# TREES OF DEFINABLE SETS IN $\mathbb{Z}_p$

IMMANUEL HALUPCZOK

Abstract. These are the notes of a talk given at the ICMS workshop *Motivic Integration and its Interactions with Model Theory and Non-Archimedean Geometry* in May 2008. We explain the main conjecture of the article *Trees of definable sets over the p-adics* by the author.

Contents



## 1. Introduction

The metric on the $p$-adic integers $\mathbb{Z}_p$ induces a metric on subsets of $\mathbb{Z}_p^n$. Suppose we have two such subsets $X$ and $X'$ and we would like to know whether there exists an isometry between them. This would be easiest if we had some invariants such that an isometry exists if and only the invariants are equal. A first step in this direction is the following: to each set $X \subset \mathbb{Z}_p^n$ one can associate a tree $\mathrm{T}(X)$

The author was supported by the Fondation Sciences mathématiques de Paris.

such that if $X, X' \subset \mathbb{Z}_p^n$ are $p$-adically closed, then isometries $X \to X'$ correspond exactly to isomorphisms of trees $\mathrm{T}(X) \xrightarrow{\sim} \mathrm{T}(X')$.

Checking whether two arbitrary trees are isomorphic is not much easier than checking whether two sets are in isometry, so now it would be helpful if we had an explicit description of the class of trees which we can obtain from subsets of $\mathbb{Z}_p^n$. For arbitrary subsets $X \subset \mathbb{Z}_p^n$, the trees may be almost arbitrary, too, but now let us restrict ourselves to sets $X$ which are algebraic. In that case, it turns out that the possible trees are very particular ones. The goal of these notes is to present the main conjecture from [3], which gives a precise description of the class of these trees. This class will be sufficiently small and explicit so that checking whether two of its trees are isomorphic should not be a difficult problem anymore.

These notes are organized as follows. First we will introduce the trees (Subsection 1.1), present some variants of the conjecture (Section 2) and give another motivation to consider the trees, namely the Poincaré series (Section 3). The biggest part of the notes will then be the definition of the class of trees which are supposed to appear. The complete definition is rather technical, so we will start by considering only the "easier half" of it (Section 4) and verify it on some examples (Section 5); in particular, we will prove the conjecture for smooth algebraic sets. Finally, we will describe the second half of the definition (Section 6) and see how this helps understanding the Poincaré series (Section 7).

1.1. **Associating trees to subsets of $\mathbb{Z}_p^n$.** Let me first fix some notation.

**Notation 1.1.**
- Fix once and for all a prime $p$.
- $\mathbb{Q}_p$ is the field of $p$-adic numbers and $\mathbb{Z}_p$ are the $p$-adic integers. We write $v\colon \mathbb{Q}_p \to \mathbb{Z} \cup \{\infty\}$ for the valuation. It will be useful to define the valuation of a tuple as the minimum of the valuations of the coordinates: for $\underline{a} = (a_1, \dots, a_n) \in \mathbb{Q}_p^n$, set $v(\underline{a}) := \min_i a_i$.
- As usual, the metric on $\mathbb{Q}_p$ is the one induced by the norm $|a| = p^{-v(a)}$. The most natural metric on $\mathbb{Q}_p^n$ is obtained from the maximum norm $|\underline{a}| = \max_i |a_i| = p^{-v(\underline{a})}$. (This explains what we mean by "isometry"; apart from that, in these notes we will prefer to use the valuation instead of the norm.)
- The set of balls in $\mathbb{Q}_p^n$ will play a crucial role. If $\underline{a} \in \mathbb{Q}_p^n$ and $\lambda \in \mathbb{Z}$, we write $B(\underline{a}, \lambda) = \underline{a} + p^\lambda \mathbb{Z}_p^n = \{\underline{x} \in \mathbb{Q}_p^n \mid v(\underline{x} - \underline{a}) \geq \lambda\}$ for the ball around $\underline{a}$ of "valuative radius" $\lambda$. (By radius, we will always mean valuative radius.)

The set of balls contained in $\mathbb{Z}_p^n$ forms a tree under inclusion: each ball of radius $\lambda$ consists of exactly $p^n$ balls of radius $\lambda + 1$; the root of this tree is the ball $\mathbb{Z}_p^n$ itself. Another way to describe this tree is the following: a ball of radius $\lambda$ contains all elements of $\mathbb{Z}_p^n$ where the last $\lambda$ $p$-adic digits of each coordinate have been fixed; going up one step in the tree means fixing one more digit; each path to infinity corresponds to one element of $\mathbb{Z}_p^n$.

We now define the tree associated to a subset of $\mathbb{Z}_p^n$:

**Definition 1.2.** Suppose $X \subset \mathbb{Z}_p^n$. Then the *tree of* $X$

$$\mathrm{T}(X) := \{B = B(\underline{a}, \lambda) \mid B \cap X \neq \emptyset\} = \{B(\underline{a}, \lambda) \mid \underline{a} \in X, \lambda \geq 0\}$$

is the tree consisting of all balls intersecting $X$, where the tree structure is given by inclusion of balls. For any $B_0 \in \mathrm{T}(X)$, the *tree of* $X$ *on* $B_0$

$$\mathrm{T}_{B_0}(X) := \{B \in \mathrm{T}(X) \mid B \subset B_0\}$$

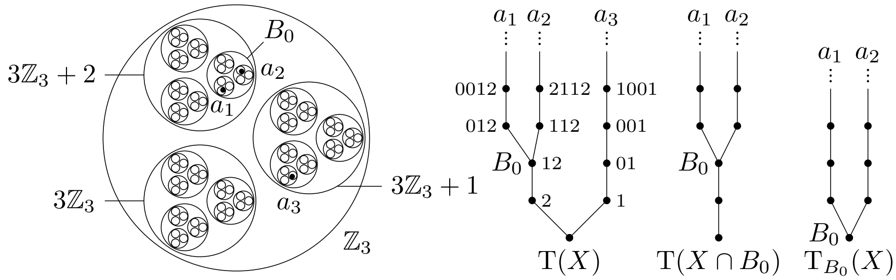


FIGURE 1. A three-point set $X = \{a_1, a_2, a_3\}$ in $\mathbb{Z}_3$, its tree, and some partial trees. Each node in the tree of $X$ corresponds to fixing some digits of the numbers $a_i$.

is the sub-tree of $\mathrm{T}(X)$ consisting of $B_0$ and everything above.

The tree $\mathrm{T}_{B_0}(X)$ is the same as $\mathrm{T}(X \cap B_0)$, except that the latter one has an additional path from $\mathbb{Z}_p^n$ to $B_0$ (see Figure 1).

**Example.**
- As we already noted, the full tree $\mathrm{T}(\mathbb{Z}_p^n)$ is the one where each node has exactly $p^n$ children.
- If $B = B(\underline{a}, \lambda) \subset \mathbb{Z}_p^n$ is a ball, then the tree $\mathrm{T}(B)$ consists of a path of length $\lambda$ (from the root to the ball $B$ itself), with a tree isomorphic to $\mathrm{T}(\mathbb{Z}_p^n)$ attached to its end. (By "a tree $\mathcal{T}$ is attached to a node $v$ of another tree", we mean that the root of $\mathcal{T}$ is identified with $v$.)
- If $X = \{\underline{a}_1, \ldots, \underline{a}_l\}$ consists of $l$ points, then $\mathrm{T}(X)$ consists of $l$ infinite paths, each one corresponding to one of the points of $X$. The paths can have a common segment at the beginning; the heights at which two paths separate is equal to $v(\underline{a}_i - \underline{a}_j)$.

More generally, one easily verifies:

**Proposition 1.3.** *There is a natural bijection between the infinite paths of* $\mathrm{T}(X)$ *and the points of the closure $\bar{X}$ of $X$ (in $p$-adic topology). In particular,* $\mathrm{T}(X) = \mathrm{T}(\bar{X})$.

This means that to get all trees $\mathrm{T}(X)$, it suffices to consider $p$-adically closed sets $X$. In the remainder of the notes, we will restrict our attention to closed sets.

Here is a first property which trees of sets obviously satisfy:

**Proposition 1.4.** *For any $X \subset \mathbb{Z}_p^n$, the tree* $\mathrm{T}(X)$ *has no leaves, i.e. no nodes without children.*

Let us now verify that isometries of sets correspond to isomorphisms of trees. Recall that a map $f\colon X \to X'$ is an isometry if and only if it satisfies $v(f(\underline{x}_1) - f(\underline{x}_2)) = v(\underline{x}_1 - \underline{x}_2)$ for any $\underline{x}_1, \underline{x}_2 \in X$.

**Proposition 1.5.** *For any two $p$-adically closed sets $X \subset \mathbb{Z}_p^n$ and $X' \subset \mathbb{Z}_p^{n'}$, we have a natural bijection*

$$\{\textit{isometries } X \longrightarrow X'\} \xleftrightarrow{1:1} \{\textit{isomorphisms } \mathrm{T}(X) \xrightarrow{\sim} \mathrm{T}(X')\}.$$

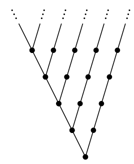

FIGURE 2. Not the tree of any algebraic set.

*Sketch of proof.* Suppose $f\colon X \to X'$ is an isometry. Then for $\underline{x} \in X$ define $f_{\mathrm{tree}}\colon B(\underline{x}, \lambda) \mapsto B(f(\underline{x}), \lambda)$. The isometry condition implies that this yields a well-defined map $\mathrm{T}(X) \to \mathrm{T}(X')$. Use $f^{-1}$ to find an inverse of $f_{\mathrm{tree}}$.

On the other hand, a map $\mathrm{T}(X) \to \mathrm{T}(X')$ defines a map on the set of infinite paths, and these correspond to the points of $X$ and $X'$, respectively. $\square$

## 2. The goal

Our goal is to understand how the tree of an algebraic set $X$ can look like. More precisely:

**Question 2.1.** *For which abstract trees $\mathcal{T}$ does there exist an algebraic set $X$ such that $\mathcal{T} \cong \mathrm{T}(V(\mathbb{Z}_p))$?*

Just to see that this is not a trivial question (i.e. that there are indeed trees which do not come from algebraic sets), consider the tree of Figure 2: the corresponding set would have infinitely many isolated points, which is impossible.

Instead of considering only algebraic subsets of $\mathbb{Z}_p^n$, we might generalize the question to some other sets. Let me present two such generalizations.

2.1. **Algebraic generalization.** Algebraic sets $X \subset \mathbb{Z}_p^n$ can be written as $X = V(\mathbb{Z}_p)$, where $V \subset \mathbb{A}^n$ is an affine embedded variety defined over $\mathbb{Z}_p$ (not necessarily irreducible). Using this, we can define $\mathrm{T}(X)$ in a more algebraic way. For $\lambda \in \mathbb{N}$, consider the following canonical maps:

$$\begin{tikzcd} V(\mathbb{Z}_p) \arrow[r] \arrow[dr, "\pi_\lambda"'] & V(\mathbb{Z}_p/p^{\lambda+1}\mathbb{Z}_p) \arrow[d, "\sigma_\lambda"] \\ & V(\mathbb{Z}_p/p^{\lambda}\mathbb{Z}_p) \end{tikzcd}$$

If $V$ is the set of zeros of some polynomials $f_i$, then

$$V(\mathbb{Z}_p/p^\lambda\mathbb{Z}_p) = \{B(\underline{x}, \lambda) \mid \underline{x} \in \mathbb{Z}_p^n, f_i(\underline{x}) \equiv 0 \mod p^\lambda\}.$$

The image of the map $\pi_\lambda$ inside this consists of those balls which do contain a point $\underline{x}$ satisfying $f_i(\underline{x}) = 0$, i.e. it is exactly the set of nodes of $\mathrm{T}(V(\mathbb{Z}_p))$ at height $\lambda$. Thus the set of nodes of $\mathrm{T}(V(\mathbb{Z}_p))$ can be defined as $\dot{\bigcup}_{\lambda=0}^{\infty} V(\mathbb{Z}_p/p^\lambda\mathbb{Z}_p)$. The tree structure on this is given by the maps $\sigma_\lambda$: they map each small ball into the bigger ball which contains the small one.

This new definition of $\mathrm{T}(V(\mathbb{Z}_p))$ has the advantage that it does not depend on the embedding of $V$ into $\mathbb{A}^n$ and that it works for non-affine $V$. So here is an algebraic generalization of the question:

**Question 2.2.** *For which abstract trees $\mathcal{T}$ does there exist a variety $V$ over $\mathbb{Z}_p$ such that $\mathcal{T} \cong \mathrm{T}(V(\mathbb{Z}_p))$?*

(Here, "variety" can also be replaced by "scheme of finite type".)

It turns out that essentially no new trees appear for non-affine varieties; this is due to the fact that whether a tree comes from an algebraic set is essentially a local question.

From the algebraic point of view, we might also consider another tree: the one whose set of nodes of height $\lambda$ is the whole set $V(\mathbb{Z}_p/p^\lambda\mathbb{Z}_p)$, instead of only the image of the map $\pi_\lambda$. These trees are quite similar in nature to the other ones; however, for these notes let us stick to the first ones.

2.2. **Model-theoretic generalization.** From a model theoretic point of view, it is also natural to replace algebraic subsets of $\mathbb{Z}_p^n$ by definable ones (say, using the two-sorted language consisting of the field $\mathbb{Q}_p$ with the ring language, the value group $\mathbb{Z}$ with the ordered group language, and the valuation map $v$). These definable sets are well understood; in particular, we have the following quantifier elimination result, which may also serve as definition of definable sets for those readers who are not so familiar with model theory.

**Proposition 2.3** (see [2], [4])**.** *The definable sets are exactly the* semi-algebraic *ones, i.e. boolean combinations of sets of the following types:*

- *algebraic sets (zero-sets of polynomials);*
- *for each polynomial $f \in \mathbb{Q}_p[X_1, \dots, X_n]$ and each $r \in \mathbb{N}$, the inverse image under $f$ of the set of $r$-th powers: $\{\underline{x} \in \mathbb{Q}_p^n \mid f(\underline{x})$ is an $r$-th power in $\mathbb{Q}_p\}$.*

(The "classical" definition of semi-algebraic sets over the reals is: boolean combinations of polynomial inequalities. In $\mathbb{R}$, being a square is equivalent to being non-negative, so there our definition of semi-algebraic yields the usual notion. The reason to generalize it to $\mathbb{Q}_p$ in the above way is that this yields precisely the smallest class of sets which contains the algebraic sets and is closed under boolean combinations and projections.)

So here is a model-theoretic version of our question:

**Question 2.4.** *For which abstract trees $\mathcal{T}$ does there exist a definable set $X$ such that $\mathcal{T} \cong \mathrm{T}(X)$?*

2.3. **The conjecture.** To state the main conjecture, we need a notion of dimension for the subsets of $\mathbb{Z}_p^n$ we are considering. For algebraic subsets, such a notion of course exists, but we have to be a little bit careful: an affine variety $V \subset \mathbb{A}^n$ might have components which have no points over $\mathbb{Q}_p$ (e.g. when $V$ is given by $X^2 = a$ and $a$ has no root in $\mathbb{Q}_p$). The following definition yields the correct dimension; moreover, even for definable subsets of $\mathbb{Z}_p^n$ we get a good notion of dimension (see [5]).

**Definition 2.5.** For $X \subset \mathbb{Q}_p^n$ definable, $\dim X$ is defined as the dimension of the Zariski closure of $X$ in $\tilde{\mathbb{Q}}_p^n$ (which is an algebraic set), where $\tilde{\mathbb{Q}}_p$ is the algebraic closure of $\mathbb{Q}_p$.

An equivalent definition would be: $\dim X$ is the maximal integer $d$ such that there exists a coordinate projection $\pi$ onto $d$ coordinates such that the image $\pi(X)$ contains an open ball.

Most of the work of these notes will be to define a class of (abstract) trees called *trees of level d*. The main conjecture is then:

**Conjecture 2.6.** *A tree $\mathcal{T}$ is of level $d$ if and only if there exists a definable set $X \subset \mathbb{Z}_p^n$ (for some suitable $n$) such that* $\mathrm{T}(X) \cong \mathcal{T}$.

The class of level $d$ trees will be surprisingly small. For this reason, the difficult direction of the conjecture is "$\Leftarrow$". Indeed, "$\Rightarrow$" can be proven by explicitly constructing, for any given tree $\mathcal{T}$ of level $d$, a corresponding definable set $X$.

The conjecture does not say when we can choose $X$ to be algebraic. However, we will see on an example that trees of algebraic sets are not essentially simpler than trees of definable sets.

Concerning "$\Leftarrow$", we will see a proof when $X$ is smooth (in a weak sense) in Subsection 5.2. Other cases in which I can prove the conjecture (see [3]) are the following: when $X$ is one-dimensional (in that case, the theorem of Puiseux gives a sufficiently explicit description of $X$), and when $X$ is an arbitrary definable subset of $\mathbb{Z}_p^2$ (use cell decomposition and apply the theorem of Puiseux to the cell centers).

## 3. Trees and Poincaré series

Before coming to the definition of level $d$ trees, let me mention one possible application of the conjecture, which was also one of the motivations to study the trees.

The *Poincaré series* of a set $X \subset \mathbb{Z}_p^n$ is defined as follows:

$$P_X(Z) := \sum_{\lambda=0}^{\infty} N_\lambda Z^\lambda \in \mathbb{Z}[[Z]],$$

where $N_\lambda$ is the number of balls of radius $\lambda$ intersecting $X$; in terms of trees, $N_\lambda$ is the number of nodes of $\mathrm{T}(X)$ at height $\lambda$.

If $X$ is a definable set, then this series is known to be a rational function in $Z$ (see [1]). This is of course a strong condition on $\mathrm{T}(X)$, and somehow it should be reflected in the structure of $\mathrm{T}(X)$. Indeed, the definition of level $d$ trees will be restrictive enough so that the rationality of the Poincaré series is easily implied; we will see this in the end of the notes. As the trees allow to compute the Poincaré series rather explicitly, one can hope that the trees will help understanding it.

## 4. Definition of trees of level $d$: part I

The definition of trees of level $d$ is inductive, so let us start defining trees of level $d = 0$.

### 4.1. Trees of level 0.

**Definition 4.1.** $\mathcal{T}$ is of level 0 iff it has no leaves and only finitely many bifurcations; in other words, it is the union of finitely many infinite paths, possibly having some segments in common at the beginning.

The 0-dimensional definable sets are exactly the finite ones. As we already saw, finite sets have trees satisfying this definition. Conversely, for any tree $\mathcal{T}$ of level 0 it is easy to construct a finite set whose points have distances given by the bifurcation heights of $\mathcal{T}$.

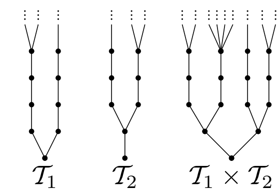


FIGURE 3. Two trees and their product.

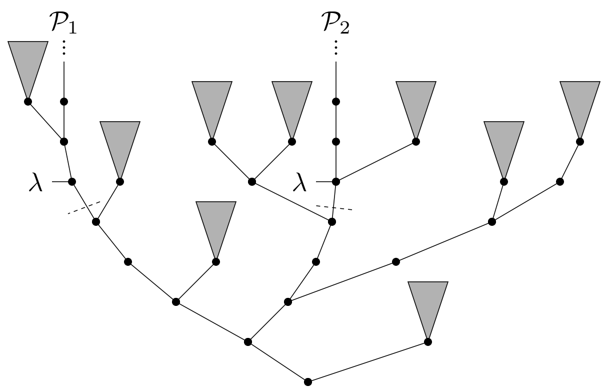


FIGURE 4. Condition (S) on trees of level $\leq d$: $\mathcal{S}_0$ consists of the two paths $\mathcal{P}_1, \mathcal{P}_2$. Each triangle is a tree of the form $\mathcal{T}' \times \mathrm{T}(\mathbb{Z}_p)$, where $\mathcal{T}'$ is of level $\leq d-1$. After cutting the paths $\mathcal{P}_i$ at any height (say, at the dotted lines), the remaining tree satisfies ($\triangledown$).

4.2. **Trees of level $d$: preliminaries.** For $d > 0$, instead of defining trees of level exactly $d$, it will be easier to define the class of trees of level $\leq d$. Then of course we can define a tree to be of level exactly $d$ if and only it is of level $\leq d$ but not of level $\leq d-1$.

The following notation will be needed in the definition:

**Notation 4.2.** If $\mathcal{T}_1$ and $\mathcal{T}_2$ are two trees, we write $\mathcal{T}_1 \times \mathcal{T}_2$ for the tree whose set of nodes at height $\lambda$ is the product of the corresponding sets of nodes of $\mathcal{T}_1$ and $\mathcal{T}_2$ (see Figure 3).

One easily checks: $\mathrm{T}(X_1 \times X_2) \cong \mathrm{T}(X_1) \times \mathrm{T}(X_2)$.

We will mainly need this notation in the form $\mathcal{T} \times \mathrm{T}(\mathbb{Z}_p)$; in this case, the product tree can be obtained from $\mathcal{T}$ using the following recursive construction: Start with a root $r$. For each child $v_i$ of the root of $\mathcal{T}$, add $p$ children to $r$. At each of these, attach a copy of $\mathcal{T}_i \times \mathrm{T}(\mathbb{Z}_p)$, where $\mathcal{T}_i$ is the sub-tree of $\mathcal{T}$ above $v$.

4.3. **Trees of level $\leq d$: the smoothness condition.** The definition of trees of level $\leq d$ can be divided into two conditions: a "smoothness condition" (S) and a "uniformity condition when approaching singularities" (U). In principle, (U) implies (S), so (S) is not really necessary. However, the complete statement of (U) is long and technical, and (S) by itself gives already a good idea on how the trees will look like. For this reason, we start by only considering (S).

**Definition 4.3.** A tree $\mathcal{T}$ is of level $\leq d$ if there exists a finite set $\mathcal{S}_0$ of infinite paths (the "singular paths") in $\mathcal{T}$ such that the following two conditions holds:

(S) For any $\lambda$, consider the tree $\tilde{\mathcal{T}}$ obtained from $\mathcal{T}$ in the following way: for each path $\mathcal{P} \in \mathcal{S}_0$, remove the node on $\mathcal{P}$ at height $\lambda$ from $\mathcal{T}$ and everything above it. We require $\tilde{\mathcal{T}}$ to satisfy the following condition:
   ($\triangledown$) $\tilde{\mathcal{T}}$ consists of a finite tree $\mathcal{F}$ with trees $\mathcal{T}'_i \times \mathrm{T}(\mathbb{Z}_p)$ attached to its leaves, where each $\mathcal{T}'_i$ is of level $\leq d-1$.

(U) For each path $\mathcal{P} \in \mathcal{S}_0$, a uniformity condition on the side branches of $\mathcal{T}$ on $\mathcal{P}$.

Condition (S) may sound complicated, but it has an easy geometric interpretation. Suppose that $\mathcal{T} = \mathrm{T}(X)$ is the tree of some $p$-adically closed set $X$. I claim that (S) translates into the following condition on $X$:

(S') There exists a finite set $S_0 \subset X$ such that the following holds: around each $x \in X \setminus S_0$, there exists a neighbourhood $B = B(x, \lambda)$ such that $X \cap B$ is isometric to a set of the form $X' \times p^\lambda \mathbb{Z}_p$, where $X'$ has a tree of level $\leq d-1$.

*Proof of the claim.* Let $S_0$ be the set of points of $X$ corresponding to the paths $\mathcal{S}_0$.

Suppose first that the sets $\mathcal{S}_0$ and $S_0$ are empty. Then the balls from (S') cover the whole set $X$. As $X$ a closed subset of $\mathbb{Z}_p^n$, it is compact, so finitely many balls suffice to cover $X$. Moreover, any two balls are either disjoint or contained in one another, so by keeping only the largest balls of our cover, we may suppose that they are all disjoint. In terms of trees, the balls $B_i$ of the cover form a "cross section" through the tree $\mathrm{T}(X)$: they are precisely the leaves of a finite sub-tree $\mathcal{F} \subset \mathrm{T}(X)$, and $\mathrm{T}(X)$ is obtained from $\mathcal{F}$ by attaching $\mathrm{T}_{B_i}(X)$ to the leaf $B_i$ (for each $i$).

The condition that $X \cap B_i$ is isometric to a set of the form $X' \times p^\lambda \mathbb{Z}_p$ (where $\lambda$ is the radius of $B_i$) translates to: $\mathrm{T}(X \cap B_i) \cong \mathcal{T}'_i \times \mathrm{T}(p^\lambda \mathbb{Z}_p)$ for some tree $\mathcal{T}'_i$ of level $\leq d-1$. After removing the path of length $\lambda$ at the lower end of the tree $\mathrm{T}(X \cap B_i)$, our condition becomes: $\mathrm{T}_{B_i}(X) \cong \mathcal{T}'_i \times \mathrm{T}(\mathbb{Z}_p)$, where $\mathcal{T}'_i$ is a tree of level $\leq d-1$. (Here, we use that removing the path of length $\lambda$ at the beginning of $\mathcal{T}'_i$ does not change whether it is of level $\leq d-1$.) This and the previous paragraph together yield exactly condition ($\triangledown$).

If we permit a non-empty singular set $S_0$ in condition (S'), then we have to take into account that $X \setminus S_0$ is not compact anymore. However, for any radius $\lambda$, the set $\tilde{X} := X \setminus \bigcup_{x \in S_0} B(x, \lambda)$ obtained from $X$ by removing balls of radius $\lambda$ around each point of $S_0$ is again compact, so we may apply the above arguments to $\tilde{X}$. The tree of $\tilde{X}$ is exactly the tree $\tilde{\mathcal{T}}$ appearing in Condition (S), so again we get that (S) and (S') are equivalent. □

## 5. Examples

Before getting to the missing part of the definition of level $\leq d$ trees, let us consider a few examples. In particular, we will prove the conjecture for smooth algebraic sets; in that case, the set $\mathcal{S}_0$ can be chosen to be empty and condition (U) becomes trivial.

5.1. **The full set $\mathbb{Z}_p^n$.** The tree of $\mathbb{Z}_p$ can be written as $\mathrm{T}(\{0\}) \times \mathrm{T}(\mathbb{Z}_p)$. As $\mathrm{T}(\{0\})$ is of level 0, $\mathrm{T}(\mathbb{Z}_p)$ is of level $\leq 1$. Inductively, one gets that $\mathrm{T}(\mathbb{Z}_p^n)$ is of level $\leq n$.

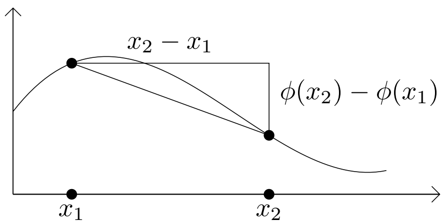


FIGURE 5. A not-too-steep curve: the vertical distance is dominated by the horizontal one, so $p$-adically the diagonal distance is equal to the horizontal one; hence the graph is isometric to the horizontal line.

To see that $\mathrm{T}(\mathbb{Z}_p^n)$ is exactly of level $n$, we use induction. It is clear that for $n > 0$, $\mathrm{T}(\mathbb{Z}_p^n)$ is not of level 0. Now suppose that $\mathrm{T}(\mathbb{Z}_p^n)$ is of level $m$ with $n > m \geq 1$ and consider condition (S). As there are only finitely many singular paths, somewhere in the tree $\mathrm{T}(\mathbb{Z}_p^n)$ there has to be a sub-tree of the form $\mathcal{T}' \times \mathrm{T}(\mathbb{Z}_p)$, where $\mathcal{T}'$ is of level $m - 1$. Now every sub-tree of $\mathrm{T}(\mathbb{Z}_p^n)$ is isomorphic to $\mathrm{T}(\mathbb{Z}_p^n)$ itself, so we get $\mathcal{T}' \times \mathrm{T}(\mathbb{Z}_p) \cong \mathrm{T}(\mathbb{Z}_p^n)$, which implies $\mathcal{T}' \cong \mathrm{T}(\mathbb{Z}_p^{n-1})$; however, $\mathrm{T}(\mathbb{Z}_p^{n-1})$ is not of level $\leq m - 1$ by induction.

With a bit of additional work, this can be turned into a general argument that if $\mathrm{T}(X)$ can not be a tree of level $d$ if $d < \dim X$.

5.2. **Smooth algebraic sets.** For simplicity, let us consider a smooth plane curve $X$; we would like to verify that it satisfies the geometric condition (S'), i.e. we have to show that $X$ is piecewise isometric to straight lines almost everywhere. It may sound surprising that this is possible: over the reals, a curved line is never isometric to a straight one. In the $p$-adics however, this is perfectly possible. Suppose first that $X = \{(x, \phi(x)) \mid x \in \mathbb{Z}_p\}$ is the graph of a "not-too-steep function", i.e. a function $\phi$ satisfying

$$v(\phi(x_1) - \phi(x_2)) \geq v(x_1 - x_2) \tag{*}$$

for any $x_1, x_2 \in \mathbb{Z}_p$. Then one easily verifies that the map

$$\mathbb{Z}_p \to X, x \mapsto (x, \phi(x))$$

is an isometry (see Figure 5). By Proposition 1.5, we get $\mathrm{T}(X) \cong \mathrm{T}(\mathbb{Z}_p)$, so in particular it is of level 1.

Now if $X$ is an arbitrary smooth curve, then by using a $p$-adic version of the implicit function theorem and choosing coordinates adequately, we can write it piecewise as graphs of not-too-steep functions. Thus it is of level 1. The same idea also works for higher-dimensional smooth algebraic sets.

If $X$ is an algebraic set which has only isolated singularities, then we let $S_0$ be the set of these singularities; in this way, we obtain that $X$ satisfies condition (S'). However, the real difficulty in proving that $X$ has a tree of level $\dim X$ lies in verifying the uniformity condition (U) at each singularity.

A remark for algebraic geometers: If we write $X = V(\mathbb{Z}_p)$, where $V$ is a scheme of finite type over $\mathbb{Z}_p$, then we did not use smoothness of $V$ in the (strong) scheme theoretic sense; in that language, what we need is only: for any $\mathbb{Z}_p$-valued point $x\colon \operatorname{spec} \mathbb{Z}_p \to V$, $V$ is smooth at $x(\eta)$, where $\eta$ is the generic point of $\operatorname{spec} \mathbb{Z}_p$.

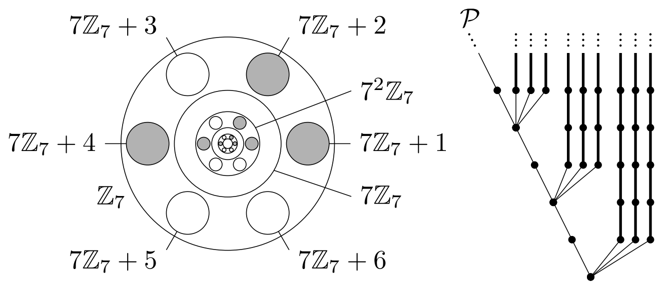


FIGURE 6. The set of squares in $\mathbb{Z}_7$ (grey area) and its tree. Each thick line in the diagram stands for a copy of the tree $\mathrm{T}(\mathbb{Z}_p)$.

5.3. **The set of squares.** Now let us consider a concrete example of a tree with a singular path. Let $X = \{x^2 \mid x \in \mathbb{Z}_p\}$ be the set of squares. (Recall that this is semi-algebraic.) For simplicity, suppose $p \neq 2$. An element of $\mathbb{Q}_p$ is a square if and only if it has even valuation and its angular component is a square in the residue field $\mathbb{F}_p$. (If $x$ has odd valuation, then a root of $x$ would have non-integer valuation. If $x$ has valuation $2\mu$, then write it as $x = (p^\mu)^2 x_0$, where $v(x_0) = 0$. By Hensels Lemma, $x_0$ is a square if and only if its residue is a square.)

In other words, $X$ is a disjoint union of balls $B(p^{2\mu}a, 2\mu + 1)$, where $\mu$ runs through $\mathbb{N}$ and $a$ runs through a set $A$ of representatives of the non-zero squares in the residue field. As the multiplicative group of $\mathbb{F}_p$ is cyclic, exactly half of its elements are squares, so $|A| = \frac{p-1}{2}$. Thus the tree of $X$ is obtained in the following way (see Figure 6):

- Start with an infinite path $\mathcal{P}$ going to 0.
- At each node of $\mathcal{P}$ of even height $\lambda$, add $\frac{p-1}{2}$ additional children (the balls $B(p^\lambda a, \lambda + 1)$ for $a \in A$).
- Attach a copy of $\mathrm{T}(\mathbb{Z}_p)$ to each of these additional children.

This tree does satisfy condition (S) if we set $\mathcal{S}_0 := \{\mathcal{P}\}$. Moreover, we see that the side branches of $\mathcal{T}$ at $\mathcal{P}$ are indeed very uniform: after fixing the height modulo 2, they are all the same. In the next example, we will see that the general situation is a bit more complicated. Moreover, we will see that complicated trees can already arise from algebraic sets, and not only from semi-algebraic ones.

5.4. **The cusp curve.** Consider the cusp curve $X = \{(x, y) \in \mathbb{Z}_p^2 \mid x^3 = y^2\}$; again suppose $p \neq 2$. The condition $x^3 = y^2$ is equivalent to: $x$ has a root and $y = \pm x\sqrt{x}$. Let $X_0$ be the set of squares from the previous example and consider one ball $B_0 = B(p^{2\mu}a, 2\mu + 1)$ of this set (again for $\mu \in \mathbb{N}$ and $a \in A$, where $A$ is our set of representatives of the non-zero residue squares). The picture of the cusp curve over $\mathbb{R}$ suggests that if $B_0$ is close to 0, then above $B_0$, $X$ should consist of two almost straight and almost parallel lines which are close together. Indeed, even when $B_0$ is not close to 0, over the $p$-adics one can show (see below) that $X \cap (B_0 \times \mathbb{Z}_p)$ is isometric to $B_0 \times \{\pm p^{3\mu}a\sqrt{a}\}$, i.e. two horizontal lines at (valuative) distance $v\left(p^{3\mu}a\sqrt{a} - (-p^{3\mu}a\sqrt{a})\right) = 3\mu$.

If $\mu = 0$, then these two lines lie in two different balls $B((a, a\sqrt{a}), 1)$ and $B((a, -a\sqrt{a}), 1)$; the sub-tree of $X$ above such a ball is isomorphic to $\mathrm{T}(\mathbb{Z}_p)$.

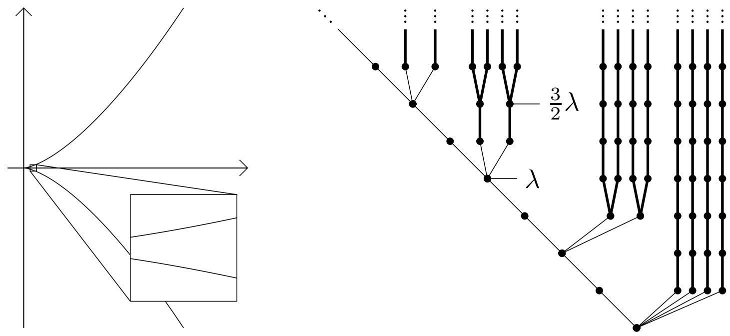


FIGURE 7. The cusp curve, a suggestive closeup, and its tree in $\mathbb{Z}_5$. The thick sub-trees in the diagram are the trees denoted by $\mathcal{T}_\lambda$ in the text; as part of the tree of the cusp curve, they stand for $\mathcal{T}_\lambda \times \mathrm{T}(\mathbb{Z}_p)$.

If $\mu > 0$, then both lines are contained in the same ball $B = B((p^{2\mu}a, 0), 2\mu+1)$; we get $\mathrm{T}_B(X) \cong \mathrm{T}(\mathbb{Z}_p) \times \mathcal{T}_\lambda$, where $\mathcal{T}_\lambda$ is the tree consisting of two paths separating at height $3\mu - (2\mu + 1) = \mu - 1$.

In other words, $\mathrm{T}(X)$ can be obtained as follows (see Figure 7):

- Start with an infinite path $\mathcal{P}$ going to 0. For each even $\lambda$, attach the following side branch $\mathcal{B}_\lambda$ to the node at height $\lambda$:
- If $\lambda = 0$, then $\mathcal{B}_\lambda$ consists of a root with $p-1$ children, above each of which the tree is isomorphic to $\mathrm{T}(\mathbb{Z}_p)$.
- If $\lambda > 0$, then $\mathcal{B}_\lambda$ consists of a root with $\frac{p-1}{2}$ children, above each of which the tree is isomorphic to $\mathcal{T}_\lambda \times \mathrm{T}(\mathbb{Z}_p)$, where $\mathcal{T}_\lambda$ is the tree consisting of two paths separating at height $\frac{\lambda}{2} - 1$.

To see that $X \cap (B_0 \times \mathbb{Z}_p)$ is isometric to two lines, we have to partition the square roots $\{\pm\sqrt{x} \mid x \in B_0\}$ into two well-defined branches. For $x, x' \in B_0$, we have $\frac{x}{x'} \in 1 + p\mathbb{Z}_p$, so by Hensels Lemma, $\frac{x}{x'}$ has exactly one root contained in $1 + p\mathbb{Z}_p$. We define that the roots $\sqrt{x}$ and $\sqrt{x'}$ lie on the same branch iff $\frac{\sqrt{x}}{\sqrt{x'}} \in 1 + p\mathbb{Z}_p$. One easily checks that this is an equivalence relation where each equivalence class contains exactly one root of each $x \in B_0$.

Now these two branches of the root function yield two branches of $X \cap (B_0 \times \mathbb{Z}_p)$, and it is an easy computation to check that each branch is the graph of a not-too-steep function. Together with the fact that they have constant distance $v(x\sqrt{x} - (-x\sqrt{x})) = 3\mu$ on $B_0$, one gets the desired isometry.

## 6. DEFINITION OF TREES OF LEVEL $d$: PART II

I will now state the missing uniformity condition (U) in the definition of trees of level $\leq d$. The cusp example gives a good idea of what it should be: for $\lambda \geq 2$ and even, each side branch $\mathcal{B}_\lambda$ consists of the same finite tree, with copies of $\mathcal{T}_\lambda \times \mathrm{T}(\mathbb{Z}_p)$ attached to it, where $\mathcal{T}_\lambda$ are trees of level 0 which are, in a certain sense, uniform in $\lambda$. This is roughly what will also happen in general. In particular, to state condition (U) we will need a notion of *uniform families* of trees of level $\leq d$; this means that we have to completely rewrite the definition.

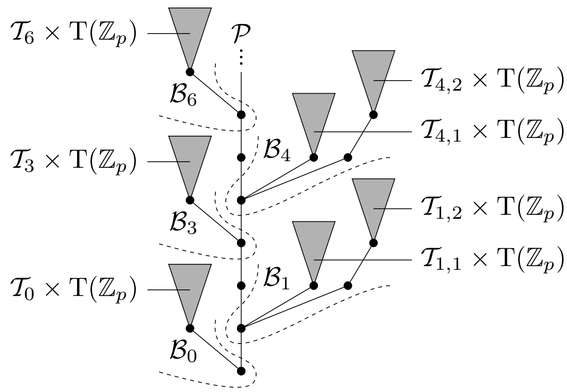


FIGURE 8. A tree which will satisfy condition (U): to the (single) singular path $\mathcal{P}$, two families of side branches have been attached: $(\mathcal{B}_{3\mu})_{\mu\in\mathbb{N}}$ and $(\mathcal{B}_{3\mu+1})_{\mu\in\mathbb{N}}$. Each of the three families $(\mathcal{T}_{3\mu})_\mu$, $(\mathcal{T}_{3\mu+1,1})_\mu$ and $(\mathcal{T}_{3\mu+1,2})_\mu$ is supposed to be a uniform family of trees of level $\leq d-1$.

When defining trees of level $\leq d$, we will need uniform families of trees of level $\leq d-1$ which are parametrized by subsets of $\mathbb{N}$. To define those families, we then need families of trees of level $\leq d-2$ parametrized by subsets of $\mathbb{N}^2$, etc. Therefore, we will consider, right from the beginning, families parametrized by subsets $M \subset \mathbb{N}^m$ for any $m \in \mathbb{N}$. (For model theorists: $M$ will be definable in the value group.)

6.1. **Uniform families of side branches.** Before I give the definition of uniform families of trees of level $\leq d$, let me define how this then yields uniform families of side branches; these side branches will then be attached to the singular paths (see Figure 8).

**Definition 6.1.** Suppose $M \subset \mathbb{N}^m$. A family of trees $(\mathcal{B}_{\underline{\kappa}})_{\underline{\kappa}\in M}$ is a *uniform family of side branches* (of level $\leq d$) if the following holds:

- Each tree $\mathcal{B}_{\underline{\kappa}}$ satisfies ($\triangledown$), i.e. it consists of a finite tree $\mathcal{F}_{\underline{\kappa}}$ with trees $\mathcal{T}'_{\underline{\kappa},i} \times \mathrm{T}(\mathbb{Z}_p)$ attached to its leaves.
- The finite trees $\mathcal{F}_{\underline{\kappa}}$ are all equal.
- For each $i$, the trees $(\mathcal{T}'_{\underline{\kappa},i})_{\underline{\kappa}}$ form a uniform family of trees of level $\leq d-1$.

Note that the side branches of the cusp example do satisfy this definition (for $\lambda \geq 2$ and even), assuming that the trees $\mathcal{T}_\lambda$ appearing there form a uniform family of level 0 trees.

6.2. **Uniform families of level 0.** Now we define uniform families of trees of level $d = 0$. Recall that a tree is of level 0 if it has finitely many bifurcations and no leaves. In the uniform version, we require that the trees of the family differ only in the lengths of the segments between the bifurcations, and moreover these lengths are linear in the parameter $\underline{\kappa}$. In other words:

**Definition 6.2.** A family of trees $(\mathcal{T}_{\underline{\kappa}})_{\underline{\kappa}\in M}$ is *uniformly of level* 0 if a single construction of the following type yields all trees $\mathcal{T}_{\underline{\kappa}}$ of the family:

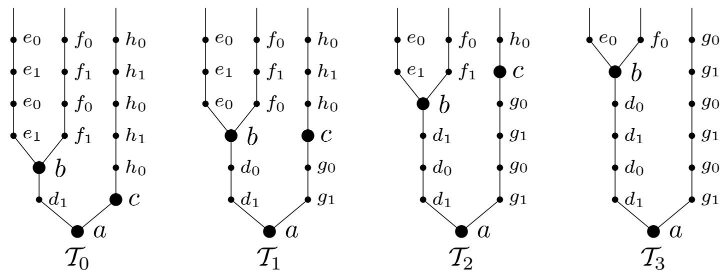


FIGURE 9. A uniform family of level 0 trees $\mathcal{T}_\lambda$ consisting of three joints $a, b, c$ (the fat points), a bone $d$ of length $\lambda + 2$, a bone $g$ of length $2\lambda + 1$, and three infinite bones $e, f, h$. If we choose $\rho = 2$ in the definition of level $\leq d$ trees, then each label (including the indices) corresponds to one uniform family of side branches.

- Start with a finite set of (not yet connected) nodes called "joints"; one of the joints will be the root of $\mathcal{T}_{\underline{\kappa}}$.
- Attach some infinite paths to some of the joints.
- Add some connections of finite length between pairs of joints; these lengths are allowed to depend linearly on $\underline{\kappa}$.

We require the result to be a tree without leaves.

Each of the infinite paths and each of the finite connections from the construction will be called a "bone".

The family of trees $\mathcal{T}_\lambda$ from the cusp example fits into this definition (for $\lambda \geq 2$ and even): start with one joint for the root and one for the (single) bifurcation, add a connection of length $\frac{1}{2}\lambda - 1$ between them, and add two infinite paths to the bifurcation joint. This example shows that although the length of a connection is an integer, as a function in $\underline{\kappa}$ it may have coefficients in $\mathbb{Q}$, as the tree $\mathcal{T}_{\underline{\kappa}}$ needs not to be defined for all $\underline{\kappa} \in \mathbb{N}^m$.

Figure 9 shows another uniform family of trees of level 0 (for the moment, ignore the small indices at the node labels). There, joint $c$ is useless; we could as well have attached an infinite path directly to joint $a$. However, such useless joints will become useful in the definition of families of trees of higher level.

6.3. **Uniform families of level $d$.** For the case $d \geq 1$, first note that if a tree $\mathcal{T}$ is of level $\leq d$, then the union of all paths in $\mathcal{S}_0$ forms a tree of level 0; let us call this union the *skeleton* of $\mathcal{T}$. Adopting this point of view, we will define a uniform family of trees of level $\leq d$ to consist of a uniform family of trees of level 0 with additional side branches. The side branches will be grouped into families, and each family will be required to satisfy Definition 6.1.

Now recall that the smoothness condition (S) required certain sub-trees $\tilde{\mathcal{T}}$ to satisfy ($\triangledown$). As Definition 6.1 implies that each side branch satisfies ($\triangledown$), the total trees will automatically satisfy (S), so we may drop condition (S) from the final definition of uniform families of level $\leq d$ trees.

Note that we now want to impose *two* kinds of uniformity on the side branches: uniformity for different side branches in the same tree (after all, this was the initial reason to introduce uniform families), and uniformity in $\underline{\kappa}$.

The uniformity inside a single tree will be required separately for each bone and each joint of the skeleton. Moreover, recall that in the cusp example, we only had uniformity separately for side branches at even heights and for side branches at odd heights. In general, we will require uniformity separately depending on the height of the side branch modulo some integer $\rho$. Here is the final and precise formulation (see Figure 9).

**Definition 6.3.** A family $(\mathcal{T}_{\underline{\kappa}})_{\underline{\kappa}\in M}$ is a *uniform family of trees of level* $\leq d$ if there exists

- a sub-tree $\mathcal{S}_{\underline{\kappa}}$ of $\mathcal{T}_{\underline{\kappa}}$ (the "skeleton")
- a natural number $\rho$

such that the following holds:

- The skeletons $(\mathcal{S}_\lambda)_{\lambda\in M}$ form a uniform family of level 0 trees.
- Fix a joint and denote, for each $\underline{\kappa}$, the side branch of $\mathcal{T}_{\underline{\kappa}}$ at that joint by $\mathcal{B}_{\underline{\kappa}}$; the family $(\mathcal{B}_{\underline{\kappa}})_{\underline{\kappa}\in M}$ has to be a uniform family of side branches. This has to hold for all joints.
- Fix a bone and denote by $N_{\underline{\kappa}}$ the set of heights at which that bone has a node in $\mathcal{T}_\kappa$ (not counting the joint(s) at the end(s) of the bone). For $\lambda \in N_{\underline{\kappa}}$, denote by $\mathcal{B}_{\underline{\kappa},\lambda}$ the side branch of $\mathcal{T}_{\underline{\kappa}}$ on the bone at height $\lambda$. Now additionally fix a congruence class $C = c + \rho\mathbb{Z}$ and define $N := \{(\underline{\kappa}, \lambda) \mid \underline{\kappa} \in M, \lambda \in N_{\underline{\kappa}} \cap C\}$. The family $(\mathcal{B}_{\underline{\kappa},\lambda})_{(\underline{\kappa},\lambda)\in N}$ has to be a uniform family of side branches (for each bone and for each congruence class).

To get the definition of a single tree of level $\leq d$, simply choose $M$ to be a one-element-set.

The congruence condition in the above definition might seem somewhat unnatural: should we really consider the absolute height of the side branch, or would it make more sense to consider the height relative to the lower end of the bone? (In the example of Figure 9, bone $e$ starts with a side branch of type $e_1$ for even $\lambda$ and with a side branch of type $e_0$ for odd $\lambda$.) Or what about the upper end of the bone, for those bones which have finite length? The answer is: it doesn't matter. More precisely, it does change the notion of uniform families, but it does not change the notion of (single) level $\leq d$ trees. We may even additionally require that in a family, the length of each bone is constant modulo $\rho$ (for all $\kappa \in M$). In that case, all the above variants become equivalent.

For simplicity, let us check this only for families indexed by a single parameter. Suppose that we have a family $(\mathcal{T}_\kappa)_{\kappa\in M}$ whose bone lengths are not constant modulo $\rho$, and suppose that this family appears in the side branches of a tree $\mathcal{T}'$. The set $M$ is a subset of a congruence class $c' + \rho'\mathbb{Z}$, where $\rho'$ is the modulus appearing in the definition of the outer tree $\mathcal{T}'$. By replacing $\rho'$ by a multiple of it, we may cut our family $(\mathcal{T}_\kappa)_{\kappa\in M}$ into several smaller families. As the bone lengths are linear in $\kappa$, we can do this in such a way that all bone lengths become constant modulo $\rho$.

## 7. Back to the Poincaré series

To finish, I will sketch how rationality of the Poincaré series can be obtained for trees $\mathcal{T}$ of level $\leq d$. Recall that the series we are interested in is the following:

$$P_{\mathcal{T}}(Z) := \sum_{\lambda=0}^{\infty} N_\lambda Z^\lambda \in \mathbb{Z}[[Z]]$$

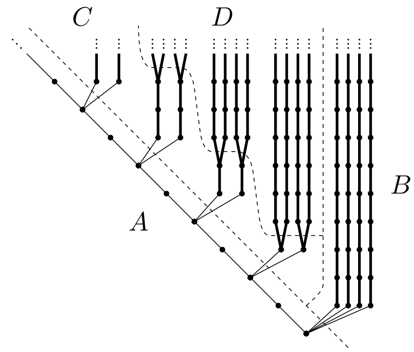


FIGURE 10. Partitioning the nodes of the tree of the cusp curve to compute its Poincaré series (don't forget that the thick lines stand for: take the product of that tree with $\mathrm{T}(\mathbb{Z}_p)$).

where $N_\lambda$ is the number of nodes of $\mathcal{T}$ at height $\lambda$.

For any subset $A$ of the nodes of $\mathcal{T}$, we can define a similar series $P_A(Z)$: use the same definition, but let $N_\lambda$ count only nodes inside $A$. For any partition $(A_i)_i$ of the nodes of $\mathcal{T}$, we then get the total series $P_\mathcal{T}(Z)$ as sum of the series $P_{A_i}(Z)$. If the partition is finite, then it suffices to prove that all $P_{A_i}(Z)$ are rational to get rationality of $P_\mathcal{T}(Z)$.

Now the idea is that our definition of level $d$ trees naturally yields a finite partition of $\mathcal{T}$ into sets, each of which is easy to describe. Recall that $\mathcal{T}$ consists of a skeleton with side branches and that the side branch come in finitely many families. Start by partitioning the nodes of $\mathcal{T}$ according to the bones and joints of the skeleton and according to the families of side branches. Each side branch family is build out of a finite tree and finitely many families of trees of level $\leq d-1$; use this to further partition the nodes of $\mathcal{T}$. Then do the same recursively for the families of trees of level $\leq d-1$ appearing in the construction. The resulting partition of $\mathcal{T}$ is still finite, as in one family, all trees together have only finitely many families of side branches.

Each of the resulting sets will have a series which can be written as nested geometric series (and such nested series are rational). Let me show this on an example. Figure 10 shows the final partition one obtains for the cusp curve (slightly simplified, in fact; by the general recipe, $B$ and $D$ would be cut into $p-1$ subsets and $C$ would be cut into $\frac{p-1}{2}$ subsets). The series of the sets $A$ and $B$ are just geometric series: $P_A(Z) = \sum_{\lambda=0}^{\infty} Z^\lambda = \frac{1}{1-Z}$ and $P_B(Z) = \sum_{\lambda=0}^{\infty}(p-1)p^\lambda Z^{\lambda+1} = \frac{(p-1)Z}{1-pZ}$. For the sets $C$ and $D$, we can write the series using an outer sum which runs over the different side branches (summand $\mu$ corresponds to the side branch starting at height $2\mu$) and an inner sum which runs over the different heights inside a side branch. The series obtained in this way are are

$$P_C(Z) = \frac{p-1}{2}\sum_{\mu=1}^{\infty}\sum_{\nu=0}^{\mu-1} p^\nu Z^{2\mu+1+\nu}$$

and

$$P_D(Z) = (p-1)\sum_{\mu=1}^{\infty}\sum_{\nu=\mu}^{\infty} p^\nu Z^{2\mu+1+\nu}.$$

Actually computing these series is left to the reader.

## References


[1] J. Denef, *The rationality of the Poincaré series associated to the p-adic points on a variety*, Invent. Math., 77 (1984), pp. 1–23.

[2] ———, *p-adic semi-algebraic sets and cell decomposition*, J. Reine Angew. Math., 369 (1986), pp. 154–166.

[3] I. Halupczok, *Trees of definable sets over the p-adics*. To appear in J. Reine Angew. Math.

[4] A. Macintyre, *On definable subsets of p-adic fields*, J. Symbolic Logic, 41 (1976), pp. 605–610.

[5] P. Scowcroft and L. van den Dries, *On the structure of semialgebraic sets over p-adic fields*, J. Symbolic Logic, 53 (1988), pp. 1138–1164.