# A Gauge-Theoretic Approach to Musical Forces

Reinhard Blutner
Emeritus member of ILLC, Universiteit van Amsterdam, Amsterdam, The Netherlands.
email: blutner@contact.uva.nl

Peter beim Graben
TU Cottbus, Institute of Electronics and Information Technology, Cottbus, Germany
email: Peter.beimGraben@b-tu.de

### Abstract

Metaphors involving motion and forces are a source of inspiration for understanding tonal music and tonal harmonies since ancient times. Starting with the rise of quantum cognition, the interactional conception of forces as developed in modern gauge theory has recently entered the field of theoretical musicology. We argue that the metaphoric conception of musical forces is unable to give a causal explanation of the essential mechanism of tonal dynamics. Rather, it describes correlation without explaining the mechanism that causes these correlations. The gauge-theoretic conception, in contrast follows a realistic foundation of physical and extra-physical forces. It identifies universal features that are necessary to the unique character and causal function of physical and extra-physical forces. In this article, we consider three gauge models of tonal attraction. The *phase model* borrows ideas from quantum electrodynamics. It is based on U(1) gauge symmetry. The *spatial deformation model*, in contrast, borrows its main idea from the general theory of relativity an can be seen being based on SO(2) gauge symmetry. In the neutral, force-free case both models agree and generate the same predictions as a simple qubit approach. However, there are several differences in the force-driven case. It is claimed that the deformation model gives a proper description of static tonal attraction (tonal hierarchies). The third model is a *combination* of the deformation model with the phase model. It is based on SU(2) gauge symmetry and yields a unifying approximation to dynamic attraction data (resolution of chords).

# 1 Introduction

The phenomenon of tonal attraction has fascinated researchers of music psychology, both empirically and theoretically (Krumhansl & Cuddy, 2010). There is a distinction between two types of tonal attraction, called static and dynamic attraction (Blutner, 2016; Parncutt, 2011). How well does a given pitch fit into a tonal scale or tonal key, being either a major or minor key? This is a question of the first type concerning the tonal centers. A question of the second, dynamic type, typically asks for the level of resolution a subject feels when she hears a probe tone following a certain chord in a serial sequence.

In an celebrated study by Krumhansl and Kessler (1982) the static type of tonal attraction was investigated. In this study, listeners were asked 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 followed by the chromatic, non-scale tones. This finding plays an essential role in Lerdahl's and Jackendoff's generative theory of tonal music (Lerdahl & Jackendoff, 1983). It clearly counts as one of the main pillars of the structural approach in music theory. A related approach of the static type is due to Bharucha (1996). The dynamic type of attraction was investigated by Krumhansl (1990, 1995), Lake (1987), Bharucha (1996), Lerdahl (1996), Larson (2004), Larson (2012), and in a recent study of Woolhouse (2009), following earlier research of Brown, Butler, and Jones (1994).

Both types of tonal attraction have not only initiated an enormous number of empirical studies but also challenged a series of different models based on static and dynamic forces. Most of these models are close in inspiration to Larson (2012). All models explicitly or implicitly consider the term "musical forces" as a metaphoric term and build a phenomenological model on this basis. The models that exhibit the metaphorical trait aim at describing correlation. They do not aim describing the causal mechanism underlying tonal attraction.

The work of Mazzola (1990, 2002) is an important exception to this widely shared methodology. His theory sees the whole conception of "musical force" as directly rooted in the basic symmetry principles of tonal music. To distinguish the phenomenological (and metaphoric) idea of forces from the idea based on fundamental symmetries and physical interaction, we will call the latter view the *structural realistic view* (or short: *realist view*). The idea is that the force conception plays a causal role in the theory. It is associated with a set of structural attributes that are necessary to the identity and function of physical and extra-physical forces within an interactional, symmetry-based setting.

The present paper is concerned with the realistic view, which explains the existence of musical forces in terms of gauge transformations. The gauge transformation used in a first step, is based on vector rotations in a two-dimensional (real) Hilbert space (founded on SO(2) symmetry). We call the corresponding model the *spatial deformation model*. It follows ideas that are borrowed from the general theory of relativity. Explicit applications to attraction phenomena were made recently (Blutner, 2016; beim Graben & Blutner, 2017). The gauge transformation used in the second step is based on local phase invariances in quantum theory (founded on U(1)). We call the corresponding model the *phase model*. Both gauge models use subgroups of the SU(2) symmetry group by transforming spinors. The *combined model* makes full use of this group.

What we aim to demonstrate empirically is that the spatial deformation model gives a proper description of static tonal attraction (tonal hierarchies as described in generative music theory, e.g. Lehrdahl 1988). In contrast, the phase model alone does not give much empirical support. However, in tandem with the deformation model it gives a fair description of dynamic attraction data, as investigated by Woolhouse (2009).

The structure of the paper is as follows. The subsequent Section 2 explains the metaphoric conception of musical forces and contrasts it with a realist conception. Section 3 will introduce the qubit model of tonal attraction. This model can be seen as a special case of the two realist models: it is explicating the force-free (neutral) case. Section 4 explains the general idea of the structural realist view and develops two local gauge theories, which subsequently are applied in music cognition. Both static and dynamic attraction phenomena are discussed. Further, we explain how the hierarchic model of tonal attraction (Lerdahl 1988, 2001) can be remodelled within gauge theory and how we can extend this model in order to grasp certain asymmetries (major/minor modes). Section 5, finally, draws some general conclusions and rises several issues for future research.

# 2 Metaphoric and realist conceptions of musical forces

In classical physics, a force is seen as the cause of any change of the motion of an object. A force has a magnitude and direction making it a vector. According to Newton's second law the force acting upon an object is equal to the rate at which its momentum (= mass times velocity of the object) changes with time. Notably, our intuitive understanding of physical forces is not exactly the same as Newton's physical understanding. This is especially visible in connection with Newton's first law. It states that physical objects continue to move in a state of constant velocity unless acted upon by an external force. This conflicts with our everyday experience assuming that objects move with constant velocity only when a constant force is applied (due to the hidden role of friction or turbulences). Aristotle, to be sure, was much closer to folk physics than Galilei, who was the first devising experiments to disprove Aristotle's theory of movement.

Within the last 100 years, the distance between theoretical physics and folk physics has increased even more. In modern particle physics, forces and the acceleration of particles are explained as a mathematical by-product of the exchange of momentum-carrying tiny particles (so-called gauge bosons). With the development of quantum field theory and general relativity, it was realized that force is a redundant concept arising from conservation of momentum (4-momentum in relativity and momentum of virtual particles in quantum electrodynamics). The conservation of momentum can be directly derived from the homogeneity or symmetry of space and so is usually considered more fundamental than the concept of a force. Hence, the modern understanding of physical forces sharply contrasts with our folk physical understanding, which is sometimes taken as a sign of progress in science (Weinberg, 1992).

Metaphors involving motion and forces are based on our folk physics and not on the modern understanding of physical forces. The former but not the latter are a source of inspiration for understanding tonal music and tonal harmonies since ancient times. The application of physical metaphors is quite common in theories of music. Physical forces are represented in our naïve (common sense) physics or folk physics. Our experience of *musical motion* is conceptualized in terms of our experience of *physical motion* and their underlying *forces*. For example, Schönberg speaks of different forces when he explains the direction of musical forces in cadences where the tonic attracts the dominant (Schönberg, 1911/1978, p. 58).

Whereas the metaphoric conception is based on analogical reasoning, the realistic conception assumes a principle-based structural mechanism and deductive reasoning. With the term "realism" we refer to a "group structural realism" (Dawid, 2017). The idea is to distinguish between token-based and type-based realism. In token-based realism (ontological realism), all entities that are postulated within the theory have a direct pendant in reality. In type-based realism, the theory as a whole is tested with reality. It is not required that all postulated entities acquire physical meaning. Perhaps, Johannes Kepler can be seen as a forerunner of the type based realism and the formulation of laws that can be tested only as a whole. Kepler also came with the idea – deeply rooted in his believe in God – that the harmonies of the world and the harmony of music could be described in a uniform way (Kepler, 1619). To be fair, we should add that some authors prefer to interpret Kepler's *Harmonices Mundi* (Kepler, 1619) and his "music of the spheres" in a different way and see Kepler as a forerunner of the metaphoric view (Hubbard, 2017).

In modern physics, gauge theories have provided our best representations of the fundamental forces of nature, including electromagnetic forces, strong and weak nuclear forces. Even when this approach seems not to be reasonable and satisfying in all respects, it is a matter of fact that since more than 50 years local gauge symmetries play an essential role in constructing the most powerful and successful physical theories. A somewhat different mechanism is applied in Einstein's general theory of relativity for describing gravity. In this theory, the manifold space itself (time and space coordinates) is structured by a non-Euclidean metric. A basic assumption is that all coordinate systems (including those that are rotating and accelerated) are equivalent. The force of gravitation relates to a deformation of the manifold space.[1]

## 2.1 Larson's metaphoric model of tonal forces

Several authors explicitly or implicitly use the ideas of musical movements and musical forces as based on conceptual metaphors in the sense of Lakoff and Johnson (1980). That means the source domain of naïve (folk) physics is assumed to constitute a conceptual network establishing main propositions about physical movements and their causes – the physical forces. Analogical reasoning is used then to transfer the physical concepts to the

[1] Several attempts were made to provide a gauge theoretic treatment for gravity. Famously, Weyl failed with his first attempt made in 1918 (Weyl, 1950). Later, others were more successful basing the gauge theory on the Lorentz group (1956) or the Poincare´ group (1961).

goal domain of tonal music. In this way, it is possible to describe the most plausible expectations a listener generates during the processing of tonal music. This includes expectations based on static and dynamic forces.

Larson (1997-98, 2004; 2012) is the most prominent author that develops this idea in detail. In particular, he proposed three musical forces that generate melodic completions. These forces are called 'gravity', 'inertia', and 'magnetism', respectively. These forces relate to conceptual metaphors (Lakoff and Johnson 1980) and structure our musical thinking per analogy with falling, inert and attracting physical bodies. Hence, physical forces are represented in our naïve (common sense) physics or folk physics.

Larson (2012) gives some examples that concern ordinary discourses about music. They demonstrate the metaphorical potential of the three forces ('gravity', 'inertia', and 'magnetism'). GRAVITY: The soprano's *high* notes rang *above.* The rising melodic line *climbed higher*. MAGNETISM: The music is *drawn* to this stable note. The *leading tone* is *pulled* to the tonic. INERTIA: The accompanimental figure, once set in motion… . This dance rhythm generates such momentum that… (citations at the end of Sect. 8).

In this subsection, we will present the basic ideas of Steve Larson as published in his last book (Larson 2012). We think that this book gives the best overview on the field of musical forces presently available. And it provides a fair discussion on related proposals such as Narmour's (1992) implication-realization model, the model of Bharucha (1996), Lerdahls (2001) algorithm, and related ideas of Margulis (2003) and others.

Larson (2012) investigates the empirical hypothesis that the average rating of each of the investigated patterns is a function of the sum of musical forces acting on that pattern. To do so, a linear regression analysis is performed testing the following hypothesis for the "net force" *F* for a probe tone *x* as reflected by the ratings:

(1) $F(x) = w_G \cdot G(x) + w_M \cdot M(x) + w_I \cdot I(x)$

Hereby, $w_G$, $w_M$, and $w_I$ are the corresponding weight factors of the three constraint functions. The constraint functions themselves reflect the intuitive content of the phenomenological forces. For instance, the constraint $G(x)$ for gravity gets the value 1 (0) if the probe tone x is lower (higher) than the preceding tone. Hence, the constraint for gravity prefers falling tones to rising ones. The results of the linear regression analysis for the investigated data (Larson & van Handel, 2005) are $w_G = 0.4$, $w_M = 0.1$, $w_I = 1.2$. The correlation between model and data is $r = 0.95$. This high $r$-value means that the three forces, taken together, can account for about 90% of the variance of the frequency data. The two weight factors for gravity and magnetism are each significantly different from zero (at a 0.1 % level), but the weight for inertia is not. Interestingly, other studies using other data sets (Larson, 2002) give a different result: gravity and inertia both make significant contributions but magnetism does not. In the 2005 study an additional analysis was performed that included in addition to GRAVITY, MAGNETISM, and INERTIA, an extra factor signaling *the ending on tonic* (= $\hat{1}$) was introduced. In this case the correlation is still a bit higher: $r = 0.977$, and the extra factor got a weight of 0.46. Interestingly, the other factors now get weights different from the former analysis: $w_G = 0.16$ (instead of 0.4), $w_M = 0.26$ (instead of 0.1), and $w_I = 1.2$ (as before). Hence, magnetism and inertia both make significant contributions to linear regression but gravity does not. This demonstrated that the contribution of single factors to the "net force" can be evaluated only when the full context of all involved factors is given.

Finally, we want to stress that even a high correlation value of the fit as found in the data analysis just described does not answer the fundamental question about constraint grounding. As we have seen, the addition of some extra factors can radically change the influence of other factors and can even marginalize some factors. Hence, a multiple regression analysis with a high overall correlation coefficient cannot be taken as argument that the involved factors are all substantiated and "symbolically grounded" in the sense of Harnad (1990). As such, we cannot expect that these factors play a causal role in explaining tonal attraction.

We think Larson (2012) was aware of these problems. Several of his careful analyses try to justify the special role of musical forces. This contrast with alternative analyses by earlier authors. Further, Larson (2012) has investigated different variants of various factors (constraints) and he found how sensitive the cognitive system reacts even on minimal variations.

### 2.2 The realistic conception of tonal forces

What we will call "realistic conception" in this article is due to ideas borrowed from theoretical physics. According to Penrose (2004), all physical interactions are governed by "gauge connections" which depend crucially on spaces having exact symmetries (p. 289). It are these gauge connections which we will take as the basic of the present realistic conception of tonal forces. From the perspective of quantum physics, the idea of gauge symmetry has been applied by pioneers such as Schrödinger, Klein, Fock and others (for an overview, see Jackson & Okun, 2001).[2] It is suitable to introduce the realist force conception by means of a simplified mechanical picture (following Harlander, 2013).

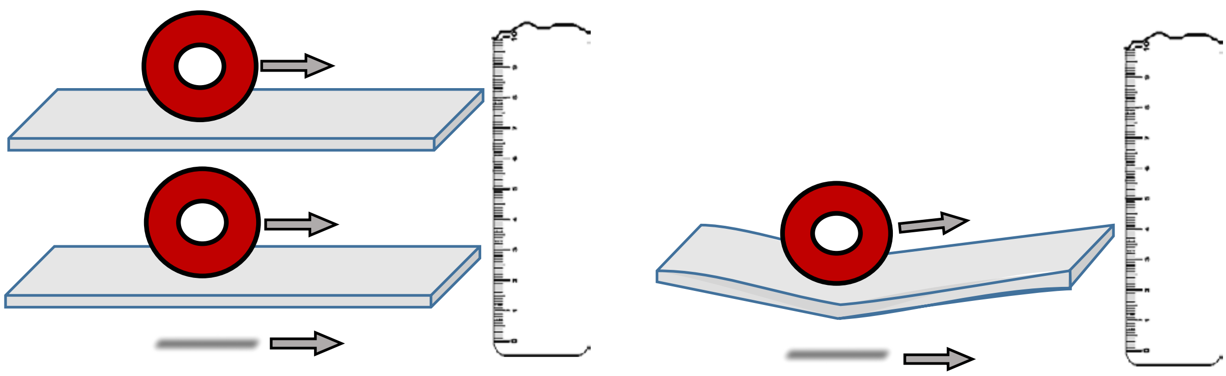


Figure 1: Left: Global Gauge. Right: Local gauge

Figure 1 gives a mechanical example of a so-called gauge symmetry provided by a tire rolling on a pane of glass. The shining sun is producing a moving shadow, which is the essential thing we can observe (similarly to Plato's allegory of the cave). For the movement of the shadow the absolute altitude of the pane is not relevant, only the velocity of the rolling tire

[2] As an example, the description of electrons as formulated by the Dirac equation can be considered. In this case, the multiplication of the wave function with a local phase factor $e^{i\varphi(x,t)}$ introduces an additional term in the transformed Dirac equations which destroys the symmetry. The crucial idea is to compensate the destroying term by an additional term modifying the original electromagnetic potential. This term is seen as describing an interaction of the original electromagnetic field with a gauge field. Obviously, this idea realizes a new dynamical principle coupling the gauge field with the electromagnetic field of the electron. There is a natural interpretation of the gauge field: it describes the interaction of a *photon* with the electron. In other words, the exchange of a photon is realizing a new force found by the idea of a gauge transformation. A more complex case is the standard model of particle physics. The model is formulated as a non-Abelian gauge theory with the symmetry group $U(1)\times SU(2)\times SU(3)$. It has twelve gauge bosons: the photon, three weak bosons and eight gluons. Between quantum electrodynamics and the full complexity of particle physics, there are symmetry groups such as $SU(2)$ which correspond to the Schrödinger-Pauli equation and $U(1)\times SU(2)$ for the Schrödinger-Pauli equation including a Higgs field to give spin-1/2 dyons their masses.

is. The fact that the whole scenery of the rolling tire can be moved vertically without changing the movement of the shadow corresponds to a *global symmetry*.

Now assume that there is a deformation of the pane resulting in a local change of the altitude of the tire. The variation of attitude is producing a breaking of the global symmetry. The dynamic effect of the symmetry breaking is that the velocity of the tire is changing by means of the deformation. The shadow at the bottom reflects this behaviour.

The request for *local symmetry* is now simple to understand. It refers to the demand that the movement of the shadow does not give any indication for the deformation of the pane of glass. Obviously, this can happen if we slow down or accelerate the tire dependent on the local deformation. In other words, the request of local symmetry demands us to introduce a varying force.[3]

Generally, the idea of founding forces by symmetries is as follows. Assume a physical system is invariant with respect to some global group of continuous transformations (for instance, independence of space and time). Then the idea of gauge invariance, is to make the stronger assumption that the basic physical equations describing the system have to be invariant when the group operations are considered locally (i.e., dependent on time and the other coordinates of the system). Normally, this principle of gauge invariance, leads to a modification of the original equation and introduces additional terms which can be interpreted as new "forces" induced by the "gauge field", which describes these local dependencies.[4]

## 2.3 The hierarchic model of tonal attraction

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

A: octave space (defined by the root tone, 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: chromatic space (including all twelve pitch classes).

Table 1 also shows the *tonal attraction* or *anchoring strength s*. This measure simply counts the number of degrees that are commonly shared across levels A to D (omitting level E that is common for all tones).

[3] Of course, pictures such as Figure 1 should be used with great caution. Moses forbid the Israelites to make any image of God. Similarly, in several respects, Dirac remarked that we should not try to make visualizations of quantum theory.

[4] As mentioned already, the idea of gauge invariance was first developed by Hermann Weyl in 1918, when he made the attempt to unify gravity and electromagnetism. Weyl assumed that the length of any single vector is arbitrary. Only the relative lengths of any two vectors and the angle between them are preserved under parallel transport. This was the birth of a new idea in physics which was called "gauge invariance" by Weyl. Even when Weyl's attempt to develop a unified theory failed, the idea survived and was extremely successful later on. It is this success that justifies the theory. The theory itself remains mysterious to a certain degree: we do not have an independent, physical or methodological motivation for it.

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A:octave | C | x | x | x | x | x | x | x | x | x | x | x |
| B: fifth | C | x | x | x | x | x | x | G | x | x | x | x |
| C: triadic | C | x | x | x | E | x | x | G | x | x | x | x |
| D: diatonic | C | x | D | x | E | F | x | G | x | A | x | B |
| E: chromatic | C | D♭ | D | E♭ | E | F | G♭ | G | A♭ | A | B♭ | B |
| Anchoring strength $s$ | 0 | 4 | 3 | 4 | 2 | 3 | 4 | 1 | 4 | 3 | 4 | 3 |

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

Temperley (2008) proposes the following formula to calculate the attraction probability:

(2) $p(j) = \frac{s(j)}{\sum_j s(j)}$

The predictions of the hierarchical model are in excellent agreement with the experimental data in the case of major keys. In case of minor keys (based on the natural or harmonic minor scale), however, there are significant deviations (Blutner, 2015; beim Graben & Blutner, 2019).

The basic tonal pitch space is easy to model within the framework of optimality theory (Prince & Smolensky, 1993/2004; Smolensky & 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 octave level) is satisfied if the considered tone is the root tone and it is violated otherwise. In Table 1, a constraint violation is marked by "x". Obviously, the number of violations agrees with the anchoring strength if all constraints are considered equally ranked.

Regarding the *function* of the tonic hierarchy in tonal music, we refer to the insights of Philip Ball, which crucially addresses the tonal dynamics:

> Although it is normally applied only to Western music, the word 'tonal' is appropriate for any music that recognizes a hierarchy that privileges notes to different degrees. That's true of the music of most cultures. In Indian music, the *Sa* note of a *that* scale functions as a tonic. It's not really known whether the modes of ancient Greece were really scales with a tonic centre, but it seems likely that each mode had at least a 'special' note the *mese*, that, by occurring most often in melodies, functioned perceptually as a tonic. This differentiation of notes is a cognitive crutch: it helps us interpret and remember a tune. The notes higher in a hierarchy offer landmarks that anchor the melody, so that we don't just hear it as a string of so many equivalent notes. Music theorists say that notes higher in this hierarchy are more *stable*, by which they mean that they seem less likely to move off somewhere else. Because it is the most stable of all, the tonic is where melodies come to rest. (Ball 2010: 95)

As we have seen, the probe tone techniques used in the experiments by Krumhansl, Kessler and others ask listeners directly to judge how well a single probe tone or chord fits an established context, and the relevant data collected by this technique represent the *static site of tonal attraction*. However, the finding that some tones are more stable than others invites some speculation about the *dynamics of attraction*. When considering sequences of pitches, "a melody is then like a stream of water that seeks the low ground" (Ball 2010: 95). The dynamic forces stipulated by Ball (2010) are forces that are directed towards the chromatically closest tones that are higher in the *static* attraction hierarchy than the trigger. In a related article (Blutner, 2019) this idea is discussed in more detail. For the present purposes, it is enough to accept that dynamic attraction phenomena cannot be understood without their static counterparts.

Importantly, at the present state of discussion we have neither a precise mathematical conception of stability nor a *realistic* idea of musical forces. The *metaphoric* view of musical forces suggest to take certain pattern of tonal continuation as expression of musical forces. These forces are assumed to determine the dynamics of tonal sequences.[5] This view is based on rather simple-minded ideas about the phenomenology of musical forces. In Sect. 2.1, we have outlined that the underlying multi-regression analysis cannot be taken as a causal explanation of the essential mechanism of tonal dynamics. Rather, it describes correlation without explaining the mechanism that causes these correlations.

In Section 4, we present a causal mechanism for explaining the phenomena of static and dynamic attraction. It is based on a realistic view of musical forces and it exploits two basic views how such forces can arise: (i) forces may emerge by a local gauge deforming the manifold space (deformation model); (ii) forces may emerge by a local gauge transforming the phases of the wave function (phase model). The deformation model accounts for the static attraction data. For explaining the dynamic attraction data, we will apply deformation model and phase model in tandem. This ensures that the dynamic attraction phenomena are based on a static attraction mechanism that is lifted to a dynamic level (by applying the phase model in the second step).

Remarkably, the realistic view of musical forces does not only give a precise definition of musical forces. It also provides a precise notion of stability, based on the ingenious work of the Russian mathematician Alexander Michailowitsch Lyapunov (Lyapunov, 1966). Before we introduce the realistic models, it is opportune to explain how basic ideas of quantum cognition can be applied to computational music theory.

# 3 The qubit model of tonal attraction

For the following, we make use of the notion of a *tonal pitch system*. 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 octave-equivalence resulting in twelve pitch classes, also called tones. Further, we assume a tuning system based on an equal temperament, i.e. a tuning system in which the fundamental frequencies between adjacent notes have the same ratio. The following numeric notation is used for defining the twelve tones of the system ("scale degrees" *j*, with *j* running from 0 to 11), in ascending order:

(3) 0 = C, 1 = D♭, 2 = D, 3 = E♭, 4 = E, 5 = F, 6 = G♭, 7 = G, 8 = A♭, 9 = A, 10 = B♭, 11 = B

For applying basic ideas of group theory it is essential that there are certain operations that allow transforming tones into other tones. For instance, we can increase the tones by a certain number of steps (0, 1, 2, …, 11). Such operations are called *transpositions*. The 1-step transposition transforms C into D♭, D♭ into D, and so on. Operations can be combined. For example, we can combine the transposition of a 2-step increase with a 3-step transposition, resulting in a 5-step transposition (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, …, 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 operations of transpositions can be described by addition (modulo 12): $x + y \bmod 12$; e.g., $2+3 \bmod 12 = 5$, $7+6 \bmod 12 = 1$. For a concise introduction of basic concepts of the mathematical theory of groups, the reader is referred to standard text books (e.g., Alexandroff, 2012).

[5] See Sect. 2.1, especially Formula (1). The regression analysis performed by (1) conforms to the dynamics of tonal music. This contrasts with the regression analysis based on the constraints A-D of the tonal pitch space, which clearly corresponds to the phenomenon of static attraction.

In the case of music based on twelve 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 $x^{-1} = (12 - x) \bmod 12$. The group consisting of the 12 tones is a cyclic group, which is called $\mathbb{Z}_{12}$. Note that 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.

In the present numerical representation of the cyclic group $\mathbb{Z}_{12}$ we have four generators conforming to the numbers 1, 11, 7, 5. 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. Figure 2 gives a visual representation of the group $\mathbb{Z}_{12}$ using the two basically different generators 1 or 11 (left hand side, 7 or 5 (right hand side).

Next, we have to look for a simple geometric representation of this symmetry group. This group could consist of linear maps as studied in linear algebra. More specifically, the group could consist of certain rotations of vectors in a two-dimensional vector space. For instance we can rotate the vector $\psi_{\rightarrow} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in 12 steps to the original vector. In linear algebra, the elementary rotation steps can be described by a rotation matrix that is rotating the state vectors by an angle of $\pi/12$.[6] In the Bloch sphere the rotation angle has to be doubled, i.e. $\pi/6$ for one rotation step. This is represented on the right hand side of Figure 2 using the generator of the circle of fifth. In contrast, the left hand side shows the 12 tones arranged in a chromatic way. In both parts of Fig. 2, the tones of the diatonic (C major) scale are shown by white circles and the other tones (called the non-diatonic ones) are represented by black circles. Obviously, the 7 diatonic tones as well as the 5 non-diatonic ones are connected (= convex) areas when the *circle of fifth* is used but they are not connected when the *chromatic ordering* is applied. It is this fact that favours the circle of fifth representation over the chromatic representation. The former can be seen cognitively more realistic than the latter (following Gärdenfors' (2000) methodology of conceptual spaces).

[6] Obviously, an explicit representation of the vectors representing the 12 tones in the 2-dimensional Hilbert space is as follows: $\psi_j = \begin{pmatrix} \cos(j\pi/12) \\ \sin(j\pi/12) \end{pmatrix}$ with j = 0, …, 11. For j = 0, we get the tonic vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and for j = 6 we get the orthogonal vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ representing the triton.

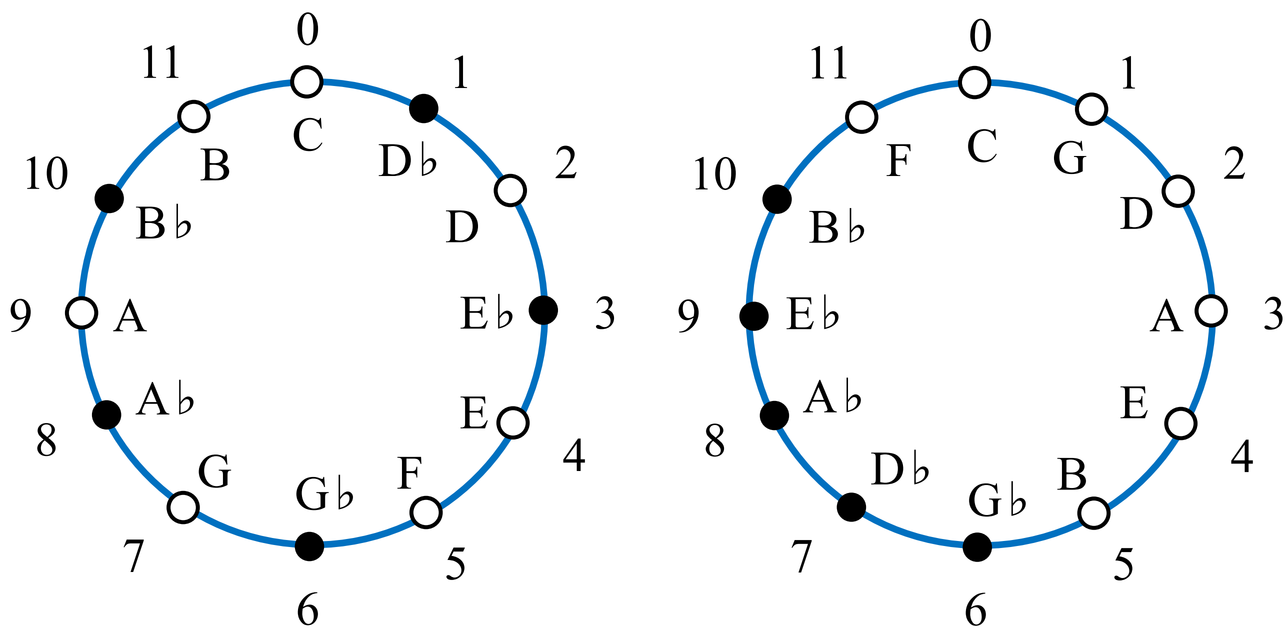



Figure 2: Visual representation of $\mathbb{Z}_{12}$. On the left hand side the different elements of the group are generated by the semi-tone generator. The white dots give an ordered subset of $\mathbb{Z}_{12}$ starting with the tone 0. It is the diatonic scale of C-major. On the right hand side, the 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.

There is a straightforward argument for the uniform distribution of the 12 tones in the Bloch circle. It is due to a fundamental symmetry principle. 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 transposition 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 assuming equal temperament.

In the following subsection, we will explain how this vector representation of the twelve tones makes it possible to derive precise attraction values (in terms of quantum probabilities).

### 3.1 The neutral case

In the case of pure states, quantum theory defines *structural* probabilities (cf. Blutner 2015). This means the probability that a state $\psi$ 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 of a given tonal context? What is the probability that it collapses into the (tonic) comparison state? The probability $P_l(j)$ of a collapse of the state $\psi_j$ into a state $\psi_l$ (rather than in an orthogonal state) can be calculated straightforwardly. It is the square of the length of the projection of state $\psi_j$ onto state $\psi_l$:

(4) $P_l(j) = \cos^2(\pi(j-l)/12) = ½(1+\cos(\pi(j-l)/6))$, where $0 \le j, l \le 11$.

For a fixed element $\psi_l$ equation (4) calculates a measure of how well each of the twelve target tones indexed by $j$ ($0 \leq k \leq 11$) fits to the contextually given comparison tone. Hence, formula (4) offers the attraction profile relative to a contextually given cue tone $\psi_l$. In the following we will set $l = 0$. This allows a simple calculation of the quantum-probabilistic profile assuming a variable $k$ referring to intervals instead of a single tones (the intervals spanned by the contextually cue tone and the target tones):

(5) $P(k) = \cos^2(\pi\, k/12) = ½(1 + \cos(\pi\, k/6))$, where $0 \leq k \leq 11$.

We can compare it with the attraction profile resulting from interval cycles (Woolhouse, 2009, 2010; Woolhouse & Cross, 2010), as presented in Fig. 4.

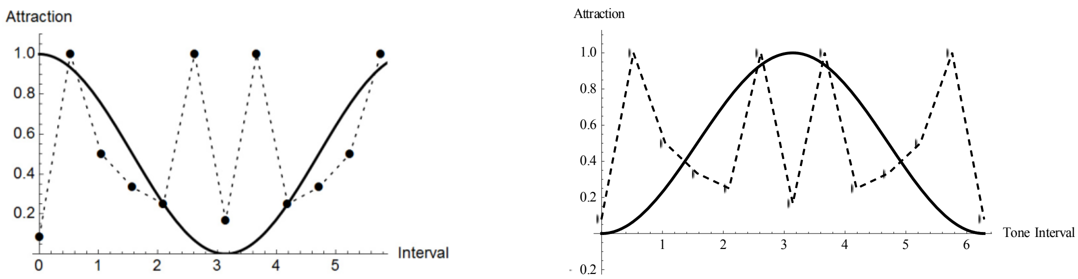



Figure 3: Comparison between the profiles (kernel functions) resulting from interval cycles (dashed) and the profile resulting from a simple quantum model (neutral kernel function; bold curve). Note that the endpoints corresponds to the tonic tone ($0 \cong 12$). Due to the chosen normalization, the maximum of the profiles is 1 in each case. On the right hand side the comparison with the phase-shifted curve is shown (phase shift by $\pi$).

The figure illustrates that the kernel resulting from interval cycles[7] and the kernel resulting from the simple quantum model are very different. In both cases shown in Figure 3, the correlation between the ICP is close to zero (- 0.1 for the kernel shown on the left hand side and + 0.1 for the kernel shown on the right hand side). This rises a series empirical problem since the ICP model gives a good fit for the dynamic attraction data (Woolhouse, 2009).

### 3.2 The role of phase parameters

When the simple Bloch circle (*real* 2-dimensional Hilbert space) is replaced by the full Bloch-sphere (*complex* 2-dimensional Hilbert space), the phase parameters come into play as shown in the following formula:

(6) $P(k) = \frac{1}{2}(1 + \cos(\Delta_k)\cos(\pi\,(k-3)/6)$.[8]

[7] 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 twelve-tone equal temperament, the *interval-cycle proximity* (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). For example, the *ICP* for the triton is 2 and the *ICP* for the fifth is 12.

[8] Taking a phase parameter into account, the 12 tones can be represented by the following vectors in the full qubit space: $\psi_j = \begin{pmatrix} \cos(j\pi/12) \\ \sin(j\pi/12)\, e^{i\,\Delta_j} \end{pmatrix}$ with j = 0, …, 11. In case we chose j=0 for the tonic, the calculated projections are identical with the neutral case and do not depend on the phase parameters of the tones. The same holds for the calculated probabilities. Therefore, we take the vector $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ for representing the tonic. Then it is easy to see that formula (6) results for the probability distribution (attraction potential).

For k= 3, the probability gets its maximum ($P(3) = 1$) if we assume a zero phase shift. In an earlier study (Blutner, 2015) all the phase parameters were fitted by the data of Krumhansl and Kessler (1982). However, it can be criticised that this procedure is not very explanatory because of the big number of parameters that have to be fitted. Further, the status of these parameters as entities that have to be learned is questionable. A more systematic solution is provided in Sect. 4.3 and 4.4 where a gauge-theoretic variant of this model is developed, the *phase model*.

### 3.3 Attraction profile = kernel + linear convolution

Formula (5) for the neutral case or formula (6) for the complex case with phases are not enough to represent the psychological attraction data as measured by Krumhansl and Kessler (1982) and by others. The point is that these formulas refer to a single contextual element only. In other words, these formulas represent kernel functions only. These kernel functions enter a procedure that takes several contextual elements into account (for example, the three tones of a chord or the seven tones of a standard cadence / diatonic scale). The idea we apply here is to handle this procedure by a linear convolution process as represented by the following formula:

(7) $P(j) = \sum_l kernel(j - l) \cdot P_{context}(l)$

One can regard this formula as an equation that describes the modification of a kernel function by a distribution of several contextual elements $l \epsilon\ context$. For example, the kernel function can be the ICP kernel or the function $P(k)$ taken from expression (5) – describing the neutral case of the qubit model. Since the kernel function in (7) depends on the difference $j - l$ only, it automatically satisfies the requirement of transposition invariance.

An important theoretical question is where the distributions $P_{context}(l)$ comes from. The general answer is that it comes from a probabilistic, Bayesian induction, which is based on the frequency of tones realized in a given piece of music. In the simplest case, it is appropriate to take the three tones of a tonic chord – assuming that these three tones realize the context, each of them having probability 1/3. Another possibility is to induce the underlying diatonic scale and to assume that the seven elements of this scale have equal probabilities.

It should be mentioned that Matthew Woolhouse and colleagues were possibly the first who used this methodology for describing attraction potentials (Woolhouse, 2009, 2010; Woolhouse & Cross, 2010) (see also Blutner, 2015). Interestingly, a similar method has been used in computational linguistics for modelling adjectival modification (de Groot, 2013).

## 4 The realist view of musical forces and three gauge models of tonal attraction

So far, we have considered tones as isolated entities, which are represented by vectors of a two-dimensional Hilbert space. From the point of view of information processing in the cochlea and the anatomy of the auditory brain this is not a very plausible assumption. Already the idea of frequency separation on the basilar membrane suggest a field model of tonal perception, which is closely related to the "place theory" of acoustic processing. The basic idea is the existence of a (one-dimensional) manifold space and the assumption that the different tones relate to discrete parts of this space.

We know that all changes of the quantum state by a unitary transformation leave all observable physical effects unchanged. A unitary transformation does not change the scalar product $\psi_1 \cdot \psi_2$ of two vectors of a Hilbert space. The following matrix describes the general form of a unitary transformation in a two-dimensional spinor-space (conforming to the group SU(2)):

(8) $E(\theta, \delta, \tau) = \begin{pmatrix} \cos\theta e^{-i\delta} & -\sin\theta e^{i\tau} \\ \sin\theta e^{-i\tau} & \cos\theta e^{i\delta} \end{pmatrix}$

Hereby, $\theta, \delta, \tau$ are real-valued parameters that determine the transformation. In a field model, these parameters can be independent of the variables that determine the manifold space (i.e. $x$ in the present case) or they can be dependent of these variables (written $\theta(x),\ \delta(x),\ \tau(x)$. In the first case, the transformation is as follows:

(9) $\psi(x) \longrightarrow \tilde{\psi}(x) = [E(\theta, \delta, \tau)]\psi(x)$

It is called *global gauge transformation*. In the second case, the transformation is dependent of the manifold space:

(10) $\psi(x) \longrightarrow \tilde{\psi}(x) = [E(\theta(x), \delta(x), \tau(x))]\psi(x)$

It is called *local gauge transformation*. The invariance of certain equations under global transformations is called *global gauge symmetry* and the invariance under local transformat-ions is called *local gauge symmetry*. The idea behind local gauge transformation is the requirement that this transformation leads to a neutral (force-free) solution of the underlying dynamics. That means the gauge transformation eliminates the forces that control the initial wave function $\psi(x)$, just as in the mechanical example of Sect. 2.2.

Here are two important examples of gauge transformations. First, we consider real-valued vectors and two-dimensional rotation matrices with the rotation parameter $\theta$.

(11) $E(\theta, 0, 0) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

This relates to the symmetry group SO(2). In the following subsection we will see that a local variant of this gauge transformation provides the *deformation model* of tonal attraction – a model of *static attraction*. The advantage of a gauge model of tonal music is that it gives a mathematical derivation of particular musical forces.

Our second example of a gauge transformation is a shift of the phases of the wave function that is illustrated here:

(12) $E(0, \delta, 0) = \begin{pmatrix} e^{-i\delta} & 0 \\ 0 & e^{i\delta} \end{pmatrix}$

The parameter $\delta$ is a phase parameter that is assumed being independent of $x$ in the global variant. The local variant of this gauge transformation assumes local dependencies $\delta(x)$ and provides a model of tonal attraction, which is called the *phase model*. At the moment we let it open whether this model is intended to describe static or dynamic attraction.[9] Later we will see that the deformation model best applies to static attraction. A *combination* of deformation and phase model describes dynamic attraction phenomena – the combined model is the third gauge model that we will discuss.

There is another subgroup of SU(2) described by the matrix

---

[9] In an earlier approach, both the deformation model and the phase model were applied to both static and dynamic attraction phenomena (Blutner, 2019).

$$(13) \quad E(0,0,\tau) = \begin{pmatrix} 0 & e^{i\tau} \\ e^{-i\tau} & 0 \end{pmatrix}.$$

The transformation triggered by this matrix is similar to those described by the matrix in (12). The sum of the two matrices manipulating the phases is given in (14)

$$(14) \quad E(0,\delta,\tau) = \begin{pmatrix} e^{-i\delta} & e^{i\tau} \\ e^{-i\tau} & e^{i\delta} \end{pmatrix}.$$

The full SU(2) transformation matrix $E(\theta,\delta,\tau)$ as defined by formula (8) is the sum of the two matrices $E(\theta,0,0)$ and $E(0,\delta,\tau)$. Using our own terminology, we refer to the first matrix (11) and the corresponding gauge field by 'deformation field' and to the second matrix (14) by 'phase field'.[10]

Obviously, in our simple mechanical picture (see Sect. 2.2), the global altitude of the plate corresponds to the global rotation parameter (in the deformation model) or the global phase parameter (in the phase model). In contrast, the locally changing value of the rotation parameter or of the phase parameter is relevant for the demand of local gauge symmetries. This demand is the structural instrument that introduces physical (and musical) forces. The fundamental equations satisfy global gauge invariance. The request for local gauge invariance can be satisfied only by introducing additional terms into the basic dynamic equations, which correspond to physical forces. In other words, we have to consider local phase changes always in tandem with emerging forces in the dynamic equations.

There are three main aspects that distinguish the gauge theoretic approaches of tonal attraction from the qubit model of the previous section. First, the qubit model considers tonal states as isolated vectors of a two dimensional vector space (qubit states). It corresponds to the force-free case. In contrast, the gauge theoretic approach analyses tonal states as resulting from a Schrödinger wave function (with its temporal and spatial dimensions). In the simplest case, the wave function is a standing wave along a one-dimensional spatial continuum $0 \leq x \leq 2\pi$, and the different tones relate to different discrete parts of the manifold space.

Second, the precise shape of the standing wave is determined either by particular local phase shifts (phase model) or by rotations of the vectors dependent of the spatial component (spatial deformation model). In both cases, the twelve tones are described by oscillations of the spin wave at particular points on the spatial axis.

Third, the origin of tonal micro-forces differs for the two models. In case of the phase model it arises from local phase shifts that are described by a parametrized phase function $\delta(x)$. In case of the deformation model it arises from a local rotation function $\theta(x)$ that rotes the spin vectors locally dependent of the spatial $x$-values. This rotation transformation of the spatial component can be interpreted as a nonlinear transformation of the manifold space and leads to a force conception similar to the idea of the force of gravitation in Einstein's general theory of relativity.[11]

Before we come to the detailed treatment of the two gauge models, we should give a concise introduction into the dynamic aspects of quantum theory. First, we consider the stationary form of the Schrödinger equation for describing objects in (one-dimensional) space and time (Schrödinger, 1926):

---

[10] Using the terminology of weak interaction physic (see Appendix 5), we refer to the (12) and the corresponding gauge theory by 'neutral current'. The other matrices relate in an indirect and less obvious way to the two variants of the 'charged current'.

[11] It goes without debating that the gauge model is not a modelling of the travelling wave in the cochlea (Terhardt, 1972, 1998). Rather, it could be seen as a third generation neural network approach approaching brain waves in the auditory cortex (Coombes, beim Graben, Potthast, & Wright, 2014).

(15) $$-\frac{\partial^2 \psi(x)}{\partial x^2} = E\,\psi(x)$$

A solution of this equation under appropriate boundary conditions is the following standing wave:

(16) $$\psi(x) = \cos(\sqrt{E}\ \cdot x)$$ [12]

For $E = 1/4$ this function is $\cos(x/2)$. It relates to a standing wave along the $x$-axis with amplitude's maxima at $x = 0$ and $x = 2\pi$ and a zero amplitude at $x = \pi$. Using standard wisdom of quantum mechanics, the probability density for each point of the manifold space is defined by the following expression:

(17) $$|\psi(x)|^2 = \cos^2\left(\frac{x}{2}\right) = 1/2(1 + \cos(x)).$$

For describing tonal music with 12 tones we need a discretization of the manifold space which we can achieve through sampling $x_k = \frac{\pi k}{6}$, for $k \in \{0, 1, \ldots, 11\}$. The corresponding distribution is then as follows:

(18) $$p_k = \cos^2\left(\frac{\pi k}{12}\right) = 1/2(1 + cos\left(\frac{\pi k}{6}\right))$$

This formula exactly corresponds to the neutral case of the qubit model described by equation (5).

The Schrödinger approach can be straightforwardly extended to the two-dimensional case of objects with spin $\psi(x) = \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix}$ assuming that the magnetic vector potential does not couple the two spinor components. Hence, we consider the familiar Schrödinger equation for an object in a purely scalar potential, except that it operates on a two-component spinor.

A force-free solution of the Schrödinger equation in this case is $\psi(x) = \begin{pmatrix}\cos(x/2)\\ \sin(x/2)\end{pmatrix}$. For calculating the probability density in the spinor case, we have to assume a projection onto the tonics. If the tonics is the unit vector $\begin{pmatrix}1\\0\end{pmatrix}$, the resulting probability density is exactly as before – see Eq. (17). In case, we take another vector as tonics, let say $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$, the result is different:

(19) $$|\psi(x)|^2 = (\sin\left(\frac{x}{2}\right) + \cos\left(\frac{x}{2}\right))^2 = \frac{1}{2}(1 + \sin(x)) = \frac{1}{2}(1 + \cos(x - \frac{\pi}{2}))$$

The difference of $\frac{\pi}{2}$ in the last cosine term of Eq. (19) corresponds to a transposition by three half-tone steps. In the vector representation, the rotation of tonics $\begin{pmatrix}1\\0\end{pmatrix}$ into the new tonic likewise reflects the difference of three half-tone steps. Generalizing this idea, we can prove the principle of transposition invariance of this simple phase model. Things are changing if phase shifts are included into the model (Sect. 4.3).

## 4.1 The spatial deformation model

[12] The solution of the full wave with the stationary part is given by $e^{-i\omega t}\ \psi(x)$

In the simplest case, the spatial deformation model (beim Graben & Blutner, 2017) can be described by the following wave function:

(20) $\psi(x) = \cos(\gamma(x)/2)$ .

This wave function yields the following probability density:

(21) $|\psi(x)|^2 = \cos(\gamma(x)/2)^2 = \frac{1}{2}(1 + \cos(\gamma(x)))$

We interpret these probability densities as attraction potential for the twelve tones localized at $x_k = \frac{\pi k}{6}$, for $k \in \{0, 1, \ldots, 11\}$c

Next, we have to consider a particular choice of the function $\gamma(x)$, which we call *deformation function*:

(22) $\gamma(x) = \pi + 1/\pi^3\,(x - \pi)^4$

The deformation function $\gamma(x)$ increases with the fourth power of the difference between $x$ and the localization of the triton in the Bloch circle (at $\pi$). The particular form of the deformation function fits static attraction by assuming that the fixpoints of the gauge transformation are the tonic (corresponding to $x = 0$) and the triton (corresponding to $x = \pi$). These assumptions reflect two plausible outcomes of static attraction experiments: (a) minimum attraction for the triton, and (b) maximum attraction for the tonics. Figure 4 shows the corresponding kernel function sing the equations (21) and (22).

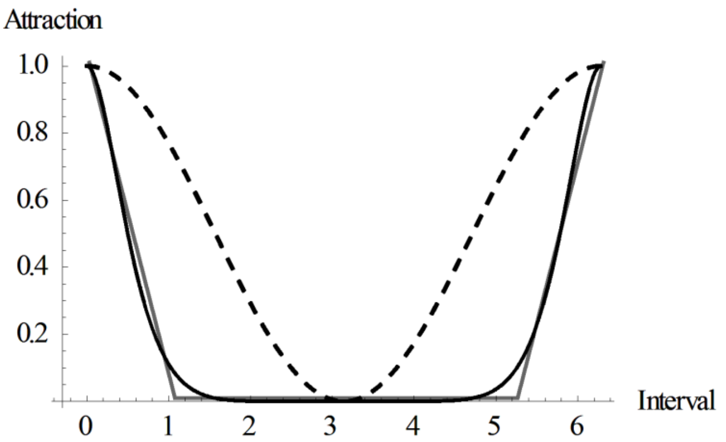



Figure 4: Static kernel (solid) and kernel resulting from the qubit model (dashed). The kernel function resulting from the hierarchic model is shown in grey. Note that the two endpoints corresponds to the tonic tone.

Earlier research (beim Graben & Blutner, 2019) has shown excellent agreement with the static attraction data of Krumhansl and Kessler (1982) using the operation of circular convolution (Sect. 3.3) for extending to musical contexts defined by triadic chords. No free parameter has been needed to fit the data (only the exponential in the nonlinear function $\gamma(x)$ can count as a parameter (the exponential 4 leads to fits that are a bit better than other choices such as 2 and 6).

At this point, it is useful to ask for the connection between the deformation model and the hierarchical model. As one sees from Fig. 5, the kernel function of static tonal attraction assigns the maximum value to the target tone (say C). The two neighbours on the circle of fifth (i.e., G and F) get an attraction value that is about half of it. The attraction values of all other tones is very low such that we can neglect them. Hence, when we construct the attraction profiles for a certain context given by a triad (say CEG), we get an approximate reconstruction of the hierarchic model. The three tones of the triad (CEG) get a very high value; C and G a bit higher than E because of the convolution operation. Next, the neighbours of the triadic tones (C: G, F; G: D, C; E: B, A) are all diatonic tones and get an attraction of about 50%. Hence, we can account for all levels of the hierarchic model shown in Tab. 1 besides the octave level (resulting in four different degrees of attraction).[13]

### 4.2 A gauge theoretic formulation of the spatial deformation model

In order to give a gauge-theoretic formulation of the spatial deformation model, we have to replace the scalar wave function given in (20) by a two-dimensional spinor function:

$$(23) \quad \begin{pmatrix} \psi_+(x) \\ \psi_-(x) \end{pmatrix} = \begin{pmatrix} \cos(\gamma(x)/2) \\ \sin(\gamma(x)/2) \end{pmatrix}$$

Now we represent the tonic by the Hilbert-space vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Then we straightforwardly obtain an expression for the probability density:

$$(24) \quad |\psi(x) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}|^2 = \cos^2\left(\frac{\gamma(x)}{2}\right) = \frac{1}{2}(1 + \cos(\gamma(x)))$$

This is exactly the result we have derived in Eq. (21). Next, let us consider the real-valued subpart of the SU(2) representation as given in (11). This is exactly the representation of two-dimensional rotations in the real Hilbert space, i.e. SO(2). With the help of this operator, we can rotate the free spinor $\begin{pmatrix} \tilde{\psi}_+(x) \\ \tilde{\psi}_-(x) \end{pmatrix} = \begin{pmatrix} \cos(x/2) \\ \sin(x/2) \end{pmatrix}$ into the spinor (23) deformed by gauge forces:

$$(25) \quad \left[E\left(\frac{1}{2}(\gamma(x) - x), 0, 0\right)\right] \begin{pmatrix} \cos(x/2) \\ \sin(x/2) \end{pmatrix} = \begin{pmatrix} \cos(\gamma(x)/2) \\ \sin(\gamma(x)/2) \end{pmatrix}$$

Intuitively, tones $x$ can be described by certain positions in the Bloch circle (angle x). The gauge transformation rotates the corresponding vectors by an angle $\gamma(x) - x$ in the Bloch

[13] At this point readers not educated with quantum mechanics may ask what is the advantage of such complex theories as quantum field theory against the much simpler and classical idea of the hierarchical model? The answer is that the quantum model allows to treat the attraction data with less stipulations than the hierarchical model where all levels of the hierarchy have to be stipulated. Further, the quantum model allows generalizations which are not obvious for the hierarchical model. For instance, the quantum model but not the hierarchical model relates static and dynamic attraction phenomena. Further, the quantum model but not the hierarchical model allows to approximate the differences between major and minor modes. The situation is similar to the situation in physic more than 100 years ago where the spectrum of the hydrogen atoms has been investigated and described by the simple Rydberg formula (Rydberg, 1890). The derivation of this formula by Bohr's atomic model and later by quantum theory counts as a milestone of physical progress, because of its explanatory value and numerous possible generalisation in the field of atomic spectra.

circle.[14] The addition theorems of cosine/sin-functions give the rotated vectors as exhibited in Eq. (25). Figure 5 illustrates the gauge transformation depictured in the circle of fifth. The twelve tones, which are uniformly distributed over the circle and localized at $x_k = \frac{\pi k}{6}$, for $k \in \{0, 1, \ldots, 11\}$, are mapped by the gauge transformation into a deformed distribution.

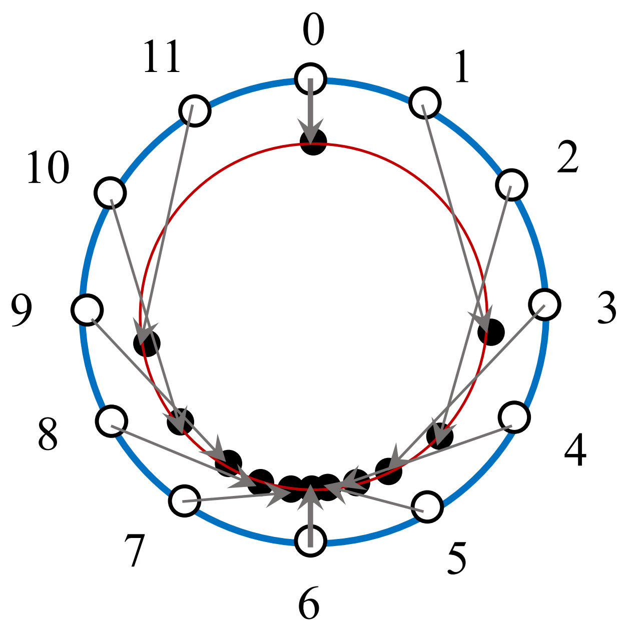



Figure 5: Illustration of SO(2) gauge transformation for the twelve tones (originally located at $x_k = \frac{\pi k}{6}$, for $k \in \{0, 1, \ldots, 11\}$ in the Bloch circle). The involved (static) gauge function is $\gamma(x) = \pi + 1/\pi^3\,(x-\pi)^4$.

Figure 5 illustrates that all tones with exception of the tonics itself and the triton are moved away from the tonics into the direction of the triton. Intuitively, it may be helpful to conceptualize this by 'gauge forces' that cause the deformation depictured in this figure.

Now consider a local gauge transformation, which converts the force-free solution of the Schrödinger equation into the deformed solution under the influence of gauge forces when we assume that $\gamma(x) = \theta(x) - x$:

(26) $$\begin{pmatrix}\tilde{\psi}_+(x)\\ \tilde{\psi}_-(x)\end{pmatrix} \to \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix} = \begin{pmatrix}\cos\theta(x) & -\sin\theta(x)\\ \sin\theta(x) & \cos\theta(X)\end{pmatrix}\begin{pmatrix}\tilde{\psi}_+(x)\\ \tilde{\psi}_-(x)\end{pmatrix} = \begin{pmatrix}\cos(\gamma(x)/2)\\ \sin(\gamma(x)/2)\end{pmatrix}.$$

Let us see the correspondence to the free Schrödinger equation (15) in case the wave function $\psi(x) = \tilde{\psi}(\gamma(x))$ is considered. As argued in detail in another paper (beim Graben & Blutner 2017, 2018), the function $\psi$ has to satisfy the following equation for both components of the spinor:

(27) $$-\frac{\partial^2\psi_i(x)}{\partial x^2} + \frac{\gamma''(x)}{\gamma'(x)}\frac{\partial\psi_i(x)}{\partial x} - \gamma'(x)^2\,\psi_i(x) = 0$$

[14] In order to avoid confusions: The variable $x$ measures the angles in the Bloch-circle. It runs from 0 to $2\pi$. The 'real' angles of the Hilbert-space vectors are half of it. Thus, in the 'real' rotation matrix (25) we have to divide the angles of the Bloch circle by the number 2 in order to get the angles in the rotation matrix.

This equation is derived by differentiating $\psi(x)$ twice and eliminating trigonometric terms. In order to get the standard stationary form of the Schrödinger equation, we decompose the Hamiltonian into three parts:

(28) $H = T + M + U$ with

a. $T = -\frac{\partial^2}{\partial x^2}$

b. $M = \frac{\gamma''(x)}{\gamma'(x)} \frac{\partial}{\partial x}$

c. $U = E - \gamma'(x)^2$ .

Then Eq. (27) takes the standard form of an eigenvalue problem:

(29) $[T + M + U]\, \psi_i(x) = E\, \psi_i(x)$.

It is obvious to call the operator $T$ the operator of kinetic energy and to call the operator $U$ the operator of potential energy. Depending of the special form of the gauge field $\theta(x)$ (or $\gamma(x)$) the kind of potential energy that is involved is alike a 'gravity' potential, a 'harmonic oscillator' potential or what else. The operator $M$ has been called magnetism in analogy to the physical examples. The details of gauge function $\gamma(x)$ specify the magnetism function.

So far, we can see it is mainly the sum of the three energies that makes sense for a musical interpretation. The decomposition into the three parts currently cannot be seen being related with any musical phenomenon. To be sure, this decomposition has absolutely nothing to do with forms of the metaphoric model that use different contributions of "folk-physical" forces in a linear regression analysis of tonal attraction (e.g. Larson 2012).

We also should stress the positive outcomes of the present gauge analysis. In contrast to the mentioned analyses by linear regression, the gauge-theoretic analysis is not ad hoc and does not require many arbitrary stipulations. The only assumption we have to make is to make a choice of the gauge field $\gamma(x)$. We have decided to approximate the gauge field by the function (22) in order to express the benefits of the hierarchical model. However, the present formulation is more than a sophisticated conversion of the hierarchic model into an oversized mathematical model. Conceptually, it is a direct and realistic introduction of musical forces based on the lead of mainstream physical ideas. Empirically, we will show how it helps to relate static and dynamic attraction models. Further, we will demonstrate that a modification of the model by exploiting the idea of symmetry breaking can help to solve the old problem of the harmonic differences between major and minor modes of tonal music.

We mentioned already that the sum of the three operators gives an energy density that is proportional to the probability density. A further plausible assumption is that the potential energy is an indicator of stability – cf. Graben & Blutner (2017) and the remark at the end of this subsection. In the following figure, we compare the overall energy density with the contribution resulting from the potential energy density (summing up $M$ and $U$).

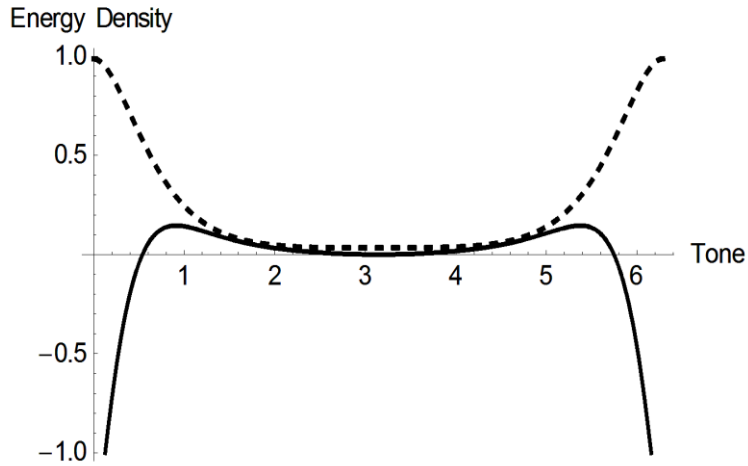



Figure 6: Energy densities of the deformation model. Dashed: total energy density; solid: density of potential energies.

As pointed out in more details in beim Graben & Blutner (2017), the overall energy density (proportional to the static attraction potential) has a maximum at the tonic tone (localized at 0 and 2 $\pi$ in Figure 6) and a minimum for the tritone localized at $\pi$). The sum of the potential energies becomes minimal toward the tonic regions. This explains the attracting force of the tonic for all other tones. Interestingly, there is a local minimum at the localization of the triton, making the triton to a kind of 'tonal trap' for tones of the region D, A …E♭, B♭ (corresponding to the interval from .8 to 5.4). Further, Figure 6 shows that there are two instable equilibria around x = .8 (D) and x= 5.4 (B♭).

At this point, we should explain the content of the mathematical conception of stability. In quantum theory, there is the distinction between *energetic stability of matter* and *dynamic stability of motion*. The former notion refers to the stability of atoms or macroscopic matter. Usually it is explained by the uncertainty relation (in a particular form) and by Pauli's exclusion principle (Lieb & Seiringer, 2010). The latter notion derives from ideas of Lyapunov and the existence of the so-called Lyapunov function the (local) minima of which describe to stable configuration of a dynamic system (Lyapunov, 1966). In the context of cognitive musicology, the stability of tonal movements is a point of interests – clearly referring to the concept of dynamic stability. Understandably, the potential energy of a system is such a Lyapunov function. And the gradients of the energy density correspond to the mathematical notion of forces, both in mechanics as in musicology. They indicate the direction of physical or tonal movements. Summarizing, the present analysis indicates that tonic appears as a center of force. In contrast, the tritone functions as a tonal trap attracting nine of the twelve tones with a moderate force.

### 4.3 Gauge models based on phase shifts

So far, we have modelled static attraction profiles by spatial deformations – an analysis that provides excellent empirical support. Hence, we can state that the deformation model is an excellent model for approaching the static attraction case. For the dynamic attraction profiles (chord resolution), in contrast, the situation is more difficult. In principle, we could follow three different approaches for finding a satisfying solution. First, we could use the phase model for approaching the dynamic attraction profiles. That means we could introduce proper gauge forces into the phase model – using a nonlinear gauge field. This variant was developed by Blutner (2019). Unfortunately, the model allows fitting the empirical data (Woolhouse, 2009) to a moderate degree only. Moreover, the model is not truly explanatory since it introduces at least two free parameters. Second, we could exclude the phase model and consider a different variants of the deformation model that allows describing the dynamic attraction data. This variant was developed in beim Graben & Blutner (2019). Unfortunately, this model is not very explanatory again since it requires fitting several parameters again. Third, we could consider a combination of the deformation model with the force-free phase model. This has the advantage to satisfy the idea of Ball (2010) that *static* forces allow to infer *dynamic* forces.

In the rest of this section, we will develop the third variant and we will explain the advantages of this approach. We will start with developing the phase model in the force-free case. For didactical reasons we start with a very naïve model considering the stationary Schrödinger equation (Subsection 4.3.1). Then we explain a spinor phase model introducing two-component wave functions based on the so-called Schrödinger-Pauli equation (Subsection 4.3.1). This approach allows a more satisfying treatment of the tonic context than the naïve approach. In Section 4.4, we use this phase model in tandem with the deformation model. This account gives a fair description of dynamic attraction data and it has a more explanatory power than the previous models for investigating the data of Woolhouse (2009). We demonstrate the explanatory value in connection with the breaking of mirror symmetry (Sections 4.5 and 4.6).

#### 4.3.1 Naïve Gauge models

The general matric form of a gauge transformation for spinors has been given in (8). A subspecies of this gauge was considered in the previous section, the SO(2) gauge. In this and the following Section we consider gauges isomorph to U(1) as specified by (12) for the spinor case. In the simpler case of scalar wave functions, this gauge relates to a single phase transformations. In the local case, we assume that all observable effects are invariant when *local* phase transformations are applied:

(30) $\psi(x) \longrightarrow \widetilde{\psi}(x) = \psi(x)\, e^{i\delta(x)}$

The idea now is to assume that local phase invariance is the gauge symmetry that we have to assume in quantum physics.

Introducing a local phase shift as defined by Eq. (30) and the idea of gauge invariance automatically converts the force-free solution of the Schrödinger equation into a solution under the influence of gauge forces. To see the idea, let us consider the following free wave function $\sin(x/2)$ which is gauged by the local phase transformation $e^{-i\delta(x)}$.

(31) $\psi(x) = \sin(x/2)\, e^{-i\delta(x)}$

A simple consequence of this choice is the following function of probability density:

(32) $\psi(x)^* \cdot \psi(x) = \sin(x/2)^2 = 1/2(1 - \cos(x))$

Obviously, the probability density does not depend on the gauge function $\delta(x)$. The form given in (32) contrasts with earlier probability density given in (17). It is based on the wave function $\cos(x/2)\, e^{-i\delta(x)}$, which is orthogonal to the wave function (31). We will see that the two wave functions based on $\cos(x/2)$ or $\sin(x/2)$ play different roles. The former approximates static attraction (with maximum probability for the tonic). The latter approximates dynamic attraction (with minimum probability for the tonic).

For performing a gauge analysis, let is differentiate the function $\psi(x)$ in (31) twice and eliminating trigonometric terms. Then the function $\psi(x)$ is a solution of the following differential equation:

(33) $$-\psi''(x) - 2i\,\delta'(x)\psi'(x) - i\,\delta''(x)\psi(x) + (E - \delta'(x)^2)\,\psi(x) = \mathrm{E}\,\psi(x).$$

Considering the stationary Schrödinger equation in the form (15) suggests a gauged Hamiltonian, which consists of a sum of three operators:

(34) $H = T + M + U$ with

a. $T = -\frac{\partial^2}{\partial x^2}$
b. $M = -2i \cdot \delta'(x)\frac{\partial}{\partial x}$
c. $U = E + \delta'(x)^2 - i\,\delta''(x)$.

As in the spatial deformation model discussed in the previous section, it is obvious that the operator $T$ is the operator of kinetic energy density (inertia) and the operator $U$ is the potential energy density. The operator $M$ ('magnetism') collects the remaining terms.

For specifying the gauge field, a simple suggestion is the following linear ansatz:

(35) $\delta(x) = x$ .

This means that the potential function (34)c is a constant function. This relates to a force-free case even when the magnetism component (34)b is considered.[15]

The sum $H$ of the three operators is Hermitian and has the expectation value $1/4\sin(x/2)^2$. Figure 7 shows the expectation values for the total energy (proportional with probability density) and for the potential energy taking the (orthogonal) stationary wave function $e^{-i\delta(x)}\cos(x/2)$ (left hand site) and $e^{-i\delta(x)}\sin(x/2)$ (right hand site).

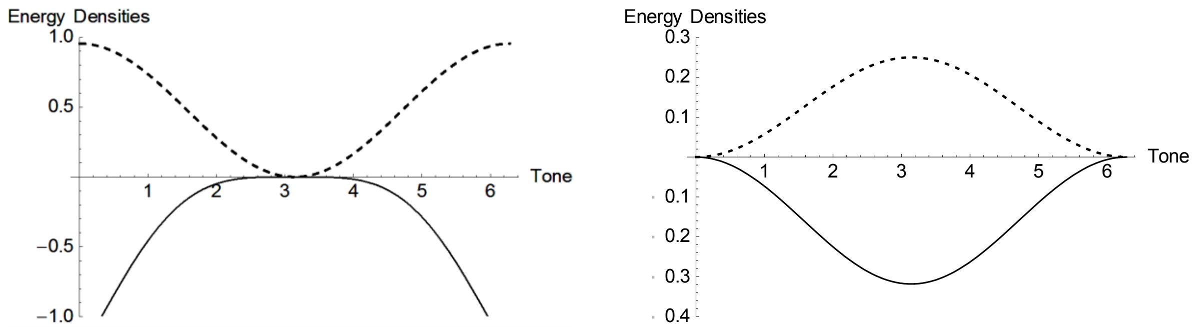



Figure 7: Left hand site: Phase model with stationary wave function $\psi(x) = \cos(\frac{x}{2})e^{-i\delta(x)}$. Right hand site: Phase model with stationary wave function $\psi(x) = \sin(\frac{x}{2})e^{-i\delta(x)}$. Dashed: total energy density; solid: density of potential energies (without the imaginary part of magnetism).

[15] Taking the wave function (31) for calculating the expectation values and (35) for the phase function, the contribution of magnetism gives two terms due to the application of the product rule for $\frac{\partial}{\partial x}\sin(x/2)\,e^{-i\delta(x)}$:
(i) a contribution that is the probability density multiplied with the factor $-2i \cdot (-i)\delta'(x)^2 = -2$;
(ii) a contribution that is purely imaginary. It compensates the imaginary part of the kinetic energy density.

What are the main properties of the calculated energy densities? First, it is obvious that for the wave function $\cos(\frac{x}{2})e^{-i\delta(x)}$ the potential energy density has the maximum for the triton and the minimum for the tonic tone. The converse holds for the total energy density, which is proportional to the probability density. Hence, the tonic tone is the stable endpoint. For the tonic, we find maximum probability density and minimum potential density. On the other hand, the triton is in an instable equilibrium state. This correspond to the case of static attraction.

Next, consider the wave function $\sin(\frac{x}{2})e^{-i\delta(x)}$. In this case, we find the situation depictured on the right hand side of Figure 7. In this case, the potential has its minimum for the triton and the maximum for the tonic. The converse holds for the total energy density (probability density), which has its maximum at the triton position and minimum for the tonic. We would like to interpret this as a first approximation to the dynamic attraction profile. However, this interpretation is rather unclear. Hence, it is fair to say that the discussed force-free phase model is not really able to approximate the dynamic attraction case.

Another important point is that introducing a local phase shift into the wave function as in (30) does not affect the resulting probability distribution. It still agrees with the results of the simple qubit model discussed before. However, a very different solution will result if we consider spinor wave functions instead of single wave functions as we did so far. The details we will consider in the next subsection. For the spinor's gauge model, which also will be based on the 'force-free' phase function (35), we will discuss total energy and potential energy in the following subsection.

#### 4.3.2 Spinor gauge models

In the case of two component spinors (applied for spin ½ particles in physics), we will look for the force-free wave equation. This is a special case of the so-called Schrödinger-Pauli equation, which sometimes has been seen as a special case of the Dirac equation in the non-relativistic limit. The following formula gives a (minimalist) solution of the Schrödinger-Pauli equation in the force-free case (cf. Blutner 2016):

(36) $$\psi(x) = \begin{pmatrix} cos(x/2) \\ sin(x/2) \end{pmatrix}$$

For a gauge-theoretic analysis, we have to consider the matrix form of the SU(2) group as given in (8). It provides a unitary transformation of spinors with the real parameters $\theta, \delta, \tau$.

If we want to simulate the properties of the tones along the circle of fifth, we have to find a proper path through this space of parameters. For the present phase model, we consider $\delta(x)$ as the only relevant parameter and chose $\theta(x) = 0, \tau(x) = 0$. Applying the corresponding unitary transformation $E(0, \delta(x), 0) = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\delta(x)} & 0 \\ 0 & e^{i\delta(x)} \end{pmatrix}$ to the spinor's wave function (36), we get the following transformed wave function (for details and motivation see Appendix 2):

(37) $$\psi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} cos(x/2)e^{-i\delta(x)} \\ sin(x/2)e^{i\delta(x)} \end{pmatrix}$$

In quantum field theory, the projections of the field vectors on certain states provides the corresponding probability densities. In the simplest case, we project onto the vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. This corresponds to the projection operator $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and we can calculate the probability density relative to this operator by calculating its expectation value for our stationary wave function (37). The result is identical with the scalar case (32) considered before.[16]

However, inspirited by the qubit model with phase factors (Sect. 3.2), we can introduce another tonic operator $\mathbb{T}$, which describes the projection onto the tonic vector $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$:

(38) $\mathbb{T} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$

With the help of this operator, the probability density can be calculated as follows:

(39) $\psi(x)^* \cdot \left(\mathbb{T}\,\psi(x)\right) = \frac{1}{2}(1 + \cos 2\delta(x)\cos(x - \frac{\pi}{2}))$

Taking the series of points $x_k = \frac{\pi(k-3)}{6} = \frac{\pi\,k}{6} - \frac{\pi}{2}$, the following probability distribution will result:

(40) $P(k) = \frac{1}{2}(1 + \cos\left(\check{\delta}(x_k)\right)\cos\left(\pi\,(k-3)/6\right))$, with $\check{\delta} = 2\delta$.

This is exactly the earlier distribution (6) of the qubit model with phase parameters. Again, the tonic is represented by $k = 3$. Hence, the spinor model produces the same probability distribution as the qubit model considering the phase parameters.

In contrast with the deformation model, the spinor phase model gives an opportunity for calculating the dynamic attraction potential. Further, it equally allows the introduction of gauge forces – for example by using a simple ansatz for the phase function $\delta(x)$ as in eq. (35).

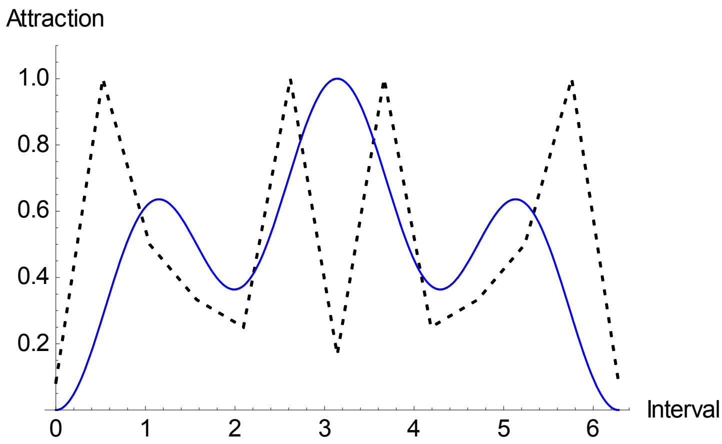



Figure 8: Dynamic kernel functions for the phase model (solid) contrasted with the ICP (dashed).

Figure 8 shows the corresponding kernel function (with tonic at 0 and 2 $\pi$ and triton at $\pi$ – using the new variable $y = x - \frac{\pi}{2}$). This could be seen as a rough approximation to the ICP attraction kernel. Unfortunately, the correlation with the ICP kernel is very low (R = 0.1) (see

[16] When we project onto the vector $\begin{pmatrix}0\\1\end{pmatrix}$, instead, we get the case described in (17) which provides an approximation to the static attraction profile. Of course, the projection operator in this case is $\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$.

Table 2 in Section 4.2). Hence, we can conclude that the force free spinor model does not give a reasonable approximation to the dynamic attraction case.

Despite of this negative result, it is opportune to consider the energy densities in the spinor model. We have to use the following operators (see Appendix 3):

(41) $H = T + M + U$ with

a. $T = -\frac{\partial^2}{\partial x^2}$

b. $M = -2i \cdot \delta'(x)\sigma_3 \frac{\partial}{\partial x}$

c. $U = E + \delta'(x)^2 - i\delta''(x)\sigma_3$ .[17]

An explicit calculation of the energy densities of these operators based on the phase shift function (35) and the tonic operator (38) is straightforward. Figure 9 shows the total energy density and the density of potential energies (left hand side). Note that there are no real forces at work given the choice of (35) as gauge function of the phase model (based on the charged current). Note further that we have re-centred the curves and localized the tonic at position x = 0 of the manifold space and the triton at x = π.

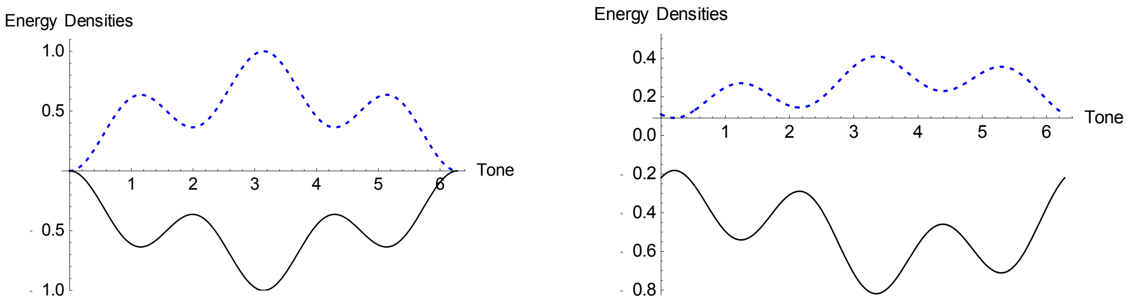



Figure 9: Phase model with stationary wave function as given in (37) in the static case. Dashed: total energy density; solid: density of potential energies (without the imaginary part of magnetism). Left hand side: kernels; right hand side: averaging for C-major triad.

As in the cases considered before, the total energy density corresponds with the probability distribution of the dynamic attraction function.

On the right hand side of Figure 9 we have averaged over the tonic triad CEG of C-major. We get three local minima of the potential function conforming to the tones D (at1.0), #F (at 3.2) , and #A (at 5.2) – an outcome that is not plausible from an empirical point of view. In the following subsection, we make an explicit comparison with the empirical data and show that the model plays a superb role when considered in tandem with the deformation model.

### 4.4 A SU(2) Gauge model for static and dynamic attraction

So far, we have modelled (i) static attraction profiles by spatial deformations and (ii) dynamic attraction profiles by local phase shifts. Further, we have shown that the deformation model is an excellent model for approaching the static attraction case. In contrast, the phase model was much less successful for approaching the dynamic attraction data.

We have already mentioned three variants of how we could improve the theory. First, we could exclude the phase model and consider a variant of the deformation model that fits the dynamic attraction data (beim Graben & Blutner, 2019). Second, we could introduce proper

[17] Note that for calculating the corresponding energy densities, we have to compose these operators with the tonic operator. That means, the relevant operators are $H\,\mathbb{T}$, $T\,\mathbb{T}$, and $U\,\mathbb{T}$ .

gauge forces into the phase model – using a nonlinear gauge field instead of the linear field described by (35) (Blutner, 2019). Third, we could consider a combination of the deformation model with the force-free phase model in order to simulate the idea of Ball (2010) that the static forces allow to infer a certain type of dynamic forces. In this subsection we will develop the third variant and we will explain the advantages of this approach.

From the theoretical point of view, the basic symmetry underlying the present analysis is the SU(2) symmetry. This symmetry can be described by the matrix form (8) transforming a two-dimensional spinor space. This transformation is determined by three generators that are connected with three real parameters $\theta$, $\delta$, and $\tau$. So far, the static case was modelled by a nonlinear SO(2) gauge function (in terms of $\theta$) introducing real gauge forces, see Eq. (11). The question is how we can best describe the dynamic attraction case, in particular, the resolution of different chords. It is suggested that the phase gauge is essential – conforming with the two generators given in (12) and (13).

The application of the phase gauge can happen in different ways. In the second variant explained above, a nonlinear phase gauge was introduced for fitting the available data of chord resolution. Unfortunately, the quality of the fit was not really satisfying (Blutner, 2019). Further, we see a conceptual problems with this approach because it does not make a distinction between forces that apply for the resolution of chords and musical forces that direct the development of melodies and go beyond the resolution of musical chords. Further, the idea of Ball (pointed out in Sect. 2.3) is not respected, which claims that the existence of static tonal forces has a direct influence on the resolution of chords. Another problem is the description of the asymmetry between major and minor chords. This phenomenon cannot be approached by kernel functions that are mirror-symmetric relative to the triton. The breaking of the mirror-symmetry can be handled in an arbitrary way only in the described variants.

In the present subsection, we therefore suggest a third variant which combines the SO(2) gauge (deformation model) with a U(1) gauge (phase model). Importantly, the phase gauge is described by a linear phase function. This leads to a constant potential function, which does not introduce their own dynamic forces. However, the combination with the SO(2) gauge of the deformation models results in the development of which can be interpreted in the spirit of Ball (2010). It suggest a mechanism that derives certain dynamic forces from purely static ones. Further, we will demonstrate that the combination of the SO(2) gauge with a weakly coupled U(1) gauge provides the asymmetry that is required for deriving the asymmetry between major and minor chords.

The following formulas collect the effects of the gauge transformations for calculating the probability densities for spinor gauge models. In all cases, the wave functions $\psi_G$ result from applications of the gauge transformations $G$ to the free spinor wave function $\psi$: $\psi_G = G(\psi)$. We consider the SO(2) gauge function applied in (26), related to the deformation field. We also consider the spinor gauge model based on phase shifts (see Sect. 4.4) with the gauge function of (37), related to the phase field. With the tonic operator (38), then we get the following distributions:

(42) $p_G(x) = \psi_G(x)^* \cdot (\mathbb{T}\, \psi_G(x))$

a. $p_D(y) = \frac{1}{2}\big(1 + cos\ \gamma(y)\big)$ (deformation field)

b. $p_P(y) = \frac{1}{2}(1 + \cos 2\delta(y + \pi/2)\ \ cos\ y)$ (phase field)

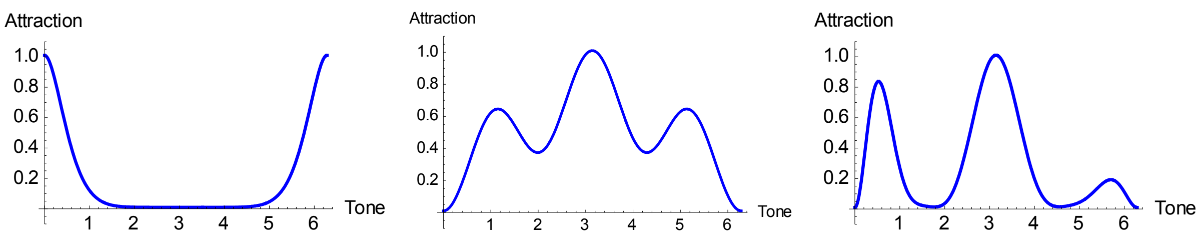
c. $p_{D+P}(y) = \frac{1}{2}(1 - \cos 2\delta(y + \pi/2)\ sin\left(y - \gamma\left(y + \frac{\pi}{2}\right)\right) \cos(y) + \frac{1}{4}\sin\Big(2y - \gamma\left(y + \frac{\pi}{2}\right)\Big) + \frac{1}{4}\sin(\gamma\left(y + \frac{\pi}{2}\right))$ (combined field)

Note that the variable $y = x - \pi/2$ ranges from 0 (tonic) to 2 $\pi$ (tonic again), the triton is marked by $\pi$. Formula (a) refers to the spatial deformation model, and (b) to the phase shift model. Both formulas are mirror-symmetric around the triton (at $\pi$) if the functions $\gamma$ and $\delta$ are chosen as in (22) and (35).

Interestingly, the combination of both models (gauge transformation *D* in tandem with gauge transformation *P*), is not mirror-symmetric around the triton. As a matter of fact, the combination is not simply the sum of the two terms given in (a) and (b). Rather, it contains complicated interference terms that violate the mirror-symmetry. In order to get a visual representation of the suggested kernel functions, we use the simple choice of the underlying gauge fields for the deformation field and the phase field introduced earlier – cf. (22) and (35), repeated here:

(43) a. $\gamma(x) = \pi + (x - \pi)^4/\pi^3$
b. $\delta(x) = x$

On the basis of these functions, the resulting kernel functions look as follows (See Figure 10):




(a) (b) (c)

Figure 10: Kernel function for (a) deformation (neutral current), (b) phase (charged current), (c) the equal combination of both gauge fields

Table 2 shows the correlation functions of the three simple models based on (a) the deformation field alone, (b) the phase field alone, and (c) a combination of both fields in equal ratios. Interestingly, the combination of the first two models shows a breaking of the centre symmetry resulting from the interference of the two involved gauge fields.

| Correlation with dynamic attraction data | ICP Model | Spinor Phase Model | | Combined Model | |
|---|---|---|---|---|---|
| | | Tonic = $\binom{0}{1}$ | Tonic = $\frac{1}{\sqrt{2}}\binom{1}{1}$ | 100% | 85% |
| C-major | .69 | -.11 | .14 | .11 | .55 |
| C-minor | .78 | .06 | .28 | .29 | .7 |
| dominant seventh | .75 | -.2 | -.08 | .37 | .76 |
| French sixth | .79 | 0. | .05 | .9 | .86 |
| half-diminished seventh | .89 | .15 | .19 | .6 | .78 |
| Correlations with ICP | | -0.1 | 0.1 | 0.4 | 0.6 |

Table 2: Dynamic attraction for single and combined models. The deformation model is based on the gauge field of the neutral current. The phase model is based on the corresponding gauge field, where different vectors for the tonic are considered. For the combination we consider two variants. In the first case, both models are coupled equally (100 %). In the second case, the phase model is coupled with 85% only. The dynamic attraction data (Woolhouse 2009) concern (a) major triad CEG, (b) minor triad CE♭G, (c) dominant seventh CEGB♭, (d) French sixth CEG♭B♭, and half-diminished seventh CE♭G♭B♭. For all models, including the ICP model, the correlation

functions with the data of Woolhouse are given. Further, in the last line, the correlations between the ICP kernel and the kernels of the different models are shown.

In order to improve the fit with the empirical data, it is suggested to vary the coupling constant of the charged gauge field. With a value of 85% we get the kernel function shown in Figure 11, which provides a satisfying fit to the data. The quality of this fit is comparable with that of the ICP model of Woolhouse (2009). The correlations shown in Table 2 substantiate this view.

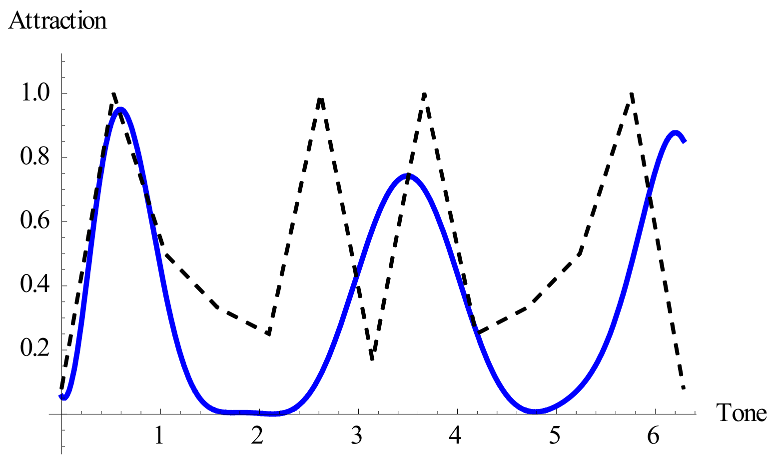



Figure 11: Kernel function for the combined model with an 85% coupling of the phase gauge field.

The energy densities in the combined model are difficult to calculate. Even when the kinetic energy density is defined as before, the magnetic density and the potential density are more complex. Besides the sum of the corresponding terms for phase and deformation models, there are very complex interference terms. We therefore resign to provide a full treatment.

### 4.5 The hierarchical model and the connection between static and dynamic attraction

It is useful to ask for the connection between the deformation model and the hierarchical model. As you see in Figure 4, the kernel function of (static) tonal attraction assigns the maximum value to the target tone (say C). The two neighbours on the circle of fifth (i.e., G and F) get an attraction value that is about half of it. The attraction values of all other tones is very low such that we can neglect them. Hence, when we construct the attraction profiles for a certain context given by a triad (say CEG), we get an approximate reconstruction of the hierarchic model. The three tones of the triad (CEG) get a very high value; C and G a bit higher than E because of the convolution operation. Next, the neighbours of the triadic tones (C→G,F vs. G→D,C vs. E→B,A) are all diatonic tones and get an attraction of about 50%. Hence, we can account for all levels of the hierarchic model shown in Table 1 with exception of the octave level. Concluding, the instrument of analytical functions – by using a strongly damped function – provides the mathematical instrument for describing the hierarchical model.

As outlined in Sect. 2.3, Philip Ball has proposed a connection between static and dynamic attraction (Ball 2010). He suggested a simple, tentative principle. It says that melodic forces are dynamic forces directed towards the chromatically closest tones that are higher in the static attraction hierarchy than the trigger. For example, consider the key of C-major and the trigger B. Then the chromatically closest tones that are higher in static attraction than B are the tones C and A.

The present gauge model (combining neutral and charged currents) equally establishes a close connection between static and dynamic attraction. Hence, the combination of deformation model and phase model (in its simplest form with a constant gauge field) gives a novel description of the dynamic attraction data of Woolhouse (2009). We have noted already that the combined model violates the mirror symmetry (relative to the position of the triton). In other words, it breaks the mirror symmetry of the dynamic kernel function. Since the deformation model itself *satisfies* mirror symmetry, we have to recognize that the combination with the gauge field of the charged current (= phase model) is responsible for the broken mirror symmetry.

Assuming that gauge fields are miracles that emerge in complex neural fields and could be studied in neural field theories (Coombes et al., 2014), one may wonder if the field of charged current has other measurable effects. If yes, this could give an independent motivation for the existence of this gauge field. In the next subsection, we will see that there are facts that suggest violations of mirror symmetry already for static attraction. In addition, we will argue that a very weak influence of the charged current is sufficient to explain the relevant phenomena. Hence, a strong coupling (85%) between neutral current and charged current explains *dynamic* attraction (Woolhouse-data) whereas a very weak coupling (around 3-4%) explains certain asymmetries related to *static* attraction.

### 4.6 The breaking of the mirror symmetry of the hierarchical model

In the preceding sections, we have considered some basic phenomena of static and dynamic attraction only. However, there are other fundamental phenomena which are studied in cognitive music theory. A very prominent phenomenon is the occurrence of graded consonance/dissonance. In an indirect way, it relates to static attraction. According to Parncutt (1989), the degree of (musical) consonance of a chord is allied to the distribution of potential root tones of a chord. Hereby, the root tone can be seen as the tone with the maximum static attraction given the chord as musical context. In cases with a single, prominent root tone, the chord sound more consonant than in cases where several root tones are in competition. Formally, we can explain the degree of consonance of a chord as the static attraction value of the (root) tone with maximum attraction after normalizing the attraction profile (i.e., the attraction values of the 12 tones sum up to 1).

The mirror symmetry (against the triton) of the spatial deformation model leads to important problems when it comes to account for the difference between major and minor modes. Important differences between major and minor modes were discussed already 90 years ago (Heinlein, 1928). Recently, Johnson-Laird, Kang, and Leong (2012) have investigated chords including major triads (CEG), minor triads (CE♭G), diminished triads (CE♭G♭), and augmented triads (CEG#). The following table shows the empirical ratings of the chord's consonance. Clearly, the major chords exhibit the highest degree of consonance followed by the minor chords. Further, the diminished chords are ranked lower and, at the bottom, we (surprisingly) find the augmented chords. It is not difficult to see that the hierarchic model and the symmetric deformation model predict the same degrees of consonance for major and minor chords.

| *Triad Class* | *Empirical Consonance Rating* | *Hierarchical Model* | *Deformation Model* | *SU(2) Model* |
|---|---|---|---|---|
| major | 5.33 | .49 | .49 | .495 |
| minor | 4.59 | .49 | .49 | .49 |
| diminished | 3.11 | .34 | .36 | .34 |
| augmented | 1.74 | .33 | .34 | .33 |

Table 3: Empirical rankings and model predictions for common triads. The predictions of the models concern the strength of the tone with strongest static attraction using normalized attraction profiles. The symmetry breaking is provided by a weak interfering phase gauge field (2 %). It gives the asymmetry between major and minor.

In order to model the ranking of the three triadic chords considered in Table 3, we have considered a modifications of the deformation model. The simple idea is to combine the deformation model, which is based on the gauge field of the neutral current, with a weakly coupled charged current. It is expected that even a weak coupling leads to interferring terms that can provide the wanted asymmetry. It should be stressed that this idea does not introduce any new parameter and is valid for a whole variety of weak couplings (from 0.1% – 8%).

The following pictures show the increase of asymmetry with rising coupling (left 3%, middle 6 %, right 8 %).

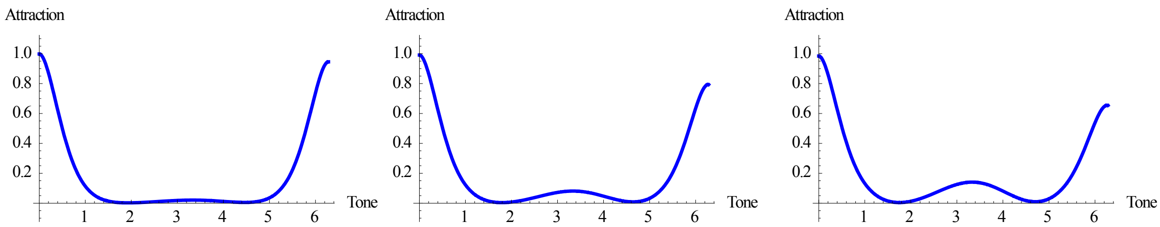



Figure 13: Kernel functions for the combined model with a very weak coupling of the phase gauge field: 3 % on the left hand side, 6 % in the middle, and 8 % and the right hand side.

It is evident that the weak couplin of the charged current does not only break the mirror symmetry relative to the tritone but also octave equivalence (the higher tonics gets a lower degree of attraction than the lower). This fact may be make sense considering the different consonance values for different inversions of a given chord. We cannot follow this issue here.

Summarizing, we have considered symmetry breaking in cognitive musicology, breaking the mirror symmetry of the tonal attraction kernel. In this way, we have overcome some weaknesses of the classical attraction model based on tonal hierarchies. This model cannot even account for the differences between major and minor modes. Consequently, we have made a proposal that accounts not only for static attraction profiles but also for graded consonance/dissonance. The ability for unification – grasping different phenomena in a systematic way – is one of the trademarks of quantum theory. It is correspondingly visible also in the domain of quantum cognition.

# 5 Discussion and Conclusions

In this article, we have contrasted the metaphoric and the realist conceptions of musical forces. Both approaches provide interesting perspectives to study static and dynamic tonal attraction profiles, including their structure, use, and acquisition. The metaphoric conception is taken from mainstream cognitive psychology initiated by work of Lakoff and Johnson (Lakoff, 1987; Lakoff & Johnson, 1980, 1999). The realist conception of musical forces is a new development within the evolving field of quantum cognition and sees musical forces as gauge forces. Here is a short presentation of both approaches.

We start with the metaphoric conception. Notably, a *conceptual metaphor* refers to the understanding of one conceptual field, in terms of another. Famous examples illustrating the idea are "life is a journey" or "time is money". According to this concept, musical forces are constructs *in analogy* to our understanding of physical forces in folk physics. Various authors have proposed different forces, which are assumed to be important for musical perception. Linear regression analysis has been applied in order to find the total effect of musical forces. Unfortunately, the realized fits with empirical data are not really convincing. A sound grounding of forces seems not to be possible in this way.

In contrast, the realist conception of musical forces constructs these forces as gauge forces, which can be derived from fundamental symmetries of the underlying theory and a gauge field. In the present case, we model tones by vectors of a 2-dimensional spinor Hilbert space. Hence, the basic symmetry group is the group of special unitary transformations (SU(2) and their subgroups). If you want, the proposed realist conception is a fresh realization of Kepler's 400 years old vision of *harmonices mundi* (Kepler, 1619) – the vision unifying physics and musicology. Recognizing the failing of Kepler's ideas – mainly for reasons the concern certain astronomical facts not known yet 400 years ago – the present approach is based on a totally different scenery centered around the insights of modern quantum cognition (e.g., Busemeyer & Bruza, 2012).

We have identified the symmetry group SU(2) as the fundamental group that directs the gauge theoretic approach. This symmetry can be described by a matrix transforming a two-dimensional spinor space. This transformation is determined by three generators. They relate to three gauge fields called the neutral current (deformation model) and two variants of the charged current (phase model). Whereas the deformation model accounts for the static attraction data, a combination of deformation and phase model accounts for the dynamic attraction data.

Concerning the results of the gauge theoretic modelling attitude, we will stress four aspects, which we see as main characteristics of the present research. First, there is the idea of discrete convolution – an operation that describes the modification of a kernel function by a distribution of several contextual elements (for instance, the tones of a single chord). Second, the present approach establishes a close connection between static and dynamic attraction. The basic idea of connecting static and dynamic attraction goes back to Ball (2010). The present gauge model (combining neutral and charged currents) establishes a close connection between static and dynamic attraction in a novel way. The gauge field of the neutral current (deformation model) explains static attraction. By adding the gauge field of the charged current (phase model), we are able to explain dynamic attraction data. Third, the combined model – in its simplest form with a linear gauge field for the charged current – violates the mirror symmetry of the dynamic kernel function. The field of charged currents has other measurable effects. Even with a very weak coupling of 3%, it breaks the mirror symmetry of static attraction kernel. This provides an unusual explanation of the fundamental asymmetry between major and minor modes. We can take this observation as an independent motivation for the existence of the gauge field of charged currents. Fourth, the notion of energy densities and musical forces was a particular outcome of the gauge theoretic approach. These concepts are directly connected with the idea of dynamic stability as established in the theory of dynamical systems.

Finally yet importantly, we will mention several unresolved issues, which should be essential points for further research. A first topic is the innateness issue. Already Leonard Bernstein has vehemently disputed the issue expressing and stressed the point that musical apperception is not possible without an innate cognitive background. At the end of his Norton lectures, he formulates his deep believe in the tonal system in the following magical phrases:

> I believe that from that Earth emerges a musical poetry which is by the nature of its sources tonal.
> I believe that these sources cause to exist a phonology of music, which evolves from the universal known as the harmonic series.
> And that there is an equally universal syntax, which can be codified and structured in terms of symmetry and repetition. (Bernstein, 1976)

In the present context, this innate background is mainly constituted by the tonal kernel function and operations such as discrete convolution that are not acquired by learning.

A second issue concerns the full dynamic attraction potential, which does not describe the resolution of chords only but the full development of melodies. Besides the metaphoric approaches we have mention in Sect. 2 there are approaches based on Bayesian networks (Temperley, 2008). The present approach suggest to exploit the dynamic evolution described by the Schrödinger equation, which can be seen as a generalization of the Kolmogorov forward equation (Busemeyer & Bruza, 2012).

A third big question, which was not fully treated in this article, concerns the nature of the distinction between consonance and dissonance. For building the distinction, learned parameters seem to play a much more important role than usually assumed (McDermott et. al., 2016). Further, a developed theory of this distinction is important for defining the 'meaning of music' and its emotional content (Blutner, forthcoming).

# Appendix 1: A brief comparison with alternative models of tonal attraction

In this subchapter, we consider three classes of models that deserve a serious comparison with the qubit model. First, there are spectral representations based on Helmholtz (1863) and his followers. Modern representatives of this approach are Milne, Laney, and Sharp (2015). Second, we have to consider cyclic pulse representations that are useful for implementing the idea of tonal fusion (Stumpf, 1883, 1890). Modern representatives of this approach are Ebeling (2008) and Stolzenburg (2015). The third kind of approach includes template-based models that approach the question how tone-like a sound is. The starting point of this research was set by Terhardt, Stoll, and Seewann (1982) – followed by authors such as Parncutt (1988), Hofmann-Engl (2004), and others.

The aim of the following discussion is not a careful discussion of the pros and contras of the three approaches. Rather, it is an identification of the conditions where the models agree with the qubit model and an indication how the qubit model could be modified to come closer to each of the three models.

## 1 Spectral representation

The spectral pitch class model of Milne and colleagues (Milne et al., 2015; Milne, Laney, & Sharp, 2016) is based on the thesis that any complex tone is represented by a spectral pitch class vector. Each of the 1,200 elements of this vector represents a different log-frequency in cents (modulo the octave to ensure octave-equivalence). The value of that element is a model of the expected number of partials (frequency components) perceived at that log-frequency. In more detail, the following assumptions are made:

(i) For each simple tone with a certain basic frequency $f$, all partials with frequencies $n{\cdot}f$ are calculated and related to one octave band.

(ii) We smear each spectral component in the log-frequency domain to model perceptual inaccuracy. The width of this smearing (called smoothing width $\sigma$) is one of the parameters of the model

(iii) We assume damping of amplitudes that is exponentially increasing with the number of partials. The steepness at which they reduce is a parameter called roll-off ($\rho$) .

(iv) Complex tones composed of several simple tones relate to vector representations that are the superposition of the simpler representations.

(v) Cosine similarity, which takes a value between 0 and 1, is taken to calculate the similarity between two vector representations. In particular, the static attraction values of a tone $k$ given a triadic context $c$ is the cosine similarity between the vector representation of tone $k$ and the vector representation of the corresponding triad $c$.

As an illustration consider Figure A1, which is self-explanatory.

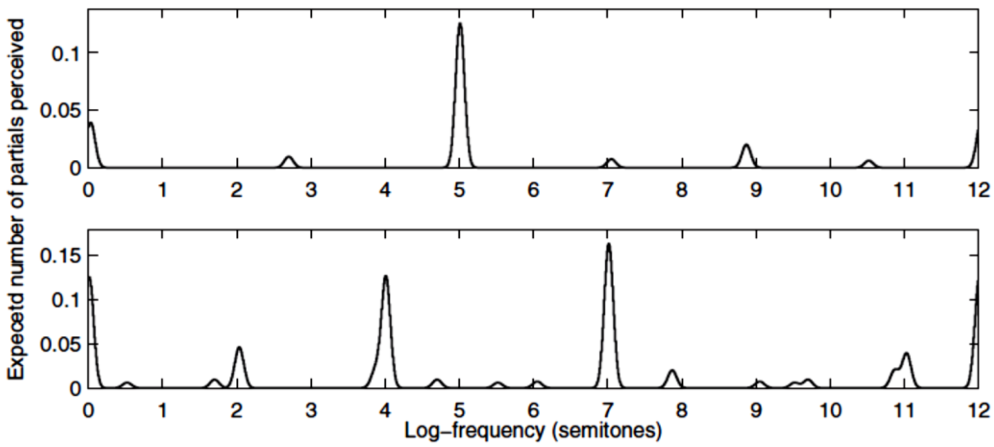



Figure A1: The spectral pitch class vectors for a major triad (bottom) and a pitch class five semitones higher than the former's root (top). The parameters are as optimized to the probe tone data ρ = 0.67 and σ = 5.95 (from Milne et al. (2015), p. 367).

## 2 Cyclic pulse representations

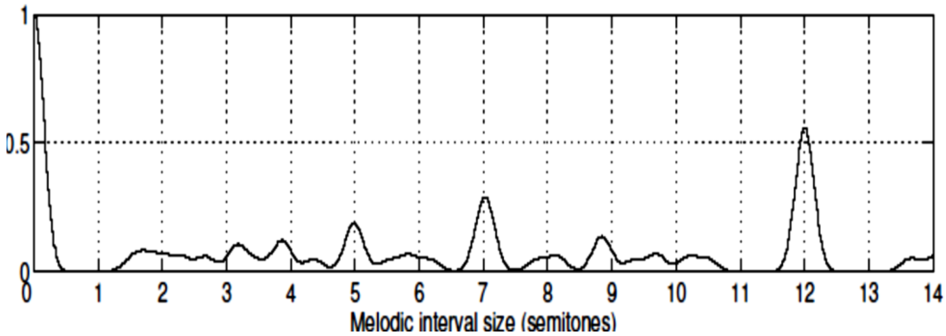
Recently, Ebeling (2008, 2009) gave a mathematical account of harmonicity and a modern explication of Stumpf's ideas in terms of autocorrelation. He defined a function called *generalized coincidence function*, which "is a measure value of overall coincidence between the two tones of the musical interval with regard to pulse forms and pulse widths" (Ebeling, 2008, p. 2325). It basically is grounded on the autocorrelation of a musical interval.

Here is a sketch of Ebeling's construction. First, the autocorrelation of the spiking sequence J(t) is calculated, where J(t) is the superposition of the spiking sequences generated by the two tones of the interval $a(t)$ and $b(t)$, i.e. $J_s(t) = a(t) + b(t)$, where the parameter $s$ describes the vibration ratio (the quotient of the two involved fundamental frequencies). Generally, the auto-correlation is defined as $R(\tau, s) = \int_{-T}^{T} J_s(t)\, J_s(t+\tau) dt$. It is a function of the period under discussion (i.e., the inverse of the frequency). Second, the general coincidence function is defined by integrating the square of autocorrelation as a function: $K(s) = \int_{-T}^{T} R^2(\tau, s) d\tau$. Ebeling (2008) explicitly performed the relevant calculation for rectangle spikes. He found three kinds of peaks if two spike trains $a(t)$ with frequency $f_a$ and $f_b$ (and corresponding periods of $T_a = 1/f_a$ and $T_b = 1/f_b$) are superposed: peaks arising from the autocorrelation of the first spike train at $\tau = m \cdot T_a$, peaks arising from the autocorrelation of the second spike train at $\tau = n \cdot T_b$, and peaks arising from the cross-correlations between the two spike trains at $\tau = m \cdot T_a + n \cdot T_b$ (with natural numbers $m$ and $n$).

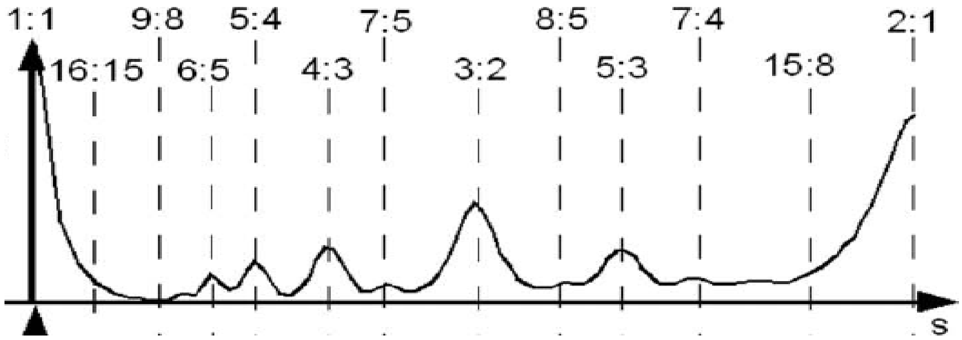
Let us consider a simple example illustrating the basic idea in establishing a fundamental pitch. Consider spiking trains with 250 and 200 Hz. The corresponding fundamental pitch has a frequency of 50 Hz and can be heard as a virtual pitch (see the corresponding example 29 of Terhardt's audio material; Terhardt, 1998). In the example, we have $T_a = 5$ ms and $T_b = 4$ ms. Resulting from the three sources, we find an accumulation of 4 simultaneous peaks at times 20 ms, 40 ms, 60 ms, …. This corresponds exactly to the period of the virtual pitch of 20 ms and a fundamental frequency of 50 Hz. Hence, the approach based on autocorrelation provides a solution to the root phenomenon discussed at the beginning of this section. Technically, the energy of the corresponding oscillation is measured by the generalized coincidence function $K(s)$ (Ebeling 2008).

Cariani and Delgutte (1996a, 1996b) have investigated the temporal discharge patterns of auditory nerve fibres in anesthetized cats. In this study, periodic complex acoustic waveforms where presented that evoked pitches at their fundamental frequencies. The authors have calculated autocorrelations for all relevant spiking trains and they have demonstrated that the maximal peaks for periods is corresponding to the pitch. These investigations give a direct underpinning of the thesis that with the help of autocorrelation the virtual pitch is calculated corresponding to the real auditory perception.

The following Figure shows that the spectral pitch similarity found by Milne et al (2016) is very similar to the found generalized coincidence function defined by Ebeling 2008). Even when the agreement between the two curves is not complete, there is an extremely high correlation between the two models.



Milne et al. (2016)



Ebeling (2008)

Figure A2. The upper curve is Milne's, the lower curve is Ebeling's. Milne is exploiting the partials directly, in the autocorrelation approach of neural spiking trains, the partials are not reflected directly. Instead of counting the fitting partials (in a frequency representation), the fitting of spikes is checked. For instance, for a vibration ratio of 3/2 (fifth) it can be said that the second partial of the upper tone agrees with the third partial of the lower tone. In the language of spike a related regularity appears: each third spike of the lower train coincides with each second spike of the upper train.

Conceptually, both models have their advantages. Ebeling (2008) cites direct neurophysiological evidence for the biological reality of autocorrelation mechanism (Cariani & Delgutte, 1996a, 1996b; Langner, 2007). Milne's model has the advantage of a proper analysis of tonal similarities. Both models are conforming to Stumpf's idea of tonal fusion and both models conform to the findings of Josh H. McDermott, Lehr, and Oxenham (2010) , suggesting to consider harmonicity as a foundation of musical preferences (consonance/dissonance). At the moment it is completely unclear whether one of the two models provides better fits to the available empirical data and what is their substantial difference concerning the empirical impact. Further, both models can be applied for modelling consonance/dissonance and both stress a biological rational for the dichotomy (in agreement with Bowling & Purves, 2015).

Template-based models of pitch and rhythm perception assume the existence of certain features which are crucial for understanding tonality and rhythm. For instance, the saliences of pitches perceived in chords may be determined using tone or chord templates. Mental representations are founded on the assumption that the harmonicity of complex tones and chords is dependent of the series of partials involved in harmonicity. The so-called *root phenomenon* means the existence of single notes that represent a whole chord. According to Terhardt (1998), a musical chord evokes a *virtual pitch* when a set of spectral pitches matches to lower elements of a harmonic series. In other words, the root of a chord is characterized by the fact that one (or more) of its partials maximally resonate with the tones of the chord.

Terhardt (1982) formulated a simple algorithm for constructing the perceptual root of a chord. The idea is that a (potential) root tone of a chord conforms to the subharmonics of the tones of the chord. My presentation of the algorithm follows Parncutt (1988). The algorithm is based on five "root supports" corresponding to the intervals between a target tone and *its first ten harmonics*. This is shown in Figure A3 with C as target tone. It is evident that there are exactly five different intervals that can be derived from the harmonics: perfect unison, perfect fifth, major third, minor seventh, and major second. (Due to conventional theory, the intervals are collapsed into a single octave).

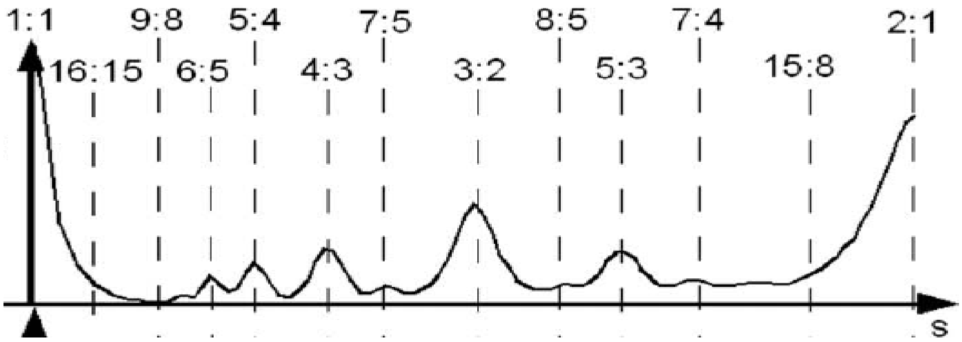



Figure A3: Harmonic pitch analysis of a single target tone (C).

From this Figure we can derive the emergence of the hierarchic model with an octave level (1:1), a fifth level (3:2 and 4:3) and possibly a third level (5:4 and 5:3).

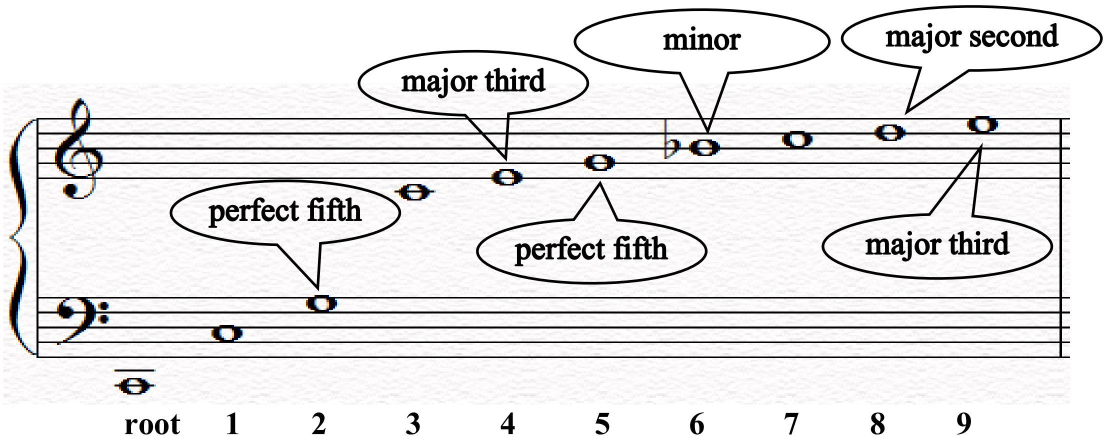



Figure A4: The first 9 partials of a given root tone (C in the present case). The intervals that are defined by the root and one of the first 9 partials roughly correspond to the perfect unison (not indicated in the

Figure), the perfect fifth (3/2), the major third (5/4), the minor seventh (7/4), and the major second (9/8).

Terhardt's algorithm simply counts the number of tones of the chord that are supported by one of the harmonics of the root candidate; i.e. we consider all five root supports and check how many of these intervals connect our root candidate with one of the tones of the chord. The candidate with the maximum number of supports is the winner of the competition. In Table A1 the number of supports is calculated for each of the 12 tones relative to the major triad. Intuitively, we can say that the number of supports is the higher the better the fit is between the chord and the harmonic series generated by the target tone (= root candidate).

| | Major Triad C E G | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **root candidates** | **C** | **D♭** | **D** | **E♭** | **E** | **F** | **G♭** | **G** | **A♭** | **A** | **B♭** | **B** |
| perfect unison | C | - | - | - | E | - | - | G | - | - | - | - |
| perfect fifth | G | - | - | - | - | C | - | - | - | E | - | - |
| major third | - | - | - | - | - | - | - | - | - | - | - | - |
| minor seventh | - | - | - | - | - | - | E | - | - | G | - | - |
| major second | - | - | E | - | - | - | - | - | - | - | - | - |
| **root supports** | 2 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 |
| | 3/2 | 0 | 1/5 | 0 | 1 | 1/2 | 1/4 | 1 | 0 | 3/4 | 0 | 0 |

Table A1: Root candidates and their root supports for the major triad. The numbers in the last two lines are measures for root support as given by Terhardt (1982) and Parncutt's (1988) modification of it.

# Appendix 2: The Gauge Manifesto

1. All musical forces are gauge forces

2. Any gauge force is founded in a symmetry group and a gauge field

3. Tones are modelled by vectors of a 2-dimensional spinor Hilbert space. Hence, the basic symmetry group is the group of unitary transformations ($SU_2$ and their subgroups)

4. The stationary Schrödinger-Pauli equation for spinors $\psi(x)$ is the fundamental equation under discussion. It has the following general form:

$$(44)\quad -\frac{\partial^2\psi(x)}{\partial x^2} + \boldsymbol{M}(x)\cdot\frac{\partial\psi(x)}{\partial x} + \boldsymbol{S}(x)\cdot\psi(x) = E\,\psi(x).$$

Hereby, the matrix function $\boldsymbol{M}(x)$ describes the magnetic vector potential, and the matrix function $\boldsymbol{S}(x)$ describes the scalar potentials such as gravity, harmonic oscillator potential, etc.

5. A special case of the Schrödinger-Pauli equation is the *free* equation:

$$(45)\quad -\frac{\partial^2\psi(x)}{\partial x^2} = E\,\psi(x).$$

The free equation corresponds to the general equation (44) with $\boldsymbol{M}(x) = 0$ and $\boldsymbol{S}(x) = 0$. A simple solution is $\psi(x) = \begin{pmatrix}\cos x/2\\ \sin x/2\end{pmatrix}$, with $E = 1/4$.

6. All gauge forces result from gauge transformations.

7. A gauge transformation is a representation of the fundamental symmetry group (or its subgroup). It is specified by a particular gauge field.

8. The forces specified in the Hamiltonian by the matrix fields $\boldsymbol{M}(x)$ and $\boldsymbol{S}(x)$ are *founded* by a gauge transformation if it transforms all solutions of the stationary Schrödinger-Pauli equation (44) (with specified matrix functions $\boldsymbol{M}(x)$ and $\boldsymbol{S}(x)$) into a force-free solution of (45).

9. *Deformation forces* and *phase forces* are defined by particular gauge transformations. Deformation forces result from the rotation group $SO_2$ and the gauge field $\theta(x)$ of rotation. Phase forces result from the unitary group $U_1$ and the gauge field of phase shifts $\delta(x)$. The combination of both gauges is possible.

10. Other symmetry groups relevant for tonal music are transposition symmetry (defines by the cyclic group) and particular modulation groups investigated by Mazzola.

## Appendix 3: Spatial Deformation Model and Spinor Phase model

This part compares the spatial deformation model with the phase model. Note that the $\psi$-functions differ for the spatial deformation model and the phase model.

Let us consider first the spatial deformation model. We consider the following local gauge transformation, which converts the force-free solution of the Schrödinger equation into the deformed solution under the influence of gauge forces when we assume that $\gamma(x) = \theta(x) - x$. We refer to our earlier formulation (26), which is repeated here:

(46) $$\begin{pmatrix}\tilde{\psi}_+(x)\\ \tilde{\psi}_-(x)\end{pmatrix} \rightarrow \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix} = \begin{pmatrix}\cos\theta(x) & -\sin\theta(x)\\ \sin\theta(x) & \cos\theta(X)\end{pmatrix}\begin{pmatrix}\tilde{\psi}_+(x)\\ \tilde{\psi}_-(x)\end{pmatrix} = \begin{pmatrix}\cos(\gamma(x)/2)\\ \sin(\gamma(x)/2)\end{pmatrix}.$$

Considering the (spinor) wave function $\psi(x) = \tilde{\psi}(\gamma(x))$, we assume the free Schrödinger equation (15) is valid, i.e. $-\frac{\partial^2\psi(x)}{\partial x^2} = E\,\psi(x)$. Differentiating the spinor $\psi(x) = \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix}$ twice, we can derive that the spinor $\psi(x)$ satisfies the following differential equation:

(47) $$-\frac{\partial^2\psi_i(x)}{\partial x^2} + \frac{\gamma''(x)}{\gamma'(x)}\frac{\partial\psi_i(x)}{\partial x} - \gamma'(x)^2\,\psi_i(x) = 0.$$

This corresponds to the general form of the stationary Schrödinger-Pauli equation for spinors $\psi(x)$ discussed in Appendix 2:

(48) $[T + M + U]\,\psi_i(x) = E\,\psi_i(x)$, with

a. $T = -\frac{\partial^2}{\partial x^2}$

b. $M = \frac{\gamma''(x)}{\gamma'(x)}\frac{\partial}{\partial x}$

c. $U = E - \gamma'(x)^2$ .

E is a constant referring to the eigenvalue of the Hamiltonian $H(x) = T(x) + M(x) + U(x)$. As usual in physics, we call the operator $T$ the operator of kinetic energy and we call the operator $U$ the operator of potential energy. The operator $M$ has been called magnetism in analogy to the physical examples. The details of the operators for $M$ and $U$ vary, of course, with the special form of the gauge field $\theta(x)$ (or $\gamma(x)$).

For the phase model, we consider the following local gauge function based on the unitary transformation $E(0, \delta(x), 0)$, which gives the following transformed wave function:[18]

(49) $$\psi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} cos(x/2)e^{-i\delta(x)} \\ sin(x/2)e^{i\delta(x)} \end{pmatrix}$$

Again, we start with the free Schrödinger equation and differentiate the spinor $\psi(x)$ twice. Note that the two components of the spinor $\begin{pmatrix} \psi_+(x) \\ \psi_-(x) \end{pmatrix}$ satisfy the following differential equations:

(50) a. $$-\frac{\partial^2\psi_+(x)}{\partial x^2} - 2i\left(\delta'(x)\frac{\partial\psi_+(x)}{\partial x} + \frac{1}{2}\delta''(x)\psi_+(x)\right) + \delta'(x)^2\,\psi_+(x) = 0.$$

b. $$-\frac{\partial^2\psi_-(x)}{\partial x^2} + 2i\left(\delta'(x)\frac{\partial\psi_-(x)}{\partial x} + \frac{1}{2}\delta''(x)\psi_-(x)\right) + \delta'(x)^2\,\psi_-(x) = 0.$$

Note the change of the signs for the middle term of the equations. In a compact form, another form of the Schrödinger-Pauli equation can be derived, where the Pauli matrix $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ describe the mentioned alteration of sign:

(51) $$[T + M + U]\,\psi_i(x) = E\,\psi_i(x), \text{ with}$$

a. $T = -\frac{\partial^2}{\partial x^2}$

b. $M = -2i \cdot \left(\delta'(x)\frac{\partial}{\partial x} + \frac{1}{2}\delta''(x)\right)\sigma_3$

c. $U = E + \delta'(x)^2$ .

In contrast to the deformation model, the operators of kinetic energy and the operator of magnetism can have imaginary expectation values. For the sum of all energies $H = T + M + U$, however, we always has a real expectation value.

For calculating the densities of the different energy operators, we have to consider the tonic operator which defines which element of the manifold space refers to the tonic. If the tonic refers to x=0, then the tonic spinor is the vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and the tonic operator is the corresponding projector, i.e. $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. If the tonic refers to x=π/2 then the tonic spinor is $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and the tonic operator is $\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. For calculating the probability density that a tone collapses into a particular tonic t (with the corresponding tonic operator $\mathbb{T}$) we have to consider the probability density $p_t(x) = |\psi(x) \cdot t|^2$, which is equivalent with $\psi(x)^*\,\mathbb{T}\,\psi(x)$. Since ψ is an Eigenfunktion of the Hamiltonian, the total energy density is proportional to the probability density $E \cdot p_t(x)$. To make the correct reference to the tonics, we always have to compose the enrgy operators with the chosen tonic operator.
Table A2 compares the energy operators for the deformation model and phase model. In all cases, the density functions are different because the $\psi$-functions differ for the two models. This is best visible for the function $H(x)$, which is proportional to the probability density. Of course, the probability densities are different for the two models. Further, note that for the

[18] The treatment for the phase model with $E(0, 0, \tau(x))$ is analogous and does not lead to different results.

deformation model, magnetism is a real-valued function whereas it is a function with purely imaginary values for the phase model.

| | Spatial deformation | Spinor phase model |
|---|---|---|
| | $\psi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}\cos(\gamma(x)/2)\\ sin(\gamma(x)/2)\end{pmatrix}$ | $\psi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}cos(x/2)e^{-i\delta(x)}\\ sin(x/2)e^{i\delta(x)}\end{pmatrix}$ |
| $T(x)$ | $\psi(x)^* \cdot -\frac{\partial^2}{\partial x^2} \cdot \mathbb{T}\,\psi(x)$ | $\psi(x)^* \cdot -\frac{\partial^2}{\partial x^2} \cdot \mathbb{T}\,\psi(x)$ |
| $M(x)$ | $\psi(x)^* \cdot \frac{\gamma''(x)}{\gamma'(x)} \cdot \frac{\delta}{\delta x} \cdot \mathbb{T}\,\psi(x)$ | $-2i \cdot \psi(x)^* \cdot\ \delta'(x) \cdot \sigma_3 \frac{\delta}{\delta x}\ \cdot \mathbb{T}\,\psi(x)$ |
| $U(x)$ | $\psi(x)^* \cdot (E - \gamma'(x)^2) \cdot \mathbb{T}\,\psi(x)$ | $\psi(x)^* \cdot (E + \delta'(x)^2 - i\delta''(x)\sigma_3) \cdot \mathbb{T}\,\psi(x)$ |
| $H(x)$ | $E\ \psi(x)^* \cdot \mathbb{T}\,\psi(x)$ | $E\,\psi(x)^* \cdot \mathbb{T}\,\psi(x)$ |
| | with $E = \frac{1}{4}$ in both cases | |

Table A2: Kinetic, magnetic and potential energy for spatial deformation model and phase model.

In the main part of this article, we discuss the shape of the energy densities in detail.

# Appendix 4: Representation theory of SU(2)

In general, the group SU(n) is the group of n×n unitary complex matrices with determinant 1. For the special case n= 2, we have claimed in Formula (8) that each unitary transformation with determinant 1 can be represented by the matrix $\begin{pmatrix}cos\,\theta e^{-i\delta} & -sin\,\theta e^{i\tau}\\ sin\,\theta e^{-i\tau} & cos\,\theta e^{i\delta}\end{pmatrix}$ with real parameters $\theta$, $\delta$, and $\tau$.
For generating this matrix, we consider the three Pauli matrices.

(52) $\sigma_1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\ \sigma_2 = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\ \sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$

Is it possibly to construct the matrix by the following generating expression:

(53) $SU(2) = \prod_{j=1}^{3} e^{-\frac{i}{2} par_j\, \sigma_j}$ , with $par_1 = \delta, par_2 = \theta, par_3 = \tau$.

For example, the middle term of the product corresponds to the matrix $\begin{pmatrix}cos\,\theta & -sin\,\theta\\ sin\,\theta & cos\,\theta\end{pmatrix}$ describing a spatial rotation, whereas the first and the last term correspond to phase shifting matrices.
In elementary particle physics, SU(2) is the symmetry group of weak interaction. It describes fermions (electron, neutrino) that act as pairs. The interaction consists in the exchange of three kinds of bosons, called Z, $W^{+}$and $W^{-}$. The ensemble of Z bosons is called neutral current, and the W bosons build the charged current.
The corresponding latter operators are given as follows:

(54) a. $\sigma_+ = \frac{1}{2}(\sigma_1 + i\,\sigma_2)$ , for $W^+$

b. $\sigma_- = \frac{1}{2}(\sigma_1 - i\,\sigma_2)$ , for $W^-$
c. $\sigma_3$, for $Z\ /\ W^0$

## Acknowledgement

We are deeply indebted to Manfred Bierwisch, the late Johannes Ranft and Harald Atmanspacher for intensive discussions. Thanks go to Gisela Ranft for explaining the crucial traits of quantum field theory and for debating several variants of it. Needless to say that for the remaining weaknesses and errors of this article, no one but ourselves can be hold responsible.

References

Alexandroff, P. (2012). *An Introduction to the Theory of Groups*. New York: Dover.
Ball, P. (2010). *The Music Instinct: How music works and why we can't do without it*. London: Bodley Head.
Bernstein, L. (1976). *The unanswered question: Six talks at Harvard*. Princeton: Harvard University Press.
Bharucha, J. J. (1996). Melodic Anchoring *Music Perception,* 13(3), 383-400.
Blutner, R. (2015). Modelling tonal attraction: Tonal hierarchies, interval cycles, and quantum probabilities. DOI 10.1007/s00500-015-1801-7. In M. L. Dalla Chiara, R. Giuntini, E. Negri & S. Smets (Eds.), *Logic, Music and Quantum Information*). Berlin: Springer.
Blutner, R. (2016). Musical Forces and Quantum Probabilities). Amsterdam: ILLC, University of Amsterdam. Available from https://www.researchgate.net/publication/301289738_Musical_Forces_and_Quantum_Probabilities.
Blutner, R. (2019). Musical Forces and Two Gauge Models of Tonal Attraction). Berlin. Available from http://www.blutner.de/Documents/musical%20forces%20and%20two%20gauge%20models.pdf
Blutner, R. (forthcoming). Towards a bare bone semantics of pure music.
Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences,* 112(36), 11155-11160.
Brown, H., Butler, D., & Jones, M. R. (1994). Musical and temporal influences on key discovery. *Music Perception,* 11(4), 371-407.
Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum Cognition and Decision*. Cambridge, UK Cambridge University Press.
Cariani, P. A., & Delgutte, B. (1996a). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *Journal of Neurophysiology,* 76(3), 1698-1716.
Cariani, P. A., & Delgutte, B. (1996b). Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *Journal of Neurophysiology,* 76(3), 1717-1734.
Coombes, S., beim Graben, P., Potthast, R., & Wright, J. (Eds.). (2014). *Neural fields: theory and application*. Berlin: Springer.
Dawid, R. (2017). Scientific Realism and High Energy Physics). LMU Munich: available from http://philsci-archive.pitt.edu/12900/.
de Groot, M. L. (2013). *Two computational methods of attributive modification in natural language semantics compared.* Unpublished Bachelor Thesis, https://esc.fnwi.uva.nl/thesis/centraal/files/f218739155.pdf Universiteit van Amsterdam, Amsterdam.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review,* 88, 503-522.
Ebeling, M. (2008). Neuronal periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion. *Journal of the Acoustical Society of America,* 124(4), 2320–2329.
Ebeling, M. (2009). Zum Wesen der Konsonanz: Neuronale Koinzidenz, Verschmelzung und Rauhigkeit. *Jahrbuch der Deutschen Gesellschaft für Musikpsychologie: Musikalisches Gedächtnis und musikalisches Lernen*, 71-93.
Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. Cambridge, Mass.: The MIT Press.
beim Graben, P., & Blutner, R. (2017). Toward a Gauge Theory of Musical Forces. In J. A. de Barros, B. Coecke & E. Pothos (Eds.), *Quantum Interaction. 10th International Conference (QI 2016), LNCS 10106* (pp. 99-111). Berlin, Heidelberg, New York: Springer.
beim Graben, P., & Blutner, R. (2019). Quantum approaches to music cognition. *Journal of Mathematical Psychology,* 91, 38-50.
Harlander, R. (2013). *Wie real ist der Higgs-Mechanismus?* , from https://www.weltderphysik.de/gebiet/teilchen/bausteine/jenseits-des-standardmodells/das-standardmodell-umfassend-aber-nicht-genug/wie-real-ist-der-higgs-mechanismus/
Heinlein, C. P. (1928). The affective characters of the major and minor modes in music. *Journal of Comparative Psychology,* 8(2), 101-142.
Helmholtz, H. v. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig: Friedrich Vieweg und Sohn.
Hofmann-Engl, L. J. (2004). Virtual pitch and its application to contemporary harmonic analysis): Chameleon Group Online Publication. Available from: http://www.chameleongroup.org.uk/research/virtual_analysis.html.
Hubbard, T. L. (2017). Momentum in Music: Musical Succession as Physical Motion. *Psychomusicology: Music, Mind, and Brain,* 27(1), 14-30.
Jackson, J. D., & Okun, L. B. (2001). Historical roots of gauge invariance. *Reviews of Modern Physics,* 73(3), 663-694.
Johnson-Laird, P. N., Kang, O. E., & Leong, Y. C. (2012). On musical dissonance. *Music Perception,* 30(1), 19-35.
Kepler, J. (1619). *Harmonices mundi*. Linz (Austria): Edited by Gottfried Tampach, printed by Johannes Plank.
Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* 11, 346-374.
Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
Krumhansl, C. L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum,* 17, 53-80.
Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies in music. In *Music perception* (pp. 51-87): Springer.
Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys. *Psychological Review of General Psychology,* 89, 334-368.
Lake, W. (1987). *Melodic perception and cognition: The influence of tonality. Unpublished doctoral dissertation.* University of Michigan, Ann Arbor.
Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.
Langner, G. (2007). Temporal processing of periodic signals in the auditory system: neuronal representation of pitch, timbre, and harmonicity (in German). *Zeitschrift für Audiologie,* 46, 8-21.
Larson, S. (1997-98). Musical Forces and Melodic Patterns. *Theory and Practice,* 22/23(8), 55-71.
Larson, S. (2002). Musical Forces, Melodic Expectation, and Jazz Melody. *Music Perception,* 19(3), 351-385.
Larson, S. (2004). Musical Forces and Melodic Expectations: Comparing Computer Models and Experimental Results. *Music Perception,* 21(4), 457-498.
Larson, S. (2012). *Musical Forces: Motion, Metaphor, and Meaning in Music*: Indiana University Press.
Larson, S., & van Handel, L. (2005). Measuring musical forces. *Music Perception,* 23(2), 119-136.
Lerdahl, F. (1988). Tonal pitch space. *Music Perception,* 5(3), 315-350.
Lerdahl, F. (1996). Calculating Tonal Tension. *Music Perception,* 13(3), 319-363.
Lerdahl, F. (2001). *Tonal Pitch Space*. NewYork: Oxford University Press.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, Mass.: The MIT Press.
Lieb, E. H., & Seiringer, R. (2010). *The stability of matter in quantum mechanics*. Cambridge: Cambridge University Press.
Lyapunov, A. M. (1966). *Stability of Motion (English Translation of "The General Problem of the Stability of Motion" -- In Russian, Doctoral dissertation, Univ. Kharkov 1892*. New York & London: Academic Press.
Margulis, E. H. (2003). *Melodic expectation: A discussion and model.* Unpublished Ph.D., Columbia University.
Mazzola, G. (1990). *Geometrie der Töne*. Basel: Birkhäuser.
Mazzola, G. (2002). *The topos of music: geometric logic of concepts, theory, and performance*. Basel: Birkhauser.
McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual Differences Reveal the Basis of Consonance. *Current Biology,* 20, 1035-1041.
McDermott, J. H., Schultz, A., Undurraga, E., & Godoy, R. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature,* 25, 21-25.
Milne, A. J., Laney, R., & Sharp, D. B. (2015). A spectral pitch class model of the probe tone data and scalic tonality. *Music Perception,* 32(4), 364-393.
Milne, A. J., Laney, R., & Sharp, D. B. (2016). Testing a spectral model of tonal affinity with microtonal melodies and inharmonic spectra. *Musicae Scientiae*, 1-30.
Narmour, E. (1992). *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. Chicago: University of Chicago Press.
Parncutt, R. (1988). Revision of Terhardt's psychoacoustical model of the root (s) of a musical chord. *Music Perception: An Interdisciplinary Journal,* 6(1), 65-93.
Parncutt, R. (1989). *Harmony: A psychoacoustical approach*. Berlin, Heidelberg, New York: Springer.
Parncutt, R. (2011). The tonic as triad: Key profiles as pitch salience profiles of tonic triads. *Music Perception: An Interdisciplinary Journal,* 28(4), 333-366.
Penrose, R. (2004). *The road to reality*. London: Jonathan Cape.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Rutgers University and University of Colorado at Boulder: Technical Report RuCCSTR-2, available as ROA 537-0802. Revised version published by Blackwell, 2004.
Rydberg, J. R. (1890). Recherches sur la constitution des spectres d'émission des éléments chimiques. *Kungliga vetenskapsakademiens handlingar,* 23(11).
Schönberg, A. (1911/1978). *Harmonielehre*. Wien Verlagsanstalt Paul Gerin (Translated by R. E. Carter as: Theory of Harmony. Berkeley: University of California Press, 1978).
Schrödinger, E. (1926). Quantisierung als Eigenwertproblem. *Annalen der Physik,* 13(385), 437-490.
Smolensky, P., & Legendre, G. (2006). *The Harmonic Mind: From neural computation to optimality-theoretic grammar*. Cambridge, Mass.: MIT Press.
Stolzenburg, F. (2015). Harmony perception by periodicity detection. *Journal of Mathematics and Music,* 9(3), 215-238.
Stumpf, C. (1883). *Tonpsychologie, Band 1*. Leipzig: Hirzel.
Stumpf, C. (1890). *Tonpsychologie, Band 2*. Leipzig: Hirzel.
Temperley, D. (2008). A probabilistic model of melody perception. *Cognitive Science,* 32(2), 418-444.
Terhardt, E. (1972). zur Tonhöhenwahrnehmung von Klängen I & II. *Acustica,* 26, 173-199.
Terhardt, E. (1982). Die psychoakustischen Grundlagen der musikalischen Akkordgrundtone und deren algorithmische Bestimmung. . In C. Dahlhaus & M. Krause (Eds.), *Tiefenstruktur der Musik*). Berlin: Technical University of Berlin.
Terhardt, E. (1998). *Akustische Kommunikation: Grundlagen mit Hörbeispielen*. Berlin: Springer-Verlag.
Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America,* 71, 679-688.
Weinberg, S. (1992). *Dreams of a final theory*. New York: Pantheon Books.
Weyl, H. (1950). *Space Time Matter*. New York: Dover Publication.
Woolhouse, M. (2009). Modelling Tonal Attraction Between Adjacent Musical Elements. *Journal of New Music Research,* 38(4), 357–379.
Woolhouse, M. (2010). Modes on the Move: Interval Cycles and the Emergence of Major-Minor Tonality. *Empirical Musicology Review,* 5(3), 62-83.
Woolhouse, M., & Cross, I. (2010). Using Interval Cycles to Model Krumhansl's Tonal Hierarchies. *Music Theory Spectrum,* 32(1), 60-78.

.