Mathematical Morphology Operators for Harmonic Analysis

Gonzalo Romero-García, Isabelle Bloch, Carlos Agon

Mathematical Morphology Operators for Harmonic Analysis*

Gonzalo Romero-García[1][0000−0003−1971−6204], Isabelle Bloch[2][0000−0002−6984−1532], and Carlos Agón[1][0000−0001−7926−2537]

[1] Sorbonne Université, CNRS, IRCAM, STMS, Paris, France
{romero,agonc}@ircam.fr
[2] Sorbonne Université, CNRS, LIP6, Paris, France
isabelle.bloch@sorbonne-universite.fr

Abstract. Mathematical Morphology provides powerful tools for image processing, analysis and understanding. In this paper, we apply these tools to analyze scores, that are image-like representations of Music. To do that, we consider *chroma rolls*, a representation of scores similar to piano rolls that use chromas instead of pitches. Endowing this representation with a lattice structure, one can define Mathematical Morphology operators, and setting a group structure to the Time-Frequency plane allows us to use the notion of *structuring element*. We show throughout some examples that this relates with the notion of pitch-class set and chord progressions, and we analyze two Chopin's Nocturnes with this technique.

Keywords: Mathematical Morphology · Harmonic Analysis · Time-Frequency · Pitch-class Set · Chord Progression · Chroma Roll.

1 Introduction

Mathematical Morphology is a theory for the analysis and processing of geometrical structures that intersects with several domains such as Topology, Lattice Theory or Integral Geometry [6, 9]. It has been developed in the second half of the 20$^{\text{th}}$ century and it has been extensively used for analyzing images. Following this approach, we will show that Mathematical Morphology operators may be useful for analyzing image-like representations of Music such as scores. In particular, we will focus on harmonic analysis and we will take as support a two-dimensional representation of the score closely related to the standard piano roll: the *chroma roll*.

There are already a few applications of Mathematical Morphology to Music [3, 4], and [5] on this topic that use piano rolls. This paper extends these earlier works in terms of both representations of music and types of operators.

There are several frameworks where Mathematical Morphology may be applied; in its deterministic form, we need to have a complete lattice as algebraic

* This research is supported by European Research Council ERC-ADG-883313 REACH, and Agence Nationale de la Recherche ANR-19-CE33-0010 MERCI.

structure. We will ask a bit more, intending to make it possible to work with structuring elements: the lattice will then consist of functions with a group as domain and a complete lattice as codomain. Moreover, since we will make abstraction of the dynamics of the score, we will select as codomain the lattice $(\{0,1\},\leq)$, and then we will be able to identify functions with subsets of the group through the characteristic function.

In this paper, we will start by presenting the algebraic framework of the Time-Frequency plane based on Group Theory. Then, we will present the Mathematical Morphology basics that we will use in this work. Finally, we will propose some applications of Mathematical Morphology operators to extract harmonic information. As a particular case, we will focus on Chopin's Nocturnes since they are well adapted to this task.

2 Algebraic Framework

Since we want to model scores, our domain is the Cartesian product of a set modeling time and a set modeling frequency. We require these sets to be groups in order to have a notion of *translation*; this will lead to straightforward definitions of the morphological operators.

We recall that the translation action of a group $(G,+)$ on itself is the function

$$\begin{aligned} T : G \times G &\to G \\ (g,x) &\mapsto T_g x = x+g \end{aligned} \quad . \tag{1}$$

In the following sections, we present the corresponding groups we chose for modeling time and frequency, that lead to a time-frequency representation of a score.

2.1 Time Group

There are several groups that may model the time: if we express it in seconds we can choose $(\mathbb{R},+)$; if we consider the time inside a MIDI file it will be useful to use $(\mathbb{Z},+)$ where each unit is a tick; if we consider an audio signal we may also use $(\mathbb{Z},+)$ where each unit is a sample.

When we measure the time inside a score, one can choose several units: among others, we may use either the bar, the beat or the *tatum*[3]. We choose to measure the time inside a score in bars. Letting $d \in \mathbb{N}$ be the ratio between the duration of a bar and the duration of the tatum, we model the time inside a score measured in bars by the group $(\frac{1}{d}\mathbb{Z},+)$, where $\frac{1}{d}\mathbb{Z} = \{\frac{n}{d} \in \mathbb{Q} : n \in \mathbb{Z}\}$. This group is a subgroup of $(\mathbb{Q},+)$ but is isomorphic to $(\mathbb{Z},+)$ using the isomorphism $\phi : \left(\frac{1}{d}\mathbb{Z},+\right) \to (\mathbb{Z},+)$, $\frac{n}{d} \mapsto n$. Nevertheless, it is interesting to use it rather than $(\mathbb{Z},+)$ because measuring the time in multiples of tatum is far less practical that doing it in bars.

[3] This name was introduced in [1] for calling the maximal note duration such that all the note durations in the score are integer multiples of it. For a deeper discussion about it and its relation with the notion of GCD (Greatest Common Divisor) see [7].

2.2 Frequency Groups

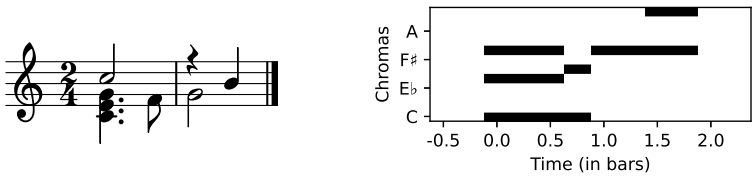
For considering the frequency inside a score, we may use the notion of pitch. Usually, pitch is measured in semitones and may be modeled by the set $\mathbb{Z}$ where each number corresponds to the MIDI number of a note (for instance, the number 60 corresponds to C_4).

This frequency group is interesting for analyzing melodies but, when considering harmony, it would be preferable to work with the set $\mathbb{Z}_{12}$ of *chromas*. Then, we use this set with the translation action of $(\mathbb{Z}_{12}, +)$. In order to make the distinction between chromas and translations, we will use letters for the chromas (C, C♯, ..., B) and numbers for the chroma shifts $(\overline{0}, \overline{1}, ..., \overline{11})$, even if chromas have an associated number for computations and are considered as elements of $\mathbb{Z}_{12}$.

2.3 Time-Frequency Groups

Now that we have set the models for time and frequency, we can couple them by means of the direct product and have a time-frequency representation of scores. We then work with the group $(\frac{1}{d}\mathbb{Z} \times \mathbb{Z}_{12}, +)$ and the action on itself. A score may then be conceived as a subset of the time-frequency plane; piano roll representations are a good way of visualizing this, but they are subsets of $\frac{1}{d}\mathbb{Z} \times \mathbb{Z}$; since we want to work with the time-frequency plane $\frac{1}{d}\mathbb{Z} \times \mathbb{Z}_{12}$, we call *chroma roll* a subset of it. The way of getting a chroma roll from a score is the following: we take the piano roll representation of the score that we call $S \subseteq \frac{1}{d}\mathbb{Z} \times \mathbb{Z}$ and we say that its chroma roll representation is the set $\pi(S) \subseteq \frac{1}{d}\mathbb{Z} \times \mathbb{Z}_{12}$, where $\pi : \mathbb{Z} \to \mathbb{Z}_{12}, n \mapsto \overline{n}$ is the canonical projection from $\mathbb{Z}$ to $\mathbb{Z}_{12}$. See Figure 1 for an illustration of a chroma roll.

Now that we have an image-like representation of the music with a base group structure, we are able to define some Mathematical Morphology operators that allow us to extract harmonic features of the score.



(a) Score (b) Chroma roll

Fig. 1. Chroma roll representation of a score

3 Mathematical Morphology

Let us now introduce some basic but useful tools of Mathematical Morphology. We start by the two basic operators, which are the *erosion* and the *dilation*.

3.1 Erosion and dilation

Erosion and dilation are defined in the context of complete lattices. A good introduction to these operators where all the subsequent definitions and propositions may be found is [2].

Definition 1. *Let $(L_1, \leq_1)$ and $(L_2, \leq_2)$ be two complete lattices, and $\wedge_1, \wedge_2$ (respectively $\vee_1, \vee_2$) the associated infimum (respectively supremum).*

An operator $\varepsilon : L_2 \to L_1$ is called an ***erosion*** *if*

$$\forall X_2 \subseteq L_2 \,,\, \bigwedge\nolimits_1 \varepsilon(X_2) = \varepsilon\left(\bigwedge\nolimits_2 X_2\right) . \tag{2}$$

An operator $\delta : L_1 \to L_2$ is called a ***dilation*** *if*

$$\forall X_1 \subseteq L_1 \,,\, \bigvee\nolimits_2 \delta(X_1) = \delta\left(\bigvee\nolimits_1 X_1\right) . \tag{3}$$

An important property of these operators is that they are *increasing*[4] and then preserve the order.

Whereas this is the most general definition for erosion and dilation, it is an implicit definition. In order to have an explicit type of erosion and dilation (the one with structuring elements), we choose as lattice the power set of a group $(G, +)$. The next proposition provides an explicit definition these types of operators.

Proposition 1. *Let $(G, +)$ be an additive group. Let $T : G \times G \to G, (x, g) \mapsto T_g x = x + g$ be the translation action of $(G, +)$ on itself. We note the translation of a set $A \subseteq G$ by an element $g \in G$ by $T_g A = \{T_g a \in G : a \in A\} \subseteq G$. Let $B \subseteq G$ and $\check{B} = \{-b \in G : b \in B\}$, where $-b$ is the inverse element of b.*

We consider the complete lattice $(\mathcal{P}(G), \subseteq)$. Then, the operators

$$\begin{aligned} \varepsilon_B : \mathcal{P}(G) &\to \mathcal{P}(G) \\ A &\mapsto \varepsilon_B(A) = \{x \in G : T_x B \subseteq A\} \end{aligned} \tag{4}$$

$$\text{and} \quad \begin{aligned} \delta_B : \mathcal{P}(G) &\to \mathcal{P}(G) \\ A &\mapsto \delta_B(A) = \{x \in G : T_x \check{B} \cap A \neq \emptyset\} \end{aligned} \tag{5}$$

are respectively an erosion and a dilation.

We use the following notations: $\forall A \subseteq G$,

$$\varepsilon_B(A) = A \ominus B \qquad\qquad \delta_B(A) = A \oplus B \,.$$

The set B is called structuring element*. The operators ε_B and δ_B are called respectively* binary erosion *and* binary dilation *with structuring element B.*

[4] $\psi : L_1 \to L_2$ is said to be increasing if $\forall X, Y \in L_1$, $X \leq_1 Y \Rightarrow \psi(X) \leq_2 \psi(Y)$.

We notice that for each choice of B we have a different erosion and a different dilation, so, what we actually have is a family of erosions and dilations.

The next property exposes an analogous way of defining these operators making use of union and intersection.

Proposition 2. *Using the notations from Proposition 1,*

$$A \ominus B = \bigcap_{x \in \check{B}} T_x A \qquad\qquad A \oplus B = \bigcup_{x \in B} T_x A \,. \tag{6}$$

Some interesting properties of these erosions and dilations are shown in the next proposition.

Proposition 3. *Let $(G, +)$ be a group. We have*

1. $\forall A \subseteq G$, $\forall B_1, B_2 \subseteq G$, $B_1 \subseteq B_2$,
   (a) $\varepsilon_{B_1}(A) \supseteq \varepsilon_{B_2}(A)$
   (b) $\delta_{B_1}(A) \subseteq \delta_{B_2}(A)$.
2. $\forall A \subseteq G$, $\forall B \subseteq G$,

$$(A \oplus B)^c = A^c \ominus \check{B} \qquad\qquad (A \ominus B)^c = A^c \oplus \check{B} \,, \tag{7}$$

where $A^c = \{x \in G : x \notin A\}$ is the complementary set of A.

Combining these operators, new morphological operators arise, in particular *opening* and *closing*.

3.2 Opening and Closing

As erosion and dilation, opening and closing have a general definition in the context of complete lattices. A good introduction to these operators is [8].

Definition 2. *Let $(L, \leq)$ be a lattice. Let $\psi : L \to L$ an operator. We say that*

1. *ψ is an* opening *if it is increasing, anti-extensive*[5] *and idempotent*[6]*;*
2. *ψ is a* closing *if it is increasing, extensive*[7] *and idempotent*[6]*.*

Since they are idempotent and increasing, openings and closings are *morphological filters*. Idempotence is particularly useful since it ensures that several applications of the same operator do not change the result (as do a lot of other filters). This can be thought as a guarantee of filtering all that we want to filter at once.

A particular case of openings and closings are the compositions of erosions and dilations with respect to a structuring element, as shown in the next proposition.

[5] $\forall X \in L$, $\psi(X) \leq X$.
[6] $\psi^2 = \psi$.
[7] $\forall X \in L$, $X \leq \psi(X)$.

Proposition 4. *Let $(G,+)$ be a group. Let $B \subseteq G$. Let ε_B and δ_B be respectively the erosion and dilation with respect to the structuring element B. Then,*

1. $\gamma_B \triangleq \delta_B \circ \varepsilon_B$ *is an opening.*
2. $\varphi_B \triangleq \varepsilon_B \circ \delta_B$ *is a closing.*

We use the following notations: $\forall A \subseteq G$,

$$\gamma_B(A) = A \circ B \qquad\qquad \varphi_B(A) = A \bullet B\,.$$

There are many other morphological operators, but in this paper we will only expose a last one: the *hit-or-miss transform*.

3.3 Hit-or-miss Transform

The hit-or-miss transform is defined as follows.

Definition 3. *Let $(G,+)$ be a group. Let $C, D \subseteq G$ with $C \cap D = \emptyset$. Then, the* hit-or-miss transform *of $A \subseteq G$ with respect to the structuring elements C and D is defined by*

$$A \odot (C,D) = (A \ominus C) \cap (A^c \ominus D). \tag{8}$$

An interesting way of seeing the hit-or-miss transform is given in the next remark and will be the one in which we will think when applying it to Music.

Remark 1. Using the notations from Definition 3,

$$A \odot (C,D) = \{p \in G : T_p C \subseteq A \subseteq T_p D^c\}\,. \tag{9}$$

We may then think the hit-or-miss transform as a pattern matching transformation with an internal condition ($T_p C \subseteq A$) and an external condition ($A \subseteq T_p D^c$); we want to find structures that are bigger than C but smaller than D^c.

4 Applications

Now that we have some Mathematical Morphology operators, let us give some applications. We propose the two following linked steps: first, we consider the lattice formed by the power set of the group $(\mathbb{Z}_{12},+)$ ordered by inclusion. Then, we extend it with time information to be able to analyze a score. We show that $\mathbb{Z}_{12}$ structuring elements correspond to pitch-class sets, and $\frac{1}{d}\mathbb{Z}\times\mathbb{Z}_{12}$ structuring elements correspond to chord progressions.

4.1 Mathematical Morphology on $\mathbb{Z}_{12}$

Musical chords may be identified with subsets of $\mathbb{Z}_{12}$. We can identify two subsets by the fact that they are both major chords and this is modeled in mathematical terms by the concept of equivalence relation, in this case $\forall A, B \subseteq \mathbb{Z}_{12}, A \sim B \Leftrightarrow \exists p \in G$ such that $T_p A = B$. Mathematical Morphology comes in handy for analyzing this fact as shown in the following.

Erosion as a structure detector Let us focus first in the way we can use erosion to detect chord types; if we want to detect, for instance, major chords, we may erode a score with the major chord structure as illustrated in the following example.

Example 1. Let us consider the following chords:

$$\mathrm{Fmaj} = \{\mathrm{C}, \mathrm{F}, \mathrm{A}\} \qquad \mathrm{Fmaj}^7 = \{\mathrm{C}, \mathrm{E}\flat, \mathrm{F}, \mathrm{A}\} \qquad \mathrm{F}^\circ = \{\mathrm{F}, \mathrm{A}\flat, \mathrm{B}\}\,.$$

We consider the structuring element $\mathrm{maj} = \{\overline{0}, \overline{4}, \overline{7}\}$. Then,

$$\mathrm{Fmaj} \ominus \mathrm{maj} = \{\mathrm{F}\} \qquad \mathrm{Fmaj}^7 \ominus \mathrm{maj} = \{\mathrm{F}\} \qquad \mathrm{F}^\circ \ominus \mathrm{maj} = \emptyset\,.$$

We see that erosion may be used as a detector: it detects the presence of a major chord with root F in Fmaj and Fmaj^7 and shows that there is no major chord in F°. The fact that we got the chroma F after the erosion is because the zero corresponds to the root in the structuring element; it serves as a reference.

An interesting case is when we deal with chords that have non trivial stabilizer[8] in $\mathbb{Z}_{12}$. An example is the diminished seventh chord: if we erode the chord $\mathrm{D}^{\circ 7} = \{\mathrm{D}, \mathrm{F}, \mathrm{A}\flat, \mathrm{B}\}$ by the diminished seventh structuring element $\mathrm{dim7} = \{\overline{0}, \overline{3}, \overline{6}, \overline{9}\}$ it remains unchanged, *i.e.*:

$$\mathrm{D}^{\circ 7} \ominus \mathrm{dim7} = \mathrm{D}^{\circ 7}\,. \tag{10}$$

This points out the fact that the diminished seventh chord may be built on each of its notes or, equivalently, that any of its notes may be considered the root. In Messiaen's terminology, it has a limited number of transpositions equal to the cardinal of the quotient group

$$\mathbb{Z}_{12}\!\diagup_{\mathrm{Stab}(\mathrm{D}^{\circ 7})} = \{\mathrm{C} + \mathrm{dim7}, \mathrm{C}\sharp + \mathrm{dim7}, \mathrm{D} + \mathrm{dim7}\}\ , \tag{11}$$

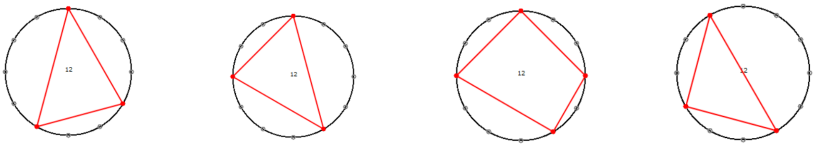
which in this case is 3.

Opening as a simplification Let us consider now opening in the same case that in Example 1; we have:

$$\mathrm{Fmaj} \circ \mathrm{maj} = \mathrm{Fmaj} \qquad \mathrm{Fmaj}^7 \circ \mathrm{maj} = \mathrm{Fmaj} \qquad \mathrm{F}^\circ \circ \mathrm{maj} = \emptyset\,.$$

We see that opening simplifies chords containing a major chord (removing the seventh in the case of Fmaj^7) and removes all the chords that do not contain a major chord. This illustrates why they are called filters.

[8] The stabilizer of a subset $A \subseteq \mathbb{Z}_{12}$ is defined by $\mathrm{Stab}(A) = \{n \in \mathbb{Z}_{12} : T_n A = A\}$ and is a subgroup of $\mathbb{Z}_{12}$.



(a) Structuring element maj (b) Fmaj (c) Fmaj^7 (d) F$^\circ$

Fig. 2. Chords and structuring element represented in $\mathbb{Z}_{12}$.

Hit-or-miss transform as a pattern detector Hit-or-miss transform is closely related to erosion; indeed, $\forall A, B \subseteq G$,

$$A \ominus B = A \odot (B, \emptyset)\,. \tag{12}$$

In addition, the second structuring element plays an interesting role: if we want to know to which major and minor scales does the Fmaj chord belong, we may proceed as in the following example.

Example 2. Let us consider the structuring elements

$$\text{majorScale} = \{\overline{0}, \overline{2}, \overline{4}, \overline{5}, \overline{7}, \overline{9}, \overline{11}\} \qquad \text{minorScale} = \{\overline{0}, \overline{2}, \overline{3}, \overline{5}, \overline{7}, \overline{8}, \overline{10}\}\,.$$

Then, we have

$$\text{Fmaj} \odot (\emptyset, \text{majorScale}^c)) = \{\text{C}, \text{F}, \text{B}\flat\} \quad \text{Fmaj} \odot (\emptyset, \text{minorScale}^c) = \{\text{D}, \text{G}, \text{A}\}\,.$$

This shows that the F major chord is present in the C, F and B♭ major scales and in the D, G and A minor scales.

It is interesting to note that the major and (natural) minor scales are equivalent in the sense that one is the translation of the other. However, the reference note is not the same and thus gives different results.

Until now, we have used the hit-or-miss transform with one of its structuring elements being the empty set. Let us show an example where we use both constraints for detecting a range of chords that hold both bigger-than and smaller-than constraints.

To detect whether a chord X is a dominant-like chord of some major scale, *i.e.* it contains the dominant third $\text{Dom}^{3\text{M}} = \{\overline{7}, \overline{11}\}$ and all its notes are contained in the major scale, the hit-or-miss transform can be used as follows:

$$X \odot (\text{Dom}^{3\text{M}}, \text{majorScale}^c)\,. \tag{13}$$

This is illustrated in the following example with three choices of X among the 4096 possible combinations[9].

[9] This number comes from the number of subsets of $\mathbb{Z}_{12}$ that is equal to 2^{12}.

Example 3. We consider the following chords:

$$\text{Emaj} = \{\text{E}, \text{G}\sharp, \text{B}\} \quad \text{Dmaj}^{7(5)} = \{\text{C}, \text{D}, \text{F}\sharp\} \quad \text{Amaj}^{7\flat 9} = \{\text{C}\sharp, \text{E}, \text{G}, \text{A}, \text{B}\flat\}\,.$$

Then, we have

- $\text{Emaj} \odot (\text{Dom}^{3\text{M}}, \text{majorScale}^c) = \{\text{A}\}$,
- $\text{Dmaj}^{7(5)} \odot (\text{Dom}^{3\text{M}}, \text{majorScale}^c) = \{\text{G}\}$,
- $\text{Amaj}^{7\flat 9} \odot (\text{Dom}^{3\text{M}}, \text{majorScale}^c) = \emptyset$,

which shows that Emaj is a dominant-like chord of the A major scale, $\text{Dmaj}^{7(5)}$ is a dominant-like chord of the G major scale and $\text{Amaj}^{7\flat 9}$ is not a dominant-like chord of any major scale.

As a last application of the hit-or-miss transform, we may use it to detect any chord type $C \subseteq \mathbb{Z}_{12}$ by choosing the structuring elements C and C^c.

4.2 Mathematical Morphology on Chroma Rolls

We now go one step further, and illustrate how Mathematical Morphology operators can be useful to analyze scores represented as chroma rolls by using the group $(\frac{1}{d}\mathbb{Z} \times \mathbb{Z}_{12}, +)$ and the lattice $\left(\mathcal{P}\left(\frac{1}{d}\mathbb{Z} \times \mathbb{Z}_{12}\right), \subseteq\right)$.

Supremum windowing for collapsing harmony When analyzing harmonically a score, we should make decisions about which notes do we keep and which ones do we discard. Also, we should decide how to handle notes spread in time: usually, we consider that several notes that do not coexist at the same time but are close enough to each other are related because they are part of the same chord, like in the case of arpeggios. To take such configurations into account, we propose a pre-processing step to modify the chroma rolls such that they become more representative of the harmony of the piece.

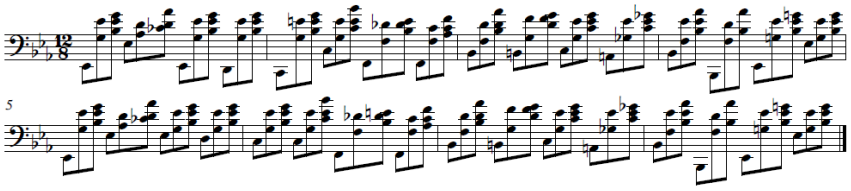
It is clear that the composer's own style and particularities could lead to different rules. In this paper, as an example, we focus on Chopin's Nocturnes since they are a good illustration of our purpose. The proposed pre-processing method consists in taking only the left hand, since the harmony is concentrated there, and applying what we call a *supremum windowing*.

The supremum windowing is a way of collapsing several notes into the same chord by taking the supremum. Its formula is presented in the next definition.

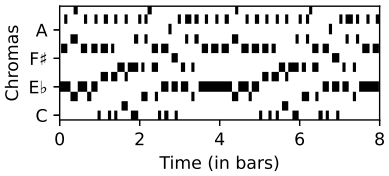
Definition 4. *Let $S \subseteq \frac{1}{d}\mathbb{Z} \times \mathbb{Z}_{12}$. Let $\chi_S \in \{0,1\}^{\frac{1}{d}\mathbb{Z}\times\mathbb{Z}_{12}}$ be the characteristic function of S. The* supremum windowing *of S with window length $L \in \mathbb{Q}$ and hop size $H \in \mathbb{Q}$ is the subset of $\frac{H}{d}\mathbb{Z} \times \mathbb{Z}_{12}$ with characteristic function*

$$\begin{aligned} W_\infty^{(L,H)}[S] : \tfrac{H}{d}\mathbb{Z} \times \mathbb{Z}_{12} &\to \{0,1\} \\ (t,c) \quad &\mapsto \sup\{\chi_S(t+x,c) : x \in [0,L)\} \end{aligned} \quad . \tag{14}$$

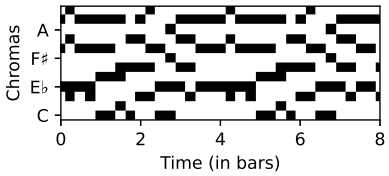
Figure 3 shows the supremum windowing of the left hand of Chopin's Nocturne Op. 9 No. 2.



(a) Original score



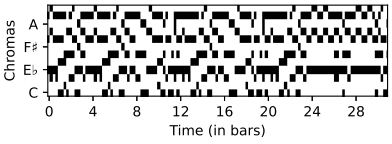
(b) Resulting chroma roll



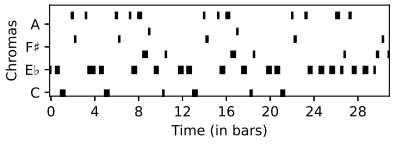
(c) Windowed chroma roll

Fig. 3. Score, chroma roll and its supremum windowed version of the first 8 bars of the left hand of Chopin's Nocturne Op. 9 No. 2 (dynamics, expressions and other features are not considered).

Chord type detection Let us now describe some applications of Mathematical Morphology operators for the detection of chord types in scores. We can extend all that we have done in the one-dimensional case to the two dimensional case; for instance, if we want to detect major chords, we may use an erosion with structuring element $\{0\}\times$maj. This is done for the (supremum windowed) chroma roll of the Nocturne Op. 9 No. 2 of Chopin in Figure 4.



(a) Chroma roll

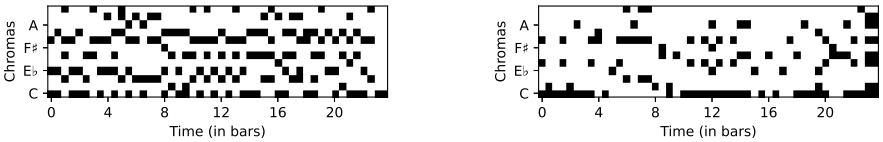


(b) Erosion of the chroma roll

Fig. 4. Chroma roll of the Nocturne Op. 9 No. 2 of Chopin and its erosion by the structuring element $\{0\} \times$ maj.

We may also apply a process analogous to the one in Example 2 to detect tonality and modulations throughout the score. This is illustrated in Figure 5 in the case of the first 24 bars of the Chopin's Nocturne Op. 48 No. 1 after supremum windowing with length and hop size equal to $\frac{1}{2}$ of the bar. We apply

a hit-or-miss transform with $C = \emptyset$ and $D = (\{0\} \times \text{minorHarmScale})^c$, where minorHarmScale $= \{\overline{0}, \overline{2}, \overline{3}, \overline{5}, \overline{7}, \overline{8}, \overline{11}\}$ is the structuring element corresponding to the harmonic minor scale. We notice several parts that are clearly in C minor and a part around bars 6 to 8 that is in G minor.

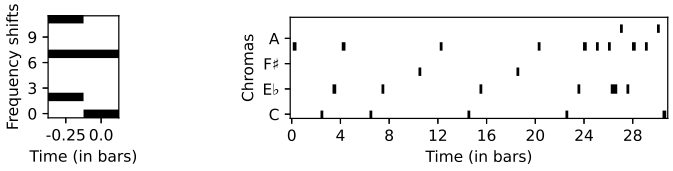


(a) Chroma roll

(b) Hit-or-miss transform with structuring elements $C = \emptyset$ and $D = (\{0\} \times \text{minorHarmScale})^c$

Fig. 5. Hit-or-miss transform of the chroma roll of the first 24 bars of the Chopin's Nocturne Op. 48 No. 1 for detecting potential harmonic minor scales.

Chord progression detection Extending structuring elements with a time component allows us to detect chord progressions. For instance, we may detect authentic cadences using the structuring element V-I $= \left(\{-\frac{1}{h}\} \times \{\overline{2}, \overline{7}, \overline{11}\}\right) \cup \left(\{0\} \times \{\overline{0}, \overline{7}\}\right)$, where $\frac{1}{h}$ models the harmonic rhythm. Here, the interval $\{\overline{0}, \overline{7}\}$ guarantees that we recover both the major and minor I degree, and we set the timestamp of the dominant chord $\{\overline{2}, \overline{7}, \overline{11}\}$ to $-\frac{1}{h}$ because we want it to precede the tonic (and thus the minus sign). This way, the result of the erosion provides the time and the chroma corresponding to the I chord. Figure 6 illustrates this for the 24 first bars of the Nocturne Op. 9 No. 2 with $h = 4$.



(a) Structuring element of V-I

(b) Erosion of the chroma roll with structuring element V-I

Fig. 6. Erosion of the chroma roll of Chopin's Nocturne Op. 9 No. 2 by the V-I structuring element.

5 Conclusions

Throughout this paper, we have seen that Mathematical Morphology, based on Group Theory, may be very useful for analyzing Music scores. This framework provides efficient tools for harmonic analysis that are lightweight and well adapted to harmonic considerations. The use of chroma rolls ensures that we do not have disposition problems in the chords. While simple, the proposed examples illustrate the power of the proposed approach.

It is interesting to notice that we have only used erosion-like operators (erosion, opening and hit-or-miss); this is because we focused on the analysis part. Other operators such as dilations and closing could be used for composition; for instance, a dilation reproduces a particular structure from a reference note.

Several paths can be followed for extending these operators to more complex lattices; for instance, if we consider the amplitude range in a score by using dynamics, morphological operators acting on functions can be leveraged. Also, we can change the base groups; for example if we want to analyze MIDI files, we may use $(\mathbb{Z}, +)$, and we may even apply these tools on signal operators such as the Short-Time Fourier Transform. In future works, we will also go deeper in the creation of an automatic analyzer that implements morphological tools based on the equivalence between chord structure and structuring element. We intend to provide a deterministic framework for the study of harmony at large scale.

References

1. Bilmes, J.A.: Timing is of the essence: Perceptual and computational techniques for representing, learning, and reproducing expressive timing in percussive rhythm. Ph.D. thesis, Massachusetts Institute of Technology (1993)
2. Heijmans, H.J.A.M., Ronse, C.: The algebraic basis of mathematical morphology I. Dilations and erosions. Computer Vision, Graphics, and Image Processing **50**(3), 245–295 (1990)
3. Karvonen, M.: Using Mathematical Morphology for Geometric Music Retrieval. Ph.D. thesis, University of Helsinky (2008)
4. Karvonen, M., Laitinen, M., Lemström, K., Vikman, J.: Error-tolerant content-based music-retrieval with mathematical morphology. In: Exploring Music Contents. pp. 321–337. Lecture Notes in Computer Science, Springer-Verlag, Germany (2011)
5. Lascabettes, P., Bloch, I., Agon, C.: Analyse de représentations spatiales de la musique par des opérateurs simples de morphologie mathématique. In: Journées d'Informatique Musicale. Strasbourg (online), France (2020)
6. Matheron, G.: Random Sets and Integral Geometry. Wiley, New-York (1975)
7. Romero-García, G.: Rhythm Transcription and Characterization for Performed Music and Arrhythmia Sequences. Master's thesis, Sorbonne Université, Paris (2020)
8. Ronse, C., Heijmans, H.J.A.M.: The algebraic basis of mathematical morphology: II. Openings and closings. CVGIP: Image Understanding **54**(1), 74–97 (1991)
9. Serra, J.: Image Analysis and Mathematical Morphology. Academic Press, New-York (1982)