MP3s – Where Did All the Bits Go?

Wu-chi Feng
Intel Systems and Networking Lab
Portland State University

PORTLAND STATE UNIVERSITY

Digital Music

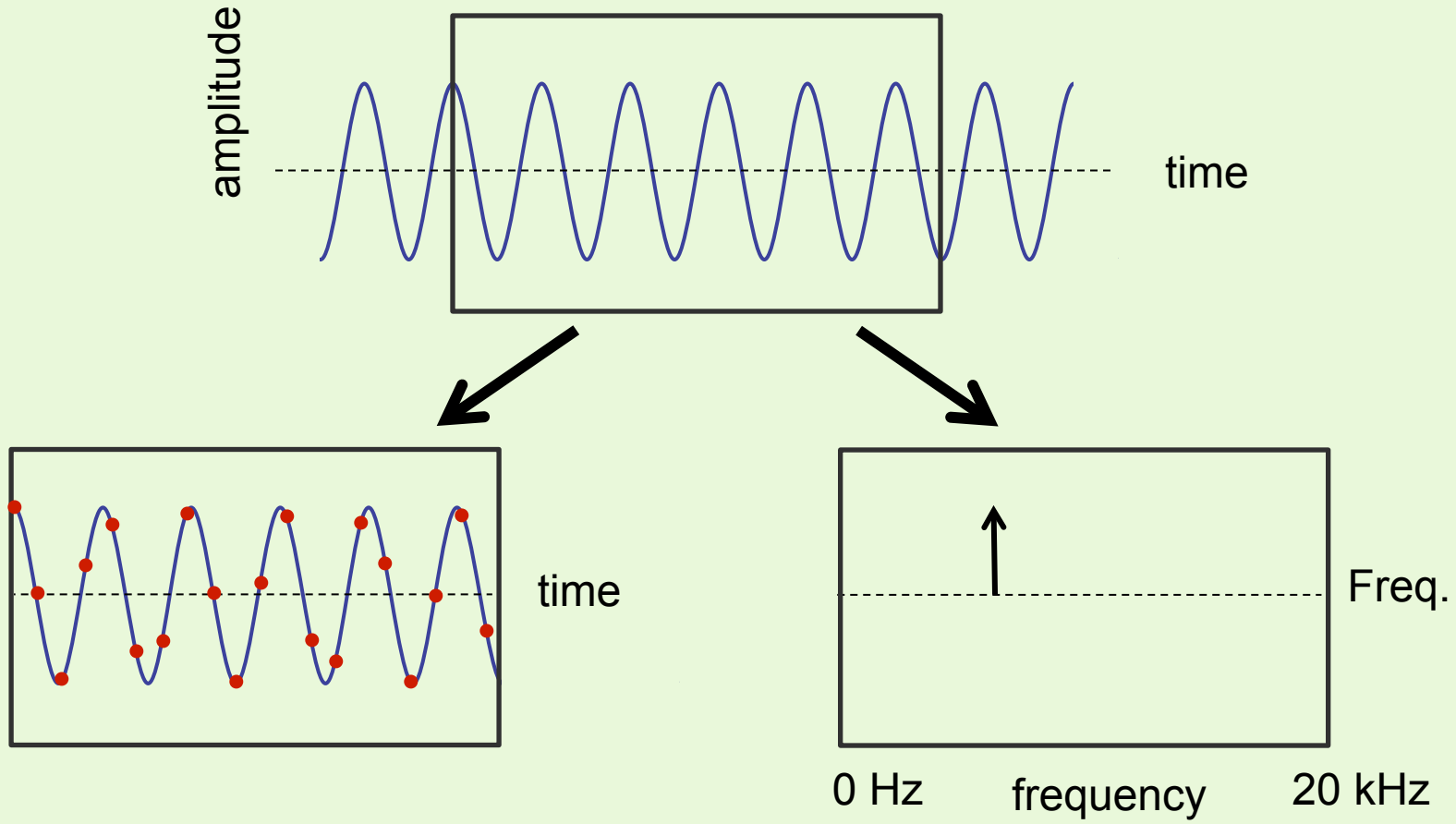
- The *gold standard* – CD Audio format
  - 44,100 samples per second (Hz)
  - 16-bit samples
  - 2 channels for stereo
  - 1.41 Mbits / second
- MP3 format
  - 64 – 256 kbits / second
  - Where did all the bits go?

Background

- *Lossless* compression
  - Examples: WinZip, gzip, etc…
  - 2-4 :1 compression maximum
  - Audio does not compress and will yield ~706 kbits / second
- *Lossy* compression
  - Purposely *throw away* data
  - For audio, images, and video – minimized perceived loss in playback / display

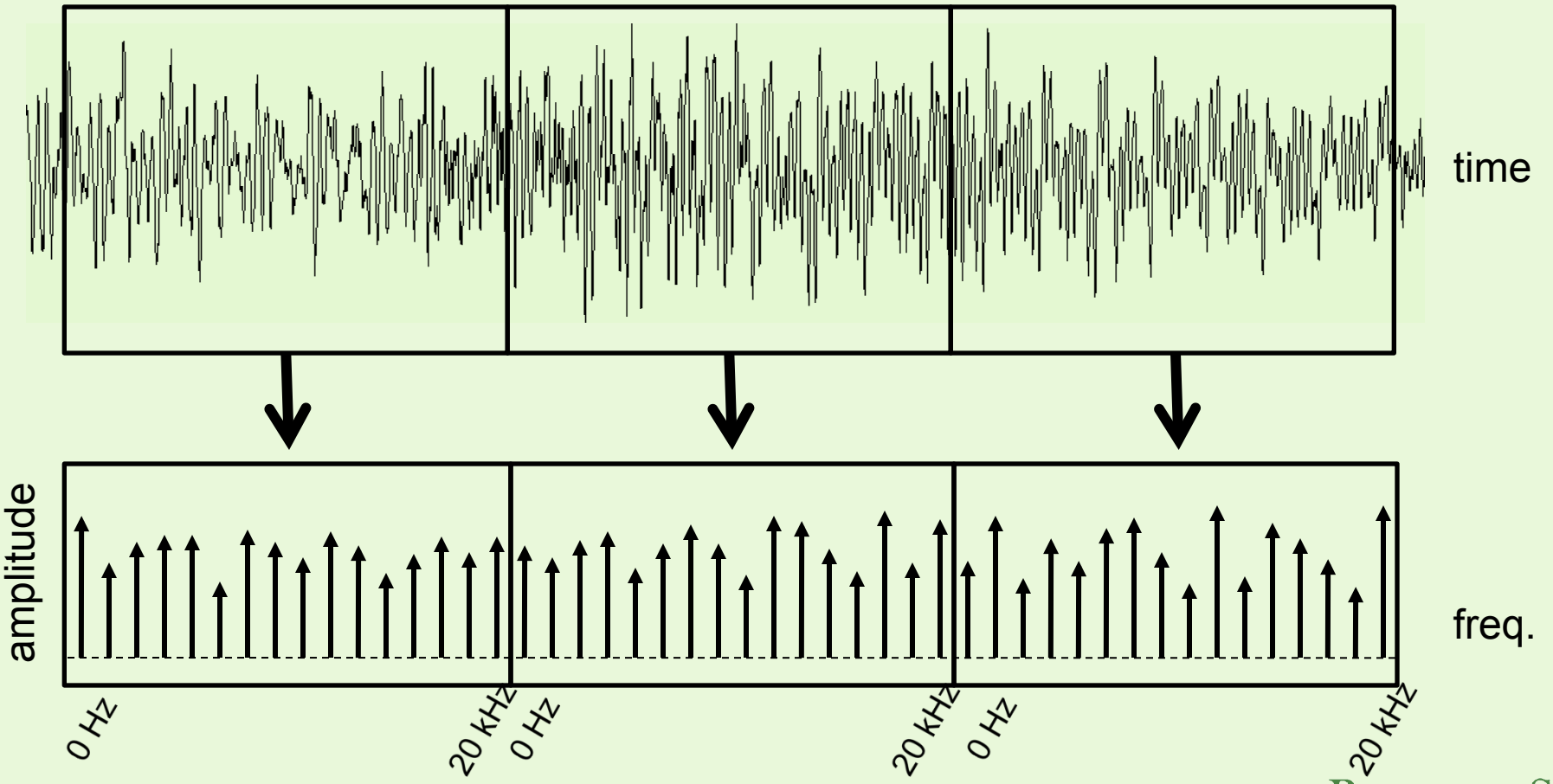
Background

- Sound can sometimes be represented more compactly



Background

- Generalizing to audio

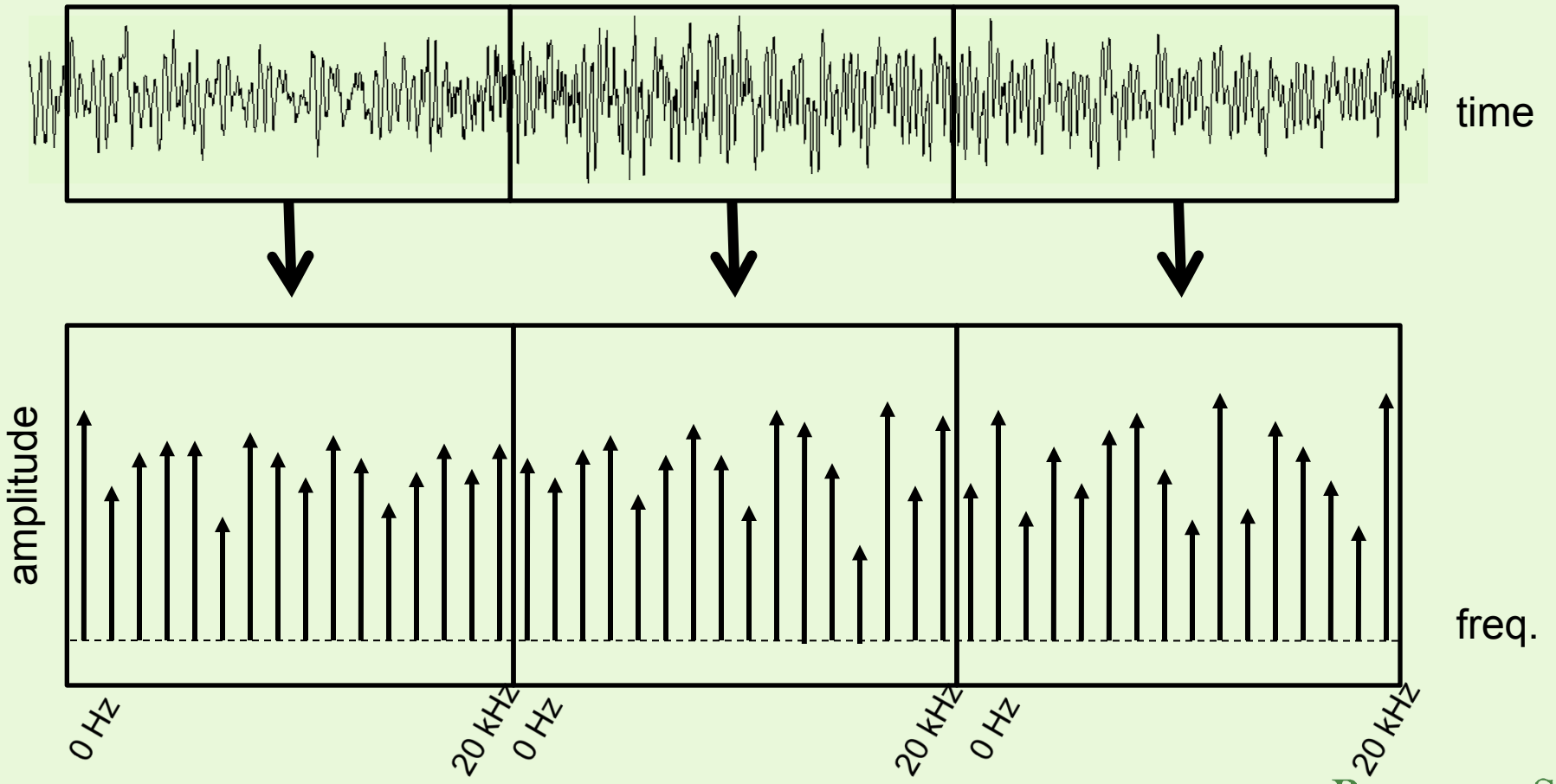


MP3 Compression

- Major Components
  - Frequency domain transform
  - Remove signals that are *perceptually irrelevant*
  - Entropy encoding

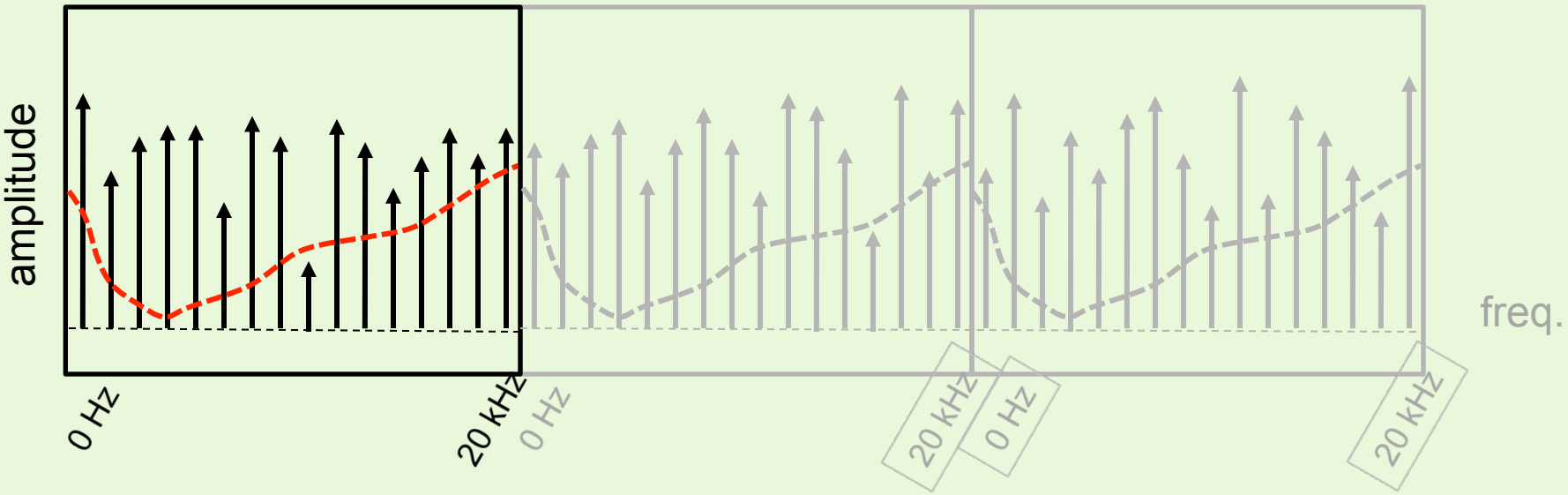
MP3 Compression

- Frequency domain transform



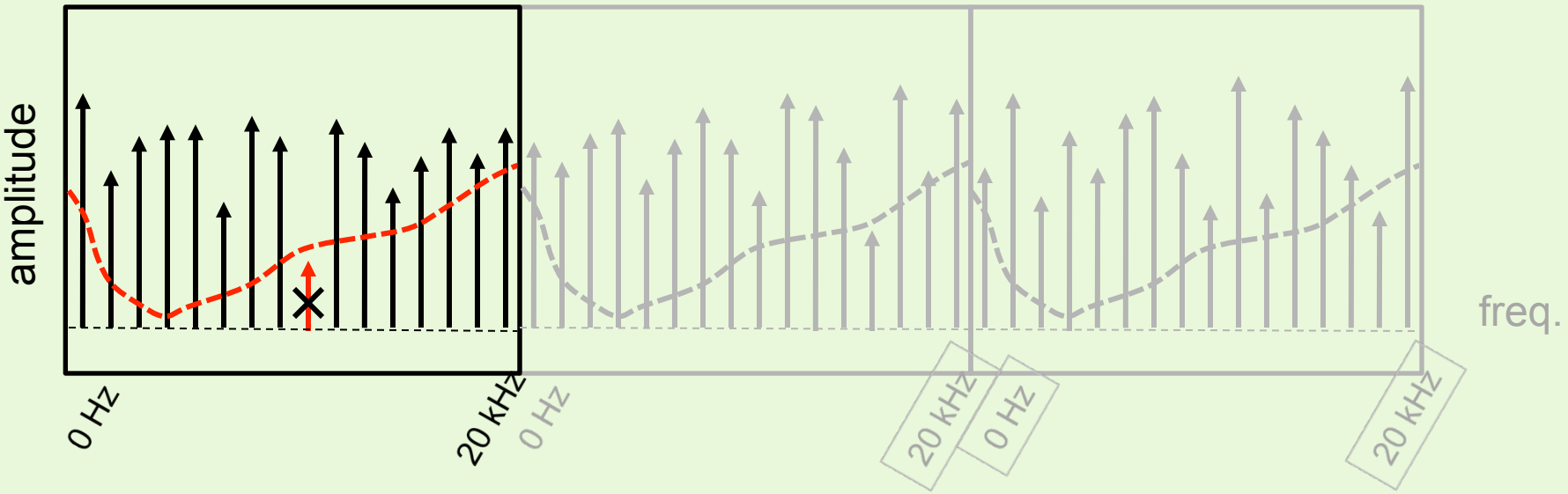
MP3 Compression

- Remove signals that cannot be heard
  Threshold of Quiet



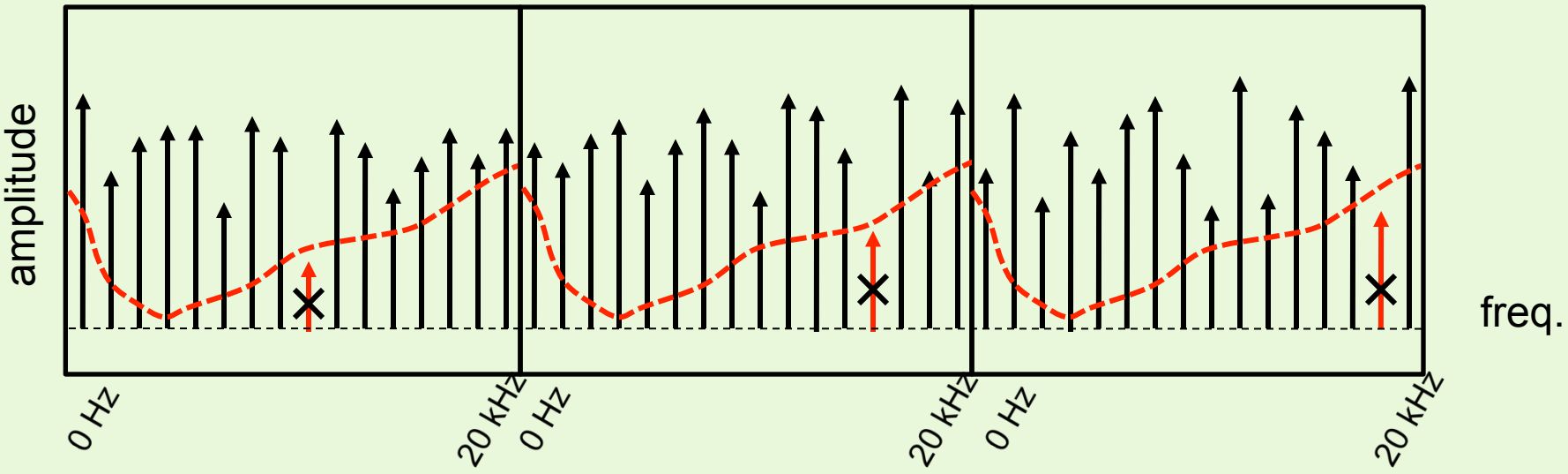
MP3 Compression

- Remove signals that cannot be heard
  Threshold of Quiet



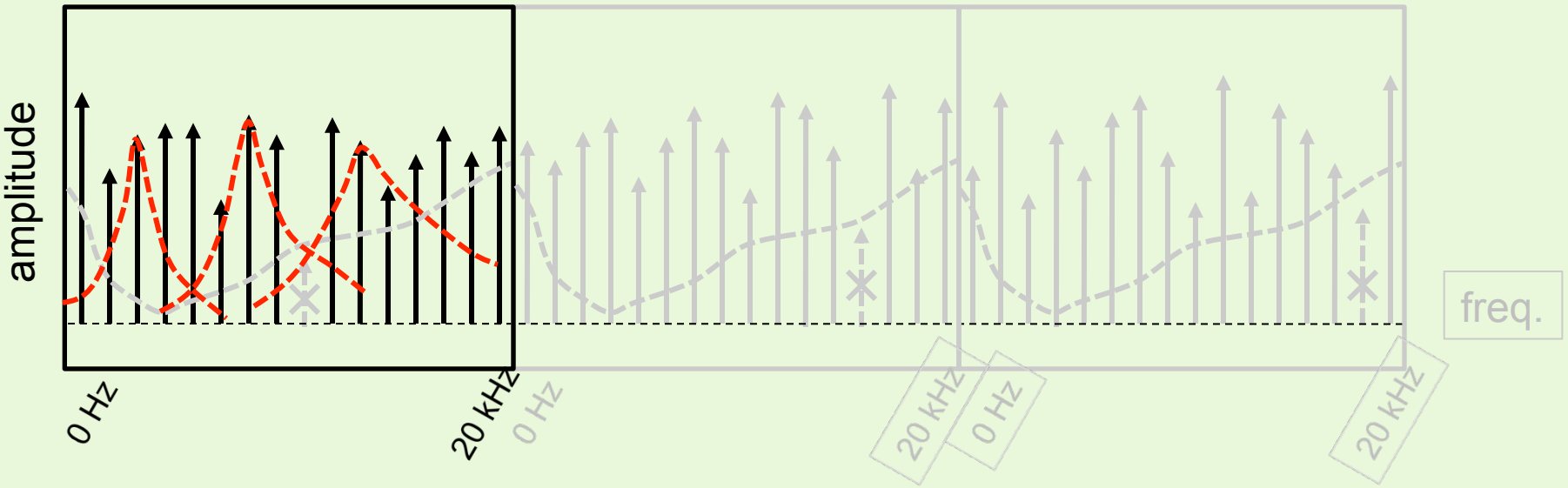
MP3 Compression

- Remove signals that cannot be heard
  Threshold of Quiet



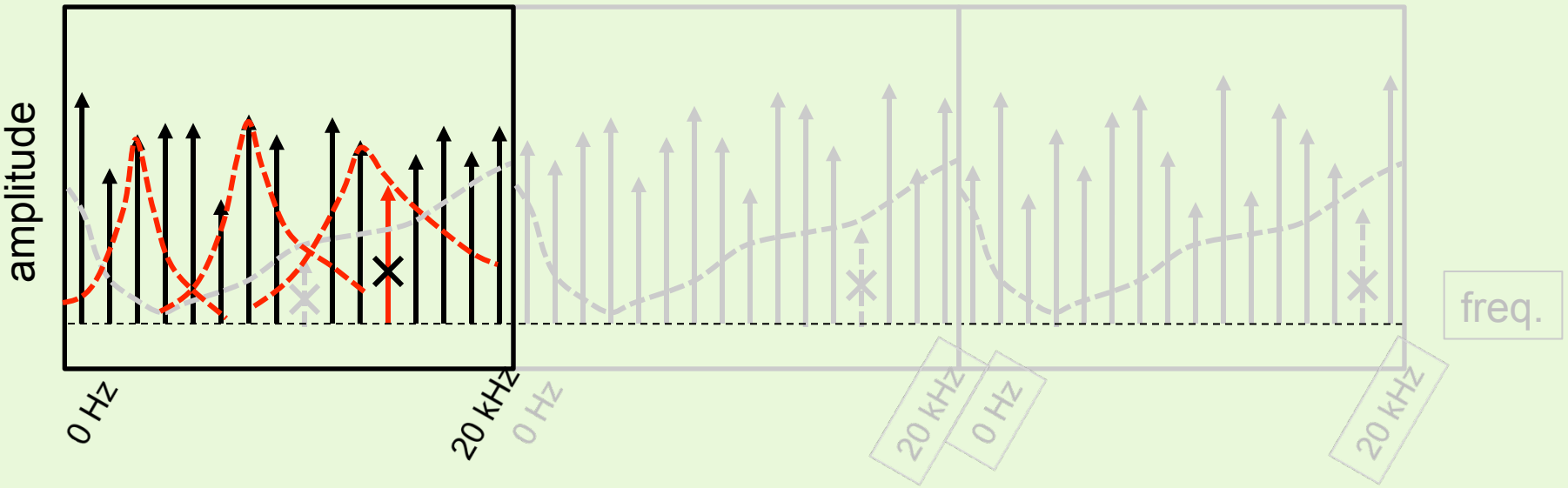
MP3 Compression

- Remove signals that cannot be heard
  Frequency Masking



MP3 Compression

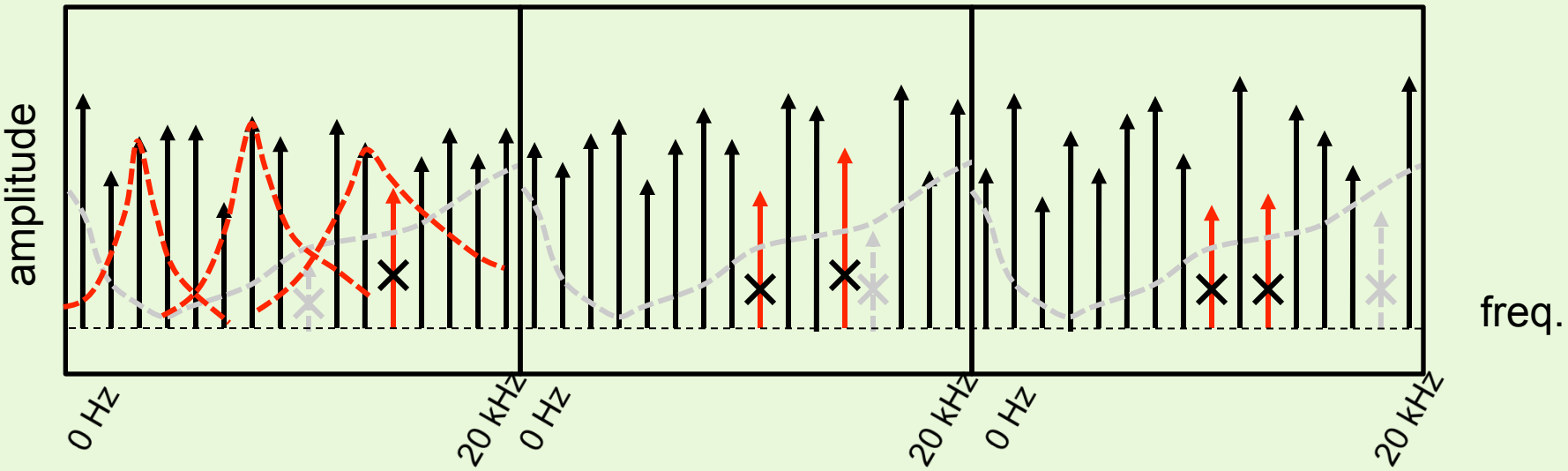
- Remove signals that cannot be heard
  Frequency Masking



MP3 Compression

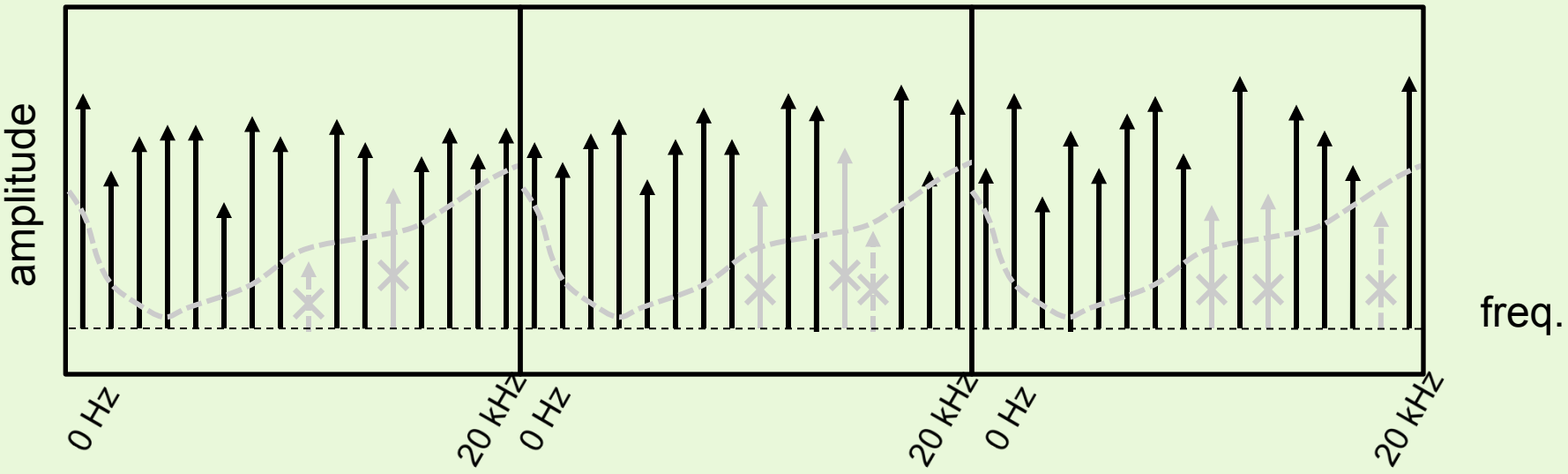
- Remove signals that cannot be heard

  Frequency Masking



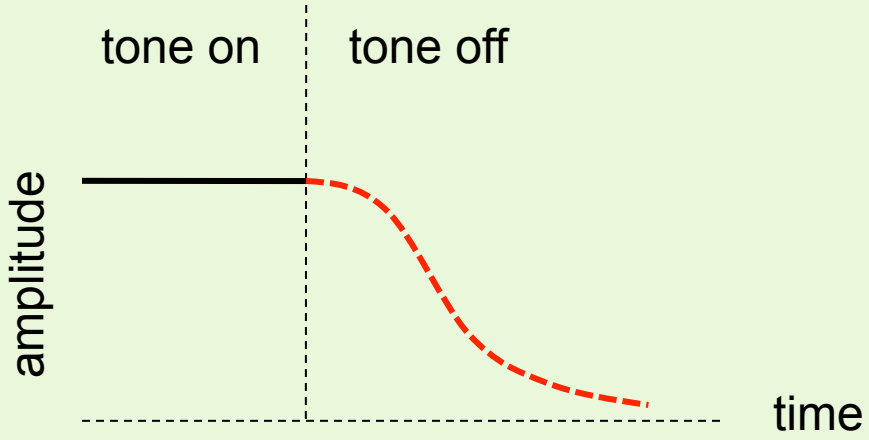
MP3 Compression

- Remove signals that cannot be heard
  Frequency Masking



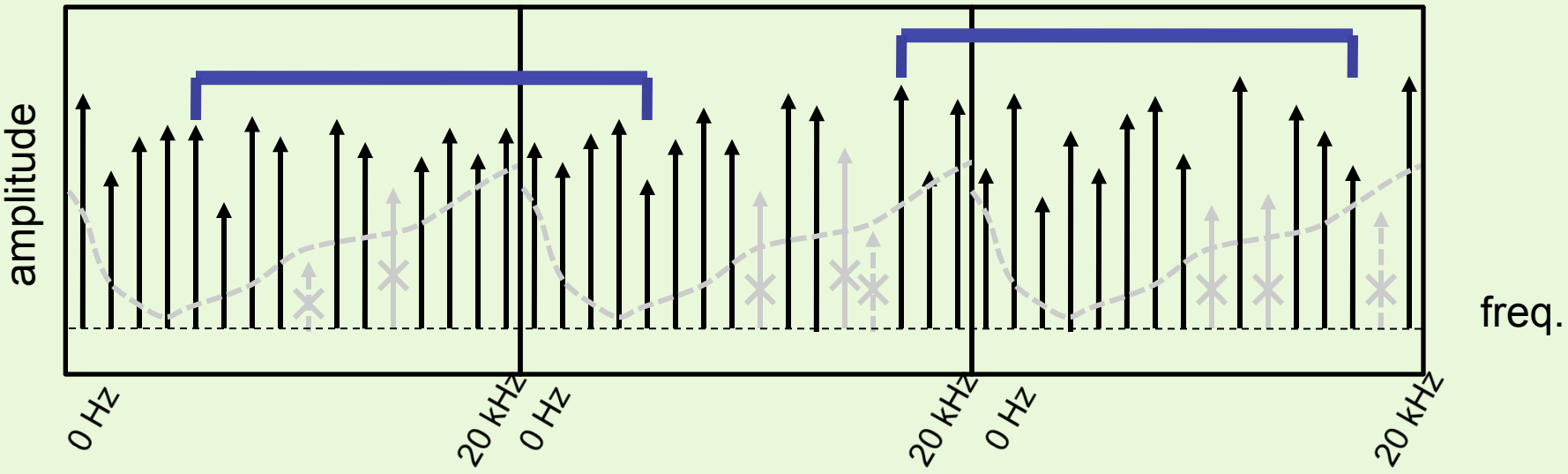
MP3 Compression

- Remove signals that cannot be heard
  Temporal Masking



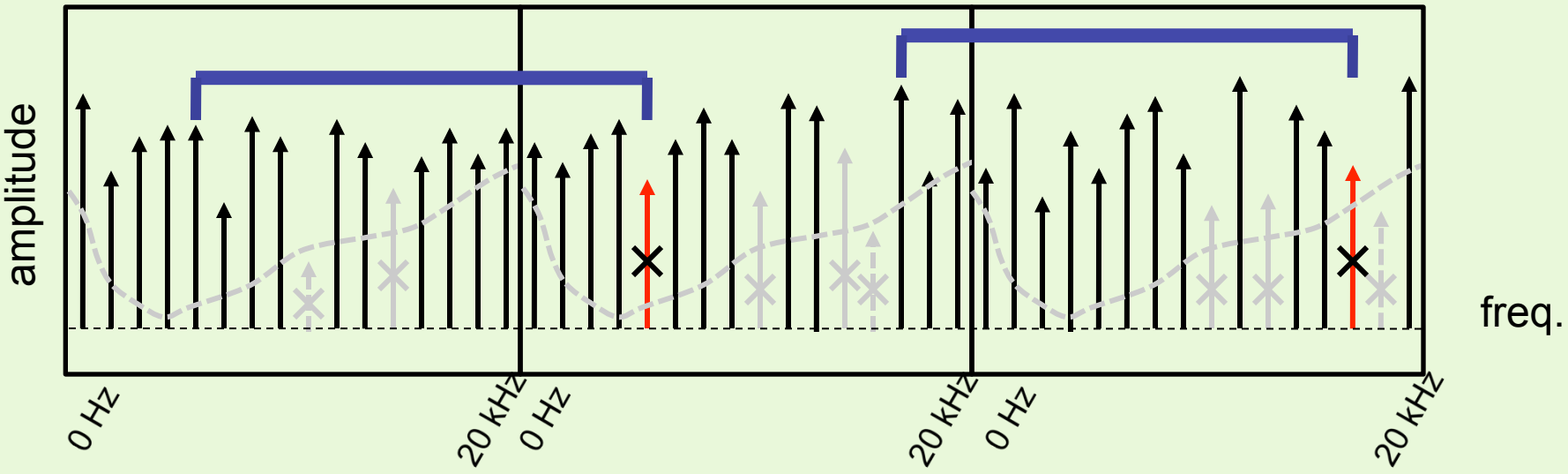
MP3 Compression

- Remove signals that cannot be heard
  Temporal Masking



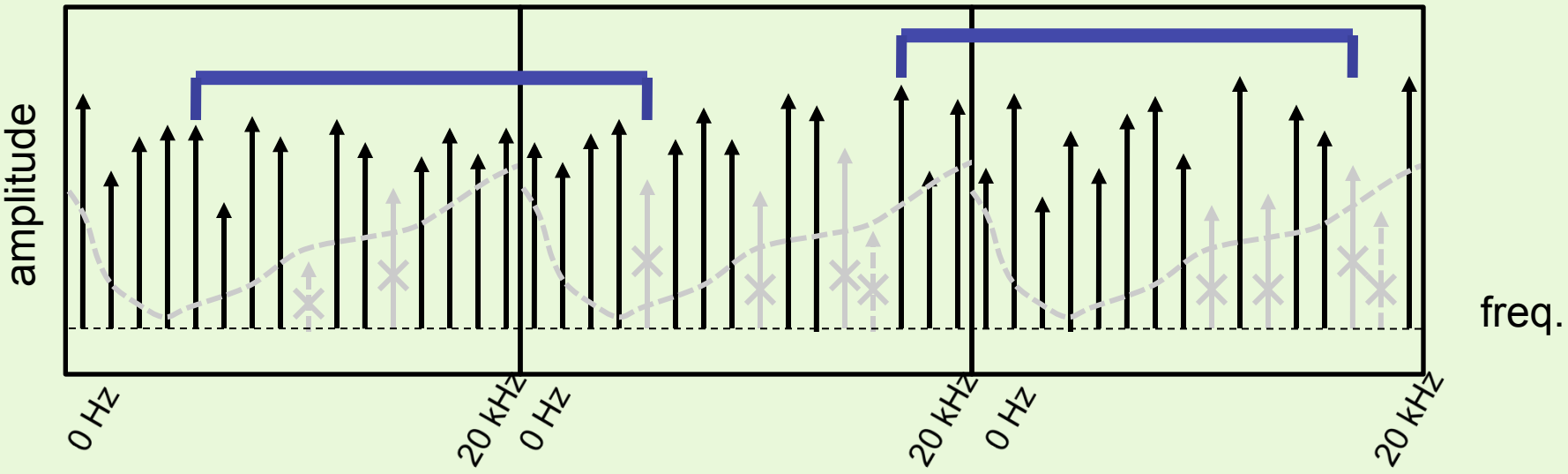
MP3 Compression

- Remove signals that cannot be heard
  Temporal Masking



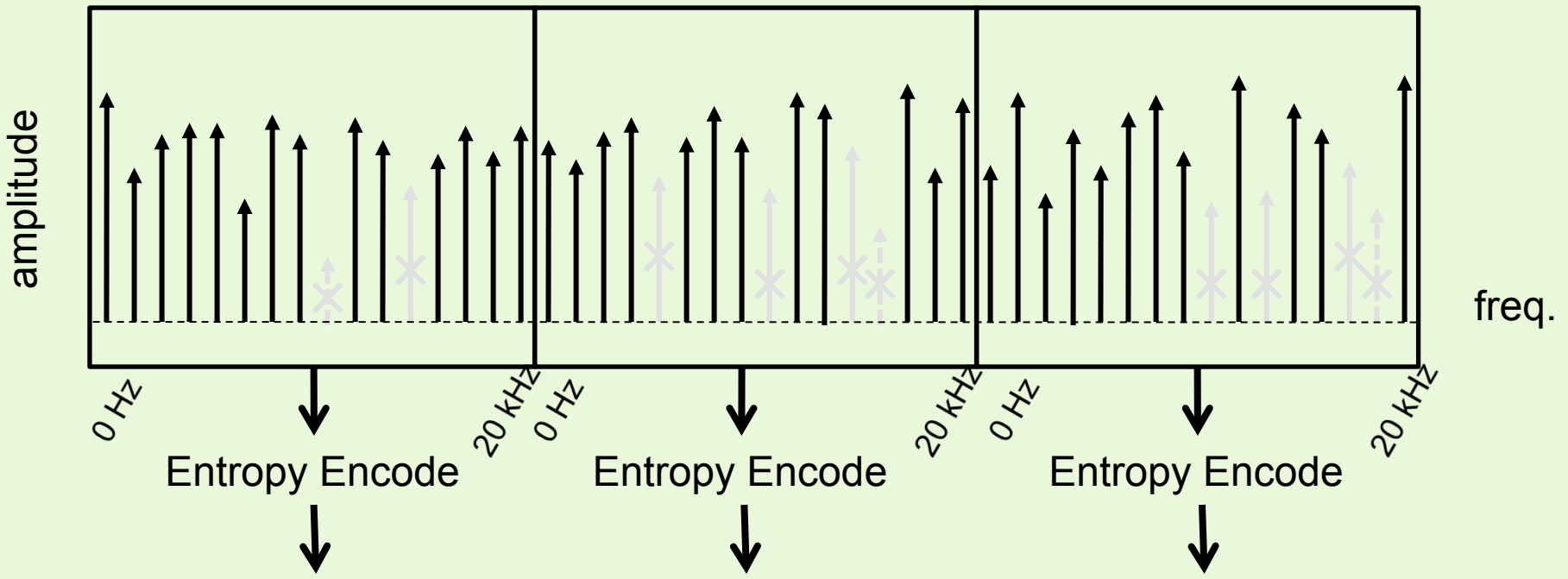
MP3 Compression

- Remove signals that cannot be heard
  Temporal Masking



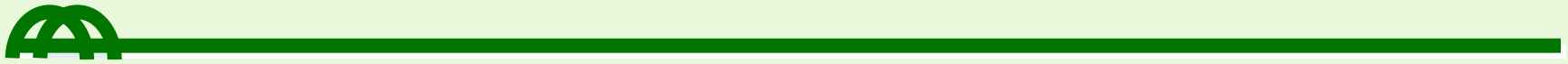
MP3 Compression

- **Entropy Encoding**
  - **Employ standard *Lossless* compression on remaining signal**



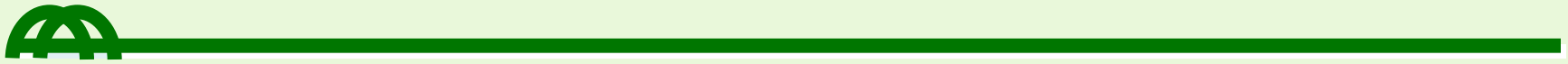
Summary

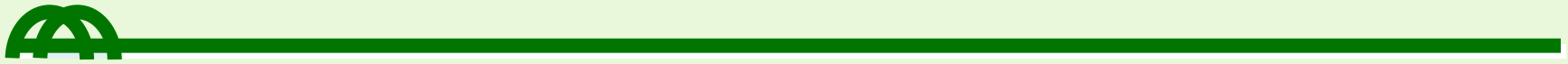
- MP3 compression yields 10:1 compression fairly easily
  - Uses frequency domain transform of data
  - Removes signals we cannot hear to minimize perceptual loss
  - Employs entropy encoding on remaining frequency data
- Interested?
  - Fall Quarters – Introduction to Multimedia Networking



My Words of Wisdom

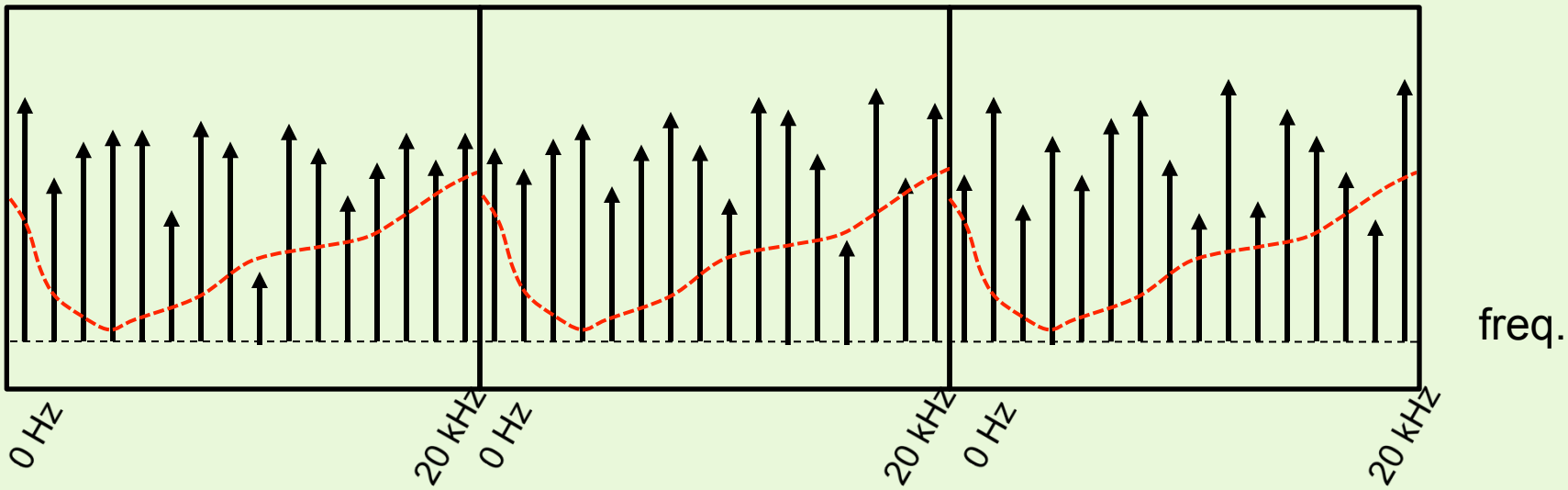
- 3 deep stack maximum
  - Organize so that the audience only needs to keep track of two to three things at a time
  - Need to summarize and finish each section to allow audience to "pop" it off of their stack
- Just because you know it doesn't mean it has to go in the presentation
  - Pitch presentation at right level of audience and for the right length of time
- Never go over time
  - Be mindful of where you are in your talk at all times
  - Know how to adjust / drop information from you talk
  - Just because you practiced it in 40 minutes does not mean it will be delivered in 40 minutes.





MP3 Compression

- Remove signals that cannot be heard
  Threshold of Quiet



MP3 Compression

- Remove signals that cannot be heard
  Threshold of Quiet