Designing a comprehensive model of tonalness as an empirical basis for a quantitative theory of musical harmony

Introduction

When Jean-Phillippe Rameau first published *Traité de l'harmonie réduite à ses principes naturels* in 1722, questions into the nature of musical harmony and harmonic structure were posed in the context of music with a much more limited harmonic vocabulary than that which came to exist three centuries later. A generalized theory of harmony in the present era should ideally encompass any continuous or discontinuous frequency distribution consisting of any given harmonic or inharmonic timbre under both static and dynamic conditions.

Fourier analysis has become an indispensable tool for the processing of frequency and phase composition of acoustic signals. However, a variety of empirical studies have demonstrated the limitations of using purely spectral models of human hearing[1,2,3,4,5,6]. Tonalness is the propensity of the human auditory system (HAS) to perceive virtual pitches as common fundamental frequencies of a given distribution of spectral pitches. This essay will show how mathematical models of virtual pitch and

---

1 Fletcher, *Loudness, Its Definition, Measurement and Calculation*.

2 McDermott, *Individual Differences Reveal the Basis of Consonance*.

3 Moore, *Cochlear hearing loss*.

4 Peretz, *Cortical deafness to dissonance*.

5 Plomp, *Tonal Consonance and Critical Bandwidth*.

6 Schellenberg, *Frequency ratios and the discrimination of pure tone sequences*.

critical band can be combined with Fourier analysis to yield a comprehensive model of tonalness which can form the basis for a quantitative theory of musical harmony.

Defining Tonalness

Fourier analysis uses a mathematical transform to break down complex audio signals into their frequency components. It is already widely incorporated into the study of acoustics and psychoacoustics. However, there are a variety of psychoacoustic phenomena such as virtual pitch and critical band that cannot be explained by Fourier analysis alone. This paper will argue that these phenomena can be used to explain a variety of features of harmonic perception in psychoacoustics and music theory.

Equation 1 defines the tonalness function as a psychoacoustic Fourier transform that incorporates a temporal window to model how the inner ear perceives virtual pitch uncertainty about a given input harmonic h. Equations 2–4 combine to create something similar to the Glasberg & Moore critical band model[7].

$$\tau(f,t) = 2\sum_{h=1}^{\infty}\frac{1}{h}\int_{-\infty}^{\infty}\frac{e^{-\frac{\left(u-\frac{t}{h}\right)^2}{2\sigma_t(f)^2}}}{\sqrt{2\pi}\sigma_t(f)}e^{-2\pi iuhf}P(u)\,du$$

EQUATION 1. TONALNESS FUNCTION.

$$\sigma_t(f) = \frac{1}{2\pi\sigma(f)}$$

EQUATION 2. TEMPORAL WINDOW FUNCTION.

7 Glasberg, 103–138.

$$\sigma(f) = \kappa\left(d_1 + d_2 f\right)$$

EQUATION 3. CRITICAL BAND MODEL.

$$\left\{d_1, d_2\right\} = \{24.7, 0.107939\}$$

EQUATION 4. CRITICAL BAND CONSTANTS.

We will set κ to 0.25 to approximate the 1/4 critical band threshold discovered in the dissonance research of Plomp and Levelt[8]. We see from Equation 3 that $\sigma(f)$ is comprised of a temporal window constant d_1 that originates from the latency of the nervous system and a frequency scaling term d_2 that scales with the resolution of pitch perception in the inner ear.

Equation 1 models virtual pitches as something that arise at the subharmonic of every spectral pitch. For a simple vibrating system such as a one-dimensional standing wave, any given harmonic h of its fundamental f_0 holds an equivalent amount of energy at an amplitude of $\frac{a_0}{h}$ where a_0 is the amplitude of the fundamental[9]. Therefore, we will adopt $\frac{1}{h}$ as the amplitude scale of a virtual pitch at subharmonic h.

To explore the implications of this tonalness model, we will first need an input sound pressure signal (SPS). Equations 5–7 collectively define the tonalness function for an additive sinusoidal SPS.

---

[8] Plomp, *Tonal Consonance and Critical Bandwidth*.

[9] See the Appendix.

$$P(t) = \sum_{i=1}^{n} a_i \cos\left(2\pi b_i t + c_i\right)$$

EQUATION 5. ADDITIVE SINUSOIDAL SOUND PRESSURE SIGNAL.

$$\tau(f, t) = \sum_{h=1}^{\infty} \sum_{i=1}^{n} U_{h,i}(f)\left(\cos\left(2\pi f t - \frac{2\pi t b_i}{h} - c_i\right) - i \sin\left(2\pi f t - \frac{2\pi t b_i}{h} - c_i\right)\right)$$

EQUATION 6. SINUSOIDAL TONALNESS FUNCTION.

$$U_{h,i}(f) = \frac{a_i}{h} e^{-\frac{(hf - b_i)^2}{2\sigma(f)^2}}$$

EQUATION 7. SINUSOIDAL SMOOTHING FUNCTION.

The tonalness function serves as a psychoacoustic Fourier transform. However, unlike the Fourier transform, tonalness varies in time. We will show that psychoacoustic phenomena such as beating, combination tones, tonal fusion, and dissonance are a mathematical consequence of the incorporation of critical band and virtual pitch into Fourier analysis.

The tonalness function provides a quantitative two dimensional description of roughness and harmonicity for a given SPS. To compare the tonalness function to a spectral frequency distribution, we will first consider a spectral input of three precise pure tones with equal loudness: 440 Hz, 554 Hz, and 659 Hz. This approximates a major triad tuned to 12 tone equal temperament. Its waveform can be represented by the equation $P(t) = \cos(2\pi 440t) + \cos(2\pi 554t) + \cos(2\pi 659t)$. Figure 1 plots the intensity of the Fourier transform in red and $|\tau(f,0)|^2$ in black for this spectral input.

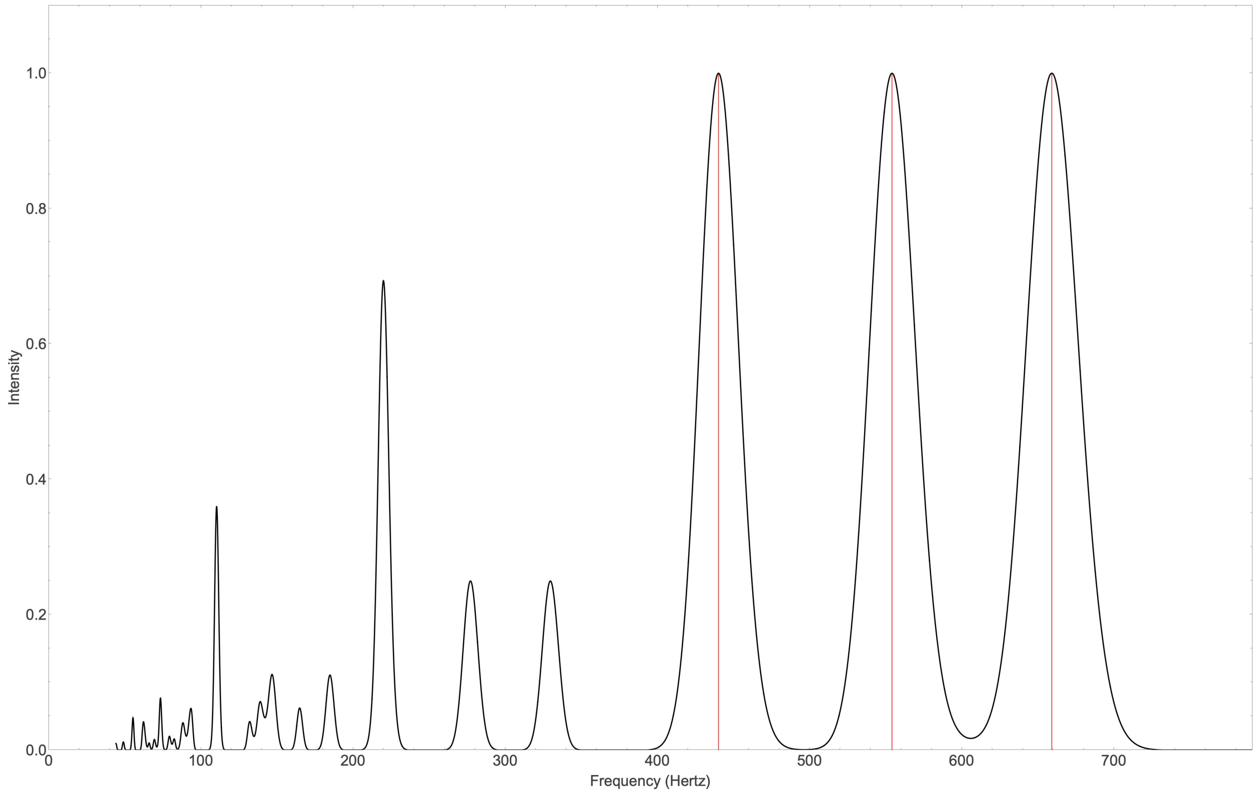

FIGURE 1. SPECTRAL (RED) VERSUS TONALNESS (BLACK) INTENSITY[10] FOR THREE PURE TONES: 440 HZ, 554 HZ, AND 659 HZ.

As we observe in Figure 1, the tonalness function models virtual pitches as a series of bell curves occurring at the subharmonics of each of the spectral pitches. The height of these curves depends primarily on the alignment of the subharmonics, frequency response, and sum of the inverse of their harmonic numbers, while the width depends on the critical band. In Figure 1 the two most prominent virtual pitches from this set of spectral pitches occur near 110 Hz and 220 Hz. This can be understood qualitatively via examination of the arrangement of spectral pitches. If we are to perceive 110 Hz as a common fundamental frequency, we would hear 440 Hz and 554 Hz as an approximate 4:5 major third and 554 Hz and 659 Hz as an approximate 5:6

[10] For our purposes, intensity is defined as the magnitude squared.

minor third. If we are to perceive 220 Hz as a fundamental frequency, we would hear 440 Hz and 659 Hz as an approximate 2:3 perfect fifth. Since 2:3 is composed of smaller harmonic numbers than 4:5 or 5:6, we should typically expect a more intense virtual pitch to be perceived at the 220 Hz location. However, since two different intervals have aligned virtual pitches near the 110 Hz location, the virtual pitch there is perceived almost as intensely as what we observe from the 220 Hz where only one interval has an aligned virtual pitch.

Virtual pitches can be understood as both a proxy for the harmonicity of a given SPS at a given spectral location and also as pseudo-pitches perceived directly by the HAS. Virtual pitches are not physically present in a SPS but are a consequence of how the HAS responds to the sounds that it hears. In general, the auditory system does not perceive just one virtual pitch but rather many of them simultaneously. By this reasoning, it is possible for the HAS to perceive a high degree of harmonicity from one part of a harmony while also perceiving a lower degree of harmonicity from another part of a harmony. The phenomenon of virtual pitch can explain the existence of combination tones, the masking of harmonic overtones into timbres, and many other aspects of harmonic perception.

While static tonalness plots contain a great deal of psychoacoustic information about a given harmony, we must remember that tonalness plots evolve with time. To better understand the dynamic nature of the tonalness function, we will invoke two additional functions. Equation 8 defines a steady state tonalness (SST) function by maximizing the intensity of the tonalness function over time interval T for each frequency value f, while Equation 9 defines a tonalness variance (TV) function by maximizing the first derivative of the frequency scaled tonalness intensity over time interval T for each frequency value f. Equations 10 and 11 define SST and TV for an additive sinusoidal SPS.

$$\psi_1(f,t) = \text{Maximize}\left[|\tau(f,u)|^2, u \in |u-t| < \frac{\sigma_t(0)}{2}\right]$$

EQUATION 8. STEADY STATE TONALNESS FUNCTION

$$\psi_2(f,t) = \frac{1}{f}\text{Maximize}\left[\frac{\partial|\tau(f,u)|^2}{\partial u}, u \in |u-t| < \frac{\sigma_t(0)}{2}\right]$$

EQUATION 9. TONALNESS VARIANCE FUNCTION

$$\psi_1(f,t) = \left(\sum_{h=1}^{\infty}\sum_{i=1}^{n} U_{h,i}(f)\right)^2$$

EQUATION 10. SINUSOIDAL STEADY STATE TONALNESS FUNCTION

$$\psi_2(f,t) = \frac{2\pi}{f}\sum_{g=1}^{\infty}\sum_{h=1}^{\infty}\sum_{i=1}^{n}\sum_{j=1}^{n} U_{g,i}(f)U_{h,j}(f)\left|\frac{b_j}{h} - \frac{b_i}{g}\right|$$

EQUATION 11. SINUSOIDAL TONALNESS VARIANCE FUNCTION

The temporal window $\sigma_t(0)$ threshold of about 0.026 seconds differentiates fluctuations that are slow enough (take longer than $\sigma_t(0)$) to be perceived rhythmically in the SST curve from those that are too fast to be perceived as rhythmic (faster than $\sigma_t(0)$), and are instead perceived as unsteadiness in the TV curve. Figure 2 plots the SST and TV for our three pure tone example.

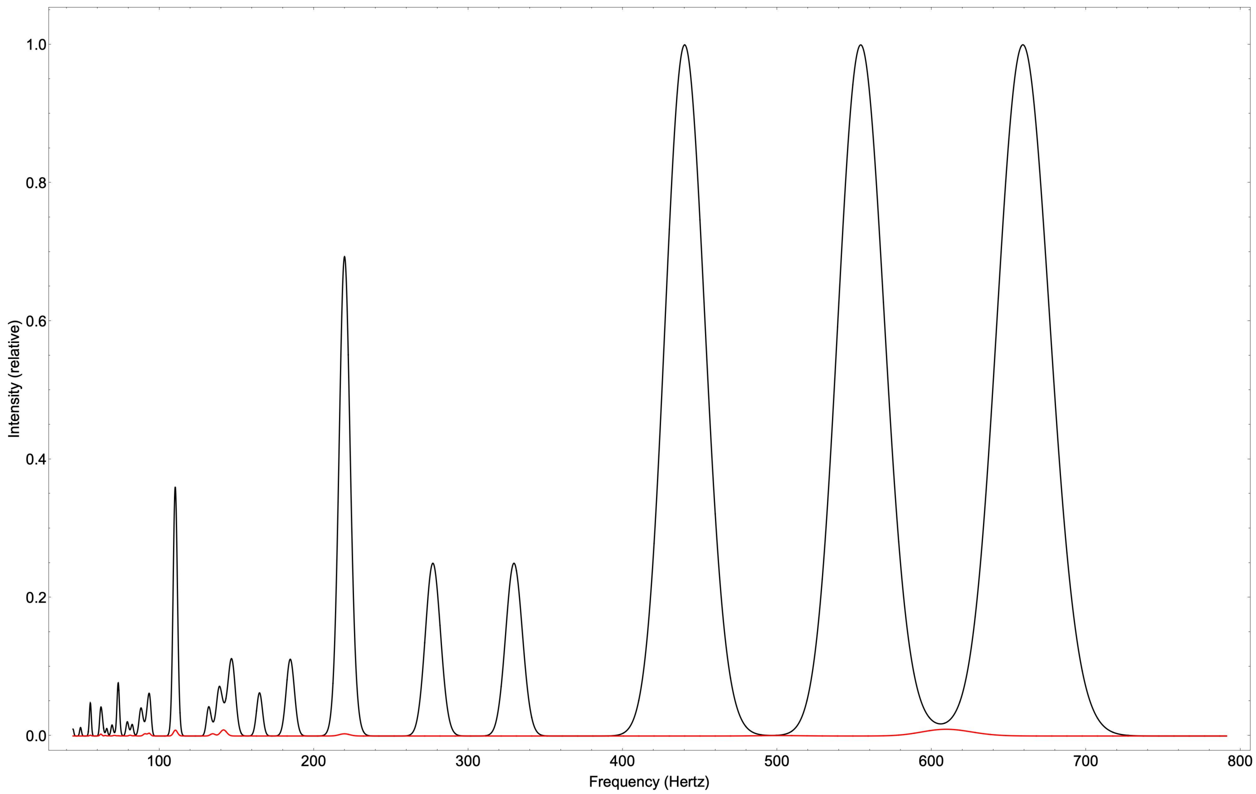

FIGURE 2. SQUARE OF STEADY-STATE TONALNESS (BLACK) AND TONALNESS VARIANCE (RED) FOR THREE PURE TONES: 440 HZ, 554 HZ, AND 659 HZ.

As we observe, Figure 2 closely mirrors the result of Figure 1 since most of the oscillations in tonalness are small and too quick to appear the SST curve. As we see from Figure 2, there is a significant intensity of variation in the TV curve near 110 Hz and 220 Hz. We should interpret this as a subtle mistuning of the simple harmonic ratios that occurs between the spectral pitches. However, we also notice that this variation is also intense in a few other locations. In general, the TV function models roughness that the HAS perceives between virtual pitches, spectral pitches, and even interference between virtual pitches and spectral pitches.

Figure 3 plots the square of SST and TV for a different three pure tone example: 440 Hz, 550 Hz, and 660 Hz. This is known musically as a justly tuned major triad. As we see, the intensity of variation in the TV curve near 110 Hz and 220 Hz is now near zero while the peak SST values at these points is slightly taller than in the previous example. We should expect this because the spectral pitches are now perfectly tuned to simple harmonic ratios.

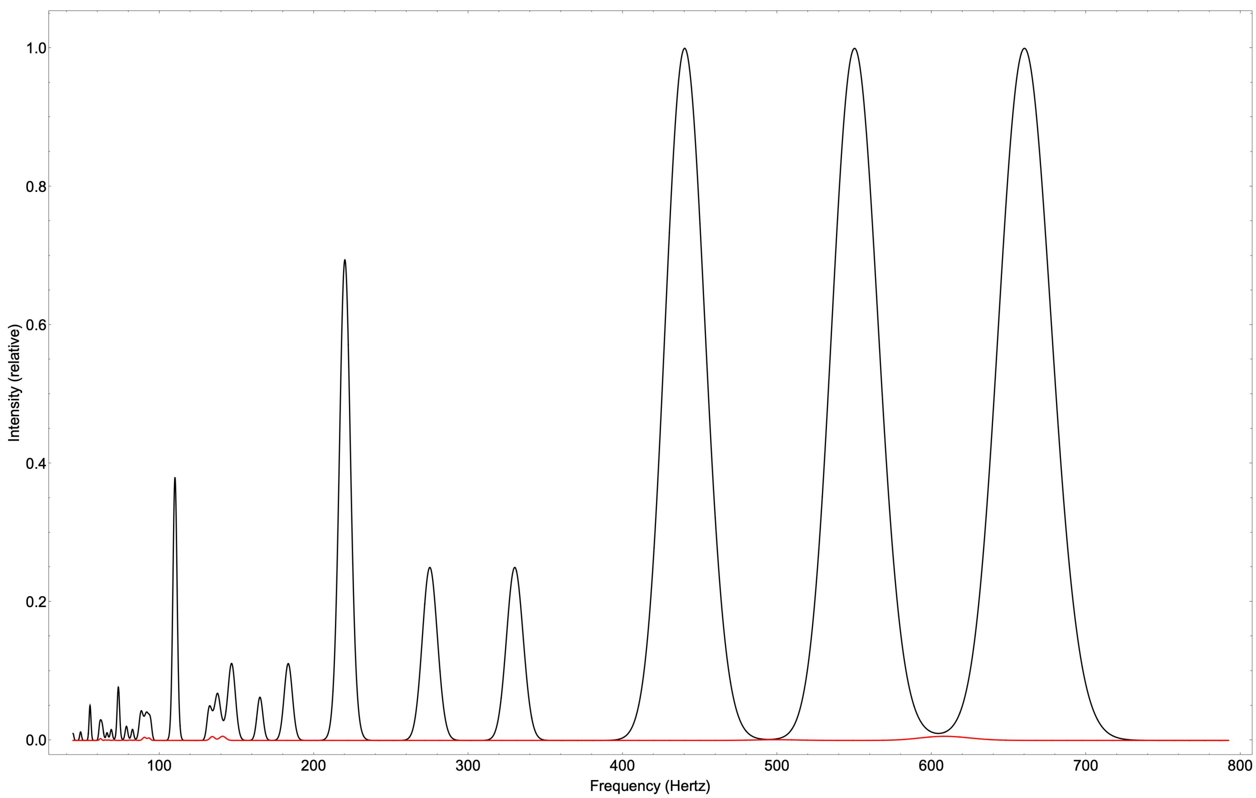

FIGURE 3. SQUARE OF STEADY-STATE TONALNESS (BLACK) AND TONALNESS VARIANCE (RED) FOR THREE PURE TONES: 440 HZ, 550 HZ, AND 660 HZ.

We will see in future examples that no matter how any interval or chord is tuned, even for unisons and octaves, there is always at least some degree of variation in the TV function due to the propensity of any given interval within that chord to be perceived as a mistuning of some other nearby interval. While the research of Plomp and Levelt suggests that no degree of roughness should be audible for pure tone intervals except inside the critical band of spectral pitches[11], the tonalness model suggests that these variations outside the critical band do exist but are quite subtle when only a few pure tones are present in a SPS. However, the tonalness model also suggests that these effects will compound as more pure tones are added to the sound. This model also helps to explain why spectral pitches, when tuned closely to a harmonic series, are perceived as timbres.

We can understand the SST function as a proxy for the harmonicity perception of a SPS, and the TV function as that for roughness perception. In this view, harmonicity and roughness are not a singular measurement of a given harmony but a complex two dimensional function that varies with pitch. It is possible for the HAS to perceive part of a harmony as having a high degree of harmonicity while other parts might have a much lower degree, and the same for roughness.

To compare the tonalness model to other models of harmonic perception, Equation 12 defines a dissonance function as an aggregate intensity of variation within the TV function and scales it against the aggregate intensity of the SST function. Equations 13–15 define the dissonance function for additive sinusoidal SPSs. We can compare this definition of dissonance to other dissonance models.

$$\eta(t) = \frac{\int_{-\infty}^{\infty} \psi_2(f,t)\,df}{\int_{-\infty}^{\infty} \psi_1(f,t)\,df}$$

EQUATION 12. DISSONANCE FUNCTION

[11] Plomp, *Tonal Consonance and Critical Bandwidth*, 553–560.

$$\eta(t) = \frac{2\pi \sum_{g=1}^{\infty} \sum_{h=1}^{\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\gamma_{g,h,i,j} \left| \frac{b_j}{h} - \frac{b_i}{g} \right|}{\phi_{g,h,i,j}}}{\sum_{g=1}^{\infty} \sum_{h=1}^{\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{g,h,i,j}}$$

EQUATION 13. SINUSOIDAL DISSONANCE FUNCTION

$$\phi_{g,h,i,j} = \frac{gb_i + hb_j}{g^2 + h^2}$$

EQUATION 14. HARMONIC AVERAGE FREQUENCY

$$\gamma_{g,h,i,j} = \frac{a_i a_j \sigma\left(\phi_{g,h,i,j}\right)\sqrt{2\pi}}{gh\sqrt{g^2 + h^2}} \exp\left(- \frac{\left(gb_j - hb_i\right)^2}{2\left(g^2 + h^2\right)\sigma\left(\phi_{g,h,i,j}\right)^2} \right)$$

EQUATION 15. ROUGHNESS FACTOR

Using Equations 13–15, we can calculate a dissonance value of 0.00423981 for the equal tempered version of the triad and 0.0027082 for the justly tuned version. As we should have expected, the justly tuned version carries less dissonance than the equal tempered version.

Before diving into the many musical predictions of Equation 12, we will introduce one last tool for the purposes of musical analysis. Equation 16 defines the tonalness progression function (TPF). The TPF calculates the degree to which the square of the SST curve shifts for a SPS between two points in time.

$$\Omega\left(t_1, t_2\right) = \sqrt{\frac{\int_{-\infty}^{\infty}\left(\psi_1\left(f, t_1\right) - \psi_1\left(f, t_2\right)\right)^2 \, df}{\int_{-\infty}^{\infty}\left(\psi_1\left(f, t_1\right) + \psi_1\left(f, t_2\right)\right)^2 \, df}}$$

EQUATION 16. TONALNESS PROGRESSION FUNCTION

Using Equation 16, we calculate the TPF value between our equal tempered and justly tuned triads to be 0.0565812. We will soon see that this is an extremely small TPF value.

Dissonance

The dissonance function is perhaps the most obvious use for tonalness theory in the analysis of musical harmony. We will first consider the dissonance of dyads. For the convenience of musical and mathematical analysis, we will express pitch as MIDI numbers where MIDI 69 corresponds to A_4 tuned to 440 Hz and a MIDI interval (MI) of 1 corresponds to 1 semitone or 100 cents. The precise relationship between a given MIDI pitch s and frequency b is defined by Equation 17 where we will set s_0 equal to MIDI 69 and b_0 equal to 440 Hz.

$$b(s) = b_0 2^{\frac{1}{12}(s - s_0)}$$

EQUATION 17. MIDI PITCH TO FREQUENCY CONVERSION

Figure 4 uses Equations 13–15 to plot dissonance values of various dyads consisting of pure tone pitches where s_1 is the lower pitch and s_2 is the higher pitch.

The black line shows the dissonance values for dyads spanning the unison (MI 0) through three octaves (MI 36) where s_1 is held to MIDI 48. The red line shows the same for dyads where s_1 is held to MIDI 60, and the blue line shows the same for dyads where s_1 is held to MIDI 72.

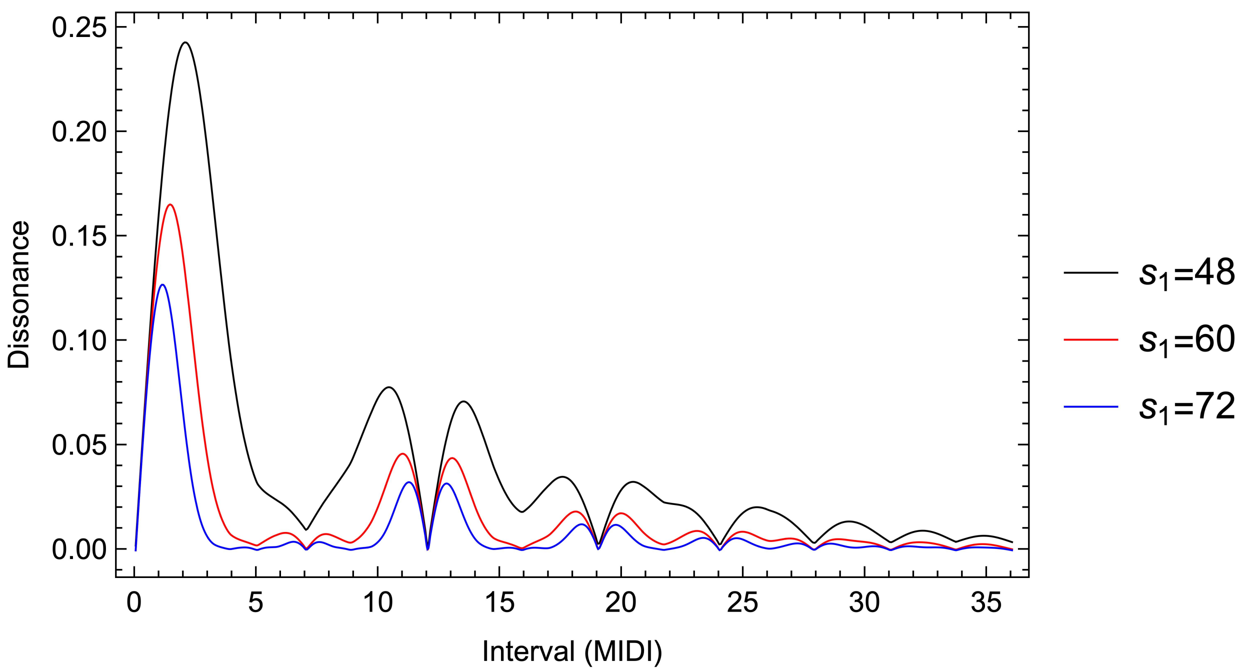

FIGURE 4. DISSONANCE PLOT FOR PURE TONE DYADS IN VARIOUS REGISTERS.

As we see from Figure 4, Equations 13–15 predict that the dissonance of a given harmonic interval is largely dependent on its register. We see that the features of this dissonance curve become increasingly detailed as intervals increase in register, but that overall dissonance levels rise as intervals decrease in register. We observe the prediction that two pitches existing within the critical band should be perceived as maximally dissonant, and this critical band dissonance produces the highest degrees of dissonance that exist on this graph regardless of register. We also notice that the relative maxima of each of the dissonance curves shifts a bit as an interval descends in register. These predictions for dissonance at the unison are largely in agreement with experimental evidence from Plomp and Levelt[12].

We may also choose to explore how timbre affects the dissonance of harmonic dyads. For this we will consider three timbres: a pure tone, a harmonic timbre consisting of harmonics 1, 2, 3, 4, and 5 with relative amplitudes of 1, 1/2, 1/3, 1/4, and 1/5 respectively, and an inharmonic timbre consisting of harmonics 1, 2.1, 3.2412, 4.41, and 5.59977 with relative amplitudes of 1, 1/2, 1/3, 1/4, and 1/5 respectively. Figure 5 uses Equations 13–15 to plot dissonance values for dyads spanning a unison (MI 0) through three octaves (MI 36) above MIDI 60 consisting of pure tones (blue), the harmonic timbre (red) as described above, and an inharmonic timbre (black) also as described above. These curves take into account dissonance due to the interactions between the spectral and virtual components within each tone and between each tone.

[12] Plomp, *Tonal Consonance and Critical Bandwidth*, 548–560.

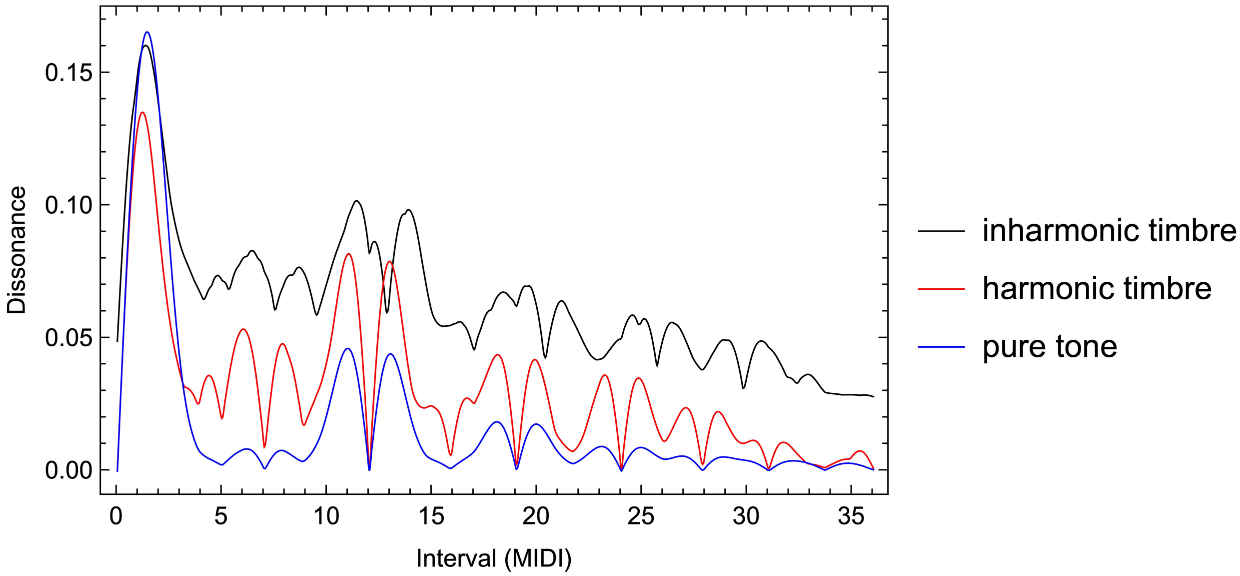

FIGURE 5. DISSONANCE PLOT FOR DYADS ABOVE MIDI 60 CONSISTING OF PURE TONES (BLACK), A HARMONIC TIMBRE (RED), AND AN INHARMONIC TIMBRE (BLUE).

As we see from Figure 5, the predicted dissonance from Equations 13–15 of a harmonic timbre dyad closely mirrors its pure tone counterpart. However, the features of the dissonance curve from a harmonic timbre are exaggerated compared to that of the pure tone. In contrast, the dissonance curve that results from the inharmonic timbre dyads is somewhat shifted and more uniform as compared to that of the harmonic timbre or pure tone dyads. Because the overtones of the inharmonic timbre were shifted to create a 2.1:1 pseudo-octave, we see a pronounced valley in the curve slightly above MI 12. Equations 13–15 predict that while we may perceive less relative dissonance for a stretched octave from an inharmonic timbre, we will never perceive anything near the very low dissonance values we can achieve for the harmonic timbre or pure tone 2:1 octave. This prediction agrees with the experimental evidence of McDermott, Lehr, and Oxenham[13].

[13] McDermott, *Individual Differences Reveal the Basis of Consonance*, 1037.

To further investigate the relationship between the dissonance and register for various intervals, Figure 6 plots the dissonance of pure tone MI 0–12 with lower pitches spanning MIDI 24–72. Figure 7 does the same using a harmonic timbre consisting of harmonics 1, 2, 3, 4, and 5 with relative amplitudes of 1, 1/2, 1/3, 1/4, and 1/5 respectively.

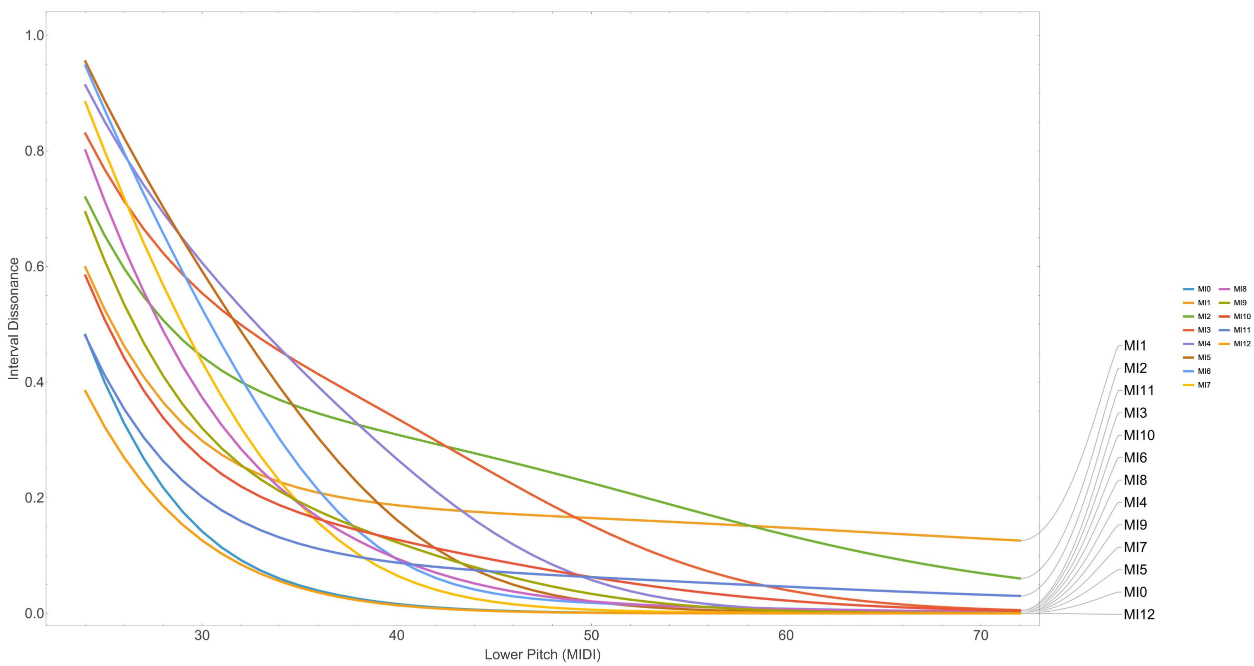

FIGURE 6. DISSONANCE PLOT FOR DYADS WITH LOWER PITCHES SPANNING FROM MIDI 24 TO MIDI 72 CONSISTING OF PURE TONES.

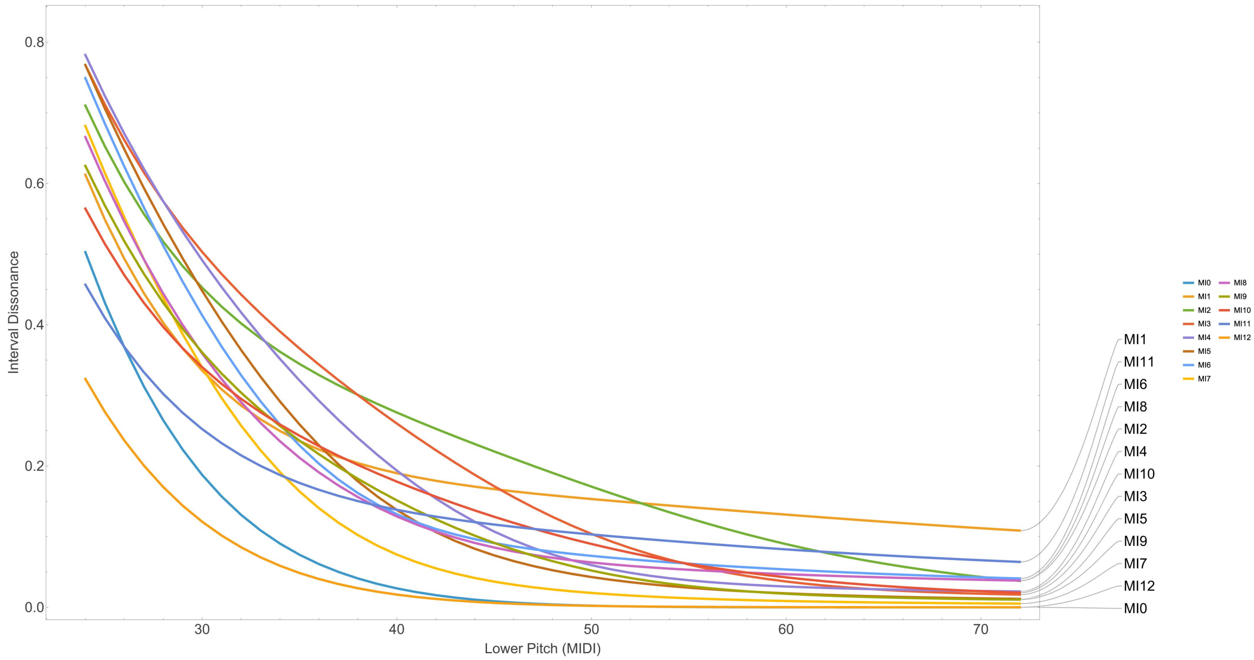

FIGURE 7. DISSONANCE PLOT FOR DYADS WITH LOWER PITCHES SPANNING FROM MIDI 24 TO MIDI 72 CONSISTING OF A HARMONIC TIMBRE.

Figures 6 and 7 show that the dissonance of a given harmonic interval varies inversely with register, although this does not occur uniformly. The relative dissonance between these intervals also varies with register and is affected significantly by timbre.

When we consider trichords, we see results that mirror those we found for our results for intervals. Figures 8–12 show continuous dissonance plots for trichords composed of intervals spanning the unison through octave.

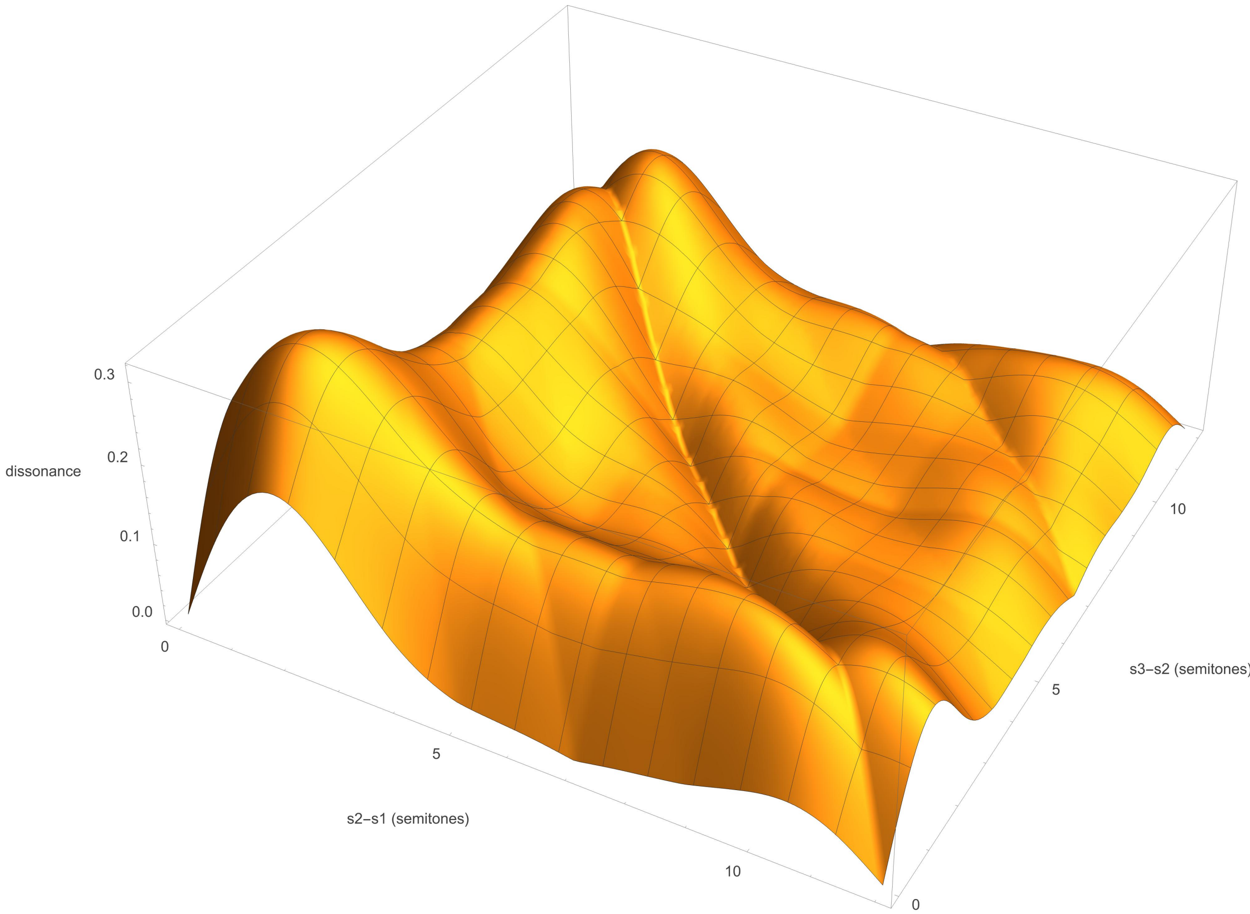

FIGURE 8. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 48. THESE TRICHORDS ARE COMPOSED OF PURE TONES.

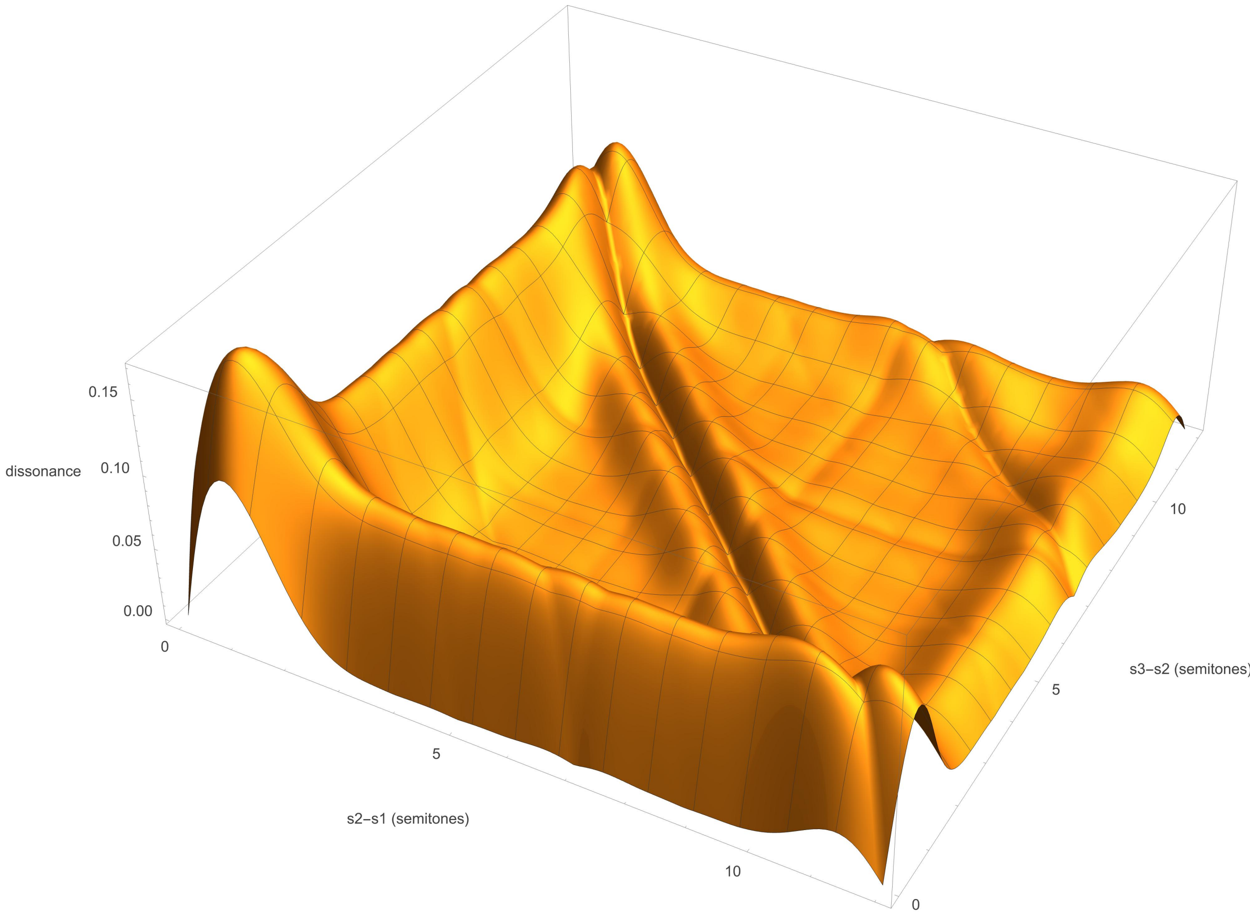

FIGURE 9. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 72. THESE TRICHORDS ARE COMPOSED OF PURE TONES.

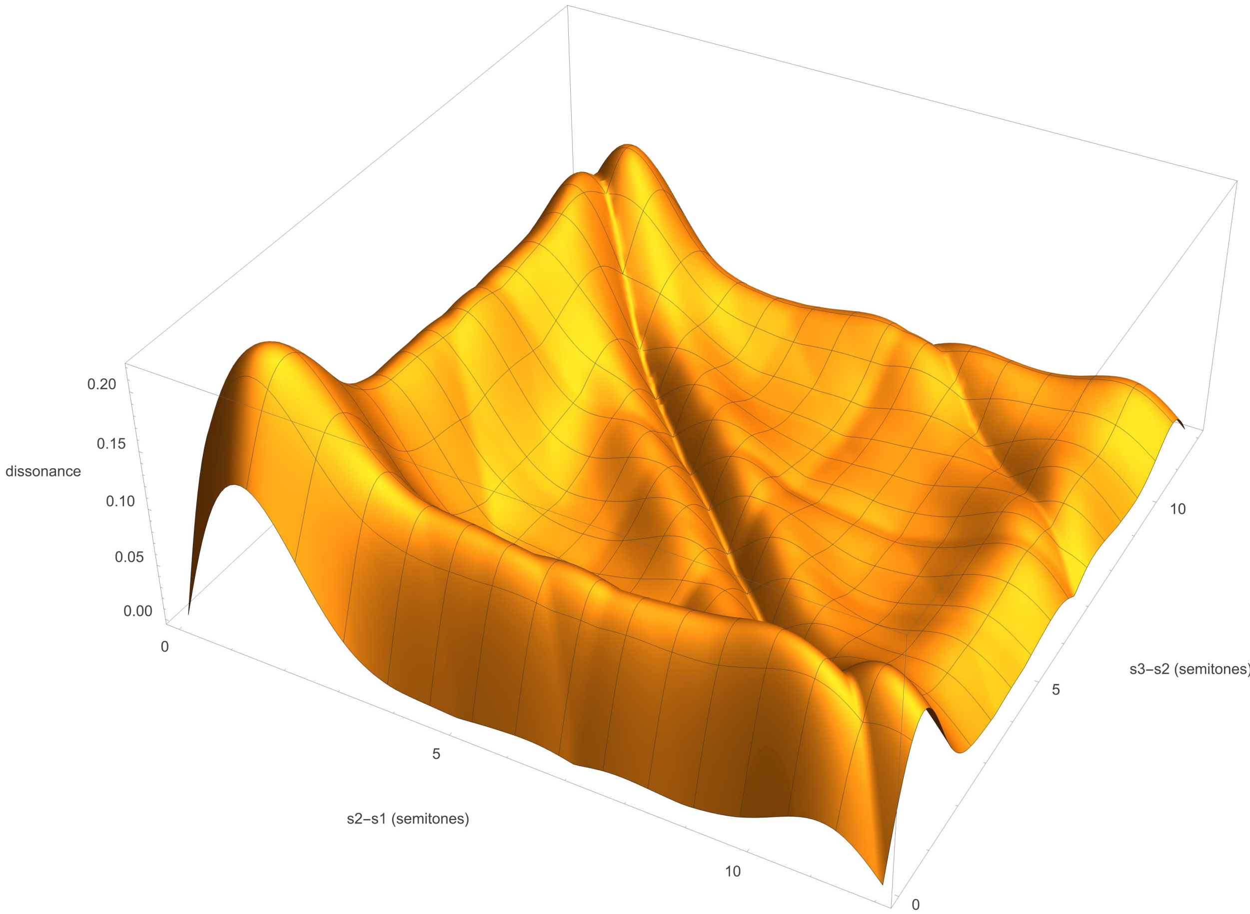

FIGURE 10. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 60. THESE TRICHORDS ARE COMPOSED OF PURE TONES.

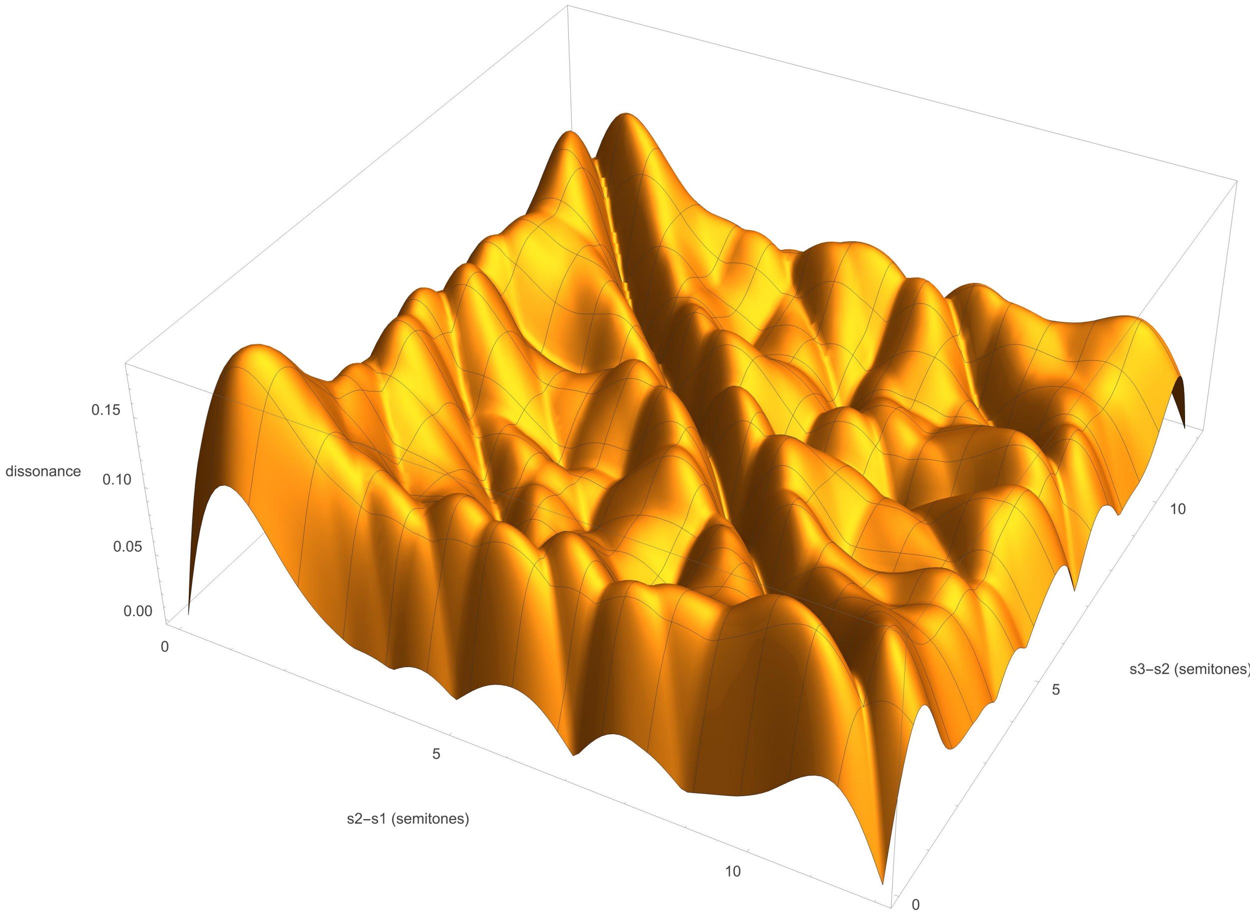

FIGURE 11. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 60. THESE TRICHORDS ARE COMPOSED OF A HARMONIC TIMBRE WITH PARTIALS 1, 2, 3, 4 AND 5 WITH RESPECTIVE RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

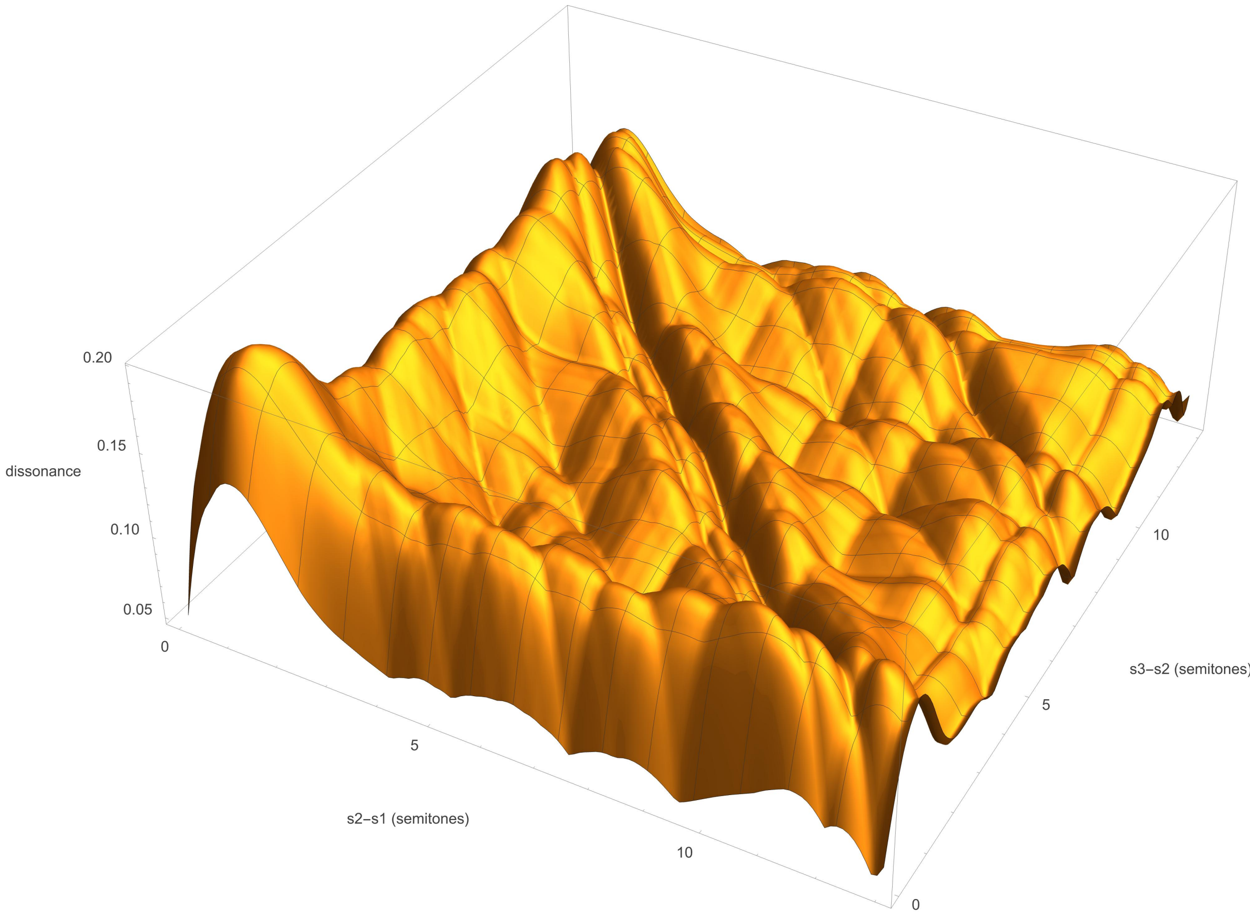

FIGURE 12. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 60. THESE TRICHORDS ARE COMPOSED OF AN INHARMONIC TIMBRE WITH PARTIALS 1, 2.1, 3.2412, 4.41, AND 5.59977 WITH RESPECTIVE RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

Figures 8–12 show that the largest dissonance peaks occur when both the upper and lower intervals within the trichord are within the critical band. The dissonance values are also larger within a critical band of various harmonic intervals, although for the inharmonic timbre we see a similar shift in dissonance values that we did for intervals composed of an inharmonic timbre. We also notice that the dissonance patterns are more pronounced for the trichords built above MIDI 72 and that more precise dissonance valleys exist for the harmonic timbre than for pure tones.

One of the other benefits of this dissonance model is its ability to compare the dissonance of chords with different numbers and intensities of pitches. Figures 13–15 show the dissonance values for a variety of two, three, and four pitch chords composed of both pure tones and a harmonic timbre composed of partials 1, 2, 3, 4, and 5 at respective relative amplitudes of 1, 1/2, 1/3, 1/4, and 1/5. Each of these chords is built above MIDI 60.

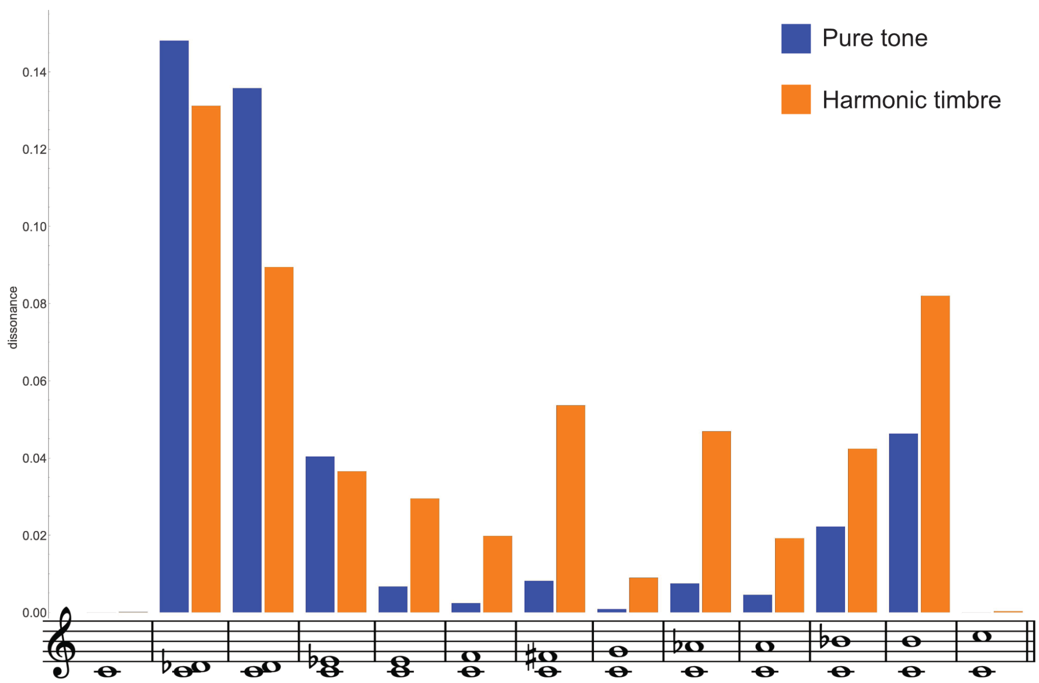

FIGURE 13. DISSONANCE VALUES OF VARIOUS TWO PITCH CHORDS COMPOSED OF PURE TONES (ORANGE) AND A HARMONIC TIMBRE (BLUE) CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

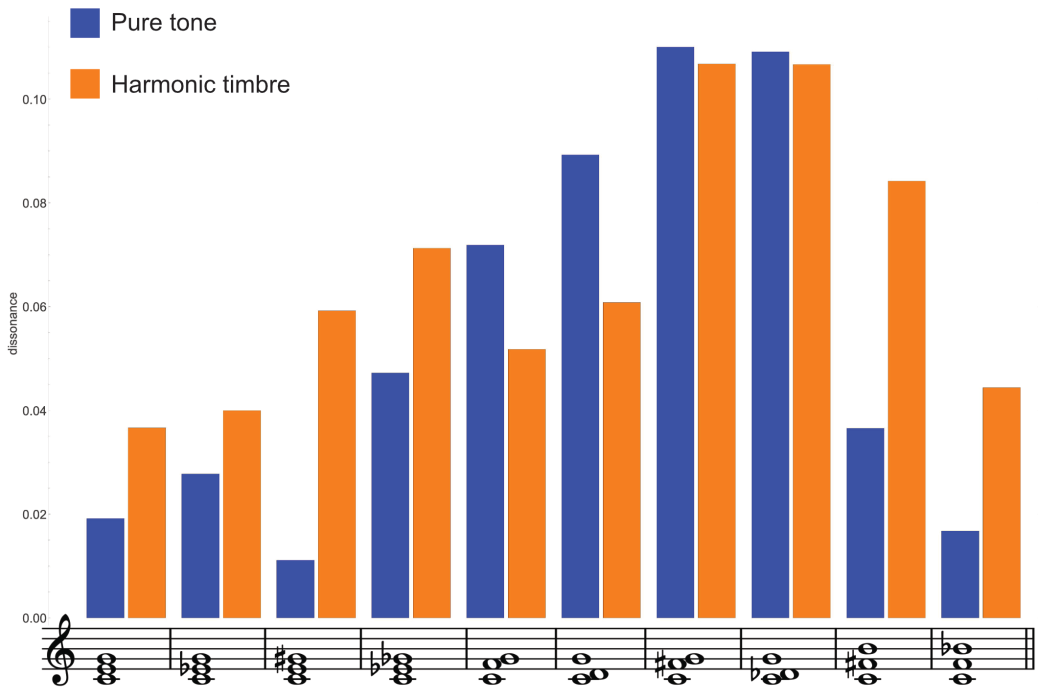

FIGURE 14. DISSONANCE VALUES OF VARIOUS THREE PITCH CHORDS COMPOSED OF PURE TONES (ORANGE) AND A HARMONIC TIMBRE (BLUE) CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

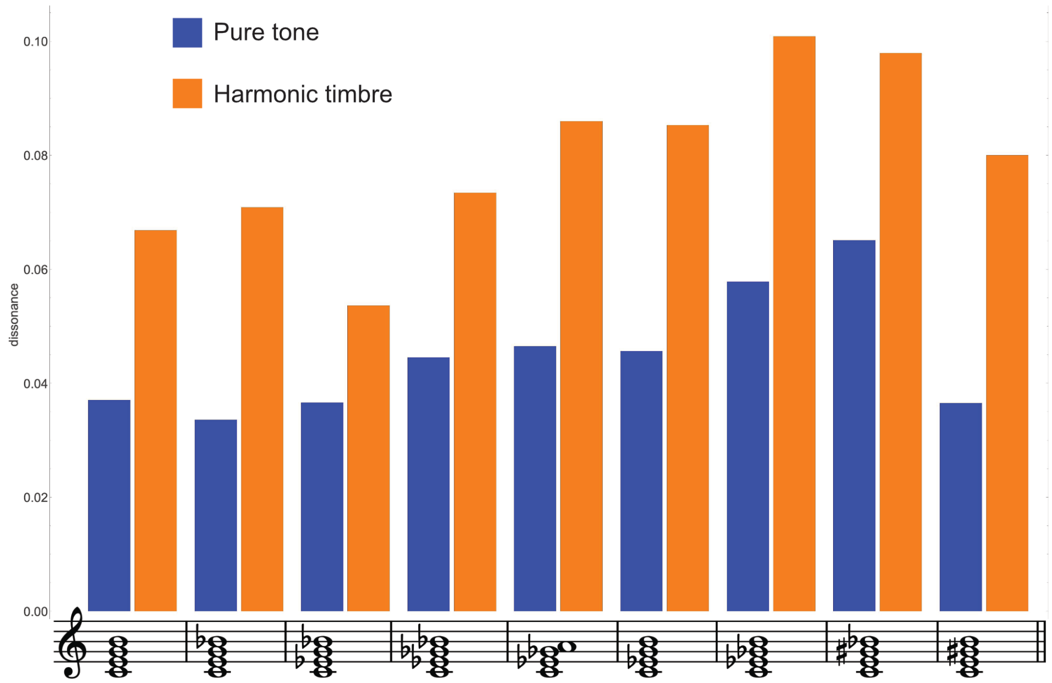

FIGURE 15. DISSONANCE VALUES OF VARIOUS FOUR PITCH CHORDS COMPOSED OF PURE TONES (ORANGE) AND A HARMONIC TIMBRE (BLUE) CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

As we see from Figures 13–15, the order of dissonance roughly follows an order that is familiar to many practitioners of Western tonal music. However, there are some surprising results on these charts. However, we should not over-interpret these results. By only considering the intensity of beating between spectral and virtual pitches, Equations 16–18 calculate a very specific type of dissonance perception. It cannot adequately serve as a replacement for a more comprehensive understanding of dissonance such as that which exists within the context of Western tonality.

Tonalness Progression

The dissonance of individual chords is only one of many factors that contributes to the musical context of harmony. We may also wish to study the perception of chord progressions. Equation 16 (TPF) measures the aggregate change in SST between two chords. The TPF accounts for both the change in spectral and virtual pitches and serves as a proxy for how smoothly or abruptly one chord progresses into another. Figure 16 uses Equations 13–16 to calculate the TPF values and changes in dissonance for a variety of common harmonic progressions composed respectively of pure tones and a harmonic timbre defined by partials 1, 2, 3, 4, and 5 with relative amplitudes of 1, 1/2, 1/3, 1/4, and 1/5.

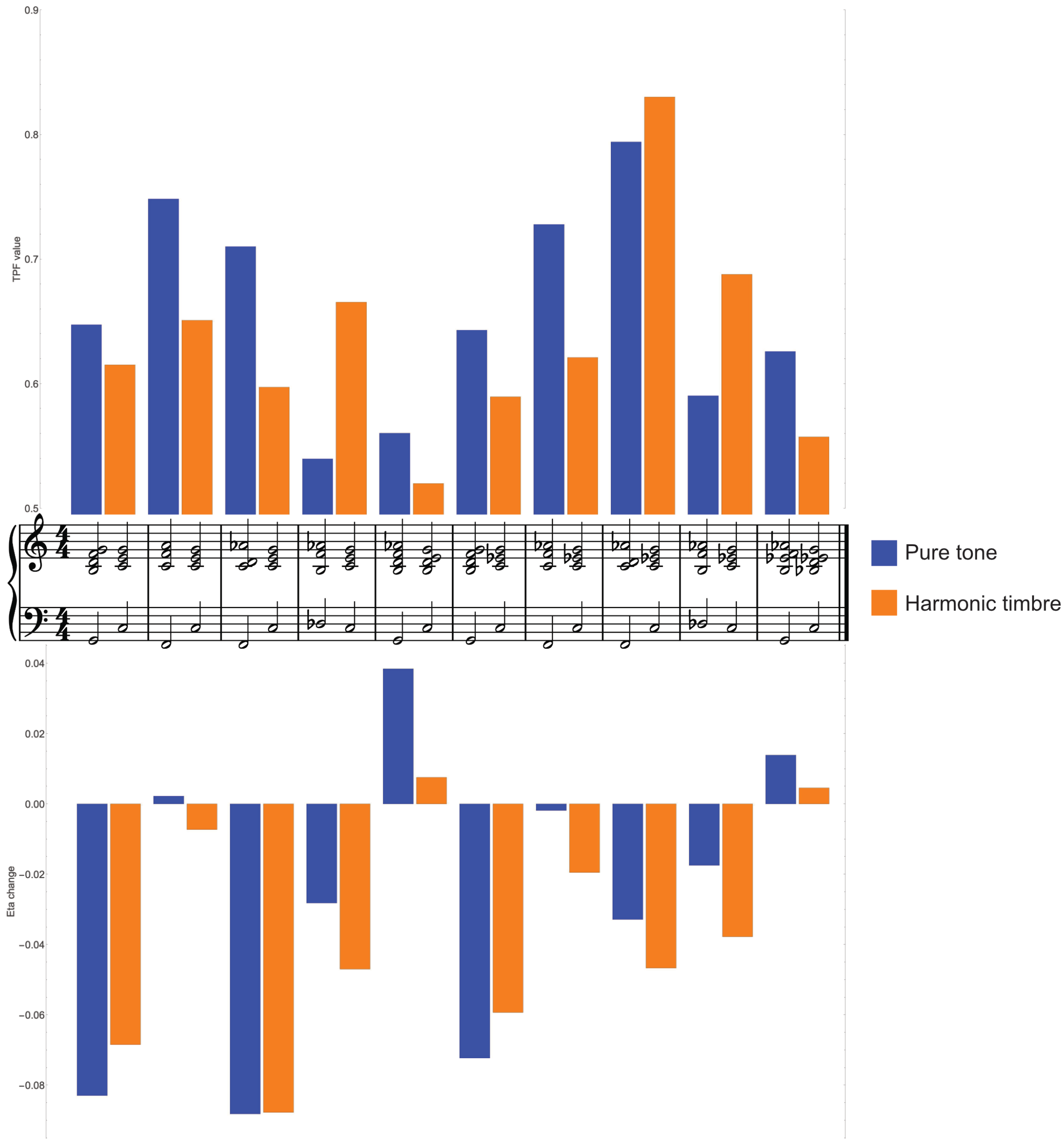

FIGURE 16. TPF VALUES AND DISSONANCE CHANGE FOR VARIOUS COMMON CHORD PROGRESSIONS IN THE KEY OF C CONSISTING OF PURE TONE PITCHES AND PITCHES WITH A HARMONIC TIMBRE CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

As we see from Figure 16, a lot of the trends that exist for pure tone chords are mirrored in the results for chords with harmonic timbres. We should again exercise caution about concluding too much from these findings. While the TPF and dissonance function can help us understand many important psychoacoustic properties of harmonic progression, to most fully model why harmonies sound the way that they do requires a return to the SST and TV functions.

Virtual Pitch Analysis

The virtual pitch theories of Ernst Terhardt[14] are in some ways the modern successor to Rameau's theories of the fundamental bass. Our SST and TV equations serve as a comprehensive model of virtual pitch. According to our formulation, there is not one fundamental bass for a given chord but rather a continuous spectrum of perceived virtual pitches. Virtual pitch analysis (VPA) studies the complex patterns of and interference between virtual pitches in an attempt to most fully understand the psychoacoustic perception of a particular harmony. Figures 17 and 18 respectively compare the square of the SST and TV for MIDI {60, 64, 67} major triad to the MIDI {60, 63, 67} minor triad, with each consisting of pure tone component pitches of equal loudness.

[14] Terhardt, *Calculating Virtual Pitch*, 155–182.

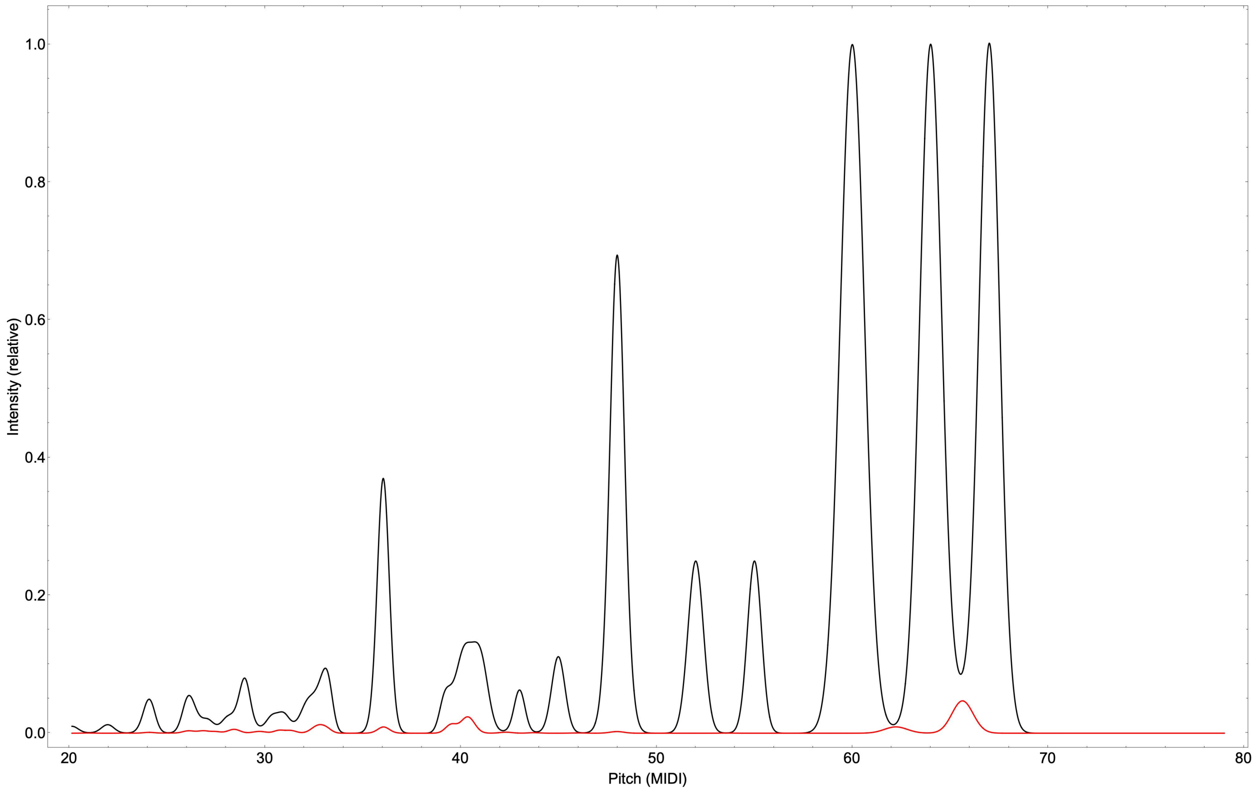

FIGURE 17. SQUARE OF SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 64, 67 MAJOR TRIAD CONSISTING OF PURE TONE PITCHES.

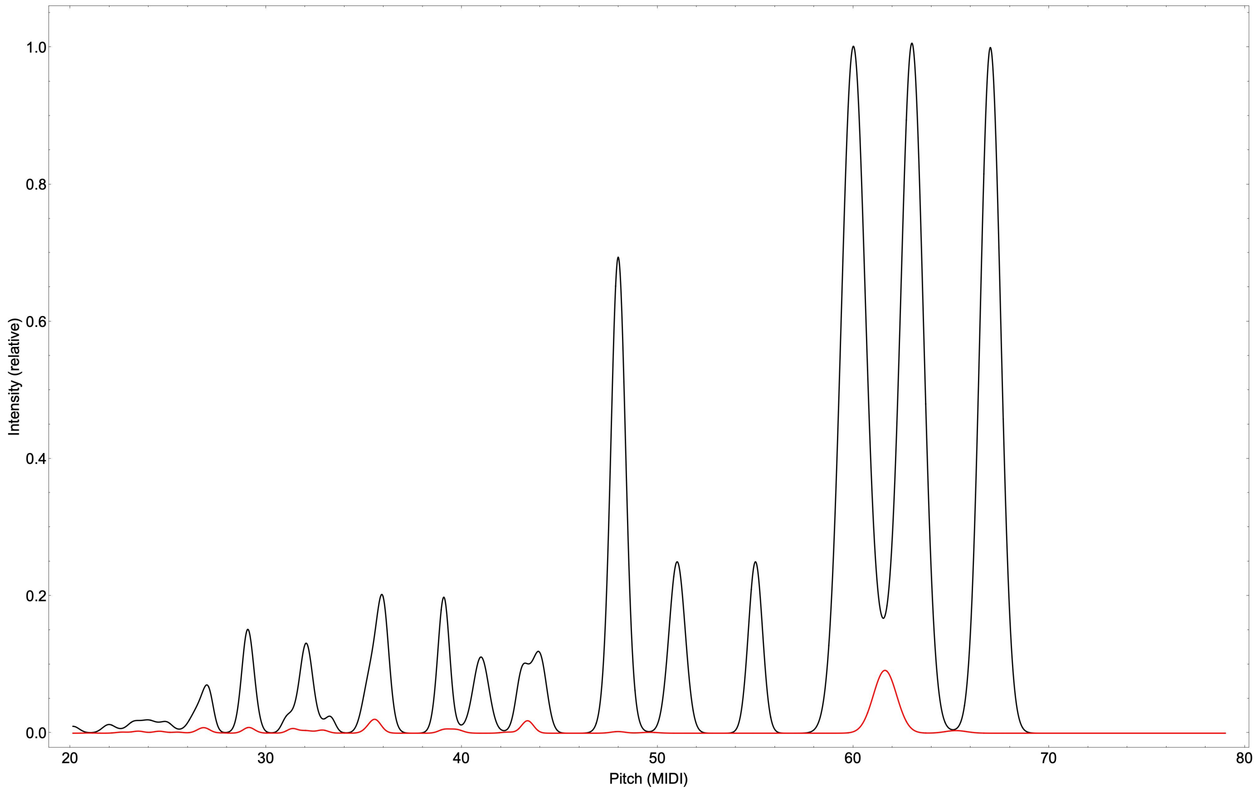

FIGURE 18. SQUARE OF SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 63, 67 MINOR TRIAD CONSISTING OF PURE TONE PITCHES.

While we might notice the slightly higher TV curve and thus greater dissonance of the minor triad, the major triad has a somewhat more prominent virtual pitch near MIDI 36. This is due to the spectral pitches in the major triad corresponding approximately to the 4th, 5th, and 6th harmonics of MIDI 36. It is the largest virtual pitch that results from all three spectral pitches. For the minor triad, there are weaker virtual pitches near MIDI 32 and 39. Its spectral pitches do not align well with a single harmonic series but rather align a bit less well with two different harmonic series—one based above MIDI 32 and the other based above MIDI 39. While both of these triads contain the same interval content and a similar dissonance value, we should expect the major triad to have a somewhat more focused and unified sound when compared to the minor triad.

We might predict a similar result when comparing the MIDI {60, 64, 67, 70} major-minor seventh chord to the MIDI {60, 63, 66, 70} half-diminished seventh chord. Figures 19 and 20 plot the square of SST and TV for these chords with pure tone component pitches.

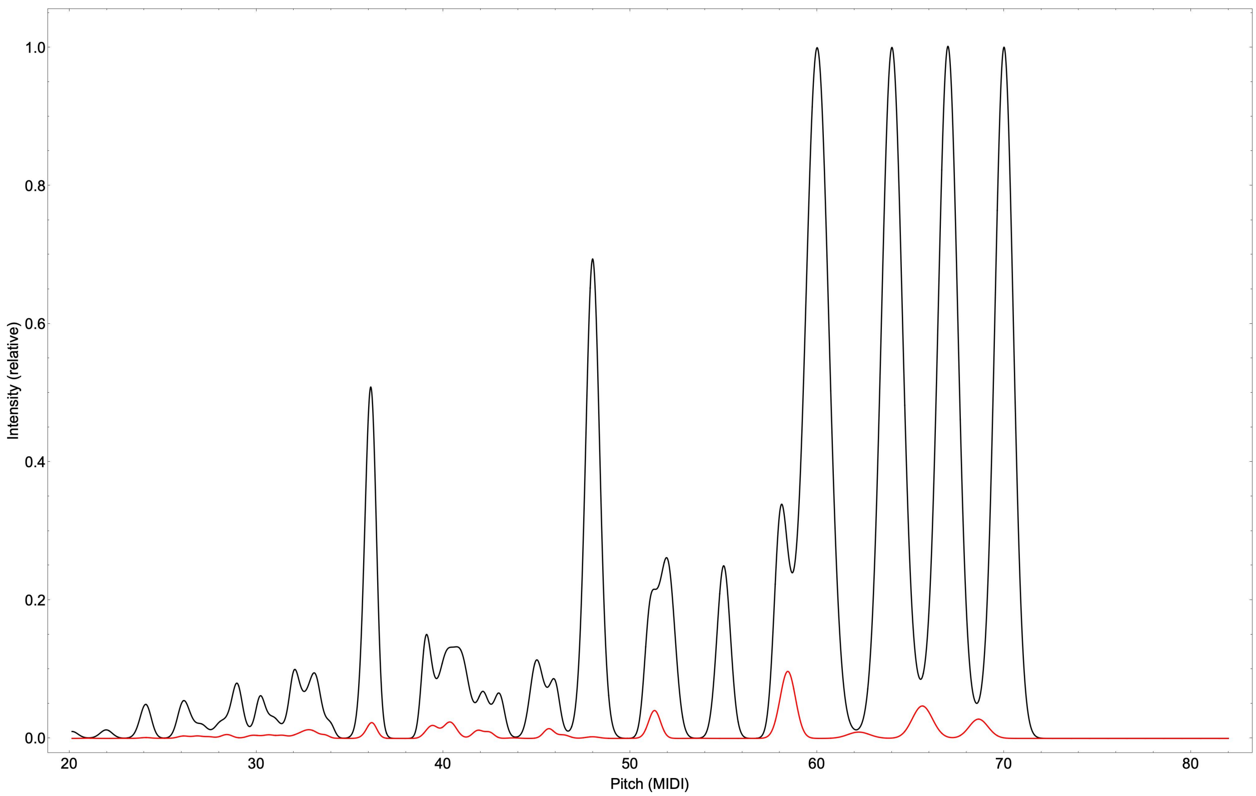

FIGURE 19. SQUARE OF SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 64, 67, 70 MAJOR-MINOR SEVENTH CHORD CONSISTING OF PURE TONE PITCHES.

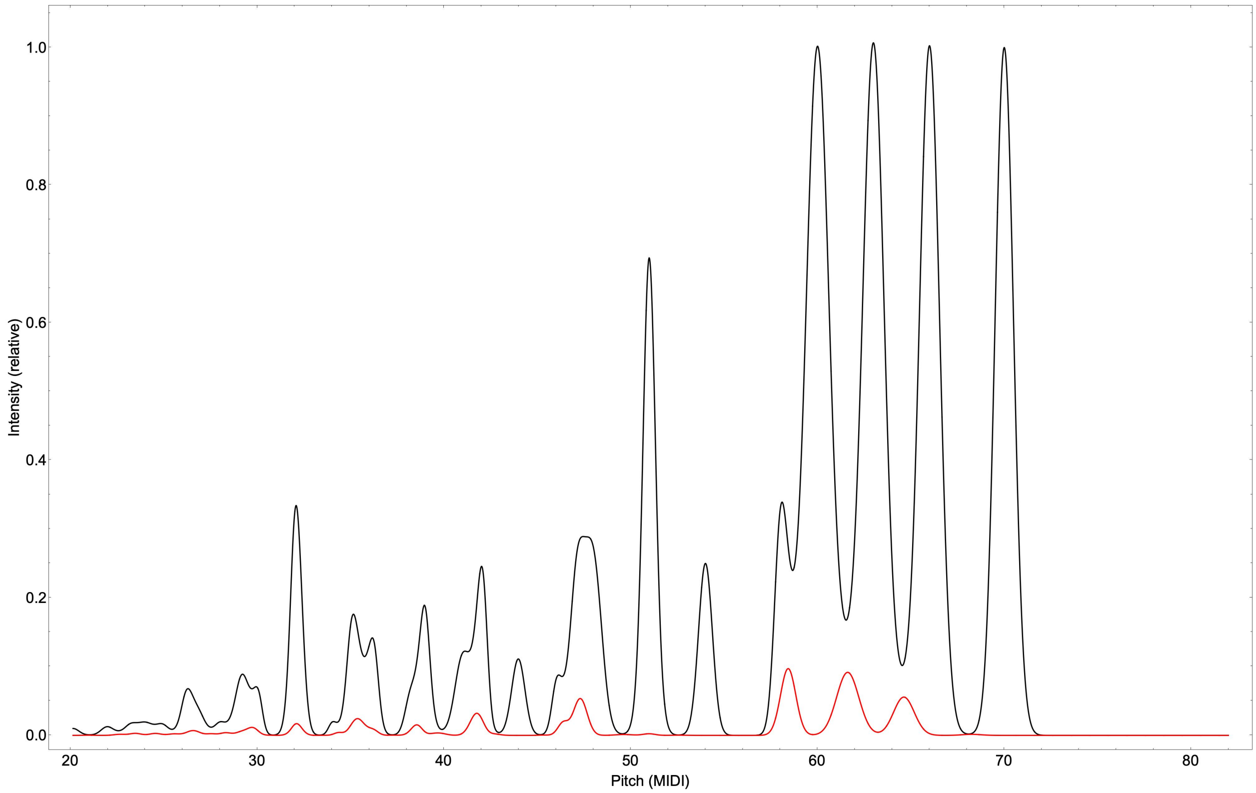

FIGURE 20. SQUARE OF SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 63, 66, 70 HALF-DIMINISHED SEVENTH CHORD CONSISTING OF PURE TONE PITCHES.

As was true for the major and minor triads, the major-minor and half-diminished seventh chords contain the same interval content. However, the arrangement of intervals makes a significant difference for the virtual pitch fingerprint of these chords. The major-minor seventh chord has a more prominent virtual pitch near MIDI 36 due to the spectral pitches in the major-minor seventh chord corresponding approximately to the 4th, 5th, 6th, and 7th harmonics of MIDI 36. It is the largest virtual pitch that results from all four spectral pitches for a major-minor seventh chord. Alternatively, the spectral pitches of the half-diminished seventh chord create more scattered virtual pitches near MIDI 32, 35, and 42. Similar to the major and minor triads, we should expect the major-minor seventh chord to have a somewhat more focused and unified sound when compared to its half-diminished counterpart.

For a more extreme example, Figures 21 and 22 compare a MIDI {60, 72, 79.0196, 84, 87.8631} chord built on the first five harmonics of an overtone series to a MIDI {60, 63.8631, 68.8435, 75.8631, 87.8631} chord built on the first five harmonics of an undertone series—each consisting of pure tone component pitches.

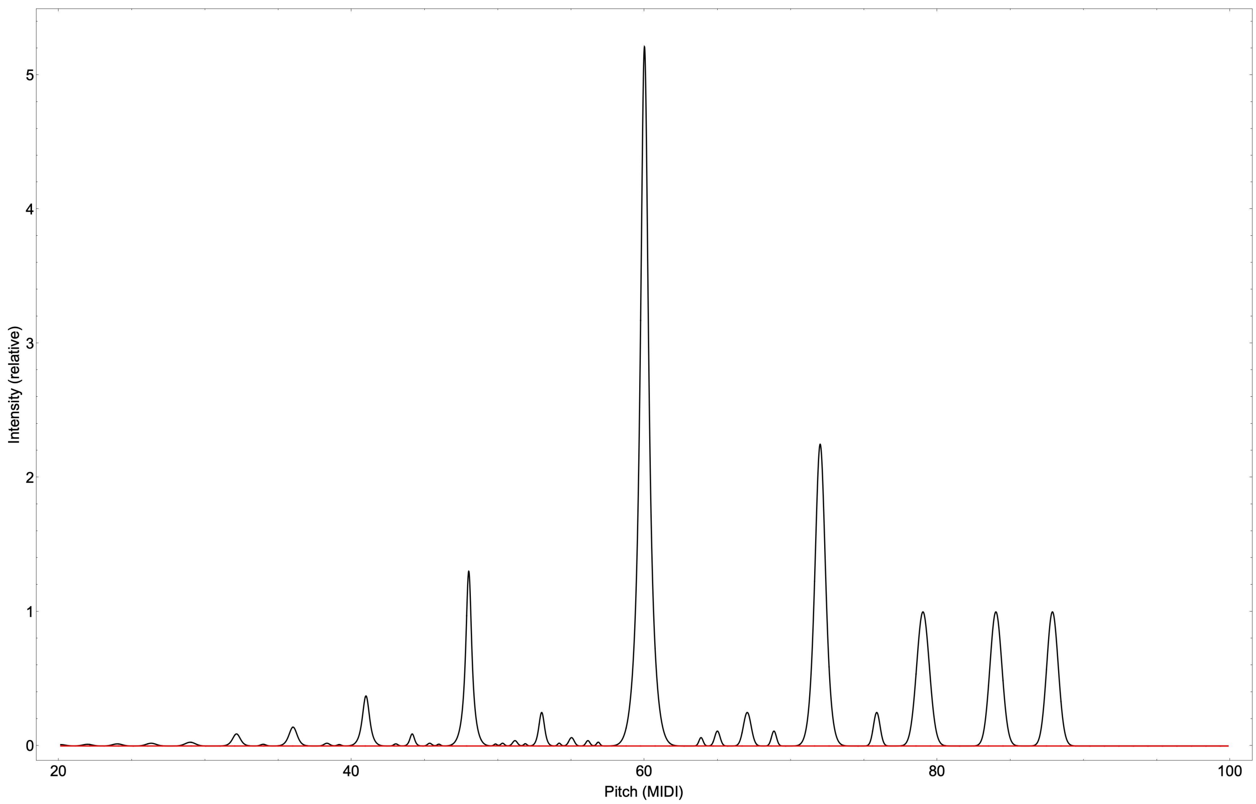

FIGURE 21. SQUARE OF SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 72, 79.0196, 84, 87.8631 CHORD CONSISTING OF PURE TONE PITCHES.

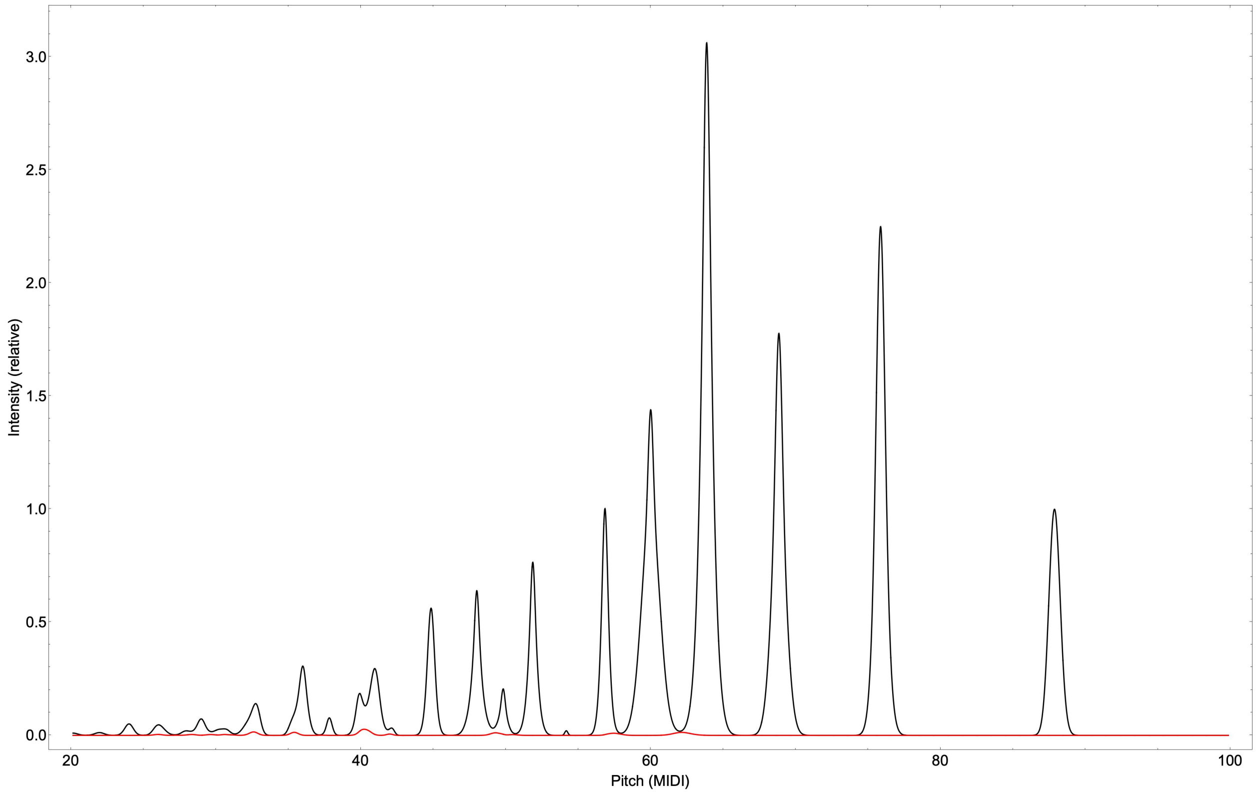

FIGURE 22. SQUARE OF SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 63.8631, 68.8435, 75.8631, 87.8631 CHORD CONSISTING OF PURE TONE PITCHES.

As we see from Figures 21 and 22, the first chord produces a very prominent peak at MIDI 60 while the second chord seems to have several less prominent peaks. The second chord should also contain a somewhat higher degree of beating and dissonance due to the peaks in the TV curve, while its counterpart seems to lack this almost entirely. As a result, we should expect the first chord to sound extremely focused and unified and the second chord to be significantly more diffuse, unstable, and less unified harmonically despite both chords containing identical interval content.

Conclusions

Many features of the perception of musical harmony can be understood via modeling virtual pitches. We can understand dissonance to be an aggregate property of the virtual pitch content of a given harmony. However, to fully grasp why a particular harmony sounds the way that it does requires us to study its unique pattern of virtual pitches. While timbre can significantly impact the perception of consonance and dissonance, harmonic timbres generally yield similar patterns of consonance and dissonance as their pure tone counterparts. Furthermore, inharmonic timbres can distort these patterns of consonance and dissonance to a point. Inharmonic timbres also tend to yield higher overall levels of dissonance than their harmonic timbre counterparts.

Tonalness theory concerns itself mainly with the immediate apprehension of harmony and harmonic progressions, not the synoptic comprehension of musical works. As a consequence, tonalness theory is not nor could ever be an all-encompassing theory of music. However, building rigorous mathematical models of harmonic perception forces us to be consistent rather than arbitrary in our application of various principles and assumptions. Tonalness theory models can help us distinguish between harmonic perceptions which are rooted mainly in a musical culture and those which are an intrinsic part of the human endowment.

Bibliography

Erlich, Paul. “On Harmonic Entropy” Accessed May 9, 2022. http://www.tonalsoft.com/enc/e/erlich/harmonic-entropy_original.aspx.

Fletcher, Harvey, and W. A. Munson. “Loudness, Its Definition, Measurement and Calculation.” Journal of the Acoustical Society of America. 5 (1933): 82–108.

Glasberg, Brian R., and Moore, Brian, C. J. “Derivation of auditory filter shapes from notched-noise data.” Hearing Research. 47, no. 1–2 (1990): 103–138.

Helmholtz, Hermann. “On the Sensations of Tone as a Physiological Basis for the Theory of Music.” Translated by Alexander J. Ellis. New York: Dover, 1954.

McDermott, Josh H., Andriana J. Lehr, and Andrew J. Oxenham. “Individual Differences Reveal the Basis of Consonance.” Current Biology. 20, no. 11 (2010): 1035–1047.

Moore, B. C. J. “Cochlear hearing loss.” London: Whurr Publishers Ltd, 1998.

Peretz, Isabelle, Anne J. Blood, Virginia Pentane, and Robert Zatorre. “Cortical deafness to dissonance.” Brain. 124 (2001): 928–940.

Plomp, Reinier, and Willem Levelt. “Tonal Consonance and Critical Bandwidth.” Journal of the Acoustical Society of America. 38, (1965): 548–560.

Rameau, Jean-Philippe. “Treatise on Harmony reduced to its natural principles.” Translated by Philip Gossett. New York: Dover, 1971.

Schellenberg, E. Glenn, and Sandra E. Trehub. “Frequency ratios and the discrimination of pure tone sequences.” Perception & Psychophysics. 56, no. 4 (1994): 472–478.

Sethares, William A. Tuning, Timbre, Spectrum, Scale. London: Springer-Verlag, 2005.

Terhardt, Ernst. “Calculating Virtual Pitch.” Hearing Research. 1, no. 2 (1979): 155–182.

Appendix

Standing wave model

We can model a standing wave on a string of length L fixed at both ends using the one-dimensional wave equation and our boundary conditions, as shown in Equations 1A and 2A.

$$\frac{\partial^2 u(x,t)}{\partial t^2} = c^2 \frac{\partial^2 u(x,t)}{\partial x^2}$$

EQUATION 1A. ONE-DIMENSIONAL WAVE EQUATION.

$$u(0,t) = u(L,t) = 0$$

EQUATION 2A. BOUNDARY CONDITIONS.

From this, Equation 3A shows the standard solution to the above equations if we assume that A is an arbitrary amplitude, h and arbitrary harmonic integer, and an initial displacement of zero.

$$u(x,t) = A \sin\left(\frac{h\pi x}{L}\right) \sin\left(\frac{hc\pi t}{L}\right)$$

EQUATION 3A. ONE-DIMENSIONAL STANDING WAVE SOLUTION.

Equation 4A defines the kinetic energy for this standing wave, where ρ is the linear mass density of the string. We observe that the kinetic energy equals the total energy for the standing wave at $t = 0$. From this assumption, Equation 5A defines the total energy for this standing wave.

$$\text{KE} = \frac{1}{2}\rho \int_0^L \left(\frac{\partial u(x,t)}{\partial t} \right)^2 dx$$

EQUATION 4A. STANDING WAVE KINETIC ENERGY.

$$\text{TE} = \frac{A^2 c^2 h^2 \pi^2 \rho}{4L}$$

EQUATION 4A. STANDING WAVE TOTAL ENERGY.

When we solve for amplitude in Equation 5A, we see that it is inversely proportionate to the harmonic number h. In other words, for equal energy across harmonics amplitudes scale as $\frac{1}{h}$. We adopt $\frac{1}{h}$ to model how the HAS weights virtual pitches.

$$A = \frac{2}{ch\pi}\sqrt{\frac{\text{TE}L}{\rho}}$$

EQUATION 5A. AMPLITUDE-HARMONIC PROPORTIONALITY.