

# ROOTED TREE STATISTICS FROM MATULA NUMBERS

EMERIC DEUTSCH

Polytechnic Institute of New York University  
emerideutsch@msn.com

ABSTRACT. There is a one-to-one correspondence between natural numbers and rooted trees; the number is called the Matula number of the rooted tree. We show how a large number of properties of trees can be obtained directly from the corresponding Matula number.

## 1. INTRODUCTION

Let  $p_m$  denote the  $m$ -th prime number ( $p_1 = 2, p_2 = 3, p_3 = 5, \dots$ ). We call  $m$  the *order* of  $p_m$ . For a positive integer  $n$ , we denote by  $\Omega(n)$  the number of prime divisors of  $n$ , counted with multiplicities.

For a rooted tree  $T$ , its *Matula number*  $\mu(T)$  is defined recursively in the following manner [23]. For the 1-vertex tree  $T$  we define  $\mu(T) = 1$ . Otherwise, let  $T_1, T_2, \dots, T_d$  be the subtrees of  $T$  rooted at the vertices  $v_1, v_2, \dots, v_d$ , adjacent to the root of  $T$  (see Fig. 1). Then, we define

$$\mu(T) = p_{\mu(T_1)} p_{\mu(T_2)} \cdots p_{\mu(T_d)}.$$

Conversely, given a natural number  $n$ , the above procedure can be done in reverse, leading to a unique rooted tree  $\tau(n)$  with Matula number  $n$ . We illustrate this on the arbitrarily chosen  $n = 987654321$ . We label the root of the desired  $\tau(n)$  by  $987654321 = 3 * 3 * 17 * 17 * 379721$ . Since we have 5 prime factors, the degree of the root is 5. We draw 5 edges from the root and we label their endpoints, for example from left to right, by the orders of these 5 primes, i.e. by 2, 2, 7, 7, 32277 since  $p_2 = 3, p_7 = 17$ , and  $p_{32277} = 379721$ . More conveniently, instead of 32277 we use the label  $32277 = 3 * 7 * 29 * 53$ . Now we repeat this procedure for each of the 5 newly labeled vertices. For example, considering one of the vertices labeled 2, we draw one edge whose endpoint will have the label 1 (the order of the prime 2). This is a leaf of the required tree  $\tau(n)$  and the procedure stops along this branch. The reader is asked to follow in detail the way the tree  $\tau(987654321)$  has been obtained (Fig. 2).

We will use the following slightly modified description of the above defined correspondence. We give the mapping  $\tau$  from natural numbers to rooted trees, i.e. the inverse of the  $\mu$  function. We define  $\tau(1)$  to be the 1-vertex tree. For a prime number  $p_t$  (the  $t$ -th prime), we define  $\tau(p_t)$  recursively as shown in Fig. 3a. If  $n$  is composite,  $n = rs$  ( $r, s \geq 2$ ), then  $\tau(rs)$  is defined recursively to be the tree shown in Fig. 3b, i.e. the trees  $\tau(r)$  and  $\tau(s)$  joined at their roots. Clearly, the obtained rooted tree does not depend on the used factorization  $rs$  of  $n$ .

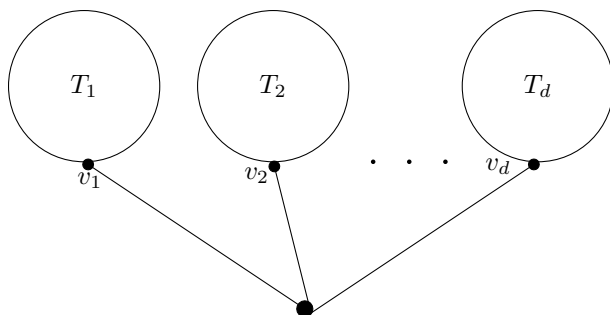


FIGURE 1.

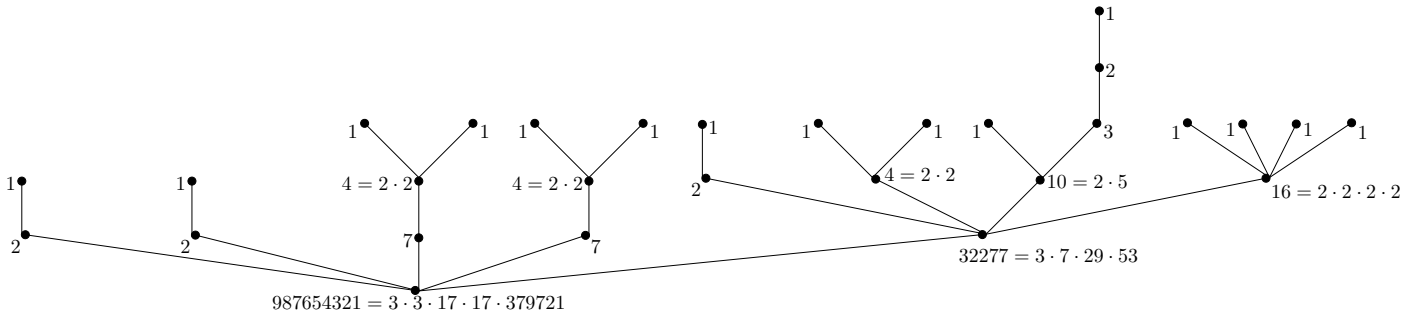


FIGURE 2.

**Remark.** The prime/composite dichotomy has led us to the same recursive construction of the rooted trees that is mentioned by Czabarka, Székely, and Wagner [4]. Although that paper does not consider the Matula bijection, there may be some overlap when the same statistic is investigated.

As pointed out by Ivan Gutman and Yeong-Nan Yeh [17], it is of interest to find statistics on rooted trees  $T$  directly from their Matula numbers  $\mu(T)$ . This has been done for several statistics in [8], [10], [13].

In this paper we will do the same in a slightly different way and also for several other properties like, for example, various topological indices (Wiener, Zagreb1, Zagreb2, Randić, etc.). We will also consider polynomial-valued statistics (i.e. a finite sequence of statistics), namely the partial Wiener polynomial, the Wiener polynomial, the degree sequence polynomial, and the exit-distance polynomial (the last two terms will be defined below).

## 2. TERMINOLOGY

A *rooted tree* is a tree having a distinguished vertex, called the *root*. In a rooted tree the *level* of a vertex  $v$  is its distance from the root, i.e. the length of the unique path from the root to  $v$ . The *height* (called also *depth*) of a rooted tree is the length of the longest path from the root. In a rooted tree if vertex  $v$  immediately precedes vertex  $w$  on the path from the root to  $w$ , then  $v$  is the *parent* of  $w$  and  $w$  is the *child* of  $v$ . In a rooted tree vertices having the same parent are called *siblings*. A *leaf* in a rooted tree is any vertex having no children. However, in the 1-vertex tree, the root is not considered to be a leaf. Given a rooted tree, the *path length* is the sum of the levels of each of the nodes. The *external path length* is the sum of the levels of each of the leaves (see, for example, [28]). The *distance* between vertices  $i$  and  $j$  of a tree  $T$  is the number of edges on the unique path from  $i$  to  $j$ ; it is denoted by  $d_T(i, j)$  (or  $d(i, j)$  when no ambiguity is possible). The *diameter* of a tree is the greatest distance between pairs of vertices. The *degree* of a vertex  $i$  of a tree  $T$  is the number of edges emanating from  $i$ ; it is denoted by  $deg_T(i)$  (or  $deg(i)$  when no ambiguity is possible). A vertex of degree 1 is called a *pendant vertex* while a vertex of degree  $\geq 3$  is called a *branching vertex*. A *root subtree* of a rooted tree is any subtree that contains the root.

The *degree sequence* of a graph is the list of vertex degrees, written in nonincreasing order. This is basically the *comparability index* of a tree introduced by Gutman and Randić [15] (see also [30]). For example, for the path on 5 vertices, the degree sequence is  $(2, 2, 2, 1, 1)$ . We define the *degree sequence polynomial* of a graph with vertex set  $\{1, 2, \dots, n\}$  as  $\sum_{i=1}^n x^{deg(i)}$  (it is the generating polynomial of the vertices of the graph with respect to vertex degree). For example, for the path on 5 vertices the degree sequence polynomial is  $2x + 3x^2$ .

The *visitation length* of a rooted tree  $T$  is defined as the sum of the number of nodes of  $T$  and of its path length [21].



FIGURE 3.

Given a vertex  $v$  in a rooted tree  $T$ , we define the *exit distance* of  $v$ , denoted  $\lambda(v)$ , to be the distance from  $v$  to the nearest leaf of  $T$  that is a descendant of  $v$ . We are interested in three statistics on a rooted tree  $T$ , connected with this new concept: (i) the sum of the exit distances of all vertices of  $T$ , (ii) the largest exit distance, and (iii) the number of vertices for which the largest exit distance is attained. To obtain all three statistics we define the *exit-distance polynomial* of  $T$  by  $\sum x^{\lambda(v)}$ , where the summation extends over all vertices  $v$  of  $T$ ; it is the generating polynomial of the vertices of the tree with respect to exit distance. To label the vertices of a rooted tree by their exit distances, label the leaves with 0, label their parents with 1, label the so far unlabelled parents of the 1's with 2, label the so far unlabelled parents of the 2's with 3, and so on. Note the monotonicity of the coefficients of this polynomial, due to the fact that each vertex with exit distance  $k$  ( $k \geq 1$ ) is the parent of some vertex with exit distance  $k - 1$ . The reader is asked to verify that the exit-distance polynomial of the rooted tree in Fig. 2 is  $15 + 9x + 5x^2$ .

The *Wiener index* of a tree  $T$  is the sum of the distances between all unordered pairs of vertices of  $T$  [31], [6].

The *terminal Wiener index* of a tree  $T$  is the sum of the distances between all unordered pairs of pendant vertices of  $T$  [12], [29].

The *hyper-Wiener index* of a tree  $T$  has been defined by M. Randić [26], [9]. Later, Klein, Lukovits, and Gutman [22] have derived the formula  $\frac{1}{2}(\sum d(i, j)^2 + \sum d(i, j))$ , where the summations extend over all unordered pairs of vertices. This is now accepted as the definition of the hyper-Wiener index of a connected graph.

The *multiplicative Wiener index* of a tree is the product of the distances between all unordered pairs of vertices of  $T$  [14].

The *Wiener polarity index* of a tree  $T$  is defined as the number of unordered pairs of vertices  $i, j$  of  $T$  such that  $d(i, j) = 3$  [31].

The *first Zagreb index* of a tree  $T$  is defined as  $\sum deg(i)^2$ , where the summation is over all the vertices  $i$  of  $T$  [16], [11], [25].

The *second Zagreb index* of a tree  $T$  is defined as  $\sum deg(i)deg(j)$ , where the summation is over all the edges  $ij$  of  $T$  [16], [25], [5].

The *Narumi-Katayama index* of a tree  $T$  is defined as  $\prod deg(i)$ , where the product is taken over all the vertices  $i$  of  $T$  [24].

The *first multiplicative Zagreb index* of a tree  $T$  is defined as  $\prod deg(i)^2$ , where the product is taken over all the vertices  $i$  of  $T$  [10]. It is the square of the Narumi-Katayama index.

The *second multiplicative Zagreb index* of a tree  $T$  is defined as  $\prod deg(i)deg(j)$ , where the product is taken over all the edges  $ij$  of  $T$  [10]. An other equivalent expression is  $\prod deg(i)^{deg(i)}$ , where the product is taken over all vertices  $i$  of  $T$  (see Lemma 3.1 in [10]).

The *Randić (connectivity) index* of a tree  $T$  is defined as  $\sum (deg(i)deg(j))^{-\frac{1}{2}}$ , where the summation is over all the edges  $ij$  of  $T$ . It was defined by Randić [26] under the name branching index. Bollobás and Erdős [2] generalized these indices by defining the *general Randić index* as  $\sum (deg(i)deg(j))^\alpha$ , where  $\alpha$  is any real number and the summation is again over all the edges  $ij$  of  $T$ . Note that for  $\alpha = 1$  this becomes the second Zagreb index, defined above.

The *Wiener polynomial* (called sometimes Hosoya polynomial or Wiener-Hosoya polynomial) of a tree  $T$  is defined as  $\sum x^{d(i,j)}$ , where the sum is taken over all unordered pairs of distinct vertices  $(i, j)$  of  $T$  [18], [27].

For a tree  $T$ , the *partial Wiener polynomial* with respect to a vertex  $w$  of  $T$  is defined as  $\sum x^{d(i,w)}$ , where the sum is taken over all vertices  $i$  of  $T$ , distinct from  $w$  [7].

### 3. STATEMENTS AND PROOFS

The proofs of our statements are based on examining how a particular statistic  $S$  on each of the two trees of Fig. 3 is obtained from the values of  $S$  and possibly of some auxiliary statistics on the branches  $\tau(t), \tau(r), \tau(s)$ . We will use frequently the fact that the degree of the root of  $\tau(n)$  is equal to the number of prime divisors of  $n$ , counted with multiplicities.

The sequences obtained by all of the following propositions can be found in OEIS [1] under the indicated number Axxxxxx, where also Maple programs are provided.

Abusing notation, given a statistic  $S$  on rooted trees, we shall denote  $S(\tau(n))$  by  $S(n)$ . In other words,  $S(n)$  is the value of the statistic  $S$  on the rooted tree having Matula number  $n$ .

We would like to point out that some of the statistics we consider have been considered previously in [8], [10] and [13]. We include them here in order that the paper be self-contained (some of them play an auxiliary role at other statistics) and because our approach is slightly different.

The symbols for the various statistics have been selected in the hope that they have at least some limited mnemonic value.

**Proposition 3.1.** *Let  $V$  denote "number of vertices". Then*

$$(1) \quad V(n) = \begin{cases} 1, & \text{if } n = 1; \\ 1 + V(t), & \text{if } n = p_t; \\ V(r) + V(s) - 1, & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Follows at once by examining Fig. 3.(A061775) □

**Proposition 3.2.** *Let  $E$  denote "number of edges". Then*

$$(2) \quad E(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1 + E(t), & \text{if } n = p_t; \\ E(r) + E(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Follows at once by examining Fig. 3. Obviously,  $E(n) = V(n) - 1$ . (A196050) □

**Proposition 3.3.** *Let  $H$  denote "height". Then*

$$(3) \quad H(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1 + H(t), & \text{if } n = p_t; \\ \max(H(r), H(s)), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Follows at once by examining Fig. 3. (A109082) □

**Proposition 3.4.** *Let  $LLL$  denote "level of lowest leaf". Then*

$$(4) \quad LLL(n) = \begin{cases} 1 + LLL(t), & \text{if } n = p_t; \\ \min(LLL(r), LLL(s)), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Follows at once by examining Fig. 3. The 1-vertex tree, corresponding to  $n = 1$ , has no leaves. (A184166) □

**Proposition 3.5.** *Let  $LV$  denote "number of leaves". Then*

$$(5) \quad LV(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1, & \text{if } n = 2; \\ LV(t), & \text{if } n = p_t, t \geq 2; \\ LV(r) + LV(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Follows at once by examining Fig. 3.(A109129) □

**Proposition 3.6.** *Let  $MD$  denote "maximum vertex degree". Then*

$$(6) \quad MD(n) = \begin{cases} 0, & \text{if } n = 1; \\ \max(MD(t), 1 + \Omega(t)), & \text{if } n = p_t; \\ \max(MD(r), MD(s), \Omega(r) + \Omega(s)), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Follows at once by examining Fig. 3. (A196046)  $\square$

**Proposition 3.7.** *Let  $DM$  denote "diameter". Then*

$$(7) \quad DM(n) = \begin{cases} 0, & \text{if } n = 1; \\ \max(DM(t), 1 + H(t)), & \text{if } n = p_t; \\ \max(DM(r), DM(s), H(r) + H(s)), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) If the diameter of  $\tau(t)$  is not the diameter of  $\tau(p_t)$ , then clearly the latter is given by the path going from the root of  $\tau(p_t)$  to a leaf of  $\tau(t)$  of maximum height. (ii) Similarly, if neither  $DM(r)$  nor  $DM(s)$  is the diameter of the entire tree  $\tau(rs)$ , then clearly this diameter is given by a path passing through the root. (A196058)  $\square$

**Proposition 3.8.** *Let  $PL$  denote "path length". Then*

$$(8) \quad PL(n) = \begin{cases} 0, & \text{if } n = 1; \\ PL(t) + V(t), & \text{if } n = p_t; \\ PL(r) + PL(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Each of the  $V(t)$  paths that contribute to  $PL(p_t)$  is 1 unit longer than those that make up  $PL(t)$ . The case  $n = rs$  follows at once from Fig. 3b. (A196047)  $\square$

**Proposition 3.9.** *Let  $EPL$  denote "external path length". Then*

$$(9) \quad EPL(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1, & \text{if } n = 2; \\ EPL(t) + LV(t), & \text{if } n = p_t, t \geq 2; \\ EPL(r) + EPL(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Each of the  $LV(t)$  paths that contribute to  $EPL(p_t)$  is 1 unit longer than those that make up  $EPL(t)$ . The case  $n = rs$  follows at once from Fig. 3b. (A196048)  $\square$

**Proposition 3.10.** *Let  $BV$  denote "number of branching vertices". Then, assuming, without loss of generality, that  $r$  is prime, we have*

$$(10) \quad BV(n) = \begin{cases} 0, & \text{if } n = 1; \\ BV(t), & \text{if } n = p_t, \Omega(t) \neq 2; \\ 1 + BV(t), & \text{if } n = p_t, \Omega(t) = 2; \\ BV(r) + BV(s), & \text{if } n = rs, r, s \geq 2, \Omega(s) \neq 2; \\ BV(r) + BV(s) + 1, & \text{if } n = rs, r, s \geq 2, \Omega(s) = 2. \end{cases}$$

*Proof.* (i) In the case  $n = p_t$  we gain a branching vertex only if the root of  $\tau(t)$  has degree 2 in  $\tau(t)$ . (ii) In the case  $n = rs$ , by joining the trees  $\tau(r)$  and  $\tau(s)$  at their roots, we create a new branching vertex only if the degree of the root of  $\tau(s)$  in the tree  $\tau(s)$  is 2, i.e.  $\Omega(s) = 2$ . (A196049)  $\square$

**Proposition 3.11.** *Let  $PV$  denote "number of pendant vertices". Then*

$$(11) \quad PV(n) = \begin{cases} 0, & \text{if } n = 1; \\ 2, & \text{if } n = 2; \\ 1 + LV(t), & \text{if } n = p_t, n > 2; \\ LV(r) + LV(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) If  $n = p_t$ ,  $n > 2$ , then the pendant vertices of  $\tau(n)$  consist of the root and the leaves of  $\tau(t)$ . (ii) If  $n$  is composite,  $n = rs$ , then the pendant vertices of  $\tau(n)$  consist of the leaves of  $\tau(r)$  and those of  $\tau(s)$ . (A196067)  $\square$

**Proposition 3.12.** *Let  $SP$  denote "number of sibling pairs". Then*

$$(12) \quad SP(n) = \begin{cases} 0, & \text{if } n = 1; \\ SP(t), & \text{if } n = p_t; \\ SP(r) + SP(s) + \Omega(r)\Omega(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) The trees  $\tau(p_t)$  and  $\tau(t)$  have the same number of sibling pairs. (ii) In addition to the sibling pairs of  $\tau(r)$  and those of  $\tau(s)$  we have to count the pairs  $(u, v)$ , where  $u$  ( $v$ ) is a child of the root of  $\tau(rs)$  in  $\tau(r)$  ( $\tau(s)$ ). (A196057)  $\square$

**Proposition 3.13.** *Let  $VL$  denote "visitation length". Then*

$$(13) \quad VL(n) = \begin{cases} 1, & \text{if } n = 1; \\ VL(t) + V(t) + 1, & \text{if } n = p_t; \\ VL(r) + VL(s) - 1, & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) When going from  $\tau(t)$  to  $\tau(p_t)$  the path length increases by  $V(t)$  and the number of nodes increases by 1. (ii) The term  $-1$  is due to the fact that in  $VL(r) + VL(s)$  the root is counted twice. (A196068)  $\square$

**Proposition 3.14.** *Let  $RST$  denote "number of root subtrees". Then*

$$(14) \quad RST(n) = \begin{cases} 1, & \text{if } n = 1; \\ 1 + RST(t), & \text{if } n = p_t; \\ RST(r)RST(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) The root subtrees of  $\tau(p_t)$  are the "extended" root subtrees of  $\tau(t)$  and the root of  $\tau(p_t)$  (a 1-vertex tree). (ii) Each root subtree of  $\tau(rs)$  is obtained by joining a root subtree of  $\tau(r)$  and a root subtree of  $\tau(s)$  at their roots. (A184160)  $\square$

**Proposition 3.15.** *Let  $ST$  denote "number of subtrees". Then*

$$(15) \quad ST(n) = \begin{cases} 1, & \text{if } n = 1; \\ 1 + ST(t) + RST(t), & \text{if } n = p_t; \\ ST(r) + ST(s) + (RST(r) - 1)(RST(s) - 1) - 1, & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) As shown in the proof of the previous proposition,  $1 + RST(t)$  is the number of subtrees of  $\tau(p_t)$  that do contain the root; those that do not contain the root are clearly the subtrees of  $\tau(t)$ , counted by  $ST(t)$ . (ii) The subtrees that belong entirely to  $\tau(r)$  or entirely to  $\tau(s)$  are counted by  $ST(r) + ST(s) - 1$ ; the remaining ones are obtained by joining at their roots a non-1-vertex root subtree of  $\tau(r)$  with a non-1-vertex root subtree of  $\tau(s)$ . (A84161)  $\square$

The statistic, depending on a real parameter  $\alpha$ , considered in the following proposition is needed at the statistics "2nd Zagreb index" and "general Randić index".

**Proposition 3.16.** *Let  $A_\alpha = \sum \deg(i)^\alpha$ , where the summation extends over all vertices at level 1. Then*

$$(16) \quad A_\alpha(n) = \begin{cases} 0, & \text{if } n = 1; \\ (1 + \Omega(t))^\alpha, & \text{if } n = p_t; \\ A_\alpha(r) + A_\alpha(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* In  $\tau(p_t)$ , there is only one vertex at level 1; its degree is  $1 + \Omega(t)$ . The other cases are obviously true. (A196052)  $\square$

**Proposition 3.17.** *Let  $W$  denote "Wiener index". Then*

$$(17) \quad W(n) = \begin{cases} 0, & \text{if } n = 1; \\ W(t) + PL(t) + E(t) + 1, & \text{if } n = p_t; \\ W(r) + W(s) + PL(r)E(s) + PL(s)E(r), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) In the case  $n = p_t$ , the first term takes care of the distances within  $\tau(t)$  while the last 3 terms give the sum of the distances between the vertices of  $\tau(t)$  and the root. (ii) In the case  $n = rs$  the first two terms take care of the distances within  $\tau(r)$  and those within  $\tau(s)$ . In the sum of the distances between the vertices of  $\tau(r)$  and those of  $\tau(s)$ , the sum  $PL(r)$  ( $PL(s)$ ) occurs for each nonroot vertex of  $\tau(s)$  ( $\tau(r)$ ) i.e.  $E(s)$  ( $E(r)$ ) times. (A196051) □

**Proposition 3.18.** *Let  $TW$  denote "terminal Wiener index". Then, assuming, without loss of generality, that  $r$  is prime, we have*

$$(18) \quad TW(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1, & \text{if } n = 2; \\ TW(t) + LV(t), & \text{if } n = p_t \text{ and} \\ & \Omega(t) = 1 \text{ i.e. } t \text{ is prime} \\ TW(t) + EPL(t) + LV(t), & \text{if } n = p_t \\ & \text{and } \Omega(t) > 1; \\ TW(r) - EPL(r) + TW(s) - \\ EPL(s) + EPL(r)LV(s) + EPL(s)LV(r), & \text{if } n = rs, r, s \geq 2, s \text{ is prime} \\ TW(r) - EPL(r) + TW(s) + \\ EPL(r)LV(s) + EPL(s)LV(r), & \text{if } n = rs, r, s \geq 2, s \text{ is not prime.} \end{cases}$$

*Proof.* (i) If both  $n$  and  $t$  are prime, then the distance between each leaf of  $\tau(t)$  and the root of  $\tau(t)$  has to be increased by 1 unit, explaining the term  $LV(t)$ . (ii) If  $n = p_t$  and  $\Omega(t) > 1$ , then the pendant vertices of  $\tau(p_t)$  consist of those of  $\tau(t)$  and the root of  $\tau(p_t)$ . Consequently, to  $TW(t)$  we have to add the sum of the additional distances, namely  $EPL(t) + LV(t)$ . (iii) In the case  $n = rs$ ,  $\Omega(s) = 1$ , the root of  $\tau(rs)$  is definitely not a pendant vertex. Consequently, the contribution to  $TW(rs)$  of the distances between the pendant vertices within  $\tau(r)$  is  $TW(r) - EPL(r)$ ; same explanation for  $TW(s) - EPL(s)$ . In the sum of the distances between the leaves of  $\tau(r)$  and those of  $\tau(s)$ , the sum  $EPL(r)$  occurs  $LV(s)$  times and the sum  $EPL(s)$  occurs  $LV(r)$  times. (iv) The case  $n = rs$ ,  $\Omega(s) > 1$  is similar to the previous one, except that now the root of  $\tau(s)$  is not a pendant vertex in  $\tau(s)$  and, therefore, there is no need to correct  $TW(s)$  by a subtraction. (A196055) □

**Proposition 3.19.** *Let  $Z1$  denote "first Zagreb index". Then*

$$(19) \quad Z1(n) = \begin{cases} 0, & \text{if } n = 1; \\ Z1(t) + 2 + 2\Omega(t), & \text{if } n = p_t; \\ Z1(r) + Z1(s) - \Omega(r)^2 - \Omega(s)^2 + \Omega(n)^2, & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) In the case  $n = p_t$ , to  $Z1(t)$  we add 1 (the squared degree of the root of  $\tau(p_t)$ ) and we take into account that the degree of the root of  $\tau(t)$  has increased by 1 unit (i.e. we add  $(1 + \Omega(t))^2 - \Omega(t)^2$ ). (ii) In the case  $n = rs$  we take into account that the degree of the root is  $\Omega(n)$ . (A196053) □

**Proposition 3.20.** *Let  $Z2$  denote "second Zagreb index". Then*

$$(20) \quad Z2(n) = \begin{cases} 0, & \text{if } n = 1; \\ Z2(t) + A(t) + \Omega(t) + 1, & \text{if } n = p_t; \\ Z2(r) + Z2(s) + A(r)\Omega(s) + A(s)\Omega(r), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Here we need the auxiliary statistic "sum of degrees of vertices at level 1", denoted by  $A_1$  and considered in proposition 3.16. (i) Assume that  $n = p_t$ . Let  $u$  be the root of  $\tau(t)$  and let  $ux$  be an edge in  $\tau(t)$ . This edge contributes  $deg(u)deg(x)$  to  $Z2(t)$  but to  $Z2(p_t)$  it has to contribute  $1 + deg(u)deg(x)$ . The difference is  $deg(x)$ , which, summed over all the edges  $ux$  in  $\tau(t)$ , yields  $A_1(t)$ . The sum  $1 + \Omega(t)$  is the contribution of the edge emanating from the root of  $\tau(p_t)$ . (ii) Assume that  $n = rs$ . Let  $v$  be the root of  $\tau(rs)$  and let  $vy$  be an edge in  $\tau(r)$ . This edge contributes  $deg_{\tau(r)}(y)deg_{\tau(r)}(v)$  to  $Z2(r)$  but to  $Z2(n)$  it has to contribute  $deg_{\tau(r)}(y)deg_{\tau(n)}(v)$ . The difference is  $deg_{\tau(r)}(y)(deg_{\tau(n)}(v) - deg_{\tau(r)}(v)) = deg_{\tau(r)}(y)deg_{\tau(s)}(v) = deg_{\tau(r)}(y)\Omega(s)$ . Summing over all the edges in  $\tau(r)$  that emanate from  $v$ , we obtain  $A_1(r)\Omega(s)$ . Same explanation for the edges within  $\tau(s)$ . (A196054) □

**Proposition 3.21.** *Let  $NK$  denote "Narumi-Katayama index". Then*

$$(21) \quad NK(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1, & \text{if } n = 2; \\ NK(t)(1 + \frac{1}{\Omega(t)}), & \text{if } n = p_t, t \geq 2; \\ NK(r)NK(s)(\frac{1}{\Omega(r)} + \frac{1}{\Omega(s)}), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) If  $n = p_t$ , then in the Narumi-Katayama index  $NK(t)$  of  $\tau(t)$  we replace the degree of the root of  $\tau(t)$  in  $\tau(t)$  by its degree in  $\tau(p_t)$ , i.e.  $\Omega(t)$  by  $1 + \Omega(t)$ . (ii) If  $n = rs$ , then in the product  $NK(r)NK(s)$  we replace the product  $\Omega(r)\Omega(s)$  of the root degrees in  $\tau(r)$  and  $\tau(s)$  by the root degree  $\Omega(n) = \Omega(r) + \Omega(s)$  in  $\tau(n)$ . (A196063)  $\square$

**Proposition 3.22.** *Let  $MZ1$  denote "first multiplicative Zagreb index". Then*

$$(22) \quad MZ1(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1, & \text{if } n = 2; \\ MZ1(t)(1 + \frac{1}{\Omega(t)})^2, & \text{if } n = p_t, t \geq 2; \\ MZ1(r)MZ1(s)(\frac{1}{\Omega(r)} + \frac{1}{\Omega(s)})^2, & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* Since  $MZ1(n) = NK(n)^2$ , the proof is entirely similar to that of proposition 3.21. (A196065)  $\square$

**Proposition 3.23.** *Let  $MZ2$  denote "second multiplicative Zagreb index". Then*

$$(23) \quad MZ2(n) = \begin{cases} 0, & \text{if } n = 1; \\ 1, & \text{if } n = 2; \\ \frac{MZ2(t)}{\Omega(t)^{\Omega(t)}}(1 + \Omega(t))^{1+\Omega(t)}, & \text{if } n = p_t, t \geq 2; \\ \frac{MZ2(r)MZ2(s)\Omega(n)^{\Omega(n)}}{\Omega(r)^{\Omega(r)}\Omega(s)^{\Omega(s)}}, & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* As mentioned in the introduction, the 2nd multiplicative Zagreb index can be expressed also as  $\prod \deg(i)^{\deg(i)}$ , where the product is taken over all vertices  $i$  of  $T$ . Making use of this, the proof is similar to the proof of proposition 3.21. (A196064)  $\square$

**Proposition 3.24.** *Let  $R_\alpha$  denote "general Randić index". Then*

$$(24) \quad R_\alpha(n) = \begin{cases} 0, & \text{if } n = 1; \\ R_\alpha(t) + A_\alpha(t)[(1 + \Omega(t))^\alpha - \Omega(t)^\alpha] + (1 + \Omega(t))^\alpha, & \text{if } n = p_t; \\ R_\alpha(r) + R_\alpha(s) + A_\alpha(r)[\Omega(n)^\alpha - \Omega(r)^\alpha] + A_\alpha(s)[\Omega(n)^\alpha - \Omega(s)^\alpha] & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* The reasoning is basically the same as that proposition 3.20 which is a special case of this one ( $\alpha = 1$ ). (i) In the case  $n = p_t$ , the single edge emanating from the root brings in the term  $(1 + \Omega(t))^\alpha$ , while the edges in  $\tau(t)$  have the same contribution to  $R_\alpha(n)$  as to  $R_\alpha(t)$ , except for those emanating from the root of  $\tau(t)$  because for these the degree of the lower endpoints has increased from  $\Omega(t)$  to  $1 + \Omega(t)$ . (ii) In the case  $n = rs$ , the only edges in  $\tau(n)$  which do not have the same contribution to  $R_\alpha(n)$  as to  $R_\alpha(r) + R_\alpha(s)$  are those emanating from the root. The difference in these contributions is taken care of by the last two terms in the right-hand side.  $\square$

**Proposition 3.25.** *5. Let  $PWP$  denote "partial Wiener polynomial with respect to the root". Then*

$$(25) \quad PWP(n) = \begin{cases} 0, & \text{if } n = 1; \\ x + xPWP(t), & \text{if } n = p_t; \\ PWP(r) + PWP(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* If  $n = p_t$ , then, in going from  $\tau(t)$  to  $\tau(p_t)$ , the distances from the vertices of  $\tau(t)$  to the root of  $\tau(t)$  are increased by 1 unit; this explains the factor  $x$  in  $xPWP(t)$ . The additional distance between the root of  $\tau(t)$  and that of  $\tau(p_t)$  explains the term  $x$ . The case  $n = rs$  ( $r, s \geq 2$ ) follows at once from Fig. 3b. (A196056)  $\square$

**Proposition 3.26.** *Let  $WP$  denote "Wiener polynomial". Then*

$$(26) \quad WP(n) = \begin{cases} 0, & \text{if } n = 1; \\ WP(t) + xPWP(t) + x, & \text{if } n = p_t; \\ WP(r) + WP(s) + PWP(r)PWP(s), & \text{if } n = rs, r, s \geq 2. \end{cases}$$



*Proof.* (i)  $WP(t)$  takes care of the distances within  $\tau(t)$ ,  $xPWP(t)$  takes care of the distances between the vertices of  $\tau(t)$  and the root of  $\tau(p_t)$ , and the term  $x$  is due to the distance between the root of  $\tau(t)$  and the root of  $\tau(p_t)$ . (ii) In the case  $n = rs$  ( $r, s \geq 2$ ), the three terms give the contributions of the distances within  $\tau(r)$ , within  $\tau(s)$ , and between the vertices in  $\tau(r)$  and those in  $\tau(s)$ , respectively. (A196059)  $\square$

**Proposition 3.27.** *Let DSP denote "degree sequence polynomial". Then*

$$(27) \quad DSP(n) = \begin{cases} 1, & \text{if } n = 1; \\ DSP(t) + x^{\Omega(t)}(x - 1) + x, & \text{if } n = p_t; \\ DSP(r) + DSP(s) - x^{\Omega(r)} - x^{\Omega(s)} + x^{\Omega(n)}, & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) When we go from  $\tau(t)$  to  $\tau(p_t)$ , the degree of the root of  $\tau(t)$  increases by 1 unit, explaining the term  $x^{\Omega(t)}(x - 1) = -x^{\Omega(t)} + x^{1+\Omega(t)}$ . We also have an additional vertex of degree 1 (the root of  $\tau(p_t)$ ), explaining the term  $x$ . (ii) We are replacing the terms corresponding to the root of  $\tau(r)$  and to the root  $\tau(s)$  by the term corresponding to the root of  $\tau(n)$ .(A182907)  $\square$

**Proposition 3.28.** *Let EDP denote "exit-distance polynomial". Then*

$$(28) \quad EDP(n) = \begin{cases} 1, & \text{if } n = 1; \\ EDP(t) + x^{1+LLL(t)}, & \text{if } n = p_t; \\ EDP(r) + EDP(s) - x^{\max(LLL(r), LLL(s))} & \text{if } n = rs, r, s \geq 2. \end{cases}$$

*Proof.* (i) The vertices in  $\tau(t)$  have the same exit distances in  $\tau(p_t)$  as in  $\tau(t)$  while the root of  $\tau(p_t)$  has exit distance  $1 + LLL(t)$ ; (ii) from the two terms corresponding to the roots of  $\tau(r)$  and  $\tau(s)$  we remove the term corresponding to the larger exit distance. (A184167)  $\square$

**Remark** From the propositions that consider polynomial-valued statistics we can re-obtain the results regarding several of the above considered statistics. Without going into details,

(i) denoting  $f(x) = PWP(n)$ , we have  $H(n) = \text{degree of } f(x)$ ,  $E(n) = f(1)$ , and  $PL(n) = \frac{df}{dx}|_{x=1}$ ; also, the number of vertices at level  $k$  ( $k \geq 1$ ) in the tree  $\tau(n)$  is equal to  $[x^k]f(x)$ ;

(ii) denoting  $g(x) = WP(n)$ , we have  $DM(n) = \text{degree of } g(x)$  and  $W(n) = \frac{dg}{dx}|_{x=1}$ .

(iii) denoting  $h(x) = DSP(n)$ , we have  $V(n) = h(1)$ ,  $MD(n) = \text{degree of } h(x)$ ,  $PV(n) = [x]h(x)$ , and  $BV(n) = h(1) - [x]h(x) - [x^2]h(x)$ .

Moreover, making use of some of these polynomial-valued statistics, we can consider also some statistics that are not included in the propositions given above.

(a) *The hyper-Wiener index.* It has been proved in [3] that the hyper-Wiener index of a connected graph is equal to  $g'(1) + \frac{1}{2}g''(1)$ , where  $g(x)$  is the Wiener polynomial of  $G$  (A196060).

(b) *The multiplicative Wiener index.* From the definition there follows easily that the multiplicative Wiener index of a graph  $G$  is equal to  $\prod_{k>1} k^{\delta(G,k)}$ , where  $\delta(G,k)$  is the number of vertex pairs in  $G$  that are at distance  $k$  (see Eq. (5a) in [14] (A196061)).

(c) *The Wiener polarity index.* This is the coefficient of  $x^3$  in the Wiener polynomial of the graph (A184156). Obviously, taking other coefficients in the Wiener polynomial, one obtains generalized Wiener polarity indices (see [19]).

(d) In [20] one introduces a family of statistics on a connected graph  $G$ , based on the graph distances. In particular,  $\text{Sum}E(1, G)$  ( $\text{Sum}O(1, G)$ ) denotes the sum of the even (odd) distances between unordered pairs of vertices of  $G$ . Clearly, it can be obtained by evaluating at  $x = 1$  the derivative of the even (odd) part of the Wiener polynomial of  $G$  (A184157, A184158).

(e) Denoting  $m(x) = EDP(n)$ , it is easy to see that (i) the sum of the exit distances of all vertices of  $\tau(n)$  is equal to  $\frac{dm}{dx}|_{x=1}$  (A184168), (ii) the maximum exit distance over the vertices of  $\tau(n)$  is the degree of the polynomial  $m(x)$  (A184169), and (iii) the number of vertices of  $\tau(n)$  having maximum exit distance is the coefficient of the highest power of  $x$  in  $m(x)$  (A184170).

## REFERENCES

- [1] *The On-Line Encyclopedia of Integer Sequences*. <http://oeis.org>, 2010.
- [2] B. Bollobás and P. Erdős. Graphs of extremal weights. *Ars Combin.*, 50:225–233, 1998.
- [3] G. G. Cash. Relationship between the Hosoya polynomial and the hyper-Wiener index. *Appl. Math. Lett.*, 15(7):893–895, 2002.
- [4] É. Czabarka, L. Székely, and S. Wagner. The inverse problem for certain tree parameters. *Discrete Appl. Math.*, 157(15):3314–3319, 2009.
- [5] K. Ch. Das and I. Gutman. Some properties of the second Zagreb index. *MATCH Commun. Math. Comput. Chem.*, 52:103–112, 2004.
- [6] A. A. Dobrynin, R. Entringer, and I. Gutman. Wiener index of trees: theory and applications. *Acta Appl. Math.*, 66(3):211–249, 2001.
- [7] T. Došlić. Vertex-weighted Wiener polynomials for composite graphs. *Ars Math. Contemp.*, 1(1):66–80, 2008.
- [8] S.B. Elk and I. Gutman. Further properties derivable from the Matula number of an alkane. *J. Chem. Inf. Comput. Sci.*, 34:331–334, 1994.
- [9] I. Gutman. Relation between hyper-Wiener and Wiener index. *Chemical Physics Letters*, 364:352–356, 2002.
- [10] I. Gutman. Multiplicative Zagreb indices of trees. *Bull. Soc. Math. Banja Luka*, 18:17–23, 2011.
- [11] I. Gutman and K. Ch. Das. The first Zagreb index 30 years after. *MATCH Commun. Math. Comput. Chem.*, 50:83–92, 2004.
- [12] I. Gutman, B. Furtula, and M. Petrović. Terminal Wiener index. *J. Math. Chem.*, 46(2):522–531, 2009.
- [13] I. Gutman and A. Ivić. On Matula numbers. *Discrete Math.*, 150(1-3):131–142, 1996.
- [14] I. Gutman, W. Linert, I. Lukovits, and Ž Tomović. The multiplicative version of the Wiener index. *J. Chem. Inf. Comput. Sci.*, 40:113–116, 2000.
- [15] I. Gutman and M. Randić. Algebraic characterization of skeletal branching. *Chem. Phys. Letters*, 47:15–19, 1977.
- [16] I. Gutman and N. Trinajstić. Graph theory and molecular orbitals. Total  $\pi$ -electron energy of alternant hydrocarbons. *Chemical Physics Letters*, 17 (4):535–538, 1972.
- [17] I. Gutman and Y. Yeong Nan. Deducing properties of trees from their Matula numbers. *Publ. Inst. Math. (Beograd) (N.S.)*, 53(67):17–22, 1993.
- [18] H. Hosoya. On some counting polynomials in chemistry. *Discrete Appl. Math.*, 19(1-3):239–257, 1988.
- [19] A. Ilić and M. Ilić. Generalizations of Wiener polarity index and terminal Wiener index. *arXiv:1106.2986v1 [math.CO]*, 15 June 2011.
- [20] O. Ivanciuc, T. Ivanciuc, D.J. Klein, W.A. Seitz, and T. Balaban. Wiener index extension by counting even/odd graph distances. *J. Chem. Inf. Comput. Sci.*, 41:536–549, 2001.
- [21] M. Keijzer and J. Foster. Crossover bias in genetic programming. *Proceedings of the 10th European Conference on Genetic Programming*, LNCS 4445:33–44, 2007.
- [22] D.J. Klein, I. Lukovits, and I. Gutman. On the definition of the hyper-Wiener index for cycle-containing structures. *J. Chem. Inf. Comput. Sci.*, 35:50–52, 1995.
- [23] D.W. Matula. A natural rooted tree enumeration by prime factorization. *SIAM Review*, 10:273, 1968.
- [24] H. Narumi and M. Katayama. Simple topological index. A newly devised index characterizing the topological nature of structural isomers of saturated hydrocarbons. *Mem. Fac. Engin., Hokkaido Univ.*, 16:209–214, 1984.
- [25] S. Nikolić, G. Kovačević, A. Milicevic, and N. Trinajstić. The Zagreb indices 30 years after. *Croatica Chemica Acta*, 76:113–124, 2003.
- [26] M. Randić. Novel molecular descriptor for structure-property studies. *Chemical Physics Letters*, 211:478–483, 1993.
- [27] B.E. Sagan, Y-N. Yeh, and P. Zhang. The Wiener polynomial of a graph. *International Journal of Quantum Chemistry*, 60:959–969, 1996.
- [28] R. Sedgewick and Ph. Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, 1996.
- [29] L. A. Székely, H. Wang, and W. Taoyang. The sum of the distances between the leaves of a tree and the ‘semi-regular’ property. *Discrete Math.*, 311(13):1197–1203, 2011.
- [30] N. Trinajstić. *Chemical Graph Theory*, volume II. CRC Press, Boca Raton, 1983.
- [31] H. Wiener. Structural determination of paraffin boiling points. *Journal of American Chemistry Society*, 69:17–20, 1947.