Tonalness Theory: Mathematical Foundations for a Signal-Domain Model of Musical Harmony

Gregory J. Strohman

Lecturer in Music, Dickinson College

July 28, 2026

Abstract

Tonalness names the propensity of the auditory system to project virtual pitches onto a distribution of spectral energy. This paper formalizes a psychoacoustic transform that folds critical band, frequency response, and virtual-pitch perception into Fourier analysis, yielding a single time-varying function of frequency from which three quantities are read: a continuous landscape of candidate fundamentals, generalizing the notion of a chord root; a normalized measure of dissonance; and a measure of the distance between successive sonorities. Unlike existing computational accounts of consonance, which take a list of partials as input and return a static value, the transform is defined for any sound pressure signal—discrete or continuous, harmonic or inharmonic, pitched or noise—and varies in time by construction. Beating, tonal fusion, timbre-dependent dissonance minima, and the register dependence of interval dissonance arise as consequences of the formalism rather than as separately modeled effects. Predictions distinguishing the model from existing roughness accounts are stated.

Tonalness Theory: Mathematical Foundations for a Signal-Domain Model of Musical Harmony

Introduction

When Jean-Philippe Rameau first published *Traité de l'harmonie réduite à ses principes naturels* in 1722[1], he posed questions about musical harmony in the context of music with a far more limited harmonic vocabulary than what emerged three centuries later. A contemporary theory of harmony should ideally accommodate any continuous or discontinuous frequency distribution, any harmonic or inharmonic timbre, and both static and dynamic conditions.

Fourier analysis is indispensable for analyzing the frequency and phase content of acoustic signals, but empirical work has repeatedly shown the limits of purely spectral models of human hearing[2][3][4]. Tonalness[5]—the propensity of the human auditory system (HAS) to perceive virtual pitches as common fundamental frequencies of a given distribution of spectral pitches—bridges this gap. This paper shows how mathematical models of virtual pitch and critical band can be combined with Fourier

[1] Jean-Philippe Rameau, *Treatise on Harmony Reduced to Its Natural Principles*, trans. Philip Gossett (New York: Dover, 1971).

[2] Isabelle Peretz, Anne J. Blood, Virginia Penhune, and Robert Zatorre, "Cortical Deafness to Dissonance," *Brain* 124, no. 5 (2001): 928–40.

[3] Reinier Plomp and Willem J. M. Levelt, "Tonal Consonance and Critical Bandwidth," *Journal of the Acoustical Society of America* 38, no. 4 (1965): 548–60.

[4] E. Glenn Schellenberg and Sandra E. Trehub, "Frequency Ratios and the Discrimination of Pure Tone Sequences," *Perception & Psychophysics* 56, no. 4 (1994): 472–78.

[5] The term has been applied to several related quantities in previous work. It is used here strictly in the sense defined as follows.

analysis that serves as the basis for a quantitative theory of musical harmony. A table of notation follows this introduction, and the model's assumptions are collected explicitly below, as is its relationship to existing computational accounts of consonance and harmonic distance.

Notation

The following symbols recur throughout, and are collected here for reference rather than to be read in sequence. Frequencies are in hertz and times in seconds unless otherwise noted. Symbols appearing only in the Appendix are defined there.

Input Signal

Symbol	Meaning	Eq.
$P(t)$	Sound pressure signal	1
n	Number of sinusoidal components	9
a_i, b_i, c_i	Amplitude, frequency, and phase of component i	9
l_i	Component loudness, $a_i R(b_i)$	12

Variables and indices

Symbol	Meaning	Eq.
f	Frequency at which tonalness is evaluated	2
t	Time	2
u	Integration variable over time	2
h, g	Harmonic index	2, 7
i	Component index	9
h_1, h_2	Paired harmonic indices in interference terms	17

Symbol	Meaning	Eq.
i_1, i_2	Paired component indices in interference terms	17
x, y	Event labels, denoting the two sonorities compared	23

Auditory model

Symbol	Meaning	Eq.
$\sigma_h(f)$	Critical bandwidth at harmonic h	4
$\hat{\sigma}_h(f)$	Temporal window width	3
κ	Critical-band fraction, set to 0.25	4
$R(f)$	Frequency-response weighting	5
T_c	Temporal window threshold, ≈ 0.05 s	14

The transform

Symbol	Meaning	Eq.
$X(f)$	Fourier transform of P	1
$Y_h(f, t)$	Psychoacoustic Fourier function at harmonic h	2
$W_h(f, t)$	Coincidence gate	7
$\Phi_h(f, t)$	Harmonic angular position	8
$\tau(f, t)$	Tonalness function	6
$U_{h,i}(f)$	Sinusoidal smoothing function	11

Derived measures

Symbol	Meaning	Eq.
$\psi_1(f, t)$	Steady-state tonalness (SST); proxy for harmonicity	14
$\psi_2(f, t)$	Tonalness variance (TV); proxy for roughness	15
$\eta(t)$	Dissonance	18

Symbol	Meaning	Eq.
$\Omega(t_1, t_2)$	Tonalness progression	22

Auxiliary quantities in the sinusoidal case

Symbol	Meaning	Eq.
ϕ	Center frequency of an interfering pair	20
γ	Roughness factor of an interfering pair	21
$G(x, y)$	Event intensity	27

Pitch conversion

Symbol	Meaning	Eq.
s	MIDI pitch number	30
s_0, b_0	Reference pitch (69) and frequency (440 Hz)	30

Defining Tonalness

The Fourier transform, shown here as Equation 1, decomposes complex audio signals into their frequency components and is already foundational in acoustics and psychoacoustics. Several psychoacoustic phenomena—most notably critical band, frequency response, and virtual pitch—cannot be explained by Fourier analysis alone, yet they underwrite many features of harmonic perception that music theory seeks to describe.

$$X(f) = 2\int_{-\infty}^{\infty} e^{-2\pi i t f} P(t)\, dt$$

EQUATION 1. FOURIER TRANSFORM.

Equation 2 defines 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 3–4 together approximate the Glasberg & Moore critical band model[6][7]. Equation 5 uses a standard model for frequency response known as A-weighting that originates from the research of Fletcher & Munson in the 1930s[8]. We set κ to 0.25 to approximate the quarter critical band threshold identified by Plomp and Levelt[9].

$$Y_h(f,t) = \frac{2R(hf)}{h} \int_{-\infty}^{\infty} \frac{e^{-\frac{(u-t)^2}{2\hat{\sigma}_h(f)^2}} e^{-2\pi i(u-t)hf} P(u)}{\sqrt{2\pi}\hat{\sigma}_h(f)} \, du$$

EQUATION 2. PSYCHOACOUSTIC FOURIER FUNCTION.

$$\hat{\sigma}_h(f) = \frac{1}{2\pi\sigma_h(f)}$$

EQUATION 3. TEMPORAL WINDOW.

$$\sigma_h(f) = \kappa\left(\frac{24.7}{h} + 0.107939f\right)$$

EQUATION 4. CRITICAL BAND.

---

6 Brian R. Glasberg and Brian C. J. Moore, "Derivation of Auditory Filter Shapes from Notched-Noise Data," *Hearing Research* 47, no. 1–2 (1990): 103–38.

7 Brian C. J. Moore, *Cochlear Hearing Loss* (London: Whurr Publishers, 1998).

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

9 Plomp and Levelt, "Tonal Consonance*,*" 548–60.

$$R(f) = \frac{12194^2 f^4}{\left(f^2 + 20.6^2\right)\left(f^2 + 12194^2\right)\sqrt{\left(f^2 + 107.7^2\right)\left(f^2 + 737.9^2\right)}}$$

EQUATION 5. A-WEIGHTING MODEL FOR FREQUENCY RESPONSE.

Equation 2 models virtual pitches as something that has the potential to 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[10]. We therefore adopt $\frac{1}{h}$ as the amplitude scale of a virtual pitch at subharmonic h. We subtract t from u inside the transform to demodulate any carrier frequencies that result from introducing a given harmonic h into the Fourier transform.

Equations 6–8 define tonalness via a coincidence gate. The coincidence gate only allows a virtual pitch to exist in the model if it originates from more than one harmonic. The perception is strongest when many harmonics are all converging on the same subharmonic and completely vanishes when it only originates from one. The exception in Equation 7 is for the spectral pitches—represented in the model as the first harmonic—which are always allowed through the gate. Equation 8—the harmonic angular position function—exists to counterbalance and unwrap the phase angle of each virtual pitch. It is a necessary adjustment to conserve time for each virtual pitch when introducing h into the Fourier transform.

[10] Appendix.

$$\tau(f,t) = \left| \sum_{h=1}^{\infty} W_h(f,t) \left| Y_h(f,t) \right| \exp\left(\frac{i\Phi_h(f,t)}{h}\right) \right|$$

EQUATION 6. TONALNESS FUNCTION.

$$W_h(f,t) = \text{If}\left[h = 1,1,1 - \frac{\left|Y_h(f,t)\right|}{\sum_{g=1}^{\infty}\left|Y_g(f,t)\right|}\right]$$

EQUATION 7. COINCIDENCE GATE.

$$\Phi_h(f,t) = \arg\left(Y_h\left(f,t_0\right)\right) + \int_{t_0}^{t} \Im\left(\frac{\frac{\partial Y_h(f,u)}{\partial u}}{Y_h(f,u)}\right) du$$

EQUATION 8. HARMONIC ANGULAR POSITION.

Well-Definedness

Equation 6 defines τ as an infinite sum. Two properties establish that it is well defined. First, the gate is bounded. For $h = 1$ the gate returns 1 by the exception in Equation 7. For $h > 1$, each $|Y_g|$ is non-negative and the sum in the denominator includes the term $|Y_h|$ itself, so the ratio $\left|Y_h(f,t)\right| / \sum \left|Y_g(f,t)\right|$ lies in [0,1] and therefore $W_h(f,t) \in$ [0,1] for every h, f, and t. The gate can neither amplify a term nor reverse its sign.

Second, virtual pitches do not exist in silence. The ratio in Equation 7 is undefined where $\sum \left| Y_g(f,t) \right| = 0$. Because the Gaussian window of Equation 2 has full support, this occurs only when P vanishes almost everywhere. We adopt the convention $\tau(f,t) = 0$ wherever the denominator vanishes: no partials are present, and no virtual pitch should be reported.

Modeling Assumptions

The aim throughout is a minimal set of foundational assumptions from which a wide range of harmonic phenomena follow as consequences, rather than a model in which each phenomenon is separately stipulated. Where an established effect can be recovered from assumptions already in place, no additional mechanism is introduced to produce it. This principle explains several of the omissions noted below, and it also sets the standard by which the model should be judged: not by the number of phenomena it addresses, but by the ratio of those phenomena to the assumptions required to reach them.

The assumptions themselves are these:

A1. Virtual pitch is generated at every subharmonic of every spectral pitch. Equation 2 admits a candidate at $f = b/h$ for each partial b and each harmonic index h. This is a deliberate over-generation, corrected downstream by the coincidence gate rather than by restricting the candidate set in advance. The alternative—matching against a fixed harmonic template—requires committing to a template, and this approach avoids that commitment.

A2. Virtual pitch amplitude scales as $1/h$. The Appendix shows that equal energy across the harmonics of a one-dimensional standing wave implies amplitudes proportional to $1/h$. The perceptual claim is distinct from the physical one: because the auditory system has developed among predominantly harmonic sources, we take $1/h$ to represent a reasonable prior expectation about how energy distributes across a harmonic series, and therefore about the weight assigned to a fundamental inferred from a partial at harmonic h. This weighting is monotonic in h and does not separately reproduce the dominance region reported for virtual pitch[11][12]. We note that in the sonorities examined below, the partials contributing to a projected virtual pitch fall at $h \geq 3$ in every case—the major triad of Figures 17 and 19 projects its principal virtual pitch from harmonics 4, 5, and 6 —so the model's treatment of the lowest harmonic indices does not bear on the results reported here. Whether a weighting that departs from 1/h would improve the model is left open.

A3. A virtual pitch requires reinforcement from more than one harmonic. The coincidence gate of Equation 7 suppresses any candidate arising from a single harmonic, while always admitting the spectral pitches at $h = 1$. This is the model's

[11] Roelof J. Ritsma, "Frequencies Dominant in the Perception of the Pitch of Complex Sounds," *Journal of the Acoustical Society of America* 42, no. 1 (1967): 191–98.

[12] Reinier Plomp, "Pitch of Complex Tones," *Journal of the Acoustical Society of America* 41, no. 6 (1967): 1526–33.

central selection mechanism and the reason it does not report a virtual pitch beneath every isolated partial.

A4. Critical bandwidth follows Glasberg and Moore, scaled by $\kappa = 0.25$. Equation 4 approximates the ERB formulation. The value of κ approximates the quarter-critical-band threshold identified by Plomp and Levelt.

A5. Frequency response follows the A-weighting curve. A-weighting derives from the 40-phon equal-loudness contour and is level-independent, and so approximates a response that is in fact level-dependent. It is adopted for its algebraic convenience. We emphasize that this choice is modular: $R(f)$ enters the model only through pointwise evaluation at partial frequencies—in Equation 12, and in Equation 2 as a scalar factor outside the integral—so any alternative response model, including tabulated ISO 226:2003 contours, may be substituted without altering a single closed form derived below. At the levels typical of musical listening the difference is modest across the mid-frequency range where musical partials predominantly fall.

A6. The temporal window $T_c \approx 0.05$ s separates rhythm from roughness. The threshold follows from taking 20 Hz as the approximate lower limit of pitch perception. Fluctuations slower than this are heard rhythmically and are captured by the SST curve; faster fluctuations register as unsteadiness and are captured by TV.

A7. Salience is modeled by the maximum rather than the mean. Equations 14 and 15 take maxima over the temporal window, following empirical evidence that the auditory system responds to peak rather than average activity[13][14][15].

A8. Sums over h are truncated at $h = 10$. This is a numerical approximation rather than a modeling choice. The smoothing function of Equation 11 is Gaussian in $hf - b_i$, while the critical bandwidth of Equation 4 is bounded below by $\kappa 0.107939f$ for every h. The offset therefore grows linearly in h against a width that does not vanish, and the exponent grows quadratically: for indices well above b_i/f it approaches $-h^2/(2\kappa^2 0.107939^2)$, independent of f and equal to roughly $-690h^2$ at $\kappa = 0.25$. A partial at b_i consequently contributes to a candidate fundamental at f only through the harmonic index nearest b_i/f, and the neighboring indices are annihilated rather than merely attenuated. At f = 220 Hz a partial at 660 Hz peaks at h = 3, and the h = 4 term carries an exponent of approximately −432. Truncation at h = 10 therefore evaluates the sum completely for candidate fundamentals down to roughly one tenth of the lowest partial present—some three and a third octaves—and incompletely below that. For the sonorities analyzed here the contributing indices lie between 2 and 7, so the truncation is not operative in any result reported.

---

13 Alva Engell et al., "Modulatory Effects of Attention on Lateral Inhibition in the Human Auditory Cortex," *PLOS ONE* 11, no. 1 (2016): 1–15.

14 Peirun Song et al., "Temporal Merging into Pitch with Click Train in the Macaque Auditory Cortex," *National Science Review* 12, no. 5 (2025): 1–13.

15 Hector Penagos, Jennifer R. Melcher, and Andrew J. Oxenham, "A Neural Representation of Pitch Salience in Nonprimary Human Auditory Cortex Revealed with Functional Magnetic Resonance Imaging," *The Journal of Neuroscience* 24, no. 28 (2004): 6810–15.

A9. Spectral masking is not modeled as a separate stage. We take the coincidence gate to perform the function masking would otherwise serve, suppressing weakly supported candidates in dense spectra.

A10. The model describes what the auditory system generates, not to what a listener attends. The tonalness function reports many simultaneous virtual pitches, of which a listener will actively notice only a few at any moment. The same holds of spectral pitches in a complex sonority. No attentional selection stage is included; τ should be read as the field from which attention selects rather than as a prediction of reported percepts.

A11. Nonlinear distortion products are not modeled. The model generates no combination tones. The combination tones of musical significance nonetheless coincide with virtual pitches the model predicts on independent grounds: for a just interval p:q, the quadratic difference tone falls at $(q-p)f_0$, which equals the projected virtual pitch f_0 exactly whenever $(q-p)=1$—that is, for the 2:3 fifth, 3:4 fourth, 4:5 major third, and 5:6 minor third. This is why the organ resultant bass and the difference tones string players use to check intonation fall where the model already places a virtual pitch.

Scope. The following lie outside the model as formulated: cultural familiarity and learned expectation, which recent work identifies as a substantial component of consonance judgment[16]; auditory streaming and scene analysis; adaptation and other time-course effects beyond the window T_c; and binaural and spatial effects.

[16] Peter M. C. Harrison and Marcus T. Pearce, "Simultaneous Consonance in Music Perception and Composition," *Psychological Review* 127, no. 2 (2020): 216–44.

Relation to Existing Models

The tonalness function is not the first attempt to model harmonic perception from acoustic first principles, and the derivations that follow will read more clearly against a sketch of the work that precedes them. The purpose of this paper is to formalize a single model and to derive its consequences. It is not to adjudicate among competing models, and no systematic comparison is offered here. A comparison worth having requires shared stimuli, shared behavioral criteria, and a common implementation, and it is properly the subject of a separate empirical study. What follows is a map of the surrounding territory and a statement of where tonalness sits within it.

Two lineages dominate the modeling of simultaneous consonance, and they have historically been pursued separately. The first descends from Helmholtz and treats dissonance as sensory interference between partials falling within a shared critical band[17]. Plomp and Levelt gave this account its modern experimental footing[18], and Hutchinson and Knopoff turned their dyad curves into a computable measure by summing pairwise roughness over all partials of a sonority[19]. Sethares extended the approach in the direction that matters most for the present work, showing that because the roughness of an interval depends on the spectrum of the tones sounding it, consonance is a joint property of tuning and timbre rather than of interval class alone—

[17] Hermann von Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, trans. Alexander J. Ellis (New York: Dover Publications, 1954), 172–97.

[18] Plomp and Levelt, "Tonal Consonance," 548–60.

[19] William Hutchinson and Leon Knopoff, "The Acoustic Component of Western Consonance," *Interface* 7, no. 1 (1978): 1–29.

and that for an inharmonic spectrum the dissonance minima migrate away from the simple frequency ratios entirely[20]. The dissonance curves of Figure 5 below reproduce this result, and it should be credited to Sethares; that the present model recovers it from a different starting point is a check on the model rather than a new finding. Vassilakis subsequently refined the roughness estimate to account for the amplitude relationship between interfering partials[21].

The second lineage treats consonance as harmonicity: the degree to which a sonority's partials conform to a single harmonic series. Terhardt's virtual pitch theory is its foundation[22], and the algorithm of Terhardt, Stoll, and Seewann made it computable, extracting both spectral and virtual pitches, with salience values, from a spectrum[23]. Parncutt then adapted the algorithm to musical harmony directly, proposing that a chord projects not a single root but a profile of candidate roots with varying salience — a formal successor to Rameau's fundamental bass, and the closest ancestor of the tonalness function defined above[24]. Leman pursued a related account through periodicity in auditory images rather than through templates[25].

---

20 William A. Sethares, "Local Consonance and the Relationship between Timbre and Scale," *Journal of the Acoustical Society of America* 94, no. 3 (1993): 1218–28; and William A. Sethares, *Tuning, Timbre, Spectrum, Scale*, 2nd ed. (London: Springer, 2005).

21 Pantelis N. Vassilakis, "Perceptual and Physical Properties of Amplitude Fluctuation and Their Musical Significance" (PhD diss., University of California, Los Angeles, 2001).

22 Ernst Terhardt, "Calculating Virtual Pitch," *Hearing Research* 1, no. 2 (1979): 155–82.

23 Ernst Terhardt, Gerhard Stoll, and Manfred Seewann, "Algorithm for Extraction of Pitch and Pitch Salience from Complex Tonal Signals," *Journal of the Acoustical Society of America* 71, no. 3 (1982): 679–88.

24 Richard Parncutt, "Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord," *Music Perception* 6, no. 1 (1988): 65–93; and *Harmony: A Psychoacoustical Approach* (Berlin: Springer-Verlag, 1989).

25 Marc Leman, "An Auditory Model of the Role of Short-Term Memory in Probe-Tone Ratings," *Music Perception* 17, no. 4 (2000): 481–509.

A third and more recent line of work addresses the distance between sonorities rather than the character of any one of them. Lerdahl's tonal pitch space assigns quantitative distances within a cognitive-structural framework[26], Tymoczko develops a continuous geometry in which voice-leading distance is a metric on chord space[27], and Milne and colleagues model the similarity of pitch collections through spectral expectation, fitted to behavioral data[28]. The tonalness progression function of Equation 22 addresses the same question these models address, and differs from all of them in the ways described below.

Finally, Harrison and Pearce have recently synthesized much of this literature, evaluating sixteen computational consonance models against four behavioral datasets and three corpora, and concluding that simultaneous consonance is composite — arising from interference, from periodicity and harmonicity, and from cultural familiarity — rather than reducible to any one of these[29]. Their open-source implementation makes the comparison deferred here practical to carry out, and it is the natural next step for this work.

Against that background, four structural properties distinguish the model formalized in this paper. The first is the domain of the input. With few exceptions, the models above take as input a list of partials—frequencies with amplitudes, supplied by the analyst. The tonalness function takes as input a sound pressure signal, and

---

26 Fred Lerdahl, *Tonal Pitch Space* (New York: Oxford University Press, 2001).

27 Dmitri Tymoczko, *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice* (New York: Oxford University Press, 2011).

28 Andrew J. Milne et al., "Modelling the Similarity of Pitch Collections with Expectation Tensors," *Journal of Mathematics and Music* 5, no. 1 (2011): 1–20.

29 Harrison and Pearce, "Simultaneous Consonance," 216–44.

Equations 2 through 8 are defined for any P(t) whatever: discrete or continuous, harmonic or inharmonic, pitched or noise. Where the existing models require that a sonority first be idealized into a set of components, tonalness computes directly on the waveform. This is why the model can be applied to a recording, or to signals for which no notated equivalent exists at all.

The second is time. The models above are static: they return a value for a sonority considered as a fixed object. The tonalness function varies in time by construction, which is what permits the steady-state and variance functions of Equations 14 and 15, and with them a treatment of beating, attack, decay, and vibrato as consequences of the model rather than as effects excluded from it.

The third is unification of roughness and harmonicity. These have generally been modeled by separate mechanisms and then combined, if at all, by weighting their outputs. Here both are read from the same function—harmonicity from ψ_1 and roughness from ψ_2—and the coincidence gate of Equation 7 is the single mechanism responsible for both.

The fourth is unification of spectral and virtual pitch. Terhardt's algorithm and Parncutt's adaptation compute the two through separate stages with separate salience formulas, and both are designed to identify a small set of pitch candidates rather than to map a full intensity distribution. In the present model a spectral pitch is simply the case $h = 1$, admitted by the exception in Equation 7, and τ returns a continuous intensity landscape over frequency rather than a ranked list of candidates. This difference should be borne in mind in any direct comparison, since the objects returned by the two approaches are not of the same kind.

These differences describe what the model is built to do; they are not claims about how well it does it. The sections that follow derive the model's consequences and check them qualitatively against established experimental findings. Quantitative evaluation against behavioral data, and against the models named above, is deferred.

Simple Sinusoidal Model

To explore the model's implications we need an input sound pressure signal (SPS). Equations 9–13 collectively define the tonalness function for an additive sinusoidal SPS.

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

EQUATION 9. ADDITIVE SINUSOIDAL SOUND PRESSURE SIGNAL.

$$\tau(f,t) = \left| \sum_{h=1}^{\infty} \sum_{i=1}^{n} W_h(f,t) U_{h,i}(f) e^{\frac{i\left(2\pi t b_i + c_i\right)}{h}} \right|$$

EQUATION 10. SINUSOIDAL TONALNESS FUNCTION.

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

EQUATION 11. SINUSOIDAL SMOOTHING FUNCTION.

$$l_i = a_i R\left(b_i\right)$$

EQUATION 12. COMPONENT LOUDNESS.

$$W_h(f,t) = \text{If}\left[h = 1, 1, 1 - \frac{\sum_{i=1}^{n} U_{h,i}(f)}{\sum_{i=1}^{n} \sum_{g=1}^{\infty} U_{g,i}(f)}\right]$$

EQUATION 13. SINUSOIDAL COINCIDENCE GATE.

The tonalness function acts as a psychoacoustic Fourier transform, but unlike an ordinary Fourier transform it varies in time. The sections below show that beating, tonal fusion, and dissonance all arise as mathematical consequences folding critical band, frequency response, 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 to a conventional spectral distribution, consider three equal loudness pure-tones at 440 Hz, 554 Hz, and 659 Hz —an approximation of a 12-tone equal tempered major triad with waveform $P(t) = \cos(2\pi 440t) + \cos(2\pi 554t) + \cos(2\pi 659t)$. Figure 1 plots the amplitude of the Fourier transform in red and $|\tau(f,0)|$ in black for this input.

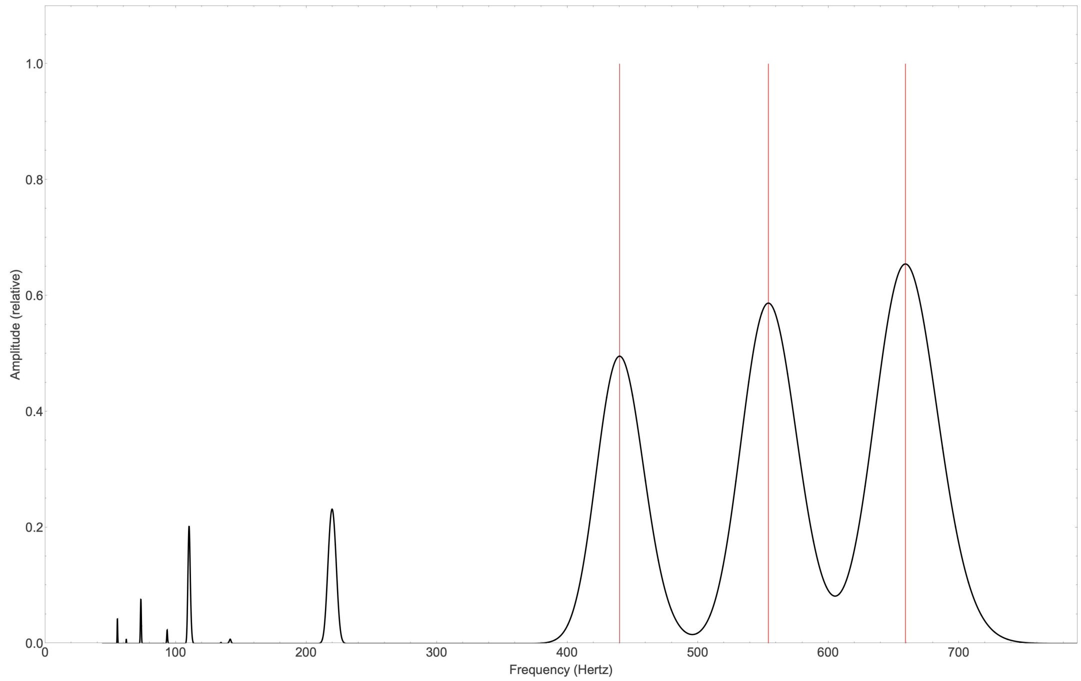

FIGURE 1. SPECTRAL INTENSITY (RED) VERSUS TONALNESS (BLACK) FOR THREE EQUAL AMPLITUDE PURE TONES: 440 HZ, 554 HZ, AND 659 HZ.

Figure 1 shows that the tonalness function models virtual pitches as a series of bell curves occurring at the subharmonics of each spectral pitch. The height of each curve depends primarily on the alignment of the subharmonics, the frequency response, and the sum of the inverse of their harmonic numbers. The width depends on the critical band. The two most prominent virtual pitches occur near 110 Hz and 220 Hz, which can be understood qualitatively from the arrangement of the spectral pitches. Perceived as a fundamental at 110 Hz, the tones form an approximate 4:5 major third (440/554) stacked with a 5:6 minor third (554/659). Perceived as a fundamental at 220 Hz, 440 Hz and 659 Hz form an approximate 2:3 perfect fifth. Because 2:3 involves smaller harmonic numbers, we would typically expect the stronger virtual pitch at 220 Hz. However, two intervals contribute aligned virtual pitches at 110 Hz, giving it an intensity nearly equal to that at 220 Hz, where only one interval aligns.

Virtual pitches function both as a proxy for harmonicity at a given spectral location and as pseudo-pitches perceived directly by the HAS. They are not physically present in the SPS but emerge from how the HAS responds to what it hears. The auditory system generally perceives many virtual pitches simultaneously, which means the HAS can register high harmonicity in one part of a harmony and lower harmonicity in another.

Static tonalness plots encode a great deal of psychoacoustic information about a harmony, but tonalness itself evolves in time. To capture that dynamic behavior we introduce two auxiliary functions. Equation 14 defines the steady state tonalness (SST) function as the maximum of the tonalness function over a temporal window T_c for each frequency f. Equation 15 defines the tonalness variance (TV) function as the maximum of the first derivative of the frequency scaled tonalness function over T_c at each frequency f. Equations 16 and 17 specialize SST and TV to an additive sinusoidal SPS.

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

EQUATION 14. STEADY STATE TONALNESS FUNCTION

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

EQUATION 15. TONALNESS VARIANCE FUNCTION

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

EQUATION 16. SINUSOIDAL STEADY STATE TONALNESS FUNCTION

$$\psi_2(f,t) = \sqrt{\frac{2\pi}{f} \sum_{h_1=1}^{\infty} \sum_{h_2=1}^{\infty} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} W_{h_1}(f,t) W_{h_2}(f,t) U_{h_1,i_1}(f) U_{h_2,i_2}(f) \left| \frac{b_{i_2}}{h_2} - \frac{b_{i_1}}{h_1} \right|}$$

EQUATION 17. SINUSOIDAL TONALNESS VARIANCE FUNCTION

The temporal window threshold $T_c \approx 0.05$ s separates fluctuations slow enough to be heard rhythmically (captured by the SST curve) from those too fast to be heard as rhythm but will instead register as unsteadiness (captured by the TV curve). We base this threshold on the assumption of 20 Hz being the approximate lowest perceivable pitch. We use maximum rather than mean values to incorporate the empirical evidence of auditory salience[30][31][32]. Figure 2 plots the SST and TV for the three pure-tone example.

30 Engell et al., "Modulatory Effects," 1–15.

31 Song et al., "Temporal Merging," 1–13.

32 Penagos, Melcher, and Oxenham, "Neural Representation," 6810–15.

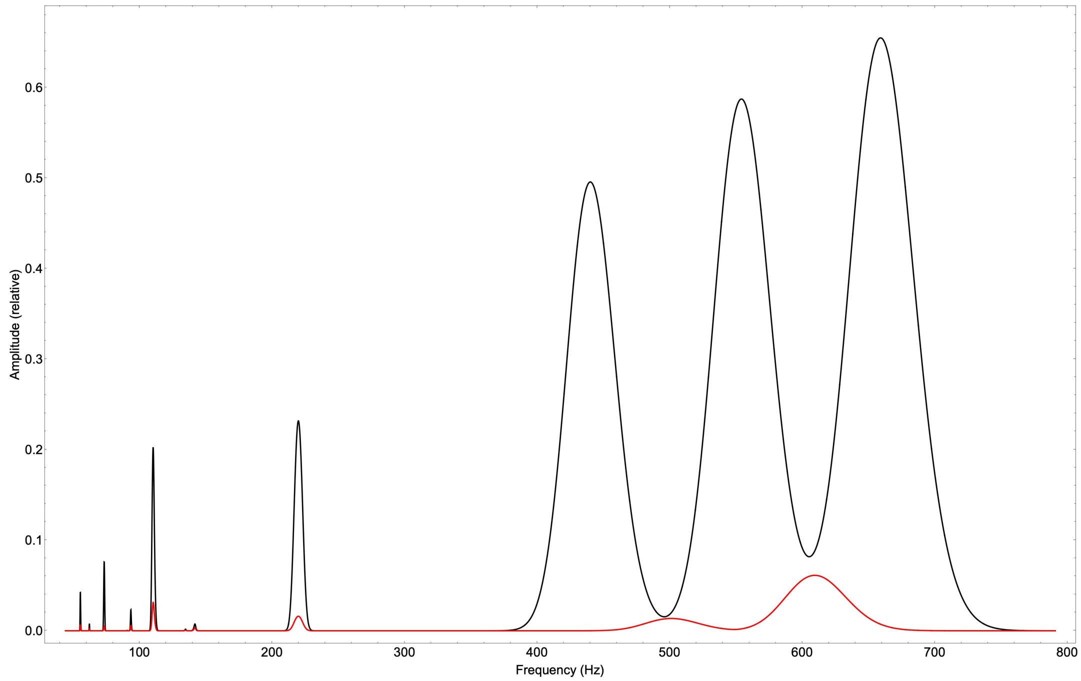

FIGURE 2. STEADY-STATE TONALNESS (BLACK) AND TONALNESS VARIANCE (RED) FOR THREE EQUAL AMPLITUDE PURE TONES: 440 HZ, 554 HZ, AND 659 HZ.

Figure 2 closely mirrors Figure 1 because most of the oscillations in tonalness are small and too rapid to appear the SST curve. Substantial variation in the TV curve near 110 Hz and 220 Hz reflects the subtle mistuning of the simple harmonic ratios among the spectral pitches. Additional peaks at other locations indicate roughness that the HAS perceives among virtual pitches, spectral pitches, and between virtual pitches and spectral pitches.

Figure 3 plots SST and TV for a justly tuned major triad—440 Hz, 550 Hz, and 660 Hz. The TV curve now drops nearly to zero at 110 Hz and 220 Hz, while the SST peaks at those frequencies rise slightly above those of the equal-tempered case, as expected when the spectral pitches are tuned to simple harmonic ratios.

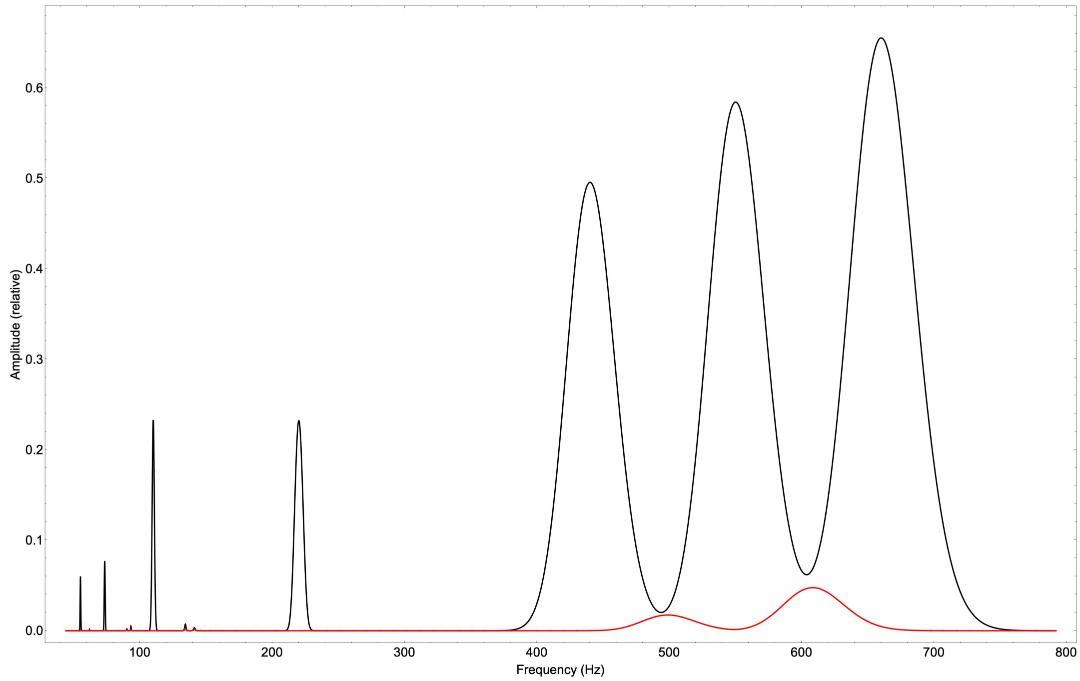

FIGURE 3. STEADY-STATE TONALNESS (BLACK) AND TONALNESS VARIANCE (RED) FOR THREE EQUAL AMPLITUDE PURE TONES: 440 HZ, 550 HZ, AND 660 HZ.

Subsequent examples will show that no matter how an interval or chord is tuned —even for unisons and octaves—the TV function retains some variation, because any interval within a chord can be heard as a mistuning of some other nearby interval. Plomp and Levelt's work suggests that roughness should be inaudible for pure-tone intervals outside the critical band[33]. The tonalness model predicts small but real variations there, variations that compound as more pure-tones are added to the SPS. This behavior also helps explain why spectral pitches tuned close to a harmonic series are perceived as a unified timbre.

SST can be read as a proxy for perceived harmonicity and TV as a proxy for perceived roughness. In this view, neither quantity is a single scalar for a given harmony but a two-dimensional function varying with pitch. The HAS can register high

[33] Plomp and Levelt, "Tonal Consonance," 553–60.

harmonicity in one region of a harmony and much lower harmonicity in another, and the same holds for roughness.

To compare the tonalness model with other models of harmonic perception, Equation 18 defines a dissonance function as the aggregate intensity of variation in the TV function, normalized by the aggregate intensity of SST. We use energy rather than amplitude to model dissonance. The introduction of virtual pitches allows us to combine roughness and harmonicity models of dissonance into a single model. Equations 19–21 specialize this definition to additive sinusoidal SPSs, enabling direct comparison with existing dissonance models.

$$\eta(t) = \frac{\int_{-\infty}^{\infty} \psi_2(f,t)^2\, df}{\int_{-\infty}^{\infty} \psi_1(f,t)^2\, df}$$

EQUATION 18. DISSONANCE FUNCTION

$$\eta(t) = \frac{2\pi \sum_{h_1=1}^{\infty} \sum_{h_2=1}^{\infty} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \frac{W_{h_1}(f,t) W_{h_2}(f,t) \gamma_{h_1,h_2,i_1,i_2}}{\phi_{h_1,h_2,i_1,i_2}} \left| \frac{b_{i_2}}{h_2} - \frac{b_{i_1}}{h_1} \right|}{\sum_{h_1=1}^{\infty} \sum_{h_2=1}^{\infty} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} W_{h_1}(f,t) W_{h_2}(f,t) \gamma_{h_1,h_2,i_1,i_2}}$$

EQUATION 19. SINUSOIDAL DISSONANCE FUNCTION

$$\phi_{h_1,h_2,i_1,i_2} = \frac{b_{i_1} h_1 + b_{i_2} h_2}{h_1^2 + h_2^2}$$

EQUATION 20. SINUSOIDAL CENTER FREQUENCY

$$\gamma_{h_1,h_2,i_1,i_2} = \frac{l_{i_1} l_{i_2} \sigma_{h_1}\left(\phi_{h_1,h_2,i_1,i_2}\right) \sigma_{h_2}\left(\phi_{h_1,h_2,i_1,i_2}\right)}{h_1 h_2 \sqrt{h_1^2 \sigma_{h_1}\left(\phi_{h_1,h_2,i_1,i_2}\right)^2 + h_2^2 \sigma_{h_2}\left(\phi_{h_1,h_2,i_1,i_2}\right)^2}} \exp\left(-\frac{\left(b_{i_2} h_1 - b_{i_1} h_2\right)^2}{2\left(h_2^2 \sigma_{h_1}\left(\phi_{h_1,h_2,i_1,i_2}\right)^2 + h_1^2 \sigma_{h_2}\left(\phi_{h_1,h_2,i_1,i_2}\right)^2\right)}\right)$$

EQUATION 21. SINUSOIDAL ROUGHNESS FACTOR

Applying Equations 19–21 yields a dissonance value of 0.00396111 for the equal tempered triad and 0.00251119 for the justly tuned triad. As expected, the justly tuned version is less dissonant than the equal tempered version.

Before turning to the musical predictions of Equation 18, we introduce one further analytical tool. Equation 22 defines the tonalness progression function (TPF), which measures how much SST shifts between two points in time for a given SPS. Equations 23–29 specialize this definition to additive sinusoidal SPSs.

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

EQUATION 22. TONALNESS PROGRESSION FUNCTION

$$\psi_1\left(f, t_x\right) = \sum_{h=1}^{\infty} \sum_{i=1}^{n_x} W_h\left(f, t_x\right) U_{h,i,x}(f)$$

EQUATION 23. SINUSOIDAL STEADY STATE TONALNESS EVENT

$$W_h\left(f, t_x\right) = \text{If}\left[h = 1,1,1 - \frac{\sum_{i=1}^{n_x} U_{h,i,x}(f)}{\sum_{i=1}^{n_x} \sum_{g=1}^{\infty} U_{g,i,x}(f)}\right]$$

EQUATION 24. SINUSOIDAL COINCIDENCE GATE EVENT

$$U_{h,i,x}(f) = \frac{l_{i,x}}{h} e^{-\frac{\left(hf - b_{i,x}\right)^2}{2\sigma_h(f)^2}}$$

EQUATION 25. SINUSOIDAL SMOOTHING EVENT

$$\Omega\left(t_t, t_2\right) = \frac{G(1{,}1) - 2G(1{,}2) + G(2{,}2)}{G(1{,}1) + 2G(1{,}2) + G(2{,}2)}$$

EQUATION 26. SINUSOIDAL TONALNESS PROGRESSION FUNCTION

$$G(x, y) = \sum_{h_1=1}^{\infty} \sum_{h_2=1}^{\infty} \sum_{i_1=1}^{n_x} \sum_{i_2=1}^{n_y} W_{h_1}\left(\phi_{h_1,h_2,i_1,i_2}(x, y), t_x\right) W_{h_2}\left(\phi_{h_1,h_2,i_1,i_2}(x, y), t_y\right) \gamma_{h_1,h_2,i_1,i_2}(x, y)$$

EQUATION 27. SINUSOIDAL EVENT INTENSITY

$$\gamma_{h_1,h_2,i_1,i_2}(x, y) = \frac{l_{i_1,x} l_{i_2,y} \sigma_{h_1}\left(\phi_{h_1,h_2,i_1,i_2}(x, y)\right) \sigma_{h_2}\left(\phi_{h_1,h_2,i_1,i_2}(x, y)\right)}{h_1 h_2 \sqrt{h_2^2 \sigma_{h_1}\left(\phi_{h_1,h_2,i_1,i_2}(x, y)\right)^2 + h_1^2 \sigma_{h_2}\left(\phi_{h_1,h_2,i_1,i_2}(x, y)\right)^2}} \exp\left(-\frac{\left(h_2 b_{i_1,x} - h_1 b_{i_2,y}\right)^2}{2\left(h_2^2 \sigma_{h_1}\left(\phi_{h_1,h_2,i_1,i_2}(x, y)\right)^2 + h_1^2 \sigma_{h_2}\left(\phi_{h_1,h_2,i_1,i_2}(x, y)\right)^2\right)}\right)$$

EQUATION 28. SINUSOIDAL EVENT FACTOR

$$\phi_{h_1,h_2,i_1,i_2}(x, y) = \frac{b_{i_{1,x}} h_1 + b_{i_{2,y}} h_2}{h_1^2 + h_2^2}$$

EQUATION 29. SINUSOIDAL CENTER EVENT FREQUENCY

Applying Equations 23–29, the TPF between the equal-tempered and justly tuned triads is 0.0416506—a very small value, as subsequent examples will show.

Formal Properties of η and Ω

η is bounded in [0,1] and also invariant under uniform gain since the numerator and denominator of Equation 18 each scale as λ^2. This invariance is what enables the comparison of η across sonorities of differing loudness and differing cardinality reported in Figures 13–15.

η is not invariant under transposition, by design. Transposing a sonority alters $\sigma_h(f)$ through Equation 4 and the component loudnesses through Equation 12, and η changes accordingly. This is a property of the model rather than a defect: the register dependence displayed in Figures 6 and 7 is a substantive prediction, and a measure invariant under transposition could not make it.

η is defined for sonorities of any cardinality. No term in Equations 19–21 depends on n except through the ranges of summation, so dyads, trichords, and dense sonorities are evaluated by the same expression and are directly comparable.

Ω is also bounded in [0,1]. The endpoints interpret directly: $\Omega = 0$ exactly when the two sonorities produce identical harmonicity landscapes, and $\Omega = 1$ exactly when two sonorities share no harmonicity anywhere in the spectrum.

Ω is symmetric but is not a metric. Symmetry is immediate from Equation 22, since both integrands are unchanged when A and B are exchanged. The triangle inequality, however, fails. The progression value between two sonorities cannot be recovered by summing progression values along a path through intermediate sonorities. A harmonically remote pair may be connected by two individually small steps, and Ω will register the direct motion as large regardless. Progression values should be read pairwise and not accumulated.

Dissonance

The dissonance function is the most direct application of tonalness theory to musical harmony. We begin with dyads. For convenience of both musical and mathematical analysis, pitch is expressed in MIDI numbers: MIDI 69 = A_4 = 440 Hz, and a MIDI interval (MI) of 1 corresponds to one semitone, or 100 cents. Equation 30 gives the MIDI-to-frequency relation, with set s_0 = 69 and b_0 = 440 Hz.

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

EQUATION 30. MIDI PITCH TO FREQUENCY CONVERSION

Figure 4 uses Equations 19–21 to plot dissonance for pure-tone dyads (s_1 lower, s_2 upper) spanning the unison (MI 0) through three octaves (MI 36) with s_1 fixed at MIDI 48 (black), 60 (red), and 72 (blue).

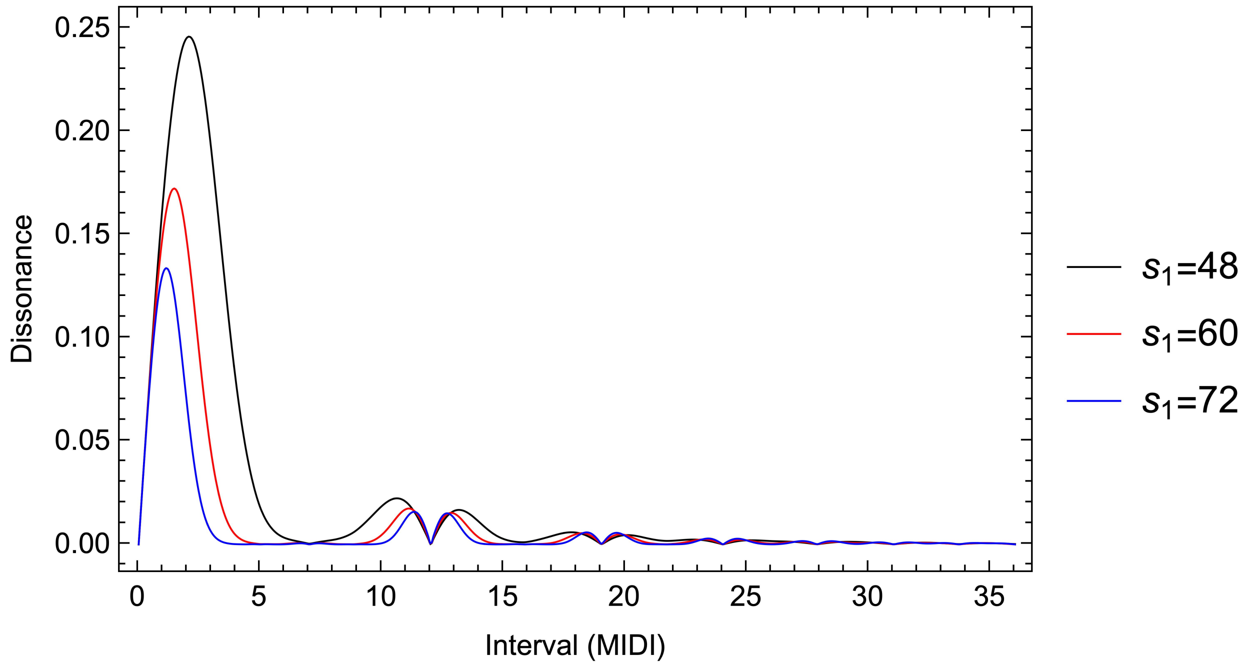

FIGURE 4. DISSONANCE PLOT FOR PURE-TONE DYADS IN VARIOUS REGISTERS.

As we see from Figure 4, Equations 19–21 predict strong register dependence: curve detail grows as the register rises, while overall dissonance grows as the register falls. Two pitches inside the critical band are predicted to be maximally dissonant, and this critical-band dissonance produces the highest values on the graph at every register. The relative maxima also shift modestly as the interval descends in register. The predictions for pure-tone dissonance agree broadly with the experimental evidence of Plomp and Levelt[34].

Timbre also affects dyad dissonance. We compare three timbres: (i) a pure-tone; (ii) a harmonic timbre with partials 1, 2, 3, 4, 5 at relative amplitudes of 1, 1/2, 1/3, 1/4, 1/5; and (iii) an inharmonic timbre with partials 1, 2.1, 3.2412, 4.41, 5.59977 at the same amplitudes. Figure 5 plots dissonance from Equations 19–21 for MI 0 through MI

[34] Plomp and Levelt, "Tonal Consonance," 548–60.

36 above MIDI 60 for pure-tones (blue), the harmonic-timbre (red), and the inharmonic-timbre (black).

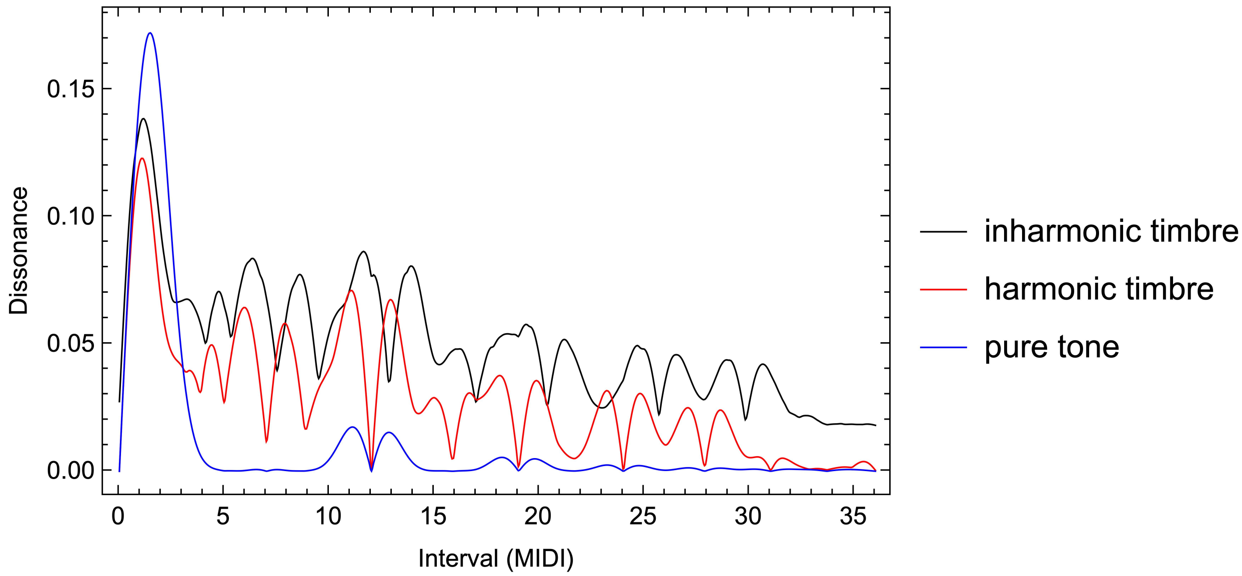

FIGURE 5. DISSONANCE PLOT FOR DYADS ABOVE MIDI 60 CONSISTING OF PURE TONES (BLUE), A HARMONIC TIMBRE (RED), AND AN INHARMONIC TIMBRE (BLACK).

The harmonic-timbre curve mirrors the pure-tone curve but with exaggerated features. The inharmonic-timbre curve is shifted and more uniform by comparison. Because its overtones produce a 2.1:1 pseudo-octave, a pronounced dissonance valley appears slightly above MI 12. Equations 19–21 predict that while a stretched octave built from an inharmonic timbre can produce locally low dissonance, it never approaches the minima available to the harmonic-timbre or pure-tone 2:1 octave— a prediction consistent with the experimental results of McDermott, Lehr, and Oxenham[35].

[35] Josh H. McDermott, Andriana J. Lehr, and Andrew J. Oxenham, "Individual Differences Reveal the Basis of Consonance," *Current Biology* 20, no. 11 (2010): 1035–47.

To probe the interval-register relationship more systematically, Figure 6 plots the dissonance for pure-tone dyads of MI 0–12 with lower pitches spanning MIDI 24–72. Figure 7 repeats the calculation for the harmonic timbre defined above.

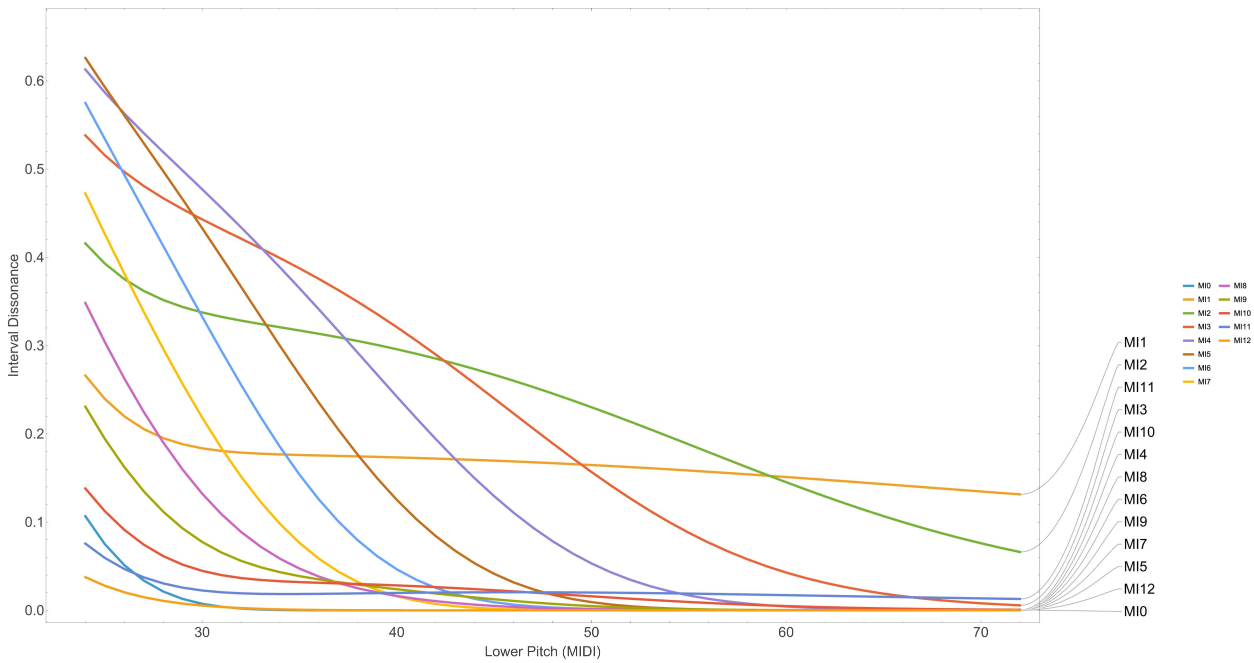

FIGURE 6. DISSONANCE PLOT FOR DYADS WITH LOWER PITCHES SPANNING FROM MIDI 24 TO MIDI 72 CONSISTING OF PURE TONES.

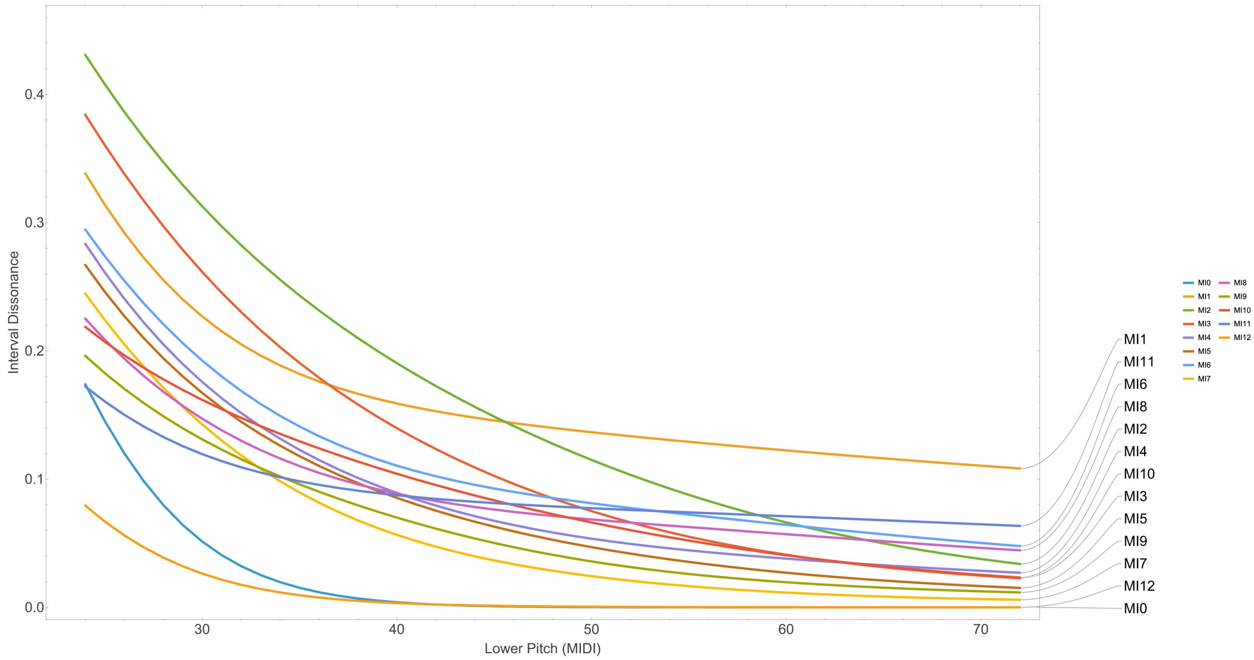

FIGURE 7. DISSONANCE PLOT FOR DYADS WITH LOWER PITCHES SPANNING FROM MIDI 24 TO MIDI 72 CONSISTING OF A HARMONIC TIMBRE.

Figures 6 and 7 show that the dissonance of a given harmonic interval varies inversely with register, though not uniformly, and that the relative ordering of intervals by dissonance depends on both register and timbre.

Trichords exhibit analogous behavior. Figures 8–12 show continuous dissonance plots for trichords with component intervals spanning the unison through octave.

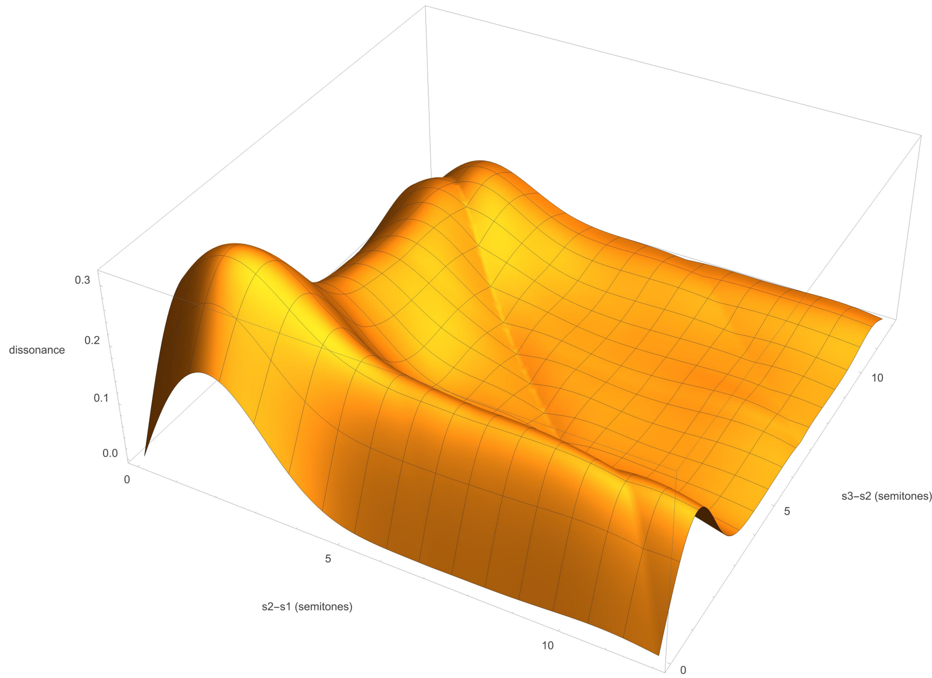

FIGURE 8. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 48. THESE TRICHORDS ARE COMPOSED OF PURE TONES.

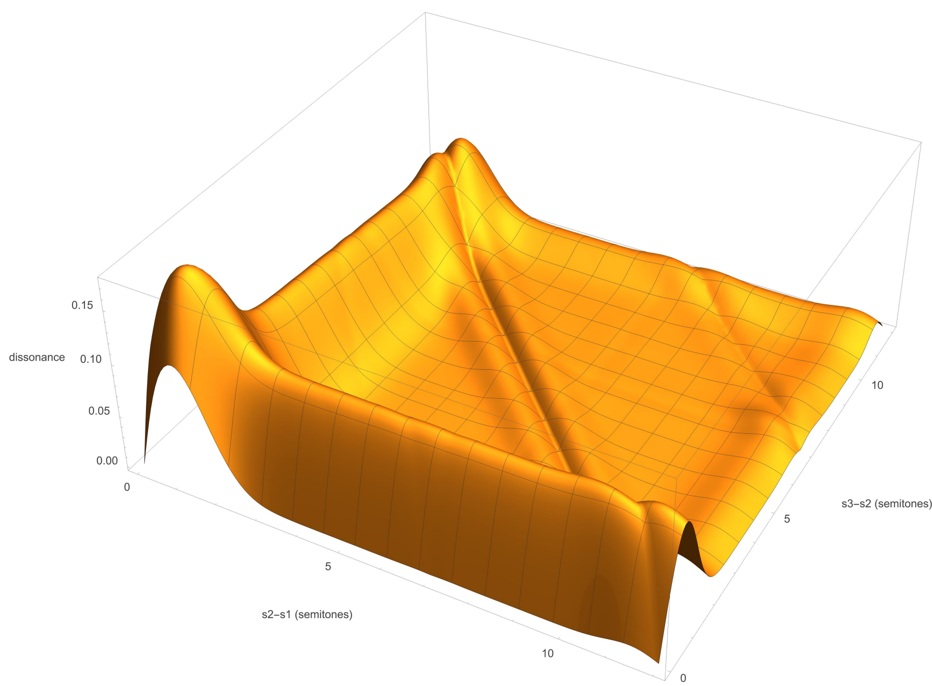

FIGURE 9. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 72. THESE TRICHORDS ARE COMPOSED OF PURE TONES.

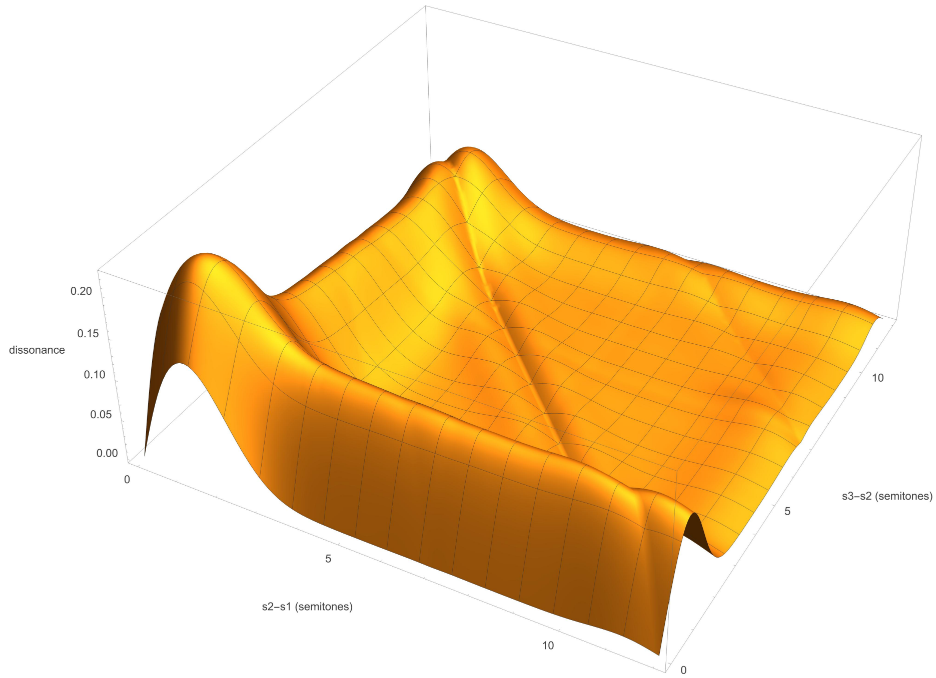

FIGURE 10. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 60. THESE TRICHORDS ARE COMPOSED OF PURE TONES.

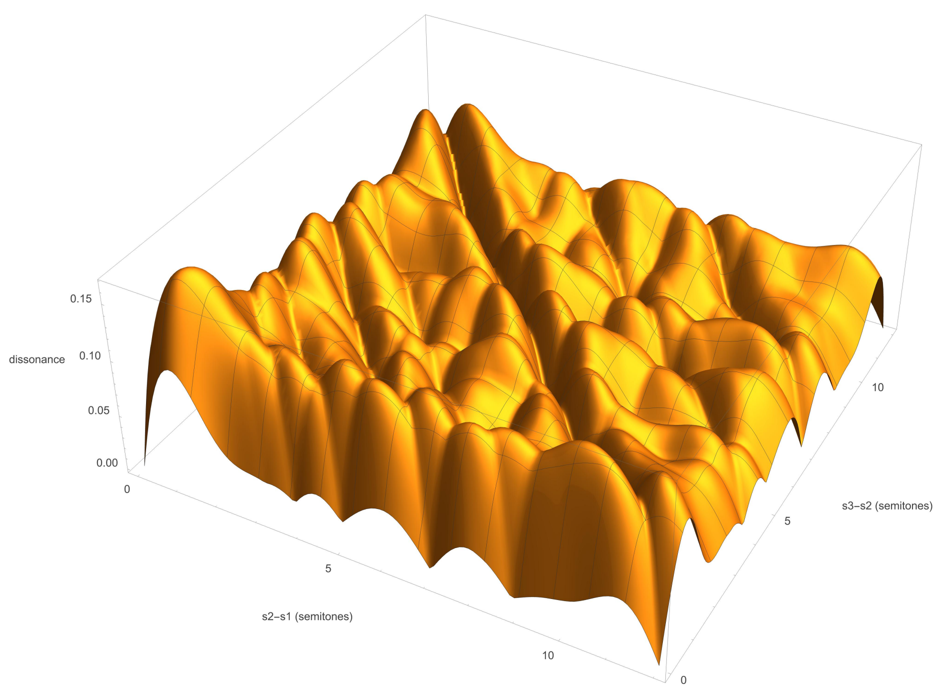

FIGURE 11. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 60. THESE TRICHORDS ARE COMPOSED OF A HARMONIC TIMBRE WITH PARTIALS 1, 2, 3, 4 AND 5 WITH RESPECTIVE RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

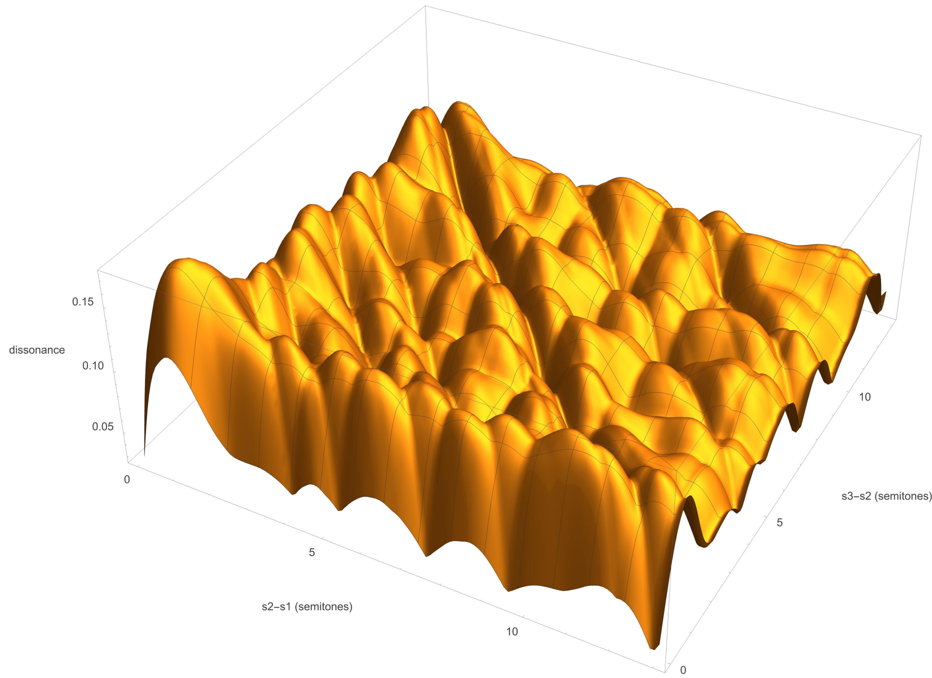

FIGURE 12. DISSONANCE PLOT FOR TRICHORDS CONSISTING OF PITCHES S1, S2, AND S3 WHERE S1 IS MIDI 60. THESE TRICHORDS ARE COMPOSED OF AN INHARMONIC TIMBRE WITH PARTIALS 1, 2.1, 3.2412, 4.41, AND 5.59977 WITH RESPECTIVE RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

Figures 8–12 show that the largest dissonance peaks occur when both component intervals of the trichord lie within the critical band. Dissonance is also elevated near harmonic intervals, and—as with dyads—the inharmonic timbre produces the same characteristic shift in these patterns. The patterns themselves are most pronounced for trichords built above MIDI 72, and the dissonance valleys are sharper for the harmonic timbre than for pure tones.

Another advantage of this dissonance model is that it allows comparison across chords of different cardinality and amplitude. Figures 13–15 show dissonance values for an assortment of two, three, and four pitch chords built above MIDI 60, using both pure tones and the harmonic timbre defined above.

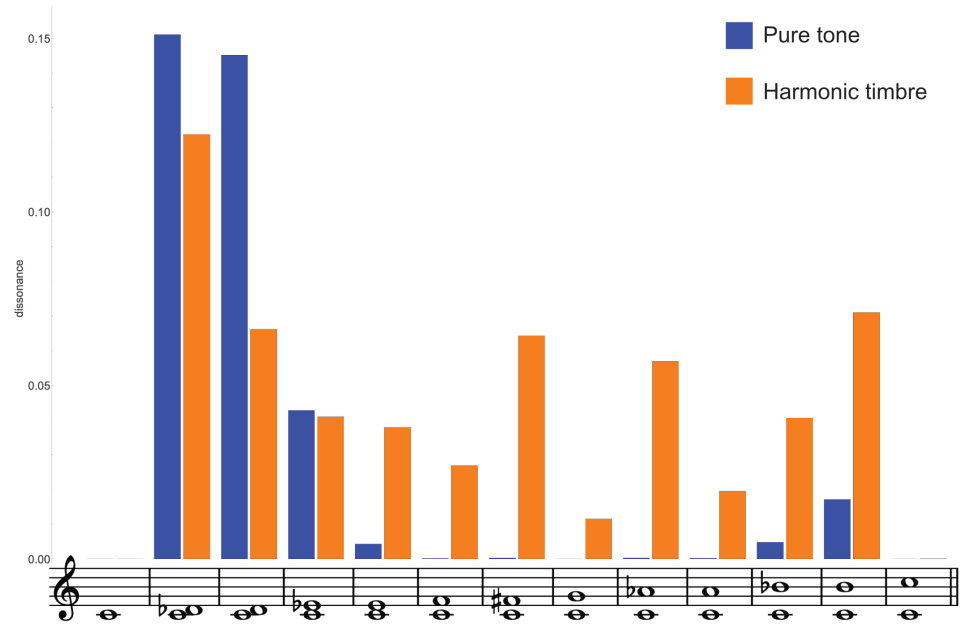

FIGURE 13. DISSONANCE VALUES OF VARIOUS TWO PITCH CHORDS COMPOSED OF PURE TONES (ORANGE) AND A HARMONIC TIMBRE (BLUE) CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

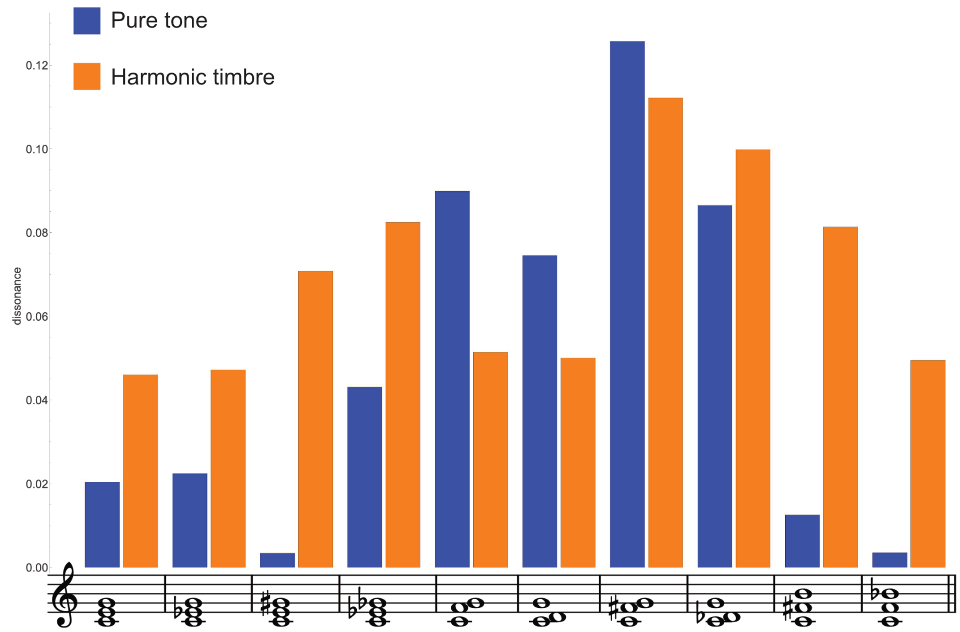

FIGURE 14. DISSONANCE VALUES OF VARIOUS THREE PITCH CHORDS COMPOSED OF PURE TONES (ORANGE) AND A HARMONIC TIMBRE (BLUE) CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

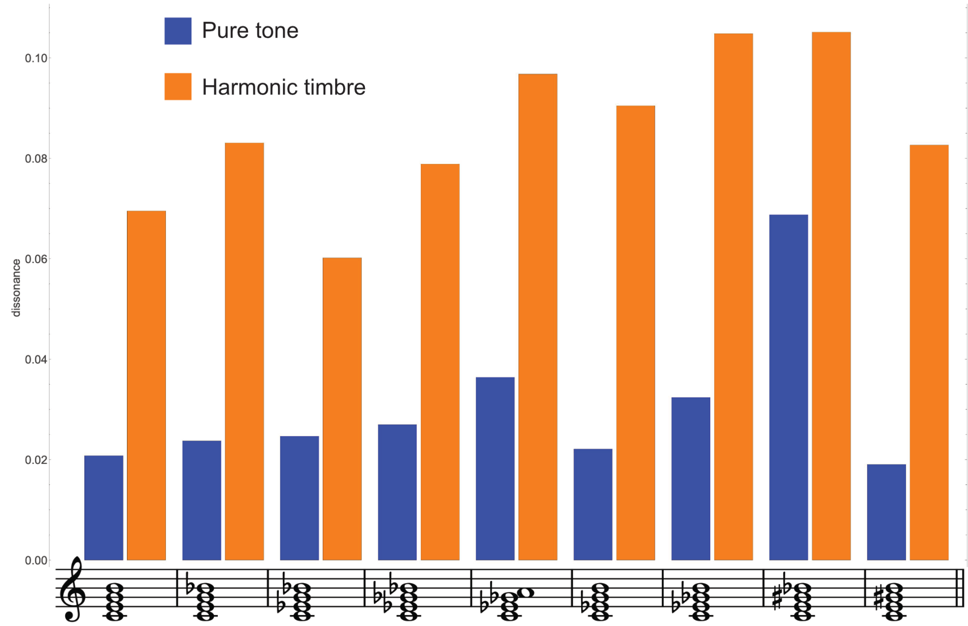

FIGURE 15. DISSONANCE VALUES OF VARIOUS FOUR PITCH CHORDS COMPOSED OF PURE TONES (ORANGE) AND A HARMONIC TIMBRE (BLUE) CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

Figures 13–15 show that the predicted dissonance ordering broadly aligns with intuitions familiar to practitioners of Western tonal music, though a few results are surprising. These results should not be over-interpreted. By considering only the intensity of beating between spectral and virtual pitches, Equations 19–21 calculate a very specific type of dissonance perception and cannot substitute for the broader notion of dissonance operative within Western tonality. This is both a feature and a shortcoming of any psychoacoustic dissonance model.

Tonalness Progression

The dissonance of individual chords is only one determinate of harmonic experience. Chord progressions are another. The TPF (Equation 22) measures the aggregate change in SST between two chords, integrating the shifts in both spectral and virtual pitches to produce a proxy for how smoothly or abruptly one chord moves to the next. Figure 16 applies Equations 19–21 and 23–29 to several common progressions, rendered with both pure tones and the harmonic timbre defined above.

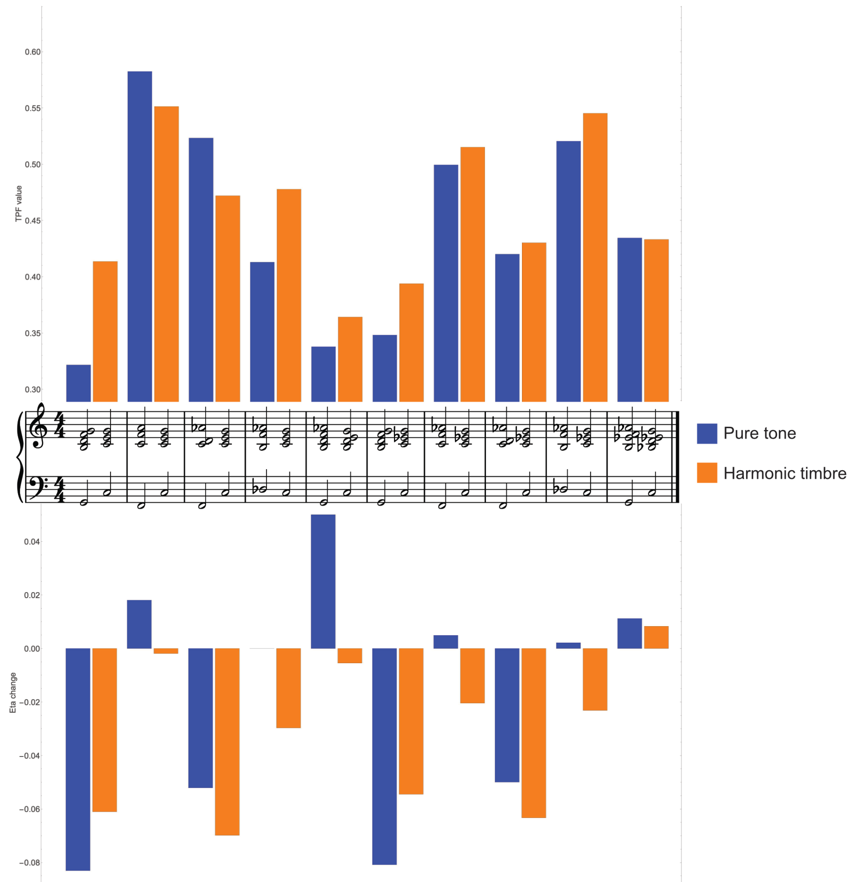

FIGURE 16. TPF VALUES AND DISSONANCE CHANGE FOR VARIOUS COMMON CHORD PROGRESSIONS IN THE KEY OF C CONSISTING OF PURE-TONE PITCHES AND PITCHES WITH A HARMONIC TIMBRE CONTAINING PARTIALS 1, 2, 3, 4, AND 5 WITH RELATIVE AMPLITUDES OF 1, 1/2, 1/3, 1/4, AND 1/5.

Figure 16 shows that most trends visible for pure-tone chords carry over to chords built from the harmonic timbre. Caution is again warranted. While the TPF and dissonance function illuminate important psychoacoustic properties of harmonic progression, a full account of why progressions sound the way that they do requires returning to the SST and TV functions directly.

Virtual Pitch Analysis

Ernst Terhardt's theories of virtual pitch[36] are, in a sense, the modern successor to Rameau's fundamental bass. The SST and TV functions defined above serve as a comprehensive model of virtual pitch, under which a chord does not possess a single fundamental bass but rather a continuous spectrum of perceived virtual pitches. Virtual pitch analysis (VPA) examines the patterns of—and interference among—these virtual pitches to characterize the psychoacoustic perception of a specific harmony. Figures 17 and 18 compare SST and TV for a MIDI {60, 64, 67} major triad and a MIDI {60, 63, 67} minor triad, each built from equal amplitude pure tones.

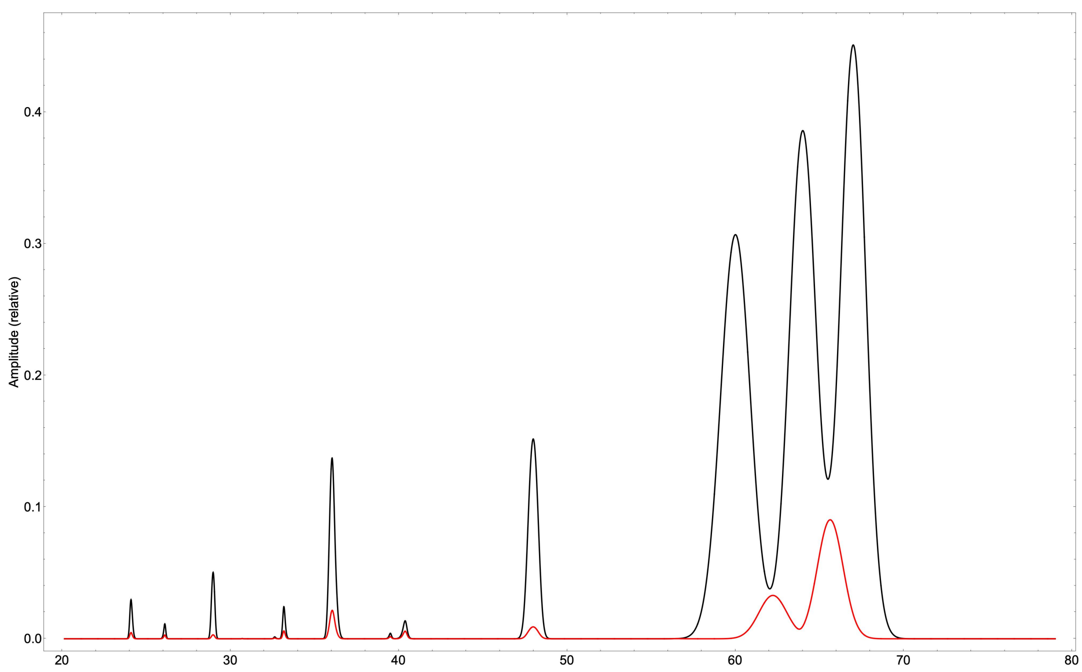

FIGURE 17. SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 64, 67 MAJOR TRIAD CONSISTING OF EQUAL AMPLITUDE PURE-TONE PITCHES.

[36] Terhardt, "Calculating Virtual Pitch," 155–82.

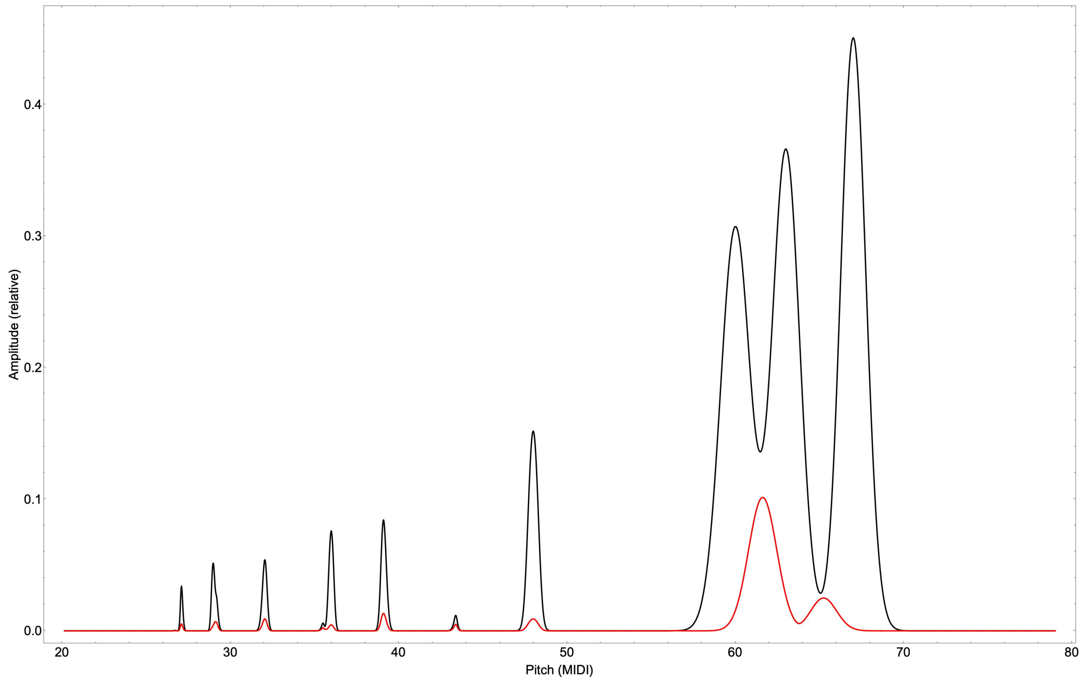

FIGURE 18. SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 63, 67 MINOR TRIAD CONSISTING OF EQUAL AMPLITUDE PURE-TONE PITCHES.

Beyond the slightly higher TV curve—and therefore greater dissonance—of the minor triad, the major triad exhibits a noticeably more prominent virtual pitch near MIDI 36. This occurs because the spectral pitches of the major triad approximate the 4th, 5th, and 6th harmonics of MIDI 36. That virtual pitch is the largest to result from all three spectral pitches jointly. The minor triad shows weaker virtual pitches near MIDI 32 and 39. Its spectral pitches align only loosely with any single harmonic series, instead splitting across two (one rooted above MIDI 32, and the other above MIDI 39). Although both triads share interval content and have similar dissonance values, the major triad should sound more focused and unified than the minor.

A similar contrast appears when comparing the MIDI {60, 64, 67, 70} major-minor seventh chord with the MIDI {60, 63, 66, 70} half-diminished seventh chord.

Figures 19 and 20 plot SST and TV for these chords, rendered with equal amplitude pure-tone component pitches.

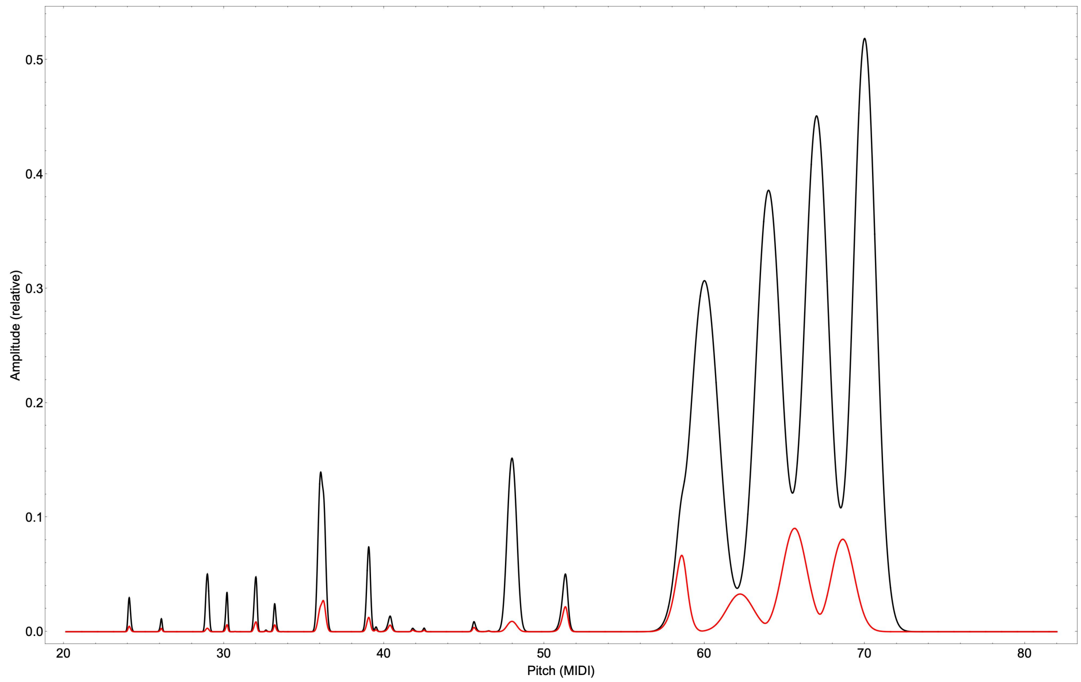

FIGURE 19. SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 64, 67, 70 MAJOR-MINOR SEVENTH CHORD CONSISTING OF EQUAL AMPLITUDE PURE-TONE PITCHES.

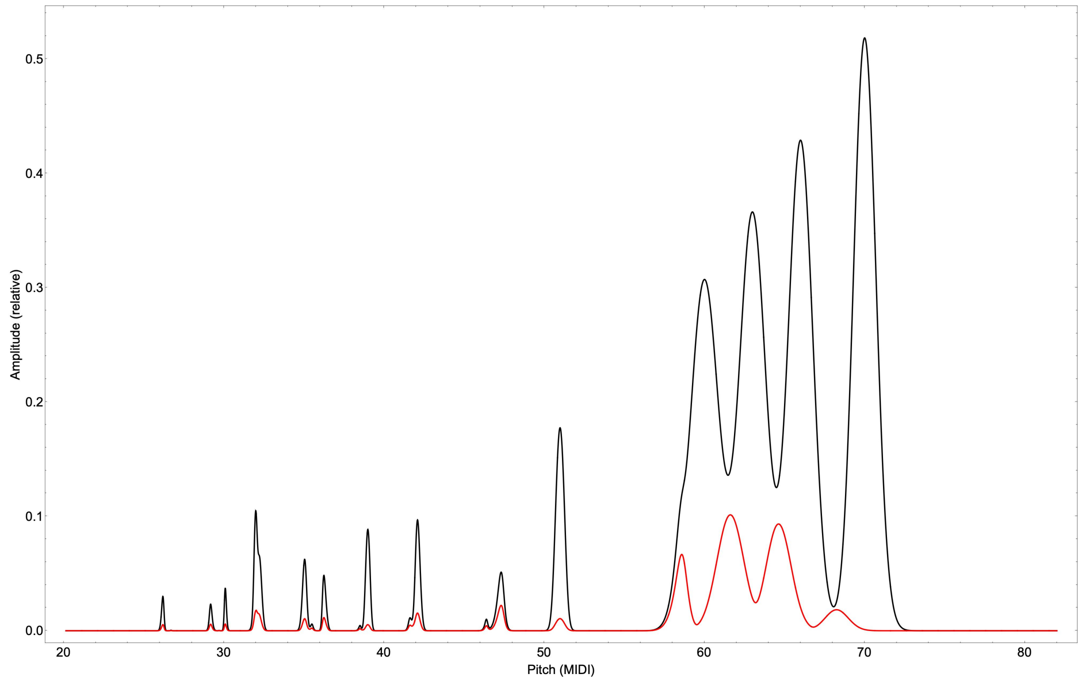

FIGURE 20. SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 63, 66, 70 HALF-DIMINISHED SEVENTH CHORD CONSISTING OF EQUAL AMPLITUDE PURE-TONE PITCHES.

As with the major and minor triad, the major-minor and half-diminished seventh chords share the same interval content, but the arrangement of those intervals produces very different virtual pitch fingerprints. The major-minor seventh has a prominent virtual pitch near MIDI 36, because its spectral pitches approximate the 4th, 5th, 6th, and 7th harmonics of that fundamental. This virtual pitch is the largest to result from all four spectral pitches jointly. The half-diminished seventh, by comparison, produces scattered virtual pitches near MIDI 32, 35, and 42. As with the major and minor triad, we should expect the major-minor seventh to sound more focused and unified than its half-diminished counterpart.

For a more extreme case, Figures 21 and 22 compare two chords, each built from equal amplitude pure tones: a MIDI {60, 72, 79.0196, 84, 87.8631} chord constructed from the first five harmonics of an overtone series, and a MIDI {60, 63.8631, 68.8435, 75.8631, 87.8631} chord constructed from the first five harmonics of an undertone series.

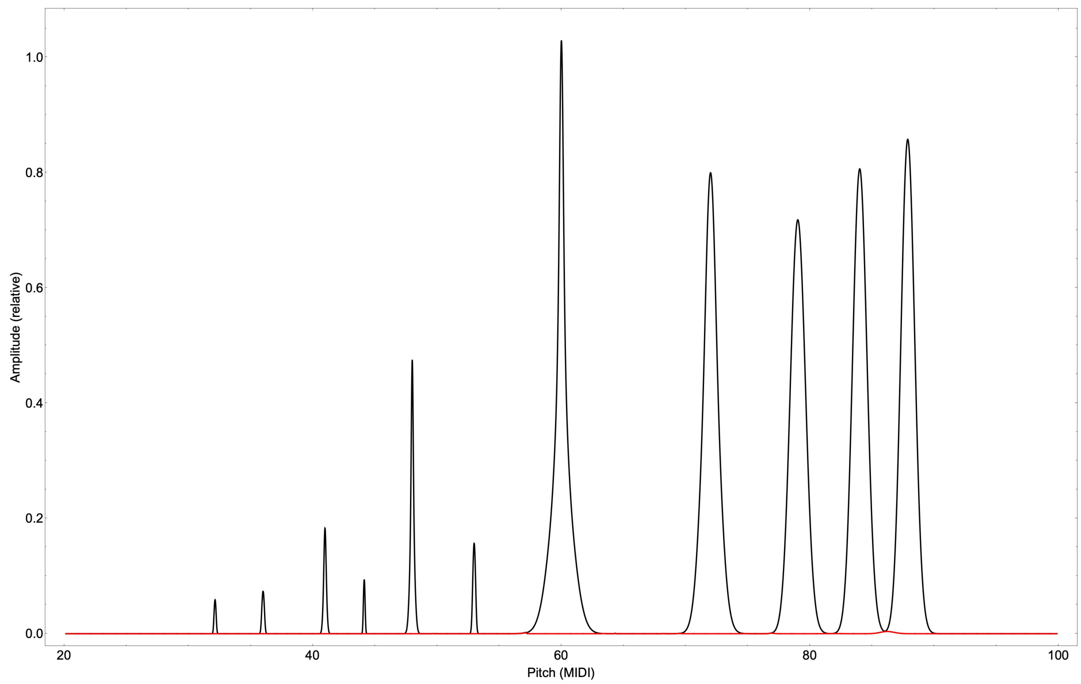

FIGURE 21. SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 72, 79.0196, 84, 87.8631 CHORD CONSISTING OF EQUAL AMPLITUDE PURE-TONE PITCHES.

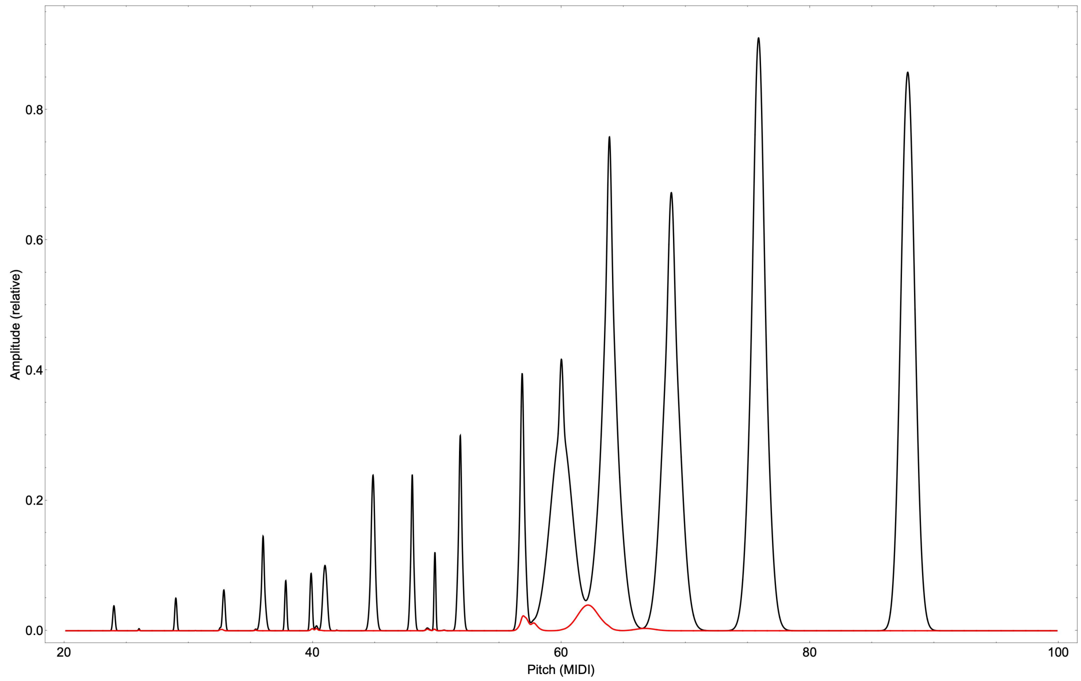

FIGURE 22. SST (BLACK) AND TV (RED) CURVES FOR A MIDI 60, 63.8631, 68.8435, 75.8631, 87.8631 CHORD CONSISTING OF EQUAL AMPLITUDE PURE-TONE PITCHES.

Figures 21 and 22 show that the overtone-based chord produces a single most prominent peak at MIDI 60, while the undertone-based chord spreads its energy across several smaller peaks. Despite having identical interval content, the overtone chord should therefore sound significantly more focused, unified, and less dissonant than its undertone counterpart.

Testable Predictions

A model of this kind earns its formalism by making claims that could turn out to be false. Several results derived above depart from existing accounts in ways that admit direct experimental test, and they are collected here.

P1. Roughness persists for pure-tone intervals outside the critical band, and compounds with the number of tones. Plomp and Levelt's account predicts that roughness vanishes once two pure tones are separated by more than a critical band[37]. The tonalness model predicts small but nonzero variation in the TV function at all separations, because any interval within a sonority can be heard as a mistuning of some nearby simpler interval, and it predicts that these residual variations accumulate as further tones are added. A discrimination experiment using widely spaced pure-tone sonorities would separate the two accounts directly. This is the model's clearest point of divergence from the standard roughness account.

P2. Inharmonic timbres cannot reach the dissonance minima available to harmonic timbres. Figure 5 shows that a stretched pseudo-octave built from an inharmonic spectrum produces a local dissonance minimum that never descends to the minimum available at a 2:1 octave in a harmonic spectrum. This is consistent with the results of McDermott, Lehr, and Oxenham[38], and is reported here as a recovered result rather than a new one.

P3. Sonorities with identical interval content differ in fusion according to their virtual-pitch distribution. Figures 21 and 22 show that a chord drawn from an overtone series and a chord drawn from an undertone series, having identical interval content, produce very different tonalness landscapes: a single dominant peak against several competing smaller ones. The model predicts the overtone chord will be rated more fused. The same prediction applies to the major/minor triad pair of Figures 17–18 and the seventh-chord pair of Figures 19–20.

---

37 Plomp and Levelt, "Tonal Consonance," 548–60.

38 McDermott, Lehr, and Oxenham, "Individual Differences," 1035–47.

P4. Dissonance ordering among intervals depends on register. Figures 6 and 7 show that the relative ordering of intervals by dissonance varies with register and with timbre.

Conclusions

Modeling virtual pitches illuminates many features of harmonic perception. Dissonance can be understood as an aggregate property of a harmony's virtual-pitch content, though a full account of why a particular harmony sounds as it does requires examining its unique virtual-pitch fingerprint. Timbre significantly shapes consonance and dissonance. Harmonic timbres yield patterns broadly consistent but more exaggerated than those of pure tones, while inharmonic timbres distort these patterns and tend to produce higher overall dissonance.

Tonalness theory addresses the immediate apprehension of harmony and harmonic progression, not the synoptic comprehension of complete musical works. It is not, and cannot be, a comprehensive theory of music. What rigorous mathematical modeling of harmonic perception does offer is consistency in the application of principles and assumptions, and with it, a basis for distinguishing harmonic perceptions rooted in musical culture from those that belong to the shared human auditory endowment.

Bibliography

Engell, Alva, Markus Junghöfer, Alwina Stein, Pia Lau, Robert Wunderlich, Andreas Wollbrink, and Christo Pantev. "Modulatory Effects of Attention on Lateral Inhibition in the Human Auditory Cortex." *PLOS ONE* 11, no. 1 (2016): 1–15.

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 Brian C. J. Moore. "Derivation of Auditory Filter Shapes from Notched-Noise Data." *Hearing Research* 47, no. 1–2 (1990): 103–38.

Harrison, Peter M. C., and Marcus T. Pearce. "Simultaneous Consonance in Music Perception and Composition." *Psychological Review* 127, no. 2 (2020): 216–44.

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

Hutchinson, William, and Leon Knopoff. "The Acoustic Component of Western Consonance." *Interface* 7, no. 1 (1978): 1–29.

Leman, Marc. "An Auditory Model of the Role of Short-Term Memory in Probe-Tone Ratings." *Music Perception* 17, no. 4 (2000): 481–509.

Lerdahl, Fred. *Tonal Pitch Space*. New York: Oxford University Press, 2001.

McDermott, Josh H., Andriana J. Lehr, and Andrew J. Oxenham. "Individual Differences Reveal the Basis of Consonance." *Current Biology* 20, no. 11 (2010): 1035–47.

Milne, Andrew J., William A. Sethares, Robin Laney, and David B. Sharp. "Modelling the Similarity of Pitch Collections with Expectation Tensors." *Journal of Mathematics and Music* 5, no. 1 (2011): 1–20.

Moore, Brian C. J. *Cochlear Hearing Loss*. London: Whurr Publishers, 1998.

Parncutt, Richard. "Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord." *Music Perception* 6, no. 1 (1988): 65–93.

———. *Harmony: A Psychoacoustical Approach*. Berlin: Springer-Verlag, 1989.

Penagos, Hector, Jennifer R. Melcher, and Andrew J. Oxenham. "A Neural Representation of Pitch Salience in Nonprimary Human Auditory Cortex Revealed with Functional Magnetic Resonance Imaging." *The Journal of Neuroscience* 24, no. 28 (2004): 6810–15.

Peretz, Isabelle, Anne J. Blood, Virginia Penhune, and Robert Zatorre. "Cortical Deafness to Dissonance." *Brain* 124, no. 5 (2001): 928–40.

Plomp, Reinier. "Pitch of Complex Tones." *Journal of the Acoustical Society of America* 41, no. 6 (1967): 1526–33.

Plomp, Reinier, and Willem J. M. Levelt. "Tonal Consonance and Critical Bandwidth." *Journal of the Acoustical Society of America* 38, no. 4 (1965): 548–60.

Rameau, Jean-Philippe. *Treatise on Harmony Reduced to Its Natural Principles*. Translated by Philip Gossett. New York: Dover, 1971.

Ritsma, Roelof J. "Frequencies Dominant in the Perception of the Pitch of Complex Sounds." *Journal of the Acoustical Society of America* 42, no. 1 (1967): 191–98.

Schellenberg, E. Glenn, and Sandra E. Trehub. "Frequency Ratios and the Discrimination of Pure Tone Sequences." *Perception & Psychophysics* 56, no. 4 (1994): 472–78.

Sethares, William A. "Local Consonance and the Relationship between Timbre and Scale." *Journal of the Acoustical Society of America* 94, no. 3 (1993): 1218–28.

———. *Tuning, Timbre, Spectrum, Scale*. 2nd ed. London: Springer, 2005.

Song, Peirun, Haoxuan Xu, Hangting Ye, et al. "Temporal Merging into Pitch with Click Train in the Macaque Auditory Cortex." *National Science Review* 12, no. 5 (2025): 1–13.

Terhardt, Ernst. "Calculating Virtual Pitch." *Hearing Research* 1, no. 2 (1979): 155–82.

Terhardt, Ernst, Gerhard Stoll, and Manfred Seewann. "Algorithm for Extraction of Pitch and Pitch Salience from Complex Tonal Signals." *Journal of the Acoustical Society of America* 71, no. 3 (1982): 679–88.

Tymoczko, Dmitri. *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice*. New York: Oxford University Press, 2011.

Vassilakis, Pantelis N. "Perceptual and Physical Properties of Amplitude Fluctuation and Their Musical Significance." PhD diss., University of California, Los Angeles, 2001.

Appendix

Standing wave model

We model a standing wave on a string of length L, fixed at both ends, using the one-dimensional wave equation with the boundary conditions 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.

Equation 3A gives the standard solution, where A is an arbitrary amplitude, h is an arbitrary positive integer harmonic, and the initial displacement is 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 of the wave, where ρ is the linear mass density. Because the kinetic energy equals the total energy at $t = 0$, Equation 5A gives the total energy of the 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.

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

EQUATION 5A. STANDING WAVE TOTAL ENERGY.

Solving Equation 5A for amplitude shows that A is inversely proportional to the harmonic number h. 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 6A. AMPLITUDE-HARMONIC PROPORTIONALITY.