Kevin Samuels

Cages

Contents:

I Definitions and notation

A **graph** consists of a vertex set and an edge set. The vertex set must be finite and nonempty; the edge set must be finite but can be empty. Each edge is associated with two vertices, known as its endpoints. A graph is typically represented as a drawing with dots representing vertices and lines representing edges.
A **loop** is an edge for which both endpoints are the same vertex. A **multiple edge** is an edge which shares both endpoints with at least one other edge. A graph is **simple** if it contains no loops or multiple edges.
Two vertices in a graph are **neighbors**, or are **adjacent**, if they are endpoints of a single edge.

A **path** is an alternating sequence of vertices and edges in a graph, beginning and ending on a vertex, such that each successive edge has the previous vertex as an endpoint, each successive vertex is an endpoint of the previous edge, and no vertex in the sequence repeats. A **cycle** is a similar sequence of vertices and edges which begins and ends on the same vertex. The **length** of a path or a cycle is the number of edges that this sequence contains. The **distance** between two vertices in a graph is the length of the shortest path between them.

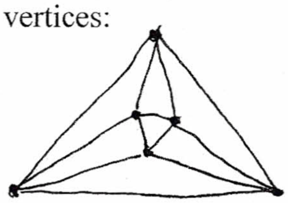
The **girth** of a graph is the length of the shortest cycle in the graph. If a graph has no cycles, its girth is defined to be infinite.
Note that a graph is simple if and only if its girth is at least 3 – a loop is a cycle of length 1, and two multiple edges form a cycle of length 2.

The **degree** of a vertex in a graph is the number of edges in the graph having endpoints at that vertex, unless the vertex is the endpoint of any loops, in which case each loop edge is considered to add 2 to the degree of the vertex.

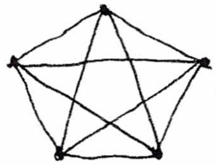
A **regular** graph is a graph in which all vertices have the same degree. A **k-regular** graph is a graph in which all vertices have degree k.
Comment: Note that adding an edge to a graph increases the sum of the degrees of all vertices by 2; therefore, the sum of the degrees of the vertices in a graph is always even. As an immediate consequence, all k-regular graphs with odd k have an even number of vertices.

The **complete graph** on n vertices, denoted K_n, is the simple graph on n vertices having the largest possible number of edges. Complete graphs on n vertices are always (n-1)-regular, and have girth 3 if $n \geq 3$.

A **(k,g)-graph** is a k-regular graph with girth g. As an example, the following graph, called the octahedron, is a (4,3)-graph on 6 vertices:



A **(k,g)-cage** is a (k,g)-graph with the fewest number of vertices of all (k,g)-graphs. The octahedron is not a (4,3)-cage, as illustrated by the following (4,3)-graph on 5 vertices:



This second graph is a (4,3)-cage, as will be shown in the next section.
Note that if a k-regular graph has n vertices, then it has kn/2 edges, if a k-regular graph has kn/2 edges then it has k vertices. So minimizing the number of vertices of a (k,g)-graph is equivalent to minimizing the number of edges of a (k,g)-graph.

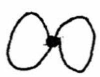
This paper will show that there exists a (k,g)-graph for any pair of k,g except when k=1. Thus, finding (k,g)-cages is always possible for any such pair of natural numbers.

II Simpler types of cages

k=1
There is no (1,g)-graph for any finite g:
Any vertex contained in a cycle has degree at least 2. Therefore, a 1-regular graph contains no cycles. So by definition a 1-regular graph has infinite girth, so there can be no (1,g)-graph for any finite g.

k=2
The (2,g)-cages are the cycles of length g:
Any graph with girth g contains a cycle of length g. So the smallest graph with girth g is itself a cycle of length g. This cycle is 2-regular, so it must be the (2,g)-cage.

g=1
A graph has girth 1 if and only if it contains a loop. If k is even, then a (k,1)-graph can be constructed using k/2 loops on a single vertex. By definition every graph contains at least 1 vertex, so these graphs are (k,1)-cages. For example, below is the (4,1)-cage:



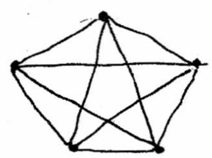
If k is odd, a (k,1)-graph can be constructed by taking 2 vertices, adding (k-1)/2 loops to each vertex, and adding a single edge between the two vertices. This graph is a (k,1)-cage, as by a previous comment, all k-regular graphs with odd k have an even number of vertices, hence at least 2.
An example of this construction for a (5,1)-cage:

This is not the unique (5,1)-cage, as illustrated by the following (5,1)-graph on 2 vertices:

These two graphs illustrate the general point that (k,g)-cages are not always unique.

g=2
In order to have a cycle of length 2, a graph must have at least 2 vertices. This lower bound is achievable for any k>1 by simply taking 2 vertices and adding k edges between them. For example, the (3,2)-cage is below:

g=3
All graphs with girth 3 are simple, and any k-regular simple graph must have at least k+1 vertices, since a vertex with degree k in a simple graph must have k distinct neighbors. This lower bound is achieved by the complete graph on k+1 vertices. An example, the (4,3)-cage was previously illustrated:



Finding the (k,g)-cage for $k \geq 2$ and $g \geq 4$ is generally very difficult (Erdős and Sachs 1963). It is always possible, however, due to a theorem proved by Erdős and Sachs.

III The Erdős-Sachs existence proof of (k,g)-graphs

Erdős and Sachs (1963) proved the following theorem:

For all $k \geq 2$, $g \geq 4$, there exists a k-regular graph on 2m vertices with no cycles of length less than g, where

$$m = 2\sum_{t=1}^{g-2}(k-1)^t.$$

The idea behind the proof is first to construct a cycle of length 2m, proving that a 2-regular graph with no cycles of length less than g exists, and then to induct on k. This induction is a bit unusual in that while it inducts only on k, the base case depends on both k and g; thus, the induction is actually different for each pair of k and g.

Base case:
The cycle of length 2m is 2-regular, and

$$2\cdot 2\sum_{t=1}^{g-2}(2-1)^t = 4(g-2) > g \quad \text{when } g \geq 4.$$

Induction step:
Suppose that there exists a (k-1)-regular graph on 2m vertices with no cycles of length less than g. Then there exists a k-regular graph on 2m vertices with no cycles of length less than g.

Let G* be a graph with the following properties:

a) G* has 2m vertices,

b) All vertices in G* have degree k-1 or k,

c) G* has no cycles of length less than g,

d) Of all graphs with properties a), b), and c), G* has the maximum possible number of edges.

The set of graphs with properties a),b), and c) is nonempty because the (k-1)-regular graph of the induction assumption has these properties. Also, the graphs have girth at least 4, so they are simple; the number of edges in a simple graph on 2m vertices is at most $\binom{2m}{2}$ so the number of edges in G* is maximizable.

Then G* is k-regular. Proof:

Suppose by way of contradiction that G* is not k-regular. Then by property b), G* has at least 1 vertex of degree k-1. Then either G* has exactly 1 vertex of degree k-1, or G* has at least 2 vertices of degree k-1.

If G* has exactly 1 vertex of degree k-1, then G* has 2m-1 vertices of degree k, by property b). Then the degree sum of G* is (2m-1)k+(k-1) = 2mk-k+k-1=2mk-1, which is odd for all k. This is impossible by the earlier comment that the degree sum of any graph is even.

So we may suppose that G* has at least 2 vertices of degree k-1. Choose 2 of these vertices and label them x_1, x_2.

Then the following lemmas hold:

Lemma 1: All vertices of degree k-1 in G* are contained in the set $D(x_1, g-2) \cap D(x_2, g-2)$, where D(x,d) denotes the set of all vertices in G* of distance less than or equal to d from x.

Proof:

Suppose by way of contradiction that there is a vertex z of degree k-1 in G* which is not contained in the above set. Then z is distance > g-2 from either x_1 or x_2. Suppose without loss of generality that z is distance > g-2 from x_1. In this case, adding the edge (x_1,z) to the graph will not create a cycle of length less than g, contradicting the assumption that G* had the maximum possible number of edges.

Lemma 2: Using the same notation as Lemma 1, $|D(x_1,r)| \leq \sum_{t=0}^{g-2}(k-1)^t$

Proof by induction:

For r=0, the formula will evaluate to 1; there is 1 vertex distance ≤ 0 from x_1, namely x_1, itself.

For r=1, there are k vertices in the set $D(x_1,r)$, namely x_1 and its k-1 neighbors, which agrees with the formula.

For r>1, each step of increasing distance from x_1 will add at most $(k-1)^r$ vertices to the set – the neighbors of the vertices from the previous set. The vertices will eventually begin to overlap, of course; this formula is merely an upper bound.

So it follows from the lemmas that

$|D(x_1, g-2) \cup D(x_2, g-2)| \leq m$:

$|D(x_1, g-2) \cup D(x_2, g-2)| = |D(x_1, g-2)| + |D(x_2, g-2)| - |D(x_1, g-2) \cap D(x_2, g-2)|$ by inclusion-exclusion,

$\leq 2\sum_{t=0}^{g-2}(k-1)^t - |D(x_1, g-2) \cap D(x_2, g-2)|$ by Lemma 2,

$=2+2\sum_{t=1}^{g-2}(k-1)^t - |D(x_1, g-2) \cap D(x_2, g-2)|$ re-indexing the sum,

$\leq m + 2 - |D(x_1, g-2) \cap D(x_2, g-2)|$ by definition of m,

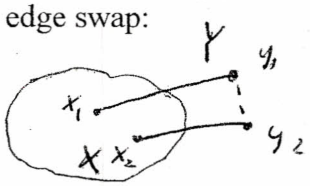
$\leq m + 2 - 2$ since $D(x_1, g-2) \cap D(x_2, g-2)$ contains at least x_1 and x_2

$=m$.

Then denote the set $D(x_1, g-2) \cup D(x_2, g-2)$ by X and denote the set of all other vertices in G* by Y. G* has 2m vertices, and by the above calculation, $|X| \leq m$. So $|Y| \geq m$.

Then there exists an edge between 2 vertices y_1, y_2 in set Y:
Suppose by way of contradiction that no such edge exists. Then by condition b) of G* and Lemma 1, all vertices in Y have degree k. So if there exists no edge between any two vertices in Y, each of the $k|Y|$ edges with endpoints in Y must have their other endpoint in X. So the average degree of vertices in X is at least $k|Y|/|X| \geq k$. Then either all vertices in X have degree at least k, or there exists a vertex in X with degree at least k+1. Both of these situations are impossible: not all vertices in X have degree k, as x_1 is in X and x_1 has degree k, and by condition b) no vertex in G* has degree greater than k.

So there is at least one edge between two vertices in set Y. Call this edge (y_1, y_2). Now, to complete the contradiction, it can be shown that G* is not maximal:
Consider the graph G' created from G* by removing edge (y_1, y_2) and adding edges (x_1, y_1) and (x_2, y_2). By definition the distance in G* from x_1, similarly x_2 to any $y \in Y$ is at least $g-1 > 1$, so the edges (x_1, y_1) and (x_2, y_2) were not in G*.
Then the graph G' meets conditions a),b) and c):
Condition a) is met, as no vertices have been added or deleted.
Condition b) is met, as deleting and adding one edge each to y_1, y_2 leaves their degrees unchanged, and by assumption x_1, x_2 had degree k-1 in G* so adding one edge to each will give them degree k in G'.
Condition c) is met as well, as by definition x_1, y_1 were at least distance g-1 apart in G*, similarly x_2, y_2 were at least distance g-1 apart in G*, and y_1, y_2 must be distance at least g-1 apart in G', as any shorter distance would imply the existence of a cycle of length < g in G*.
Illustration of the edge swap:

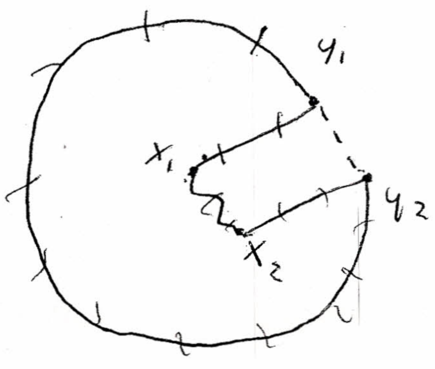


so adding the edges (x_1, y_1) and (x_2, y_2) to G* will not create any cycles of length < g. This augmentation of G* contradicts condition d), the maximality of G*.

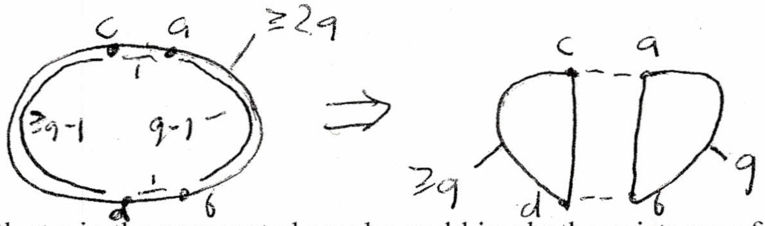
Therefore, G* must be k-regular. This completes the induction step.

IV Refining the Erdős-Sachs proof

The Erdős-Sachs proof does not establish the existence of a (k,g)-graph; it establishes the existence of a k-regular graph with girth *at least* g. However, it is possible to show the existence of a k-regular graph with girth *exactly* g using the Erdős-Sachs result:
To construct a k-regular graph with girth exactly g, first construct a k-regular graph G with girth at least 2g, which exists by the Erdős-Sachs result. Then note that such a graph contains at least 1 cycle: the base case of the induction is a cycle, and the only point in the induction at which an edge was removed was the final edge swap, which in fact preserves the existence of a cycle in the graph:
The edge swap removes one edge, breaking one cycle, but it creates a new cycle: by Lemma 1, x_1, x_2 are distance < g-1 apart, so there is a path from x_1 to x_2. The new cycle is illustrated below:



So G contains at least one cycle. Choose any cycle in G. This cycle has length at least 2g by assumption. Then a cycle of length g may be inserted into G by choosing 2 vertices a, b in the cycle that are distance g-1 apart and adding an edge between them, then deleting the neighboring edges (a,c), (b,d) in the cycle to preserve k-regularity, then adding another edge (b,d) to preserve k-regularity as illustrated below:



Any cycle of length < g in the augmented graph would imply the existence of a cycle of length < 2g in the original graph.

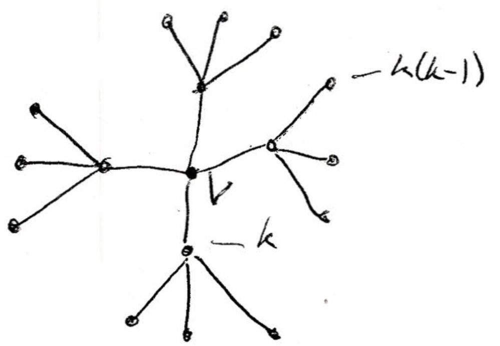
In fact, it has been shown, as mentioned in Exoo and Jajcay (2008), that the k-regular graph with girth at least g with the minimum possible number of vertices contains a cycle of length exactly g. This result not only refines the Erdős-Sachs proof as above, but also establishes the validity of 2m as an upper bound for the number of vertices in a (k,g)-cage.

V The Lower Bound and Moore Graphs

Erdős and Sachs also showed a lower bound for the number of vertices in a (k,g)-cage, where g>2. The lower bound is

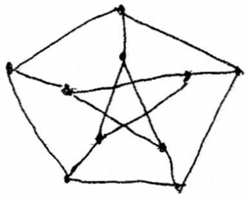
$$1 + \sum_{t=0}^{\lfloor \frac{g-3}{2} \rfloor} k(k-1)^t$$

Proof: Suppose that G is a (k,g)-cage. Then choose a vertex v in G. Of course there is 1 vertex distance 0 from v, namely v. By definition G is k-regular, so v has k neighbors. Then these k vertices have k neighbors in turn – v, and k-1 other vertices. So there are k vertices distance 1 from v, and k(k-1) vertices distance 2 from v. These successive circles of vertices spreading out from v cannot overlap until reaching distance $\lfloor \frac{g-3}{2} \rfloor$ from v; any earlier overlap would result in a cycle of length <g. The lower bound immediately follows.
Illustration of the tree surrounding v for k=4, g=5:



This lower bound immediately gives rise to the question of whether cages always achieve this lower bound, and if not, when the lower bound can be achieved. The lower bound cannot always be achieved, as demonstrated by Hoffman and Singleton (1960).

Cages which achieve the lower bound are known as **Moore graphs.** Hoffman and Singleton (1960) showed that in the case of girth 5, the the (k,5)-cage is a Moore graph only if k=2,3,7, or possibly 57. The (2,5)-cage is of course the cycle of length 5. The (3,5)-cage is the Petersen graph:



The (7,5)-cage is called the Hoffman-Singleton graph, which is too large to draw here; the lower bound formula shows that this graph has 1+7+6*7=50 vertices. The question of whether or not the (57,5)-cage is a Moore graph remains open; such a graph would have 1+57+56*57=3250 vertices.

VI Bibliography

- Paul Erdős and Horst Sachs: *Reguläre Graphen gegebener Taillenweite mit minimaler Knotenzahl.* Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe 12 (1963), 251-257
- A.J. Hoffman and R. R. Singleton: *On Moore graphs with diameters 2 and 3*, IBM J. Res. Develop. 4 (1960) 497-504
- Geoffrey Exoo and Robert Jajcay: *Dynamic Cage Survey.* The Electronic Journal of Combinatorics 15 (2008), #DS16