A Scientific Theory of Music

Guangyu Xu, CID 00813507

December 19, 2014

Abstract

Music is the art of manipulating sound, which is ubiquitous across human society. In particular, harmony is one of the key phenomena of music. However, current music theory seems to lack the power to explain it. This article intends to explain the mysterious harmony by deriving the basic building block of western music, the Major Scale, via first principles of physics. Along with a few postulates about how the brain "computes" music, a possible explanation for the difference of the emotional quality between different Diatonic Scales and Chords is discussed. This scientific approach to explain music provides an insight into a unified theory of music based on solid science.

Introduction

Music is capable of delivering emotions, but the reason behind this association between a sound and its emotional quality remains a mystery. Music theory books simply describe the empirical rules coined by generations of musicians as a guideline to which note or note combinations sound "right". However, the frustrating fact is that none of them really explains why these rules work. [1] As a physicist, I would prefer a scientific theory of music with *explanatory* power rather than a *descriptive* formula sheet.

While music is a miracle of human creativity, sound is strictly governed by laws of physics. In essence, music is the combinations and arrangements of sound. Therefore it is reasonable to ask for a scientific explanation of music.

In pursuit of a unified scientific theory of music, we will build up the foundation in this article in three steps. We will start with the first principles of physics about the nature of sound in section one. Then five postulates about the mechanism of sound perception will be discussed in section two. In section three, we will derive the Major Scale from harmonic series, whereby the Chromatic Scale could be assembled. A brief explanation of the reason behind the emotional quality difference between different Diatonic Scales[1] and Chords[2] is discussed in the end.

[1] Diatonic Scale is a Scale composed of seven pithes including five whole-steps and two half-steps in each octave, in which the two half-steps are separated by either two or three whole-steps. [2]

[2] Chord is a collection of three or more harmonically related notes. [3]

1 Physics of Sound

1.1 Sound Waves

Sound is a propagating wave triggered by vibrations. The behaviours of these waves are encapsulated in the wave equation,

$$\frac{\partial^2\psi}{\partial t^2} = c^2\frac{\partial^2\psi}{\partial x^2}, \ \psi = \psi(x,t), \qquad (1)$$

where ψ is displacement, c is speed of propagation, x and t are space and time. [4] It is essentially an application of Newton's Second Law, $F = ma$, to a string.

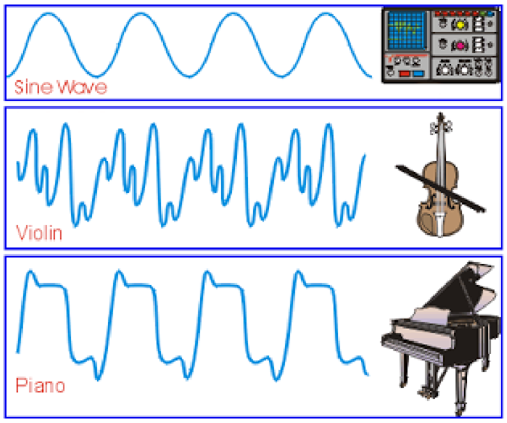

Figure 1: Harmonic Wave and Waveforms of Real Instruments. [5]

Sound wave can be of any form as long as it satisfies the wave equation, e.g., all the waveforms shown in Fig. 1 are valid solutions.

1.2 Fourier Transform

The wave equation is a *linear* differential equation as the highest power of ψ is one. Hence, any linear combination, i.e., sum of its solutions is also a valid solution. Conversely, a complicated waveform can be decomposed into a sum of individual harmonic, i.e., sinusoidal waves, each of different frequencies and amplitudes. [6]

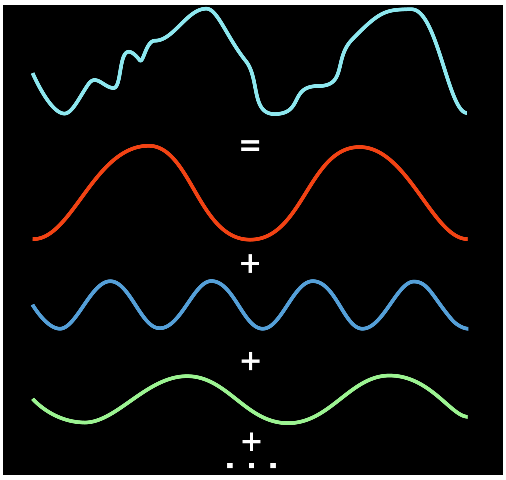

Figure 2: Schematic Diagram of Fourier Transform. [7]

Fourier transform is the algorithm to express an arbitrary waveform as a series sum of harmonic waves.[3] Human auditory system virtually implements a Fourier analyser in the cochlea. [8] So the pitches we hear are the frequencies of the individual harmonic components.

1.3 Harmony in Nature

A harmonic series is defined differently in mathematics and music theory, although they share the same origin. In mathematics, a harmonic series refers to an infinite sum,

$$\sum_{n=1}^{\infty} \frac{1}{n}, \; n = 1, 2, 3 \ldots, \qquad (2)$$

which is derived from the wavelengths of the vibrating modes of a string, shown in Fig. 3. [9] The mode with longest wavelength is called the fundamental. For a string with both ends fixed, its displacement must be zero at both ends. This is a boundary condition which implies that the string can only vibrate at wavelengths that are 1, $\frac{1}{2}$, $\frac{1}{3}$, …of the fundamental wavelength. A real waveform is a mixture of *all* the harmonics of different wavelengths with different amplitudes.

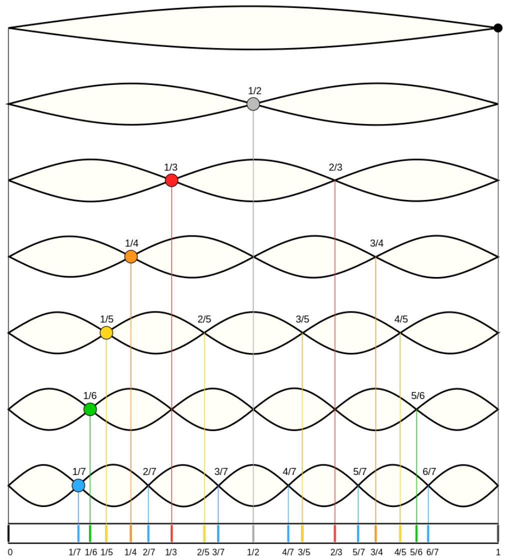

Figure 3: Harmonic Series On String. [11]

Frequency is proportional to the reciprocal of wavelength; hence frequencies of the harmonics are integral multiples of the fundamental frequency,

$$f_n = n f_1, \; n = 1, 2, 3 \ldots, \qquad (3)$$

where f_n is the frequency of the n^{th} harmonic and f_1 is the frequency of the 1^{st} harmonic, i.e., the fundamental. This sequence of frequencies is the harmonic series defined in music theory, which is the definition we will use in this article. [12]

The harmonic series does not only apply to strings, it is omnipresent in nature. In fact, the harmonic series can be regarded as an *abstraction* of the sound of nature. For

[3] For a non-periodic wave, the sum becomes an integral of harmonic waves over all frequencies. [10]

example, the relation is exactly the same for a column of air with both ends open or both closed. For an air column with one end open and the other closed, it only has odd harmonics, i.e., $n = 1, 3, 5 \ldots$ due to its asymmetric boundary conditions. [13]

1.4 Timbre of Instruments

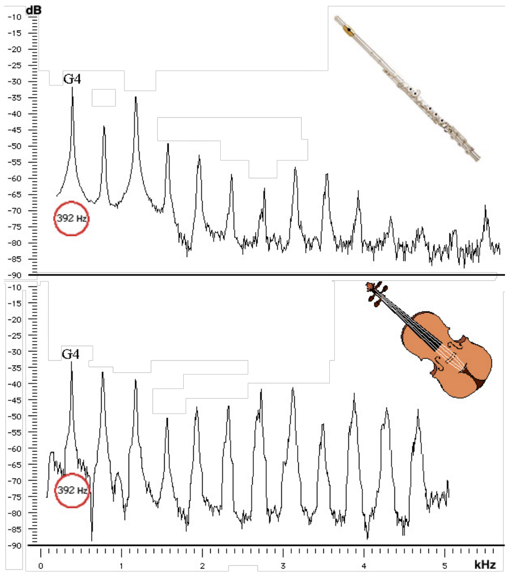

Figure 4: Overtone Profiles of Flute and Violin. [14]

In reality, when a note is played on an instrument, frequencies higher than the fundamental are produced simultaneously, which are called *overtones*. [12]

However, an overtone series does not necessarily resemble a harmonic series exactly, e.g., the overtones of piano strings have higher frequencies than the harmonic series they approximate. [15] It is the overtone profile that differentiates one instrument from another and yields the *timbre* of an instrument. For example, clarinets are characterised by higher energy in odd harmonics while trumpets have relatively even amounts of energy in both odd and even harmonics. [16]

2 Machinery of the Brain

2.1 Harmonic Detector

The brain is a machine optimised by evolution. As the harmonic series is a universal phenomenon in nature, in order to effectively recognise it, the auditory system is likely to be equipped with a harmonic series detector.

Postulate 1 The brain has a hard-wired harmonic detector.[4]

2.2 Nature is Sweet

Virtual pitch is a well-known acoustic phenomenon. If the fundamental of a harmonic series is removed, people will still hear the entire harmonic series including the fundamental. [17] This phenomenon does not only advocate postulate one, but also suggest that the brain is *actively* searching for harmonic series, possibly by reconstructing the fundamental as the greatest common divisor of higher harmonics. It indicates that the brain might enjoy the sound of harmonic series.

Postulate 2 The harmonic series sounds sweet to the brain.

I think many people would agree that the sound of a violin is generally sweeter than a brass instrument. In fact, the overtones of a violin approximate a harmonic series very accurately, in comparison, the approximation of a brass is much poorer. [1]

[4] All five postulates in this section are inspired by Wilkerson (2014). [1]

2.3 Optimisation of Algorithm

The computational power of the brain is very constrained compared to modern computers. To conserve the precious computing power, the algorithm executed by the brain must have been optimised by evolution.

People perceive frequencies whose ratios are integer powers of 2, e.g., 220Hz, 440Hz and 880Hz, as sounding very similar. [18] In fact, they are all called musical note A but one Octave apart. This suggests that the brain might be normalising the input harmonic tone from the cochlea by doubling or halving its frequency, until it is within a specific frequency range spanned by a factor of 2, e.g., 220–440Hz, which is one Octave. The brain could save computing resources consumed for recognising the absolute pitch.

Postulate 3 The brain normalises the input harmonic tone by doubling or halving its frequency until it is within one Octave.

Furthermore, people could recognise a tune no matter how high or low the first note is. As long as the frequency ratios of consecutive notes are the same, people perceive them as the same melody, which is known as the phenomenon of relative pitch. [19] It is an indication that the brain might recognise musical intervals via the frequency ratios of the input tones.

Postulate 4 The brain computes the frequency ratios of the tones and regards the same ratio as the same musical interval.

For example, we hear three consecutive notes C4, G4 and D5, which are about 262Hz, 392Hz and 587Hz respectively. First, the brain normalises D5 by halving its frequency to D4 $\approx$ 293Hz to recognise it as a D note. Then the frequency ratios are calculated and the brain notices that

$$\frac{392\text{Hz}}{262\text{Hz}} \approx \frac{3}{2} \approx \frac{587\text{Hz}}{392\text{Hz}}, \qquad (4)$$

so the interval between C4 and G4 is deemed as the same as the interval between G4 and D5, in which case both are the Perfect Fifth.

2.4 Disambiguation Engine

The real world has a lot of ambiguities that can be understood in multiple ways. The brain implements a disambiguation engine that tries resolve these ambiguities via personal experience and knowledge. [1] For example, a joke often consists of a preceding story introducing some ambiguity and a punchline resolving the ambiguity in the opposite way to common understanding. [20] People enjoy jokes as the disambiguation engine of the brain is teased by surprise.

Postulate 5 The brain enjoys its disambiguation engine being teased.

3 Harmony of Music

3.1 The Major Triad

As the brain enjoys the sound of harmonic series, we shall examine what musical notes harmonic series generate.

Shown in Fig. 5 is the harmonic series with Middle C[5] as the fundamental. The frequency ratio of a harmonic relative to the

[5] Middle C is C4 in scientific music notation, corresponding to 261.6Hz.

Harmonic Order	Musical Note	Normalised Frequency Ratio	Musical Interval
1	C4	1	Unison
2	C5	1	Unison
3	G5	3/2	Perfect Fifth
4	C6	1	Unison
5	E6	5/4	Major Third

Figure 5: Harmonic Series of Middle C.

fundamental is simply just its order, n. By executing normalisation to render all ratios between 1 and 2, i.e., within one Octave, C Major Triad[6] spontaneously arises from the harmonic series, in which the musical intervals corresponding to $\frac{5}{4}$ and $\frac{3}{2}$ are the Major Third and Perfect Fifth respectively. [21]

Interval \ Harmonic	1st	2nd	3rd	4th	5th
Root/Unison	1	2	3	4	5
Major Third	5/4	2*5/4	3*5/4	4*5/4	5*5/4
Perfect Fifth	3/2	2*3/2	3*3/2	4*3/2	5*5/4

Figure 6: Three Harmonic Series Present in the Major Triad.

When the three notes of a Major Triad are played together, there are three harmonic series present at the same time. The frequency ratios relative to the fundamental of the Root are shown in Fig. 6.

By inspecting each column of the table, the structure of intervals is found to be preserved for each order of the harmonics. For instance, the frequency ratio of the 3rd harmonic of the Perfect Fifth relative to the 3rd harmonic of the Root is still $\frac{3}{2}$, a Perfect Fifth. This argument is in fact valid for any chord.

[6]The Major Triad is a chord built on the Root, the Major Third and the Perfect Fifth intervals.

In this particular case, the Major Triad effectively produces a “matrix” of harmonic series, represented both vertically and horizontally in the table. Each column and each row forms a harmonic series itself. According to postulate two, the Major Triad should sound even sweeter than the already sweet harmonic series.

3.2 The Major Scale

The Major Scale is the basic building block of western music and all the other Diatonic Scales can be deemed as the Major Scale starting from different positions.[7] [22] As the Major Fifth is the first musical interval produced by harmonic series, we will start with three notes one Perfect Fifth apart. The Major Scale can then be constructed via three interlocking Major Triads based on these three notes. [1]

To construct the Major Scale, we take an arbitrary note as the Root of the Scale. All intervals in the Scale are calculated relative to the Root, which means all frequency ratios are computed relative to that arbitrary note as shown in Fig. 7. In the first row of the diagram, the ratio of the Root with itself, 1, is multiplied by $\frac{3}{2}$ to get the the note a Perfect Fifth higher and divided by $\frac{3}{2}$ to get the note a Perfect Fifth lower. Then three Major Triads are constructed using the three notes obtained in the 1st row as the Roots of the Triads. Finally the frequency ratios are normalised to be between 1 and 2 and re-ordered ascendingly to get the Major Scale.

In musical notation, intervals are expressed as *additive* distances between two notes.

[7]All Diatonic Scales can be divided into seven Modes, equivalent to the Major Scale starting from its seven different notes. [23]

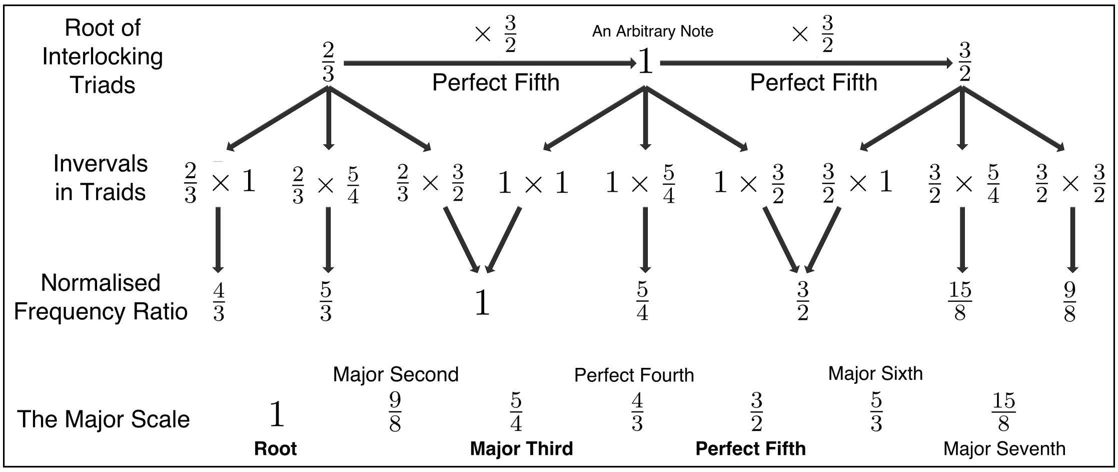

Figure 7: The Major Scale Built on Three Interlocking Triads.

However, the intervals perceived by the brain are based on *multiplicative* ratios. [1] To turn multiplicative factors into additive increments, the only method is logarithm. In Fig. 8, the frequency ratios are plotted against their logarithm with respect to base 2.

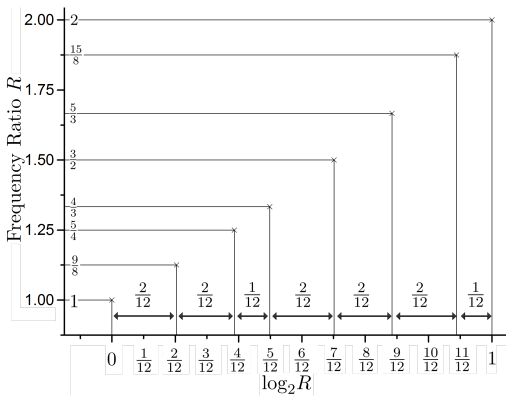

Figure 8: Intervals of Major Scale on Logarithm Space.

Now some old familiar patterns are formed on the $\log_2 R$ axis. Specifically, the separations between consecutive notes are roughly $\frac{2}{12}, \frac{2}{12}, \frac{1}{12}, \frac{2}{12}, \frac{2}{12}, \frac{2}{12}, \frac{1}{12}$. If the intervals between 2 adjacent white keys on a piano are

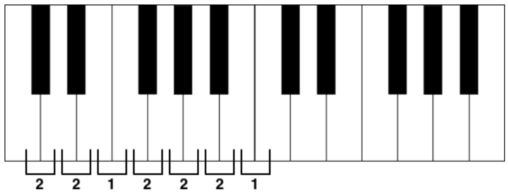

Figure 9: Two Octaves on Piano Keyboard.

written in half-steps, as shown in Fig. 9, the above two patterns precisely resemble each other. Indeed, the white keys on a piano are the notes of C Major Scale.

3.3 The Chromatic Scale

The black keys on a piano are placed in-between the consecutive white keys with wider intervals, i.e., those separated by $\frac{2}{12}$ on $\log_2 R$ axis. There are a total of twelve keys on a piano within one Octave corresponding to the Chromatic Scale, upon which all western music is written.

The Chromatic Scale can be constructed by combining the Major Scale with its inversion. If all the intervals are calculated relative the last note in the Major Scale, i.e.,

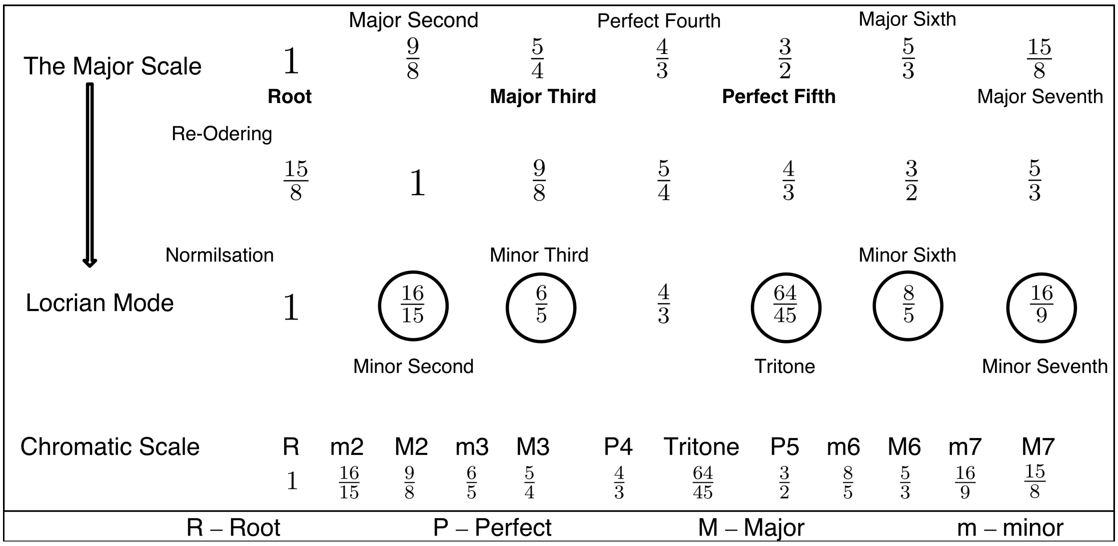

Figure 10: Construction of Chromatic Scale via Locrian Mode.

the Major Seventh, a Scale called Locrian Mode is generated. Shown in Fig. 10, the notes in the Major Scale are re-ordered and then normalised with respect to the Major Seventh, $\frac{15}{8}$, resulting in the Locrian Mode. There appear five new intervals circled in the diagram. Finally the Chromatic Scale is built by collecting all the intervals present in the two Scales.[8]

3.4 Emotional Quality

The Major Scale presents a bright and happy feeling as it pleases the brain with three interlocking Major Triads built on harmonic series. It is the Scale that the brain expect to hear naturally. However, if the starting note is changed, the *pairwise* intervals will remain the same but appear in different positions in other Diatonic Scales.

For example, the Natural Minor Scale, or Aeolian Mode, is built by the same method as we build the Locrian Mode but now the 6^{th} note of the Major Scale is used as the Root. The brain could recognise the Natural Minor as a scale built on harmonic series, but the order of intervals is "wrong".

According to postulate five, the deviation from the expected "right" order teases the disambiguation engine of the brain, which might be the reason behind the quality difference between different Diatonic Scales. The Minor Scale sounds a bit "off" or dissonant, but still interesting to the brain. In fact, there exist several Minor Scales including the Natural Minor, Harmonic Minor and Melodic Minor. It might be due to the fact that while there is only one way to be right, there can be many ways to be wrong. [1]

A Chord is constructed with three or more notes from the Chromatic Scale, which can be thought as a miniature Scale. Similar to the argument for Diatonic Scales, the

[8] There are different ways to construct the Chromatic Scale with just intonation, which lead to similar ratios. This method is my original work.

	R	m2	M2	m3	M3	P4	Tritone	P5	m6	M6	m7	M7	P8
Just Intonation	1	1.067	1.125	1.200	1.250	1.333	1.422	1.500	1.600	1.667	1.778	1.875	2
Equal Temperament	1	1.059	1.122	1.189	1.260	1.335	1.414	1.498	1.587	1.682	1.782	1.888	2

Figure 11: Just Intonation vs. Equal Temperament.

combination and ordering of different intervals in a Chord might be the source of its quality. For example, the C Major Seventh chord is C-E-G-B containing two Triads. [1] C-E-G constitute C Major Triad in which the intervals are in the "right" order, while E-G-B forms E Minor Triad in which, the intervals are ordered "incorrectly".

3.5 Just and Equal Tunning

In the above argument, an instrument is assumed to be tuned according to *just intonation*, which means the intervals are adjusted to the exact ratios in the Chromatic Scale shown in Fig. 10. But this tuning method has a severe problem in application. Assuming an instrument is tuned using C as the Root, the harmony will be perfect for a piece in Key of C. However, the harmony will break down when a piece in other Keys is played.[9] For example, the "Perfect" Fifth in D Major Scale will correspond to the interval between the 6$^{\text{th}}$ and 2$^{\text{nd}}$ note in C Major Scale, corresponding to $\frac{5}{3} \div \frac{9}{8} \approx 1.48$, which has a noticeable deviation from the "perfect" ratio, $\frac{3}{2}$. This problem becomes more prominent with Keys further away from the Key according to which the instrument is tuned as the deviations are accumulative.

To mitigate the problem of just intonation, *equal temperament* is often used as a compromise between the technical limitations and pursuit of harmony. All the intervals between consecutive notes in the Chromatic Scale are tuned equally to $2^{1/12} \approx 1.059$. As a result, it makes none of the Keys sound perfect, but it allows every Key sound acceptable. [24] A comparison between frequency ratios in equal temperament and the just intonation derived in this article is shown in Fig. 11.

[9]The Key of a piece is the tonic note, which is usually the Root of the scale used.

4 Summary

In pursuit of a scientific music theory, we started from the first principles of physics and derived the Major Triad from the abstraction of natural sound, harmonic series. Then the Major Scale was constructed via three interlocking Major Triads. Subsequently, Locrian Mode was constructed by re-ordering the Major Scale and the Chromatic Scale was assembled using the intervals present in both the Major Scale and Locrian Mode. Finally, a possible explanation of the emotional quality of Diatonic Scales and Chords is discussed.

Although this theory of music is far from complete, at least it provides some insight into the viability of a scientific approach towards a unified music theory.

References

[1] Wilkerson, D.S. (2014) *Harmony Explained: Progress Towards A Scientific Theory of Music.* arXiv:1202.4212v2 [cs.SD].

[2] Johnson, T.A. (2008) *FOUNDATIONS OF DIATONIC THEORY: A Mathematically Based Approach to Music Fundamentals.* The Scarecrow Press, Inc.

[3] Benward, B., Saker, M. (2003) *Music in Theory and Practice, Vol. 1.* 7th edition. McGraw-Hill. 67.

[4] French, A.P. (1971) *Vibrations and Waves.* Massachusetts: M.I.T. Press.

[5] Jhavar, S. (2010) *Design: Physics of Sound.* [online] Scotland: D stands for Design. Available from: http://swastidesign.blogspot.co.uk/2010/04/physics-of-sound.html [Accessed 5 December 2014].

[6] Körner, T. W. (1988) *Fourier Analysis.* Cambridge, England: Cambridge University Press.

[7] Hardesty, L. (2012) *The faster-than-fast Fourier transform.* [online] PHYS.ORG. Available from: http://phys.org/news/2012-01-faster-than-fast-fourier.html [Accessed 5 December 2014].

[8] Wolfe, J.M., Kluender, K.R., Levi, D.M., Bartoshuk, L.M., Herz, R.S., Klatzky, R.L., Lederman, S.J., Merfeld, D.M. (2012) *Sensation and Perception.* 3rd edition. Massachusetts: Sinauer Associates, 15-17.

[9] Arfken, G. (1985) *Mathematical Methods for Physicists.* 3rd edition. Orlando: Academic Press, 279-280.

[10] Riley, K.F., Hobson, M.P., Bence, S.J. (2006) *Mathematical Methods for Physics and Engineering.* 3rd edition. New York: Cambridge University Press, 422-423.

[11] Landman, Y. (2010) *Just intonation.* [online] Wikimedia Commons. Available from: http://commons.wikimedia.org/wiki/File:Moodswingerscale.jpg [Accessed 5 December 2014].

[12] Bain, R. (2003) *The Harmonic Series.* [online] Columbia, USA: University of South Carolina. Available from http://in.music.sc.edu/fs/bain/atmi02/hs/hs.pdf [Accessed 9 December 2014].

[13] Nave, R. (2012) *Resonance Concepts.* [online] HyperPhysics. Available from http://hyperphysics.phy-astr.gsu.edu/hbase/sound/rescon.html#c1 [Accessed 5 December 2014].

[14] Anon. (2012) *All of Rhythm.* [online] Evolutional Groove. Available from http://www.groovequest.ca [Accessed 6 December 2014].

[15] Livelybrooks, D. (2007) *Physics of Sound and Music.* [online] University of Oregon. Available from http://hendrix2.uoregon.edu/ dlivelyb/phys152/L16.html [Accessed 9 December 2014].

[16] Levitin, D.J. (2006) *This Is Your Brain on Music: The Science of a Human Obsession.* New York: Dutton.

[17] Terhardt, E. (1974) *Pitch, consonance, and harmony.* Journal of the Acoustical Society of America, 55(5), 1061-1069.

[18] Braun, M. (2010) *The octave - history of a discovery.* [online] Neuroscience of Music. Available from http://www.neuroscience-of-music.se/Octave-History.htm [Accessed 10 December 2014].

[19] Zatorre, R.J., Perry, D.W., Beckett, C.A., Westbury, C.F., Evans, A.C. (1998) *Functional Anatomy of Musical Processing in Listeners with Absolute Pitch and Relative Pitch.* Proceedings of the National Academy of Sciences of the United States of America, 95, 3172-3177.

[20] Lew, R. (1996) *Exploitation of linguistic ambiguity in Polish and English jokes.* Papers and Studies in Contrastive Linguistics, 31, 127-133.

[21] Tobey, F. (2012) *A Feeling for Harmony.* [online] Earlham College. Available from: http://legacy.earlham.edu/∼tobeyfo/musictheory/Book1/home1.html [Accessed 11 December 2014].

[22] Hewitt, M. (2013) *Musical Scales of the World.* The Note Tree.

[23] Allen, A. (2014) *Modal Diatonicism 3: Harmonic and Melodic Music Theory and Method for the 21st Century.* CreateSpace Independent Publishing Platform.

[24] Gann, K. (1997) *Just Intonation Explained.* [online] Kyle Gann's Home Page. Available from: http://www.kylegann.com/tuning.html [Accessed 16 December 2014].