Title
Paths of given length in tournaments

Journal
Combinatorial Theory, 3(2)

ISSN
2766-1334

Authors
Sah, Ashwin
Sawhney, Mehtaab
Zhao, Yufei

Publication Date
2023

DOI
10.5070/C63261981

PATHS OF GIVEN LENGTH IN TOURNAMENTS

Ashwin Sah*[1], Mehtaab Sawhney†[2], and Yufei Zhao‡[3]

[1,2,3]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, U.S.A.*
asah@mit.edu , msawhney@mit.edu , yufeiz@mit.edu

Submitted: Dec 14, 2020; Accepted: Mar 13, 2023; Published: Sep, 15 2023

Abstract. We prove that every n-vertex tournament has at most $n\left(\frac{n-1}{2}\right)^k$ walks of length k.
Keywords. Paths, tournaments
Mathematics Subject Classifications. 05C38, 05D99

We determine the maximum density of directed k-edge paths in an n-vertex tournament. Our focus is on the case of fixed k and large n. The expected number of directed k-edge paths in a uniform random n-vertex tournament is $n(n-1)\cdots(n-k)/2^k = (1+o(1))n(n/2)^k$. In this short note we show that one cannot do better, thereby confirming an unpublished conjecture of Jacob Fox, Hao Huang, and Choongbum Lee. The *length* of a path or walk refers to its number of edges.

Theorem 1. *Every n-vertex tournament has at most $n\left(\frac{n-1}{2}\right)^k$ walks of length k.*

Every regular tournament (with odd n) has exactly $n\left(\frac{n-1}{2}\right)^k$ walks of length k, thereby attaining the upper bound in the theorem. On the other hand, the transitive tournament minimizes the number of k-edge paths (or walks) among n-vertex tournaments. Indeed, a folklore result (with an easy induction proof) says that every tournament contains a directed Hamilton path. So every $(k+1)$-vertex subset contains a path of length k. Hence every n-vertex tournament contains at least $\binom{n}{k+1}$ paths of length k, with equality for a transitive tournament.

Figure 1 provides a "proof by picture" of Theorem 1. A more detailed proof is given later. A different proof, using entropy, by Dingding Dong and Tomasz Ślusarczyk, is given in the appendix.

Let us mention some related problems and results. The most famous open problem with this theme is Sidorenko's conjecture [ES83, Sid93], which says that for a fixed bipartite graph H,

*Supported by NSF Graduate Research Fellowship Program DGE-1745302.
†Supported by NSF Graduate Research Fellowship Program DGE-1745302.
‡Supported by NSF Award DMS-1764176, the MIT Solomon Buchsbaum Fund, and a Sloan Research Fellowship.

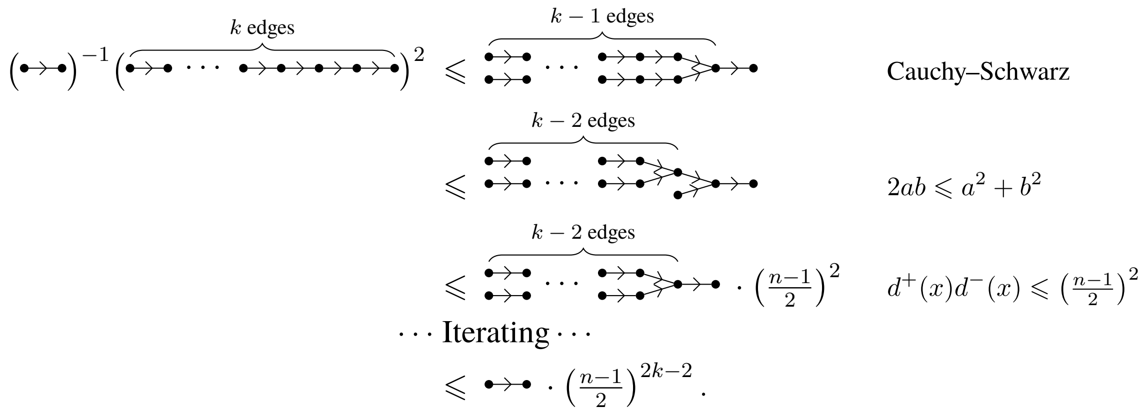

Figure 1: A "proof by picture" of Theorem 1.

among graphs of a given density, quasirandom graphs minimize H-density. For recent progress on Sidorenko's conjecture see [CKLL18, CL21].

Zhao and Zhou [ZZ20] determined all directed graphs that have constant density in all tournaments; they are all disjoint unions of trees that are each constructed in a recursive manner, as conjectured by Fox, Huang, and Lee. As discussed at the end of [ZZ20], it would be interesting to characterize directed graphs H where is the H-density in tournaments maximized by the quasirandom tournament (such H is called *negative*), and likewise when "maximized" is replaced by "minimized" (such H is called *positive*). Our main result here implies that all directed paths are negative. It would be interesting to see what happens for other edge-orientations of a path. Starting with a negative (resp. positive) digraph, one can apply the same recursive construction as in [ZZ20] to produce additional negative (resp. positive) digraphs, namely by taking two disjoint copies of the digraph and adding a single edge joining a pair of twin vertices.

The problem of maximizing the number of directed k-cycles in a tournament is also interesting and not completely understood. Recently, Grzesik, Král', Lovász, and Volec [GKLV23] showed that quasirandom tournaments maximize the number of directed k-cycles whenever k is not divisible by 4. On the other hand, when k is divisible by 4, quasirandom tournaments do not maximize the density of directed k-cycles. The maximum directed k-cycle density is known for $k = 4$ [BH65, Col64] and $k = 8$ [GKLV23] but open for all larger multiples of 4. See [GKLV23] for discussion.

A related problem is determining the maximum number $P(n)$ of Hamilton paths in a tournament (the problem for Hamilton cycles is related). By considering the expected number of Hamilton paths in a random tournament, one has $P(n) \geqslant n!/2^{n-1}$. This result, due to Szele [Sze43], is considered the first application of the probabilistic method. This lower bound has been improved by a constant factor [AAR01, Wor]. Alon [Alo90] proved a matching upper bound of the form $P(n) \leqslant n^{O(1)}n!/2^{n-1}$ (also see [FK05] for a later improvement).

Proof of Theorem 1. We may assume that $k \geqslant 1$. Let $f(x, y) = 1$ if (x, y) is a directed edge in the tournament, and 0 otherwise. Let $g_t(x)$ denote the number walks of length t ending at x.

Let $d^+(x)$ and $d^-(x)$ denote the out-degree and the in-degree of x, respectively. We have, for each $t \geqslant 1$,

$$g_t(y) = \sum_x g_{t-1}(x) f(x, y).$$

Define

$$A_t \coloneqq \sum_y g_t(y)^2 d^+(y). = \sum_{y,z} g_t(y)^2 f(y, z).$$

We have, for each $t \geqslant 1$,

$$\begin{aligned}
A_t &= \sum_{x,x',y} g_{k-1}(x) f(x,y) g_{k-1}(x') f(x', y) d^+(y) \\
&\leqslant \sum_{x,x',y} \left(\frac{g_{k-1}(x)^2 + g_{k-1}(x')^2}{2} \right) f(x,y) f(x',y) d^+(y) \\
&= \sum_{x,x',y} g_{k-1}(x)^2 f(x,y) f(x',y) d^+(y) \\
&= \sum_{x,y} g_{k-1}(x)^2 f(x,y) d^-(y) d^+(y) \\
&\leqslant \left(\frac{n-1}{2} \right)^2 \sum_{x,y} g_{k-1}(x)^2 f(x,y) \qquad \text{[since } d^-(y)d^+(y) \leqslant \left(\tfrac{n-1}{2}\right)^2\text{]} \\
&= \left(\frac{n-1}{2} \right)^2 A_{t-1}.
\end{aligned}$$

So, for all $t \geqslant 0$,

$$A_t \leqslant A_0 \left(\frac{n-1}{2} \right)^{2t} \leqslant n \left(\frac{n-1}{2} \right)^{2t+1}.$$

Let W_k be the number of walks of length k. Applying the Cauchy–Schwarz inequality,

$$W_k = \sum_y g_{k-1}(y) d^+(y) \leqslant \sqrt{\sum_y g_{k-1}(x)^2 d^+(y)} \sqrt{\sum_y d^+(y)} \leqslant \sqrt{A_{k-1} A_0} \leqslant n \left(\frac{n-1}{2} \right)^k . \square$$

The above proof also gives the following stability result.

Theorem 2 (Stability)**.** *For $k \geqslant 2$, an n-vertex tournament satisfying*

$$\sum_x \left| d^+(x) - \frac{n-1}{2} \right| \geqslant \varepsilon \binom{n}{2}$$

has at most $(1 - \frac{\varepsilon^2}{2}) n \left(\frac{n-1}{2}\right)^k$ walks of length k.

Note that by symmetry, we can replace d^+ by d^- in the hypothesis of Theorem 2.

Proof. We use the notation from the earlier proof. We have

$$\begin{aligned}
W_2 = \sum_x d^+(x)d^-(x) \leqslant \sum_x d^+(x)(n-1-d^+(x)) \\
&= \sum_x \left(\left(\frac{n-1}{2}\right)^2 - \left(\frac{n-1}{2} - d^+(x)\right)^2\right) \\
&\leqslant n\left(\frac{n-1}{2}\right)^2 - \frac{1}{n}\left(\sum_x \left|\frac{n-1}{2} - d^+(x)\right|\right)^2 \\
&\leqslant (1-\varepsilon^2)n\left(\frac{n-1}{2}\right)^2.
\end{aligned}$$

From the proof of Theorem 1, we have

$$W_k^2 \leqslant A_{k-1}A_0 \leqslant \left(\frac{n-1}{2}\right)^{2(k-2)} A_1 A_0.$$

Using $A_0 \leqslant n(n-1)/2$ and $A_1 = \sum_x d^-(x)^2 d^+(x)$, we obtain

$$W_k^2 \leqslant n\left(\frac{n-1}{2}\right)^{2k-3} \sum_x d^-(x)^2 d^+(x).$$

In the proof of Theorem 1, we defined $g_k(x)$ to be the number of k-edge walks ending at x. By running the same proof for the number of k-edge walks starting at x, we deduce

$$W_k^2 \leqslant n\left(\frac{n-1}{2}\right)^{2k-3} \sum_x d^-(x)d^+(x)^2.$$

Taking the average of the two bounds, we obtain

$$\begin{aligned}
W_k^2 &\leqslant n\left(\frac{n-1}{2}\right)^{2k-3} \sum_x d^-(x)d^+(x)\left(\frac{d^-(x)+d^+(x)}{2}\right) \\
&\leqslant n\left(\frac{n-1}{2}\right)^{2k-2} \sum_x d^-(x)d^+(x) \\
&\leqslant (1-\varepsilon^2)n^2\left(\frac{n-1}{2}\right)^{2k} \leqslant \left(\left(1-\frac{\varepsilon^2}{2}\right)n\left(\frac{n-1}{2}\right)^k\right)^2.
\end{aligned}$$

□

References

[AAR01] Ilan Adler, Noga Alon, and Sheldon M. Ross. On the maximum number of Hamiltonian paths in tournaments. *Random Structures Algorithms*, 18(3):291–296, 2001. doi:10.1002/rsa.1010.

[Alo90] Noga Alon. The maximum number of Hamiltonian paths in tournaments. *Combinatorica*, 10(4):319–324, 1990. doi:10.1007/BF02128667.

[BH65] Lowell W. Beineke and Frank Harary. The maximum number of strongly connected subtournaments. *Canad. Math. Bull.*, 8:491–498, 1965. doi:10.4153/CMB-1965-035-x.

[CKLL18] David Conlon, Jeong Han Kim, Choongbum Lee, and Joonkyung Lee. Some advances on Sidorenko's conjecture. *J. Lond. Math. Soc. (2)*, 98(3):593–608, 2018. doi:10.1112/jlms.12142.

[CL21] David Conlon and Joonkyung Lee. Sidorenko's conjecture for blow-ups. *Discrete Anal.*, pages Paper No. 2, 13, 2021. doi:10.19086/da.21472.

[Col64] Umberto Colombo. Sui circuiti nei grafi completi. *Boll. Un. Mat. Ital. (3)*, 19:153–170, 1964.

[ES83] Paul Erdős and Miklós Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983. doi:10.1007/BF02579292.

[FK05] Ehud Friedgut and Jeff Kahn. On the number of Hamiltonian cycles in a tournament. *Combin. Probab. Comput.*, 14(5-6):769–781, 2005. doi:10.1017/S0963548305006863.

[GKLV23] Andrzej Grzesik, Daniel Král', László M. Lovász, and Jan Volec. Cycles of a given length in tournaments. *J. Combin. Theory Ser. B*, 158(part 1):117–145, 2023. doi:10.1016/j.jctb.2022.07.007.

[Sid93] Alexander Sidorenko. A correlation inequality for bipartite graphs. *Graphs Combin.*, 9(2):201–204, 1993. doi:10.1007/BF02988307.

[Sze43] Tibor Szele. Kombinatorische Untersuchungen über den gerichteten vollständigen Graphen. *Mat. Fiz. Lapok*, 50:223–256, 1943.

[Wor] N. C. Wormald. Tournaments with many Hamilton cycles. manuscript https://users.monash.edu.au/~nwormald/papers/hamtourn.pdf.

[ZZ20] Yufei Zhao and Yunkun Zhou. Impartial digraphs. *Combinatorica*, 40(6):875–896, 2020. doi:10.1007/s00493-020-4280-0.

A. An entropy proof by Dingding Dong and Tomasz Ślusarczyk

Here is another proof of Theorem 1 using entropy. Given a discrete random variable X taking values in Ω, its entropy is defined as

$$H(X) = -\sum_{x\in\Omega} \mathbb{P}(X=x)\log\mathbb{P}(X=x).$$

We have the uniform bound

$$H(X) \leqslant \log|\Omega|.$$

The chain rule says that if X and Y are jointly distributed random variables, then

$$H(X,Y) = H(X) + H(Y|X),$$

where the conditional entropy $H(Y|X)$ is defined as

$$H(Y|X) = \sum_{x\in\Omega} \mathbb{P}(X = x) H(Y|X = x).$$

Here $H(Y|X = x)$ is the entropy of the conditional distribution of Y given $X = x$.

Entropy proof of Theorem 1. Consider a random walk $X_1, \dots, X_{k+1}$ chosen uniformly from the set of all W_k walks of length k in the given tournament. This random walk is Markovian in the sense that the distribution of $(X_i, \dots, X_{k+1})$ conditional on $(X_1, \dots, X_i)$ is the same as the distribution of $(X_i, \dots, X_{k+1})$ conditional on X_i. Indeed, this conditional distribution is uniform over all walks $(X_i, \dots, X_{k+1})$ with a given starting vertex X_i. In particular, $H(X_j|X_{j-1}, \dots, X_1) = H(X_j|X_{j-1})$.

Applying the chain rule, we have

$$\begin{aligned}\log W_k = H(X_1, \dots, X_{k+1}) &= H(X_1, X_2) + \sum_{j=2}^{k} H(X_{j+1}|X_1, \dots, X_j) \\ &= H(X_1, X_2) + \sum_{j=2}^{k} H(X_{j+1}|X_j).\end{aligned}$$

Likewise,

$$H(X_1, \dots, X_{k+1}) = H(X_{k+1}, X_k) + \sum_{j=2}^{k} H(X_{j-1}|X_j).$$

Taking the average of the two bounds, we obtain

$$H(X_1, \dots, X_{k+1}) = \frac{H(X_1, X_2) + H(X_{k+1}, X_k)}{2} + \frac{1}{2}\sum_{j=2}^{k}(H(X_{j-1}|X_j) + H(X_{j+1}|X_j)).$$

For each $2 \leqslant j \leqslant k$ and vertex x, by the uniform bound, $H(X_{j-1}|X_j = x) \leqslant \log d^-(x)$ and $H(X_{j+1}|X_j = x) \leqslant \log d^+(x)$. Also, $d^-(x)d^+(x) \leqslant (n-1)^2/4$. Thus

$$H(X_{j-1}|X_j = x) + H(X_{j+1}|X_j = x) \leqslant \log d^-(x) + \log d^+(x) \leqslant 2\log\left(\frac{n-1}{2}\right).$$

Thus

$$H(X_{j-1}|X_j) + H(X_{j+1}|X_j) \leqslant 2\log\left(\frac{n-1}{2}\right).$$

Also $H(X_j, X_{j+1}) \leqslant \log\binom{n}{2}$ by the uniform bound. Therefore

$$\log W_k = H(X_1, \dots, X_{k+1}) \leqslant \log\binom{n}{2} + (k-1)\log\left(\frac{n-1}{2}\right) = \log\left(n\left(\frac{n-1}{2}\right)^k\right). \quad \square$$