Modelling tonal attraction: tonal hierarchies, interval cycles, and quantum probabilities

Reinhard Blutner

500

Soft Computing

A Fusion of Foundations, Methodologies and Applications

Springer

Springer

CrossMark

FOCUS

Modelling tonal attraction: tonal hierarchies, interval cycles, and quantum probabilities

Reinhard Blutner[1]

Abstract How well does a given pitch fit into a tonal scale or tonal key, let it be a major or minor key? A similar question can be asked regarding chords and tonal regions. Structural and probabilistic approaches in computational music theory have tried to give systematic answers to the problem of tonal attraction. We will discuss two previous models of tonal attraction, one based on tonal hierarchies and the other based on interval cycles. To overcome the shortcomings of these models, both methodologically and empirically, I propose a new kind of models relying on insights of the new research field of quantum cognition. I will argue that the quantum approach integrates the insights from both group theory and quantum probability theory. In this way, it achieves a deeper understanding of the cognitive nature of tonal music, especially concerning the nature of musical expectations (Leonhard Meyer) and a better understanding of the affective meaning of music.

Keywords Quantum cognition · Tonal attraction · Computational music theory · Generative music theory · Interval cycles · Musical expectation · Affective meaning of music

1 Introduction

In the past 20 years, there has been an enormous progress in the development of cognitive theories of tonal music.

✉ Reinhard Blutner
blutner@contact.uva.nl

[1] Universiteit van Amsterdam, Amsterdam, The Netherlands

We can distinguish two different kinds of models that have been developed within the years—structural and probabilistic accounts. The label "structural" applies to approaches that follow the mechanics of symbol manipulation systems (Harnad 1990; Newell 1980) prominently represented by the generative music theory (Lerdahl and Jackendoff 1983). The label "probabilistic" applies to all approaches that see (subjective) probabilities expressing expectations of the music listener as crucial for the understanding of tonal music (Huron 2006; Meyer 1956; Temperley 2007).

One central claim of the present paper is that structural and probabilistic approaches can and must be integrated to construct a proper approach to tonal music. This idea of integration is not new in cognitive science. It was developed successfully in cognitive linguistics, most prominently by the writings of Paul Smolensky and his colleagues (Prince and Smolensky 1993/2004; Smolensky and Legendre 2006). In recent years, proponents of quantum cognition have developed another kind of integration. This kind of approach uses insights from the mathematics developed in quantum physics, especially the role of symmetries and the formulation of a special kind of probability theory that differs from the standard Kolmogorovian calculus. It integrates the structural insights of the relevant symmetry groups with the insights of quantum probabilistics.

As the central problem of this paper, we will investigate the question of tonal attraction. How well does a given pitch fit into a tonal scale or tonal key, let it be a major or minor key? In an celebrated study, Krumhansl and Kessler (1982) asked listeners to rate how well each note of the chromatic octave fitted with a preceding context, which consisted of short musical sequences in major or minor keys. The results of this experiment clearly show a kind of hierarchy: the tonic pitch received the highest rating, followed by the pitches completing the tonic triad (third and fifth), followed by the

remaining scale degrees, and finally the chromatic, non-scale tones. This finding plays an essential role in Lerdahl's and Jackendoff's generative theory of tonal music (Lerdahl and Jackendoff 1983) and is one of the main pillars of the structural approach in music theory.

Several models aim to account for the mechanism of tonal attraction. We will discuss two of them: the hierarchical approach and the interval cycles account. The hierarchic approach to tonal pitches was developed by Krumhansl and Kessler (1982), Lerdahl and Jackendoff (1983), and Lerdahl (1988, 2001). In this approach, tonal attraction is proportional to the number of levels in the tonal hierarchy (tonic root, tonic triad, diatonic series) the pitch belongs to. The account based on interval cycles is due to Woolhouse and colleagues (Woolhouse 2009, 2010; Woolhouse and Cross 2010) and was criticized by Quinn (2010).[1] The model assumes the hypothesis of interval cycle proximity (ICP):

> "Tonal attraction in music is proportional to the sum of the interval cycles formed between sequential pairs of tones and/or chords. Higher interval cycles produce strong tonal attraction; low interval cycles produce weak tonal attraction." (Woolhouse 2010: 66)

I will demonstrate that both models are insufficient for empirical and conceptual reasons. Moreover, I will develop a very simple but new model based on quantum probabilities, combined with the idea of symmetries grounded on the cyclic group C_{12}. The model highlights certain geometric aspects of a theory of tonal music (Tymoczko 2011) and it integrates its structural and probabilistic aspects.

In the following section, I will discuss some previous models of tonal attraction in more detail including Lerdahl's hierarchic model and Woolhouse's model of interval cycles. Section 3 develops the quantum model of tonal attraction. In Sect. 4, I compare the three models, discuss their empirical and methodological impact, and draw some general conclusions.

2 Previous models of tonal attraction: tonal hierarchies and interval cycles

Geometrical (or topological) models have a long tradition in music theory. They can relate to pitches and pitch classes, to chords, and to tonal regions (defining the known key systems in European music). A tonal pitch system consists of a number of pitches where pitches are sounds defined by a certain fundamental frequency. In this paper, we assume 12

pitch classes, also called tones[2], and we will use a numeric notation to define the 12 tones of the system, in ascending order:

$$0 = C, 1 = C\sharp, 2 = D, 3 = D\sharp, 4 = E, 5 = F, 6 = F\sharp,$$
$$7 = G, 8 = G\sharp, 9 = A, 10 = B\flat, 11 = B. \tag{1}$$

Pitches, and on a more abstract level, tones are objects of our acoustic perception. Basically, perception is based on similarities. Consequently, pitches can be ordered by this similarity relation. Geometric models simulate the perceived similarity of pitches by geometric distances in a spatial model.

A standard example is the *tonnetz* first proposed by Euler (1739), later adapted by Riemann, Longuet-Higgins and many others. In the Euler *tonnetz*, the tones are organized along two axes. The horizontal axis consists of a fifth cycle and the vertical axis of a major third cycle (and the diagonals yield semitone and minor third cycles). Balzano (1980) proposed another kind of *tonnetz* grounded on group theory. The horizontal axis consists of a minor third cycle and the vertical axis of a major third cycle (the diagonals yield semitone and fifth cycles).

One significant problem with both kinds of spatial representations is that they generate absolute, context-independent distances (or similarities) between pitches. However, the similarity of pitches is not absolute—it is dependent of a given scale that underlies the tonal system. In Western systems of tonal music, for instance, a common scale is a diatonic scale based on a certain root tone. If C is the root tone, the diatonic scale (C major) consists of the seven pitches C, D, E, F, G, A, B. Based on the C-major scale, the perceived distances between E–F on the one hand and C–D on the other hand are equal. However, this is not reflected by equal distances in the *tonnetz*.[3] Once more based on the C-major scale, the perceived distances between E–F♯ on the one hand and C–D on the other hand are perceived differently even when the corresponding distances in the *tonnetz* are the same.[4] When the underlying scale is changed to G-major we get the opposite pattern even when the distances within the *tonnetz* (and the corresponding frequency quotients) are not changed: the perceived distances between E–F and C–D are different but the perceived distances between E–F♯ and C–D are identi-

---

[2] Tones can be seen as equivalent classes of pitches. Two pitches with fundamental frequencies $f_1 \geq f_2$ are equivalent if f_1/f_2 is a natural number (i.e., the two pitches are equal or have a distance of one or more octaves). Hence, the concept of tones as equivalence classes of pitches abstracts from the octave level.

[3] Similar observation can be based on acoustic considerations by comparing the fundament frequency quotients (assuming the octave is equally divided into 12 parts, i.e. the tuning is equal temperament). For E–F the quotient is about 16/15. In contrast, for C–D the corresponding quotient is 9/8.

[4] Likewise, the frequency quotients are identically in both cases (9/8).

cal. The situation is like the situation in natural language. The similarity relations between different phonemes depend on the underlying language (assuming the considered phonemes do really appear in all the considered languages).

Lerdahl (1988) states another problem with the *tonnetz*. The problem is that this structural approach does not provide a consistent theory of similarity considering the level of tones, chords, and tonal regions.

> A comprehensive model must incorporate all three levels into one framework, representing perceived proximity at each level and showing how the levels interconnect. (Lerdahl 1988: 319)

In the following subsection, we present Lerdahl's model of the tonal pitch space and look how it generalizes to chords and tonal regions.

2.1 Tonal hierarchies

Lerdahl (1988, 2001) has developed a model of tonal attraction based on a tonal hierarchy. Forerunners of this approach are Krumhansl (1979), Krumhansl and Kessler (1982) and Deutsch and Feroe (1981). Lerdahl has presented the model in a way that allows stringent generalisations and that invites for comparisons with the linguistic domain—using the common nominator "optimality theory" (Prince and Smolensky 1993/2004; Smolensky and Legendre 2006).

A numerical representation of Lerdahl's basic space for C-major is given in Table 1. It shows the 12 tones at their levels in the tonal hierarchy. In all, five levels are considered:

A: octave space (defined by the root tone, $0 = C$ in the present case),
B: open fifth space,
C: triadic space,
D: diatonic space (including all diatonic pitches of C-major in the present case),
E: all (including all 12 pitch classes).

Table 1 also shows the *embedding distance c*, which is calculated by counting the number of levels down that a pitch class first appears. The smaller the embedding distance, the

Table 1 The basic tonal pitch space as given in Lerdahl (1988)

Level A	0										
Level B	0						7				
Level C	0			4			7				
Level D	0	2		4	5		7		9		11
Level E	0	1	2	3	4	5	6	7	8	9	10	11
Embedding distance c	0	4	3	4	2	3	4	1	4	3	4	3

higher its tonal attraction (i.e. the better it fits into the given tonal scale).

The basic tonal pitch space is easy to model within the framework of optimality theory (Prince and Smolensky 1993/2004; Smolensky and Legendre 2006). In this framework, the tonal levels have to be interpreted by tonal constraints. The constraints simply express whether a given tone is a member of the considered tonal level. For example, the constraint A (related to the tonal level A) is satisfied if the considered tone is the root tone and it is violated otherwise.

From Table 2 it is easy to see that the embedding distance is exactly the sum of the constraint violations. Hence, all constraints have to be considered as equally ranked to yield identical numerical values for identical numbers of constraint violations. Table 2 also exhibits a measure of tonal abstraction, which is a linear function of embedding distance c. I have chosen the form $6.5 - c$ since it best fits the data of Krumhansl and Kessler (1982) for the C major scale. Figure 1 presents the best fit for the major scale and Fig. 2 for the (harmonic) minor scale.[5]

The most obvious empirical finding of the study is that both for major and minor keys the seven tones of the scale have higher values of tonal attraction than the five tones which are not part of the scale. This is clearly seen in the left part of Fig. 1, where we almost have a complete agreement for data and model around the average rating of 2.3. In the right part of the figure, the tones are ordered by the ascending circle of fifth. In this case, the five tones before the last are the non-scale members and approximate a horizontal line. The ordering qua circle of fifths makes a special shape of the data visible, which will be important when we discuss the quantum model in Sect. 3. A second general finding is that all tones of the tonic triad have higher values than the

Table 2 The basic tonal pitch space as given by an optimality theoretic tableau

Constraints\pitch class	0	1	2	3	4	5	6	7	8	9	10	11
A		x	x	x	x	x	x	x	x	x	x	x
B		x	x	x	x	x	x		x	x	x	x
C		x	x	x		x	x		x	x	x	x
D		x		x			x		x		x	
Embedding distance c	0	4	3	4	2	3	4	1	4	3	4	3
Tonal attraction $(6.5 - c)$	6.5	2.5	3.5	2.5	4.5	3.5	2.5	5.5	2.5	3.5	2.5	3.5

The embedding distance now is simply the number of violations of the involved constraints A–D

---

[5] There are three minor scales. If C is the root tone, these are the three scales: (1) natural: C D E♭ F G A♭ B♭ C; (2) harmonic: C D E ♭ F G A♭ B C; (3) melodic: C D E♭ F G A B C (ascending) and C B♭ A♭ G F E♭ D C (descending). In the following, we consider the harmonic scales only (in agreement with Krumhansl and Kessler (1982).
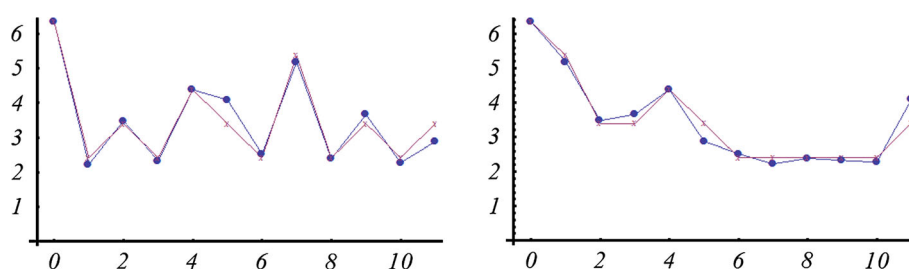
Fig. 1 Distribution marked by *filled circle*: data of Krumhansl and Kessler (1982) for the key C major; distribution marked by *x*: score of constraint violations d fitted to the data of Krumhansl and Kessler (1982) by using the linear approximation $6.5 - c$. The fit gives the value 6.5 for pitch class 0 (minimal violations) and the value 2.5 for pitch class 1 (maximal violations). On the *left hand side*, the tones are ordered chromatically; on the *right hand side*, they are ordered by the ascending circle of fifth

Fig. 2 Related distributions for the key C minor. The harmonic minor scale is chosen for defining the level D violations. On the *left hand side*, the tones are ordered chromatically; on the *right hand side*, they are ordered by the ascending circle of fifth
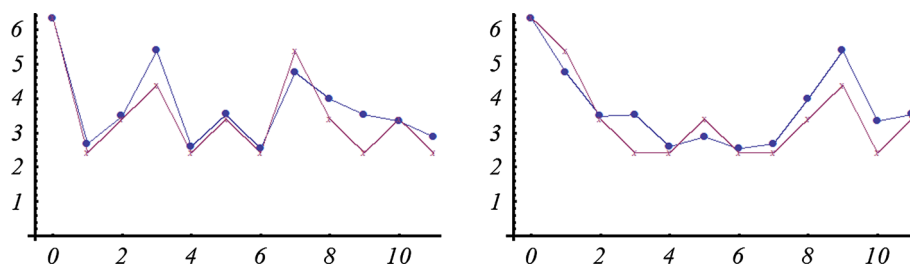
other tones of the scale. On the right-hand side of Fig. 1, we find a local maximum at the element 4 of the circle of fifth (the major third of the triad), which a little bit disturbs the charming shape of the curve.

The data for the minor key are similar but the fit with the model is far from being complete. The problem arises because we have three minor scales. The one which leads to the best agreement is the harmonic minor scale. On the left-hand side, you see that for the penultimate two tones there is the highest disagreement between model and data. These are the tones A and ♭B, which are no elements of the harmonic C-minor scale. Now consider the right-hand side of Fig. 2. A local maximum appears at point 9 representing the note E♭ in the harmonic minor triad.

Next, we consider charts of region and key relations. The most famous of these graphic representations is Weber's regional chart (Weber 1824), which was later adopted by Schönberg (1969). The chart represents the regions by key, and it is shown in Fig. 3.

The general idea of a spatial chart is that distances and similarities between tones, chords, and regions are expressed by spatial distances and similarities. For instance, Fig. 3 shows that the C-major region is closest to the G-major and F-major regions, as well as to the A-minor and C-minor regions. This is an intuitively correct outcome. However, when comparing C-major with F-minor, the perceptual distance is predicted to be larger than that between C-major and F-major. As Tymoczko (2011) notes in appendix C: F-minor frequently appears as a passing note between F-major and C-major and


Fig. 3 Weber's regional chart

should be closer to both, i.e. the distances f–F and f–C should be smaller than the distance C–F. Unfortunately, this is not expressed by the spatial chart.[6]

A more general problem is that the distance and similarity relation are assumed to be symmetric relations. However, Temperley (2007: 104) stresses the point that key relations are asymmetric. For instance, consider the most closely related key to C-major, which is G-major (dominant). However, when taking G-major as a starting point, the most closely related key is not C-major (subdominant) but D-major. This

---

[6] A similar problem arises with Euler's *tonnetz*. Here, triadic chords can be represented by triangles on the *tonnetz*. Again, the spatial distance between the two triangles for C-major and F-minor is larger than the distance between the two triangles for C-major and F-major (cf. Tymoczko 2011, appendix C).

Fig. 4 The chord circle of fifths (*bottom*) and the region circle of fifths (*top*) are used to calculate the distance between two chords in Lerdahl's ([1988](#)) model of chord proximity
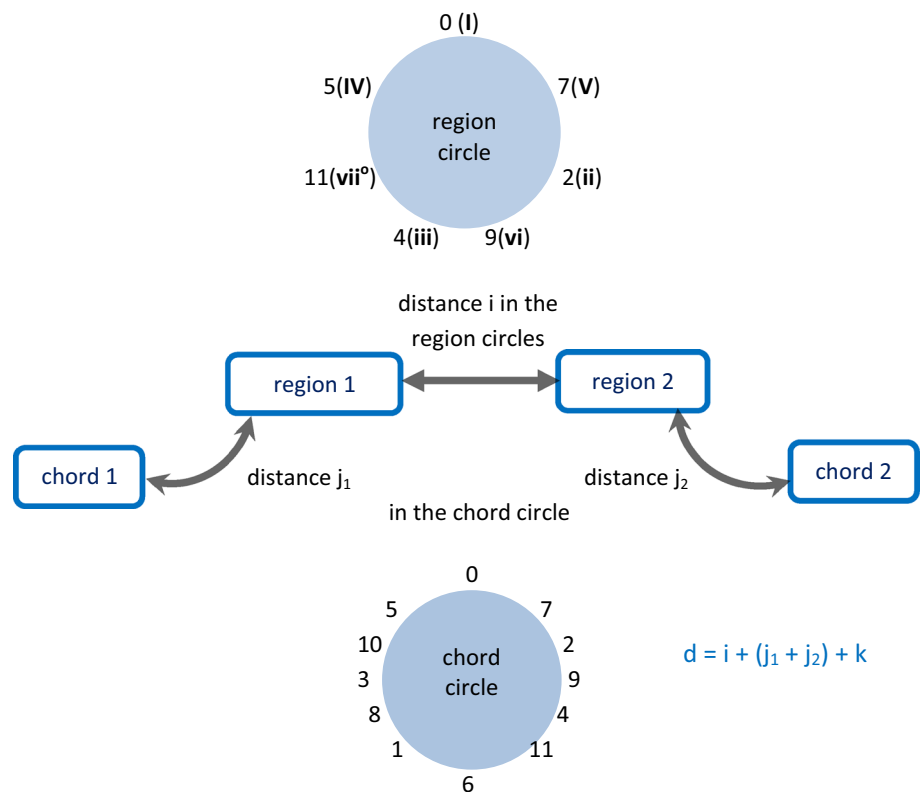
exemplifies the asymmetry of the distance relation between keys and Temperley (2007) gives empirical evidences for it. The basic findings of asymmetry are in agreement with general work on prototype effects in conceptual judgments. For example, Tversky (1977) found that the similarity between China and Korea is less than the similarity between Korea and China. Similarly, people have a tendency to say 999 is about 1000 but not 1000 is about 999. For a new treatment of asymmetric similarity judgements within the framework of quantum cognition the reader is referred to Pothos et al. (2013).

Lerdahl's model is able to account for Weber's regional chart (Lerdahl 1988: 331ff) but it also shares the shortcomings just outlined. The basic idea is that this chart can be constructed by combining the fifths cycle and the relative and parallel major-minor cycle. With each application of one-step in the circle of fifth, the distance rises by one-step (transforming C–G or C–F, for instance). The same holds when applying the major–minor cycle (transforming c to E♭ or C to a, for instance). In an alternative model, Krumhansl and Kessler (1982) suggest to quantify the similarities between two keys (regions) by calculating the correlations between their key profiles. They construct a four-dimensional spatial representation of keys corresponding to the correlation values and show how closely it resembles Weber's regional chart.

The application of Lerdahl's model to distances between chords is more complicated. An important insight is that chords are always relative to a tonal region (scale). Hence, the chord of C-major is not absolute; it can be I/**I** (if it is seen in the context of C-major) or IV/**V** (in the context of G-major) or even III/**VI** (in the context of A-minor). The distance measure is based on the chord circle of fifths and the region circle of fifths. The chord circle relates the seven chords within a tonal scale and the region circle relates the 12 regions. Note that the region circle does not distinguish between major and minor regions (Fig. 4).

Hence, the distance between C-major and A-minor is zero in the region circle but three in the chord circle. Lerdahl (1988, 2001) makes use of the following linear formula to calculate the distance d between two chords (given in the context of a certain region):

$$d = i + (j_1 + j_2) + k \tag{2}$$

Hereby, i is the distance (number of steps) in the region circle; j_1 is the distance from the first chord to the tonic chord of the first region (calculated in the corresponding chord circles); j_2 is the distance from the second chord to the tonic chord of the second region; k is the number of distinct tones of the two chords. For example, when comparing iv/**vi** (D-minor chord in the context of A-minor region) with I/**I** (C-major chord in the context of C-major region), the calculated distance is $d = 0 + (3 + 1) + 6 = 10$. This is different from the distance between ii/**I** and I/**I,** where D-minor and C-major

Fig. 5 Stacked chart based on Lerdahl's (1988) model. The calculated conditioned probabilities are inversely related to the distances predicted by the model
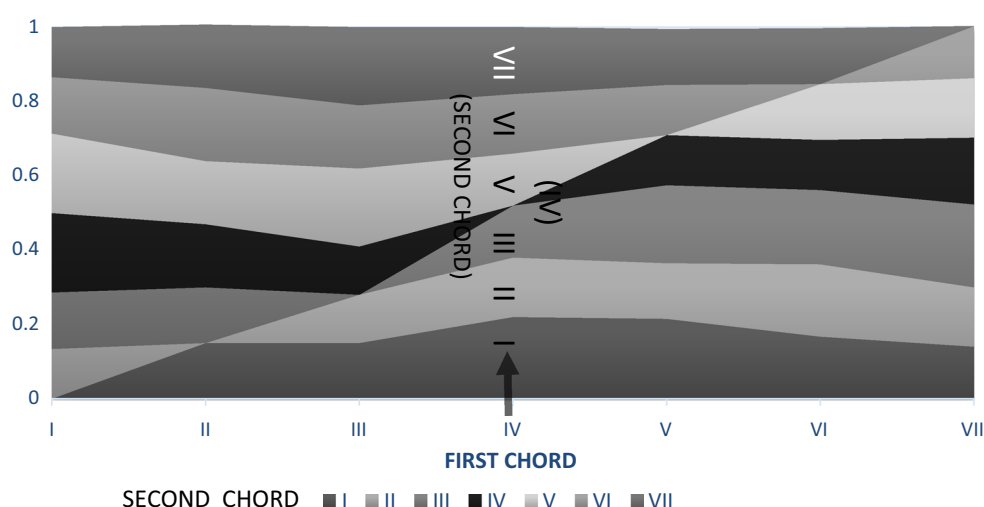
are both compared in the context of C-major: $d = 0 + 2 + 4 = 6$.

Recently, Huron (2006) has presented data based on corpus studies (Huron 2006: 251). These data consist of the frequencies of various chord progressions in a sample of baroque music. From these data the probabilities of a chord given some antecedent chord are derived [i.e., we consider the conditioned probabilities P(target chord/antecedent chord)]. The following stacked chart shows these conditioned probabilities.

Note that the conditioned probabilities for each chord sum up to 1 in the diagram. The broader the considered "second chord strip" for a given "first chord", the higher the probability of the considered target chord.

If we take Lerdahl's distance model and relate distances inversely to probabilities and normalize appropriately, then we get the stacked chart shown in Fig. 5. Even though the visual impression is counterintuitive against a strict correlation between the model's prediction and the corpus data, there is some positive correlation ($r = 0.21$ in the average).

In agreement with Quinn (2010), I consider the (averaged) correlation test as a very week instrument to establish causal connections between models and data. It is fair to say that Lerdahl's distance model is not able to explain Huron's (2006) corpus data of a sample of baroque music. However, this is not really a disproof of Lerdahl's model.

Lerdahl makes a careful distinction between tonal hierarchies and event hierarchies. The latter are "part of the structure that listeners infer from temporal musical sequences" (Lerdahl 1988: 316). Data that concern "chord progression" should be explained in terms of such event hierarchies. According to Lerdahl, consequently, we should not expect that the distances calculated by his model, which obviously is based on tonal hierarchies, conform to the chord progression

Table 3 Interval-cycle proximity as a function of interval length

Interval length	0	1	2	3	4	5	6	7	8	9	10	11
Interval-cycle proximity (ICP)	1	12	6	4	3	12	2	12	3	4	6	12

data. Unfortunately, there is no alternative model available that is created on event hierarchies.

2.2 Interval cycles

In recent research, Matthew Woolhouse has proposed to explain tonal attraction in terms of interval cycles (Woolhouse 2009, 2010; Woolhouse and Cross 2010). The basic idea is that the attraction between two pitches is proportional to the number of times the interval spanned by the two pitches must be multiplied by itself to produce some whole number of octaves. Assuming 12-tone equal temperament, the ICP of the interval can be defined as the smallest positive number *ICP* such that the product with the interval length (i.e. the number of half tone steps spanned by the interval) is a multiple of 12 (maximal interval length). The following table lists the *ICP*s for all intervals spanned by a given interval length. For example, you see that the *ICP* for the triton is 2 and the *ICP* for the fifth is 12. This has the plausible consequence that, relative to a root tone, the fifth has higher tonal attraction than the triton (Table 3).

A more general consequence is the kind of symmetry that arises: an interval of n semitones will have the same *ICP* as an interval of $12 - n$ semitones. Unfortunately, this consequence is wrong empirically. In fact, Krumhansl (1979) found that subjects rated the same pairs of notes differently when the notes were presented in different orders. For understanding this result it is essential that Krumhansl presented the note

pairs in a tonal context (say C-major or C-minor). Such a tonal context requires more than a root tone in order to be defined. A defining context can consist of a whole scale, a chord, or a cadential sequence of chords. In any case, it requires more than just *one* root tone.

Woolhouse proposes to overcome the problem of symmetry by taking a linear combination of the *ICP*s of the note pairs considering all elements of the tonal context. In the simplest case, this is the straight sum.[7] Instead of the straight sum, I suggest to take the arithmetic mean. This makes it easier to compare the effect of different tonal contexts (chords, cadences, scales), which can have a quite different number of tones. In cases with the same number of notes, the results agree (up to a scaling factor) with Woolhose's values. I will call the arithmetic mean of *ICP*s relative to a given context *the context-driven ICP*.

In the following, I will take the context-driven *ICP* as a measure for tonal attraction in a given tonal context. Context-driven *ICP* can also be calculated for chords. To get an *ICP* for a chord, we simply add the values of the tones of the chord. Now let us consider some results presented by Woolhouse (2010). First, taken the C-major scale as context, the tones with the highest context-driven *ICP* (i.e. the highest tonal attraction) are C and E. Of the seven possible diatonic triads, C-major and A-minor have the highest context driven *ICP*. Second, consider the *natural* A-minor scale (same tones as for C-major but in a different order). Again, C and E are the tones with the highest context-driven *ICP*, and—as before— of the seven possible diatonic triads, C-major and A-minor have the highest context driven *ICP*. Third, considering the *harmonic* A-minor scale as context, the tone with the highest context-driven *ICP* is A, and the optimal triad is the A-minor triad. All these results are plausible findings. Importantly, they were found without any additional stipulation. Quinn (2010) notes as a problem for Woolhouse's model that it gives counterintuitive results when the melodic minor scale is considered. In this case, contrasting with the harmonic minor scale, the tone with the highest context-driven *ICP* is B, and the optimal triad is the E-major triad.

Next, we will consider an example presented by Woolhouse and Cross (2010). In this example, the G7 chord is taken as context and the ICP-profile is considered for the harmonic minor scale C D Eb F G Ab B (= 0, 2, 3, 5, 7, 8, 11). The chord G[7], the dominant seventh in the key of C minor, has been chosen because it strongly attracts C minor. The following figure shows the ICP-profile for the seven tones of the the harmonic C-minor scale.

Figure 7 shows the full context-driven *ICP* profile, i.e. it considers the tonal attraction of all 12 pitch classes relative to the given context.
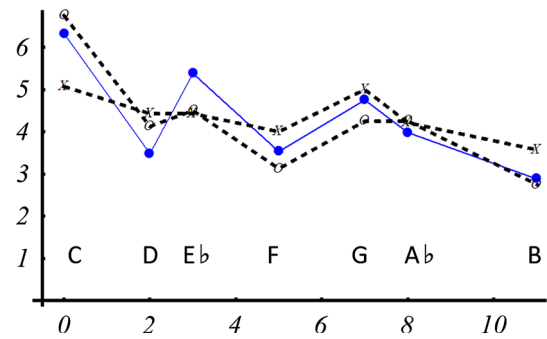

Fig. 6 Context-driven ICP-profile for the harmonic minor scale. The contexts are (1) the harmonic C-minor scale (*dashed profile* marked by *x*) and (2) the chord G[7] (*dashed profile* marked by *o*). The distribution marked by *filled circle* represents the data of Krumhansl and Kessler (1982) for the harmonic minor scale (C-minor). The theoretical curves are scaled for best agreement with the empirical data
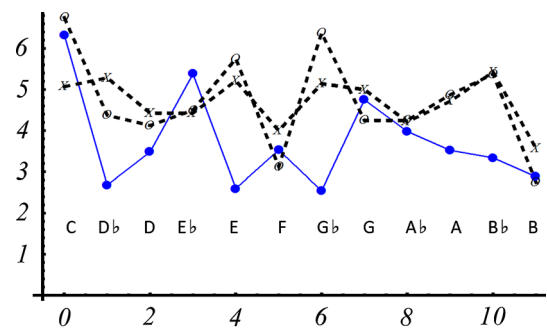

Fig. 7 Full context-driven ICP-profile, where the contexts are (1) the harmonic C-minor scale (*dashed profile* marked by *x*) and (2) the chord G[7] (dashed profile marked by *o*). The *solid curve* represents the data of Krumhansl and Kessler (1982) for all 12 pitch classes. The scaling factors are identical with those of Fig. 6

Figure 7 illustrates two important findings. First, it shows that the agreement between the Krumhansl–Kessler data and the theoretically calculated attraction values is very unsatisfying for all pitches that are not members of the (harmonic) C-minor scale. Second, it demonstrates that Woolhouse's approach in terms of interval cycles is very sensitive to the contextual triggers, even if they conform to a given tonal scale.

Both aspects are essential for of a sound methodological perspective. If the attraction value a listener feels in a given tonal context for a particular pitch is causally connected with the numbers of interval cycles, then we could expect that this connection is effective for all 12 pitches of the scales, not only for the subset that optimally fits the given tonal context. Second, we could expect that there are significant differences between contextual triggers conforming to a tonal scale and contextual triggers non-conforming to it. Ample evidence suggests that these differences are not predicted theoretically. In the next section, I will show how this aspect can be assessed.

---

[7] In tonal contexts consisting of chords, a weighted sum can be considered that gives extra weight to intervals involving a chord root.

Fig. 8 Stacked chart reflecting Piston's table of chord progression
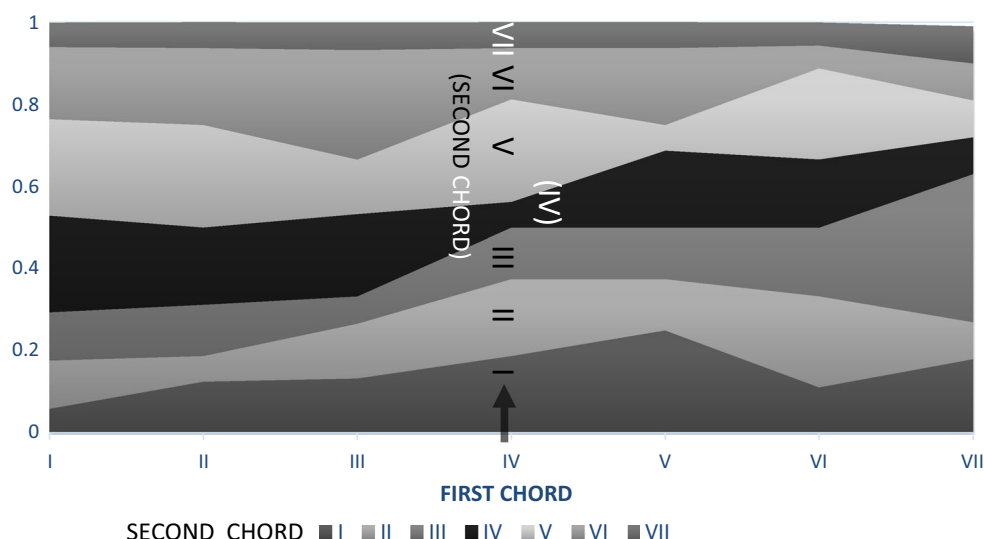
Quinn (2010) calculated the correlations between the Krumhansl–Kessler data and the predictions of the context-driven ICP model for fully chromatic key profiles. In contrast to the suggestions made by Woolhouse and Cross (2010), the average correlations between the full context-driven *ICP* profile and Krumhansl–Kessler profiles are much closer to zero than the correlations for scale-restricted key profiles. In the major-scale case, the mean correlation is 0.089 and in the minor case, the mean correlation is –0.045.[8]

Quinn (2010) further notes an important methodological issue. It concerns the question of a deeper conceptual motivation for the (causal) connection between interval cycles and tonal attraction.[9]

> Why should we expect that the attraction a listener feels between two pitches should have anything to do with the number of times the interval separating them needs to be multiplied to produce an octave-equivalent interval? On the face of it, these two properties of an interval do not seem to have anything to do with one another. Woolhouse does not provide much theoretical discussion of the matter, largely confining himself to attempts at showing correlational and anecdotal links between ICP and various aspects of tonality. (2010: 173)

Woolhouse (2010) applied his model of tonal attraction to calculate values for all chord pairs of a given region (key) and compared it with Piston's semi-empirical table of expectation in chord progression (Piston 1979). Piston's table consists of statements like "IV is followed by V, sometimes I or II, less often III or VI." Woolhouse quantified such statements by identifying four levels of chord-progression frequency: "is followed by" was rated 4, "sometimes" was rated 3, "less often" was rated 2, and a progression not mentioned was rated 1. It was found that the predictions of the ICP model were significantly correlated with the Piston-derived frequency ratings ($r_s = 0.66$, $p < 0.005$). An attempt to replicate Woolhouse's findings with data about chord progressions in musical corpora from Bach and Mozart showed a significantly degraded performance (Quinn 2010). "That study showed that the ICP model was better at explaining Piston's table than actual music and suggested further that Woolhouse's model and Piston's table suffered from structurally similar distortions of tonal harmonic syntax." (Quinn 2010: 175). Figure 8 shows a stacked chart with scaled data reflecting Piston's table.

Even a shallow comparison with the data of Fig. 9 shows capital discrepancies. The correlation value with the Huron data presented in this Figure is considerably low ($r = 0.29$).

---

[8] For scale-restricted key profiles, the values are 0.171 and 0.410, respectively.

[9] What I mean by "deeper" motivation can be made clear by an analogy to the old problem of analysing consonance/dissonance. Both the model of Helmholtz (1877) and the Gestalt-theoretic model of psychologist Stumpf (1890) are attractive because they both are motivated by physical-acoustic and physiological arguments. With regard to Stumpf's tonal fusion idea, it was only recently that some authors could derive the underlying mechanism from certain properties of the neuronal periodicity detection mechanism (Ebeling 2008; Tramo 2001). An analogous argument is missing for Woolhouse's approach.
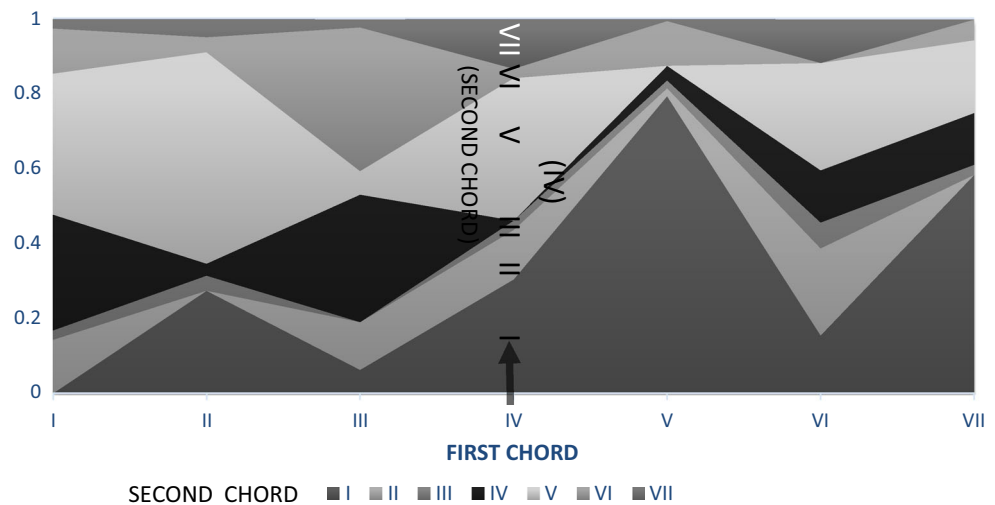
2.3 Comparing the models

The main problem of the traditional *tonnetz* approach is that it does not provide a consistent theory of similarity considering the three levels of analysis concerning tones, chords, and tonal regions. Further, there are many empirical issues which are discussed in length in the existing literature (Lerdahl 2001; Tymoczko 2011). The two models we have discussed in this section overcome some of the methodological and

Fig. 9 Stacked chart using corpus data of a sample of baroque music (data from Huron 2006: 251)

empirical issues. Both kinds of models have advantages and disadvantages I will discuss now (Fig. 8).

Empirically, I think the hierarchical model gives an adequate description for the attraction profiles for tones. However, there are many stipulations in the model. They concern the number of levels and the precise content of some levels. For instance, they concern the question of which chords constitute the triadic level. For Western music, the decision is easy to make by assuming that we have a clear distinction between major and minor systems. Non-Western kinds of music need not conform to the major/minor system and can be based on tonal scales quite different from those of Western music. Alternative scales such as Indian ragas or the scales underlying traditional Japanese music are widely used in world music. It is completely unclear how we can modify or extend the hierarchical model to account for the traits of these kinds of music.

A common criticism against generative linguistics is that it is founded on an English-biased view of the nature of language (such as the extended projection principle or Burzio's generalisation), which hinders sound typological work (Babby 2009). Similarly, a criticism against generative theories of music may include a powerful attack against the assumption of a universal tonal hierarchy, which is built on a bias towards Western music. I admit that—at least on a first glance—many assumptions of Lerdahl's and Jackendoff's (1983) generative theory of tonal music seem to be very plausible. This is not necessarily an advantage.[10]

Some authors, e.g. Katz and Pesetsky (2011) recommend optimality theory (Prince and Smolensky 1993/2004; Smolensky and Legendre 2006) as a means of exploring the

similarities between language and music. The presentation of Lerdahl's and Jackendoff's model of the tonal pitch space in an optimality-theoretic style has shown, unfortunately, that the instrument ceases to be useful in the present case. This is a consequence of the fact that the assumed constraints represent an inclusion hierarchy, where violations of lower constraints always entail violations of higher constraints. Hence, all that counts is the total number of constraint violations; the ranking of the constraints does not matter. In this vein, the conceptual core of optimality theory—constraint interaction and combinatorial typologies—cannot play any role. As a consequence, the core of the generativist strategy—to make a distinction between universal (innate) and learned knowledge—remains obscure in the musical domain.

Concerning the empirical facts, there is general agreement that for both major and minor profiles, scalar tones have higher values of tonal attraction than non-scalar tones. With reference to a piano this means that the white tones have higher values than the black tones (when considering C-major or A-minor). A second general finding is that all tones of the tonic triad have higher values than other tones of the scale (Temperley 2007: 84). These two important empirical facts are directly stipulated by the hierarchic model: by assuming a "diatonic space" (level D) which includes all scalar notes and by assuming a higher order "triadic space" (level C) that includes the tones of the triadic space.[11] The important methodological insight of the interval cycle model is that we have to "explain" these empirical findings rather than to describe them. Even if the ICP model finally fails for empirical reasons, its way of theorizing is important.

---

[10] Tucholsky recognized "the greatest enemy of truth is plausibility". Muns (2014) listed some of the "plausible" mistakes. Further, a recent article of Bierwisch (2014) criticizes an extreme way of generativist's argumentation.

[11] Another finding is that for major scales the tonic pitch has a higher attraction value than the fifth. In the hierarchic model, this suggests the assumption of an open fifth space (level B). Unfortunately, this conflicts with the Kostka–Payne corpus data (Kostka and Payne 1995). Consequently, the existence of this level of fifth space is questionable, for it does not apply for all assumed tonal scales.

The ICP model is important from a methodological point of view. The model seeks to derive the basic traits of major and minor attraction profiles, rather than to stipulate them. In the ICP model, *absolute* profiles are defined, taking interval-cycle proximity as an absolute function of interval length. These absolute profiles are key-independent. Absolute profiles are theoretical entities, i.e. they cannot directly be observed empirically. They abstract from the underlying tonal context, which in Western music is defined by a major or a minor scale.[12] A capital advantage of this approach is that it can also be applied to non-Western kinds of music. From the empirical point of view, the model is not really convincing, especially if we consider the full context-driven ICP-profiles.

An interesting possibility to check the models is to apply them to predict the proximity values of chords and tonal regions. As we have seen the ICP-model is able to approximate rather artificial and semi-empirical data (Piston's table), but it fails to account for real corpus data. The hierarchic model, on the other hand, can predict the basic impact of Weber's regional chart, but it fails to account for the similarity data for chord pairs, as investigated for instance by Huron (2006). Further, the hierarchical model predicts a symmetric similarity relation, which is a clear failure.

Summing up, the hierarchical model has some serious conceptual and empirical flaws. In contrast, the ICP-model makes an interesting methodological point. It tries to derive the observed phenomena and to fit the empirical data by assuming only *one* important principle: the principle of *interval-cycle proximity*. Unfortunately, the model is descriptively inadequate on all levels of investigation—tones, chords and regions.

3 The quantum model of tonal attraction

In this section, I will seek for a model that successfully combines the attractive methodological point of the ICP-model with the issue of empirical adequacy. Hence, similar to the ICP model of Woolhouse, I will start with defining *absolute* profiles, and then I will extend this theoretic core to derive context-dependent profiles of pitch attraction, and measures for regional and chordal similarities. These similarity relations will clearly be asymmetric. However, in contrast to Woolhouse's approach, the absolute profiles will not be based on interval cycles but on a new idea of a mathematical interpretation of the circle of fifth. This new understanding will be based on the idea of interpreting cognitive states as vectors in a Hilbert space[13] and of constructing a probability measure by means of projecting such states.

The new understanding of probabilities as projection properties of cognitive states is one of the most important developments in theoretical and mathematical psychology. In a series of papers, quantum probabilities are discussed as providing an alternative to classical probabilities for the understanding of cognition (Aerts 2009; Aerts et al. 2005; Blutner 2009; Blutner et al. 2013; Bruza et al. 2009a, b; Busemeyer and Bruza 2012; Busemeyer et al. 2006; Conte et al. 2008; Gabora and Aerts 2002; beim Graben 2004; Kitto 2008). In considerable detail, they point out several cognitive phenomena of perception, decision, and reasoning, which cannot be explained based on classical probability theory, and they demonstrate how quantum probabilities can account for these phenomena.

In their recent book, Busemeyer and Bruza (2012), give several arguments why quantum models are necessary for cognition. Some arguments relate to the cognitive mechanism of judgments. Judgments normally do not take place in definite situations. Rather, judgments create the context where they take place. This is the dynamic aspect of judgments also found in dynamic models of meaning (beim Graben 2013). An alternative aspect is the logical issue. The logic of judgments does not obey classical logic. Rather, the underlying logic is very strange with asymmetric conjunction and disjunction operations. When it comes to considering probabilities and conditioned probabilities the principle of *unicity* is violated, i.e. it is impossible to assume a single sample space with a fixed probability distribution for judging all possible events. Another line of argumentation seeks to answer the question "why quantum models of cognition" by speculating about implications for brain neurophysiology. In an algebraic approach, even classical dynamical systems such as neural networks could exhibit quantum-like properties, for example in the case of coarse-graining measurements, when testing a property cannot distinguish between epistemically equivalent states (beim Graben 2004).

3.1 Qubit states

In classical information science, a *bit* is the basic unit of information in computation referring to a choice between two discrete states, say {0, 1}. In contrast, a *qubit* is the

---

[12] Assume we perform an empirical study asking for the proximity between two pitches without explicitly presenting a context. It would be wrong to assume that we can measure the absolute profiles in this way. Instead, the subjects will automatically *infer* a tonal scale, i.e. they will decide about a scale which fits best into the presented pitch interval (and normally, the two pitches are assumed both to be elements of the inferred scale). Hence, we will get a profile relative to an inferred scale.

[13] A Hilbert space is a vector space upon which an inner product is defined and which makes use of complex numbers. If the vector state is assumed to describe a wave function, then the multiplication of the vector with a complex number allows an elegant description of phase shifts of the wave.

basic of information in quantum computing referring to a choice between two orthogonal unit-vectors in a two-dimensional Hilbert space. For instance, the orthogonal states $\varphi_\rightarrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\varphi_\uparrow = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ can be taken to represent *true* and *false*, the vectors in between are appropriate for modeling degrees of truth (vagueness) or degrees of expectation (probabilities).

The simplest non-trivial physical system is a two-state system, also called a *qubit* system. In such a system, each proper observable has exactly two (orthogonal) eigenvectors, say $\varphi_\rightarrow$ and $\varphi_\uparrow$. In the eigenstates of the observable the question asked by the observable has a certain outcome. Of course, a qubit can realize an infinite set of states but only two orthogonal states relate to eigenstates of the observable.

Formally, an arbitrary state of a qubit can be written as

$$\psi = \alpha\varphi_\uparrow + \beta\varphi_\rightarrow \text{ with } |\alpha|^2 + |\beta|^2 = 1 \qquad (3)$$

Making use of a particular parameterization of the states ψ every state of a qubit can be realized as the point on a three-dimensional sphere, the so-called Bloch sphere (Fig. 6, left-hand side).

$$\psi = \cos(\theta/2)\varphi_\uparrow + \sin(\theta/2)e^{+i\Delta}\varphi_\rightarrow \qquad (4)$$

The parameters θ and Δ are nothing but spherical polar coordinates, $0 \leq \Delta < 2\pi$ and $0 \leq \theta < \pi$.[14] One example of the realization of qubits is the spin of electrons. Hereby, it is possible to measure the spin in three "spatial" directions x, y and z. Another example is the polarization of photons. Hereby the state $\varphi_\uparrow$ represents a state with definite polarization in $\uparrow$-direction; the state $\frac{1}{\sqrt{2}}(\varphi_\uparrow - \varphi_\rightarrow)$ represents definite polarization in $\nwarrow$-direction; the superposition of states including a $\pi/2$ phase shift, such as $\frac{1}{\sqrt{2}}(\varphi_\uparrow - i\varphi_\rightarrow)$, represents circularly polarized light (Fig. 10).

For a simple illustration, consider a photon in a qubit state ψ and take $\varphi_\uparrow$ as indicating vertical polarization and $\varphi_\rightarrow$ as indicating horizontal polarization. Then the probability that the object is vertically polarized (i.e. it collapses into the state $\varphi_\uparrow$) is

$$P_\uparrow(\psi) = |\varphi_\uparrow \cdot \psi|^2 = \cos^2(\theta/2) = 1/2(1 + \cos(\theta)) \qquad (5)$$

[14] That means we can represent each unit vector ψ as a point on the unit sphere in the three-dimensional space by the coordinates $x = \sin(\theta)\cos(\Delta), y = \sin(\theta)\sin(\Delta), z = \cos(\theta)$. If the phase parameter Δ is zero, then the y-component vanishes and the three-dimensional phase can be reduced to a circle in the x, z-plane with coordinates $x = \sin(\theta)$ and $z = \cos(\theta)$. Note that the angle θ in the Bloch sphere corresponds to the angle $\theta/2$ in the original parameterization (4). Hence, two orthogonal vector states correspond to opposite points on the Bloch sphere.
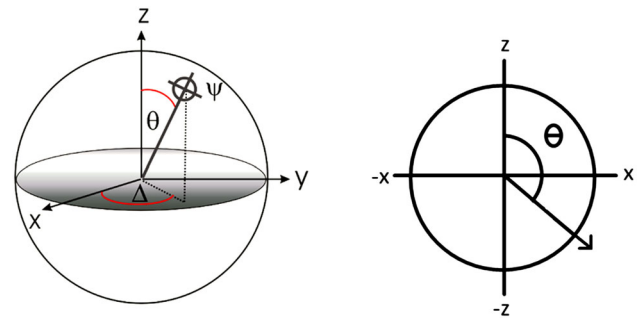

Fig. 10 *Bloch sphere*. Using Eq. (4) an arbitrary (normalized) state of the two-dimensional Hilbert space can be parameterized by the two spherical polar coordinates θ and Δ. Hereby, Δ corresponds to a phase shift of the two superposing states $\varphi_\uparrow$ and $\varphi_\rightarrow$. On the *right hand side*, a *Bloch circle* is shown resulting from the assumption of a zero phase shift ($\Delta = 0$)

Further, we can also calculate the probability that the object is polarized into a direction given by the superposition of $\varphi_\uparrow$ and $\varphi_\rightarrow$, say $\nwarrow = \frac{1}{\sqrt{2}}(\varphi_\uparrow - \varphi_\rightarrow)$. Interestingly, if the photon is described by ψ and collapses into the state $\varphi_\nwarrow$, then the calculated probabilities for the collapse also depend on the phase shift Δ:

$$\begin{aligned} P_\nwarrow(\psi) &= 1/2|(\varphi_\uparrow - \varphi_\rightarrow) \cdot \psi|^2 \\ &= 1/2(1 + \cos(\theta) \cdot \sin(\Delta/2)) \end{aligned} \qquad (6)$$

For the understanding of quantum cognition it is not required to give an interpretation in terms of some mysterious properties relating to the spin of electrons, the polarization of photons or other entities. In contrast to quantum mind theory (Hameroff and Penrose 1995), quantum cognition does not follow strategies of reducing mental entities to physical ones. Not unlike representatives of artificial intelligence who try to analyse big corpora using vector states modelling their distributive semantics (Widdows 2004), representatives of quantum cognition also work with vector states. The role of projections and quantum probabilities is essential in both cases.

Before we can apply the projection of states to calculate (probabilistic) key profiles, we have to introduce some basic ideas for symmetry groups.

3.2 Symmetry, group theory, and the principle of translation invariance

One of the fundamental ideas of quantum cognition is to apply the mathematics of the physical formalism to the domain of cognition. For example, we can use a series of qubit states to represent the 12 pitch classes used in tonal music. In addition, we can use the probability that one of these qubit state collapses into another one as a measure for the tonal attraction between the corresponding tones.

In Sect. 2, we have introduced a numeric notation to define the 12 tones of the system. For convenience, it is repeated here:

$$0 = C, 1 = C\sharp, 2 = D, 3 = D\sharp,$$
$$4 = E, 5 = F, 6 = F\sharp, 7 = G, 8 = G\sharp, \qquad (1)$$
$$9 = A, 10 = B\flat, 11 = B$$

There are certain actions or operations that allow transforming tones into other tones. For instance, we can increase the tones by a certain number of steps $(0, 1, 2, \ldots, 11)$. Such actions are called *translations*. The 1-step translation transforms c into c♯, c♯ into d, and so on. Actions can be combined. For example, we can combine the translation of a 2-step increase with a 3-step translation, resulting in a 5-step translation (in other words, a major second combined with a minor third gives a fifth). We will denote these operations likewise with the numbers $0, 1, 2, \ldots, 11$. Normally, the context makes clear what the numbers denote: a pitch class or the operation of increasing tones by a number of elementary steps. It is obvious that the combination of acts of translations can be described by addition (modulo 12): $x + y \bmod 12$; e.g., $2 + 3 \bmod 12 = 5, 7 + 6 \bmod 12 = 1$.

At this point, it is indispensable to introduce some basic concepts of group theory.[15] Generally, a group consists of a set of (abstract) elements and a binary operation defined on it. Usually, this operation is written with a product sign, for example $g_1 \cdot g_2 \in G$ (the product sign "·" can be left out). The following properties have to be satisfied:

1. all elements of G are connected by the group operation, i.e., for all elements $g_1, g_2 \in G$ it holds that $g_1 \cdot g_2 \in G$
2. There is a particular element $e \in G$ (the neutral element) such that for all elements $g \in G$ it holds that $e \cdot g = g \cdot e = g$.
3. The associative law is valid, i.e. for all elements $g_1, g_2, g_3 \in G$ we have $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$.
4. For each element $g \in G$ there exists an inverse element g^{-1}, which has the property $g \cdot g^{-1} = g^{-1} \cdot g = e$.

In the case of music based on 12 tones, we have to consider the set of group elements $\{0, 1, 2, \ldots, 11\}$, and the group operation is $x \cdot y = x + y \bmod 12$. The neutral element is the element denoted by $0 : (0 + x) \bmod 12 = (x + 0) \bmod 12 = x$. For the inverse element x^{-1}, we have $x^{-1} = (12 - x) \bmod 12$.

A group G is called *cyclic* if there exists a single element $g \in G$ such that every element in G can be represented as a composition of g's. The element g is called a *generator* of the group. If a cyclic group has n elements (i.e. the group is of order n), the group can be represented as $\mathbb{Z}_n = \{e, g, g^2, g^3, \ldots, g^{n-1}\}$, where $g^n = e$. In the present numerical representation of the cyclic group $\mathbb{Z}_{12}$ we have four generators conforming to the numbers 1, 11, 7, 5.[16] Hence, 1 (upward) and 11 (downward) generate the sequence of semitones. In addition, the elements 5 and 7 enumerate the group elements in successive fifths or fourths—representing the circle of fifths. If x is an integer variable running from 0 to 11, we can generate the group elements in the respective four cases in the following way, where the variable x runs from 0 to 11:

(a) $x + 1 \bmod 12$
(b) $x + 11 \bmod 12 (= 12 - x \bmod 12)$
(c) $x + 7 \bmod 12$ $\qquad (7)$
(d) $x + 5 \bmod 12 (= x - 7 \bmod 12)$.

Figure 11 gives a visual illustration of two ways to generate the cyclic group $\mathbb{Z}_{12}$. On the left-hand side, the generator of the group is the action of increasing the tones by one semitone. On the right-hand side, the generator is the action of increasing tones by seven semi-tones. Hence, when we apply the group generator to the tone C in the first case, we get C♯. In the second case we get the tone G. The construction in the second case does exactly what usually is represented by the circle of fifth.

Next, let us introduce the concept of symmetry. This concept plays an important role in many areas of science, including classical mechanics, quantum mechanics, chemistry, crystallography, and theoretical biology. In music, it is indispensable for a mathematical understanding of modulation theory and counterpoint (Mazzola 2002; Mazzola et al. 1989). Mathematically, symmetry is simply a set of transformations applied to given structural states such that the transformations preserve the properties of the states. In music, the most basic symmetry principle is the *principle of translation invariance*. It says that the musical quality of a musical episode is essentially unchanged if it is transposed into a different key, i.e. if the operations of the cyclic group $\mathbb{Z}_{12}$ are applied. Therefore, we can say that $\mathbb{Z}_{12}$ is the symmetry group of (Western) music.

In mathematics, the word *representation* means a structure-preserving function. In group theory, a representation is simply a homomorphism. The object of our investigation is the symmetry group of translations. The homomorphism we seek for should map this group to a more concrete group that is in some sense easier to understand

---

[15] For the interested reader, we refer to standard textbooks of group theory (Alexandroff 2012; Jones 1998) or good introductory sections or appendixes in music-theoretic treatises (Balzano 1980; Honingh 2006)

[16] The general rule is that the generators for a cyclic group Z_n are exactly the numbers i in the range $0 < i < n$ such that n and i are relatively prime (i.e., they have no common factor differently from 1).
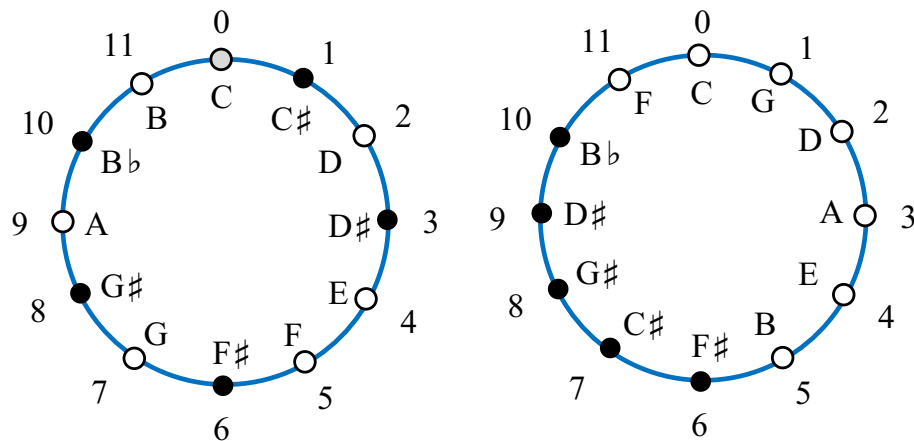
Fig. 11 Visual representation of $\mathbb{Z}_{12}$. On the *left hand side* the different elements of the group are generated by the semi-tone generator. The *white dots* give an ordered subset of $\mathbb{Z}_{12}$ starting with the tone 0. It is the diatonic scale of C-major. On the *right hand side*, the group elements are generated by a generator that transposes by seven semi-tones (resulting in the circle of fifth). The *numbers* indicate how often the generator is applied recursively. The *tones* in the *inner circles* are the results of application of the corresponding group element to the basic pitch class C

than the original one. For example, this group could consist of linear maps as studied in linear algebra. More concretely, the group could consist of certain rotations of vectors in a two-dimensional vector space. For instance, we can rotate the vector $\varphi_{\rightarrow} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in n steps to the original vector. In linear algebra, the elementary rotation steps can be described by the following rotation matrix γ:

$$\gamma = \begin{pmatrix} \cos(2\pi/n) & \sin(2\pi/n) \\ -\sin(2\pi/n) & \cos(2\pi/n) \end{pmatrix} \tag{8}$$

One application of this matrix to the vector $\varphi_{\uparrow} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ results in the vector $\gamma \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin(2\pi/n) \\ \cos(2\pi/n) \end{pmatrix}$. This is a rotation of the original vector by an angle of $2\pi/n$. It is not difficult to see that the generator γ as defined in (8) generates the cyclic group $\mathbb{Z}_n$. For $n = 12$, the group elements of this group can be enumerated as follows, where k runs from 0 to 11:

$$\gamma^k = \begin{pmatrix} \cos(2\pi k/12) & \sin(2\pi k/12) \\ -\sin(2\pi k/12) & \cos(2\pi k/12) \end{pmatrix} \tag{9}$$

In this way, we can generate a series of vector states ψ_k representing the 12 tones. In (10a) these states are given as vectors in a two-dimensional real Hilbert space (we have assumed zero phases). In the Bloch sphere, these vectors are represented as in (10b). The y-component is zero because of the zero phase. Note that the angles in (10a) are half of the ones in (10b). Hence, the triton in the vector picture is orthogonal to the tonic tone (angle $\pi/2$). But in the Bloch sphere the two points are on opposite sides of the sphere; hence, their angle is π.

(a) $\psi_k = \gamma^k \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin(\pi k/12 \\ \cos(\pi k/12) \end{pmatrix}$

(b) $x_k = \sin(\pi k/6), z_k = \cos(\pi k/6).$ (10)

Importantly, we have to consider two different ways of enumeration, corresponding to two generators of the group $\mathbb{Z}_{12}$. One enumerates the pitches in a chromatic (ascending) way; the other enumerates the tones according to the (ascending) circle of fifth. In this way, we get two Bloch circles, which exactly correspond to the two circles shown by Fig. 11. Which of these two representations of tones is the preferred one depends of an empirical decision. This decision is not difficult in the present case because we intend to express the similarity relation between tones, tonal regions, or chords. In the next section, I will demonstrate that this clearly favours the circle of fifth.

3.3 Key profiles in the quantum model

In the case of pure states, quantum theory defines *structural* probabilities. This means the probability that a state ψ collapses into another state depends exclusively on the geometric, structural properties of the considered states. How well does a given tone fit with the tonic pitch? What is the probability that it collapses into the (tonic) comparison state? The probability of a collapse of the state ψ_k into a state ψ_l can be calculated straightforwardly:

$$P_{\psi_l}(\psi_k) = \cos^2(\pi(k-l)/12)$$
$$= 1/2(1+\cos(\pi(k-l)/6)), \text{ where } 0 \le k, l < 12. \tag{11}$$

For a fixed element ψ_l the probabilities of the 12 tones indexed by k ($0 \le k < 12$) sum up to 1. Hence, formula (11)
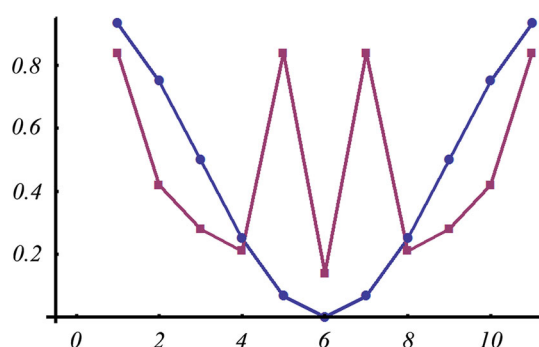
Fig. 12 Comparison between the profile resulting from interval cycles (*squares*) and the quantum-probabilistic profile (*circles*)

offers a (probabilistic) attraction profile relative to a given tone ψ_l. We can compare it with the absolute attraction profile resulting from tonal cycles (see Fig. 12).

Figure 12 illustrates that the profile resulting from interval cycles and the profile resulting from the quantum model are very different. The correlation between both profiles is very weak. The correlation coefficient is $r = 0.27$, i.e. the correlation explains $r^2 = 0.063$ or 6.3 % of the variance of each data set. Hence, we can conclude that the two models are based on two quite different assumptions about the absolute profiles.

If the comparison state is not a single tone, but a tonal region, a chord, or a series of chords, then I will consider the mixture of all the states conforming to all the involved single tonal elements. For simplicity, I will take all tones that go into this mixture as equivalent and give them the common weight $1/N$ (assuming N tonal elements are to consider). This assumption is rather similar to Woolhouse's treatment of the problem of context effects in tonal attraction (Woolhouse 2009, 2010; Woolhouse and Cross 2010).

Figure 13 shows the key profiles for major and minor keys using the quantum model and scales it to the Krumhansl and Kessler data considered before.

The correlation coefficient between the predicted profile and the Krumhansl–Kessler profile is $r = 0.78$ in the case of major keys and $r = 0.69$ in the case of minor keys. Remember the correlation coefficients for the full chromatics scales using the ICP model: $r = 0.089$ in the major case, and $r = 0.045$ in the minor case.[17]

3.4 Symmetry breaking and the learning of key profiles

A qubit can be characterized by two parameters θ and Δ as described by formula (4). The parameter Δ describes the phase shift between the two orthogonal "wave functions" $\varphi_\uparrow$

---

[17] Since we do not know the detailed statistics of the Krumhansl–Kessler data, it cannot be decided at what level the differences to the quantum model are significant.

and $\varphi_\rightarrow$ (representing the tonic and the corresponding triton). Δ was set to be zero so far. Now we will generalize the earlier model by assuming that non-zero phase shifts can be involved. That means we replace formula (10a) by the following expression for the states ψ_k expressing the tones (in an enumeration conforming to the circle of fifth):

$$\psi_k = \begin{pmatrix} \sin(\pi k/12) \\ e^{i\Delta_k} \cos(\pi k/12) \end{pmatrix} \tag{12}$$

We will consider the phase parameters Δ_k as free parameters that are determined by learning processes. In general, the parameters can break the symmetry that originally conformed to the symmetry group $\mathbb{Z}_{12}$. This is a case of symmetry breaking by learning, which is prominently investigated in connectionist modelling (e.g., Földiák 1991).

In the present context, we fit the phase parameters with the Krumhansl–Kessler data. The result of the fit is shown in Fig. 14.

The correlation coefficient between the model fit and the Krumhansl–Kessler profile is $r = 0.95$ in the case of major keys and $r = 0.97$ in the case of minor keys (explaining 90 and 94 % of the variance, respectively). This result is comparable to the hierarchical model ($r = 0.97$ for major keys and $r = 0.93$ for harmonic minor keys).

3.5 Similarity between regions

For the calculation of the similarity between regions we use a measure that is commonly used in quantum information science: the Kullbeck–Leibler distance (also called *relative entropy*). It is defined as follows, where p and q denote two probability distributions:

$$KL(p/q) = \Sigma_k\, p_k \log_2(p_k/q_k) \tag{13}$$

The index k ranges over all events of a given partition of the sampling space. Typically, one of the distributions represents empirical observations, the other an approximating model. Intuitively, the Kullbeck–Leibler distance is the expected number of bits required to code samples for p when using a code optimized to code samples for q. The Kullbeck–Leibler distance is closely related to cross entropy $H(p/q) = -\Sigma_k\, p_k \log_2(q_k)$, which was introduced into music theory by Temperley (2007). The connection is $KL(p/q = H(p/q) - H(p)$. The Kullbeck–Leibler distance has interesting mathematical properties. For instance, it is a convex function of p_k, is always nonnegative, and equals zero only if $p_k = q_k$, for all k (Cover and Thomas 1991). It is not really a distance in the strict sense, for it is an asymmetric function.

Table 4 shows two different distance models for regions based on the Kostka–Payne profiles. One is based on the

Fig. 13 The *dashed curves* show the Krumhansl–Kessler profiles (*left* major keys; *right* minor keys). The *bold curves* are the theoretical predictions of the quantum model
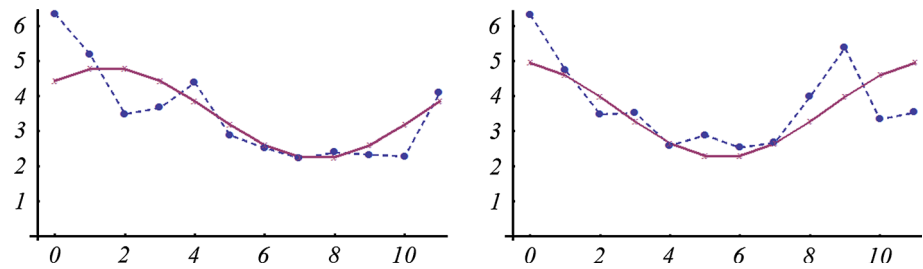

Fig. 14 Comparison of the KK key profiles with the data from the quantum model with fitted phase factors. The *dashed curves* show the KK key profiles (*left* major keys; *right* minor keys). The *bold curves* are the theoretical predictions
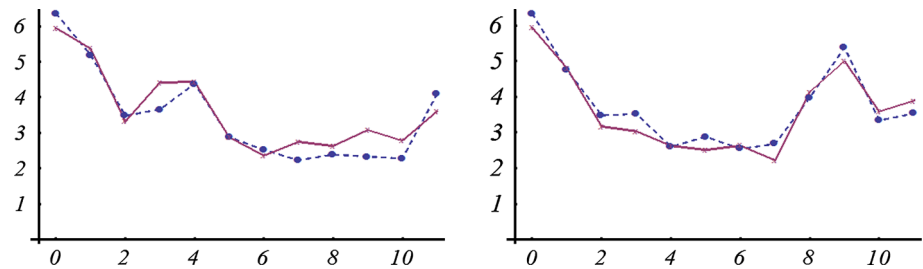

Table 4 Distances from C major and A minor to all other keys using correlation and relative entropy

Key	Kostka–Payne profiles correlation method		Kostka–Payne profiles relative entropy		Quantum model relative entropy	
	C major	A minor	C major	A minor	C major	A minor
C	1.000	0.646	0.00	1.34	0.00	0.017
C♯/D♭	−0.575	−0.448	6.93	6.17	0.242	0.260
D	0.146	0.198	3.19	3.78	0.064	0.017
D♯/E♭	−0.065	−0.421	4.84	5.20	0.130	0.195
E	−0.224	0.251	4.85	3.41	0.194	0.130
F	0.638	0.521	1.51	1.99	0.017	0.065
F♯/G♭	−0.840	−0.709	7.93	7.27	0.260	0.242
G	0.638	0.343	1.24	2.63	0.017	0.00
G♯/A♭	−0.224	−0.355	5.77	5.30	0.194	0.242
A	−0.065	0.436	4.08	3.01	0.130	0.065
A♯/B♭	0.146	−0.009	3.64	3.84	0.064	0.130
B	−0.575	−0.414	6.70	6.13	0.242	0.195
c	0.436	−0.072	2.34	3.44	0.065	0.130
c♯/d♭	−0.355	−0.002	5.64	4.48	0.242	0.195
d	0.343	0.374	2.29	2.51	0.00	0.017
d♯/e♭	−0.709	−0.542	6.77	5.73	0.242	0.260
e	0.521	0.374	1.74	2.66	0.065	0.017
f	0.251	−0.002	3.45	3.65	0.130	0.195
f♯/g♭	−0.421	−0.072	5.20	4.56	0.195	0.130
g	0.198	−0.165	3.40	4.44	0.017	0.065
g♯/a♭	0.488	−0.364	6.15	5.18	0.260	0.242
a	0.646	1.000	0.89	0.00	0.017	0.00
a♯/b♭	−0.414	−0.364	6.00	5.69	0.195	0.242
b	−0.009	−0.165	3.91	4.38	0.123	0.065

It also makes the comparison between the quantum model ($\Delta = 0$) and the Kostka–Payne data

correlation method; the other is based on relative entropy. The distances are calculated from C major and A minor to all other keys.

The two columns on the left-hand side give the correlation coefficients based on the Kostka–Payne profiles. The two columns in the middle are the relative entropies for the

Fig. 15 Stacked chart based on the quantum model with zero phases
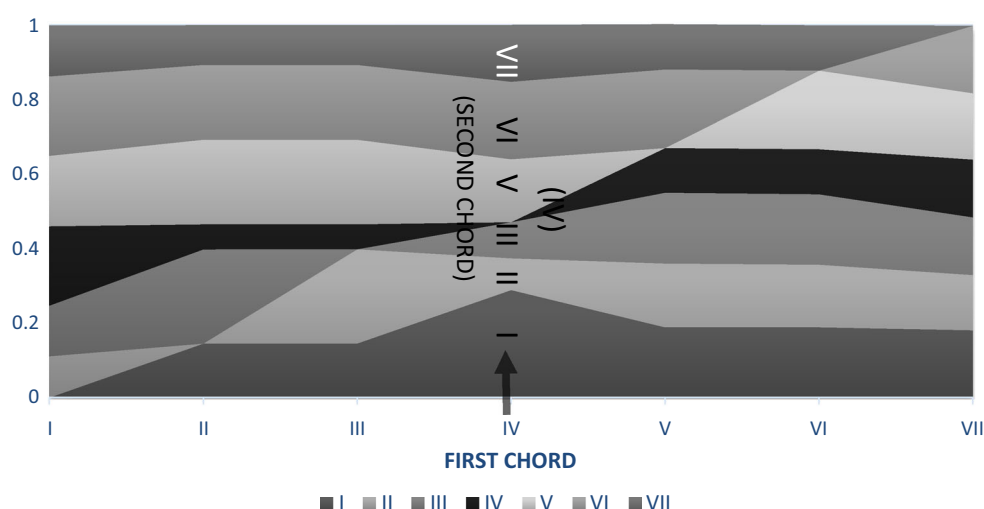
corresponding profiles. On the right-hand side, the empirical predictions of the quantum model (with zero phases) are presented.

Note that the Kostka–Payne profiles do not really present probability distributions. For each key, the numbers for the 12 pitch classes do not add up to 1. In contrast, the relative entropies according to the quantum model are calculated on the basis of probability distributions. Therefore, you should not expect similar values for the empirical and theoretical predictions. Rather, the number columns should be similar when taking into account a certain scaling factor. This can be tested via the correlation values between the two columns of relative entropies, which are independent of scaling factors.

The correlations for the relative entropies according the Kostka–Payne data and the quantum model are $r = 0.97$ (between C-major and the major keys), $r = 0.91$ (between C-major and the minor keys), $r = 0.88$ (between A-minor and the major keys), $r = 0.77$ (between A-minor and the minor keys). A comparison with the Krumhansl–Kessler profiles shows very similar results.[18] Further, it is notable that the correlation values increase to almost 1 if the fitted phase factors are included. Without going into a detailed comparison, it can be said that the quantum model gives a good approximation to the distances established by Weber's regional chart (see Fig. 3).

3.6 Similarity between chords

Finally, let us compare the predictions of the quantum model with the similarity between chords as found by Huron's (2006) corpus analysis and presented in Fig. 8. Using the

quantum model, I first calculated the conditioned probabilities of all tones of the considered scale triggered by each of the seven chords. To get the conditioned probabilities for chords (instead of single tones) I simply used the product of the corresponding conditioned probabilities for each of the three tones of the chord. The stacked chart for the quantum model representing the conditioned probabilities for all chords is given in Fig. 15.

Obviously, the comparison with Fig. 8 shows that the predictions of the quantum model with zero phases are far from fitting the Huron (2006) data. The averaged correlation between the Huron data and the quantum model with zero phases is $r = 0.48$. Remember the corresponding correlation value for the Lerdahl-model: $r = 0.21$; and for the interval cycles model: $r = 0.29$. Surprisingly, we do not get an improvement of the fit when the (learned) phase parameters are involved: $r = 0.35$.

I mentioned already the important distinction between tonal hierarchies and event hierarchies introduced by Lerdahl (1988). Empirical data that concern "chord progression" should be explained in term of such event hierarchies. The learning of event structures seems to be quite different from the learning of key attraction profiles, which relates to tonal hierarchies in Lerdahl's model. If we assume that there is a partial overlap of innate knowledge about temporal musical sequences and tonal attraction, then we are able to understand why the symmetry breaking by learning attraction values can challenge the prediction of musical sequences and weaken the degree of correlation between the quantum model and Huron's (2006) chord progression data.

4 General discussion and conclusions

Structural and probabilistic approaches in computational music theory have tried to give systematic answers to the

---

[18] The correlations are: $r = 0.95$ between C-major and the major keys; $r = 0.91$ between C-major and the minor keys; $r = 0.90$ between A-minor and the major keys; $r = 0.82$ between A-minor and the minor keys.

problem of tonal attraction. I have discussed two previous models of tonal attraction, one based on tonal hierarchies (Lerdahl 1988, 2001) and the other based on interval cycles (Woolhouse 2009, 2010; Woolhouse and Cross 2010). Both models aim to account for the phenomenon of tonal attraction at the level of pitches, regions, and chords.

Unfortunately, both models have serious limitations. The hierarchical model has serious conceptual flaws because it stipulates the empirical generalisations rather than predicting them. Further, it envisages symmetric similarity relations between regions and chords, whoch cannot be correct empirically. The ICP-model, on the other hand, has interesting methodological advantages. Unfortunately, it is descriptively inadequate on all levels of investigation—pitches, chords, and regions.

To overcome the shortcomings of these models, both methodologically and empirically, I proposed a new probabilistic model relying on insights of quantum cognition. I have argued that the quantum approach integrates the insights from both group theory and quantum probability theory. In some sense, the model integrates the conceptual advantages of the ICP model with the empirical prospects of Lerdahl's tonal hierarchies. The present model does not incorporate information that listeners infer from temporal musical sequences. According to Lerdahl (1988), these are the effects of event hierarchies, which deserve a special treatment—possibly along the lines of Mazzola (2002) and Mazzola et al. (1989).

In his recent book, Philip Ball outlines that at the heart of any scientific explanation of music is an understanding of how and why it affects us (Ball 2010). From generativist theory building, we can learn that a basic pillar of scientific explanations is a careful distinction between different levels of description. I think it is a good idea to start with the perceptual and the cognitive level and to add a level of affective meaning. The perceptual level is investigated in psychoacoustics. It relates the relevant physical properties of sensory stimuli and the psychological responses evoked by them. The cognitive level refers to psychological processes which go beyond the purely sensual processes such as in the musical context of event hierarchies or counterpoint. The distinction is standard in approaches to consonance/dissonance where it relates to the distinction between perceptual (or sensory) consonance/dissonance and musical consonance/dissonance, following Rasch and Plomp (1999). If I consider affective meaning as the third level of musical representation, I have in mind the form of meaning which Meyer (1956) called *embodied meaning*. This term refers to the significance a musical event can have for a listener in terms of its own structure and in interaction with the listener's musical expectations. In his seminal book, Meyer (1956) pointed out that the principal emotional content of music arises through the composer's arranging of expectations. Composers sometimes satisfy our

expectations, sometimes delay an expected outcome or even thwart it, and sometimes composers play with ambiguities avoiding any clear expectations to be established. The secret to composing a likeable song is to balance predictability and surprise. Because most music has a beat and is based on repetition, we know when the next musical event is likely to happen, but we do not always know what it will be. Our brains are working to predict what will come next. The skillful composer rewards our expectations often enough to keep us interested, but violates those expectations the rest of the time in interesting ways.

The mathematical treatment of expectations is in terms of probabilities, let it be classical Bayesian probabilities (Oaksford and Chater 2007) or non-classical quantum probabilities (Busemeyer and Bruza 2012). In this paper, we have investigated the problem of tonal attraction and we have argued in favour of a probabilistic approach in terms of quantum probabilities. In this way, we have presented a framework for expressing and handling expectations. Looking at future work, this could be one of the building blocks for realising the mapping between music and its affective (emotional) answer. Another building block relevant for realizing the ultimate aim of connecting musical structures with affective meaning is the proper characterization of the qualitative character of chords in terms of consonance and dissonance.

A surprising outcome of this paper is that we can make a judgment on the percentage of variance that comes from the symmetry conserving quantum model (and possibly can be seen as innate and not learnable) and that part of the quantum model that deals with symmetry breaking and fixing the phases. The first part of the variance is about 50 %, the second part is about 40, and 10 % cannot be explained. The symmetry breaking parts introduce significant differences between the various keys. In the history of music theory such differences have been significantly doubted by authors such as Helmholtz (1877). Others have vehemently argued for them, e.g. Beckh (1937). Of course, we have to leave this issue unresolved here.

If the model's distinction between learned (phases) and more or less innate knowledge (symmetric structures) contains a bit of truth, then this is a powerful argument for the idea that the quantum model can realize important issues of the generative tradition. Relating to the innateness issue, I do not see any generativist wisdom in the model of tonal hierarchies. Without a careful treatment of symmetry principles such as the *principle of translation invariance* a cognitive theory of tonal music is not possible (Balzano 1980; Honingh 2006; Mazzola 2002).

References

Aerts D (2009) Quantum structure in cognition. J Math Psychol 53:314–348

Aerts D, Czachor M, D'Hooghe B (2005) Do we think and communicate in quantum ways? On the presence of quantum structures in language. In: Gontier N, Bendegem JPV, Aerts D (eds) Evolutionary epistemology, language and culture. Studies in language, companion series. John Benjamins Publishing Company, Amsterdam

Alexandroff P (2012) An introduction to the theory of groups. Dover, New York

Babby LH (2009) The syntax of argument structure, vol 120. Cambridge University Press, Cambridge

Ball P (2010) The music instinct: how music works and why we can't do without it. Bodley Head, London

Balzano GJ (1980) The group-theoretic description of 12-fold and microtonal pitch systems. Comput Music J 4:66–84

Beckh H (1937) Die Sprache der Tonart in der Musik von Bach bis Bruckner: Vom geistigen Wesen der Tonarten. Urachhaus, Stuttgart

Beim Graben P (2004) Incompatible implementations of physical symbol systems. Mind Matter 2:29–51

Beim Graben P (2013) Order effects in dynamic semantics. Top Cognit Sci 6:67–73

Bierwisch M (2014) The role of arbitrariness from a minimalist point of view. In: Kosta P, Franks SL, Radeva-Bork T, Schürcks L (eds) Minimalism and beyond: radicalizing the interfaces. John Benjamins Publishing Company, Amsterdam, pp 392–415

Blutner R (2009) Concepts and bounded rationality: an application of Niestegge's approach to conditional quantum probabilities. In: Accardi L, Adenier G, Fuchs C, Jaeger G, Khrennikov AY, Larsson J-Å, Stenholm S (eds) Foundations of probability and physics-5. American Institute of Physics conference Proceedings, vol 1101, New York, pp 302–310

Blutner R, Pothos EM, Bruza P (2013) A quantum probability perspective on borderline vagueness. Top Cognit Sci 5:711–736

Bruza P, Busemeyer JR, Gabora L (2009a) Special issue on quantum cognition. J Math Psychol 53:303–305

Bruza P, Kitto K, Nelson D, McEvoy CL (2009b) Is there something quantum-like about the human mental lexicon? J Math Psychol 53:362–377

Busemeyer JR, Bruza PD (2012) Quantum cognition and decision. Cambridge University Press, Cambridge

Busemeyer JR, Wang Z, Townsend JT (2006) Quantum dynamics of human decision-making. J Math Psychol 50:220–241

Conte E, Khrennikov AY, Todarello O, Robertis RD, Federici A, Zbilut JP (2008) A preliminary experimental verification on the possibility of bell inequality violation in mental states. NeuroQuantology 6:214–221

Cover TM, Thomas JA (1991) Elements of information theory. Wiley, New York

Deutsch D, Feroe J (1981) The internal representation of pitch sequences in tonal music. Psychol Rev 88:503–522

Ebeling M (2008) Neuronal periodicity detection as a basis for the perception of consonance: a mathematical model of tonal fusion. J Acoust Soc Am 124:2320–2329

Euler L (1739) Tentamen novae theoriae musicae. Teubner, St. Petersburg (Stuttgart 1926)

Földiák P (1991) Learning invariance from transformation sequences. Neural Comput 3:194–200

Gabora L, Aerts D (2002) Contextalizing concepts using a mathematical generalization of the quantum formalism. J Exp Theor Artif Intell

Hameroff SR, Penrose R (1995) Orchestrated reduction of quantum coherence in brain microtubules: a model for consciousness? Neural Netw World 5:793–804

Harnad S (1990) The symbolic grounding problem. Physica D 42:335–346

Helmholtz HV (1877) On the sensations of tones (A. J. Ellis, Trans.). Dover, New York (reprinted 1885 and 1954)

Honingh A (2006) The origin and well-formedness of tonal pitch structures. Ph.D., Universiteit van Amsterdam

Huron D (2006) Sweet anticipation: music and the psychology of expectation. The MIT Press, Cambridge

Jones H (1998) Groups, representations and physics, 2nd edn. Taylor & Francis, New York

Katz J, Pesetsky D (2011) The identity thesis for language and music

Kitto K (2008) Why quantum theory? In: Bruza P, Lawless W, van Rijsbergen C, Sofge D, Coecke B, Clark S (eds) Proceedings of second quantum interaction symposium, Oxford, pp 11–18

Kostka S, Payne D (1995) Tonal harmony. McGraw-Hill, New York

Krumhansl CL (1979) The psychological representation of musical pitch in a tonal context. Cognit Psychol 11:346–374

Krumhansl CL, Kessler EJ (1982) Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. Psychol Rev Gen Psychol 89:334–368

Lerdahl F (1988) Tonal pitch space. Music Percept 5:315–350

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

Lerdahl F, Jackendoff R (1983) A generative theory of tonal music. The MIT Press, Cambridge

Mazzola G (2002) The topos of music: geometric logic of concepts, theory, and performance. Birkhauser, Basel

Mazzola G, Wieser H-G, Bruner V, Muzzulini D (1989) A symmetry-oriented mathematical model of classical counterpoint and related neurophysiological investigations by depth EEG. Comput Math Appl 17:539–559

Meyer LB (1956) Emotions and meaning in music. Chicago University Press, Chicago

Muns L (2014) Classical music and the language analogy. Ph.D., Humboldt-University

Newell A (1980) Physical symbol systems. Cognit Sci 4:135–183

Oaksford M, Chater N (2007) Bayesian rationality: the probabilistic approach to human reasoning. Oxford University Press, Oxford

Piston W (1979) Harmony. W. W. Norton, New York

Pothos EM, Busemeyer JR, Trueblood JS (2013) A quantum geometric model of similarity. Psychol Rev 120:679–696

Prince A, Smolensky P (1993/2004) Optimality theory: constraint interaction in generative grammar. Technical report RuCCSTR-2, available as ROA 537-0802. Revised version published by Blackwell, 2004, Rutgers University and University of Colorado at Boulder

Quinn I (2010) On Woolhouse's interval-cycle proximity hypothesis. Music Theory Spectr 32:172–179

Rasch R, Plomp R (1999) The perception of musical tones. In: Deutsch D (ed) The psychology of music, vol 2. Academic Press, London, pp 89–112

Schönberg A (1969) Structural functions of harmony. Norton, New York (originally published 1954)

Smolensky P, Legendre G (2006) The harmonic mind: from neural computation to optimality-theoretic grammar. MIT Press, Cambridge

Stumpf C (1890) Tonpsychologie. Hirzel, Leipzig

Temperley D (2007) Music and probability. MIT Press, Cambridge

Tramo MJP, Cariani A, Delgutte B, Braida LD (2001) Neurobiological foundations for the theory of harmony in Western tonal music. In: Zatorre J, Peretz I (eds) The biological foundations of music. Wiley, Online Library, New York, pp 92–116

Tversky A (1977) Features of similarity. Psychol Rev 84:327–352

Tymoczko D (2011) A geometry of music. Oxford University Press, Oxford, New York

Weber G (1824) Versuch einer geordneten Theorie. B. Schotts Söhne, Mainz

Widdows D (2004) Geometry and Meaning. CSLI, Stanford University

Woolhouse M (2009) Modelling tonal attraction between adjacent musical elements. J New Music Res 38:357–379

Woolhouse M (2010) Modes on the move: interval cycles and the emergence of major–minor tonality empirical musicology. Review 5:62–83

Woolhouse M, Cross I (2010) Using interval cycles to model Krumhansl's tonal hierarchies music theory. Spectrum 32:60–78

Woolhouse M, Cross I (2011) To The editor: A response to Ian Quinn. Music theory spectrum 33:99–105