Audio Engineering I

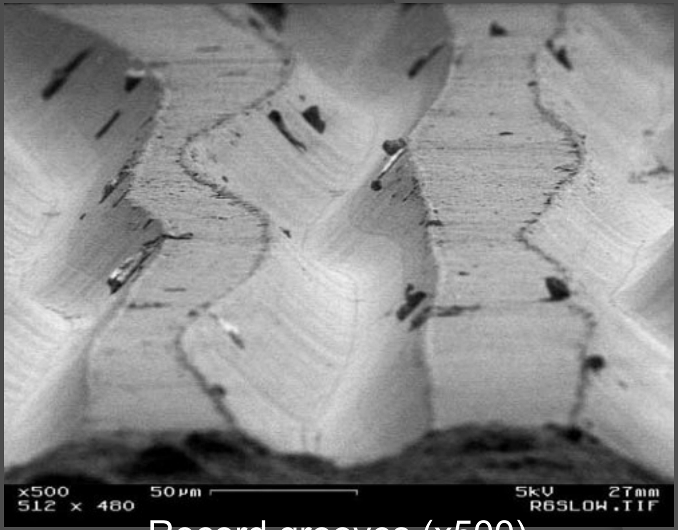
digital audio technology

binary number system

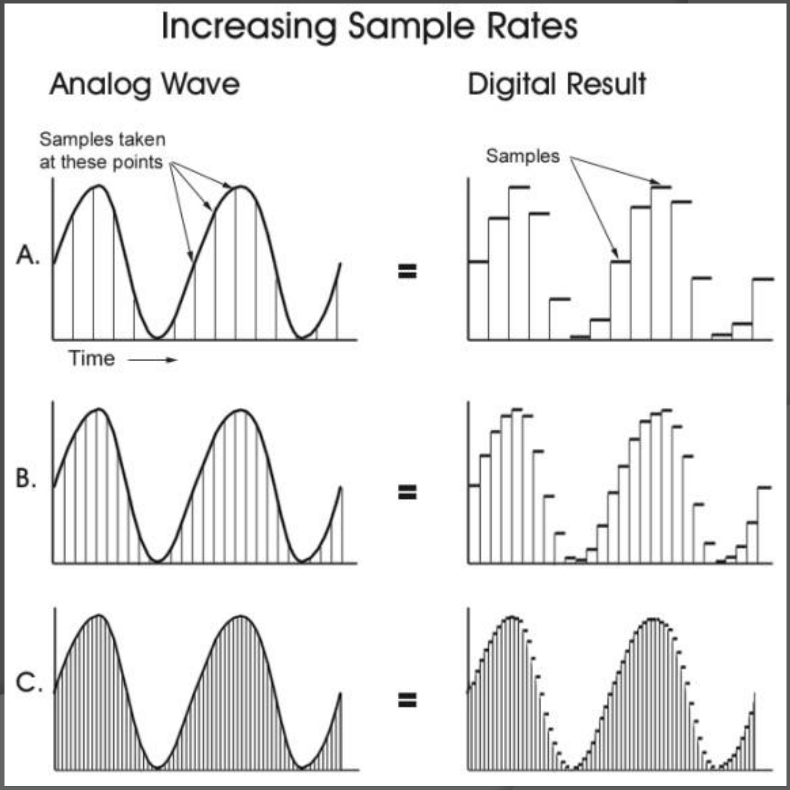
- Computers do everything in binary: a "language" that has an alphabet of only two characters: 0 and 1.
- One "bit" (***b**inary dig**it***) is one 0 or 1 value
- ASCII code: 7-bit (or later, 8-bit) code for characters
  - 010 0001 = !
  - 100 0001 = A
  - 011 0001 = 1
- How many different values are possible when using 2 bits? 3? 8?
  - 2 bits: 00, 01, 10, 11 = 4 possible values ($2^2 = 4$)
  - 3 bits: 000, 001, 010, 011, 100, 101, 110, 111 = 8 values ($2^3 = 8$)
  - 8 bits: $2^8 = 256$ values

sampling

- Analog systems mirror signals almost exactly ("analogously") to a physical medium.
- Digital systems take *samples* (i.e. snapshots) of the signal's voltage level at regular intervals.
- The ***sample rate*** is how frequently samples are taken.
  - CD audio standard SR = 44,100 Hz
  - Pro audio and film SR = 48,000 Hz
  - 96 and 192 Hz also!
- Sample and hold



Record grooves (x500)



quantization

- Quantization = rounding
- The translation of voltage levels (amplitude) into binary words
  - The more digits the system's binary word has, the more values possible to express the voltage level (i.e. higher resolution).
- **Bit Depth**
  - Higher bit depth = larger *dynamic range*
  - 16 bit = 2^{16} possible values = 65,536 different points on the amplitude axis
  - 24 bit = 2^{24} = 16,777,216 values
  - 32 bit = 4,294,967,296!
- Higher bit depth = lower *noise floor*

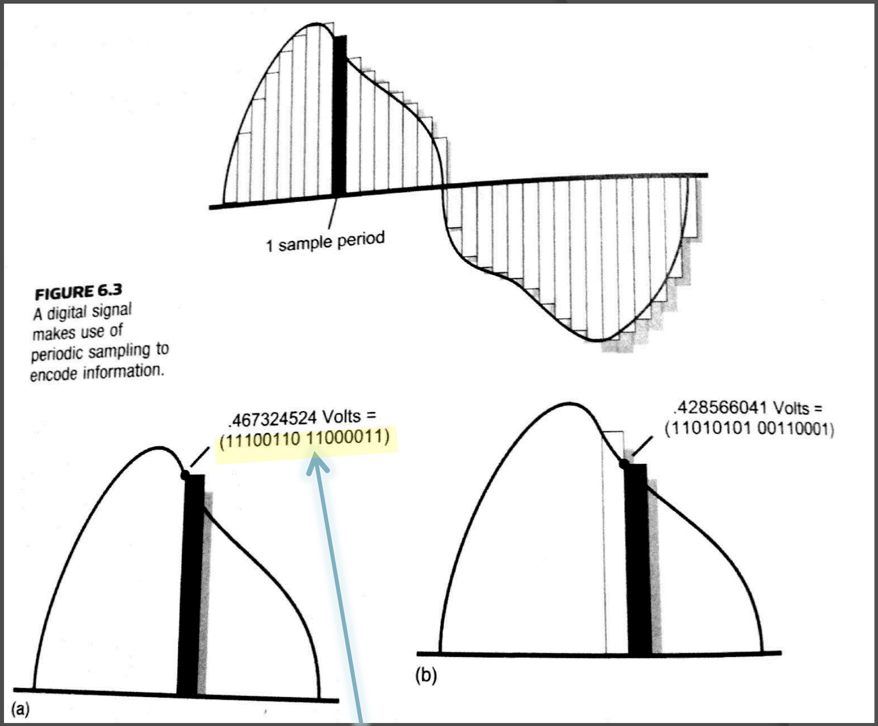
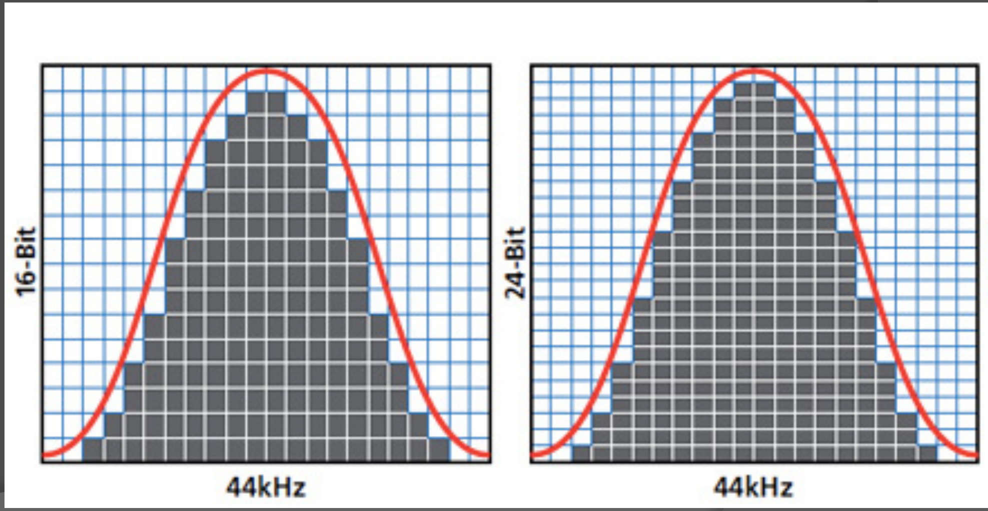


FIGURE 6.3
A digital signal makes use of periodic sampling to encode information.

16-bit word length



bit depth and dynamic range

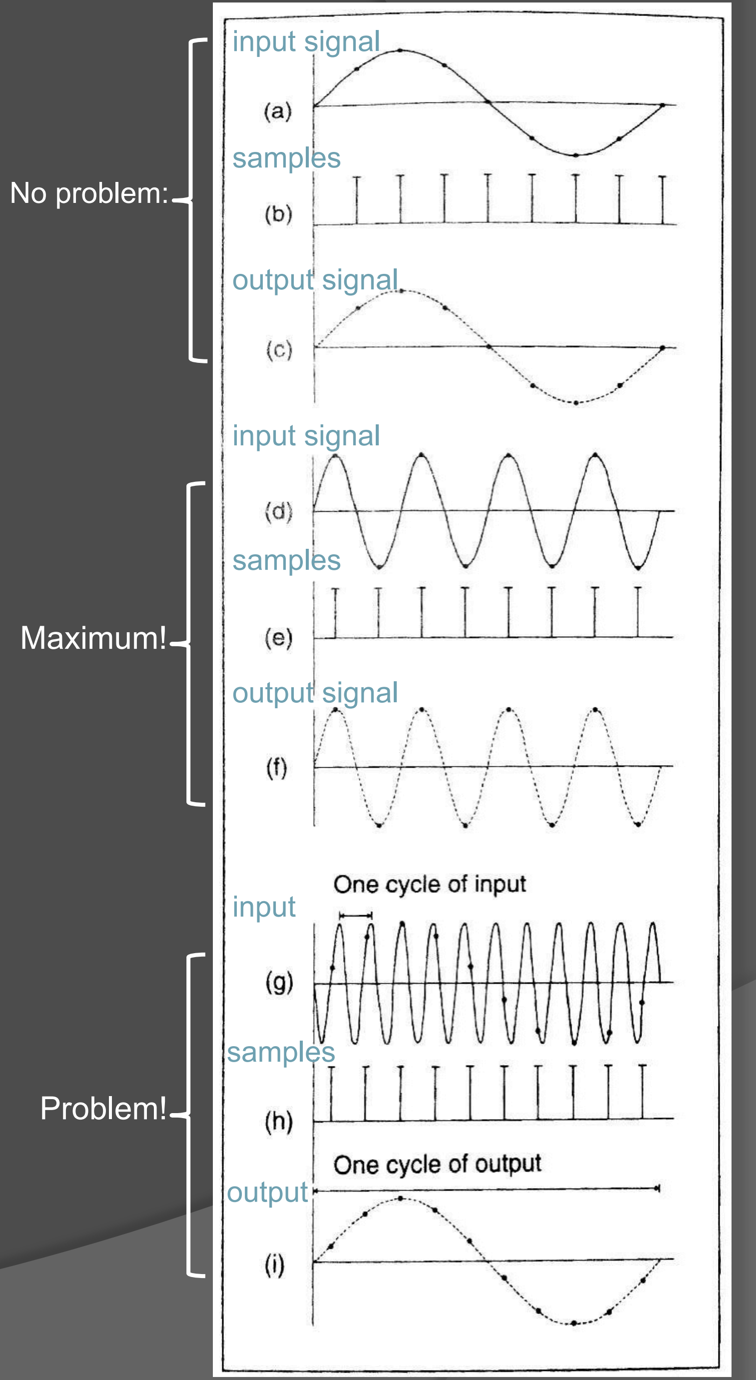
- Most popular music has a dynamic range of ~24 dB
- Even wildly dynamic classical music - ~60 dB range

- Early 78 RPM records: 30-40 dB range (above noise floor)
- Cassette tapes: 40 dB
- Pro audio tape: up to 110 dB (with saturation)
- 16 bit audio: 100-120 dB (noise floor is completely inaudible)
- 24 bit audio: 144 dB range!
  - The difference between barely perceptible rustling and a jet engine.
  - If you turn up the audio to the point that you can hear the quietest sound above the noise floor, the maximum peak above that would be so loud it could send you into a coma (theoretically!)
  - You can not hear the difference between 16-bit and 24-bit.
- 32 bit audio?!
  - Actually, 32-bit is just a different encoding of 24-bit. It's no different in terms of quality or dynamic range.
  - Why use 32-bit if there's no perceivable difference? If *lots* of channels will be added together, if extreme processing will be applied to the audio, or if you just want to worry less about keeping levels from clipping.

bit depth / sample rate demo

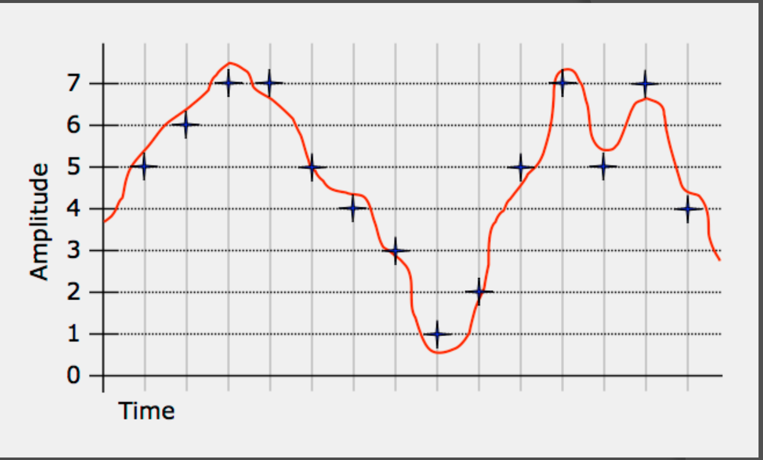
nyquist theorem

- The Nyquist Frequency is the highest frequency a digital system can handle.
  - Nyquist frequency = SR/2
  - At 44.1 kHz sample rate, the Nyquist frequency is 22.05 kHz.
- Any frequency higher than this can't be reproduced: the sampling actually causes a *lower* frequency: this is called aliasing or foldover
- Example of aliasing
- To prevent aliasing, a low-pass filter, cutting out everything above 22k, is placed before the analog to digital converter (ADC) for incoming signals.



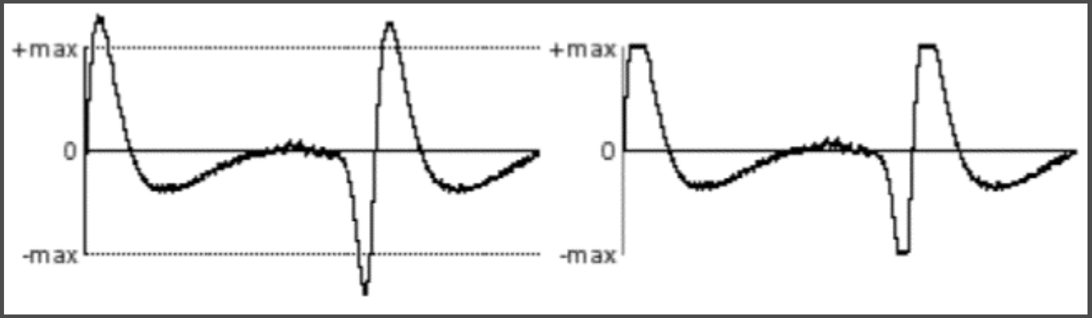
quantization error

- Because quantization rounds a signal's voltage levels (amplitude) to the nearest bit value, there will be some amount of *error* no matter what the bit rate. (Although in high bit rates, it's incredibly low)
- This error comes out as low-level noise added to the signal by the digital system.
  - Higher bit rate, = lower quantization error noise
  - "Noise floor"
- **Signal to Noise Ratio** – total amplitude range above the noise floor
  - ~6 dB per bit
  - 8 bit: ~48 dB; 16 bit: ~96 dB; 24 bit: ~144 dB



clipping

- If the amplitude of a signal goes above the system's maximum amplitude, it clips off the waveform.

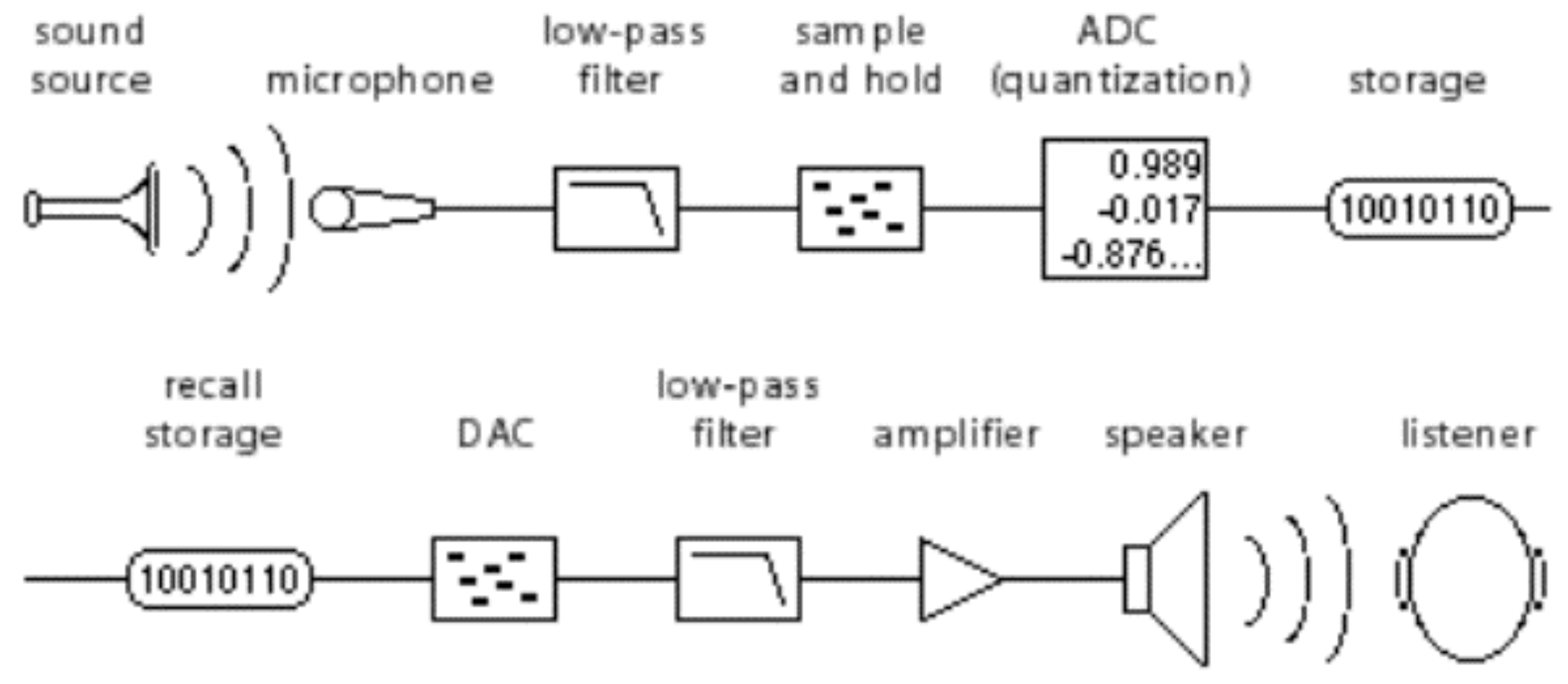


- Analog distortion is a warm, often desirable sound. Not so with digital!
- This can happen at ADC (i.e. microphones) or DAC (i.e. speakers)
- If a signal is recorded with clipping, it can't be fixed!

Audio Engineering I

digital audio technology

the digital recording chain



source: http://cycling74.com/docs/max5/tutorials/msp-tut/mspdigitalaudio.html

dither

- Because of quantization error, any digital signal has noise distortion added by the digital sampling process
  - The noise is present at the lowest-level bit.
- Dither minimizes the distortion by adding noise
  - Ironically, adding an ultra-low-level noise into the signal, the distortion caused by quantization error is removed.
- Usually done when converting from a higher bit depth (i.e. 24-bit) to a lower bit depth (i.e. 16-bit).

- Dithering takes place at the mastering stage, and usually is only done once.
- [Video demo]

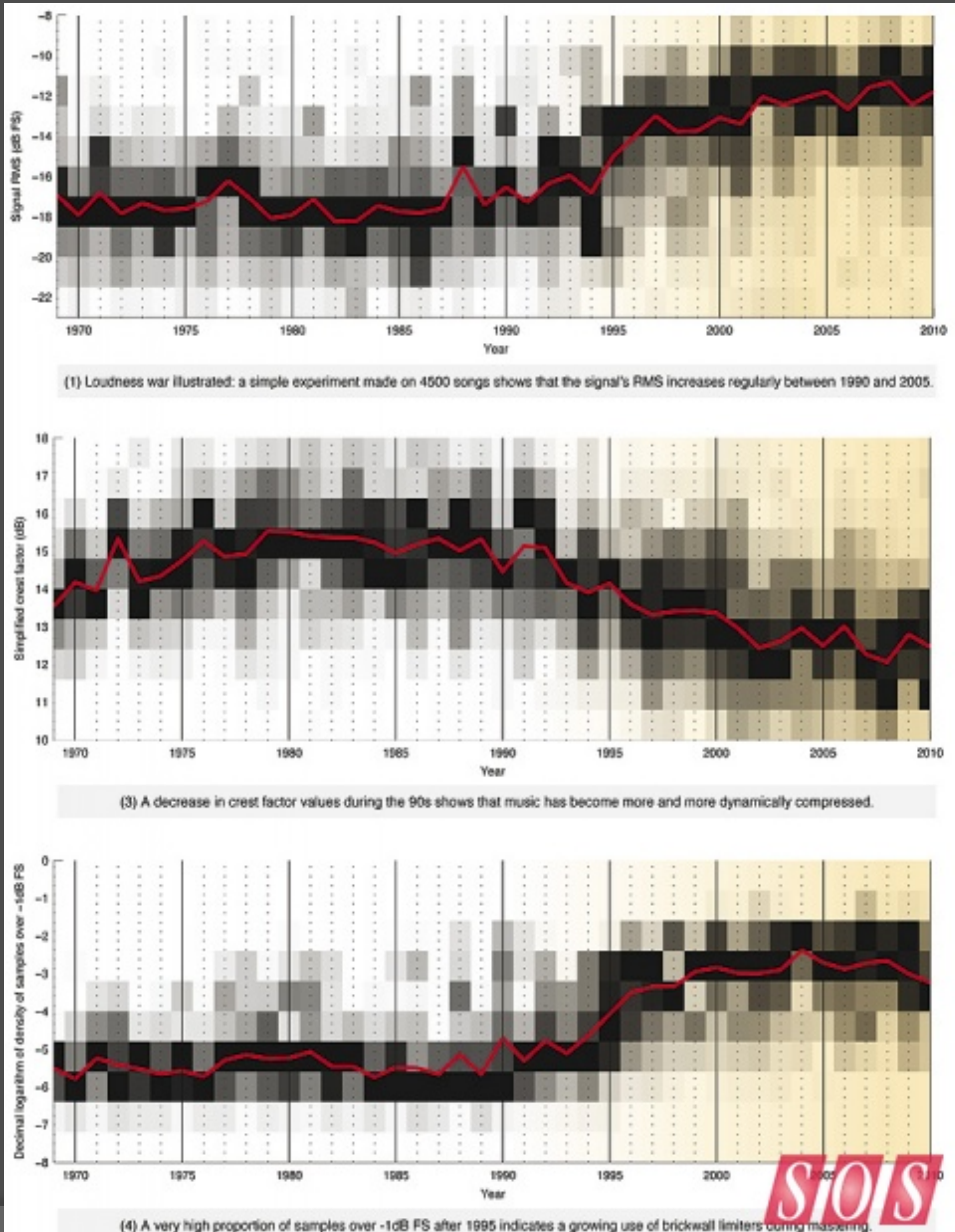
sample rates

- 32k: used in satellite broadcast communications
- 44.1k: CD-audio standard
  - The minimum pro audio sample rate
  - 20k bandwidth
- 48k: Standard rate for pro audio applications and film/DVD production
- 88.2k: double of 44.1k. High resolution purposes.
- 96k: double of 48k. Standard sample rate for high resolution recordings
- 192k: double of 96k. Very uncommon

- **The higher the sample rate, the larger the file size!**

the loudness war

- In the late 1990s and on, a trend began in the music industry: mixing music at the maximum possible loudness
  - Compressing/limiting so much that the whole mix hovers near the maximum level
  - Producers, record companies, radio, etc. pressured artists, engineers, etc. for a bigger sound that would grab more attention
  - Musical dynamic range suffers
- Video – 1989 vs. 2007
- Video – What famous engineers say



4500 songs from 1970 to 2010: Average RMS value

Average “crest factor”: the difference between RMS level and peak level over the course of a song. Measures the amplitude of the peaks in the audio stream.

Proportion of samples over -1 dbFS

Frequency and EQ