AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **48** (2010), Pages 87–109

On the sizes of the graphs G, G^r, $G^r \setminus G$: the directed case

David Auger Irène Charon Olivier Hudry

Institut Telecom
Telecom ParisTech & Centre National de la Recherche Scientifique
LTCI UMR 5141, 46, rue Barrault, 75634 Paris Cedex 13
France
{david.auger, irene.charon, olivier.hudry}@telecom-paristech.fr

Antoine Lobstein

Centre National de la Recherche Scientifique
LTCI UMR 5141 & Telecom ParisTech
46, rue Barrault, 75634 Paris Cedex 13
France
antoine.lobstein@telecom-paristech.fr

Abstract

Let G be a directed graph and G^r be its r-th power. We study different issues dealing with the number of arcs, or size, of G and G^r: given the order and diameter of a strongly connected digraph, what is its maximum size, and which are the graphs achieving this bound? What is the minimum size of the r-th power of a strongly connected digraph, and which are the graphs achieving this bound? Given all strongly connected digraphs G of order n such that $G^r \neq K_n$, what is the minimum number of arcs that are added when going from G to G^r, and which are the graphs achieving this bound?

1 Introduction

Before we expound our study, we first give, for directed graphs, some very basic definitions and notation, which can be found, e.g., in [3], [4].

1.1 Definitions and Notation

We shall denote by $G = (V, A)$ a *directed graph*, or *digraph*, or simply *graph*, with vertex set V and arc set A, where an *arc* (i.e., a directed edge) from $x \in V$ to $y \in V$ is denoted by the pair (x, y). We say that x is the *origin* of the arc, and y its *end*; the origin and the end are the *extremities* of the arc. We require the graph to have no loops nor double arcs — but the arcs (x, y) and (y, x) can simultaneously exist; in this case, we say that there is a *symmetric arc* between x and y. The *size* of a digraph is its number of arcs, and its *order* is its number of vertices.

A *directed path*, or simply *path*, $P = x_0x_1 \dots x_\ell$, is a sequence of vertices x_i, $0 \leq i \leq \ell$, such that $(x_i, x_{i+1}) \in A$, $0 \leq i \leq \ell - 1$. The *length* of P is its number of arcs, ℓ. A digraph is called *strongly connected* if for any two vertices x and y there is a path going from x to y.

In a strongly connected graph G, we can define the distance from any vertex x to any vertex y, denoted by $d_G(x, y)$, as the number of arcs in any shortest directed path from x to y, since such a path exists. Note that in general, $d_G(x, y)$ is not equal to $d_G(y, x)$. We shall say that two vertices x, y are at distance in G greater (respectively, smaller) than s from one another if $d_G(x, y) > s$ or $d_G(y, x) > s$ (respectively, if $d_G(x, y) < s$ and $d_G(y, x) < s$). The *diameter* of a strongly connected graph G is the maximum distance in the graph:

$$\varphi(G) = \max_{x \in V, y \in V} d_G(x, y).$$

Given an integer $r \geq 1$, the *r-th power*, or *r-th transitive closure*, of the graph $G = (V, A)$ is the graph $G^r = (V, A^r)$, where, for two distinct vertices x and y, the arc (x, y) belongs to A^r if and only if $d_G(x, y) \leq r$.

The *clique*, or *complete graph*, K_n, is the digraph of order n with all possible $n(n-1)$ arcs. Finally, an *induced subgraph* of $G = (V, A)$ is a graph $G^* = (V^*, A^*)$ where $V^* \subseteq V$ and $A^* = \{(x, y) : x \in V^*, y \in V^*, (x, y) \in A\}$; a *subgraph* is such that A^* is included in $\{(x, y) : x \in V^*, y \in V^*, (x, y) \in A\}$.

1.2 Scope of the Paper

We are interested in the following related problems on sizes and powers, for digraphs:

(a) Given the order and diameter of a strongly connected digraph, what is its maximum size, and which are the graphs achieving this bound?

The first part of this question received its answer in [5] as far back as 1960; in Section 2, we characterize the graphs reaching this bound.

(b) Given an integer $r \geq 2$, what is the minimum size of a digraph of order n, of which it is known that it is the r-th power of a strongly connected digraph, and which are the graphs achieving this bound?

The answer is in Section 3.

(c) Given an integer $r \geq 2$ and all strongly connected digraphs G of order n such that $G^r \neq K_n$, what is the minimum number of arcs that are added when going from G to G^r, and which are the graphs achieving this bound?

We give the answer in Section 4.

Similar issues for undirected graphs are treated in [2], the motivation coming from previous studies by Aingworth et al. [1] and Ore [6]: for instance, Ore determined the maximum number of edges for undirected connected graphs with given order and diameter, and characterized the graphs reaching this bound, which is exactly Question (a) above for undirected graphs.

2 The Size of a Digraph with Given Order and Diameter

The following result is the corrected version of the theorem which is proved in [5] and contains a very small inaccuracy.

Theorem 1 *[5], [3, Th. 2.4.6] Let $G = (V, A)$ be a strongly connected digraph of order n and size m. Then*

$$\varphi(G) \leq \begin{cases} n-1 & \text{if } n \leq m \leq \frac{n(n+1)}{2} - 1, \\ \left\lfloor n + \frac{1}{2} - \sqrt{2m - n^2 - n + \frac{17}{4}} \right\rfloor & \text{if } \frac{n(n+1)}{2} - 1 \leq m \leq n(n-1). \end{cases}$$

From this we can immediately derive the following theorem. We still give a proof of Theorem 2, because we use it when characterizing the graphs which reach the bound.

Theorem 2 *Let $G = (V, A)$ be a strongly connected digraph of order n and diameter $\varphi \geq 2$. Then the size of G is at most*

$$\frac{\varphi(\varphi+3)}{2} + (n - \varphi - 1)(n + 2). \tag{1}$$

Proof. Let $z_1, z_2 \in V$ be such that $d_G(z_1, z_2) = \varphi$, and C be a shortest directed path from z_1 to z_2: $C = x_0 x_1 \dots x_\varphi$, with $x_0 = z_1$ and $x_\varphi = z_2$; there are no more arcs (x_i, x_j), $i < j$, but any arc (x_i, x_j), $i > j$, can exist. In G, the remaining vertices y_j, $1 \leq j \leq n - \varphi - 1$, can at most constitute the clique $K_{n-\varphi-1}$, and each y_j can be part of at most $\varphi + 4$ arcs with ends or origins in C: this is clear if $\varphi = 2$; if $\varphi \geq 3$ and there are $\varphi + 5$ arcs (y_j, x_k) or (x_k, y_j), then there are at least four vertices x_i such that both (y_j, x_i) and (x_i, y_j) belong to A. This in turn implies that in A there are two arcs (x_{i_1}, y_j) and (y_j, x_{i_2}) with $i_1 + 3 \leq i_2$. This is impossible, since the path $x_0 \dots x_{i_1} y_j x_{i_2} \dots x_\varphi$ would be shorter than C. All in all, we have at most

$$\varphi + \frac{\varphi(\varphi+1)}{2} + (n - \varphi - 1)(n - \varphi - 2) + (\varphi + 4)(n - \varphi - 1)$$

arcs in G, which amounts to (1). □

We shall see that this theorem is also a direct consequence of Theorem 5. We set

$$\sigma(\varphi, n) = \frac{\varphi(\varphi+3)}{2} + (n - \varphi - 1)(n + 2), \tag{2}$$

and we are going to characterize the graphs $G = (V, A)$ reaching $\sigma(\varphi, n)$. The previous proof shows that necessarily G consists of the path $C = x_0 x_1 \dots x_\varphi$, all the arcs (x_i, x_j), $i > j$, the clique $K_{n-\varphi-1}$ and exactly $\varphi + 4$ arcs between every vertex $y \in K_{n-\varphi-1}$ and the vertices of C. All we have to determine is how these $(\varphi + 4)(n - \varphi - 1)$ arcs are located.

We observe that, in particular, for each $y \in K_{n-\varphi-1}$, there are in A at least three arcs (x_i, y) and three arcs (y, x_j). Let h be the smallest subscript such that there is a vertex $y_1 \in K_{n-\varphi-1}$ with $(x_h, y_1) \in A$; the parameter h can vary from 0 to $\varphi - 2$. Let k be the largest subscript such that $(y_1, x_k) \in A$. Because $x_h y_1 x_k$ must not allow

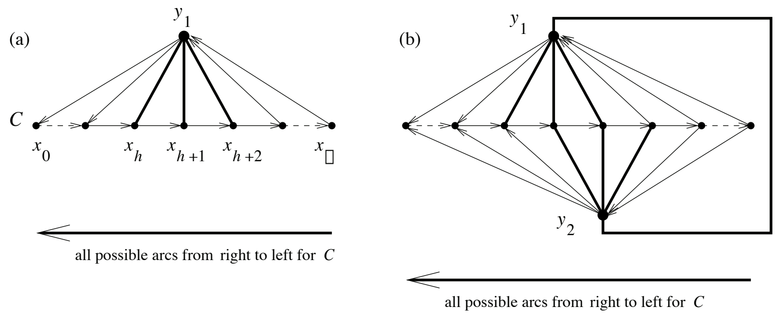

Figure 1: The locations of the arcs between the vertices in C and the clique. A bold line without arrow represents a symmetric arc.

shortcuts with respect to C, we have $k \le h+2$. Now one can see that necessarily: $k = h+2$, and we must have in A all the $\varphi - h + 1$ arcs (x_i, y_1), $h \le i \le \varphi$, and all the $h+3$ arcs (y_1, x_j), $0 \le j \le h+2$, see Figure 1(a).

Now consider another vertex y_2 in the clique. If $(x_h, y_2) \in A$, then everything goes as with y_1. If $(x_h, y_2) \notin A$, let ℓ (respectively, m) be the largest (respectively, smallest) subscript such that $(y_2, x_\ell) \in A$ (respectively, $(x_m, y_2) \in A$); immediately, $m \ge h+1$, and, because of the forbidden shortcut $x_h y_1 y_2 x_\ell$, $\ell \le h+3$. And again, there is no choice left for ℓ and m: $\ell = m+2$ and by the previous inequalities on ℓ and m, we have $\ell = h+3$, $m = h+1$, and we must have in A all the $h+4$ arcs (y_2, x_j), $0 \le j \le h+3$, and all the $\varphi - h$ arcs (x_i, y_2), $h+1 \le i \le \varphi$, see Figure 1(b).

So the clique is divided into at most two types of vertices, those with $(x_h, y) \in A$, and the others, with $(x_h, y) \notin A$ but $(x_{h+1}, y) \in A$. When h varies from 0 to $\varphi - 2$, we obtain the description of all the graphs achieving the bound $\sigma(\varphi, n)$ (since we did not introduce shortcuts, their diameter is indeed φ, the distance from x_0 to x_φ). Note that if $h = \varphi - 2$, we can have only one type of vertex in the clique.

3 Size of the Power of a Digraph

We address the following issue: given an integer $r \ge 2$ and all strongly connected digraphs of order n, $G = (V, A)$, what is the smallest number of arcs in G^r? and which are the graphs which meet this bound?

We give the complete answer in the next two theorems. If $r \ge n-1$, then $G^r = K_n$ and the problem is trivial, so we assume that $r \le n-2$.

Theorem 3 *If $r \le n-2$ and $G = (V, A)$ is a strongly connected digraph of order n, then the size of G^r is at least nr.*

Proof. Let $x \in V$. If for all $y \in V$, $d_G(x, y) \le r$, then the $n-1$ arcs (x, y), $y \in V \setminus \{x\}$, are in G^r. If there is a vertex y such that $d_G(x, y) > r$, consider a shortest path $xz_1z_2\dots z_r\dots y$ from x to y. Then the r arcs (x,z_i), $1 \le i \le r$, are in G^r. In both cases, since $r \le n-1$, we see that each of the n vertices x brings at least r arcs with origin x to G^r. □

Theorem 4 *If $r \le n-2$, the only strongly connected digraph $G = (V,A)$ with order n such that G^r has size exactly nr is the* circuit $x_0x_1\dots x_{n-1}x_0$.

Proof. Obviously, the circuit achieves the bound rn. If G meets rn, that is, if $|A^r| = rn$, then, following the proof of Theorem 3, each vertex x must contribute exactly r to G^r with arcs originating in x. Since $n-1>r$, there is a shortest path with length greater than r, $P = xz_1z_2\dots z_r\dots y$, from x to some $y \in V$ (otherwise, all $n-1$ vertices in $V \setminus \{x\}$ being within distance r from x, there would be more than r arcs with origin x in G^r). If there is an arc $(x,w) \in A$, $w \ne z_1$, then x gives at least $r+1$ arcs with origin x to G^r. So in A the only arc with origin x is (x,z_1) and we have just proved that every vertex is the origin of exactly one arc in G. Since G is strongly connected, the only possibility is the circuit. □

4 From G to G^r

We consider an integer $r \ge 2$ and all strongly connected digraphs $G = (V,A)$, of order n, such that $G^r \ne K_n$, and we want to determine what is the minimum number of arcs that have to be added to go from G to G^r, i.e., what is the minimum cardinality of $A^r \setminus A$: we shall denote this number by $\Lambda(r,n)$. Once we know the value of $\Lambda(r,n)$ (Theorem 5), we shall characterize the graphs reaching it in Section 4.2.

4.1 Minimum Number of Arcs

Observe that the condition $G^r \ne K_n$ is equivalent to the fact that G has diameter at least $r+1$, so

$$n > \varphi(G) \ge r+1.$$

In the following theorem, we can see that $\Lambda(r,n)$ is linear in n, with the factor $r-1$.

Theorem 5 *For $r \ge 2$ and $n \ge r+2$, we have:*

$$\Lambda(r,n) = (r-1)(n-1-\frac{r}{2}). \tag{3}$$

Proof. In the course of this proof, we shall prove Lemmas 6–9 and 11, and Corollaries 10 and 12.

Note that equality (3) contains the case $r=1$ (no arc is added). We set

$$b(r,n) = (r-1)(n-1-\frac{r}{2}).$$

First, we exhibit a digraph $G_0 = (V_0,A_0)$ showing that $\Lambda(r,n) \le b(r,n)$. This graph has vertex and arc sets defined by

$$V_0 = \{x_i : 0 \le i \le n_0 - 1\}, \tag{4}$$

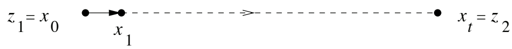

Figure 2: The path $C = x_0x_1 \dots x_t$ used in the proof of Theorem 5.

$$A_0 = \{(x_i, x_{i+1}) : 0 \leq i \leq n_0 - 2\} \cup \{(x_j, x_k) : 0 \leq k < j \leq n_0 - 1\}. \qquad (5)$$

In other words, G_0 consists of a directed path going from x_0 to x_{n_0-1}, plus all arcs going from a vertex x_j to a vertex of smaller subscript.

Since G_0 has diameter at least $r + 1$, we have $n_0 \geq r + 2$. The arcs in $A_0^r \setminus A_0$ are the arcs (x_i, x_j) with $0 \leq i \leq n_0 - 3$ and $i + 2 \leq j \leq \min\{i + r, n_0 - 1\}$. So there are $r - 1$ additional arcs starting from x_i as long as $i \leq n_0 - 1 - r$, and $(r - 2), (r - 3), \dots, 1$ additional arcs starting from the subsequent vertices, $x_{n_0-r}, x_{n_0-r+1}, \dots, x_{n_0-3}$, respectively. All in all, we have $(n_0 - r)(r - 1) + \frac{1}{2}(r - 1)(r - 2) = b(r, n_0)$ new arcs, which proves the upper bound for $\Lambda(r, n)$.

Now, let $G = (V, A)$ be any digraph fulfilling the hypotheses, and $G^* = (V^*, A^*)$ be a strongly connected induced subgraph of G, having minimum order, n^*, among all the strongly connected induced subgraphs of G which have two vertices at distance in G greater than r from one another — since G has diameter at least $r + 1$, such two vertices exist and if necessary we take $G^* = G$, so such a graph G^* always exists. We name z_1 and z_2 these two vertices, so that $z_1 \in V^*$, $z_2 \in V^*$ and, say, $d_G(z_1, z_2) = \theta > r$. Obviously, if $t = d_{G^*}(z_1, z_2)$, then $t \geq \theta > r$.

It is useful to give a name to the following property, which simply uses the very definition of G^*:

> (P) If H is a strongly connected subgraph of G with order n_H such that $2 \leq n_H < n^*$, then any two vertices x, y in H are at distance in G at most r from one another; consequently, $(x, y) \in A^r$ and $(y, x) \in A^r$.

(1) In a first step, we are going to show that V^* contains at least $b(r, n^*)$ couples of vertices u, v such that the arc (u, v) belongs to $A^r \setminus A$. Such couples are called *G^*-friendly couples*, with origin u and end v. This first step will stop when (13) is established, and will use Lemmas 6 and 7.

Let C be a shortest path in G^* from z_1 to z_2:

$$C = x_0x_1 \dots x_t, \qquad (6)$$

with $x_0 = z_1, x_t = z_2$ and $t \geq \theta > r \geq 2$, see Figure 2.

The path C is a subpath of the graph G_0 defined by (4) and (5), with $n_0 - 1 = t$: same vertices, fewer arcs; therefore, there are at least the $b(r, t + 1)$ arcs (x_i, x_j), $i + 2 \leq j$, to be added when going from G^* to $(G^*)^r$, cf. the above study of G_0. Moreover, these couples x_i, x_j are G^*-friendly: (a) $x_i, x_j \in V^*$, (b) $(x_i, x_j) \notin A$ because $(x_i, x_j) \notin A^*$, and (c) $(x_i, x_j) \in A^r$ because $(x_i, x_j) \in (A^*)^r \subseteq A^r$. So:

$$\text{there are at least } b(r, t + 1) \;\; G^*\text{-friendly couples } x_i, x_j \text{ in } C, \text{ with } i < j. \qquad (7)$$

If every vertex in G^* belongs to C, i.e., $n^* = t+1$, then we have proved the existence of at least $b(r, n^*)$ G^*-friendly couples, which is what we wanted. So from now on, we assume that there is at least one vertex in G^* which is not a vertex $x_i, 0 \le i \le t$. We denote by Y^* the set of these vertices:

$$Y^* = V^* \setminus \{x_i : 0 \le i \le t\}, \text{ with } |Y^*| = n^* - t - 1.$$

(1a) In an intermediate step, our goal is to prove the following property (Q):

(Q) for any vertex $y \in Y^*$, we add, when going from G to G^r, at least $r-1$ distinct arcs such that:

either (i) the origin is y and the end is in G^*,

or (ii) the end is y and the origin is in C.

As a consequence, we never count twice the same arc for two different y_1, y_2 in Y^*. Obviously, these arcs yield G^*-friendly couples. To prove that (Q) is true, we state a first lemma.

Lemma 6 *Let $y \in Y^*$. If for all x_i in C, $d_G(x_i, y) \le r$ and $d_G(y, x_i) \le r$, then property (Q) holds.*

Proof of Lemma 6. By assumption, all the arcs (x_i, y) and (y, x_i) belong to A^r, so it is sufficient to show that at least $r-1$ of them do not belong to A.

If no i exists such that $(x_i, y) \in A$, then (Q) holds, because $t > r$. So we can assume that there is a smallest k, $0 \le k \le t$, such that $(x_k, y) \in A$. We use the fact, which is true for all j between $k+3$ and t, that $(y, x_j) \notin A$: otherwise the two arcs $(x_k, y), (y, x_j)$, belonging to $A^* \subseteq A$, would contribute to provide, in G^*, a path shorter than C from z_1 to z_2.

Therefore, the arcs (y, x_j), $k+3 \le j \le t$, and (x_i, y), $0 \le i \le k-1$, do not belong to A, i.e., all in all, $t-2 \ge r-1$ arcs, which proves Lemma 6. $\square$

Back to the proof of Theorem 5, we consider a shortest path W in G^* from $y \in Y^*$ to x_0:

$$W = w_0 w_1 \dots w_q,$$

with $w_0 = y$ and $w_q = x_0$; see Figure 3. Note that the intersection between C and W is not necessarily reduced to x_0. For i between 2 and q, the arc (y, w_i) does not belong to A, because W is a shortest path, and for i between 2 and $\min\{q, r\}$, the arc (y, w_i) belongs to A^r and satisfies (i) in (Q). If $q \ge r$, then (Q) holds, so from now on we assume that $q < r$, and we have just shown that

$$\text{in } A^r \setminus A, \text{ there are } q-1 \text{ arcs } (y, w_i),\ 2 \le i \le q, \text{ satisfying (i) in (Q).} \tag{8}$$

We consider the vertices $x_0, x_1, \dots, x_{r-q}$ in C. By the triangle inequality, for i between 0 and $r-q$, we have $d_{G^*}(y, x_i) \le q + (r-q) = r$ and so

$$d_G(y, x_i) \le r \text{ for } 0 \le i \le r-q. \tag{9}$$

We are now ready to prove Lemma 7.

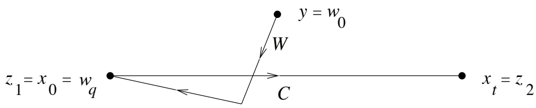

Figure 3: The paths C and W used in the proof of Theorem 5.

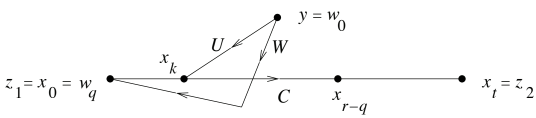

Figure 4: The paths C, W and U used in the proof of Lemma 7.

Lemma 7 *If there exists a path in G^* from y to x_k, $1 \leq k \leq r-q$, which does not go through x_0, then (Q) is true.*

Proof of Lemma 7. Among these paths, we choose a path U for which k is minimum; see Figure 4. For i between 1 and $k-1$, the arcs (y, x_i) are in $A^r \setminus A$, because of (9) and the minimality of k, and these arcs satisfy (i) in (Q), so we have just shown that

$$\text{in } A^r \setminus A, \text{ there are } k-1 \text{ arcs } (y, x_i),\ 1 \leq i \leq k-1, \text{ satisfying (i).} \tag{10}$$

Let M be a path in G^* from x_t to y; see Figure 5. Then M does not go through x_0: otherwise, consider the subpath M^* of M going from x_t to x_0, and $C \cup M^*$, by which we mean the induced subgraph of G with vertices in $C \cup M^*$, so that there can be more arcs than simply the arcs of the path C and the arcs of the subpath M^*. Now this graph is strongly connected, contains $x_0 = z_1$ and $x_t = z_2$ which are at distance greater than r in G, and does not contain y, thus contradicting the minimality of G^*.

Consider next the subpath C^* of C going from x_k to x_t; see Figure 6, and $U \cup C^* \cup M$; this graph is strongly connected and contains fewer vertices than G^* (because it does not contain x_0), and so, by property (P), we have

$$d_G(y, x_i) \leq r \text{ and } d_G(x_i, y) \leq r, \text{ for } k \leq i \leq t. \tag{11}$$

If for all i between $k+1$ and $t-1$, the arcs (x_i, y) are not in A, then they are in $A^r \setminus A$ and satisfy (ii) in (Q). So, together with our first $k-1$ arcs from (10), we

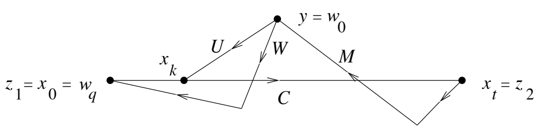

Figure 5: The paths C, W, U and M used in the proof of Lemma 7.

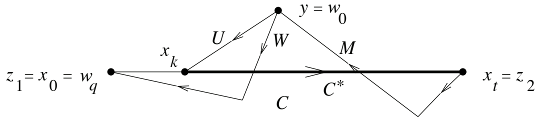

Figure 6: The paths C, W, U, M, and the path C^* (in bold) used in the proof of Lemma 7.

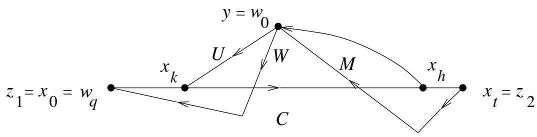

Figure 7: The paths C, W, U, M, and the arc (x_h, y) used in the proof of Lemma 7.

have at least $(k-1)+(t-k-1) = t-2 \geq r-1$ suitable arcs. We assume finally that there is an arc (x_h, y) in A, with $k+1 \leq h \leq t-1$; see Figure 7.

Consider, with abuse of notation, the graph $\mathcal{G} = (x_h, y) \cup W \cup \{x_0, x_1, \ldots, x_h\}$. This strongly connected graph cannot go through x_t, otherwise it would do so along W, and, since W goes back to x_0, we would again have the vertices in C strongly connected between themselves, contradicting the minimality of G^*. So this graph $\mathcal{G}$ has fewer vertices than G^*, and therefore by property (P), for all i between 0 and h, we have: $d_G(y, x_i) \leq r$, $d_G(x_i, y) \leq r$. Using (11) and $h \geq k+1$, we see that we are in the conditions of Lemma 6, which shows that (Q) is true and ends the proof of Lemma 7. □

Back to the proof of Theorem 5, we can conclude about property (Q), assuming finally that the hypothesis of Lemma 7 is not fulfilled: in particular, in A^*, hence in A, there is no arc (y, x_i) for i between 1 and $r-q$, but, by (9), all these arcs are in A^r. So:

$$\text{in } A^r \setminus A, \text{ there are } r-q \text{ arcs } (y, x_i),\ 1 \leq i \leq r-q, \text{ satisfying (i).} \tag{12}$$

Also because the conditions of Lemma 7 do not apply, and unlike in Figure 4, a vertex w_i in W, $2 \leq i \leq q$, cannot coincide with a vertex x_j, $1 \leq j \leq r-q$, for otherwise the beginning of the path W from y to x_0 would be a path from y to x_j not going through x_0. Therefore we can add the arcs obtained in (8) and (12), which proves that property (Q) holds in all cases, and marks the end of our intermediate step.

(1b) The first step is now almost complete: by (7) and because (Q) provides G^*-friendly couples which are counted only once, we have in G^* at least

$$b(r, t+1) + (r-1)|Y^*| = b(r, t+1) + (r-1)\big(n^* - t - 1\big) = b(r, n^*) \tag{13}$$

G^*-friendly couples, which ends our first step.

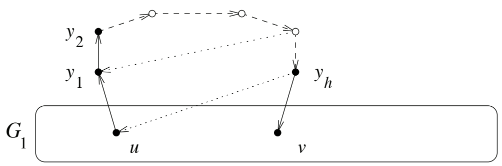

Figure 8: The graph G_1 and a smallest ear induce the graph G_2.

(2) The sequel of the proof of Theorem 5 will use Lemmas 8, 9 and 11, as well as Corollaries 10 and 12.

Now we consider all the strongly connected induced subgraphs $\widehat{G}$ of G, of order $\widehat{n}$, containing two vertices at distance in G greater than r from one another, and containing at least $b(r, \widehat{n})$ $\widehat{G}$-friendly couples — we have just proved that such graphs exist; among them, we take one, $G_1 = (V_1, A_1)$, with order n_1. If $n_1 = n$, then $G_1 = G$, there are at least $b(r, n)$ G-friendly couples, i.e., at least $b(r, n)$ arcs in $A^r \setminus A$, and Theorem 5 is proved. So from now on, we assume that $V \setminus V_1 \neq \emptyset$.

Since G is strongly connected, there is a smallest set of vertices $Y = \{y_1, y_2, \ldots, y_h\} \subseteq V \setminus V_1$ such that $H = y_1 y_2 \ldots y_h$ is a directed path in G and the arcs (u, y_1), (y_h, v) are in A, where u and v are two, non necessarily distinct, vertices in V_1; see Figure 8. We call Y an *ear* and set $G_2 = (V_2, A_2)$, the induced subgraph of G with vertex set $V_2 = V_1 \cup Y$; the arc set A_2 contains, among others, the arcs of A_1 and of H, as well as (u, y_1) and (y_h, v). Because of the minimality of Y, there is in A_2, hence in A, no arc (y_i, y_j) with $1 \leq i < j \leq h$, no arc with origin in V_1 and end y_i, $2 \leq i \leq h$, and no arc with origin y_i, $1 \leq i \leq h-1$, and end in V_1.

Trivially, every G_1-friendly couple is G_2-friendly, and G_1- and G_2-friendly couples all give arcs in $A^r \setminus A$.

We are going to show that every vertex in Y belongs to $r-1$ G_2-friendly couples, except possibly one vertex; in this case however, this will be compensated by one vertex in Y belonging to $2(r-1)$ G_2-friendly couples. In any case, no G_2-friendly couple will be counted twice.

We first assume that $h \geq 2$.

Let $y \in Y$. If there is a vertex x in V_2 such that $d_{G_2}(y, x) \geq r+1$, then

$$\text{in } A^r \setminus A, \text{ there are at least } r-1 \text{ arcs with origin } y \text{ and end in } V_2. \tag{14}$$

Indeed, a shortest path in G_2 from y to x: $z_0 z_1 \ldots z_t$, with $z_0 = y$, $z_t = x$ and $t > r$, shows that the $r-1$ arcs (y, z_i), $2 \leq i \leq r$, belong to $A^r \setminus A$.

Let $y \in Y \setminus \{y_h\}$. If all vertices x in V_2 are such that $d_{G_2}(y, x) \leq r$, then

$$\text{in } A^r \setminus A, \text{ there are at least } r-1 \text{ arcs with origin } y \text{ and end in } V_1. \tag{15}$$

Indeed, for all $x \in V_1$, the arc $(y, x) \notin A$, as mentioned above, and $(y, x) \in A^r$ because $d_{G_2}(y, x)$, hence $d_G(y, x)$, is at most r. Since G_1 contains two vertices at distance in G greater than r from one another, its order is at least $r-1$ and claim (15) is true.

Gathering (14) and (15), we obtain immediately the following lemma.

Lemma 8 *If $h \geq 2$, then, for all $y \in Y \setminus \{y_h\}$,*

$$\textit{there are at least } r-1 \ \ G_2\textit{-friendly couples with origin } y. \tag{16}$$

□

Lemma 9 *If $h \geq 2$, then*

$$\textit{there are at least } r-1 \ \ G_1\textit{-friendly couples with end } y_2. \tag{17}$$

Proof of Lemma 9. Take any vertex $x \in V_1$ and a shortest path C in G_2 from x to y_2. By the minimality of the ear Y, this path goes through vertices in V_1, then goes to y_1, and then it goes to y_2, by minimality of C.

First, we assume that there is a vertex $w \in V_1$ such that $d_{G_2}(w, y_2) \geq r$. By the above remark and because $r \geq 2$, there is a vertex $z \in V_1$ with $d_{G_2}(z, y_2) = r$ (and so $d_G(z, y_2) \leq r$), and the first $r-1$ vertices in a shortest path in G_2 from z to y_2 belong to V_1. If we call these $r-1$ vertices z_0, z_1, …, z_{r-2}, then the arcs (z_i, y_2) belong to $A^r \setminus A$, and claim (17) holds.

If, on the other hand, for all $w \in V_1$, $d_{G_2}(w, y_2) < r$, then for all vertices w in V_1, $(w, y_2) \in A^r \setminus A$, and (17) follows, which proves Lemma 9. □

Corollary 10 *If $h \geq 2$, then there are at least $|Y|(r-1)$ distinct arcs in $A^r \setminus A$ with one end or one origin in Y and the other extremity in V_2, i.e., at least $|Y|(r-1)$ distinct G_2-friendly couples with one extremity in Y.*

Proof of Corollary 10. Simply add up the arcs, or friendly couples, obtained in (16) and in (17): if $h \geq 3$, then y_h gives no arc, y_2 gives $r-1$ arcs with origin y_2 and end in V_2, and $r-1$ arcs with origin in V_1 and end y_2, and the remaining vertices $y_i \in Y$ each give $r-1$ arcs with origin y_i and end in V_2; if $h = 2$, then y_1 gives $r-1$ arcs with origin y_1 and end in V_2, and y_2 gives $r-1$ arcs with origin in V_1 and end y_2. All these arcs are distinct, which proves Corollary 10. □

Back to the proof of Theorem 5, we are left with the case $h = 1$, that is, $V_2 = V_1 \cup \{y\}$.

Lemma 11 *If $Y = \{y\}$, then there are at least $r-1$ G_1-friendly couples whose origin or end is y.*

Proof of Lemma 11. Assume first that there is a vertex x in V_1 such that $d_{G_2}(y, x) \geq r$; then the argument leading to (14) still works, and we obtain $r-1$ arcs in $A^r \setminus A$ with origin y and end in V_1. Similarly, if there is a vertex x in V_1 such that $d_{G_2}(x, y) \geq r$, then there exist $r-1$ arcs in $A^r \setminus A$ with end y and origin in V_1.

Finally, we treat the case when for all $x \in V_1$, $d_{G_2}(x, y) < r$ and $d_{G_2}(y, x) < r$. We know that in V_1 there are two vertices at distance in G at least $r+1$ from one another: if we denote them by z_1 and z_2, with $d_G(z_1, z_2) > r$, there is in G_1 a shortest path $x_0 x_1 \dots x_t$ with $x_0 = z_1$, $x_t = z_2$ and $t > r$. Mimicking the proof of Lemma 6, we see that y is the origin or the end of $r-1$ arcs (y, x_i) or (x_i, y) belonging to $A^r \setminus A$, with $x_i \in V_1$, which proves Lemma 11. □

Corollary 12 *For all $h \geq 1$, there are at least $h(r-1)$ distinct arcs in $A^r \setminus A$ with one end or one origin in the ear $Y = \{y_1, y_2, \ldots, y_h\}$ and the other extremity in V_2, i.e., at least $h(r-1)$ G_2-friendly couples with one extremity in Y.* □

The proof of Theorem 5 is now almost complete: in G_2, which has $n_2 = n_1 + h$ vertices, there are two vertices at distance in G at least $r+1$ from one another, and there are at least

$$b(r, n_1) + h(r-1) = b(r, n_1 + h) = b(r, n_2)$$

G_2-friendly couples. If $G_2 = G$, we are done; if not, we can act again on G_2 as we did on G_1, ..., until we eventually reach G, which proves that there are at least $b(r, n)$ arcs in $A^r \setminus A$ and ends the proof of Theorem 5. □

Theorem 5 implies directly Theorem 2: let $G = (V, A)$ be a strongly connected digraph of order n and diameter $\varphi \geq 2$. Then $G^{\varphi-1} \neq K_n$, $|A^{\varphi-1}| \leq n(n-1)-1$, and $|A| \leq |A^{\varphi-1}| - b(\varphi-1, n)$. Calculations show that $n(n-1)-1-b(\varphi-1, n) = \sigma(\varphi, n)$, including the case $\varphi = 2$.

4.2 Characterization

We now characterize the graphs which attain the bound $|A^r \setminus A| = \Lambda(r, n)$; we have already seen at the beginning of the proof of Theorem 5 that the graph G_0 with vertex set $V_0 = \{x_i : 0 \leq i \leq n_0 - 1\}$ and arc set A_0 given by (5) is such that there are exactly $\Lambda(r, n_0)$ arcs in $A_0^r \setminus A_0$.

We consider a strongly connected digraph $G = (V, A)$ of order n, such that $G^r \neq K_n$ and $|A^r \setminus A| = \Lambda(r, n)$; the diameter φ of G is at least $r+1$.

In the process of proving Theorem 5, we considered the graph G^* and, in G^*, a shortest directed path $C = x_0 x_1 \ldots x_t$, $t > r$, from $z_1 = x_0$ to $x_t = z_2$, cf. (6); this path will provide at least $\Lambda(r, t+1)$ G^*-friendly couples, cf. (7). Each vertex in $Y^* = V^* \setminus \{x_i : 0 \leq i \leq t\}$ will bring at least $r-1$ G^*-friendly couples, thanks to property (Q), and all in all G^* will give at least $\Lambda(r, n^*)$ G^*-friendly couples, cf. (13). Then, switching from G^* to G_1, we proved (Corollary 12) that each vertex in $V \setminus V_1$ gives, in average, at least $r-1$ arcs to $A^r \setminus A$, finally leading to at least $\Lambda(r, n)$ arcs in $A^r \setminus A$.

If G attains the bound, then in the previous paragraph, we can replace each occurrence of "at least" by "exactly". In particular, G^* achieves the bound $\Lambda(r, n^*)$ and C achieves the bound $\Lambda(r, t+1)$, for the number of G^*-friendly couples. The following easy lemma will be used repeatedly.

Lemma 13 *If $C = x_0 x_1 \ldots x_t$, the shortest path in G^* from $z_1 = x_0$ to $x_t = z_2$, meets the bound $\Lambda(r, t+1)$, and if $i > j$ and $(x_i, x_j) \in A^r$, then $(x_i, x_j) \in A$.*

Proof. The $\Lambda(r, t+1)$ G^*-friendly couples from (7) are of type x_k, x_ℓ, with $k < \ell$, so no arc (x_i, x_j), $i > j$, can belong to $A^r \setminus A$. □

Note that we are in the conditions of Lemma 13 as soon as G meets the bound $\Lambda(r, n)$. We now show that all vertices of G^* are in C.

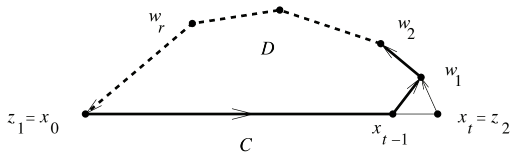

Figure 9: The paths C and D and the path F (in bold).

Lemma 14 *If G achieves the bound $\Lambda(r,n)$, then the set Y^* is empty.*

Proof. Let D be a shortest path in G^* going from x_t to x_0 (such a path exists since G^* is strongly connected). Then all the vertices in Y^* are vertices of D, for otherwise the set of vertices of C and D would violate the minimality of G^*. We now show that D has no vertices outside of C.

Assume the contrary.

If D intersects C in a vertex x_h, $h \notin \{0, t\}$, we consider the couple x_t, x_h. The arc (x_t, x_h) belongs to A^r: the graph consisting of $x_h x_{h+1} \ldots x_t$ and the part of D from x_t to x_h is strongly connected and smaller than G^*, and we use property (P). Therefore, by Lemma 13, $(x_t, x_h) \in A$. Similarly, $(x_h, x_0) \in A$, and, because $r \geq 2$, (x_t, x_0) is in A^r, hence in A, which yields a path shorter than D.

Assume now that D does not intersect C, except on x_0 and x_t.

If D, apart from x_t and x_0, has at most $r-1$ vertices, then the distance in D from x_t to x_0 is at most r, from which we can conclude that $(x_t, x_0) \in A$, again a contradiction.

So we assume that D, apart from x_t and x_0, has at least r vertices, $w_1, w_2, \ldots,$ $w_r, \ldots,$ with (x_t, w_1) the first arc in D. Then $(w_1, w_i) \in A^r \setminus A$ for $3 \leq i \leq r$, and the same is true with either (w_1, w_{r+1}) or (w_1, x_0); by property (Q), which must be satisfied with equality, there is no other arc in $A^r \setminus A$ having w_1 as an extremity; in particular, because $r \geq 2$, the arc (x_{t-1}, w_1) must belong to A, see Figure 9. It follows that the path F which goes from x_0 to x_0 using the shortcut (x_{t-1}, w_1) is strongly connected and smaller than G^*, so we can conclude that $(x_{t-1}, x_0) \in A$ by Lemma 13, since $(x_{t-1}, x_0) \in A^r$ by (P).

Let y be any vertex in $D \setminus \{x_0, x_t\}$. If there is an arc (y, x_i) in A with $1 \leq i \leq t-1$, see Figure 10, the same argument with $x_t w_1 \ldots y x_i x_t$ shows that $(x_t, x_i) \in A$. Then we can see in Figure 11 a path going from x_t to x_i, then to x_{t-1}, then to x_0, that is, a path from x_t to x_0 which uses only vertices in C; this contradicts the minimality of G^*.

So we can assume that there is no arc (y, x_i), $1 \leq i \leq t-1$, in A. Using again the property of the path F, this means however that all the arcs (y, x_i), $1 \leq i \leq t-1$, belong to $A^r \setminus A$, which represents more than $r-1$ arcs, since $t > r$. This contradiction completes the proof of Lemma 14. $\square$

Thus, G^* is made of the path C, plus some arcs (x_i, x_j), $i > j$, which make G^* strongly connected. We now show that G^* contains all the arcs (x_i, x_j), $i > j$.

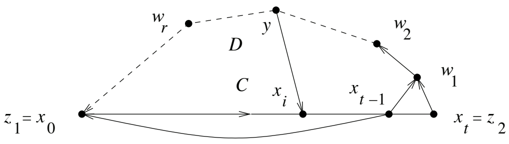

Figure 10: The paths C and D and the arc (y, x_i).

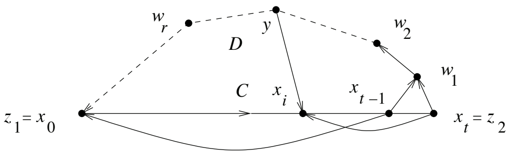

Figure 11: A new path from x_t to x_0.

Let D still be a shortest path in G^* from x_t to x_0. If $D \neq x_t x_0$, let (x_t, x_h), $h \neq 0$, be the first arc in D. Let us consider, in D, the next arc which goes "from right to left", that is, which reads (x_j, x_k) with $k < j$ and $j \geq h$: see Figure 12 for an example where $j > h$. Because D is a shortest path, we have $k < h$.

Since the path $x_h \ldots x_j \ldots x_t x_h$ has fewer vertices than G^*, all its vertices are within distance r from each other, and as before, using Lemma 13, we can conclude that $(x_t, x_i) \in A$ for $h \leq i \leq t-1$. But if $(x_t, x_i) \in A$ for $h < i \leq j$, then $x_t x_i \ldots x_j x_k$ yields a path from x_t to x_0 shorter than D ; so $h = j$, and (x_t, x_h) and (x_h, x_k) belong to A, which in turn implies that $(x_t, x_k) \in A^r$, because $r \geq 2$, and $(x_t, x_k) \in A$ by Lemma 13, again yielding a path shorter than D. Therefore we have shown that $D \neq x_t x_0$ is impossible: actually, $(x_t, x_0) = (z_2, z_1) \in A$.

This implies that $d_G(x_t, x_1) \leq 2$, so, still using $r \geq 2$ and Lemma 13, $(x_t, x_1) \in A$, and step by step, $(x_t, x_j) \in A$ for $0 \leq j \leq t-1$. Similarly, $(x_i, x_j) \in A$ for $0 \leq j < i \leq t-1$, and we have proved the following result.

Lemma 15 *If G achieves the bound $\Lambda(r, n)$, then G^* has vertex set $V^* = \{x_i : 0 \leq i \leq n^* - 1\}$ and arc set given by (5) with $n^* = t + 1 = n_0$.* □

Lemma 16 *If G achieves the bound $\Lambda(r, n)$, then $d_G(z_1, z_2) = d_{G^*}(z_1, z_2)$.*

Proof. When we introduced z_1, z_2, G^* and C, we remarked that if $\theta = d_G(z_1, z_2)$ and $t = d_{G^*}(z_1, z_2)$, then obviously $t \geq \theta$. Assume that $t > \theta$ and consider a shortest

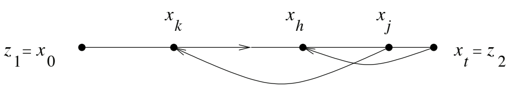

Figure 12: Going from x_t to x_0.

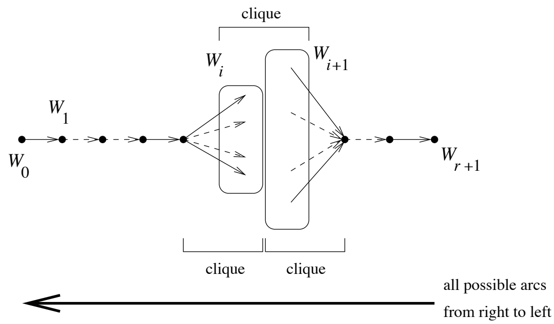

Figure 13: Type 1 digraphs. All arcs going from right to left exist.

path in G from z_1 to z_2: $J = w_0w_1 \dots w_\theta$, with $w_0 = x_0 = z_1$ and $z_2 = x_t = w_\theta$. Then, because we have just shown that $(x_t, x_0) \in A$, the subgraph induced by J is strongly connected, so that it contradicts the minimality of G^*. $\square$

We are now ready to describe the graphs which meet the bound $\Lambda(r, n)$. We say that a graph $G = (V, A)$ is a *directed path* $(W_0, W_1, \dots, W_q)$ *of cliques* if it meets the following conditions:

- the sets $W_0, W_1, \dots, W_q$ partition V;
- for i between 1 and q, the subgraph of G induced by $W_{i-1} \cup W_i$ is a clique;
- for i and j between 0 and q, $i \leq j-2$, there is no arc from W_i to W_j in A, and all arcs from W_j to W_i belong to A.

Note that a graph meeting these conditions has diameter q, and that if $y \in W_i$ and $z \in W_j$, $i < j$, then $d_G(y, z) = j - i$ and $d_G(z, y) = 1$.

Next, we define graphs of type 1 and of type 2 in the following way (see Figures 13 and 14):

- a digraph $G = (V, A)$ is of type 1 if it is a directed path $(W_0, W_1, \dots, W_{r+1})$ of cliques such that if, for $0 \leq i < j \leq r+1$, one has $|W_i| \geq 2$ and $|W_j| \geq 2$, then $j = i+1$; in other words, there are at most two values of i such that $|W_i| \geq 2$, and if they exist, these two values are consecutive. Moreover, we ask that:

$$\text{at least one set } W_i \text{ is not a singleton;} \tag{18}$$

$$\text{if } |W_0| \geq 2, \text{ then } |W_1| \geq 2; \text{ if } |W_{r+1}| \geq 2, \text{ then } |W_r| \geq 2; \tag{19}$$

- a digraph $G = (V, A)$ is of type 2 if it is a directed path $(W_0, W_1, \dots, W_q)$ of cliques with $|W_i| = 1$ for $2 \leq i \leq q-2$.

Moreover, we ask that if $q = r+2$, then $|W_1| = 1$ or $|W_{q-1}| = 1$, and that if $q = r+1$, then $|W_1| = |W_{q-1}| = 1$.

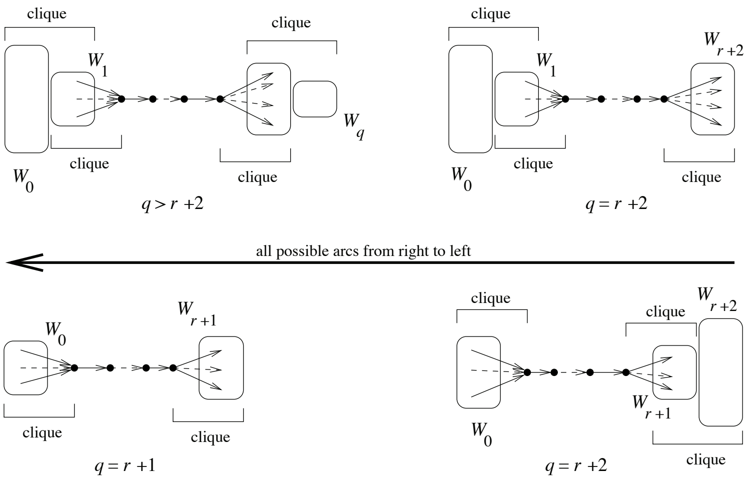

Figure 14: Type 2 digraphs. All arcs going from right to left exist.

Note that, thanks to (18), (19), a graph of type 1 cannot be of type 2, and thus cannot simply be a graph of the type defined by (4), (5) (a shortest path from left to right and all arcs from right to left), which is of type 2.

The digraphs of type 1 and 2 just described are very similar to the undirected graphs of type 1 and 2 depicted in [2, Th. 8].

Theorem 17 *Let $G = (V, A)$ be a connected digraph of order n such that $G^r \neq K_n$. Then $|A^r \setminus A| = \Lambda(r, n)$ if and only if G is of type 1 or of type 2.*

Proof. In the course of this proof, we shall prove and use Lemmas 18–20.

First, it has to be checked that these digraphs do satisfy $|A^r \setminus A| = \Lambda(r, n)$. This can be seen using the following argument: first, in all cases there are at least r singleton sets W_i; second, if all sets W_i are singletons, i.e., we have the graph G_0 defined by (4) and (5), then we know that the graph meets $\Lambda(r, n_0)$, where n_0 is the order of G_0; third, we observe that, starting from G_0, if we add one by one the vertices y belonging to the sets W_j which are not singletons, each vertex y brings exactly $r-1$ new arcs in $A^r \setminus A$, arcs (y, z) or (z, y) according to the position of the non singleton set it belongs to, with $W_i = \{z\}$ and $r \geq |i - j| \geq 2$; these arcs are counted only once; finally, we use that $\Lambda(r, n)$ is linear in n, with the factor $r - 1$: we obtain

$$\Lambda(r, n_0) + (n - n_0)(r - 1) = \Lambda(r, n)$$

edges.

Next, we consider a graph $G = (V, A)$ of order n and diameter φ which meets the bound $\Lambda(r, n)$; thanks to Lemmas 15 and 16, we know that there exists an induced subgraph G^* of G with the following properties: G^* is of type 1 or 2, and is such that $r + 1 \leq d_G(x_0, x_t) = d_{G^*}(x_0, x_t) = t \leq \varphi$, with the notation of Lemma 15.

This means that the set of induced subgraphs of G, of type 1 or 2, with vertex set $W_0 \cup W_1 \cup \cdots \cup W_q$, $r+1 \leq q \leq \varphi$, such that

$$\forall y \in W_0,\ \forall z \in W_q,\ d_{G^*}(y,z) = d_G(y,z) = q, \tag{20}$$

is a nonempty set, and we can take one, $G_1 = (V_1, A_1)$, with maximum order, n_1. In particular, (20) implies that

$$\text{for all } y \in W_i,\ z \in W_j,\ 0 \leq i < j \leq q, \quad d_{G_1}(y,z) = d_G(y,z) = j - i. \tag{21}$$

If $G_1 = G$, we are done, so we assume that $V \setminus V_1 \neq \emptyset$. As in the proof of Theorem 5, there is a smallest ear $Y = \{y_1, y_2, \ldots, y_h\} \subseteq V \setminus V_1$ such that $H = y_1 y_2 \ldots y_h$ is a directed path in G and the arcs (u, y_1), (y_h, v) are in A, where u and v are two, non necessarily distinct, vertices in V_1; cf. Figure 8.

Lemma 18 *There is a smallest ear such that*
- *either $u \in W_i$, $v \in W_j$ and $i \neq j$,*
- *or $u = v$.*

Proof of Lemma 18. If we can choose one vertex x_i in each set W_i, $0 \leq i \leq q$, such that $u = x_j \in W_j$ and $v = x_k \in W_k$ for some j, k in $\{0, 1, \ldots, q\}$, see Figure 15 where $u = v$ is possible, we are done; if not, this means that u and v belong to the same set W_{i_0} and are distinct, in which case we choose $u = x_{i_0}$, see Figure 16. We show that we can come down to the former case: if in G there is the arc (y_h, u) or an arc (y_h, x_ℓ) with $x_\ell \in W_\ell$, $\ell \neq i_0$, then we simply rename v: $v = u = x_{i_0}$ or $v = x_\ell$. So we assume that $(y_h, x_i) \notin A$ for $i = 0, 1, \ldots, q$. For $i = 0, 1, \ldots, i_0 + 1$, we have $d_G(y_h, x_i) = 2$, which implies that $(y_h, x_i) \in A^r \setminus A$. If $i_0 + 2 > r - 1$, we get a contradiction with the proof of Corollary 10 or with Lemma 11 when G achieves the bound $\Lambda(r, n)$, so we can assume that $i_0 \leq r - 3$. For $i = i_0 + 2, i_0 + 3, \ldots, \min\{q, i_0 + r - 1\}$, we have $d_G(y_h, x_i) \leq r$, so the $\min\{q - i_0 - 1, r - 2\}$ arcs (y_h, x_i), $i \geq i_0 + 2$, belong to $A^r \setminus A$, and all in all, we have $\min\{q + 1, r + i_0\} > r - 1$ arcs in $A^r \setminus A$. This is again a contradiction, and Lemma 18 is proved. □

Back to the proof of Theorem 17, we consider from now on a smallest ear with vertices $u = x_j$ and $v = x_k$ in (possibly equal) sets W_j and W_k. Note that, if y_2 exists, there is no arc with origin in V_1 and end y_2. We call C the path $x_0 x_1 \ldots x_q$, which is a shortest path from x_0 to x_q, in G_1 as well as in G.

We are going to show that the ear Y contains only one vertex; suppose on the contrary that $h \geq 2$.

Lemma 8 and the proof of Corollary 10 show that y_1 is the origin of $r - 1$ arcs belonging to $A^r \setminus A$, and, consequently, that, if G achieves the bound $\Lambda(r, n)$, then y_1 cannot be the end of any arc in $A^r \setminus A$.

Since $(x_j, y_1) \in A$, we have $(x_{j-1}, y_1) \in A^r$ and therefore $(x_{j-1}, y_1) \in A$; step by step, we obtain that for all i between 0 and $j - 1$, $(x_i, y_1) \in A$. For $i > j$, the arc (x_i, x_j) is in A, so as before we must have $(x_i, y_1) \in A$: we have just proved that all arcs (x_i, y_1), $0 \leq i \leq q$, are in A, which implies that $d_G(x_i, y_2) \leq 2$. So $(x_i, y_2) \in A^r$, and we have observed that $(x_i, y_2) \notin A$. This represents $q + 1 > r - 1$ arcs in $A^r \setminus A$

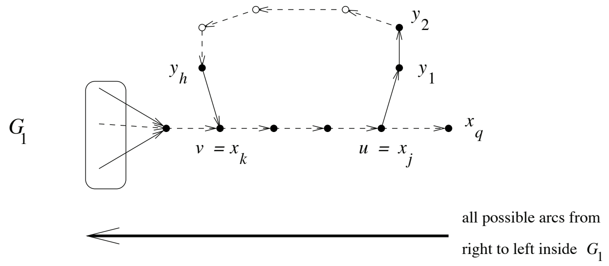

Figure 15: The ear Y: $u = x_j$, $v = x_k$.

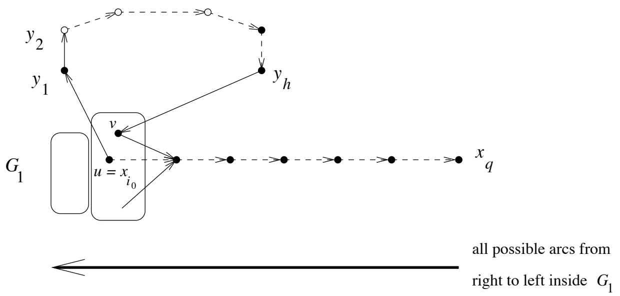

Figure 16: The ear Y: $u = x_{i_0}$, $v \in W_{i_0}$.

with origin in V_1 and end y_2, which is more than stated in Lemma 9, and so the bound cannot be achieved: the assumption $h \geq 2$ led to a contradiction, and we have, setting $y = y_1$, $Y = \{y\}$. By Lemma 11 and to satisfy the bound with equality, we have in $A^r \setminus A$ exactly $r-1$ arcs with end or origin y.

We denote by $G_2 = (V_2, A_2)$ the induced subgraph of G with vertex set $V_2 = V_1 \cup \{y\}$. Knowing that $G_1 = (V_1, A_1)$ is a path of cliques $(W_0, W_1, \dots, W_q)$ of type 1 or 2, satisfying (21) with $q \geq r+1$, we are going to show that G_2 is a sum of cliques of type 1 or 2, also satisfying (21), with:

- $G_2 = (W_0, \dots, W_{i_0}, \dots, W_q)$ and $y \in W_{i_0}$;
- or, with abuse of notation, $G_2 = (W_{-1} = \{y\}, W_0, \dots, W_q)$;
- or $G_2 = (W_0, \dots, W_q, W_{q+1} = \{y\})$.

Let h be the smallest subscript such that $(x_h, y) \in A$ and k be the largest subscript such that $(y, x_k) \in A$; then in $A^r \setminus A$, there are the $\min\{r-1, h\}$ arcs $(x_{h-1}, y), (x_{h-2}, y), \dots$, and the $\min\{r-1, q-k\}$ arcs $(y, x_{k+1}), (y, x_{k+2}), \dots$

Therefore, in $A^r \setminus A$, we have at least

$$\Gamma = \min\{r-1, h\} + \min\{r-1, q-k\} \tag{22}$$

arcs with one extremity in C and one extremity on y, and so $\Gamma \leq r-1$.

Lemma 19 *(a)* $k \leq h+2$.

(b) The following three cases, which are not excluding each other, are the only possible cases:

- $q = r+1$ *and* $k = h+2$*;*
- $h = 0$ *and* $k \in \{0, 1, 2\}$*;*
- $k = q$ *and* $h \in \{q-2, q-1, q\}$.

(c) Every arc in $A^r \setminus A$ *with one extremity on* y *has its other extremity in* $\{x_0, x_1, \dots, x_q\}$.

(d)
- *If* $i \geq h$ *and* $w \in W_i$*, then* $(w, y) \in A$*; in particular,* $(x_i, y) \in A$.
- *If* $i \leq k$ *and* $w \in W_i$*, then* $(y, w) \in A$*; in particular,* $(y, x_i) \in A$.

Proof of Lemma 19. (a) Since C is a shortest path, both in G_1 and in G, we cannot have shortcuts going through y, which means that the existence of (x_h, y) in A implies the nonexistence in A of $(y, x_{h+3}), (y, x_{h+4}), \dots$, and we have $k \leq h+2$.

(b, c) Examining Γ in (22), we see that four cases are possible:
(i) $h \leq r-1$ and $q-k \leq r-1$. Then $\Gamma = q + (h-k)$ and, because $\Gamma \leq r-1$, $q \geq r+1$ and $k-h \leq 2$, we must have: $q = r+1$, $k = h+2$, and $\Gamma = r-1$ counts exactly all the arcs in $A^r \setminus A$ with one extremity on y: these are the h arcs (x_{h-1}, y), $(x_{h-2}, y), \dots, (x_0, y)$, plus the $q-k$ arcs (y, x_{k+1}), (y, x_{k+2}), $\dots$, (y, x_q).
(ii) $h \leq r-1$ and $q-k \geq r$. Then $\Gamma = h+r-1$ and $h = 0$, $k \leq 2$, and $\Gamma = r-1$ counts exactly all the arcs in $A^r \setminus A$ with one extremity on y, which here are the arcs (y, x_{k+1}), (y, x_{k+2}), $\dots$, (y, x_{k+r-1}).

(iii) $h \geq r$ and $q-k \leq r-1$. Then $\Gamma = r-1+q-k$ and $q=k$, $h \geq q-2$; again, $\Gamma = r-1$ counts exactly all the arcs in $A^r \setminus A$ with one extremity on y. This time, these are the arcs (x_{h-1}, y), (x_{h-2}, y), $\ldots$, (x_{h-r+1}, y).
(iv) $h \geq r$ and $q-k \geq r$. This case is impossible, since this would imply that $\Gamma = 2(r-1)$, so $\Gamma > r-1$ when $r \geq 2$.

(d) Since in G_1 all arcs (z, y) exist for $z \in W_j$, $y \in W_i$ and $i < j$, we have for all $w \in W_i$, $i \geq h$: $d_G(w, y) \leq d_{G_1}(w, y) \leq d_{G_1}(w, x_h) + d_{G_1}(x_h, y) \leq 2$. Therefore, $(w, y) \in A^r$, and we have just seen, in the three cases (i)–(iii) of the previous step, that (w, y) does not appear in $A^r \setminus A$ when $i \geq h$, hence $(w, y) \in A$. Finally, if $w \in W_i$, $i \leq k$, then $d_G(y, w) \leq d_{G_1}(y, w) \leq d_{G_1}(y, x_k) + d_{G_1}(x_k, w) \leq 2$, and, exactly as above, this leads to $(y, w) \in A^r$ with $(y, w) \notin A^r \setminus A$, and ultimately $(y, w) \in A$.

This ends the proof of Lemma 19. □

Back to the proof of Theorem 17, we give the following definition of i_0, which covers all cases of Lemma 19(b) and will be used in Lemma 20; note that i_0 depends on y through h and k.

- if $h = k = 0$, then $i_0 = -1$;
- if $h = 0$, $k = 1$, then $i_0 = 0$;
- if $k - h = 2$, then $i_0 = h + 1 = k - 1$;
- if $h = q - 1$, $k = q$, then $i_0 = q$;
- if $h = k = q$, then $i_0 = q + 1$.

Lemma 20 *Let $w \in W_i$, $0 \leq i \leq q$.*

(1) If $i \leq i_0 - 1$, then $(y, w) \in A$.
(2) If $i \geq i_0 + 1$, then $(w, y) \in A$.
(3) If $i_0 \notin \{q, q+1\}$ and $i \leq i_0 + 1$, then $(y, w) \in A$.
(4) If $i_0 \notin \{-1, 0\}$ and $i \geq i_0 - 1$, then $(w, y) \in A$.
(5) If $-1 \leq i_0 < i \leq q$, then $d_G(y, w) = i - i_0$.
(6) If $0 \leq i < i_0 \leq q + 1$, then $d_G(w, y) = i_0 - i$.
(7) If $0 \leq i_0 = i \leq q$, then $(y, w) \in A$ and $(w, y) \in A$.
(8) If $|W_i| \geq 2$ and $i - i_0 \geq 2$, then $i - i_0 \geq r + 1$.
(9) If $|W_i| \geq 2$ and $i_0 - i \geq 2$, then $i_0 - i \geq r + 1$.

Proof of Lemma 20. (1,2) The definition of i_0 shows that in all cases, $k \geq i_0 - 1$ and $h \leq i_0 + 1$, and we simply apply Lemma 19(d).

(3,4) We apply the same argument as above, noting that if $i_0 \notin \{q, q+1\}$, then $k = i_0 + 1$, and if $i_0 \notin \{-1, 0\}$, then $h = i_0 - 1$.

(5) If $i = i_0 + 1$, then by Lemma 20(3), we have immediately $(y, w) \in A$, i.e., $d_G(y, w) = 1 = i - i_0$. So from now on, we assume that $i - i_0 \geq 2$. Now

$$d_G(y, w) \leq d_G(y, x_{i_0+1}) + d_G(x_{i_0+1}, w) \leq 1 + (i - i_0 - 1) = i - i_0, \qquad (23)$$

thanks to (21). We distinguish between three cases, $i_0 \geq 1$, $i_0 = 0$ and $i_0 = -1$.

We first assume that $i_0 \geq 1$. We have

$$d_G(x_0, y) \leq d_G(x_0, x_{i_0-1}) + d_G(x_{i_0-1}, y) \leq (i_0 - 1) + 1 = i_0, \qquad (24)$$

using (21) and Lemma 20(4). Since $d_G(x_0,w)=i$, one has $i \le d_G(x_0,y)+d_G(y,w) \le i_0+(i-i_0)=i$, by (24), (23), so that $d_G(x_0,y)=i_0$ and $d_G(y,w)=i-i_0$, which is what we wanted.

Next, we assume that $i_0=0$, so that $i \ge 2$ and $h=0$, $k=1$. Let C be a shortest path in G from y to w, with length $\ell(C)$; then $\ell(C) \le i-i_0=i$ by (23), and $\ell(C) \ge d_G(x_0,w)-d_G(x_0,y)$, which is equal to $i-1$, by (21) and Lemma 19(d). If C does not go through the vertex x_2, we set

$$\widehat{C}=C \text{ if } i=q \text{ and } \widehat{C}=Cx_{i+1}\dots x_q=y\dots wx_{i+1}\dots x_q \text{ if } i<q;$$

this new path has length $\ell(C)+(q-i) \ge (i-1)+(q-i)=q-1 \ge r$. Because C is a shortest path, y belongs to only one arc in C, and because $k=1$, none of the arcs $(y,x_{i+1}),\dots,(y,x_q)$ belongs to A. Consequently, there exist at least $r-1$ arcs in $A^r \setminus A$, originating in y and ending in $\widehat{C}$. Moreover, (y,x_2) belongs to A^r (because $k=1$, $(y,x_1)\in A$, $(x_1,x_2)\in A$ and $r\ge 2$) and $(y,x_2)\notin A$ (because $k=1$), and this contradicts the assumption that G meets the bound $\Lambda(r,n)$. So from now on, we can assume that C goes through x_2; in this case, since $(y,x_2)\notin A$, we see that $\ell(C) \ge 2+d_G(x_2,w)=2+(i-2)=i$, which means that the distance from y to w is exactly $i=i-i_0$ in G, which is what we wanted.

Finally, we consider the case $i_0=-1$, which implies that $i\ge 1$, $h=k=0$. In this case, $d_G(y,w)\le i+1$ by (23) and $d_G(y,w)\ge d_G(x_0,w)-d_G(x_0,y)=i-1$ because $h=0$. We proceed as in the previous case: if C is a shortest path from y to w, not going through x_1, we extend it to $\widehat{C}$, which has length at least r. This provides more than $r-1$ arcs in $A^r\setminus A$ with extremity on y, including the arc (y,x_1). So C goes through x_1, $\ell(C)\ge 2+d_G(x_1,w)=2+(i-1)=i+1$, and the distance from y to w is exactly $i+1=i-i_0$.

(6) This case is symmetric to the previous case (5).

(7) If $k=i_0$ or $k=i_0+1$, then $i=i_0\le k$ and by Lemma 19(d), $(y,w)\in A$; similarly, if $h=i_0$ or $h=i_0-1$, then $i=i_0\ge h$ and $(w,y)\in A$. Due to the definition of i_0 showing that $\{h,k\}\subset\{i_0-1,i_0,i_0+1\}$, the only cases that we have to investigate are when $k=i_0-1$ and when $h=i_0+1$. But $k=i_0-1$ implies $i_0=q+1$, and $h=i_0+1$ implies $i_0=-1$, a contradiction.

(8) We can assume that $w\ne x_i$: otherwise, we consider another vertex in W_i. Since $i-i_0\ge 2$ and $d_G(y,w)=i-i_0$, we have $d_G(y,w)\ge 2$, and $(y,w)\notin A$. Lemma 19(c) then implies that $(y,w)\notin A^r$, and so $d_G(y,w)=i-i_0\ge r+1$.

(9) This case is symmetric to the previous case (8) and Lemma 20 is proved. □

We return to the proof of Theorem 17: starting from the graph G_1 which is a path of cliques $(W_0,W_1,\dots,W_q)$, of type 1 or 2, satisfies (21) and has maximum order n_1, we are led to add one vertex y, following the different conditions on h and k. Case by case, we are going to check, using Lemmas 19 and 20, that the resulting graph G_2 is necessarily still of type 1 or 2 and satisfies (21), contradicting the maximality of n_1.

There are five cases, according to the values of i_0, defined just before Lemma 20. In each case we say that $y\in W_{i_0}$ and check, in a very straightforward way, that G_2 has the desired properties.

• If $h = k = 0$ and $i_0 = -1$, we add a set $W_{-1} = \{y\}$. Then, by Lemma 20(2), all arcs from the sets W_i, $i \geq 0$, to y are in A. By Lemma 20(5), $d_G(y, w) = i - i_0$, so the only arcs in A which originate in y go to the vertices in W_0, and property (21) is preserved in G_2 (for $-1 \leq i < j \leq q$). Finally, by Lemma 20(8), the only sets W_i which can have size at least 2 in G_1 are W_0 and W_i, $i \geq r$.

We can see that either G_1 is of type 1 with, using (18) and (19), $|W_0| = 1$, $|W_r| \geq 2$ and $|W_{r+1}| \geq 1$, or G_1 is of type 2 (with $|W_1| = 1$). In both cases, the addition of y yields a graph of type 2 with $q+2 \geq r+3$ sets W_i (starting, with abuse of notation, at W_{-1}).

• The case $h = k = q$ and $i_0 = q + 1$ is symmetric.

• If $h = 0$, $k = 1$ and $i_0 = 0$, we put y into W_0. Then, by Lemma 20(2,7), all arcs from the sets W_i, $i \geq 0$, to y are in A, plus all arcs from y to the vertices of W_0; by Lemma 20(5,7), the only arcs in A which originate in y go to the vertices in W_0 and W_1; also, property (21) is preserved in G_2. Finally, Lemma 20(8) shows that the only sets W_i which can have size at least 2 in G_1 are W_1 and W_i, $i \geq r + 1$.

If G_1 is of type 1, then by (18) and (19), $|W_1| \geq 2$, and the addition of y to W_0 yields a graph of type 1 with $|W_0| \geq 2$, $|W_1| \geq 2$ and $q = r + 1$. If G_1 is of type 2, then G_2 simply has one additional element in W_0, and remains of type 2, with the same number, $q + 1$, of sets W_i.

• The case $h = q - 1$, $k = q$ and $i_0 = q$ is symmetric.

• Finally, if $k - h = 2$ and $1 \leq i_0 = h + 1 = k - 1 \leq q - 1$, we put y into W_{i_0}. By Lemma 20(1,2,7), all arcs from the sets W_i, $i \geq i_0$, to y are in A, and all arcs from y to the sets W_i, $i \leq i_0$, are in A. By Lemma 20(5,6), the arcs from W_{i_0-1} to y and from y to W_{i_0+1} are in A, there are no more arcs with extremity y in A, and property (21) still holds in G_2. Besides, by Lemma 20(8,9), the only sets W_i which can have size at least 2 in G_1 are W_{i_0-1}, W_{i_0-r-1}, W_{i_0-r-2}, $\ldots$, and W_{i_0+1}, W_{i_0+r+1}, W_{i_0+r+2}, $\ldots$

If G_1 is of type 1, then $q = r + 1$ and $1 \leq i_0 \leq r$, implying that the only sets W_i which can have size at least 2 in G_1 are W_{i_0-1} and W_{i_0+1}, and consequently exactly one of them is not a singleton. Together with $W_{i_0} \cup \{y\}$, this gives exactly two consecutive sets of size at least 2, showing that G_2 is of type 1, with $q + 1$ sets W_i. So we assume that G_1 is of type 2.

If $q = r + 1$, as before the only sets W_i which can have size at least 2 in G_1 are W_{i_0-1} and W_{i_0+1}, and consequently at most one is not a singleton. If none of them has size at least 2, then $W_{i_0} \cup \{y\}$ is the only non singleton, and G_2 is of type 1, since $1 \leq i_0 \leq q - 1$. If $W_{i_0-1} = W_0$ (respectively, $W_{i_0+1} = W_{r+1}$) is the only non singleton in G_1, then we have in G_2 exactly two, consecutive, sets of size at least 2, W_0 and W_1 (respectively, W_r and W_{r+1}), and again G_2 is of type 1.

From now on, we assume that $q > r + 1$. By Lemma 19(b), this, plus the assumption $k - h = 2$, implies that $h = 0$, $k = 2$ or $k = q$, $h = q - 2$, and so $i_0 = 1$ or $i_0 = q - 1$.

If $q = r + 2$, then $i_0 = 1$ or $i_0 = r + 1$, and by Lemma 20(8,9), the only sets W_i which can have size at least 2 in G_1 are W_0, W_2, W_r and W_{r+2}. But since G_1 is of type 2, we can have only $|W_0|, |W_1|, |W_{r+2}| \geq 2$ or $|W_0|, |W_{r+1}|, |W_{r+2}| \geq 2$, so that only $|W_0| \geq 2$ and $|W_{r+2}| \geq 2$ are possible. After addition of y in W_{i_0}, the only possible non singletons in G_2 are $|W_0|, |W_1|, |W_{r+2}|$ for $i_0 = 1$, and $|W_0|, |W_{r+1}|, |W_{r+2}|$ for $i_0 = r + 1$, so G_2 is of type 2 (with $q + 1$ sets W_i).

If $q > r + 2$, then the only sets W_i of size at least 2 in G_1 are $|W_0|$, $|W_1|$, $|W_{q-1}|$, and $|W_q|$, and adding y to W_1 or to W_{q-1} does not change anything.

So, starting from an induced subgraph of G, $G_1 \neq G$, which was assumed to have maximum order, we exhibited an induced subgraph of G with one more vertex and same properties, G_2. This ends the proof of Theorem 17. $\square$

Acknowledgment

We wish to thank the two referees for pointing out the reference [5] to us.

References

[1] D. Aingworth, R. Motwani and F. Harary, The difference between a graph and its square, *Utilitas Mathematica* 54, (1998), 223–228.

[2] D. Auger, I. Charon, O. Hudry and A. Lobstein, On the sizes of the graphs G, G^r, $G^r \setminus G$: the undirected case, *Discrete Applied Math.* (to appear).

[3] J. Bang-Jensen and G. Gutin, *Digraphs: Theory, Algorithms and Applications*, Springer, 2001.

[4] C. Berge, *Graphes*, Paris: Gauthier-Villars, 1983.

[5] A. Ghouila-Houri, Un résultat relatif à la notion de diamètre, *Comptes-Rendus de l'Académie des Sciences*, (1960), 4254–4256.

[6] O. Ore, Diameters in graphs, *J. Combin. Theory* 5, (1968), 75–81.

(Received 8 Oct 2009; revised 5 May 2010)