

Audio systems and processing

Aldo G. Cugnini

AGC Systems LLC

www.agcsystems.com



BroadcastEngineering

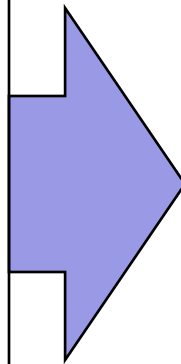
Outline

- Sampling
- Aliasing
- Digital interfaces
- Fourier transform
- Types of transforms
- Frequency domain
- Reconstruction
- Quantization
- FIR filters
- Sample rate conversion
- Critical sampling
- Audio compression
- Psychoacoustic masking
- Artifacts
- Lip sync
- Level compression
- Loudness processing
- Surround processing



Sampling

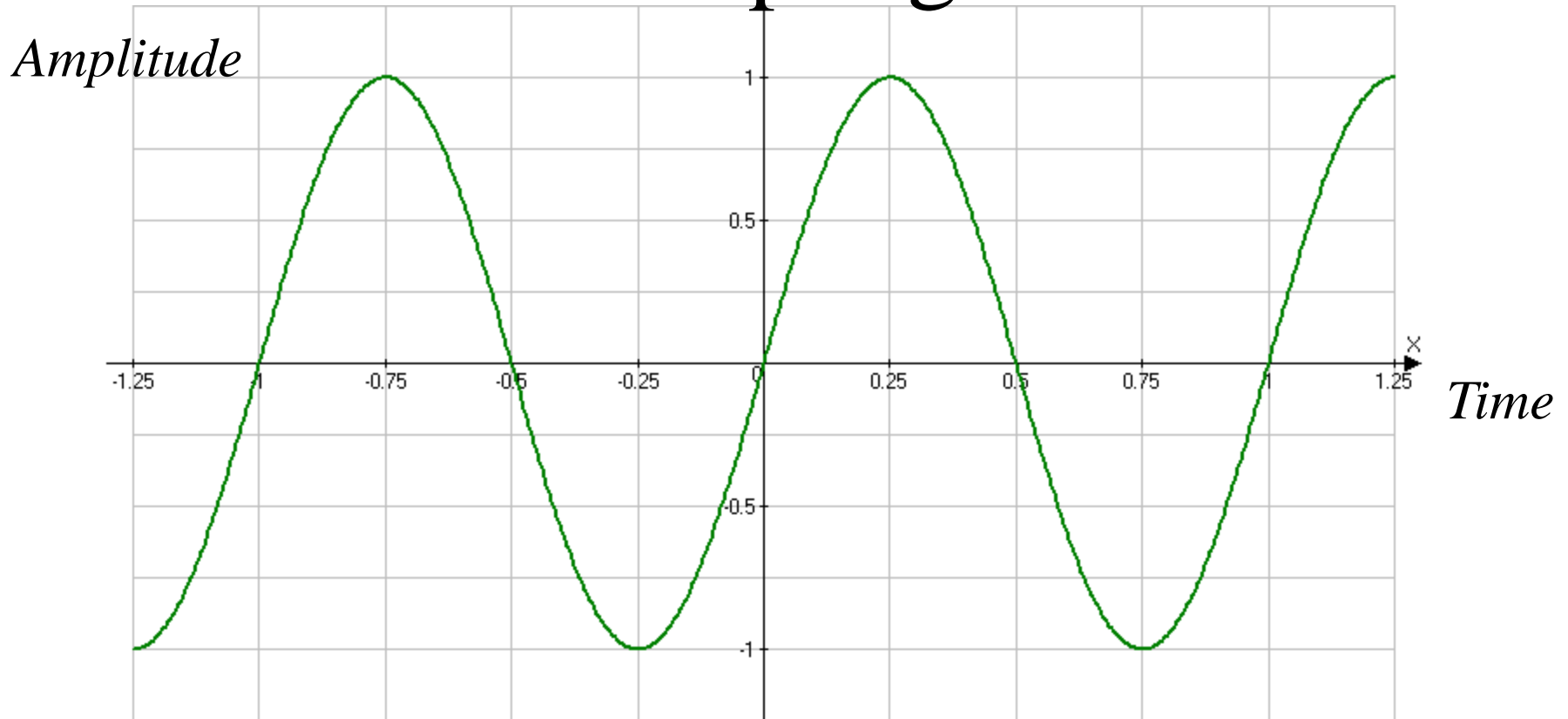
- Continuous-time signal
- Continuous-amplitude signal



- Discrete-time signal
- Quantized-level signal



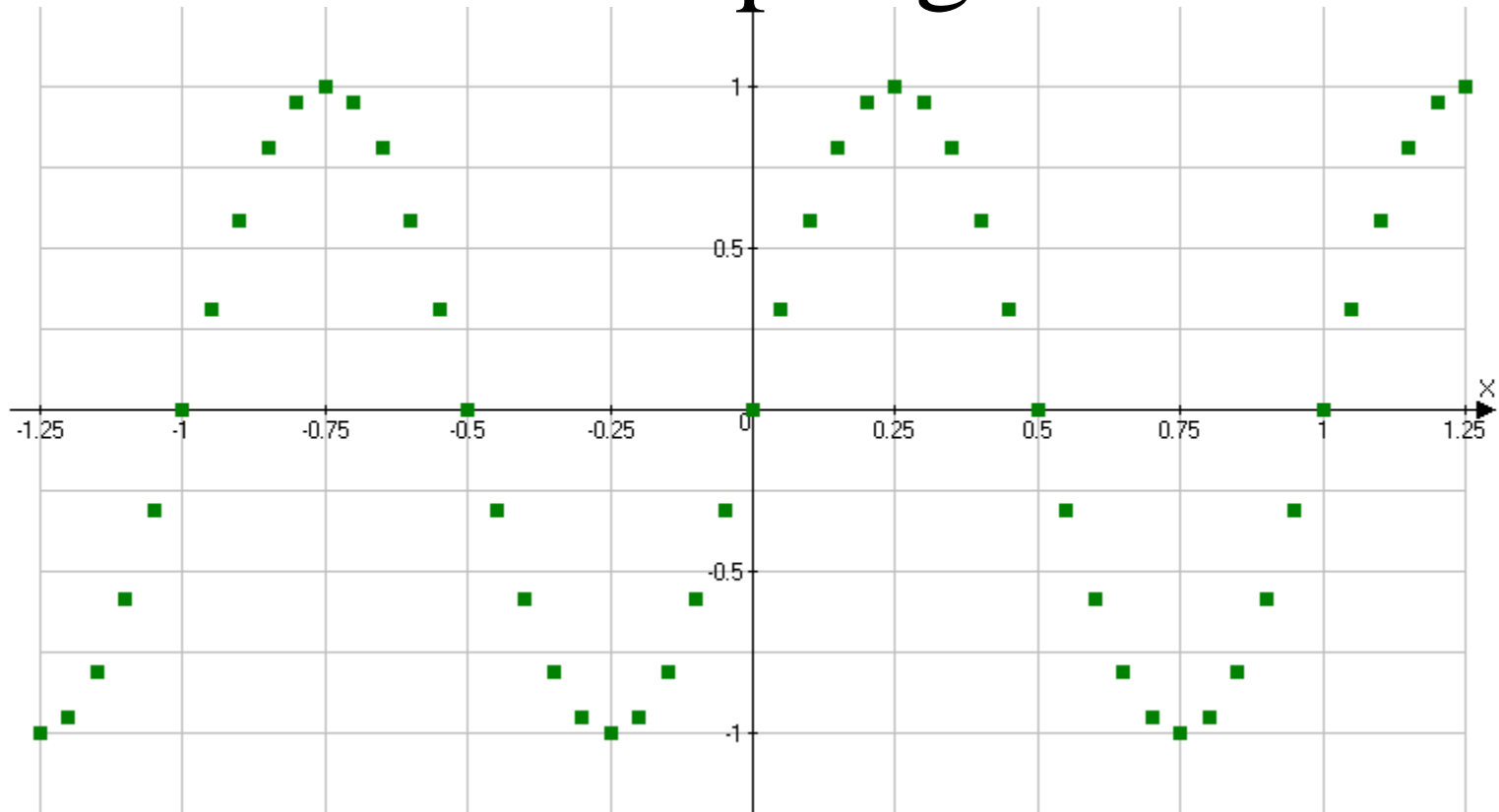
Sampling



Continuous-time signal



Sampling

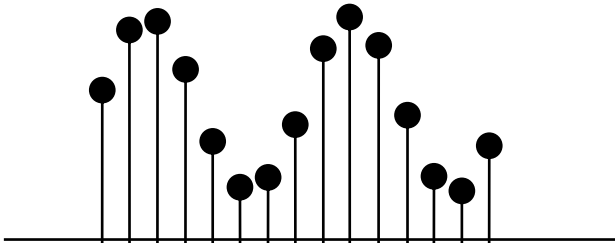


Discrete-time signal



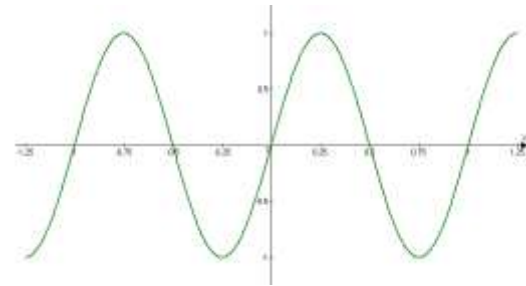
Reconstruction

- Discrete-time signal
- Digital signal



Lowpass
filter

- Continuous-time signal
- Analog signal

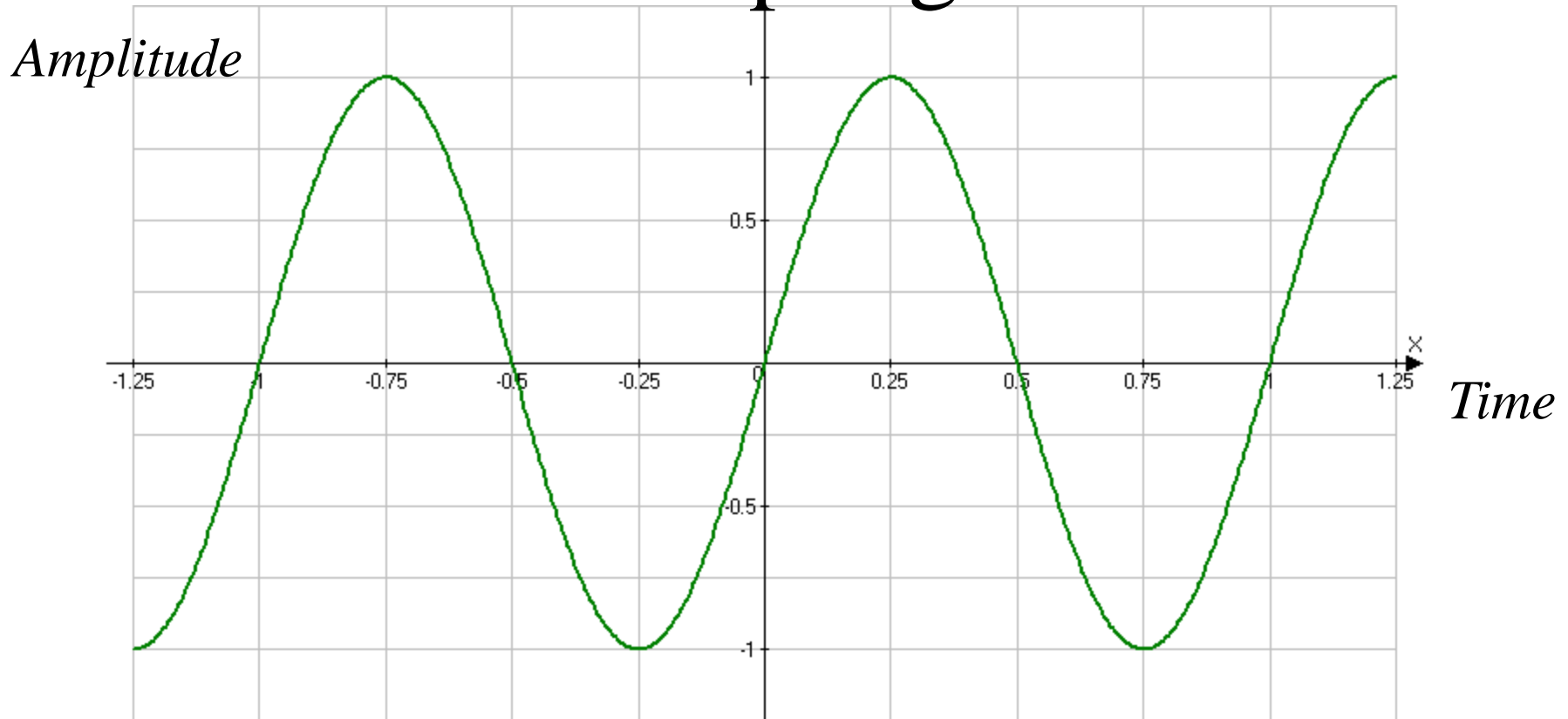


Sampling

- Nyquist Theorem
 - Signal must be sampled at least at twice the highest frequency component.



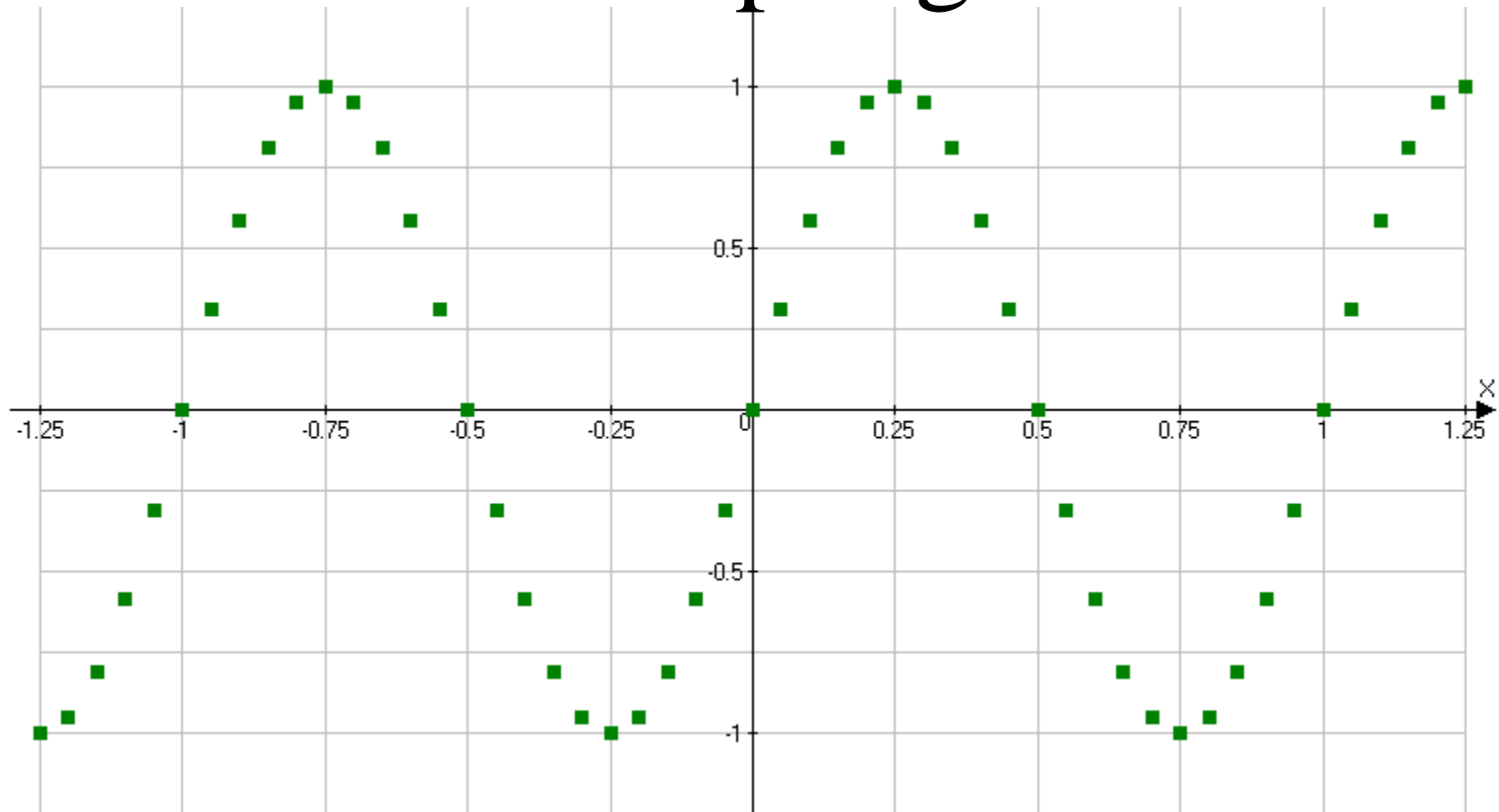
Sampling



Continuous-time signal



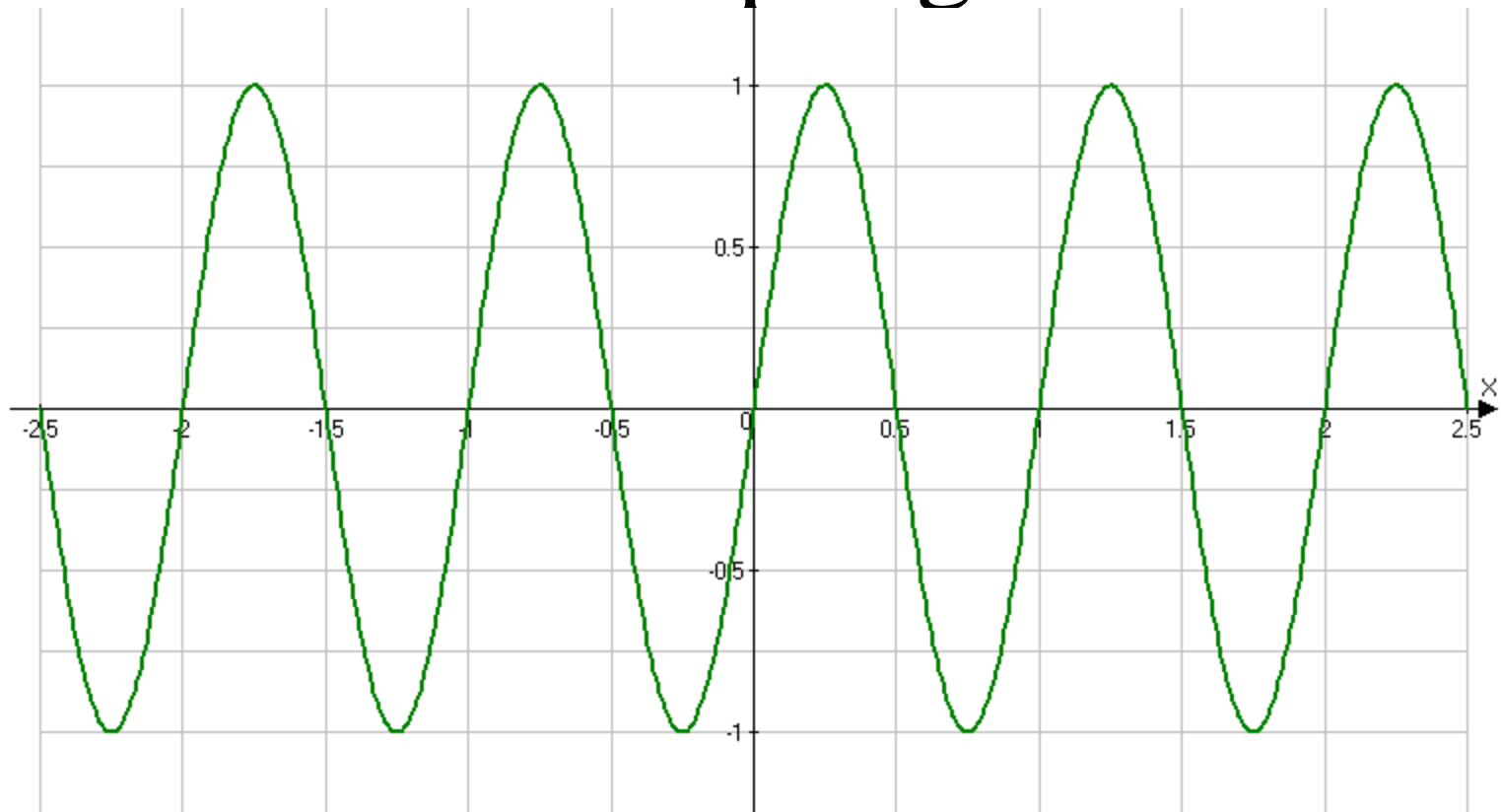
Sampling



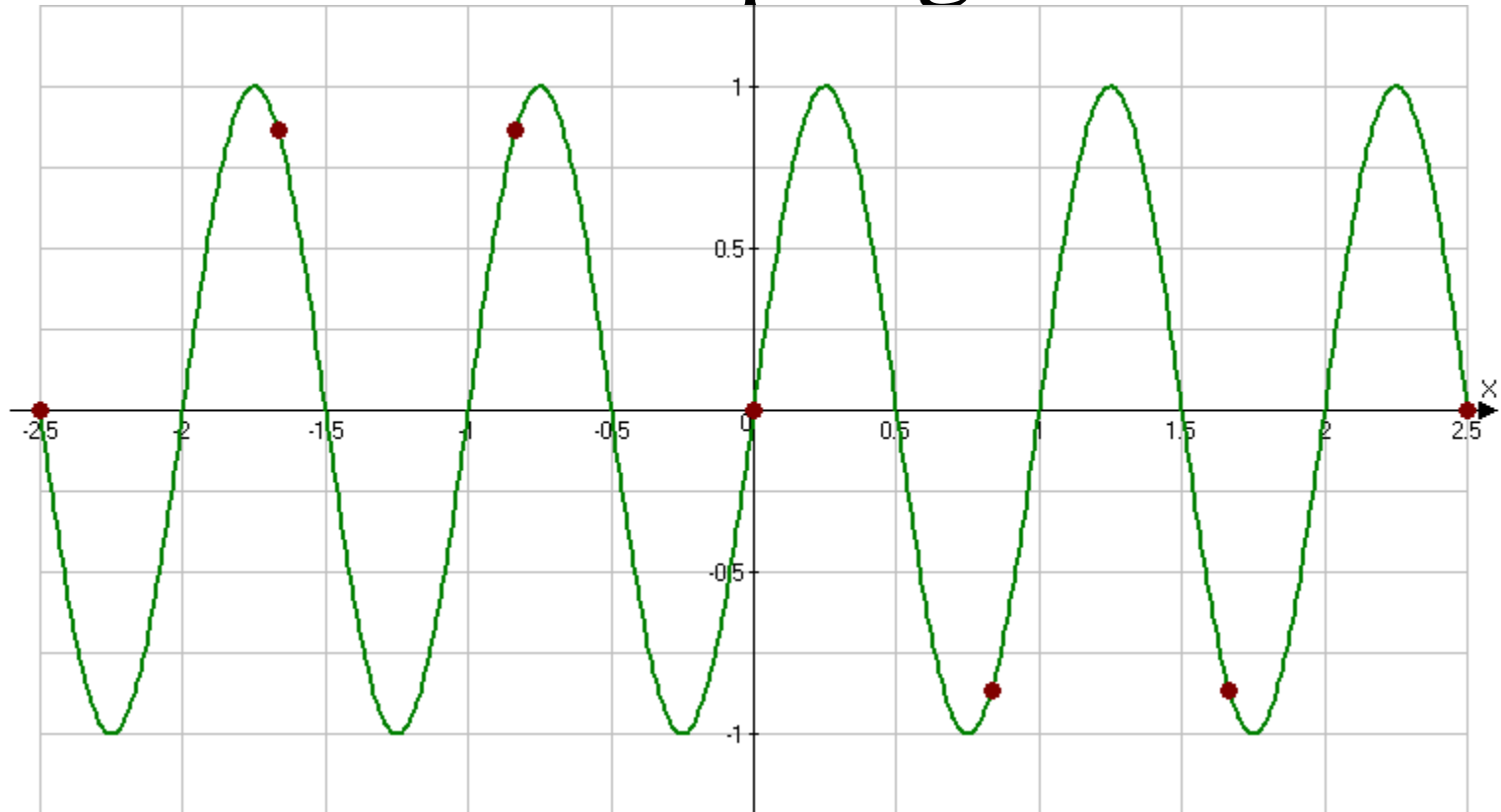
Discrete-time signal



Sampling



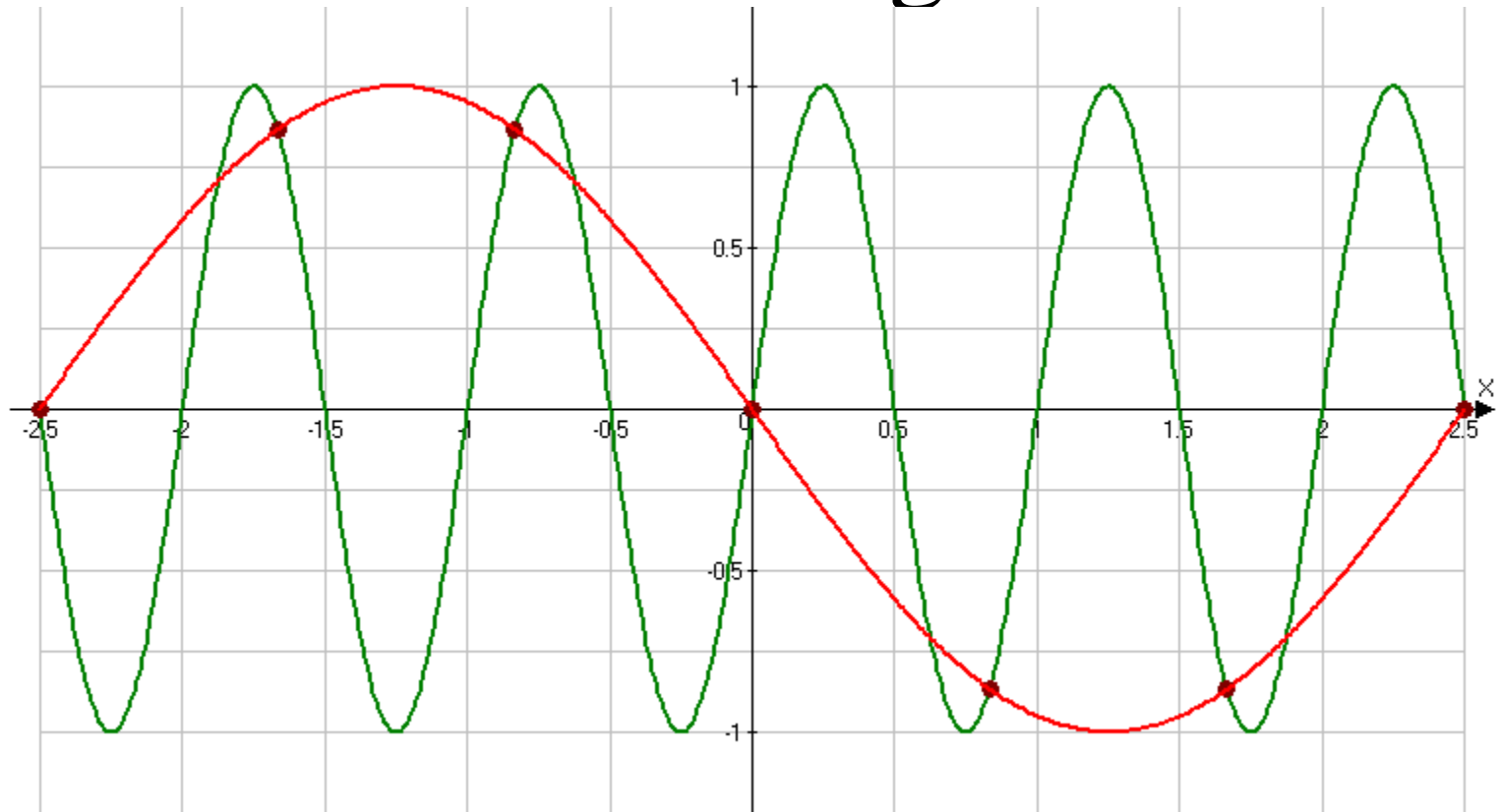
Sampling



Undersampling



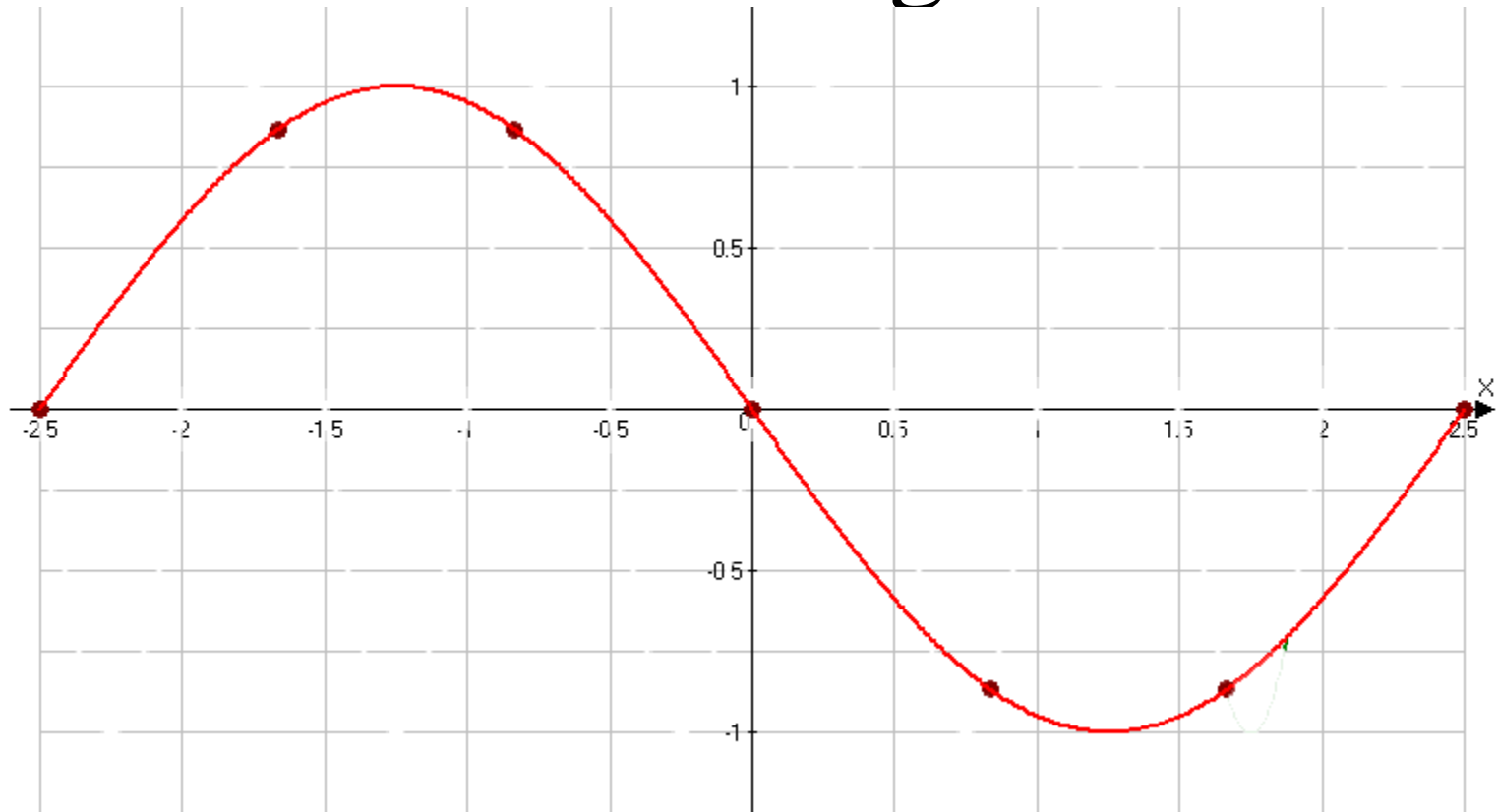
Aliasing



Signal is not sampled at a high enough rate



Aliasing



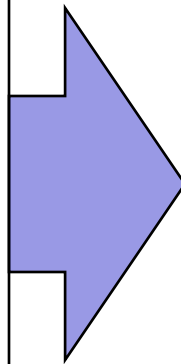
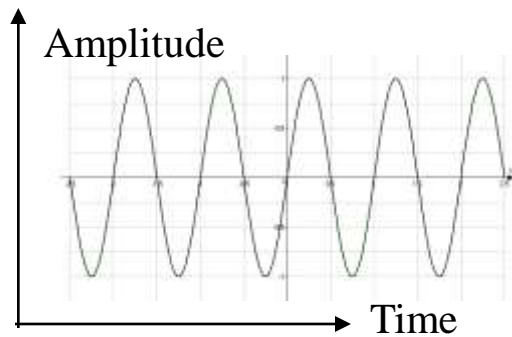
Signal is not sampled at a high enough rate



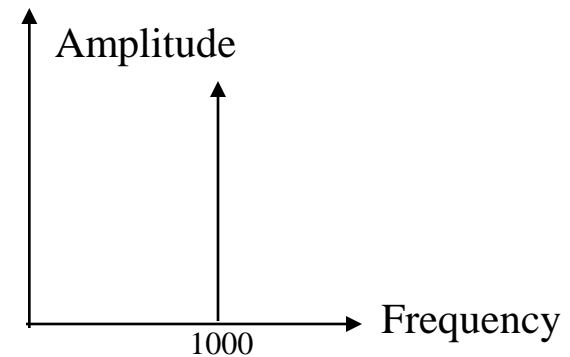
Fourier transform

- Time-varying signal representation

- $f(t) = \sin(2\pi \cdot 1000t)$



- Frequency representation
- $F(\omega) = \delta(1000)$



Types of transforms

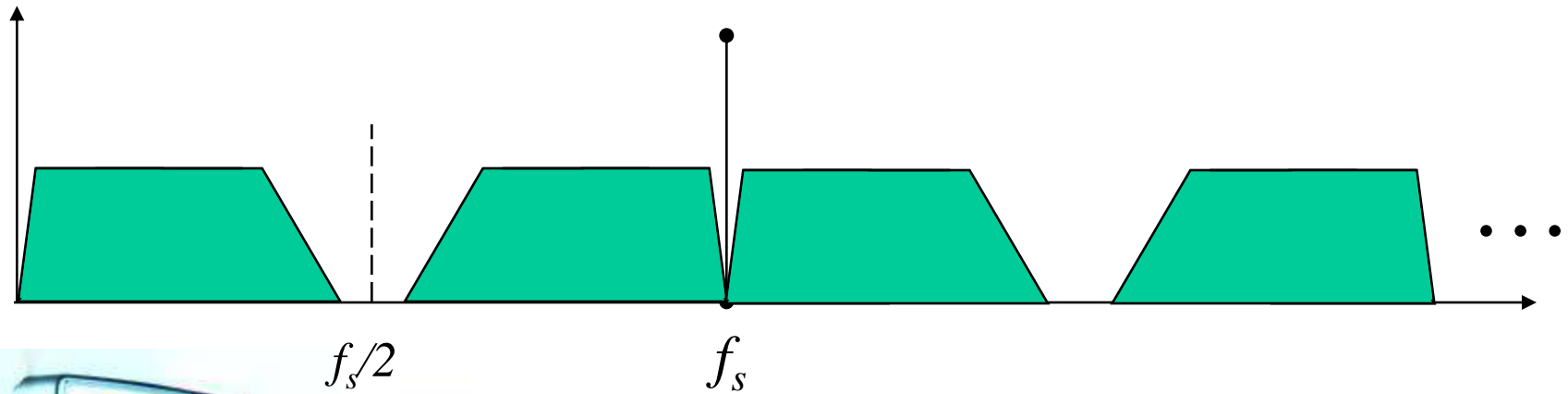
- Fast Fourier Transform (FFT)
- Discrete Fourier Transform (DFT)
 - Converts discrete samples into frequency samples
- Discrete Cosine Transform (DCT)
 - Special case of DFT
 - Used in MP3, AAC and Dolby compression



Sampling: Frequency domain



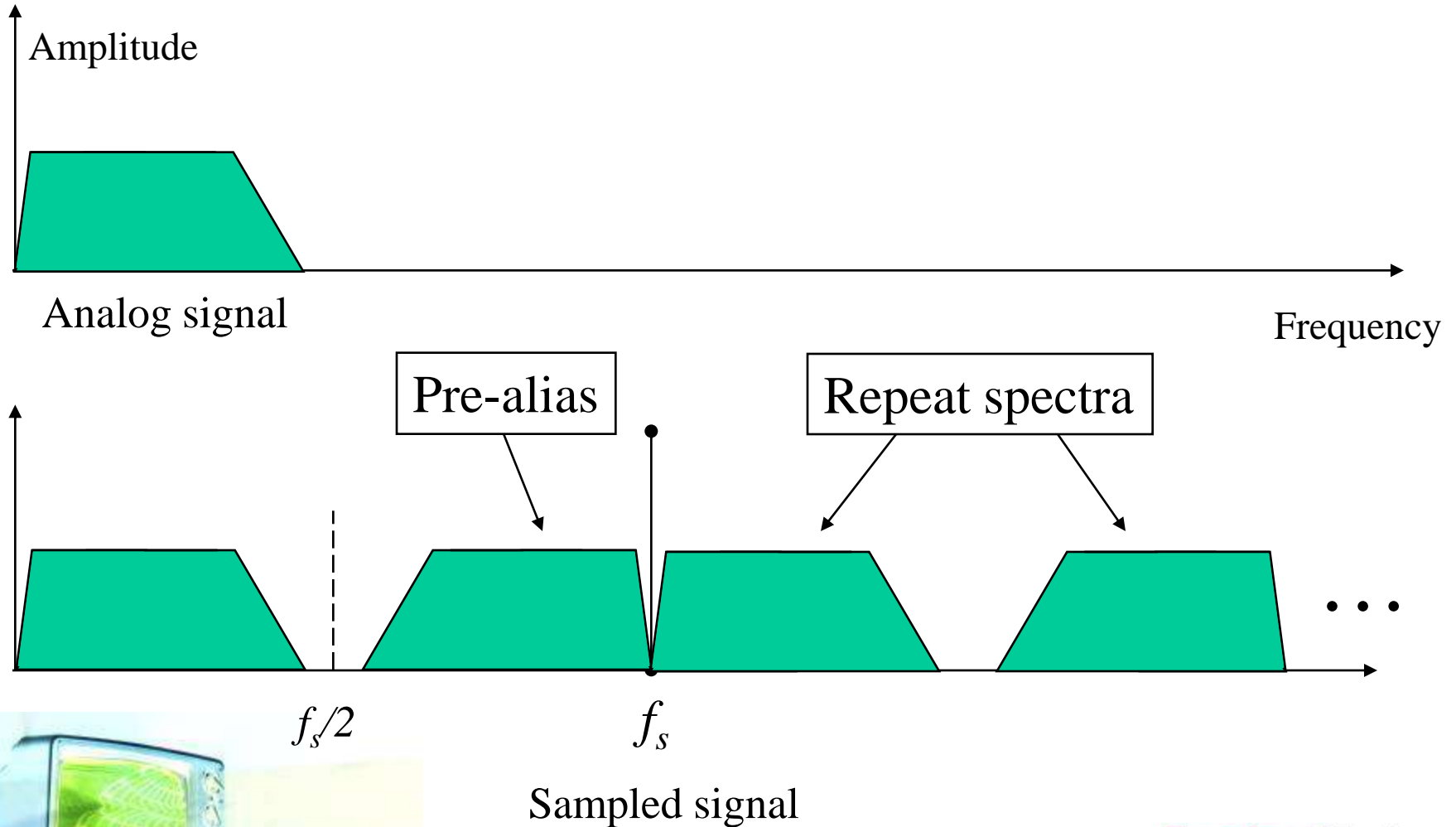
Sampling: Frequency domain



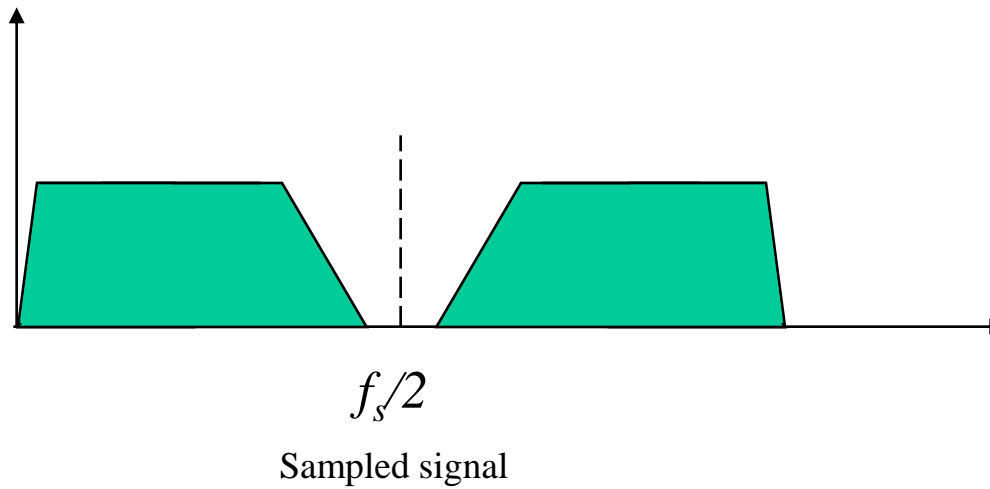
Sampled signal



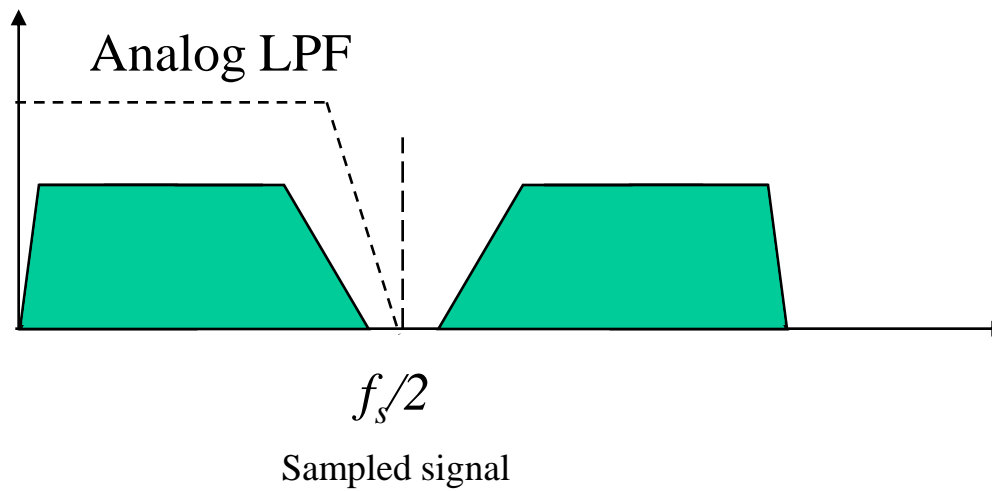
Sampling: Frequency domain



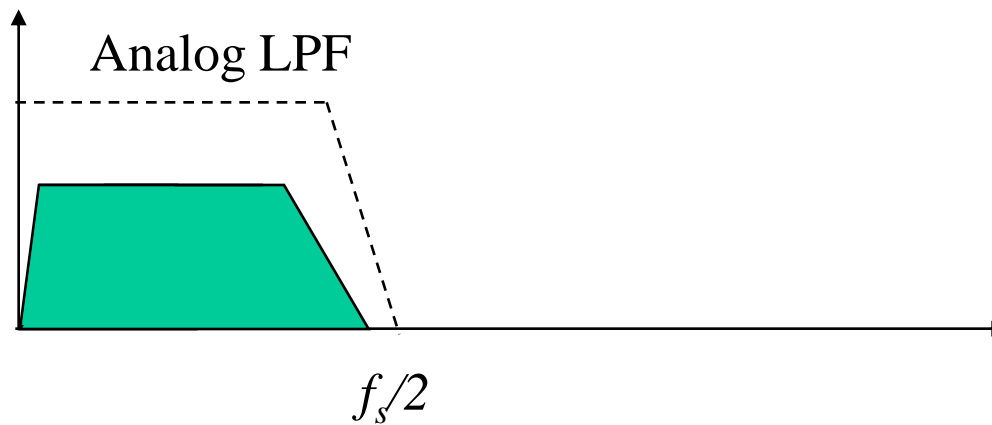
Reconstruction



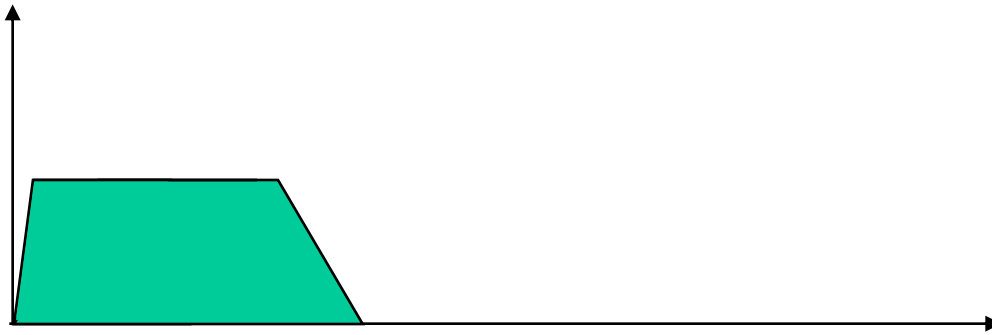
Reconstruction



Reconstruction



Reconstruction



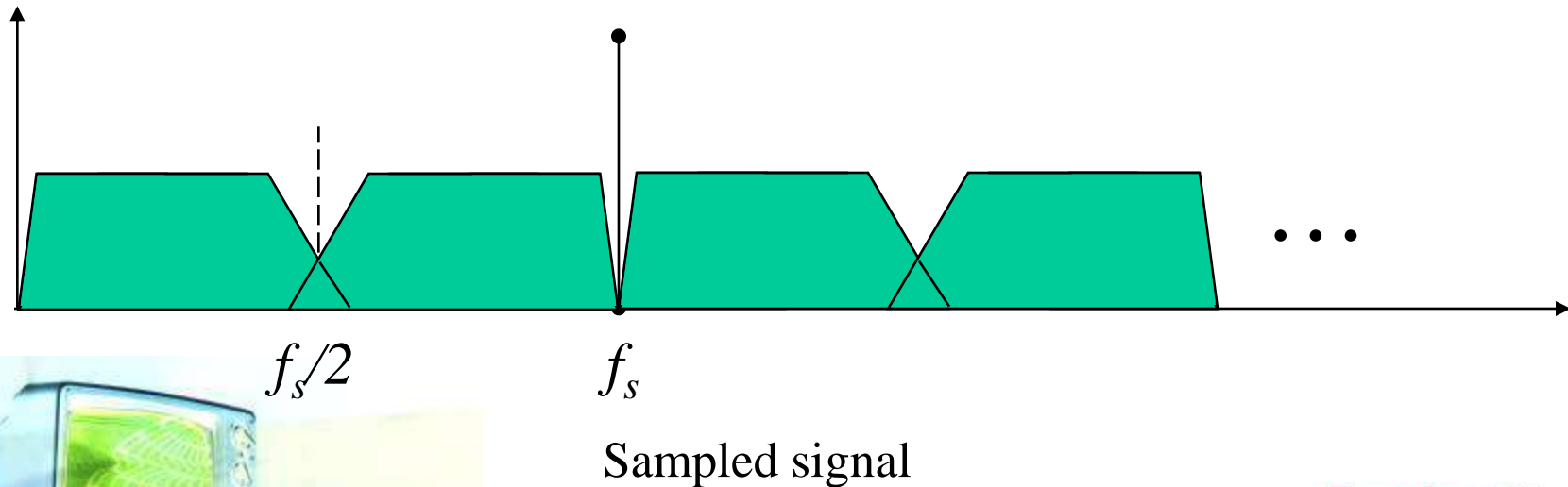
Analog signal



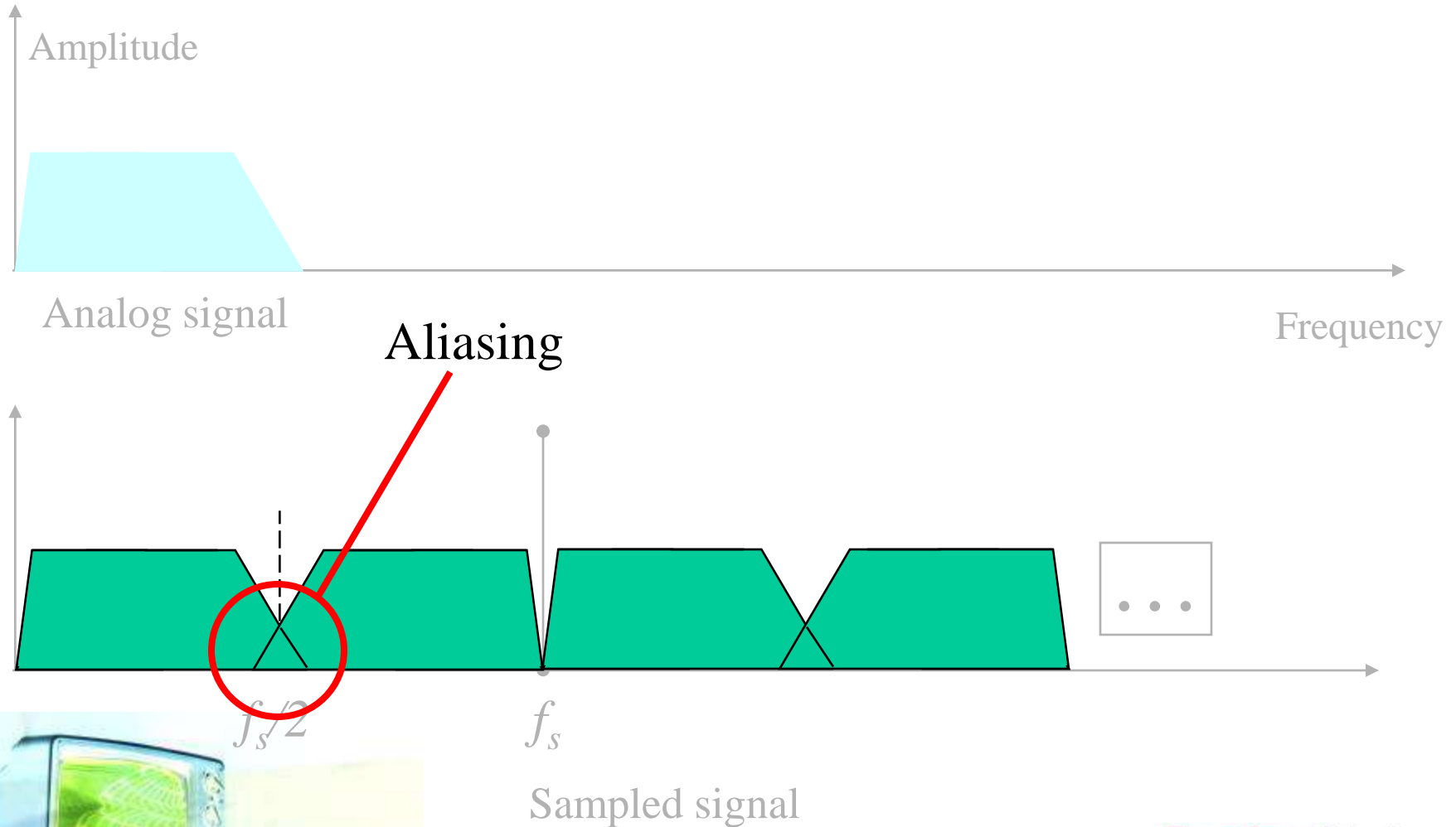
Aliasing: Frequency domain



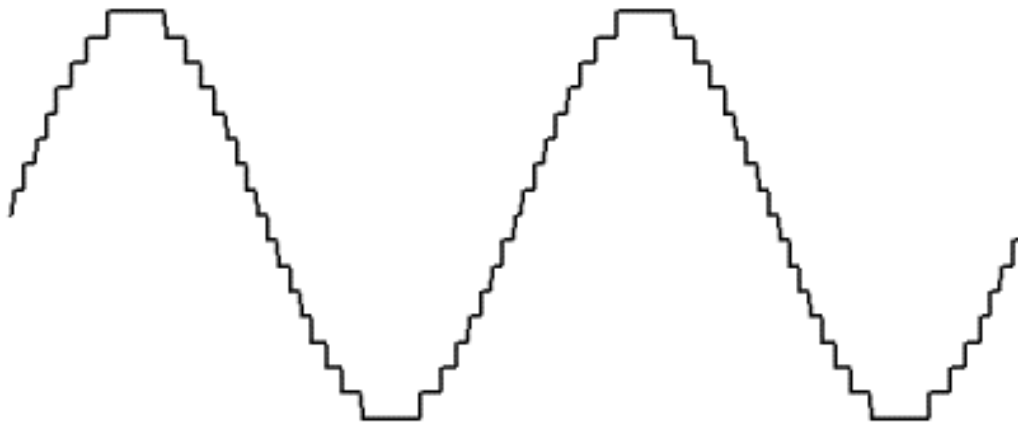
Aliasing: Frequency domain



Aliasing: Frequency domain



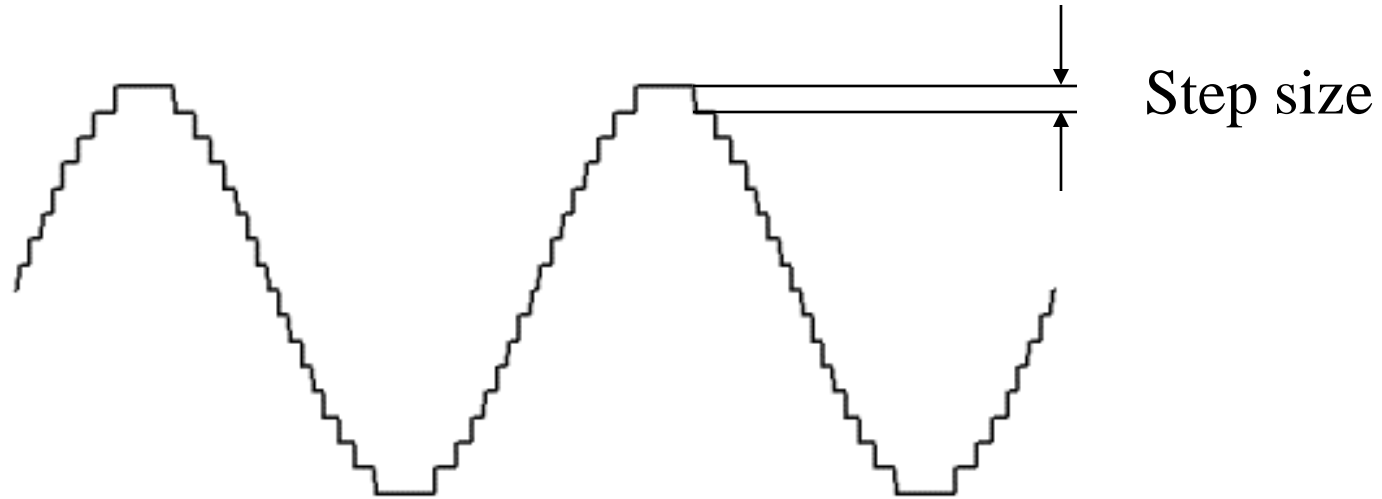
Quantization



Quantized sine wave



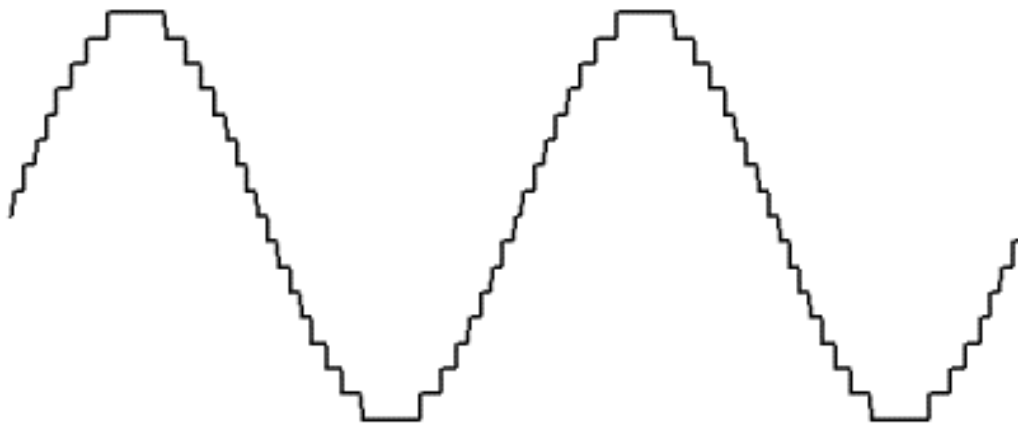
Quantization



- 16 steps total = 2^4
- 4 bits are needed to code signal



Quantization



Quantization error

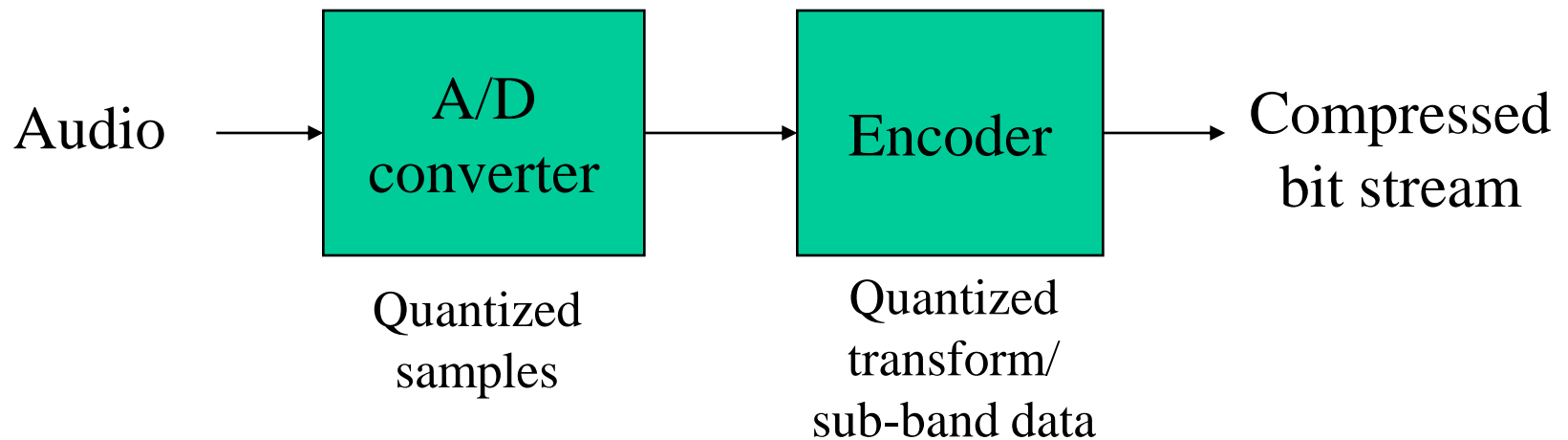


Quantization noise

- Depends on signal distribution, # of bits (n)
- $\text{SNR}_{\text{sine}} \approx (1.76 + 6.02 n) \text{ dB}$
- $\text{SNR}_{\text{video}} \approx (10.8 + 6.02 n) \text{ dB}$



Quantization



- Audio can be quantized in several places in signal processors.

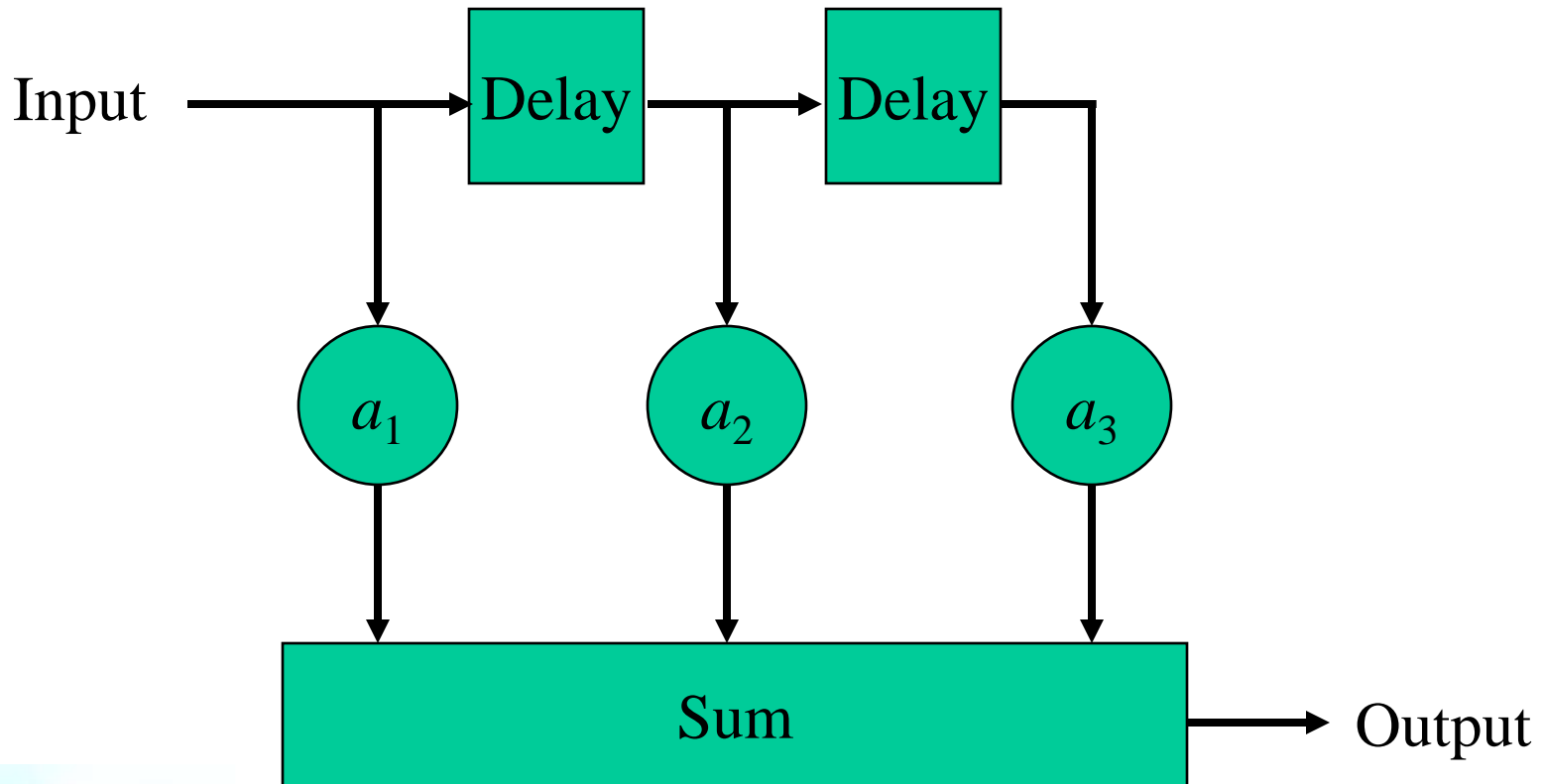


FIR filter

- Finite Impulse Response filter
- Filter is stable
- Output settles to zero



FIR filter: “tapped delay line”



Filter example

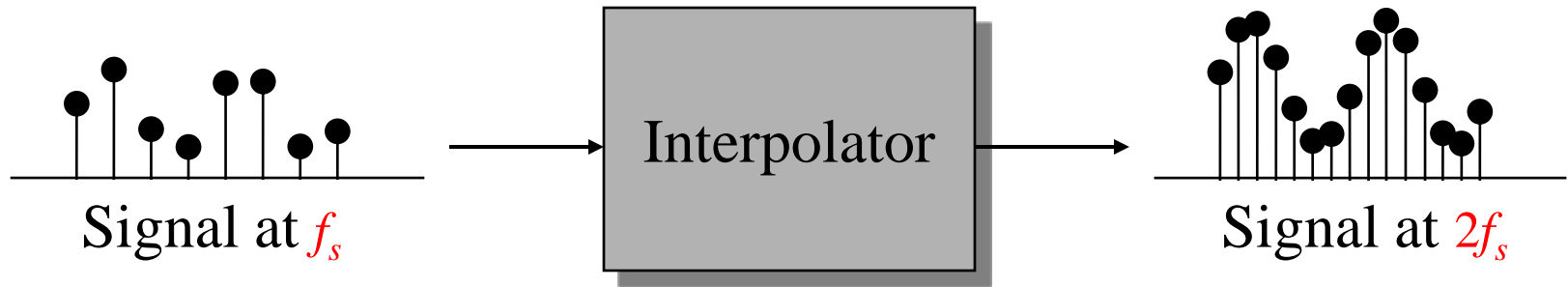
- Low-pass filter
- $y[n] = \frac{1}{4} x[n-1] + \frac{1}{2} x[n] + \frac{1}{4} x[n+1]$
- “ $\frac{1}{4}$ - $\frac{1}{2}$ - $\frac{1}{4}$ ” filter
- Moving average filter



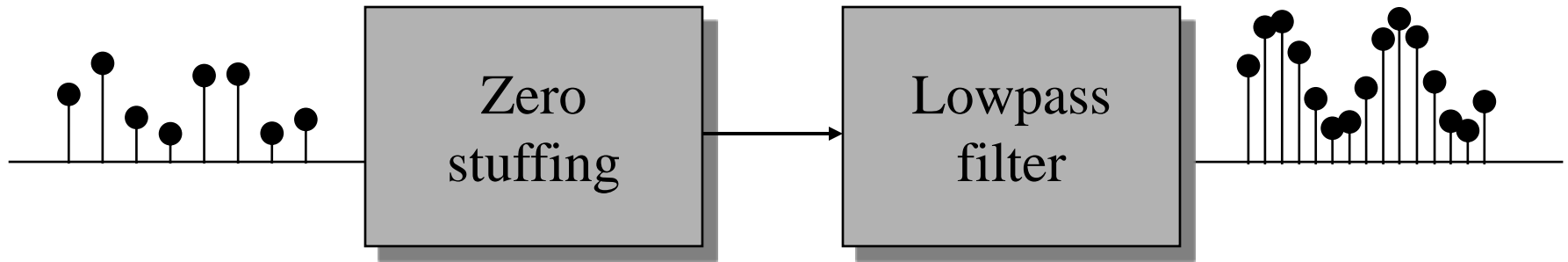
Sample rate conversion



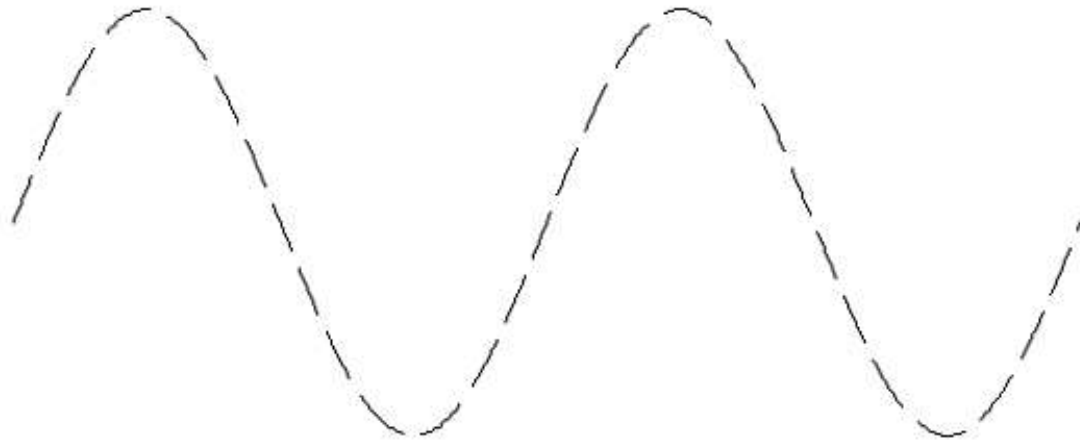
Upsampling



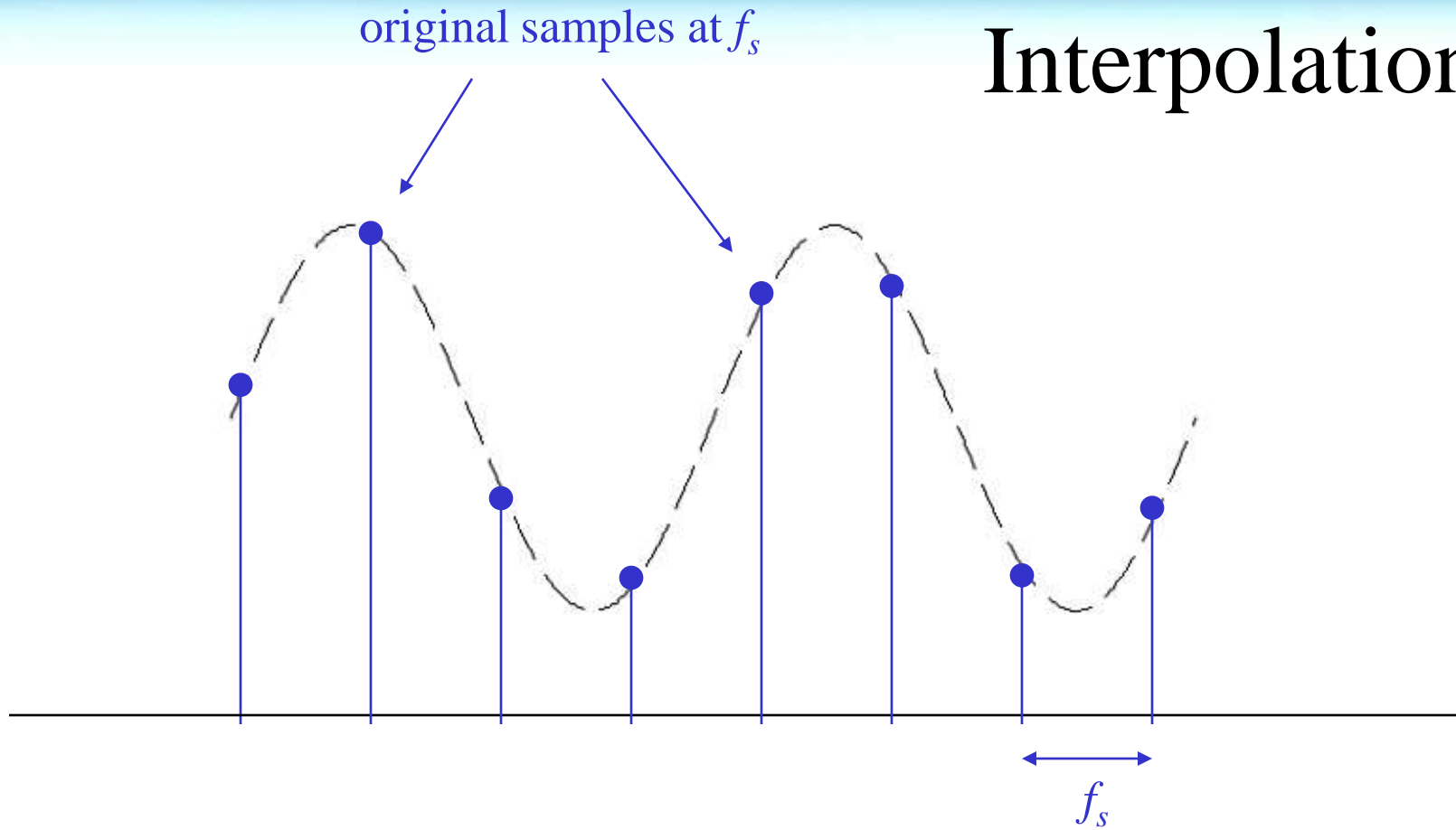
Interpolation (upsampling)



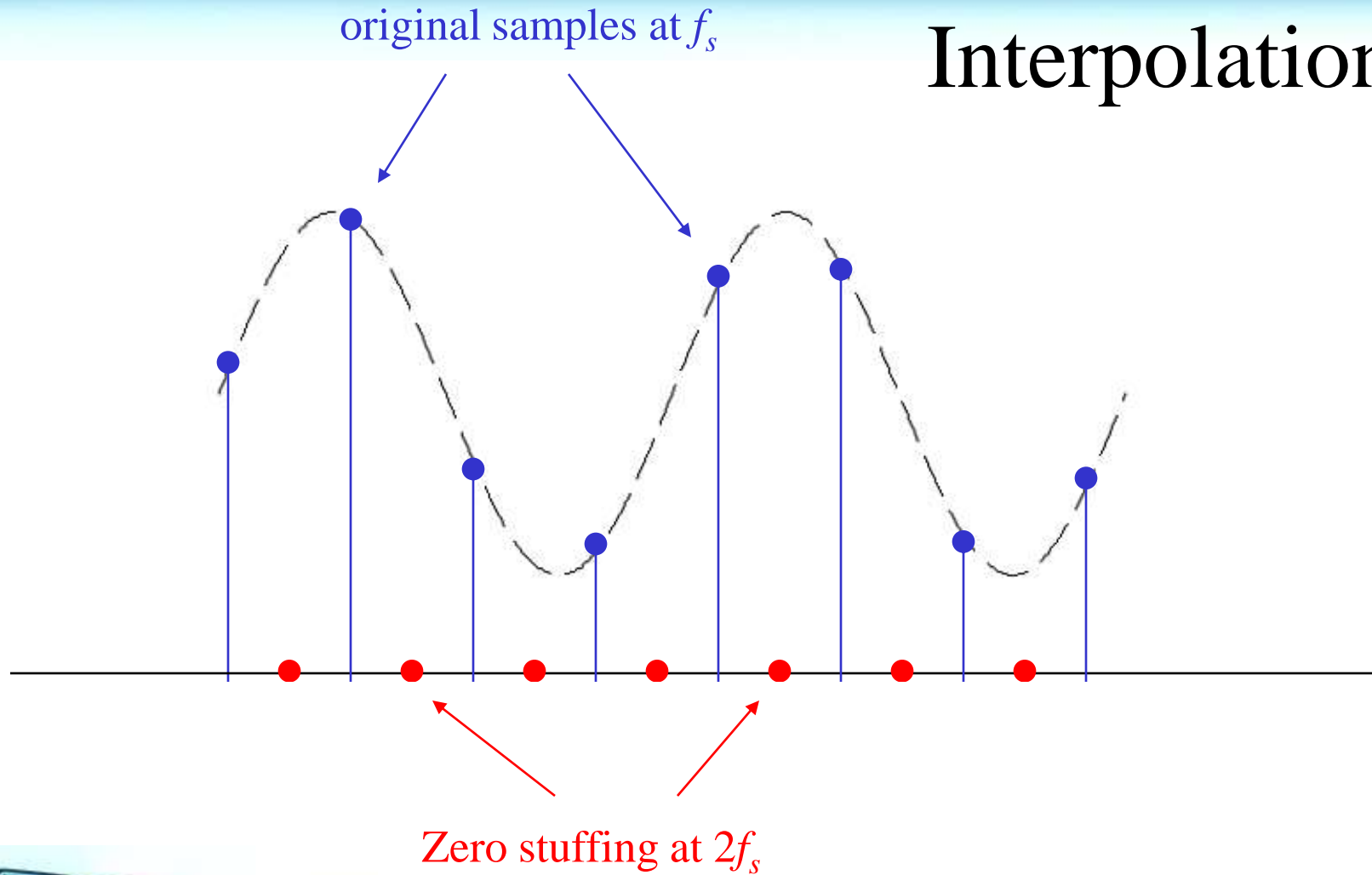
Interpolation (upsampling)



Interpolation

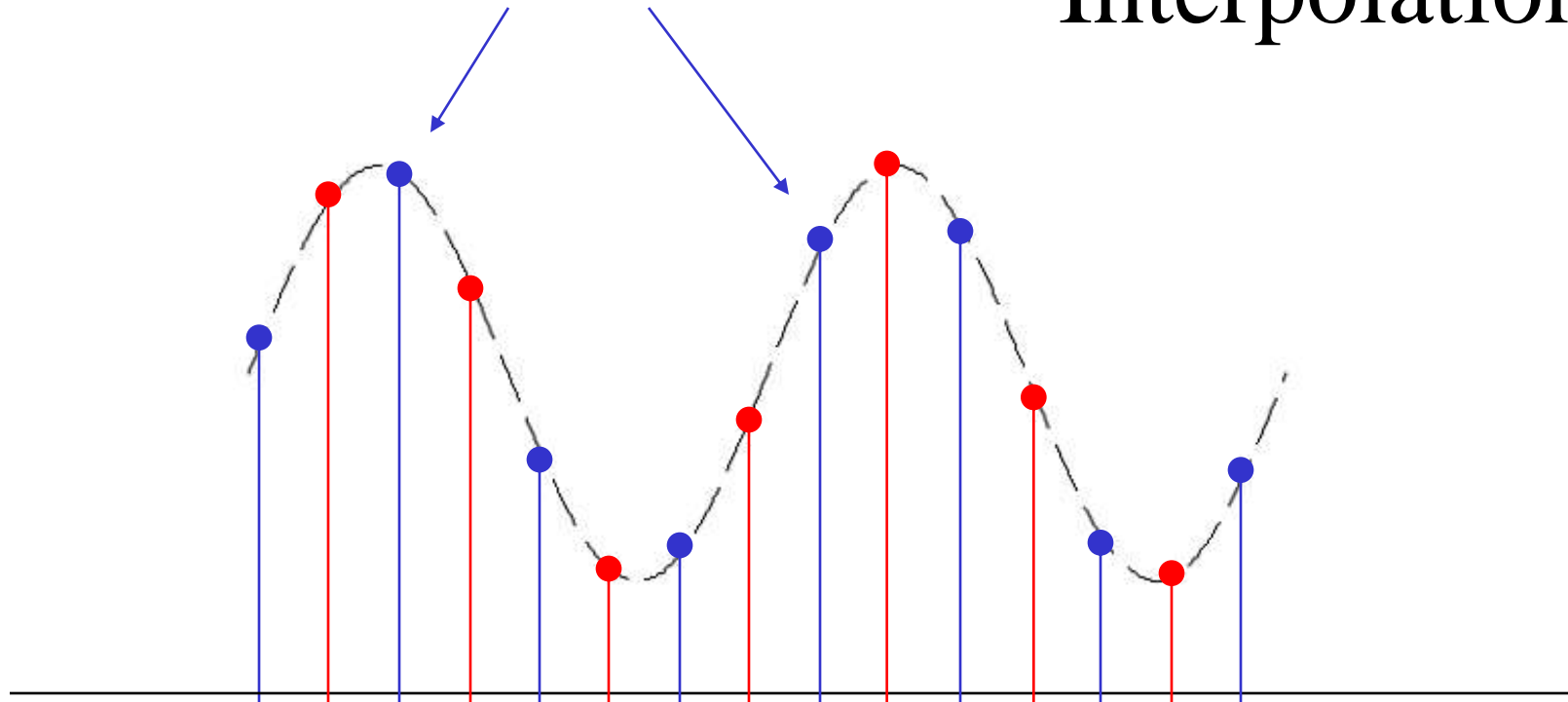


Interpolation



Interpolation

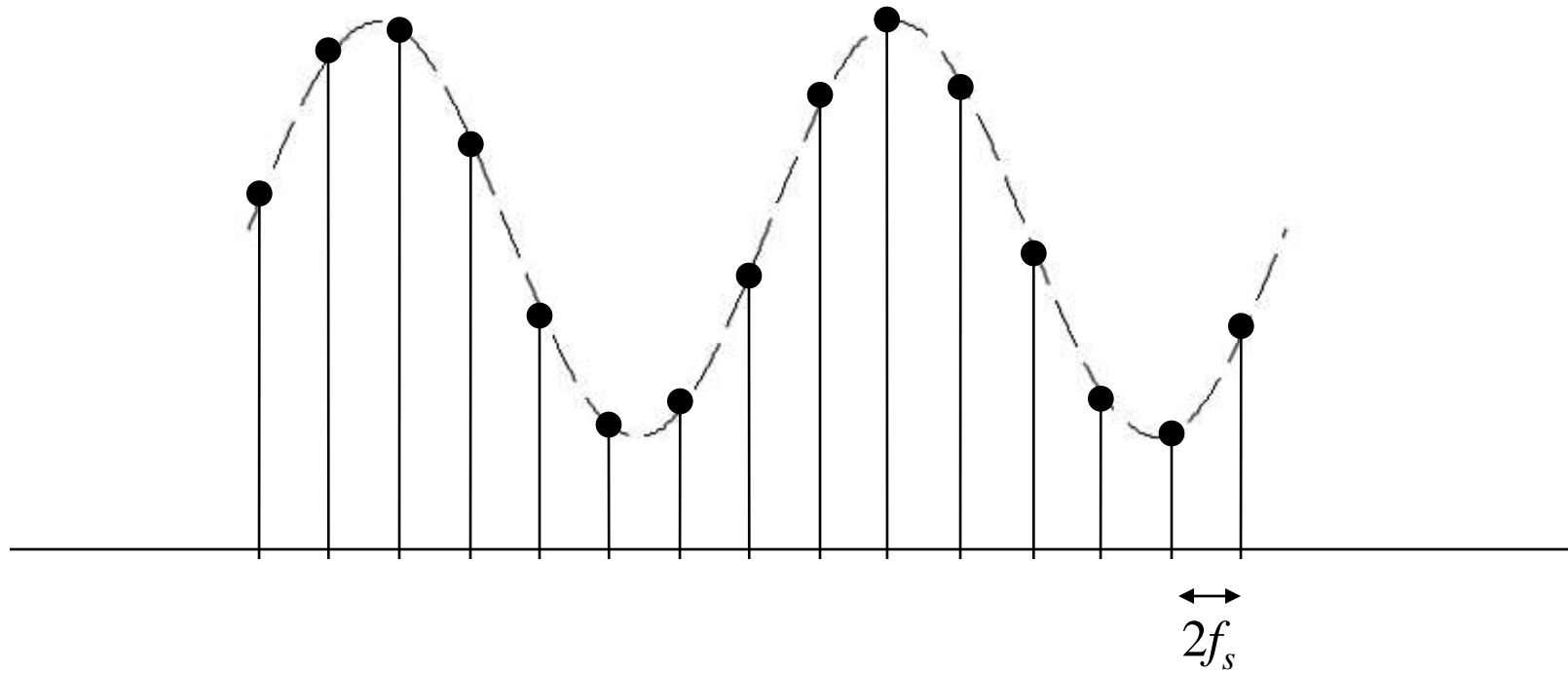
original samples at f_s



After lowpass filtering at $2f_s/2$



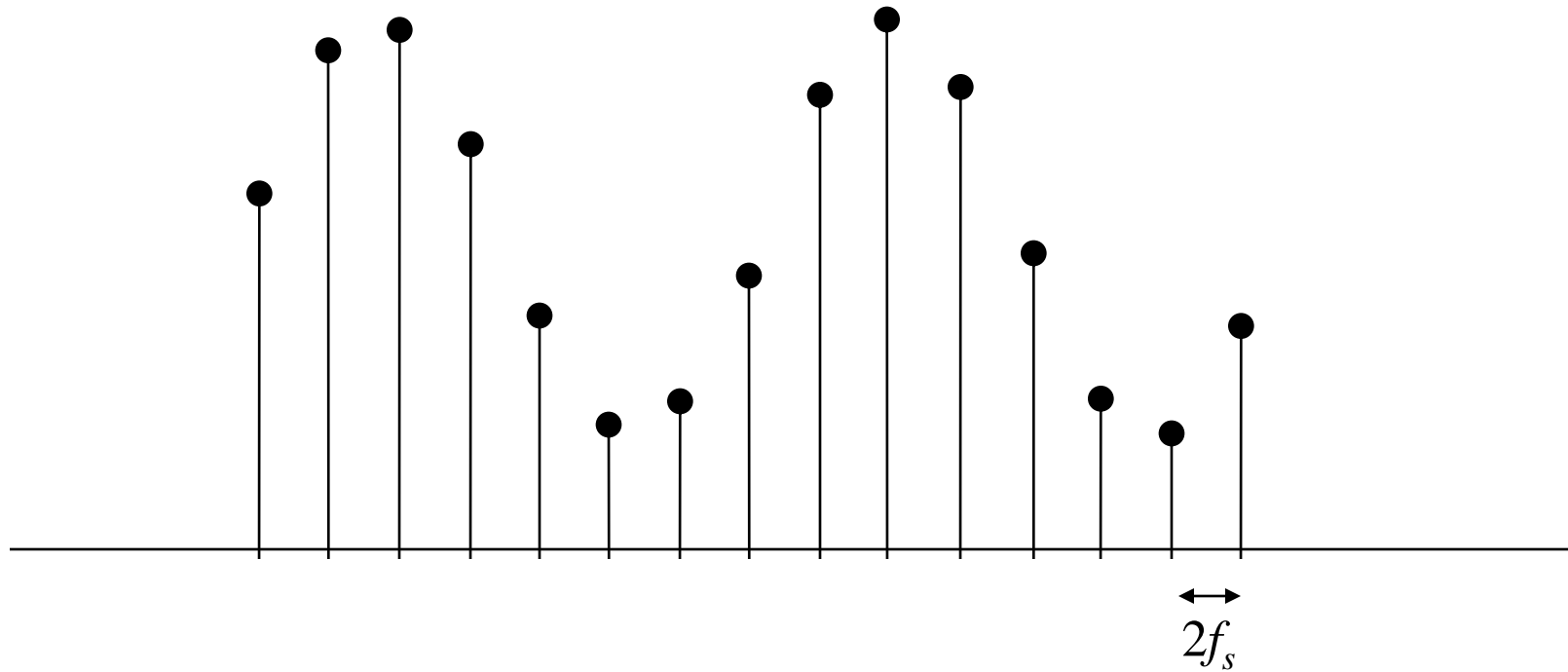
Interpolation



Signal at new sample rate = $2f_s$



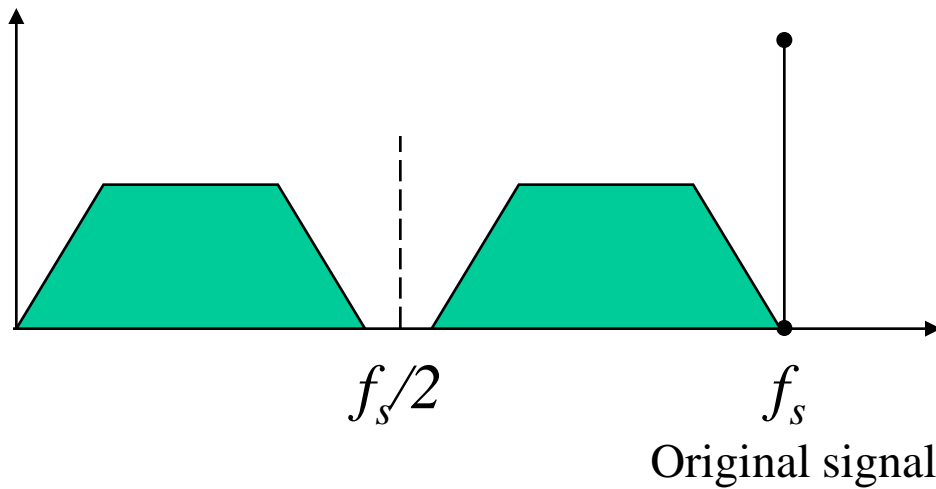
Interpolation



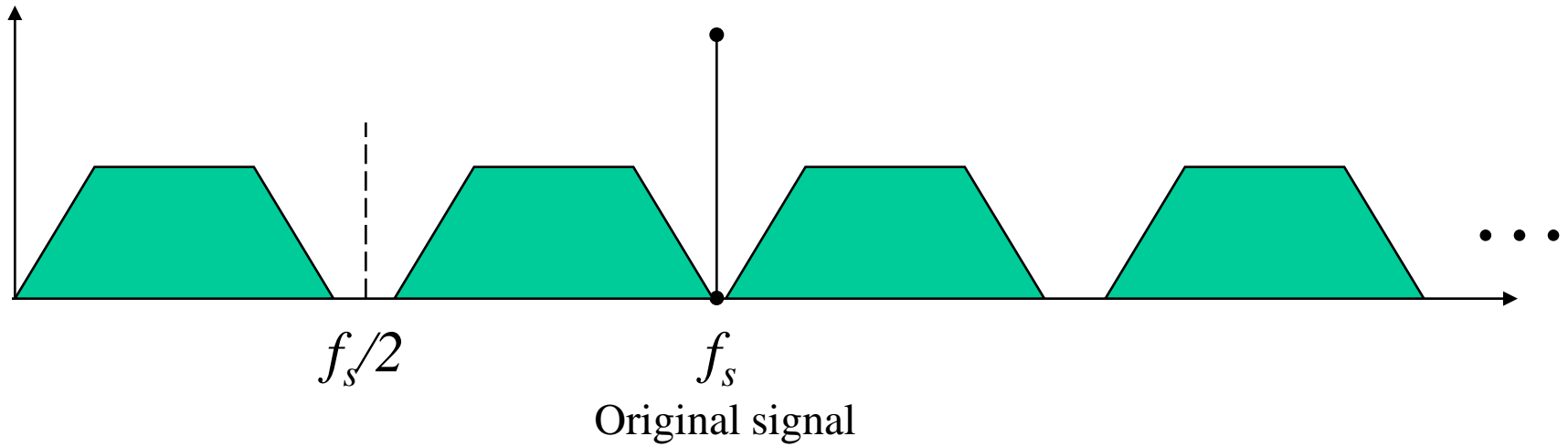
Signal at new sample rate = $2f_s$



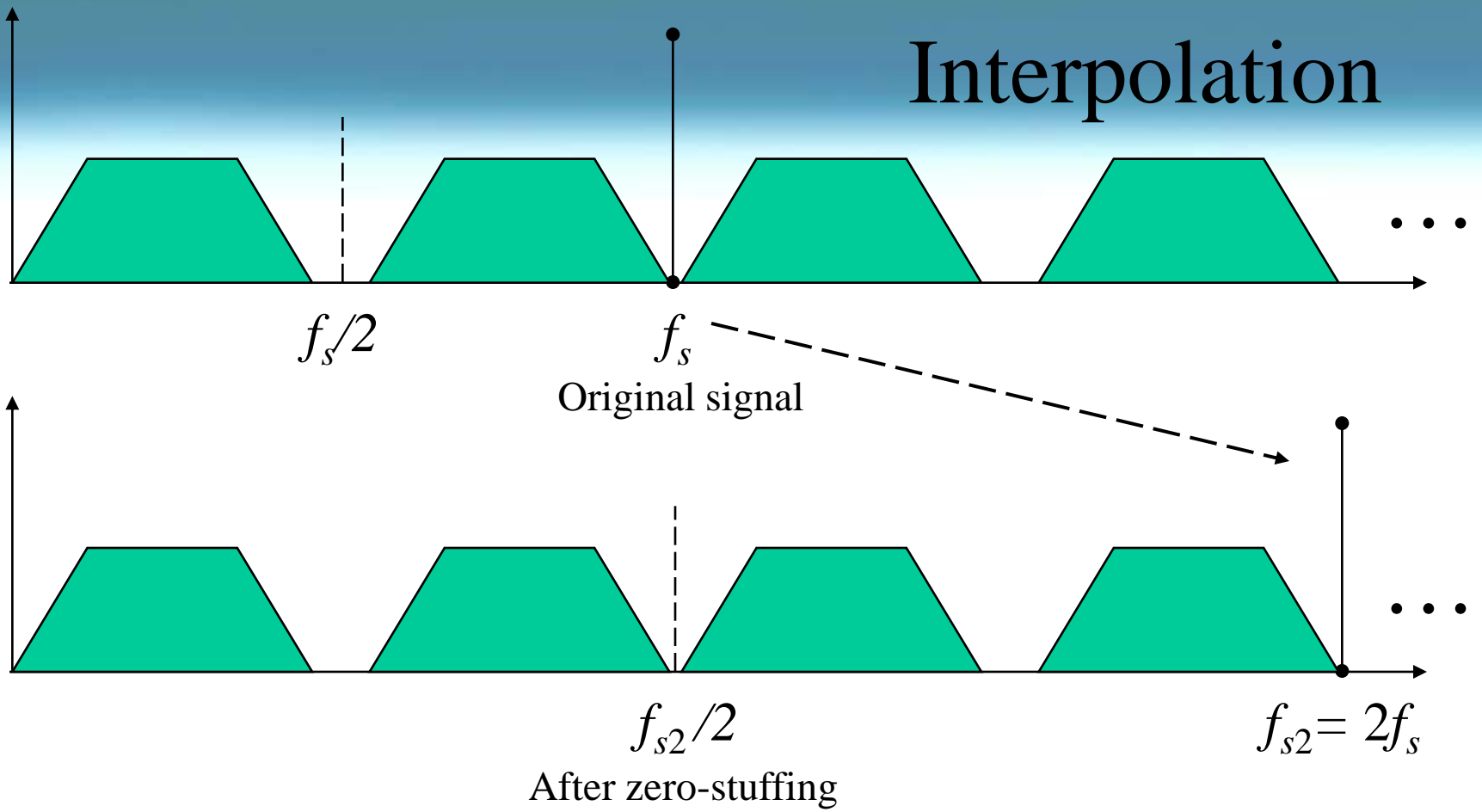
Interpolation



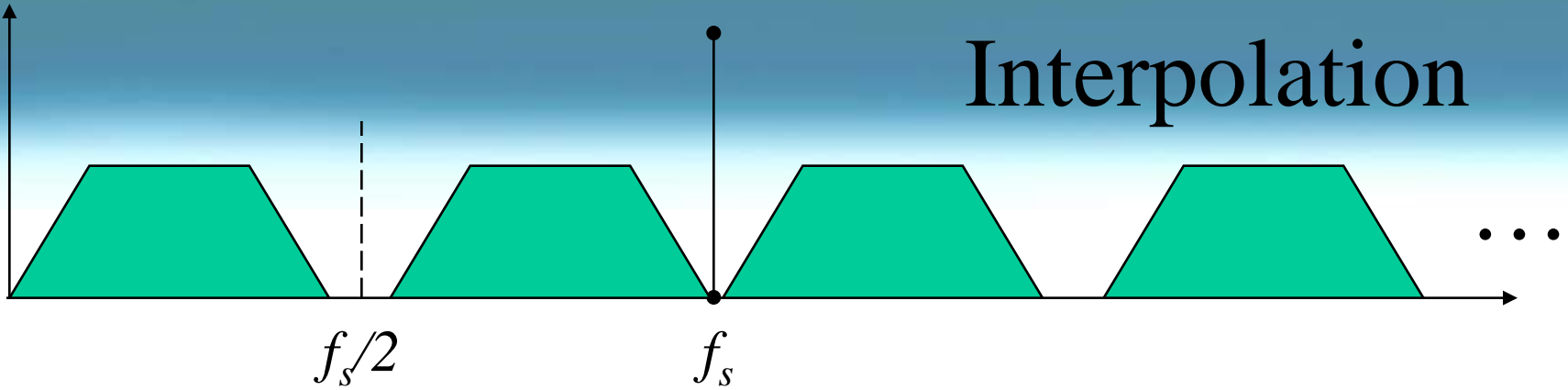
Interpolation



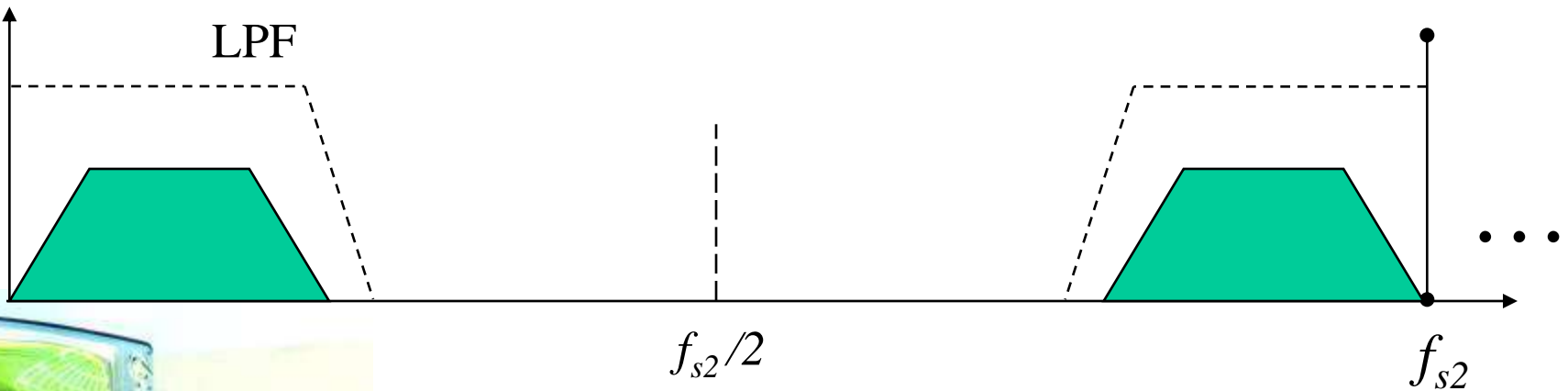
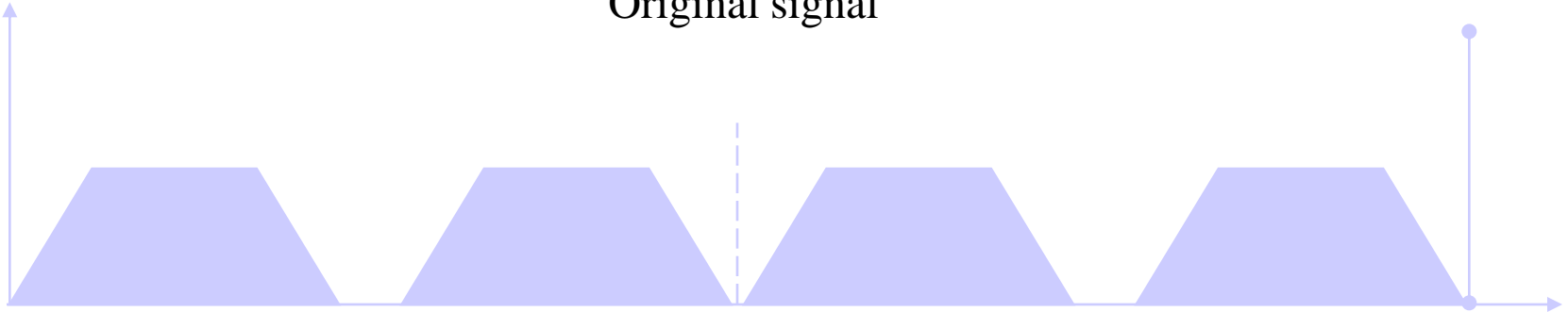
Interpolation



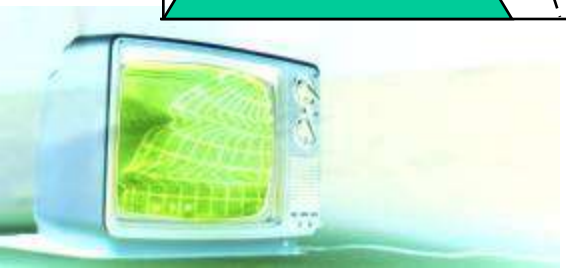
Interpolation



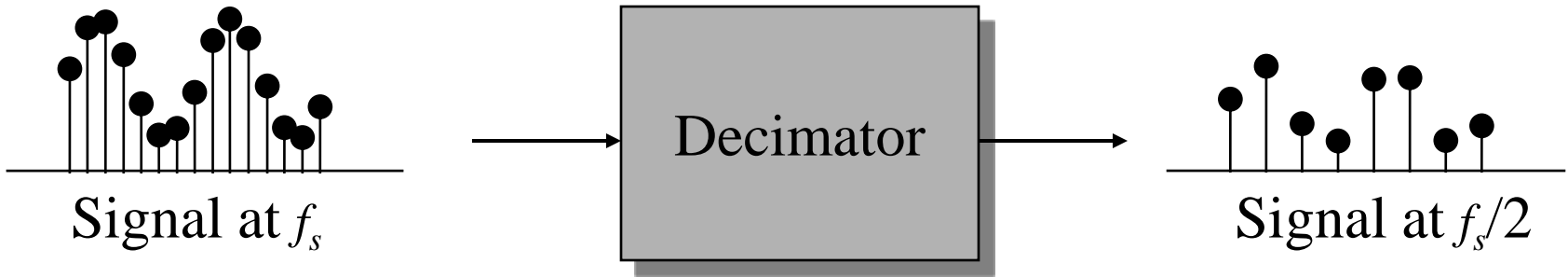
Original signal



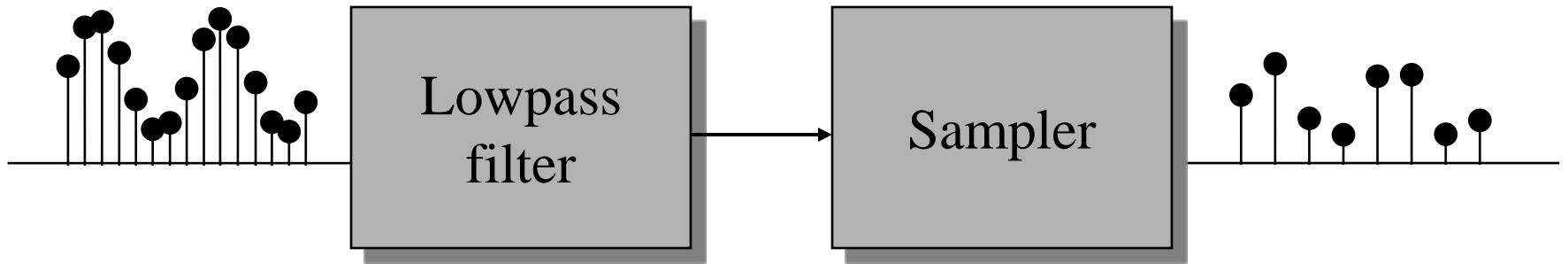
Final signal after digital LPF



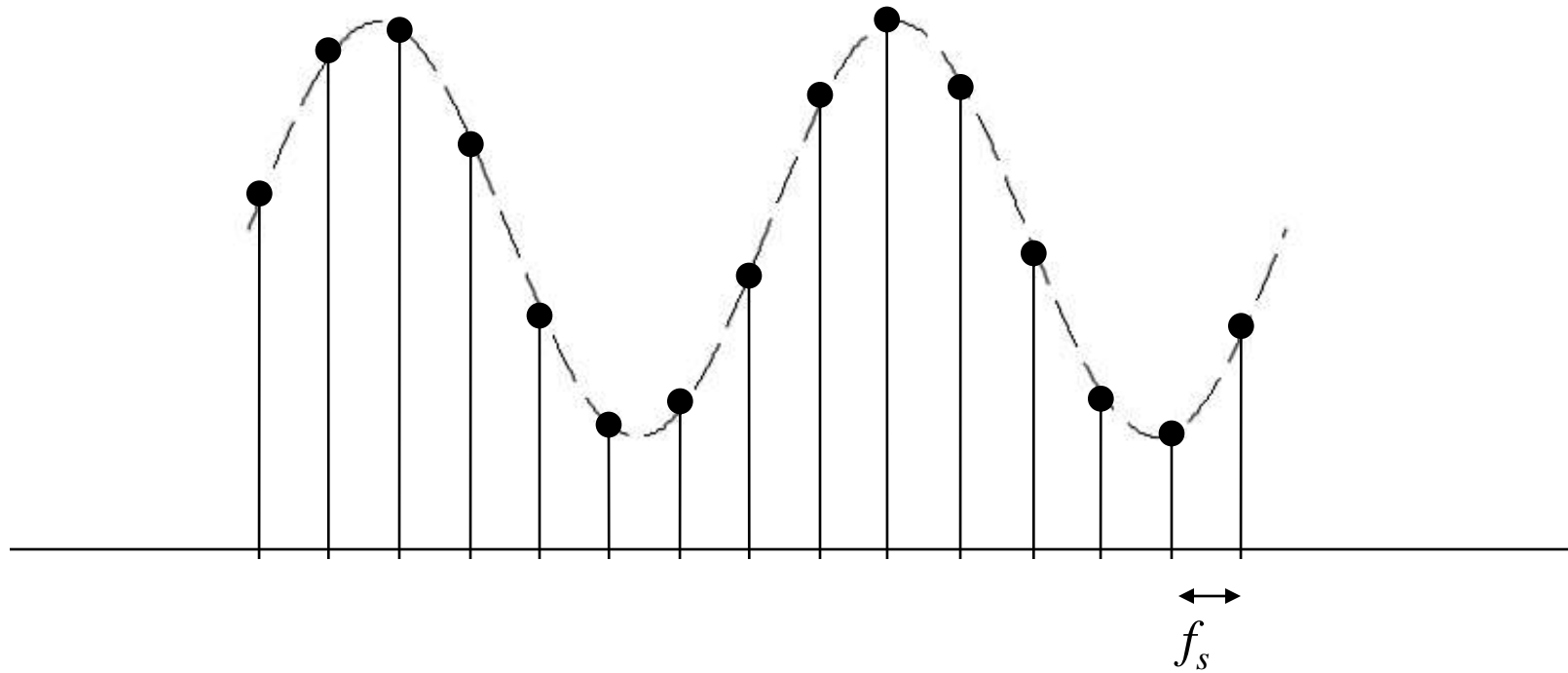
Downsampling



Decimation (downsampling)



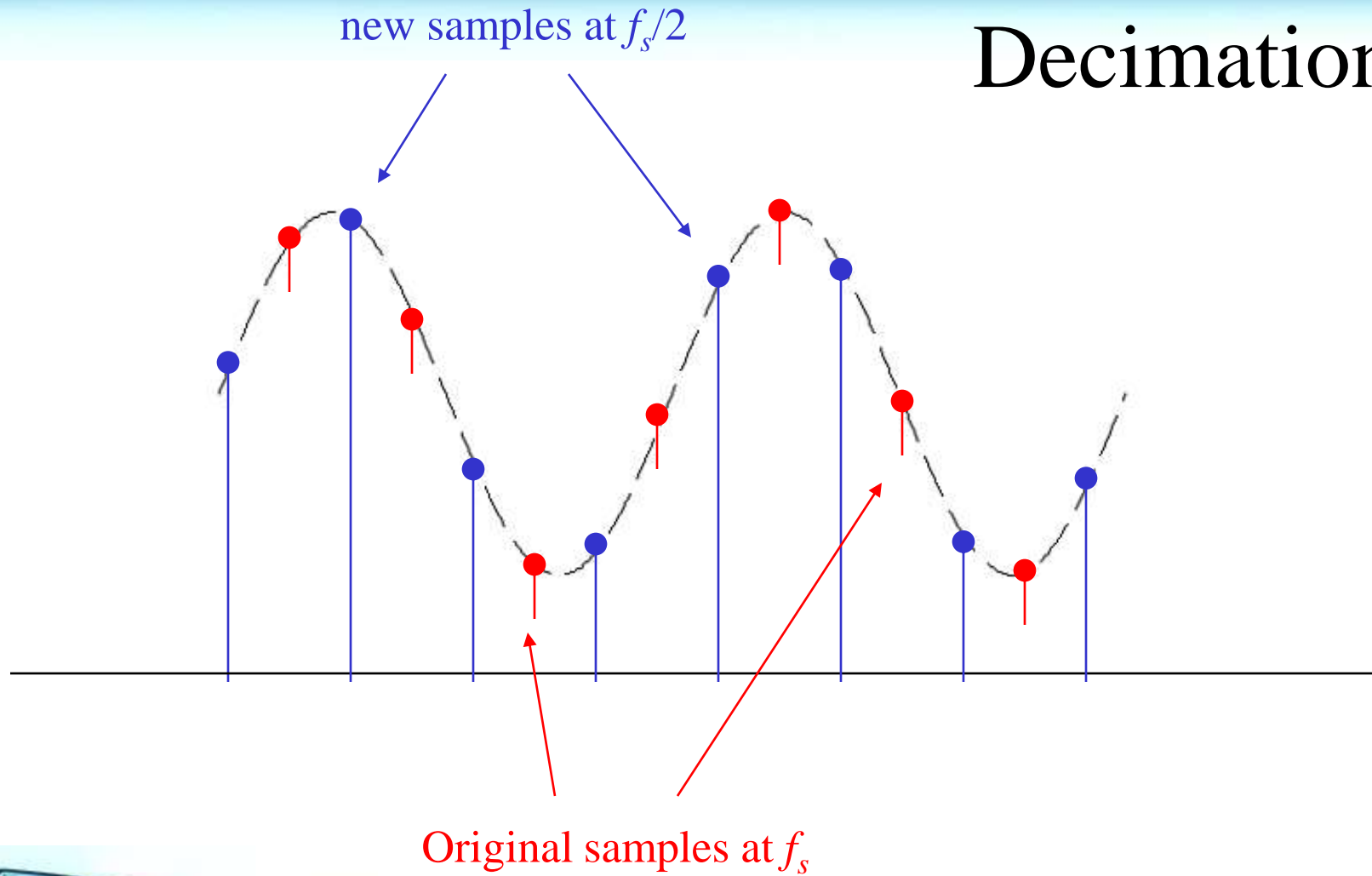
Decimation



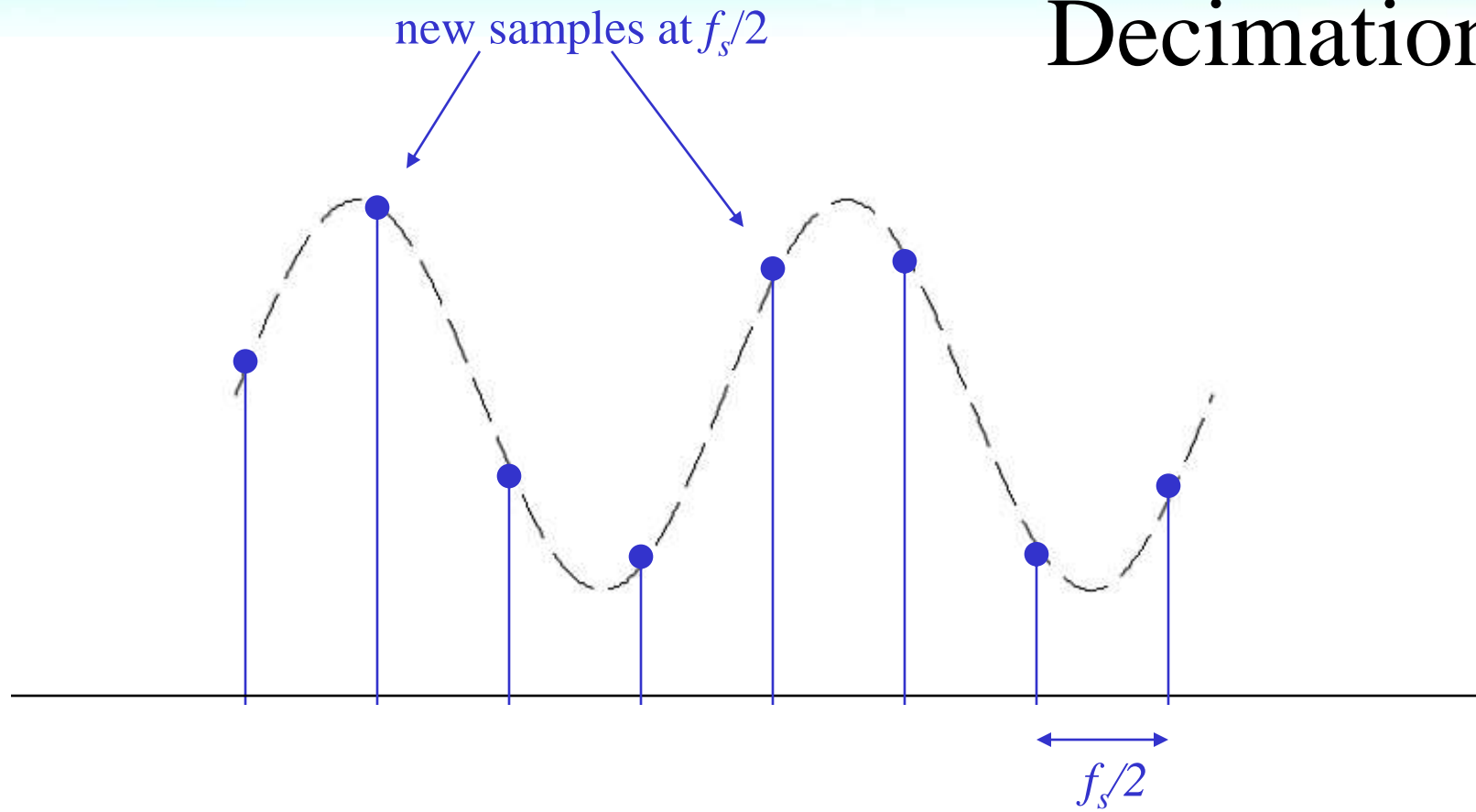
Signal at sample rate = f_s

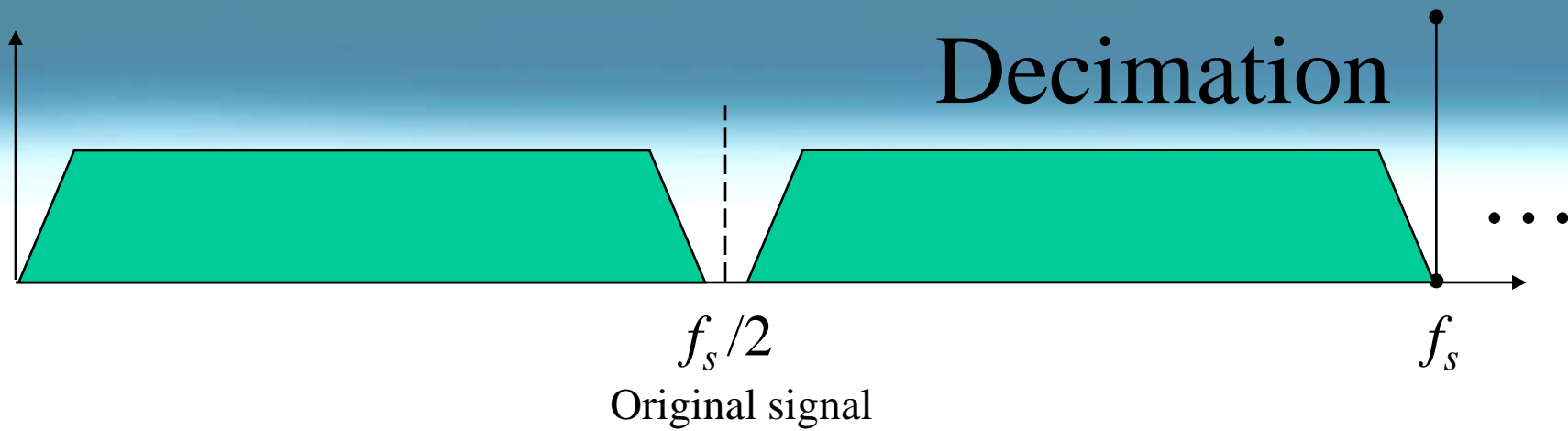


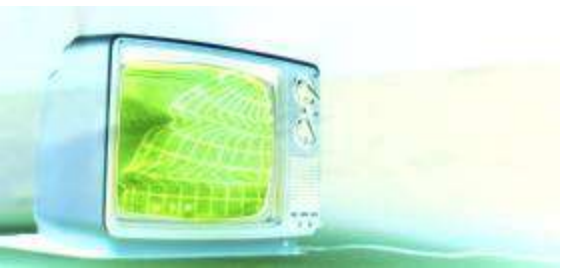
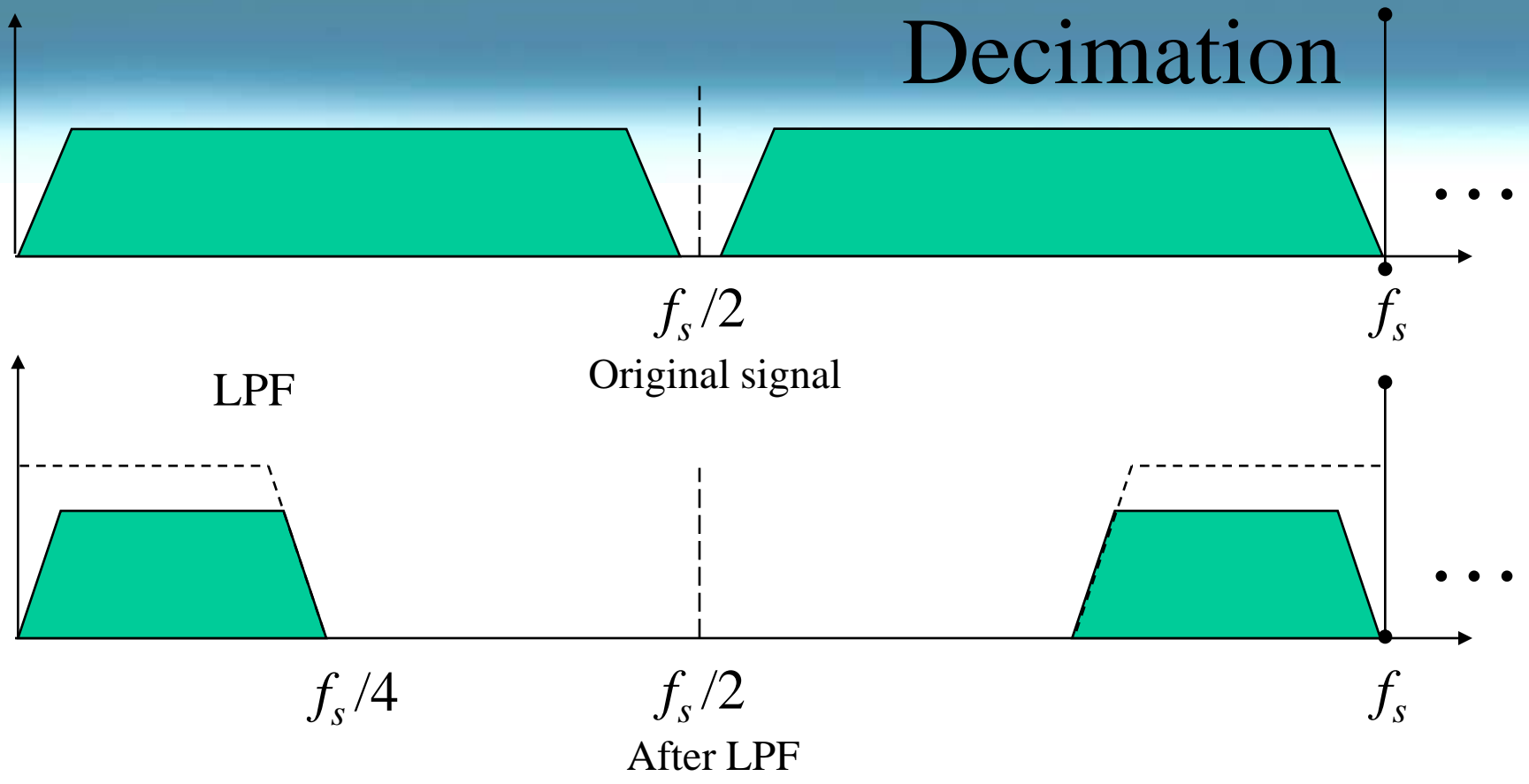
Decimation



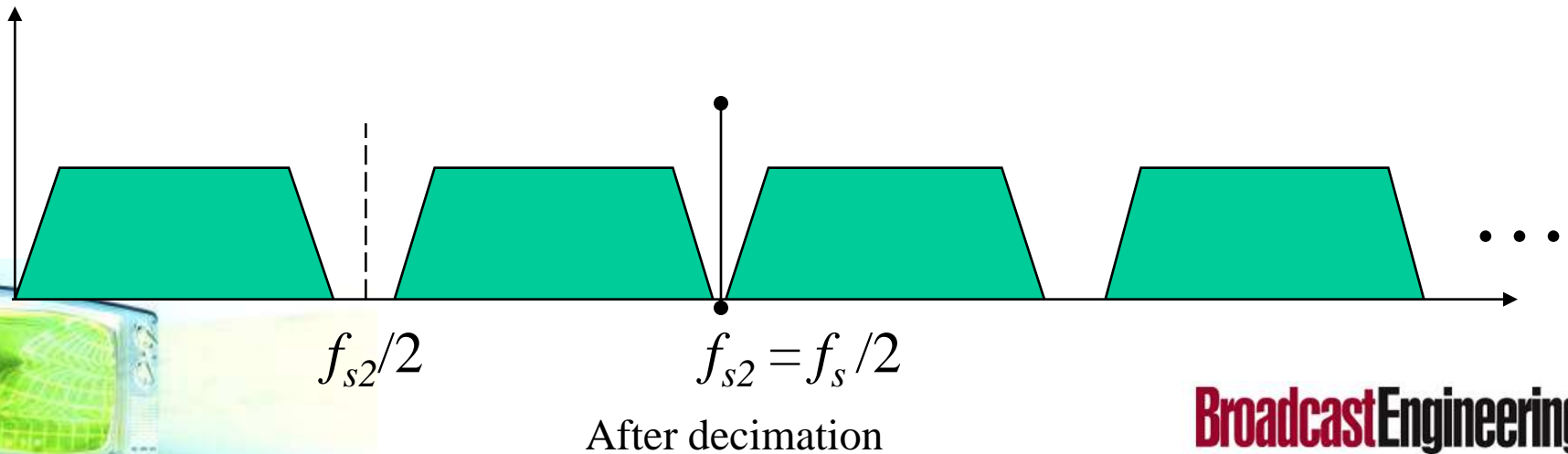
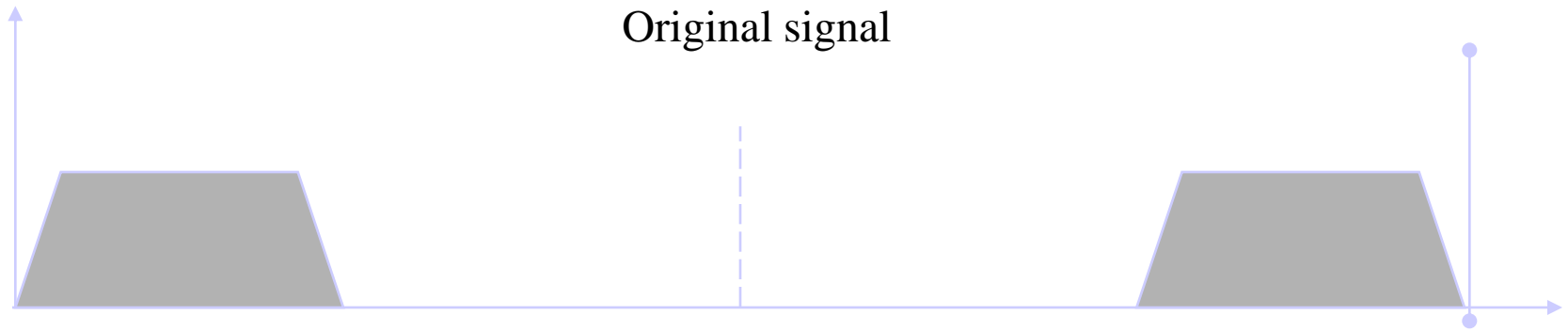
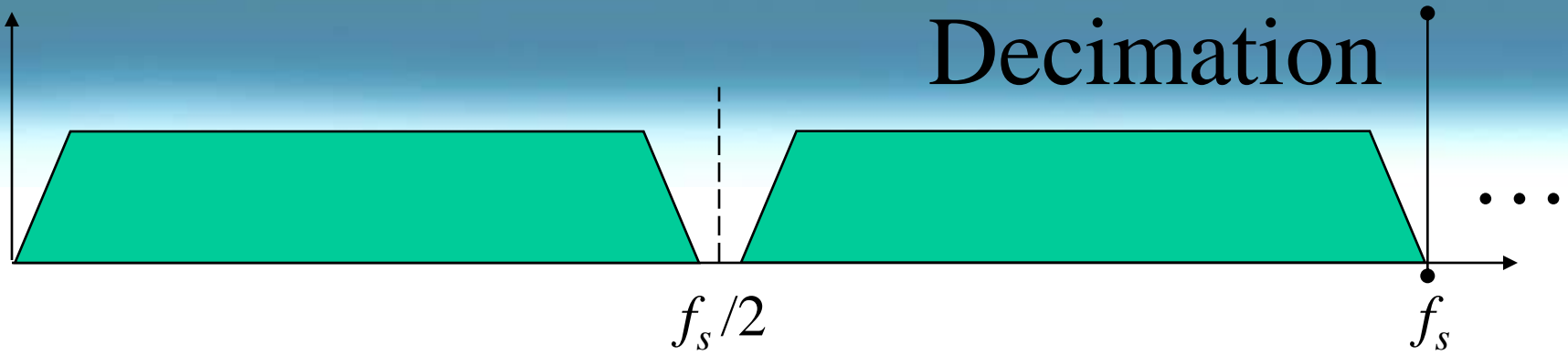
Decimation



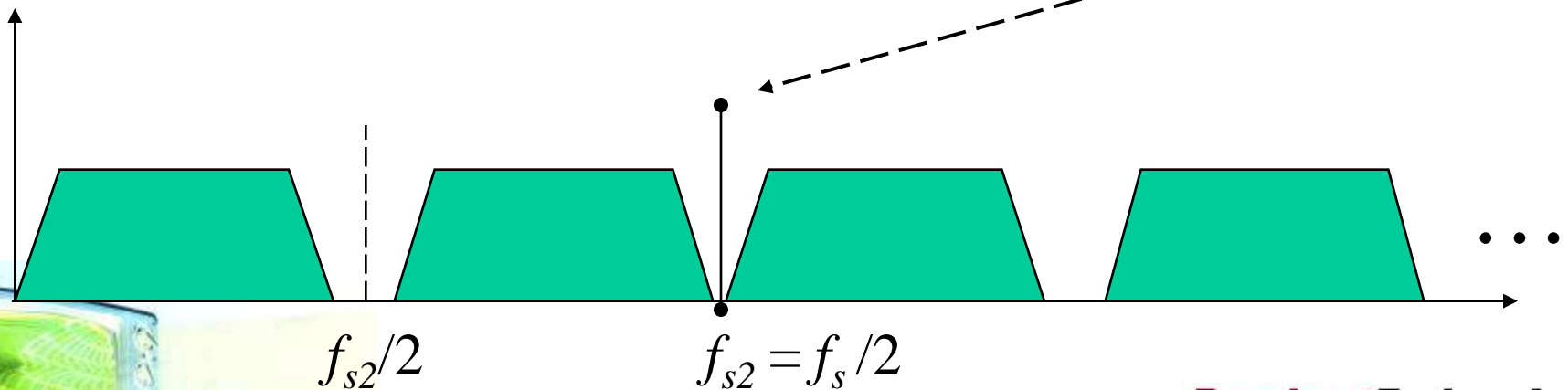
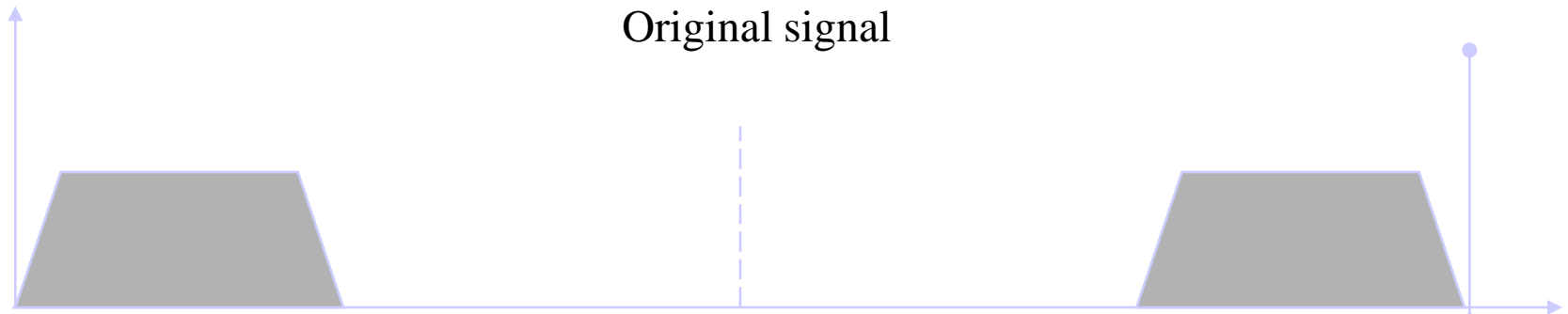
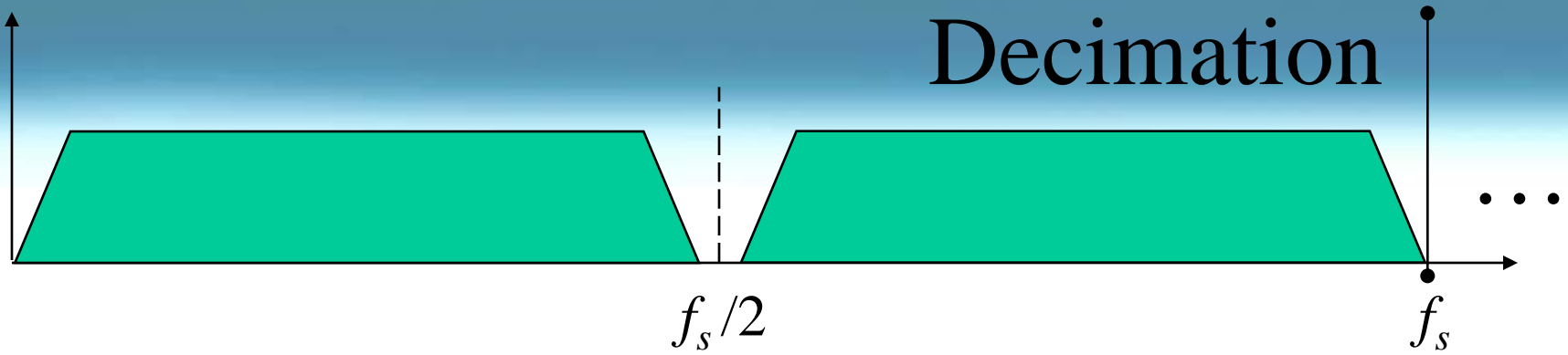




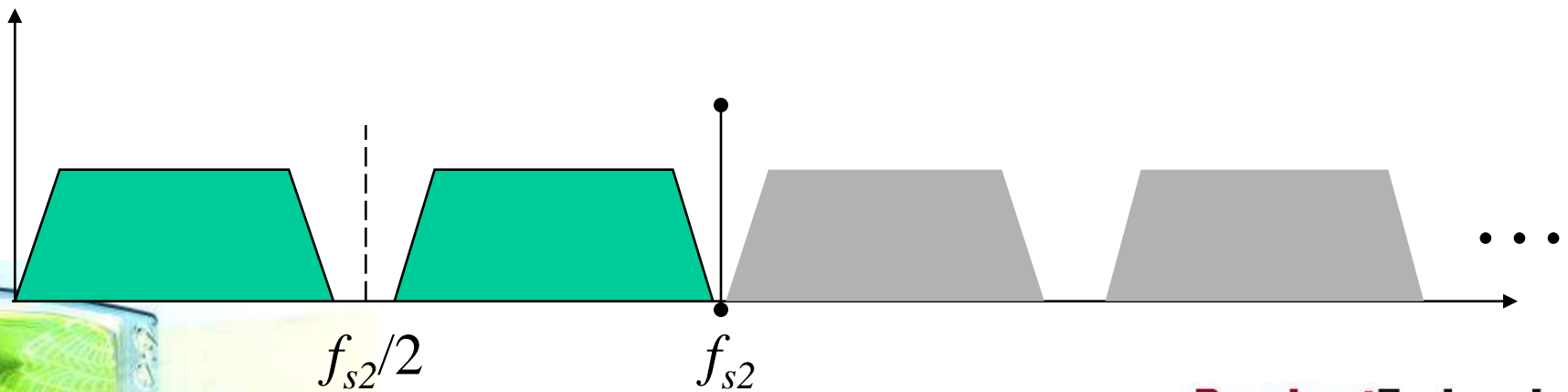
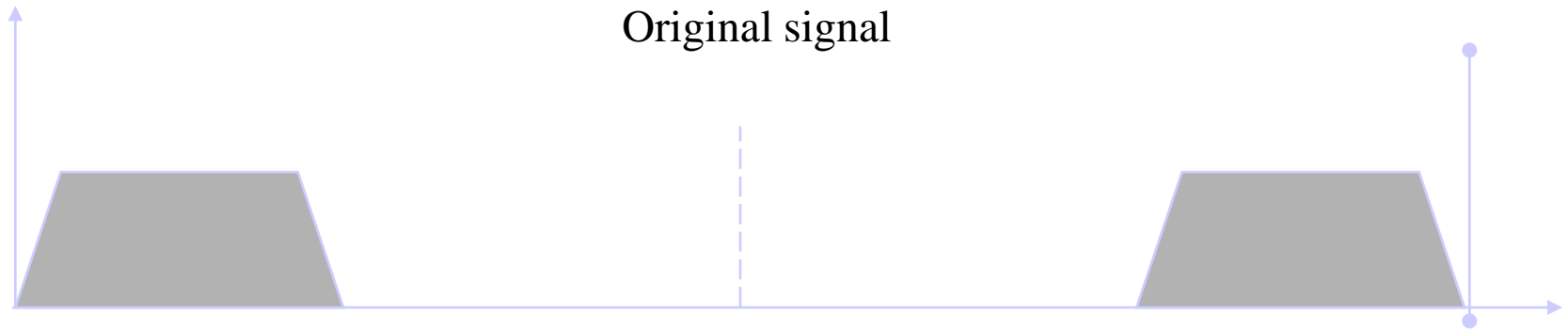
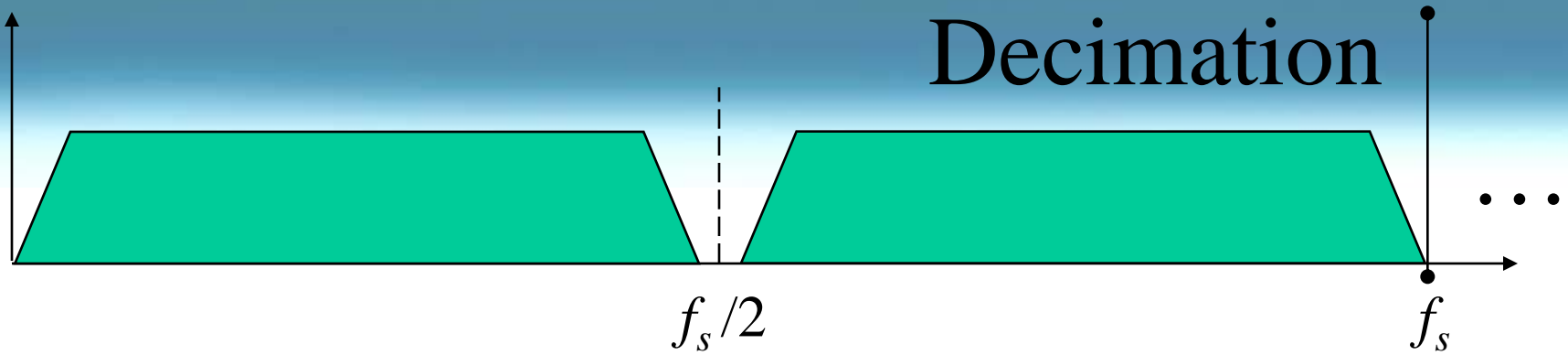
Decimation



Decimation



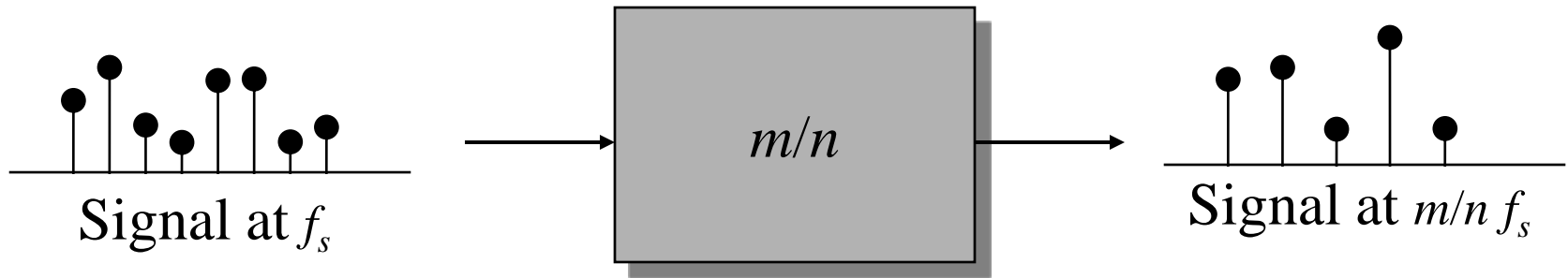
Decimation



After decimation



Fractional rate conversion

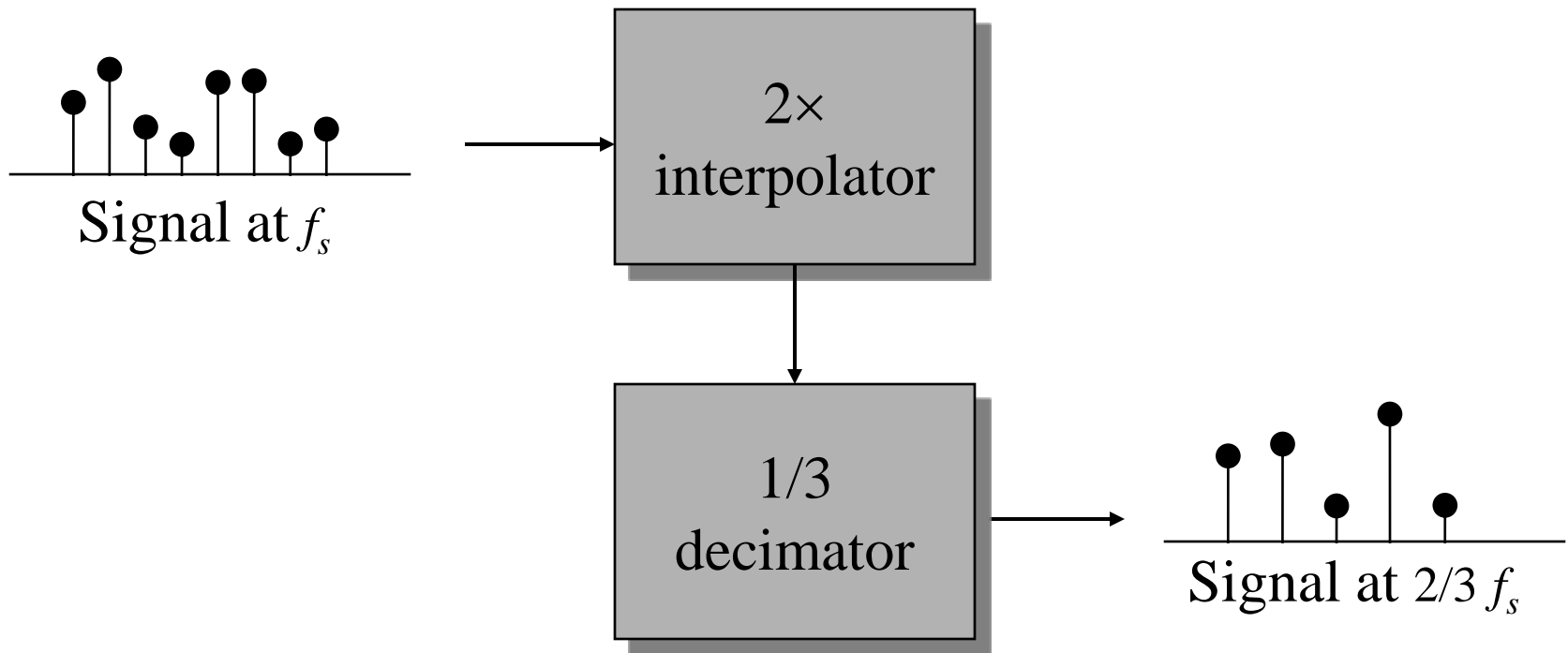


Fractional rate conversion

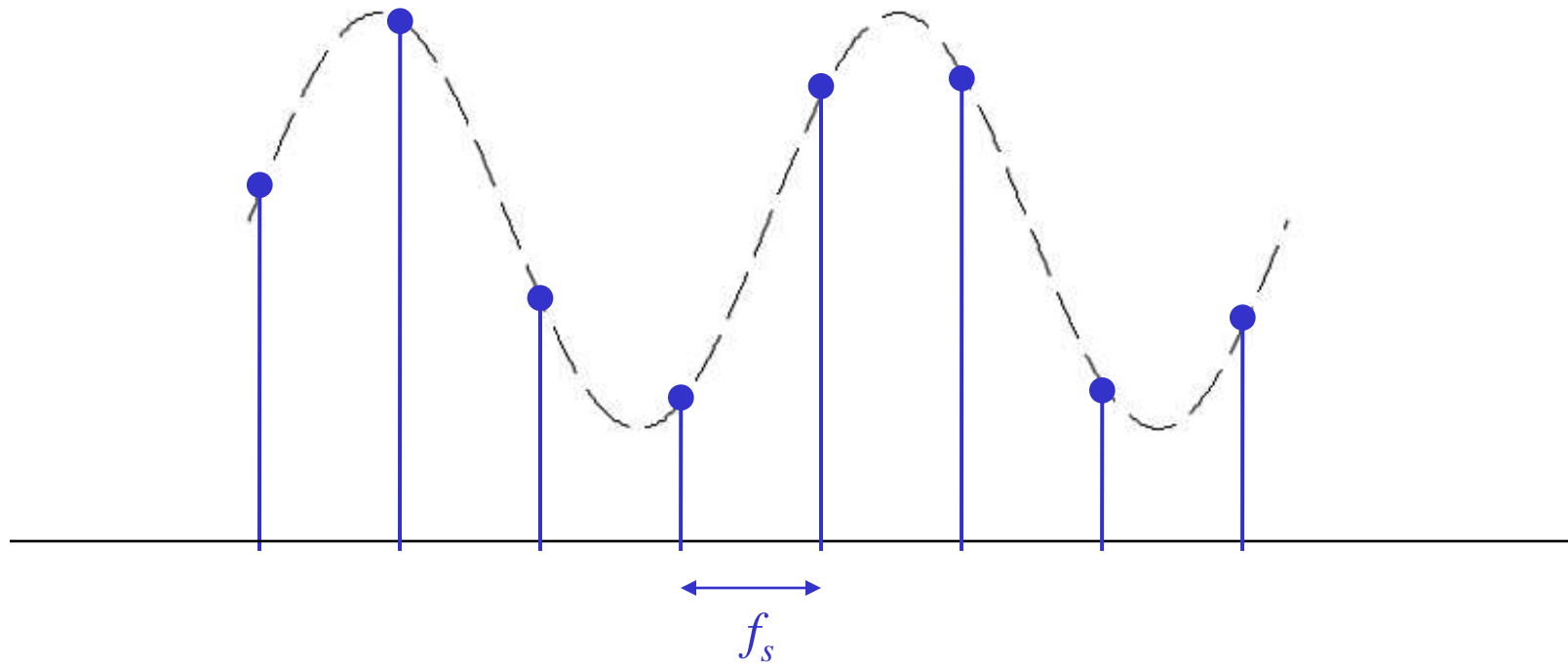
- Sample rates not integrally related
- Audio
 - 48kHz \rightarrow 44.1kHz
 - 147/160



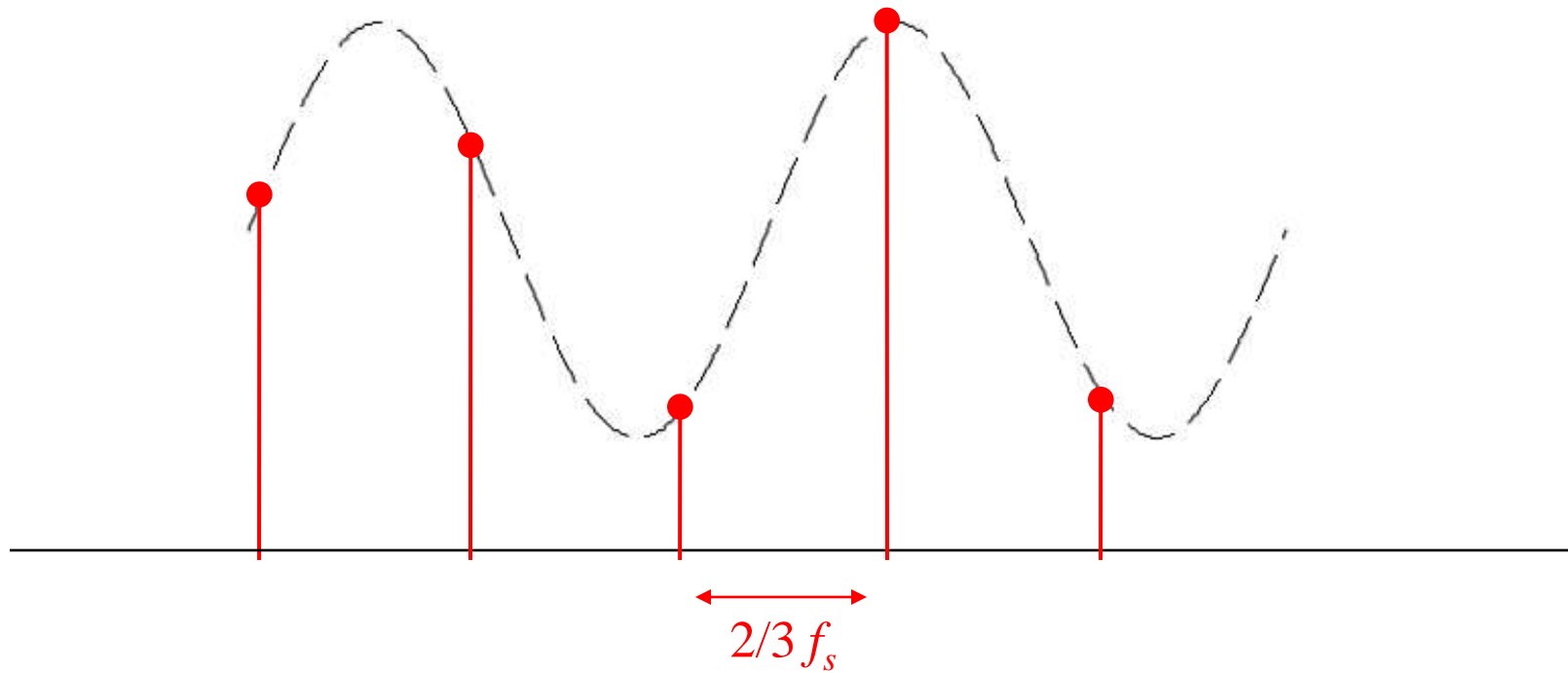
Fractional rate conversion



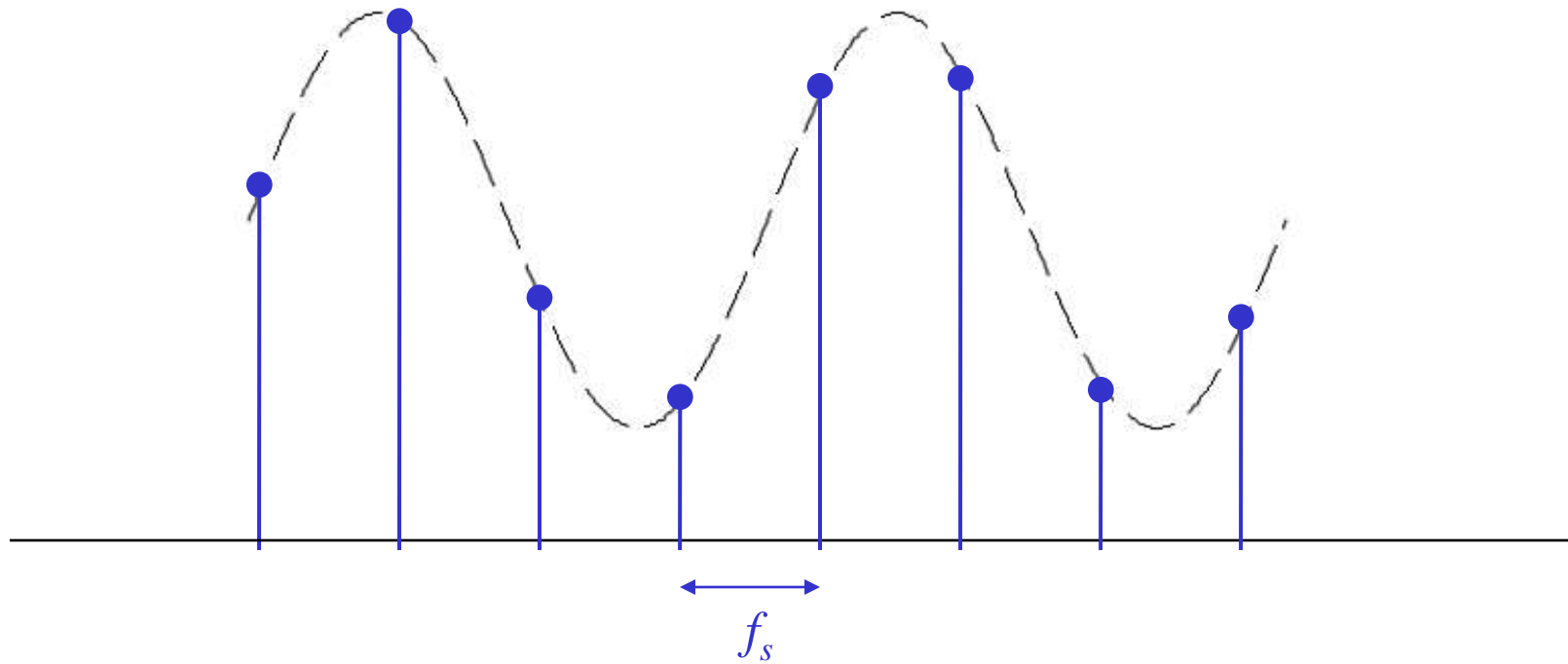
Fractional rate conversion



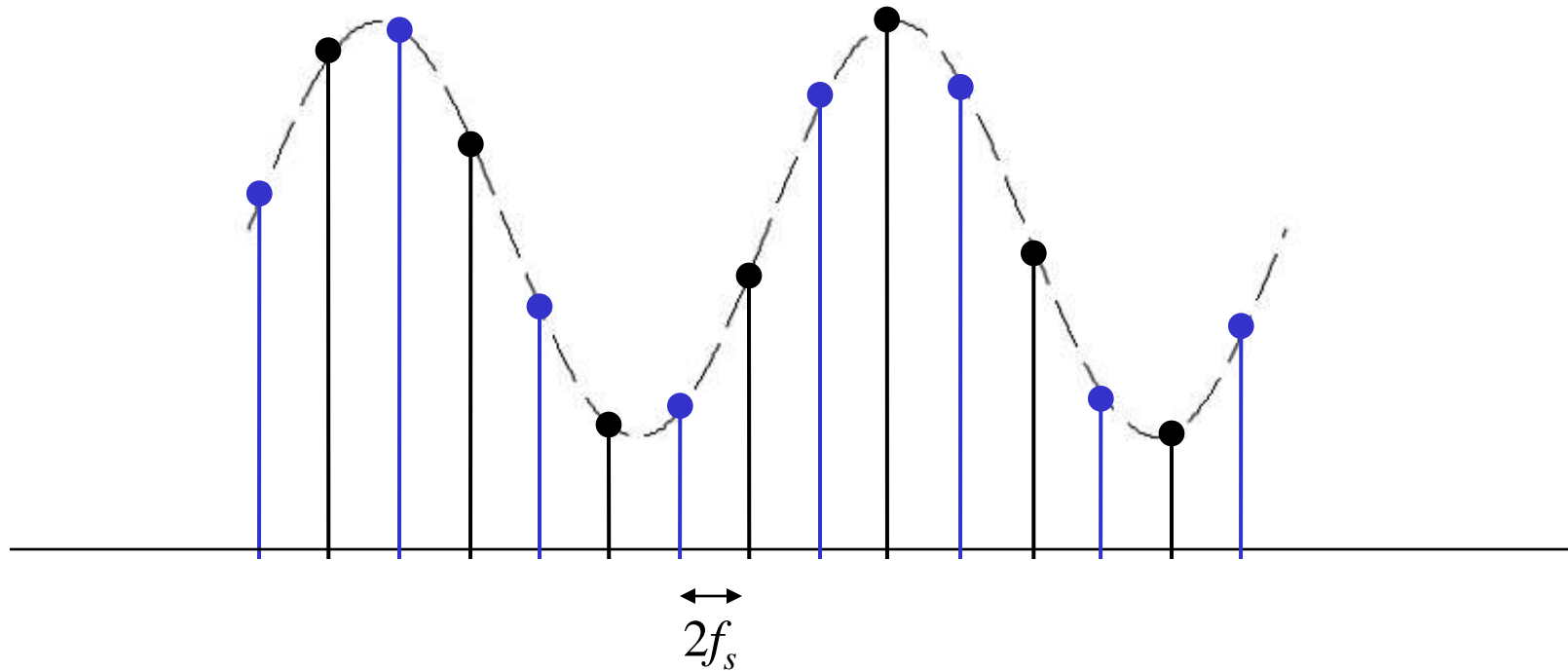
Fractional rate conversion



Fractional rate conversion



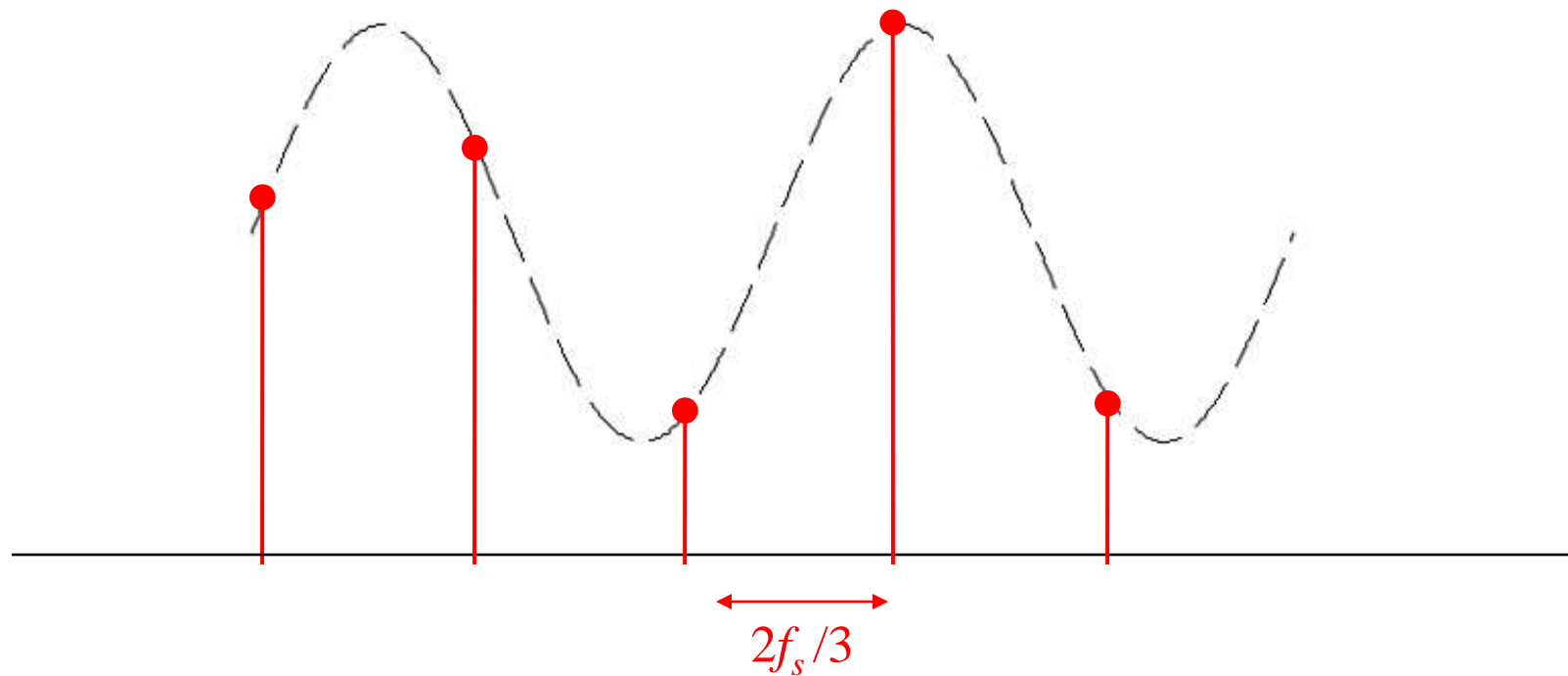
Fractional rate conversion



Interpolation



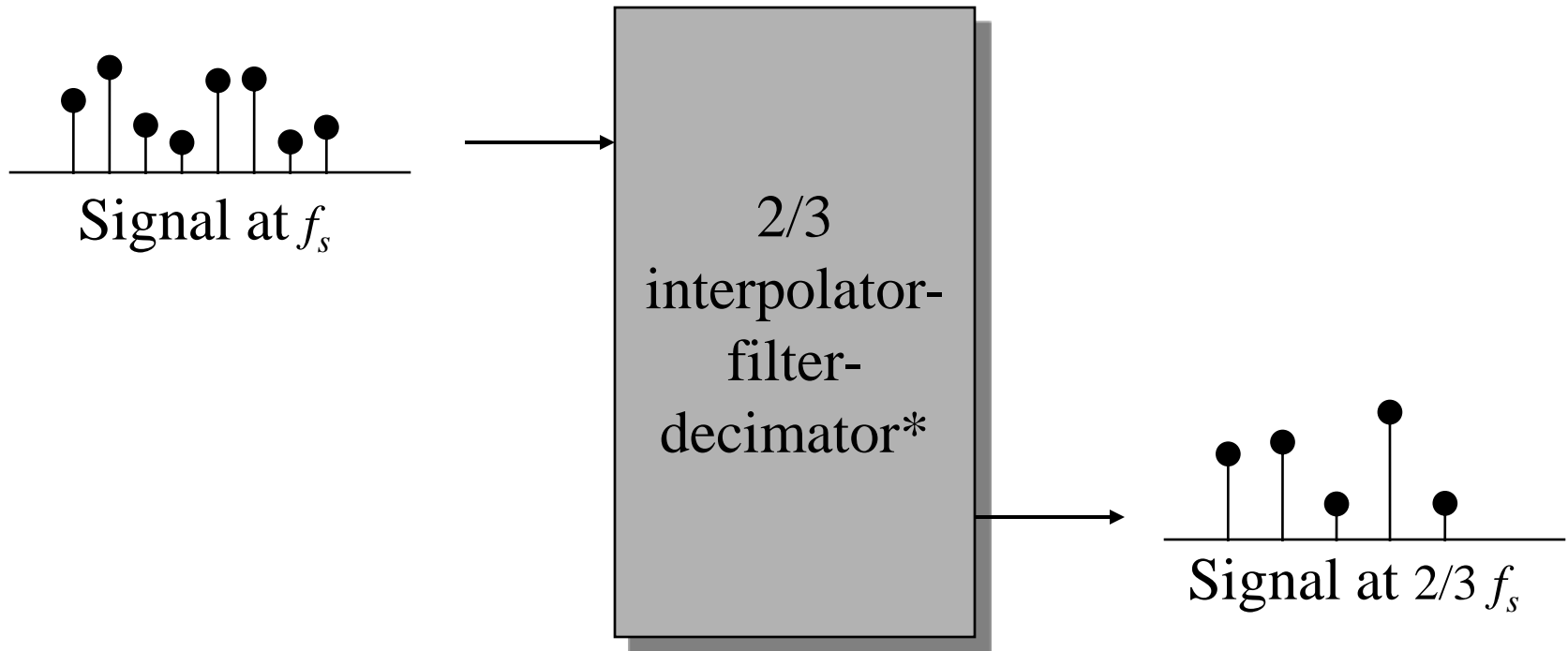
Fractional rate conversion



Decimation



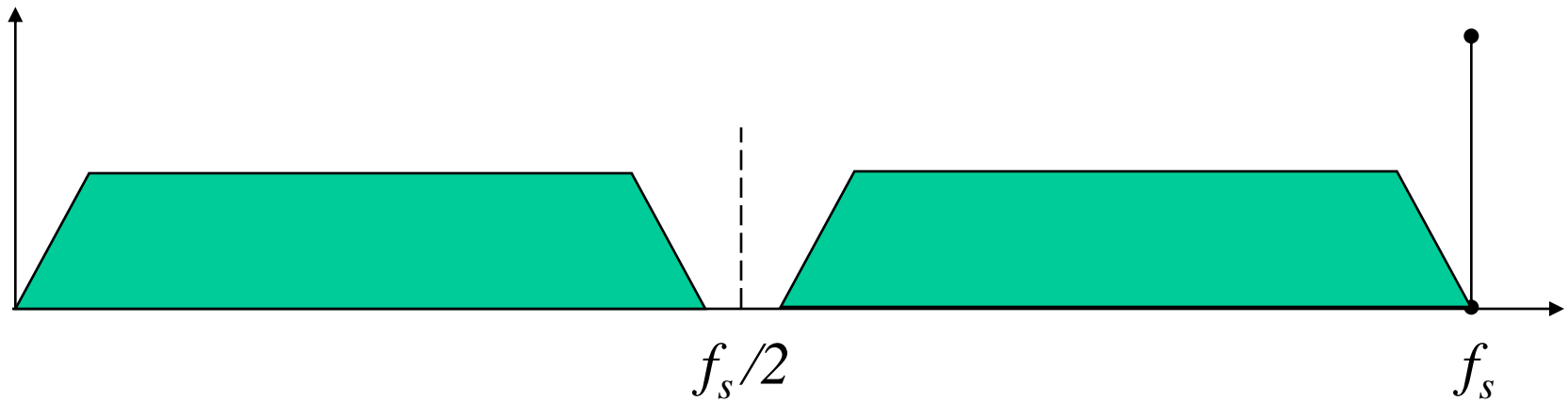
Fractional rate conversion



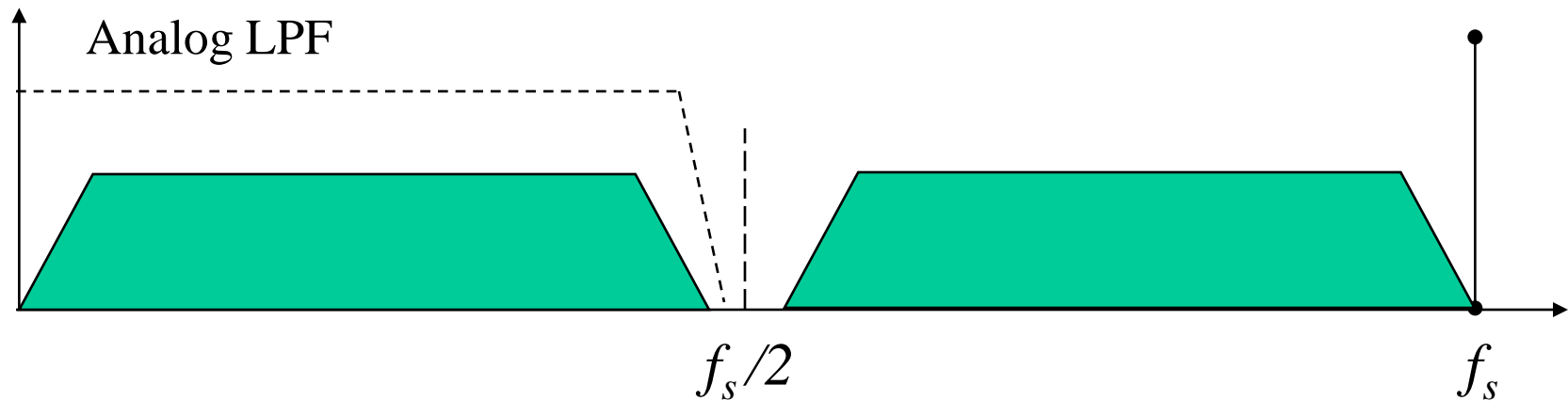
*w/ trade-offs



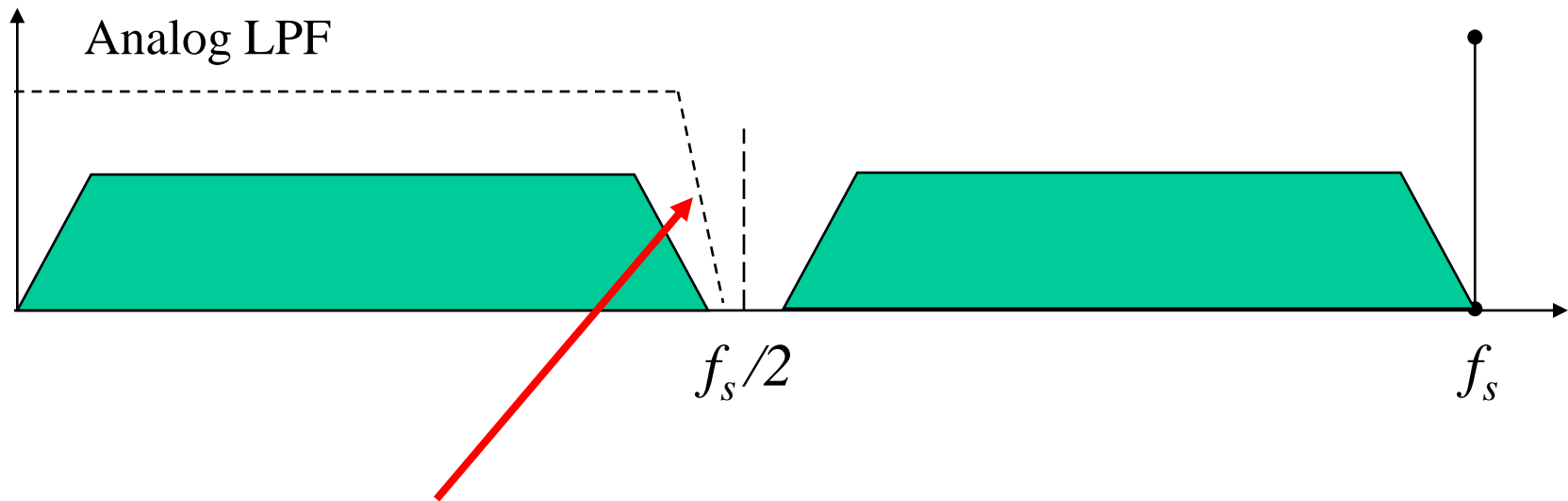
Oversampling



Critical sampling



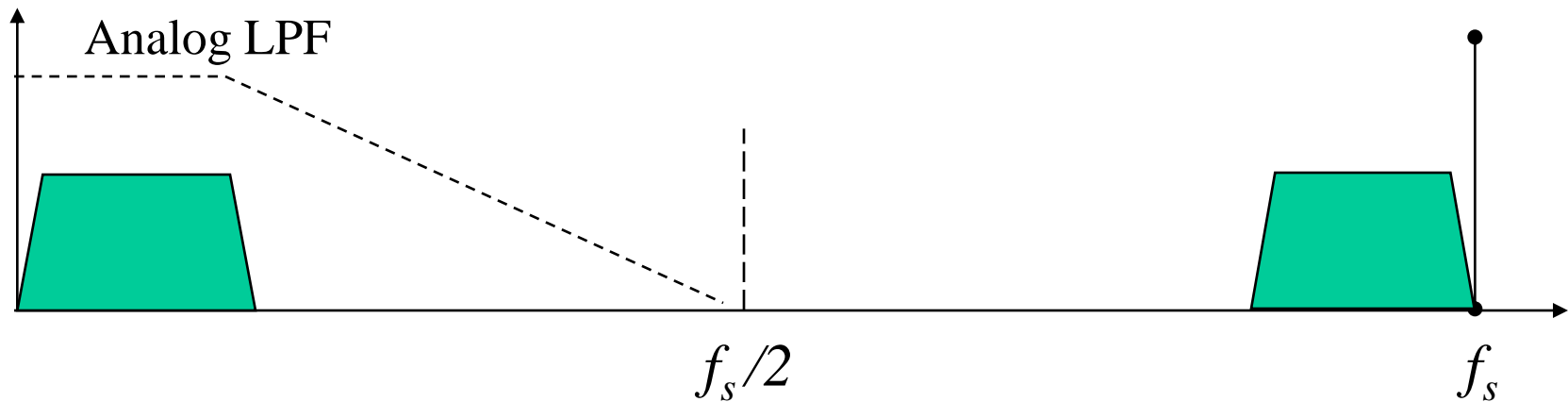
Critical sampling



- LPF must have very sharp cutoff
 - Complex
 - Expensive



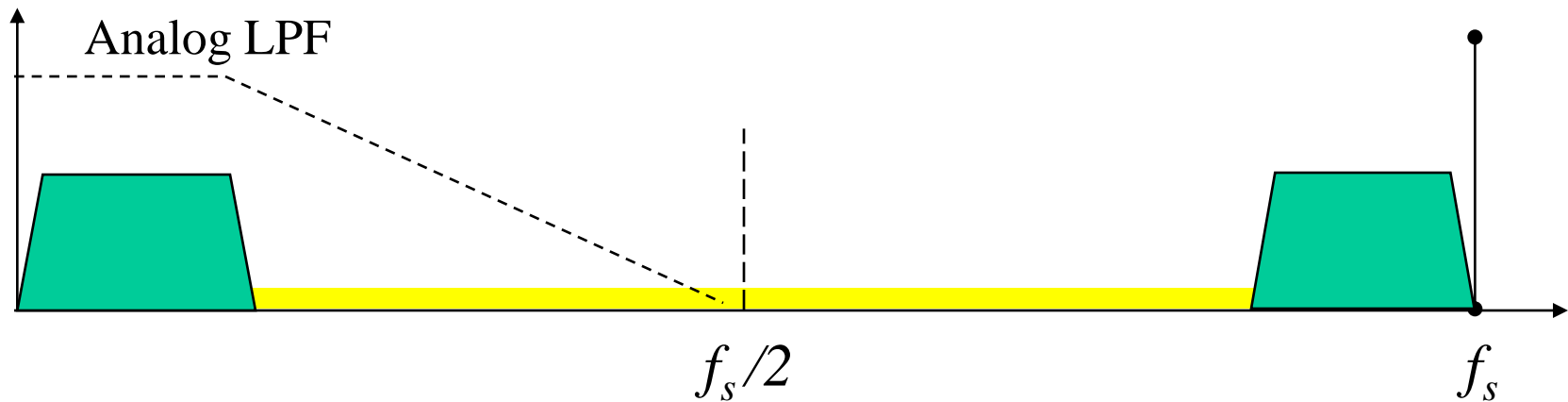
Oversampling



- Requirements on LPF are greatly relaxed



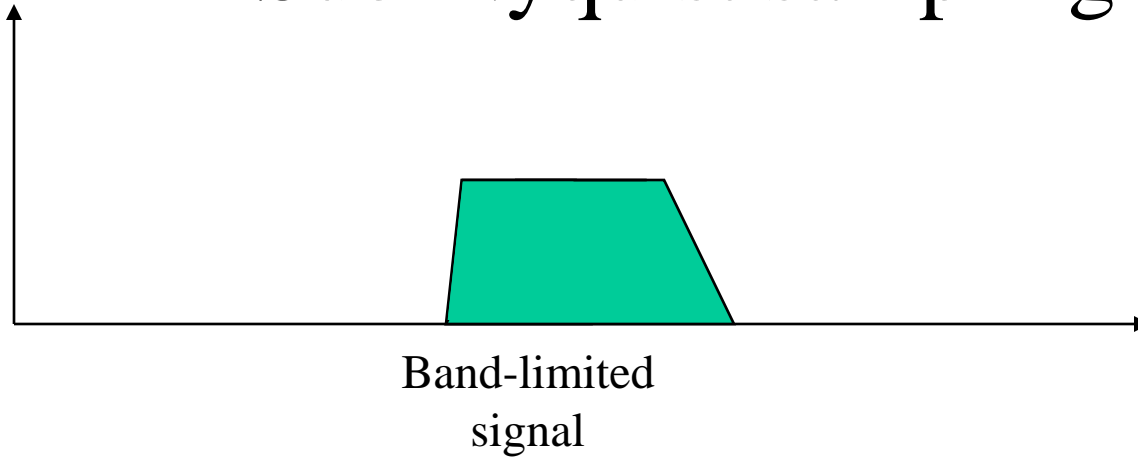
Oversampling



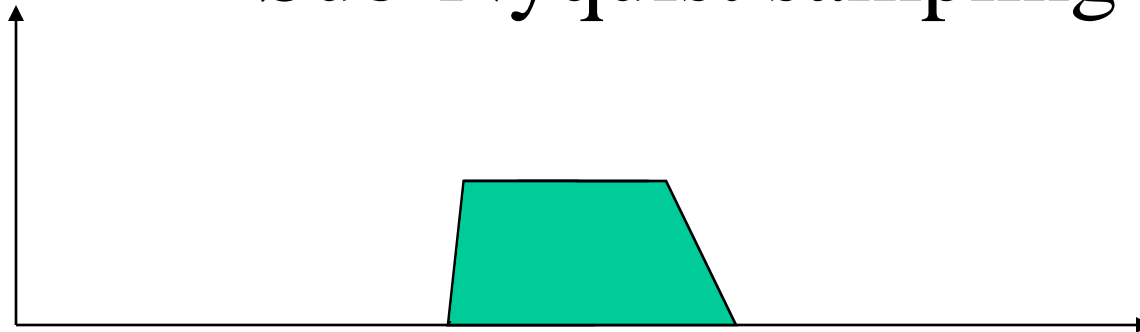
- Quantization noise level is reduced



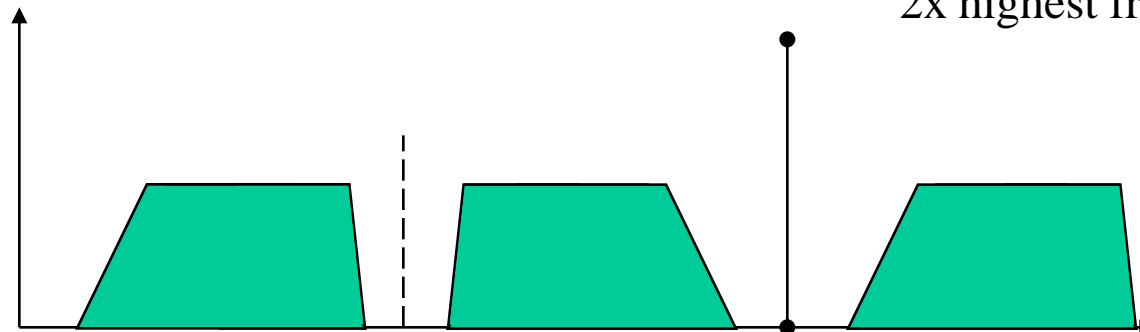
Sub-Nyquist sampling



Sub-Nyquist sampling



Band-limited
signal



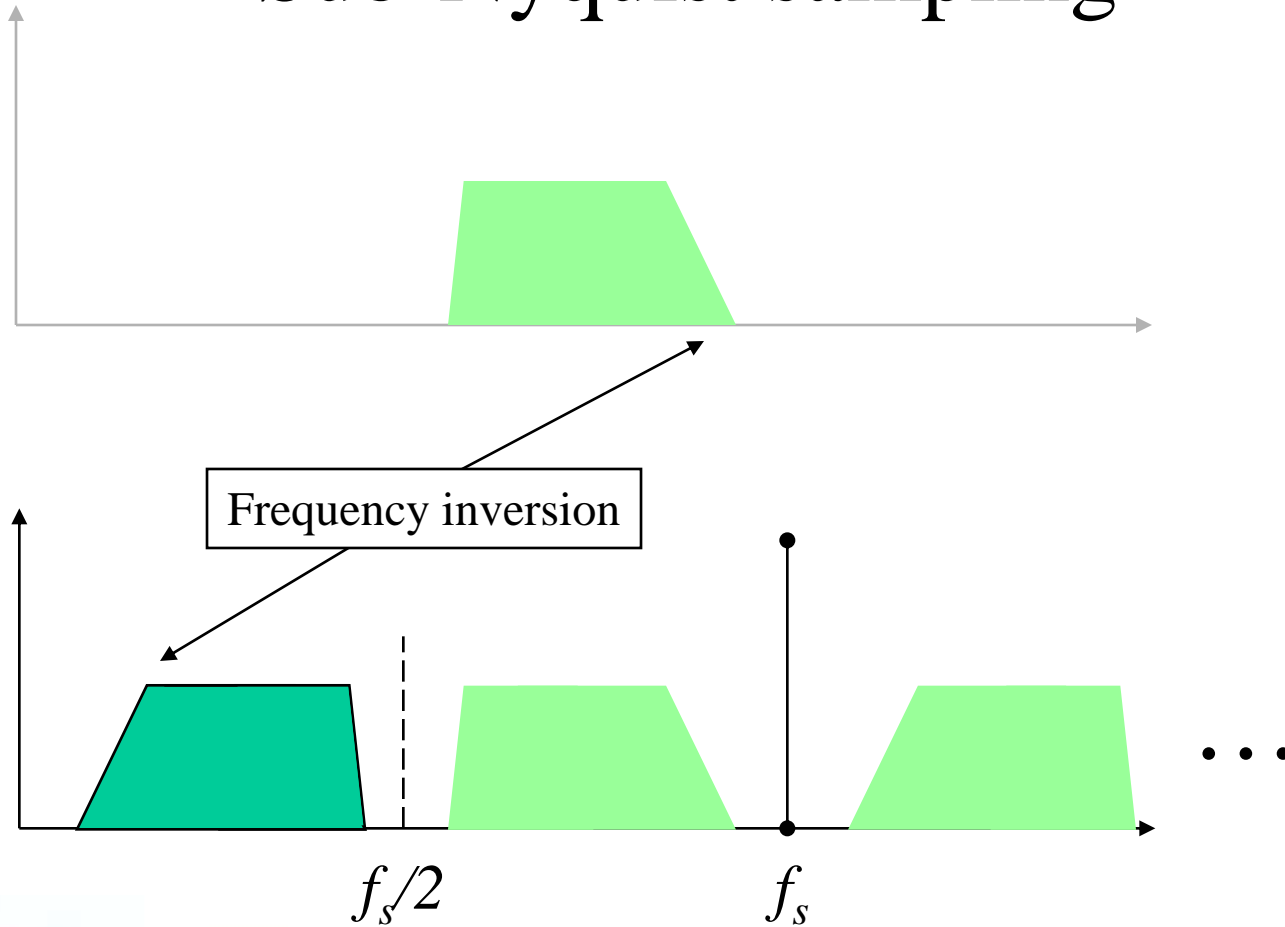
f_s is less than
2x highest frequency

$f_s/2$

f_s



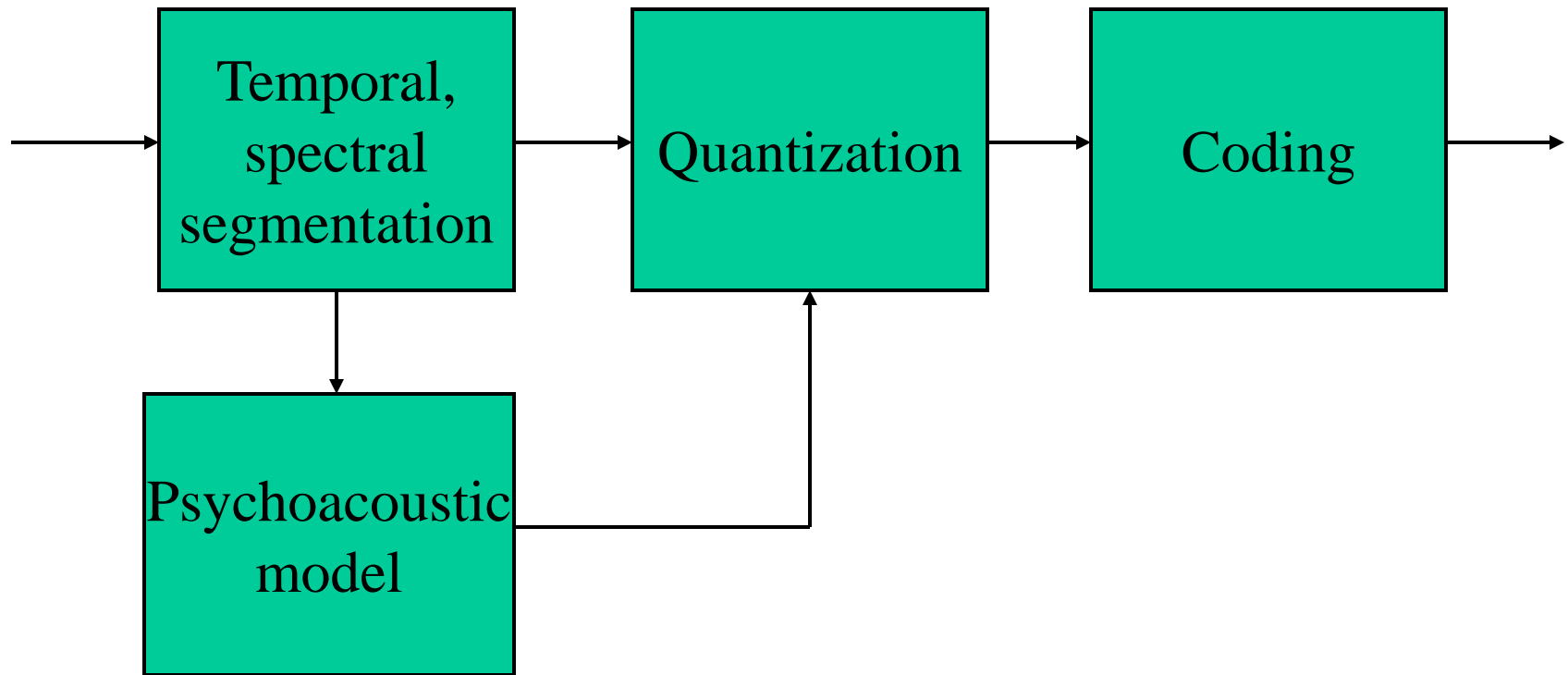
Sub-Nyquist sampling



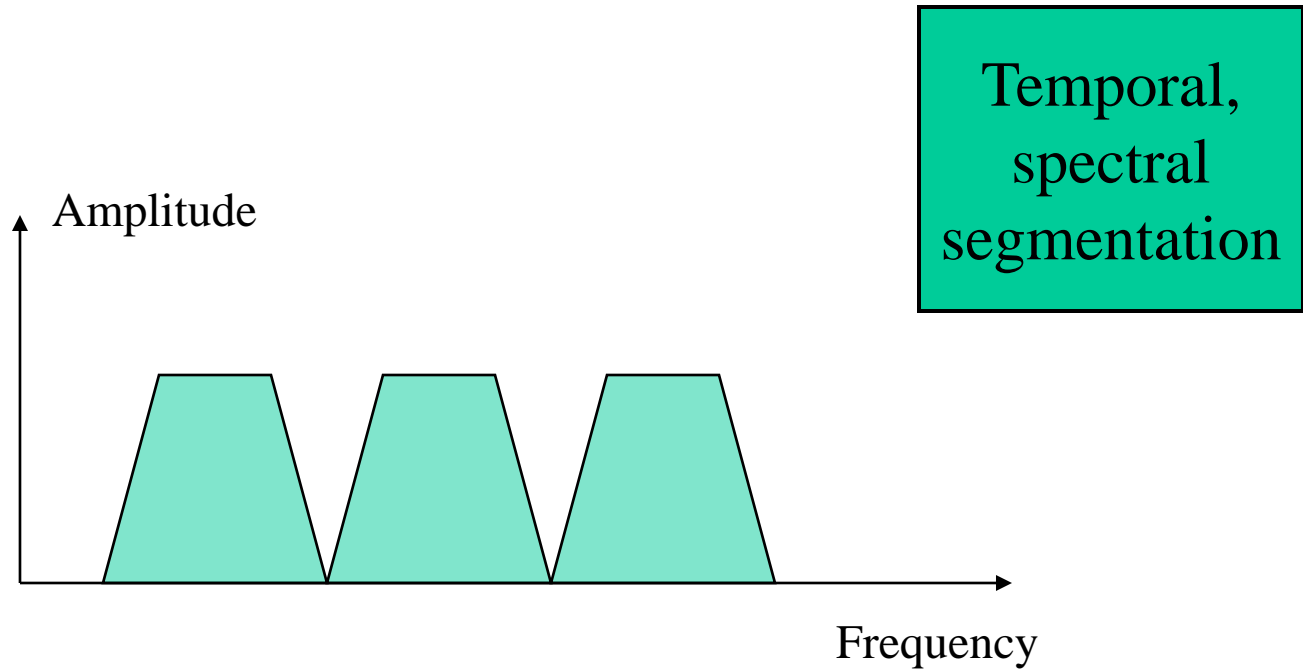
Audio compression



Audio compression



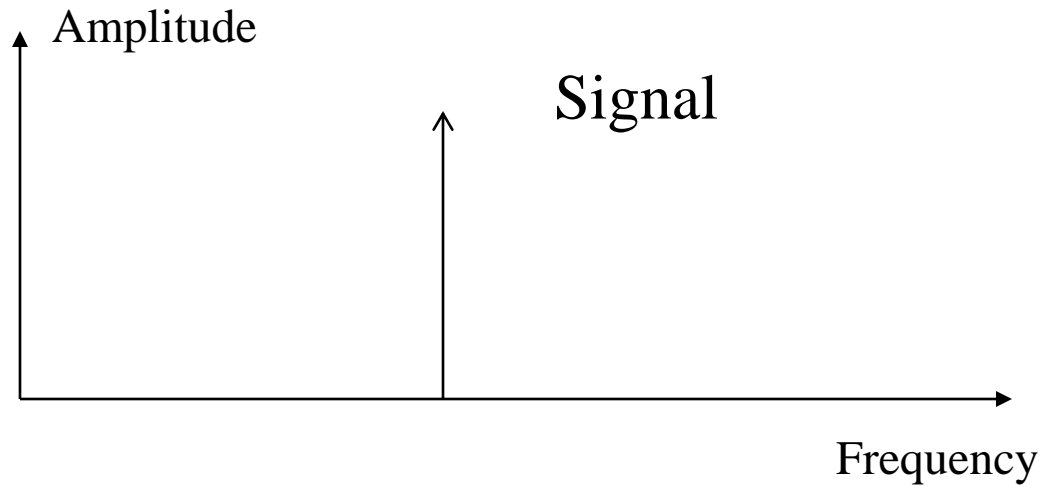
Audio compression



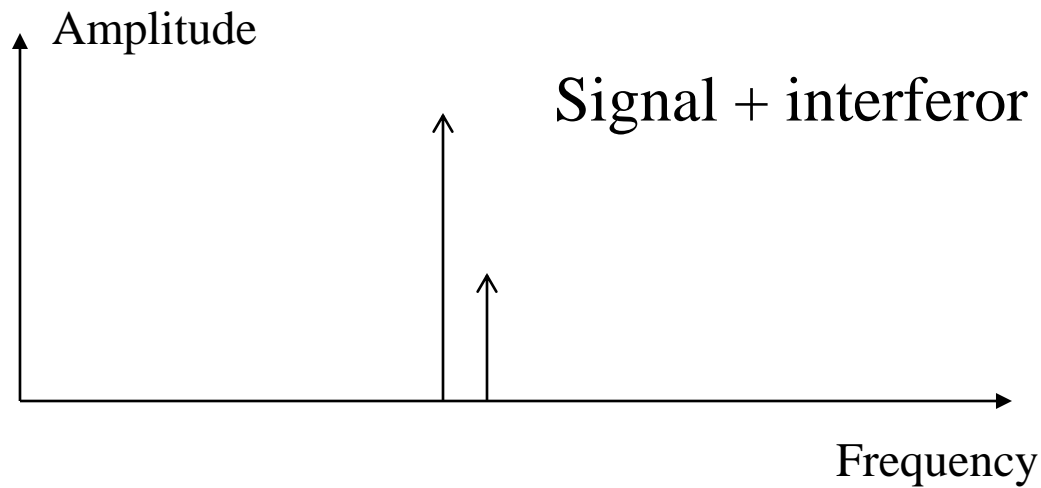
- Signal is broken up into smaller bands



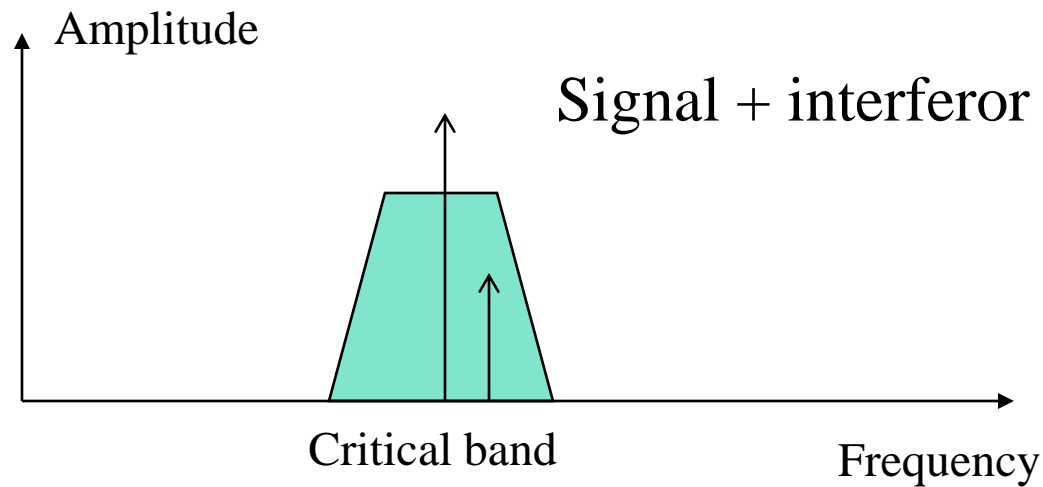
Psychoacoustic masking



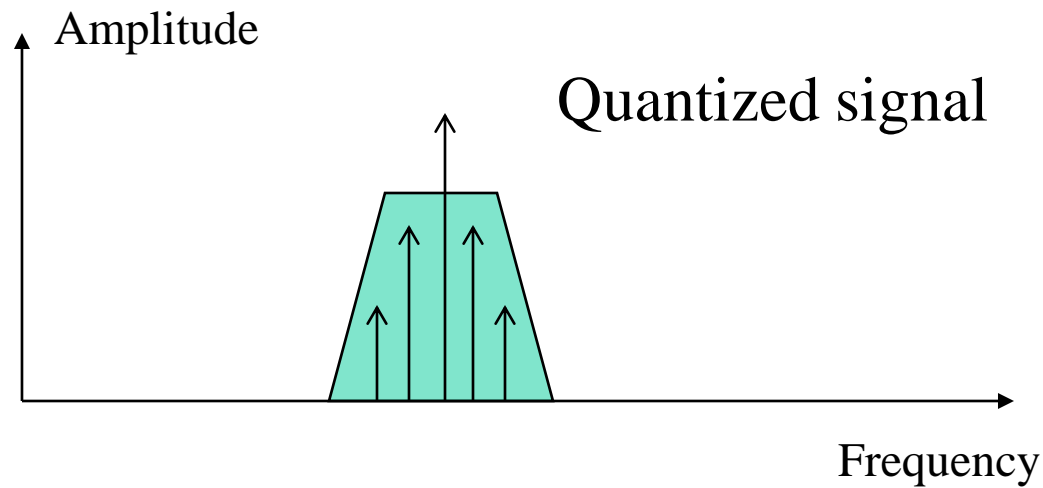
Psychoacoustic masking



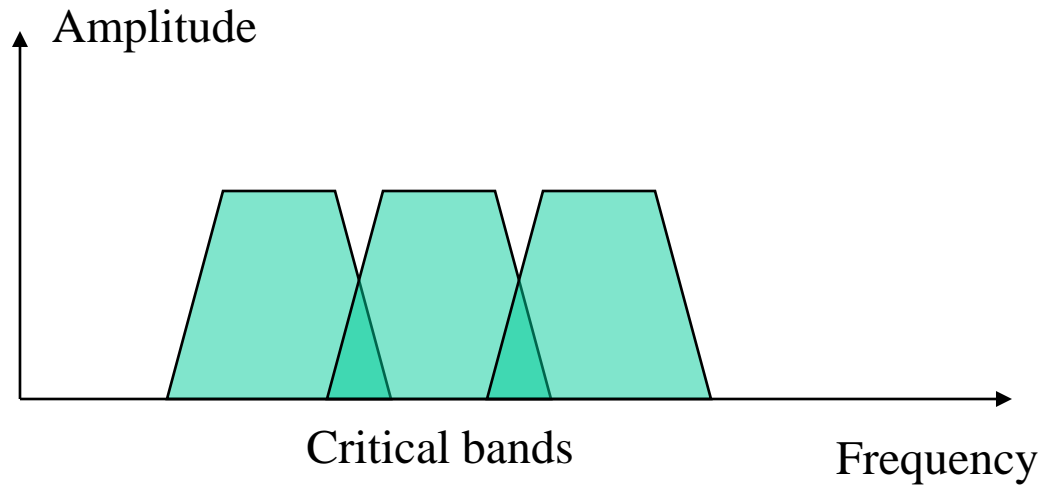
Psychoacoustic masking



Psychoacoustic masking



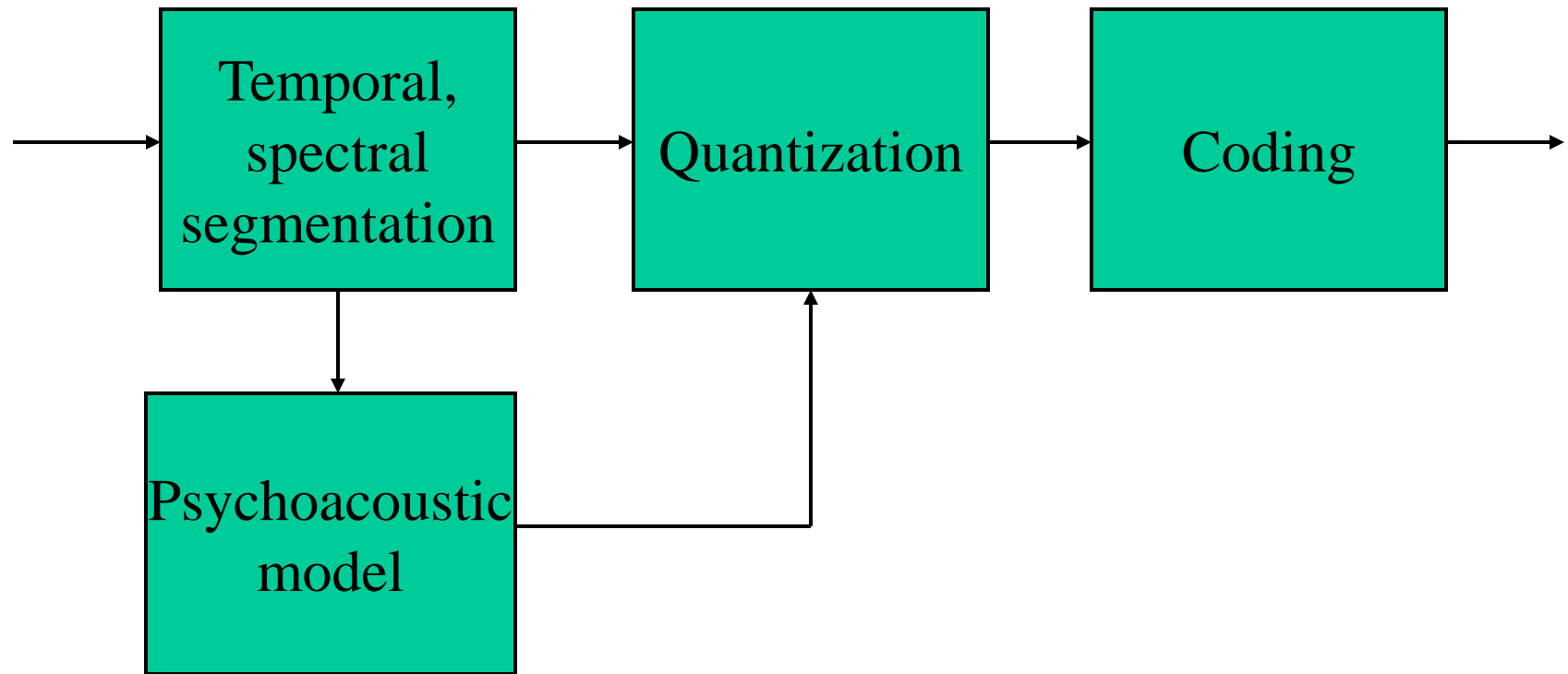
Psychoacoustic masking



- Quantization within critical bands discards “irrelevant” information



Audio compression



Compression artifacts

- Compression
 - Windowing
 - Quantization
 - Intermodulation



Compression artifacts

- Audibility
 - Warbling / swishing
 - “Underwater” sounds
 - Pre-echo and smearing
 - Samples: *www.ff123.net/training*



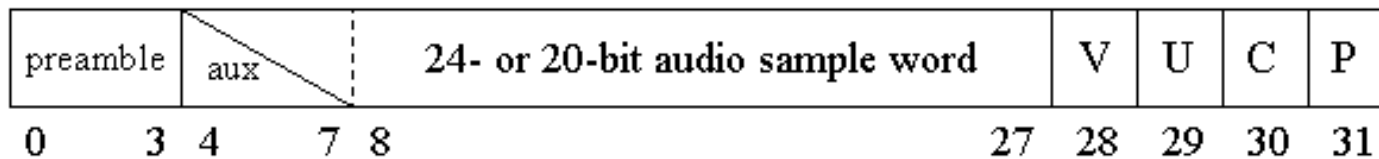
Other audio artifacts

- Aliasing
 - Sample rate conversion
 - Poor filtering



Digital interfaces

- AES3
 - Serial, arbitrary clock rate (3MHz typ.)
 - Typically 44.1kHz or 48kHz sampling
 - 16–24 bit quantization
 - Sub-frame protocol with channel status



Digital interfaces

- AES3
 - Optical (TOSLINK)
 - S/PDIF

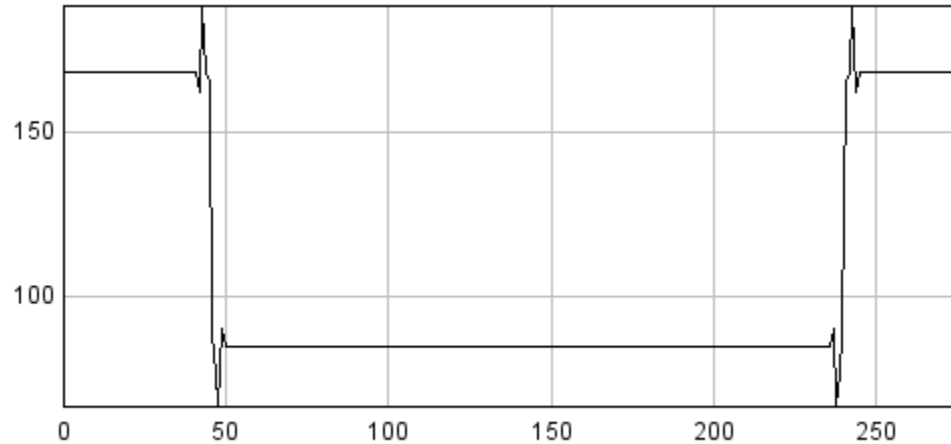


Digital interfaces

- AES10
 - MADI, Multichannel Audio Digital Interface
 - Serial digital
 - 28, 56, or 64 channels
 - Coaxial cable or fiber-optic lines
 - Sampling rates up to 96kHz
 - Resolution up to 24 bits per channel



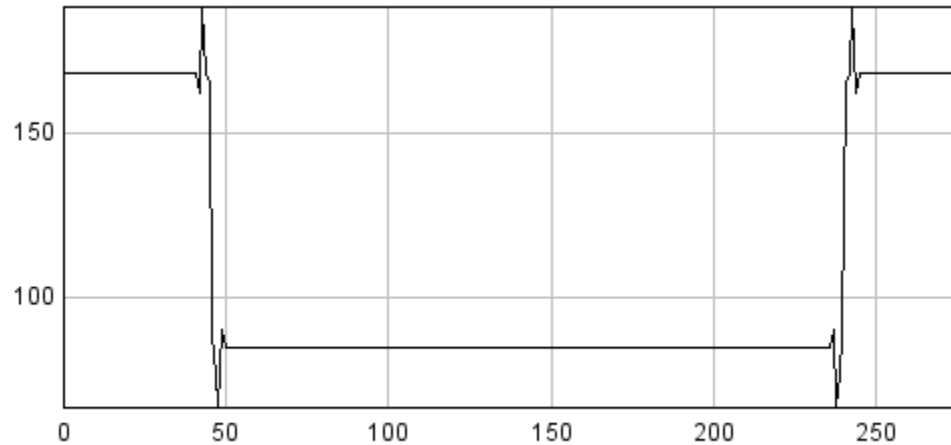
Ringling



Edge transitions due to poor filters



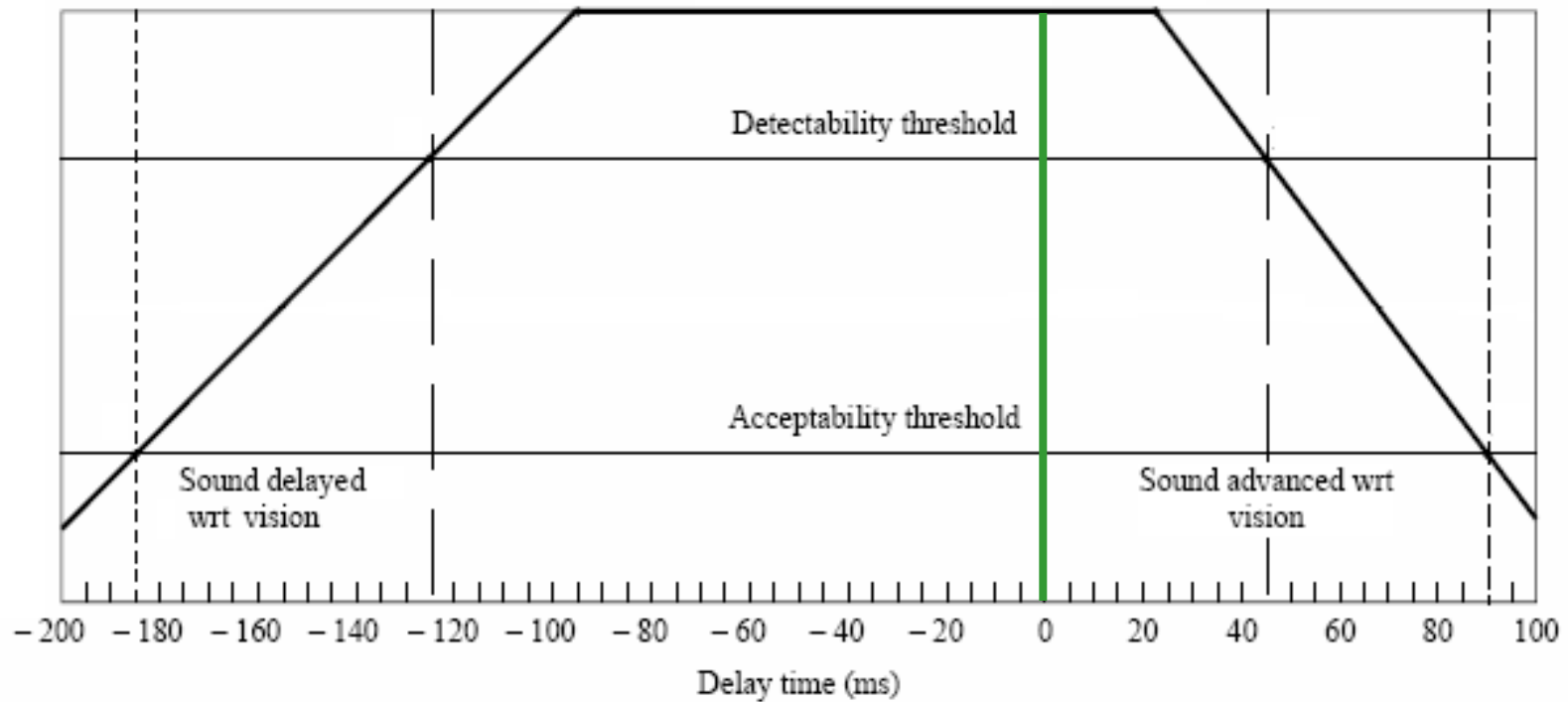
Ringling



Serial digital signals can be degraded



A/V sync (Lip sync)



A/V sync

- ATSC recommends:
 - Sound should never lead the video by more than 15 ms
 - Sound should never lag the video by more than 45 (± 15) ms



A/V Sync

- ATSC recommends:
 - sound should never lead the video by more than 15 ms
 - sound should never lag the video by more than 45 (± 15) ms
- Still not met by many systems

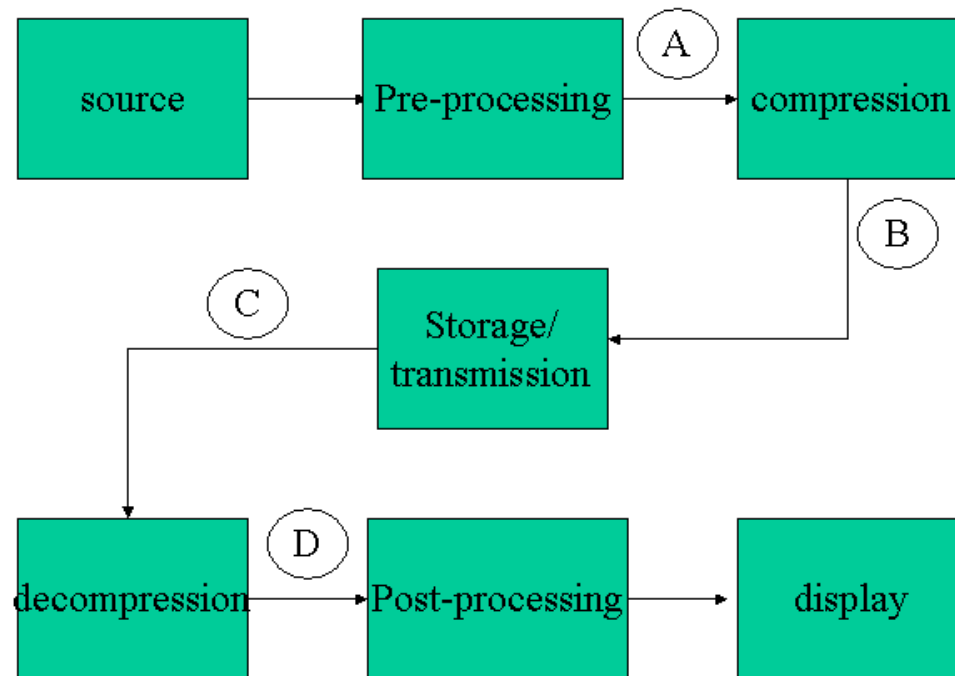


A/V sync delays

- Video elements have long and variable delays
- Re-routing of video and audio problematic



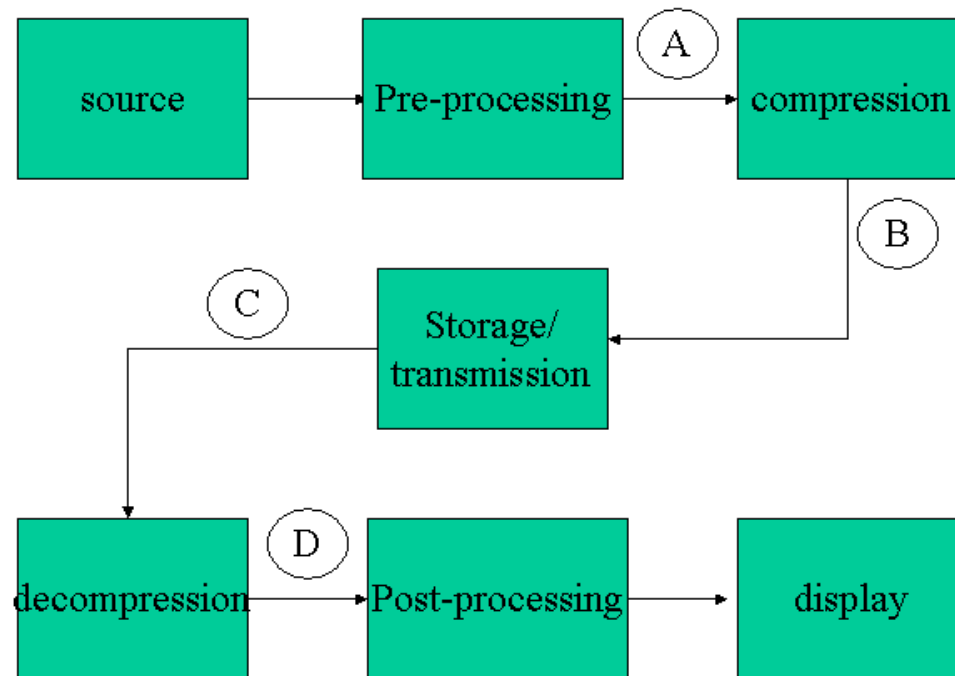
A/V sync delays



- Compression causes variable delays



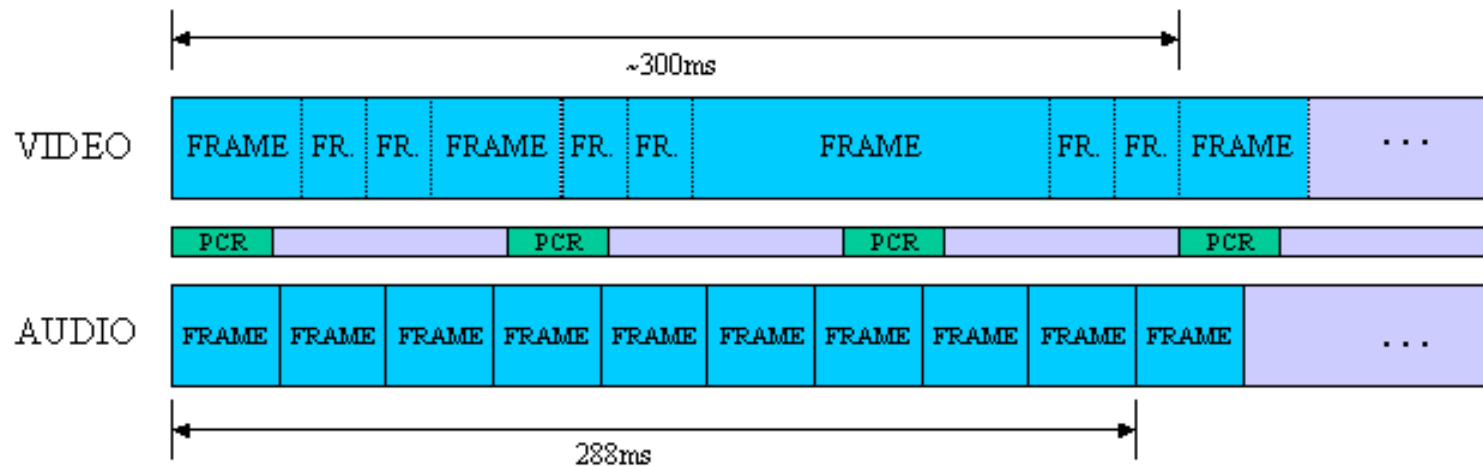
A/V sync delays



$$(t_B - t_A) + (t_D - t_C) = k (?)$$



Bit stream delays



- Video and audio streams have different frame-size characteristics



Bit stream issues

- MPEG Program Clock Reference (PCR)
 - Sync's encoder and decoder clocks without need for transmitting clock
 - Allows decoder to re-synthesize master encoding clock
 - Allows bit stream to be asynchronous with video and audio
 - Allows file storage, etc.



Bit stream issues

- MPEG Program Clock Reference (PCR)
 - Sync's encoder and decoder clocks without transmitting clock
 - Allows decoder to re-synthesize master encoding clock
 - Allows bit stream to be asynchronous with video and audio
 - Allows file storage, etc.
- MPEG Presentation Time Stamp (PTS)
 - Tells decoder when to present video and audio
 - Critical to A/V sync



Bit stream issues

- PCR and PTS monitoring varies in decoders
 - Problem if data is corrupted or missing
 - Resynchronization questionable
 - No formal requirement for decoder behavior

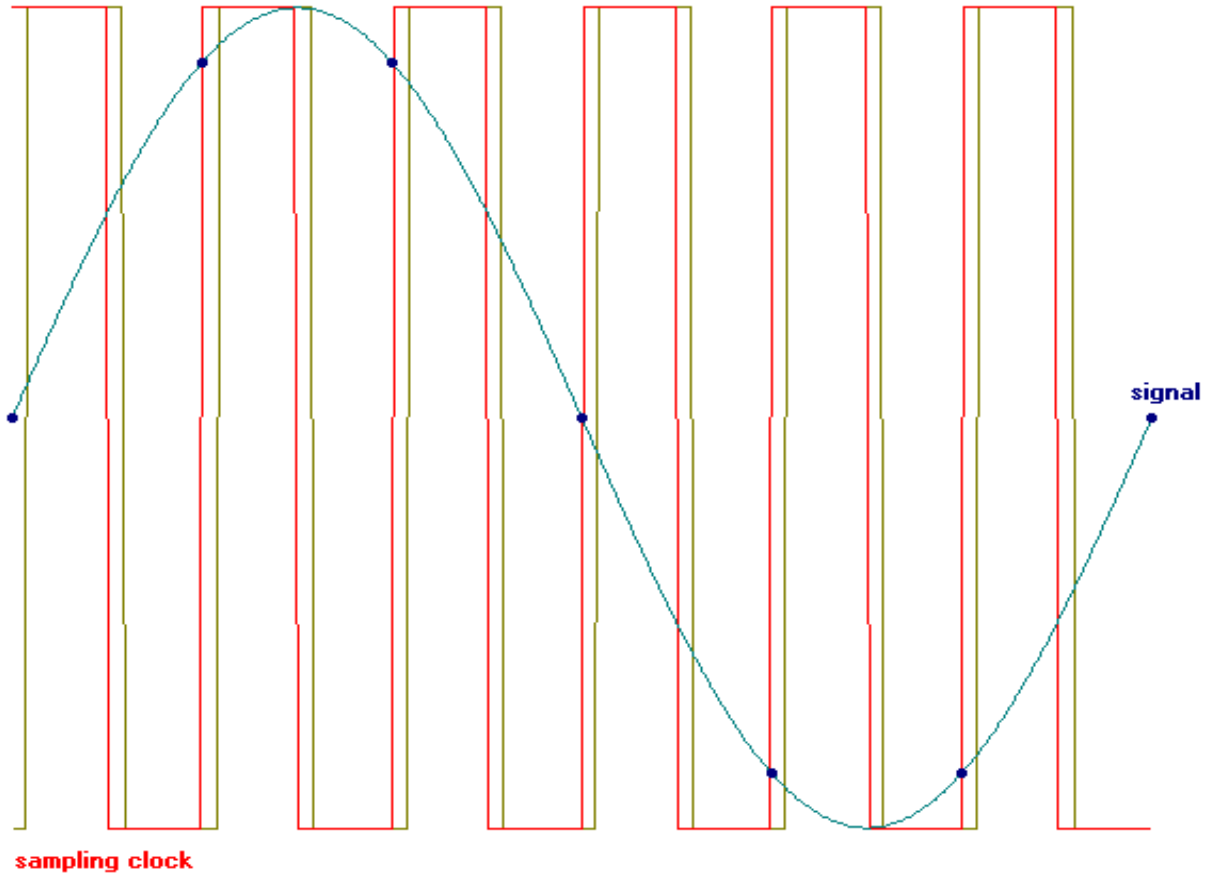


A/V issues

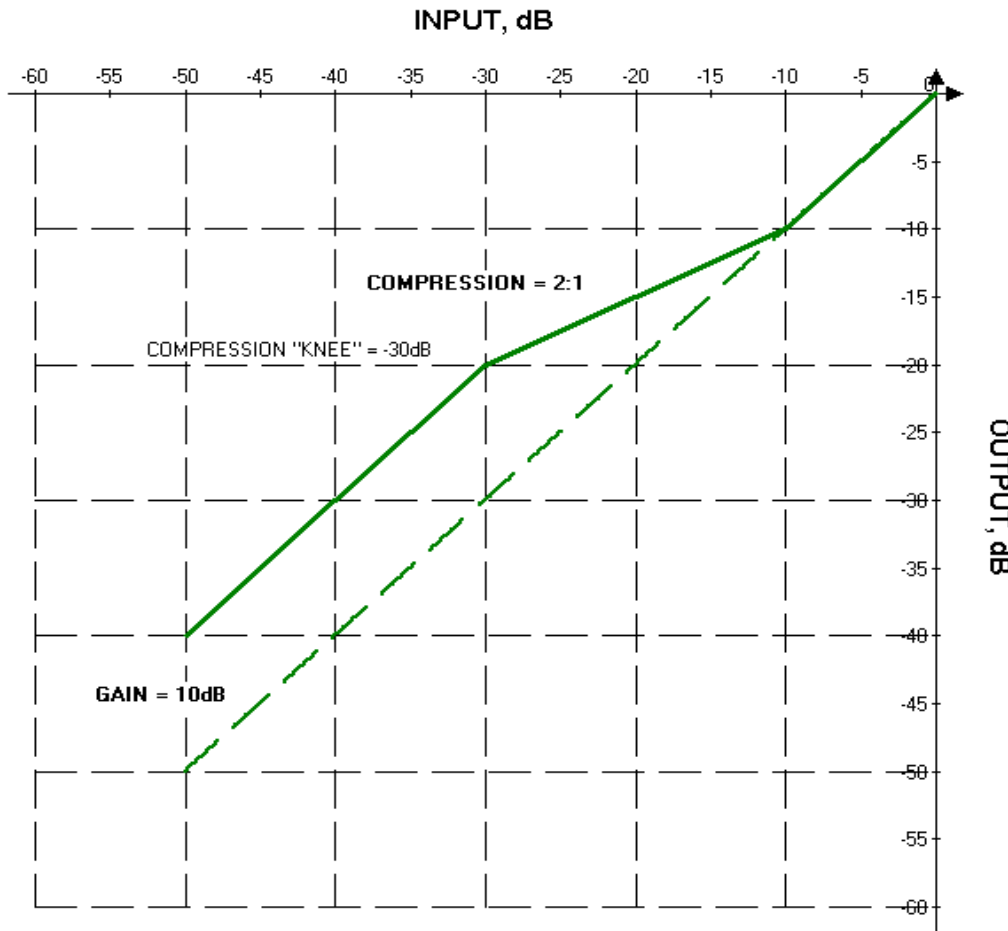
- Cross-industry collaboration needed between broadcasters and consumer electronics manufacturers



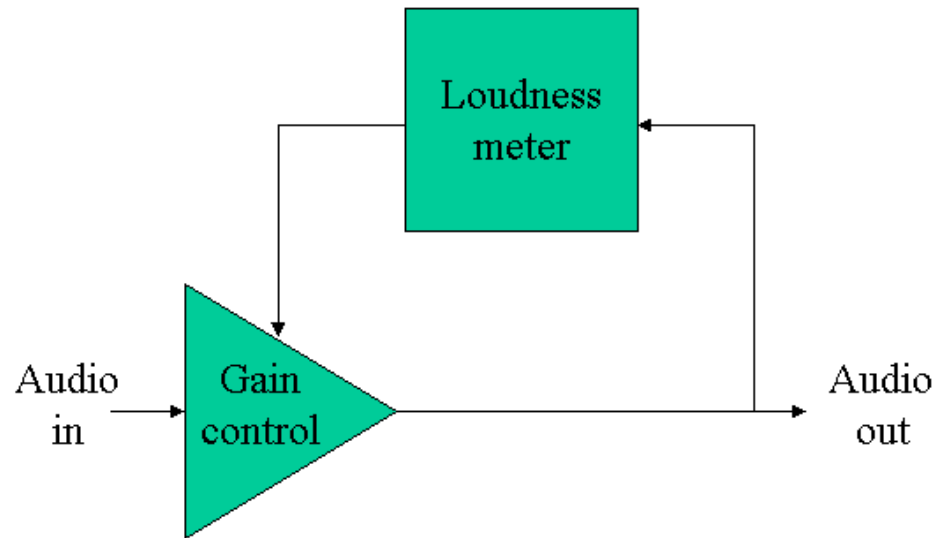
Sampling jitter



Level compression



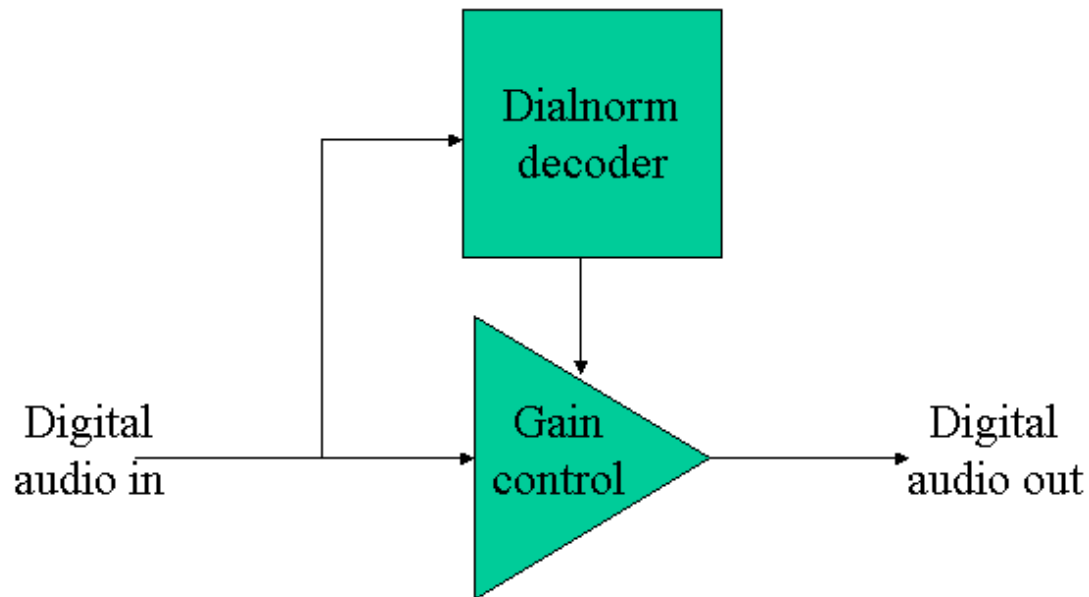
Loudness processing



Loudness controller



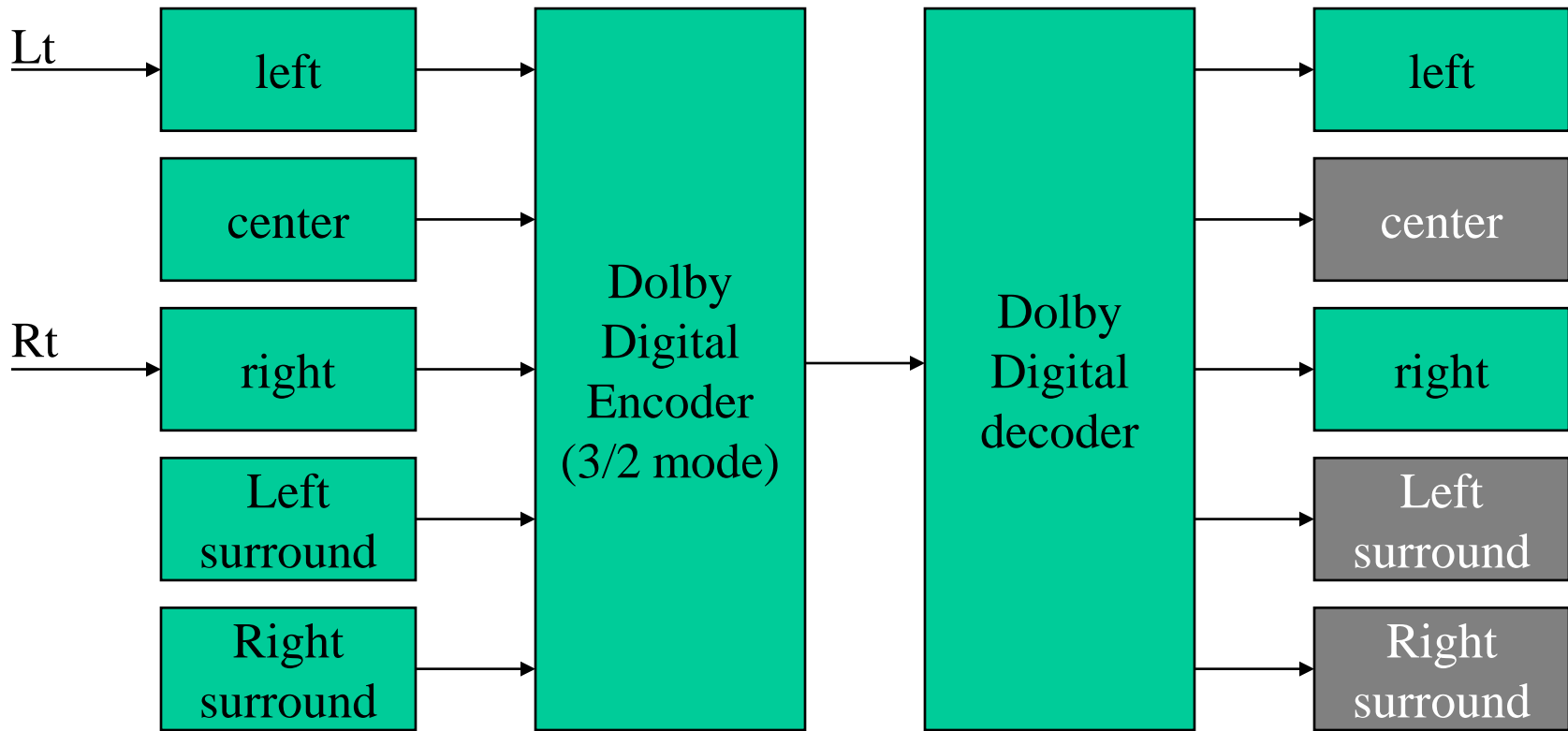
Loudness processing



Dialnorm at the DTV receiver adjusts loudness on a program-to-program basis



Surround processing



Beware of the “whole in the middle”



Summary

- Artifacts can arise from:
 - Signal processing
 - Sampling
 - Quantization
 - Compression



Audio systems

Questions?

acugnini@agcsystems.com

