Continuous Chromagrams and Pseudometric Spaces of Sound Spectra

Jordan Lenchitz[1](✉) and Anthony Coniglio[2]

[1] Florida State University, Tallahassee FL 32306, USA jlenchitz@fsu.edu
[2] Columbia University, New York NY 10027, USA coniglio@math.columbia.edu

Abstract. In this paper we extend the ubiquitous music information retrieval technology of the quantized chromagram (or chroma feature) to propose a continuous chromagram, an octave-reduced spectrogram that logarithmically reduces the frequencies of a sound spectrum onto a half-open interval $[a, 2a)$ we call a chroma octave. We prove that for any real number $r > 1$, any two logarithmically reduced spectrograms onto intervals of reduction $[a_1, ra_1)$ and $[a_2, ra_2)$ with $a_1 \neq a_2$ are equivalent up to logarithmic scaling and rotation. In the case $r = 2$ this proof shows why all chroma octaves bounded by both the upper and lower frequencies of the sound spectrum in question yield essentially the same continuous chromagram. We then propose a family of pseudometrics on sound spectra and discuss potential applications to analysis and composition.

Keywords: Continuous chromagram · Generalized octave reduction · Pseudometric spaces

1 Introduction

1.1 Pitch and Chroma

Although nowhere near as slippery a notion as timbre, pitch—relative to its centrality in theoretical discourse—can easily be taken for granted definitionally. Depending on who was asked and when, pitch has been variously defined as "that attribute of auditory sensation by which sounds are ordered on the scale used for melody in music" [1], "that attribute of sensation whose variation is associated with musical melodies" [2], "[that which] can be reliably matched by adjusting the frequency of a pure tone of arbitrary amplitude" [3], or "a subjective measure [of] relative highness or lowness" [4].

Whatever definition one prefers, pitch in the most abstract clearly merits a continuous rather than a discrete understanding. Shepard's [5] foundational results on the perceptual representation of pitch suggested that the pitch p of a sound can be decomposed into values of chroma c and tone height h: $p = 2^c \cdot 2^h$, where c is any real number $[0, 1)$ and h is any integer in order for the decomposition to be unique. Patterson [6] later generalized Shepard's two-dimensional representation of pitch into a two-dimensional representation of frequency: $f = 2^c \cdot 2^h$, where once again c is any real number $[0, 1)$ and h

is any integer in order for the decomposition to be unique. And recent neuroscience [4] suggests that such two-dimensional representations are compatible with the tonotopic organization of pitch in human auditory cortex.

Wakefield [7] presented an ad hoc method for producing chroma-time representations of sounds in the context of analyzing the resonance of an individual human singing voice before subsequently [8] presenting limited formalization of his ad hoc method. Building on the models of Shepard [5] and Patterson [6], in this paper we provide a non-ad hoc formalization of a continuous chromagram, an octave-reduced spectrogram that sums all amplitudes found in a sound spectrum onto a half-open interval $[a, 2a)$ we call a chroma octave. Working from any sound spectrum, our formalization generalizes the notion of octave reduction to any interval of logarithmic reduction and fosters the formulation of pseudometric spaces of sound spectra for measuring distances between musical sounds. After presenting illustrative technical aspects of our theoretical apparatus, we conclude with brief discussions of potential applications of the present work to analysis and composition as well as future directions for continued exploration of chroma perception.

1.2 Motivations

There are a number of reasons for the consideration of continuous chromagrams. Firstly, the 12-bin equal quantization of frequency underlying chromagrams (also known as chroma features) in most music information retrieval applications does a disservice to the reality of musical diversity. As musical inquiry more broadly begins to come to terms with how to expand its purview, continuous models of chroma provide a natural basis for consideration of musical traditions with different underlying pitch structures than twelve-tone equal partitions of 2:1 octaves.

Secondly, chroma-time representations naturally erase boundaries between timbre and pitch in the consideration of the aggregate quality of a sound. Computational models of chord spacing and chord quality may embrace the possibility of pitch percepts of spectral components other than those of fundamentals. Continuous chromagrams privilege only those octave equivalence classes with more amplitude than others. In light of pitch's longstanding status as a fertile field of inquiry and more recent explorations of timbre using multidimensional models, considering them holistically is attractive from the perspective of perceptually informed theorizing.

Thirdly, chroma-time representations represent a possibility for new and broader directions in the theoretical exploration of parsimony. Fundamental frequencies are obvious places to listen for pitch percepts but they are by no means the only frequencies that a listener might perceive as pitches, especially in harmonically rich polyphonic timbres. To this end, the continuous chromagram facilitates thinking about chromatic parsimony in the very literal sense of chroma and can provide new directions for analysis and composition of musical sounds.

2 Continuous Chromagrams

2.1 Illustration and Definition

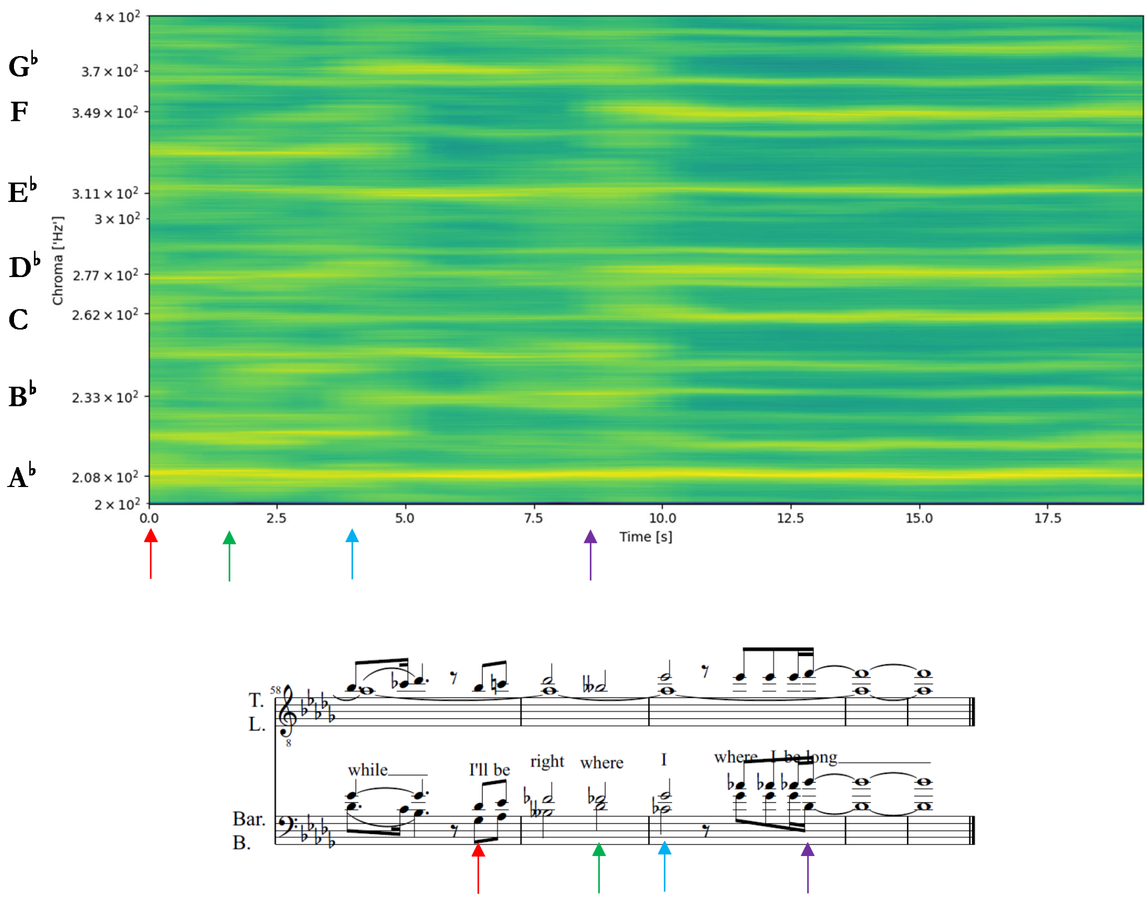

Fig. 1. The ending of the tag of Vocal Spectrum's arrangement of "Go the Distance" as represented through a continuous chromagram and via staff notation. A recording may be accessed at https://youtu.be/DNmW4t22v0s?t=201.

Figure 1 presents a representative example of a continuous chromagram contextualized by staff notation as well as the recording of the excerpt in question. This continuous chromagram presents the ending of the tag to "Go the Distance" from the 1997 Disney film *Hercules* as arranged by the international champion barbershop quartet Vocal Spectrum and is discussed further in Section 4.1. In the meantime, it will suffice to present foundational notions for defining and understanding continuous chromagrams in general.

Logarithmically Reduced Spectra Let $n, N \in \mathbb{R}$ with $0 < n < N$ and let $f : [n, N] \to \mathbb{R}$ represent a sound spectrum. Choose a *reduction parameter* $r \in \mathbb{R}$ with $r > 1$ and suppose $a \in \mathbb{R}$ satisfies $n \leq a < ra \leq N$. We define the *logarithmically reduced spectrum* $g : [a, ra) \to \mathbb{R}$ by

$$g(x) := \sum_{\substack{\{y \in [n,N] \,:\, \exists k \in \mathbb{Z} \\ \text{s.t. } x = r^k y\}}} f(y), \qquad \text{for } x \in [a, ra), \tag{1}$$

where $[a, ra)$ is the *interval of reduction*. In particular, $[a, 2a)$ yields a spectrum reduced onto a *chroma octave*, while $[a, ra)$ for $r \approx 2$ reduces onto a *chroma pseudo-octave*. For $r = 2$ the logarithmically reduced spectrum may be said to be octave reduced in the conventional musical sense. If the chosen $r \neq 2$ we define the process of producing the associated logarithmically reduced spectrum as *generalized octave reduction.*

Logarithmically Reduced Spectrograms and Continuous Chromagrams Implicit in the above formulation of logarithmically reduced spectra is the elimination of time variance. In practice, this would tend to stem from either only looking at one minuscule temporal window or taking the average of the spectrum over some time interval that ideally does not feature too many transients. In computational work on *a cappella* singing, for instance, one might well compute continuous chromagrams in the middle of sustained chords where vowels and formants are most stable. Incorporating time into our formulation of logarithmically reduced spectra is nonetheless worthwhile because of the number of applications which motivate working with transients and/or other phenomenon where time variance matters strongly.

Suppose $[0, T] \subset \mathbb{R}$ for some $T \in \mathbb{R}$ represents time. Let $f : [n, N] \times [0, T] \to \mathbb{R}$ represent a time variant sound spectrum. For a chosen reduction parameter r and interval of reduction $[a, ra)$ as defined above, we define the *logarithmically reduced spectrogram* $g : [a, ra) \times [0, T] \to \mathbb{R}$ by

$$g(x, t) := \sum_{\substack{\{y \in [n,N] \,:\, \exists k \in \mathbb{Z} \\ \text{s.t. } x = r^k y\}}} f(y, t), \qquad \text{for } x \in [a, ra), \quad t \in [0, T]. \tag{2}$$

For $r = 2$, g is a *continuous chromagram* providing a chroma-time representation of a sound analogous to the frequency-time representation of a conventional spectrogram but with the amplitudes of all frequencies sharing a chroma value summed onto a single representative in the chroma octave. Continuous in this context is therefore used to contrast with the conventional chromagram (or chroma feature), which can most accurately be described as a quantized chromagram since it reduces amplitudes into twelve logarthmically equal bins corresponding to the twelve chroma of twelve-tone equal temperament.

2.2 Invariance to Interval of Reduction

We next present a proof justifying why one's choice of interval of reduction in producing a logarithmically reduced spectrum does not essentially change the result. For quantized chromagrams the bins are almost always determined in relation to a fixed reference frequency such as A4 = 440 Hz using the convenience of multiplying by powers of $\sqrt[12]{2}$. In the case of logarithmically reduced spectra, on the other hand, there are no structural *a priori* assumptions that would motivate the choice of any particular interval of reduction absent any more specific context. One may nonetheless choose any appropriately bounded interval of reduction for a given sound spectrum and produce the same logarithmically reduced spectrum up to logarithmic scaling and rotation.

Theorem 1. *For a fixed reduction parameter $r > 1$, let g and h be logarithmically reduced spectra with intervals of reduction $[a_1, ra_1)$ and $[a_2, ra_2)$, respectively and with $a_1 \neq a_2$. Then g and h are equivalent up to logarithmic scaling and rotation.*

Proof. Without loss of generality, we will assume that $a_1 < a_2$. (Otherwise, swap the roles of g and h in what follows.)

We will consider the following two cases separately and prove the Theorem separately for each case:

1. Case 1: there exists some $m \in \mathbb{N}$ such that $\frac{a_2}{a_1} = r^m$; and
2. Case 2: there does *not* exist any $m \in \mathbb{N}$ such that $\frac{a_2}{a_1} = r^m$.

Proof of the Theorem for Case 1: Suppose that $\frac{a_2}{a_1} = r^m$ for some $m \in \mathbb{N}$. We claim that h can be written in terms of g as

$$h(x) = g(r^{-m}x) \quad \text{for } x \in [a_2, ra_2). \tag{3}$$

Before verifying Equation 3, we need to check that the expression on the right-hand side of Equation 3 is well-defined; that is, we need to check that $r^{-m}x \in [a_1, ra_1)$ if $x \in [a_2, ra_2)$, otherwise it would not make sense to apply g to $r^{-m}x$ as shown on the right-hand side of Equation 3 (since g is defined only on $[a_1, ra_1)$ *a priori*).

Indeed, if $x \in [a_2, ra_2)$ then $a_2 \leq x$ and thus, using the equality $\frac{a_2}{a_1} = r^m$, we conclude $a_1 = r^{-m}a_2 \leq r^{-m}x$. Furthermore, the inclusion $x \in [a_2, ra_2)$ implies $x < ra_2$ and thus, using the equality $\frac{a_2}{a_1} = r^m$, we conclude $r^{-m}x < r^{-m}(ra_2) = r \cdot r^{-m}a_2 = ra_1$. Combining these results, we see that $a_1 \leq r^{-m}x < ra_1$. Since our choice of $x \in [a_2, ra_2)$ was arbitrary, we conclude that $r^{-m}x \in [a_1, ra_1)$ for all $x \in [a_2, ra_2)$ and hence that the right-hand side of Equation 3 is indeed well-defined.

So, now we proceed to verify the equality in Equation 3. We compute, for $x \in [a_2, ra_2)$, that

$$g(r^{-m}x) = \sum_{\substack{\{y\in[n,N]\,:\,\exists k\in\mathbb{Z} \\ \text{s.t. } r^{-m}x=r^k y\}}} f(y) = \sum_{\substack{\{y\in[n,N]\,:\,\exists k\in\mathbb{Z} \\ \text{s.t. } x=r^{k+m} y\}}} f(y) = \sum_{\substack{\{y\in[n,N]\,:\,\exists l\in\mathbb{Z} \\ \text{s.t. } x=r^{l} y\}}} f(y) = h(x).$$

Therefore, Equation 3 indeed holds. Equation 3 states, in other words, that h is obtained from g by a logarithmic scaling (by r^{-k}). In this case, no rotation is necessary in order to obtain h from g.

Proof of the Theorem for Case 2: Suppose that there does *not* exist any $m \in \mathbb{N}$ with $\frac{a_2}{a_1} = r^m$. Let $m_0 := \min\left\{m \in \mathbb{Z} : r^m > \frac{a_2}{a_1}\right\}$. That is, m_0 is the unique integer such that $r^{m_0}a_1 \in [a_2, ra_2)$. So, $a_2 \leq r^{m_0}a_1 < ra_2$. We claim that h can be written in terms of g as

$$h(x) = \begin{cases} g(r^{1-m_0}x) & \text{if } x \in [a_2, r^{m_0}a_1) \\ g(r^{-m_0}x) & \text{if } x \in [r^{m_0}a_1, ra_2). \end{cases} \tag{4}$$

Before verifying Equation 4, we need to check that the expressions on the right-hand side of Equation 4 are well-defined; that is, we need to check that $r^{1-m_0}x \in [a_1, ra_1)$ if $x \in [a_2, r^{m_0}a_1)$ and that $r^{-m_0}x \in [a_1, ra_1)$ if $x \in [r^{m_0}a_1, ra_2)$, otherwise it would not make sense to apply g to those values as shown on the right-hand side of Equation 4 (since g is defined *a priori* only on $[a_1, ra_1)$):

- First suppose that $x \in [a_2, r^{m_0}a_1)$. Then $a_2 \leq x$ and thus, using the inequality $r^{m_0}a_1 < ra_2$, we conclude $a_1 = r^{-m_0}(r^{m_0}a_1) < r^{-m_0}ra_2 = r^{1-m_0}a_2 \leq r^{1-m_0}x$. Furthermore, the inclusion $x \in [a_2, r^{m_0}a_1)$ implies $x < r^{m_0}a_1$ and thus $r^{1-m_0}x < r^{1-m_0}r^{m_0}a_1 = ra_1$. Combining these results, we conclude that $a_1 \leq r^{1-m_0}x < ra_1$, i.e., $r^{1-m_0}x \in [a_1, ra_1)$.
- Next, suppose that $x \in [r^{m_0}a_1, ra_2)$. Then $r^{m_0}a_1 \leq x$ and thus $a_1 = r^{-m_0}(r^{m_0}a_1) \leq r^{-m_0}x$. Furthermore, the inclusion $x \in [r^{m_0}a_1, ra_2)$ implies $x < ra_2$ and thus, using the inequality $a_2 \leq r^{m_0}a_1$, we conclude $r^{-m_0}x < r^{-m_0}ra_2 = r^{1-m_0}a_2 \leq r^{1-m_0}(r^{m_0}a_1) = ra_1$. Combining these results, we conclude that $a_1 \leq r^{-m_0}x < ra_1$, i.e., $r^{-m_0}x \in [a_1, ra_1)$.

Now that we have shown that the right-hand side of Equation 4 is well-defined, we now proceed to verify the equality. We first verify it for $x \in [a_2, r^{m_0}a_1)$ and then verify it for $x \in [r^{m_0}a_1, ra_2)$.

If $x \in [a_2, r^{m_0}a_1)$, then

$$g(r^{1-m_0}x) = \sum_{\substack{\{y\in[n,N]\,:\,\exists k\in\mathbb{Z} \\ \text{s.t. } r^{1-m_0}x = r^k y\}}} f(y) = \sum_{\substack{\{y\in[n,N]\,:\,\exists k\in\mathbb{Z} \\ \text{s.t. } x = r^{k+m_0-1} y\}}} f(y) = \sum_{\substack{\{y\in[n,N]\,:\,\exists l\in\mathbb{Z} \\ \text{s.t. } x = r^l y\}}} f(y) = h(x),$$

and so Equation 4 holds for $x \in [a_2, r^{m_0}a_1)$. On the other hand, if $x \in [r^{m_0}a_1, ra_2)$, then

$$g(r^{-m_0}x) = \sum_{\substack{\{y\in[n,N]\,:\,\exists k\in\mathbb{Z} \\ \text{s.t. } r^{-m_0}x = r^k y\}}} f(y) = \sum_{\substack{\{y\in[n,N]\,:\,\exists k\in\mathbb{Z} \\ \text{s.t. } x = r^{k+m_0} y\}}} f(y) = \sum_{\substack{\{y\in[n,N]\,:\,\exists l\in\mathbb{Z} \\ \text{s.t. } x = r^l y\}}} f(y) = h(x),$$

and so Equation 4 holds for $x \in [r^{m_0}a_1, ra_2)$.

Therefore, Equation 4 indeed holds for all $x \in [a_2, ra_2)$. Equation 4 states, in other words, that h is obtained from g by a logarithmic scaling of, separately,

$[a_2, r^{m_0}a_1)$ by r^{1-m_0} and of $[r^{m_0}a_1, ra_2)$ by r^{-m_0}. This can be equivalently thought of as a rotation of $[r^{m_0}a_1, ra_2) \subset [a_2, ra_2)$ to be immediately to the left of $[a_2, r^{m_0}a_1)$ while scaled logarithmically followed by a logarithmic scaling of the new full interval down until it precisely overlays $[a_1, ra_1)$.

□

For the reduction parameter $r = 2$, this proof justifies why any two chroma octaves bounded by both the upper and lower frequencies of the sound spectrum in question yield essentially the same continuous chromagram. Additionally, the proof's independence of reduction parameter r demonstrates how our notion of a continuous chromagram can be extended to intervals other than the octave in the process we defined as generalized octave reduction. Consider for instance a recording of monophonic music composed using the Bohlen–Pierce scale (BP scale) which repeats at the tritave (the frequency ratio 3:1) rather than the 2:1 octave and performed by a timbre such as the clarinet whose spectrum consists overwhelmingly of odd harmonics outside of higher registers. A continuous chromagram of such a sound signal onto a chroma octave such as $[200, 400)$ Hz could be arguably unhelpful at best in understanding the music relative to its compositional structure. Choosing a chroma tritave of $[200, 600)$ Hz instead would make for a more parsimonious mapping between the representation and the compositional structure. Logarithmically reduced spectrograms additionally offer new possibilities for investigating musical sounds with inharmonic timbres. Indeed, generalized octave reduction of sound spectra onto intervals of reduction $[a, ra)$ with $r \neq 2$ would represent a logical choice for exploring sounds such as chords of tones with stretched or compressed partials produced by Slaymaker [9] and Mathews [10] as well as the consonant nonoctaves used as illustrations by Sethares [11]. As the above proof has shown, such generalized octave reductions may also pick any interval of reduction within the bounds of the sound spectrum in question and obtain the same result up to logarithmic scaling and rotation.

2.3 Computational Implementations

A tension between the terms continuous and chromagram is inevitable in signal processing applications which necessarily entail sampling. Continuous may thus be understood in computational implementations as referring to the highest resolution possible via one's spectral analysis of a given signal. In a context of either continuous chromagrams or generalized octave reduction, one may observe two limitations of the fast Fourier transform (FFT) for spectral analysis. Because its linear frequency bins result in logarithmic increases in resolution in higher octaves, the logarithmic reduction of fibers to their image in the chroma octave (or other interval) sums amplitudes unevenly. The presence of a zero-frequency/DC term in the spectrum violates one of the hypotheses of our previous theorem (namely that the spectrum is defined on $[n, N]$ for $n > 0$). A Python implementation of a continuous chromagram generated via the FFT (and thus bearing the two aforementioned limitations) is freely available on Github [12]. For future work, the constant-Q transform—oft-maligned for its lack of total invertibility although extendible to least-squares invertibility [13]—may address both of these limitations at once.

3 Pseudometric Spaces of Sound Spectra

One application of continuous chromagrams and generalized octave reduction is to measuring the similarity or dissimilarity of musical sounds. If comparisons based on octave reduction in the usual sense measure distances between chords in a notes-on-a-page sense, then comparisons based on continuous chromagrams instead measure distances along the logarithmically reduced chroma helix [4]. We present as an illustration a family of pseudometrics arising from the theory of L^p, metric, and pseudometric spaces [14,15] on sets of sound spectra and then observe how it can be used to derive metric spaces of sound spectra. This family is by no means the only potentially fruitful source of pseudometrics for comparing musical sounds. Among myriad options, one could for instance imagine adapting earth mover's distance-based similarity analysis (EMDSA) [16] to pseudometric comparisons of sound spectra.

3.1 L^p Pseudometrics

Suppose $n, N \in \mathbb{R}$ with $0 < n < N$. Let $r \in \mathbb{R}$ satisfy $r > 1$, and suppose $a \in \mathbb{R}$ satisfies $n \leq a < ra \leq N$. For a function $f : [n, N] \to \mathbb{R}$, define its logarithmically reduced function $\hat{f} : [a, ra) \to \mathbb{R}$ by

$$\hat{f}(x) := \sum_{\substack{\{y \in [n,N] \,:\, \exists k \in \mathbb{Z} \\ \text{s.t. } x = r^k y\}}} f(y), \qquad \text{for } x \in [a, ra).$$

L^p pseudometrics with $p \in [1, \infty)$. Let $p \in [1, \infty)$.

The L^p norm $||\cdot||_{L^p([a,ra))}$ on the vector space

$$L^p\big([a, ra)\big) := \left\{ g : [a, ra) \to \mathbb{R} \,|\, g \text{ is Lebesgue-measurable and } \int_a^{ra} |g(x)|^p \, dx < \infty \right\},$$

where two functions which are equal almost everywhere are considered to define the same element of $L^p\big([a, ra)\big)$, is defined by

$$||g||_{L^p([a,ra))} := \left(\int_a^{ra} |g(x)|^p \, dx \right)^{\frac{1}{p}}. \tag{5}$$

The metric $d_{L^p([a,ra))}$ on $L^p\big([a, ra)\big)$ induced from $||\cdot||_{L^p([a,ra))}$ is

$$d_{L^p([a,ra))}(g_1, g_2) := ||g_1 - g_2||_{L^p([a,ra))} \tag{6}$$

for $g_1, g_2 \in L^p\big([a, ra)\big)$. From Equation 6 we may define on the set $\widetilde{S}^p_{[a,ra)}\big([n, N]\big) := \Big\{f : [n, N] \to \mathbb{R} | \hat{f} \in L^p\big([a, ra)\big)\Big\}$ an associated pseudometric $\tilde{d}_{\widetilde{S}^p_{[a,ra)}([n,N])}$ by

$$\tilde{d}_{\widetilde{S}^p_{[a,ra)}([n,N])}(f_1, f_2) := d_{L^p([a,ra))}(\hat{f}_1, \hat{f}_2) \tag{7}$$

for $f_1, f_2 \in \widetilde{S}^p_{[a,ra)}\big([n, N]\big)$; we refer to $\tilde{d}_{\widetilde{S}^p_{[a,ra)}([n,N])}$ as an L^p pseudometric.

L^p pseudometric with $p = \infty$. One may obtain for $p = \infty$ a pseudometric $\tilde{d}_{\widetilde{S}^\infty_{[a,ra)}([n,N])}$ in the same manner as for the above L^p pseudometrics by replacing the instances of the integral $\int_a^{ra} |g(x)|^p dx$ with the essential supremum

$$\text{ess sup}_{x \in [a,ra)} |g(x)|$$

and leaving the rest of the formulation essentially unchanged [14,15,17].

3.2 Metrics Induced from Pseudometrics

If $\tilde{X}$ is any set and $\tilde{d}$ is a pseudometric on $\tilde{X}$, one can canonically obtain a corresponding metric space as follows [15,18]. Put the equivalence relation $\sim$ on $\tilde{X}$ given by $x \sim y$ if $\tilde{d}(x, y) = 0$. If one lets X be the set of equivalence classes of $\tilde{X}$ under this equivalence relation and defines $d([x], [y]) := \tilde{d}(x, y)$ for the equivalence classes $[x]$ and $[y]$ of x and y respectively, then d is a well-defined metric on X, and thus (X, d) is a metric space. This may be applied, for each $p \in [1, \infty]$, to the pseudometric space $\Big(\widetilde{S}^p_{[a,ra)}\big([n, N]\big), \tilde{d}_{\widetilde{S}^p_{[a,ra)}([n,N])}\Big)$, with the aforementioned equivalence relation $\sim$, to produce a metric on the set

$$S^p_{[a,ra)}\big([n, N]\big) := \bigg(\Big\{f : [n, N] \to \mathbb{R} | \hat{f} \in L^p\big([a, ra)\big)\Big\} \Big/ \sim\bigg),$$

i.e., on the set of spectra being reduced into $L^p\big([a, ra)\big)$ wherein two spectra are considered to be equivalent if their logarithmically reduced spectra are equal almost everywhere on $[a, ra)$.

3.3 Comparing Time Variant Spectra

Pseudometric measures of similarity or dissimilarity can be extended analogously to our generalized octave reduction apparatus to account for time variance as follows. If we consider logarithmically reduced spectrograms of the form $f : [n, N] \times [0, T] \to \mathbb{R}$, then analogously to above, for $p \in [1, \infty)$ one may yield metrics and pseudometrics by replacing the instances of the integral $\int_a^{ra} |g(x)|^p dx$ with

$$\int_0^T \int_a^{ra} |g(x, t)|^p \, dx \, dt$$

and leaving the rest of the formulation essentially unchanged.

4 Applications

4.1 Visualizations of Musical Sounds

Returning to Figure 1, the color-coded arrows provide a temporal reference while the equal-tempered letter names are for reference within the chroma octave $[200, 400)$ 'Hz'. The ordinate is labeled as 'Hz' because these are not frequencies corresponding to single pitches but rather representatives of their chroma within that chroma octave. This continuous chromagram visualization highlights the salience of several prominent spectral components beyond the four sung fundamental frequencies. These components stem from the combination of formant spread and formant tuning [19] compounded with precise vertical intonation to maximize the production of ring [20].

This illustration demonstrates how continuous chromagrams can facilitate new framings of parsimonious voice-leading. 'Above' the lead's sustained A♭, for instance, we can observe a veritable thicket of chroma across which it would not be unreasonable to trace smooth voice-leading trajectories. The perceptual implications of such trajectories are not to be overstated until they might be validated by ecologically valid experiments on listening. In the meantime, if we listen again we can find that this chroma-time representation can not only represent such phenomena but also (and perhaps even more importantly) facilitate new ways of listening.

4.2 Creative and Theoretical Possibilities

Continuous chromagrams, generalized octave reduction, and the various pseudometrics we have presented also have further, more speculative applications. Among these are creative sonic practices and possibilities for future theorizing. One elegant formulation of a compositional constraint could easily stem from the fact that there are infinitely many sounds in the fiber of any logarithmically reduced spectrum.

Remark 1. Suppose that we are given a function $g : [a, ra) \to \mathbb{R}$. Let us construct a function $f : [n, N] \to \mathbb{R}$ such that g is the logarithmically reduced version of f, i.e.,

$$\sum_{\substack{\{y \in [n,N] \,:\, \exists k \in \mathbb{Z} \\ \text{s.t. } x = r^k y\}}} f(y) = g(x), \qquad \text{for } x \in [a, ra). \tag{8}$$

For each $x \in [n, N]$, let $\tilde{x}$ denote the unique number in $[a, ra)$ such that $x = 2^k \tilde{x}$ for some $k \in \mathbb{Z}$. Then,

$$f(x) := \frac{g(\tilde{x})}{|\{y \in [n, N] \,:\, \exists k \in \mathbb{Z} \text{ s.t. } x = r^k y\}|} \tag{9}$$

is one such function, where the vertical bars denote cardinality.

Although it is by construction impossible to regain the information lost through the process of logarithmically reducing a sound spectrum, this does not mean that one could not generate sounds from a logarithmically reduced spectrogram. From a musical perspective, a sort of chromatic (in the sense of chroma) minimalism could stem from treating none of pitch, timbre or rhythm alone as a basis for structural repetition and instead repeating the entire image of fibers within the chroma octave of a continuous chromagram. Another possible compositional constraint derived from logarithmically reduced spectra would be to limit oneself to sounds all equidistant under a given metric or pseudometric, giving at once many more possibilities than similar constraints on dots-on-a-page representations of musical similarities while also posing substantial limitations on the timbral possibilities for the sounds in question. In entirely acoustic contexts such constraints are probably more readily posited than realized. Nonetheless, through the use of digital tools speculative applications of chromatic (again in the sense of chroma) thinking might well prove musically rewarding.

One additional compositional approach motivated by our exploration of generalized octave reduction is in following with Slaymaker [9], Mathews [10], and Sethares [11] in designing timbres based on non-octave intervals. If we, for instance, sought to produce a given number of discrete amplitudes in an $\left[a, \frac{7}{4}a\right)$ interval of reduction through purely additive synthesis in just intonation, we would necessarily want to think differently both in terms of relationships between spectral components and in terms of where they lie to begin with than if we were reducing to the chroma octave $[a, 2a)$. In a similar vein, composers interested in historical keyboard temperaments might analogously consider limitations imposed on an $\left[a, \sqrt{5}a\right)$ interval of reduction if they wished to partake in contemporary meantone composition.

Finally, another theoretical possibility would be to use logarithmically reduced spectra to quantitatively characterize spectral aggregates. Recent work by Kahrs [21] on the music of Gubaidulina explored "dissonance" in relationship to distance from white noise using the relatively common music information retrieval measures of spectral centroid and spectral flatness. One can imagine how entropy-related or otherwise probabilistic measurements of sound signals derived from logarithmically reduced spectra could extend these notions of distance from noise as possibilities for both analysis and composition. It is our hope that these illustrations of the possibilities afforded by logarithmically reduced spectra and logarithmically reduced spectrograms will begin a conversation that sheds light on further future prospects.

References

1. Pitch, https://asastandards.org/terms/pitch/. Last accessed 8 Jan 2022
2. Plack, C., Oxenham, A.: Overview: the present and future of pitch. In: Plack, C., Oxenham, A., Fay, R., Popper, A. (eds.) Pitch, pp. 1–6. Springer, New York (2005) https://doi.org/10.1007/0-387-28958-5_1
3. Hartmann, W.: Signals, sound, and sensation. AIP Press, Woodbury, NY (1997)

4. Langner, G.: The neural code of pitch and harmony. Cambridge University Press (2015) https://doi.org/10.1017/CBO9781139050852
5. Shepard, R.: Circularity in judgements of relative pitch. J. Acoust. Soc. Am **36**(12), 2346–2353 (1964) https://doi.org/10.1121/1.1919362
6. Patterson, R.: Spiral detection of periodicity and the spiral form of musical scales. Psych. of Music **14**, 44–61 (1986) https://doi.org/10.1177/0305735686141004
7. Wakefield, G.: Chromagram visualization of the singing voice. In: Proceedings of the International Workshop on Models and Analysis of Vocal Emissions for Biomedical Applications, MAVEBA 1999, pp. 24–29 (1999)
8. Wakefield, G.: Mathematical representation of joint chroma-time distributions. In: Proceedings of the 9th SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations, pp. 637–645 (1999)
9. Slaymaker, F. H.: Chords from tones having stretched partials. J. Acoust. Soc. Am. **47**(2), 1569–1571 (1970) https://doi.org/10.1121/1.1912089
10. Mathews, M., Roads, C.: Interview with Max Mathews. Computer Music J. **4**(4), 15–22 (1980) https://doi.org/10.2307/3679463
11. Sethares, W.: Tuning, timbre, spectrum, scale. 2nd edn. Springer, New York (2005)
12. Lenchitz, J.: Continuous chromagram, https://github.com/jordan-lenchitz/continuous-chromagram. Last accessed 14 March 2022
13. Ingle, A., Sethares, W.: The least-squares invertible constant-Q spectrogram and its application to phase vocoding. J. Acoust. Soc. Am. **132**(2), 894–903 (2012) https://doi.org/10.1121/1.4731466
14. Folland, G.: Real analysis. 2nd edn. Wiley, New York (1999)
15. Howes, N.: Modern analysis and topology. Springer, New York (1995)
16. Ruzon, M., Tomsasi, C.: Edge, junction, and corner detection using color distributions. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(11), 1281–1295 (2001) https://doi.org/10.1109/34.969118
17. Evans, L.: Partial differential equations. 2nd edn. American Mathematical Society, Providence, RI (2010)
18. Simon, B.: A comprehensive course in analysis. American Mathematical Society, Providence, RI (2015)
19. Kalin, G.: Formant frequency adjustment in barbershop quartet singing. Master's thesis, KTH Royal Institute of Technology, Stockholm, Sweden (2005)
20. Averill, G.: Bell tones and ringing chords: sense and sensation in barbershop harmony. The World of Music **41**(1), 37–51 (1999)
21. Kahrs, N.: Consonance, dissonance, and formal proportions in two works by Sofia Gubaidulina. Mus. Theor. Online **26**(2) (2020) https://doi.org/10.30535/mto.26.2.7