Available online at www.sciencedirect.com
SCIENCE DIRECT®
Advances in Mathematics 203 (2006) 476–496
ADVANCES IN Mathematics
www.elsevier.com/locate/aim

On the Turán number for the hexagon

Zoltan Füredi[a,1], Assaf Naor[b,*], Jacques Verstraëte[c]

[a]Department of Mathematics, University of Illinois at Urbana-Champaign 1409 W. Green Street Urbana, Illinois 61801-2975, USA
[b]Theory Group, Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA
[c]Department of Combinatorics and Optimization, Math Faculty, University of Waterloo, 200 University Avenue West, Waterloo, Ont., Canada N2L 3G1

Received 14 April 2004; accepted 27 April 2005

Communicated by Laszlo Lovasz
Available online 9 June 2005

Abstract

A long-standing conjecture of Erdős and Simonovits is that $\mathrm{ex}(n, C_{2k})$, the maximum number of edges in an n-vertex graph without a $2k$-gon is asymptotically $\frac{1}{2}n^{1+1/k}$ as n tends to infinity. This was known almost 40 years ago in the case of quadrilaterals. In this paper, we construct a counterexample to the conjecture in the case of hexagons. For infinitely many n, we prove that

$$\mathrm{ex}(n, C_6) \quad > \quad \frac{3(\sqrt{5}-2)}{(\sqrt{5}-1)^{4/3}}n^{4/3} + O(n) \quad > \quad 0.5338n^{4/3}.$$

We also show that $\mathrm{ex}(n, C_6) \leqslant \lambda n^{4/3} + O(n) < 0.6272n^{4/3}$ if n is sufficiently large, where λ is the real root of $16\lambda^3 - 4\lambda^2 + \lambda - 3 = 0$. This yields the best-known upper bound for the number of edges in a hexagon-free graph. The same methods are applied to find a tight bound for the maximum size of a hexagon-free $2n \times n$ bipartite graph.

MSC: 05C35; 05C38

Keywords: Excluded cycles; Turán numbers; Extremal graph theory

* Corresponding author. Fax: +1 425 936 7329.
E-mail addresses: z-furedi@math.uiuc.edu (Z. Füredi), anaor@microsoft.com (A. Naor), jverstra@math.uwaterloo.ca (J. Verstraëte).
[1] Research supported in part by a Hungarian National Science Foundation Grant OTKA T 032452, and by a National science Foundation Grant DMS 0140692.

0001-8708/$ - see front matter
doi:10.1016/j.aim.2005.04.011

1. Introduction

The forbidden subgraph problem involves the determination of the maximum number of edges that an n-vertex graph may have if it contains no isomorphic copy of a fixed graph H. This number is called the *Turán number* for H, and denoted $\mathrm{ex}(n, H)$. In this paper, we study the Turán problem for the hexagon, that is, the cycle of length six, C_6.

The densest constructions of $2k$-cycle-free graphs for certain small values of k arise from the existence of rank two geometries called *generalized k-gons*, first introduced by Tits [18]. These may be defined as rank two geometries whose bipartite incidence graphs are regular graphs of diameter k and girth $2k$, and are known to exist only when k is three, four or six. This fact is a consequence of a fundamental theorem of Feit and Higman [11]. It is therefore of interest to examine the extremal problem for cycles of length four, six and ten.

In these cases, Lazebnik et al. [15] used the existence of polarities of generalized polygons to construct dense $2k$-cycle-free graphs when $k \in \{2, 3, 5\}$. In particular, for $k = 3$, their construction shows that

$$\mathrm{ex}(n, C_6) \geqslant \tfrac{1}{2}n^{4/3} + O(n)$$

for infinitely many n. Erdős and Simonovits [10] conjectured the asymptotic optimality of these graphs, asking whether $\mathrm{ex}(n, C_{2k})$ is asymptotic to $\frac{1}{2}n^{1+1/k}$ as n tends to infinity. This was known to hold for quadrilaterals almost 40 years ago, as proved by Erdős et al. [9], and independently by Brown [4], but was recently disproved in [15] for cycles of length ten. The only remaining case allowed by the Feit–Higman theorem is that of hexagons. In this paper, we refute the Erdős–Simonovits conjecture for hexagons:

Theorem 1.1. *For infinitely many positive integers n, there exists an n-vertex hexagon-free graph of size at least*

$$\frac{3(\sqrt{5}-2)}{(\sqrt{5}-1)^{4/3}}n^{4/3} + O(n) \;>\; 0.5338n^{4/3}.$$

For all n, $\mathrm{ex}(n, C_6) \leqslant \lambda n^{4/3} + O(n) < 0.6272n^{4/3}$ *if n is sufficiently large, where λ is the real root of* $16\lambda^3 - 4\lambda^2 + \lambda - 3 = 0$.

This theorem gives the best-known upper bound for $\mathrm{ex}(n, C_6)$. The proof of Theorem 1.1 requires a statement about hexagon-free bipartite graphs, which is interesting in its own right (see [6]). Let $\mathrm{ex}(m, n, C_6)$ be the maximum number of edges amongst all $m \times n$ bipartite hexagon-free graphs. We are able to determine $\mathrm{ex}(m, 2m, C_6)$ asymptotically for all m:

Theorem 1.2. *Let m, n be positive integers. Then*

$$\mathrm{ex}(m, n, C_6) < 2^{1/3}(mn)^{2/3} + 16(m + n).$$

Furthermore, if $n = 2m$ then as n tends to infinity,

$$\mathrm{ex}(m, n, C_6) = \begin{cases} 2^{1/3}(mn)^{2/3} + O(n) & \text{for infinitely many } m, \\ 2^{1/3}(mn)^{2/3} + o(n^{4/3}) & \text{for all } m. \end{cases}$$

Another natural question is the maximum number of edges in a quadrilateral-free subgraph of a hexagon-free graph. We answer this question completely, generalizing a theorem of Györi [12] and of Kühn and Osthus [14]:

Theorem 1.3. *Let G be a hexagon-free graph. Then there exists a subgraph of G of girth at least five, containing at least half the edges of G. Furthermore, equality holds if and only if G is a union of edge-disjoint complete graphs of order four or five.*

Throughout the paper, $G = (V, E)$ denotes a graph with vertex set $V = V(G)$ and edge set $E = E(G)$. We write $uv \in E$ instead of $\{u, v\} \in E$. If uv is assigned an orientation, then $u \to v$ means the edge uv is oriented from u to v. A sequence $(x_1, x_2, \ldots, x_{k+1})$ of distinct vertices of G is used to denote a path of length k or, if $x_{k+1} = x_1$, a cycle of length k. The vertices x_1 and x_{k+1} are referred to as *endvertices* of the path. If A is a set of vertices of G, then $G[A] = \{uv \in E(G) : u, v \in A\}$ is the subgraph of G *induced by* A. Finally, $d_G(v)$ and $\Gamma_G(v)$ denote the *degree* and *neighborhood* of a vertex $v \in V$ in the graph G. In general, the subscript G is omitted.

2. Constructions

The aim of this section is to give a construction of hexagon-free graphs achieving the lower bound in Theorem 1.1, and bipartite hexagon-free graphs achieving the lower bound in Theorem 1.2. We start with a known construction of generalized quadrangles, and then modify it.

2.1. Bipartite construction

It is known that there exist $(q+1)$-regular bipartite graphs $H = H_q$, with $q^3 + q^2 + q + 1$ vertices in each part, and girth eight, where q is a prime power. These are the incidence graphs of certain rank two geometries known as *generalized quadrangles*. Let L and R be the parts of H. To construct $m \times 2m$ bipartite graphs achieving the lower bound in Theorem 1.2, we add a new set L^* of $|L|$ vertices to H, and let $\phi : L^* \leftrightarrow L$ be a bijection. Define a bipartite graph $H_q^* = H^*$ with parts R and $L \cup L^*$ as follows:

$$E(H^*) = E(H) \cup \{uv : u \in L^*, v \in R, \phi(u)v \in E(H)\}.$$

In words, we take two edge-disjoint copies of H and identify them on the vertices in R. To see that H^* contains no hexagons, suppose $C = (v_1, v_2', v_3, v_4', v_5, v_6', v_1)$ is a hexagon in H^*. By symmetry, we may assume C contains at least four edges of H. Form the closed walk $\omega = (v_1, v_2, v_3, v_4, v_5, v_6, v_1)$ in H, where $v_i = v_i'$ if $v_i' \in L$ and $v_i = \phi(v_i')$ if $v_i' \in L^*$. Since H contains no cycles of length at most six, ω traverses each edge of a tree $T \subset H$ twice. However, T has at least four edges, since C contains at least four edges of H. This implies ω has length at least eight, which is a contradiction. Therefore H^* contains no hexagons. Also, H^* has $2(q+1)(q^3+q^2+q+1)$ edges and, with $n = 2(q^3+q^2+q+1)$ and $m = q^3+q^2+q+1$, we have

$$|E(H^*)| \; > \; 2^{1/3}(mn)^{2/3} + \tfrac{2}{3}n - O(n^{2/3}).$$

It is known that there is a prime number in the interval $\{n, n+1, \ldots, n+cn^{\theta}\}$ for some $\theta \in [\frac{1}{2}, 1)$. The most recent advance is due to Baker et al. [2], who proved that $\theta = 21/40$ works. Then the infinite sequence of graphs H_q^* shows that for all m and $n = 2m$,

$$\mathrm{ex}(m, n, C_6) > 2^{1/3}(mn)^{2/3} - O(n^{\theta+5/6}).$$

This shows Theorem 1.2 is asymptotically optimal when $n = 2m$.

2.2. Non-bipartite construction

A polarity of a bipartite graph is a bijection between the parts of the bipartite graph which is an involutary automorphism of the graph. The existence of polarities for incidence graphs of certain rank two geometries, known as generalized polygons, was used by Lazebnik, Ustimenko and Woldar to construct dense graphs without cycles of length $2k$ for $k \in \{2, 3, 5\}$. In particular, it is known that the bipartite graphs H_q possess a polarity (see [16]) if and only if $q = 2^{2\alpha+1}$, for some positive integer α. This was used in [15] to construct from H_q a graph G_q with $N = q^3+q^2+q+1$ vertices, $\frac{1}{2}[(q+1)(q^3+q^2+q+1) - q^2 - 1]$ edges, and no cycles of length three, four or six. For our construction, we start with the graph $G_q = G = (V, E)$. Let (A, B) be a partition of V into two sets, A and B, let $G[A]$ be the subgraph of G induced by A, and suppose that each edge in $G[A]$ is given an orientation. Let W be a new set of vertices, disjoint from V, with $|W| = |A|$, let ϕ be a bijection $W \leftrightarrow A$, and let G^* be the graph on $V \cup W$ defined as follows:

$$E(G^*) \; = \; E \; \cup \; \{uv : u\phi(v) \in E, u \in B, v \in W\} \; \cup \; \{uv : u \in A, v \in W, \phi(v) \to u\}.$$

We claim that G^* is hexagon-free. To see this, suppose $C = (v_1, v_2, v_3, v_4, v_5, v_6, v_1)$ is a hexagon in G^*, and form the closed walk

$$\omega = (u_1, u_2, u_3, u_4, u_5, u_6, u_1),$$

where $u_i = v_i$ if $v_i \in V$ and $u_i = \phi(v_i)$ if $v_i \in W$. Note that ω is a closed walk in G. Since G contains no cycles of length three, four or six, ω is a walk on a tree

$T \subset G$. Now C contains at least two vertices in W, otherwise T contains four edges of C and ω has length at least eight. So there are two vertices $u, v \in W$ of C, each incident with two edges of $E(G^*) \backslash E(G)$, since W is an independent set in G^*. The four edges of C incident with u and v correspond to two subpaths of ω of length two whose center vertices are $\phi(u)$ and $\phi(v)$, by construction. These subpaths must share an edge, otherwise T has at least four edges, and ω has length at least eight, which is not possible. Furthermore, the edge they have in common must be $\phi(u)\phi(v) \in G[A]$, by construction. However, this implies both $\phi(u) \to \phi(v)$ and $\phi(v) \to \phi(u)$, which is impossible, since $\phi(u)\phi(v)$ was given only one orientation. This contradiction completes the proof.

We now choose A so as to maximize $|E(G^*)|$. Fix a positive integer $K < N$ and let $A \subset V$ be a subset of size K, chosen uniformly at random among all such subsets. Observe that in expectation, the number of edges incident with A is at least

$$\frac{\binom{N-2}{K-2} + 2\binom{N-2}{K-1}}{\binom{N}{K}} = \frac{K}{N} \cdot \left(2 - \frac{K-1}{N-1}\right) |E|.$$

Therefore we can choose such an $A \subset V$ for which

$$|E(G^*)| \geqslant |E| + \frac{K}{N}\left(2 - \frac{K-1}{N-1}\right)|E|.$$

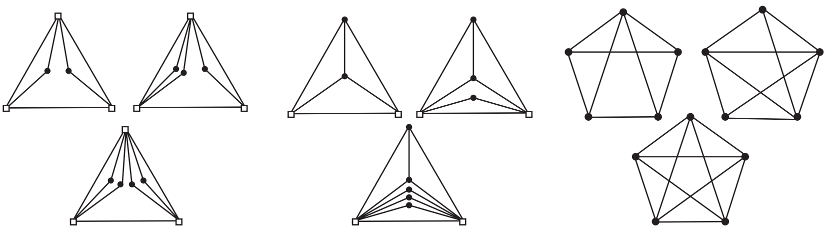
The number of vertices in G^* is $n = N + K$. Choosing $K = \lfloor(\sqrt{5}-2)N\rfloor$ we find

$$|E(G^*)| \geqslant \frac{3(\sqrt{5}-2)}{(\sqrt{5}-1)^{4/3}} n^{4/3} + \frac{2(\sqrt{5}-2)}{(\sqrt{5}-1)} n - O(n^{2/3}).$$

This completes the construction for the lower bound in Theorem 1.1.

3. The structure of hexagon-free graphs

The *quadrilateral relation* on the edge-set of a hexagon-free graph $G = (V, E)$ is the symmetric relation ϕ under which two edges are related if some quadrilateral in G contains both of them. In bounding $\mathrm{ex}(n, C_6)$, we require a description of the components of ϕ. In this section, we give a complete description of these components. The set of edges in each component of ϕ is the edge-set of a subgraph of G. We let $\Phi(G)$ denote the set of all subgraphs of G formed in this way. Let us say that a subgraph H of G is *strongly induced* if every path of length at most four with both endvertices in H is contained in H. A *maximal complete bipartite* subgraph of G is a complete bipartite subgraph of G which contains a cycle and is not contained in any other complete bipartite subgraph of G. Since G is hexagon-free, these consist of a pair of vertices and all their common neighbors in G. We say that a collection of subgraphs



Type (1) Type (2) Type (3)

Fig. 1. Strongly induced subgraphs.

of a graph G *decomposes* G if the subgraphs are pairwise edge-disjoint and every edge of G is in one of the subgraphs. The main theorem we prove is as follows:

Theorem 3.1. *Let G be a hexagon-free graph. Then $\Phi(G)$ decomposes G into single edges, maximal complete bipartite graphs and strongly induced subgraphs of types* (1), (2) *or* (3).

We give a description of the graphs in the illustrations. A subgraph of G of type (1) is a subgraph of G consisting of a complete graph with vertex set $\{a, b, c\}$ (the vertices a, b and c are shown as white squares in the leftmost drawing), such that a and b have at least two common neighbors, a and c have at least two common neighbors, and the set of common neighbors of a and b is disjoint from the set of common neighbors of a and c. A subgraph of type (2) is a subgraph consisting of a complete graph with vertex set $\{a, b, c, d\}$, together with all common neighbors of $\{b, c\}$ (the vertices b and c are shown as white squares in the center Fig. 1). Finally, a subgraph of type (3) is a graph of minimum degree at least three on exactly five vertices.

Proof of Theorem 3.1. One observes that if V is the vertex set of a subgraph of G of type (i), then $G[V]$ is of type (j) for some $j \geqslant i$. Let us prove that $J = G[V]$ is a strongly induced subgraph of G. Let P be a shortest path, with both endpoints in V, such that $P \not\subset J$. Then P has length at least two, since J is an induced subgraph of G. Let k be the length of P. By inspection, any pair of vertices of V is joined by a path of length two, three, and four in J, excepting those pairs of vertices (drawn as white squares in the illustration), all of whose common neighbors are already included in J. If $k \leqslant 4$, then we find a path $P' \subset J$, of length $6-k$, with the same pair of endvertices as P. However, $P \cup P'$ is a hexagon in G. This contradiction shows $k \geqslant 5$, so J is a strongly induced subgraph of G. In particular, if $J \subset H$ for some $H \in \Phi(G)$, then $H = J$: since J is strongly induced, there can be no quadrilateral in G containing an edge in $E(J)$ and an edge in $E(G) \backslash E(J)$.

Now we complete the proof. Suppose $H \in \Phi(G)$ is not a complete bipartite graph (in particular, not a single edge). Then H contains quadrilaterals, Q_1 and Q_2, such that $E(Q_1) \cap E(Q_2) \neq \emptyset$ and $Q_1 \cup Q_2$ is non-bipartite. For if $Q_1 \cup Q_2$ is bipartite, then $Q_1 \Delta Q_2$ is a cycle of length six. Let V consist of the union of $V(Q_1) \cup V(Q_2)$ and

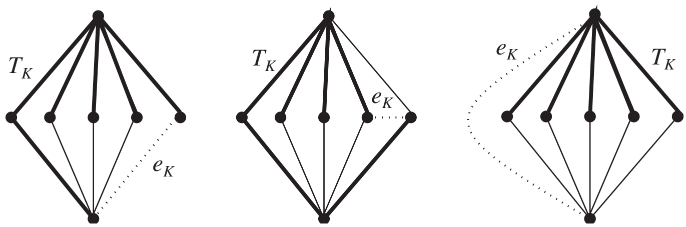


Fig. 2. K-triangles.

all vertices of G with at least two neighbors in $V(Q_1) \cup V(Q_2)$. Since G is hexagon-free, Q_1 and Q_2 share at least three vertices. If $V(Q_1) = V(Q_2)$, then $Q_1 \cup Q_2$ is a complete graph on four vertices. It follows that $G[V]$ contains a subgraph, J, of type (2) or (3). By the first part of the proof, $H = J$. Now suppose Q_1 and Q_2 share three vertices. Then $Q_1 \cup Q_2$ is of type (1), so $G[V]$ contains a subgraph J of type (1), (2) or (3). By the first part of the proof, $H = J$. Therefore H has type (1), (2) or (3). □

3.1. Quadrilateral-free subgraphs

In this section, we prove Theorem 1.3, that every hexagon-free graph contains a subgraph of girth at least five containing at least half its edges.

Proof of Theorem 1.3. By Theorem 3.1, $\Phi(G)$ decomposes G into single edges, maximal complete bipartite graphs and strongly induced subgraphs of types (1)–(3). Let F be the subgraph consisting of all single edges in the decomposition given by $\Phi(G)$. By inspection, each $H \in \Phi(G)$ of types (1)–(3) has a subgraph T_H of girth at least five containing more than half the edges of H, except if H is a complete graph of order four or five in which case $|E(T_H)| = \frac{1}{2}|E(H)|$. If F is non-empty, then we can choose a bipartite subgraph of F of girth at least five with a least $\frac{1}{2}|E(F)|+1$ edges. It remains to deal with those components of $\Phi(G)$ which are complete bipartite graphs.

Let K be such a component. A triangle in G is defined to be a *K-triangle* if it contains an edge of K. We claim that there is an edge $e_K \in E(G)$ joining two vertices of K such that every K-triangle in G contains e_K. One checks the claim using the fact that every quadrilateral in G is either contained in K or edge-disjoint from K, which follows since K is a component of $\Phi(G)$. We now define a tree $T_K \subset K$ so that every triangle containing at least one edge of K also contains an edge of $E(K)\backslash E(T_K)$. Then the graph consisting of the union over K of all trees T_K has girth at least five. The tree T_K is chosen according to the location of the edge e_K, as shown in Fig. 2 (the tree T_K is shown in bold and the edge e_K is dotted). In all cases except one, T_K has $\frac{1}{2}|E(K)| + 1$ edges. The exceptional case is when e_K joins two vertices of degree $\frac{1}{2}|E(K)|$ in K (the figure on the right in the illustration above).

Let J be the union of all the subgraphs T_H and T_K, together with F. Then J contains no quadrilateral (each is contained in an element of $\Phi(G)$) and no triangle, since a triangle not contained in a type (1), (2) or (3) subgraph and not contained in F consists of at least one edge of a maximal complete bipartite subgraph of G. Finally, J has size at least

$$\sum_{H\in\Phi(G)} \tfrac{1}{2}\,|E(H)| \;=\; \tfrac{1}{2}|E(G)|.$$

Equality holds only if every element of Φ is a complete graph of order four or five (which are subgraphs of types (2) and (3), respectively) or a maximal complete bipartite graph K with two vertices of degree $\frac{1}{2}|E(K)|$ joined in G, and $E(F)=\emptyset$. Now suppose every subgraph of G of girth at least five has size at most $\frac{1}{2}|E(G)|$. We claim that all components of $\Phi(G)$ are complete graphs of order four or five. Suppose this is not true. Then there is a component of $\Phi(G)$ which is a maximal complete bipartite graph with two vertices of degree $\frac{1}{2}|E(K)|$, joined by the edge e_K. Note that there is no triangle in G containing e_K with is not already contained in $E(K)\cup\{e_K\}$. Since $E(F)=\emptyset$ and $|E(J)\cup\{e_K\}| > \frac{1}{2}|E(G)|$, e_K is contained in a quadrilateral in $E(J)\cup\{e_K\}$, and this quadrilateral is contained in a complete bipartite component L of $\Phi(G)$. It follows that $e_L=e_K$ and so the tree $T_L\subset L$ has size $\frac{1}{2}|E(L)|+1$. Therefore $|E(J)| > \frac{1}{2}|E(G)|$, which is a contradiction. So all components of $\Phi(G)$ are complete graphs of order four or five. □

In fact, for bipartite graphs, the following stronger statement is true.

Theorem 3.2. *Let $G=(V,E)$ be a bipartite hexagon-free graph. Then G contains a sub-graph F, of girth at least eight, such that $d_F(v)\geqslant\frac{1}{2}d_G(v)$ for every vertex $v\in V$.*

Proof. By Theorem 3.1, each $H\in\Phi(G)$ is either a complete bipartite graph containing a quadrilateral, or a single edge. For each $H\in\Phi(G)$ containing a quadrilateral, we choose a spanning tree T_H in H containing at least half the edges on each vertex of H. The union of all these trees together with all single edges in $\Phi(G)$ is the required subgraph of G. □

Corollary 3.1. *Let G be a hexagon-free bipartite graph with parts A and B. Suppose every vertex of A has degree at least $\delta(A)$ and every vertex of B has degree at least $\delta(B)$, and G has maximum degree Δ. Then*

$$\Delta(\delta(A)-2)(\delta(B)-2)\leqslant 8\max\{|A|,|B|\}.$$

For any n-vertex hexagon-free graph of minimum degree δ and maximum degree Δ, the inequality $\Delta(\delta-4)^2\leqslant 64n$ holds.

Proof. Let F be a subgraph of G as in the last theorem. Then every vertex of A has degree at least $\delta(A)/2$ in F and every vertex of B has degree at least $\delta(B)/2$ in F. Let v be a vertex of maximum degree in F. Then $d_F(v)\geqslant\Delta/2$. We assume $v\in A$, as similar arguments are applied in the case $v \in B$. Since F has girth at least eight, the number of vertices of F at distance three from v is at least

$$\frac{\Delta}{2} \cdot \left(\frac{\delta(A)}{2} - 1\right)\left(\frac{\delta(B)}{2} - 1\right).$$

As all these vertices are distinct, this expression is at most $|B|$. For the second statement of the corollary, we use an observation of Erdős [7]: any graph G contains a bipartite subgraph F such that $d_F(v) \geqslant \frac{1}{2} d_G(v)$ for all vertices v of G. By the first part of the corollary,

$$\frac{\Delta}{2}\left(\frac{\delta}{2} - 2\right)^2 \leqslant 8n.$$

This completes the proof. □

4. Matrix inequalities

A very general Hölder-type inequality for matrices was proved by Blakley and Roy [3]: they showed that if A is an $n \times n$ symmetric pointwise non-negative matrix and v is a non-negative unit vector, then

$$\left\langle A^k v, v \right\rangle \geqslant \langle Av, v\rangle^k .$$

In the current context, we note that the Blakley–Roy inequality implies that $\|A^k\|_1 \geqslant (\|A\|_1)^k$, where $\|A\|_1$ is the sum of the entries of the matrix A divided by n. If A is the adjacency matrix of an n-vertex graph, then $\|A^k\|_1$ is precisely the (normalized) number of walks of length k in the graph, where the walks $(\omega_0, \omega_1, \omega_2, \dots, \omega_k)$ and $(\omega_k, \omega_{k-1}, \dots, \omega_1, \omega_0)$ are considered distinct. It is not hard to deduce that the number of paths of length three in a graph $G = (V, E)$ with n vertices is at least

$$\frac{1}{2}(W_3 - 6T - 4P_2 - 2P_1) \geqslant \frac{4|E|^3}{n^2} - 3\Delta|E|, \tag{1}$$

where Δ is the maximum degree of G, P_i is the number of paths of length i in G, W_3 is the number of walks of length three in G, and T is the number of triangles in G. The above inequality follows from the fact that G contains at most $(\Delta - 1)|E|/3$ triangles, at most $(\Delta - 1)|E|$ paths of length two, and $|E|$ paths of length one. Inequality (1) may also be deduced from a Moore-type bound, established recently by Alon et al. [1].

In the case of bipartite graphs, Sidorenko [17] proved a similar statement to (1) for matrices which are not necessarily symmetric. Extending the results in [1], Hoory [13] gave a tight lower bound on the number of paths of length k in an $m \times n$ bipartite graph with prescribed average degree. A consequence of these inequalities is that the number of paths of length three in an $m \times n$ bipartite graph $G = (V, E)$ with maximum degree Δ is at least

$$\frac{1}{2}(W_3 - 4P_2 - P_1) \geqslant \frac{|E|^3}{mn} - 2\Delta|E|. \qquad (2)$$

These inequalities will be used in the proofs of Theorems 1.1 and 1.2.

5. Degenerate and non-degenerate pairs

Throughout this section, $G = (V, E)$ is a fixed hexagon-free graph and π is a pair of vertices of G. We denote by $|\pi|$ the number of paths of length three in G with both endvertices in π. Suppose $P_1, P_2, \ldots, P_k$ are the paths of length three joining the two vertices of π. We say that π is *degenerate* if the distance between the vertices of π in $P_1 \cup P_2 \cup \ldots \cup P_k$ is two, and π is *non-degenerate* otherwise. We emphasize that the distance in the definition of degeneracy is taken in $P_1 \cup P_2 \cup \ldots \cup P_k$, and not in G.

Lemma 5.1. *Let $\pi \in \binom{V}{2}$ with $|\pi| \geqslant 2$, and let $P_1, P_2, \ldots, P_k$ be the paths of length three with both endvertices in π. Then π is non-degenerate if and only if*

$$K_\pi = \bigcup_{i,j=1}^{k} E(P_i) \,\Delta\, E(P_j)$$

is a complete bipartite graph, and $E(P_1) \cap E(P_2) \cap \cdots \cap E(P_k)$ consists of a single edge, e_π, incident with a vertex of maximum degree in K_π.

Proof. It is clear that if K_π is a complete bipartite subgraph of G, then the distance between the vertices of π in $P_1 \cup P_2 \cup \ldots \cup P_k$ is three, so π is non-degenerate. Now suppose π is non-degenerate. Let E_π be the set of edges of $P_1 \cup P_2 \cup \ldots \cup P_k$ which are not incident with π. Since G is hexagon-free, E_π is an intersecting family, so E_π is a star or a triangle. In the latter case it is not hard to see that π is degenerate. Therefore E_π is a star. Let w be the center of E_π. Then exactly one vertex v of π is adjacent to w. If the other vertex u of π is not adjacent to some endvertex x of E_π, then the edge wx is not in any of the paths P_i, which is a contradiction. Therefore u is adjacent to all endvertices of E_π, which means that the union of all $E(P_i) \,\Delta\, E(P_j)$ is a complete bipartite subgraph of G. Furthermore,

$$E(P_1) \cap E(P_2) \cap \ldots \cap E(P_k) = \{vw\},$$

so we take the edge e_π to be vw (Fig. 3). $\square$

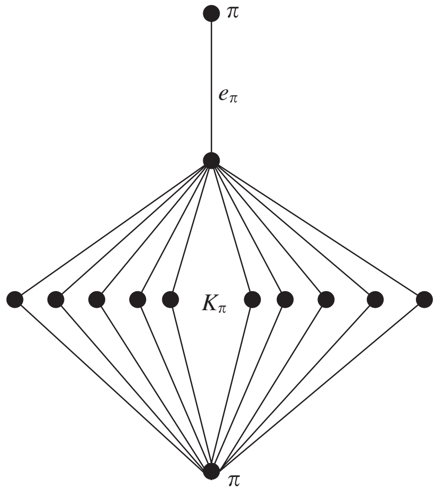


Fig. 3. A non-degenerate pair.

We define a graph H to be *central* if some vertex of H is adjacent to all other vertices in H. In other words, H has a spanning star.

Lemma 5.2. *Let π be a pair of vertices of G with $|\pi| \geqslant 2$, and suppose $P_1, P_2, \ldots, P_k$ are the paths of length three in G joining the vertices in π. Then π is degenerate if and only if the graph $P_1 \cup P_2 \cup \cdots \cup P_k$ is central.*

Proof. Suppose $J = P_1 \cup P_2 \cup \ldots \cup P_k$ is central. Since π is not an edge of any P_i, π is certainly degenerate. Now suppose π is degenerate, and let E_π be the set of edges of J which are not incident with a vertex of π. Then E_π is an intersecting family, so E_π is a star or a triangle. If E_π is a triangle, then J has five vertices, and it is straightforward to verify that J is central. Suppose E_π is a star with center w. Since π is degenerate, there is a path (u, x, v) in J with both endvertices in $\pi = \{u, v\}$. If $x = w$, then u and v are adjacent to w, so w is central. If $x \neq w$, then since E_π is a star with center w, (u, x, w, v) and (v, x, w, u) are paths of length three in J, so u and v are adjacent to w. It follows that J is central. □

6. Counting paths: bipartite graphs

We now prove that very few pairs of vertices in a bipartite hexagon-free graph are joined by at least three paths of length three. More precisely:

Lemma 6.1. *Let $G = (V, E)$ be a hexagon-free bipartite graph of maximum degree Δ, and let Π be the set of pairs π for which $|\pi| \geqslant 2$ and $\pi \in E$, or $|\pi| \geqslant 3$. Then*

$$\sum_{\pi \in \Pi} |\pi| \leqslant \Delta |E|.$$

Proof. A complete bipartite subgraph of G is called *maximal* if it is not properly contained in any other complete bipartite subgraph, and contains a cycle. By Theorem 3.1, $\Phi(G)$ is a decomposition of G into maximal complete bipartite graphs and single edges. Moreover, every maximal complete bipartite graph in G is a component of $\Phi(G)$, since every two edges of this subgraph are related in ϕ. Now let $\pi \in \Pi$. By Lemma 5.2, the union of all paths of length three with endvertices in a degenerate pair is a central graph, and such a graph cannot be bipartite. So every pair $\pi \in \Pi$ is non-degenerate. Now we use the notation K_π and e_π of Lemma 5.1. By this lemma, for each $\pi \in \Pi$, K_π is a complete bipartite graph, and all paths of length three with both endvertices in π comprise a subpath of K_π of length two, together with a fixed edge e_π, incident with a vertex of maximum degree in K_π. For $\pi \in \Pi$, define

$$h(\pi) = \begin{cases} K_\pi \cup \{\pi, e_\pi\} & \text{if } \pi \in E, \\ K_\pi & \text{if } \pi \notin E. \end{cases}$$

Note that $h(\pi)$ is a complete bipartite subgraph of G, with exactly two vertices of degree at least three. We claim that $h(\pi)$ is a maximal complete bipartite subgraph of G. Suppose, for a contradiction, that this is false for some $\pi \in \Pi$. Then there is a path $P \subset G$, of length two, such that $P \not\subset h(\pi)$, and P joins the two vertices of $h(\pi)$ of degree at least three. However, $P \cup \{e_\pi\}$ is a path of length three with both endvertices in π, which contradicts Lemma 5.1. This proves the claim, and we conclude that $h(\pi) \in \Phi(G)$. Let $\Phi^\circ(G)$ consist of all elements of $\Phi(G)$ which are not single edges. Now define

$$h^{\dashv} : \Phi^\circ(G) \to \Pi \quad \text{where} \quad h^{\dashv}(H) = \{\pi \in \Pi : h(\pi) = H\}.$$

Thus $h^{\dashv}$ is the inverse image of h. Our next claim is $|h^{\dashv}(H)| \leqslant 2\Delta$ for all $H \in \Phi^\circ(G)$. For a fixed $H \in \Phi^\circ(G)$, if $h(\pi) = H$ and $\pi \in E$, then $\pi \in E(H)$, by definition of $h(\pi)$. So the equation $h(\pi) = H$ has $|E(H)|$ solutions $\pi \in E$. The equation $h(\pi) = H$ has at most $2\Delta - |E(H)|$ solutions $\pi \notin E$, since $K_\pi = H$ in this case, and there are at most $2\Delta - |E(H)|$ choices of the edge $e_\pi \in E(G)\setminus E(H)$. We conclude that $|h^{\dashv}(H)| \leqslant 2\Delta$ for all $H \in \Phi^\circ(G)$, proving the claim.

Finally, for all $\pi \in \Pi$, we have $2|\pi| = |K_\pi|$, so $|\pi| \leqslant \frac{1}{2}|E(H)|$ for all $\pi \in h^{\dashv}(H)$. Therefore

$$\begin{aligned} \sum_{\pi \in \Pi} |\pi| &= \sum_{H \in \Phi^\circ(G)} \sum_{\pi \in h^{\dashv}(H)} |\pi| \\ &\leqslant \sum_{H \in \Phi^\circ(G)} \sum_{\pi \in h^{\dashv}(H)} \tfrac{1}{2}\,|E(H)| \end{aligned}$$

$$= \sum_{H \in \Phi^\circ(G)} |h^{-1}(H)| \cdot \tfrac{1}{2}\,|E(H)|$$
$$\leqslant \sum_{H \in \Phi^\circ(G)} \Delta|E(H)|.$$

This is at most $\Delta|E|$, since the subgraphs in $\Phi^\circ(G)$ are pairwise edge-disjoint. □

Remark. The condition $\pi \in \Pi$ in this lemma is necessary. For example, the bipartite graph H^*, defined in Section 2, has maximum degree $2(q+1)$ and $2(q+1)(q^3+q^2+q+1)$ edges, and

$$\sum_{\substack{|\pi|=2 \\ \pi \notin \Pi}} |\pi| \;=\; \sum_{\substack{|\pi|=2 \\ \pi \notin E}} |\pi| \;\sim\; 4q^6 \;\sim\; q\Delta|E|.$$

The reason the lemma does not apply to pairs $\pi \notin \Pi$ with $|\pi| = 2$ is that if two non-adjacent vertices π with $|\pi| = 2$ are joined by paths P and Q of length three, then $P \cup Q$ contains a quadrilateral by Lemma 5.1, but this quadrilateral (the subgraph $h(\pi)$ in the proof above) might not be a maximal complete bipartite subgraph of G.

7. Proof of Theorem 1.2

We prove, by induction on $m+n$, that if $G = (V, E)$ is a bipartite graph with parts A and B of sizes $m \geqslant 3$ and $n \geqslant 3$ respectively, and $|E| \geqslant 2^{1/3}(mn)^{2/3} + 16(m+n)$, then G contains a hexagon. Let $f(m,n) = 2^{1/3}(mn)^{2/3} + 16(m+n)$. First note that the statement of the induction hypothesis is vacuously true if $m+n \leqslant 16$, since $f(m,n) > mn$ in this case. Now let $m+n > 16$, and suppose, for a contradiction, that some m by n bipartite graph $G = (V, E)$, with $|E| \geqslant f(m,n)$, contains no hexagon. Let

$$\delta(A) = \tfrac{2}{3} \cdot 2^{1/3} m^{-1/3} n^{2/3} + 16,$$
$$\delta(B) = \tfrac{2}{3} \cdot 2^{1/3} n^{-1/3} m^{2/3} + 16.$$

Note that $\delta(A) \leqslant f(m,n) - f(m-1,n)$ and $\delta(B) \leqslant f(m,n) - f(m,n-1)$. It follows that every vertex of A has degree at least $\delta(A)$, otherwise we remove a vertex of degree less than $\delta(A)$ to obtain an $(m-1)$ by n bipartite graph with at least $f(m-1,n)$ edges, and this graph contains a hexagon, by induction. Similarly, every vertex of B has degree at least $\delta(B)$. By Corollary 3.1,

$$\Delta \cdot \tfrac{4}{9}(m^{1/3} n^{1/3}) \;<\; \Delta(\delta(A) - 2)(\delta(B) - 2) \;\leqslant\; 8n.$$

It follows that $\Delta < 18 \cdot 2^{1/3} n^{2/3} m^{-1/3}$. Now the number of paths of length three in G is $\sum |\pi|$, where the sum is over all pairs $\pi \in \binom{V}{2}$. By Lemma 6.1 and inequality (2),

$$\frac{|E|^3}{mn} - 2\Delta|E| \leqslant \sum |\pi| \leqslant 2mn + 2\Delta|E|.$$

Let $g(x) = x^3/mn - 4\Delta x - mn$ and $z = 27(mn)^2 + 3\sqrt{(3mm)^4 - 192(\Delta mn)^3}$. Then $g(x)$ has a unique real root x_0 given by

$$\begin{aligned} x_0 &= \tfrac{1}{3} z^{\frac{1}{3}} + 4\Delta mn z^{-\frac{1}{3}} \\ &< 2^{\frac{1}{3}} (mn)^{\frac{2}{3}} + 4\Delta mn (27mn)^{-\frac{2}{3}} \\ &< 2^{\frac{1}{3}} (mn)^{\frac{2}{3}} + \frac{4 \cdot 18 \cdot 2^{\frac{1}{3}}}{27^{\frac{2}{3}}} n \\ &< 2^{\frac{1}{3}} (mn)^{\frac{2}{3}} + 16n \end{aligned}$$

In this calculation we used the upper bound $\Delta < 18 \cdot 2^{1/3} n^{2/3} m^{-1/3}$. The above expression is an upper bound for $|E|$, which contradicts the lower bound $|E| \geqslant f(m, n)$. $\square$

8. Counting paths: non-bipartite graphs

In the case of non-bipartite hexagon-free graphs, we prove that $\sum |\pi| \leqslant 35\Delta|E|$ where the sum is over all pairs π of vertices which have $|\pi| \geqslant 3$ or $|\pi| \geqslant 2$ and $\pi \in E$ or $|\pi| \geqslant 2$ and π is degenerate:

Lemma 8.1. *Let $G = (V, E)$ be a hexagon-free graph of maximum degree Δ, and let Π^* be the set of degenerate pairs $\pi \in \binom{V}{2}$ with $|\pi| \geqslant 2$. Then*

$$\sum_{\pi \in \Pi \cup \Pi^*} |\pi| \leqslant 35\Delta|E|.$$

Proof. We first consider the contribution of pairs $\pi \in \Pi^*$ to the sum on the left. Let $p[v]$ denote the number of paths of length three in G consisting only of vertices in $A = \Gamma(v) \cup \{v\}$. By Lemma 5.2, for $\pi \in \Pi^*$, the union of all the paths of length three with endvertices in π is a central subgraph of G. Therefore

$$\sum_{\pi \in \Pi^*} |\pi| \leqslant \sum_{v \in V} p[v].$$

Let $e[v]$ be the number of edges of $G[A]$. If we choose a pair of disjoint edges in $G[A]$, then there are at most four choices for a path of length three containing these edges. Therefore $p[v] < 2e[v]^2$. Now since G is hexagon free, no path of length four consists entirely of edges in $\Gamma(v)$. By the Erdős–Gallai theorem [8], there are then at most $3d(v)/2$ edges in $\Gamma(v)$, so $e[v] \leqslant 5d(v)/2$. It follows that

$$\sum_{\pi \in \Pi^*} |\pi| \ < \ \sum_{v \in V} 2e[v]^2 \ \leqslant \ \sum_{v \in V} \tfrac{25\Delta}{2} d(v) \ \leqslant \ 25\Delta|E|.$$

This takes care of the sum over Π^*.

Now let $\pi \in \Pi$ and define $h(\pi) = K_\pi$ if $\pi \notin E$ and $h(\pi) = K_\pi \cup \{e_\pi, \pi\}$ if $\pi \in E$, as in the proof of Lemma 6.1. Recall that a maximal complete bipartite subgraph of G is a complete bipartite subgraph of G which contains a cycle and it not properly contained in any other complete bipartite subgraph of G. Following the proof of Lemma 6.1, $h(\pi)$ is a maximal complete bipartite subgraph of G. By Theorem 3.1, $h(\pi)$ is contained in a unique subgraph $H_\pi \in \Phi^\circ(G)$, where $\Phi^\circ(G)$ is the set of subgraphs of G in $\Phi(G)$ which contain a cycle. By inspection, using Theorem 3.1, each $H \in \Phi^\circ(G)$ contains at most ten maximal complete bipartite subgraphs of G. Define

$$g : \Phi^\circ(G) \to \Pi \text{ by } g(H) = \{\pi : H_\pi = H\}.$$

Fix $H \in \Phi^\circ(G)$. Then the equation $g(\pi) = H$ with $\pi \in E$ has at most $|E(H)|$ solutions, since $\pi \in E$ implies $\pi \in E(H_\pi)$. The equation $g(\pi) = H$ with $\pi \notin E$ has at most 20Δ solutions, since there are ten choices for $h(\pi)$, which is a maximal complete bipartite subgraph of H, and then at most 2Δ choices for e_π, which is incident with one of the two vertices of $h(\pi)$ of maximum degree. Finally, $|\pi| \leqslant \frac{1}{2}|E(H)|$ for all $\pi \in g(H)$. Therefore

$$\begin{aligned}\sum_{\pi \in \Pi} |\pi| &= \sum_{H \in \Phi^\circ(G)} \sum_{\pi \in g(H)} |\pi| \\ &\leqslant \sum_{H \in \Phi^\circ(G)} |g(H)| \cdot \tfrac{1}{2}|E(H)| \ \leqslant \ \sum_{H \in \Phi^\circ(G)} 10\Delta|E(H)|.\end{aligned}$$

This is at most $10\Delta|E|$, by Theorem 3.1. □

8.1. Maximum directed cuts

A *maximum directed cut* in an oriented graph G is a partition of the vertex set into two sets X and Y such that the number of directed edges from X to Y is as large as possible. This maximum is denoted $\mathrm{mdc}(G)$.

Lemma 8.2. *Let $G = (V, E)$ be an oriented graph on n vertices, and suppose that $\mathrm{mdc}(G) \leqslant \gamma|E|$. Then G contains at least $(1-\gamma)^2|E|^2/n$ directed paths of length two.*

The proof of Lemma 8.2 is based on the following numerical inequality:

Lemma 8.3. *For $n \geqslant 1$ let V be an n-point set and $f, g : V \to [0, \infty)$ be two functions such that $\sum f(v) = \sum g(v)$. Then*

$$\sum_{v \in V} f(v)g(v) \geqslant \frac{1}{n} \left[\sum_{v \in V} \min\{2f(v) - g(v), f(v)\} \right]^2.$$

Proof. If $f = g$, then the inequality becomes the Cauchy–Schwartz inequality

$$\sum_{v \in V} f(v)^2 \geqslant \frac{1}{n} \left(\sum_{v \in V} f(v) \right)^2.$$

Suppose, therefore, that there exist $u, v \in V$ such that $f(u) > g(u)$ and $f(v) < g(v)$, and let $\varepsilon = \min\{|f(u) - g(u)|, |f(v) - g(v)|\}$. Then define functions f' and g' which are identical to f and g on $V \setminus \{u, v\}$ and such that $f'(u) = f(u) - \varepsilon$ and $g'(v) = g(v) - \varepsilon$, and $f'(v) = f(v)$ and $g'(u) = g(u)$. Then f' and g' are non-negative functions with the same sum. Furthermore, the left-hand side of the inequality decreases, whereas the right-hand side stays the same. Repeating the transformation $f \to f'$ and $g \to g'$ finitely many times, we arrive at the Cauchy–Schwartz inequality. Since the inner product $\sum f(v)g(v)$ only decreased under the transformation, this completes the proof. $\square$

Proof of Lemma 8.2. We assume the vertex set of G is V. For a vertex $v \in V$, let $f(v)$ and $g(v)$ be the number of edges incident with v which are oriented into v, and the number of edges oriented out of v, respectively. The number of directed paths of length two is exactly $\sum f(v)g(v)$. The restriction on the size of the maximum directed cut implies that for some $a \leqslant \gamma |E|$,

$$\sum_{v \in I} [g(v) - f(v)] = a \quad \text{where} \quad I = \{v : g(v) > f(v)\}.$$

Also $\sum f(v) = |E| = \sum g(v)$. By Lemma 8.3,

$$\sum_{v \in V} f(v)g(v) \geqslant \frac{1}{n}(|E| - a)^2 \geqslant \frac{1}{n}(1 - \gamma)^2 |E|^2.$$

Therefore the number of directed paths of length two in G is at least $(1 - \gamma)^2 |E|^2 / n$. $\square$

8.2. Pairs joined by a unique path

Lemma 8.1 shows that less than $18\Delta|E|$ pairs of vertices of a hexagon-free graph G are joined by at least three paths of length three. The question remains: does there exist an extremal n-vertex hexagon-free graph for which almost all pairs are joined by two paths of length three? In this section, we prove a lemma which shows that such graphs are far from extremal. This should be compared with the bipartite construction in Section 2.

An oriented path of length three is *a forward path* if it contains a directed path of length two *from* one of its endvertices. For example, $u \to v \to w \leftarrow x$ and $u \to v \leftarrow w \leftarrow x$ are forward paths, whereas $u \to v \leftarrow w \to x$ is not. For a given orientation of a graph G, let Π^{**} denote the set of pairs of vertices joined by at least one forward path and at least two paths of length three.

Lemma 8.4. *Let $G = (V, E)$ be a hexagon-free graph of maximum degree Δ. Then there is an orientation of G such that*

$$\sum_{\pi \in \Pi^{**}} |\pi| < 43\Delta|E|.$$

Proof. The orientation is described as follows. For subgraph $H \in \Phi(G)$ which is not a quadrilateral and not a single edge, we orient the edges of $E(H)$ so that each vertex of degree two in H has indegree two. The remaining edges of G are oriented arbitrarily. By Theorem 3.1, since $\Phi(G)$ decomposes G, this orientation is consistent. Then, by Lemma 8.1,

$$\sum_{\pi \in \Pi^* \cup \Pi} |\pi| \leqslant 35\Delta|E|.$$

For the rest of the proof, $\pi \in \Pi^{**} \backslash (\Pi^* \cup \Pi)$, so that $|\pi| = 2$ and π is non-degenerate. If $P = P_\pi$ and $Q = Q_\pi$ are the paths of length three with endvertices in π, then $E(P)\Delta E(Q)$ is a quadrilateral, K_π, by Lemma 5.1. By Theorem 3.1, K_π is contained in a unique subgraph $H_\pi \in \Phi(G)$. It is not hard to see, via Theorem 3.1, that there are at most $4\Delta|E|$ choices of $\pi \in \binom{V}{2}$ such that at least one of the two vertices incident in K_π with a vertex of π has degree at least three in H_π, or such that $K_\pi = H_\pi$. Now suppose π is not such a pair and $|\pi| = 2$. Then we can assume $\pi = \{u, v\}$,

$$K_\pi = (u, x, w, y) \quad P_\pi = (u, x, w, v) \quad Q_\pi = (u, y, w, v)$$

and $K_\pi \neq H = H_\pi$, and $d_H(x) = d_H(y) = 2$. Since H is not a quadrilateral, K_π has the orientation $u \leftarrow x \to w \leftarrow y \to u$. If P_π is a forward path, then $u \to x \to w$ or $v \to w \to x$, however, these orientations both conflict with the orientation of K_π. We conclude that neither P_π nor Q_π is a forward path. So there are less than $4\Delta E|$ pairs $\pi \in \Pi^{**} \backslash (\Pi^* \cup \Pi)$. It follows that

$$\sum_{\pi \in \Pi^{**}} |\pi| \leqslant \sum_{\pi \in \Pi^{**} \backslash (\Pi^* \cup \Pi)} |\pi| + \sum_{\pi \in \Pi^* \cup \Pi} |\pi| < 35\Delta|E| + 8\Delta|E| < 43\Delta|E|.$$

This completes the proof. $\square$

Corollary 8.1. *Let $G = (V, E)$ be a hexagon-free graph of minimum degree δ and maximum degree Δ, where $\delta \geqslant 2$, and suppose G has no cut of size more than $\gamma|E|$. Then the number of pairs of vertices of G joined by exactly one path of length three is greater than*

$$(1-\gamma)^2 \frac{(\delta-2)|E|^2}{n} - 43\Delta|E|.$$

Proof. By Lemma 8.2, the number of directed paths of length two in any orientation of the edges of G is at least $(1-\gamma)^2|E|^2/n$. In particular, the number of forward paths in any orientation of G is at least $(1-\gamma)^2(\delta-2)|E|^2/n$, since there are at least $\delta - 2$ ways of extending a given directed path of length two to a forward path. By Lemma 8.4, there is an orientation of G such that fewer than $43\Delta|E|$ forward paths are the unique path of length three joining their endvertices. This completes the proof. □

9. Proof of Theorem 1.1

Let $G = (V, E)$ be an n-vertex graph with at least $\lambda n^{4/3} + cn$ edges and no hexagon, such that n is as small as possible. Let $\delta = \lfloor \frac{4}{3}\lambda n^{1/3} \rfloor + c$, and suppose G has maximum degree Δ. We may assume $d_G(v) > \delta - 1$ for all $v \in V$, otherwise we remove a vertex of smallest degree to obtain a graph with at least $\lambda(n-1)^{4/3} + c(n-1)$ edges, and this graph contains a hexagon, by the minimality of G. Consider a maximum cut H with parts of sizes m and $n-m$, where $m = \lfloor n/2 \rfloor$. By Theorem 1.2, if $c > 32\lambda$, then

$$|E(H)| \leqslant 2^{1/3}(m(n-m))^{2/3} + 16n \leqslant \frac{1}{2}n^{4/3} + 16n < \frac{1}{2\lambda}|E|.$$

Let Π_i be the set of pairs of vertices of G joined by exactly i paths of length three. Using the lower bound for $|\Pi_1|$ given by Corollary 8.1, with $\gamma = 1/2\lambda$,

$$|\Pi_1| + 2|\Pi_2| \;<\; n^2 - \left(1 - \frac{1}{2\lambda}\right)^2 \frac{(\delta-2)|E|^2}{n} + 43\Delta|E|.$$

By inequality (1), G has at least $\frac{4|E|^3}{n^2} - 3\Delta|E|$ paths of length three. So, by Lemma 8.1,

$$\begin{aligned}
\frac{4|E|^3}{n^2} - 3\Delta|E| \;&<\; \sum |\pi| \;<\; |\Pi_1| + 2|\Pi_2| + 35\Delta|E| \\
&< n^2 - \left(1 - \frac{1}{2\lambda}\right)^2 \frac{(\delta-2)|E|^2}{n} + 78\Delta|E| \\
&< n^2 - \left(1 - \frac{1}{2\lambda}\right)^2 \frac{4|E|^3}{3n^2} + \frac{(c+9)|E|^2}{3n} + 78\Delta|E|.
\end{aligned}$$

By Corollary 3.1, Δ is within a constant factor of the average degree of G:

$$\Delta \leqslant (6/\lambda)^2 n^{1/3}.$$

Using this inequality, we may rearrange the preceding inequality by replacing $|E|$ by $\lambda n^{4/3} + cn$, multiplying by n^2, and moving all terms to the left-hand side. A small calculation shows that after this rearrangement, the coefficient of n^4 is $\frac{1}{3}(16\lambda^3 - 4\lambda^2 + \lambda - 3)$. By definition of λ, this is zero. The rest of the inequality above becomes:

$$\left(c + 3\lambda^2 + \frac{49}{3}\lambda^2 c - \frac{2916}{\lambda} - 4\lambda c\right) n^{\frac{11}{3}} + \left(\frac{c^2}{\lambda} + \frac{50}{3}\lambda c^2 + 6\lambda c - \frac{2916c}{\lambda^2} - 4c^2\right) n^{\frac{10}{3}}$$
$$+ \left(\frac{17c^3}{3} - \frac{4c^3}{3\lambda} + \frac{c^3}{3\lambda^2} + 3c^2\right) n^3 \;<\; 0.$$

The left-hand side is positive if c is large enough, which is a contradiction. Therefore if c is large enough constant, then every n-vertex graph with at least $\lambda n^{4/3} + cn$ edges contains a hexagon. This proves Theorem 1.1. $\square$

10. Concluding remarks

- Using Corollary 8.1, one can deduce the following statement:

  if $o(n^2)$ pairs of vertices in an n-vertex hexagon-free graph G are joined by exactly one path of length three, then G has a bipartite subgraph with $|E(G)| - o(n^{4/3})$ edges.

  This is proved as follows. First we observe that the statement is trivial if $|E(G)| = o(n^{4/3})$, so we assume $|E(G)| \geqslant \varepsilon n^{4/3}$ where $\varepsilon > 0$ is an absolute constant. As in the proof of Theorem 1.1, one proves that the minimum degree of G is at least $c|E|/n$ and the maximum degree of G is at most $d|E|/n$ where $0 < c < d$ are absolute constants (for example see the upper bound on Δ given in the proof of Theorem 1.1). If the maximum cut in G has size $\gamma|E(G)|$ then Corollary 8.1 shows that at least

  $$(1-\gamma)^2 c\varepsilon^3 n^2 - O(d\varepsilon^2 n^{5/3})$$

  pairs of vertices of G are joined by exactly one path of length three. On the other hand, the number of such pairs of vertices of G is $o(n^2)$, by assumption. Therefore $\gamma = 1 - o(1)$, and G has a maximum cut of size $|E(G)| - o(n^{4/3})$.
- The non-bipartite construction in Section 2 above gives a large family of hexagon-free graphs with roughly the same number of edges. Indeed, the base graph G is known to be pseudorandom (see [5]), and it is known that in such graphs, the number

of edges between disjoint sets A and B of vertices and the number of edges induced by A are roughly the same as in a random graph with the same density. Therefore we may choose *any* subset A of size $K = \lfloor(\sqrt{5}-2)N\rfloor$ and *any* orientation of the edges within A, and the construction will have about the same density. On the other hand, the distribution of the degrees of the new vertices in W depends explicitly on the orientation of A, so many of the graphs obtained are non-isomorphic.

- There is not much evidence to suggest whether

$$\lim_{n\to\infty} \mathrm{ex}(n, C_6)/n^{4/3}$$

exists. If the limit does exist, then determining its value is an interesting problem. We are only able to show that the value is in the interval $(0.5338, 0.6272)$. We believe that its value should be closer to 0.5338 than 0.6272, but do not have a conjecture as to the exact value, even under the assumption that our graphs are regular. For these reasons, we have not attempted to optimize the constants in the terms of lower order in all of our bounds.

- Another interesting question, which might shed light on the asymptotic behavior of $\mathrm{ex}(n, C_6)$, is to determine whether or not there are infinitely many n such that

$$\mathrm{ex}(n, n, C_6) \geqslant cn^{4/3} - o(n^{4/3})$$

for some constant $c > 1$. One way of proving this would be to find infinitely many $n/2$ by n bipartite graphs of girth at least eight, and with at least $\frac{1}{2}cn^{4/3} - o(n^{4/3})$ edges; then the bipartite construction in Section 2 can be applied to give an $n \times n$ bipartite graph with $cn^{4/3} - o(n^{4/3})$ edges. Note that any bipartite graph of girth at least eight and with parts of sizes $n/2$ and n has at most $2^{-2/3}n^{4/3} + O(n)$ edges, by Theorem 1.2.

References

[1] N. Alon, S. Hoory, N. Linial, The Moore bound for irregular graphs, Graphs Combin. 18 (1) (2002) 53–57.

[2] R.C. Baker, G. Harman, J. Pintz, The difference between consecutive primes II, Proc. London Math. Soc. 83 (3) (2001) 532–562.

[3] G. Blakley, P. Roy, A Hölder type inequality for symmetric matrices with nonnegative entries, Proc. Amer. Math. Soc. 16 (1965) 1244–1245.

[4] W.G. Brown, On graphs that do not contain a Thomsen graph, Canad. Math. Bull. 9 (1966) 281–285.

[5] F. Chung, R. Graham, Sparse quasi-random graphs. Special issue: Paul Erdős and his mathematics, Combinatorica 22 (2) (2002) 217–244.

[6] D. de Caen, L.A. Székely, The Maximum Size of 4- and 6-Cycle Free Bipartite Graphs on m, n Vertices. Sets, Graphs and Numbers (Budapest, 1991), Colloquium Mathematical Society János Bolyai, vol. 60, North-Holland, Amsterdam, 1992, pp. 135–142.

[7] P. Erdős, On the graph theorem of Turán, Mat. Lapok 21 (1970) 249–251.

[8] P. Erdős, T. Gallai, On maximal paths and circuits of graphs, Acta Math. Acad. Sci. Hungar. 10 (1959) 337–356.

[9] P. Erdős, A. Rényi, V.T. Sós, On a problem of graph theory, Studia Sci. Math. Hungar. 1 (1966) 215–235.
[10] P. Erdős, M. Simonovits, Compactness results in extremal graph theory, Combinatorica 2 (3) (1982) 275–288.
[11] W. Feit, G. Higman, The nonexistence of certain generalized polygons, J. Algebra 1 (1964) 114–131.
[12] E. Györi, C_6-free bipartite graphs and product representation of squares. Graphs Combin. (Marseille, 1995), Discrete Math. 165/166 (1997) 371–375.
[13] S. Hoory, The size of bipartite graphs with a given girth, J. Combin. Theory (Ser. B) 86 (2) (2002) 215–220.
[14] D. Kühn, D. Osthus, 4-cycles in graphs without a given even cycle, J. Graph Theory 48 (2) (2005) 147–156.
[15] F. Lazebnik, V.A. Ustimenko, A.J. Woldar, Polarities and $2k$-cycle-free graphs, Discrete Math. 197/198 (1999) 503–513.
[16] L.P. Peykre, An automorphism of the geometry of the group $B_2(2^n)$ as a polynomial mapping, Ukrainian Math. J. 44 (11) (1992) 1530–1534 (in Russian).
[17] A.F. Sidorenko, Inequalities for functionals generated by bipartite graphs, Diskret. Mat. 3(3) (1991) 50–65 (in Russian); translation in Discrete Math. Appl. 2(5) (1992) 489–504.
[18] J. Tits, Théorème de Bruhat et sous-groupes paraboliques, C. R. Acad. Sci. Paris 254 (1962) 2910–2912.