

MDFourier

Artemio Urbina

June 27, 2019

Abstract

MDFourier is an *open source* software solution created to compare *audio signatures* and generate a series of graphs that show how they differ. The software consists of two separate programs, one that generates the signal for recording from the console and the other that analyses and displays the audio comparisons.

The information gathered from the comparison results can be used in a variety of ways: to identify how *audio signatures* vary between *systems*, to detect if the *audio signals* are modified by *audio equipment*, detect if *modifications* resulted in audible changes, to help tune *emulators*, *FPGA* implementations or *mods*, etc.

This document serves as an introduction, a manual and a description of the methods used and how they work. Its intent is to guide new users in making simple comparisons while still providing advanced users with information to perform a more detailed analysis of audio signatures.

Contents

1	Introduction	6
1.1	Possible applications	7
1.2	Disclaimer	8
2	Audio Basics	9
2.1	Frequency	9
2.2	Amplitude	10
2.3	Digital representations of sound	11
2.3.1	Sample rate	11
2.3.2	Bit depth	12
2.3.3	Decibels in the digital domain	12
3	How <i>MDFourier</i> works	13
3.1	Workflow	14
3.2	File alignment	14
3.3	The heart of the process	15
3.4	Minimal significant amplitude	16
4	How to interpret the graphs	17
4.1	Scenario 1: Comparing the same file against itself	17
4.2	Scenario 2: Comparing two different recordings from the same console	19

4.3	Scenario 3: Comparing against a modified file	21
4.4	Scenario 4: Comparing against digital low pass and high pass filters	22
4.5	Scenario 5: Comparing two recordings from the same console made with different Audio Cards	27
4.6	Scenario 6: Comparing <i>WAV</i> vs <i>MP3</i>	29
5	Results from vintage retail hardware	32
5.1	Sega Genesis Model 1 VA3 (US) vs Mega Drive Model 1 VA1 (JP)	33
5.2	Sega Genesis Model 1 VA3 (US): Headphone out vs AV out	34
5.3	Sega Genesis Model 1 VA3 (US) vs Sega Genesis Model 1 VA6 (JP)	35
5.4	Sega Genesis Model 1 VA6 (JP) vs Mega Drive Model 1 VA6 (US)	35
5.5	Sega Genesis Model 1 VA6 (US) vs Sega Genesis Model 1 VA6 (US)	36
5.6	Sega Genesis Model 1 VA3 (US) vs Sega Genesis Model 2 VA1.8 (US)	37
5.7	Sega Genesis Model 1 VA3 (US) vs Sega Nomad (US)	38
5.8	Sega Genesis Model 2 VA1.8 (US) vs Sega Nomad (US)	40
5.9	Sega Genesis Model 1 VA3 (US) vs Sega CDX (US)	41
5.10	Sega CDX (US) Line-out vs Sega CDX (US) Headphone port	42
6	Future developments	44
7	Conclusions	45
Appendices		46
A	Downloads	47

B How to use the Front End	48
B.1 Front End Options	50
B.1.1 Window Functions	50
B.1.2 Color Filter Functions	50
B.1.3 Align Transform to 1Hz	51
B.1.4 Output Graphs	52
B.1.5 Extra Command	53
B.1.6 Verbose Log	54
C Configuration file	55
D MDWave	57
E Normalization and amplitude matching	59
E.1 Frequency domain normalization	59
E.2 Time domain normalization	60
E.3 Highest fundamental average normalization	61
F Frame rates and their importance	62
G Stereo Audio Balancing	65
G.1 Possible causes	65
H Differences due to temperature	67
I Corner cases	69
J Known Issues	72
K Window Function equations and graphs	73
K.1 Tukey	73
K.2 Hann	74

K.3	Flattop	75
K.4	Hamming	75
K.5	No Window	76
L	Color Filter Function details	77
L.1	None	77
L.2	\sqrt{dBFS}	78
L.3	$\beta(3, 3)$	79
L.4	Linear	80
L.5	$dBFS^2$	81
L.6	$\beta(16, 2)$	82
M	Requirements	84
M.1	Audio capture device	84
M.1.1	Sampling clocks	85
M.2	Computer	85
M.3	Game Consoles or emulators	85
M.4	Flash cart, or means to run the binary	85
M.5	Cables and adapters	86
M.6	Audio capture software	86
N	Compiling from source code	87
N.1	Dependencies	87
O	Colors available for graphs	89
P	Licensing	90
Q	Contact the author	91
R	Acknowledgments	92

Chapter 1

Introduction

MDFourier is a free and open source software suite for analyzing and comparing *audio signatures*.

It can be used to identify the differences between two *audio recordings*. For instance, a *Sega Genesis Model 1 VA3* and a *Sega Genesis Model 2 VA 1.8* can be compared in order to verify how different they are across the human hearing frequency spectrum, using objective and repeatable results.

The process is not limited to one specific system, and the software can be used to compare any other platform by adding new configuration files.¹ At the moment, profiles for *Mega Drive/Genesis* and *Mega/Sega CD* are functional—with more to follow.

The intention is not to disparage any particular system, but to help understand and improve them whenever possible by identifying their differences.

A secondary but not less important goal is to create a community driven catalog of audio signatures from these systems, in order to help preservation efforts achieve their objectives.

I believe it would be very useful to have configurations tuned to replicate the audio signatures of vintage retail hardware variations in emulators and *Field-programmable gate array (FPGA)* implementations. But that doesn't mean there isn't room for improved interpretations; reducing noise while keeping the reference sound signature is one such scenario.

Although powerful industrial solutions for this kind of analysis must exist, they do not seem to have reached the enthusiast community. I'm not aware of any other similar effort to create public software analysis tools that are aimed at vintage console hardware. It must be noted that there have been detailed

¹The term *system* will be used to cover vintage retail hardware, emulators, *FPGA* implementations and any other possible variant that executes binaries for the target hardware.

comparisons made in particular for the *MegaDrive/Sega Genesis*.².

1.1 Possible applications

- Figure out if *vintage hardware* retail variations really have different audio signatures when compared against each other.
- Help emulator and *FPGA* implementation authors with tools to tune *filter profiles* that match *vintage retail* hardware audio signatures.
- Help the modding community to compare how audio signatures change for each *synthesizer* within the targeted system while developing audio modifications.
- Determine if there were changes in the audio *spectrum* after modifications to a console, like *recapping* or changing the audio circuitry.
- Evaluate equipment, such as *switchers* and *upscalers* with audio *passthrough*. It can help figure out if they have any effect on the signal, by comparing a recording with and without the device connected in the *Audio/Video* chain.
- Help to recreate specific audio signatures.
- Help to tune a system to a particular taste.

²See the forum post at [2]

1.2 Disclaimer

MDFourier and this documentation are work in progress. Although I tried to adhere to the best practices known to me, my experience in *digital signal processing* was almost non-existent before this project.

Please contact me if you have corrections, improvements, suggestions or comments. These are encouraged and welcome. Contact information is available in appendix Q.

This project was born out of my curiosity to compare the audio signatures of different revisions of the *Mega Drive/Genesis*, and verify them with a tool assisted analysis. Due to my interest in *game preservation* it naturally grew into its current form, and will hopefully continue to change.

Chapter 2

Audio Basics

Sound is a series of *vibrations* that travel through a medium. In our case this medium is typically air, but it can be water or even a solid. These vibrations are *waves of pressure* and—as any other *wave*—they have characteristics that we can measure in order to classify and understand them.

2.1 Frequency

We use the term *frequency* to refer to a repeating event—how *frequent* it is. In this case it means how many times the *wave* we are measuring repeats in a second. We measure this in *cycles per second*, and use a unit named *Hertz* (*Hz*).

As you can imagine these vibrations in the air can have a wide range of *frequencies*, but we can't perceive them all due to biological limitations. The range that humans can typically perceive is called the *human hearing spectrum*. A *spectrum* refers to a continuous range of frequencies, and in this case it is defined to start at *20Hz* and end at *20000Hz*—we commonly use the term *20kHz*, since *kHz* refers to multiples of *1000Hz*.¹ These numbers are of course an average, and every one of us has a different *hearing spectrum* that also changes as we age.

We perceive *frequency* variations as *pitch* differences: lower frequencies are what we call *bass* and higher frequencies are *treble*, corresponding to lower and higher *pitch* tones.

¹This is the *decimal-based multiplier* system defined by the *International System of Units* (SI).

2.2 Amplitude

The other *wave* characteristic we use to measure sound is called *amplitude*. This is how *powerful* the *wave* is when compared to the rest of the signal that is being analyzed.

We perceive this *amplitude* as *volume*, and the terms are directly related. The higher the *amplitude* of a *sound pressure wave*, the higher the *volume* we perceive. Since the software can only use *amplitude* for its calculations, we'll use that term throughout the whole document.

The unit used to measure *amplitude* in *sound pressure waves* is called *decibel* (*dB*). These use a *logarithmic scale*, which means that different values in the scale correspond to *orders of magnitude* instead of *linear quantities*, each value is the result of multiplying by a constant.

We use this scale since it closely models the way we perceive sound, and also because *sound amplitudes* have a very wide range. For instance, the use of a logarithmic scale allows us to compare the variations of the ambient sound from a library against those from a rock concert. You might be familiar with this concept from other *logarithmic scales*, like the *Richter scale* used to measure earthquake strength.

This scale starts at *0dB* which is the inferior limit of human hearing, and goes up to *194dB*; where the sound barrier is broken and it becomes a *shock wave*. The following table has some common values:²

<i>dBA</i>	Source
0	healthy hearing threshold
10	a pin dropping
20	rustling leaves
30	whisper
40	babbling brook
50	refrigerator
60	conversational speech
70	vacuum cleaner at 1 meter
80	alarm clock
85	passing diesel truck
90	squeeze toy
95	inside subway car
100	handheld drill
110	rock band
115	emergency vehicle siren
120	thunderclap
125	balloon popping
130	peak stadium crowd noise
140	jet engine at takeoff
145	firecracker
150	fighter jet launch
180	rocket launch
194	sound waves become shock waves

²Table values taken from [33]—you might notice they are listed as *dBA*, this is tuned to our combined *frequency* and *amplitude* response curve.

The use might be confusing at first, so we'll use an example in order to show how the values behave in this scale. Imagine we have two sources that are individually measured as having a *sound pressure level* of $9dB$. If both sources produce sound at the same time, we will measure a combined *sound pressure level* of $12dB$ instead of the $18dB$ that one might have expected when using a *linear scale*.

This doesn't mean you'll need to use *logarithm algebra* when dealing with this document—or with sound in general. In practical terms whenever you see a $3dB$ difference in sound *amplitude*, it means that you'll perceive *twice* the *volume*.

2.3 Digital representations of sound

Sound is a continuous signal that varies in time. This means that if you could zoom into it at the smallest fraction of time imaginable, you'd also have a tiny bit of sound happening during that lapse.

Unfortunately computers don't have the unlimited storage that would be able to replicate a continuous signal—and they also have speed limitations. In order to represent sound in the *digital domain* we use a process called *sampling*.

2.3.1 Sample rate

Sampling involves measuring the *amplitude* of the *sound pressure wave* at *regular cycles*, and assigning a value to each one of these measurements. The number of cycles we use to take each of these *samples* from the continuous *audio signal* is called the *sampling rate*.

The *sample rate* defines the *frequencies* we can represent in the digital domain—as the *sampling rate* is increased, so does the *frequency range* that can be portrayed by this digital representation of the signal.

If we want to represent a certain *frequency range* reliably, we must choose a *sample rate* that is *twice* as big as the highest *frequency* value in the desired *spectrum*.

Since we want to cover the whole *human hearing range*—which goes from $20Hz$ to $20kHz$ —we must choose a *sample rate* of at least $40kHz$. This is known as the *Nyquist frequency*, and is the reason why some of the most common *sample rates* in digital audio are $44.1kHz$ and $48kHz$ —which are the ones supported by the *Analysis Software* in *MDFourier*.³

³See [19] for a detailed video demonstration of how *sampling rates* work.

2.3.2 Bit depth

When we create a *digital representation* of the *amplitude* value for each *sample*, we must define a *bit depth*. This term refers to the amount of *bits*—the minimal piece of data in our current computing systems—that will be used to store each number.

2.3.3 Decibels in the digital domain

When we were measuring amplitudes before, we used the *dB*. In that scenario we were analyzing amplitudes that start at the minimum amplitude our human ears can detect and assigned a value of *0dB* to that level. Values can theoretically go up without limits—aside from that pesky sound barrier we mentioned before when using air as a transmission medium.

However when dealing with digital values that have a defined *bit depth*, our limit is fixed from the other side—the biggest amplitude we can represent with a limited number of *bits*. This number is easily calculated, for n bits we get 2^n possible values.

Because of this, when using digital files we use a variation of the *decibel* named *decibels relative to full scale (dBFS)*. The highest possible amplitude starts at *0dBFS* and values go down into negative numbers, until they reach the lowest amplitude that can be represented by the selected *bit depth*. When using a *bit depth* of *16*—as used in *Compact Discs* and by the *Analysis tool* in *MDFourier*—the minimum *amplitude* that can be represented is *-96dBFS*.

This is commonly referred to as the *dynamic range*, and it tells us how far apart the minimum and maximum *amplitudes* that can be represented in a *audio recording* are. For example, if we have a dynamic range of *96dB*, we can represent the relative amplitudes of *thunderclap* and *rustling leaves* within the same recording; but we could have no lower or higher amplitudes.

Chapter 3

How *MDFourier* works

The primary thing to keep in mind is what *MDFourier* does. It takes two signals, the first one is designated as the *Reference* file. The second one is the file under scrutiny, and is referred to as *Comparison* file.

The *Reference* file is used as a control. This means that its characteristics are considered the true values to be expected and against which the *Comparison* file will be evaluated. In consequence, all results are relative between the signals.

Having these files, the software analyzes and compares their audio characteristics and generates graphs that visually show how they differ.

These files are audio recordings from the desired hardware, preferably captured with a *flat frequency* audio *capture card* and generated by a *custom binary* that runs on the target platform.¹ This *custom binary* is specifically crafted for the particular hardware capabilities and frequency range of the target. Whenever possible, this binary will be part of the *240p Test Suite*.²

Although the analysis software is *command line* based—in order to be multi-platform and offer the tool on every operating system that has an *ANSI C99 compiler*—a *Graphical User Interface (GUI)* front end for *Microsoft Windows* is provided for simplicity and accessibility.³ Full Source code can be downloaded from *github*.⁴

¹ Flat frequency response refers to the capability of capturing the whole audible spectrum with little or no variation, regardless of frequency. A non flat frequency card can be used to gather relative information, but the captures won't match the ongoing catalog in order to make further comparisons. See *Audio Cards* in appendix M for more information.

² A homebrew software suite for video game consoles developed to help in the evaluation of TVs, upscalers, upscan converters, line doublers and video processing in general.[6]

³ See appendix B for a guide on how to use the software.

⁴ See download link [5].

3.1 Workflow

The first step is loading the custom binary to your console. This varies for each platform, from placing the binary in a *flash cart*, burning a *CD-ROM* or using a *custom loader*.

The next step is setting up your *audio capture card* and computer, in order to record a $44.1/48\text{kHz}$ 16 bit stereo *WAV* file.⁵

Once the *capture card* is ready and the cables are hooked up from the console to the *capture card*, start recording on the computer and execute the *MDFourier* test from the console.

Wait for the console to show a message indicating you can stop recording, this typically takes less than 1 minute. After you have at least two files: a *reference*—which can be one of the provided files—and a *comparison* file, you are ready to go.⁶

3.2 File alignment

MDFourier takes both files and auto detects the starting and ending point of the recording. These are identified by a series of 8820Hz pulses in the current *Mega Drive/Genesis* implementation. From these, a *frame rate*⁷ is calculated in order to trim each file into the segments that are defined in the *configuration file*⁸ for further comparison.

Most importantly, it guarantees that the *Reference* and *Comparison* files are logically aligned, and that each note—or segment—is compared to its corresponding one, with no overlap and without any audio editing or trimming skills required from the user. Current pulse detection accuracy is around $1/4$ of a millisecond.

After alignment is accomplished, the software reports the starting and end points of both signals, in seconds and bytes.

⁵At the moment only 16 bit *WAV* and *FLAC* files with *PCM* encoding and at 44.1 or 48kHz can be used. *MP3* is not supported since its compression method alters the results in this type of analysis, as shown in section 4.6.

⁶Available at the web page of the project appendix A

⁷The *frame rate* is the ratio at which the console sends video frames to a display. For more details see appendix F

⁸This file specifies the operating parameters for each hardware configuration. It is described in appendix C

3.3 The heart of the process

A process called *Discrete Fourier Transform*⁹ is used in order to analyze and compare the *amplitudes* from each of the *frequencies* and *harmonics*—the *spectrum*—that compose the audio signal. The software uses the *FFTW*¹⁰ library in order to accomplish this.

In order to compare the frequencies and amplitudes of each file against each other, a reference point must be set. This is achieved via a relative *normalization*¹¹ based on the *maximum amplitude* of the *Reference* file. A *local* maximum search is done at the same frame in the *Comparison* signal, and that amplitude is then used to normalize both files.

This is done in the *frequency domain*¹², in order to reduce amplitude imprecision when comparing recordings with different *frame rates*¹³. The software does have the option to do this in the *time domain*¹⁴, but quantization and amplitude imprecision is to be expected if used.

After normalization is finished, the frequencies of each *block*¹⁵ are sorted out by amplitude, from highest to lowest. *MDFourier* can be configured to compare a range of these frequencies—by default 2000 of them are compared for each *block* defined in the *mfn* configuration file¹⁶. I've found such number to be more than enough, and usually the *minimum significant amplitude*¹⁷ limits this parameter to an even lower number. In case a different comparison of frequencies is needed, the amount can be increased or decreased via the command line¹⁸.

Sometimes it is helpful to listen to the results of these limiting filters, in order to evaluate if—for instance—2000 frequencies are either enough or too little for the current application. For this and other purposes I made an extra tool named *MDWave*¹⁹, which creates segmented audio files as internally processed by *MDFourier*, even including the effects of *window*²⁰ filters and amplitude limits.

After comparing these frequencies between both files, matches are made and the differences in amplitude are plotted to a graph. Please read chapter 4 for a detailed description of what the various graphs created by the program

⁹ Discrete Fourier Transform (DFT) is the variant of *Fourier Transform* applied to discrete values, such as the ones we have in the audio file. See [1].

¹⁰ *The Fastest Fourier Transform in the West*. [7]

¹¹ This process is detailed in appendix E.

¹² Basically all analysis from this point and forward takes place after the signal has been decomposed into its fundamental and harmonic sine waves by the Fourier Transform.

¹³ For more details see appendix F.

¹⁴ In other words, normalize based on the samples and their values. This is detailed in appendix E.

¹⁵ A *block* is the basic segment of audio to be compared, such a specific note by a synthesizer as generated by the custom binary.

¹⁶ Described in appendix C.

¹⁷ See section 3.4.

¹⁸ See appendix B.1.5.

¹⁹ Described in appendix D.

²⁰ Window functions are described in section B.1.1.

represent.

If you are interested in learning what the *Fourier Transform* does and how its "*magic*" works, there are several resources online. Here are a few of them:

- But what is the Fourier Transform? A visual introduction.
<https://www.youtube.com/watch?v=spUNpyF58BY>
- An Interactive Introduction to Fourier Transforms.
<http://www.jezzamon.com/fourier/>
- The Uncertainty Principle and Waves.
<https://www.youtube.com/watch?v=VwGyqJMPmvE>

3.4 Minimal significant amplitude

Although the whole frequency spectrum can be compared, there is little practical use in doing so—due to execution time and extra noise from low amplitude harmonics. As a result, rule of thumb defaults are set in order to minimize these issues.

One of these values is a *minimum amplitude* at which to stop comparing the fundamental frequencies that are found after decomposing the signal with the *Fourier Transform*. This is the *minimal significant amplitude* and it is the cutoff at which comparisons made by the software are stopped.

Currently *MDFourier* can recognize three scenarios to define the *minimal significant amplitude* to compare the signals, and the three are derived from the first *silence block* in the file.

The first scenario is the *electrical grid* frequency noise, which is searched for at either *60Hz* for *NTSC* or *50Hz* for *PAL*.²¹ This is automatically selected from the previously calculated *frame rate*, if available.

The second one is *refresh rate noise*, again derived from the frame rate, in *NTSC* it is between *15697-15698Hz*.

If neither is found, which would be surprising for a file generated by recording from a vintage console via analogue means, the frequency with the highest amplitude within the *silence block* is used.

Finally, in case none of the previous scenarios is met—or if those values are lower than *-60dBFS*—a default level of *-60dBFS*²² is used.

²¹ *PAL* needs to be tested yet, since I don't have console for that video system.

²² This can be changed via *command line* options if needed. See appendix B.1.5.

Chapter 4

How to interpret the graphs

The main output of the program is a set of different graphs that vary in quantity based on the definitions made in the *mfn* file¹, and the selected options².

The graph files are saved under the folder *MDFourier* and a sub-folder named after the input *WAV* file names. They are stored in *Portable Network Graphic (PNG)* format. Currently *1600x800* graphs are used, although this can be changed via options.

For the current document *800x400* graphs were used in order to fit within a *PDF* or *HTML* presentation. The output graphs that are created by the software are listed and described in appendix B.1.4.

In our examples for the *Mega Drive/Genesis*, there are three *active blocks*³: *FM*, *PSG* and *Noise*. These will result in a graph of each type, plus a general one named *ALL*.

We'll follow a series of results from different input files to *MDFourier*, starting with cases that have either none or a few differences and build on top of each one, so you can familiarize with what to expect as output.

4.1 Scenario 1: Comparing the same file against itself

The first scenario we'll cover is the basic one, the same file against itself. Let's keep in mind that *MDFourier* is designed to show the relative differences be-

¹Detailed in appendix C.

²See appendix B.

³The *Mega Drive/Sega Genesis* has two audio synthesizers: a *Yamaha 2612* (YM2612) for *Frequency Modulated* (FM) audio, and a *Texas Instruments SN76489 Programmable Sound Generator* (PSG)—or equivalent in an *ASIC*—for *Sega Master System* compatibility and extra audio channels.

tween two audio files.

So, what is the expected result of comparing a file to itself? No differences at all. An empty graph file as shown in figure 4.1.

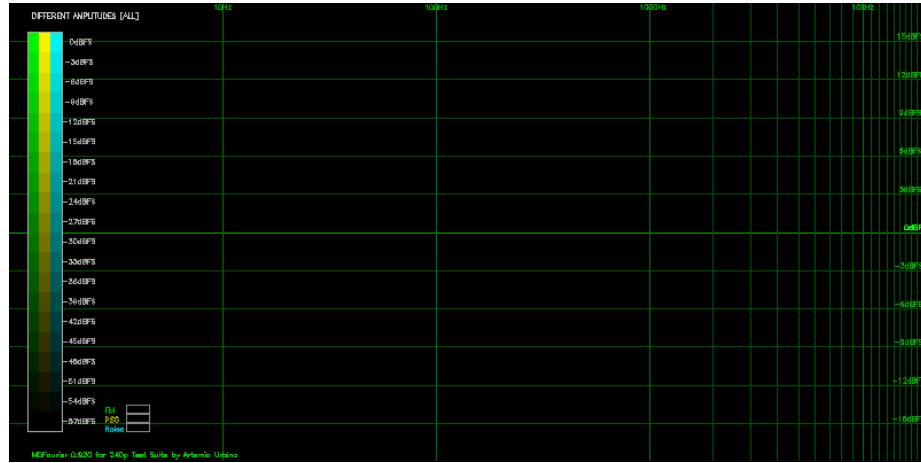


Figure 4.1 Different Amplitudes result file when comparing the same file against itself.

Of course all of the *Differences* and *Missing* graphs will only have the grid and reference bars, with no plotted information since both input files are identical.

All graphs use the horizontal axis for frequency. Values start from *1Hz* on the left and end at *20kHz* on the right, which covers the human hearing spectrum⁴.

For *Difference* graphs the vertical axis is the amplitude in *dBFS*, and a central axis for *0dBFS*. Values rise to *18dBFS* towards the top and fall to *-18dBFS* on the lower part of the graph.

There will be two sets of *Spectrograms*, one for the *Reference* file and one for the *Comparison* file, with one graph for each defined *type* plus the general one.

⁴ Although the plotted range can be altered to cover only a fraction of the spectrum, the resulting graph will hold the same scale.

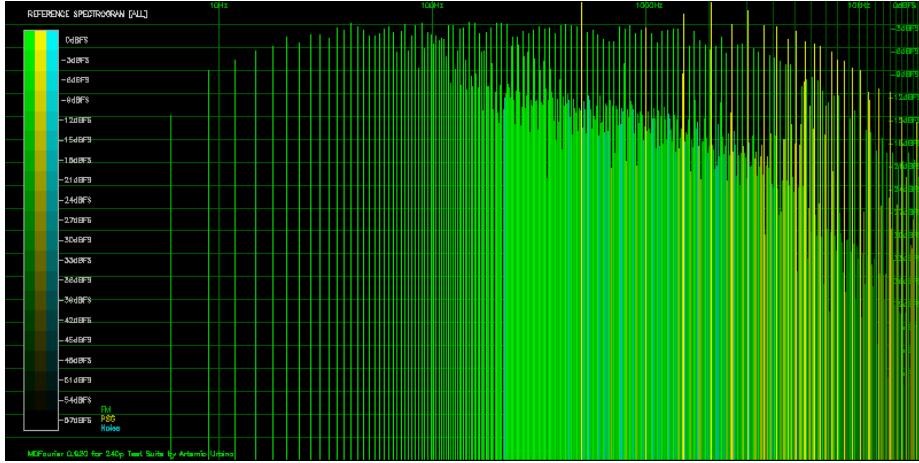


Figure 4.2 The Spectrogram for a Genesis 1 VA3 via headphone out.

The Amplitude of each of the fundamental *sine waves* that compose the original signal is represented by vertical lines that reach from the bottom to the point that corresponds to the amplitude in *dBFS*. The line is also colored to represent that amplitude with the scale on the left showing the equivalence.

Three colors—as defined from the *mfn file*⁵—are used to plot the graph, with each one of them plotting the frequencies from each corresponding *type* from the *WAV* file.

The top of the graph corresponds to the maximum possible amplitude, which is *0dBFS*. The bottom of the graph corresponds to the *minimum significant amplitude*, as described in section 3.4.

As expected, both sets of *spectrograms* are identical in this case, since we used the same file compared against itself.

4.2 Scenario 2: Comparing two different recordings from the same console

This is another control case. What should we expect to see if we record two consecutive audio files from the same console using the same sound card?

⁵The configuration file is described in appendix C.

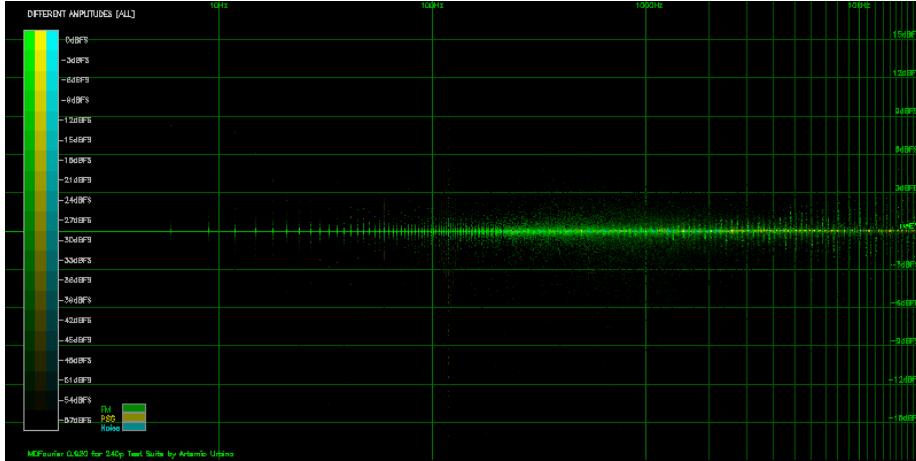


Figure 4.3 The same console, two different recordings.

As you can see, we have basically a flat line around zero. This means that there were no meaningful differences found.

But wait, there *are* differences. Why is that? Due to many reasons: analogue recordings are not always identical. There are also variations from the analogue part of console itself, and probably from the internal states and clocks from the digital side of the process. It can also be noise generated by differences in frequency bins when performing the *Discrete Fourier Transform (DFT)* after calculating the frame rates—we have that $1/4$ alignment error after all.

We now know that there will be certain *fuzziness*, or variation, around each graph due to this subtle recording and performance nuances. It is a normal situation that is to be expected, and a baseline for further results.

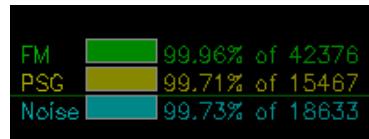


Figure 4.4 Bars for percentage matched.

Until now we haven't discussed the small bars at the lower left quadrant of the graphs. These represent the percentage of matches that were made. In this case we get all above 99.8% as shown in figure 4.4. This serves as a quick guide on how relevant the plotted differences are when comparing both signals. The numbers to the right⁶ is the amount of values that were compared.

⁶Only drawn at higher resolutions, since they aren't readable in 600x400 mode.

4.3 Scenario 3: Comparing against a modified file

For demonstration purposes, the same *Reference* file was modified to add a 500Hz 6dBFS parametric equalization across all the signal. This is an artificially modified and controlled scenario in order to demonstrate what the graphs mean.

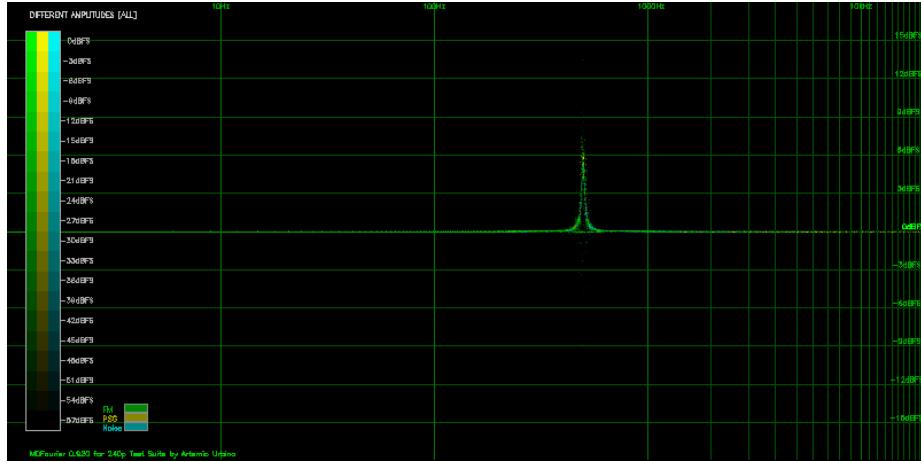


Figure 4.5 Compared against itself modified with a 500Hz 6dBFS equalization.

As expected all three *types* (*FM*, *PSG* and *Noise*) were affected and show a spike, exactly 6dBFS tall and centered around 500Hz .

It is interesting to contrast both spectrograms (figures 4.6 and 4.7), since the 500Hz spike is also shown there.

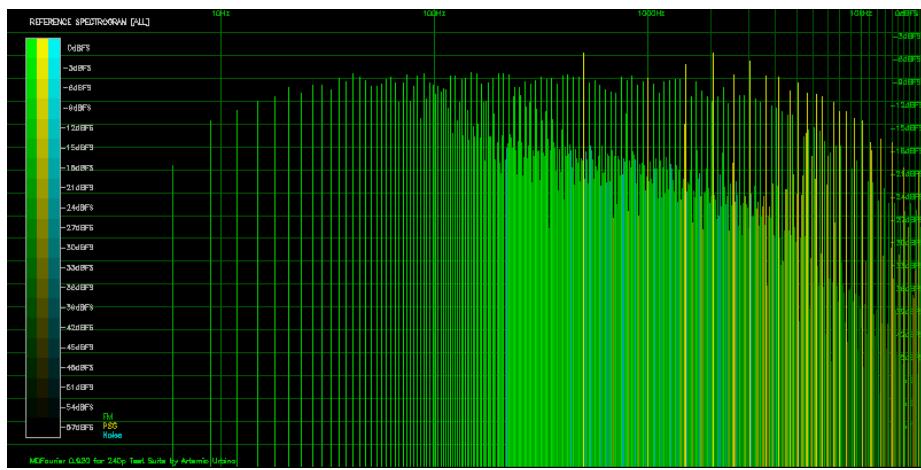


Figure 4.6 Spectrogram for the *Reference* File.

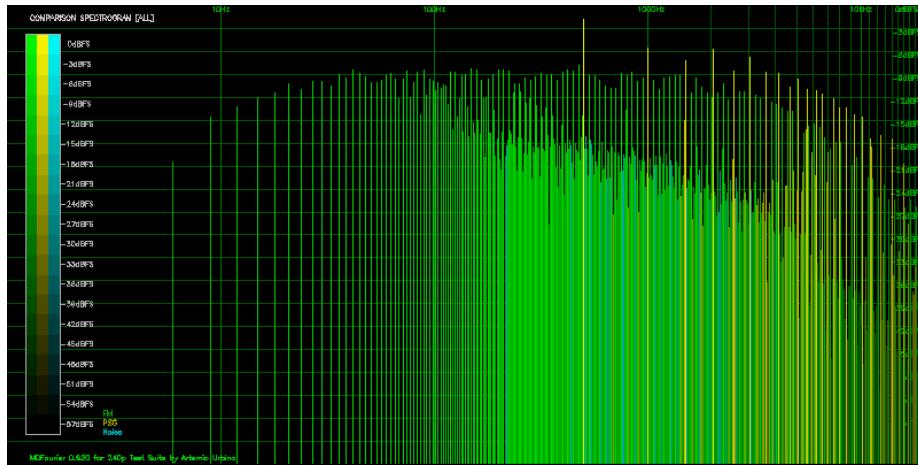


Figure 4.7 *Comparison* file modified with 500Hz peak.

And the *Missing Frequencies* graphs are basically empty, since no relevant frequencies are missing from the *Comparison* file.

4.4 Scenario 4: Comparing against digital low pass and high pass filters

We will use the same *Reference* file, and compare it to a file modified by several digital filters that were inserted via an audio editor:

- A *low pass filter*⁷ to the *FM* section of the file.
 - A steeper *low pass filter* at a different cutoff frequency to the *PSG* section.
 - A *high pass filter*⁸ to the *Noise* section at a different frequency.

Figure 4.8 shows the general graph for the three types:

⁷A low pass filter allows only frequencies lower than a specified value pass through.

⁸A *high pass filter* allows only frequencies higher than a specified value pass through.

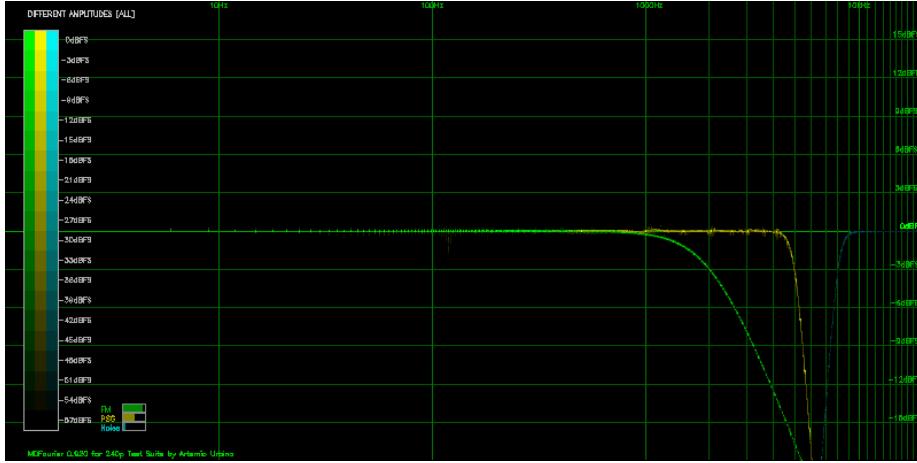


Figure 4.8 *FM* with a low pass filter (in green), *PSG* with a low pass filter (in yellow) and *Noise* with a high pass filter (in aqua).

We can now see that the higher frequencies above 1kHz in the *FM* graph quickly fall down towards $-\infty\text{dBFS}$, so the first low pass filter is there.

The second low pass filter for *PSG* is at 3kHz , and it is steeper.

But we can barely see what is going on with the *Noise* part of the graph. We can see that there is some black dots on top of the 0dBFS line.

In order to better see what is going on, we'll change the *color filter function* to $\sqrt{\text{dBFS}}$ so we can have higher contrast against the background.⁹

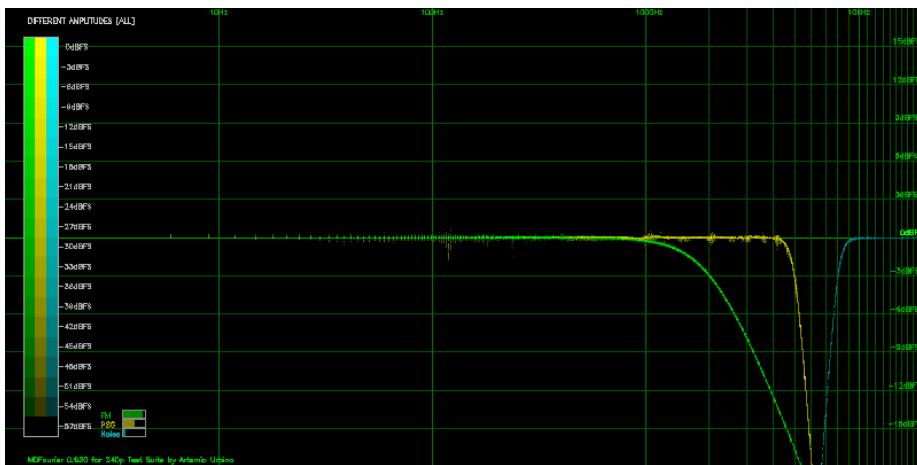


Figure 4.9 Using the $\sqrt{\text{dBFS}}$ color filter function.

With the new emphasis, we can now make out the curve that raises from

⁹Described in appendix B.1.2.

$-\infty dBFS$ to $0dBFS$, and it aligns with $8kHz$.

We can still do better than that, by using the *Average Graphs* option.¹⁰

Here is the resulting graph for only the *Noise* section of the signal, with average enabled:

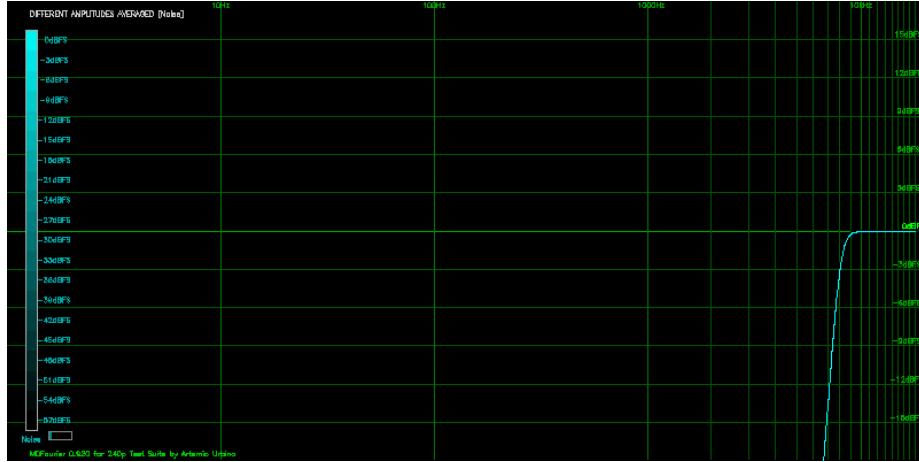


Figure 4.10 Noise plot with the *Average Graphs* option enabled.

There are some other interesting graphs that result from this experiment. For example, the *Missing* graphs now show all the frequencies the *low/high pass* filters cutoff.

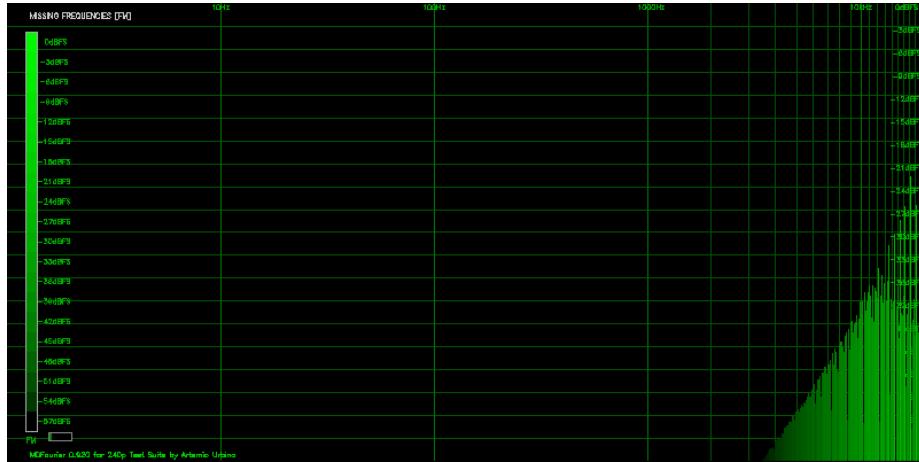


Figure 4.11 Missing frequencies in *FM* cutoff by low pass filter.

As shown in Figure 4.11, there is a curve in the *spectrogram* and only frequencies above $1kHz$ show up, slowly rising in amplitude. Many of these are

¹⁰See section B for details.

not shown since they are below our defined *minimum significant amplitude*.

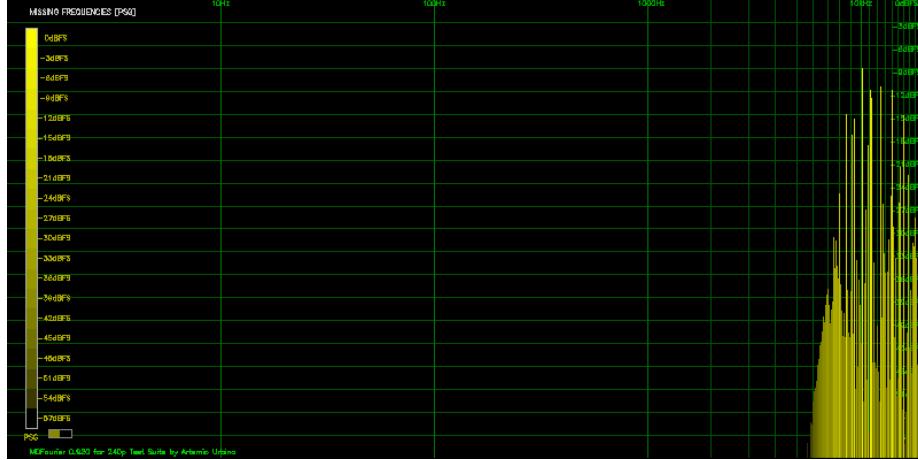


Figure 4.12 Missing frequencies in *PSG* cutoff by low pass filter.

The same behavior can be observed in in Figure 4.12. This is the *PSG spectrogram*, and it shows a different curve that starts at 4kHz .

As you might have noticed, a bar is also present in the *Missing graph*. But in this case it represents the percentage of unmatched frequencies, as shown in Figure 4.13.

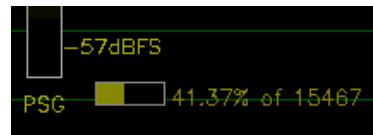


Figure 4.13 Bar for missing frequencies in *PSG* cutoff by low pass filter.

And finally, figure 4.14 shows the opposite kind of curve, the *high pass filter* that cuts off everything higher than 8kHz in the *Noise* section.

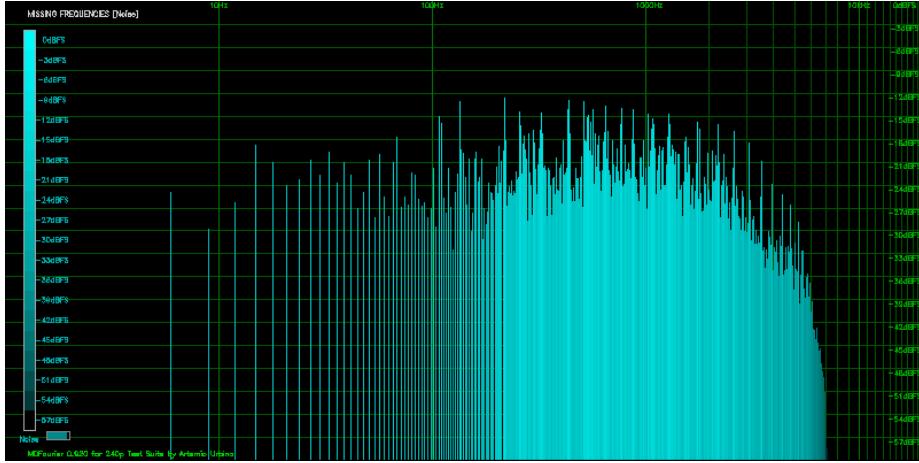


Figure 4.14 Missing frequencies in *Noise* cutoff by high pass filter.

It is a good moment to emphasize that these are relative graphs. They show how different the *Comparison* signal is to the *Reference* signal. And so far we've compared the same signal to itself, although modified though very precise digital manipulations. An analog filter would look the same, but a bit fuzzier.

However, some interesting ideas arise. What would happen if we take this *low/high pass filter* signal and use it as *Reference* and the original one as *Comparison*?

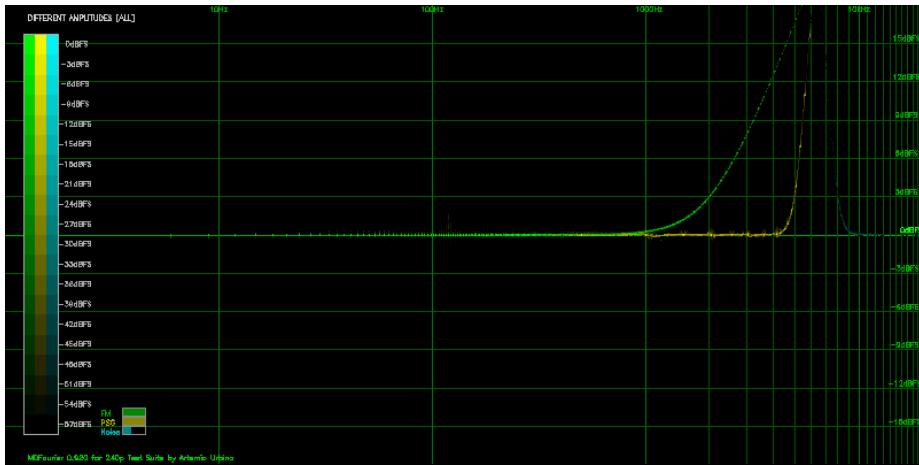


Figure 4.15 Results when using modified signal as *Reference*.

Based on this, one could jump to the conclusion that everything will simply be inverted. After all, the original signal now rises to $+\infty dBFS$ at the same spots—and that makes complete sense—since those frequencies now have a higher amplitude.

Although the *Differences* graphs will indeed be inverted under these controlled conditions, the *Missing* graphs are different. Most of them are now empty:



Figure 4.16 *Missing Frequencies* graph for *FM* is empty.

This happens because although we cut a lot of frequencies with such steep *low* and *high pass filters*, all the frequency content from this modified signal is present in the original, but not the other way around as shown above.

4.5 Scenario 5: Comparing two recordings from the same console made with different Audio Cards

We'll now compare the same console using two different recordings, one made with the *USB Lexicon Alpha* and the other with a *USB M-Track*¹¹. Figure 4.17 shows the results:

¹¹See appendix M.1 for details on the sound cards used.

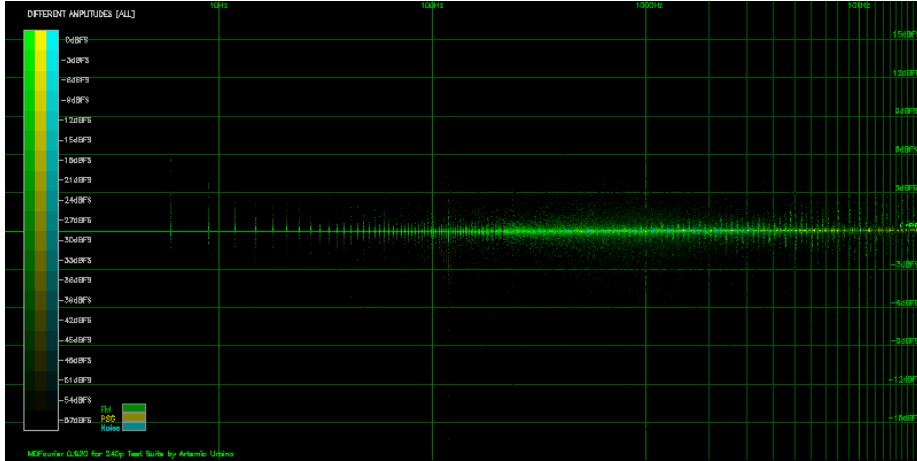


Figure 4.17 Differences using same system and cables, but different audio capture cards.

There is some scatter, as it was to be expected from the results of *Scenario 2*¹². We can tell that the scatter is centered around the $0dBFs$ line, which means that even using different sound cards we discern differences between systems¹³.

Also, there was a slight difference in the detected frame rates. This happens since the sampling clock is not exactly the same in both audio cards¹⁴. Here is the output text from the analysis that shows the detected frame rates:

```
* Loading 'Reference' audio file A-MD1UTVA3-LA.wav
- WAV file is PCM 48000Hz 16bits and 63.76 seconds long
- Starting sync pulse train: 1.88142s [361232 bytes]
- Trailing sync pulse train: 60.2921s [11576080 bytes]
- Detected 59.9204 Hz video signal (16.6888ms per frame) from WAV file

* Loading 'Comparison' audio file A-MD1UTVA3-MT.wav
- WAV file is PCM 48000Hz 16bits and 62.5774 seconds long
- Starting sync pulse train: 1.44967s [278336 bytes]
- Trailing sync pulse train: 59.8602s [11493160 bytes]
- Detected 59.9208 Hz video signal (16.6887ms per frame) from WAV file

- Reference signal stereo imbalance: left channel is higher by 0.570059%
- Comparison signal stereo imbalance: right channel is higher by 0.487369%
```

At $48kHz$ both cards have very accurate sampling clocks with a minimal difference¹⁵ and *MDFourier* compensates for such issues.¹⁶ Differences can be

¹²See section 4.2.

¹³Under the assumption both cards have a relatively flat frequency response, like the ones used here. For specifications see the *Audio Cards* appendix M.1.

¹⁴More information in appendix M.1.1.

¹⁵We are talking 0.0001 ms . In this case, that is 0.0000001 seconds !

¹⁶The expected frame rate from the vintage console analyzed here is 16.688 ms per frame, or 59.92 Hz as measured with a scope. See appendix F

more prominent when using 44.1kHz , but the software can still compensate for them without issues.

Here is the graph with the *Average Graphs* option turned on:

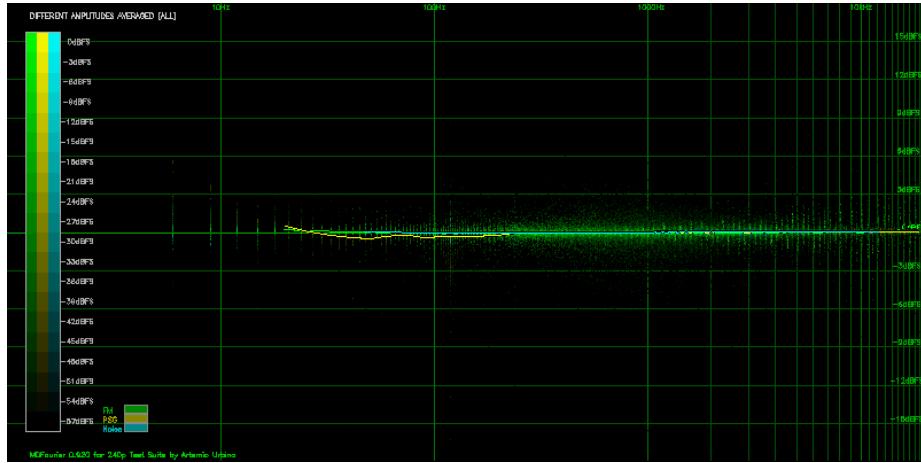


Figure 4.18 Differences using same hardware and cables, different capture cards with average enabled.

The results obtained are quite similar to the ones from *scenario 2*, even to the point of getting almost the same percentages as shown in Figure 4.19. So as long as we are using a relatively flat frequency response audio card, we'll be able to compare results from different systems even when using a variety of audio cards.

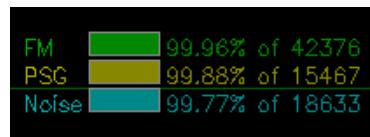


Figure 4.19 Percentage match Bars using same hardware and cables, different capture cards.

4.6 Scenario 6: Comparing *WAV* vs *MP3*

The audio file from *Scenario 1* will be used as *Reference* and *Comparison* file, but this time encoded as *MP3*¹⁷ and decompressed to *WAV* in order to use it with the software.

If the files were identical, an empty graph would result as shown in Figure 4.1 from *Scenario 1*. However we get something slightly different:

¹⁷The file was created with the *MP3 LAME* encoder using *variable bit rate* in high quality mode, with slow encoding enabled and using real stereo.

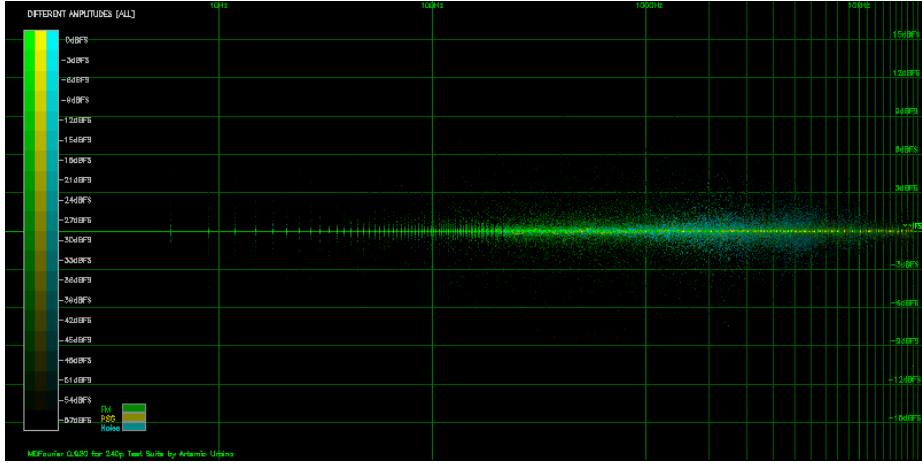


Figure 4.20 Differences between *WAV* as *reference* and *MP3* as *Comparison*.

We get everything centered around the *0dBFS*—as it should be—which indicates the *audio signatures* are very similar. Variations are correspondent to what we found in *Scenario 5* while we were comparing the same console with different audio cards. It must be noted that we used high quality compression for the test, and lower settings will provide different results. The reader is invited to experiment with the tools.

However in this case some more differences show up if we enable the *Average Graphs* option:

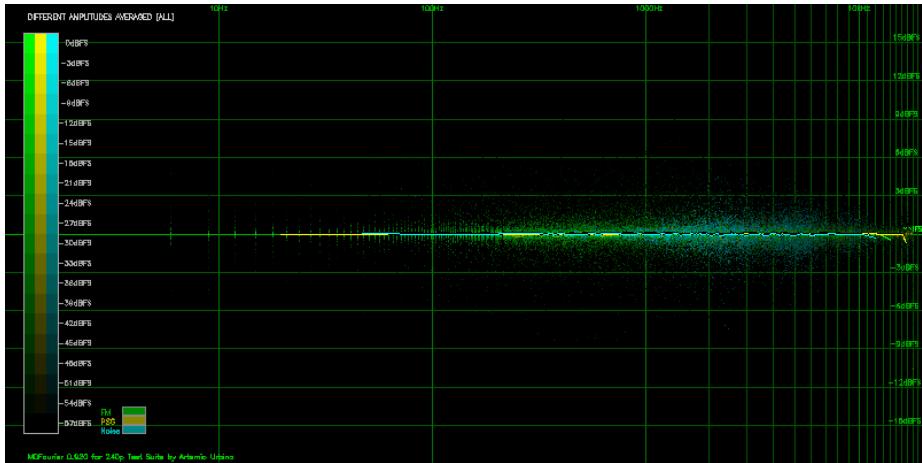


Figure 4.21 Differences between *WAV* as *reference* and *MP3* as *Comparison*, averaged.

As it can be appreciated in the rightmost part of the graph, there is a low pass filter around *18kHz*. These frequencies are at the border of acoustic human limits, and most adults won't even be able to hear them. In part this is what *MP3* compression relies on for reducing the required storage size of each audio

file: cutting off frequencies that human beings are unlikely to hear. This is how the *Fourier Transform* helps to create smaller audio-and video-files.

Another side effect of using *lossy* compression¹⁸—such as *MP3*—is the fuzzy cloud of amplitude differences present in the $500Hz$ to $8kHz$ range. We are highly tuned to this part of the spectrum, since most of the frequencies that typically compose the human voice reside here¹⁹. This is most likely what some people perceive as lower quality audio when using these type of files.

Although impressive this is not good enough for our purpose, since we are trying to measure small differences between signals. This kind of noise would make it harder to identify the real source of the differences, specially with different encoding settings, audio cards and between consoles.

Another graph that shows the same differences is the *Missing* frequencies graph:

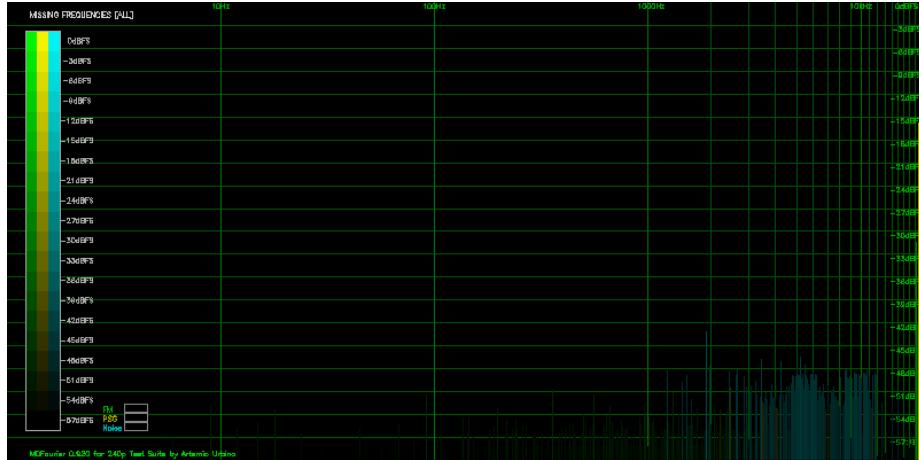


Figure 4.22 Missing frequencies from *MP3*.

It can be appreciated that the missing frequencies were detected after the $18kHz$ line, aside from some harmonics below $-45dBFS$.

¹⁸This means it loses some information, as contrasted to lossless.

¹⁹See [32].

Chapter 5

Results from vintage retail hardware

The following table lists all the hardware used to make the recordings for the graphs in this chapter. All systems had stock parts at the time of recording, no modifications and used original power supplies.¹.

The *MDFourier ID* value in the following table refers to the file name in the catalog².

Type	Model	Revision	FCCID	Serial	Region	Made in	MDFourier ID	Recorded from
Model 1	HAA-2510	VA1		89N61751	Japan	Japan	A-MD1JJVA1	Headphone Out
Model 1	1601	VA3	FJ846EUSASEGA	30W59853	USA	Taiwan	A-MD1UTVA3	Headphone Out
Model 1	1601	VA3	FJ846EUSASEGA	30W59853	USA	Taiwan	A-MD1UTVA3-AV	AV Out
Model 1	1601	VA6	FJ8USASEGA	B10120356	USA	Japan	A-MD1UJVA6	Headphone Out
Model 1	1601	VA6	FJ8USASEGA	59006160	USA	Taiwan	A-MD1UTVA6-1	Headphone Out
Model 1	1601	VA6	FJ8USASEGA	31X73999	USA	Taiwan	A-MD1UTVA6-2	Headphone Out
Model 1	HAA-2510	VA6		A10416197	Japan	Japan	A-MD1JJVA6	Headphone Out
Model 2	MK-1631	VA1.8	FJ8MD2SEGA	151014280	USA	China	A-MD2UCVA18	AV Out
Nomad	MK-6100		50059282		USA	Taiwan	A-NMUT-AV	AV Out
Nomad	MK-6100		50059282		USA	Taiwan	A-NMUT	Headphone Out
CDX	MK-4121		Y40 014198		USA	Japan	A-CDXUJ-LO	Line Out
CDX	MK-4121		Y40 014198		USA	Japan	A-CDXUJ-HP	Headphone Out
CDX	MK-4121		Y40 014198		USA	Japan	A-CDXUJ-AV	AV Out

It must be considered that this is a very small sample size, and no results or interpretations should be extrapolated as generalized conclusions about differences between models.

However, since several common trends are found and the systems were sourced from very different providers across a long period, it would be too high a coincidence for these results to be by pure chance. I expect that we—as

¹They were all connected to a PVM-5041Q—a 4" *CRT*—via *RGB*, although the *CRT* was turned off while recording. Since *video refresh* noise is a very isolated frequency range it is not filtered out for comparisons, however it is detected and reported by the software for several purposes.

²All the recordings are available from the *downloads page* from the *MDFourier* website. See appendix A.

a community—can grow a bigger data set of recordings, in order to make better assessments in the future³.

For simplicity, the recordings used were made with the *USB Lexicon Alpha* audio card⁴. The first system is the *Reference*, and the second one is the *Comparison* signal.

Even if the system volume settings won't affect the analysis since *normalization* is performed in software⁵, the volume control was set to its maximum when using the *headphone out* port. This was made in order to get consistent results, since variations could arise from the performance of the system amplifiers.

5.1 Sega Genesis Model 1 VA3 (US) vs Mega Drive Model 1 VA1 (JP)

These models are an interesting starting point since the *Reference* system—the *VA3*⁶—is usually considered⁷ one of the best sounding versions of the system family line.

On the other hand, the Japanese *VA1* is the first revision to the launch *JP VA0* system. This will give us an idea of how much the audio signatures changed from the launch units in Japan to the *US VA3*—which is also the first revision to the USA system, where the launch unit was the *VA2*⁸.

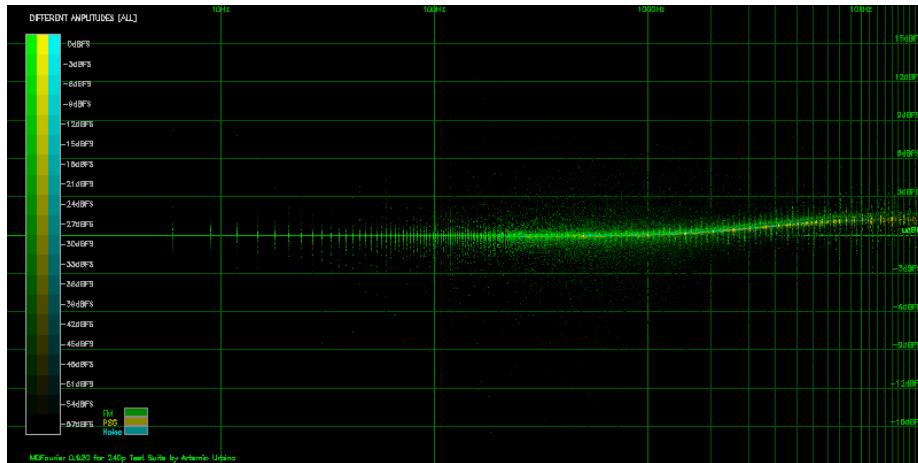


Figure 5.1 Differences between A-MD1UTVA3-LA and A-MD1JJVA1-LA.

³I've consider the viability of creating "averaged" system profiles in the future—using an extra tool yet to be developed—that could give us a generalized profile created from multiple samples.

⁴Reasons are: flat frequency response, available in the market at the time of writing, works externally via USB and is not expensive. See appendix M.1.

⁵For details in the *normalization* procedures used, see appendix E.

⁶This refers to the motherboard revision in the system, see [2].

⁷See [3] for more information.

⁸See [4] for information on hardware revisions.

The resulting graph shows that the signals are quite similar, where only the higher frequencies—the treble—exhibit some deviation. The curve starts around 2kHz and slowly rises about 1dBFS above the 0dBFS line.

It is generally accepted that the minimum perceivable difference—under normal conditions—is around 3dB ⁹. So although a measurable difference exists—which is possibly noticeable under ideal listening conditions—the systems are very similar.

It is also noticeable that the percentage match bars are full, meaning that practically all the frequencies compared were matched.

5.2 Sega Genesis Model 1 VA3 (US): Headphone out vs AV out

Ever wondered if something was lost—besides stereo—when using the *AV out* port from the console instead of the *headphone out*? Figure 5.2 shows both ports them compared using the same system.

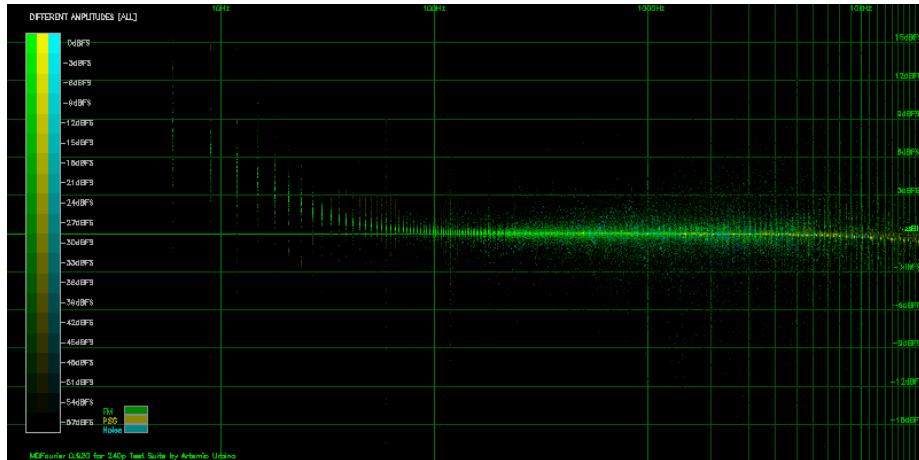


Figure 5.2 Differences between *Headphone out* and *AV out* for the A-MD1UTVA3-LA.

Differences are present only in *bass*¹⁰ and *treble*¹¹. In the higher end of the spectrum, differences are less than 1dBFS and higher than 17kHz . The low end differences are more interesting—but still at a frequency range that would hardly matter for most setups, aside from the lost stereo. These reach a 3dBFS difference around 30Hz , so you get deeper bass when using the *AV out* if the music¹² reaches those notes and your audio system has the means to reproduce

⁹Under ideal conditions 1dB can be achieved, see [12].

¹⁰Low frequency tones.

¹¹high frequency tones.

¹²Yes, *Bare Knuckle/Streets of Rage* does reach this part of the spectrum. It must be noted that most of the bass range is really between 60Hz - 500Hz . See [34].

it.

5.3 Sega Genesis Model 1 VA3 (US) vs Sega Genesis Model 1 VA6 (JP)

We'll again be using the *US VA3* as *reference*, now contrasted against a *Japanese VA6*. The *VA6* is regarded as a good sounding unit in par with the *VA3*¹³, and is also very common.

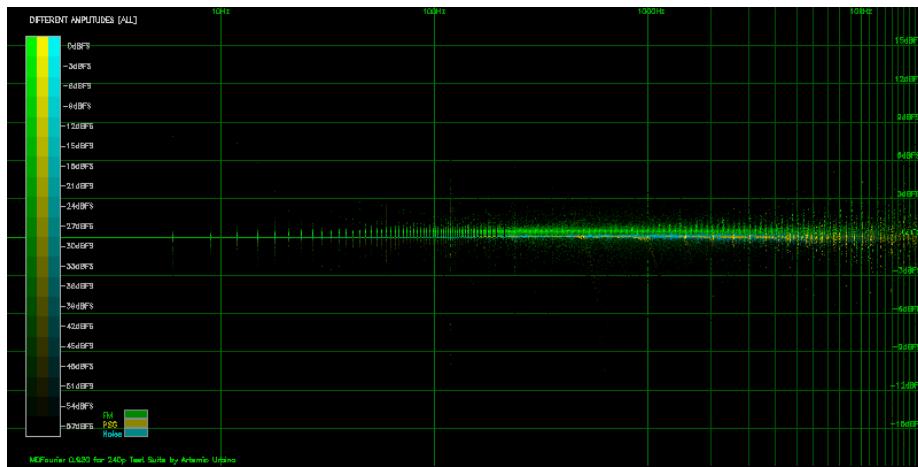


Figure 5.3 Differences between A-MD1UTVA3-LA and MD1JJVA6-LA.

The curves closely follow the $0dBFs$ line, however the *YM2612*¹⁴ amplitude is higher than the *PSG* amplitude by a really low margin, around $0.4dBFS$. Although measurable, this would be even less noticeable than the difference against the *VA1*.

5.4 Sega Genesis Model 1 VA6 (JP) vs Mega Drive Model 1 VA6 (US)

We'll now compare two systems with the same *VA6* motherboard revision. However, one is Japanese and the other is from North America. Results are displayed in Figure 5.4.

¹³See [2] for details.

¹⁴The *Yamaha 2612* synthesizer—also known as *YM2612*—is responsible for the *FM* sound in this system.

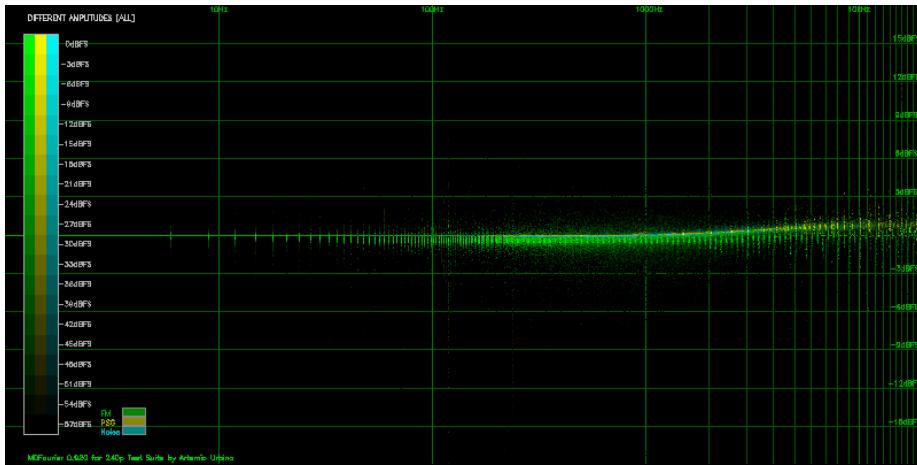


Figure 5.4 Differences between A-MD1JJVA6-LA and A-MD1UJVA6-LA.

Once again, differences are minimal and quite similar to some previous results—in particular to the ones between the *US VA3* and the *JP VA1*. But now *FM* seems to be closer and *PSG* is the one that is less than *1dBFS* higher on the *US* system.

5.5 Sega Genesis Model 1 VA6 (US) vs Sega Genesis Model 1 VA6 (US)

These are two other *US VA6* systems, both made in *Taiwan*. And here are the results:

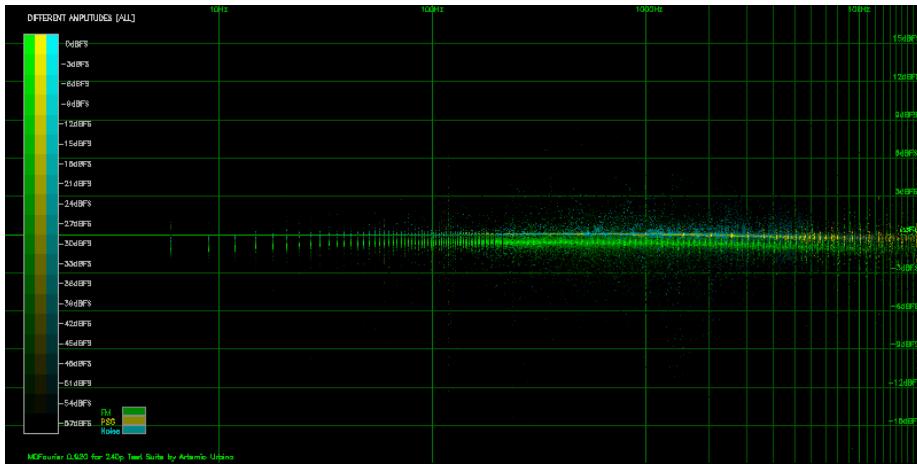


Figure 5.5 Differences between A-MD1UTVA6-1-LA and A-MD1UTVA6-2-LA.

Again, these are very close matches with differences below $1dBFS$. In this case FM is lower by that amount and PSG matches the reference unit.

5.6 Sega Genesis Model 1 VA3 (US) vs Sega Genesis Model 2 VA1.8 (US)

And now for something a bit more interesting. The same American $VA3$ unit we've compared to a few others, but contrasted against a *Model 2 VA1.8*. These *Model 2* systems don't use the *YM2612*, but an *Application-specific integrated circuit (ASIC)* *Yamaha YM3438*, which is a variation used for some later models.

These *Model 2* versions—grouped with $VA7$ and other revisions—are regarded as having a lower quality sound output. This is due to various factors: the variation in FM synthesizer, the amplifier circuits used and the general audio circuit design.

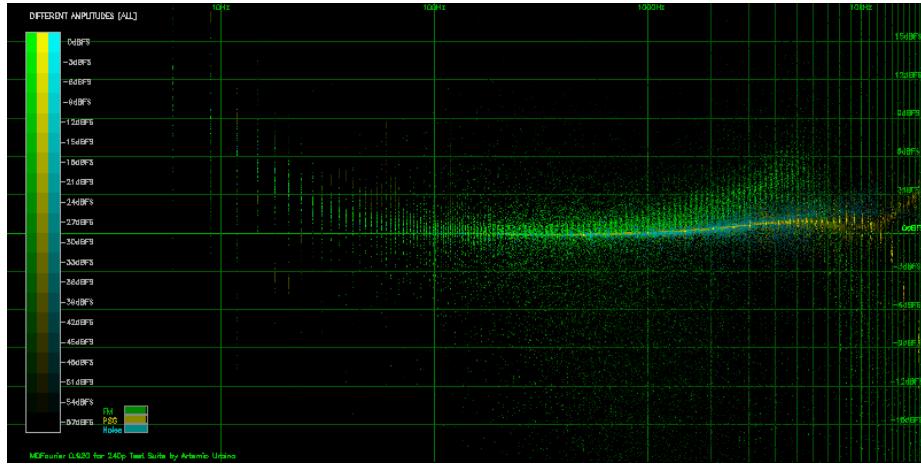


Figure 5.6 Differences between A-MD1UTVA3-LA and A-MD2UCVA18-LA.

The differences in this case are more noticeable: FM raises above $3dBFS$ —which is generally accepted as a discernible threshold¹⁵—and does this around frequencies to which our hearing is quite sensitive¹⁶, from $2kHz$ to around $8kHz$.

But since a lot of data points are highly scattered, we can gather more accurate information for analysis using the *Average Graph* option as shown in Figure 5.7.

¹⁵See [12] for results from different studies.

¹⁶Since this range falls within the spectrum for human voice, see [32].

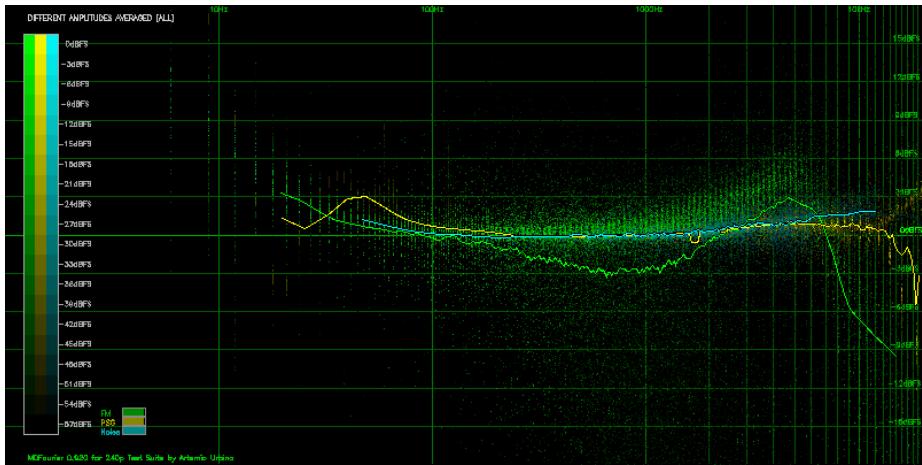


Figure 5.7 Averaged Differences between A-MD1UTVA3-LA and A-MD2UCVA18-LA.

We can now appreciate that there are some relevant frequencies—that is, *high amplitude* in the *reference* signal—that pull the averaged graph lower between *200Hz* and *3kHz-FM* spectrum, and these go down up to *3dBFS*.

We can also make out that peak in the human voice spectrum that goes *3dBFS* higher, and then it dips in the treble above *6kHz* to amplitudes lower by *6dBFS* after *10kHz*.

On the other hand, *PSG* seems quite similar aside from a bump around *60Hz* and dipping in amplitude after *14kHz*.

It must also be noted that, for the first time, there is a drop in the percentage of matched frequencies, shown in the bars detailed in Figure 5.8.

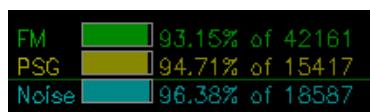


Figure 5.8 Percentage Match Bars between A-MD1UTVA3-LA and A-MD2UCVA18-LA.

So far we had been getting results above *99.7%*, and this time it dropped to *93.15%* for *FM*.

5.7 Sega Genesis Model 1 VA3 (US) vs Sega Nomad (US)

The *Sega Nomad* is a portable version of the *Sega Genesis*, and it had only a single motherboard revision. Recordings were made via both: *AV out* and *headphone out*.

Our first contrast example will be using the *headphone out* port from both consoles, shown in Figure 5.9.

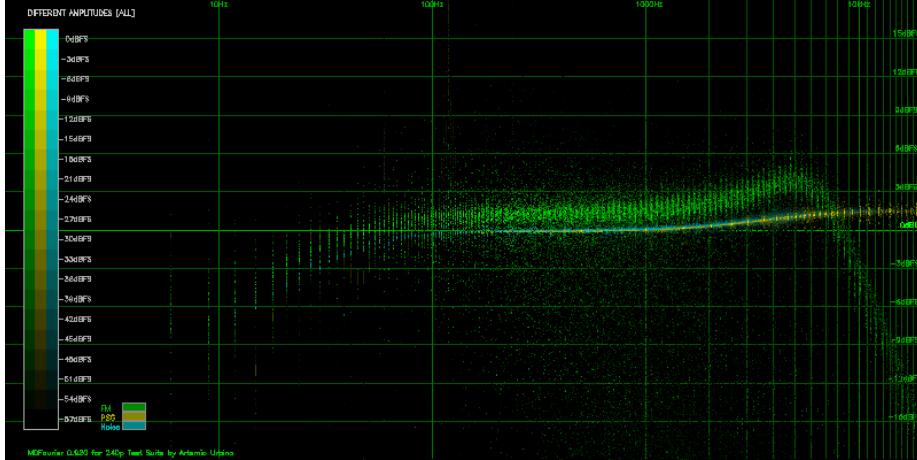


Figure 5.9 Differences between A-MD1UTVA3-LA and A-NMUT-LA.

Some resemblance to the previous comparison to the *Model 2 VA1.8* can be seen. In order to contrast both signals, an *averaged graph* is shown below.

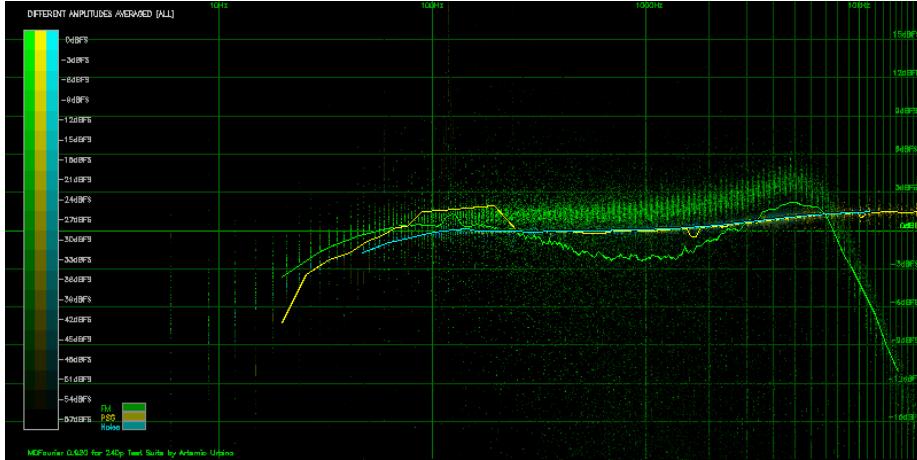


Figure 5.10 Averaged Differences between A-MD1UTVA3-LA and A-NMUT-LA.

The same general behavior from figure 5.7 can be seen in *FM*, but *PSG* doesn't fall in the same way in the high frequency range. This must be due to differences in the *mixing circuit* or amplifiers.

And if you've wondered about the difference between using the headphone out or the AV out in a Sega Nomad, here are the results:

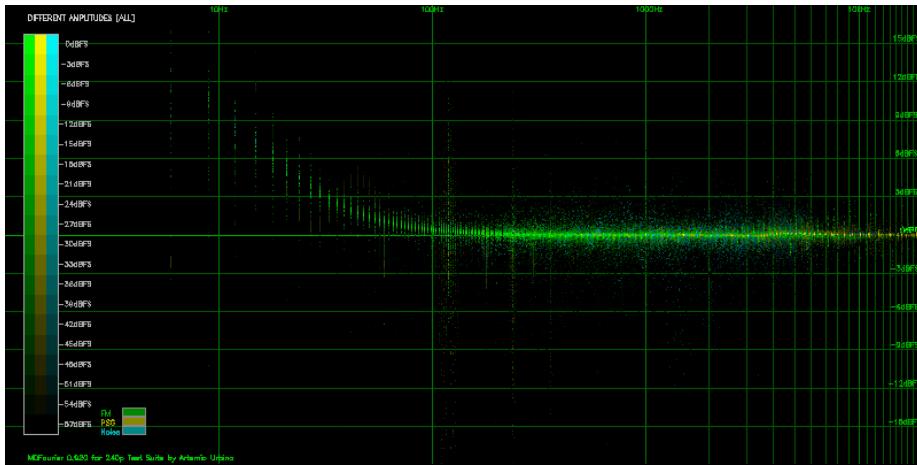


Figure 5.11 Sega Nomad: *headphone* vs *AV out*.

In this case there are no differences above 100Hz . However the *AV out* has a more low frequency emphasis that is 3dBFS higher below 40Hz .

5.8 Sega Genesis Model 2 VA1.8 (US) vs Sega Nomad (US)

Due to the similarities shown between the *Nomad* and the *Model 2 VA 1.8* when contrasted against the *Model 1 VA3*, here is the comparison between them using the *AV out*.

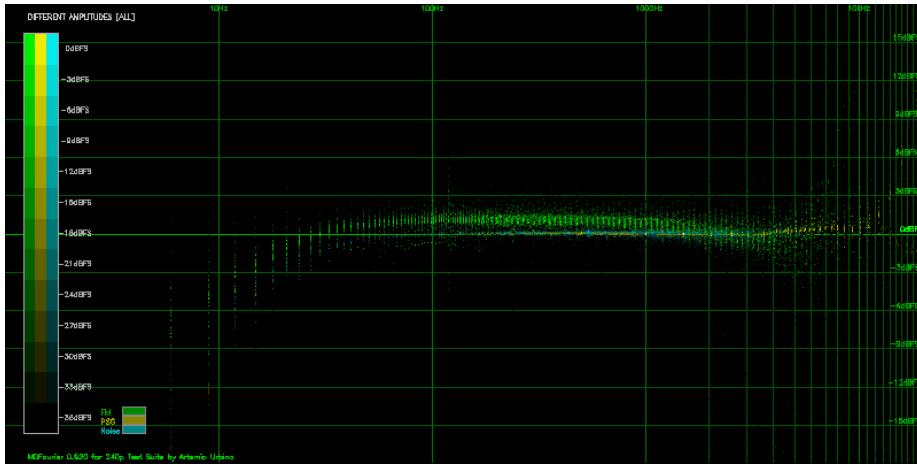


Figure 5.12 Differences between A-MD2UCVA18-LA vs A-NMUT-LA.

The consoles are indeed similar between themselves, with differences prob-

ably arising from amplification circuits. The *Low pass filter* is steeper in the *Sega Nomad* for *FM*, among other differences in the high range regarding *PSG*.

Another interesting point is that both of these systems exhibit video refresh noise at a higher amplitude via the *AV out* than their *Model 1* counterparts: around $-45dBFS$ instead of $-70dBFS$.

5.9 Sega Genesis Model 1 VA3 (US) vs Sega CDX (US)

The *Sega CDX* is an all-in-one system that includes the *Sega CD* unit. It has three different output options: *headphone out*, *line out* and *AV out*.

We'll start by with the results of the *Model 1* as *reference* contrasted to the *CDX line out*.

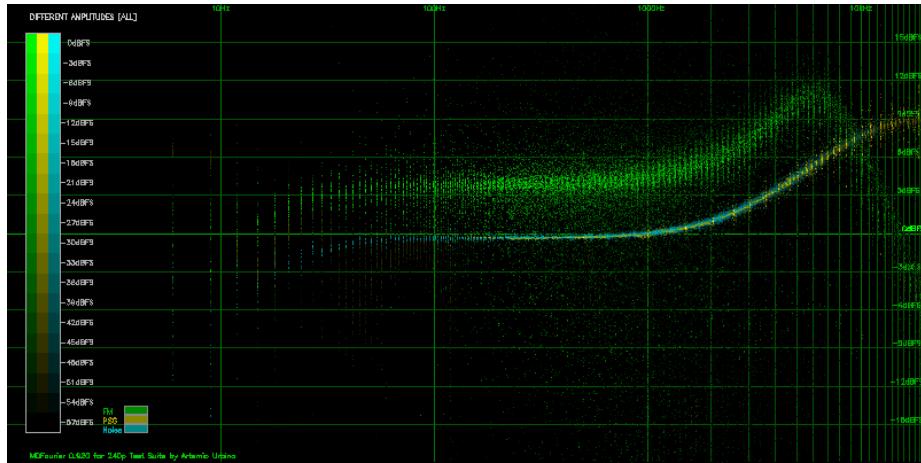


Figure 5.13 Differences between A-MD1UTVA3-LA and A-CDXUJ-LO-LA.

These are easily the most notorious differences between *vintage retail hardware* I've found so far.

The *CDX PSG* does not have the low pass filter present in *Model 1* systems. The *FM* output is higher by at least $3dBFS$, and it has a *low pass filter* around $9kHz$.

Figure 5.14 shows the results when the *average* option is enabled.

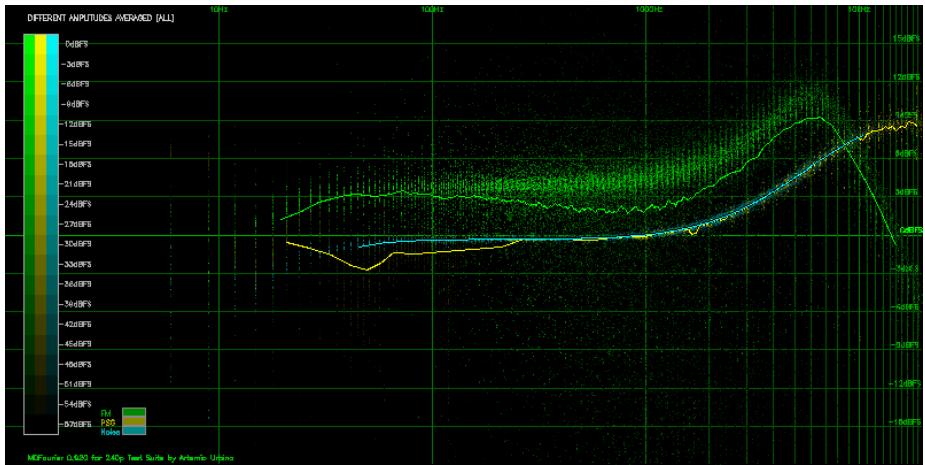


Figure 5.14 Differences between A-MD1UTVA3-LA and A-CDXUJ-LO-LA Averaged.

This makes it easier to discern the curve that the *FM* audio follows when compared to the *Model 1*. It reaches as up to *9dBFS* higher between the *4kHz-9kHz* range.

We must remember that the *Model 1*, used here as reference, has it's highest amplitude in the *PSG* signal with the current test. That's why the results show *FM* higher instead of a lower *PSG*, since *MDFourier* does relative comparisons based on the *Reference* signal.

5.10 Sega CDX (US) Line-out vs Sega CDX (US) Headphone port

And for completeness, here is the *differences* graph when contrasting the *Sega CDX line out* vs the *headphone out*.

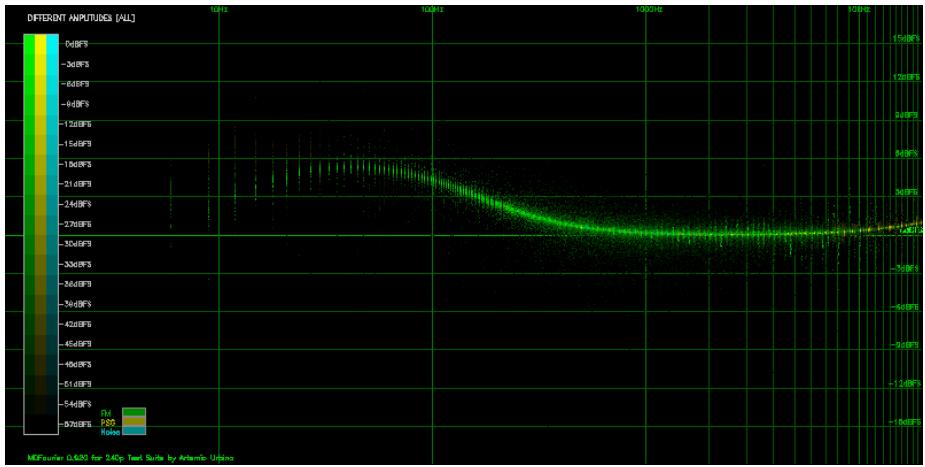


Figure 5.15 Differences between A-CDXUJ-LO-LA and A-CDXUJ-HP-LA.

The *headphone out* has clear emphasis in the low frequency range, and even below 1.5kHz.

Chapter 6

Future developments

PC Engine/Turbo Grafx-16 support would be the next target, as well as all systems that are currently supported by the *240p Test Suite*.

Chapter 7

Conclusions

So far *MDFourier* has shown results that are consistent with previous empirical evaluations, and can support many such claims with data. I hope it can provide the community with a process that is repeatable and objective that can be used in any project where someone wants to preserve, compare, replicate or play with the results.

I believe these cases support the idea that this process has value, and that it can be replicated in as many platforms as possible. Audio has been historically relegated to a secondary position, when the audio visual experience is one—as a whole. Games are not the same without their acoustic experience.

This could help to document the original audio signatures of old systems. And of course, to enable the community to play with variations and modifications that are as valid as the ones the systems originally had at launch.

Appendices

Appendix A

Downloads

The project web page is available at <http://junkerhq.net/MDFourier/>, where the latest version of this document is available in *PDF* and *HTML*.

The *Downloads* page contains the following items:

- A pre-compiled *Microsoft Windows* executable of the latest build with *GUI* front end.
- The custom binaries for supported platforms.
- Updated source code for all tools.
- Sample recordings used for the creation of this document as detailed in chapter 5.

Appendix B

How to use the Front End

The current version of the front end allows access to the main options of *MDFourier*. It is a *Windows* executable and all corresponding files must be placed in the same folder. *Uncompressing* the package to a folder should have all that is necessary to run the program.

After executing *MDFourierGUI.exe* you should be presented with the following interface:

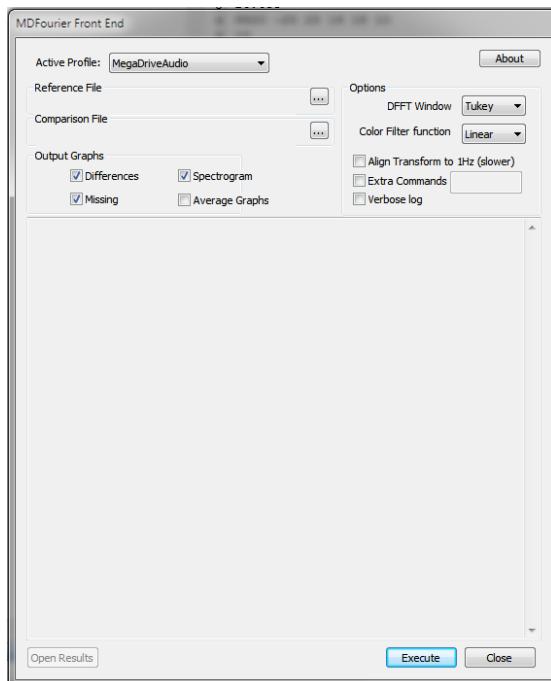


Figure B.1 MDFourier Windows Front End.

In order to generate the output graphs, two files must be selected to compare them. One as a *Reference* and the other as the *Comparison* file, as detailed in section 3.

The following sequence of steps indicates the typical work flow within the *GUI*:

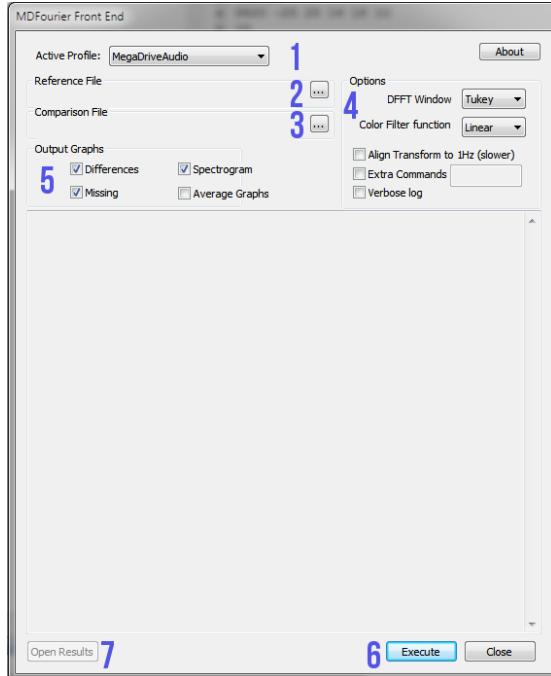


Figure B.2 Typical sequence of steps.

1. Select a *Profile* for the analysis.
2. Select a *Reference* file.
3. Select a *Comparison* file.
4. Change the default *options* if needed.
5. Select the desired output *graphs*.
6. Execute *MDFourier*.
7. When execution ends, open the *results folder*.

The *Front End* will display the output text from the command line tool, including any errors or progress as it becomes available.

Keep in mind that the *Open Results* button will only be enabled after a successful comparison between files has finished, and it won't open a second instance of the window if you have one already present.

The default options will generate graphs that work on most situations. In some cases fine tuning the results might highlight specific aspects. These options will be described in the following sections.

B.1 Front End Options

The currently available options in the *Front End* are:

B.1.1 Window Functions

In order to reduce *spectral leakage*¹ when applying the *DFT*, a *filtering window*² is applied to each element to be compared between both signals. Since we are generating the signal ourselves from the *custom binary*, the signal can be analyzed as *periodic*³, and has a natural *attack* and *decay* rate if possible.

By default we use a custom *Tukey window*⁴ with very steep slopes. *MD-Fourier* does offer alternate windows as options for further analysis.

B.1.2 Color Filter Functions

Each dot in the *Differences* graph uses the *X axis* for the frequency range and the *Y axis* for the amplitude difference between the *Comparison* and *Reference* signals.⁵

Color intensity of each dot is used to represent the amplitude for that frequency in the *Reference* signal. In other words, how relevant it was to create the original signal in that note. Please refer to chapter 4 in order to see examples of their use.

A color scale is presented in each graph, with the color graduation and the corresponding amplitude level.

¹This term describes new frequency components created when applying a Fourier Transform to a segment of a signal. See [20].

²All details regarding the windows used, their formulas and graphs are in appendix K.

³In theory this means we could get away without applying a window, but in practice the signal is not perfect and *windowing* the data helps to eliminate noise.

⁴This is not a common choice for window function, but was selected due to the above reasons after comparing results with the rest of the available options.

⁵See section B.1.4 for output file details.

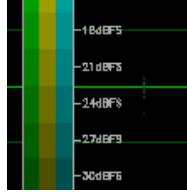


Figure B.3 Detail of color scale in graph.

The options are useful to *highlight* or *attenuate* these differences by applying the range to one of the following functions.

They are sorted in descending order. The topmost option will highlight all differences; and the bottom one will attenuate most of them, and show just the ones with highest amplitudes in the *Reference* signal.

All filters, their graphs and effects are listed in appendix L.

B.1.3 Align Transform to 1Hz

When designing the *audio signal* for use during analysis, one consideration is how to balance gathering more information versus the duration of each recording.

For practical reasons, it is desirable to have a short test that will be recorded for analysis. This reduces the time it takes to digitize audio from several systems, the storage used, analysis time by the software, and makes distributing such files easier.

In contrast, when applying the *Fourier transform* a compromise is made between frequency detail and time accuracy, very similar to the Heisenberg's uncertainty principle. If we compare a longer signal for each element, we end up with more frequency information. In our case, we don't care much about time accuracy but we do care about the length of the test. Nobody wants to record a 5 or 10 minute signal for each test to be made.

The compromise made is to use *sub second signals* for each element to be compared. Since the time and sample rate determine how the *DFT frequency bins*⁶ are spaced after analysis—and how much information we end up analyzing—the end result is a lower graph resolution.

Zero padding the input signal for Fourier analysis is a controversial subject⁷, but for the present application no adverse effects have been found. This might be the case since we control how the source signal is generated, with a predetermined *attack* and *decay*. However, it is disabled by default so the results presented can be free from questioning regarding the effects, and available

⁶The size of each slice in Hertz is the derived unit of frequency in the International System of Units (SI). It is defined as one cycle per second for the frequency spectrum.

⁷Zero-padding a signal does not reveal more information about the spectrum. See [21] [22].

for cases where more precise frequency information is needed with no spectral leakage to adjacent *frequency bins*.

B.1.4 Output Graphs

Several graphs will be generated as a result of the analysis. There will be several graphs of each type in the *output folder*⁸. One for each type, and one that summarizes all types in a single graph.

All graphs are saved in the *PNG*⁹ format, which is lossless and open source. Please read section 4 for examples and a guide to interpret their meaning.

Different Amplitudes

Enabling this option creates the most relevant output graphs from *MDFourier*. These contain the amplitude difference for the frequencies common to both files across the hearing spectrum using the *Reference* file as control.

If the files are identical, the graph will be a perfect line across the *0dBFS* line. In case the signals differ, a scatter graph will show how it behaves across the typical human hearing spectrum¹⁰.

The file names for these graphs have the format:
DA_Reference_vs_Comparison_block_options.png

Missing Frequencies

Graphs the frequencies available in the *Reference* file but not found in the *Comparison* file within the significant amplitude range. This is in effect a *Spectrogram*, but limited to the frequencies that were expected to be in the intersection but that were not present in the *Comparison* file.

The file names for these graphs have the format:
MIS_Reference_vs_Comparison_block_options.png

Average Graphs

This option traces an *average line* on top of the *Difference* graphs, making it easier to follow the trend when the output has severely scattered data. It is *off by default*.

⁸It is named *MDFourier* and a subfolder named as the compared files. It can be accessed from the *Front End* with the *Open Results* button.

⁹*Portable Network Graphic* support via *libpng* [8]

¹⁰Typically defined as 20-20000Hz, but the higher end tends to be lost with age.

The curve is created by averaging time segments from the frequency data sorted in ascending order. A *Simple Moving Average*¹¹ is then calculated to smooth out the results.

The curve is weighted according to the *Color Filter Functions* described in section B.1.2, by repeating each data point by the amount mapped in the 0-1 interval described by the function.

As a result, the average will follow the relative amplitudes from the *Reference* signal proportional to the selected *filter function*.

If there is need for a graph without weighting, please disable the *Color Filtering function*.

The file names for these graphs have the format:

DA_AVG_Reference_vs_Comparison_block_options.png

Spectrograms

Plots all the frequencies available in each file¹². Two sets of spectrograms are generated, one for the *Reference* file and one for the *Comparison* file.

The file names for these graphs have the format:

SP_AVG_Reference_vs_Comparison_block_options.png

B.1.5 Extra Command

This checkbox enables the text field to send any extra commands that are not available via the *GUI* to *MDFourier*.

Current options as of version 0.918 are:

```
MDFourier 0.918 [240p Test Suite Fourier Audio compare tool]
Artemio Urbina 2019 free software under GPL
```

```
usage: mdfourier -r reference.wav -c compare.wav
FFT and Analysis options:
-a: select <a>udio channel to compare. 's', 'l' or 'r'
-w: enable <w>indowing. Default is a custom Tukey window.
'n' none, 't' Tukey, 'h' Hann, 'f' FlatTop & 'm' Hamming
-f: Change the number of analyzed frequencies to use from FFTW
-s: Defines <s>tart of the frequency range to compare with FFT
-e: Defines <e>nd of the frequency range to compare with FFT
-i: <i>nores the silence block noise floor if present
```

¹¹This is a type of average that is intended to smooth out sudden peaks in the data. Often used in exchange rate graphs. See [31].

¹²As limited by the analysis parameters, such as amplitude and frequency count per block.

```

-t: Defines the <t>olerance when comparing amplitudes in dBFS
-z: Uses <z>ero Padding to equal 1 hz FFT bins
-n: <N>ormalize: 't' Time Domain Max, 'f' Frequency Domain Max or 'a' Average
-B: Do not do stereo channel audio <B>alancing
Output options:
-l: <l>og output to file [reference]_vs_[compare].txt
-v: Enable <v>erbose mode, spits all the FFTW results
-g: Create avera<g>e points over the plotted graphs
-A: Do not weight values in <A>veraged Plot (implies -g)
-L: Create 800x400 plots as shown in the manual
-H: Create 1920x1080 plots
-D: Don't create <D>ifferences Plots
-M: Don't create <M>issing Plots
-S: Don't create <S>pectrogram Plots
-d: Max <d>BFS for plots vertically
-k: cloc<k> FFTW operations
-j: (text) Cuts per block information and shows <j>ust total results
-x: (text) Enables e<x>tended log results. Shows a table with all matches
-m: (text) Enables Show all blocks compared with <m>atched frequencies
-h: Shows command line help

```

Sending `-h` in this field will enlist all the currently supported options for your version.

B.1.6 Verbose Log

A log is always created by default when using the *Font End*, however this option enables a verbose version with the whole frequency analysis and many other details dumped to the file.

Useful for reporting errors or unexpected behavior. *Please send the audio files if possible as well!*¹³

¹³Contact details are available in appendix Q.

Appendix C

Configuration file

All these parameters are defined in the file *mdfblocks.mfn*¹. Here is the current file we are using to compare *Mega Drive/Genesis* audio characteristics:

```
MDFourierAudioBlockFile 1.0
MegaDriveAudio
16.688
8820 -25 25 14 18 10
7
Sync s 1 20 red s
Silence n 1 20 red s
FM 1 96 20 green s
PSG 2 60 20 yellow m
Noise 3 16 20 aqua m
Silence n 1 20 red s
Sync s 1 20 red s
```

This file defines what *MDFourier* must do and how to interpret the *WAV* files. For now it can read *44.1kHz* and *48kHz* files, in Stereo *PCM* format.

The first line is just a header, so that the program knows it is a valid file and in the current format.

The second line is the name of the current configuration, since I plan to add support for different consoles or arcade hardware in the future. This would imply creating a new *mfn* file for each configuration, and a specific binary to be run on the hardware.

The third line is the expected *frame rate*². This is only used as a reference

¹This is the default file name, a different file can be used and selected via the command line. In future *Front End* revisions it might be selected via a combo box. One for each supported hardware profile.

²For more details regarding frame rate, see appendix F.

to estimate the placement for the blocks within the file before calculating the precise *frame rate* of the recording as captured by the *sound card*. After that is calculated, each signal uses its own timing in order to be fully aligned. *Frame rate* variations in the order of 0.001ms are natural, since we have an error of $1/4$ of a millisecond³ during alignment, and differences also occur by the deviations in *sample rates* and audio card limitations⁴.

The fourth line defines the characteristics of the pulse tone used to identify the starting and end points of the signal within the *WAV* file. Its frequency, relative *amplitude difference* to the *background noise* (silence), and length *intervals* that will be better explained in future revisions of this document.

The fifth line defines how many different blocks are to be identified within the files. There are seven blocks in this case.

Each block is composed of six characteristics: A *Name*, a *type*, the *total number of elements* that compose it, the *duration* for each element specified in frames, the *color*⁵ to be used for identifying it when plotting the results and a letter marking if it is a *stereo* or *mono* element⁶. Each block must correspond to a line with these parameters.

For example, FM audio has been named "*FM*", type *1*, *96* elements of *20* frames each and will be colored in *green*. Definition is in frames since emulators and *FPGA* implementations tend to run at different frame rates than the vintage retail platform, which result in different durations. The only way to align them, is by respecting the driving force that tied this up in the old days: the video signal.

There are currently two special types, identified by the letters '*s*' and '*n*'. The first one defines a *sync pulse*, which is used to automatically recognize the starting and ending points of the signal inside the *WAV* file.

The second one is for null audio, or silence. This *silence* is used to measure the *background noise*⁷ as recorded by the audio card.

³For reference $1/4$ of a millisecond corresponds to 0.00025 seconds. Modern systems have different frame rates adapted for modern displays, small differences are more likely caused by the audio capture hardware [16].

⁴See appendix M.1.

⁵Available colors are listed in Appendix O.

⁶This is important for stereo audio balancing, see appendix G.

⁷How analysis is affected by this is described in section 3.4.

Appendix D

MDWave

MDWave is a companion command line tool to *MDFourier*. During development and while learning about *DSP*, I needed to check what I was doing in a more tangible way. So, in order to visualize the files in an audio editor and listen to the results *MDWave* was born.

It takes a single *WAV* file as argument, and loads all the parameters defined in the configuration file in order to verify the same environment¹.

The output is stored under the folder *MDWave*, and a subfolder with the name of the input audio file. The default output is a *WAV* file named *Used* which has the reconstructed signal from the original file after removing all frequencies that were discarded by the parameters used.

This means that it does a *Fourier Transform*, applies the selected *window*² and estimates the noise floor. The highest amplitude frequencies are identified and limited by range for each element defined in the configuration file, and rest are discarded. An *Inverse Fourier Transform* is applied in order to reconstruct the *WAV* file and the results are saved.

The opposite can be done as well by specifying the *-x* option, and the result is a *Discarded WAV* file, that has all the audio information that was deemed irrelevant and discarded by the specified options. With this you can listen to these and determine if a more severe comparison is needed.

In addition, the *-c* option creates a *WAV* file with the chunk that corresponds to each element from the *Reference* file being used, trimmed using the detected frame rate. Two chunks are created for each element, the *Source WAV* chunk has the element trimmed without modification and the *Processed WAV* chunk has the same element but with the windows and frequency trimming applied.

It has a few more command line options, which I'll detail in later revisions

¹See appendix C.

²See appendix B.1.1.

of the document. You can type `mdwave -h` in your `mdfourier` folder for details.

```
MDWave 0.920 (MDFourier Companion)
[240p Test Suite Fourier Audio compare tool]
Artemio Urbina 2019 free software under GPL
```

```
usage: mdwave -r reference.wav
FFT and Analysis options:
-a: select <a>udio channel to compare. 's', 'l' or 'r'
-c: Enable Audio <c>hunk creation, an individual WAV for each block
-w: enable <w>indowing. Default is a custom Tukey window.
'n' none, 't' Tukey, 'h' Hann, 'f' FlatTop & 'm' Hamming
-x: e<x>cludes results to generate discarded frequencies wav file
-i: <i>nores the silence block noise floor if present
-f: Change the number of <f>requencies to use from FFTW
-s: Defines <s>tart of the frequency range to compare with FFT
-e: Defines <e>nd of the frequency range to compare with FFT
-t: Defines the <t>olerance when comparing amplitudes in dBFS
-z: Uses Zero Padding to equal 1 hz FFT bins
-B: Do not do stereo channel audio <B>alancing
-C: Use <C>omparison framerate profile in 'No-Sync' compare mode
Output options:
-v: Enable <v>erbose mode, spits all the FFTW results
-l: <l>og output to file [reference]_vs_[compare].txt
-k: cloc<k> FFTW operations
-h: Shows command line help
```

Appendix E

Normalization and amplitude matching

Since each signal is probably at its own distinct amplitude, we need to perform a *normalization* process in order to have a common compare point between them.

Each type of *normalization* has its own strengths and weaknesses, but the *frequency domain* normalization designed for this process is always accurate with respect to the *Reference* signal, but might be confusing in some corner cases¹ if the underlying causes are not well understood when interpreting the results.

Since silence can't be used as reference², the only other option is fixing points from within the signal. Every normalization process used follows a different logic for fixing one such point for comparison.

The default is to do this in the *frequency domain*³, but there are two other options available via command line⁴. All three methods are described in the following sections.

E.1 Frequency domain normalization

This is the default option used by the software. It involves finding the *highest magnitude*⁵ from the *Fourier Transform* of the *Reference* signal before ampli-

¹One such example is available in appendix I.

²Noise floor will vary by console and by audio card.

³This means it is performed after the *Discrete Fourier Transform* and using the data generated by it.

⁴These can be enabled via the *extra command* option from the *GUI* as explained in section B.1.5.

⁵At this point in the analysis there are no amplitudes defined, since we have no reference points. Hence, raw magnitude values from each transform are used.

tudes are calculated. Then a corresponding match in the *Comparison* signal's frequency spectrum for the same block is searched for.

This means that the exact same fundamental frequency with the highest magnitude value is searched for, occurring at the same position in time—which corresponds to the block. Having both points, a meaningful *reference point* is set for the comparison, and the relative amplitudes between the signals can be calculated.

In order to calculate the amplitudes, $0dBFS$ is matched against the absolute highest magnitude from the *Comparison* file⁶, after both signals have been relatively normalized in amplitude against the reference point.

This method has shown to be always accurate within the tests. However, results can be unexpected in certain corner cases as the one shown in appendix I.

E.2 Time domain normalization

The process is very similar to the *frequency domain* variant but it is done using the raw audio samples directly from the *WAV* file⁷.

The highest amplitude is searched for within the samples of the meaningful audio signal, as dictated by the configuration file⁸. The corresponding segment in time is then located in the *Comparison* file, and a pre-defined duration⁹ is searched for the sample with the *maximum local* value.

The *Comparison* signal is then absolutely normalized to $0dBFS^{10}$, and the *Reference* signal is then relatively normalized using the adjusted *local maximum* value. This follows the same rationale described for the *frequency domain* equivalent.

Although the results can sometimes be deceptively familiar, they are not correctly referenced in corner cases¹¹, and they do not represent the real relation between the signals. However, they can be useful for analysis while in the process of understanding how to interpret the graphs.

⁶Since the highest point of the *Reference* signal was matched to one other point in the *Comparison* file, there are only two options: either both peaks are in the exact same spot, or there is a higher peak in the *Comparison* file.

⁷Values are changed only in RAM during execution, input files are never modified.

⁸Described in section C.

⁹One frame in the current implementation.

¹⁰Relative to $0xFFFF$ —the top value in *16 bit* samples in the *WAV* file.

¹¹See appendix I for an example.

E.3 Highest fundamental average normalization

This normalization option also takes place in the *frequency domain*. The idea is to average the highest magnitudes—the fundamentals—from all segments and use the resulting ratio calculated between both signals to normalize them.

The results are always centered around the *0dBFS* line in the *Differences graph*, allowing a globalized view. However, the amplitude differences are not to be relied upon for calibration, since they are not relative to a fixed point from the *Reference* signal. This option is just available for cases when both signals present a high difference rate between them, which is unexpected but possible.

Appendix F

Frame rates and their importance

The *frame rate* is the amount of frames per second a system sends to the display. Since the *custom binary* runs from within the target system, we are subject to the internal timing. Every time a frame starts the process for being sent to the display, an event called *vertical sync* occurs. This is the driving clock for the whole console¹.

It is vital for the process to have a basic unit of time, in order to segment the file in the chunks needed for analysis and to send the correct values to the *Discrete Fourier Transform*.

Signals generated from different sources can have dissimilar *frame rates*, even when running the same programs and representing the same platform. This can be due to several reasons: having a different display technology as target, hardware inconsistencies between revisions, etc.

Another source for variation is the *audio capture device*, since these can vary slightly in their sample rate clocks. But usually these are lower variations that are reported and compensated for internally. It has also been found that some audio cards use a more inaccurate clock when sampling at 44.1kHz and a very accurate one when using 48kHz ².

The frame duration is the basic unit of measurement, since we are dealing with signals that are well below one second. Because of this, the configuration file³ defines the expected frame duration in milliseconds as measured from a vintage console. One such example is the frame duration as measured from an *NTSC Sega Genesis* using an *oscilloscope*⁴ as shown in Figure F.1.

¹ As a matter of fact, almost all the code for running games—or the custom binary in our case—is executed during a segment called *vertical blank*.

² The internal *M-Audio 192* exhibits this behavior. See appendix M.1.1.

³ See appendix C.

⁴ Electronic tool for measuring electric signals.



Figure F.1 Frame duration as measured with a scope from a *Model 1 VA3 system*.

The oscilloscope shows a frame duration of *16.888 ms*, and this is the value used in the configuration file⁵ for the initial frame rate calculations.

After the *sync pulses* are detected⁶, a new frame rate value is calculated from the audio file. This new frame rate is used across the whole process, and can vary from the vintage console due to a combination of different reasons.

Every audio card can have a slightly different *sample clock*⁷, and this variation will affect the starting and ending points within the recording. Such variation can be detected and estimated from the audio signal, and it is reported if the *verbose log*⁸ option is enabled.

The selected *sample rate* for recording can also affect the *sample clock* precision even while using the same card⁹.

The most important case, and the one for which the *sync pulse* solution was implemented, is when comparing a modern system such as an *emulator* or a *FPGA* implementation. These usually run at modern *refresh rate*¹⁰ for better compatibility with current display technology, such as *HDTVs*, since the vintage hardware has *refresh rates* that are slightly off-spec¹¹.

Since *frame rates* differ and the system uses the frame rate as a master clock

⁵Described in appendix C.

⁶Described in appendix C.

⁷See appendix M.1.

⁸See section B.1.6.

⁹This is the case when using the *M-Audio 192* [26] in *44100Hz*, but when using *48000Hz* the sample clock is spot on.

¹⁰The refresh rate is the inverse of the frame rate. For example with *16.888* it would be $1/16.888$ which is *59.92 frames per second*.

¹¹*NTSC* is *59.97 frames per second*.

for program execution, the resulting audio recordings also have a difference in length. For example when using an implementation adjusted for *NTSC*, there is a *0.001* second difference in each block of 20 frames. Although this seems to be negligible, it adds. In a one minute recording, such as the ones used for *MDFourier*, every note was misaligned due to this difference.

The software discards the extra duration in these cases, adjusting for fractions of a second every time they reach the next sample. Since the notes are not played during the whole 20 frames and the window filters help to compensate the signals, the comparison is perfectly aligned for the tests. The reader can verify these results via the *MDWave*¹² tool.

¹²See appendix D.

Appendix G

Stereo Audio Balancing

Sometimes the audio recording has different amplitude levels between the right and left audio channels, this is referred to as *stereo imbalance*. Whenever possible, *MDFourier* detects and auto compensates this imbalance in order to give a more uniform graph result. The default behavior can be disabled if desired with option `-B`¹.

This should only be a concern if an imbalance in the source system is suspected.

G.1 Possible causes

Two causes for imbalance have been identified. At the time of analysis *MDFourier* cannot possibly discern between them, and both can be present at the same time.

- Amplification deficiencies: The system being recorded has an imbalance due to aging capacitors or other issues in its internal amplification circuitry.
- Sound card balance: Another common source is caused by imbalance while capturing the audio data. This is regularly present in USB audio cards that have individual knobs to adjust gain per channel as shown in Figure G.1.

It must be noted that the internal *PCI audio card*² used during testing has minimal audio imbalance³, and this allows detecting *system amplification deficiencies* with consistency.

¹See appendix B.1.5 for details on how to use *extra commands*.

²See appendix M.1 for information regarding *audio cards*.

³In this case the detected unbalance is 0.0018% in average.



Figure G.1 Lexicon Alpha volume knobs.

The sound card balance issue can be minimized by adjusting the knobs carefully. *MDFourier* can be useful for compensating this via repeated testing, since any imbalance is reported in the output. It has also been found that leaving gain at zero in some cards, like the *M-Track*, produces more pronounced imbalance than at 50%. However this requires recording from a source that is known to be balanced.



Figure G.2 M-Track volume knobs at 50%.

In order to detect and compensate audio imbalance, at least one *monaural* block must exist within the *mfn*⁴ configuration profile being used.

The software assumes that the monaural signal must be the exact same amplitude between left and right audio channels, and compensates by normalizing the *samples* from one unbalanced channel by adjusting them via the proportional ratio.

⁴See appendix C for details.

Appendix H

Differences due to temperature

During the initial development of the software there was a single instance where unexpected variations were present between different recordings of the same hardware using the same audio capture card as shown in Figure H.1.

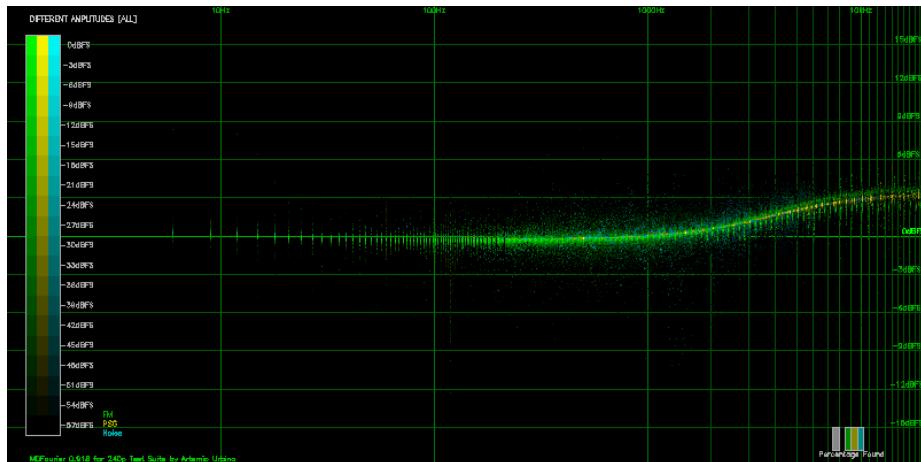


Figure H.1 Audio difference caused by heat in *Model 1 VA3 system*.

After analysis, it was found that the difference was present only in the left channel and was caused by a severe difference in temperature while recording. The room was at 29°C ¹ as shown in Figure H.2, and the *Sega Genesis Model 1 VA3* under testing had been powered on for around 30 minutes.

¹ 84.2°F .

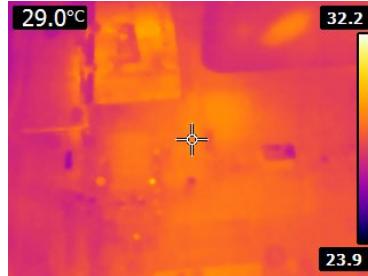


Figure H.2 Console at room temperature *Model 1 VA3 system.*

At first warm up differences were suspected, but after some cooling, *cold boot*² and measurements using a thermal camera it was determined that the room temperature on top of the warm up was the culprit, since the results could not be replicated when the room had cooled down to 22°C³.

In order to replicate the results, its vents were obstructed to prevent heat dissipation⁴ for a brief period. The results could be replicated at will under these conditions after just a few minutes of use.

It must be noted that the temperature of the headphone amplifier⁵ of the console reaches 82.9°C⁶ with cover removed as shown in Figure H.3, so higher temperatures are expected with it closed. This is important since some of the stock capacitors around the headphone amplifier are radiated due to proximity, and can probably reach temperatures beyond their intended rating in warm climates.



Figure H.3 *Model 1 VA3 system* with amp at 80°C.

²Powering on the console after it had cool down, and immediately making the recording.
³71.6°F.

⁴Ventilation were covered with electrical tape.

⁵CXA1034

⁶181.22°F.

Appendix I

Corner cases

While testing the default *frequency domain* normalization¹, I found that one artificially generated case would produce results that were unexpected and surprising. After analysis, I believe them to be the correct handling of the data, and the appropriate comparison results.

For this example we'll modify a file as detailed in *Scenario 3*², but this time a *1kHz 6dBFS* tone will be added instead of the *500Hz* from that previous case.

I was expecting the result detailed in *Scenario 3*, but was greeted with the graph shown in Figure I.1:

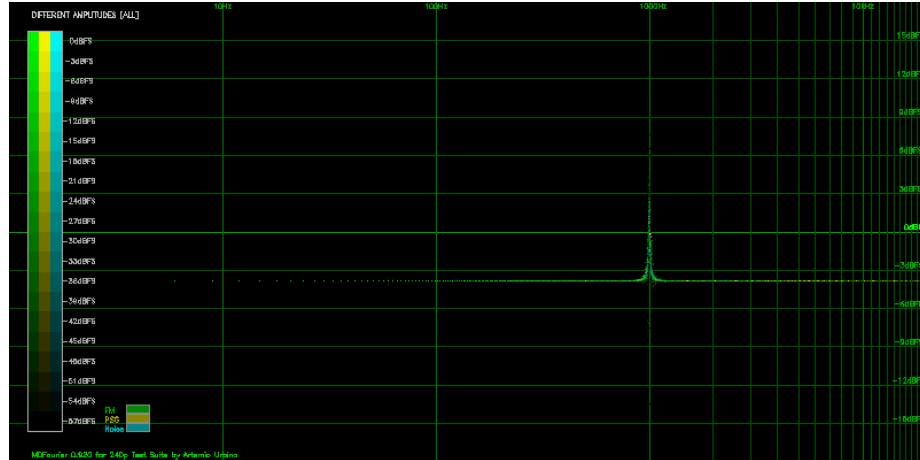


Figure I.1 Signal compared to itself with *1kHz 6dBFS* artificially inserted.

I thought it was a bug in the processing of *frequency domain normalization*.

¹See appendix E.

²See section 4.3.

But after careful analysis, this is the correct result from such process.

It is all due to a big coincidence: the maximum global *amplitude* of the original signal used for this scenario is located at 1005.99Hz , at just 5.99Hz from the 1kHz peak that was artificially inserted. As you might sharply point out, figure I.1 doesn't show the peak 6dBFS below zero. But that is because it is not centered at 1005.99Hz .

Here is the result when centering the peak at 1005.99Hz :

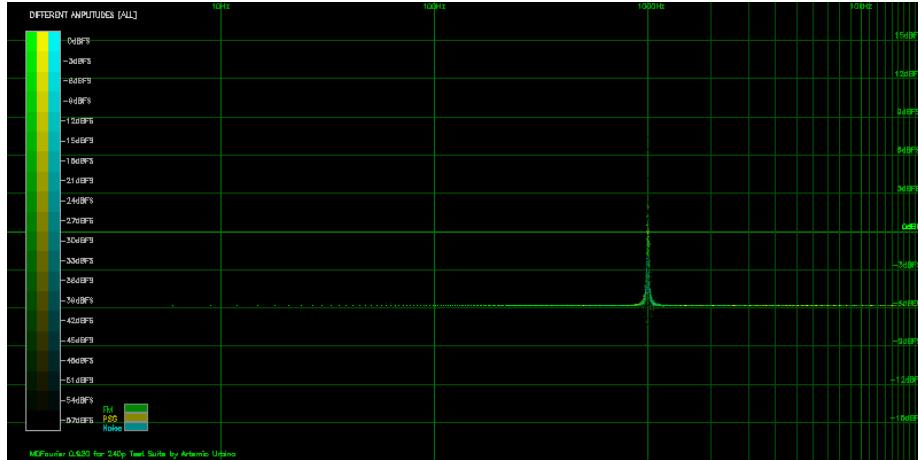


Figure I.2 Signal compared to itself with 1005.99Hz 6dBFS artificially inserted.

Now the majority of the curve is 6dBFS below zero, and the peak is at zero. But why?

As explained in appendix E, the process of frequency normalization consists in finding the highest amplitude in the signal and use that as a reference point for matching both signals. In this case, both signals have their maximum at 1005.99Hz , but we also modified the *comparison* file to have a 6dBFS 1005.99Hz peak. As a result when using the original *reference* file as a model and judging the *comparison* file against it, the rest of the signal has been lowered from this perspective.

This would mean that if two systems had these audio signatures, the resulting graph would tell us that the *comparison* console is—in general—producing the frequencies at lower amplitudes. If the *comparison* system had to be modified to match the reference system, the rest of the frequencies in the signal should be raised by 6dBFS , since the single point at 1005.99Hz already matches the reference signal. It is just a matter of perspective. It helps keeping in mind that *MDFourier* does relative comparisons, and in this framework the *reference* file is the absolute model.

In case we'd like to see a graph that would probably be easier to digest—but which I now believe is not as accurate in representing the relative differences—one could enable the *time domain normalization*. The resulting graph would be

the familiar now result from figure I.3:

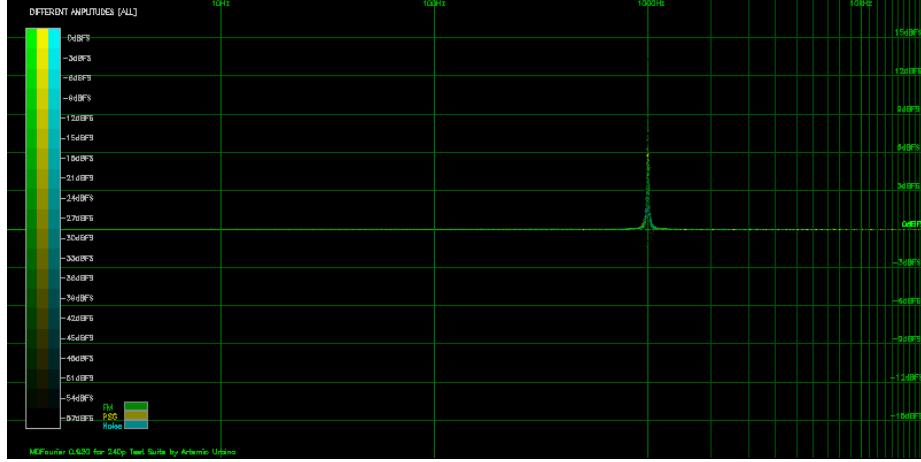


Figure I.3 Signal compared to itself with 1005.99Hz 6dBFS artificially inserted, but normalized in the time domain.

This graph tells us the exact same thing: if the *comparison* signal is to be modified to match the *reference* signal, the peak at 1005.99Hz must be lowered by 6dBFS . Since the results in both cases would be the same, both interpretations are equivalent.

However, I am inclined to adopt the first position, since it has an absolute reference point. Whereas the second graph and interpretation must be modified depending on the input. For this second scenario to work, our framework needs to be modified for special cases—that are highly unlikely to be found in the wild. And since that creates unnecessary inconsistencies and results in imprecision, the *time domain normalization* is not used by default.

Appendix J

Known Issues

I'm positive several deficiencies in my implementation still elude me. However here are a few I'm aware of and have not addressed yet.

- *PCM* generated by the *FM* chip is not covered in the current *Mega Drive/Sega Genesis* implementation. It is assumed the response will follow the *YM2612* curve.
- A better signal could be generated to cover either more or specific cases.
- When comparing *Mega CD/Sega CD*, the *CDDA* uses the detected frame rate for the comparison. It should use the expected one even if comparing to an *emulator/FPGA*. This causes slight ringing in the results, but they are at least 95% correct.

Please contact¹ me if you are aware of any details that have eluded me.

¹See appendix Q.

Appendix K

Window Function equations and graphs

This appendix lists the equation and curve of each *Window Function* used to limit *spectral leakage* as described in section B.1.1.

K.1 Tukey

The default is a *Tukey* window selected for this purpose. It uses $\alpha = 0.6$, zeroing just a few samples, with minimal *spectral leakage* and good amplitude response.

The following equation is used to create the slopes:

$$tukey(x) = \frac{1}{2} \left(1 + \cos\left(\frac{\pi(|x - \frac{N-1}{2}| - \alpha \frac{N-1}{2})}{(1 - \alpha) \frac{N-1}{2}}\right) \right) \quad (\text{K.1})$$

And this is the resulting graph of the *Tukey* window, ranges are *0.0* to *1.0* horizontally and *-0.1* to *1.1* vertically.

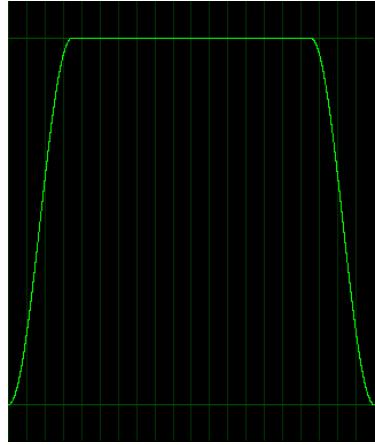


Figure K.1 Tukey window used by *MDFourier*, vertical lines are frames on a 20 frame signal.

Detailed information can be found in the reference webpage [25].

K.2 Hann

When selected, a typical *Hann* window is used. This should be used to get the last *spectral leakage*, with a very small trade off in amplitude accuracy.

$$hann[x] = \frac{1}{2}(1 - \cos(\frac{2\pi(x + 1)}{n + 1})) \quad (\text{K.2})$$

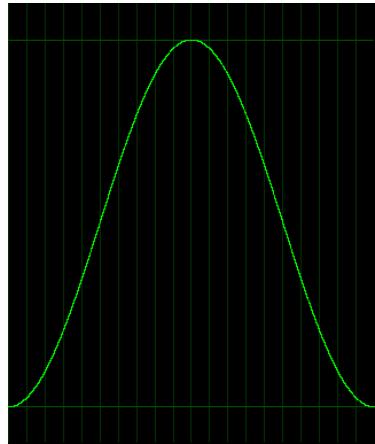


Figure K.2 Hann Window.

K.3 Flattop

A typical *Flat top* window is used, selecting this will target amplitude accuracy against frequency bin precision.

$$\begin{aligned} flattop(x) = & 0.21557895 - 0.41663158 \cos(2\pi \frac{x}{n-1}) + 0.277263158 \cos(4\pi \frac{x}{n-1}) \\ & - 0.083578947 \cos(6\pi \frac{x}{n-1}) + 0.006947368 \cos(8\pi \frac{x}{n-1}) \end{aligned}$$

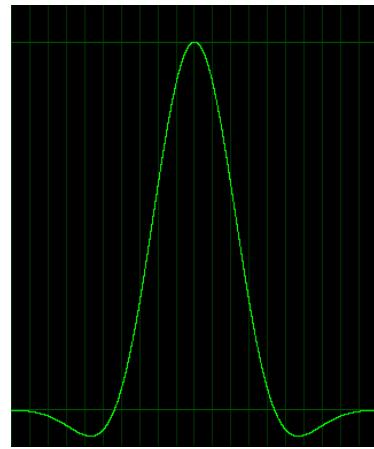


Figure K.3 Flat Top window.

K.4 Hamming

A typical *Hamming* window is used, presented for completeness and reference since the samples are never zeroed out.

$$hamming[x] = 0.54 - 0.46 \cos\left(\frac{2\pi x}{n-1}\right) \quad (K.3)$$

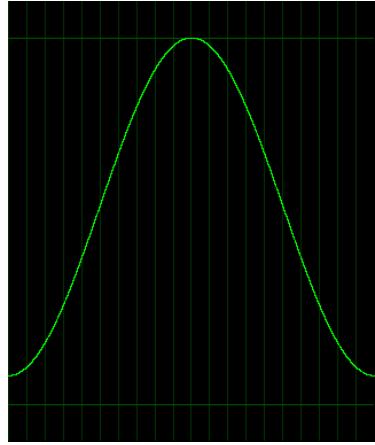


Figure K.4 Hamming window.

K.5 No Window

No window is applied to the signal before applying the *DFT*, equivalent to a *rectangular window*. This leaves the signal unprocessed and any uncontrolled decay and audio card noise will be factored in as part of the periodic signal.

There is more information on windows and their usage in the reference webpage [14].

Appendix L

Color Filter Function details

This appendix contains a description, graph and example of each *Color Filter Function* from section B.1.2.

L.1 None

No filtering is applied, as a result all differences are plotted with the brightest color.



Figure L.1 No Filter.

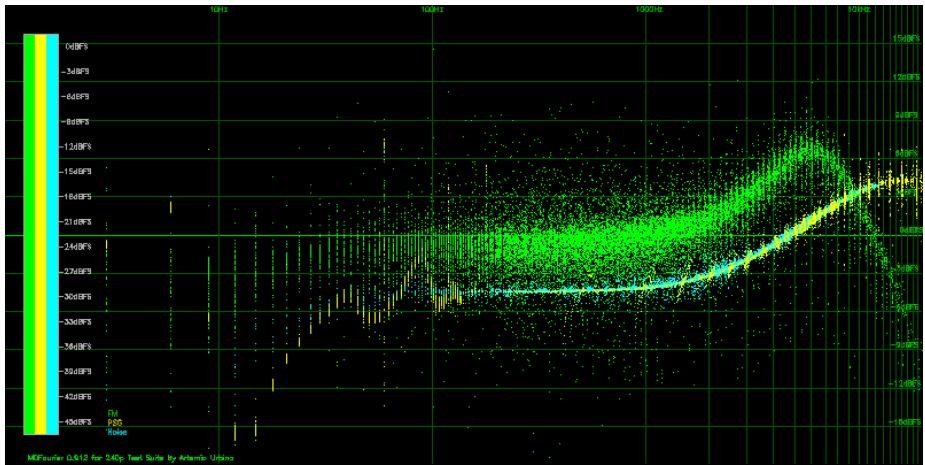


Figure L.2 No Filter Applied.

L.2 \sqrt{dBFS}

A square root function will only attenuate the lowest amplitude differences.

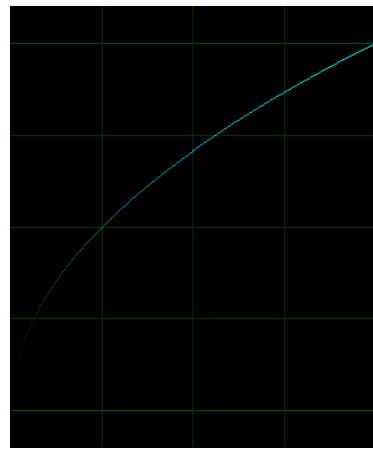


Figure L.3 Square Root filter.

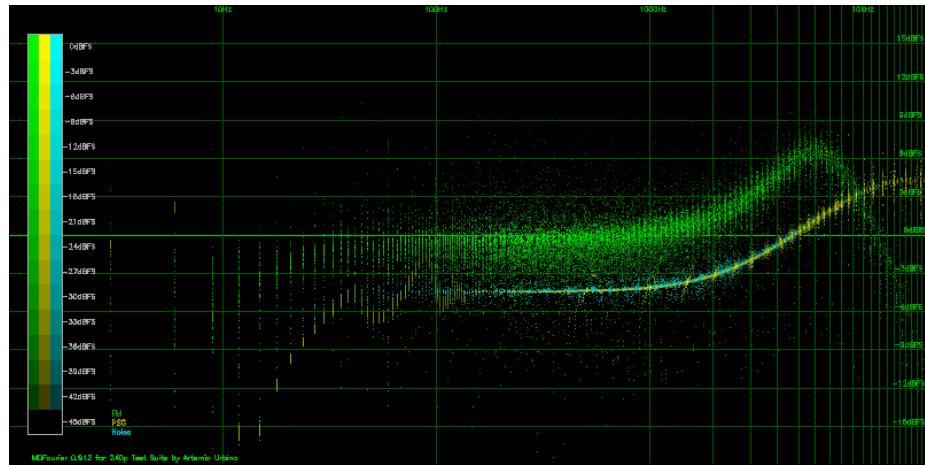


Figure L.4 Square Root filter Applied.

L.3 $\beta(3, 3)$

A Beta Function filter with parameters (3, 3) will attenuate a bit more from the lower range, still showing most of the differences.

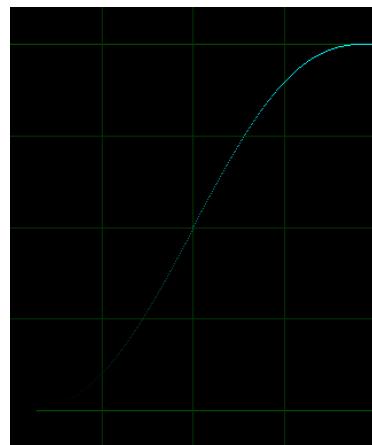


Figure L.5 Beta Function(3,3).

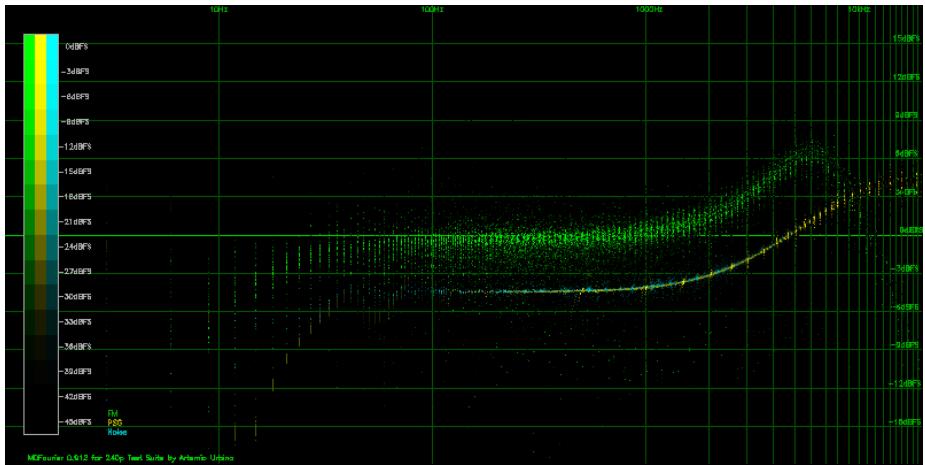


Figure L.6 Beta Function(3,3) Applied.

L.4 Linear

The linear function is the default, and has no bias. Half the dinamic range corresponds to half the color rage.

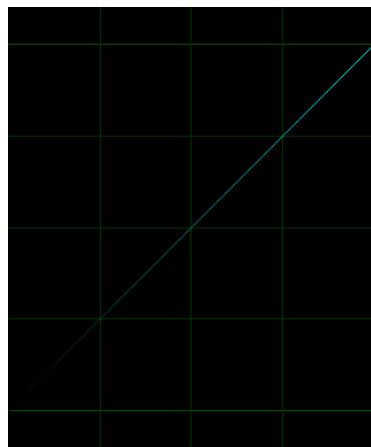


Figure L.7 Linear Function.

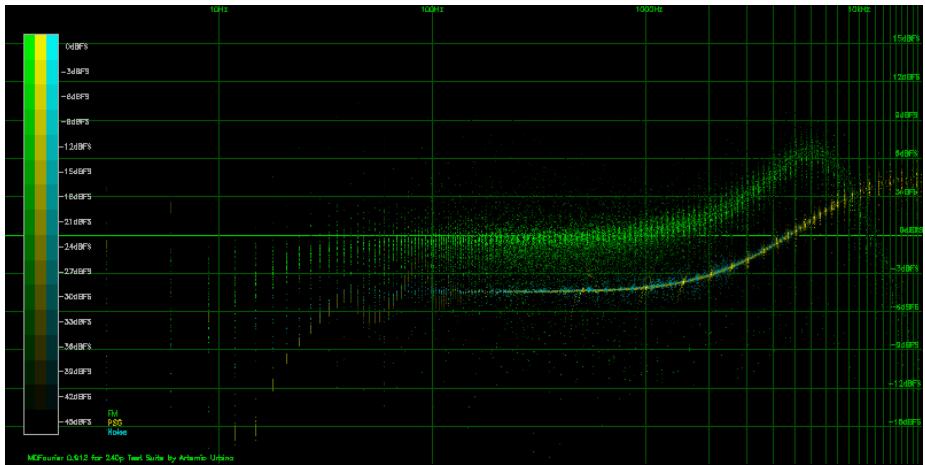


Figure L.8 Linear Function Applied.

L.5 $dBFS^2$

A squared function will attenuate a lot more differences, as a result frequencies with the highest amplitude in the reference signal will be brighter.

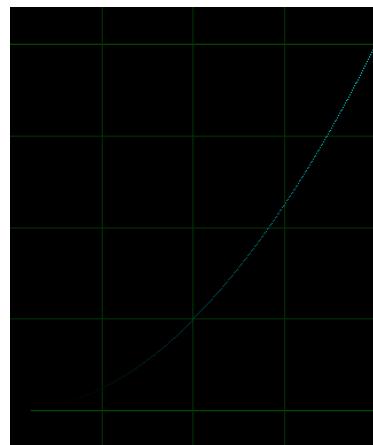


Figure L.9 Linear Function.

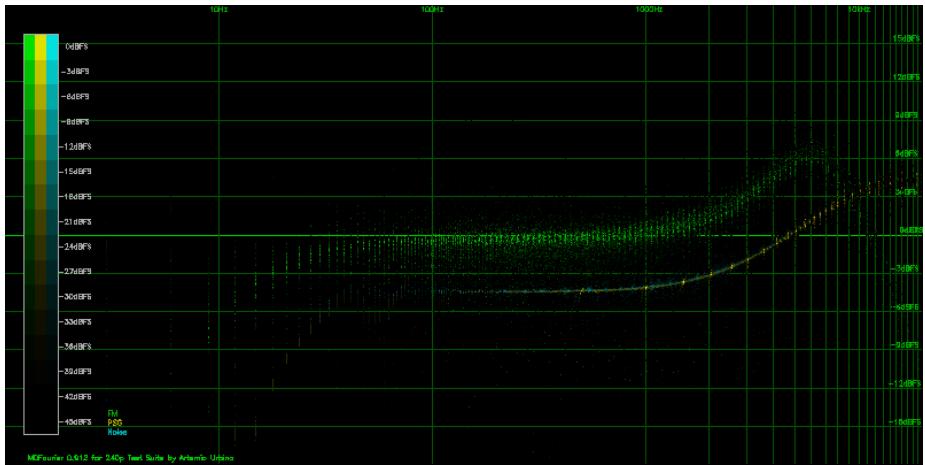


Figure L.10 Linear Function Applied.

L.6 $\beta(16, 2)$

A Beta Function filter with parameters (16,2) will attenuate almost all the differences, and only the frequencies with the highest amplitude in the reference signal will be brighter.

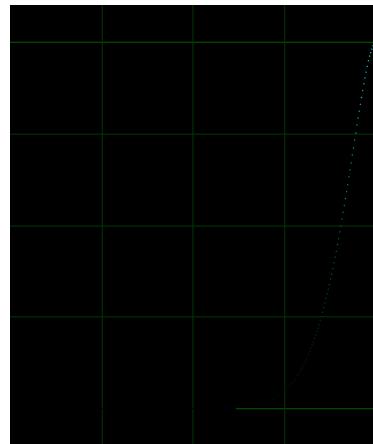


Figure L.11 Beta Function(16,2).

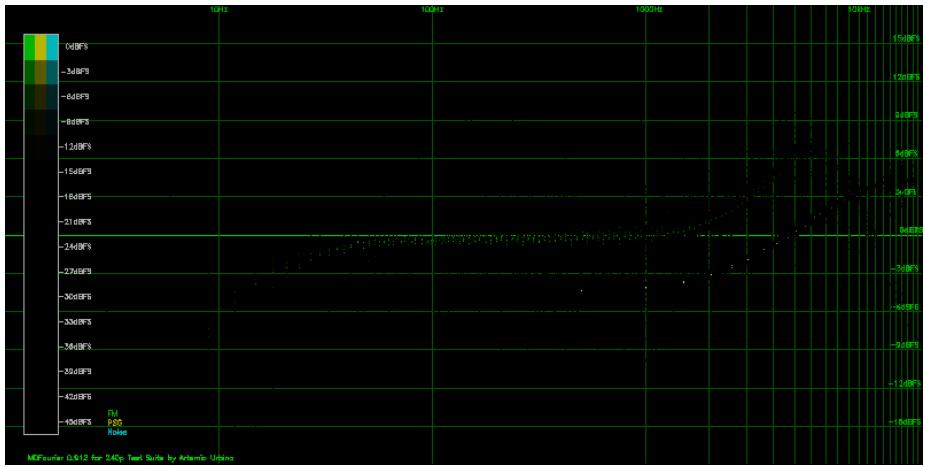


Figure L.12 Beta Function(16,2) Applied.

Appendix M

Requirements

M.1 Audio capture device

For capturing the audio files, an *audio capture device* is needed. It is recommended to use a musical grade *audio card* in order to get a *flat frequency* response across the human hearing spectrum. Fortunately, these devices have become more affordable.

So far four audio cards have been used in my personal setup. The sample recordings available for download¹ were made with three cards of them, a complete set with each one of them.

I have no association or business relationship with these products, they are just presented as working examples. As more people use the software and we—as a community—compare files, this list can be expanded with recommendations.

- **Lexicon Alpha:** An USB audio card².
- **M-Audio M-Track:** A USB audio card³.
- **M-Audio M-Track Pro:** A USB audio card, adds *Sony/Philips Digital Interface (S/PDIF)* over the regular model⁴. [28]
- **M-Audio Audiophile 192:** An internal audio card⁵ that is no longer available in the market [26]

¹See appendix A.

²Uses a CS4270 24-Bit, 192 kHz Stereo Audio CODEC and a TAS1020B for USB communications. Verified by opening the device. [27].

³Uses a CS5341 105 dB, 192 kHz, Multi-Bit Audio A/D Converter and a TAS1020B for USB communications. Verified by opening the device. [28].

⁴Not opened for component verification, but assumed to use the same components as the regular version.

⁵Uses a VT1720 Envy24PT PCI Multi-Channel Audio Controller.

Some cards marketed for podcasting and audio cassette digitization were tested⁶, but they didn't show a flat frequency response. You should try to use a sound card that is aimed to musicians or instrument recording.

M.1.1 Sampling clocks

Sound cards have their own internal clock for sampling⁷, which can sometimes deviate enough so that *MDFourier* can report *frame rate* differences. From my tests, this is more prominent when using the *44.1kHz* sample rate, and minimized when using *48kHz*. The effect is compensated for while loading the file and during trimming, so it shouldn't be a problem due to the small variations this produces in the frequency domain.

M.2 Computer

Any computer and operation system can be used if you are compiling the source code from scratch⁸, but a statically linked *Microsoft Windows executable* is provided for convenience, alongside a front end. See chapter B for instructions on using the *GUI*.

M.3 Game Consoles or emulators

You'll need either the provided example audio files⁹ or to create your own, by recording from your own source. This is probably the desired route, since you will want to compare against your files.

M.4 Flash cart, or means to run the binary

The console needs to run a *custom binary* that is provided with the rest of *MDFourier*. If possible this will be integrated into each version of the *240 test suite* [6].

In order to run these, you'll either need a *flash cart* or a custom loading solution compatible with the target platform.

⁶Gear of this type could exist with a *flat frequency* response, current technology should allow any product to offer this basic requirement. But alas, the ones we've tested failed in our tests.

⁷For more details, see [16] [17] [18].

⁸See appendix A for download links.

⁹See appendix A for download urls.

M.5 Cables and adapters

You'll need cables—and maybe some adapters—to connect the audio output from the console to the input of your *audio capture card*.

M.6 Audio capture software

Your *audio capture card* will probably be bundled with some audio editing software, or you can use *Audacity* [29] or *Goldwave* [30] depending on your operating system.

Appendix N

Compiling from source code

N.1 Dependencies

MDFourier needs a few libraries to be compiled. In *Linux*, *UN*X* based systems and *Cygwin* [24]; you can link it against the latest versions of the libraries.

- Fastest Fourier Transform in the West. (fftw) [7]
- The GNU plotutils package. [10]
- *PNG* Reference Library: libpng. [8]
- *FLAC* Reference Library: libFLAC. [9]

The following implementations are also used and included with the source files:

- sort.h for tim sort. [11]
- Incomplete Beta Function. [13]

The pre-compiled binary for *Windows* is created with *MinGW* [23] and statically linked for distribution against these libraries:

- fftw-3.3.8 [7]
- plotutils-2.6 [10]
- libpng-1.5.30¹

¹This older version was used to simplify the build process in this statically linked executable. Sources at <https://sourceforge.net/projects/libpng/files/libpng15/1.5.30/>.

The *makefiles* to compile either version are provided with the source code [5].

Appendix O

Colors available for graphs

This is the list of colors that can be used in the *MFN configuration file* described in appendix C:

- red
- green
- blue
- yellow
- magenta
- aqua
- orange
- purple

Appendix P

Licensing

MDFourier Copyright (C)2019 Artemio Urbina

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program. If not, see <https://www.gnu.org/licenses/>.

Appendix Q

Contact the author

You can contact me via twitter <http://twitter.com/Artemio> or e-mail me at aurbina@junkerhq.net.

Appendix R

Acknowledgments

I'd like to thank the following people for helping me so these tools could be completed. First and foremost to my family, who helped me find the time to learn and make progress.

Bibliography

- [1] Bracewell, Ronald N. *The Fourier Transform and Its Applications* (2 ed.). McGraw-Hill. ISBN 978-0-07303938-1.
- [2] Ace, *GUIDE: A complete overview of the many Sega Genesis/MegaDrive models*, Sega-16, <http://www.sega-16.com/forum/showthread.php?7796-GUIDE-Telling-apart-good-Genesis-1s-and-Genesis-2s-from-bad-ones>
- [3] The Best Version of the Genesis, RetroRGB <https://www.retrorgb.com/genesisversions.html>
- [4] Sega Mega Drive/Hardware revisions, Sega Retro, https://segaretro.org/Sega_Mega_Drive/Hardware_revisions
- [5] Github, *MDFourier source code C99*, <https://github.com/ArtemioUrbina/MDFourier/>.
- [6] 240p test Suite, *Wiki web page*, <http://junkerhq.net/240p/>.
- [7] Fastest Fourier Transform in the West., *web page*, <http://fftw.org/>.
- [8] libPNG, *web page*, <https://libpng.sourceforge.io/>
- [9] libFLAC, *web page*, <https://xiph.org/flac/>
- [10] GNU Plot Utils, *web page*, <https://www.gnu.org/software/plotutils/>
- [11] An implementation of a ton of sorting algorithms in C with a user-defined type, *web page*, <https://github.com/swenson/sort/>
- [12] *Human Hearing: Amplitude Sensitivity*, Mark Sanfilipo — April 04, 2005, <https://www.audioholics.com/room-acoustics/human-hearing-amplitude-sensitivity-part-1>
- [13] incbeta, *Incomplete Beta Function in C*, <https://codeplea.com/incomplete-beta-function-c/>.
- [14] Window Types: Hanning, Flattop, Uniform, Tukey, and Exponential , *web page*, <https://community.plm.automation.siemens.com/t5/Testing-Knowledge-Base/Window-Types-Hanning-Flattop-Uniform-Tukey-and-Exponential/ta-p/445063/>.

- [15] dB Full Scale, <https://www.sweetwater.com/insync/dbfs/>
- [16] Sound Card Sampling clock variation http://www.stu2.net/wiki/index.php/Calibrate_Sound_Card
- [17] So how accurate is a typical PC sound card? How stable is the output? How would one measure this? Experiments with a PC sound card by leapsecond <http://www.leapsecond.com/pages/sound-1pps/>
- [18] Each sound card has it's own sampling clock, Goldwave support forums <http://www.goldwave.ca/forums/viewtopic.php?p=17470>
- [19] D/A and A/D | Digital Show and Tell (Monty Montgomery @ xiph.org) <https://www.youtube.com/watch?v=cIQ9IXSUzU>
- [20] Harris, Fredric j. (Jan 1978). *"On the use of Windows for Harmonic Analysis with the Discrete Fourier Transform"*. Proceedings of the IEEE. 66 (1): 51–83. <https://web.mit.edu/xiphmont/Public/windows.pdf>
- [21] Spectral Leakage and Zero-Padding of the Discrete Fourier Transform <https://dspillustrations.com/pages/posts/misc/spectral-leakage-zero-padding-and-frequency-resolution.html>
- [22] Why is it a bad idea to filter by zeroing out FFT bins? <https://dsp.stackexchange.com/questions/6220/why-is-it-a-bad-idea-to-filter-by-zeroing-out-fft-bins>
- [23] MinGW, *Minimalist GNU for Windows*, <http://mingw.org/>.
- [24] Cygwin, *a large collection of GNU and Open Source tools which provide functionality similar to a Linux distribution on Windows*. <https://www.cygwin.com/>
- [25] Tukey window, Recording Blogs <https://www.recordingblogs.com/wiki/tukey-window>
- [26] M-Audio Audiophile 192, *Specifications*, <https://www.soundonsound.com/reviews/m-audio-audiophile-192/>.
- [27] Lexicon Alpha USB card, *Product web page*, <https://lexiconpro.com/en/products/alpha/>.
- [28] M-Audio M-Track USB card, *Product web page*, <https://www.m-audio.com/products/view/m-track>.
- [29] Audacity *web page*, <https://www.audacityteam.org/>.
- [30] Goldwave *product web page*, <https://www.goldwave.com/>.
- [31] Adam Hayes, Simple Moving Average - SMA Definition <https://www.investopedia.com/terms/s/sma.asp>
- [32] Human Speech Spectrum, Frequency Range, Formants, *web site*, <http://www.bnoack.com/index.html?http&&www.bnoack.com/audio/speech-level.html>.

[33] Noise Level Chart, *web site*, <https://www.noisehelp.com/noise-level-chart.html>

[34] Audio Spectrum Explained *web site*, <https://www.teachmeaudio.com/mixing/techniques/audio-spectrum/>.

Glossary

ASIC Application-specific integrated circuit. 17, 37

CDDA Compact Disc Digital Audio. 72

dB decibel. 10–12

dBFS decibels relative to full scale. 4, 12, 16, 18, 19, 21, 23, 24, 26, 28, 30, 31, 34–38, 40–42, 52, 60, 61, 69–71, 78, 81

DFT Discrete Fourier Transform. 15, 20, 50, 51, 76

FLAC Free Lossless Audio Codec. 14, 87

FPGA Field-programmable gate array. 1, 6, 7, 56, 63

GUI Graphical User Interface. 13, 47, 49, 53, 59, 85

Hz Hertz is the derived unit of frequency in the International System of Units (SI). It is defined as one cycle per second. 9, 11, 14, 16, 18, 21, 22, 28, 31, 34, 38, 40, 51, 52, 63, 69–71

kHz Kilohertz, refers to 1000 Hertz. 9, 11, 14, 18, 23–25, 28–31, 34, 37, 38, 41–43, 55, 62, 69, 70, 85

MP3 MPEG-2 Audio Layer III, a popular coding format for digital audio. 14, 29–31

ms Millisecond ($\frac{1}{1000}$ of a second). 28, 56, 63

NTSC National Television System Committee, a colour encoding system for analogue television. 16, 62–64

PAL Phase Alternating Line, a colour encoding system for analogue television. 16

PCM Pulse-code modulation, a method used to digitally represent sampled analog signals. 14, 55, 72

PNG Portable Network Graphic. 17, 52, 87

S/PDIF Sony/Philips Digital Interface. 84

WAV Waveform Audio File Format, also referred to as WAVE. 14, 17, 19, 29, 30, 55–57, 60