Even Cycles in Directed Graphs

F. R. K. Chung
Bellcore
445 South Street
Morristown, NJ 07962

Wayne Goddard
D. J. Kleitman
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract

It is proved that every strongly connected directed graph with n nodes and at least $\lfloor (n+1)^2/4 \rfloor$ edges must contain an even cycle. This is best possible, and the structure of extremal graphs is discussed.

1 Introduction

A directed graph G is a set of nodes $N(G)$ together with an edge set $E(G)$ consisting of ordered pairs of $N(G)$. A *path* in G from a node u to a node v is a sequence of distinct nodes $u = v_0, v_1, \ldots, v_t = v$ so that (v_i, v_{i+1}), $i = 0, \ldots, t-1$, are in $E(G)$. A path from u to v together with the edge (v, u) is called a *cycle*. If a cycle contains an even number of edges, it is said to be an even cycle. A *hamiltonian* cycle is a cycle which contains every node in the graph; a hamiltonian graph is one which has such a cycle. A graph G is said to be *strongly connected* if for every pair of nodes u and v in $N(G)$, there is a path from u to v. In this paper, we consider directed graphs containing no loop (i.e. $(v, v) \notin E(G)$ for any v).

In this paper, we prove that every strongly connected graph on n nodes and at least $\lfloor (n+1)^2/4 \rfloor$ edges must contain an even cycle, thus settling a conjecture of Brualdi and Shader [3, 4]. This is best possible and we give several examples of *edge-critical* graphs which are strongly connected directed graphs on n nodes and $\lfloor (n+1)^2/4 \rfloor - 1$ edges containing no even cycle. In particular we characterize the edge-critical graphs which are hamiltonian.

Two simple examples of edge-critical graphs are the following:

1. The node set of H_n can be partitioned into three parts A, B and a node v, where $|A| = \lfloor (n-1)/2 \rfloor$ and $|B| = \lceil (n-1)/2 \rceil$. $E(H_n) = \{ (a, b) : a \in A,\ b \in B \} \cup \{ (v, a) : a \in A \} \cup \{ (b, v) : b \in B \}$.

2. The node set of L_n consists of $v_0, v_1, \ldots, v_{n-1}$ which form a cycle. In addition, (v_i, v_j) is an edge if $i - j$ is a positive even number.

For a node v, the in-degree of v is the cardinality of $\{ u : (u, v) \in E(G) \}$, and the out-degree of v the cardinality of $\{ u : (v, u) \in E(G) \}$. If for all v the in- and out-degree of v is d, we say G is d-regular. We say nodes u and v are adjacent or u is adjacent with v or u is a neighbor of v if either (u, v) or (v, u) is in $E(G)$. The adjacency matrix of a directed graph G on n nodes is the 0-1 $n \times n$ matrix M with $M(u, v) = 1$ if and only if $(u, v) \in E(G)$. For a subset S of $N(G)$, the induced subgraph of G on S has node set S and edge set $\{ (a, b) \in E(G) : a, b \in S \}$. A node v is said to be a *cut point* if by removing v and edges incident to v from G the resulting graph is no longer strongly connected.

A related result of S. Friedland [8] states that every 7-regular directed graph contains an even cycle. Recently, C. Thomassen [17] showed that every d-regular directed graph, $d \geq 3$, contains an even cycle. Earlier, C. Thomassen [14] proved that for each positive integer k, there is a directed graph of minimum out-degree k and with no even cycle. Furthermore, every directed graph on n nodes and minimum out-degree $\lfloor \log_2 n \rfloor + 1$ contains an even cycle and this could be improved to $\frac{1}{2}\log_2 n$. (In the weighted version $\lfloor \log_2 n \rfloor + 1$ is best possible.)

To determine if a directed graph contains an even cycle is a surprisingly difficult problem, which was first raised by D. H. Younger (see [22]). In spite of attempts by many researchers [14, 15, 21], it is not known if this problem is NP-complete or not. The result in this paper provides a simple solution to the above problem when the graph contains "many" edges.

The problem of finding even cycles in a directed graph is closely related to the problem of converting some 1-entries of a (0,1)-matrix A to -1 so the resulting matrix B satisfies the property that the determinant of B is equal to the permanent of A [1]. Therefore a hard problem [20] of computing the permanent is transformed into an easy one for some graphs. C. H. C. Little [10] and Seymour and Thomassen [13] give characterizations for matrices for which such conversion is possible. In particular, a directed graph G containing no even cycles satisfies

$$\text{determinant}(I + A) = \text{permanent}(I + A),$$

where A is the adjacency matrix of G and I is the identity matrix.

The result in this paper implies that the largest number of 1's in an irreducible $n \times n$ (0,1)-matrix B with the same value for its determinant and permanent (see [2, 3, 21]) is $\lfloor (n+1)^2/4 \rfloor + n - 1$.

2 Hamiltonian Graphs

Let G be a directed graph without an even cycle.

Lemma 1. *Suppose there are three nodes a, b and c and $(a,b),(b,c)$ and (a,c) are edges of G (so that (a,b,c) forms a "transitive" triple). Then all paths from c to a contain b and thus b is a cut point.*

Proof: Suppose there is a path P from c to a which does not contain b. Then there must be an even cycle which is either adding (a, c) to P or adding $(a, b), (b, c)$ to P. Therefore all paths from c to a must contain b. □

The following establishes the result for hamiltonian graphs:

Lemma 2. *Let T be a cycle in G. Then any path in T on m nodes spans at most $f(m) = \lfloor (m^2 + 2m - 3)/4 \rfloor$ edges.*

Proof: By induction on m. Let P denote a path $v, v_1, \ldots, v_{m-1}$ in cycle T. If v is adjacent with at most $\lceil m/2 \rceil$ nodes of P, then we are done since $f(m) - f(m-1) = \lceil m/2 \rceil$. We may therefore assume that v is adjacent with at least $\lceil m/2 \rceil + 1$ nodes of P. Then v must be connected to two consecutive nodes in P of the form v_{2i} and v_{2i+1}. The edge between v and v_{2i} must be (v_{2i}, v), and the edge between v and v_{2i+1} must be (v, v_{2i+1}), else there is an even cycle.

Observe that $P' = v, v_1, \ldots, v_{2i}$ and $P'' = v, v_{2i+1}, \ldots, v_{m-1}$ form paths and may be extended to cycles. Further, (v_{2i}, v, v_{2i+1}) is a transitive triple, so that the removal of v from G disconnects v_{2i} from v_{2i+1}. This means that all edges between the two paths P' and P'' not involving v must be directed from P' to P''. Now an edge (v_a, v_b) with $a < b$ requires a and b to have opposite parities, or it produces an even cycle. Thus there are at most $i(m - 2i - 1)$ such edges.

Since every edge in P is either of this form or within P' or P'', P can have at most $f(2i + 1) + f(m - 2i) + i(m - 2i - 1) \leq f(m)$ edges. □

We now construct all edge-critical hamiltonian graphs on $2k + 1$ nodes. Start with a directed triangle T_1 on nodes v_1, v_2, v_3. Then for $i = 2, \ldots k$, create triangle T_i by introducing two new nodes, say v_{2i} and v_{2i+1}, and three edges, so that v_{2i} and v_{2i+1} form a directed triangle with some node a_{i-1} of triangle T_{i-1} (in that order). This yields a spine S_k on $2k + 1$ nodes and $3k$ edges. Then to construct G_k, take S_k and for every pair of distinct nodes $v \in T_i$ and $w \in T_j$, $i \leq j$, add (if necessary) the edge (v, w) iff the path from w to v in S_k has even length (number of edges).

It is easy to show by induction on k that G_k so constructed has $k(k + 2)$ edges: except for a_{k-1}, every other node is adjacent with exactly one of v_{2k} and v_{2k+1}. One may also show that G_k is hamiltonian by induction on k. Let H denote the hamiltonian cycle of the subgraph on the first $k-1$ triangles and x the predecessor of a_{k-1} in H. Then the path from a_{k-1} to x in the spine has even length; thus the same is true of the path from v_{2k} to x in S_k, and hence (x, v_{2k}) is an edge of G_k. Hence the edge (x, a_{k-1}) in H can be replaced by the three edges (x, v_{2k}), (v_{2k}, v_{2k+1}) and (v_{2k+1}, a_{k-1}) yielding the hamiltonian cycle of G_k.

It is less trivial to show that G_k has no even cycle. Observe that every a_i is a cut node: any path from T_{i+1} (and beyond) to T_i (and before) must pass through a_i. Let C be any cycle, and let v and w be nodes of C in the triangles of least and greatest index respectively. Then the w-to-v portion of C is the path along S_k; say it has $p+1$ nodes. Let $w = h_0, h_1, \ldots, h_j = v$ denote the remainder of C. Let s_i be the distance from h_i to h_{i-1} in the spine ($i = 1, \ldots, j$). It may be shown by induction on j that $\sum_i s_i = p + 3(j-1)$. But all the s_i are even. Hence $p + j$ is odd; but this is the length of C.

We now show that every graph G on $2k+1$ nodes that is hamiltonian and edge-critical has the above form. This we do by induction on k, so assume $k \geq 2$. G has maximum degree at least $k+2$, so by the proof of the above lemma, G has a cut point v whose removal partitions the nodes into parts S_1 and S_2 such that all edges between the parts are directed from S_1 to S_2. Further, the proof shows that $S_1 \cup \{v\}$ and $S_2 \cup \{v\}$ induce edge-critical hamiltonian graphs G_1 and G_2. By the inductive hypothesis G_1 and G_2 have the requisite form.

We must show that v is in the last triangle of G_1. This is immediate if G_1 is a triangle, so let the last two triangles of the spine S_m of G_1 be T_{m-1} and T_m. Let a denote the node that T_{m-1} and T_m have in common, y the out-neighbor of a in T_m, and x the in-neighbor of a in T_{m-1}. Then (x, a, y) forms a transitive triple in G_1 (the path from y to x in S_m has length four). Thus a must be on every path in G from y to x; in particular on such a path that goes via S_2. Hence v must be in the last triangle of G_1. Similarly v must be in the first triangle of G_2.

There can be an edge from $v \in S_1$ to $w \in S_2$ only if the path from w to v has even length. But to obtain the correct number of edges, we must then have all possible edges from S_1 to S_2. This means that G is one of the G_k constructed above.

3 The General Result

We will prove the following:

Main Theorem. *A strongly connected directed graph G on n nodes without an even cycle contains at most $f(n) = \lfloor (n^2 + 2n - 3)/4 \rfloor$ edges.*

Further, such a G with $f(n)$ edges consists of a maximal hamiltonian edge-critical subgraph C (on $2r + 1$ nodes say), and an acyclic complete bipartite graph on the remaining nodes, having as equal as possible size parts, each node of which is connected to $r + 1$ of the nodes of C.

Roughly speaking, the proof of the main theorem can be described as follows. We partition the node set of our directed graph into a sequence of subsets, which we call *layers*, so that the subgraph induced by each initial segment of layers is strongly connected. Each layer consists of either the nodes of a directed cycle, the nodes of a (non-trivial) directed path, or a single node. Thus we refer to *cycle*, *path* and *singleton* layers.

We establish first an upper bound on the number of edges inside a layer. We then establish an upper bound B on the number of edges between any two layers. Finally, we show that this bound B cannot be attained between every pair of layers; indeed we derive an upper bound on the number of pairs of layers for which the bound B is attained.

We assume throughout that the directed graph G has n nodes and no even cycle. The first layer L_0 consists of the nodes of a maximum cycle in G. Thereafter, we construct a sequence of layers L_j ($j \geq 1$) of nodes of G as follows. Let S_j denote $\bigcup_{i \leq j} L_i$.

1. If there is a cycle T of $G - S_{j-1}$ such that there are edges between T and S_{j-1} in both directions, then we let the nodes of L_j be those of one such cycle of maximum length.

2. Otherwise, if there is a node v of $G - S_{j-1}$ such that there are edges between v and S_{j-1} in both directions, then we let L_j be $\{v\}$.

3. Otherwise, we let the nodes of L_j be those of a maximum length path $P = x \ldots y$ in $G - S_{j-1}$, subject to there being an edge from S_{j-1} to x, and an edge from y to S_{j-1}.

Observe that L_j is (part of) a cycle in S_j and that S_j is strongly connected.

Now, we define the *excess* between two layers of cardinalities a and b as the number of edges between them minus $ab/2$. Similarly we define the excess inside a layer of cardinality a as the number of edges within it minus $\binom{a}{2}/2$.

To analyze the number of edges between layers, we consider an auxiliary undirected graph H whose nodes correspond to the layers in G. There is an edge between two nodes in H if and only if there is positive excess between the corresponding layers in G (i.e. the number of edges exceeds the product of the layers' cardinalities). We will establish a tight upper bound on the number of edges in H. In this regard we define a forest F of the edges of H as follows: for each node w of H, we include in F the edge of H linking w to the node corresponding to the layer of smallest index such that there are edges in both directions between the corresponding layers in G, if such a layer exists.

In order to prove the main theorem we will prove the following three claims:

Claim 1: The number of edges of G within any layer on m nodes is at most $f(m)$.

Claim 2: The excess in G between any two layers is at most 1.
If equality holds then there is an edge between the corresponding nodes in F. The excess between the first layer and a non-singleton layer is at most -1.

Claim 3: The graph $H-F$ has no triangles, and the first layer is isolated in $H-F$.

Actually, we have already established Claim 1 (see Lemma 2).

Proof of the main theorem: We show that the above three claims are sufficient to establish that the excess in G is at most $3n/4-3/4$. Let x denote the number of layers of odd cardinality, excluding L_0, and y the number of even cardinality. The contribution to the number of edges in G is in two parts:

- *Inside a layer:* In a layer of size m there is an excess of at most $3m/4-3/4$ if m is odd and at most $3m/4-1$ if m is even, by Claim 1. Hence the excess inside layers is at most
$$3n/4 - y - 3x/4 - 3/4.$$
- *Between layers:* We say that a potential edge e between two node of H is "odd" if the layers corresponding to both its ends have odd cardinality; otherwise it is "even". By

Claim 2, if e is odd, then it contributes at most $+1/2$ to the excess in G if it is in H and $-1/2$ otherwise. If e is even, then it contributes $+1$ if in F and 0 otherwise. We may reassign the contributions and say: Any edge contributes $+1$ if it is in F; and an odd edge contributes $+1/2$ if it is in $H-F$ and $-1/2$ otherwise.

There are $\binom{x+1}{2}$ potential odd edges. By Claim 3, at most $x^2/4$ of these are in $H-F$. There are at most $x+y$ edges in F since it is a forest. Hence the excess between layers is at most

$$\frac{1}{2} \cdot \frac{x^2}{4} - \frac{1}{2}\left(\binom{x+1}{2} - \frac{x^2}{4}\right) + (x+y) = 3x/4 + y.$$

Thus the overall excess is at most $3n/4 - 3/4$.

We can also deduce the partial characterization of edge-critical graphs. The first layer L_0 is a cycle layer. By Claim 2, equality in the main theorem requires all layers except L_0 to be singleton layers. Also, a singleton layer must have positive excess with L_0; indeed it must be joined to L_0 by an edge in F, which means that it has edges in both directions with L_0 in G. Further $G - L_0$ must be complete bipartite having as equal as possible size parts (by Turán's theorem [18]). And by the construction of the layers, $G - L_0$ must be acyclic. □

It remains to establish Claims 2 and 3.

4 Proof of Claim 2

Lemma 3. *Let T be any cycle $w_0, w_1, \ldots, w_{2r}, w_0$ in G, and let v be a node in $G - T$. Then v can connect to only one pair of consecutive nodes in T, say (w_0, w_1). Thus there are at most $r+1$ edges between v and T. If the equality holds, then v connects to w_0 and to $w_1, w_3, \ldots, w_{2r-1}$.*

Proof: Let there be edges between v and both w_0 and w_1 in G. Of the four possible pairs of directions for these edges, one (viz. (w_0, v) and (v, w_1)) produces an even cycle with the rest of T.

Suppose that the two edges are (w_1, v) and (v, w_0) so that $\{v, w_0, w_1\}$ forms a directed cycle. It is immediate that, independent of direction, to avoid even cycles all other edges between T and v involve only odd-indexed nodes of T.

Suppose now that both edges are directed from T towards v. Then by Lemma 1, removal of w_1 destroys all paths from v to w_0. Thus any path from v to w_1 is internally disjoint from T. Hence, all edges between v and T are directed toward v, and moreover can exist only from odd-indexed nodes w_a in T. □

We define the *in-node* v of a layer L_j ($j \geq 1$) as follows: if L_j is a cycle layer then v is any node such that there is an edge from S_{j-1} to v; if L_j is a path layer then v is the first node on the path; and if L_j is a singleton layer then v is the node itself. The *out-node* of L_j is defined analogously.

Lemma 4. *Let node v in $G - S_j$ be adjacent with two consecutive nodes x and y in layer L_j.*
1) If both edges are directed towards v, then L_j is a cycle layer, $j \geq 1$, y is the unique in-node of L_j, and any path from v to S_{j-1} avoids L_j.
2) If the edges go in opposite directions then $\{v, x, y\}$ forms a directed triangle.

Proof: If the edges go in opposite directions then $\{v, x, y\}$ forms a directed triangle, else there is an even cycle. So suppose both edges are directed towards v. Then any path P from v to x must go via y (by Lemma 1). Consider the first node w of P that is in S_j. If w is in L_j, then it must be y. But then we can replace the edge (x, y) by (x, v) and P, thereby contradicting the maximality of L_j. Thus w must be in S_{j-1}. Hence y must be the unique in-node from S_{j-1}, and so L_j must be a cycle layer. □

We say a node v of $G - S_j$ is *special* with respect to layer L_j if it is adjacent with more than half the nodes in L_j. Further, we say that v is *backtrack-special* if it is in a directed triangle with consecutive nodes of L_j. It is *up-special* if all the edges from L_j are directed towards v, and *down-special* if they are directed away from v. The above lemma shows that there is no up- or down-special node with respect to the first layer L_0.

Lemma 5. *Let L_j be a path layer with path $x \ldots y$ and assume that node v of $G - S_j$ is special with respect to L_j but not backtrack-special. Then L_j has an odd number of nodes, and v is adjacent with every alternate node on L_j with (x, v) and (v, y) being edges. We say such a node v is detour-special.*

Proof: By Lemma 4, if v is not backtrack-special then it has no consecutive neighbors on L_j. Thus L_j is odd, and v is adjacent with every alternate node starting with x.

Suppose edges between v and L_j go only one way; say towards v. Then consider a path P from v to S_{j-1}. If P is disjoint from L_j, then it contradicts the maximality of the path L_j since edge (y, v) is in G. If it meets L_j after x, then again L_j can be lengthened. But if it meets x first, then there is a longer cycle—by using L_j, (y, v) and P—and this cycle is anchored (with y and x) to S_{j-1}. This is a contradiction.

So edges go both ways. Then the edges directed away from v come after those directed towards v, else an even cycle results; thus (x, v) and (v, y) are edges. □

Lemma 6. *If nodes v and w of $G - S_j$ are backtrack-special with respect to L_j, then they do not lie together in a cycle outside S_j.*

Proof: Let nodes v and w be backtrack-special w.r.t. L_j. We will show that, for some parity π, there are paths in both directions between v and w which use only S_j and whose lengths have parity π. This contradicts the non-existence of even cycles in G if v and w are in a cycle in $G - S_j$.

Suppose L_j is a path layer. Then we may assume that the path is $\ldots x_0 x_1 x_2 \ldots$, that v is connected to x_0 and x_1, and that w is connected to x_i and x_{i+1} for $i \geq 0$. If $i = 0$ then there are odd paths between v and w in both directions through L_j; if $i = 1$ then even paths. Otherwise suppose that i is even. Then v is connected to x_{i+1} (by Lemma 3). It is not possible to orient this edge without producing paths of the same parity in both directions between v and w through S_j. The case when i is odd, or when L_j is a cycle layer, is similar. □

Lemma 7. *Let nodes v and w be consecutive in layer L_k and have a common neighbor x in lower layer L_j. (I.e. $k > j$.)*
1) If both edges are directed away from x, then L_k is a cycle layer and v is the unique out-node of L_k.
2) If the edges go in opposite directions and L_k is a cycle layer, then $\{v, w, x\}$ forms a directed triangle.

Proof: Suppose (x, v) and (x, w) are in G. Then any path P from w to x goes via v (by Lemma 1). Thus v is the unique out-node to S_{k-1} and thus L_k must be a cycle layer. If the edges go in opposite directions and L_k is a cycle layer, then $\{v, w, x\}$ forms a directed triangle, else an even cycle results. □

We define a (potential) edge between two nodes in H as *one-way* or *two-way* depending on whether there are edges in one direction or in both directions between the corresponding layers in G.

We now prove Claim 2:

Lemma 8. *Let layer L_k be above layer L_j (i.e. $k > j$). Then the following table gives upper bounds on the excess between L_j and L_k in G, depending on whether there are edges in G between the layers in one direction or in both directions:*

L_k	L_j	*one-way*	*two-way*	*Comments*
C	C	$+1/2$	$-1/2$	$-3/2$ *excess if* $j = 0$
C	P	0	$+1$	0 *excess in* $H - F$
P	C	$+1/2$	0	-1 *excess if* $j = 0$
P	P	0	0	
C	S	$+1/2$	$-1/2$	
S	C	$+1/2$	$+1/2$	$+ve$ *excess for* $j = 0$ *requires* 2*-way*
S	P	0	$+1$	0 *excess in* $H - F$
P	S	$+1/2$	$+1/2$	
S	S	$+1/2$	–	

Proof:

1. *Cycle above Cycle:* Suppose L_k has an up-special node u. Then by Lemma 4 all edges are directed towards L_k. Further, u must be adjacent with the unique in-node of L_j. So there are at most $(|L_k| + 1)/2$ up-specials (by Lemma 3) and the bound follows. Otherwise, the only chance of positive excess is to have a backtrack-special node, but there can be at most one such node by Lemma 6.

   Suppose there is $-1/2$ excess and $j = 0$. Then L_k is a triangle $abca$ with c (say) backtrack-special. Let cycle L_0 be $x_0, x_1, \ldots, x_{2r}, x_0$. Then a and b must have r

neighbors on L_0. We claim that a cannot have two consecutive neighbors in L_0; for by Lemma 4, they would form a directed triangle with a, which yields a contradiction as in the proof of Lemma 6. Let c's consecutive neighbors in L_0 be x_0 and x_1. Then, by the lack of even cycles and the maximality of L_0, a cannot be adjacent with x_0 or x_2. Thus a must be adjacent with x_1. Similarly b must be adjacent with x_0. These two edges cannot be in opposite directions (e.g. $x_1abx_0x_1$ would be a 4-cycle). So without loss of generality we may assume that (b, x_0) and (a, x_1) are edges. Thus (b, c, x_0) is a transitive triple, and thus all edges are directed from b to L_0. It then follows that (b, x_3) is an edge (indices modulo $2r+1$). But that means there is a cycle of length $2r+3$ in G, viz. $x_3x_4 \dots x_1cabx_3$, which contradicts our choice for L_0.

2. *Cycle above Path:* There is at most one backtrack-special node. Further, detour-special nodes, which are only possible if the path layer L_j has odd cardinality, are nonadjacent (by the lack of even cycles). So positive excess requires edges to go both ways in G. By the layering strategy, L_k cannot be connected both ways to S_{j-1}; thus this edge is in F.

3. *Path above Cycle:* Observe that by the layering strategy, no node in the path layer L_k can have edges both ways to L_j. Up-specials w.r.t. L_j are non-adjacent and the same with down-specials. Suppose L_k has a down-special node d and an up-special node u. By the maximality of L_j, u must come after d. By Lemma 4, the path from u to S_{j-1} must avoid L_j, and the path from S_{j-1} to d must avoid L_j, so that u and d lie in an odd cycle disjoint from L_j. But then there is an even cycle, since there is a path of length two and one of length three from d to u using only L_j. Furthermore, if $j = 0$, then L_k can have no special node and so the excess is at most $-|L_k|/2$.

4. *Path above Path:* By Lemma 5, special nodes have edges in both directions with L_j, which is impossible by the layering strategy.

5. *Cycle above Singleton:* By Lemma 3 there is at most $+1/2$ excess. If there is positive excess and edges both ways between the singleton s and the cycle layer L_k, then s is in a directed triangle with L_k in G, which is impossible by the layering strategy.

6. *Singleton above Cycle:* By Lemma 3 there is at most $+1/2$ excess. By Lemma 4, for $j = 0$ positive excess requires a two-way edge.

7. *Singleton above Path:* By Lemma 5, if the node s in layer L_k is special, then it is connected both ways with L_j. By the layering strategy, s cannot be connected both ways to S_{j-1}; so this edge is in F.

8. *Path above Singleton:* Suppose the node s in the singleton layer is adjacent with two consecutive nodes v and w in L_k. Node s cannot lie in a cycle outside S_{j-1}. Thus by Lemma 7, the only possibility is that the edges are (v, s) and (s, w), so that the removal of s disconnects v from w (by Lemma 1). But then there must be a path either from w to s or from s to v outside S_{j-1}, so that s is in a cycle outside S_{j-1}, a contradiction. □

5 Proof of Claim 3

Lemma 9. *Assume all edges are directed from layer L_j to layer L_k in G, $j < k$, and there is positive excess. Let y be the in-node of L_j and v the out-node of L_k.*
1) Nodes v and y are unique.
2) It holds that $j \geq 1$. Any path from L_k to S_{j-1} avoids L_j. In particular there is a path from v to S_{j-1} avoiding L_j and the rest of L_k.
3) Nodes v and y are adjacent. If L_j (L_k) is a cycle layer then v (y) is adjacent with the predecessor of y (successor of v).

Proof:
1) Node v is unique by definition if L_k is a singleton or path layer. If L_k is a cycle layer then it is unique by Lemma 7. Similarly for y.
2) Suppose there is a path from L_k to L_j in $G - S_{j-1}$. Then this path either contradicts the maximality of L_j, if that is a cycle layer, or it contradicts the fact that the node of L_j is not in a cycle outside S_{j-1}, if L_j is a singleton layer. There is a path from v to S_{j-1} avoiding the rest of L_k, since there is an edge from v to S_{k-1}.
3) This follows from Lemmas 4 and 7. Node v is up-special since more than half of the nodes of L_k are. All up-special nodes connect to y, and to its predecessor if L_j is a cycle layer. A similar argument holds looking from L_j to L_k. □

Lemma 10. *There is no triangle in H with all one-way edges.*

Proof: Suppose there is a triangle in H using one-way edges among nodes A, B, C with indices a, b, c. If it is a directed triangle, say $A \to B \to C \to A$ with c the maximum, then the B to A path through C contradicts the above avoidance property.

So assume they form a transitive triple (A, B, C) with $a < c$. Let v denote the in-node of A, and y the out-node of C. Then there is an edge from v to y by the above lemma. Further, there must be an even-length path from v to y through B (by part 3 of above lemma). Thus the removal of B must disconnect A from C.

But this is impossible. For, if $b > c$, then A and C lie together in the strongly connected subgraph S_c. If $a < b < c$, then by the above lemma there is a path from C to S_{b-1} (and hence A) which avoids B. And if $b < a$, then by the above lemma, any path from C to S_{b-1} avoids B, as do the ones from S_{b-1} to A. $\square$

By Lemma 8, there are only two possibilities for two-way edges in $H - F$: odd path layer above singleton layer (PS); and singleton layer above cycle layer (SC).

Lemma 11. *There is no triangle in $H - F$ with a PS two-way edge.*

Proof: Assume that path layer L_k (of odd cardinality) is above singleton layer L_j and they are connected by a two-way edge in H. Then (by the proof of Lemma 8), the node s of L_j is connected to every alternate node in L_k; in particular to the in-node x and out-node y of L_k. Since the edges directed towards s must come before those directed away from s (by the lack of even cycles), (x, s) and (s, y) must be edges in G. Further, any path from y to S_{j-1}, or from S_{j-1} to x must avoid s, since s is not in a cycle outside S_{j-1}.

Suppose L_j and L_k are in a triangle with layer L in $H - F$. If L is also a singleton layer, with node t say, and connected to L_k by a two-way edge in H, then by the above observation L_k must lie in an odd cycle disjoint from s and t. But x and y are also connected by a path of length three through s and t, which yields an even cycle.

So we may assume that all edges between L and L_k go the same way; say towards L_k. Then there is a path from y directly to the subgraph S below all three layers. Also since there is a path from L to s in $G - S$, by Lemma 9 all edges are directed from L to s. This yields a transitive triple (v, s, y) where v is the in-node of L. By Lemma 9 (or otherwise) there is a path direct from S to v. This yields an even cycle. $\square$

Lemma 12. *There is no triangle in $H-F$ with an SC two-way edge.*

Proof: Assume singleton layer L_k, with node s, and cycle layer L_j $(k > j)$ are joined in $H-F$ by a two-way edge, and that this edge is in a triangle with layer L in $H-F$. Then we claim that all edges between L and $L_j \cup L_k$ go the same way. For if L_k is joined to L by a two-way (SC) edge in H, then L must be one-way with L_j, and then either the L_j to L or the L to L_j path through s in G contradicts Lemma 9. A similar contradiction results if L is two-way with L_j. Or if L_k and L_j are both one-way with L but in different directions. So we may assume that all edges are directed towards L.

The edge of F incident with L_k connects to a layer in the subgraph S below all three layers. (By definition the edge goes below L_j; by Lemma 9 it must go below L.) Also there is a path direct from the out-node y of L to S, and an edge from s to y. This means that (s, y) lies in an odd cycle disjoint from L_j. But there is also an even-length path from s to y using only L_j (by part 3 of Lemma 9). This yields a contradiction. □

We have established Claim 3 and hence the proof of the main theorem is complete.

6 Problems and remarks

There are many unsolved problems about even cycles in directed graphs, some of which we mention here:

1. A directed graph is said to be k-strongly-connected if for every pair of two nodes u and v there are at least k (node) disjoint paths joining u to v. P. Seymour found [12] a 2-strongly-connected directed graph on 7 nodes containing no even cycle. Is it true that every 3-strongly-connected directed graph contains an even cycle? This was recently proved by C. Thomassen [17] to be true.

2. Erdős and Pósa [6, 7] proved that every *undirected* graph contains either k disjoint cycles or contains $ck \log k$ nodes which must meet all cycles for some constant c. Recently, W. McCuaig [11] proved a conjecture of T. Gallai [9] by showing that every directed graph contains either two disjoint cycles or three nodes meeting all cycles. Does there exist [21] a number $f(k)$ for every integer k such that a directed graph contains either k disjoint cycles or a set of $f(k)$ nodes meeting all cycles?

3. Of course, the problem of deciding if a directed graph contains an even cycle or not remains open. In [16] C. Thomassen gives a polynomial algorithm for deciding if a planar directed graph contains an even cycle.

7 Acknowledgement

The authors wish to thank John Goldwasser, C. Q. Zhang, Hong-Jia Lai and Zoltan Furedi for very helpful discussions, and Paul Seymour for mentioning some of the research problems.

References

[1] R. A. Brualdi, Counting permutations with restricted positions: permanents of (0,1) matrices, *Linear Algebra Appl.,* 104 (1988) 173–183.

[2] R. A. Brualdi, Graphs and matrix and graphs, Proceedings of the 22nd Southeastern Conference on Combinatorics, Graph Theory and Computing (1991).

[3] R. A. Brualdi and B. L. Shader, Cutsets in bipartite graphs, *Linear Multilin. Algebra,* to appear.

[4] R. A. Brualdi and B. L. Shader, On sign-nonsingular matrices and the conversion of the permanent into the determinant, *Victor Klee Festschrift,* P. Gritzmann and B. Sturmfels eds., Amer. Math. Soc., Providence, to appear.

[5] J. C. Bermond and C. Thomassen, Cycles in digraphs—A survey, *J. Graph Theory,* 5 (1981) 1–43.

[6] P. Erdős and L. Pósa, On the maximal number of disjoint circuits in a graph, *Publ. Math. Debrecen,* 9 (1962) 3–12.

[7] P. Erdős and L. Pósa, On independent circuits contained in a graph, *Canad. J. Math.,* 17 (1965) 347–352.

[8] S. Friedland, Every 7-regular digraph contains an even cycle, *J. Combin. Theory Ser. B,* 46 (1989) 249–252.

[9] T. Gallai, Problem 6 in Theory of Graphs, *Proc. Colloq. Tihany 1966,* Academic Press, New York, 362.

[10] C. H. C. Little, A characterization of convertible (0,1)-matrices, *J. Combin. Theory Ser. B.,* 18 (1975) 187–208.

[11] W. McCuaig, preprint.

[12] P. Seymour, On the two-coloring of hypergraphs, *Quart. J. Math. Oxford,* 25 (1974) 303–312.

[13] P. Seymour and C. Thomassen, Characterization of even directed graphs, *J. Combin. Theory Ser. B,* 42 (1987) 36–45.

[14] C. Thomassen, Even cycles in directed graphs, *European J. Combin.,* 6 (1985) 85–89.

[15] C. Thomassen, Sign-nonsingular matrices and even cycles in directed graphs, *Linear Algebra Appl.,* 75 (1986) 27–41.

[16] C. Thomassen, The even cycle problem for planar directed graphs, to appear.

[17] C. Thomassen, The even cycles problem for directed graphs, to appear.

[18] P. Turán, Egy gráfelméletei szélsöékfeladatról, *Matem. Physikai Lapok,* 48 (1941) 436–452.

[19] P. Turán, On the theory of graphs, *Coll. Math.,* 3 (1954) 19–30.

[20] L. G. Valiant, The complexity of computing the permanent, *Theor. Comp. Sci.,* 8 (1979) 189–201.

[21] V. V. Vazirani and M. Yannakakis, Pfaffian orientations, 0-1 permanents and even cycles in directed graphs, *Discrete Appl. Math.,* 25 (1989) 179–190.

[22] D. H. Younger, Graphs with interlinked directed circuits, *Proc. Midwest Symp. on Circuit Theory,* 2 (1973) XVI2.1–XVI2.7.

---

wdg LaTeX source: "/NO.EVEN.CYCLE/final.tex"