Journal of New Music Research, 2016
Vol. 45, No. 4, 281–294, http://dx.doi.org/10.1080/09298215.2016.1182192

Routledge
Taylor & Francis Group

A multi-level tonal interval space for modelling pitch relatedness and musical consonance

Gilberto Bernardes[1], Diogo Cocharro[1], Marcelo Caetano[1], Carlos Guedes[1,2] and Matthew E.P. Davies[1]

[1]*INESC TEC, Portugal;* [2]*New York University Abu Dhabi, United Arab Emirates*

(Received 12 March 2015; accepted 18 April 2016)

Abstract

In this paper we present a 12-dimensional tonal space in the context of the *Tonnetz*, Chew's Spiral Array, and Harte's 6-dimensional Tonal Centroid Space. The proposed Tonal Interval Space is calculated as the weighted Discrete Fourier Transform of normalized 12-element chroma vectors, which we represent as six circles covering the set of all possible pitch intervals in the chroma space. By weighting the contribution of each circle (and hence pitch interval) independently, we can create a space in which angular and Euclidean distances among pitches, chords, and regions concur with music theory principles. Furthermore, the Euclidean distance of pitch configurations from the centre of the space acts as an indicator of consonance.

Keywords: tonal pitch space, consonance, tonal hierarchy

1. Introduction

A number of tonal pitch spaces have been presented in the literature since the 18th century (Euler, 1739/1968). These tonal spaces relate spatial distance with perceived proximity among pitch configurations at three levels: pitches, chords, and regions (or keys). For example, a tonal space that aims to minimize distances among perceptually-related pitch configurations should place the region of C major closer to G major than B♭ major because the first two regions are understood to be more closely related within the Western tonal music context. For similar reasons, within the C major region, a G major chord should be closer to a C major chord than a D minor chord, and the pitch G should be closer to A than to G#.

The intelligibility and high explanatory power of tonal pitch spaces usually hide complex theories, which need to account for a variety of subjective and contextual factors. To a certain extent, the large number of different, and sometimes contradictory, tonal pitch spaces presented in the literature help us understand the complexity of such representations. Existing tonal spaces can be roughly divided into two categories, each anchored to a specific discipline and applied methods. On the one hand we have models grounded in music theory (Cohn, 1997, 1998; Lewin, 1987; Tymoczko, 2011; Weber, 1817–1821), and on the other hand, models based on cognitive psychology (Krumhansl, 1990; Longuet-Higgins, 1962; Shephard, 1982).

Tonal pitch spaces based on music theory rely on musical knowledge, experience, and the ability to imagine complex musical structures to explain which of these structures work. Tonal pitch spaces based on cognitive psychology intend to capture the mental processes underlying musical activities such as listening, understanding, performing, and composing tonal music by interpreting the results of listening experiments. Despite their divergence in terms of specific methods and goals, music theory and cognitive psychology tonal pitch spaces share the same motivation to capture intuitions about the closeness of tonal pitch configurations, which is an important aspect of our experience of tonal music (Deutsch, 1984). Recent research has attempted to bridge the gap between these two approaches by proposing models that share methods and compare results from both disciplines, such as the contributions of Balzano (1980, 1982), Lerdahl (1988, 2001), and Chew (2000, 2008).

Both music theory and cognitive tonal pitch spaces have been implemented computationally to allow computers to better model and generate sounds and music. Among the

Correspondence: Gilberto Bernardes, INESC TEC (Sound and Music Computing Group), FEUP campus, Rua Dr. Roberto Frias, 4200 - 465 Porto, Portugal. E-mail: gba@inesctec.pt
This article was originally published with errors. This version has been corrected. Please see Corrigendum (http://dx.doi.org/10.1080/09298215.2016.1198463).

computational problems that have been addressed by tonal pitch spaces we can highlight key estimation (Chew, 2000, 2008; Temperley, 2001; Bernardes et al., 2016), harmonic change detection (Harte, Snadler, & Gasser, 2006; Peiszer, Lidy & Rauber, 2008), automatic chord recognition (Mauch, 2010), and algorithmic-assisted composition (Behringer & Elliot, 2009; Gatzsche, Mehnert, & Stöcklmeier, 2008; Bernardes et al., 2015).

Following research into tonal pitch spaces, we present the Tonal Interval Space, a new tonal pitch space inspired by the *Tonnetz*, Chew's (2000) Spiral Array, and Harte et al.'s (2006) 6-dimensional (6-D) Tonal Centroid Space. We describe the mathematical formulation of the Tonal Interval Space and we discuss properties of the space related to music theory. The innovations introduced in this paper constitute a series of controlled distortions of the chroma space calculated as the weighted Discrete Fourier Transform (DFT) of normalized 12-element chroma vectors, in which we can measure the proximity of multi-level pitch configurations and their level of consonance.

Primarily, our approach extends the Tonnetz, as well as the work of Chew (2000, 2008) and Harte et al. (2006), in four fundamental aspects. First, it offers the ability to represent and relate pitch configurations at three fundamental levels of Western tonal music, namely pitch, chord and region within a single space. Second, we compute the space by means of the DFT and furthermore demonstrate how Harte et al.'s 6-D space can also be calculated in this way. Third, it allows the calculation of a tonal pitch consonance indicator. Fourth, it projects pitch configurations that have a different representation in the chroma space as unique locations in our space—thus expanding the Harte et al.'s 6-D space to include all possible intervallic relations.

The remainder of this paper is structured as follows. In Section 2, we review the problems and limitations of existing tonal pitch spaces. In Section 3, we detail the three most related tonal pitch spaces to our work, which form the basis of our approach. In Section 4, we describe the computation of Tonal Interval Vectors (TIVs) that define the location of pitch configurations in a 12-dimensional (12-D) tonal pitch space. In Section 5, we detail the representation of multi-level pitch configurations in the space as well as the implications of the symmetry of the DFT for defining a transposition invariant space. In Section 6, we detail distance metrics used in the Tonal Interval Space. In Section 7, we describe a strategy to adjust the distances among pitch configurations in the 12-D space in order to better represent music theory principles. In Section 8, we discuss the relations among different pitch configurations on three fundamental tonal pitch levels, namely pitch classes, chords, and regions and we compare the effect of different DFT weightings on the measurement of consonance. Finally, in Sections 9 and 10 we reflect on the original contributions of our work, draw conclusions and propose future directions.

2. Tonal pitch spaces: existing approaches and current limitations

The relations among tonal pitch structures, fundamental to the study of tonal pitch spaces, have been a research topic extensively investigated in different disciplines including music theory (Schoenberg, 1969; Weber, 1817–1821), psychology (Deutsch, 1984), psychoacoustics (Parncutt, 1989), and music cognition (Krumhansl, 1990; Longuet-Higgins, 1962; Shepard, 1982). Different models and interpretations of the same phenomena have been presented in these disciplines. We argue that their discrepancy is due to historical, cultural, and aesthetic factors. Therefore, tonal pitch spaces cannot be disassociated from the context where they have been presented, and, more importantly, their understanding requires exposure to tonal schemes (Deutsch, 1984).

Even though recent cognitive psychology research has managed to reduce confounding factors and offer a more general view on the subject of perceptual proximity of tonal pitch (Krumhansl, 1990), it is not averse to the idiosyncratic factors that regulate listening expectancies within the Western tonal music context. For example, many musical idioms that exist at the edge of tonality are clearly misrepresented by tonal spaces resulting from empirical studies, such as the post-romantic works of Richard Strauss and Gustav Mahler (Kross, 2006). Therefore, it is important to bear in mind that, whichever applied method, a tonal space is only a partial explanation of the entire Western tonal music corpus.

Given the limitations of tonal spaces to provide a universal explanation for the cognitive foundations of pitch perception, related research must necessarily clarify their basis, applied methodology, and most importantly their limitations. For the purpose of this work we follow Lerdahl (2001) and most cognitive psychological studies in the area, which position themselves in the extrapolation of 'hierarchical relations that accrue to an entire tonal system beyond its instantiation in a particular piece' (Lerdahl, 2001, p. 41). In other words, we are concerned with 'tonal hierarchy' that diverges from the concept of 'event hierarchy' (Bharucha, 1984) in the sense that basic tonal structures of the first apply to the majority of Western tonal music rather than a specific response to a particular style or composer's idiom.

Tonal music structures result from the interaction of several levels of pitch configuration, most importantly pitch, chords and regions (in increasing order of abstraction). In the resulting tonal hierarchy, the upper levels embed lower ones and all levels are inter-dependent. Therefore, as Lerdahl (2001) claims, a tonal pitch space must account for the proximity of individual pitches, chords and regions in the same framework, as well as explain their interconnection.

In his Tonal Pitch Space theory, Lerdahl (2001) contextualizes all low-level pitch configurations with top-level regions by representing pitch classes, chords, and regions according to a similar method and all in the same space. Nevertheless, in Lerdahl's space, in order to represent low-level pitch configurations, we must define the top-level region(s) to measure distances among their lower level pitch configurations. Therefore, in order to measure the distance between two chords, for example, we must define their region(s) in advance. Despite this compelling solution, Lerdahl's theory cannot be used in contexts where the regional level is unknown, such as in Music Information Retrieval (MIR) problems like the automatic estimation of keys and chords from a musical input. Therefore, we strive for a model that explains all fundamental tonal pitch levels in a single space, without the need to define *a priori* information.

Another commonly raised issue in the tonal pitch space literature, particularly when discussing tonal spaces grounded in music theory, is their symmetry, which does not equate with how humans perceive pitch distances (Krumhansl, 1990, pp. 119–123). However, the cyclical nature of the tonal system embeds operations like transposition, which naturally create symmetrical spaces. By disregarding the cyclical nature of the tonal system, we risk lacking an explanation for some of its most fundamental operations. Additionally, the human ability to understand, abstract, and group pitch contours invariant to their key or transposition factor stresses the importance of relational distances that account for these operations, resulting in symmetrical pitch space organizations (Shepard, 1982).

Consonance and dissonance, so closely related to the perception of pitch proximity and musical tension, are poorly addressed in all theories supporting tonal pitch spaces. Consonance and dissonance are at best implicitly considered in Lerdahl's (1988, 2001) Tonal Pitch Space, but never explicitly modelled (or measurable) as a property of the space. Similarly, Krumhansl's (1990, pp. 59–60) analysis comparing Krumhansl and Kessler's (1982) 24 major and minor key profiles with several ratings of consonance and dissonance show poor results for intervals formed within the minor keys.

3. The *Tonnetz* and its derivations

The Tonnetz is a planar representation of pitch relations first attributed to the eighteenth-century mathematician Leonhard Euler (Cohn, 1998). In its most traditional representation, the Tonnetz organizes (equal-tempered) pitch on a conceptual plane according to intervallic relations, favouring perfect fifths, major thirds and minor thirds (see Figure 1). Fifths run horizontally from left to right, minor thirds run diagonally from bottom left to top right, and major thirds run diagonally from top left to bottom right.

Despite its original basis as a pitch class space, the Tonnetz has been extensively used as a chordal space since the 19th century by music theorists such as Riemann and Oettingen and more recently by neo-Riemannian music theorists (Cohn, 1997; Hyer, 1995; Lewin, 1987). Chords are represented on the Tonnetz as patterns formed by adjacent pitches, whose shapes are constant for chords with the same quality. For example, major triads always form a downward pointing triangle, whereas minor triads always form an upward pointing triangle (see Figure 1).

Music theorists following the Riemannian tradition adopted the Tonnetz to explain significant tonal relationships between harmonic functions, which are near one another in the Tonnetz (Cohn, 1998). For example, the dominant and the subdominant chords are at close distances on either side of the chord of the tonic in a given key. In Figure 1, if we draw a horizontal line traversing the centre of the C major region tonic (**I**) we find its dominant (**V**) and subdominant (**IV**) chords in the neighbourhood and its relative (C minor triad), mediant (**iii**), and submediant (**vi**) chords in edge-adjacent triangles. Moreover, in the Tonnetz, chord distances also equate with the number of common tones. The closer chord configurations are, the greater their number of common tones. In addition to the large amount of music theory literature on the Tonnetz, Krumhansl (1998) presented experimental support for the psychological reality of one of its most important theoretical branches, the neo-Riemannian theory.

Various derivations and models of the Tonnetz have been proposed. Of interest here are those that have a mathematical formulation and that can be computationally modelled, notably Chew's (2000) Spiral Array and Harte et al.'s (2006) 6-D space. Chew's Spiral Array results from wrapping the Tonnetz into a tube in which the line of fifths becomes a helix on its surface and major third intervals are directly above each other. Chew's model allows chords and keys to be projected into the interior of the tube by the centre of mass of their constituent pitches.

The spatial location of pitches on the Spiral Array ensures that some pitch configurations understood as perceptually related within the Western tonal music context correspond to small Euclidean distances. That is, pitch distances are minimized for intervals that play an important role in tonal music, such as unisons, octaves, fifths and thirds. These distances result from the helix representation of pitch locations in the Tonnetz and from further defining the ratio of height to diameter, akin to stretching out a spring coil. The Spiral Array has been applied to problems such as key estimation (Chew, 2000) and pitch spelling (Chew & Chen, 2003) from music encoded as symbolic data.

Following Chew's research, Harte et al. (2006) proposed a tonal space that projects pitch configurations encoded as 12-element chroma vectors to the interior of a 6-D polytope visualized as three circles. Inter pitch-class distances in the 6-D space mirror the spatial arrangement for the perfect fifth, major thirds, and minor thirds of the Tonnetz, weighted in a similar fashion to Chew's Spiral Array to favour perfect fifths and minor thirds over major thirds. The fundamental difference from Chew's Spiral Array is the possibility to represent harmonic information in a single octave by invoking enharmonic equivalence. Distances between pitch configurations with variable numbers of notes are represented in the space by the centroid of their component pitches, whose distances emphasize harmonic changes in musical audio (Harte et al., 2006). Additionally, the 6-D tonal space has been applied in

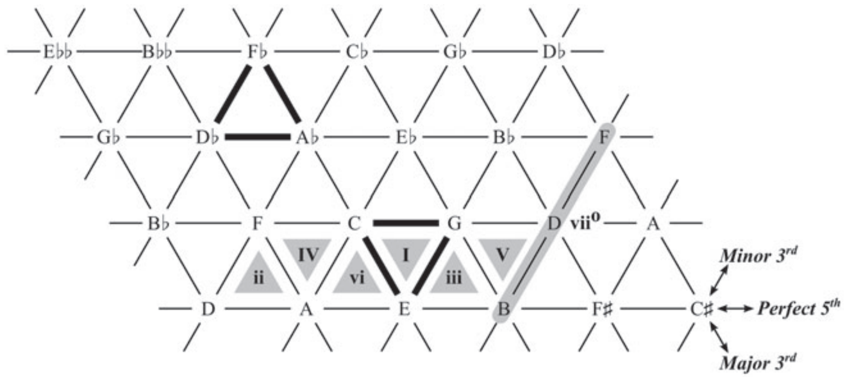

Fig. 1. Representation of the Tonnetz or harmonic network, in which triangular heavier strokes emphasize major/minor triads' formation and shaded areas the complete set of diatonic triads within the C major region—represented by their degree in Roman numerals.

a variety of MIR problems, including chord recognition (Lee, 2007), key estimation (Lee & Slaney, 2007) and structural segmentation (Peiszer et al., 2008).

In the following section, we introduce the Tonal Interval Space which inherits features from the Tonnetz and its derivative spaces (Chew's Spiral Array and Harte et al.'s 6-D space), concerning the organization of pitch classes. We extend Harte et al.'s 6-D space by including all intervallic relationships, reinforcing and controlling the contribution of each interval in the space according to empirical consonance and dissonance ratings (Hutchinson & Knopoff, 1979; Kameoka & Kuriyagawa, 1969; Malmberg, 1918). This allows us to measure the interpreted proximity of pitch configurations within Western tonal music at various levels of abstraction as well as measuring their level of consonance in a single space.

4. Tonal interval space

The Tonal Interval Space maps 12-D chroma vectors to complex-valued TIVs with the DFT.[1] On the one hand, the chroma vector can be used to represent different levels of pitch configurations such as pitches, chords and regions. On the other hand, Fourier analysis has been widely used to explore the harmonic relations between pitch classes, primarily to investigate intervallic differences between two pitch class sets and expand on the notion of maximal evenness (Amiot, 2013; Amiot & Sethares, 2011; Callender, 2007; Clough & Douthett, 1991; Lewin, 2001; Quinn, 2006, 2007) and to a lesser extent tonal pitch relations (Bernardes et al., 2015; Yust, 2015). In this paper, we explore the effect of all coefficients of the DFT of chroma vectors, including coefficients discarded by Harte et al. (2006) towards enhancing the description of tonal pitch and the computation of a tonal pitch consonance indicator.

[1]The use of the DFT in the context of our work was inspired by Ueda, Uchiyama, Nishimoto, Ono and Sagayama (2010), who identified a correspondence between the DFT coefficients of a chroma vector and Harte et al.'s (2006) 6-D space.

4.1 Chroma vectors

In this work, we restrict our analysis of musical notation to symbolic representations, and hence we consider chroma vectors $c(n)$ which express the pitch class content of pitch configurations as binary activations in a 12-element vector. Each element corresponds to a pitch class of the equal-tempered chromatic scale. The chroma vector $c(n)$ in Table 1 represents the C major chord, so it activates pitch classes [0, 4, 7] with the value 1. Table 1 supposes enharmonic and octave equivalence characteristic of equal tempered tuning. There is no information about pitch height encoded in $c(n)$. Consequently, the octave cannot be represented by $c(n)$ with binary encoding because all the octaves are collapsed into one.

The chroma vector $c(n)$ allows the representation of multi-level pitch configuration by simply indicating the presence of the respective pitch classes. For example, for the pitch class, C is [0], for the G major chord it is [2, 7, 11], and for the diatonic C major scale (or diatonic scale of A harmonic minor) it is [0, 2, 4, 5, 7, 9, 11].

The chroma vector $c(n)$ occupies a 12-D space independently of the pitch configuration it represents. However, the geometric properties of the space spanned by the chroma vector do not capture harmonic or musical properties of the pitch configurations that it represents. In other words, chroma vectors $c(n)$ that represent perceptually similar harmonic relations are not necessarily close together in the space. For example consider the following three dyads: a minor second [0, 1], a major third [0, 4] and a perfect fifth [0, 7]. While all three share a single pitch class [0] and the Euclidean distance between all of their chroma representations is the same, from a perceptual standpoint, the minor second is perceptually further from the other two. The DFT maps chroma vectors to TIVs into a space that exhibits useful properties to explore the harmonic relationships of the tonal system, which we detail in Section 8.

4.2 Tonal interval vectors

TIVs $T(k)$ are calculated as the DFT of the chroma vector $c(n)$ as follows

Table 1. Chroma vector $c(n)$ representation of the C major chord.

	Chroma vector $c(n)$											
Position n	0	1	2	3	4	5	6	7	8	9	10	11
Pitch class	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
Value	1	0	0	0	1	0	0	1	0	0	0	0

$$T(k) = w(k) \sum_{n=0}^{N-1} \bar{c}(n) e^{-\frac{j2\pi kn}{N}}, \quad k \in \mathbb{Z} \text{ with } \bar{c}(n) = \frac{c(n)}{\sum_{n=0}^{N-1} c(n)}, \tag{1}$$

where $N = 12$ is the dimension of the chroma vector and $w(k)$ are weights derived from empirical dissonance ratings of dyads used to adjust the contribution of each dimension k of the space, which we detail at length in Section 7. $T(k)$ uses $\bar{c}(n)$, which is $c(n)$ normalized by the DC component $T(0) = \sum_{n=0}^{N-1} c(n)$ to allow the representation of all levels of tonal pitch represented by $c(n)$ in the same space. In doing so, $T(k)$ can be compared amongst different hierarchical levels of tonal pitch.

Equation 1 can be interpreted from the point of view of Fourier analysis or complex algebra. The Fourier view is useful to visualize TIVs and interpret k as musical intervals, whereas the algebra view (explained in Section 6) is used to define objective measures that capture perceptual features of the pitch sets represented by the TIVs. The Fourier view interprets $T(k)$ as a sequence of complex numbers with $k \in \mathbb{Z}$. When $0 \leq n \leq 11$, k is usually set $0 \leq k \leq 11$. In practice, $1 \leq k \leq 6$ for $T(k)$ since the coefficients for $7 \leq k \leq 12$ are equal to $T(k)$ for $1 \leq k \leq 6$ because of the symmetry properties of the DFT (Oppenheim, Schafer, & Buck, 1989). In this section, $T(k)$ is represented as magnitude $|T(k)|$ versus k and phase $\varphi(k)$ versus k. For each index k, we have

$$|T(k)| = \sqrt{\Re\{T(k)\}^2 + \Im\{T(k)\}^2} \quad 1 \leq k \leq 6, \tag{2}$$

$$\varphi(k) = \tan^{-1} \frac{\Im\{T(k)\}}{\Re\{T(k)\}} \quad 1 \leq k \leq 6, \tag{3}$$

where $\Re\{T(k)\}$ and $\Im\{T(k)\}$ denote the real and imaginary parts of $T(k)$ respectively.

The Tonal Interval Space uses the interpretation in Table 2, which we show in Figure 2, using a strategy borrowed from Harte et al. (2006) to depict their 6-D space. Each circle in Figure 2 corresponds to $T(k)$ when $\leq k \leq 6$ in Equation 1. The circle representing the intervals of m2/M7 has the real part of $T(1)$ on the x axis and the imaginary part of $T(1)$ on the y axis and so on. The integers around each circle represent $0 \leq n \leq N-1$ for $N = 12$, corresponding to the positions in the chroma vector $c(n)$. A fixed k in Equation 1 generates $N = 12$ points equally spaced by $\varphi(k) = -2\pi k/N$. Both in Table 2 and Figure 2, a musical nomenclature is adopted to denote each of the DFT coefficients that arise from the interpretation of these points as musical intervals. The musical interpretation assigned to each coefficient corresponds to the musical interval that is furthest from the origin of the plane (i.e. the centre of the circles shown in Figure 2). For $k = 1$ and $k = 5$, the furthest musical interval from the centre is formed between adjacent positions. For $k = 2$, $k = 3$, $k = 4$ and $k = 6$, the furthest interval from the centre is formed between overlapping positions.

5. Multi-level pitch configurations and transposition

Section 4.1 demonstrated that the chroma vector $c(n)$ can represent multi-level pitch configurations as the sum of $c(n)$ for each single pitch class. For example, the chroma vector of the C major chord can be obtained as the sum of the chroma vectors of its constituent pitch classes C, E and G. Mathematically, $c_{1,2,3}([0, 4, 7]) = c_1([0]) + c_2([4]) + c_3([7])$. Due to the linearity of the DFT, multi-level pitch configurations in the Tonal Interval Space can be represented as a linear combination of the DFT of its component pitch classes. Mathematically, $T_{1,2,3}(k) = T_1(k) + T_2(k) + T_3(k)$.

Figure 2 illustrates $T(k)$ for the C major chord as a *convex* combination of $T(k)$ for its component pitch classes. Convex combinations are linear combinations $\sum_{k=1}^{K} \alpha(k)T(k)$, where the coefficients $\alpha(k)$ are non-negative (i.e. $\alpha(k) \geq 0$) and $\sum_k \alpha(k) = 1$. Geometrically, a convex combination always lies within the region bounded by the elements being combined. So the convex combination of TIVs lies inside the shaded regions shown in Figure 2 due to the normalization of $\bar{c}(n)$ in Equation 1. These regions can be obtained by connecting the adjacent TIVs of isolated pitch classes.

An important feature of Western tonal music arising from 12 tone equal-tempered tuning is the possibility to modulate across regions. This attribute establishes hierarchies in tonal pitch, in which low-level components relate to, and are commonly defined by their regional level. For example, we commonly define the chords formed by the diatonic pitch set of the C major region by the function they play within that region, such as the chords of the tonic, sub-dominant, dominant, etc. Perceptually, Western listeners also understand interval relations in different regions as analogous (Deutsch, 1984). For example, the intervals from C to G in C major and from C# to G# in C# major are perceived as equivalent. As Shepard (1982) claims, this theoretical and perceptual aspect of Western tonal music is an important attribute that should be modelled by tonal pitch spaces, which 'must have properties of great regularity, symmetry, and transformational invariance' (p. 350). Briefly, a tonal space must be transposition invariant.

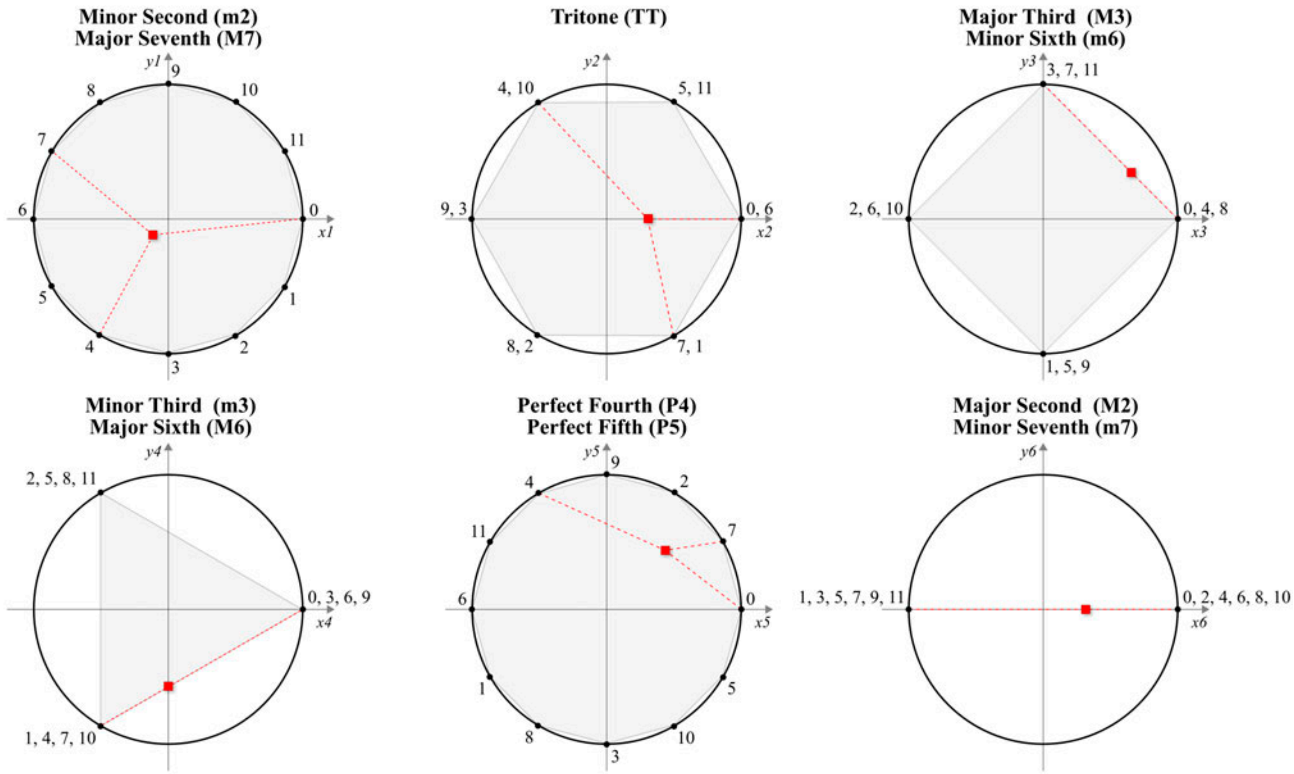

Fig. 2. Visualization of the TIV for the C major chord (pitch classes 0, 4 and 7) in the 12-D Tonal Interval Space. Each circle corresponds to the coefficients of $T(k)$ labelled according to the complementary musical intervals they represent. The TIVs of isolated pitch classes lie on the circumference and the TIV corresponding to the linear combination lies inside the region bounded by the straight lines connecting the points. Shaded grey areas denote the regions that TIVs can occupy for each circle.

Table 2. Intervallic interpretation of k for $T(k)$.

Position k	1	2	3	4	5	6
Steps n Musical interval	Adjacent m2/M7	Overlap TT (A4/D5)	Overlap M3/m6	Overlap m3/M6	Adjacent P4/P5	Overlap M2/m7

In the Tonal Interval Space, transpositions by p semitones result in rotations of $T(k)$ by $\varphi(n) = -2\pi kp/N$ radians. Transpositions of $c(n)$, which by definition are circular in the chroma domain, are represented as $c(n-p)$. So, transposing C by $p = 7$ results in G and by $p = 12$ results in C. Using the properties of the Fourier transform (Oppenheim et al., 1989), the pair $c(n) \overset{\mathcal{F}}{\Rightarrow} T(k)$ becomes $c(n-p) \overset{\mathcal{F}}{\Rightarrow} T(k)e^{-\frac{j2\pi k}{N}p}$, where $\mathcal{F}$ represents the DFT. Denoting $T_p(k)$ as the TIV of $c(n-p)$ we have

$$T_p(k) = |T(k)|e^{-\frac{j2\pi(k+p)}{N}}. \tag{4}$$

Hence, any transposition $c(n-p)$ resulting in $T_p(k)$ has the same magnitude $|T(k)|$ as the original sequence $c(n)$ and a linear phase component $e^{-\frac{j2\pi k}{N}p}$. Figure 3 illustrates the rotation of the TIV of the C major chord by one semitone.

6. Distance metrics in the tonal interval space

This section illustrates the properties of the Tonal Interval Space which rely on the complex algebra view of Equation 1, where $T(k) \in \mathbb{C}^M$; $M = 6$. Here, $T(k)$ is interpreted as a 6-D complex-valued vector in the space spanned by the Fourier basis when $1 \leq k \leq 6$. Note that 6 complex dimensions correspond to 12 real dimensions because the real and imaginary axes are orthogonal. Using the norm L_2 in $\mathbb{C}^M$, we can define the inner product between $T_1(k)$ and $T_2(k)$, the norm of $T_1(k)$, and the Euclidean distance between $T_1(k)$ and $T_2(k)$ as follows

$$T_1(k) \cdot T_2(k) = ||T_1(k)||\,||T_2(k)|| \cos\theta = \sum_{k=1}^{M} T_1(k)\overline{T_2(k)}, \tag{5}$$

$$d\{T_1(k), T_2(k)\} = \sqrt{||T_1(k)|| - ||T_2(k)||}$$

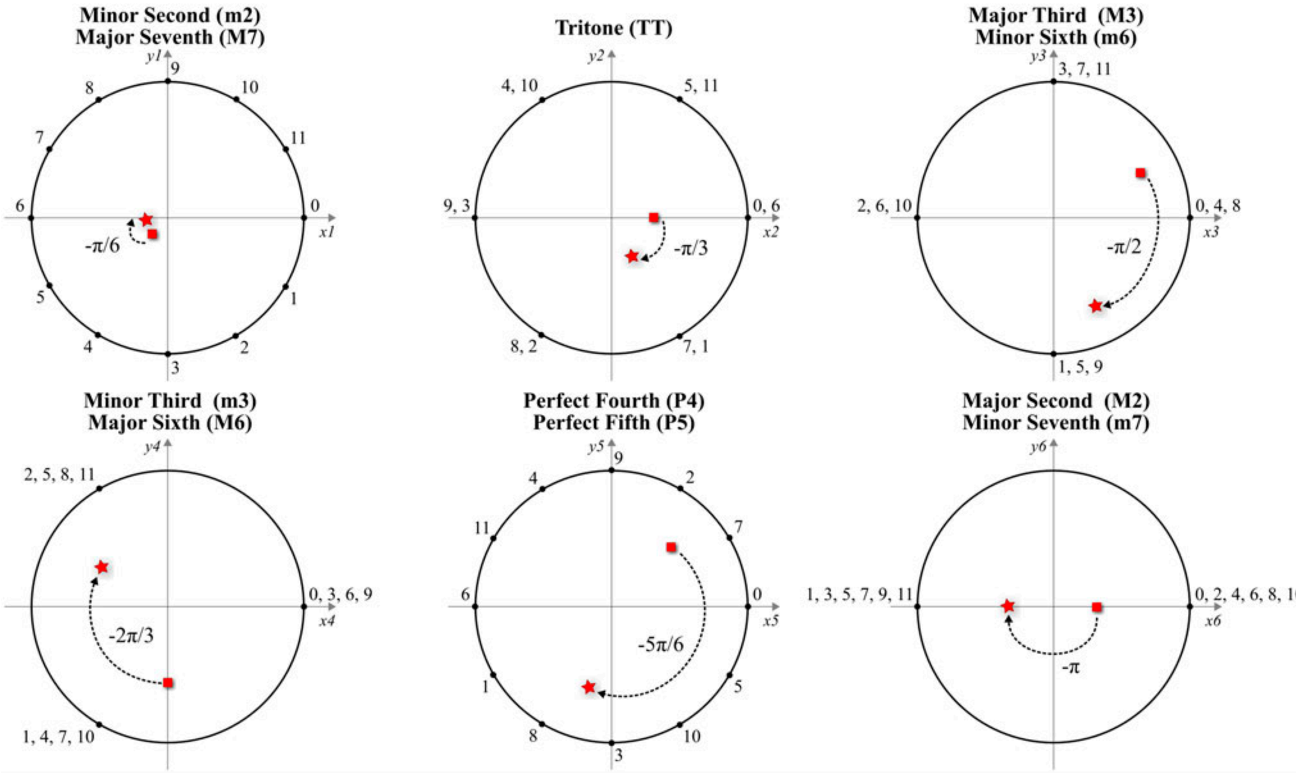

Fig. 3. Visualization of the C major triad (pitch classes 0, 4 and 7—represented as a square) and the rotation of its TIV to transpose it one semitone higher (i.e. pitch classes 1, 5 and 8—represented as a star).

$$= \sqrt{\sum_{k=1}^{M} |T_1(k) - T_2(k)|^2}, \quad (6)$$

$$||T_1(k)|| = \sqrt{T_1(k) \cdot T_1(k)} = \sqrt{\sum_{k=1}^{M} |T_1(k)|^2}, \quad (7)$$

where $M = 6$ is the dimension of the complex space, θ is the angle between $T_1(k)$ and $T_2(k)$, and $\overline{T_2(k)}$ denotes the conjugate transpose of $T_2(k)$. Equation 5 is the inner product and Equation 6 is the Euclidean distance between $T_1(k)$ and $T_2(k)$. Equation 7 is the norm of $T_1(k)$, which can also be calculated as the Euclidean distance from the centre of the Tonal Interval Space $\vec{0}$ as $||T_1(k)|| = d\{T_1(k), \vec{0}\}$.

We use Equations 5 to 7 within the Tonal Interval Space in order to measure tonal pitch relations and consonance using complex algebra. The musical interpretation of the algebraic properties are detailed at length in Section 8.

7. Improving the perceptual basis of the space

Following Chew (2000) and Harte et al. (2006), we apply a strategy to adjust pitch class distances in our space. To this end, we apply weights $w(k)$ to each circle when calculating $T_1(k)$ using Equation 1. By controlling the weights we can regulate the contribution of the musical intervals associated with each of the DFT coefficients, as described in Section 4. Specifically, we intend to use the weights as a means to allow the computation of consonance of pitch configurations in the Tonal Interval Space, which we calculate as the norm of a TIV (see Section 6).

We rely on two complementary sources of information to derive the set of weights. First, the set of composite consonance ratings of dyads (Huron, 1994), as shown in Table 3 and second, the relative ordering of triads according to increasing dissonance (Cook, Fujisawa, & Konaka, 2007): {maj/min, sus4, aug, dim}.[2] Our goal is to find a set of weights which both maximizes the linear correlation with Huron's composite consonance ratings of dyads and preserves Cook et al.'s relative ordering of triads.[3] While the search for weights can be considered a multidimensional optimization problem, by applying two simplifying constrains we can perform an exhaustive brute force search and thus consider all possible combinations of weights. In this way, we can guarantee a near optimal result subject to our constraints.

To allow a computationally tractable search for weights, we restrict the properties of the weights as follows: we allow only integer values in a defined range, such that each $w(k)$

[2]Because the pitch configurations for major and minor triads contain identical relative intervals when represented as chroma vectors, the Tonal Interval Space cannot disambiguate them, hence we must consider them equally ranked.

[3]While Roberts (1986) provides consonance ratings of triads, these were obtained from an experimental design that relied heavily on a preceding musical context and not listener judgements of isolated triads. Therefore we do not attempt to directly incorporate these absolute ratings when determining the weights $w(k)$.

Table 3. Composite consonance ratings based on normalized data from Malmberg (1918), Kameoka and Kuriyagawa (1969), and Hutchinson and Knopoff (1979) (as presented in Huron, 1994).

Interval class	Consonance
m2/M7	−1.428
M2/m7	−0.582
m3/M6	0.594
M3/m6	0.386
P4/P5	1.240
TT	−0.453

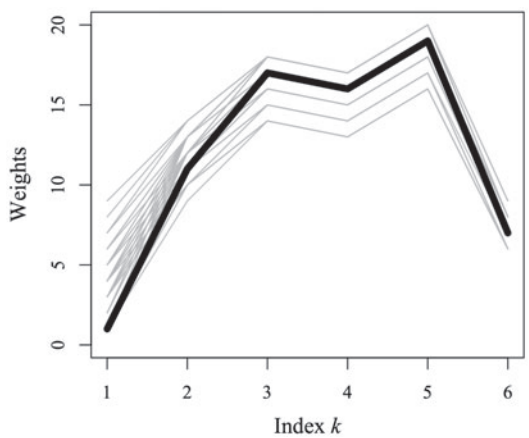

Fig. 4. The set of weights that maximize the linear correlation with the composite consonance ratings of dyads as shown in Table 3 while simultaneously preserving the relative ordering of triads shown in Table 7. The bold black line corresponds to the set of weights $w(k)$ used in Equation 1.

can only take values between 1 and 20. Consequently this creates a secondary constraint that the largest weight (i.e. the most important interval) can be, at most, 20 times the smallest (i.e. the least important interval). Given that each individual weight $w(k)$ can take any value between 1 and 20 independently of all the others, this provides a total of $B = 20^6 - 1$ possible combinations of weights (i.e. 64 million) ranging from $w_1(k) = \{1, 1, 1, 1, 1, 1\}$ to $w_B(k) = \{20, 20, 20, 20, 20, 20\}$. For simplicity, we do not discard any sets of weights that are trivially related to one another in terms of scalar multiples.

For each set of weights $w_b(k)$ we first calculate the corresponding TIV, $T_b(k)$, using Equation 1 for each dyad interval in Table 3 and the following set of triads {maj, sus4, aug, dim}. We then calculate the consonance (i.e. the distance to the centre of the Tonal Interval Space) as the magnitude $||T_b(k)||$ using Equation 7. We then measure the linear correlation to Huron's dyad consonance ratings, and verify the ordering of the triads' consonance according to Cook et al. From the complete set of $20^6 - 1$ combinations of weights, we found 46 solutions (each of which is plotted in Figure 4) that resulted in a linear correlation greater than 0.995 and preserved the triad consonance ordering. Given the inherent similarity in shape of the different sets of weights, we do not believe the choice over exactly which set of weights to be critical. However, we ultimately selected the weights with the greatest mutual separation between the triads according to consonance, thus $w(k) = \{2(m2/M7), 11(TT), 17(M3/m6), 16(m3/M6), 19(P4/P5), 7(M2/m7)\}$.

8. Musical properties of the multi-level Tonal Interval Space

Pitch configurations are separated in the 12-D tonal pitch space by spatial and angular distances whose metrics were presented in Section 6. In this section, we discuss how these distances translate into musical properties within the most salient hierarchical layers of the tonal system from lower to higher levels of abstraction, i.e. starting with the spatial relations between pitch classes, then chords, and finally keys or regions.

The musical properties of the Tonal Interval Space can be split into two major groups. The first is detailed in Section 8.1 and reports the ability of the space to place pitch configurations that share harmonic relations close to one another. The second is reported in Section 8.2 and explains how $||T(k)||$ can be used as a measure of consonance.

8.1 Perceptual similarity among multi-level pitch configurations

Proximity in the Tonal Interval Space equates with how pitch structures are understood within the Western tonal music context rather than objective pitch frequency ratios. In other words, the closeness between pitch classes in our space corresponds to interpreted proximity between pitch classes as used in the context of Western tonal music rather than distances on a keyboard. For example, pitches placed at a close distance on the keyboard, such as C and C#, are quite distant in our space. In fact, objective frequency ratios among pitch classes are immediately misrepresented in the chroma vector by collapsing all octaves into one, and even further distorted in the weighted DFT of chroma vectors expressed by the TIVs.

Similar to Harte et al.'s (2006) 6-D space, the resulting structure of the Tonal Interval Space inherits the pitch organization of the Tonnetz by wrapping the plane into a toroid, see Harte et al. (2006) for a detailed explanation and illustration of this operation. Therefore, in the Tonal Interval Space, as in the Tonnetz, the proximity of dyads using both the angular and Euclidean distances computed by Equations 5 and 6 are ranked as follows: unisons; perfect fourths/fifth; minor thirds/major sixths; major thirds/minor sixth; tritone (augmented fourth or diminished fifth); major second/minor seventh; and finally, minor second/major seventh (see Table 4). Additionally, as a result of the symmetry of the Tonal Interval Space imposed by the DFT, complementary intervals are at equidistant locations (see Section 5).

At the chordal level, the major/minor triad formation, which groups close pitch classes in the representations, equates with

Table 4. Angular and Euclidean distances of complementary dyads in the Tonal Interval Space presented from left to right in descending order of consonance.

Distance	P1	P4/P5	m3/M6	M3/m6	TT	M2/m7	m2/M7
Angular θ (in rad)	0.00	1.39	1.49	1.60	1.78	1.80	1.98
Euclidian d	0.00	42.13	44.59	47.18	51.15	51.50	54.88

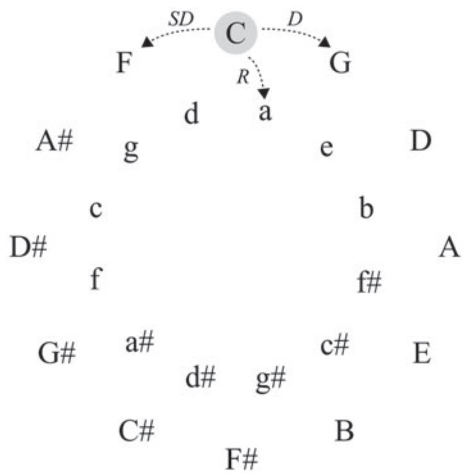

Fig. 5. 2-D visualization of the interkey distances in the Tonal Interval Space amongst all major and minor regions using multidimensional scaling (De Leeuw & Mair, 2009).[4] The neighbour dominant (D), subdominant (SD), and relative (R) regions of C major are emphasized.

the triads formed and commonly highlighted in the Tonnetz (see Figure 1). Therefore, motions between adjacent triads in the Tonal Interval Space indicate a chord progression that maximizes the number of common tones while minimizing the displacement of moving voices (known as voice-leading parsimony). For all regions we find, in the neighbourhood of the chord of the tonic, the mediant and submediant chords, which each share two pitch classes with the tonic. Neo-Riemannian theorists refer to these motions as primary transformations (Cohn, 1997, 1998). Motions between chords that share fewer pitch classes are placed further apart in the space.

Within the context of Western tonal music, we can also say that close harmonic functions are depicted in our space as chord substitutions. Typical harmonic progressions in Western tonal music remain at relatively close distances but are not explicitly minimized in the space. Briefly, the chordal level in our space minimizes distances for common-tone chord progressions, which commonly substitute themselves, rather than typical harmonic sequences.

Therefore, the Tonal Interval Space shows great potential to explore voice-leading parsimony (as applied in Bernardes et al., 2015) and offers the possibility to explore formal transformations that have been derived from Riemann's fundamental harmonic theory (Cohn 1997, 1998; Hyer, 1995; Kopp, 1995; Lewin, 1982, 1987, 1992; Mooney, 1996).

Interkey distances in the Tonal Interval Space result in two concentric layers which position keys by intervals of fifths. The outer layer (corresponding to vectors with larger magnitude) contains the circle of fifths for all major keys and an inner layer (corresponding to vectors with smaller magnitude) contains the circle of fifths for all minor keys. Figure 5[4] illustrates interkey distances on a 2-D space. There the spatial proximity of each key to its dominant, subdominant and relative keys, corresponds to our expectation of the proximity between the 24 major and minor keys and adheres to Schoenberg's (1969) map of key regions, which is a geometrical representation of proximity between keys (Lerdahl, 1988, 2001).

The next consideration concerns the degree to which our space can explain the interconnection of the three tonal pitch levels, and particularly the relation of the lower abstraction levels with the top regional ones. This aspect is especially relevant within the Western tonal music context because our understanding of pitch classes and chords is dependent on their upper hierarchical levels (Krumhansl, 1990, pp. 18–21). Ideally, the three tonal pitch levels should interconnect and the distances among pitch classes, chords and regions should be meaningful.

In the Tonal Interval Space, the pitch class set of different diatonic regions occupies a compact neighbourhood. The same property applies to the set of diatonic triads within a region because their location is the convex combination of the $T(k)$ for its component pitch classes as explained in Section 5. Table 5 reinforces the validity of this assertion by showing the angular and Euclidean distances between all individual pitch classes from the C major and C harmonic minor regions. The set of diatonic pitch classes (in bold) of each region are at smaller distances than the remaining pitch classes. Due to the transposition invariance of the Tonal Interval Space, these results hold true for all remaining major and minor regions in our space.

[4]In order to illustrate distances among pitch configurations in the 12-D Tonal Interval Space, we use nonmetric multidimensional scaling (MDS) to plot it into a 2-dimensional plane. Shepard (1962) and Kruskal (1964) first used this method, which has been extensively applied to visualize representations of multidimensional pitch structures (Barlow, 2012; Krumhansl & Kessler, 1982; Lerdahl, 2001). Briefly, nonmetric MDS attempts to transform a set of n-dimensional vectors, expressed by their distance in the item-item matrix, into a spatial representation that exposes the interrelationships among a set of input cases. We use the `smacof` library from the statistical analysis package 'R' to compute dimensionality reduction using a nonmetric MDS algorithm. More specifically, we use the function `smacofSym`, with 'ordinal' type and 'primary' ties.

Table 5. Angular (θ) and Euclidean (d) distances between C major and C harmonic minor regions TIVs (labeled as upper and lower case 'c', respectively) and the entire set of pitch classes. The diatonic pitch class set of each region is presented in bold.

		C	C#	D	D#	E	F	F#	G	G#	A	A#	B
C	θ	**1.22**	2.12	**1.09**	2.12	**1.22**	**1.40**	2.02	**1.15**	1.97	**1.15**	2.01	**1.40**
	d	**30.9**	39.63	**29.50**	39.63	**30.91**	**32.79**	38.80	**30.07**	38.41	**30.07**	38.80	**32.79**
c	θ	**1.19**	2.24	**1.33**	**1.13**	1.97	**1.25**	2.06	**1.24**	**1.23**	1.98	1.78	**1.33**
	d	**30.58**	39.75	**31.95**	**31.62**	37.81	**31.14**	38.54	**31.13**	**30.99**	37.87	36.45	**32.02**

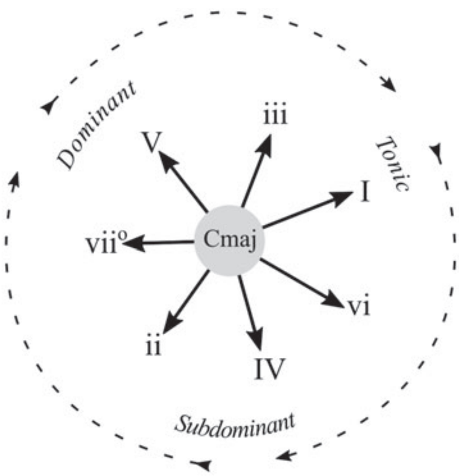

Fig. 6. 2-D visualization of the diatonic triads of the C major region in the Tonal Interval Space using nonmetric MDS. Riemann's harmonic categories (tonic, subdominant and dominant) are well represented in the space and typical motions between these are denoted by dashed lines.

Table 6. Consonance level of all interval dyads within an octave.

	P1	m2/M7	M2/m7	m3/M6	M3/m6	P4/P5	TT
Consonance	32.86	18.09	20.41	24.15	22.88	25.23	20.64

Finally, as illustrated in Figure 6, the diatonic set of chords around a key TIV is organized according to Riemman's categorical harmonic functions and distributed in roughly equal angular distances around its key centre. Chords common to more than one region, also referred to as pivot chords, are located at the edge of the regions. This allows the Tonal Interval Space to explain the modulation between keys or regions as these chords are typically used to smoothly transition between them. See (Bernardes et al., 2015) for a more comprehensive explanation of the angular distance between key TIVs and their diatonic chordal set and an application of this property to generate musical harmony and estimate the key of a musical input.

8.2 Measuring consonance

The Tonal Interval Space follows the pitch organization of the Tonnetz and expands this geometric pitch representation with the possibility to compute indicators of tonal consonance. Two important elements in Equation 1 allow the computation

Table 7. Consonance level of chords measured by our model (presented by increasing order of consonance).

	maj/min	sus4	dim	aug	min7	maj7	dom7
Consonance	20.36	19.77	18.64	18.38	17.46	16.35	15.88

of consonance, the normalization by $T(0)$ and the weights $w(k)$. The former was discussed in Section 4.2 and constrains the space to a limited area for all possible (multi-)pitch configurations that a chroma vector can represent. The latter was discussed in Section 7 and distorts the DFT coefficients to regulate the contribution of each interval according to empirical ratings of consonance. These two elements create a space in which pitch classes (at the edge of the space and furthest from the centre) are considered the most consonant configurations. A chroma vector $c(n)$ with all active elements will be located in the centre of space $\vec{0}$, which we consider the most dissonant. Within this range, the consonance of any pitch configuration can then be measured. Hence, we extrapolated the consonance measure of the TIV by the norm $||T(k)||$ given by Equation 7.

Due to the symmetry of the Tonal Interval Space, complementary intervals and transposition share the same level of consonance as indicated in Equation 4. In fact, 12 transpositions of $T(k)$ by $p = 1$ semitone creates a concentric layer of 12 instances with the same magnitude $||T(k)||$. Given this formulation, we present the level of consonance for all interval dyads within an octave in Table 6 and the consonance level of common triads and tetrads in tonal music in Table 7.

By comparing the values presented in Tables 5 and 6, we note that our consonance measure contradicts a limitation of the sensory dissonance models—one of the most popular models to measure innate aspects of consonance. As Huron emphasizes (cited in Mashinter, 2006), in sensory dissonance models adding spectral components *always* results in an increase of sensory dissonance. In the Tonal Interval Space, sonorities with fewer notes or partials may have a higher level of dissonance than sonorities with more notes or partials, depending on the level to which it 'fits' triadic harmony and tonal structures.

9. Evaluation and discussion

The consonance level modelled in the Tonal Interval Space constitutes an innovative aspect that has not been investigated in any other Tonnetz-derived spaces. While our method provides the possibility to compute a consonance indicator in the space by design, we now investigate whether other spaces are equally adept to that task. In particular, given the resemblance of the Tonal Interval Space with Harte et al.'s (2006) Tonal Centroid Space, we assess if the latter embeds properties for computing tonal pitch consonance. Additionally, we further assess the role of the weights $w(k)$ in the Tonal Interval Space by comparing the consonance measurement in a uniform version of the space. Both aforementioned spaces can be computed using Equation 1 by assigning different weights $w(k)$. In Harte et al.'s 6-D space $w_H(k) = \{0, 0, 1, 0.5, 1, 0\}$ whose non-zero weights correspond to the musical interpretation of major thirds/minor sixth, minor third/major sixth, and perfect fourth/perfect fifth, respectively (see Figure 3). In the uniform version of the Tonal Interval Space $w_U(k) = \{1, 1, 1, 1, 1, 1\}$. To investigate the behaviour of the spaces in measuring tonal pitch consonance, we will adopt the same consonance measure used in the Tonal Interval Space, computed by Equation 7, for all dyads and common triads.

To analyse the results we will use the Pearson correlation coefficient to compare the tonal pitch consonance indicators computed in the spaces with empirical ratings of dyads' consonance (used to build the model and shown in Table 3), and the ranking order of common triads' consonance derived from both listening experiments (Cook, 2012; Roberts, 1986) along with psychoacoustic models of sensory dissonance (Parcutt, 1989; Plomp & Levelt, 1965; Sethares, 1999). Our hypothesis is that, since we explicitly choose weights to control consonance, $w_H(k)$ and $w_U(k)$ will be less effective in highlighting the consonance of pitch configurations, respectively due to the exclusion of three intervals in $w_H(k)$ and the omission of any meaningful distortion of the weights in $w_U(k)$.

Figure 7 shows the correlation between the spaces under evaluation, to which we included our proposed Tonal Interval Space for the purpose of visual comparison. The correlation between empirical data and the uniform Tonal Interval Space ($r = -0.201$, $p = 0.703$) shows the DFT of chroma vectors carry no information about tonal consonance, reinforcing the positive impact of explicitly designing the weights in the Tonal Interval Space. The correlation between empirical data and Harte et al.'s 6-D space ($r = 0.741$, $p = 0.09$) shows that the space while positively correlated, is limited as an indicator of tonal consonance. In particular, this is shown in Figure 7 by the outlier corresponding to the consonance of the tritone interval in Harte et al.'s 6-D space, which is explicitly not modelled in their space.

We additionally assess how the consonance level of common triads measured in Harte et al.'s (2006) 6-D space and the uniform version of the Tonal Interval Space compares to empirical studies (Cook et al., 2007; Roberts, 1986) and psychoacoustic models of sensory dissonance (Parcutt, 1989; Plomp & Levelt, 1965; Sethares, 1999).[5] To this end, we

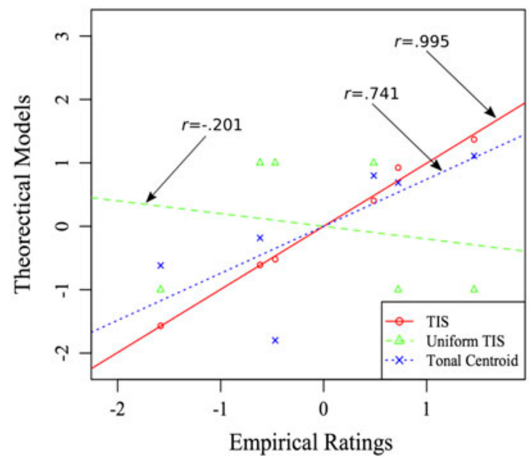

Fig. 7. Scatter plot exposing the correlation between empirical consonance ratings for complementary dyads (Huron, 1994) and their consonance level in three theoretical models: Tonal Interval Space (bold line), uniform Tonal Interval Space (dashed line) and Harte et al.'s (2006) 6-D space (dotted line). Plotted data is normalized to zero mean and unit variance for enhanced visualization.

compared the ranking order of common triads' consonance of the three theoretical models and both empirical ratings and psychoacoustic models of sensory dissonance. As shown in Table 8, the Harte et al.'s 6-D space and the uniform Tonal Interval Space fail to predict the relative consonance of common triads' consonance, further reinforcing the impact the weights and extended intervallic representations have on the Tonal Interval Space. Table 8 additionally shows that psychoacoustic models of sensory dissonance also fail to predict the relative dissonance of common triads as expressed by the results of the empirical listening tests.

While preserving the pitch organization of the Tonnetz, the Tonal Interval Space constitutes an extension of Tonnetz-derived spaces towards the possibility to compute a consonance indicator in the space. Furthermore, by expanding the number of dimensions in relation to similar spaces, and in particular to Harte et al.'s (2006) 6-D space, we obtain a finer definition of the intervallic content of chroma vectors, whose contribution we were able to fine-tune by adopting the set of weights $w(k)$. In doing so, our model not only ensures that all information from the chroma vector is retained in the TIV, but also guarantees that each TIV occupies a unique location in the 12-D Tonal Interval Space.[6] Both these properties are not found in any of the existing tonal pitch spaces. Finally, despite its larger number of dimensions and the increased complexity

[5] A major difference between the two empirical studies conducted relies on their population. While Roberts' (1986) study was conducted among Western listeners, Cook et al.'s (2007) study involved East Asian listeners. The population of both studies included individuals with and without music training. The remaining psychoacoustic-based models aim at measuring auditory roughness, which largely equates with sensory dissonance (Sethares, 1999).

[6] By guaranteeing uniqueness, our space avoids the overlap between relevant tonal pitch configurations as in the Harte et al.'s (2006) Tonal Centroid Space, such as the pair of dyads F#-B (P4) and D-G# (D5)—pitch classes [5, 11] and [2, 8])—and the D diminished seventh chord and the dyad D-A# (A4)—pitch classes [2, 5, 8, 11] and [2, 8].

Table 8. Ranking order of chord consonance based on Cook et al. (2007) comparing empirical data derived from listening experiments and theoretical models. 1 corresponds to the most consonant chord and 5 the most dissonant.

	Empirical ratings		Theoretical models					
			Sensory dissonance models			Tonal pitch spaces		
Chord quality	Roberts (1986)	Cook et al. (2007)	Plomp & Levelt (1965)	Parcutt (1989)	Sethares (1999)	Tonal Interval Space	Uniform Tonal Interval Space	Tonal Centroid (Harte et al., 2006)
major	1	**1**	2	2	2	**1**	1	1
minor	2	**2**	2	3	2	**1**	1	1
sus4	–	**3**	1	–	1	**2**	1	3
dim	3	**4**	5	4	4	**3**	1	4
aug	4	**5**	4	1	5	**4**	2	2

of the Tonal Interval Space in relation to similar spaces, we believe that using the DFT makes it particularly accessible to the music, signal processing, and MIR communities.

10. Conclusions and future work

In this paper we presented a 12-D Tonal Interval Space that represents pitch configurations by the location of Tonal Interval Vectors, which are calculated as the DFT of 12-element chroma vectors. A visualization of the 12-D space is provided by six circles, each representing a DFT coefficient, from which we devised a musical interpretation. The contribution of each DFT coefficient (or circles in the visualization) is then weighted according to empirical ratings of dyads consonance to improve the relationship among pitch configuration at the three most important levels of tonal pitch in Western music, i.e. pitches, chords and regions, as well as allowing the computation of a consonance indicator in the space.

While preserving the pitch organization and common-tone logic of the Tonnetz, our 12-D space expands its range of representable pitch configurations beyond major and minor triads. In relation to Chew's Spiral Array, the input of our space is more flexible in the sense that it allows the codification of any sonority representable as a chroma vector albeit subject to enharmonic equivalence. In relation to Harte et al.'s (2006) research, we expand their 6-D space to include all possible interval relations within one octave, and hence the ability to represent all pitch configurations by a unique location in space.

Two major indicators can be computed in the Tonal Interval Space. The first, explains the relation among pitch configurations in light of the Western tonal music theory principles by the angular and Euclidean distances among TIVs. Additionally, due to the possibility to represent all hierarchical levels of tonal music in the same space, given by the normalization strategy applied in Equation 1, we can equally compare and relate multi-level TIVs.

The second, and most innovative aspect of the Tonal Interval Space is the possibility to compute indicators of tonal consonance for multi-level pitch configurations as the norm of the TIVs. To the best of our knowledge, this attribute has not been considered in any other Tonnetz-derived spaces, nor any other tonal pitch spaces. By encoding all intervallic content of chroma vectors, distorted by both the DFT and weights derived from dyads and triads consonance, we enhance the pitch organization by allowing the measurement of consonance without disrupting the Tonnetz-like pitch organization.

Our goal in this paper was to present the Tonal Interval Space from a theoretical perspective, hence aspects concerning its scope and wider applicability are somewhat superficially treated. Nonetheless, the space has been successfully used in different application areas within the scope of generative music and MIR. In generative music, we have explored its potential to generate a corpus of chords related to a user-defined region (Navarro et al., 2015) as well as the possibility to smoothly transition (or modulate) between regions in real-time (Bernardes et al., 2015). In (Bernardes et al., 2015) we further explored the capabilities of the Tonal Interval Space to harmonize a given input using its ability to generate tonal harmony with consonance and perceptual relatedness as parameters. Additionally, in order to identify the region of the musical input we proposed a key induction algorithm which outperforms the current state of the art.

Despite the robustness of our consonance measurement in the context of Western tonal music, we are aware that our measure may fail to capture some aspects of consonance and dissonance, because it does not take into account the physical or physiological aspects of this phenomenon, which are directly related with frequency ratios among the partials of a sonority (Sethares, 1999). Despite these limitations, our consonance measure sheds some light on the future development of musical consonance models that consider both schemata learned culturally and innate physical and physiological principles.

Another limitation of our space is that it currently ignores the temporal dimension of music, or simply put, the order of musical events. Therefore, even though the perceived relation of tonal pitch events is known to depend on the order in which they are sounded (Krumhansl, 1990, pp. 121–123), we cannot yet account for that feature in our space due to its symmetry, which is inherent to Fourier spaces. On the other hand, the symmetry of the space imposed by the DFT is particularly relevant to create a transposition invariant space, seminal to tonal pitch structures. Additionally, we believe that many other mathematical properties of the DFT may have useful musical counterparts, and we plan to study these further in the future. Among them, we can highlight the capability to transform the Tonal Interval Space back to the chroma space by computing the inverse Fourier transform.

In future work, we aim to assess the level to which the Tonal Interval Space conforms to empirical judgments of tonal pitch relatedness, with the ultimate goal of improving the distances among multi-level TIVs. For example, despite the current possibility to compute the set of diatonic pitch classes of a given region, the distances among pitch classes and key TIVs do not express the goodness of fit of the pitch classes into that region.

Finally, the initial experiments reported here were conducted under very controlled conditions by manually encoding pitch configurations as binary chroma vectors. However, despite the possibility to represent musical audio (e.g. with chroma vectors calculated from audio signals), further tests must be conducted in order to understand the robustness of our space under such a non-binary input. In doing so, we aim to study and expand our model with relevant dimensions in musical practice, notably timbre/spectral and amplitude information. Ultimately, we want to describe musical audio as robustly as symbolic music representations and apply our model within the realm of performed music.

Funding

This research was supported by the Project “NORTE-01-0145-FEDER-000020”, financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF) and by the Portuguese Foundation for Science and Technology under the post-doctoral grants SFRH/BPD/109457/2015 and SFRH/BPD/88722/2012.

ORCID

Gilberto Bernardes http://orcid.org/0000-0003-3884-2687
Marcelo Caetano http://orcid.org/0000-0002-5119-8964
Matthew E.P. Davies http://orcid.org/0000-0002-1315-3992

References

Amiot, E. (2013). The torii of phases. In J. Yust, J. Wild, & J. A. Burgoyne (Eds.), *Mathematics and computation in music* Lecture Notes in Computer Science, Vol. 7937 pp. 1–18). Berlin Heidelberg: Springer.

Amiot, E., & Sethares, B. (2011). An algebra for periodic rhythms and scales. *Journal of Mathematics and Music, 5*(3), 149–169.

Balzano, G.J. (1980). The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal, 4*(4), 66–84.

Balzano, G.J. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, mind, and brain* (pp. 321–351). New York: Springer.

Barlow, C. (2012). *On musiquantics*. Mainz: Johannes Gutenberg-Universität Mainz.

Behringer, R., & Elliot, J. (2009). Linking physical space with the Riemann Tonnetz for exploration of Western tonality. In J. Hermida & M. Ferrero (Eds.), *Music education* (pp. 131–143). Hauppauge, NY: Nova Science Publishers.

Bernardes, G., Cocharro, D., Guedes, C., & Davies, M.E.P. (2015). *Conchord: An Application for Generating Musical Harmony by Navigating in a Perceptually Motivated Tonal Interval Space*. Proceedings of the 11th International Symposium on Computer Music Modeling and Retrieval (CMMR), (pp. 71–86). Plymouth, UK,

Bernardes, G., Cocharro, D., Guedes, C., & Davies, M.E.P. (2016). Harmony generation driven by a perceptually motivated tonal interval space. ACM Computers in Entertainment. In Press.

Bharucha, J.J. (1984). Event hierarchies, tonal hierarchies, and assimilation: A reply to Deutsch and Dowling. *Journal of Experimental Psychology: General, 113*, 421–425.

Callender, C. (2007). Continuous harmonic spaces. *Journal of Music Theory, 51*(2), 277–332.

Chew, E. (2000). *Towards a mathematical model of tonality* (Ph.D. dissertation), MIT, Cambridge, MA, USA.

Chew, E. (2008). Out of the grid and into the spiral: Geometric interpretations of and comparisons with the spiral-array model. *Computing in Musicology, 15*, 51–72.

Chew, E., & Chen, Y. (2003). Determining context-defining windows: Pitch spelling using the spiral array. Poster presented at the International Society for Music Information Retrieval Conference; Baltimore, MD.

Clough, J., & Douthett, J. (1991). Maximally even sets. *Journal of Music Theory, 35*, 93–173.

Cohn, R. (1997). Neo-Riemannian operations, parsimonious trichords, and their tonnetz representations. *Journal of Music Theory, 41*, 1–66.

Cohn, R. (1998). Introduction to neo-Riemannian theory: A survey and a historical perspective. *The Journal of Music Theory, 42*(2), 167–180.

Cook, N.D. (2012). *Harmony, perspective, and triadic cognition*. New York: Cambridge University Press.

Cook, N. D., Fujisawa, T., & Konaka, H. (2007). Why not study polytonal psychophysics? *Empirical Musicology Review, 2*(1), 38–44.

De Leeuw, J., & Mair, P. (2009). Multidimensional scaling using majorization: SMACOF in R. *Journal of Statistical Software, 31*(3), 1–30.

Deutsch, D. (1984). Two issues concerning tonal hierarchies: Comment on Castellano, Bharuch, and Krumhansl. *Journal of Experimental Psychology: General, 113*(3), 413–416.

Euler, L. (1968). Tentamen novae theoriae musicae. *St. Petersburg*. New York: Broude. (Original published 1739).

Gatzsche, G., Mehnert, M., & Stöcklmeier, C. (2008). Interaction with tonal pitch spaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 325–330). Genova: NIME.

Harte, C., Sandler, M., & Gasser, M. (2006). Detecting harmonic change in musical audio. In *Proceedings of the 1st ACM Workshop on Audio and Music Computing Multimedia* (pp. 21–26). New York: ACM.

Huron, D. (1994). Interval-class content in equally tempered pitch-class sets: Common scales exhibit optimum tonal consonance. *Music Perception: An Interdisciplinary Journal, 11*(3), 289–305.

Hutchinson, W., & Knopoff, L. (1979). The acoustic component of Western consonance. *Interface, 10*(2), 129–149.

Hyer, B. (1995). Re-imagining Riemann. *Journal of Music Theory, 39*(1), 101–138.

Kameoka, A. & Kuriyagawa, M. (1969). Consonance theory. Part I: Consonance of dyads. *Journal of Acoustical Society of America, 45*, 1451–1459.

Kopp, D. (1995). *A comprehensive theory of chromatic mediant relations in mid-nineteenth-century music* (Ph.D. dissertation). Brandeis University, Waltham, MA.

Kross, D. (2006). Modernism and tradition and the traditions of modernism. *Muzikologija, 2006*(6), 19–42.

Krumhansl, C. (1990). *Cognitive foundations of musical pitch*. Oxford: Oxford University Press.

Krumhansl, C.L. (1998). Evidence supporting the psychological reality of neo-Riemannian transformations. *Journal of Music Theory, 42*, 265–281.

Krumhansl, C.L., & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334–368.

Kruskal, J.B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika, 29*, 28–42.

Lee, K. (2007). A system for automatic chord transcription from audio using genre-specific HMM. In N. Boujemaa, M. Detyniecki, & A. Nürnberger (Eds.), *Adaptive multimedial retrieval: Retrieval, user, and semantics* (pp. 134–146). Berlin-Heidelberg: Springer-Verlag.

Lee, K., & Slaney, M. (2007). A unified system for chord transcription and key extraction using hidden Markov models. In *Proceedings of International Society for Music Information Retrieval* (pp. 245-250).Vienna: Austrian Computer Society (OCG).

Lerdahl, F. (1988). Tonal pitch space. *Music Perception, 5*(3), 315–350.

Lerdahl, F. (2001). *Tonal pitch space*. New York: Oxford University Press.

Lewin, D. (1982). A formal theory of generalized tonal functions. *Journal of Music Theory, 26*, 23–60.

Lewin, D. (1987). *Generalized musical intervals and transformations*. New Haven: Yale University Press.

Lewin, D. (1992). Some notes on analyzing Wagner: The Ring and Parsifal. *Nineteenth Century Music, 16*, 49–58.

Lewin, D. (2001). Special cases of the interval function between pitch-class sets X and Y. *Journal of Music Theory, 45*, 1–29.

Longuet-Higgins, H.C. (1962). Two letters to a musical friend. *Music Review*, *23*, 244–248 and 271–280.

Malmberg, C.F. (1918). The perception of consonance and dissonance. *Psychological Monographs, 25*(2), 93–133.

Mashinter, K. (2006). Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. *Empirical Musicology Review, 1*(2), 65–84.

Mauch, M. (2010). *Automatic chord transcription from audio using computational models of musical context* (Ph.D. dissertation). University of London, London, UK.

Mooney, M.K. (1996). *The table of relations and music psychology in Hugo Riemann's harmonic theory* (Ph.D. dissertation). Columbia University, New York, NY.

Navarro, M., Caetano, M., Bernardes, G., de Castro, L.N., & Corchado, J.M. (2015). Automatic generation of chord progressions with an artificial immune system. *Evolutionary and Biologically Inspired Music, Sound, Art and Design*. Springer International Publishing. (pp. 175–186)

Oppenheim, A., Schafer, R., & Buck, J. (1989). *Discrete-time signal processing* (2nd ed.). Englewood Cliffs, NJ: Prentice-hall.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach*. Berlin: Springer.

Peiszer, E., Lidy, T., & Rauber, A. (2008). Automatic audio segmentation: Segment boundary and structure detection in popular music. In *Proceedings of the International Workshop on Learning the Semantics of Audio Signals* (pp. 45–59). Paris: LSAS.

Plomp, R., & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38*, 548–560.

Quinn, I. (2006). General equal-tempered harmony (introduction and part I). *Perspectives of New Music, 44*(2), 114–158.

Quinn, I. (2007). General Equal-tempered Harmony: Parts 2 and 3. *Perspectives of New Music, 45*(1), 4–63.

Roberts, L. A. (1986). Consonance judgements of musical chords by musicians and untrained listeners. *Acta Acustica United with Acustica, 62*(2), 163–171.

Schoenberg, A. (1969). *Structural functions of harmony* ([1954], 2nd ed., revised by Leonard Stein). New York: W.W. Norton, Inc.

Sethares, W. (1999). *Tuning, timbre, spectrum, scale*. London: Springer-Verlag.

Shepard, R.N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I & II. *Psychometrika*, *27*, 125–140 and 219–246.

Shepard, R.N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 335–353). New York: Academic Press.

Temperley, D. (2001). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.

Tymoczko, D. (2011). *A geometry of music: Harmony and counterpoint in the extended common practice*. New York: Oxford University Press.

Ueda, Y., Uchiyama, Y., Nishimoto, T., Ono, N., & Sagayama, S. (2010). HMM-based approach for automatic chord detection using refined acoustic features. In *Proceedings of the 2010 IEEE International Conference on Acoustics Speech and Signal Processing* (pp. 5518–5521). Piscataway, NJ: IEEE Press.

Weber, G. (1817–1821). *Versuch einer geordeneten theorie der tonsetzkunst* [Theory of musical composition]. Mainz: B. Schotts Söhne.

Yust, J. (1983). Schubert's Harmonic Language and Fourier Phase Space. *Journal of Music Theory* 59(1): 121–181.