On graphs of high girth

Thesis submitted for the degree
Doctor of Philosophy

by

Shlomo Hoory

Submitted to the Senate of the Hebrew University
March 2002

On graphs of high girth

Thesis submitted for the degree
Doctor of Philosophy

by

Shlomo Hoory

Submitted to the Senate of the Hebrew University
March 2002

This work was carried out under the supervision of

Professor Nati Linial

I would like to thank Nati
for his encouragement and patience
throughout the years.

Abstract

The purpose of this work, is the study of extremal properties of graphs with high girth. More specifically, most of the work is centered around the Moore bound ([6], p.180), a simple lower bound on the number of vertices for a d-regular graph of a specified girth. The effort throughout the years, toward improving this trivial bound, has resulted only in minor improvements. Namely, the bound has been increased by 2 for girth larger that 12 and degree large than 2 (Biggs, Bannai and Ito [4, 7]). Graphs achieving equality in the Moore bound are called Moore graphs, or generalized polygons. The spectrum of these graphs is uniquely determined ([6], Theorem 23.6). These improvements are far from settling the extremal problem, since the best known upper bound (a construction of Lubotzky, Phillips and Sarnak, [16]) exceeds the Moore bound by a power of $3/2$.

The body of this work includes four chapters.

- About 30 years ago, Bollobás has asked whether the Moore bound holds also for irregular graphs. This problem appears in his classical book *Extremal Graph Theory* [8]. In this chapter, we answer this long standing open question in the affirmative. We prove that if the degree of regularity d in the Moore bound is replace by the average degree of the graph, the bound still holds. This is joint work with N. Alon and N. Linial, and was published in a paper titled *The Moore bound for irregular graphs*.

- The results of *The Moore bound for irregular graphs* are improved for bipartite graphs. The previously know results for girth 6 and 8 (of Reiman, de Caen, Székely and Neuwirth [20, 9, 18]) are generalized to an arbitrary (even) girth. This work appears in a paper titled *The size of bipartite graphs with a given girth*.

- We improve the spectral analysis of graphs whose size is at most twice the Moore bound. Also, we improve the best known lower bound (Friedman [12], and Winnie Li and Solé [15]) on the least possible second eigenvalue of a d-regular graph with n vertices and girth g. This work is based on a classical theory [14], characterizing distributions through their moments. This is a joint work with A. Amit and N. Linial, and was published as a paper titled *A Continuous Analogue of the Girth Problem*.

- A different approach to the girth problem, is through the study of lifts of d-regular graphs to their universal cover - the d-regular infinite tree. We show that the existence of small d-regular graphs of a specified girth, is equivalent to the existence of certain codes in the d-regular infinite tree. We prove that in the tree "perfect" codes exist, but these codes are usually not "graphic". We also give an explicit construction that is "nearly perfect" and "nearly graphic".

Contents

Chapter 1

Introduction

Definition: The girth of a graph G is the length of the shortest cycle in G (or infinity, if there are no cycles in G).

Let $g = g(n, d)$ be the largest possible girth of a d-regular graph of order $\leq n$.

Obviously $g(n, 1) = \infty$, and $g(n, 2) = n$. However, giving good bounds on $g(n, d)$ for any $d \geq 3$ is a notoriously difficult problem in extremal graph theory.

If we consider $d \geq 3$ fixed and growing n, the best asymptotic estimates known are:

$$2 \log_{d-1} n \cdot (1 + o(1)) \geq g(n, d) \geq \frac{4}{3} \log_{d-1} n \cdot (1 - o(1)). \tag{1.1}$$

The lower bound is due to Lubotzky, Phillips and Sarnak [16]. The Ramanujan graph they construct, in the bipartite case, has high girth. However, their construction (and its extensions) works only when the degree d is a prime power plus one. For other values of d, there are two lower bounds. The first lower bound is obtained by taking a d'-regular LPS graph of girth g, where $d' > d$ is the smallest possible (prime power plus 1), and deleting $d' - d$ perfect matchings. The bound obtained this way is:

$$g(n, d) \geq \frac{4}{3} \cdot \frac{\log(d-1)}{\log(d'-1)} \cdot \log_{d-1} n \cdot (1 - o(1)).$$

The second is due to Erdös and Sachs [11], and is given in Theorem A.2. Their lower bound gives:

$$g(n, d) \geq \log_{d-1} n \cdot (1 - o(1)). \tag{1.2}$$

By known results about the density of primes, the first lower bound is always the better one. More about the construction of graphs with high girth can be found in Appendix 6.

The upper bound in (1.1) is the Moore bound, which is a simple counting argument. To present this bound, it is more convenient to turn (1.1) around, and define $n(d, g)$ as the least possible number of vertices n in a d-regular graph of girth g. In this form, the Moore bound has the following form:

Theorem 1.1 *(The Moore bound)*

$$n(d, g) \geq n_0(d, g), \tag{1.3}$$

where

$$n_0(d, 2r+1) = 1 + d \cdot \sum_{i=0}^{r-1} (d-1)^i,$$
$$n_0(d, 2r) = 2 \cdot \sum_{i=0}^{r-1} (d-1)^i.$$

Proof:

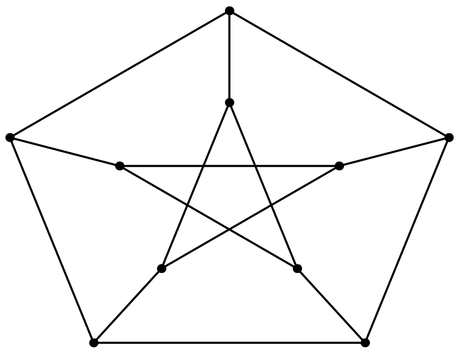
Consider a d-regular graph with n vertices and girth g. The bound immediately follows from the following observation:

- If the girth is odd $g = 2r + 1$, then a radius r ball around any vertex is a tree.
- If the girth is even $g = 2r$, then a radius $r - 1$ ball around any edge is a tree.

∎

A d-regular graph of girth g and $n_0(d, g)$ vertices is called a *Moore Graph* for odd g and a *Generalized Polygon* if g is even. Excluding trivial cases, we assume that $d \geq 3$ and $g \geq 5$. It is known (and this requires considerable work) that Moore graphs exist only for $g = 5$. Generalized Polygons exist only for $g \in \{4, 6, 8, 12\}$ (see [6] Theorem 23.6, and the survey [26]). Moore graphs of girth 5 can exist only for $d \in \{2, 3, 7, 57\}$ (a nice proof can be found in [3]). Of these possibilities, there are graphs with $d = 2$ (a pentagon), $d = 3$ (Petersen), $d = 7$ (Hoffman Singleton [26]), and the existence of such a graph with $d = 57$ is still open.

The smallest non-trivial example for a Moore graph is the Petersen graph, a 3-regular graph of girth 5 with 10 vertices.



The Petersen Graph

The known proof for the non-existence of Moore graphs and Generalized Polygons, uses algebraic properties of the adjacency matrix to show that their spectra are completely determined by the girth requirement. In fact, when one derives explicit formulas for the eigenvalues and their multiplicities, one finds that for almost all values of g the multiplicities are not integral, and therefore no such graphs exist. Later works [4, 7, 13] that improve these bounds slightly (by additive 1 or 2) use similar methods.

We now turn to several generalizations of the Moore bound. It is natural to ask for the situation in non-regular graphs.

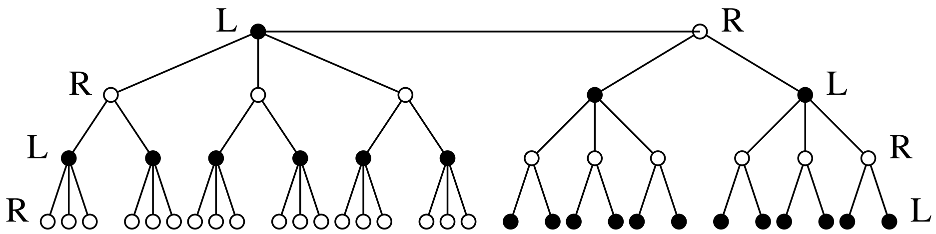
1.1 The Moore bound for bi-regular bipartite graphs

The simplest generalization of the Moore bound to irregular graphs, is for graphs that are bipartite and bi-regular.

Consider a bipartite graph G where all the vertices on the right and left have degree d_R and d_L respectively, and let $g = 2r$ be the girth of G. We can repeat the argument of the Moore bound also in this scenario and get the following bound on n_R and n_L, the number of vertices on the right and left side of G, respectively:

Theorem 1.2

$$\begin{aligned} n_L &\geq \sum_{i=0}^{r-1} (d_R - 1)^{\lceil \frac{i}{2} \rceil} \cdot (d_L - 1)^{\lfloor \frac{i}{2} \rfloor}, \\ n_R &\geq \sum_{i=0}^{r-1} (d_R - 1)^{\lfloor \frac{i}{2} \rfloor} \cdot (d_L - 1)^{\lceil \frac{i}{2} \rceil}. \end{aligned}$$



A radius 3 ball around an edge ($d_L = 4$ and $d_R = 3$).

The proof is straightforward. It follows from the fact that the radius $r - 1$ ball around an edge is a tree. Since G is bipartite, the number left (or right) vertices of G can be counted separately in the tree to give the required lower bound.

1.2 The Moore bound for irregular graphs

When the graph is a general graph the situation is more complex. Ball sizes are no longer uniform, and may depend strongly upon the choice of their center. It is not at all obvious how to find a good lower bound even on the size of the largest ball of radius r.

One obvious lower bound is:

$$n \geq n_0(\delta, g), \tag{1.4}$$

where δ is the minimal degree of the graph. But such a bound is far from satisfactory.

A better argument is obtained from the well known fact ([10] proposition 1.2.2):

Theorem 1.3 *Every graph G with average degree $d > 0$ has a subgraph H with minimal degree $> d/2$.*

From (1.4) and Theorem 1.3, we deduce that:

$$n \geq n_0(\frac{d}{2}, g), \tag{1.5}$$

where d is the average degree of the graph.

Chapter 1 of this work deals with this question, and there we prove the following theorem:

Theorem 1.4 *Any graph with average degree $d \geq 2$, girth g and n vertices, satisfies:*

$$n \geq n_0(d, g), \tag{1.6}$$

a significant improvement over (1.5). Any further improvement of the theorem would yield an improvement of the Moore bound for regular graphs (1.3). The inequality (1.6) was conjectured about 25 years ago by Bollobás, and appears in his book "Extremal graph theory" [8], p.163 problem 10.

Also, it should be further remarked, that our results in Chapter 1 are stronger than (1.6), if there is variance in G's vertex degrees. As an example, consider a graph with average degree 2.5. This graph is necessarily far from being regular, and our bound indeed implies here:

$$n \geq n_0(2.5157\ldots, g) > n_0(2.5, g).$$

1.3 The Moore bound for irregular bipartite graph

Let G be a bipartite graph where the average degree of the vertices on the right and left side is $\overline{d_R}$ and $\overline{d_L}$ respectively, and let $g = 2r$ be the girth of G. Let n_R and n_L be the number of vertices on the right and left sides respectively.

In chapter 2, we prove that the analogue of Theorem 1.2 holds, also with $\overline{d_R}$ and $\overline{d_L}$ instead of d_R and d_L.

1.4 A continuous version of the Moore bound

We now return to the d-regular scenario, and we would like to examine the Moore bound from a different perspective.

Let G be a d-regular graph of girth g and size n, and let A be the adjacency matrix of G. Since A is a real and symmetric matrix, it has a basis of eigenvectors, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. $\lambda_1 = d$ is Perons eigenvalue of A, and $\lambda_n \geq -d$.

Consider the expression trace(A^k). For integers $k < g$, this value is completely determined:

$$\text{trace}(A^k) = n \cdot t_k,$$

where t_k is the number of closed walks of length k from a fixed vertex to itself. This is also, the number of such walks in the d-regular infinite tree that start and end at some fixed vertex v.

These *tree numbers* t_i play an important role in the following discussion. We give here a list of the first tree numbers. Note, that only even tree numbers are given, since the odd ones vanish.

$$\begin{aligned} t_0 &= 1 \\ t_2 &= d \\ t_4 &= 2d^2 - d \\ t_6 &= 5d^3 - 6d^2 + 2d. \end{aligned}$$

Since,

$$\text{trace}(A^k) = \sum_{i=1}^{n} \lambda_i^k,$$

we know the k'th moment of the eigenvalues for $k = 0, 1, \ldots, g-1$.

We ask, for what values of n, d, and g this set of equations in $\lambda_1, \ldots, \lambda_n$ has a solution. Since it is difficult to work with such a set of equations, we consider *distributions*, and ask:

Is there a distribution μ supported on $[-d, d]$ satisfying:

$$\int_{x=-d}^{d} x^k d\mu(x) = nt_k \qquad \text{for } k = 0, 1, \ldots, g-1? \tag{1.7}$$

The problem with this formulation, is that it has a solution for any d, g and n. Such a solution can be obtained by taking the eigenvalues of any d-regular graph with girth $\geq g$ and size n', and defining μ to have a mass n/n' concentrated at each of the n' eigenvalues.

To get a formulation of the problem that doesn't have such a trivial solution, we recall that $\lambda_1 = d$ and subtract it from nt_k. This yields:

$$\int_{x=-d}^{d} x^k d\mu(x) = nt_k - d^k \qquad \text{for } k = 0, 1, \ldots, g-1. \tag{1.8}$$

We prove in chapter 3 the following theorem:

Theorem 1.5 *For any integers d and g, the smallest n for which there is a measure μ satisfying (1.8) is equal to $n_0(d, g)$.*

The proof uses tools from a classical theory, of *the Markov moment problem.* These tools enable one to determine, whether there exists a distribution with a specified moments that is supported on a segment $[a, b]$.

1.5 The Alon Boppana lower bound

During the last two decades, much progress has been made in understanding the spectra of graphs, and in particular the relation between the second eigenvalue and the graph's expansion properties. It seems, therefore, natural to ask for the least possible value Λ of the second largest eigenvalue (in absolute value), in a d-regular graph of girth g, size n, and diameter D.

A. Nilli [19], following earlier work by N. Alon and R. Boppana, proved that:

$$\Lambda \geq 2\sqrt{d-1}(1-o(1)),$$

where the $o(1)$ term, refers to growing n and fixed d.

The best known lower bound on Λ is due to Friedman [12]:

$$\Lambda \geq 2\sqrt{d-1}(1-O(\frac{1}{D^2})). \tag{1.9}$$

Since $D \geq \Omega(\log n)$, and $g \leq 2D+1$, the error term of (1.9) is also $O(1/\log^2 n)$ or $O(1/g^2)$. For graphs of high girth, (1.9) was improved by Solé and Li [23, 15]. Their bound is:

$$\Lambda \geq 2\sqrt{d-1}(1-O(\frac{1}{g^2})). \tag{1.10}$$

The constants in the $O()$ notation are such, that (1.10) is a better bound than (1.9) when $g > D$. In section 4 of chapter 4, we improve both (1.10) and (1.9) in the range $2D+1 > g \geq D$. The bound follows from the general framework, described in section 1.4, considering the moments of the eigenvalue distribution for graphs of high girth. If $g = 2D+1$, the graph must be a Moore graph, and its eigenvalue distribution is completely determined, and our lower bound coincides with (1.10). Keeping g and d constant, our lower bound is monotone in n, and therefore, is indeed an improvement over (1.9) at least as long as $g \geq D$. However, for large n, (1.9) seems to be the better bound.

1.6 Does high girth imply expansion ?

The question in the title must certainly be severely qualified, to have an affirmative answer. Indeed, high girth does not even imply connectivity, as demonstrated by taking two disjoint copies of a d-regular graph with high girth (Connected examples that are still poor expanders, can be easily obtained by modifying this example). We discuss the question, "How large can the second eigenvalue be in a d-regular graph of size n?" in section 5 of chapter 4. Phrased differently, we ask, for the minimal number of vertices n in a d-regular graph of girth g that has λ as (some) eigenvalue. The best known lower bound for $g = 2r+1$, is due to Biggs [5]:

$$n - n_0(d,g) \geq |P_r(\lambda)|, \tag{1.11}$$

where P_r is a certain degree r polynomial, and $n_0(d,g)$ is the Moore bound (1.3). It is not difficult to see that as long as $n < n_0(d,g+1)$, (1.11) gives an upper bound on λ, and therefore we know the graph is an expander. We explore a similar problem for the eigenvalue distribution. Our bound coincides with (1.11) when $\lambda = \pm d$, but is poorer for other values of λ. For small values of g (we can determine the exact value of bound), the bound is weaker. It seems that in order to improve (1.11), moment information alone is not enough, and more "graphical" information must be incorporated.

1.7 Girth and Codes

We recall here, a simple fact from Topology:
The d-regular infinite tree is the universal covering space of any d-regular undirected graph (both are regarded as one dimensional complexes).

Therefore, given a d-regular graph G with n vertices and girth g, there is a cover map $\phi : V(\mathbf{T_d}) \to V(G)$, where $\mathbf{T_d}$ denotes the d-regular infinite tree. Such a map can be regarded as a vertex coloring of $\mathbf{T_d}$ using n colors where each $x \in V(\mathbf{T_d})$ is colored $\phi(x)$. This coloring ϕ has the property that:

$$\text{if } \phi(x) = \phi(y) \text{ for some } x \neq y \text{ then } d_T(x, y) \geq g. \tag{1.12}$$

Therefore, a necessary condition for the existence of a graph with parameters d, g and n, is that there exists a vertex coloring of $\mathbf{T_d}$ satisfying (1.12).

We show in Chapter 5, that the least number of colors n, needed to color $\mathbf{T_d}$ in a way that satisfies (1.12) is exactly $n_0(d, g)$. In other words, we show that the Moore bound holds also here ($n \geq n_0(d, g)$), and we prove that there are colorings of the tree satisfying (1.12) that use only $n_0(d, g)$ colors.

Although such "perfect" colorings exist, we know that there are no Moore graphs of girth > 5 and degree > 2. Therefore, a natural question that arises is, when is a coloring of $\mathbf{T_d}$ "graphic". Let ϕ be a coloring of $\mathbf{T_d}$, and define $E(\phi)$ as the set of all unordered pairs $(\phi(x), \phi(y))$ over x, y that are neighbors in $\mathbf{T_d}$. As we observe in Chapter 5, an n-vertex coloring ϕ of $\mathbf{T_d}$ satisfying (1.12) for some $g \geq 3$ is graphic, iff

$$|E(\phi)| = \frac{nd}{2}.$$

We further exhibit an explicit coloring ψ satisfying (1.12), that uses only $n = g \cdot (d-1)^{\lfloor \frac{g}{2} \rfloor}$ colors, for which $E(\psi)$ has size $(d-1) \cdot n$ (only about twice as many as a graphic coloring).

The approach discussed above, can be viewed as a problem of constructing good codes. In particular, we seek a "large" $C \subset V(\mathbf{T_d})$ of which any two distinct vertices $x, y \in C$ satisfy $d_T(x, y) \geq g$. (The meaning of "large" would be made more precise in Chapter 5.) Viewed this way, the non-existence of d-regular Moore graphs is equivalent to the non-existence of certain "perfect" codes in $\mathbf{T_d}$.

Chapter 2

The Moore bound for irregular graphs

Graphs and Combinatorics (2002) 18:53–57

Graphs and Combinatorics

The Moore Bound for Irregular Graphs

Noga Alon[1*], Shlomo Hoory[2], and Nathan Linial[3†]

[1] Mathematics and Computer Science, Sackler Faculty of Exact Sciences,
Tel Aviv University, Tel Aviv, Israel. e-mail: noga@math.tau.ac.il
[2,3] Institute of Computer Science, Hebrew University, Jerusalem 91904, Israel
e-mails: [2]shlomoh; [3]nati@cs.huji.ac.il

Abstract. What is the largest number of edges in a graph of order n and girth g? For d-regular graphs, essentially the best known answer is provided by the Moore bound. This result is extended here to cover irregular graphs as well, yielding an affirmative answer to an old open problem ([4] p. 163, problem 10).

What is the maximal number of edges in a graph with n vertices and girth g? Put differently, what is the least number of vertices $n = n(d, g)$ in a graph of girth g and an average degree $\geq d$?

If d is an integer, it is natural to compare the lower bound on n, to what is known for d-regular graphs. The lower bound for d-regular graphs arises from the fact the ball of radius $\lfloor \frac{g-1}{2} \rfloor$ around a vertex or an edge (depending on the parity of g) is a tree. This simple argument yields the so-called *Moore bound* denoted by $n_0(d, g)$ (see [2], p. 180), where:

$$n_0(d, 2r+1) = 1 + d\sum_{i=0}^{r-1}(d-1)^i$$

$$n_0(d, 2r) = 2\sum_{i=0}^{r-1}(d-1)^i$$

This simple lower bound is difficult to improve upon. There are graphs that achieve equality, but only for $g \in \{3, 4, 5, 6, 8, 12\}$, (see [2] Theorem 23.6, and the survey [7]). Some further improvement (only by +2) was achieved in [1, 3, 6]. In this note we show that the Moore bound holds also for irregular graphs. This

* Supported in part by a US-Israel BSF grant, by a grant from the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University
† Supported in part by grants from the US-Israel Binational Science Fund and from the Israel Science Foundation

result has been conjectured long ago and it appears in Bollobás' "Extremal graph theory" [4], p. 163 problem 10.

Theorem 1. *The number of vertices n in a graph of girth g and average degree at least $d \geq 2$, satisfies*:

$$n(d,g) \geq n_0(d,g).$$

Note 1. When d is an integer Theorem 1 is exactly the Moore bound. If d is not an integer, or when the degree sequence of the graph is known, our proof yields a better bound (see the proof for details).

Note 2. This bound is exponentially bigger (in g) than the previously known lower bound on n given in [5] (the result there is stated as an upper bound on g):

$$n \geq 4\left\lfloor \frac{d}{2} \right\rfloor^{\frac{g-2}{2}}.$$

Proof. Let G be a graph with n vertices, girth g and an average degree $\bar{d} \geq d$. Vertices of degree 0 and 1 can be ignored: If we delete all the vertices of degree <2, the resulting graph G' has $n' \leq n$ vertices, girth $g' = g$ and an average degree $\bar{d}' \geq \bar{d} \geq d$. If we prove the theorem for G', then by monotonicity of n_0:

$$n \geq n' \geq n_0(\bar{d}', g') \geq n_0(d,g),$$

proving the theorem for G. Henceforth we assume then w.l.o.g. that all vertex degrees in G are at least 2.

Consider each edge of G as a pair of directed edges, and let A be the $n\bar{d} \times n\bar{d}$ matrix indexed by directed edges:

$$(A)_{\vec{uv},\vec{wz}} = \begin{cases} 1 & v = w \text{ and } u \neq z \\ 0 & \text{otherwise} \end{cases}$$

Also, let the matrix P be defined by:

$$(P)_{\vec{uv},\vec{wz}} = \begin{cases} \frac{1}{d_v - 1} & v = w \text{ and } u \neq z \\ 0 & \text{otherwise} \end{cases}$$

where d_v is the degree of the vertex v. The matrix P is the transition matrix of the *non-returning random walk*. The state space of this walk is the set of directed edges, and if at time t the walk is at the edge $\vec{uv}$ (we sometime say that the walk is at the vertex v coming from u), then at time $t+1$ we move to an edge $\vec{vw}$, where w is a random neighbor of v that is not u.

Let $x = \frac{1}{n\bar{d}}\bar{1}$ be the uniform distribution on the directed edges. It can be easily verified that x is a stationary distribution for the non-returning random walk, i.e., $xP = x$.

Let $\Omega_{e,l}$ be the set of all length l non-returning walks starting at the edge e, and let the number of these walks be $N_{e,l} = |\Omega_{e,l}|$. Define also, N_l as:

$$N_l = \sum_e x_e N_{e,l} = \langle xA^l, \bar{1} \rangle. \tag{1}$$

For any walk $\omega = \{e_0 = e, e_1, \ldots, e_l\} \in \Omega_{e,l}$, denote by $n_f(\omega)$, the number of occurrences of f among $e_0, e_1, \ldots, e_{l-1}$. Finally, let $p(\omega)$ be the probability of the walk ω, namely:

$$p(\omega) = \prod_{\vec{uv}} (d_v - 1)^{-n_{\vec{uv}}(\omega)}. \tag{2}$$

We can now give a lower bound on N_l. First we rewrite (1) and then use the inequality of the arithmetic and geometric means:

$$N_l = \sum_e x_e \sum_{\omega \in \Omega_{e,l}} p(\omega) \frac{1}{p(\omega)} \geq \prod_e \prod_{\omega \in \Omega_{e,l}} \left(\frac{1}{p(\omega)} \right)^{x_e p(\omega)}.$$

Using (2), this becomes:

$$N_l \geq \prod_{\vec{uv}} (d_v - 1)^{\sum_e x_e \sum_{\omega \in \Omega_{e,l}} n_{\vec{uv}}(\omega) p(\omega)}.$$

The sum in the exponent of the rhs is exactly the expected number of visits to the edge $\vec{uv}$, made by a non-returning random walk of length l starting with the initial distribution x:

$$\begin{aligned} N_l &\geq \prod_{\vec{uv}} (d_v - 1)^{\sum_{k=0}^{l-1} (xP^k)_{\vec{uv}}} = \prod_{\vec{uv}} (d_v - 1)^{\sum_{k=0}^{l-1} \frac{1}{n\bar{d}}} \\ &= \prod_{\vec{uv}} (d_v - 1)^{\frac{l}{n\bar{d}}} = \left(\prod_v (d_v - 1)^{\frac{d_v}{n\bar{d}}} \right)^l = \Lambda^l \end{aligned}$$

where $\Lambda = \prod_v (d_v - 1)^{\frac{d_v}{n\bar{d}}}$. Since the function $f(d) = (d-1)^d$ is log convex for $d \geq 2$, it implies that $\Lambda \geq \bar{d} - 1 \geq d - 1$.

To prove the theorem, we consider two cases:
• If the girth is odd $g = 2r + 1$, consider the matrix:

$$C = \sum_{i=0}^{r} A^i + M \sum_{i=0}^{r-1} A^i,$$

where M is the $n\bar{d} \times n\bar{d}$ permutation matrix that matches each edge to its inverse:

$$(M)_{\vec{uv},\vec{wz}} = \begin{cases} 1 & v = w \text{ and } u = z \\ 0 & \text{otherwise} \end{cases}$$

For every $\vec{uv}$ and z, the number $\sum_{w:w\sim z} C_{\vec{uv},\vec{wz}}$ is either zero or one. This follows from the fact that $\sum_{w:w\sim z}\left(\sum_{i=0}^{r} A^i\right)_{\vec{uv},\vec{wz}}$ is the number of non-returning walks from v to z of length $\leq r$ whose first move is not $\vec{vu}$, and that $\sum_{w:w\sim z}\left(M\sum_{i=0}^{r-1} A^i\right)_{\vec{uv},\vec{wz}}$ is the number of non-returning walks from v to z of length $\leq r$ whose first move is $\vec{vu}$. Therefore $\sum_{w:w\sim z} C_{\vec{uv},\vec{wz}}$ is the total number of non-returning walks from v to z of length $\leq r$. There is at most one such walk since no cycle of length $<g$ exists in the graph.

Therefore for every edge $\vec{uv}$,

$$n \geq \sum_{\vec{wz}} C_{\vec{uv},\vec{wz}},$$

and therefore,

$$\begin{aligned} n &\geq \langle xC, \bar{1}\rangle = \sum_{i=0}^{r} N_i + \sum_{i=0}^{r-1} N_i \\ &\geq \sum_{i=0}^{r} \Lambda^i + \sum_{i=0}^{r-1} \Lambda^i = n_0(\Lambda + 1, g) \geq n_0(d, g). \end{aligned}$$

The last inequality follows from our observation that $\Lambda \geq d - 1$ and the monotonicity of n_0.

• If the girth is even $g = 2r$, then consider the matrix:

$$D = (I + M)\sum_{i=0}^{r-1} A^i.$$

By similar considerations:

$$n \geq \langle xD, \bar{1}\rangle \geq n_0(\Lambda + 1, g) \geq n_0(d, g). \qquad \square$$

Note 3. The above proof actually shows that $n \geq n_0(\Lambda + 1, g)$, where $\Lambda = \prod_v (d_v - 1)^{\frac{d_v}{nd}}$. The graph of n/g cycles of length g having a common vertex, demonstrates that this bound is nearly tight.

References

1. Bannai, E., Ito, T.: Regular graphs with excess one. Discrete Math. **37**, 147–158 (1981)
2. Biggs, N.: Algebraic graph theory. (2nd edn.) Cambridge: Cambridge University Press 1993

3. Biggs, N., Ito, T.: Graphs with even girth and small excess. Math. Proc. Cambridge Philos. Soc. **88**, 1–10 (1980)
4. Bollobás, B.: Extremal graph theory. London: Academic Press [Harcourt Brace Jovanovich Publishers] 1978
5. Dutton, R. D., Brigham, R. C.: Edges in graphs with large girth. Graphs Comb. **7**, 315–321 (1991)
6. Kovács, P.: The nonexistence of certain regular graphs of girth 5. J. Comb. Theory Ser. B **30**, 282–284 (1981)
7. Wong, P. K.: Cages—a survey. J. Graph Theory **6**, 1–22 (1982)

Received: June 27, 2000
Final version received: July 3, 2001

Chapter 3

The size of bipartite graphs with a given girth

The size of bipartite graphs with a given girth

Shlomo Hoory *

February 3, 2002

Abstract

What is the maximum number of edges in a bipartite graph of girth g whose left and right sides are of size n_L, n_R ? We generalize the known results for $g = 6, 8$ to an arbitrary girth.

Running head: Girth & Bipartite graphs.

Keywords: graph, bipartite, girth, random walk.

*Institute of Computer Science, Hebrew University Jerusalem 91904 Israel *shlomoh@cs.huji.ac.il*

1 Introduction

Consider a bipartite graph with n_L vertices on the left side, n_R vertices on the right, m edges, and girth g. In this paper we present a new lower bound for n_L and n_R that holds for an arbitrary girth g. Our bound (Theorem 1) is best presented as a function of the left and right average degrees, $d_L = \frac{m}{n_L}$ and $d_R = \frac{m}{n_R}$ and of the girth $g = 2r$:

$$n_L \geq \sum_{i=0}^{r-1} (d_R - 1)^{\lceil i/2 \rceil} (d_L - 1)^{\lfloor i/2 \rfloor} \tag{1}$$

$$n_R \geq \sum_{i=0}^{r-1} (d_L - 1)^{\lceil i/2 \rceil} (d_R - 1)^{\lfloor i/2 \rfloor}. \tag{2}$$

This bound can be viewed as a generalization of the Moore bound for d-regular graphs (see [2], p. 180). If the given bipartite graph is bi–regular, i.e. all the vertices on the left (resp. right) are of degree d_L (resp. d_R), then (1), (2) follow from an easy argument similar to the Moore bound. In this work we show that this bound applies also to general bipartite graphs that are not necessarily bi–regular. Our bound and its proof can be viewed as an adaptation of the main argument of [1] to the case of bipartite graphs.

Previous relevant bounds for bipartite graphs were known for girth $4, 6, 8$. For $g = 4$, (1) reduces to the trivial statement $n_L \geq d_R$. For $g = 6$ the best lower bound is due to Reiman ([6], see [3], p.312 Theorem 2.6). His bound is the same as the bound given by (1):

$$n_L \geq 1 + (d_R - 1) + (d_R - 1) \cdot (d_L - 1).$$

For the case of $g = 8$ the best bound is due to Neuwirth [5] (following work of de Caen and Székely [4]). This bound can be shown to be the same as (1):

$$n_L \geq 1 + (d_R - 1) + (d_R - 1) \cdot (d_L - 1) + (d_R - 1)^2 \cdot (d_L - 1).$$

2 The Theorem

Theorem 1 *Let $G = (V_L, V_R, E)$ be a bipartite graph of girth $g = 2r$, with $n_L = |V_L|$ and $n_R = |V_R|$ the number of vertices on the left and right sides, and $m = |E|$ the number of edges. Assume further that all vertex degrees in G are ≥ 2. Then:*

$$n_L \geq \sum_{i=0}^{r-1} \Lambda_R^{\lceil i/2 \rceil} \Lambda_L^{\lfloor i/2 \rfloor}$$

$$n_R \geq \sum_{i=0}^{r-1} \Lambda_L^{\lceil i/2 \rceil} \Lambda_R^{\lfloor i/2 \rfloor},$$

where $\Lambda_R = \prod_{v \in V_R} (d_v - 1)^{d_v/m}$, $\Lambda_L = \prod_{v \in V_L} (d_v - 1)^{d_v/m}$ and d_v is the degree of the vertex v.

Corollary 1 *(1) and (2) follow from Theorem 1. Because the function $(x-1)^x$ is log-convex for $x \geq 2$, and all vertex degrees are ≥ 2, we get $\Lambda_L = \prod_{v \in V_L}(d_v - 1)^{d_v/m} \geq d_L - 1$ and $\Lambda_R = \prod_{v \in V_R}(d_v - 1)^{d_v/m} \geq d_R - 1$.*

Proof:

The main tool in this proof are *non-returning walks* on the vertices of the graph. If at time t the walk is at vertex x_t, then at time $t+1$ it can advance to any neighbor of x_t other than x_{t-1}. To translate this description into algebraic terms, we record the state of the walk at time t (it is at x_t having arrived from x_{t-1}) by the directed edge $\overrightarrow{x_{t-1}x_t}$.

Regard each (undirected) edge of G as a pair of directed edges, and let E_{RL} and E_{LR} be the sets of right to left and left to right edges respectively. Then $|E_{LR}| = |E_{RL}| = m$. We define the $m \times m$ matrice A:

$$A_{\vec{uv},\vec{wz}} = \begin{cases} 1 & v = w \text{ and } u \neq z \\ 0 & \text{otherwise} \end{cases}$$

where $\vec{uv} \in E_{LR}$ and $\vec{wz} \in E_{RL}$. The matrix A (resp. A^t) records the possible right to left (resp. left to right) steps of the non-returning walk. Also, we define the length of a walk $\omega = (e_0, e_1, \ldots, e_l)$ to be l.

For any $\vec{uv} \in E_{LR}$, consider the set of all non-returning walks that either:

1. start at $\vec{uv}$ and have an even length $\leq r-1$,
2. start at $\vec{vu}$ and have an odd length $\leq r-1$.

All these walks must end in distinct vertices in V_R because of the girth requirement. Therefore the number of these walks is a lower bound on n_R.

To proceed, denote by $\Omega_{e,2i}$ the set of length $2i$ non-returning walks starting from $e \in E_{LR}$, and the number of such walks by $N_{e,2i}$. Then:

$$N_{e,2i} = |\Omega_{e,2i}| = ((AA^t)^i 1_{LR})_e,$$

where 1_{LR} is the all ones vector defined on E_{LR}. Similarly, denote by $\Omega_{e,2i+1}$ the set of length $2i+1$ non-returning walks starting from $e \in E_{RL}$, and the number of such walks by $N_{e,2i+1}$. Then:

$$N_{e,2i+1} = |\Omega_{e,2i+1}| = (A^t(AA^t)^i 1_{LR})_e$$

We can now write our lower bound as:

$$n_R \geq \sum_{i=0}^{\lfloor \frac{r-1}{2} \rfloor} N_{\vec{uv},2i} + \sum_{i=0}^{\lfloor \frac{r-2}{2} \rfloor} N_{\vec{vu},2i+1} \tag{3}$$

Since this inequality holds for every $\vec{uv} \in E_{LR}$, it follows that:

$$n_R \geq \sum_{i=0}^{\lfloor\frac{r-1}{2}\rfloor} \sum_{e\in E_{LR}} (\pi_{LR})_e N_{e,2i} + \sum_{i=0}^{\lfloor\frac{r-2}{2}\rfloor} \sum_{e\in E_{RL}} (\pi_{RL})_e N_{e,2i+1} \tag{4}$$

where $\pi_{LR} = 1^t_{LR}/m$ is the uniform probability distribution on E_{LR}, and π_{RL} is the uniform probability distribution on E_{RL}. When no confusion may arise, we will drop the LR and RL subscripts.

The required lower bound on n_R is obtained using the following Lemma. The lower bound on n_L follows by symmetry. ∎

Lemma 1 *For any $i \geq 0$ the following two inequalities hold:*

$$\sum_{e\in E_{LR}} \pi_e N_{e,2i} \geq \Lambda_R^i \Lambda_L^i \tag{5}$$

$$\sum_{e\in E_{RL}} \pi_e N_{e,2i+1} \geq \Lambda_R^i \Lambda_L^{i+1} \tag{6}$$

Proof:

We regard $\Omega_{e,2i}$ as a probability space. For $\omega = (e = e_0, e_1, \ldots, e_{2i}) \in \Omega_{e,2i}$ we let $p_e(\omega)$ be the probability that a length $2i$ non-returning random walk starting at e will be ω. The value of $p_e(\omega)$ is given by:

$$p_e(\omega) = \prod_{\vec{wz}\in E_{LR}} (d_z - 1)^{-n_{\vec{wz}}(\omega)} \cdot \prod_{\vec{zw}\in E_{RL}} (d_w - 1)^{-n_{\vec{zw}}(\omega)}, \tag{7}$$

where $n_f(\omega)$ is the number of times the walk ω exited the edge f. In other words, if $\omega = (e_0, e_1, \ldots, e_l)$, then $n_f(\omega)$ is the number of indices $j < l$ for which $e_j = f$.

Analogously, $\Omega_{e,2i+1}$ can be regarded as a probability space. For $\omega \in \Omega_{e,2i+1}$, the probability $p_e(\omega)$ is the probability that a length $2i+1$ non-returning random walk starting at e will be ω. The value of $p_e(\omega)$ is again given by (7).

After these preliminaries we can carry on with the proof of the Lemma.

$$\sum_{e\in E_{LR}} \pi_e N_{e,2i} = \sum_{e\in E_{LR}} \sum_{\omega\in\Omega_{e,2i}} \pi_e p_e(\omega) \cdot \frac{1}{p_e(\omega)}$$

By applying the inequality of the arithmetic and geometric means, we get:

$$\sum_{e\in E_{LR}} \pi_e N_{e,2i} \geq \prod_{e\in E_{LR}} \prod_{\omega\in\Omega_{e,2i}} \left(\frac{1}{p_e(\omega)}\right)^{\pi_e p_e(\omega)}$$

Using (7) we rewrite the $p_e(\omega)$ in the denominator, and get:

$$\begin{aligned}\sum_{e\in E_{LR}} \pi_e N_{e,2i} \quad &\geq \quad \prod_{\vec{wz}\in E_{LR}} (d_z-1)^{\sum_{e\in E_{LR}}\sum_{\omega\in\Omega_{e,2i}} \pi_e p_e(\omega) n_{\vec{wz}}(\omega)} \\ &\quad \cdot \prod_{z\vec{w}\in E_{RL}} (d_w-1)^{\sum_{e\in E_{LR}}\sum_{\omega\in\Omega_{e,2i}} \pi_e p_e(\omega) n_{z\vec{w}}(\omega)}. \end{aligned} \tag{8}$$

We now define the matrix P as the $m\times m$ transition matrix of the non-returning random walk, where in the transition from $\vec{uv}\in E_{LR}$ to $\vec{vz}\in E_{RL}$ the vertex z is chosen uniformly among the d_v-1 possible choices.

$$P_{\vec{uv},\vec{wz}} \;=\; \begin{cases} 1/(d_v-1) & v=w \text{ and } u\neq z \\ 0 & \text{otherwise} \end{cases}$$

Since in (8) the exponent of (d_z-1) is exactly the expected number of times the non-returning random walk starting with the distribution π_{LR} exits the edge $\vec{wz}$, and a similar statement holds also for the exponent of (d_w-1). We conclude that:

$$\sum_{e\in E_{LR}} \pi_e N_{e,2i} \;\geq\; \prod_{\vec{wz}\in E_{LR}} (d_z-1)^{\sum_{j=0}^{i-1}(\pi(PP^t)^j)_{\vec{wz}}} \cdot \prod_{z\vec{w}\in E_{RL}} (d_w-1)^{\sum_{j=0}^{i-1}(\pi P^t\cdot(PP^t)^j)_{z\vec{w}}}.$$

It is easy to verify that π is the stationary distribution of the non-returning random walk ($\pi P=\pi P^t=\pi$), and therefore:

$$\begin{aligned}\sum_{e\in E_{LR}} \pi_e N_{e,2i} \;&\geq\; \prod_{\vec{wz}\in E_{LR}} (d_z-1)^{i/m} \cdot \prod_{z\vec{w}\in E_{RL}} (d_w-1)^{i/m} \\ &=\; (\prod_{z\in V_R} (d_z-1)^{d_z/m})^i \cdot (\prod_{w\in V_L} (d_w-1)^{d_w/m})^i = \Lambda_R^i\cdot\Lambda_L^i,\end{aligned}$$

proving (5).

The proof of (6) is very similar.

$$\begin{aligned}\sum_{e\in E_{RL}} \pi_e N_{e,2i+1} \;&=\; \sum_{e\in E_{RL}}\sum_{\omega\in\Omega_{e,2i+1}} \pi_e p_e(\omega)\cdot\frac{1}{p_e(\omega)} \\ &\geq\; \prod_{e\in E_{RL}}\prod_{\omega\in\Omega_{e,2i+1}} \left(\frac{1}{p_e(\omega)}\right)^{\pi_e p_e(\omega)} \\ &=\; \prod_{\vec{wz}\in E_{LR}} (d_z-1)^{\sum_{e\in E_{RL}}\sum_{\omega\in\Omega_{e,2i+1}} \pi_e p_e(\omega) n_{\vec{wz}}(\omega)} \\ &\quad\cdot \prod_{z\vec{w}\in E_{RL}} (d_w-1)^{\sum_{e\in E_{RL}}\sum_{\omega\in\Omega_{e,2i+1}} \pi_e p_e(\omega) n_{z\vec{w}}(\omega)} \\ &=\; \prod_{\vec{wz}\in E_{LR}} (d_z-1)^{i/m} \cdot \prod_{z\vec{w}\in E_{RL}} (d_w-1)^{(i+1)/m} \\ &=\; \Lambda_R^i\cdot\Lambda_L^{i+1}.\end{aligned}$$

▮

References

[1] N. Alon, S. Hoory, and N. Linial. The Moore bound for irregular graphs. *Graphs Combin.*, to appear.

[2] N. Biggs. *Algebraic graph theory.* Cambridge University Press, Cambridge, second edition, 1993.

[3] B. Bollobás. *Extremal graph theory.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1978.

[4] D. de Caen and L. A. Székely. On dense bipartite graphs of girth eight and upper bounds for certain configurations in planar point-line systems. *J. Combin. Theory Ser. A*, 77(2):268–278, 1997. doi:10.1006/jcta.1997.2745.

[5] I. Neuwirth. The size of bipartite graphs with girth eight. *arXiv:math.CO/0102210*, 2001.

[6] I. Reiman. Über ein Problem von K. Zarankiewicz. *Acta. Math. Acad. Sci. Hungar.*, 9:269–273, 1958.

Chapter 4

A Continuous Analogue of the Girth Problem