THE NOTATION OF SOUND FOR COMPOSITION AND TRANSCRIPTION–AN ADAPTATION OF LASSE THORESEN'S SPECTROMORPHOLOGICAL ANALYSIS

Mattias Sköld
KMH Royal College of Music, Stockholm
KTH Royal Institute of Technology, Stockholm
mattias.skold@kmh.se

ABSTRACT

This paper details my adaptation of Lasse Thoresen's spectromorphological analysis notation for the sake of composition and transcription, re-imagining the analysis symbols for use over a spectrum staff system over which pitch and spectra can be indicated with great detail, and possibly interpreted by musicians and computers for performance. A sound object is notated with regard to its spectral width, density, centroid frequency, significant sound components, modulation and amplitude envelope. It can also have a spectrum reference. The symbols are placed over a spectrum grand-staff with a frequency scale to show each parameter both from a frequency and pitch perspective. Also included are suggestions for the visual representation of spatialisation where positions and movements are displayed in two or three dimensions above the sound notation while constant rotations are notated as modulations.

1. INTRODUCTION

Lasse Thoresen's spectromorphological analysis was the result of years of teaching Schaeffer's typomorphology at the Norwegian State Academy of Music [1]. As part of the development of early electroacoustic music, Schaeffer developed a vocabulary and a typology culminating in the famous TARTYP [2] diagram with 28 categories that in combinations were to describe most if not all sounds of a music not only concerned with pitch structures. This was an important, if not necessary, development with regard to a traditional musicology not ready for a music of recorded trains and saucepans [1]. As the name suggests, Thoresen's research also builds on Denis Smalley's influential theory of *spectromorphology*–a highly developed framework for studying structural relations in music over time.[3].

In order to make practical use of Schaeffer's typomorphology Thoresen streamlined Schaeffer's diagram and reduced the 28 categories to 15, keeping the nine core categories as well as the six extremes of unpredictable nature. New categories were then introduced to fill the gaps between the core categories and the extremes. Also, Schaeffer's normative vocabulary was removed since he had, in his typomorphology, incorporated ideas relating to the suitability of sounds for musical use. But most importantly, graphic symbols were introduced making possible graphic analyses of music with a detail and consistency not possible before. To my knowledge, the *Spectromorphological analysis of sound objects* (hereinafter SASO), as described in [1] by Thoresen assisted by Hedman, remains the most detailed and developed symbolic system for analysis of sound structures to this date.

The background for my adaptation of SASO comes from teaching electroacoustic music since 2004, first at EMS and later at the Royal College of Music in Stockholm. When teaching composition students how to analyse the organisation of sound through aural sonology, I found it fruitful to also focus on the act of organising sound. However, neither Schaeffer's research nor Thoresen and Hedman's development of the same were aimed for composition. Their tools were developed to describe what is heard. Categories and symbols used for the composition of musical structures, on the other hand, are not only descriptive, but also meant to communicate a musical composition for performance. I found that in order to make full use of Thoresen and Hedman's notation in a compositional context, their symbols needed to be translated into acoustic properties that can be communicated and interpreted by musicians and computers. This translation means both reducing and expanding the symbol palette while developing a practice for placing the symbols over a fixed time-frequency-oriented staff system.

2. THE NOTATION SYSTEM IN DETAIL

Unless indicated otherwise, the notation symbols described here were all originally developed by Thoresen assisted by Hedman. Refer to [1] for a more detailed description of their system, here abbreviated as SASO.

2.1 Sound spectrum

2.1.1 The spectrum staff system

The first major adaptation of SASO is the placement of symbols over a hybrid frequency-staff system where specific pitches are easily identified while a frequency scale

[1] *Étude aux chemins de fer* and *Étude aux casseroles* are two parts of Pierre Schaeffer's genre-originating work *Cinq études de bruits*

helps relating spectral data to the actual frequency contents of the sound (see Figure 1). The system covers the complete listening range of our ears but may of course be decimated to using fewer staves if the notated information does not make use of the full spectral range. This kind of grand staff system is common in software for computer-assisted composition for control and/or display of data as pitch, e.g. the *nslider* in Max [2]. One can of course remove the staff systems and rely solely on the frequency scale for music where exact pitch relations have no significance. For such passages I would still use grey horizontal lines to mark the boundaries of each octave.

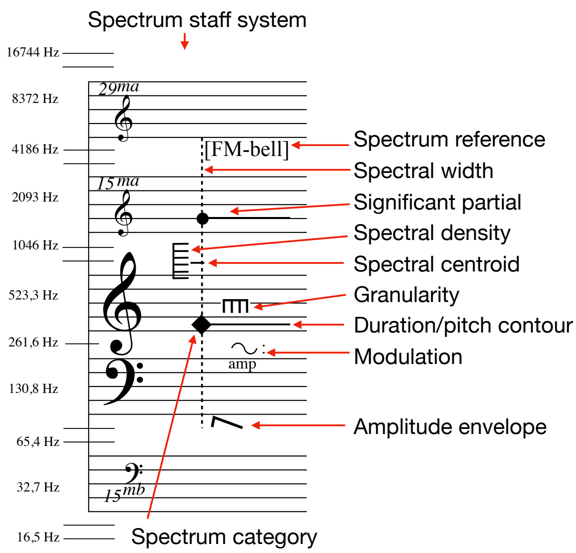

Figure 1. Spectrum staff and an example sound object with several indicators of sound spectrum and energy articulation.

2.1.2 *Spectrum categories*

The three core spectrum categories, pitched, dystonic and complex retain their meaning: *pitched* sounds are sounds with pitch, *dystonic* sounds have inharmonic spectra or are clusters while *complex* sounds have no pitch. We need different strategies for notating the "pitch contour" of these types of spectra since we perceive them differently. Pitched sounds, which have harmonic spectra, are (naturally) notated at the position of the root frequency of the spectrum. Dystonic sounds, which have inharmonic spectra or are clusters of pitched sounds, are placed at the position of the most significant partial. Complex sounds are notated at the position of the spectral centroid, marking the centre frequency of the spectral content of the sound. Figure 2 shows the three main spectrum category symbols and what spectral features that define their placements over the staff system.

2.1.3 *Spectral width*

Different from SASO, spectral width indicates the significant frequency range of a sound's spectrum with a dashed

[2] https://cycling74.com/products/max

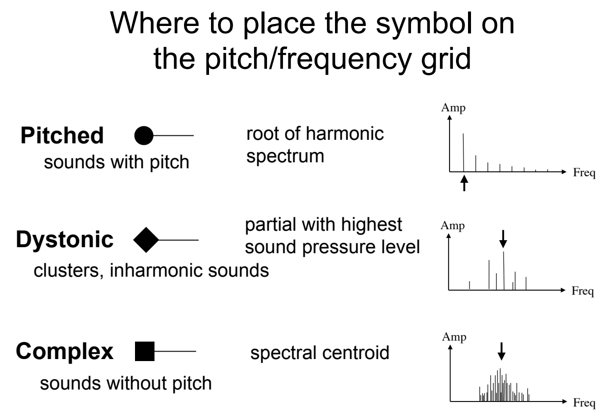

Figure 2. The three spectrum categories, their notation symbols and what spectral feature that determines their vertical positions on the spectrum staff

vertical line across the spectral staff system. For filtered sounds this may be equal to the full frequency range of its spectrum while for non-treated acoustic sounds it makes sense to indicate a frequency range of the relatively louder portion of the spectrum. Figure 3 shows three examples of spectral width: A is a pitched sound with spectral width 440 Hz–2,6 kHz. Since a pitched sound is notated at its root, the spectral width is never below its symbol. B is a dystonic sound with spectral width 164 Hz–2,1 kHz and C is a complex sound of width 147 Hz–1 kHz. The straight horizontal line marking the higher limit of the width of example C indicates that there is no spectral content above this line. In SASO, spectral width is presented as a continuum of different sound spectra from sine tone to white noise [1].

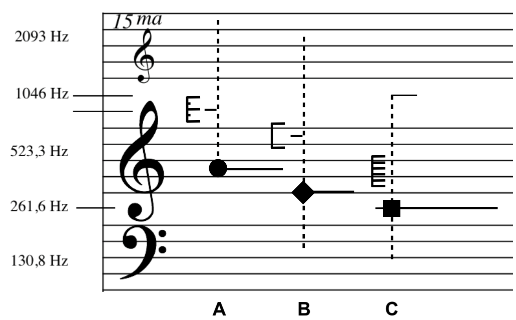

Figure 3. Three examples of spectral width in the pitched (A), dystonic (B) and complex (C) spectrum categories, which are the vertical dashed lines. Spectral density is indicated with the comb-like symbols to the left of the width lines, while spectral centroid is shown as small line indicators on the left side of the width lines for A and B.

2.1.4 *Spectral centroid*

The spectral centroid is the centre frequency of a sound's spectral energy. It is one valuable descriptor for the perceived brightness of a sound, though not the only one. Other factors, such as frequency range, also play a part. It is indicated as a small horizontal line indicator on the left side of the spectral width vertical line. Combined, the width and centroid indicators resemble the symbol for spectral brightness in SASO denoting the perceived brightness of a sound. Figure 3 includes three examples of spectral centroid indications: examples A and B have specific centroid indicators, while C is from the complex spectrum category and has therefore its symbol (a solid square shape) at the centroid frequency position.

2.1.5 Spectral density

A comb-like symbol represents spectral density, as a value of the density of partials with high amplitude in a sound's spectrum, contingent on the sound's spectrum category. For purely pitched sounds (with harmonic spectra) maximum density in terms of positions of partials is dependent on the root frequency ant its multiples, while the spectra of dystonic (inharmonic or clusters) and complex sounds can be saturated to the extent that they eventually turn into noise (both ending up in the complex category). The number of teeth of the vertical comb-like indicator provides a relative value of density from lowest possible (two teeth) to maximum (six teeth). There are also two special cases: a maximum saturated spectrum, i.e. noise, indicated with a thick toothless comb, implying that the teeth are too close to separate, and a particular comb-symbol for indicating spectra with only every other partial–a spectral phenomenon commonly referred to as having a "hollow" sound quality. See Figure 4 for an overview of the density symbols. These are placed over and to the left of the spectrum category symbol, unless there is a spectral centroid indicator in which case the symbol is placed to the left of this indicator. Figure 3 shows three examples of density: A has only every other partial, B has a sparse spectrum while C has a very dense spectrum though it is not pure noise which would have yielded the maximum density symbol. The every-other-partial symbol can also be configured to convey different degrees of density as one sees fit. SASO's equivalent symbol is placed on the extension line of the sound object to indicate spectral saturation [1].

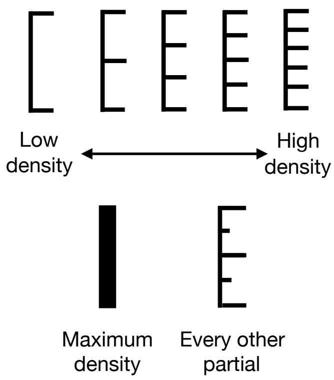

Figure 4. Overview of spectral density symbols

2.1.6 Significant partials and components

Particularly dystonic sounds, like bells with inharmonic spectra, can have multiple significant partials that are clearly audible and whose frequencies are highly relevant when composing for them. While the partial considered the most significant provides the position of the main spectrum symbol, any other significant partials are indicated with small symbols. Any spectrum category symbol can be used to represent significant components of a sound. E.g. an electric fan may have both a noise component and a humming pitched component. When used as a pitched sound for composition, the noise would be considered a complex component of the pitched sound.

2.1.7 Spectrum reference

Not part of SASO, spectrum reference is a text label for the spectrum to indicate a particular spectrum identity. It may be the spectrum recognised from our shared bank of culturally conditioned references such as the sound of an alto saxophone or a large church bell, or it may be a spectrum identity established during the course of a single musical work. These references can be indicated in two ways, either as *being* the spectrum of a known reference or as something *resembling* a certain reference. Both are indicated as text labels within square brackets as shown in Figure 5. The difference is that a reference of resemblance is italicized, within simple quotes and has no capital letter. The reference is positioned to the right of the higher limit of the spectral width vertical line as shown in Figure 1.

[Trumpet]	is the spectrum of a trumpet
['*trumpet*']	sounds like the spectrum of a trumpet

Figure 5. Spectrum reference, indicated as either being or resembling a known spectrum

2.2 Energy articulation

2.2.1 Pitch/spectral contour and extension lines

The extension line from the spectrum category symbol both serves to show how the parameter indicated by the symbol's vertical position changes over time, and the duration of the sound.

Any frequency-dependent sound spectrum indicator, such as the high and/or low value of the spectral width, can have dashed extension lines to indicate changes. These are then treated as time-dependent breakpoints when represented as data.

2.2.2 Granularity and iterative sounds

Granular or iterative sounds is the phenomenon when a chain of rapidly repeated sound grains form one continuous sound. The symbols used are basically the original horizontal comb-like symbols from SASO with different numbers of teeth for different granularity speeds, though with a five-step scale rather than three. The roundness of the symbol's angles can be varied for large, small and moderate granularity coarseness. However, I introduce another level of severe granular coarseness, when there are perceived silences between the grains. This is indicated with the back of the comb-symbol removed to clarify the separation of grains. For Thoresen, when adapting Schaeffer's ideas, granularity and iterative sounds are two different concepts [1] but I find it more useful in this context to treat them as one and the same. See Figure 6 for an overview of granularity symbols of different degrees of coarseness and velocity and Figure 1 for its placement over the extension line. Different sound components may have different granularity symbols.

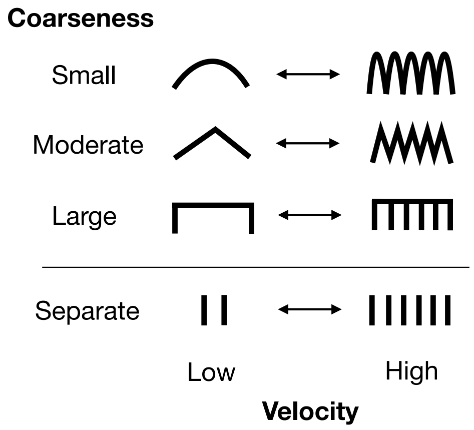

Figure 6. Granularity symbols of different degrees of coarseness and velocity

2.2.3 Accumulation

Accumulations are hordes of sound objects, not to be confused with sounds with granularity where a chain of grains form one continuous sound. The sounds involved in an accumulation are notated as a group of small spectrum category note heads embraced by a bracket with an extension line. The number of note heads included depends on what amount of information is necessary to understand the behaviour of the accumulation. The major difference from SASO is how the placement over the spectrum staff system affects the positioning of the individual sounds included in the accumulation. Figure 7 shows three examples of accumulations. A represents a horde of very short complex sounds, B consists of slightly longer pitched sounds. C also has a random indicator specified for the vertical axis, in this case representing 100 % random positions of the individual sound particles. The extension lines are positioned at resulting spectral centroid frequency of the accumulation. Spectral width is indicated in the same manner as individual sounds, but for the whole accumulation.

For accumulations, levels of randomness can be indicated both for the spectrum and time axes using a percentage value and a question mark with arrows indicating the axis affected as seen in Figure 7. Introducing randomness to the description of textural sounds is in line with the findings of Grill et al where ordered-chaotic and homogenuous-heterogenuous were found to be defining characteristics relating to the perception of sound textures with over 50 % agreement among expert listeners [4].

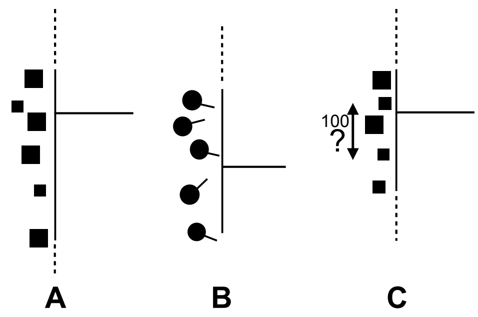

Figure 7. Three examples of accumulations–these can look very different depending on what sound objects are accumulated

2.3 Variation

Godøy and Thoresen both suggest *gait* as the English equivalent of Schaeffer's *allure* - a way of moving forward [5]. Besides granular gait covered above as granularity, SASO has indicators for variation with regard to pitch gait, dynamic gait and spectral gait. Approaching these concepts from a sound synthesis perspective, I choose to treat them as different forms of modulation with a standardised and flexible mode of representation:

2.3.1 Modulation

Modulation is change as articulation rather than structural changes of values. These can be of any kind but common in the music literature are vibrato and tremolo though these terms are not used here since they are ambiguous because of their connections to music instrument practice. A small line shape placed below the extension line of the sound component affected indicates the modulation curve with a short written label below describing the nature of the modulation. The line shape is to be interpreted as describing a change covering the full duration as indicated by the extension line under which it is placed. A small colon mark means a repeated curve/wave which would be the case for a vibrato. Further information can of course be introduced as one sees fit, e.g. a frequency value next to the colon mark specifying the speed of a repeated variation and the height of the symbols can be used to indicate various degrees of modulation. Also, one sound object can have several modulations and these may vary over time.

Since the main contour for pitch (for pitched and dystonic sounds) and centroid (for complex sounds) is indicated by the extension line of the object, modulation of these parameters are notated on the extension line. See Figure 8 for examples of modulation.

2.3.2 Amplitude envelope

The amplitude envelope of a sound is a special case of variation, indicated with a line shape thicker than the modulation shapes and is placed below the whole sound object. No text label is necessary. Figure 8 D shows an example of an amplitude envelope.

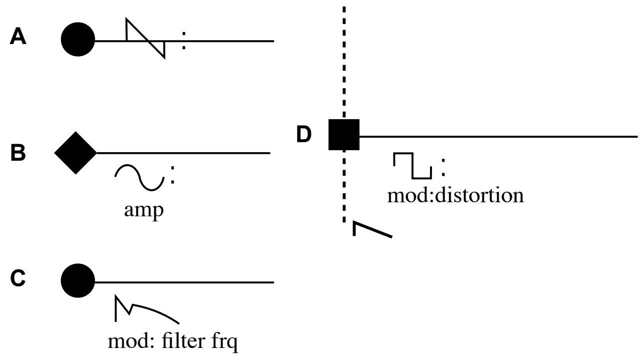

Figure 8. Examples of modulated sounds: A has a repeated sawtooth modulation of pitch, B has a sine wave modulation of amplitude, C has an envelope modulation of a filter extending over the length of the extension line, and D has a pulse wave modulation of distortion as well as an indicator of amplitude envelope at the bottom.

2.4 Rhythm

The visual representation of rhythm was covered in more detail in [6]. I recommend notating rhythm using traditional notation on separate staves below the spectrum staves so that each system of sound notation is mirrored by a layer of rhythm notation on the separate rhythm staff. Traditional notation is still the best way for communicating complex rhythmical relations in terms of note onsets. However, for the individual durations of each note I rely on the extension lines of each sound object and its components. These lines are used to indicate the duration variations treated as articulation in traditional notation, such as staccato and legato. Figure 9 shows an example of rhythm notated below the spectrum staff.

2.5 Dynamics

While the amplitude envelope can be used to define the amplitude shape of an individual sound object, we also need to address the overall dynamic development of the contents of a staff system. Depending on the notation purpose I suggest using either traditional relative dynamic notation as in the example analysis in the Addendum II of the Thoresen and Hedman paper [1], or a continuous line graph below the staff system, similar to track volume automation in Digital Audio Workstations.

2.6 Spatialisation

The representation of spatialisation is not part of SASO, and there is no common standard for notating spatialised sound [7] though there are some interesting solutions, see e.g. [8]. There is much to take into consideration when visually representing the spatial aspects of music, which I covered in greater detail in [9]. An important aspect is how a space and its characteristics can be described from different perspectives. For notating structurally significant movements and changes of position I suggest a 2D image placed above the staves displaying all numbered/labelled notated layers from a top view for horizontal movements. New images are introduced when necessary and/or whenever a layer starts or ends a trajectory, which is indicated using arrows to indicate the change to be performed until next indicator appears. As in some graphical user interfaces for surround panning, positions can either be introduced as coordinates on a cartesian coordinate system or as angles and distance related to the listener position. This 2D-notation was influenced by the work of Ellberger et al [8] [10] and Garcia et al [11].

For 3D positioning one can either introduce a front view image below the top view to account for the added dimension, or (when applicable) use a combination of colour and brightness for elevation when more exact readings of elevation are not necessary. My suggested colour scheme was inspired by the artificial horizon of airplane controls where blue and brown represent the areas above and below the horizon respectively. In my notation, symbols have brighter shades of blue as they ascend above the centre position, while they have brighter shades of brown as they descend. Figure 10 shows examples of a cartesian style representation (A), 3D positions using top and front views (B) and 3D positions using colour shades for elevation. The colour scheme itself is also represented next to the indicator (C). Figure 11 is a short example of notated movements of two numbered layers of sound notation.

Important for the notation of a musical parameter is to distinguish structural changes from elements of articulation. As with traditional musical parameters, certain spatial aspects of sound can also be considered articulations of a sound rather than positional changes, such as a sound rotating around the listener. The movement is experienced as a sound in orbit rather than a sound changing from one position to the other and should therefore be notated as a case of modulation (See Figure 12).

Sometimes sounds have their own dedicated reverb effects, functioning as resonators for those sounds rather than providing artificial room acoustics for the entire sound world. This is notated as a grey shadow behind the main extension line of the sound, reflecting the amplitude envelope of the effect. A shadow of lesser width than the spectrum category symbol (as shown in Figure 13 A) represents reverb with lower amplitude than the original sound while a reverb with the same width as the symbol (Figure 13 B) represents a reverb of equal amplitude to the original sound.

2.7 Change and transformation

What Denis Smalley introduced with spectromorphology in contrast to Schaeffer's typomorphology is a framework for describing sounds even as they change, in all their dimensions [3]. SASO accounts for this in various ways for the sake of making detailed analyses possible. Since my adaptation is supposed to work also for algorithmic composition, all individual parameters defining a sound object can be thought of as an array of one or several breakpoints. This is reflected in the continuation of spectrum category extension lines and/or dashed lines extending from the indicators subject to change. For example, in electroacous-

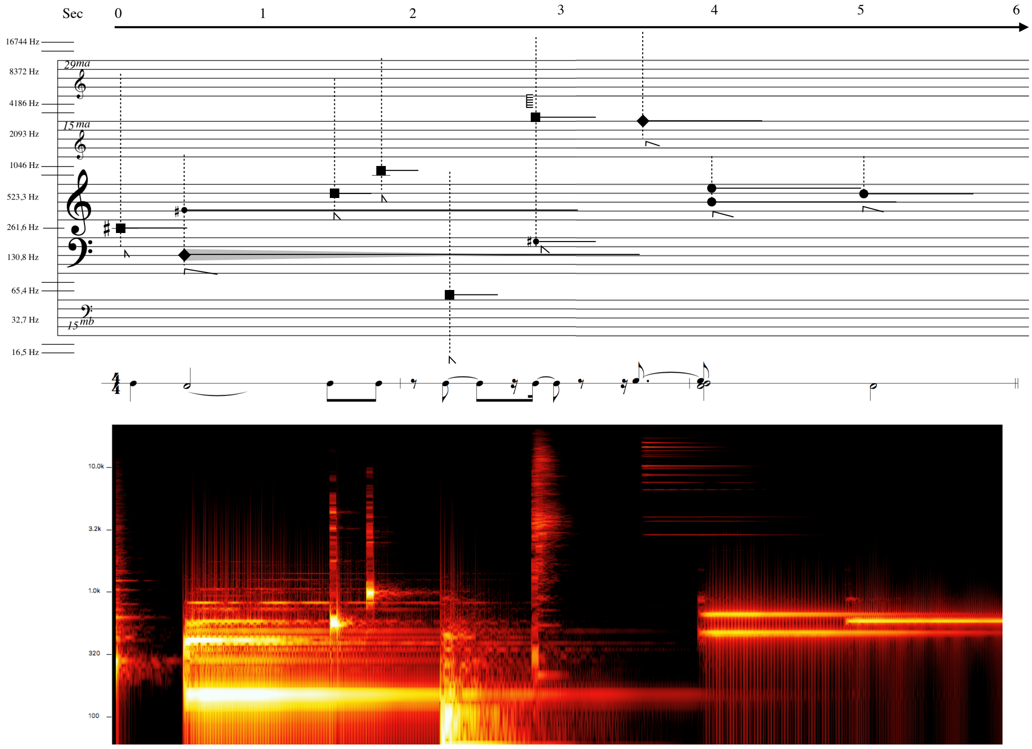

Figure 9. Short notation examples placed over a sonogram to show the correlation.

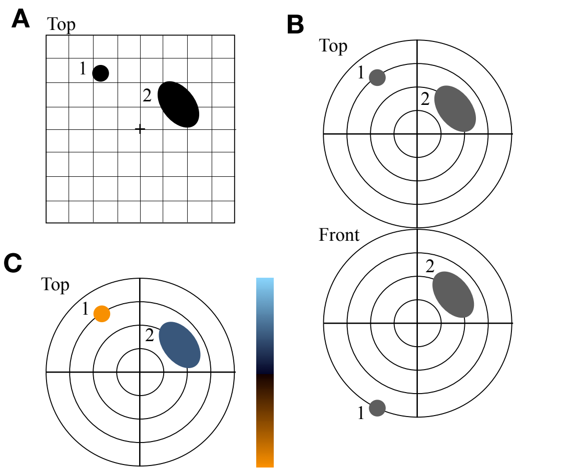

Figure 10. Examples of representation of the positions of two numbered sound layers: A) a cartesian coordinate system for horizontal placement, B) as angles relative to the centre positions in three dimensions with a top view and a front view, and C) like B but with elevation represented as colours with the colour scale shown to the right.

tic music changes to the spectral centroid and the spectral width of a sound object would be expected results of the use of automated filters which would result in dashed lines tracing the changes of width values and centroid frequency for the duration of the changes. That being said, the composer must decide what is the reasonable level of detail for notating a sound object.

2.8 Performance

For performance scores of traditional music notation, symbols are added and/or reinterpreted to accommodate the various features of each instrument. Similarly, the notation presented here can not necessarily be performed as is, but needs to be adapted and in many cases translated to action notation that makes sense for the performers and/or sound sources involved. For computer playback this means converting the notation to midi-like data to be interpreted by the computer's sound-producing software. For performances with acoustic instruments, an exploratory work may be necessary to find the actions needed to produce the sounds prescribed by the notation. Such explorations may invite the musician to take a more active part in the final design of the work. The actions can be indicated on a separate line below the sound notation in the same way that guitar tablature is often positioned below a staff of traditional notation. Transcriptions of music with unusual playing techniques as in the works of Helmut Lachenmann [12], will result in scores with major differences between the notated sounds and their indicated actions.

3. CLOSING REMARKS

What has been described here is not exhaustive but an introduction to my work with adapting Thoresen's analysis tools for composition and transcription related to the acoustic properties the system could be imagined to represent. As with any use of composition tools I expect that users of this system will make the necessary tweaks and additions to make it useful for the situation at hand. Exper-

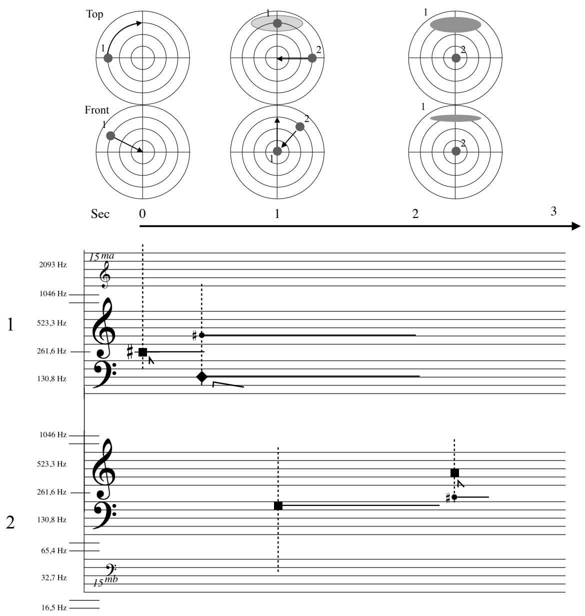

Figure 11. The notation of two numbered layers' movements

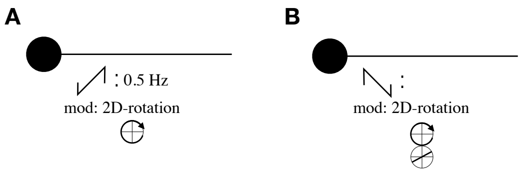

Figure 12. Spatialisation as articulation expressed as a form of modulation: A) is a clockwise steady rotation in two dimensions with one rotation every other second (0.5 Hz) while B) is a counterclockwise 2D rotation tilted in the 3D space.

imental music composers have shown to be quite inventive in terms of modifying traditional music notation in order to convey their musical ideas to musicians. Nevertheless the foundation of traditional notation remains the Western chromatic scale and their positions over a staff system constructed with diatonic scales in mind though there are plenty of suggestions for replacement systems [13].

Starting from a notation system aimed for analysis, the most important difference from notation used for composition is not the symbols themselves, but their interpretation. Neither in analysis nor composition do graphic symbols serve as a complete manifestation of the sound. But for analysis it can sometimes suffice that certain significant features of the music can be assessed, while a composed score may be the carrier of a work's identity and must contain sufficient information for its performance. A consequence of this difference in level of detail may be that the composed score needs to be broken down into different layers on several systems, while analysis scores of electroacoustic music are often displayed as one score.

There are several ways in which I will explore these ideas further, one being the development of software tools for using this notation for algorithmic composition in similar

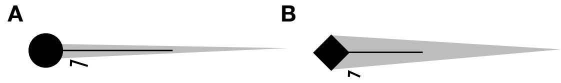

Figure 13. Two examples of reverb added to specific sounds: A) represents a reverb tail of weaker amplitude than the original pitched sound while B) represents a reverb of equal amplitude to the dystonic sound.

ways and/or in conjunction with existing solutions like the MaxScore [14] and Bach [15] libraries for Max.

I will presently compose new music with this notation system to further explore its strengths and weaknesses in relation to the music creation process together with singers, musicians and electronic devices. Also, pedagogy remains an important goal for this work. Following three years of case studies with first year bachelor students in composition at the Royal College of Music in Stockholm using this notation at different stages of its development, I have found this use of a hybrid system for pitch, inharmonic and noise components a useful and intuitive way of bringing composition students with classical and electroacoustic background together. The results of the first case study was presented at TENOR 2018 [16] and a more detailed account of the three studies will be presented in an upcoming article.

Acknowledgments

I would like to thank the anonymous reviewers for their valuable suggestions and comments. The work presented in this paper was made possible through a collaboration between The Royal Institute of Technology in Stockholm and The Royal College of Music in Stockholm.

4. REFERENCES

[1] L. Thoresen and A. Hedman, "Spectromorphological analysis of sound objects: an adaptation of pierre schaeffer's typomorphology," *Organised Sound*, vol. 12, no. 2, pp. 129–141, 2007.

[2] P. Schaeffer, *Treatise on Musical Objects: An Essay Across Disciplines*. Univ of California Press, 2017, vol. 20.

[3] D. Smalley, "Spectromorphology: Explaining Sound-Shapes," *Organised Sound*, vol. 2, no. 2, pp. 107–126, 1997.

[4] T. Grill, A. Flexer, and S. Cunningham, "Identification of perceptual qualities in textural sounds using the repertory grid method," in *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound, Coimbra, Portugal*. ACM, 2011, pp. 67–74.

[5] J. Dack and C. North, "Translating pierre schaeffer: Symbolism, literature and music," in *Proceedings of EMS06 Conference, Beijing*, 2006.

[6] M. Sköld, "Visual representation of musical rhythm in relation to music technology interfaces-an overview," in *CMMR 2019: 14th International Symposium on Computer Music Multidisciplinary Research, Oct 2019, Marseille, France*, 2019, pp. 725–736.

[7] N. Peters, G. Marentakis, and S. McAdams, "Current technologies and compositional practices for spatialization: A qualitative and quantitative analysis," *Computer Music Journal*, vol. 35, no. 1, pp. 10–27, 2011.

[8] E. Ellberger, G. T. Perez, J. Schuett, G. Zoia, and L. Cavaliero, "Spatialization symbolic music notation at icst," in *Proceedings of the joint International Computer Music Conference / Sound and Music Computing conference ICMC—SMC—2014, Athens, Greece*, 2014.

[9] M. Sköld, "The visual representation of spatialisation for composition and analysis," in *Nordic Sound and Music Computing Conference, Stockholm, Sweden*, 2019, pp. 70–77.

[10] E. B. Ellberger, G. T. Pérez, L. Cavaliero, J. Schütt, G. Zoia, and B. Zimmermann, "Taxonomy and notation of spatialization," in *Proceedings of the International Conference on Technologies for Notation and Representation, Cambridge, UK*, 2016.

[11] J. Garcia, T. Carpentier, and J. Bresson, "Interactive-compositional authoring of sound spatialization," *Journal of New Music Research*, vol. 46, no. 1, pp. 74–86, 2017.

[12] D. Alberman, "Abnormal playing techniques in the string quartets of helmut lachenmann," *Contemporary Music Review*, vol. 24, no. 1, pp. 39–51, 2005.

[13] "The music notation project," http://musicnotation.org, accessed: 2019-12-30.

[14] G. Hajdu and N. Didkovsky, "Maxscore: Current state of the art." in *International Computer Music Conference, Ljubljana, Slovenia*, 2012, pp. 156–162.

[15] A. Agostini and D. Ghisi, "Bach : an Environment for Computer-Aided Composition in Max," in *Proceedings of the International Computer Music Conference*, Ljubjana, Slovenia, 2012, pp. 373–378.

[16] M. Sköld, "Combining sound-and pitch-based notation for teaching and composition," in *TENOR'18–Fourth International Conference on Technologies for Music Notation and Representation, Montreal, Canada*, 2018, pp. 1–6.