Rainbow matchings and connectedness of coloured graphs

Alexey Pokrovskiy

Methods for Discrete Structures,
Freie Universität,
Berlin, Germany.
Email: alja123@gmail.com

May 30, 2015

Abstract

Aharoni and Berger conjectured that every bipartite graph which is the union of n matchings of size $n+1$ contains a rainbow matching of size n. This conjecture is a generalization of several old conjectures of Ryser, Brualdi, and Stein about transversals in Latin squares. When the matchings are all edge-disjoint and perfect, an approximate version of this conjecture follows from a theorem of Häggkvist and Johansson which implies the conjecture when the matchings have size at least $n + o(n)$.

Here we'll discuss a proof of this conjecture in the case when the matchings have size $n + o(n)$ and are all edge-disjoint (but not necessarily perfect). The proof involves studying connectedness in coloured, directed graphs. The notion of connectedness that we introduce is new, and perhaps of independent interest.

1 Introduction

A Latin square of order n is an $n \times n$ array filled with n different symbols, where no symbol appears in the same row or column more than once. Latin squares arise in many branches of mathematics such as algebra (where Latin squares are exactly the multiplication tables of quasigroups), experimental design (where some row-column designs come from Latin squares), and coding theory (where some error-correcting codes are constructed from mutually orthogonal Latin squares). They also occur in recreational mathematics—for example completed Sudoku puzzles are Latin squares.

In this paper we will look for *transversals* in Latin squares—a transversal in a Latin square of order n is a set of n entries such that no two entries are in the same row, same column, or have the same symbol. One reason transversals in Latin squares are interesting is that a Latin square has an orthogonal mate if, and only if, it has a decomposition into disjoint transversals. See [12] for a survey about transversals in Latin squares. It is easy to see that not every Latin square has a transversal (for example the unique 2×2 Latin square has no transversal), however perhaps every Latin square contains a large *partial transversal* (a partial transversal of size m is a set of m entries such that no two entries are in the same row, same column, or have the same symbol)?

There are some old and difficult conjectures which guarantee large partial transversals in Latin squares. One is a conjecture of Ryser that every Latin square of odd order contains a transversal [10]. Brualdi [4] and Stein [11] independently made the following conjecture.

Conjecture 1.1 (Brualdi and Stein, [4, 11])**.** *Every Latin square contains a partial transversal of size $n-1$.*

There have been many partial results about this conjecture. It is known that every Latin square has a partial transversal of size $n - o(n)$—Woolbright [13] and independently Brower, de

Vries, and Wieringa [3] proved that ever Latin square contains a partial transversal of size $n-\sqrt{n}$. This has been improved by Hatami and Schor [7] to $n-C\log^2 n$. A remarkable result of Häggkvist and Johansson shows that if we consider $(1-\epsilon)n \times n$ Latin rectangles rather than Latin squares, then it is possible to decompose all the entries into disjoint transversals (for $m \leq n$ a $m \times n$ Latin rectangle is an $m \times n$ array of n symbols where no symbol appears in the same row or column more than once. A transversal in a Latin rectangle is a set of m entries no two of which are in the same row, column, or have the same symbol).

Theorem 1.2 (Häggkvist and Johansson, [6]). *For every ϵ, there is an $m_0 = m_0(\epsilon)$ such that the following holds. For every $n \geq (1+\epsilon)m \geq m_0$, every $m \times n$ Latin rectangle can be decomposed into disjoint transversals.*

This theorem is proved by a probabilistic argument, using a "random greedy process" to construct the transversals. The above theorem gives yet another proof that every sufficiently large $n \times n$ Latin square has a partial transversal of size $n - o(n)$—indeed if we remove ϵn rows of a Latin square we obtain a Latin rectangle to which Theorem 1.2 can be applied.

In this paper we will look at a strengthening of Conjecture 1.1. The strengthening we'll look at is a conjecture due to Aharoni and Berger which takes place in a more general setting than Latin squares—namely coloured bipartite graphs. To see how the two settings are related, notice that there is a one-to-one correspondence between $n \times n$ Latin squares and proper edge-colourings of $K_{n,n}$ with n colours—indeed to a Latin square S we associate the colouring of $K_{n,n}$ with vertex set $\{x_1, \dots, x_n, y_1, \dots, y_n\}$ where for every i, j the edge between x_i and y_j receives colour $S_{i,j}$. It is easy to see that in this setting transversals in S correspond to perfect rainbow matchings in $K_{n,n}$ (a matching is *rainbow* if all its edges have different colours). Thus Conjecture 1.1 is equivalent to the statement that "in any proper n-edge-colouring of $K_{n,n}$, there is a rainbow matching of size $n-1$".

One could ask whether a large rainbow matching exists in more general bipartite graphs. Aharoni and Berger posed the following conjecture, which generalises Conjecture 1.1.

Conjecture 1.3 (Aharoni and Berger, [1]). *Let G be a bipartite graph consisting of n matchings, each with at least $n+1$ edges. Then G contains a rainbow matching with n edges.*

In the above conjecture we think of the n matchings forming G as having different colours, and so "rainbow matching" means a matching containing one edge from each matching in G. It is worth noting that the above conjecture does not require the matchings in G to be disjoint i.e. it is about bipartite multigraphs rather than simple graphs. This above conjecture was posed in a different form in [1] as a conjecture about matchings in tripartite hypergraphs (Conjecture 2.4 in [1]). It was first stated as a conjecture about rainbow matchings in [2].

The above conjecture has attracted a lot of attention recently, and there are many partial results. One very natural approach to Conjecture 1.3 is to prove it when the matchings have size much larger than $n+1$. When the matchings have size $2n$ then it is easy to see that the conclusion of the conjecture is true (by greedily choosing disjoint edges one at a time). Aharoni, Charbit, and Howard [2] proved that matchings of size $7n/4$ are sufficient to guarantee a rainbow matching of size n. Kotlar and Ziv [8] improved this to $5n/3$. Clemens and Ehrenmüller [5] further improved this to $3n/2 + o(n)$ which is currently the best known bound.

An approximate version of Conjecture 1.3 can be obtained from Theorem 1.2. It is easy to see that Theorem 1.2 is equivalent to the following "let G be a bipartite graph consisting of n edge-disjoint perfect matchings, each with at least $n + o(n)$ edges. Then G can be decomposed into disjoint rainbow matchings of size n" (to see that this is equivalent to Theorem 1.2, associate an m-edge-coloured bipartite graph with vertex set $\{x_1, \dots, x_n, y_1, \dots, y_n\}$ with any $m \times n$ Latin rectangle by placing a colour k edge between x_i and y_j whenever (k, i) has symbol j in the rectangle).

The result that we'll discuss is an approximate version of Conjecture 1.3 in the case when the matchings in G are disjoint, but not necessarily perfect.

Theorem 1.4. *For all $\epsilon_0 > 0$, there exists an $N_0 = N_0(\epsilon_0)$ such that the following holds. Let G be a bipartite graph consisting on $n \geq N_0$ edge-disjoint matchings, each with at least $(1+\epsilon_0)n$ edges. Then G contains a rainbow matching with n edges.*

Unlike the proof of Theorem 1.2 which can be used to give a randomised process to find a rainbow matching, the proof of Theorem 1.4 is algorithmic i.e. the matching in Theorem 1.4 can be found in polynomial time.

The proof of Theorem 1.4 will appear in [9]. For the remainder of this extended abstract we will sketch some of the ideas which go into the proof. At a very high level the proof of Theorem 1.4 consists of associating an auxiliary directed graph D to the graph G, such that directed paths in D give some kind of information about rainbow matchings in G. Then we apply results about connectedness in directed graphs in order to prove the theorem. In Section 1.1 we sketch how the directed graph D is constructed. In Section 1.2 we state the result about directed which is used in the proof of Theorem 1.4.

1.1 From bipartite graphs to directed graphs

Let G be a graph consisting of n disjoint matchings each of size $(1+\epsilon_0)n$ as in the statement of Theorem 1.4. Let X and Y be the two parts of the bipartition of G. Let M be a rainbow matching in G. Let $X_0 = X \setminus V(M)$ be the subset of X consisting of vertices not touched by M. Let c^* be a colour missing from M.

We construct an edge-labelled directed graph D as follows. The vertex set of D will be the set of colours in G. Every edge of D is labelled by a vertex $x \in X_0$. For two colours $u, v \in V(D)$ we set uv to be an edge of D labelled by $x \in X_0$ whenever there is a colour u edge from x to the colour v edge of M in G.

Why might this graph be useful? It turns out that if there is a vertex of small out-degree in D which is close to c^* then we can find a larger rainbow matching in G.

Lemma 1.5. *Let $P = (c^*, p_1, p_2, \ldots, p_k)$ be a directed rainbow path from c^* to some $p_k \in V(D)$. If we have $d^+(v) < \epsilon_0 n - |P|$, then there is a rainbow matching of size $|M|+1$ in G.*

Proof. Let $e_1, e_2, \ldots, e_k$ be the edges of G corresponding to the edges c^*p_1, $p_1p_2, \ldots, p_{k-1}p_k$ of D. For $i = 1, \ldots, k$, let m_i be the colour p_i edge of M. From the definition of D, we have that e_i and m_i intersect in Y, and that e_{i+1} and m_i have the same colour. Notice that for distinct i and j, the edges e_i and e_j are disjoint (since P is rainbow) as are m_i and m_j (since P is a path). Therefore $M' = M - m_1 - \cdots - m_k + e_1 + \cdots + e_k$ is a rainbow matching of size $|M|$ in G, missing colour p_k.

Notice that since there are $(1+\epsilon_0)n$ colour p_k edges in G, there must be at least $\epsilon_0 n$ colour p_k edges touching X_0. Each of these gives rise to an edge leaving p_k unless it goes through Y_0. Therefore if $d^+(v) < \epsilon_0 n - |P|$, then there are at least $|P|+1$ colour p_k edges between X_0 and Y_0. One of these must be disjoint from $e_1, \ldots, e_k$ and so can be added to M' to give a matching of size $|M|+1$. □

Therefore rainbow paths in D can give useful information about rainbow matchings in G. In the full proof of Theorem 1.4 we use more complicated directed graphs than the one constructed above. We also use a result about coloured directed graphs which we discuss in the next section.

1.2 Rainbow connectedness

The key idea in the proof of Theorem 1.4 seems to be a new notion of connectedness of coloured graphs.

Definition 1.6. *An edge-coloured graph G is said to be strongly rainbow k-connected if for any set of at most k colours S and any pair of vertices u and v, there is a rainbow u to v path whose edges have no colours from S.*

The above definition differs from usual notions of connectedness, since generally the avoided set S is a set of *edges* rather than colours. In some ways Definition 1.6 is perhaps *too strong*. In particular, there doesn't seem to be a natural analogue of Menger's Theorem for strongly rainbow k-edge-connected graphs. Nevertheless, strongly rainbow k-connected graphs turn out to be very useful for studying rainbow matchings in bipartite graphs. The following lemma is a key part of the proof of Theorem 1.4. It shows that every properly coloured directed graph D with big out-degree has a large, highly connected subset.

Lemma 1.7. *For all $\epsilon > 0$ and $k \in \mathbb{N}$, there is an $N = N(\epsilon, k)$ such that the following holds.*

Let D be a properly edge-coloured directed graph on at least N vertices. Then there is a strongly rainbow k-connected subset $A \subseteq V(D)$ satisfying

$$|A| \geq \delta^{+}(D) - \epsilon|D|.$$

Here "strongly rainbow k-connected *subset*" means a set of vertices $A \subseteq V(D)$ such that for any set of at most k colours S and any pair of vertices u and v, there is a rainbow u to v path in D whose edges have no colours from S. This is different from just saying that $D[A]$ is strongly rainbow k-connected because the paths connecting u and v are allowed to leave A. The above lemma is proved in [9].

References

[1] R. Aharoni and E. Berger. Rainbow matchings in r-partite r-graphs. *Electron. J. Combin.*, 16, 2009.

[2] R. Aharoni, P. Charbit, and D. Howard. On a generalization of the Ryser-Brualdi-Stein conjecture. *J. Graph Theory*, 78:143–156, 2015.

[3] A. E. Brouwer, A. J. de Vries, and R. M. A. Wieringa. A lower bound for the length of partial transversals in a Latin square. *Nieuw Archief Voor Wiskunde*, 26:330–332, 1978.

[4] R. A. Brualdi and H. J. Ryser. *Combinatorial matrix theory.* Cambridge University Press, 1991.

[5] D. Clemens and J. Ehrenmüller. An improved bound on the sizes of matchings guaranteeing a rainbow matching. *arXiv:1503.00438*, 2015.

[6] R. Häggkvist and A. Johansson. Orthogonal Latin rectangles. *Combin. Probab. Comput.*, 17:519–536, 2008.

[7] P. Hatami and P. W. Shor. A lower bound for the length of a partial transversal in a Latin square. *J. Combin. Theory Ser. A*, 115:1103–1113, 2008.

[8] D. Kotlar and R. Ziv. Large matchings in bipartite graphs have a rainbow matching. *European J. Combin.*, 38:97–101, 2014.

[9] A. Pokrovskiy. Rainbow matchings and rainbow connectedness. *preprint*, 2015.

[10] H. J. Ryser. Neuere probleme der kombinatorik. *Vorträge über Kombinatorik, Oberwolfach*, pages 69–91, 1967.

[11] S. K. Stein. Transversals of Latin squares and their generalizations. *Pacific J. Math.*, 59:567–575, 1975.

[12] I. M. Wanless. Transversals in Latin squares: A survey. In R. Chapman, editor, *Surveys in Combinatorics 2011.* Cambridge University Press, 2011.

[13] D. E. Woolbright. An $n \times n$ latin square has a transversal with at least $n - \sqrt{n}$ distinct symbols. *J. Combin. Theory Ser. A*, 24:235–237, 1978.