POLYCHROMATIC HAMILTON CYCLES

Alan Frieze
Department of Mathematics, Carnegie-Mellon University,
Pittsburgh, U.S.A.*

and

Bruce Reed
Department of Combinatorics and Optimization
University of Waterloo
Waterloo
Canada

January 30, 2006

Abstract

The edges of the complete graph K_n are coloured so that no colour appears more than $k = k(n)$ times, $k = \lceil n/(A \ln n) \rceil$, for some sufficiently large A. We show that there is always a Hamiltonian cycle in which each edge is a different colour. The proof technique is probabilistic.

*Supported by NSF grant CCR-89000112

1 Introduction

Let the edges of the complete graph K_n be coloured so that no edge is coloured more than $k = k(n)$ times. We refer to this as a k-bounded colouring. We say that a Hamilton cycle of K_n is **polychromatic** if each edge is of a different colour. We say that the colouring is **good** if a polychromatic Hamilton cycle exists. Clearly the colouring is good if $k = 1$ and may not be if $k \geq n/2$, since then we may only use $n-1$ colours. The question we address here then is that of how fast can we allow k to grow and still *guarantee* that a k-bounded colouring is good.

The problem is mentioned in Erdös, Nestril and Rödl [1]. There they mention it as an Erdös - Stein problem and show that k can be any constant. Hahn and Thomassen [3] were the next people to consider this problem and they showed that k could grow as fast as $n^{1/3}$ and conjectured that the growth rate of k could in fact be linear. In unpublished work Rödl and Winkler [5] in 1984 improved this to $n^{1/2}$. In this paper we make further progress and prove

Theorem 1 *There is an absolute constant A such that if n is sufficiently large and k is at most $\lceil n/(A \ln n) \rceil$ then any k-bounded colouring is good.*

Proof Throughout the proof assume that A is a large constant, n is large and that we have some fixed k-bounded colouring of K_n.

Let

$$B = 10^{1/3} A^{2/3} \text{ and } D = \frac{4B^2}{A} + 20.$$

Let $p = \frac{B \ln n}{n}$ and construct a random graph H as follows:

Step 1: let $G = G_{n,p} = ([n], E)$.

(Recall that $G_{n,p}$ is the random graph with vertex set $[n] = \{1, 2, \ldots n\}$ in which each possible edge occurs independently with probability p.)

Step 2: let Y denote the set of edges whose colour appears more than once in E.

Let $H = ([n], E/Y)$.

Thus no two edges of H are of the same colour. We prove our theorem by showing that

$$\mathbf{Pr}(H \text{ is Hamiltonian }) = 1 - o(1).$$

as $n \to \infty$.

This clearly implies that K_n *must* have at least one polychromatic Hamilton cycle, provided n is sufficiently large. The proof can be broken into two lemmas.

For $v \in [n]$ let d_v denote the number of edges in Y which are incident with v.

Lemma 1 $\mathbf{Pr}(\exists v \in [n] : d_v \geq D \ln n) = o(1)$

Lemma 2 *If starting with $G = G_{n,p}$ we delete an arbitrary set of edges Y to obtain a graph H and in the process no vertex loses more than $D \ln n$ edges then H is almost surely Hamiltonian.*

Our Theorem is clearly an immediate consequence of these two lemmas.

2 Proof of Lemma 1

Let $d = d_1$ and let $S_1, S_2, \ldots, S_m$ be the partition of the edges of K_n incident with vertex 1 into sets of the same colour $i = 1, 2, \ldots, m$. Let E_i be the set of edges of K_n which have colour i. Let $|S_i| = l_i$ and $|E_i| = k_i \leq k$ for $i = 1, 2, \ldots, m$.

An edge $e \in S_i$ is deleted in Step 2 if either

(a) $E \cap S_i = \{e\}$ and $E_i/S_i \neq \emptyset$

or

(b) $e \in E$ and $|E \cap S_i| \geq 2$.

Let

$D_x = \{$ edges incident with vertex 1 which are deleted via case (x)$\}$,

x=a or b.

Observe that if $i \neq j$ then the sets $D_x \cap S_i$ and $D_x \cap S_j$ are independent (as random sets.)

The size of D_a

Clearly

$$|D_a \cap S_i| = 0 \text{ or } 1, \qquad i = 1, 2, \ldots, m.$$

Also

$$\begin{aligned} \mathbf{Pr}(|D_a \cap S_i| = 1) &= l_i p(1-p)^{l_i - 1}(1 - (1-p)^{k_i - l_i}) \\ &\leq l_i(k_i - l_i)p^2 \\ &\leq (k-1)l_i p^2. \end{aligned}$$

Thus

$$\begin{aligned}
\mathbf{E}(|D_a|) &\le (k-1)p^2\sum_{i=1}^{m} l_i \\
&= (k-1)(n-1)p^2 \\
&< \frac{B^2 \ln n}{A} \\
&= 10^{2/3}A^{1/3}\ln n.
\end{aligned}$$

Now by Theorem 1 of Hoeffding [2]

$$\begin{aligned}
\mathbf{Pr}\left(|D_a| \ge \frac{2B^2\ln n}{A}\right) &\le \exp\left\{-\frac{B^2\ln n}{3A}\right\} \\
&\le n^{-2}.
\end{aligned}$$

The size of D_b

Let $X_i = |E \cap S_i|$ and $\delta_i = 1_{X_i \ge 2}$. Thus

$$|D_b| = \sum_{i=1}^{m} X_i\delta_i.$$

Now fix $i \in [m]$. Unfortunately X_i and δ_i are correlated (positively). So let $Y_i (= BIN(l_i, p))$ be distributed as X_i but be independent of it. Then we claim that

$$X_i\delta_i \text{ is majorised by } (2+Y_i)\delta_i$$

i.e. for all $u \ge 0$

$$\mathbf{Pr}(X_i\delta_i \ge u) \le \mathbf{Pr}((2+Y_i)\delta_i \ge u). \tag{1}$$

To see this we take 2 independent sequences $A_1, A_2, \dots A_l, B_1, B_2, \dots B_l, l = l_i$ of Bernouilli random variables where each is 1 with probability p and zero with probability $1-p$.

Let

$$\rho = \begin{cases} \min\{r : A_1 + A_2 + \ldots + A_r = 2\} & \text{if } A_1 + A_2 + \ldots + A_l \geq 2 \\ \infty & \text{if } A_1 + A_2 + \ldots + A_l \leq 1 \end{cases}$$

Let

$$Z_1 = \begin{cases} 2 + B_{\rho+1} + \ldots + B_l & \text{if } \rho < \infty \\ 0 & \text{if } \rho = \infty. \end{cases}$$

Z_1 has the same distribution as $X_i \delta_i$.

Let

$$Z_2 = \begin{cases} 2 + B_1 + \ldots + B_l & \text{if } \rho < \infty \\ 0 & \text{if } \rho = \infty. \end{cases}$$

Z_2 has the same distribution as $(2 + Y_i)\delta_i$ and (1) follows immediately.

Thus $|D_b|$ is majorised by $\sum_{i=1}^m (2 + Y_i)\delta_i$.

Now

$$\mathbf{Pr}(\delta_i = 1) \leq \binom{l_i}{2} p^2$$

and so

$$\begin{aligned} \mathbf{E}(\sum_{i=1}^m \delta_i) &\leq p^2 \sum_{i=1}^m \binom{l_i}{2} \\ &\leq p^2 \frac{n}{k} \binom{k}{2} \\ &\leq \frac{B^2}{2A} \ln n. \end{aligned}$$

Hence

$$\begin{aligned} \mathbf{Pr}\left(\sum_{i=1}^n \delta_i \geq \frac{B^2}{A} \ln n\right) &\leq \exp\left\{-\frac{B^2}{6A \ln n}\right\} \\ &\leq n^{-2}. \end{aligned}$$

Consider now the distribution of $\sum_{i=1}^{m}(2+Y_i)\delta_i$ conditional on $\sum_{i=1}^{m}\delta_i \leq m_0 = \lfloor (B^2\ln n)/A \rfloor$. This is majorised by

$$\frac{2B^2}{A}\ln n + \sum_{i=1}^{m_0} Z_i$$

where $Z_1, Z_2, \ldots, Z_{m_0}$ are independent binomials $BIN(k,p)$ and so $Z = \sum_{i=1}^{m_0} Z_i = BIN(m_0 k, p)$. Thus

$$\begin{aligned}
\mathbf{E}(Z) &\leq (1+o(1))\frac{B^2}{A}\ln n \frac{n}{A\ln n}\frac{B\ln n}{n} \\
&= (1+o(1))\frac{B^3}{A^2}\ln n \\
&\leq 11\ln n
\end{aligned}$$

So

$$\begin{aligned}
\mathbf{Pr}(Z \geq 20\ln n) &\leq \exp\left\{-\frac{1}{3}\left(\frac{9}{11}\right)^2 11\ln n\right\} \\
&= O(n^{-2}).
\end{aligned}$$

Hence

$$\mathbf{Pr}\left(d \geq \frac{2B^2}{A}\ln n + \frac{2B^2}{A}\ln n + 20\ln n\right) = O(n^{-2}).$$

Multiplying by a factor n to account for all vertices gives the lemma. □

3 Proof of Lemma 2

We modify the proof of Posá [4] to account for the deletion of edges. So assume now that $G = G_1 \cup G_2 \cup G_3$ where G_1 and G_2 are independent copies of $G_{n,p/2}$ and where G_3 is an independent copy of $G_{n,p'}$, where p' satisfies the equation $1-p = (1-p/2)^2(1-p')$. G_3 plays no further role in the analysis.

We first show that G_1/Y almost surely contains a Hamilton path. If it doesn't then there exists $i \in [n]$ such that

there exists a longest path of G_1/Y *which does not go through* i

which implies

no longest path of $\Gamma_i = (G_1/Y)/\{i\}$ *has an end-vertex adjacent to* i *in* G_1.

Let this final event be denoted by $\mathcal{E}_i$. Then

$$\mathbf{Pr}(G_1/Y \text{ has no Hamilton path }) \leq n\mathbf{Pr}(\mathcal{E}_n). \tag{2}$$

Given a longest path Q with end-vertices x_0, y and an edge yv where v is an internal vertex of Q, we obtain a new longest path $Q' = x_0..vy..w$ where w is the neighbour of v on P between v and y. We say that Q' is obtained from Q by a rotation.

So now let P be a longest path of Γ_n and let x_0 be one of its end-vertices. Let END be the set of end-vertices of longest paths of Γ_n which can be obtained from P by a sequence of *rotations* keeping x_0 as a fixed end-vertex.

It follows from Posá [4] that

$$|N(\Gamma_n, END)| < 2|END|, \tag{3}$$

where for a graph Γ and a set $S \subseteq V(\Gamma)$

$$N(\Gamma, S) = \{w \notin S : \exists v \in S \text{ such that } vw \in E(\Gamma)\}.$$

CLAIM: with probability 1-o(n^{-1})

$$S \subseteq [n-1], |S| \leq \frac{n}{4D \ln n} \text{ implies } |N(G_1/\{n\}, S)| \geq 3D(\ln n)|S|.$$

(The proof of this claim is deferred to the end of the proof of the lemma.)

Hence in Γ_n we have with probability 1-o(n^{-1})

$$S \subseteq [n-1], |S| \leq \frac{n}{4D\ln n} \text{ implies } |N(\Gamma_n, S)| \geq D(\ln n)|S|.$$

It follows from (3) that with probability 1-o(n^{-1})

$$|END| \geq \frac{n}{12}.$$

Now consider the edges of G_1 from vertex n to END. They are independent of END and so are distributed as $B(|END|, p/2)$. Thus their expected number is at least $(B \ln n)/24$. Thus if A and hence B is large there will be at least $(B \ln n)/48$ such edges with probability 1-o(n^{-1}). But for large $A, D < B/48$ and so not all of these edges can be included in Y. Thus $\mathbf{Pr}(\mathcal{E}_n) = o(n^{-1})$ and (2) implies that G_1/Y almost surely has a Hamilton path.

To finish the proof take a Hamilton path P of G_1 and fix one of its end-vertices, x_0 say, and using rotations create a set of end-vertices END of Hamilton paths with one end-vertex x_0. The above analysis shows that $|END| \geq \frac{n}{12}$ almost surely. Now add the edges of G_2, which are independent of x_0 and END. Again we can argue that there are almost surely too many $x_0 - END$ edges in G_2 for them all to be included in Y and the lemma follows since the existence of any one not in Y means that H is Hamiltonian.

Proof of CLAIM

If the condition in the claim does not hold then there exist disjoint sets $S, T \subseteq [n-1], s = |S| \leq n/(4D \ln n), t = |T| \leq 3D(\ln n)s \leq 3n/4$ such that each vertex of T is adjacent to at least one vertex in S and no vertex in $[n-1]/(S \cup T)$ is adjacent to any vertex of S.

Fix s, t and let $t_0 = 3sD(\ln n)$ Then the probability of the above event is bounded by

$$\begin{aligned}
\binom{n-1}{s}\binom{n-1}{t}\left(\frac{sp}{2}\right)^t\left(1-\frac{p}{2}\right)^{s(n-1-s-t)} &\leq \left(\frac{ne}{s}\right)^s\left(\frac{ne}{t}\right)^t\left(\frac{sp}{2}\right)^t e^{-snp/10} \\
&= \left(\frac{ne}{s}\right)^s\left(\frac{e}{t}\right)^t\left(\frac{sB\ln n}{2}\right)^t n^{-sB/10} \\
&\leq \left(\frac{ne}{s}\right)^s\left(\frac{e}{t_0}\right)^{t_0}\left(\frac{sB\ln n}{2}\right)^{t_0} n^{-sB/10} \\
&= \left(\frac{ne}{s} n^{3D\ln(Be/6D)-B/10}\right)^s \\
&= o(n^{-3})
\end{aligned}$$

for large A. Now multiply this upper bound by n^2, which bounds the number of possible s, t, in order to prove the claim. □

References

[1] P.Erdös, J.Nesetril and V.Rödl, *Some problems related to partitions of edges of a graph* in graphs and other Combinatorial topics, Teubner, Leipzig (1983) 54-63.

[2] W.Höeffding,*Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association 58 (1963) 13-30.

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

[4] L.Pósa,*Hamilton circuits in random graphs*,Discrete Mathematics 14 (1976) 359-64.

[5] P.Winkler, Private Communication.