Technical Report

TR-2010-020

A note on Thomassen's conjecture

by

Domingos Dellamonica Jr., Vojtech Rodl

Mathematics and Computer Science

EMORY UNIVERSITY

A note on Thomassen's conjecture

Domingos Dellamonica Jr.[1], Vojtěch Rödl[2]

Emory University – Department of Mathematics and Computer Science
400 Dowman Dr, W401
Atlanta, GA 30322
USA

Abstract

In 1983 C. Thomassen conjectured that for every $k, g \in \mathbb{N}$ there exists d such that any graph with average degree at least d contains a subgraph with average degree at least k and girth at least g. A result of Pyber–Rödl–Szemerédi implies that the conjecture is true for every graph G with average $d(G) \geq c_{k,g} \log \Delta(G)$.

We strengthen this and show that the conjecture holds for every graph G with average $d(G) \geq \alpha(\log\log \Delta(G))^{\beta}$ for some constants α, β depending on k and g.

1. Introduction

An old question of P. Erdős and A. Hajnal [3, 4] asks whether for every k and g there exists $\chi = \chi(k, g)$ such that any graph with chromatic number at least χ contains a subgraph with chromatic number at least k and girth (length of the shortest cycle) at least g. It was shown in [8] that for any k and n there exists $\varphi(k, n)$ such that, if G is a graph with chromatic number at least $\varphi(k, n)$, then G contains either a complete subgraph on n vertices or a triangle free k-chromatic subgraph. This result implies an affirmative answer to the case $g = 4$, while all other cases remain open to this date.

In 1983, C. Thomassen [9] stated a related conjecture.

Conjecture 1. *For every k and g there exists $d = d(k, g)$ such that any graph G with average degree $d(G) = \frac{2|E(G)|}{|V(G)|}$ at least d contains a subgraph $H \subset G$ with $d(H) \geq k$ and $\mathrm{girth}(H) \geq g$.*

Conjecture 1 is known to hold for $g \leq 6$ due to a result of Kühn and Osthus [5] (in [2] the authors, Koubek, and Martin gave an alternative proof of this result). It is also known that any regular graph with sufficiently large degree satisfies the conclusions of the conjecture:

Email addresses: `ddellam@mathcs.emory.edu` (Domingos Dellamonica Jr.), `rodl@mathcs.emory.edu` (Vojtěch Rödl)

[1]Supported by a CAPES/Fulbright scholarship.

[2]Partially supported by NSF grant DMS0800070.

Proposition 2 ([5]). *For any k and g there exists $D = D(k, g)$ such that any D-regular graph G contains a subgraph H with $d(H) \geq k$ and* $\mathrm{girth}(G) \geq g$.

Determining sufficient conditions for a graph to contain a regular subgraph is therefore relevant to this conjecture. In this direction, Pyber, Rödl, and Szemerédi proved the following result.

Theorem 3 ([7]). *For every $r \geq 1$ there exists $c > 0$ and $d_0 > r$ for which the following holds.*

Suppose that G is a graph with average degree $d(G) \geq \max\{c \log \Delta(G), d_0\}$. Then G contains an r-regular subgraph.

In view of Theorem 3 and Proposition 2, for any fixed k and g, Conjecture 1 holds for graphs G of average degree $d(G) \geq c \log \Delta(G)$. Namely, we have the following corollary.

Corollary 4. *For every $k \geq 1$, $g \geq 3$ there exists $c > 0$ and $d_0 > k$ for which the following holds.*

Suppose that G is a graph with average degree

$$d(G) \geq \max\big\{c \log \Delta(G), d_0\big\}.$$

Then G contains a subgraph H with $d(H) \geq k$ and $\mathrm{girth}(H) \geq g$.

Here we will strengthen this corollary as follows.

Theorem 5 (Main result). *For every $k \geq 1$, $g \geq 3$ there exists $\alpha, \beta > 0$ and $d_0 > k$ for which the following holds.*

Suppose that G is a graph with average degree

$$d(G) \geq \max\big\{\alpha\big(\log\log \Delta(G)\big)^{\beta}, d_0\big\}.$$

Then G contains a subgraph H with $d(H) \geq k$ and $\mathrm{girth}(H) \geq g$.

Remark 6. In [7] it is also shown that there are graphs G with average degree $d(G) \geq c \log\log \Delta(G)$ that do not contain any r-regular subgraphs ($r \geq 3$).

In Theorem 5 the subgraph H is not necessarily regular. In particular, our result improves Corollary 4 but not Theorem 3.

Definition 7 (Notation). We denote by $\deg_G(v)$ the degree of a vertex v in a graph G and let $N_G(v)$ denote its neighborhood. For a set $X \subset V(G)$ we denote by $\deg_G(v, X)$ the number of neighbors of v which are contained in the set X. For disjoint sets $A, B \subset V(G)$ let $e_G(A, B)$ denote the number of edges in G between A and B. Denote by $G[A, B]$ the bipartite graph containing all the edges in G between A and B.

2. A result on Thomassen's conjecture

In order to prove Theorem 5 we will require some results for hypergraphs.

Definition 8. A hypergraph is called *linear* if no two edges intersect in more than one vertex.

The *degree* of a vertex v in a hypergraph $\mathcal{H}$ is the number of edges $e \in \mathcal{H}$ incident to v.

A *cycle* (of length $i \geq 3$) in a hypergraph $\mathcal{H}$ is a sequence of vertices and edges $(v_0, e_0, v_1, e_1, \ldots, v_{i-1}, e_{i-1})$ where $v_0, \ldots, v_{i-1} \in V(\mathcal{H})$ are distinct vertices; $e_0, \ldots, e_{i-1} \in \mathcal{H}$ are distinct edges; and $v_\ell \in e_{\ell-1} \cap e_\ell$ for all $\ell \in \mathbb{Z}_i$.

Our key idea is the simple observation that an "almost regular" linear hypergraph of large average degree contains a sub(hyper)graph with large average degree and large girth.

Lemma 9. *Let $g \geq 3$, $k \geq 1$ and $d \geq 2$ be given.*

Suppose that $\mathcal{H}$ is a linear hypergraph with edges of size at most d,

$$\delta(\mathcal{H}) \geq (2kd^3)^{2g-2} \quad \textit{and} \quad \Delta(\mathcal{H}) \leq \delta(\mathcal{H})^{\frac{2g-1}{2g-2}}.$$

Then $\mathcal{H}$ contains a spanning subgraph with average degree at least k and girth at least g.

The proof of Lemma 9 is based on the Probabilistic Method (see [1]) and will be given in Section 2.1. It is a straightforward extension of a similar result known to hold for graphs.

Remark 10. Any improvement in the upper bound on $\Delta(\mathcal{H})$ in the statement of Lemma 9 yields an improvement on Theorem 5. For instance, if one could allow $\Delta(\mathcal{H}) \leq e^{c\delta(\mathcal{H})}$ then Theorem 5 would hold with $\log^* \Delta(G)$ (the iterated log) in place of $\log\log \Delta(G)$.

Unfortunately, a straightforward attempt to improve Lemma 9 by means of extending Theorem 3 to hypergraphs is not possible. More formally, there are hypergraphs $\mathcal{H}$ with, say, $d(\mathcal{H}) \geq \Delta(\mathcal{H})^c$, that contain no almost regular subgraph of large average degree [10].

In §2.2 we will prove Theorem 5 by applying a strategy inspired by the idea originally used in [5]: In a graph G of large average degree we either find a large complete bipartite subgraph or a C_4-free bipartite subgraph of large average degree. If a complete bipartite graph is found then, by Proposition 2, we obtain a subgraph of large girth and large average degree. Otherwise we apply Lemma 11 below to the C_4-free subgraph to obtain a subgraph with large girth and average degree.

Lemma 11. *For every k and g there exists $\gamma > 0$, $d_{\min} > k$ such that the following holds.*

Let G be a C_4-free bipartite graph with classes A, B, $|A| \le |B| = n$ such that

$$\text{for all } w \in B, \quad \deg_G(w) = d \ge \max\{\gamma \log\log \Delta(G), d_{\min}\}. \tag{1}$$

Then there exists $H \subset G$ such that $\mathrm{girth}(H) \ge g$ *and* $d(H) \ge k$.

Proof. Set $\varepsilon = 1/(2g)$ and $\gamma = \frac{8k}{\log(1+\varepsilon)}$. Let $c = c(k,g) > k$ and $d_0 = d_0(k,g) > k$ be the constants from Corollary 4 and let $d_{\min} \ge 2d_0$ be large enough so that any $d \ge d_{\min}$ satisfies (for k, g and c fixed)

$$\Delta := e^{\frac{d}{2c}} > 4d^{4g}. \tag{2}$$

Consider the following partition $A = A_0 \cup A_1 \cup \cdots \cup A_t$, where

$$A_0 = \{v \in A \,:\, \deg(v) \le \Delta\}, \quad \text{and}$$

$$A_j = \{v \in A \,:\, \Delta^{(1+\varepsilon)^{j-1}} < \deg(v) \le \Delta^{(1+\varepsilon)^j}\},$$

for $j = 1, \dots, t$.

We may assume that $\Delta^{(1+\varepsilon)^{t-1}} < \Delta(G) \le e^{e^{d/\gamma}}$ since otherwise $A_t = \emptyset$. Therefore

$$t \le \frac{d}{\gamma \log(1+\varepsilon)} = \frac{d}{8k}$$

by our choice of γ.

Suppose first that $e_G(A_0, B) \ge dn/2$. In this case, the graph $G_0 = G[A_0, B]$ has average degree $d(G_0) \ge d/2 \ge d_{\min}/2 \ge d_0$ and maximum degree $\Delta(G_0) \le \Delta = e^{\frac{d}{2c}}$. Consequently,

$$d(G_0) \ge \frac{d}{2} \ge \max\{c \log \Delta(G_0), d_0\}.$$

Applying Corollary 4 to the graph G_0 yields $H \subset G_0$ with $d(H) \ge k$ and $\mathrm{girth}(H) \ge g$.

Hence, let us assume that $e_G(A_0, B) < dn/2$. By the Pigeon-hole principle, there exists $j \in [t]$ such that $e_G(A_j, B) \ge dn/(2t)$. In the graph $G_j = G[A_j, B]$ the average degree of the vertices in the class A_j must be at least

$$D = \Delta^{(1+\varepsilon)^{j-1}} \tag{3}$$

and the maximum degree at most $D^{1+\varepsilon}$ while the average degree of the vertices in the class B is at least $d/(2t) \ge 4k$ and the maximum degree at most d (since by assumption $\deg_G(w) = d$ for all $w \in B$).

Next we remove sequentially all the vertices of B that have degree less than k in G_j and all the vertices of A_j that have degree less then $D/4$ in G_j. Once this process ends, the resulting graph $\tilde{G} \subset G_j$ is non-empty since we may remove at most $k\,|B| + \frac{D}{4}|A_j| \le |G_j|/2$ edges from G_j. Moreover, the graph $\tilde{G}$ is bipartite withs classes $X \subset A_j$ and $Y \subset B$ such that

$$\min\{\deg_{\tilde{G}}(x) \,:\, x \in X\} \ge D/4 \quad \text{and} \quad \min\{\deg_{\tilde{G}}(y) \,:\, y \in Y\} \ge k. \tag{4}$$

Define a hypergraph $\mathcal{H}$ with vertex set X and edges $\{N_{\tilde{G}}(y) : y \in Y\}$. Notice that $\deg_{\mathcal{H}}(x) = \deg_{\tilde{G}}(x)$ for all $x \in X$ and that $\mathcal{H}$ is linear because $\tilde{G} \subset G$ and G is C_4-free by assumption. Moreover, by (1) and (4), every edge in $\mathcal{H}$ has between k and d vertices. By our choice of $d_{\min}$, we have

$$\delta(\mathcal{H}) \overset{(4)}{\geq} D/4 \overset{(3)}{\geq} \Delta/4 = e^{\frac{d}{2c}}/4 \overset{(2)}{\geq} d^{4g} \geq (2kd^3)^{g-2},$$

and by our choice of ε, we have

$$\Delta(\mathcal{H}) \leq D^{1+\varepsilon} = D^{1+1/2g} \leq (D/4)^{\frac{g-1}{g-2}} \leq \delta(\mathcal{H})^{\frac{g-1}{g-2}}.$$

We now apply Lemma 9 to $\mathcal{H}$ and obtain $\mathcal{H}^* \subset \mathcal{H}$ such that the girth of $\mathcal{H}^*$ is at least $g/2$ and the average degree of $\mathcal{H}^*$ is at least k. Let $X^* = V(\mathcal{H}^*) \subset V(\mathcal{H}) = X$ and $Y^* \subset Y$ be the set of vertices of Y corresponding to edges of $E(\mathcal{H}^*)$, more precisely,

$$E(\mathcal{H}^*) = \{N_{\tilde{G}}(y) : y \in Y^*\}.$$

We claim that the induced subgraph $H = \tilde{G}[X^*, Y^*] \subset G$ has average degree $d(H) \geq k$. This follows because

$$e(H) = \sum_{v \in X^*} \deg_H(v) = \sum_{v \in V(\mathcal{H}^*)} \deg_{\mathcal{H}^*}(v) \geq k\,|V(\mathcal{H}^*)| = k\,|X^*|$$

and

$$e(H) = \sum_{y \in Y^*} \deg_H(y) \overset{(4)}{\geq} k\,|Y^*|,$$

which implies that $2e(H) \geq k\,|V(H)|$ and thus $d(H) \geq k$.

The graph H has girth at least g because $\mathcal{H}^*$ has girth at least $g/2$. Therefore H satisfies the conclusions of the lemma.

2.1. Proof of Lemma 9

For given $g \geq 3$, $k \geq 1$ and $d \geq 2$, let $\mathcal{H}$ be a hypergraph satisfying the conditions of Lemma 9. Set $\delta = \delta(\mathcal{H})$ and $\varepsilon = \frac{1}{2g-2}$. Notice that

$$\Delta(\mathcal{H}) \leq \delta^{\frac{2g-1}{2g-2}} = \delta^{1+\varepsilon} \quad \text{and} \quad \delta^{\varepsilon} \geq 2kd^3.$$

Denote by X the vertex set of $\mathcal{H}$.

For all $3 \leq i \leq g$, let N_i be the number of cycles of length i in $\mathcal{H}$. We claim that

$$N_i \leq |X|\,(\delta^{1+\varepsilon} d)^{i-1}. \tag{5}$$

Indeed, every cycle may be described by a sequence $(v_0, e_0, v_1, e_1, \ldots, v_{i-1}, e_{i-1})$, where $v_0, \ldots, v_{i-1} \in X$ are distinct vertices and $e_0, \ldots, e_{i-1} \in \mathcal{H}$ are distinct edges satisfying $v_\ell \in e_{\ell-1} \cap e_\ell$ for all ℓ (with indices in $\mathbb{Z}_i$). The number of choices for v_0 is $|X|$, then there are at most $\Delta(\mathcal{H}) \leq \delta^{1+\varepsilon}$ choices for $e_0 \ni v_0$.

Within e_0 one must choose $v_1 \neq v_0$, which can be done in at most $|e_0| - 1 < d$ ways. Repeating this argument we observe that there are at most $\delta^{1+\varepsilon} d$ choices for $(e_\ell, v_{\ell+1})$, $\ell = 0, 1, \ldots, i-2$. Once v_{i-1} is determined, one must choose e_{i-1} containing both v_0 and v_{i-1}. Given the assumption that $\mathcal{H}$ is linear, there can be at most one such edge e_{i-1}. Hence (5) holds.

Let

$$p = \delta^{\varepsilon-1}/(2d^2) \in (0,1)$$

and $\mathcal{H}'$ be the hypergraph obtained by selecting each edge of $\mathcal{H}$ to be in $\mathcal{H}'$ independently and uniformly with probability p. Notice that

$$\mathbf{E}|\mathcal{H}| \geq p\frac{1}{d}\sum_{v\in X} \deg_{\mathcal{H}}(v) \geq \frac{p\,|X|\,\delta}{d}. \tag{6}$$

For all $3 \leq i \leq g$, consider the number N'_i of cycles of length i in $\mathcal{H}'$. The probability that a fixed cycle of length i in $\mathcal{H}$ belongs to $\mathcal{H}'$ is p^i. Hence, we have $\mathbf{E}N'_i = p^i N_i$. By linearity of expectation

$$\mathbf{E}\left[\sum_{i=3}^{g} N'_i\right] = \sum_{i=3}^{g} p^i N_i \overset{(5)}{\leq} p\,|X| \sum_{i=3}^{g} (p\delta^{1+\varepsilon} d)^{i-1} \leq 2p\,|X|(p\delta^{1+\varepsilon}d)^{g-1}, \tag{7}$$

where the last inequality follows since the sum is a geometric series and $p\delta^{1+\varepsilon}d > 2$.

By our choice of p, we have $p\delta/d = \delta^{\varepsilon}d^{-3}/2$ and since $\delta^{\varepsilon} \geq 2kd^3$, it follows that $p\delta/d \geq k$. Consequently, by (6) and (7),

$$\begin{aligned}\mathbf{E}\left[|\mathcal{H}'| - \sum_{i=3}^{g} N'_i\right] &\geq p\,|X|\left(\frac{\delta}{d} - 2\big(p\delta^{1+\varepsilon}d\big)^{g-1}\right)\\ &\geq p\,|X|\left(\frac{\delta}{d} - 2\left(\frac{\delta^{2\varepsilon}}{2d}\right)^{g-1}\right)\\ &\geq p\,|X|\frac{\delta}{2d} \geq \frac{k}{2}\,|X|.\end{aligned} \tag{8}$$

We conclude (by the first moment method) that there exists a spanning hypergraph $\mathcal{H}' \subset \mathcal{H}$ which satisfies the inequality (8) above. In particular, by deleting one edge incident to each cycle of length $3 \leq i \leq g$ in $\mathcal{H}'$ we obtain a spanning subgraph $\mathcal{H}^*$ with at least $k\,|X|/2$ edges and girth at least g.

2.2. Proof of Theorem 5

A slight adaptation of Theorem 2 from [5] yields the following.

Lemma 12. *Let $r \geq 2^{16}$, $s > 1$ be integers. Then every graph of average degree at least $sr^{2\cdot 11^s}$ either contains a complete bipartite graph $K_{s,s}$ or a C_4-free bipartite subgraph with average degree at least r.* □

For given k and g, let γ and $d_{\min}$ be the constants in Lemma 11. Also let $s = (16k)^{2g-2}$, $\beta = 2 \cdot 11^s$, and $\alpha = s\,\gamma^\beta$. Set d_0 large enough so that

$$r := 2\gamma \log\log \Delta(G) \geq 2\gamma \log\log d_0 \geq \max\{2d_{\min}, 2^{16}\}.$$

Suppose that G is a graph with average degree

$$d(G) \geq \alpha\big(\log\log \Delta(G)\big)^\beta = sr^{2\cdot 11^s}.$$

By Lemma 12 the graph G either contains

(i) The complete bipartite graph $K_{s,s}$ or

(ii) A C_4-free bipartite subgraph G' with $d(G') \geq r$.

In case (i), by our choice of $s = (16k)^{2g-2}$, we may apply Lemma 9 with $d = 2$ to the s-regular graph $K_{s,s}$ to obtain a graph $H \subset K_{s,s} \subset G$ with $d(H) \geq k$ and $\mathrm{girth}(H) \geq g$.

In case (ii), we will use the fact that every bipartite graph of average degree r contains an $(r/2)$-half-regular graph, namely, a graph in which every vertex in the larger class has the same degree $r/2$ (see [6, Lemma 3]). Applying this fact to G' yields a subgraph $G^* \subset G'$ with classes A and B such that $|B| \geq |A|$ and every vertex in B has degree $d \geq \frac{1}{2}d(G') \geq \frac{r}{2}$.

By our choice of r,

$$d \geq \max\{\gamma \log\log \Delta(G^*), d_{\min}\}.$$

Applying Lemma 11 to the C_4-free graph $G^* \subset G$ and obtain a subgraph $H \subset G^*$ with average degree at least k and girth at least g. This concludes the proof of Theorem 5.

3. Concluding remarks

In this paper we have shown that even graphs with a substantial gap between average and maximum degrees must contain a subgraph of large girth and average degree. Indeed, the bound we obtained in Theorem 5, $d(G) \leq \alpha\big(\log\log \Delta(G)\big)^\beta$, is asymptotically stronger than the previously known bound given by Corollary 4. Short of solving Conjecture 1 completely, a very natural problem is to improve the bound on the gap between average and maximum degrees that is sufficient for the conjecture to hold. As a step in this direction, one could perhaps attempt proving Lemma 9 with a weaker assumption on $\Delta(\mathcal{H})$.

References

[1] N. Alon and J. H. Spencer. *The probabilistic method.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

[2] D. Dellamonica, V. Koubek, D. M. Martin, and V. Rödl. On a conjecture of Thomassen concerning subgraphs of large girth. To appear in Journal of Graph Theory.

[3] P. Erdős. Problems and results in chromatic graph theory. In *Proof Techniques in Graph Theory (Proc. Second Ann Arbor Graph Theory Conf., Ann Arbor, Mich., 1968)*, pages 27–35. Academic Press, New York, 1969.

[4] P. Erdős. Some unsolved problems in graph theory and combinatorial analysis. In *Combinatorial Mathematics and its Applications (Proc. Conf., Oxford, 1969)*, pages 97–109. Academic Press, London, 1971.

[5] D. Kühn and D. Osthus. Every graph of sufficiently large average degree contains a C_4-free subgraph of large average degree. *Combinatorica*, 24(1):155–162, 2004.

[6] L. Pyber. Regular subgraphs of dense graphs. *Combinatorica*, 5(4):347–349, 1985.

[7] L. Pyber, V. Rödl, and E. Szemerédi. Dense graphs without 3-regular subgraphs. *J. Combin. Theory Ser. B*, 63(1):41–54, 1995.

[8] V. Rödl. On the chromatic number of subgraphs of a given graph. *Proc. Amer. Math. Soc.*, 64(2):370–371, 1977.

[9] C. Thomassen. Girth in graphs. *J. Combin. Theory Ser. B*, 35(2):129–141, 1983.

[10] J. Verstraete. Personal communication, 2010.