The axiom system of classical harmony

Tóbiás András

October 2, 2014

The main goal of this article is to provide the mathematical axiomatization of the strictly homophonic four-part model of classical harmony. This model is mainly based on J. S. Bach's four-part chorales, and almost unequivocal, prescriptive compositional principles were given for it in the second half of the 18th century. To understand this procedure itself one has to be able to follow Mathematics on the level of a Bachelor's course, but the results of the formalization can be directly applied by musicians. A logical ordering of the basic notions in harmony can also be used for a new highschool music theory coursebook, which seems to be necessary in Hungary. This motivated me to write my Bachelor's thesis about this topic at TU Budapest[1]. About my work I also consulted with Prof. Robert S. Sturman and Prof. Scott McLaughlin at the University of Leeds (UK), where a *Mathematics and Music* Bachelor's course takes place.

The main steps and new results of our axiomatization work have been the following:

1. Deduction of the most important basic notions in music theory which are closely related to harmonics, such as the overtone system, the equal-tempered piano, the enharmonic equivalence, the musical intervals, triads and seventh chords. We applied the mathematical-physical axiomatization results of Dave Benson ([1]) about overtones and chords for this part.

2. Defining the musical key from the perspective of functional tonality. According to this, we have been able to prove that there is exactly one key type different from the traditionally well-known two types, the major and the minor.

3. Defining a strictly four-part piece using the topological properties of the real line. Here we use the mathematical thoughts of the Hungarian music theorist and composer Ligeti György ([2]). This model of four-part edition gives an opportunity to define the genre of Bach's chorals mathematically precisely.

4. After giving the definitions of musical functions and tonality, proving our most important proposition, the so-called *fundamental theorem of tonality* about the relation between tonality and compliance with classical harmonics, in the strictly four-part case.

5. Providing the usual chord-changing compositional principles (e.g. the principle of least motion; princile of keeping the common voice; forbiding paralel octaves, paralel fifths and augmented second steps etc.) embedded in our axiom system of classical harmony. This is closely related to the research of Dmitri Tymoczko regarding the principle of least motion in a wider context ([6]). The compositional principles are given following the Hungarian music theory coursebook of Kesztler Lörinc ([3]).

6. Standing a simplified model for strictly four-part pieces for negotiating *modulations* complying with classical harmonics. Providing the modulational principles of classical harmony (still according to [3]).

[1] T.: *A klasszikus összhangzattan axiomatikája*, 65p. Budapest University of Technology and Economics, 2014. Thesis advisor: Dr. G. Horváth Ákos, associate professor, leader of the Faculty of Geometry.
According to the goal of preparing for a music theory coursebook in Hungarian, my thesis follows the conventional Hungarian–German notation and treatment system of classical harmony, and not the English one. In this paper we will use the English notation wherever possible, although the perspective and therefore also some definitions (e.g. of the tonic, subdominant and dominant function) come from Hungarian music theory education. However, this article does not only extract or sum up the methods and the main results of my degree work, but shows up some further results and generalisations, e.g. the fact that the considered musical phenomena can be directly derived from the ZFC axiom system (page 2) or the notion of *successive pauseless extension* (page 20).

Contents

1 Basic notions of music theory: overtones, equal-tempered piano, enharmonic equivalence and consonance

Our exact goal is to provide the axiom system for composing four-part classical harmony examples, in a first-order language. We will not construct a new language but use the one in set theory, supposing the ZFC axiom system. We will use simple physical properties of the overtone system, but formally these only have arithmetic meaning. When we consider *a tone X with frequency $f(X)$*, it can be handled as $X = (f(X), f(X)) \in (\mathbb{R}^+)^2$, these way the exact definitions of this paper can be derived consequently from ZFC .

During the whole article let $B_r(x)$ denote the open ball with radius r around the point x in an arbitrary metric space, and $\overline{A}$ the closure of A in an arbitrary topological space.

As usual in music theory, *tone* Y is a longitudinal wave moving in an elastic medium with frequency $f(Y) > 0$. For a tone X with frequency $f(X) > 0$, X is *audible* if 20 Hz $< f(X) <$ 20 000 Hz. When talking about tones, we always mean that the tone consists of all of its *overtones*. The set of overtones of the tone X is $\{Y|\ Y \text{ is a tone}, \exists n \in \mathbb{N}^+ :\ f(Y) = n\ f(X)\}$. The overtone of X with frequency $n\ f(X)$ is called the nth overtone of X.

Musical *intervals* are equal distances in the (base 2) logarithmic frequency scale. The most important intervals can be derived from the overtone system. The interval of a tone and its 2nd overtone is called *perfect octave*, the one of a tone's 2nd and 3rd overtone is called *perfect fifth*, the one of a tone's 3rd and 4th overtone is called *perfect fourth*, the one of a tone's 4th and 5th overtone is called *major third* and the one of a tone's 5th and 6th overtone is called *minor third*. The interval of a tone and itself is called *perfect prime*.

If an interval of tones X and Y ($f(X) < f(Y)$) is not greater than an octave, then its *inverse interval* is the interval completing it to a whole octave, this is the interval of Y and X's second overtone. It can easily be seen that the inverse property is a symmetric relation amongst tones. The perfect octave and the perfect prime are the inverses of each other, so are the perfect fifth and the perfect fourth. The inverse of major third is called *minor sixth* and the inverse of minor third is *major sixth*.

We say that the tone X is higher than the tone Y, or Y is lower than X if $f(X) > f(Y)$. Intervals can be summed and hence we can talk about *octave-equivalent* tones X and Y, the interval of which is n octaves, where $n \in \mathbb{Z}$. If X is n octaves higher than Y, it means that $f(X) = 2^n f(Y)$. It is obvious that octave-equivalence is an equivalence relation on the set of tones, and therefore we can consider *the octave-eqivalency class* of the tone X, we will denote it by $[X]$.

On the following figure[2] see the first 11 overtones of the tone C on the piano[3]. See that the 7th and 11th overtones are not members of the C major scale, they will have specific roles at *modulations*, i.e. changings between musical keys with different bases[4].Now we give the basic definitions for these relations.

Let X be a tone with its 3rd overtone Y. The *leading tone* of $[Y]$ is the octave-equivalency class of X's 11th overtone. The *seventh tone belonging to* $[X]$, by other words the *upper leading tone of X's fifth overtone's equivalency class* is the octave-equivalency class of Y's 7th overtone. These properties can be also defined for the tones themselves. If we consider the tones $U \in [U]$ and $V \in [V]$ and, for example, $[U]$ is the leading tone of $[V]$, then we say that U is the leading tone of V.

We define the *perfect X major scale* for a tone X. Generally, a seven-degree scale with base X_1 is a set of tones $\{X_1, X_2, \ldots, X_7$ where $f(X_i) > f(X_j) \Leftrightarrow i > j$ and $f(X_7) < 2f(X_1)$ (this is, every member of the scale is strictly less then one octave higher than the base). X_i is called the ith degree scale tone of the scale[5]. This means that originally we denote the degrees and the operations among them with the elements of the primefield $\mathbb{Z}_7$, but we use the capital roman numeral for the integer $(n \pmod 7)) + 1$ instead of $n \in \mathbb{Z}_7$. The *perfect X-major scale* is a seven degree scale with base X, where the frequency ratios of the neighbouring degree tones are respectively:

$$\frac{9}{8}, \frac{10}{9}, \frac{16}{15}, \frac{9}{8}, \frac{10}{9}, \frac{9}{8}, \frac{16}{15},$$

where the last ratio is the ratio of the VIIth degree scale tone and the second (one octave higher) overtone of X.

Observation 1.1. Consider a perfect major scale with all elements in the audible area for the human ear. Then it can be observed that for the audience the following are perceptually true about the scale tones:

(i) the degree VII is the leading tone of the degree I, III is the leading tone of IV,

(ii) IV is the upper leading tone of III, I is the upper leading tone of VII,

(iii) the interval between IV and I is a perfect fifth, the one of I and V is also,

(iiii) IV is the seventh tone belonging to I, and I the one belonging to V (see on page 3).

Consider the sum of *twelve perfect fifths* and the one of *seven perfect octaves* from a tone X. The first interval results a tone with frequency 128 $f(X)$, while the second one gives a tone with frequency $\frac{531441}{4096} f(X)$. The difference of the two tones is noticeable by an average person. However, if it is equally spread along the whole interval, it locally cannot be percieved. Having 12 quasi-fifths perceptually equal to 7 octaves on a musical instrument such that each tone of it is a base of a seven-degree scale perceptually equivalent to a perfect major scale could fease the concept of *the circle of fifths*. We could use the leading tone and upper leading tone/seventh tone connections between the quasi-perfect major scales to make it possible to move

[2]The figures come from the appendix of the Hungarian paper and were made using Lilypond.

[3]This paragraph is only for musical illustration, all the new notions in it will be defined precisely in the following pages.

[4]Note that these overtones are out of tune in equal temperament. The frequency of the seventh overtone of C is 7 times as many as the frequency of C, while the frequency of its enharmonic equivalent on the equal-tempered piano, $b\flat^1$, is $2^{2+\frac{10}{12}} = 7,1272$ times as many as the one of C. Similarly, the frequency of the 11th overtone of C is 11 times as many as the one of C, while the frequency of its enharmonic equivalent, $f\sharp^2$, is $2^{3+\frac{6}{12}} = 8\sqrt{2} = 11,3137$ times as many as the one of C. Therefore if we started from C on the equal-tempered piano, and used these actual overtones for modulating to G, we would get a G significantly different from G of the equal-tempered scale.

[5]According to the Hungarian denotation, we will use capital roman numerals for the degrees.

from each major scale to the two with a base one quasi-perfect fifth higher and lower. The one perfect fifth higher scale is called the *dominant* scale, while the one perfect fifth lower is the *subdominant* scale. This is the basic idea for *equal-tempered piano*[6].

Definition 1.1. K, a countable set of tones $X_0, X_1, \ldots$ is an equal-tempered piano if

(i) $A \in K$, where A is the normal a^1 tone with frequency 440 Hz,

(ii) $f(X_i) > f(X_j) \Leftrightarrow i > j$,

(iii) $\forall i \leq 0$: if $\exists X_{i+1}$, then $f(X_{i+1}) = \sqrt[12]{2} f(X_i)$,

(iiii) K has two tones X and Y the interval of which is at least 7 octaves.

Hence A is the member of every – finite and infinite – equal-tempered piano. Choosing A to be fixed, the octave-equivalency classes of the piano's white keys $(A, B, \ldots, G)$ can be given. J.S. Bach showed that every tone of the equal-tempered piano can serve as a base of a major scale, by composing his *Das Wohltemperierte Klavier*, which contains one piece for every major key of the well-tempered piano. This piano was not yet equal-tempered, it was an earlier stage in actualising the concept of the circle of fifths.

Enharmonic equivalence in the context of the 12-tone equal tempered scale, between audible tones generally means that two different tones come from two different perfect major scales, but they have a common member of the equal tempered piano from which they are not significally different for the human ear. Enharmonic equivalence therefore depends of the listener's individual properties and cultural background. In Europe, the audience of music has got used to equal temperament within three centuries and generally accepts the following two propositions.

1. If we call the interval of two neighbouring tones of the equal tempered piano a *semitone*, the sequence of tones 0, 2, 4, 5, 7, 9 and 11 semitones higher than an *arbitrary* piano tone is a good approximation of a perfect major scale.

2. If we consider X and Y, two piano tones where Y is two semitones (i.e. *one whole tone*) higher than X, then the (only) piano tone Z such that $f(X) < f(Z) < f(Y)$ is the leading tone of Y and the upper leading tone of X.

Consider the C major scale on the equal-tempered piano —it consists of the seven white keys. Move *stepwise upwards* in the circle of fifths, in the direction of the *upper circle of fifths*, reaching the G, D, A, E, B major scales respectively. In this sequence, there is always exactly one degree in the current major scale that is a non-scale tone in the following scale. The new tone is the leading tone of the first degree in the new scale. Denote this with the letter of the original C major scale tone which becomes a semitone higher in the new scale and attach a ♯ to the letter, this way you reach F♯, C♯, G♯, D♯, A♯, the leading tones to G, D, A, E, B respectively. The base of the sixth scale is also a black key which has already been denoted F♯.

Now go back to the C major scale and move *stepwise downwards*, in the direction of the *lower circle of fifths*, this way you reach the F major scale first. There is exactly one scale tone that the C major scale does not contain: instead of B, there is a semitone lower tone, the seventh tone with regard to F, by other words the upper leading tone of A. Denote it B♭. If, similarly, X♭ denotes the one semitone lower piano tone than the white key X, moving downwards one fifth by one on the equal-tempered piano, we reach F, B♭, E♭, A♭, D♭ and G♭.

Note that G♭ and F♯ refer to the same (black) piano keys: these two tones are *enharmonic* in the context of equal-tempered piano. We emphasize that they are not the same tones: if we build a *perfect* A major scale, F♯ is the VI degree scale tone there, while G♭ can be reached if we move down in the D♭ major scale with 8 *perfect* fifth steps, and take the IV degree scale tone. It is a well-known experimental result that the actual G♭ and F♯ that we receive this way are significantly different [?], but there is no difference between their actualisation on the equal-tempered piano.

[6] About why we consider exactly 12 fifths and 7 octaves, see [1, p. 190–193.]

$\sharp$ and $\flat$ marks can be multiplied. The degree n scale tone's original name with one more $\sharp$ means always the $\pmod 7$ $n+1$th degree scale tone's leading tone and the degree $n+1$ scale tone's original name with one more $\flat$ means always the nth degree scale tone's upper leading tone. Note that $\flat\sharp = \sharp\flat = \natural$ means the identity of the C major scale, and "multiplication" of $\flat$'s and $\sharp$'s is commutative.

Consequently, we will not describe the properties of clear major scales but the ones on the equal-tempered piano; a piano tone will always mean a class of enharmonic equivalency of countably infinite tones connected to the piano's C major scale by the system of $\flat$s and $\sharp$s. Usually we will not consider more than twice altered tones. Note that one major scale cannot contain sharpened and flattened tones in the same time. The fifth-by-fifth sequence of sharpened scale tones of a major scale on the equal-tempered piano: F$\sharp$, C$\sharp$, ... or the sequence of flattened scale tones of the major scale: B$\flat$, E$\flat$, ... is called the major scale's *key signature*.

After this, if X and Y are arbitrary —not by all means audible— tones of the equal-tempered piano and there is an $n \in \mathbb{Z}$ so that the audible tone X' n octaves higher than X is enharmonic to the (also audible) Y', which is n octaves higher than Y, then we call X and Y audible, too. We also can talk about enharmonic equivalence of seven-degree scales, octave-equivalency classes, intervals of an equal-tempered piano K and elements of its Cartesian product powers K^n $(n \in \mathbb{N})$ (its elements are the *chords*) too. Enharmonic equivalence of the tones A and B is denoted as $A \sim B$, and it is easy to prove that $\sim$ is an equivalence relation.

According to the construction of the twelve enharmonic equivalency classes of major scales, we can define all the musical intervals that are in use in classical harmony. In this, we always follow the international terminology (i.e. diminished/minor/perfect/major/augmented prime/first/second/.../seventh/octave), therefore we do not detail this. In the context of classical harmony, we call an interval of the equal-tempered piano *consonant* if the following two properties apply to it:

1. It is enharmonic with the piano version of the interval between the kth and the mth overtone of a tone X, where $1 \le k, m < 7$, up to octave-equivalency,
2. it is neither a diminished nor an augmented interval.

Corollary 1.1. *An interval is consonant if and only if it is n octave plus 0, 3, 4, 5, 7, 8 or 9 semitones and its name is neither diminished nor augmented.*

Note that this definition for consonance cannot be applied for earlier ages than Viennese classicism, because in those ages the conditions of consonance had been stricter, e.g. perfect fourth was considered as a dissonant interval. *Dissonant* means not consonant. Again, questions of consonance and dissonance depend on human perception and can be modified by the recipient's cultural–musical environment.

We have declared that an equal-tempered piano has to be *at least as wide as a real piano*: 7 octaves. This is for making it possible to contain 7 octaves (12 fifths). For certain parts of the observation it is more comfortable to consider an equal-tempered piano that is *infinite into both directions*. Let $d_2(X,Y)$ mean the interval between the notes X and Y of an arbitrary equal-tempered piano K, measured in semitones. (The interval consisting of two semitones is called a wholetone.) Then the following is true:

Proposition 1.1. (K, d_2) *is a metric space and* d_2 *generates the discrete topology on* K.

Proof. It is trivial that d_2 has the metric properties. Because there is no smaller d_2 distance between piano elements than 1, $\forall X \in K$ $B_1(X) \subseteq \{X\}$, which shows that the one-point set $\{X\}$ is open, therefore K is a discrete topological space. $\square$

Hence, we can talk about *Borel-measurable functions* $M : \mathbb{R}_0^+ \to K^n$, where $n > 1$ integer and K^n is a finite topological direct product $\underbrace{K \times K \times \ldots \times K}_{n \text{ instances}}$. Considering the real half line as continuous time, we call this function M *n-part piece*. More exact definitions about strict four-part composition are coming in section 4. The *kth voice* of M is $\mathrm{pr}_k \circ M$, where pr_k is the projection to the kth instance of the equal-tempered piano. In the four-part case, the voices (from the first to the fourth respectively) are called, as conventionally, *bass, tenor, alto* and *soprano*. Often in the case of more than one voice, the first voice is called bass and the last soprano.

2 Triads and seventh chords

Triad names are special elements of the factor space $K^3/\equiv$ on an arbitrary equal-tempered piano K, where $\equiv$ is the octave equivalency relation. These contain scale or once modified tones from a certain seven-degree scale on K, and their main characteristic is that they consist of a kth, a $k+2$th and a $k+4$th degree tone $\pmod 7$ of the given major scale based on one of the twelve enharmonic equivalence classes of the equal-tempered piano. With this notation we call the triad degree k. There are four kind of triad names which we say that *comply with classical harmony*, according to the following chart:

Name	Notation	Interval of the degree k and $k+2$	Interval of the degree $k+2$ and $k+4$ tones	Interval of the degree k and $k+4$
Major triad	M	*major third*	minor third	perfect fifth
Minor triad	m	*minor third*	major third	perfect fifth
Diminished triad	d	minor third	minor third	*diminished fifth*
Augmented triad	A	major third	major third	*augmented fifth*

There is a major and a diminished triad in the set of the first 7 overtones of an arbitrary tone: the 4th, 5th and 6th overtones form a major one and the 5th, 6th and 7th overtones form a diminished one. Minor and augmented triads cannot be found among the triplets of these tones. On the other hand, major and minor are the consonant triad types—because all the three intervals among their tones are consonant—while diminished and augmented are dissonant—because their fifths are not perfect and therefore dissonant.

A *four-part version of a triad* – in other words, later in this article: a *triad* – is an element of the piano power K^4 which consists of the tones of a triad name, exactly one of them in two voices. Triads with the same name are considered to be equal if and only if all of their voices are the same (and not only octave-equivalent!). If the triad consists of the kth, $k+2$nd and $k+4$th degree scale tone of a seven-degree scale – these are called the *base*, the *third* and the *fifth* of the triad, respectively – on the equal-tempered piano, its *position* is determined by which tone it has in the bass. If in the bass there is the k degree tone, the triad is in *root position* (with Hungarian notation: k), if it is the $k+2$ degree tone in the bass, the triad is in *first inversion* (k^6), and if it is the $k+4$ degree tone, the triad is in *second inversion* (k^6_4).

If the name of the triad complies with classical harmony, the duplication of this triad name's voices is appropriate[7] and there is a preliminarily fixed frequency interval (of some musical instruments or vocal voices) for each voice and the contributing voice of the triad is inside this, then we say that the triad *complies with classical harmony*. The previous sentence is true with a little correction, which we will give after defining tonality topologically, on page 16.

Consider the union of a degree k and a degree $k+2$ triad name on an arbitrary seven-degree scale. This is virtually an element of $K^4/\equiv$, and it is called *seventh chord name*. If $H \in K^4$ consists of the tones of a seventh chord name, with an arbitrary permutation of the voices, then H is called a *seventh chord*. This name comes from the fact that there is a seventh interval between the k and the $k+6$ degree scale tones. The degree k and degree $k+2$ triads are the *partial triads* of the seventh chord. The position of a degree k seventh chord inverson can be: *(root position) seventh chord* (k^7 in the Hungarian notation), *first inversion* (k^6_5), *second inversion* (k^4_3) and *third inversion* (k^2 or sometimes k^4_2), if in the bass there is the kth, $k+2$nd, $k+4$th and $k+6$th degree tone of the seventh chord name, respectively. The origin of these notations is the following. If a 7th chord does not have the kth degree tone in the bass and the $k+6$th degree tone in the soprano, then there are two voices between which the interval is a second, by other words: between these two voices there is a *second friction*. The upper and lower indices on the right of the degree's roman numeral show the interval of these two voices from the bass (e.g. in the case of k^6_5 the lower voice contributing to the friction is a fifth higher than the bass, the higher one is a sixth higher than the bass), up to octave-equivalency. The kth degree tone is the seventh chord's *base*, the $k+2$nd degree one is its *third*,

[7] This means: if the triad is in root position, the kth degree tone participates in two voices, and if the triad is in second inversion, then the $k+4$th degree. Furthermore, if the triad is diminished and it is in root position, then the $k+2$nd degree tone is not in the tenor and the $k+4$th degree tone is not in the alto (in order to avoid too strong dissonances). There is no limitation for voice duplication in the case of first inversion.

the $k+4$th degree one is its *fifth* and the $k+6$th degree one is its *seventh*. We also refer to the interval between the base and the third, fifth or seventh by the words 'third', 'fifth' or 'seventh', respectively. The name of a four-part version of a triad or a seventh chord H is signed as $\mathfrak{N}(H)$ (pronounce: *nomen H*).

If both partial triads of a seventh chord name H comply with classical harmony, and there is no triad which is voicewise enharmonic to H, then we say that H *complies with classical harmony*. This second condition is for excluding the *augmented triad* from the set of seventh chord names, which can be represented by a triad name in which there is a major third between the kth and $k+2$nd, between the $k+2$nd and $k+4$th and also between the $k+4$th and the $k+6$th degree tone, but this time the kth and $k+6$th degree tone are enharmonic on the equal-tempered piano. If a four-part version H_0 of a seventh chord H complying with classical harmony has it voices inside a preliminarily fixed frequency interval, we—similarly to the case of triads—say that H_0 complies with classical harmony. The following chart shows the seventh chord (name) types those comply with classical harmony. (Definitions of the major and the minor keys will follow in the next section, the examples for seventh chord types in the last two columns are only for musical illustration.)

Third	Fifth	Seventh	Partial triads	Name (...) *seventh*	Examples in major (degrees)	Example in minor (degree)
major	augmented	major	major, augmented	augmented major	none	III
major	perfect	major	major, minor	major minor	I, IV	VI
major	perfect	minor	major, diminished	major/dominant	V	V
minor	perfect	major	minor, augmented	harmonic minor	none	I
minor	perfect	minor	minor, major	minor major	II, III, VI	IV
minor	diminished	minor	diminished, minor	semi-diminished	VII	II
minor	diminished	diminished	diminished, diminished	diminished	none	VII

Remark: a root position dominant seventh that complies with classical harmony may be *fifth deficient*, which means that it need not contain its fifth in any voice but instead the base in two voices (one of these voices is necessarily the bass).

3 Trichotomy of musical keys

After introducing the basic musical notions we define keys. Our goal is to provide a definition of key based on the idea of key stability and tonality in classical harmony. We preliminarily ensure that it accepts the major and minor keys, which we have been using in Europe for five centuries, to be keys, but we will see that it does not exclude every other seven-degree scale. What's more, this definition gives the ability to find all possible key types, and we will see that apart from major and minor there is exactly one more kind of key.

Definition 3.1. Let H be a seven-degree scale on the equal-tempered piano (with scale tones and modified tones from the C major scale), with seven pairwise non-enharmonic scale tones. We say that H is a key if:

(i) the Vth degree triad of H (built up from scale tones) is major and the Vth degree seventh is dominant,

(ii) all triad and seventh chord names that consist of the scale tones of H comply with classical harmony,

(iii) if the kth degree seventh is dominant, then the degree $k+3 \mod 7$ triad is major or minor $\pmod 7$, with the $k+3$th degree scale tone one perfect fourth higher than the kth degree one[8].

It is obvious that the definition implies the next two properties:

Proposition 3.1. *In a key the first degree triad is major or minor, and the VIIth degree scale tone is the leading tone of the Ist degree scale tone.*

The next two are examples of easily provable propositions in the area of keys.

[8]In other words: the kth degree dominant seventh, with its seventh dissonance, can *resolve to its tonic*. This will follow in detail in section 5.

Example.

1. Consider a major scale and duplicate it one octave higher. This way you get a sequence of 15 tones – 0, 2, 4, 5, 7, 9, 11, 12, 14, 16, 17, 19, 21 and 23 semitones higher than the lowest one. The seven-degree scales of seven neighbouring elements (in frequency's order) are called *modal scales*. Show that among these only the major scale is a key.

2. Now modify these modal scales to the so-called harmonic modal scales: if the VIIth degree scale tone is not enharmonic to the leading tone of the Ist degree scale tone, substitute this VIIth degree tone with the leading tone. Show that this way we receive one more key: the harmonic Aeolian or harmonic minor, —this is the harmonic version of the Aeolian or natural minor scale— which has got the following intervals between its first degree and the other degrees: 0, 2, 3, 5, 7, 8, 11 and 12 semitones, respectively.

According to this example, we can introduce the major and minor keys. The next lemma helps us find the third key type and prove that there are no other possible types.

Lemma 3.1 (The Minor Lemma)**.** *In an arbitrary key H the following properties are equivalent:*

(i) the VIth degree scale tone is 8 semitones higher than the Ist degree one,

(ii) every type of seventh chord complying with classical harmony can be built up from scale tones of H,

(iii) the VIIth degree seventh chord (built up from scale tones) is diminished,

(iv) the IVth degree triad is minor,

(v) the IInd degree triad is diminished,

(vi) there are two neighbouring degree scale tones of H between which the interval is an augmented second.

Proof. The definition of keys implies that the sequence of intervals of the Ist degree scale tone and the other scale tones is: $(0, 2, ?, 5, 7, ?, 11)$, where the ?'s refer to unknown intervals. Using the key's definition it is easy to prove that the lemma's conditions are equivalent to the condition (i), which states that the sequence of intervals is $(0, 2, ?, 5, 7, \mathbf{8}, 11)$. The IIIrd degree scale tone can be either 3 or 4 semitones higher than the Ist degree one in order to satisfy the key's definition. □

Proposition 3.2 (The trichotomy of keys)**.** *Let X be an enharmonic equivalence class on an equal-tempered piano K. Then there are exactly three keys with first degree X, up to enharmonic equivalence. These are the major, —with interval sequence* $(0, 2, 4, 5, 7, 9, 11)$*— the minor —with interval sequence* $(0, 2, 3, 5, 7, 8, 11)$*— and a third type, the so-called minor with a Picardian third*[9] *—with interval sequence* $(0, 2, 4, 5, 7, 8, 11)$*. The second and third types are the two ones that satisfy the Minor Lemma.*

Proof. As in the previous proof, the key's definition implies that the interval sequence of an arbitrary key's scale is $(0, 2, ?, 5, 7, ?, 11)$.

If the degree VI scale tone has sign 9, then from the requirement that every triad and seventh chord built up from scale tones has to comply with classical harmony follows that the sign of the IIIrd degree tone cannot be different from 3 or 4. If this sign is 4, then the scale is the X-major scale. If the sign is 3, then the IVth degree seventh chord built up from scale tones is a dominant seventh, but the interval between the IVth and the VIIth degree scale tones is not a perfect fourth but a tritone (enharmonic with 6 semitones/3 whole tones). Therefore in this case we do not get a key[10].

If the VIth degree scale tone has sign 8, then the Minor Lemma implies that we have got two possible keys, exactly which are named in the proposition. □

[9] Referred to as Picardian minor from now on.

[10] This scale is also a harmonic modal scale: the *harmonic Dorian*, which begins with the second degree of the original major scale.

From the 1500s European music is determined by the major–minor duality. The *Picardian minor key*, as we have seen, has the basic key stability properties, but it differs in only one scale tone (degree VI) from the major scale and also in only one scale tone (degree III) from the minor one. Therefore the listener who thinks in only major and minor tries to decide which of this two key types a given music piece is in, even if it is actually in a Picardian minor, and hence it has become perceptually very hard to preserve the stability and individuality of this key type. In music history, this third key type usually appears as a transitional state between the major and the minor key, not as an individual key. It has got another important application: when J. S. Bach harmonizes a chorale with the given soprano melody ending in minor key with the first degree tone, he always *makes the last triad major* (makes its third *Picardian*). The reason of that is that there is a major but no minor triad consisting of a tone's overtones, and the tones of a long minor triad would cause dissonance with the overtones' major triad among certain acoustic conditions (e.g. in churches).

As the minor key comes from a modal scale but its VIIth degree tone is altered (lifted with a semitone), this alteration has to be signed in the triads including this tone. Therefore we use modified notations in Hungarian music theory for the triads and seventh chords built up from scale tones: we refer to the interval from the base to the voice that contains the altered voice with an Arabic letter on the right side of the degree number, —where we also write the sign of the inversion— and the alteration of the bass under the degree number. However, we do not write 3$\sharp$ but only $\sharp$ when referring to an altered third. E.g. $\mathrm{V}^{\sharp}$ – root position fifth degree triad, $\mathrm{V}^{7}_{\sharp}$ – root position fifth degree seventh, $\underset{\sharp}{\mathrm{VII}}{}^{7}$ – root position seventh degree seventh, $\mathrm{III}^{5\sharp}$ – root position third degree triad, $\mathrm{V}\,{}^{6\sharp}_{\substack{4\\3}}$ – seventh degree seventh in second inversion etc. If there are different $\flat$ or $\sharp$ alterations, we will use the same notation system. The *key signature of a (possibly Picardian) minor key* is the key signature of the major scale which has the degree I scale tone of the minor key as degree VI scale tone. These major and minor scales are called *parallel*.

4 Topological features of four-part edition

In this whole section let K denote an arbitrary equal-tempered piano. Remember that *n-part pieces* are $\mathbb{R}_0^+ \to K^4$ Borel-measurable functions, where we consider $\mathbb{R}$ with the standard topology and K^4 with the discrete topology. Now we give the definitions of *strictly four-part pieces*, which we use for modeling classical harmony. Our notions will permit some non-feasible musical phenomena, such as infinite pieces and chords accumulating in one point in time (referred to as *packing point*, see later). Why don't we use only chord sequences, which are enough to describe the four-part examples for classical harmony? Basically for two reasons. On the one hand, we would like to make it possible to talk about *periodic pieces* without ending in finite time, which is often aimed for in 20^{th}–21^{st} century popular music. On the other hand, topological description will make it technically easier to prove the *main theorem of tonality* and also possible to define the genre of *Bach's chorale harmonizations* mathematically precisely and musically almost correctly. When talking about *modulations* (i.e. movements between keys), we will simplify the model and consider only the chord sequences.

Definition 4.1. Let $M : \mathbb{R}_0^+ \to K^4$ be a four-part piece. M is called a strictly four-part piece if:

(i) Each element of Ran M is a four-part version of a triad or a seventh chord (in some inversion) that complies with classical harmony,

(ii) each voice of each element of Ran M only contains tones that can be derived from the C major scale on K using the system of $\flat$'s and $\sharp$'s,

(iii) $\forall H \in \text{Ran } M$: $M^{-1}(H) = \{x \in \text{Dom } M | M(x) = H$ is a disjoint union of intervals closed on the left and open on the right.

Definition 4.2. If M is a strictly four-part piece, $B(M)$, the narrowest interval closed on the left and open on the right that contains Dom M is called the cover of M.

Strictly four-part pieces, which we use for composing examples for showing and teaching compositional principles of classical harmony, and which are very close to J.S.Bach's chorales from real music, are *completely homophonic*. This means that if in a point in time one voice starts to play a new tone, all other voices do so. From this fact, the general definition of n-part pieces and the point (iii) of definition 4.1 follows that if M is a strictly four-part piece, then there is a pause in the point in time $t \in \mathbb{R}_0^+$ in at least one voice of M ($t \notin \text{Dom } pr_i \circ M$), then this is actually a *general pause*, i.e. pause in all voices. Another consequence of these definitions is that the connected components of pauses are also intervals closed on the left and open on the right.

It is easy to prove that every strictly four-part piece M is *continuous from the right*, which means:

$$\forall t_0 \in \text{Dom } M: \lim_{t \to t_0+0} M(t) = M(t_0). \tag{1}$$

More precisely, with respect to the fact that M is a discrete-valued function:

$$\forall t_0 \in \text{Dom } M \ \exists \delta > 0: \ \forall t \in [t_0, t_0 + \delta[\ \ M(t) = M(t_0). \tag{2}$$

Definition 4.3. Let M be a strictly four-part piece.

(i) $t = \inf \ \text{Dom } M$ is the beginning/starting point of M,

(ii) $t = \sup \ \text{Dom } M$ is the ending point of M, if this point is finite,

(iii) t is a melody starting point of M if $\exists \varepsilon > 0$ such that t is the beginning point of $M|_{[t-\varepsilon,\infty[}$,

(iv) t is a melody ending point of M if $\exists \varepsilon > 0$ such that t is the ending point of $M|_{[0,t+\varepsilon[}$,

(v) $t \in \text{Dom } M$ is a chord changing point of M if $\exists \varepsilon > 0$, $\exists H_1 \neq H_2 \in K^4$ such that $\forall x \in [t-\varepsilon, t[\ M(x) = H_1$ and $\forall x \in [t, t+\varepsilon[\ M(x) = H_2$. Let $A(M)$ denote the set of the chord changing points of M.

Definition 4.4. Let M be a strictly four-part piece.
$\inf\limits_{t \in \mathcal{D}(M)} \sup \{r_1 + r_2 | \ r_1, r_2 \geq 0 \ \wedge \ \forall x \in [t - r_1, t + r_2[: \ M(x) = M(t)\}$ is the infimum of chord lengths of M, while $\sup\limits_{t \in \mathcal{D}(M)} \sup \{r_1 + r_2 | \ r_1, r_2 \geq 0 \ \wedge \ \forall x \in [t - r_1, t + r_2[: \ M(x) = M(t)\}$ is the supremum of chord lengths of M.

The proofs of the following two propositions are not very complex but rather technical and therefore we leave them to the reader in this article.

Proposition 4.1. *Let M be a strictly four-part piece and λ the Lebesgue measure. Then $|A(M)| \leq \aleph_0$. If the infimum of the chord lengths of M is positive, then $A(M)$ does not have any accumulation point. If $\sup\limits_{H \in \mathcal{R}(\mathcal{M})} \lambda(M^{-1}(\mathfrak{N}(H))) < \infty$ is also true, then $A(M)$ is finite.*

Proposition 4.2. *Let M be a strictly-four part piece with positive infimum of chord lengths. Then*

(i) $\forall t \in \text{Dom } M$*:* $t \notin \text{Int Dom } M$ *if and only if t is a melody starting point of M,*

(ii) $\forall t \in \overline{\text{Dom } M}$*:* $t \notin \text{Dom } M$ *if and only if t is a melody starting point of M,*

(iii) if $\text{Dom } M = B(M)$ *(then we say that M is pauseless), then M has only one melody starting point: the starting point and only one melody ending point: the ending point.*
M has a finite number of chord-changing points, all other points of $\text{Dom } M$ *are in the interior of an interval on which the image of M is a constant triad or seventh chord.*

We do prove the next proposition:

Proposition 4.3. *Every ω-accumulation point of the domain of a strictly four-part piece M is a complete accumulation point.*

Proof. We want to see that if infinitely many points of Dom M accumulate to $t \in \overline{\text{Dom } M}$, then the actual cardinality of points of Dom M accumulating to t is continuum.

The definition of accumulation point says that for such $t \in \overline{Dom\ M}$, $\forall n \in \mathbb{N}:\ (B_{\frac{1}{n+1}}(t) \setminus \{t\}) \cap \text{Dom } M \neq \emptyset$. Then the axiom of choice guarantees that $\prod\limits_{n\in\mathbb{N}} \left(B_{\frac{1}{n+1}}(t) \setminus \{t\}\right) \cap \text{Dom } M \neq \emptyset$. Therefore there is a sequence $a:\ \mathbb{N} \to \text{Dom M}$ such that $\forall n \in \mathbb{N}\ a_n \in B_{\frac{1}{n+1}}(t) \setminus \{t\}$. Let I_n be the widest open interval containing a_n on which $M[I_n] = \{M(t) | t \in I_n\} = M(a_n)$. This sequence of intervals accumulates to t, because $\forall n \in \mathbb{N}$ I_n has a point, namely a_n, such that $|t - a_n| < \frac{1}{n+1}$. Therefore the cardinality of the points of Dom M accumulating to t is $2^{\aleph_0} = |\text{Dom } M|$, which means that t is a complete accumulation point of Dom M.[11] □

As we anticipated in the beginning of this section, the topological model of strictly four-part pieces gives us opportunity to compose not completely feasible music. On the one hand, domains of strictly four-part pieces are not by all means bounded, on the other hand the homophonic edition does not exclude the existence of *packing points*: accumulation of infinitely many chords to one point —in other words: executing infinite music in finite time.

Definition 4.5. Let M be a strictly four-part piece. $t \in \overline{\text{Dom } M}$ is a packing point of M if $\forall \varepsilon > 0$ $[t - \varepsilon, t[$ contains infinitely many melody starting or chord changing points.

Hence, for feasible music we have to forbid packing points and limit the domain to a bounded interval:

Definition 4.6. A strictly four-part piece M is feasible if it satisfies the following two conditions:

(i) the infimum of chord lengths of M is positive, and if Dom $M \neq B(M)$, then the infimum of general pause interval lengths is also positive,

(ii) $\overline{\text{Dom } M}$ is compact.

This, from the beginning of the current section, has been the construction of *natural parametrization* of strictly four-part pieces. *Playing functions* describe the performance of these pieces with non-constant velocity.

Definition 4.7. Let M be an n-part piece for a certain $n \in \mathbb{N}^+$ and $\theta:\ [0, \infty[\to [0, \infty[$ a continuous, strictly increasing function, for which $[0, \infty[$ can be divided into countably many disjoint intervals $(I_i)_{i\in\omega}$ joining each other and altogether covering $[0, \infty[$, such that in the interior of each interval $\theta \in C^2$, θ' nowhere vanishes and $\inf\limits_{n\in\omega} \lambda(I_n) > 0$.
Then θ is called a playing function. The name of $M \circ \theta|_{B(M)}$ is the playing of M that belongs to θ. For $t \in B(M)$ $\theta'(t)$ is the playing velocity and $\theta''(t)$ the playing acceleration in the point in time t, if they exist. $\theta \equiv 1$ gives the naturally parameterized n-part piece. The set of playing functions is denoted as $PL(\mathbb{R})$.

It is easy to verify that on bounded intervals playing functions are absolutely continuous with respect to the Lebesgue measure. The following definition is based on this.

Definition 4.8. If $\theta \in PL(\mathbb{R})$, M is a strictly four-part piece and A is a Lebesgue-measurable subset of $B(M)$, then the length of the part A of piece M by the playing function θ is $\mu_\theta(A) = \int\limits_A 1 \mathrm{d}\theta x = \int\limits_A \theta'(x)\mathrm{d}x$.

Definition 4.9. For a strictly four-part piece M $A \subseteq B(M)$ has the Ligeti–Boulez measure zero[12] if $\forall \theta \in PL(\mathbb{R})$ $\mu_\theta(A) = 0$.

We leave the proofs of the next two propositions —they can also be found in my thesis.

Proposition 4.4. *$X \subseteq [a, b[\subseteq B(M)$ has the Ligeti–Boulez measure 0 if and only if it has the Lebesgue measure 0.*

[11] This proof does not work only if Dom $M = \emptyset$, but the proposition is also true in this case.
[12] About the origin of this definition see: [2, part I.; Decision]

Proposition 4.5. *$PL(\mathbb{R})$ is a group under the composition of playing functions.*

Now we turn our attention to the mathematical definition of the genre of Bach's chorales. We emphasize that chorale is a live-music genre with originally non-prescriptive characteristics in Baroque, and therefore however precisely we define it, our definition will only be correct with certain exceptions. Therefore our goal is not to determine what a four-part chorale is but to show that the topological description of strictly four-part pieces gives an opportunity to form a definition that is correct for the great majority of these pieces written by Bach for harmonising with a given soprano voice. In preparation for this, we need one more definition.

Definition 4.10. Let M be a strictly four-part piece with $x \in \operatorname{Dom} M$ and $M(x) = H$. Then the connected component of $M^{-1}(H)$ containing x is the area of $M(x)$.
The halving of $I = [a, b[\subseteq B(M)$ in the playing belonging to $\theta \in PL(\mathbb{R})$ is dividing I into two disjoint intervals closed on the left and open on the right I_1, I_2 which together cover I and $\mu_\theta(I_1) = \mu_\theta(I_2)$.

Using this, our definition for four-part chorale is the following. Generally, 'keeping the common tone' means staying at the same degree of a seven-degree scale (moving by a perfect or an augmented first interval), 'step' means moving one (diminished/minor/major/augmented) second upper or lower and 'skip' means any motion that is bigger than a second.

Definition 4.11. A four-part piece $\mathfrak{K}$ is a four-part chorale if there is a feasible, naturally parameterized, pauseless strictly four-part piece M such that $\exists c > 0$: $\forall x \in \operatorname{Dom}(M)$ the length of the area of x by the identic playing function is c, and
(1) $\mathfrak{K}$ can be derived from M with using the following steps, the so-called *figurations*. They are used for a finite number of $x \in \operatorname{Dom} M$ and the figurations excluding each other are not done at the same time.
Types of figurations are:

Chord duplication Halve the area of $M(x)$ by the identic playing function (natural parametrization), and in the first part of the area keep $M(x)$ for $\mathfrak{K}(x)$, in the other part $\mathfrak{K}(x)$ is one constant triad or seventh chord different from $M(x)$.

Suspension Halve the area of $M(x)$ by natural parametrization, in the second part of the area keep $M(x)$, in the first part, in one or two voices change the appropriate tone of $M(x)$ one step higher and keep the remaining voices.

Advancement Halve the area of $M(x)$ by natural parametrizaton, in the first part of the area keep $M(x)$, in the second part, in exactly one voice write a step higher or lower tone, which is equal to the tone in the same voice of the next chord after $M(x)$.

Accented passing tone Suppose that there is a third skip in some voice(s) of M at arriving at or departing from $M(x)$. Then halve the area of $M(x)$, and on the part which is closer to the interval of the neighbouring chord that is involved in the third skip, instead of $M(x)$, write a tone the degree of which is between these two tones' degree. Note that only one accented passing tone per one chord area of M is accepted.

(2) consider the given four-part piece $\mathfrak{K}$ with a playing function θ that differs from the identic playing in the following: $\exists m \in \mathbb{N}$ that θ increases the length of every mth chord interval of M k times greater than originally, where $k \in]1, 2[$ is a conventionally accepted factor. Then we say that there is a pause on every mth chord.

5 Convergence area of a key. Functions and tonality

Let T be a key with degree I scale tone X in an enharmonic equivalence class on the equal-tempered piano K. *Convergence area* of T is $CA(T) = \{(G_i, L_i) | i = 1, \ldots, N\}$, where $N \in \mathbb{N}$ and $\forall i$ G_i is a fixed triad or seventh chord name of K in a certain inversion and L_i is the list of the accepted four-part versions of G_i. L_i is usually described by a formula that is only meant for four-part versions of triads or seventh-chords that comply with classical harmony —see page 6. In simplified word usage we also call G_i an element of the convergence area, $CA(T)$ is the set of chords belonging to T, our goal is to make this meaningful. Before exact definitions we give a little explanation.

Since keys are determined by their scale tones, in classical harmony in an arbitrary key T almost every triad and seventh chord inversion is an element of $CA(T)$. Dissonance of diminished triads and certain classical compositional principles exclude some inversions and four-part versions. All other chords that can be led to the chords built up from scale tones 'appropriately for classical harmony' —this means not only 'complying with classical harmony' but actually 'in a way that is not unusual in the works of Viennese classical authors'— are also elements of $CA(T)$ and are called *altered chords*.

Triads and seventh chords built up from scale tones of T in four-part versions complying with classical harmony are elements of $CA(T)$ with the following exceptions. *Diminished triads* are only accepted in first inversion —in order to avoid strong tritone dissonance—, and especially if they are degree VII ones, then they have to be *third-duplicated* (the third of the triad has to participate in two voices[13]). In classical harmony in every key *second inversions*of only degree I and IV triads are used (and hence only these are convergent). Every non-diminished triad's root position and every triad's first inversion is an element of $CA(T)$.

Example. Which degree triads' root positions are elements of $CA(T)$ if T is a (a) major, (b) minor, (c) Picardian minor key?

After introduction of convergence areas, we can give the definition of weak tonality.

Definition 5.1. Let M be a strictly four-part piece and t an accumulation point of Dom M. We say that M is k-tonal[14] in the point t with keys $T_1, \ldots, T_k$ if there is a connected open neighbourhood U of t such that $\forall x \in U \setminus \{t\} \cap \text{Dom } M$ $M(x)$ is the element of $CA(T_i)$ for some $0 < i \leq k$.

This definition connects the elements of $CA(T)$ to the tonality with key T, but for actual (not only weak) tonality we need to introduce new musical notions, the *functions*: the tonic, subdominant and dominant. Again, we try to give the general characteristics for some well-known musical notion —instead of only naming the tonic, dominant and subdominant chords built up from scale tones in every key. This method will later help us classify the altered elements of the keys' convergence areas by their functions.

In the following, the kth degree triad/seventh chord of a tone will always mean the appropriate triad/seventh chord built up from scale tones and, in these words, *the leading tone/seventh tone of a key T* will refer to the leading tone/seventh tone of the key's Ist degree scale tone. We say that a triad or seventh chord H is *a major chord* if H is a major triad or a dominant seventh. *The leading tone of a major chord* is its third. We say that a triad or seventh chord L is *a diminished chord* if L is a diminished triad or a diminished seventh. *The leading tone of a diminished chord* is its base.

Definition 5.2. Let T be a key, $H \in CA(T)$ be a major or diminished chord, i.e. major triad, diminished triad, major (dominant) seventh or diminished seventh, and $G \in CA(T)$ a major or minor third. We say that H resolves to G if H contains the leading tone of (the base of) G and if there is a tone x belonging to H that is not a scale tone in the major key built on the base of G, then x is the upper leading tone of the fifth of G.

Definition 5.3 (Dominant function (D) and secondary dominant property.)**.** $X \in CA(T)$ has the dominant function in the key T if it resolves to the first degree triad of T. $Y \in CA(T)$ is a secondary dominant chord if it resolves to any other major or minor chord built up from the scale tones of T.

In order to show the most typical dominant chords, we require that the *seventh degree diminished seventh chord*, which is built on the leading tone of the key's first degree scale tone, and which obviously satisfies

[13]VIIth degree triad inversions cannot contain two fifths, because then there would be two tritones among the intervals of the triad's voices, which would be too much dissonance. Two classical compositional principles, the *leading tone axiom* and *avoiding parallel octaves* exclude the usage of the base-duplicated first inversion of the VIIth degree triad.

[14]For $k = 1$: 'weakly tonal', for $k = 2$: 'weakly bitonal', for $k = 3$: 'weakly tritonal' etc.

Type of T	Tonic chords	Dominant chords	Subdominant chords
major	I, VI, VI^7, I^7	V, VII, V^7	II, IV, II^7
minor	I, VI, VI^7	$\mathrm{V}^\sharp$, $\underset{\sharp}{\mathrm{VII}}$, $\mathrm{V}^7_\sharp$, $\underset{\sharp}{\mathrm{VII}}^7$	II, IV, II^7
Picardian minor	$\mathrm{I}^\sharp$, $\mathrm{VI}^{5\sharp}$, $\mathrm{VI}^7_{5\sharp}$	$\mathrm{V}^\sharp$, $\underset{\sharp}{\mathrm{VII}}$, $\mathrm{V}^7_\sharp$, $\underset{\sharp}{\mathrm{VII}}^7$	II, IV, II^7

the definition of dominant function, is also the element of $CA(T)$ if T is a major key. The Minor Lemma (lemma 3.1) guarantees that this chord is built up from scale tones in a minor key, also in the case of the Picardian minor. (We will not name different altered chords in this article, contrary to my thesis, where the well-known secondary dominants, altered diminished seventh chords, minor subdominants and the chords with augmented sixth interval are described as well.)

Definition 5.4 (Tonic function (T))**.** $X \in CA(T)$ has the tonic function in the key T if

(i) X contains a Ist and IIIrd degree tone of T, the first one from the scale T,

(ii) if X contains the leading tone of T, then it is the seventh tone of X,

(iii) if X is secondary dominant, then X is the Ist degree major triad —with a lifted third in the minor case,

(iiii) X has got no augmented and no diminished partial triad.

Definition 5.5 (Subdominant function (S))**.** $X \in CA(T)$ has the subdominant function in the key T, if

(i) X contains the IVth and VIth degree scale tone of T, possibly both altered,

(ii) if X has an altered leading tone (in this case X is a major or diminished chord), then this is the leading tone of the dominant key (in other words: the degree V triad's base),

(iii) the intersection of X and the tones of the Ist degree seventh chord of T is either empty or the Ist degree scale tone.

Proposition 5.1. *Consider the elements of $CA(T)$ which are built up from scale tones. Then the following chart shows which of these chords have the dominant, the subdominant and the tonic function. In this case, every chord has the same function as its inversions. The chords that cannot be found in the chart have got no certain function.*

The degree I triad is the *tonic main triad* of the key T, the degree IV one is the *subdominant main triad* and the degree V one is the *dominant main triad*. *Authentic step* means basically two things in classical harmony. On the one hand, it means modulating (changing key) to the one fifth higher —the dominant— key, without changing the type of key. Among triads this means a V→I or I→IV chord progression. On the other hand, it means function changing $D \to T$ in a certain key, not necessarily between main triads. Similarly, *plagal step* means two things. On the one hand modulation to the one fifth lower —the subdominant— key, among triads making a I→V or IV→I step, on the other hand function changing $T \to D$ in certain key.

A *cadence* is a chord progression consisting of at least two chords that is appropriate for finishing a piece in classical harmony. This way we can call the dominant→tonic or tonic→subdominant steps authentic cadences and the tonic→dominant and subdominant→tonic steps plagal cadences in certain cases, and in a key a *complete authentic cadence* is a chord progression with $T \to S \to D \to T$ function sequence, while a *complete plagal cadence* is a chord progression with $T \to D \to S \to T$. From the overtone system, the connection of leading tones, seventh tones and the circle of fifths (see on pages 19 and 9) it can be explained in great detail why complete authentic cadences are the most applicable for finishing a piece[15]. Most of

[15]'It is not fully understood why V–I imparts such a feeling of finality, but it cannot be denied that it does.' [1, p. 174]

classical, romantic and also recent popular music is based on $D \to T$ resolutions, coloured and strenghened by complete authentic cadences using the function S[16].

In complete authentic cadences, especially at finishing a piece, the *second inversion of the degree I triad* and the *first inversion of the degree III triad*, which are actually *suspensions of the degree I triad* (see definition 4.11 and note that figuration does not mean dissonance in every case!), are used for leading from a subdominant chord to the degree V triad or seventh chord. The most typical authentic chord progression in Viennese classicism is $\mathrm{I} \to \mathrm{IV} \to \mathrm{I}_4^6 \to \mathrm{V}^{(7)} \to \mathrm{I}$. I_4^6 and III^6 are also good for leading to degree V from a dissonant, altered subdominant chord, from which moving directly to V is not possible without violating classical compositional principles. For further information about classical chord progression see [1, p. 173–176.].

We need to give the first definition and the first compositional principle —the first axiom of classical harmony given in this article— about *modulations* in order to become able to define functional tonality. We will sum up description of modulations later, in section 7. What is very important to emphasise: we demand that modulations themselves be *feasible* and *pauseless*: they need to take place during a bounded interval, without general pauses and packing points.

Definition 5.6 (Modulation)**.** If there are keys T_1 and T_2 for the strictly four-part piece M such that Dom M has got a subset $Z = [a, b[$, for which $M|_Z$ is feasible, and $\exists r_1 > 0, r_2 > 0$ such that on the whole set $B_{r_1}(a) \cap \text{Dom } M \setminus Z$ M is weakly (1-)tonal with key T_1 and on the whole set $B_{r_2}(b) \cap \text{Dom } M \setminus Z$ M is weakly tonal with key T_2, then $\forall W \subseteq Z$ we say that W belongs to a $T_1 \to T_2$ modulation. We also say that there is a modulation on Z.

Compositional principle 1 (First modulational axiom)**.** Let M be a strictly four-part piece. If M complies with classical harmony and M contains a $T_1 \to T_2$ modulation, then there exists $[a, b[\subseteq \text{Dom } M$ such that $M(a)$ is the degree I triad of T_1(built up from scale tones), $M(b-) = \lim_{x \to b-0}$ is the degree I triad of T_2 (similarly), and in a M is weakly tonal with key T_1, in b M is weakly tonal with key T_2, and $[a, b[$ is the widest interval which belongs to this $T_1 \to T_2$ modulation.

We have got all notions that we need for defining functional tonality. The basic idea for tonality of a piece in one point is to assign a key to the point as a limit, requiring that all three functions can be found in the neighbourhood of the point.

Definition 5.7. Let M be a (not by all means strictly) four-part piece and t an accumulation point of Dom M. M is tonal in the point t with key T , if there is a connected open neighbourhood U of t such that $U \setminus \{t\} \cap \text{Dom } M$ $M(x)$ satisfies the following conditions:

(i) M is weakly tonal with key T in every point of V[17],

(ii) $M[V]$ contains at least one tonic, one subdominant and one dominant chord of T,

(iii) if $t \notin \text{Int Dom } M$, then only triad-valued points of Dom M accumulate to t.

Definition 5.8. Let M be a strictly four-part piece and t an accumulation point of Dom M. M is tonal in t and modulates from a key T_1 to another key T_2 if there is a connected open neighbourhood U of t such that

[16] As we already mentioned in the proposition 5.1, it is not true that every chord built up from scale tones in a key has a determined function. It can be observed in the chart that function is closely related to common tones with the main triads. The diminished degree VII triad is definitely dominant in every key, as it has got the leading tone in the bass and it contains every tone of the degree V dominant seventh except its bass. The minor or diminished degree II triad has got two common tones with the subdominant main triad. Degree VI has got two common tones with either the tonic and the subdominant main triad, but lack of degree IV tone implies that this triad is an almost completely stable tonic chord. In certain cases in minor (possibly Picardian) it can occur that the degree VI triad also has subdominant characteristics. The step between triads of degree V→VI in any key is called *deceptive cadence* and it is often used before beginning the actual closing complete authentic cadence. The function of the degree III triad is incertain: it has got two common voices with the tonic and the dominant main triad. However, in minor $\mathrm{III}^{5\sharp}$ is augmented, its dissonance makes it lead to the tonic, and therefore this triad is almost completely dominant, although it is neither a major nor a diminished chord.

[17] We can also say that M is weakly tonal in V.

$\exists\, [a,b[\, = V \supseteq U$, where the whole interval V belongs to a modulation as in definition 5.6, which complies with classical harmony apart from the chord-changing points (according to the modulational compositional principles, which we will give later).

Definition 5.9 (Tonal piece)**.** Let M be a strictly four-part piece and $A \subseteq \overline{\text{Dom } M}$. We say that M is tonal on A if it is tonal in every point of $\overline{A}$ by definition 5.7.

Now we can give the missing point from the conditions for compliance with classical harmony to triads on page 6. If a triad H sounds in a point t of a strictly four-part piece M which is tonal in t with key T, then if H complies with classical harmony, then every tone which is altered in T appears in only one voice of M.

6 Structure of the axiom system of classical harmony. The fundamental theorem of tonality

When we think of classical compositional principles, probably *chord-changing or voice-leading rules* come to our mind. The basic goal of the four-part model of classical harmony is to describe applicable chord progression (which chords can be used and how they *should* follow each other) and applicable voice-leading between chords which should follow each other. According to the prescriptive perspective of classical harmony we emphasize these 'should's —however, there are few cases when classical compositional principles tell exactly what and how to do. It is more typical that compositional principles *exclude* some kind of chord progression —e.g. they forbid some dominant to subdominant chord changes— or voice-leading, —e.g. they forbid parallel octaves, parallel fifths and in vocal pieces augmented second steps— but they do not actually determine what music should be composed. Virtually this property of the axiom system makes it possible to actually write pieces of art and not only 'regular examples' complying with classical harmony.

Basic chord-changing (and amongst these: detailed chord-changing) principles have been accepted for more than two hundred years and they are no subject of debate. Decision about some less frequent cases is less equivocal. Therefore our goal is, here and in more detail in Hungarian in my thesis, to give a frame for chord-changing rules so that our whole axiom system for classical harmony is *mathematically consistent and complete*. That is, if we get a strictly four-part piece, using our axioms we will be able to decide if this piece violates any of them or not. We also aim to make the axiom system *extendable* with new chord-changing rules if necessary. This is the cause of not beginning immediately with a description of typical correct and incorrect chord-changing but something more general. Modulational compositional principles also belong to the axiom system of classical harmony, so we will sum them up in the following section.

Definition 6.1 (Correctable piece)**.** Let M be a strictly four-part piece, $a \in \text{Dom } M$, $b \in \overline{\text{Dom } M} \cup \{\infty\}$, $N = [a,b[\, \subseteq B(M)$. If N satisfies all the following conditions:

(i) $M|_N$ has a positive infimum of chord lengths,

(ii) N as an interval (possibly without an upper bound) is the disjoint union of a finite number $2n{+}1$ ($n \leq 0$) of intervals closed on the left and open on the right $(I_1, I_2, \ldots I_{2n+1})$ so that $\forall i \in \{1, 2, \ldots 2n+1\}$ $I_i \cap \text{Dom } M \neq \emptyset$, $\forall 0 < k \leq n$ in the whole interval I_{2k} there is a modulation complying with classical harmony apart from the chord-changing points, and $\forall 0 \leq k \leq n$ for $I_{2k+1} \cap \text{Dom } M$ there is exactly one key T_k such that $\forall G \in M[I_{2k+1}] = \{H \in K^4 |\ \exists x \in I_{2k+1} :\ M(x) = H\} :\ G \in CA(T_k)$ and $M[I_{2k+1}]$ contains at least one tonic, one dominant and one subdominant chord of T_k,

(iii) on N to every melody starting and melody ending point only triad-valued points accumulate,

then M is called a correctable piece,
a) and $t \in \text{Int Dom } M \cap N$, then if there is no chord-changing which is forbidden by the compositional principles according to chord-changing, then we say that M complies with classical harmony in t and N is a classical neighbourhood of t.

b) if $t_0 \in M$ is a melody starting point of $M|_N$, then according to the definition of strict four-part edition $\lim\limits_{t \to t_0+0} M(t) = M(t_0)$. In this case, if $\exists r > 0$ such that $\forall x \in \,]t_0, t_0 + r[\;\; x \in N$ and in x M complies with classical harmony, then we say that M complies with classical harmony in t_0. If t_0 is not the least element of N, then we also call N a classical neighbourhood of t_0.

Definition 6.2. Let M be a strictly four-part piece, $X \subseteq \text{Dom } M$. M complies with classical harmony on X if M complies with classical harmony on each point of X according to the previous definition and every point of X has a classical neighbourhood which is a superset of X. If M complies with classical harmony on Dom M, then we say that M complies with classical harmony.

Now we turn to the most important result of our axiomatisation process, the fundamental theorem of tonality. We can claim that this proposition has been used for centuries in music theory, —this ensures that our definitions for tonality and the definition of correctable piece are rightful— but it cannot be claimed and proved without mathematical tools. The idea of the proposition came from Máté Vécsey, my university yearmate at the Budapest University of Technology, who has also helped me edit and correct this article in English.

Theorem 6.1 (The fundamental theorem of tonality)**.** *Let M be a strictly four-part piece that is pauseless (*$\text{Dom } M = B(M)$*) and feasible. Then M is tonal (on $\overline{\text{Dom } M}$) if and only if it is correctable, i.e. it complies with classical harmony (on* Dom M*) apart from its chord-changing points.*

Proof. First, we show that the condition of the theorem is sufficient for tonality. If M complies with classical harmony apart from the finite set $A(M)$ of its isolated chord-changing points, then every point of Dom M, except the beginning point, has a classical neighbourhood containing Dom M. Therefore $B(M)$ can be divided into an interval system as in the definition of correctability, let $I_1, \ldots, I_{2n+1}$ be such an interval system.

If $0 < k \le n$ is an integer, then in I_{2k} there is a modulation from T_k to T_{k+1}. According to the first modulational axiom (compositional principle 1), in I_{2k-1} or I_{2k} one can find a point x_k in which the first degree triad of T_k that starts the modulation begins, and in I_{2k} or I_{2k+1} a point in which the first degree triad of T_{k+1} that closes the modulation ends. Now consider the division of $B(M)$ into the interval system determined by all these x_k and y_k points: $J_1, \ldots, J_{2n+1}$. It is clear that $\forall k (0 < k \le n)\ J_{2k} \supseteq I_{2k}$, and this implies that $\forall k\ (0 \le k \le n)\ J_{2k+1} \subseteq I_{2k+1}$. $\forall t \in \text{Int } J_{2k+1} \cap \overline{\text{Dom } M}$ Int J_{2k+1} is an open neighbourhood of t that guarantees tonality and key T_k in t, while $\forall u \in \text{Int } J_{2k} \cap \overline{\text{Dom } M}$ Int J_{2k} is an open neighbourhood of u that guarantees tonality and modulation from T_k to T_{k+1} in u (k's as before). With this, tonality has been shown for each point of $\overline{\text{Dom } M}$ except the x_k and y_k points. The first modulation axiom guarantees tonality and key T_k in x_k, because the first degree triad starting the modulation and beginning in x_k is an element of $CA(T_k)$ and $\exists \varepsilon > 0$ such that in $]x_k - \varepsilon, x_k[$ M is tonal with key T_k. It can similarly be derived from the first modulation axiom that in each y_k M is tonal with key T_{k+1}. Therefore M is tonal (on the whole set $\overline{\text{Dom } M}$.

Now we show that the condition is necessary for the tonality. Let M be tonal, feasible and pauseless. $\forall t \in \overline{\text{Dom } M}$ take an open neighbourhood U_t of t that shows its tonality. If possible, let us choose U_t so that it shows key and not modulation. Because Dom M is a bounded subset of $\mathbb{R}$, it can be supposed that $\forall t \in \overline{\text{Dom } M}$ U_t is a bounded open interval. These neighbourhoods give an open cover of M: $\overline{\text{Dom } M} \subseteq \bigcup\limits_{t \in \overline{\text{Dom } M}} U_t$. $\overline{\text{Dom } M}$ is compact and therefore we can choose a finite number of U_t's which still cover it: $\overline{\text{Dom } M} \subseteq \bigcup\limits_{i=1}^{n} U_i$. We can suppose that $U_i = \,]a_i, b_i[$, where $a_i < a_j \Leftrightarrow i < j$ and $b_i < b_j \Leftrightarrow i < j$. The tonality of M guarantees that in U_1 there is a key T_1. Let us start a sequential process with $V = U_1$, $j = 1$ and $T = T_1$ in order to divide $B(M)$ into an interval system that shows that M is correctable. Of the following two possibilities, one is true:

1. If $\forall k > j$ in U_k there is the same key as in V (or $\sup V$ is the ending point of M), then let $V \cup \bigcup\limits_{k>j} U_k \cap$ Dom M be the last interval for showing correctability. In this interval M is tonal with key T, to the

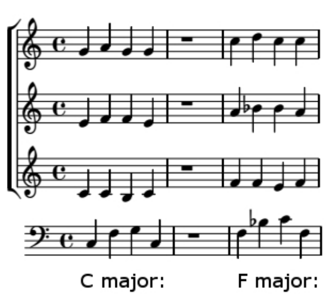
only possible melody starting point, which is the beginning point and to the only possible melody ending point, which is the ending point only triad-valued points of Dom M accumulate.

2. If the previous case is not true, then $\exists k > j$ such that in U_k there is a key T', because definition of tonality implies that in $\sup \mathrm{Dom}\ M$ there is a key. Then let s be the supremum of the points of $]\inf V, \sup U_k[\ (\cap \overline{\mathrm{Dom}\ M})$ in which the key is T and i the infimum of the points of $]\inf V, \sup U_k[\ (\cap \overline{\mathrm{Dom}\ M})$ in which the key is T'. Because $\mathbb{R}$ is a complete ordered field, these points exist, and tonality of M implies weak tonality with key T in s and weak tonality with key T' in i, therefore $s \leq i$. Adding the first modulation axiom to these, $s < i$ follows. Then on the whole interval $[s, i[$ there is a modulation from T to T' complying with classical harmony apart from the chord-changing points. Let us add $[\inf V, s[$ (as an interval with key T) and $[s, \sup U_k[$ (as an interval of a $T \to T'$ modulation) to the already given set of the intervals showing the correctability of M (this set is empty in the beginning). Let $T = T'$, $j = k$ and $V = [i, \sup U_k[$, then go back to the beginning alternative of the sequential process.

Every time we arrive back to the beginning of the process, the ending point of the actual V is the ending point of U_k for a k that is at least one more than the one in the previous turn. This ensures that the process is finite, furthermore the number of turns is not more than n: when $\sup V = \sup \mathrm{Dom}\ M$ holds, the process is finished. The intervals given by the process show that M is correctable: the intervals with an odd index are intervals with key and the ones with an even index contain modulation from the previous interval's key to the following one's. Therefore each point of $\mathrm{Dom}\ M \setminus A(M)$ has got a classical neighbourhood containing $B(M)$. This finishes the proof of the fundamental theorem of tonality. □

When proving that the condition of the theorem is sufficient for tonality, we did not use that Dom M is bounded; therefore every pauseless and correctable —not by all means feasible— strictly four-part piece is tonal. The situation is different with the necessity of the conditions: each condition is necessary for ensuring that the piece complies with classical harmony apart from chord-changing points:

- A tonal piece may have a packing point, in this case it is sure that it has no point in which it complies with classical harmony.
- A tonal piece M with $B(M) = \mathbb{R}_0^+$ may have no packing point but the infimum of the lengths of chord intervals can be still zero (in this case the sum of the chord lengths is infinite).
- A tonal piece M with $B(M) = \mathrm{R}_0^+$ and positive infimum of chord lengths can contain infinitely many modulation intervals. In this case it can occur that $\forall t \in]\inf \mathrm{Dom}\ M, \infty[\ M|_{[0,t[}$ is correctable but M itself is not.
- A tonal piece, even if it is feasible, is not by all means correctable if it is not pauseless. The next music score example is tonal in every accumulation point of the domain, its two connected components by themselves comply with classical harmony, but the whole piece is not correctable because the C major to F major modulation is completely missing.



The technical reason why we needed the fundamental theorem of tonality is to find the exact role of basic chord-changing compositional principles (namely: obligatory moving direction of leading tones, seventh tones and altered tones; avoidance of parallel fifths, parallel octaves, V $\rightarrow$ IV steps and augmented second steps; keeping the common voice; the principle of the least motion etc.). In this article we do not detail these one by one; we have done this following the perspective of the Hungarian music theory coursebooks, e.g. [3, p. 30–183.] —but we show the scheme how they can be claimed according to our theorem.

Compositional principle 2 (Scheme for chord-changing compositional principles)**.** Let M be a strictly four-part piece that satisfies the conditions of the fundamental theorem of tonality and t a chord-changing point of M. If M complies with classical harmony in t, then ... [here come the detailed conditions for the chord-changing in t].

This structure of the compositional principles guarantees that if correctable pieces actually exist and our compositional principles for modulation also do not cause inconsistence, then new chord-changing compositional principles can be added to the axsiom system of classical harmony as long as they do not contradict each other.

7 Chord sequences: the feasible model of the four-part edition. Introduction to the modulational rules.

In this, last section we give an introduction to the modulational rules, without giving the exact details of the technically rather complicated compositional principles themselves and also without naming the *altered chords*, the missing elements of the convergence areas of the keys. Note that our whole model for modulations that can be found in my thesis at [8, p. 55–59.], in which we show all the altered chords of the keys and state the modulational axioms, is mathematically complete but still much less than universal. Authors often use very complex modulations, with visiting many keys before arriving in the actual target key. It is technically quite impossible to give a clear definition for modulations which is based on the properties of every piece of the Viennese classical composers. While four-part pieces without modulations are very close to the real-music Bach chorales and also to some genres of the Classicism, education of music theory has always been using a strongly simplified model for describing modulations. This is the concept of *seven-chord modulations*, which still gives freedom to the $T_1 \rightarrow T_2$ modulations if T_1 and T_2 are close in the circle of fifths (the difference in their key signatures is small) but for example it excludes reaching a third key between the beginning and the ending key. After giving the exact modulational roles the following proposition, which —together with its proof— has been well-known and used for two centuries in music theory, can be proven by giving the structure or music score of exact strictly four-part modulations (e.g. see a simple argument at [3, p. 234–235.]):

Theorem 7.1 (The fundamental theorem of modulations.)**.** *Let T_1 and T_2 be both major or both minor keys. Then there is a $T_1 \rightarrow T_2$ modulation consisting of seven chords which complies with classical harmony.*

The seven chords of the modulation do not have to ensure that the key is T_1 in the beginning point of the modulations, but the new key T_2 has to be strengthened by a complete authentic cadence, according to the compositional principles.

Definition 5.6 for modulations guarantees that in the context of modulations it is enough to consider feasible and pauseless strictly four-part pieces. Pauselessness ensures that chords that cannot follow each other by actual chord-changing also cannot come after each other, separated by a pause. Pause can weaken the impact of irregular chord progression and there are some examples in music history when composers use this. But at modulations our basic aim is to make the key change as smooth as possible and to find some connection between the beginning and the target key, therefore such trickery is not advised in these cases.

Let M be an arbitrary strictly four-part piece and $Pp(M)$ be the set of the packing points of M. If $t \in \overline{\mathrm{Dom}\ M}$ is a packing point, $\exists \varepsilon > 0$ for which the interval $]t, t+\varepsilon[$ does not contain any packing points. This ensures that the following construction gives a well-defined function f. Let $f(x) = x$ for each packing point x of M, and for $y \in B(M) \setminus Pp(M)$ let $f(x) = \inf_{x \in Pp(M), x<y} f(y)$. This function is increasing, and therefore the cardinality of its jumps is countable. Hence, M has only countably many packing points.

Therefore there is a countable ordinal number α for which the set of the starting points of every chord interval and every general pause interval of M can be enumerated in increasing order in α-type: $(t_i, i < \alpha)$. It can be supposed that if t_i is the starting point of a general pause interval, then the starting point of the successor interval $t_{S(i)}$ begins a chord interval (that is, finitely many general pause intervals following each other are united). This way we can define the ***successive pauseless extension***, denoted by $|M$, of an arbitrary strictly four-part piece M from Dom M to $B(M)$: let $|M : B(M) \to K^4$,

$$|M(t) = \begin{cases} M(t), \text{ if } t \in \text{Dom } M, \\ M\left(\inf\limits_{u>t, u \in \text{Dom } M} u\right), \text{if } t \notin \text{Dom } M. \end{cases}$$

It can be derived from the definition of strictly four-part piece that for such a piece M $B(M)$ can be divided into a system of intervals closed on the left and open on the right $(I_i, i < \alpha)$, where α is a countable ordinal number, so that $\forall i < \alpha$ the ending point of I_i is the beginning point of $I_{S(i)}$[18], and on each interval I_i the image of M is constant: $I_i \subseteq \text{Dom } M$, $M[i] = H_i$ $(\exists H_i \in K^4)$ or I_i is a general pause interval of M. It is clear that these intervals can be chosen so that if I_i is a general pause interval, then $I_{S(i)}$ is a subset of Dom M. In this case, the successive pauseless extension of M can be given as

$$|M(t) = \begin{cases} M(t), \text{ if } t \in \text{Dom } M, \\ M[I_{S(i)}] = M(\inf\{u \in \text{Dom } M | u > t\}), \text{ if } t \in I_i \setminus \text{Dom } M. \end{cases}.$$

$|M$ is a pauseless piece which is equal to M on the domain of the original piece. If M has no packing point apart from its ending point, then $\alpha \leq \omega$ and (I_i) is a (possibly finite) sequence of intervals and therefore $\forall i < \alpha$ i is successor ordinal (virtually a natural number). In this case, another invariant extension $\overline{M}$ of M can be given from Dom M to $B(M)$, this is called the *right-invariant pauseless extension* of M:

$$\overline{M}(t) = \begin{cases} M(t), \text{ if } t \in \text{Dom } M, \\ M[I_i] = M(\sup\{u \in \text{Dom } M | u < t\}), \text{ if } t \in I_{S(i)},\ I_{S(i)} \cap \text{Dom } M = \emptyset. \end{cases}$$

In this case $(\overline{M}(t_i), t_i \in A(\overline{M}))$ is the *chord sequence* of M, where t_i's follow each other in their order in $\mathbb{R}_0^+$. We omit the proof of the following proposition, which shows the role of the playing function group $PL(\mathbb{R})$ in the topology of four-part pieces.

Proposition 7.1. *Let M_1 and M_2 be strictly four-part pieces such that the values of the chord sequences of M_1 and M_2 are the same. Then $\exists \theta \in PL(\mathbb{R}):\ M_2 = M_1 \circ \theta$.*

When describing modulations, we will not make any difference between feasible strictly four-part pieces which have the same chord sequence. 'A chord sequence is tonal/is correctable/complies with classical harmony' will mean that every strictly four-part piece with the given chord sequence has this property.

Firstly, we have to ensure the connection of modulational axioms to the definition of the correctable piece and the fundamental theorem of tonality. *Basic modulations* are the modulations which satisfy not only definition 5.6 but also the first modulational axiom (compositional principle 1). For a $T_1 \to T_2$ modulation that complies with classical harmony, even if apart from chord changes, a next necessary condition is to be tonal in the beginning point of the first degree triad of T_1 that opens the modulation, with key T_1, and to also be tonal in the ending point of the first degree triad of T_2 that opens the modulation (these chords are guaranteed by the first modulational axiom).

Now we can turn to the basic idea of modulations complying with classical harmony: the chord sequence of the —pauseless, feasible— $T_1 \to T_2$ modulation section has to be able to be divided into three disjoint segments that cover the whole chord sequence [4, p. 36]:

Neutral phase *(N)* In this segment, which is opened by the degree I triad the key is still T_1 (each element of N is a member of $CA(T_1)$), but there are no secondary dominant chords. In the whole modulation after the first chord there is neither in T_1 nor in T_2 any root position degree I triad until the tonic main triad of T_2 that closes the modulation.

Fundamental step *(F)* If T_1 and T_2 are of the same type and they are neighbours in the circle of fifths, this whole segment may be empty. Otherwise here dominant chords of different keys follow each other.

[18] $S(i)$ is the successor of the ordinal number i, this is $i \cup \{i\}$.

Only the last can be a (major/diminished/in the case of minor T_1 and minor T_2 augmented) triad, the ones before have to be seventh chord inversions. These seventh chords have to follow each other by *elision*[19]. The last chord of F can be a triad and the previous chord may resolve to it.

Cadence ***(C)*** The modulation has to be finished with a complete authentic cadence in the new key T_2, this shows and ensures the tonality in the new key. It may occur that we do not write a cadence in each key, but a modulation progress is only finished when we reach a cadence in some key. Actually, the only sure thing is that the last chord before the closing degree I triad is the degree V triad/dominant seventh chord of T_2 in the segment C: the tonic and subdominant chords before these belong to C (and not F) if and only if they are built up from scale tones in T_2, otherwise they belong to F.

If a basic modulation chord sequence has all the properties that we have introduced in this section, and it is the member of one of the following three modulation types, then we say that it complies with classical harmony apart from its chord-changing points. If its chord-changings are also acceptable, the modulation complies with classical harmony. The modulation types are:

Diatonic For the last chord H of N $H \in CA(T_2)$, and every chord after this is convergent in T_2. This time F usually consists of at most one chord. This is the smoothest possible key change, but it is often not possible between keys further from each other.

Enharmonic The last chord of N or the first chord of F is an element of T_1 that is enharmonic with some element of $CA(T_2)$. The most common enharmonic modulation types use the enharmonic equivalence of degree VII diminished sevenths or different minor keys' degree III augmented triads in different keys. After playing this chord, we consider it as an element of $CA(T_2)$, and make a cadence in T_2.

Chromatic There is elision in the modulation chord sequence. Very far modulations, such as C major $\rightarrow$ F$\sharp$ major can be feased this way. In most of the cromatic modulations $\sharp F \geq 2$ holds.

While it is very difficult to accomplish a modulation that is diatonic and enharmonic at the same time, these three categories do not exclude each other pairwise. In the music score collection of my thesis we show examples of both enharmonic and chromatic and both diatonic and chromatic modulations.

Modulational compositional principles finish our axiomatization work. By giving more and more detailed chord-changing compositional principles based on the research of real classical music pieces and Bach's chorales, the four-part model can be refined. Some of the chord-changing rules, e.g. the *principle of least motion* are very hard to formalize mathematically precisely. However, there are significant results about this compositional principle. Dmitrij Tymoczko has shown in a wider context than strictly four-part edition that voice leadings complying with this principle are *crossing-free*, while for every voice leading that violates it there is a voice-leading that is closer to the least motion but it does contain voice crossing [5, p. 4–6.].

Finishing our mathematical axiomatization of classical harmony, we name our two basic goals for the close future. On the one hand, as we have already mentioned, we would like to write a new Hungarian classical harmony coursebook for high schools in cooperation with professionals of music theory, according to the logical ordering and the mathematically simpler results of our approach. On the other hand, we would like to do experiments on the possibilities and barriers of composing four-part chorales by Markov models, according to the experience of the paper [6].

References

[1] Benson, Dave. *Music: A Mathematical Offering*. University of Aberdeen, 2008.

[19] We generally use the word 'elision' for chord progression of inversions of different seventh chords, without the first seventh chord resolving to its tonic. The lack of resolution is expressed by 'elision', the greek word for omission. Chord progression using elision always has to use *chromatic* for making it possible to comply with classical harmony. Chromatic means the sequence of at least two semitone steps after each other in one voice, so that there is a key in the scale of which every other step changes degree (i.e, every other step is a minor second and the remaining steps are augmented primes) —unless the direction of motion changes, because in this case two steps with the same interval name can follow each other.

[2] Ligeti, György. Decision and automatism in the piece *Structure 1a* of Pierre Boulez. (Original German title: Entscheidung und Automatik in der *Structure 1a* von Pierre Boulez.) *die Reihe* 4 (1958), p. 38–63. *GS* 1, p. 413–446.

[3] Lörinc, Kesztler. *Összhangzattan* (Harmony). Music theory coursebook for Hungarian high schools. Editio Musica Budapest, 1952.

[4] Mazzola, Guerino. Mathematical Music Theory, Status Quo 2000. Conference lecture, ETH Zürich, Departement GESS, and Universität Zürich, Institut für Informatik, 15th October 2001. Online reference: http://recherche.ircam.fr/equipes/repmus/mamux/documents/status.pdf.

[5] Tymoczko, Dmitri. The Geometry of Musical Chords. In: *Science* 313, 72 (2006).

[6] Thorpe, Christopher A. *C.P.U. Bach: Using Markov Models for Chorale Harmonization*. Harvard College, Cambridge, Massachusetts, 1998.

[7] Bach, Johann Sebastian. *Vierstimmige Choralgesänge*. Editor: Imre, Sulyok. Editio Musica Budapest, 1982.

[8] Tóbiás, András. *A klasszikus összhangzattan axiomatikája (The axiom system of classical harmony)*, Mathematics BSc degree work, Budapest University of Technology, 2014. Online reference: `math.bme.hu/~tobiasaj/analosszhang_szakdolg.pdf`