

Project Number: FB IQP HB10

HIGH BANDWIDTH ACOUSTICS – TRANSIENTS/PHASE/ AND THE HUMAN EAR

An Interactive Qualifying Project
Submitted to the Faculty
of the
WORCESTER POLYTECHNIC INSTITUTE
In partial fulfillment of the requirements for the
Degree of Bachelor of Science
By
Samir Zutshi
8/26/2010

Approved:
Professor Frederick Bianchi, Major Advisor
Daniel Foley, Co-Advisor

Abstract

Life over 20 kHz has been a debate that resides in the audio society today. The research presented in this paper claims and confirms the existence of data over 20 kHz. Through a series of experiments involving various microphone types, sampling rates and six instruments, a short-term Fourier transform was applied to the transient part of a single note of each instrument. The results are graphically shown in spectrum form alongside the wave file for comparison and analysis. As a result, the experiment has given fundamental evidence that may act as a base for many branches of its application and understanding, primarily being the effect of this lost data on the ears and the mind.

Table of Contents

1. List of Figures	3
2. Introduction	5
3. Terminology	5
4. Background Information	11
a. Timeline of Digital Audio	12
5. Preparatory	17
6. Theory	17
7. Zero-Degree Phase Shift Environment	17
8. Short Term Fourier Transform	20
9. Experimental Data	22
10. Analysis and Discussion	34
11. Conclusion	36
12. Bibliography	36

List of Figures

1. Amplitude and Phase diagram	6
2. Harmonic Series	8
3. Reduction of continuous to discrete signal	9
4. Analog Signal	9
5. Resulting Sampled Signal	9
6. Sample Waveform	10
7. Diagram of the ear	15
8. Sampling Error Formula	20
9. Masking Curves	21
10. Recording Studio Setup	22
11. Table of 192 kHz data	22
12. FlexPro STSF Settings	24
13. Conga 192 kHz GRAS ½ “	25
14. Conga 192 kHz GRAS ¼ “	25
15. Conga 192 kHz SHURE	26
16. Guitar 192 kHz GRAS ½ “	26
17. Guitar 192 kHz GRAS ¼ “	27
18. Guitar 192 kHz GRAS SHURE	27
19. Trumpet 192 kHz GRAS ½”	28
20. Trumpet 192 kHz GRAS ¼ “	28
21. Trumpet 192 kHz SHURE	29
22. Violin 192 kHz GRAS ½ “	29

23. Violin 192 kHz GRAS ¼ “	30
24. Violin 192 kHz SHURE	30
25. Vocals 192 kHz GRAS ½ “	31
26. Vocals 192 kHz GRAS ¼ “	31
27. Vocals 192 kHz SHURE	32
28. Guitar 192 kHz various comparison	34

Introduction

There has been a growing amount of interest surrounding frequencies above 20 kHz and its influence on human perception, especially in music. Currently there is some amount of conflict regarding the significance of this high frequency audio, but little of the existing research is concerned about closely examining the *transient* portion of the signals, instead making the more steady-state aspects of musical signals the foundation of the analysis. While there may be a considerable amount of research into explaining if and how sampling rate will impact fidelity, there is not a substantial amount of investigation into how *sampling rate* choice could impact important characteristics of a transient such as its rise and decay times, peak sound pressure levels, and the integrity of the impulse's phase across its frequency range. We intend to examine how the sampling rate and microphone bandwidth can impact the phase characteristics of the transient parts of these waveforms and how the resulting changes manifest themselves acoustically.

There hasn't been significant research that analyzes music using high bandwidth equipment to address the phase shift due to the analog electronics in the recording chain and the anti-aliasing filters. Therefore, there is a push to explore the signal path and more closely analyze these waves focusing on the transients and the phase shift that so vividly exists in the sound.

Terminology

All periodic waves have three primary attributes that help describe them. The simplest of these is amplitude, which describes the magnitude of the wave. A higher amplitude wave has a greater peak-to-peak distance than a low one. Frequency is the number of oscillations per unit of time (e.g. one second) that a wave goes through. In the context of audio, each wave refers to the local, instantaneous variations of air pressure away from the ambient atmospheric pressure which

produce the sensation of sound when experienced by the auditory system. A higher amplitude sound wave means a louder sound, whereas a higher frequency wave makes a higher pitched sound. A sinusoidal wave is the only waveform with a single “pure” frequency, with all other waveforms involving multiple frequencies, such as sawtooth, square, etc...

The third primary attribute of waves is normally less perceivable to humans. This attribute is phase, the point along a waveform some fraction of a complete oscillation between the start of a given wave cycle and the start of the next one. There are 360 degrees of phase along a wavelength. Position within a cycle is usually thought of as relative to some benchmark such as “at 0° there is no local instantaneous pressure deviation from the ambient.” When a wave’s phase is altered by repositioning this benchmark by feeding it forward or lagging it behind the reference, this is known as a *phase shift*. Phase shift is frequently the result of time delay in audio systems. A frequently used tool for analyzing audio signals is Fourier analysis, a mathematical technique which reduces a complex, multi-frequency waveform into an array of sine waves which at each component frequency has its own amplitude and phase (shown below).

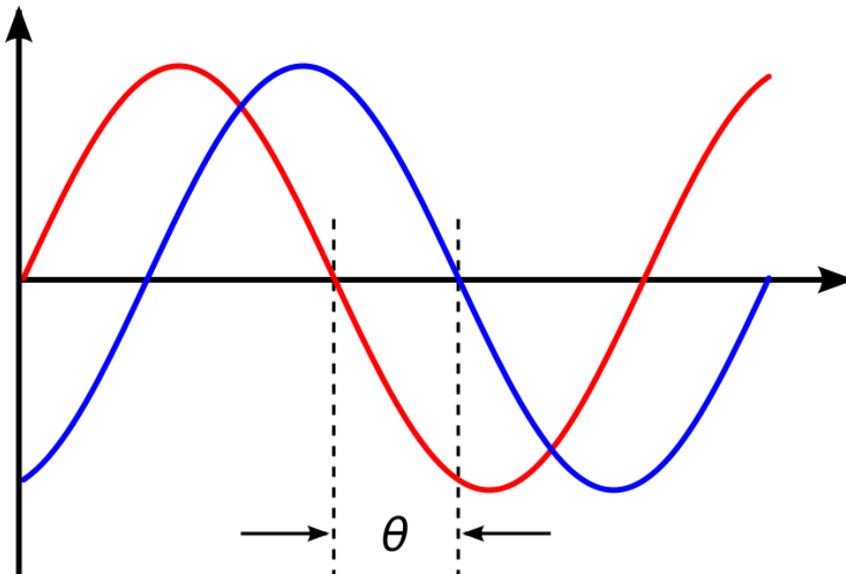


FIGURE 1: (Source: wikipedia.org)

Any complex tone "can be described as a combination of many simple periodic waves (i.e., sine waves) or partials, each with its own frequency of vibration, amplitude, and phase."(Source: Thompson 2008). A partial is any of the sine waves by which a complex tone is described.

A harmonic (or a harmonic partial) is any of a set of partials that are whole number multiples of a common fundamental frequency.(Source: Pierce 2001) This set includes the fundamental, which is a whole number multiple of itself (1 times itself). The musical pitch of a note is usually perceived as the lowest partial present (the fundamental frequency), which may be the one created by vibration over the full length of the string or air column, or a higher harmonic chosen by the player. The musical timbre of a steady tone from such an instrument is determined by the relative strengths of each harmonic. (Source:Wiki)

The harmonic series is an arithmetic series ($1 \times f$, $2 \times f$, $3 \times f$, $4 \times f$, $5 \times f$, ...). In terms of frequency (measured in cycles per second, or hertz (Hz) where f is the fundamental frequency), the difference between consecutive harmonics is therefore constant and equal to the fundamental. But because our ears respond to sound nonlinearly, we perceive higher harmonics as "closer together" than lower ones. On the other hand, the octave series is a geometric progression ($2 \times f$, $4 \times f$, $8 \times f$, $16 \times f$, ...), and we hear these distances as "the same" in the sense of musical interval. In terms of what we hear, each octave in the harmonic series is divided into increasingly "smaller" and more numerous intervals.(Wiki) A musical representation of the harmonic series is shown below.

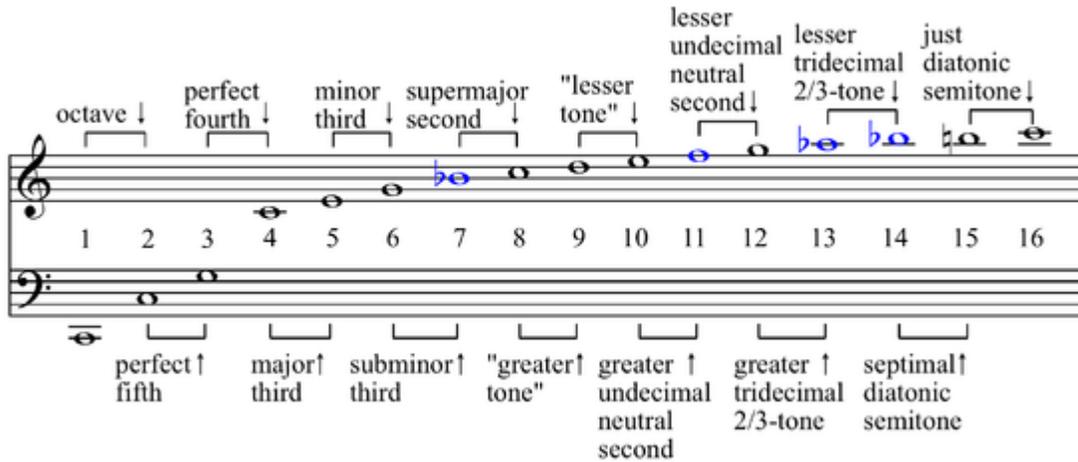


Figure 2: (Source: wikipedia.org)

Bandwidth is the difference between the upper and lower frequencies in a contiguous set of frequencies. Similarly, frequency response is the measure of a system's output spectrum in response to an input signal, typically characterized by output magnitude or phase. This information is sometimes shown as a graph, and frequency response curves are often used to indicate the accuracy of audio systems and components, with a "flat" response being one where all desired input signals are reproduced without attenuation or emphasis of a particular frequency band.

The range of frequencies that can be heard by humans is known as the hearing range. It is often said that the audible range of frequencies of humans is 20 Hz to 20 kHz. There is some variation pertaining to sensitivity by age and gender, with children under 8 able to hear as high as 22 kHz and the elderly only sensitive to about 12-17 kHz, and women typically being more sensitive to higher frequencies than similar-aged men. The "20 kHz debate" is the difference in opinion as to whether sound above approximately 20 kHz is perceptible or meaningful when storing or playing back audio information. The word 'ultrasonic' describes the cyclical pressure variations that are beyond the upper limit of human hearing range. In air, ultrasonic frequencies begin around 20 kHz

and continue up through about 1 megahertz, at which point the wavelength size approaches that of the size of its constituent molecules.

The reduction of a continuous signal to a discrete signal is called sampling. In **FIGURE 3**, the green line represents the source continuous signal, where the blue dots represent the sampled discrete signal.

Sampling rate can be defined as the number of samples per some unit of time taken from a continuous signal to make the resulting discrete signal.

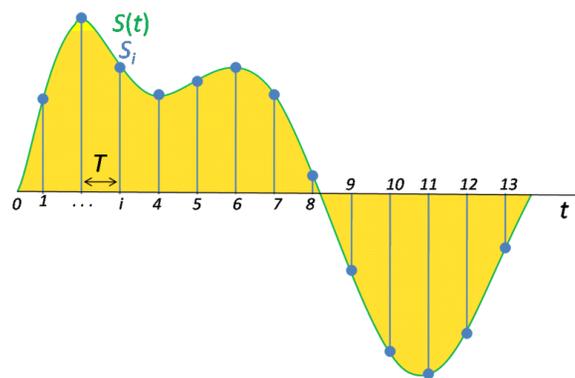


FIGURE 3. Source: wikipedia.org

If you take a look at **FIGURE 5** you can see the resulting sampled signal from the original analog signal

FIGURE 4 where T in **FIGURE 5** is one interval of time.

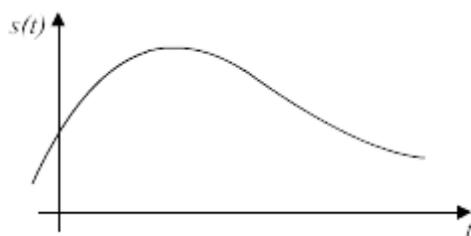


FIGURE 4

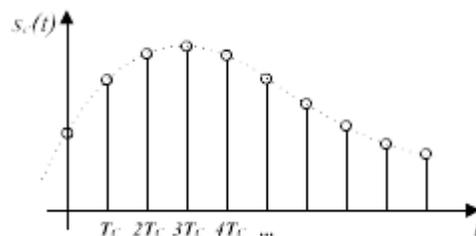


FIGURE 5

The Nyquist frequency is half the sampling frequency of a discrete signal processing system, sometimes referred to as the folding frequency. The Nyquist frequency is not the same as the Nyquist rate. It is important to understand that this frequency is a property of a discrete-time system, not of a signal. The Nyquist rate, however, is a property of a continuous-time signal and not of the system responsible for recording (sampling) it. (Source: 1)

The segments with near vertical rise, relatively high magnitudes, and fast decays shown in **FIGURE 6** are known as transients. These are very short duration signals that represent a portion of musical sound much higher in amplitude and more broadband in nature than the harmonic content of that sound.

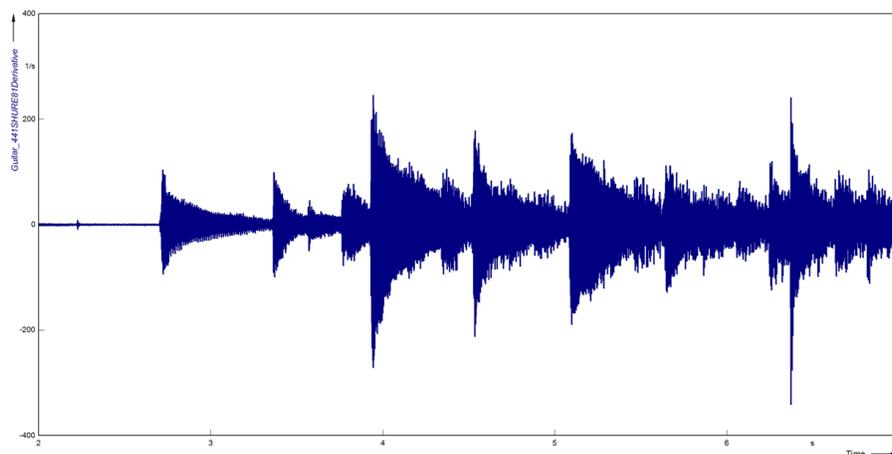


FIGURE 6. Experimental data

The zero-degree phase shift environment is an environment where the signal from an acoustic source (e.g. musical instrument or voice) exhibits no phase shift between that source and the listener's ear drum. This excludes any phase shift due to the listening environment (e.g. room reverberation). On the other hand, a non-zero degree phase shift environment is an environment where man-made devices (such as microphones, audio amps, A/D and D/A converters,

compressors, etc) introduce phase shift of the acoustic signal (potentially more prevalent at certain frequencies than others) between the source and the listener's ear drum.

Background Information

Nyquist began looking into negative-feedback amplifiers and how to determine when they become stable. His findings resulted in the development of the Nyquist Stability Theorem (Source: Wikipedia), which during World War II helped control artillery that used electromechanical feedback systems. Nyquist determined that the number of independent pulses that could be put through a telegraph channel per unit time is limited to twice the bandwidth of the channel. His rule essentially states what is more commonly known as the Nyquist-Shannon sampling theorem. Nyquist's work shows up prominently when working with feedback loops in what is known as the "Nyquist plot" (Source: Wikipedia) which relates the magnitude of a frequency response with its phase. His work is important in research concerning auditory perception because of his insight into the relationship of the potential frequency and phase of signals in bandlimited channels.

As the capabilities of technology continue to advance, the ability to record acoustical signals is advancing along with it. The first digital recordings were crude, short melodies played by massive supercomputers, but increases in all aspects of the digital technology domain have helped digital audio become a high fidelity medium. One of the most attractive aspects of digital audio is that when processing a digital signal, one can implement any audio effect that one understands the mathematics behind, without having to build a discrete circuit. All operations on the sound are arithmetical calculations on the detected voltage values, instead of having the voltages pass through hardwired analog audio circuitry. The process of detecting sound, converting it into a discrete

stream of bits, storing them in one form or another, and playing them back is desirable for many reasons when computationally feasible.

Timeline of the development of digital audio technology

1951 – The first digital recordings are played, by the CSIRAC in Australia and then the Ferranti Mark 1 in the UK, only a few weeks apart.

1959 – The integrated circuit (IC) or microchip is developed, revolutionizing the world of electronics and unlocking the processing power needed for viable digital audio to become a reality.

1962 – The first PCM system is developed at Bell Labs. This concept is still very relevant in the way digital music is stored today.

1967 – Richard C. Heyser devises the “TDS” (Time Delay Spectrometry) acoustical measurement scheme, which paves the way for the revolutionary “TEF” (Time Energy Frequency) technology.

1969 – Dr. Thomas Stockham begins to experiment with digital tape recording after experimenting on the TX-0 mainframe at MIT Lincoln Laboratory.

1971 – Denon demonstrates 18-bit PCM stereo recording using a helical-scan video recorder

1975 – Digital tape recording begins to take hold in professional studios

1976 – Dr. Stockham, now having founded Soundstream, makes a pioneering 16-bit digital recording at a 37 kHz sample rate of the Santa Fe Opera.

1978 – The first EIAJ standard for the use of 14-bit PCM adaptors with VCR decks is embodied in Sony's PCM-1 consumer VCR adaptor.

1980 – Many manufactures begin introducing multi-track digital recorders.

The first edition of the Compact Digital Audio Disc (CD) technical specification is released by

Philips and Sony. CD recordings are stored at audio bit depth of 16 at a sample rate of 44.1 kHz. This yields an audio signal with 96 dB of dynamic range and digital bit rate of 1.35 megabits per second.

1981 – IBM introduces a 16-bit personal computer, the PC 5150. The Personal Computer architecture's flexible modular design and open technical standard help the architecture to become incredibly popular.

1982 – Sony releases the first CD player, the Model CDP-101. They also release the PCM-F1, the first 14- and 16-bit digital adaptor for VCRs. Intended for the consumer market; it is eagerly snapped up by professionals, sparking the digital revolution in recording equipment.

1984 – The Apple Corporation markets the Macintosh computer.

1986 – The first digital recording consoles begin to appear.

1987 – Digidesign markets "Sound Tools," a Macintosh-based digital workstation using Digital Audio Tape as its source and storage medium.

1988 – Apple, alongside Electronic Arts, develops the Audio Interchange File Format (AIFF), one of the first standards for storing audio bitstreams on PCs.

1991 – Alesis unveils the ADAT, the first "affordable" digital multitrack recorder, capable of recording 8 tracks of 16-bit audio.

1992 – The premiere of Dolby Digital which featured sample rates of up to 48 kHz and 5.1 channel audio. This technology begins to enter the home by 1995.

1993 – Improving storage capabilities of personal computer hardware, such as internal hard drives and the Iomega Zip drive render digital tape—and the analog alternative—obsolete.

1994 – Yamaha unveils the ProMix 01, the first "affordable" digital multitrack console.

1996 – Experimental digital recordings are made at 24 bits and 96 kHz, such as those by the Nagra D from the Kudelski Group.

1997 – DVD videodiscs and players are introduced. An audio version with 6-channel surround sound is expected to eventually supplant the CD as the chosen playback medium in the home.

1998 – MPEG Audio Layer 3 (MP3) is developed, allowing the compression of CD-quality sound by a factor of 812.

1999 – Audio DVD Standard 1.0 agreed upon by manufacturers.

Today devices supporting audio playback at 192 kHz sample rates at 24-bit depth are computer commodities, and recording interfaces which support capture at that rate are becoming more prevalent as well. Digital storage to hold recorded music is cheaper than ever, and the processing power needed for high sample rate/high bit-depth recording and playback continues to plummet.

Auditory Perception

The hearing mechanisms of the human ear are remarkably complex. Our incomplete understanding of how the auditory system works gives rise to the growing amount of skepticism by some that sounds above 20 kHz are irrelevant.

The (in)ability of our auditory organs to perceive sounds above 20 kHz is not a fate sealed by inherent flaws in their design. The same basic auditory structures are found in almost all mammals, with dogs able to hear at around 60 kHz and bats to 100 kHz, and these are only a few prominent examples. Thus it is unlikely that the upper limits of our hearing are restricted by our body's use of the specialized hair cells, *stereocilia*, found on the basilar membrane. Ultrasonic emissions can exist at frequencies up to about 1 megahertz in air, where the limiting factor becomes the size of oxygen molecules. When the wavelength of the density change becomes approximately equal to or shorter

than the size of the molecule itself, that wave can no longer be propagated, because the molecule is not itself compressible (Moulton 2000).

“At very high frequencies, the inertia of the ear drum limits our ability to detect sounds. The change in pressure away from average pressure and back happens for such a short period of time that the ear drum does not have time to move in response.” David Moulton

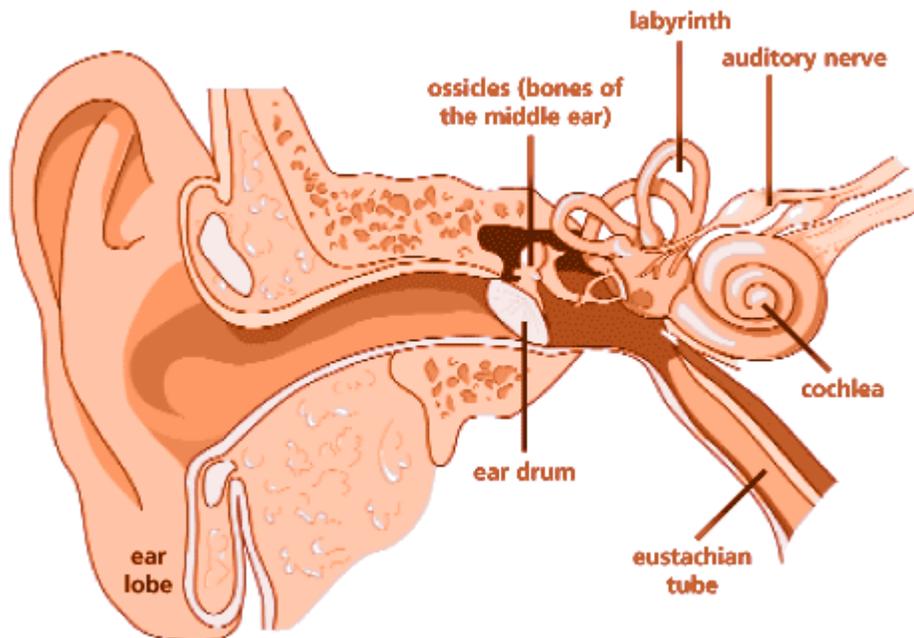


FIGURE 7: DIAGRAM OF THE EAR (SOURCE:HANS WORDPRESS)

There are researchers who immediately contest the influence of frequencies above 20kHz and those who believe these high frequencies are important to accurate auditory perception.

One of the strongest arguments for the influence of high frequency content was presented by Tsutomu Oohashi and claims that sound reproduced above 26 kHz “induces activation of alpha-EEG (electroencephalogram) rhythms that persist in the absence of high frequency stimulation and can affect perception of sound quality” (Oohashi et al., 2000).

Oohashi attempts to provide evidence that sounds containing high-frequency components (HFC) above 20 kHz significantly affect the activity listeners' brains using "noninvasive psychological measurements of brain responses." Though the subjects could not recognize the HFC when presented alone, their brain activity altered significantly when presented with music containing the HFC simultaneously with their corresponding low frequency components (LFC) compared to just those low frequency components alone.

Lenhardt et al. (Source:Lenhardt) report that "bone-conducted ultrasonic hearing has been found capable of supporting frequency discrimination and speech detection in normal, older hearing-impaired, and profoundly deaf human subjects" in a paper published in *Science*. This research concerned neural connections between the "saccul" ("an otolithic organ responsible for transduction of sound after destruction of the cochlea, and that responds to acceleration and gravity" (Source: Boyk:caltech) and the cochlea. Similarly an experiment done by Kimio Hamasaki et al. found that the duration of the stimulus (how long the sounds were listened to; in other words, a short split-second burst versus a longer musical passage) might affect discrimination between sound perception with and without high frequency components (Hamasaki 2004). Though they were not able to find significant evidence through [\[for?\]](#) this paper, Nishiguchi and Hamasaki conjectured that if the extension of the frequency range by high sampling formats affects the perception of sound, it seemed to be caused by reproduction of very high frequency components (Hamasaki 2005).

There has also been research attempting to explain some of these "audible differences" at high and low frequencies. Mike Story suggests that some of the audible differences may be related to defocusing caused by anti-alias filtering, and that the ear is sensitive to energy as well as spectrum (Story 1997).

Preparatory

Several factors can cause signal distortion such as phase shift. Microphones have physical limits, and even then additional factors such as the geometry of the coil, gap, and magnet, as well as the physical properties of its other parts, all play a role in reducing the fidelity of the source signal.

Other factors frustrate researcher's ability to pin down exact performance parameters of audio equipment and the human auditory system.

Theory

Starting from a mathematical aspect, if we take a complex signal such as a musical note from a 440 trumpet in concert A pitch. The content is going to be made up of two elements of the signal, the initial attack of the transient and the steady state part. The steady state part will have the 440 part of the signal and ensuing harmonics that give the listener the characteristics of a trumpet.

Transients have the highest amplitude and highest frequency content. By the definition of a transient, they go from lock background sound level amp to peak sound pressure level in what could be tens of microseconds.

Zero-degree phase shift environment

A zero-degree phase shift environment keeps the transient intact, meaning it is not "distorted" by phase changes caused by man-made devices excluding the listening environment itself. It represents in a "natural setting," a source of music (from instrument or voice) and the receiver which is a person's ear (eardrum). This zero-degree phase shift environment is where there is nothing that alters the phase characteristics (anything man-made) in any manner, whether it be

phase response, frequency response, or amplitude response between the source of the music and the receiver, even a hearing aid is man-made “interference.” The zero-degree phase shift environment does include the listening environment, as even though we say zero-degree, we include the environmental acoustics. So any man-made device that alters phase, amplitude, or frequency between source and receiver would be known as a non-zero-degree phase shift environment.

This leads into the recording of music. Any acoustic music signals have to be recorded using man made devices that will necessarily be in that recording chain, such as microphones, microphone electronics, associated electronics for power supply, perhaps an anti-aliasing filter, and A/D converters. These pieces of the signal chain will modify the phase and or frequency/amplitude content of the signal. Therefore they are elements of this non-zero-degree phase shift environment. In fact, every recording station is by definition a non-zero-degree phase shift environment, because of the man-made devices in between source acoustic signals and the eventual perception of these signals by the eardrum of a listener.

Therefore first we must ascertain the actual characteristics of the musical instruments, in their natural state. In this case our zero-degree phase shift environment will be a point in space where the recording microphone is placed. In case of recording, the microphone is positioned about one inch from the instrument. What this study is looking to achieve is first to characterize the transient characteristics of the musical signals in the zero-degree phase shift environment. We achieve this using a quarter-inch lab microphone with flat frequency response ± 1 dB up/down at 100 kHz, the microphone power pre-amp and microphone power supply are “ruler flat” to within tenths of a dB out to 200 kHz. A high end Orpheus ADD converter hardware was used at its 192 kHz sampling rate. This study involved five musical instruments that were recorded using this equipment: a violin, a trumpet, female vocals, a guitar and conga.

The other element to study is using those data acquisitions as our benchmark, or reference point. The files captured with this data acquisition system we represent as being the “best available.” Using this as our reference, we make the claim that this particular data acquisition system—which is also doubling as a recording system—will have the least amount of phase shift within audio range. This research was done to analyze how the musical transients from these recordings of these five instruments are different. Studying characteristics such as temporal rise and decay, peak levels, frequency content, and phase, as well as how it relates to the standard (the Orpheus reference). We will also examine software with which we can mathematically modify characteristics of phase, frequency content, and amplitude of the reference recordings. The method through which the transients were captured can help us make the 192 kHz recording have the same characteristics as the other recordings. If significant evidence is found, we can say that we know that since the higher frequency range microphone at a higher frequency rate is capturing more data, and is thus more accurately capturing the physical sound wave being produced by the instrument.

This ties into whether or not an instrument produces sound over 20 kHz, if there is actual energy being produced, then this energy will be propagating, and there is chance that a human might respond. Evidence can be shown by putting in various phase shift to the frequencies. There hasn't been enough attention to the fact that, if we mathematically manipulate phase, frequency, and relative amplitude then reference recordings can look like other recordings. We know exactly how these modified transients with less information came to be, and what characteristics they are lacking, both mathematically and aesthetically.

Therefore, there is speculation of evidence of “effect” on listener's ears when listening to a 192 kHz file versus a 44.1 kHz file. If there were a double-blind listening test with the two samples and a sample of at least a thousand people, significant evidence may be found. By the formula below by Gary Klass, the sample size can be determined.

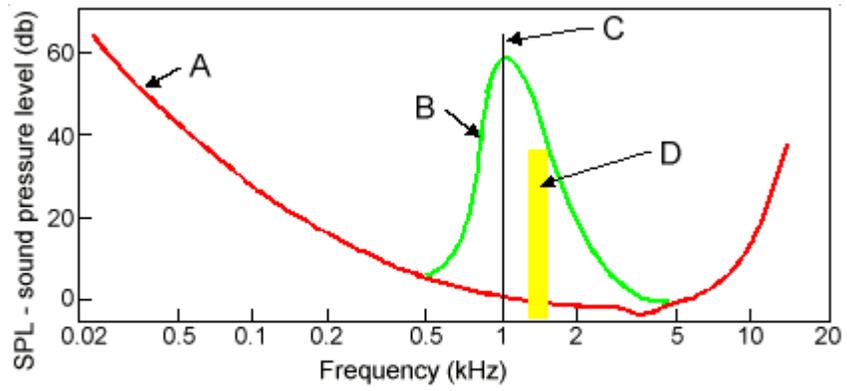
$$\mathbf{S.E.} = \left(\frac{1}{\mathbf{N}} \right) \times 100$$

Figure 8: Sampling Error Formula (Source:Klass)

A sample size of 1600 would produce 2.5% error rate, thus at least a thousand people would yield significant evidence. There have been similar tests done in the past, and if there is no data found, the next step would be to utilize EEG tests to see if there is any extraneous activity in the brain showing evidence of a difference between the two recordings.

Short-term Fourier Transform (STSF)

The **short-time Fourier transform (STFT)**, is a related transform used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. A key factor in the analysis done to-date is that the amplitude axis is displaying the sound levels on a linear scale and to a log (decibel) scale. This will severely mask amplitudes that can be 1/100 of a peak value but a ratio of 1:100 is -40 dB and acoustic signals 40 dB below a peak level are quite audible. There is a need for data analysis on a log, and not linear, amplitude scale. The analysis of transient portions of the recorded signals is possible using shareware audio editing programs like Audacity. Also, the FFT spectrum analysis provided by Audacity will display data on a log (dB) scale so this will enable visibility on what energy exists (if any) above 20 kHz on the transient portions of the wav files.



- A Normal threshold of hearing
- B Modified threshold due to tone C
- D Band of noise rendered inaudible by the presence of tone C

Figure 9: Masking Curves (Source:MPEG Audio)

Experimental Data

The setup used during the recording sessions is shown below.

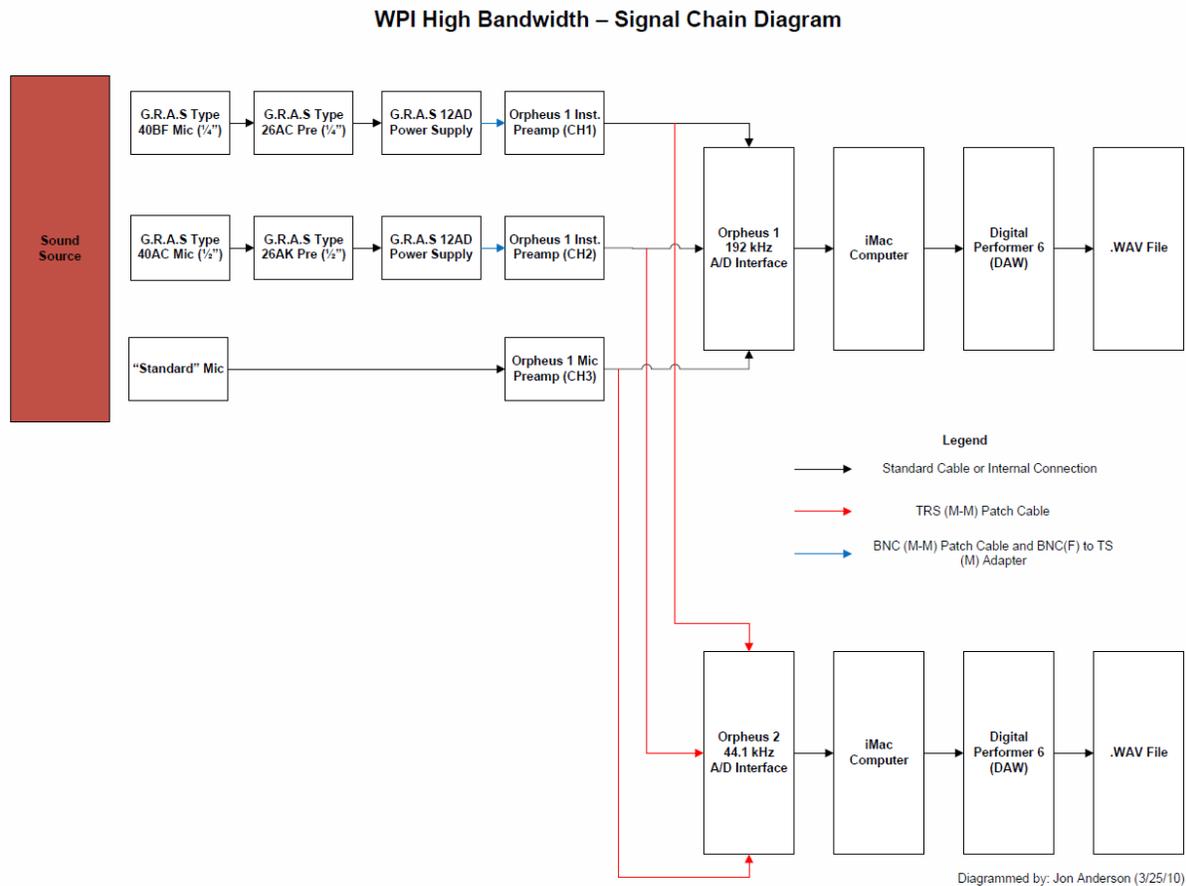


Figure 10: Recording Studio Setup (Source: Jon Anderson)

Recording Type:	Evidence of information over 20kHz?(With STSF Transform)
Conga 192 kHz GRAS ½ "	Yes
Conga 192 kHz GRAS ¼ "	Yes
Conga 192 kHz SHURE	Yes
Guitar 192 kHz GRAS ½ "	Yes

Guitar 192 kHz GRAS ¼ “	Yes
Guitar 192 kHz SHURE	Yes
Trumpet 192 kHz GRAS ½ “	Yes
Trumpet 192 kHz GRAS ¼ “	Yes
Trumpet 192 kHz SHURE	Yes
Violin 192 kHz GRAS ½ “	Yes
Violin 192 kHz GRAS ¼ “	Visible but not significant
Violin 192 kHz SHURE	Yes
Vocals 192 kHz GRAS ½ “	Yes
Vocals 192 kHz GRAS ¼ “	Yes
Vocals 192 kHz SHURE	Yes

Figure 11: 192 kHz files over 20 kHz info

The data below was generated by extracting a small portion of the wave file and importing into a software used for analysis and presentation of scientific data called FlexPro by Weisang GmbH. Once imported, each file was treated to a Time-Frequency Analysis by method of the short-term Fourier Transform. The general parameters for the transform are shown in the following image, though the db range varies between image. This was done in order to convey the best representation of the experimental findings.

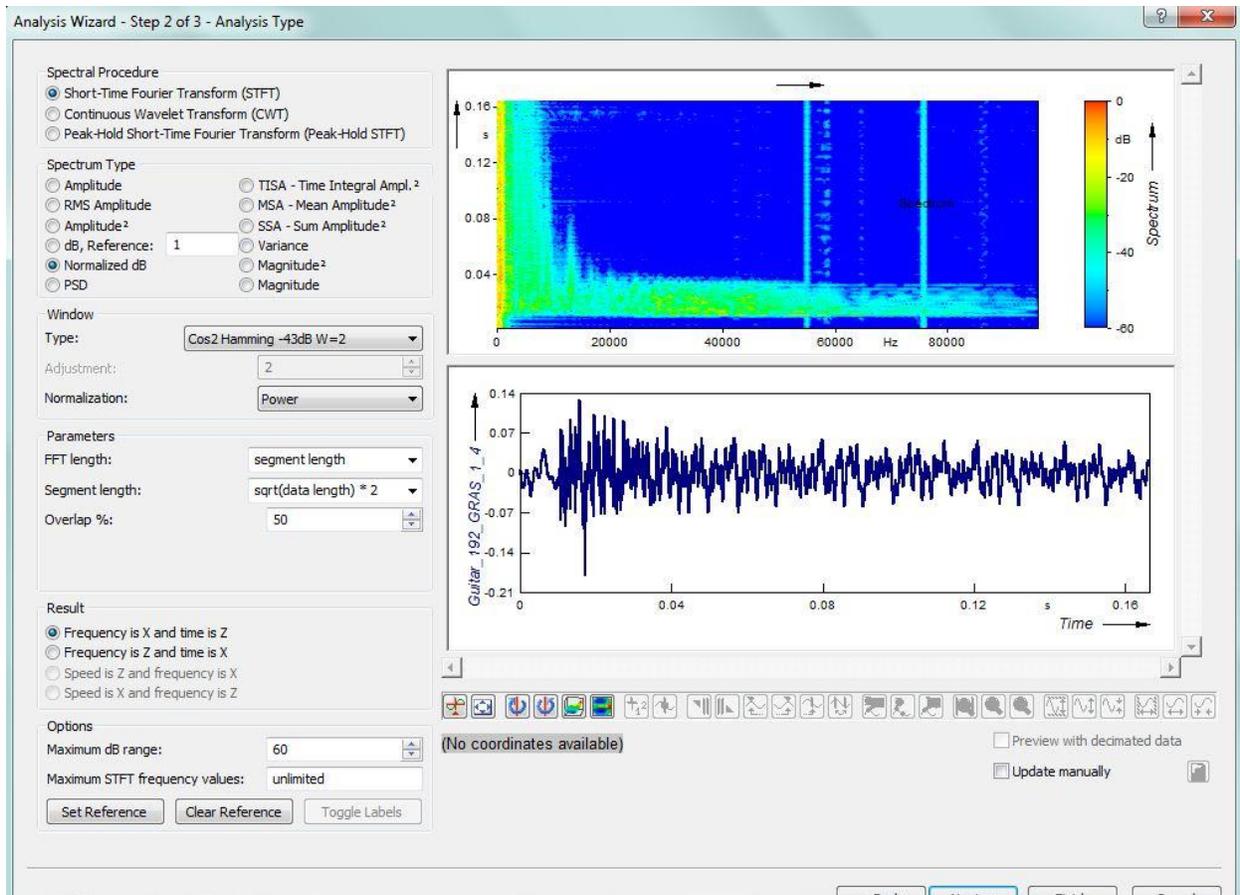


Figure 12: FlexPro STSF Settings

The experimental data below draws a parallel between a wave file and the STSF of the wave file. These tests were done using all 6 microphones and data regarding the 192 kHz files are displayed below. To recall, the goal of the tests were to discover and analyze information that existed over the 20 kHz bound and attempt to define its purpose. Naturally the 44.1 kHz files will not contain the same information, therefore the recordings were done as such to make a comparison to show that there is a difference with the information that is left out and that this information is valuable to us. Unfortunately, the tests did not provide consistent evidence regarding this theory, yet did not leave us completely stranded either. The tests merely granted direction to pursue more testing in this

field to help define more clearly the purpose and benefit of preserving the higher frequencies that may be found in these files.

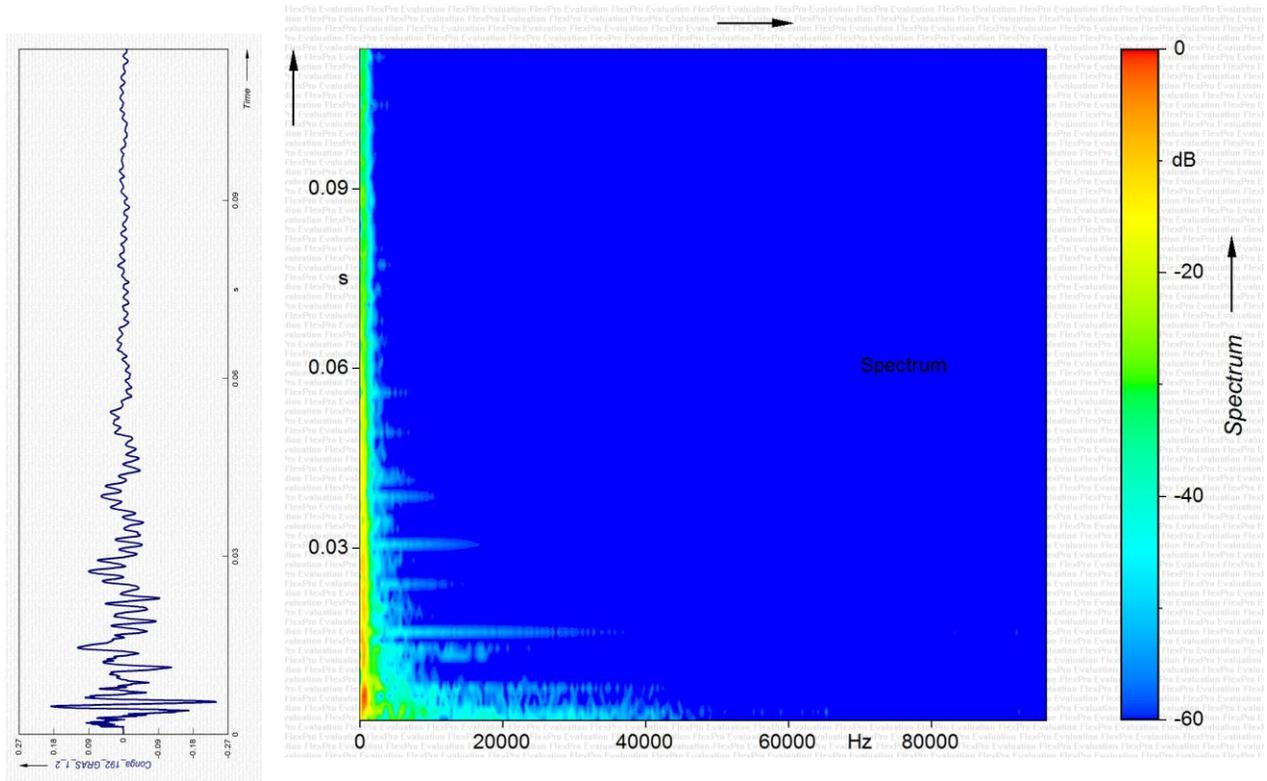


Figure 13: Conga 192 kHz GRAS ½ " Max dB: 60

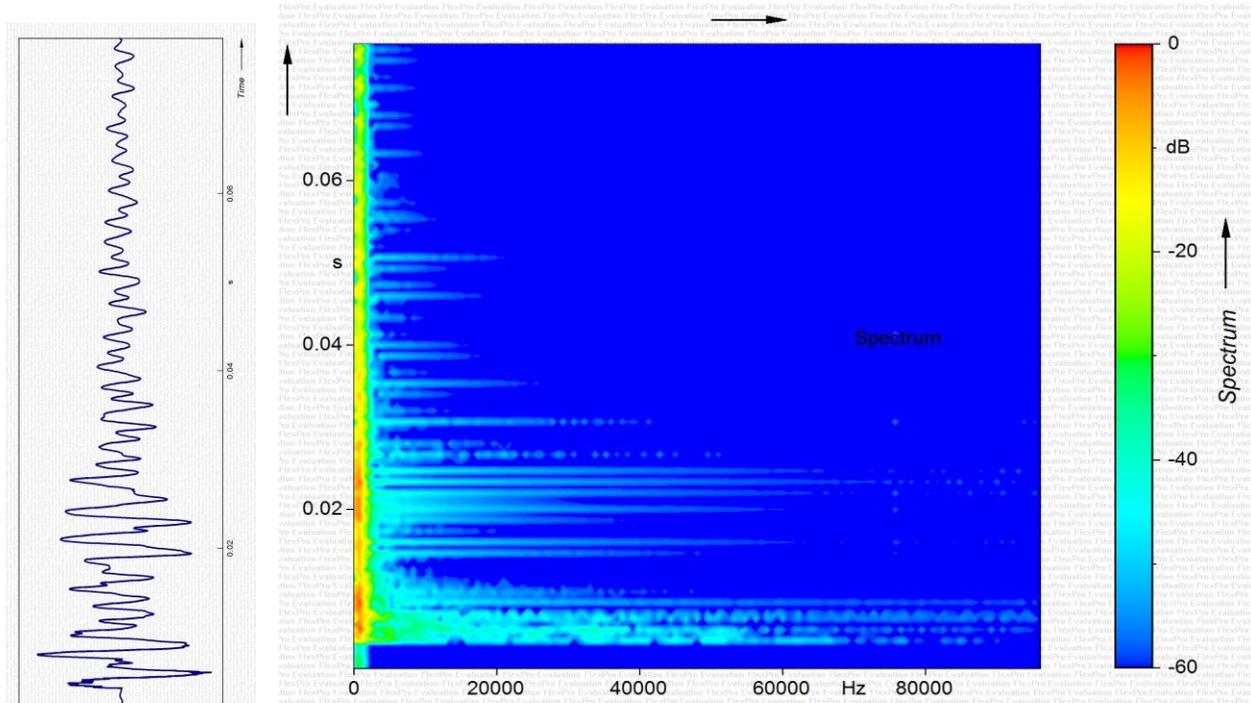


Figure 14: Conga 192 kHz GRAS 1/4 "Max dB: 60

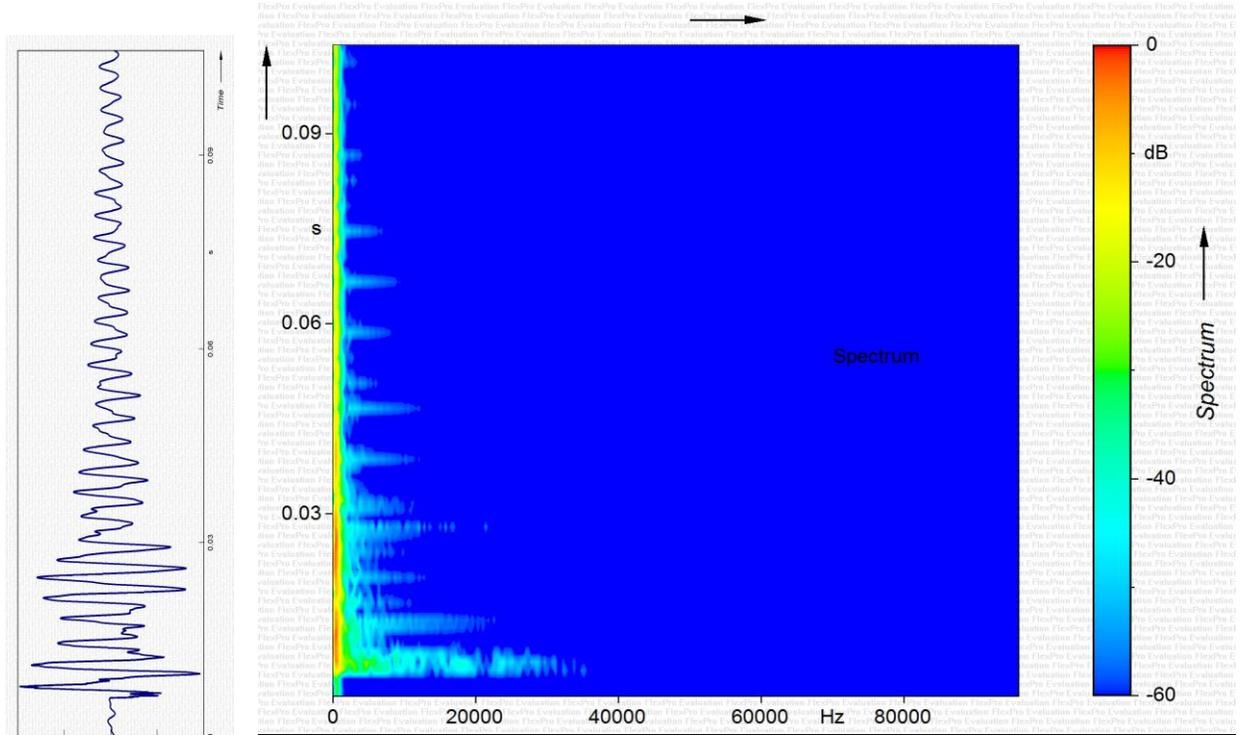


Figure 15: Conga 192 kHz SHURE Max dB: 60

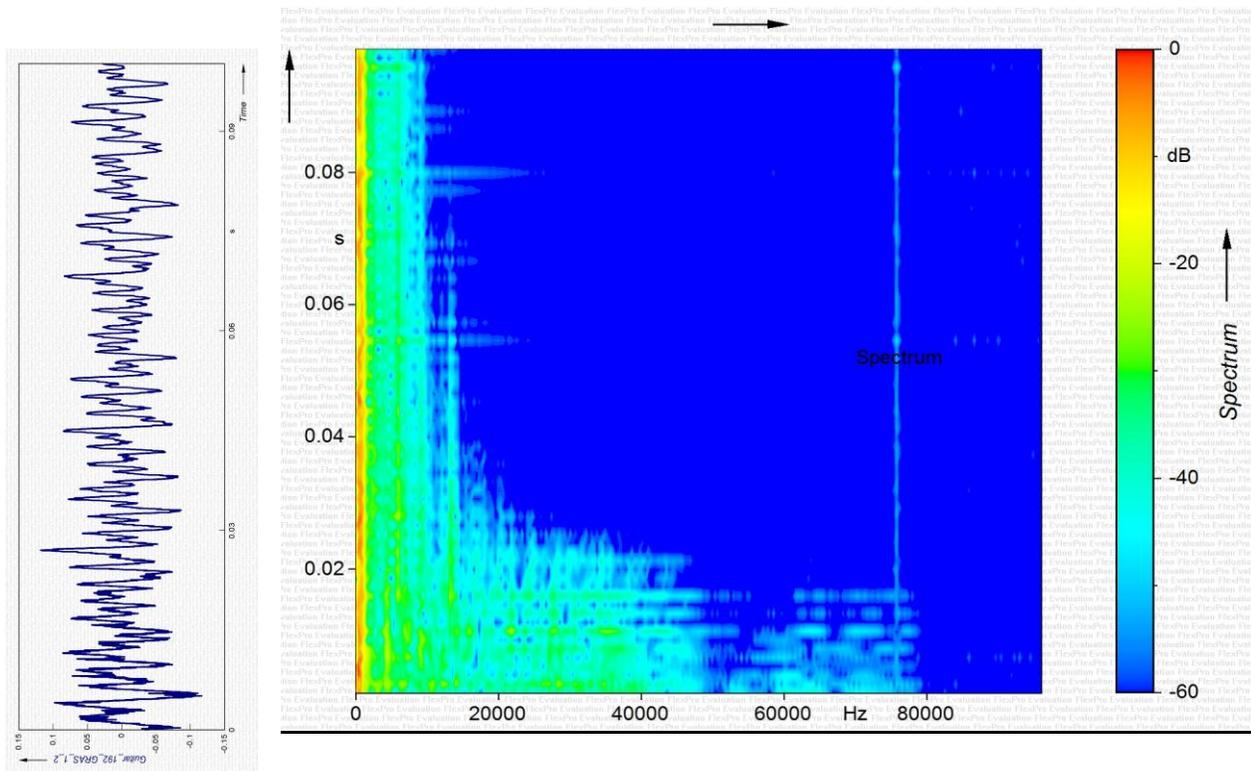


Figure 16: Guitar 192 kHz GRAS 1/2 "Max dB: 60

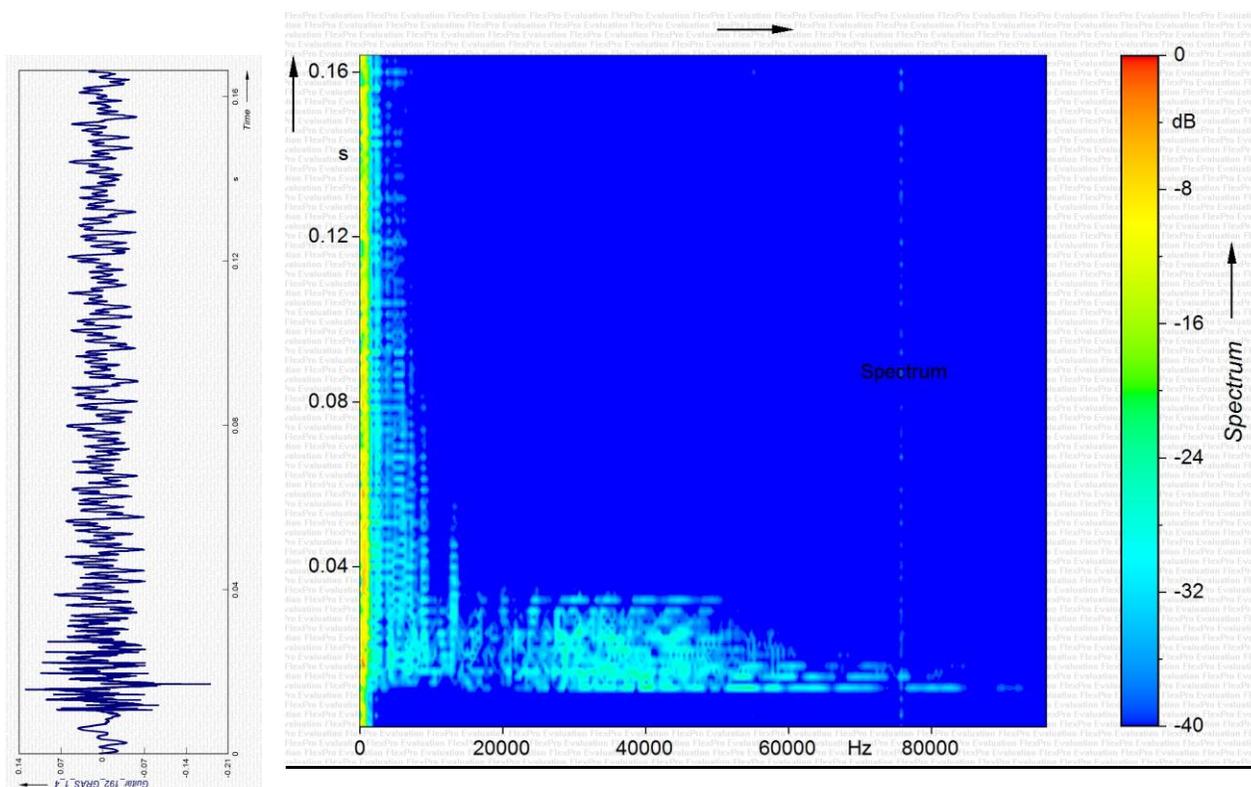


Figure 17: Guitar 192 kHz GRAS 1/4 "Max dB: 60

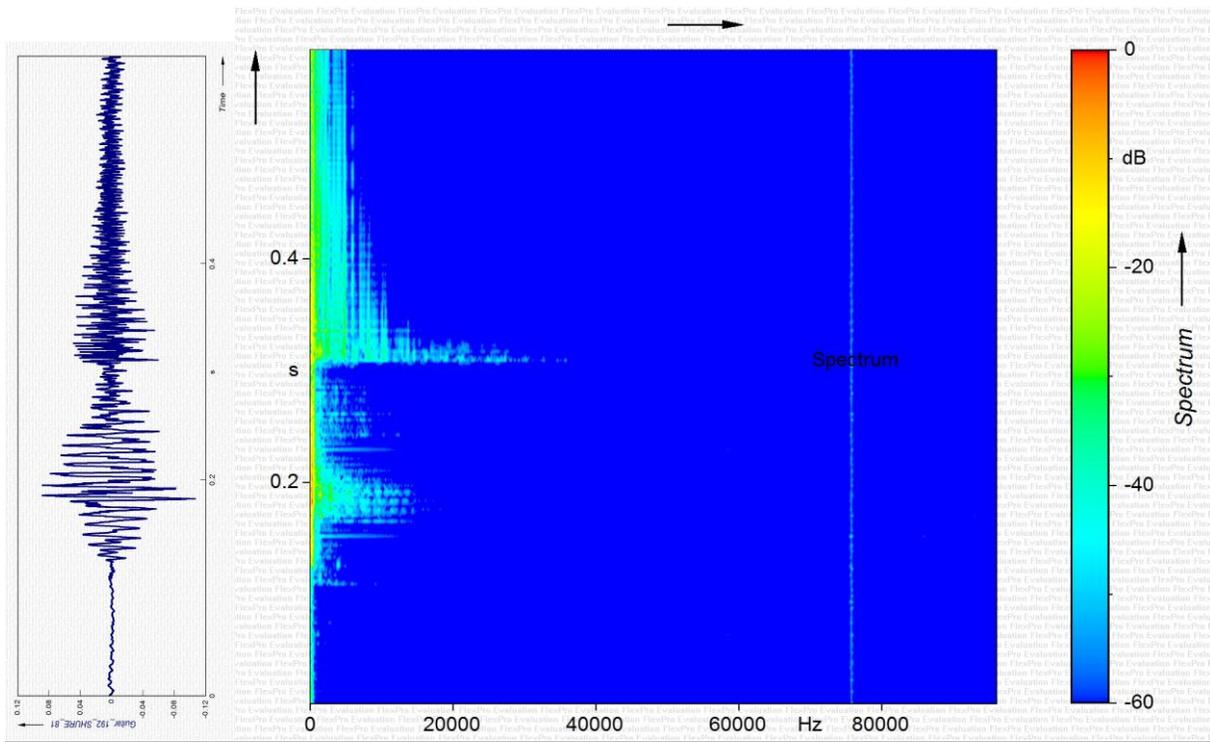


Figure 18: Guitar 192 kHz SHURE Max dB: 60

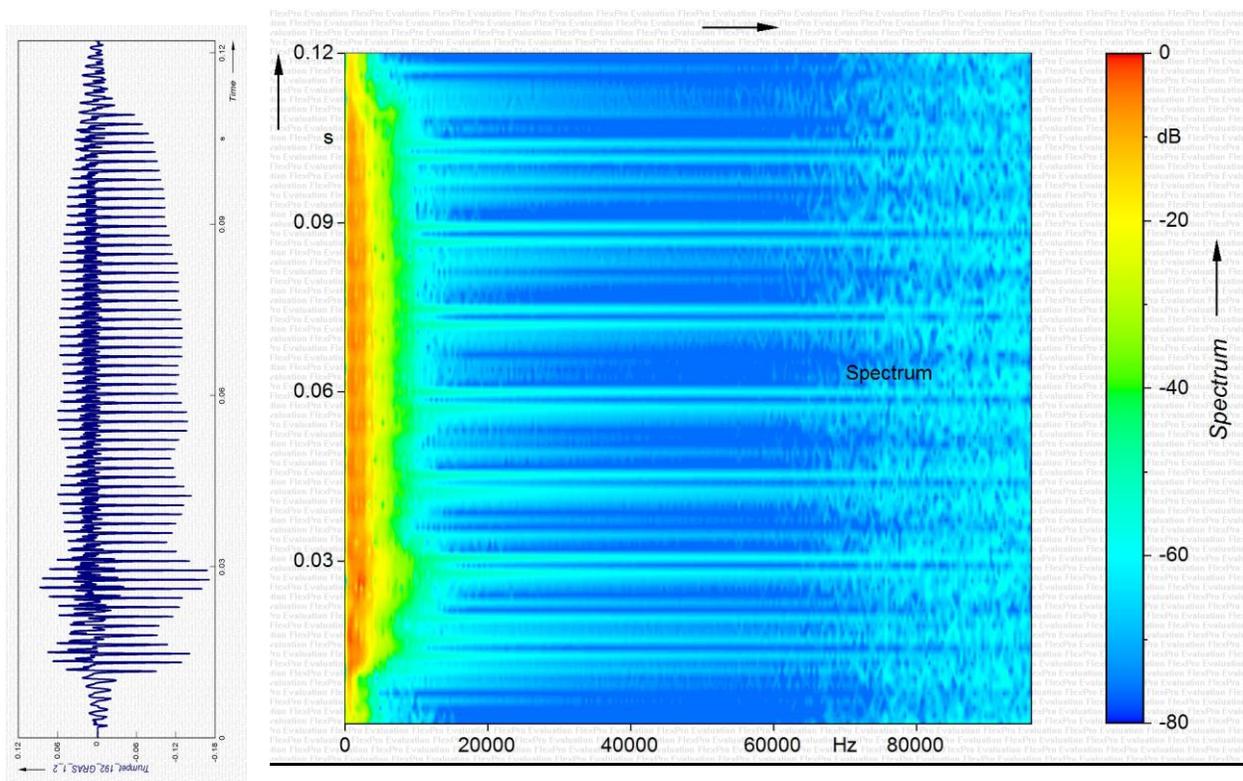


Figure 19: Trumpet 192 kHz GRAS 1/2 Max dB: 60

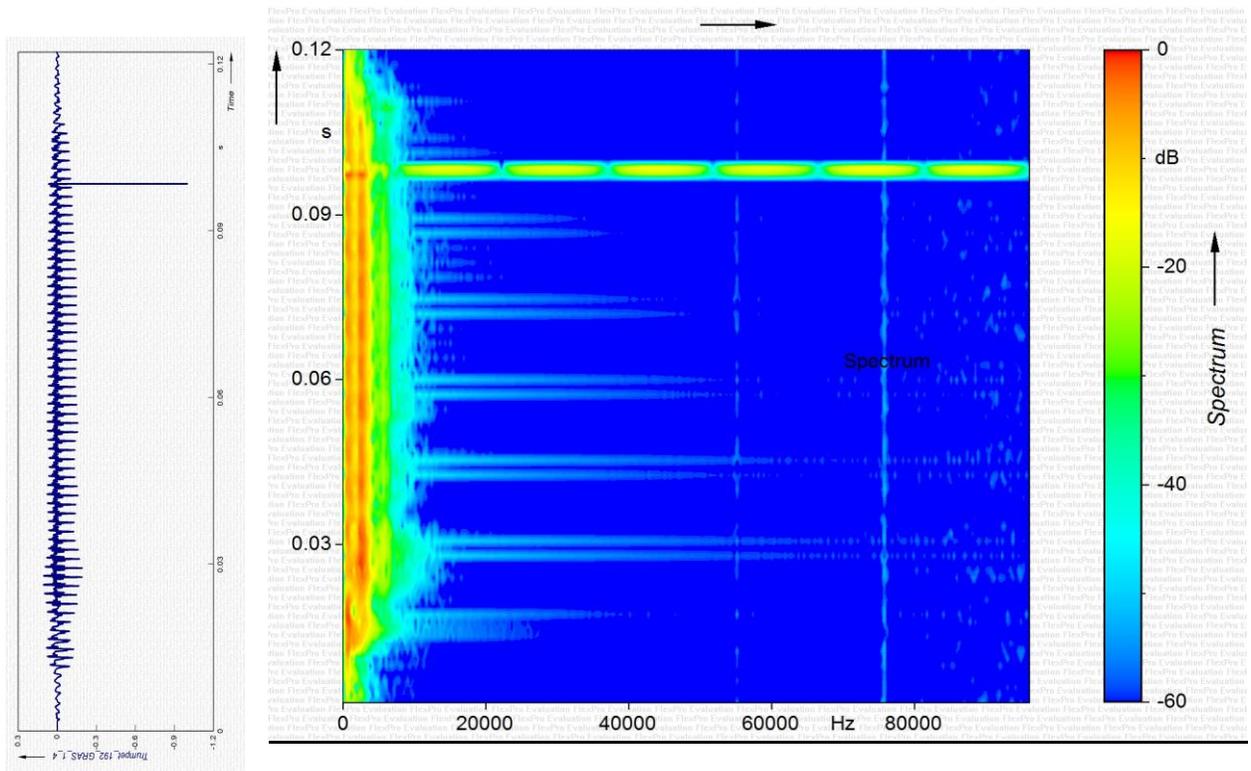


Figure 20: Trumpet 192 kHz GRAS ¼ "Max dB: 60

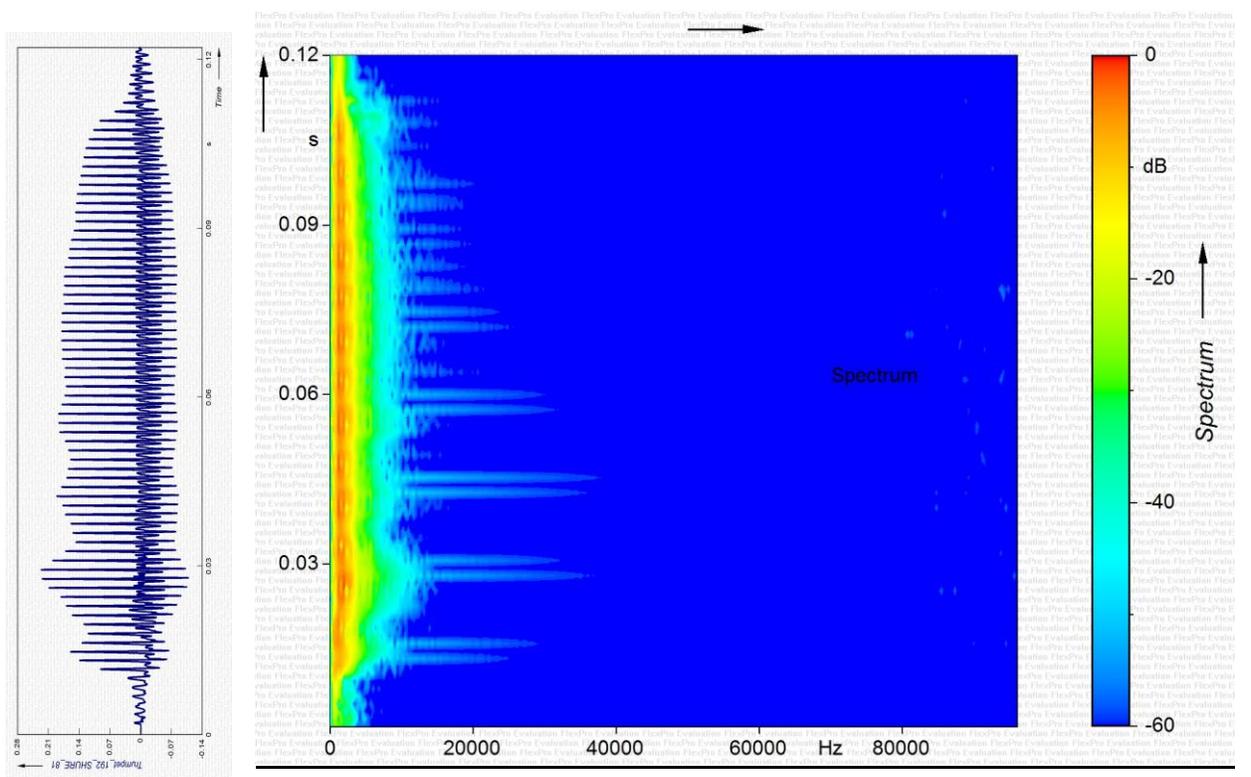


Figure 21: Trumpet 192 kHz SHURE Max dB: 60

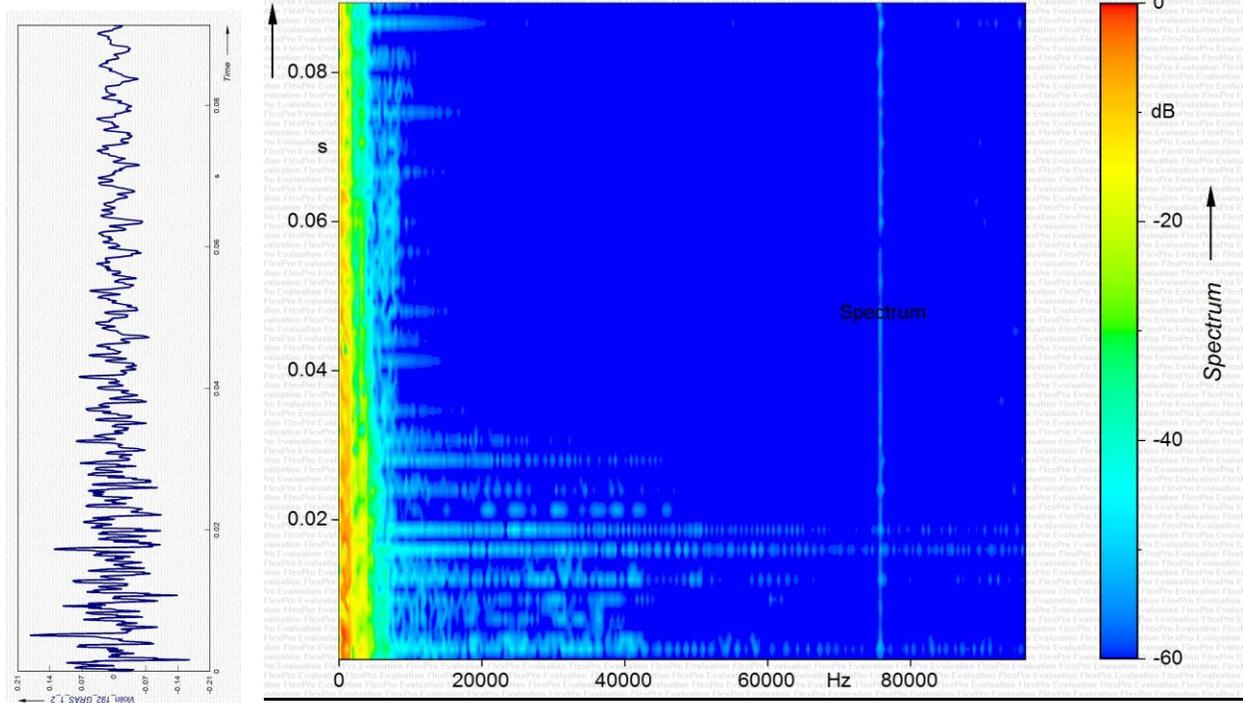


Figure 22: Violin 192 kHz GRAS 1/2 "Max dB: 60

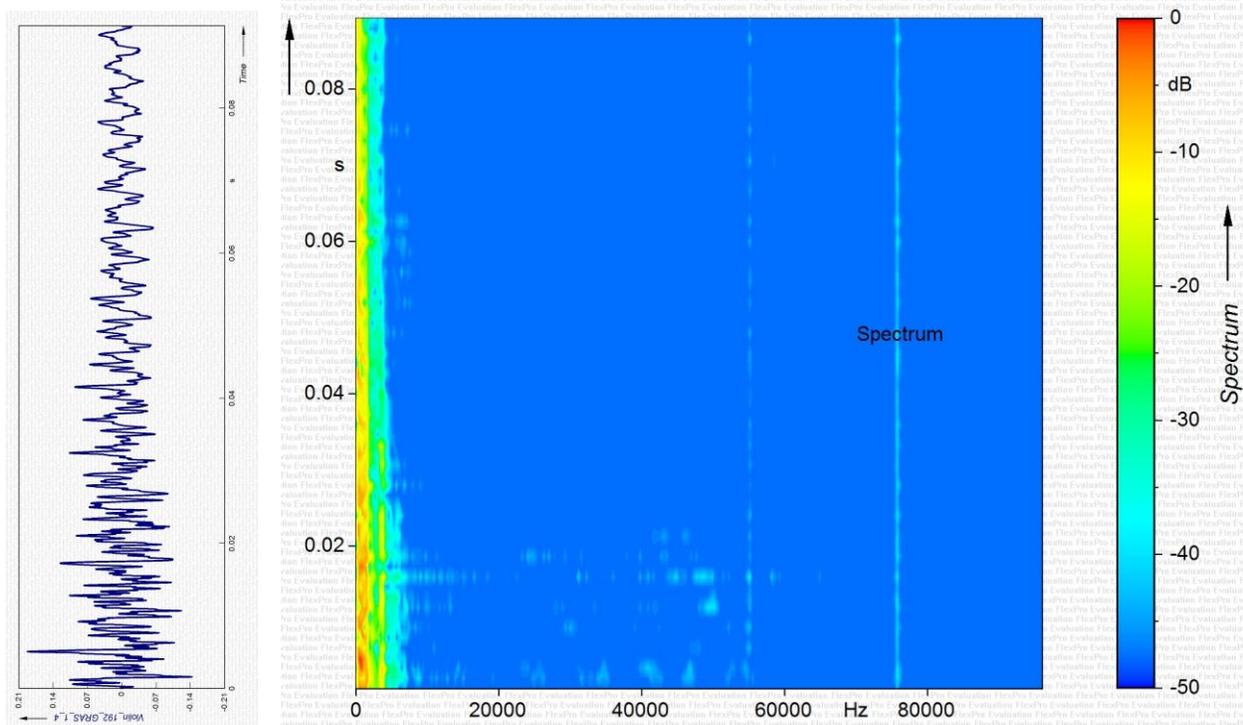


Figure 23: Violin 192 kHz GRAS 1/4 "Max dB: 47

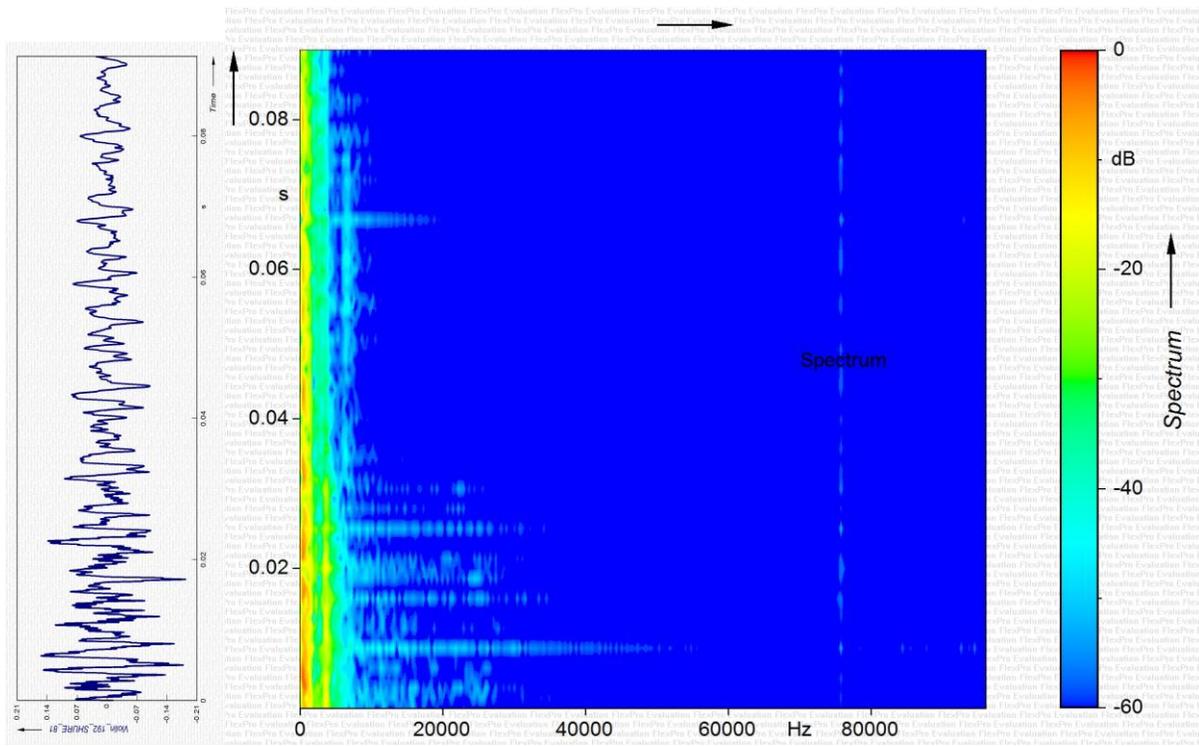


Figure 24: Violin 192 kHz SHURE Max dB: 60

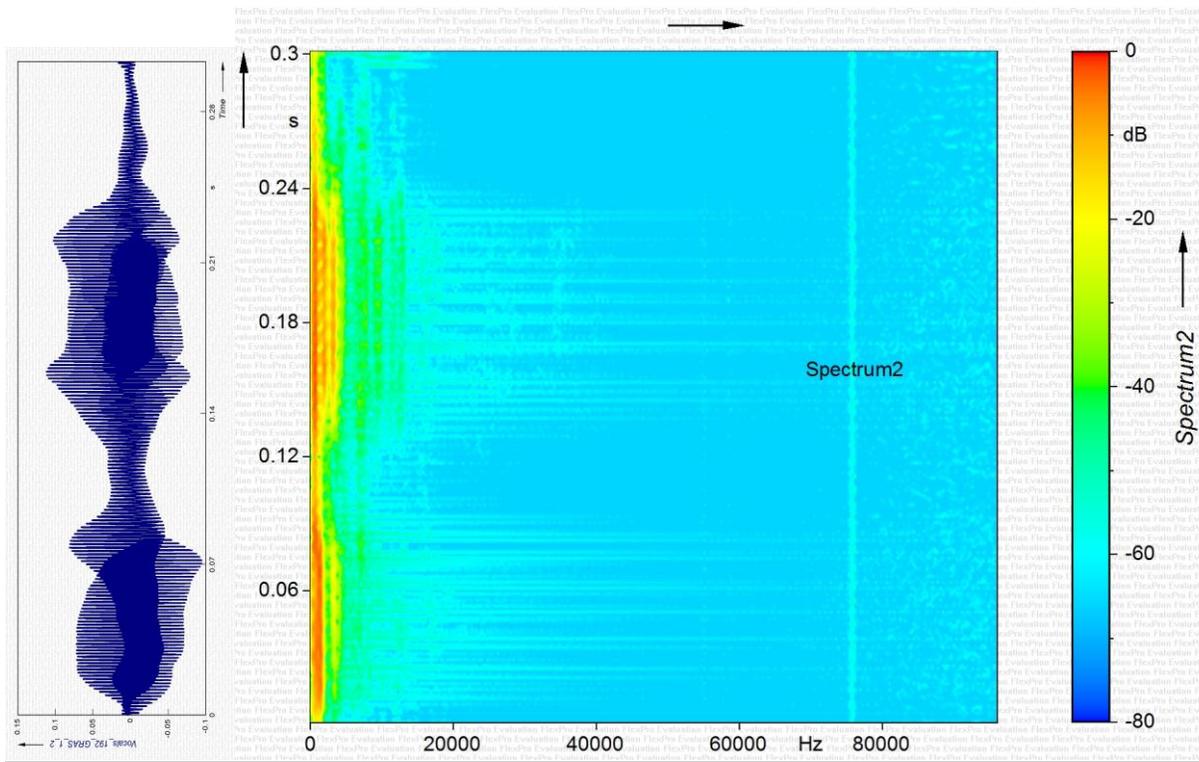


Figure 25: Vocals 192 kHz GRAS 1/2 Max dB: 65

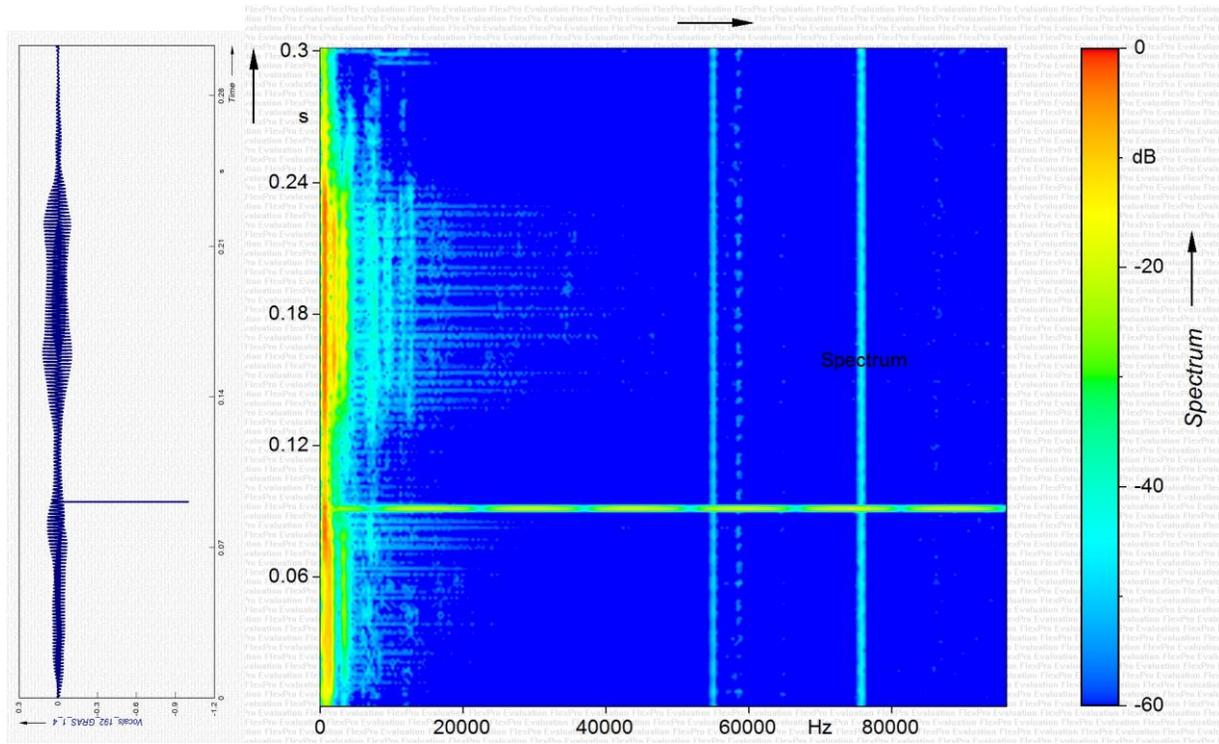


Figure 26: Vocals 192 kHz GRAS ¼ "Max dB: 60

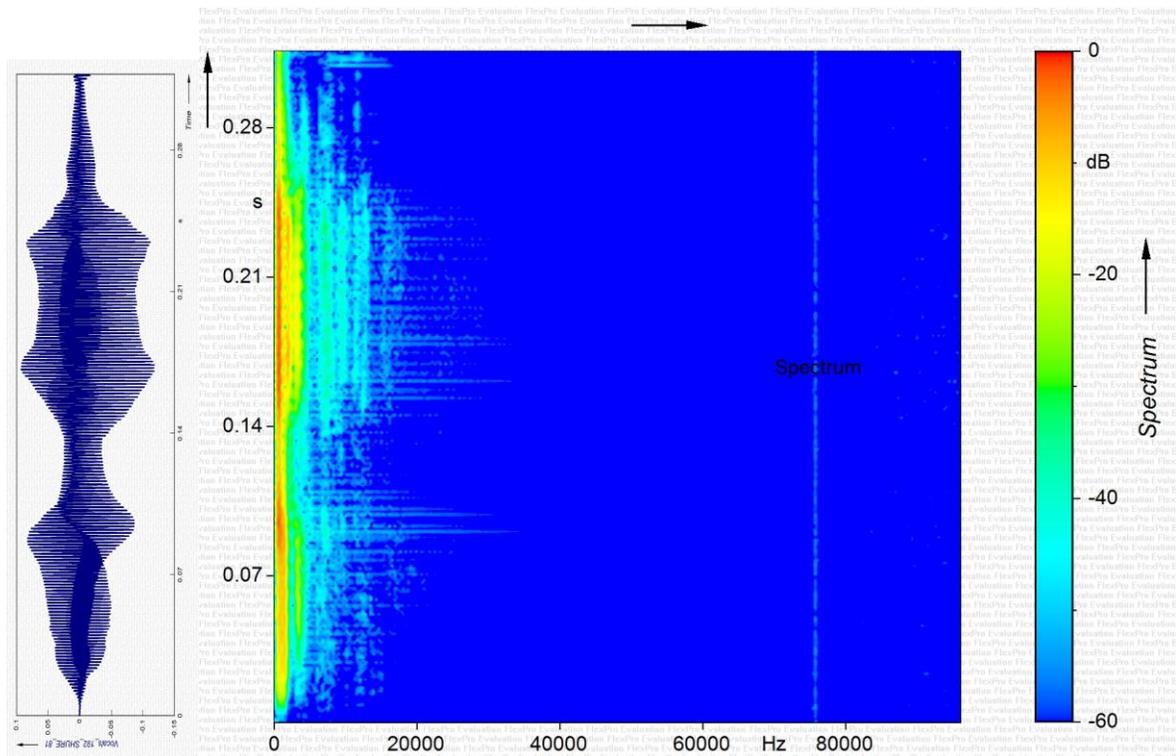


Figure 27: Vocals 192 kHz SHURE Max dB: 60

Analysis and Discussion

Without listening tests it can only be shown that the different captures do in fact contain different information, with some being more faithful reproductions than others. Also, a tentative assumption is that the 192 kHz GRAS ¼” microphone is our best reference. Is there a perceivable sound difference between captures, ever? As stated before, each of the instruments above have three different images that are related to the three different microphones that recorded them. For each of the three 192 kHz samples above, the exact same part of the wave file was taken. As seen above, there are differences in the types of microphones used.

Here below is a comparison of the wave files recorded by the different microphones at different sampling rates.

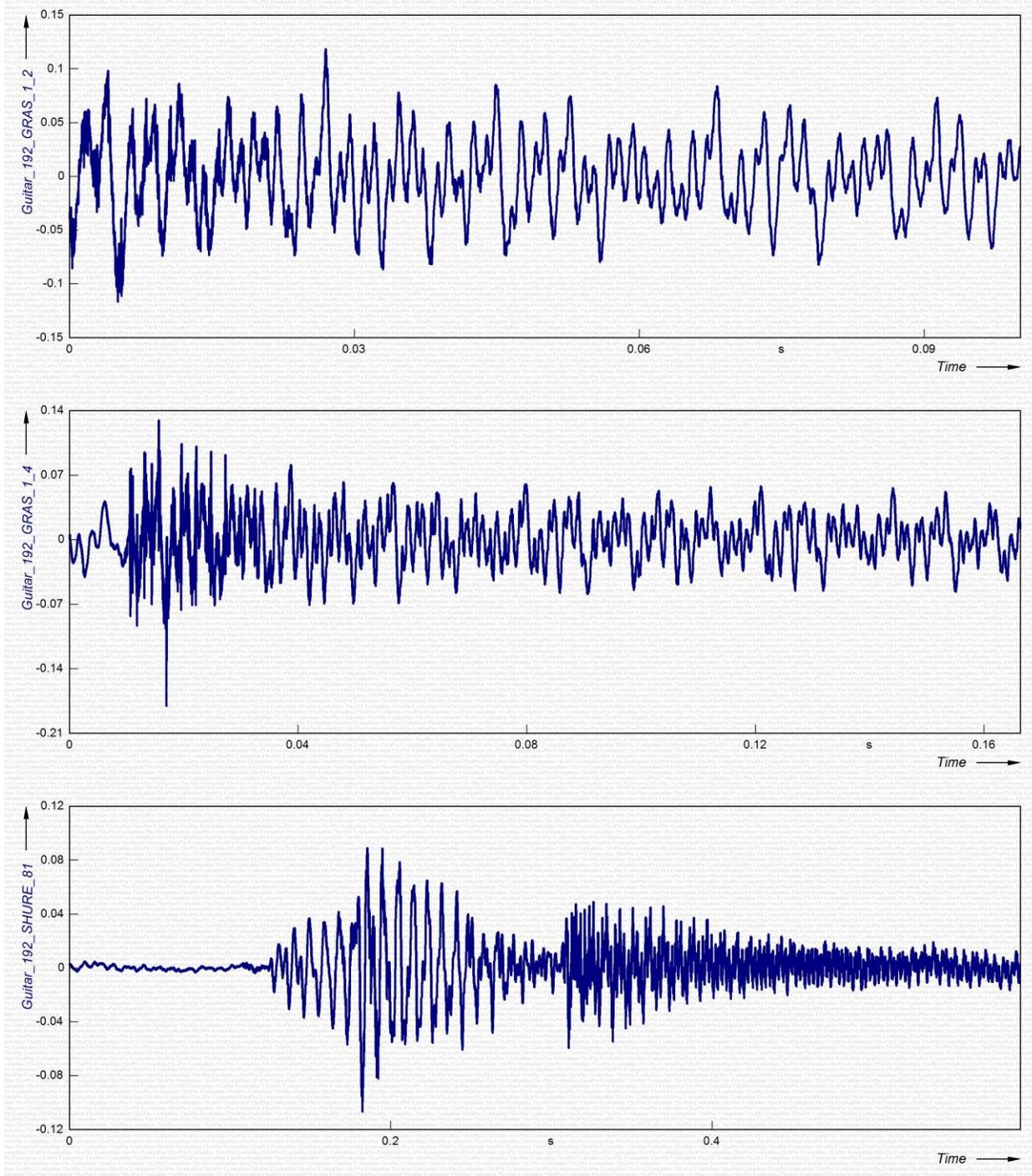


Figure 28: Guitar 192 kHz ½" (top) ¼" (mid) SHURE(bot) microphones

Conclusion

The results above put to rest the question of the existence of data over 20 kHz when recorded at a sampling rate of 192 kHz. From the data acquired, it is safe to claim that this experiment will stand as a foundation for future experiments further investigating the effects of the data that exists above 20 kHz.

Bibliography

- T. Oohashi, E. Nishina, M. Honda, Y. Yonekura, Y. Fuwamoto, N. Kawai, T. Maekawa, S. Nakamura, H. Fukuyama, and H. Shibasaki. Inaudible high-frequency sounds affect brain activity: Hypersonic effect. Journal of Neurophysiology, 83(6):3548–3558, 2000.
- K. Hamasaki, T. Nishiguchi, K. Ono, and A. Ando. Perceptual Discrimination of Very High Frequency Components in Musical Sound Recorded with a Newly Developed Wide Frequency Range Microphone. 117th AES Convention, 2004.
- K. Hamasaki and T. Nishiguchi. Differences of Hearing Impressions Among Several High Sampling Digital Recording Formats. 118th AES Convention, 2005.
- Lenhardt et al., Human ultrasonic speech perception. 1991
- Story, Mike. A Suggested Explanation For (Some Of) The Audible Differences Between High Sample Rate And Conventional Sample Rate Audio Material. 1997.
- ANSI S1.1-1994. American National Standard Acoustical Terminology. 1994.
- Elert, Glenn. Frequency Range of Human Hearing. The Physics Factbook, 2003.
<http://hypertextbook.com/facts/2003/ChrisDAmbrose.shtml>
- Jerry Bruck, Al Grundy, and Irv Joel. An Audio Timeline. The Committee for the Fiftieth Anniversary of the AES. V--1999-10-17. 1997.
- Moulton, David. The Audio Window. Moulton Laboratories, December, 2000.\
- S.Carlile, Auditory Space, http://cns.bu.edu/~shinn/Carlile%20book/chapter_1.pdf
- Wikipedia.org.
<http://www.wikipedia.org>

Boyk, James. There's Life Above 20 Kilohertz! A Survey of Musical Instrument Spectra to 102.4 KHz.
<http://www.cco.caltech.edu/~boyk/spectra/spectra.htm>

Klass, Gary. A note on Sampling error
<http://lilt.ilstu.edu/gmklass/pos138/datadisplay/sampling.htm>

MPEG Audio Coding,
http://www.digitalradiotech.co.uk/mpeg_coding.htm

William Forde Thompson (2008). Music, Thought, and Feeling: Understanding the Psychology of Music. p. 46

John R. Pierce (2001). "Consonance and Scales". In Perry R. Cook. Music, Cognition, and Computerized Sound. MIT Press

Hans Weblog Wordpress
<http://hans27026.wordpress.com/2008/11/23/sound/>