Industrial Mathematics Institute

2005:01

All but 49 numbers are Wiener indices of trees

H. Wang and G. Yu

IMI

Preprint Series

Department of Mathematics
University of South Carolina

All but 49 numbers are Wiener indices of trees

Hua Wang [1] and Guang Yu

Department of Mathematics,
University of South Carolina, Columbia, SC 29208
{hwang0, yu}@math.sc.edu

Abstract

The Wiener index is one of the main descriptors that correlate a chemical compound's molecular graph with experimentally gathered data regarding the compound's characteristics. A long standing conjecture on the Wiener index ([4], [5]) states that for any positive integer n (except numbers from a given 49 element set), one can find a tree with Wiener index n. In this paper, we prove that every integer $n > 10^8$ is the Wiener index of some short caterpillar tree with at most six non-leaf vertices. The Wiener index conjecture for trees then follows from this and the computational results in [8] and [5].

1 Introduction

The structure of a chemical compound is usually modelled as a polygonal shape, which is often called the *molecular graph* of this compound. The biochemical community has been using topological indices to correlate a compound's molecular graph with experimentally gathered data regarding the compound's characteristics.

In 1947, Harold Wiener [7] developed a topological index, the *Wiener Index*. This concept has been one of the most widely used descriptors in quantitative structure activity relationships, as Wiener index has been shown to have a strong correlation with the chemical properties of the chemical compound. Therefore, in order to construct a compound with a certain property, one may want to build some structure that has the corresponding Wiener index. From

[1] This research was supported in part by the NSF contracts Nr. 0072187 and 0302307.

this arises the important *inverse Wiener problem* ([4], [8]): Given a positive integer n, can we find a structure (graph) with Wiener index n?

Goldman *et al.* [2] solved the inverse Wiener problem for general graphs: they showed that for every positive integer n there exists a graph G such that Wiener index of G is n.

Since the majority of the chemical applications of the Wiener index deal with chemical compounds that have acyclic organic molecules, whose molecular graphs are trees, the inverse Wiener problem for trees attracts more attention and, actually, most of the prior work on Wiener indices deals with trees ([1]). When the graph is restricted to trees, the problem is more complicated. Gutman and Yeh [4] conjectured that:

Conjecture 1.1 *For all but a finite set of integers n, one can find a tree with Wiener index n.*

Lepović and Gutman [5] checked the integers up to 1206 and found that the following numbers are not Wiener indices of any trees:

2, 3, 5, 6, 7, 8, 11, 12, 13, 14, 15, 17, 19, 21, 22, 23, 24, 26, 27, 30, 33, 34, 37, 38, 39, 41, 43, 45, 47, 51, 53, 55, 60, 61, 69, 73, 77, 78, 83, 85, 87, 89, 91, 99, 101, 106, 113, 147, 159.

They claimed that the listed were the only "forbidden" integers and posed the following stronger version of Conjecture 1.1.

Conjecture 1.2 *There are exactly* 49 *positive integers that are not Wiener indices of trees, namely the numbers listed above.*

A recent computational experiment by Ban, Baspamyatnikh and Mustafa [8] shows that every integer $n \in [10^3, 10^8]$ is the Wiener index of some caterpillar tree (see Figure 1). Thus, the conjectures will be proved if one is able to show that every integer greater than 10^8 is the Wiener index of a tree.

Ban, Baspamyatnikh and Mustafa [8] conjectured that caterpillar trees with fixed diameter will represent all sufficiently large integers as their Wiener indices. For such trees, the Wiener index is a quadratic polynomial of the degrees of the non-leaf vertices, they recalled Lagrange's Theorem about representing positive integers as a sum of 4 squares as a related phenomenon.

In this short note, we prove the following

Theorem 1.1 *Every integer $n > 10^8$ is the Wiener index of a caterpillar tree* *of at most six non-leaf vertices.*

Combined with the computational results of [5] and [8], the theorem yields the trueness of the two conjectures. It also shows that Ban, Baspamyatnikh and Mustafa's conjecture is correct: every sufficiently large integer is the Wiener index of a caterpillar tree.

2 Preliminaries

For a tree $T = (V, E)$, denote by $d(v_i, v_j)$ the length of the path between two distinct vertices $v_i, v_j \in V$. The Wiener index $W(T)$ is then defined as

$$W(T) = \sum_{v_i, v_j \in V} d(v_i, v_j).$$

We define a family $\mathcal{T}$ of trees each of which is in the form of $T(x_1, x_2, x_3, x_4, x_5, x_6)$ where

$$V = \{s_1, \ldots, s_6\} \cup (\cup_{i=1}^{6} \{t_{i,1}, \ldots, t_{i,x_i}\}),$$

$$E = \{(s_i, s_{i+1}), 1 \le i \le 5\} \cup \{(t_{i,j}, s_i), 1 \le j \le x_i, 1 \le i \le 6\},$$

where $x_i, 1 \le i \le 6$, are nonnegative integers.

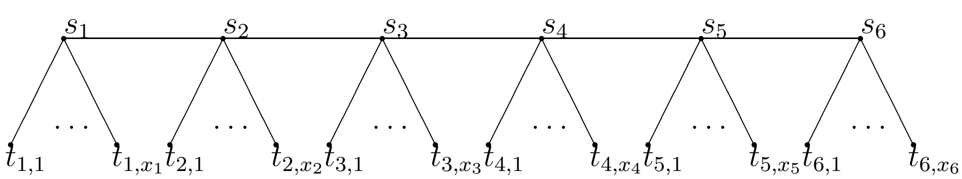

Fig. 1. tree $T(x_1, x_2, x_3, x_4, x_5, x_6)$

To find the Wiener index of such a tree, we quote the following result from [1]:

Lemma 2.1 *Given a tree $T = (V, E)$, one can assign weight $w(e)$ for each $e \in E$ such that $W(T) = \sum_{e \in E} w(e)$.*

Proof. For a sketchy proof of the lemma, one considers for any $e \in E$ the vertex sets for the two components of $T \setminus \{e\}$, say V_1 and V_2, lets $w(e) = |V_1||V_2|$, and then Lemma 2.1 follows. $\square$

In fact, the idea involved in the proof provides a way to calculate $W(T)$. For $T = T(x_1, x_2, x_3, x_4, x_5, x_6)$ as shown in figure 1, we have

$$W(T) = \sum_{i=1}^{5}\Big((\sum_{l=1}^{i} x_l + i)(\sum_{j=i+1}^{6} x_j + 6 - i)\Big) + (\sum_{i=1}^{6} x_i + 5)(\sum_{i=1}^{6} x_i). \tag{1}$$

Our proof of Theorem 1.1 relies on the well known result on the representation of integers as a sum of three squares ([3]).

Lemma 2.2 *Every integer $n \neq 4^k(8m-1)$ (where $k \geq 0$ and $m > 1$ are integers) is a sum of three integer squares.*

3 Proof of Theorem 1.1

In the expression (1), let

$$\begin{aligned} &x_1 + 1 = E - A,\ x_2 + 1 = A,\ \ x_3 + 1 = B, \\ &x_4 + 1 = C - B,\ \ x_5 + 1 = D - C,\ \ x_6 + 1 = E - D, \end{aligned} \tag{2}$$

then we get

$$W(T) = (9E^2 - 14E + 6) - (A^2 + B^2 + C^2 + D^2). \tag{3}$$

With this change of variables, it suffices to prove the following lemma.

Lemma 3.1 *Every integer $N > 10^8 - 6$ has a representation*

$$N = (9E^2 - 14E) - (A^2 + B^2 + C^2 + D^2) \tag{4}$$

with A, B, C, D and E all being integers and satisfying $E > D > C > B \geq 1$ and $E > A \geq 1$.

Let

$$\alpha = 1.2, \quad \beta = 1.208 \quad \text{and} \quad \gamma = 0.36. \tag{5}$$

For a proof of Lemma 3.1, we consider the possible solutions of (4) satisfying

$$E \in \mathcal{E}(N) := (E_1(N), E_2(N)] \quad \text{and} \quad D \in \mathcal{D}(N) := (\gamma\sqrt{N} - 1, \gamma\sqrt{N} + 1], \tag{6}$$

where

$$E_1(N) = \frac{14 + \sqrt{196 + 36\alpha N}}{18}, \quad E_2(N) = \frac{14 + \sqrt{196 + 36\beta N}}{18}.$$

We note that $E \in \mathcal{E}(N)$ is equivalent to $\alpha N < 9E^2 - 14E \leq \beta N$. We also note that, if $N > 10^8 - 6$, then

$$E_2(N) - E_1(N) > 12, \tag{7}$$

$$E_1(N) - (\gamma\sqrt{N} + 1) > 1, \tag{8}$$

$$(\gamma\sqrt{N} - 1)^2 > \beta N - (\gamma\sqrt{N} - 1)^2 - N, \tag{9}$$

and

$$\alpha N - N - (\gamma\sqrt{N} + 1)^2 > 0. \tag{10}$$

These inequalities guarantee that, if $N > 10^8 - 6$ and there is a solution to (4) with $E \in \mathcal{E}(N)$ and $D \in \mathcal{D}(N)$, then such a solution satisfies $E > D > \max\{A, B, C\}$.

Suppose $N > 10^8 - 6$. We choose D and E in accordance with $N(\text{mod} 4)$.

(i). $N \equiv 0, 1 \pmod 4$

Note that $\mathcal{D}(N)$ has length 2, we can fix an odd integer in $D \in \mathcal{D}(N)$. From (7), we see that $\mathcal{E}(N)$ contains at least 12 consecutive integers, among which there are at least two integers $e, e+3 \in \mathcal{E}(N)$ satisfying that, if $E = e$ or $e+3$, then $3|(9E^2 - 14E - D^2 - N)$ but $9 \nmid (9E^2 - 14E - D^2 - N)$. We choose from $\{e, e+3\}$ the odd integer for the value of E.

(ii). $N \equiv 2, 3 \pmod 4$

Similar to case (i), we can fix an even $D \in \mathcal{D}(N)$, then choose an even $E \in \mathcal{E}(N)$ such that $3|(9E^2 - 14E - D^2 - N)$ but $9 \nmid (9E^2 - 14E - D^2 - N)$.

In either case, with our choices of D and E, we always have

$$9E^2 - 14E - D^2 - N \equiv 1, 2 (\text{mod} 4) \quad \text{and} \quad 3||(9E^2 - 14E - D^2 - N). \tag{11}$$

From Lemma 2.2, we see that $9E^2 - 14E - D^2 - N$ is a sum of three squares. We further claim that $9E^2 - 14E - D^2 - N$ is actually the sum of three non-zero squares, at most two of which are equal. To see this, we first notice that $9E^2 - 14E - D^2 - N$ can not be a square or a sum of two (equal or unequal) square since $3||(9E^2 - 14E - D^2 - N)$. Moreover, if $9E^2 - 14E - D^2 - N = 3a^2$ for some integer a, then we have $a^2 \equiv 3(9E^2 - 14E - D^2 - N) \equiv 2, 3 (\text{mod} 4)$ which is impossible.

Thus, we conclude that

$$9E^2 - 14E - D^2 - N = a^2 + b^2 + c^2 \tag{12}$$

with integers $0 < a < b < c$, or

$$9E^2 - 14E - D^2 - N = 2a^2 + b^2 \tag{13}$$

with $a, b > 0$ and $a \neq b$.

In the first case, we take $A = a$, $B = b$ and $C = c$. In the latter case, let $A = a$, $C = \max\{a, b\}$ and $B = \min\{a, b\}$. Together with our choices for E and D, a solution to (4) found this way satisfies $E > D > C > B \geq 1$ and $E > A \geq 1$. □

Remark. With more careful choices of the values of α, β and γ in (5), together with some tedious discussion on the existence of D and E which satisfy our requirements (7)–(11) in shorter intervals, our proof for Theorem 1.1 works for $N > 2.9 \times 10^5$.

Acknowledgment. The authors thank Dr. L. A. Székely for letting us know about this problem and his help in writing this paper.

References

[1] Dobrynin, A. A., Entringer, R., Gutman, I., Wiener index of trees: Theory and applications. *Acta Applicandae Mathematicae* **66** (2001), 211–249.

[2] Goldman, D., Istrail, S., Lancia, G., Piccolboni, A., Algorithmic strategies in combinatorial chemistry, *Proc. 11th ACM-SIAM Sympos. Discrete Algorithms, 2000*, 275–284.

[3] Grosswald, Emil, Representations of integers as sums of squares, *Springer-Verlag*, 1985.

[4] Gutman, I., Yeh, Y., The sum of all distances in bipartite graphs, *Math. Slovaca* **45** (1995), 327–334.

[5] Lepović, M., Gutman, I., A collective property of trees and chemical trees, *J. Chem. Inf. Comput. Sci.* **38** (1998), 823–826.

[6] Li, Xueliang, Wang, Lusheng, Solutions for two conjectures on the inverse problem of the Wiener index of peptoids, *SIAM J. Discrete Math.* **17** (2003), No. 2, 210–218.

[7] Wiener, H., Structural determination of paraffin boiling points, *J. Amer. Chem. Soc.* **69** (1947), 17–20.

[8] Yih-En Andrew Ban, Sergei Bespamyatnikh, Nabil H. Mustafa, On a conjecture on wiener indices in combinatorial chemistry, *Proc. of the 9th International Computing and Combinatorics Conference '03 , 2003*, 509-518. (The jornal version will appear in Algorithmica, 2004.)