A PERIODICITY-BASED THEORY FOR HARMONY PERCEPTION AND SCALES

Frieder Stolzenburg

Hochschule Harz, Automation & Computer Sciences Department, 38855 Wernigerode, GERMANY
fstolzenburg@hs-harz.de

ABSTRACT

Empirical results demonstrate, that human subjects rate harmonies, e.g. major and minor triads, differently with respect to their sonority. These judgements of listeners have a strong psychophysical basis. Therefore, harmony perception often is explained by the notions of dissonance and tension, computing the consonance of one or two intervals. In this paper, a theory on harmony perception based on the notion of periodicity is introduced. Mathematically, periodicity is derivable from the frequency ratios of the tones in the chord with respect to its lowest tone. The used ratios can be computed by continued fraction expansion and are psychophysically motivated by the just noticeable differences in pitch perception. The theoretical results presented here correlate well to experimental results and also explain the origin of complex chords and common musical scales.

1. INTRODUCTION

1.1 Motivation

Music perception and composition seem to be influenced not only by convention or culture, manifested by musical styles or composers, but also by the psychophysics of tone perception [1–3]. Thus, in order to better understand the process of musical creativity and information retrieval, the following questions should be addressed:

- What are underlying (psychophysical) principles of music perception?

- How can the perceived sonority of chords and scales, in particular of western music, be explained?

Therefore, in the rest of this section (Sect. 1), we will introduce basic musical notions and results. After that, we will briefly review existing psychophysical theories on harmony perception (Sect. 2), which are often based on the notions dissonance and tension, taking harmonic overtone spectra into account. In contrast to this, the approach presented here (Sect. 3) is simply based on the periodicity of chords. Applying this theory to common musical chords and also scales (Sect. 4), shows a very good correlation to empirical results, that e.g. most subjects prefer major to

minor chords. Finally, we will highlight the psychophysical basis of the proposed approach, by reviewing some recent results from neuro-science on periodicity detection of the brain, and end up with conclusions (Sect. 5).

1.2 Basic Musical Notions

Before we are able to address the problem of harmony perception, we should clarify the terminology we use. For this, we follow the lines of [2]. The basic entity we have to deal with is a *tone*: A *pure* tone is a tone with a sinusoidal waveform. It has a specific pitch, corresponding to its perceived *frequency* f, usually measured in Hertz (Hz), i.e. periods per second. In practice, pure tones almost never appear. The tones produced by real instruments like strings, tubes, or the human voice have harmonic or other *overtones*. The frequencies of harmonic overtones are integer multiples of a fundamental frequency f. For the frequency of the n-th overtone ($n \geq 1$), it holds $f_n = n \cdot f$, i.e. $f_1 = f$. The amplitudes of the overtones define the *spectrum* of a tone or sound and account for its loudness and specific timbre.

A *harmony* in an abstract sense can be identified by a set of tones forming an interval, chord, or scale. Two tones define an *interval*, which is the distance between two pitch categories. The most prominent interval is the *octave*, corresponding to a frequency ratio of 2/1. Since the same names are assigned to notes an octave apart, they are assumed to be *octave equivalent*. An octave is usually divided into 12 semitones in western music, corresponding to a frequency ratio of $\sqrt[12]{2}$ in equal temperament (cf. Sect. 3.3). Thus, intervals may also be defined by the number of semitones between two tones. A *chord* is a complex musical sound comprising three or more simultaneous tones, while a *scale* is a set of musical notes, whose corresponding tones usually sound consecutively. Both can be identified by the numbers of semitones in the harmony.

A *triad* is a chord consisting of three tones. Classical triads are built from major and minor thirds, i.e., the distance between successive pairs of tones are 3 or 4 semitones. For example, the major triad consists of the semitones $\{0, 4, 7\}$, which is the *root position* of this chord. An *inversion* of a chord is obtained by transposing the currently lowest tone by an octave. Fig. 1 (a) shows the three inversions of the E major chord, including the root position. Fig. 1 (b)–(e) shows all triads that can be build from thirds including their inversion, always with e' as lowest tone. Fig. 1 (f) shows the suspended chord, built from perfect fourths (5 semitones). Its last inversion, consisting of the semitones $\{0, 5, 10\}$, reveals this.

(a) triads   (b) major   (c) minor   (d) diminished   (e) augmented   (f) suspended
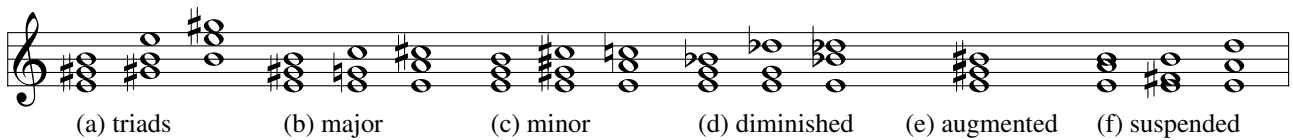
Figure 1. Triads and their inversions.

2. THEORIES ON HARMONY PERCEPTION

Chord classes lead to different musical modes. The major chord is often associated with emotional terms like *happy*, *strong*, or *bright*, and, in contrast to this, the minor chord with terms like *sad*, *weak*, or *dark*. Empirical results (see e.g. [4]) reveal a preference ordering on the perceived sonority of the triads as follows: major $\prec$ minor $\prec$ diminished $\prec$ augmented. Since all these triads are built from thirds, thirds do not provide an explanation of this preference ordering on its own. Therefore, let us now review existing theories on harmony perception, discussing some of their merits and drawbacks.

2.1 Explanation by Overtones

Overtones can explain the origin of the major triad and hence its high perceived sonority. The major triad appears early in the sequence, namely overtones 4, 5, 6 (root position) and —even earlier— 3, 4, 5 (second inversion). But it is well-known, that overtones fail to explain the origin of the minor chord.

2.2 Dissonance and Tension

Since the origin of harmony and scales cannot be explained well by overtones, newer explanations base upon the notions of dissonance [2,5] and tension [6]. In general, *dissonance* is the opposite to consonance, meaning how well tones sound together. Although this approach correlates better to the empirical results on harmony perception, it does not explain the low perceived sonority of the diminished or the augmented triad, which are built from two minor or major thirds, respectively. Therefore, [6] adopts the argument from psychology that neighboring intervals of equivalent size are instable and produce a sense of tonal tension, that is resolved by pitch changes leading to unequal intervals. Since lowering any tone in an augmented triad by one semitone leads to a major triad and raising to a minor triad, [6] assumes sound symbolism, where the major triad is associated with social strength and the minor triad with social weakness. But on the contrary, a minor triad becomes a major triad by raising the third. In addition, it is unclear whether suspended triads, built from two perfect fourths, also have a low perceived sonority. Finally, most of the empirical experiments on harmony perception present only single chords to the tested subjects. This means, there is actually no pitch movement at all.

3. A PERIODICITY-BASED THEORY

The approaches discussed so far more or less take the frequency spectrum of a sound as their starting point. Obvi-
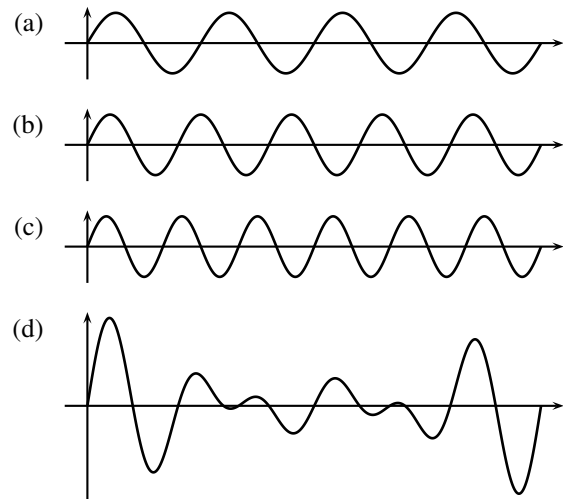

Figure 2. Sinusoids of the major triad.

ously, analyzing the frequency spectrum is closely related to analyzing the time domain (periodicity). Fourier transformation allows to translate between both mathematically. However, subjective pitch detection, i.e., the capability of our auditory system to identify the repetition rate (periodicity) of a complex tone sensation, only works for the lower but musically important frequency range up to about 1.500 Hz [3]. In consequence, a *missing fundamental* tone can be assigned to each interval. The tone with the respective frequency, called *virtual pitch* of the interval, is not present as an original tone component. It has nothing to do with (first-order) beats and is perceived not directly in the ear, but in the brain.

3.1 Periodicity Pitch of Chords

For intervals, i.e. two tones, the concept of virtual pitch has been studied many times in the literature (see [3] and references therein). The idea in this paper now is to transfer this concept to chords by considering *relative periodicity*, i.e. the period length of complex sinusoids relative to the period length of the frequency of the lowest tone component (cf. [7, Sect. 7.1]). For example, the *A* major triad in just intonation consists of three tones with (absolute) frequencies $f_1 = 440\,\text{Hz}$, $f_2 = 550\,\text{Hz}$, and $f_3 = 660\,\text{Hz}$. The respective frequency ratios wrt. the lowest tone (a') are $F_1 = 1/1$, $F_2 = 5/4$ (third), and $F_3 = 3/2$ (fifth), corresponding to the semitones $\{0, 4, 7\}$. Fig. 2 (a)–(c) show the sinusoids for the three pure tone components and Fig. 2 (d) their superposition, i.e. the graph of the function $\sin(\omega_1 t) + \sin(\omega_2 t) + \sin(\omega_3 t)$, where $\omega_i = 2\pi f_i$ are the respective angular frequencies, and t is the time.

As one can see, the period length of the chord is (only)

four times the period length of the lowest tone for this example. In the following, we call this ratio h. It depends on the frequency ratios $\{a_1/b_1, \ldots, a_k/b_k\}$ of the given chord. We assume, that each frequency ratio F_i is a fraction a_i/b_i (in its lowest terms), because otherwise no finite period length can be found in general, and it holds $F_i \approx f_i/f_1$ for $1 \leq i \leq k$. This means, all frequencies are relativized to the lowest frequency f_1, and $F_1 = 1$. The value of h then can be computed as $\mathsf{lcm}(b_1, \ldots, b_k)$, i.e., it is the *least common multiple* (lcm) of the denominators of the frequency ratios. This can be seen as follows: Since the relative period length of the lowest tone $T_1 = 1/F_1$ is 1, we have to find the smallest integer number that is an integer multiple of all relative period lengths $T_i = 1/F_i = b_i/a_i$ for $1 < i \leq k$. Since after a_i periods of the i-th tone, we arrive at the integer b_i, h can be computed as the least common multiple of all b_i.

3.2 A Hypothesis on Harmony Perception

We now set up the following hypothesis on harmony perception: The perceived sonority of a chord, called *harmonicity* in this context, decreases with the value of h. For the major triad in root position we have $h = 4$ (see above), which is quite low. Therefore, its predicted sonority is high. This correlates well to the empirical results, in general better than the approaches discussed in the previous section (Sect. 2), as we will see later on (in Sect. 4). In addition, the periodicity-based theory presented here is computationally simple, because it needs no assumptions on parameters, such as harmonic overtone spectra. Neither complex summation nor computing local extrema is required. Only the frequency ratios of the tone components in the chord are needed as input parameters. But we still have to answer the question, which frequency ratios should be used in the computation of h. Since this is done in a special way here, we present this now in more detail.

3.3 Tuning and Frequency Ratios

The frequencies for the k-th semitone in *equal temperament* with twelve tones per octave can be computed as $f_k = \sqrt[12]{2}^k \cdot f_1$, where f_1 is the frequency of the lowest tone. The respective frequency ratios are shown in Tab. 1 (a). The values grow exponentially and not linearly, following the *Weber-Fechner law* in psychophysics, which says that, if the physical magnitude of stimuli grows exponentially, then the perceived intensity grows only linearly. In equal temperament, all keys sound equal. This is essential for playing in different keys on one instrument and for modulation, i.e. changing from one key to another within one piece of music. Since this seems to be universal, at least in western music, we will adopt the equal temperament as *reference system* for other tunings.

The frequency ratios in equal temperament are irrational numbers (except for the ground tone and its octaves), but for periodicity detection they must be fractions, as mentioned above. Let us thus consider other tunings with rational frequency ratios. The oldest tuning with this property is probably the *Pythagorean tuning*, shown in Tab. 1 (b). Here, frequency relationships of all intervals have the form $3^m/2^n$ for some integers m and n, i.e., they are based on fifths, strictly speaking, a stack of perfect fifths (frequency ratio $3/2$), applying octave equivalence. However, although huge numbers appear in the numerators and denominators of the fractions in Pythagorean tuning, the relative errors compared to equal temperament (shown in brackets in Tab. 1) grow up to more than 1%.

In fact, the Pythagorean tuning does not follow results of psychophysics, namely that human subjects can distinguish frequency differences for pure tone components only up to a certain resolution, namely 0.5% under optimal conditions. For the musically important low frequency range, especially the tones in (accompanying) chords, this so-called *just noticeable difference* is worse, namely only below about 1% [3]. Therefore, we should look for tunings, where the relative error is approximately 1%. In addition, the frequency ratios should be simple integer ratios, i.e. fractions with small numerators and denominators. In order to achieve the latter, we can look in the harmonic overtone sequence, when a tone of the chromatic scale appears for the first time, applying again octave equivalence. The result of this procedure, which we will call *overtonal tuning*, leads to frequency ratios of the form $m/2^n$ for some integers m and n as shown in Tab. 1 (c). However, as one can see, the relative error compared to equal temperament again is sometimes high.

In the literature (see e.g. [5] and references therein), other historical and modern tunings are listed, e.g. *Kirnberger III*, see Tab. 1 (d). However, they are also only partially useful in this context, because they do not take into account the fact on just noticeable differences explicitly. In principle, this also holds for the adaptive tunings in [5], where simple integer ratios are used and scales are allowed to vary. An adaptive tuning can be viewed as a generalized dynamic *just intonation*, which fits well to musical practice, because the frequencies for one and the same pitch category may vary significantly during the performance of a piece of music. Trained musicians try to intonate e.g. a perfect fifth with the frequency ratio $3/2$, and listeners are hardly able to distinguish this frequency ratio from others that are close to the value in equal temperament, namely $\sqrt[12]{2}^7 \approx 1.498$. In consequence, also the *rational tuning*, which we introduce now, primarily should not be considered as a tuning, but more as the basis for intonation and perception of intervals. We will use the frequency ratios of the rational tuning, shown in Tab. 1 (e), in our analyses of harmonicity. They are fractions with smallest possible denominator, such that the relative error wrt. equal temperament is just below 1%. They can be computed by means of *Farey sequences*, i.e. ordered sequences of completely reduced fractions between 0 and 1 which have denominators less than or equal to some (small) n, or by continued fraction expansion.

3.4 Continued Fraction Expansion

In mathematics, a (regular) *continued fraction* is an expression as shown in Fig. 3 (a), where the c_i are integer numbers that must be positive for $i > 0$. For a given rational or

interval	k	(a) equal temperment	(b) Pythagorean		(c) overtonal		(d) Kirnberger III		(e) rational	
prime, unison	0	1.000	1/1	(0.00%)	1/1	(0.00%)	1/1	(0.00%)	1/1	(0.00%)
minor second	1	1.059	$3^7/2^{11}$	(0.79%)	17/16	(0.29%)	25/24	(−1.68%)	16/15	(0.68%)
major second	2	1.122	9/8	(0.23%)	9/8	(0.23%)	9/8	(0.23%)	9/8	(0.23%)
minor third	3	1.189	$3^9/2^{14}$	(1.02%)	19/16	(−0.14%)	6/5	(0.91%)	6/5	(0.91%)
major third	4	1.260	81/64	(0.45%)	5/4	(−0.79%)	5/4	(−0.79%)	5/4	(−0.79%)
perfect fourth	5	1.335	$3^{11}/2^{17}$	(1.25%)	21/16	(−1.67%)	4/3	(−0.11%)	4/3	(−0.11%)
tritone	6	1.414	$3^6/2^9$	(0.68%)	23/16	(1.65%)	45/32	(−0.56%)	17/12	(0.17%)
perfect fifth	7	1.498	3/2	(0.11%)	3/2	(0.11%)	3/2	(0.11%)	3/2	(0.11%)
minor sixth	8	1.587	$3^8/2^{12}$	(0.91%)	25/16	(−1.57%)	25/16	(−1.57%)	8/5	(0.79%)
major sixth	9	1.682	27/16	(0.34%)	27/16	(0.34%)	5/3	(−0.90%)	5/3	(−0.90%)
minor seventh	10	1.782	$3^{10}/2^{15}$	(1.14%)	7/4	(−1.78%)	16/9	(−0.23%)	16/9	(−0.23%)
major seventh	11	1.888	243/128	(0.57%)	15/8	(−0.68%)	15/8	(−0.68%)	15/8	(−0.68%)
octave	12	2.000	2/1	(0.00%)	2/1	(0.00%)	2/1	(0.00%)	2/1	(0.00%)

Table 1. Table of relative frequencies for different tunings.

(a)
$$x \approx c_0 + \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cfrac{1}{c_3 + \cfrac{1}{\ddots}}}}$$

(b) $c_0 = \lfloor x \rfloor \quad c_n = \lfloor 1/x_{n-1} \rfloor$
    $x_0 = x - c_0 \quad x_n = 1/x_{n-1} - c_n$

(c) $a_{-1} = 1 \quad a_0 = c_0 \quad a_{n+1} = a_{n-1} + c_{n+1}a_n$
    $b_{-1} = 0 \quad b_0 = 1 \quad b_{n+1} = b_{n-1} + c_{n+1}b_n$

Figure 3. Continued fractions and Euclidean algorithm.

real number x, the values c_i can be computed recursively by the (extended) *Euclidean algorithm*, stated in Fig. 3 (b), where the floor function $\lfloor x \rfloor$ is used, which yields the largest integer less than or equal to x. The sequence of the c_i induces a sequence of fractions a_i/b_i, called convergents or fraction expansion of x, which can be computed by the equations in Fig. 3 (c). Continued fractions obey many interesting properties (see [8]), for instance:

- Any finite continued fraction represents a rational number.

- Every convergent a_i/b_i of a continued fraction is in its lowest terms, i.e. , a_i and b_i have no common divisors.

- Each convergent is nearer to x than the preceding convergent and also than any other fraction whose denominator is less than that of the convergent.

The most important property in this context is the last one, because it provides a procedure for computing the frequency ratios of the rational tuning as follows. For the k-th semitone, we consider the fraction expansion of $x = \sqrt[12]{2}^k$, i.e. the frequency ratio in equal temperament, until the relative error of the convergent $y = a_n/b_n$ wrt. x, i.e. the term $|y/x - 1|$, is less than 1%.

Continued fractions may help us explain the origin of the chromatic twelve-tone scale. For this, we look for a tuning in equal temperament with n tones per octave, such that the perfect fifth in just intonation (frequency ratio $3/2$) is approximated as good as possible. Thus, we develop a fraction m/n with $2^{m/n} \approx 3/2$, where m is the number of the semitone representing the fifth. Hence, we have to approximate $x = \log_2(3/2) \approx 0.585$. In this case, the sequence of convergents is $0/1$, $1/1$, $1/2$, $3/5$, $7/12$, $24/41$, $31/53, \ldots$, showing $m/n = 7/12$ as desired, because semitone $m = 7$ gives the perfect fifth in the chromatic scale with $n = 12$ tones per octave.

4. APPLICATION OF THE THEORY

4.1 Comparison of Different Approaches

Let us now apply the periodicity-based theory to common musical chords and correlate the obtained results with empirical results. Tab. 2 shows the perceived and computed relative sonority of basic chord classes (cf. Fig. 1). Tab. 2 (a) shows the ranking for the perceived sonority according to empirical experiments reported in [4], which have been repeated by many others with similar results. Unfortunately, [4] does not consider the suspended triad. Therefore, it is not ranked in the table. Tab. 2 (b) provides the ranking for *complex tonalness* [2], whose numerical values are shown in brackets. The model according to [2] builds on earlier work [9]. However, especially the dissonance of the augmented triad is not reflected in this model by its calculated tonalness: It appears on rank 2, right after the major triad in root position. Therefore, [2] argues, that this has cultural rather than sensory origin. Tab. 2 (c) shows the ranking wrt. *instability* [6]. The notion of tension used in this model produces the desired low sonority of the diminished and the augmented triad (cf. Sect. 2.2). The correlation with the empirical results is good, but can still be improved, e.g., the minor triad in root position (rank 2) scores better than the inversions of the major triad (ranks 4 and 5), which is not as desired.

Tab. 2 (d)–(e) shows the ranking wrt. the harmonicity values h. As one can see, there is almost a one-to-

chord class		(a) empirical [4]	(b) tonality [2]		(c) instability [6]		(d) harmonicity		(e) harmonicity*	
major	$\{0,4,7\}$	1	1	(0.48)	1	(0.624)	2	(4)	2	(4.0)
	$\{0,3,8\}$	2	6	(0.38)	5	(0.814)	3	(5)	3	(5.0)
	$\{0,5,9\}$	3	3	(0.43)	4	(0.780)	1	(3)	1	(3.0)
suspended	$\{0,5,7\}$				8	(1.175)	4	(6)	4	(6.0)
	$\{0,2,7\}$				11	(1.219)	5	(8)	5	(8.0)
	$\{0,5,10\}$				9	(1.191)	6	(9)	6	(9.0)
minor	$\{0,3,7\}$	4	4	(0.42)	2	(0.744)	7	(10)	7	(10.0)
	$\{0,4,9\}$	5	7	(0.38)	3	(0.756)	8-9	(12)	8	(12.0)
	$\{0,5,8\}$	6	10	(0.32)	6	(0.838)	10-11	(15)	9	(15.0)
diminished	$\{0,3,6\}$	7	9	(0.35)	12	(1.431)	13	(60)	13	(26.0)
	$\{0,3,9\}$	8	5	(0.40)	7	(1.114)	10-11	(15)	10	(16.6)
	$\{0,6,9\}$	9	8	(0.37)	10	(1.196)	8-9	(12)	12	(19.9)
augmented	$\{0,4,8\}$	10	2	(0.44)	13	(1.998)	12	(20)	11	(19.7)

Table 2. Ranking relative sonorities of common triads.

one correspondence with the empirical results. The numbers in brackets are the respective harmonicity values h and h^*, where the latter are averaged over all inversions. For this, we compute the harmonicity of the given chord (cf. Sect. 3), e.g. the first inversion of the diminished triad $\{0,3,9\}$, that is $h_0 = \mathsf{lcm}(1,5,3) = 15$. In addition, we adopt each tone as reference tone, not only the lowest tone. Thus, we consider also the chords with the semitones $\{-3,0,6\}$ and $\{-9,-6,0\}$. For semitones associated with a negative number n, we take the frequency ratio of semitone $12-n$ according to Tab. 1 (e) and halve it, i.e., we do not apply octave equivalence here. Therefore, we get the frequency ratios $\{5/6, 1/1, 17/12\}$ and $\{3/5, 17/24, 1/1\}$ with harmonicity values $h_1 = 12$ and $h_2 = 120$, respectively. Since periodicity of chords is related to the lowest tone, we multiply the h values by the lowest frequency ratio in the chord, obtaining $h_0' = 15$, $h_1' = 5/6 \cdot 12 = 10$, and $h_2' = 3/5 \cdot 120 = 72$. We then average the virtual chord frequencies f_1/h, where h appears in the denominator. Hence, we calculate the harmonic average of all harmonicity values h_0', h_1', and h_2', which yields $h^* \approx 16.6$.

Tab. 2 (a) and (e) differ only in two respects: First, the most consonant chord according to harmonicity (rank 1) is the second inversion of the major triad with semitones $\{0,5,9\}$ and not the root position. Its calculated harmonicity is $h = 3$, which however coincides with the fact, that the second inversion appears before the root position in the harmonic overtone sequence (cf. Sect. 2.1). Second, the augmented triad appears late as expected (rank 11 of 13), but the root position and the second inversion of the diminished triad appear still later. However, the continued fraction expansion for the tritone (semitone 6, frequency ratio $\sqrt{2}$), occurring in both triads, yields first $7/5$, which is only slightly mistuned. This would lead to a significantly lower h value of the two chords – as desired. Thus, in summary, the periodicity-based approach on harmony perception fits best to empirical results.

4.2 Overtones and Periodicity

Harmonic overtone spectra are irrelevant for determining relative periodicities. The period length of such complex waveforms is identical with that of its fundamental tone. We obtain $h = 1$, since the frequencies of harmonic overtones are integer multiples of the fundamental frequency, hence all frequency ratios $\{1/1, 2/1, 3/1, \dots\}$ have 1 as denominator. Therefore, harmonicity is independent from concrete amplitudes and phase shifts of the sinusoids of the pure tone components. This seems plausible, because harmony perception only partially depends on loudness and timbre of the sound. It should not matter much, whether a chord is played e.g. on guitar, piano, or pipe organ. Of course, this argument only holds for tones with harmonic overtone spectra. If we have inharmonic overtones in a complex tone such as in *gamelan music* (cf. [5]), then it holds $h > 1$ for the harmonicity value of a single tone, i.e., we have an inherently increased *harmonic complexity* (cf. [2]).

4.3 From Chords to Scales

The harmonicity value h can be determined for harmonies, consisting of far more than three tones, without any computational problems. Thus, let us apply the formulae from Sect. 3 to general chords and scales. Fig. 4 (a)–(b) shows harmonies with 5 tones, that have low h values. The pentachord Emaj7/9 with $h = 8$, classically built from a stack of thirds, is standard in jazz music. Alternatively, it may be understood as superposition of the major triads E and B, which are in a tonic-dominant relationship according to classical harmony theory. Fig. 4 (b) shows the pentatonic scale ($h = 24$), which could alternatively be viewed as the standard jazz chord E6/9. All harmonies shown in Fig. 4 have low, i.e. good harmonicity values h, ranking among the top 5% in their tone multiplicity category. This also holds for the diatonic scale (7 tones, $h = 24$) and the blues scale (8 tones, $h = 24$) in Fig. 4 (c)–(d). Furthermore, according to their h^* value, all church modes, i.e. the diatonic scale and its inversions, rank among the top 11 of 462 possible scales with 7 tones. Therefore, the periodicity-based theory can contribute significantly to the discussion about the origin of scales of western music. There are other mathematical explanations for the origin of scales, e.g. by group theory [10], ignoring however the sensory psychophysical

Figure 4. Harmonies (scales) with more than three tones.

basis for the musical importance of the perfect fifth.

5. CONCLUSIONS

As we have seen in this paper, harmony perception can be explained well by considering relative periodicities of chords, that can be computed from the frequency ratios of the intervals in the chord. The approach shows a good correlation to empirical studies on perceived sonority. Even the origin of scales can be described with this approach. It is mathematically simple, employing Farey sequences or the Euclidean algorithm for computing continued fractions. The approach has a strong psychophysically basis. It takes into account that human pitch perception is limited by a just noticeable difference of about 1% and assumes that virtual pitch of chords (chord periodicity) can be detected. The latter is indeed possible, as results from neuro-science prove, which we briefly review now.

5.1 Periodicity and Neuro-Science

From a spectral point of view, sounds are combinations of a fundamental frequency and certain overtones. Spectral analysis is performed in the cochlea. When a pure tone is detected, waves travel along the basilar membrane, which the cochlea houses, reaching a maximum amplitude at a point depending on the frequency of the tone [1–3]. Thus, the ear works as a spectral analyzer. This function of the ear is used in the explanations of harmony perception, based on overtones or dissonance (Sect. 2).

Periodicity-based explanations use missing fundamental tones, i.e. tones that are physically not present and hence cannot perceived by the ear directly. It has been well-known for years that periodicity can be detected in the brain. For example, two pure tones forming a mistuned octave cause so-called *second-order beats*, although no exact octave is present [3]. Recently, neuro-science found the mechanism for being able to perceive periodicity. As a result of a combined frequency-time analysis, i.e. some kind of *auto-correlation* by comb-filtering, pitch and timbre are mapped temporally and also spatially and orthogonally to each other in the auditory midbrain and auditory cortex [1] (see also [11]). [12] reviews neuro-physiological evidence for interspike interval-based representations for pitch and timbre in the auditory nerve and cochlear nucleus. Timings of discharges in auditory nerve fibers reflect the time structure of acoustic waveforms, such that the interspike intervals (i.e. the period lengths) that are produced convey information concerning stimulus periodicities, that are still present in short-term memory [1].

5.2 Summary and Open Questions

From the good correlation of the periodicity-based theory with the empirical results presented here, one may conclude, that there is a strong psychophysical basis for harmony perception and the origin of musical scales. As underlying principle for this, periodicity detection turns out to be more important than spectral analysis, although cultural and other aspects certainly must not be neglected. The question, how different harmonies cause different emotions or subjective effects like happiness or sadness is not yet answered by this, of course.

6. REFERENCES

[1] Gerald Langner. Die zeitliche Verarbeitung periodischer Signale im Hörsystem: Neuronale Präsentation von Tonhöhe, Klang und Harmonizität. *Zeitschrift für Audiologie*, 46(1):8–21, 2007.

[2] Richard Parncutt. *Harmony: A Psychoacoustical Approach*. Springer, Berlin, Heidelberg, New York, 1989.

[3] Juan G. Roederer. *The Physics and Psychophysics of Music: An Introduction*. Springer, Berlin, Heidelberg, New York, 4th edition, 2008.

[4] L. A. Roberts. Consonant judgments of musical chords by musicians and untrained listeners. *Acustica*, 62:163–171, 1986.

[5] William A. Sethares. *Tuning, Timbre, Spectrum, Scale*. Springer, London, 2nd edition, 2005.

[6] Norman D. Cook and Takashi X. Fujisawa. The psychophysics of harmony perception: Harmony is a three-tone phenomenon. *Empirical Musicology Review*, 1(2), 2006.

[7] James Beament. *How we hear music: The relationship between music and the hearing mechanism*. The Boydell Press, Woodbridge, UK, 2001.

[8] Carl D. Olds. *Continued Fractions*, volume 9 of *New Mathematical Library*. L.W. Singer Company, 1963.

[9] Ernst Terhardt, Gerhard Stoll, and Manfred Seewann. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 71(3):679–688, 1982.

[10] Gerald J. Balzano. The group-theoretic description of twelvefold and microtonal pitch systems. *Computer Music Journal*, 4(4):66–84, 1980.

[11] Ray Meddis and Michael J. Hewitt. Virtual pitch and phase sensivity of a computer model of the auditory periphery: I. Pitch identification, II. Phase sensivity. *Journal of the Acoustical Society of America*, 89(6):2866–2894, 1991.

[12] Peter A. Cariani. Temporal coding of periodicity pitch in the auditory system: An overview. *Neural Plasticity*, 6(4):147–172, 1999.