PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 79(93) (2006), 29–36 DOI: 10.2298/PIM0693029S

DETERMINATION OF LARGE FAMILIES AND DIAMETER OF EQUISEPARABLE TREES

Zoran Stanić

Communicated by Slobodan Simić.

Abstract. We consider the problem of determining all the members of an arbitrary family of equiseparable trees. We introduce the concept of saturation (based on the number partitions). After that, we use the same concept to obtain the least upper bound for the difference in the diameters of two equiseparable trees with m edges. We prove that this bound is equal to $\frac{m-4}{3}$, where m is the size of trees.

1. Introduction

Let T be a tree and e an arbitrary edge of T. Then $T-e$ consists of two components with $n_1(e)$ and $n_2(e)$ vertices. Conventionally, $n_1(e) \leqslant n_2(e)$. If T' and T'' are two trees of the same order n and if their edges $e'_1, e'_2 ..., e'_{n-1}$ and $e''_1, e''_2 ..., e''_{n-1}$ can be labelled so that $n_1(e'_i) = n_1(e''_i)$ holds for all $i = 1, 2, ..., n-1$, then T' and T'' are said to be equiseparable. For the notion of equiseparable trees and for definition of the values of $n_1(e)$ and $n_2(e)$ we refer to Gutman (see [**1**]). The definitions of quantities $n_1(e)$ and $n_2(e)$ in the case of an arbitrary graph is more complex and it is given also in [**1**]. It is shown (see [**6**]) that almost all trees have equiseparable mates (in the sense that the ratio of the cardinality of trees without any equiseparable mate to the ones with at least one equiseparable mate tends to zero when their order converges to infinity). One general method for constructing the equiseparable trees is presented in [**2**]. Let T, X and Y be arbitrary trees, each on more than two vertices. Let u and v be two vertices of T, p a vertex of X and q a vertex of Y. Let tree T' be obtained from T, X and Y by identifying the vertices u and p and by identifying the vertices v and q. Let T'' be obtained in the same way when p and q change places. In order that T' differs from T'', the fragments X and Y (when attached via vertices p and q) are required to be different. If X

2000 *Mathematics Subject Classification.* 05C05; 11P81.
Key words and phrases. equiseparable tree, saturation, number partition, diameter.
Partially supported by the Grant 144032D of the Serbian Ministry of Science, Technology and Development.

and Y have an equal number of vertices, then trees T' and T'' are equiseparable. The statistics of equiseparable trees of order n, ($7 \leqslant n \leqslant 20$) is presented in [**5**]. It is easy to show that there are no equiseparable mates for $n < 7$, and the only pair of such trees for $n = 7$ is depicted in Figure 1. The largest family of equiseparable trees for $n = 20$ has cardinality 603. The notion of equiseparable trees is closely connected to the Wiener index (mentioned for the first time in [**7**], and later well studied in chemistry and graph theory) and the Zenkevich index (see [**8**]). These indices can be expressed as

$$W(G) = \sum_{i=1}^{n} n_1(e_i)n_2(e_i)$$

and

$$Z(G) = \sqrt{gn} \sum_{i=1}^{n} \frac{1}{\sqrt{(gn_1(e_i) + h)(gn_2(e_i) + h)}},$$

where G is an arbitrary graph on n vertices, and $g \approx 14$, $h \approx 1$. For more details on equiseparable trees and their implementations one can see [**2**], [**3**] and [**5**].

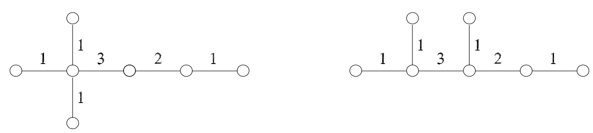

FIGURE 1. Pair of equiseparable trees of the smallest order. Each edge e is labelled by the corresponding value $n_1(e)$.

In Section 2, we present a method for the determination of large families of equiseparable trees. This method is not only based on graph theory but also on some notions of number theory. By observating known families of equiseparable trees we saw that the diameters of two equiseparable trees do not differ too much. This motivates us to consider (in Section 3), the problem of determining the upper bound for $|\mathrm{dm}(T') - \mathrm{dm}(T'')|$, where T' and T'' are two such trees. We prove that this bound is equal to $\dfrac{m-4}{3}$, where m is the size of trees.

2. Determination of Equiseparable Trees of Large Order

Furthermore, let m denote the size of an arbitrary tree. Obviously, every family of equiseparable trees is uniquely determined by the values

$$n_1(e_1), n_1(e_2), ..., n_1(e_m), \tag{2.1}$$

which can be considered as a non–decreasing sequence. We consider the problem of determining all equiseparable trees for a given sequence (2.1). Firstly we ensure that the given sequence is *graphable* (in the sense that it generates at least one tree). Let $e = xy$ be an arbitrary edge of a tree T and let $n_1(e)$ be the corresponding value. This means that one component of $T - e$ has $n_1(e) - 1$ edges, while the other component has at least $n_1(e) - 1$ edges. Suppose that the vertex x belongs to the first component. Then, the sum of values corresponding to edges different from e and incident to x, equals $n_1(e) - 1$. In other words, these values represent one partition of the number $n_1(e) - 1$. (The partition of the number $n \in \mathbf{N}$ is a set of positive integers $\{a_1, a_2, ..., a_k\}$, such that $a_i \neq 0$, $a_i \leqslant n$, $i = 1, 2, ..., k$ and $\sum_{i=1}^{k} a_i = n$. For more details see [**4**].) We say that the element k of the sequence (2.1) is *saturated*, if there is a partition of the number $k - 1$ among the other elements of the same sequence, such that the same element cannot be used in more than one saturation. Note that the element $k = 1$ is saturated by default. By the recursion, we get that for saturation of element k we need exactly $k - 1$ elements which are less than k. Clearly, the sequence is graphable if and only if each of its elements is saturated. Also, having a saturation of all elements, we can easily reconstruct the corresponding tree. For example, in the sequence $[1, 1, 1, 1, 2, 3]$, 3 is saturated by 2, and 2 is saturated by 1, while ones are saturated by default. This can be written as $3(2(1)), 1, 1, 1$ (from which we reconstruct the first tree of Figure 1). Also, the elements of the same sequence can be saturated in the following way: 3(1, 1), 2(1), 1 (which leads us to the second tree of Figure 1).

In the following two lemmas we give a necessary and sufficient conditions for the sequence (2.1) be graphable.

Lemma 2.1. *If a non–decreasing sequence $n_1(e_i)$, $i = 1, 2, ..., m$, is graphable, then the following inequality holds:*

$$|E_k| \leqslant \frac{|E_1| + |E_2| + ... + |E_{k-1}|}{k - 1},\ k = 2, 3, ..., \left\lfloor \frac{m+1}{2} \right\rfloor,$$

where E_k denotes the set of edges e_i such that $n_1(e_i) = k$, $i \in \{1, 2, ..., m\}$.

Proof. For the saturation of $n_1(e_i) = k$ we need $k - 1$ elements which are less than k (all these elements correspond to the edges of the same component of $T - e_i$). Therefore, we have

$$(k - 1)|E_k| \leqslant |E_1| + |E_2| + ... + |E_{k-1}|.$$

Also, since $n_2(e_i) \geqslant n_1(e_i)$, we need at least $k - 1$ elements less than k, which correspond to the edges of the other component. This implies $k \leqslant \left\lfloor \frac{m+1}{2} \right\rfloor$.

□

Lemma 2.2. *The non–decreasing sequence $n_1(e_i)$, $i = 1, 2, ..., m$, is graphable if these two conditions hold:*

(i): *there is a saturation of each element* $n_1(e_i)$ *of the corresponding sequence (all these elements correspond to the edges of one component of* $T - e_i$*);*

(ii): *there are at least* $k-1$ *elements less than* k *which correspond to the edges of the other component.*

PROOF. Obviously, a non–decreasing sequence $n_1(e_i)$, $i = 1, 2, ..., m$ corresponds to some tree if the edges of that tree can be labelled by elements of the given sequence such that for every edge e_i its label $n_1(e_i)$ determines the order of one component of $T - e_i$ (this is provided by condition (i)), where the order of this component is less than or equal to the order of the other component (this is provided by condition (ii)). This completes the proof. □

Recall that the saturation of an arbitrary element $n_1(e_i)$ is based on the partition of the number $n_1(e_i) - 1$ into the smaller elements of the given sequence. Therefore, we immediately obtain the following lemma.

LEMMA 2.3. *The number of different saturations of the element* n *is less than or equal to the number of all partitions* $p(n)$ *of* n.

The following table can, therefore, be used to obtain an upper bound for the number of saturations of an arbitrary element.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	
$p(n)$	1	2	3	5	7	11	15	22	30	42	56	77	101	...

TABLE 1. The cardinality of different partitions of the number $n \in \mathbf{N}$.

The concept of saturation yields all equiseparable trees for a given graphable sequence. For example, the sequence $[1, 1, 1, 1, 2, 2, 3, 3, 4]$ generates three equiseparable trees, since we can saturate all its elements in three different ways. These saturations are

$$\left\{\begin{array}{l} 4(3(2(1))) \\ 3(2(1)) \\ 1 \\ 1 \end{array}\right. , \qquad \left\{\begin{array}{l} 4(3(2(1))) \\ 3(1,1) \\ 2(1) \end{array}\right. , \qquad \left\{\begin{array}{ll} 4(3(1,1)) & \longleftarrow 1.\,\text{tail} \\ 3(2(1)) & \longleftarrow 2.\,\text{tail} \\ 2(1) & \longleftarrow 3.\,\text{tail} \end{array}\right.$$

Note that some elements are not included in the saturation of other elements (in the third representation these elements are 4, 3 and 2). The edges which correspond to these elements are incident to the same vertex. One can observe that every representation can be naturally partitioned into *tails* (see the third representation). Each tail is determined by its largest element k and consists of exactly k elements. It is an easy exercise to prove that this is a property of all tails. An element of a tail together with its saturation we call a *subtail*.

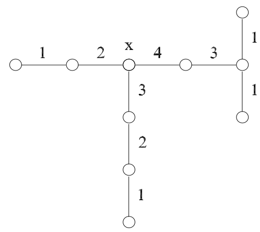

FIGURE 2. One of the three trees corresponding to the sequence $[1,1,1,1,2,2,3,3,4]$. Three tails are determined by edges labelled by 4, 3 and 2 and these edges are incident to the same vertex x.

The concept of saturation enables us to avoid searching among all trees of the same order, and therefore, presents a method for obtaining very large families of equiseparable trees. Also, the concept of saturation will be a useful tool in the next Section.

3. Diameter of Equiseparable Trees

Here we consider the difference between the diameters of the equiseparable trees. Furthermore we do not make any distinction between an element of some (sub)tail and the corresponding edge. Therefore, when we talk about paths in a (sub)tail, we think of paths in the corresponding tree.

LEMMA 3.1.

(i): *The longest path in a (sub)tail which includes the largest element of the (sub)tail is shorter than or equal to the number of distinct elements of the (sub)tail.*

(ii): *For the elimination of element k of a (sub)tail, different from the largest element, (such that the saturation is preserved) we need k new elements.*

PROOF. The longest path (in a (sub)tail) which contains the largest element also contains some of the elements labelled by 1 (this path must ends by some terminal vertex). Such a path is represented by a strictly decreasing subsequence of the elements in a (sub)tail. Since the length of this path equals to the number of elements in a such subsequence, statement (i) follows. Furthermore, every element k is saturated by $k-1$ other elements. Also, the element k is included in the saturation of the largest element of the (sub)tail. If we eliminate the element k from this (sub)tail (i.e. remove it together with its saturation in some other (sub)tail, and put the other elements in its place), we must put exactly k new elements instead of the eliminated element (the saturation of the largest element must be preserved), and the statement (ii) is proved. $\square$

We illustrate the statements of the previous lemma by an example. Consider the tail 8(4(3(2(1))), 3(2(1))). The longest path which includes the edge labelled by 8 consists of edges labelled by 8, 4, 3, 2 and 1, respectively (these edges are labelled by strictly decreasing elements, and the length of the path is 5). We can eliminate the element 4 from this tail by forming a new tail in the following way (the added elements are underlined):

$$\begin{cases} 8(\underline{1}, \underline{1}, \underline{1}, \underline{1}, 3(2(1))) \\ 4(3(2(1))) \end{cases}$$

Naturally, the elimination of some element is possible only if we have a sufficient number of new elements in the remaining tails for that and it is not uniquely determined. Also, the elimination of some element will change the tree, but, after elimination, we will obtain an equiseparable mate of the given tree. When eliminating some element k, at most $k-2$ elements, which are less than k, can be eliminated. Incidentally, in the previous elimination, beside the element 4, we transfer the elements 3 and 2 in the other tail.

THEOREM 3.1. *If T' and T'' are equiseparable mates with m edges, then* $|\mathrm{dm}(T') - \mathrm{dm}(T'')| \leqslant \dfrac{m-4}{3}$.

PROOF. We have seen that a tree can be represented by its subtails. The longest path in a tree (whose length represents a diameter) is determined by two (sub)tails, so that the largest elements of these (sub)tails correspond to the edges which belong to this path. With no loss of generality, we can assume that diameter $\mathrm{dm}(T') = d'$ $(d' \geqslant 2)$ is determined by the two tails:

$$\begin{cases} a(...1) & \longleftarrow \text{tail}\, A \\ b(...1) & \longleftarrow \text{tail}\, B \end{cases}$$

We have that these tails correspond to exactly $a+b$ edges. Consider the problem: which is the minimal number of edges in tree T' that can be transferred so that decreasing the diameter (by elimination of some elements) to d'' $(d'' < d')$ is possible? In this case, we can assume that $a+b = d'$ (in other words: there are no equal elements in each of the chosen tails or, equivalently, the tails have a form $A = a(a-1(a-2(...2(1))...)))$ and $B = b(b-1(b-2(...2(1))...)))$). If $a+b > d'$, then we have indeed elements which are in excess in tails A and B, in sense that the corresponding edges do not belong to the largest path.

Suppose that $d'' = a+b-(k+l)$, $1 \leqslant k \leqslant a$, $1 \leqslant l \leqslant b$. In order to decrease the diameter $a+b$ to $a+b-(k+l)$, we, certainly, need to decrease the number of different elements in A to $a-k$ and in B to $b-l$. Therefore, by Lemma 3.1, the minimal number of new edges is $k+l+2$. After decreasing, we obtain two new tails: A' (with $k+1$ elements) and B' (with $l+1$ elements) and repeat the procedure in order to decrease the number of different elements in each of them to $\min\{a-k,\, b-l\}$ and so on. Repeation of the procedure stops when the last tails obtained consist of at most $\min\{a-k,\, b-l\}$ different elements, each. Therefore, to decrease the number of different elements in A and B we need $\alpha \geqslant k+1$ and $\beta \geqslant l+1$ new elements, respectively.

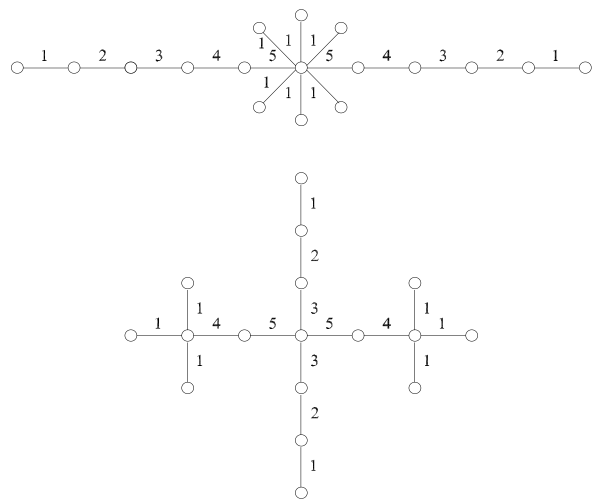

FIGURE 3. Equiseparable mates for $m = 16$. The diameters differ by $\left\lfloor \dfrac{m-4}{3} \right\rfloor = 4$.

The difference between diameters is equal to $d' - d'' = k + l$. In order to prove that this difference is less than or equal to $\dfrac{m-4}{3}$, it is sufficient to prove:

$$k + l \leqslant \frac{(a + b + \alpha + \beta) - 4}{3},$$

and, since $\dfrac{m-4}{3}$ monotonically increases, it is sufficient to prove the case for $\alpha = k + 1$, $\beta = l + 1$:

$$k + l \leqslant \frac{a + b + k + 1 + l + 1 - 4}{3},$$

In this case $a \geqslant 2k+1$ and $b \geqslant 2l+1$ hold. Hence, if we substitute a by $2k+1$, and b by $2l + 1$ in the inequality above we get $k + l \leqslant k + l$, and the proof follows. □

Let d' and d'' be the diameters of two equiseparable trees, and let $d' \geqslant d''$. By using the same notation as in the previous theorem we get

$$d' = a + b \geqslant 2(k + l) + 2 \quad \text{and} \quad \alpha + \beta \geqslant k + l + 2. \tag{3.1}$$

Since, for the common size of both trees we have $m = d' + \alpha + \beta$, we get $m \geqslant 3(k + l) + 4$, where $d' - d'' = k + l$. Therefore, if we have the equalities in (3.1), we obtain $d' - d'' \leqslant \dfrac{m-4}{3}$, and this explains our upper bound. Equiseparable trees with extremely different diameters can be obtained in a similar way as those in Figure 3

(Here, we have $k = l = 2$ and the mentioned equalities hold.). However, for $m = 13$, we have $\left\lfloor \frac{m-4}{3} \right\rfloor = 3$, but this bound is not reached for any pair of equiseparable trees on 13 edges. One can show that the same holds for $m = 13 + 6s,\ s = 1, 2, ...$ In these cases, the integer value of our bound can be decreased by 1.

References

[1] I. Gutman, *A formula for the Wiener number of trees and its extension to graphs containing cycles*, Graph Theory Notes, New York, **27** (1994) (9–15).

[2] I. Gutman, B. Furtula, D. Vukičević, B. Arsić, *Equiseparable molecules and molecular graphs*, Indian J. Chem., **43** (2004) (7–10).

[3] I. Gutman, D. Vidović, B. Arsić, Ž. Bošković, *On the relation between Zenkevich and Wiener indices of alcanes*, J. Serb. Chem. Soc., vol. 69, no. 4 (2004) (265–272).

[4] G. H. Hardy, E. M. Wright, *An Introduction to the Theory of Numbers*, Calarendon Press, Oxford, 1979.

[5] O. Miljković, B. Furtula, I. Gutman, *Statistics of equiseparable trees and chemical trees*, MATCH Commun. Math. Comput. Chem. **51** (2004) (179–184).

[6] D. Vukićević, I. Gutman, *Almost all trees and chemical trees have equiseparable mates*, J. Comput. Chem. Jpn., **3** (2004) (109–112).

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

[8] I. G. Zenkevich, *Dependence of chromatographic retention indices on the dynamic characteristics of molecules*, Rus. J. Phy. Chem. **73** (1947) (797–801).

Matematički fakultet
Studentski trg 16
11000 Beograd
Serbia
zstanic@matf.bg.ac.yu

(Received 21 06 2005)
(Revised 23 09 2006)