Almost given length cycles in digraphs

Raphael Yuster *

Department of Mathematics, University of Haifa, Haifa 31905, Israel

Abstract. A digraph is called k-cyclic if it cannot be made acyclic by removing less than k arcs. It is proved that for every $\epsilon > 0$ there are constants K and δ so that for every $d \in (0, \delta n)$, every ϵn^2-cyclic digraph with n vertices contains a directed cycle whose length is between d and $d + K$. A more general result of the same form is obtained for blow-ups of directed cycles.

Key words. directed graph, directed cycle, regularity lemma.

1. Introduction

All graphs and directed graphs (digraphs) considered here are finite and simple. For standard terminology on graphs and digraphs the reader is referred to [3]. A *feedback arc set* of a digraph is a set of arcs whose removal makes the digraph acyclic. A digraph is called *k-cyclic* if it does not have a feedback arc set whose size is less than k. Thus, acyclic digraphs are 0-cyclic and the directed triangle is 1-cyclic. It is not difficult to see that a random n-vertex tournament is $\frac{1}{4}n^2(1 - o(1))$-cyclic, almost surely (see, e.g., [5]).

An *r-blowup* of a directed cycle is obtained by replacing each vertex with an independent set of size r, and each arc (u, v) with the r^2 arcs connecting the vertices blown up from u to the vertices blown up from v. Let C_p^r denote the r-blowup of a p-cycle. Our main result is the following.

Theorem 1. *For every $\epsilon > 0$, and every positive integer r, there are constants K and δ so that for every $n > K$, and for every $d \in (0, \delta n)$, every n-vertex digraph that is ϵn^2-cyclic has a C_p^r where $d \le p \le d + K$.*

In the case $r = 1$ a simpler statement immediately follows.

Corollary 1. *For every $\epsilon > 0$, there are constants K and δ so that for every $d \in (0, \delta n)$, every n-vertex digraph that is ϵn^2-cyclic has a cycle whose length is between d and $d + K$.*

The proof of Theorem 1 is based on a version of Szemerédi's regularity lemma for directed graphs. There are some obvious bounds for the constants δ and K appearing in Theorem 1, already for the case $r = 1$. In Section 3 we will show that we must have $K = \Omega(\epsilon^{-1/2})$ and $\delta = O(\epsilon)$.

Theorem 1 can be used to prove the following theorem.

* e-mail: raphy@math.haifa.ac.il

Theorem 2. *For every $\epsilon > 0$ there are constants K and δ so that for every n-vertex digraph and for every $d \in (0, \delta n)$, the set of vertices can be partitioned to parts $V_0, V_1, \ldots, V_t$ so that V_0 induces a subgraph that can be made acyclic by removing ϵn^2 arcs, and V_i induces a hamiltonian digraph where $d \le |V_i| \le d + K$ for $i = 1, \ldots, t$.*

We call a digraph with pr vertices *r-hamiltonian* if it contains a C_p^r. In Theorem 2 we can replace the requirement that each V_i is hamiltonian with the stronger requirement of being r-hamiltonian (assuming, of course, that n is sufficiently large and that δ and K are also functions of r).

The rest of this paper is organized as follows. In Section 2 we introduce the regularity lemma for directed graphs which is the main tool in the proof of Theorem 1. In Section 3 we prove Theorem 1. Section 4 consists of the proof of Theorem 2. Section 5 contains some concluding remarks.

2. The regularity lemma for digraphs

An important tool used in the proof of Theorem 1 is the following version of Szemerédi's regularity lemma for directed graphs, that has been used implicitly in [4] and proved in [2]. The proof is a modified version of the proof of the standard regularity lemma given in [7]. We now give the definitions necessary to state this version of the regularity lemma.

Let $G = (V, E)$ be a digraph, and let A and B be two disjoint subsets of $V(G)$. If A and B are non-empty and $e(A, B)$ is the number of arcs from A to B, the *density of arcs from A to B* is

$$d(A, B) = \frac{e(A, B)}{|A||B|}.$$

For $\gamma > 0$ the pair (A, B) is called *γ-regular* if for every $X \subset A$ and $Y \subset B$ satisfying $|X| > \gamma|A|$ and $|Y| > \gamma|B|$ we have

$$|d(X, Y) - d(A, B)| < \gamma \qquad\qquad |d(Y, X) - d(B, A)| < \gamma.$$

An *equitable partition* of a set V is a partition of V into pairwise disjoint classes $V_1, \ldots, V_m$ whose sizes are as equal as possible. An equitable partition of the set of vertices V of a digraph G into the classes $V_1, \ldots, V_m$ is called *γ-regular* if $|V_i| \le \gamma|V|$ for every i and all but at most $\gamma\binom{m}{2}$ of the pairs (V_i, V_j) are γ-regular. The directed regularity lemma states the following:

Lemma 1. *For every $\gamma > 0$, there is an integer $M(\gamma) > 0$ such that every digraph G with $n > M$ vertices has a γ-regular partition of the vertex set into m classes, for some $1/\gamma < m < M$.* □

3. Proof of the main result

Proof of Theorem 1. Let $\epsilon > 0$, and let r be a positive integer. Let $\mu = 3\epsilon$ and let γ be sufficiently small so that $(\mu - \gamma)^r > 2\sqrt{\gamma}$. Let $M = M(\gamma)$ be the constant from Lemma 1. Let $K > rM/\gamma$ and let $\delta = 1/(2rM)$. Let $G = (V, E)$ be a digraph with n vertices where $n > K$, and assume that G is ϵn^2-cyclic. Finally, let $d \in (0, \delta n)$. We need to prove that G has a C_p^r where $d \le p \le d + K$.

We apply Lemma 1 to G and obtain a γ-regular partition of $V(G)$ into m parts, where $1/\gamma < m < M$. Denote the parts by $V_1, \dots, V_m$. Notice that the size of each part is either $\lfloor n/m \rfloor$ or $\lceil n/m \rceil$. For simplicity we may and will assume that $\ell = n/m$ is an integer, as this assumption does not affect our result since it is asymptotic. We say that the set of arcs $E(V_i, V_j)$ is *good* if (V_i, V_j) is a γ-regular pair and also $d(V_i, V_j) \geq \mu$. Notice that it is possible that $E(V_i, V_j)$ is good while $E(V_j, V_i)$ is not good (because it is too sparse).

Claim If $E(V_i, V_j)$ is good, then for every $X \subset V_i$ with $|X| \geq 2\gamma\ell$ and $Y \subset V_j$ with $|Y| \geq \sqrt{\gamma}\ell$, there exists $X^* \subset X$ with $|X^*| = r$ and and $Y^* \subset Y$ with $|Y^*| \geq |Y|(\mu - \gamma)^r$ so that for each $x \in X^*$ and $y \in Y^*$, $(x, y) \in E(V_i, V_j)$.

The proof of the claim is reminiscent of the *key embedding lemma* for applying the regularity lemma in undirected graphs (see [6]). We prove the claim by induction on r. Namely, we prove that for $t = 0, \dots, r$ there exists $X_t \subset X$ with $|X_t| = t$ and $Y_t \subset Y$ with $|Y_t| \geq |Y|(\mu - \gamma)^t$ so that for each $x \in X_t$ and $y \in Y_t$, $(x, y) \in E(V_i, V_j)$. This is clearly true for $t = 0$ by setting $X_0 = \emptyset$ and $Y_0 = Y$. Assuming this holds for t, we prove it for $t+1$. Let $X' \subset X \setminus X_t$ be those vertices that have less than $(\mu - \gamma)|Y_t|$ outgoing neighbors in Y_t. We claim that $|X'| \leq \gamma\ell$. Indeed, if this were not the case, the ordered pair (X', Y_t) would have density less than $\mu - \gamma$, while both $|X'| > \gamma\ell$ and

$$|Y_t| \geq |Y|(\mu - \gamma)^t \geq \sqrt{\gamma}\ell(\mu - \gamma)^r > \gamma\ell$$

thereby violating the γ-regularity of the pair (V_i, V_j). Now,

$$|X| - |X_t| - |X'| \geq 2\gamma\ell - t - \gamma\ell = \gamma\ell - t \geq \gamma\ell - r > \gamma\frac{K}{M} - r > 0.$$

Thus, let $x \in X \setminus (X_t \cup X')$. Let $Y_{t+1} \subset Y_t$ be the outgoing neighbors of x in Y_t. Hence, $|Y_{t+1}| \geq |Y_t|(\mu - \gamma) \geq |Y|(\mu - \gamma)^{t+1}$, and setting $X_{t+1} = X_t \cup \{x\}$ completes the induction step. Finally, setting $X^* = X_r$ and $Y^* = Y_r$ the claim follows.

Let G^* be the spanning subgraph of G consisting of the union of the good sets of arcs. That is, G^* is obtained from G by discarding arcs inside the parts, within non-regular pairs, within sparse pairs, or one-sided sparse pairs (a one-sided sparse pair is a pair with only one direction having density less than μ). We claim that G^* is not acyclic. To see this, we must show that that $|E(G)| - |E(G^*)| < \epsilon n^2$. Indeed, there are at most $m(n^2/m^2)$ arcs with both endpoints in the same vertex class, there are at most $\gamma\binom{m}{2}\frac{2n^2}{m^2}$ arcs within non-regular pairs, and there are at most $m(m-1)\mu\frac{n^2}{m^2}$ arcs within sparse pairs or one-sided sparse pairs. Thus,

$$|E(G)| - |E(G^*)| \leq m\frac{n^2}{m^2} + \gamma\binom{m}{2}\frac{2n^2}{m^2} + m(m-1)\mu\frac{n^2}{m^2} < 3\mu n^2 = \epsilon n^2.$$

Let R denote the m-vertex digraph whose vertices are $\{1, \dots, m\}$ and $(i, j) \in E(R)$ if and only if $E(V_i, V_j)$ is good. Notice that since G^* is not acyclic, R is not acyclic either. Assume, without loss of generality, that the shortest cycle of R consists of $(1, 2, \dots, s)$. Notice that, trivially, $2 \leq s \leq m$. We will prove that for all $2 \leq k \leq \lfloor \delta n/s + 2 \rfloor$, G has a C^r_{ks}. This clearly suffices, since, given $d \in (0, \delta n)$, choose $k \geq 2$ to be the smallest integer so that $p = ks \geq d$. Then, $d \leq p \leq d + 2s \leq d + 2m \leq d + 2M \leq d + K$ and $k = p/s \leq d/s + 2 \leq (\delta n)/s + 2$.

Let, therefore k be an integer satisfying $2 \le k \le \delta n/s+2$ and let $p = ks$. We will prove, by induction on $i = 2, \ldots, p-1$ that there are *disjoint* subsets of vertices $U_0, U_1, \ldots, U_i$ so that the following conditions hold.

(1) $U_j \subset V_{j \bmod s}$ for $j = 0, \ldots, i$ (for simplicity define $V_0 = V_s$).
(2) $|U_j| = r$ for $j = 1, \ldots, i-1$.
(3) $|U_0| \ge \sqrt{\gamma}\ell$, $|U_i| \ge 2\gamma\ell$.
(4) For $j = 1, \ldots, i$ and for every $x \in U_{j-1}$ and every $y \in U_j$, $(x, y) \in E(G)$.

We first prove the case $i = 2$. We will show that for all $t = 0, \ldots, r$, there is a set X_t of t vertices of V_1, a set $Y_t \subset V_2$ with $|Y_t| \ge \ell(\mu-\gamma)^t$ and a set $Z_t \subset V_s$ with $|Z_t| \ge \ell(\mu-\gamma)^t$ so that for each $x \in X_t$, for each $y \in Y_t$ and for each $z \in Z_t$, $(x, y) \in E$ and $(z, x) \in E$. Indeed, this is trivially true for $t = 0$. Assuming this is true for $t < r$ with sets X_t, Y_t Z_t, we prove it for $t+1$. Let $X' \subset V_1 \setminus X_t$ be those vertices with less than $(\mu-\gamma)|Y_t|$ outgoing neighbors in Y_t. We must have $|X'| \le \gamma\ell$ since otherwise, the pair (X', Y_t) would have density less than $\mu - \gamma$, while both X' and Y_t are larger than $\gamma\ell$, thus violating the γ-regularity of the pair (V_1, V_2). Similarly, if $X'' \subset V_1 \setminus X_t$ are those vertices with less than $(\mu-\gamma)|Z_t|$ incoming neighbors in Z_t we must have $|X''| \le \gamma\ell$. Since

$$|X_t| + |X'| + |X''| < r + 2\gamma\ell < \ell = |V_1|$$

we can choose $w \in V_1 \setminus (X_t \cup X' \cup X'')$. Letting $Y_{t+1} \subset Y_t$ be the outgoing neighbors of w in Y_t and $Z_{t+1} \subset Z_t$ be the incoming neighbors of w in Z_t we have $|Y_{t+1}| \ge \ell(\mu-\gamma)^{t+1}$ and $|Z_{t+1}| \ge \ell(\mu-\gamma)^{t+1}$. Defining $X_{t+1} = X_t \cup \{w\}$ the induction step is completed. Now, since we have

$$|Y_r| \ge \ell(\mu-\gamma)^r \ge 2\gamma\ell$$

we may take U_2 to be any subset of Y_r with size $\lceil 2\gamma\ell \rceil$. Defining $U_0 = Z_r \setminus U_2$ (we must be careful since we may have $s = 2$) we have

$$|U_0| \ge \ell(\mu-\gamma)^r - \lceil 2\gamma\ell \rceil > \sqrt{\gamma}\ell.$$

We have therefore proved the case $i = 2$.

Assuming that our claim holds for i, we prove it for $i+1$. Let $Y = V_{i+1} \setminus U_0 \cup \cdots \cup U_i$. Clearly, V_{i+1} can only intersect U_j for $j = i+1-ts$ for some positive integer t. It follows that

$$|Y| \ge \ell - r\frac{p}{s} - \sqrt{\gamma}\ell \ge \ell(1-\gamma) - r(\frac{\delta n}{s}+2) > \frac{\ell}{2}$$
$$= \ell(1-\gamma-\frac{r\delta m}{s}) - 2r > \sqrt{\gamma}\ell.$$

Putting $X = U_i$ we have $|X| \ge 2\gamma\ell$. Thus, we have set $X^* \subset X$ and $Y^* \subset Y$ as guaranteed in Claim 3. We may therefore redefine $U_i = X^*$ and define $U_{i+1} = Y^*$ and, noticing that

$$|U_{i+1}| = |Y^*| \ge |Y|(\mu-\gamma)^r \ge \sqrt{\gamma}\ell(\mu-\gamma)^r \ge 2\gamma\ell,$$

we have completed our induction step.

To complete the proof of Theorem 1 we consider U_0 and U_{p-1}. Since $s \ne 1$, they are disjoint. By our construction, if we define $U_{p-1} = X$ and $U_0 = Y$ we can again apply Claim 3 and obtain set $X^* \subset U_{p-1}$ and $Y^* \subset U_0$. Since $|X^*| = r$ we redefine $U_{p-1} = X^*$, and we redefine U_0 to be any r-subset of Y^*. It follows that $U_0, \ldots, U_{p-1}$ are now each of size precisely r, and induce a copy of C_p^r in G. We have thus completed the proof of Theorem 1. $\square$

The constant $K = K(\epsilon, r)$ in Theorem 1 is huge, and the constant $\delta = \delta(\epsilon, r)$ is very small. It is not difficult to see that we must have $\delta = O(\epsilon)$, even if $r = 1$. Consider a random regular tournament of degree $2\epsilon n$. It has $4\epsilon n+1$ vertices, and it is $4\epsilon^2 n^2(1-o(1))$-cyclic. Hence, by taking roughly $\frac{1}{4}\epsilon^{-1}$ vertex-disjoint copies of such a tournament we get an n-vertex digraph that is $\epsilon n^2(1-o(1))$-cyclic and no cycle has length greater than $4\epsilon n + 1$. Thus, $\delta \leq 4\epsilon$.

It is not difficult to see that we must have $K = \Omega(\epsilon^{-1/2})$, even if $r = 1$. Consider the digraph obtained by taking $\epsilon^{-1/2}$ disjoint sets $V_1, \ldots, V_{\epsilon^{-1/2}}$ of size $\epsilon^{1/2}n$ each, and adding all possible arcs from V_i to V_{i+1} (indices modulo $\epsilon^{-1/2}$). The length of every cycle in this graph is a multiple of $\epsilon^{-1/2}$ and it cannot be made acyclic by removing less than ϵn^2 arcs. Thus, $K \geq \epsilon^{-1/2}$.

4. An application

As mentioned in the introduction, we prove the following stronger version of Theorem 2.

Theorem 3. *For every $\epsilon_1 > 0$, and every positive integer r, there are constants K_1 and δ_1 so that for all $n > K_1$, and for all $d \in (0, \delta_1 n)$, the set of vertices of every digraph with n vertices can be partitioned into parts $V_0, V_1, \ldots, V_t$ so that V_0 induces a subgraph that can be made acyclic by removing at most $\epsilon_1 n^2$ arcs and V_i induces an r-hamiltonian digraph where $d \leq |V_i| \leq d + K_1$ for $i = 1, \ldots, t$.*

Proof: First notice that in the case $r = 1$ the requirement that $n > K_1$ is not necessary. This simply follows from the fact that if $n \leq K_1$ we can simply take as many vertex-disjoint cycles as we can, and then remain with an acyclic subgraph. Thus, the case $r = 1$ in Theorem 3 simply amounts to Theorem 2.

Let $\epsilon_1 > 0$, and let r be a positive integer. Define $\epsilon = \epsilon_1/2$. Let $K = K(\epsilon)$ and $\delta = \delta(\epsilon)$ be as in Theorem 1. Define $K_1 = Kr/\epsilon$ and $\delta_1 = r\delta\epsilon$. Let $G = (V, E)$ be a digraph with $n > K_1$ vertices, and let $d \in (0, \delta_1)$. Greedily pick vertex-disjoint r-blowups of cycles as long as their number of vertices is between d and $d + K_1$. Assume that when the process ends, we remain with an induced subgraph $G[W]$ of G on the vertex set $W \subset V$. Starting with $i = 1$, as long as $G[W]$ has a vertex w_i with minimum out-degree (in $G[W \setminus \{w_1, \ldots, w_{i-1}\}]$) less than ϵn, we remove w_i and continue in the same manner. Once this process ends, we remain with an induced subgraph $G[U]$ of G on the vertex set $U \subset W$. If U can be made acyclic by removing at most ϵn^2 arcs we are done since by removing these arcs and the arcs going from w_i to $U \cup \{w_{i+1}, \ldots, w_{|W|-|U|}\}$ (there are at most ϵn^2 such arcs), we get a spanning acyclic sub-digraph of $G[W]$ showing that $G[W]$ can be made acyclic by removing at most $\epsilon_1 n^2$ arcs. We claim that, indeed, U *can* be made acyclic by removing at most ϵn^2 arcs. Indeed, assuming otherwise, we must have, in particular, $|U| \geq \epsilon n$. This implies that $|U| > \epsilon K_1 \geq Kr \geq K$ and that, trivially, U is $\epsilon |U|^2$-cyclic. Furthermore,

$$0 < \frac{d}{r} \leq \frac{\delta_1 n}{r} = \delta\epsilon n \leq \delta |U|.$$

Thus, by Theorem 1, $G[U]$ contains a C_p^r with $d/r \le p \le d/r + K$, that is, an r-blowup of a cycle whose number of vertices is between d and $d + K_1$, a contradiction. □

5. Concluding remarks

- It would be interesting to obtain a direct proof for the case $r = 1$ in Theorem 1 that *does not* use the regularity lemma, and that yields Corollary 1 with the correct orders of magnitude of δ and K as functions of ϵ. For the more general case $r > 1$ proved in Theorem 1 we suspect that the regularity lemma is indispensable.
- The proof of Theorem 1 is algorithmic, and can be implemented in polynomial time. Given an n-vertex graph, and $d \in (0, \delta n)$, the algorithm either finds a set of ϵn^2 arcs whose removal makes the graph acyclic, or else finds a C_p^r with $d \le p \le d+K$. The only non-constructive part in the proof of Theorem 1 is obtaining the γ-regular partition. This, in turn, can be done in polynomial time using the method from [1].
- Using a proof similar to the proof of Theorem 1 we can obtain the following result.

  Proposition 1. *For every $\epsilon > 0$, and every positive integer r, there are constants K and δ so that for every $n > K$, and for every $d \in (0, \delta n)$, every n-vertex digraph with minimum out-degree at least ϵn has a C_p^r where $d \le p \le d + K$.*

- Unlike Theorem 1 and corollary 1 that give conditions guaranteeing a C_p^r whose size deviates from a given number only by a constant, the problem of finding *long* cycles in kn-cyclic digraphs is significantly easier. Indeed, every kn-cyclic digraph has a subdigraph with minimum out-degree greater than k; as long as there is a vertex with out-degree at most k, delete it, and continue. The process must halt while there are vertices still remaining, forming a subdigraph with minimum out-degree at least $k+1$. This subgraph has a cycle of length at least $k + 2$.

Acknowledgements. I wish to thank Noga Alon and the referees for insightful suggestions.

References

1. Alon N., Duke R.A., Lefmann H., Rödl V., Yuster R.: The algorithmic aspects of the Regularity Lemma. Journal of Algorithms **16**, 80–109 (1994)

2. Alon N., Shapira A.: Testing subgraphs in directed graphs. Proc. 35th ACM STOC, ACM Press 700–709 (2003)

3. Bollobás B.: *Modern Graph Theory*, Springer-Verlag (1998)

4. Czygrinow A., Poljak S., Rödl V.: Constructive Quasi-Ramsey numbers and tournament ranking. SIAM J. Disc. Math. **12**, 48–63 (1999)

5. de la Vega W.F.: On the maximal cardinality of a consistent set of arcs in a random tournament. J. Combinatorial Theory, Ser. B **35**, 328-332 (1983)

6. Komlós J., Simonovits M.: Szemerédi regularity lemma and its application in Graph Theory. *Paul Erdős is 80, Proc. Coll. Bolyai Math. Soc. Vol 2*, (Bolyai Math. Soc., Keszthely) 295–352 (1993)

7. Szemerédi E.: Regular partitions of graphs. *Proc. Colloque Inter. CNRS 260*, (CNRS, Paris) 399–401 (1978)

Received:
Final version received: