Periods in Partial Words: An Algorithm

By: F. Blanchet-Sadri, Travis Madel, Gautam Sisodia

Blanchet-Sadri, T., Mandel, T., Sisodia, G. (2012). Periods in Partial Words: An Algorithm. *Journal of Discrete Algorithms, 16,* 113-128. doi: 10.1016/j.jda.2012.04.001

Abstract:

Partial words are finite sequences over a finite alphabet that may contain some holes. A variant of the celebrated Fine–Wilf theorem shows the existence of a bound L=L(h,p,q) such that if a partial word of length at least L with h holes has periods p and q, then it also has period gcd(p,q). In this paper, we associate a graph with each p - and q -periodic word, and study two types of vertex connectivity on such a graph: modified degree connectivity and r -set connectivity where $r = q \bmod p$. As a result, we give an algorithm for computing L(h,p,q) in the general case and show how to use it to derive the closed formulas.

Keywords: Automata and formal languages | Combinatorics on words | Partial words | Fine and Wilf's theorem | Strong periods | Graph connectivity | Optimal lengths

Article:

1. Introduction

The problem of computing periods in *words*, or finite sequences of symbols from a finite alphabet, has important applications in several areas including data compression, coding, computational biology, string searching and pattern matching algorithms. Repeated patterns and related phenomena in words have played over the years a central role in the development of combinatorics on words [5], and have been highly valuable tools for the design and analysis of algorithms. In many practical applications, such as DNA sequence analysis, repetitions admit a certain variation between copies of the repeated pattern because of errors due to mutation, experiments, etc. Approximate repeated patterns, or repetitions where errors are allowed, are playing a central role in different variants of string searching and pattern matching problems [13]. *Partial words* , or finite sequences that may contain some holes, have acquired importance in this context. A (*strong*) *period* of a partial word u over an alphabet A is a positive integer p such that $u(i)=u(j)$ whenever $u(i),u(j)\in A$ and $i \equiv j \bmod p$ (in such a case, we call u *p-periodic*). In other words, p is a period of u if for all positions i and j congruent modulo *p,* the letters in these positions are the same or at least one of these positions is a hole. For example, the word aabaabaa has period 3 but not 4, while the partial word a⋄⋄aabaa, with holes at positions 1 and 2, has periods 3 and 4 (note that our words are starting at position 0 rather than 1).

There are many fundamental results on periods of words. Among them is the well-known periodicity result of Fine and Wilf [8], which determines how long a p - and q -periodic word needs to be in order to also be $\gcd(p,q)$-periodic. More precisely, any word having two periods p,q and length at least $p+q-\gcd(p,q)$ has also $\gcd(p,q)$ as a period. Moreover, the length $p+q-\gcd(p,q)$ is optimal since counterexamples can be provided for shorter lengths, that is, there exists an *optimal* word of length $p+q-\gcd(p,q)-1$ having p and q as periods but not having $\gcd(p,q)$ as period [5]. Extensions of Fine and Wilf's result to more than two periods have been given. For instance, in [6], Constantinescu and Ilie give an extension for an arbitrary number of periods and prove that their lengths are optimal.

Fine and Wilf's result has been generalized to partial words [1], [2], [3], [10], [11], [12] and [14]. Some of these papers are concerned with *weak* periodicity, a notion not discussed in this paper (a *weak period* of a partial word u over an alphabet A is a positive integer p such that $u(i)=u(i+p)$ whenever $u(i),u(i+p)\in A$). The papers that are concerned with strong periodicity refer to the basic fact, proved by Shur and Konovalova (Gamzova) in [12], that for positive integers h , p and q , there exists a positive integer l such that a partial word u with h holes, two periods p and q , and length at least l has period $\gcd(p,q)$. The smallest such integer is called the optimal length and it will be denoted by $L(h,p,q)$. They gave a closed formula for the case where $h=2$ (the cases $h=0$ or $h=1$ are implied by the results in [8] and [1]), while in [11], they gave a formula in the case where $p=2$ as well as an optimal asymptotic bound for $L(h,p,q)$ in the case where h is "large." In [3], Blanchet-Sadri et al. gave closed formulas for the optimal lengths when q is "large," whose proofs are based on connectivity in the so-called (p,q)-periodic graphs. The (p,q)-*periodic graph of size* l is the graph $G=(V,E)$, with $V=\{0,1,\ldots,l-1\}$, such that $\{i,j\}\in E$ if and only if $i \equiv j \bmod p$ or $i \equiv j \bmod q$.

In this paper, we study two types of vertex connectivity in the (p,q)-periodic graphs: the modified degree connectivity and r -set connectivity where $r = q \bmod p$. Although the graph-theoretical approach is not completely new, our paper gives insights into periodicity in partial words and provides an algorithm for determining $L(h,p,q)$ in *all* cases. Our paper also shows how the closed formulas can be derived from our methods.

We end this section by reviewing basic concepts on partial words. Fixing a nonempty finite set of letters or an *alphabet A,* finite sequences of letters from A are called (*full*) *words* over A .

The number of letters in a word *u* , or *length* of *u* , is denoted by $|u|$. The unique word of length 0, denoted by ε , is called the *empty* word. A word of length *n* over *A* can be defined by a total function $u:\{0,\dots,n-1\}\to A$ and is usually represented as $u=a_0a_1\dots a_{n-1}$ with $a_i\in A$. The set of all words over *A* of finite length (greater than or equal to zero) is denoted by A^{*}. A *partial word u* of length *n* over *A* is a partial function $u:\{0,\dots,n-1\}\to A$. For $0\leqslant i<n$, if $u(i)$ is defined, then *i* belongs to the *domain* of *u* , denoted by $i\in D(u)$, otherwise *i* belongs to the *set of holes* of *u* , denoted by $i\in H(u)$. The set of distinct letters of *A* occurring in *u* is denoted by $\alpha(u)$. For convenience, we will refer to a partial word over *A* as a word over the enlarged alphabet $A_\diamond=A\cup\{\diamond\}$, where $\diamond\notin A$ represents a "do not know" symbol or hole. So a partial word *u* of length *n* over *A* can be viewed as a total function $u:\{0,\dots,n-1\}\to A_\diamond$ where $u(i)=\diamond$ whenever $i\in H(u)$.

2. (p,q)-Periodic graphs

In this section, we discuss the fundamental property of periodicity, our goal, and some initial results. We can restrict our attention to the case where *p* and *q* are coprime, that is $\gcd(p,q)=1$, since it is well known that the general case can be reduced to the coprime case (see, for example, [1] and [11]). Also, we assume without loss of generality that $1<p<q$.

Fine and Wilf show that $L(0,p,q)=p+q-\gcd(p,q)$[8], Berstel and Boasson that $L(1,p,q)=p+q$[1], and Shur and Konovalova prove $L(2,p,q)$ to be $2p+q-\gcd(p,q)$[12]. Other results include the following.

Theorem 1.

(See [3] and [11].*) Let* $q>2$ *be an integer satisfying* $\gcd(2,q)=1$. *Then*

$$L(h,2,q)=h+q\left(1+\left\lfloor\frac{h}{q}\right\rfloor\right)+1.$$

Theorem 2.

(See [3].) Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$. *If* $q>p\lfloor\frac{h+1}{2}\rfloor$, *then*

$$L(h,p,q)=\begin{cases}p(\frac{h+2}{2})+q-1, & \text{if } h \text{ is even;}\\ p(\frac{h+1}{2})+q, & \text{if } h \text{ is odd.}\end{cases}$$

The problem of finding $L(h,p,q)$ is equivalent to a problem involving the vertex connectivity of certain graphs, as described in [3], which we now discuss.

Definition 1.

Let p and q be integers satisfying $1<p<q$ and $\gcd(p,q)=1$. The (p,q)-*periodic graph* of size l is the graph $G=(V,E)$ where $V=\{0,1,\ldots,l-1\}$ and for $i,j\in V$, the pair $\{i,j\}\in E$ if and only if $i \equiv j \bmod p$ or $i \equiv j \bmod q$.

The *p-class of vertex i* is $\{j \in V \mid j \equiv i \bmod p\}$. A *p-connection* (or *p-edge*) is an edge $\{i,j\}\in E$ such that $i \equiv j \bmod p$. If an edge $\{i,j\}$ is a p-connection, then i and j are *p-connected*. Similar statements hold for q-classes, q-connections and pq-classes, pq-connections.

Fig. 1 illustrates a (p,q)-periodic graph.

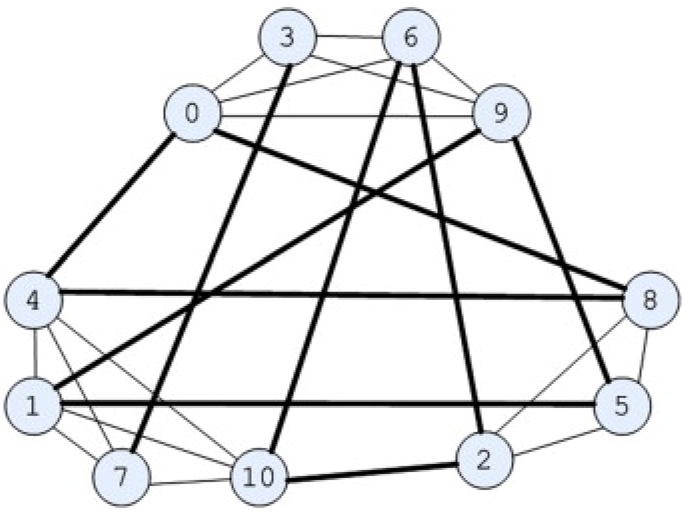

Fig. 1. The (3,4)-periodic graph of size 11. The bold connections are q-edges, while the lighter ones are p-edges.

The (p,q)-periodic graph G of size l can be thought to represent a full word u of length l with periods p and $q$as well as a partial word w with h holes of length l with periods p and q. Key observations are:

- Positions in u correspond to vertices in G.
- If there is a path from vertex i to vertex j, then $u(i)=u(j)$ (so if G is connected, then u has period 1).
- A hole in w corresponds to the removal of the associated vertex from G.
- If the h vertex removals disconnect G, then w need not have period 1.

Recall that a graph has vertex connectivity κ if it can be disconnected with a suitable choice of κ vertex removals, but cannot be disconnected by any choice of $\kappa-1$ vertex removals [9]. Thus, our goal, which is to determine $L(h,p,q)$ in all cases (when $\gcd(p,q)=1$), can be restated in terms of vertex connectivity.

Lemma 1.

The length L(h,p,q)*is the smallest size of a* (p,q)-*periodic graph with vertex connectivity at least* h+1.

Throughout the paper, we will find it useful to group together *p* -classes whose smallest elements are congruent modulo *r* where $r = q \bmod p$. We do so by introducing the *r-set* of vertex *i* , where i∈{0,1,…,r−1}, which is the set

$$\bigcup_{0 \leqslant j < p \text{ and } j \equiv i \bmod r} p\text{-class of vertex } j = \bigcup_{j=0}^{\lfloor \frac{p-i-1}{r} \rfloor} p\text{-class of vertex } jr+i.$$

Fig. 2 shows some (p,q)-periodic graphs in terms of $r = q \bmod p$.

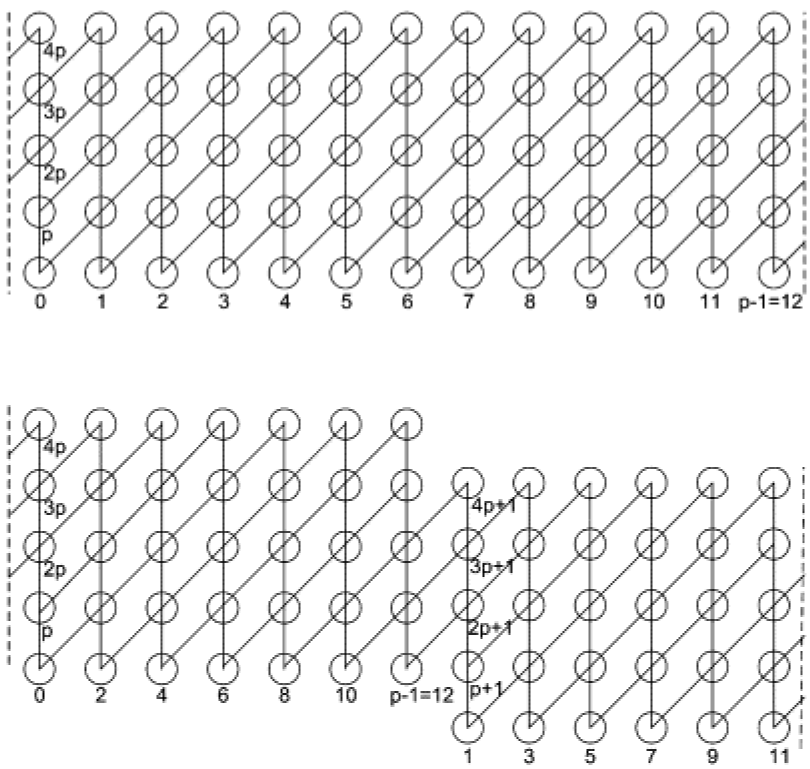

Fig. 2. Some (*p*,*q*)-periodic graphs where the vertical lines represent *p*-classes, while the diagonal lines represent *q*-classes. The*q*-edges wrap around at the dashed lines. All vertices in vertical and diagonal lines are connected to each other. In other words, lines represent several “normal” edges. In the first graph, *p* = 13, *q* = 14, and *r* = 1; this is the (13,14)-periodic graph of size 65 where the *p*-classes are grouped into one *r*-set (the *r*-set of vertex 0). In the second graph, *p* = 13, *q* = 15 and*r* = 2; this is the (13,15)-periodic graph of size 65 where the *p*-classes are grouped into two *r*-sets (the *r*-set of vertex 0 and the *r*-set of vertex 1).

3. Connectivity in (p,q)-periodic graphs

Our algorithm to calculate L(h,p,q) is based on connectivity in (p,q)-periodic graphs. In this section, we discuss modified degree connectivity and r -set connectivity in these graphs, where $r = q \bmod p$. Using Theorem 1 and Theorem 2, we can restrict our discussion to the case where $p \neq 2$ and $q \leqslant p\lfloor \frac{h+1}{2} \rfloor$.

Let G=(V,E) be a graph. A *disconnection* of G is a partition $\{V_1,V_2,H\}$ of V (that is, $V=V_1 \cup V_2 \cup H$ and V_1, V_2, H are mutually disjoint), such that neither V_1 nor V_2 is empty, and for $v_1 \in V_1$, $v_2 \in V_2$, $\{v_1,v_2\} \notin E$. An *optimal disconnection* is a disconnection such that the cardinality of H is κ, where κ is the vertex connectivity of G. The set H represents the vertices removed in a disconnection, while the sets V_1 and V_2 represent the vertices disconnected from each other in a disconnection.

If G is the (p,q)-periodic graph of size l for some p, q and l and $\{V_1,V_2,H\}$ is an optimal disconnection of G, then we cannot disconnect G within a p -class since p -classes form complete subgraphs. In other words, a p -class cannot contain elements of both V_1 and V_2, that is, for a p -class C, either $C \subset V_1 \cup H$ or $C \subset V_2 \cup H$. We say that a disconnection $\{V_1,V_2,H\}$ of G disconnects a union of p -classes P if $V_1 \subset P$ and $P \subset V_1 \cup H$, or $V_2 \subset P$ and $P \subset V_2 \cup H$. Similarly, a q -class cannot both contain elements in V_1 and V_2.

Suppose we want to disconnect a single p -class C from G. For a q -class C' of G, all of the vertices of C' within C or all of the vertices of C' outside of C must be removed. For $l \geqslant 2q$, a vertex $i \in C$ has q -connections with vertices outside of C. Each of these q -connections must be broken in order to disconnect C from G. The most efficient way to do so is to remove i itself, since i may have more than one q -connection. However, if we remove all of C from G, we have not formed a disconnection (V_1 or V_2 is empty). Thus, we do not remove the vertex in C contained in the smallest q -class in order to minimize the number of vertex removals required to disconnect C. So, if each vertex $i \in C$ is q-connected to some vertex j outside of C such that no other vertex in C is q-connected to j (no vertex in C is q-connected to i), then the most efficient way of disconnecting C from G is to disconnect a vertex of lowest degree in C.

When $l \leqslant pq$, any two distinct vertices within the same p-class belong to different q-classes. In this case, the most efficient way to disconnect a single p-class from G is to disconnect a single vertex of lowest degree in G (this is called a minimum degree disconnection).

When $l > pq$, distinct vertices within the same p-class may belong to the same q-class (that is to say, distinct vertices may be both p- and q-connected, or pq-connected). In this case, it is more efficient to disconnect the entire pq-class in order to disconnect a single p-class from G. For a vertex i in V, vertices that are pq-connected to i share all other connections with i, and thus should not be counted in the number of vertices required to disconnect i as they are disconnected when i is disconnected. Thus, we introduce the idea of "modified" degree.

Definition 2.

Let p and q be integers satisfying 1<p<q and gcd(p,q)=1. Let G=(V,E) be the (p,q)-periodic graph of size l, and let i∈V.

- The *degree* of i, denoted **d**(i), is the number of vertices connected to *i*, that is,

equation(1)

$$\left\lfloor \frac{l-1-i \bmod p}{p} \right\rfloor + \left\lfloor \frac{l-1-i \bmod q}{q} \right\rfloor - \left\lfloor \frac{l-1-i \bmod pq}{pq} \right\rfloor.$$

The first term gives the number of *p*-connections, the second term the number of *q*-connections, and the third term the number of *pq*-connections.

- The *modified degree* of i, denoted **d**□(i), is the number of vertices that are either *p*- or *q*-connected to *i*, but not *pq*-connected to *i*, that is,

equation(2)

$$\left\lfloor \frac{l-1-i \bmod p}{p} \right\rfloor + \left\lfloor \frac{l-1-i \bmod q}{q} \right\rfloor - 2\left\lfloor \frac{l-1-i \bmod pq}{pq} \right\rfloor.$$

In (2), we subtract 2 times the number of pq -connections: once because we double counted them, and again because vertices that are pq -connected are connected to the same vertices, so disconnecting one vertex will also disconnect all the vertices pq -connected to it. Note that when l⩽pq, **d**(i)=**d**□(i). When l>pq, minimum degree disconnections are replaced by minimum modified degree disconnections. Fig. 3 illustrates a minimum modified degree disconnection in some (p,q)-periodic graph.

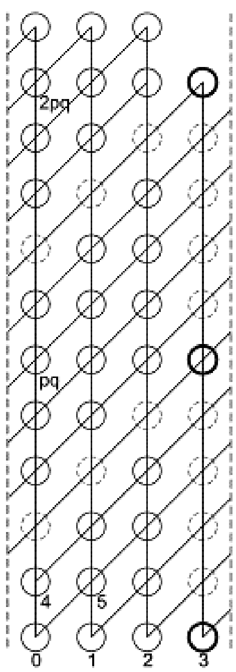

Fig. 3. The (4,5)-periodic graph of size 47. This figure depicts an optimal disconnection where the dashed vertices are in *H*, the bold vertices are in V_2, and the rest of the vertices are in V_1. Notice that the vertices in V_2 have the minimum modified degree. They are all *pq*-connected to each other, and are *p*- or *q*-connected to the vertices in *H*. Increasing the size of the graph by 1 gives this *pq*-class one more *p*-connection, thereby increasing the connectivity of the graph by 1.

Definition 3.

For a (p,q)-periodic graph *G*, a *minimum modified degree disconnection* is a disconnection that disconnects vertices of lowest modified degree in an entire *pq* -class from the other *pq* -classes. We define the *modified degree connectivity* of *G*, denoted κ_d, to be the smallest number of vertex removals required to make a minimum modified degree disconnection, and denote the minimum size of *G* such that $\kappa_d=h+1$ by $l_d(h,p,q)$.

Usually, disconnecting more than one *p*-class takes more holes than individually disconnecting any one *p*-class, because in general, a set of *p*-classes has more connections with the rest of the graph than any single *p*-class. However, disconnecting entire *r*-sets may prove to be efficient when *l* is small, as the graph "bottlenecks" between *r*-sets (that is, fewer *q*-classes span *r*-sets than connect *p*-classes within an *r*-set).

Definition 4.

For a (p,q)-periodic graph *G*, let $r = q \bmod p$. An *r -set disconnection* is a disconnection that disconnects an entire *r* -set from the other *r* -sets. We define the *r -set connectivity* of *G*, denoted κ_r, to be the smallest number of vertex removals required to make an *r* -set disconnection, and denote the minimum size of *G* such that $\kappa_r=h+1$ by $l_r(h,p,q)$.

Thus, if *G* is the (p,q)-periodic graph of size *l* for $l>2q$, then either a modified degree disconnection or an *r*-set disconnection will give an optimal disconnection of *G*.

Note that the sizes at which our graphs change connectivity are the optimal lengths in question. If the (p,q)-periodic graph of size *l* has vertex connectivity κ while the (p,q)-periodic graph of size l+1 has vertex connectivity $\kappa+1$, then $L(\kappa,p,q)=l+1$. Similarly, if the (p,q)-periodic graph of size *l* has modified degree connectivity κ_d (respectively, *r* -set connectivity κ_r) while the (p,q)-periodic graph of size l+1 has modified degree connectivity κ_d+1 (respectively, *r* -set connectivity κ_r+1), then $l_d(\kappa_d,p,q)=l+1$ (respectively, $l_r(\kappa_r,p,q)=l+1$).

Algorithm 2, which will be described in Section 5, will find L(h,p,q) when $1<p<q$ and $\gcd(p,q)=1$, based on the calculation of both $l_r(h,p,q)$ and $l_d(h,p,q)$ lengths. As mentioned earlier, if p=2 then L(h,p,q) is already known by Theorem 1. Otherwise, if $q > p\lfloor \frac{h+1}{2} \rfloor$, then L(h,p,q) is also already known by Theorem 2. And if $q \leq p\lfloor \frac{h+1}{2} \rfloor$, then $l_r(h,p,q)$ will be calculated using Theorem 3 and $l_d(h,p,q)$ using Theorem 4 (and Algorithm 1).

if $h+2 \leqslant \lfloor \frac{q}{p} \rfloor$ then $l_d(h,p,q)=(h+2)p$
else solve for $f(\omega,p,q)=2$ solutions for
  $l_d(h-1,p,q)$ and $l_d(h,p,q)$
  if the $f(\omega,p,q)=2$ value for $l_d(h,p,q)$ is $n_1 p = \lceil \frac{h+3}{1+\frac{p}{q}} \rceil p$ then
    find the maximum value of $n'_1 p \bmod q$ for $0 < n'_1 < n_1$
    if this vertex has a q-connection between
    $f(\omega,p,q)=2$ solutions for $l_d(h-1,p,q)$ and $l_d(h,p,q)$ then
      $l_d(h,p,q)$ is the position of this q-connection
    else $l_d(h,p,q)=n_1 p$
  if the $f(\omega,p,q)=2$ value for $l_d(h,p,q)$ is $n_2 q = \lceil \frac{h+3}{1+\frac{q}{p}} \rceil q$ then
    find the maximum value of $n'_2 q \bmod p$ for $0 < n'_2 < n_2$
    if this vertex has a p-connection between
    $f(\omega,p,q)=2$ solutions for $l_d(h-1,p,q)$ and $l_d(h,p,q)$ then
      $l_d(h,p,q)$ is the position of this p-connection
    else $l_d(h,p,q)=n_2 q$

Algorithm 1.

Find $l_d(h,p,q)$ when $1 < p < q$, $\gcd(p,q)=1$, and $h < p + q - 2$.

4. *r*-Set connectivity

Fig. 4 depicts an r -set disconnection in some (p,q)-periodic graph of size some multiple of p , whereq=mp+r with 0<r<p. This figure will be useful in understanding the arguments provided in the proof of the following theorem which gives a formula for $l_r(h,p,q)$.

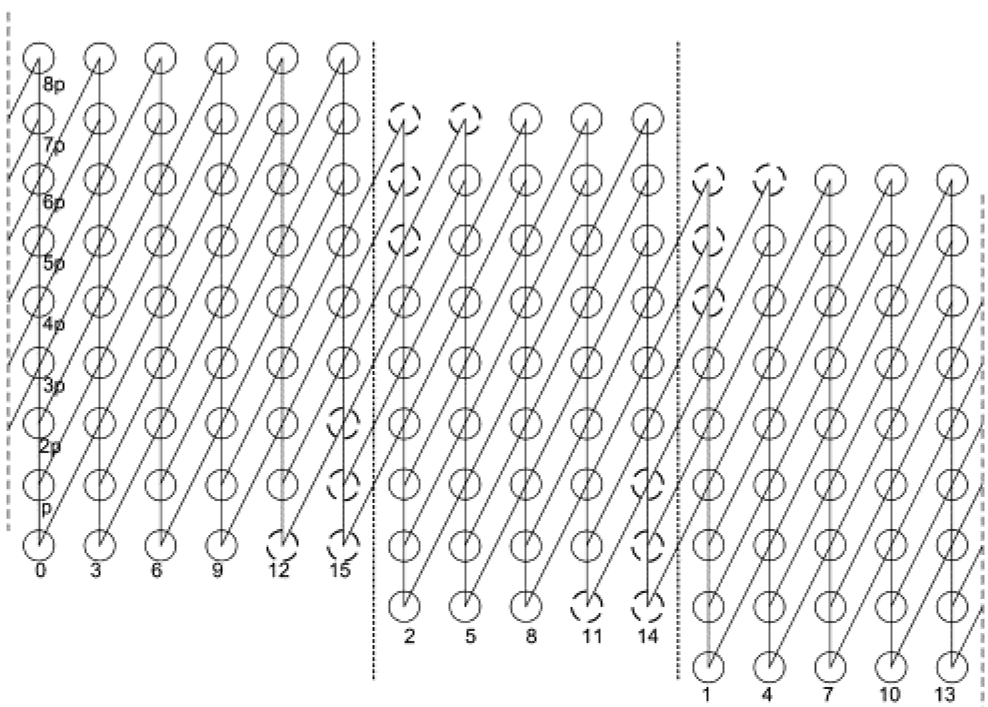

Fig. 4. An *r*-set disconnection for $p = 16$, $q = 35 = 2p + 3$, $r = 3$, and $l = 9p = 144$ (this length is not optimal). This is the (16,35)-periodic graph of size 144 where the *p*-classes are grouped into

three *r*-sets (the *r*-set of vertex 0, the *r*-set of vertex 1, and the *r*-set of vertex 2). Here we are disconnecting the *r*-set of vertex 2 from the two other *r*-sets.

Theorem 3.

Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$, *and set* $q=mp+r$ *where* $0<r<p$. *Then*

$$l_r(h,p,q)=(\beta+m+1)p+\left\lfloor\frac{\delta+1}{2}\right\rfloor r-(h+1)\bmod 2,$$

where

- $\beta=2m\gamma+\phi$;
- γ *is the greatest integer strictly less than* $\frac{\sqrt{m^2+2m(h+1)}-m}{2m}$;
- ϕ *is the greatest integer strictly less than* $\frac{h+1}{2(\gamma+1)}-m\gamma$;
- $\delta=h+1-2(m\gamma+\phi)(\gamma+1)$.

Proof.

Consider the (p,q)-periodic graph of size l where $q=mp+r$ with $0<r<p$.
Set $l=kp+r'$ where $0\leqslant r'<p$. There are k complete rows in each r -set (and an additional partial row when $r'>0$). In the columns on either side of any r -set, there are $m+1$ vertices which do not have q -connections to the adjacent r -set, so exactly $\beta=k-(m+1)$ vertices are q-connected to the adjacent r-set.

Consider two adjacent r -sets. Looking at the q -classes that connect these r -sets, the bottom m of these q -classes have 1 vertex in the left r -set. The next m q -classes have 2 vertices in the left r -set, and so on for the first $k-(m+1)$ q -classes. The left side of the right r -set is anti-symmetric to this: the top m q -classes each have 1 vertex in the right r -set, and the next m q -classes each have 2 vertices and so on working down. When breaking these q -connections it is best to remove all the vertices from the smaller side of the q -class. Thus, for the bottom half of the q -classes we remove vertices from the left side, and for the top half we remove the same number of vertices from the right side. If $\beta=\gamma(2m)+\phi$ for $0\leqslant\phi<2m$, then the number of vertices we must remove to separate these adjacent r -sets is

$$2m\sum_{i=1}^{\gamma} i+\phi(\gamma+1)=2m\frac{\gamma(\gamma+1)}{2}+\phi(\gamma+1)=m\gamma(\gamma+1)+\phi(\gamma+1).$$

Since an r -set disconnection requires separating adjacent r -sets twice, we have

$$\frac{\kappa_r}{2}=m\gamma(\gamma+1)+\phi(\gamma+1)=(m\gamma+\phi)(\gamma+1).$$

Since γ is an integer and $\phi<2m$, we can find γ in terms of κ_r and m by solving for when ϕ is equal to zero and then taking the floor. Using the quadratic formula, we calculate

$$\gamma = \left\lfloor \frac{\sqrt{m^2 + 2m\kappa_r} - m}{2m} \right\rfloor.$$

We solve for ϕ and find $\phi = \frac{\kappa_r}{2(\gamma+1)} - m\gamma$. From the definition of β we have $k=2m\gamma+\phi+m+1$.

The length is never optimal when $r'=0$ because κ_r only increases for nonzero values of r', as described below. We therefore want to select γ and ϕ such that they give us a value of κ_r that is strictly less than$h+1$. We will make room for the remaining vertex removals by adding r' vertices.

Now we need to calculate r' by determining at exactly which sizes the r -set connectivity actually increases. Starting with size $l=kp$, if we increase the size by r , then the number of vertex removals required to break any r -set connection increases by 1 because between each connected pair of r -sets there is one more q -connection. Thus, the r -set connectivity increases by 2. Notice that every connected pair of r -sets requires the same number of vertex removals to separate them. Thus, if we remove the last vertex we added, then the$r$ -set connectivity will have only increased by 1 from the previous size. After decreasing the size by one more vertex the r -set connectivity will be back down to where it was for $l=kp$. The same thing happens if we add another r vertices and continue until we reach the r -set connectivity of the graph of size $l=(k+1)p$.

If we have calculated k for a given p , q and h and define δ to be the difference between the r -set connectivity that we are looking for and the r -set connectivity at length $l=kp$, then $\delta=h+1-2(m\gamma+\phi)(\gamma+1)$ and we can calculate $r' = \lfloor \frac{\delta+1}{2} \rfloor r - (h+1) \bmod 2$. So

$$l_r(h,p,q) = kp + r' = (\beta + m + 1)p + \left\lfloor \frac{\delta+1}{2} \right\rfloor r - (h+1) \bmod 2$$

as desired. □

Using Theorem 3 we have calculated the lengths in Table 1.

Table 1. Some $l_r(h,p,q)$ lengths. The empty entries of the table are where $q > p\lfloor \frac{h+1}{2} \rfloor$ (see Theorem 2).

	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$
$p < q < 2p$	$2p + q$	$3p + q - 1$	$3p + q$	$2p + 2q - 1$	$2p + 2q$
$2p < q < 3p$			$3p + q$	$4p + q - 1$	$4p + q$
$3p < q < 4p$					$4p + q$

Let us show how Table 1’ s entry corresponding to $2p<q<3p$ and $h=5$ is calculated. Set $q=mp+r=2p+r$ where $0<r<p$. Here γ is the greatest integer strictly less than

$$\frac{\sqrt{m^2+2m(h+1)}-m}{2m}=\frac{\sqrt{2^2+2(2)(5+1)}-2}{2(2)}=\frac{3.29}{4}$$

and so $\gamma=0$. Also ϕ is the greatest integer strictly less than

$$\frac{h+1}{2(\gamma+1)}-m\gamma=\frac{5+1}{2(0+1)}-2(0)=3$$

and so $\phi=2$. Thus

$\beta=2m\gamma+\phi=2(2)(0)+2=2$

and

$\delta=h+1-2(m\gamma+\phi)(\gamma+1)=5+1-2(2(0)+2)(0+1)=2.$

Then

$$\begin{aligned} l_r(5,p,q) &= (\beta+m+1)p+\left\lfloor\frac{\delta+1}{2}\right\rfloor r-(h+1) \bmod 2 \\ &= (2+2+1)p+\left\lfloor\frac{2+1}{2}\right\rfloor r-(5+1) \bmod 2 \\ &= 5p+r=3p+2p+r=3p+q \end{aligned}$$

as desired.

By comparing the $l_r(h,p,q)$ lengths in Table 1 calculated using Theorem 3 to the $l_d(h,p,q)$ lengths that can be calculated using Theorem 4 and Algorithm 1 from Section 5, it turns out that r -set disconnections are only more efficient when $h=4$ and $q<\frac{3p}{2}$ (detailed calculations for the cases where $3\leqslant h\leqslant 7$ are provided in the proofs of Theorem 9, Theorem 10, Theorem 11, Theorem 12 and Theorem 13). As we increase the length beyond the values shown in Table 1, experimental evidence suggests that *r*-set disconnections will continue to become less efficient because *r*-sets now gain *q*-connections faster than any *pq*-class gains connections.

5. Modified degree connectivity

Let G be the (p,q)-periodic graph of size l . We now reduce the size of G modulo pq , that is, we reduce the case where $l\geqslant pq$ to that where $l<pq$. The idea is to write $l=\tau pq+\omega$ for some nonnegative integers τ,ω satisfying $\omega<pq$ and then show that the number of vertices we must remove to disconnect vertex i and all the vertices pq -connected to it is given by

equation(3)

$$\mathbf{d}^*(i) = \tau(p+q-2) + \mathbf{d}^*_{G'}(i),$$

where we use the formula in (2) for the modified degree of i in G, $\mathbf{d}^*(i)$, and where we denote by $\mathbf{d}^*_{G'}(i)$ the modified degree of i in the subgraph G' of G that contains only the last ω vertices. Solving $\mathbf{d}^*(i)=h+1$ gives the following theorem.

Theorem 4.

Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$. *Then the equality* $l_d(h,p,q)=\tau pq+\omega$ *holds, where* $\tau = \lfloor \frac{h+1}{p+q-2} \rfloor$ *and* $0\leqslant\omega<pq$. *Moreover,*

$$\omega = \begin{cases} l_d((h+1) \bmod (p+q-2) - 1, p, q), & \text{if } (h+1) \bmod (p+q-2) \neq 0; \\ 0, & \text{otherwise.} \end{cases}$$

Proof.

Let G be the (p,q)-periodic graph of size l. Suppose $l=\tau pq+\omega$ for nonnegative integers τ,ω with $\omega<pq$. If $\omega=0$ then, using (2), every vertex i has the same modified degree:

$\mathbf{d}^*(i)=(\tau q-1)+(\tau p-1)-2(\tau-1)=\tau(p+q-2)$.

If $\omega>0$ then define G' to be the subgraph of G that contains only the last ω vertices, that is, the vertices $\tau pq,\ldots,\tau pq+\omega-1$. Each of them has $\tau(p+q-2)$ vertices among the first τpq vertices, $0,1,\ldots,\tau pq-1$, to which it is either p-connected or q-connected but not pq-connected. Thus, the modified degree of a vertex i in G' is equal to $\tau(p+q-2)+\mathbf{d}_{G'}(i)$, where $\mathbf{d}_{G'}(i)$ is the degree of i in G'. In other words, we can find the degree of the vertex i within the subgraph G', and add this degree to $\tau(p+q-2)$ to get its modified degree in G. Thus, we have Eq. (3). The positions of these last ω vertices modulo pq are all less than $\omega = l \bmod pq$, and any two positions in the same pq-class have the same modified degree. Thus we know that one of them will have the lowest modified degree of the graph.

We want $\mathbf{d}^*(i)=h+1$. Since τ is an integer and $\mathbf{d}^*_{G'}(i) < p+q-2$, we can use the division algorithm and Eq. (3) to get $\tau = \lfloor \frac{h+1}{p+q-2} \rfloor$ and $\mathbf{d}^*_{G'}(i) = (h+1) \bmod (p+q-2)$. The length $l_d(h,p,q)$ being the smallest one at which the minimum modified degree is $h+1$, the result follows. □

We have now reduced cases where $l\geqslant pq$ to those cases where $l<pq$, so now we will assume $l=\omega<pq$.

Theorem 5.

Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$. *Define the function*

$$f(\omega,p,q)=\begin{cases}2, & \text{if there exists } i\in[0\ldots\omega-1] \text{ such that } i \bmod p \geqslant \omega \bmod p \text{ and } i \bmod q \geqslant \omega \bmod q;\\ 1, & \text{otherwise.}\end{cases}$$

Then the (p,q)-*periodic graph of size* ω *has a modified degree connectivity*

$$\kappa_d=\left\lfloor\frac{\omega}{p}\right\rfloor+\left\lfloor\frac{\omega}{q}\right\rfloor-f(\omega,p,q).$$

Proof.

A vertex i in a (p,q)-periodic graph of size ω has $\lfloor\frac{\omega}{p}\rfloor-1$ p -connections if $i\geqslant\omega \bmod p$ and $\lfloor\frac{\omega}{p}\rfloor$ p -connections if $i<\omega \bmod p$. Similarly, i has $\lfloor\frac{\omega}{q}\rfloor-1$ q -connections if $i\geqslant\omega \bmod q$ and $\lfloor\frac{\omega}{q}\rfloor$ q -connections if $i<\omega \bmod q$. We add together the number of p -connections and the number of q -connections to find that the degree of i is $\lfloor\frac{\omega}{p}\rfloor$ plus $\lfloor\frac{\omega}{q}\rfloor$ minus either 0,1 or 2 depending on the value of i . We can assume that $\omega\geqslant p$ because there will never be an optimal length with $0<l_d(h,p,q) \bmod pq<p$, since there are no p - or q -connections within this range. Thus, we can assume that the $p-1$ vertex exists and we know that it satisfies the condition $p-1\geqslant\omega \bmod p$. □

We now state our algorithm for finding $l_d(h,p,q)$.

Theorem 6.

Given a number of holes h and two periods p and q satisfying $1<p<q$, $\gcd(p,q)=1$, *and* $h<p+q-2$, *Algorithm* 1 *computes the length* $l_d(h,p,q)$.

Proof.

From Theorem 5, we can see that κ_d increases whenever $f(\omega,p,q)$ changes from 2 to 1, or whenever ω increases to a multiple of either p or q while $f(\omega,p,q)$ stays constant.

Remark 1.

If $l_d(h,p,q)=\omega$ and $f(\omega,p,q)=2$, then $\omega=n_1p$ or $\omega=n_2q$ for some positive integers n_1 and n_2.

Since adding a new vertex never decreases the modified degree connectivity of these (p,q)-periodic graphs, $f(\omega,p,q)$ can only change from 1 to 2 at multiples of p and q . If $\omega=n_1p$ for a positive integer n_1, then a vertex i in the q -class of $q-1$ satisfies $i \bmod p>\omega \bmod p$ and $i \bmod q>\omega \bmod q$, so$f(n_1p,p,q)=2$ for $n_1p>q$ and $f(n_1p,p,q)=1$ for $n_1p<q$. Similarly, $f(n_2q,p,q)=2$ for any positive integer n_2.

To calculate n_1 when f(ω,p,q)=2 we use the formula $\kappa_d = h+1 = \lfloor \frac{n_1 p}{p} \rfloor + \lfloor \frac{n_1 p}{q} \rfloor - 2$. We can solve as follows: $n_1 + \lfloor \frac{n_1 p}{q} \rfloor = (h+1)+2$ or $\lfloor n_1(1+\frac{p}{q}) \rfloor = h+3$. So if a solution exists, it is

equation(4)

$$n_1 = \left\lceil \frac{h+3}{1+\frac{p}{q}} \right\rceil.$$

If there is no solution for n_1 satisfying $n_1 + \lfloor \frac{n_1 p}{q} \rfloor - 2 = h+1$, then there must be a solution for $n_2$satisfying $\kappa_d = h+1 = \lfloor \frac{n_2 q}{p} \rfloor + \lfloor \frac{n_2 q}{q} \rfloor - 2$ and we calculate

equation(5)

$$n_2 = \left\lceil \frac{h+3}{1+\frac{q}{p}} \right\rceil.$$

We now consider the f(ω,p,q)=1 case. Note that f(ω,p,q)=1 for all ω<q. For these cases, the vertices can only have *p* -connections, and we can see that l_d(h,p,q)=(h+2)p so long as $h+2 \leqslant \lfloor \frac{q}{p} \rfloor$.

For $h+2 > \lfloor \frac{q}{p} \rfloor$, optimal l_d(h,p,q) lengths occur when vertices of minimum degree gain a new *p* - or *q* -connection. First, there is always a vertex of minimum degree in either the *p* -class of p−1 or the *q* -class ofq−1. This is because if we pick any vertex *i* , other than p−1 or q−1, that has minimum degree then there is some vertex $i+i'$ in either the *p* -class of p−1 or the *q* -class of q−1 that has no more *p*- and *q*-connections than *i* itself. This leads to the following remark.

Remark 2.

If l_d(h,p,q)=ω, f(ω,p,q)=1, and $h+2 > \lfloor \frac{q}{p} \rfloor$, then $\omega = n'_1 p + n'_2 q$ for some positive integers n'_1 and n'_2. For $\omega = n'_1 p + n'_2 q - 1$, the vertices of lowest degree are in the symmetric positions $n'_1 p - 1$ and $n'_2 q - 1$.

We now focus on finding these positions $n'_1 p - 1$ and $n'_2 q - 1$. If f(ω,p,q) changes from 2 to 1 when the $n'_1 p - 1$ vertex gains a *q* -connection, then we see from the definition of f(ω,p,q) that the $n'_1 p - 1$ vertex must have a larger value modulo *q* than the other vertices in the *p* -class of p−1. Thus, we can say that $(n'_1 p - 1) \bmod q > (n''_1 p - 1) \bmod q$ for all positive integers $n''_1 \neq n'_1$ where $n''_1 p < n'_1 p + n'_2 q$. Similarly, we must have $(n'_2 q - 1) \bmod p > (n''_2 q - 1) \bmod p$ for all positive integers $n''_2 \neq n'_2$ where $n''_2 q < n'_1 p + n'_2 q$. Also, $n'_1 p + n'_2 q$ must fall between the f(ω,p,q)=2 solutions for l_d(h−1,p,q) andl_d(h,p,q). □

For $m = \lfloor \frac{q}{p} \rfloor$, the $mp-1$ vertex has the lowest degree in a large number of cases when the size of the(p,q)-periodic graph is less than pq (keep in mind that we can reduce any case to one where the size is less than pq). The following lemma identifies many of these cases. We then use this knowledge to find a large number of $l_d(h,p,q)$ lengths in the theorem that follows.

Lemma 2.

Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$. *Let G be the* (p,q)-*periodic graph of size l, let* $q=mp+r$ *where* $0<r<p$, *and let* $l=nq+r_1$ *where* $0\leqslant r_1<q$. *Let* $mp\leqslant l\leqslant pq$. *If* $l \bmod q < mp$ *or* $nr-1 < l \bmod p$, *then the* $mp-1$ *vertex has minimum degree.*

Proof.

We require $l\geqslant mp$ so the $mp-1$ vertex exists, and we require $l\leqslant pq$ so we do not have vertices that are both p - and q -connected to each other. We have that $l=nq+r_1=n(mp+r)+r_1=mnp+nr+r_1$, so $l \equiv (nr+r_1) \bmod p$. A vertex in the p -class of i has $\lfloor \frac{l}{p} \rfloor p$ -connections if $i \bmod p < (nr+r_1) \bmod p$ or $\lfloor \frac{l}{p} \rfloor - 1 p$ -connections if $i \bmod p \geqslant (nr+r_1) \bmod p$. Similarly, the number of q -connections for a vertex in the q -class of j is n if $j \bmod q < r_1$ or $n-1$ if $j \bmod q \geqslant r_1$. The $mp-1$ vertex is in the p -class of $p-1$ so it always has $\lfloor \frac{l}{p} \rfloor - 1 p$ -connections since $p-1 \geqslant (nr+r_1) \bmod p$. The $mp-1$ vertex is in the q -class of $mp-1$ and so it has $n-1 q$ -connections if $r_1\leqslant mp-1$ and has n q -connections if $mp\leqslant r_1<q$. The degree of the $mp-1$ vertex is clearly minimal when $r_1<mp$, that is, when $l \bmod q < mp$.

However, if $mp\leqslant r_1\leqslant mp+s$ for some $0\leqslant s<r$, then the vertices in the q -class of $mp+s$ have one fewer q -connections than any other vertex, and may have the same number of p -connections as the $mp-1$ vertex, giving them a lower degree than the $mp-1$ vertex. These vertices are of the form $(mp+s)+tq=mp+s+t(mp+r)=(t+1)mp+tr+s$ for some nonnegative integer t satisfying $mp+s+tq\leqslant l-1$. Thus, a vertex $mp+s+tq$ falls in the p -class of $(tr+s) \bmod p$. Thus, vertices in the q -class of $mp+s$ have $\lfloor \frac{l}{p} \rfloor p$ -connections if and only if $(tr+s) \bmod p < l \bmod p$ for all integers $t\in\{0,\ldots,n-1\}$ and $s\in\{r_1-mp,\ldots,r-1\}$. If this is the case, then these vertices have one more p -connection than the $mp-1$ vertex and, therefore, do not have lower degree.

Since $t\leqslant n-1$ and $s\leqslant r-1$, we have that $tr+s\leqslant nr-1$. Note that if $nr-1 < l \bmod p$, then $(tr+s) \bmod p = (tr+s) < l \bmod p$ for all $t\in\{0,\ldots,n-1\}$ and $s\in\{r_1-mp,\ldots,r-1\}$. Thus, if $nr-1 < l \bmod p$, then the $mp-1$ vertex has lowest degree in G. □

The following theorem gives $l_d(h,p,q)$ when the $mp-1$ vertex has the minimum degree in the graph of size $l_d(h,p,q)-1$.

Theorem 7.

Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$. *Let* $q=mp+r$, *where* $0<r<p$. *Define* n_1 *as calculated using Eq.* (4) *and* n_2 *as calculated using Eq.* (5), *and define* $\omega'=\min\{n_1p, mp+(n_2-1)q\}$. *Let* $mp\leqslant\omega'\leqslant pq$. *If* $\omega' \bmod q < mp$ *or* $\lfloor \frac{\omega'}{q} \rfloor r - 1 < \omega' \bmod p$, *then* $l_d(h,p,q)=\omega'$.

Proof.

Let G denote the (p,q)-periodic graph of size l. If we restrict the size so that $mp\leqslant l\leqslant pq$ with $l \bmod q < mp$ or $nr-1 < l \bmod p$, then by Lemma 2 the vertex $mp-1$ of G has lowest degree. Thus, within these ranges, optimal $l_d(h,p,q)$ lengths occur whenever the $mp-1$ vertex gains a *p*- or *q*-connection.

The $mp-1$ vertex gains a p -connection exactly when $l=n_1p$ with $n_1>m$.

The $mp-1$ vertex gains a q -connection exactly when $l=mp+n'_2q$ for some positive integer n'_2. This fits the form described in Remark 2 where $n'_1=m$. We search for n'_2 satisfying $\max\{(n_1-1)p,(n_2-1)q\} < mp+n'_2q < n_1p$. Then, $l_d(h,p,q)=mp+n'_2q$ if and only if such an integer n'_2 exists. Since $mp<q$ and n_2q is the smallest multiple of q greater than n_1p, any such n'_2 satisfying the inequalities must be equal to n_2-1. We then know that $mp+n'_2q > \max\{(n_1-1)p,(n_2-1)q\}$, so we can now say that $l_d(h,p,q)=mp+n'_2q$ if and only if $mp+n'_2q$ is less than n_1p. Otherwise, $l_d(h,p,q)=n_1p$. □

We now state our algorithm for finding L(h,p,q).

Theorem 8.

Given a number of holes h and two periods p and q satisfying $1<p<q$ *and* $\gcd(p,q)=1$, *Algorithm*2 *computes the optimal length* $L(h,p,q)$. *The time for computing* $l_d(h,p,q)$ *is linear in p and q and constant in h.*

---

if $p=2$ then $L(h,p,q)=(2\lfloor \frac{h}{q} \rfloor+1)q+(h \bmod q)+1$ by Theorem 1
else
    if $q>p\lfloor \frac{h+1}{2} \rfloor$ then $L(h,p,q)=p\lfloor \frac{h+2}{2} \rfloor+q-(h+1) \bmod 2$
        by Theorem 2
    else
        compute $l_r(h,p,q)$ using Theorem 3
        compute $l_d(h,p,q)$ using Theorem 4 (and Algorithm 1)
        $L(h,p,q)=\max\{l_r(h,p,q),l_d(h,p,q)\}$

Algorithm 2.

Find $L(h,p,q)$ when $1 < p < q$ and $\gcd(p,q)=1$.

6. Closed formulas

Using the ideas of r -set and modified degree connectivities described in this paper, our methods can be used to prove closed formulas for any given number of holes (however, as the number of holes increases, the proofs become very tedious). Our calculations, performed for h=3 to h=7, show that an r -set disconnection is strictly more efficient than any modified degree disconnection, or $l_r(h,p,q)>l_d(h,p,q)$, if and only if h=4, p>2, and $q < \frac{3p}{2}$, in which case, L(h,p,q)=q+3p−1.

We now provide details for the closed formulas in the cases where $3\leqslant h\leqslant 7$. These five results confirm five conjectures stated in [3].

Theorem 9.

Let p and q be integers satisfying 2<p<q*and* gcd(p,q)=1. *Then* L(3,p,q)*is* p+2q*if* $q < \frac{3p}{2}$, 4p *if* $\frac{3p}{2} < q < 2p$*, and* 2p+q*if* q>2p.

Proof.

Let q=mp+r for some positive integers m and r such that 0<r<p, and let G denote the (p,q)-periodic graph of arbitrary size. The case where q>2p falls within the domain of Theorem 2. To find L(3,p,q) in the case where q<2p, we must find the sizes of G at which r -set and modified degree connectivities change from 3 to 4 (that is, $l_r(3,p,q)$ and $l_d(3,p,q)$), and take the maximum.

First we consult Table 1 and find that $l_r(3,p,q)=2p+q$ when q<2p. Now we find $l_d(3,p,q)$ using the algorithm described in Section 5. Let $l_d(3,p,q)=\tau pq+\omega$ where $0\leqslant\omega<pq$. From Theorem 4, we see that $\tau = \lfloor \frac{h+1}{p+q-2} \rfloor$. Since in this case h+1=4 and $5\leqslant p+q-2$, τ=0 (we also see that (h+1)mod(p+q−2)=4). Thus $l_d(3,p,q)=\omega$ for some ω<pq.

First, suppose f(ω,p,q)=2. We find that in this case, ω=4p when $\frac{4p}{3} < q < 2p$ and ω=3q when $q < \frac{4p}{3}$.

Now, suppose f(ω,p,q)=1. Note that $h+2 > \lfloor \frac{q}{p} \rfloor$. For $q < \frac{4p}{3}$, the multiple of q smaller than 3q that is largest modulo p is 2q . Since 3p<p+2q<3q we see that $l_d(3,p,q)=p+2q$. For $\frac{4p}{3} < q < \frac{3p}{2}$, the multiple of p smaller than 4p that is largest modulo q is p . Since 3p<p+2q<4p we see that$l_d(3,p,q)=p+2q$. For $\frac{3p}{2} < q < 2p$, we find that the f(ω,p,q)=2 case is optimal. Indeed, the multiple of$p$ smaller than 4p that is largest modulo q is 3p . Since 3p has no q -connections between 2q and 4p ,$l_d(3,p,q)=4p$.

Since $L(3,p,q)=\max\{l_d(3,p,q),l_r(3,p,q)\}$ and 2p+q<2q+p when $q < \frac{3p}{2}$ and 2p+q<4p when $\frac{3p}{2} < q < 2p$, we have that $l_d(3,p,q)$ is greater on these intervals. The result follows. □

Theorem 10.

Let p and q be integers satisfying $2<p<q$ *and* $\gcd(p,q)=1$. *Then* $L(4,p,q)$ *is* $q+3p-1$ *if* $q<\frac{3p}{2}$, $q+3p$ *if* $\frac{3p}{2}<q<2p$, *and* $q+3p-1$ *if* $q>2p$.

Proof.

Let $q=mp+r$ for some positive integers m and r such that $0<r<p$, and let G denote the (p,q)-periodic graph of arbitrary size. The case where $q>2p$ falls within the domain of Theorem 2. To find $L(4,p,q)$ in the case where $q<2p$, we must find the sizes of G at which r -set and modified degree connectivities change from 4 to 5 (that is, $l_r(4,p,q)$ and $l_d(4,p,q)$), and take the maximum.

First we consult Table 1 and find that $l_r(4,p,q)=3p+q-1$ when $q<2p$. Now we find $l_d(4,p,q)$ using the algorithm described in Section 5. Let $l_d(4,p,q)=\tau pq+\omega$ where $0\leqslant\omega<pq$. From Theorem 4, we see that $\tau=\lfloor\frac{h+1}{p+q-2}\rfloor$. In this case, $h+1=5$. Note that $p+q-2=5$ when $p=3$ and $q=4$, and $p+q-2>5$ otherwise. When $p=3$ and $q=4$, we have that $\tau=1$ and, since in this case $(h+1)\bmod(p+q-2)=0$, $\omega=0$, and $l_d(4,3,4)=\tau pq=12$. Since $l_r(4,3,4)=3p+q-1=12$, we have that $L(4,3,4)=12$.

When $p>3$ or $q>4$, since $p+q-2>5$, we have $\tau=0$, and $(h+1)\bmod(p+q-2)=5$. Thus $l_d(4,p,q)=\omega$ for some $\omega<pq$.

First, suppose $f(\omega,p,q)=2$. We find that in this case, $\omega=5p$ when $\frac{5p}{3}<q<2p$, $\omega=3q$ when $\frac{4p}{3}<q<\frac{5p}{3}$, and $\omega=4p$ when $q<\frac{4p}{3}$. Since $4p\leqslant l_r(4,p,q)=q+3p-1$, we have no need to check the $f(\omega,p,q)=1$ solution when $q<\frac{4p}{3}$, and we see that in this case, $L(4,p,q)=q+3p-1$.

Now, suppose $f(\omega,p,q)=1$. Note that $h+2>\lfloor\frac{q}{p}\rfloor$. For $\frac{4p}{3}<q<\frac{5p}{3}$, the multiple of q smaller than $3q$ that is largest modulo p is q if $q>\frac{3p}{2}$ and $2q$ if $q<\frac{3p}{2}$. If $q<\frac{3p}{2}$, we see that $2q$ has no p -connection between $4p$ and $3q$, and so in this case $l_d(4,p,q)=3q$. If $q>\frac{3p}{2}$, q has a p -connection between $4p$ and $3q$, namely $3p+q$, and so in this case $l_d(4,p,q)=3p+q$.
For $\frac{5p}{3}<q<2p$, the multiple of p smaller than $5p$ that is largest modulo q is $3p$. We see that $3p$ has a q -connection between $4p$ and $5p$, namely $3p+q$, and so in this case $l_d(4,p,q)=3p+q$.
Since $L(4,p,q)=\max\{l_d(4,p,q),l_r(4,p,q)\}$, and $3q\leqslant 3p+q-1$ when $q<\frac{3p}{2}$ and $3p+q>3p+q-1$, we have our result. □

Theorem 11.

Let p and q be integers satisfying $2<p<q$ *and* $\gcd(p,q)=1$. *Then* $L(5,3,4)=18$. *If* $p\neq3$ *or* $q\neq4$, *then* $L(5,p,q)$ *is* $3q+p$ *if* $q<\frac{4p}{3}$, $5p$ *if* $\frac{4p}{3}<q<\frac{5p}{3}$, $3q$ *if* $\frac{5p}{3}<q<2p$, $6p$ *if* $2p<q<3p$, *and* $q+3p$ *if* $3p<q$.

Proof.

Let q=mp+r for some positive integers *m* and *r* such that 0<r<p, and let *G* denote the (p,q)-periodic graph of arbitrary size. The case where q>3p falls within the domain of Theorem 2. To find L(5,p,q) in the case where q<3p, we must find the sizes of *G* at which *r* -set and modified degree connectivities change from 5 to 6 (that is, $l_r(5,p,q)$ and $l_d(5,p,q)$), and take the maximum.

First we consult Table 1 and find that $l_r(5,p,q)$ is 3p+q. Now we find $l_d(5,p,q)$ using the algorithm described in Section 5. Let $l_d(5,p,q)=\tau pq+\omega$ where $0\leqslant\omega<pq$. From Theorem 4, we see that $\tau=\lfloor\frac{h+1}{p+q-2}\rfloor$. In this case, h+1=6. Note that p+q−2=5 when p=3 and q=4, p+q−2=6 when p=3 andq=5, and p+q−2>6 otherwise. When p=3 and q=4, we have that τ=1 and, since in this case(h+1)mod(p+q−2)=1, $\omega=l_d((h+1)\bmod(p+q-2)-1,p,q)=l_d(0,p,q)=p+q-1=6$. We have that $l_d(5,3,4)=\tau pq+\omega=18$. Since $l_r(5,3,4)=3p+q=13$, we have thatL(5,3,4)=max{18,13}=18. When p=3 and q=5, we have that τ=1 and, since in this case(h+1)mod(p+q−2)=0, ω=0, and $l_d(5,3,5)=\tau pq=15$. Since $l_r(5,3,5)=3p+q=14$, we have thatL(5,3,5)=max{15,14}=15.

When p>3 or q>5, since p+q−2>6, we have τ=0, and (h+1)mod(p+q−2)=6. Thus $l_d(5,p,q)=\omega$for some ω<pq.

First, suppose f(ω,p,q)=2. We find that in this case, ω is 4*q* when $q<\frac{5p}{4}$, 5*p* when $\frac{5p}{4}<q<\frac{5p}{3}$, 3*q* when $\frac{5p}{3}<q<2p$, and 6*p* when 2p<q<3p.

Now, suppose f(ω,p,q)=1. Note that $h+2>\lfloor\frac{q}{p}\rfloor$. For $q<\frac{5p}{4}$, the multiple of *q* smaller than 4*q* that is largest modulo *p* is 3*q* . Since 3*q* has a *p* -connection between 4*p* and 4*q* , namely 3q+p,$l_d(5,p,q)=3q+p$.

For $\frac{5p}{4}<q<\frac{4p}{3}$, the multiple of *p* smaller than 5*p* that is largest modulo *q* is *p* . We see that *p* has a *q* -connection between 4*p* and 5*p* , namely p+3q, and so $l_d(5,p,q)=p+3q$.

For $\frac{4p}{3}<q<\frac{5p}{3}$, the multiple of *p* smaller than 5*p* that is largest modulo *q* is 3*p* if $q>\frac{3p}{2}$ and 4*p* if $q\leqslant\frac{3p}{2}$. If $q\leqslant\frac{3p}{2}$, we see that 4*p* has no *q* -connection between 3*q* and 5*p* , and so $l_d(5,p,q)=5p$. If $q>\frac{3p}{2}$, 3*p* has no *q* -connection between 3*q* and 5*p* , and so $l_d(5,p,q)=5p$.

For $\frac{5p}{3}<q<2p$, the multiple of *q* smaller than 3*q* that is largest modulo *p* is *q* . We see that *q* has no *p* -connection between 5*p* and 3*q* , and so in this case $l_d(5,p,q)=3q$.

For $2p<q<3p$, the multiple of p smaller than $6p$ that is largest modulo q is $5p$ if $q > \frac{5p}{2}$ and $2p$ if $q \leqslant \frac{5p}{2}$. If $q \leqslant \frac{5p}{2}$, $2p$ has no q -connection between $q+3p-1$ and $6p$, and so $l_d(5,p,q)=6p$. If $q > \frac{5p}{2}$, $5p$ has no q -connection between $q+3p-1$ and $6p$, and so $l_d(5,p,q)=6p$.

Since $L(5,p,q)= \max\{l_d(5,p,q),l_r(5,p,q)\}$, and $3p+q\leqslant 3q+p$, $3p+q\leqslant 5p$ when $\frac{4p}{3} < q < \frac{5p}{3}$,$3p+q\leqslant 3q$ when $\frac{5p}{3} < q < 2p$, and $3p+q\leqslant 6p$ when $2p<q<3p$, the result follows. □

Theorem 12.

Let p and q be integers satisfying $2<p<q$ *and* $\gcd(p,q)=1$.
Then $L(6,3,4)=19, L(6,4,5)=20$ *and* $L(6,3,5)=21$. *Otherwise,* $L(6,p,q)$ *is* $5p$ *if* $q < \frac{5p}{4}$, $4q$ *if* $\frac{5p}{4} < q < \frac{3p}{2}$, $6p$ *if* $\frac{3p}{2} < q < 2p$, $2q+2p$ *if* $2p < q < \frac{5p}{2}$, $7p$ *if* $\frac{5p}{2} < q < 3p$,
and $q+4p-1$ *if* $3p<q$.

Proof.

Let $q=mp+r$ for some positive integers m and r such that $0<r<p$, and let G denote the (p,q)-periodic graph of arbitrary size. The case where $q>3p$ falls within the domain of Theorem 2. To find $L(6,p,q)$ in the case where $q<3p$, we must find the sizes of G at which r -set and modified degree connectivities change from 6 to 7 (that is, $l_r(6,p,q)$ and $l_d(6,p,q)$), and take the maximum.

First we consult Table 1 and find that $l_r(6,p,q)$ is $4p+q-1$ when $2p<q<3p$,
and $2p+2q-1$ when $q<2p$. Now we find $l_d(6,p,q)$ using the algorithm described in Section 5.
Let $l_d(6,p,q)=\tau pq+\omega$ where $0\leqslant\omega<pq$. From Theorem 4, we see that $\tau = \lfloor \frac{h+1}{p+q-2} \rfloor$. In this case, $h+1=7$. Note
that $p+q-2=5$ when $p=3$ and $q=4$, $p+q-2=6$ when $p=3$ and $q=5$, $p+q-2=7$ when $p=4$ and $q=5$, and $p+q-2>7$ otherwise. When $p=3$ and $q=4$ for instance, we have that $\tau=1$ and, since in this case $(h+1)\bmod(p+q-2)=2$, $\omega=l_d((h+1)\bmod(p+q-2)-1,p,q)=l_d(1,p,q)=p+q=7$. We have that $l_d(6,3,4)=\tau pq+\omega=19$. Since $l_r(6,3,4)=2p+2q-1=13$, we have that $L(6,3,4)=\max\{19,13\}=19$.

When $p+q-2>7$, we have $\tau=0$, and $(h+1)\bmod(p+q-2)=7$. Thus $l_d(6,p,q)=\omega$ for some $\omega<pq$.

First, suppose $f(\omega,p,q)=2$. We find that in this case, ω is $5p$ when $q < \frac{5p}{4}$, $4q$ when $\frac{5p}{4} < q < \frac{3p}{2}$, $6p$ when $\frac{3p}{2} < q < 2p$, $3q$ when $2p < q < \frac{7p}{3}$, and $7p$ when $\frac{7p}{3} < q < 3p$.

Now, suppose $f(\omega,p,q)=1$. Note that $h + 2 > \lfloor \frac{q}{p} \rfloor$. For $q < \frac{5p}{4}$, the multiple of p smaller than $5p$ that is largest modulo q is p. Since p has no q -connection between $4q$ and $5p$, $l_d(6,p,q)=5p$.

For $\frac{5p}{4} < q < \frac{4p}{3}$, the multiple of q smaller than $4q$ that is largest modulo p is $3q$, which has no p-connection between $5p$ and $4q$, and so $l_d(6,p,q)=4q$. For $\frac{4p}{3} < q < \frac{3p}{2}$, the multiple of q smaller than $4q$ that is largest modulo p is $2q$, which has no p-connection between $5p$ and $4q$, and so $l_d(6,p,q)=4q$.

For $\frac{3p}{2} < q < 2p$, the multiple of p smaller than $6p$ that is largest modulo q is $5p$ if $q > \frac{5p}{3}$ and $3p$ if $q \leqslant \frac{5p}{3}$. For $q > \frac{5p}{3}$, we see that $5p$ has no q-connection between $3q$ and $6p$, and so in this case $l_d(6,p,q)=6p$. If $q \leqslant \frac{5p}{3}$, $3p$ has no q-connection between $5p$ and $6p$, and so $l_d(6,p,q)=6p$.

For $2p < q < \frac{7p}{3}$, the multiple of q smaller than $3q$ that is largest modulo p is $2q$, which has a p-connection between $6p$ and $3q$, namely $2q+2p$, and so $l_d(6,p,q)=2q+2p$.

For $\frac{7p}{3} < q < 3p$, the multiple of p smaller than $7p$ that is largest modulo q is $5p$ if $q > \frac{5p}{2}$ and $2p$ if $q \leqslant \frac{5p}{2}$. If $q \leqslant \frac{5p}{2}$, $2p$ has a q-connection between $6p$ and $7p$, namely $2p+2q$, and so $l_d(6,p,q)=2p+2q$. If $q > \frac{5p}{2}$, $5p$ has no q-connection between $6p$ and $7p$, and so $l_d(6,p,q)=7p$.

Since $L(6,p,q)= \max\{l_d(6,p,q),l_r(6,p,q)\}$, and $2p+2q-1 \leqslant 5p$ when $q < \frac{5p}{4}$, $2p+2q-1 \leqslant 4q$ when $\frac{5p}{4} < q < \frac{3p}{2}$, $2p+2q-1 \leqslant 6p$ when $\frac{3p}{2} < q < 2p$, $4p+q-1 \leqslant 2p+2q$ when $2p < q < \frac{5p}{2}$, and $4p+q-1 \leqslant 7p$ when $\frac{5p}{2} < q < 3p$, the result follows. □

Theorem 13.

Let p and q be integers satisfying $2<p<q$ *and* $\gcd(p,q)=1$.
Then $L(7,3,4)=L(7,3,7)=21$, $L(7,3,5)=23$ *and* $L(7,4,5)=28$.
Otherwise, $L(7,p,q)$ *is* $4q+p$ *if* $q < \frac{5p}{4}$, $6p$ *if* $\frac{5p}{4} < q < \frac{3p}{2}$, $4q$ *if* $\frac{3p}{2} < q < \frac{5p}{3}$, $q+5p$ *if* $\frac{5p}{3} < q < 2p$, $7p$ *if* $2p < q < \frac{7p}{3}$, $3q$ *if* $\frac{7p}{3} < q < \frac{5p}{2}$, $q+5p$ *if* $\frac{5p}{2} < q < 3p$, $8p$ *if* $3p<q<4p$, *and* $q+4p$ *if* $4p<q$.

Proof.

Let $q=mp+r$ for some positive integers m and r such that $0<r<p$, and let G denote the (p,q)-periodic graph of arbitrary size. The case where $q>4p$ falls within the domain of Theorem 2. To find $L(7,p,q)$ in the case where $q<4p$, we must find the sizes of G at which r-set and modified degree connectivities change from 7 to 8 (that is, $l_r(7,p,q)$ and $l_d(7,p,q)$), and take the maximum.

First we consult Table 1 and find that $l_r(7,p,q)$ is $4p+q$ when $2p<q<4p$, and $2p+2q$ when $q<2p$. Now we find $l_d(7,p,q)$ using the algorithm described in Section 5.

Let $l_d(7,p,q)=\tau pq+\omega$ where $0\leqslant\omega<pq$. From Theorem 4, we see that $\tau=\lfloor\frac{h+1}{p+q-2}\rfloor$. In this case, $h+1=8$. Note

that $p+q-2=5$ when $p=3$ and $q=4$, $p+q-2=6$ when $p=3$ and $q=5$, $p+q-2=7$ when $p=4$ and $q=5$, $p+q-2=8$ when $p=3$ and $q=7$, and $p+q-2>8$ otherwise. When $p=3$ and $q=4$, we have that $\tau=1$ and, since in this case $(h+1)\bmod(p+q-2)=3$, $\omega=l_d((h+1)\bmod(p+q-2)-1,p,q)=l_d(2,p,q)=2p+q-1=9$. We have that $l_d(7,3,4)=\tau pq+\omega=21$. Since $l_r(7,3,4)=2p+2q=14$, we have that $L(7,3,4)=\max\{21,14\}=21$. When $p=3$ and $q=5$, we have that $\tau=1$ and, since in this case $(h+1)\bmod(p+q-2)=2$, $\omega=l_d((h+1)\bmod(p+q-2)-1,p,q)=l_d(1,p,q)=p+q=8$, and $l_d(7,3,5)=\tau pq+\omega=23$. Since $l_r(7,3,5)=2p+2q=16$, we have that $L(7,3,5)=\max\{23,16\}=23$. When $p=3$ and $q=7$, we have that $\tau=1$ and, since in this case $(h+1)\bmod(p+q-2)=0$, $\omega=0$. We have that $l_d(7,3,7)=\tau pq=21$. Since $l_r(7,3,7)=4p+q=19$, we have that $L(7,3,7)=\max\{21,19\}=21$. When $p=4$ and $q=5$, we have that $\tau=1$ and, since in this case $(h+1)\bmod(p+q-2)=1$, $\omega=l_d((h+1)\bmod(p+q-2)-1,p,q)=l_d(0,p,q)=p+q-1=8$. We have that $l_d(7,4,5)=\tau pq+\omega=28$. Since $l_r(7,4,5)=2p+2q=18$, we have that $L(7,4,5)=\max\{28,18\}=28$.

When $p+q-2>8$, we have $\tau=0$, and $(h+1)\bmod(p+q-2)=8$. Thus $l_d(7,p,q)=\omega$ for some $\omega<pq$.

First, suppose $f(\omega,p,q)=2$. We find that in this case, ω is $5q$ when $q<\frac{6p}{5}$, $6p$ when $\frac{6p}{5}<q<\frac{3p}{2}$, $4q$ when $\frac{3p}{2}<q<\frac{7p}{4}$, $7p$ when $\frac{7p}{4}<q<\frac{7p}{3}$, $3q$ when $\frac{7p}{3}<q<\frac{8p}{3}$, and $8p$ when $\frac{8p}{3}<q<4p$.

Now, suppose $f(\omega,p,q)=1$. Note that $h+2>\lfloor\frac{q}{p}\rfloor$. For $q<\frac{6p}{5}$, the multiple of q smaller than $5q$ that is largest modulo p is $4q$. Since $4q$ has a p-connection between $5p$ and $5q$, namely $4q+p$, we get $l_d(7,p,q)=4q+p$.

For $\frac{6p}{5}<q<\frac{4p}{3}$, the multiple of p smaller than $6p$ that is largest modulo q is p if $q\leqslant\frac{5p}{4}$ and $5p$ if $q>\frac{5p}{4}$. If $q\leqslant\frac{5p}{4}$, we see that p has a q-connection between $5p$ and $6p$, namely $p+4q$, and so $l_d(7,p,q)=p+4q$. If $q>\frac{5p}{4}$, $5p$ has no q-connection between $4q$ and $6p$, and so $l_d(7,p,q)=6p$. For $\frac{4p}{3}<q<\frac{3p}{2}$, the multiple of p smaller than $6p$ that is largest modulo q is $4p$, which has no q-connection between $4q$ and $6p$, and so $l_d(7,p,q)=6p$.

For $\frac{3p}{2}<q<\frac{7p}{4}$, the multiple of q smaller than $4q$ that is largest modulo p is $3q$ if $q\leqslant\frac{5p}{3}$ and q if $q>\frac{5p}{3}$. If $q\leqslant\frac{5p}{3}$, $3q$ has no p-connection between $6p$ and $4q$, and so $l_d(7,p,q)=4q$. If $q>\frac{5p}{3}$, q has a p-connection between $6p$ and $4q$, namely $q+5p$, and so $l_d(7,p,q)=q+5p$.

For $\frac{7p}{4}<q<2p$, the multiple of p smaller than $7p$ that is largest modulo q is $5p$, which has a q-connection between $6p$ and $7p$, namely $5p+q$, and so $l_d(7,p,q)=5p+q$. For $2p<q<\frac{7p}{3}$, the multiple of p smaller than $7p$ that is largest modulo q is $2p$, which has no q-connection between $3q$ and $7p$, and so $l_d(7,p,q)=7p$.

For $\frac{7p}{3} < q < \frac{8p}{3}$, the multiple of q smaller than $3q$ that is largest modulo p is q if $q > \frac{5p}{2}$ and $2q$ if $q \leqslant \frac{5p}{2}$. If $q \leqslant \frac{5p}{2}$, $2q$ has no p-connection between $7p$ and $3q$, and so $l_d(7,p,q)=3q$. If $q > \frac{5p}{2}$, q has a p-connection between $7p$ and $3q$, namely $q+5p$, and so $l_d(7,p,q)=q+5p$.

For $\frac{8p}{3} < q < 3p$, the multiple of p smaller than $8p$ that is largest modulo q is $5p$, which has a q-connection between $7p$ and $8p$, namely $5p+q$, and so $l_d(7,p,q)=5p+q$. For $3p<q<4p$, the multiple of p smaller than $8p$ that is largest modulo q is $7p$ if $q > \frac{7p}{2}$ and $3p$ if $q \leqslant \frac{7p}{2}$. If $q \leqslant \frac{7p}{2}$, $3p$ has no q-connection between $q+4p-1$ and $8p$, and so $l_d(7,p,q)=8p$. If $q > \frac{7p}{2}$, $7p$ has no q-connection between $q+4p-1$ and $8p$, and so $l_d(7,p,q)=8p$.

Since $L(7,p,q)= \max\{l_d(7,p,q),l_r(7,p,q)\}$, and $2p+2q\leqslant 4q+p$ when $q < \frac{5p}{4}$, $2p+2q\leqslant 6p$ when $\frac{5p}{4} < q < \frac{3p}{2}$, $2p+2q\leqslant 4q$ when $\frac{3p}{2} < q < \frac{5p}{3}$, $2p+2q\leqslant 5p+q$ when $\frac{5p}{3} < q < 2p$, $4p+q\leqslant 7p$ when $2p < q < \frac{7p}{3}$, $4p+q\leqslant 3q$ when $\frac{7p}{3} < q < \frac{5p}{2}$, $4p+q\leqslant q+5p$ when $\frac{5p}{2} < q < 3p$, and $4p+q\leqslant 8p$ when $3p<q<4p$, the result follows. □

7. Conclusion

Our goal was to give an algorithm for determining the minimum length $L(h,p,q)$ which guarantees that $\gcd(p,q)$ is also a period of any partial word having periods p and q, having h holes, and having at least that length, and to show how to use it to derive the closed formulas.

A topic of future research is to use our approach to study partial words with number of holes h, periods p and q, and length $L(h,p,q)-1$. We let $W_{h,p,q}$ denote the set of all such words, and we let $V_{h,p,q}$ denote the set of all such words which do not have $\gcd(p,q)$ as a period. The sets PER_h and $VPER_h$ are defined as follows:

$$\mathrm{PER}_h = \bigcup_{\gcd(p,q)=1} \mathcal{W}_{h,p,q} \quad \text{and} \quad \mathcal{V}\mathrm{PER}_h = \bigcup_{\gcd(p,q)=1} \mathcal{V}_{h,p,q}.$$

It turns out that $VPER_0$ has remarkable combinatorial properties [7]. The following is a result from [3] concerning PER_1, the proof of which we have simplified with the use of (p,q)-periodic graphs.

Theorem 14.

Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$.

1. *Given a singleton set H satisfying* $H \subset \{0,\ldots,p+q-2\} \setminus \{p-1,\ldots,q-1\}$, $W_{1,p,q}$ *contains a unique partial word u* (*up to a renaming*) *such that the cardinality of* $\alpha(u)$ *is* 2 *and* $H(u)=H$.

2. *Given a singleton set H satisfying* $H \subset \{p-1,\ldots,q-1\}$, $W_{1,p,q}$ *contains a unique word u such that* $\|\alpha(u)\|=1$ *and* $H(u)=H$.

Proof.

Let *G* be the (p,q)-periodic graph of size $p+q-1$. We have from Fine and Wilf's theorem that *G* is connected. In *G*, we have *p p* -classes connected by $p-1$ *q* -connections, so removing any vertex that has a *q* -connection will disconnect *G* into two components, whereas removing a vertex with no *q* -connections will not disconnect *G*. We see that the vertices in $\{0,\ldots,p+q-2\} \setminus \{p-1,\ldots,q-1\}$ each have a *q* -connection while the vertices in $\{p-1,\ldots,q-1\}$ do not have any *q*-connections. □

The following theorem, which gives a characterization of $VPER_2$, answers positively a conjecture of [3].

Theorem 15.

Let p and q be integers satisfying $1<p<q$ *and* $\gcd(p,q)=1$. *The membership* $u \in V_{2,p,q}$ *holds if and only if*

- $H(u)=\{p-2,p-1\}$ *or* $H(u)=\{q+p-1,q+p-2\}$ *or* $H(u)=\{p-2,q+p-1\}$ *when* $q-p=1$;
- $H(u)=\{p-2,p-1\}$ *or* $H(u)=\{q+p-1,q+p-2\}$ *or* $H(u)=\{p-2,q+p-1\}$ *or* $H(u)=\{p-1,q+p-2\}$ *when* $q-p>1$.

Proof.

Let $q=mp+r$ where *m* is an integer and $0<r<p$, and let *G* be the (p,q)-periodic graph of size $L(2,p,q)-1=2p+q-2$. We will first consider the case when $r=1$. We can form a cycle in *G* as follows: the 0 vertex is *q* -connected to the *q* vertex, which is *p* -connected to the 1 vertex, which is *q* -connected to the $q+1$ vertex, and so on, until we have the $q+p-1$ vertex *p* -connected to the 0 vertex. Note that this cycle visits all the *p* -classes, as it visits the $0,1,2,\ldots,p-1$ vertices. The cycle can be seen in Fig. 5. The *p*-connections are shown as dotted lines, while the *q*-connections are shown as full lines.

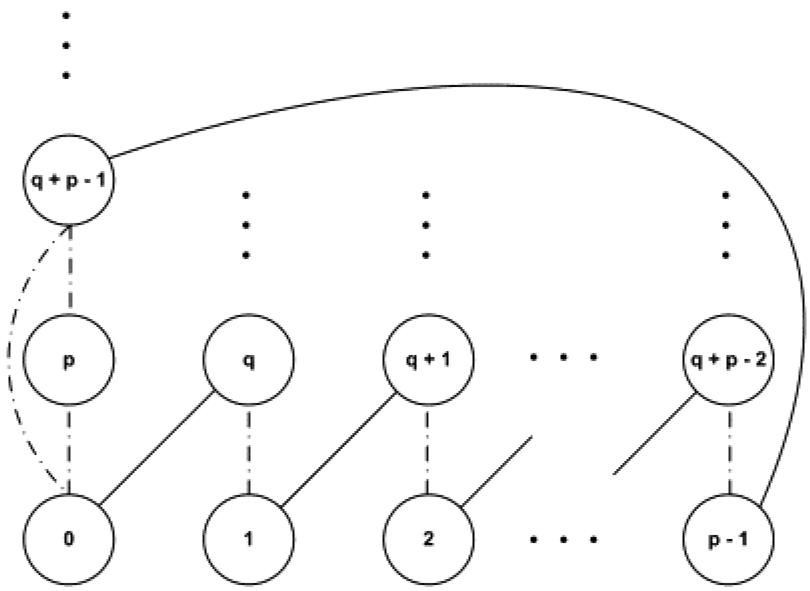

Fig. 5. A part of the graph *G*.

Thus, in order to disconnect *G* , this cycle must be broken. This requires two vertex removals. However, if one of the 0,1,…,p−3,q,q+1,…,q+p−3 vertices is removed, the cycle can be "fixed" around that vertex as follows:

If 0 is removed, we had previously for the cycle around 0:

$$q + p - 1 \xrightarrow{p} 0 \xrightarrow{q} q.$$

We fix it with:

$$q + p - 1 \xrightarrow{p} p \xrightarrow{q} p + q \xrightarrow{p} q.$$

If i∈{1,2,…,p−3} is removed, we had previously for the cycle around *i*:

$$i + q - 1 \xrightarrow{p} i \xrightarrow{q} i + q.$$

We fix it with:

$$i + q - 1 \xrightarrow{p} i + p \xrightarrow{q} i + p + q \xrightarrow{p} i + q.$$

If i∈{q,q+1,…,q+p−3} is removed, we had previously for the cycle around *i*:

$$i - q \xrightarrow{q} i \xrightarrow{p} i - q + 1.$$

We fix it with:

$i-q \xrightarrow{p} i-q+p \xrightarrow{q} i+p \xrightarrow{p} i-q+1.$

Fixing the cycle in two of the cases is shown in Fig. 6 and Fig. 7.

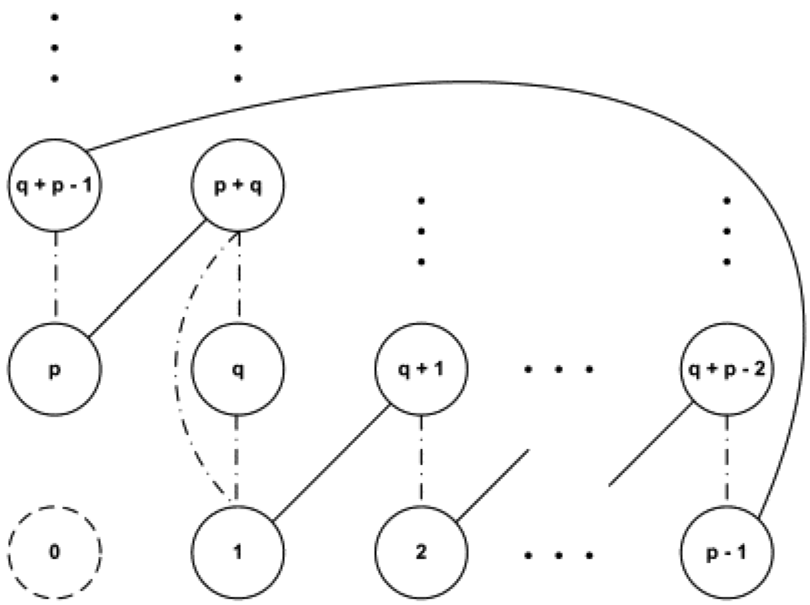

Fig. 6. Fixing the cycle when the 0 vertex is removed.

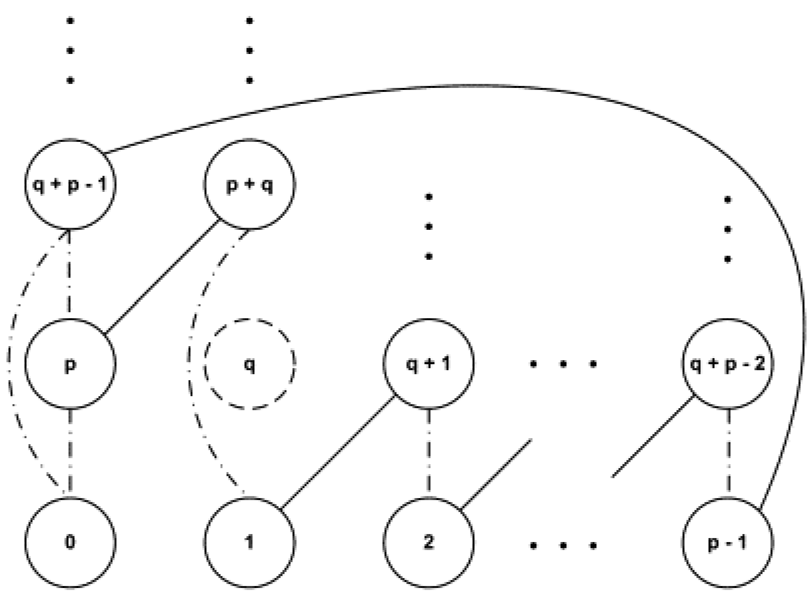

Fig. 7. Fixing the cycle when the *q* vertex is removed.

Thus, if one of the 0,1,…,p−3,q,q+1,…,q+p−3 vertices is removed, at least another two vertex removals are required to disconnect the graph.

Consider, however, the p -class of vertex p−1. There are only two q -connections in G such that one of the adjacent vertices is in the p -class of vertex p−1, namely the edge between the p−1 and q+p−1 vertices and the edge between the p−2 and q+p−2 vertices. Thus, removal of the p−2, p−1 vertices, the p−2,q+p−1 vertices, the p−1, q+p−2 vertices, or the q+p−1, q+p−2 vertices disconnects the p -class of vertex p−1 from the rest of the graph G . In the m=1 case, the p−1, q+p−2 vertices form the entire p -class of vertex p−1; otherwise, removal of the p−1, q+p−2 vertices disconnects the graph G.

Let us now consider the case when r>1. We can form a cycle between the p -classes of G similar to the cycle in the case above by making use of the r -sets. The path

$$i \xrightarrow{q} i+q \xrightarrow{p} i+r \xrightarrow{q} i+q+r \xrightarrow{p} i+2r \xrightarrow{q} \cdots \xrightarrow{p} i+\left\lfloor \frac{p-i-1}{r} \right\rfloor r,$$

denoted by **p**(i), visits all the *p*-classes in the *r*-set of vertex *i*. We also have that

$$i+\left\lfloor \frac{p-i-1}{r} \right\rfloor r \xrightarrow{q} i+\left\lfloor \frac{p-i-1}{r} \right\rfloor r+q \xrightarrow{p} \left(i+(-p \bmod r)\right) \bmod r.$$

Note that r and $-p \bmod r$ are coprime. Thus, we have that the cycle

p(0)→**p**(−pmodr)→⋯→**p**(r−(−pmodr)),

which starts and ends at 0, visits all the *p*-classes in *G*.

If this cycle is broken by removing one of the 0,1,…,p−3,q,q+1,…,q+p−3 vertices, it can be fixed in much the same way as above. Also, we have that removing the p−2,p−1 vertices, the q+p−1,q+p−2vertices, the p−2,q+p−1 vertices, or the p−1,q+p−2 vertices disconnects at least the r -set of vertex $(p-1) \bmod r$ from the rest of *G*. Thus, we have our result. □

Another topic of future research is to extend our approach to any number of periods.

A World Wide Web server interface has been established at www.uncg.edu/cmp/research/finewilf4 for automated use of a program which given as input a number of holes h and two periods p and q , outputsL(h,p,q) and an optimal word for that length.

References

[1] J. Berstel, L. Boasson, Partial words and a theorem of Fine and Wilf, Theoretical Computer Science 218 (1999) 135–141.

[2] F. Blanchet-Sadri, Algorithmic Combinatorics on Partial Words, Chapman & Hall/CRC Press, Boca Raton, FL, 2008.

[3] F. Blanchet-Sadri, D. Bal, G. Sisodia, Graph connectivity, partial words, and a theorem of Fine and Wilf, Information and Computation 206 (5) (2008)

676–693.

[4] F. Blanchet-Sadri, T. Mandel, G. Sisodia, Periods in partial words: an algorithm, in: C.S. Iliopoulos, W.F. Smyth (Eds.), 22nd International Workshop on

Combinatorial Algorithms, IWOCA 2011, Victoria, British Columbia, Canada, in: Lecture Notes in Computer Science, vol. 7056, Springer-Verlag, Berlin,

Heidelberg, 2011, pp. 57–70.

[5] C. Choffrut, J. Karhumäki, Combinatorics of words, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, vol. 1, Springer-Verlag, Berlin,

1997, pp. 329–438, Chapter 6.

[6] S. Constantinescu, L. Ilie, Generalised Fine and Wilf's theorem for arbitrary number of periods, Theoretical Computer Science 339 (2005) 49–60.

[7] A. de Luca, On the combinatorics of finite words, Theoretical Computer Science 218 (1999) 13–39.

[8] N.J. Fine, H.S. Wilf, Uniqueness theorems for periodic functions, Proceedings of the American Mathematical Society 16 (1965) 109–114.

[9] J.L. Gross, J. Yellen, Handbook of Graph Theory, CRC Press, 2004.

[10] V. Halava, T. Harju, T. Kärki, Interaction properties of relational periods, Discrete Mathematics and Theoretical Computer Science 10 (2008) 87–112.

[11] A.M. Shur, Y.V. Gamzova, Partial words and the interaction property of periods, Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya 68 (2)

(2004) 191–214.

[12] A.M. Shur, Y.V. Konovalova, On the periods of partial words, in: J. Sgall, A. Pultr, P. Kolman (Eds.), 26th International Symposium on Mathematical

Foundations of Computer Science, MFCS 2001, in: Lecture Notes in Computer Science, vol. 2136, Springer-Verlag, London, UK, 2001, pp. 657–665.

[13] W.F. Smyth, Computing Patterns in Strings, Pearson/Addison–Wesley, 2003.

[14] W.F. Smyth, S. Wang, A new approach to the periodicity lemma on strings with holes, Theoretical Computer Science 410 (2009) 4295–4302.