Edge-Coloring Cliques with Three Colors on All 4-Cliques

Nabiha Asghar
Department of Combinatorics & Optimization
University of Waterloo, Ontario, Canada
nasghar@uwaterloo.ca

1. Objectives and Layout

This discourse aims to provide an in-depth summary of one of the early works (published in 1998) by Dhruv Mubayi on Ramsey Theory. His article answers a question posed by P. Erdős 17 years earlier.

The article in question is highly concise and consists of just the proof of one theorem. In this paper, we aim to put the article in context by providing a background and a summary of the later works on the same topic. We also endeavor to improve its presentation and exposition by dividing the theorem into a lemma and a theorem, and explaining the proofs in greater detail by filling in the gaps that were (intentionally or otherwise) left by the author. In terms of technical rigor, a high premium is placed on the proof of the lemma and a major portion of the proof of theorem.

The layout of this paper is as follows: Section 2 gives a background of how and when the problem was posed, and gives a summary of the preliminary results. Section 3 defines the problem by stating the lemma and the theorem to be proved. Section 4 provides a detailed proof of the lemma, and section 5 shows how the theorem follows easily from the lemma. Section 6 gives a brief overview of later works on the topic.

2. Background and Motivation

Ramsey's Theorem states that in any coloring of the edges of a sufficiently large complete graph, one will find monochromatic complete graphs. In formal terms: for positive integers p and k, there exists a positive integer n so that every k-edge-coloring of K_n has a monochromatic K_p (n depends on p and k in general). Many different proofs of this statement exist, of varying elegance. The Classical Ramsey Problem (commonly known as the Party Problem) takes this problem one step ahead and asks for the minimum number n for which the statement holds:

Classical Ramsey Problem: Find the minimum value of n such that every k-coloring of the edges of K_n yields a monochromatic K_p.

This threshold value of n still remains unknown, but one can deduce that for every value of n below this threshold, there exists a k-edge-coloring in which every K_p receives at-least two colors. In 1981, P. Erdős generated an interesting variation of this problem in [1] by fixing *n* and varying *k*, and asked the following question:

For some fixed n, p and q, find $f(n,p,q)$, the minimum number of colors that are needed to color the edges of a given K_n such that every K_p receives atleast q colors

[The original statement of the problem in [1] was slightly different but equivalent].

**Note: From here on, the term 'coloring' denotes edge-coloring, and the term 'k-coloring' denotes edge-coloring a graph with k colors.*

Erdős gave a few interesting preliminary results, for small values of p and q, to see if any patterns emerged. Two of them were:

- $f(n,5,3) \leq 1 + \frac{\log n}{\log 2}$ (i.e. K_n can be colored with $1 + \frac{\log n}{\log 2}$ colors such that every K_5 receives at-least 3 colors)
- $f\ (n,4,3) = O(n)$.

It was found that determining the value of $f(n,p,q)$ for small values of p, q led to problems of varying difficulty. For example, for $q = 2$, the problem is equivalent to determining the classical Ramsey number for multi-colorings (as explained above), which is an open and hard problem. On the other hand, $f(n,3,3)$ is equivalent to giving a proper edge-coloring to K_n, which is exactly equal to the chromatic number of K_n (it is equal to n when n is odd, and $n-1$ when n is even). So we see that $f(n,3,2)$ is hard to compute, while computing $f(n,3,3)$ is very easy.

So it was seen that the bounds on the results for small p, q were not good enough to shed light on the solution of the general problem. For this reason, research in this area was abandoned for 15 years, before Erdős and Gyárfás made another attempt in [2] at improving the bounds for small values of p and q. They used the Local Lemma technique, proposed in [3] to show that

$f(n,p,q) \leq cn^{\frac{p-2}{1-q+\binom{p}{2}}}$ (this is a probabilistic bound). In particular, they improved the result of [1] and showed that that $f\ (n,4,3) = O\left(\sqrt{n}\right)$. They also determined, for each p, the smallest q such that $f(n,p,q)$ is linear in n and the smallest q such that $f\ (n,p,q)$ is quadratic in n. However, they were unable to solve many of the small cases, notably $f(n,4,3)$, which they termed as *'the most annoying problem among all the small cases'*.

This is where Dhruv Mubayi entered the arena. In the paper under discussion, Mubayi tackles this *annoying* small case of $f(n,4,3)$ and significantly improves its bound from a probabilistic $O(\sqrt{n})$ to a deterministic $e^{O(\sqrt{\log n})}$ as n grows large.

3. Definition of the Problem

Definition: For integers $n, p, q,$ a (p,q)-coloring of K_n is an edge-coloring of K_n in which every p-clique (i.e. K_p) receives at-least q colors. Let $f(n,p,q)$ denote the minimum number of colors in a (p,q)-coloring of K_n.

Mubayi presents the proofs of the following lemma and theorem:

Lemma: Let G be a complete graph on $\binom{m}{t}$ vertices, for $m, t \in \mathbb{Z}^+, t < m$. Then there exists a coloring of G that uses at most $(2^t - 1)(m-1)$ colors.

Theorem: $f(n,4,3) < e^{\sqrt{c \log n}\,(1+o(1))}, where\ c = 4\log 2.$

Note: The probabilistic bound $f(n,p,q) = O(\sqrt{n})$ by Erdős used a random (4,3)-coloring of K_n. Mubayi shows that the optimal (4,3)-coloring of K_n uses much fewer colors, and he shows how to construct this optimal coloring.

4. The Lemma

In this section, we give an explicit way of coloring a complete graph G, such that at most $(2^t - 1)(m-1)$ colors are used. First we give some definitions, followed by the method of construction and coloring of graph. This is followed by the actual proof of the lemma, which consists of two parts.

Definitions

1. $[n]$ denotes the set $\{1, 2, \ldots\ldots, n\}$.
2. The symmetric difference of sets A and B is $A \vartriangle B = (A - B) \cup (B - A)$.
3. For integers $t < m$, $\binom{[m]}{t}$ denotes the family of all t-sized-subsets (called t-subsets) of $[m]$.

Construction and coloring of the graph

Let G be the complete graph on $\binom{m}{t}$ vertices. Let $V(G) = \binom{[m]}{t}$ and for each t-subset T of $[m]$, rank the $2^t - 1$ proper subsets of T by some linear order. (One possible way to do this is: for each T, map its set of proper subsets to the set of first $2^t - 1$ Natural numbers.) For any two vertices A and B in G (each of which is a t-subset of m), let the color of edge AB be a two-dimensional vector

$$c(AB) = (c_0(AB), c_1(AB))$$

where

$$c_0(AB) = min\{i: i \in A \triangle B\}.$$

Set

$$S = \begin{cases} A & if\ c_0(AB) \in A \\ B & if\ c_0(AB) \in B \end{cases}$$

Let $c_1(AB)$ be the rank of $A \cap B$ in the linear order associated with the proper subsets of S.

<u>Claim 1:</u> This coloring is a (4,3)-coloring of G.

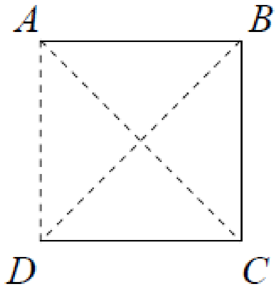
<u>Claim 2:</u> In this construction, the number of colors used is at most $(2^t - 1)(m - 1)$.

<u>Proof of Claim 1:</u>

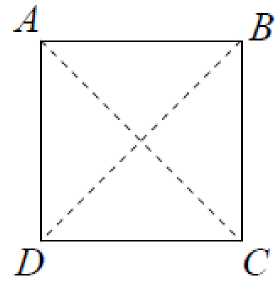
We first show that there cannot be monochromatic triangles in G. We then proceed to show that there can only be two types of 2-colored K_4's possible in G, and that G does not contain either of them. This proves that we have a (4,3)-coloring on G.

As a first step, we observe that there are no monochromatic triangles in G. Otherwise, if ABC was one such triangle, where $A, B, C \in V(G)$, then let $c_0(AB) = i$. Without loss of generality, suppose $i \in A$. Then $i \notin B$ (otherwise $i \in A \cap B$, hence $i \notin A \triangle B$, a contradiction to the definition of $c_0(AB)$). Since $c(AB) = c(BC)$, we have that $c_0(AB) = c_0(BC) = i$. Also, $i \notin B$ means $i \in C$. Now $i \in A$ and $i \in C$, which implies $i \notin A \triangle C$, so $c_0(AC) \neq i$. Hence $c_0(AC) \neq c_0(AB)$, which contradicts the fact that ABC is a monochromatic triangle. Hence there are no monochromatic triangles in G.

Since monochromatic triangles are forbidden, the only types of 2-colored K_4's that can occur are those shown in the figure below:



Type 1



Type 2

We show by proofs of contradiction that both types of K_4's cannot occur in G.

Type 1: Here one color class is the path $ABCD$, and the other is the path $BDAC$. Suppose $c_0(AB) = i$. Then there are two possible cases:

Case 1: $i \in A$. Then by the argument made above, $i \notin B$. This implies that $i \in C$, and in turn $i \notin D$.

We now show that $A \cap [i-1] = B \cap [i-1]$.

Suppose $g \in A \cap [i-1]$. Since i is the smallest element $A \triangle B$, $g \in [i-1]$ implies $g \notin A \triangle B$. Now $g \in A$ and $g \notin A \triangle B$, so $g \in A \cap B$. Therefore $g \in [i-1]$ and $g \in B$ imply that $g \in B \cap [i-1]$. Hence $A \cap [i-1] \subseteq B \cap [i-1]$.

Now let $g \in B \cap [i-1]$. Since i is the smallest element $A \triangle B$, $g \in [i-1]$ implies $g \notin A \triangle B$. Now $g \in B$ and $g \notin A \triangle B$, so $g \in A \cap B$. Therefore $g \in [i-1]$ and $g \in A$ imply that $g \in A \cap [i-1]$. Hence $B \cap [i-1] \subseteq A \cap [i-1]$. So we have shown that

$$A \cap [i-1] = B \cap [i-1].$$

Using the same argument, since $c(AB) = c(BC)$, $i \notin B$ and $i \in C$, we can deduce that

$$B \cap [i-1] = C \cap [i-1].$$

Finally, $c(BC) = c(CD)$ and $i \in C$ and $i \notin D$ imply that

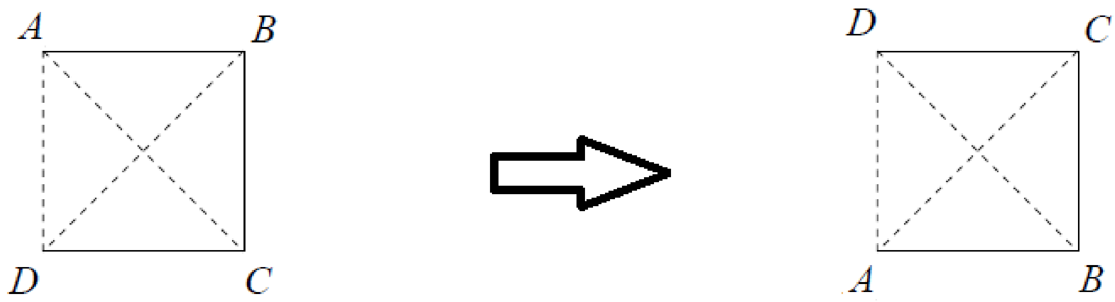
$$C \cap [i-1] = D \cap [i-1].$$

Therefore,

$$A \cap [i-1] = B \cap [i-1] = C \cap [i-1] = D \cap [i-1] \tag{1}$$

We have that $i \in A$ and $i \in C$, so (1) implies $A \cap [i] = C \cap [i]$. So $[i] \subseteq A \cap C$, hence $[i] \cap (A \triangle C) = \phi$. Thus the minimum element in $(A \triangle C)$ must be greater than i, so $c_0(AC) > i$. On the other hand, we know that $i \in A$ and $i \notin D$, and by (1) we have $[i-1] \subseteq A \cap D$, so i is the minimum element in $A \triangle D$. Hence $c_0(AD) = i$. So $c_0(AC) > i = c_0(AD)$. Hence $c(AC) \neq c(AD)$. This contradicts the fact that the path $BDAC$ is monochromatic.

<u>Case 2:</u> $i \in B$. Then by the argument made above, $i \in D$ and $i \notin A, C$. We reverse the labels on the path $ABCD$:



Type 1 Type 1

Then we get $i \in A$ and $i \notin B, D$, which puts us back in case 1.

Therefore, a 2-colored K_4 of Type 1 cannot occur in G.

<u>Type 2:</u>

Here one color class is the 4-cycle $ABCD$, while the other contains the edges AC and BD. By symmetry, we may assume that $c_0(AB) \in A - B$ (if $c_0(AB) \in B - A$, then just reverse the labels of A and B). Similarly, we may assume that $c_0(CD) \in C - D$. Since $c_0(AB) = c_0(CD)$, $c_0(AB) \in C - D$. Thus $c_0(AB) \in (A - B) \cap (C - D)$.

Since $A - B = A \cap \bar{B}$, we have

$$(A - B) \cap (C - D) = (A \cap \bar{B}) \cap (C \cap \bar{D})$$

$$= (A \cap C) \cap (\bar{B} \cap \bar{D})$$

$$= (A \cap C) \cap (\overline{B \cup D}) \text{ (by DeMorgan's Law)}$$

$$= (A \cap C) - (B \cup D) \text{ (by definition of set subtraction)}$$

Hence $c_0(AD) = c_0(AB) \in (A \cap C) - (B \cup D)$. This implies that

1. $c_0(AB)$ lies in A and does not lie in B, therefore $S = A$, hence $c_1(AB)$ is the rank $A \cap B$ in A. Similarly,
2. $c_0(AD)$ lies in A and does not lie in D, therefore $S = A$, hence $c_1(AD)$ is the rank $A \cap D$ in A.

Since $c_0(AB) = c_0(AD)$, we see that the two sets $A \cap B$ and $A \cap D$ have equal rank in A. Recall that our definition of ranks imposed a linear order on distinct subsets of a given set. Hence, two subsets in A may have equal rank if and only if they are equal. This implies

$$A \cap B = A \cap D \qquad \textbf{(2)}$$

We now repeat the process with interchanging A's and C's roles: assume that $c_0(CB) \in C - B$ and $c_0(AD) \in A - D$. Since $c_0(CB) = c_0(AD)$, $c_0(CB) \in A - D$. Thus $c_0(CB) \in (C - B) \cap (A - D)$.

Since $(C - B) \cap (A - D) = (C \cap A) - (B \cup D)$, we have that $c_0(CD) = c_0(CB) \in (C \cap A) - (B \cup D)$. This implies that

1. $c_0(CB)$ lies in C and does not lie in B, therefore $S = C$, hence $c_1(CB)$ is the rank of $C \cap B$ in C. Similarly,
2. $c_0(CD)$ lies in C and does not lie in D, therefore $S = C$, hence $c_1(CD)$ is the rank of $C \cap D$ in C.

Since $c_0(CB) = c_0(CD)$, we see that the two sets $C \cap B$ and $C \cap D$ have equal rank in C. This implies

$$C \cap B = C \cap D \qquad \textbf{(3)}$$

Because $c(AC) = c(BD)$, let $c_0(AC) = c_0(BD) = i$. We know that

$i\ (= c_0(AC)) \in A - C$ or $C - A$

and

$i\ (= c_0(BD)) \in B - D$ or $D - B$

Hence, there are 4 possibilities for i:

1. $i \in (A - C) \cap (B - D) = (A \cap B) - (C \cup D)$.
   (2) implies $i \in (A \cap D) - (C \cup D) = \emptyset$, which is a contradiction.

2. $i \in (A - C) \cap (D - B) = (A \cap D) - (C \cup B)$.
   (2) implies $i \in (A \cap B) - (C \cup B) = \emptyset$, which is a contradiction.

3. $i \in (C - A) \cap (B - D) = (C \cap B) - (A \cup D)$.
   (3) implies $i \in (C \cap D) - (A \cup D) = \emptyset$, which is a contradiction.

4. $i \in (C - A) \cap (D - B) = (C \cap D) - (A \cup B)$
   (3) implies $i \in (C \cap B) - (A \cup B) = \emptyset$, which is a contradiction.

Hence a 2-colored K_4 of Type 2 cannot occur in G.

This completes the proof of Claim 1.

<u>Proof of Claim 2:</u>

i. First we show that the function c_0 can attain atmost $m-1$ distinct values in G.

Observe that for any edge $AB, A \triangle B$ cannot be empty (otherwise $A-B = B-A$, hence $A = B$ which contradicts the fact that all vertices of G are distinct). This implies $A \triangle B$ has at-least 1 element. Furthermore, $A \triangle B$ cannot have exactly one element. To see this, suppose $A \triangle B$ has exactly one element $k \in A-B$. Then $|A| = 1 + |A \cap B|$ and $|B| = |A \cap B|$, so $|A| > |B|$, which contradicts the fact that every vertex of G has equal cardinality (i.e. t). Therefore $A \triangle B$ must have at-least two elements.

Recall that the function c_0 maps an edge AB onto the minimum element in $A \triangle B \subseteq [m]$. c_0 can attain any value from 1 to $m-1$ (Note: c_0 can never equal m, because if $m \in A \triangle B$, then m will always be the largest element in $A \triangle B$, and we just showed that $A \triangle B$ always contains at-least 2 elements, so c_0 will take the other value). Hence the total number of possible values of c_0 is $m-1$.

ii. Now we show that the function c_1 can attain at-most $2^t - 1$ distinct values in G.

Observe that c_1 maps an edge AB onto the rank of $A \cap B$ in either A or B. Since the total number of proper subsets of a t-subset T of $[m]$ is $2^t - 1$, the total number of possible rank values of AB is at-most $2^t - 1$. The range of ranks is the same for every edge, hence the function c_1 attains at most $2^t - 1$ distinct values.

Conclusion: Since the coloring function $c(AB)$ is an ordered pair of $c_0(AB)$ and $c_1(AB)$, the total number of distinct values it can attain is $(2^t - 1)(m-1)$. This completes the proof of Claim 2.

5. The Theorem

To prove the theorem statement for a given n, we choose the optimal values of t and m such that the final result falls within the desired bound. These optimal values are as follows: set $t = \left\lceil \frac{\sqrt{\log n}}{\sqrt{\log 2}} \right\rceil$ and choose m such that $\binom{m}{t} < n \leq \binom{m+1}{t}$. To proceed with the proof, we make note of the following facts:

<u>Fact 1</u>: f is a non-decreasing function of n.

<u>Proof:</u> Let p, q be some fixed integers. Suppose f is a decreasing function of n. Then for some u and v, where $u < v$, $f(u,p,q) > f(v,p,q)$. Let G_u and G_v be the corresponding graphs on $|u|$

and $|v|$ vertices respectively. This means a lesser number of colors is required to give the bigger graph G_v a (4,3)-coloring. However, this coloring is a (4,3)-coloring on the smaller graph G_u as well, which is a contradiction to $f(u,p,q) > f(v,p,q)$. Hence f is a non-decreasing function of n.

<u>Fact 2:</u> For $t < m$, $(\frac{m}{t})^t < \binom{m}{t}$.

<u>Proof:</u> $t < m$ implies $kt < km$ for any $k \in \mathbb{Z}^+$, $k < t$.

Hence $-km < -kt$. Adding mt on both sides gives us $mt - km < mt - kt$. So

$m(t-k) < t(m-k)$. Therefore

$$\frac{m}{t} < \frac{m-k}{t-k}, \text{ for } k < t \qquad \textbf{(4)}$$

Now $\binom{m}{t} = \frac{(m)(m-1)(m-2)\ldots\ldots(m-t+1)}{t!} = \left(\frac{m}{t}\right)\left(\frac{m-1}{t-1}\right)\left(\frac{m-2}{t-2}\right)\left(\frac{m-3}{t-3}\right)\ldots\ldots(\frac{m-t+1}{1})$

$$> \left(\frac{m}{t}\right)\left(\frac{m}{t}\right)\left(\frac{m}{t}\right)\ldots\ldots..(\frac{m}{t}) \quad \text{by (4)}$$

$$= (\frac{m}{t})^t$$

So $(\frac{m}{t})^t < \binom{m}{t}$. This completes the proof of Fact 2.

The proof of the theorem now follows easily from the Lemma, Fact 1, Fact 2 and our choices for t and m. We have:

$$f(n,4,3) \leq f(\left(\binom{m+1}{t}\right),4,3) \text{ by Fact 1}$$

$$\leq (2^t - 1)m \quad \text{by the Lemma}$$

$$< (2^t)m$$

Now $n > \binom{m}{t} > (\frac{m}{t})^t$ by Fact 2, therefore $n^{\frac{1}{t}} > \frac{m}{t}$, and this gives us $m < tn^{\frac{1}{t}}$.

So we have $$f(n,4,3) < (2^t)m$$

$$< 2^t t n^{\frac{1}{t}}$$

Finally it can be shown, using non-graph-theoretic techniques, that

$$2^t t n^{\frac{1}{t}} = \big(1+o(1)\big)e^{2\sqrt{\log 2 \log n} + \frac{\log\log n - \log\log 2}{2}} = e^{\sqrt{4\log 2 \log n}\,(1+o(1))}.$$

Therefore, $f(n,4,3) < e^{\sqrt{c\log n}(1+o(1))}, where\ c = 4\log 2$. This completes the proof of the theorem.

6. Subsequent Works

Not long after this article was published, Mubayi came out with another article ([7]), in which he proved that the construction used in this article has the property that at-least $2\lceil \lg p \rceil - 2$ colors appear on the edges of every copy of K_p in G, for $p \geq 5$.

These results by Mubayi helped prove many other cases of the original problem posed by Erdős, as well as other related problems. For instance, it was proved in [5] that $f(n,4,4) = O(n^{\frac{1}{2}+o(1)})$. The proof used similar construction and coloring techniques as in this article. [4] also used Mubayi's result to show that if $n > (\log k)^{ck}$ for a fixed positive constant c, then no matter how the edges of K_n are colored, there is a copy of K_4 that receives at-most two colors. In [6], M. Axenovich proved a tight bound on $f(n,5,9)$, another small case of the original problem and termed as *annoying*.

Recently in [8], Fox and Sudakov rephrased the $f(n,4,3)$ problem in terms of another, more convenient, function as follows: Let $g(k)$ be the largest n for which there is a k-coloring of K_n such that every K_4 receives at least 3 colors, i.e., for which $f(g(k),4,3) \leq k$. Then Mubayi's result is the best known lower bound on $g(k)$ (i.e. $g(k) \geq e^{c\log k\log k}$, for a positive constant c). [4] provides an upper bound on it: $g(k) < (\log k)^{ck}$. [8] provides a better upper bound on it: $g(k) < 2^{ck}$. However, there is still a very large gap between the lower and upper bound for this problem, and as research continues on this particular problem, the conjecture is that the correct growth of $g(k)$is likely to be subexponential in k.

References

[1] P Erdős: Solved and unsolved problems in combinatorics and combinatorial number theory, *Congressus Numerantum*, 32 (1981), 49-62.

[2] P Erdős and A. Gyárfás: A variant of the classical ramsey problem, *Combinatorica*, 17 (1997), 459-467.

[3] P. Erdős and L. Lovász: Problems and results on 3-chromatic hypergraphs and some related problems, in: *Infinite and Finite Sets, Colloquia Math. Soc. J. Bolyai* vol. 10 (A. Hajnal et al eds.) North-Holland, Amsterdam, 609-617.

[4] A. Kostochka and D. Mubayi: When is an almost monochromatic K4 guaranteed?, *Combinatorics, Probability and Computing* 17, (2008), no. 6, 823-830.

[5] D. Mubayi: An explicit construction of a Ramsey Problem, *Combinatorica* 24 (2004), no. 2, 313-324.

[6] M. Axenovich: A generalized Ramsey problem, *Discrete Math.* 222 (2000), no. 1-3, 247-249.

[7] D. Mubayi and D. Eichhorn: Edge-coloring cliques with many colors on subcliques. *Combinatorica*, 20 (3) (2000) 441-444.

[8] J. Fox and B. Sudakov: Dependent random choice. *Random Structures Algorithms* 38 (2011), no. 1-2, 68-99.