A Short Proof of the Hajnal-Szemerédi Theorem on Equitable Coloring

H.A. Kierstead* A.V. Kostochka†

July 26, 2006

1 Introduction

An *equitable k-coloring* of a graph G is a proper k-coloring, for which any two color classes differ in size by at most one. Equitable colorings naturally arise in some scheduling, partitioning, and load balancing problems [1, 15, 16]. Pemmaraju [13] and Janson and Ruciński [6] used equitable colorings to derive deviation bounds for sums of dependent random variables that exhibit limited dependence. In 1964 Erdős [3] conjectured that any graph with maximum degree $\Delta(G) \leq r$ has an equitable $(r+1)$-coloring. This conjecture was proved in 1970 by Hajnal and Szemerédi [5] with a surprisingly long and complicated argument. Recently, Mydlarz and Szemerédi [11] found a polynomial time algorithm for such coloring.

In search of an easier proof, Seymour [14] strengthened Erdős' conjecture by asking whether every graph with minimum degree $\delta(G) \geq \frac{k}{k+1}|G|$ contains the k-th power of a hamiltonian cycle. (If $|G| = (r+1)(s+1)$ and $\Delta(G) \leq r$ then $\delta(\bar{G}) \geq \frac{s}{s+1}|G|$; each $(s+1)$-interval of a s-th power of a hamiltonian cycle in $\bar{G}$ is an independent set in G.) The case $k = 1$ is Dirac's Theorem and the case $k = 2$ is Pósa's Conjecture. Fan and Kierstead [4] proved Pósa's Conjecture with cycle replaced by path. Komlós, Sarkozy and Szemerédi [7] proved Seymour's conjecture for graphs with sufficiently many (in terms of k) vertices. Neither of these partial results has a simple proof. In fact, [7] uses the Regularity Lemma, the Blow-up Lemma and the Hajnal-Szemerédi Theorem.

A different strengthening was suggested recently by Kostochka and Yu [9, 10]. In the spirit of Ore's theorem on hamiltonian cycles [12], they conjectured that every graph in which $d(x) + d(y) \leq 2r$ for every edge xy has an equitable $(r+1)$-coloring.

In this paper we present a short proof of the Hajnal-Szemerédi Theorem and present another polynomial time algorithm that constructs an equitable $(r+1)$-coloring of any

*Department of Mathematics and Statistics, Arizona State University, Tempe, AZ 85287, USA. E-mail address: kierstead@asu.edu. Research of this author is supported in part by the NSA grant MDA 904-03-1-0007

†Department of Mathematics, University of Illinois, Urbana, IL, 61801, USA and Institute of Mathematics, Novosibirsk, 630090, Russia. E-mail address: kostochk@math.uiuc.edu. Research of this author is supported in part by the NSF grant DMS-0400498.

graph G with maximum degree $\Delta(G) \leq r$. Our approach is similar to the original proof, but a discharging argument allows for a much simpler conclusion. Our techniques have paid further dividends. In another paper we will prove the above conjecture of Kostochka and Yu [9, 10] in a stronger form: with $2r+1$ in place of $2r$. They also yield partial results towards the Chen-Lih-Wu Conjecture [2] about equitable r-colorings of r-regular graphs and towards a list analogue of Hajnal-Szemerédi Theorem (see [8] for definitions).

Most of our notation is standard; possible exceptions include the following. For a vertex y and set of vertices X, $N_X(y) := N(y) \cap X$ and $d_X(y) = |N_X(y)|$. If μ is a function on edges then $\mu(A,B) := \sum_{xy \in E(A,B)} \mu(x,y)$, where $E(A,B)$ is the set of edges linking a vertex in A to a vertex in B. For a function $f : V \to Z$, the restriction of f to $W \subseteq V$ is denoted by $f|W$. Functions are viewed formally as sets of ordered pairs. So if $u \notin V$ then $g := f \cup \{(u,\gamma)\}$ is the extension of f to $V \cup \{u\}$ such that $g(u) = \gamma$.

2 Main proof

Let G be a graph with $s(r+1)$ vertices. A *nearly equitable* $(r+1)$-coloring of G is a proper coloring f, whose color classes all have size s except for one *small* class $V^- = V^-(f)$ with size $s-1$ and one *large* class $V^+ = V^+(f)$ with size $s+1$. Given such a coloring f, define the auxiliary digraph $H = H(G,f)$ as follows. The vertices of H are the color classes of f. A directed edge VW belongs to $E(H)$ iff some vertex $y \in V$ has no neighbors in W. In this case we say that y is *movable* to W. Call $W \in V(H)$ *accessible*, if V^- is reachable from W in H. So V^- is trivially accessible. Let $\mathcal{A} = \mathcal{A}(f)$ denote the family of accessible classes, $A := \bigcup \mathcal{A}$ and $B := V(G) \setminus A$. Let $m := |\mathcal{A}| - 1$ and $q := r - m$. Then $|A| = (m+1)s - 1$. Then $|B| = (r-m)s + 1$. Each vertex $y \in B$ cannot be moved to A and so satisfies

$$d_A(y) \geq m+1 \text{ and } d_B(y) \leq q-1. \tag{1}$$

Lemma 1 *If G has a nearly equitable $(r+1)$-coloring f, whose large class V^+ is accessible, then G has an equitable $(r+1)$-coloring.*

Proof. Let $\mathcal{P} = V_1, \dots, V_k$ be a path in $H(G,f)$ from $V_1 := V^+$ to $V_k := V^-$. This means that for each $j = 1, \dots, k-1$, V_j contains a vertex y_j that has no neighbors in V_{j+1}. So, if we move y_j to V_{j+1} for $j = 1, \dots, k-1$, then we obtain an equitable $(r+1)$-coloring of G. ■

Suppose $V^+ \subseteq B$. If $A = V^-$ then $|E(A,B)| \leq r\,|V^-| = r(s-1) < 1 + rs = |B|$, a contradiction to (1). Thus $m+1 = |\mathcal{A}| \geq 2$. Call a class $V \in \mathcal{A}$ *terminal*, if V^- is reachable from every class $W \in \mathcal{A} \setminus \{V\}$ in the digraph $H - V$. Trivially, V^- is non-terminal. Every non-terminal class W partitions $\mathcal{A} \setminus \{W\}$ into two parts $\mathcal{S}_W$ and $\mathcal{T}_W \neq \emptyset$, where $\mathcal{S}_W$ is the set of classes that can reach V^- in $H - W$. Choose a non-terminal class U so that $\mathcal{A}' := \mathcal{T}_U \neq \emptyset$ is minimal. Then every class in $\mathcal{A}'$ is terminal and no class in $\mathcal{A}'$ has a vertex movable to any class in $(\mathcal{A} \setminus \mathcal{A}') \setminus \{U\}$. Set $t := |\mathcal{A}'|$ and $A' := \bigcup \mathcal{A}'$. Thus every $x \in A'$ satisfies

$$d_A(x) \geq m - t. \tag{2}$$

Call an edge zy with $z \in W \in \mathcal{A}'$ and $y \in B$, a *solo edge* if $N_W(y) = \{z\}$. The ends of solo edges are called *solo vertices* and vertices linked by solo edges are called *special neighbors* of each other. Let S_z denote the set of special neighbors of z and S^y denote the set of special neighbors of y in A'. Then at most $r-(m+1+d_B(y))$ color classes in A have more than one neighbor of y. Hence

$$|S^y| \geq t-q+1+d_B(y). \tag{3}$$

Lemma 2 *If there exists $W \in \mathcal{A}'$ such that no solo vertex in W is movable to a class in $\mathcal{A} \setminus \{W\}$ then $q+1 \leq t$. Furthermore, every vertex $y \in B$ is solo.*

Proof. Let S be the set of solo vertices in W and $D := W \setminus S$. Then every vertex in $N_B(S)$ has at least one neighbor in W and every vertex in $B \setminus N_B(S)$ has at least two neighbors in W. It follows that $|E(W,B)| \geq |N_B(S)| + 2(|B| - |N_B(S))|)$. Since no vertex in S is movable, every $z \in S$ satisfies $d_B(z) \leq q$. By (2), every vertex $x \in W$ satisfies $d_B(x) \leq t+q$. Thus, using $s = |W| = |S| + |D|$,

$$qs + q\,|D| + 2 = 2(qs+1) - q\,|S| \leq |E(W,B)| \leq q\,|S| + (t+q)\,|D| \leq qs + t\,|D|$$

It follows that $q+1 \leq t$. Moreover, by (3) every $y \in B$ satisfies $|S^y| \geq t - q + d_B(y) \geq 1$. Thus y is solo. ■

Lemma 3 *If $V^+ \subseteq B$ then there exists a solo vertex $z \in W \in A'$ such that either z is movable to a class in $\mathcal{A} \setminus \{W\}$ or z has two nonadjacent special neighbors in B.*

Proof. Suppose not. Then by Lemma 2 every vertex in B is solo. Moreover, S_z is a clique for every solo vertex $z \in A'$. Consider a weight function μ on $E(A', B)$ defined by

$$\mu(xy) := \begin{cases} \frac{q}{|S_x|} & \text{if } xy \text{ is solo,} \\ 0 & \text{if } xy \text{ is not solo.} \end{cases}$$

For $z \in A'$ we have $\mu(z,B) = |S_z|\,\frac{q}{|S_z|} = q$ if z is solo; otherwise $\mu(z,B) = 0$. Thus $\mu(A',B) \leq q\,|A'| = qst$. On the other hand, consider $y \in B$. Let $c_y := \max\{|S_z| : z \in S^y\}$, say $c_y = |S_z|\,, z \in S^y$. Using that S_z is a clique and (1), $c_y - 1 \leq d_B(y) \leq q-1$. So $c_y \leq q$. Together with (3) this yields

$$\mu(A',y) = \sum_{z \in S^y} \frac{q}{|S_z|} \geq |S^y|\,\frac{q}{c_y} \geq (t-q+c_y)\frac{q}{c_y} = (t-q)\frac{q}{c_y} + q \geq t.$$

Thus $\mu(A',B) \geq t\,|B| = t(qs+1) > qst \geq \mu(A',B)$, a contradiction. ■

We are now ready to prove the Hajnal-Szemerédi Theorem.

Theorem 4 *If G is a graph satisfying $\Delta(G) \leq r$ then G has an equitable $(r+1)$-coloring.*

Proof. We may assume that $|G|$ is divisible by $r+1$. To see this, suppose that $|G| = s(r+1) - p$, where $p \in [r]$. Let $G' := G + K^p$. Then $|G'|$ is divisible by $r+1$ and $\Delta(G') \leq r$. Moreover, the restriction of any equitable $(r+1)$-coloring of G' to G is an equitable $(r+1)$-coloring of G.

Argue by induction on $\|G\|$. The base step $\|G\| = 0$ is trivial, so consider the induction step $\|G\| \geq 1$. Let $e = xy$ be an edge of G. By the induction hypothesis there exists an equitable $(r+1)$-coloring f_0 of $G-e$. We are done, unless some color class V contains both x and y. Since $d(x) \leq r$, there exists another class W such that x is movable to W. Doing so yields a nearly equitable $(r+1)$-coloring f of G with $V^-(f) = V \setminus \{x\}$ and $V^+(f) = W \cup \{x\}$. We now show by a secondary induction on $q(f)$ that G has an equitable $(r+1)$-coloring.

If $V^+ \in \mathcal{A}$ then we are done by Lemma 1; in particular, the base step $q = 0$ holds. Otherwise, by Lemma 3 there exists a class $W \in \mathcal{A}'$, a solo vertex $z \in W$ and a vertex $y_1 \in S_z$ such that either z is movable to a class $X \in \mathcal{A} \setminus \{W\}$ or z is not movable in $\mathcal{A}$ and there exists another vertex $y_2 \in S_z$, which is not adjacent to y_1. By (1) and the primary induction hypothesis, there exists an equitable q-coloring g of $B^- := B \setminus \{y_1\}$. Let $A^+ := A \cup \{y_1\}$.

Case 1: z is movable to $X \in \mathcal{A}$. Move z to X and y_1 to $W \setminus \{z\}$ to obtain a nearly equitable $(m+1)$-coloring φ of A^+. Since $W \in \mathcal{A}'(f)$, $V^+(\varphi) = X \cup \{z\} \in \mathcal{A}(\varphi)$. By Lemma 1, A^+ has an equitable $(m+1)$-coloring φ'. Then $\varphi' \cup g$ is an equitable $(r+1)$-coloring of G.

Case 2: z is not movable to any class in $\mathcal{A}$. Then $d_{A^+}(z) \geq d_A(z) + 1 \geq m + 1$. Thus $d_{B^-}(z) \leq q - 1$. So we can move z to a color class $Y \subseteq B$ of g to obtain a new coloring g' of $B^* := B^- \cup \{z\}$. Also move y_1 to W to obtain an $(m+1)$-coloring ψ of $A^* := V(G) \setminus B^*$. Set $\psi' := \psi \cup g'$. Then ψ' is a nearly equitable coloring of G with $A^* \subseteq A(\psi')$. Moreover, y_2 is movable to $W^* := W \cup \{y_1\} \setminus \{z\}$. Thus $q(\psi') < q(f)$ and so by the secondary induction hypothesis, G has an equitable $(r+1)$-coloring ψ''. ■

3 A polynomial algorithm

Our proof clearly yields an algorithm. However it may not be immediately clear that its running time is polynomial. The problem lies in the secondary induction, where we may apply Case 2 $O(r)$ times, each time calling the algorithm recursively. Lemma 2 is crucial here; it allows us to claim that when we are in Case 2 (doing lots of work) we make lots of progress. As above G is a graph satisfying $\Delta(G) \leq r$ and $|G| =: n =: s(r+1)$. Let f be a nearly equitable $(r+1)$-coloring of G.

Theorem 5 *There exists an algorithm $\mathcal{P}'$ that from input (G, f) constructs an equitable $(r+1)$-coloring of G in $c(q+1)n^3$ steps.*

Proof. We shall show that the construction in the proof of Theorem 4 can be accomplished in the stated number of steps. Argue by induction on q. The base step $q = 0$ follows immediately from Lemma 1 and the observation that the construction of H and the recoloring can be carried out in $\frac{1}{4}cn^3$ steps. Now consider the induction step. In $\frac{1}{4}cn^3$ steps construct $\mathcal{A}, \mathcal{A}', B, W, z, y_1$. Using the induction hypothesis on the input $(G[B^-], f|B^-)$, construct the coloring g of B^- in $c(q(f|B^-)+1)(qs)^3 \leq cqn^3$ steps. In $\frac{1}{4}cn^3$ steps determine whether Case 1 or Case 2 holds.

If Case 1 holds, construct the recoloring φ' in $\frac{1}{4}cn^3$ steps. This yields an equitable $(r+1)$-coloring $g \cup \varphi'$ in a total of $\frac{3}{4}cn^3 + cqn^3 \leq c(q+1)n^3$.

If Case 2 holds then, by Lemma 2, $q+1 \leq t$. Thus we used only $\frac{1}{8}cqn^3$ steps to construct g. Use an additional $\frac{1}{4}cn^3$ steps to extend g to ψ'. Notice that W^* is non-terminal in ψ'. Thus we can choose $\mathcal{A}'(\psi')$ so that $A'(\psi') \subseteq B$. If Case 1 holds for ψ' then as above we can construct an equitable coloring in an additional $\frac{1}{4}cn^3 + \frac{1}{8}cqn^3$ steps. So the total number of steps is at most $c(q+1)n^3$. Otherwise by Lemma 2 $q(\psi') < \frac{1}{2}q$. Thus by the induction hypothesis we can finish in $c\frac{qn^3}{16}$ additional steps. Then the total number of steps is less than $c(q+1)n^3$. ∎

Theorem 6 *There is an algorithm $\mathcal{P}$ of complexity $O(n^5)$ that constructs an equitable $(r+1)$-coloring of any graph G satisfying $\Delta(G) \leq r$ and $|G| = n$.*

Proof. As above, we may assume that n is divisible by $r+1$. Let $V(G) = \{v_1, \ldots, v_n\}$. Delete all edges from G to form G_0 and let f_0 be an equitable coloring of G_0. Now, for $i = 1, \ldots, n-1$, do the following:
(i) Add back all the edges of G incident with v_i to form G_i;
(ii) If v_i has no neighbors in its color class in f_{i-1}, then set $f_i := f_{i-1}$.
(iii) Otherwise, move v_i to a color class that has no neighbors of v_i to form a nearly equitable coloring f'_{i-1} of G_i. Then apply $\mathcal{P}'$ to (G_i, f'_{i-1}) to get an equitable $(r+1)$-coloring f_i of G_i.

Then f_{n-1} is an equitable $(r+1)$-coloring of $G_{n-1} = G$. Since we have only $n-1$ stages and each stage runs in $O(n^4)$ steps, the total complexity is $O(n^5)$. ∎

Acknowledgement. We would like to thank J. Schmerl for many useful comments.

References

[1] J. Blazewicz, K. Ecker, E. Pesch, G. Schmidt, J. Weglarz, *Scheduling computer and manufacturing processes. 2nd ed.*, Berlin: Springer. 485 p. (2001).

[2] B.-L. Chen, K.-W. Lih, and P.-L. Wu, Equitable coloring and the maximum degree, *Europ. J. Combinatorics*, **15** (1994), 443–447.

[3] P. Erdős, Problem 9, in "Theory of Graphs and Its Applications" (M. Fieldler, Ed.), 159, Czech. Acad. Sci. Publ., Prague, 1964.

[4] G. Fan and H. A. Kierstead, Hamiltonian square paths, *J. Combin. Theory Ser. B* **67** (1996) 167-182.

[5] A. Hajnal and E. Szemerédi, Proof of a conjecture of P. Erdős, in "Combinatorial Theory and its Application" (P. Erdős, A. Rényi, and V. T. Sós, Eds.), pp. 601-623, North-Holland, London, 1970.

[6] S. Janson and A. Ruciński, The infamous upper tail, *Random Structures and Algorithms*, **20** (2002), 317–342.

[7] J. Komlós, G. Sarkozy and E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Annals of Combinatorics*, **1** (1998), 43-60.

[8] A. V. Kostochka, M. J. Pelsmajer, and D. B. West, A list analogue of equitable coloring, *J. of Graph Theory*, **44** (2003), 166–177

[9] A. V. Kostochka and G. Yu, Extremal problems on packing of graphs, *Oberwolfach reports*, No 1 (2006), 55–57.

[10] A. V. Kostochka and G. Yu, Ore-type graph packing problems, to appear in *Combinatorics, Probability and Computing.*

[11] M. Mydlarz and E. Szemerédi, Algorithmic Brooks' Theorem, manuscript.

[12] O. Ore, Note on Hamilton circuits, *Amer. Math. Monthly*, **67** (1960), 55.

[13] S. V. Pemmaraju, Equitable colorings extend Chernoff-Hoeffding bounds, *Proceedings of the 5th International Workshop on Randomization and Approximation Techniques in Computer Science (APPROX-RANDOM 2001)*, 2001, 285–296.

[14] P. Seymour, Problem section, in "Combinatorics: Proceedings of the British Combinatorial Conference 1973" (T. P. McDonough and V. C. Mavron, Eds.), pp. 201-202, Cambridge Univ. Press, Cambridge, UK, 1974.

[15] B. F. Smith, P. E. Bjorstad, and W. D. Gropp, *Domain decomposition. Parallel multilevel methods for elliptic partial differential equations*, Cambridge: Cambridge University Press, 224 p. (1996).

[16] A. Tucker, Perfect graphs and an application to optimizing municipal services, *SIAM Review*, 15(1973), 585–590.