FORMI: A NEW AGE OF MUSIC LISTENING

By

Drew Stogner

Kshithij Kamal

Final Report for ECE 445, Senior Design, Fall 2019

TA: Jon Hoff

11 December 2019

Project No. 28

Abstract

The device described within this report is a proof of concept Bluetooth speaker system that adapts to the music being played. The hardware portion is designed for a wide array of genres or, rather, as a "blank slate" to adequately showcase the music augmentation from the software. The software augments music for enhanced listening experience utilizing meta-data and extracted singer vocal characteristics which are paired with a realtime music/EQ augmentation program. By changing the equalization of the music, the device changes the sound signature or in colloquial terms, the feel, of the music. Within this document, there is a study of the design process as well as the verification of our device. Furthermore, we will talk about both technical and project management success and failures.

Contents

1. Introduction

1.1 Purpose for Project

With technology constantly advancing, music has become available to the masses in numerous forms. In fact, music has grown to become an imperative part of people's lives with Americans spending on average 4.5 hours per day listening to music. [1] For over hundreds of years, it has been used as a stress reliever yet each modern speaker, headphone, etc. optimizes for a limited range of music dictated by its sound signature.[2] [3]

Sound signature emanates from two elements:

1. Programmed equalization in the DACs (Digital to Analog Converters)
2. Circuit Components

Speaking with musicians, producers, other audiophiles and music lovers, there is a common problem of how to best enjoy music. With every headphone having different sound signatures, each one works best for a specific genre or use case and may even cause fatigue when matched with the wrong song. In order to enjoy a wide and varied range of music genres, people must end up buying multiple high-end headphones. This becomes extremely cost ineffective with the price of high-end headphones ranging from eighty dollars to thousands of dollars. The problem becomes more poignant if you compose or record music for a living. Modern pop(ular) music is incorporating multiple genres into one. ex: With artists like Lil NasX, Billy Eilish, and Twenty-One Pilots all in one genre, the differences in sound can be quite extreme.

To target this problem, niche high-end headphone companies are introducing adjustable EQ settings allowing musical experts to make minute sound adjustments. Our solution advances on this innovation by optimizing from the ground up. Starting with the circuit components, we built for adaptability and utilized a baseline/neutral DAC equalization setting. Concurrently, we created a computer application (with Bluetooth connectivity) that augments equalization based on the music playing. For example, it would analyze a song and make certain EQ parameter decisions based on the stored genre of the song, and the vocal range of the singer. [4]

1.2 High Level Requirements

1) The Hardware device must take musical data through a bluetooth signal and convert it into an analog signal that can be played on the POC speaker,

2) The software must be able to obtain file embedded (basic) genre and ML derived genre/characteristics at 70% accuracy

3) The software must exhibit an equalization change based on song genre/characteristics through means of a dB vs frequency augmentation (must be transferred to our device).

1.3 Introductory View of High Level Design

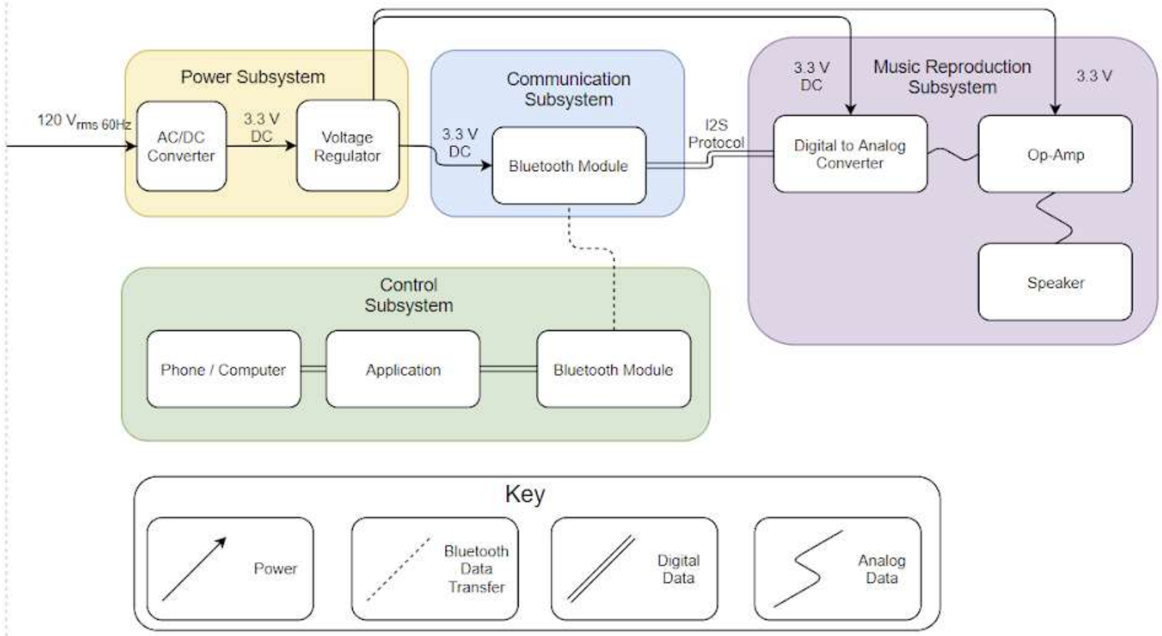


Figure 1: Block Diagram

The Design is split into four distinct subsystems, three of which are hardware and one is software.

The power subsystem is responsible for taking a 120 V AC input at 60 Hz and convert it to a 3.3 V DC +/- .1 V output to power the other hardware subsystems. Previously, this was 5V DC but was changed as a result of a Bluetooth module part swap. There is also the communication subsystem that is comprised of the Bluetooth module that is responsible for receiving the digital sonic data from the third party device, a phone or computer, and output using I2S data bus to the digital to analog converter. The next and final subsystem for hardware is the music reproduction subsystem. The digital to analog converter takes the digital signal from the I2S data bus and outputs a stereo analog output. The class D operational amplifier takes the stereo analog output and then amplifies the signal by an order of 50 dB.

2 Design

2.1 Power Subsystem

2.1.1 Design Procedure

To accomplish the task of creating an AC – DC voltage converter from 120 V AC at 60 Hz to 3.3 V DC, there are multiple solutions. Starting with the least efficient converter, we could use a transformer to step down the input voltage and then rectify using a full-bridge rectifier. Then use capacitors to lower voltage ripple or in more colloquial terms smooth out the rectified AC signal to a signal closer to typical DC. This allows for a Zener diode to be placed with a breakdown voltage of 3.3 V DC for the output. The use of Zener diodes would cause massive efficiency problems, but the benefit to this design is cheap cost for production. The design that we ultimately went with was very similar to the previous design with the exception that instead of a Zener diode, we used a linear voltage regulator to produce the 3.3 V DC. This is more expensive, but also more efficient. Another formidable approach would have been utilizing a Boost active rectifier in contingency with a Buck converter. This is ultimately the most efficient and cheapest solution, but the circuit complexity is far greater. Concurrently, this solution would utilize a much smaller physical space due to the smaller component sizes. For future design, we would use this method.

2.1.2 Design Details

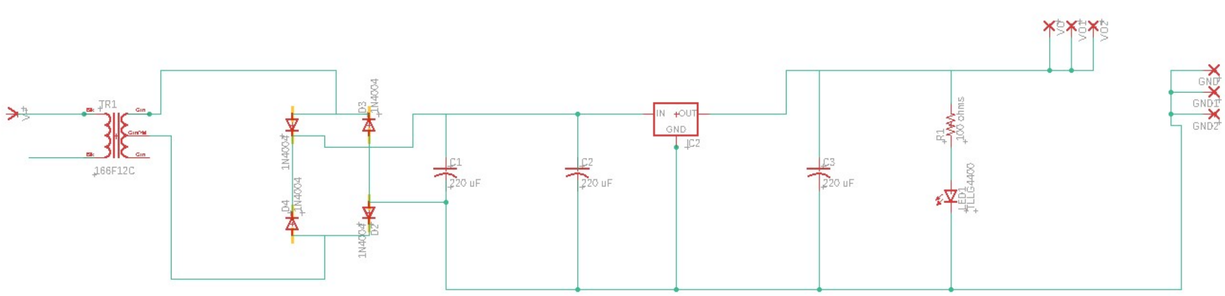


Figure 2: Power Subsystem Schematic

The power subsystem, as seen in figure 2, transforms 120 V AC at 60 Hz into 3.3 V DC to power the rest of the subsystems. All the V_{in} and Ground ports to each subsystem come from this subsystem. With this design, we first had to step the voltage down from the wall outlet voltage at approximately 120 V AC at 60 Hz frequency. Our solution was to use a 115:12 voltage ratio transformer.

$$1.\ V_{out} = \frac{12 * V_{in}}{115}$$

This gives us an output of 12.5217 V AC RMS.

$$2.\ V_m = V_{rms} * \sqrt{2}$$

With a sinusoidal output from the transformer with an amplitude of 17.7084 V, the next step was rectification of the sinusoidal input. we used a full-bridge rectifier that would allow the positive voltage to go to the top leg of the output, while nullifying the bottom leg of the circuit so that the bottom leg of the circuit would act as ground. Following this, we had to convert this new AC signal into a DC one with limited ripple to ensure the linear regulator chip would function properly. To ensure this, we placed two capacitors in parallel with the voltage regulator. [2]

$$3.\ V_{pk-pk\ ripple} = 4 * I_{load} * f * C$$

With the above equation, we can now solve for the capacitance by setting I_{load} to .6 A, $V_{pk-pk\ ripple}$ to 2 V and frequency to 60 Hz. We get a value for capacitance of 1.25 mF. It is important to note that this amount of capacitance is excessive due to the functionality of the voltage regulator. If the input voltage is regulated between 8 V and 25 V, the circuitry will function as intended. To ensure this, we utilized 440 µF total capacitance. Next, we knew we had to regulate the ~17 V DC voltage down to 3.3 V DC so that we could power other subsystems. Hence, we added a 3.3 V DC and 600 mA voltage regulator. The output side of the regulator is connected a capacitor in parallel with a 100 Ω resistor in series with a LED to show that the circuit is functioning.

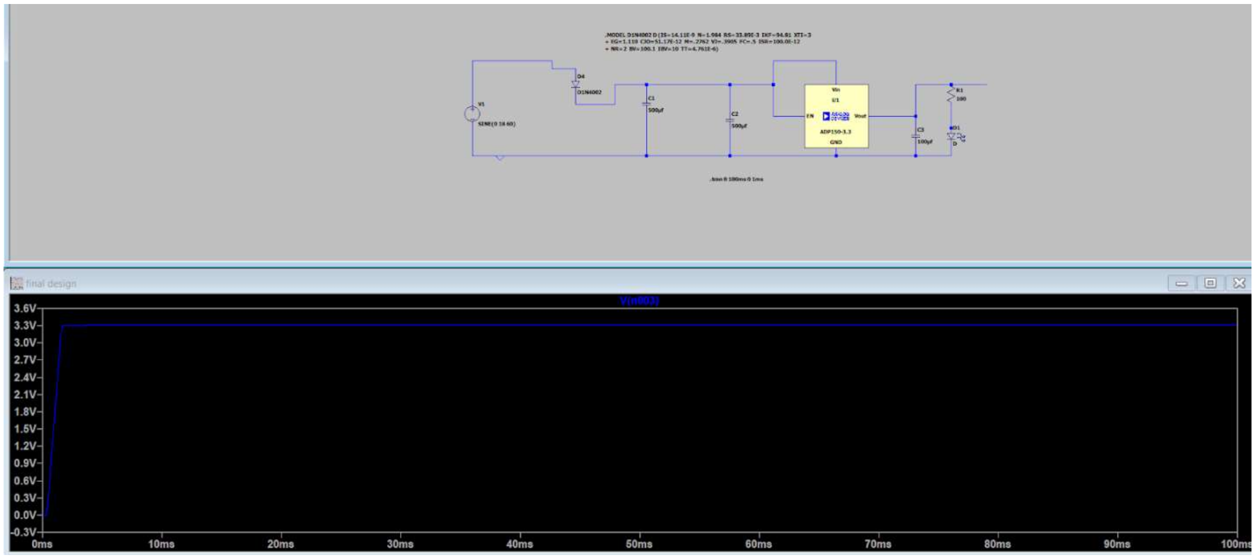


Figure 3: Simulation for Power Subsystem

For this simulation in figure 3, we had to change the schematic a bit for the program to run it. we assumed an ideal transformer and replaced it with an already stepped down sinusoidal voltage source. As well, the program we used demands that a ground be placed before it simulates, so we used a half-bridge rectifier instead of the full bridge. This would increase the amount of capacitance needed for the ripple constraints, so due to our overdesign for capacitance, this wouldn't be a problem. As can be seen in the voltage vs. time graph, the simulation shows a correct functionality in that the output voltage is regulated to 3.3 V DC.

2.2 Bluetooth and Digital to Analog Converter

2.2.1 Design Procedure

These two components are linked by a certain type of data bus. It is important to note that any Bluetooth module used for this application must have a digital to analog converter with the same type of data bus. There are four types of data buses typically used for Bluetooth modules: I2S, I2C, SPI and UART. Starting with I2S, the key point is that data is outputted to a synchronous clock and word select that communicates with the digital to analog converter. We chose the Bluetooth and digital to analog converter with this data bus due to the synchronous nature of this type. This allows for reliable communication and this is important for musical and sonic applications. As well, this is a common type of data bus used for those types of applications. I2C is like I2S except I2C is asynchronous meaning that the Bluetooth and digital to analog converter must agree on frequency for the clock else data loss is inevitable. This makes the design process unnecessarily more difficult by having to create a clock to feed into both whereas in I2S the clock is generated by the Bluetooth module and feed into the digital to analog converter. UART is asynchronous as well, so it has the same drawbacks as I2C. UART also has an abysmal data transfer rate and thus for real time sonic data rate would not be well suited. SPI is another synchronous data bus and well suited for sonic applications but is much more complex than I2S due to the number of output pins used for SPI tripling the amount used in I2S. For most applications for music, one should either choose SPI or I2S, so we chose I2S on the basis of simplicity of design.

2.2.2 Design Details

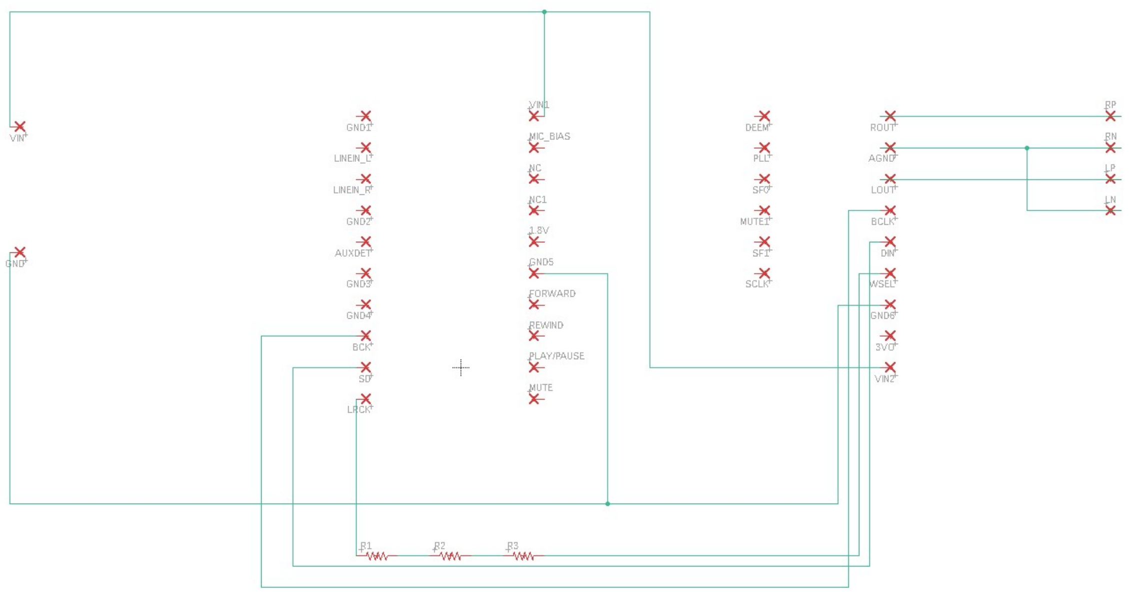


Figure 4: Bluetooth and Digital to Analog Converter Schematic

As can be noticed in figure 4, there are two modules used in our design. The Bluetooth module, AudioB I2S Bluetooth Digital Audio Receiver Module, and the digital to analog converter, Adafruit I2S Stereo Decoder - UDA1334A Breakout, were chosen due to several reasons. [10][16] The first and most important is that they share a common data bus, I2S. The second reason is that both are powered by 3.3 V DC power source and thus can be powered by the power subsystem. This specific digital to analog converter was chosen because it outputs the data in stereo. This means that musical data is outputted in analog for speaker left and speaker right thus allowing for separation of sound. As well, the digital to analog converter module came with an auxiliary port that allows for simple verification which added to our decision for this specific component as a third party powered speaker could be used for verification of the Bluetooth and digital to analog converter.

2.3 Class D Operational Amplifier

2.3.1 Design Procedure

Although there are many different integrated circuits for class D operational amplifiers, they all essentially work with the same principle in mind. The class D operational amplifier is the amplifier used for musical applications meaning that other amplifiers simply won't work or suffice for our application. The main design choice used here was choosing to use two mono application based operational amplifiers instead of a single stereo based one. They are around the same cost and both can be used equally well in our application. The stereo based operational amplifier will save space but will be more complicated in implementation. The opposite is true for the two mono based operational amplifiers. We deemed the best choice was the simplest.

2.3.2 Design Details

The Class-D amplifier circuit in figure 4 was designed with two separate Class-D amplifier ICs because each IC was designed for mono applications.

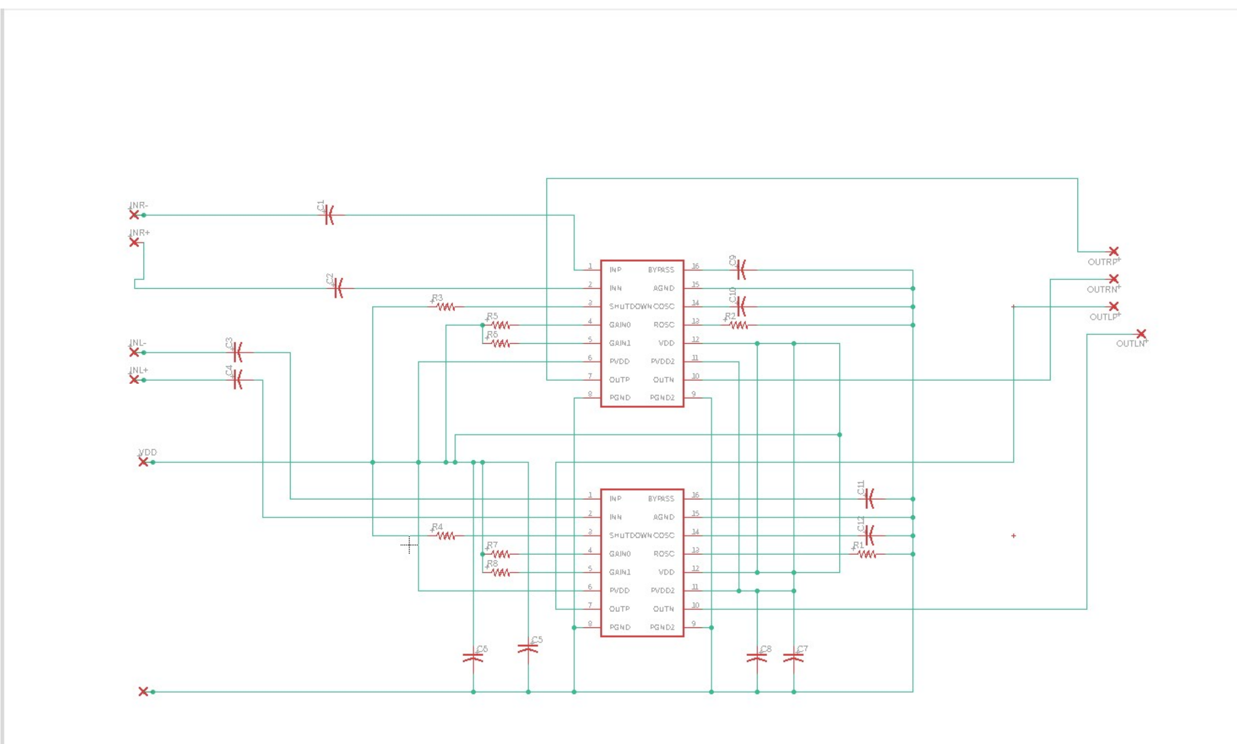


Figure 4: Class-D Operational Amplifier

[5] Essentially, we used the data sheet to figure out the possible max gain, 50 dB, so that the signal would be boosted the highest amount. This was done by putting equal resistance in series with Gain0 and Gain1 so that they would both be high and thus the gain would be maximized according to the datasheet. Since the digital to analog converter outputted stereo signals, we knew that each op-amp would get either the left or right analog signal. As well, we knew that Class-D op-amps invert their signals and thus as an example the INR+ signal would go to the IN- port on the bottom IC and INR- would go to IN+. As well, placing capacitors in series with each of the inputs analog signals would block out unwanted noise.

$$4.\ I = C\frac{dV}{dt}$$

As can be seen in the equation above, if there is no change in the voltage of the analog signal, then no current will flow and this will block out V DC bias noise, a baseline voltage that will cause noise if carried through to the speaker and especially if amplified. Similarly, the other capacitors connected in parallel with PVdd are used for noise reduction as well. V_{in}, the 3.3 V DC from the power subsystem is connected to the pins according to the requirements set by the data sheet. This goes the same for the ground input. The four outputs, two of which will go to each speaker and these carry the fully amplified sound.

2.4 Software Subsystem

2.4.1 Design

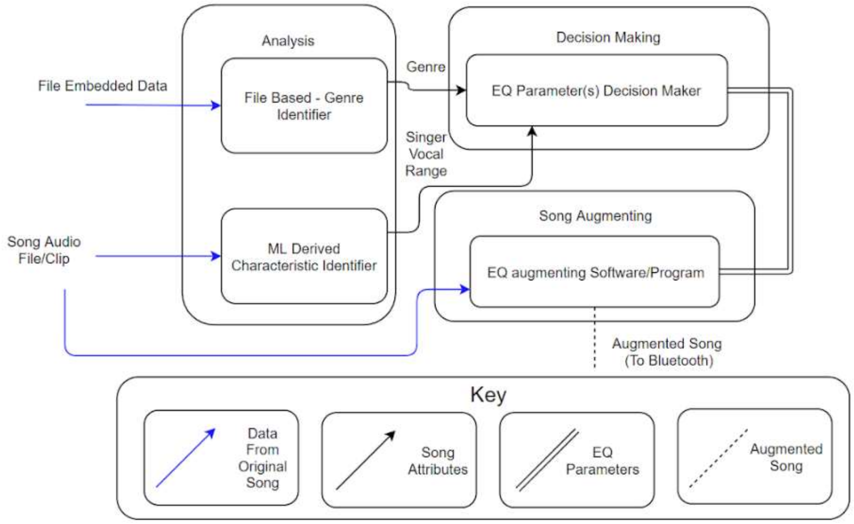


Figure 5: Software Block Diagram

The software portion of the project is best showcased in figure 5 (above). Using the song genre metadata and ML prediction of singer Vocals, the program makes and executes on EQ decisions to augment the song in realtime. The ML extracts 25 features from the song so it can predict the vocal characteristics of a singer. The 25 features can be summed up as follows:

1. Chroma Frequencies: Chroma Frequencies are generally used when training music-related ML algorithms or music analysis. It separates the frequencies into 12 segments (representing the musical octave).
2. Zero-Crossing Rate: [14] The Zero crossing rate seems like it would be very useful as it is used in many speech recognition projects. It is a simple calculation of the number of times the audio signal's sign changes.
3. Spectral Centroid: It is a weighted calculation of the frequency range.
4. Spectral Bandwith: Looks at the range of frequencies across the whole song clip
5. Spectral Rolloff: The spectral roll-off can be used to find the "roll-off" frequency at each frame of the song
6. (20) Mel-frequency Cepstral Coefficients: Mel-frequency Cepstral Coefficients are derived from an audio signal's spectral envelope shape. It is widely used in speech recognition projects, so it was very promising.

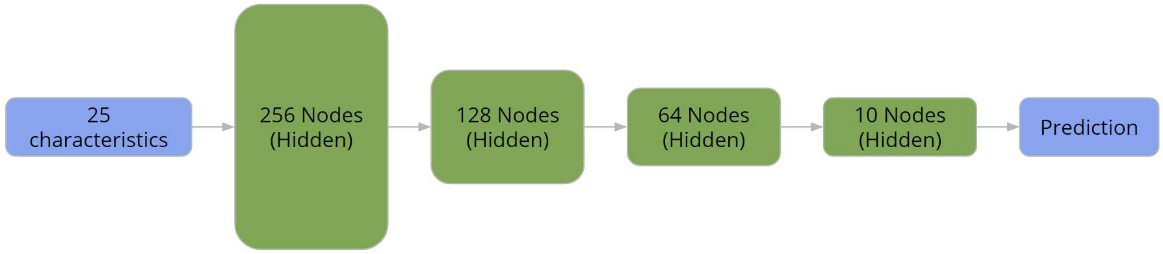


Figure 6: Neural Net Design

Using these 25 features, we used a Neural Net with 4 hidden layers to make a prediction on the vocal type of the singer. While we are not limited in time when extracting song features to train the algorithm. We are limited in time when someone wants to play a new song with our project. Every feature used to train the ML algorithm needs to be extracted from a song that we want to play so the ML algorithm can make decisions. However, I found many of the above feature derivations to be time-consuming. Making our goal of 0.7 seconds for ML decision making impossible.

Initially, the time to extract features and predict on a song took ~90s. To combat this, we worked with music producers to allow the ML algorithm to make predictions much faster. The key goal was to extract the longest song clip that contained vocals yet did not exceed the 0.7s prediction time goal. We figured out that taking a 12 second clip 30 seconds into the song gave the best results and never exceeded the 0.7 second threshold.

The song augmentation or Realtime EQ program was based on an open source EQ program. Not only did we rebuild many facets of the backend from adding stereo (rather than mono) playback capability, we modified and addended the program so it functions seamlessly with our EQ decision program, genre meta data extractor and ML vocal predictor. [17] Using the 3 band EQ, the decisions for the low and high frequency EQ bands were based on genre and the Vocal prediction determined the mid band.

3. Design Verification

3.1 Power subsystem

The verification was completed by plugging the power subsystem into a standard wall outlet and a multimeter was connected to the output to measure output voltage. Resistance was changed linearly from 0 to 1000 Ω in increments of 200 Ω in order to ensure load independence.

Table 1: Power Subsystem Verification Results

Resistance (***Ω***)	Output Voltage (V)
0	3.269
200	3.256
400	3.281
600	3.272
800	3.267
1000	3.273

These results were consistent with our requirement of error less than .1 V DC. As well, there appears to be no load dependence for the output voltage. The low error rate is showcased in graph 1 found in appendix C.

3.2 Bluetooth and Digital to Analog Converter

Using a phone or computer and an application that produces sinusoidal signals to be sent to the Bluetooth and digital to analog converter want can then use a powered speaker plugged into the auxiliary port on the digital to analog converter to play the pure tone. Using a musical tuner, we can measure the cents, or distance between each note. This tells us if the conversion is lossy or out of tune.

Table 2: Bluetooth and Digital to Analog Converter Verification Results

Input Frequency (Hz)	Output Frequency (Hz)	Musical Cents Difference
440	440	0
660	660	0
880	880	0

Shown in our results in table 2, there was no loss in frequency of the tone. This is important because loss in music is unacceptable due to the transmission especially in terms of frequency.

3.3 Class D Operational Amplifier

Using a function generator to produce sinusoidal signals to send into the operational amplifier and powering it with a separate DC power supply. Then by using an oscilloscope we can measure average output voltage. This will allow us to find the gain of the amplifier.

Table 3: Class D Operational Amplifier

Input Frequency (Hz)	dB Gain Measured
440	49.1
660	49.2
880	49.1

We expected a gain of 50 dB and from the results shown in table 3, we can see an approximate error of 2%. This is acceptable and in totality our circuit functioned and was able to reproduce music sent from a computer or a phone and play it through the speaker. There was an issue with this section of our design. If volume or amplitude of the signal was raised beyond a certain point for one of the two operational amplifiers, there existed clipping of the system. Several reasons could be the cause of this, but the reason that we expect to be the most likely is that these integrated circuit chips were unreliable as many had quirks in functionality. In future design, we would likely redesign this section and avoid using the Texas Instruments, *1-W FILTERLESS MONO CLASS-D AUDIO POWER AMPLIFIER.* [11]

3.4 Meta Data Extraction

Works for Flac,WAV, MP3

```
get_genre('Data_111019/music/02 - too late.flac')
3]: 'R-clsc'

get_genre('wonderfult.wav')
4]: 'R-clsc'

get_genre('audio/All Falls Down.mp3')
5]: 'Pop'
```

Figure 7: Meta Data Example

The Meta Data extractor that we created successfully works for flac, wav and mp3 file formats. Above you can see a test run on all three file types.

3.5 ML Accuracy

Table 4: ML Accuracy Results

Accuracy of Predictions (%)

High Voice	**79%**
Low Voice	**94%**
Weighted Average	**90%**

As seen above, the ML prediction accuracy was above our goal of 70%. The program tends to better predict low voices. However, the overall accuracy was 90%.

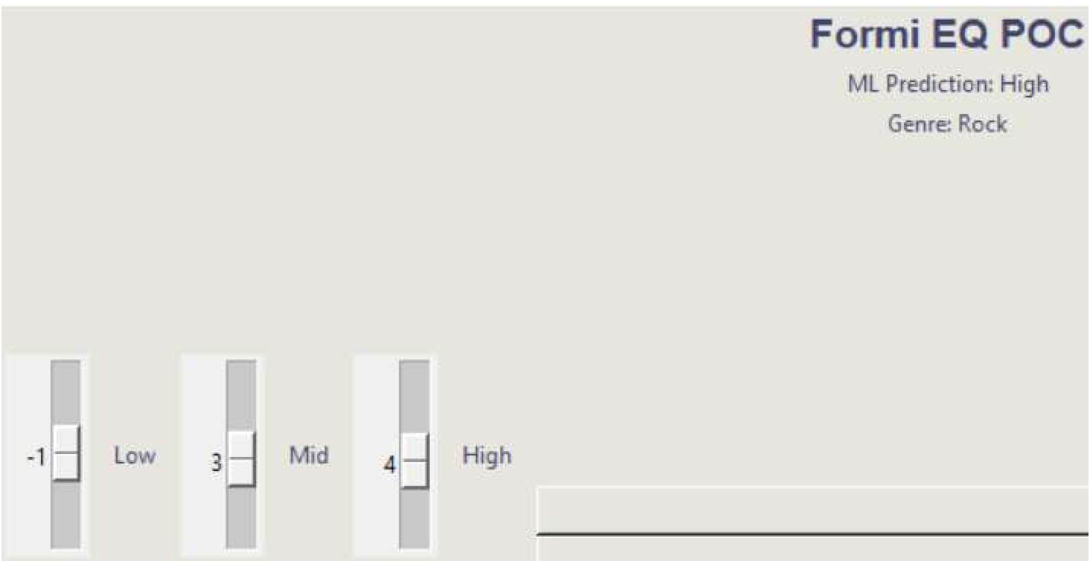
3.6 ML Latency

```
0.6173481941223145 seconds
The prediction for this song is:  Low (should be low)
0.6337363719940186 seconds
The prediction for this song is:  High (should be high)
```

Figure 8: ML Latency Example

As mentioned previously, we worked with producers to reduce our initial time latency of up to 90s. With the revised method, we never exceeded 0.7s. Above, is a iteration of our testing showcasing the time to extract features and make ML predictions on two songs.

3.7 EQ Change



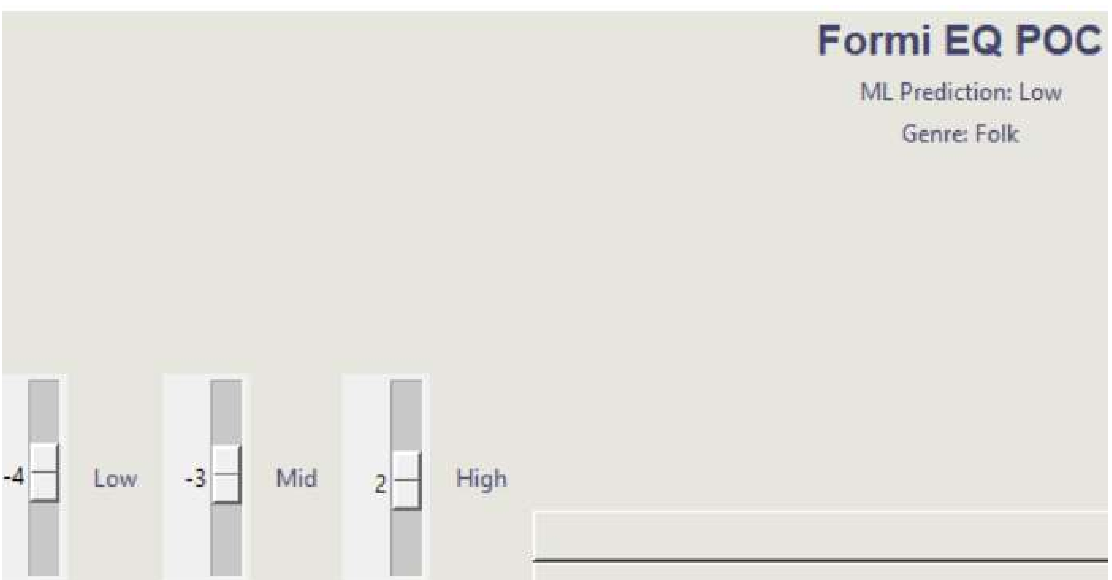


Figure 8: EQ Augmentation Example

As you can see in the above screenshots, the EQ changes for genre and vocal type. The genre determines the low and high frequency EQ bands. The Vocal prediction determines the mid band.

4. Cost

Starting with the assumption that a competitive salary would be approximately $35/hr and that we spent approximately 15 hours per week on this project individually, an approximated labor cost can be calculated as such:

5. 2 People x 2.5 x $35/hr x 15 hrs/week x 16 weeks

= $42,000

The total cost for our labor for this project can be estimated to be $42,000. Comparing to this the average value for an electrical engineering graduate at $67,000 and calculating a yearly salary for this project, we would get $109,000 for labor alone. [12] This is greater than the typical starting salary, but given that we are the engineering, marketing and research team, we feel as though this is a fair salary. Adding this with the cost of parts as shown in table 4 in appendix B. The entirety of cost for this project is $109,201.82.

5. Conclusion

5.1 Accomplishments

This project was a resounding success in the hardware system function as intended. The hardware successfully passed all our requirements and although in some areas didn't suffice to our satisfaction, such as the operational amplifier clipping above a certain volume or noise from soldering done poorly, we ultimately left with a product with that was to our satisfaction. The software generated ML exceeded our expectation in not only greatly exceeding the success rate that we had set for ourselves in our requirements, but also in the processing time needed for the software to run. The software enabled a definite audible change from a flat equalization, or no equalization, to the equalization set by the software. Although, enjoyment of music is subjective and ultimately left to the listener, both of us enjoyed the equalized versions better than the flat ones. This ultimately has the potential to change the way that many listeners will interact with music and improve the experience of something that many people already love so much.

5.2 Uncertainties

The main uncertainty that exist is that of reproducing our own hardware system. Ultimately, the hardware system although functional is not exceptional and many products that already exist perform better with far less noise and far more reliability. This is due to reasons previously mentioned such as the faulty operational amplifier functionality. Also, size would be a problem in practical usage. Our design if far too cumbersome for most consumers especially due to the small nature of competing products. The use of a wall outlet to power the hardware was also a poor choice in retrospect due to the fact that it eliminates a key benefit of Bluetooth technology, mobility. For all these reasons, we remain uncertain of the prospects of the hardware system.

5.3 Ethical considerations

There are numerous potential safety hazards that we may face when executing this project. There is some risk involved with using wall power since the voltage coming from the wall is 120Vrms. [8] To ensure safe practice, someone else should always be present in the laboratory when utilizing wall power. Accordingly, having a TA check the circuit prior to plugging in or powering the circuit, should be adhered to. The one-hand rule can be useful in ensuring that the person working on the circuit is never the quickest path to ground. [ACM 1.2] There is a potential fire hazard if we raise the decibel levels through volume or EQ such that clipping occurs. We should place hard upper limits on the decibel levels in the code for the equalization. Listen for signs of clipping or distortion in the noise being produced by the speaker and act appropriately such as powering off the circuit. As well, in case of emergency, a protocol for the laboratory fire emergency must be followed. Listening to loud music for long periods can damage eardrums. [5] Use ear protection when nuanced listening is not imperative (this is for internal testing). Or limit long periods of listening to music.

While the ethical concerns are limited in the scope of this project, there are some key ones to point out. [13] Mismatching EQs with Songs or distorting music can cause fatigue or dizziness. Do not create EQs purposefully to cause harm. Try not to pair songs with non-conforming EQs. Allow for plenty of breaks if the operator starts to feel fatigued or dizzy.

5.4 Future work

We remain committed to continue the work that we have started with this project. With more focus on the software side, we will focus on expanding the machine learning to not only dynamically change EQ depending on pitch of the voice, but with many more qualities of music. These could include rhythm, tempo, instruments used, etc. Developing a more well design user interface would also be a key goal as well as further improvements to both accuracy of the decisions made by the ML and lowering the process time. As well, this could either be partnered with an existing speaker or headphone producer or be made as a third-party application. Truly the expanse of the future work is endless, and we very much look forward to see how expansive we can make this project.

References

[1] We Listen to Music For More Than 4 1/2 Hours A Day, Nielsen Says. (2017, November 13). Retrieved September 9, 2019, from https://www.marketingcharts.com/industries/media-and-entertainment-81082.

[2] Releasing Stress Through the Power of Music. (n.d.). Retrieved September 30, 2019, from https://www.unr.edu/counseling/virtual-relaxation-room/releasing-stress-through-the-power-of-music.

[3] Kisla, R. (2017, August 1). Sound Signatures Explained. Retrieved September 30, 2019, from https://www.headphonezone.in/blogs/audiophile-guide/sound-signatures-explained.

[4] Newman, M. (2019, April 9). Country or Not? Lil Nas X's 'Old Town Road' Raises Big Questions About Genre in the Streaming Age. Retrieved September 30, 3n.d., from https://www.billboard.com/articles/business/8506428/lil-nas-x-old-town-road-splits-country-music-industry-opinion-genre-streaming-age.

[5] Hearing loss and music: MedlinePlus Medical Encyclopedia. (n.d.). Retrieved from https://medlineplus.gov/ency/patientinstructions/000495.htm

[6] Ethics Resources and Organizations. (n.d.). Retrieved from https://www.ieee.org/about/ethics/resources.html

[7]ACM Code of Ethics. (n.d.). Retrieved from https://www.acm.org/code-of-ethics

[8]IEEE Safety Recommendations For High Voltage Testing. (n.d.). Retrieved from https://www.spellmanhv.com/en/Technical-Resources/Articles/IEEE-Std-510-1983-IEEE-Recommended-Practices-for-Safety-in-High-Voltage-and-High-Power-Testing

[9] LABORATORY SAFETY TRAINING (n.d.). Retrieved from https://ovcrportal.research.illinois.edu/Training/Overview.aspx?TrainingId=67

[10] Adafruit Industries. (n.d.). Adafruit I2S Stereo Decoder - UDA1334A Breakout. Retrieved September 30, 2019, from https://www.adafruit.com/product/3678?gclid=EAIaIQobChMIj-LPpff35AIVUL7ACh0xwwx3EAAYASAAEgJ_WvD_BwE

[11] TPA2001D1PWR. (n.d.). Retrieved September 30, 2019, from https://www.ti.com/store/ti/en/p/product/?p=TPA2001D1PWR&utm_source=google&utm_medium=cpc&utm_campaign=ecm-tistore-invf-GMC_US-cpc-store-google-

wwe&utm_content=GMC&ds_k=PRODUCT_GROUP&DCM=yes&gclid=EAIaIQobChMI3JHK6rz45AIVxsDACh2SdAYrEAQYBSABEgLGKvD_BwE&gclsrc=aw.ds.

[12] Services, E. I. T. S. (n.d.). Salary Averages. Retrieved September 30, 2019, from https://ece.illinois.edu/admissions/why-ece/salary-averages.asp.

[13] Why Do Some Sounds Make People Sick? (2015, November 2). Retrieved from http://mentalfloss.com/article/69578/why-do-some-sounds-make-people-sick.

[14] zero-crossing rate. (n.d.). Retrieved from https://www.sciencedirect.com/topics/engineering/zero-crossing-rate.

[15] Spectral Centroid. (n.d.). Retrieved from https://www.sciencedirect.com/topics/engineering/spectral-centroid.

[16] AudioB I2S Bluetooth Digital Audio Receiver Module. (n.d.). Retrieved December 11, 2019, from https://www.tinyosshop.com/audiob-i2s-bluetooth-digital-audio-receiver-module.

[17] Jayu8. (n.d.). Jayu8/Audio-Equalizer-. Retrieved from https://github.com/Jayu8/Audio-Equalizer-.

Appendix A Requirement and Verification Table

	Requirement	Verification
Hardware		
a. Power Subsystem	Must maintain a 3.3 VDC +- .1V output	Utilizing a voltmeter attached to the output of this subsystem, ensure the DC Voltage is 3.3 VDC +- .1V.
b. Bluetooth and Digital to Analog Converter	Must be able to receive sinusoidal inputs and retain the frequency.	Sending audio signals of 440 Hz, 660 Hz, and 880 Hz, or A4, E4, A5 in terms of musical notes respectively from a paired Bluetooth device, use a tuner or similar device to verify the output frequency matches.
c. Operational Amplifier	Must be able to amplify the analog signal such that it will play and retain frequency through the speakers.	Send signals from a device of 440 Hz, 660 Hz and 880 Hz, or A4, E4, A5 in terms of musical notes respectively and then use a tuner or similar device to measure the frequency of the output.
Software (Genre Extraction)	-	-
a. Metadata Extractor	The Metadata Extractor must accurately extract the IDV3 “genre” tag from the song file.	Run the Metadata extractor on WAV, MP3, and Flac files and verify “genre” extracted matches the stored Metadata.
Software (Machine Learning)	-	-
a. Accuracy	The Machine Learning algorithm must be 70% accurate in determining the vocal characteristics of a singer(high/low).	Using a Testing dataset of songs (separate from the songs used to train the algorithm), verify that the precision for both High and Low Vocal characteristic songs is >= 70%
b. Latency	The Program must be able to make a prediction on a Song’s Vocal characteristics in 0.7 seconds or less. (Includes both feature extraction from song and ML prediction)	Using the time-it function on python, ensure that the total time for feature extraction and ML prediction does not exceed 0.7 seconds over multiple runs.

Software (Dynamic EQ/ Software Integration)	-	-
a. Software Integration/ Dynamic EQ	The Final Program must make an ML Prediction on Song Vocals and extract song genre to make and execute on EQ decisions. (High-Level explanation: tie all software together) Note: Our UI implements a display that shows genre extracted and ML prediction	1. Change the EQ Sliders to verify that the song is being modified in realtime/dynamically (not pre-processed) 2. Using songs of different genres, and singers/vocals, verify that the EQ settings, and/or displayed genre and ML Prediction are different.

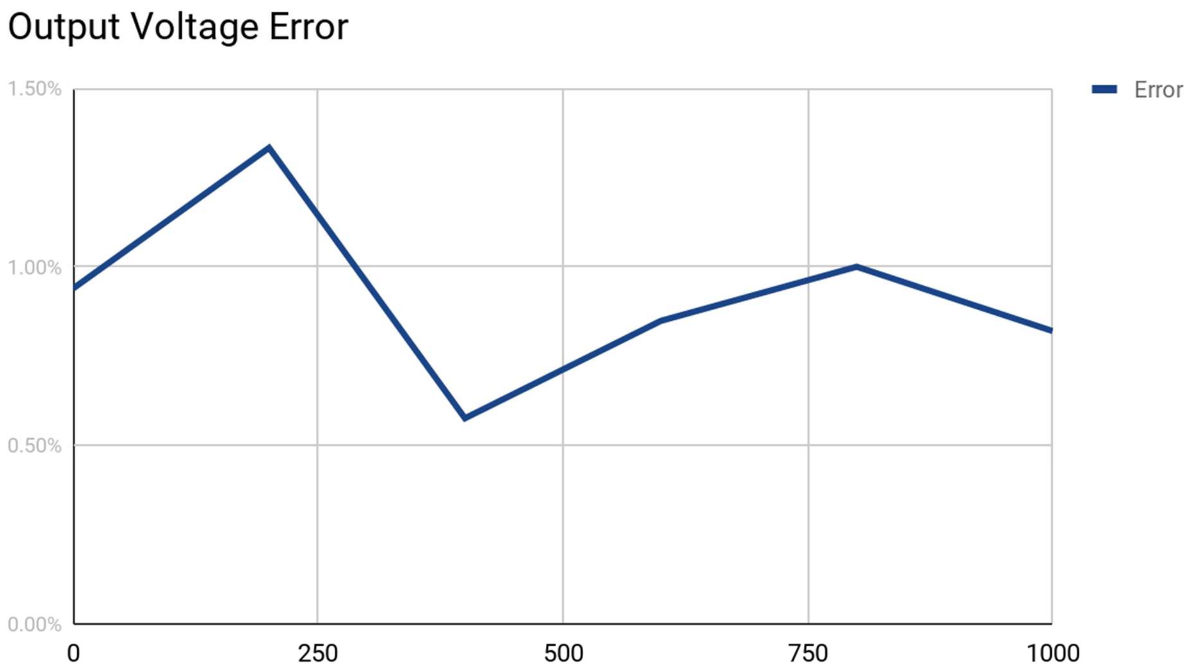
Appendix B Parts Table

Table 4: Parts Cost Table

Part	Cost
10 1-W FILTERLESS MONO CLASS-D AUDIO POWER AMPLIFIER	$9.60
2 Adafruit I2S Stereo Decoder - UDA1334A Breakout	$13.90
2 AudioB I2S Bluetooth Digital Audio Receiver Module	$23.90
2 Bluetooth 5.0 APTX Audio Module - TS8670	$25.90
2 Speakers - 3" Diameter - 8 Ohm 1 Watt	$3.52
Assorted Components. E.G. Resistors, Inductors, Capacitors, Transformers, Etc.	$45
Printed Circuit Boards	$80
Total	$201.82

Appendix C Graphs

Graph 1: Output Voltage Error for Power Subsystem



Appendix D Abbreviations

DC	Direct Current
AC	Alternating Current
ML	Machine Learning
RMS	Root Mean Squared
EQ	Equalization
Ω	Ohms
V	Voltage
A	Current
dB	Decibels
Hz	Hertz
C	Capacitance