Sampling & Quantization

by Erol Seke

For the course "**Communications**"

ESKİŞEHİR OSMANGAZİ UNIVERSITY

General Communication System

DIGITAL Stages

RF Stages

DAC

ADC

hamming RS
convolutional
viterbi parity
interleaver ...

ISI matched-filter
waveshaping
raised-cosine ...

PSK ASK FSK
QPSK AM FM
VSB DSB-SC ...

DSSS FHSS
THSS OFDM
PN-codes ...

CDMA
FDMA
CSMA
aloha ...

PDMA SDMA
mixer
upconverter
antenna ...

1110011001

1110011001011010110100...

data
compression

ciphering

Other
sources

info → | Formatting | → | Source Encoding | → | Encryption | → | Channel Encoder | → Mux → | Pulse Modulation | → | Bandpass Modulation | → | Frequency Spreading | → | Multiple Access | → | Transmitter |

sampling &
quantization

PGP triple-DES
public-key ...

delta DPCM
Huffman VQ
entropy-coding
jpeg mpeg RLE ...

interleaver
sync

decoder
sync

frame
sync

bit/symbol
sync

carrier
sync

pn-code/
freq. hop
sync

carrier

carrier

error
detection/correction

noise

CHANNEL

AWGN BER
noise-figure
link-budget

info ← | Formatting | ← | Source Decoding | ← | Decryption | ← | Channel Decoder | ← De Mux ← | Pulse Demod. | ← | Bandpass Demod. | ← | Frequency Despreading | ← | Multiple Access | ← | Receiver |

data
decompression

deciphering

DAC

ADC

Other
sinks

polarization
multipath
fading ...

Problem : We have a signal at the hand which we would like to replicate at another location, but we are not allowed to transmit it.

Is it possible? : Someone looks at the signal and tells about the shape of it to someone at the other end and tries to make him construct the signal. ☺
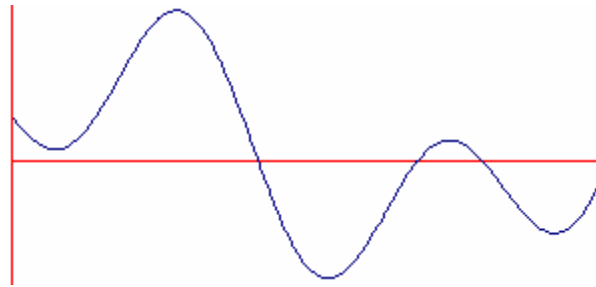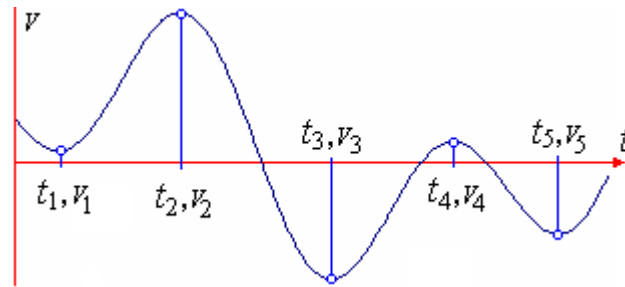
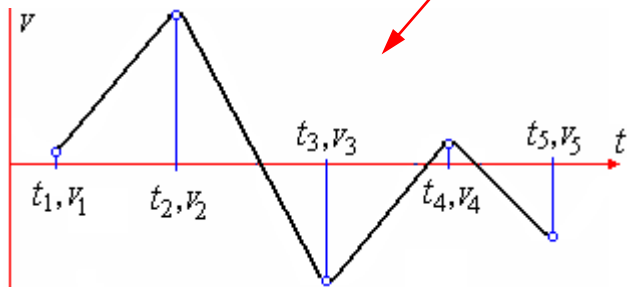Example : Tell some values at some critical points. How about min-max points?
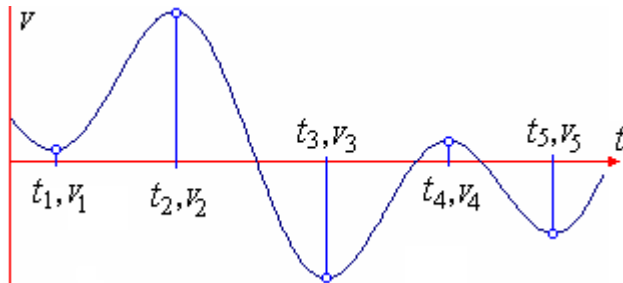
…and transmit time and value
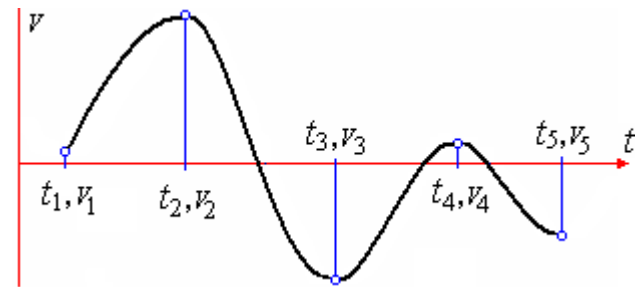
$$(t_1,V_1)(t_2,V_2)(t_3,V_3)(t_4,V_4)(t_5,V_5)\ldots$$

…and reconstruct the signal from its irregular samples.

reconstruct from samples

linear interpolation

Higher order curve fit techniques

Any difficulty ? ☹

How about taking regular samples in which sampling time need not be measured and transmitted everytime a sample is taken?
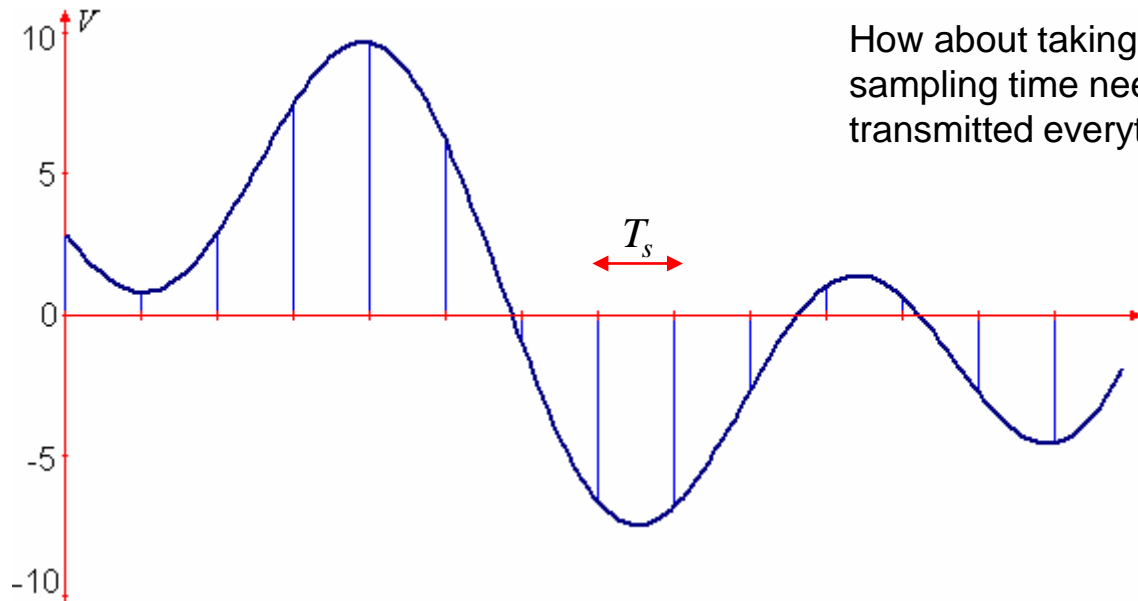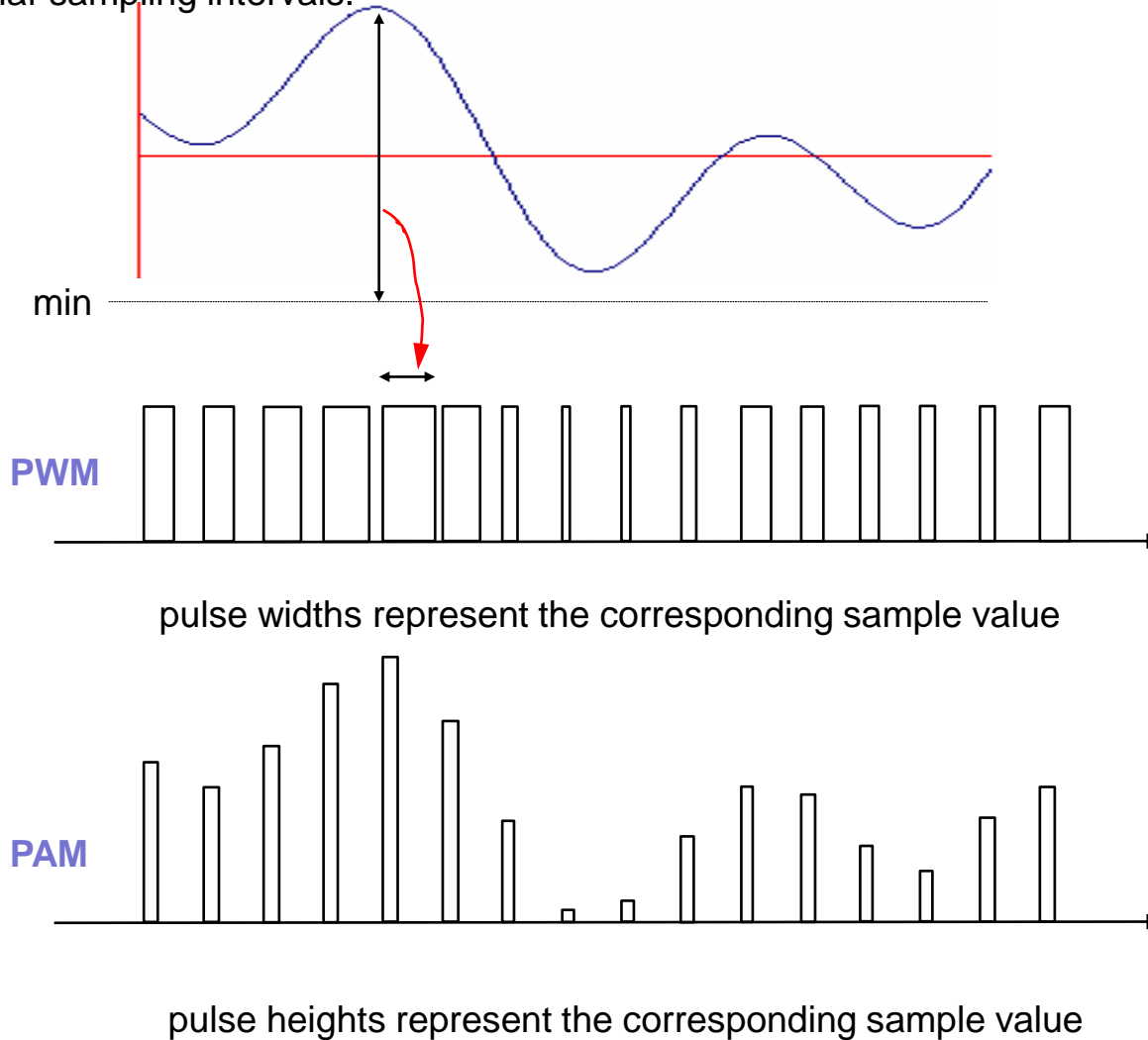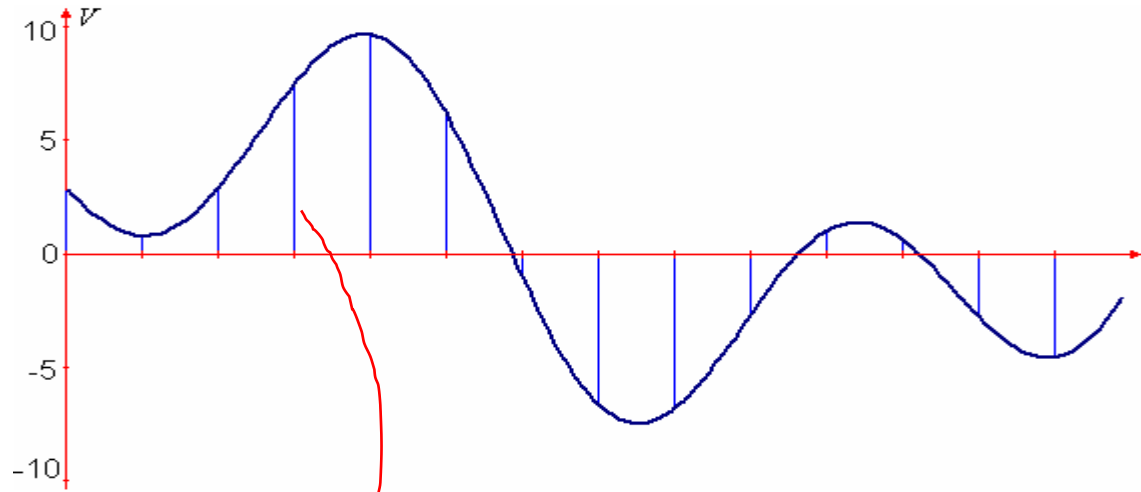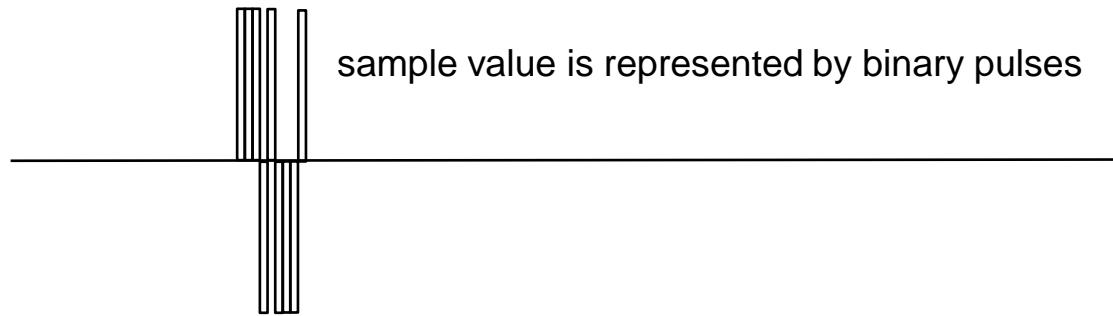
Q: How large the intervals can be?
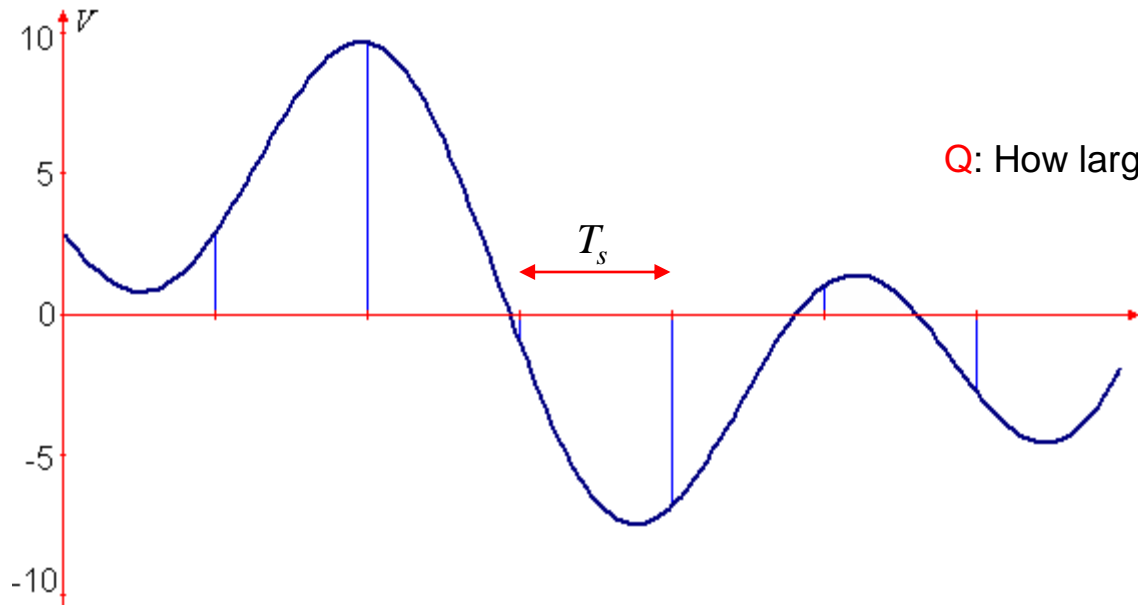
Pulsed Representations

Let us assume that we have an analog baseband signal that we want to represent using pulses with regular sampling intervals.

PWM

min

pulse widths represent the corresponding sample value

PAM

pulse heights represent the corresponding sample value

PCM

sample value is represented by binary pulses

Most known binary pulse code system is the "straight binary"

Nyquist's sampling criterion : A baseband signal with the highest frequency component at f_m can be reconstructed from its samples if the interval between samples is less than $1/2f_m$

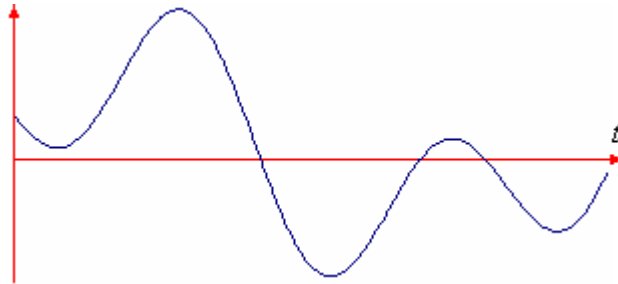That is; sampling frequency must be higher than twice the highest frequency of the signal

$x[n]$
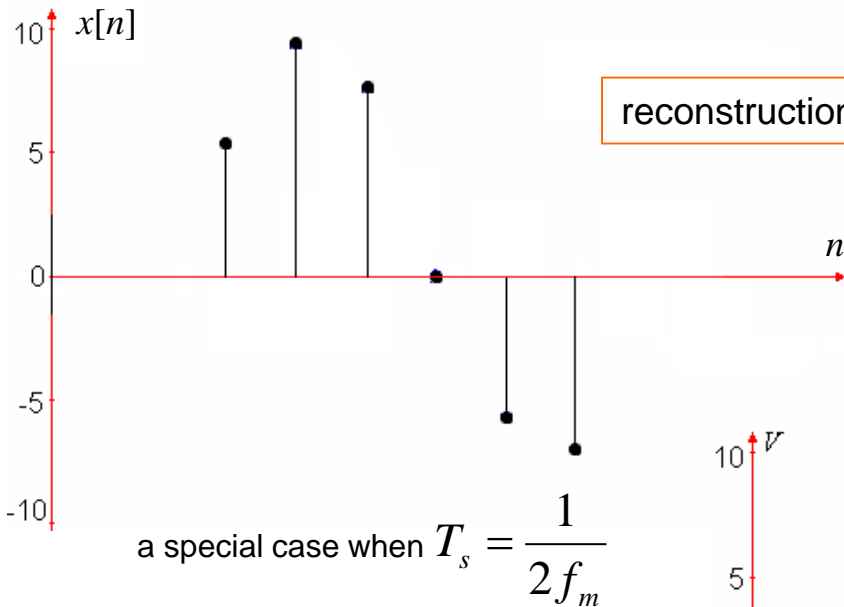
$$\boxed{\text{reconstruction formula}} \quad x(t) = \sum_{n=-\infty}^{n=\infty} x(nT_s)\operatorname{sinc}(2f_m(t-nT_s))$$

The sum of all **sinc**s

a special case when $T_s = \dfrac{1}{2f_m}$
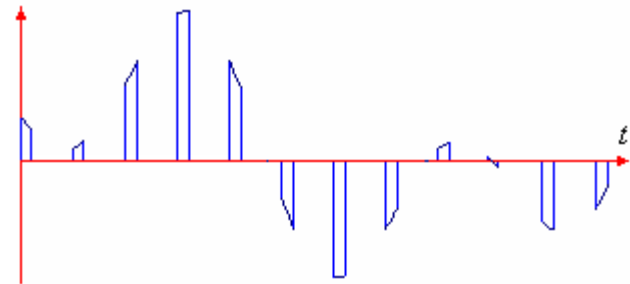
truncated sinc function
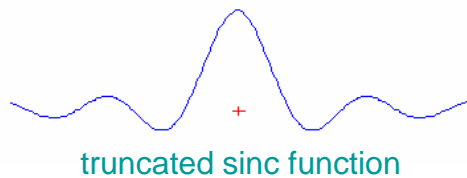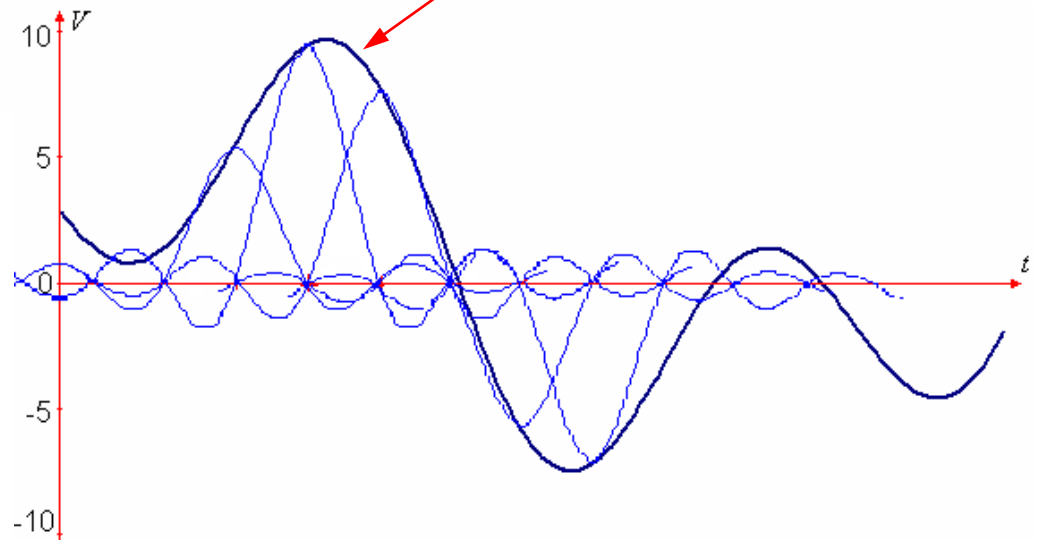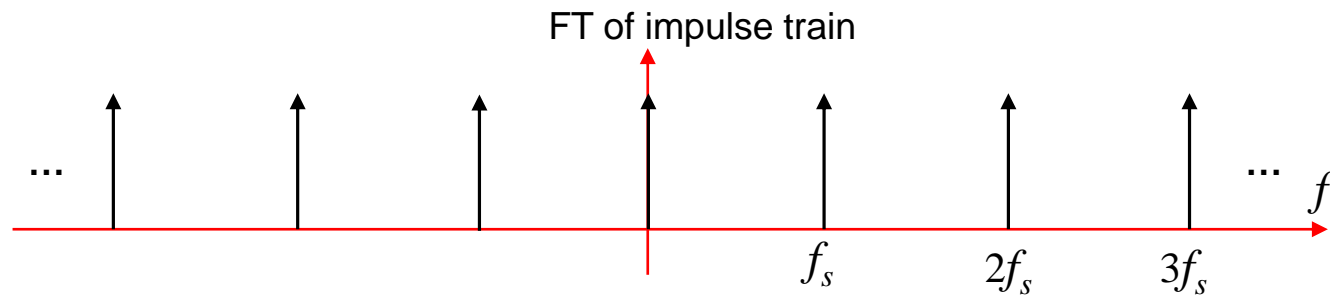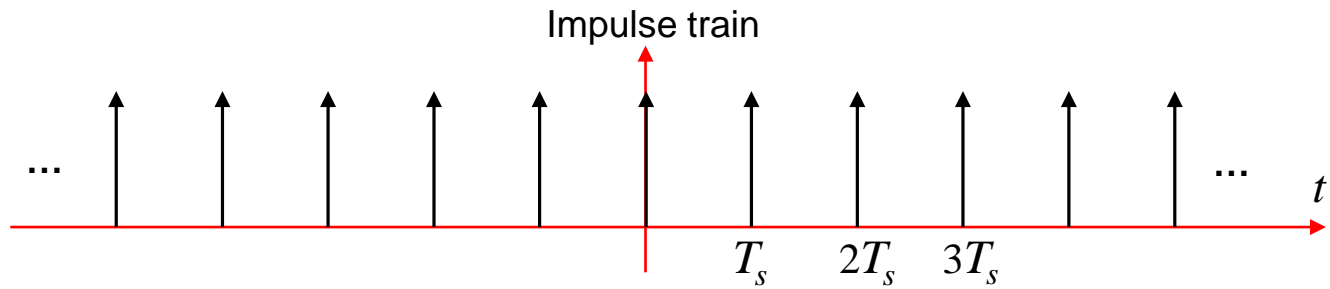
Original signal

Sampled signal

Sampling signal
(impulse train)

the signal values when the measurement
is initiated and ended may be different
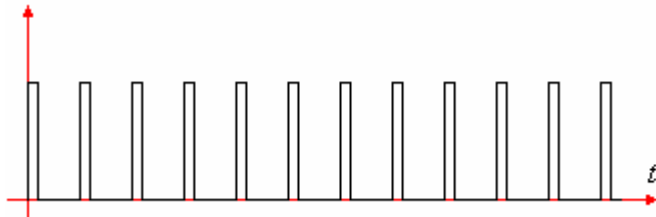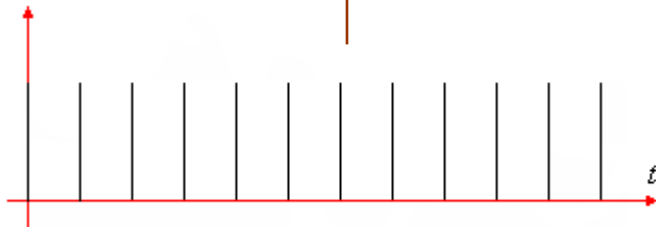
more realistic sampling signal

Impulse train

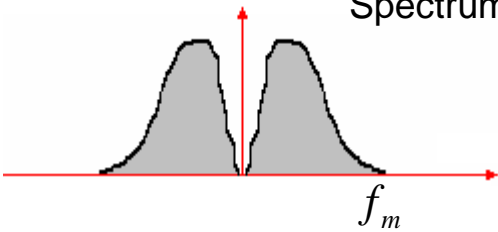... T_s $2T_s$ $3T_s$ t

FT of impulse train
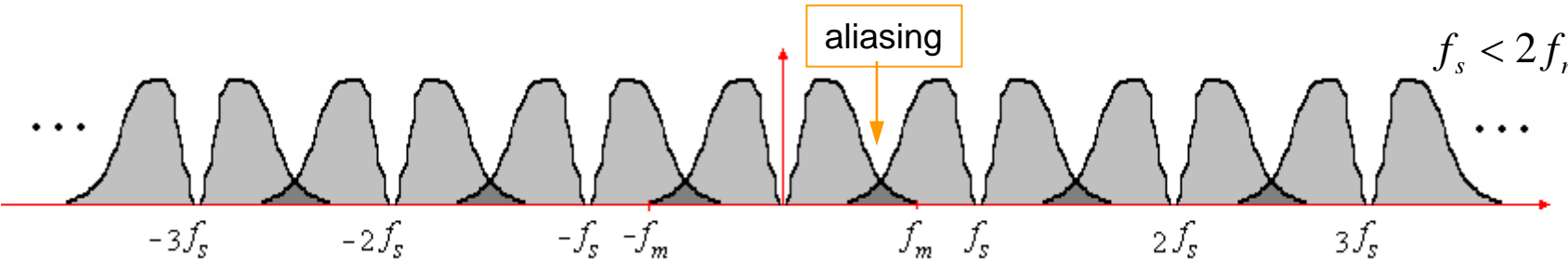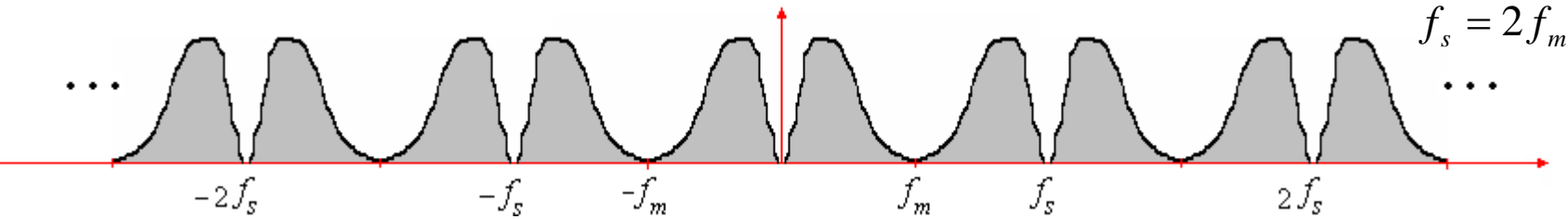
... f_s $2f_s$ $3f_s$ f

$f_s = 1/T_s$
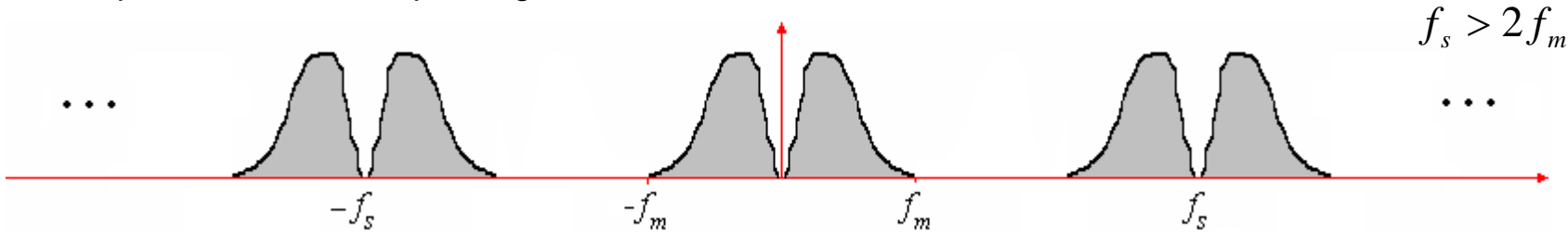
inversely proportional : remember $\mathcal{F}\{x(at)\} = \dfrac{1}{|a|} X\left(\dfrac{\omega}{a}\right)$

Spectrum of the baseband signal

f_m

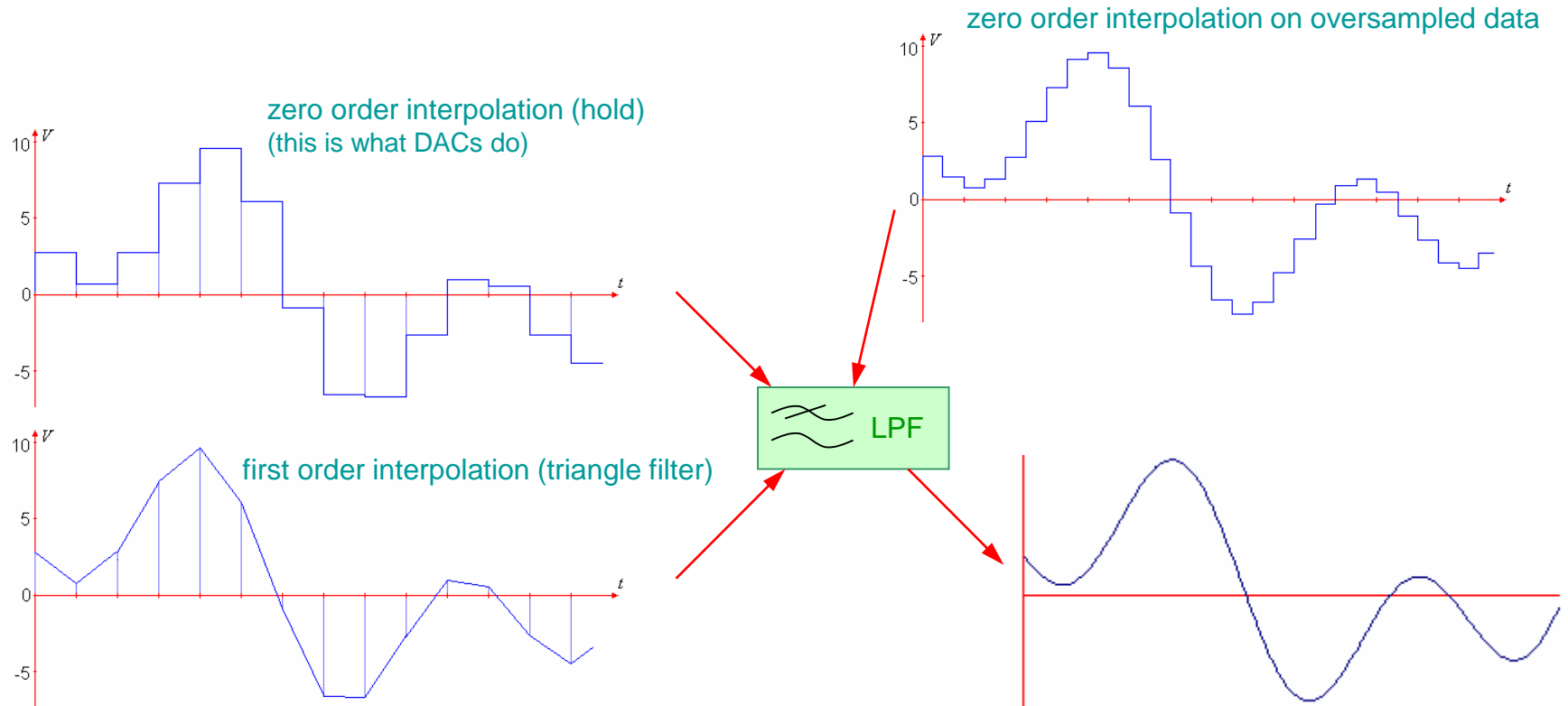Spectrum of the sampled signal

$f_s > 2f_m$

$-f_s$ $-f_m$ f_m f_s

$f_s = 2f_m$

$-2f_s$ $-f_s$ $-f_m$ f_m f_s $2f_s$

aliasing

$f_s < 2f_m$

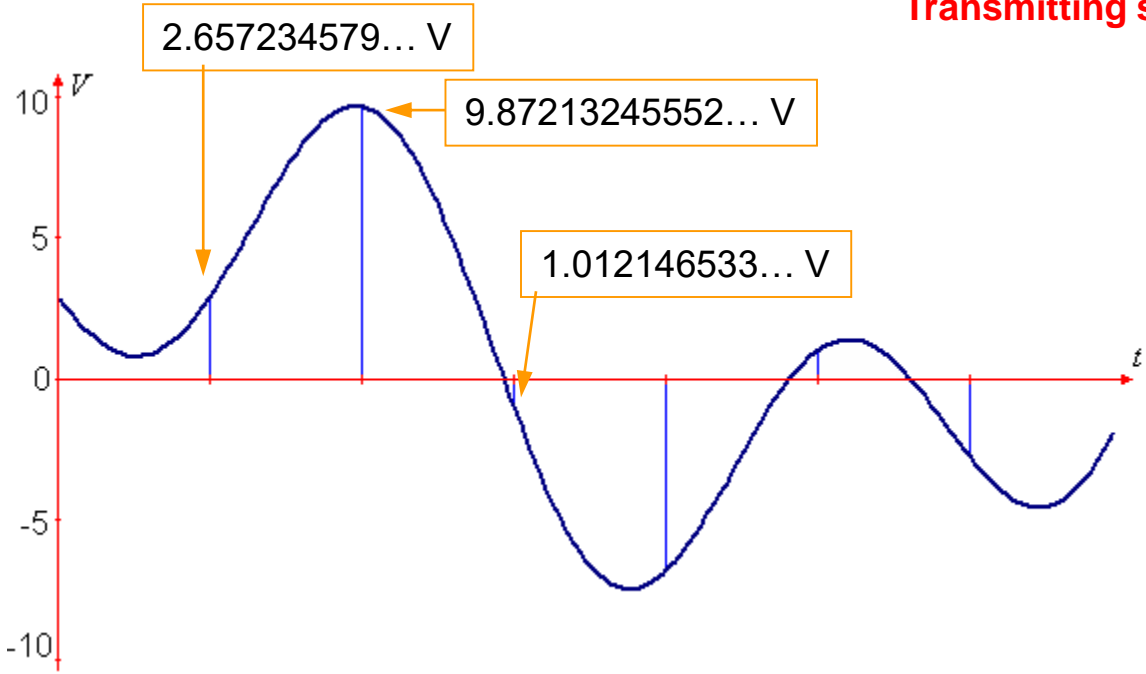$-3f_s$ $-2f_s$ $-f_s$ $-f_m$ f_m f_s $2f_s$ $3f_s$

Ways to reconstruct the baseband signal from its samples

The actual problem is to "fill in the blanks" between the sample values and make it look like the original signal as good as possible. The ideal way is to use an ideal LPF which passes no components beyond f_m and disturbs none of the lower components. Ideal filter (Π), however, has infinite time duration (sinc). Truncated sinc is still impractical in most cases.

Filling the gaps some other way and passing the signal through an ordinary LPF is practical but that requires higher sampling rates than Nyquist rate (oversampling).

zero order interpolation on oversampled data

zero order interpolation (hold)
(this is what DACs do)

first order interpolation (triangle filter)

LPF

2.657234579… V

9.87213245552… V

1.012146533… V

Transmit the sequence of the sample values as 2.657234579… V   9.87213245552… V   …?

Is it getting costlier than transmitting the original signal ?

It looks like we have to give up some precision… and possibly transmit 2.6… V   9.8… V   etc.

The truncation done is called **quantization**. (Nyquist's rule is **not** involved here)

Let us divide (-13.5,+10.5) range into 3 V sub-ranges and map all values in a sub-range to a single value

…and instead of transmitting 1.71323426148… V let us transmit 3 as shown
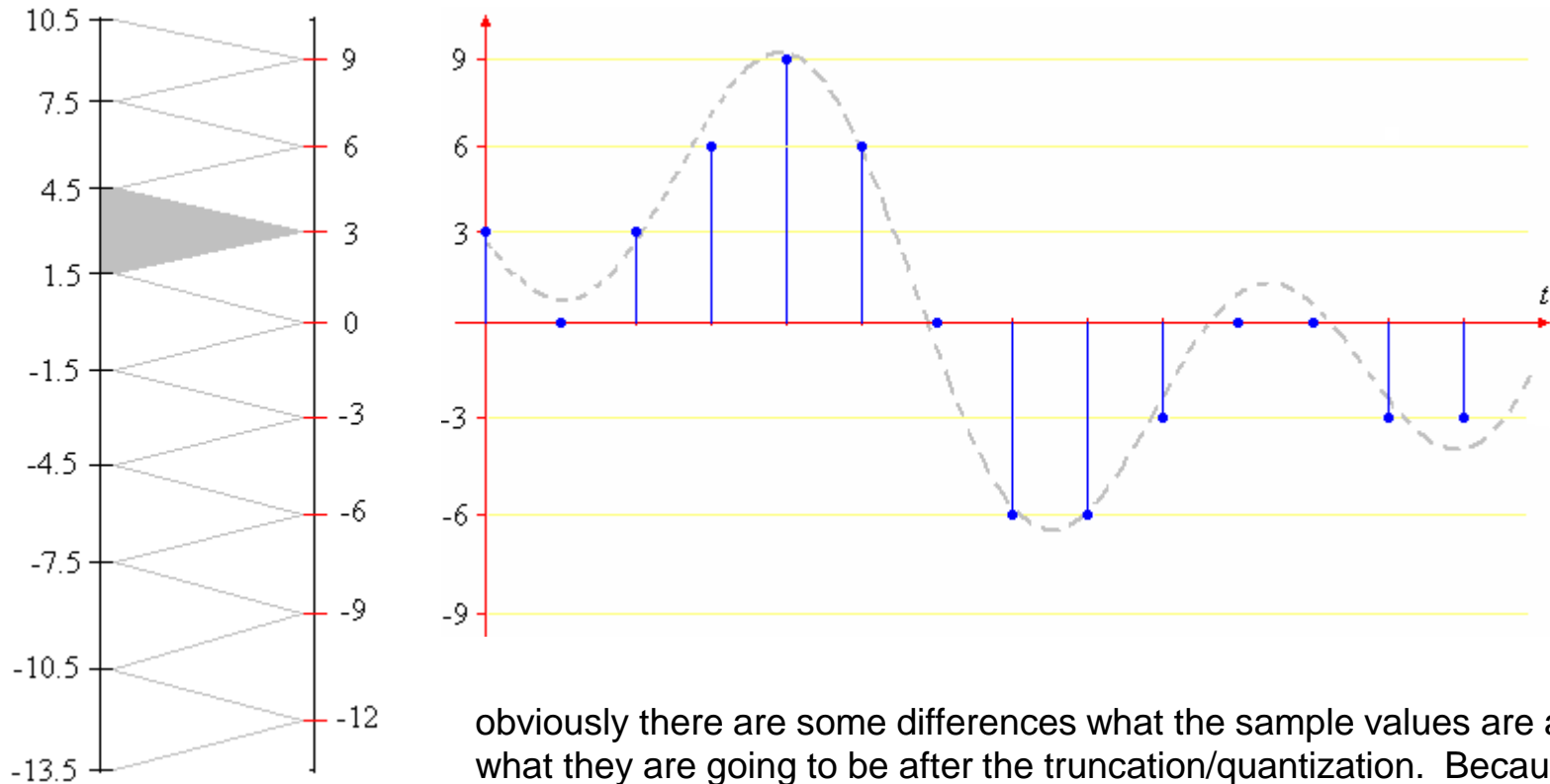
obviously there are some differences what the sample values are and what they are going to be after the truncation/quantization. Because of this differences, the original value can never be recovered.

The transmitted sequence would be 3, 0, 3, 6, 9, 6, 0, -6, -6, -3, 0, 0, -3, -3 for the above example
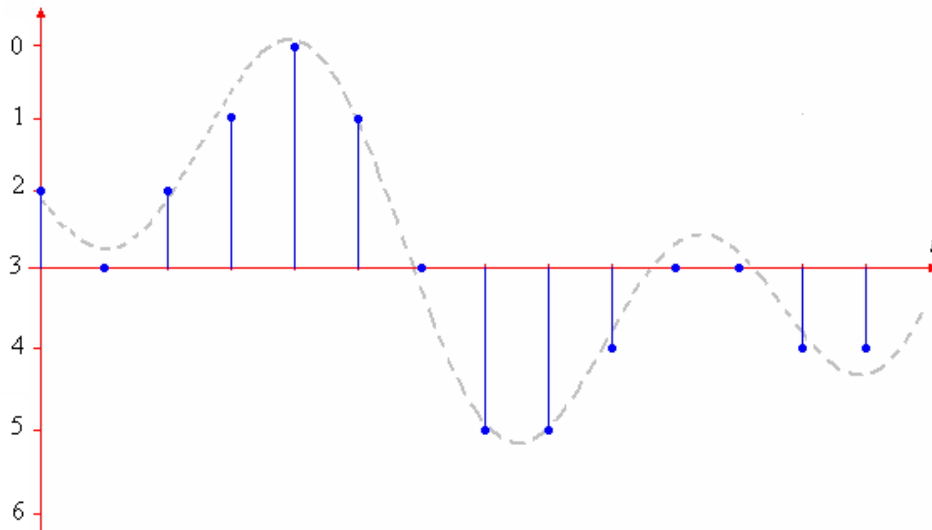
In order to transmit the value of 9 in binary we would need 4 bits,
however for -9 5-bits are required (one for the sign).
For generality, we would say that all values in the set {-9, -6, -3, 0, 3, 6, 9} can be transmitted by
using 5 bits each. This is a terrible waste of bits.

Instead, we can assign an integer for each possible value so the set {0, 1, 2, 3, 4, 5, 6} would
represent all the possible values. Instead of transmitting the actual values, we transmit the
corresponding integer which is merely 3 bits.

So the sequence is 2, 3, 2, 1, 0, 1, 3, 5, 5, 4, 3, 3, 4, 4 using the mapping

provided that the receiver knows the inverse of the map.

$$\begin{bmatrix} 9 & \rightarrow & 0 \\ 6 & \rightarrow & 1 \\ 3 & \rightarrow & 2 \\ 0 & \rightarrow & 3 \\ -3 & \rightarrow & 4 \\ -6 & \rightarrow & 5 \\ -9 & \rightarrow & 6 \end{bmatrix}$$


It is obvious that the number of
quantization levels is usually
chosen to be a *power of 2* so that
bits are used most efficiently.

Once the receiver gets the integer number, it first converts it to its original scale which might be just another integer or a real number, but the truncated portion is lost and cannot be recovered.

Samples with some error
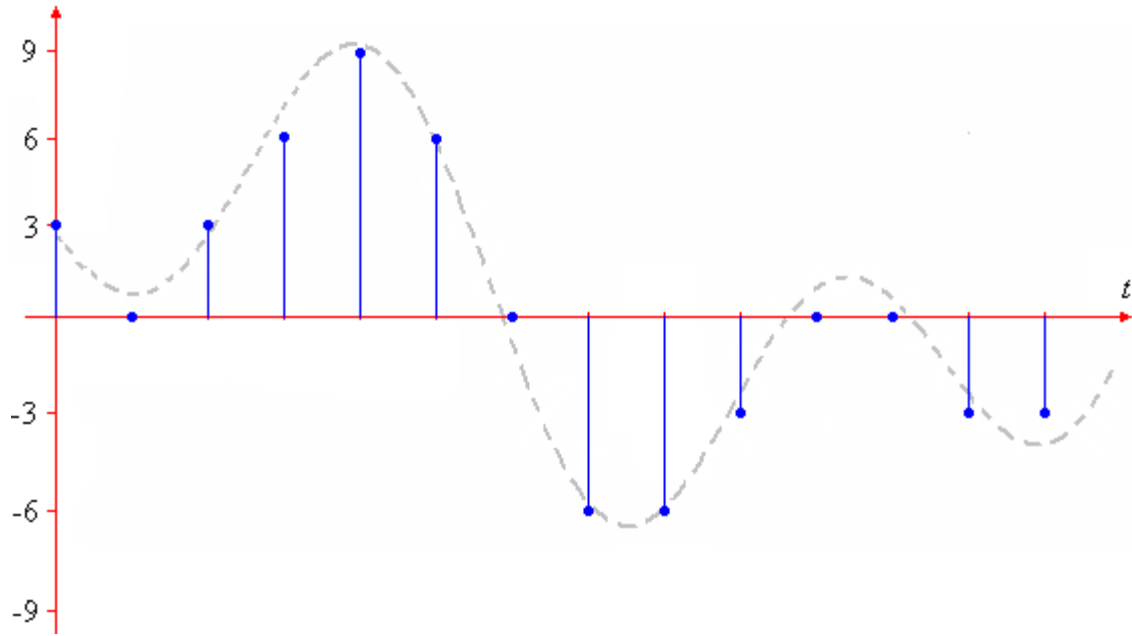
Quantization noise
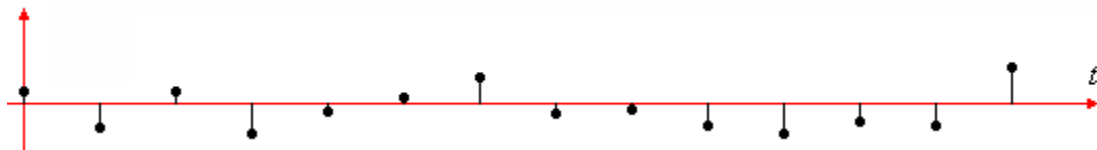
More quantization levels mean less quantization noise
One more bit reduces quantization error to half

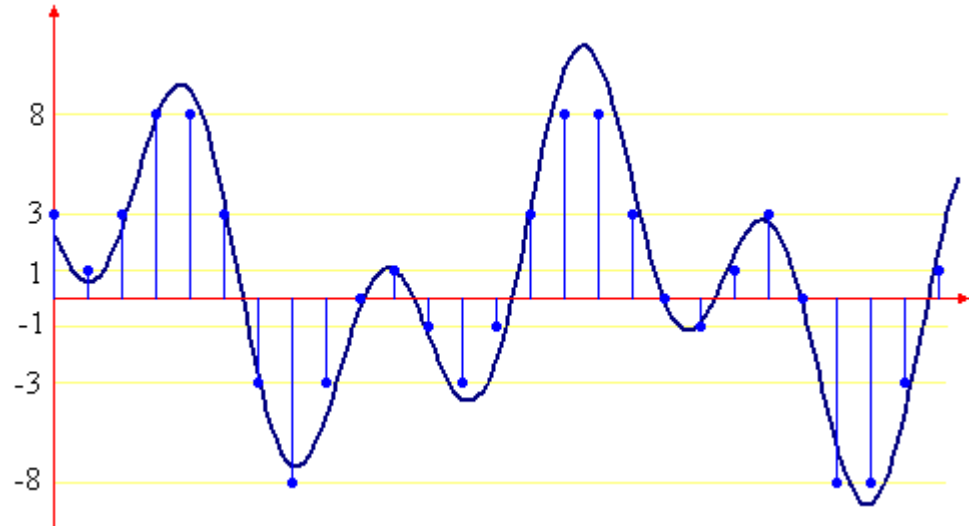Uniform quantization levels for voice signal samples are not quite efficient, because

1. Voice signal statistics say that the signal roams around zero, most of the time. Uniform quantization would waste many quantization levels for high values although the signal rarely assumes such values.
2. One could not tell the difference between two high levels by just hearing them, nor it is necessary to distinguish them if we are just trying to understand what is being spoken (phone system, for example).
3. We need more quantization levels in lower levels in order to tell apart low signal samples. We wouldn't understand what is being whispered, otherwise. (now that's a big problem in a phone conversation)

Consider the quantization levels in the figure

speech signal level (normalized) vs its probability
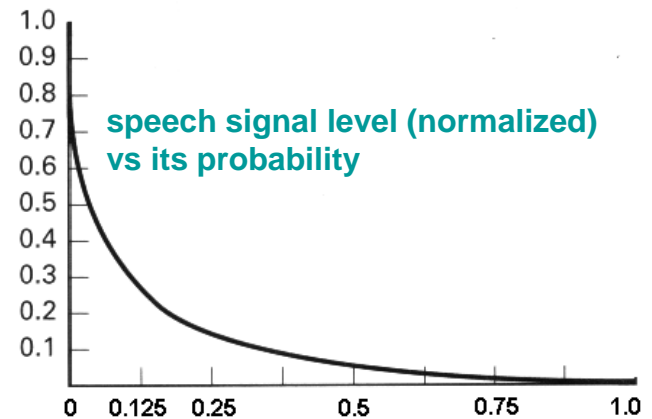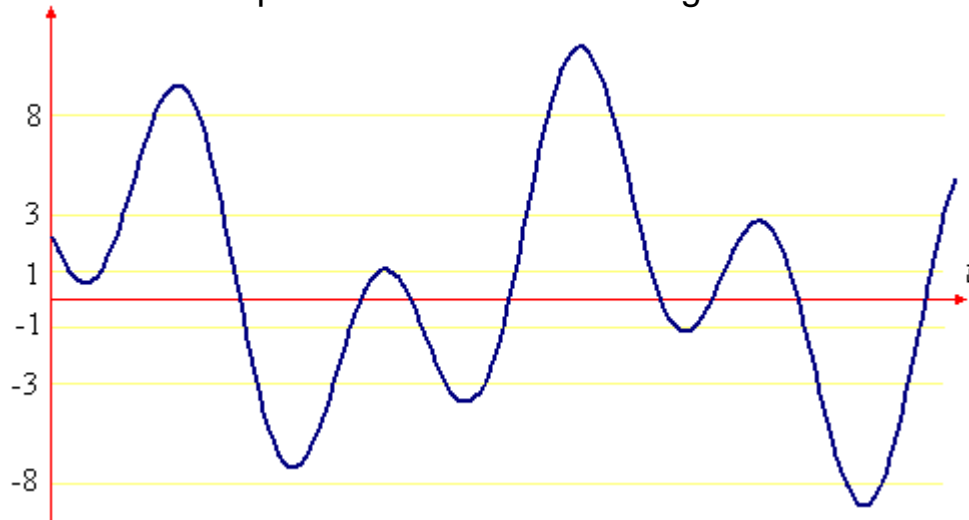
We have same number of levels but low voltage levels are more accurately measured than the high levels.

Such a schema would allow low levels to be measured more accurately.
But this increases the quantization errors made in the measurement of higher level samples.


Since the signal is mostly around (and close to) zero, on the average we reduced the total error.

Voice sample statistics can be analyzed and quantization levels can be selected accordingly so that minimum amount of total error is achieved.

But, it is difficult (or meaningless) to manufacture Analog-to-Digital Converter chips (which will generate binary values corresponding analog voltages) with such non-uniform levels
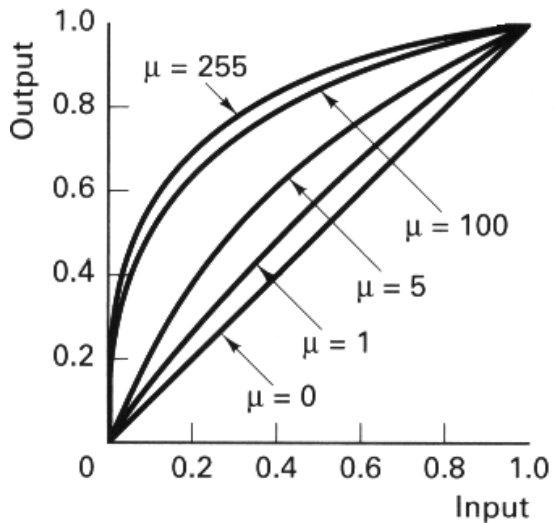
Instead of assigning non-uniform quantization levels, it is practical to pass the signal through a nonlinear device/amplifier which amplifies lower voltage levels more than higher voltage levels so that a uniform quantization afterwards shall have the same effect of non-uniform quantization

Voice(t)

compression

uniform Q

channel

expansion

$\overline{\text{Voice}}(t)$
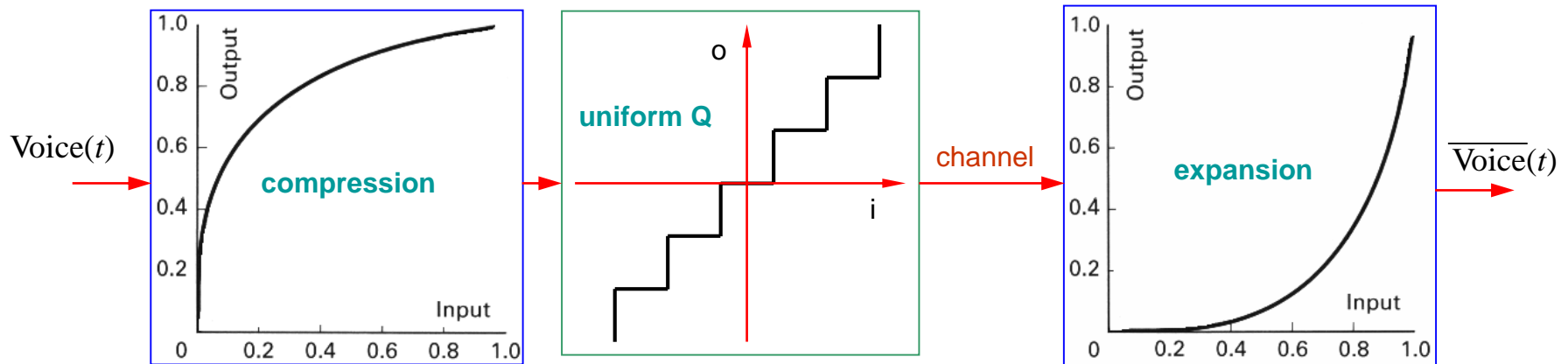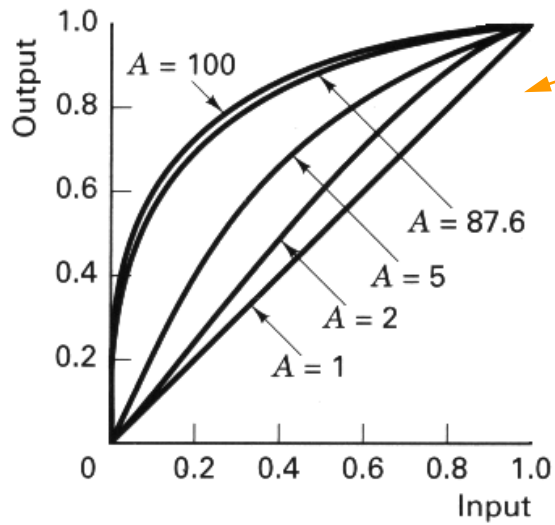
companding standard used in North America (μ-Law)
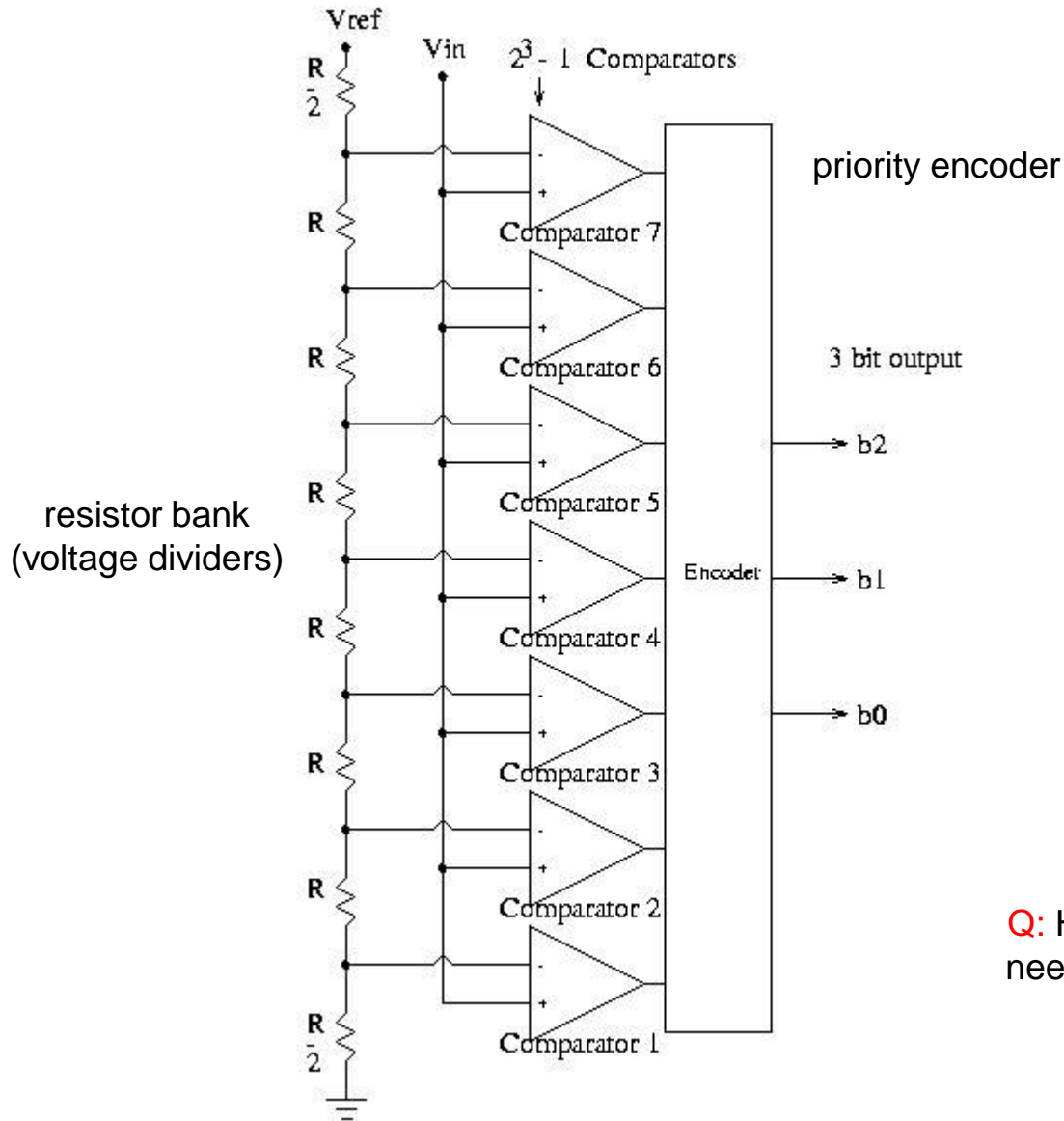
$$y = y_{max} \frac{\ln(1 + \mu(|x|/x_{max}))}{\ln(1+\mu)} \mathrm{sgn}(x)$$
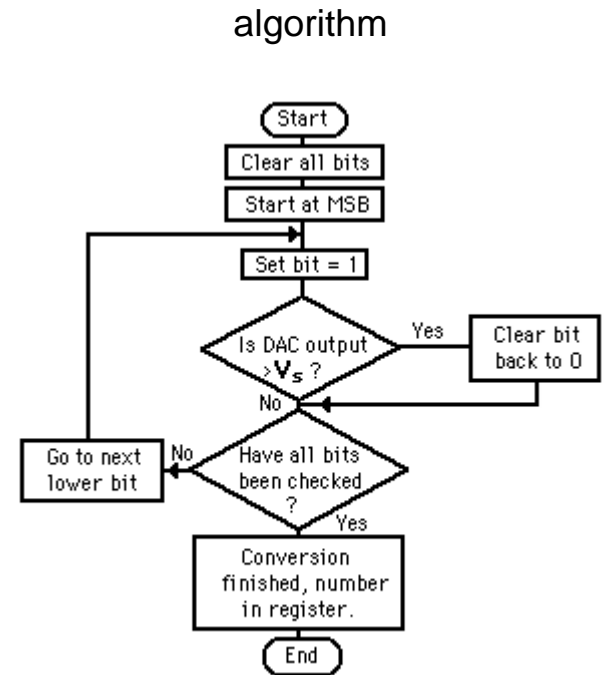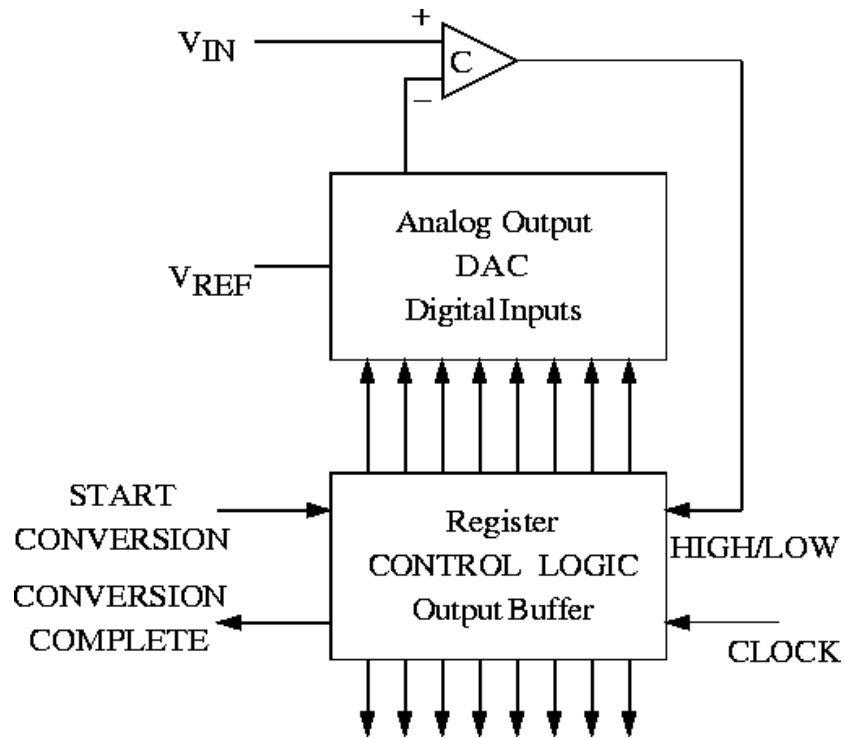
companding standard used in Europe (A-Law)

$$y = \begin{cases} y_{max} \dfrac{A(|x|/x_{max})}{1 + \ln A} \mathrm{sgn}(x) & , \; 0 < \dfrac{|x|}{x_{max}} < \dfrac{1}{A} \\[3ex] y_{max} \dfrac{1 + \ln(A(|x|/x_{max}))}{1 + \ln A} \mathrm{sgn}(x) & , \; \dfrac{1}{A} < \dfrac{|x|}{x_{max}} < 1 \end{cases}$$

Parallel Analog-To-Digital Conversion

priority encoder

3 bit output

→ b2

→ b1

→ b0

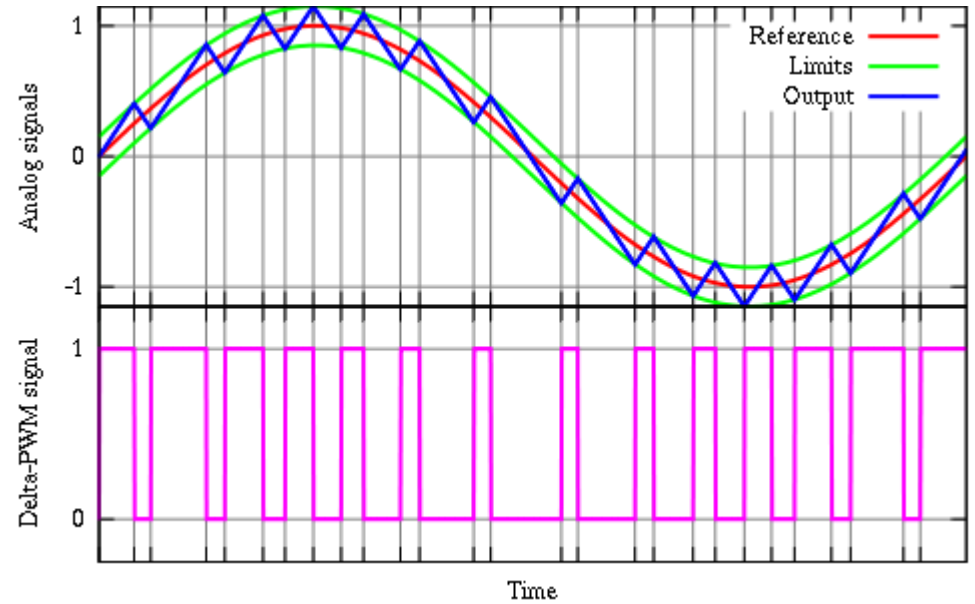resistor bank
(voltage dividers)
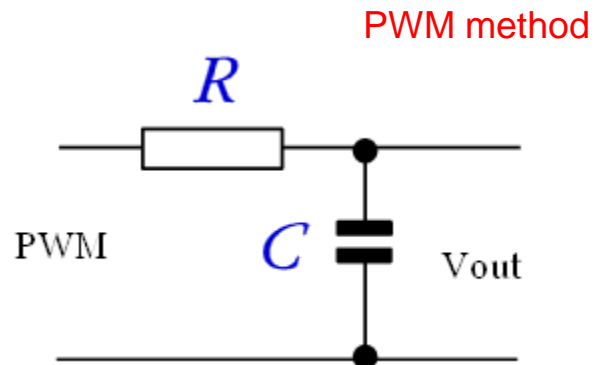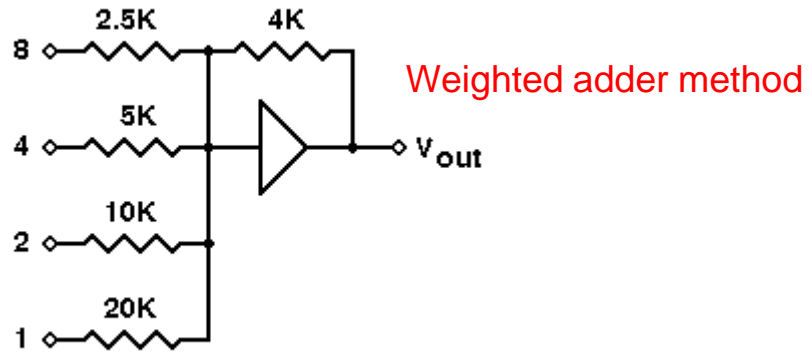
Q: How many comparators are
needed for 14-bits conversion?

Analog-To-Digital Conversion via Succesive Approximation

algorithm

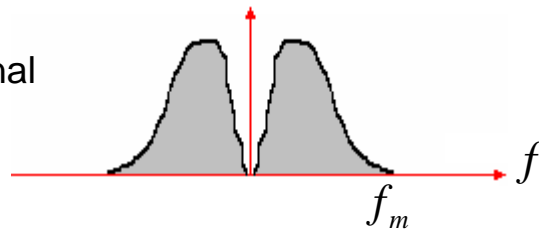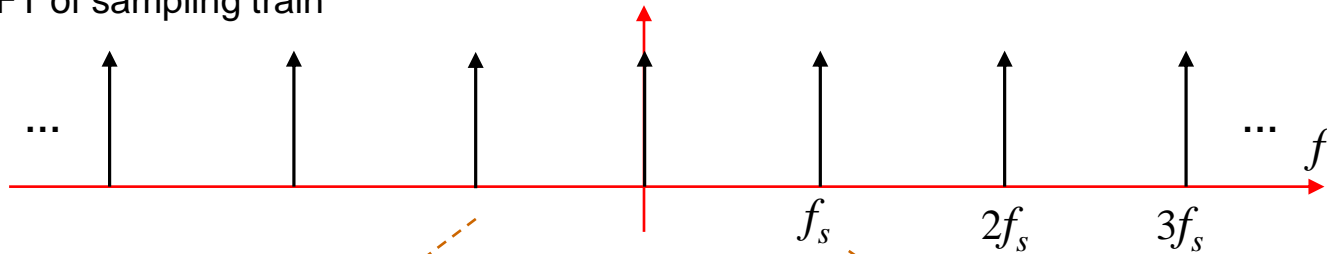# Digital-to-Analog Conversion

Weighted adder method

PWM method

Remember the appropriately sampled baseband signal?

Spectrum of the baseband signal

f_m

f

FT of sampling train

...

...

f

f_s

$2f_s$

$3f_s$

Spectrum of the sampled signal

$f_s > 2f_m$

...

...

f

$-f_s$

$-f_m$

f_m
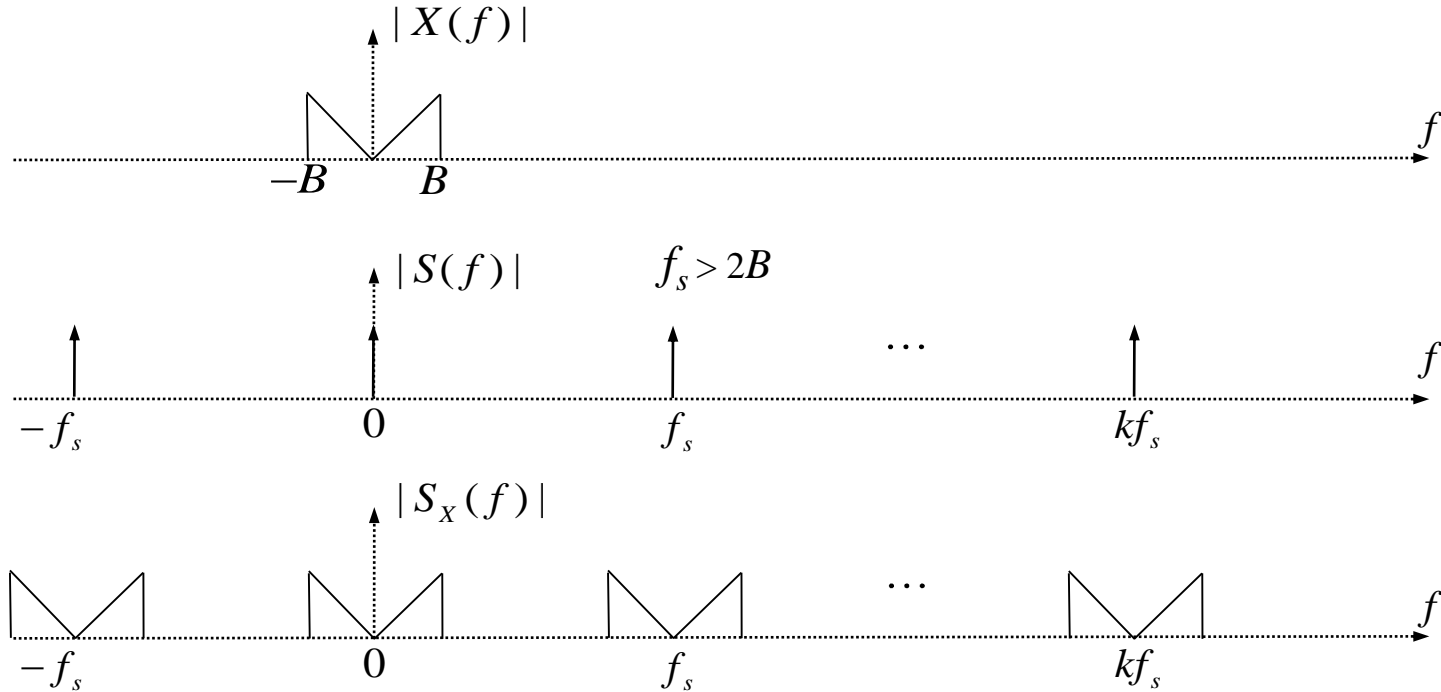
f_s
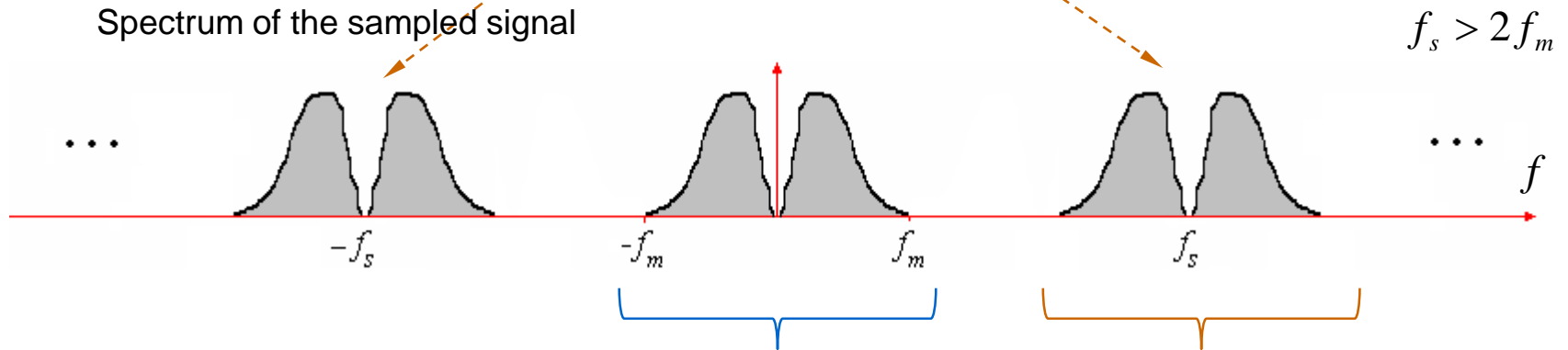
we would use a LPF in order to extract this
and obtain original baseband signal

what is we use a BPF
and extract this part?

Frequency Up Conversion

$|X(f)|$

$-B$ B

f

$|S(f)|$ $f_s > 2B$

$\cdots$

$-f_s$ 0 f_s kf_s

f

$|S_X(f)|$

$\cdots$

$-f_s$ 0 f_s kf_s

f

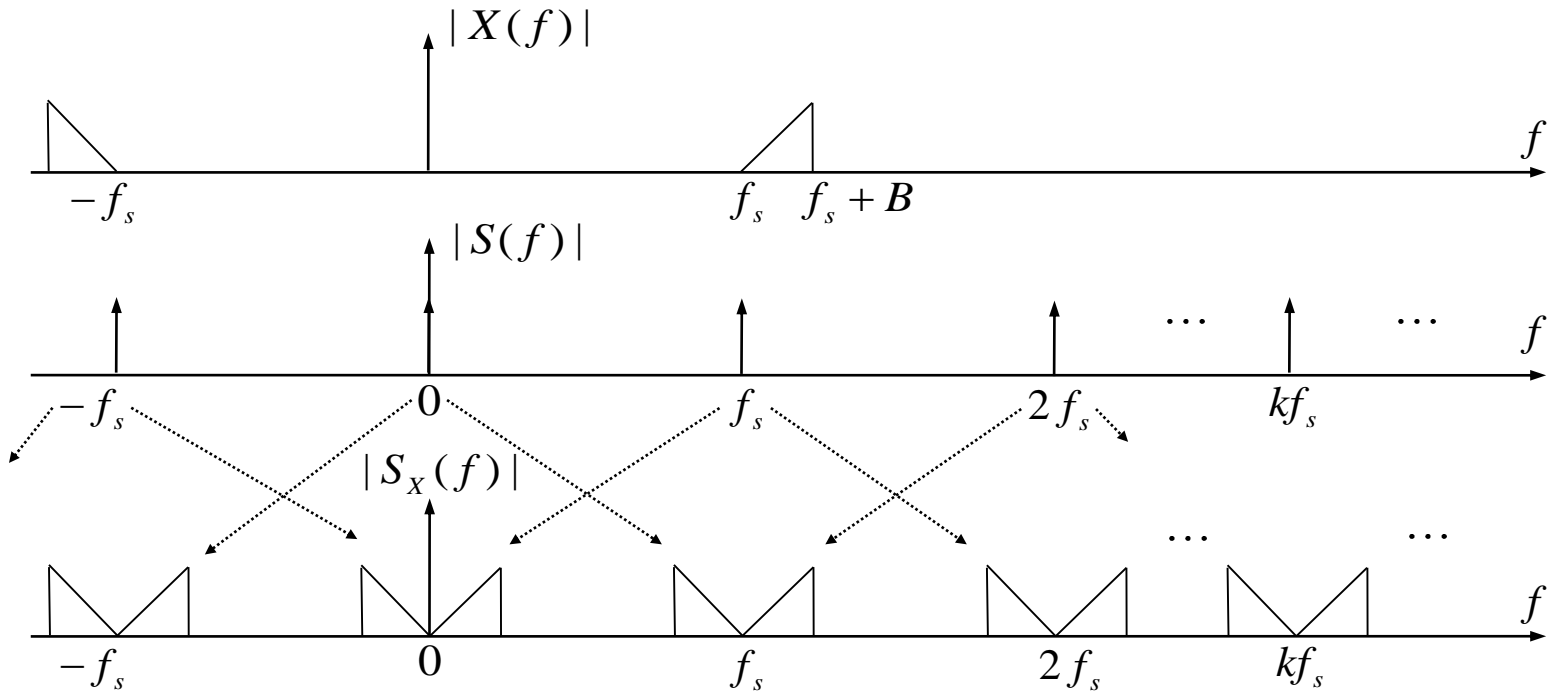One can conveniently extract one of the HF bands with an suitable BPF

Frequency Down Conversion
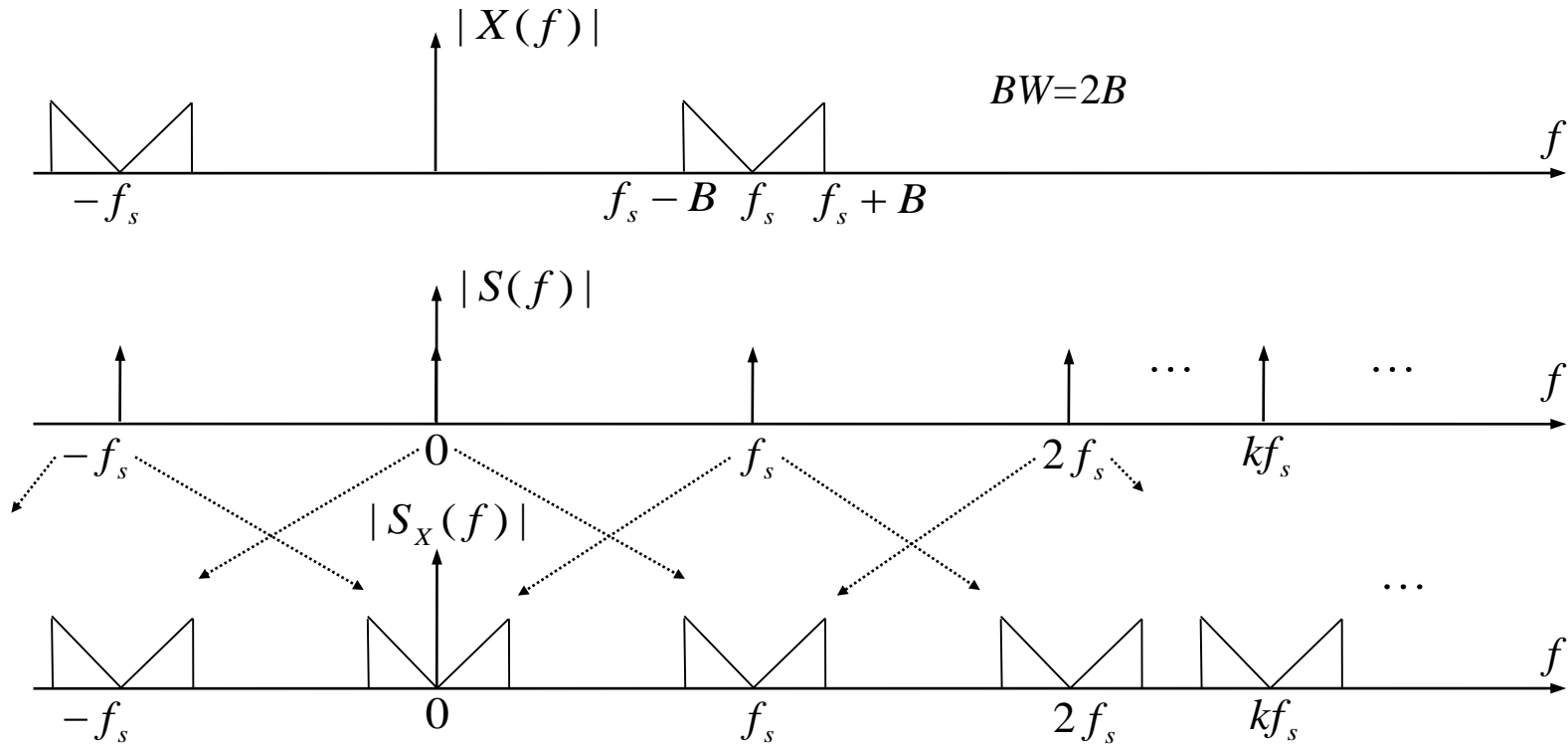
Refinement on Nyquist's sampling criterion : A signal with the bandwidth of BW Hz can be reconstructed from its samples if it is sampled at a rate greater than 2BW s/s
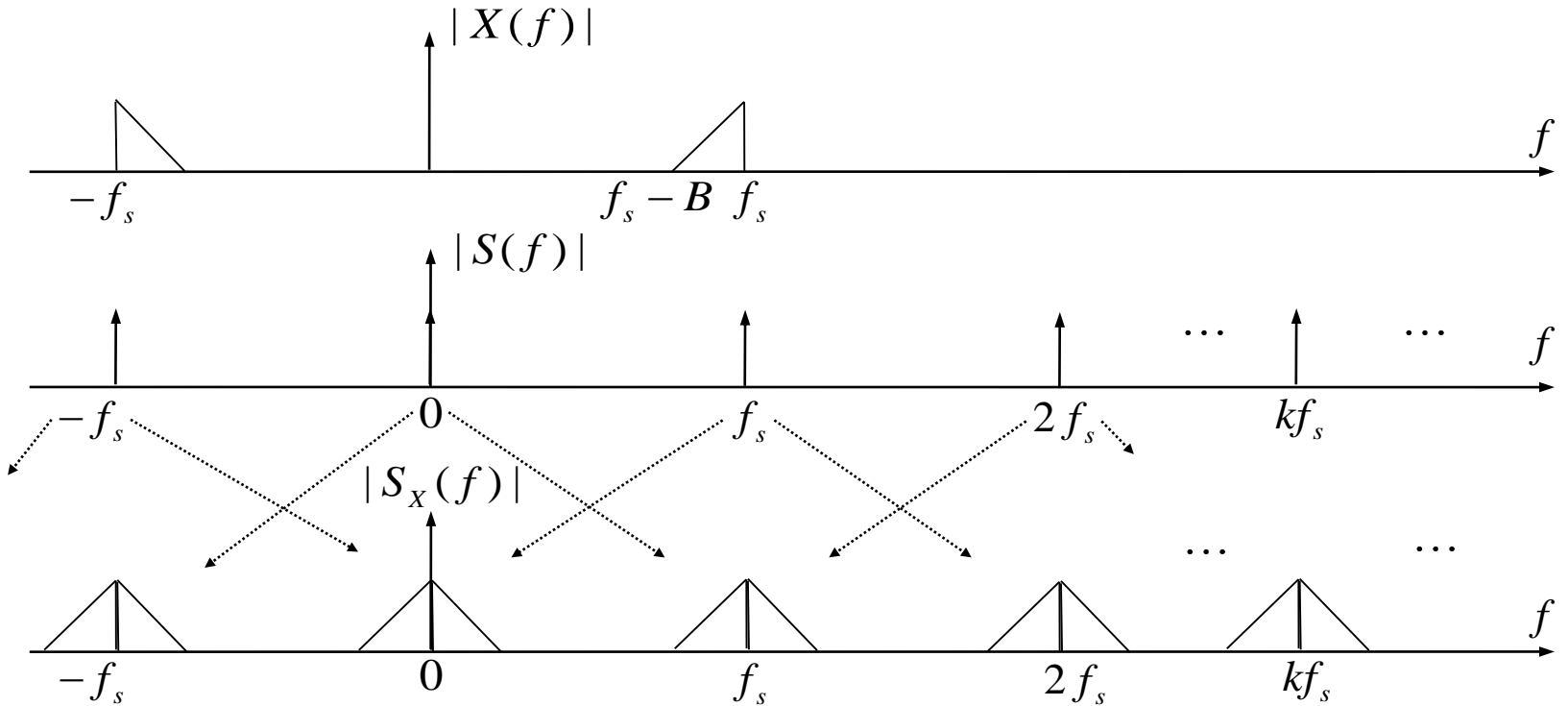
$$f_s > 2BW$$

Down Conversion of a Single Side Band

corollary: A SSB signal can be both demodulated and digitized at the same time
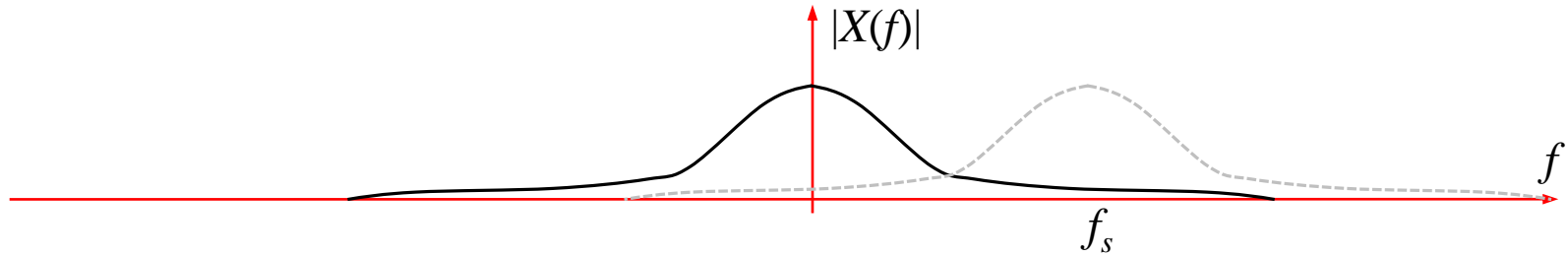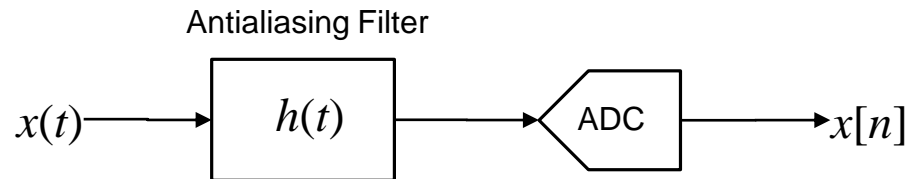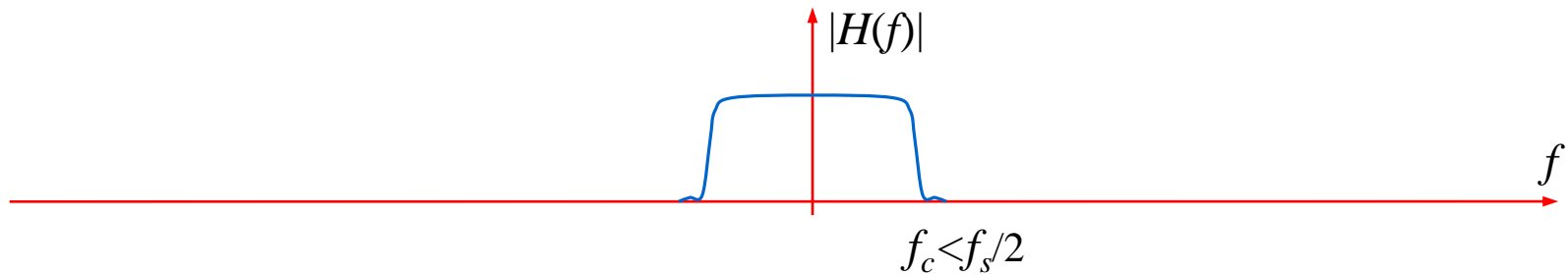
Spectral Inversion

corollary: We need to be careful about whether it is USB or LSB

If the BW of the Signal is Wider than the Sampler can Handle

(which is the case, most of the time)

$|X(f)|$

f

f_s

One wishes to limit the bandwidth in order to prevent aliasing

$|H(f)|$

f

$f_c < f_s/2$

Antialiasing Filter

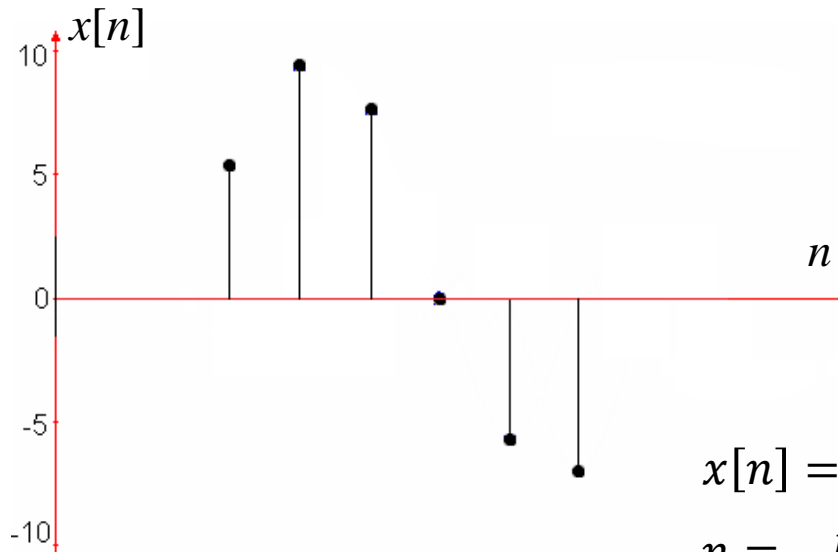$x(t)$ → $h(t)$ → ADC → $x[n]$

Not something we actually want, but something we are forced to have, in order to protect the sampled portion

Another Problem Solver
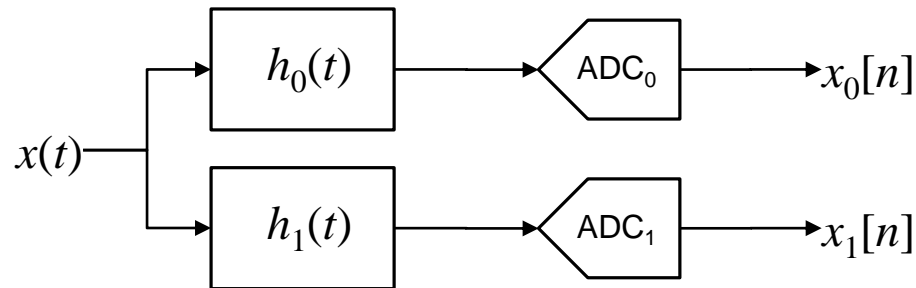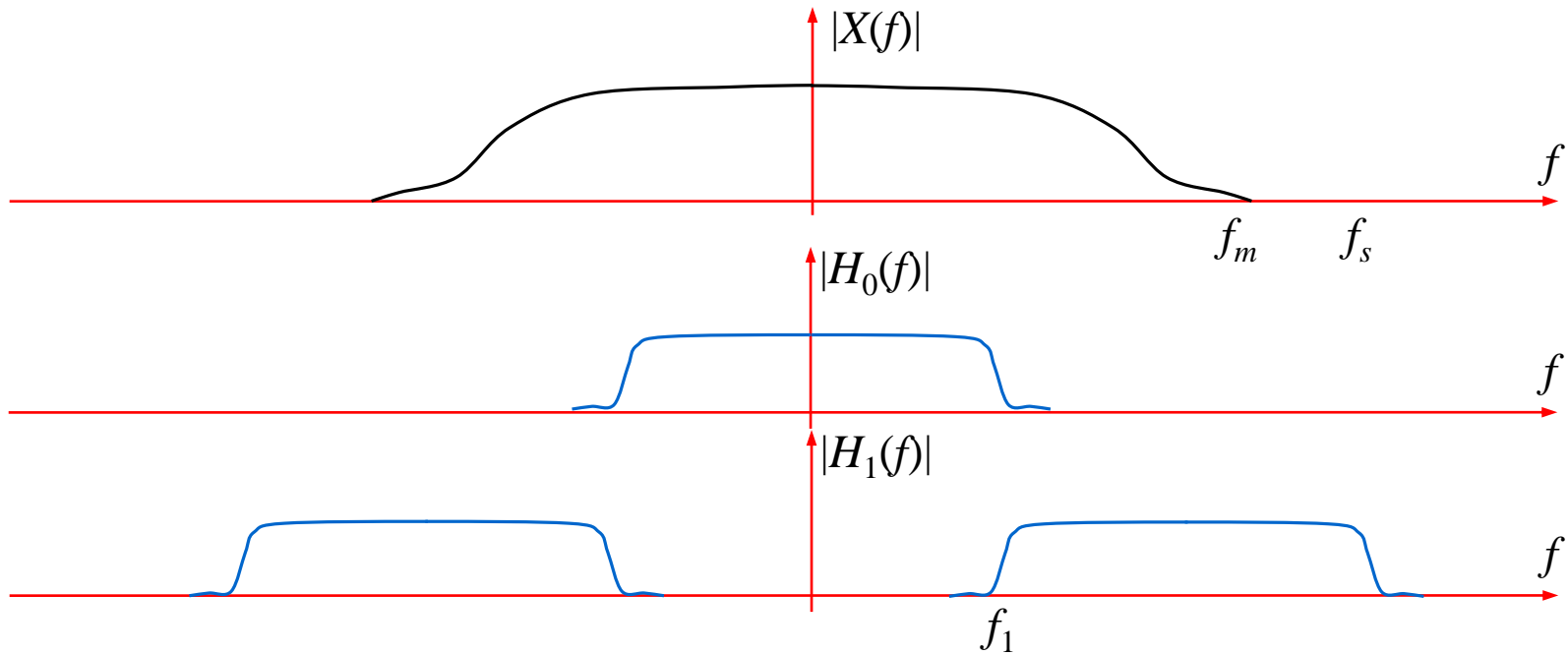
Hmw: Let f_m=100 kHz, and maximum sample rate of ADCs are 100 ksps. Determine filters.

We have the samples, so what?

$$x[n] = \{\cdots, 5,9,7,0,-6,-8,\cdots\}$$

$$n = -N_{min} \cdots - 1,0,1,2,\cdots, N_{max}$$

or $\quad n = 0,1,2,\cdots, N-1$

Now, we can do;
1. Digital signal procesing on the data, including
   a. data compression (source coding)
   b. error correction (channel coding)
2. Modulations / signal shaping in digital domain

We are not going to delve into digital signal processing in this course,
but only summarize how some things are done in digital

Discrete Fourier Transform (DFT)

DFT is the discrete counterpart of FT.
Although they are distinct transforms, their properties are verrry similar.

DFT is defined as

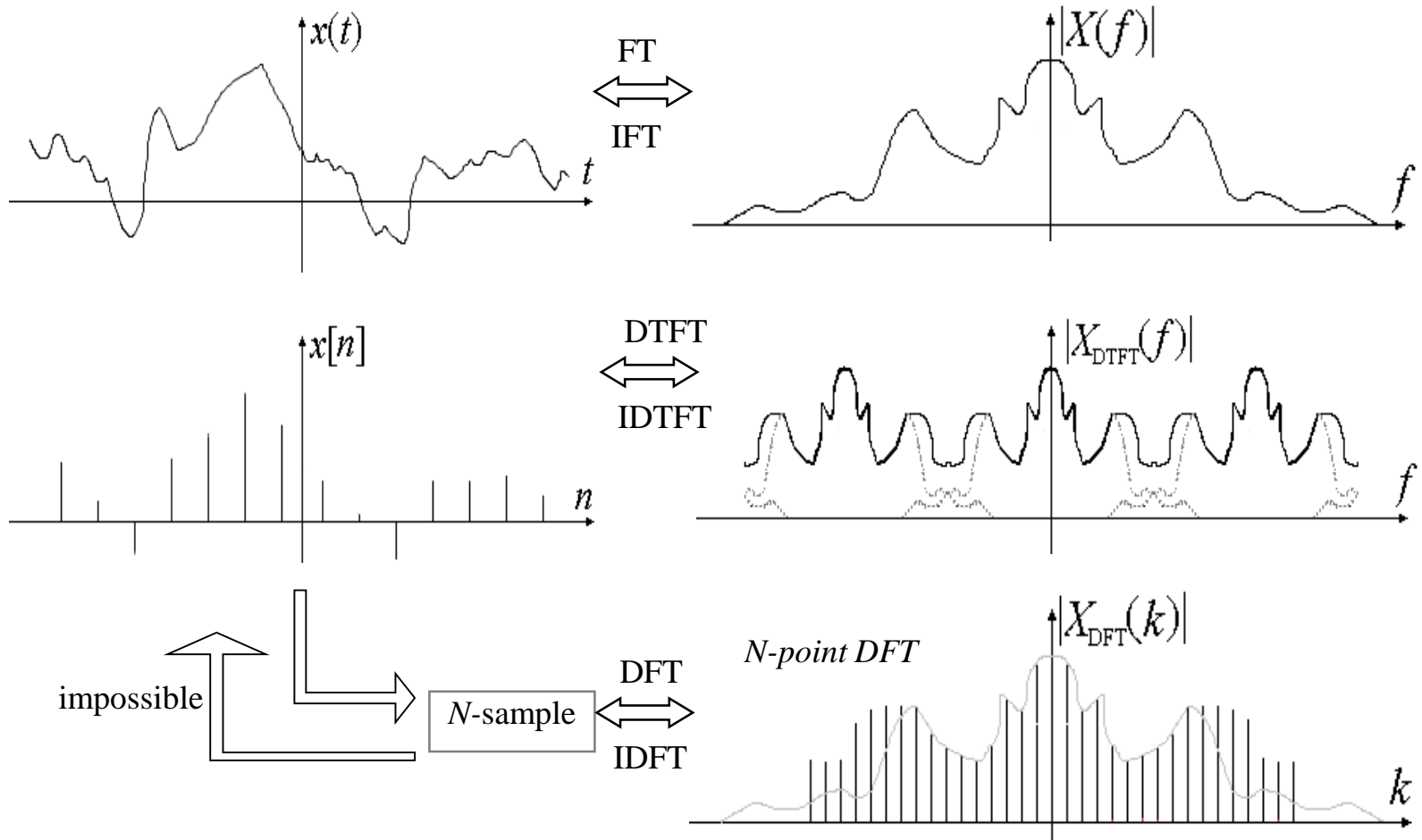$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}$$

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X[k]e^{j2\pi kn/N}$$

compare them with the continuous FT

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt$$

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft}df$$

where $x[\cdot]$, $X[\cdot]$, $x(\cdot)$ and $X(\cdot)$ are complex valued
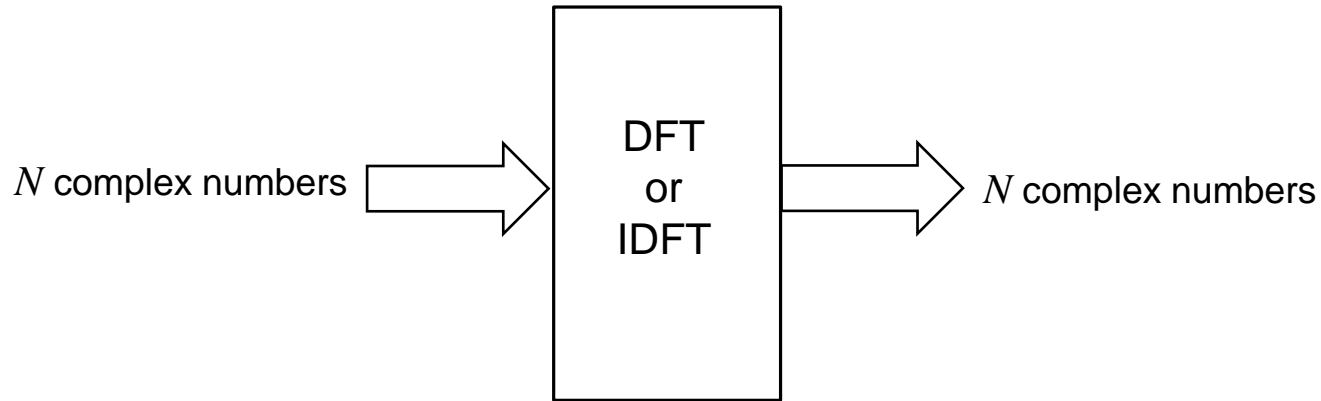
Relations

N samples can not represent entire waveform

DFT assumes that the sampled signal is periodic with period $\quad T = NT_s = N/f_s$

In that case DFT gives the Fourier Series coefficients   (assumption: $x(t)$ is bandlimited)

Calculation of DFT

N complex numbers → DFT or IDFT → N complex numbers

Think of DFT as one of several orthogonal linear transforms.
That is, the input data set is represented by the same number of orthogonal functions and the output data set is the coefficients (weights) of these functions whose weighted sum forms the original continuous signal (whose samples are the input data)

DFT and IDFT can be calculated using their definitions.
However, this is almost never the case in practice.

For all real-time signal processing, DFT/IDFT is calculated using fast calculation methods in which symmetry properties of FT are used
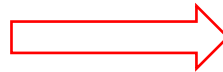
Fast Fourier Transform (FFT)

FFT is not a different transform, it is identical to DFT.
The only difference is that, FFT uses symmetry properties of the transform kernel ($e^{-j2\pi kn/N}$) in order to reduce the number of arithmetic operations (additions, multiplications).

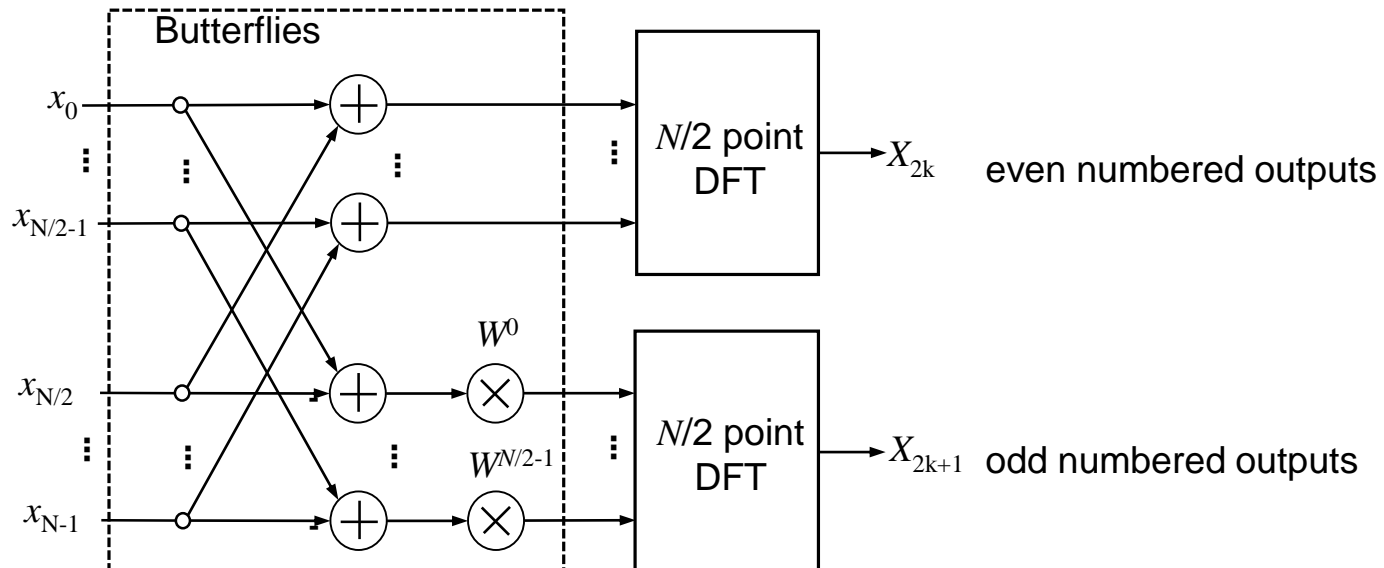$$X_k = \sum_{n=0}^{N-1} x_n W_N^{nk}$$

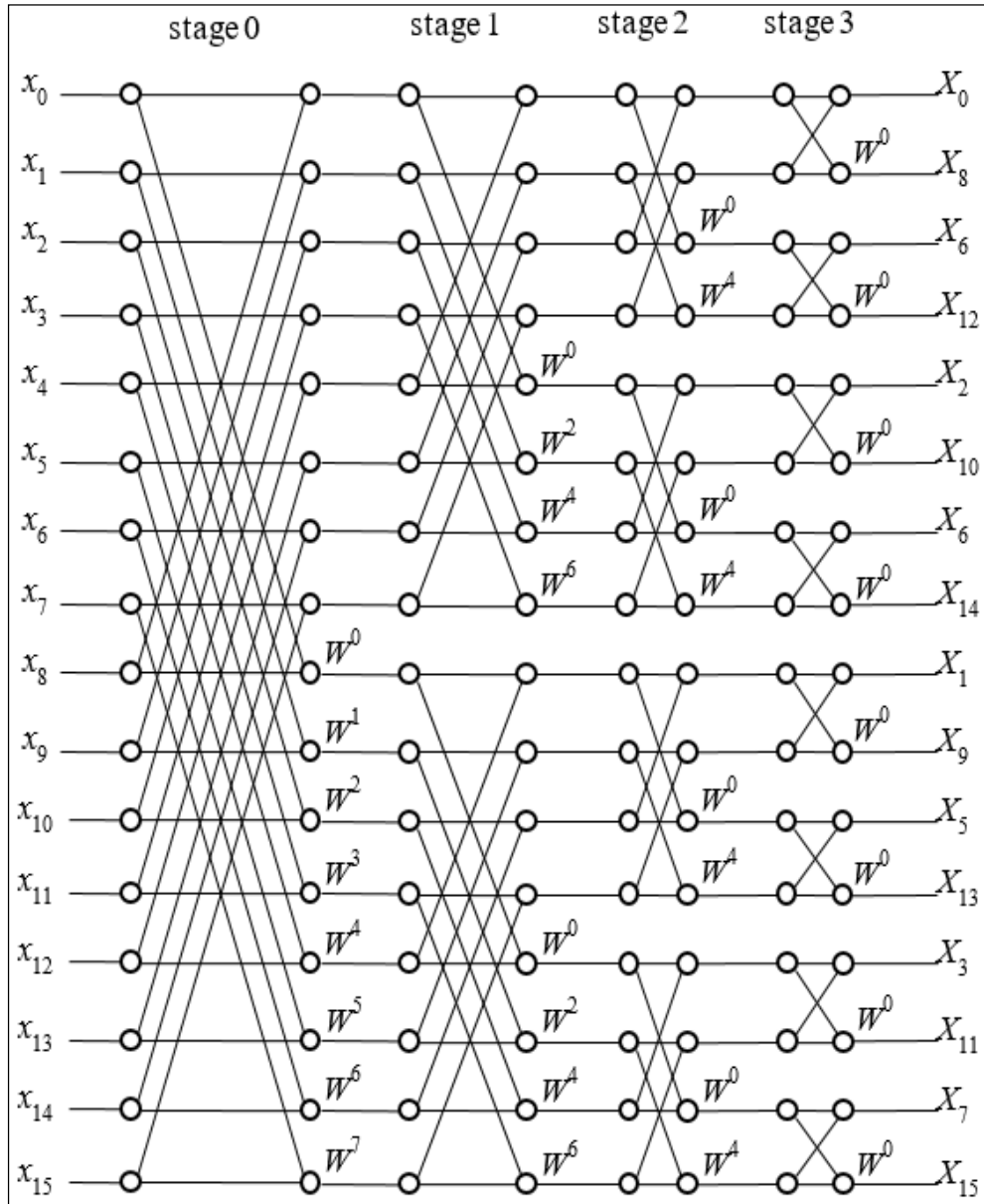where

$$W_N^{nk} = e^{-j2\pi kn/N}$$

odd and even numbered outputs are calculated separately

$$X_{2k} = \sum_{n=0}^{\frac{N}{2}-1} [x_n + x_{n+N/2}] W_{N/2}^{nk}$$

$$X_{2k+1} = \sum_{n=0}^{\frac{N}{2}-1} W_N^n [x_n + x_{n+N/2}] W_{N/2}^{nk}$$

N=16 point, Radix-2, Decimation in Frequency FFT

FFT reduces the complexity of DFT from

$$N^2$$

to

$$N/2 \, log_2(N)$$

(5120 instead of 1048576 for N=1024. Good savings)

There is much more in sampling / DFT than we summarized here.

Summary

1. What happens when we sample a continuous signal
2. Requirements for reconstructing the sampled signal
3. Frequency UpConversion / DownConversion while sampling / reconstruction
4. Relations between FS, FT, DFT, FFT

END