Pattern avoiding colorings of Euclidean spaces

Jarosław Grytczuk

Dedicated to the memory of Wacław Sierpiński

ABSTRACT. Given a coloring f of Euclidean space $\mathbb{R}^n$ and some group of its transformations, its subsets A and B are said to be colored *similarly*, if there exists $g \in G$, such that $B = g(A)$ and $f(a) = f(g(a))$, for all $a \in A$. From our earlier result [12] it follows that there are 2-colorings of $\mathbb{R}^n$, in which no two different line segments are colored similarly with respect to isometries. The main purpose of this paper is to investigate other types of such *pattern avoiding* colorings. In particular, we consider topological as well as measure theoretic aspects of the above scene. Our motivation for studying this topic is twofold. One is that it extends square-free colorings of $\mathbb{R}$, introduced in [2] as a continuous version of the famous *non-repetitive sequences* of Thue. The other is its relationship to some exciting problems and results of Euclidean Ramsey Theory, especially those concerning avoiding distances.

1. Introduction

In this paper we study colorings of Euclidean spaces avoiding some specified regularities. Perhaps the most famous problem of this type concerns unit distances. A k-coloring $f : \mathbb{R}^n \rightarrow \{0, 1, ..., k-1\}$ *avoids* unit distances if $f(x) \neq f(y)$ for any two points x, y distance one apart. The problem is to determine the minimal number of colors needed for such a coloring. Sometimes this number is called *the chromatic number of* $\mathbb{R}^n$ and is denoted by $\chi(\mathbb{R}^n)$. It is easy to see that $\chi(\mathbb{R}) = 2$; simply, color the intervals of the form $[2m, 2m+1)$, $m \in \mathbb{Z}$, red, and the rest of the line blue. Surprisingly, for higher dimensions the problem is much harder and is still

---

1991 *Mathematics Subject Classification.* Primary 68R15, 05C55; Secondary 04A20.

Key words and phrases. non-repetitive sequences, pattern avoiding colorings.

open, even for the plane. By simple constructions, one can get quickly the inequalities $4 \leq \chi(\mathbb{R}^2) \leq 7$, but finding the exact value is certainly one of the most challenging problems of discrete geometry (see [4]).

Of course, there are many possible variations on the above theme. For instance, suppose that we want to avoid more different distances with as few colors as possible. A beautiful theorem of Erdős et al. [10] states that there exists a set S of cardinality 2^ω and a 2-coloring of $\mathbb{R}$ such that $f(x) \neq f(y)$, if the distance between x and y belongs to S. On the other hand , such an S can not have positive measure.

Much stronger "avoidability" can be achieved having more colors at a disposal. In 1943, Erdős and Kakutani [11] proved, assuming a validity of Continuum Hypothesis, that there exists an ω-coloring of $\mathbb{R}$ in which *any* given distance appears *at most once*, that is, if $f(x) = f(y)$, $x \neq y$, then $f(x+t) \neq f(y+t)$, for every $t \neq 0$. Actually, they proved that this statement is equivalent to Continuum Hypothesis. Their result was then generalized for the plane by Davies [7], and finally, for all finite dimensional spaces, by Kunen [13].

The next theorem seems to fit nicely into the above collection, although our primer motivation was of somewhat different nature.

THEOREM 1. *(Grytczuk and Śliwa [12]) There exists a 2-coloring f of the real line, such that for any two points x, y and any $\varepsilon > 0$ there exists $0 \leq t < \varepsilon$ such that $f(x+t) \neq f(y+t)$.*

It was found as a generalization of the result of Bean, Ehrenfeucht and Mc Nulty [2] concerning square-free colorings of $\mathbb{R}$. We say that two intervals I and its translation $t + I$, $t \neq 0$, are colored *in the same way*, if for every $x \in I$, colors of x and $x + t$ agree. A coloring f of $\mathbb{R}$ is called *square-free* if no two *adjacent* intervals (i.e. such that $t = |I|$) are colored in the same way. This notion was introduced in [2] as a continuous version of the famous *non-repetitive sequences* discovered in 1906 by Thue [18] (see section 3). It was proved there that there are square-free 2-colorings of $\mathbb{R}$. Clearly, Theorem 1 extends this result by dropping the assumption of adjacency.

The proof of Theorem 1 is very simple by the use of transfinite induction. However, we would like to present here two other proofs, one of which provides an ingenious example of an explicit coloring function found by Rote [17] (see section 2). The purpose of this paper is also to consider other kinds of situations in which similarly colored objects do not appear. Many directions and generalizations are possible (see section 3). Suppose, for example, that instead of translations and intervals on the line we are interested rather in *homeomorphisms* and *topological disks* on the plane. Now, two disks A and B are colored *similarly,* if there exists a homeomorphism h transforming A onto B, such that every point $x \in A$ has the same color as $h(x)$. From Theorem 3 of section 3, it follows easily that there exists a 2-coloring of $\mathbb{R}^2$ in which no two disks are colored similarly.

New problems arise if we impose some restrictions on the coloring. The most natural ones are those of set theoretic or measure theoretic nature. For instance, a coloring of $\mathbb{R}^n$ is measurable if every color class is a measurable subset of $\mathbb{R}^n$. The coloring function of Rote (see section 2) satisfies this condition, but one of the color classes is countable, hence of measure zero. In consequence, two intervals differ only in points forming a null set. So, one tempts to ask whether there is possible a measurable k-coloring of $\mathbb{R}$ such that for any interval I the set $\{x \in I;\ f(x) \neq f(x+t)\}$ has positive measure, for all $t \neq 0$. A concluding discussion of this and other open problems is contained in the last section of the paper.

2. Three proofs of Theorem 1

2.1. Proof 1 (by transfinite induction). Let A be the set consisting of all triples $\{x, y, \varepsilon\}$, where $x < y$ and $\varepsilon > 0$. Clearly, A is of cardinality 2^ω, and, by the Well Ordering Theorem, can be enumerated by ordinals less then 2^ω. Say, $A = \{a_\alpha;\ \alpha < 2^\omega\}$. We will proceed inductively, coloring at each step *at most two* different points. In the beginning step we can simply color the points of the first triple differently. In further steps each time we have to pick such a number $t \in [0, \varepsilon)$ that neither a point $t + x$ nor $t + y$ have been already colored, and then color them differently. To see that it is always possible, suppose that all triples a_β, $\beta < \alpha$ are done, for some $\alpha < 2^\omega$, and we have to deal now with the element $a_\alpha = \{x, y, \varepsilon\}_\alpha$. The number of points that have been colored by this procedure so far is at most equal to 2α, which is strictly less then 2^ω. Hence, there are still free points of the form $t + x$ and $t + y$, with $t \in [0, \varepsilon)$, and we are able to do our job again. We obtain in this way a partial coloring of $\mathbb{R}$ possessing the desired property. Of course, it can be extended to the whole line arbitrarily, so, the proof is complete. ■

The above proof is highly non-constructive as it relies on the Well Ordering Theorem, which, in turn, is equivalent to the Axiom of Choice. Therefore, we would like to present another proof of a bit more algebraic nature, in which transfinite induction is omitted. Unfortunately, the Axiom of Choice is still present, because our construction uses representatives of cosets of the subgroup $\mathbb{Q}$ of rational numbers in the additive group $\mathbb{R}$.

2.2. Proof 2 (by cosets). First, we shall color only rational points such that the assertion of the theorem restricted to the set $\mathbb{Q}$ will be satisfied. Note that now the translation summand t must be rational. It can be done the same way as before, by considering the countable set of rational triples $\{x, y, \varepsilon\}$ and applying the usual induction. Hence, assume that $f : \mathbb{Q} \rightarrow \{0, 1\}$ is a function that do the job, and consider an arbitrary coset H of $\mathbb{Q}$, in the additive group $\mathbb{R}$. Let $r > 0$ be a fixed representative of H. We shall copy the coloring of $\mathbb{Q}$ on H by translation. So, if $h \in H$ and $h = r + q$, where $q \in \mathbb{Q}$, then we set $f(h) = f(q)$. Since cosets form a partition of $\mathbb{R}$, this definition extends f to the whole line.

We will show now that f possesses the property postulated in the theorem. Let there be given real numbers $x < y$ and $\varepsilon > 0$. Then $x = r + q_1$ and $y = s + q_2$, $q_1, q_2 \in \mathbb{Q}$, where $r, s \geq 0$ are fixed representatives of cosets containing x and y, respectively. If $q_1 \neq q_2$, then it suffices to choose a rational $0 \leq t < \varepsilon$, such that colors of $q_1 + t$ and $q_2 + t$ are different. Indeed, in this case we have $x + t = r + (q_1 + t)$, $y + t = s + (q_2 + t)$, with r and s still being the representatives of cosets containing $x + t$ and $y + t$, respectively. Hence, $f(x + t) \neq f(y + t)$.

If $q_1 = q_2$, then first we have to shift x and y by some small number $0 \leq \alpha < \varepsilon$, so as to place $y + \alpha$ into the set $\mathbb{Q}$. Then certainly $x + \alpha \notin \mathbb{Q}$ and we get the same situation as in the previous case. Namely, $x + \alpha = r' + q_3$, $r' > 0$ and $y + \alpha = 0 + q_4$, with $q_3 \neq q_4$, by the initial assumption that $x < y$. Thus, the proof is complete. ■

Finally, we would like to present third proof of Theorem 1, discovered by Günter Rote [17]. This elegant argument not only omits induction, but also provides a simple example of an explicit coloring function. However, on the other hand, it is based on a powerful number theoretic result - the Lindemann-Weierstrass theorem (see [1]). It goes as follows.

2.3. Proof 3 (by guessing an explicit function). Consider a function $f : \mathbb{R} \to \{0, 1\}$ assigning 0 to a real number x, if $\ln|x| \in \mathbb{Q}$, and 1 in all other cases. As we shall soon see, the signs of considered numbers do not matter, so, assume, only for convenience, that $0 < x < y$. If colors of x and y agree, shift them slightly to get $x + t_1 = e^{q_1}$, where $0 \leq t_1 < \varepsilon$ and $q_1 \in \mathbb{Q}$. Then $f(x + t_1) = 0$, by the definition of f. If at the same time $f(y + t_1) = 0$, then there must exist a rational number $q_2 \neq q_1$, such that $y + t_1 = e^{q_2}$. In such a bad situation we have to make another shift by some $t_1 < t_2 < \varepsilon$ to get $x + t_2 = e^{q_3}$, for some $q_3 \in \mathbb{Q}$, different from q_1 and q_2. If it would again happened that $y + t_2 = e^{q_4}$ is a rational power of e, then we will get the equality $e^{q_1} - e^{q_2} = e^{q_3} - e^{q_4}$, with all q_i *different* rational numbers. However, this contradicts the well-known result of algebraic number theory, which we formulate below as a lemma.

LEMMA 1. *(Lindemann-Weierstrass theorem) If $a_1, ..., a_n$ are non-zero algebraic numbers and $b_1, ..., b_n$ are pairwise distinct algebraic numbers, then $a_1 e^{b_1} + ... + a_n e^{b_n} \neq 0$.*

This finishes the proof. ■

Let us remark here that from the above lemma it follows, in particular, that the function e^x takes only transcendental values at rational points

$x \neq 0$. Hence, the example of Rote does not work if we restrict ourselves to the set $\mathbb{Q}$, and we could ask for explicit construction in that case, too. One such, rather complicated example, was found by Wiesław Śliwa [12], but it would be nice to have something simpler.

3. Some generalizations

Theorem 1 can be generalized in many diverse directions. For instance, we can consider any finite or even countable configuration A of points on the line, and look for copies of A with prescribed color patterns, lying arbitrarily close to A. We can also allow more colors and expect multicolored copies of A. The following striking fact covering all these cases can be deduced by the use of transfinite induction.

THEOREM 2. *For every cardinal $k \leq \omega$ there exists a k-coloring f of the real line, such that given any countable configuration $A \subset \mathbb{R}$, any function $p : A \to k$ (a color pattern), and any $\varepsilon > 0$, there exists $0 \leq t < \varepsilon$, such that $f(a+t) = p(a)$, for all $a \in A$.*

PROOF. The idea is the same as in the first proof of Theorem 1. It suffices to note, that there are only 2^ω all possible triples $\{A, p, \varepsilon\}$, and that at most $\alpha \times \omega < 2^\omega$ points have been colored before the step $\alpha < 2^\omega$. ∎

It is worth mentioning now, that this argument can be applied in much more general situations. It works, for instance, in all finite dimensional Euclidean spaces. Actually, Theorems 1 and 2 can be regarded as special cases of a purely set theoretical result that can be found in [12]. The set theoretic approach is also fruitful when we would like to consider other types of transformations in Euclidean spaces (or even other types of "spaces"). In fact, isometry is not the only way of mapping one line segment onto the other. So, let us make now a general setting, which will be convenient for our further discussion. Suppose we consider some space S with a group G of transformations acting on S. Let f denote a k-coloring of S. We say that two subsets $A, B \subseteq S$ are colored *similarly* with respect to G, if there is some transformation $g \in G$ mapping A onto B and preserving the coloring f, i.e. $f(a) = f(g(a))$, for all a in A. If S is the real line and G is the group of translations, then Theorem 1 establishes the possibility of coloring points of the line red and blue, with no two distinct segments colored similarly. This concerns, of course, only segments of equal positive length. If we would like to consider segments of different lengths, we must allow other kinds of mappings. It is somewhat surprising that, even if we take the group of all homeomorphisms as G, then an analogous result holds for topological line segments (arcs) in any finite dimensional Euclidean space.

THEOREM 3. *There exists a 2-coloring of $\mathbb{R}^n$ such that no two different arcs are colored similarly.*

PROOF. Again, the essential point is to apply transfinite induction argument. As before, we consider the set of triples $\{A, B, h\}$, where A and B are different arcs, and h is a homeomorphism mapping A onto B. The space $\mathbb{R}^n$ is separable, in consequence, this set has cardinality of continuum. Thus, we can make the same conclusions as before noting that, by continuity of h, there are also 2^ω non-fixed points in A. ∎

This time we have no hope for any explicit coloring. So, to make things more concrete, let us concentrate now on translations over the set of integers $\mathbb{Z}$. It is obvious that an analog of Theorem 1 can not hold here. However, restricting attention to consecutive intervals leads to very interesting problems. A coloring of $\mathbb{Z}$ is called *square-free* (or, more generally, *n-th power-free*), if no two (no n) consecutive intervals $A = [a, a+m)$ and $B = [a+m, a+2m)$, $a, m \in \mathbb{Z}$, $m > 0$, are similarly colored. The following unexpected results were discovered by Axel Thue in 1906, [18].

THEOREM 4. *(Thue [18]) There exist a square-free 3-coloring and a cube-free 2-coloring of the set of integers $\mathbb{Z}$.*

The method invented by Thue relies on iterating substitutions on words. For example, iterating the substitution $0 \to 01$, $1 \to 10$ gives an infinite sequence of words 0, 01, 0110, 01101001, ... defining in a natural way the infinite word $T = 0110100110010110...$, in which no finite block B appears three times consecutively. In other words, no fragment of T looks like BBB, hence the name - a *cube-free* word. Actually, the above sequence was discovered earlier by Prouhet [16] in connection to a number theoretic problem concerning sums of powers, but Thue was the first who recognized its avoidability properties. The sequence T was rediscovered later many times in different contexts, especially intriguing being its appearance in symbolic dynamics. For this and much other relevant information, one can consult survey articles in [15] and [3].

An infinite square-free word over three symbols can be generated by substitution $a \to abc$, $b \to ac$, $c \to b$, which was found by Dekking, who considered also a stronger avoidance property [8]. An infinite word is *strongly square-free*, if no two of its consecutive blocks are permutations of each other. In other words, for any two adjacent blocks the number of appearances of some letter in these blocks must be different. It is checked easily that the minimal possible number of symbols for an infinite word with this property is 4, and it was Erdős, who first asked the question of the existence of such a sequence [9]. This was not answered in the affirmative until 1992 by Keränen, who found the appropriate (rather large) substitution. Although the literature concerning avoidable properties of words is rather huge, there are still many exciting unsolved problems (see [2], [3], [5], [6],

[15]). In the next section we would like to add one more to this collection, together with some continuous analogs of the above.

4. Open problems

Let us return once more to Theorem 1. Given a coloring f of $\mathbb{R}$ denote, for any $t > 0$, the set $D(t) = \{x \in \mathbb{R};\ f(x) \neq f(x+t)\}$, and, by $D(t, I)$, its intersection with an interval I. In this terms it can be formulated as follows: *there exists a 2-coloring of $\mathbb{R}$ such that the set $D(t)$ is dense for every $t > 0$*. The size of the set $D(t, I)$ expresses how much the intervals I and its translate $t+I$ differ. So, we may try to impose stronger conditions on its cardinality or structure to obtain colorings of higher degrees of irregularity. Our first question is the following.

PROBLEM 1. *Is it possible for the set $D(t)$ to be everywhere of second category, for some 2-coloring of $\mathbb{R}$, and all $t > 0$?*

Another natural problem appears if we want the set $D(t, I)$ to be large in the measure theoretic sense. Here we have to consider only measurable colorings. This problem can be seen as a continuous analog of strong avoidance for the integers.

PROBLEM 2. *Can all of the sets $D(t, I)$ be of positive measure, for some finite (countable) measurable coloring of $\mathbb{R}$?*

Finally, we would like to go the opposite direction and state our last problem that arose as a discrete analog of Theorem 2. If true, it would be a far reaching extension of the theorem of Thue.

PROBLEM 3. *Do there exist, for every integer $k \geq 2$, a k-coloring of the set of positive integers such that, given any k term arithmetic progression A with positive difference d, there is an integer $0 \leq t < d$, such that $t + A$ is multicolored ?*

References

[1] A. Baker, Transcendental number theory, Cambridge Univ. Press, 1975.

[2] D.R. Bean, A. Ehrenfeucht, G.F. McNulty, Avoidable patterns in strings of symbols, Pacific J. Math. **85** (1979), 261-294.

[3] Ch. Choffrut, J. Karhumäki, Combinatorics of Words, in Handbook of Formal Languages (G. Rozenberg, A. Salomaa eds.) Springer-Verlag, Berlin Heidelberg, 1997, 329-438.

[4] H.T. Croft, K.J. Falconer, R.K. Guy, Unsolved problems in geometry, Springer-Verlag, New York, 1991.

[5] J.D. Currie, Open problems in pattern avoidance, Amer. Math. Monthly **100** (1993), 790-793.

[6] J.D. Currie, Words avoiding patterns: Open problems, manuscript.

[7] R.O. Davies, Partitioning the plane into denumerably many sets without repeated distances, Proc. Camb. Phil. Soc. **72** (1972), 179-183.

[8] F.M. Dekking, Strongly non-repetitive sequences and progression free sets, J. Combin. Theory Ser. A **16** (1974), 159-164.
[9] P. Erdös, Some unsolved problems, Magyar Tud. Akad. Mat. Kutato. Int. Kozl. **6** (1961), 221-254.
[10] P. Erdös, R.L. Graham, P. Montgomery, B.L. Rothschild, J. Spencer, E.G. Straus, Euclidean Ramsey theorems II, in: Infinite and Finite Sets, Keszthely (Hungary), 1973, Coll. Math. Soc. J. Bolyai **10**, 529-557.
[11] P. Erdös, S. Kakutani, On non-denumerable graphs, Bull. Amer. Math. Soc., **49** (1943), 457-461.
[12] J. Grytczuk, W. Śliwa, Pattern-free partitions of Euclidean spaces, manuscript.
[13] K. Kunen, Partitioning Euclidean space, Proc Camb. Phil. Soc. **102** (1987), 379-383.
[14] V. Keränen, Abelian squares are avoidable on 4 letters, Automata, Languages and Programming: Lecture notes in Computer science **623** (1992) Springer-Verlag 4152.
[15] M. Lothaire, Combinatorics on Words, Addison-Wesley, Reading MA, 1983.
[16] E. Prouhet, Mémoire sur quelques relations entre les puissances des nombres, Inst. France Acad. d. Sci., Comptes Rendus **33** (1851), 225.
[17] G. Rote, personal comunication.
[18] A. Thue, Über unendliche Zeichenreichen, Norske Vid Selsk. Skr. I. Mat. Nat. Kl. Christiana, **7** (1906), 1-22.

INSTITUTE OF MATHEMATICS, TECHNICAL UNIVERSITY OF ZIELONA GÓRA, 65-246 ZIELONA GÓRA, POLAND
E-mail address: jarek@axel.im.pz.zgora.pl