Unraveling Harmony: an Analysis of the Evolution of Temperament Systems and a Proposal of a Mathematical Framework for Musical Consonance

A MUSC 115 Final Research Paper

Instructor: Prof. Rotem Gilbert

Written by Qilin Ye

May 4, 2023

1 Introduction

At the core of musical expression, harmony serves as the bedrock that supports the intricate interplay of melodies and emotions. These complex relationships between notes weave a rich tapestry that enables composers to convey their artistic intent. Throughout history, musicians and theorists have endeavored to understand and achieve this harmony by developing a variety of temperament systems. As new systems emerged, perceptions of what constituted "harmony" evolved; for example, the perfect fifth interval has always been considered harmonious, but it took several centuries for the major third and major sixth intervals to gain widespread acceptance among composers (Tenney 1988). This essay will trace the progression of these systems, examining early tuning methods such as Pythagorean Tuning and Just Intonation before delving into the contemporary prevalence of Equal Temperament. To further demystify harmony, this paper will introduce a mathematical framework that encapsulates its essential aspects, applying it to explain specific harmonic progressions and resolutions. Lastly, this paper will touch upon the enduring debate surrounding Equal Temperament, which has persisted since its inception.

2 History of Temperament Systems

2.1 Pythagorean Tuning (PT)

One of the earliest attempts to rigorously quantify musical intervals was by Pythagoras, where he observed that the ratio of frequencies of two notes that form an octave is roughly 2:1. Similarly, the ratio for a perfect fifth is 3:2, perfect fourth 4:3, and whole tone 9:8 (Barker 1989). In 6th century A.D., Roman philosopher Boethius realized that five whole tones overshoot an octave, and six tones overshoot, but dividing a whole tone in half would lead to a geometric mean in frequencies, inducing an irrational proportion (Boethius 1989). To tackle this issue, Boethius subdivided a whole tone into a "major semitone" and a "minor semitone."

Unfortunately, this technique, known as *Pythagorean Tuning*, has a few drawbacks. To implement this method, one iteratively increase or decreases the frequency by a factor of 1.5 to create a complete circle of fifths, and multiply or divide by 2 to adjust for octaves. Some simple mathematical calculation shows that enharmonically equivalent notes can have different frequencies. Assuming the frequency of D is x Hz, by traversing downward in the circle of fifths via D $\mapsto$ G $\mapsto$ C $\mapsto$ ... $\mapsto$ ♭E $\mapsto$ ♭A, we obtain a low ♭A. Normalizing it into the center octave, the frequency ratio therefore becomes $(2/3)^6 \cdot 2^4 \approx 1.405$. On the other hand, by traversing upward and obtaining D $\mapsto$ A $\mapsto$ E $\mapsto$ ... $\mapsto$ ♯C $\mapsto$ ♯G, after normalizing, we obtain a frequency ratio of $(3/2)^6 \cdot 2^{-3} \approx 1.424$. This means pairs of enharmonic notes are not well-defined! And consequently, all notes are not well-defined. (Starting with D, we get an ill-defined ♯G/♭A, so starting with C, we can end up with an ill-defined enharmonic ♯F, and so on.) This slight difference, called the *Pythagorean comma*, is precisely the difference between a major semitone and a minor semitone. Furthermore, while

Pythagorean tuning ensures that fifths maintain a perfect 3/2 frequency ratio (and consequently, a perfect 4/3 for fourths), the thirds are represented by a much uglier fraction: 81/64.

2.2 *Just* Intonation (JI)

Pythagorean Tuning maintained its popularity until the 16th century, when Nicola Vicentino (1511-c.1576) invented *Just* Intonation. Instead of maintain a fixed ratio between certain intervals, Vicentino manually assigned a ratio between the frequency of any note and that of the tonic. He followed Pythagoras' idea and assigned simple fractions to whole tone, thirds, fourths, fifths, and octaves, and additionally set the frequency ratio of a perfect thirds to be 5/4. For other intervals, he assigned fractions based on a series of multiplication of his existing fractions. For example, a major seventh can be broken down into a major third plus a perfect fifth, so he sets the frequency of a major seventh to 5/4 * 3/2 = 15/8.

In addition to fixing the issue of a somewhat dissonant thirds, Just Intonation is also invariant under modulation, so a piece of music would maintain its harmony even when played under a different key. Within the same key, however, translations of identical chords may create different effects. For example, setting C as tonic, both C-D and D-E are whole tones, but the frequency ratio between C-D is 9/8 whereas that for D-E is 10/9, leaving a clearly audible 25% difference.

Previous literature has pointed out one similarity between the two tuning systems — both are results of human prejudice. Boethius and Vicentino manipulated the fractions so that what they considered harmonious had a simple fraction, and what they considered dissonant had much more complex ones. Back in the 6th century, thirds and sixths were considered dissonances, which might justify why he only assigned a nice proportion to the fifths. This could very well explain why Vicentino assignsed a stunning 45/32 to a tritone, while this ugly-looking ratio is in fact only 0.4% away from a much simpler 7/5 (Pesic 2010).

2.3 Equal Temperament (ET)

To tackle with the limitations of the previous temperament systems, a few decades later, scholars from around the world — Chinese musicologist Zhu Zaiyu in 1596, French philosopher Marin Mersenne in 1636, among others — would independently invent an Equal Temperament tuning system. In this system, notes are tuned in terms of *cents*, a logarithmic unit used to measure musical intervals. An octave, where the upper note is twice the frequency of the lower note, is defined to be 1200 cents, and each semitone is measured at 100 cents. Therefore, in each semitone step, the frequency increases by a factor of $2^{1/12} \approx 1.059$. Since this definition applies to any particular key, it naturally enjoys the privilege of invariance under modulation: the relationships between intervals are consistent regardless of the key in which the music is being played. Furthermore, since the frequency of each note is unambiguously defined, the issue of Pythagorean comma naturally resolves.

Previous tuning systems have all assumed that a nice-sounding interval needs to maintain a frequency ratio representable by sufficiently nice, simple fractions. ET, on the other hand, breaks this tradition, as it is well-known that any power of $2^{1/12}$ other than a multiple of 12 results in an irrational number. Yet, in 1988, Joos Vos conducted a series of experiments on 24 professional musicians, in which they were asked to rate their impression on music snippets generated based on different tuning schemes. Using regression analysis, Vos concluded that ET was among the few most positively rated temperament systems (Vos, 1988). Vos also pointed out that ET's enharmonic consistency as well as streamlined manufacturing and tuning framework helped it quickly accumulate

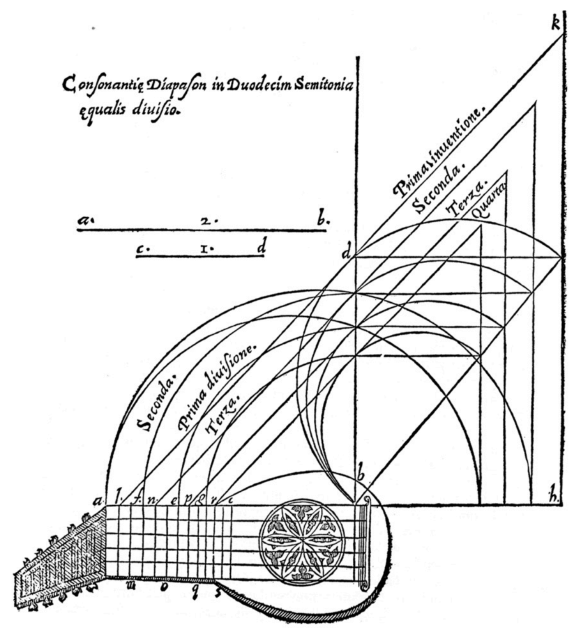

Figure 1: An illustration of the geometric derivation of a variant of ET (Zarlino 1588). Caption translates to "*the equal division of the consonance of a diapason into twelve semitones.*"

popularity, eventually becoming one of the most widely accepted tuning scheme today.

3 Quantifying Harmony via Rational Approximation of Frequency Ratio

3.1 The Simple Fraction Proximity Coefficient

Historically, it is believed that a consonance between two notes happens when the frequency ratio is a simple fraction, for example, 2:1 for an octave and 3:2 for a perfect fifth, as observed by Pythagoras. Intuitively, this makes sense: if the sound waves of two notes share a rather common short cycle, then their superimposed sound wave still exhibits a simple, short cycle. In this case, human ear would perceive the combined sound as pleasing and harmonious, as they create a highly periodic sound wave. In contrast, a dissonance occurs when the frequency ratio is a more complex fraction, or even worse, irrational. In this case, the resulting superimposed sound wave either display less periodicity due to longer period or is simply aperiodic.

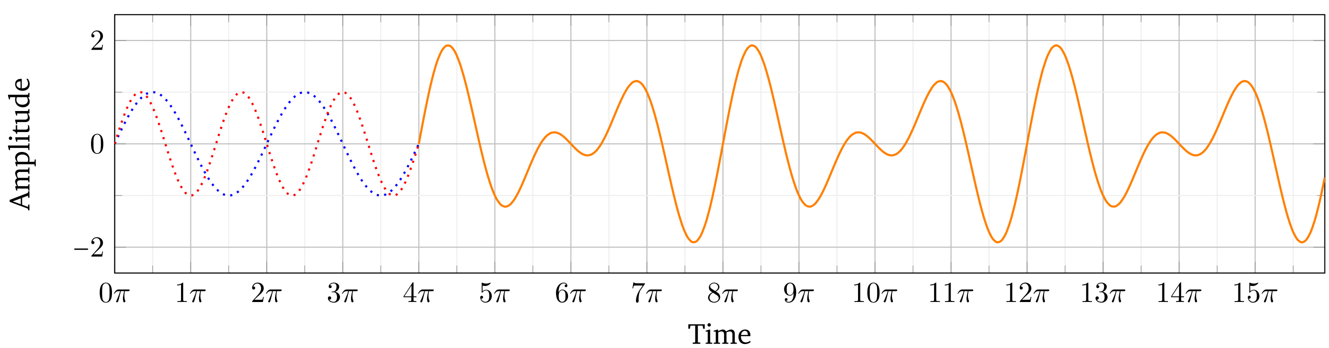

Figure 2: Plots of $y = \sin x$ [dotted blue], $\sin(1.5x)$ [dotted orange], and $\sin x + \sin(1.5x)$ [solid].

However, the irrational frequency ratios used in ET suggest that two sound waves do not need to have a perfectly rational proportion in order to sound harmonious. After all, human ears are not designed to perfectly identify the exact frequency of a sound. Therefore, it is more natural to introduce a margin of error, say 1%. Extending our previous observations, a two-note interval can be considered “nice” if its frequency ratio is *sufficiently close* to a nice, simple fraction. These settings are formalized in the following definition:

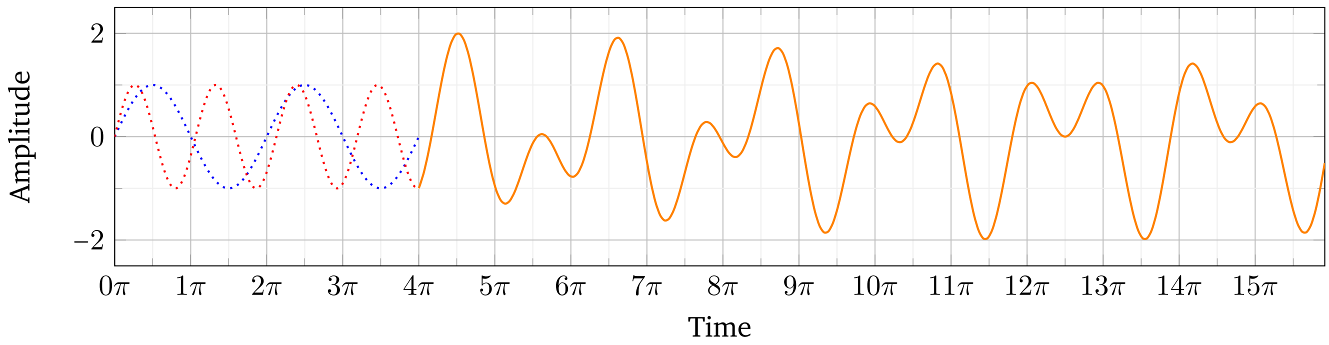

Figure 3: Plots of $y = \sin x, \sin(1.88x)$, and $\sin x + \sin(1.88x)$. Note that the superimposed wave does not display any particular periodic pattern, unlike in the previous figure where the period is 4π.

Definition: Simple Fraction Proximity Coefficient

Define $\mathcal{S} := \bigcup_{1\leqslant n\leqslant 10} \{m/n : 1 \leqslant m/n \leqslant 2\}$ to be the set of "nice, simple" fractions inside $[1,2]$. For any real number $1 \leqslant x \leqslant 2$, we define its **simple fraction proximity coefficient** to be

$$\rho(x) := \min_{q\in\mathcal{S}} d(q,x) \qquad \text{where } d(a,b) := (\max(a,b)/\min(a,b) - 1)\cdot 100\%.$$

In other words, given an arbitrary fraction inside $[1,2]$ that represents the frequency ratio of two notes within one octave from each other, we answer the question "*how close is this frequency ratio to the closest simple fraction?*" Under ET and using C as root note, the following table can be computed, where r represents the frequency ratio:

Note	r	$\text{argmin}_{q\in\mathcal{S}}\, d(q,r)$	$\rho(r)$	Note	r	$\text{argmin}_{q\in\mathcal{S}}\, d(q,r)$	$\rho(r)$
♭D	1.0595	16/15*	0.67%	D	1.1225	9/8	0.22%
♭E	1.1892	6/5	0.89%	E	1.2599	5/4	0.79%
F	1.3348	4/3	0.11%	♯F	$\sqrt{2}$	7/5	1.01%
G	1.4983	3/2	0.11%	♭A	1.5874	8/5	0.78%
A	1.6818	5/3	0.90%	♭B	1.7818	16/9	0.22%
B	1.8877	17/9	0.06%	C'	2.000	2/1	0%

Table 1: Simple fraction proximity coefficient of each note when compared against C. A semitone is so small that it is not representable by fractions with small demoniators; however, it is very close to 16/15.

3.2 Dissonance of Tritone, Quantified

It can be seen immediately that the tritone indeed deviates the most from any simple fraction. In fact, if we visualize all the intervals in which a fraction is considered approximately "nice," then a really interesting pattern appears: $\sqrt{2}$ appears exactly between two of these "nice" intervals, as if it was almost intentional. Since $\mathcal{S}$ covers more than half of $[1, 2]$ (i.e., $\mu(S) > 0.5$ where μ denotes the Lebesgue measure), under such metric, more than half of the times, if we randomly pick a tune within one octave from the base note, the resulting interval will sound "nicer" compared to a tritone.

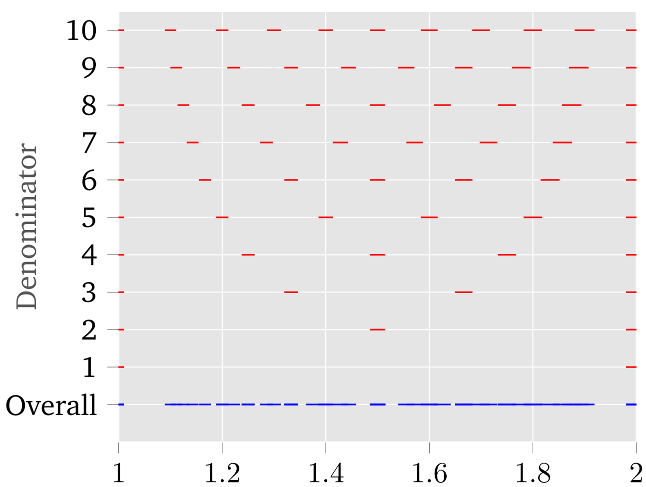

Figure 4.1: a visualization of the set $\mathcal{S}$.

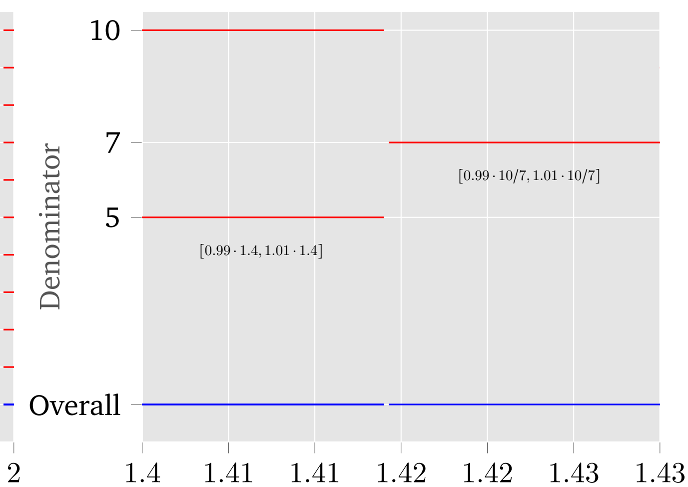

Figure 4.2: a visualization of $\mathcal{S} \cap [1.4, 1.43]$.

Furthermore, if we compute the proximity coefficient of the tritone under PT, obtaining the frequency ratio of $(3/2)^6 \cdot 2^{-3} = 729/512$, this error increases to 1.7%, which may partially explain why it was considered even more dissonant in the early centuries.

However, tritone is not "beyond salvage," even from classical and a pure mathematical persepctive. Consider a dominant seventh chord, C-E-G-♭B. The frequency ratios between C and each other note can be approximated by 5/4, 3/2, and 16/9, who share a common denominator of 36. Similarly, a diminished triad C-♭E-♭G, where the frequency ratios with respect to C are 6/5 and 7/5, has a surprisingly simple common denominator 5. Previous studies have also pointed out that when two or notes are played together, the resulting complex sound wave often masks the slight inharmonicity of individual notes, making everything sound purer (Fletcher and Rossing, 2012). In other words, a tritone, when played by itself, may sound more dissonant than when combined with other notes, for example when being part of a dominant seventh or a diminished triad.

3.3 Harmonic Resolution & Common Denominator

Another interesting observation compatible with this mathematical framework is the harmonic resolutions of classical and some romantic compositions. In general, chords that express dire needs of suspension tend to have a higher common denominator, and as they converge toward resolution, this common denominator decreases significantly.

For example, consider the first four measures of Schubert/Liszt's *Ständchen*, D.957/S.560:

Figure 5: Beginning of Schubert-Liszt's Ständchen (Schubert-Liszt 1840).

A measure-by-measure chord analysis is shown in the following table, where CD denotes common denominator, and frequency ratios are computed with respect to the base note in that measure:

Measure	Notes (ascending)	Chord	Frequency Ratios	CD
1	D-F-A	D minor (I)	$6/5, 3/2$	10
2	♭B-D-F	♭B major (VI)	$5/4, 3/2$	4
3	G-♭B-D-E	G minor sixths (VI)	$6/5, 3/2, 5/3$	30
4	A-♯C-E	A major (V)	$5/4, 3/2$	4

Table 2: A harmonic and frequency ratio analysis of the first four measures of *Ständchen*.

The piece starts with a dark and gloomy D-minor chord that transitions into a submediant ♭B-major, depicting the tranquility of a Senerade. But in the third measure, the introduction of a minor sixths brings forth additional colors, complicating the original simple chord progression, and longing for resolution, which eventually takes place as the D-minor return in measure 5. Unsurprisingly, this minor sixth is significantly more complicated in terms of frequency ratios. When playing them, most pianists also choose to start soft, and slightly add some momentum in the third measure while again fading away in the fourth.

Another more involved example is a beautiful three-measure modulation in Schubert's Impromptu in ♭G-major, Op.90, D.899, No.3: Andante.

Figure 6: Measures 21-24 of Schubert's Impromptu Op.90 No.3 (Schubert 1827).

Measure/beat	Notes	Chord	Frequency Ratios	CD
21	♭C-♭E-♭G	major (root)	$5/4, 3/2$	4
21, beat 3	♭C-♭E-F-♭A	diminished 7th with 4-3 suspension	$5/4, 7/5, 5/3$	60
21, beat 3	♭C-D-F-♭A	diminished 7th	$6/5, 7/5, 5/3$	15
21, beat 4	♭C-F-♭A	augmented 7th (3rd inv)	$7/5, 5/3$	15
22	♭B-♭D-♭G	major (1st inv)	$6/5, 8/5$	5
22, 3rd beat	♭C-♭E-G	augmented fifth	$5/4, 8/5$	20
22, 3rd beat cont.	♭C-♭E-♭A	minor (1st inv)	$5/4, 5/3$	12
22, 4th beat	♭C-♭E-♭A	minor (1st inv)	$5/4, 5/3$	12
22, 4th beat cont.	C-♭E-♯F-♭A	augmented 7th (1st inv)	$6/5, 7/5, 8/5$	5
23	♭D-♭G-♭B	major (2nd inv)	$4/3, 5/3$	3
23, beat 3	♭D-♭G-♭A-♭C	dominant 7th (root) with 4-3 suspension	$4/3, 3/2, 16/9$	18
23, beat 3 cont.	♭D-F-♭A-♭C	dominant 7th (root)	$5/4, 3/2, 16/9$	36
23, beat 4	♭D-♭G-♭B	major (2nd inv)	$4/3, 5/3$	3
23, beat 4 cont.	♭D-F-♭A-♭C	dominant 7th (root)	$5/4, 3/2, 16/9$	36
24 (resolution)	♭G-♭B-♭D	major (root)	$5/4, 3/2$	4

Table 3: A harmonic and frequency ratio analysis of measures 21-23.

The surprising suspensions, the dense usage of sevenths, as well as a quick modulation (♭C major ↦ ♭G-major ↦ ♭G-minor ↦ ♭G-major, not listed in the table above) make these three measures arguably the most beautiful place in the first section of the piece, and every time when a change is about to happen, CD surges, but eventually it settles down as the harmonies resolve to the tonic major chord.

4 "No Free Lunch" in Temperament Systems

In machine learning, the *No Free Lunch Theorem* states that no optimization technique is superior in every aspect. In other words, if an algorithm A achieves a better performance on a problem X over an algorithm B, then there must exist another problem Y on which algorithm B outperforms A.

I would like to draw an analogy to different tuning schemes, as there has long been scholarly debate over the prevalence of ET. In his book "How Equal Temperament Ruined Harmony," Duffin claims that ET, despite convenient and mathematically justified, is not the perfect recipe to creating harmonious music (Duffin 2008). As the title suggests, he believes that ET fails to appeal to nature and is rather a manipulated artifact that is bland in taste and color. One of the most well-known piano repertoires, the Well-Tempered Clavier by Bach, was specifically composed to exploit the varying colors in different keys under the temperament he chose. Had Bach composed the WTC in ET, there would have been no difference in character between different keys.

The strength of ET being a practical and easily reproducible solution to the old, unsolved issue of temperament has now made it almost the "standard" way of tuning. Unfortunately, amateurs rarely have opportunities to dive deeper into theory, and few have prior exposure to different tuning systems. Consequently, to most people, me included, ET feels like a second nature. We take ET for granted, and we assume all notes must abide by the laws of physics and mathematics. It is as if we are so accustomed to eating mass-produced burgers served in fast food restaurants that we never know tasteful a handcrafted wagyu burger can be. Duffin suggests that we should not be tempted by the convenience of ET and that we should step up to help restore the balance of a diverse heritage of various tempered scales that we had centuries ago.

But in the end, a tuning system is but one component of a musical composition. Unless explicitly demanded by the composer, for any piece of music, a good tempered scale can be anything as long as it keeps the music in tune. We must also not forget that traditional, physical instruments inevitably induce errors in precision when they are manually tuned. Theories justify empirical observations and feelings. Ultimately, what matters most is the final product: we know a temperament system is successful if the audience cannot “feel” the presense of the tuning. As long as this criterion is satisfied, I venture to say that there is little practical need in determining which systems are superior and inferior, respectively. There does not exist a one-size-fits-all solution to the temperament issue, and there need not to be one.

5 Bibliography

Boethius. *Fundamentals of Music.* Translated by Calvin M. Bower. Edited by Claude V. Palisca. New Haven, Connecticut: Yale University Press, 1989. 3.11.

Duffin, Ross W. *How Equal Temperament Ruined Harmony (and Why You Should Care)*. New York: W.W. Norton & Company, 2008.

Fletcher, Neville H., & Thomas D. Rossing. "*The Physics of Musical Instruments.*" Springer New York, 2012.

Pesic, Peter. *Hearing the Irrational: Music and the Development of the Modern Concept of Number.* Isis 101, No. 3 (2010): 50130. https://doi.org/10.1086/655790.

Pythagoras. "*On the Harmonics of the Universe.*" Translated by Andrew Barker. In *Greek Musical Writings: Volume 2, Harmonic and Acoustic Theory*, edited by Andrew Barker, 111-148. Cambridge University Press, 1989.

Schubert, Franz. "4 Impromptus, D.899." In IMSLP Petrucci Music Library. First published 1827. Accessed April 23, 2023. https://imslp.org/wiki/4_Impromptus,_D.899_(Schubert,_Franz)

Schubert, Franz. "Ständchen." In *Schwanengesang*, D.957, arranged by Franz Liszt, S.560. In IMSLP Petrucci Music Library. First published 1840. Accessed April 23, 2023. https://imslp.org/wiki/Schwanengesang,_D.957_(Schubert,_Franz).

Tenney, James. *A history of consonance and dissonance.* New York: Excelsior, 1988.

Vos, Joos. Subjective Acceptability of Various Regular TwelveTone Tuning Systems in Two-Part Musical Fragments. *The Journal of the Acoustical Society of America* 83, no. 6 (1988): 238392. https://doi.org/10.1121/1.396317.

Zarlino, Gioseffo. "*Sopplimenti musicali.*" In IMSLP Petrucci Music Library. First published 1588. Accessed April 11, 2023. https://imslp.org/wiki/Sopplimenti_musicali_(Zarlino,_Gioseffo).