Extended Essay: Mathematics

Session: May 2025

Wordcount: 3896

Topic of Interest: Fractal Geometry and Music

Research Question: To what extent can the Fourier and Wavelet transforms be used to identify fractal patterns in an audio signal?

Investigating Fractal Patterns in Music

Research Question: To what extent can the Fourier and Wavelet transforms be used to identify fractal patterns in an audio signal?

Table of Contents

1 | Introduction

Theorised by Benoit Mandelbrot in 1975, Fractals are commonly known as self-similar patterns on different scales found in the natural world around us("fractal"). Examples can be seen in galaxies, flowers, leaves, and rivers.



Fig 1. Fractal Patterns in Galaxies("Whirpool Galaxy") Fig 2. Fractal Patterns in Plants (9 Amazing) Fig 3. Fractal Patterns in rivers(9 Amazing)

Fractals can also be found through computerised algorithms such as the Mandelbrot set where groups of complex numbers are graphed in a manner that is also self-similar at varying scales.

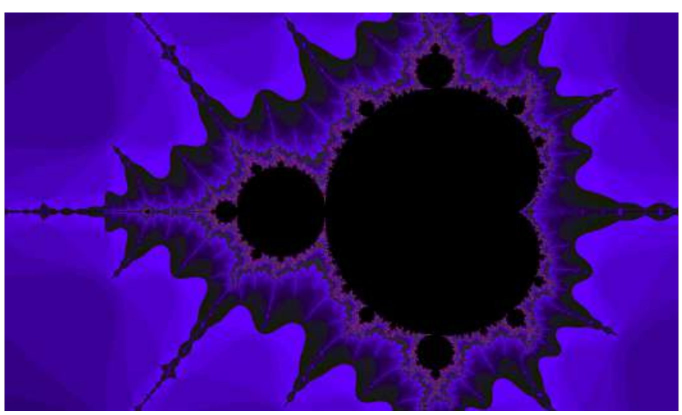

Fig 4. The Mandelbrot Set

The fractal dimension is a measure of the degree of self-similarity an object has. A higher dimension suggests that an object has greater self-similarity. A fractal dimension is also a measure of the complexity/roughness of an object. The fractal dimension of a line is 1, the fractal dimension of a square is 2 and a cube has a fractal dimension of 3 ("Fractal Dimension"). A fractal dimension less than 1 means the object has properties similar to a line but has a roughness in some form that does not permit it to be strictly 1-dimensional. Similarly, a fractal dimension that is between the values of 1 and 2 implies that the object looks two-dimensional but has some roughness to it. An example of this is the coastline of Britain or the Sierpinski Triangle. An object that is three-dimensional but has roughness would yield a fractal dimension between 2 and 3("Fractals Are Typically").

Currently, there is no commonly accepted method through which fractals manifest in music and what a fractal pattern in music comprises but different approaches have been taken in the past. These include Temporal Scaling and Tonal Complexity.

Temporal scaling relies on fractal-like variations in tempo where a single note is scaled down by a constant factor and duplicated to produce a repeating pattern, thus exhibiting fractal properties. Tonal complexity analyses fractals through the variation of pitch. It helps distinguish between musical patterns that vary not in their tempo but in their tones.

1.1 | Aim

This investigation aims to identify existing fractal patterns in audio signals. This will help better understand the tonal and rhythmic complexities within music and contribute to existing musical knowledge.

This leads to the research question: **To what extent can Fourier transform and Wavelet transform be used to identify fractal patterns in an audio signal?**

2 | Fractals in Geometry and Music: A Comparison

2.1 | Fractal Patterns and Dimensions

To explain the logic behind the fractal dimension, the case of known bodies will be taken. If a straight line were to be scaled down by a factor of 2, it would yield two identical smaller lines. If the same were to be done with a square, one would obtain four smaller squares of equal area, and if a cube side length were to be scaled down by a factor of 2, one would obtain 8 smaller cubes of equal volume. Since all these broken-down pieces have the same exact shape, we call them “self similar pieces.” The data has been presented in tabular form below and visualised through the image provided:

Geometric Body	Scale factor	Number of self-similar pieces
Line	2	$2 = (2)^1$ - Line is 1D
Square	2	$4 = (2)^2$ - Square is 2D
Cube	2	$8 = (2)^3$ - Cube is 3D

Table 1: Scaling factor and Self Similar Pieces

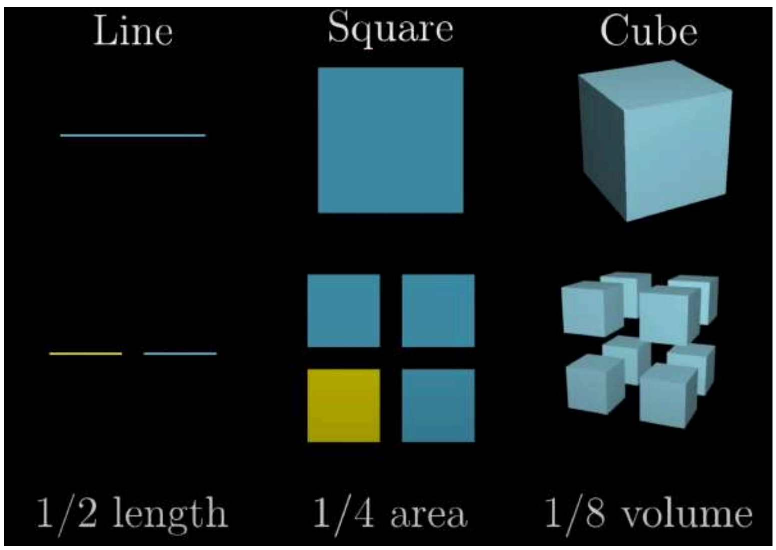

Fig 5. Fractal Dimension Derivation ("Fractals Are Typically")

A relation between the scaling factor and a number of self-similar pieces is present. In fact, the obtained number of self-similar pieces is the **scaling factor raised to the dimension of the geometric body.**

This presents us with a definition of a dimension ("Fractals Are Typically").

Neither lines, squares, or cubes are true fractals(objects that have a non-integral dimension). Two common examples of patterns considered to be fractals are Koch's Snowflake and Sierpinski's triangle.

The Sierpinski's triangle is a shape considered to be a fractal.

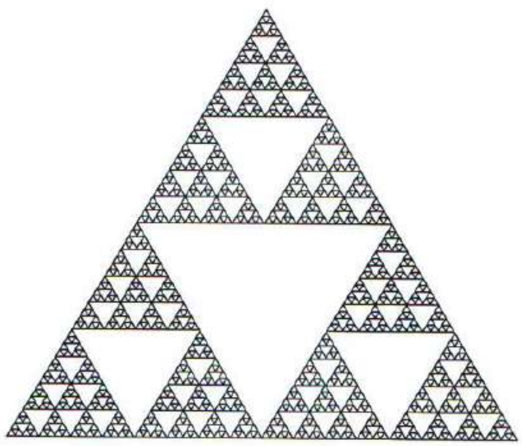

Fig 6. Sierpinski's Triangle ("Fractal Dimension")

If the shape is scaled down by a factor of 2 (all the side lengths are halved), three self-similar copies of following shape are obtained.

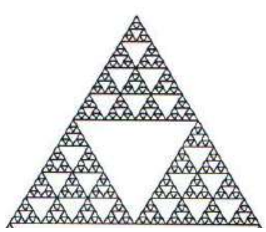

Fig 7. Scaled Sierpinski's Triangle ("Fractal Dimension")

Using the previous definition of dimension, the following equation is obtained:

$(2)^D = 3$ where D is the dimension. This leads to the formulation of a logarithmic relationship:

$D \times log(2) = log(3)$

$D = \frac{log(3)}{log(2)} \approx 1.585.$

The dimension of Koch's snowflake can be found in a similar manner:

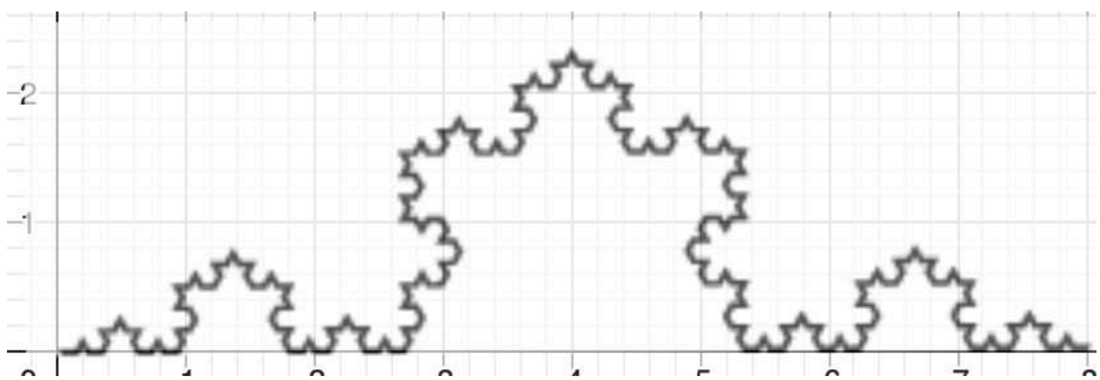

Fig 8. Koch's snowflake

In this case the length (horizontal axis) has been reduced to a third of its initial size. However, the scaling of the object yields four self-similar copies of the snowflake. Hence the fractal dimension is:

$4 = (3)^D$

From this, we get

$log(4) = log(3^D)$

Using log properties:

$log(4) = D\, log(3)$

Hence:

$D = \frac{log(4)}{log\ (3)} \approx 1.262$ ("Fractals Are Typically")

A generalised equation can be formed relating the fractal dimension D; the factor by which the object is being scaled, S; and the total number of self-similar pieces created after scaling, N.

Thus, we get: $D = \frac{log(N)}{log\ (S)}$ ("Fractal Dimension").

2.3 | Literature Review - Fractals in Music: Temporal Scaling and Tonal Complexity

In temporal scaling, each copy of the music is created such that the notes of the motif (the melody) are shortened by a given factor. Temporal scaling works similarly to traditional geometric fractals where objects are scaled and tiled. Initial suggestions of this method were done by Henderson Sellers and Cooper. A sample motif begins with a four-note melody and is followed up with the same melody but tiled 4 times with quarter notes. This melody is further copied with 64 eighth notes to form the fractal pattern.

In the image below, a musical pattern comprising a two-note and two-rest pattern has been composed. In the second line, this pattern has been repeated only this time with three notes that are half the duration.

Fig 10. Temporal Scaling (McDonough and Herczyński)

The resulting fractal dimension is:

$$D = \frac{log(3)}{log(2)} = \frac{ln(3)}{ln\,(2)}\ 1.58$$ (McDonough and Herczyński)

This method of creating sequences with notes shorter in length have been defined as nested sequences. Thus it is possible to create a non-integral fractal dimension in a musical composition.

Tonal Complexity

On the other hand, tonal complexity is a method that manipulates the pitch of a scoresheet. This can be used to analyse music containing minimal temporal variation. Tonal complexity helps distinguish between musical patterns that vary not in their tempo but in their tones. In order to quantify this, the 12 note scale, the difference between two notes is given as a ratio of their frequencies. As a baseline, two notes separated by an octave are in the ratio 2:1. Consequently, the ratio between two notes differing by a specific interval of j notes is given by $2^{\frac{j}{12}}$. The ratio obtained is further analysed for fractal patterns through the application of log rules (McDonough and Herczyński).

Temporal scaling works in a similar method to geometric fractals. As previously elaborated, the temporal scaling relies on the scaling factors and the number of total self-similar pieces that appear due to the scaling down in the duration of a note. However, tonal complexity in music, as defined in the paper, only exists for specific cases when the notes on the scoresheet resemble a visual fractal. Since this method of calculating fractals cannot be generalised for any musical motif, a new method to decipher the fractal dimension in music must be used.

3 | Methodology and Data Collection

3.1 | Methodology

The approach taken is one that will analyse the **waveforms** created by music.

This first step would be to understand precisely how a **change in musical characteristics affects** the signal produced.

This would be followed by an exploration of two different mathematical tools: **The Fourier and Wavelet Transforms** in an attempt to obtain a fractal dimension of the audio signal.

3.2 | Data Collection

To conduct this study, a sample music audio known to exhibit qualitative fractal patterns was used. The soundtrack used was created by YouTuber Woochia - Charly Sauret. The audio comprised a melody that was initially played at a given pace. The melody was then layered with repeats of itself at shorter paces. Each layer was also played at a higher octave such that when it was played, it sounded like the music was ridden with many smaller, self similar layers of itself.

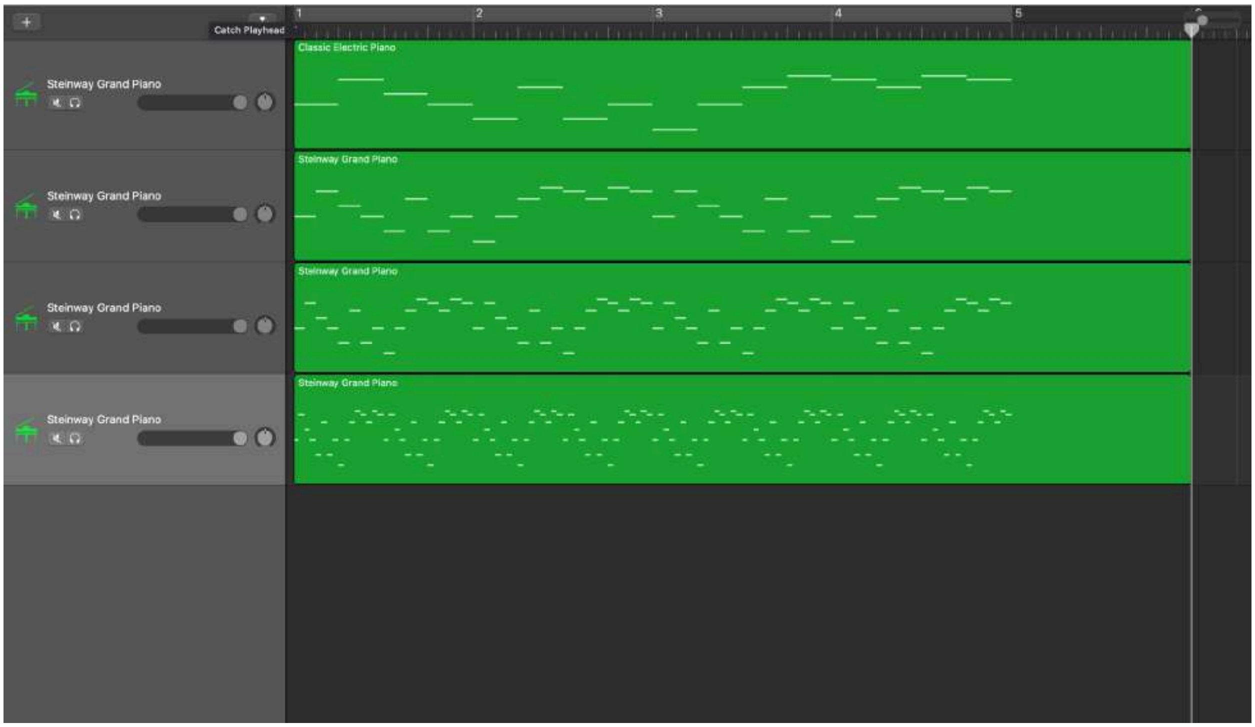

Fig 11. GarageBand File with the piece of music being analysed

5 | Using the Fourier Transform to Investigate Fractal Patterns in Music

The Fourier Transform is a mathematical tool that breaks down a waveform (function or signal) into its sine and cosine components. The transform converts a waveform from a time vs amplitude graph to a frequency vs amplitude graph.

The Fourier Transform is given as follows:

$$\widehat{g}(f) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i f t}dt$$

Where t is the time interval and f is the frequency ("But What").

From the Fourier Transform, any given signal can be represented as the sum of a group of sinusoidals.

When graphed, the waveform obtained from the Fourier Transform is a graph with varying peaks that signify the strength (amplitude) of the signal at given frequencies. When applied to music, the peaks observed help understand the notes that comprise the audio signal.

The Fourier Transform has an imaginary sine component. However, graphical representations of the Transform omit this part of the wave and hence the individual component graphs only include the cosine component and are represented in the following way: $acos(2\pi f t + c)$ where a is the amplitude of the wave, f is the frequency and c is the phase response of the signal at the given frequency.

The amplitude of a sound signal is usually given in decibels. The conversion between decibels and metres is:

Metres (amplitude) = $10^{\frac{dB}{20}}$ ("Decibel Conversion")

For example, if the amplitude of the sound wave was -12 dB:

$10^{\frac{-12}{20}} = 2.51 \times 10^{-1}$ metres.

To epitomise this process, the Fourier Transform has been applied to the sample audio created:

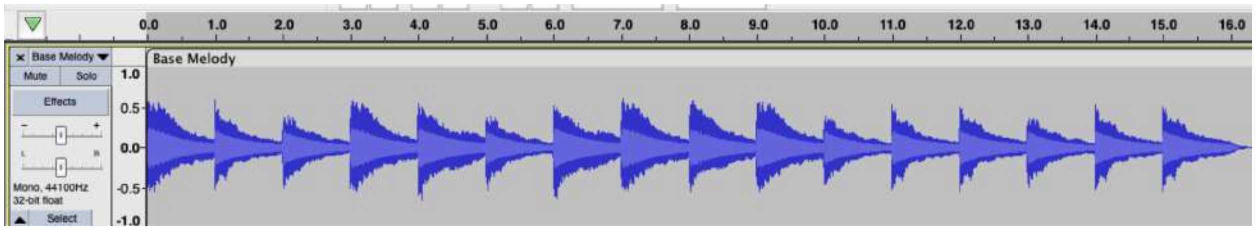

Fig 14. The waveform of the created audio file on Audacity Software

Figure 14 depicts the wave signal from the audio file. This signal is currently in the time domain where amplitude is graphed against time. The Fourier Transform is then computed and the signal is converted into the frequency domain. This is done using a software called Audacity.

The Fourier Transform (Fig 15) highlights the frequencies at which the sound wave has the strongest power (amplitude). The peak frequencies and their associated amplitudes are recorded and substituted into the cosine equation given above. Given the complex nature of the sinusoid, there will be an argument (known as the phase) of the waveform. As a result, each frequency is accompanied by a phase response. In the component sinusoids, the phase plot depicts the horizontal translation of the cosine wave.

.

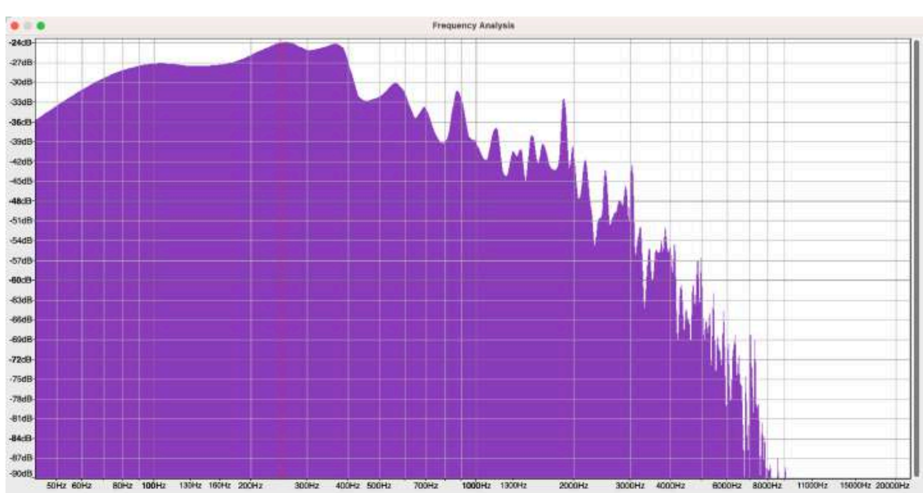

Fig 15. The Fourier Transform of the created audio file on Audacity Software

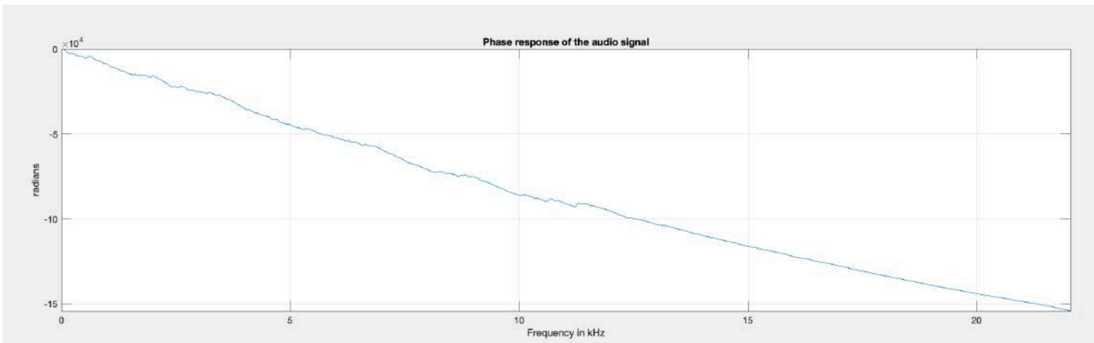

Fig 16. The Phase plot of the created audio file on Audacity Software

From the Fourier and phase diagrams of the sound signal, one can find the individual sinusoids that comprise the sound signal. The following are the cosine components obtained from the graph above.

Note	Frequency	Amplitude (dB)	Sound Wave Equation
F#2	96	-26.0	$0.05012 cos((2\pi)(96)(t) - 1302.02)$
B3	248	-23.8	$0.06531 cos((2\pi)(245)(t) - 2075.12)$
F#4	367	-24.2	$0.06166 cos((2\pi)(367)(t) - 3296.94)$
C#5	553	-29.8	$0.03236 cos((2\pi)(553)(t) - 3658.00)$
F5	680	-33.5	$0.02113 cos((2\pi)(680)(t) - 4056.30)$
A5	872	-31.2	$0.02754 cos((2\pi)(872)(t) - 5733.22)$

A#6	1862	-32.4	$0.02399cos((2\pi)(1862)(t) - 11173.50)$

Table 2: Equations of the waveform from sample audio

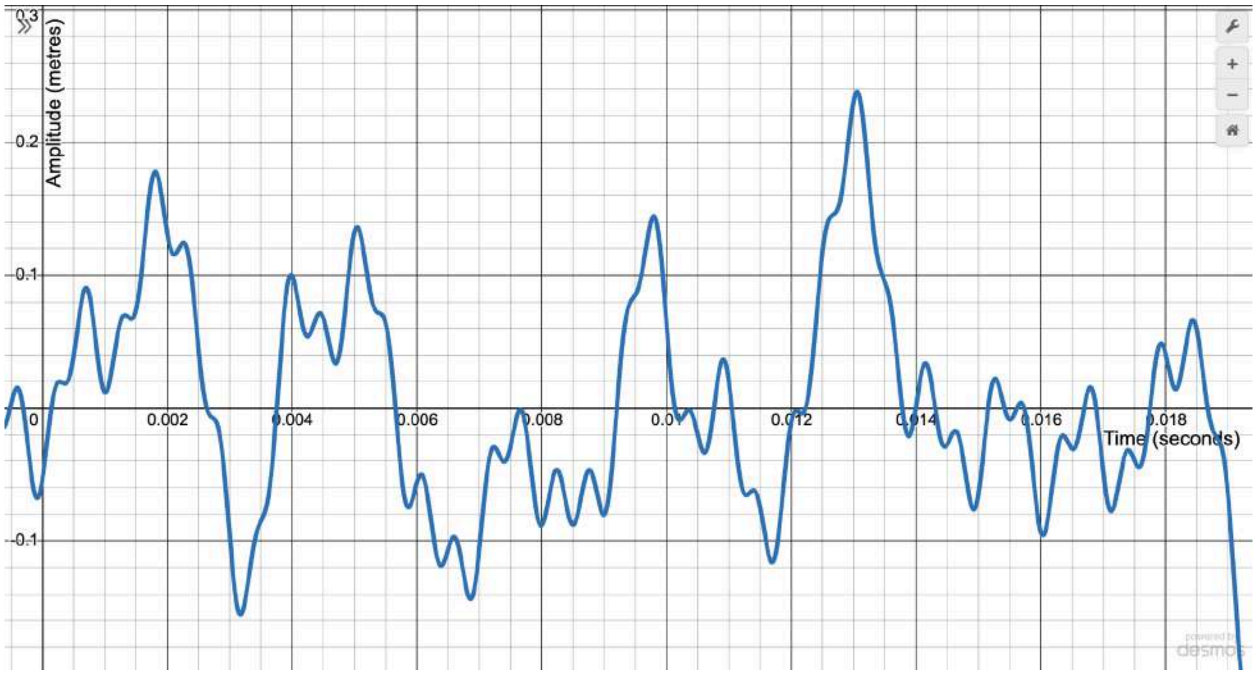

Fig 17. The reconstructed waveform of the audio graphed on desmos using the information from the Fourier Transform

The graph above depicts the waveform for a sample sound file that has been analysed. When the equations derived from the Fourier Transform representation of the file were graphed, the initial waveform in the time domain was obtained.

5.1 | Manipulation of Sound and its Impact on the Fourier Decomposition

There are two main characteristics of sound that can be manipulated: the pitch (how high or low a note sounds) and the tempo (how fast the sound may be repeated) of the signal, all of which have been demonstrated below. Since this approach includes the analysis of waveforms, seeing how musical characteristics such as pitch, volume and tempo affect waveforms is important.

5.1.1 | Impact of Pitch on Waveforms:

Below are the Fourier Transform graphs of the C4 and C5 notes (two C notes separated by an octave). The Fourier Transforms of the Graphs are shown below. To investigate the impact of pitch, the loudness(-12dB) and the tempo will be kept constant.

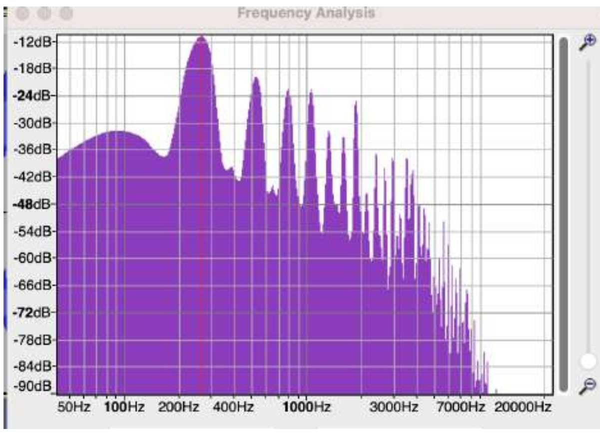

Fig 18. C4 note Fourier transform

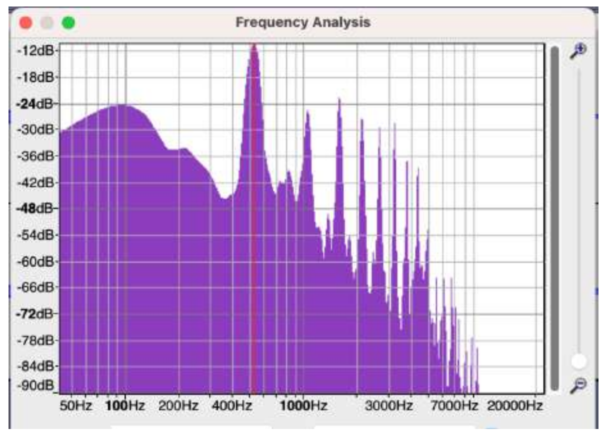

Fig 19. C5 note Fourier transform

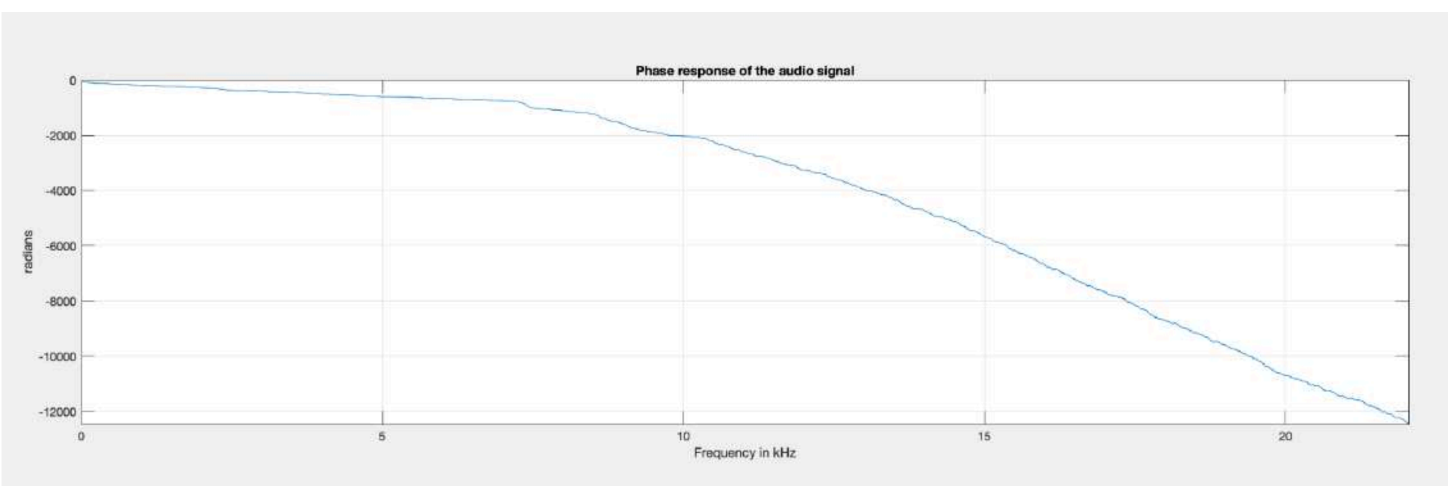

Fig 20. C4 note Phase Response

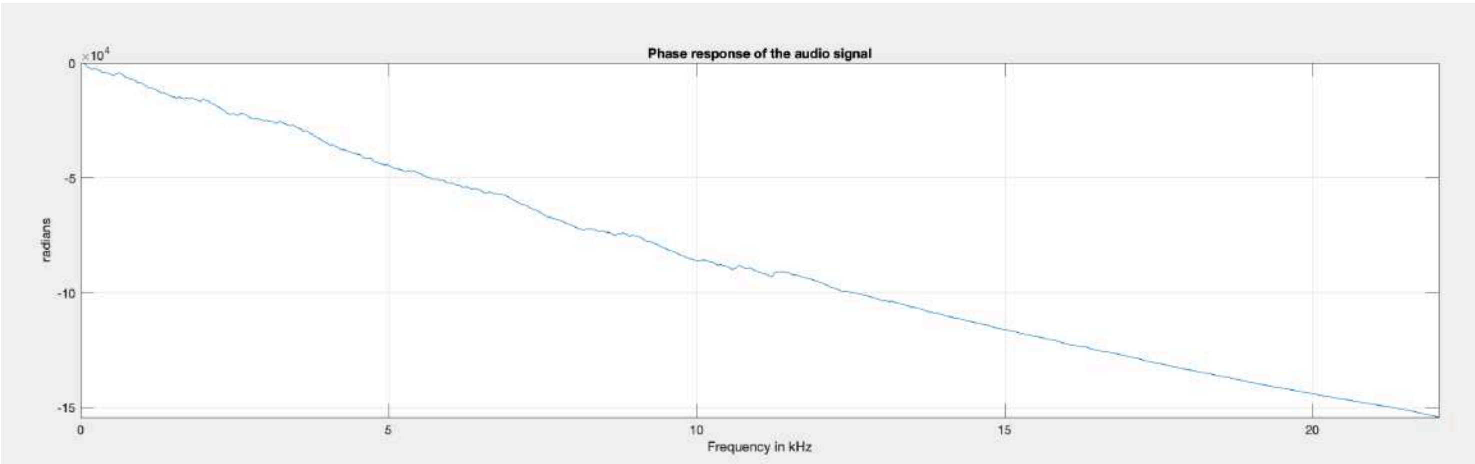

Fig 21. C5 note Phase Response

In the images above, the Fourier Transform and phase plots of the two notes can be seen. When a single note is played, there is usually one single frequency where a spike is observed. This is known as the principal frequency and is the most prominent component of the sound signal. However, the peak is also accompanied with harmonics (smaller and weaker components of the sound wave). The C4 note has a frequency of approximately 256 Hz (as indicated in the diagram) and is the strongest component of this sound signal (fundamental frequency).

In the second figure, a C5 note has been played at the same controlled volume. From the Fourier Transform, it can be seen that the C5 note is the principal frequency at 525Hz due to the major spike in the Transform.

The equations pertaining to both C4 and C5 notes can be seen below:

C4 note equations	C5 note equations

$0.15488\, cos((2\pi)(261)(t) - 98.56)$	$0.16898\, cos((2\pi)(525)(t) - 130.783)$
$0.04121\, cos((2\pi)(515)(t) - 131.288)$	$0.034674\, cos((2\pi)(1049)(t) - 199.600)$
$0.044157\, cos((2\pi)(384)(t) - 112.56)$	$0.047863\, cos((2\pi)(1547)(t) - 249.473)$
$0.064565\, cos((2\pi)(652)(t) - 149.892)$	$0.01698\, cos((2\pi)(2648)(t) - 373.228)$
$0.057544\, cos((2\pi)(919)(t) - 180.892)$	$0.02786\, cos((2\pi)(89)(t) - 320.438)$

Table 3: Waveform constituent parts equations for C4 and C5 notes

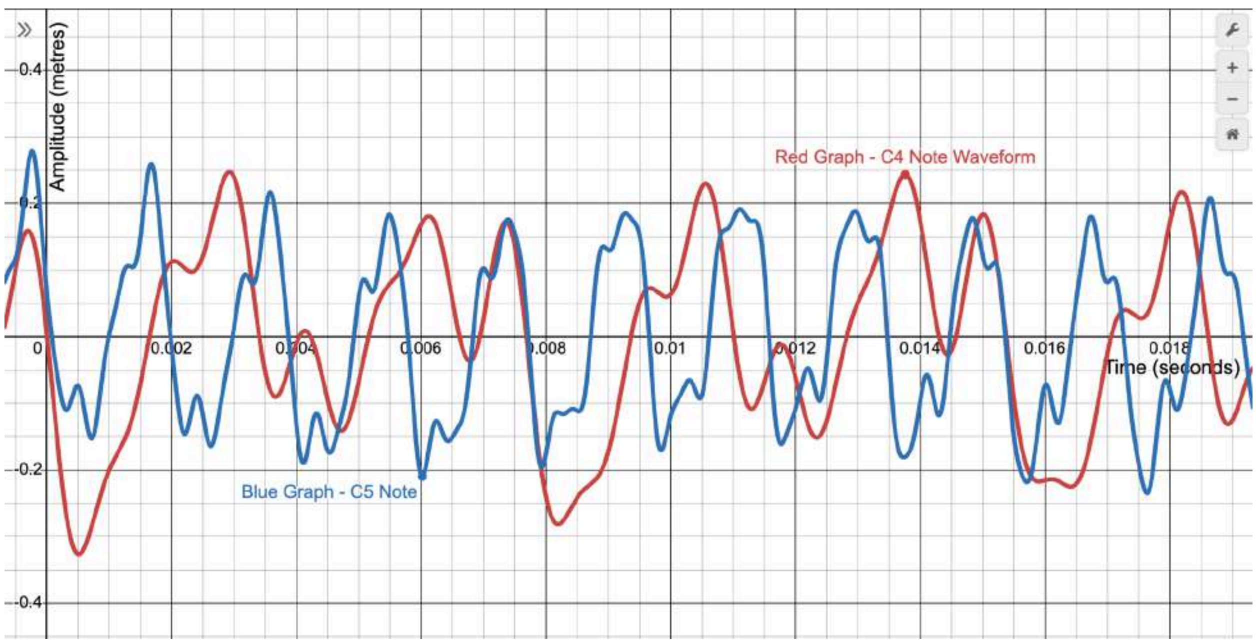

Fig 22. The constructed waveforms of the C4 and C5 note using Desmos

By comparing the Fourier representations of the two signals, it is clear that an increase in pitch leads to an increase in the frequency of the principal note. Also, the frequency of the C5 note was observed to be approximately double the frequency of the C4 note. Thus, it can be concluded that the frequency of notes separated by octaves is multiples of each other. Important to note on these graphs is the domain restriction where $x > 0$ as time is positive.

5.1.2 | Impact of Tempo on Waveforms:

The final characteristic of sound that is investigated is the tempo of the sound. This is important to investigate because previous literature on fractals in music uses temporal scaling (manipulation of the music's tempo) in a way that mimics the Cantor set and Koch's snowflake to create fractal patterns. In order to understand the effect of tempo on the Fourier Transform and the signal in general, a C4 note was repeated a varying number of times over a fixed time interval.

The equations derived from the Fourier Tansform of the C4 note have already been given above. To evaluate the tempo, a time interval of two seconds over which the sound wave would be analysed was kept constant. Instead of playing a single note, three notes were played in the fixed time interval. The equations obtained for them have been given below:

C4 Slower Tempo (1 rep/2 second interval)	C4 Faster Tempo (3 reps/2 second interval)
Principal: $0.15488\ cos((2\pi)(261)(t) - 98.56)$	Principal: $0.3981cos((2\pi)(261)(t) - 829.454)$
C5: $0.04121\ cos((2\pi)(515)(t) - 131.288)$	C5: $0.07413\ cos((2\pi)(515)(t) - 1556.91)$
G4: $0.044157\ cos((2\pi)(384)(t) - 112.56)$	G4: $0.08035\ cos((2\pi)(384)(t) - 1173.29)$
E5: $0.064565\ cos((2\pi)(652)(t) - 149.892)$	E5: $0.09226\ cos((2\pi)(652)(t) - 1954.08)$
A#5: $0.057544\ cos((2\pi)(919)(t) - 180.892)$	A#5: $0.05754\ cos((2\pi)(919)(t) - 2741.92)$
A#2: $0.056234cos((2\pi)(118)(t) - 53.5889)$	A#2 note: $0.18621cos((2\pi)(118)(t) - 604.29)$
G5: $0.01995cos((2\pi)(779)(t) - 161.876)$	G5: $0.06531cos((2\pi)(779)(t) - 2341.47)$

Table 4: Constituent Equations of the C4 note at varying tempi

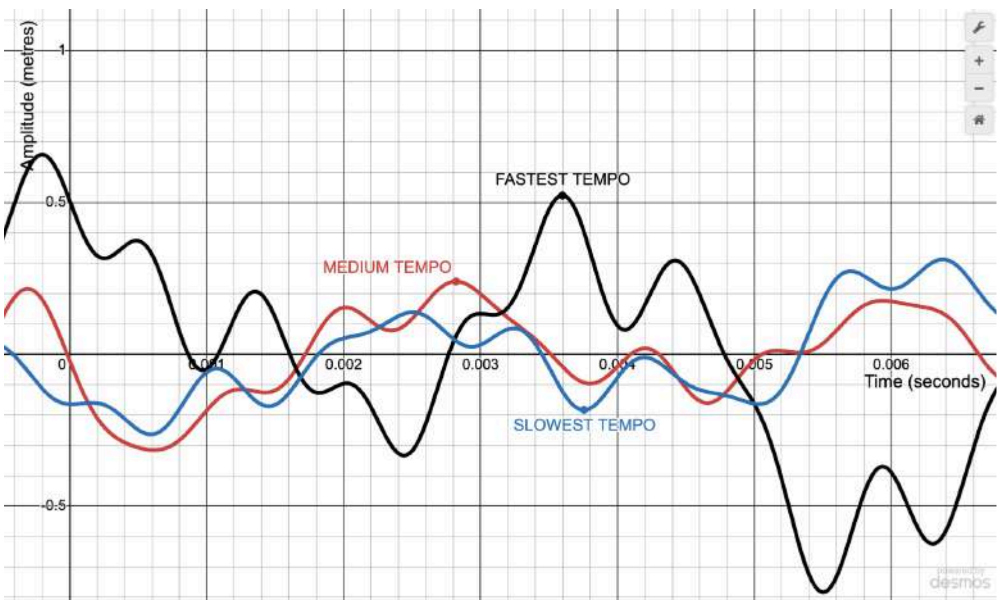

Fig 23. Comparisons of the waveforms of notes at varying tempi

From the analysis of a tempo that is even faster than the previous analysis, the phase value at each frequency continues to increase. In the equation of the cosine wave, this is seen as a change in the horizontal translation of the waveform as the phase of the wave is denoted by ‘c’ in the equation: $y = a\cos(2\pi f t + c)$

Also, the amplitude is increasing. To understand the 3 audio files in comparison to each other, their equations have been plotted below, thus creating the waveform each audio exhibits:

Despite seeing a change in the phase, there is no evident pattern regarding the amount by which the phase or amplitude changes which makes interpretation of any applications of the Fourier Transform difficult.

5.2 | Fractal Dimension of a piece of music using the Fourier Transform

From the initial investigation, the following conclusions have been drawn:

1. An increase in the pitch of the sound file leads to an increase in the frequency of the waveform. Moreover, increasing the pitch by a perfect octave leads to the exact doubling of the waveform's fundamental frequency
2. An increase in the loudness of the sound file leads to an increase in the amplitude of the waveform
3. An increase in the tempo of the waveform leads to an increase in the phase and amplitude

After using the Fourier Transform in an attempt to identify fractal patterns in music, it has been concluded that the Transform does not have the complexity required to help find fractal patterns. Tempo, which is an important factor for fractals in music cannot be manipulated in a controlled manner. This irregularity makes finding any type of self-similarity difficult. An alternative is the wavelet transform: a more intricate mathematical tool which represents an audio signal in the time and frequency domain will be explored further.

6 | Using the Wavelet Transform to Investigate the Fractal Patterns Music

The Wavelet Transform is another mathematical tool used in waveform analysis. This mathematical tool goes a step further from the Fourier Transform and helps analyse a signal in both the time and frequency domain and is expressed in the following manner:

$F(\tau, s) \ = \ \frac{1}{\sqrt{|s|}} \times \int_{-\infty}^{\infty} f(t) \ \times \ \psi \ * \ (\frac{t-\tau}{s}) dt$ (M.G. Manisha Milani and De Silva)

Where:

$F(\tau, s)$ are the decomposition coefficients - representation of the amplitude and phase response of a signal at different scales and positions (parameters) in the transform - obtained at scaling and translation parameters s, τ and

$\psi \ *$ is the complex conjugate of the binary Haar wavelet.

A scaling parameter, expressed as $\psi \ * \ (\frac{t}{s})$ allows the signal to be compressed ($0 \ < \ s \ < \ 1$) depicting lower frequencies of the signal or stretched ($s \ > \ 1$) depicting higher frequencies and more detailed parts of the wavelet. Thus, the scaling factor permits the analysis of the waveform broadly or in great detail.

The translation parameter, expressed as $\psi \ * \ (t \ - \ \tau)$ allows the signal to be shifted horizontally across the time axis. However, since this investigation is about fractal patterns, only an analysis of the scaling parameter (the signal at varying magnification levels) is necessary as the translation parameter does not change the waveform.

The transform, however, does have a limitation. At a given time, accuracy can only be obtained in either the time or frequency domains.

When values for the scaling and translational parameters are substituted into the transform, the resulting set of values obtained are known as decomposition coefficients. In MATLAB, the coefficients obtained are from increasing “levels.” The exact scaling values are not known but it is known that each level yields progressively more detailed readings.

6.1 | Investigating Fractal Patterns using Wavelet Transform

Exploring fractal patterns using the wavelet transform has been done before by Guoxi Lia, Kai Zhanga, Jingzhong Gongb, and Xin Jin. To do so, the decomposition coefficients at varying levels will be taken and analysed. Since each level yields a large number of coefficients, the variance of these coefficients will be calculated and graphed. The graph will be the level or scaling parameter "s" at which the decomposition coefficients are taken (1,2,3,4,5,6,7) vs $log_2(variance\ at\ level)$. The variance levels will be plotted until the relationship received is non-linear. The mathematical proof for this relationship is the following:

As previously mentioned, the decomposition coefficients of the wavelet are given by:

$$d^{s,\tau}(t) = F(s,\tau) = \int_{-\infty}^{\infty} f(t)(\psi^{*})(t)dt$$

Where ψ^{*} is the complex conjugate of $\psi^{s,\tau}$

The binary wavelet satisfies the condition of zero mean value (the mean of the decomposition coefficients obtained at particular scaling parameters is 0).

The power spectral density given by $S(\omega)$ is the measurement of the power a signal carries (the sum of the absolute squares of its time-domain samples divided by the signal length) at various frequencies. The spectral density at a given frequency is given by the following formula:

$$S(2^{-s}\omega_0) = \frac{1}{N_s}\sum_{k=1}^{N_s}(d^{s,\tau})^2$$

Where ω_0 is the reference frequency of the function at which the power is calculated and N_s is the number of decomposition coefficients at scaling and translation parameters s, τ respectively.

At high frequencies, the surface of the signal is zoomed into, thus providing information about the finer details of the wavelet. On the other hand, at lower frequencies, a broader overview of the wavelet is seen. As a result, low frequencies, due to the proportion of the signal they cover, dominate the power spectrum and have the highest powers while higher frequencies have lower power.

Thus, the power spectral density of a signal is inversely proportional to the frequency. Through empirical observations, a relationship known as the $\frac{1}{f^{\beta}}$ law is obtained where the power spectrum is inversely proportional to f^{β} where β is a constant and is known as the power-law exponent (developed through empirical evidence).

Thus, in this case:

$$S(\omega) \propto \frac{1}{\omega^{\beta}}$$

The power law exponent is related to the Fractal Dimension D through empirical observation in the following way:

$\beta = 2(2 - D) + 1$

$\beta = 4 - 2D + 1$

$\beta = 5 - 2D$

We can now substitute this into:

$$S(\omega) \propto \frac{1}{\omega^{(5-2D)}}$$

From this, we get:

$$\frac{1}{N_s}\sum_{k=s}^{N_s}(d^{s,\tau})^2 \propto \frac{1}{(2^{-s}\omega_0)^{(5-2D)}}$$

The variance, of the decomposition coefficients a scaling paramaters s, τ is given as:

$$var[d^{s,\tau}] = \frac{1}{N_s}\sum_{k=1}^{N_s}(d^{s,\tau} - \overline{d^{s,\tau}})^2$$

Since the binary wavelet satisfies the condition of the zero mean value, the mean becomes 0. Therefore,

$$var[d^{s,\tau}] = \frac{1}{N_s}\sum_{k=1}^{N_s}(d^{s,\tau} - 0)^2$$

As a result, the power spectrum and variance at parameters τ, s is given as

$$var[d^{s,\tau}] = a \times \frac{1}{(2^{-s}\omega_0)^{(5-2D)}} = a \times \frac{1^{(5-2D)}}{(2^{-s}\omega_0)^{(5-2D)}} = a \times (\frac{1}{(2^{-s}\omega_0)})^{(5-2D)}$$

Where a is the constant of proportionality

Taking the base two logarithms of both sides (base 2 has been taken as the wavelet being used is a binary wavelet).

$$log_2(var[d^{s,\tau}]) = log_2((\frac{1}{(2^{-s}\omega_0)})^{(5-2D)} \times a)$$

$$log_2(var[d^{s,\tau}]) = log_2((\frac{1}{(2^{-s}\omega_0)})^{(5-2D)}) + log_2(a)$$

$$log_2(var[d^{s,\tau}]) = (5 - 2D) \times log_2(\frac{1}{(2^{-s}\omega_0)}) + log_2(a)$$

$$log_2(var[d^{s,\tau}]) = (5 - 2D)log_2(\frac{2^s}{\omega_0}) + log_2(a)$$

$$log_2(var[d^{s,\tau}]) = (5 - 2D)[log_2(2^s) - log_2(\omega_0)] + log_2(a)$$

$$log_2(var[d^{s,\tau}]) = (5 - 2D) \times [s\, log_2(2) - log_2(\omega_0)] + log_2(a)$$

$$log_2(var[d^{s,\tau}]) = (5 - 2D) \times s - (5 - 2D)log_2(\omega_0) + log_2(a)$$

Taking $- (5 - 2D)log_2(\omega_0) + log_2(a)$ as a constant "b":

$$log_2(var[d^{s,\tau}]) = (5 - 2D) \times s + b$$

This creates a linear relationship correlating the level at which the scaling parameter s is taken with the binary log of the variance of the decomposition coefficients at the given level

In the function, b is the y-intercept of the graph.

Therefore the fractal dimension is given as:

$D = \frac{(5-M)}{2}$ (Li et al.)

Where $M = 5 - 2D$ is the gradient of the graph.

From the method proved above, the fractal dimension of the previously used audio file is once again called upon.

The next step was to analyse the fractal dimension of the layered audio.

Since there are a massive number of decomposition coefficients at each scaling level, the variance cannot be calculated manually. Instead, MATLAB is used and the variance is obtained through the software's processing. The commands imputed into MATLAB can be found in the appendix attached.

From the method previously detailed, the variances obtained from the decomposition coefficients at the various levels were obtained as:

Level (scaling vector)	Variance of Decomposition Coefficients	$log_2(variance)$
1	$8.190900004133981 \times 10^{-5}$	-13.5756184926
2	$6.179638006060353 \times 10^{-4}$	-10.6601900499
3	0.004087743510074	-7.93447960902
4	0.020450087231319	-5.61174919268
5	0.067383453650594	-3.89146181598
6	0.119855603927593	-3.06063073017
7	0.122479993752166	-3.029381980536

Table 5: Variance of the decomposition coefficients of the Wavelet at varying scaling parameters (for layered audio)

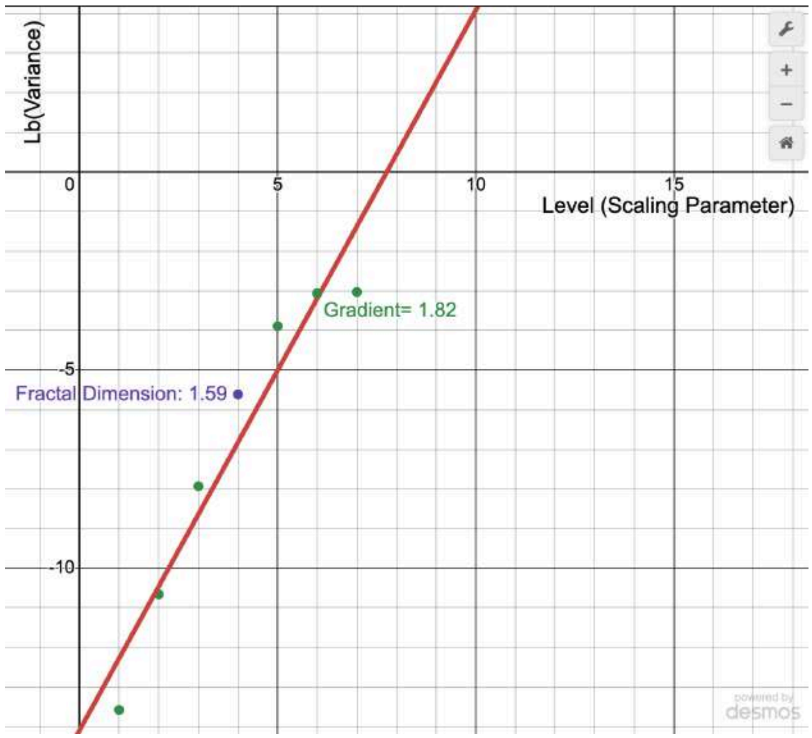

Fig 24. The obtained linear plot of the level of scaling parameter vs binary log of variance of the decomposition coefficients for the layered audio

A fractal pattern is only obtained if the graph formed is linear. The graph here has been limited to solely 7 data points because adding more points reduces the accuracy of the linear function given as the R^2 value reduces significantly.

From the proof, the fractal dimension of the signal is $\frac{5-M}{2}$. In this case, the fractal dimension obtained is 1.59.

However, in order to make a conclusion, the fractal dimension of the audio without the extra fractal-like layers needs to be analysed. The variance at different levels for the base audio has been given below:

Level (scaling vector)	Variance of Decomposition Coefficients	$log_2(variance)$
1	$5.508023009471281 \times 10^{-5}$	-14.1481058882
2	$4.216274983871875 \times 10^{-4}$	-11.2117434175
3	0.002892597078669	-8.43341890612
4	0.014820984965500	-6.07621486105
5	0.053262113098713	-4.23074652321
6	0.185377521599113	-2.4314617778
7	0.399187062853260	-1.32486312937

Table 6: Variance of the decomposition coefficients of the Wavelet at varying scaling parameters (for base audio)

From the information given above, the same plot was created:

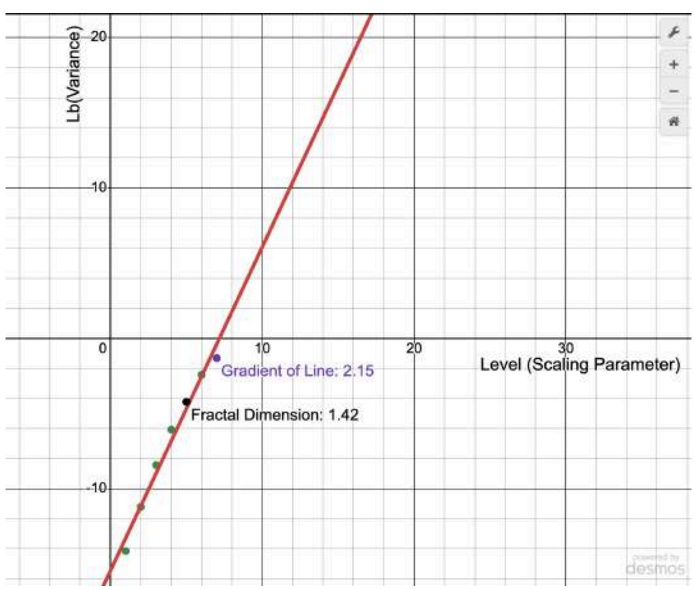

Fig 25. The obtained linear plot of the level vs log (variance) of the decomposition coefficients for the base audio

The same formula involving the slope of the line was used. The fractal dimension ultimately obtained was 1.42.

7 | Evaluation

7.1 | Limitations

While this investigation has yielded the intended results, there are limitations to the method. Given that MATLAB has been used to process the math, there could be processes in the code that have not been accounted for. Since there is a lack in programming knowledge, the processing and accuracy of the commands imputed into MATLAB cannot be analysed in detail As a result, the obtained results may not be entirely accurate.

7.2 | Improvements

Possible improvements to the method include performing the same analysis on pieces of music that range between different genres and observing whether a pattern of fractal dimensions emerges when the genre of music is changed. This would help compare and contrast the complexity in the structures of specific music genres. Moreover, if genres were to differ in their fractal dimensions, the fractal dimension could serve as an indicator of style, perhaps providing a greater insight into why composers adhere with/deviate from particular patterns. Finally, different music evokes varied emotional responses from an audience. Deciphering the difference in fractal characteristics could further help explain why this happens in general.

To increase the overall accuracy of results, other platforms such as Python could be used to generate all the results that were otherwise obtained through MATLAB.

8 | Conclusion

To conclude, finding the fractal dimension in audio files using the Fourier and Wavelet Transforms is possible to a large extent. Through the use of the boxcount method and the decomposition coefficients of the wavelet transform, the fractal dimension of a layered audio that has been temporally scaled and tonally manipulated has been obtained. When compared to the base melody created by the audio, the fractal dimension obtained is a higher value indicating greater self-similarity. While using the Fourier Transform was instrumental in understanding the way musical characteristics influenced the waveform, obtaining the fractal dimension was not possible as the effect of tempo manipulation did not provide a concrete reflection in the waveform. Since the fractal patterns largely relied on tempo variations, the desired results were not obtained. The lack of results prompted the use of the Wavelet Transform which allowed for the analysis of time and frequency simultaneously.

This investigation proves that layering an audio with temporally scaled versions of itself does, in fact, increase the audio's self-similarity index and thus its fractal dimension. The fact that the Wavelet Transform has been used to obtain the fractal dimensions shows that the transform can be used to obtain the fractal dimension of any audio through the methodology employed in this investigation.

Works Cited

"But What Is the Fourier Transform? A Visual Introduction." *YouTube*, uploaded by 3Blue1Brown, 26 Jan. 2018, www.youtube.com/watch?v=spUNpyF58BY&t=668s.

"The Fourier Transform." *The Fourier Transform*, www.thefouriertransform.com/. Accessed 17 Mar. 2024.

"Fourier Transforms." *MathWorks*, MATLAB, www.mathworks.com/help/matlab/math/fourier-transforms.html. Accessed 30 Aug. 2024.

"Fractal Dimension." *Boston University Math*, Boston University, math.bu.edu/DYSYS/chaos-game/node6.html. Accessed 30 Aug. 2024.

"Fractals Are Typically Not Self-similar." *YouTube*, uploaded by 3Blue1Brown, 27 Jan. 2017, www.youtube.com/watch?v=gB9n2gHsHN4&t=1040s. Accessed 17 Mar. 2024.

"FRACTALS IN MUSIC - Exploration & Experiment." *YouTube*, uploaded by Woochia Charly Sauret, 22 Mar. 2002, www.youtube.com/watch?v=6XTe5EsPE20&t=176s. Accessed 30 June 2024.

"Hausdorff (Box-Counting) Fractal Dimension." *MATLAB*, www.mathworks.com/matlabcentral/fileexchange/30329-hausdorff-box-counting-fractal-dimension. Accessed 30 Aug. 2024.

"how to interpret the fractal dimesion" prompt. *Phind*, Michael Royzen, 5 July 2024, www.phind.com/search?cache=o4gmmjubdx94a9ldtmbixgrp&source=sidebar.

Li, Guoxi, et al. "Calculation method for fractal characteristics of machining topography surface based on wavelet transform." *Elsevier*. *Science Direct*, http://dx.doi.org/10.1016/j.procir.2019.02.109. Accessed 30 July 2024.

McDonough, John, and Andrzej Herczyński. "Fractal Patterns in Music." *Chaos, Solitons & Fractals*, vol. 170, May 2023, p. 113315, https://doi.org/10.1016/j.chaos.2023.113315. Accessed 30 Jan. 2024.

M.G. Manisha Milani, and Liyanage C. De Silva. "Wavelet Transforms." *Science Direct*, www.sciencedirect.com/topics/computer-science/wavelet-transforms. Accessed 30 July 2024.

9 Amazing Fractals Found in Nature.

"Wavedec." *MathWorks*, MATLAB, mathworks.com/help/wavelet/ref/wavedec.html. Accessed 30 Aug. 2024.

"what is a reference frequency" prompt. *Phind*, Michael Royzen, 8 July 2024, www.phind.com/search?cache=u7u6xlv4a2fcs9bs8wjzgu6f&source=sidebar.

"what is the difference between the detail coefficients and decomposition coefficents in the wavelet transform" prompt. *Phind*, Phind-70B version, Michael Royzen, 9 July 2024, www.phind.com/search?cache=k70qi6f15tyq76eznudw8i0f&source=sidebar.

"Whirpool Galaxy." fractalfoundation.org/OFC/OFC-1-7.html. Accessed 5 Mar. 2024.

Wu, Jiaxin, et al. "An Effective Method to Compute the Box-counting Dimension Based on the Mathematical Definition and Intervals." *Results in Engineering*, vol. 6, June 2020, p. 100106, https://doi.org/10.1016/j.rineng.2020.100106. Accessed 30 Aug. 2024.