Counting all parity realizable trees

Stephan Wagner
Department of Mathematical Sciences
Mathematics Division
Stellenbosch University
Private Bag X1, Matieland 7602
South Africa
swagner@sun.ac.za

January 30, 2007

Abstract

The parity dimension of a graph G is defined as the dimension of the null space of its closed neighborhood matrix N. A graph with parity dimension 0 is called *all parity realizable* (APR). In this paper, a simple recursive procedure for calculating the parity dimension of a tree is given, which is more apt to be used in the context of enumeration than the graph-theoretical characterizations due to Amin, Slater and Zhang. Applying the recursive relation, we find asymptotic formulas for the number of APR trees and for the average parity dimension of a tree.

1 Introduction

The open neighborhood of a vertex v in a graph G is the set $N(v) = \{u \in V(G) \mid uv \in E(G)\}$; the closed neighborhood is given by $N[v] = N(v) \cup \{v\}$. The classical domination problem asks for a set S (typically, of minimal cardinality or weight) with the property that $|N[v] \cap S| \geq 1$ for all vertices v. It is a natural generalization to require $|N[v] \cap S| \geq k$ for some k, which is also known as the *k-tuple domination problem* [11]. Even more generally, one may ask for a set S such that $|N[v] \cap S| \in R_v$ for all v and given sets R_v of integers. Clearly, classical domination corresponds to the special case $R_v = \{1, 2, 3, \ldots\}$, and k-tuple domination corresponds to $R_v = \{k, k+1, \ldots\}$. These and other variants, such as $R_v = \{1\}$, are treated in the book of Haynes, Hedetniemi and Slater [14].

One motivation to consider domination with parity restrictions is the following remarkable result of Sutner [21]:

Theorem 1 (Sutner [21]) *For every graph G, there is a set $S \subseteq V(G)$ such that $|N[v] \cap S|$ is odd for every $v \in V(G)$.*

This means that the domination problem for $R_v = \{1, 3, 5, \ldots\}$ is always solvable. Thus it is pretty natural to consider a general parity assignment problem, where each R_v is either $\{1, 3, 5, \ldots\}$ or $\{0, 2, 4, \ldots\}$. Clearly, $S = \emptyset$ solves the problem when $R_v = \{0, 2, 4, \ldots\}$ for every v, but this might not be the only solution.

Generally, we call a set S a *D-parity set* if $|N[v] \cap S|$ is odd precisely for $v \in D$. The parity domination problem has been investigated in a series of papers by Amin, Slater and others. [2] mainly deals with algorithmic questions, whereas papers [3, 1, 4] investigate the problem from a graph-theoretical point of view.

Note that the problem can be reformulated in the world of matrix algebra. Let A be the open neighborhood matrix or adjacency matrix of G; if $V(G) = \{v_1, \ldots, v_n\}$, then A is the symmetric

$n \times n$-matrix whose entries are $a_{ij} = 1$ if $v_i v_j \in E(G)$ and $a_{ij} = 0$ otherwise. Furthermore, let $N = I + A$ (I denoting the identity matrix) be the closed neighborhood matrix. Then a D-parity set S corresponds to a vector s over the field $\mathbb{F}_2$ such that the entries of $t = N \cdot s$ are $t_i = 1$ if $v_i \in D$ and $t_i = 0$ otherwise.

From elementary linear algebra, we obtain that a D-parity set exists for all D if and only if N has full rank over $\mathbb{F}_2$, or equivalently, if the dimension of the null space of N is 0. This dimension will be called the *parity dimension* $\mathrm{PD}(G)$ [1, 4]. $2^{\mathrm{PD}(G)}$, the cardinality of the null space, is precisely the number of $\emptyset$-parity sets. Note also that for every $D \subseteq V(G)$, the D-parity sets, if there are any, form an affine space of dimension $\mathrm{PD}(G)$.

Similar notions are known for the adjacency matrix A. The rank of A is also known as the *rank* of the graph G, and it is related to many other parameters of a graph. A very interesting theorem due to Bevis, Domke and Miller [6] states that the rank of a tree equals twice the size of a maximal matching.

In this paper, we study the parity dimension of trees in detail. First of all, a simple algorithm for calculating the parity dimension of a rooted tree is given. This algorithm yields a recursive characterization of APR trees which is more apt to the treatment of enumeration problems than the graph-theoretical characterizations due to Amin, Slater and Zhang:

Theorem 2 (Amin and Slater [3]) *A tree T is an APR tree if and only if $T = K_1$ or if T can be obtained by one of the following two operations:*

- *take two APR trees T_1, T_2 and vertices $v_i \in V(T_i)$ such that v_1 is not a member of the unique $\{v_1\}$-parity set in T_1, and join T_1 and T_2 by an edge between v_1 and v_2,*
- *take $2k$ APR trees $T_1, \ldots, T_{2k}$ and vertices $v_i \in V(T_i)$ such that v_i is a member of the unique $\{v_i\}$-parity set in T_i for all i, and join the trees $T_1, \ldots, T_{2k}$ by $2k$ edges between the v_i and a new vertex w.*

Theorem 3 (Amin, Slater and Zhang [4]) *A tree T is an APR tree if and only if $T = K_1$ or if T can be obtained from an APR tree F by one of the following three operations:*

- *add two vertices and two edges connecting the new vertices to the same vertex in $V(F)$,*
- *add a path P_3 and an edge connecting one of its endpoints to a vertex in $V(F)$,*
- *add a path P_4 and an edge connecting one of its center vertices to a vertex in $V(F)$.*

The new characterization enables us to enumerate APR trees by means of a generating function approach. A priori, it is not clear at all whether there are even infinitely many APR trees. However, the following theorem of Amin and Slater immediately shows that this is the case:

Theorem 4 (Amin and Slater [3]) *If a tree T has exactly one vertex of even degree, then T is APR.*

Finally, we are going to determine the average parity dimension of a tree on n vertices asymptotically. A similar result has been provided by Amin, Clark and Slater in [1] for random graphs. There, the authors also determine the parity dimension for certain special trees, including the following interesting result:

Theorem 5 (Amin, Clark and Slater [1]) *If every vertex of a tree T has odd degree, then $\mathrm{PD}(T) = 1$.*

2 An algorithm for rooted trees

We consider a tree T rooted at some vertex r. We want to calculate the parity dimension of T from the parity dimensions of the branches $T_1, T_2, \ldots, T_k$, which are rooted at the neighbors of r. Unfortunately, there is no simple way to do so, and so we have to define three different classes of rooted trees:

- A rooted tree T is said to be of type A if there exists an $\emptyset$-parity set S such that $r \in S$ and if there is no $\{r\}$-parity set.
- A rooted tree T is said to be of type B if there is no $\emptyset$-parity set S such that $r \in S$ and if there is an $\{r\}$-parity set such that $r \in S$, but no $\{r\}$-parity set such that $r \notin S$.
- A rooted tree T is said to be of type C if there is no $\emptyset$-parity set S such that $r \in S$ and if there is an $\{r\}$-parity set such that $r \notin S$, but no $\{r\}$-parity set such that $r \in S$.

Of course, it is not obvious at all why these three types should cover all possible rooted trees. However, this will be a consequence of the following two lemmas. Note, for the moment, that the rooted tree consisting of a single vertex is of type B.

Lemma 6 *Let $T_1, T_2, \ldots, T_k$ be the branches of a tree T that is rooted at r. Furthermore, let $r_1, \ldots, r_k$ be the roots of $T_1, \ldots, T_k$ (the neighbors of r in T). If one of the branches T_i is of type A, then T is of type C, and*

$$\mathrm{PD}(T) = \sum_{i=1}^{k} \mathrm{PD}(T_i) - 1.$$

Proof: Let S be an $\emptyset$-parity set and let $S_i = S \cap V(T_i)$ be its restriction to T_i. Then S_i is either an $\emptyset$-parity set or an $\{r_i\}$-parity set on T_i.

Let T_j be a branch of type A. Then S_j has to be an $\emptyset$-parity set on T_i, but this implies that $r \notin S$. Therefore, all other S_i's have to be $\emptyset$-parity sets on the respective T_i's as well. Conversely, the union of an arbitrary collection of $\emptyset$-parity sets on each of the T_i yields an $\emptyset$-parity or $\{r\}$-parity set on T. Thus the total number of $\emptyset$-parity and $\{r\}$-parity sets together is

$$2^{\sum_{i=1}^{k} \mathrm{PD}(T_i)}.$$

Finally, note that there are $\{r\}$-parity sets (but, by the same argument, none that contain r). Choose any $\emptyset$-parity set containing r_j on an A-type branch T_j. This readily gives us an $\{r\}$-parity set. Altogether, we see that T has to be of type C and that the number of $\emptyset$-parity (or $\{r\}$-parity) sets is $\frac{1}{2} \cdot 2^{\sum_{i=1}^{k} \mathrm{PD}(T_i)}$, so that

$$\mathrm{PD}(T) = \sum_{i=1}^{k} \mathrm{PD}(T_i) - 1.$$

$\square$

Lemma 7 *With the same notation as in Lemma 6, suppose that all branches are either of type B or type C. If the number of B-type branches is even, then T is of type B and*

$$\mathrm{PD}(T) = \sum_{i=1}^{k} \mathrm{PD}(T_i).$$

On the other hand, if the number of B-type branches is odd, then T is of type A and

$$\mathrm{PD}(T) = \sum_{i=1}^{k} \mathrm{PD}(T_i) + 1.$$

Proof: Let S be an $\emptyset$- or $\{r\}$-parity set and take S_i as in the proof of Lemma 6. If $r \notin S$, then all S_i have to be $\emptyset$-parity sets, and there are precisely $2^{\sum_{i=1}^{k} \mathrm{PD}(T_i)}$ possibilities for that. Since all T_i are of type B or C, we have $r_i \notin S_i$ for all i, and thus S is an $\emptyset$-parity set.

On the other hand, if $r \in S$, then all S_i have to be $\{r_i\}$-parity sets, and $|N(r) \cap S|$ is the number of type-B branches. Therefore, in this case, every such S is an $\{r\}$-parity set if the number of type-B branches is even, implying that T has to be of type B and that $\mathrm{PD}(T) = \sum_{i=1}^{k} \mathrm{PD}(T_i)$.

Similarly, if the number of B-type branches is odd, every such S is an $\emptyset$-parity set (so that there are no $\{r\}$-parity sets). Then, T has to be of type A, and $\mathrm{PD}(T) = \sum_{i=1}^{k} \mathrm{PD}(T_i) + 1$. $\square$

Corollary 8 *Every rooted tree is either of type A, or B, or C.*

Proof: This follows by induction from the preceding lemmas. $\square$

Now, we have an easy procedure to determine the parity dimension of a rooted tree in a bottom-to-top manner. This algorithm is depicted in the following figure (letters denote the type, numbers the parity dimension of the branch rooted at each vertex):

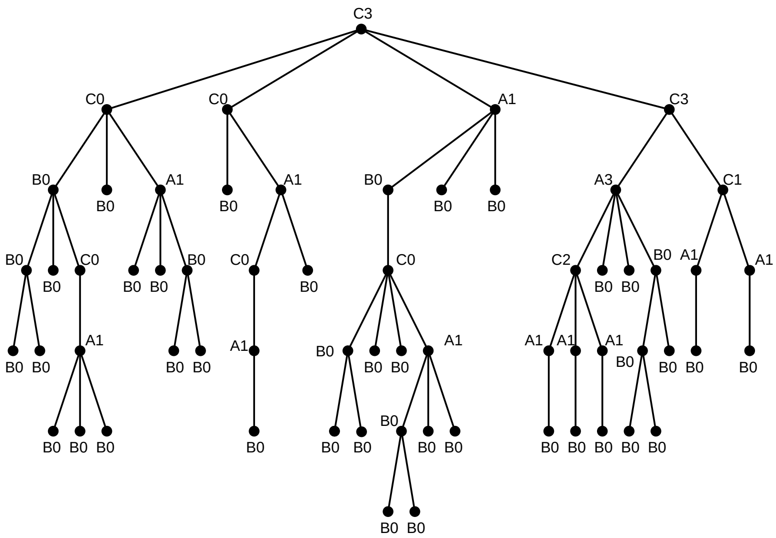

Figure 1: An example for determining the type and parity dimension of a rooted tree.

For the enumeration problems of the following section, we need yet another theorem, which provides information on the parity dimension of a tree that results from joining two rooted trees by an edge between the roots:

Theorem 9 *Let T_1, T_2 be two rooted trees, and let T be the tree that results from joining T_1, T_2 by an edge between their roots r_1, r_2. Then we have*

$$\mathrm{PD}(T) = \begin{cases} \mathrm{PD}(T_1) + \mathrm{PD}(T_2) & \text{if either } T_1 \text{ or } T_2 \text{ is of type C,} \\ \mathrm{PD}(T_1) + \mathrm{PD}(T_2) - 2 & \text{if } T_1 \text{ and } T_2 \text{ are of type A,} \\ \mathrm{PD}(T_1) + \mathrm{PD}(T_2) + 1 & \text{if } T_1 \text{ and } T_2 \text{ are of type B,} \\ \mathrm{PD}(T_1) + \mathrm{PD}(T_2) - 1 & \text{if } T_1 \text{ is of type A, } T_2 \text{ of type B or vice versa.} \end{cases}$$

Proof: As an example, we consider the case when T_1 is of type A and T_2 is of type C, the other cases being similar. An $\emptyset$-parity set S in T induces an $\emptyset$- or $\{r_i\}$-parity set S_i in T_i $(i = 1, 2)$. Since there are no $\{r_1\}$-parity sets in T_1, S_1 is a $\emptyset$-parity set, and $r_2 \notin S$. However, S_2 may be an $\emptyset$- or $\{r_2\}$-parity set ($2^{\mathrm{PD}(T_2)}$ possibilities in each case), and $r_1 \notin S_1$ resp. $r_1 \in S_1$ ($2^{\mathrm{PD}(T_1)-1}$ possibilities in each case). Therefore, there are

$$2 \cdot 2^{\mathrm{PD}(T_1)-1} \cdot 2^{\mathrm{PD}(T_2)} = 2^{\mathrm{PD}(T_1)+\mathrm{PD}(T_2)}$$

$\emptyset$-parity sets, which means that $\mathrm{PD}(T) = \mathrm{PD}(T_1) + \mathrm{PD}(T_2)$. □

Note that Theorems 2, 3, 4 and 5 can be obtained as corollaries of Lemmas 6, 7 and Theorem 9. Moreover, one can prove the following result of Amin, Clark and Slater [1] by means of our recursive procedure:

Corollary 10 *For a tree T on $n \geq 5$ vertices, we have $\mathrm{PD}(T) \leq \lfloor \frac{n-3}{2} \rfloor$. Furthermore, for every $0 \leq k \leq \lfloor \frac{n-3}{2} \rfloor$, there is a tree T on n vertices such that $\mathrm{PD}(T) = k$.*

Proof: Note that we have $\mathrm{PD}(T) \leq \frac{n}{2}$ for trees with an arbitrary number of vertices $n \leq 4$, and that equality only holds for the type-A tree with two vertices. Furthermore, $\mathrm{PD}(T) \leq \frac{n-3}{2}$ is valid for all trees on $n \leq 4$ vertices except the type-B tree consisting of a single vertex and the type-A trees on 2 or 4 vertices. Now, we prove by means of induction that $\mathrm{PD}(T) \leq \frac{n-3}{2}$ holds for all trees on $n \geq 5$ vertices, and that the stronger inequality $\mathrm{PD}(T) \leq \frac{n-4}{2}$ holds for type-B trees. It is easy to check that this is true for $n = 5$. Now, we consider three cases:

- Let T be a tree on $n \geq 6$ vertices of type A. If all branches are single vertices, we have $\mathrm{PD}(T) = 1 \leq \frac{n-3}{2}$. If there is only one branch T_1, it has to be of type B, so that we have

$$\mathrm{PD}(T) = \mathrm{PD}(T_1) + 1 \leq \frac{|T_1| - 4}{2} + 1 = \frac{n-5}{2} + 1 = \frac{n-3}{2}.$$

  Finally, let there be at least two branches (of type B or C), at least one of which has ≥ 3 vertices. This branch T_1 satisfies $\mathrm{PD}(T_1) \leq \frac{|T_1|-3}{2}$, and all others satisfy $\mathrm{PD}(T_i) \leq \frac{|T_i|-1}{2}$. Hence,

$$\mathrm{PD}(T) = \sum_i \mathrm{PD}(T_i) + 1 \leq \frac{\sum_i |T_i| - 4}{2} + 1 = \frac{n-5}{2} + 1 = \frac{n-3}{2}.$$

- If T is of type B, then there is either one branch T_1 (of type B or C) with ≥ 3 vertices, which satisfies $\mathrm{PD}(T_1) \leq \frac{|T_1|-3}{2}$, yielding

$$\mathrm{PD}(T) = \sum_i \mathrm{PD}(T_i) \leq \frac{\sum_i |T_i| - 3}{2} = \frac{n-4}{2},$$

  or all branches are single vertices, yielding $\mathrm{PD}(T) = 0 \leq \frac{n-4}{2}$.

- Finally, let T be of type C. Then all branches T_i satisfy $\mathrm{PD}(T_i) \leq \frac{|T_i|}{2}$, so that we have

$$\mathrm{PD}(T) = \sum_i \mathrm{PD}(T_i) - 1 \leq \frac{\sum_i |T_i|}{2} - 1 = \frac{n-1}{2} - 1 = \frac{n-3}{2}.$$

This finishes the induction, so that the inequality is proved. Finally, note that it is easy to construct a tree on $n \geq 5$ vertices with $\mathrm{PD}(T) = k$ for $0 \leq k \leq \lfloor \frac{n-3}{2} \rfloor$: we just have to consider a tree with $n - 2k - 3$ single-vertex branches and $k + 1$ two-vertex branches. This tree is of type C, and Theorem 6 shows that $\mathrm{PD}(T) = k$.

3 The number of APR trees

Note that, by definition, a type-A tree cannot be an APR tree. Therefore, every APR tree is either of type B or type C. Furthermore, we see from Lemmas 6 and 7 that

- The branches of an APR tree of type B are APR trees of type B and C, and the number of type-B branches is even.
- Exactly one of the branches of an APR tree of type C is of type A with parity dimension 1. All other branches are APR trees of type B or C.
- The branches of a type-A tree with parity dimension 1 are APR trees of type B and C, and the number of type-B branches is odd.

3.1 Rooted ordered trees

The recursive descriptions can be easily translated to the world of generating functions by standard methods (see [9, 22]). For instance, we determine the functional equations in the case of rooted ordered trees; however, analogous equations hold for all simply generated families of trees (in the sense of Meir and Moon [16]):

$$\begin{aligned}
A(x) &= \frac{x}{2}\left(\frac{1}{1-B(x)-C(x)} - \frac{1}{1+B(x)-C(x)}\right),\\
B(x) &= \frac{x}{2}\left(\frac{1}{1-B(x)-C(x)} + \frac{1}{1+B(x)-C(x)}\right),\\
C(x) &= \frac{xA(x)}{(1-B(x)-C(x))^2},
\end{aligned}$$

where $A(x)$, $B(x)$, $C(x)$ denote the generating functions for type-A trees of parity dimension 1, APR trees of type B and APR trees of type C respectively. Using the method of Gröbner bases [10] and the power of a computer algebra system such as Mathematica, one can reduce these to a single equation for the generating function $R(x) = B(x) + C(x)$ of all rooted ordered APR trees:

$$\begin{aligned}
&R(x)^8 - 5R(x)^7 + (x+9)R(x)^6 - (4x+5)R(x)^5 + (4x^2+5x-5)R(x)^4 - (13x^2-9)R(x)^3\\
&+ (x^3+15x^2-5x-5)R(x)^2 - (2x^3+7x^2-4x-1)R(x) + (x^4+x^3+x^2-x) = 0.
\end{aligned}$$

Now, the coefficients of $R(x)$ can be easily computed:

$$R(x) = x + 2x^3 + 3x^4 + 10x^5 + 25x^6 + 86x^7 + 252x^8 + 842x^9 + 2706x^{10} + \dots$$

A routine singularity analysis (cf. Flajolet and Odlyzko [8]) yields the asymptotic number of rooted ordered APR trees: it is known (see [15]) that every singularity x_0 of an algebraic function $f(x)$, defined by

$$F(x, f(x)) = \sum_{j=0}^{k} p_j(x) f(x)^j = 0$$

for some polynomials p_j is either a zero of p_k or a zero of the system

$$F(x,y) = 0, \;\; F_y(x,y) = 0. \tag{1}$$

In our case, p_k has no zeros, and solving the latter system yields the dominant singularity $x_0 = 0.259371$ (an algebraic number of degree 6). Now we apply a well-known theorem of Bender and Canfield [5, 7], which states that the asymptotic behavior of the coefficients a_n of $f(x)$ is

$$a_n \sim \sqrt{\frac{x_0 F_x(x_0,y_0)}{2\pi F_{yy}(x_0,y_0)}}\, n^{-3/2} x_0^{-n}$$

under the conditions stated above, where x_0, y_0 are solutions of the system (1) and x_0 is the dominant singularity of f. Then we obtain that the number a_n of rooted ordered APR trees with n vertices is asymptotically given by

$$a_n \sim 0.116269 \cdot n^{-3/2} \cdot 3.855482^n.$$

Since the number of rooted ordered trees is known to be

$$t_n = \frac{1}{n}\binom{2n-2}{n-1} \sim \frac{1}{\sqrt{\pi}} n^{-3/2} 4^{n-1},$$

this means that the ratio of APR trees among all rooted ordered trees is asymptotically

$$0.824328 \cdot 0.963870^n,$$

which is rather large, compared to the number of perfect matching trees (cf. [17]) for instance.

3.2 Pólya trees

Things are a little more involved if one is interested in rooted unordered trees (also known as Pólya trees in view of Pólya's groundbreaking work [19]) and free (unrooted) trees. For the former, the recursive structure translates to the following functional equation (where the notation is taken analogously to the previous chapter):

$$\begin{aligned}
A(x) &= x \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} C(x^k) + \sum_{k=1}^{\infty} \frac{1}{2k} B(x^{2k})\right) \sinh\left(\sum_{k=1}^{\infty} \frac{1}{2k-1} B(x^{2k-1})\right), \\
B(x) &= x \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} C(x^k) + \sum_{k=1}^{\infty} \frac{1}{2k} B(x^{2k})\right) \cosh\left(\sum_{k=1}^{\infty} \frac{1}{2k-1} B(x^{2k-1})\right), \\
C(x) &= x A(x) \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} \left(B(x^k) + C(x^k)\right)\right).
\end{aligned}$$

Again, the generating function for the total number of rooted APR trees is given by

$$R(x) = B(x) + C(x) = x + 2x^3 + 2x^4 + 6x^5 + 12x^6 + 30x^7 + 64x^8 + 167x^9 + 390x^{10} + \dots.$$

Now, one can essentially follow the lines of Harary, Robinson and Schwenk [13]. Note first that

$$C(x) = A(x)(A(x) + B(x)),$$

which simplifies the equations. From what we already know, it is clear that A, B, C have radius of convergence $\rho < 1$. Furthermore, we see that

$$\log\left(\frac{2B(x)}{x}\right) \geq \log\left(2\cosh(B(x))\right) \geq B(x)$$

or

$$\frac{2B(x)/x}{\log(2B(x)/x)} \leq \frac{2}{x},$$

which shows that $B(x)$ is bounded (and monotonous) on the interval $(0, \rho)$, so the limit $b_0 = \lim_{x\to\rho} B(x)$ exists. Since $A(x) \leq B(x)$, the same holds for $a_0 = \lim_{x\to\rho} A(x)$ as well. We write $L_1(x)$ and $L_2(x)$ for the functions

$$\sum_{k=2}^{\infty} \frac{1}{k} C(x^k) + \sum_{k=1}^{\infty} \frac{1}{2k} B(x^{2k}) \text{ resp. } \sum_{k=2}^{\infty} \frac{1}{2k-1} B(x^{2k-1}),$$

which are analytic for $|x| < \rho^{1/2}$. The Jacobian determinant of the system

$$\begin{aligned}
F_1(A(x), B(x), x) &= x \exp\left(C(x) + L_1(x)\right) \sinh\left(B(x) + L_2(x)\right) - A(x) \\
&= x \exp\left(A(x)^2 + A(x)B(x) + L_1(x)\right) \sinh\left(B(x) + L_2(x)\right) - A(x) \\
F_2(A(x), B(x), x) &= x \exp\left(C(x) + L_1(x)\right) \cosh\left(B(x) + L_2(x)\right) - B(x) \\
&= x \exp\left(A(x)^2 + A(x)B(x) + L_1(x)\right) \cosh\left(B(x) + L_2(x)\right) - B(x)
\end{aligned}$$

has to vanish at a singularity of $A(x)$ and $B(x)$. Otherwise, by the implicit function theorem, they would have a unique analytic continuation in a certain neighborhood. Now, the Jacobian matrix of $F_1(y_1, y_2, x)$ and $F_2(y_1, y_2, x)$ is

$$\begin{aligned}\frac{\partial F}{\partial y} &= \begin{pmatrix}(F_1+y_1)(2y_1+y_2)-1 & (F_1+y_1)y_1+(F_2+y_2)\\ (F_2+y_2)(2y_1+y_2) & (F_2+y_2)y_1+(F_1+y_1)-1\end{pmatrix}\\ &= \begin{pmatrix}y_1(2y_1+y_2)-1 & y_1^2+y_2\\ y_2(2y_1+y_2) & y_1y_2+y_1-1\end{pmatrix},\end{aligned}$$

since both F_1 and F_2 must vanish. The determinant is thus given by

$$\left|\frac{\partial F}{\partial y}\right| = 1 - y_1 - 2y_1(y_1+y_2) - (y_1+y_2)(y_2-y_1)(2y_1+y_2),$$

which means that

$$a_0 + 2a_0(a_0+b_0) + (a_0+b_0)(b_0-a_0)(2a_0+b_0) = 1.$$

Note also that

$$\begin{aligned}&A(x) + 2A(x)(A(x)+B(x)) + (A(x)+B(x))(B(x)-A(x))(2A(x)+B(x))\\ &\quad = A(x) + 2A(x)(A(x)+B(x)) + x^2\exp\left(2C(x)+2L_1(x)\right)\left(2A(x)+B(x)\right)\end{aligned}$$

has only positive coefficients, so that ρ must be the only value of x such that $|x| = \rho$ and the Jacobian determinant vanishes. Therefore, ρ is the only singularity on the circle of convergence, and we may make use of the following well-known theorem (cf. [12]):

Theorem 11 *Let $F(x,y)$ be analytic in each variable separately in some neighborhood of (x_0, y_0) and suppose that the following conditions are satisfied:*

1. $F(x_0, y_0) = 0$,
2. $y = f(x)$ *is analytic in* $|x| < |x_0|$ *and* x_0 *is the unique singularity on the circle of convergence,*
3. *if* $f(x) = \sum_{n=0}^{\infty} f_n x^n$ *is the expansion of* f *at the origin, then* $y_0 = \sum_{n=0}^{\infty} f_n x_0^n$,
4. $F(x, f(x)) = 0$ *for* $|x| < |x_0|$,
5. $\frac{\partial F}{\partial y}(x_0, y_0) = 0$,
6. $\frac{\partial^2 F}{\partial y^2}(x_0, y_0) \neq 0$.

Then $f(x)$ *may be expanded about* x_0*:*

$$f(x) = f(x_0) + \sum_{k=1}^{\infty} a_k (x_0 - x)^{k/2},$$

and if $a_1 \neq 0$,

$$f_n \sim \frac{-a_1}{2\sqrt{\pi}} x_0^{-n+1/2} n^{-3/2}.$$

If $a_1 = 0$ *and* $a_3 \neq 0$,

$$f_n \sim \frac{3a_3}{4\sqrt{\pi}} x_0^{-n+3/2} n^{-5/2}.$$

Note that the conditions of the theorem are satisfied (for instance) for $f(x) = B(x)$ and

$$F(x,y) = x\exp\left(y^2\tanh(y+L_2(x))(1+\tanh(y+L_2(x))) + L_1(x)\right)\cosh(y+L_2(x)) - y,$$

which can be deduced from the fact that $A(x) = B(x)\tanh(B(x) + L_2(x))$. So we may write

$$A(x) = a_0 + a_1\sqrt{\rho - x} +, \; B(x) = b_0 + b_1\sqrt{\rho - x} + \ldots, C(x) = c_0 + c_1\sqrt{\rho - x} + \ldots.$$

Here, a_0, b_0, c_0 and ρ are given by

$$\begin{aligned} a_0 &= \rho e^{c_0 + L_1(\rho)} \sinh(b_0 + L_2(\rho)), \\ b_0 &= \rho e^{c_0 + L_1(\rho)} \cosh(b_0 + L_2(\rho)), \\ c_0 &= a_0(a_0 + b_0), \\ 1 &= a_0 + 2a_0(a_0 + b_0) + (a_0 + b_0)(b_0 - a_0)(2a_0 + b_0). \end{aligned} \tag{2}$$

For a numerical solution, one calculates the coefficients of L_1 and L_2 up to some point and estimates the error. In this case, the numerical values are given by $a_0 = 0.281894$, $b_0 = 0.550337$, $c_0 = 0.234601$ and $\rho = 0.349484$.
Moreover, a_1, b_1, c_1 can be determined by means of another routine calculation, making use of the fact that

$$\begin{aligned} s_1 &= \lim_{x \to \rho} A'(x)(1 - A(x) - 2A(x)(A(x) + B(x)) - (A(x) + B(x))(B(x) - A(x))(2A(x) + B(x)) \\ &= \frac{a_1}{2}\left(a_1 + 4a_0a_1 + 2a_1b_0 + 2a_0b_1 - 6a_0^2a_1 - 2a_0a_1b_0 - a_0^2b_1 + 2a_1b_0^2 + 4a_0b_0b_1 + 3b_0^2b_1\right) \end{aligned}$$

and

$$\begin{aligned} s_2 &= \lim_{x \to \rho} B'(x)(1 - A(x) - 2A(x)(A(x) + B(x)) - (A(x) + B(x))(B(x) - A(x))(2A(x) + B(x)) \\ &= \frac{b_1}{2}\left(a_1 + 4a_0a_1 + 2a_1b_0 + 2a_0b_1 - 6a_0^2a_1 - 2a_0a_1b_0 - a_0^2b_1 + 2a_1b_0^2 + 4a_0b_0b_1 + 3b_0^2b_1\right) \end{aligned}$$

can be expressed in terms of a_0, a_1, ρ and the functions L_1, L_2 by differentiating the functional equations for $A(x), B(x), C(x)$ and solving the resulting system of linear equations:

$$\begin{aligned} A'(x) &= \frac{A(x)}{x} + (C'(x) + L_1'(x))A(x) + (B'(x) + L_2'(x))B(x), \\ B'(x) &= \frac{B(x)}{x} + (C'(x) + L_1'(x))B(x) + (B'(x) + L_2'(x))A(x), \\ C'(x) &= 2A'(x)A(x) + A'(x)B(x) + A(x)B'(x). \end{aligned} \tag{3}$$

Numerical calculations yield $a_1 = -0.813926$, $b_1 = -0.886469$ and $c_1 = -1.156707$. Furthermore, expanding the system

$$\begin{aligned} A(x) &= x\exp(C(x) + L_1(x))\sinh(B(x) + L_2(x)), \\ B(x) &= x\exp(C(x) + L_1(x))\cosh(B(x) + L_2(x)), \\ C(x) &= A(x)(A(x) + B(x)) \end{aligned}$$

about ρ and comparing coefficients yields

$$\begin{aligned} a_1 &= a_0c_1 + b_0b_1, \\ b_1 &= b_0c_1 + a_0b_1, \\ c_1 &= 2a_0a_1 + a_0b_1 + b_0a_1. \end{aligned} \tag{4}$$

These identities will be of use in the following section. For now, we note that the generating function $R(x) = B(x) + C(x)$ for the total number of rooted APR trees has the expansion

$$R(x) = 0.784939 - 2.043176\sqrt{x - \rho} + \ldots,$$

so that one obtains the asymptotic formula

$$r_n \sim 0.340733 \cdot n^{-3/2} \cdot 2.861365^n,$$

where r_n is the number of rooted APR trees. Finally, we observe that the ratio of APR trees among rooted trees (whose asymptotic number is given by Otter's well-known formula [12, 18]) is asymptotically

$$0.774527 \cdot 0.968062^n,$$

a result similar to that for rooted ordered trees.

3.3 Free trees

In order to deal with trees rather than rooted trees, we use, as usual, Otter's dissymmetry theorem [12, 18]. It states that the number of representations of a tree as a rooted tree equals the number of representations as a pair of two distinct rooted trees, joined by an edge between the roots, increased by 1. From Theorem 9, we see that joining two rooted trees T_1, T_2 by an edge between their roots yields an APR tree in one of the following four cases:

- T_1, T_2 are of type A, with $\mathrm{PD}(T_1) = \mathrm{PD}(T_2) = 1$,
- T_1, T_2 are APR trees of type C,
- T_1, T_2 are APR trees of type B and C (or vice versa),
- T_1 is of type A, with $\mathrm{PD}(T_1) = 1$, and T_2 is an APR tree of type B (or vice versa).

Using the same notation as before, we obtain an equation for the counting function $T(x)$ of (unrooted) APR trees:

$$\begin{aligned} T(x) &= R(x) - \frac{1}{2}\left(A(x)^2 + C(x)^2 + 2A(x)B(x) + 2B(x)C(x) - A(x^2) - C(x^2)\right) \\ &= x + x^3 + x^4 + 2x^5 + 3x^6 + 7x^7 + 12x^8 + 27x^9 + 54x^{10} + \dots. \end{aligned}$$

We expand $T(x)$ about ρ:

$$T(x) = t_0 + t_1\sqrt{\rho - x} + t_2(\rho - x) + t_3(\rho - x)^{3/2} + \dots.$$

Here, we have the following expression for t_1:

$$t_1 = b_1 + c_1 - a_0a_1 - a_1b_0 - a_0b_1 - b_1c_0 - b_0c_1 - c_0c_1.$$

Using the identities (2) and (4), we see that $t_1 = 0$, a typical phenomenon that appears in the analysis of unrooted trees. Hence, in order to obtain the desired asymptotics, we have to determine t_3 and apply Theorem 11 again. To obtain t_3, we differentiate (3) once again, solve the resulting system for A_2'', B_2'' and C_2'' and use the solution to express

$$T''(x) = \frac{3t_3}{4}(\rho - x)^{-1/2} + \dots$$

in terms of $A(x), B(x), L_1(x)$ and $L_2(x)$ (note also that $A(x^2)$ and $C(x^2)$ are holomorphic at ρ, so they don't contribute to the singularity). This gives us the numerical value $t_3 = 4.678700$, and hence the asymptotic number of APR trees on n vertices is given by

$$t_n \sim 0.409027 \cdot n^{-5/2} \cdot 2.861365^n.$$

The ratio of APR trees among all trees (see again [12, 18]) is thus asymptotically

$$0.764608 \cdot 0.968062^n.$$

4 The average parity dimension

The recursive process we introduced to compute the parity dimension can also be used to obtain the average parity dimension (at least asymptotically) by means of a generating function approach. In the following two sections, we are going to do this for rooted ordered trees and for Pólya trees resp. free trees.

4.1 Rooted ordered trees

Again, we use three different generating functions $A(x, y)$, $B(x, y)$, $C(x, y)$, one for each of the three types of trees. Since we are interested in the behavior of the parity dimension, we have to use a bivariate generating function. For instance, $A(x, y)$ is defined by

$$A(x, y) = \sum_{T\colon T \text{ is of type A}} x^{|T|} y^{\mathrm{PD}(T)}.$$

From the recursive characterization, we obtain three functional equations:

$$\begin{aligned}
A(x, y) &= \frac{xy}{2}\left(\frac{1}{1 - B(x, y) - C(x, y)} - \frac{1}{1 + B(x, y) - C(x, y)}\right),\\
B(x, y) &= \frac{x}{2}\left(\frac{1}{1 - B(x, y) - C(x, y)} + \frac{1}{1 + B(x, y) - C(x, y)}\right),\\
C(x, y) &= \frac{x}{y}\left(\frac{1}{1 - A(x, y) - B(x, y) - C(x, y)} - \frac{1}{1 - B(x, y) - C(x, y)}\right).
\end{aligned}$$

Of course, $R(x, y) = A(x, y) + B(x, y) + C(x, y)$ is the generating function for all rooted ordered trees. In order to find the asymptotics for the average parity dimension, we have to consider

$$\frac{\partial}{\partial y}\left(A(x, y) + B(x, y) + C(x, y)\right)\Big|_{y=1}. \tag{5}$$

To this end, we differentiate the functional equations for $A(x, y)$, $B(x, y)$, $C(x, y)$ with respect to y, solve for $A_y(x, y)$, $B_y(x, y)$, $C_y(x, y)$, and plug in $y = 1$. Then we obtain an expression for (5) in terms of $A = A(x, 1)$, $B = B(x, 1)$, $C = C(x, 1)$, namely

$$\frac{\partial}{\partial y}\left(A(x, y) + B(x, y) + C(x, y)\right)\Big|_{y=1} = \frac{x(1 - A - B - C)(A - B + 2AB + B^2 - AC + BC)}{(1 - B - C)(1 + B - C)(x - (1 - A - B - C)^2))}.$$

As in Section 3.1, we expand $A(x, 1), B(x, 1), C(x, 1)$ around the dominating singularity, which has to be $\frac{1}{4}$, since we know that

$$A(x, 1) + B(x, 1) + C(x, 1) = R(x, 1) = \frac{1 - \sqrt{1 - 4x}}{2}.$$

Using these expansions and the Flajolet-Odlyzko singularity analysis once again, we obtain the following results:

- Of all rooted ordered trees, 24.21% are of type A, 27.64% are of type B, and 48.15% are of type C.
- The average parity dimension of a rooted ordered tree on n vertices is asymptotically $0.036148n + 0.126778$. The constant in the main term is a zero of the polynomial $x^4 + 8x^3 + 9x^2 - 28x + 1$.

Furthermore, by plugging in $y = 2$, we obtain the average size of the null space of the closed neighborhood matrix (note that this size is given by $2^{\mathrm{PD}(T)}$) by the same method as in Section 3.1. Asymptotically, the average value of $2^{\mathrm{PD}(T)}$ for a random rooted ordered tree on n vertices is given by

$$1.093108 \cdot 1.035965^n.$$

4.2 Pólya trees and free trees

For Pólya trees, the same approach yields the following system of equations:

$$A(x,y) = xy \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} C(x^k,y^k) + \sum_{k=1}^{\infty} \frac{1}{2k} B(x^{2k},y^{2k})\right) \sinh\left(\sum_{k=1}^{\infty} \frac{1}{2k-1} B(x^{2k-1},y^{2k-1})\right),$$

$$B(x,y) = x \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} C(x^k,y^k) + \sum_{k=1}^{\infty} \frac{1}{2k} B(x^{2k},y^{2k})\right) \cosh\left(\sum_{k=1}^{\infty} \frac{1}{2k-1} B(x^{2k-1},y^{2k-1})\right),$$

$$C(x,y) = \frac{x}{y} \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} \left(B(x^k,y^k) + C(x^k,y^k)\right)\right) \left(\exp\left(\sum_{k=1}^{\infty} \frac{1}{k} A(x^k,y^k)\right) - 1\right).$$

Again, we are interested in the derivative at $y = 1$, i.e. $A_y(x,1) + B_y(x,1) + C_y(x,1)$. By differentiating and plugging in $y = 1$, we obtain

$$A_y(x,1) = A(x,1)\left(1 + \sum_{k=1}^{\infty} C_y(x^k,1) + \sum_{k=1}^{\infty} B_y(x^{2k},1)\right) + B(x,1) \sum_{k=1}^{\infty} B_y(x^{2k-1},1),$$

$$B_y(x,1) = B(x,1)\left(\sum_{k=1}^{\infty} C_y(x^k,1) + \sum_{k=1}^{\infty} B_y(x^{2k},1)\right) + A(x,1) \sum_{k=1}^{\infty} B_y(x^{2k-1},1),$$

$$C_y(x,1) = C(x,1)\left(\sum_{k=1}^{\infty} \left(B_y(x^k,1) + C_y(x^k,1)\right) - 1\right) + \left(A(x,1) + B(x,1) + C(x,1)\right) \sum_{k=1}^{\infty} A_y(x^k,1)$$

upon some simplifications. In the same manner as in section 3.2, we use the system of equations to expand $A = A(x,1)$, $B = B(x,1)$ and $C = C(x,1)$ around the singularity $\rho = 0.338322$ (Otter's well-known tree-enumeration constant [18]):

$$A(x) = a_0 + a_1\sqrt{\rho - x} + a_2(\rho - x) + a_3(\rho - x)^{3/2} + \dots,$$
$$B(x) = b_0 + b_1\sqrt{\rho - x} + b_2(\rho - x) + b_3(\rho - x)^{3/2} + \dots,$$
$$C(x) = c_0 + c_1\sqrt{\rho - x} + c_2(\rho - x) + c_3(\rho - x)^{3/2} + \dots.$$

Here, we may make use of the fact that the expansion of $R(x,1) = A(x,1) + B(x,1) + C(x,1)$, the generating function for the number of rooted trees, is well-known (and its coefficients have already been calculated with high precision). As a first result, we see that 26.55% of all rooted trees are of type A, 30.08% of type B, and 43.36% of type C.

Furthermore, we add the equations for $A_y = A_y(x,1)$, $B_y = B_y(x,1)$ and $C_y = C_y(x,1)$ to obtain

$$A_y(x,1) + B_y(x,1) + C_y(x,1) = \frac{A(x,1) - C(x,1) + (A(x,1) + B(x,1) + C(x,1)) \sum_{k\geq 2} \left(A_y(x^k,1) + B_y(x^k,1) + C_y(x^k,1)\right)}{1 - A(x,1) - B(x,1) - C(x,1)}$$

or

$$R_y(x,1) = \frac{A(x,1) - C(x,1) + R(x,1) \sum_{k\geq 2} R_y(x^k,1)}{1 - R(x,1)}.$$

It is a known fact (see [12]) that R takes on the value 1 at its singularity ρ. Hence, the denominator vanishes at the singularity. Now, we easily obtain the expansion of $R_y(x,1)$ around that singularity (note that $\sum_{k\geq 2} R_y(x^k,1)$ is holomorphic within the open circle of radius $\rho^{1/2} > \rho$), yielding the average parity dimension of a rooted tree on n vertices: it is asymptotically equal to

$$0.032040n + 0.195710.$$

Applying Theorem 9 again, it is easy to perform the step from rooted trees to trees. The corresponding generating function $T(x,y)$ is given by

$$T(x,y) = R(x,y) - \tfrac{1}{2}\Big(2C(x,y)(A(x,y)+B(x,y)) + C(x,y)^2 + \tfrac{1}{y^2}A(x,y)^2 + yB(x,y)^2 \\ + \tfrac{2}{y}A(x,y)B(x,y) - \tfrac{1}{y^2}A(x^2,y^2) - yB(x^2,y^2) - C(x^2,y^2)\Big).$$

After some simplifications, we arrive at

$$\begin{aligned} T_y(x,1) &= A(x,1)^2 + A(x,1)B(x,1) - \tfrac{1}{2}B(x,1)^2 + (1 - A(x,1) - B(x,1) - C(x,1))R_y(x,1) \\ &\quad - A(x^2,1) + A_y(x,1) + \tfrac{1}{2}B(x^2,1) + B_y(x^2,1) + C_y(x^2,1) \\ &= A(x,1)^2 + A(x,1)B(x,1) - \tfrac{1}{2}B(x,1)^2 + A(x,1) - C(x,1) + R(x,1)\sum_{k\geq 2} R_y(x^k,1) \\ &\quad - A(x^2,1) + \tfrac{1}{2}B(x^2,1) + R_y(x^2,1), \end{aligned}$$

and the asymptotic average parity dimension of a tree on n vertices follows:

$$0.032040n + 0.213217.$$

The coefficients of the linear term are indeed the same for rooted trees and trees, and this can be shown as follows: let the expansion of $R(x,1)$ and $T(x,1)$ be

$$R(x) = r_0 + r_1\sqrt{\rho - x} + r_2(\rho - x) + r_3(\rho - x)^{3/2} + \dots.$$

Then we already know that $1 = r_0 = a_0 + b_0 + c_0$, $r_1 = a_1 + b_1 + c_1$, $\dots$. By the formulas for $R_y(x,1)$ and $T_y(x,1)$ and the equation

$$T(x,1) = R(x,1) - \tfrac{1}{2}(R(x,1)^2 - R(x^2,1)),$$

the coefficient for rooted trees is given by

$$\frac{2(a_0 - c_0 + \sum_{k\geq 2} R_y(\rho^k,1))}{\rho r_1^2},$$

whereas the coefficient for trees is given by

$$\frac{2(2a_0a_1 + a_0b_1 + a_1b_0 - b_0b_1 + a_1 - c_1 + r_1\sum_{k\geq 2} R_y(\rho^k,1))}{3\rho r_1 r_2}.$$

It is well-known [12] that $r_2 = \frac{r_1^2}{3}$, which follows easily from the functional equation for R, so that we only have to prove that

$$2a_0a_1 + a_0b_1 + a_1b_0 - b_0b_1 + a_1 - c_1 = r_1(a_0 - c_0).$$

This follows upon combination of $a_1 + b_1 + c_1 = r_1$, $a_0 + b_0 + c_0 = r_0 = 1$ and the identities

$$\begin{aligned} a_1 &= a_0c_1 + b_0b_1, \\ b_1 &= b_0c_1 + a_0b_1, \end{aligned}$$

which are deduced in the same way as in Section 3.2.

5 Final remarks and conclusion

It has already been observed that the asymptotic number of APR trees is remarkable large—the proportion of APR trees among all trees on n vertices from some given family decreases asymptotically like $C \cdot \alpha^n$, where α is a constant close to 1. Accordingly, the average parity dimension grows like $a \cdot n$, where a is a small constant, close to 0 (note that the possible parity dimensions range between 0 and approximately $\frac{n}{2}$ by Corollary 10).

Finally, let us remark that, using the same recursive characterization, it is also possible to investigate different parameters on APR trees, such as the number of leaves or the total height [20].

References

[1] A. T. Amin, L. H. Clark, and P. J. Slater. Parity dimension for graphs. *Discrete Math.*, 187(1-3):1–17, 1998.

[2] A. T. Amin and P. J. Slater. Neighborhood domination with parity restrictions in graphs. In *Proceedings of the Twenty-third Southeastern International Conference on Combinatorics, Graph Theory, and Computing (Boca Raton, FL, 1992)*, volume 91, pages 19–30, 1992.

[3] A. T. Amin and P. J. Slater. All parity realizable trees. *J. Combin. Math. Combin. Comput.*, 20:53–63, 1996.

[4] A. T. Amin, P. J. Slater, and G.-H. Zhang. Parity dimension for graphs—a linear algebraic approach. *Linear Multilinear Algebra*, 50(4):327–342, 2002.

[5] E. A. Bender. Asymptotic methods in enumeration. *SIAM Rev.*, 16:485–515, 1974.

[6] J. H. Bevis, G. S. Domke, and V. A. Miller. Ranks of trees and grid graphs. *J. Combin. Math. Combin. Comput.*, 18:109–119, 1995.

[7] E. R. Canfield. Remarks on an asymptotic method in combinatorics. *J. Combin. Theory Ser. A*, 37(3):348–352, 1984.

[8] P. Flajolet and A. Odlyzko. Singularity analysis of generating functions. *SIAM J. Discrete Math.*, 3(2):216–240, 1990.

[9] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. 2006. available online at `http://algo.inria.fr/flajolet/Publications/books.html`.

[10] R. Fröberg. *An introduction to Gröbner bases*. Pure and Applied Mathematics (New York). John Wiley & Sons Ltd., Chichester, 1997.

[11] F. Harary and T. W. Haynes. The k-tuple domatic number of a graph. *Math. Slovaca*, 48(2):161–166, 1998.

[12] F. Harary and E. M. Palmer. *Graphical enumeration*. Academic Press, New York, 1973.

[13] F. Harary, R. W. Robinson, and A. J. Schwenk. Twenty-step algorithm for determining the asymptotic number of trees of various species. *J. Austral. Math. Soc. Ser. A*, 20(4):483–503, 1975.

[14] T. W. Haynes, S. T. Hedetniemi, and P. J. Slater. *Fundamentals of domination in graphs*, volume 208 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1998.

[15] E. Hille. *Analytic function theory. Vol. II.* Introductions to Higher Mathematics. Ginn and Co., Boston, Mass.-New York-Toronto, Ont., 1962.

[16] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canad. J. Math.*, 30(5):997–1015, 1978.

[17] J. W. Moon. The number of trees with a 1-factor. *Discrete Math.*, 63(1):27–37, 1987.

[18] R. Otter. The number of trees. *Ann. of Math. (2)*, 49:583–599, 1948.

[19] G. Pólya. Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.*, 68:145–254, 1937.

[20] J. Riordan and N. J. A. Sloane. The enumeration of rooted trees by total height. *J. Austral. Math. Soc.*, 10:278–282, 1969.

[21] K. Sutner. Linear cellular automata and the Garden-of-Eden. *Math. Intelligencer*, 11(2):49–53, 1989.

[22] H. S. Wilf. *generatingfunctionology*. A K Peters Ltd., Wellesley, MA, third edition, 2006.