Long rainbow cycles in proper edge-colorings of complete graphs *

András Gyárfás, Miklós Ruszinkó[†]

Computer and Automation Research Institute
Hungarian Academy of Sciences
Budapest, P.O. Box 63
Budapest, Hungary, H-1518
gyarfas,ruszinko@sztaki.hu

Gábor N. Sárközy[‡]

Computer Science Department
Worcester Polytechnic Institute
Worcester, MA, USA 01609
gsarkozy@cs.wpi.edu
and
Computer and Automation Research Institute
Hungarian Academy of Sciences
Budapest, P.O. Box 63
Budapest, Hungary, H-1518

Richard H. Schelp[§]

University of Memphis
Department of Mathematical Sciences
Memphis, TN 38152, USA
rschelp@memphis.edu

Dedicated to the memory of our colleague and friend Dick Schelp.

*2000 Mathematics Subject Classification: 05C70, 05C38, 05C15.

[†]Research supported in part by OTKA Grant No. K68322.

[‡]Research supported in part by OTKA Grant No. K68322, by NSF Grant DMS-0968699 and by a Janos Bolyai Research Scholarship.

[§]Research supported by the FIST (Finite Structures) project, in the framework of the European Community's "Structuring the European Research Area programme".

Abstract

We show that any properly edge-colored K_n contains a rainbow cycle with at least $(4/7 - o(1))n$ edges. This improves the lower bound of $n/2 - 1$ proved in [1].

We consider *properly edge-colored* complete graphs K_n, where two edges with the same color cannot be incident to each other, so each color class is a matching. An important and well investigated special case of proper edge-colorings is a *factorization* where each color class forms a perfect (if n is even) or nearly perfect (if n is odd) matching. A colored subgraph of K_n is called *rainbow* if its edges have different colors.

The size of rainbow subgraphs of maximum degree two, i.e. union of paths and cycles in proper colorings are well investigated. A consequence of Ryser's well-known conjecture ([12] stating that every Latin square has a transversal) would be that for odd n in every factorization of K_n there is a rainbow 2-factor (and for even n a 2-factor covering all but one vertices). Although this is not known, there were several results that made advances towards Ryser's conjecture and show the existence of a 2-factor covering $n - o(n)$ vertices, [4, 10, 13, 14]. Andersen [3] applied the method of [4] to prove that in every proper coloring of K_n there is a rainbow subgraph with at least $n - \sqrt{2n}$ vertices whose components are paths.

Another line of research looked for rainbow Hamiltonian cycles from the assumption that there is an upper bound k on the number of colors in each color class. This problem is mentioned in Erdős, Nesetril and Rödl [5]. Hahn and Thomassen [9] showed that k could grow as fast as $n^{1/3}$ and in fact Hahn conjectured (see [9]) that the growth of k could be linear in n. After further improvements [7], Albert, Frieze and Reed [2] proved the Hahn Conjecture by showing that k could be $\lceil cn \rceil$, for any constant $c < 1/32$ if $n \geq n_0(c)$. See also [6] for related results.

Although it is widely believed that in every proper coloring of K_n there is a rainbow path and cycle with length almost n (the obstacle to a spanning rainbow path or cycle comes from a special factorization, see [1], [9], [11]), the above mentioned results do not imply such a bound. As far as we know the best lower bounds are $2n/3$ for the path ([8]) and $n/2-1$ for the cycle ([1]). The purpose of this note is the improvement of the latter result to $(1 - o(1))\frac{4n}{7}$.

Theorem 1. *For arbitrary ε, where $1/2 > \varepsilon > 0$, there exists an $n_0(\varepsilon)$ such that if $n \geq n_0(\varepsilon)$, then in any proper edge-coloring of K_n there is a rainbow cycle with length at least $\left(\frac{4}{7} - \varepsilon\right) n$.*

Proof: The vertex-set and the edge-set of a graph G are denoted by $V(G)$ and $E(G)$. C_l is the cycle with l vertices and P_l is the path with l vertices.

Fix ε, such that $1/2 > \varepsilon > 0$ and choose constants $d = d(\varepsilon)$ and $n_0 = n_0(\varepsilon)$ in the following way:

$$d = d(\varepsilon) = \left(\frac{48}{7\varepsilon}\right)^2, \; n_0 = n_0(\varepsilon) = \frac{8(d+1)}{\varepsilon}. \tag{1}$$

Assume that $n \geq n_0$. Let us take an arbitrary proper edge-coloring of K_n and let $C_t = \{v_1, \ldots, v_t\}$ be a rainbow cycle with t edges such that t is maximum. We will show that

$$t \geq \left(\frac{4}{7} - \varepsilon\right) n.$$

During the proof we will try to increase the length of C_t using rainbow "detours". More precisely a segment of the cycle C_t will be deleted and replaced by a new part. If the vertices added to the cycle are greater in number than those removed, a longer rainbow cycle is obtained contradicting the fact that C_t has maximum length. The colors already used on C_t will be called old colors and the set containing them will be denoted by OLD. The colors not used yet are called new colors and the set containing them will be denoted by NEW, i.e., we start with $OLD = \{\textit{colors used along } C_t\}$ and NEW consists of the remaining colors. These sets of colors, however, may vary during the proof according to the detours along which we will try to enlarge C_t. For $x \in V$, $R \subseteq V$ we denote by $deg_{NEW}(x, R)$ the number of edges adjacent to x and $u \in R$ having color from NEW.

To make the presentation more transparent, we avoid using floors and ceilings. Since the obtained result is probably far from the best possible these "inaccuracies" do not have any impact.

Case 1: There exists a pair of vertices y_1 and y_2 in C_t which are within distance d along the cycle and which are adjacent to two different vertices, say x_1 and x_2, in two different new colors in the remaining part of the vertex set $R = V \backslash V(C_t)$. Here we will try to delete this short segment of C_t between y_1 and y_2 and replace it with a longer rainbow path, as outlined above. Move the two new colors used on the edges (x_1, y_1) and (x_2, y_2) from NEW to OLD. Notice, that no vertex $x \in R$ is connected to two consecutive v_i, v_{i+1} vertices along C_t in new colors, since otherwise we obtain a longer cycle by substituting the edge (v_i, v_{i+1}) by the path $P = \{v_i, x, v_{i+1}\}$. Therefore, for an arbitrary vertex $x \in R$

$$deg_{NEW}(x, R) \geq n - t - 1 - t/2 - 2 = n - 3t/2 - 3. \tag{2}$$

Next we find a rainbow path P_d with d vertices in R in new colors starting at x_1 and avoiding x_2. This is always possible assuming

$$deg_{NEW}(x, R) - 2d \geq n - 3t/2 - 3 - 2d \geq 0, \quad \text{i.e.,} \quad t \leq 2n/3 - 4d/3 - 2.$$

Let x_1' be the other endpoint of P_d and set $R' = R \setminus (V(P_d) \setminus x_1')$. Move the colors along the path P_d from NEW to OLD, i.e.,

$$|NEW| \geq n - t - 2 - d.$$

Similar to (2), for an arbitrary vertex $x \in R'$

$$deg_{NEW}(x, R') \geq n - 3t/2 - 3 - (2d - 1) = n - 3t/2 - 2d - 2, \tag{3}$$

where we have to subtract $d-1$ colors used in P_d and d other colors (possibly) going to vertices in P_d. Let $N_{NEW}(x, R')$ be the set of those vertices in R' which are adjacent to x in new colors. Set

$$\Gamma_1 = N_{NEW}(x_1', R'), \Gamma_2 = N_{NEW}(x_2, R').$$

If there exists an $z \in \Gamma_1 \cap \Gamma_2$, then we could substitute the path $\{y_1, \ldots, y_2\}$ of length $\leq d$ along the cycle by the path $\{y_1, x_1, P_d, x_1', z, x_2, y_2\}$ of length $> d$ and obtain a longer rainbow cycle, a contradiction. So assume $\Gamma_1 \cap \Gamma_2 = \emptyset$. Then, since $|R'| = n - t - d + 1$, for

$$S = R' \setminus (\Gamma_1 \cup \Gamma_2),$$

by (3) we have

$$|S| \leq n - t - d + 1 - 2(n - 3t/2 - 2 - 2d) = 2t - n + 3d + 5.$$

Without loss of generality we may assume that $|\Gamma_1| \geq |\Gamma_2|$ and then, clearly,

$$\frac{n - t - d + 1}{2} \leq \frac{|R'|}{2} \leq |\Gamma_1 \cup S|,$$

and by (3)

$$|\Gamma_1 \cup S| \leq |R'| - (n - 3t/2 - 2 - 2d) = n - t - d + 1 - (n - 3t/2 - 2 - 2d) = \frac{t}{2} + d + 3.$$

Notice, that if $x \in \Gamma_1$ is adjacent to a vertex $z \in \Gamma_2$ in a new color then (x_1', x) and (x_2, z) must have the same color. Otherwise we could substitute the path of length $\leq d$ $\{y_1, \ldots, y_2\}$ in the cycle by the path $\{y_1, x_1, P_d, x_1', x, z, x_2, y_2\}$ of length $> d$ and obtain a longer rainbow cycle, a contradiction. And since the coloring is proper, every $x \in \Gamma_1$ is adjacent to at most one vertex $z \in \Gamma_2$ in a new color. Therefore, every vertex $x \in \Gamma_1$ has all but at most one of its neighbors in new colors in $\Gamma_1 \cup S$, i.e., by (3) for every $x \in \Gamma_1$

$$deg_{NEW}(x, \Gamma_1 \cup S) \geq n - 3t/2 - 2d - 3.$$

If twice this degree is greater than $|\Gamma_1 \cup S| + 3$, i.e.,

$$2(n - 3t/2 - 2d - 3) \geq \frac{t}{2} + d + 6 \geq |\Gamma_1 \cup S| + 3, \tag{4}$$

then two arbitrary vertices in Γ_1 can be joined by 3 different paths of length two in new colors. If (4) does not hold, then we have

$$t > \frac{4n}{7} - \frac{10d + 24}{7}, \tag{5}$$

i.e., the original cycle is sufficiently large. Therefore, we will assume that two arbitrary vertices in Γ_1 can be joined by three paths of length two in new colors.

To finish this case we will try to find two vertices of distance 1, 2 or 3, say v_i and v_j, $|j - i| \leq 3$, along the cycle such that they are adjacent to two different vertices $x_i, x_j \in \Gamma_1$ in two different new colors. If such two edges exist, then one of the 3 existing paths, say P, of length 2 between x_i and x_j in new colors contains neither the color of the edge (v_i, x_i), nor the color of the edge (v_j, x_j). Replacing the path of length ≤ 3 $\{v_i, \ldots, v_j\}$ by the path $\{v_i, x_i, P, x_j, v_j\}$ of length four we obtain a longer rainbow cycle, a contradiction.

Notice that every $x \in \Gamma_1$ satisfies

$$deg_{NEW}(x, C_t) \geq |NEW| - |P_d| - (|\Gamma_1 \cup S| - 1) - 1 \geq n - t - 2d - 4 - (|\Gamma_1 \cup S| - 1) \geq$$

$$\geq n - t - 2d - \left(\frac{t}{2} + d + 3\right) - 3 = n - \frac{3t}{2} - 3d - 6, \tag{6}$$

and therefore, for the number of edges in new colors $|E_{NEW}[\Gamma_1, C_t]|$ in the bipartite graph with parts C_t and Γ_1 by (3) and (6) we have

$$|E_{NEW}[\Gamma_1, C_t]| \geq deg_{NEW}(x'_1, R') \cdot (n - \frac{3t}{2} - 3d - 6). \tag{7}$$

Next we get an upper bound for the number of these edges with respect to the degrees of the vertices in C_t. In order to have this, partition the vertices along C_t into consecutive quadruples $\{v_1, v_2, v_3, v_4\}$, $\{v_5, v_6, v_7, v_8\}$, $\ldots$. (If 4 does not divide t then let the last part contain one, two or three vertices.)

Claim 2. *If for some i and for some quadruple $\{v_{4i+1}, v_{4i+2}, v_{4i+3}, v_{4i+4}\}$ the sum of the degrees*

$$s_i = \sum_{j=1}^{4} deg_{NEW}(v_{4i+j}, \Gamma_1) \geq |\Gamma_1| + 3, \tag{8}$$

then there exist v_{4i+k} and $v_{4i+\ell}$, $1 \leq k < l \leq 4$, such that they are adjacent to two different vertices x_i and x_j in Γ_1 in two different new colors.

Proof. Indeed, if (8) holds then there have to be vertices in Γ_1 which are covered 2 or 3 times by the four sets in

$$T_i = \bigcup_{j=1}^{4} N_{NEW}(v_{4i+j}, \Gamma_1).$$

If $\exists x \in \Gamma_1$ which is covered 3 times, then x is connected to 3 vertices out of four consecutive ones in C_t. Out of these three vertices two have to be consecutive, contradicting the maximality of C_t. If $\nexists x \in \Gamma_1$ which is covered 3 times, then there must be (at least) three vertices, say, $x_1, x_2, x_3 \in \Gamma_1$ which are covered twice by $\cup_{j=1}^{4} N_{NEW}(v_{4i+j}, \Gamma_1)$. Consider the bipartite graph G_i with parts $A_i = \{v_{4i+j} : j = 1, \dots, 4\}$ and $B = \{x_1, x_2, x_3\}$ with the edges defined by T_i. All vertices in B are of degree 2. A trivial case analysis shows that there always exists a rainbow matching formed by two edges of G_i. $\square$

So we may assume that for each i, inequality (8) does not hold. But then for the number of new edges $|E_{new}[\Gamma_1, C_t]|$ between Γ_1 and C_t by Claim 2

$$|E_{new}[\Gamma_1, C_t]| \leq \frac{t}{4}(deg_{NEW}(x_1', R') + 2) \tag{9}$$

holds. Combining estimates (7, 9) we get

$$deg_{NEW}(x_1', R') \cdot (n - \frac{3t}{2} - 3d - 6) \leq \frac{t}{4}(deg_{NEW}(x_1', R') + 2),$$

which implies (dividing by $deg_{NEW}(x_1', R')$ and using (3))

$$t \geq \frac{4n}{7} - \frac{12d + 24}{7} - \frac{2t}{7(n - \frac{3t}{2} - 2d - 2)}. \tag{10}$$

Here for the last term we have

$$\frac{2t}{7(n - \frac{3t}{2} - 2d - 2)} \leq \frac{8}{7}. \tag{11}$$

Indeed, if (11) does not hold, then we have

$$t > \frac{4n}{7} - \frac{8d + 8}{7}, \tag{12}$$

i.e. again we have a lower bound similar to (5). Thus otherwise from (10) we get

$$t \geq \frac{4n}{7} - \frac{12d + 32}{7}. \tag{13}$$

Case 2: Assume that no pair of vertices y_1, y_2 exists within distance d along the cycle that are adjacent in two different new colors to two different vertices, say $x_1, x_2 \in R$. This implies easily that in each interval of length d along the cycle there is at most one vertex x with $deg_{NEW}(x, R) \geq 3$. Therefore, the number of edges in new colors between C_t and R is at most

$$\frac{t}{d}|R| + 2t \leq \frac{2t}{d}|R|,$$

since $2d \leq n/4 \leq |R|$ (using (1)).

Thus, if we denote by B the set of those bad vertices $x \in R$ for which

$$deg_{NEW}(x, C_t) \geq \frac{2t}{\sqrt{d}},$$

then we have

$$|B|\frac{2t}{\sqrt{d}} \leq \frac{2t}{d}|R| \quad \text{i.e.,} \quad |B| \leq \frac{|R|}{\sqrt{d}}.$$

Set $R^* = R \setminus B$. We have $|R^*| \geq (1 - \frac{1}{\sqrt{d}})|R|$. Moreover, R^* is almost complete in new colors. For every $x \in R^*$ we have:

$$deg_{NEW}(x, R^*) \geq |R| - 1 - \frac{2t}{\sqrt{d}} - \frac{|R|}{\sqrt{d}} \geq |R| - \frac{3t}{\sqrt{d}} - \frac{|R|}{\sqrt{d}} \geq$$

$$\geq |R|\left(1 - \frac{10}{\sqrt{d}}\right) \geq |R^*|\left(1 - \frac{10}{\sqrt{d}}\right), \tag{14}$$

where the third inequality is equivalent (through $|R| + t = n$) to $t \leq 3n/4$. We can assume this, otherwise we have nothing to prove.

Lemma 1. *Suppose k, l are given integers with $l < k/2$ and G is a properly edge colored k-vertex graph with minimum degree at least $k/2 + l$. Then an arbitrary pair of vertices $x_1, x_2 \in V(G)$ can be joined by a rainbow path of length at least $\frac{2l}{3}$.*

Proof. Starting at x_1, build a greedy path by extending the current endpoint $y \neq x_1$ with an edge yz such that $z \neq x_2$ and yz has a color not used on the current path. Assume that at a certain point we have $P = \{x_1, \ldots, y\}$. Call a color new, if it does not appear on P. Set $Q = V(G) \setminus (V(P) \cup \{x_2\})$ and $m = k/2 + l - 2|P|$. Observe that $deg_{NEW}(y, Q) \geq m$ and $deg_{NEW}(x_2, Q) \geq m$. Thus, if

$$2m = k + 2l - 4|P| > |Q| = k - |P| - 1,$$

i.e., if equivalently $\frac{2l+1}{3} > |P|$, then $M = N_{new}(y,Q) \cap N_{new}(x_2,Q) \neq \emptyset$. Thus with $w \in M$, the path $P^+ = Pwx_2$ is a rainbow path from x_1 to x_2 so there exists a path P^* such that

$$|P^*| = \left\lfloor \frac{2l+1}{3} \right\rfloor - 1 + 2 \geq \frac{2l}{3},$$

as desired. $\square$

Choose G as the subgraph induced by the edges with new colors in R^*, set $k = |R^*|$ and notice that using Lemma 1 and (14) we can join an arbitrary pair of vertices in R^* by a rainbow path in all new colors of length at least

$$|R^*| \left(\frac{1}{3} - \frac{20}{3\sqrt{d}}\right) \geq |R| \left(1 - \frac{1}{\sqrt{d}}\right)\left(\frac{1}{3} - \frac{20}{3\sqrt{d}}\right) \geq |R| \left(\frac{1}{3} - \frac{7}{\sqrt{d}}\right).$$

For some ℓ, move the colors of the edges of the path $v_1, \ldots, v_{l+1}$ along the cycle from OLD to NEW, now $|NEW| \geq n - t - 1 + \ell$. If

$$n - t - 1 + \ell \geq t + |B| + 3 \geq t + \frac{|R|}{\sqrt{d}} + 3 \quad \text{i.e.,}$$

$$\ell \geq 2t - n + \frac{|R|}{\sqrt{d}} + 4, \tag{15}$$

then v_1 and v_{l+1} both send at least 3 new colors to R^* out of which we can find a rainbow matching of two edges, say, (v_1, x_1), (v_{l+1}, x_2), where $x_1, x_2 \in R^*$. But if in addition

$$\ell \leq |R| \left(\frac{1}{3} - \frac{7}{\sqrt{d}}\right), \tag{16}$$

then we could substitute the path $\{v_1, \ldots, v_{l+1}\}$ by the path $\{v_1, x_1, P, x_2, v_{l+1}\}$ of length $\ell + 2$, where P is a path of length ℓ joining x_1 and x_2 which must exist by Claim 1, a contradiction. Therefore no ℓ satisfies both (15) and (16), so we may assume

$$|R| \left(\frac{1}{3} - \frac{7}{\sqrt{d}}\right) < 2t - n + \frac{|R|}{\sqrt{d}} + 4,$$

and substituting $|R|$ by $(n - t)$ we conclude that

$$7t > 4n + \frac{24t}{\sqrt{d}} - \frac{24n}{\sqrt{d}} - 12 > 4n - \frac{24n}{\sqrt{d}} - 12, \quad \text{i.e.,}$$

$$t > \frac{4n}{7} - \frac{24n}{7\sqrt{d}} - \frac{12}{7}. \tag{17}$$

We finish the proof by observing that with our choice of d and n_0 (see (1)) all the obtained lower bounds on t (namely (5), (12), (13) and (17)) are at least $(4/7-\varepsilon)n$. $\square$

References

[1] S. Akbari, O. Etesami, H. Manini, M. Mahmoody, On rainbow cycles in edge colored complete graphs, *Australasian Journal of Combinatorics,* 37 (2007), 33-42.

[2] M. Albert, A. Frieze, B. Reed, Multicoloured Hamilton cycles, *Electronic Journal of Combinatorics,* 2 (1995), #R10.

[3] L.D. Andersen, Hamilton circuits with many colours in properly edge-colored complete graphs, *Math. Scand.,* 64 (1989), 5-14.

[4] A.E. Brouwer, A.J. de Vries, R.M.A. Wieringa, A lower bound for the length of partial transversals in a Latin square, *Nieuw Arch. Wisk.,* (3) 24 (1978), 330-332.

[5] P. Erdős, J. Nesetril, V. Rödl, Some problems related to partitions of edges of a graph, in *Graphs and Other Combinatorial Topics*, Teubner, Leipzig (1983), 54-63.

[6] A. Frieze, M. Krivelevich, On rainbow trees and cycles, *Electronic Journal of Combinatorics,* 15(1) (2008), #R59.

[7] A. Frieze, B. Reed, Polychromatic Hamilton cycles, *Discrete Mathematics,* 118 (1993), 69-74.

[8] A. Gyárfás, M. Mhalla, Rainbow and orthogonal paths in factorizations of K_n, *Journal of Combinatorial Designs,* 18(3) (2010), 167-176.

[9] G. Hahn, C. Thomassen, Path and cycle sub-Ramsey numbers and an edge-colouring conjecture, *Discrete Mathematics,* 62 (1986), 29-33.

[10] P. Hatami, P.W. Shor, A lower bound for the length of a partial transversal in a Latin square, *Journal of Combinatorial Theory, Series A,* 115 (2008), 1103-1113.

[11] M. Maamoun, H. Meyniel, On a problem of G. Hahn about coloured Hamiltonian paths in K_{2n}, *Discrete Mathematics,* 51 (1984), 213-214.

[12] H.J. Ryser, Neuere Probleme der Kombinatorik, in: Vorträge über Kombinatorik, pp. 69 - 91, a collection of papers presented at the Matematischen Forschungsinstitute, Oberwolfach, July 24 - 29, 1967.

[13] P.W. Shor, A lower bound for the length of a partial transversal in a Latin square, *Journal of Combinatorial Theory, Series A,* 33 (1982), 1-8.

[14] D.E. Woolbright, An $n \times n$ Latin square has a transversal with at least $n - \sqrt{n}$ distinct symbols, *Journal of Combinatorial Theory, Series A,* 24 (1978), 235-237.