



**HAL**  
open science

## Avalanche polynomials

Robert Cori, Anne Micheli, Dominique Rossin

► **To cite this version:**

| Robert Cori, Anne Micheli, Dominique Rossin. Avalanche polynomials. 2006. hal-00384994

**HAL Id: hal-00384994**

**<https://hal.science/hal-00384994>**

Preprint submitted on 18 May 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Avalanche polynomials

Robert Cori\*, Anne Micheli and Dominique Rossin†

May 18, 2009

## Abstract

The avalanche polynomial on a graph, introduced in [5], capture the distribution of avalanches in the abelian sandpile model. Studied on trees, this polynomial could be defined by simply considering the size of the subtrees of the original tree. In this article, we study some properties of this polynomial on plane trees. In [5], they show that two different trees could have the same avalanche polynomial. We show here that the problem of finding a tree with a prescribed polynomial is NP-complete. In a second part, we study the average and the variance of the avalanche distribution on trees and give a closed formula.

## 1 Introduction

Self-organized criticality is a concept introduced by Bak, Tang and Wiesenfeld [1] to describe the behavior of natural systems like earthquakes [4, 16], forest fires. A simple model that verifies this paradigm is the Abelian Sandpile Model on the 2-dimensional lattice [11, 7, 10]. This model is based on a cellular automaton where each cell has a number of sand grains on it and one cell topples whenever the number of grains is greater or equal to four. In this toppling, the cell gives one grain to each of its neighbor. Thus, some other cells may topple and the sequence of toppling is called an avalanche.

This model was also considered by combinatorists [2, 3, 6, 14] but on general graphs. In [5], a polynomial was introduced to encode the distribution of avalanches that was previously studied from another point of view [1, 9, 8, 13, 15]. In this paper, we study the inverse problem. Given the distribution  $D$  of avalanches, can we find a tree whose avalanche distribution is  $D$ ? Moreover we study the average and the variance of this distribution on plane trees.

---

\*LABRI, Domaine Universitaire, 351 cours de la Libération, 33405 Talence Cedex

†LIAFA, Université Paris Diderot - Paris 7, Case 7014, 75205 Paris Cedex 13

## 2 Avalanche polynomial on a plane tree

### 2.1 Definition and main result

The *avalanche polynomial* [5] encodes the size of avalanches in the abelian sandpile model. Let  $T = (V, E, r)$  be a tree rooted at  $r \in V$  whose vertex-set is  $V = \{v_1, \dots, v_{n+1}\}$ , and edge-set  $E$ . Without loss of generality we can take  $r = v_{n+1}$ . Let  $v \in V$  be a vertex. A *subtree*  $T'$  rooted at  $v$  is a tree whose vertices are descendants of  $v$  in  $T$ . The *maximal* subtree is the subtree containing all descendants of  $v$ . The size of a tree is its number of vertices.

We label the vertices  $v_i$  according to the following algorithm [5] (see Figure 1):

1. Label the root 0
2. The label of a child  $v$  of a vertex labeled  $\mu$  is  $\mu + |\text{maximal subtree rooted at } v|$ .

In [5], the authors show that the *avalanche polynomial* is also defined by  $Av_T(q) = \sum_{i \geq 1} p_i q^i$ , where  $p_i$  is the number of vertices labeled  $i$  in  $T$ . In the sequel we will either speak of the polynomial or its sequence of labels.

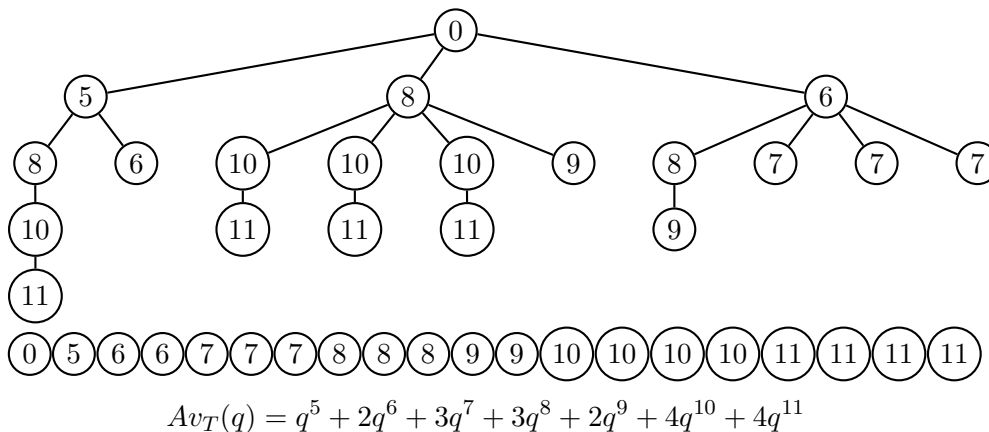


Figure 1: Example of an avalanche polynomial of a tree

In [5], they show that the avalanche polynomial is not a tree invariant by exhibiting two different trees with the same avalanche polynomial. In the sequel, we study the problem of finding a tree for a given polynomial.

**Theorem 1** *Given a polynomial  $P(q)$ , finding whether there exists a tree  $T$  whose avalanche polynomial is  $P(q)$  is NP-complete.*

## 2.2 Proof of theorem 1

We show that the 3-PARTITION problem can be reduced to our problem.

**Definition 1** *The 3-PARTITION problem  $I_{n,C}$  is the following:*

- *INPUT* :  $C \in \mathbb{N}$ ,  $(a_1, a_2, \dots, a_{3n})$  be non-negative integers such that  $\frac{C}{4} < a_i < \frac{C}{2}$ .
- *OUTPUT* :
  1. *A partition of the integers into  $n$  parts of equal sum  $C$ .*
  2. *NO if there is no such partition.*

This is a well-known example of NP-complete problem ([12]). Our proof has several steps:

1. We associate a polynomial  $P$  to each instance of 3-PARTITION problem (see Definition 2).
2. We show that  $P$  is an avalanche polynomial by exhibiting a tree  $T_P$  with  $Av_{T_P}(q) = P$ . Furthermore, the structure of  $T_P$  gives a solution to 3-PARTITION. (see Lemma 1)
3. Then, we prove that  $T_P$  is the unique tree whose avalanche polynomial is  $P$  (see Lemma 2, 3, 4, 5).

**Definition 2** *Let  $I_{n,C}$  be an instance of the 3-PARTITION problem on  $a_1, \dots, a_{3n}$ .  $P(q)$  is the polynomial associated to  $I_{n,C}$ :*

$$P(q) = nq^{C+1} + \sum_{i=1}^{3n} q^{C+1+a_i} + \sum_{i=1}^{3n} (a_i - 1)q^{C+a_i+2}$$

**Lemma 1** *The polynomial  $P(q)$  associated to an instance  $I_{n,C}$  of the 3-PARTITION problem is an avalanche polynomial. Moreover, the tree given in Figure 2 yields a solution to the 3-PARTITION problem.*

**Proof 1** *Let  $\Pi_k = (\alpha(k), \beta(k), \gamma(k))$  be a partition of  $\{1, \dots, 3n\}$  in  $n$  parts such that*

$$a_{\alpha(k)} + a_{\beta(k)} + a_{\gamma(k)} = C$$

*Note that  $P(q)$  is the avalanche polynomial of the tree  $T_P$  in Figure 2.*

*The nodes at height 1 in  $T_P$  have 3 children, whose labels are a part of  $\Pi_k$ . Hence, there exists at least a solution to the Avalanche Polynomial Problem which yields a solution to the 3-PARTITION problem.*

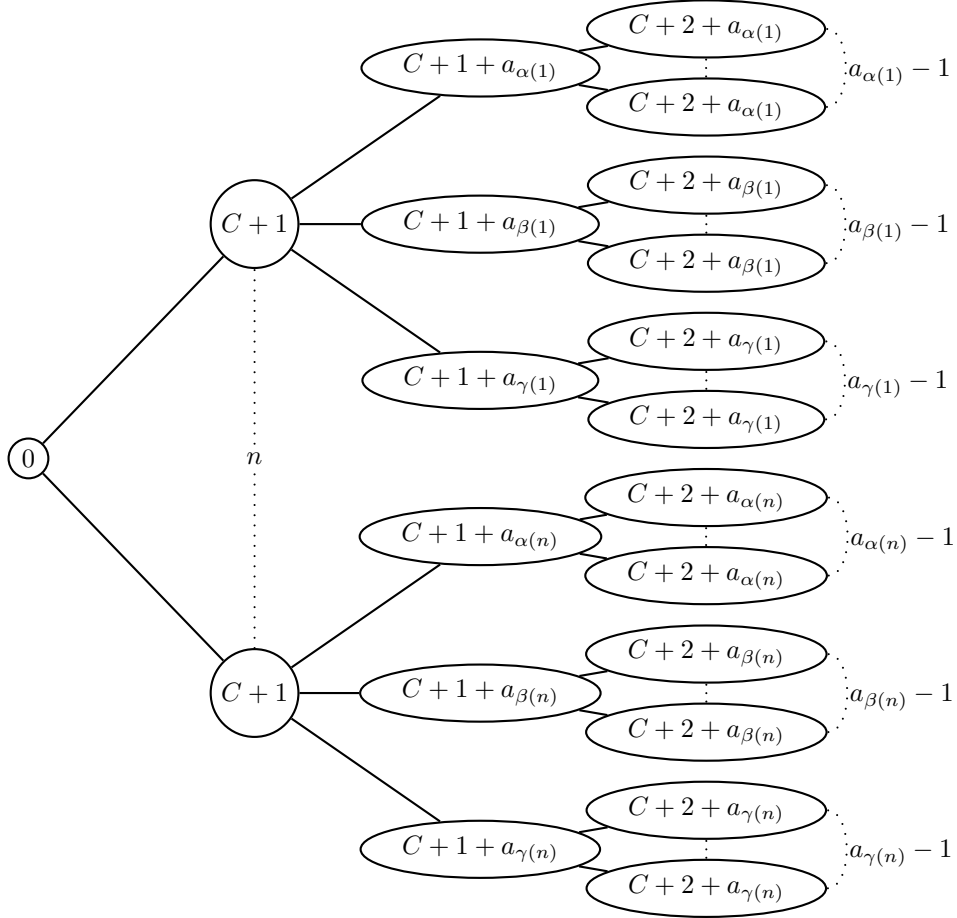


Figure 2: A solution to the 3-PARTITION problem

Nevertheless, many solutions (different non-isomorphic trees) to the **Avalanche Polynomial Problem** might exist.

To prove Theorem 1, we ensure the unicity thanks to the following remark: each value  $a_i$  (and  $C$ ) can be multiplied by a polynomial integer factor  $\lambda = \lambda(n)$ . The avalanche polynomial becomes:

$$P_\lambda(q) = nq^{\lambda C+1} + \sum_{i=1}^{3n} q^{\lambda C+1+\lambda a_i} + \sum_{i=1}^{3n} (\lambda a_i - 1)q^{\lambda C+\lambda a_i+2}$$

Note that  $\lambda$  is polynomial ensures that our transformation is a polynomial-time reduction. With the following substitutions ( $a_i \rightarrow \lambda a_i$  and  $C \rightarrow \lambda C$ ), the tree of Figure 2 is a solution of  $P_\lambda$ . We are going to prove that, taking

$\lambda$  large enough, every tree such that  $Av_T(q) = P_\lambda(q)$  yields a solution to the 3-PARTITION problem.

**Lemma 2** *The root has  $n$  children, each one labeled  $\lambda C + 1$ .*

**Proof 2** *By definition, the smallest value is the label of a child of the root. Hence the  $n$  vertices labeled  $\lambda C + 1$  are children of the root. But  $n(\lambda C + 1) = (\sum \lambda a_i) + n$ , the number of vertices. Hence all other vertices are children of one of the vertices labeled  $\lambda C + 1$ .*

Let  $I$  be the set of all vertices labeled  $\lambda C + 1 + \lambda a_i$  and  $L$  the set of those labeled  $\lambda C + \lambda a_i + 2$ ,  $\lambda \geq 2$ . Note that if  $\lambda > 1$  then  $|I| = 3n$ ,  $|L| = n(\lambda C - 3)$  and  $I \cap L = \emptyset$ .

**Lemma 3** *If  $\lambda > 2$  then the leaves of every solution tree are in  $L$ .*

**Proof 3** *Let  $v$  be a leaf of a solution tree labeled  $l_v$ . Then its parent  $p$  is labeled  $l_v - 1$ . If  $l_v = \lambda C + 1 + \lambda a_v$  there does not exist any vertex labeled  $\lambda C + \lambda a_v$ . Otherwise  $\lambda C + \lambda a_v = \lambda C + \lambda a_p + 1 + \delta$  ( $\delta \in \{0, 1\}$ ). Hence,  $\lambda(a_v - a_p) = 1 + \delta$  which contradicts  $\lambda > 2$ .  $p$  cannot be labeled  $\lambda C + 1$  since  $\lambda a_v \neq 1$ .*

**Lemma 4** *If  $\lambda > 3n$  then all vertices of  $L$  are leaves.*

**Proof 4** *Suppose that there exists an internal vertex in  $L$ . Take  $v$  as the one with the greatest label  $l_v = \lambda C + \lambda a_v + 2$ . Let  $\nu$  be a child of  $v$ . If  $\nu$  is a leaf, then its label is  $l_\nu + 1 = \lambda C + \lambda a_\nu + 3$ . But there is no such vertex since  $\lambda > 3$ . Hence  $\nu$  is an internal node. But  $v$  is the greatest internal node of  $L$  and the label of  $\nu$  is greater than  $l_v$ , hence  $\nu \in I$ .*

*Then  $v$  is either a child of the node labeled  $\lambda C + 1$  or not. The two cases are shown in Figure 3.*

*Suppose that the parent  $p$  of  $v$  is labeled  $\lambda C + 1 + \lambda a_p + \delta$ ,  $\delta = 0, 1$ . The size  $\sigma$  of the subtree rooted at  $v$  is given:*

1. *either by the difference between its label and the label of its parent,*

$$\sigma = \lambda C + \lambda a_v + 2 - (\lambda C + 1 + \lambda a_p + \delta) = \lambda(a_v - a_p) + (1 - \delta),$$

2. *or by the sum of the size of the  $t$  child-trees of  $v$  plus 1 (to take  $v$  into account):*

$$\sigma = 1 + \sum_{i=1}^t ((\lambda C + \lambda a_{s_i} + 1) - (\lambda C + \lambda a_v + 2)) = 1 - t + \sum_{i=1}^t \lambda(a_{s_i} - a_v)$$

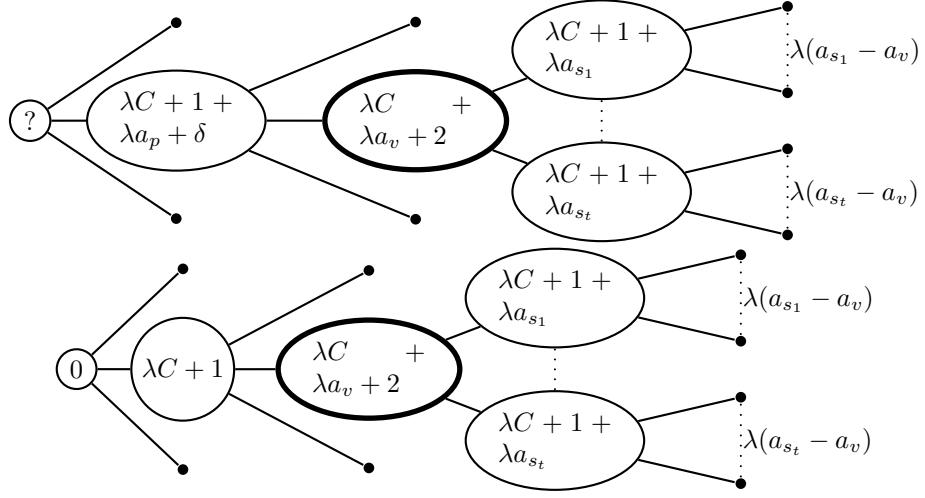


Figure 3: Two different shapes of trees that appear in Lemma 4.

Hence  $\lambda(a_v - a_p) + (1 - \delta) = 1 - t + \sum_{i=1}^t \lambda(a_{s_i} - a_v)$ , and  
 $\lambda(\sum_{i=1}^t a_{s_i} - (t+1)a_v + a_p) = t - \delta$ .

But  $t \leq |I| = 3n$  (the number of values  $a_i$ ). Hence  $t - \delta < 3n$ . Taking  $\lambda > 3n$  concludes whenever  $p$  is labeled  $\lambda C + \lambda a_p + 1 + \delta$ ,  $\delta = 0, 1$ .

Note that when  $p$  is labeled  $\lambda C + 1$ , taking  $a_p = 0$  and  $\delta = 0$  in the preceding proof concludes.

**Lemma 5** *If  $\lambda > 3n$ , then the nodes of  $I$  are children of one of the vertices labeled  $\lambda C + 1$ .*

**Proof 5** *Let  $v \in I$  be a vertex labeled  $\lambda C + \lambda a_v + 1$ . By Lemmas 3 and 4, the  $\lambda a_v - 1$  vertices labeled  $\lambda C + \lambda a_v + 2$  are leaves, hence children of  $v$ . Note that if  $k$  vertices  $v_1, \dots, v_k$  have the same label than  $v$  then there are  $k(\lambda a_v - 1)$  vertices labeled  $\lambda C + \lambda a_v + 2$ . These vertices are children of the  $v_i, 1 \leq i \leq k$ . The size of each maximal subtree rooted at  $v_i$  is less or equal than  $\lambda a_v$  since the size of the subtree is the label of  $v$  minus the label of its parent  $p$ , and the label of  $p$  is at least  $\lambda C + 1$ . Hence each  $v_i$  has exactly  $\lambda a_v - 1$  children labeled  $\lambda C + \lambda a_v + 2$ .*

*Suppose that the parent  $p$  of  $v$  is labeled  $\lambda C + 1 + \lambda a_p$ , then the size of the subtree rooted at  $v$  is  $\lambda(a_v - a_p) < \lambda a_v$ .*

*Thus the parent of  $v$  is labeled  $\lambda C + 1$ .*

### 2.3 Avalanche polynomial on plane tree of height at most 2.

We showed in the last section that finding a tree whose avalanche polynomial is given is NP-complete. In this part, we show that considering only trees of height at most 2, then the problem becomes polynomial. Moreover we give a linear time and space algorithm to solve it.

A tree of height 2 and its labeling is given in Figure 4 and its associated polynomial is  $P(q) = \sum_{i=1}^n (q^{a_i} + (a_i - 1)q^{a_i+1})$ . Note that if one of the  $a_i$  is equal to 1 then it is a leaf rooted at the vertex labeled 0.

The following algorithm takes as input a polynomial  $P(q) = \sum_{i=1}^n a_i q^i$  and output either:

- A tree  $T$  of height at most 2 whose avalanche polynomial is  $P$ .
- NO if no such tree exists.

1. Create the root with label 0.
2. Find a child of the root by taking the first non-nul coefficient  $a_j$ . If  $j = 1$  we have found  $a_j$  leaves of the root. If  $j > 1$  we have  $a_j$  nodes. Each of these nodes must have  $j - 1$  children labeled  $j + 1$ . Hence  $a_{j+1}$  must be greater than  $a_j(j - 1)$ . If not, output NO and exit.
3. Subtract  $a_j q^j + a_j(j - 1)q^{j+1}$  from the polynomial; we take this new polynomial as  $P$
4. If  $P = 0$  output the tree and exit.
5. Goto 2.

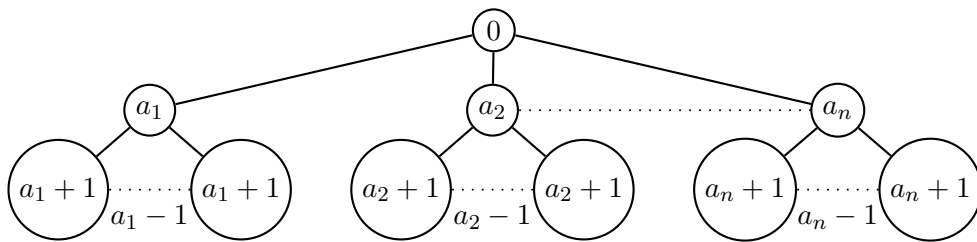


Figure 4: Labeled tree of height 2



### 3 Avalanche polynomial on plane trees

In this section, we study the avalanche polynomial on plane trees. The aim of this section is to find the average distribution of avalanches on a random tree. We first give the closed expression for the number of avalanches of a given size in plane trees. Unfortunately, this formula does not help to retrieve informations on the distribution, thus we give a recursive formula for the distribution and deduce the mean and the variance of this distribution.

#### 3.1 Definition

Let us note  $C_k = \frac{1}{k+1} \binom{2k}{k}$  the  $k$ -th Catalan number. Let  $\mathcal{T}_n$  be the set of rooted plane trees with  $n$  edges.  $|\mathcal{T}_n| = C_n$ . We extend the definition of avalanche polynomial to the set of plane trees (see Figure 5):

$$A(t, q) = \sum_{i \geq 0} \sum_{T \in \mathcal{T}_i} Av_T(q) t^i = \sum_{p \geq 0} A_p(q) t^p \quad (1)$$



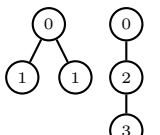
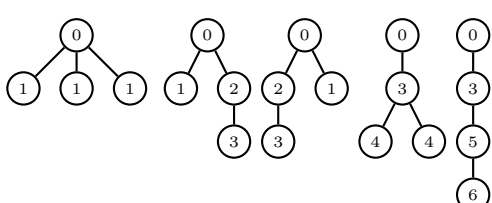
Size	Trees	Polynomial
1		0
2		$qt$
3		$(2q + q^2 + q^3)t^2$
4		$(5q + 2q^2 + 4q^3 + 2q^4 + q^5 + q^6)t^3$

Figure 5: First terms of  $A(t, q)$

**Remark 1** Note that :

$$A(t, 1) = \sum_{p \geq 0} A_p(1) t^p = \sum_{p \geq 0} p C_p t^p$$

In fact, the coefficient of  $t^p$  in  $A(t, 1)$  is the number of rooted plane trees of size  $p + 1$  times the number of vertices (excepting the roots). Hence it is  $pC_p$ .

### 3.2 Closed formula

In this section, we exhibit a decomposition of the rooted plane trees which yields a closed formula for the avalanche polynomial.

**Theorem 2** *The coefficients of the polynomials  $A_n(q)$  are given by:*

$$[q^v]A_n(q) = \sum_{k=1}^{\lfloor \frac{-1+\sqrt{1+8v}}{2} \rfloor} \sum_{p \in \Pi_{n,v,k}} C_{p_1-1} \left( \prod_{i=2}^k C_{p_i-p_{i-1}} \right) C_{n-p_k+1} \quad (2)$$

where

$$\Pi_{n,v,k} = \{(p_1, \dots, p_k); p_i < p_{i+1}, p_k \leq n, \sum_{i=1}^k p_i = v\}$$

**Proof 6** *The coefficient  $[q^v]A_n(q)$  represents the number of vertices labeled  $v$  in every plane trees of size  $n + 1$ .*

*Let  $x$  be such a vertex labeled  $v$  in a tree  $T$  and  $y_0 = x, y_1, \dots, y_k = x_{n+1}$  be the path joining the root  $x_{n+1}$  labeled 0 to the vertex  $x$  in  $T$ . Let  $lab$  be the function which maps any vertex  $y$  of  $T$  onto  $lab(y)$ , the label of  $y$  in  $T$ . The proof is based on the study of rooted plane trees which have a given path.*

- *Note that  $k \leq \lfloor \frac{-1+\sqrt{1+8v}}{2} \rfloor$  because the labels of the vertices of the path from the root to  $x$  take minimal values when the maximal subtree of  $T$  rooted at  $y_{k-1}$  is reduced to the path  $y_0, \dots, y_{k-1}$ .*

*Then the labels of the vertices on this path are  $k, k + (k - 1), k + (k - 1) + (k - 2), \dots, \frac{k(k+1)}{2}$  and if  $k > \lfloor \frac{-1+\sqrt{1+8v}}{2} \rfloor$  then the last label is strictly greater than  $v$ .*

- *The sequence of labels of the vertices  $y_l$  is strictly decreasing by definition.*
- *The sequence of the differences  $(lab(y_l) - lab(y_{l+1}))$  is strictly increasing. This relies on the fact that each difference is the size of the maximal subtree rooted at  $y_{l+1}$  and that the subtrees are nested. Moreover,  $lab(y_{k-1}) - lab(y_k) \leq n$  as this is the size of the maximal subtree rooted at  $y_{k-1}$ . Finally,  $\sum_{i=1}^k lab(y_{i-1}) - lab(y_i) = lab(y_0) - lab(y_k) = v$ .*

Let  $k$  be an integer such that  $k \leq \left\lfloor \frac{-1+\sqrt{1+8v}}{2} \right\rfloor$  and let  $(p_1, \dots, p_k)$  be a strictly increasing sequence such that  $p_k \leq n$  and  $\sum_{i=1}^k p_i = v$ . Let  $P$  be a path of vertices  $\{y_0, \dots, y_k\}$  labeled  $\text{lab}(y_0) = v, \text{lab}(y_1) = \text{lab}(y_0) - p_1, \text{lab}(y_2) = \text{lab}(y_1) - p_2, \dots, \text{lab}(y_k) = \text{lab}(y_{k-1}) - p_k = 0$  (see Figure 6). Denote by  $F_{i,1}$  (resp.  $F_{i,2}$ ) the forest hanging on  $y_i$  at the left (resp. right) of the edge  $(y_i, y_{i-1})$ .

**Lemma 6** The number of plane trees of size  $n + 1$  having  $P$  as subpath is  $C_{p_1-1} \left( \prod_{i=2}^k C_{p_i-p_{i-1}} \right) C_{n-p_k+1}$ .

**Proof 7** Note first that the size of the maximal subtree hanging on  $y_i$  is  $y_i - y_{i+1} = p_{i+1}$ . Then the number of vertices of the forest  $F_{i,1} \cup F_{i,2}$  is  $p_{i+1} - p_i - 1$  for  $i \in \{1, \dots, k-1\}$ . The number of such pair of forests  $\{F_{i,1}, F_{i,2}\}$  is  $C_{p_{i+1}-p_i}$ .

The size of the maximal subtree rooted at  $y_0$  is  $y_0 - y_1 = p_1$ . There are  $C_{p_1-1}$  such trees.

The remaining nodes are in the forests  $F_{k,1}$  and  $F_{k,2}$  which union has  $n - p_k$  vertices. So that the number of such forests is  $C_{n-p_k+1}$ .

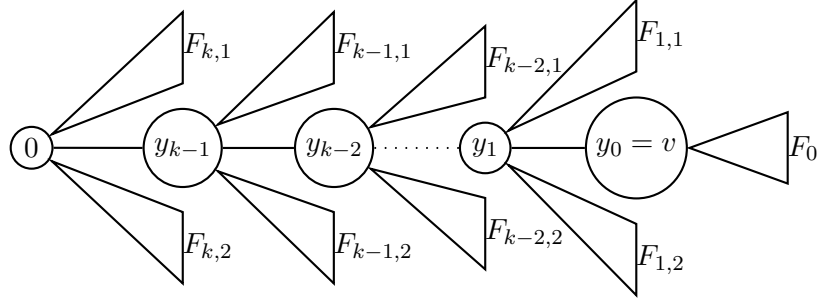


Figure 6: A tree with its subpath  $P$

**Remark 2** The coefficient of  $t^n q$  in  $A(t, q)$  is  $C_n$  and the coefficient of  $t^n q^2$  is  $C_{n-1}$ .

**Proof 8** There are two different proofs of these results, the first one consists in taking  $v = 1$  or  $v = 2$  in Equation (2).

The second one relies on the following property:

**Lemma 7** Let  $T'$  be a tree. Let  $S_{T'}$  be the set of trees  $T$  of size  $n + 1$  where  $T'$  is a subtree of  $T$  rooted at the root of  $T$ . Then  $|S_{T'}| = C_{n-|T'|+2}$ .

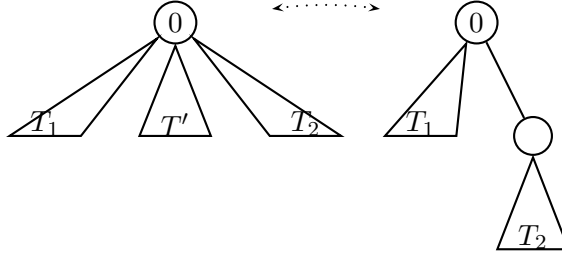


Figure 7: Bijection between  $S_{T'}$  and plane trees

**Proof 9** The bijection between the trees of  $S_{T'}$  and plane trees of size  $n - |T'| + 1$  is given in Figure 7.

For  $v = 1$  take  $T'$  as a single edge. For  $v = 2$  take  $T'$  as a path of length 2.

### 3.3 Functional equation

**Lemma 8** The polynomial  $A_p(q)$ ,  $p \geq 0$  - the distribution of avalanches on plane trees of size  $p + 1$  - is defined by the following recurrence:

$$A_{p+1}(q) = \sum_{k=0}^p C_k C_{p-k} q^{k+1} + C_{p-k} q^{k+1} A_k(q) + C_k A_{p-k}(q) \quad (3)$$

This yields a functional equation for  $A(t, q)$ :

$$A(t, q) = \frac{(1 - \sqrt{1 - 4t})(1 - \sqrt{1 - 4qt})}{4t} + q \frac{1 - \sqrt{1 - 4t}}{2} A(qt, q) + \frac{1 - \sqrt{1 - 4t}}{2} A(t, q) \quad (4)$$

**Proof 10** The proof relies on the following decomposition of rooted plane trees (see Figure 8). A tree  $T_{p+1}$  of size  $p + 2$  can be decomposed into a tree  $T_k$  of size  $k + 1$  and a tree  $T_{p-k}$  of size  $p + 1 - k$  as shown in Figure 8. Hence, the root of the tree  $T_k$  is the leftmost child of the root of  $T_{p-k}$  in  $T_{p+1}$ . Furthermore, the root of  $T_k$  is labeled  $k + 1$  the number of vertices of  $T_k$ . This vertex contributes to the term  $C_k C_{p-k} q^{k+1}$  in  $A_{p+1}(q)$ .

The avalanche polynomial of  $T_k$  is  $A_k(q)$ . In  $T_{p+1}$  the label of a vertex of the subtree  $T_k$  is increased by  $k + 1$ . They contribute to the term  $C_{p-k} q^{k+1} A_k(q)$  in  $A_{p+1}(q)$ . The last term takes into account the labels of

the vertices of  $T_{p-k}$  which stay unchanged since the root of  $T_{p-k}$  is the root of  $T_{p+1}$ .

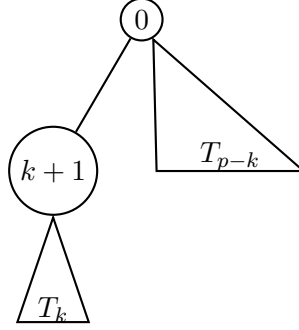


Figure 8: Decomposition of a plane tree

Substituting Equation (3) in Equation (1) yields a functional equation for  $A(t, q)$ :

$$\begin{aligned}
A(t, q) &= \frac{(1 - \sqrt{1 - 4t})(1 - \sqrt{1 - 4qt})}{4t} \\
&\quad + \sum_{p \geq 0} \sum_{k=0}^p (C_{p-k} q^{k+1} A_k(q) t^{p+1} + C_k A_{p-k}(q) t^{p+1}) \\
&= \frac{(1 - \sqrt{1 - 4t})(1 - \sqrt{1 - 4qt})}{4t} \\
&\quad + q \frac{1 - \sqrt{1 - 4t}}{2} A(qt, q) + \frac{1 - \sqrt{1 - 4t}}{2} A(t, q)
\end{aligned}$$

### 3.4 Average of the distribution of avalanches

**Proposition 1** *The average size of an avalanche on a plane tree of size  $n$  is asymptotically equal to  $\frac{\sqrt{\pi}}{4} n^{3/2}$ .*

**Proof 11**

$$A(t, q) = \sum_n \sum_m a_{m,n} q^m t^n$$

The mean size of avalanches on all plane trees of size  $n$  is given by

$$M_n = \frac{\sum_m m a_{m,n}}{\sum_m a_{m,n}} = \frac{\left[ \frac{\partial A}{\partial q}(t, 1) \right]_n}{[A(t, 1)]_n} = \frac{\left[ \frac{\partial A}{\partial q}(t, 1) \right]_n}{nC_n} \quad (5)$$

Equation (4) yields:

$$\frac{\partial A}{\partial q}(t, 1) = \frac{1 - \sqrt{1 - 4t}}{2\sqrt{1 - 4t}} + \frac{1 - \sqrt{1 - 4t}}{2} \left( A(t, 1) + \frac{\partial A}{\partial q}(t, 1) + t \frac{\partial A}{\partial t}(t, 1) + \frac{\partial A}{\partial q}(t, 1) \right)$$

Factorization of  $\frac{\partial A}{\partial q}(t, 1)$  and substitution of  $A(t, 1)$  and  $\frac{\partial A}{\partial t}(t, 1)$  with their values of Remark 1 yield:

$$\frac{\partial A}{\partial q}(t, 1) = \frac{(\sqrt{1 - 4t} - 1)(2t - 1)}{2(4t - 1)^2}$$

Since

$$\frac{(\sqrt{1 - 4t} - 1)(2t - 1)}{2} = t + \sum_{i \geq 0} (C_{i+1} - 2C_i)t^{i+2},$$

$$\begin{aligned} \frac{\partial A}{\partial q}(t, 1) &= t \sum_{i \geq 0} 4^i (i+1)t^i + \sum_{i, j \geq 0} (C_{i+1} - 2C_i)4^j (j+1)t^{i+j+2} \\ &= t \sum_{i \geq 0} 4^i (i+1)t^i + \sum_{k \geq 1} 4^k t^{k+1} \sum_{i=0}^{k-1} (C_{i+1} - 2C_i) \frac{k-i}{4^{i+1}} \end{aligned}$$

$$\text{Since } \sum_{i=0}^{k-1} \frac{C_{i+1} - 2C_i}{4^{i+1}} = -\frac{kC_k}{4^k}, \sum_{i=0}^{k-1} \frac{i(C_{i+1} - 2C_i)}{4^{i+1}} = \frac{(2 + 5k + k^2)C_k}{4^k} - 2,$$

$$\frac{\partial A}{\partial q}(t, 1) = \sum_{i \geq 1} (4^{i-1}(i+2) - (-1 + i + 2i^2)C_{i-1}) t^i$$

Substitution of  $\frac{\partial A}{\partial q}(t, 1)$  in Equation (5) yields:

$$M_n = \frac{4^{n-1}(n+2)}{nC_n} - \frac{(n+1)^2}{2n}$$

Since  $C_n \sim \frac{4^n}{\sqrt{\pi n^3}}$ ,  $M_n \sim \frac{\sqrt{\pi}}{4} n^{3/2}$ .

### 3.5 Variance

The variance of the distribution is given by  $\frac{[\frac{\partial^2 A}{\partial q^2}(t,1)]_n}{nC_n} + M_n - M_n^2$  where  $M_n$  denotes the mean size (computed in the last section).

Hence we only need the expression of  $[\frac{\partial^2 A}{\partial q^2}(t,1)]_n$ .

$$\begin{aligned} \frac{\partial^2 A}{\partial q^2}(t,1) &= \frac{t(1-\sqrt{1-4t})}{(1-4t)^{3/2}} + \frac{1-\sqrt{1-4t}}{2} \left( 2t \frac{\partial A}{\partial t}(t,1) + 2 \frac{\partial A}{\partial q}(t,1) \right. \\ &\quad \left. + t^2 \frac{\partial^2 A}{\partial t^2}(t,1) + 2 \frac{\partial^2 A}{\partial q^2}(t,1) + 2t \frac{\partial^2 A}{\partial t \partial q}(t,1) \right) \end{aligned}$$

Thus we need the other derivatives  $\frac{\partial A}{\partial t}(t,q)$ ,  $\frac{\partial A}{\partial q}(t,q)$ ,  $\frac{\partial^2 A}{\partial t^2}(t,1)$  and  $\frac{\partial^2 A}{\partial t \partial q}(t,1)$ . Substituting their expression in  $\frac{\partial^2 A}{\partial q^2}(t,1)$  yields:

$$\frac{\partial^2 A}{\partial q^2}(t,1) = -1/2 \frac{(-1 + \sqrt{1-4t})(10t + \sqrt{1-4t} - 1 - 32t^2 + 32t^3)}{(-1 + 4t)^4}$$

Since :

$$\begin{aligned} (1-4t)^{-4} &= \sum_{n \geq 0} \frac{4^n}{6} (n+1)(n+2)(n+3)t^n \\ (1-4t)^{-7/2} &= \sum_{n \geq 0} \frac{(n+1)(2n+1)(2n+3)(2n+5)4^n C_n}{15 \cdot 2^{2n}} t^n \end{aligned}$$

We obtain:

$$\begin{aligned} V_n &= \frac{4}{15}n^3 + \frac{73}{60}n^2 + \frac{26}{15}n + \frac{8}{15} - \frac{1}{2n} - \frac{1}{2n} - \frac{1}{4n^2} \\ &\quad - \frac{16^{n-1}}{n^2(n+1)^2 C_n^2} (n^4 + 6n^3 + 13n^2 + 12n + 4) \\ &\quad + \frac{4^{n-1}}{n^2(n+1)C_n} (n^3 + 4n^2 + 5n + 2) \end{aligned}$$

The leading term is:

$$\frac{4}{15}n^3 - \frac{16^{n-1}}{n^2(n+1)^2 C_n^2} n^4 \sim \left( \frac{4}{15} - \frac{\pi}{16} \right) n^3$$

In fact the distribution show strange peaks and after a renormalization (scaling in  $x$  by a factor  $1/n$  for the distribution on trees of size  $n$  and rescaling the ordinate between 0 and 1) it gives the curves in Figure 9 and 9. In the

following figures, we draw a curve for  $A_{10}(q)$ ,  $A_{20}(q)$ ,... Moreover for such a polynomial  $A_n(q) = \sum_{i \geq 0} p_i q^i$  we put points at coordinates  $(i/n, p_i/p_1)$  since  $p_1$  is the maximum. Let's recall from Remark 2 that  $p_1 = C_n$ . The grey scale goes from white (for the first polynomial) to black (for the last one).

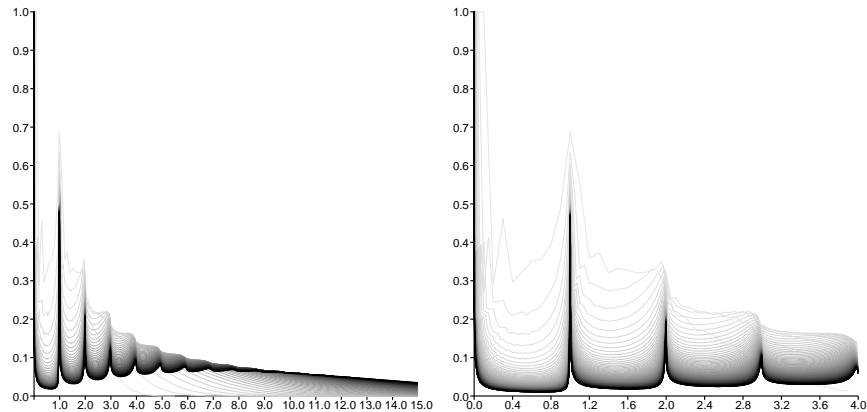


Figure 9: Distribution for the first 500 and 1000 polynomials

## 4 Conclusion

The Figures (9) and (9) point out some convergence of the renormalized distribution. A question arises naturally:

- Is it possible to retrieve the asymptotic of the coefficients from their closed form ?

Note first that the peaks that appear at integer abscissa  $x$ ,  $x \ll \sqrt{n}$  do not converge to 0. In fact consider the trees made of a path  $\pi_0 = 0, \pi_1, \dots, \pi_x$  of length  $x$  and a plane tree of size  $n - x$  rooted in  $\pi_x$ . The label of  $\pi_x$  is  $n + (n-1) + (n-2) + \dots + (n-x+1)$  thus tends to  $x$  after the renormalization by  $n$ . Moreover there are  $C_{n-x-1}$  such trees. Since  $p_1 = C_n$  the peak is higher than  $C_{n-x-1}/C_n \sim (1/4)^{x-1}$ .

## References

- [1] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: an explanation of  $1/f$  noise. *Phys. Rev. Lett.*, 59, 1987.



- [2] N. Biggs. Chip firing and the critical group of a graph. Technical report, LSE, 1996.
- [3] N. Biggs and P. Winkler. Chip-firing and the chromatic polynomial. Technical report, LSE, 1997.
- [4] K. Chen, P. Bak, and S.P. Obukhov. Self-organized criticality in a crack-propagation model of earthquakes. *Physical Review A*, 43(2):625–630, 1991.
- [5] R. Cori, A. Dartois, and D. Rossin. Avalanche polynomials on some families of graphs. *Trends in Mathematics*, pages 81–94, 2004.
- [6] R. Cori and D. Rossin. On the sandpile group of dual graphs. *European Journal of Combinatorics*, 21:447–459, 2000.
- [7] M. Creutz. Abelian sandpiles. *Computers in Physics*, pages 198–203, Mar/Apr 1991.
- [8] F. Daerden and C. Vanderzande. Sandpiles on a sierpinski gasket. *cond-mat/9712183*, 1997.
- [9] D. Dhar and S.N. Majumdar. Abelian sandpile model on the bethe lattice. *Journal of Physics A*, 23:4333–4350, 1990.
- [10] D. Dhar and S.S. Manna. Inverse avalanches in the abelian sandpile model. *Physical Review E*, 49(4), 1994.
- [11] D. Dhar, P. Ruelle, S. Sen, and D. Verma. Algebraic aspects of abelian sandpile models. *Journal of Physics A*, 28:805–831, 1995.
- [12] M.R. Garey and D.S. Johnson. Complexity results for multiprocessor scheduling with resource constraints. *SIAM J. Comp.*, 4:397–411, 1975.
- [13] K.I. Goh, D.S. Lee, B. Kahng, and D. Kim. Cascading toppling dynamics on scale-free networks. *Physica A*, 2004.
- [14] E. Goles and M. Margenstern. Universality of the chip-firing game. *Theoretical Computer Science*, 172:121–134, 1997.
- [15] C. Maes, F. Redig, and E. Saada. The abelian sandpile model on an infinite tree. *Ann. Probab.*, 30(4):2081–2107, 2002.
- [16] A. Sornette and D. Sornette. Self-organized criticality and earthquakes. *Europhysics Letters*, 9(3):197–202, 1989.