ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Multimedia Technology

Section # 7: Audio

Instructor: George Xylomenos

Department: Informatics

Contents

- Nature of sound
- Digitization via transform
- Digitization via sampling
- Sample quantization
- Pulse code modulation
- Symbolic representation and sound synthesis

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Nature of sound

Class: Multimedia Technology, **Section # 7:** Audio

Instructor: George Xylomenos, **Department:** Informatics

What is sound? (1 of 3)

- Sound is generated by oscillating materials
  - Example: hitting a drum skin
  - Oscillation creates pressure waves
    - Also called sound waves
  - The pressure waves reach our ears
  - The eardrums oscillate based on pressure
  - A set of bones move accordingly
  - Eventually, our brain senses sound

What is sound? (2 of 3)

- Sound waves require a carrier
  – Gasses (like air) are the best carriers
  – Solids and fluids can also carry sound
    - Only some frequencies
    - And with a lot of attenuation
    - Depending on the material
  – Sound does not travel in a vacuum
    - Space battles are silent!

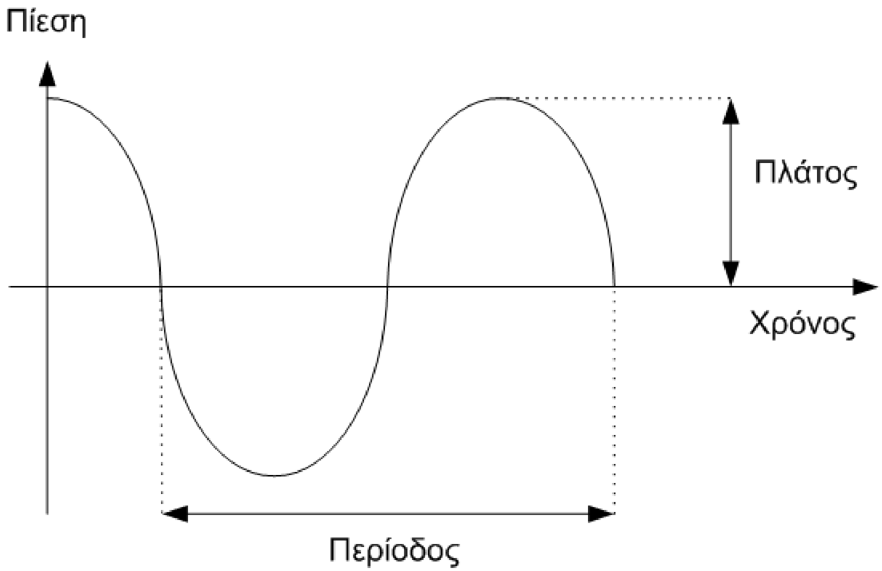
What is sound? (3 of 3)

- Sound waves can use multiple paths
  - The wave can be reflected
  - What we hear depends on the environment
  - Direct wave and multiple reflections
- Sound is sensed by both ears
  - This creates a sense of direction
    - Multiple speakers needed to recreate this
    - And multiple microphones to capture it

Sound categories

- Audible sounds: 20 Hz - 20 kHz
  - Also called acoustic signals
  - Below 20 Hz: infrasound
  - Above 20 kHz: ultrasound
  - In practice, each human is different
    - With age, we lose sensitivity to higher frequencies
  - Humans are most sensitive around 3 kHz
    - This is where speech is!

Audio waveforms (1 of 2)



- Sound can be represented as a wave
  - Showing how pressure varies with time
  - The y axis shows the pressure variation
    - The energy is the square of this variation

Audio waveforms (2 of 2)

- Amplitude: determines the level
  - Divergence from normal pressure
- Period of the wave
  - Periodic sounds are more musical
- Frequency: inverse of the period
  - Measured in cycles per second (Hertz, Hz)
  - Pitch of the sound

Audio capture

- Microphone: contains a diaphragm
  - The pressure wave excites the diaphragm
  - Its motion is converted to electric current
    - With magnetic or other means
  - Many types of microphone exist
    - Dynamic, Condenser, Ribbon
    - Each type responds somewhat differently to sound
    - Audio engineers care a lot about microphones!

Audio playback

- Speaker: an oscillating cone
  - Current is converted to oscillations
    - Usually with magnetic means
  - The cone oscillation excites the air
  - A sound wave is created
  - Many different types of speaker exist
    - Each can reproduce different ranges of sound
    - Loudspeakers typically have multiple speakers

Digitization

- A soundwave is an analog signal
  - Amplitude as a function of time g(t)
  - Can take any value at any time
  - A computer cannot represent this!
    - Limited number of values
    - Limited accuracy of values
- Digitization is needed
  - Conversion of analog to digital

ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ

ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

Digitization via transform

Class: Multimedia Technology, **Section # 7:** Audio

Instructor: George Xylomenos, **Department:** Informatics

Fourier transform (1 of 4)

- The Fourier theorem
  - Decomposition of g(t) to simpler functions
    - It must be periodic
    - But the simpler functions may be infinite!
  - f=1/T: fundamental frequency
  - a_n and b_n: harmonics of the signal

$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \sin(2\pi nft) + \sum_{n=1}^{\infty} b_n \cos(2\pi nft)$$

Fourier transform (2 of 4)

$$a_n = \frac{2}{\mathrm{T}} \int_0^{\mathrm{T}} g(t) \sin(2\pi nft)\, dt$$

$$b_n = \frac{2}{\mathrm{T}} \int_0^{\mathrm{T}} g(t) \cos(2\pi nft)\, dt$$

$$c = \frac{2}{\mathrm{T}} \int_0^{\mathrm{T}} g(t)\, dt$$

- Calculation of g(t)’s component coefficients
  - Can recreate g(t) by superimposing them
  - The coefficients are just numbers

Fourier transform (3 of 4)

$$\sqrt{a_n^2 + b_n^2}$$

- RMS energy of harmonic n
  - Does not decrease with n
  - The harmonics can be too many
  - Dropping harmonics distorts the sound
- The coefficients also need digitization!
  - Only some coefficients can be stored
  - Each coefficient has finite accuracy
  - Calculations also have limited accuracy

Fourier transform (4 of 4)

- So, Fourier is not good for digitization
- But, it is good for many other things!
  - Analyzing the nature of a sound
  - Finding most important frequency components
    - Those with large coefficients
    - Heavily used in audio compression
  - Modifying specific harmonics
    - Used in audio processing

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

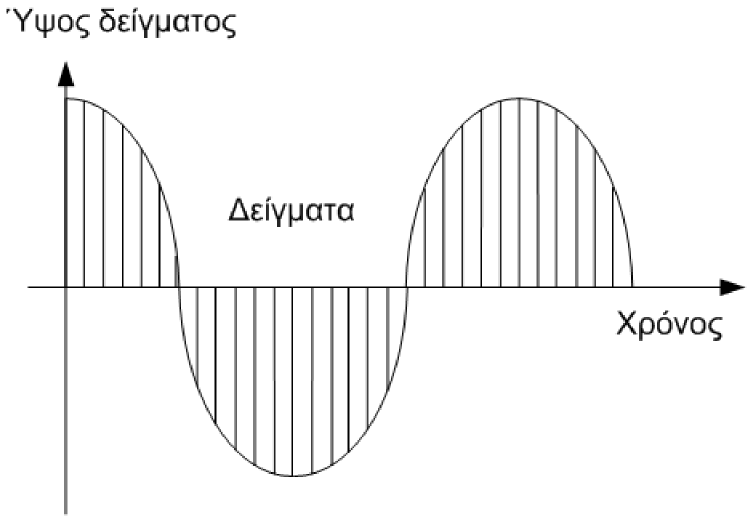
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Digitization via sampling

Class: Multimedia Technology, **Section # 7:** Audio

Instructor: George Xylomenos, **Department:** Informatics

Sampling (1 of 3)



- Digitization via sampling
  - Periodic measurement of amplitude
  - Each measurement is a sample
  - Makes the signal discrete in the time dimension

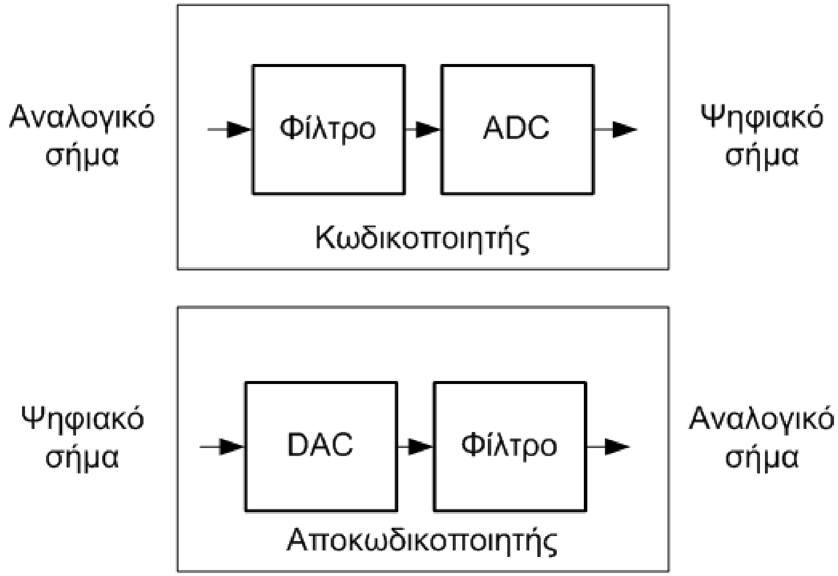
Sampling (2 of 3)

- Sampling rate (sampling frequency)
  – Measured Hz (samples / sec)
  – 44,1 kHz in CD Audio, 8 kHz in telephony
- Nyquist's sampling theorem
  – Assume a signal where f is the highest frequency
    - We can calculate f using a Fourier transform
    - Or just use a lowpass filter with f as the cutoff
  – A sampling rate of 2f is enough to avoid distortion
    - Specifically, to avoid aliasing

Sampling (3 of 3)

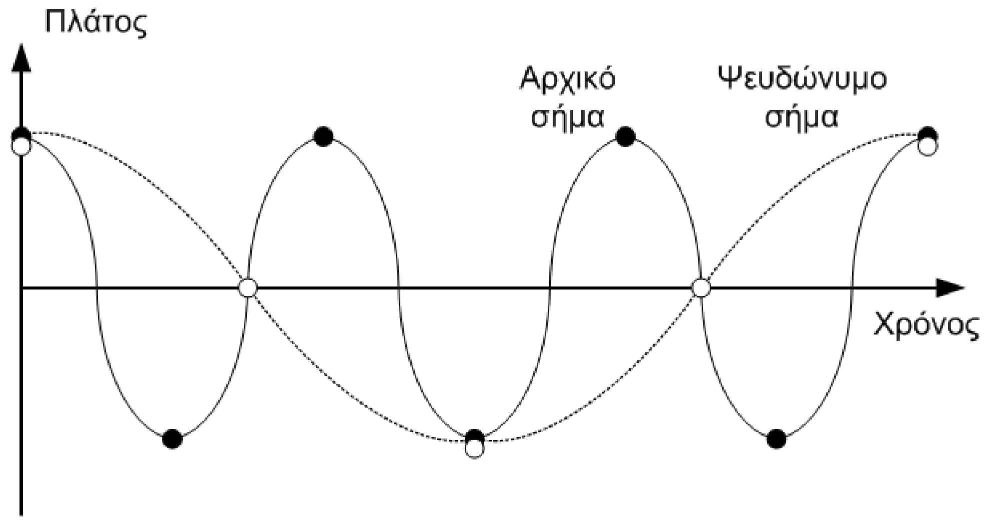
- Alternate version with frequency range
  - Assume a signal with frequencies [f1,f2]
    - Using a bandpass filter with f1 and f2 as cutoffs
  - Sampling rate must be 2(f2-f1)
    - We can move the signal to [0,f2-f1]
    - Without losing any information
- Also called Nyquist-Shannon theorem
  - And Whittaker-Nyquist-Shannon theorem

ADC and DAC (1 of 4)



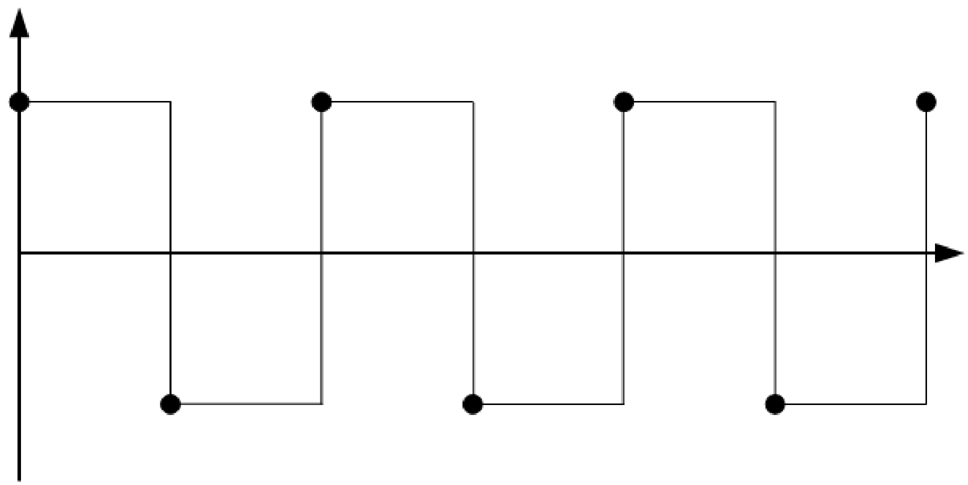
- Analog to digital and vice versa
  - ADC: Analog to digital converter
  - DAC: Digital to analog converter
  - Filters before ADC and after DAC

ADC and DAC (2 of 4)



- Why a filter before a DAC?
  - Alias signals
    - Black dots: adequate sampling rate
    - White dots: inadequate sampling rate
  - The filter cuts off frequencies above f Hz

ADC and DAC (3 of 4)



- Why a filter after the ADC?
  - Audio pulses are generated periodically
  - The output is a square wave
    - Contains frequency way above f Hz
  - Filtering removes the high frequency components

ADC and DAC (4 of 4)

- Sound waves may contain high frequencies
  - Need to filter them before they are sampled
- Audio pulses contain high frequencies
  - Need to filter them before the speaker
- Note: filters are not perfect!
  - They attenuate but not eliminate frequencies
  - We allow for some slack in the signals
    - Filter at 20 KHz, sample at 44.1 KHz

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Sample quantization

Μάθημα: Τεχνολογία Πολυμέσων, **Ενότητα # 4:** Ήχος

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής

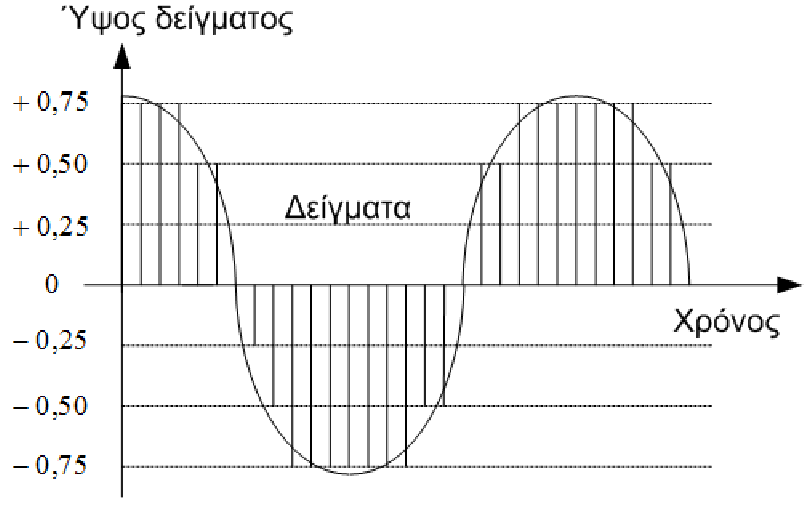
Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Quantization (1 of 2)



- (Sample) quantization
  - Transforms a continuous to a discrete value
    - Each sample is approximated
  - Makes the signal discrete in terms of amplitude

Quantization (2 of 2)

- Quantization affects quality
  - 16 bits: 65536 values, 8 bits: 256 values
  - Fewer values -> higher distortion
  - Called the quantization error
- How many discrete values are enough?
  - No Nyquist theorem for this!
  - We can instead bound the resulting error
  - How many bits do we need to achieve it?

Quantization error

- Calculating the quantization error
  - Assume a signal with an amplitude –V to +V
  - Assume n bits are available for quantization
  - Quantized values differ by $q = 2V / 2^n$
    - Assuming linear quantization (equal distances)
  - Worst case quantization error
    - $q/2 = V / 2^n$ (half of the distance)
  - How large is the acceptable error?

Quantization levels (1 of 4)

- 1^{st} approach: dynamic range
  - Amplitude of smallest audible signal: v
  - Amplitude of largest tolerable signal: V
  - Their difference is the dynamic range
    - $10 \log_{10}(V^2/v^2)$ dB = $20 \log_{10}(V/v)$ dB
    - Recall that power is the square of the amplitude
    - Measured in orders of magnitude (dB)

Quantization levels (2 of 4)

- Exploiting the dynamic range
  - Quantization error < smallest audible signal
    - A reasonable bound for this error
  - $V / 2^n < v \Rightarrow V / v < 2^n$
  - $\log_{10}(V/v) < n \log_{10}2 = 0,3\ n$
  - $20 \log_{10}(V/v) < 6\ n$
  - For a dynamic range of 40 dB, we need n >= 7
    - Hence, at least 7 bits / sample

Quantization levels (3 of 4)

- 2^{nd} approach: signal to noise ratio (SNR)
  - Assume signal with amplitude S
  - Assume noise with amplitude N
  - SNR: $10 \log_{10}(S^2/N^2)$ dB = $20 \log_{10}(S/N)$ dB
    - Power is the square of the amplitude
    - If $S^2 = 10N^2$, 10dB
    - If $S^2 = 100N^2$, 20dB
    - If $S^2 = 1000N^2$, 30dB

Quantization levels (4 of 4)

- Exploiting the SNR
  - Assume signal with maximum amplitude V
  - Assume that the noise is the quantization error q/2
  - The SNR is 20 $\log_{10}(V/(q/2))$
  - = 20 $\log_{10}(V/(V / 2^n))$
  - = 20 $\log_{10}2^n$ = 6 n
  - For an SNR > 40 dB, we need n >= 7
    - Hence, at least 7 bits / sample

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Pulse Code Modulation

Class: Multimedia Technology, **Section # 7:** Audio

Instructor: George Xylomenos, **Department:** Informatics

PCM types (1 of 6)

- Pulse Code Modulation (PCM)
  - Simple periodic sampling
    - Each sample is played back as a pulse
  - Memoryless: each sample is independent
    - Can start playback from any sample
- Linear quantization
  - Equal intervals between quantized values
  - One value per interval

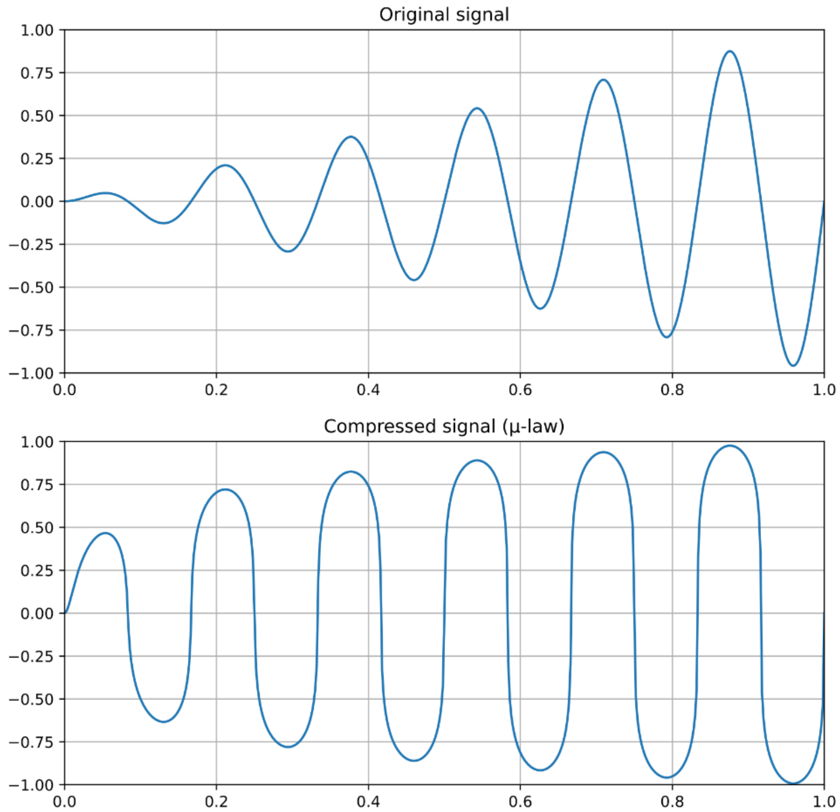
PCM types (2 of 6)

- Audio CD: CD-DA standard
  - Uses linear quantization
    - Makes no assumptions about the input signal
  - Assumes a frequency range of 20 kHz
  - Sample rate 44,1 kHz (48 kHz in DAT)
  - Quantization with 16 bit values
  - Bitrate equal to 1,411 Mbps
    - Assuming stereo sound (two independent channels)

PCM types (3 of 6)

- Logarithmic quantization
  - Pass the signal through a log function
    - Linear quantization of the resulting signal
  - Reverse the process in the decoder
  - Uses analog circuit called a compander
    - Compressor + expander = compander
  - "Compresses" one end of the signal range
  - Allows for higher resolution on the other end

PCM types (4 of 6)



- Signal compression (μ-law)

PCM types (5 of 6)

- Why emphasize one end of the range?
  - Either the signal has more information there
  - Or our perception is more sensitive there
- Our ears hear “logarithmically”
  - As the volume grows higher…
  - …we can only distinguish larger differences
  - So, we need more accuracy at lower volumes

PCM types (6 of 6)

- Voice telephony: ITU G.711 standard
  - Logarithmic quantization: compress higher end
    - Modify signal before / after quantization
    - A-law (Europe), μ-law (US / Japan)
  - Signal range 3.1-3.5 kHz, 8 kHz sampling
  - 8 bit samples (or 7 bit)
    - Equivalent to 12-14 bit with linear (for voice)
  - 64 Kbps bit rate

ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ

ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

Symbolic representation and sound synthesis

Class: Multimedia Technology, **Section # 7:** Audio

Instructor: George Xylomenos, **Department:** Informatics

The MIDI standard (1 of 5)

- Music is not random sound
  - It consists of instruments playing notes
  - Musical scores are a symbolic notation
  - Each symbol corresponds to a sound
- Musical Instruments Digital Interface (MIDI)
  - A specification for connecting instruments
  - Can also be used as a music representation

The MIDI standard (2 of 5)

- What does MIDI cover?
  - Interconnection specifications
    - Connectors, electrical signals
  - Data formatting over the channel
- Allows communication between instruments
  - Keyboards (controller)
  - Synthesizers (sound modules)
  - Storage (sequencers)

The MIDI standard (3 of 5)

- MIDI messages describe events
  - Actions that musicians perform
  - Press or release a specific key
  - Pitch and amplitude of the key
  - Each message refers to a channel
  - Devices are connected in daisy-chain mode
  - Each device filters its own messages (channels)
  - There are also system wide messages

The MIDI standard (4 of 5)

- 16 event channels
  - Can control one or more synthesizers
  - Monophonic or polyphonic synthesizers
- Each channel must be mapped to a sound
  - The actual sound is called a patch
    - Terminology from modular synthesizers
  - General MIDI: 128 predefined instruments
    - Each synth tries to produce reasonable sounds

The MIDI standard (5 of 5)

- Advantages and disadvantages
  - Very compact representation
  - Allows symbolic processing
  - The output depends on the hardware
    - The “piano” can be wildly different
  - Suitable only for musical instruments
    - Assumes the production of musical notes
    - Does not cover (say) the human voice

Sound synthesis (1 of 4)

- How can we create sounds electronically?
  - From a MIDI sequence in a computer?
- Sounds synthesis predates computing
  - Started with analog synthesizers
  - Digital synthesis followed
    - Initially for instruments, then for computers
  - Evolved into sampling
    - When memory and processors allowed it

Sound synthesis (2 of 4)

- Subtractive synthesis
  - Starts with an oscillator producing a signal
    - Sine wave, square wave, sawtooth wave
  - Applies a set of filters to the signal
    - High pass, low pass, bandpass
  - Combines additional oscillators
    - Low Frequency Oscillator (LFO)
  - Modular synths combine these with patch cables

Sound synthesis (3 of 4)

- Frequency modulation synthesis
  - Two (at least) sine waves
  - One wave modulates the other
  - Waves can have different phases
  - More waves can be combined
  - May have polyphony
  - This takes place entirely digitally (DSP)
    - So, it can be implemented in computer

Sound synthesis (4 of 4)

- Sampling
  - Starts with sampling an instrument
    - Either modulate the sample to create notes
    - Or store separate samples per instrument
  - A synth can be based around a sampler
  - Requires memory and processing power
  - But it can reproduce any sound
  - Sound cards today are basically samplers

End of Section # 7

Class: Multimedia Technology, **Section # 7:** Audio

Instructor: George Xylomenos, **Department:** Informatics