The Number of Spanning Trees in Regular Graphs

Noga Alon*
School of Mathematical Sciences, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel

ABSTRACT

Let $C(G)$ denote the number of spanning trees of a graph G. It is shown that there is a function $\epsilon(k)$ that tends to zero as k tends to infinity such that for every connected, k-regular simple graph G on n vertices $C(G) = \{k[1 - \sigma(G)]\}^n$, where $0 \le \sigma(G) \le \epsilon(k)$.

1. INTRODUCTION

What is the minimum possible number of spanning trees of a k-regular connected simple graph on n vertices? This problem was suggested to me by P. Sarnak. His motivation came from number theory; there are many constructions of regular, connected graphs coming from number theoretic considerations (see, e.g., [8], [9]). In some of these constructions, one can obtain an expression for the class number of a certain function field in terms of the number of spanning trees of the corresponding graph. Therefore, the study of class numbers of function fields leads to a study of the number of spanning trees in regular, connected graphs.

All graphs considered here are finite, undirected, and simple, i.e., have no loops and no multiple edges, unless otherwise specified. The *complexity* of a graph G, denoted by $C(G)$, is the number of spanning trees of G (0 if G is disconnected). As will be shown in the next section, for $k \ge 3$ the complexity of any k-regular connected graph G on n vertices is exponential in n. We thus

* Research supported in part by a USA-Israel BSF grant and by a Bergmann Memorial grant.

Random Structures and Algorithms, Vol. 1, No. 2 (1990)

define, for a graph G with n vertices, its *complexity exponent* by $c(G) = [C(G)]^{1/n}$. In [10], McKay proved that for every fixed $k \geq 3$, when $G = (V(G), E(G))$ ranges over all k-regular connected graphs

$$\limsup_{|V(G)| \to \infty} c(G) = \frac{(k-1)^{k-1}}{(k^2 - 2k)^{k/2-1}} .$$

The determination of the corresponding lim inf seems more difficult. For $k \geq 3$ define $c(k) = \liminf_{|V(G)| \to \infty} c(G)$, where $G = (V(G), E(G))$ ranges over all k-regular connected graphs.

Our main result is the following.

Theorem 1.1. *The number of spanning trees of any k-regular connected graph on n vertices is at least*

$$\frac{k^n}{2^{O(n(\log \log k)^2/\log k)}} .$$

Therefore, $c(k) \geq k - O(k(\log \log k)^2/\log k)$.

We also observe that for every $k \geq 3$, $\sqrt{2} \leq c(k) \leq [(k+1)^{k-2}(k-1)]^{1/(k+1)} = k - \Theta(\log k)$ and that for infinitely many values of k

$$c(k) \leq k - \Omega(\sqrt{k}) .$$

Thus, when k tends to infinity, $c(k) = k - o(k)$.

Moreover, since it is easy to show that any k-regular connected graph on n vertices contains less than k^n spanning trees, it follows that as k grows the complexity of any k-regular connected graph on n vertices is $(k - o(k))^n$, i.e., in some (weak) sense, the complexity of any k-regular connected graph G on n vertices does not depend too much on the structure of G and its asymptotic behavior is only a function of n and k.

2. THE COMPLEXITY OF REGULAR CONNECTED GRAPHS

It is not too difficult to show that the number $c(k)$ defined as a lim inf is in fact

$$c(k) = \inf c(G) , \tag{2.1}$$

where the infimum is taken over all k-regular, connected graphs G. To see this, consider an arbitrary, fixed, k-regular, connected graph G on r vertices and let uv be an edge of G such that $G - uv$ is connected (there is always such an edge as G contains a cycle). For each integer $m \geq 2$, let $H_0, H_1, \ldots, H_{m-1}$ be m copies of $G - uv$ and let u_i and v_i be the vertices of H_i corresponding to u and v, respectively, $0 \leq i < m$. Let G_m be the graph obtained from the disjoint union of

$H_0, \ldots, H_{m-1}$ by adding to it the edges $v_i u_{i+1}$ $(0 \le i < m)$, where the indices are reduced mod m. Obviously G_m is a k-regular, connected graph on $m \cdot r$ vertices. Its complexity $C(G_m)$ satisfies

$$C(G_m) = m[C(G-uv)]^m + m[C(G-uv)]^{m-1} \cdot R(G)\,,$$

where $R(G)$ is the number of spanning forests of $G - uv$ with two connected components, one containing u and the other containing v. Since $C(G-uv) \le C(G)$ and $R(G) \le C(G)$, this implies that $C(G_m) \le 2m[C(G)]^m$ and hence $c(G_m) \le (2m)^{1/mr} c(G)$. By letting m tend to infinity, we conclude that $c(k) \le c(G)$, and (2.1) follows. Observe that since $R(G)^{1/mr}$ tends to 1 as m tends to infinity the proof actually implies that for every k-regular, connected G and every edge uv of G which is not a bridge,

$$c(k) \le c(G-uv)\,. \tag{2.2}$$

By the well-known Kirchhoff formula (see, e.g., [3]), for every connected graph G on n vertices $C(G) = (1/n)\prod_{i=1}^{n-1} \lambda_i$, where $\lambda_1, \ldots, \lambda_{n-1}$ are the nonzero eigenvalues of the Laplace matrix $(q_{uv})_{u,v\in V}$ of $G=(V,E)$ defined by $q_{uu} = \deg u$ and $q_{uv} = -1$ for $u \ne v$, $uv \in E$, $q_{uv} = 0$ for $u \ne v$, $uv \notin E$. Let $H = (h_{ij})$ be a Hadamard matrix of order n, that is, an $n \times n$ matrix of ± 1's satisfying $HH^T = nI$. It is well known that such matrices exist for all $n = 2^r$ (and it is conjectured that they exist for all n divisible by 4; see, e.g. [7]). By inverting, if necessary, the signs of some of the rows or columns of H, we may assume that $h_{1j} = 1$ for $1 \le j \le n$ and $h_{i1} = 1$ for $1 \le i \le n$. We can now define a bipartite graph G on the classes of vertices $A = B = \{2, \ldots, n\}$ where $i \in A$ is joined to $j \in B$ iff $h_{ij} = -1$. This graph has $2(n-1)$ vertices, is regular of degree $k = n/2$ and it is easy to show (see, e.g., [1]) that the nonzero eigenvalues of its Laplace matrix are n (with multiplicity 1) and $n/2 \pm \sqrt{n}/2$, each with multiplicity $n-2$. In view of Kirchhoff's formula this implies that $c(G) = [1 + o(1)](n/2 - \sqrt{n}/4) = [1 + o(1)](k - \Omega(\sqrt{k}))$. Other examples showing that $c(k) \le k - \Omega(\sqrt{k})$ can be obtained from other symmetric designs, including, e.g., the hyperplanes versus points incidence graphs of projective geometries, or the well-known Paley graphs (see, e.g., [1]). This, together with (2.1), implies the following statement.

Proposition 2.1. *For infinitely many values of k*

$$c(k) \le k - \Omega(\sqrt{k})\,. \qquad \blacksquare$$

Next we make some simple observations that supply bounds for $c(k)$ for small values of k. First we observe that for every $k \ge 3$ (and in particular for $k = 3$) $c(k) \ge \sqrt{2}$. To see this, let G be a k-regular connected graph on n vertices, with $k \ge 3$. We claim that $C(G) > 2^{n/2}$. Indeed, by the well-known theorem of Nash-Williams [11], all edges of G can be covered by $\lceil (k+1)/2 \rceil$ trees. Let T_1 be one of these trees. Then $E(G) - E(T_1)$ is a set of $\frac{1}{2}kn - (n-1)$ edges, and it is contained in the union of the $\lceil (k-1)/2 \rceil$ other trees. Thus, at least one tree has at least $[\frac{1}{2}kn - (n-1)]/\lceil (k-1)/2 \rceil > n/2$ of these edges. Let F be the set of these edges contained in this tree. Every subset $S \subseteq F$ can be completed to a spanning tree of G by adding only edges of T_1. It follows that for every $S \subseteq F$ there is a spanning tree T_S of G so that $E(T_S) \cap F = S$. Thus $C(G) \geq 2^{|F|}$ and hence $c(k) \geq \sqrt{2}$ for all $k \geq 3$, as claimed.

Another simple observation is that for every $k \geq 3$, $c(k) \leq [(k+1)^{k-2}(k-1)]^{1/(k+1)}$. In view of (2.2), this would follow from the existence of a k-regular graph G such that $G - uv$ is connected for some edge uv and $c(G - uv) = [(k+1)^{k-2}(k-1)]^{1/(k+1)}$. Let G be the complete graph on $k+1$ vertices. By the well-known theorem of Cayley (cf., e.g., [3]), $C(G) = (k+1)^{k-1}$. By symmetry, for every edge uv of G,

$$C(G - uv) = C(G) \cdot (1 - k/\tbinom{k+1}{2}) = (k+1)^{k-1} \frac{(k(k+1) - 2k)}{k(k+1)} = (k+1)^{k-2}(k-1)$$

and hence $c(G - uv) = [(k+1)^{k-2}(k-1)]^{1/(k+1)}$, as needed.

We summarize the last two observations in the following claim.

Proposition 2.2. *For every* $k \geq 3$

$$\sqrt{2} \leq c(k) \leq [(k+1)^{k-2}(k-1)]^{1/(k+1)} . \qquad \blacksquare$$

Finally, we prove Theorem 1.1, which shows that in fact for large k, $c(k) = [1 - o(1)]k$.

Proof of Theorem 1.1. Throughout the proof we assume, whenever it is needed, that k is sufficiently large. All logarithms in this proof are in base 2.

Let $G = (V, E)$ be an arbitrary k-regular, connected graph on n vertices. For each vertex $v \in V$ choose, randomly (with a uniform distribution on the k edges incident with v) and independently, an edge incident with v and orient it from v outwards. This gives a random oriented subgraph H of G (in which some edges may appear twice; oriented in the opposite directions). One can easily check that every connected component K of H is a connected graph with a unique directed cycle (which may be a cycle of length 2). Moreover, given the edges of K (including the one that appears twice, if there is such an edge), there are at most two possible orientations of the edges of K corresponding to the actual choice of edges. (Only the orientation of the directed cycle need be chosen, as the other edges are necessarily oriented towards the cycle).

Put $g = \log k/(10 \log\log k)$. For every integer r, $2 \leq r \leq g$, the expected number of connected components with r vertices in the randomly chosen graph H does not exceed

$$\frac{n(kr)^{r-1} \cdot r^2}{k^r} < \frac{n}{\sqrt{k}} . \tag{2.3}$$

This is because there are n choices for the first vertex of the component. The other $r - 1$ vertices can be chosen so that each vertex is a neighbor (in G) of one of the previous ones. Consequently, there are less than kr choices for each such new vertex. This covers the choice of all r vertices of the component and $r - 1$ of its edges. The last edge has at most $\binom{r}{2}$ choices, and the number of possible orientations is either 1 or 2. Hence, there are less than $n(kr)^{r-1} \cdot r^2$ choices for an oriented component, and the probability that its r edges will indeed be chosen is $1/k^r$. This implies (2.3). In view of (2.3), the expected number of connected components of size at most g in H does not exceed $gn/\sqrt{k} < n/2g$. Thus, the probability that H has less than n/g such components is at least $\frac{1}{2}$. This means that there are at least $\frac{1}{2}k^n$ possible choices for the oriented H, so that it has less than n/g connected components of size at most g. For each such H we associate a forest F_H by deleting an arbitrary edge from the unique cycle of each connected component. Notice that F_H has less than $2n/g$ connected components (as it has less than n/g components of size $\leq g$, and, of course, less than n/g components of size $>g$). Thus, the number of edges $|E(F_H)|$ of F_H satisfies

$$|E(F_H)| \geq n - 2n/g\,. \tag{2.4}$$

We claim that the same forest F_H cannot be obtained from too many oriented subgraphs H. Indeed, given F_H, in order to know H we have to know the edge deleted from each component of H and the orientation in each component. If a component has r vertices, then there are less than $\binom{r}{2}$ choices for the deleted edges in it, and at most 2 possible orientations. Thus the number of subgraphs H that may give the same F_H is at most the product of the squares of the sizes of the connected components. The contribution of the small components to this product is at most $g^{2n/g}$. The contribution of the large ones is also at most $g^{2n/g}$ (as $g \geq 3$, the product of numbers whose sum is at most n and each exceeds g is at most $g^{n/g}$). It follows that the number of forests of size at least $n - 2n/g$ in G is at least

$$\tfrac{1}{2}k^n/g^{4n/g}\,. \tag{2.5}$$

However, each forest is contained in a spanning tree of G, and each tree contains at most $\Sigma_{i=0}^{2n/g}\binom{n-1}{i}$ forests of size at least $n - 2n/g$. By the standard estimates for binomial distributions (see, e.g., [5])

$$\sum_{i=0}^{2n/g}\binom{n-1}{i} \leq 2^{H(2/g)\cdot n} \leq 2^{(4\log g/g)\cdot n}\,, \tag{2.6}$$

where here $H(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function.

Combining (2.5) and (2.6), we conclude that the total number of spanning trees of G is at least

$$\frac{1}{2}\,\frac{k^n}{g^{8n/g}} = \frac{k^n}{2^{1+8n\log g/g}}\,.$$

Substituting $g = \log k/10\log\log k$, the assertion of the theorem follows. ∎

3. CONCLUDING REMARKS

1. The first step in the proof of Theorem 1.1 shows that for every k-regular graph G on n vertices $C(G) \leq (1/(n-1))k^n$. This is because by choosing a

single edge emanating from every vertex we obtain each tree as an oriented subgraph, with one edge taken in both directions exactly $n-1$ times. It is well known (cf., e.g., [2]) that Kirchhoff's formula implies the slightly better bound

$$C(G) \le \frac{1}{n}\left(\frac{nk}{n-1}\right)^{n-1} < \frac{e}{n}\, k^{n-1}$$

Indeed, the trace of the Laplace matrix of G is kn and hence the sum of its $n-1$ nonnegative eigenvalues $\lambda_1, \ldots, \lambda_{n-1}$ is kn, implying, by the arithmetic-geometric inequality, that

$$C(G) = \frac{1}{n}\prod_{i=1}^{n-1} \lambda_i \le \frac{1}{n}\left(\frac{nk}{n-1}\right)^{n-1}$$

The bound of McKay [10], stated in the Introduction, shows that in fact, for large n, $C(G)$ is a little smaller. Each of these estimates together with Theorem 1.1 shows that for any k-regular, connected graph G on n vertices, $C(G) = [k - o(k)]^n$.

2. One may study the minimum possible complexity of nonsimple k-regular connected graphs. If loops are allowed then trivially, for any odd k and any even n there is a k-regular connected graph on n vertices with a unique spanning tree. We are thus left with the question of estimating $c'(k) = \liminf_{|V(G)|\to\infty}[C(G)]^{1/n}$, where G ranges over all k-regular, connected graphs on n vertices which contain no loops but may have parallel edges. The graph on two vertices with k edges between them [and the fact that the equality analogous to (2.1) holds for $c'(k)$ too] shows that $c'(k) = O(\sqrt{k})$. On the other hand, we can show that $c'(k) = \Omega(\sqrt{k})$. To see this, observe, first, that since any spanning tree contains less than 2^n forests and any forest is contained in a spanning tree, it suffices to show that any k-regular connected multigraph on n vertices with no loops contains at least $[\Omega(\sqrt{k})]^n$ forests. This can be shown by a simple modification of the first part in the proof of Theorem 1.1. A somewhat stronger result is the fact that any such graph G contains at least $[\Omega(\sqrt{k})]^n$ linear forests, i.e., forests in which every connected component is a path. This is a simple consequence of the Van der Waerden conjecture, proved in [6] and [4]. By applying it to the permanent of the adjacency matrix of G, we conclude that G contains at least $[\Omega(k)]^n$ spanning 2-factors, and by deleting an edge from each cycle in each of these 2-factors we obtain the desired estimate for the number of linear forests. We omit the details.
3. It would be interesting (and seems to be difficult) to determine $c(k)$ precisely for each $k \ge 3$. □

ACKNOWLEDGMENT

I would like to thank Peter Sarnak for helpful comments.

REFERENCES

[1] N. Alon, Eigenvalues, geometric expanders, sorting in rounds and Ramsey Theory, *Combinatorica*, **6**, 207–219 (1986).

[2] N. L. Biggs, *Algebraic Graph Theory*, Cambridge University Press, London, 1974.

[3] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, Macmillan, London, 1978.

[4] G. P. Egorichev, The solution of the Van der Waerden problem for permanents, *Dokl. Math. Nauk SSSR*, **258**, 1041–1044 (1981).

[5] P. Erdös and J. Spencer, *Probabilistic Methods in Combinatorics*, Academic, New York and London, 1974.

[6] D. I. Falikman, A proof of Van der Waerden's conjecture on the permanent of a doubly stochastic matrix, *Mat. Zametki*, **29**, 931–938 (1981).

[7] M. Hall, *Combinatorial Theory*, Wiley, New York, 1967.

[8] Y. Ihara, Discrete subgroups of $PL(2, k_p)$, *Proc. Symp. Pure Math.* IX, 272–278 (1968).

[9] A. Lubotzky, R. Phillips, and P. Sarnak, Ramanujan graphs, *Combinatorica*, **8**, 261–278 (1988).

[10] B. D. McKay, Spanning trees in regular graphs, *Euro. J. Combinatorics*, **4**, 149–160 (1983).

[11] C. St. J. A. Nash-Williams, Decomposition of finite graphs into forests, *J. London Math. Soc.*, **39**, 12 (1964).

[12] J. Sedlacek, Regular graphs and their spanning trees, *Čas. Pro. Pěst Mat.*, **95**, 420–426 (1970).

Received July 15, 1989