Research Article

The Science of Harmony: A Psychophysical Basis for Perceptual Tensions and Resolutions in Music

Paul Yaozhu Chan [iD] [1,2]**, Minghui Dong**[1]**, and Haizhou Li**[1,3]

[1]*Institute for Infocomm Research, A∗STAR, Singapore*
[2]*Nanyang Technological University, Singapore*
[3]*National University of Singapore, Singapore*

Correspondence should be addressed to Paul Yaozhu Chan; ychan@i2r.a-star.edu.sg

This paper attempts to establish a psychophysical basis for both stationary (tension in chord sonorities) and transitional (resolution in chord progressions) harmony. Harmony studies the phenomenon of combining notes in music to produce a pleasing effect greater than the sum of its parts. Being both aesthetic and mathematical in nature, it has baffled some of the brightest minds in physics and mathematics for centuries. With stationary harmony acoustics, traditional theories explaining consonances and dissonances that have been widely accepted are centred around two schools: rational relationships (commonly credited to Pythagoras) and Helmholtz's beating frequencies. The first is more of an attribution than a psychoacoustic explanation while electrophysiological (amongst other) discrepancies with the second still remain disputed. Transitional harmony, on the other hand, is a more complex problem that has remained largely elusive to acoustic science even today. In order to address both stationary and transitional harmony, we first propose the notion of interharmonic and subharmonic modulations to address the summation of adjacent and distant sinusoids in a chord. Based on this, earlier parts of this paper then bridges the two schools and shows how they stem from a single equation. Later parts of the paper focuses on subharmonic modulations to explain aspects of harmony that interharmonic modulations cannot. Introducing the concept of stationary and transitional subharmonic tensions, we show how it can explain perceptual concepts such as tension in stationary harmony and resolution in transitional harmony, by which we also address the five fundamental questions of psychoacoustic harmony such as why the pleasing effect of harmony is greater than that of the sum of its parts. Finally, strong correlations with traditional music theory and perception statistics affirm our theory with stationary and transitional harmony.

1. Introduction

Even though it is one of the most important components in music, and possibly the most widely studied [1], the definition of harmony differs vastly across time, genre, and individuals, reflecting how little is understood about it [2, 3].

There are three aspects to the complete understanding of our perception of harmony, which we will, for brevity, refer to as *what*, *why*, and *when*. The *what* of harmony refers to an attribution to a defining quality. Its *why* goes further to explain the means by which such a quality ascribes to consonance or dissonance (or even sentiment or emotions). Finally, it should be recognized that the same harmony perceived as consonant in one context can be perceived as dissonant in another. This takes the *what* and *why* of stationary harmony (sonorities) into the context of transitional

harmony (progression). We refer to this as the *when* of harmony and it has remained largely unaddressed by acoustic science.

1.1. Background. Early works effectively attributed the *what* of harmony to rational relationships [1, 4]. This ascribes a chord's consonance to the ratio amongst its contributing string lengths (and consequently, wave periods and fundamental frequencies), being fractional with integer numerators and denominators. A fascinating number of esteemed mathematicians, physicists, and philosophers have made different contributions in this aspect. The development of the Pythagorean tuning system is commonly credited to Pythagoras in the fourth century BC [3, 5, 6]. Euclid wrote the earliest surviving record on the tuning of the monochord [7] and documented numerous experiments on rational tuning

[8]. Aristotle and Plato made various contributions to the development of ancient Grecian (rationally scaled) music that was later integrated into the diatonic system [8, 9]. Ptolemy developed the syntonic diatonic system as early as the second century [10]. Euler proposed a grading system of chord aesthetics based on the assertion that the notes have a least common multiple (i.e., that they are rational) [11]. Since string lengths correspond to wavelengths, which correspond to wave period, and since notes used in harmony are taken from the scale, it can be said that the Pythagorean school effectively attributes harmony to temporal features.

It was not until 1877 that Helmholtz pioneered the psychoacoustic approach [3, 8, 12, 13]. Isolating adjacent harmonic sinusoids from different notes using specifically devised acoustic resonators, he was able to record how amplitude modulation that resulted from their summation grew perceptually unpleasant as their modulation frequency increased towards a certain threshold [8], thus attributing dissonance to what he called beating frequencies and addressing the questions *what sounds bad and why*. Numerous others [14–25] conducted further studies in this approach, while others raised several questions with Helmholtz's theory [13, 17, 26]. For example, Plomp and Levelt [12] and Schellenberg and Trehub [27] have separately shown that consonances and dissonances are still perceived in harmonies with pure tones (tones without harmonics). Itoh [28] and Bidelman [29], amongst others, also showed that electrophysiological responses to pure-tone intervals did not agree with Helmholtz. All in all, the Helmholtz school attributes harmony to frequency features and comprises a large part of what is referred to in this paper as interharmonic modulations.

In 1898, a notable but short-lived [3] attempt at *what sounds good and why* was seen in Stumpf's tonal fusion theory [30], which theorized that harmony was the effect the harmonics of its component notes fusing together to sound like a single note with a common fundament [12, 13, 26, 30].

Because of the nonlinear relationship between tonal scale and frequency, scales derived from rational lengths of a string tended to leave certain intervals more rational than others. With this realization, Western music eventually adopted 12-tone equal temperament scale. This equally segments the octave in the log-frequency scale [31] such that each semitone interval is a factor of $2^{1/12}$, evenly redistributing the dissonances to accommodate to different keys. Despite its late adoption, original development of this scale predates Helmholtz to the 1500s. Vincenzo Galilei (father of Galileo Galilei) made the earliest known estimate of this in the West by approximating $2^{1/12}$ with 18/17 [32], while Zhu was credited for perfecting it in the East by computing it to accurately to the 25th decimal, both in the 1580s [12]. The earliest recorded estimate of this in the East was by He in the 5th century, whose estimate was already about as accurate as Galilei's [33, 34].

In Rameau's Treatise on Harmony [1], which paved the foundations of harmony in modern music theory, notes of basic chords are derived from the division of the length of a common string [35]. However, this remains disjoint with the rest of the treatise, and modern music theory remains more of a compilation of rules and deductions from the pattern clustering of perceptual experiences [36–42], addressing the questions *what sounds good and when* without the scientific reasoning of *why* [37].

More recently, several studies have found high correlations between harmony and periodicity measures of the resultant signal [43, 44]. This novel leap advances the Pythagorean school while presenting a persuasive attribute of *what sounds good and why*.

Several notable studies have also been conducted that relate harmony to nonacoustic attributes such as statistics and geometry. An example is Tymoczko's exploration of how multidimensional geometric patterns correlate strongly with patterns that exist in historic harmony use, addressing *what sounds good and when* [45–47]. Authors in [48] explored properties of musical scales on the Euler lattice, addressing the *what* of harmony. Numerous others such as [49–51] have worked on other mathematical relationships in harmony, addressing its *what*.

Yet others have looked towards a biological rationale towards our perception of harmony to address *what sounds good and why*. A recent example is Purves' attribution of the effect of the tonal scale to the familiarity of excited or subdued speech [14, 52–54]. Other examples are the works of [43, 55, 56] in the neuronal mechanism of harmony perception.

1.2. Scope. In this work, we first seek a mathematical resolution across both acoustic schools by a single psychophysical theory. To start off somewhere familiar, we first describe the concept of *interharmonic modulations* (which adopts and encompasses Helmholtz's beating frequencies), from which we then introduce the concept of *subharmonic* [57] *modulations* and show how the two categories of modulations relate. (At some point after which, we also show how a specific case of subharmonic modulations addresses Pythagoras, thus integrating the two schools.) After explaining how perceptual tensions [18, 36, 58, 59] in musical harmony may be identified in subharmonic tension in the stationary context, we continue to explain how perceptual tension resolutions [18, 42] in transitional harmony (chord progressions) may be visualized in subharmonic trajectories. By these, we address the *what*, *why*, and *when* of harmony. Numerical results show strong to near-complete correlations with perception and chord-use statistics that are presented towards the end of the paper.

By applying our theory and equations, we will answer the five fundamental questions of psychoacoustic harmony. These are as follows.

(1) The phenomenon that the effect of harmony is greater than the sum of its parts [18, 60]:

$$\varepsilon\{x_1 + x_2 + x_3\} \gg \varepsilon\{x_1\} + \varepsilon\{x_2\} + \varepsilon\{x_3\} \qquad (1)$$

where ε denotes *the harmonious effect of* x_1, x_2, and x_3 representing notes of the chord and '+' denotes simultaneous presentation or cumulation.

(2) There are the definition and explanation of stationary harmony, i.e., *what sounds good* and *why*, or, mathematically, to quantify $\varepsilon\{X_n\}$, where X_n represents chord *n*.

(3) There are the definition and explanation of transitional harmony, i.e., *what sounds good*, *why*, and *when*, or, mathematically, to quantify $\varepsilon\{X_1 \longrightarrow X_2\}$, where ' $\longrightarrow$' denotes transition from one chord to another.

(4) We have the following phenomena.

   (a) A chord that sounds better than another out of context can sound worse than being in context [42]. Given $\varepsilon\{X_2\} > \varepsilon\{X_3\}$ this shows that $\varepsilon\{X_1 \longrightarrow X_2\} < \varepsilon\{X_1 \longrightarrow X_3\}$

   (b) A chord that sounds better than another in one context can sound worse than being in another context [42]. Given $\varepsilon\{X_4 \longrightarrow X_2\} > \varepsilon\{X_4 \longrightarrow X_3\}$ this shows that $\varepsilon\{X_1 \longrightarrow X_2\} < \varepsilon\{X_1 \longrightarrow X_3\}$

(5) We have the phenomenon that the transition from a low-tension chord to a high-tension one can still bring about the effect of tension release (resolution). Given $\varepsilon\{X_1\} < \varepsilon\{X_2\}$ this shows that $\varepsilon\{X_1 \longrightarrow X_2\} > 0$

Apart from Pythagoras [3, 5, 6] and Helmholtz [8], we will, in closing, also briefly explain how our theory mathematically bridges other subsidiary psychophysical theories such as Stumpf [30], Euler [11], Galilei [33, 34, 61], and Zhu [12].

2. A Universal Theory of Harmony

In this section, a psychophysical basis for harmony is proposed as follows.

The human perception of harmony is composed of auditory events produced by the combination of sinusoids that make up each note in the harmony. These may be classified into interharmonic and subharmonic modulations.

First-order interharmonic modulations are those produced by the interplay amongst adjacent sinusoids across differing notes. These are loosely categorized by the frequency of the resultant amplitude modulation into dissonant beating frequencies [8] and consonant low-frequency modulations, triggering a variety of emotions according to their modulation and carrier frequencies. Second-order interharmonic modulations are produced by the alignment of first-order ones. The consonance types of different intervals may be identified according to patterns cast by interharmonic modulations on the interharmonic plot.

Despite the significance of interharmonic modulations, the effect of consonances and dissonances is still experienced in the absence of harmonics with pure tone harmonies. This implies that interharmonic modulations are not exclusive in our perception of harmony [12, 13, 17, 26–29]. From this, it may be deduced that subharmonic modulations also play a significant role.

Subharmonic modulations are produced by the interplay of sinusoids much further apart than interharmonic modulations. Unlike interharmonic modulations, which are analysed primarily in the frequency domain, subharmonic modulations are analysed primarily in the temporal domain and they are comprised of two parts. The first part is subharmonic wave formation, which occurs with the summation of component waveforms from each note to produce a waveform largely periodic to a common subharmonic frequency. The second is subharmonic wave deformation (an example is provided in Supplementary Video S1.), which is a distortion to every successive period of this composite subharmonic waveform due to the imperfect alignment of contributing wave periods. Stationary tension and transitional resolution may both be derived from subharmonic features which serve as measures of stationary and transitional harmony.

In order to explain interharmonic and subharmonic modulations in detail and how they unify the two prevailing schools of harmony, we will start from first principle by looking at the notes of a chord as the sum of their composite sinusoids.

2.1. Modulations in Sinusoidal Summation. When waveforms of two notes, $x_1(t)$ and $x_2(t)$, at amplitudes α and β, respectively, are presented together, the result may be expressed as a sum of their composite sinusoids such that

$$
\begin{aligned}
\alpha x_1(t) + \beta x_2(t) = \alpha \sum_{n=1}^{N} q_n \cos\left(2\pi n f_1 t + \rho_n\right) \\
+ \beta \sum_{m=1}^{M} r_m \cos\left(2\pi m f_2 t + \varphi_m\right)
\end{aligned}
\tag{2}
$$

where, respectively, n and m represent the individual harmonics from each note, N and M represent the highest harmonics that need to be considered because of audible range, q_n and r_m represent the amplitude coefficients of each harmonic, nf_1 and mf_2 represent the frequencies of each harmonic with f_1 and f_2 representing the fundamental frequency of each note, ρ_n and φ_m represent the starting phases of each harmonic, and t represents monotonically increasing time.

Isolating a single pair of adjacent sinusoids from differing notes we get

$$
A h_1(t) + B h_2(t) = A \cos \omega_1 t + B \cos \omega_2 t
\tag{3}
$$

where $h_1(t)$ and $h_2(t)$ are the pair of harmonics from differing notes, $A = \alpha q_n$, $B = \beta r_m$, $\omega_1 = 2\pi n f_1$, and $\omega_2 = 2\pi m f_2$.

Since we are considering the modulating frequency resultant of the summation of both sinusoids spanning all phase combinations, it no longer matters which starting phase we take reference from. Hence, ρ_n and φ_m can both be set to zero.

In the case of A=B, the resultant amplitude modulation is trivial and, as illustrated in Figure 1 (left), is given by the sum-to-product rule

$$
\cos \omega_1 t + \cos \omega_2 t = 2 \cos \frac{\Delta\omega}{2} t \cos \overline{\omega} t
\tag{4}
$$

where $\Delta\omega/2$ is the normalized modulating frequency and is given by

$$
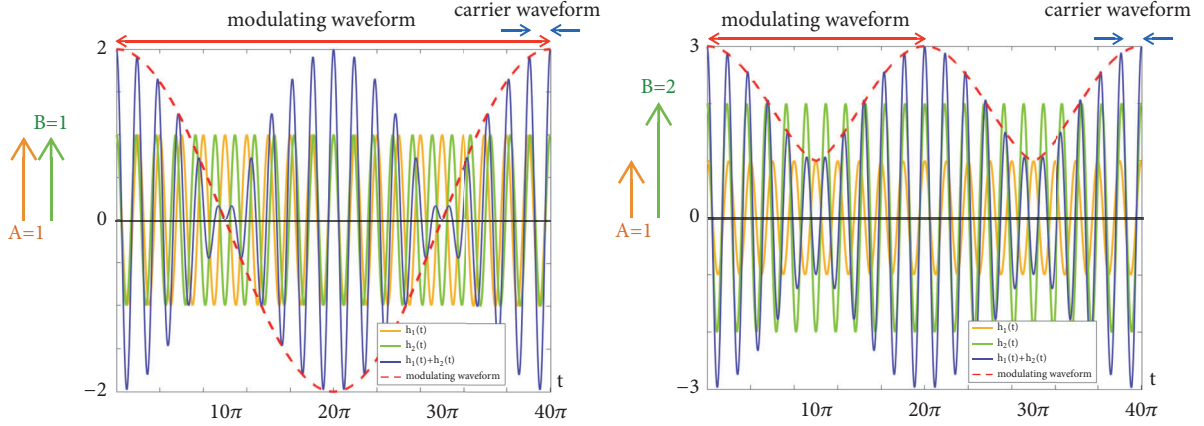\frac{\Delta\omega}{2} = \frac{|\omega_1 - \omega_2|}{2}
\tag{5}
$$

FIGURE 1: Summation of sinusoids of equal (left) versus unequal (right) amplitudes. Notice the difference in modulating frequencies even though frequencies of component sinusoids remain unchanged.

$\overline{\omega}$ is the normalized carrier frequency given by

$$\overline{\omega} = \frac{\omega_1 + \omega_2}{2} \quad (6)$$

and the values of A and B are normalized to 1.

However, in most cases, $A \neq B$, and the problem becomes nontrivial, because of the change in modulation frequency as the modulating waveform no longer crosses zero. This can be seen in Figure 1 (right).

We approximate the summation of these sinusoids to be

$$A \cos \omega_1 t + B \cos \omega_2 t \approx B - A$$
$$+ 2A \left\| \cos^{2-A/B} \frac{\Delta\omega}{2} t \right\| \cos \omega_c t \quad (7)$$

where ω_c is bounded by ω_1 and ω_2 and is approximated to be $\overline{\omega}$ (which denormalizes to $\overline{f}$); $\|\cos^{2-A/B}(\Delta\omega/2)t\|$ denotes the magnitude of $\cos^{2-A/B}(\Delta\omega/2)t$ signed according to the quadrant of $(\Delta\omega/2)t$. B denotes the larger of the amplitudes and A and B are normalized to $A = 1$.

When $A = B$, this simplifies to (4), where the modulating frequency is $\Delta\omega/2$.

However, as B increases with respect to A, $2 - A/B$ gravitates towards 2, and

$$A \cos \omega_1 t + B \cos \omega_2 t \approx B - A + 2A \cos^2 \frac{\Delta\omega}{2} t \cos \omega_c t \quad (8)$$
$$\approx B + A \cos \Delta\omega t \cos \omega_c t$$

for which the modulating frequency is $\Delta\omega$.

We can see from the plots in Supplementary Figure S1 that this estimation is accurate for values of B marginally larger than A to much larger than A.

For consistency, the effective modulating frequency for the case of $A = B$ will be considered by the frequency of its rectified modulating waveform which is then, similarly, $\Delta\omega$. In music, we are interested in this frequency in hertz. Hence, we denormalize this to be

$$\Delta f = |f_1 - f_2| \quad (9)$$

In the next two sections, we will move on to see how this is applicable not only to the summation of adjacent harmonics in interharmonic modulations but also to distant sinusoids in subharmonic modulations.

3. Interharmonic Modulations

Interharmonic modulation refers to modulations across adjacent pairs of sinusoids from different notes that fall within a certain threshold, with modulation frequency corresponding to Δf in (9).

Figure 2 shows a plot of all harmonics of notes c_3 (blue) and e^b_3 (red) under 3 kHz. All adjacent sinusoids less than 120 Hz apart are identified in the figure, with their centre, $\overline{f}$, and modulating, Δf, frequencies labeled accordingly.

3.1. Beating Frequencies and Low-Frequency Modulations. Interharmonic modulations with Δf that increase towards a certain threshold are known to become increasingly dissonant, and, as coined by Helmholtz, are known as beating frequencies [8]. Interharmonic modulations with small Δf, on the other hand, contribute to the harmonious effect perceived in consonance [65]. Figure 3 illustrates this.

3.2. Perceptual Responses across the Δf-$\overline{f}$ Feature Space. It is known that different combinations of notes contribute to different emotive valences [66]. This too may be decomposed into a sum of its harmonics. Hence, further to the consonances and dissonances, emotive responses may also be mapped onto the interharmonic plot. Although, as one might imagine, such responses would be different for every individual, we can plot the response for an individual as an example. Figure 4 shows an example of auditory responses triggered in the mind of the (first) author when exposed to frequencies in the horizontal ($\overline{f}$) axis modulated by frequencies in the vertical (Δf) axis. The value of $\overline{f}$ is indicated in the horizontal axis in both Hz and its corresponding note names. The degree of pleasure derived from interharmonic modulation is coded in the colored background as a reference. The green regions
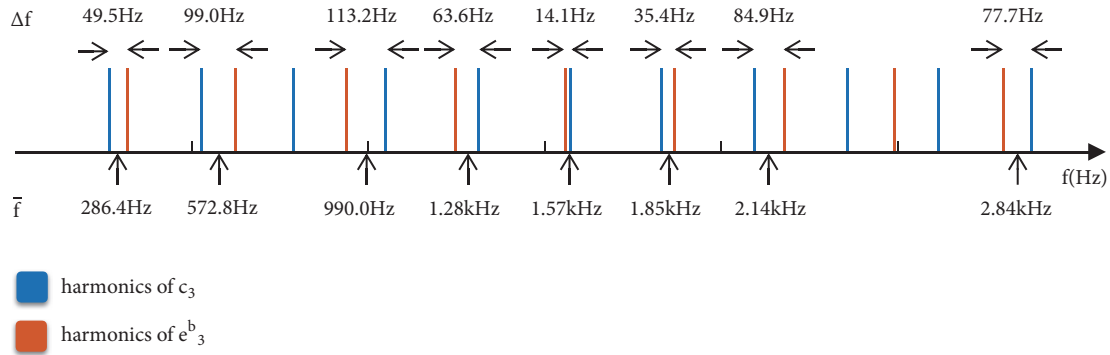
FIGURE 2: Identifying the interharmonic modulations across c_3 and e^b_3.

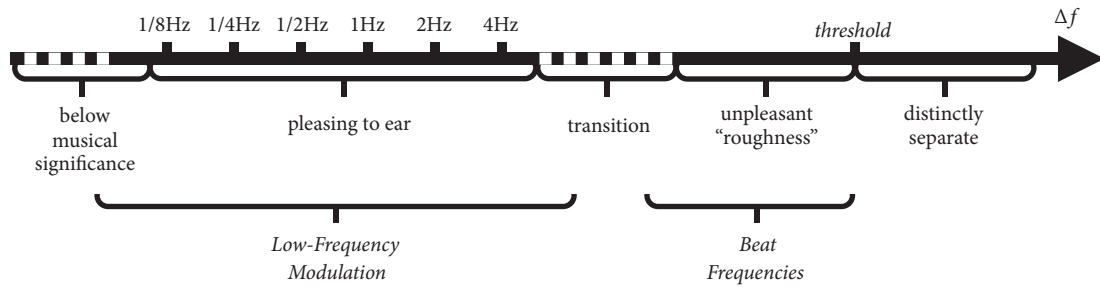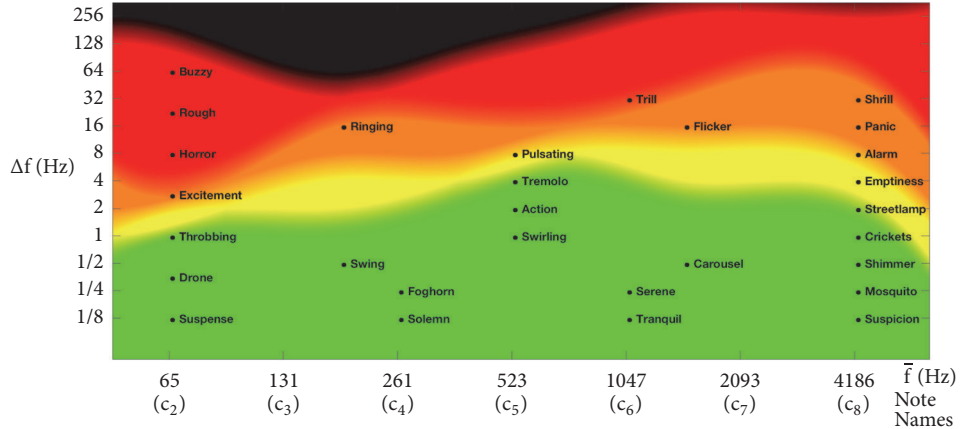FIGURE 3: Types of interharmonic modulation on the scale of Δf.

FIGURE 4: Example of auditory responses triggered by pure-tone frequencies on the horizontal axis modulated at frequencies on the vertical axis. Green, yellow, orange, red, and black indicate pleasing, somewhat pleasing, unpleasant, dissonant, and beyond beating range, respectively.

are perceived to be pleasing, yellow as somewhat pleasing, orange as unpleasant, but not to the point of annoying, red as dissonant, and black as beyond beating range. The black dots mark the locations of the thoughts or emotions labelled. This shows that interharmonic modulations bring about a large variety of thoughts or emotions. If several of these are triggered simultaneously when just one pair of notes sound simultaneously, one can imagine how ten fingers on a piano or all the instruments in an orchestra could combine several (thoughts or emotions) to paint stories on the interharmonic feature-space over time.

3.3. *Intervals and Second-Order Modulations on the Δf-$\overline{f}$ Feature Space.* The interharmonic modulations of each interval within an octave are similarly plotted in Figures 5, 6, and 7. However, this time, the plots are in the linear scale. Green, yellow, orange, and red, again, represent regions of different degrees of consonance or dissonance according to the same color scheme as Figure 4. However, because this time both horizontal and vertical axes are in the linear scale, consonance-dissonance levels that populate the space on the nonlinear plot in Figure 4 now populate lower right regions of these linear plots. The remaining upper left regions are
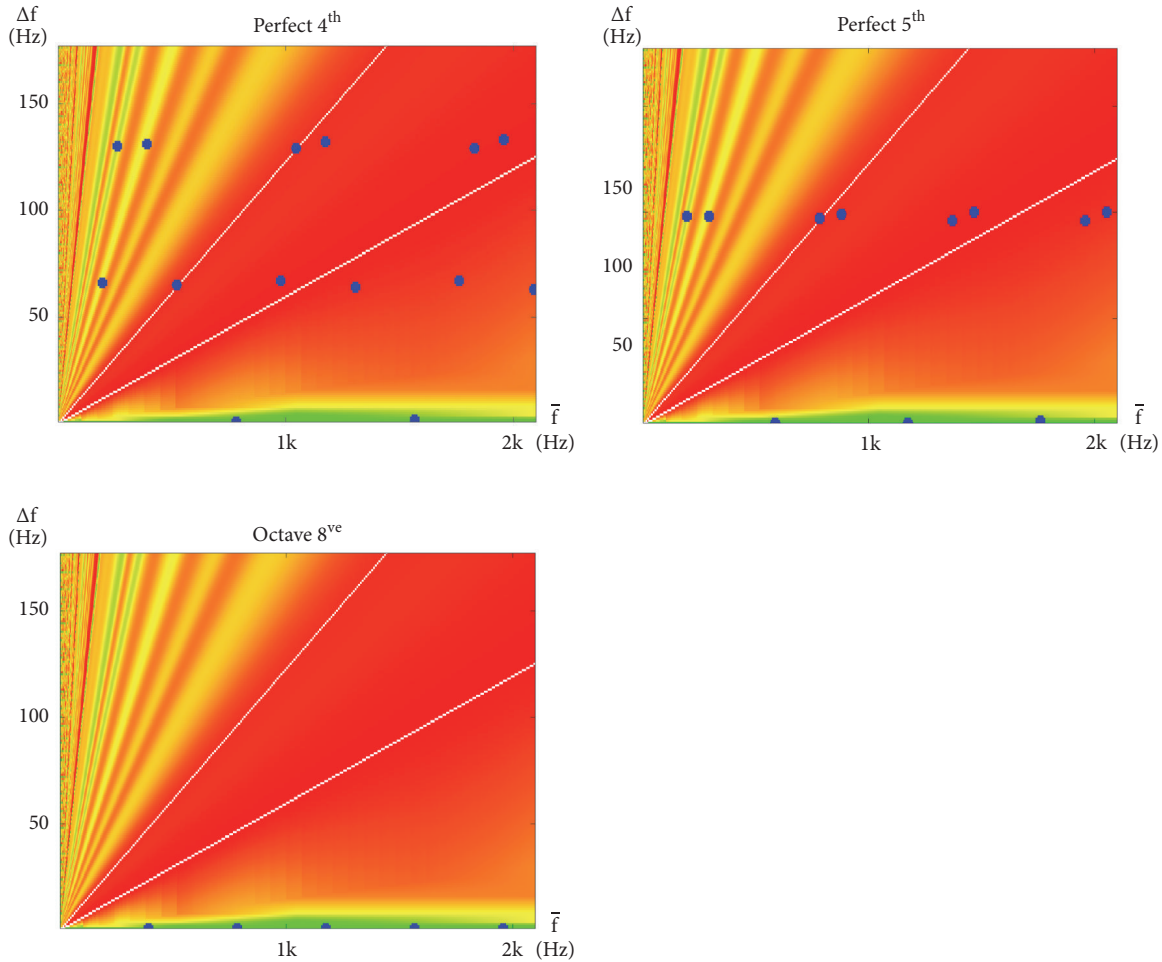
FIGURE 5: Interharmonic plots for all intervals within an octave regarded, in classical music theory, to be perfectly consonant with a root of g_3. These are, namely, the Perfect 4th (g_3 and c_4), Perfect 5th (g_3 and d_4), and Octave (g_3 and g_4) intervals.

then populated with dissonance levels from [12]. These colors provide a simple background reference for the dark blue dots that each represent a modulation at their corresponding Δf and $\overline{f}$ values, which results from the summation of neighboring pairs of sinusoids (at frequencies $\overline{f} + \Delta f/2$ and $\overline{f} - \Delta f/2$) of the notes specified by the indicated interval. Also, for reference, are the two white lines that run across each plot, indicating the locations where the values of Δf coincide with a semitone (gentler slope) and a tone (steeper slope) of the corresponding values of $\overline{f}$ (where $\Delta f = (2^{1/12} - 1)\overline{f}$ and $\Delta f = (2^{2/12} - 1)\overline{f}$, resp.). The semitone and the tone are regarded as the most dissonant intervals up to halfway in either direction around the cyclic chroma [12, 21, 54].

The plots of perfect consonances are presented in Figure 5. These intervals are described with a bit of a dilemma in classical music theory [67]. They may be described as so consonant that they sound almost like one note. As such, their use contributes in a limited way to harmony [15]. For example, the use of perfect fifths is forbidden in parallel motion and

octaves are regarded as the same note in a different register [42].

The interharmonic plot reveals the perceived traits of each category of intervals in a way that explains why they sound the way they do, and in a way music theory alone has never been able to. As shown in Figure 5, the constellations formed by interharmonic modulations of perfect intervals line up almost horizontally (While the methods used in this study are applicable with any form of tuning, only equitempered tuning is assumed in the computations in this section. This is consistent throughout this paper, unless otherwise stated.). Since each point that falls on the same horizontal has the same Δf, this means that they modulate synchronously and may be perceived collectively as a single modulation. This may be interpreted as fewer modulating microevents taking place, making them less interesting than other consonance intervals.

Dissonant intervals are presented in Figure 7. As can be seen in the figure, these intervals have points that fall mostly within the central dissonant region and line up along the
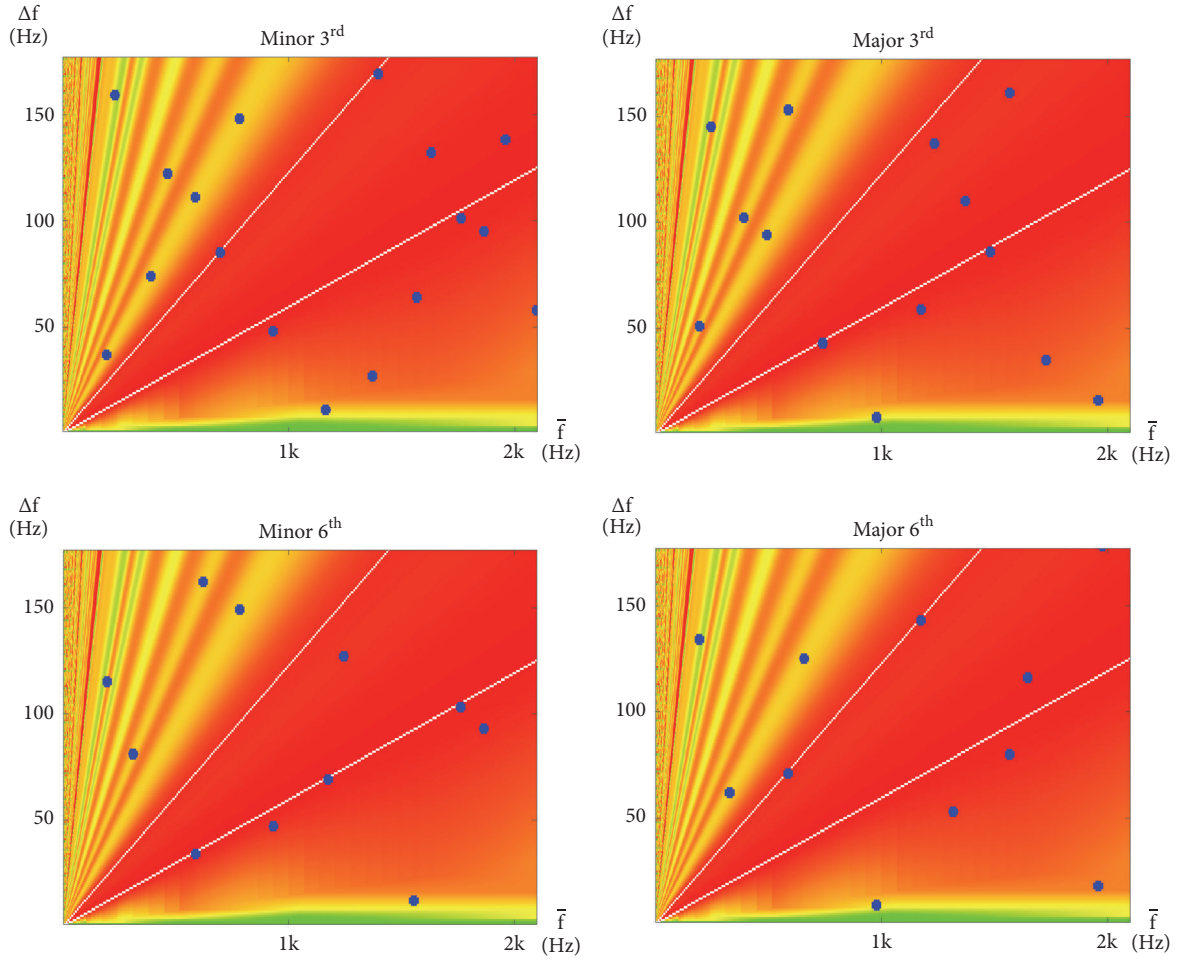
FIGURE 6: Interharmonic plots for all intervals within an octave regarded, in classical music theory, to be imperfectly consonant with a root of g_3. These are, namely, the Minor 3rd (g_3 and b^b_3), Major 3rd (g_3 and b_3), Minor 6th (g_3 and e^b_4), and Major 6th (g_3 and e_4) intervals.

two dissonant lines. Evenly spaced points along a line that passes through the origin also reveal that their Δf share a harmonic relationship. This has a similar (although this is somewhat lesser) redundant effect to that of the synchronous modulation described with perfect consonances.

Consonances that properly contribute to harmony are called imperfect consonances [67] and are presented in Figure 6. As can be seen in the figure, imperfectly consonant intervals have points better distributed. This may be interpreted as erratic modulations that create a continuous stream of unpredictable events to stimulate aural attention, and thus, interest.

A lot of work has already been done on interharmonics since Helmholtz [12, 19–21, 24, 25]. While the main focus of this work is not interharmonics, one purpose of this section is, nevertheless, to provide sufficient background to complete our theory of how the human experience of stationary harmony is based around modulations of both interharmonic and subharmonic nature. From the interharmonic plots in Figures 5–7, a simple predictor of dissonance may be identified to be

$$C\left(\Delta\hat{f}\right) = C\left(\frac{\Delta f}{\overline{f}}\right)$$

$$= \sum_{i=1}^{n}\left[2^{r_{lower}/12} - 1 < \frac{\Delta f_i}{\overline{f}_i} < 2^{r_{upper}/12} - 1\right]$$

(10)

where $\Delta\hat{f}$ will be our shorthand for $\Delta f/\overline{f}$, $C(\Delta\hat{f})$, or $C(\Delta f/\overline{f})$ referring to the number of interharmonic modulations that fall within the central region of dissonance region, i iterates through all interharmonic modulations on the plot, n is the total number of modulations considered, Δf_i and $\overline{f}_i$ refer to the pair of Δf and $\overline{f}$ that describe the ith interharmonic modulation, respectively, and r_{lower} and r_{upper} define the lower and upper boundaries of the region on the interharmonic plot, respectively.

In this section, we have seen how interharmonic modulations are significant to our perception of consonance, dissonance, and emotive response in music. When listening to a duet of instruments with no overtones such as a sinewave theremin or a very pure musical saw, we realize that
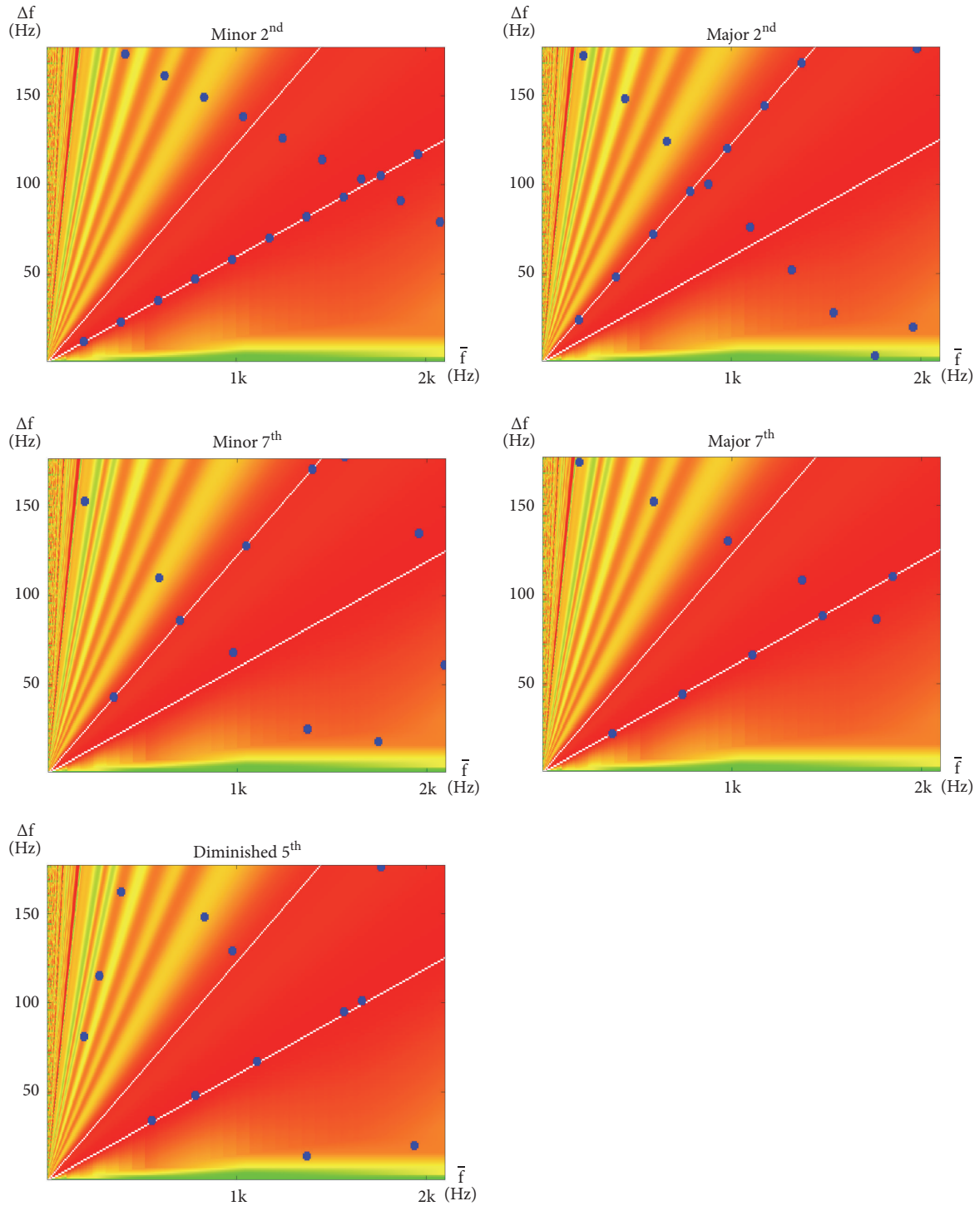
FIGURE 7: Interharmonic plots for all intervals within an octave regarded, in classical music theory, to be dissonant with a root of g_3. These are, namely, the Minor 2nd (g_3 and a^b_3), Major 2nd (g_3 and a_3), Minor 7th (g_3 and f_4), Major 7th (g_3 and $f^\sharp_4$), and Diminished 5th (g_3 and d^b_4) intervals.

consonance, dissonance, and emotion remain present even in harmony without harmonics (i.e., across a well-spaced pair of fundamental frequencies alone). This is just one amongst the several different ways [12, 13, 17, 28, 68, 69] from which

we can deduce that interharmonic modulations cannot be the only determinant of our perception of harmony, which thereby leads to our hypothesis on subharmonic modulations.
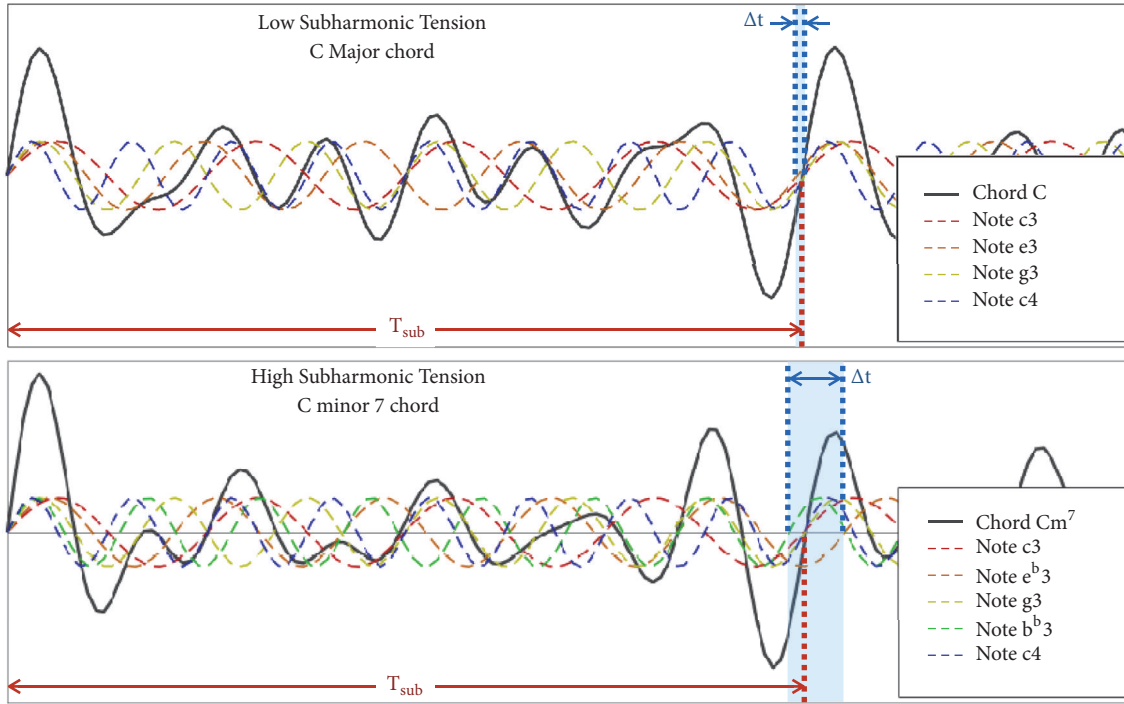
FIGURE 8: Subharmonic wave formation and deformation in the C and Cm^7 chords.

4. Subharmonic Modulations

Apart from the modulations that arise from the summation of adjacent harmonic sinusoids across differing notes, we can (as explained above) deduce that another category of modulations is significant to our perception of harmony. We call these subharmonic modulations. There are two levels of subharmonic modulations, which we dub subharmonic wave formation and subharmonic wave deformation. In this section, we will show how these are significant to our perception of not only stationary harmony, but also transitional harmony.

Figure 8 shows the waveforms of a C Major chord (C) and a C minor 7 chord (Cm^7) composed of the fundamental sinusoids of each composite note. We let each sinusoid start at phase zero since; for purpose of example, we are only interested in wave period. Only the fundament needs to be considered for the same reason. In both cases, the waveform resultant of this summation repeats at a frequency approximately subharmonic to all its composite waveforms. In the figure, its period is marked T_{sub}. We call this subharmonic wave formation and say that T_{sub} is a common subharmonic to all its composite waveforms.

In the case of the C chord, as shown in the figure, each composite sinusoid crosses zero at nearly the same point around $t = T_{sub}$. As marked in the figure, Δt (which is the difference between the first and the last negative-to-positive zero-crossing around the $t = T_{sub}$ region) is small. However, in the case of the Cm^7 chord, Δt is much larger. One can imagine that each successive period of the resultant waveform looks less and less like the first as it gets more and more deformed. This happens slowly for the C chord because of

the small Δt but faster for the Cm^7 because of the large Δt. We call this subharmonic wave deformation. Supplementary Video S1 compares subharmonic wave deformation in a low-tension C chord to that in a high tension Cm7 chord.

Recalling our wave equation from (3), we can rewrite $A\cos\omega_1 t + B\cos\omega_2 t$, or $A\cos 2\pi f_1 t + B\cos 2\pi f_2 t$, as

$$A\cos 2\pi f_1 t + B\cos 2\pi f_2 t$$
$$= A\cos 2\pi \left(k_1 f_{sub} + \Delta f_1\right) t \tag{11}$$
$$+ B\cos 2\pi \left(k_2 f_{sub} + \Delta f_2\right) t$$

where f_{sub} is an approximate common factor of f_1 and f_2, k_1 and k_2 are integer multipliers, and Δf_1 and Δf_2 are small values that balance the equation by making up for the discrepancies that arise with finding a common factor.

In (11), two fundamental frequencies f_1 and f_2 are described as the multiple of a lower subharmonic frequency that is common to them (f_{sub}). We call this their *common subharmonic*.

Since all harmonics are multiples of their fundamental, a subharmonic to any fundamental would inherently be subharmonic to all its harmonics. For this reason, only the fundamental of each note needs to be considered.

Since harmony in music is commonly composed of more than just two notes, we generalize this to describe fundamentals and common subharmonics from any number of notes to get

$$\sum_{i=1}^{N} A_i \cos 2\pi f_i t = A_1 \cos 2\pi \left(k_1 f_{sub} + \Delta f_1\right) t$$
$$+ A_2 \cos 2\pi \left(k_2 f_{sub} + \Delta f_2\right) t + \cdots$$

$$+ A_i \cos 2\pi \left(k_i f_{sub} + \Delta f_i\right) t + \cdots$$
$$+ A_N \cos 2\pi \left(k_N f_{sub} + \Delta f_N\right) t$$
$$(12)$$

where N is the number of notes in the chord, i cycles through each of them, and A_i is the amplitude coefficient of note i.

Beyond this point, it would be easier to visualize subharmonics in the time domain. With the fundamental frequency of note i given by

$$f_i = k_i f_{sub} + \Delta f_i \qquad (13)$$

the fundamental period of each note i is then

$$t_i = \frac{1}{f_i} = \frac{1}{k_i f_{sub} + \Delta f_i} \qquad (14)$$

where t_i is the fundamental period of the note.

Hence, the period of any common subharmonic can be expressed as $k_i t_i$. We can then compensate for nonintegral discrepancies in period rather than in frequency. In doing so, we get

$$T_{sub} = \frac{1}{f_{sub}} = k_i t_i + \Delta t_i \qquad (15)$$

for all i, where T_{sub} is the common subharmonic wave period (we will simply say common subharmonic) of the chord. What carries over as $k_i t_i$ is essentially just the kth subharmonic of note i which lies in the region of T_{sub}. Since this is true for all pairs of k_i and t_i across all values of i when they are each balanced by appropriate t_i, i may be dropped from the left hand side of the equation.

Although the common subharmonic was introduced as the period between primary zero crossings as in Figure 8, we shall, for computational simplicity, redefine it as the mean of $k_i t_i$ across all notes of the chord. Hence,

$$T_{sub} = \overline{k_i t_i} \qquad (16)$$

Figure 9 shows how the period of each subharmonic in the C Major chord from Figure 8 may be plotted. The left column first shows how the period of each subharmonic of c_3 may be plotted in red. The right column then extends this to every remaining note in the chord, with orange, yellow, and blue for the notes e_3, g_3, and c_4, respectively. It may be seen in the right column that a subharmonic period from every note in the chord nearly coincides at around 30 ms. Hence, we say that this is its common subharmonic, T_{sub}, as defined in (16).

Having reduced the waveform plot to subharmonic periods in the vertical axis, we can represent time spanned by each subharmonic in the horizontal axis. We will do this for a song stanza in the next section, in a subharmonic plot.

4.1. Subharmonic Modulations in Stationary Harmony. Figure 10 shows an example of a subharmonic plot. In the horizontal axis there is time in bars and in the vertical axis there is the subharmonic wave period in milliseconds. Note that the subharmonic axis runs top down to put shorter wave periods at the top because they correspond to higher

frequencies. Larger wave periods, which correspond with lower frequencies sit conversely at the bottom. The tails that run horizontally represent the span of time covered by each note. Subharmonics are colored to match their corresponding notes on the music score. For example, in the first bar, all subharmonics of $f^\#_5$ are marked out in red, followed by d_5 in orange, a_4 in yellow, d_4 in green, a_3 in blue, and d_3 in purple. The musical score runs in parallel at the bottom of the plot as reference. Once again, all plots and computations in our examples assume equal temperament unless stated otherwise. This example shows the opening stanza of Pachelbel's Cannon in D [70] and focuses on stationary harmony, leaving transitional harmony to a later example.

Subharmonics. For every bar, the dashes that flush with the reference point at 0 ms mark $0 \times t_0$. Carrying on top down with each bar in accordance to color, we get subharmonics at $1 \times t_0$, $2 \times t_0$, $3 \times t_0$, $4 \times t_0$, etc.

Notes and Melody Line. Since the topmost dash of each color for every bar below the 0 ms reference represents $1 \times t_0$, they relate to the fundamental period of each note; of these, the topmost ones of every bar across all colors mark the melody line, $f^\#_5$-e_5-d_5-$c^\#_5$-b_4-a_4-b_4-$c^\#_5$. (They are red in this particular example.) Hence, it is easy to interpret the melody line in a subharmonic plot. The periods, t_i, of each note of the melody are marked against the vertical axis in milliseconds as well as their common note names.

Chords and Coincidence. Common subharmonics may be visualized in regions with the (approximate) coincidence of dashes of every color. Again, the common subharmonics (T_{sub}) of each chord in the stanza are marked out against the vertical axis in both milliseconds and their respective chord names.

Key. Every note of the diatonic shares a common subharmonic. Hence, it is possible to identify the key of a song by its common subharmonic, assuming minimal deviations from its key. The common subharmonic associated with the key of this song is marked out much further down the plot. Dotted lines indicate discontinuity. (This part of the figure is plotted in just intonation to avoid the snowballing of Δt_i to better illustrate this.)

Stationary Tension. Most of the time, contributing subharmonics from different notes are not precisely coincident. Major chords have better coincidence than minor chords, and triads coincide better than sevenths and extended chords. With subharmonic modulations, perceptual tension arises with the noncoincidence of common subharmonics. Noncoincidence is measured by an overall Δt as reflected in Figures 8 and 10. We call this its (stationary) subharmonic tension.

This Δt is given by the difference between the largest and smallest subharmonics in the chord that coincides around T_{sub}.

$$\Delta t = [k_i t_i]_{max} - [k_i t_i]_{min} \qquad (17)$$

Period of
Fundament
of c_3
7.7ms

Period of
2nd Subharmonic
of c_3
15.3ms

Period of
3rd Subharmonic
of c_3
22.9ms

Period of
4th Subharmonic
of c_3
30.6ms

Period of
5th Subharmonic
of c_3
38.2ms

- 0ms -

- 10ms -

- 20ms -

- 30ms -

- 40ms -

T_{sub}

Period of
Common Subharmonic
across c_3, e_3, g_3, and c_4

C
c_4
e_4
g_4
c_5

harmonics of c_3
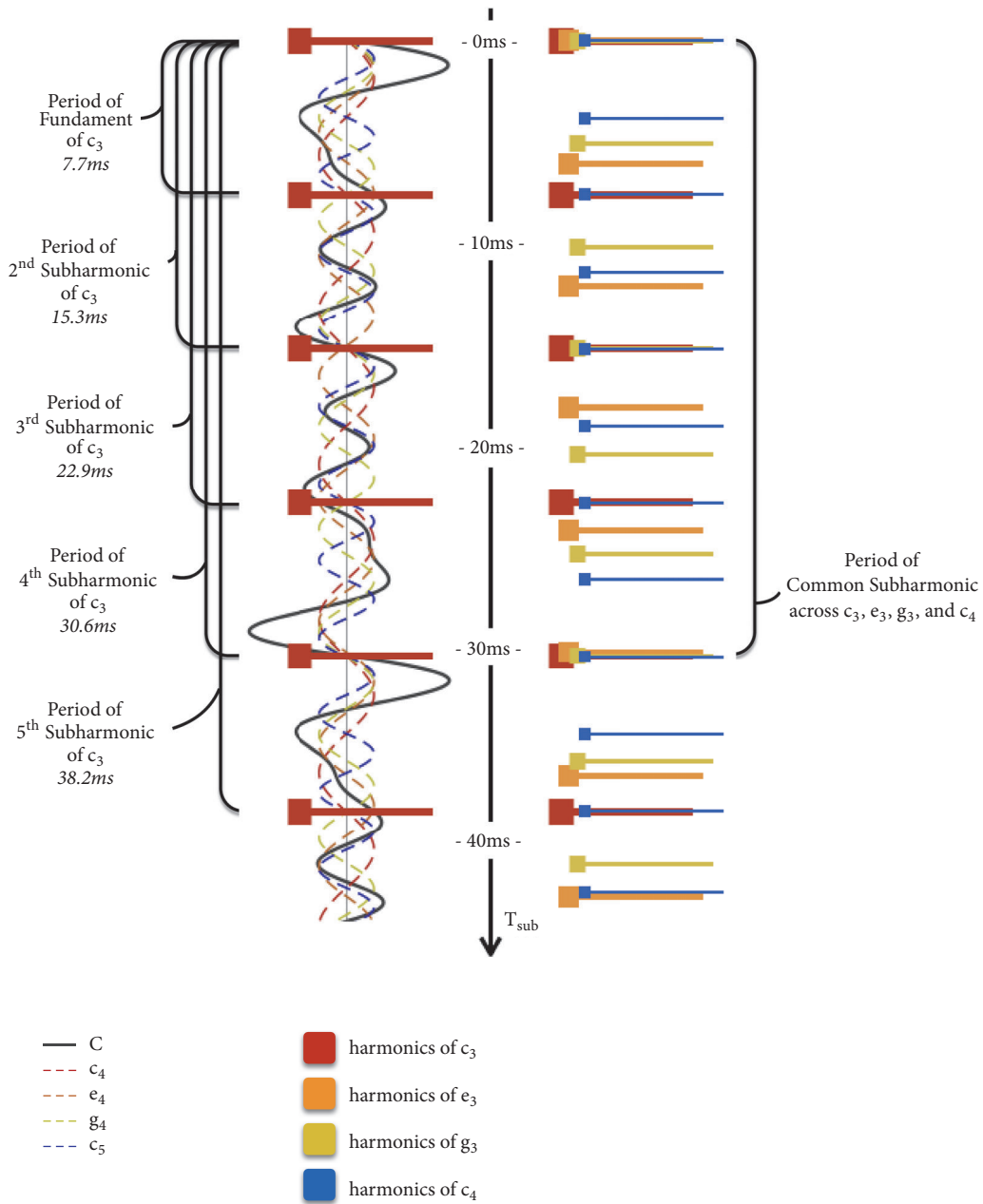
harmonics of e_3

harmonics of g_3

harmonics of c_4

FIGURE 9: Plotting subharmonic wave periods of the C Major chord.

where $[k_i t_i]_{max}$ and $[k_i t_i]_{min}$ denote the largest and the smallest subharmonics in the chord that (nearly) coincides around T_{sub} (mathematically, they are the maximum and minimum values of $k_i t_i$, resp.).

Δt and T_{sub} are the primary features of stationary tension. Δt may be normalized by expressing it like a duty cycle by taking

$$\Delta \hat{t} = \frac{\Delta t}{T_{sub}} \tag{18}$$

From Figure 3 in the section on interharmonic modulation, recall that dissonances increased and decreased with interharmonic modulation frequency while consonances behaved inversely. This happens only within a certain range. When interharmonic modulation frequency shrinks to the brink of zero, it falls below musical significance.

Subharmonic tension behaves similarly. Figure 11 describes different types of harmony on the subharmonic tension scale. As can be seen in the figure, our response to subharmonic tension is likewise. Perceived dissonances increase and decrease with subharmonic tension while
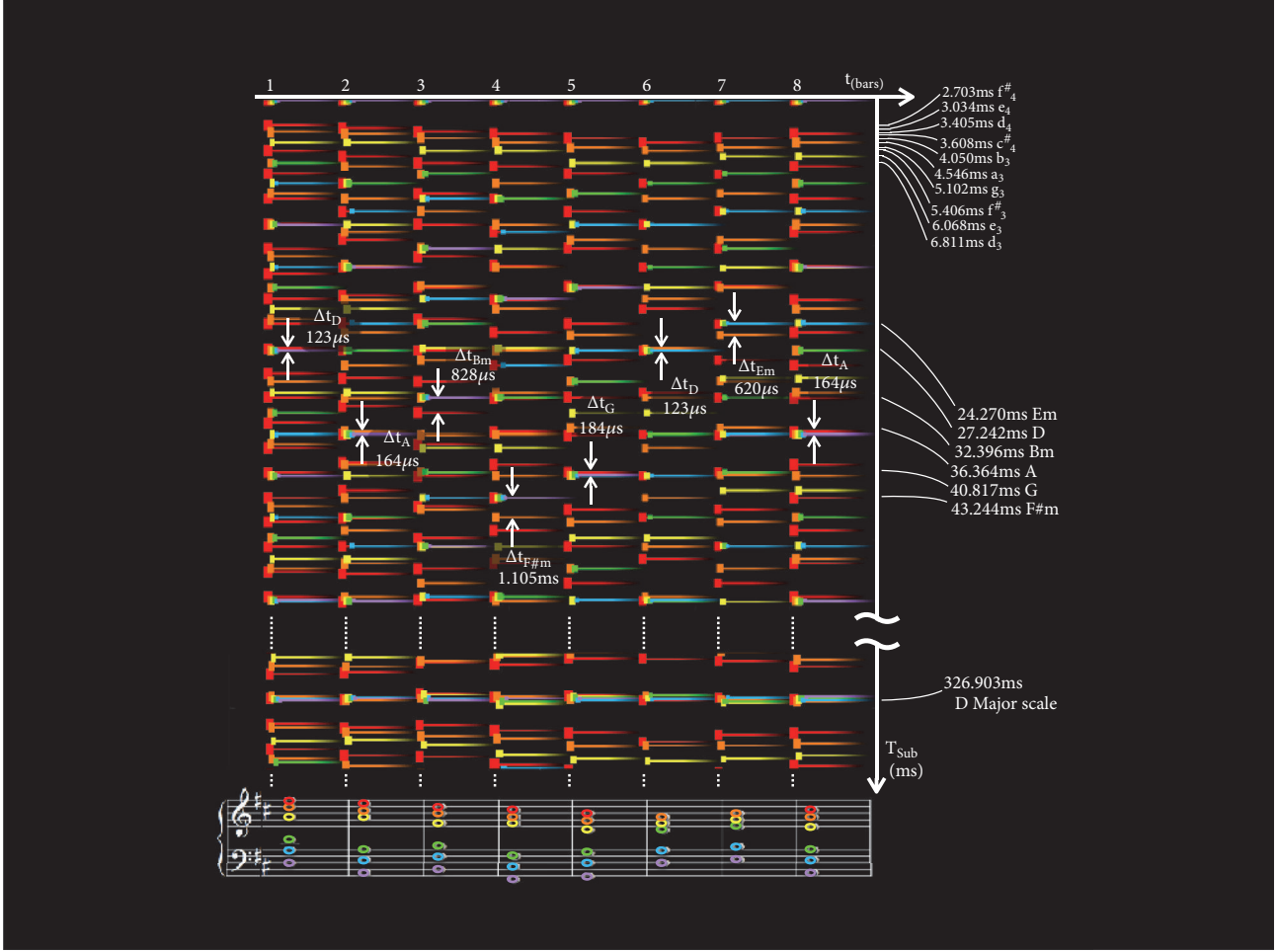
FIGURE 10: Subharmonic plot of the opening stanza of Pachelbel's Cannon in D with period in milliseconds on the vertical axis and time in bars on the horizontal axis. Subharmonics are colored to match the color of their corresponding notes on the music score below. The subharmonic tensions of each chord, Δt, are marked out on the plot with white arrows. Significant wave periods, along with common subharmonic periods, T_{sub}, are marked against the vertical axis on the right. In the interest of visiting all common chords of the key, Em is used in the 7th bar instead of G, which already occurs in the 5th bar. Considering the fact that this example is not used for transitional harmony, all chords are presented in its root position at the expense of introducing parallel 5^{th}s in the interest of normalization for fairer comparison.

perceived consonances behave inversely within common range. Mathematically,

$$\varepsilon\{X\} \propto \frac{1}{\widehat{\Delta t}_X} \tag{19}$$

where $\varepsilon\{X\}$ is the *harmonious effect of chord X* and $\widehat{\Delta t}_X$ is its stationary subharmonic tension (its $\widehat{\Delta t}$).

However, as described in the figure, modulations from subharmonic tension fall below musical significance; the effect of harmony drops to zero as modulations from subharmonic tension fall below musical significance. Hence, where $\widehat{\Delta t}_{threshold}$ is the said threshold of musical significance, as $\widehat{\Delta t} < \widehat{\Delta t}_{threshold}$,

$$\lim_{\widehat{\Delta t} \to 0} \varepsilon\{X\} = 0 \tag{20}$$

Thus, perceptual tensions and consonances are experienced in slew-like modulations of the waveform at common subharmonic locations. (This is the effect of periodically changing phase relationships amongst the contributing waveforms, for which Δt is a measure.) While there may be several common subharmonics for every chord within reasonable range, we theorize that our ears identify most with the shortest few. Subharmonic consonances are described by gentler modulations (small Δt) at the shortest common subharmonic locations (short T_{sub}), while subharmonic dissonances are described by more turbulent ones (associated with absence of small Δt at short T_{sub}).

The sensation of a chord can be highly complex, with different tensions and consonances perceived simultaneously, an experience inadequately represented by a single term for dissonance. Attempting to rate every chord by its dissonance level alone can be compared to rating every
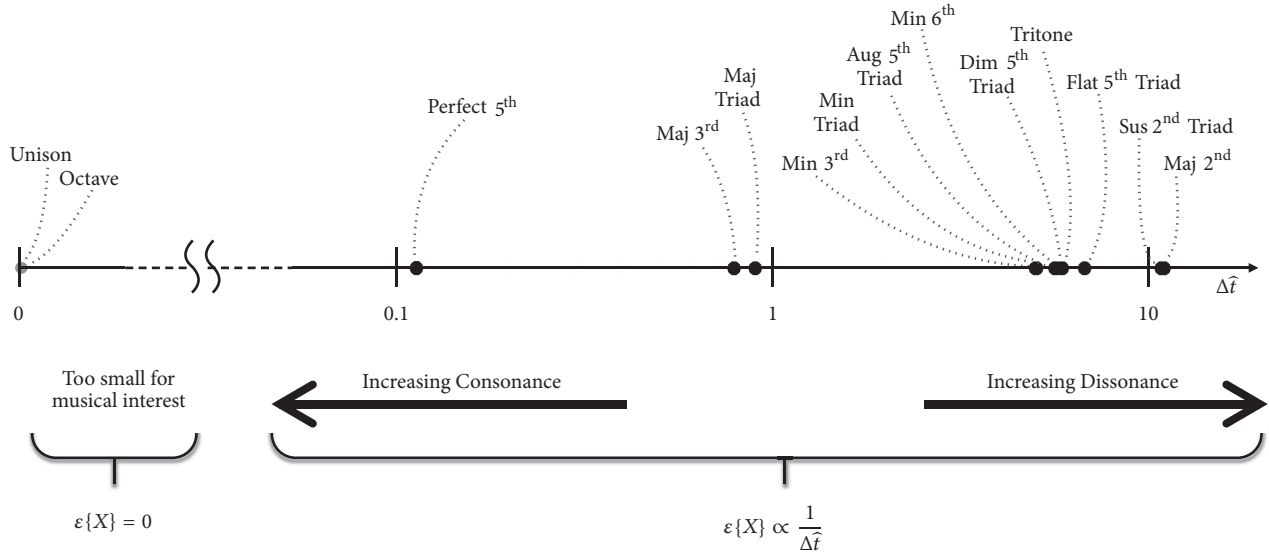
FIGURE 11: The effect of harmony, $\varepsilon\{X\}$, on the scale of subharmonic tension, $\Delta\hat{t}$.

variety of chocolate in a candy store by only how sweet or bitter it is. The advantage of Δt, as opposed to existing correlates of harmony [3, 13, 43, 54], is the way it explains abstract notions of perceptual tensions and consonances by ascribing them to regions across the subharmonic spectrum with a strong sense of attribution or identification. While, for purpose of illustration, Figures 9 and 10 have shown examples where a modal T_{sub} (shortest T_{sub} with smallest Δt) is easiest to identify, we theorize complex chords with ambiguous T_{sub} (where it is difficult to attribute the collection of modulations experienced to a single modal); our ears often identify with several common subharmonics simultaneously. In other words indeterminate cases could possibly arise with particularly discordant harmonies without small Δt at short T_{sub}. Thus, for programmatic analysis of a large number of chords, it is, nevertheless, useful to have a single term to represent the overall dissonance of each chord. For this, we use

$$\widetilde{\Delta t} = n \left(\frac{1}{\sum_{n:m} \left(1/T_{sub,j} \left(\Delta t_j \right)^c \right)} \right)^{1/c} \quad (21)$$

where a single term, $\widetilde{\Delta t}$, represents the overall subharmonic tension, $T_{sub,j}$ and Δt_j refer to individual candidates of T_{sub} and Δt with j iterating through each candidate pair, c is the preemphasis (while $1/c$ serves as "post de-emphasis"), and $\Sigma_{n:m}$ denotes summing over the n smallest values out of a range of m values considered. In our work, n is always chosen to be half of m unless stated otherwise. Note that $T_{sub,j}$ here serves as a weighting factor to weight down higher subharmonics, which, as aforementioned, are less significant. Inverting before (and rectifying after) summation mimics our hearing by allowing smaller values of Δt_j to contribute better towards a smaller $\widetilde{\Delta t}$.

We will see how representative $\widetilde{\Delta t}$ is of stationary harmony in the next section. But before that, we will first explain subharmonic modulations in transitional harmony.

4.2. Subharmonic Modulations in Transitional Harmony. While stationary harmony studies chord sonorities (how a chord sounds on its own), transitional harmony deals with chord progressions and resolutions (how chords transit from one to another). It is remarkable how a low tension (consonant) chord can transit to a high tension (dissonant) one yet still bring about the perceptual effect of tension release (resolution) [18]. From this it may be deduced that transitional harmony stands largely independent of stationary harmony, even though both are considered when assigning harmony in composition. Even though numerous studies have been conducted on stationary harmony from the psychoacoustic approach, work on transitional harmony remains primarily nonpsychophysical.

Traditional classical music theory uses the term resolution to describe the perception of tension released when a chord is suitably followed by another chord [18]. With subharmonic modulation, we theorize that these abstract perceptions of tensions released may be identified and quantified in the perceived trajectories of subharmonics as one chord progresses to the next. Figure 12 illustrates this.

Figure 12 shows the opening line of Beethoven's Moonlight Sonata [71]. Before we begin our analysis, one should note that unlike Pachelbel's Cannon the use of *arpeggios* (broken chords) means that notes contributing to the harmony may not necessarily start at the same time, but, when the sustain pedal on the piano is applied, they sustain and overlap until the end of each bar. The names of the chords formed by the notes are labelled along the top of the score to aid the reader in this analysis. Another thing to note would be the fact that this piece maintains a strong sense of *voice leading*
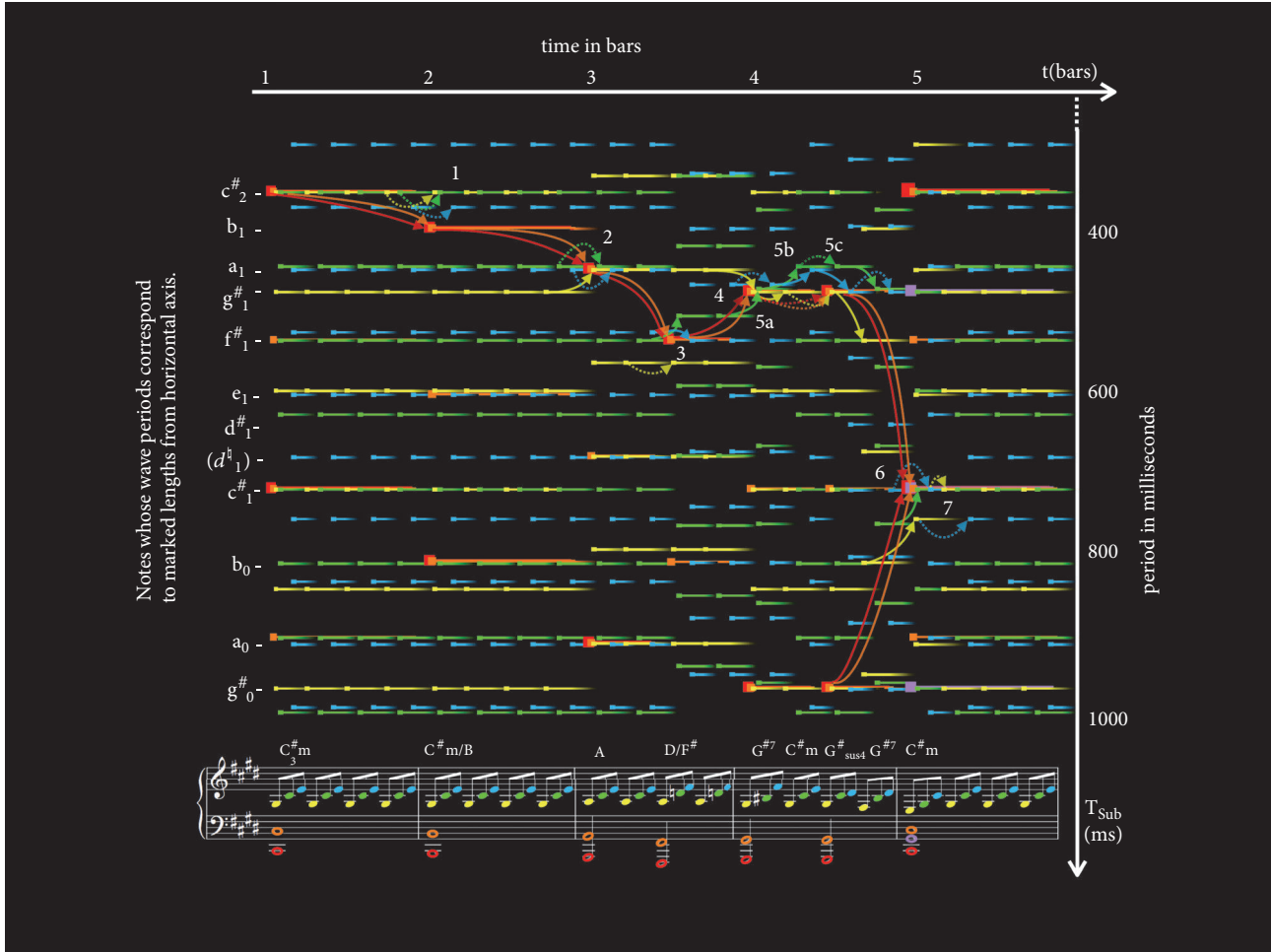
FIGURE 12: Subharmonic plot of the opening line of Beethoven's Moonlight Sonata with period in milliseconds on the vertical axis and time in bars in the horizontal axis. Subharmonics are colored to match their corresponding notes on the music score. Names of relevant notes are marked out on the left, at T_{sub} values corresponding to their wave period. The region of each transition is numbered in white. Colored arrows follow voice leading along the notes across chord changes.

[72], which means that each note from a chord has strong progressive associations with a note from the previous and another from the succeeding chord. The subharmonics of all notes that are associated in this way (i.e., of the same voicing) across the song are coded with the same color to aid the reader in this analysis. For example, all notes in red on the music score represent the bass (lowest) notes throughout the song, and every subharmonic of these notes is portrayed in red.

We theorize that in chord transitions every subharmonic $(k_i t_i)$ that (nearly) coincides around the common subharmonic (T_{sub}) of a succeeding chord is perceived to transit from the nearest corresponding (i.e., of the same voicing) subharmonics in the preceding chord. These transitions are marked out by the arrows in Figure 12, which are colored according to the notes they are associated with. Arrows are usually convergent (with the exception of, for example, a basic triad progressing onto an extended chord of the same root) because the subharmonics of the succeeding chord always identify with a common subharmonic whereas those of the preceding chord usually do not.

The central hypothesis of transitional subharmonic theory is that perceptual tension resolution, which is so often described in traditional music theory but never physically identified in acoustics, lies in the degree of convergence seen here.

Assuming transition to be abrupt (since notes do not commonly glide from one pitch to another in music) we compute a Δt for the succeeding common subharmonic and a Δt for its preceding corresponding subharmonics and simply measure this degree of convergence as the difference between the two. As such,

$$\Delta\Delta t = \Delta t_p - \Delta t_s \tag{22}$$

where Δt_s refers to the Δt of the succeeding chord and Δt_p refers to the Δt defined by its nearest preceding subharmonics.

This can be normalized by dividing by T_{sub} such that

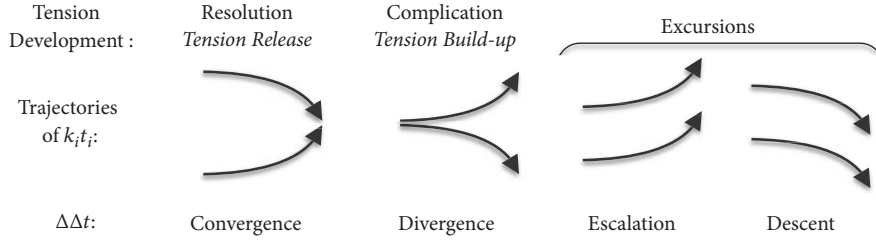$$\Delta\Delta\hat{t} = \frac{\Delta t_p - \Delta t_s}{T_{sub}} \tag{23}$$

FIGURE 13: Trajectories of $k_i t_i$ in different states of tension development (states of convergence).

where $\Delta\Delta\hat{t}$ denotes normalized $\Delta\Delta t$ and T_{sub} refers to that of its succeeding chord.

$\Delta\Delta t$ is, thus, a quantification of the tension; Δt is released over the transition at the wave period of the succeeding common subharmonic.

According to our theory, tension resolution is perceived in the release of this tension across each transition. Thus, mathematically,

$$\varepsilon\left\{X_1 \longrightarrow X_2\right\} \propto \Delta\Delta\hat{t}_{X_1 \rightarrow X_2} \qquad (24)$$

where ε denotes the perceptual resolving effect of tension release and $\Delta\Delta\hat{t}_{X_1 \rightarrow X_2}$ denotes the $\Delta\Delta\hat{t}$ across the transition of chord X_1 to chord X_2.

Since resolution (tension release) [18, 42] in harmony progression is perceived in the convergence of $\Delta\hat{t}$, what we will refer to as *complication* (build-up of tension or negative resolution) is seen in its divergence, where $\Delta\Delta\hat{t} < 0$ and $\varepsilon\{X_1 \longrightarrow X_2\}$ is negative.

Three possibilities arise when looking at T_{sub} and Δt from this perspective, by which we can divide transitional harmony into three classes. As illustrated in Figure 13, these are as follows.

(1) Resolution, also called tension release: this is the most common occurrence and occurs with the *convergence* of Δt (i.e., $\Delta t_p > \Delta t_s$) and a positive $\Delta\Delta t$. The larger the $\Delta\Delta t$, the larger the perceptual tension release.

(2) Complication, also called tension buildup: this is the least common occurrence and occurs with the *divergence* of Δt (i.e., $\Delta t_p < \Delta t_s$) and a negative $\Delta\Delta t$. *Just as negative aesthetics may be used expressively in a painting, it may similarly be used in music* [73]. The larger the magnitude of $\Delta\Delta t$, the larger the perceptual tension buildup. Complications usually only occur when the preceding T_{sub} is equal or nearly equal to the succeeding T_{sub}. Musically speaking, it usually occurs when a simpler chord is followed by a more complex chord of the same root.

(3) Excursion: Because of the circular nature of the musical chroma, the preceding T_{sub} and the succeeding T_{sub} may be computed to differ by up to 6 semitones in either direction. When the difference is 1 or 2 semitones, this corresponds to a neighboring note, and the collective (uplifting or detrimental) effect of melodic movement (i.e., melody) across each note of the chord can overpower the effect of harmony. In

such cases, our ears are persuaded to identify Δt_p with $[k_i t_i]_{max} - [k_i t_i]_{min}$ of the nearest preceding T_{sub}. When this happens, $[k_i t_i]_{max}$ and $[k_i t_i]_{min}$ move in the same direction; hence, *neither convergence nor divergence is perceived*. There are 2 such cases as follows.

(a) Escalation: this occurs when each $[k_i t_i]$ shortens simultaneously, T_{sub} shortens by a factor equivalent to 1 or 2 semitones ($2^{1/12}$ to $2^{2/12}$ times), and f_{sub} rises, producing the uplifting effect of melodies rising by 1 or 2 semitones.

(b) Descent: this occurs when each $[k_i t_i]$ lengthens simultaneously, T_{sub} lengthens by a factor equivalent to 1 or 2 semitones ($2^{1/12}$ to $2^{2/12}$ times), and f_{sub} falls, producing the detrimental effect of melodies falling by 1 or 2 semitones.

It is fascinating to note how the perceptual development (build-up and resolution) of tension that is so often described in music [18, 42] but never identifiable with an acoustic attribute may here be visualized in the convergence and divergence of common subharmonics. Figure 13 further illustrates how $k_i t_i$ trajectories reflect the development of tension build-up and release. Additionally, trajectories for excursions are illustrated in the same figure.

Returning to Figure 12, the transitions between each chord are labeled 1 to 7 in the figure and correspond to 1 to 7 as follows.

(1) The song starts off with a $C^{\#}m$ chord. Hence, the common subharmonic is observed around a wave period of $c^{\#}$. Our ears adhere especially to the shortest one, which is at $c^{\#}{}_2$. Large Δt is attributed to the complex tensions within a minor chord. At the region marked 1, this transits to a $C^{\#}m/B$ chord. The tension built up with the divergence of Δt may be visualized in the divergence of the arrows in the figure (of which the dotted ones across the plot are used to indicate the continuation of subharmonics, i.e., $k_i t_i$ that do not change). Both perceptually in music and acoustically, as defined above, this translates to a further complication to the existing minor tension.

(2) At region 2, there is a convergence to a momentary (half-bar) low-tension A chord. The uplifting effect of a large tension release, $\Delta\Delta\hat{t} \gg 0$, is counterbalanced by the detrimental effect of a falling melodic sequence

(lengthening T_{sub}), adding to the complexity of the song.

(3) At region 3, A transits to a D/F$^{#}$, which is a Neapolitan chord. The low f$^{#}$ bass extends over 2 octaves below the treble notes, putting a strong T_{sub} at a nonroot period of f$^{#}_1$ and creating an amount of stationary tension that is unusual for a major chord. (In such cases, there is usually another common subharmonic with lower Δt but at a wave period corresponding to a root at a much larger T_{sub}.)

(4) At region 4, the Neapolitan chord resolves to the Dominant 7th, marked G$^{#7}$ in the figure, with a large perceptual resolution that is signature to ^{b}II^{6}-V^{7} transitions in music [42]. This large tension release is visualized as a large convergence in the subharmonic plot as indicated by the arrows.

(5) Musically, the Dominant 7th typically plays the role of building an anticipation for the upcoming return to the Tonic [42]. Beethoven enhanced this function particularly well with a double suspension with staggered resolutions in regions 5a through 5c. The subharmonic plot gives tangibility to the perceptual details with suspension-resolution long theorized about in music that can now be affirmed with visualization.

   (a) At region 5a, the transition from the G$^{#7}$ progresses to what is labeled C$^{#}$m. However, this C$^{#}$m is functionally still a G$^{#}$ with a double suspension of the 3rd (b$^{#}$) to a 4th (c$^{#}$) and the 5th (d$^{#}$) to a 6th (e), respectively. The perceptual complication that arises with this transition can be visualized in the subharmonic plot as indicated by the divergence of the green and cyan arrows, respectively. The deviation of the suspended notes from the primary triad is visualized as a deviation of their $k_i t_i$ from T_{sub}.

   (b) At region 5b, the tension resolution with the 6th being resolved back down to the 5th can be visualized in the subharmonic plot by its $k_i t_i$ resolving back to T_{sub} as indicated by the convergent cyan arrow. The continuation of the suspended 4th is visualized in the dotted green arrow.

   (c) At region 5c, the tension resolution with the 4th being resolved back down to the 3rd can be visualized in the subharmonic plot by its $k_i t_i$ resolving back to T_{sub} as indicated by the solid green arrow. In preparation for a major resolution back to the upcoming tonic, Beethoven's touch of genius combines this resolution with a simultaneous complication in the introduction of the 7th at this point. This is visualized in the deviation of its $k_i t_i$ away from T_{sub} as indicated by the divergent solid yellow arrow.

(6) At region 6, the Dominant 7th is resolved back to the Tonic with a tension release unique to V$_7$-tonic cadences that is so immense that it is has been long established as the de facto cadence for the end of musical passages [42]. This immense perceptual release of tension, too, is identifiable in the subharmonic plot. From the figure, it may be seen that the common subharmonic, T_{sub}, of C$^{#}$m (located at the period of c$^{#}_1$ this time, because of the g$^{#}_2$ in purple) lies right in the middle of two common subharmonics of G$^{#7}$ (located at the periods g$^{#}_1$ and g$^{#}_0$). This unique subharmonic behavior allows our ears to quite possibly identify with both $k_i t_i$ for the preceding Δt making $\Delta \hat{t}_p$ significantly larger than its $\Delta \hat{t}_s$. Its staggering convergence produces an immense sense of tension resolution with this transition.

(7) A final landmark that is interesting to note is at region 7, where the triad in the treble flips from the 1st inversion to the 2nd inversion while the chord remains unchanged. Notice that this brings about no change to both T_{sub} and $\Delta \hat{t}$ while $\Delta \Delta \hat{t} = 0$. This, again, shows how subharmonic analysis agrees with music theory where, despite the change of notes, harmony remains the same at this point.

In this section, we have seen how, even in the context of transitional harmony, perceptual tensions and resolutions in a song may be visualized in its subharmonic modulation. We will move on to see how well numerical values computed with such modulations verify against listening tests and chord use statistics.

5. Experiment and Results

For both stationary and transitional harmony, tensions computed from our models show strong correlations with consonance rankings and historical chord use statistics. Table 1 tabulates a summary of the results of our experiment.

We will explain each of these results in detail in the following subsections.

5.1. Stationary Harmony. For stationary harmony, we take the overall tension of a chord to be a simple weighted sum of $T_{\Delta f}$ and $T_{\Delta t}$

$$T_{\Delta f | \Delta t} = w_i T_{\Delta f} + w_s T_{\Delta t} \tag{25}$$

where $T_{\Delta f | \Delta t}$ is overall tension, $T_{\Delta f}$ and $T_{\Delta t}$ are taken to represent the tensions contributed by interharmonic and subharmonic modulations, respectively (normalized by linearly scaling to fit between 0 and 1), and w_i and w_s are their weights, or summing coefficients respectively, where $w_i + w_s = 1$ and 0.61 and 0.39 are found to provide a good distribution.

We use a simple estimate of $T_{\Delta f}$, taking

$$T_{\Delta f} = C_1 \left(\Delta \hat{f} \right) + C_2 \left(\Delta \hat{f} \right) \tag{26}$$

where $C_1(\Delta \hat{f})$ and $C_2(\Delta \hat{f})$ are a tally of interharmonic modulations (given by (10)). By visual inspection of the

TABLE 1: Summary of correlations with consonance rankings and historical chord use.

Stationary harmony		Transitional harmony		
Dyads/intervals (2 notes)	Triads (3 notes)		Triads & tetrads (3 or 4 notes)	
		All transitions	All transitions *Excl. comp.*	Resolutions
r= 0.922	r=0.907	r= 0.903	r= 0.970	r= 0.996
p=0.0001	p=0.0000	p=0.0000	p=0.0000	p=0.0000

interharmonic plot, regions of dissonance are defined by r_{lower} = 0.95 and r_{upper} = 1.1 for $C_1(\Delta \hat{f})$ and r_{lower} = 1.5 and r_{upper} = 2.8 for $C_1(\Delta \hat{f})$.

For $T_{\Delta t}$, we use $(\widetilde{\Delta t})^2$, where $\widetilde{\Delta t}$ is given by (21) preemphasized with c = 2.1 across a range of m = 5. (A preemphasis of just over 2 provided the sufficient discrimination without driving data into saturation. A broad range of m-values are suitable but we settled on a smaller value of 5 for computational simplicity.)

Numerous previous authors have performed notable work for stationary harmony both within and outside the psychophysical context [8, 12, 13, 18, 21–25, 43, 53, 62–64]. For dyads (intervals, or two-note chords) and triads (three-note chords), we the use precollated information in Tables 2–5 from Stolzenburg [43] for comparison. Dyads (intervals) are compared against the results of an average across 7 notable studies collated by Schwartz et al. [54] on a ranking of 12 chords. Stolzenburg adds the unison to Schwartz's list, which he reasonably assumes to be the most consonant, hence, we have appropriately included it as well. Triads are compared to results from an experiment by Johnson-Laird, Kang, and Leong [13] as cited in Stolzenburg [43]. For consistency with Stolzenburg's statistics in the comparison, these were first converted to ordinal rankings before computing the correlation as practised by Stolzenburg [43]. Table 2 lists our correlations for dyads and triads in stationary harmony against known relevant work as taken from Stolzenburg's [43]. A detailed tabulation of all available values for each chord is provided in the appendix.

5.2. Transitional Harmony. For transitional harmony, $\Delta\Delta t$ from (22) is suitable for hand-computation of transitional harmony across individual locations of succeeding common subharmonics, Δt_s, across the soundscape. While this is advantageous for visualizing individual complications and resolutions at multiple locations across the tensional soundscape, it requires manual identification of a modal Δt_s for every transition which can be ambiguous for particularly discordant harmonies. For a consistent programmatic approach with larger datasets, we take the measure of *overall $\Delta\Delta t$* of a transition defined by

$$\widetilde{\Delta\Delta t} = \left(\frac{1}{n} \sum_{j=1, \, \forall \Delta t_{s,j} < (1/2)\Delta T_{sub}}^{N} \left(\frac{\Delta\Delta \hat{t}_j}{\Delta t_{s,j} T_{sub,j}} \right)^c \right)^{1/c} \quad (27)$$

where $\widetilde{\Delta\Delta t}$ is representative of overall tension resolved, $\Delta\Delta\hat{t}_j$, $\Delta t_{s,j}$, and $T_{sub,j}$ refer to individual candidates of $\Delta\Delta\hat{t}$, Δt_s, and T_{sub}, respectively, N is the range of nodes considered, j iterates through all relevant common subharmonics of the succeeding chord, ΔT_{sub} denotes the distance between two adjacent $T_{sub,j}$, $\Sigma_{j=1, \, \forall \Delta t_{s,j} < (1/2)\Delta T_{sub}}^{N}$ denotes summing across all values of $1 < j < N$ wherever $\Delta t_{s,j}$ is less than half the distance between the adjacent $T_{sub,j}$ on either side, n is the number of nodes summed, and c is the preemphasis as explained with (21).

This effectively computes the preemphasized, weighted, and compensated mean $\Delta\Delta t$ across all eligible common subharmonics within a range of N for a given transition. T_{sub} weights down larger subharmonics which are less significant according to the theory. (It is a reciprocal as opposed to (21) because greater pleasure is associated with larger tension released.) $\Delta t_{s,j}$ compensates for the fact that, apart from tension resolution alone, stationary consonance also affects one's preference for the succeeding chord. $\Delta t_{s,j} < (1/2)\Delta T_{sub,j}$ effectively sets the criterion for a node to be considered a common subharmonic. In our experiments, we set N = 9. (A broad range of N will work, but we choose a smaller value for computational simplicity. Larger values may be required with larger range or dataset size.) In consideration of divergent transitions in the dataset, we set c = 1 (no preemphasis) because divergent transitions have negative $\Delta\Delta t$ which can be distorted by preemphasis.

With transitional harmony, conducting an accurate listening test is less straightforward. Rather than attempting to acquire a small number of fresh unproven opinions, it is reasonable to use statistics from a large number of well-esteemed premade decisions. A simple way to measure how well numerical values of subharmonic transition agree with the music theorists' school is to compare them with statistics of an expert music theorist's chord use. Capturing chord-use statistics from music score is again, however, a labor-intensive process requiring domain expertise [46, 47, 74]. Details such as melody-harmony discrimination, transition onset, and root ambiguity (e.g., Dm^7/F versus F^6) are often not precisely defined in a song. We find the largest relevant data readily available that also meets chord-spelling precision requirements in Tymoczko's *Study on the Origins of Harmonic Tonality* [45]. In this study, Tymoczko interpreted and recorded the statistics of 11,000 chord transitions from Palestrina's [75] corpus. Palestrina was highly regarded for his style of harmony by Helmholtz himself [76]. He is widely

TABLE 2: Proposed and existing correlates of stationary harmony.

Method	Dyads r (p)	Triads r (p)
$T_{\Delta f \mid \Delta t}$ *(Proposed) Equal Temperament*	0.922 (0.0000)	0.907 (0.0000)
Log Periodicity *Just* [43]	0.982 (0.0000)	0.831 (0.0002)
Rel. Periodicity *Just* [43]	0.982 (0.0000)	0.846 (0.0001)
Log Periodicity *Rational* [43]	0.936 (0.0000)	0.813 (0.0004)
Rel. Periodicity *Rational* [43]	0.936 (0.0000)	0.808 (0.0004)
Rel. Periodicity *Pythagorean* [43]	0.817 (0.0003)	-
Rel. Periodicity *Kirnberger III* [43]	0.796 (0.0006)	-
Ω measure [62]	0.886 (0.0000)	-
Consonance Raw Value / Degree[*][†]	0.978 (0.0000)	0.826 (0.0016)
Dual Process [13][‡]	-	0.791 (0.0006)
Percentage Similarity [53][‡]	0.977 (0.0000)	0.802 (0.0005)
Instability [18][‡]	-	0.698 (0.0040)
Tension [18][‡]	-	0.599 (0.0153)
Sonance Factor[$]	0.982 (0.0000)	0.434 (0.0692)
Generalized Coincidence [63][‡]	0.841 (0.0002)	-
Consonance Value[‖]	0.940 (0.0000)	0.755 (0.0014)
Dissonance Curve [21][‡]	0.905 (0.0000)	0.723 (0.0026)
Pure Tonality [22][‡]	0.938 (0.0000)	0.675 (0.0162)
Complex Tonality [22][‡]	0.738 (0.0020)	-
Roughness [23][‡]	0.967 (0.0000)	0.352 (0.1193)
Sensory Dissonance [24, 25][‡]	-	0.607 (0.0139)
Critical Bandwidth [12][‡]	-	0.570 (0.0210)
Temporal Dissonance [8][‡]	-	0.503 (0.0399)
Gradus Suavitatis[#]	0.941 (0.0000)	0.690 (0.0045)

[*] Raw Value was used for Dyads and Degree was used for Triads.

[†] Fotlyn, 2012, as cited in [43]

[‡] as cited in [43]

[$] Dyads from [64] and Triads from Hofmann-Engl, 2004, both as cited in [43]

[‖] Brefeld, 2005, as cited in [43]

[¶] Dyads from [23] and Triads from Hutchinson & Knopoff, 1979, both as cited in [43]

[#] Euler, 1739, as cited in [43]

considered amongst music theorists to be the pinnacle of contrapuntal harmony [77].

Table 3 lists $\widetilde{\Delta\Delta t}$ against frequencies of occurrence for each of the 17 most frequently used chords that follow V as read-off Tymoczko [45]'s chord tendency histogram. C, D, X↑, and X↓ indicate the convergence type of the progression. Just intonation was used as being opposed to equal temperament in this case to be consistent with Palestrina.

Their correlations are listed in Table 4. $\widetilde{\Delta\Delta t}$ shows a significantly strong positive correlation of 0.903 with Palestrina's chord tendencies in general. It is close to perfect at 0.996 for resolutions since the programmatic version of the model was designed with resolutions in mind. Complications may be interpreted as the negative release of tension. Even though a large number of contributing $\Delta\Delta \hat{t}_j$ are negative, only one negative $\widetilde{\Delta\Delta t}$ can be seen in the table due to the influence of nonnegative candidates. Nevertheless, $\widetilde{\Delta\Delta t}$ shows a strong negative correlation of -0.761 with [45] for complications (agreeing with the fact that this resolution is negative). As earlier explained, with excursions the perception of a

succeeding chord is also influenced by the rising or falling of parallel melodies. Unfortunately, descending excursions were insufficiently popular in Palestrina and only V-IV was being tallied. For escalating excursions, however, we have enough statistics to compute a correlation of 0.863. We have also computed the correlation across all other chords separately from complications (because, as explained, they correlate negatively) to be 0.970.

6. Discussion

Addressing the Fundamental Questions of Psychoacoustic Harmony. At this point, let us address the fundamental questions of psychoacoustic harmony as promised at the start of this paper in the context of subharmonic modulations. We will begin with question 2 and leave the first question for the last.

(2) We discussed the definition and explanation of stationary harmony, i.e., *what sounds good* and *why*, or, mathematically, to quantify $\varepsilon\{X_n\}$, where $\varepsilon\{\}$ denotes *the harmonious effect of* and X_n represents chord *n*.

TABLE 3: Tabulation of $\widetilde{\Delta\Delta t}$ for chords that most commonly follow V against Palestrina's chord tendencies as cited in [45].

	I	V^7	iii^6	V/V	V^2	vi	V^6/V	vi^7	i	iii	vi^6	I^6	ii^6	vii°	V^6	ii	IV
Convergence*	C	D	D	C	D	X↑	C	X↑	C	C	X↑	C	C	D	D	C	X↓
Frequency†	42	7	6	2	2	11	1	1	0	4	1	5	2	0.5	2	2	6
$\widetilde{\Delta\Delta t}$	444	-1.5	1.8	1.9	0.4	3.4	2.3	1.9	6.9	1.2	0.3	39.6	1.9	7.4	4.3	2.5	192

*States of convergence:
C denotes convergence of $\Delta\hat{t}$.
D denotes divergence of $\Delta\hat{t}$.
X↑ denotes escalating excursion of $\Delta\hat{t}$.
X↓ denotes descending excursion of $\Delta\hat{t}$.
†In percent, as read off the histogram of chord tendencies from [45] computed over a dataset of 11000 chords from Palestrina.

TABLE 4: Tabulation of correlations between $\widetilde{\Delta\Delta t}$ and Palestrina's chord use statistics as collated in [45]. Correlations are listed in the top row with corresponding significance in brackets below.

Resolutions*	Complications†	Excursions‡		All excl. comp.†	All
		Escalating	Descending§		
0.996	-0.761	0.863	-	0.970	0.903
(0.0000)	(0.1353)	(0.3366)		(0.0000)	(0.0000)

*Our model is designed to compute tension release in resolution.
†Complications in music may be interpreted as negative tension resolutions; hence, correlation seen is negative.
‡Excursions usually encompass tension release; however, apart from resolution alone, the perception of succeeding chords are also influenced by the rising or falling of parallel melodies.
§Apart from the descending excursions leading to IV, insufficient other descending transitions are recorded to compute its correlation.

With large subharmonic tension being perceived as dissonance while small subharmonic modulations are perceived as consonance, the aesthetics of a chord may be visualized in the subharmonic tension acting on its shortest common subharmonics. Mathematically, they are inversely related. As described by (19), $\varepsilon\{X\} \propto 1/\Delta\hat{t}$.

(3) We have the definition and explanation of transitional harmony, i.e., *what sounds good*, *why*, and *when*, or, mathematically, to quantify $\varepsilon\{X_1 \longrightarrow X_2\}$, where ' $\longrightarrow$ ' denotes transition from one chord to another.

The aesthetics of a chord transition may be visualized in the release of subharmonic tension at the shortest common subharmonics of the succeeding chord. As explained in (22) and indicated by the arrows in Figure 12, this refers to the transition to the shortest common subharmonics of the succeeding chord from the nearest subharmonics of the preceding chord. Thus, resolution (tension release) in a chord transition is perceived in the convergence of $\Delta\hat{t}$ (where $\Delta\Delta\hat{t} > 0$) while what we call complication (build-up of tension or negative resolution) is seen in its divergence (where $\Delta\Delta\hat{t} < 0$). Mathematically, as described by (24), $\varepsilon\{X_1 \longrightarrow X_2\} \propto \Delta\Delta\hat{t}_{X_1 \longrightarrow X_2}$.

(4) We have the following phenomena.

(a) A chord that sounds better than another out of context can sound worse than being in context [42]. Given $\varepsilon\{X_2\} > \varepsilon\{X_3\}$ this shows that $\varepsilon\{X_1 \longrightarrow X_2\} < \varepsilon\{X_1 \longrightarrow X_3\}$

The section on subharmonic modulations differentiates between stationary tension and transitional tension. The tension release brought about by a transition to a chord may be large even for high tension succeeding chords. To prove this, we will use an example with E^7, G, and Am^7. Taking $E^7 = \{b_3, d_4, e_4, g^{\#}_4\}$, $G = \{g_3, b_3, d_4, g_4\}$, and $Am^7 = \{a_3, c_4, e_4, g_4, a_4\}$, the stationary subharmonic tension for G and Am^7 may be computed by (18) to be $\Delta\hat{t}_G = 0.902\%$ and $\Delta\hat{t}_{Am^7} = 6.849\%$, respectively. Thus, $\varepsilon\{G\} > \varepsilon\{Am^7\}$, whereas the transitional subharmonic resolution (tension resolution) for $E^7 \longrightarrow G$ and $E^7 \longrightarrow Am^7$ may be computed by (22) to be $\Delta\Delta\hat{t}_{E^7 \longrightarrow G} = 8.783\%$ and $\Delta\Delta\hat{t}_{E^7 \longrightarrow Am^7} = 10.748\%$, respectively. Thus, $\varepsilon\{E^7 \longrightarrow G\} < \varepsilon\{E^7 \longrightarrow Am^7\}$ despite the fact that $\varepsilon\{G\} > \varepsilon\{Am^7\}$.

(b) A chord that sounds better than another in one context can sound worse than being in another context [42]. Given $\varepsilon\{X_4 \longrightarrow X_2\} > \varepsilon\{X_4 \longrightarrow X_3\}$ this shows that $\varepsilon\{X_1 \longrightarrow X_2\} < \varepsilon\{X_1 \longrightarrow X_3\}$

With reference to (22) and our answer in question 3, since our ears identify the subharmonics of preceding notes that correspond to the succeeding common subharmonic, transitional harmony is contextual. Continuing from our answer to question 4a, we take D^7 to be $D^7 = \{c_4, d_4, f^{\#}_4, a_4\}$. The transitional subharmonic resolution (tension resolution) for $D^7 \longrightarrow G$ and $D^7 \longrightarrow Am^7$ may be computed by (22) to be $\Delta\Delta\hat{t}_{D^7 \longrightarrow G} = 11.421\%$ and $\Delta\Delta\hat{t}_{D^7 \longrightarrow Am^7} = 4.540\%$,

respectively. Thus, $\varepsilon\{D^7 \longrightarrow G\} > \varepsilon\{D^7 \longrightarrow Am^7\}$ despite the fact that $\varepsilon\{E^7 \longrightarrow G\} < \varepsilon\{E^7 \longrightarrow Am^7\}$.

(5) phenomenon that the transition from a low-tension chord to a high-tension one can still bring about the effect of tension release (resolution). Given $\varepsilon\{X_1\} < \varepsilon\{X_2\}$ this shows that $\varepsilon\{X_1 \longrightarrow X_2\} > 0$.

The answer to this is in the independence of stationary and transitional tension, as established in our answer to Question 4a.

Taking $E = \{b_3, e_4, g^{\#}{}_4\}$ and $Am^7 = \{a_3, c_4, e_4, g_4, a_4\}$, the transitional subharmonic resolution (tension resolution) for $E \longrightarrow Am^7$ may be computed by (22) to be $\Delta\Delta\hat{t}_{E \longrightarrow Am^7} = 4.323\%$. The stationary subharmonic tension for E and Am^7 may be computed by (18) to be $\Delta\hat{t}_E = 0.902\%$ and $\Delta\hat{t}_{Am^7} = 6.849\%$, respectively. Hence, $\varepsilon\{E \longrightarrow Am^7\} > 0$ despite the fact that $\varepsilon\{Am^7\} < \varepsilon\{E\}$.

(6) There is the phenomenon that the effect of harmony is greater than the sum of its parts [18, 60]. $\varepsilon\{x_1 + x_2 + x_3\} \gg \varepsilon\{x_1\} + \varepsilon\{x_2\} + \varepsilon\{x_3\}$

Apart from certain exceptions with rational intonation and octaves, the stationary tension of any combination of unique notes is observed to be larger than zero on the subharmonic plot. Hence, $\Delta\hat{t}_{\{x_1+x_2+x_3\}} > 0$. Likewise, the stationary tension of each note on its own is observed to be zero on the subharmonic plot. Hence, $\Delta\hat{t}_{\{x_1\}} = 0$, $\Delta\hat{t}_{\{x_2\}} = 0$, and $\Delta\hat{t}_{\{x_3\}} = 0$ for all x_1, x_2, and x_3 within musical range. Thus, by (19), $\varepsilon\{x_1 + x_2 + x_3\} \gg 0$, whereas by (20) $\varepsilon\{x_1\} = 0$, $\varepsilon\{x_2\} = 0$, $\varepsilon\{x_3\} = 0$, and $\varepsilon\{x_1\} + \varepsilon\{x_2\} + \varepsilon\{x_3\} = 0$. Therefore, $\varepsilon\{x_1 + x_2 + x_3\} \gg \varepsilon\{x_1\} + \varepsilon\{x_2\} + \varepsilon\{x_3\}$.

7. Conclusion

In this paper the notion of interharmonic and subharmonic modulations was proposed as a psychophysical basis for both stationary and transitional harmony.

In the domain of stationary harmony (tension in chord sonorities), this work presents subharmonic modulations as an integral complement to interharmonic modulations and shows how perceptual tensions [18, 36, 58, 59] and consonances [17, 19, 44] may be visualized through which.

In the domain of transitional harmony (resolution in chord progression), it unlocks the means of physically identifying, quantizing, and, thus, verifying perceptual resolutions and complications [18, 42] in acoustic features that have until now remained abstract and nontangible.

This work can be seen to bind prevailing psychoacoustic schools into a single theory. The Helmholtz school [3, 8, 12–17, 19, 20, 23] is represented by the interharmonic Δf in (11). The Pythagorean school [5, 6, 11] generally seeks small values of integer k_i in (15) and (16) while requiring Δt_i to be zero. Taking this further, if Δt_i is ignored, f_{sub} in (15) would then correspond to the fusion tone in Stumpf's tonal fusion theory [3, 30]. Euler's *gradus suavitatis* [11] graded the goodness of k_i-combinations for $\Delta t_i = 0$. The adoption of 12-tone equal temperament [12, 33, 34] sought

to evenly distribute interharmonic Δf in (11). Since the aforementioned conditions may be generalized by a central theory of modulations across adjacent (interharmonic) and distant (subharmonic) sinusoids which stems from (3), this effectively integrates them into a general theory.

Computed values correlate strongly with perception and harmony-use statistics for both stationary (tension) and transitional (resolution) harmony.

Finally, this paper presented a psychoacoustic solution to the five fundamental questions of harmony.

Conflicts of Interest

The authors declare no conflicts of financial interest.

Supplementary Materials

Supplementary 1. Supplementary Figure S1: Sinusoidal Summation across 8 Amplitude Ratios. $A \cos \omega_1 t + B \cos \omega_2 t$ for various values of B normalized to $A=1$.

Supplementary 2. Supplementary Table S1: Correlation for Intervals. There is tabulation of ordinal ranking of dyads (intervals) using $T_{\Delta f | \Delta t}$ against available rankings collated in [43].

Supplementary 3. Supplementary Table S2: Correlation for Triads. There is tabulation of ordinal ranking of triads using $T_{\Delta f | \Delta t}$ against available rankings collated in [43].

Supplementary 4. Supplementary Audio S1: Low-Frequency Modulation. There is audio example of consonant low-frequency modulation with $\overline{f} = 440Hz$ and $\Delta f = 0.5Hz$.

Supplementary 5. Supplementary Audio S2: Beating Frequency. There is audio example of dissonant high-frequency modulation with $\overline{f} = 440Hz$ and $\Delta f = 70Hz$.

Supplementary 6. Supplementary Video S1: Low versus High Subharmonic Tension. There is video example comparing subharmonic wave deformation in low tension C chord against high tension Cm^7 chord.

Supplementary 7. Supplementary Video S2: Stationary Tensions in Pachelbel's Canon. We visualize stationary harmony with subharmonic tension in Pachelbel's Canon [[70]].

Supplementary 8. Supplementary Video S3: Transitional Tensions in Moonlight Sonata. We visualize transitional harmony with subharmonic tension in Beethoven's Moonlight Sonata [71].

References

[1] J. P. Rameau, *Treatise on Harmony*, Courier Corporation, 1722.

[2] P. Hindemith, *The Craft of Musical Composition: Theoretical Part*, vol. 1, Schott Co Ltd, 1970.

[3] N. McLachlan, D. Marco, M. Light, and S. Wilson, "Consonance and pitch," *Journal of Experimental Psychology: General*, vol. 142, no. 4, pp. 1142–1158, 2013.

[4] J. Sauveur, *Principes d'acoustique et de musique: ou, Système général des intervalles des sons*, Editions Minkoff, 1701.

[5] K. J. Hsü and A. J. Hsü, "Fractal geometry of music," *Proceedings of the National Academy of Sciences*, vol. 87, no. 3, pp. 938–941, 1990.

[6] B. V. Rivera, "Theory Ruled by Practice: Zarlino's Reversal of the Classical System of Proportions," *Indiana Theory Review*, vol. 16, pp. 145–170, 1995.

[7] G. Pont, "Philosophy and Science of Music in Ancient Greece," *Nexus Network Journal*, vol. 6, no. 1, pp. 17–29, 2004.

[8] H. v. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, Longmans, Green, 1912.

[9] E. Wellesz and J. A. Westrup, *Ancient and Oriental Music*, vol. 1, Oxford University Press, 1957.

[10] J. M. Barbour, *Tuning and Temperament: A Historical Survey*, Courier Corporation, 2004.

[11] A. Gräf, "On musical scale rationalization," in *Proceedings of the International Computer Music Conference, ICMC 2006*, pp. 91–98, USA, November 2006.

[12] R. Plomp and W. J. Levelt, "Tonal Consonance and Critical Bandwidth," *The Journal of the Acoustical Society of America*, vol. 38, no. 4, pp. 548–560, 1965.

[13] P. N. Johnson-Laird, O. E. Kang, and Y. C. Leong, "On musical dissonance," *Music Perception*, vol. 30, no. 1, pp. 19–35, 2012.

[14] D. L. Bowling and D. Purves, "A biological rationale for musical consonance," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 112, no. 36, pp. 11155–11160, 2015.

[15] H. E. White and D. H. White, *Physics and Music: The Science of Musical Sound*, Courier Corporation, 2014.

[16] G. Dillon, "Calculating the dissonance of a chord according to Helmholtz theory," *The European Physical Journal Plus*, vol. 128, no. 8, Article 90, 2013.

[17] I. S. Lots and L. Stone, "Perception of musical consonance and dissonance: An outcome of neural synchronization," *Journal of the Royal Society Interface*, vol. 5, no. 29, pp. 1429–1434, 2008.

[18] N. D. Cook and T. X. Fujisawa, *The Psychophysics of Harmony Perception: Harmony Is a Three-Tone Phenomenon*, 2006.

[19] Y. I. Fishman, I. O. Volkov, M. D. Noh et al., "Consonance and dissonance of musical chords: Neural correlates in auditory cortex of monkeys and humans," *Journal of Neurophysiology*, vol. 86, no. 6, pp. 2761–2788, 2001.

[20] E. Zwicker and H. Fastl, *Psycho-Acoustics: Facts and Models*, Springer, Berlin, Germany, 2nd edition, 1999.

[21] W. A. Sethares, "Local consonance and the relationship between timbre and scale," *The Journal of the Acoustical Society of America*, vol. 94, no. 3, pp. 1218–1228, 1993.

[22] R. Parncutt, *Harmony: A Psychoacoustical Approach*, vol. 19 of *Springer Science & Business Media*, Springer, Berlin, 1989.

[23] W. Hutchinson and L. Knopoff, "The acoustic component of Western consonance," *Journal of New Music Research*, vol. 7, no. 1, pp. 1–29, 1978.

[24] A. Kameoka and M. Kuriyagawa, "Consonance Theory Part I: Consonance of Dyads," *The Journal of the Acoustical Society of America*, vol. 45, no. 6, pp. 1451–1459, 1969.

[25] A. Kameoka and M. Kuriyagawa, "Consonance Theory Part II: Consonance of Complex Tones and Its Calculation Method," *The Journal of the Acoustical Society of America*, vol. 45, no. 6, pp. 1460–1469, 1969.

[26] P. Lalitte, "The theories of Helmholtz in the work of Varese," *Contemporary Music Review*, vol. 30, no. 5, pp. 327–342, 2011.

[27] E. G. Schellenberg and S. E. Trehub, "Frequency ratios and the perception of tone patterns," *Psychonomic Bulletin & Review*, vol. 1, no. 2, pp. 191–201, 1994.

[28] K. Itoh, S. Suwazono, and T. Nakada, "Cortical processing of musical consonance: An evoked potential study," *NeuroReport*, vol. 14, no. 18, pp. 1061–1069, 2003.

[29] G. M. Bidelman, "The role of the auditory brainstem in processing musically relevant pitch," *Frontiers in Psychology*, vol. 4, no. 264, 2013.

[30] C. Stumpf, "Konsonanz und dissonanz [Consonance and dissonance]," *Beiträge zur Akustik und Musikwissenschaft*, vol. 1, pp. 1–108, 1898.

[31] W. S. B. Woolhouse, *Essay on Musical Intervals, Harmonics, and The Temperament of The Musical Scale*, 1835.

[32] D. D. Nolte, *Galileo Unbound: A Path Across Life, the Universe and Everything*, Oxford University Press, 2018.

[33] H. L. Goodman and Y. E. Lien, "A Third Century AD Chinese System of Di-Flute Temperament: Matching Ancient Pitch-Standards and Confronting Modal Practice," *The Galpin Society Journal*, pp. 3–24, 2009.

[34] G. J. Cho, *The Discovery of Musical Equal Temperament in China and Europe in The Sixteenth Century*, vol. 93, Edwin Mellen Press, 2003.

[35] T. Christensen and J. P. Rameau, "Eighteenth-century science and the" corps sonore:" the scientific background to rameau's" principle of harmony," *Journal of Music Theory*, vol. 31, no. 1, pp. 23–50, 1987.

[36] E. Bigand, R. Parncutt, and F. Lerdahl, "Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training," *Perception & Psychophysics*, vol. 58, no. 1, pp. 125–141, 1996.

[37] P. F. Broman, C. Geertz, and O. Neurath, Music Theory Art, Science, or What? What Kind of Theory Is Music Theory? 17 (2007).

[38] E. Bigand and B. Poulin-Charronnat, "Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training," *Cognition*, vol. 100, no. 1, pp. 100–130, 2006.

[39] T. M. Fiore, Music and mathematics (2007). Recuperado de: http://www-personal.umd.umich.edu/~tmfiore/1/musictotal.pdf.

[40] R. Parncutt, "Revision of Terhardt's psychoacoustical model of the root (s) of a musical chord. Music Perception," *An Interdisciplinary Journal*, vol. 6, no. 1, pp. 65–93, 1988.

[41] R. Scruton, *The Aesthetics of Music*, Oxford University Press, 1999.

[42] P. I. Tchaikovsky, *Guide to The Practical Study of Harmony*, Courier Corporation, (1872/2005).

[43] F. Stolzenburg, "Harmony perception by periodicity detection," *Journal of Mathematics and Music*, vol. 9, no. 3, pp. 215–238, 2015.

[44] L. Hofmann-Engl, "Consonance/DissonanceA historical Perspective," in *Proceedings of the 11th International Conference on Music Perception and Cognition*, pp. 852–856, 2010.

[45] D. Tymoczko, *A Study on the Origins of Harmonic Tonality. Paper delivered to the national meeting of the Society for Music Theory*, Indianapolis, Indiana, USA, 2014.

[46] D. Tymoczko, *A Geometry of Music: Harmony and Counterpoint in The Extended Common Practice*, Oxford University Press, 2010.

[47] D. Tymoczko, "Scale theory, serial theory and voice leading," *Music Analysis*, vol. 27, no. 1, pp. 1–49, 2008.

[48] A. Honingh and R. Bod, "In search of universal properties of musical scales," *Journal of New Music Research*, vol. 40, no. 1, pp. 81–89, 2011.

[49] G. J. Balzano, "What are musical pitch and timbre?" *Music Perception: An Interdisciplinary Journal*, vol. 3, no. 3, pp. 297–314, 1986.

[50] G. J. Balzano, "The pitch set as a level of description for studying musical pitch perception," in *Music, Mind, and Brain*, pp. 321–351, Springer, Boston, MA, USA, 1982.

[51] N. Carey and D. Clampitt, "Aspects of well-formed scales," *Music Theory Spectrum*, vol. 11, no. 2, pp. 187–206, 1989.

[52] D. Purves, *Music as Biology*, Harvard University Press, 2017.

[53] K. Z. Gill and D. Purves, "A biological rationale for musical scales," *PLoS ONE*, vol. 4, no. 12, Article ID e8144, 2009.

[54] D. A. Schwartz, C. Q. Howe, and D. Purves, "The statistical structure of human speech sounds predicts musical universals," *The Journal of Neuroscience*, vol. 23, no. 18, pp. 7160–7168, 2003.

[55] G. Langner, M. Sams, P. Heil, and H. Schulze, "Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: Evidence from magnetoencephalography," *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology*, vol. 181, no. 6, pp. 665–676, 1997.

[56] G. Langner and C. E. Schreiner, "Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms," *Journal of Neurophysiology*, vol. 60, no. 6, pp. 1799–1822, 1988.

[57] T. Houtgast, "Subharmonic pitches of a pure tone at low S/N ratio," *The Journal of the Acoustical Society of America*, vol. 60, no. 2, pp. 405–409, 1976.

[58] M. M. Farbood, "A parametric, temporal model of musical tension," *Music Perception*, vol. 29, no. 4, pp. 387–428, 2012.

[59] C. K. Madsen and W. E. Fredrickson, "The experience of musical tension: a replication of nielsen's research using the continuous response digital interface," *Journal of Music Therapy*, vol. 30, no. 1, pp. 46–63, 1993.

[60] E. Terhardt, G. Stoll, and M. Seewann, "Algorithm for extraction of pitch and pitch salience from complex tonal signals," *The Journal of the Acoustical Society of America*, vol. 71, no. 3, pp. 679–688, 1982.

[61] S. Drake, "Renaissance music and experimental science," *Journal of the History of Ideas*, pp. 483–500, 1970.

[62] F. Stolzenburg, "Harmony perception by periodicity and granularity detection," *Cambouropolos*, pp. 958-959, 2012.

[63] M. Ebeling, "Neuronal periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion," *The Journal of the Acoustical Society of America*, vol. 124, no. 4, pp. 2320–2329, 2008.

[64] L. J. Hofmann-Engl, *Virtual Pitch and The Classification of Chords in Minor and Major Keys*, 2008.

[65] S. Ternström, "Physical and acoustic factors that interact with the singer to produce the choral sound," *Journal of Voice*, vol. 5, no. 2, pp. 128–143, 1991.

[66] D. Temperley and D. Tan, "Emotional connotations of diatonic modes," *Music Perception*, vol. 30, no. 3, pp. 237–257, 2013.

[67] E. C. Bairstow, *Counterpoint and harmony*, Read Books Ltd, 2013.

[68] M. J. Tramo, "Music of the hemispheres," *Science*, vol. 291, no. 5501, pp. 54–56, 2001.

[69] D. Harrison, *Harmonic Function in Chromatic Music: A Renewed Dualist Theory and An Account of Its Precedents*, University of Chicago Press, 1994.

[70] J. Pachelbel, *Canon And Gigue for 3 Violins and Basso Continuo*, 1680-1706.

[71] L. v. Beethoven, Piano Sonata No. 14 in C♯ minor, "Quasi una fantasia", Op. 27, No. 2 (1801).

[72] E. Aldwell and A. Cadwallader, *Harmony and voice leading*, Cengage Learning, 2018.

[73] M. Guernsey, "The Role of Consonance and Dissonance in Music," *The American Journal of Psychology*, vol. 40, no. 2, pp. 173–204, 1928.

[74] D. Tymoczko, "The Geometry of Musical Chords," *Science*, vol. 313, no. 5783, pp. 72–74, 2006.

[75] M. Farbood and B. Schöner, "Analysis and Synthesis of Palestrina-Style Counterpoint Using Markov Chains," in *Proceedings of the ICMC*, 2001.

[76] J. Kursell, "A Third Note: Helmholtz, Palestrina, and the Early History of Musicology," *Isis*, vol. 106, no. 2, pp. 353–366, 2015.

[77] C. Marvin, *Giovanni Pierluigi da Palestrina: A Research Guide*, Routledge, 2013.