

RESEARCH ARTICLE

Metrics for Scales and Tunings

Andrew J. Milne^{a*} and William A. Sethares^{b†}

^a*Department of Music, P.O. Box 35, FIN-40014, University of Jyväskylä, Finland;*

^b*1415 Engineering Drive, Department of Electrical and Computer Engineering,
University of Wisconsin, Madison, WI 53706, USA*

(Received d mmmm yyyy; final version received d mmmm yyyy)

We present a class of novel metrics for measuring the distance between any two periodic scales whatever their precise tuning or cardinality. The metrics have some important applications: (1) finding effective lower-dimensional temperaments of higher-dimensional tunings (such as just intonation); (2) finding simple scales that effectively approximate complex scales (e.g., using equally-tuned and well-formed scales to approximate Fokker periodicity blocks and pairwise well-formed scales); (3) finding ways to map the notes of any arbitrary scale to a button lattice controller (a generalized keyboard) so as to maximise geometrical consistency and playability; (4) comparing the distance of various scales (such as equal tunings with different numbers of notes per octave) for analytical, maybe even compositional, purposes; (5) a generalized method to determine the similarity of different “pitch class sets” that is not dependent on the use of a low cardinality equal tunings such 12-tone, thus hinting at an approach towards a generalized theory of chord progression.

Keywords: metric; tuning; scale; pitch class set; interval class vector; fourier transform; autocorrelation

MCS/CCS/AMS Classification/CR Category numbers: ?????; ?????

*Corresponding author. Email: andymilne@tonalcentre.org

†Email: sethares@ece.wisc.edu

1. Introduction

An n -note tuning or scale may be written as a set of n intervals that define an element in \mathbb{R}^n . For instance, the 7-note major scale, a subset of the standard 12-equal divisions of the octave (12-*edo*) can be represented

$$\mathbf{major12} = \{0, 200, 400, 500, 700, 900, 1100\} \in \mathbb{R}^7$$

where the intervals are expressed in cents and reduced to a single octave (i.e., the values are expressed mod 1200). Similarly, 12-*edo* can be written

$$\mathbf{12edo} = \{0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100\} \in \mathbb{R}^{12},$$

while 10-*edo* is

$$\mathbf{10edo} = \{0, 120, 240, 360, 480, 600, 720, 840, 960, 1080\} \in \mathbb{R}^{10}.$$

The syntonic Just major scale [1] is

$$\mathbf{syntonicJI} = \{0, 204, 386, 498, 702, 884, 1088\} \in \mathbb{R}^7.$$

From a musical perspective, it is clear that some such scales can be thought of as closer than others. For instance, a piece written in **syntonicJI** can be played in a subset of **12edo** (such as **major12**) without undue strain, yet may not be particularly easy to perform when the pitches are translated to a subset of **10edo**. Thus it is desirable to have a metric that allows a statement such as **syntonicJI** is closer to **12edo** than to **10edo**.

It is intuitively plausible that tunings may be “close together” or “far apart” based on similarities and differences in the set of intervals that define the tunings. When two tunings have the same number of notes (and hence the same number of intervals), any reasonable metric can be used to describe the distance between them. For example, Chalmers [2] measures the distances between tetrachords using a variety of metrics such as the Euclidean ℓ_2 , the taxicab metric ℓ_1 , and the max-value ℓ_∞ distance. Similarly, the distance between **major12** and **syntonicJI** can be calculated by writing the two tunings as elements of \mathbb{R}^7 and then calculating any of the standard metrics.

However, when two tunings have different cardinalities, there is no obvious way to define a metric since this would require a direct comparison of elements in \mathbb{R}^n with elements in \mathbb{R}^m for $n \neq m$. A common strategy is to identify subsets of the elements of the tuning and then try to calculate a distance in this reduced space. For instance, one might attempt to calculate the distance between **syntonicJI** and **12edo** by first identifying the seven nearest elements of the 12-*edo* scale, and then calculating the distance in \mathbb{R}^7 . Besides the obvious problems with identifying corresponding intervals in ambiguous situations, the triangle inequality will fail in such schemes. For example, let tuning A be 12-*edo*, tuning B be any seven note subset drawn from 12-*edo* (such as the major scale), and tuning C be a different seven note subset of 12-*edo*.

The identification of intervals is clear since B and C are subsets of A . The distances $d(A, B)$ and $d(A, C)$ are zero under any reasonable metric since $B \subset A$ and $C \subset A$, yet $d(B, C)$ is nonzero because the intervals in the two scales are not the same. Hence the triangle inequality $d(B, C) \leq d(B, A) + d(A, C)$ is violated. Analogous counter-examples can be constructed whenever $n \neq m$.

Another strategy used in musical set theory and neo-Reimannian approaches, effectively counts the number of intervals in **12edo** that is contained in each tuning or chord [3]. These interval vectors can then be treated fruitfully as elements of \mathbb{R}^6 (\mathbb{R}^6 can be used in place of \mathbb{R}^{12} due to the symmetry of intervals about the octave) and forms a true metric. This allows rigorous statements (such as those describing the distance between two chords) but it is inherently limited to subsets of equal temperaments.

This paper shows how metrics for tunings can be constructed by embedding the elements of \mathbb{R}^n and \mathbb{R}^m in a single larger space \mathbb{R}^L . Any metric on \mathbb{R}^L can be used, and the “distance” between the tunings is inherited from the corresponding distance in \mathbb{R}^L . The metrics will be stated for periodic (mostly octave-based) tunings and some important special cases are shown.

2. An embedding

A octave-based tuning \mathbf{x} is an element $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$ where each $x_i \in [0, L)$ defines an interval in cents [4] ($L = 1200$ would be the most common value.) The element \mathbf{x} is mapped to a function $x(t)$ on $t \in [0, L)$ by

$$x(t) = \sum_{i=1}^n \delta(t - x_i) * g(t) \quad (1)$$

where $\delta(t)$ is the Kronecker delta function, $g(t)$ is a smoothing kernel, and $*$ is convolution. It is reasonable to interpret $x(t)$ as a single period of a periodic function $x_p(t)$ defined on the real line. Figure 1 shows $x(t)$ for the **syntonicJI** and **12edo** tunings.

Example 2.1 Euclidean metric. Perhaps the most straightforward way to compare two octave based tunings elements \mathbf{x} and \mathbf{y} is via the corresponding functions $x(t)$ and $y(t)$ defined by (1) on $[0, L)$. The Euclidean L_2 distance is

$$d_{L_2}(\mathbf{x}, \mathbf{y}) = \sqrt{\int_0^L (x(t) - y(t))^2 dt}. \quad (2)$$

While the Euclidean distance function (and others such as the L_1 absolute value, and the L_∞ metric) may be useful for comparing chords or pitch class sets within a single scale (see Example 5.1), they do not mimic many of the features usually desired in a distance function between musical scales or tunings. For instance, the periodicity of $x(t)$ (with period $L = 1200$) parallels the octave repetition of the tuning. Under the Euclidean metric, $d_{L_2}(\mathbf{x}, \mathbf{y}) > 0$ when \mathbf{x}

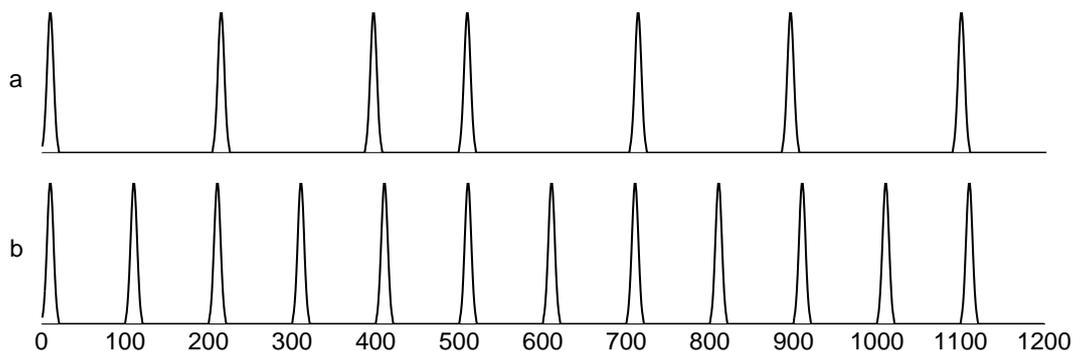


Figure 1. Elements of \mathbb{R}^n are mapped to \mathbb{R}^L ($L = 1200$ is used for illustrative purposes) where they may be sensibly compared. Each component in the space is mapped to a δ function which is smoothed by a kernel $g(t)$ (a finite-length approximation to a Gaussian shape is shown). (a) shows $x(t)$ for the **syntonicJI** tuning and (b) shows $x(t)$ for **12edo**.

and \mathbf{y} are related by a circular shift, i.e., when $y(t-T) = x(t)$. The remaining distance functions are constructed so that tunings which are rotations of each other are considered the same.

Example 2.2 Fourier transform metric. Let p define any of the usual norms $\|\cdot\|_p$ on the set of bounded functions with support on $t \in [0, L)$. Given tunings \mathbf{x} and \mathbf{y} , define the corresponding $x(t)$ and $y(t)$ as in (1), and normalize so that $x_n(t) = x(t)/\|x(t)\|_p$ and $y_n(t) = y(t)/\|y(t)\|_p$. Let $|X_n(f)|$ and $|Y_n(f)|$ be the magnitudes of the Fourier Transforms. Then the distance between \mathbf{x} and \mathbf{y} can be defined as

$$d_{FT}(\mathbf{x}, \mathbf{y}) = \| |X_n(f)| - |Y_n(f)| \|_p. \quad (3)$$

This metric is insensitive to rotations (circular shifts) because it depends only on the magnitude (and not the phase) of the Fourier Transforms.

Circular shifts in the vector \mathbf{x} correspond to circular shifts in $x(t)$ and represent the “same” tuning element. Formally, let $y_p(t) = x_p(t - T)$ for some shift T , where $x_p(t)$ is the periodic extension of $x(t)$, and let $y(t)$ be the single period of $y_p(t)$ with support in $[0, L)$. Then $x(t)$ and $y(t)$ are members of the same equivalence class, written $x(t) \sim y(t)$. This can also be expressed directly in terms of the tunings as $\mathbf{x} \sim \mathbf{y}$. Under the Fourier Transform metric, $d_{FT}(\mathbf{x}, \mathbf{y}) = 0$.

Example 2.3 The equivalence class of the **syntonicJI** tuning contains all circular shifts of the underlying intervals. For example,

$$\begin{aligned} \mathbf{x1} &= \{0, 204, 386, 498, 702, 884, 1088\}, \mathbf{x2} = \{10, 214, 396, 508, 712, 894, 1006\} \\ \mathbf{x3} &= \{0, 112, 316, 498, 702, 814, 1018\}, \mathbf{x4} = \{0, 204, 386, 590, 702, 906, 1088\} \end{aligned}$$

all represent the same tuning with different (shifted) starting points, that is, $\mathbf{x1} \sim \mathbf{x2} \sim \mathbf{x3} \sim \mathbf{x4}$. These elements are all considered the same by the Fourier transform metric of Example 2.2

since $d_{FT}(\mathbf{x1}, \mathbf{x2}) = d_{FT}(\mathbf{x2}, \mathbf{x3}) = d_{FT}(\mathbf{x3}, \mathbf{x4}) = 0$.

An alternative measure is based on the autocorrelation function. Like the Fourier Transform distance, this also respects the equivalence of circular shifts so that $d(\mathbf{x}, \mathbf{y}) = 0$ whenever $\mathbf{x} \sim \mathbf{y}$.

Example 2.4 The autocorrelation metric. Let p define any of the usual norms $\|\cdot\|_p$ on the set of bounded functions with support on $t \in [0, L)$. Given tunings \mathbf{x} and \mathbf{y} , define the corresponding $x(t)$ and $y(t)$ as in (1), and normalize so that $x_n(t) = x(t)/\|x(t)\|_p$ and $y_n(t) = y(t)/\|y(t)\|_p$. Let \mathcal{A}_x be the circular autocorrelation of $x_n(t)$ and let \mathcal{A}_y be the circular autocorrelation of $y_n(t)$. Then the distance between \mathbf{x} and \mathbf{y} can be defined as

$$d_{\mathcal{A}}(\mathbf{x}, \mathbf{y}) = \|\mathcal{A}_x - \mathcal{A}_y\|_p. \quad (4)$$

This metric essentially counts the number of intervals between peaks of the $x_n(t)$ and $y_n(t)$.

For more rapid implementation, observe that the autocorrelation can be written in terms of the Fourier Transform as $\mathcal{A}_x = \text{Re}\{IFT\{|X_n(f)| |X_n(f)|^*\}\}$, where IFT represents the inverse Fourier Transform, $*$ is the complex conjugate, and $\text{Re}\{\}$ takes the real part of its argument. Again, because the metric depends only on the magnitudes (and not the phases) of the Fourier Transforms, it is insensitive to rotations (circular shifts).

The final metric for tunings is used when it is desired to have no explicit peak at the period.

Example 2.5 The centered autocorrelation metric. Define $p, \|\cdot\|_p, \mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^m, x(t), y(t), x_n(t), y_n(t), \mathcal{A}_x$, and \mathcal{A}_y as in Example 2.4. Let $\mathbf{q}_x \in \mathbb{R}^n$ and $\mathbf{q}_y \in \mathbb{R}^m$ be the vectors with a “1” in the leading position followed by $n - 1$ (or $m - 1$) zeros, and define the corresponding $q_x(t)$ and $q_y(t)$ as in (1). Normalize so that $\bar{q}_x(t) = \frac{q_x(t)}{\sqrt{n}\|q_x(t)\|_p}$ and $\bar{q}_y(t) = \frac{q_y(t)}{\sqrt{m}\|q_y(t)\|_p}$, and let $\mathcal{A}_{\bar{q}_x}$ be the circular autocorrelation of $\bar{q}_x(t)$ and $\mathcal{A}_{\bar{q}_y}$ be the circular autocorrelation of $\bar{q}_y(t)$. Then the distance between \mathbf{x} and \mathbf{y} can be defined as

$$d_{c\mathcal{A}}(\mathbf{x}, \mathbf{y}) = \|(\mathcal{A}_x - \mathcal{A}_{\bar{q}_x}) - (\mathcal{A}_y - \mathcal{A}_{\bar{q}_y})\|_p. \quad (5)$$

This metric essentially counts the number of intervals between peaks of the $x_n(t)$ and $y_n(t)$, removing the “extra” peak at the period of repetition L of the tuning by centering about the unit vectors \mathbf{q}_x and \mathbf{q}_y .

Applications of these metrics, and examples of their use, are given in Sections 4 and 5. MATLAB functions to calculate all the above metrics, and a number of the routines used to calculate the examples, can be downloaded from eceserv0.ece.wisc.edu/~sethahares/tuningmetrics.html. The next section examines some interval properties of the metrics.

3. Interval invariances in the tuning metrics

Let $\mathbf{t} \in \mathbb{R}^n$ be an n -element tuning vector. Let m be a $(n^2 - n)$ by n matrix where each row contains a unique permutation of the numbers $+1, -1$, and $n - 2$ zeroes. The set of all intervals

contained in the n -element tuning vector $\mathbf{t} \in \mathbb{R}^n$ is

$$I_{\mathbf{t}} = m\mathbf{t} \pmod{L} \in \mathbb{R}^{n^2-n}$$

and $I_{\mathbf{t}}$ is called the *interval set* or *interval vector* of the tuning \mathbf{t} since it contains all possible differences (intervals) between elements of the tuning. The first result shows that the set of intervals in $I_{\mathbf{t}}$ remains unchanged if the tuning is transposed by adding a constant value to each element.

THEOREM 3.1 *Let $\mathbf{1}_n$ be the n -vector with all elements equal to 1. The interval vector $I_{\mathbf{t}+c\mathbf{1}_n}$ corresponding to the transposed tuning $\mathbf{t} + c\mathbf{1}_n$ is equal to $I_{\mathbf{t}}$.*

Proof:

$$\begin{aligned} I_{\mathbf{t}+c\mathbf{1}_n} &= m(\mathbf{t} + c\mathbf{1}_n) \pmod{L} \\ &= m\mathbf{t} + cm\mathbf{1}_n \pmod{L} \\ &= m\mathbf{t} \pmod{L} = I_{\mathbf{t}} \end{aligned}$$

where $m\mathbf{1}_n = 0$ because m contains exactly one $+1$ and one -1 in each row. ■

The second result shows that permutations of a tuning vector \mathbf{t} leave the intervals in the vector $I_{\mathbf{t}}$ unchanged, except that they may be reordered. A permutation matrix p is a square matrix where all elements are zero except for one 1 in each column and one 1 in each row.

THEOREM 3.2 *Let \mathbf{s} be any permutation of \mathbf{t} , i.e., suppose there is a permutation matrix $p \in \mathbb{R}^{n \times n}$ with $\mathbf{s} = p\mathbf{t}$. Then there is a permutation matrix $P \in \mathbb{R}^{(n^2-n) \times (n^2-n)}$ such that $I_{\mathbf{s}} = PI_{\mathbf{t}}$.*

Proof: By definition

$$I_{\mathbf{s}} = m\mathbf{s} \pmod{L} = mp\mathbf{t} \pmod{L}.$$

The product mp is a permutation of the columns of m . Since each row in m contains exactly two nonzero elements, each row of mp contains the same two nonzero elements (though they may be reordered). Hence, the rows of m and the rows of mp are the same only in a different order. Consequently, there is a permutation P such that $mp = Pm$. Substituting this into the above yields

$$I_{\mathbf{s}} = Pm\mathbf{t} \pmod{L} = PI_{\mathbf{t}}.$$

■

In particular, Theorem 3.2 shows that any circular shift (rotation) of a tuning has the same interval vector. The third result relates the interval vectors of a tuning to the distance under the various metrics of Section 2.

THEOREM 3.3 *If two tunings \mathbf{s} and \mathbf{t} have the same interval vectors (i.e., if there is a permutation matrix P with $I_{\mathbf{s}} = PI_{\mathbf{t}}$), then $d_{FT}(\mathbf{s}, \mathbf{t}) = d_{\mathcal{A}}(\mathbf{s}, \mathbf{t}) = d_{c\mathcal{A}}(\mathbf{s}, \mathbf{t}) = 0$.*

Proof: We begin by showing the result when $\mathbf{s} = (s_1, s_2, s_3)$ and $\mathbf{t} = (t_1, t_2, t_3)$ are three dimensional, and we show it for the discrete case (using the DFT). By construction, $t[k]$ is a vector that is zero everywhere except at $k = t_1, t_2$, and t_3 where it is 1. The DFT is defined as

$$W[n] = \sum_{k=0}^{N-1} t[k] e^{-\frac{j2\pi nk}{N}}.$$

Since there are only three nonzero terms in the sum,

$$W[n] = e^{-\frac{j2\pi nt_1}{N}} + e^{-\frac{j2\pi nt_2}{N}} + e^{-\frac{j2\pi nt_3}{N}}.$$

Multiply both sides by $e^{\frac{+j2\pi nt_1}{N}}$ shows that

$$e^{\frac{+j2\pi nt_1}{N}} W[n] = 1 + e^{-\frac{j2\pi n(t_2-t_1)}{N}} + e^{-\frac{j2\pi n(t_3-t_1)}{N}} \triangleq W_{t_1}[n].$$

Clearly, the magnitudes are equal:

$$|W[n]| = |e^{\frac{+j2\pi nt_1}{N}} W[n]| = |W_{t_1}[n]|.$$

This shows that the magnitude of the transform values are only a function of the differences (intervals) $t_2 - t_1$ and $t_3 - t_1$ and not of the particular values of the t_i themselves. Similarly, $|W[n]| = |W_{t_1}[n]| = |W_{t_2}[n]| = |W_{t_3}[n]|$, and so the magnitude is a function of all of the intervals

$$t_2 - t_1, t_3 - t_1, t_1 - t_2, t_3 - t_2, t_1 - t_3, t_2 - t_3.$$

These are precisely the terms in the interval vector $I_{\mathbf{t}}$. Now suppose there is another tuning vector \mathbf{s} which has the same set of intervals in $I_{\mathbf{s}}$ (i.e., for which there is a permutation matrix P with $I_{\mathbf{s}} = PI_{\mathbf{t}}$). Then there is a reordering of the interval set $s_i - s_j$ for which $|W_{s_i}[n]| = |W_{t_j}[n]|$. Hence the magnitudes of the DFTs are the same, and the difference between the magnitudes is zero.

Observe that the same argument (but with more terms) can be used for any size tuning vector. Observe also that this holds not only for the Fourier distance function (3), but for the autocorrelation distance (4) and the centered autocorrelation (5), since they also can be computed from the magnitude of the transform. ■

There are several ways that two tunings might have the same interval set. Theorem 3.1 shows that (circular) shifts of a tuning $\mathbf{t} + c\mathbf{1}_n$ have the same interval vector as the (unshifted) tuning while Theorem 3.3 shows that these are in the same equivalence class with respect to the three transform-based metrics. Theorem 3.2 shows that tunings that are related by a permutation matrix $\mathbf{s} = p\mathbf{t}$ also have the same interval matrix (and hence are in the same equivalence class).

In fact, these two classes of equivalence do not exhaust all the possibilities. The next example shows a pair of tunings which are not shifts and not permutations of each other, yet which have the same interval set (and hence are at zero distance).

Example 3.4 Let $\mathbf{s} = [0, 100, 400, 600]$ and $\mathbf{t} = [0, 100, 300, 700]$ with $L = 1200$. The interval vector for both $I_{\mathbf{s}}$ and $I_{\mathbf{t}}$ is $[100, 200, 300, 400, 500, 600, 600, 700, 800, 900, 1000, 1100]$ yet \mathbf{s} is not a permutation nor shift of \mathbf{t} . Direct calculation shows $d_{FT}(\mathbf{s}, \mathbf{t}) = d_{\mathcal{A}}(\mathbf{s}, \mathbf{t}) = d_{c\mathcal{A}}(\mathbf{s}, \mathbf{t}) = 0$, in agreement with Theorem 3.3.

4. Lower-dimensional temperaments of higher-dimensional tunings

The *dimensionality* of a tuning is equivalent to the minimum number of unique intervals (expressed in $\log(f)$) that are required to generate, by linear combination, all of that tuning's intervals. One- and two-dimensional tunings have a number of musical advantages:

- (1) They are melodically important because they produce, with no wastage, relatively simple and comprehensible scales [5]. One-dimensional tunings can produce equally-tuned scales such as the 12-edo scale used in contemporary Western music, the 7-edo scale used in traditional Thai music [6], and the 5-edo scale used in Indonesian Slendro [7]. Two dimensional tunings can produce *well-formed* scales, such as the familiar pentatonic and diatonic [8].
- (2) They can be isomorphically mapped to a two-dimensional lattice controller, such as the *Thummer* [9] or *Hex* [10]. This ensures that the geometrical shape of all intervals is invariant over different transpositions, and that the pitch height order of any *privileged* set of shapes/intervals (such as the harmonic consonances of Western music) is invariant over a continuum of tunings [1]. It is speculated that these invariances can significantly enhance the playability and flexibility of the instrument [1, 11].
- (3) Only two control interfaces (such as sliders or rotary knobs) are required to change the tuning of all notes in an ordered way. This enables easily controllable tuning bends that can simulate the variable intonation used by string and aerophone players [12], as well as more extravagant tuning bends used as a novel compositional technique [11, 13].
- (4) The many-to-one mapping implied by dimensionality-reduction enables the use of musical “puns”, whereby different higher-dimensional intervals can be expressed by the same lower-dimensional interval. Such puns can enable enharmonic modulation, and ensure that the familiar I-vi-ii-V-I cadence ends at the same pitch it started (by tempering out the syntonic comma). Each unique choice of intervals (commas) that are tempered out, determines a unique set of voice-leading in the resulting temperament.

However, there are reasons why a higher-dimensional system might serve as a starting point for a given tuning. For example, we may wish to define a set of linearly independent privileged intervals that need to occur, in abundance, within a scale of a given cardinality. The most familiar example is the set of harmonic consonances found in just intonation—when using traditional

Western instruments, the harmonic consonances are generally considered to be those of low-numbered ratios (e.g., $2/1$, $3/2$, $4/3$, $5/3$, $5/4$, etc.) [4], though with non-harmonic spectra quite different intervals may form a useful set of consonances [14]. If the set of privileged intervals has a dimensionality of three (as in the example just given) or more, finding a one- or two-dimensional tempering that does not do too much “damage” to the tunings may be a general method to combine the advantages of both higher and lower-dimensional systems—that is, lots of privileged intervals embedded in a well-formed scale that can be mapped to a button lattice controller, whose tuning can be easily controlled, and which embodies unique voice-leading.

A familiar example of a well-formed scale rich in consonances is the diatonic when tuned to meantone, which contains numerous intervals that closely approximate $3/2$, $5/4$, $5/3$, and their inversions. A less familiar example is the well-formed scale of ten notes that is generated by an octave and an interval of approximately 380 cents, which also contains many good approximations of $3/2$, $5/4$, and $5/3$, and so forth.

Another reason for dimensionality-reduction is to find an effective spatial mapping from the notes of an arbitrary scale to a button lattice controller. Mapping a two-dimensional scale to a two-dimensional button lattice is straightforward [1] so, given an arbitrary scale, it is useful to find similar two-dimensional scales to guide the positioning of the arbitrary notes. The tunings of the notes are not changed, but their spatial layout ensures that intervals with similar size have the same geometric shape, and that the overall shape of the scale is the same as similar scales.

Scales whose spatial mappings can benefit from this process include those with three or more dimensions, or those with an unspecified generator—examples include: five- or higher-limit *Fokker periodicity blocks* [15] and *pairwise well-formed scales* [5] such as the just intonation diatonic scale $1/1$, $9/8$, $5/4$, $4/3$, $3/2$, $5/3$, $15/8$ [4]; *circulating* (irregular) temperaments that are commonly used to accommodate a meantone-like tuning on a twelve-note instrument such as the piano [16]; scales whose tunings have been derived from ethnographic research; scales designed for a specific aesthetic/compositional purpose; regular scales whose generator is unspecified.

Using the metrics defined in Section 2, it is relatively straightforward to find one- and two-dimensional tunings that can approximate any arbitrary set of privileged intervals or scale tunings. A number of examples are given in the following sections.

4.1. Equal (one-dimensional) temperaments

Given a set of privileged intervals, we can test its distance from a series of n -edos (up to any given value of n). The Fourier and autocorrelation metrics are both suitable choices for this task.

Example 4.1 1-D approximations to 4:5:6 (JI major triad). The just intonation major triad contains all and only the common practice harmonic consonances (i.e., the perfect fifth and fourth, and the major and minor thirds and sixths). It is, therefore, useful to find tunings that produce simple scales containing many of these intervals. The just intonation major triad with tuning ratios of $4 : 5 : 6$ is approximated by $\{0, 386, 700\}$ cents. Figure 2 shows the distances between this chord and n -edos with $3 \leq n \leq 55$ using the autocorrelation metric and a window

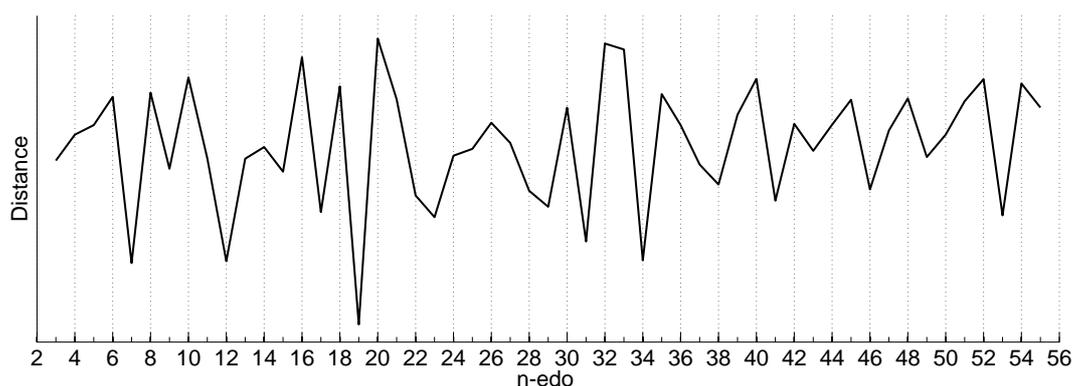


Figure 2. The distance (using the autocorrelation metric with a window of 20 cents) between a just intonation major triad $\{0, 386, 702\}$ and all n -edos from $n = 3$ to $n = 55$.

width of 10 cents.

It is interesting to note that the metric automatically trades the number of notes in the tuning with the accuracy of the tuning. So an n -edo with a large n needs to be more accurate than one with a low n to be rated as similarly close; this would seem to correspond to any naïve, but sensible, notion of distance because most of the intervals in the higher n -edo are superfluous to the task of matching the set of privileged intervals.

4.2. Two-dimensional temperaments

A two-dimensional tuning has two generating intervals with pitch heights, in $\log(f)$, denoted α and β . All intervals in the tuning can be generated by α and β . A β -chain is generated by stacking integer multiples of β for all integers in a finite range of values, so a 19-note β -chain might consist of the notes $j\alpha - 9\beta, j\alpha - 8\beta, \dots, j\alpha + 8\beta, j\alpha + 9\beta$.

Given an arbitrary set of privileged intervals with a period of repetition p (typically 1200 cents), how do we find similar two-dimensional tunings? It is logical to make the tuning of $\alpha = p/n$, for $n \in \mathbb{N}$. This means that, given a choice of α , we can generate a β -chain (of a given cardinality), and then iterate the size of β so that it covers a range from 0 to $\alpha/2$ (any β -tuning outside this range simply duplicates tunings within this range). For each iteration we measure the distance to our set of privileged intervals using the centered autocorrelation metric. The reason we require this metric is because α is already approximating the period, so it is not the function of β to do the same—the centered autocorrelation metric removes the period peak from the autocorrelation of the privileged interval set, so the period plays no part in the distance measure.

Example 4.2 2-D approximations to 4:5:6 (JI major triad). Using the centered autocorrelation metric with a window width of 20 cents gives a number of candidate tunings for a 19-note

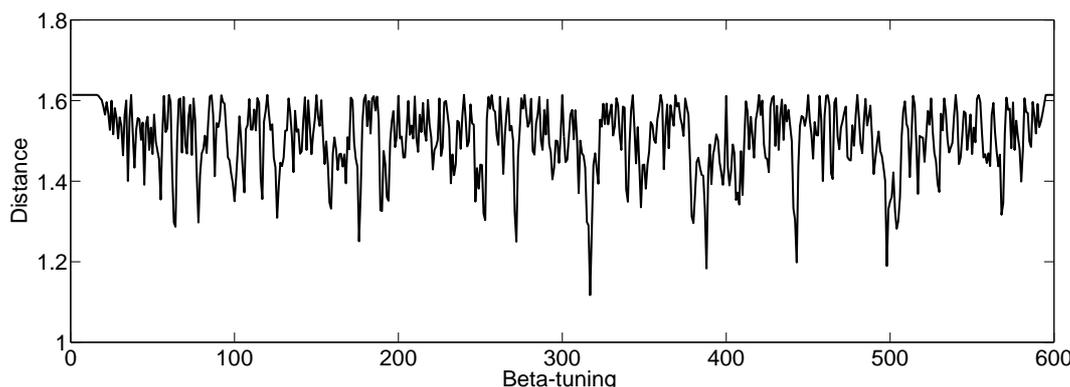


Figure 3. The distance (using the centered autocorrelation metric with a window of 10 cents) between a just intonation major triad $\{0, 386, 702\}$ and a 19-note β -chain whose β -tuning ranges from 0 to 600 cents.

β -chain, as shown in Figure 3.

For example, the minimum at 504 cents corresponds to the familiar meantone temperament; the minimum at 498 cents to the *schismatic* temperament; the minimum at 443 cents to the *sensipent* temperament; the minimum at 388 cents to the *würschmidt* temperament; the minimum at 380 cents to the *magic* temperament; the minimum at 317 to the *hanson* temperament; the minimum at 272 cents to the *orson* temperament; the minimum at 176 cents to the *tetracot* temperament (the names for each of these temperaments has been taken from Erlich [17]).

Example 4.3 2-D approximations to 4:5:6:7 (7-limit JI major tetrad).

If we wish to approximate higher-limit just intonation consonances, the 7-limit chord tuned to $\{0, 386, 702, 969\}$ is an obvious template. Using the same window width and β -chain cardinality as in Example 4.2, gives the distance curve shown in Figure 4.

The local minimum at 503 cents corresponds to 7-limit meantone, the minimum at 317 cents to *keemun*, the minimum at 310 cents to *myna*, the minimum at 271 cents to *orwell*, the minimum at 232 cents to *cynder*, the minimum at 117 cents to *miracle*, the minimum at 88 cents to *nautilus* (the designations have been taken from Erlich [17]). There are also a number of additional (previously unidentified) 7-limit temperaments, notably those at 498 and 194 cents.

Example 4.4 2-D approximations to 4:5:6:7 with a period of 600 cents. This example assumes a scale with a 600 cent period, so α is set to 600 cents. Figure 5 shows the distance between a 19-note β -chain and a tetrad tuned to $\{0, 386, 102, 369\}$, which is an α -reduced version of 4 : 5 : 6 : 7. The closest tuning with a 10 cent window is at $\beta = 217$ cents.

Example 4.5 2-D approximations to 3:5:7 (7-limit Bohlen-Pierce triad). The Bohlen-Pierce scale is intended for spectra with just odd harmonics. It has a period of $3/1$ (the “tritave”), which is approximated by 1902 cents, and the 3:5:7 triad, which is approximated by $\{0, 884, 1467\}$ cents, is treated as a consonance. Figure 6 shows the distance of a β -chain of 20 notes with $0 < \beta < 951$ cents with a 20 cent window. The closest tuning is found at 439 cents, which

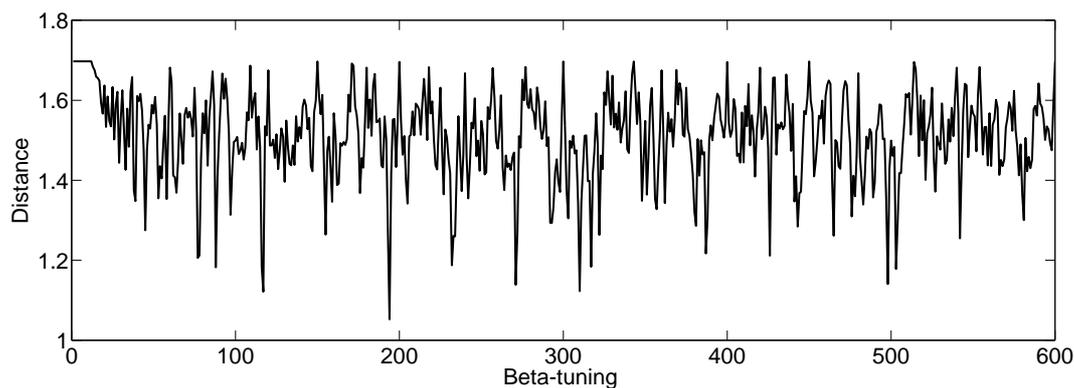


Figure 4. The distance (using the centered autocorrelation metric with a window of 10 cents) between a 7-limit just intonation major tetrad $\{0, 386, 702, 969\}$ and a 19-note β -chain whose β -tuning ranges from 0 to 600 cents.

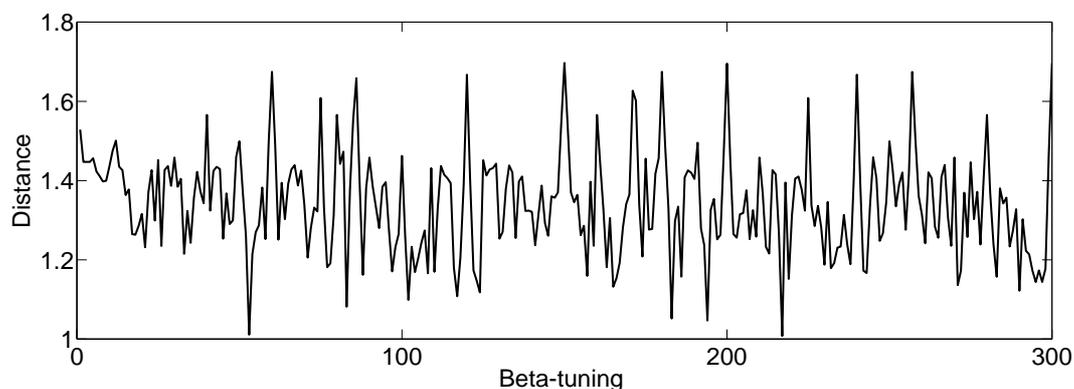


Figure 5. The distance (using the centered autocorrelation metric with a window of 10 cents) between a 7-limit just intonation major tetrad $\{0, 386, 702, 969\}$ with a 600 cent period, and a 19-note β -chain whose β -tuning ranges from 0 to 600 cents.

corresponds to the 13-equal divisions of the tritave tuning suggested by Bohlen and Pierce.

Example 4.6 2-D approximations to an ethnic scale. This task is closely related to the above except the arbitrary scale may have a much higher dimensionality than is typically the case when considering only consonant intervals. It is also safe to assume that if we are given a scale, the period will be explicit (and typically 1200 cents).

The principal purpose of finding the closest two-dimensional tuning to an ethnic scale is to determine an effective spatial mapping to a button lattice controller. We can then map the arbitrary scale as if it were equivalent to its two-dimensional neighbour. As long as the underlying layout mapping is consistent (on the *Thummer* and *Hex*, the default layout is Wicki [1, 18]), this ensures that similar intervals in the arbitrary scale have the same shape on the button lattice,

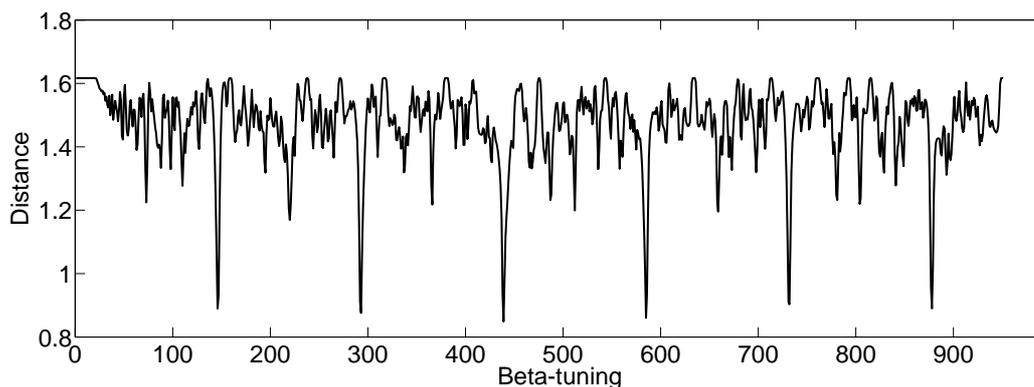


Figure 6. The distance (using the centered autocorrelation metric with a window of 20 cents) between a just intonation Bohlen-Pierce “major” triad $\{0, 884, 1467\}$ and a 20-note β -chain whose β -tuning ranges from 0 to 600 cents.

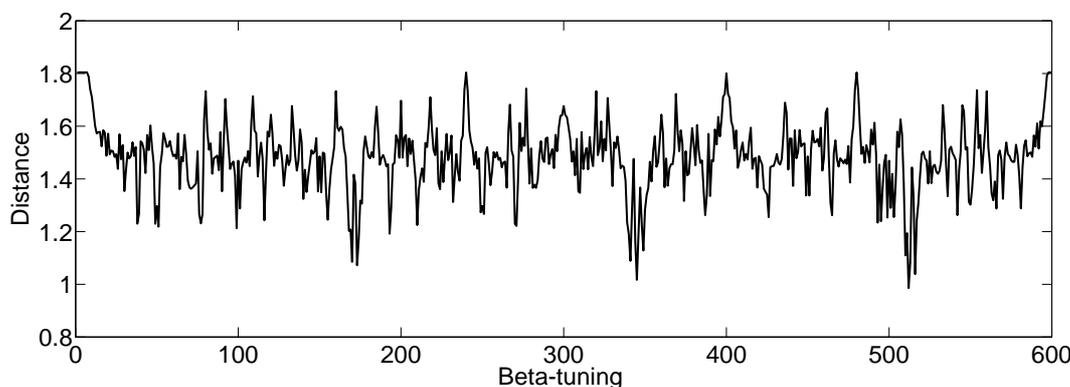


Figure 7. The distance (using the centered autocorrelation metric with a window of 10 cents) between a bagpipe scale $\{0, 197, 341, 495, 703, 853, 1009\}$ and a 19-note β -chain whose β -tuning ranges from 0 to 600 cents. The global minimum is found at 512 cents.

and that the arbitrary scale’s overall note layout is the same as other similar scales (such as its two-dimensional neighbour).

The most useful metric for this purpose is the centered autocorrelation because we already know the period (typically 1200), and this is assigned to α . For this example we will use the scale tunings given for a “highland bagpipe made by MacDonald of Edinburgh” as quoted by Ellis [4], which is $\{0, 197, 341, 495, 703, 853, 1009\}$ with a period of 1200 cents. Using a 19-note β -chain with a window of 10 cents, the closest match is found at $\beta = 512$ cents (see Figure 7).

Example 4.7 2-D approximations to a regular scale with unknown generator. We may be given an arbitrary scale that is regular but we are not given any information about

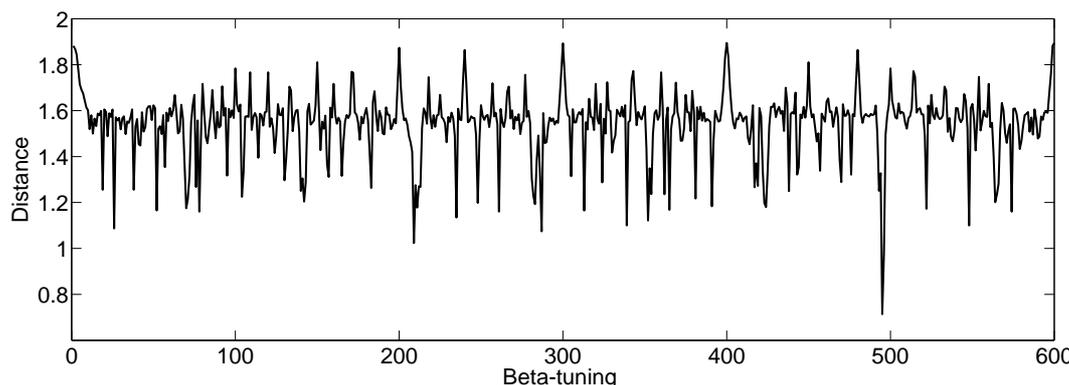


Figure 8. The distance (using the centered autocorrelation metric with a window of 10 cents) between $\{0, 77, 132, 209, 286, 341, 418, 495, 551, 572, 628, 705, 782, 837, 914, 991, 1046, 1068, 1123\}$ and a 19-note β -chain whose β -tuning ranges from 0 to 600 cents. The generator has been correctly identified as 495 cents.

the size of its generator. Figure 8 shows the distance between a 19-note β -chain, with the values of β ranging from 0 to 600, and an arbitrary scale with the following notes $\{0, 77, 132, 209, 286, 341, 418, 495, 551, 572, 628, 705, 782, 837, 914, 991, 1046, 1068, 1123\}$. The generator is correctly identified as 495 cents.

5. Distances between different scales and different tunings

In addition to finding optimal approximations to arbitrary intervals, the metrics can also be used to measure the distance between different tunings, different scales, and different set classes.

For example we can use the metrics to determine how close one n -edo is to another, to give a measure of the potential “strain” produced by switching from one such tuning to another.

There is a rich heritage of measures used to determine the distance between set classes in musical set theory [19]. However, all these measures are predicated on the use of an equal temperament (with a low number of steps per octave). Our proposed tuning metrics can measure the distance between any set of pitch classes from any tuning—equal, regular, or irregular. The Fourier, and autocorrelation metrics defined in this paper are of the $T_n I$ type—that is, they measure the distance between sets that are related by transposition or inversion as 0, they also treat the members of Z -related sets as equivalencies.

However, the use of gaussian smoothing makes the simple Euclidean metric, which is not transposition or inversion invariant, particularly interesting. We can use it to measure the distance between any two pitch class sets, chords, or scales, in any tuning, such as meantone, or even just intonation, as well as being applied to non-Western musical practices. Indeed, to expand on Rahn’s original aims in “Towards a Theory of Chord Progression” [20], such a metric in combination with a theory of “instances (of chords)” hints at a theory of chord progressions, even

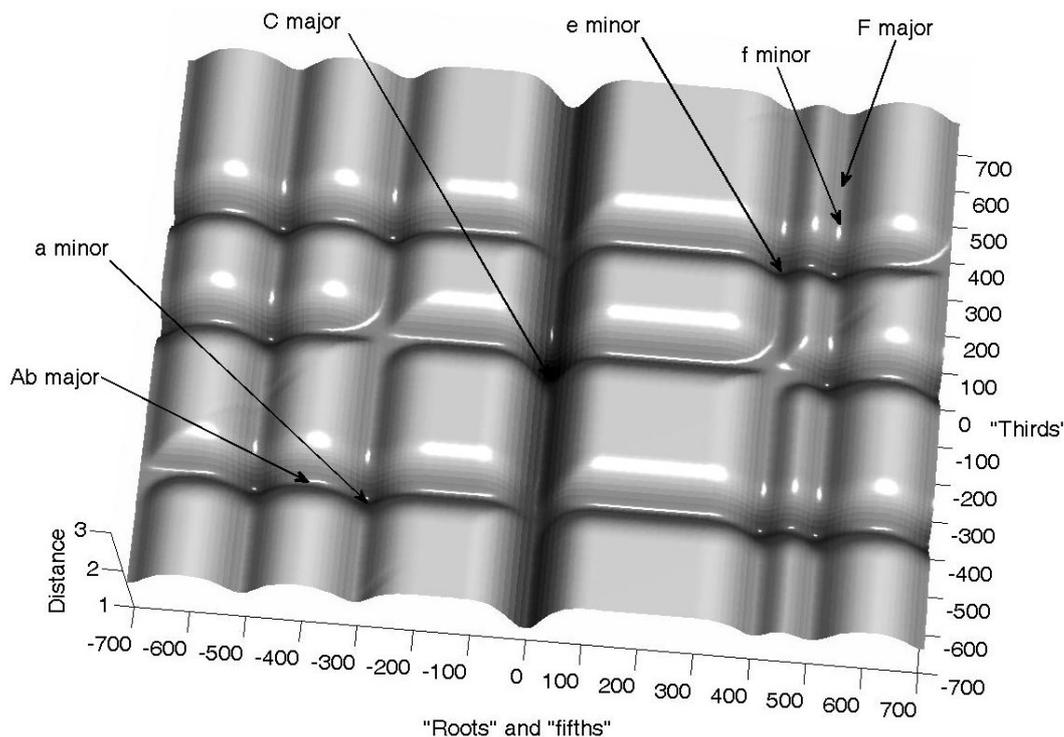


Figure 9. The distance (using the Euclidean metric with a window of 100 cents) between all possible “example” triads and a “reference” major triad tuned to 12-edo. The distance between each example triad and the reference triad is shown as a function of the pitch difference (in cents) between the notes of the example triad and the notes of the reference triad: the y -axis is the cents difference between the thirds of the two triads, the x -axis is the cents difference between the root and fifth of both triads, the z -axis shows the distance as measured by the Euclidean metric. If the reference triad is C major, the positions of a selection of triads are labeled. Observe how triads with common tones (e.g., C and a) lie at local distance minima.

tonality, that could be generalized beyond the low n -edo (i.e., 12-edo) milieu of contemporary Western music.

Example 5.1 Distances between triad-pairings. Using the Euclidean metric, it is possible to measure the distance between any two chords as a function of the pitch differences between their notes. This is illustrated in Figure 9. It is interesting to observe how there are distance minima at those pairings with common tones, with the deepest minima at pairings with two common tones (i.e., between chords related by the Riemannian transformations P, L, and R).

6. Discussion and conclusion

In this paper, we have introduced a number of related metrics and provided a number of examples to demonstrate some of their uses. We believe these metrics to be novel in three principal

respects:

- (1) Although it is an established method to embed interval classes into a 6-vector (and some theorists have explored higher cardinality vectors to represent higher cardinality n -edos), we believe this is the first description of using a vector with a high enough cardinality to represent tunings rather than scale degrees.
- (2) The use of gaussian smoothing to model the inaccuracies of human pitch perception (e.g., the way that an interval of 300 cents may be heard as identical or very similar to another interval of say 301 or 310 cents) is a novel method to make the final distance measure a more accurate model of perceived similarity.
- (3) The use of such high cardinality vectors benefits from an efficient means to represent these scales in a rotation invariant manner. The use of discrete Fourier transforms to achieve this—a technique that is usually the reserve of audio signal processing—is a novel application of this well understood function.

A future line of research would be to correlate the metrics and their variables (such as window shape and width) to empirical ratings of similarity. Furthermore, new window shapes could be derived from collecting ratings of the similarity of tone pairs with different pitch distances, or derived from even more basic psychoacoustic data such as auditory filter shapes. It may also be found that the width of the window may correlate with different musical contexts—for example, melodic similarity may be better modeled with a wider window than would be used for harmonic similarity.

References

- [1] A.J. Milne, W.A. Sethares, and J. Plamondon, *Tuning continua and keyboard layouts*, *J. Math. and Music* 2, (2008), pp. 1–19.
- [2] J. Chalmers, *Divisions of the Tetrachord*, Frog Peak Music, Hanover NH, 1993.
- [3] A. Forte, *The Structure of Atonal Music*, Yale University Press, New Haven and London, 1973.
- [4] H.L.F. Helmholtz, *On the Sensations of Tones*. Trans. Ellis, A.J.. Reprint by Dover in 1954 (New York: Dover).
- [5] N. Carey, *Coherence and sameness in well-formed and pairwise well-formed scales*, *J. Math. Music* 1, (2006), pp. 79–98.
- [6] D. Morton, *The music of Thailand*, in E. May, ed. *Music of Many Cultures*, University of California Press, Berkeley, California, 1980, pp. 63–82.
- [7] W. Surjodiningrat, P. J. Sudarjana, and A. Susanto, *Tone measurements of outstanding Javanese gamelan in Yogyakarta and Surakarta*, Gadjah Mada University Press, Yogyakarta, Indonesia, 1993.
- [8] N. Carey, and D. Clampitt, *Aspects of well-formed scales*, *Music Theory Spectrum* 11, (1989), pp. 187–206.
- [9] Thumtronics Inc., *Thumtronics: The new shape of music*, available online at

- <http://www.thummer.com> (accessed 16 December 2008).
- [10] A.J. Milne, and A. Prechtl, *2032 and Hex: A Dynamic Tonality modal synthesizer and lattice roll sequencer*, in preparation.
 - [11] A.J. Milne, W.A. Sethares, and J. Plamondon, *Isomorphic controllers and dynamic tuning: Invariant fingering over a tuning continuum*, *Computer Music J.* 31 (2007), 15–32.
 - [12] J. Sundberg, A. Friberg, and L. Frydén, *Rules for automated performance of ensemble music*, *Contemporary Music Review* 3 (1989), 89–109.
 - [13] W.A. Sethares, A.J. Milne, S. Tiedje, A. Prechtl, and J. Plamondon, *Spectral tools for Dynamic Tonality and audio morphing*, *Computer Music J.* (in press).
 - [14] W.A. Sethares, *Tuning, Timbre, Spectrum, Scale*, Springer-Verlag, London, 2nd edn, 2004.
 - [15] A.D. Fokker, *Unison vectors and periodicity blocks in the three-dimensional (3-5-7) harmonic lattice of notes*, *Proceedings of the KNAW, Series B*, vol. 72, Amsterdam, (1969), pp. 153–168.
 - [16] J.M. Barbour 1951. *Tuning and Temperament: A Historical Survey*, East Lansing, Michigan: Michigan State College Press. Reprint, New York: Da Capo Press, 1973. Reprint, New York, Dover, 2004.
 - [17] P. Erlich, *A middle path between just intonation and the equal temperaments*, Part 1, *Xenharmonikon* 18, (2006), pp. 159–199.
 - [18] K. Wicki, *Tastatur für Musikinstrumente*. Swiss patent Nr. 13329, in Neumatt (Münster, Luzern, Schweiz) 30 October 1896.
 - [19] M. Castrén, *RECREL: A Similarity Measure for Set-classes*, Sibelius Academy, Helsinki, 1994.
 - [20] J. Rahn, *Toward a theory of chord progression*, *In Theory Only*, 11, (1989), pp. 1–10.