Multimedia Coding & Communications

Course Outline

- **Multimedia Overview**
- **Component of Multimedia**
- **Lossless Compression Techniques**
- **Lossy Compression Techniques**
- **Elements of Image Compression System and Standards**
- **Video Coding and Compressing Standards**
- **Audio Compression Standards**
- **Multimedia Communication and Retrieval**
- **Multimedia Architecture**

Course Outline

- **Multimedia Overview**
- **Component of Multimedia**
  - Text – types, Unicode standard on file format
  - Image and graphics - data types, file formats, color science and color model;
  - **Audio- digitization, midi, quantization and transformation of audio;**
  - Video- types of video signals, analog and digital video, television broadcast standards, pc video;
  - animation- types, principals and techniques, 3D animation, camera, special effects, rendering

Sub-course Outline

- Sound
- Digitization of Sound
- Quantization of Audio
  - Nyquist Theorem
  - Signal to Noise Ratio (SNR)
  - Signal to Quantization Noise Ratio (SQNR)
  - Linear and Non-linear Quantization
  - Audio Filtering
  - Audio Quality vs. Data Rate
  - Synthetic Sounds
- Midi
  - MIDI Overview
  - MIDI Concepts
  - Hardware Aspects of MIDI
  - Structure of MIDI Messages
    - Channel messages
    - System messages
  - General MIDI
  - MIDI to WAV Conversion
- Transformation of audio
  - Pulse Code Modulation
  - PCM in Speech Compression
  - Differential Coding of Audio
  - Lossless Predictive Coding
  - Differential Pulse Code Modulation (DPCM)
  - Delta Modulation (DM)
  - Adaptive DPCM

SOUND

Sound

- **What is Sound?**

Sound is a wave phenomenon like light, but is macroscopic and involves molecules of air being compressed and expanded under the action of some physical device.

(a) For example, a speaker in an audio system vibrates back and forth and produces a *longitudinal* pressure wave that we perceive as sound.

(b) Since sound is a pressure wave, it takes on continuous values, as opposed to digitized ones.

Sound

c) Even though such pressure waves are longitudinal, they still have ordinary wave properties and behaviors, such as reflection (bouncing), refraction (change of angle when entering a medium with a different density) and diffraction (bending around an obstacle).

d) If we wish to use a digital version of sound waves we must form digitized representations of audio information.

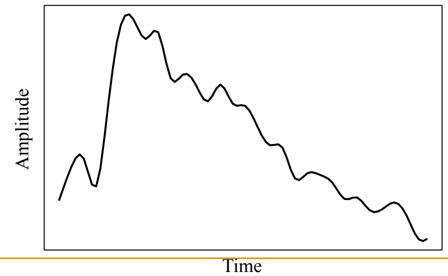
Sub-course Outline

- Sound
- Digitization of Sound
- Quantization of Audio
  - Nyquist Theorem
  - Signal to Noise Ratio (SNR)
  - Signal to Quantization Noise Ratio (SQNR)
  - Linear and Non-linear Quantization
  - Audio Filtering
  - Audio Quality vs. Data Rate
  - Synthetic Sounds
- Midi
  - MIDI Overview
  - MIDI Concepts
  - Hardware Aspects of MIDI
  - Structure of MIDI Messages
    - Channel messages
    - System messages
  - General MIDI
  - MIDI to WAV Conversion
- Transformation of audio
  - Pulse Code Modulation
  - PCM in Speech Compression
  - Differential Coding of Audio
  - Lossless Predictive Coding
  - Differential Pulse Code Modulation (DPCM)
  - Delta Modulation (DM)
  - Adaptive DPCM

DIGITIZATION OF SOUND

Digitization of Sound

- **Digitization** means conversion to a stream of numbers, and preferably these numbers should be integers for efficiency.
- Below Figure shows the 1-dimensional nature of sound:
  - **amplitude** values depend on a 1D variable, time.



An analog signal: continuous measurement of pressure wave.

Sub-course Outline

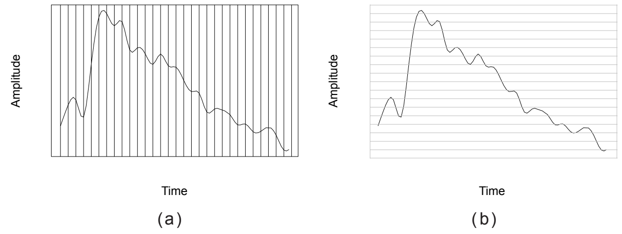
- Sound
- Digitization of Sound
- Quantization of Audio
  - Nyquist Theorem
  - Signal to Noise Ratio (SNR)
  - Signal to Quantization Noise Ratio (SQNR)
  - Linear and Non-linear Quantization
  - Audio Filtering
  - Audio Quality vs. Data Rate
  - Synthetic Sounds
- Midi
  - MIDI Overview
  - MIDI Concepts
  - Hardware Aspects of MIDI
  - Structure of MIDI Messages
    - Channel messages
    - System messages
  - General MIDI
  - MIDI to WAV Conversion
- Transformation of audio
  - Pulse Code Modulation
  - PCM in Speech Compression
  - Differential Coding of Audio
  - Lossless Predictive Coding
  - Differential Pulse Code Modulation (DPCM)
  - Delta Modulation (DM)
  - Adaptive DPCM

QUANTIZATION OF AUDIO

Quantization of Audio

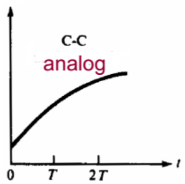
- To digitize, the signal must be **sampled** in each dimension: in time, and in amplitude.
  - (a) Sampling means measuring the quantity we are interested in, usually at evenly-spaced intervals.
  - (b) The first kind of sampling, using measurements only at evenly spaced time intervals, is simply called, *sampling*.
  - (c) The rate at which it is performed is called the *sampling frequency* (see the next Figure (a)).
  - (d) For audio, typical sampling rates are from 8 kHz (8,000 samples per second) to 48 kHz. This range is determined by Nyquist theorem.
  - (e) Sampling in the amplitude or voltage dimension is called **quantization**. Next Figure (b) shows this kind of sampling.

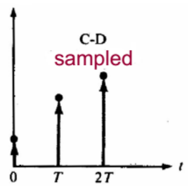
Quantization of Audio



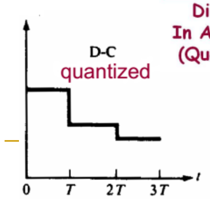
Sampling and Quantization.
(a): Sampling the analog signal in the time dimension.
(b): Quantization is sampling the analog signal in the amplitude dimension.

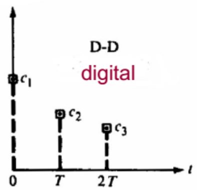
Computer-Controlled System Signals

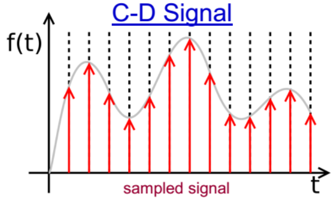


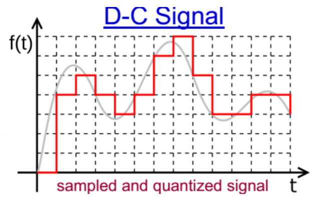


	Continuous In Time	Discrete In Time (Sampled)
Continuous In Amplitude	C-C	C-D
Discrete In Amplitude (Quantized)	D-C	D-D





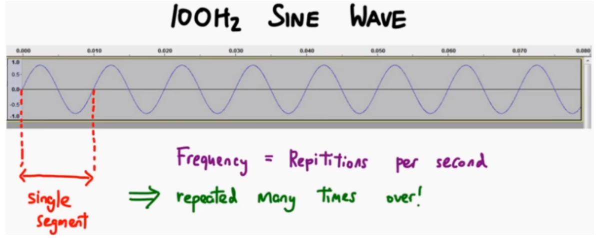




Quantization of Audio

- Thus to decide how to digitize audio data we need to answer the following questions:
  1. What is the sampling rate?
  2. How finely is the data to be quantized, and is quantization uniform?
  3. How is audio data formatted? (file format)

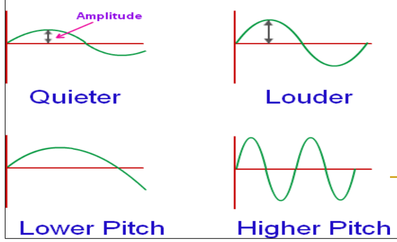
Basic Terms - Frequency

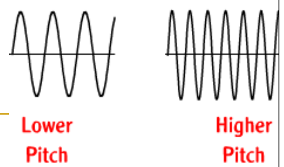


- **1 Hertz = 1 vibration/second or** 440 Hz = 440 cycles of this per second
- The frequency of a wave refers to how often the particles of the medium vibrate when a wave passes through the medium.
- The frequency of a wave is measured as the number of complete back-and-forth vibrations of a particle of the medium per unit of time.

Basic Terms - Pitch

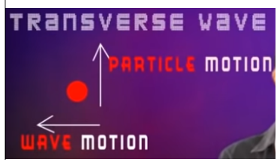
- The sensation of a frequency is commonly referred to as the **pitch** of a sound.
- A high pitch sound corresponds to a high frequency sound wave and a low pitch sound corresponds to a low frequency sound wave.
- Since **pitch** is such a close proxy for frequency, it is almost entirely determined by how quickly the **sound** wave is making the air vibrate and has almost nothing to do with the intensity, or amplitude, of the wave.
- "high" **pitch** means very rapid oscillation, and "low" **pitch** corresponds to slower oscillation.

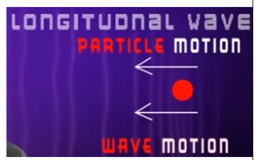


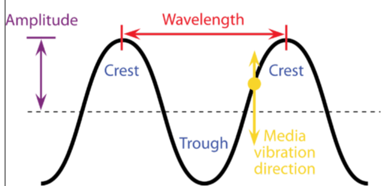


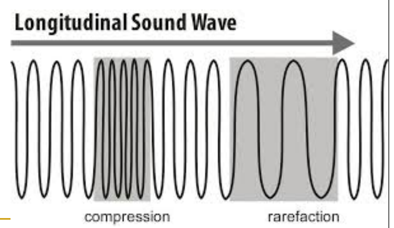
Basic Terms - Wave

- Particle goes up and down and wave goes forwards



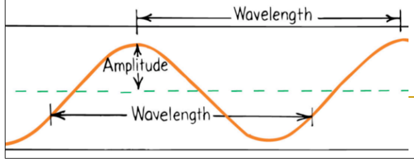






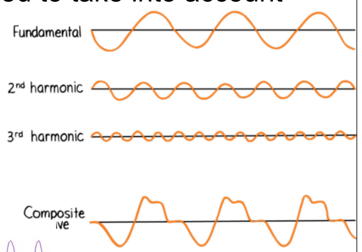
Basic Terms – Sign Wave

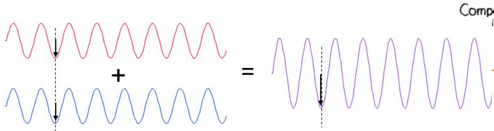
- Any analog signal consists of components at various frequencies.
- The simplest case is the sine wave, in which all the signal energy is concentrated at one frequency.
- In practice, analog signals usually have complex waveforms, with components at many frequencies.
- The highest frequency component in an analog signal determines the bandwidth of that signal.
- The higher the frequency, the greater the bandwidth, if all other factors are held constant.



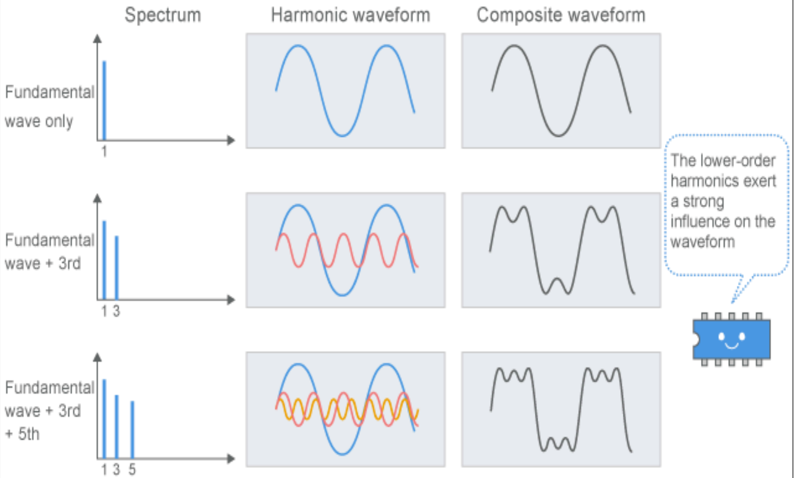
Basic Terms - Composite Signal/ Wave

- When we combine two waves to form composite waves, it is the algebraic sum of the two original waves, point by point in space.
- When we the two waves we need to take into account their:
  - Direction
  - Amplitude and
  - Phase



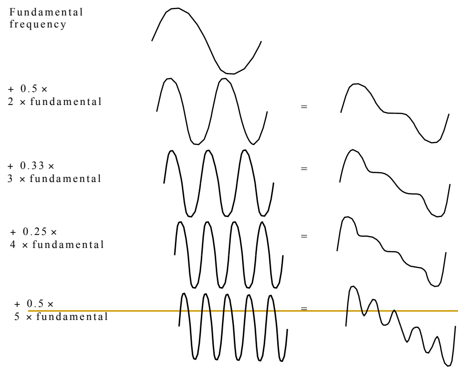


Basic Terms - Composite Signal/ Wave



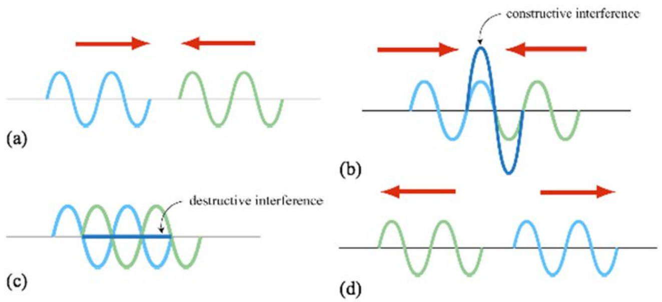
Quantization of Audio

- Below Figure shows how weighted sinusoids/sign wave can build up quite a complex signal.

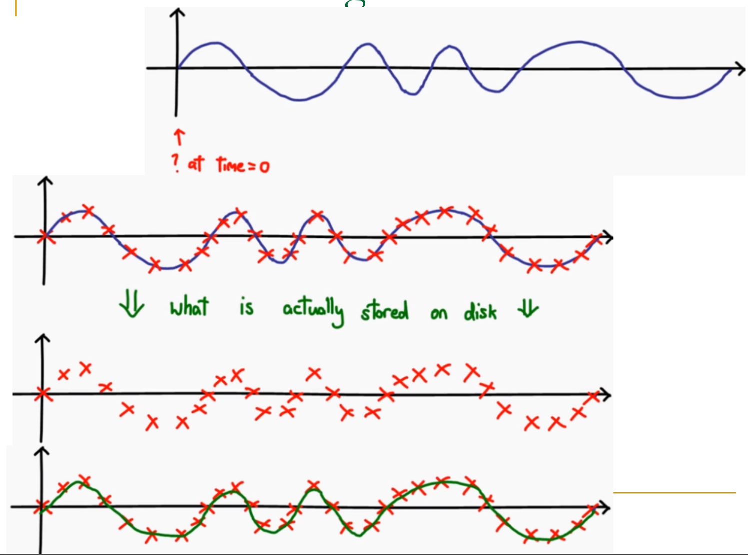


Building up a complex signal by superposing sinusoids

Interference is the combination of two or more waves to form a composite wave, based on the principle of superposition



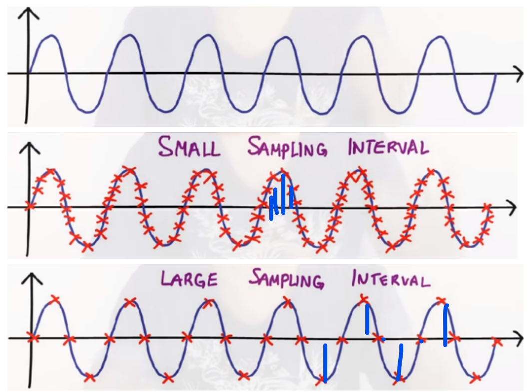
Basic Terms – Digitization of Audio



Quantization of Audio

- For analog-to-digital conversion (ADC) to result in a faithful reproduction of the signal, slices, called *samples*, of the analog waveform must be taken frequently.
- The number of samples per second is called the **sampling rate** or **sampling frequency**.

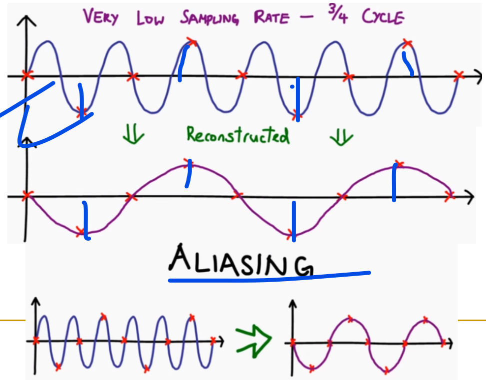
Basic Terms - Sampling



Basic Terms – Aliasing

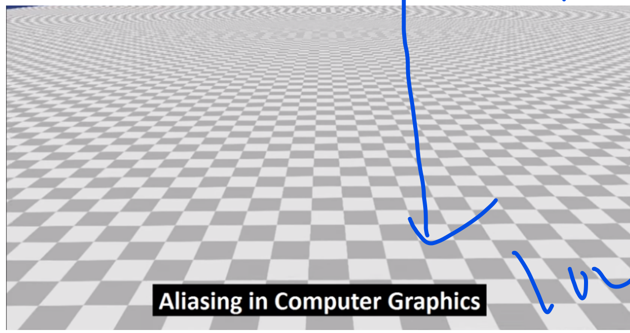
High Low

- High frequency signal appearing low frequency after sampling at a sampling rate that is too low.



Basic Terms – Aliasing

high



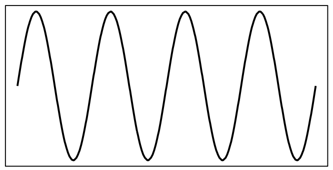
low

Quantization of Audio

- Nyquist Theorem
  - Signals can be decomposed into a sum of sinusoids or sign wave.
- It is also known as **Sampling Theorem. It** states that
  - An analog signal waveform can be converted to digital format and be reconstructed without error from samples taken at equal time intervals if the sampling rate is equal to, or greater than, twice the highest frequency component in the analog signal.
- The Nyquist theorem forms the basis for pulse code modulation (PCM), the fundamental method for converting analog voice to digital format.

Quantization - Nyquist Theorem

- The Nyquist theorem states how frequently we must sample in time to be able to recover the original sound.
  - (a) Figure (a) shows a single sinusoid: it is a single, pure, frequency (only electronic instruments can create such sounds).

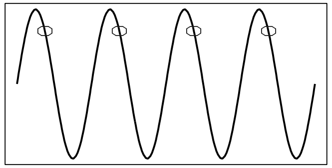


Aliasing. (a): A single frequency.

(a)

Quantization - Nyquist Theorem

- The Nyquist theorem states how frequently we must sample in time to be able to recover the original sound.



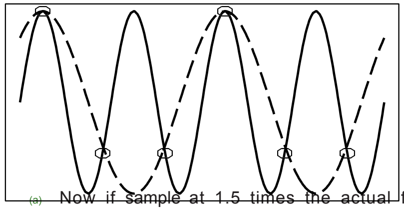
(b)

Aliasing: (b): Sampling at exactly the frequency produces a constant.

(a) If sampling rate just equals the actual frequency, Figure (b) shows that a false signal is detected: it is simply a constant, with zero frequency.

Quantization - Nyquist Theorem

- The Nyquist theorem states how frequently we must sample in time to be able to recover the original sound.



(c)

Aliasing: (c): Sampling at 1.5 times per cycle produces an *alias* perceived frequency.

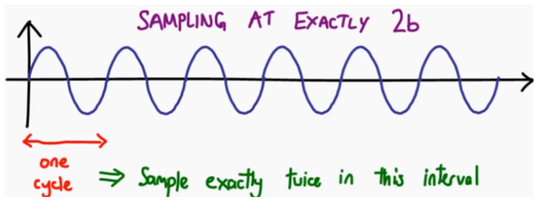
(a) Now if sample at 1.5 times the actual frequency, Figure (c) shows that we obtain an incorrect (**alias**) frequency that is lower than the correct one — it is half the correct one (the wavelength, from peak to peak, is double that of the actual signal).

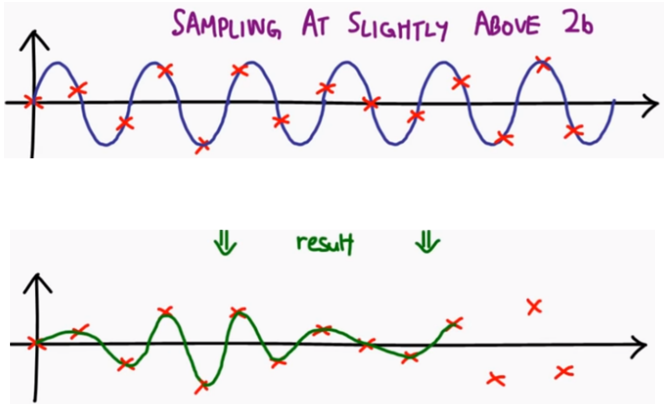
(b) Thus for correct sampling we must use a sampling rate equal to at least *twice the maximum frequency* content in the signal. This rate is called the **Nyquist rate**.

Quantization - Nyquist Theorem

- **Nyquist Theorem**: If a signal is **band-limited**, i.e., there is a lower limit f_1 and an upper limit f_2 of frequency components in the signal, then the sampling rate should be at least $2(f_2 - f_1)$.
- **Nyquist frequency**: half of the Nyquist rate.
  - Since it would be impossible to recover frequencies higher than Nyquist frequency in any event, most systems have an **antialiasing filter** that restricts the frequency content in the input to the sampler to a range at or below Nyquist frequency.
- The relationship among the Sampling Frequency, True Frequency, and the Alias Frequency is as follows:

$$F_{alias} = f_{sampling} - f_{true}, \quad for \quad f_{true} < f_{sampling} < 2 \times f_{true}$$





Nyquist Rate vs Nyquist Frequency

- Nyquist Rate: Sampling rate required for a frequency to not alias
- Minimum sampling rate to avoid aliasing problem
- The *Nyquist rate* is the minimum sampling rate that satisfies the Nyquist sampling criterion for a given signal or family of signals.
- The Nyquist rate is twice the maximum component frequency of the function being sampled.
- Fs = 2 x Fmax -> Nyquist Rate

Nyquist Rate vs Nyquist Frequency

- Nyquist Frequency: Maximum frequency that will not alias given a sampling rate
- The **Nyquist frequency**, is half of the sampling rate of a discrete signal processing system
- Fs/2 is the Nyquist frequency or Folding Frequency or Cutoff Frequency

- [-Fs/2, Fs/2] = Nyquist Interval

SIGNAL TO NOISE RATIO (SNR)

Signal to Noise Ratio (SNR)

- The ratio of the power of the correct signal and the noise is called the *signal to noise ratio* (**SNR**) — a measure of the quality of the signal.
- The SNR is usually measured in decibels (**dB**), where 1 dB is a tenth of a **bel**.
- The SNR value, in units of dB, is defined in terms of base-10 logarithms of squared voltages, as follows:

$$SNR = 10 \log_{10} \frac{V_{signal}^2}{V_{noise}^2} = 20 \log_{10} \frac{V_{signal}}{V_{noise}}$$

Signal to Noise Ratio (SNR)

a) The power in a signal is proportional to the square of the voltage.

b) For example, if the signal voltage V_{signal} is 10 times the noise, then the SNR is 20 $*\log_{10}(10)$ = 20dB.

c) In terms of power, if the power from ten violins is ten times that from one violin playing, then the ratio of power is 10dB, or 1B.

Signal to Noise Ratio (SNR)

- The usual levels of sound we hear around us are described in terms of decibels, as a ratio to the quietest sound we are capable of hearing.
- Below table shows approximate levels for these sounds.

Threshold of hearing	0
Rustle of leaves	10
Very quiet room	20
Average room	40
Conversation	60
Busy street	70
Loud radio	80
Train through station	90
Riveter	100
Threshold of discomfort	120
Threshold of pain	140
Damage to ear drum	160

Magnitude levels of common sounds, in decibels

AUDIO FILTERING

Audio Filtering

- Prior to sampling and AD conversion, the audio signal is also usually *filtered* to remove unwanted frequencies.
- The frequencies kept depend on the application:
  (a) For speech, typically from 50Hz to 10kHz is retained, and other frequencies are blocked by the use of a **band-pass filter** that screens out lower and higher frequencies.
  (b) An audio music signal will typically contain from about 20Hz up to 20kHz.
  (c) At the DA converter end, high frequencies may reappear in the output — because of sampling and then quantization, smooth input signal is replaced by a series of step functions containing all possible frequencies.
  (d) So at the decoder side, a **lowpass** filter is used after the DA circuit.

Audio Quality vs. Data Rate

- The uncompressed data rate increases as more bits are used for quantization.
- Stereo: double the bandwidth. to transmit a digital audio signal.

Quality	Sample Rate (KHz)	Bits per Sample	Mono/ Stereo	Data Rate (uncompressed) (kB/sec)	Frequency Band (KHz)
Telephone	8	8	Mono	8	0.200-3.4
AM Radio	11.025	8	Mono	11.0	0.1-5.5
FM Radio	22.05	16	Stereo	88.2	0.02-11
CD	44.1	16	Stereo	176.4	0.005-20
DAT	48	16	Stereo	192.0	0.005-20
DVD Audio	192 (max)	24 (max)	6 channels	1,200.0 (max)	0-96 (max)

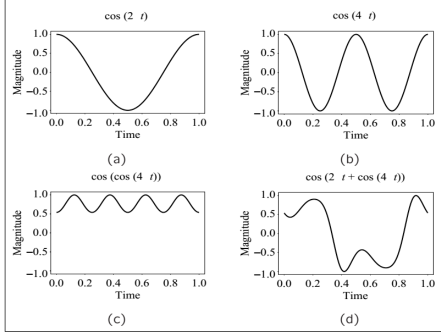
Data rate and bandwidth in sample audio applications

SYNTHETIC SOUND

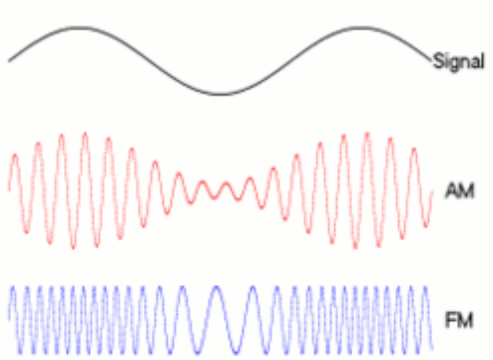
Synthetic Sounds

- **FM** (Frequency Modulation): one approach for generating synthetic sound:

$$x(t) = A(t)\cos[\omega_c \pi t + I(t)\cos(\omega_m \pi t + \varphi_m) + \varphi_c]$$



Frequency Modulation.
(a): A single frequency.
(b): Twice the frequency.
(c): Usually, FM is carried out using a sinusoid argument to a sinusoid.
(d): A more complex form arises from a carrier frequency, $2\pi t$ and a modulating frequency $4\pi t$ cosine inside the sinusoid.



Synthetic Sounds

- 2. **Wave Table synthesis**: A more accurate way of generating sounds from digital signals. Also known as **sampling**.
  - In this technique, the actual digital samples of sounds from real instruments are stored.
  - Since wave tables are stored in memory on the sound card, they can be manipulated by software so that sounds can be combined, edited, and enhanced.
  - The sound of an existing instrument (a single note) is sampled and parsed into a sequence of circular **tables** of samples or wavetables, each having one period or cycle per **table**.

Wave Table synthesis

- A set of wavetables with user specified harmonic content can also be generated mathematically.
- Upon playback, these wavetables are used to fetch samples (table-lookup) in the same manner as in a numerically-controlled oscillator to produce a waveform.
- In wavetable synthesis, the output waveform is not normally static and evolves slowly in time as one wavetable is mixed with another, creating a changing waveform.
- Looping occurs when the wavetable evolution is halted, slowed, or reversed in time.

Sub-course Outline

- Sound
- Digitization of Sound
- Quantization of Audio
  - Nyquist Theorem
  - Signal to Noise Ratio (SNR)
  - Signal to Quantization Noise Ratio (SQNR)
  - Linear and Non-linear Quantization
  - Audio Filtering
  - Audio Quality vs. Data Rate
  - Synthetic Sounds
- Midi
  - MIDI Overview
  - MIDI Concepts
  - Hardware Aspects of MIDI
  - Structure of MIDI Messages
    - Channel messages
    - System messages
  - General MIDI
  - MIDI to WAV Conversion
- Transformation of audio
  - Pulse Code Modulation
  - PCM in Speech Compression
  - Differential Coding of Audio
  - Lossless Predictive Coding
  - Differential Pulse Code Modulation (DPCM)
  - Delta Modulation (DM)
  - Adaptive DPCM