(\omega_{c_2}(n-M/2)\right)}{\pi(n-M/2)} - \dfrac{\sin\left(\omega_{c_1}(n-M/2)\right)}{\pi(n-M/2)}, & 0 < n < M,\ n \neq \dfrac{M}{2} \\ \dfrac{\omega_{c_2}}{\pi} - \dfrac{\omega_{c_1}}{\pi}, & n = \dfrac{M}{2} \end{cases}$$

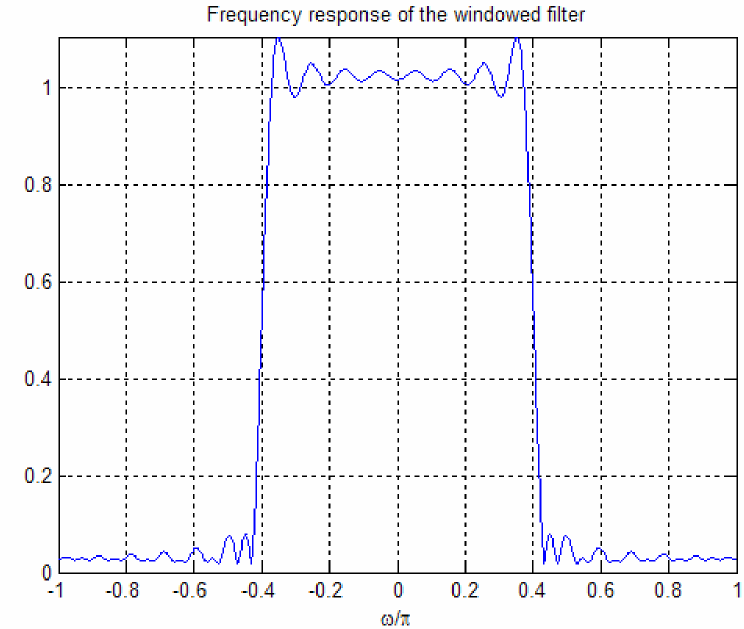
Similarly,

$$H_{BS}(\omega) = 1 - H_{BP}(\omega) \Longleftrightarrow h_{BS}[n] = \delta[n] - h_{BP}[n]$$

HOWEVER…

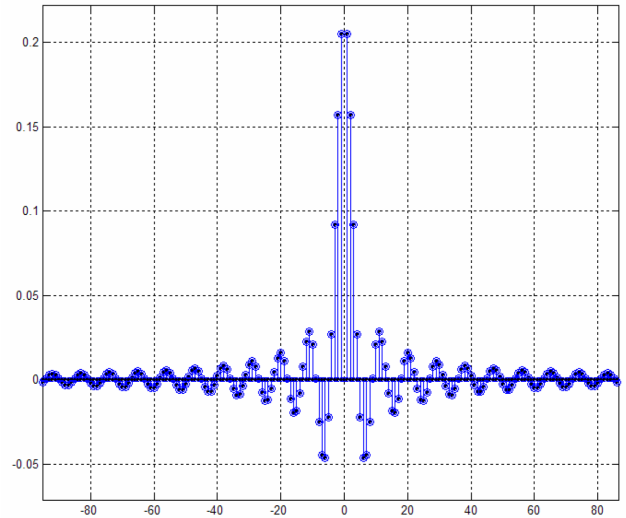




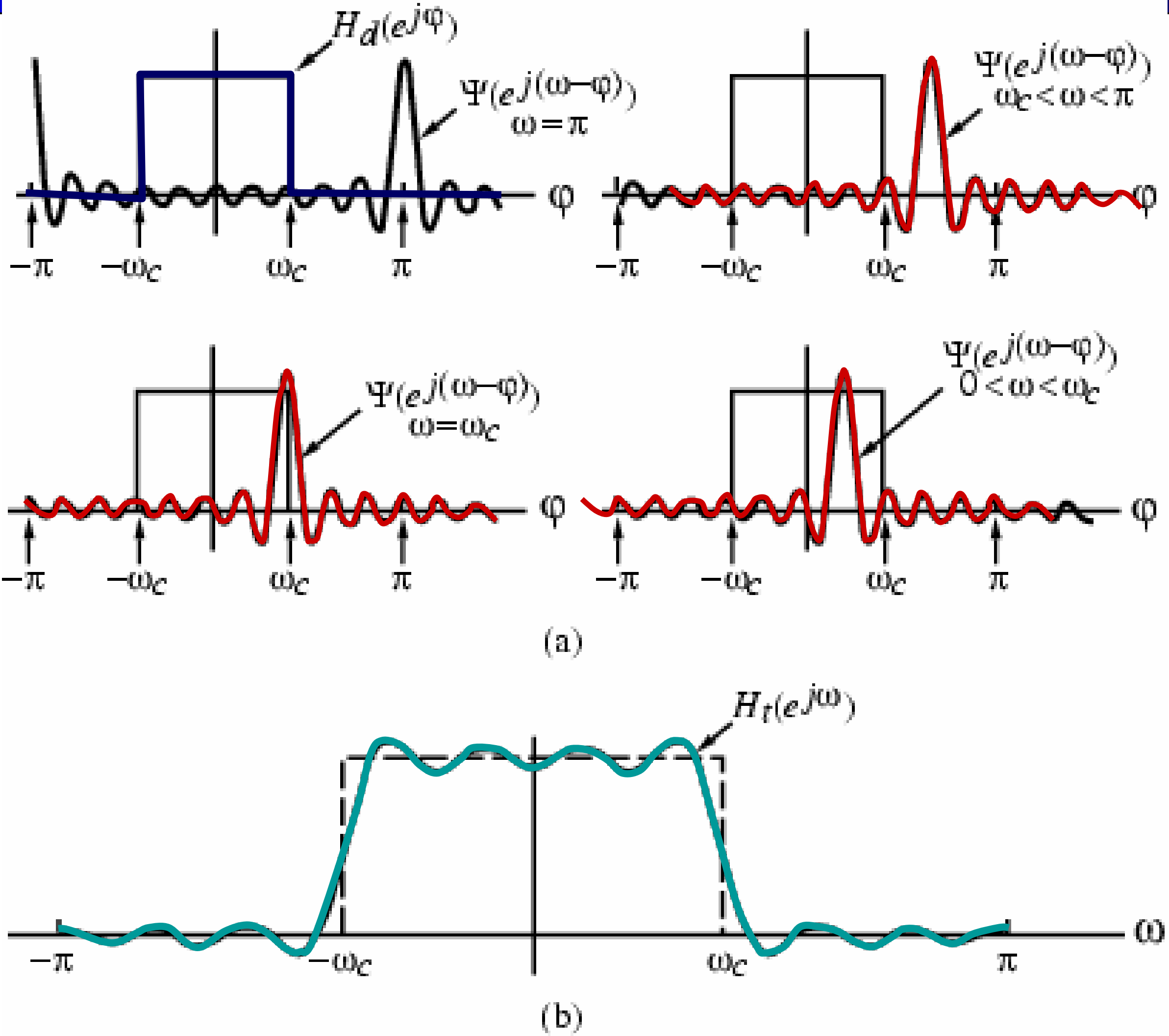


?

What happened…?



Gibbs Phenomenon



H_d: Ideal filter frequency response

Ψ: Rectangular window frequency response

H_t: Truncated filter's frequency response