ELEC2100: Signals and Systems

Lecture 16 Sampling Theorem and Aliasing (Application)

(Reference: O&W Chapter 7)

I. Introduction to Nyquist's Sampling Theorem

II. Proof of the Sampling Theorem

III. Example - Sampling by Gating

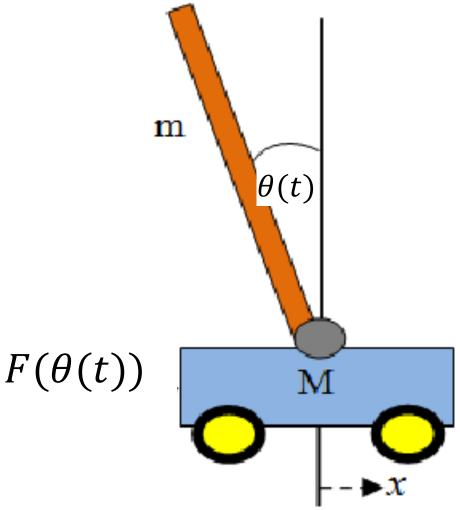
IV. Aliasing

I. Introduction to Nyquist's Sampling Theorem

- 50 years ago, engineering systems were primarily ***analog*** systems that processed CT signals.

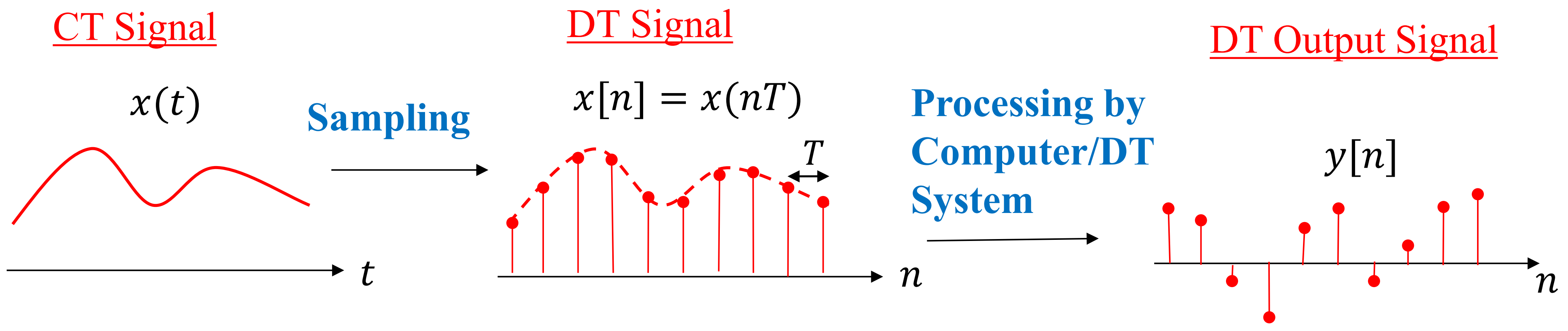
- Today, real-world CT signals are often converted into digital DT signals via ***sampling*** to be processed by digital computing systems. Resulting DT signals are then converted back to CT signals via ***interpolation*** to actuate physical systems.

- Advantages in processing by digital systems:
  1. Compact and low cost because of transistors and integrated circuits
  2. Immunity to noise and distortion; can be made error-free
  3. Parameter values can be made exact in computer program
  4. Easier to build more complex (higher order) digital systems giving better performance
  5. Can program and re-program.

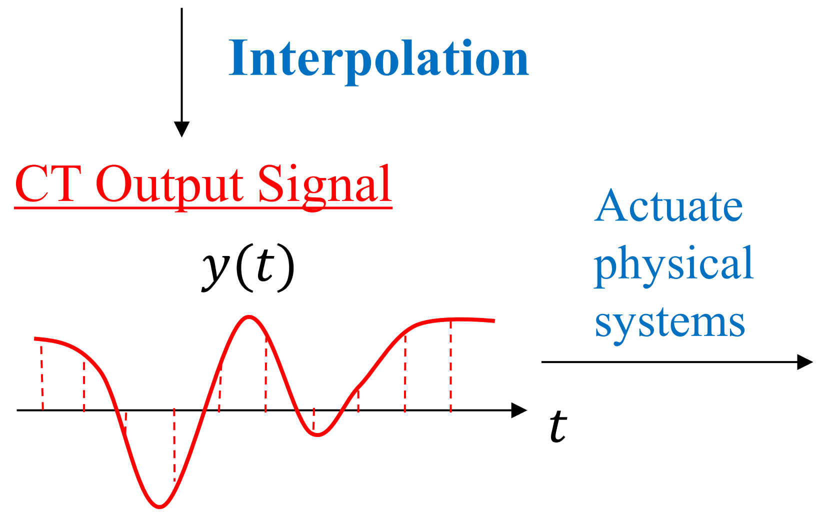


$$\theta(t) \rightarrow \begin{matrix}\mathit{Control}\\ \mathit{System}\end{matrix} \rightarrow F(\theta(t))$$

Sampling, Processing, and Interpolation



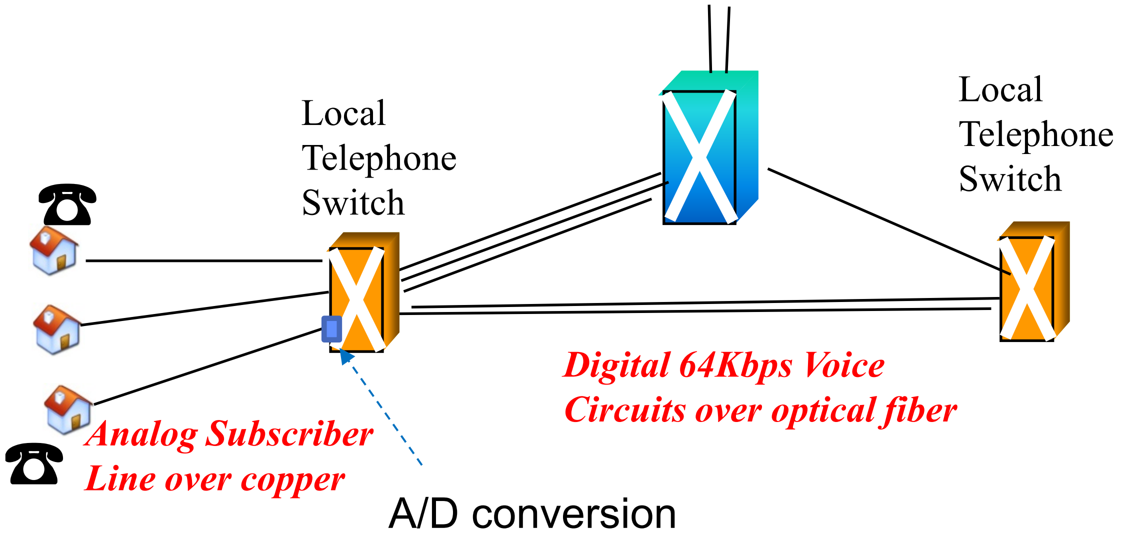
Intuitively, the DT signal $x[n] = x(nT)$ can represent the CT signal $x(t)$ very well if the sampling interval T is "small enough".

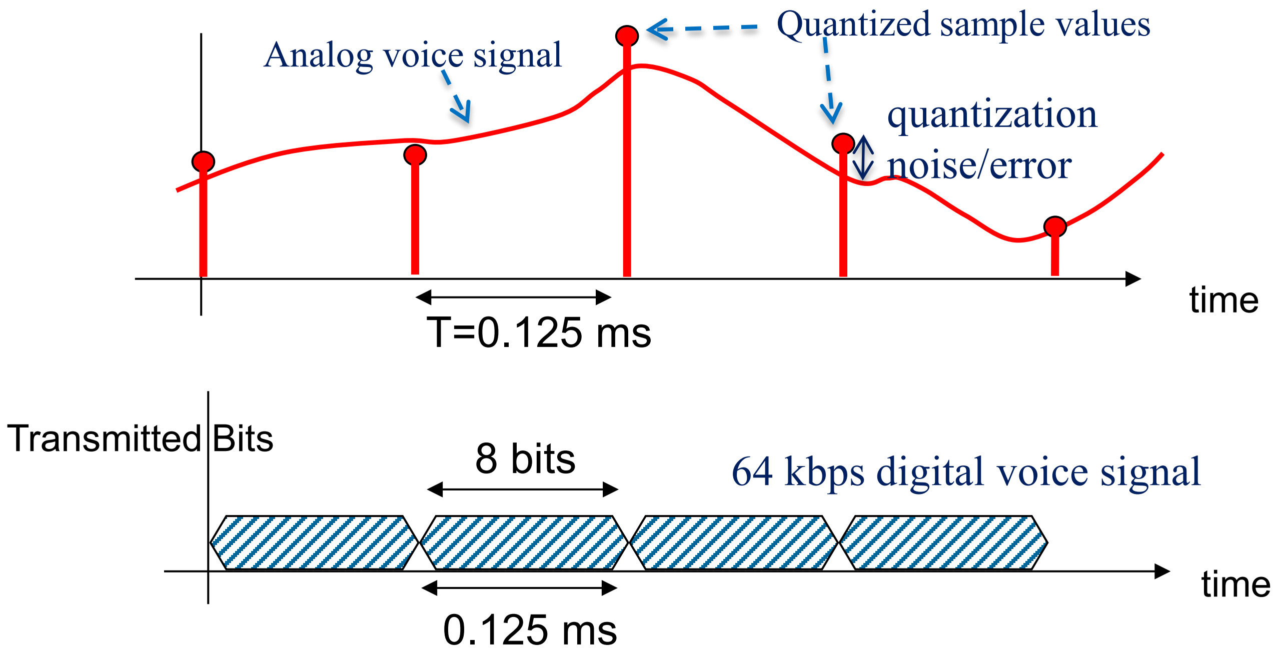


Examples – DT processing of CT Signals

- **Example 1 –Digital Telephony** The telephone network became digital in the 1970's.

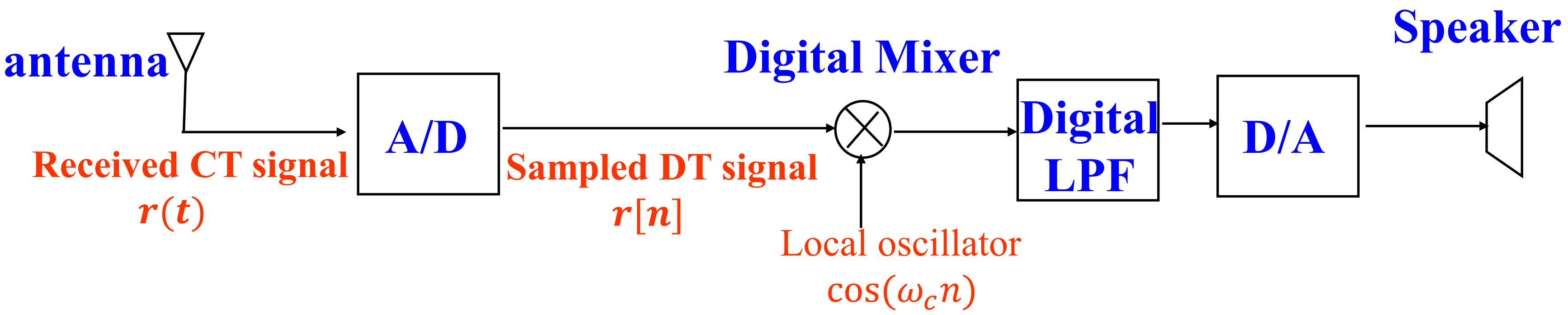
CT signal from each telephone subscriber line is sampled at 8,000Hz (T=125 μsec) by the telephone switch. 8 bits are used to represent each sample. The resulting 64Kbps digital signal is transmitted across the telephone network.





- **Example 2 – Compact Disc Music** CD arrived around 1990, replacing vinyl albums.
- In CD, music is sampled at 44.1KHz ($T=22.68\ \mu sec$), using 16 bits per sample for each of two stereo channels, for a total data rate of 1.41 Mbps (44.1Kx16x2). Data rate is 25 times of the digital telephony signal.

- **<u>Example 3 – Digital Radio Receiver</u>** Recall that to demodulate an AM radio channel, the receiver mixes the received signal with a local oscillator at the carrier frequency, and then applies a low-pass filter. In your cell phone, iPod, etc., the mixing and filtering are now all done digitally in a tiny electronic "chip".



Treating $x[n]$ as Real Number and Focusing on the Sampling Rate

- As can be seen in the telephony and compact disc example, for a digital signal, one can only use a finite number of bits to represent a sample. But,
  - In ELEC2100, we assume that $x[n]$ has infinite accuracy and focus on understanding how we should choose the sampling interval T.
  - That is why we call $x[n]$ a "*discrete-time signal (DT)*" instead of a "*digital signal*"
- T is the sampling interval and $\frac{1}{T}$ is the sampling rate
- Intuitively, for a DT signal $x[n]$=$x(nT)$ to represent the CT signal $x(t)$ accurately, we need the sampling rate to be high.
- But the smaller T is, the more data we need to store, transmit and process.
- What determines the sampling rate that is needed?

Representing CT Signals from Samples

- In 1927, Harry Nyquist of Bell Labs derived the ***sampling theorem***, which states that if a CT signal is "smooth" enough, and the sampling interval T is small enough, then:

  In theory the CT signal can be uniquely recovered from its samples!

- Implies that it is theoretically possible to convert signals back-and-forth from CT to DT without information loss.
- Nyquist's result is surprising! Consider Fig. 7.1 of the text – since different CT signals that produce the same set of samples, how can we tell which CT signal the samples represent?

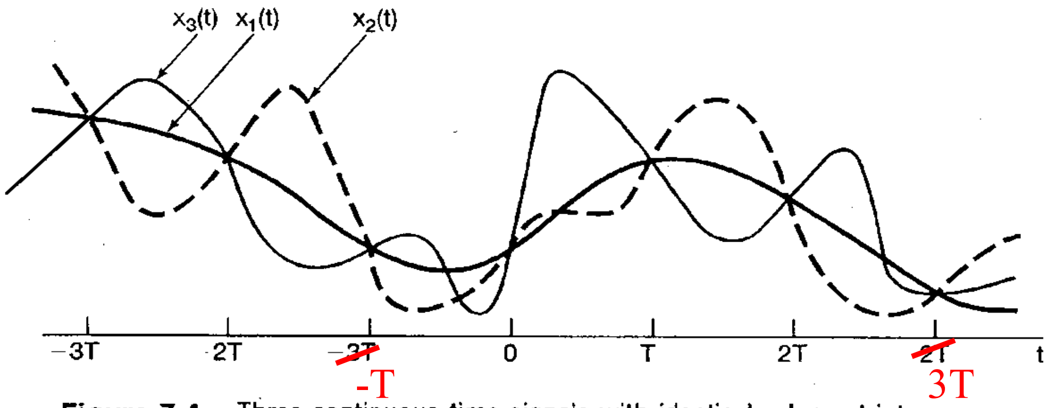


Figure 7.1 Three continuous-time signals with identical values at integer multiples of T.

$$x_1(t) \neq x_2(t) \neq x_3\ (t)$$

but

$$x_1(nT) = x_2(nT) = x_3\ (nT) \quad \forall n$$

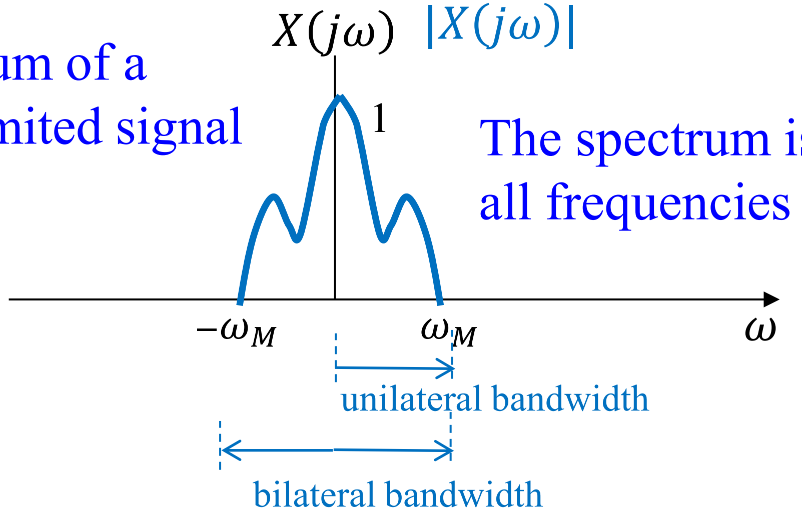
Bandwidth of a Signal

The key in the definition of "smoothness" of a signal:

- In Lecture 15, we introduced the notion that signal is often ***bandlimited*** with ***bandwidth*** often defined by its ***maximum frequency***.
- Specifically, $x(t)$ with FT $X(j\omega)$ is band-limited with ***unilateral bandwidth*** ω_M if :

$$\boxed{X(j\omega) = \ 0 \qquad \forall\, \omega \ \ \text{s.t.}\ |\omega| > \ \omega_M}$$

Spectrum of a bandlimited signal



The spectrum is zero for all frequencies $|\omega| > \omega_M$

- As explained in lecture 15, the bilateral bandwidth of a real signal is twice its unilateral bandwidth since spectrum of a real signal is conjugate symmetric.

Nyquist's Sampling theorem

Stated as follow:

- Let $x(t)$ be a ***band-limited*** signal so that $X(j\omega) = 0$ for $|\omega| > \omega_M$.

  Then $x(t)$ *can be uniquely reconstructed* by its samples $x(nT)$, $n = -\infty, \dots \infty$, if T is *small enough* such that $\omega_s = \frac{2\pi}{T} > 2\omega_M$, by:

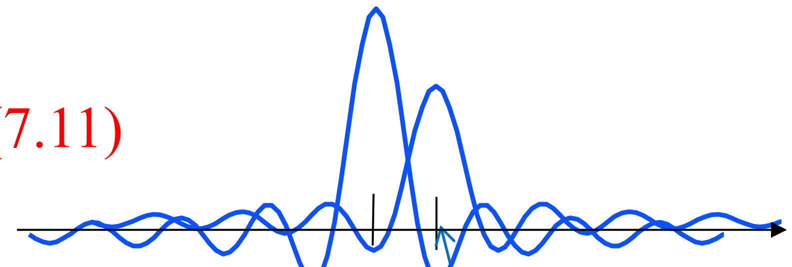
  bilateral bandwidth

Value of CT signal at any time t

Sampled values of $x(t)$ at instances nT.

A shifted sinc function centered at nT.

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT) \frac{T\sin(\omega_c(t - nT))}{\pi(t - nT)} \quad \dots .(7.11)$$



for any ω_c such that $\omega_M < \omega_c < \omega_s - \omega_M$

- $\omega_S = \frac{2\pi}{T}$ is the *sampling frequency* in angular frequency.

  $f_S = \frac{1}{T}$ is the *sampling frequency* in ordinary frequency

The Nyquist's Rate

- While we use angular frequency in mathematical expressions, mentally it is often simpler to think of sampling rate in ordinary frequency: $f_S = \frac{1}{T} = \frac{\omega_S}{2\pi}$;

  e.g., for digital telephony, f_S=8 kHz; for CD music, f_S=44.1 kHz

- Likewise, it is more convenient to talk about signal bandwidth in ordinary frequency $f_M = \frac{\omega_M}{2\pi}$.
- Nyquist theorem states that the sampling frequency needs to be greater than two times the signal's maximum frequency (=bilateral bandwidth):

$$f_S > 2f_M \text{ or } \omega_s > 2\omega_M$$

- This minimum sampling rate required is called the ***Nyquist rate*** or ***Nyquist frequency***.
  - Telephone network: maximum frequency in voice signal passed by subscriber loop ≅ 3.4 KHz ⟹ sampling rate standardized at 8 KHz.
  - CD music: maximum audio frequency supported ≅ 20 KHz ⟹ sampling rate = 44.1 KHz.

ELEC2100: Signals and Systems

Lecture 16

Chapter 7: Sampling Theorem and Aliasing

I. Introduction to Nyquist's Sampling Theorem

II. Proof of the Sampling Theorem

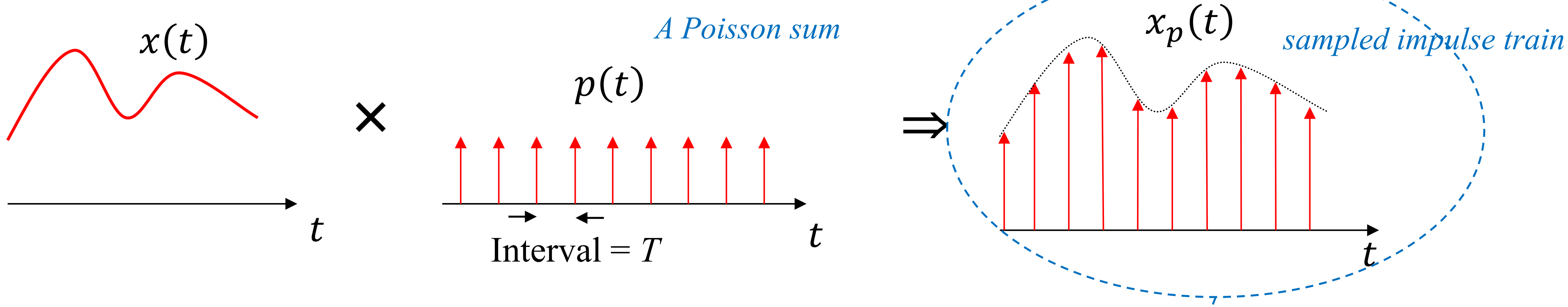
III. Sampling by Gating

IV. Aliasing

II. Proof of the Sampling Theorem

- Consider the periodic impulse train $p(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)$

A Poisson sum



- Multiplying a CT signal $x(t)$ by $p(t)$ produces a sequence of scaled impulses called a *sampled impulse train*:

$$x_p(t) = x(t)p(t) = \sum_{n=-\infty}^{\infty} x(t)\delta(t - nT) = \sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT)$$

$x(t)\delta(t - t_0) = x(t_0)\delta(t - t_0)$

- When multiplying $x(t)$ by $p(t)$, we are throwing away all values of $x(t)$ except the sample values at $t = nT$. This is the essence of sampling!

- Recall from the multiplication property that multiplication in time domain corresponds to convolution in frequency domain:

$$x_p(t) = x(t)p(t) \quad \leftrightarrow \quad X_p(j\omega) = \frac{1}{2\pi} X(j\omega) * P(j\omega)$$

$$= \frac{1}{2\pi} X(j\omega) * \sum_{k=-\infty}^{\infty} \frac{2\pi}{T} \delta(\omega - k\omega_s)$$

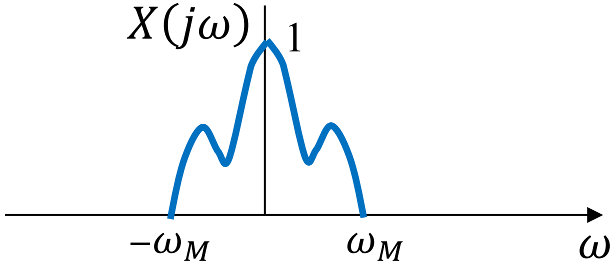
$X(j\omega) * \delta(\omega - k\omega_s) = X(j(\omega - k\omega_s))$

$$\Rightarrow X_p(j\omega) = \sum_{k=-\infty}^{\infty} \frac{1}{T} X(j(\omega - k\omega_s))$$

which is a *Poisson sum ,* or *periodic extension* of $X(j\omega)$ in frequency (with scaling by $1/T$)!

Are you able to draw $X_p(j\omega)$?

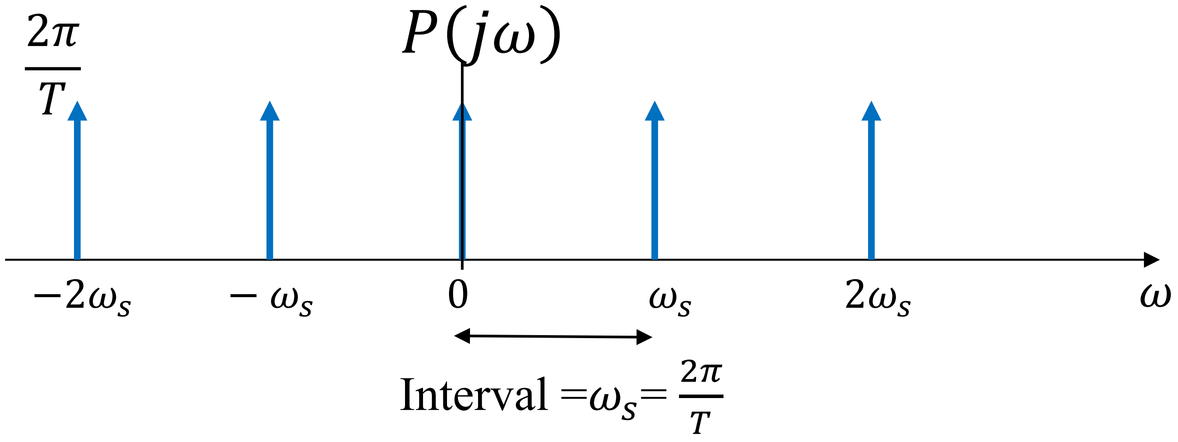
- Assume $x(t)$ is band-limited with FT as shown:

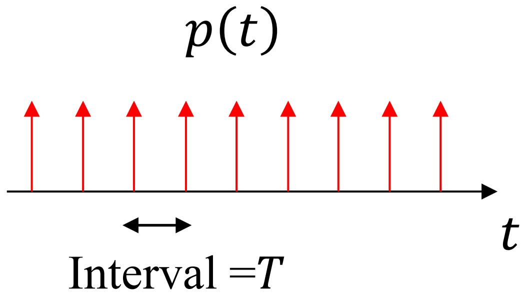


what does $X_p(j\omega)$, the FT of $x_p(t)$ look like? $\boxed{X_p(j\omega) = ??}$

- We recall that the FT of an impulse train in time domain is an also an impulse train in frequency domain, i.e.,

$$P(j\omega) = \Im\{p(t)\} = \frac{2\pi}{T}\sum_{k=-\infty}^{\infty}\delta(\omega - k\omega_s)$$



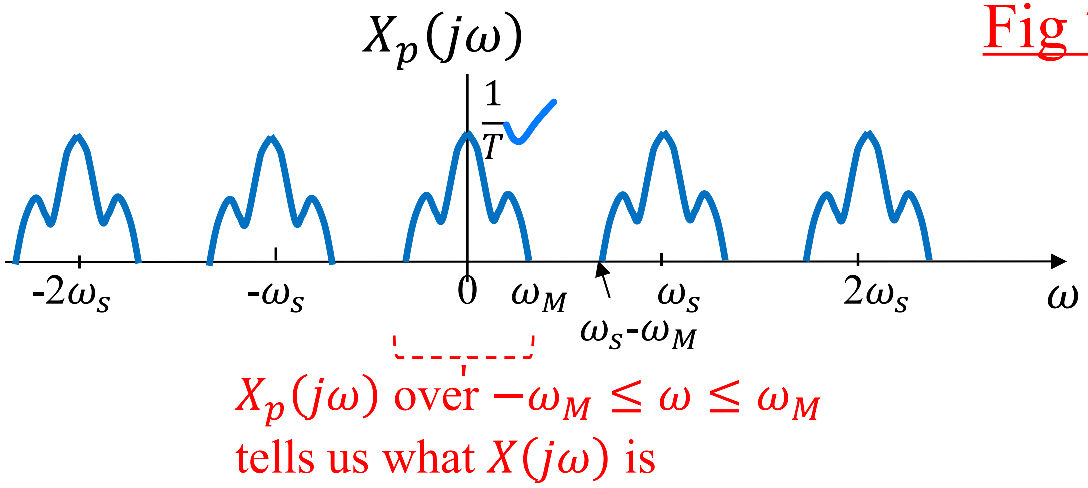


Shown below is $X_p(j\omega)$ under two situations:

1. When sampling frequency is high enough ($\omega_s > 2\omega_M$), there is no overlapping of individual pieces in the Poisson sum.

1. $\omega_s > 2\omega_M$

Fig 7.3

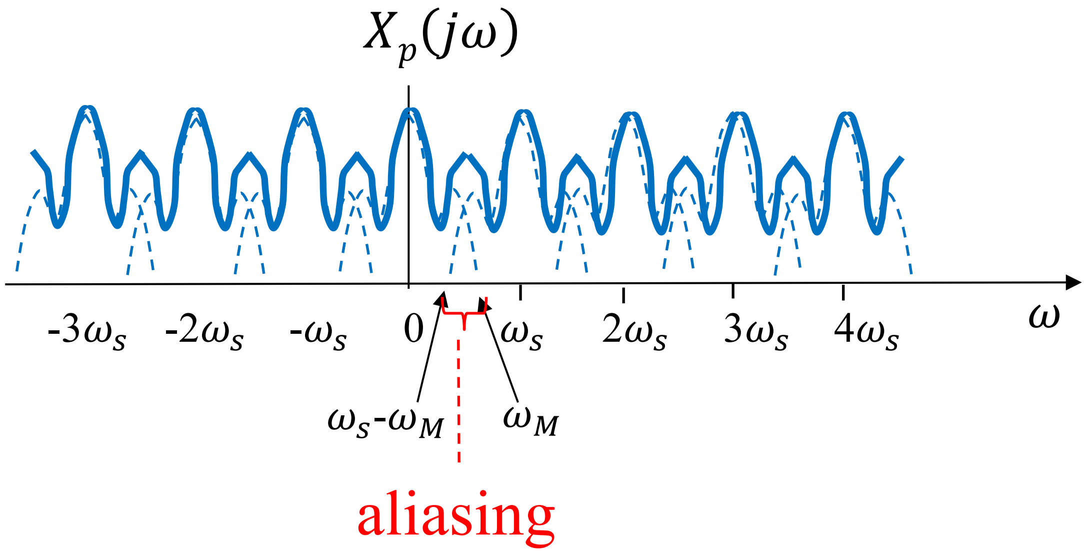


Hence $X(j\omega) = TX_p(j\omega)$ in the frequency range $-\omega_M \le \omega \le \omega_M$

and we can recover $X(j\omega)$ from $X_p(j\omega)$ by low-pass filtering.

2. When sampling frequency is too low ($\omega_s \leq 2\omega_M$), individual pieces in the Poisson sum overlap.

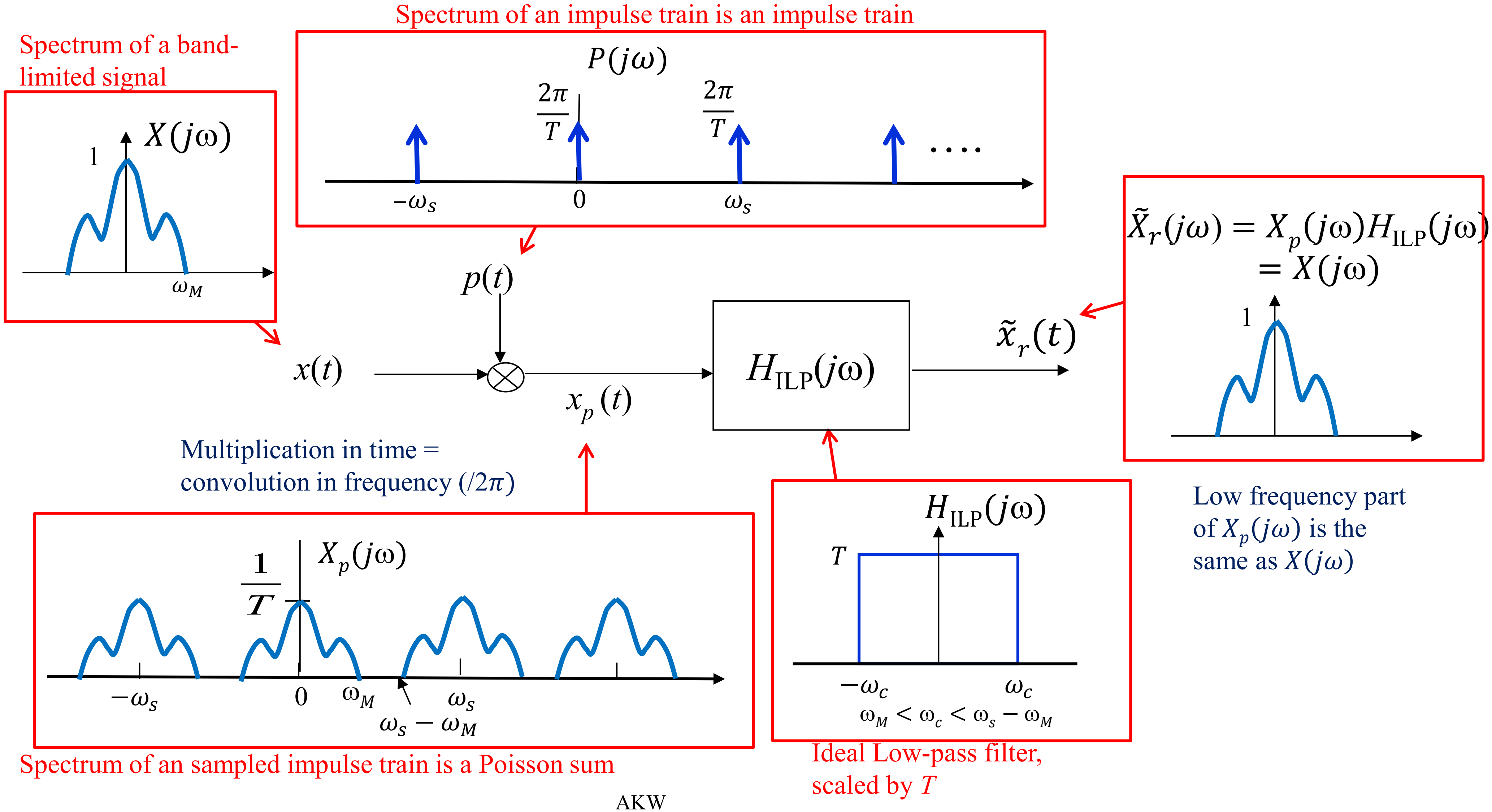
2. $\omega_s \leq 2\omega_M$



Where there is overlap, we cannot determine $X(j\omega)$ from $X_p(j\omega)$.

The effect of this overlapping is called ***aliasing***, it means the disguising of one frequency as another frequency.

- If $\omega_s > 2\omega_M$, we can recover $x(t)$ by passing $x_p(t)$ through an idea low-pass filter with cut-off frequency ω_c s.t. $\omega_M < \omega_c < \omega_s - \omega_M$ (and a scaling factor T), as shown below:



AKW

The Sampling Theorem – Final Explanation

- From table 4.2, the impulse response of an ILPF with cut-off frequency ω_c is

$$h_{lp}(t) = \frac{\sin \omega_c t}{\pi t}$$

- Therefore:

Low-pass filtering of $x_p(t)$ yields $x(t)$ if no aliasing

$$x(t) = T x_p(t) * h_{lp}(t) = T \underbrace{\left(\sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT) \right)}_{x_p(t)} * h_{lp}(t)$$

$$= T \sum_{n=-\infty}^{\infty} x(nT)\, h_{lp}(t - nT) = T \sum_{n=-\infty}^{\infty} x(nT) \frac{\sin \omega_c (t - nT)}{\pi (t - nT)}$$

which is eqn. (7.11).

$\delta(t - nT) * h_{lp}(t) = h_{lp}(t - nT)$

- Eqn. (7.11) represents the convolution of a sampled impulse train with the ILPF.

- You can also view eq. (7.11) as an *interpolation formula* that interpolates $x(t)$ from all of its sampled values.

ELEC2100: Signals and Systems

Lecture 16

Chapter 7: Sampling Theorem and Aliasing

I. Introduction to Nyquist's Sampling Theorem

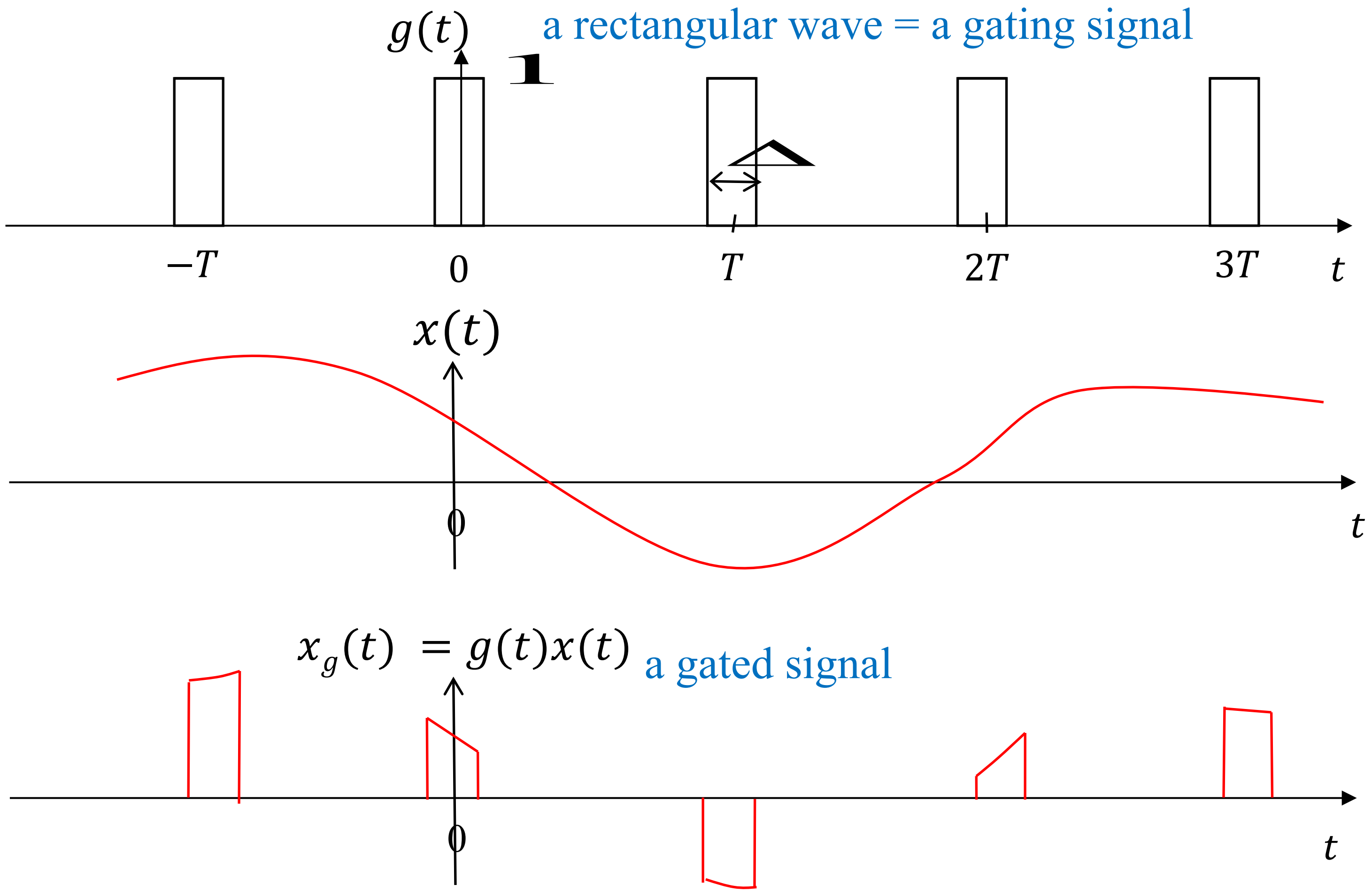
II. Proof of the Sampling Theorem

III. Example - Sampling by Gating

IV. Aliasing

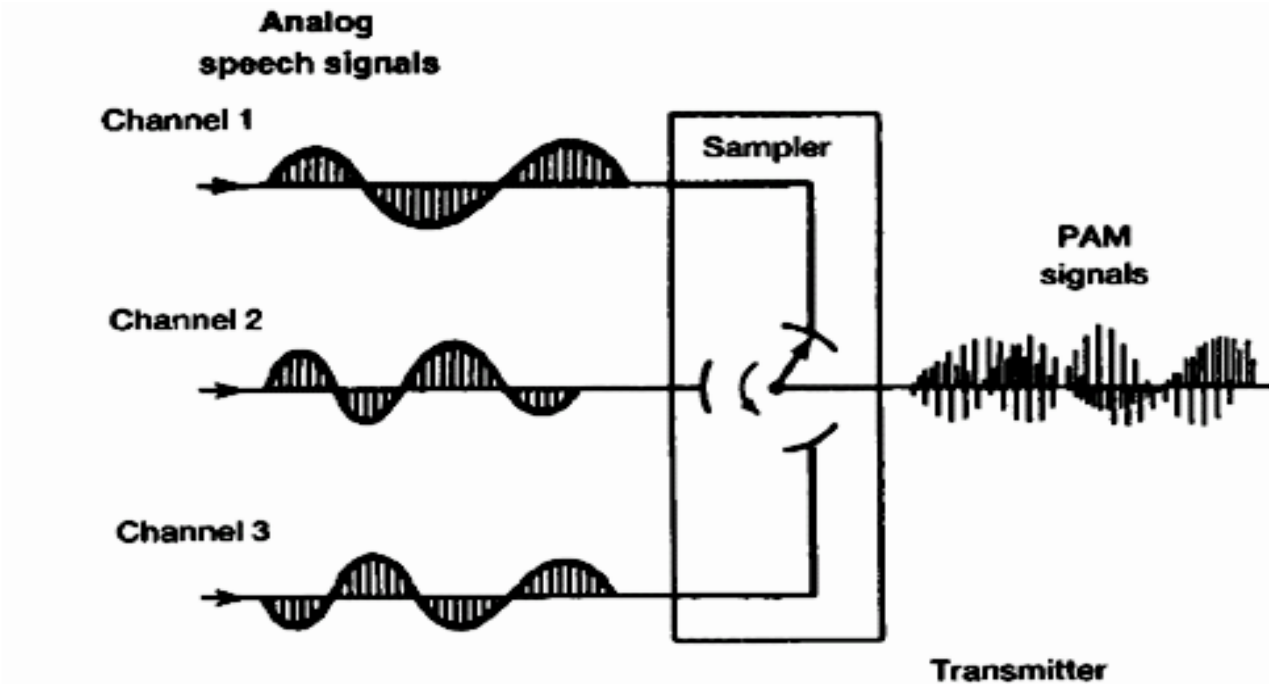
III. Example: Sampling By Gating

Instead of multiplying $x(t)$ by $p(t)$, the periodic impulse train, what if we simple multiply $x(t)$ by a gating function $g(t)$ which is a periodic rectangular wave with narrow pulses?



Early Time-Division Multiplexed System for Telephony

In fact, sampling by gating was done in early analog telephone systems. Here, several analog telephone signals are gated by a rotating contact switch.



In effect, each signal is multiplied by a time-shifted gating function $g(t)$, and are time-multiplexed and transmitted together.

How to show that we can recover $x(t)$ from $x_g(t)$?

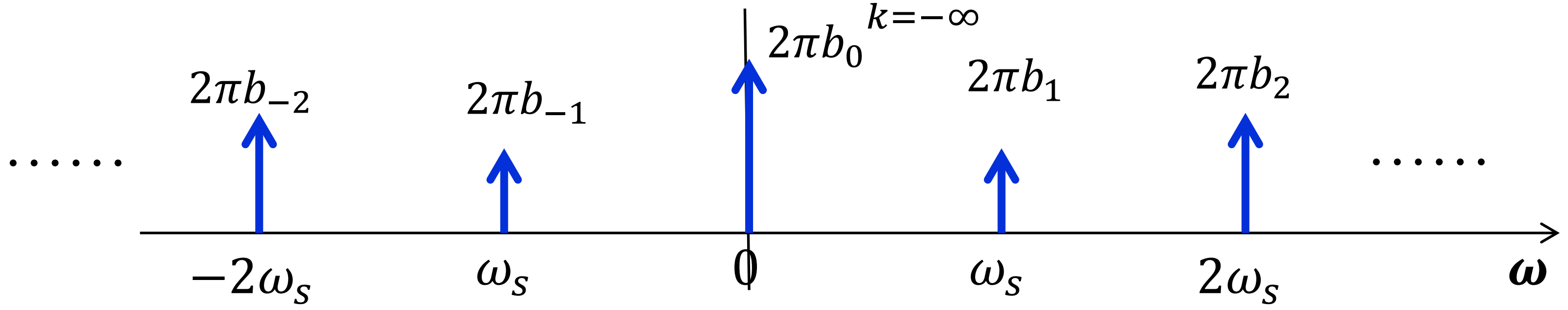
Recovering from the gated signal

First, what is the FT of $g(t)$?

$g(t)$ is T-periodic, so it has a FS: $g(t) = \sum_{k=-\infty}^{\infty} b_k e^{jk\omega_s t}$

This means the FT of $g(t)$ is made up of impulses at the harmonic frequencies:

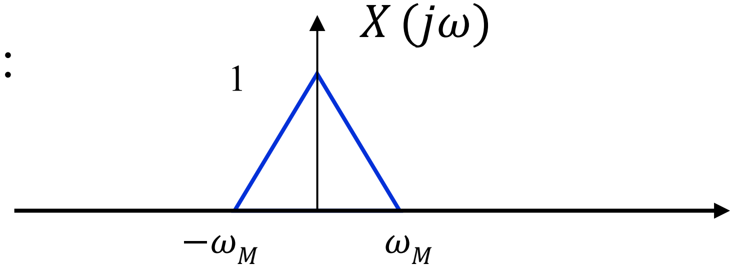
$$G(j\omega) = \Im\{g(t)\} = 2\pi \sum_{k=-\infty}^{\infty} b_k \delta(\omega - k\omega_s)$$



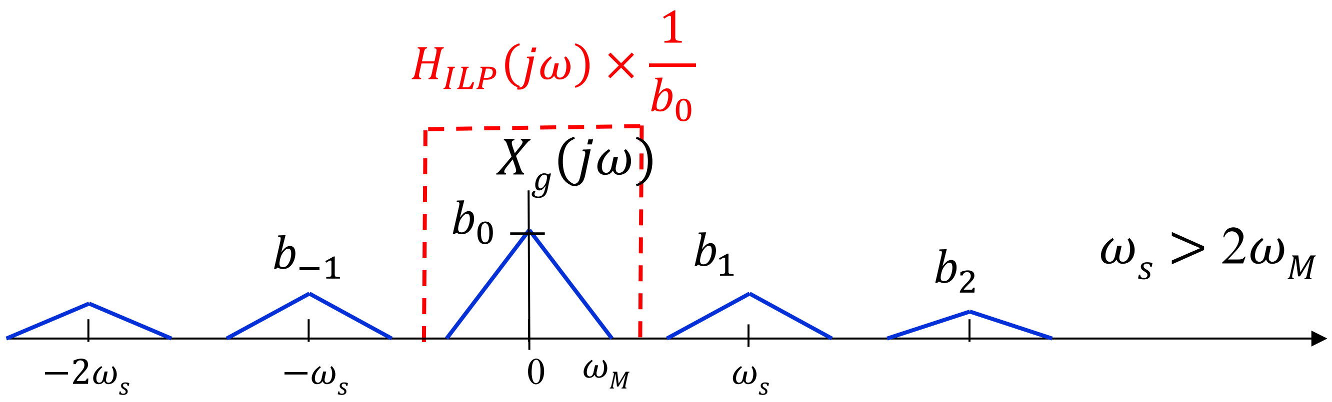
We multiply $x(t)$ by $g(t)$, so we convolve in frequency domain:

$$X_g(j\omega) = \Im\{x_g(t)\} = \frac{1}{2\pi} X(j\omega) * G(j\omega) = \sum_{k=-\infty}^{\infty} b_k X(j(\omega - k\omega_s))$$

So, if $x(t)$ is a bandlimited signal as shown:



$X_g(j\omega)$ is the sum of many shifted and weighted copies of $X(j\omega)$:

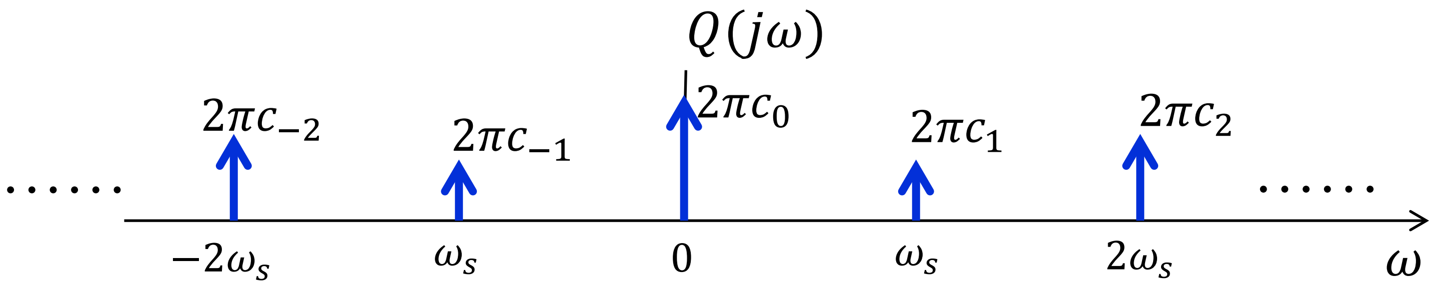


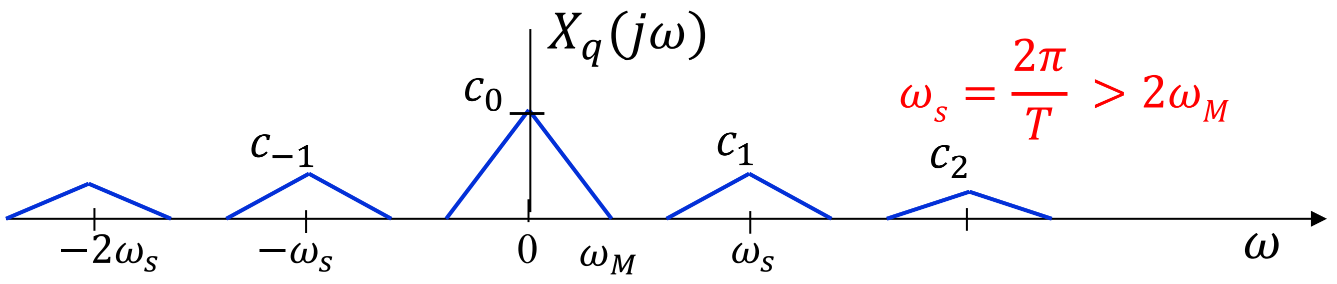
Therefore, as long as $\omega_s > 2\omega_M$, we can recover $X(j\omega)$ from $X_g(j\omega)$ by applying a low-pass filter (with scaling if so desired)

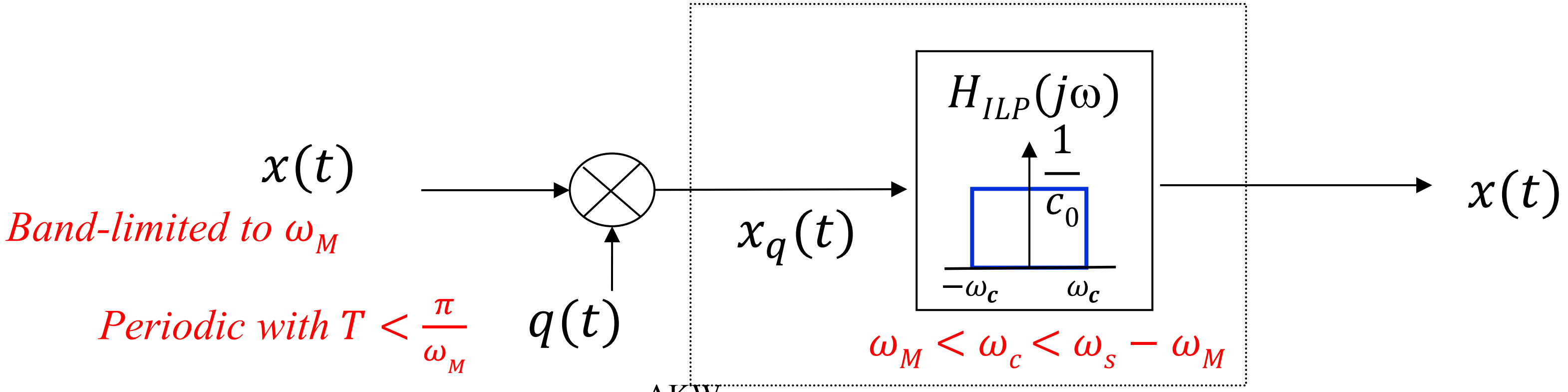
Multiplying with any Periodic Signal

In fact, the story remains the same if we multiply $x(t)$ with any T-periodic signal $q(t)$.

We can recover $x(t)$ from $x_q(t) = x(t)q(t)$, as long as $T < \frac{\pi}{\omega_M}$.







AKW

ELEC2100: Signals and Systems

Lecture 16

Chapter 7: Sampling Theorem and Aliasing

I. Introduction to Nyquist's Sampling Theorem

II. Proof of the Sampling Theorem

III. Example - Sampling by Gating

IV. Aliasing

IV. Aliasing

- Recall that a DT complex sinusoid is unchanged if we change its frequency by $m2\pi$:

$$e^{j(\Omega+m2\pi)n} \equiv e^{j\Omega n}$$

Here we use Ω to represent frequency of DT signals because we will also be talking about the frequency of CT signals at the same time

- Another key concept is that if we sample at frequency $\omega_s = 2\pi f_s = \frac{2\pi}{T}$, two CT complex sinusoids with a frequency difference of $m\omega_s$ will produce exactly the same set of samples, which means:

while $e^{j(\omega+m\omega_s)t} \neq e^{j\omega t}$

$$e^{j(\omega+m\omega_s)nT} \equiv e^{j\omega nT}$$

This is because $e^{j(\omega+m\omega_s)nT} = e^{j(\omega nT+\cancel{mn2\pi})} = e^{j\omega nT}$

since $m\omega_s nT = m\frac{2\pi}{T}nT = mn2\pi$

- In terms of ordinary frequency, the above means that complex sinusoids at ordinary frequency f and $f + mf_s$ will produce the same set of samples.

m additional revolutions around the circle between two sampling instances

Propeller of an Aircraft by Video Camera

propeller of a Bombardier Q400 airplane appearing stationary or rotating at a low speed:

Sampling of a real sinusoid

- In many physical problems, we are concerned with the real sinusoid (e.g., our ears hear real sinusoids).
- A real sinusoid is a conjugate pair of complex sinusoids, and therefore we can say even more.

Consider the cosine:

Like the complex sinusoid, if sampled at ω_s, increasing frequency by $m\omega_s$ does not change any sample value:

$+\cancel{mn2\pi}$

$$\cos((m\omega_s + \omega)nT) = \cos\left(\left(m\frac{2\pi}{T} + \omega\right)nT\right) = \cos(\omega nT)$$

$\cos(-\theta) = \cos(\theta)$

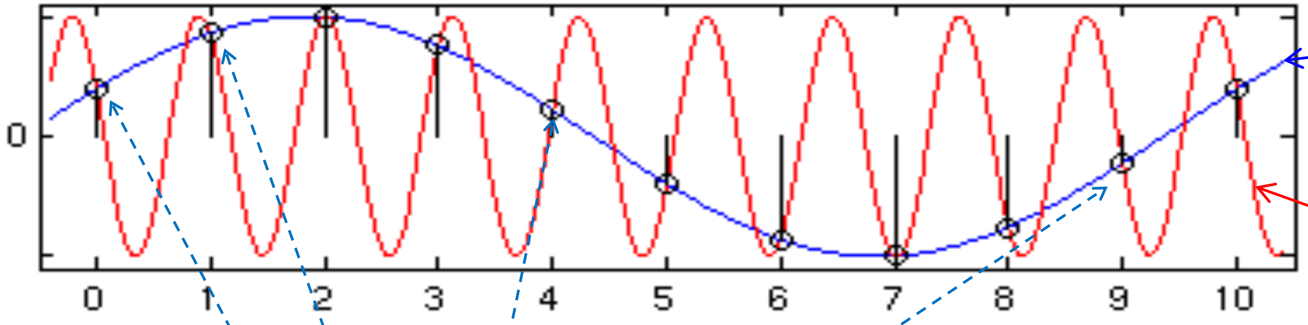
But cosine is even; meaining: $\cos(\omega nT) = \cos(-\omega nT) = \cos\big((m\omega_s - \omega)nT\big) = \cos\big((m\omega_s + \omega)nT\big)$

Therefore $\Rightarrow \cos(m\omega_s \pm \omega)nT = \cos(\omega nT)$

This means if we sample at 1Hz, real sinusoids at 1.1 Hz, 1.9 Hz, 2.1 Hz, 2.9 Hz, -0.1 Hz, -0.9Hz …, etc., will all become the same as 0.1 Hz!

Sampling of a real sinusoid

The diagram below illustrates that two sinusoids, one at frequency 0.1 Hz and the other at frequency 0.9 Hz, give identical samples when sampled at T=1 (f_s = 1):



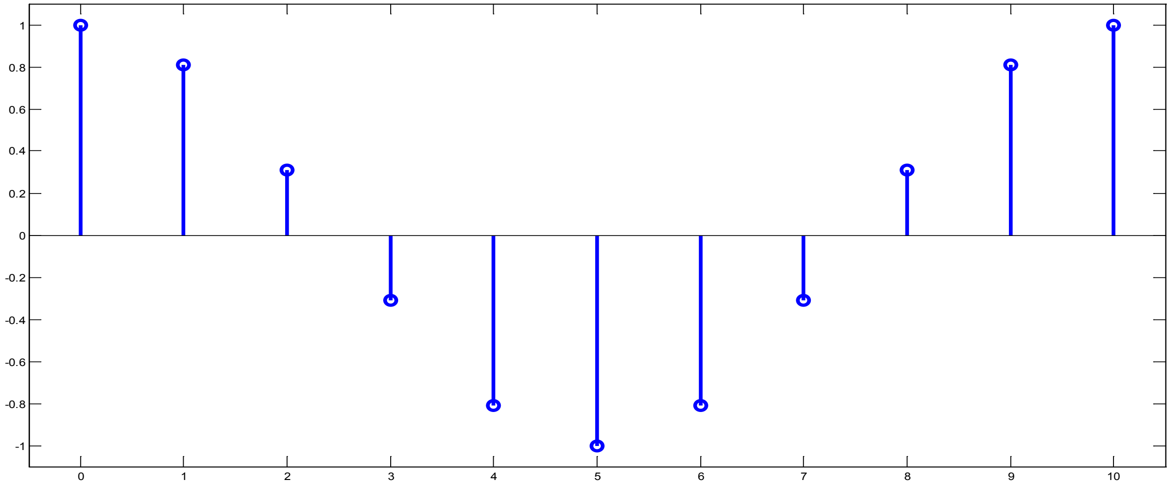
$$\cos\left(2\pi\frac{1}{10}n\right)$$

$$\cos\left(2\pi\frac{9}{10}n\right)$$

All sampled values are identical!

Aliasing for real sinusoid

If I just show you the sampled DT sequence from the previous slide, what frequency will you perceive?

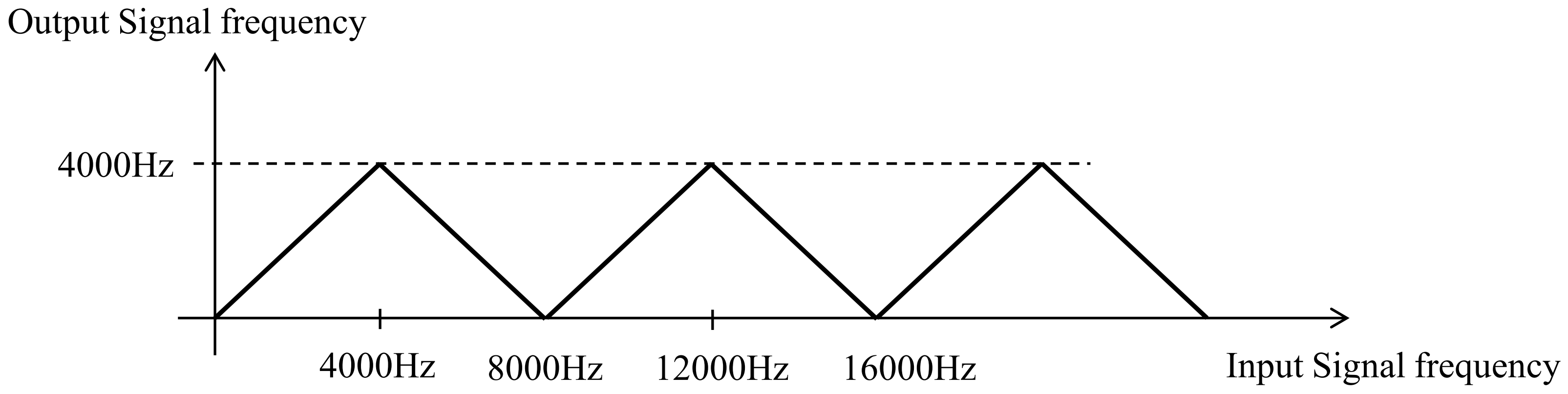


- We perceive 0.1 Hz.
- But CT sinusoids at 0.9 Hz, 1.1 Hz, 1.9 Hz, 2.1 Hz, etc., will all produce the DT sequence above and be perceived as 0.1 Hz!
- This is ***aliasing***, the *disguising of high frequency signals as low frequency*.

Demo – Sampling for Digital Telephony

- We mentioned that the modern digital telephone network samples voice signal at 8,000 Hz. Input frequencies above 4,000 Hz will alias into lower frequencies. For example, 5000 Hz and 11000 Hz will appear as 3000 Hz, 7000 Hz and 9000 Hz will appear as 1000 Hz, etc.
- We can plot the perceived output frequency against the input CT signal frequency. As we sample at 8 KHz, the highest perceived frequency possible is 4 KHz.

In the telephone network, we sample at 8 KHz. That means input signals at 1 KHz, 7 KHz, 9 KHz, 15 KHz, 17 KHz, 23 KHz, etc., will all produce identical samples and



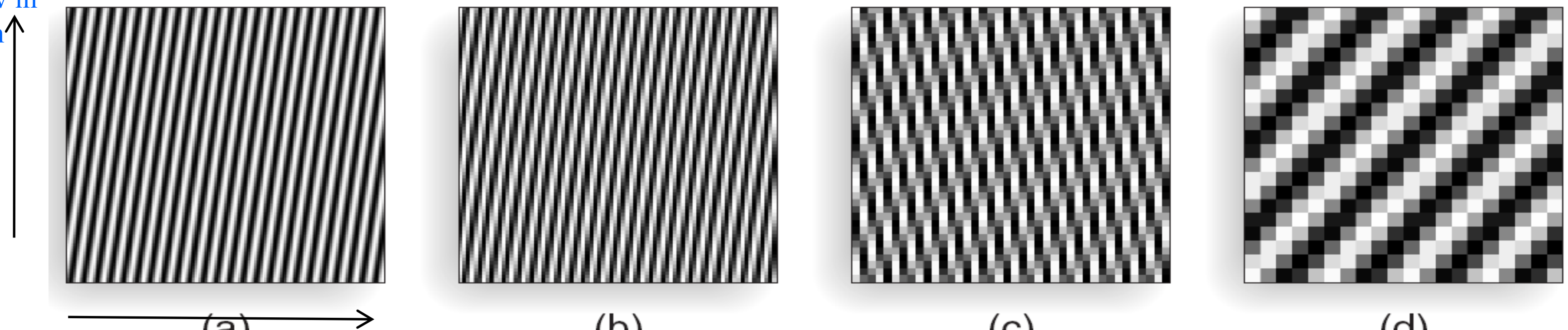
Aliasing in Images

The effect of aliasing is common in images as well.

- Picture below contains slanted finely-spaced strips, with high (spatial) frequency horizontally (24 cycles along the edge) but low frequency vertically (4 cycles along the edge).
- As we sample the image by reducing the resolution, aliasing occurs in the horizontal direction first when the sampling frequency drops below 48, changing the perceived orientation of the strips.

signal freq is low in vertical direction ($f_{vertical} = 4$)

$f_s = 76$ $f_s = 32$ $f_s = 20$



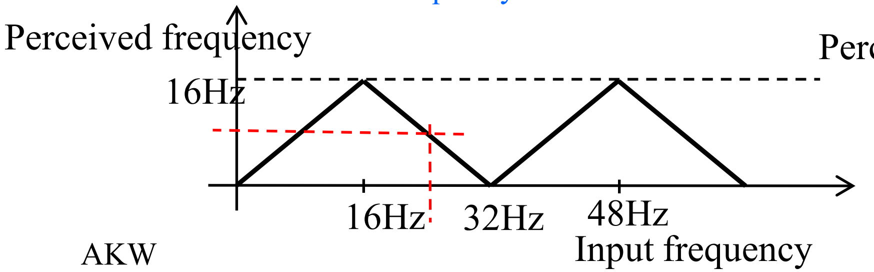
(a) (b) (c) (d)

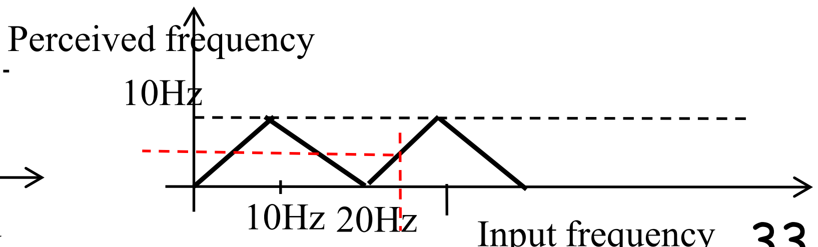
signal freq is high in horizontal direction (24 black and white cycles per unit length; $f_{horizontal} = 24$)

Sampling freq is greater than 2 times horizontal or vertical frequency; no aliasing

Sampling freq reduced to 32, so perceived horizontal frequency becomes 8

Sampling freq reduced to 20, and perceived horizontal frequency becomes 4





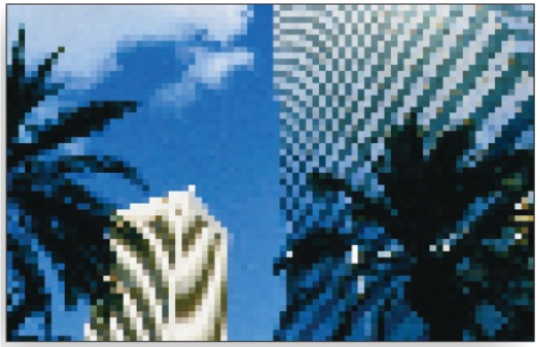
AKW

Aliasing in Images - Moire Effect

Appearance of odd patterns in images because of aliasing is called the Moire effect.

Picture is a signal in space (spatial)

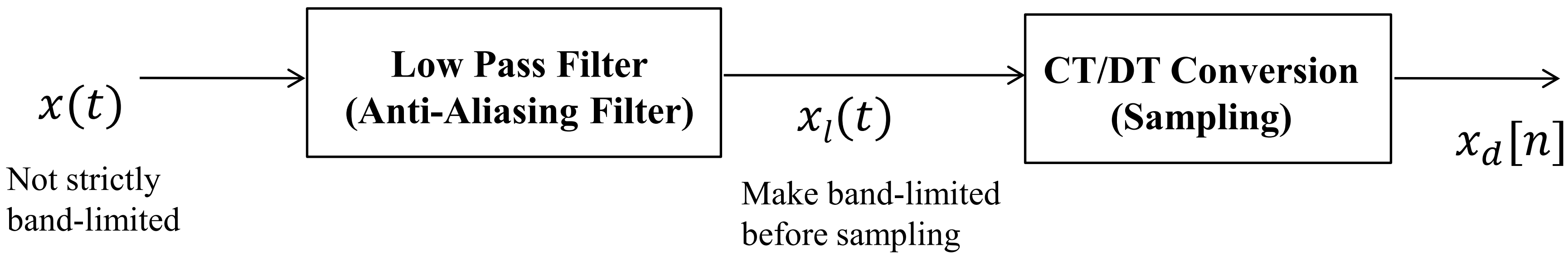
Reduced resolution leads to aliasing causing errors in “spatial” frequencies



Anti-Aliasing Filter

In many CT to DT conversion systems, we want to keep the sampling rate low, but the input CT signal may not be strictly band-limited to ½ of the sampling rate.

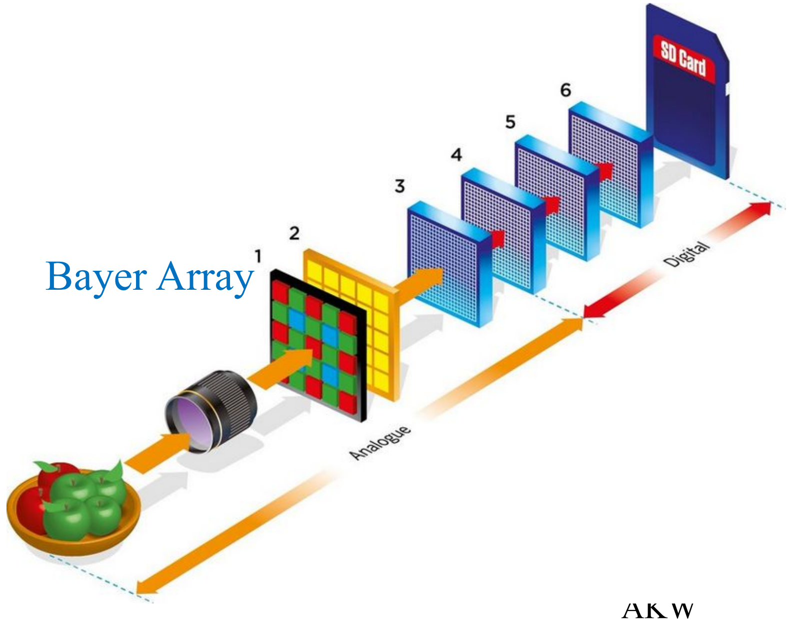
- A ***anti-aliasing filter*** (**AAF**), which is simply a low-pass filter, is often used to remove the high frequency components in the CT signal before sampling.



- With anti-aliasing filter, the error is *only* in the high frequency components lost.
- Without anti-aliasing filter, *the high frequency components are not only lost, they turn into low frequency signals that should not be there*!

Anti-Aliasing Filter Examples

- The digital telephone network samples voice signals at 8 KHz. Therefore, filter is applied to first remove voice signal components above 4 KHz before sampling.
- When we create the thumbnail of a picture, we must first low pass filter it to avoid the Moire effect.
- Behind the lens of your digital camera is a filter that blurs the optical signal before spatial sampling by the CMOS sensor.



???