Europ. J. Combinatorics (1980) **1**, 107–112

On a Diophantine Equation Arising in Graph Theory

R. L. GRAHAM

INTRODUCTION

In a recent paper of Schwenk and Watanabe [6], trees* whose adjacency matrices have only integer eigenvalues are investigated. In particular, they show that the "double star" $S_{m,n}$ consisting of a vertex of degree $m+1$, a vertex of degree $n+1$, and all other vertices having degree 1, has all eigenvalues integral if and only if the polynomial $x^2-(m+n+1)x+mn$ has roots which are perfect squares (of integers), say, a^2 and b^2. This means that the system

$$\begin{aligned} m+n+1&=a^2+b^2\\ mn&=a^2b^2 \end{aligned} \tag{1}$$

should have integral solutions. Solving for m and n, we obtain

$$\begin{aligned} m&=\tfrac{1}{2}(a^2+b^2-1\pm\sqrt{(a^2+b^2-1)^2-4a^2b^2}),\\ &=\tfrac{1}{2}(a^2+b^2-1\pm\sqrt{((a+b)^2-1)((a-b)^2-1)}),\\ n&=\tfrac{1}{2}(a^2+b^2-1\mp\sqrt{((a+b)^2-1)((a-b)^2-1)}). \end{aligned}$$

Setting

$$A=b-a,$$

$$B=b+a$$

we see that (1) is solvable if and only if

$$(A^2-1)(B^2-1)=C^2 \tag{2}$$

is solvable.

In this paper we determine all integer solutions to (2). It turns out somewhat unexpectedly that all solutions are just given by the values of Chebyshev polynomials evaluated at integers.

CHEBYSHEV POLYNOMIALS

The Chebyshev polynomials (of the first kind) $T_n(x)$ form a well-studied (e.g. see [4]) family of orthogonal polynomials which perhaps are most conveniently defined by:

$$T_n(\cos\theta)=\cos n\theta \tag{3}$$

for $n=0,1,2,\ldots$. The first few polynomials are:

$$\begin{aligned} T_0(x)&=1,\\ T_1(x)&=x,\\ T_2(x)&=2x^2-1,\\ T_3(x)&=4x^3-3x,\\ T_4(x)&=8x^4-8x^2+1,\\ T_5(x)&=16x^5-20x^3+5x. \end{aligned}$$

* For undefined graph theory terminology, see [1].

0195–6698/80/020107+06 \$01.00/0

In general, the $T_n(x)$ satisfy the linear recurrence

$$T_{n+2}(x) = 2xT_{n+1}(x) - T_n(x), \tag{4}$$

and, in particular, have integer coefficients. Associated with the $T_n(x)$ are the so-called Chebyshev polynomials of the second kind, denoted by $U_n(x)$. They are given by

$$U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta}$$

and satisfy the same linear recurrence (4) as do the $T_n(x)$, but with the starting values

$$U_0(x) = 1, \qquad U_1(x) = 2x.$$

Among the many identities (cf. [4]) relating the T_m's and U_n's are the following (which we will need later):

$$T_m^2(x) - 1 = (x^2-1)U_{m-1}^2(x), \tag{5}$$

$$T_{m+n}(x) - T_{m-n}(x) = 2(x^2-1)U_{m-1}(x)U_{n-1}(x), \qquad m \geq n. \tag{6}$$

It is already clear from (5) that if we choose $A = T_m(x)$, $B = T_n(x)$, then (2) will hold with $C = (x^2-1)U_{m-1}(x)U_{n-1}(x)$. What we next show is that this is the only way (2) can hold.

Determining the Solutions of (2)

Suppose (A, B, C) is a solution of (2), i.e.

$$(A^2-1)(B^2-1) = C^2.$$

We assume for now (without loss of generality) that $A \leq B$. Of course, we always assume A, B and C are non-negative. Define B' by

$$B' = AB - C. \tag{7}$$

Note that $B' > 0$ for $A > 0$. Thus

$$\begin{aligned}(A^2-1)(B^2-1) = C^2 &= (AB-B')^2 \\ &= A^2B^2 - 2ABB' + B'^2, \\ B^2 - 2ABB' + B'^2 + A^2 - 1 &= 0, \\ (B-AB')^2 &= (A^2-1)(B'^2-1).\end{aligned}$$

Setting

$$C' = B - AB', \tag{8}$$

we can rewrite this as

$$(A^2-1)(B'^2-1) = C'^2. \tag{9}$$

Furthermore, if we assume $A > 1$ then

$$B' < B. \tag{10}$$

For suppose not, i.e. suppose

$$B' = AB - C \geq B.$$

Then

$$(A-1)B \geqslant C,$$
$$(A-1)BC \geqslant C^2 = (A^2-1)(B^2-1),$$
$$BC \geqslant (A+1)(B^2-1),$$

since $A > 1$. Consequently

$$(A-1)B^2 \geqslant BC \geqslant (A+1)(B^2-1),$$
$$A+1 \geqslant 2B^2 \geqslant 2A^2 \quad \text{since } A \leqslant B,$$

which is a contradiction since $A > 1$. A similar argument shows that $C' > 0$ in this case.

Therefore, any solution (A, B, C) to (2) with $1 < A \leqslant B$ generates another solution (A, B', C') to (2) with

$$0 < B' < B, \qquad C' > 0.$$

By the symmetry of A and B in (2), if (A, B, C) is a solution then so is $(B, A, C) \equiv \rho(A, B, C)$. Hence, we can keep applying the transformation

$$\sigma : (A, B, C) \to (A, B', C') \tag{11}$$

(interchanging A and B when necessary to keep $A \leqslant B$) until we eventually reach a solution of the form $(1, x, 0)$. We call such a solution *primitive*. It is important to note that the transformation (11) is invertible, the inverse to

$$\sigma : \begin{array}{l} B' = AB - C, \\ C' = B - AB' = AC - (A^2-1)B, \end{array}$$

being given by

$$\tau : \begin{array}{l} B = AB' + C', \\ C = AC' + (A^2-1)B'. \end{array} \tag{12}$$

Therefore, we have shown that *every solution to* (2) *can be generated by starting with a primitive solution* $(1, x, 0)$ *and applying a sequence of transformations* τ *and* ρ *to it.*

Note that $(1, x, 0) \xrightarrow{\tau} (1, x, 0)$. Thus, the solutions begin as:

$$\begin{tikzcd}[column sep=tiny]
 & (1, x, 0) \arrow[loop left, "\tau"] \arrow[d, "\rho"'] & \\
 & (x, 1, 0) \arrow[u, bend right=60, "\rho"'] \arrow[d, "\tau"] & \\
 & (x, x, x^2-1) \arrow[loop right, "\rho"] \arrow[d, "\tau"] & \\
 & (x, 2x^2-1, 2x(x^2-1)) \arrow[dl, "\tau"'] \arrow[dr, "\rho"] & \\
(x, 4x^3-3x, (x^2-1)(4x^2-1)) & & (2x^2-1, x, 2x(x^2-1)) \\
\vdots & & \vdots
\end{tikzcd}$$

(where x is an arbitrary positive integer).

Suppose now that we have a solution to (2) with $A = T_m(x)$, $B = T_n(x)$. Thus, by (5),

$$\begin{aligned}(A^2-1)(B^2-1) &= (T_m^2(x)-1)(T_n^2(x)-1) \\ &= ((x^2-1)U_{m-1}(x)U_{n-1}(x))^2\end{aligned}$$

and so, by (6),

$$C=(x^2-1)U_{m-1}(x)U_{n-1}(x)$$
$$=\tfrac{1}{2}(T_{m+n}(x)-T_{|m-n|}(x)).$$

Applying τ to (A, B, C), we obtain

$$\begin{gathered}(T_m(x),\ T_n(x),\ \tfrac{1}{2}(T_{m+n}(x)-T_{|m-n|}(x)))\\ \downarrow^{\tau}\\ (T_m(x),\ T_{m+n}(x),\ \tfrac{1}{2}(T_{2m+n}(x)-T_n(x)))\end{gathered}$$

since

$$\begin{aligned}AB+C &= T_m(x)T_n(x)+\tfrac{1}{2}(T_{m+n}(x)-T_{|m-n|}(x))\\ &=\tfrac{1}{2}(T_{m+n}(x)+T_{|m-n|}(x))+\tfrac{1}{2}(T_{m+n}(x)-T_{|m-n|}(x))\\ &= T_{m+n}(x).\end{aligned}$$

Since all primitive solutions $(1, x, 0)$ can be written as $(T_0(x), T_1(x), 0)$ then the following result holds by induction.

THEOREM. *The integral solutions to*

$$(A^2-1)(B^2-1)=C^2$$

with A, $B>0$ *are exactly given by*

$$A=T_m(x), \qquad B=T_n(x)$$

for some choice of integers $x>0$, m, $n\geqslant 0$. *For these* A *and* B,

$$C=\tfrac{1}{2}(T_{m+n}(x)-T_{|m-n|}(x)).$$

EXTENSIONS

Equation (2) can be rewritten as

$$C^2+A^2+B^2-A^2B^2=1. \tag{2'}$$

It is natural to consider the more general equation

$$C^2+A^2+B^2-A^2B^2=\Delta. \tag{13}$$

It is not hard to show that the following modular restrictions hold:

$$\left.\begin{aligned}&\Delta\neq 8k+3, 6, 7\\ &\Delta\neq 4(8k+3)\\ &\Delta\neq 4^t(8k+7), \qquad t\geqslant 2\\ &\Delta\neq 9k+3, 6.\end{aligned}\right\} \tag{14}$$

Furthermore, these are the *only* modular restrictions which hold. Asymptotically this leaves $49/108=0{\cdot}454\ldots$ of the integers which are not sieved out by these expressions. However, not all of these can occur as a Δ. The least positive such integer is 88, i.e. 88 is not ruled out by any of the forms in (14) and yet $C^2+A^2+B^2-A^2B^2=88$ has no integer solutions. In fact, only $0{\cdot}388\ldots$ of the integers up to 10 000 are represented by $C^2+A^2+B^2-A^2B^2$. This ratio is surprisingly constant over relatively small intervals in this range. It is not known whether the set of possible Δ in (13) has a density or, if so, whether it is positive. The same techniques used in proving the Theorem can be used to show that whenever (13) has one solution, it has infinitely many, and further, all these solutions can be generated from a *finite* set of starting solutions, using the transformations τ and ρ.

A simple transformation reduces (2) to the equivalent equation

$$x^2+y^2+z^2-2xyz=1.$$

This is similar to the classic equation

$$x^2+y^2+z^2-3xyz=0 \tag{15}$$

studied by Markoff [5], Frobenius [3], Cassels [2] and others in connection with certain fundamental questions in diophantine approximation. Regarding (15), it is still unresolved whether for two solutions (x, y, z) and (x', y', z'), $x=x'$ implies $\{y, z\}=\{y', z'\}$. It is conjectured (see [2]) that it does. A similar conjecture for (2) is discussed in the next section.

Concluding Remarks

It follows from the Theorem that the possible values of a and b in the original system (1) are given by

$$\begin{aligned} a &= \tfrac{1}{2}(T_n(x)-T_m(x)),\\ b &= \tfrac{1}{2}(T_n(x)+T_m(x)) \end{aligned} \tag{16}$$

with $m \leqslant n$, where we have (arbitrarily) assumed that $a \leqslant b$. In order for a and b to be integers, we must choose x odd or $m \equiv n \pmod 2$.

As noted in [6], the system (1) is equivalent to the system

$$\begin{aligned} pq &= a^2,\\ (p+1)(q+1) &= b^2. \end{aligned} \tag{17}$$

It follows from above that the solutions to (17) (assuming w.l.o.g. that $p \leqslant q$) are given by

$$\begin{aligned} p &= \tfrac{1}{2}(T_{n-m}(x)-1),\\ q &= \tfrac{1}{2}(T_{n+m}(x)-1) \end{aligned}$$

where, as before, $(x-1)(m-n)$ must be even. The first few non-trivial solutions are:

$$\begin{array}{llll} 3.48=12^2 & 1.49=7^2 & 2.242=22^2 & 8.288=48^2\\ 4.49=14^2, & 2.50=10^2, & 3.243=27^2, & 9.289=51^2. \end{array}$$

It follows from (16) that the number of $b \leqslant N$ which satisfy (1) non-trivially (i.e. with $a<b$) is asymptotic to $\sqrt{N}$, thus verifying a conjecture in [6]. (The main contribution comes from taking $m=1$, $n=2$, and x odd).

No example is currently known of two non-trivial solutions (a, b) and (a, b') with $b \neq b'$. It is possible to generate the same solution (a, b) in many different ways. This can be done using the fact that the $T_n(x)$ obey the following composition rule:

$$T_{rs}(x)=T_r(T_s(x)).$$

Thus, the choices

$$m=rt, \qquad n=st, \qquad x=y$$

and

$$m = r, \qquad n = s, \qquad x = T_t(y)$$

give the same values of a and b. Perhaps this is the only way this can happen.

Finally, returning to the original graph theoretic question which motivated (1), it is still not known if there are trees having arbitrarily large diameter which have all integral eigenvalues. The double stars $S_{m,n}$ have diameter 3; no example with diameter exceeding 10 is currently known.

Acknowledgements

The author wishes to acknowledge stimulating discussions on this topic he has had with Allen Schwenk, who originally raised the question, and with F. R. K. Chung.

References

1. J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, American Elsevier, New York (1976).
2. J. W. S. Cassels, An Introduction to Diophantine Approximation, *Cambridge Tracts in Math. and Math. Physics*, No. 45, Cambridge University Press, (1957).
3. G. Frobenius, Über die Markoffschen Zahlen, *Preuss. Akad. Wiss. Sitzungsberichte* (1913), 458–487.
4. W. Magnus, F. Oberhettinger and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics*, Springer-Verlag, New York (1966).
5. A. Markoff, Sur les formes quadratiques binaires indéfinies, *Math. Ann.* **15** (1879), 381–406.
6. A. J. Schwenk and M. Watanabe, Integral starlike trees, *J. Australian Math. Soc.* (to appear).

Received 1 October 1979

R. L. Graham
Bell Laboratories, 600 Mountain Avenue,
Murray Hill, New Jersey 07974, U.S.A.