GDC 09

learn
network
inspire

www.GDConf.com

Game Developers Conference®

March 23-27, 2009 | Moscone Center, San Francisco

Digital Audio 101

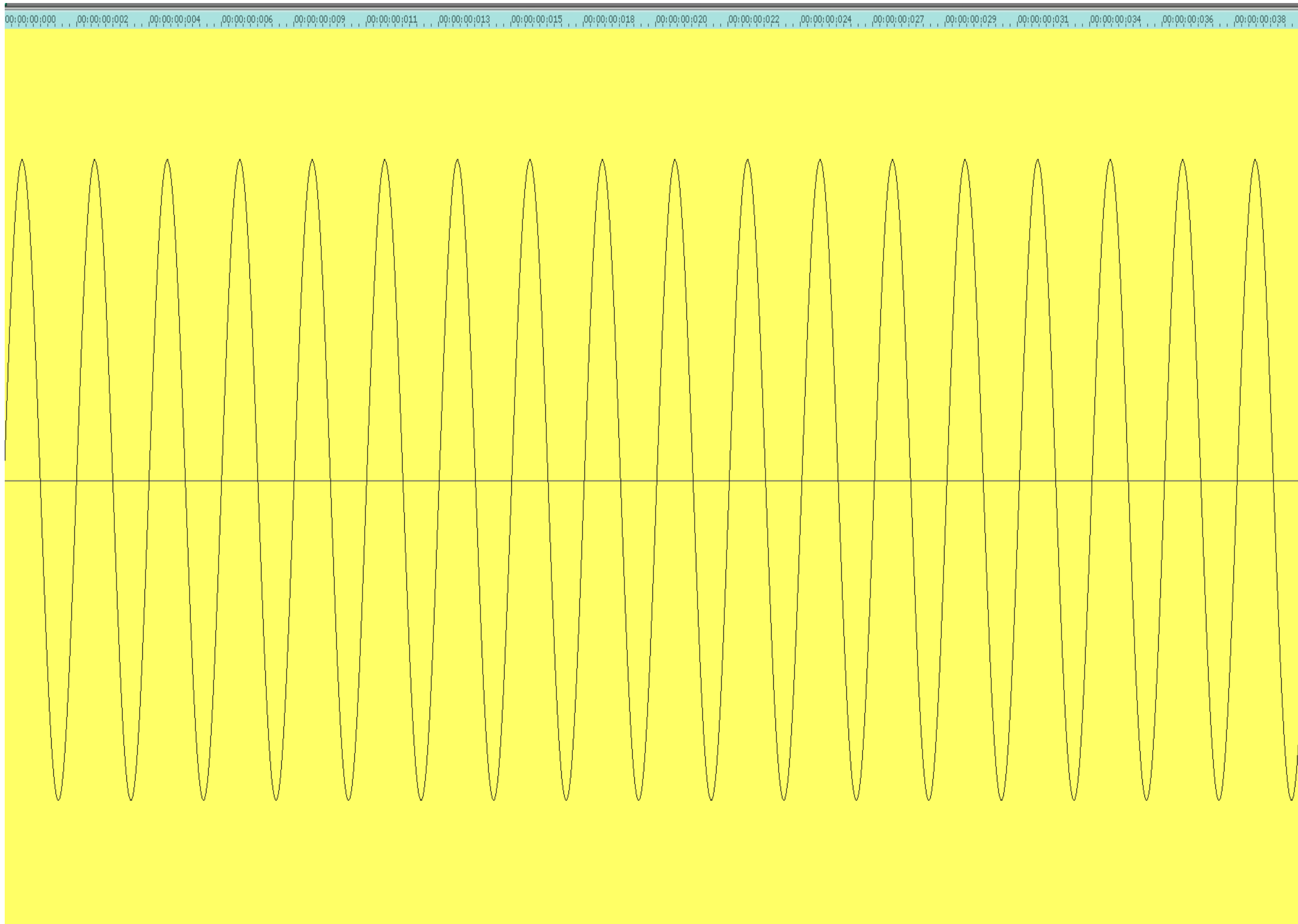
Theory and Practice

Alistair Hirst
OMNI Audio

Sine Wave

» Most basic waveform
» 1 frequency

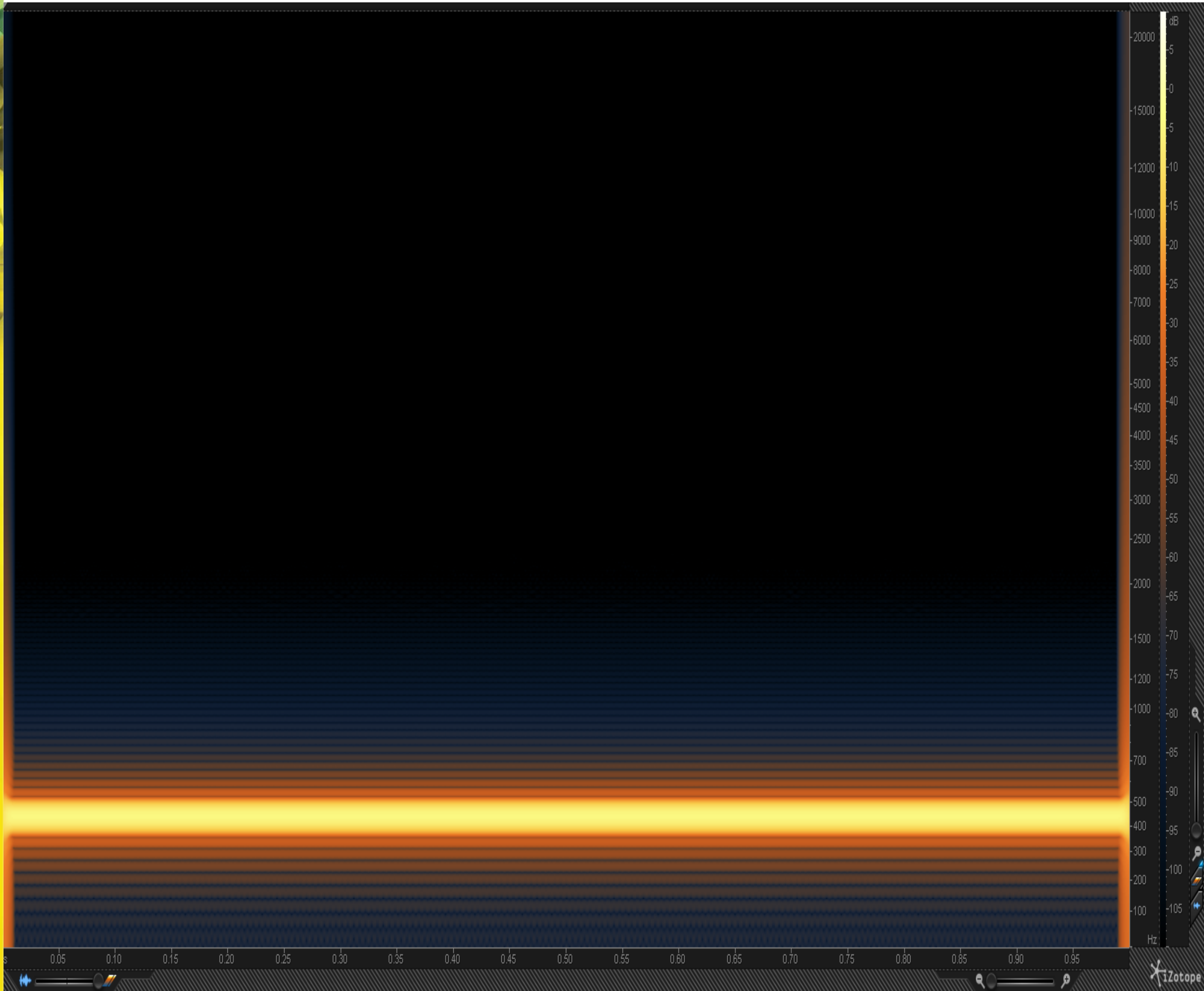
Sine Waveform



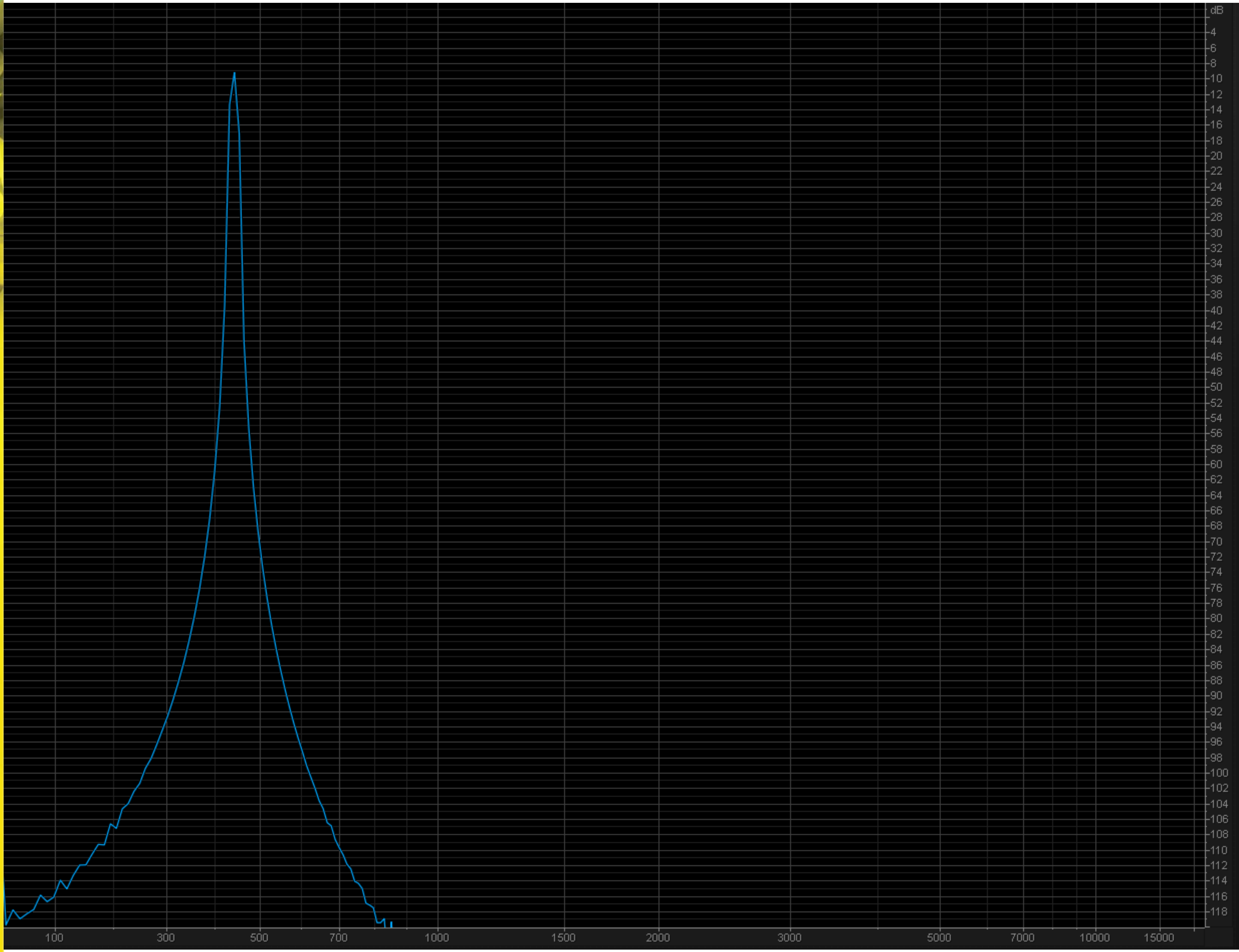
learn
network
inspire

www.GDConf.com

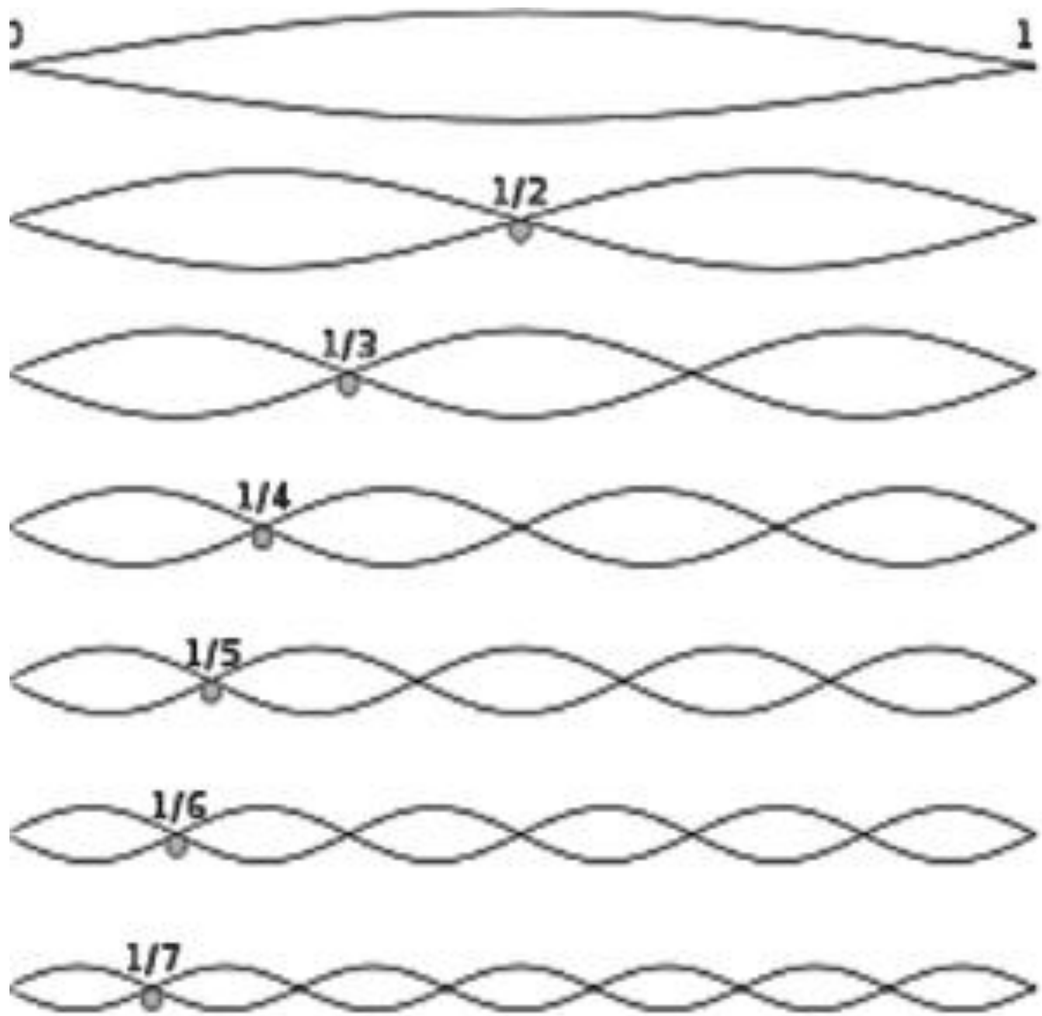
Sine Spectragram



Sine Spectragraph

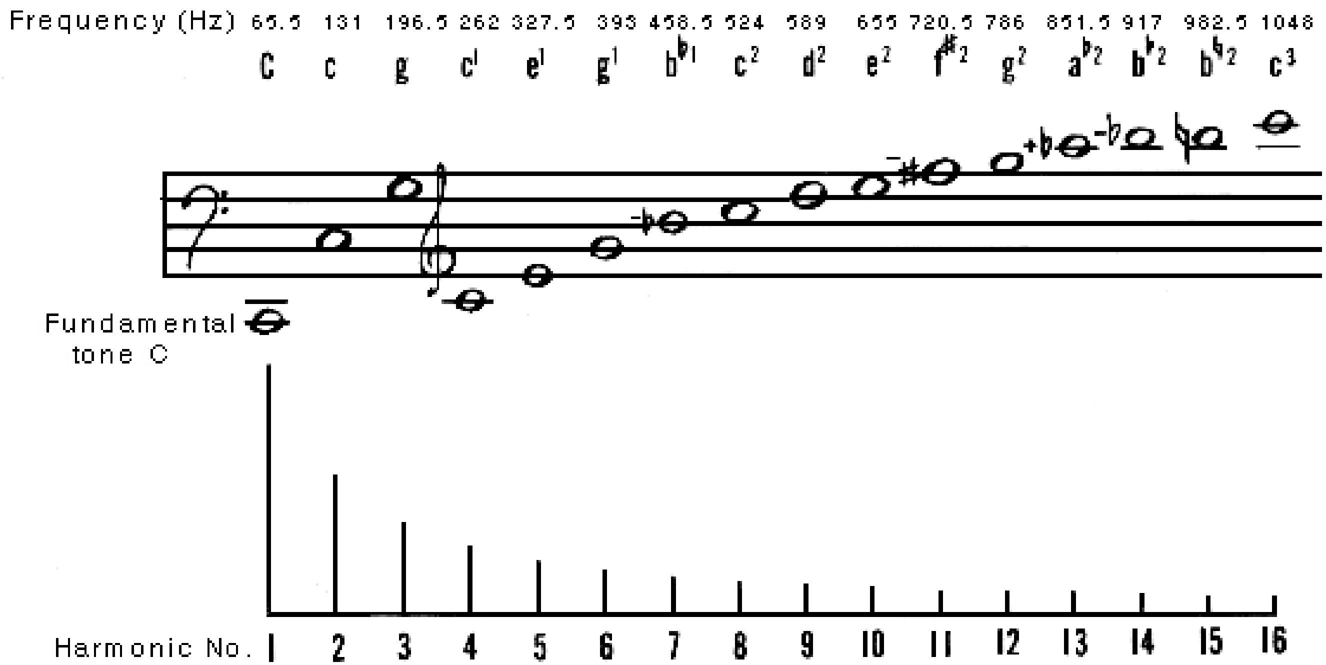


Harmonic Series



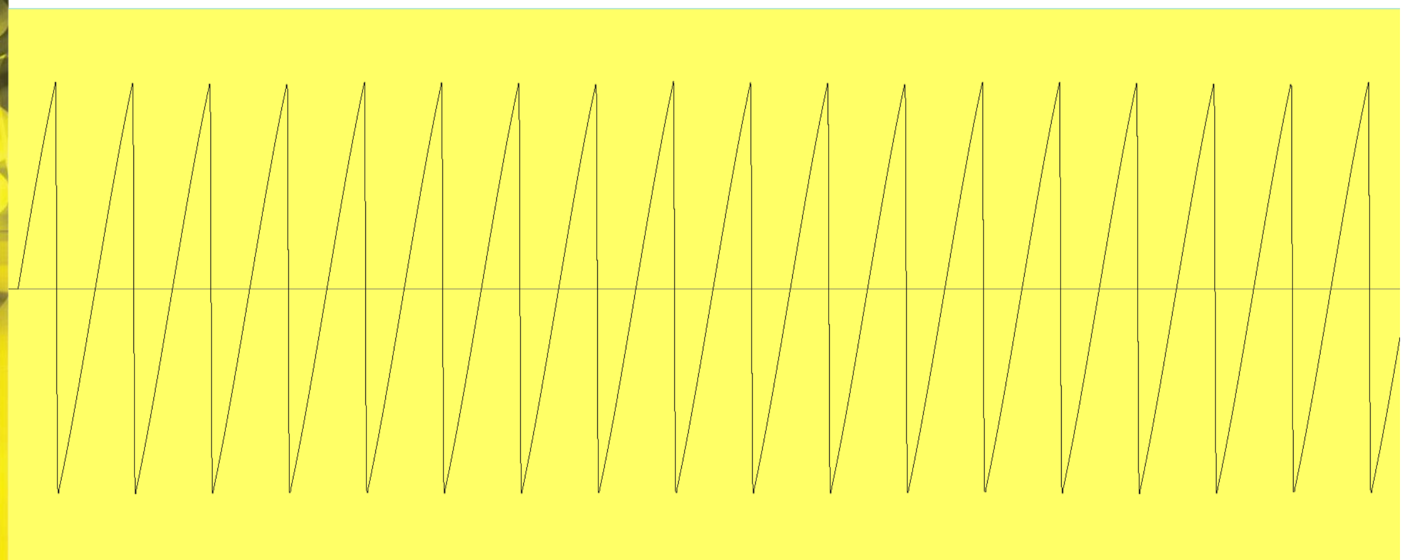
Guitar string

Harmonic series

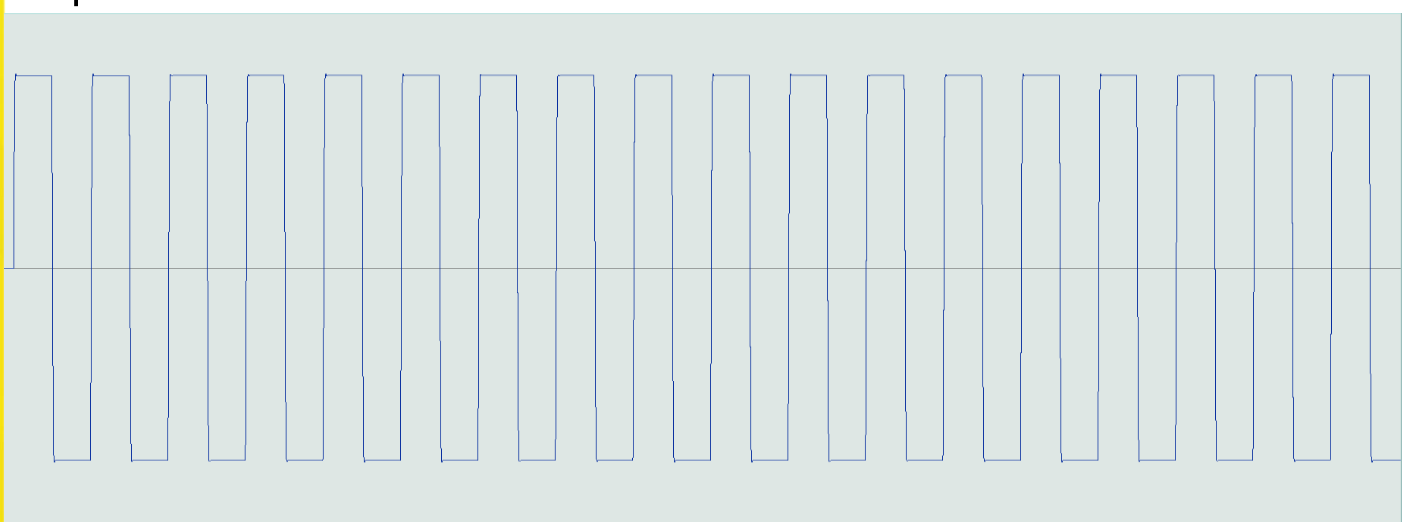


Timbre:

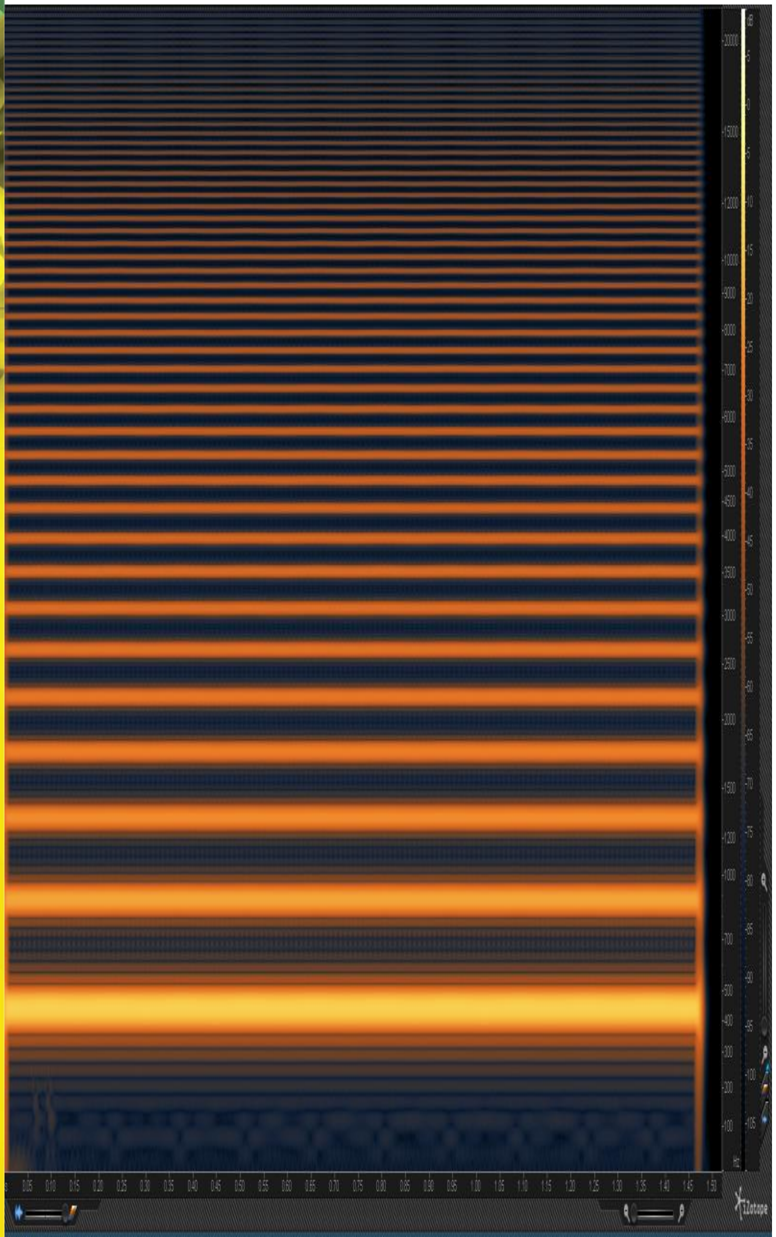
Sawtooth wave

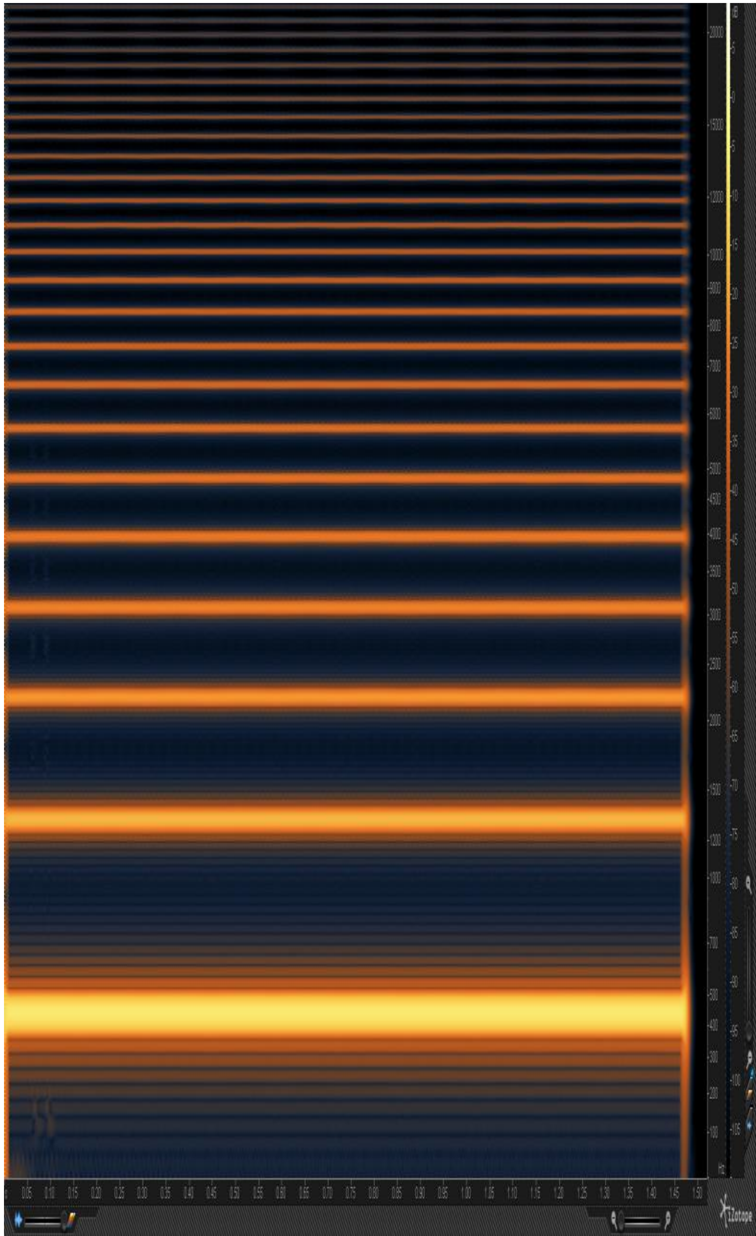


Square wave

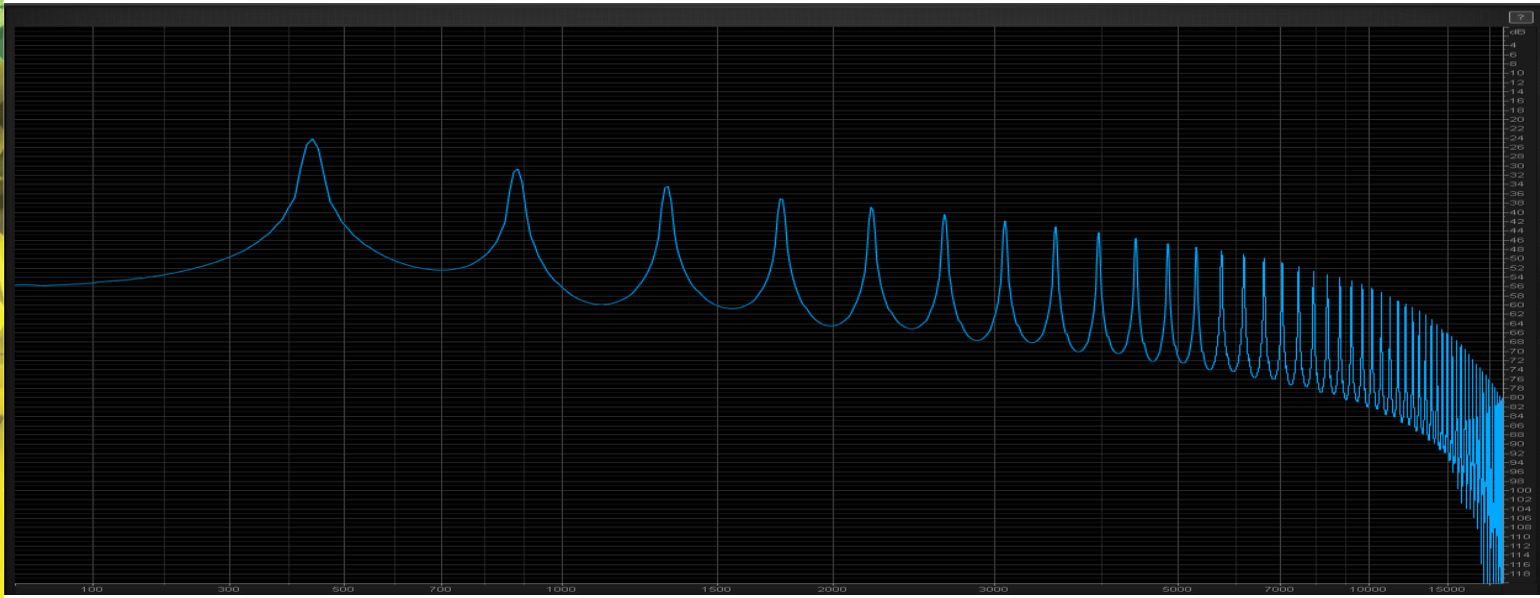


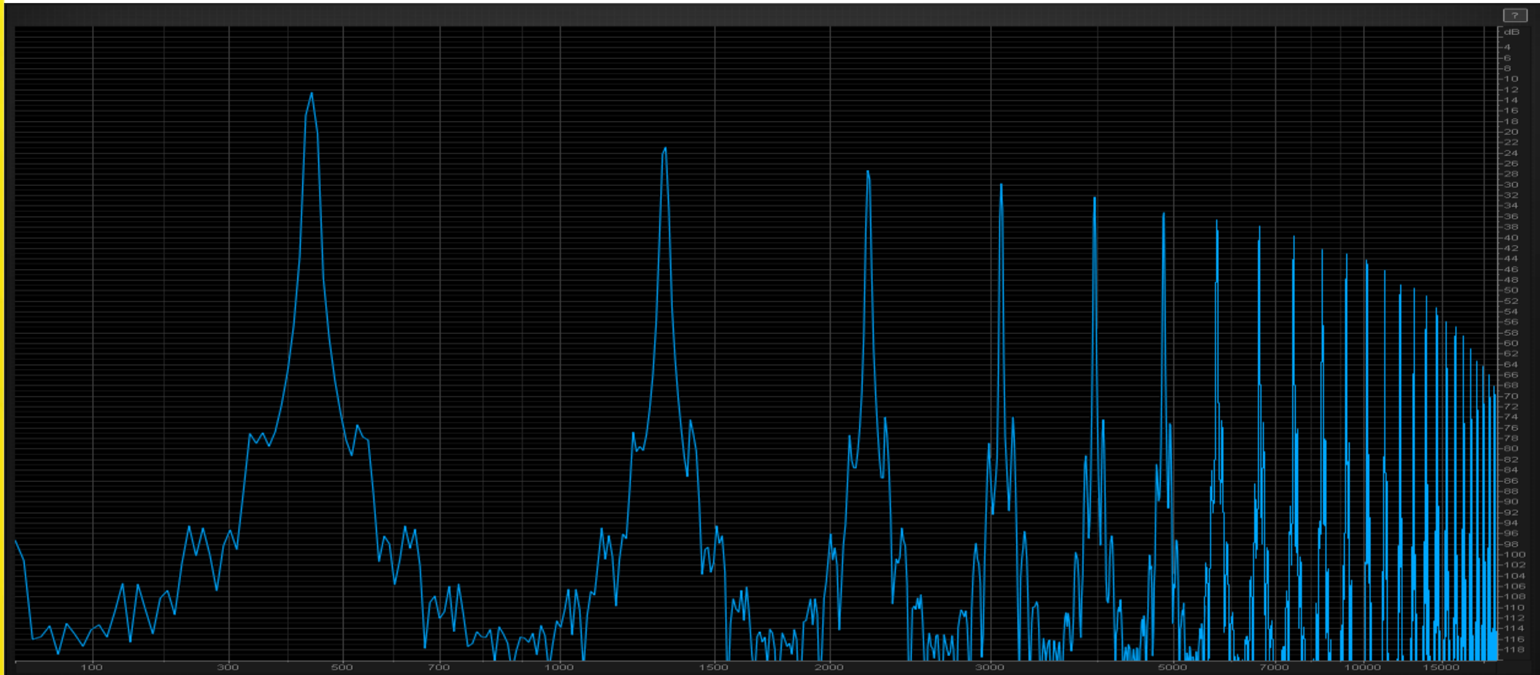
Sawtooth & square spectrogram



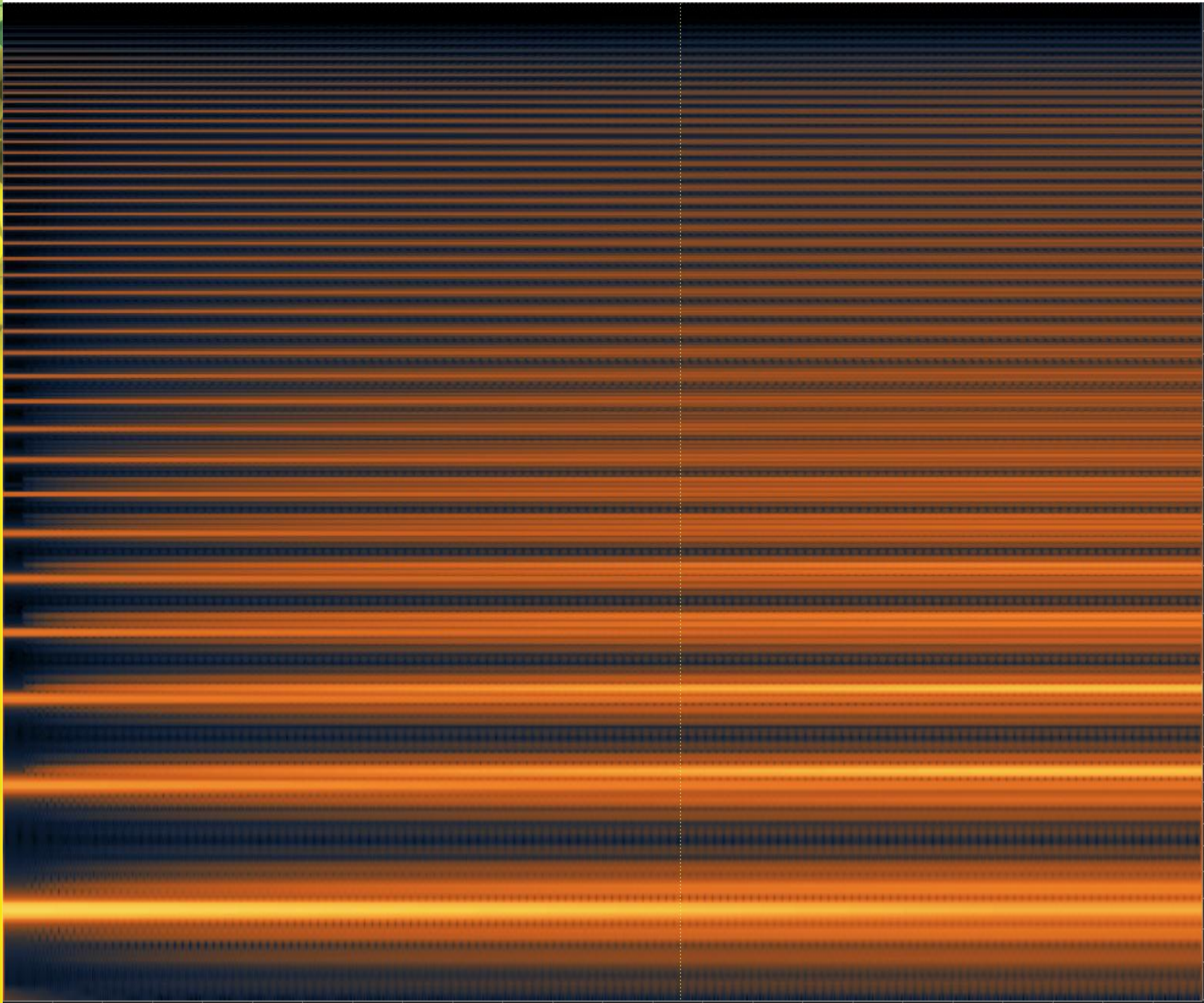


Sawtooth & Square spectrograph

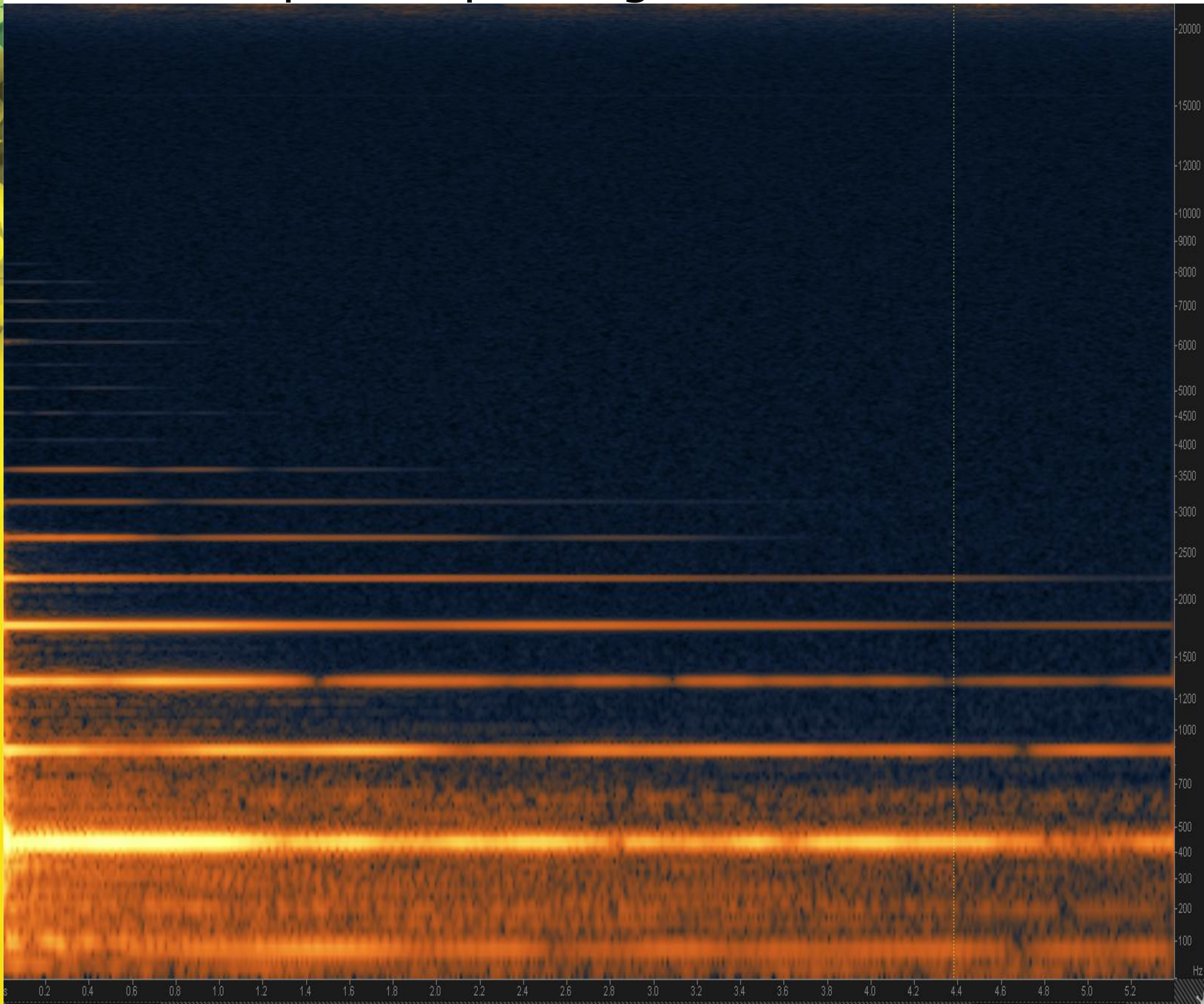




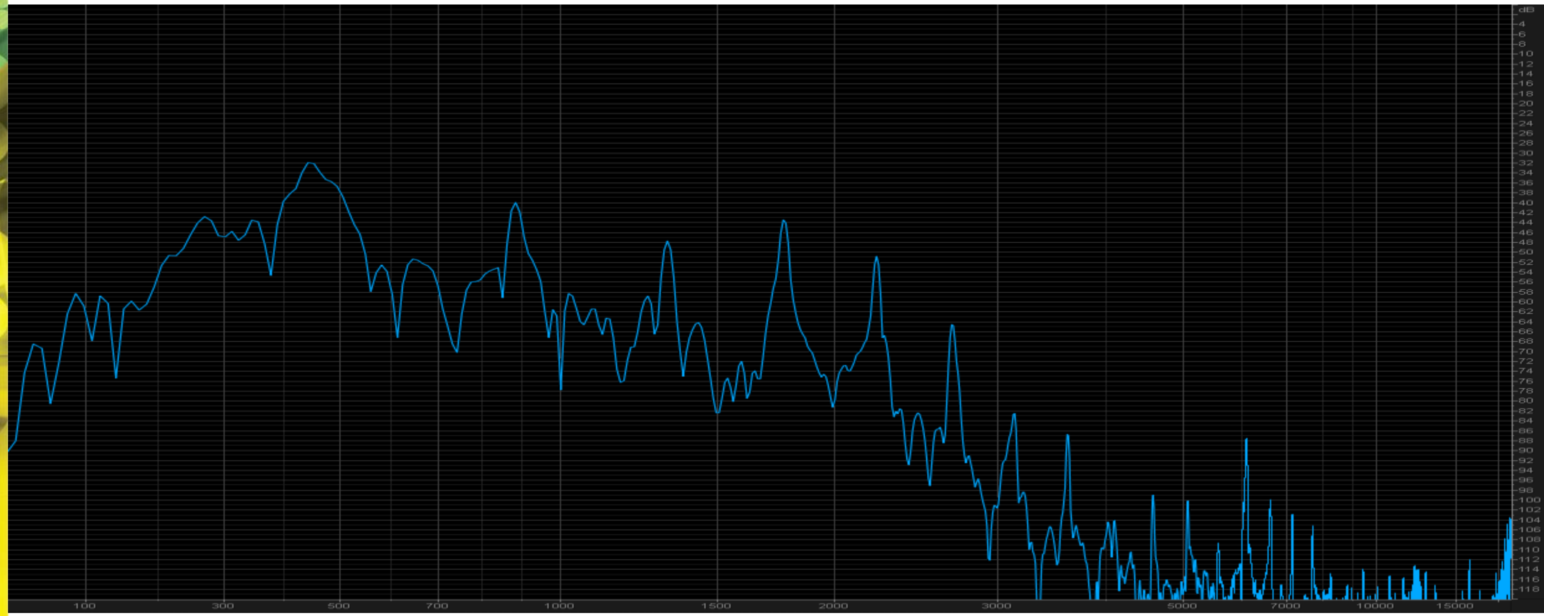
Enharmonics

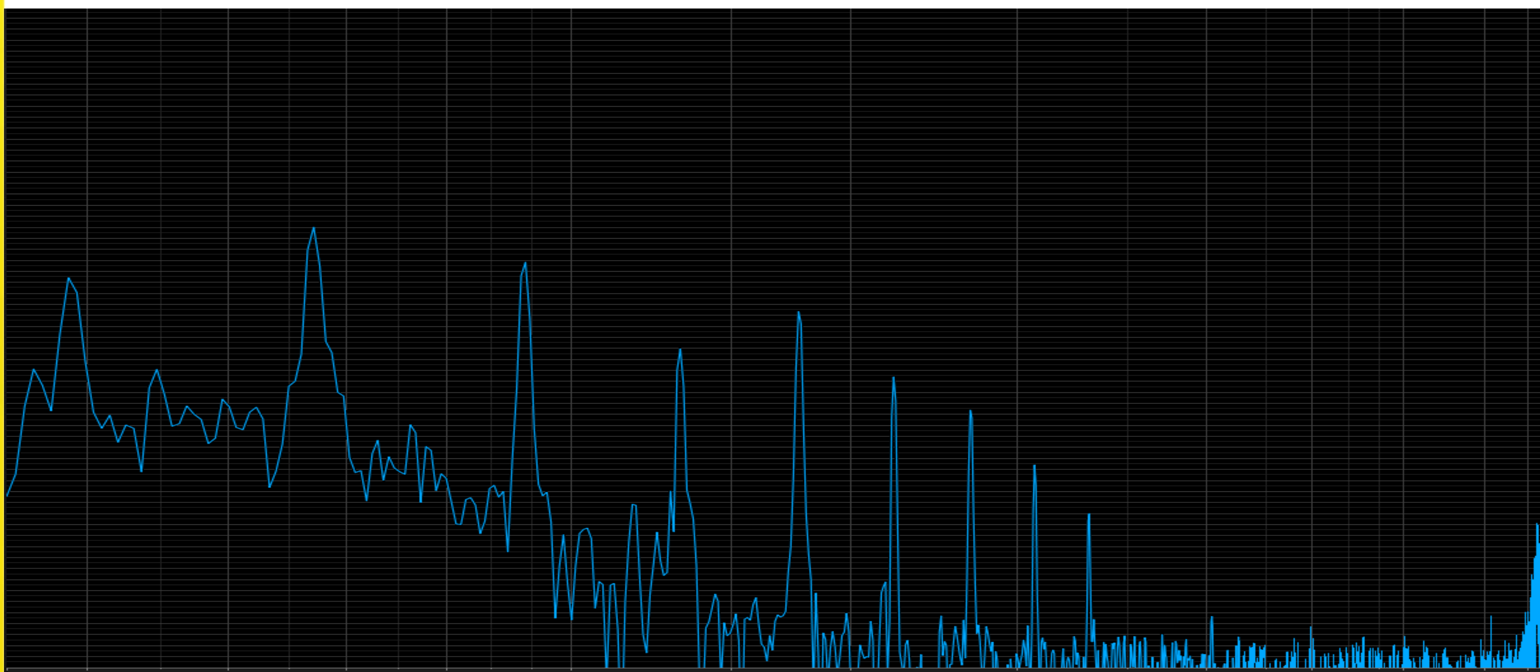


Acoustic piano spectrogram

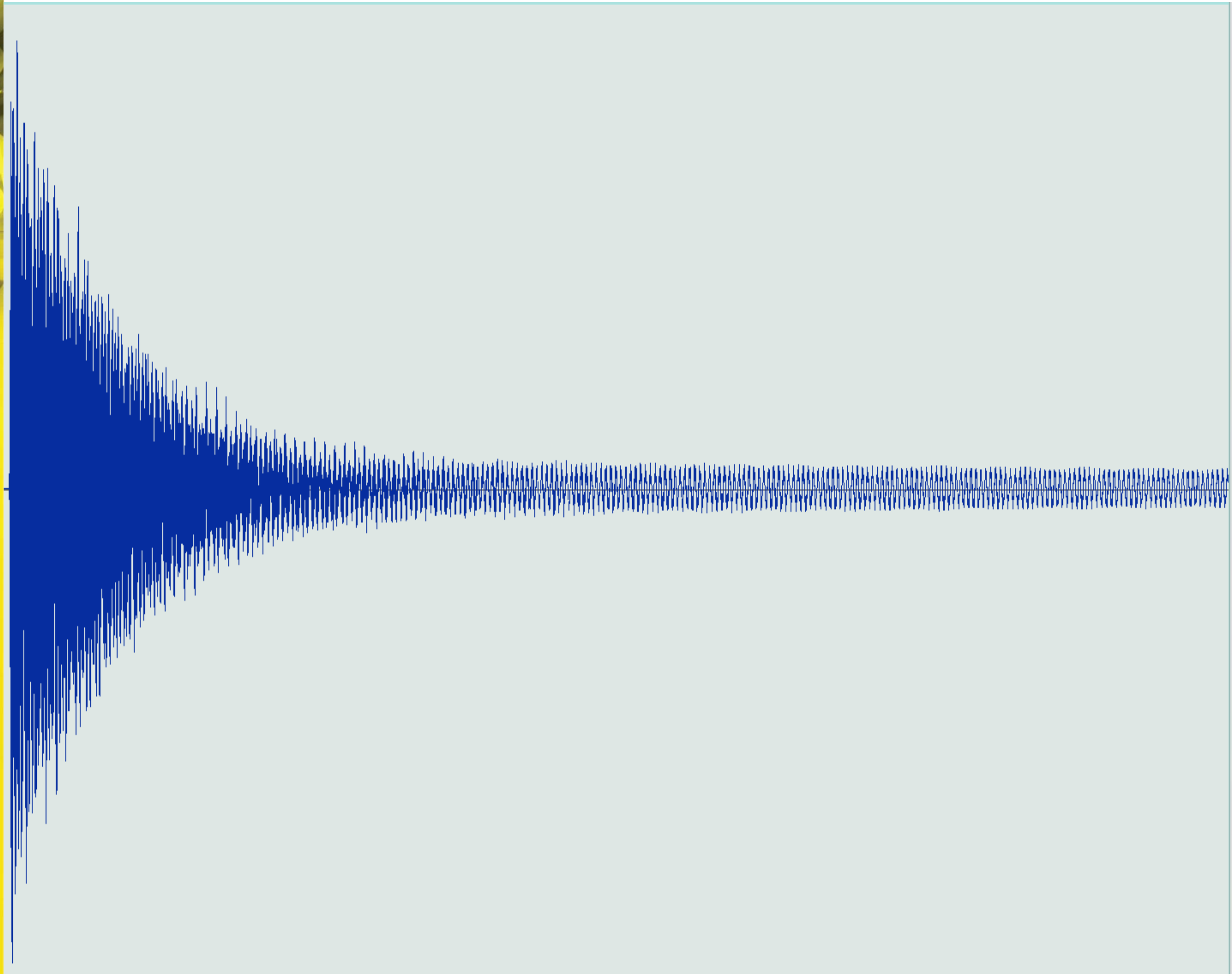


Piano attack, sustain

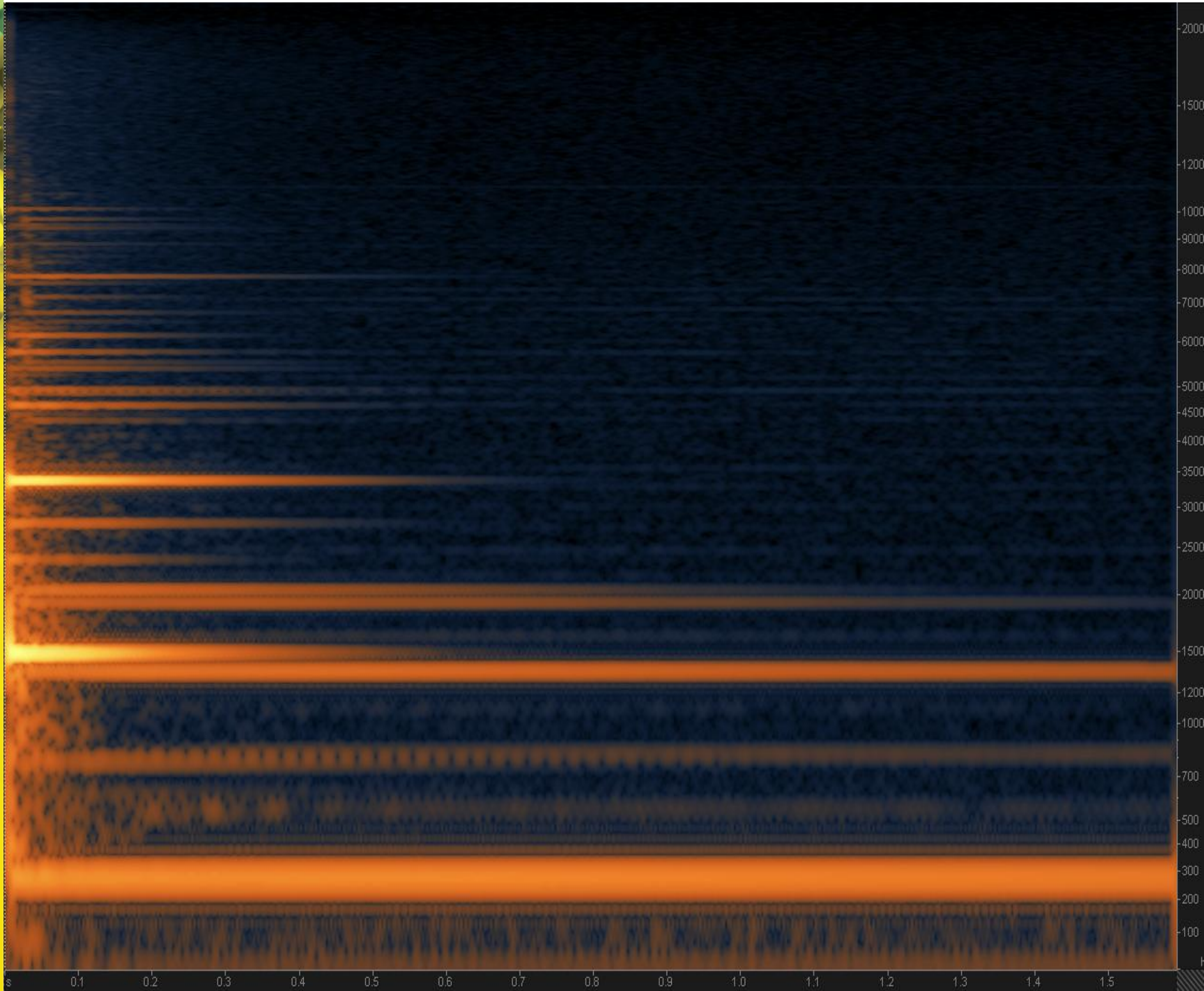




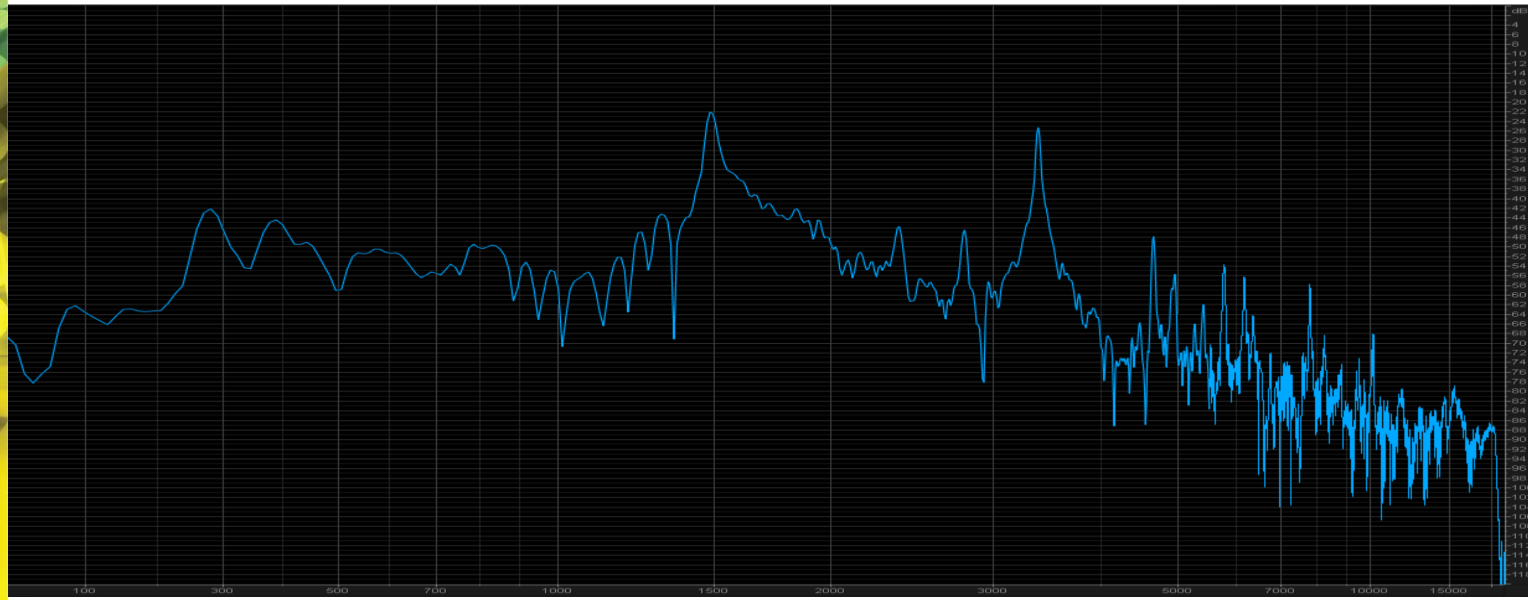
Gamelan (Demung)

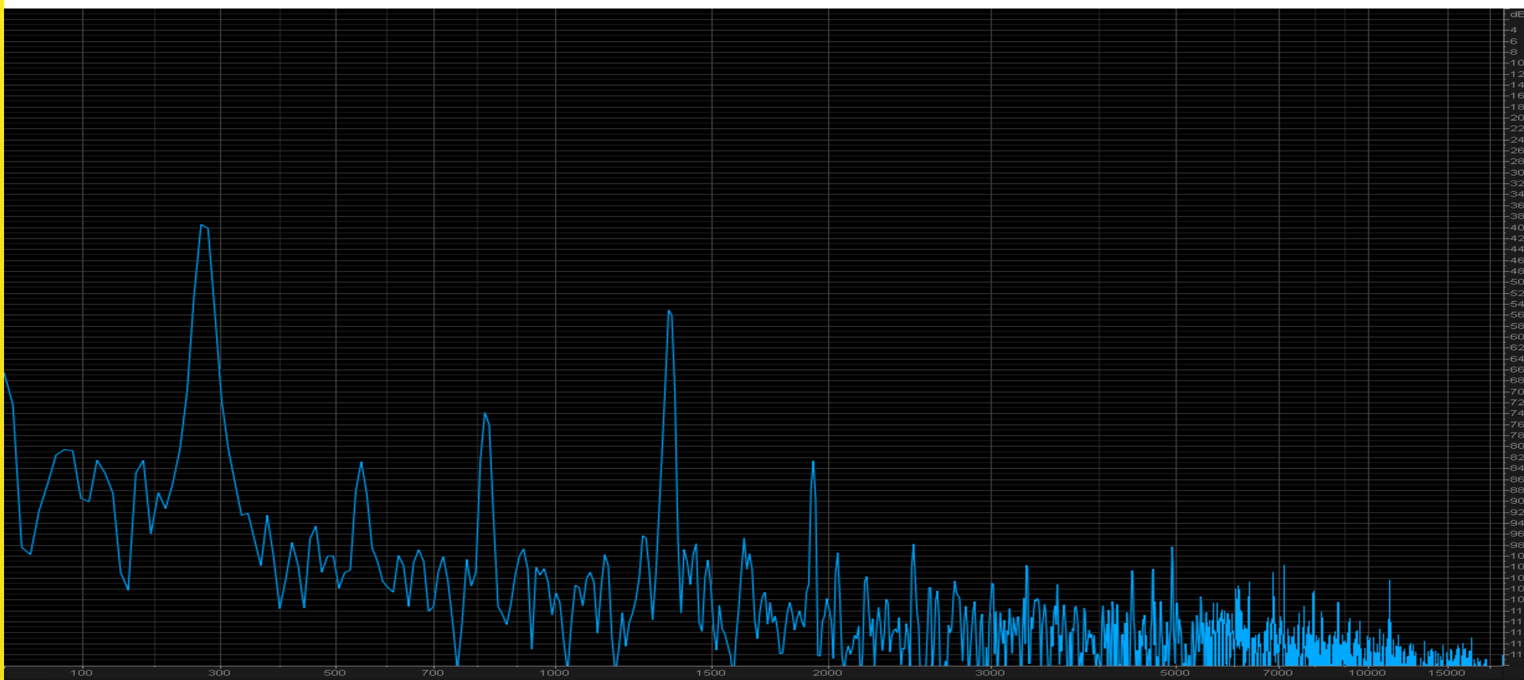


Gamelan spectrogram

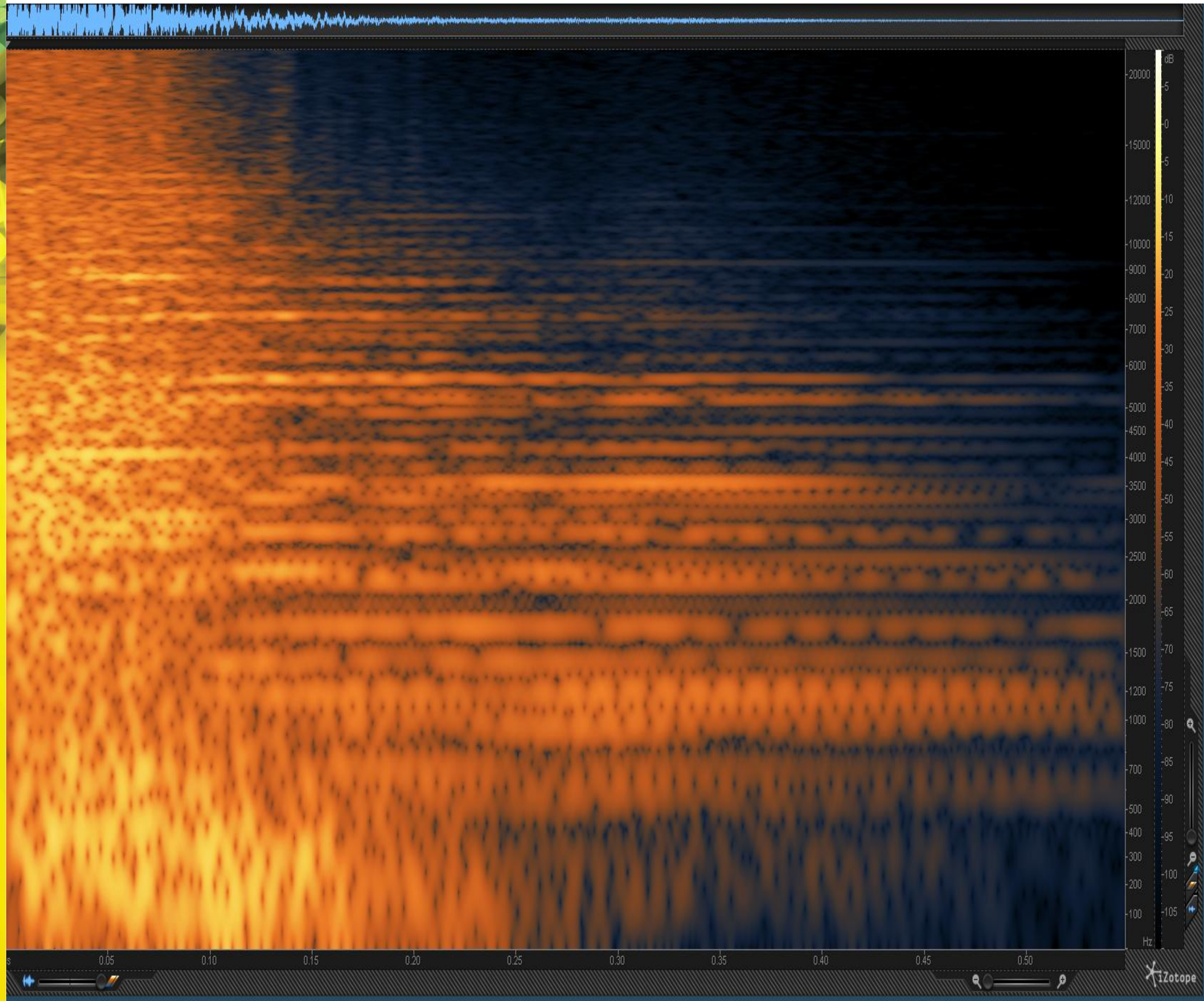


Gamelan spectrograph - attack, sustain

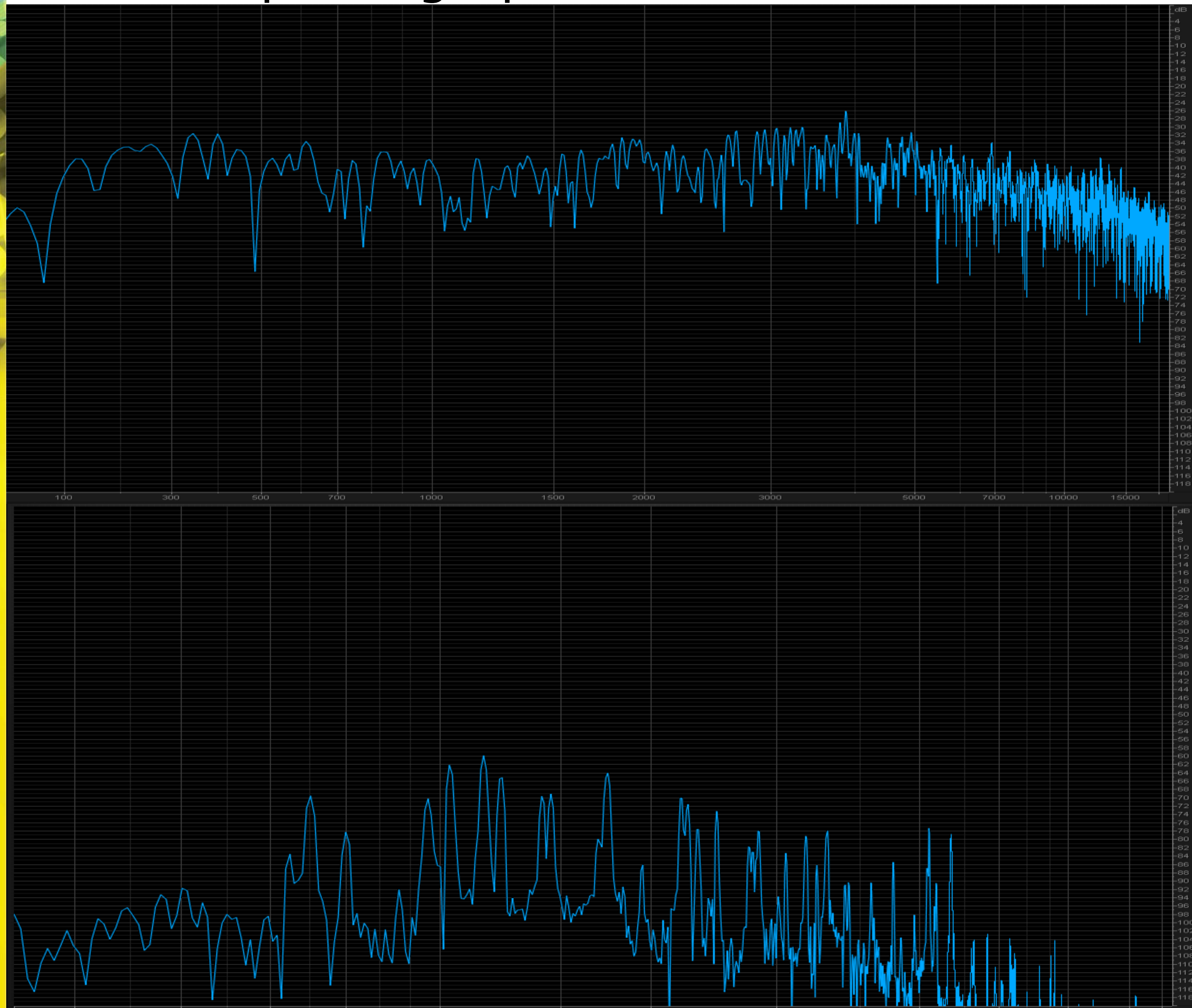




Sword spectrogram

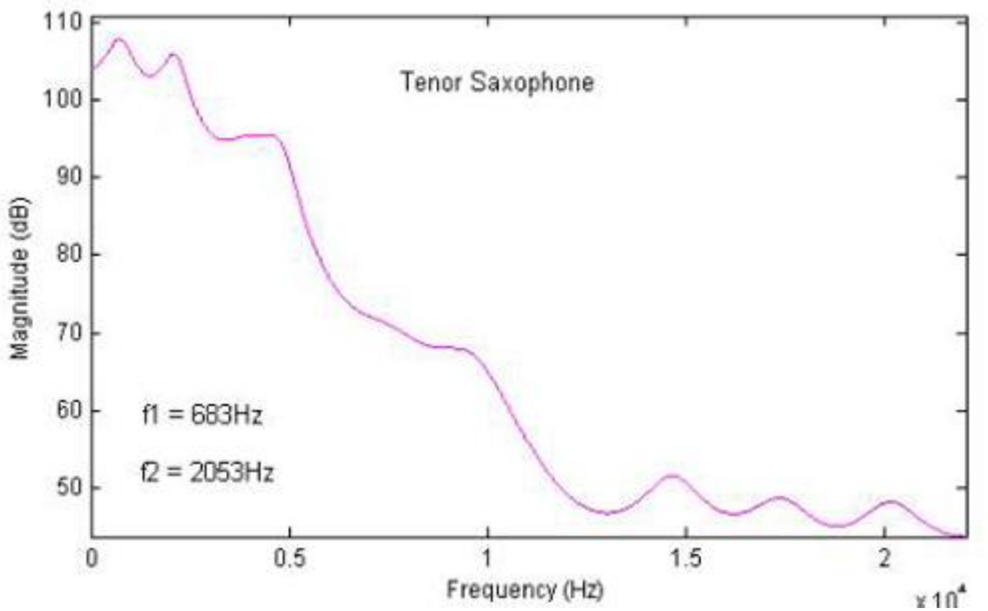


Sword spectrograph attack



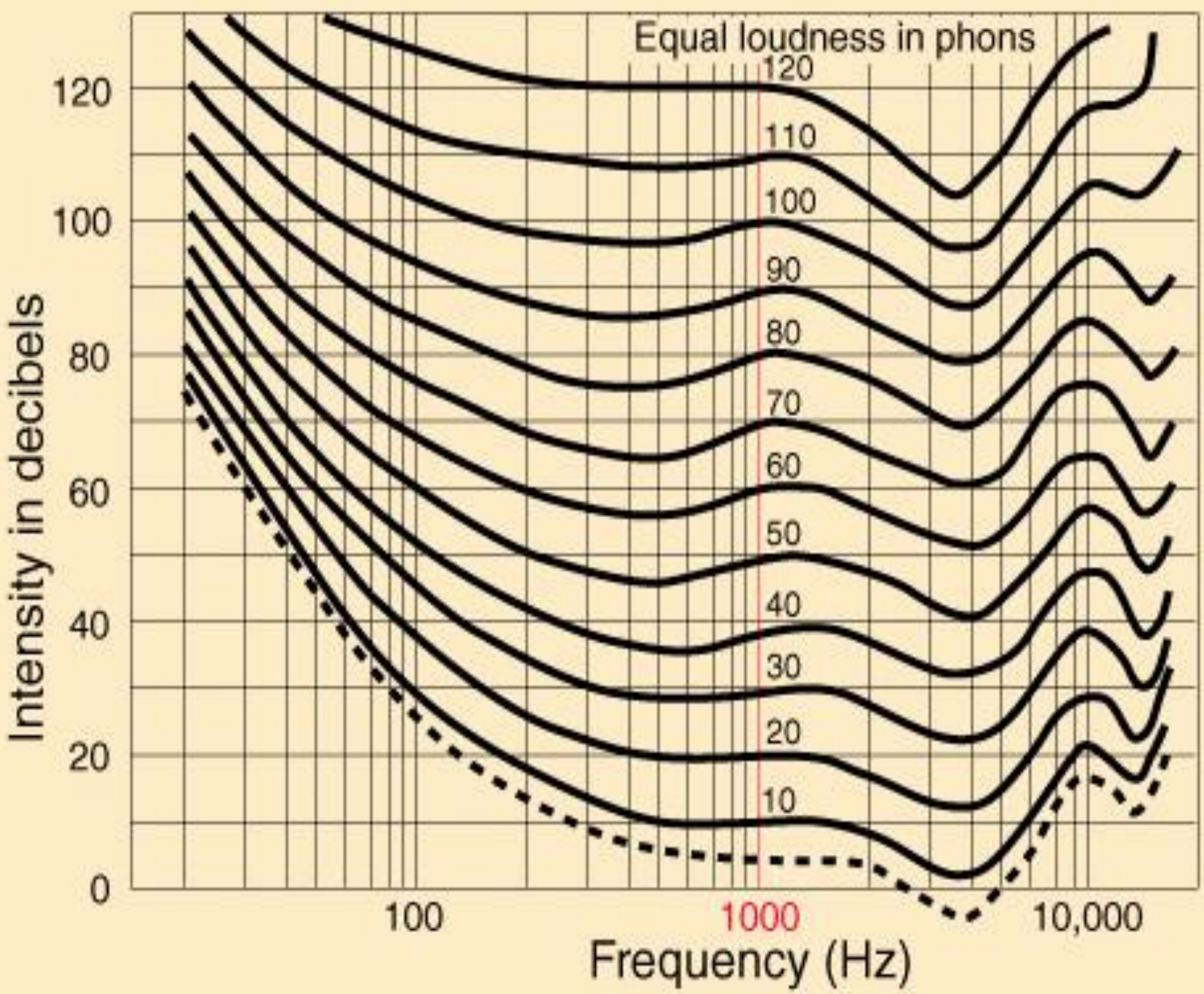
Formants

- The frequency spectrum of a sound caused by acoustic resonance.
- Examples:
    - Vocal tract
    - Violin
- Independent of pitch.



http://ccrma.stanford.edu/~jmccarty/formant.htm

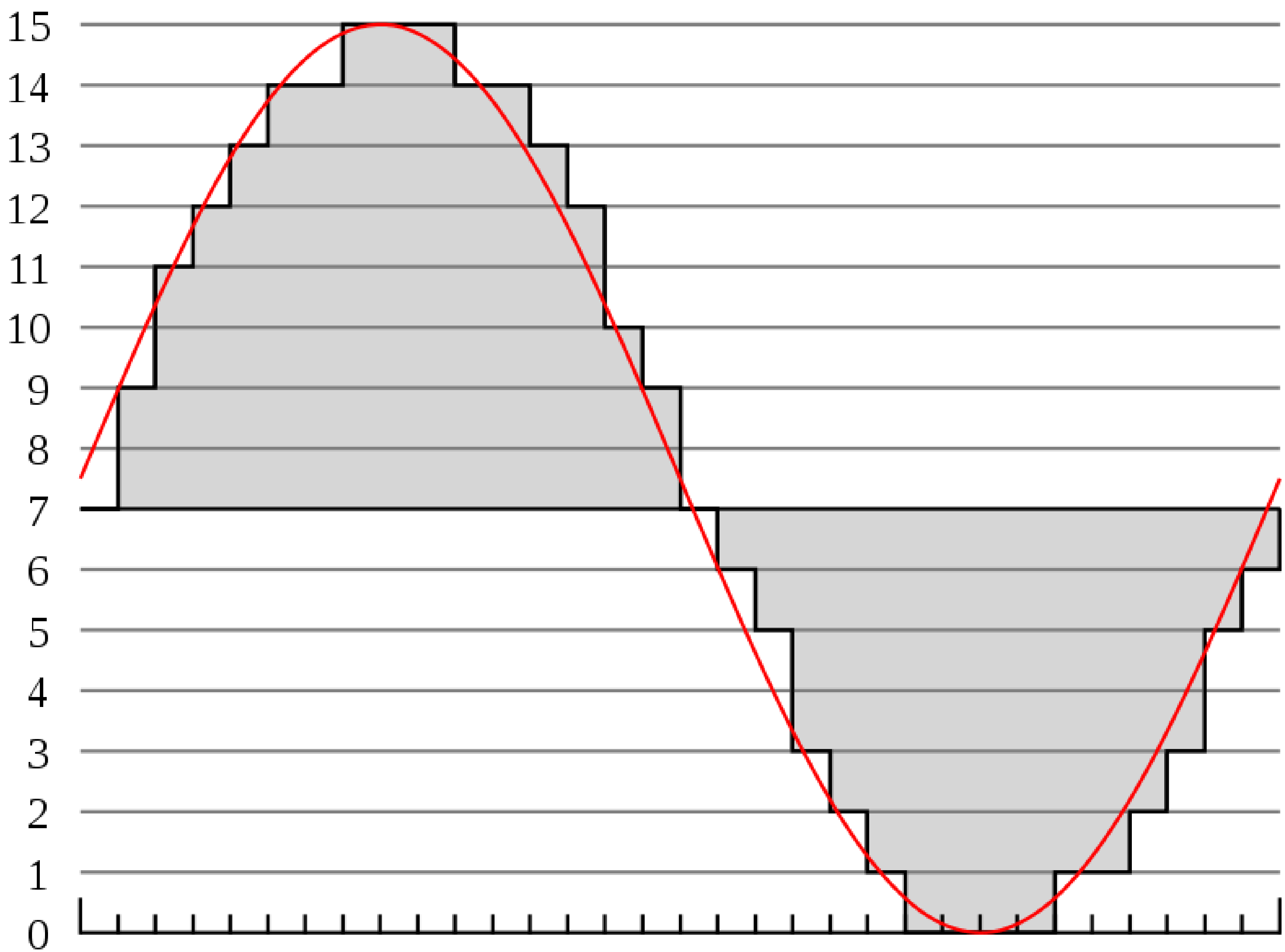
Equal loudness contours



The Golden Rule of Digital Audio

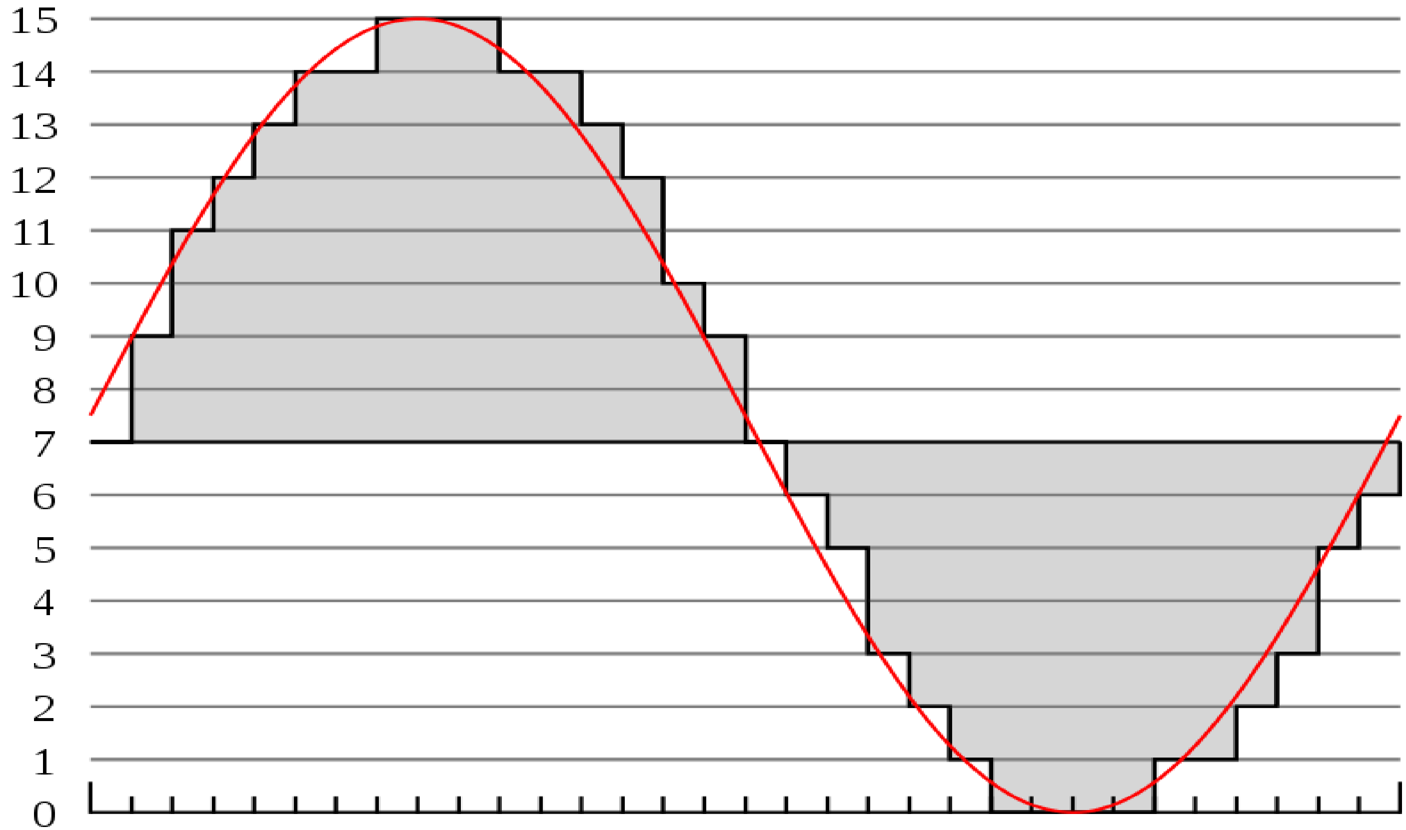
The higher quality you keep your sound at, the less it will degrade in the final stage when it goes to delivery.

Pulse Code Modulation



http://en.wikipedia.org/wiki/File:Pcm.svg

Resolution



The higher the resolution, the more accurate the representation of the analog waveform.

16 bits vs 24 bits

- Each bit of additional resolution lowers quantization noise by 6dB.

- 24 bits = Max 144dB (theoretical)

- 16 bits = Max 96dB (theoretical)

Analog to digital conversion

- Analog consoles have a dynamic range of about 115dB
- Aim for -12dB to -6dB on meters when recording.
  - Digital meters can miss brief overages

- Digital to Analog converters limited to about 120dB

So why use 24 bits?

Answer:

DSP

Digital Signal Processing

The Math (simplified)

Example: Reduce the volume by 6dB (1/2 volume because of logarithms)

Sample	X half	=
0.9	X 0.5	= 0.45

Increased number decimal places.

More Math:

A 1 dB gain boost involves multiplying by 1.122018454 (to 9 place accuracy)

Source: www.digido.com

Truncating (dropping) decimal points is cutting off detail:

- Ambience
- Warmth
- Stereo separation

Multiply that by many calculations in a signal processing chain and mixing, and it adds up

More bits = more information = more detail

Take away:

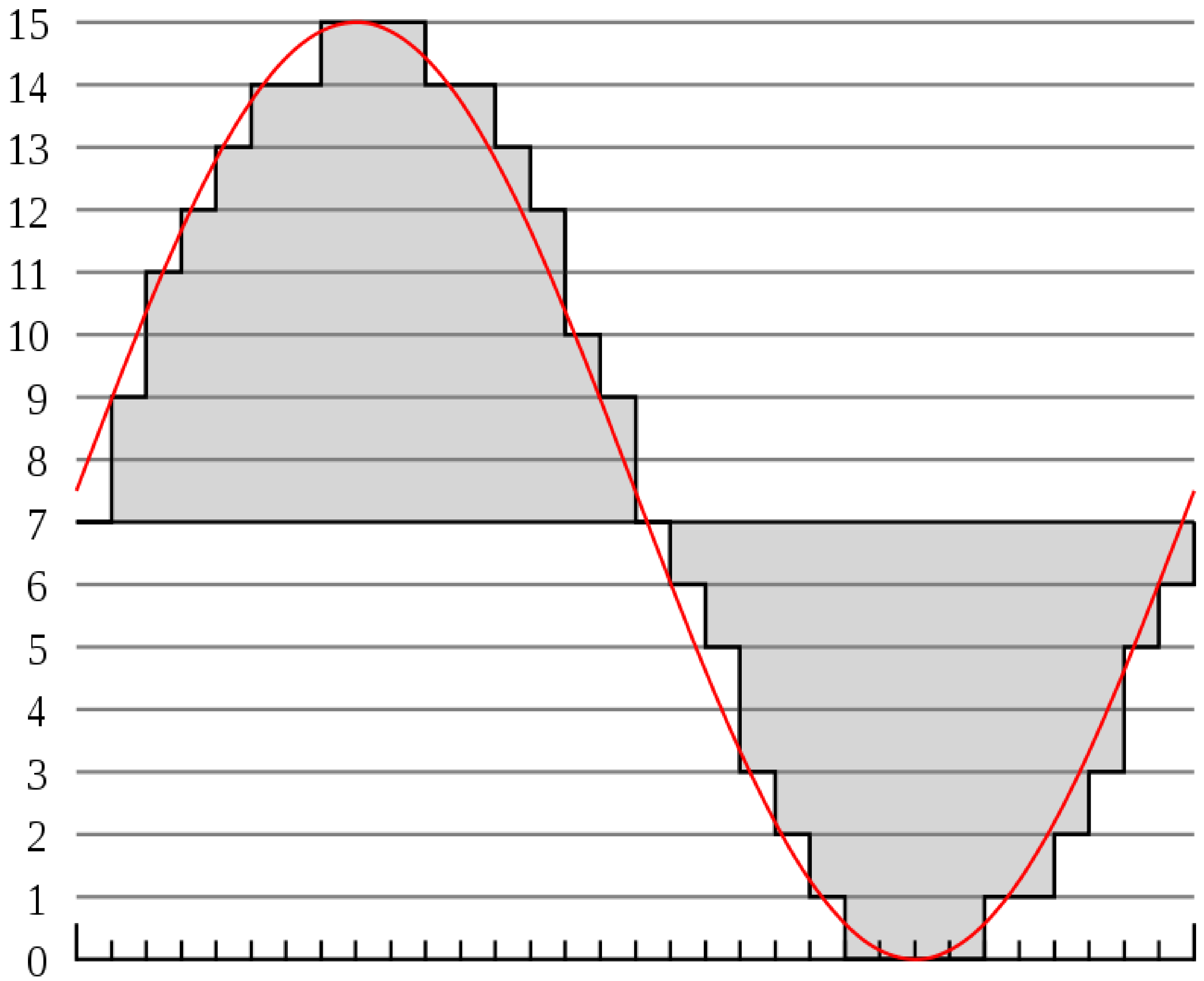
Cold sound comes from cumulative quantization distortion, which produces nasty inharmonic distortion.

- Bob Katz

Take away:

Keep all of your sounds and music at 24 bit, at healthy levels, until the final stage of delivery, in case they need to be processed again.

Sample Rate



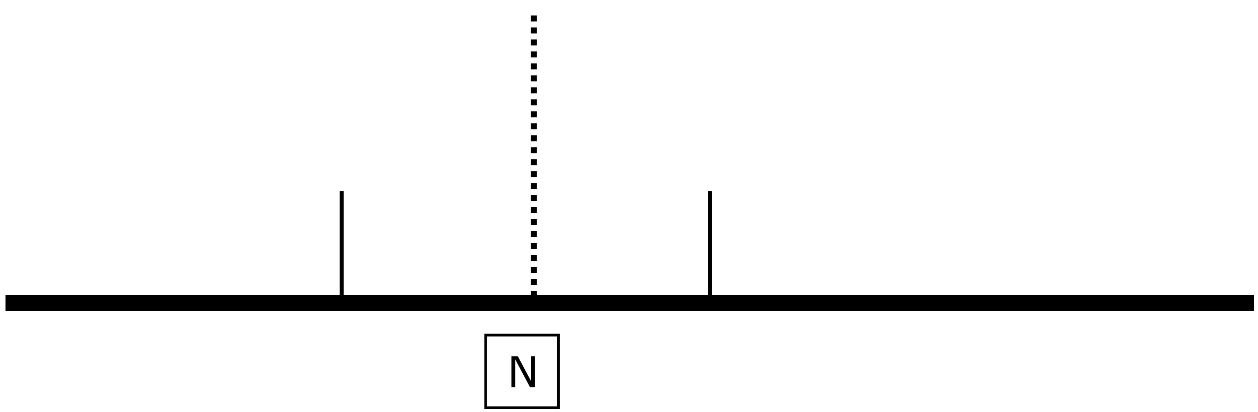
Nyquist Frequency

= 1/2 the sampling rate

At 44.1 kHz Nyquist frequency = 22.05kHz

Nyquist frequency

A frequency sampled above the Nyquist frequency will be mirrored centered on the Nyquist frequency

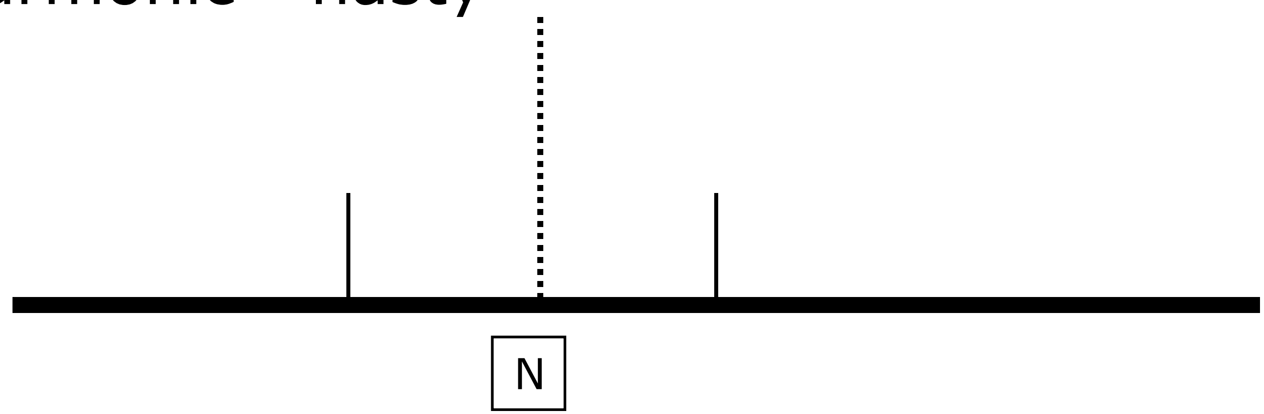


Nyquist frequency

Example:
Sample rate = 2000Hz
Nyquist frequency = 1000 Hz

A frequency of 1500Hz would wrap around to 500 Hz.

Enharmonic - nasty



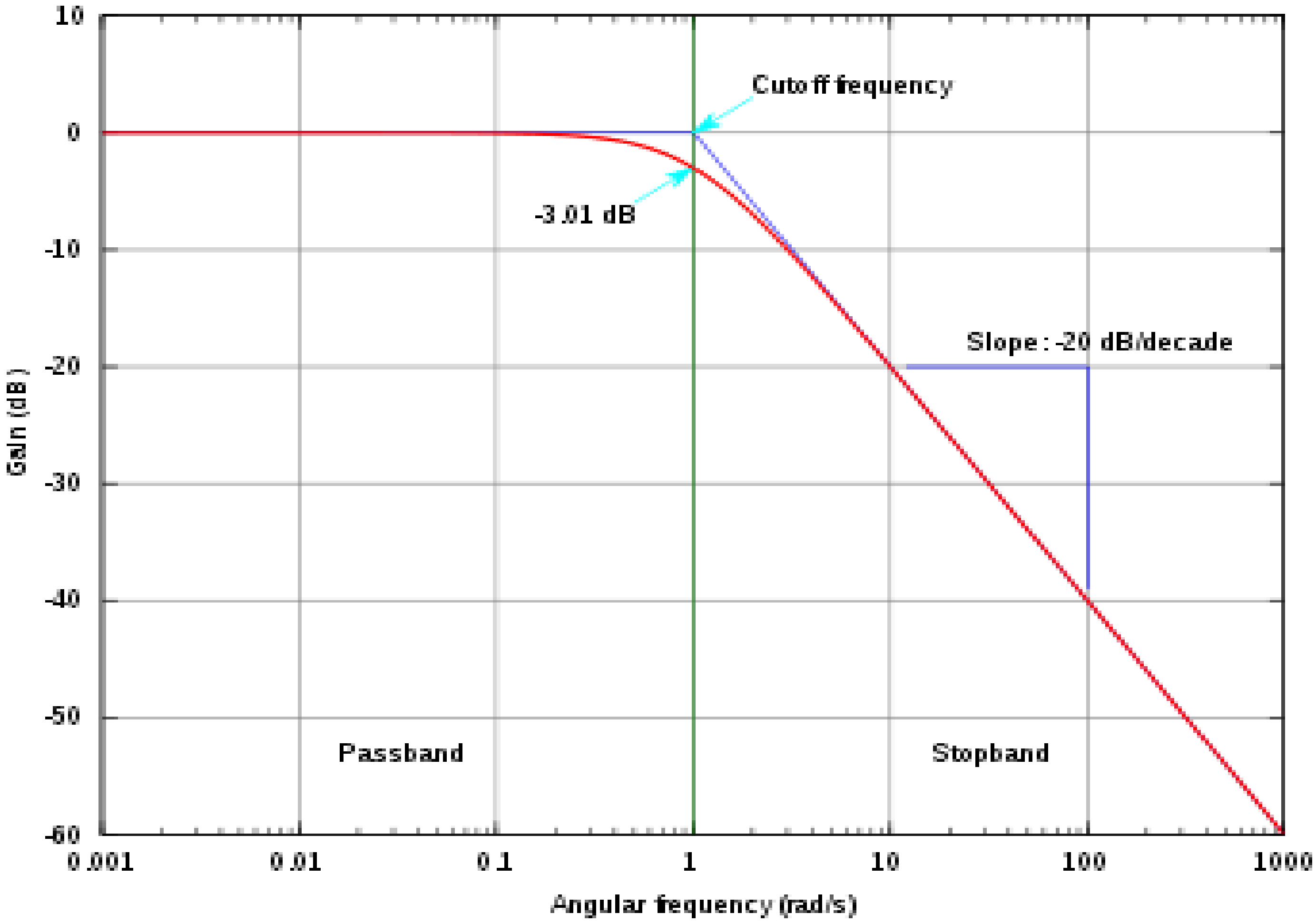
Sampling - filters

Solution is a filter to remove frequencies above the Nyquist frequency.

No such thing as a brickwall filter (instant drop off of frequencies above the cutoff)

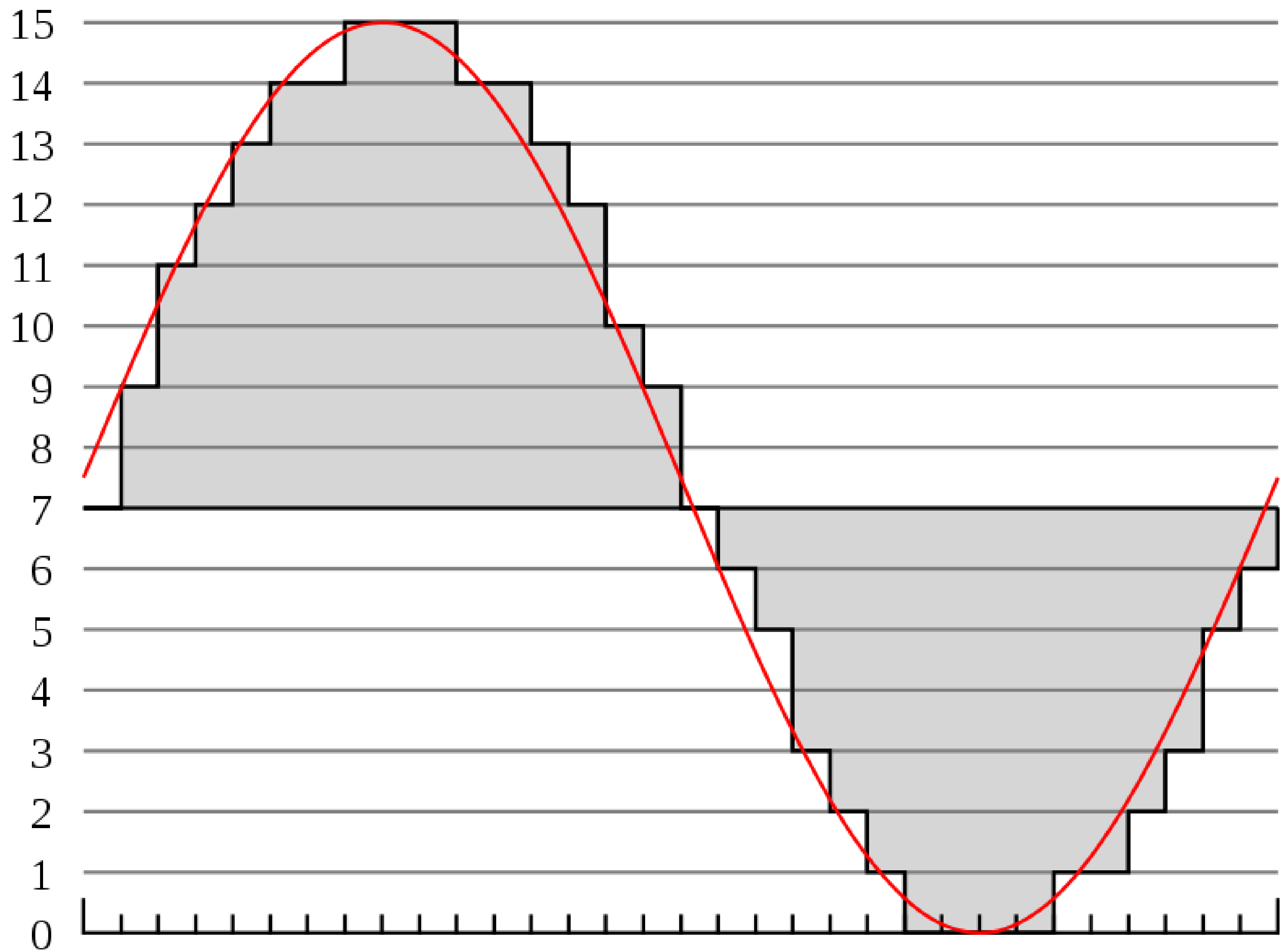
Filter needs to have the cutoff set below the Nyquist to allow for rolloff.

Sampling - filters

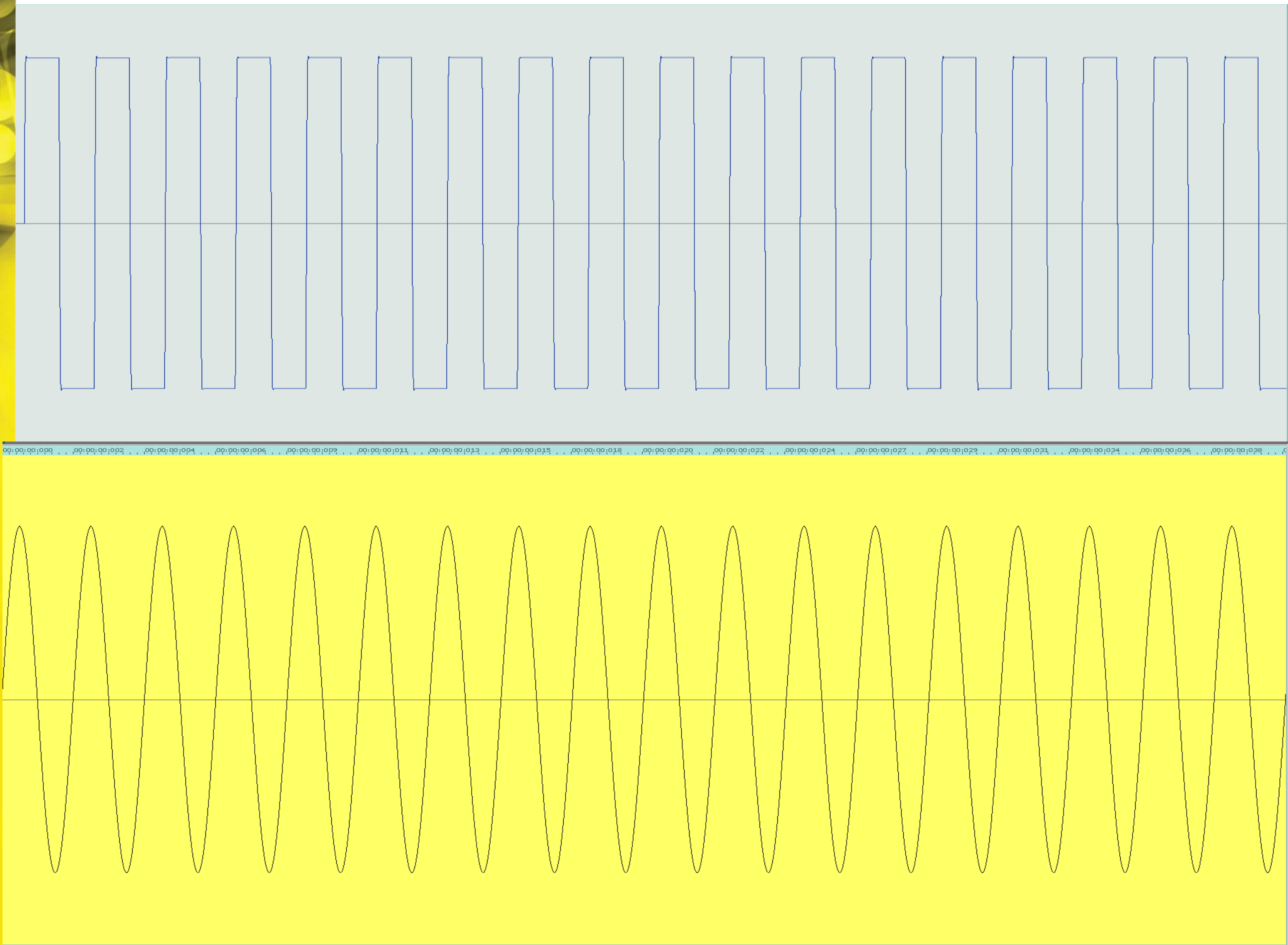


Source: Wikipedia

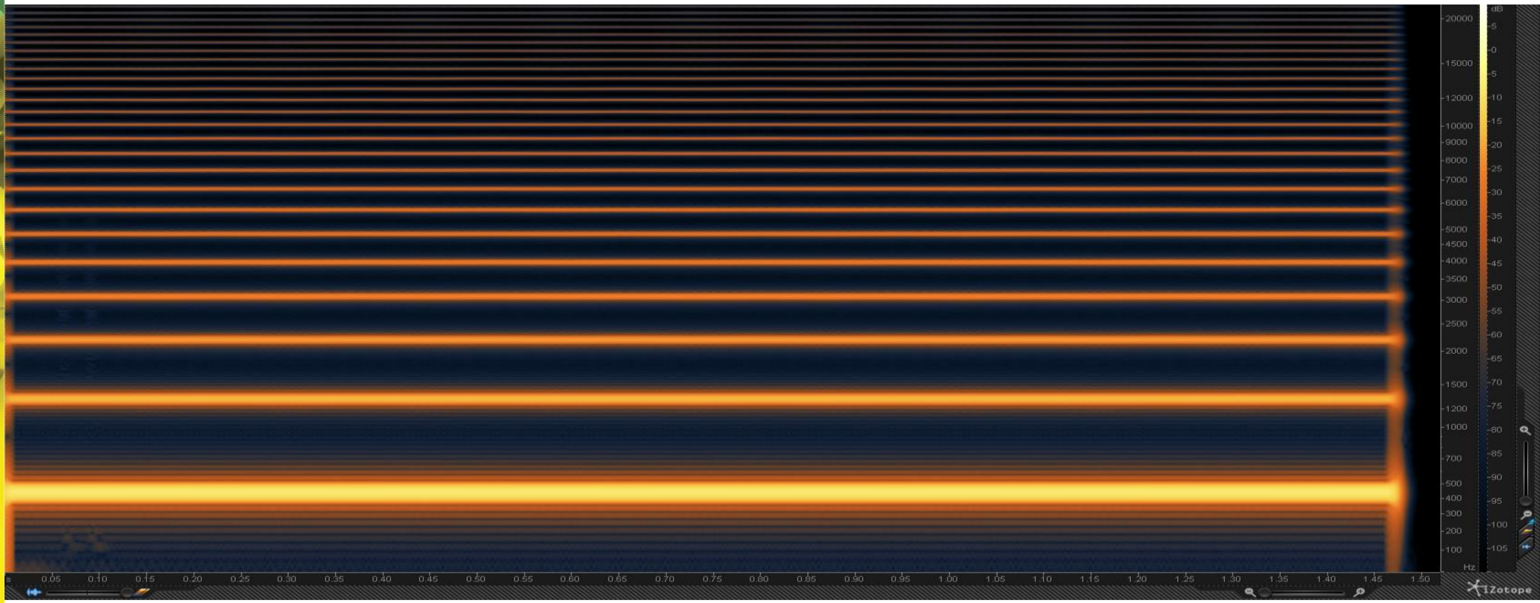
Playback: Reconstruction filter

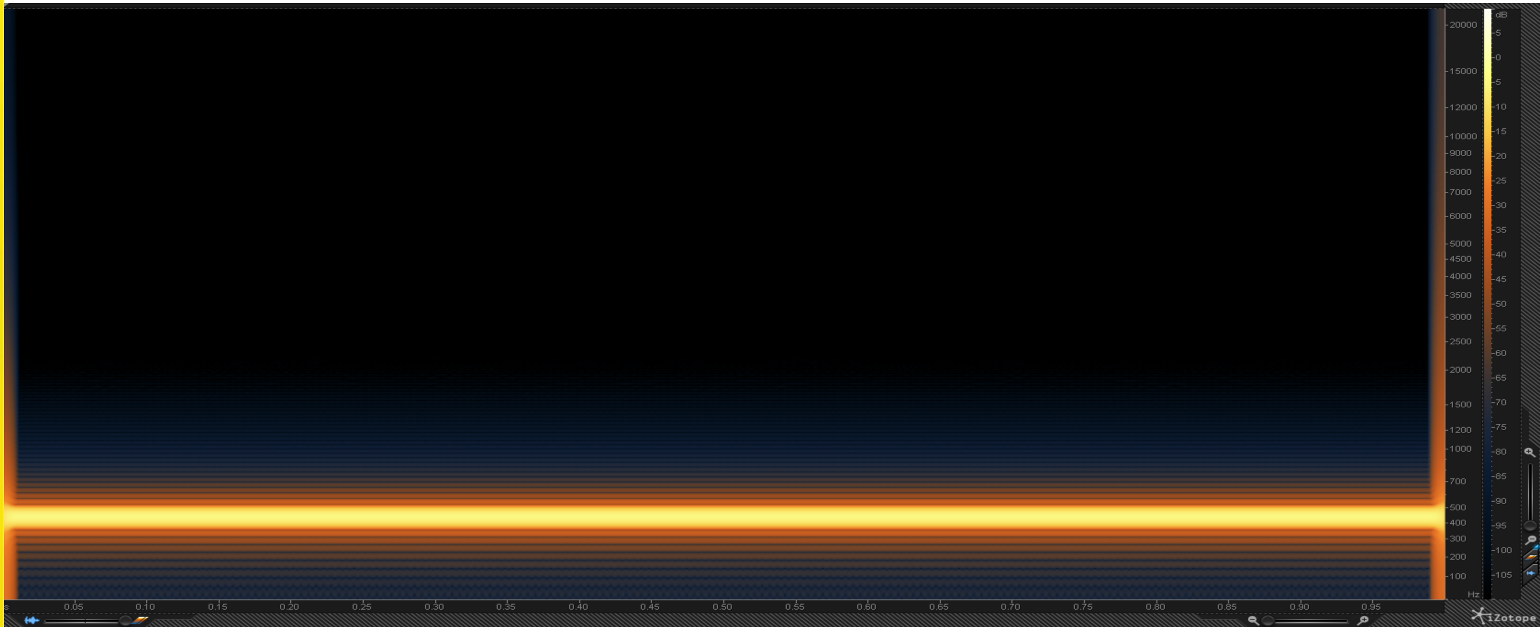


Playback: remove harmonics



Playback - remove harmonics





Digital Audio Best Practices

Use the bits:

Record and mix at 24 bit
Leave headroom of -12db Peak

Keep the levels healthy all the way through the chain through proper gain staging.

Digital Audio Best Practices

Conversely, avoid clipping which will add digital distortion

Clipping will add odd harmonics beyond the Nyquist, (after the input filter), introducing aliasing

- You would have to drop a 24-bit recording by 48 dB to reduce it to 16-bit resolution

Digital Audio Best Practices

Individual sound assets should be kept at healthy levels when they’re put into the game

You can always turn them down in the audio engine.

If you have to boost levels, you’ve already lost the extra bits of information

Digital Audio Best Practices

Normalizing raises the noise floor with it.

But

Converting 24 bit to 16 drops the last 8 bits, so you may want to move the information into the top 16 bits

Just chopping off the last 8 bits is bad (truncation) - adds quantization distortion

Digital Audio Best Practices

Solution:

Dither - random low level noise added to a signal

Should only be used when converting to a lower bit depth (for integration into game)

Repeated dithering will add a veil of noise to the sound - do it only once.

That’s why we keep our sounds at 24 bit

Normalization pitfalls

Groups of sounds that will playback as randomly chosen samples (eg. footsteps) should be batch normalized.

Normalizing individual speech files is a bad idea (whisper, shout)

Solution:

Wavelab batch normalize, or Dolby DP600

Digital Audio Best Practices

Watch out for unnecessary low frequencies!

- They have a lot of energy, take up a lot of bits
- DSP processes can add them
- DC offset can add it
- Your speakers may not reproduce them
- They muddy up your mix
- Cheap subwoofers will sound boxy and overloaded

Digital Audio Best Practices

- Consider putting a high pass filter at the end of a processing chain on your channel strip

- Put a high pass filter on your Master Bus in your sound design template to catch any stray subsonic frequencies. Set to ~100Hz

- Leave the low frequencies for sounds that really need them (explosions), which will leave room for them, and they'll sound bigger

Compression (Data reduction)

Data compression such as Zip don’t work well on audio.

-better suited for text files

Compression (Data reduction)

WMA, mp3 are optimized for audio

Use psychoacoustics to identify parts of the spectrum that listener can't hear

- -low level frequencies at the same time as loud frequencies in another part of the spectrum
- -Frequencies getting masked.

-That information is thrown out, or coded with less accuracy.

-They are LOSSY compression schemes

Compression (Data reduction)

Help the algorithm by filtering out irrelevant frequencies
-especially high or low frequencies

-A distant sound may not need crisp high frequencies for example.

-Leaves more information in the frequencies that matter

Questions?

Presentation at
http://www.omniaudio.com/GDC09

Alistair Hirst
AH@OmniAudio.com

Useful links

Bit depth http://en.wikipedia.org/wiki/Audio_Bit_Depth
Aliasing: http://en.wikipedia.org/wiki/Aliasing
Oversampling: http://en.wikipedia.org/wiki/Oversampling
Distortion: http://www.geofex.com/effxfaq/distn101.htm

http://www.soundonsound.com/sos/feb08/articles/digitalaudio.htm

http://www.sfu.ca/sonic-studio/handbook/Alphabet_list.html#Q_Anchor

http://www.digido.com/media/articles-and-demos.html

Formants: http://ccrma.stanford.edu/~jmccarty/formant.htm

http://www.sounddevices.com/notes/recorders/real-world-24-bits/