ATOMIC NOTATION AND MELODIC SIMILARITY

Ludger Hofmann-Engl
The Link
+44 (0)20 8771 0639
ludger.hofmann-engl@virgin.net

Abstract. Musical representation has been an issue as old as music notation itself. There has been considerable development over the centuries since medieval times. Contemporary music particularly has seen many attempts to extend traditional notation. However, it appears that none of the existing systems has paid sufficient attention to the notation of durations and loudness. This article proposes a notation in form of atomic chains, where durational values are represented according to their lengths. This is, the longer a duration is, the longer will be the line representing this value. This does not allow only for the transformation of a given melody into any other melody, but provides a conceptual framework for melodic similarity.

1. INTRODUCTION

Clearly, the way we represent music will have a great impact on how music is composed, performed, analysed and music information is stored and retrieved. This is an issue which has been recognized by composers like Schoenberg (1925), theorists like Brennink (1976) and more recently in the field of music information retrieval (e.g. Selfridge-Field, 1999). We might for instance think of early music notation in form of neumes, which were sufficient for a person as an aid to remember a known melody, but not sufficient to learn an unknown melody. Over the centuries musical notation has developed to a complex system which has been undergoing rapid extensions within the composing community of the 20th century. While music representation provides marly a tool for the composer ensuring adequate performances, music theorists, psychologists and music information retrieval researchers are concerned with the representation of music ensuring its appropriate description, storage and efficient retrieval. Maybe not surprising, the pitch aspect of music representation has received much attention and so has the durational aspect to some extend, but dynamics has so far received little or no attention. It is the intention of this paper to present a novel form of music representation. However, before this system will be presented, we will briefly discuss some existing approaches to music representation.

2. REPRESENTATIONAL SYSTEMS

2.1. Conventional music notation

It is an impossible task to give an exhaustive description of our conventional music notational system, but it generally incorporates the three parameters pitch, duration and loudness. We find, if we projected this notational system into a x,y-coordinate system, that pitch coincides somewhat with the x-axis and the duration somewhat with the y-axis, while loudness is treated mainly in form of verbal description. In fact, while pitch is highly correlated to the vertical axis (except that for instance a *c* and a *c#* fetch the same vertical position although they differ in pitch), we find a lower correlation between duration and the horizontal axis. This is, the representation of durations only indicates the order of durations rather than the exact length. The exact length however, is only indicated by the shape of a tone (e.g. tones with or without stems).

2.2. The Ailler-Brennink Chromatic Notation

Brennink published his essay "Die Halbton-Schrift/Chromatic Notation" in 1976. Clearly, being disturbed by the fact that the pitch information is not perfectly correlated to the vertical axis, he proposed a notational system whereby no sharps or flats occur. He later realized that Ailler had attempted to build a similar system at the beginning of the 20th century, hence Brennink called his notation the Ailler-Brennink chromatic notation. We give an example below where a piece (Debussy's Brouillards from Préludes pour piano, No. 1) written in Ailler-Brennink notation (figure 1).

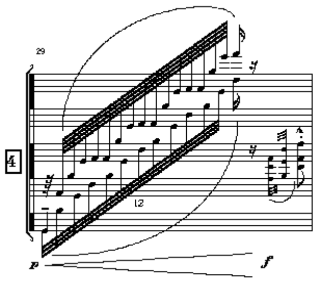


Fig. 1. Debussy's Brouillards from Préludes pour piano, No. 1 in Ailler-Brennink chromatic notation. Each system consists of four lines, where each system contains the span of an octave with 'a' being in the middle space of each system.

While pitch is now strictly correlated to the y-axis, durations are not correlated to the y-axis.

2.3. Abstract representations

There have been several attempts to produce abstract representations of music (such as MIDI or XML) driven by the need to process music within computer technology. Generally, we find that music is represented within these environments in form of sequences, such as: (pitch 1, duration 1), (pitch 2, duration 2), ..., (pitch *n*, duration *n*). Although MIDI includes loudness information, in general little attention is given to loudness. If we represented such sequences graphically in a 3-dimensional space (x-axis as the event, y-axis as the pitch and z-axis as the duration), we find an improved situation as durations are more than secondary qualities attached to a tone. However, a 3D representation is clearly not feasible in the context of musical composition , performance or music cognition.

Before we will introduce now an alternative notation, we will briefly discuss what exactly we intend to represent.

3. THE RELEVANT PARAMETERS

It seems sensible to ask before we develop a music representation what exactly we intent to represent.We have to stress that we are not interested in establishing a notational system which enhances performance practices, but a system which represents cognitive aspects in an adequate way. The terms pitch, loudness, duration and melody are highly ambiguous (what one person considers to be a melody is not regarded as a melody by another person). Thus, we will consider a new terminology (compare Hofmann-Engl, 2001, 2002, 2003).

3.1. Meloton versus pitch

According to models of virtual pitch (e.g. Terhardt, 1979) a tone does not only produce one distinct pitch, but a series of possible candidates which might serve as the pitch of the tone. Musical tones such as a piano 'a' produce often one strong candidate (e.g. 'a') and a series of very weak candidates (e.g. 'd' or 'f'), but other tones such as tones as produced by a drum instrument do not produce such a clear distinction between one strong and other weak candidates. Thus we define the term meloton as such:

Definition:

The meloton is the psychological concept whereby a listener listens to a sound directing her/his attention to the sound with the intention to decide whether the sound is high or low.

True, this does not deliver a quantity we can represent, and hence we will have to define the value of a meloton somehow without using a physicalistic concept. In this context it seems most appropriate to consider the following experimantal setting. A group of listeners is asked to tune in a (sinusoidal) comparison tone with variable frequency to match a test tone best according to each listener's individual judgment. We expect to obtain a distribution of different responses from which we then derive melotonic values (compare Hofmann-Engl, 2001).

3.2. Dynamon versus loudness

In analogy to the term meloton, we define the term dynamon:

Definition:

The dynamon is the psychological concept whereby a listener listens to a sound directing her/his attention to the sound with the intention to decide whether the sound is loud or soft.

The experimental measurement of the dynamic value follows the same idea as did the measurement of the melotonic value.

3.3. Chronoton versus duration

In order to ensure an equal treatment of the parameters we expect to be of importance in the context of melodies, we introduce the following definition:

Definition:

The chronoton is the psychological concept whereby a listener listens to a sound directing her/his attention to the sound with the intention to decide whether the sound is short or long.

The measurement of the chronotonic value will have to be conducted this time in a different way. A group of listeners will be presented with the test tone for which we intend to obtain the chronotonic value. After the test tone is heard the listener will be asked to adjust a control tone (by switching it on and off) so as to match the duration of the test tone best according to each listener's individual judgment. We expect to obtain a distribution of different responses which again will serve as the basis for the calculation of chrontonic values (compare Hofmann-Engl, 2002).

3.4. Melodic chains versus melodies

We are now in the position to define what we will call melodic chains.

Definition:

A chain is a sequence of a finite amount of tones.

True, this definition does not deviate from the more conventional concept of what we consider melodies except that the term melody is highly ambivalent and implies some form of musical judgment. We denote a chain in the form:

$$ch = [t_1, t_2, ..., t_n] \tag{1}$$

If we are interested in the melotonic contents of a chain, we write $M(ch)$, $D(ch)$ and $C(ch)$.

4. ATOMIC BEATS AND ATOMIC NOTATION

There seems to be one main deficiency with all notational systems so far, even after we introduced chains and clarified the meaning of the relevant parameters. We therefore suggest to approach this issue from a different angle. We give an example: $C(ch)$ = {quarter note, half note, quarter note, eights note, eights note). Now we find that the smallest chronoton is an eighth. In fact all other chronota in our example are multiples of an eighth. The 1/4 is twice as long as the 1/8 and the ½ is four times as long as the 1/8. Hence, we call the 1/8 the atomic beat of this chain. Without going into any detail, atomic beats can be generally found, even if c-chains contain triplets or more complex time ratios. Even if time ratios are irrational, we will find atomic beats, as irrational time ratios can be approximated by rational time ratios. We now return to our example: Quantesizing time into 1/8 beats, we obtain the length for $|C(ch)| = [2 + 4 + 2 + 1 + 1]\,(1/8) = 10\,(1/8)$. Thus, the c-chain consists of 10 atomic beats where 1 beat is given by 1/8. On the first atomic beat a 1/4 commences, on the second atomic beat the 1/4 comes to its end, on the third atomic beat the ½ commences, it continues through the forth and fifth atomic beat and finds its completion on the sixth atomic beat. The seventh and eighth atomic beats are occupied by a 1/4 and the last two atomic beats both are 1/8. We write: $C(ch) = [2, 2; 4, 4, 4, 4; 2, 2; 1; 1](1/8)$. This is the first two beats are 2/8 = 1/4. The 3rd to 6th beats are 4/8 = ½ etc. Generally we write:

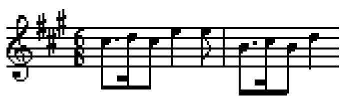
$$C(ch) = [c_{11}, c_{12}, \ldots, c_{1m_1}; c_{21}, c_{22}, \ldots, c_{2m_2}; \ldots; c_{n1,} c_{n2}, \ldots, c_{nm_n}](1/a) \qquad (2)$$

where $C(ch)$ is a chronotonic chain in atomic notation, $c_{i1}, c_{i2}, \ldots c_{im_i}$ the *ith* chronoton in atomic notation, m_i the number of atomic beats covered by the ith chronoton, n the length of the the c-chain and $1/a$ the atomic beat.

Dealing with melodic chains in general, we obtain:

$$ch = [t_{11}, t_{12}, \ldots, t_{1m_1}; t_{21}, t_{22}, \ldots, t_{2m_2}; \ldots; t_{n1}, t_{n2}, \ldots t_{nm_n}](1/a) \qquad (3)$$

We will give an example. The opening of the theme from Mozart's a -major sonata (Köchel Nr. 331) is:



The chain*ch*is given as:

ch = [(c#, 3/16, 4), (d, 1/16, 1), (c#, 1/8, 2), (e, 1/4, 3), (e, 1/8, 2), (b, 3/16, 4), (c#, 1/16, 1), (b, 1/8, 2), (d, 1/4, 3)]

in standard notation and in atomic notation:

ch = [(c#,3,4), (c#,3,4), (c#,3,4); (d,1,1); (c#,2,2), (c#,2,2);(e,4,3), (e,4,3), (e,4,3,), (e,4,3,); (e,2,2), (e,2,2); (b,3,4), (b,3,4), (b,3,4); (c#,1,1); (b,2,2), (b,2,2); (d,4,3), (d,4,3), (d,4,3), (d,4,3)](1/16)

Note, we are dealing here with a score rather than a transcript of a sound source. Hence, we are dealing with predicted values rather than measured values. Still, for the purpose of illustration, we will assume that the score does represent approximated measured values. However, this is not true for the dynama. Clearly, Mozart did not intend the piece to be performed without any dynamic variation, but following 18th century notational practice, Mozart did not feel the need to be more specific about the dynamics of the piece. Rating dynama on a scale from 1 to 9 with 1 = soft, 9 = loud), we might assume that the above stated interpretation is suitable.

This form of atomic notation preserves the information of all three parameters equally well relating them to quantasized time events. The main advantage of this method might only become fully apparent in the context of melodic similarity, however the transformation theory as presented next will enable us to see some useful aspects of this notation.

5. MELODIC TRANSFORMATIONS

There are several ways of introducing melodic transformations. However, the author decided it would be most appropriate to refer to an example, and then to explain the underlying concept in more detail.

Let us consider the beginning of the 1st variation of the Mozart piece we were talking about. This variation opens as:

As before, we have no sufficient dynamic information. We assume that the following d-chain represents an acceptable interpretation of the piece: $D(ch_v)$ = [4; 3; 0; 2; 4; 2; 3; 2; 0; 2; 3; 1; 4; 3; 0; 2; 4; 2; 3; 2; 0; 2]**.** It is interesting to note, that in this case standard and atomic notation produce the same chains. This is, because now all chronota are 1/16 notes. We further obtain the m-chain $M(ch_v)$ = [b#; c#; -; c#; b#; c#; d#; e; -; e; f#; e; e; b; -; b; a#; b; c#; d; -; d] and the c-chain $C(ch_v)$ = [1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1].

Now transforming ch into ch_v, will require some form of transformation process. We follow the mechanism as proposed by Hofmann-Engl (2001, 2002) whereby ch will be mapped onto ch_v via a chain of reflection points. In order to illustrate this, we will consider the m-chain of the theme compared to the m-chain of the first variation. We find:

$M(ch)$ = [c#, c#, c#; d; c#, c#; e, e, e, e; e, e; b, b, b; c#; b, b; d, d, d, d](1/16)
and
$M(ch_v)$ = [b#; c#; -; c#; b#; c#; d#; e; -; e; f#; e; e; b; -; b; a#; b; c#; d; -; d](1/16)

We will depict the chains in log frequencies (where a rest will be mapped onto a 0 value) Mapping the values of each atomic beat of $M(ch)$ onto $M(ch_v)$, we obtain the reflection line (middpoints such as c# = 2.75 and b# = 2.73, the middpoint is 2.73):

$R_m(ch)$ = [2.74, 2.75, 1.38, 2.76, 2.74, 2.75, 2.80, 2.81, 1.41, 2.81, 2.83, 2.81, 2.76, 2.71, 1.34, 2.73, 2.70, 2.71, 2.76, 2.77, 1.39, 2.77]

This reflection chain makes it possible to map two m-chains onto each other. Moreover it delivers some information on how closely the two m-chains in question are inversions to each other (the more straight the reflection line the closer the transformation to an inversion). Clearly, in our case we can see that Mozart did not have the concept of an inversion in mind when he devised the first variation. In fact we will later see, that his intention must have been very different.

We now consider the c-chains. Rewriting them in $\log_2$ notation, we map both c-chains onto each other via the reflection chain (quarter note = 1.6 and an eigths note = 0. The middpoint is 0.8):

$R_c(ch)$ = [0.8, 0.8. 0.8, 0, 0.5, 0.5, 1,1 ,1, 1, 0.5, 0.5, 0.8, 0.8, 0.8, 0, 0.5, 0.5, 1, 1, 1, 1]

As we can observe, Mozart's variation and theme are not close chronotonic inversions (as it does not produce a straight line).

We finally map the dynamic chains onto each other via the reflection chain:

$R_d(ch)$ = [4, 3.5, 2, 1.5, 3, 2, 3, 2.5, 1.5, 2.5, 2.5, 1.5, 4, 3.5, 2, 1.5, 3, 2, 3, 2.5, 1.5, 2.5]

As we can see, the d-chains are also not close inversions to each other either (as it does not produce a straight line). We will refrain from defining the exact mechanisms underlying these reflections, as this is very tedious. However, generally we calculate the reflection points as:

$$p_{ri} = \frac{p_{1i} + p_{2i}}{2} \tag{4}$$

where p_{ri} is the reflection point at the place *i*, p_{1i} the value of the first chain at the place *i* and p_{2i} the value of the second chain at the place *i*.

We will now consider a mechanism, which will enable us to estimate the similarity relation between the theme and the variation.

6. A CONCEPT OF MELODIC SIMILARITY

Just as we introduced melodic reflections by referring to the Mozart example, so will we now introduce a concept of melodic similarity by referring to the Mozart example.

Reflecting (=inverting) the theme along the x-axis, we obtain:

$-M(ch)$ = [-2.75, -2.75,- 2.75;- 2.77;- 2.75;- 2.75;- 2.81,- 2.81,- 2.81,- 2.81;- 2.81,- 2.81;- 2.71,- 2.71,- 2.71; -2.75;- 2.71,- 2.71;- 2.77,- 2.77,- 2.77,- 2.77]

Reflecting $-M(ch)$ onto $M(ch_v)$, we obtain the similarity chain $S_m(ch, ch_v)$:

$S_m(ch, ch_v)$ = [0.02, 0, 2.75, 0.02, 0.02, 0, 0.02, 0, 2.81, 0, 0.04, 0, 0.1, 0, 2.71, 0.04, 0.02, 0, 0.02, 0, 2.77, 0]

We illustrate this in figure 2.

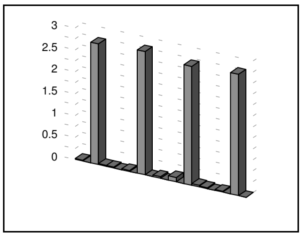


Fig. 2. The x-axis represent the atomic beat and the y-axis the corresponding value of the melotonic similarity chain, where Mozart's theme and variation are compared

As we can see, the lines reach a peak each time the first variation of the Mozart theme is on a rest. All other times, we find small deviations between the theme and the variation. Variation and theme even coincide in 9 atomic beats. Clearly, theme and variation fetch high melotonic similarity.

If we form the similarity chain comparing the c-chains, we obtain:

$S_c(ch)$ = [1.6, 1.6, 1.6; 0; 1, 1; 2, 2, 2, 2; 1, 1; 1.6, 1.6, 1.6; 0; 1, 1; 2, 2, 2, 2]

As we can see $S_c(ch, ch_v)$ and $C(ch_v)$ are identical. This is, because all chronotonic values of the variation fetch the value 0.

As we can see in figure 3, we obtain a pattern, whereby the chronotonic similarity is highest on the 4th atomic beat and lowest on the 7th to 10th atomic beat. Clearly, the theme and the variation seem to be little similar in terms of chronotonic similarity.

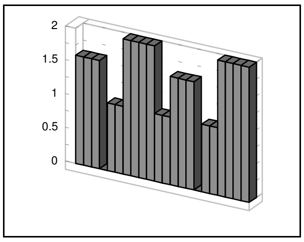


Fig. 3. Here, the x-axis represents the atomic beats and the y-axis the corresponding values of the chronotonic similarity chain, where Mozart's theme and variation are compared.

Before we will discuss more general features of the similarity chains, we will also produce the dynamic similarity chain. However, it must be stressed that this is somewhat problematic as the dynamic chains are the result of interpretation by the author rather than based on Mozart's score. Even more, going by Mozart's instructions, both the theme and the variation are to be performed piano. Accordingly, all values of the similarity chain would fetch the value 0. However, if we considered our earlier interpretation as acceptable, we obtain the following similarity chain:

$S_d(ch)$ = [0, 1, 4, 1, 2, 0, 0, 1, 3, 1, 1, 1, 0, 1, 4, 1, 2, 0, 0, 1, 3, 1]

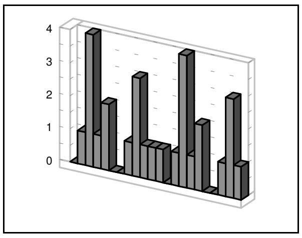


Fig. 4. Again, the x-axis represents the atomic beats while the y-axis represents the corresponding values of the dynamic similarity chain, where Mozart's theme and variation are compared

As shown in figure 4, the similarity between theme and variation is smallest on the atomic beats 3 and 15. This is a rest in the variation (dynamon value 0) is compared to a tone on an down beat. However, we find, unlike in the context of chronotonic similarity that the dynamic similarity chain still fetches a 0 value on 6 atomic beats. Hence, we assume that the dynamic similarity is higher than the chronotonic similarity. This does not surprise: Mozart clearly overrides the original rhythm in his first variation by implementing a monotonous c-chain, while the dynamics would remain somewhere similar, because the m-chains are similar and because the accents will be somewhat the same as both pieces are written in a 6/8 time.

We finally point out two specific features of similarity chains: (A) A straight line such as *S*(*ch*) = [1, 1, ..., 1] indicates in the context of melotonic similarity a transposition, in the context of dynamic similarity a volume change and in the context of chronotonic similarity a split or fusion of chronota (compare Hofmann-Engl, 2002). (B) The curvier the line (e.g. *S*(*ch*) = [1, 4, 2, 9 ....], the more the two compared chains differ in shape. As shown by Hofmann-Engl & Parncutt (1998), melotonic similarity can be predicated by referring to the transposition interval and the interval difference (shape). Further (Hofmann-Engl 2002) produced data which indicate that the distance of the chrontonic similarity chain is the sole predictor for chronotonic similarity.

The points of a similarity chain are given as:

$$p_{s_i} = \left| p_{1_i} - p_{2_i} \right| \quad (5)$$

where p_{si} is the ith point of the similarity chain, p_{1i} is the ith point of the first chain and p_{2i} is the ith point of the second chain

As shown by Hofmann-Engl (2003), the points of the similarity chain can serve as the basis for a similarity predictor based upon an algebra of similarity and interval vectors.

7. CONCLUSION

We briefly demonstrated some shortcomings of conventional notational systems and proposed the usage of a novel notation for the purpose of cognitive analysis and composition of music. We found that the representation of music in form of atomic notation has many advantages including the treatment of all relevant parameters in an equal fashion. We also were able to deduct from our notation a conceptual framework for estimating melodic similarity ratings.

REFERENCES

Brennink, A. (1976). Die Halbton-Schrift/The Chromatic Notation. Edition Chroma, Montreux

Hofmann-Engl, L. & Parncutt, R. (1998). Computational modelling of melodic similarity judgments - two experiments on isochronous melodic fragments. Online: http://www.chameleongroup.org.uk/research/sim.html

Hofmann-Engl, L. (2001). Towards a cognitive model of melodic similarity. In: Proceedings of the 2nd annual ISMIR, Bloomington

Hofmann-Engl, L. (2002). Rhythmical similarity: A theoretical and empirical approach. In

Proceedings of ICMPC 7, Sydney

Hofmann-Engl, L. (2003). Melodic transformations and similarity: a theoretical and empirical approach. PhD thesis, Keele Univerisity

Schoenberg, A. (1925). Eine Neue Zwölftonschrift. Manuscript.
available on http://www.schoenberg.at/6_archiv/texts/T_20_38/T_20_38.htm#

Selfridge-Field, E. (1999). Conceptual and Representational Issues in Melodic Comparison. In: Melodic Similarity - Concepts, Procedures, and Applications (ed. Hewlett & Selfridge-Field). Computing in Musicology, vol 11.

Terhardt, E. (1979). Die psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmische Bestimmung. In Tiefenstruktur der Musik, TU Berlin, 1979