Surveying some notions of complexity for finite and infinite sequences

J.-P. Allouche
CNRS, Institut de Mathématiques
Équipe Combinatoire et Optimisation
Université P. et M. Curie
F-75252 Paris Cedex 05
France
allouche@math.jussieu.fr

1 Introduction

Given a finite or infinite sequence, how to recognize that it is "complicated" (or "complex"), "random", "chaotic"? How to compare the "complexity", the "randomness", the "chaos" of two sequences? Is it reasonable to consider that the sequence 0000... is "simpler" than the sequence 010101..., and that this last sequence is simpler than a binary eventually periodic sequence like 000001001001? Also should a "random" sequence be "complicated"? Among the many studies devoted to these... complicated questions, we will survey quickly some notions of complexity for sequences taking only finitely many values; then we will end with a notion related to an old theorem of Cauchy and Crofton and recently adressed in a paper of L. Maillard-Teyssier and the author: the *inconstancy* of a sequence.

2 Algorithmic complexity

One way of measuring the complexity of a sequence is to evaluate how difficult it is to generate that sequence. The difficulty is measured, e.g., by the quantity of instructions needed to generate the sequence (typically the length of a computer program), or by the time needed to compute the nth term or the first n terms of the sequence (typically the time for running a computer program). Of course the interesting measure is the *smallest* quantity of instructions, or the *smallest* computing time. In particular it is usually the case that only upper bounds can be given for such a measure of complexity.

The *Kolmogorov-Solomonoff-Chaitin complexity* of a sequence is the size of the shortest program (on a Turing machine) generating that sequence. In particular this notion is related to the *compressibility* of the sequence. To give a simple example, the shortest way of describing the "complicated" finite sequence 0 0 1 0 1 0 1 1 0 0 0 1 0 might well be just to pronounce it, while the finite sequence 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 can be quickly described as "repeat the pattern 01 ten times". A standard reference for the *Kolmogorov-Solomonoff-Chaitin complexity* is the book [45].

In a similar vein, the *algorithmic complexity* or the *computational complexity* determine the time and space resources needed for a computation using a given algorithm and given data sizes. Two references are [25] on the analysis of algorithms, and [8] more focused on the complexity classes. In connection with analysis of algorithms, it is impossible not to cite the fascinating book on analytic combinatorics by Flajolet and Sedgewick [32]

Remark 1 We give two more references: [36] compares several notions related to the computational complexity or to the Kolmogorov-Solomonoff-Chaitin complexity (also see the references given there), while [27] shows how to approximate the Kolmogorov-Solomonoff-Chaitin complexity of words of small length. Finally we cite the theory of *simplicity* (J.-L. Dessalles) for which we refer to `http://www.simplicitytheory.org/` and the references therein.

3 Combinatorial complexities

Another way of looking at whether a sequence is complicated is to look for its combinatorial properties (combinatorial in the sense of *combinatorics on words*), e.g., counting subwords, palindromes, squares, etc., occurring in the sequence. The first example of such a complexity is the block complexity (the sequences are supposed to take finitely many values).

- The *block – or factor – complexity* (sometimes just called *complexity*) counts how many distinct blocks (also called factors/subwords) of each length occur in the sequence. In particular, the complexity of any ultimately periodic sequence is ultimately constant, while "random" sequences (i.e., almost all sequences for the Lebesgue measure) have the maximal number (i.e., all) blocks of each length. Three surveys on the block complexity are [3, 30, 31].

Variations on the block complexity are the following.

- The *repetition complexity* counts the amount of repetitions in a sequence, see, e.g., [44]. (Many other papers deal with related questions, see, e.g., [54, 34, 35]...)

- The *palindrome complexity* counts the number of palindromes of each length, see, e.g., [5].

- The *arithmetical complexity* counts the number of words of each length occurring in the arithmetical subsequences of the sequence, see, e.g., [9].

- The *pattern complexity* of a sequence for a given pattern S (i.e., a finite subset of the integers) is the number of distinct restrictions of the sequence to $S + n$ (with $n = 0, 1, 2, ...$), while the *maximal pattern complexity* is the function of k defined as the maximal value of the pattern complexity for all patterns of length k, see, e.g., [41, 42]; also see the papers [43, 40].

- The *window complexity* of a sequence counts the number of *contiguous* factors of each length, see [38, 20].

Let us also mention properties of sequences that are related to some sort of complexity: quasi-periodicity, recurrence and uniform recurrence (also called repetetivity and uniform repetitivity), some definitions of pseudo-randomness, e.g., [49], the study of certain subsequences of the given sequence – in particular the *measure of automaticity* introduced by Shallit et al., see, e.g., Chapter 15 of [7].

Remark 2

- Among the works comparing these combinatorial complexities, we would like to cite the paper [39] that looks at block complexity, maximal pattern complexity and minimal pattern complexity.

- One can define the *VC-dimension* (i.e., the *Vapnik-Chervonenkis dimension*) of a sequence; this dimension is related to the above complexities, see in particular [52].

- The definition of block-complexity for a sequence inspired the notion of *permutation complexity*, see in particular [33] and [48]; also see [61] for the case of the Thue-Morse word.

- Another interesting notion of complexity for sequences was defined in relation with symbolic dynamics, namely the *heaviness*, see [12, 55, 56].

- Finally we would like to mention the *visibility graph* of a sequence studied in particular in [46, 47].

4 Number-theoretical complexity

With a sequence of numbers (say integers) taking finitely many values one can associate a real number whose digits in some integer base are given by the sequence, a formal power series whose coefficients are given by the sequence or by its reduction modulo some prime number, or a real number whose continued fraction expansion is given by the sequence (or the sequence translated by some integer to have only positive integers).

It is tempting to relate complexity properties of the sequence to algebraic properties (e.g., being rational or algebraic) of these associated real numbers, formal power series, or continued fractions. But what should be expected? Algebraic irrational real numbers are conjectured to have very complicated expansions in integer bases: these expansions are suspected to be *normal* in any integer base b, i.e., to be such that any block of length n occurs with the frequency $1/b^n$; see in particular the paper [11] of Borel about the digits of $\sqrt{2}$ in base 10. In other words the expansion of an irrational algebraic number is expected to look like the expansion of a random number. On the other hand, generating power series of combinatorial objects are expected to be algebraic when these objects have a strong structure. Finally the algebraicity of formal power series with coefficients in a finite field is equivalent to a combinatorial property of the sequence of their coefficients, namely to be *automatic*: this is a theorem of Christol [23], completed by Christol, Kamae, Mendès France and Rauzy [24].

In other words, algebraicity of reals or formal power series associated with a given sequence depends on the type of expansion considered, as well as on the ground field, thus yielding quite distinct notions of complexity. For more on these questions we refer to the survey [4] and to the references therein. For more on the block complexity of expansions of algebraic irrationals, one can read in particular [2, 1, 14] and the references therein.

5 Inconstancy

Now another kind of complexity of a sequence is whether it varies "softly" or "crazily". What is a *fluctuating* (finite or infinite) sequence? How is it possible to detect and define sequences that admit large variations or fluctuations? A classical criterion is the *residual variance* of a sequence: this is a measure of the "distance" between the piecewise affine curve associated with the sequence and its linear regression line. Residual variance does not discriminate between a sequence that oscillates wildly and a sequence that grows very rapidly. We thus proposed in [6] to bring to light an old result of Cauchy and Crofton, in order to define what we called the *inconstancy* of a sequence. This definition is based upon the idea that a complicated curve is cut by a "random" straight line in many more points than a "quasi-affine" curve. We give below some more details. (For the original work concerning the Cauchy-Crofton theorem for curves, see [21, 22, 26], see also [57, 29] and the references in [6].)

5.1 Cauchy-Crofton's theorem. Inconstancy of a curve

Let Γ be a plane curve. Let $\ell(\Gamma)$ denote its length and let $\delta(\Gamma)$ denote the perimeter of its convex hull. Any straight line in the plane can be defined as the set of (x, y) such that $x \cos\theta + y \sin\theta - \rho = 0$, where θ belongs to $[0, 2\pi)$ and ρ is a positive real number, and hence is completely determined by (ρ, θ). Letting μ denote the Lebesgue measure on the set $\{(\rho, \theta),\ \rho \geq 0,\ \theta \in [0, 2\pi)\}$, the *average* number of intersection points between the curve Γ and straight lines is defined to be the quantity

$$\int_{D\in\Omega(\Gamma)} \sharp(\Gamma \cap D) \frac{\mathrm{d}\rho\ \mathrm{d}\theta}{\mu(\Omega(\Gamma))}$$

where $\Omega(\Gamma)$ is the set of straight lines which intersect Γ.

The following result can be found in [26, p. 184–185], see also the papers of Cauchy [21, 22].

Theorem 1 (Cauchy-Crofton) *The average number of intersection points between the curve Γ and the straight lines in $\Omega(\Gamma)$ satisfies the equality*

$$\int_{D\in\Omega(\Gamma)} \sharp(\Gamma \cap D) \frac{\mathrm{d}\rho\ \mathrm{d}\theta}{\mu(\Omega(\Gamma))} = \frac{2\ell(\Gamma)}{\delta(\Gamma)}.$$

Remark 3 The reader will have noted the relation between this theorem and the *Buffon needle* problem (see [13, p. 100–104]).

The theorem of Cauchy-Crofton suggests the following definition.

Definition 1 Let Γ be a plane curve. Let $\ell(\Gamma)$ be its length and $\delta(\Gamma)$ the perimeter of its convex hull. The *inconstancy* of the curve Γ, denoted $\mathcal{I}(\Gamma)$, is defined by

$$\mathcal{I}(\Gamma) := \frac{2\ell(\Gamma)}{\delta(\Gamma)}.$$

Remark 4 The minimal value of $\mathcal{I}(\Gamma)$ is 1. It is obtained when Γ is a segment.

5.2 Inconstancy of sequences. First results

The inconstancy of the sequence $(u_n)_{n\geq 0}$ is defined as the inconstancy of the broken line joining the points $(0, u_0)$, $(1, u_1)$, $(2, u_2)$... In [6] we compare inconstancy with residual variance for very simple sequences. Then we compute the inconstancy of classical infinite sequences. We give some of our results below.

Theorem 2 *Let $(u_n)_{n\geq 0}$ be an infinite sequence taking two values 0 and $h > 0$, with $u_0 = 0$. We make the assumption that the frequencies of occurrences of the blocks 00, hh, $0h$, $h0$ in the sequence exist and are respectively equal to $\mathcal{F}_{00}, \mathcal{F}_{hh}, \mathcal{F}_{0h}, \mathcal{F}_{h0}$. Then*

$$\mathcal{I}((u_n)_{n\geq 0}) = \mathcal{F}_{00} + \mathcal{F}_{hh} + (\sqrt{1+h^2})(\mathcal{F}_{0h} + \mathcal{F}_{h0}) = 1 + (\sqrt{1+h^2} - 1)(\mathcal{F}_{0h} + \mathcal{F}_{h0}).$$

Remark 5 A similar result holds with sequences taking any finite number of values.

Before stating a corollary, we recall the definitions of three classical sequences (see, e.g., [7] for more on these sequences).

- The *Thue-Morse sequence* $(m_n)_{n\geq 0}$ is defined by $m_n := s_n \bmod 2$, where s_n is the sum of the binary digits of the integer n.
- The *Shapiro-Rudin sequence* $(r_n)_{n\geq 0}$ is defined by $r_n := e_n \bmod 2$, where e_n is the number of (possibly overlapping) blocks 11 in the binary expansion of the integer n.
- The *regular paperfolding sequence* is the sequence $(z_n)_{n\geq 0}$ of "hills" (1) and "valleys" (0) obtained by unfolding a strip of paper folded an infinite number of times. It can be defined by the relations: for all $n \geq 0$, $z_{4n} = 0$, $z_{4n+2} = 1$, $z_{2n+1} = z_n$.

Corollary 1 *We have in particular the following results for binary sequences (taking only values 0 and 1). Let $(m_n)_{n\geq 0}$ be the Thue-Morse sequence; let $(r_n)_{n\geq 0}$ be the Shapiro-Rudin sequence; let $(z_n)_{n\geq 0}$ be the regular paperfolding sequence. Then*

$$\begin{array}{lllll}
\mathcal{I}((01)^\infty) & = & \sqrt{2} & = & 1.414... \\
\mathcal{I}((0^2 1)^\infty) & = & \frac{1+2\sqrt{2}}{3} & = & 1.276... \\
\mathcal{I}((m_n)_{n\geq 0}) & = & \frac{1+2\sqrt{2}}{3} & = & 1.276... \\
\mathcal{I}((0^3 1)^\infty) & = & \frac{1+\sqrt{2}}{2} & = & 1.207... \\
\mathcal{I}((r_n)_{n\geq 0}) & = & \frac{1+\sqrt{2}}{2} & = & 1.207... \\
\mathcal{I}((z_n)_{n\geq 0}) & = & \frac{1+\sqrt{2}}{2} & = & 1.207... \\
\mathcal{I}((u_n)_{n\geq 0}) & = & \frac{1+\sqrt{2}}{2} & = & 1.207... \quad \text{(for almost all sequences } (u_n)_{n\geq 0}) \\
\mathcal{I}(0^\infty)) & = & 1. & &
\end{array}$$

Remark 6 For binary sequences taking values 0 and 1, the minimal inconstancy is obtained for the constant sequences 0000... and 1111..., while the inconstancy of the sequences 01010101... and 10101010... is maximal.

5.3 Possible applications

We began checking whether inconstancy is a pertinent measure of fluctuation, or even a prediction tool in different domains: variations of BMI (*body mass index*) and metabolic syndrome (trying to address, using inconstancy, the relation with cardio-vascular diseases, studied, e.g., in [60]), smoothness of musical themes, and fluctuations of the stockmarket.

6 Miscellaneous addenda

We give in this section a sample of other notions of complexity or entropy that the reader might find interesting.

- The authors of [37, 19, 18] study the *folding complexity* of finite binary sequences: this is the minimal number of folds required to obtain a given pattern of "mountains and valleys" by repeatedly folding a strip of paper.
- In [53] the authors propose and study an "aperiodicity measure", which compares a given sequence with its (non-identical) shifts.
- The authors of [16] study the "conditional entropy" of classical sequences. This work was extended in [10], then in e.g., [58, 59, 17, 15].

7 Conclusion

We could only mention a few notions in the huge field of complexity which concerns both mathematics and computer science. Furthermore we restricted this survey to (some) *complexities of sequences*. It would be interesting to compare some of these notions with more "philosophical" notions, for example the work of Edgar Morin, in particular the *eight avenues of complexity* that he proposes in [50]: one of these avenues is "the irreducible character of randomness or disorder" which can be compared with the incompressibility of sequences in relation with the Kolmogorov-Solomonoff-Chaitin complexity. To know more about the work of Morin, the reader can read [51], see also [28].

8 Acknowledgments

We want to thank heartily the organizers of the conference *Functions in Number Theory and Their Probabilistic Aspects* for their kind invitation and their warm hospitality in Kyoto; we also want to thank warmly S. Laplante and L. Maillard-Teyssier, for interesting discussions, and for pointing some references.

References

[1] B. Adamczewski, Y. Bugeaud, On the complexity of algebraic numbers II. Continued fractions, *Acta Math.* **195** (2005) 1–20.

[2] B. Adamczewski, Y. Bugeaud, On the complexity of algebraic numbers I. Expansions in integer bases, *Annals of Math.* **165** (2007) 547–565.

[3] J.-P. Allouche, Sur la complexité des suites infinies, *Bull. Belg. Math. Soc.* **1** (1994) 133–143.

[4] J.-P. Allouche, Automates et algébricités, *J. Théor. Nombres Bordeaux* **17** (2005) 1–11.

[5] J.-P. Allouche, M. Baake, J. Cassaigne, D. Damanik, Palindrome complexity, *Theoret. Comput. Sci.* **292** (2003) 9–31.

[6] J.-P. Allouche, L. Maillard-Teyssier, Inconstancy of finite and infinite sequences, *Theoret. Comput. Sci.* **412** (2011) 2268–2281.

[7] J.-P. Allouche, J. Shallit, *Automatic Sequences. Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, 2003.

[8] S. Arora, B. Barak, *Computational Complexity: a Modern Approach*, Cambridge University Press, 2009.

[9] S. V. Avgustinovich, D. G. Fon-Der-Flaass, A. E. Frid, Arithmetical complexity of infinite words, in M. Ito and T. Imaoka, editors, *Words, Languages & Combinatorics III*, ICWLC 2000, Kyoto, Japan, March 14-18, 2000, World Scientific Publishing, Singapore, 2003, pp. 51–62.

[10] V. Berthé, Conditional entropy of some automatic sequences, *J. Phys. A* **27** (1994) 7993–8006.

[11] É. Borel, Sur les chiffres décimaux de $\sqrt{2}$ et divers problèmes de probabilités en chaîne, *C. R. Acad. Sci. Paris* **230** (1950) 591–593. Reprinted in *Œuvres d'É. Borel*, vol. 2, Éditions du CNRS, Paris, 1972, pp. 1203–1204.

[12] M. Boshernitzan, D. Ralston, Continued fractions and heavy sequences, *Proc. Amer. Math. Soc.* **137** (2009) 3177–3185.

[13] G. L. Leclerc Buffon, *Histoire naturelle générale et particulière*, Supplément, Tome quatrième, Imprimerie Royale, 1777.

[14] Y. Bugeaud, J.-H. Evertse, On two notions of complexity of algebraic numbers, *Acta Arith.* **133** (2008) 221–250.

[15] B. L. Burrows, J. I. Millington, Recursive procedures for measuring disorder in non-periodic sequences and lattices, *Physica A* **295** (2001) 488–506.

[16] B. L. Burrows, K. W. Sulston, Measurement of disorder in nonperiodic sequences, *J. Phys. A* **24** (1991) 3979–3987.

[17] B. L. Burrows, K. W. Sulston, Generalized measures for physical properties of nonperiodic chains, *Phys. Rev. B* **51** (1995) 5732–5738.

[18] J. Cardinal, E. D. Demaine, M. L. Demaine, S. Imahori, T. Ito, M. Kiyomi, S. Langerman, R. Uehara, T. Uno, Algorithmic Folding Complexity, *Graphs Combin.* **27** (2011) 341–351.

[19] J. Cardinal, E. D. Demaine, M. L. Demaine, S. Imahori, S. Langerman, R. Uehara, Algorithmic folding complexity, *Lecture Notes in Comput. Sci.* **5878** (2009) 452–461.

[20] J. Cassaigne, I. Kaboré, T. Tapsoba, On a new notion of complexity on infinite words, *Acta Univ. Sapientiae Math.* **2** (2010) 127–136.

[21] A. Cauchy, Notes sur divers théorèmes relatifs à la rectification des courbes, et à la quadrature des surfaces, *C. R. Acad. Sci. Paris* **13** (1841) 1060–1063. (Also in *Œuvres complètes* **6**, Gauthier-Villars, Paris, pp. 369–375, 1888.)

[22] A. Cauchy, Mémoire sur la rectification des courbes et la quadrature des surfaces courbes, *Mém. Acad. Sci. Paris* **22** (1850) 3–15. (Also in *Œuvres complètes* **2**, Gauthier-Villars, Paris, pp. 167–177, 1908.)

[23] G. Christol, Ensembles presque périodiques k-reconnaissables, *Theoret. Comput. Sci.* **9** (1979), 141–145.

[24] G. Christol, T. Kamae, M. Mendès France, G. Rauzy, Suites algébriques, automates et substitutions, *Bull. Soc. Math. France* **108** (1980), 401–419.

[25] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein, *Introduction to Algorithms*, Third edition, MIT Press, 2009.

[26] M. W. Crofton, On the theory of local probability, applied to straight lines drawn at random in a plane; the methods used being also extended to the proof of certain new theorems in the Integral Calculus, *Philos. Trans. R. Soc. Lond.* **158** (1868) 181–199.

[27] J.-P. Delahaye, H. Zenil, Numerical evaluation of algorithmic complexity for short strings: a glance into the innermost structure of randomness, Preprint, 2011, available at `http://arxiv.org/abs/1101.4795`

[28] L. Dobuzinskis, Where is Morin's road to complexity going, *World Futures* **60** (2004) 433–455.

[29] Y. Dupain, T. Kamae, M. Mendès France, Can one measure the temperature of a curve?, *Arch. Rational Mech. Anal.* **94** (1986) 155–163. Corrigenda, *Arch. Rational Mech. Anal.* **98** (1987) 395.

[30] S. Ferenczi, Complexity of sequences and dynamical systems, *Discrete Math.* **206** (1999) 145–154.

[31] S. Ferenczi, Z. Kása, Complexity for finite factors of infinite sequences, *Theoret. Comput. Sci.* **218** (1999) 177–195.

[32] P. Flajolet, R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009.

[33] D. G. Fon-Der-Flaass, A. E. Frid, On periodicity and low complexity of infinite permutations, *Eur. J. Comb.* **28** (2007) 2106–2114.

[34] A. S. Fraenkel, R. J. Simpson, How many squares must a binary sequence contain?, *Elec. J. Comb.* **2** (1995) #2.

[35] A. S. Fraenkel, J. Simpson, How many squares can a string contain?, *J. Combin. Theory Ser. A* **82** (1998) 112–120.

[36] P. Grassberger, Toward a quantitative theory of self-generated complexity, *Internat. J. Theoret. Phys.* **25** (1986) 907–938.

[37] T. Ito, M. Kiyomi, S. Imahori, R. Uehara, Complexity of pleat folding, in *PROC. 25th European Workshop on Computational Geometry (EuroCG'09), March 2009 (Brussels, Belgium)*, 2009, pp. 53–56.

[38] I. Kaboré, T. Tapsoba, Complexité par fenêtre, in *Proceedings of the 6th International Conference on Words (WORDS 2007)*, Inst. Math. Luminy, Marseille (France), 2007, pp. 184–188.

[39] T. Kamae, Behavior of various complexity functions, Preprint, 2011, available at `http://www14.plala.or.jp/kamae/variouscomplexity.pdf`

[40] T. Kamae, P. V. Salimov, On maximal pattern complexity of some automatic words, *Ergodic Theory Dynam. Systems*, to appear (DOI: 10.1017/S0143385710000453).

[41] T. Kamae, L. Q. Zamboni, Sequence entropy and the maximal pattern complexity of infinite words, *Ergodic Theory Dynam. Systems* **22** (2002) 1191–1199.

[42] T. Kamae, L. Q. Zamboni, Maximal pattern complexity of discrete dynamical systems, *Ergodic Theory Dynam. Systems* **22** (2002) 1201–1214.

[43] L. Ilie, Combinatorial complexity measures for strings, in *Recent Advances in Formal Languages and Applications*, Studies in Computational Intelligence, 2006, vol. 25/2006, pp. 149–170.

[44] L. Ilie, S. Yu, K. Zhang, Word complexity and repetitions in words, Computing and Combinatorics Conference—COCOON'02, *Internat. J. Found. Comput. Sci.* **15** (2004) 41–55.

[45] M. Li, P. Vitanyi, *An Introduction to Kolmogorov Complexity and its Applications*, Third edition, Springer Verlag, 2008.

[46] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J. C. Nuño, From time series to complex networks: the visibility graph, *Proc. Natl. Acad. Sci. USA* **105** (2008) 4972–4975.

[47] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, Horizontal visibility graphs: exact results for random time series, *Phys. Rev. E* **80** (2009) 046103-1–046103-11.

[48] M. Makarov, On permutations generated by binary infinite words, *Siberian Electron. Math. Rep.* **6** (2006) 304-311 (in Russian), available at `http://semr.math.nsc.ru/v3/p304-311.pdf`

[49] C. Mauduit, A. Sárközy, On finite pseudorandom binary sequences. I. Measure of pseudorandomness, the Legendre symbol, *Acta Arith.* **82** (1997) 365–377.

[50] E. Morin, De la complexité : complexus, in *Les Théories de la Complexité. Autour de l'œuvre d'Henri Atlan*, ed. F. Fogelman-Soulié, Seuil, Paris, 1991, pp. 283–296.

[51] E. Morin, *Introduction à la pensée complexe*, Points, Essais, 2005.

[52] M.-P. Muller, Suites de Sturm, complexité et dimension de Vapnik-Chervonenkis, Preprint, 2005, available at `http://hal.archives-ouvertes.fr/hal-00129709/fr/`

[53] Y. Pritykin, J. Ulyashkina, Aperiodicity measure for infinite sequences, *Lecture Notes in Comput. Sci.* **5675** (2009) 274–285.

[54] H. Prodinger, F. J. Urbanek, Infinite $0-1$-sequences without long adjacent identical blocks, *Discrete Math.* **28** (1979) 277–289.

[55] D. Ralston, Heaviness: an extension of a lemma of Y. Peres, *Houston J. Math.* **35** (2009) 1131–1141

[56] D. Ralston, Heaviness in symbolic dynamics: substitution and Sturmian systems, *Discrete Contin. Dyn. Syst. Ser. S* **2** (2009) 287–300.

[57] H. Steinhaus, Length, shape, and area, *Colloquium Math.* **3** (1954) 1–13.

[58] K. W. Sulston, B. L. Burrows, A. N. Chishti, Recursive procedures for measuring disorder in non-periodic sequences, *Physica A* **217** (1995) 146–160.

[59] P. Q. Tong, Measurement of disorder in three-component nonperiodic sequences, *Phys. Lett. A* **207** (1995) 159–164.

[60] A.-C. Vergnaud, S. Bertrais, J.-M. Oppert, L. Maillard-Teyssier, P. Galan, S. Hercberg, S. Czernichow, Weight fluctuations and risk for metabolic syndrome in an adult cohort, *Int. J. Obesity* **32** (2008) 315–321.

[61] S. Widmer, Permutation complexity of the Thue-Morse word, *Adv. Appl. Math.* **47** (2011) 309–329.