Journal of Mathematics and Music

Mathematical and Computational Approaches to Music Theory, Analysis, Composition and Performance

ISSN: 1745-9737 (Print) 1745-9745 (Online) Journal homepage: https://www.tandfonline.com/loi/tmam20

Combinatorics of words and morphisms in some pieces of Tom Johnson

Jean-Paul Allouche & Tom Johnson

To cite this article: Jean-Paul Allouche & Tom Johnson (2018) Combinatorics of words and morphisms in some pieces of Tom Johnson, Journal of Mathematics and Music, 12:3, 248-257, DOI: 10.1080/17459737.2018.1524028

To link to this article: https://doi.org/10.1080/17459737.2018.1524028

Published online: 03 May 2019.

Submit your article to this journal

Article views: 20

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tmam20

Journal of Mathematics and Music, 2018
Vol. 12, No. 3, 248–257, https://doi.org/10.1080/17459737.2018.1524028

Combinatorics of words and morphisms in some pieces of Tom Johnson

Jean-Paul Allouche[a*] and Tom Johnson[b]

[a]*CNRS, IMJ-PRG, Sorbonne Université, Paris, France;*
[b]*Éditions 75, Paris, France*

(*Received 23 April 2018; accepted 11 September 2018*)

We survey several occurrences of combinatorics of words and morphisms in the pieces of Tom Johnson, showing in particular that some of the sequences of notes that he used intuitively can be interpreted or constructed through combinatorial objects such as morphisms. Furthermore some of these sequences have an independent interest in number theory and theoretical computer science.

Keywords: combinatorics on words; morphisms; creative process; composition; sequence (structural); form; rhythm and meter

2010 Mathematics Subject Classification: 11B85; 68R1

1. Introduction

Tom Johnson is a minimalist composer who explored morphic sequences intuitively in his musical compositions beginning in the late 1970s. After meeting Michel Waldschmidt and Jean-Paul Allouche in Paris around 1987, he became more conscious of the mathematics of what he was doing, as one can see in *Formulas for String Quartet*, *Narayana's Cows*, *Pascal's Triangle Modulo Seven*, *Automatic Music*, and many other works. Without necessarily being exhaustive we will briefly present some of his pieces with a quick description of the underlying mathematical objects related to Combinatorics on words.

2. Two 'morphic' pieces: *Formulas for String Quartet* and *Narayana's Cows*

As described in Allouche and Johnson (1995a), when Tom Johnson was composing his *Formulas for String Quartet*, he was using heuristic algorithmic rules. Wanting to make these rules clearer and more formal he questioned several mathematicians. Hints were given by D. Feldmann at the University of New Hampshire and M. Waldschmidt at Université Paris 6: in particular Waldschmidt suggested that there could be 'automatic sequences' hidden behind these rules and that it could be interesting to ask Allouche about this point. Indeed automatic and/or morphic sequences

*Corresponding author. Email: jean-paul.allouche@imj-prg.fr

© 2019 Informa UK Limited, trading as Taylor & Francis Group

proved useful: they provide a unified description of these rules; they also helped the composer to finish some of the *Formulas*.

A basic example of sequence 'generated by a morphism' is the Thue–Morse sequence. First define on the set of *words* (i.e. finite sequences) and infinite sequences the *morphism* (i.e. the rewriting rule) defined by $0 \to 01$, $1 \to 10$. Starting with 0 and iterating this morphism gives successively

$$\begin{array}{l} 0 \\ 0\,1 \\ 0\,1\,1\,0 \\ 0\,1\,1\,0\,1\,0\,0\,1 \\ \dots \end{array}$$

Continuing to iterate provides an infinite sequence that is, by construction, invariant under the morphism:

$$\underbrace{0\,1}_{0}\quad\underbrace{1\,0}_{1}\quad\underbrace{1\,0}_{1}\quad\underbrace{0\,1}_{0}\quad\underbrace{1\,0}_{1}\quad\dots$$

The two sequences above are the same infinite sequence. Grouping the terms pairwise in the first one recalls where they come from when the morphism was applied: in other words the upper sequence appears as the image of the lower one by the morphism, or equivalently, the lower sequence is obtained by 'decoding' the upper one. More properties of morphisms, morphic sequences, and *automatic sequences* – which are a particular case of morphic sequences – can be found, e.g. in Allouche and Shallit (2003).

The details of the (morphic) constructions generating the *Formulas for String Quartet* can be seen in the score (Johnson 1994), also see Allouche and Johnson (1995a). We only give here an explanation for Movement 1 of the *Formulas*, whose first lines are given in Figure 1. This movement uses the morphism $+ \to + - +$, $- \to - - +$ so that the corresponding infinite sequence begins

$$+ - + - - + + - + - - + - - + \dots$$

The formula is followed simultaneously by the four instruments with tempo proportions 1:3:9:27 and interval proportions of 1:2:3:4. The symbol $+$ (or $+1$) indicates a melodic ascent, while the

Figure 1. The first lines of Movement 1 of the *Formulas for String Quartet*.

symbol $-$ (or -1) indicates a melodic descent.

VN1	+	−	+	−	−	+	+	−	+	−	−	+	−	−	+	
	+	−	+	+	−	+	−	−	+	+	−	+				
VN2	+2			−2			+2			−2			−2			+2
			+2			−2			+2							
VLA	+3									−3						
									+3							
VCL	+4															

Narayana's Cows is a piece of Tom Johnson, dated 1989. Narayana was an Indian mathematician in the fourteenth century, who proposed a variant of the Fibonacci rabbit, where the cows cannot produce a calf before their fourth year. It is possible to define a notion of *delayed morphism* to describe the successive generations of cows (see Allouche and Johnson 1995b).

Remark 2.1 Renormalizing the successive steps of the iterative construction of a sequence generated by a morphism of constant length (like the morphism generating the Thue–Morse sequence where all single letters give words of length 2) produces a one-dimensional self-similar set, which is actually a fractal object. Tom Johnson also used this more general concept of self-similarity, see, e.g. Johnson (1996) (also see Feldman (1998) and Johnson (2006)).

3. Generalized locally catenative formulas

A classical way of defining a sequence of words that can converge to a structured infinite word is the use of *locally catenative formulas*. For a precise definition, the reader can consult Rozenberg and Lindenmayer (1973) whose abstract gives a simplified definition: *A locally catenative sequence of strings of letters is such that each string in the sequence, after an initial stretch, is formed by concatenating strings which occurred at some specified distances previously in the sequence*. This means that the nth string (i.e. the nth word) is obtained by concatenating the $(n-i_1)$st, the $(n-i_2)$nd, $\ldots$, and the $(n-i_k)$th strings,where $i_1, i_2, \ldots, i_k$ is a finite increasing sequence of fixed integers. A simple example is as follows: fix $x_1 = 1$, $x_2 = 0$, and define, for $n \geq 3$, $x_n = x_{n-1}x_{n-2}$. We thus have $x_3 = x_2x_1 = 01$, $x_4 = x_3x_2 = 010$, $x_5 = x_4x_3 = 01001 \cdots$ It is easily checked that the sequence of strings $(x_n)_{n\geq 1}$ tends to an infinite binary sequence, namely 0 1 0 0 1 0 1 0 … which is known as the binary Fibonacci sequence (because of the lengths of the consecutive x_n's). A link between these sequences and morphic sequences is given in Shallit (1988), where 'generalized locally catenative formulas' are introduced: they differ from locally catenative formulas in that codings are allowed before concatenating strings.

In his *Doublings for Double Bass* (1980, and a more general version in Johnson (2018)), Tom Johnson gave six formulas.

- The first movement is given by

  1 1
  1 1 1 2
  1 1 1 2 1 1 1 3
  1 1 1 2 1 1 1 3 1 1 1 2 1 1 1 4
  …

  where each number represents another note on a specially defined scale.

- The second movement is given by

1 1
1 2 1 1
1 2 1 3 1 2 1 1
1 2 1 3 1 2 1 4 1 2 1 3 1 2 1 1
...

The score only gives the first four levels and defines the infinite scale. The bass player has to calculate the rest.

- The fifth formula is given by

1 1
1 2 1 2
1 3 2 3 1 3 2 3
1 4 3 4 2 4 3 4 1 4 3 4 2 4 3 4
...

It is worth mentioning that there are six different formulas, but only one of them was well played, and that was just last year, 37 years after the piece was composed! Tom Peters succeeded in going on with the fifth formula for over 20 minutes and he made a fine recording (Peters 2017).

The reader can check that the three formulas above can be obtained by iterating the following locally catenative formulas:

- *The first formula.* Take the positive integers as (infinite) alphabet. Define the coding f by $f(n) := n+1$. Define three sequences of words (u_n), (v_n), and (z_n) by $u_1 = 11$, $v_1 = 1$, $z_1 = 1$, and for $n \geq 1$, $u_{n+1} = v_n z_n v_n f(z_n)$, $v_{n+1} = v_n z_n v_n$, and $z_{n+1} = f(z_n)$. So that letting $X_n = (u_n, v_n, z_n)$, we have $X_1 = (11, 1, 1)$ and, for $n \geq 1$, $X_{n+1} = (u_{n+1}, v_{n+1}, z_{n+1}) = (v_n z_n v_n f(z_n), v_n z_n v_n, f(z_n))$. Iterating we obtain

$$
\begin{aligned}
X_1 &= (11, 1, 1)\\
X_2 &= (1112, 111, 2)\\
X_3 &= (11121113, 1112111, 3)\\
X_4 &= (1112111311121114, 111211131112111, 4)\\
&\ldots
\end{aligned}
$$

The first component u_n of X_n (taking the first component is again a coding) gives the first formula of the *Doublings for Double Bass*.

- *The second formula.* Take the positive integers as alphabet again. Define the coding f by $f(n) := n+1$. Define the triple of sequences $Y_n = (a_n, b_n, z_n)$ by $Y_1 = (11, 1, 1)$ and, for $n \geq 1$, $Y_{n+1} = (a_{n+1}, b_{n+1}, z_{n+1}) := (b_n f(z_n) b_n 1, b_n f(z_n) b_n, f(z_n))$. Iterating we obtain

$$
\begin{aligned}
Y_1 &= (11, 1, 1)\\
Y_2 &= (1211, 121, 2)\\
Y_3 &= (12131211, 1213121, 3)\\
Y_4 &= (1213121412131211, 121312141213121, 4)\\
&\ldots
\end{aligned}
$$

The first component a_n of Y_n gives the second formula of the *Doublings*.

- *The fifth formula.* This formula seems different from the other formulas in the *Doublings*. The following 'simple' construction appears: at step $n \geq 2$ interleave letter n between the letters of the previous word, at every second place.

Step 1	1 1
Step 2	1 **2** 1 **2**
Step 3	1 **3** 2 **3** 1 **3** 2 **3**
Step 4	1 **4** 3 **4** 2 **4** 3 **4** 1 **4** 3 **4** 2 **4** 3 **4**
	...

  Actually this sequence of words can also be generated by a generalized locally catenative formula. Take the positive integers as alphabet. Define f by $f(n) := n+1$. Extend f to a morphism on all words with integer letters (e.g. $f(142) = f(1)f(4)f(2) = 253$). Now define a sequence of words (u_n, v_n) by $(u_1, v_1) = (1,1)$ and, for $n \geq 1$, $(u_{n+1}, v_{n+1}) = (1f(v_n)1f(v_n), f(v_n)1f(v_n))$. It is clear that $u_n = 1v_n$ for any $n > 1$. Furthermore it is not difficult to prove by induction that the sequence (u_n) is exactly the sequence of consecutive words occurring in the fifth formula of the *Doublings*.

Remark 3.1 The famous Encyclopedia of Integer Sequences (Sloane 1964) contains the three sequences above that occur in other contexts. Namely:

- The sequence in the first formula is sequence A204988 in Sloane (1964) (the nth term of the sequence is the index $j < k$ such that n divides $2^k - 2^j$, where k is the least index for which such a j exists).
- The sequence in the second formula is sequence A001511 in Sloane (1964). It is called the *ruler sequence* or the *Gros sequence*. The nth term of the sequence is the 2-adic valuation of $2n$ (i.e. the largest integer r such that 2^r divides $2n$).
- Write the sequence in the fifth formula on a single line by gluing together the successive words. This gives: 1 1 1 2 1 2 1 3 2 3 1 3 2 3 1 4 3 4 2 4 3 4 1 4 3 4 2 4 3 4 ... This sequence –up to gluing 01 in front of it– happens to be sequence A238845 in Sloane (1964): its nth term is the length of the longest common prefix of the binary expansions of n and $n+1$ (starting at 0).

The claims in this remark can be proved by induction. The details can be found in Allouche, Dekking, and Queffélec (2018).

4. The Pascal triangle

Recall that the Pascal triangle is a table whose first row is equal to 1 0 0 0 ... and where each entry is the sum of the entry to its north west and of the entry to its north.

1							
1	1						
1	2	1					
1	3	3	1				
1	4	6	4	1			
1	5	10	10	5	1		
1	6	15	20	15	6	1	
1	7	21	35	35	21	7	1

The first eight lines of the Pascal triangle

The Pascal triangle modulo d is obtained by replacing each number in the Pascal triangle with its remainder in the division by 7. For example we show the Pascal triangle modulo 2 and the

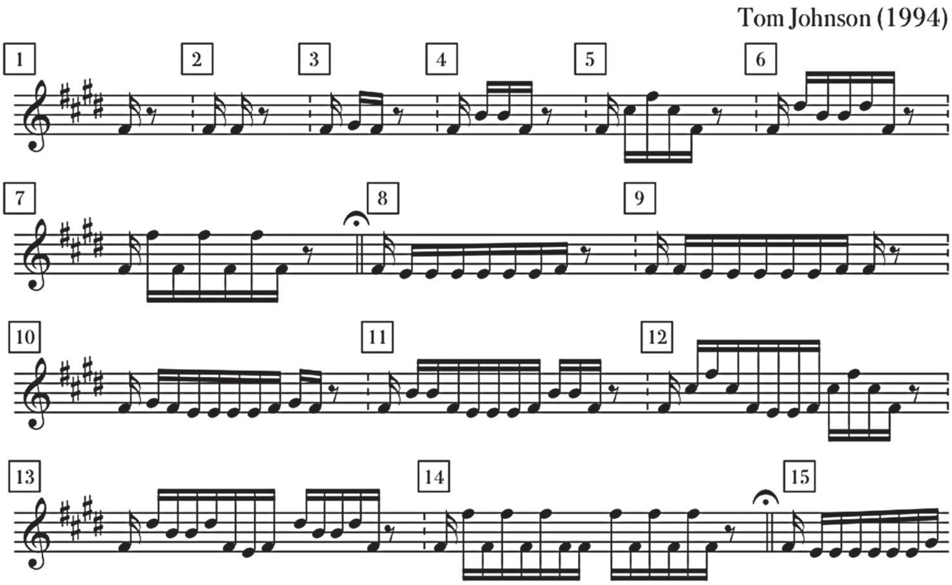

Figure 2. The first four lines of the score: Le Triangle de Pascal modulo 7.

Pascal triangle modulo 7.

1							
1	1						
1	0	1					
1	1	1	1				
1	0	0	0	1			
1	1	0	0	1	1		
1	0	1	0	1	0	1	
1	1	1	1	1	1	1	1

The first eight lines of the Pascal triangle reduced modulo 2
(to see a link with the Sierpinski triangle, draw all straight lines that contain only 1's)

1							
1	1						
1	2	1					
1	3	3	1				
1	4	6	4	1			
1	5	3	3	5	1		
1	6	1	6	1	6	1	
1	0	0	0	0	0	0	1

The first eight lines of the Pascal triangle reduced modulo 7.

In 1995 Tom Johnson used the Pascal triangle reduced modulo 7 to generate music as follows. Each of the 7 congruence classes of integers modulo 7 was associated with a note in a diatonic scale. Then each row was played with notes of equal duration, there was a brief stop after each row, and a longer pause after row 7, row 14 etc. The beginning of the score is given on Figure 2. This gave rise to a radio program, see Figure 3, where the music generated by the Pascal triangle modulo 7 was played by a computer with comments by Johnson and two other people (Johnson and Farabet 1995).

France Culture, 20 h 30 le 12 février, Atelier de Création Radiophonique

```
1
11
121
1331
14641
153351
1616161
```

LE TRIANGLE DE PASCAL

```
10000001
110000011
1210000121
13310001331
146410014641
1533510153351
16161611616161
```

par Tom Johnson et René Farabet

```
100000020000001
1100000220000011
12100002420000121
133100026620001331
1464100215120014641
15335102366320153351
161616125252521616161
```

Le Triangle de Pascal modulo sept
musique électronique de Tom Johnson

```
1000000300000030000001
11000003300000330000011
121000036300003630000121
1331000322300032230001331
14641003545300354530014641
153351031221303122130153351
1616161343434334343431616161
```

Jean-Paul Allouche, mathématicien

```
10000004000000600000040000001
110000044000006600000440000011
1210000414000065600004140000121
13310004554000644600045540001331
146410042324006313600423240014641
1533510465564062442604655640153351
16161614343434616161643434341616161
```

Bernadette Le Saché, voix

```
100000050000003000000300000050000001
1100000550000033000003300000550000011
12100005350000363000036300005350000121
133100051150003223000322300051150001331
1464100562650035453003545300562650014641
15335105411450312213031221305411450153351
161616152525253434343343434352525251616161
```

De l'esprit géométrique
Blaise Pascal

```
1000000600000010000006000000100000060000001
11000006600000110000066000001100000660000011
121000065600001210000656000012100006560000121
1331000644600013310006446000133100064460001331
14641006313600146410063136001464100631360014641
153351062442601533510624426015335106244260153351
1616161616161616161616161616161616161616161616161
```

techniciens
Jean-Michel Cauquy
et Marc Didier

```
10000000000000000000000000000000000000000000000001
110000000000000000000000000000000000000000000000011
1210000000000000000000000000000000000000000000000121
13310000000000000000000000000000000000000000000001331
146410000000000000000000000000000000000000000000014641
1533510000000000000000000000000000000000000000000153351
16161610000000000000000000000000000000000000000001616161
100000010000000000000000000000000000000000000000010000001
1100000110000000000000000000000000000000000000000110000011
12100001210000000000000000000000000000000000000001210000121
133100013310000000000000000000000000000000000000013310001331
```

Figure 3. Announcement of the radio program *Le Triangle de Pascal*.

Could moduli other than 7 be tested? It is for example well known that the picture of the Pascal triangle modulo 2, once rescaled, is a famous fractal, namely the Sierpinski fractal. Since the structure of the Pascal triangle modulo d is 'simpler' when d is a prime power and more and more complicated when the number of distinct prime factors of d increases, – see, e.g. Allouche and Berthé (1994) – Allouche suggested that it could be interesting to compare the Pascal triangles modulo 6 and modulo 8. Tom Johnson tried the reduction modulo 6, which indeed gave something seemingly "chaotic." A surprise came with the reduction modulo 8. Johnson first concocted a musical scale with 8 degrees. Then he let his computer play the triangle modulo 8; as a sort of 'built in' rhythm seemed to emerge, he added a regular beat. The result amazingly sounded "jazzy." To conclude this section let us note that the Pascal triangle modulo 7 became quite recently a piece for piano.

5. Exploring exhaustive sortings: *The Catalogue of Chords*

Combinatorics of words also deals with *orders* on finite or infinite families of words. The *lexicographical order* (also called the *alphabetical order*) and the *genealogical order* (also called the *radix order*) are the most frequently used orders on words. Recall that the genealogical order can be defined by saying that a word is smaller than another one for this order if either its length is smaller, or if they have the same length and it is smaller according to the lexicographical order. In his *Catalogue of Chords* Johnson (1986) proposes to play all possible chords on the thirteen notes of the chromatic scale: a chord is defined as a set of at least 2 distinct notes, so that there are $\binom{13}{j}$ chords on j notes for $j \in [2, 13]$, the total number of chords thus being $\sum_{2 \le j \le 13} \binom{13}{j} = 2^{13} - \binom{13}{0} - \binom{13}{1} = 8178$. The order in which these chords are played consecutively is as follows: the lowest-pitched note that can go to the next semi-tone (without collision) rises one semi-tone and the notes below return to their starting points; every time that the highest note rises, there is a pause. Here is an extract of the order for four-note chords (read vertically as on a score)

$$\begin{array}{cccccccccc}
\ldots & 6 & 6 & 6 & 6 & 6 & & 7 & 7 & \ldots \\
\ldots & 5 & 5 & 5 & 5 & 5 & & 3 & 4 & \ldots \\
\ldots & 3 & 3 & 4 & 4 & 4 & \text{pause} & 2 & 2 & \ldots \\
\ldots & 1 & 2 & 1 & 2 & 3 & & 1 & 1 & \ldots
\end{array}$$

Playing this piece on a piano is a real *performance*. For a long time, only Tom Johnson and Samuel Vriezen played it successfully (Johnson 2013). Since then Samuel Boré and Christopher Adler learned the piece, and probably others; and many people have programmed it for computer, player piano, and other mechanical instruments.

Remark 5.1 Other pieces of Johnson use ordered exhaustive lists, for example *Six-note Melody* (1987), *360 Chords* and *844 Chords* (both dated 2005).

6. Conclusion

Along with this brief description of some of the pieces of Tom Johson related to combinatorics of words, we gave some examples of what he calls "Found mathematical objects." A common feature of these pieces is that they can be played "automatically," or even that they could be, in some cases, potentially infinite: in this vein one could have cited, e.g. the pieces: *Infinite Melodies*, *Rational Melodies* (see Figure 4), *Counting Keys*, *Music for 88* with Mersenne numbers, *Tick*

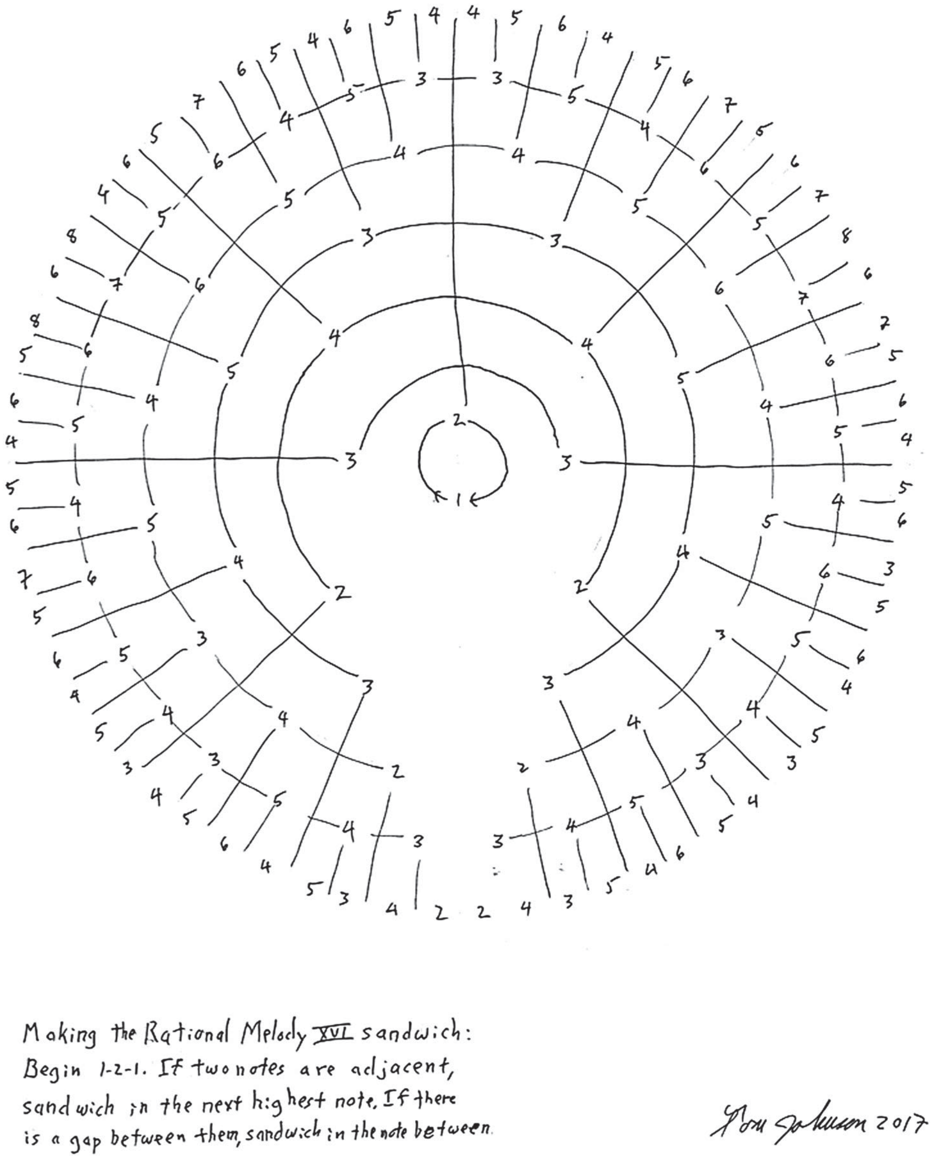

Figure 4. Rational Melodies, ♯16.

Tock Rhythms, *Counting Duets* (the reader is invited to try to describe, e.g. the second movement with one of the mathematical tools above), *Counting languages*, *Eggs and Baskets*··· and many of the pieces that can be found at http://brahms.ircam.fr/tom-johnson#works_by_date (a more complete catalogue is available at http://www.editions75.com). We chose not to be exhaustive: for example we did not mention pieces related to rhythmic (miniscule) canons, self-replicating structures, one-dimensional tilings, block designs, homometric pairs, logical series of harmonies unlike the lexicographical sequence of the Chord Catalogue, nor 'twisted' morphic sequences as in *Automatic Music for Six Percussionists*. What should be perhaps underlined as a *caveat* is the difference that Johnson sees between a mathematical object (as general and universal as possible) and a musical object (often simply a singular phenomenon particularly adapted to musical transformation).

Acknowledgments

We would like to thank the referees for their suggestions which helped to improve this article. The first named author thanks É. Allouche for her remarks.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Allouche, J.-P., and V. Berthé 1994. "Triangle de Pascal, Complexité et Automates." *Bulletin of the Belgian Mathematical Society – Simon Stevin* 4: 1–23.

Allouche, J.-P., F. M. Dekking, and M. Queffélec 2018. "On Some Automatic and Morphic Sequences in the OEIS", in preparation.

Allouche, J.-P., and T. Johnson 1995a. "Finite Automata and Morphisms in Assisted Musical Composition." *Journal of New Music Research* 24: 97–108.

Allouche, J.-P., and T. Johnson 1995b. "Narayana's Cows and Delayed Morphisms." *3èmes Journées d'Informatique Musicale (JIM '96), Ile de Tatihou* 2–7.

Allouche, J.-P., and J. Shallit 2003. *Automatic Sequences. Theory, Applications, Generalizations*. Cambridge: Cambridge University Press.

Feldman, D 1998. "Review of Self-Similar Melodies by Tom Johnson." *Leonardo Music Journal, Ghosts and Monsters: Technology and Personality in Contemporary Music* 8: 80–84.

Johnson, T. 1986. *Le Catalogue des Accords*. Score available at http://www.editions75.com/FreeScores/TheChordCatalogue.pdf.

Johnson, T. 1994. *Formulas for String Quartet*. Score published by Éditions 75, 75 rue de la Roquette, 75011 Paris, and Two-Eighteen Press, POB 218, 5004 Applewood Circle, Carmel NY 10512–2640.

Johnson, T. 1996. *Self-similar Melodies*. Éditions 75, 75 rue de la Roquette, 75011 Paris, and Two-Eighteen Press, POB 218, 5004 Applewood Circle, Carmel NY 10512–2640.

Johnson, T. 2006. "Self-Similar Structures in My Music: An Inventory", talk given a IRCAM. http://repmus.ircam.fr/_media/mamux/saisons/saison06-2006-2007/johnson-2006-10-14.pdf.

Johnson, T. 2018. *Doublings*. (New more general version of Doublings for Double Bass) Éditions 75, 75 rue de la Roquette, 75011 Paris, and Two-Eighteen Press, POB 218, 5004 Applewood Circle, Carmel NY 10512–2640.

Johnson, T., and R. Farabet 1995. "France Culture, Atelier de Création Radiophonique," *Le triangle de Pascal*, edited by Tom Johnson and René Farabet, with Jean-Paul Allouche and Bernadette Le Saché, 12 February 1995.

Johnson, T., and S. Vriezen 2013. *The Chord Catalogue Within Fourths Within Fifths*. CD, Edition Wandelweiser, Ref: EWR 1304.

Peters, T. 2017. "Failing." In *Plucking*, edited by T. Johnson. Microfest Records.

Rozenberg, G., and A. Lindenmayer 1973. "Developmental Systems with Locally Catenative Formulas." *Acta Informatica* 2: 214–248.

Shallit, J 1988. "A Generalization of Automatic Sequences." *Theoretical Computer Science* 61: 1–16.

Sloane, N. J. A. (founded by). 1964. *On-Line Encyclopedia of Integer Sequences*. Electronically available at http://oeis.org.