

Structural Characterization of Musical Harmonies

María Rojo González

Simone Santini

Escuela Politécnica Superior
Universidad Autónoma de Madrid

Abstract

Understanding the structural characteristics of harmony is essential for an effective use of music as a communication medium. Of the three expressive axes of music (melody, rhythm, harmony), harmony is the foundation on which the emotional content is built, and its understanding is important in areas such as multimedia and affective computing.

The common tool for studying this kind of structure in computing science is the formal grammar but, in the case of music, grammars run into problems due to the ambiguous nature of some of the concepts defined in music theory.

In this paper, we consider one of such constructs: modulation, that is, the change of key in the middle of a musical piece, an important tool used by many authors to enhance the capacity of music to express emotions. We develop a hybrid method in which an evidence-gathering numerical method detects modulation and then, based on the detected tonalities, a non-ambiguous grammar can be used for analyzing the structure of each tonal component. Experiments with music from the XVII and XVIII centuries show that we can detect the precise point of modulation with an error of at most two chords in almost 97% of the cases. Finally, we show examples of complete modulation and structural analysis of musical harmonies.

1 Introduction

Music is one of the fundamental media through which the human experience is created and shared. All cultures, from before the emergence of Homo Sapiens, have felt the need to emit harmonious sounds to communicate, to ritualize, to reinforce human bonds. We, as a species, came to life in a world that was already full of music. From the traditional dances of the early human culture to the mating rituals of contemporary young people in clubs and raves, music has been a constant presence in the life of our species.

The very dawn of western science coincides with the pythagorean discovery of the relation between music and mathematics: sounds with a pitch ratio of 2:1 are perceived as the same note in different octaves, while the ratio 3:2 (the perfect fifth) is pleasing to the ear and generates a sense of expectation: the first confirmed relation between mathematics and emotions; affective computing *ante litteram*. In the ancient world, the relation between mathematics and music was so well established that in medieval universities mathematics consisted of two major branches: geometry and music.

Tonal music, created between the XVI and the XVII century mainly in Germany, is one of the pinnacles of European civilization, a great audible mathematical cathedral whose principles are exposed in Bach's "The art of the Fugue". For this is the unique position of music: it is a very direct emotional medium, one that speaks to us without the intervening mediation of language (Isaac Asimov once said that Beethoven is greater than Shakespeare because his work did not need translation), a truly emotional medium but also, at the same time, one of the media whose basis is more firmly mathematical.

The idea of using computers to formalize music pre-dates the existence of actual computers, and can be traced back to an observation by Ada Lovelace [8, 12]. For a contemporary computer scientist, such an attempt is interesting from two points of view. On the one hand, music is the perfect testbed for analyzing the formalization of emotionally "charged" media [10], and the limits of such formalization: music is mathematical but, like any truly open system, musical creativity is (fortunately) impossible to formalize. On the other hand, music (and automatic musical generation) is a fundamental part of any multimedia experience (try to remove the music from the shower scene in Hitchcock's *Psycho*) and, in order to understand how to use this medium properly, one has to realize that music is not merely a sequence of sounds, but has an harmonic structure. The characterization of this harmonic structure is the objective of this paper.

The interest of computing scientists for music goes back to the days of yore of computing. Surveys on the early work in the field can be found, for example, in [1, 13].

Work on computer music has developed in various directions. Some work has concentrated on the sequential aspects of music, the most common techniques in this direction being Markov models [2] and machine learning [6]. This work has obtained good results, especially in music composition [3, 4, 9], but has fallen short of capturing the structural aspects of harmony and, in general, of creating a bridge between mathematical formalization and musical theory as developed, for example, in [14]. Similar techniques have also enjoyed considerable success in the area of music retrieval and recommendation [5, 11].

A second line of work has concentrated on the structural characteristics of harmony. A lot of this work has used formal grammars as a modeling tool [15, 17]. The problem with this approach is that many musical structures are, as we shall see, strongly contextual, when not ambiguously defined. The result is that grammars that cover a significant range of musical styles turn out to be ambiguous [16]. In most cases, this ambiguity is eliminated by restricting the grammar to a specific genre, such as the 12-bar blues [18] or other forms of jazz [7].

In this paper, we propose to follow a different route. We considered a fairly general grammar such as that in [16], analyzed the various types of ambiguity, and looked for ways to remove them. We concentrated in particular on *modulation*, the process of changing the key in the middle of a piece. Modulation cannot be disambiguated in a context-free grammar and, rather than resorting to a much more complex context-sensitive grammar, we use a numerical method to detect modulation and integrate it into the analysis. The result is, we believe, a better understanding of a process that has a great impact on the effectiveness of the musical medium to convey emotions and to highlight the contents of other media.

2 Musical Background

Harmony is the study of chords and their construction and chord progressions and the principles of connection that govern them. In this paper we examine the structural relation between chords. Depending on the time period and the musical style these principles may vary. Here, we focus on the Western classical music, in particular on the tonal music of the XVII, XVIII and early XIX century.

We can divide the harmony in 24 *tonalities* (or *keys*, we shall use the two terms interchangeably). Each tonality is an arrangement of notes, scales and chords with different functions and relations. They relate with each other in a hierarchy where the tonic chord is the one with the greatest stability and the rest of chords play around it. Depending on the scales there are 12 major tonalities (C, B \flat ,...) and 12 minor tonalities (Cm, B \flat m,...). Functionally, the major and minor mode of a tonality

are very similar and some composers of the XVIII and XIX century even consider the two modes as two different aspects of the same tonality. So we are going to consider that a tonality is the union of its major and its minor modes (C is the union of C major and C minor), considering a total of 12 keys.

Tonal music is based on diatonic scales, composed of seven notes each. These notes are called the scale degrees, they are represented with roman numerals (I, II, \dots, VII) and have different harmonic functions. The main functions are *tonic* (I), *dominant* (V, VII), *sub-dominant* (IV, II) and *modal* (III, VI). Following standard musical theory, we consider the latter as part of the tonic function. The chords of a tonality are built from the notes of the corresponding diatonic scales. In this paper we use two different notations: the Roman numerals notation (IV of C, or IV_C) and the English one (F). The Roman numerals one is the notation corresponded with the scale degree where the chord is formed from, determining its function. This notation is used along with the tonality to determine the chord. On the other hand, the English notation determines a particular chord without requiring the specification of a key.

Chords are played sequentially and each progression of chords can lead to a more or less stable chord, creating musical interest. The movement of a chord from an unstable sound to a more stable chord is called a *resolution*. The most common and stable resolutions are called *regular* resolutions, whereas an *irregular* resolutions are less common and less stable than a regular one.

The tonal center of a musical fragment is the note or chord with more stability and it usually corresponds to the tonic of the key. We mostly use chords near this tonal center, that is, chords that belong to the corresponding key. But we can briefly leave the tonal center (without leaving the tonality) in order to make music more interesting, in order words, we can use chords that do not correspond to the current key. An example of these chords are the secondary dominants. As their name suggests, the secondary dominants are dominant chords (V) of a scale degree that is not the tonic, considering this scale degree as a new temporal tonic. For example, in the key of C the secondary dominant of the chord II (Dm), is the dominant of the key of Dm (A), and we represent it as V^{II} . The regular resolution of a secondary dominant is the chord used to form it: $V^{II} - II$.

Lastly, *modulation* is the process of changing the tonality within a musical fragment, that is to say, changing the tonal center for a significant amount of time. There are different types of modulation depending on the duration of the new tonal center and its stability.

In essence, a tonality is formed by a determined set of chords. If we only use those chords within a musical fragment, music will get monotonous and lack of interest. For this reason, we usually change from one tonality to another (modulation) and within a tonality we occasionally use chords from other nearby tonalities (secondary dominants).

3 Grammar Analysis

At the higher hierarchical level, a musical *piece* is composed of *regions* that play, each one, one of the three major harmonic roles: tonic (TR), dominant (DR), or sub-dominant (SR). In order to avoid ambiguities in the parse tree, we define separate terminals for the initial regions and their continuation (TR, CTR, DR, CDR, SR, CSR). With these definitions, the top-level of

our grammar is:

$$\begin{aligned}
 \text{piece} &\rightarrow TR \\
 TR &\rightarrow CTR \mid CTR TR \mid CTR DR \\
 CTR &\rightarrow DR t \mid t \\
 DR &\rightarrow CDR \mid CDR DR \\
 CDR &\rightarrow SR d \mid d \\
 SR &\rightarrow CSR \mid CSR SR \\
 CSR &\rightarrow s
 \end{aligned} \tag{1}$$

These regions resolve in sub-trees rooted at the non-terminals t , d , s , which represent the localized tonic, dominant, and sub-dominant functions, respectively. The harmonic functions expand as

$$\begin{aligned}
 t &\rightarrow dI \mid dI dIV dI \mid dVI \mid dIII \\
 s &\rightarrow dIV \mid dII \mid \flat II \\
 d &\rightarrow dV \mid dVII
 \end{aligned} \tag{2}$$

These non-terminals are local harmonic functors that expand either to the corresponding chord or to a secondary dominant with a regular (that resolves to that chord) or an irregular resolution. For example, $dIII$ resolves either to III or to a secondary dominant with its resolution:

$$dIII \rightarrow III \mid V^{III} III \mid V^V III \mid VII^{III} III \mid VII^V III \tag{3}$$

and similarly for the others. For example, if we are in the key of C , this production would be equivalent to

$$dIII \rightarrow Em \mid B Em \mid D Em \mid D\sharp^o Em \mid F\sharp^o Em \tag{4}$$

The objective of this and the analogous expansions of $dI, \dots, dVII$ is to allow the introduction of secondary dominants at a limited depth. Resolutions like $V^{III} III$ (the dominant of the 3rd followed by the 3rd, or $B Em$ in the key of C) allows the introduction of chords far from the tonal center of the piece, adding interest to otherwise flat harmonies. In our grammar we have deliberately chosen to limit the depth of nesting of secondary dominants. In [16], for example, secondary dominants are defined by a recursive rule, allowing constructions such as $V^V V$, $V^{VV} V^V V$, $V^{VV^V} V^{VV} V^V V$, and so on. On the one hand, this results in an ambiguous grammar; on the other hand, this allowed the specification of sequences like $V^{VV^V} V^{VV} V^V V$ that, from the musical point of view, make no sense as they go too far from the tonal center without settling on a new one. Note that the secondary dominant rules are very low in the hierarchy of productions, reflecting the fact that they are local phenomena that do not span more than two or three chords.

Ambiguity also emerges when trying to introduce other common musical constructs, such as the *descending fifths*: a progression where each chord is a fifth above the next one, all chords being in the same key. For example, the sequence $Em Am Dm G C$ is a descending fifths sequence in the key of C ($III VI II V I$). These progressions can in principle be analyzed using the productions of [16]:

$$\begin{aligned}
 dI &\rightarrow dV \mid I \\
 dII &\rightarrow dVI \mid II \\
 dIII &\rightarrow dVII \mid III \\
 dIV &\rightarrow dI \mid IV \\
 dV &\rightarrow dII \mid V \\
 dVI &\rightarrow dIII \mid VI \\
 dVII &\rightarrow dIV \mid VII
 \end{aligned} \tag{5}$$

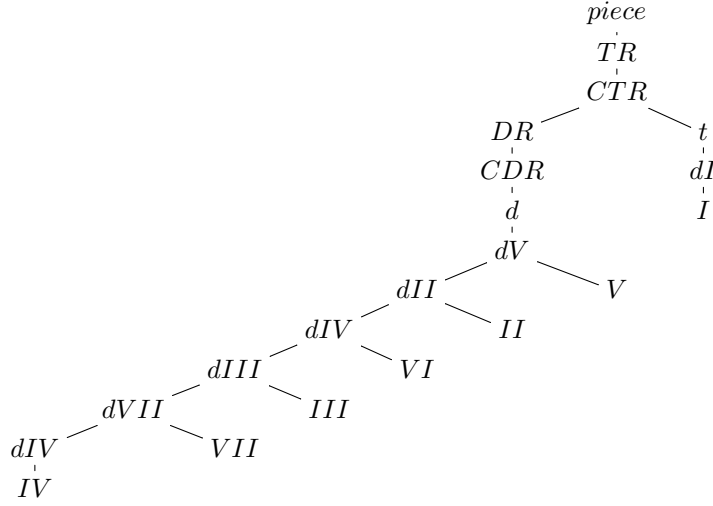


Figure 1: Parse tree that results from the analysis of the sequence *IV VII III VI II V I* using the descending fifths sequences productions.

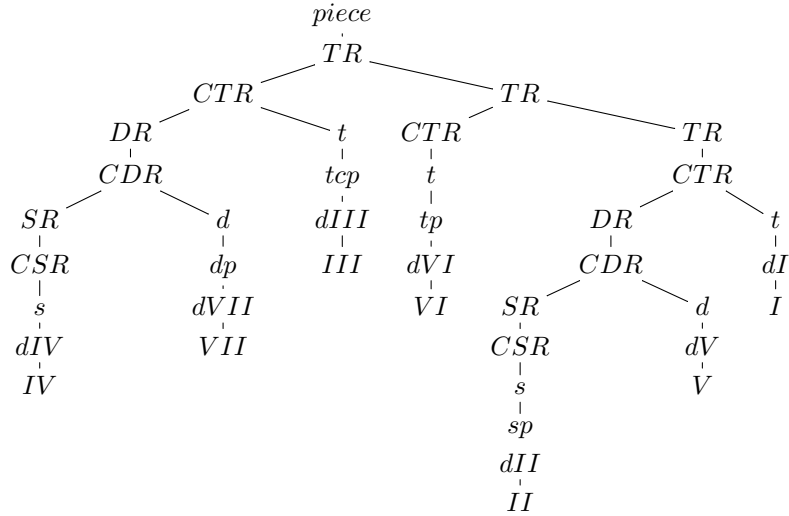


Figure 2: Parse tree that results from the analysis of the sequence *IV VII III VI II V I* using the rest of the productions of the grammar, proving that the descending fifths sequence productions make the grammar ambiguous.

obtaining parse trees as the one shown in Figure 1.

However, these sequences could be also analyzed using the rest of productions of the grammar, as shown in Figure 2, which means that the rules of 5 make the grammar ambiguous. Furthermore, the definition of this kind of sequences is, like many concepts in music, a fuzzy one: it is not clear what is the minimum number of chords needed to determine the existence of a

piece	\rightarrow	TR	dI	\rightarrow	$V^{III} I \mid V^V I$
TR	\rightarrow	$CTR \mid CTR TR \mid CTR DR$	dII	\rightarrow	$V^{II} II \mid VII^{II} II \mid V^{IV} V^{II} II$
CTR	\rightarrow	$DR t \mid t$	$dIII$	\rightarrow	$V^{III} III \mid V^V III \mid VII^{III} III \mid VII^V III$
DR	\rightarrow	$CDR \mid CDR Dr$	dIV	\rightarrow	$V^{IV} IV \mid V^{VI} IV \mid VII^{IV} IV \mid VII^{VI} IV \mid V^V V^{IV} IV$
CDR	\rightarrow	$SR d \mid d$	dV	\rightarrow	$V^V V \mid V^{II} V \mid VII^V V \mid VII^{II} V$
SR	\rightarrow	$CSR \mid CSR SR$	dVI	\rightarrow	$V^{VI} VI \mid VII^{VI} VI$
CSR	\rightarrow	s	$dVII$	\rightarrow	$V^{VII} VII \mid VII^{VII} VII$
t	\rightarrow	$tp \mid tcp dI \mid dI dV dI$	dI	\rightarrow	I
s	\rightarrow	$sp \mid dIV$	dII	\rightarrow	II
d	\rightarrow	$dp \mid dV$	$dIII$	\rightarrow	III
tp	\rightarrow	dVI	dIV	\rightarrow	IV
dp	\rightarrow	$dVII$	dV	\rightarrow	V
sp	\rightarrow	$dII \mid bII$	dVI	\rightarrow	VI
tcp	\rightarrow	$dIII$	$dVII$	\rightarrow	VII

Figure 3: The grammar that we use in our analysis. Here, $I, II, III, IV, V, VI, VII$ are the terminals, while "piece" is the start symbol.

descending fifths sequence, nor whether it makes sense to talk of such minimum, independent of the harmonic context in which the sequence is placed. For these reasons, in this work we disregard descending fifths sequences, dropping the corresponding productions from the grammar (see the conclusion for more on this subject).

The complete grammar is shown in Figure 3. Figure 4 shows an example of an analysis: we consider the sequence $I II V I VI V^V V I$ from J. S. Bach's *Prelude 1 in C major* (BWV 846).

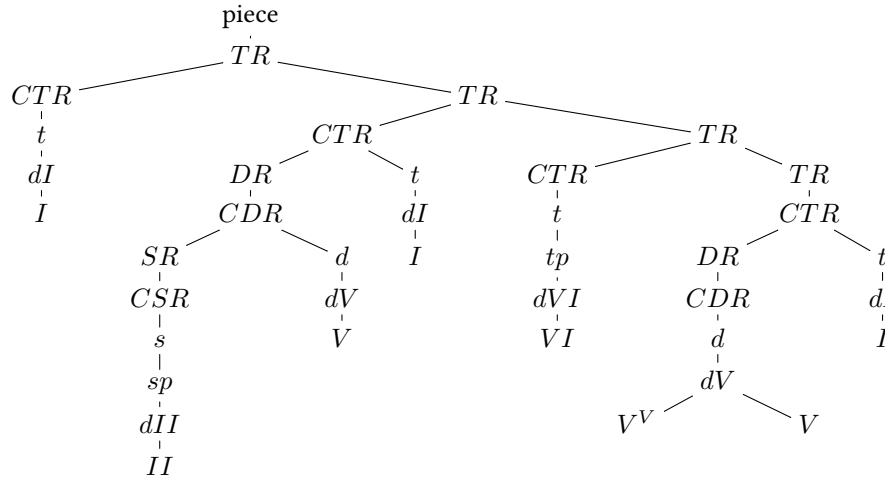
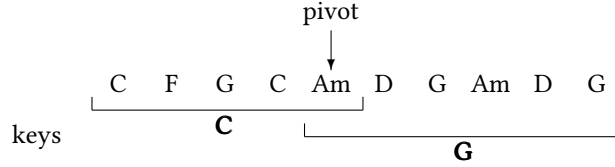


Figure 4: An example of harmonic analysis from the fragment $I II V I VI V^V V I$ from J. S. Bach's *Prelude 1 in C major* (BWV 846). The example shows a typical structure of tonic, dominant, and sub-dominant region, and a secondary dominant (V^V).

4 Modulation

One of the major hurdles in grammar-based approaches is the detection of modulation, that is, of the change of key in the middle of a piece. Modulation is typically articulated around a *pivot chord*, a chord that belongs to both keys and that marks the transition point from one to the other:



This process is inherently ambiguous for two reasons. First, each chord has a harmonic interpretation in any key, so that sequences of chords can be interpreted in exponentially many ways. In Figure 5, for example, we interpret the same sequence of chords with a modulation (the most natural way, from the point of view of music theory), without modulation (in a single key), and with an arbitrary sequence of modulations (the subscript is the key in which we interpret the chord, resulting in the harmonic function indicated). Second, the presence of a chord far away from the tonal center is not enough to define a

	C	F	G	C	Am	D	G	Am	D	G
(A)	I_C	IV_C	V_C	I_C	VI_C/II_G	V_G	I_G	II_G	V_G	I_G
(B)	I_C	IV_C	V_C	I_C	VI_C	V_C^V	V_C	VI_C	V_C^V	V_C
(C)	V_F	I_F	III_E	VI_E	IV_E	I_D	V_D	I_{Am}	V_G	I_G

Figure 5: A sequence of chords interpreted in three different ways: with a modulation (in (A), the most natural way, from the point of view of music theory), without modulation (in (B), in a single key), and with an arbitrary sequence of modulations (in (C)). The subscripts represent the key in which we are interpreting the chord to give it the indicated harmonic function.

modulation: it is necessary to have a reasonably long sequence of chords reasonably close to a different tonal center. The double use of "reasonably" shows that the definition is fuzzy, and not easy to analyze with a formal grammar, at least, not with a context-free one. We have therefore decided to decouple the analysis of the modulation from the grammar. Modulation is detected by gathering evidence on the local key by giving each chord a score that depends on the centrality of that chord. For example, Table 1 shows the centrality of various chords for the key of C. There are 12 such tables, one for each key. In order to

		<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>VI</i>	<i>VII</i>	<i>bII</i>
C	major	C (5)	Dm (3)	Em (2)	F (3)	G/G7 (5)	Am (3)	B° (3)	D♭ (1)
	secondary dominant	-	A/A7 (1)	B/B7 (1)	C7 (1)	D/D7 (1)	E/E7 (1)	-	
Cm	ascending melodic	-	Dm (3)	-	F (3)	-	-	-	
	harmonic minor	Cm (5)	D° (2)	-	Fm (3)	G/G7 (5)	A♭ (2)	B° (3)	
	descending melodic	-	-	E♭ (2)	-	Gm (2)	-	B♭ (2)	
	secondary dominant	-	-	B♭7 (1)	C7 (1)	D/D7 (1)	E♭7 (1)	F7 (1)	

Table 1: Table with the scores that determine the chord centrality for the key of C. Chords not represented in the table have a score of 0.

	C	F	C	Dm	G7	C	G7	C	Tot.
C	5	3	5	3	5	5	5	5	36
G	3	2	3	2	1	3	1	3	18
D	2	2	2	5	1	2	1	2	17
A	2	2	2	3	1	2	1	2	15
E	2	1	2	0	1	2	1	2	11
B	1	0	1	0	0	1	0	1	4
F#	0	1	0	0	0	0	0	0	1
C#	1	1	1	0	0	1	0	1	5
Ab	1	1	1	0	1	1	1	1	7
Eb	1	1	1	0	1	1	1	1	7
Bb	1	5	1	2	1	1	1	1	13
F	5	5	5	3	1	5	1	5	30

Table 2: Determination of the key of the sequence $C F G Dm G7 C G7 C$ (top row). Each chord is scored in all keys and the scores are accumulated (last column). The highest score indicates that the fragment is in the key of C.

determine the key of a sequence of chords, we consider the score of each chord in each of the 12 keys and accumulate the scores: the key with the highest score will be considered as the key of the sequence of chords. Consider, for example, the sequence $C F G Dm G7 C G7 C$. Table 2 shows the process of detecting the key: each column represents the score of that chord in each of the keys (e.g., C has a weight of 5 in the key of C, 3 in the key of G, 2 in the key of D, etc.). At the end of each row we indicate the accumulated score for the corresponding key. The highest score indicates that the fragment is in the key of C. Based in this general idea, modulation is detected by sliding a window on the musical piece and detecting, in each window, the dominant key. When the dominant key changes, we detect a modulation, and the central chord of the window in which the key changes is taken as the pivot. Table 3 shows the analysis of the sequence $C F C Dm G7 C G7 C C C Fm E^o Fm E^o Fm$; When the window covers chords 6 to 13, a key change is detected. Correspondingly, we place the pivot chord in position 9 (position $\lfloor (13 - 6 + 1)/2 \rfloor$ of the window), the central C of the series of three:

	C	F	C	Dm	G7	C	G7	C	C	C	Fm	E ^o	Fm	E ^o	Fm
C	I	IV	I	II	V7	I	V7	I	I						
F										V	V	I	VII	I	VII

Different types of modulations are detected as different patterns of score changes in the sliding window. The most common examples are shown in Figure 6. In Box 1 we have a regular modulation: the key of the piece changes from C to F; both the key of C (prior to the modulation) and that of F (afterwards) are maintained for a significant amount of time and are well established to the ear. In Box 2 we have an example of *passing* (or *false*) *modulation*: we have a sequence of three keys, $A\flat$, $D\flat$, and $B\flat$. $A\flat$ and $B\flat$ are dominant for some time, and are thus well established, but $D\flat$ is dominant for a short time, never becoming well established and working only as a nexus between $A\flat$ and $B\flat$. Finally, in Box 3, we have an example of *tonicization*: here too we have a key (F) that stays dominant for a short period but, in this case, before and after the modulation we are in the same key: $B\flat$. This solution is used by composers when they want to add chromaticism to the music: here, F is the dominant

	1-8	2-9	3-10	4-11	5-12	6-13	7-14	8-15
C	36	36	38	36	33	31	26	24
G	18	18	19	16	14	13	10	9
D	17	17	17	15	12	11	11	10
A	15	15	15	13	10	9	7	6
E	11	11	12	10	10	9	7	6
B	4	4	5	4	4	4	3	3
F#	1	1	0	0	0	0	0	0
C#	5	5	5	6	6	8	7	9
A#	7	7	7	9	9	11	10	12
E#	7	7	7	9	9	11	10	12
B#	13	13	9	10	8	9	8	9
F	30	30	30	30	30	34	32	36
C	C	C	C	C	C	F	F	F

Table 3: Determination of the modulation of a sequence using a sliding window of 8 chords. Each column represents the accumulated scores of the 12 keys for the position of the window shown in the first row. The last row shows the dominant key for each position of the window. A modulation is detected in window 6-13, thus identifying the 9th chord (position $\lfloor (13 - 6 + 1)/2 \rfloor$ of the window) as the pivot.

of B \flat and in that fragment there are chords that are related to F, not with the purpose of changing key, but to have a more "dominantly-colored" harmony on B \flat .

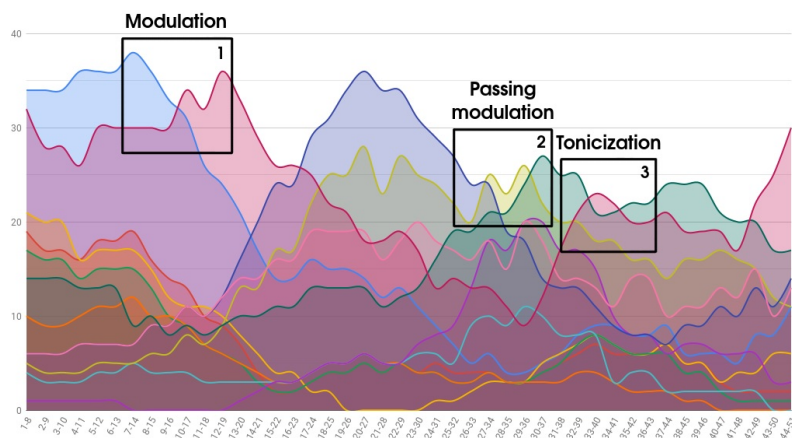


Figure 6: Analysis of the modulation of Chopin's 4 Ballade. The boxes indicate different types of modulation (see text). For the color codes of the curves and their correspondence with tonality, see Figure 7.

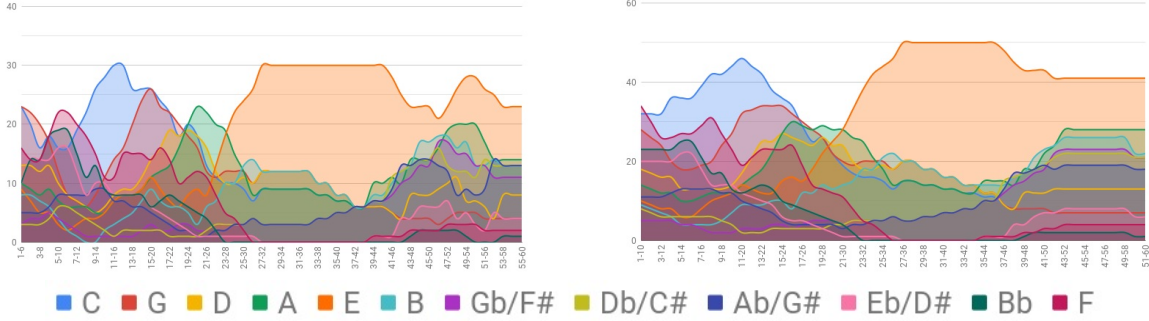


Figure 7: Analysis of the modulation of Beethoven's *Waldstein Sonata* using windows length $W = 6$ (left) and $W = 10$ (right).

5 Results

The most important parameter of our method is the window length W . Large values of W ($W \sim 10 - 12$) give a more firm establishment of the dominant key, as they allow the method to gather more evidence. On the other hand, large values of W act as a low-pass filter on the sequence of chords, which could lead to overlook fast tonicizations or false modulations. Short windows sizes ($W \sim 2 - 4$), on the other hand, could lead to over-detection of tonicization simply due to the presence of a few chords far from the tonal center. The problem is made more complicated by the fact that there is, in these cases, no established ground truth: whether a series of chords are a tonicization (or a false modulation) or simply a few chords far from the tonal

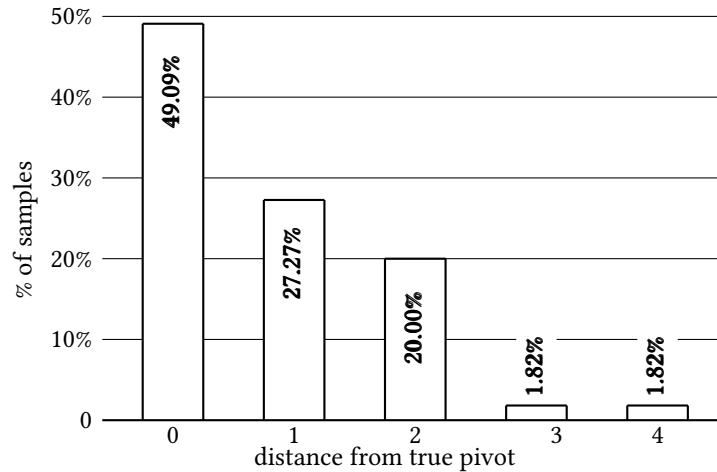


Figure 8: Precision in the placement of the pivot chords, from a sample of 55 modulations. In about 97% of the cases the method places the pivot chord within two chords of the theoretical one determined by harmonic analysis.

center, is a matter of interpretation, and different analysts may (and will) differ in their analysis.

An important factor in determining the value of W is the style of the music that we are analyzing. We have observed that a value of $W = 6$ is adequate for music from the XVII to the middle of the XIX century. Modern music, up to the early XX century, is more experimental with key variations, and authors do resort to rapid, short changes of key to the point that sometimes the very concept of tonal center is lost. In these cases, smaller values of W must be used. Large values of W are useful also when we want to ignore the rapid variations of key and detect the main harmonic themes of a piece.

The beginning of Beethoven's *Waldstein Sonata* is an example (Figure 7): the key values oscillate very rapidly, and do not settle on any established key—the fragment is harmonically ambiguous. With a longer value of W (Figure 7b) we can appreciate that the dominant key of the beginning of the *Waldstein Sonata* is C. An important measure to evaluate our method is the precision with which we can identify the pivot chord. Figure 8 shows the distance between the position of the pivot chord as estimated by our method and that determined by musical analysis. The data are based on a sample of 55 modulations from musical pieces of the XVII, XVIII, and XIX centuries. In about 97% of the cases, the estimated pivot is within two chords of the actual one.

Finally, in Figure 9 we show a complete analysis of a musical fragment, from Mozart's *Eine kleine Nachtmusik*. At the bottom we show the weights of the main keys involved (we do not show the score of all the 12 keys for the sake of clarity: the other keys receive a score well below the two we show): the piece begins in G and, after 10 chords, switches to D. The G chord marked in bold is the pivot, and belongs to both parts; consequently, it has two different harmonic functions: it is a tonic (I) for the key of G and a sub-dominant (IV) for the key of D. Each key is analyzed by a separate tree, and the two trees have one leaf in common: the pivot.

6 Conclusions

The original objective of this investigation was to use formal grammars to analyze the harmonic structure of a musical fragment, studying their scope and limitations. Nevertheless, during the project we discovered that most of the limitations of the grammar come from the inherent ambiguous nature of music. Thus, it was not possible to adapt the grammar to analyze these ambiguous musical concepts. Particularly, we studied deeply the notion of modulation and, after concluding that formal context-free grammars could not represent it, we decided to use other methods to analyze modulations in order to complement the formal grammars for the harmonic analysis. Due to the different functions and importances of chords within the different tonalities, we decided to use an evidence-gathering numerical method to detect the different tonalities of a musical fragment and therefore the modulations.

We ended with a hybrid method where we first detect the modulations with the numerical method and then, based on the detected tonalities, we analyze their harmonic structure using the formal grammars. The results were really successful as we could analyze different harmonic properties using the methods that better fit each one and combine them obtaining really accurate analysis. This mixed approach could be extended using the methods described on this paper or other methods (as Markov chains, genetic algorithms, machine learning, etc.) to analyze other ambiguous musical concepts. Hence, this is a very promising research area that could allow us to have more powerful analysis tools.

With regard to the numerical method for detecting modulations, it allowed us to detect tonalities with a small error using

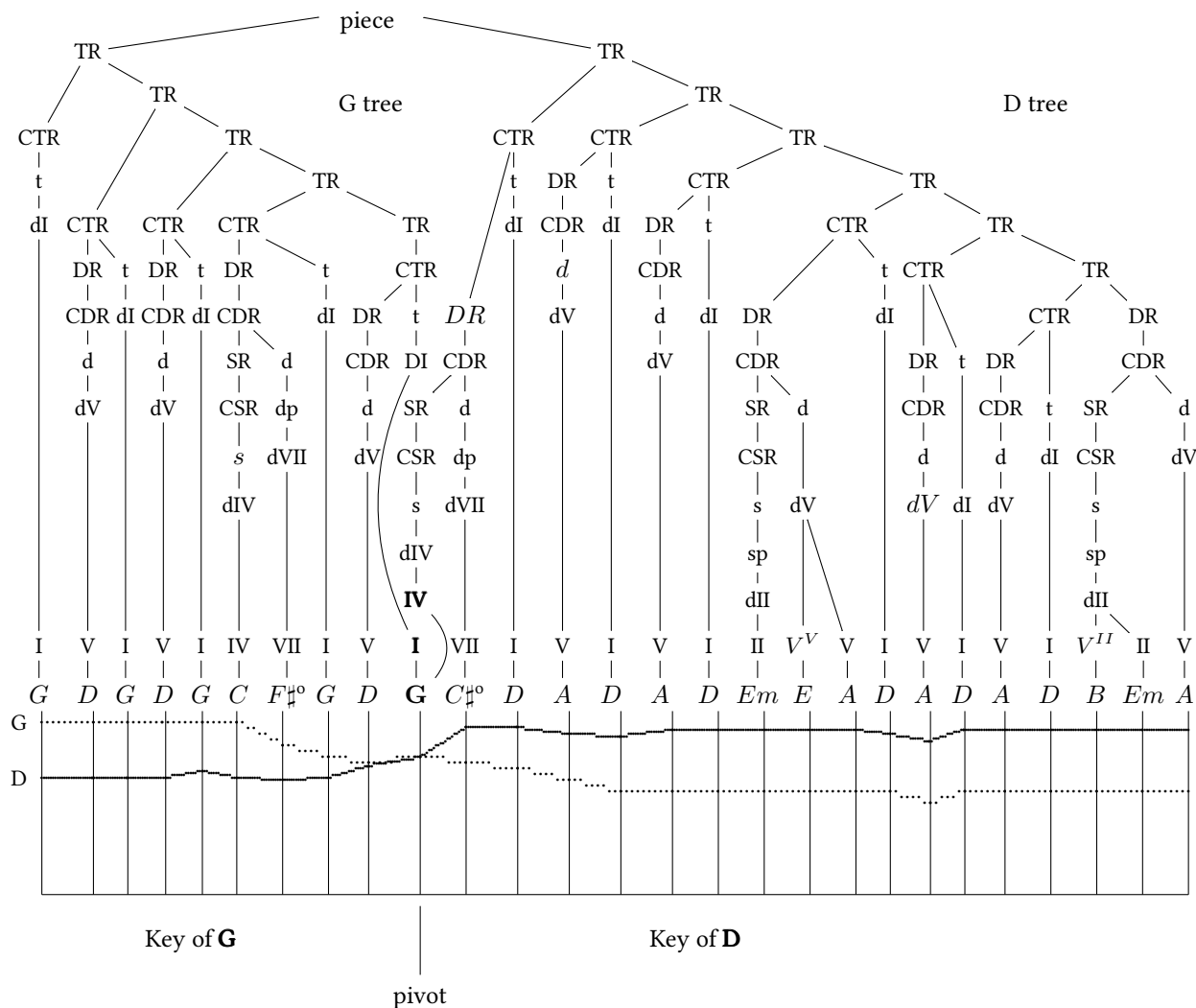


Figure 9: Complete analysis of Mozart's *Eine kleine Nachtmusik*. At the bottom of the figure we show the scores at the various chords of the music. For the sake of clarity, we only show the scores of the keys of **G** and **D**. The chord indicated as *pivot* marks the change in key from **G** to **D**. The top part of the figure contains the trees that analyze the two parts. The harmonic functions (line just above the chords) have been assigned considering the chord and the key using a simple correspondence table. Note that the pivot is assigned two harmonic functions, as it is analyzed in both keys: it is *I* (the tonic) in the key of **G** and *IV* (the sub-dominant) in the key of **D**. This causes that both trees have a branch ending in that chord.

the context: how well the chords were combined with each other in each of the 12 tonalities. This is a very encouraging method that could be adapted to analyze other musical properties that depend deeply on the context, as the descending fifths and harmonic sequences. We would build up evidence for the various competing hypothesis and adopt the solution best supported by the evidence. Once the sequence has been divided into unambiguous parts, grammars can be used to analyze them.

It is important to recall that the musical target of the investigation was the tonal music of the XVII, XVIII and early XIX century, developed mostly in Europe. The harmony of this music is organized around tonalities and follows a particular set of rules that differ from other musical styles. This makes the developed methods of this project fail when they analyze musical fragments from styles that are very different from the chosen one. The grammar rules are based on the common chord progressions and relations of tonal music, so it cannot properly analyze musical fragments that follow other relations, as jazz music, and even that have different chords, as Arab music. Moreover, the numerical method for detecting modulations is based on the tonalities and hierarchy of chords in them, so applying it to other musical styles, as modal music (or quasi-modal music, as flamenco), does not make any sense, as modal music is based on modes instead of tonalities. Thus, the method was thought and works great with tonal music, with a certain melodic and harmonic structure.

Lastly, it is important to highlight the usefulness of understanding the harmonic structure that underlies music and the importance that harmonic analysis has in music. This way, we are able to know how music, as media, communicates and therefore we could be able to use it better along other mediums such as video and audio. It is not only a way of better understanding the nature of music and feelings, it could also be applied in the process of automatic generation of music and we could combine it better with other multimedia systems.

References

- [1] Charles Ames. Automated composition in retrospect: 1956-1986. *Leonardo*, 20(2):169–185, 1987.
- [2] Charles Ames. The markov process as a compositional model: A survey and tutorial. *Leonardo*, 22(2):175–187, 1989.
- [3] Torsten Anders and Eduardo R Miranda. Constraint-based composition in realtime. In *ICMC*, 2008.
- [4] Jean Bresson, Carlos Agon, and Gérard Assayag. Openmusic–visual programming environment for music composition, analysis and research. In *ACM MultiMedia (MM’11)*, 2010.
- [5] Wei Chai and Barry Vercoe. Music thumbnailing via structural analysis. In *Proceedings of the eleventh ACM international conference on Multimedia*, pages 223–226. ACM, 2003.
- [6] Jon Gillick, Kevin Tang, and Robert M Keller. Machine learning of jazz grammars. *Computer Music Journal*, 34(3):56–66, 2010.
- [7] Daniel Harasim, Martin Rohrmeier, and Timothy O’Donnell. A generalized parsing framework for generative models of harmonic syntax. In *Proceedings of the 19th ISMIR Conference*, Paris, France, September 2018.
- [8] Rowan Hooper. Ada lovelace: My brain is more than merely mortal. *New Scientist*, 216(2886):29, 2012.

- [9] Robert M Keller and David R Morrison. A grammatical approach to automatic improvisation. In *Proceedings of the 4th Sound and Music Computing Conference*, Lefkada, Greece, July 2007.
- [10] Fang-Fei Kuo, Meng-Fen Chiang, Man-Kwan Shan, and Suh-Yin Lee. Emotion-based music recommendation by association discovery from film music. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 507–510. ACM, 2005.
- [11] Michael S Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 2(1):1–19, 2006.
- [12] Luigi Federico Menabrea and Ada Lovelace. Sketch of the analytical engine invented by charles babbage, 1842.
- [13] George Papadopoulos and Geraint Wiggins. AI methods for algorithmic composition: A survey, a critical view and future prospects. In *AISB Symposium on Musical Creativity*, volume 124, pages 110–117. Edinburgh, UK, 1999.
- [14] Walter Piston. *Harmony*. W W Norton & co., fifth edition, March 1987. Trad. Esp. Walter Piston. Armonía. Idea Books, 2001.
- [15] Martin Rohrmeier. A generative grammar approach to diatonic harmonic structure. In *Proceedings of the 4th sound and music computing conference*, pages 97–100, Lefkada, Greece, July 2007.
- [16] Martin Rohrmeier. Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1):35–53, 2011.
- [17] Horacio Alberto García Salas, Alexander Gelbukh, and Hiram Calvo. Music composition based on linguistic approach. In *Mexican International Conference on Artificial Intelligence*, pages 117–128. Springer, 2010.
- [18] Mark J Steedman. A generative grammar for jazz chord sequences. *Music Perception: An Interdisciplinary Journal*, 2(1):52–77, 1984.