Available online at www.sciencedirect.com
ScienceDirect
Discrete Mathematics 308 (2008) 4419–4429
DISCRETE MATHEMATICS
www.elsevier.com/locate/disc

Thue type problems for graphs, points, and numbers

Jarosław Grytczuk[a, b]

[a]*Theoretical Computer Science Department, Faculty of Mathematics and Computer Science, Jagiellonian University, 30-387, Kraków, Poland*
[b]*Institute of Engineering, State Higher Vocational School in Nowy Sącz, 33-300 Nowy Sącz, Poland*

Received 13 January 2004; accepted 8 August 2007
Available online 26 November 2007

Dedicated to Miki Simonovits on the occasion of his 60th birthday

Abstract

A sequence $S=s_1s_2\dots s_n$ is said to be *nonrepetitive* if no two adjacent blocks of S are the same. A celebrated 1906 theorem of Thue asserts that there are arbitrarily long nonrepetitive sequences over the set $\{0, 1, 2\}$. This result is the starting point of Combinatorics on Words—a wide area with many deep results, sophisticated methods, important applications and intriguing open problems.

The main purpose of this survey is to present a range of new directions relating Thue sequences more closely to Graph Theory, Combinatorial Geometry, and Number Theory. For instance, one may consider graph colorings avoiding repetitions on paths, or colorings of points in the plane avoiding repetitions on straight lines. Besides presenting a variety of new challenges we also recall some older problems of this area.

MSC: Primary, 05C38;; 15A15; Secondary, 05A15;; 15A18

Keywords: Nonrepetitive sequence; Graph coloring; Avoidable pattern

1. Introduction

A sequence $r_1r_2\dots r_{2n}$ such that $r_i = r_{n+i}$ for all $i = 1, \dots , n$, is called a *repetition*. A sequence S is *nonrepetitive* if none of its blocks forms a repetition, where by a *block* we mean any subsequence of consecutive terms of S. In other words, no two adjacent blocks (of any length) in S are the same. It is trivial to notice that every binary sequence of length more than three contains a repetition. However, as proved about 100 years ago by Thue [67,68], there are arbitrarily long nonrepetitive sequences over three symbols. This result is the starting point of Combinatorics on Words—a wide area with many interesting connections and applications (cf. [13,18,47,48]).

The proof of Thue's theorem is based on a peculiar family of binary sequences

$$0, 01, 0110, 01101001, \dots,$$

where the next member is the previous one followed by its negation. This construction was actually known earlier to Prouhet [61], who used it in a number-theoretic problem on sums of powers (cf. Section 2). Later it was rediscovered by Morse [53] on the occasion of his work on recurrent geodesics, leading to the birth of Symbolic Dynamics (cf. [14,36]). Another spectacular application of Thue sequences was made by Novikov and Adjan [56,57] in a solution

E-mail address: grytczuk@tcs.uj.edu.pl.

0012-365X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.disc.2007.08.039

of the famous Burnside problem for groups (cf. [44]). The same idea was used earlier by Morse and Hedlund [54] in the easier case of semigroups. Curiously, in [56] nonrepetitive sequences are attributed to Arszon [7], who probably discovered them while working on Khintchine's problem on continued fractions (see Section 6.2). Actually, the list of independent discoverers of Thue sequences is long and diverse, including some famous nonmathematicians like Max Euwe (chess master) and Per Norgård (composer) (cf. [3]). Other historical facts and striking connections can be found in several excellent expository articles and monographs (cf. [2,3,9,11–13,18,47,48]).

Our aim in this paper is to consider "nonrepetitiveness" in a wider mathematical context, with the intention of showing its universal combinatorial appeal. We hope to achieve this goal by presenting a variety of new challenges, as well as by recalling some older open problems of this area.

2. Genesis

In this section we describe briefly the method of generating nonrepetitive sequences discovered by Thue. The whole story goes back to the following problem on sums of powers: given positive integers r and N, find all non-negative integer solutions x_i, y_i of the system of $N+1$ equations

$$x_1^k + \cdots + x_r^k = y_1^k + \cdots + y_r^k,$$

where $k = 0, 1, \ldots, N$. The result of Prouhet [61] deals with the case $r = 2^n$ and $N = n - 1$. It asserts that for every $n \geqslant 1$ there is a partition of the interval $\{1, \ldots, 2^n\}$ into two parts A_n and B_n such that

$$\sum_{x \in A_n} x^k = \sum_{y \in B_n} y^k$$

for $k = 0, 1, \ldots, n-1$. These partitions are defined inductively as follows. Let $A_1 = \{1\}$, $B_1 = \{2\}$ and

$$A_{n+1} = A_n \cup (2^n + B_n), \quad B_{n+1} = B_n \cup (2^n + A_n)$$

for $n > 0$. Thus, $A_2 = \{1, 4\}$, $B_2 = \{2, 3\}$, $A_3 = \{1, 4, 6, 7\}$, $B_3 = \{2, 3, 5, 8\}$, and so on. It is not hard to show by induction that the resulting sets A_n, B_n satisfy the desired property for every $n \geqslant 1$.

Prouhet's partitions are conveniently represented by binary sequences $T_n = t_1 t_2 \ldots t_{2^n}$, where $t_i = 0$ if $i \in A_n$, and $t_i = 1$ for $i \in B_n$. So, $T_1 = 01$, $T_2 = 0110$, $T_3 = 01101001$. In general we have $T_{n+1} = T_n T_n'$, where T_n' is a "negation" of T_n. Note that this process defines uniquely the *infinite* sequence $T = \lim_{n\to\infty} T_n$.

In [67] Thue observed the following striking property of the sequence T. Let $A = t_i t_{i+1} \ldots t_{i+n}$ and $B = t_j t_{j+1} \ldots t_{j+m}$ be two blocks of T, with $i \neq j$. We say that A and B are *overlapping* if the intervals $\{i, \ldots, i+n\}$ and $\{j, \ldots, j+m\}$ have nonempty intersection. Thue's result asserts that no two overlapping blocks of T are the same. The proof starts with an observation that T_{n+1} is an effect of substituting a block 01 for each 0 in T_n, and a block 10 for each 1. For instance,

$$\begin{array}{ccccc} T_2 = & 0 & 1 & 1 & 0 \\ & \downarrow & \downarrow & \downarrow & \downarrow . \\ T_3 = & 01 & 10 & 10 & 01 \end{array}$$

Next, one shows that the property of being *overlap-free* is preserved under this substitution, that is, if S is *any* finite binary sequence without identical overlapping blocks then its substitution image also possesses this property. Since $T_1 = 01$ is overlap-free the same is true of each T_n, and consequently of the infinite sequence T. Note that this property implies that no 3 consecutive blocks of T are the same.

Finally, consider the sequence $S = 210201\ldots$ obtained by counting the number of 1's between consecutive 0's in T. By the property of T the sequence S is an infinite nonrepetitive sequence over symbols 0, 1, 2.

3. Variations

This section contains several generalizations of Thue sequences and the related open problems. Most of them are well-known and were extensively studied.

3.1. Anagrams and sums

A stronger form of "nonrepetitiveness" is defined as follows. An *anagram* of a sequence $r_1 r_2 \dots r_n$ is a sequence $r_{\sigma(1)} r_{\sigma(2)} \dots r_{\sigma(n)}$, where σ is any permutation of the elements $1, 2, \dots, n$. A sequence S is said to be *strongly nonrepetitive* if no two adjacent blocks of S are anagrams of the same sequence.

One checks easily that no analog of Thue's theorem is possible over the set $\{0, 1, 2\}$. However, in [34] Evdokimov showed that there are arbitrarily long strongly nonrepetitive sequences over 25 symbols. Later Pleasants [60] lowered this number to 5. Whether 4 symbols will do was an open question posed by Erdős [31] and restated by Brown [16], until Keränen [43] found a substitution over 4 symbols preserving the property of avoiding anagram repetitions.

Keränen's result may be interpreted in arithmetic terms as follows: there is a function $f : \mathbb{N} \to \{2, 3, 5, 7\}$ such that

$$\prod_{i \in I} f(i) \neq \prod_{j \in J} f(j),$$

for any two adjacent segments of positive integers I and J. But what if we try to avoid equal *sums* instead of equal products in adjacent segments? A simple application of van der Waerden's theorem on arithmetic progressions shows that for any function mapping $\mathbb{N}$ into a finite subset of the integers there will be arbitrarily many consecutive segments with equal sums (cf. [41,47]). However, the following question looks intriguing.

Problem 1. *Is there a natural number k and a function $f : \mathbb{N} \to \{1, 2, \dots, k\}$ such that*

$$\sum_{i \in I} f(i) \neq \sum_{j \in J} f(j)$$

for any two adjacent segments I and J, with $|I| = |J|$?

The problem was posed independently in [59,41]. As I learned from Lorenz Halbeisen (personal communication), the question is actually due to his wife Stephanie Halbeisen.

A weaker form of the problem has a positive answer: there are infinite binary sequences without 4 consecutive blocks that are anagrams of the same sequence (at least 2 of any 4 consecutive blocks of the same length have different sums), as proved by Dekking [28]. This fact has a striking anti-Ramsey interpretation: there exists an infinite sequence Q_n of lattice points in the plane such that $Q_{n+1} - Q_n \in \{(1, 0), (0, 1)\}$, yet there are no 5-term arithmetic progressions among terms of Q_n. A positive answer to Problem 1 would yield a similar result for 3-term arithmetic progressions of lattice points in the plane.

Another equivalent formulation of the problem using arithmetic progressions goes as follows: is there an increasing sequence $S = (s_i)_{i \in \mathbb{N}}$ of positive integers such that: (1) there is a constant L such that $s_{i+1} - s_i \leqslant L$, for all $i \in \mathbb{N}$, (2) the number of terms of S between s_i and s_j is different than the number of terms of S between s_j and s_k, whenever s_i, s_j, s_k form an arithmetic progression?

3.2. Dejean conjecture

In this variation one looks for infinite sequences over a fixed number of symbols in which identical blocks are as far apart as possible. Suppose S is an infinite sequence over k symbols and let $d_S(n)$ be the minimum number of terms separating two identical blocks of length n in S. If $d_S(n) \geqslant 1$ for all n, then S is nonrepetitive.

The problem of maximizing $d_S(n)$ goes back to Thue [68], who constructed for every $k \geqslant 3$ an infinite sequence S over k symbols satisfying $d_S(n) \geqslant k - 2$. For $k = 3$ Dejean [27] extended this result significantly by giving a sequence S in which $d_S(n) \geqslant \frac{1}{3}n$. Moreover, she made the following conjecture which, if true, is best possible.

Problem 2 (*Dejean Conjecture*). *For every $k \geqslant 5$ there is an infinite sequence S over k symbols such that any two identical blocks of length n in S are separated by at least $(k - 2)n$ terms.*

The conjecture has been confirmed so far only for $k = 5, 6, \dots, 11$ by Ollagnier [58]. Surprisingly, if we neglect short blocks then much stronger property holds already for binary sequences. The theorem of Beck [10] asserts that for any $\varepsilon > 0$ there is a constant $n_0(\varepsilon)$ and an infinite binary sequence S such that $d_S(n) > (2-\varepsilon)^n$, provided $n > n_0(\varepsilon)$. The proof of this striking fact uses the Lovász Local Lemma—a powerful weapon in the arsenal of the probabilistic method (cf. [6,52]). A similar result appears also as an exercise in [6]. It asserts that for any $\varepsilon > 0$ there is $m_0(\varepsilon)$ and an infinite binary sequence S such that any two adjacent blocks of S of length $m > m_0(\varepsilon)$ differ in at least $(\frac{1}{2} - \varepsilon)m$ places. Note that both results imply the existence of infinite nonrepetitive sequences over a finite number of symbols.

3.3. Pattern avoidance

The following generalization of Thue sequences was introduced independently by Zimin [69] and Bean et al. [9] (cf. [22,48,65]). A *pattern* is an arbitrary finite sequence $P = p_1p_2 \ldots p_n$ over any set of symbols. A sequence S *resembles* a pattern P if S may be divided into nonempty blocks $S = B_1B_2 \ldots B_n$ such that $p_i = p_j$ implies $B_i = B_j$ for all $i, j = 1, 2, \ldots, n$. For instance, the sequence $S = 0120121010$ resembles the pattern $P = xxyy$ via factorization

$$S = \overset{x}{\boxed{012}}\overset{x}{\boxed{012}}\overset{y}{\boxed{10}}\overset{y}{\boxed{10}}\,.$$

A sequence S *avoids* a pattern P if no block of S resembles P. The minimum number of symbols in an infinite sequence S avoiding P is called the *avoidance index* of P, denoted by $\mu(P)$. For instance, $\mu(xx) = 3$ and $\mu(xxx) = 2$, by Thue's theorem. A pattern P is *avoidable* if its index $\mu(P)$ is finite. Otherwise it is *unavoidable*.

One natural family of unavoidable patterns is

$$Z = \{x, xyx, xyxzxyx, \ldots\}.$$

It was proved in [9,69] that each unavoidable pattern (up to renaming of symbols) is a block of some member of Z.

Problem 3. *Is there an absolute constant c such that $\mu(P) \leqslant c$ for any avoidable pattern P?*

The problem was posed in [8]. Surprisingly, it is not so easy to produce avoidable patterns with large avoidance index. The simplest pattern with index 4 is $xytyzuzxvyxwxz$, and for a long time no pattern with $\mu(P) > 4$ was known. However, several patterns of index 5 were finally found by Clark in [19] which is a current record.

3.4. Arithmetic progressions

Another variant of Thue sequences appeared in a recent paper of Currie and Simpson [26]. Given a positive integer k, a sequence S is called nonrepetitive up to mod k if all subsequences of S "situated" on arithmetic progressions of differences $d = 1, \ldots, k$ are nonrepetitive. Let $M(k)$ be the minimum number of symbols needed to built arbitrarily long sequences which are nonrepetitive up to mod k. Thus $M(1) = 3$ by Thue's theorem.

Problem 4. *Determine $M(k)$ for all $k \geqslant 1$.*

It is easy to see that $M(K) \geqslant k + 2$ and explicit constructions with $k + 2$ symbols was found for $k = 2, 3$ and 5 in [25,26]. Maybe $M(k) = k + 2$ for all k? The Local Lemma gives only a linear bound (cf. [38]) and it is not even clear if $M(k) = k + \mathrm{O}(1)$.

4. Graphs

In this section we present graph theoretic variations on the theme of Thue. The first problem involving graphs and nonrepetitive sequences was studied by Currie [20,21]. A result proved in [21] asserts that every connected graph (except a path with at most 4 vertices) contains arbitrarily long nonrepetitive walks. Below we present several new challenges relating Thue sequences and graph colorings.

4.1. Nonrepetitive colorings

A coloring of the vertices of a graph G is *nonrepetitive* if the sequence of colors along any simple path in G is nonrepetitive. The minimum number of colors needed is called the *Thue chromatic number* of G, denoted by $\pi(G)$. Thue's theorem asserts that $\pi(P_n) = 3$, for all $n \geqslant 4$, where P_n is the path with n vertices. It follows that $\pi(C_n) \leqslant 4$ for any cycle C_n. Actually, $\pi(C_n) = 4$ for only six exceptional values $n = 5, 7, 9, 10, 14, 17$, as proved by Currie [23].

Denote by $\pi(d)$ the supremum of $\pi(G)$ over all graphs G of maximum degree at most d. For instance, $\pi(2) = 4$ by the above remarks.

Problem 5. *Determine* $\pi(3)$.

Notice that it is not clear *a priori* that $\pi(d)$ is finite for $d \geqslant 3$. However, Alon et al. [5] proved that there are absolute constants $c_1, c_2 > 0$ such that

$$c_1 \frac{d^2}{\ln d} \leqslant \pi(d) \leqslant c_2 d^2.$$

The proof is probabilistic and makes use of the Local Lemma (for the upper bound) and Random Graphs (for the lower bound) (cf. [6,51]). No constructive method for $d \geqslant 3$ is known.

Perhaps the most intriguing problem involving Thue chromatic number concerns planar graphs.

Problem 6. *Is there an absolute constant k such that any planar graph has a nonrepetitive vertex k-coloring*?

Using Thue sequences without palindromes one shows easily that $\pi(T) \leqslant 4$ for any tree T. Moreover, as proved by Kündgen and Pelsmajer [45], $\pi(G) \leqslant 12$ for any outerplanar graph G, and $\pi(G) \leqslant 4^t$ for a graph of treewidth t.

Let H be a fixed graph and let $Forb_{\preccurlyeq}(H)$ denote a class of graphs not containing a minor isomorphic to H. A famous result of Robertson and Seymour [63] asserts that $Forb_{\preccurlyeq}(H)$ has bounded treewidth if and only if H is planar. This implies that planar graphs form the smallest class closed under taking minors for which boundedness of Thue chromatic number is not clear.

4.2. Thue threshold

Let $k \geqslant 2$ be a fixed integer. A sequence consisting of k identical blocks is called a *k-power*. A coloring of the vertices of a graph G is *k-power-free* if no path in G looks like a k-power. Let $\pi_k(G)$ denote the minimum number of colors in k-power-free coloring of G. For instance, $\pi_3(P_n) = 2$ for $n \geqslant 3$, by Thue's theorem.

Let $\mathscr{P}$ be a class of graphs. Let $\pi_k(\mathscr{P}) = \sup\{\pi_k(G) : G \in \mathscr{P}\}$ and let $t(\mathscr{P}) = \inf\{\pi_k(\mathscr{P}) : k \in \mathbb{N}\}$. In words, $t(\mathscr{P})$ is the smallest number of colors allowing to avoid k-powers in a class $\mathscr{P}$ for some (possibly huge) k. We call it the *Thue threshold*of a class of graphs $\mathscr{P}$. Notice that $t(\mathscr{P})$ is infinite if and only if $\pi_k(\mathscr{P})$ is infinite for all k.

Problem 7. *Determine the Thue threshold for planar graphs.*

At present it is not even known if $t(\mathscr{P})$ is finite in this case. We only know that at least 4 colors are necessary, since there are planar graphs containing arbitrarily long monochromatic paths in any 3-coloring of the vertices.

Let $t(d)$ denote the Thue threshold of a class of graphs with maximum degree at most d. A result of Currie and Fitzpatrick [24] imply that $t(2) = 2$ and this is the only known value of the function $t(d)$. In general, we have

$$\tfrac{1}{2}(d+1) \leqslant t(d) \leqslant d+1,$$

as proved in [4]. The following question was asked by Łuczak.

Problem 8. *Is it true that* $\lim_{d\to\infty} t(d)/d$ *exists*?

It seems plausible that the answer is in the affirmative. More risky would be to conjecture that $t(d) = d$ for all d.

Problem 9. *Determine $t(3)$.*

4.3. Edge colorings

In a similar way we may consider nonrepetitive *edge* colorings of graphs, where repetitions (or k-powers) are forbidden on paths. Note that in this case repetitions forming full cycles are allowed. We denote the minimum number of colors in a nonrepetitive edge coloring of G by $\pi'(G)$. Let $\pi'(d) = \sup\{\pi'(G) : \Delta(G) \leqslant d\}$, where $\Delta(G)$ denotes the maximum degree of G. The probabilistic proof of the upper bound for $\pi'(d)$ goes the same as for the vertex case (cf. [5]), hence $\pi'(d) = \mathrm{O}(d^2)$. However, we expect a much better estimate.

Problem 10. *There is an absolute constant c such that $\pi'(d) \leqslant cd$.*

As noted by Łuczak (personal communication), such constant c must be greater than $\frac{4}{3}$. In a special case of complete graphs we have $\pi'(K_n) \leqslant 2n$ for all n (cf. [5]).

A seemingly related problem for palindromes was considered earlier (without any Thue type motivation) by Erdős and Nešetřil. A *palindrome* is a sequence that looks the same when written backward, for instance, 0123210. An edge coloring of a graph G is *palindrome-free* if no simple path with at least two edges in G looks like a palindrome. An equivalent condition is that each color class forms an *induced* matching in G. In this form the concept was introduced by Erdős and Nešetřil in [32]. The minimum number of colors needed is called the *strong chromatic index* of G, denoted by $s\chi_e(G)$ (cf. [52]). It is conjectured that $s\chi_e(G) \leqslant \frac{5}{4}\Delta(G)^2$, which is sharp, since there are graphs of arbitrarily large maximum degree for which the equality holds (cf. [32]).

4.4. Nonrepetitive walks

In [15] Brešar and Klavžar introduced the following extension of nonrepetitive colorings of graphs. Let $w = v_1 v_2 \ldots v_n$ be a walk in a graph G, that is, $e_i = v_i v_{i+1}$ is an edge of G for $i = 1, \ldots, n-1$. A walk w is *open* if $v_1 \neq v_n$. If f is a coloring of the edges of a graph G then a walk w determines a sequence $S(w) = f(e_1) \ldots f(e_{n-1})$. A coloring f is *square-free* if for no open walk w the associated sequence $S(w)$ is a repetition. Let $\pi_w(G)$ denote the minimum number of colors needed for a square-free coloring of G.

Problem 11. *Is $\pi_w(G)$ bounded for graphs of bounded maximum degree*?

Another variation of this type was proposed by Currie (cf. [15]). Consider an edge coloring of a graph G such that the sequence of colors $S(w)$ on any walk w is nonrepetitive whenever w itself is nonrepetitive (as a sequence of uncolored edges). A similar question for Currie's colorings is also open.

4.5. 123-conjecture

Suppose f is a coloring of the edges of a graph $G = (V, E)$. For a vertex $u \in V$ denote by $M_f(u)$ the *multiset* of colors "around" u, that is, $M_f(u) = \{f(ux) : x \in N(u)\}$, where $N(u)$ is the set of all neighbors of u. A coloring f *distinguishes* neighbors in G if $M_f(u) \neq M_f(v)$ for any pair of adjacent vertices u and v. This notion was introduced by Karoński et al. [42] as a variant of a general vertex distinguishing problem. In abstract sense the concept resembles strongly nonrepetitive sequences (Section 3.1) where adjacent intervals are distinguished by multisets of symbols.

The main result of [42] establishes the existence of a finite number k such that any connected graph G (with at least two edges) has a neighbors distinguishing k-coloring of the edges. The proof relies on the Local Lemma and gives a bound of 183. On the other hand, the cycle C_6 shows that the minimum possible value of k is at least 3. The authors of [42] suspect that 3 colors suffice even for the following additive version of the problem.

Problem 12. *Every connected graph $G \neq K_2$ has an edge labelling f by integers* 1, 2, 3 *such that for every pair of adjacent vertices u and v*,

$$\sum_{x \in N(u)} f(ux) \neq \sum_{y \in N(v)} f(vy).$$

This holds for complete graphs and 3-colorable graphs (cf. [42]). Recently, Addario-Berry et al. [1] proved that 16 integer labels do the job in general.

5. Points

Thue's theorem is also a geometric result in a natural way. Therefore, we may ask for its continuous or multidimensional version. In this section we present a few Thue type problems with geometric flavor.

5.1. Coloring the real line

A coloring of the real line is *square-free* if no two adjacent intervals are colored the same. More precisely, for any intervals $I=[a,b]$ and $J=[b,c]$ of the same length $L>0$, there is a point $x \in I$ of different color than $x+L$. Bean et al. [9] proved that there is a square-free 2-coloring of the real line. The proof is a simple application of the principle of transfinite induction. Actually, the same argument gives a stronger result: there is a 2-coloring of the real line such that no two different segments (whether adjacent or not) are colored the same (cf. [39]). Rote (personal communication) found an explicit coloring with this striking property: a point x is black if $\ln|x|$ is rational, otherwise it is white. A validity of this simple coloring follows from the Lindemann–Weierstrass theorem, which says that algebraic powers of e are linearly independent over the field of algebraic numbers (cf. [39]). A simpler argument omitting this strong result has been found by Klazar (personal communication).

More mysterious is the following continuous version of strongly nonrepetitive sequences. A coloring f of $\mathbb{R}$ is *measurable* if for every color i the set $f^{-1}(i)$ is measurable in the sense of Lebesgue. For a measurable subset $X \subseteq \mathbb{R}$ denote its measure by $\lambda(X)$. We say that two intervals $I, J \subset \mathbb{R}$ are *distinguished* by a coloring f if

$$\lambda(I \cap f^{-1}(i)) \neq \lambda(J \cap f^{-1}(i)),$$

for at least one color i.

Problem 13. *Is there a finite measurable coloring of the real line distinguishing any pair of nontrivial intervals*?

Alon (personal communication) proved that no such 2-coloring of $\mathbb{R}$ is possible even if the condition was restricted to adjacent intervals. On the other hand it seems plausible that $\aleph_0$ colors suffice.

5.2. Topological disks

A *topological disk* in the plane is any homeomorphic image of the unit circle $U=\{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leqslant 1\}$. Suppose f is a coloring of the plane and D_1, D_2 are two different topological disks. We say that f *distinguishes* D_1 and D_2 if for every homeomorphism $h: D_1 \to D_2$ there is at least one point $x \in D_1$ such that $f(x) \neq f(h(x))$. It was proved in [37] that there exists a 2-coloring of $\mathbb{R}^2$ distinguishing each pair of topological disks. The proof uses transfinite induction.

Problem 14. *Find an explicit finite coloring of the plane distinguishing each pair of topological disks*.

5.3. Lattice points

Let $\mathbb{Z}^n$ be the set of integer lattice points in n-dimensional Euclidean space. A coloring of $\mathbb{Z}^n$ is *nonrepetitive* if a sequence of colors of consecutive points on any straight line segment is not a repetition. Let $\pi(\mathbb{Z}^n)$ denote the minimum number of colors in a nonrepetitive coloring of $\mathbb{Z}^n$. By Thue's theorem we have $\pi(\mathbb{Z})=3$.

Problem 15. *Determine* $\pi(\mathbb{Z}^2)$.

Notice that it is not obvious that $\pi(\mathbb{Z}^2)$ is finite. However, Carpi [17] proved that

$$3 \cdot 2^{n-1} \leqslant \pi(\mathbb{Z}^n) \leqslant 4^n$$

for every $n > 1$. Here is a brief sketch of the proof of the upper bound. Define a function $f : \mathbb{Z} \to \{1, 3, 5, 7\}$ by: (1) $f(i)$ is a (mod 4) residue of i, if i is odd, (2) $f(i) = h_i + 5$, where h_i is a (mod 4) residue of the highest odd factor of i, if $i \neq 0$ is even, and, (3) $f(0) = 5$. The values of f for the first 16 positive integers are

1, 5, 3, 5, 1, 7, 3, 5, 1, 5, 3, 7, 1, 7, 3, 5.

A key property of the function f is that any sequence of the form $f(a), f(a+d), \ldots, f(a+nd)$, spanned by an arithmetic progression of *odd* difference $d > 0$, is nonrepetitive (for arbitrary $a \in \mathbb{Z}$ and $n \in \mathbb{N}$). In consequence, an induced coloring of $\mathbb{Z}^n$ given by $f(x_1, \ldots, x_n) = (f(x_1), \ldots, f(x_n))$ is nonrepetitive.

As for graphs we may define the *Thue threshold* $t(n)$ of $\mathbb{Z}^n$ as the minimum number of colors such that any $(t(n) - 1)$-coloring of $\mathbb{Z}^n$ contains a k-power for each $k \in \mathbb{N}$. By Carpi's result $t(n)$ is finite for every n.

Problem 16. *Determine $t(2)$.*

Answering a question of Pach, Dumitrescu and Radoičić [29] proved that there is a 2-coloring of $\mathbb{Z}^2$ in which one can find at most 3 consecutive points on a line in the same color.

5.4. Unit distance graph

A famous problem relating graphs and geometry asks for the minimum number of colors in a coloring of the plane such that the ends of each unit segment have different colors. It is well known that the truth lies between 4 and 7. Let U be the *unit distance graph* of the plane, that is, the vertex set of U is the set of all points of the Euclidean plane, and the edge set of U consists of pairs of endpoints of all unit segments in the plane.

Problem 17. *Determine the Thue chromatic number of U.*

Unfortunately, $\pi(U)$ is not finite. Indeed, it can be proved that $\pi(G)$ is greater than the minimum degree of G, for any finite graph G (cf. [4]). It thus remains to decide whether $\pi(U) = \aleph_0$.

6. Numbers

The problems of this section are of different nature and are rather loosely related to Thue sequences. We include this short collection just for sentimental reasons.

6.1. Prime factors of consecutive integers

In 1857 Terquem and Prouhet (cf. [55]) made a conjecture that a product of any $k > 1$ consecutive integers cannot be a perfect power. The conjecture was confirmed first for squares by Erdős [30] and Rigge [62], and then proved in general by Erdős and Selfridge [33]. The proof relied on detailed analysis of *square-free* parts of consecutive integers.

Let $p(n)$ denote the product of all distinct primes dividing $n \geqslant 2$, and let S be an infinite sequence formed of these products. The first 20 terms of S are

2, 3, 2, 5, 6, 7, 2, 3, 10, 11, 6, 13, 14, 15, 2, 17, 6, 19, 10.

The following problem is due to Erdős and Woods (cf. [40]). It is also known to be equivalent to Robinson's problem on definability of Arithmetic in terms of operations of addition and greatest common divisor.

Problem 18. *There exists a constant k such that no two different blocks of S of length at least k are the same.*

By Bertrand's Postulate it is easy to see that S is nonrepetitive. More general version of Erdős–Woods conjecture follows from the celebrated *abc*-conjecture (cf. [46]).

6.2. Continued fractions

Let α be an algebraic number of degree deg α. It is well known that the sequence $P(\alpha)$ of partial quotients of the continued fraction expansion of α is eventually periodic if and only if deg $\alpha = 2$. In 1936 Khintchine conjectured that if deg $\alpha > 2$ then $P(\alpha)$ is actually unbounded (cf. [66]).

Surprisingly, there is an equivalent formulation of Khintchine's conjecture using Thue properties of certain binary sequences. Suppose $\alpha \in (0, 1)$ and consider the associated *billiard sequence* $B(\alpha)=(b_i)_{i\in\mathbb{N}}$, defined by $b_i = \lfloor (i+1)\alpha \rfloor - \lfloor i\alpha \rfloor$ for $i \geqslant 1$. For instance, if $\alpha = (-1+\sqrt{5})/2$ then $B(\alpha)$ is the famous *Fibonacci string*

$$F = 1011010110110\ldots .$$

The sequence F can be obtained, similarly as the Thue sequence $T = 01101001\ldots$, by iterating the substitution $1 \to 10$ and $0 \to 1$. Moreover, F is known to be 4-power-free (cf. [48]).

Sequences $B(\alpha)$ and $P(\alpha)$ are combinatorially related (cf. [3,48,64]). By this relation it is seen immediately that unboudedness of $P(\alpha)$ implies that $B(\alpha)$ contains k-powers for any k. What is more surprising is that the converse also holds, as proved by Mignosi [50] (cf. [48]). Thus, Khintchine's conjecture reduces to the following statement.

Problem 19. *If α is an algebraic number of degree* deg $\alpha \geqslant 3$ *then the billiard sequence $B(\alpha)$ contains k-powers for every positive integer k.*

6.3. Primes and sums of digits

Let $q \geqslant 2$ be an integer and let $s_q(n)$ denote the sum of digits of base q expansion of n. Let A_r be the set of all integers n such that $s_q(n) \equiv r \pmod{q}$, $r = 0, 1, \ldots, q-1$. In analogy to the famous Dirichlet's theorem on primes in arithmetic progressions we may pose the following.

Problem 20. *Let p_i be the ith prime. Then for every fixed $q \geqslant 2$ and $0 \leqslant r < q$,*

$$\lim_{n\to\infty} \frac{|\{i \leqslant n : p_i \in A_r\}|}{n} = \frac{1}{q}.$$

The conjecture was stated in 1968 by Gelfond [35]. It is not known whether any particular set A_r contains infinitude of primes (cf. [49]). Notice that the case $q = 2$ corresponds to Prouhet partitions defined in Section 2.

Acknowledgments

I would like to thank Noga Alon, Adrian Bondy, Boštjan Brešar, Arturo Carpi, James Currie, Michael Dekking, Lorenz Halbeisen, Michał Karoński, Sandi Klavžar, Martin Klazar, André Kündgen, Tomasz Łuczak, Michael Pelsmajer, Narad Rampersad, Oliver Riordan, Günter Rote, Jeff Shallit, Jamie Simpson, Vera Sós and Carsten Thomassen, for valuable comments, suggestions and materials. Also I would like to acknowledge a support from Grant KBN 1P03A 017 27.

References

[1] L. Addario-Berry, K. Dalal, B.A. Reed, Degree constrained subgraphs, Discrete Appl. Math. (to appear).
[2] J.-P. Allouche, J. Shallit, The ubiquitous Prouhet-Thue-Morse Sequence. Sequences and their Applications (Singapore, 1998), Springer Series in Discrete Mathematics and Theoretical Computer Science, Springer, London, 1999, pp. 1–16.
[3] J.-P. Allouche, J. Shallit, Automatic Sequences. Theory, Applications, Generalizations, Cambridge University Press, Cambridge, 2003.
[4] N. Alon, J. Grytczuk, Power-free colorings of graphs, manuscript.
[5] N. Alon, J. Grytczuk, M. Hałuszczak, O. Riordan, Non-repetitive colorings of graphs, Random Struct. Algor. 21 (2002) 336–346.
[6] N. Alon, J.H. Spencer, The Probabilistic Method, second ed., Wiley, New York, 2000.
[7] E.S. Arszon, Proof of the existence of n-valued infinite asymmetric sequences, Mat. Sb. 2 (44) (1939) 769–779 (in Russian).
[8] K.A. Baker, G.F. McNulty, W. Taylor, Growth problems for avoidable words, Theoret. Comput. Sci. 69 (1989) 319–345.
[9] D.R. Bean, A. Ehrenfeucht, G.F. McNulty, Avoidable patterns in strings of symbols, Pacific J. Math. 85 (1979) 261–294.
[10] J. Beck, An application of Lovász local lemma: there exists an infinite 01-sequence containing no near identical intervals, in: Finite and Infinite Sets, Eger 1981, Colloquia Mathematica Societatis János Bolyai, 1983, pp. 103–107.

[11] J. Berstel, Axel Thue's work on repetitions in words, in: P. Leroux, C. Reutenauer (Eds.), Séries formelles et combinatoire algébrique Publications du LaCIM, Université du Québec a Montréal, 1992, pp. 65–80.
[12] J. Berstel, Axel Thue's papers on repetitions in words: a translation, Publications du LaCIM, vol. 20, Université du Québec a Montréal, 1995.
[13] J. Berstel, J. Karhumäki, Combinatorics on words—a tutorial, Bull. Eur. Assoc. Theoret. Comput. Sci. EATCS 79 (2003) 178–228.
[14] R. Bott, Marston Morse and his mathematical works, Bull. Amer. Math. Soc. (New Series) 3 (1980) 907–950.
[15] B. Brešar, S. Klavžar, Square-free colorings of graphs, Ars Combin. 70 (2004) 3–13.
[16] T.C. Brown, Is there a sequence on four symbols in which no two adjacent segments are permutations of one another?, Amer. Math. Monthly 78 (1971) 886–888.
[17] A. Carpi, Multidimensional unrepetitive configurations, Theoret. Comput. Sci. 56 (1988) 233–241.
[18] C. Choffrut, J. Karhumäki, Combinatorics of words, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, Springer, Berlin, Heidelberg, 1997, pp. 329–438.
[19] R.J. Clark, Avoidable formulas in combinatorics on words, Ph.D. Thesis, University of California, Los Angeles, 2001.
[20] J.D. Currie, Non-repetitive walks in graphs and digraphs, Ph.D. Thesis, University of Calgary, Alberta, Canada, 1987.
[21] J.D. Currie, Which graphs allow infinite nonrepetitive walks?, Discrete Math. 87 (3) (1991) 249–260.
[22] J.D. Currie, Open problems in pattern avoidance, Amer. Math. Monthly 100 (1993) 790–793.
[23] J.D. Currie, There are ternary circular square-free words of length n for $n \geqslant 18$, Electron. J. Combin. 9 (10) (2002) 7.
[24] J.D. Currie, D.S. Fitzpatrick, Circular words avoiding patterns, in: Developments in Language Theory, Lecture Notes in Computer Science, vol. 2450, Springer, Berlin, 2003, pp. 319–325.
[25] J.D. Currie, E. Moodie, A word on 7 letters which is non-repetitive up to mod5, Acta Inform. 39 (6–7) (2003) 451–468.
[26] J.D. Currie, J. Simpson, Non-repetitive tilings, Electron. J. Combin. 9 (2002).
[27] F. Dejean, Sur un théoreme de Thue, J. Combin. Theory Ser. A 13 (1972) 90–99.
[28] F.M. Dekking, Strongly non-repetitive sequences and progression free sets, J. Combin. Theory Ser. A 27 (1979) 181–185.
[29] A. Dumitrescu, R. Radoicic, On a coloring problem for the integer grid, in: Towards a Theory of Geometric Graphs, vol. 342, Contemporary Mathematics, AMS, Providence, RI, 2004, pp. 67–74.
[30] P. Erdős, Note on products of consecutive integers, J. London Math. Soc. 14 (1939) 194–198.
[31] P. Erdős, Some unsolved problems, Magyar Tud. Akad. Mat. Kutato. Int. Kozl. 6 (1961) 221–254.
[32] P. Erdős, J. Nešetřil, Problem, in: G. Halász, V.T. Sós (Eds.), Irregularities of Partitions (Papers from the meeting held in Fertöd, July 7–11, 1986), Algorithms and Combinatorics: Study and Research Texts, vol. 8, Springer, Berlin, 1989, pp. 162–163.
[33] P. Erdős, J.L. Selfridge, The product of consecutive integers is never a power, Illinois J. Math. 19 (1975) 292–301.
[34] A.A. Evdokimov, Strongly asymmetric sequences generated by finite number of symbols, Dokl. Akad. Nauk. SSSR 179 (1968) 1268–1271 Soviet Math. Dokl. 9 (1968) 536–539.
[35] A.O. Gelfond, Sur les nombres qui ont des propriétés additives et multiplicatives données, Acta Arith. 13 (1967/1968) 259–265.
[36] W.H. Gottschalk, G.A. Hedlund, A characterization of the Morse minimal set, Proc. Amer. Math. Soc. 15 (1964) 70–74.
[37] J. Grytczuk, Pattern avoiding colorings of Euclidean spaces, Ars Combin. 64 (2002) 109–116.
[38] J. Grytczuk, Thue-like sequences and rainbow arithmetic progressions, Electron. J. Combin. 9 (2002).
[39] J. Grytczuk, W. Śliwa, Non-repetitive colorings of infinite sets, Discrete Math. 265 (2003) 365–373.
[40] R.K. Guy, Unsolved Problems in Number Theory, second ed., Springer, New York, 1994.
[41] L. Halbeisen, N. Hungerbühler, An application of van der Waerden's theorem in additive number theory, Integers: Electr. J. Combin. Numer. Theory 0 (2000).
[42] M. Karoński, T. Łuczak, A. Thomason, Edge weights and vertex colours, J. Combin. Theory Ser. B 91 (2004) 151–157.
[43] V. Keränen, Abelian squares are avoidable on 4 letters, Automata, Languages and Programming: Lecture Notes in Computer science, vol. 623, Springer, Berlin, 1992, pp. 41–52.
[44] A.I. Kostrikin, Around Burnside, "Nauka", Moscow, 1986.
[45] A. Kündgen, M.J. Pelsmajer, Nonrepetitive colorings of graphs of bounded tree-width, Discrete Math., in press, doi:10.1016/j.disc.2007.08.043.
[46] M. Langevin, Sur quelques conséquences de la conjecture *abc* en arithmétique et en logique. Symposium on Diophantine Problems (Boulder, CO, 1994), Rocky Mountain J. Math. 26 (3) (1996) 1031–1042.
[47] M. Lothaire, Combinatorics on Words, Addison-Wesley, Reading, MA, 1983.
[48] M. Lothaire, Algebraic Combinatoric on Words, Cambridge, 2001.
[49] C. Mauduit, Multiplicative properties of the Thue-Morse sequence, Period. Math. Hungar. 43 (2001) 137–153.
[50] F. Mignosi, Infinite words with linear subword complexity, Theoret. Comput. Sci. 82 (1989) 221–242.
[51] M. Molloy, B. Reed, A bound on the strong chromatic index of a graph, J. Combin. Theory Ser. B 69 (2) (1997) 103–109.
[52] M. Molloy, B. Reed, Graph Colouring and the Probabilistic Method, Algorithms and Combinatorics, vol. 23, Springer, Berlin, 2002.
[53] M. Morse, A one-to-one representation of geodesics on a surface of negative curvature, Amer. J. Math. 43 (1921) 35–51.
[54] M. Morse, G. Hedlund, Unending chess, symbolic dynamics, and a problem in semigroups, Duke Math. J. 11 (1944) 1–7.
[55] W. Narkiewicz, Classical Problems in Number Theory, Monografie Matematyczne, PWN Warszawa, 1986.
[56] P.S. Novikov, On periodic groups, Dokl. Akad. Nauk SSSR 127 (1959) 749–752.
[57] P.S. Novikov, S.I. Adjan, On infinite periodic groups I, II, III, Math. USSR Izv. 32 (1968) 212–244 251–524, 709–731.
[58] M.J. Ollagnier, Proof of Dejean's conjecture for alphabets with 5, 6, 7, 8, 9, 10 and 11 letters, Theoret. Comput. Sci. 95 (1992) 187–205.
[59] G. Pirillo, S. Varricchio, On uniformly repetitive semigroups, Semigroup Forum 49 (1) (1994) 125–129.
[60] P.A.B. Pleasants, Non-repetitive sequences, Proc. Cambridge Philos. Soc. 68 (1970) 267–274.
[61] E. Prouhet, Memoire sur quelques relations entre les puissances des nombres, C. R. Acad. Sci. Paris 33 (1851) 31.
[62] O. Rigge, Über ein diophantisches Problem, in: IX Scandinavian Congress of Mathematicians, 1939, pp. 155–160.

[63] N. Robertson, P.D. Seymour, Graph minors V: excluding a planar graph, J. Combin. Theory Ser. B 41 (1986) 92–114.
[64] J. Rosenblatt, The sequence of greatest integers of an arithmetic progression, J. London Math. Soc. (2) 17 (1978) 213–218.
[65] M.V. Sapir, Combinatorics on Words with Applications, Birkhäuser (to appear).
[66] J. Shallit, Real numbers with bounded partial quotients: a survey, Enseign. Math. (2) 38 (1–2) (1992) 151–187.
[67] A. Thue, Über unendliche Zeichenreichen, Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiania 7 (1906) 1–22.
[68] A. Thue, Über die gegenseitigen Lage gleicher Teile gewisser Zeichenreihen, Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiania 1 (1912) 1–67.
[69] A.I. Zimin, Blocking sets of terms, Mat. Sb. (N.S.) 119 (1982) 363–375.