On a bound for the maximum number of $C_4$ in a 4-cycle free bipartite graph

GENE FIORINI and FELIX LAZEBNIK
Department of Mathematical Sciences
University of Delaware
Newark, DE 19716

Abstract. Let $G$ be a 4-cycle free bipartite graph on 2n vertices with parts of equal cardinality $n$ having $e$ edges. Let $a(G)$ denote the number of cycles of length 8 in $G$. We prove that for $n \geq 4$, $a(G) \leq 3 (\frac{e}{n} - \frac{n}{2})^2$, where $t(t-1)(t-2)$, if $t \geq 2$. If $G$ is extremal with respect to the number of 8-cycles, then $r_n < \frac{2}{3}$, where $r_n = \frac{1}{2} + \frac{1}{n}$. This implies that $a(G) \leq 3 (\frac{e}{n} - \frac{n}{2}(r_n - \frac{3}{2}))$. Furthermore, if $G_n$ is the incidence point-line graph of a finite projective plane of order $q$, and $n_k = q^2 + q + 1$, then $a(G_n) = \frac{(q^2 + 1)(q^2 + q)}{2} \leq 3 (\frac{e}{n} - \frac{n}{2}(r_n - \frac{3}{2}))$, and $G_n$ is "close" to being extremal in this sense.

Section 1: Introduction.

Let $G = G_n$ denote a family of simple graphs of order $n$. For a simple graph $H$ and $G \subseteq G_n$, let $(G, H)$ denote the number of subgraphs of $G$ isomorphic to $H$. Let $h(n) = h(G, H) = \max \{ (G, H) | G \subseteq G_n \}$. If $G \subseteq G_n$, is fixed $h(G, H)$ as an extremal. The problem of finding $h(G, H, n)$, has been studied extensively and is considered as central in extremal graph theory. Though it is hopeless in whole generality, some of its instances have been solved. Often the results are concerned with bounds on $h_n$ and partial description of the extremal graphs. For example, if $A_n$ denotes the complete graph of order $n$, $H = K_2$, and if $A_n = (G, H, n)$, then the solution is given by the famous Turán Theorem. For the same $H$, if $A_n$ denotes the complete bipartite graph with partition classes sizes $n$ and $t$, and $G$ is the family of all $(m, n)$-bipartite graphs with no $A_m$, we have the so-called, Zarankiewicz problem. These and many other examples can be found in [2]. For some later results see [4,5,6].

All missing definitions can be found in [2]. Let $V(G) = E(G)$ denote the set of vertices and edges of a graph $G$, $\varepsilon = \varepsilon(G) = |E(G)|$. The neighborhood of a vertex $v \in V(G)$ is denoted by $N(v) = \{v \in G(v)\}$, and the degree of vertex $v$ in $G$ by $deg(v)$.

For $S \subseteq V(G)$, define $N(S) = \bigcup \{N(v) \mid v \in S\}$. If $G$ contains a cycle, the girth of $G$ is the length of a shortest cycle in $G$. For a positive integer $n$, $n \geq 2$, let $G_n(n, n)$ be the class of bipartite graphs on 2n vertices with partitions of equal cardinality $n$ and girth at least 6 (i.e. $4$-cycle free). Let $G \subseteq G_n(n, n)$ have partition $(V_1(G), V_2(G))$ such that $V_1(G) = \{v_1, \ldots, v_{2n}\}$, $V_2(G) = \{u_1, \ldots, u_{2n}\}$. Let $x_i = \deg(v_i)$, $i = 1, \ldots, 2n$ and $u_i = \deg(u_i)$, $i = 1, \ldots, 2n$. A subset $S \subseteq \{2n\}$ is $3$-free if $\forall i, j, k \in S \subseteq \{2n\}$, $i \neq j \neq k$, $i \neq j \neq k$, $v_i \neq v_j \neq k$.

Let $S \subseteq \{2n\}$ be a partition of $2n$ and $P = \{p_1, \ldots, p_n\}$ and $L = \{l_1, \ldots, l_n\}$ be the point set and the line set of $\pi$, respectively. A bipartite graph $G_\pi$ with partition $(P, L)$ is said to be the incidence point-line graph of the projective plane $\pi$. Let $n_i = \deg(p_i)$, $i = 1, \ldots, n$.

Let $a(G)$ denote the number of 8-cycles in $G$. The main goal of this paper is to find a nontrivial upper bound for $a(G)$, where $G \subseteq G_n(n, n)$. The results are summarized below.

CONGRESSUS NUMER untium 99 (1994), pp. 191-197
Theorem 1. Let $G \in \mathcal{G}(n,n)$ be a 4-cycle free bipartite graph on 2n vertices with partition classes of size n. Then

(i) $\omega(G) \leq \left\lceil \frac{n}{2} \left( \frac{n}{2} - 1 \right) \right\rceil$, where $f(t) = t(t-1)(2t-2)(4t-3)-3$, $n_1 = \frac{n}{2}$, and $n \geq 4$.

(ii) If $G$ has $e$ edges and is extremal with respect to the number of 8-cycles, then $n_1 - 2 \leq e \leq n_2$, with $n_2$ as in (i) above.

Using the terminology of finite geometries (see [1]), Theorem 1 provides an upper bound for the number of quadrilaterals in near-linear spaces with n points and n lines. To prove Theorem 1, we will rely on the following additional facts. The first fact, Lemma A below, was motivated by the statement Lemma 3 from [2]. Its proof is immediate and we omit it.

Lemma A. Let $G$ be a connected graph and $H$ be a subgraph of $G$. Then $|V(H)| \leq |E(H)| / |E(G)|$.

\[ \omega(G) \leq \left\lceil \frac{n}{2} \left( \frac{n}{2} - 1 \right) \right\rceil, \]

Lemma B.[2] Suppose $G$ is a 4-cycle free bipartite graph with partition classes of cardinality $n \geq 2$. Let $e = \omega(G)$ be the size of $G$. Then $e \leq \frac{n}{2} \left( \frac{n}{2} - 1 \right) \frac{n}{2} - 3$ with equality if and only if $n = n_1 = n_2 = 2$ and $G = G_2$ is the perfect matching on a graph of a finite projective plane of order $q$. Then $G_2$ has the greatest number of edges among all graphs in $\mathcal{G}(n_1, n_2)$. In [3] it was shown that $G_2$ has greatest number of 8-cycles among all graphs in $\mathcal{G}(n_1, n_2)$. Theorem 1 above grew out of our attempts to prove that $G_2$ also has the greatest number of 8-cycles among all graphs in $\mathcal{G}(n_1, n_2)$.

Section 2: Proof of Theorem 1.

Let $G \in \mathcal{G}(n,n)$ with partition classes $V_1(G) = \{v_1, \ldots, v_n\}$, $V_2(G) = \{v_{n+1}, \ldots, v_n\}$. To count the number of 8-cycles in $G$ note that each 8-cycle in $G$ determines a 3-set of vertices $\{v_i, v_{i+j}, v_{i+k}, v_{i+j+k}\}$ in $V_1(G)$ such that no three of the vertices form an intersecting 3-set. Furthermore, each 8-cycle identifies two distinct partitions of $\{v_1, v_2, \ldots, v_n\}$ into pairs of intersecting 2-sets. That is, if an 8-cycle is identified by the sequence $v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8$, then $\{(v_1, v_2), (v_3, v_4), (v_5, v_6), (v_7, v_8)\}$ are two distinct partitions of $\{v_1, v_2, \ldots, v_n\}$ into pairs of intersecting 2-sets of vertices.

Conversely, it is obvious that each 4-set of vertices $\{v_i, v_{i+j}, v_{i+k}, v_{i+j+k}\} \in V_1(G)$ can be partitioned in at most three ways into pairs of intersecting 2-sets. This implies that the set $\{v_1, v_2, \ldots, v_n\}$ in $V_1(G)$ can be contained in the vertex sets of at most $3 \times 3 = 9$ cycles of length eight in $G$.

Clearly, the number of intersecting 4-sets in $V_2(G)$ is $\sum_{i,j,k}^n \delta_{\bar{G}}(v_1, v_2, v_3, v_4) = \sum_{i,j,k}^n (\sum_{i,j,k}^n \delta_{\bar{G}}(v_1, v_2, v_3, v_4))$. Also, the number of 4-sets in $V_2(G)$ such that exactly three of the four vertices form an intersecting 3-set of vertices $\sum_{i,j,k}^n \delta_{\bar{G}}(v_1, v_2, v_3, v_4) = \sum_{i,j,k}^n (\sum_{i,j,k}^n \delta_{\bar{G}}(v_1, v_2, v_3, v_4))$.

Therefore, the number of 4-sets in $V_2(G)$ such that no three vertices form an intersecting 3-set is

\[ \binom{n}{2} - \sum_{i,j,k}^n \delta_{\bar{G}}(v_1, v_2, v_3, v_4) \]

and $\delta_{\bar{G}}(v_1, v_2, v_3, v_4) = \binom{n-2}{2}$, where $f(t) = t(t-1)(2t-2)(4t-3)-3$, $n_1 = \frac{n}{2}$, and $n \geq 4$.

Hence, we have proved the following

Lemma 2.1. For $n \geq 2$, let $G \in \mathcal{G}(n,n)$ with $V_1(G) = \{v_1, \ldots, v_n\}$ and $z_i = \delta_{\bar{G}}(v_i)$, $i = 1, \ldots, n$. Then

\[ \omega(G) \leq \frac{1}{4} \left( \frac{n}{2} \right) \left( \frac{n}{2} - 1 \right) \sum_{i=1}^n f(z_i) \]

where $f(t) = t(t-1)(2t-2)(4t-3)-3$. Moreover, equality holds if every 4-set of $V_1(G)$ for which no subset of three vertices is intersecting determines exactly three 8-cycles of $G$. Since every pair of vertices in $V_1(G)$ is an intersecting pair of vertices, it is easy to see that every 4-set of vertices in $V_1(G)$, no three vertices of which form an intersecting set determines exactly three 8-cycles of $G$. Therefore, Lemma 2.1 implies that

\[ \omega(G) = \frac{1}{4} \left( \frac{n}{2} \right) \left( \frac{n}{2} - 1 \right) \sum_{i=1}^n f(z_i) \]

Equation (2.2) provides us with a lower estimate of the upper bound on the max $\omega(G)$. $G \in \mathcal{G}(n_1, n_2)$.

In our next step, we find an upper bound on $\omega(G)$ in terms of $e = \omega(G)$ and $n$. Define the function $F(x) = x(x-1)x \in \mathbb{R}$, by $F(x) = \sum_{i=1}^n f(z_i)$, where $f(t) = t(t-1)(2t-2)(4t-3)-3$, $\sum_{i=1}^n f(z_i) \leq 1$, $n \geq 4$.

Theorem 2.2. Let $G \in \mathcal{G}(n,n)$, $n \geq 4$. Then $\omega(G) \leq \frac{1}{3} \left( \frac{n}{2} \right) \left( \frac{n}{2} - 1 \right) \sum_{i=1}^n f(z_i)$, where $F(x) = x(x-1)$.

Proof: Let $G \in \mathcal{G}(n,n)$ have partition classes $V_1(G) = \{v_1, \ldots, v_n\}$, $V_2(G) = \{v_{n+1}, \ldots, v_n\}$. It is easy to show that for all $n \geq 2$, there exist real values $\alpha = \omega(n)$, $\beta = \beta(n)$, $0 < \alpha < 1$, $\frac{1}{2} < \beta < \frac{1}{3}$, such that $F(x) \leq \frac{1}{6} |(n-\alpha)(n-1-\beta)|$. This implies $f(t)$ is concave up on the interval $(\alpha, \beta)$. We consider two cases:

Case 1: $\alpha < z_i < \beta$ for all $i = 1, \ldots, n$. In this case, the equation $\sum_{i=1}^n f(z_i)$ is minimized when $z_i = \beta$ for all $i = 1, \ldots, n$. Therefore, $\omega(G) \leq \frac{1}{3} \left( \frac{n}{2} \right) \left( \frac{n}{2} - 1 \right) \sum_{i=1}^n f(z_i)$.
Since \( f \) is concave up on \((a, b)\) and \( y < x_i < b, i = 1, \ldots, n \), Jensen's inequality implies 
\[
F(\bar{x}) = \frac{1}{n} \sum_{i=1}^{n} f(x_i) \geq n \cdot f \left( \frac{1}{n} \sum_{i=1}^{n} x_i \right) = n \cdot f \left( \frac{y}{n} \right) = F(y).
\]

Therefore 
\[
\alpha(G) \leq \sum_{i=1}^{n} \frac{\left( \frac{n}{y} \right) - \frac{1}{n} f(\frac{y}{n})}{\frac{n}{y} - y} = \left[ \left( \frac{n}{y} \right) - \frac{1}{n} f(\frac{y}{n}) \right] \left( \frac{n}{y} - y \right).
\]

Case 2: \( z_i > \beta \) for some \( i \in \{1, \ldots, n\} \).

Suppose \( G \) has a vertex \( v \) of degree at least \( \beta \) and, relabeling if necessary, assume \( u = v_1 \). For all \( n \) for which there exists a vertex \( u \neq v_i \), we do not require \( N(v_i) \neq 0 \) for \( i = 1, \ldots, n \). Then with \( N(v_i) \neq 0 \), \( i = 1, \ldots, n \), which means that \( G \) contains a cycle of length \( n \). Therefore, Case 2 is equivalent to \( z_i > \beta \) and \( z_i < \beta \) for \( i = 1, \ldots, n \). We will address this case later.

First, suppose \( z_i = \infty \). Since \( G \) is a tree, \( |N(v_i)| \leq d(v_i) = \alpha(G) \leq \alpha(G_1) \leq 1 \) for all \( u \in V(G_1), \ i = 1, \ldots, n \). Therefore, \( \alpha(G_1) = 0 \) and, trivially, we have that \( \alpha(G) \leq 0 \) and \( \frac{1}{n} f(\frac{y}{n}) \leq 0 \) and the result is proved.

So, suppose \( \beta \leq z_i \leq \alpha(G_1) \leq 1 \). For all \( n \geq 2, \) there exist real values \( y = n(n) \) and \( \zeta = (\zeta, \zeta) \), \( 1 < \zeta < n, n-1 < \zeta < n \), such that \( f(x) = 2x^2 - 2n + \beta + (3 - 12n^2 - 4n - 3 \zeta) \zeta \zeta < 0 \) for all \( x \in (0, \infty) \). That implies \( f(x) \) is strictly increasing on the interval \((0, \zeta)\). Hence, \( \frac{1}{n} z_i \leq \alpha(G) \leq \alpha(G_1) \leq 1 \). Therefore, \( f(x) \geq f(\frac{\zeta}{n}) \) for all \( x \in (0, \zeta) \). Hence, \( \alpha(G) > 0 \) and, trivially, we have that \( \alpha(G) \leq 0 \) and \( \frac{1}{n} f(\frac{y}{n}) \leq 0 \) and the result is proved.

Finally, \( z_i \leq \frac{1}{2} + \sqrt{4n - 2} \). Theorem 2.4 implies \( \frac{z_i}{n} \leq \frac{1}{n} \) and so \( \frac{z_i}{n} \leq \frac{1}{n} \). Therefore, \( \frac{f(\frac{z_i}{n})}{n} \leq f(\frac{1}{n}) \) and so \( \frac{z_i}{n} \leq \frac{1}{n} \). Hence, \( \alpha(G) \leq \frac{1}{n} f(\frac{1}{n}) \) and so we have, 
\[
\alpha(G) \leq \frac{1}{n} \left( 1 - \frac{1}{n} f(\frac{y}{n}) \right) < \left[ \left( \frac{n}{y} \right) - \frac{1}{n} f(\frac{y}{n}) \right] \left( \frac{n}{y} - y \right).
\]

\[\text{Lemma 2.2. For } n \geq 2, \text{ suppose } G \in \mathcal{G}_n \text{ satisfies } \alpha(G) \leq \left( \frac{1}{2} \right)^{\frac{1}{n}} \text{ for all } \alpha(G) \text{ and } \frac{1}{n} \leq \frac{1}{n}. \text{ Then } r_n \geq \frac{m}{4}. \]

\[\text{Proof:} \text{ Theorem 2.2 implies } \frac{1}{n} \leq \frac{1}{n}. \text{ Here we show that } \frac{1}{n} \leq \frac{1}{n}. \text{ It is sufficient to prove the result for connected graphs of girth at least six and partition classes of equal cardinality. Indeed, suppose } G \text{ is disconnected. By induction on the number of components of } G \text{ it is easy to see that } G \text{ can be embedded in a connected graph } G' \in \mathcal{G}_{n \cdot v(n)} \text{ such that } \alpha(G) \leq \alpha(G'). \text{ Thus, suppose } G' \in \mathcal{G}_{n \cdot v(n)} \text{ is connected.} \]

\[\text{Let } P(G) \text{ represent the set of all paths of length three in } G. \text{ Define the function } g \text{ by } g(x) = x - 2 \text{ for } x \text{ such that } G \text{ is a bipartite graph. Therefore, the distance between } u \text{ and } v \text{ in } V(G), \text{ } (u, v) \text{ is at least } 2. \text{ If the distance is exactly } 2, \text{ since } G \text{ has girth at least six, the path of length } 2 \text{ joining } u \text{ and } v \text{ is unique. Therefore, the path from } u \text{ to } v \text{ through } u \text{ is unique and of length } 1. \text{ This implies there are } x_i = n - z_i \text{ paths of length } 3 \text{ with } u_i \text{ as one end vertex of those paths. Therefore, } |P(G)| \leq \sum \frac{1}{x_i}. \text{ and since } g \text{ is convex on the interval } (1, \infty), \text{ by Jensen's inequality, we have} \]

\[
|P(G)| \leq \sum_{i=1}^{n} g(x_i) \leq g \left( \frac{1}{n} \sum_{i=1}^{n} x_i \right) = g \left( \frac{1}{n} \right).
\]

\[\text{where } x = \frac{1}{n} \sum_{i=1}^{n} x_i. \text{ Since } g \text{ is convex on } [1, 2], \text{ and } \frac{1}{n} \leq \frac{1}{2} \text{ for } n \geq 2, \text{ inequality (2.4) implies} \]

\[|P(G)| \leq \frac{1}{n} \left( \frac{n}{2} - 1 \right) = \frac{1}{n} \left( \frac{n}{2} - 1 \right).
\]

\[\text{Clearly, } F(x) = \frac{n}{x} \text{ depends on } x \text{ which is strictly increasing on the interval } (0, \infty), \text{ and so } n \frac{1}{n} f(\frac{1}{n}) \leq F(\frac{1}{n}) \text{ and so } n \frac{1}{n} f(\frac{1}{n}) \leq F(\frac{1}{n}) = n \frac{1}{n} f(\frac{1}{n}) \text{ in terms of } n \text{. Indeed, we should like to determine a bound on } \alpha(G) \text{ that depends only on } n, \text{ the number of vertices in a partition class of } G. \text{ When } n = \frac{1}{2} + \sqrt{4n - 2} \text{, equation (2.3) and inequality (2.3) imply that any bound on } \alpha(G) \text{ must be at least } \left( \frac{1}{2} \right)^{\frac{1}{n}}. \text{ With this in mind we can determine an appropriate range } \alpha(G). \]

\[\text{Figure 2.1} \]

\[\text{Figure 2.1} \]

\[\text{Figure 2.1} \]
Let \( a = u v u' \in \mathcal{P}(G) \). Let \( S = N(u) \setminus \{v\}, T = N(v') \setminus \{u\} \). At this point we use a construction (see Figure 2.1) similar to one which appears in the proof of Lemma 2 in [7] and which allows us to estimate the number of \( S \)-cycles passing through \( a \) presently. We describe it presently.

Let \( V_a = N(u) \cup N(v') \cup N(u) \cup N(T) \). Define \( G_a = (V_a, E) \) to be the subgraph of \( G \) induced by the vertex set \( V_a \). Let \( S_a = \{E(G \setminus E(G_a)) \} \) and \( a_0 = \{a_1 \setminus \{a_0\} \} \).

Lemma A implies \( a_0 = a_1 \geq a_2 \geq \cdots \).

Let \( A_a = \{(v_i, v_j) \in E(G) | v_i \in T \text{ and } v_j \in S\} \) and \( B_a = \{(v_i, v_j) \in E(G) | v_i \in V(T) \text{ and } v_j \in S\} \). Define \( A_a \setminus (A_a \cup B_a) \) as the graph obtained from \( G_a \) by removing the edges of \( A_a \cup B_a \). Note that the graph \( G_a \setminus A_a \cup B_a \) is a tree with \( 2n - 2a_0 \) vertices, where \( |A_a| \) is the number of \( S \)-cycles through \( a \) and \( |E(G_a)| \) is the number of \( S \)-cycles through \( a \). Therefore \( a = \{E(G \setminus E(G_a)) \} + \{E(G_a) \} \), which implies for each \( e \in \mathcal{P}(G) \),

\[
e - 2n + 2 \geq |B_a|.
\]

Now, summing both sides of (2.6) over all paths of length three, the left hand side expression becomes

\[
\sum_{e \in \mathcal{P}(G)} (e - 2n + 1) = |B_a| (e - 2n + 1) \geq (e - 2n + 1)(n - a_1) \frac{2}{2}.
\]

(2.7a)

and the right hand side expression becomes

\[
\sum_{e \in \mathcal{P}(G)} |B_a| = 3|B_a| \geq \frac{3}{2} \frac{(n - a_1)^2}{2}.
\]

(2.7b)

Therefore, combining both expressions we have \( 2(e - 2n + 1)(n - a_1) \geq \frac{3}{2} \frac{(n - a_1)^2}{2} \).

Solving this inequality for \( e \) we see that \( e \geq r_2^0 - 3r_2 + 6r_a - 3 > r_2^0 - 3r_2 + 3r_a - 2 = (r_2 - 2)\frac{r_2 - 2}{2} - r_a - 2, \) which implies \( \frac{4}{2} > r_a - 2. \)

We are now ready to prove the first statement of Theorem 1.

Proof of Theorem 1, part (ii): If \( a < a \leq \left( \frac{r_2}{2} \right)^{r_2} \), then clearly \( \left( \frac{r_2}{2} \right)^{r_2} < 3 \frac{r_2}{2} \frac{r_2 - 2}{2} \) and the result is proved. If \( a > \left( \frac{r_2}{2} \right)^{r_2} \), Theorem 2.2 implies that \( a < 3 \frac{r_2}{2} \frac{r_2 - 2}{2} \). Since \( f \) is strictly increasing on the interval \( (n, c) \), where \( 1 < c < 2 \) and \( d - 1 \leq c < n \), Lemma 2.3 implies \( f \left( \frac{r_2}{2} \right) > f (r_2 - 2) \) and so \( \mathcal{C}(G) < 3 \frac{r_2}{2} \frac{r_2 - 2}{2} \frac{r_2 - 2}{2} \) and the result is proved.

We make a final comment on the lower bound on \( c \) appearing in Lemma 2.3. Central to the proof of Lemma 2.3 is the calculation of \( | \mathcal{P}(G) | \), the number of paths of length three in \( G \setminus \{a\} \). Clearly, there exists \( a \in \mathcal{P}(G) \) for which \( 0 < a < a \) and \( |A_a| > 0 \). Hence, it would seem that if we could find functions \( \alpha \) and \( \beta \) such that \( \sum_{|A_a| > 0} |A_a| > \alpha |n| \) and \( \sum_{|A_a| > 0} |A_a| > \beta |n| \), then \( \alpha |n| \) and \( \beta |n| \) could be inserted into the inequalities (2.6), (2.7a), and (2.7b) to produce a tighter bound on \( \mathcal{C}(G) \). However, we have been unable to find such functions so far. This would imply that the incidence point-line graph of order \( \beta \) instead is the best candidate for the extremal graphs for those values of \( n \) for which it exists. However, such a proof would require another technique.

Acknowledgment. The authors are grateful to Professors R. D. Baker and G. L. Ebert for discussions on the topics related to this paper.

References


