On rainbow connection

Yair Caro

Department of Mathematics
University of Haifa-Oranim, Tivon 36006, Israel
yacaro@kvgeva.org.il

Arie Lev

Department of Computer Sciences
Tel Aviv University and The Academic College of Tel-Aviv-Yaffo, Tel-Aviv 61161, Israel

Yehuda Roditty

Department of Computer Sciences
Tel Aviv University and The Academic College of Tel-Aviv-Yaffo, Tel-Aviv 61161, Israel
jr@post.tau.il

Zsolt Tuza*

Hungarian Academy of Sciences and University of Pannonia, Budapest, Hungary
jr@post.tau.il

Raphael Yuster

Department of Mathematics
University Haifa, Haifa 31905, Israel
raphy@math.haifa.ac.il

Submitted: Nov 14, 2007; Accepted: Apr 6, 2008; Published: Apr 18, 2008
Mathematics Subject Classification: 05C15, 05C40

Abstract

An edge-colored graph G is *rainbow connected* if any two vertices are connected by a path whose edges have distinct colors. The *rainbow connection number* of a connected graph G, denoted $rc(G)$, is the smallest number of colors that are needed in order to make G rainbow connected. In this paper we prove several non-trivial upper bounds for $rc(G)$, as well as determine sufficient conditions that guarantee $rc(G) = 2$. Among our results we prove that if G is a connected graph with n vertices and with minimum degree 3 then $rc(G) < 5n/6$, and if the minimum degree is δ then $rc(G) \leq \frac{\ln \delta}{\delta} n(1 + o_\delta(1))$. We also determine the threshold function for a random graph to have $rc(G) = 2$ and make several conjectures concerning the computational complexity of rainbow connection.

*Research supported in part by the Hungarian Scientific Research Fund, OTKA grant T-049613

1 Introduction

All graphs in this paper are finite, undirected and simple. We follow the notation and terminology of [2]. Connectivity is perhaps the most fundamental graph-theoretic property. There are many ways to strengthen the connectivity property, such as requiring hamiltonicity, k-connectivity, imposing bounds on the diameter, requiring the existence of edge-disjoint spanning trees, and so on.

A natural and interesting quantifiable way to strengthen the connectivity requirement was recently introduced by Chartrand et al. in [6]. An edge-colored graph G is *rainbow connected* if any two vertices are connected by a path whose edges have distinct colors. Clearly, if a graph is rainbow connected, then it is also connected. Conversely, any connected graph has a trivial edge coloring that makes it rainbow connected; just color each edge with a distinct color. Thus, one can properly define the *rainbow connection number* of a connected graph G, denoted $rc(G)$, as the smallest number of colors that are needed in order to make G rainbow connected. An easy observation is that if G has n vertices then $rc(G) \leq n-1$, since one may color the edges of a given spanning tree with distinct colors, and color the remaining edges with one of the already used colors or, as we shall equivalently and conveniently assume throughout this paper, leave the remaining edges uncolored. Chartrand et al. computed the precise rainbow connection number of several graph classes including complete multipartite graphs [6]. We note also the trivial fact that $rc(G) = 1$ if and only if G is a clique, the (almost) trivial fact that $rc(G) = n-1$ if and only if G is a tree, and the easy observation that a cycle with $k > 3$ vertices has rainbow connection number $\lceil k/2 \rceil$. Also notice that, clearly, $rc(G) \geq diam(G)$ where $diam(G)$ denotes the diameter of G.

Our goal in this paper is to study the extremal graph-theoretic behavior of rainbow connection. Motivated by the fact that there are graphs with minimum degree 2 and with $rc(G) = n-3$ (just take two vertex-disjoint triangles and connect them by a path of length $n-5$), and by the fact that cliques have $rc(G) = 1$, it is interesting to study the behavior of $rc(G)$ with respect to the minimum degree $\delta(G)$. Our main results relate the minimum degree of a graph with its rainbow connection number.

Is it true that minimum degree at least 3 guarantees $rc(G) \leq \alpha n$ where $\alpha < 1$ is independent of n? This turns out to be correct, although certainly not trivial. Indeed, we prove:

Theorem 1.1 *If G is a connected graph with n vertices and $\delta(G) \geq 3$ then $rc(G) < 5n/6$.*

The constant $5/6$ appearing in the proof of Theorem 1.1 is not optimal, but we are unable to improve it significantly. In fact, it provably *cannot* be replaced with a constant smaller than $3/4$, since, as we shall see, there are 3-regular connected graphs with $rc(G) = diam(G) = (3n-10)/4$. In fact, we conjecture:

Conjecture 1.2 *If G is a connected graph with n vertices and $\delta(G) \geq 3$ then $rc(G) < 3n/4$.*

We are able to assert Conjecture 1.2 in the case of 3-regular class-1 graphs (recall that a graph G is class-1 if its chromatic index $\chi'(G)$ is equal to its maximum degree). This is a special case of the following theorem that improves the bound in Theorem 1.1 in many cases.

Theorem 1.3 *Suppose G is a connected graph with n vertices, and assume that there is a set of vertex-disjoint cycles that cover all but s vertices of G. Then $rc(G) < 3n/4 + s/4 - 1/2$. In particular:*

1. *If G has a 2-factor then $rc(G) < 3n/4$.*

2. *If G is k-regular and k is even then $rc(G) < 3n/4$.*

3. *If G is k-regular and $\chi'(G) = k$ then $rc(G) < 3n/4$.*

Not surprisingly, as the minimum degree increases, the rainbow connection number decreases. Specifically, we can prove the following upper bound.

Theorem 1.4 *Let G be a connected graph with n vertices and minimum degree δ. Then,*

$$rc(G) \le \min\{\ n\frac{\ln \delta}{\delta}(1+o_\delta(1))\ ,\ n\frac{4\ln \delta + 3}{\delta}\ \}.$$

Already for $\delta = 18$, Theorem 1.4 gives a better bound 0.81 for $rc(G)$ than the bound 0.833 from Theorem 1.1. We do not know how far Theorem 1.4 is from being tight, but in any case it cannot be improved below $\frac{3n}{\delta+1} - \frac{\delta+7}{\delta+1}$ as there are connected n-vertex graphs with minimum degree δ and this diameter.

Our next two results give non-trivial sufficient conditions for having $rc(G) = 2$. As noted earlier, having diameter 2 is a necessary requirement for having $rc(G) = 2$, although certainly not sufficient (e.g., consider a star). Clearly, if $\delta(G) \ge n/2$ then $diam(G) = 2$, but we do not know if this guarantees $rc(G) = 2$. The next theorem shows that by slightly increasing the minimum degree assumption, $rc(G) = 2$ follows.

Theorem 1.5 *Any non-complete graph with $\delta(G) \ge n/2 + \log n$ has $rc(G) = 2$.*

Another intriguing question is the *random graph* setting. Let $G = G(n,p)$ denote, as usual, the random graph with n vertices and edge probability p. In the extensive study of the properties of random graphs, many researchers observed that there are sharp *threshold functions* for various natural graph properties. For a graph property A and for a function $p = p(n)$, we say that $G(n,p)$ satisfies A *almost surely* if the probability that $G(n,p(n))$ satisfies A tends to 1 as n tends to infinity. We say that a function $f(n)$ is a *sharp threshold function* for the property A if there are two positive constants c and C so that $G(n, cf(n))$ almost surely does not satisfy A and $G(n,p)$ satisfies A almost surely for all $p \ge Cf(n)$. It is well known that all monotone graph properties have a sharp threshold (see [3] and [7]). Since having $rc(G) \le 2$ is a monotone graph property (adding edges does not destroy this property), it has a sharp threshold function. The following theorem establishes it.

Theorem 1.6 *$p = \sqrt{\log n/n}$ is a sharp threshold function for the property $rc(G(n,p)) \le 2$.*

The rest of this paper is organized as follows. The next section is devoted to the proofs relating minimum degree with the rainbow connection number. In particular, Theorems 1.1, 1.3, and 1.4 are proved. Section 3 consists of the proofs of Theorems 1.5 and 1.6. The final section contains some concluding remarks and open problems, mainly concerning the computational complexity aspects of rainbow connection.

2 Bounded degree graphs

2.1 Minimum degree 3

In this subsection we prove Theorem 1.1. We start with a simple lemma that will be useful in many of the results of this paper.

Lemma 2.1 *If G is a connected graph and $H_1, \ldots, H_k$ is a partition of the vertex set of G into connected subgraphs then $rc(G) \leq k - 1 + \sum_{i=1}^{k} rc(H_i)$.*

Proof: Contracting each H_i to a single vertex we obtain a connected minor of G with k vertices. This minor has rainbow connection number at most $k - 1$. We color each edge connecting vertices in distinct H_i with the color of the corresponding minor edge. For each $i = 1, \ldots, k$, the edges inside each H_i are colored with a dedicated set of $rc(H_i)$ colors. ■

Notice that in Lemma 2.1 we allow the H_i to be singletons (and the rainbow connection number of singletons is 0).

The following proposition is an important ingredient in the proof of Theorem 1.1.

Proposition 2.2 *If G is a 2-connected graph with n vertices then $rc(G) \leq 2n/3$.*

Proof: If G is a 5-cycle then $rc(G) = 3$ so the proposition clearly holds in this case. Otherwise, let H be a maximal connected subgraph of G having the property that $rc(H) \leq 2h/3 - 2/3$, where h is the number of vertices of H. We first claim that H exists. Indeed, if G has a triangle then already taking H to be a triangle we obtain $rc(H) = 1 \leq 2 - 2/3$. Otherwise, if G has any cycle of length $k \geq 4$ and $k \neq 5$, then already taking H to be such a cycle we obtain $rc(H) = \lceil k/2 \rceil \leq 2k/3 - 2/3$. Otherwise, if each cycle of G is a C_5 then taking H to be a C_5 attached to one additional edge we obtain $rc(H) = 3$, $h = 6$, and $3 \leq 4 - 2/3$.

We next claim that $h \geq n - 2$. Indeed, assume first that there are three distinct vertices outside of H, say x_1, x_2, x_3, each having two neighbors in H (the neighbors of x_i do not have to be distinct from the neighbors of x_j). We can add x_1, x_2, x_3 to H and form a larger subgraph H' with $h + 3$ vertices. Suppose e_i, f_i are two edges connecting x_i with H. We use only two new colors to color the 6 designated edges; e_1, e_2, e_3 all get the same color and f_1, f_2, f_3 all get the same color. We now have

$$rc(H') \leq rc(H) + 2 \leq 2h/3 - 2/3 + 2 = 2(h + 3)/3 - 2/3$$

contradicting the maximality of H. It follows that if there are three vertices outside of H then at least one of these vertices, say x, has the property that a shortest path from H to H passing through x has length at least 3 (notice that there must be such a path as the graph is 2-connected). Let, therefore, $a, x_1, \ldots, x_t, b$ be a path with $a, b \in H$, with $x_1, \ldots, x_t \notin H$, and $t \geq 2$. We can add $x_1, \ldots, x_t$ to H and form a larger subgraph H' with $h + t$ vertices. If t is odd we can color the $t + 1$ edges of the path with $(t + 1)/2$ new colors. In the first half of the path the colors are all distinct, and the same ordering of colors is repeated in the second half of the path. It is straightforward to verify that H' is rainbow connected. If t is even, we can color the $t + 1$ edges of the path with $t/2$ colors as follows. The middle edge $(x_{t/2}, x_{t/2+1})$ receives any color that already appears in H. The first $t/2$ edges of the path all receive distinct new colors and in the last $t/2$ edges of the path this coloring is repeated in the same order. Again, it is straightforward to verify that H' is rainbow connected. We now have

$$rc(H') \leq rc(H) + \lceil t/2 \rceil \leq 2h/3 - 2/3 + \lceil t/2 \rceil \leq 2(h + t)/3 - 2/3$$

contradicting the maximality of H.

Having proved that $h \geq n - 2$ we now clearly have $rc(G) \leq 2(n-2)/3 - 2/3 + 2 = 2n/3$ as claimed. ■

Although the following asymptotic improvement of Proposition 2.2 does not yield a better bound for Theorem 1.1 it does yield the best possible coefficient of n if we look for an upper bound of the form $cn + o(n)$ on $rc(G)$ for 2-connected graphs.

Theorem 2.3 *If G is a 2-connected graph on n vertices, then $rc(G) \leq n/2 + O(\sqrt{n})$.*

Proof: A suitable edge coloring of G will be constructed sequentially, starting with a cycle on some number n_0 of vertices and making it rainbow connected with $\lceil n_0/2 \rceil \leq n_0/2 + 1/2$ colors. We assume that until the current stage of the procedure, a subgraph G' of G has been made rainbow connected, with n' vertices and $n'/2 + c$ colors.

We have seen in the proof of Proposition 2.2 that if a path P of length t has both endpoints in G' but is internally disjoint from G', then the t edges of P require as few as $\lfloor t/2 \rfloor$ new colors to make $G' \cup P$ rainbow connected. This step keeps the value of c unchanged if t is odd, and increases c with $1/2$ if t is even. Our goal is to prove that $c = O(\sqrt{n})$ holds when the entire G is edge-colored.

The key idea is to proceed with adding paths in *non-increasing* order of length. Let t denote the maximum path length that can currently be added, and let $P = a, x_1, \ldots, x_{t-1}, b$ be such a path, having its endpoints a, b in G' and its internal vertices x_i outside G'. Observe that no path of length t (or more) and being internally disjoint from $G' \cup P$ can join any x_i to any vertex of $G' \cup P$, otherwise t would not be maximum. Consequently, inserted paths of the same length are internally vertex-disjoint, and once we finish with t, we can never return to length t or more.

If t is odd, the value of c remains the same after path insertion. Suppose that t is even. If just one or two paths of length t can be added before we continue the procedure with length $t-1$, then, applying the coloring pattern described earlier, the value of c increases with $1/2$ or 1, respectively. If three or more paths of length t are added, we multicolor all those paths with the same t new colors (in the same order). In this way a rainbow-connected subgraph is obtained. For three or more paths we insert at least $3t-3$ internal vertices, but use just $t < (3t-3)/2$ colors if $t \geq 4$, what makes c decrease. If $t = 2$, c does not increase unless three paths are inserted, in which case c increases with $1/2$.

The worst case for counting this upper bound on $rc(G) - n/2$ is when each even path length $2, 4, \ldots, \ell$ occurs precisely two times (and no odd lengths). Even then, $rc(G) - n/2 \leq \ell/2$ holds. Since $\sum_{i=1}^{\ell/2}(2i-1) < n/2$ is valid for this particular sequence, we obtain $\ell = O(\sqrt{n})$ and the theorem follows. ■

Recall that a graph is *bridgeless* if the removal of an edge does not increase the number of connected components.

Proposition 2.4 *If G is a connected bridgeless graph with n vertices then $rc(G) \leq 4n/5 - 1$.*

Proof: Consider first the case where G is 2-connected. By Proposition 2.2, $rc(G) \leq \lfloor 2n/3 \rfloor \leq 4n/5 - 1$ for all $n \geq 7$. Since a 2-connected graph contains a cycle we have by Lemma 2.1 that $rc(G) \leq n-2$ and notice that $n - 2 \leq 4n/5 - 1$ for $n = 3, 4, 5$. The only remaining case is, therefore $n = 6$, and we need to show that $rc(G) \leq 3$ in this case. If the longest cycle in G has length 6 then $rc(G) \leq 3$. If the longest cycle has length 5 then G is the graph consisting of the 5-cycle and another vertex of degree 2 whose neighbors are two non-adjacent vertices of the cycle. It is straightforward to check that this graph has $rc(G) = 3$. If the longest cycle has length 4 then G is $K_{2,4}$ which has $rc(G) = 2$. A 2-connected graph with 6 vertices always has a cycle of length at least 4.

Having proved the proposition for 2-connected graphs, we prove it for connected bridgeless graphs by induction on the number of 2-connected components. Let X be the set of vertices of a 2-connected component of G so that X contains only one cut-vertex, say x (recall that such a 2-connected component always exists). Consider the subgraph H of G induced by $(V(G) \setminus X) \cup \{x\}$. It has $n - |X| + 1$ vertices and is connected, bridgeless, and with one less 2-connected component. By the induction hypothesis, $rc(H) \leq 4(n - |X| + 1)/5 - 1$. Since X induces a 2-connected graph, we have $rc(X) \leq 4|X|/5 - 1$. Hence, $rc(G) \leq 4n/5 - 2 + 4/5 < 4n/5 - 1$, completing the proof. ■

Proof of Theorem 1.1: We are given a connected graph $G = (V, E)$ with n vertices and minimum degree at least 3. Let $B \subset E$ denote the set of bridges of G. If $B = \emptyset$ we are done by Proposition 2.4. Thus, we may assume $B \neq \emptyset$. Let $\mathcal{C}$ denote the set of connected components of $G' = (V, E \setminus B)$. Notice that $\mathcal{C}$ contains at least two elements. There are two types of elements in $\mathcal{C}$, singletons and connected bridgeless subgraphs of G. Let $\mathcal{S} \subset \mathcal{C}$ denote the singletons and let $\mathcal{D} = \mathcal{C} \setminus \mathcal{S}$. Each element of $\mathcal{S}$ is, therefore, a vertex, and each element of $\mathcal{D}$ is a subset of vertices that induces a connected bridgeless subgraph.

We construct a rooted tree T whose nodes are the elements of $\mathcal{C}$ using the following standard recursive definition. The root of T is some arbitrarily chosen element of $\mathcal{C}$. The *children* of a node C are all the elements of $\mathcal{C}$ that are reachable from C via a single bridge, other than the parent of C in T (the root of T is the only node without a parent). The leaves of T are the nodes without children. Notice that since $\mathcal{C}$ contains at least two elements, so does T, and hence the root is not a leaf. For convenience, we order the children of a node C from left to right and denote by $\ell(C)$ the leftmost child and by $r(C)$ the rightmost child. It may be that $\ell(C) = r(C)$ if C has only one child, and for leaves we define $\ell(C) = r(C) = \emptyset$. Finally, we define $L(C)$ to be the leftmost leaf in the subtree of T rooted at C (if C is a leaf then $L(C) = C$).

Let $\mathcal{L} \subset \mathcal{C}$ denote the set of leaves of T. Notice that $\mathcal{L}$ is, in fact, a subset of $\mathcal{D}$ because each singleton in $\mathcal{S}$ is incident with at least three bridges, while the leaves of T are incident with only one bridge; the bridge connecting them to their parent in T. The same reasoning shows that for $s \in \mathcal{S}$, we have that s has at least two children in T and hence $\ell(s) \neq r(s)$. It now follows that $|\mathcal{S}| \leq |\mathcal{L}|$. Indeed, the mapping $s \to L(r(s))$ is one-to-one from $\mathcal{S}$ to $\mathcal{L}$.

Another important feature of the elements of $\mathcal{L}$ is that they each contain at least four vertices. Indeed, if $X \in \mathcal{L}$ then X is a connected bridgeless subgraph of G that is incident with only one bridge. Any vertex of X other than the one incident with that bridge has all of its neighbors in X, and as G has minimum degree at least 3, we have $|X| \geq 4$. We further partition $\mathcal{S}$ into two parts, $\mathcal{S}'$ are all those singletons s for which $L(r(s))$ has cardinality 4, and $\mathcal{S}''$ are all those singletons s for which $L(r(s))$ has cardinality at least 5. We now have:

$$|\mathcal{S}| + 4|\mathcal{S}'| + 5|\mathcal{S}''| = 5|\mathcal{S}'| + 6|\mathcal{S}''| \leq n. \tag{1}$$

We now color E making G rainbow connected. The edges of each element $X \in \mathcal{D}$ are colored using $4|X|/5 - 1$ dedicated colors. This can be achieved using Proposition 2.4. We also color the unique bridge connecting X to its parent in T (unless X is the root of T) using one dedicated color, altogether using $4|X|/5$ colors. Likewise, for each $s \in \mathcal{S}$ we color the unique bridge connecting s to its parent in T (unless X is the root of T) using one dedicated color. Notice that this process colors *all* the edges of G. Altogether we have used at most

$$|\mathcal{S}| - 1 + \frac{4}{5}\sum_{X \in \mathcal{D}} |X| = \frac{4}{5}n + \frac{1}{5}|\mathcal{S}| - 1$$

colors. There is, however, some spare in the estimate above. Recall that there are at least $|\mathcal{S}'|$ elements of $\mathcal{D}$ that have cardinality 4. Since 2-connected graphs with cardinality $|X| = 4$ can be rainbow connected using only two colors, we actually *do not* use $4|X|/5 - 1$ colors in this case; rather, we use only $4|X|/5 - 6/5$ colors, sparing an additional $1/5$ in the above calculation at least $|\mathcal{S}'|$ times. It follows that

$$rc(G) \leq \frac{4}{5}n + \frac{1}{5}|\mathcal{S}''| - 1.$$

By (1), $|\mathcal{S}''| \leq n/6$. It follows that $rc(G) < 5n/6$, as claimed. ∎

2.2 General minimum degree

Proof of Theorem 1.4: A set of vertices S of G is called a *connected* 2*-dominating set* if S induces a connected subgraph of G, and, furthermore, each vertex outside of S has at least two neighbors in S. Let $\gamma_{2c}(G)$ denote the smallest size of a connected 2-dominating set of G. Notice that the parameter is well-defined since, trivially, $V(G)$ is a connected 2-dominating set of G; hence $\gamma_{2c}(G) \le n$.

We claim that $rc(G) \le \gamma_{2c}(G)+1$. Indeed, let S be a connected 2-dominating set with $|S| = \gamma_{2c}(G)$. As S induces a connected subgraph, we have $rc(S) \le \gamma_{2c}(G) - 1$. Using two additional colors, say, red and blue, we can color the rest of the graph. For a vertex $v \in V \setminus S$, let (v, x) and (v, y) be two distinct edges with $x, y \in S$. We color (v, x) red and (v, y) blue, and do the same for each $v \in V \setminus S$. The resulting coloring clearly makes G rainbow connected.

A special case of the main result in [5], which was also implicitly proved earlier in [4] asserts that $\gamma_{2c}(G) \le n\frac{\ln\delta}{\delta}(1 + o_\delta(1))$. Together with the argument in the previous paragraph we have that $rc(G) \le n\frac{\ln\delta}{\delta}(1 + o_\delta(1))$, as required.

For the second, non-asymptotic part of the theorem, consider the following probabilistic argument (see, e.g. [1]). We create a subset of vertices X by randomly and independently choosing each vertex to X with probability p. Hence, the expected number of elements of X is $E[|X|] = np$. Let Y be the set of vertices not belonging to X and having no neighbor in X. The probability that $v \in Y$ is at most $(1-p)^{\delta+1}$. Hence, $E[|Y|] \le n(1-p)^{\delta+1}$. Let Z be the set of vertices not belonging to X and having precisely one neighbor in X. The probability that $v \in Z$ is at most $\delta p(1-p)^{\delta}$. Hence, $E[|Z|] \le n\delta p(1-p)^{\delta}$. Notice that $W = X \cup Y \cup Z$ is a 2-dominating set, although not necessarily a connected one. However, notice that the number of connected components in W is at most $|X| + |Y|$ (vertices of Z already have neighbors in X so they do not contribute additional connected components). Now, to make W connected, notice that it suffices to add just $2(|X| + |Y| - 1)$ additional vertices to W, forming a connected 2-dominating set S of G. The expected size of S is, henceforth,

$$E[|S|] \le 3np + 3n(1-p)^{\delta+1} + n\delta p(1-p)^{\delta} - 2.$$

Since $\gamma_{2c}(G) \le E[|S|]$ we have that

$$rc(G) \le 3np + 3n(1-p)^{\delta+1} + n\delta p(1-p)^{\delta}.$$

Choosing $p = \ln\delta/\delta$ we get that

$$rc(G) < \frac{n}{\delta}(4\ln\delta + 3).$$

■

As noted in the introduction, Theorem 1.4 cannot be improved below $\frac{3n}{\delta+1} - \frac{\delta+7}{\delta+1}$. To see this, we construct a connected n-vertex graph with minimum degree δ and diameter $\frac{3n}{\delta+1} - \frac{\delta+7}{\delta+1}$. Take m copies of $K_{\delta+1}$, denoted $X_1, \ldots, X_m$ and label the vertices of X_i

with $x_{i,1}, \ldots, x_{i,\delta+1}$. Take two copies of $K_{\delta+2}$, denoted X_0, X_{m+1} and similarly label their vertices. Now, connect $x_{i,2}$ with $x_{i+1,1}$ for $i = 0, \ldots, m$ with an edge, and delete the edges $(x_{i,1}, x_{i,2})$ for $i = 0, \ldots, m+1$. The obtained graph has $n = (m+2)(\delta+1)+2$ vertices, and minimum degree δ (and maximum degree $\delta+1$). It is straightforward to verify that a shortest path from $x_{0,1}$ to $x_{m+1,2}$ has length $3m+5 = \frac{3n}{\delta+1} - \frac{\delta+7}{\delta+1}$. If δ is odd, we can, in fact, make the graph regular. Just delete a maximum matching within X_0 where the only non-matched vertex is $x_{0,1}$ and delete a maximum matching within X_{m+1} where the only non-matched vertex is $x_{m+1,2}$.

2.3 Regular graphs

Proof of Theorem 1.3: We are given a connected graph G with n vertices, and a set of pairwise vertex-disjoint cycles $C_1, \ldots, C_t$ that cover $n-s$ vertices. Let c_i denote the length of C_i for $i = 1, \ldots, t$. As in Lemma 2.1 we notice that by adding to the set of cycles an additional set of $s+t-1$ edges we obtain a connected *spanning* subgraph of G that contains t cycles and $s+t-1$ bridges. We color the bridges using $s+t-1$ dedicated colors. Also, each even cycle of length c_i is colored using $c_i/2$ dedicated colors and each C_i which is a triangle is colored using one dedicated color. Now, let us arbitrarily pair the odd cycles of length at least 5. Without loss of generality assume that $C_1, \ldots, C_p$ are the odd cycles of length at least 5. We pair C_{2i-1} with C_{2i} for $i = 1, \ldots, \lfloor p/2 \rfloor$ (if p is odd then C_p remains unpaired, and we color it using $(c_p+1)/2$ dedicated colors). We will color these pairs of odd cycles using the following procedure. While there remains at least one pair that is yet uncolored, pick a pair (C_{2i-1}, C_{2i}) with the property that there is an already colored path between a vertex x of C_{2i-1} and a vertex y of C_{2i} (there must be at least one such pair). Let e be the unique edge of C_{2i-1} opposite to x (in an odd cycle each vertex has a unique edge opposite to it). Let f be the unique edge of C_{2i} opposite to y. Assign to e and f the same dedicated color, assign $(c_{2i-1}-1)/2$ dedicated colors to properly color the remaining $c_{2i-1}-1$ edges of C_{2i-1}, and assign $(c_{2i}-1)/2$ dedicated colors to properly color the remaining $c_{2i}-1$ edges of C_{2i}. Notice that we used $(c_{2i-1}+c_{2i})/2$ new colors to color this pair, and that after this coloring every connected component of the subgraph of colored edges is properly rainbow connected.

Altogether, if we let ℓ denote the number of triangles in our set of cycles we have that the number of colors used to rainbow-connect G is

$$\ell + \sum_{i=1}^{t-\ell} c_i/2 + s + t - 1/2 = \ell + \frac{1}{2}(n - s - 3\ell) + s + t - 1/2.$$

(If p is even then we even have 1 instead of 1/2 in both sides of the last equality.) Now, since $t \le \ell + (n-s-3\ell)/4$ we have, together with the last equality, that

$$rc(G) < 3n/4 + s/4 - 1/2$$

and the result follows. For the specific parts of the theorem notice that the first case follows from the fact that in a graph with a 2-factor we can assume $s = 0$ in the last inequality.

For the second and third cases we recall Petersen's Theorem [8], which states that if k is even, every k-regular graph is the union of $k/2$ 2-factors. Recalling the definition of the chromatic index, if a k-regular graph has $\chi'(G) = k$ then it is the union of k perfect matchings, and, in particular, the union of any two perfect matchings is a 2-factor. ■

2.4 Small graphs

We begin this subsection with the following proposition, which, although asymptotically inferior to the result of Theorem 1.1, is more useful for small graphs.

Proposition 2.5 *If G is a connected graph with minimum degree δ then $rc(G) \le n - \delta$.*

Proof: The proof is trivial in the case $\delta = 1$. Fixing δ, we prove the proposition by induction on n where the base case $n = \delta + 1$ is trivial since cliques have rainbow connection number 1. So, we assume $n > \delta + 1$.

Le K be a maximal clique of G consisting only of vertices whose degree is δ. Since there is at least one vertex with degree δ and since G is connected we we have $1 \le k = |K| \le \delta$.

Consider the graph G' obtained from G by deleting the vertices of K. Suppose the connected components of G' are $G_1, \ldots, G_t$ where G_i has n_i vertices and minimum degree δ_i for $i = 1, \ldots, t$. Let $K_i \subset K$ be the vertices of K with a neighbor in G_i, and assume that $|K_1| \ge |K_i|$ for $i = 2, \ldots, t$ (notice that it may be that $t = 1$ and G' is connected). Consider first the case where $K_1 = K$. By the induction hypothesis, $rc(G_i) \le n_i - \delta_i$. Clearly, we may give the edges of K and the edges from K to G_1 the same color. Hence,

$$rc(G) \le t + \sum_{i=1}^{t}(n_i - \delta_i) = t + n - k - \sum_{i=1}^{t}\delta_i.$$

By the maximality of K, each vertex of G_i has degree at least $\delta - k + 1$ in G_i. Hence, $\delta_i \ge \delta - k + 1$, and therefore

$$rc(G) \le t + n - k - t\delta + tk - t \le n - t\delta + (t-1)k \le n - \delta.$$

Now assume that $K_1 \subsetneq K$ but that $|K_1| = k_1 > 1$. By contracting K_1 to a single vertex v, we obtain a contraction G^* of G with $n - k_1 + 1$ vertices and minimum degree $\delta - k_1 + 1$, so by induction $rc(G^*) \le n - \delta$. Now, going back to G, any edge with both endpoints not in K_1 receives the same color it had in G^*. Any edge with one endpoint in K_1 receives the color of the edge of G^* from v to that other endpoint. Any edge with both endpoints in K_1 receives the color of an edge of G^* from v to another vertex in $K \setminus K_1$. The resulting coloring makes G rainbow connected and $rc(G) \le n - \delta$.

Finally, if $k_1 = 1$ (and since $K_1 \subsetneq K$ we have $k \ge 2$), contract all of K into a single vertex v and notice that the contracted graph G^* also has minimum degree δ, and $n-k+1$ vertices. Hence, by induction, $rc(G^*) \le n - k - \delta + 1$. Going back to G and coloring the edges of the clique K with another new color, we obtain $rc(G) \le n - k - \delta + 2 \le n - \delta$. ■

Corollary 2.6 *If G is a connected graph with n vertices and average degree $d \geq 2$ then $rc(G) \leq n - d/2 - 1$.*

Proof: It is well known (see, e.g., [2]) that a graph G with average degree $d \geq 2$ has a connected subgraph H with minimum degree at least $d/2 + 1$. By proposition 2.5, $rc(H) \leq |H| - d/2 - 1$. By Lemma 2.1, $rc(G) \leq n - |H| + rc(H) \leq n - d/2 - 1$. ■

3 Ensuring 2-rainbow-connection

3.1 Dense graphs

Proof of Theorem 1.5: Given a graph G with n vertices and $\delta(G) \geq n/2 + \log n$, we randomly color the edges with two colors, red and blue. We show that with positive probability, such a random coloring makes G rainbow connected. Consider two non-adjacent vertices x, y. The minimum degree requirement forces x and y to have more than $2 \log n$ common neighbors. Now, for each such common neighbor z, the probability that the path x, z, y is not a rainbow path is precisely $1/2$. Since the paths corresponding to distinct common neighbors are edge-disjoint, the probability that *all* these paths are not rainbow is less than $(1/2)^{2 \log n} = 1/n^2$. Since there are less than $\binom{n}{2}$ pairs x, y to consider, it follows from the union bound that with positive probability, each pair of non-adjacent vertices are connected by a rainbow path. ■

A bipartite graph which is not complete has diameter at least 3. A proof similar to the proof of Theorem 1.5 gives the following result.

Theorem 3.1 *Let $c = 1/\log(9/7)$. If G is a non-complete bipartite graph with n vertices and any two vertices in the same vertex class have at least $2c \log n$ common neighbors in the other vertex class, then $rc(G) = 3$.*

Proof: The only modification needed is that now the edges are colored with three colors. For two vertices with distance 2 the probability that a common neighbor is the center of a non-rainbow path is only $1/3$. For two vertices with distance 3 (clearly the diameter is at most 3), say, x, y, fixing a neighbor z of x, and the color c of (x, z), the probability that a common neighbor of y and z is the center of a non-rainbow path, or a path that contains c, is only $7/9$. Since $(7/9)^{2c \log n} = 1/n^2$ the result follows as before. ■

3.2 Random graphs

Proof of Theorem 1.6: For the first part of the theorem, we need to prove that for a sufficiently large constant C, the random graph $G(n, p)$ with $p = C\sqrt{\log n/n}$ almost surely has $rc(G(n, p)) = 2$. By the proof of theorem 1.5, it suffices to show that almost surely any two vertices of $G(n, p)$ have at least $2 \log n$ common neighbors. By the union bound, it suffices to prove that a fixed pair of vertices does not have at least $2\log n$ common neighbors with probability $o(1/n^2)$. Fixing a pair x, y, the probability of the event that z is a common neighbor is $C^2 \log n/n$. Thus, the expected number of common neighbors is $\frac{n-2}{n}(C^2 \log n) > 0.5c^2 \log n$, and this is the sum of $n-2$ indicator random variables of *independent* events. By the standard Chernoff estimates (cf. [1] Theorem A.1.13), the probability that the number of common neighbors is less than half of its expectation (in this case, less than $0.25c^2 \log n$) is $o(1/n^2)$ for C sufficiently large.

For the other direction, it suffices to prove that for a sufficiently small constant c, the random graph $G(n,p)$ with $p = c\sqrt{\log n/n}$ almost surely does not have diameter two. Fix a set X of $n^{1/5}$ vertices (we may and will assume that $n^{1/5}$ is an even integer), and let Y be the remaining $n - n^{1/5}$ vertices. Let $\mathcal{A}$ be the event that X induces an independent set, and let $\mathcal{B}$ be the event that there exists a pair of vertices of X with no common neighbor in Y. Clearly, if both $\mathcal{A}$ and $\mathcal{B}$ occur then the diameter is at least 3. Hence, it suffices to prove (separately) that for a small c, $\mathcal{A}$ occurs with probability approaching 1 and $\mathcal{B}$ occurs with probability approaching 1.

We start with $\mathcal{A}$ first. For c sufficiently small we indeed have:

$$\Pr[\mathcal{A}] = (1-p)^{\binom{n^{1/5}}{2}} = (1 - c\sqrt{\log n/n})^{\binom{n^{1/5}}{2}} = 1 - o_n(1).$$

To estimate $\mathcal{B}$, let us first partition X into $n^{1/5}/2$ pairs, arbitrarily. For a pair x, y, the probability that x, y have a common neighbor in Y is precisely

$$1 - \left(1 - \frac{c^2 \log n}{n}\right)^{n-n^{1/5}}.$$

Thus, for c sufficiently small, the probability that all $|X|/2$ pairs have a common neighbor is

$$\left(1 - \left(1 - \frac{c^2 \log n}{n}\right)^{n-n^{1/5}}\right)^{n^{1/5}/2} = o_n(1).$$

We use here the fact that the event that a pair x, y has a common neighbor in Y is *independent* of all the events that other pairs have common neighbors in Y. Since the last inequality is $o_n(1)$ we have that with probability $1 - o_n(1)$ at least one of the pairs does not have a common neighbor in Y, and, in particular, $\mathcal{B}$ holds. ■

4 Concluding remarks and open problems

- The computational status of rainbow connection is not yet determined. In fact, we do not even know whether deciding if $rc(G) = 2$ is polynomial or is NP-complete. Recall that it is straightforward to determine the diameter of a graph in polynomial time. The main difficulty, however, lies in the fact that even if we know that $diam(G) = 2$ (and, in particular, know that any two non-adjacent vertices have at least one common neighbor), we still cannot even use this fact to bound the rainbow connection number. We do conjecture the following:

Conjecture 4.1 *The following decision problem is NP-Complete: "Given an input graph G, determine if $rc(G) = 2$".*

Notice that the problem trivially belongs to NP. More general, and therefore perhaps easier, is the following conjecture:

Conjecture 4.2 *Computing $rc(G)$ is NP-hard.*

- Theorem 1.5 asserts that minimum degree $n/2+\log n$ guarantees $rc(G) = 2$. Clearly, minimum degree $n/2 - 1$ does not, as there are connected graphs with minimum degree $n/2 - 1$ and diameter 3 (just take two vertex-disjoint cliques of order $n/2$ each and connect them by a single edge. It is therefore interesting to raise:

  Problem 4.3 *Determine the minimum degree threshold that guarantees $rc(G) = 2$.*

- As noted in Section 2, finding the optimal bound for $rc(G)$ given that a graph has minimum degree k seems to be an intriguing open problem.

References

[1] N. Alon and J. H. Spencer, **The Probabilistic Method**, *Second Edition*, Wiley, New York, 2000.

[2] B. Bollobás, **Extremal Graph Theory**, Academic Press, London, 1978.

[3] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica 7 (1986), 35–38.

[4] Y. Caro, D. West, and R. Yuster, *Connected domination and spanning trees with many leaves*, Siam J. Disc. Math. 13 (2000), 202–211.

[5] Y. Caro and R. Yuster, *Dominating a family of graphs with small connected subgraphs*, Combin. Probab. Comput. 9 (2000), 309–313.

[6] G. Chartrand, G. L. Johns, K. A. McKeon, and P. Zhang, *Rainbow connection in graphs*, Math. Bohem., to appear.

[7] E. Friedgut and G. Kalai, *Every monotone graph property has a sharp threshold*, Proc. Amer. Math. Soc. 124 (1996), 2993–3002 .

[8] J. Petersen, *Die theorie der regularen graphen*, Acta Math. 15 (1891), 193–220.