2023

Relative Periodicity of Empirical Audio Samples With Application to Dissonance Perception

Linden Faye
lfaye1@pride.hofstra.edu

Spencer Kuhn
spencerkuhn@mail.adelphi.edu

Anil Venkatesh
Adelphi University, avenkatesh@adelphi.edu

Follow this and additional works at: https://digitalcommons.montclair.edu/laser-journal

Recommended Citation
Faye, Linden; Kuhn, Spencer; and Venkatesh, Anil (2023) "Relative Periodicity of Empirical Audio Samples With Application to Dissonance Perception," LASER Journal: Vol. 1: Iss. 1.
Available at: https://digitalcommons.montclair.edu/laser-journal/vol1/iss1/6
Relative Periodicity of Empirical Audio Samples With Application to Dissonance Perception
Linden Faye¹, Spencer Kuhn², and Anil Venkatesh³

Abstract

The concept of dissonance in music perception has been variously associated with the physical concepts of roughness, instability, and tension by appealing to a subjective cross-sensory analogy. For all the subjectivity around the term, it is noteworthy that non-musical test subjects in clinical experiments have produced remarkably consistent rank-orderings of musical sounds, according to the perceived dissonance of those sounds. Further studies in psychology, neuroscience, and mathematics show that consonance perception appears to be influenced not only by convention and culture but by the psychoacoustics of sound perception, likely by some combination of the Helmholtz theory of roughness and the theory of harmonicity originally contemplated by Galilei. In this work, we elaborate on Stolzenburg’s relative periodicity, a harmonicity metric that can be computed in the time domain by neuronal systems. We aim to determine whether relative periodicity can be extended from pure sinusoids to empirical audio signals, and whether such extension would still yield dissonance rankings that comport with the known experimental perceptions of dissonance. By drawing on the theory of quasi-periodic signals, we show that the relative periodicities of many familiar dyads and triads agree with Stolzenburg’s calculations and experimental dissonance rankings. However, we also observe substantial departures from Stolzenburg’s work in the case of the most highly empirically dissonant dyads and the diminished triads. This discrepancy likely supports the work of Masina et al. that found periodicity unable to fully explain empirical studies of dyad dissonance.

1 Introduction

The concept of dissonance in music perception has been variously associated with the physical concepts of roughness, instability, and tension by appealing to a subjective cross-sensory analogy. For all the subjectivity around the term, it is noteworthy that non-musical test subjects in clinical experiments have produced remarkably consistent rank-orderings of musical sounds, according to the perceived dissonance of those sounds. Further studies in psychology, neuroscience, and mathematics show that consonance perception appears to be influenced not only by convention and culture but by the psychoacoustics of sound perception, likely by some combination of the Helmholtz theory of roughness and the theory of harmonicity originally contemplated by Galilei. In this work, we elaborate on Stolzenburg’s relative periodicity, a harmonicity metric that can be computed in the time domain by neuronal systems. We aim to determine whether relative periodicity can be extended from pure sinusoids to empirical audio signals, and whether such extension would still yield dissonance rankings that comport with the known experimental perceptions of dissonance. By drawing on the theory of quasi-periodic signals, we show that the relative periodicities of many familiar dyads and triads agree with Stolzenburg’s calculations and experimental dissonance rankings. However, we also observe substantial departures from Stolzenburg’s work in the case of the most highly empirically dissonant dyads and the diminished triads. This discrepancy likely supports the work of Masina et al. that found periodicity unable to fully explain empirical studies of dyad dissonance.

¹lfaye1@pride.hofstra.edu, Undergraduate student at Hofstra University, Hempstead, New York, USA
²spencerkuhn@mail.adelphi.edu, Graduate student at Adelphi University, Garden City, New York, USA
³avenkatesh@adelphi.edu, Assistant Professor, Department of Mathematics and Computer Science, Adelphi University, Garden City, NY, USA

Published by Montclair State University Digital Commons, 2023
cross-sensory analogy. This concept has also been marked by affective terms such as unpleasant and sad. For all the subjectivity around the term, it is noteworthy that non-musical test subjects in clinical experiments have produced remarkably consistent rank-orderings of musical sounds, according to the perceived dissonance of those sounds [1]. Taking into account the numerous studies of this kind, the variety of plausible mathematical models, and the discovery of neuronal systems that are theoretically capable of realizing these models [2] [3] [4], it seems likely that consonance perception is influenced not only by convention and culture but by the psychoacoustics of tone perception [5] [6].

While many studies of dissonance perception center on categorizing or ranking musical intervals, chords, or sequence of notes, this approach risks eliding the influence of the partials, i.e. the numerous higher frequencies produced by any vibrating body, the superposition of which is perceived as a single sound by the human ear. Indeed, Sethares showed that subtle adjustments to the timbre of an instrument could cause the octave to sound out-of-tune, demonstrating that the partials can have a dramatic effect on the overall perception of dissonance [7]. Considering that the octave is empirically ranked as the most consonant musical interval with the possible exception of the non-interval of unison, Sethares’ result proves that the partials—their pitches and relative amplitudes—must be considered in any complete theory of dissonance perception. In spite of this, the majority of dissonance studies have been conducted using pure tones or familiar timbres (e.g. piano, clarinet), while dissonance perception in the context of unusual timbres or outright inharmonicity has only seen concerted investigation in recent times [8].

Indeed, the Helmholtzian [9] theory of dissonance as roughness which dominated the field until the late 20th century has ultimately struggled to account for the influence of timbre. This theory and its elaborations center on the sensation of roughness caused by the rapid amplitude modulation in the signal that occurs when tones of similar but distinct frequencies are superimposed. Cook and Fujisawa note that harmony is more than the summation of interval dissonance among tones and their partials [10], a principle that is central to the classical theory of Helmholtz. Further issues with the roughness theory arise in its treatment of rapid sequences of notes; since the theory rests on the superposition of clashing frequencies, perception of dissonance should not extend to arpeggiated chords, yet this is observed empirically in subjective preferences for certain melodies and scales [11] and in the processing of pitch and time sequences [12] [13]. It is clear, then, that a fully computational theory of dissonance perception must involve processing in the time domain.

The theory of dissonance as harmonicity is a primary contender to address this question. The premise of this theory, which has existed in some form since Galilei’s commensurability theory, is that the period of a superposition of musical notes can serve as a proxy for the perceived dissonance of that set of notes. More specifically, Stolzenburg defines the relative periodicity of a set of musical notes as the period of their superposition divided by the longest period of the individual notes [11]. He observes that this quantity is only well defined for sets of notes whose frequencies are all rational multiples of the lowest frequency; given that, an alternative definition of relative pe-
riodicity is the least common multiple of the denominators of those rational numbers in lowest terms. Quantifying dissonance in this way is referred to as harmonicity as it ultimately measures the extent to which the overtones collectively produce a harmonic sound wave [14]. Stolzenburg observes that the relative periodicity of a signal can be computed in the time domain by means of the autocorrelation of the signal, the dot product of the signal with a phase-delayed copy of itself. He notes not only that the period of the autocorrelation function is equal to the period of the original signal, but that the autocorrelation function peaks at the origin and achieves the same peak at each of its periods. Therefore, the task of measuring the period of an audio signal is reduced to locating the first maximum of the associated autocorrelation function. Stolzenburg’s model is biologically plausible due to the discovery of a neuronal autocorrelation mechanism that was shown to be capable of performing this computation for frequencies up to 1000 Hz [15] and which has more recently resulted in a full computational model of periodicity pitch detection [16]. Stolzenburg goes on to show that the relative periodicity metric comports well with empirical dissonance rankings of dyads and triads, lending credence to the position that dissonance perception in the musically important low frequency range is dominated by a time-domain autocorrelation process. This finding is further bolstered by Trulla et al. who achieved the same ranking of dyads by recurrence quantification analysis [17].

The debate over roughness versus harmonicity in dissonance perception is far from over. Eerola and Lahdelma analyzed many leading dissonance metrics spanning roughness, harmonicity, and familiarity to produce a linear mixed model that achieved a modest prediction rate of 73% on empirical data, with familiarity—not roughness or harmonicity—dominating the prediction [18]. Masina et al. found that harmonicity models alone were insufficient to explain empirical studies of dyad dissonance and investigated linear interpolations between roughness and harmonicity models in an attempt to close this gap [19]. Masina and Lo Presti went on to show that a linear interpolation of roughness and harmonicity indicators yielded improved alignment with triad empirical dissonance rankings as well, although with some caveats relating to the major and minor triads [20]. These studies offer a kind of upper bound on the extent to which harmonicity can explain dissonance perception since the linear models they present have no known biological basis; moreover, both studies adopt the limitation of rational frequency relationships in their harmonicity metrics, further idealizing their findings compared to the empirical data. Friedman et al. explored a different set of rational frequency multiples by generating chord stimuli according to the Bohlen-Pierce chromatic just scale which divides a tritave (an octave plus a perfect fifth) into 13 intervals based on a series of odd integer frequency ratios. This nonstandard scale allowed them to present unfamiliar tone combinations to the study participants in order to gauge the influence of harmonicity without familiarity as a confounding variable. They found that harmonicity had a positive association with consonance ratings, yet highlighted the likely influence of roughness on participants’ dissonance perception [14].

This work aims to answer the following question: can we extend the concept of relative periodicity to empirical audio signals, complete with non-rational frequency
relationships, dozens of partials, and inharmonicity? As a corollary to this question, we wish to determine whether any such extension of relative periodicity would still comport with empirical dissonance rankings of common musical intervals and chords. In Section 2, we lay out the numerous technical obstacles associated with these questions, and our proposed methods of addressing these obstacles. In Section 3, we present our computationally derived estimates of the relative periodicity of common dyads and triads and compare our results with Stolzenburg’s theoretical values of relative periodicity. In Section 4, we discuss our findings and present evidence that relative periodicity is a strong candidate for dissonance perception in major and minor triads but may be insufficient to account for the empirical dissonance of certain dyads and diminished triads. Our conclusions regarding dyads reinforce the recent work by Masina et al. which found that periodicity was insufficient to explain empirical studies of dyad dissonance [19], as well as Masina’s recent work that found periodicity to be an effective but not fully satisfactory indicator of triad consonance [21] [20]. Moreover, our model lays the groundwork for future investigations of the relative periodicity of inharmonic sound such as that found in the contemporary spectral music tradition [8].

2 Methods

Given a superposition of pure tones that are all rational multiples of a certain fundamental, Stolzenburg’s relative periodicity is computed as the least common multiple of the denominators of those rational numbers (in lowest terms). This definition does not immediately extend to chords in an equal tempered scale since notes in such a scale do not generally have a rational relationship to each other. The standard treatment is to equate each note to a convenient nearby just interval, a substitution for which Stolzenburg contends there may exist some biological basis, in spite of the fact that several notes in twelve-tone equal temperament deviate from the nearest familiar just interval by more than the just noticeable difference of 1% Hz [11]. We approach the problem of non-rational frequency multiples in a more concrete way: since Stolzenburg shows that the relative periodicity of a signal can also be computed via the location to the first maximum of the associated autocorrelation function, we compute the autocorrelation of the signal and attempt to locate this maximum.

In this section, we lay out our approach to computing relative periodicity directly from synthesized piano audio samples. We used 1,020 synthesized piano dyads spanning eight octaves from the jazznet data set [22] and 432 synthesized piano triads from the Piano Triads Wavset [23] containing major, minor, and diminished chords in their standard forms and their first inversions across six octaves. Each file contains three seconds of audio sampled at a rate of 44.1KHz for triads and 16KHz for dyads. Figure 1 shows the waveform of the major triad with A4 root note.

Initially, we were concerned that the timbre of the audio sample might greatly influence the computation of relative periodicity, as Stolzenburg only carries out a simplified treatment of the question on an unweighted sum of three sinusoids. However, we determined that the autocorrelation of a weighted sum of sinusoids follows a simi-
Figure 1: This figure displays the waveform of the major triad with root note A4 as found in the Piano Triad Wavset.

Harmonic pattern to the simplified case and therefore conclude that the influence of timbre on harmonicity must be situated deeper in the process than at the stage of autocorrelation. For completeness’ sake, we include below a sketch of the derivation of this finding.

Let $s(t)$ be the sum of periodic signals of different frequencies and varying amplitudes and $\rho(\tau)$ be the autocorrelation of $s(t)$, as follows:

$$s(t) = a_1 \sin(\omega_1 t) + a_2 \sin(\omega_2 t) + \cdots + a_n \sin(\omega_n t)$$

$$\rho(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} s(t)s(t - \tau)dt.$$

Expanding the integrand, we observe that the integral over any full period of $s(t)$ is zero except for those terms in which $i = j$:

$$s(t)s(t - \tau) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \sin(\omega_i t) \sin(\omega_j (t - \tau)).$$

Since the autocorrelation function is computed as a limit over arbitrarily many periods of the signal, the terms of the integrand for which $i \neq j$ are dominated by $\frac{1}{2T}$, leaving the following identity:

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} s(t)s(t - \tau)dt = \sum_{i=1}^{n} a_i^2 \cos(\omega_i \tau).$$

Similar to Stolzenburg’s proof that the autocorrelation of unweighted periodic signals converges to a sum of cosines, the autocorrelation of weighted periodic signals converges to a sum of cosines weighted by the squares of their amplitudes. The resulting weights do not prevent the autocorrelation function from being periodic with the same period as the signal, so we can conclude that any effect of timbre on harmonicity does not manifest at the level of autocorrelation.
When extending Stolzenburg’s model of relative periodicity to empirical audio samples, several factors cause the autocorrelation of the signal to be quasi-periodic instead of periodic: the sampling rate of the audio, the inharmonic attack, and the attenuation of the sustained waveform. To overcome these obstacles, we must determine the best candidate for the period of a function that is not exactly periodic in the first place. In order to compute the relative periodicity of empirical audio samples, we have developed a novel approach to estimating the periods of such quasi-periodic signals.

The first step in our process was to trim the sample to limit the influence of attack and attenuation. For dyads, the audio was trimmed to the interval between 0.6 and 0.85 seconds for a total of 0.25 seconds of data, whereas for triads the audio was trimmed to the period between 1.4 and 1.9 seconds for a total of 0.5 seconds of data. Trimming windows were chosen qualitatively based on the perceived interval of sustained sound after the conclusion of the attack. Each chord’s fundamental period was set according to the theoretical value for its root note under the assumption that the piano was accurately tuned. Figure 2 shows the trimmed audio sample associated with the A4 major triad.

![Trimmed Piano Triad Audio: A4 Major](image)

Figure 2: This figure displays the trimmed audio sample associated with the A4 major triad.

The autocorrelation of the trimmed sample was then computed numerically as a function of phase delay ranging across half the width of the trimming window with a step size equal to the reciprocal of the sampling frequency. Figure 3 shows the autocorrelation function observed for the A4 major triad.

The attenuation of the original signal and the finite sampling rate prevent the autocorrelation from returning to its value at $\tau = 0$, resulting in the quasi-periodicity clearly visible in Figure 3. However, it is also clear to the naked eye that the analogous value of tau lies around 0.02, where the graph exhibits its first pulse or “heartbeat” feature. We have developed three different methods for estimating the location of this pulse in the autocorrelation function. Each method is described below with application to the A4 major triad. We used these methods to produce a total of five estimates of relative periodicity for each of the 1,020 dyads and 432 triads in our data set.
First Method: Simple Maximum

The first method estimates the period of $\rho(\tau)$ by the $\tau$-value corresponding to the function’s absolute maximum, starting at $\tau$ equal to the fundamental period of the signal. Since the lowest possible relative periodicity possible is 1.0, a maximum within the first fundamental period would not be informative. For the A4 major triad, this estimate is shown with a dotted red line in Figure 4.

Second Method: Smoothed Maximum

For the second method, the function’s positive values were averaged over a sliding window in order to provide a smoothing mechanism that weakened the influence of random fluctuations in the autocorrelation. We produced two estimates in this method: one with the window size equal to twice the fundamental period and one...
with the window size equal to the fundamental period. For the A4 Major triad, Figure 5 shows the estimate obtained using the wider window option and the resulting period estimate.

Figure 5: This figure illustrates the smoothed maximum method (2x window version) for estimating relative periodicity, applied to the autocorrelation of the A4 major triad.

Third Method: Low Pass Relative Extrema

For the third method, we first identified every value of \( \tau \) at which \( \rho(\tau) \) achieved a relative maximum compared to its values within a fixed window (with a width either equal to the fundamental period of twice this value). Our estimate for the period of the autocorrelation was then set to the location of the first relative maximum that itself was greater than the immediately preceding and succeeding maxima. This method was specifically designed to account for attenuation in autocorrelation that was particularly observed in the higher octaves. Figure 6 shows the result of this smoothing for the A4 Major triad (using the wider window option) as a line plot in blue, with the period estimate marked with a red dotted line.

3 Results

All five methods of relative periodicity estimation yielded very similar results for dyads, which was unsurprising due to the much simpler geometry of the autocorrelation function of two notes compared to three. For the sake of brevity, we therefore report our findings for dyads only in terms of the first method (simple maximum). By examining each of the twelve dyadic intervals across every root note, we visualize each interval’s relative periodicity across seven octaves in Figures 7–9. Table 1 displays for each dyad Stolzenburg’s computed relative periodicity and the median relative periodicity that we observed in our study.

For the seven most empirically consonant intervals, our relative periodicity estimates yielded very close to twice the theoretical values in Stolzenburg’s work, conse-
Figure 6: This figure illustrates the relative maxima method (2x window version) for estimating relative periodicity, applied to the autocorrelation of the A4 major triad.

<table>
<thead>
<tr>
<th>Dyad Interval</th>
<th>Smoothed Relative Per.</th>
<th>Method 1</th>
<th>Dyad Interval</th>
<th>Smoothed Relative Per.</th>
<th>Method 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>octave</td>
<td>1.0</td>
<td>1.9834</td>
<td>minor 3rd</td>
<td>5.0</td>
<td>9.9846</td>
</tr>
<tr>
<td>perfect 5th</td>
<td>2.0</td>
<td>3.9774</td>
<td>tritone</td>
<td>6.0</td>
<td>9.8979</td>
</tr>
<tr>
<td>perfect 4th</td>
<td>3.0</td>
<td>5.9779</td>
<td>minor 7th</td>
<td>7.0</td>
<td>7.9232</td>
</tr>
<tr>
<td>major 3rd</td>
<td>4.0</td>
<td>7.9289</td>
<td>major 2nd</td>
<td>8.5</td>
<td>8.0115</td>
</tr>
<tr>
<td>major 6th</td>
<td>3.0</td>
<td>5.9369</td>
<td>major 7th</td>
<td>8.0</td>
<td>2.0246</td>
</tr>
<tr>
<td>minor 6th</td>
<td>5.0</td>
<td>9.9418</td>
<td>minor 2nd</td>
<td>15.0</td>
<td>1.9455</td>
</tr>
</tbody>
</table>

Table 1: This table displays the twelve intervals in increasing order of empirical dissonance along with Stolzenburg’s computed relative periodicity and the median relative periodicity that we observed in our study.

Subsequently giving the same rank-ordering by dissonance. This pattern did not extend to the tritone, minor 7th, or major 2nd, which were surprisingly placed between the major 3rd and minor 3rd in terms of dissonance. Stranger still, the major 7th and minor 2nd obtained an estimated relative periodicity quite close to that of the octave in spite of having much greater empirical dissonance.

Examining the estimated relative periodicity of each dyad across the entire range of root notes, we observe in Figures 7–9 that most dyads exhibited rather consistent measurements in all but the highest octave in which the attack-to-sustain ratio was high, leading to extreme inharmonicity of the samples. The most notable exception is the major 2nd, whose median level appears to undershoot the plurality value of approximately 19 (notably, around twice the value that Stolzenburg predicted for this dyad). Among the empirically more dissonant intervals, some notable fluctuation in measurement is observed in the minor 6th, major 7th, and especially the minor 7th.

For triads, we found that the two variants of method 3 yielded very similar results so we have discarded the variant of method 3 that uses the wider window option in

Published by Montclair State University Digital Commons, 2023
Figure 7: This figure displays the relative periodicity estimates across all root notes for the intervals of minor 2nd, major 2nd, minor 3rd, and major 3rd as found in the jazznet data set.

For the interest of brevity. Figures 10–12 display the four remaining relative periodicity estimates applied to each chord shape (no inversion) with root note ranging across seven octaves. Table 2 displays all periodicity results for triads, including first inversion chords.

<table>
<thead>
<tr>
<th>Triad</th>
<th>Inversion</th>
<th>Smoothed Relative Per.</th>
<th>Method 1</th>
<th>Method 2 (1x)</th>
<th>Method 2 (2x)</th>
<th>Method 3 (1x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major</td>
<td>{0, 4, 7}</td>
<td>4.0</td>
<td>7.9551</td>
<td>7.8308</td>
<td>7.7037</td>
<td>7.9551</td>
</tr>
<tr>
<td></td>
<td>{0, 3, 8}</td>
<td>5.0</td>
<td>10.0131</td>
<td>10.4686</td>
<td>9.8398</td>
<td>10.0131</td>
</tr>
<tr>
<td>Minor</td>
<td>{0, 3, 7}</td>
<td>10.0</td>
<td>31.9487</td>
<td>20.9597</td>
<td>31.6292</td>
<td>31.9487</td>
</tr>
<tr>
<td></td>
<td>{0, 4, 9}</td>
<td>12.0</td>
<td>37.8228</td>
<td>20.2975</td>
<td>33.2169</td>
<td>37.8228</td>
</tr>
<tr>
<td>Diminished</td>
<td>{0, 3, 6}</td>
<td>17.0</td>
<td>9.9463</td>
<td>9.4365</td>
<td>9.7212</td>
<td>9.9509</td>
</tr>
<tr>
<td></td>
<td>{0, 3, 9}</td>
<td>15.3</td>
<td>10.0154</td>
<td>10.7119</td>
<td>11.1996</td>
<td>10.0154</td>
</tr>
</tbody>
</table>

Table 2: This table displays Stolzenburg’s computed relative periodicity and the four median relative periodicity estimates we observed in our study (for minor triads, third quartile replaces median).
Figure 8: This figure displays the relative periodicity estimates across all root notes for the intervals of perfect 4th, tritone, perfect 5th, and minor 6th as found in the jazznet data set.

**Major Triads**

Among major triads, the relative periodicity estimates of method 1 and method 3 yielded highly consistent results with many estimates clustered around 7.9551, especially for root notes between the third and sixth octaves. In fact, the only significant variations in periodicity occurred within the second and seventh octaves, although there were a few exceptions within the fourth octave. Method 2 yielded similarly consistent results for both window sizes although these plots show noticeably higher variance by root note, even within the interior octaves. The median periodicities for this method were 7.7037 and 7.8308 for the smaller and larger windows, respectively, a slight departure from the identical medians found for the other two methods. Figure 10 displays our findings for major triads. For first inversions of the major chords, the same respective patterns were observed for each method, although median periodicities were found consistently around 10.0 instead of 7.9.

**Minor Triads**

Relative periodicity estimates for minor triads varied much more widely than for major triads, although within the interior octaves, several quantized groups of similar measurements can be seen from the scatter plots. Since median values often failed to capture the largest plurality of measurement groups, the third quartile value was
used to capture a more representative summary of the data. Thus, for method 1 a third quartile of 31.9487 was observed, which was again identical to the low pass relative extrema method’s third quartile. For method 2, the observed third quartiles were 31.6292 and 20.9597 for the smaller and larger windows, respectively, again indicating more variance among estimates for that method. Figure 11 displays our findings for minor triads. The first inversions of minor chords consistently demonstrated similar or higher periodicity estimates, with third quartile values of 37.8228, 33.2169, 20.2975, and 37.8228 for method 1, method 2 (1x), method 2 (2x), and method 3 (1x), respectively.

**Diminished Triads**

Among diminished triads, the relative periodicity estimates were more consistent than those of the minor triads, with a median value 9.9463 for method 1 and medians of 9.7212, 9.4365, and 9.9509 for method 2 (1x), method 2 (2x), and method 3 (1x), respectively. Method 2 again exhibited more variation than the other two methods. As with the major triads, measurements in the second and seventh octaves were far less consistent than those in the interior octaves, with some periodicities differing by an entire order of magnitude or more. This inconsistency is expected due to increasing inharmonicity of the audio samples in the extreme octaves. Figure 12 displays our findings for diminished triads. Much more variation was observed among the first inversions of diminished triads, with quantized groups of measurements emerging that resembled
Figure 10: This figure displays the four relative periodicity estimates across all root notes for the major triad interval as found in the Piano Triads Wavset.

Figure 11: This figure displays the four relative periodicity estimates across all root notes for the minor triad interval as found in the Piano Triads Wavset.

the patterns seen among minor triads. The medians observed among the four reported methods were 10.0154, 11.1996, 10.7119, and 10.0154, respectively.
Figure 12: This figure displays the four relative periodicity estimates across all root notes for the diminished triad interval as found in the Piano Triads Wavset.

4 Discussion

In this work, we have attempted to extend the concept of harmonicity from the realm of non-attenuating pure sinusoids to that of empirical audio samples. Our aim has been to gauge the extent to which the harmonicity theory of dissonance can predict empirically perceived dissonance when applied to real waveforms as opposed to mathematically idealized sound. While the classical definition of harmonicity cannot be extended in this way, Stolzenburg offers a possible solution by means of the period of the autocorrelation function [11]. However, his work does not trivially generalize to the case of empirical audio samples because such waveforms yield merely quasi-periodic autocorrelation functions due to attenuation and other factors. Therefore, we have developed a variety of methods for producing a principled estimate of the period of a quasi-periodic autocorrelation function and applied these methods to many hundreds of synthesized piano dyads and triads. For the vast majority of dyads and triads, we observed a high level of measurement consistency independent of choice of root note, with the exception of root notes in the highest and lowest octaves. Since the audio samples in those ranges exhibited a high level of inharmonicity, this was expected.

In our analysis of 1,020 dyads from the jazznet data set [22], we found that the seven most empirically consonant intervals showed a high degree of agreement between Stolzenburg’s smoothed relative periodicity and our estimates, which were consistent across a wide range of octaves. Conceivably, this finding also holds for the major 2nd, for which the median estimate appeared to miss the plurality value which would have
been consistent with Stolzenburg’s calculation. However, we did observe a substantial unexplained departure between Stolzenburg’s findings and ours with the tritone and minor 7th, whose relative periodicity we estimated substantially lower than empirical results would dictate. Additionally, we observed a massive unexplained departure from the empirical dissonance rankings in our estimates for the major 7th and minor 2nd, both of which received a comparable score to the octave! In the case of these latter two dyads, we believe that the error is attributable to the attenuation of the autocorrelation function outweighing the eventual idealized peak. The ideal peak location is expected after 15 (or perhaps 30) “heartbeats” of the autocorrelation function, which for C4 is around 57 ms (or perhaps 114 ms). By contrast, our window of 250 ms for dyads is reduced to an interval of 125 ms for the autocorrelation function. It is therefore likely that for root notes below C4, the window of 250 ms is too short to obtain a periodicity estimate for the major 7th and minor 2nd. Regardless of this limitation, we observe no better performance for root notes about C4, not even with method 3 which was specifically designed to detect peaks among relative maxima. Instead, the first “heartbeat” was consistently identified as the period of the autocorrelation function, resulting in similar performance to the octave. The fact that for these dyads the idealized peak of the autocorrelation function evidently cannot be observed suggests that their high empirical dissonance is not exclusively based on a time-domain judgment. In sum, our investigation of relative periodicity of dyads supports the previous work of Masina et al. who found that harmonicity alone cannot explain the empirical dissonance perception of dyads [19] [20]. We speculate that cultural familiarity plays a particularly strong role in dyad dissonance perception since the relatively sparse spectrum contains much less roughness than a typical triad; in this sense, we concur with Eerola and Lahdelma on the prominent role of familiarity in dissonance perception, at least for dyads [18].

Our analysis of the 432 triads in the Piano Triad Wavset [23] yielded similarly mixed results when compared to Stolzenburg’s theory. For the major triad, we observed highly consistent results across many octaves that yielded almost exactly double the value predicted by Stolzenburg. While we have no rigorous interpretation of this doubling phenomenon, we observe that this same pattern was evident with the seven most empirically consonant dyads. We therefore conclude that Stolzenburg’s theory of relative periodicity via autocorrelation was successfully extended to empirical audio samples in the cases of those seven dyads and the major triad (including its first inversion).

Our findings with the minor triad are much less clear as we obtained relative periodicity estimates ranging between 1.7 and 3.2 times Stolzenburg’s values with much less clustering of the estimates across methods than was observed with the major triad, and much more variation by root note as well. While our findings clearly support the empirically established fact that minor triads are more dissonant than major triads, it is difficult for us to contextualize our results beyond this simple observation. In fact, the autocorrelation functions of minor triad waveforms were highly variable in their geometry and more work is needed in order to consistently measure the autocorrelation periods.

Lastly, our analysis of diminished triads poses a different type of interpretive chal-
challenge. Much akin to the major triads, the diminished triads exhibit rather orderly estimates across root note choice, clustering notably around their median values over many octaves. Unlike with the major triads, however, the resulting median estimates for relative periodicity depart substantially from the emergent doubling rule that was observed not only with major triads but with seven of the dyads as well. In fact, our findings for diminished triads would suggest that they are roughly as consonant as major triads in the first inversion, a finding that substantially breaks with empirical dissonance data. Simply put, we cannot offer a convincing explanation for this finding and more research is required to make sense of it. While progress on autocorrelation period measurement would surely shed some light on this question, the consistency of our measurements across root note choice suggests the presence of a systemic phenomenon.

The limitations of applying Stolzenburg’s theory to empirical audio samples were largely overcome by our study, especially with regard to major triads and the most common dyads. Our results serve as a proof of concept for quantifying the harmonicity of arbitrary empirical audio by means of studying the quasi-periodic autocorrelation of the signal. Moreover, our analysis of the most highly empirically dissonant dyads comports with recent literature on the shortcomings of harmony and the role of cultural familiarity in dissonance perception. However, there are major unresolved questions with our approach, particularly with the tritone and minor 7th dyads and the diminished triads. While advances in the study of quasi-periodic autocorrelation functions will surely help to unravel the mysteries of those chords, our findings suggest the presence of some other systemic phenomenon, whether based on roughness or something else entirely.

References


