New Bounds for the Acyclic Chromatic Index

Anton Bernshteyn*
University of Illinois at Urbana-Champaign

Abstract

An edge coloring of a graph $G$ is called an acyclic edge coloring if it is proper (i.e. adjacent edges receive different colors) and every cycle in $G$ contains edges of at least three different colors (there are no bichromatic cycles in $G$). The least number of colors needed for an acyclic edge coloring of $G$ is called the acyclic chromatic index of $G$ and is denoted by $\alpha'(G)$. It is conjectured by Fiamčik [7] and independently by Alon et al. [2] that $\alpha'(G) \leq \Delta(G) + 2$, where $\Delta(G)$ denotes the maximum degree of $G$. However, the best known general bound is $\alpha'(G) \leq 4(\Delta(G) - 1)$ due to Esperet and Parreau [3]. We apply a generalization of the Lovász Local Lemma that we call the Local Action Lemma [3] to show that if $G$ contains no copy of a given bipartite graph $H$, then $\alpha'(G) \leq 3\Delta(G) + o(\Delta(G))$. Moreover, for every $\varepsilon > 0$ there exists a constant $c$ such that if $g(G) \geq c$, then $\alpha'(G) \leq (2 + \varepsilon)\Delta(G) + o(\Delta(G))$, where $g(G)$ denotes the girth of $G$.

1 Introduction

We consider only finite simple undirected graphs. An edge coloring of a graph $G$ is called an acyclic edge coloring if it is proper (i.e. adjacent edges receive different colors) and every cycle in $G$ contains edges of at least three different colors (there are no bichromatic cycles in $G$). The least number of colors needed for an acyclic edge coloring of $G$ is called the acyclic chromatic index of $G$ and is denoted by $\alpha'(G)$. The notion of acyclic (vertex) coloring was first introduced by Grünbaum [8]. The edge version was first considered by Fiamčik [7], and independently by Alon et al. [1].

Like for many other graph parameters, it is quite natural to ask for an upper bound on the acyclic chromatic index of a graph $G$ in terms of its maximum degree $\Delta(G)$. Since $\alpha'(G) \geq \chi'(G) \geq \Delta(G)$, this bound must be at least linear in $\Delta(G)$. The first linear bound was given by Alon et al. [1], who showed that $\alpha'(G) \leq 64\Delta(G)$. Although it resolved the problem of determining the order of growth of $\alpha'(G)$ in terms of $\Delta(G)$, it was conjectured that the sharp bound should be much lower.

Conjecture 1.1 (Fiamčik [7], Alon et. al. [2]). For every graph $G$ we have $\alpha'(G) \leq \Delta(G) + 2$.

Note that the bound in Conjecture 1.1 is only one more than Vizing’s bound on the chromatic index of $G$. However, this elegant conjecture is still far from being proven.

The first major improvement to the bound $\alpha'(G) \leq 64\Delta(G)$ was made by Molloy and Reed [11], who proved that $\alpha'(G) \leq 16\Delta(G)$. This bound remained the best for a while, until Ndreca et al. [14] managed to improve it to $\alpha'(G) \leq \lceil 9.62(\Delta(G) - 1) \rceil$. This estimate was recently lowered further to $\alpha'(G) \leq 4(\Delta(G) - 1)$ by Esperet and Parreau [3].

All the bounds mentioned above were derived using probabilistic arguments, and recent progress was stimulated by discovering more sophisticated and powerful analogues of the Lovász Local Lemma, namely the stronger version of the LLL due to Bissacot et al. [4] and the entropy compression method of Moser and Tardos [12].

The probability that a cycle would become bichromatic in a random coloring is less if the cycle is longer. Thus it should be easier to establish better bounds on the acyclic chromatic index for graphs with high enough girth. Indeed, Alon et al. [2] showed that if $g(G) \geq c_1\Delta(G)\log\Delta(G)$, where $c_1$ is some universal constant, then $\alpha'(G) \leq \Delta(G) + 2$. They also proved that if $g(G) \geq c_2\log\Delta(G)$, then $\alpha'(G) \leq 2\Delta(G) + 2$. That was lately improved by Muthu et al. [13] in the following way: For every $\varepsilon > 0$ there exists a constant $c$ such that if $g(G) \geq c\log\Delta(G)$, then $\alpha'(G) \leq (1 + \varepsilon)\Delta(G) + o(\Delta(G))$.

*Supported by the Illinois Distinguished Fellowship.
We are now turning to the case when \( g(G) \) is bounded below by some constant independent on \( \Delta(G) \), which would be the main topic of this paper. The first bounds of such type were given by Muthu et al. \([13]\), who proved that \( a'(G) \leq 9\Delta(G) \) if \( g(G) \geq 9 \), and \( a'(G) \leq 4.5\Delta(G) \) if \( g(G) \geq 220 \). In their already mentioned work \([6]\) Esperet and Parreau not only improved both these estimates even in the case of arbitrary \( g(G) \), they also showed that \( a'(G) \leq \lceil 3.74(\Delta(G) - 1) \rceil \) if \( g(G) \geq 7 \), \( a'(G) \leq \lceil 3.14(\Delta(G) - 1) \rceil \) if \( g(G) \geq 53 \), and in fact that for every \( \varepsilon > 0 \) there exists a constant \( c \) such that if \( g(G) \geq c \), then \( a'(G) \leq (3 + \varepsilon)\Delta(G) + o(\Delta(G)) \).

The result that we present here consists in further improvement of the latter bounds. Namely, we establish the following.

**Theorem 1.2.** Let \( G \) be a graph with maximum degree \( \Delta \), and let \( H \) be some bipartite graph. If \( G \) does not contain \( H \) as a subgraph, then \( a'(G) \leq 3\Delta + o(\Delta) \).

**Remark 1.3.** In our original version of Theorem 1.2 we considered only the case where \( H \) was the 4-cycle. That almost the same proof in fact works for any bipartite \( H \) was observed by Esperet and de Verclos.

**Remark 1.4.** The function \( o(\Delta) \) in the statement of Theorem 1.2 depends on \( H \).

**Theorem 1.5.** For every \( \varepsilon > 0 \) there exists a constant \( c \) such that for every graph \( G \) with maximum degree \( \Delta \) and \( g(G) \geq c \) we have \( a'(G) \leq (2 + \varepsilon)\Delta + o(\Delta) \).

**Remark 1.6.** The bound of the last theorem was recently improved further to \( a'(G) \leq (1 + \varepsilon)\Delta + o(\Delta) \) by Cai et al. \([3]\), using a different (and much more sophisticated) argument.

To prove Theorems 1.2 and 1.5 we use a generalization of the Lovász Local Lemma that we call the Local Action Lemma (the LAL for short). The LAL is inspired by recent combinatorial applications of the entropy compression method. Many examples of applying the LAL as well as its proof are given in \([3]\). We provide all the required definitions and the statement of the LAL in Section 2. In Section 3 we prove Theorem 1.2 and in Section 4 we prove Theorem 1.5.

## 2 The Local Action Lemma

### 2.1 Informal discussion

We first informally describe the general framework for the LAL. Suppose that we are given a certain set \( \mathcal{X} \), and a subset \( \mathcal{G} \subseteq \mathcal{X} \) of “good” elements. Then we (according to some distribution) pick a random element \( X \) from \( \mathcal{X} \). Our goal is to show that \( \Pr(X \in \mathcal{G}) > 0 \). To do so, we need to use some information about the structure of the set \( \mathcal{X} \). In the case of the LAL, this structure is given by a collection \( M \) of operations that map \( \mathcal{X} \) to itself and satisfy certain properties. We need \( M \) to be closed under taking compositions, so it is useful to think of \( M \) as an abstract monoid (i.e. an algebraic structure with a single associative binary operation and an identity element) that acts on \( \mathcal{X} \) (i.e. each element \( \alpha \in M \) is interpreted as a map \( \alpha : \mathcal{X} \to \mathcal{X} \) in a way that is compatible with the monoid operation). If \( \alpha \in M \) and \( x \in \mathcal{X} \), then we write \( \alpha(x) \) to denote the result of application of \( \alpha \) (as a map) to the element \( x \).

We require that the action of \( M \) on \( \mathcal{X} \) “respects” the set \( \mathcal{G} \), namely the two following conditions must be satisfied. First, \( \mathcal{G} \) must be \( M \)-invariant, in other words, if \( x \in \mathcal{G} \) and \( \alpha \in M \), then \( \alpha(x) \in \mathcal{G} \). Second, for some \( \alpha \in M \) we must have \( \Pr(\alpha(X) \in \mathcal{G}) > 0 \).

Now that we have this structure on \( \mathcal{X} \), we need to define the notion of “locality” in this case. For that we fix a generating set \( B \subseteq M \) (i.e. a subset of \( M \) such that the smallest set containing \( B \) that is closed under the monoid operation is \( M \) itself). The elements of \( B \) should be thought of as “local” operations on \( \mathcal{X} \).

Consider the following situation of “local improvement”: an element \( x \in \mathcal{X} \setminus \mathcal{G} \) and a generator \( \beta \in B \) such that \( \beta(x) \in \mathcal{G} \) (i.e. a “bad” element that becomes “good” after a certain local change). We “classify” such situations by assigning to each of them an element \( g_\beta(x) \in M \). The way of choosing \( g_\beta(x) \) depends on a particular problem. Roughly speaking, it should somehow reflect the reason preventing \( x \) from being a “good” element.
Fix some $\alpha \in M$, $\beta \in B$, and $\gamma \in M$. Suppose we are given that $\alpha \beta \gamma (X) \in \mathcal{G}$ (i.e. the randomly chosen element becomes good after applying the operation $\alpha \beta \gamma$). Then we are interested in estimating the conditional probability of the event that $\gamma (X) \notin \mathcal{G}$, but $\beta \gamma (X) \in \mathcal{G}$ (i.e. $\gamma (X)$ is not “good” yet, while $\beta \gamma (X)$ already is), and that $g_\beta (\gamma (X)) = \alpha$ (note that $g_\beta (\gamma (X))$ is defined in this case). We denote this probability (more precisely, the supremum of these probabilities over all $\gamma \in M$) by $P(\beta, \alpha)$. The LAL, roughly speaking, asserts that if the quantities $P(\beta, \alpha)$ are small enough, then necessarily $\Pr (X \in \mathcal{G}) > 0$. We make this statement precise in the next section.

2.2 Precise statement

We will need the following definitions. Recall that a monoid is an algebraic structure with a single associative binary operation and an identity element. We will use the multiplicative notation for the monoid operation, and denote its identity element by $1$. For a monoid $M$, a subset $B \subseteq M$ is called a generating set of $M$ if $M$ is the smallest set containing $B$ that is closed under the monoid operation. If $B$ is a fixed generating set of $M$, then we refer to the elements of $B$ as the generators of $M$.

A (left) action of a monoid $M$ on a set $\mathcal{X}$ is a map $\varphi : M \times \mathcal{X} \to \mathcal{X}$ that is compatible with the monoid operation, i.e.

$$\varphi (\alpha \beta, x) = \varphi (\alpha, \varphi (\beta, x)),$$

and

$$\varphi (1, x) = x$$

for all $\alpha, \beta \in M$ and $x \in \mathcal{X}$. If a monoid action $\varphi$ is fixed, then we say that $M$ acts on $\mathcal{X}$ and use the notation $\alpha (x)$ for $\varphi (\alpha, x)$.

Let $M$ be a monoid with a generating set $B$. If $\zeta : B \to \mathbb{R}_+$, then for each element $\alpha \in M$ let

$$\zeta (\alpha) := \inf \left\{ \prod_{i=1}^k \zeta (\beta_i) : \beta_1, \ldots , \beta_k \in B, \prod_{i=1}^k \beta_i = \alpha \right\} .$$

Suppose that $\mathcal{X}$ is a non-empty set equipped with a $\sigma$-algebra $\Sigma$. If a monoid $M$ acts on $\mathcal{X}$, then this action is measurable if for every $\alpha \in M$ the map $x \mapsto \alpha (x)$ is measurable. We also use the following notational convention for conditional probabilities: If $F$ is a random event and $\Pr (F) = 0$, then $\Pr (E|F) = 0$ for all events $E$. Now we are ready to state the LAL.

**Theorem 2.1** (Local Action Lemma, [3]). Suppose that $\mathcal{X}$ is a non-empty set equipped with a $\sigma$-algebra $\Sigma$, and let $\mathcal{G} \in \Sigma$. Let $M$ be an at most countable monoid with a generating set $B \subseteq M$. Suppose that $M$ acts measurably on $\mathcal{X}$, and for every $x \in \mathcal{G}$ and $\alpha \in M$ we have $\alpha (x) \in \mathcal{G}$.

Let $\Omega$ be a probability space, and let $X : \Omega \to \mathcal{X}$ be a random variable. For every $x \in \mathcal{X} \setminus \mathcal{G}$ and $\beta \in B$ such that $\beta (x) \in \mathcal{G}$ choose an arbitrary element $g_\beta (x) \in M$ in such a way that the maps $x \mapsto g_\beta (x)$ are measurable. For $\beta \in B$ and $\alpha \in M$ define

$$P(\beta, \alpha) := \sup_{\gamma \in M} \Pr (g_\beta (\gamma (X)) = \alpha | \alpha \beta \gamma (X) \in \mathcal{G}) .$$

Suppose that for some $\alpha \in M$ we have $\Pr (\alpha (X) \in \mathcal{G}) > 0$. If there exists a function $\zeta : B \to \mathbb{R}_+$ such that for every $\beta \in B$ we have

$$\zeta (\beta) \geq 1 + \sum_{\alpha \in M} P(\beta, \alpha) \zeta (\alpha),$$

then $\Pr (X \in \mathcal{G}) > 0$.

3 Graphs with forbidden bipartite subgraph

3.1 Combinatorial lemmata

For this section we assume that a bipartite graph $H$ is fixed. In particular, all constants that we mention depend on $H$. We will use the following version of the Kővari–Sós–Turán theorem.
Theorem 3.1 (Kővari et al. [10]). Let $G$ be a graph with $n$ vertices and $m$ edges that does not contain the complete bipartite graph $K_{k,k}$ as a subgraph. Then $m \leq O(n^{2-1/k})$ (assuming that $n \to \infty$).

Corollary 3.2. There exist positive constants $\alpha$ and $\delta$ such that if a graph $G$ with $n$ vertices and $m$ edges does not contain $H$ as a subgraph, then $m \leq \alpha n^{2-\delta}$.

In what follows we fix the constants $\alpha$ and $\delta$ from the statement of Corollary 3.2. We say that the length of a path $P$ is the number of edges in it. Using Corollary 3.2 we obtain the following.

Lemma 3.3. There is a positive constant $\beta$ such that the following holds. Let $G(V,E)$ be a graph with maximum degree $\Delta$ that does not contain $H$ as a subgraph. Then for any two vertices $u,v \in V$ the number of $uv$-paths of length 3 in $G$ is at most $\beta 2^{3-\delta}$.

Proof. Suppose that $u-x-y-v$ is a $uv$-path of length 3 in $G$. Then $x \in N_{G}(u)$, $y \in N_{G}(v)$, and hence $xy \in E(G[N_{G}(u) \cup N_{G}(v)])$. Note that any edge $xy \in E(G[N_{G}(u) \cup N_{G}(v)])$ can possibly give rise to at most two different $uv$-paths of length 3 (namely $u-x-y-v$ and $u-y-x-v$). Therefore, the number of $uv$-paths of length 3 in $G$ is not greater than $2|E(G[N_{G}(u) \cup N_{G}(v)])|$. Since $|V(G[N_{G}(u) \cup N_{G}(v)])| \leq 2\Delta$, by Corollary 3.2 we have that $|E(G[N_{G}(u) \cup N_{G}(v)])| \leq \alpha (2\Delta)^{2-\delta}$, so the number of $uv$-paths of length 3 in $G$ is at most $(2^{3-\delta}\alpha) \Delta^{2-\delta}$. 

In what follows we fix the constant $\beta$ from the statement of Lemma 3.3. The following fact is crucial for our proof.

Lemma 3.4. Let $G(V,E)$ be a graph with maximum degree $\Delta$ that does not contain $H$ as a subgraph. Then for any edge $e \in E$ and for any integer $k \geq 4$ the number of cycles of length $k$ in $G$ that contain $e$ is at most $\beta 2^{k-2-\delta}$.

Proof. Suppose that $e = uv \in E$. Note that the number of cycles of length $k$ that contain $e$ is not greater than the number of $uv$-paths of length $k-1$. Consider any $uv$-path $u-x_1-\ldots-x_{k-2}-v$ of length $k-1$. Then $u-x_1-\ldots-x_{k-4}$ is a path of length $k-4$, and $x_{k-4}-x_{k-3}-x_{k-2}-v$ is a path of length 3. There are at most $\Delta^{k-4}$ paths of length $k-4$ starting at $u$, and, given a path $u-x_1-\ldots-x_{k-4}$, the number of $x_{k-4}$-paths of length 3 is at most $\beta 2^{3-\delta}$. Hence the number of $uv$-paths of length $k-1$ is at most $\Delta^{k-4} \cdot \beta 2^{3-\delta} = \beta \Delta^{k-2-\delta}$. 

3.2 Probabilistic set-up

We use the following general scheme for turning coloring problems into instances of the LAL. If $A$ and $B$ are sets, then a partial function from $A$ to $B$ (notation: $f : A \to B$) is a map $f : A' \to B$ for some $A' \subseteq A$. Denote the set of all partial functions from $A$ to $B$ by $\text{PF}(A,B)$. For $f \in \text{PF}(A,B)$ let $\text{dom}(f) \subseteq A$ be the domain of $f$. If $\text{dom}(f) = A$, then to emphasize this fact we would sometimes call $f$ a total function from $A$ to $B$. If $A$ is finite and $B$ is at most countable, then $\text{PF}(A,B)$ is at most countable, so we may equip it with the $\sigma$-algebra $\Sigma_{A,B} := \mathcal{P}(\text{PF}(A,B))$ of all its subsets. The power set $\mathcal{P}(A)$ can be turned into a commutative monoid with the multiplication given by the union operation that acts on $\text{PF}(A,B)$ with

$$S(f) = f|_{\text{dom}(f) \setminus S}.$$ 

Clearly, this action is measurable with respect to $\Sigma_{A,B}$. Observe that if $A$ is finite, then as a monoid $\mathcal{P}(A)$ is generated by the set $\{\{a\} : a \in A\}$ of singletons. Slightly abusing the notation, we would indentify the singletons with the corresponding elements of $A$. In particular, we would say that $\mathcal{P}(A)$ is generated by $A$, and write $a(f)$ instead of $\{a\}(f)$.

Let $G(V,E)$ be a graph with maximum degree $\Delta$ that does not contain $H$ as a subgraph. Let $\mathcal{C}$ be a set of colors of cardinality $|\mathcal{C}| = (2 + c)\Delta$. Then the set $\text{PF}(E,\mathcal{C})$ is the set of all partial edge colorings of $G$. Let $\mathcal{G} \subseteq \text{PF}(E,\mathcal{C})$ be the set of partial acyclic colorings. Note that for every $S \in \mathcal{P}(E)$ and $f \in \mathcal{G}$ we have $S(f) \in \mathcal{G}$ (in other words, any restriction of an acyclic coloring is acyclic). Choose a total edge coloring of $G$ uniformly at random. That gives us a random variable $X \in \text{PF}(E,\mathcal{C})$, and we want to show that $\Pr(X \in \mathcal{G}) > 0$ provided that $c = 1 + o(1)$. Also note that $E(X) = \emptyset \in \mathcal{G}$, in particular, $\text{Pr}(E(X) \in \mathcal{G}) = 1 > 0$.

If $f : E \to \mathcal{C}$ and $f \not\in \mathcal{G}$, but $e(f) \in \mathcal{G}$ for some $e \in E$, then one of the two following conditions must hold.
1. There is an edge $h \in \text{dom}(f)$ that is adjacent to $e$ and $f(e) = f(h)$. In this case let $g_*(f) := \emptyset$.

2. There is a cycle $K$ contained in $\text{dom}(f)$ that passes through $e$ and is colored bichromatically by $f$. In this case let $g_*(f) := K'_e$, where $K'_e$ is a subset of $E(K)$ that contains $e$ and all other edges of $K$ except for two arbitrary adjacent edges.

If there is some ambiguity in the above definition of $g_*(f)$ (e.g. if there is more than one bichromatic cycle passing through $e$), then choose any available option.

Let $e \in E$ and $S \in \mathcal{P}(E)$. Then $P(e, S) \neq 0$ only if either $S = \emptyset$ or $S = K'_e$ for an even cycle $K$ that contains $e$. In the former case, $P(e, \emptyset) = \sup_{F \subseteq E} \Pr \left( g_*(X|_{E \setminus F}) = \emptyset \mid X|_{E \setminus \{f(e)\}} \in \mathcal{G} \right)$

\[
\leq \sup_{F \subseteq E} \Pr \left( \text{there is } h \in E \setminus F \text{ that is adjacent to } e \text{ and } X(e) = X(h) \mid X|_{E \setminus \{f(e)\}} \in \mathcal{G} \right)
\]

\[
\leq \frac{2\Delta}{|C|} = \frac{2}{2 + c},
\]

where the last inequality follows from the fact that there can be at most $2\Delta$ (in fact, at most $2(\Delta - 1)$) colors that are “forbidden” for $e$ in $X$.

Now suppose that $S = K'_e$ for some cycle $K$ of length $2t$, where $2t = |S| + 2$. Let $\Theta(S)$ be the set of all cycles $K$ such that $K'_e = S$. Then

\[
P(e, S) = \sup_{F \subseteq E} \Pr \left( g_*(X|_{E \setminus F}) = S \mid X|_{E \setminus \{f(e)\}} \in \mathcal{G} \right)
\]

\[
\leq \sup_{F \subseteq E} \sum_{K \in \Theta(S)} \Pr \left( K \text{ is colored bichromatically in } X|_{E \setminus F} \mid X|_{E \setminus \{f(e)\}} \in \mathcal{G} \right)
\]

\[
\leq \sum_{K \in \Theta(S)} \frac{1}{|C|^{2t-2}} = \frac{1}{(2 + c) \Delta^{2t-2}},
\]

where the last inequality follows from the fact that a given acyclic coloring of $E \setminus (F \cup K'_e)$ can be extended to a coloring of $E \setminus F$ in $|C|^{2t-2}$ ways, but $K$ is colored bichromatically in only one of them (the only bichromatic coloring of $K$ is determined by the two colors used on the edges in $E(K) \setminus K'_e$).

Now we can apply the LAL. Assuming that a function $\zeta : E \to \mathbb{R}_+$ is actually a constant $\zeta \in \mathbb{R}_+$ independent of $e \in E$, and using Lemma 3.4 it is enough to show

\[
\zeta \geq 1 + \sum_{t=2}^{\infty} \frac{\beta \Delta^{t-2} \delta \zeta^{t-2}}{(2 + c) \Delta^{t-2}} + \frac{2\zeta}{2 + c} = 1 + \beta \Delta^{-\delta} \frac{\zeta/(2 + c)}{1 - \zeta/(2 + c)} + \frac{2\zeta}{2 + c},
\]

where the last equality holds under the assumption that $\zeta/(2 + c) < 1$. If we denote $y = \zeta/(2 + c)$, then \(2\) turns into

\[
c \geq \frac{1}{y} + \beta \Delta^{-\delta} \frac{y^2}{1 - y^2}.
\]

Now if $c = 1 + \varepsilon$ for any given $\varepsilon > 0$, then we can take $1/y = 1 + \varepsilon/2$. For this particular value of $y$ we have

\[
\beta \Delta^{-\delta} \frac{y^2}{1 - y^2} \xrightarrow{\Delta \to \infty} 0,
\]

so for $\Delta$ large enough $\beta \Delta^{-\delta} y^2/(1 - y^2) \leq \varepsilon/2$, and \(3\) is satisfied. This observation completes the proof of Theorem 1.2.

4 Graphs with large girth

4.1 Breaking short cycles

The proof of Theorem 4.1 proceeds in two steps. Assuming that the girth of $G$ is large enough, we first show that there is a proper edge coloring of $G$ by $(2 + \varepsilon)/2 + \Delta$ colors with no “short” bichromatic cycles.
Lemma 4.3. For every \( L \) bichromatic cycles of length at most 2

Subsection 3.2 for the notation used), but this time

We work in a probabilistic setting similar to the one used in the proof of Theorem 1.2 (see

at most \( 2 \)

G

Lemma 4.1. Let \( G(V, E) \) be a graph with maximum degree \( \Delta \) and girth \( g > 2r \), where \( r \geq 2 \). Then

for any two vertices \( u, v \in V \) the number of \( uv \)-paths of length \( r \) in \( G \) is at most 1.

Proof. If there are two \( uv \)-paths of length \( r \), then their union forms a closed walk of length \( 2r \), which

means that \( G \) contains a cycle of length at most \( 2r \). ■

Lemma 4.2. Let \( G(V, E) \) be a graph with maximum degree \( \Delta \) and girth \( g > 2r \), where \( r \geq 2 \). Then

for any edge \( e \in E \) and for any integer \( k \geq 4 \) the number of cycles of length \( k \) in \( G \) that contain \( e \) is at most \( \Delta^{k-r-1} \).

Proof. Suppose that \( e = uv \in E \). Note that the number of cycles of length \( k \) that contain \( e \) is not greater than the number of \( uv \)-paths of length \( k - 1 \). Consider any \( uv \)-path \( u-x_1-\ldots-x_{k-2}-v \) of length \( k - 1 \). Then \( u-x_1-\ldots-x_{k-2} \) is a path of length \( k - r - 1 \), and \( x_{k-r-1}-x_{k-r}-x_{k-2} \) is a path of length \( r \). There are at most \( \Delta^{k-r-1} \) paths of length \( k - r - 1 \) starting at \( u \), and, given a path \( u-x_1-\ldots-x_{k-1} \), the number of \( x_{k-1}v \)-paths of length \( r \) is at most 1. Hence the number of \( uv \)-paths of length \( k - 1 \) is at most \( \Delta^{k-r-1} \). ■

Lemma 4.3. For every \( \varepsilon > 0 \) there exists a positive constant \( a_\varepsilon \) such that the following holds. Let

\( G(V, E) \) be a graph with maximum degree \( \Delta \) and girth \( g > 2r \), where \( r \geq 2 \). Then there is a proper edge coloring of \( G \) using at most \( (2 + \varepsilon)\Delta + o(\Delta) \) colors that contains no bichromatic cycles of length at most \( 2L \), where \( L := a_\varepsilon(r - 2) \log \Delta + 1 \).

Proof. We work in a probabilistic setting similar to the one used in the proof of Theorem 1.2 (see Subsection 3.2 for the notation used), but this time \( G \) is the set of all proper edge colorings with no bichromatic cycles of length at most \( 2L \). Then, taking into account Lemma 4.2, (2) turns into

\[
\zeta \geq 1 + \sum_{t=r+1}^{L} \frac{\Delta^{2t-r-1} \zeta^{2t-2}}{(2 + c)\Delta^{2t-2}} + \frac{2\zeta}{2 + c} = 1 + \Delta^{-r+1} \sum_{t=r+1}^{L} \left( \frac{\zeta}{2 + c} \right)^{2t-2} + \frac{2\zeta}{2 + c}. \tag{4}
\]

If \( y = \zeta/(2 + c) \), then (4) becomes

\[
c \geq \frac{1}{y} + \Delta^{-r+1} \sum_{t=r+1}^{L} y^{2t-3}.
\]

Note that if \( y > 1 \) and \( L \leq \Delta \), we have

\[
\sum_{t=r+1}^{L} y^{2t-3} \leq \sum_{t=r+1}^{L} y^{2L-3} = (L - r)y^{2L-3} \leq \Delta y^{2L-3},
\]

so it is enough to get

\[
c \geq \frac{1}{y} + \Delta^{-r+2} y^{2L-3}.
\]

Now take \( y = 2/\varepsilon \) and \( c = \varepsilon \). We need

\[
\frac{\varepsilon}{2} \geq \Delta^{-r+2} \left( \frac{2}{\varepsilon} \right)^{2L-3},
\]

i.e.

\[
L \leq \left( 2 \log \frac{2}{\varepsilon} \right)^{-1} (r - 2) \log \Delta + 1,
\]

and we are done. ■
4.2 Breaking long cycles

To deal with “long” cycles we need a different random procedure. A similar procedure was analysed in [13] using the LLL.

**Lemma 4.4.** For every $\varepsilon > 0$ there exist positive constants $b_\varepsilon$ and $d_\varepsilon$ such that the following holds. Let $G(V, E)$ be a graph with maximum degree $\Delta$ and let $\varphi : E \to \mathcal{C}$ be a proper edge coloring of $G$. Then there is a proper edge coloring $\psi : E \to \mathcal{C} \cup \mathcal{C}'$ such that

- $|\mathcal{C}'| = \varepsilon \Delta + o(\Delta)$;
- if a cycle is bichromatic in $\psi$, then it was bichromatic in $\varphi$;
- there are no bichromatic cycles of length at least $L$ in $\psi$, where $L := b_\varepsilon \log \Delta + d_\varepsilon$.

**Proof.** Let $\mathcal{C}'$ be a set of colors disjoint from $\mathcal{C}$ with $|\mathcal{C}'| = c\Delta$. Let $\mathcal{X}$ be the set of partial colorings $\psi : E \to \mathcal{C} \cup \mathcal{C}'$ such that for all $e \in \text{dom}(\psi)$ either $\psi(e) \in \mathcal{C}'$, or $\psi(e) = \varphi(e)$ (in other words, we allow recoloring some of the edges using new colors $\mathcal{C}'$). Let $\mathcal{G} \subseteq \mathcal{X}$ be the set of colorings that satisfy the conditions of the lemma. Fix some $0 < p < 1$ and define a random variable $X$ in the following way: For every edge $e \in E$ either do not change its color with probability $1 - p$, or choose for it one of the new colors, each with probability $p/|\mathcal{C}'| = p/(c\Delta)$.

If $\psi \in \mathcal{X} \setminus \mathcal{G}$, but $e(\psi) \in \mathcal{G}$ for some edge $e \in E$, then one of the following situations must happen. Again, these situations are not mutually exclusive. If two or more of them happen simultaneously, then we choose any available option for the definition of $g_e(\psi)$. Also we assume that the function $\zeta : E \to \mathbb{R}_+$ is a constant $\zeta \in \mathbb{R}_+$.

**Case 1:** There is an edge $h \in \text{dom}(\psi)$ that is adjacent to $e$ such that $\psi(h) = \psi(e) \in \mathcal{C}'$. In this case let $g_e(\psi) := h$. Let us estimate the corresponding terms on the right-hand side of (1). For an edge $h \in E$ incident to $e$ we have

$$P(e, h) \leq \sup_{F \subseteq E} \Pr(X(h) = X(e) \in \mathcal{C}' \mid X|_{E \setminus (F \cup \{e, h\})} \in \mathcal{G}) = \frac{p^2}{c\Delta}.$$  

Since there are less than $2\Delta$ edges adjacent to any given edge $e$, we get at most

$$2\Delta \cdot \frac{p^2}{c\Delta} \cdot \zeta^2 = 2c \left(\frac{K}{c}\right)^2$$  

on the right-hand side of (1).

**Case 2:** There is a cycle $K$ of length at least $L$ such that $e \in E(K) \subseteq \text{dom}(\psi)$ that is colored bichromatically by $\varphi$ and $\psi(h) = \varphi(h)$ for all $h \in E(K)$. In this case let $g_e(\psi) := E(K)$. Then

$$\sup_{F \subseteq E} \Pr(X(h) = \varphi(h) \text{ for all } h \in E(K) \mid X|_{E \setminus (F \cup E(K))} \in \mathcal{G})$$

$$= (1 - p)^{|E(K)|}.$$  

Note that there are less than $\Delta$ cycles that are bichromatic in $\varphi$ containing any given edge $e$ (because the second edge on the cycle specifies it uniquely). Hence if we assume that $(1 - p)\zeta < 1$, then we get at most

$$\Delta((1 - p)\zeta)^L$$  

on the right-hand side of (1).

**Case 3:** There is a cycle $K$ of length $2t$ such that $e \in E(K) \subseteq \text{dom}(\psi)$ that is colored bichromatically by $\psi$ and $\psi(h) \in \mathcal{C}'$ for all $h \in E(K)$. In this case let $g_e(\psi) := E(K)$. We have

$$\sup_{F \subseteq E} \Pr(K \text{ is bichromatic in } X \text{ and } X(h) \in \mathcal{C}' \text{ for all } h \in E(K) \mid X|_{E \setminus (F \cup E(K))} \in \mathcal{G})$$

$$= \frac{p^{2t}}{(c\Delta)^{2t-2}}.$$  

7
There are at most $\Delta^{2t-2}$ cycles of length $2t$ containing a given edge $e$. Hence if we assume that $p\zeta/c < 1$, then we get at most

$$\sum_{t=2}^{\infty} \Delta^{2t-2} \cdot \frac{p^t}{(c\Delta)^{2t-2}} \cdot \zeta^{2t} = c^2 \sum_{t=2}^{\infty} \left( \frac{p\zeta}{c} \right)^{2t} = c^2 \frac{(p\zeta/c)^4}{1 - (p\zeta/c)^2}$$

(7)
on the right-hand side of (1).

**Case 4:** There is a cycle $K$ of length $2t$ such that $e \in E(K) \subseteq \text{dom}(\psi)$ that is colored bichromatically by $\psi$ and half of the edges of $K$ retain the color assigned to them by $\varphi$, while the other edges are colored by a color from $C'$. Moreover, $\psi(e) = \varphi(e)$. In this case let $g_e(\psi) := E(K)$. We have

$$\sup_{F \subseteq E} \Pr(\text{Case 4 happens with } K \mid |X[K \setminus (F \cup E(K))]| \in \mathcal{G}) = \frac{p^t(1-p)^t}{(c\Delta)^{t-1}}.$$ 

There are at most $\Delta^{t-1}$ cycles $K$ of length $2t$ containing a given edge $e$ such that every second edge in $K$ is colored the same by $\varphi$. (See Fig. 1) Solid edges retain their color from $\varphi$ (this color must be the same for all of them). Arrows indicate $t-1$ edges that must be specified in order to fully determine the cycle. Hence if we assume that $(1-p)\zeta < 1$ and $p\zeta/c < 1$, then we get at most

$$\sum_{t=2}^{\infty} \Delta^{t-1} \cdot \frac{p^t(1-p)^t}{(c\Delta)^{t-1}} \cdot \zeta^{2t} \leq c \sum_{t=2}^{\infty} \left( \frac{p\zeta}{c} \right)^{t} = c \frac{(p\zeta/c)^2}{1 - p\zeta/c}$$

(8)
on the right-hand side of (1).

**Case 5:** There is a cycle $K$ of length $2t$ such that $e \in E(K) \subseteq \text{dom}(\psi)$ that is colored bichromatically by $\psi$ and half of the edges of $K$ retain the color assigned to them by $\varphi$, while the other edges are colored by a color from $C'$. Moreover, $\psi(e) \in C'$. In this case let $g_e(\psi) := E(K)$. The same analysis as in Case 4 (see Fig. 2) gives as at most

$$c \frac{(p\zeta/c)^2}{1 - p\zeta/c}$$

(9)
on the right-hand side of (1), provided that $(1-p)\zeta < 1$ and $p\zeta/c < 1$.

Figure 1: Case 4

Figure 2: Case 5

Adding together (5), (6), (7), (8), and (9), it is enough to have the following inequality:

$$\zeta \geq 1 + 2c(p\zeta/c)^2 + \Delta((1-p)\zeta) + c^2 \frac{(p\zeta/c)^4}{1 - (p\zeta/c)^2} + 2c \frac{(p\zeta/c)^2}{1 - p\zeta/c}. $$

(10)

under the assumption that $(1-p)\zeta < 1$ and $p\zeta/c < 1$. Denote $y := p\zeta/c$. Then (10) turns into

$$\frac{c}{p} \geq \frac{1}{y} + 2cy + c^2 \frac{y^3}{1-y^2} + 2c \frac{y}{1-y} + \frac{\Delta}{y} \left( \frac{c(1-p)}{p} y \right)^L,$$
and we have the conditions $y < 1$ and $y < p/(c(1-p))$. Let $c = \varepsilon$. We can assume that $\varepsilon$ satisfies

$$2\varepsilon^2 + \frac{\varepsilon^5}{1-\varepsilon^2} + \frac{2\varepsilon^2}{1-\varepsilon} \leq \frac{\varepsilon}{4}.$$ 

Take $y = \varepsilon$. Then it is enough to have

$$\frac{\varepsilon}{p} \geq \frac{1}{\varepsilon} + \frac{\Delta}{\varepsilon} \left( \frac{\varepsilon^2(1-p)}{p} \right)^L. \quad (11)$$

Let $p_\varepsilon := \varepsilon/(\varepsilon/2 + 1/\varepsilon)$. Note that

$$\frac{p_\varepsilon}{\varepsilon(1-p_\varepsilon)} = \frac{1}{(\frac{\varepsilon}{2} + \frac{1}{2})\left(1 - \varepsilon/(\frac{\varepsilon}{2} + \frac{1}{2})\right)} = \frac{1}{\frac{\varepsilon}{2} - \frac{1}{2}} > \varepsilon,$$

so this choice of $p_\varepsilon$ does not contradict our assumptions. Then (11) becomes

$$\frac{\varepsilon}{4} \geq \frac{\Delta}{\varepsilon} \left( \frac{\varepsilon^2(1-p_\varepsilon)}{p_\varepsilon} \right)^L,$$

which is true provided that

$$L \geq \left( \log \left( \frac{p_\varepsilon}{\varepsilon^2(1-p_\varepsilon)} \right) \right)^{-1} \left( \log \Delta + \log \frac{4}{\varepsilon^2} \right),$$

and we are done. \hfill \Box

### 4.3 Finishing the proof

To finish the proof of Theorem 1.5 fix $\varepsilon > 0$. By Lemma 4.3 if $g(G) > 2r$, where $r \geq 2$, then there is a proper edge coloring $\varphi$ of $G$ using at most $(2 + \varepsilon/2)\Delta + o(\Delta)$ colors that contains no bichromatic cycles of length at most $L_1 := 2a_{2\varepsilon/2}(r-2)\log \Delta + 2$. Applying Lemma 4.4 to this coloring gives a new coloring $\psi$ that uses at most $(2 + \varepsilon)\Delta + o(\Delta)$ colors and contains no bichromatic cycles of length at most $L_1$ (because there were no such cycles in $\varphi$) and at least $L_2 := b_{2\varepsilon/2} \log \Delta + d_{2\varepsilon/2}$. If $r - 2 > b_{2\varepsilon/2}/(2a_{2\varepsilon/2})$ and $\Delta$ is large enough, then $L_1 > L_2$, and $\psi$ must be acyclic. This observation completes the proof.

### 5 Concluding remarks

We conclude with some remarks on why it seems difficult to get closer to the desired bound $a'(G) \leq \Delta(G) + 2$ using the same approach as in the proof of Theorem 1.5. Observe that in the proof of Theorem 1.5 (specifically in the proof of Lemma 4.3) we reserve $2\Delta$ colors for making a coloring proper and use only $c\Delta$ “free” colors to make this coloring acyclic. Essentially, Theorem 1.5 asserts that $c$ can be made as small as $\varepsilon + o(1)$ provided that $g(G)$ is big enough. It means that the only way to improve the linear term in our bound is to reduce the number of reserved colors, in other words, to implement in the proof some Vizing-like argument. Unfortunately, we do not know how to prove Vizing’s theorem by a relatively straightforward application of the LLL (or any analog of it). On the other hand, as was mentioned in the Introduction, using a more sophisticated technique (similar to the one used by Kahn [9] in his celebrated proof that every graph is $(1 + o(1))\Delta$-edge-choosable) Cai et al. managed to obtain the bound $a'(G) \leq (1 + \varepsilon)\Delta + o(\Delta)$, which is very close to the desired bound $a'(G) \leq \Delta(G) + 2$.

### Acknowledgements

I would like to thank Louis Esperet and Rémi de Verclos for the observation that the proof of Theorem 1.2 works not only for excluding $C_4$, but for any bipatite graph $H$ as well.

This work is supported by the Illinois Distinguished Fellowship.
References

[1] N. Alon, C. McDiarmid, and B. Reed. Acyclic coloring of graphs. Random structures and algorithms, Volume 2, No. 3, 1991. Pages 277–288.

[2] N. Alon, B. Sudakov, and A. Zaks. Acyclic edge colorings of graphs. J. Graph Theory, Volume 37, 2001. Pages 157–167.

[3] A. Bernshteyn. The Local Action Lemma. arXiv:1410.1591.

[4] R. Bissacot, R. Fernández, A. Procacci, and B. Scoppola. An improvement of the Lovász Local Lemma via cluster expansion. J. Combinatorics, Probability and Computing, Volume 20, Issue 5, 2011. Pages 709–719.

[5] X.S. Cai, G. Perarnau, B. Reed, and A.B. Watts. Acyclic edge colourings of graphs with large girth. arXiv:1411.3047.

[6] L. Esperet, A. Parreau. Acyclic edge-coloring using entropy compression. European J. Combin., Volume 34, Issue 6, 2013. Pages 1019–1027.

[7] J. Fiamčík. The acyclic chromatic class of a graph (in Russian). Math. Slovaca, Volume 28, 1978. Pages 139–145.

[8] B. Grünbaum. Acyclic colorings of planar graphs. Israel Journal of Mathematics, Volume 14, Issue 4, 1973. Pages 390–408.

[9] J. Kahn. Asymptotics of the list chromatic index for multigraphs. Random Structures & Algorithms, Volume 17, Issue 2, 2000. Pages 117–156.

[10] T. Kövari, V. Sós, and P. Turán. On a problem of K. Zarankiewicz. Colloquium Math. 3, (1954). Pages 50–57.

[11] M. Molloy, B. Reed. Further algorithmic aspects of the Local Lemma. Proceedings of the 30th Annual ACM Symposium on Theory of Computing, 1998. Pages 524–529.

[12] R. Moser, G. Tardos. A constructive proof of the general Lovász Local Lemma. J. ACM, Volume 57, Issue 2, 2010.

[13] R. Muthu, N. Narayanan, and C.R. Subramanian. Improved bounds on acyclic edge colouring. J. Discrete Math., Volume 307, Issue 23, 2007. Pages 3063–3069.

[14] S. Ndreca, A. Procacci, and B. Scoppola. Improved bounds on coloring of graphs. European J. Combin., Volume 33, Issue 4, 2012. Pages 592–609.