PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 137, Number 5, May 2009, Pages 1593–1599
S 0002-9939(08)09699-8
Article electronically published on November 25, 2008

SPLITTING NECKLACES AND MEASURABLE COLORINGS OF THE REAL LINE

NOGA ALON, JAROSŁAW GRYTCZUK, MICHAŁ LASOŃ, AND MATEUSZ MICHAŁEK

(Communicated by Jim Haglund)

Abstract. A (continuous) *necklace* is simply an interval of the real line colored measurably with some number of colors. A well-known application of the Borsuk-Ulam theorem asserts that every k-colored necklace can be fairly split by at most k cuts (from the resulting pieces one can form two collections, each capturing the same measure of every color). Here we prove that for every $k \geq 1$ there is a measurable $(k+3)$-coloring of the real line such that no interval can be fairly split using at most k cuts. In particular, there is a measurable 4-coloring of the real line in which no two adjacent intervals have the same measure of every color. An analogous problem for the integers was posed by Erdős in 1961 and solved in the affirmative by Keränen in 1991. Curiously, in the discrete case the desired coloring also uses four colors.

1. Introduction

In 1906 Thue [23] proved that there is a 3-coloring of the integers such that no two adjacent intervals are colored exactly the same. This result has had lots of unexpected applications in distinct areas of mathematics and theoretical computer science (see [1], [6], [8], [19]). Many variations and generalizations of this property have been considered, specifically in other combinatorial settings such as Euclidean spaces [6], [14], [15] and graph colorings [3], [4], [16].

In particular, in 1961 Erdős [11] (cf. [9], [10], [17]) asked whether there is a 4-coloring of the integers such that no two adjacent segments are identical, even after arbitrary permutation of their terms. (In other words, there is always a color whose number of occurrences in one segment differs from those in other segments.) It is not hard to check by hand that four colors are needed for this property, but on the other hand, it is not obvious that any finite number of colors is enough. This fact was first established by Evdokimov [12], who found a 25-coloring with the desired property. Another construction, provided by Pleasants [22], improved the bound to 5. That 4 colors actually suffice was finally proved by Keränen [18], with some verifications made by computer.

In the present paper we study a continuous variant of the Erdős problem. In particular, we prove that there exists a measurable 4-coloring of the real line such that no two adjacent segments contain equal measure of every color. Actually our result is more general and relates to the continuous version of the *necklace*

Received by the editors March 25, 2008, and, in revised form, July 4, 2008.
2000 *Mathematics Subject Classification.* Primary 05C38, 15A15; Secondary 05A15, 15A18.
Key words and phrases. Measurable coloring, splitting necklaces.

splitting problem. Let $f : \mathbb{R} \to \{1, 2, \ldots, k\}$ be a k-coloring of the real line such that for every $i \in \{1, 2, \ldots, k\}$, the set $f^{-1}(i)$ of all points in color i is Lebesgue measurable. A *splitting* of size r of an interval $[a, b]$ is a sequence of points $a = y_0 < y_1 < \ldots < y_r < y_{r+1} = b$ such that it is possible to partition the resulting collection of intervals $F = \{[y_i, y_{i+1}] : 0 \leq i \leq r\}$ into two disjoint subcollections F_1 and F_2, each capturing exactly half of the total measure of every color. The partition $F = F_1 \cup F_2$ will be called a *fair* partition of F. For instance, in the continuous analog of the Erdős problem, intervals with splitting of size one are forbidden.

Goldberg and West [13] proved that every k-colored interval has a splitting of size at most k (see also [5] for a short proof using the Borsuk-Ulam theorem, and [20] for other applications of the Borsuk-Ulam theorem in combinatorics). This result is clearly the best possible, as can be seen in a necklace where colors occupy consecutively full intervals. Our result goes in the other direction and provides an upper bound for the number of colors in a general version of the Erdős problem on the line.

Theorem 1. *For every $k \geq 1$ there is a $(k + 3)$-coloring of the real line such that no interval has a splitting of size at most k.*

The proof is based on the Baire category theorem applied to the space of all measurable colorings of $\mathbb{R}$, equipped with a suitable metric (Section 2). By the same argument one may obtain other versions of the result. For instance, one of these asserts that there is a 5-coloring of $\mathbb{R}$ such that no two intervals (not necessarily adjacent) contain the same measure of every color (Section 3). We do not know whether the bound in Theorem 1 is optimal, even in the simplest case of $k = 1$, though one can show that two colors are not enough to avoid intervals with splitting of size one.

2. Proof of the main result

Recall that a set in a metric space is *nowhere dense* if the interior of its closure is empty. A set is said to be of *first category* if it can be represented as a countable union of nowhere dense sets. In the proof of Theorem 1 we apply the Baire category theorem in the following form (see [21]).

Theorem 2 (Baire)**.** *If X is a complete metric space and A is a set of first category in X, then $X \setminus A$ is dense in X (and in particular is nonempty).*

Our plan is to construct a suitable metric space of colorings and demonstrate that the subset of *bad* colorings is of first category.

2.1. **The setting.** Let k be a fixed positive integer and let $\{1, 2, \ldots, k\}$ be the set of colors. Let f, g be two measurable k-colorings of $\mathbb{R}$. Let n be another positive integer and consider the set

$$D_n(f, g) = \{x \in [-n, n] : f(x) \neq g(x)\}.$$

Clearly $D_n(f, g)$ is measurable, and we may define the normalized distance between f and g on $[-n, n]$ by

$$d_n(f, g) = \frac{\lambda(D_n(f, g))}{n},$$

where $\lambda(D)$ is the Lebesgue measure of D. Then we may define the distance between any two measurable colorings f and g by

$$d(f,g) = \sum_{n=1}^{\infty} \frac{d_n(f,g)}{2^{n+1}}.$$

Identifying colorings whose distance is zero gives the metric space $\mathcal{M}$ of equivalence classes of all measurable k-colorings. Clearly, equivalent colorings preserve the anti-splitting properties we are looking for.

Lemma 1. *$\mathcal{M}$ is a complete metric space.*

This lemma is a straightforward generalization of the fact that sets of finite measure in any measure space form a complete metric space, with symmetric difference of sets as the distance function (see [21]).

Let $t \geq 1$ be a fixed integer. Let $\mathcal{D}_t$ be the subspace of $\mathcal{M}$ consisting of those k-colorings that avoid intervals having a splitting of size at most t. Our task is to show that $\mathcal{D}_t$ is not empty provided $t \leq k-3$. By *granularity* of a splitting we mean the length of the shortest subinterval in the splitting. For $r, n \geq 1$ let $\mathcal{B}_n^{(r)}$ be the set of those colorings from $\mathcal{M}$ for which there exists at least one interval contained in $[-n,n]$ having a splitting of size exactly r and granularity at least $1/n$. Finally let

$$\mathcal{B}_n(t) = \bigcup_{r=1}^{t} \mathcal{B}_n^{(r)}.$$

These are the bad colorings. Clearly we have

$$\mathcal{D}_t = \mathcal{M} \setminus \bigcup_{n=1}^{\infty} \mathcal{B}_n(t).$$

So, our aim is to show that the sets $\mathcal{B}_n(t)$ are nowhere dense, provided $k \geq t+3$, and hence that the union $\bigcup_{n=1}^{\infty} \mathcal{B}_n(t)$ is of first category.

2.2. **The sets $\mathcal{B}_n(t)$ are closed.** We show that each set $\mathcal{B}_n^{(r)}$ is a closed subset of $\mathcal{M}$. Since $\mathcal{B}_n(t)$ is a finite union of these sets, it must be closed too. For any family F of measurable subsets of $\mathbb{R}$ and any coloring f, we denote by $\lambda_i(f,F)$ the measure of color i in the union of all members of F.

Lemma 2. *For every $r, n \geq 1$, the set $\mathcal{B}_n^{(r)}$ is a closed subset of the space $\mathcal{M}$.*

Proof. Fix $n, r \geq 1$. Let f_m be an infinite sequence of colorings from $\mathcal{B}_n^{(r)}$ tending to the limit coloring f. Let $I^{(m)} = [a^{(m)}, b^{(m)}]$ be a sequence of intervals in $[-n,n]$, each having a splitting $a^{(m)} = y_0^{(m)} \leq y_1^{(m)} \leq \ldots \leq y_r^{(m)} \leq y_{r+1}^{(m)} = b^{(m)}$ of granularity at least $1/n$. Let $F^{(m)} = \{[y_i^{(m)}, y_{i+1}^{(m)}] : 0 \leq i \leq r\}$ be the resulting family of intervals and let $F^{(m)} = F_1^{(m)} \cup F_2^{(m)}$ be the related fair partition of $F^{(m)}$. Finally, let $\{0,1,\ldots,r\} = A^{(m)} \cup B^{(m)}$ be the associated partition of indices, that is, $F_1^{(m)} = \{[y_i^{(m)}, y_{i+1}^{(m)}] : i \in A^{(m)}\}$ and $F_2^{(m)} = \{[y_i^{(m)}, y_{i+1}^{(m)}] : i \in B^{(m)}\}$.

Since there are only finitely many index patterns $(A^{(m)}, B^{(m)})$, one of them must appear infinitely many times in the sequence. Hence, without loss of generality, we may assume that all the $A^{(m)}$ and all the $B^{(m)}$ are equal, say $A^{(m)} = A$ and $B^{(m)} = B$ for every $m \geq 1$. Since $[-n,n]$ is compact, there are subsequences of the $r+1$ sequences $y_i^{(m)}$, $0 \le i \le r+1$, which are convergent to some points $y_i \in [-n,n]$, respectively. Clearly the splitting $a = y_0 \le y_1 \le \ldots \le y_r \le y_{r+1} = b$ has granularity at least $1/n$. Moreover, we claim that the related fair partition $F = F_1 \cup F_2$ of the family $F = \{[y_i, y_{i+1}] : 0 \le i \le r\}$ has the same index pattern (A, B). Indeed, if this were not the case, then there would be a color i such that

$$|\lambda_i(f, F_1) - \lambda_i(f, F_2)| > \varepsilon,$$

for some $\varepsilon > 0$. Taking m large enough we can make the distances $d(f_m, f)$ arbitrarily small, so that

$$|\lambda_i(f_m, F_1) - \lambda_i(f_m, F_2)| > \varepsilon_1$$

for some $\varepsilon_1 > 0$. Now, for sufficiently large m, the symmetric difference between the unions of intervals in F_1 and $F_1^{(m)}$ can also be made arbitrarily small; hence

$$\left|\lambda_i(f_m, F_1^{(m)}) - \lambda_i(f_m, F_2^{(m)})\right| > \varepsilon_2,$$

for some $\varepsilon_2 > 0$. This contradicts the assumption that $F^{(m)} = F_1^{(m)} \cup F_2^{(m)}$ is a fair partition. Consequently, the limit coloring f must be in $\mathcal{B}_n^{(r)}$. □

2.3. **The sets $\mathcal{B}_n(t)$ have empty interiors.** Next we prove that each $\mathcal{B}_n(t)$ has empty interior provided the number of colors k is at least $t+3$. For this purpose, let us call f an *interval* coloring on $[-n,n]$ if there is a partition of the segment $[-n,n]$ into some number of (half-open) intervals of equal length, each filled with only one color. Let $\mathcal{I}_n$ denote the set of all colorings from $\mathcal{M}$ that are interval colorings on $[-n,n]$.

Lemma 3. *Let $f \in \mathcal{M}$ be a coloring. Then for every $\varepsilon > 0$ and $n \in \mathbb{N}$ there exists a coloring $g \in \mathcal{I}_n$ such that $d(f,g) < \varepsilon$.*

Proof. Let $C_i = f^{-1}(i) \cap [-n,n]$ and let $C_i^* \subset [-n,n]$ be a finite union of intervals such that

$$\lambda((C_i^* \setminus C_i) \cup (C_i \setminus C_i^*)) < \frac{\varepsilon}{2k^2}$$

for each $i = 1, 2, \ldots, k$. Define a coloring h so that for each $i = 1, 2, \ldots, k$, the set $C_i^* \setminus (C_1^* \cup \ldots \cup C_{i-1}^*)$ is filled with color i, the rest of interval $[-n,n]$ is filled with any of these colors, and h agrees with f everywhere outside $[-n,n]$. Then $d(f,h) < \varepsilon/2$ and clearly each $h^{-1}(i) \cap [-n,n]$ is a finite union of intervals. Let $A_1, \ldots, A_t$ be the whole family of these intervals. Now split the interval $[-n,n]$ into $N \ge 8tn/\varepsilon$ intervals $B_1, \ldots, B_N$ of equal length and define a new coloring g so that $g(B_i) = h(A_j)$ if $B_i \subset A_j$ and $g(B_i)$ is any color otherwise. Hence, g differs from h on at most $2t$ intervals of total length $2t2n/N \le \varepsilon/2$, and we get $d(f,g) \le \varepsilon$. □

Lemma 4. *If $k \ge t+3$, then each set $\mathcal{B}_n(t)$ has empty interior.*

Proof. Let $f \in \mathcal{B}_n(t)$ be any bad coloring. Let $U(f,\varepsilon)$ be the open ε-neighborhood of f in the space $\mathcal{M}$. Assume the assertion of the lemma is false: there is some $\varepsilon > 0$ for which $U(f,\varepsilon) \subset \mathcal{B}_n(t)$. By Lemma 3 there is a coloring $g \in \mathcal{I}_n$ such that $d(f,g) < \varepsilon/2$, so $U(g, \varepsilon/2) \subset \mathcal{B}_n(t)$. The idea is to modify slightly the interval coloring g so that the new coloring will still be close to g, but there will be no intervals inside $[-n,n]$ possessing a splitting of size at most t and granularity at least $1/n$. Without loss of generality we may assume that there are equally spaced points $-n = a_0 < a_1 < \ldots < a_N = n$ in the interval $[-n,n]$ such that each interval

$A_i = [a_{i-1}, a_i)$ is filled with a unique color in the interval coloring g. Let $\delta > 0$ be a real number satisfying

$$\delta < \min\left\{\frac{\varepsilon}{2N}, \frac{n}{N^2}\right\}.$$

Consider another collection of intervals $S_i = [a_{i-1}, b_i)$, with $a_{i-1} < b_i < a_i$, such that $\lambda(S_i) = \delta$ for all $i = 1, 2, \ldots, N$.

Now randomly split each S_i into k subintervals $I_i^{(m)}$, $1 \leq m \leq k$. Then, with probability one, for every fixed m, $1 \leq m \leq k$, the set $\{\lambda(I_i^{(m)}) : 1 \leq i \leq N\}$ is linearly independent over the rationals, and in particular no two nonempty disjoint subsets of it can have the same total sum. Let h_i be a k-coloring of S_i defined by $h_i^{-1}(m) = I_i^{(m)}$ for every color $m \in \{1, 2, \ldots, k\}$. Finally, let h be a coloring such that $h = h_i$ on S_i and $h = g$ everywhere outside the union of S_i.

First notice that

$$d(g, h) \leq \sum_{i=1}^{N} \lambda(S_i) = \delta N < \frac{\varepsilon}{2}$$

by the choice of δ. Hence h is in $\mathcal{B}_n(t)$. Let $[a, b] \subset [-n, n]$ be an interval with a splitting $a = y_0 < y_1 < \ldots < y_r < y_{r+1} = b$ of size $r \leq t$ and granularity at least $1/n$. Let $F = F_1 \cup F_2$ be the fair partition of the family $F = \{[y_i, y_{i+1}] : 0 \leq i \leq r\}$; that is, we have

$$\lambda_j(h, F_1) = \lambda_j(h, F_2)$$

for every $1 \leq j \leq k$. Since $k \geq t + 3 > r + 2$, there is a color $s \in \{1, 2, \ldots, k\}$ that does not appear on any of the points $y_0, y_1, \ldots, y_{r+1}$ in coloring h. Hence for every open interval I filled with color s in coloring h and for every member Y of the family F, either I is completely contained in Y or I and Y are disjoint.

We distinguish two types of intervals filled with color s: those of the form (b_i, a_i) (*large* intervals) and those of the form $I_i^{(s)}$ (*small* intervals). Let l_i denote the number of large intervals contained in the union of all members of F_i, $i = 1, 2$. We claim that $l_1 = l_2$. Indeed, suppose that this is not the case and assume (without loss of generality) that $l_1 > l_2$. Denote by $L = \frac{2n}{N}$ the common length of intervals $A_i = [a_{i-1}, a_i)$. Then

$$\lambda_s(h, F_1) \geq l_1 (L - \delta) = l_1 L - l_1 \delta \geq l_1 L - N\delta > l_1 L - \frac{n}{N}$$

by the choice of δ. On the other hand,

$$\begin{aligned} \lambda_s(h, F_2) &\leq l_2 L + N\delta \leq (l_1 - 1)L + N\delta = l_1 L - L + N\delta \\ &< l_1 L - \frac{2n}{N} + \frac{n}{N} = l_1 L - \frac{n}{N}, \end{aligned}$$

again by the initial choice of δ. This is a contradiction, so we have $l_1 = l_2$.

Since $\lambda_s(h, F_1) = \lambda_s(h, F_2)$, the sum of lengths of small intervals of color s in F_1 equals the sum of lengths of small intervals of color s in F_2. However this contradicts the choice of the numbers $\lambda(I_i^{(m)})$ as rationally independent. The proof of the lemma is complete. $\square$

3. Generalizations and open problems

In [2] the necklace splitting theorem was generalized to fair partitions into more than just two collections. A *q-splitting* of size r of the necklace $[a, b]$ is a sequence $a = y_0 < y_1 < \ldots < y_r < y_{r+1} = b$ such that it is possible to partition the resulting collection of intervals $F = \{[y_i, y_{i+1}] : 0 \leq i \leq r\}$ into q disjoint subcollections $F_1, F_2, \ldots, F_q$, each capturing exactly $1/q$ of the total measure of every color. So, $F = F_1 \cup F_2 \cup \ldots \cup F_q$ is a *fair* partition of F into q parts. The result of [2] asserts that every k-colored necklace has a q-splitting of size at most $(q-1)k$, which is clearly the best possible.

Notice that we may speak about fair partitions of any collection F of intervals on the line (not necessarily adjacent) whose interiors are pairwise disjoint. Modifying slightly the proof of Theorem 1, one may obtain the following more general result.

Theorem 3. *For every $k \geq 4$ there exists a measurable k-coloring of the real line with the following property: there is no family F of closed intervals whose members have pairwise disjoint interiors and at most $k-1$ endpoints in total such that F has a fair partition into q parts, for any $q \geq 2$.*

For instance, there is a 5-coloring of the real line such that there is no pair of intervals (not necessarily adjacent) having the same measure of every color (this solves an open problem posed in [15]). It is not known whether the constant 5 is optimal for this property, but it is not hard to show that two colors are not enough.

Let us go back at the end to the discrete case. By the discrete necklace splitting theorem (see [13], [5], [20]), any k-colored necklace has a splitting of size at most k (provided the number of beads in each color is even). It is not clear, however, if the following discrete version of Theorem 1 holds.

Problem 1. Is it true that for every $k \geq 1$ there is a $(k+3)$-coloring of the integers such that no segment has a splitting of size at most k?

References

[1] J-P. Allouche, J. Shallit, Automatic sequences. Theory, applications, generalizations, Cambridge University Press, Cambridge, 2003. MR1997038 (2004k:11028)

[2] N. Alon, Splitting necklaces, Advances in Math. 63 (1987), 247-253. MR877785 (88f:05010)

[3] N. Alon, J. Grytczuk, M. Hałuszczak, O. Riordan, Nonrepetitive colorings of graphs, Random Struct. Alg. 21 (2002), 336-346. MR1945373 (2004a:05047)

[4] N. Alon, J. Grytczuk, Breaking the rhythm on graphs, Discrete Math. 308 (2008), 1375-1380. MR2392054

[5] N. Alon, D. West, The Borsuk-Ulam theorem and bisection of necklaces, Proc. Amer. Math. Soc. 98 (1986), 623-628. MR861764 (88b:05017)

[6] D.R. Bean, A. Ehrenfeucht, G.F. McNulty, Avoidable patterns in strings of symbols, Pacific J. Math. 85 (1979), 261-294. MR574919 (81i:20075)

[7] 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. MR1536459

[8] Ch. Choffrut, J. Karhumäki, Combinatorics of words, in: Handbook of Formal Languages, G. Rozenberg, A. Salomaa (Eds.), Springer-Verlag, Berlin-Heidelberg, 1997, 329-438. MR1469998 (98k:68141)

[9] J.D. Currie, Unsolved problems: Open problems in pattern avoidance, Amer. Math. Monthly 100 (1993), 790-793. MR1542406

[10] F.M. Dekking, Strongly nonrepetitive sequences and progression-free sets, J. Combin. Theory Ser. A 27 (1979), 181-185. MR542527 (81b:05027)

[11] P. Erdős, Some unsolved problems, Magyar Tud. Akad. Mat. Kutató Int. Közl. 6 (1961), 221-254. MR0177846 (31:2106)

[12] A. A. Evdokimov, Strongly asymmetric sequences generated by a finite number of symbols, Dokl. Akad. Nauk. SSSR 179 (1968), 1268-1271; Soviet Math. Dokl. 9 (1968), 536-539. MR0234842 (38:3156)

[13] C. H. Goldberg, D. B. West, Bisection of circle colorings, SIAM J. Algebraic Discrete Methods 6 (1985), 93-106. MR772181 (86c:05010)

[14] J. Grytczuk, Pattern avoiding colorings of Euclidean spaces, Ars Combin. 64 (2002), 109-116. MR1914201 (2003e:05087)
[15] J. Grytczuk, W. Śliwa, Non-repetitive colorings of infinite sets, Discrete Math. 265 (2003), 365-373. MR1969386 (2004g:68135)
[16] J. Grytczuk, Nonrepetitive colorings of graphs—a survey, Int. J. Math. Math. Sci. (2007), Art. ID 74639, 10 pp. MR2272338 (2008f:05060)
[17] J. Grytczuk, Thue type problems for graphs, points, and numbers, Discrete Math. (2008).
[18] V. Keränen, Abelian squares are avoidable on 4 letters, Automata, Languages and Programming, Lecture Notes in Computer Science 623, Springer, Berlin, 1992, 41-52. MR1250629 (94j:68244)
[19] M. Lothaire, Combinatorics on Words, Addison-Wesley, Reading, MA, 1983. MR675953 (84g:05002)
[20] J. Matoušek, Using the Borsuk-Ulam theorem, Springer-Verlag, Berlin, 2003. MR1988723 (2004i:55001)
[21] J. C. Oxtoby, Measure and Category, Grad. Texts in Math. 2, Springer-Verlag, New York-Berlin, 1980. MR584443 (81j:28003)
[22] P.A.B. Pleasants, Non-repetitive sequences, Proc. Cambridge Philos. Soc. 68 (1970), 267-274. MR0265173 (42:85)
[23] A. Thue, Über unendliche Zeichenreichen, Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiania 7 (1906), 1-22.
[24] A. Thue, Über die gegenseitigen Lage gleicher Teile gewisser Zeichenreihen, Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiania 1 (1912), 1-67.

Schools of Mathematics and Computer Science, Raymond and Beverly Sacler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel – and – Institute for Advanced Study, Princeton, New Jersey 08540

E-mail address: nogaa@tau.ac.il

Theoretical Computer Science Department, Faculty of Mathematics and Computer Science, Jagiellonian University, 30-387 Kraków, Poland

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

Faculty of Mathematics and Computer Science, Jagiellonian University, 30-387 Kraków, Poland

E-mail address: mlason@op.pl

Faculty of Mathematics and Computer Science, Jagiellonian University, 30-387 Kraków, Poland

E-mail address: wajcha2@poczta.onet.pl