On the Number of Arcs in Primitive Digraphs with Large Exponents

Jian Shen*
Department of Mathematics
University of Illinois
Urbana, Illinois 61801
jshen@math.uiuc.edu

Cynthia J. Wyels
Mathematics/ Physics
California Lutheran University
Thousand Oaks, CA 91360
wyels@clunet.edu

June, 2000

Abstract

A digraph G is called primitive if for some positive integer k, there is a walk of length exactly k from each vertex u to each vertex v (possibly u again). If G is primitive, the smallest such k is called the exponent of G, denoted by $\exp(G)$. For any real number r, $0 < r < 1$, let $f(n, r)$ be the maximum number of arcs in a primitive digraph with n vertices having exponent greater than or equal to r^2n^2. We show that $f(n, r)/n^2$ is asymptotically $(1 - r)^2/3$ whenever $r \geq \sqrt{2}/2$.

Let $G = (V, E)$ denote a digraph on n vertices. Loops are permitted but no multiple arcs. A $u \to v$ *walk* in G is a sequence of vertices $u, u_1, \ldots, u_p = v$ and a sequence of arcs $(u, u_1), (u_1, u_2), \ldots, (u_{p-1}, v)$ where the vertices and the arcs are not necessarily distinct. A *path* is a walk with distinct vertices. A *cycle* is a closed $u \to v$ walk with distinct vertices except for $u = v$. The *length* of a walk is the number of arcs in the walk.

A digraph G is called *primitive* if, for some positive integer k, there is a walk of length exactly k from each vertex u to each vertex v (possibly u again). If G is primitive, the smallest such k is called the *exponent* of G, denoted by $\exp(G)$. It is well-known that (for example see [2]) G is primitive if and only if G is strongly connected and the greatest common divisor of all the cycle lengths of G is 1.

A 1950 paper of Wielandt [8] gives the maximum possible exponent of a primitive digraph on n vertices as $W_n = (n-1)^2 + 1$, and exhibits a primitive digraph that achieves this bound as its exponent. Further work by Dulmage and Mendelsohn [3], Lewin and

*Supported by an NSERC Postdoctoral Fellowship.

Vitek [5], Shao [6] and Zhang [9] completely characterizes those exponents which are actually achievable. In particular, Lewin and Vitek's, Shao's and Zhang's papers show that all exponents not exceeding $\left\lfloor \frac{1}{2}W_n \right\rfloor + 1$ are achievable (with the single exception of 48 when $n = 11$). A key step in this characterization is the following lemma.

Lemma 1 (Lewin and Vitek [5]) *Let G be a primitive digraph with more than two different cycle lengths. Then*

$$\exp(G) \leq \left\lfloor \frac{1}{2}W_n \right\rfloor + 1 = \left\lfloor \frac{1}{2}(n^2 - 2n + 4) \right\rfloor.$$

From Lemma 1, a connection between primitive digraphs having small exponents and primitive digraphs with many arcs may be inferred. (Adding arcs to a primitive digraph does not increase its exponent.) Accordingly, it is of interest to ask how many arcs a primitive digraph on n vertices may contain and yet have an exponent exceeding $\left\lfloor \frac{1}{2}W_n \right\rfloor + 1$. We consider a general problem in terms of $f(n, r)$.

Definition. For any real number r, $0 < r < 1$, let $f(n, r)$ be the maximum number of arcs in a primitive digraph G on n vertices for which $\exp(G) \geq r^2n^2$.

In this paper, we find the asymptotic value for $f(n, r)/n^2$ (n^2 being the maximum number of arcs in a digraph on n vertices) when $r \geq \sqrt{2}/2$. In particular, when $r = \sqrt{2}/2$, Theorem 1 below shows that the maximum number of arcs in a primitive digraph on n vertices with exponent exceeding $\left\lfloor \frac{1}{2}W_n \right\rfloor + 1$ is asymptotically $(\sqrt{2} - 1)^2n^2/6 + O(n) \approx 0.0286n^2 + O(n)$.

Theorem 1 *For any real number $r \geq \sqrt{2}/2$,*

$$\lim_{n \to +\infty} \frac{f(n, r)}{n^2} = \frac{1}{3}(1 - r)^2.$$

Proof. We assume n is sufficiently large. To simplify the notation and calculation, we avoid the use of floor and ceiling signs since they are not crucial when we are interested in the asymptotic value of $f(n, r)/n^2$. We will prove

$$(1 - r)^2n^2 + 7n - 4rn + 1 \leq 3f(n, r) \leq \left(n - \sqrt{r^2n^2 - n}\right)^2 + 6n - 3\sqrt{r^2n^2 - n} - \frac{3}{4}.$$

The lower bound for $f(n, r)$ is based on the following construction. We construct $G_1 = (V_1, E_1)$ as follows: $V_1 = \{i : 1 \leq i \leq n - 3k + 3\}$ and $E_1 = \{(i, i + 1) : 1 \leq i \leq n - 3k + 2\} \cup \{(n - 3k + 3, 1), (n - 3k + 3, 2), (1, 3)\}$. It is known [3, Theorem 6] that G_1 is primitive with $\exp(G_1) = (n - 3k + 2)^2$. Now we "blow-up" G_1 by replacing each vertex i, $i = 1, 2, 3$, by k copies of i (preserving the adjacency relation). Let $G = (V, E)$

denote the resulting digraph. Then G contains n vertices and $3k^2 + n - 1$ arcs. Also $\exp(G) = \exp(G_1) = (n - 3k + 2)^2$ since adding copies of vertices does not change the primitivity and the exponent of a digraph. Letting $k = (n - rn + 2)/3$ implies that $\exp(G) = r^2n^2$ and $|E| = (n-rn+2)^2/3+n-1 = (1-r)^2n^2/3+7n/3-4rn/3+1/3$. This proves the lower bound for $f(n, r)$.

To prove the upper bound for $f(n, r)$, now we suppose G is a primitive digraph on n vertices with $\exp(G) \geq r^2n^2 \geq n^2/2$. By Lemma 1, G contains cycles with exactly two different lengths, say p and q with $p \leq q$. Let C_q be a cycle of length q in G.

Claim 1: $p + q \geq 2\sqrt{r^2n^2 - n} + 2$.
By [5, Lemma 3.2 and Theorem 3.3], we have

$$\exp(G) \leq \begin{cases} 2n + (p-1)(q-2) - 2 & \text{if } p + q \leq n, \\ n + p(q-2) & \text{if } p + q > n. \end{cases}$$

If $p + q \leq n$, then $n^2/2 \leq \exp(G) \leq 2n + (p-1)(q-2) - 2 \leq 2n - 2 + (p+q-3)^2/4 \leq 2n - 2 + (n-3)^2/4$, a contradiction. Thus $p + q > n$ and $r^2n^2 \leq \exp(G) \leq n + p(q-2) \leq n + (p+q-2)^2/4$, which implies Claim 1.

Let $G \setminus C_q$ be the subdigraph of G induced by $V(G) \setminus V(C_q)$.

Claim 2: $G \setminus C_q$ is acyclic; that is, it contains no cycle.
Suppose, on the contrary, that $G \setminus C_q$ contains a cycle C of length p or q. Then C_q and C are vertex-disjoint, and so $p + q \leq |C| + |C_q| \leq n$, a contradiction to Claim 1.

Claim 3: Each vertex in $G \setminus C_q$ is adjacent to or from at most 3 vertices on C_q.
Otherwise suppose, on the contrary, that some vertex u in $G \setminus C_q$ is adjacent to s vertices on C_q and is adjacent from t vertices on C_q with $s + t \geq 4$. If both $s \geq 1$ and $t \geq 1$, then the subdigraph induced by $V(C_q) \cup \{u\}$ contains cycles with more than two different lengths, and then so does G, which is a contradiction. If $s = 0$ or $t = 0$, G still contains cycles with more than two different lengths since there exists a shortest path from u to the cycle C_q and a shortest path from the cycle C_q to u. This proves Claim 3.

Claim 4: The cycle C_q has at most $q - p + 1$ chords.
Otherwise the subdigraph induced by $V(C_q)$ contains cycles of length other than p or q.

Claim 5: $G \setminus C_q$ contains as a subdigraph no tournament on 4 vertices.
Suppose, on the contrary, that $G \setminus C_q$ contains as a subdigraph a tournament on 4 vertices u_i, $1 \leq i \leq 4$. By Claim 2, this tournament is a transitive tournament [4, Corollary 5a]; that is, $(u_i, u_j) \in E(G)$ whenever $1 \leq i < j \leq 4$. Let P_1 and P_2 denote a shortest path from C_q to u_1 and a shortest path from u_4 to C_q, respectively. By Claim 2, P_1 does not contain u_2, u_3 or u_4, and P_2 does not contain u_1, u_2 or u_3. Let C_l be the cycle of length l formed by the path P_1, the arc (u_1, u_4), the path P_2 and the path along the cycle C_q. Then two more cycles, of lengths $l + 1$ and $l + 2$ respectively, can be obtained by replacing the

arc (u_1, u_4) by the path $u_1 \to u_2 \to u_4$ and by the path $u_1 \to u_2 \to u_3 \to u_4$, respectively. This contradicts that G contains cycles with exactly 2 different lengths.

Claim 6: $G \setminus C_q$ contains at most $(n-q)^2/3$ arcs.
By Claim 2, the digraph $G \setminus C_q$ contains no loops or digons. We may consider a digraph without loops or digons as a undirected graph by replacing each arc by a undirected edge. By Claim 5, the corresponding undirected graph of the digraph $G \setminus C_q$ contains as a subgraph no K_4, the complete graph on 4 vertices. Therefore Claim 6 follows from Turan's theorem [7, Theorem 4.1].

By Claims 3, 4, 6, the digraph G contains at most

$$(n-q)^2/3 + 3(n-q) + q + (q-p+1) = (n-q)^2/3 + 3n - (p+q) + 1$$

arcs. By Claim 1, $p + q \geq 2\sqrt{r^2n^2 - n} + 2$, which gives $q \geq (p+q+1)/2$, so $q \geq \sqrt{r^2n^2 - n} + 1.5$. So G contains at most

$$\frac{\left(n - \sqrt{r^2n^2 - n} - 1.5\right)^2}{3} + 3n - 2\sqrt{r^2n^2 - n} - 1 = \frac{\left(n - \sqrt{r^2n^2 - n}\right)^2}{3} + 2n - \sqrt{r^2n^2 - n} - \frac{1}{4}$$

arcs. This proves the upper bound for $f(n,r)$. Therefore Theorem 1 holds. □

Now we raise the following problem: Find the asymptotical value for $f(n,r)/n^2$ when $0 < r < \sqrt{2}/2$. We obtain in Theorem 2 a lower bound for the asymptotical value by constructing a set of primitive digraphs which seem to have the most number of arcs. We need some definitions before we state Theorem 2.

The *local exponent of* G *from vertex* u *to vertex* v, denoted $\exp(G\!:u,v)$, is the smallest integer k such that there is a walk of length l from u to v for all $l \geq k$. It is known [2, Lemma 3.5.1] that $\exp(G) = \max_{u,v \in V} \exp(G\!:u,v)$.

Let $a_1 < a_2 < \cdots < a_p$ be positive integers. The Frobenius-Schur index, $\phi(a_1, a_2, \ldots, a_p)$, is the smallest integer such that the equation $x_1a_1 + x_2a_2 + \cdots + x_pa_p = l$ has a solution in non-negative integers $x_1, x_2, \ldots, x_p$ for all $l \geq \phi(a_1, \ldots, a_p)$. The following result was due to Brauer in 1942.

Lemma 2 (Brauer [1]) *Let m be a positive integer. Then,*

$$\phi(m, m+1, \ldots, m+j-1) = m \cdot \left\lfloor \frac{m+j-3}{j-1} \right\rfloor.$$

Now we are ready to prove a lower bound for the asymptotical value of $f(n,r)/n^2$.

Theorem 2 *For any real number r, $0 < r < 1$,*

$$\lim_{n\to+\infty} \frac{f(n,r)}{n^2} \geq \max_{t:2\leq t<\frac{1}{r^2}+1} \frac{t}{2(t+1)}\left(1-r\sqrt{t-1}\right)^2.$$

Proof. Again we assume n is sufficiently large, and we avoid the use of floor and ceiling signs since they are not crucial when we are interested in the asymptotic value of $f(n,r)/n^2$.

We construct as follows a primitive digraph $G_t = (V_t, E_t)$ for each t, $2 \leq t < \frac{1}{r^2}+1$: Let $V_t = \{i : 1 \leq i \leq n-(t+1)(k-1)\}$ and $E_t = \{(n-(t+1)(k-1), 1)\} \cup \{(i, i+1) : 1 \leq i \leq n-(t+1)(k-1)-1\} \cup \{(i,j) : 1 \leq i < j \leq t+1\}$, where k will be decided later. It is easy to see that

$$\begin{aligned}
\exp(G_t\!:\!t,2) &= \phi\left(\{i : n-(t+1)k+2 \leq i \leq n-(t+1)(k-1)\}\right) + n-(t+1)k+3 \\
&= \tfrac{n-(t+1)(k-1)-2}{t-1}(n-(t+1)k+2) + n-(t+1)k+3 \\
&= \tfrac{(n-(t+1)(k-1)-1)^2-1}{t-1} - t + 4 \\
&> \tfrac{(n-(t+1)k)^2}{t-1},
\end{aligned}$$

where the second equation follows from Lemma 2. Now we "blow-up" G_k by replacing each vertex i, $1 \leq i \leq t+1$, by k copies of i (preserving the adjacency relation). Let $\bar{G}_t = (\bar{V}_t, \bar{E}_t)$ denote the resulting digraph. Then $\bar{G}_t$ contains n vertices and $\binom{t+1}{2}k^2+n-(t-1)k-1$ arcs. Also $\exp(\bar{G}_t) = \exp(G_t)$ since adding copies of vertices does not change the primitivity and the exponent of a digraph. Letting $k = (1-r\sqrt{t-1})n/(t+1)$ implies that $\exp(\bar{G}_t) = \exp(G_t) \geq \exp(G_t\!:\!t,2) > r^2n^2$ and that

$$\begin{aligned}
f(n,r) &\geq |\bar{E}_t| = \tbinom{t+1}{2}\tfrac{(1-r\sqrt{t-1})^2n^2}{(t+1)^2} + n - (t-1)\tfrac{(1-r\sqrt{t-1})n}{(t+1)} - 1 \\
&= \tfrac{t}{2(t+1)}\left(1-r\sqrt{t-1}\right)^2 n^2 + \tfrac{2}{t+1}n + \tfrac{t-1}{t+1}r\sqrt{t-1}n - 1,
\end{aligned}$$

from which Theorem 2 follows. □

To conclude the paper, we raise the following conjecture:

Conjecture 1 *For any real number r, $0 < r < 1$,*

$$\lim_{n\to+\infty} \frac{f(n,r)}{n^2} = \max_{t:2\leq t<\frac{1}{r^2}+1} \frac{t}{2(t+1)}\left(1-r\sqrt{t-1}\right)^2.$$

It is interesting to comment here that

$$\lim_{n\to+\infty}\lim_{r\to0} \frac{f(n,r)}{n^2} \neq \lim_{r\to0}\lim_{n\to+\infty} \frac{f(n,r)}{n^2}.$$

Indeed, $\lim_{r\to 0} f(n,r)/n^2 = 1$ for all $n \geq 1$. This follows from the following fact: If $\exp(G) = 1$, then G is a complete digraph, i.e. $(u,v) \in E$ for all vertices $u, v \in V$ (possibly $u = v$). Since $\exp(G) = r^2n^2 = 1$ means $r = 1/n$, we have $f(n, 1/n) = n^2$ for all $n \geq 1$. Therefore $\lim_{n\to+\infty} \lim_{r\to 0} f(n,r)/n^2 = 1$. We show the next theorem in support of Conjecture 1.

Theorem 3

$$\lim_{r\to 0} \lim_{n\to+\infty} \frac{f(n,r)}{n^2} = \frac{1}{2}.$$

Proof. Let $t = \lfloor 1/r \rfloor + 1$. Then $2 \leq t < 1/r^2 + 1$. By Theorem 2,

$$\begin{aligned} \lim_{n\to+\infty} \tfrac{f(n,r)}{n^2} &\geq \tfrac{t}{2(t+1)} \left(1 - r\sqrt{t-1}\right)^2 \\ &\geq \tfrac{\lfloor 1/r \rfloor + 1}{2(\lfloor 1/r \rfloor + 2)} (1 - \sqrt{r})^2, \end{aligned}$$

from which we have

$$\lim_{r\to 0} \lim_{n\to+\infty} \frac{f(n,r)}{n^2} \geq \frac{1}{2}.$$

Now we suppose, contrary to the theorem, that the inequality above is strict. Then there exist some r, $0 < r < 1$, and a sufficiently large n (independent of r) such that $f(n,r) > n^2/2$. Thus there exists a digraph $G = (V, E)$ with n vertices such that $\exp(G) \geq r^2n^2$ and $|E| > n^2/2$. Then G contains either a loop or a digon as G contains more than $n^2/2$ arcs. By the Dulmage-Mendelsohn bound in [3], we have $r^2n^2 \leq \exp(G) \leq n + 2(n-2)$, a contradiction since n (independent of r) is sufficiently large. This completes the proof of Theorem 3. $\square$

Acknowledgment. Part of this research was carried out when both authors met at the Mini-conference on Combinatorial Matrix Theory held at Utah State University in May, 1999. We thank Professor LeRoy Beasley and the mathematics department of USU for organizing the conference.

References

[1] A. Brauer, On a problem of partitions, *Amer. J. Math.* **64** (1942), 299-312.

[2] R. A. Brualdi and H. Ryser, *Combinatorial Matrix Theory, Encyclopedia of Mathematics and its Applications* 39, Cambridge University Press, Cambridge, 1991.

[3] A. L. Dulmage and N. S. Mendelsohn, Gaps in the exponent set of primitive matrices, *Illinois J. Math.* **8** (1964) 642–656.

[4] F. Harary and L. Moser, The theory of round robin tournaments, *Amer. Math. Monthly* **73** (1966), 231-246.

[5] M. Lewin and Y. Vitek, A system of gaps in the exponent set of primitive matrices, *Illinois J. Math.* **25** (1981), 87–98.

[6] J. Y. Shao, On a conjecture about the exponent set of primitive matrices, *Linear Algebra Appl.* **65** (1985), 91–123.

[7] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, Cambridge University Press, Cambridge, 1992.

[8] H. Wielandt, Unzerlegbare, nicht negative Matrizen, *Math. Zeit.* **52** (1950), 642–645.

[9] K. M. Zhang, On Lewin and Vitek's conjecture about the exponent set of primitive matrices, *Linear Algebra Appl.* **96** (1987), 101–108.