AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **82(2)** (2022), Pages 182–196

Minimal graphs with a prescribed number of spanning trees

Richard Stong

Center for Communications Research
San Diego, CA, U.S.A.
stong@ccrwest.org

Abstract

For $n \geq 3$, let $\alpha(n)$ denote the minimum number of vertices in a graph with exactly n spanning subtrees. This notion was introduced by Sedláček and has been studied by Azarija and Škrekovski. In particular, they have conjectured that $\alpha(n) = o(\log n)$.

This paper will prove a weak version of this conjecture: specifically we will show that $\alpha(n) = O((\log n)^{3/2}/(\log\log n))$. This bound is substantially larger than the conjectured upper bound; it is at least in the same ballpark.

1 Introduction

For $n \geq 3$, let $\alpha(n)$ denote the minimum number of vertices in a graph with exactly n spanning subtrees. (In this paper, by a graph we will always mean a simple graph without loops or multiple edges.) This notion was introduced by Sedláček [5] and has been studied by Nebeský [3] and Azarija and Strekovski [1],[4]. In particular, the last two authors have conjectured that $\alpha(n) = o(\log n)$.

Note that for a graph with m vertices, the number of spanning subtrees is maximized for a complete graph K_m which by Cayley's theorem has m^{m-2} spanning subtrees. Therefore $\alpha(m^{m-2}) = m$. For general n, we get a lower bound from this maximality. Specifically, if we let m be the least integer such that $m^{m-2} \geq n$, then we have $\alpha(n) \geq m$. With a little manipulation, we can write this as $\alpha(n) \geq \frac{\log n}{\log\log n}(1 + o(1))$.

If G and G' are graphs, then we can form the wedge $G \vee G'$ by taking the disjoint union and identifying one vertex of G with one vertex of G'. The number of spanning subtrees multiplies under taking wedges. Therefore we have fairly tight upper bounds on $\alpha(n)$ if n is a product of a small number of terms of the form m^{m-2} and a slightly weaker upper bound $\alpha(a^r) = O(r)$ for any fixed a (so $\alpha(n) = O(\log n)$ at least for this very sparse family of ns).

These considerations make the conjectured upper bound seem reasonable and one can also write down some crude heuristics that make the conjecture seem plausible.

ISSN: 2202-3518

Despite this all the reported upper bounds on $\alpha(n)$ seem to be very weak. (The bounds I saw were all linear in n, though I am hardly an expert on this question.) The goal of this paper is to present a much better upper bound, which while not as strong as the conjecture is at least within the same ballpark. The best result we will be able to prove is that $\alpha(n) = O((\log n)^{3/2}/(\log\log n))$. However this best bound requires some fairly heavy algebraic number theory.

To compensate, we will begin rather slowly. We will first prove a very elementary linear bound $\alpha(n) \leq n/17+O(1)$. Even this weak bound seems to be an improvement on the existing literature. This construction will naturally generalize to a family of linear bounds $\alpha(n) \leq n/r + O(1)$ for arbitrarily large r, and by balancing the linear term and the constant term we will give an almost elementary proof that $\alpha(n) = O(n^{1/2})$. These initial examples will all be in some sense built of just two pieces, but for the sharpest results we will need examples built from chains of arbitrary length and we will be forced to use some fairly serious number theory. We will give an argument which shows $\alpha(n) = O((\log n)^2/(\log\log n))$ but requires less sophisticated number theory (namely, Dirichlet's theorem on primes in arithmetic progressions) and then our strongest bound, which requires very serious number theory.

2 The toolkit

Since this paper is devoted to proving upper bounds, all our work will be devoted to just presenting some basic examples. These examples will be built from very simple graphs by relatively simple operations, so we will first collect up these tools.

The basic building blocks will be just two kinds of graphs, cycles C_p of prime length p, and theta graphs $\theta(a, b, c)$ which consist of two vertices joined by paths of length a, b, c. (Since we do not allow multiple edges, we must insist that at most one of the parameters a, b, c is equal to 1.) These are among the simplest graphs in a topological sense, and they arise in many graph theoretic examples. In particular the theta graphs were used in [1], though in an unrelated way.

For a graph G, let $s(G)$ denote the number of spanning subtrees of G, and if u, v are vertices of G let $s_{u,v}(G)$ denote the number of spanning forests with two components and u, v in different components. Note that $s_{u,v}(G) = s(G/(u = v))$, where by $G/(u = v)$ we mean the multigraph where u and v have been collapsed to a single vertex. (An edge joining u and v might be viewed under this collapse as producing a loop, but since a loop cannot be part of a spanning subtree we may without loss assume that such an edge is also collapsed.)

Obviously $s(G)$ is the main quantity of interest in this paper, but $s_{u,v}(G)$ should be thought of as some sort of derivative of $s(G)$ in the following sense. Consider the graphs Γ_k obtained from G by adding a path of length k (counting edges) between two selected vertices $u, v \in V(G)$. Note that since we want Γ_k to have neither loops nor multiple edges, we need $k \geq 2$ if u, v are adjacent in G, and $k \geq 1$ if $u \neq v$ are non-adjacent in G. Since any spanning subtree of Γ_k either deletes one edge of the path (and hence intersects G in a spanning subtree) or has the entire path (and hence meets G in a 2-component spanning forest with u and v in different components), we have

$$s(\Gamma_k) = ks(G) + s_{u,v}(G).$$

There is a related construction where we simply take the wedge of G with a cycle of length $k \geq 3$. This will have $s(G \vee C_k) = ks(G)$. We can regard this as a special case of the construction above where we set $u = v$ and we take the convention that $s_{u,u}(G) = 0$. We will take this attitude in most of this paper, but in Sections 3 and 5 we will regard taking a wedge with a loop as a separate part of our toolkit.

Note that $s(C_p) = p$ since deleting any of the p edges gives a spanning subtree and conversely. Also $s(\theta(a, b, c)) = ab + bc + ca$ since choosing two of the three paths and deleting one edge from each gives a spanning subtree and conversely. In this paper we will consider only theta graphs where $ab + bc + ca = p$ is a prime.

3 The prime 17 and the first upper bound

Focus on $p = 17$ and just the two graphs C_{17} and $\theta(1, 2, 5)$. These graphs have $s(C_{17}) = s(\theta(1, 2, 5)) = 17$. We will use these two basic examples and the construction of Γ_k from the previous section to build 17 different families of graphs. Each family will have $s(\Gamma_k)$ in a particular congruence class modulo 17, will cover all sufficiently large elements of its congruence class, and the 17 families will cover between them all 17 congruence classes. Thus we see that all sufficiently large n can be written as $s(\Gamma_k)$ for one of these 17 families. Since this section is just a warm-up for the more complete argument and 17 cases gets a little tedious, we will omit many of the specific computations.

For the first family, we will take $\theta(1, 2, 5) \vee C_k$ for $k \geq 3$, which again can be viewed as a degenerate case of the Γ_k construction. Such a graph has $k + 6$ vertices and

$$s(\theta(1, 2, 5) \vee C_k) = 17k.$$

Thus for any $n \equiv 0 \pmod{17}$ with $n \geq 51$ we can find a graph Γ with $n/17 + 6$ vertices and $s(\Gamma) = n$.

The next eight families will all be built from C_{17}. We choose u and v to be vertices of C_{17} separated by j edges, for $j = 1, 2, \ldots, 8$, and let $\Gamma_{j,k}$ be the graph obtained by adding a path of length k joining u and v (which will actually give a theta graph $\theta(j, 17 - j, k)$). Since we produce a 2-component spanning forest by deleting one edge of the path of length j from u to v and one edge of the complementary path of length $17 - j$, we see that $s_{u,v}(C_{17}) = j(17 - j)$. Thus for each $j = 1, 2, \ldots, 8$, we have a graph with $k + 16$ vertices and

$$s(\Gamma_{j,k}) = 17k + j(17 - j).$$

Note that for $j = 1$ we must insist on $k \geq 2$ to avoid multiple edges, but otherwise we can allow $k \geq 1$. The numbers $j(17-j)$ are congruent to $16, 13, 8, 1, 9, 15, 2, 4$ modulo 17 (which not coincidentally are exactly the nonzero quadratic residues modulo 17).

Turning this around, from the specific case $j = 1$, we see that for any $n \equiv 16 \pmod{17}$ with $n \geq 2 \cdot 17 + 1(17 - 1) = 50$, there is a graph $\Gamma_{1,(n-16)/17}$ with $(n - 16)/17 + 16 = n/17 + 256/17$ vertices and $s(\Gamma_{1,(n-16)/17}) = n$. For $j = 2, \ldots, 8$, we get similar families. More explicitly, for each of the congruence classes d above, we will find constants $B_1(d)$ and $B_2(d)$ such that for every $n \equiv d \pmod{17}$ with $n > B_1(d)$, there is a graph Γ with $n/17 + B_2(d)$ vertices and $s(\Gamma) = n$.

The final eight families will all be built from $\theta(1, 2, 5)$ using the nondegenerate version of the Γ_k construction from the previous section. The graph $\theta(1, 2, 5)$ can be viewed as a 7-cycle, say with vertices v_i for $i \in \mathbb{Z}/7\mathbb{Z}$, with one additional edge added joining two vertices 2 apart, say v_1v_6. To avoid writing double subscripts, we will write $s_{i,j}$ for $s_{v_i,v_j}(\theta(1, 2, 5))$. One can compute $s_{1,5} = 20 \equiv 3 \pmod{17}$, $s_{1,3} = 22 \equiv 5 \pmod{17}$, $s_{0,2} = 23 \equiv 6 \pmod{17}$, $s_{1,4} = 24 \equiv 7 \pmod{17}$, $s_{1,6} = 10$, $s_{0,1} = 11$, $s_{0,3} = 29 \equiv 12 \pmod{17}$, and $s_{1,2} = 14$. These congruence classes modulo 17 are exactly the quadratic non-residues, and hence they complete our menagerie. We should go through all these cases individually, but since they will be covered by the general computation in Section 5, we will focus on just one of them (the easiest), $s_{1,6} = 10$. In this case the two vertices chosen, v_1 and v_6, are the two endpoints of the three paths. Hence $s_{1,6} = 10$ since we get a 2-component spanning forest by deleting one edge from each of the three paths, and conversely. Since the vertices v_1, v_6 are adjacent, we must join them by a path of length $k \geq 2$ to construct a Γ_k without multiple edges. Thus for each $k \geq 2$, we get a graph Γ_k with $k + 6$ vertices and

$$s(\Gamma_k) = 17k + 10.$$

Turning this around, we see that if $n \equiv 10 \pmod{17}$ and $n \geq 2 \cdot 17 + 10 = 44$, then the graph $\Gamma_{(n-10)/17}$ has $(n-10)/17+6 = n/17+92/17$ vertices and $s(\Gamma_{(n-10)/17}) = n$. The other cases are similar, the resulting values for $B_1(d)$ and $B_2(d)$ are smaller than those for C_{17}, largely because $\theta(1, 2, 5)$ has fewer vertices.

Let $B_1 = \max_d B_1(d)$ and $B_2 = \max_d B_2(d)$. If one accepts the omitted calculations and accepts that the maxima are attained for the $(C_{17}, j = 8)$ and $(C_{17}, j = 1)$ families above, respectively, then we have proven the following theorem.

Theorem 3.1. *For all $n > 72$, we have $\alpha(n) \leq n/17 + 256/17$.*

4 The general construction

It is hopefully clear that the construction of the previous section is just a specific case of what should be a fairly powerful construction. Let us introduce some notation to help us access this power.

Suppose we have fixed a triple (G, u, v) of a graph G and two (possibly equal) vertices u, v of G. Recall that if $u = v$, we set $s_{u,u}(G) = 0$. Let G have $s(G) = r$, and $s_{u,v}(G) \equiv d \pmod{r}$. Then we can define the graph Γ_k by joining u to v by a path of length k. The graphs Γ_k will all have $s(\Gamma_k) \equiv d \pmod{r}$ and since we only have a lower bound on k they will cover all sufficiently large values in this congruence class. Let k_{min} be the smallest allowed value of k. Then Γ_k has $|V(G)| + k - 1$ vertices and $s(\Gamma_k) = kr + s_{u,v}(G) \equiv d \pmod{r}$. To turn this around as we did in the previous section, we define $B_1(G, u, v) = (k_{min} - 1)r + s_{u,v}(G)$. Since we have seen that $k_{min} - 1 = 2, 1, 0$ depending on whether $u = v$, u, v are adjacent, or otherwise, we can write this as

$$B_1(G, u, v) = \begin{cases} 2r & \text{if } u = v \\ r + s_{u,v}(G) & \text{if } u, v \text{ are adjacent} \\ s_{u,v}(G) & \text{if } u, v \text{ are neither equal nor adjacent} \end{cases}.$$

(Note that as required all these values are congruent to d modulo r.) If $n \equiv d \pmod{r}$ and $n > B_1(G, u, v)$, then the graph $\Gamma_{(n-s_{u,v}(G))/r}$ will have $s(\Gamma_{(n-s_{u,v}(G))/r}) = n$ and

$$\frac{n}{r} + |V(G)| - 1 - \frac{s_{u,v}(G)}{r}$$

vertices. Let $B_2(G, u, v) = |V(G)| - 1 - s_{u,v}(G)/r$ be the constant term of this expression.

Thus from the triple (G, u, v) and our basic construction, we can conclude that for any $n > B_1(G, u, v)$ with $n \equiv d \pmod{r}$, there is a graph Γ with n vertices and $s(\Gamma) = n$.

Since a single triple gives all sufficiently large n in one congruence modulo r, we obviously want to do exactly what we did in the previous section and assemble a collection of r triples (G, u, v) all with the same value $s(G) = r$ and such that the numbers $s_{u,v}(G)$ collectively represent all the different congruence classes modulo r. As in Section 3, our first and simplest examples will use just two different graphs G and many pairs of vertices within these graphs. The later and more difficult examples will use large collections of graphs. If we have such a collection we will write (G_d, u_d, v_d) for the triple with $s_{u_d,v_d}(G_d) \equiv d \pmod{r}$, and we will abuse notation slightly and write $B_1(d)$ and $B_2(d)$ for the values of $B_1(G_d, u_d, v_d)$ and $B_2(G_d, u_d, v_d)$. As in the previous section, we will also define $B_1 = \max_d B_1(d)$ and $B_2 = \max_d B_2(d)$. We will call such a family $\mathcal{G} = \{(G_d, u_d, v_d)\}$ a representative family of triples (RFT) modulo r with parameters $B_1 = B_1(\mathcal{G})$ and $B_2(\mathcal{G})$. Then we have the following theorem, which is arguably more of a restatement of the definition.

Theorem 4.1. *If there is a RFT modulo r with parameters B_1, B_2, then for all $n > B_1$, we have $\alpha(n) \leq n/r + B_2$.*

5 Analysis of cycles and thetas

In the previous section, we saw that to any triple (G, u, v) we can assign the parameters $s(G)$, $s_{u,v}(G)$, $B_1(G, u, v)$ and $B_2(G, u, v)$, and that by assembling enough such triples with the same value of $s(G)$ we can prove the sorts of theorems we want to prove. In the section before that, we tried to indicate that cycles and theta graphs were good examples to consider, but we omitted some of the key details. Now it is time to pay the piper and compute these parameters for our fundamental examples.

Again we will only look at two kinds of graphs, cycles C_p of prime length p, and theta graphs $\theta(a, b, c)$ where the number of spanning subtrees $ab + bc + ca = p$ is prime. For each of these cases, we will look at the range of values of $s_{u,v}(G)$ as u, v vary over distinct vertices of G. We will see that these values are either all (non-zero) quadratic residues modulo p or all (non-zero) quadratic non-residues modulo p. The zero congruence class in an RFT is always provided by the degenerate case (G, u, u), so we can ignore it. Therefore as long as both cases occur, we will have a RFT modulo p and hence a linear upper bound on α with slope $1/p$.

Suppose $G = C_p$ is a cycle of prime length p. If we choose two vertices u, v on this cycle j apart, then $C_p/(u = v)$ will be a wedge of a cycle of length j and a cycle of length $p - j$ and therefore $s_{u,v}(C_p) = j(p - j)$. Thus we see that the values of $-s_{u,v}(C_p) \pmod p$ are exactly the quadratic residues modulo p. Hence the values of $s_{u,v}(C_p)$ are all quadratic residues if $p \equiv 1 \pmod 4$ and all quadratic non-residues if $p \equiv 3 \pmod 4$.

Next consider $B_1(G, u, v)$ for this family. For $j = 1$ (and only in this case) we are choosing adjacent vertices and the construction above applies for $k > 1$, hence for $n > p + (p - 1) = 2p - 1$. For all other j, we need $n > s_{u,v}(C_p) = j(p - j)$, and the lower bound is maximized for $j = (p \pm 1)/2$ when it equals $(p^2 - 1)/4$. For $p > 7$, which is the only case where our results will apply, this means that $B_1 = (p^2 - 1)/4$. Since B_2 is the maximum of $|V(G)| - 1 - s_{u,v}(G)/p$ and $s_{u,v}(G)$ is minimized for $j = 1$, we have

$$B_2 = p - 1 - (p - 1)/p = (p - 1)^2/p.$$

Suppose $G = \theta(a, b, c)$ where $s(G) = ab + bc + ca = p$ is a prime. View G as being built from three paths (of lengths a, b, and c) with vertices $(x_i)_{i=0}^{a}$, $(y_i)_{i=0}^{b}$ and $(z_i)_{i=0}^{c}$ by identifying $x_0 = y_0 = z_0$ and $x_a = y_b = z_c$. Suppose we choose $u = x_i$ and $v = y_j$. Then identifying u and v will produce a multigraph with three vertices $X = x_0 = y_0 = z_0$, $Y = x_a = y_b = z_c$, and $Z = u = v$ where X and Y are joined by a path of length c, X and Z are joined by paths of length i and j, and Y and Z are joined by paths of length $a - i$ and $b - j$. Hence we compute

$$\begin{aligned} s_{u,v}(\theta(a, b, c)) &= c(i + j)(a + b - i - j) + ij(a + b - i - j) \\ &\qquad + (i + j)(a - i)(b - j) \\ &= (i + j)(ab + bc + ca) - j^2a - i^2b - (i + j)^2c. \end{aligned}$$

Arguably, the argument above breaks down if $i = 0, a$ or $j = 0, b$. One can either convince oneself that it does apply (if we interpret a path of length 0 as meaning we have identified the endpoints) or one can directly compute $s_{u,v}$ in these cases (the resulting graphs are wedges of a theta and a cycle) and check that the formula is still correct.

There is another case where u, v lie on the same path of the theta, but it is easy to see that in this case $G/(u = v)$ is also a wedge of a cycle and a theta and hence the values of $s_{u,v}$ produced are the same as those produced by the degenerate case above. Hence we will ignore this possibility.

Note that the formula above for $s_{u,v}(G)$ gives

$$-(a+b)s_{u,v}(\theta(a,b,c)) \equiv (ja-ib)^2 \pmod{p}$$

and since $(b+c)(a+c) \equiv c^2 \pmod{p}$

$$-(a+b)(b+c)(c+a)s_{u,v}(\theta(a,b,c)) \equiv (c(ib-ja))^2 \pmod{p}.$$

Thus as in the case of a cycle all the values of $s_{u,v}(G)$ are either quadratic residues modulo p or all the values are quadratic non-residues modulo p. Further as the following lemma shows we get all such values.

Lemma 5.1. *Suppose $ab+bc+ca=p$ is a prime. Then every congruence class modulo p can be written in (at least) one of the following forms:*

1. *$c(ib-ja)$ for $0 \le i \le a$ and $0 \le j \le b$,*
2. *$a(ic-jb)$ for $0 \le i \le b$ and $0 \le j \le c$, or*
3. *$b(ia-jc)$ for $0 \le i \le c$ and $0 \le j \le a$.*

Proof. Consider the standard triangular lattice in the plane. For ease of identifying points on this lattice, we will think of the plane as the complex plane and we will identify the lattice points with the Eisenstein lattice $\mathbb{Z}[\omega]$, where ω is a primitive cube root of unity. Starting at the origin and moving along the positive real axis draw an equiangular hexagon with side lengths read cyclically a,b,c,a,b,c by turning 120 degrees counterclockwise at each angle. Label the vertices inside and on this hexagon by numbers modulo p by adding bc if we move by $+1$, ab if we move by $+\omega$ and ca if we move by $+\omega^2$. Since $1+\omega+\omega^2=0$ is the minimal polynomial of ω and $ab+bc+ca \equiv 0 \pmod{p}$, we can unequivocably label all vertices in this way.

Decompose this hexagon into three $60-120$ degree parallelograms as follows. Cut the perimeter at the three points obtained by starting at the origin and taking every other vertex. Explicitly, these are the points with coordinates 0, $a-b\omega^2$, and $c\omega-b\omega^2$. Note that all three of these vertices are labelled 0 in our scheme. This divides the perimeter of the hexagon into three pieces which (read clockwise) have side lengths (a,b), (b,c), and (c,a) with a 120 angle between the two sides. Extending each pair to a parallelogram gives the desired decomposition. Points in the first parallelogram have coordinates $i-j\omega^2$ for $0 \le i \le a$ and $0 \le j \le b$ and hence are labelled by the numbers $ibc-jca=c(ib-ja)$. Thus they give exactly case (1) of the statement. Since all three cutting vertices have label 0, the other two parallelograms differ by just a cyclic rotation of the parameters a,b,c. Hence case (2) is the values given by the second parallelogram, and case (3) the third. Thus the labels assigned within the hexagon are exactly the aggregate of the numbers of the three forms given.

The values assigned to the vertices of this hexagon alternate between 0 and $abc \pmod{p}$, so if we translate the hexagon taking any side to the opposite side, the values will agree. Iterating this we tile the entire plane with translates of the hexagon with each copy having the same values assigned. We see that from any lattice point steps in the three directions add ab, bc, or ca modulo p and therefore any value of the form $bcx + aby + caz \pmod p$ for integer x, y, z occurs. Since a, b, c are pairwise relatively prime, this implies every value modulo p occurs. □

Next consider the values of B_1 and B_2 for this collection. The required maxima and minima will depend on the relative sizes of a, b, c so assume without loss of generality that $a \leq b \leq c$. If $i = 0$ and $j = 1$ (and up to symmetry only in this case), we are choosing adjacent vertices and taking the most extreme case, we need $n > (p - a - b) + p = 2p - a - b$. For all other i, j, we need only $n > s_{u,v}(\theta(a, b, c))$. By convexity, this lower bound is maximized when u, v are on the two longest paths and i and j as close to $b/2$ and $c/2$ as possible (given that they must be integers). Ignoring the integrality condition and hence settling for just an upper bound (that is, just plugging in $i = b/2$ and $j = c/2$ and simplifying), we find

$$B_1 \leq \max(2p - a - b, (b + c)p/4).$$

If $(a, b, c) \neq (1, 2, 3)$ (and in this exceptional case both C_{11} and $\theta(1, 2, 3)$ give values of $s_{u,v}$ that are non-residues modulo 11 so the theorem stated below will not apply anyway), then we can replace this by the easier upper bound

$$B_1 \leq (a + b + c)p/4.$$

For B_2 using the crudest lower bound $s_{u,v} \geq 0$, we get

$$B_2 \leq a + b + c - 2.$$

Summarizing, we have seen that if G is either a cycle or a theta graph and $s(G) = p$ is prime, then the values of $s_{u,v}(G)$ as u, v vary over pairs of distinct vertices of G will either be every quadratic residue or non-residue modulo p. As long as we can find two choices of G, one of which gives the residues and the other the non-residues, we will have built an RFT modulo p. The simplest case of this is the following theorem.

Theorem 5.2. *Suppose the prime p can be written as $p = ab + bc + ca$ for positive integers a, b, c at most one of which is* 1 *and suppose $a + b$ is a quadratic non-residue modulo p. Then for every $n > (p^2 - 1)/4$, we have $\alpha(n) \leq \frac{n+(p-1)^2}{p}$.*

Proof. Since $(a+c)(b+c) \equiv c^2 \pmod p$, we see that if $a+b$ is a quadratic non-residue modulo p, then so is $(a+b)(b+c)(c+a)$ (and so are $a+c$ and $b+c$). Since $-s_{u,v}(C_p)$ and $-(a + b)(b + c)(s + a)s_{u,v}(\theta(a, b, c))$ are quadratic residues modulo p, it follows that one of $s_{u,v}(C_p)$ and $s_{u,v}(\theta(a, b, c))$ gives quadratic residues modulo p and the other gives non-residues. Hence the graphs C_p and $\theta(a, b, c)$ and various choices of vertices u, v provide an RFT modulo p. For the 0 class, we use $\theta(a, b, c)$ with $u = v$ which gives $B_1 = 2p$ and $B_2 = a + b + c - 2$. Since the hypotheses force $p \geq 17$ and $p > a + b + c$, the maxima for B_1 and B_2 both come from the cycle of length p. □

Corollary 5.2.1. *If $p > 5$ is a prime congruent to 5 modulo 12, then the above theorem applies.*

Proof. Write $p = 3m + 2$ and note that $s(\theta(1, 2, m)) = p$ and since p is 1 modulo 4 quadratic reciprocity shows that $1 + 2 = 3$ is a quadratic nonresidue modulo p. □

Either of these last two results generalizes Section 3 (and fills in the missing details in that section), and shows that we can give completely elementary proofs of upper bounds of the form $\alpha(n) \leq n/p + O(1)$ for a sequence of primes p. Going further seems to require some fairly deep number theory. For example, from the (relatively) elementary fact that there are infinitely many primes congruent to 5 modulo 12 we get the following.

Corollary 5.2.2. $\alpha(n) = o(n)$.

There is a weak version of Bertrand's postulate for primes congruent to 5 modulo 12: for all sufficiently large x we can choose a prime p congruent to 5 modulo 12 with $x < p < 2x$. (One proof of this, albeit using some serious machinery, is to invoke Dirichlet's theorem, which in its full form says that the primes congruent to 5 modulo 12 have Dirichlet density $\frac{1}{4}$. Positive Dirichlet density implies this result and much stronger ones. Still this claim feels more elementary than Dirichlet.) Choosing a prime p congruent to 5 modulo 12 with $\sqrt{n} < p < 2\sqrt{n}$, Theorem 5.2 gives

$$\alpha(n) < \frac{n}{p} + p.$$

Since the right-hand side is an increasing function of p for $p > \sqrt{n}$, we get the following result.

Corollary 5.2.3. *For all sufficiently large n, we have $\alpha(n) < 5/2 \cdot n^{1/2}$.*

One could of course replace the constant by any value greater than 2 by only slightly modifying this discussion.

6 Assembling chains

Now we come to one of the really powerful tricks in the current paper. Suppose we RFTs $\mathcal{G} = \{(G_d, u_d, v_d)\}$ modulo r and $\mathcal{G}' = \{(G'_d, u'_d, v'_d)\}$ modulo r' for relatively prime numbers r and r'. Look at what happens if we form the rr' triples $(G_d \vee G'_{d'}, v_d, v'_{d'})$, where the wedge $G_d \vee G'_{d'}$ is formed by identifying u_d and $u'_{d'}$. To avoid excessive subscripts, we will drop them, relying on the reader to understand that all unprimed objects correspond to a fixed choice of d and similarly for primed objects and d'. These graphs all have $s(G \vee G') = rr'$. If we have any two component spanning forest of $G \vee G'$ with v and v' in different components, then exactly one of v and v' will be in the same component as the wedge vertex $u = u'$. Thus we see that

$$s_{v,v'}(G \vee G') = rs_{u',v'}(G') + r's_{u,v}(G).$$

Since r and r' were assumed to be relatively prime, and we assumed that our triples were chosen so that $s_{u,v}(G)$ and $s_{u',v'}(G')$ cover every congruence class modulo r and r', respectively, it follows from the Chinese remainder theorem that $s_{v,v'}(G \vee G')$ covers every congruence class modulo rr'. Thus we have produced an RFT modulo rr' which we will denote by $\mathcal{G} \vee \mathcal{G}'$.

Now let us look at how the parameters B_1 and B_2 for this new collection relate to the analogous parameters for the two components. Write the formula above as

$$\frac{s_{v,v'}(G \vee G')}{rr'} = \frac{s_{u,v}(G)}{r} + \frac{s_{u',v'}(G')}{r'}.$$

Recalling that $B_2(G,u,v) = |V(G)| - 1 + \frac{s_{u,v}(G)}{s(G)}$, and using the fact that "number of vertices minus 1" is additive over taking wedges, it follows that $B_2(G \vee G', v, v') = B_2(G,u,v) + B_2(G',u',v')$. That is, B_2 is additive over this construction. Recall that $\frac{B_1(G,u,v)}{s(G)} = (k_{min} - 1) + \frac{s_{u,v}(G)}{s(G)}$. The only way v and v' can be equal or adjacent in $G \vee G'$ is if either $u = v$ or $u' = v'$. It follows that $k_{min} - 1$ is subadditive, and hence

$$\frac{B_1(G \vee G', v, v')}{rr'} \leq \frac{B_1(G,u,v)}{r} + \frac{B_1(G',u',v')}{r'}.$$

That is, B_1/r is subadditive over this construction. (Since the examples of Section 5 all have the maximum for B_1 attained in the case where u, v are not adjacent, it will follow that the maximum over triples of B_1/r is actually additive for all our examples.)

Thus we have proven the following theorem.

Theorem 6.1. *Suppose $\mathcal{G}$ and $\mathcal{G}'$ are RFTs modulo relatively prime numbers r and r', respectively. Then $\mathcal{G} \vee \mathcal{G}'$ is an RFT modulo rr' and the parameters of these RFTs are related by*

$$\frac{B_1(\mathcal{G} \vee \mathcal{G}')}{rr'} \leq \frac{B_1(\mathcal{G})}{r} + \frac{B_1(\mathcal{G}')}{r'}$$

and

$$B_2(\mathcal{G} \vee \mathcal{G}') = B_2(\mathcal{G}) + B_2(\mathcal{G}').$$

Iterating this will allow us to build up linear upper bounds where B_2 and B_1/r are very small compared to r. This and a little analysis and number theory will let us prove much better upper bounds on α. But before we begin this, let me digress slightly.

Since we are proving an upper bound on $\alpha(n)$, we must be at least implicitly constructing a family of graphs Γ_n with $s(\Gamma_n) = n$. One weakness to the way in which we are presenting the proof is that this family is perhaps not as explicit as one would like. However unwinding the discussion above gives a fairly nice description of the family.

I think of the family as "necklaces", which consist of a cyclic sequence of "beads". Each bead is a graph G_i together with two distinguished vertices u_i and v_i and the necklace Γ is assembled by taking the disjoint union of the G_i and identifying v_i with u_{i+1}, interpreted cyclically. With one exception the beads are either cycles or theta graphs. The one exception, which can be thought of as a "clasp", is simply a path with its endpoints as the distinguished vertices. If we index the G_i with $0, 1, \ldots, m-1$ modulo m so that G_0 is the clasp and is a path of length k, then any spanning tree must either break the clasp (which can be done in k ways, each leading to $s(G_1)\cdots s(G_{m-1})$ spanning subtrees) or break one of the beads into a two component forest. If it breaks bead i into two components (which can be done in $s_{u_i,v_i}(G_i)$ ways), then there are $s(G_1)\cdots s(G_{i-1})s(G_{i+1})\cdots s(G_{m-1})$ ways to choose subtrees of the remaining beads. Thus we to see that

$$s(\Gamma) = s(G_1)\cdots s(G_{m-1})\left(k + \frac{s_{u_1,v_1}(G_1)}{s(G_1)} + \cdots + \frac{s_{u_{m-1},v_{m-1}}(G_{m-1})}{s(G_{m-1})}\right).$$

The term k ensures that in some sense we only need to build examples modulo the product $s(G_1)\cdots s(G_{m-1})$, and the key observation in this paper is that with a little cleverness we can arrange the other terms to cover all congruence classes. Note that this also means that an RFT constructed using wedges will involve a very large number of different graphs, unlike our starting examples which involved only two distinct graphs and many pairs of vertices.

Now look what happens when we apply Theorem 6.1 to the examples from Corollary 5.2.1. Fix a a set T of primes $p > 5$ congruent to 5 modulo 12. For any prime $p \in T$ we have an RFT with $B_1/p < p/4$ and $B_2 < p$. Let P_T and S_T denote the product and sum of the elements of T. Then applying Theorem 6.1 to all these RFTs will give an RFT modulo P_T with $B_1/P_T < S_T/4$ and $B_2 < S_T$. Thus Theorem 4.1 gives the following result.

Theorem 6.2. *For any set T of primes $p > 5$ congruent to 5 modulo 12 and any $n > S_T P_T/4$, we have $\alpha(n) \le n/P_T + S_T$.*

Thus we have a family of linear upper bounds where B_1/r and B_2, which both grow like S_T, are much smaller than $r = P_T$. This and a fair amount of number theorey implies a very strong bound on $\alpha(n)$.

Corollary 6.2.1. $\alpha(n) = O((\log n)^2/(\log\log n))$.

Proof. Call a prime q which is greater than 5 but congruent to 5 modulo 12 "good". Let $\pi_g(x)$ be the number of good primes less than or equal to x. Then Dirichlet's theorem on primes in arithmetic progressions, in its full power, says that $\pi_g(x) \sim \frac{x}{4\ln x}$, that is, good primes have Dirichlet density 1/4. Standard manipulations of this asymptotic result, give

$$\prod_{q<x} q = e^{x/4(1+o(1))}, \quad \sum_{q<x} q = \frac{x^2}{8\log x}(1+o(1)),$$

where the product and sum run over good primes smaller than x. Therefore if we take T to be all the good primes up to about $4\log X$, we will have $P_T S_T \approx X$. Turning this around, we define m to be the least value such that for T the first m good primes, we have $P_T S_T > 4n$. Then $m \sim \frac{\log n}{\log \log n}$ and hence the largest of these primes is $q_m \sim 4 \log n$. (Note that if we fixed a hard cut-off of $4 \log X$ for our good primes, then because the error term in our estimate for P_T is in the exponent, it would be hard to get a precise asymptotic for $P_T S_T$. However estimating m involves taking a logarithm, so we can get a precise asymptotic for m.) Now we give up exactly one of these primes to produce T'. Specifically, we give up the smallest prime such that after giving it up we have $P_{T'} S_{T'} < 4n$.

Since the ratio between consecutive good primes is bounded, there is a constant c so that $P_{T'} S_{T'} > 4n/c$. Roughly c is the bound by the ratio since that is the amount P decreases by when we replace the removed prime by the next smaller. The only catch is that S will also decrease, but much more slightly. Since S decreases by at most a factor of 2 (a much smaller fudge factor would work, but we will just use 2), we can take c to be twice the maximum ratio. Therefore

$$\alpha(n) \leq n/P_{T'} + S_{T'} < (c/4+1)S_{T'} < (c/4+1)(m-1)q_m = O\left(\frac{(\log n)^2}{\log \log n}\right).$$

□

7 A technical improvement

Corollary 6.2.1, which gives an upper bound that matches the conjectured upper bound (and the hard lower bound) to within a power, could have been taken as the main theorem of this paper. However with some fairly serious number theory we can push this bound a little bit further.

The basic examples we used to prove Corollary 6.2.1 came from a family of primes modulo which we built RFTs with B_1/p and B_2 both growing like p. This relatively fast growth rate is an artefact of the fact that both of the graphs we used had $O(p)$ vertices. We can do better, though it will require more sophistication. Since $ab+bc+ca = p$, the best we can hope for is to choose two graphs $\theta(a, b, c)$ both with $O(\sqrt{p})$ vertices, and we will see below that we can do this. Notice that this bound holds even more generally. Suppose we build an RFT modulo r which uses only a bounded number of basic graphs, say K of them, but many pairs of points u, v on these graphs. Since our RFT must cover all r congruence classes modulo r, at least one of these basic graphs must yield at least r/K pairs u, v. Hence at least one of the graphs must have at least $\sqrt{2r/K}$ vertices. Thus B_1/r and B_2 must both be of size at least $O(r^{1/2})$. Thus the result of this section cannot be improved without some further idea.

First we turn to the problem of finding an RFT modulo a prime p with B_1/p and B_2 both of size $\sqrt{p}$. There are many ways to do this, but we will focus on just one very specific construction. Note that

$$s(\theta(a+2b, 2a+3b, 3a+b)) = s(\theta(a+3b, 2a+b, 3a+2b)) = 11a^2 + 25ab + 11b^2.$$

Assume $p = 11a^2 + 25ab + 11b^2$ is a prime, and further assume that $p \equiv 1 \pmod 4$. Let $\left(\frac{x}{y}\right)$ denote the Jacobi symbol. Since

$$p = (4a+3b)\left(\frac{31}{9}a + \frac{11}{3}b\right) - \frac{25a^2}{9} = (4a+5b)\left(\frac{81}{25}a + \frac{11}{5}b\right) - \frac{49a^2}{25},$$

quadratic reciprocity gives (assuming without loss of generality that b is odd)

$$\begin{aligned}\left(\frac{(4a+3b)(4a+5b)}{p}\right) &= \left(\frac{p}{4a+3b}\right)\left(\frac{p}{4a+5b}\right)\\ &= \left(\frac{-25a^2/9}{4a+3b}\right)\left(\frac{-49a^2/25}{4a+5b}\right)\\ &= \left(\frac{-1}{4a+3b}\right)\left(\frac{-1}{4a+5b}\right) = -1.\end{aligned}$$

Hence one of $4a + 3b = (a + 2b) + (3a + b)$ and $4a + 5b = (a + 3b) + (3a + 2b)$ is a quadratic residue modulo p and the other is not. Thus these two theta graphs represent the two different cases. Thus the above results give the following.

Theorem 7.1. *Suppose a and b are positive integers such that $p = 11a^2+25ab+11b^2$ is a prime and is congruent to 1 modulo 4. Then for $n > 3(a + b)p/2$, we have $\alpha(n) \leq \frac{n}{p} + 6a + 6b - 2$.*

As desired, this is an improvement over Theorem 5.2 in the sense that B_1/p and B_2 are both of order $O(\sqrt{p})$. What is not obvious is that there are infinitely many such primes and that they have positive Dirichlet density. However, this follows from some (moderately heavy) algebraic number theory. For this the author is heavily indebted to two mathoverflow posts [6] and [7] (with a lovely answer by Brunault) and additional feedback given by one of the referees.

Theorem 7.2. *The set of primes p which can be written as $p = 11a^2+25ab+11b^2$ for positive integers a, b and are congruent to 1 modulo 4 has positive Dirichlet density.*

Proof. First, we drop the positivity condition on a, b and just look for primes of the form $11a^2+25ab+11b^2$. The discriminant of this quadratic is $141 = 3\cdot 47$ and we can write $11a^2 + 25ab + 11b^2 = 11(a + b)^2 + 3ab = 11(a - b)^2 + 47ab$. Thus we see after a little quadratic reciprocity that for a prime to be of this form it must be a quadratic nonresidue modulo both 3 and 47. Conversely, if p is such a prime, then 141 is a square modulo p and hence the prime ideal (p) splits in the number ring of $\mathbb{Q}[\sqrt{141}]$. But this number ring has class number 1, hence is a unique factorization domain. Thus this splitting means that we have integers a and b such that $11a^2+25ab+11b^2 = p$. (Specifically, the splitting means that there is an element of the number ring of norm $\pm p$, say $c + d\sqrt{141}$, where $c + d$ and $c - d$ are both integers. Looking at the congruence modulo 3 implies that the norm must be $-p$ so $c^2 - 141d^2 = -p$. Then setting $a = 59d - 5c$ and $b = 3c - 35d$, which are both integers, gives $c = (35a+59b)/2$ and $d = (3a + 5b)/2$ and $c^2 - 141d^2 = -p = -(11a^2 + 25ab + 11b^2)$.) The pair (a, b)

is not unique since we can multiply $c + d\sqrt{141}$ by any power of the fundamental unit $\alpha = 95 + 8\sqrt{141}$ and get another solution. However by Hecke equidistribution (see [2], Theorem 6 of Chapter XV), the logarithm $\log|c + d\sqrt{141}|$ modulo $\log\alpha$ viewed as an element of $\mathbb{R}/(\log\alpha)\mathbb{Z}$ is equidistributed. The condition that a and b are positive translates into an inequality $\frac{35}{3}d < c < \frac{59}{5}d$, and this in turn translates into $\log|c + d\sqrt{141}|$ lying in some open interval in $\mathbb{R}/(\log\alpha)\mathbb{Z}$. Thus a non-zero fraction of all such primes will have both a and b positive, and hence will be of the desired form. □

The argument given in the previous section using Corollary 5.2.1 and Dirchlet's theorem to conclude positive Dirichlet density now applies mutatis mutandis to the primes from Theorem 7.1 using Theorem 7.2 (hence Hecke equidistribution and the prime number theorem) to get positive Dirichlet density. Specifically, we call a prime of the required type "better". For any set T of better primes, we define P_T to be their product and

$$S'_T = \sum_{q\in T} \sqrt{q}.$$

Since $9(a + b)^2 < p$, the basic examples have $B_1/p < \sqrt{p}/2$ and $B_2 < 2\sqrt{p}$, so Theorem 4.1 implies

Theorem 7.3. *For any set T of primes of the form $11a^2+25ab+11b^2$ and congruent to 1 modulo 4 and any $n > S'_T P_T/2$, we have $\alpha(n) \leq n/P_T + 2S'_T$.*

From this theorem we argue almost verbatim the same way we did in the proof of Corollary 6.2.1. Letting $\pi_b(x)$ be the number of better primes up to and including x, Theorem 7.2 says $\pi_b(x) \sim \frac{x}{K\ln x}$ for some constant K. Again, this implies

$$\prod_{q<x} q = e^{x/K(1+o(1))}, \quad \sum_{q<x} q^{1/2} = \frac{x^{3/2}}{3/2\cdot K\log x}(1 + o(1)),$$

where the product and sum run over better primes smaller than x. Therefore if we take T to be all the better primes up to about $K\log X$, we will have $P_T S'_T \approx X$. Turning this around, we define m to be the least value such that for T the first m better primes, we have $P_T S'_T > 2n$. Then $m \sim \frac{\log n}{\log\log n}$ and hence the largest of these primes is $q_m \sim K\log n$. Now we give up exactly one of these primes to produce T'. Specifically, we give up the smallest prime such that after giving it up we have $P_{T'}S'_{T'} < 2n$ (and hence Theorem 7.3 applies).

Since the ratio between consecutive better primes is bounded, if we let c be twice this maximum ratio, then $P_{T'}S'_{T'} > 2n/c$. Therefore

$$\alpha(n) \leq n/P_{T'} + 2S'_{T'} < (c/2 + 2)S'_{T'} < (c/2 + 2)(m - 1)\sqrt{q_m} = O\left(\frac{(\log n)^{3/2}}{\log\log n}\right).$$

Thus we can end on our strongest result:

Corollary 7.3.1. $\alpha(n) = O((\log n)^{3/2}/(\log\log n))$.

Acknowledgements

The author is indebted to the referees for their help in cleaning up the presentation.

References

[1] J. Azarjia and R. Škrekovski, Euler's idoneal numbers and an inequality concerning minimal graphs with a prescribed number of spanning trees, *Math. Bohem.*, October 2012.

[2] S. Lang, "Algebraic Number Theory", Graduate Texts in Mathematics, Springer-Verlag, New York (1994).

[3] L. Nebeský, On the minimum number of vertices and edges in a graph with a given number of spanning trees, *Časopis pro pestování matematiky* 98 (1973), 95–97.

[4] Open Problem Garden, Minimal graphs with a prescribed number of spanning trees, `http://garden.irmacs.sfu.ca/category/graph_theory`, Accessed: 2022-01-04.

[5] J. Sedláček, On the minimal graph with a given number of spanning trees, *Canad. Math. Bull.* 13 (1970), 515–517.

[6] `https://mathoverflow.net/questions/133410/hecke-equidistribution`; Accessed: 2022-01-04.

[7] `https://mathoverflow.net/questions/65059/does-the-quadratic-form-x2-7y2-represent-infinitely-many-primes-with-`; Accessed: 2022-01-04.

(Received 19 Dec 2020; revised 20 Nov 2021, 15 Dec 2021)