Optimizing Properties of Balanced Words

Nikita Sidorov
School of Mathematics
The University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom
sidorov@manchester.ac.uk

In the past few decades there has been a good deal of papers which are concerned with optimization problems in different areas of mathematics (along 0-1 words, finite or infinite) and which yield – sometimes quite unexpectedly – balanced words as optimal. In this note we list some key results along these lines known to date.

Firstly, we recall that a finite or infinite 0-1 word $w = w_1 w_2 \dots$ is called *balanced* if for every pair of finite subwords u, v such that $|u| = |v|$, we necessarily have $||u|_1 - |v|_1| \le 1$, where $|u|_1 = \#\{j : u_j = 1\}$ stands for the 1-*length* of u. An infinite balanced word which is not eventually periodic is called *Sturmian*.

There are several equivalent definitions of Sturmian sequences. Let $w = w_1 w_2 \dots$ be an infinite 0-1 sequence and put $p_w(n) = \#\{w_j \dots w_{j+n-1} : j \ge 1\}$ – the *complexity function* of w. Then w is Sturmian if and only if $p_w(n) = n+1$. The 1-*ratio* $\gamma = \lim_{n\to\infty} |w_1 \dots w_n|_1/n$ is well defined for any Sturmian sequence w; furthermore, any Sturmian sequence with the 1-ratio γ can be obtained by the formula

$$w_n = \lfloor (n+1)\gamma + \delta \rfloor - \lfloor n\gamma + \delta \rfloor \tag{1}$$

for some $\delta \in [0,1)$. For more details see, e.g., [10].

This survey paper is concerned with some optimization problems from various areas of mathematics and physics, in which balanced words (and, in some cases, Sturmian sequences) turn out to be optimizing.

1 Multimodular functions and queuing

We will begin with optimization problems in mathematics. Our first example comes from the seminal paper [5].

Define vectors $f_0, f_1, \dots, f_m$ in $\mathbb{Z}^m$ as follows: $f_0 = (-1,0,\dots,0,0)$, $f_1 = (1,-1,0,0,\dots,0)$, $f_2 = (0,1,-1,0,\dots,0), \dots, f_m = (0,0,\dots,0,1)$. Let now $\mathscr{F} = \{f_0, \dots, f_m\}$. We say that a function $J : \mathbb{Z}^m \to \mathbb{R}$ is *multimodular* if for any $u \in \mathbb{Z}^m$ we have

$$J(u+v) + J(u+w) \ge J(u) + J(u+v+w)$$

for all $v, w \in \mathscr{F}$ with $v \ne w$. The function

$$\underline{J}(z) = \sup\{A(z) : A \text{ is affine and } A(u) \le J(u) \text{ for all } u \in \mathbb{Z}^m\}$$

is called the *lower convex envelope* of J.

P. Ambrož, Š. Holub and Z. Masáková (Eds.):
8th International Conference WORDS 2011
EPTCS 63, 2011, pp. 240–246, doi:10.4204/EPTCS.63.30

© Nikita Sidorov
This work is licensed under the
Creative Commons Attribution License.

Theorem 1 ([5]) *Let J be a multimodular function on $\mathbb{Z}^m$ and let $\underline{J}$ denote its lower convex envelope. If x is any infinite 0-1 sequence with the 1-ratio γ, then*

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} J(x_k, x_{k+1}, \ldots, x_{k+m-1}) \geq \underline{J}(\widehat{\gamma}),$$

where $\widehat{\gamma} = (\gamma, \ldots, \gamma)$. Moreover, if $x = w$ given by (1) for some δ, then we have the equality.

The author then applies this result to the following queuing problem: consider a sequence of customers arriving at a fixed rate in such a way that the interarrival times are i.i.d. Poisson random variables with the same finite mean. A 0-1 input sequence $x = (x_1, x_2, \ldots)$ determines what happens to the kth customer, namely, if $x_k = 1$, the customer is admitted and if $x_k = 0$, he is sent elsewhere. The mean service time is assumed to be fixed. The number in the queue after customer k arrives, is $N_k + x_k$. Then the expectation of $\max N_k$ is a multimodular function.

Consequently, it follows from the above theorem that if a fraction γ of customers is sent to a server queue according to a splitting sequence x, then the long-term average is minimized when $x_k = \lfloor (n+1)\gamma \rfloor - \lfloor n\gamma \rfloor$, i.e., along Sturmian sequences.

2 Cyclic permutations of binary expansions

Let $w = w_1 \ldots w_m$ be a finite 0-1 word and put

$$b(w) = \sum_{k=1}^{m} w_k 2^{m-k}.$$

Now define

$$B(w) = b(w) \cdot b(w^{(2)}) \cdot b(w^{(m)}),$$

where $w^{(1)} = w, w^{(2)}, \ldots, w^{(m)}$ are the cyclic permutations of w. Let now $\mathscr{W}_{p,q}$ denote the set of 0-1 words of length q with the 1-length p. As is well known (see, e.g., [10]), there are precisely q balanced words in $\mathscr{W}_{p,q}$, all of which are in the same *orbit* (= all cyclic permutations of a word), so if $\mathbb{W}_{p,q}$ is defined to be the set of all orbits of words in $\mathscr{W}_{p,q}$, there is a unique balanced orbit in $\mathbb{W}_{p,q}$.

Theorem 2 ([8]) *Suppose $1 \leq p < q$ are coprime integers. For $w \in \mathbb{W}_{p,q}$, the product $B(w)$ is maximized precisely when w is balanced.*

For instance, put $p = 2, q = 5$. Here there are only two possible orbits, namely, those of $w = 10100$ and $v = 11000$. We have $B(w) = b(10100) \times b(01001) \times b(10010) \times b(00101) \times b(01010) = 20 \times 9 \times 18 \times 5 \times 10 = 162000$, whereas $B(v) = b(11000) \times b(10001) \times b(00011) \times b(00110) \times b(01100) = 24 \times 17 \times 3 \times 6 \times 12 = 88128$.

3 Maximizing measures

Let $T : [0,1) \to [0,1)$ and denote the space of all T-invariant measures by $\mathscr{M}(T)$. We say that a measure μ is *majorated* by ν (notation: $\mu \prec \nu$) if $\int_0^1 f \, d\mu \leq \int_0^1 f \, d\nu$ for any convex function $f : [0,1) \to [0,1)$. Put for any $\gamma \in (0,1)$,

$$\mathscr{M}_\gamma = \{\mu \in \mathscr{M}(T) : \mathrm{bar}(\mu) = \gamma\},$$

where

$$\mathrm{bar}(\mu) = \int_0^1 x \, d\mu(x),$$

i.e., the *barycentre* of μ.

Theorem 3 ([7]) *Let $Tx = 2x \bmod 1$. For any $\gamma \in [0,1]$, the partially ordered set $(\mathscr{M}_\gamma, \prec)$ has a least element. This least element is the Sturmian measure S_γ of rotation number γ.*

Here S_γ is the following. Let $\varphi : [0,1) \to [0,1)$ be defined as follows:

$$\varphi_\gamma(x) = \sum_{n=0}^{\infty} \frac{\chi_{[1-\gamma,1)}(x + n\gamma \bmod 1)}{2^{n+1}},$$

i.e., the binary sum of the standard symbolic sequence associated with the rotation by γ with a starting point x. Then the Sturmian measure S_γ is the push forward of the Lebesgue measure on $[0,1)$ under φ_γ.

If γ is rational, then S_γ sits on a finite set. For instance, the support of $S_{2/5}$ is the orbit of the binary sum of 00101 00101 00101… under T, i.e., the set $\{\frac{5}{31}, \frac{10}{31}, \frac{20}{31}, \frac{9}{31}, \frac{18}{31}\}$, each point having the $S_{2/5}$-measure of $1/5$.

If γ is irrational, then S_γ is supported by a Cantor set. For instance, if $\gamma = \frac{3-\sqrt{5}}{2}$, then $\mathrm{supp}(S_\gamma)$ is the closure of the set of all shifts of the Fibonacci word $f = 0010100100101\ldots$ (It is indeed a Cantor set because f has such a low complexity!)

There are other papers in this area which produce Sturmian sequences in similar optimization problems. For instance, one may replace the class of convex functions with increasing functions and consider the β-*transformation* given by $\tau_\beta x = \beta x \bmod 1$ (with $\beta > 1$) and the β-shift – the subshift on the alphabet $\{0,1,\ldots,\lceil\beta\rceil - 1\}$ which corresponds to the natural partition $[0,1) = [0,\beta^{-1}) \cup [\beta^{-1}, 2\beta^{-1}) \cup \cdots \cup [(\lceil\beta\rceil - 1)\beta^{-1}, 1)$ for τ_β.

It has been shown in [2] that the β-shift has a largest shift-invariant measure if and only if β is an algebraic integer of a special form. (In particular, if $1 < \beta < 2$, then it has to be *multinacci*, i.e., the dominant root of $x^m = x^{m-1} + \cdots + x + 1$ for some $m \geq 2$.) In this case the largest shift-invariant measure on the β-subshift is the unique one supported by the periodic shift-orbit generated by its lexicographically largest element. The supporting measure is always Sturmian.

Another direction in this line of research is concerned with imposing no extra conditions of the class of functions but instead considering a specific (usually, one-parameter) family of those. For example, let, as above, T be the doubling map, and let $g_\theta(x) = \cos 2\pi(x - \theta)$ or $f_\theta(x) = 1 - 4\mathrm{dist}_{\mathbb{T}}(x, \theta)$, where $\mathrm{dist}_{\mathbb{T}}$ is the distance on the circle $\mathbb{R}/\mathbb{Z}$. In both cases maximizing measures are Sturmian – see [3, 1] and references therein.

Thus, in questions concerning maximizing measures, the Sturmian measures seem to be a very robust class.

4 Tetris heaps

Consider a version of popular Tetris game with two pieces, 0 and 1 – see Fig. 1. We will be interested in stacking these pieces in such a way that the height of the heap is minimal.

More precisely, for a finite 0-1 word w we define $h(w)$ to be the height of the heap specified by w and put

$$\rho_{\min} = \liminf_{n\to\infty} \min_{w \in \{0,1\}^n} \frac{h(w)}{n}.$$

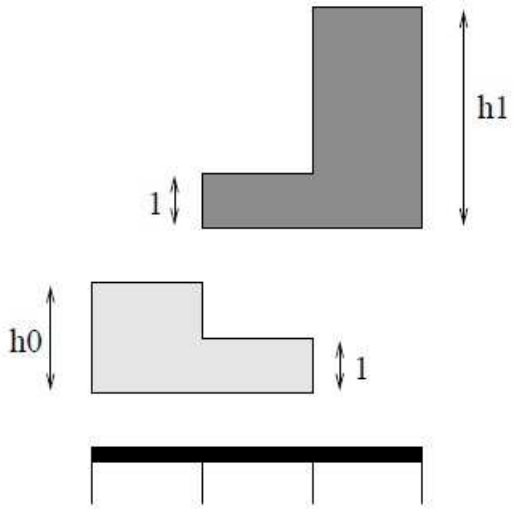

Figure 1: Tetris heaps

An *optimal schedule* is an infinite 0-1 word u such that

$$\lim_{n\to\infty} \frac{h(u[n])}{n} = \rho_{\min},$$

where $u[n]$ is the prefix of length n of u.

Theorem 4 ([11]) *Let us consider a heap model with two pieces. There exists an optimal schedule which is balanced – either periodic or Sturmian.*

The authors of [11] characterize the cases where the optimal is periodic and the ones where it is Sturmian. The proof is constructive, providing an explicit optimal schedule. For a more general approach (using a special class of iterated function systems) see [4].

5 Joint spectral radius

Given a finite set of $d \times d$ real matrices $\mathscr{A} = \{A_0, \dots, A_{r-1}\}$, we define the *joint spectral radius* $\rho(\mathscr{A})$ to be the quantity

$$\rho(\mathscr{A}) := \limsup_{n\to\infty} \max \left\{ \|A_{i_1} \cdots A_{i_n}\|^{1/n} : i_j \in \{0, \dots, r-1\} \right\},$$

a definition introduced by G.-C. Rota and G. Strang in 1960.

A $\{0, \dots, r-1\}$-sequence $(i_1, i_2, \dots)$ is called *maximizing* if

$$\lim_{n\to\infty} \|A_{i_1} \cdots A_{i_n}\|^{1/n} = \rho(\mathscr{A}).$$

This is somewhat similar to the Tetris model considered above, which is reflected in the general model accounting for both set-ups considered in [4].

Example 5 *Put* $\mathscr{A} = \{A_0, A_1\}$, *where*

$$A_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \; A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then it turns out that $010101\ldots$ *is maximizing and furthermore, any maximizing sequence for* $\mathscr{A}$ *has the same growth rate of the corresponding sequence of matrix products. Consequently,* $\rho(\mathscr{A}) = (\rho(A_0A_1))^{1/2} = (1+\sqrt{5})/2$.

Now consider the one-parameter family of pairs $\mathscr{A}_\alpha := \{A_0, \alpha A_1\}$ with $\alpha \in [0,1]$. Is it true that for any fixed α any maximizing sequence is "essentially periodic" like in the case $\alpha = 1$?

More precisely, a set of matrices $\mathscr{A}$ is said to have the *finiteness property* if there exists an eventually periodic maximizing sequence for $\mathscr{A}$. It was shown in the PhD thesis [14] that if $\alpha \in [4/5, 1]$, then we have the same conclusion as for $\alpha = 1$. What about the case $\alpha < 4/5$?

We say that an infinite 0-1 word w is *recurrent* if any of its subwords occurs in w infinitely often. For each $\gamma \in [0,1]$, let X_γ denote the set of all recurrent balanced infinite words whose 1-ratio is equal to γ. For any rational γ the set X_γ is finite and it is a continuum (called a *Sturmian system*) for any irrational γ.

We say that γ is an *optimal 1-ratio* for $\mathscr{A}_\alpha$ if there exists a maximizing sequence for $\mathscr{A}_\alpha$ with the 1-ratio γ.

Theorem 6 ([9]) *There exists a continuous, non-decreasing surjection* $\mathfrak{r} : [0,1] \to [0, \frac{1}{2}]$ *such that for each* α, $\gamma = \mathfrak{r}(\alpha)$ *is the unique optimal* 1*-ratio of* $\mathscr{A}_\alpha$.

Furthermore, for each $\alpha \in [0,1]$, *every element of* X_γ *is a maximizing sequence.*

Thus, if one takes any irrational $\gamma \in (0, 1/2)$, then there exists (in fact, unique – see [12]) $\alpha \in [0,1]$ such that any maximizing sequence for $\mathscr{A}_\alpha$ has the 1-ratio γ, i.e., cannot be periodic. This disproves the *Finiteness Conjecture* which asserts that any maximizing sequence for an arbitrary set of matrices should be periodic. (The first counterexample to the FC appeared in [4], however it was not explicit.)

Moreover, there is an explicit formula for such an α in terms of the elements of the continued fraction expansion of γ. Namely, let

$$\gamma = [a_1, a_2, \ldots]$$

denote the continued fraction expansion of γ with p_n/q_n being the nth convergent. Recall that the sequence of *standard words* specified by γ is given by $s_{-1} = 1, s_0 = 0, s_{n+1} = s_n^{a_{n+1}} s_{n-1}$, $n \geq 0$. It is obvious that s_n is a prefix of s_{n+1} and that the length of s_n tends to the infinity. It is also well known that the 1-ratio of s_n is p_n/q_n (see [10]).

Put $s_\infty = \lim_{n\to\infty} s_n$ (its 1-ratio is thus γ) and define the sequence of 2×2 matrices as follows: $B_{-1} = A_1, B_0 = A_0$, and

$$B_{n+1} = B_n^{q_{n+1}} B_{n-1}, \quad n \geq 0.$$

Put $\rho_n = \rho(B_n)$ (the spectral radius of B_n) and $\tau_n = \mathrm{tr}(B_n)$.

Theorem 7 ([9, 12]) *If* $\mathfrak{r}^{-1}(\gamma) = \{\alpha\}$, *we have*

$$\alpha = \lim_{n\to\infty} \left(\frac{\rho_n^{q_{n+1}}}{\rho_{n+1}^{q_n}} \right)^{(-1)^n} = \prod_{n=0}^{\infty} \left(\frac{\rho_n^{d_{n+1}} \rho_{n-1}}{\rho_{n+1}} \right)^{(-1)^n q_n}.$$

In particular, if $\gamma_* = \frac{3-\sqrt{5}}{2}$ (i.e., $d_n \equiv 1$ for all n), then (see [9])

$$\alpha_* := \lim_{n\to\infty} \left(\frac{\tau_n^{F_{n+1}}}{\tau_{n+1}^{F_n}} \right)^{(-1)^n} = \prod_{n=1}^{\infty} \left(1 - \frac{\tau_{n-1}}{\tau_n \tau_{n+1}} \right)^{(-1)^n F_{n+1}},$$

where $F_0 := 0, F_1 := 1$ and $F_{n+1} := F_n + F_{n-1}$ is the Fibonacci sequence and, as it turns out, $\tau_0 := 1$, $\tau_1, \tau_2 := 2$ and $\tau_{n+1} := \tau_n \tau_{n-1} - \tau_{n-2}$. The infinite product converges superexponentially fast, and

$$\alpha_* \simeq 0.749326546330367557943961948091344672091327\ldots$$

Similarly, one can easily compute $\mathfrak{r}^{-1}(\gamma)$ for any irrational γ with a very high precision. For more detail see [12].

6 One-dimensional Wigner lattices

Consider the one-dimensional lattice $\mathbb{Z}_+$ whose each node is either occupied by an electron or is empty ("occupied by a hole"). The electrons' interaction is given by a potential V which is assumed to be convex and vanish at the infinity, which is a rather weak assumption (such is, for instance, a Coulomb potential). More precisely, the energy of the system is given by

$$E = \frac{1}{2} \sum_{x_i, x_j} V(|x_i - x_j|).$$

Let $\gamma \in \mathbb{Q}$ denote the (fixed) ratio of electrons on the lattice.

Then, as noticed by Hubbard [6], the lowest-energy configuration with respect to this potential is attained at balanced words (see Table II from the cited paper). It should be mentioned that the author does not use this terminology and provides a proof of his claim only for certain special cases of γ.

The same result has been rediscovered by several authors for various physical models – see, e.g., [13].

Summing up, there are numerous – seemingly unrelated – areas of mathematics and physics in which optimization problems yield balanced words. Gaining a better understanding of this phenomenon looks like a perspective line of research.

References

[1] V. Anagnostopoulou, K. Diaz-Ordaz, O. Jenkinson and C. Richard, *Sturmian maximizing measures for the piecewise-linear cosine family*, preprint, see http://www.maths.qmul.ac.uk/˜omj/.

[2] V. Anagnostopoulou and O. Jenkinson, *Which beta-shifts have a largest invariant measure?* J. London Math. Soc. **79** (2009), 445–464. doi:10.1112/jlms/jdn070

[3] T. Bousch, *Le poisson n'a pas d'arêtes*, Ann. Inst. H. Poincaré Probab. Statist. **26** (2000), 489-508. doi:10.1016/S0246-0203(00)00132-1

[4] T. Bousch and J. Mairesse, *Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture*, J. Amer. Math. Soc. **15** (2002), 77–111. doi:10.1090/S0894-0347-01-00378-2

[5] B. Hajek, *Extremal splittings of point processes*, Math. Oper. Res. **10** (1985), 543–556. doi:10.1287/moor.10.4.543

[6] J. Hubbard, *Generalized Wigner lattices in one dimension and some applications to tetracyanoquinodimethane (TCNQ) salts*, Phys. Rev. B **17** (1978), 494–505. doi:10.1103/PhysRevB.17.494

[7] O. Jenkinson, *A partial order on $\times 2$-invariant measures*, Math. Res. Lett. **15** (2008), 893–900.

[8] O. Jenkinson, *Balanced words and majorization*, Discr. Math. Alg. Appl. **1** (2009), 485–498. doi:10.1142/S179383090900035X

[9] K. Hare, I. Morris, N. Sidorov and J. Theys, *An explicit counterexample to the Lagarias-Wang finiteness conjecture*, Adv. Math. **226** (2011), 4667–4701. doi:10.1016/j.aim.2010.12.012

[10] M. Lothaire, *Algebraic combinatorics on words*, *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, 2002. doi:10.1145/1412700.1412706

[11] J. Mairesse and L. Vuillon, *Asymptotic behavior in a heap model with two pieces*, Theoret. Comp. Sci. **270** (2002), 525-560. doi:10.1016/S0304-3975(01)00004-4

[12] I. Morris and N. Sidorov, *A Devil's staircase associated to the joint spectral radius of a pair of matrices*, in preparation.

[13] V. L. Pokrovsky and G. V. Uimin, *On the properties of monolayers of adsorbed atoms*, J. Phys. C **11** (1978), 3535–3549. doi:10.1088/0022-3719/11/16/022

[14] J. Theys, *Joint Spectral Radius: theory and approximations*, PhD thesis, Université Catholique de Louvain, 2005.