Congruences, Trees, and p-Adic Integers
Author(s): Wolfgang M. Schmidt and C. L. Stewart
Source: *Transactions of the American Mathematical Society*, Vol. 349, No. 2 (Feb., 1997), pp. 605-639
Published by: American Mathematical Society
Stable URL: https://www.jstor.org/stable/2155389

TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 349, Number 2, February 1997, Pages 605–639
S 0002-9947(97)01547-X

CONGRUENCES, TREES, AND p-ADIC INTEGERS

WOLFGANG M. SCHMIDT AND C. L. STEWART

ABSTRACT. Let f be a polynomial in one variable with integer coefficients, and p a prime. A solution of the congruence $f(x) \equiv 0(\bmod p)$ may branch out into several solutions modulo p^2, or it may be extended to just one solution, or it may not extend to any solution. Again, a solution modulo p^2 may or may not be extendable to solutions modulo p^3, etc. In this way one obtains the "solution tree" $T = T(f)$ of congruences modulo p^λ for $\lambda = 1, 2, \ldots$. We will deal with the following questions: What is the structure of such solution trees? How many "isomorphism classes" are there of trees $T(f)$ when f ranges through polynomials of bounded degree and height? We will also give bounds for the number of solutions of congruences $f(x) \equiv 0(\bmod p^\lambda)$ in terms of p, λ and the degree of f.

INTRODUCTION

The tree $U = U(p)$ of residue classes modulo powers of a given prime p is defined as follows. Consider the diagram

$$\{0\} = \mathbb{Z}/p^0\mathbb{Z} \overset{\Phi_1}{\leftarrow} \mathbb{Z}/p^1\mathbb{Z} \overset{\Phi_2}{\leftarrow} \mathbb{Z}/p^2\mathbb{Z} \leftarrow \cdots$$

where the Φ_λ are the natural homomorphisms. The vertices of U are the elements of $\mathbb{Z}/p^\lambda\mathbb{Z}$ for $\lambda = 0, 1, \ldots$, and the directed edges are $u \to v$ where $u \in \mathbb{Z}/p^\lambda\mathbb{Z}$, $v \in \mathbb{Z}/p^{\lambda-1}\mathbb{Z}$ and $\Phi_\lambda(u) = v$ for some $\lambda > 0$. Thus U is a rooted tree with root $\{0\}$. Exactly one directed edge emanates from each vertex of U, except for the vertex $\{0\}$, from which no edge emanates. On the other hand, every vertex is the end point of precisely p directed edges. The notation $u > v$ for vertices will mean that there is a sequence of vertices and edges $u \to u^{(1)} \to \cdots \to u^{(\kappa)} = v$. Thus v is obtained from u by a finite number of applications of homomorphisms Φ_λ. We will write $u \geqq v$ if $u > v$ or $u = v$. The *level* $\lambda(u)$ of a vertex u is λ if $u \in \mathbb{Z}/p^\lambda\mathbb{Z}$. A *subtree*, or simply a *tree*, is defined as a nonempty subset T of the vertices of U such that when $u \in T$ and $u > v$, then $v \in T$. Thus T together with the directed edges $u \to v$ where $u, v \in T$ is again a tree with root $\{0\}$. But we will identify subtrees with their set of vertices. An example of a tree is the tree L^ν consisting of all vertices $u \in U$ of level $\leqq \nu$.

Given $x \in \mathbb{Z}$ (or $\in \mathbb{Z}_p$ where $\mathbb{Z}_p$ is the ring of p-adic integers), we will write x_λ for its residue class modulo p^λ, i.e., its image under the canonical isomorphism $\mathbb{Z} \to \mathbb{Z}/p^\lambda\mathbb{Z}$ (or $\mathbb{Z}_p \to \mathbb{Z}_p/p^\lambda\mathbb{Z}_p$). Every vertex of U is of the type x_λ with $x \in \mathbb{Z}$,

---

Received by the editors August 30, 1994.
1991 *Mathematics Subject Classification.* Primary 11A12, 11S05.
The first author was supported in part by NSF grant DMS–9108581.
The second author was supported in part by Grant A3528 from the Natural Sciences and Engineering Research Council of Canada.

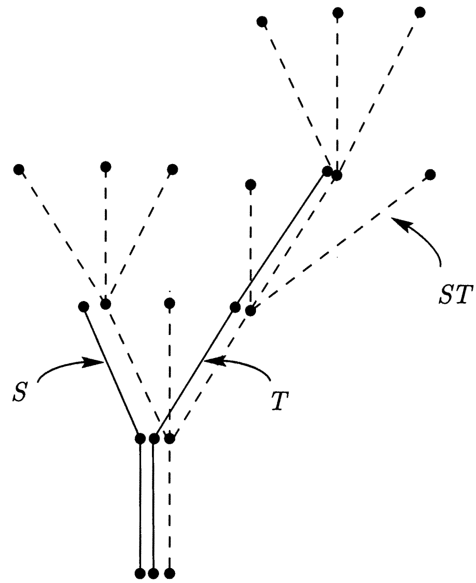

FIGURE 1. $(p = 3)$

$\lambda \in \mathbb{Z}_{\geqq 0}$. Now when $0 \leqq \mu \leqq \lambda$, we have $x_\mu \leqq x_\lambda$, and for x, y in $\mathbb{Z}$ or $\mathbb{Z}_p$ we note that $x_\lambda = y_\lambda$ implies $x_\mu = y_\mu$. We have

(1) $(xy)_\lambda = 0_\lambda$ *if and only if there exist integers* μ *and* ν *with* $\mu + \nu \geqq \lambda$ *for which* $x_\mu = 0$ *and* $y_\nu = 0$.

Now let $f \in \mathbb{Z}[X]$ or $\in \mathbb{Z}_p[X]$. We define $T(f)$ to consist of vertices x_λ where $x \in \mathbb{Z}$, $\lambda \geqq 0$, such that $(f(x))_\lambda = 0_\lambda$, i.e., $f(x) \equiv 0 \pmod{p^\lambda}$. When $x_\lambda \in T(f)$ and $\mu \leqq \lambda$, then $x_\mu \in T(f)$, so that $T(f)$ is a tree, the *solution tree* of f.

Basic to our investigation will be a notion of products of trees. Given trees $S, T \subseteq U$, we define ST to consist of vertices u for which there are $s \in S$, $t \in T$ with $s \leqq u$, $t \leqq u$ and $\lambda(u) \leqq \lambda(s) + \lambda(t)$. See Figure 1, where the underlying prime is $p = 3$.

Put differently, $x_\lambda \in ST$ if there are μ, ν with $x_\mu \in S$, $x_\nu \in T$ and $\mu \leqq \lambda$, $\nu \leqq \lambda$, $\lambda \leqq \mu + \nu$. Clearly ST is again a tree, and ST contains $S \cup T$. The conditions $\mu \leqq \lambda$, $\nu \leqq \lambda$ are not necessary, i.e., $x_\lambda \in ST$ if there are μ, ν with $x_\mu \in S$, $x_\nu \in T$ and $\lambda \leqq \mu + \nu$. We now observe that

$$T(fg) = T(f)T(g) \tag{2}$$

for any polynomials f, g : for $x_\lambda \in T(fg)$ precisely if $(f(x)g(x))_\lambda = 0_\lambda$, which by (1) is the same as $(f(x))_\mu = 0_\mu$, $(g(x))_\nu = 0_\nu$ for μ, ν with $\mu + \nu \geqq \lambda$, i.e., $x_\mu \in T(f)$, $x_\nu \in T(g)$ for μ, ν having $\mu + \nu \geqq \lambda$, and this is the same as $x_\lambda \in T(f)T(g)$.

In part I of this paper we will study the set $\mathcal{T}$ of trees $T \subseteq U$ as a structure with the binary operations of product and intersection. In part II we will study solution trees of polynomials using this structure. Among our results will be the following.

A *stalk* is defined as a tree K having at most one vertex at each level. Thus a stalk is either finite, of the type $\{0\} \leftarrow u^{(1)} \leftarrow \cdots \leftarrow u^{(\nu)}$, or infinite, of the type $\{0\} \leftarrow u^{(1)} \leftarrow \cdots$. Clearly a finite stalk may be written as $\{0\} \leftarrow x_1 \leftarrow \cdots \leftarrow x_\nu$

with $x \in \mathbb{Z}$, and an infinite stalk as $\{0\} \leftarrow x_1 \leftarrow \cdots$ with $x \in \mathbb{Z}_p$. Note that there is a 1–1 correspondence between infinite stalks and p-adic integers. In fact p-adic integers could be *defined* as infinite stalks.

Theorem A. *Let $\mathcal{P}$ consist of U and of products of finitely many stalks, i.e., trees of the form $K_1 \cdots K_m$ with stalks $K_1, \dots, K_m$ (with the empty product interpreted to be the tree $L^0 = \{0\}$). Then $\mathcal{P}$ is a substructure of $\mathcal{T}$, i.e., it is closed under intersections.*

Note that $\mathcal{P}$ is trivially closed under products. Theorem A will be proved in the context of rather more general trees than the trees $U(p)$.

We will show that every solution tree $T(f)$ lies in $\mathcal{P}$. It is U when $f = 0$, and is a finite product of stalks when $f \neq 0$. Given an ideal I in $\mathbb{Z}_p[X]$, define $T(I)$ as the intersection of the trees $T(f)$ with $f \in I$.

Theorem B. *The set of trees $T(I)$ where I runs through the ideals of $\mathbb{Z}_p[X]$ is identical with $\mathcal{P}$.*

Two trees T, T' will be called *isomorphic* if there is a bijective map $\Psi : T \to T'$ such that $t_1 < t_2$ on T precisely when $\Psi(t_1) < \Psi(t_2)$ on T'. Clearly such a map Ψ (an "isomorphism") will have $\lambda(\Psi(t)) = \lambda(t)$ for $t \in T$.

Now let $f \in \mathbb{Z}_p[X]$ be a non-constant polynomial with discriminant $D = D(f)$. Define $\delta(f)$ as the order of D with respect to p, i.e., the largest integer δ with $p^\delta \mid D$ when $D \neq 0$, and $\delta = \infty$ when $D = 0$.

Theorem C. *Given $d \geqq 2$ and $\delta \geqq 1$, the number of isomorphism classes of trees $T(f)$ where f runs through the polynomials in $\mathbb{Z}_p[X]$ of degree d and with $\delta(f) \leqq \delta$ is*

$$\leqq c_1(d)\delta^d, \tag{3}$$

where $c_1(d)$ is a positive number which depends only on d. Except for the determination of $c_1(d)$ this bound is best possible.

A variation is

Theorem C′. *Given $d \geqq 1$ and $H \geqq 1$, the number of isomorphism classes of trees $T(f)$ where f runs through the polynomials in $\mathbb{Z}[X]$ of degree $\leqq d$ and with coefficients of modulus $\leqq H$ is*

$$\leqq c_2(d)((\log H/\log p)^d + 1), \tag{4}$$

where again $c_2(d)$ is positive and depends on d. Except for the determination of $c_2(d)$ this bound is best possible.

Lastly, we mention the following.

Theorem D. *Let λ, d be positive integers, and write λ/d as a regular continued fraction:*

$$\lambda/d = c_0 + \cfrac{1}{c_1 + \cfrac{}{\ddots + 1/c_n}} \tag{5}$$

with n odd. Suppose $p \geqq d$, and $f \in \mathbb{Z}_p[X]$ is a primitive (in the sense that not all coefficients are divisible by p) polynomial of degree d. Then the number of solutions of the congruence

$$f(x) \equiv 0 \pmod{p^\lambda} \tag{6}$$

does not exceed

$$c_1 p^{\lambda-c_0-1} + c_3 p^{\lambda-c_0-c_2-1} + \cdots + c_n p^{\lambda-c_0-c_2-\cdots-c_{n-1}-1}. \tag{7}$$

This bound is best possible for every λ, d and prime $p \geqq d$.

We will deduce the

Corollary. *Define d_1 by $\lambda = dc_0 + d_1$, so that $\lambda/d = c_0 + (d_1/d)$. Note that $1 \leqq d_1 \leqq d$. Then the number of solutions of the congruence (6) is at most*

$$(d/d_1)p^{\lambda-c_0-1}.$$

In particular, the number is $\leqq dp^{\lambda-c_0-1}$. This is the bound given by Stewart in [2, (44)], since (with [ ] denoting the integer part) his exponent is $[\lambda(d-1)/d] = \lambda - c_0 + [-d_1/d] = \lambda - c_0 - 1$.

The proof of Theorem D, being the most technical, is given in the last two sections of the paper. However, sections 1,3,4, 6 up to (6.2), 8 up to (8.4) (but excluding Theorem 8.1) should provide enough background for this proof.

I. The Structure of Trees under Products and Intersections

1. General Universal Trees and the Structure $\mathcal{T}$. We will develop our theory within the framework of a more general rooted tree U. A rooted tree U consists of vertices, with a distinguished vertex $\{0\}$, its "root," together with directed edges $u \to v$ where u, v are vertices, such that exactly one edge emanates from each vertex, except that no edge emanates from $\{0\}$, and such that starting from any vertex $u \neq \{0\}$, there is a finite diagram $u \to u^{(1)} \to \cdots \to u^{(\ell)} = \{0\}$ of vertices and directed edges. It is then clear that this diagram is unique; the *level* of u is $\lambda(u) = \ell$, whereas the level of $\{0\}$ is 0. We will write $u \in U$ if u is a vertex of U. The relations $u > v$ and $u \geqq v$ are defined as before. We will call u, v *compatible* if $u \geqq v$ or $v \geqq u$; otherwise they are *incompatible.* The *valence* of a vertex u is the number (possibly an infinite cardinality) of directed edges terminating at u. We say that U is of valence ϕ (or $\geqq \phi$) if every vertex is of valence ϕ (or $\geqq \phi$). *We will suppose throughout that U is of valence* $\geqq 2$, i.e., that for every $u \in U$ there are $v \neq v'$ with $v \to u$, $v' \to u$. From now on, U will be fixed and will be called the *universal tree.*

The definition of subtrees T of U will be as in the Introduction. A vertex t of T will also be called an *element* of T and we will write $t \in T$. Again L^ν will be the tree of all vertices of U of level $\leqq \nu$. The definition of products of trees also will be as before. We are going to study the set $\mathcal{T}$ of all subtrees (briefly "trees") under the operations of product and intersection. (It is true that $\mathcal{T}$ is also closed under union, but union will play only an auxiliary rôle.) It is clear that both the product and the intersection satisfy the commutative and the associative law. In fact, a vertex u lies in a product $T_1 \cdots T_m$ precisely if there are t_i, $1 \leqq i \leqq m$, with $t_i \in T_i$, $t_i \leqq u$, such that $\lambda(u) \leqq \sum_{i=1}^m \lambda(t_i)$. We have

$$TU = U, \qquad T \cap U = T, \tag{1.1}$$

as well as

$$TL^0 = T, \qquad T \cap L^0 = L^0 \tag{1.2}$$

for every tree T, where L^0 is the tree consisting just of $\{0\}$. Finally, the distributive law

$$T(R \cap S) = TR \cap TS \tag{1.3}$$

holds: clearly the left hand side is contained in the right hand side, and when u lies in the right hand side, there are $t, t' \in T$, $r \in R$, $s \in S$, all of them $\leqq u$, having $\lambda(u) \leqq \lambda(t) + \lambda(r)$, $\lambda(u) \leqq \lambda(t') + \lambda(s)$. But since both r, s are $\leqq u$, they are compatible, say $s \leqq r$, so that $s \in R \cap S$. Now in view of $\lambda(u) \leqq \lambda(t') + \lambda(s)$ we have $u \in T(R \cap S)$, so that u lies in the left hand side of (1.3).

We will call a set with two binary operations $\cdot$ and $\cap$ a *pseudolattice* if these operations are commutative and associative, if the distributive law (1.3) holds, and if there are elements U, L^0 with (1.1), (1.2). Thus $\mathcal{T}$ is a pseudolattice.

The elements U, L^0 of a pseudolattice are clearly unique. The best known example of a pseudolattice is the set $\mathcal{J}$ of ideals in a commutative ring R with identity, under the operations $\cdot$ and $+$, and with $(0), R$ playing the rôles of U, L^0.

Given a tree T, let T_λ be the set of its vertices of level λ. In particular, $U_0 = L^0 = \{0\}$ and $U_\lambda = L^\lambda \backslash L^{\lambda-1}$ for $\lambda > 0$. Given $u \in U_1$, let $\mathcal{T}(u)$ consist of all trees $T \in \mathcal{T}$ such that T_1 is empty or consists of u only. Thus $T = L^0$ or T contains u but no other element of U_1. It is then clear that $\mathcal{T}(u)$ is a subpseudolattice of $\mathcal{T}$ (with the rôle of U played by the tree $U(u)$ consisting of $\{0\}$ and all elements $\geqq u$). Further $\mathcal{T}$ is a direct product:

$$\mathcal{T} = \prod_{u \in U_1} \mathcal{T}(u). \tag{1.4}$$

We end this section with some more notation. When $u \to v$ write $u^- = v$, so that u^- is the element directly below u. Again a *stalk* is a tree with at most one vertex at each level. A finite stalk is of the type $K(u)$, consisting of all t with $t \leqq u$. We define the level $\lambda(K(u))$ to be $\lambda(u)$. When $K = K(u)$ is finite of level $\lambda > 0$, let K^- be the stalk $K(u^-)$ of level $\lambda - 1$. An infinite stalk consists of vertices u_ν ($\nu \in \mathbb{Z}_{\geq 0}$) with $u_{\nu+1} \to u_\nu$; for such a stalk K we set $\lambda(K) = \infty$.

Given $u \in U$ and a tree T, let u_T be the element of highest level in the stalk $K(u) \cap T$. Then u lies in a product of trees $T_1, \dots, T_m$ precisely if

$$\lambda(u) \leqq \sum_{i=1}^{m} \lambda(u_{T_i}). \tag{1.5}$$

Given $u \in U$ and α, where α is a nonnegative integer or $+\infty$, write $F(u, \alpha)$ for the "fan" consisting of vertices $t \geqq u$ with $\lambda(u) \leqq \alpha$. When $\alpha \geqq \lambda(u)$, we may consider $F(u, \alpha)$ to be a tree with root u, but it is not a subtree of U as defined above unless $u = \{0\}$. When $\alpha = \lambda(u)$, then $F(u, \alpha)$ consists only of u, and when $\alpha < \lambda(u)$, then $F(u, \alpha)$ is empty. When u, v are incompatible, fans $F(u, \alpha), F(v, \beta)$ are disjoint.

When $T \subseteq U$ is a tree, we define a function α_T on U as follows. If $u \in T$, we let $\alpha_T(u)$ be the largest number (or possibly ∞) with $F(u, \alpha_T(u)) \subseteq T$. When $u \notin T$, then $\alpha_T(u) = \lambda(u_T)$. Thus $u \in T$ *precisely when* $\lambda(u) \leqq \alpha_T(u)$. It is easily seen that for trees S, T we have

$$\alpha_S(u) + \alpha_T(u) \leqq \alpha_{ST}(u). \tag{1.6}$$

The following will not be used in the sequel. In the case when U is the tree $U(p)$ of the Introduction, set $\mu(u) = p^{-\lambda(u)}$ for $u \in U$. Then $\mu(u)$ is the Haar measure of the set of elements $x \in \mathbb{Z}_p$ whose image modulo $p^{\lambda(u)}$ lies in u. Given a tree $T \subseteq U$, set

$$\mu(T) = \sum_{u \in T} \mu(u).$$

Since T is not a subset of $\mathbb{Z}_p$, $\mu(T)$ is not a Haar measure, and in fact many trees T will have $\mu(T) = \infty$. It may be shown that

$$\mu(ST) = \mu(S) + \mu(T) \tag{1.7}$$

for any subtrees S, T.

2. Lean Trees. A tree T will be called *lean* if $\alpha_T(u) < \infty$ for every $u \in U$. This means that for every u there is a vertex v of U with $v > u$, $v \notin T$.

Proposition 2.1. *T is lean if and only if cancellation by T holds, i.e., if $TR = TS$ implies $R = S$.*

Proof. Suppose that T is lean and $TR = TS$. By symmetry it will suffice to show that $R \subseteq S$. So let $r \in R$; we will show that $r \in S$.

Assume at first that $r \notin T$. Set $t = r_T$. Pick $u \geqq r$ with $\lambda(u) = \lambda(t) + \lambda(r)$; this is possible since U is of positive valence. Then $u \in TR = TS$. Therefore there exist $t' \leqq u$, $s \leqq u$ with $t' \in T$, $s \in S$ and $\lambda(u) \leqq \lambda(t') + \lambda(s)$. Now since $u \geqq t'$, $u \geqq r$, the vertices t', r are compatible, and since $r \notin T$, we have $r > t'$. By the maximal choice of t in $K(r) \cap T$, we have $\lambda(t') \leqq \lambda(t)$. Combining our relations we obtain $\lambda(r) \leqq \lambda(s)$, and since $u \geqq r$, $u \geqq s$ this gives $r \leqq s$; therefore $r \in S$.

Next, suppose that $r \in T$. We clearly may suppose that $r \neq \{0\}$. Since T is lean, $\alpha_T(r) < \infty$. There is then a $t^* > r$ having $\lambda(t^*) = \alpha_T(r)+1$ and $t^* \notin T$. But $t^* \to t$ with $\lambda(t) = \alpha_T(r)$ and $t \in T$. On the other hand, $\alpha_{TR}(t) \geqq \lambda(t) + \lambda(r) > \lambda(t)$, so that there is a $u \in TR = TS$ with $u \geqq t^*$ and $\lambda(u) = \lambda(t) + \lambda(r)$. Since $u \in TS$, we have elements $t' \in T$, $s \in S$ with $t' \leqq u$, $s \leqq u$ and $\lambda(u) \leqq \lambda(t') + \lambda(s)$. Now $t \geqq t'$ (otherwise $t' \geqq t^*$ with $t^* \notin T$), which yields $\lambda(t') \leqq \lambda(t)$. Combining our relations we get $\lambda(r) \leqq \lambda(s)$, and therefore $r \leqq s$ in view of $u \geqq s$ and $u \geqq t \geqq r$. Thus $r \in S$.

Now suppose that T is not lean. Pick $v \in T$ with $\alpha_T(v) = \infty$. Set $S = K(v)$ and $R = K(v) \cup F(v, \infty)$. Then $S \subsetneqq R$ are trees, and we will show that

$$TR = TS \tag{2.1}$$

For suppose that $u \in TR$. When $u \geqq v$, then $u \in T \subseteq TS$. When $u \ngeqq v$, we note that $u \geqq t$, $u \geqq r$ with certain $t \in T$, $r \in R$ having $\lambda(u) \leqq \lambda(t) + \lambda(r)$. But since $u \ngeqq v$, we have $r \ngeqq v$; therefore $r \in S$ and thus $u \in TS$. We have shown that $TR \subseteq TS$, and since the reverse inclusion is obvious, (2.1) follows. Therefore cancellation by T does not hold. □

Let $\mathcal{L}$ consist of U and all the lean trees.

Proposition 2.2. *$\mathcal{L}$ is a subpseudolattice of $\mathcal{T}$, i.e., $\mathcal{L}$ is closed under products.*

Note that $\mathcal{L}$ is trivially closed under intersections.

Proof. It suffices to verify that ST is lean when both S, T are lean. But if cancellation by S and T holds, then clearly cancellation by ST holds.

Every stalk is lean, and therefore so is every finite product of stalks. □

3. The Semigroup Generated by Stalks. Consider the abelian semigroup $\mathcal{K}_0$ generated by nonzero stalks, i.e., by stalks $\neq L^0$. The elements of $\mathcal{K}_0$ are of the form

$$\underline{\underline{K}} = \sum_K \oplus n(K) \cdot K, \tag{3.1}$$

where the sum is over nonzero stalks K and $n(K) \in \mathbb{Z}_{\geq 0}$, with $n(K) = 0$ for all but finitely many K. Elements of $\mathcal{K}_0$, called *stalk sums*, will also be written as

$$\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m, \tag{3.2}$$

where $K_1, \ldots, K_m$ are nonzero stalks and where the empty sum is the zero element $\underline{\underline{0}}$ of $\mathcal{K}_0$. This zero element will also be symbolized by L^0. Let $\mathcal{K}$ be the union of $\mathcal{K}_0$ and an element we will denote by $\underline{\underline{U}}$. Define $\underline{\underline{U}} \oplus \underline{\underline{K}} = \underline{\underline{U}}$, $\underline{\underline{U}} \oplus \underline{\underline{U}} = \underline{\underline{U}}$. Given a sum (3.2), define $\tau(u) = \tau_{\underline{\underline{K}}}(u)$ as the number of stalks K_i with $u \in K_i$; in particular $\tau(\{0\}) = \tau_{\underline{\underline{K}}}(\{0\}) = m$. Then

(i) $\tau(u) \geqq 0$ *for* $u \in U$,
(ii) $\sum_{u \in U_1} \tau(u) = \tau(\{0\})$,
(iii) *when* $u_i \to u$ $(i = 1, \ldots, \ell)$ *with distinct* $u_1, \ldots, u_\ell$, *then*

$$\sum_{i=1}^{\ell} \tau(u_i) \leqq \tau(u).$$

The reason for (ii) is that $K_1, \ldots, K_m$, being different from L^0, each contain a vertex of level 1. The reason for (iii) is that no stalk can contain more than one of the vertices $u_1, \ldots, u_\ell$, and every stalk containing one of them also contains u.

An integer valued function τ on U with (i), (ii), (iii) will be called a *stalk function.* We will show that for such a function there is a unique $\underline{\underline{K}} \in \mathcal{K}_0$ with $\tau = \tau_{\underline{\underline{K}}}$. Set $m = \tau(\{0\})$. When $m = 0$, then $\tau = \tau_{\underline{\underline{0}}}$. When $m > 0$, we certainly can construct $K_1, \ldots, K_m$ such that each u of level 1 is contained in precisely $\tau(u)$ of the stalks $K_1, \ldots, K_m$. In fact this determines $K_1, \ldots, K_m$ up to level 1. Suppose we have constructed $K_1, \ldots, K_m$ up to level λ such that for each u of level $\leqq \lambda$, exactly $\tau(u)$ of them contain u. Now when $\lambda(u) = \lambda$, then by (iii) there are only finitely many $u_i \to u$ with $\tau(u_i) > 0$; say $u_1, \ldots, u_\ell$. Of the $\tau(u)$ stalks containing u we continue $\tau(u_i)$ up to u_i. This is possible by (iii), and $\tau(u) - \sum_{i=1}^{\ell} \tau(u_i)$ stalks will terminate at u. In this way $K_1, \ldots, K_m$ are constructed up to level $\lambda + 1$. It is now apparent by induction on λ that there is exactly one $\underline{\underline{K}}$ with $\tau = \tau_{\underline{\underline{K}}}$. Thus there is a 1–1 correspondence between stalk functions and elements $\underline{\underline{K}} \in \mathcal{K}_0$. This correspondence can be extended to $\mathcal{K} = \mathcal{K}_0 \cup \underline{\underline{U}}$ by setting $\tau_{\underline{\underline{U}}}(u) = +\infty$ for each u. (But $\tau_{\underline{\underline{U}}}$ is not a stalk function as defined above.)

Given a stalk function τ, we set

$$\sigma(u) = \sum_{\{0\} < s \leqq u} \tau(s).$$

Then

(i) $\sigma(\{0\}) = 0$, *and* $\sigma(u') \geqq \sigma(u)$ *when* $u' \geqq u$,
(ii) $\sigma(u) > 0$ *for only finitely many* $u \in U_1$,

(iii) *when* $u_i \to u$ $(i = 1, \dots, \ell)$ *with distinct* $u_1, \dots, u_\ell$, *and* $u \to u^-$, *then*

$$\sigma(u^-) + \sum_{i=1}^{\ell} \sigma(u_i) \leqq (\ell+1)\sigma(u). \tag{3.3}$$

For (iii) is the same as

$$\sum_{i=1}^{\ell} (\sigma(u_i) - \sigma(u)) \leqq \sigma(u) - \sigma(u^-),$$

which amounts to condition (iii) for τ (restricted to vertices $u > \{0\}$). An integer valued function σ on U with (i), (ii), (iii) will be called a *sum function.*

When σ is a sum function, define τ by $\tau(u) = \sigma(u) - \sigma(u^-)$ when $u \neq \{0\}$, and by

$$\tau(\{0\}) = \sum_{u \in U_1} \sigma(u).$$

Then τ is easily seen to be a stalk function. There is a 1–1 correspondence between stalk functions and sum functions, therefore between elements $\underline{\underline{K}}$ of $\mathcal{K}_0$ and sum functions. Given $\underline{\underline{K}}$, let $\sigma_{\underline{\underline{K}}}$ be the corresponding sum function. The correspondence can be extended to $\mathcal{K} = \mathcal{K}_0 \cup \underline{\underline{U}}$ by setting $\sigma_{\underline{\underline{U}}}(u) = +\infty$ for $u \in U$. It is clear that

$$\tau_{\underline{\underline{K}}_1 \oplus \underline{\underline{K}}_2} = \tau_{\underline{\underline{K}}_1} + \tau_{\underline{\underline{K}}_2}, \quad \sigma_{\underline{\underline{K}}_1 \oplus \underline{\underline{K}}_2} = \sigma_{\underline{\underline{K}}_1} + \sigma_{\underline{\underline{K}}_2}. \tag{3.4}$$

In particular the set of sum functions is closed under addition.

We define a binary operation $\wedge$ on sum functions by

$$(\sigma \wedge \sigma')(u) = \min(\sigma(u), \sigma'(u)).$$

To see that $\sigma \wedge \sigma'$ is in fact a sum function we have to check (iii). Say without loss of generality $\sigma(u) \leqq \sigma'(u)$. Then writing $\sigma'' = \sigma \wedge \sigma'$ we have

$$\sigma''(u^-) + \sum_{i=1}^{\ell} \sigma''(u_i) \leqq \sigma(u^-) + \sum_{i=1}^{\ell} \sigma(u_i) \leqq (\ell+1)\sigma(u) = (\ell+1)\sigma''(u).$$

We consider the binary operations $+$ and $\wedge$ on the sum functions, where we include $\sigma_{\underline{\underline{U}}}$. Then the set of sum functions becomes a pseudolattice. For clearly $+$ and $\wedge$ are commutative and associative, and for any sum function σ,

$$\sigma + \sigma_{\underline{\underline{U}}} = \sigma_{\underline{\underline{U}}}, \qquad \sigma \wedge \sigma_{\underline{\underline{U}}} = \sigma,$$
$$\sigma + \sigma_{\underline{\underline{0}}} = \sigma, \qquad \sigma \wedge \sigma_{\underline{\underline{0}}} = \sigma_{\underline{\underline{0}}}.$$

It remains for us to check the distributive law

$$\sigma + (\sigma' \wedge \sigma'') = (\sigma + \sigma') \wedge (\sigma + \sigma''),$$

but this holds since

$$\sigma(u) + \min(\sigma'(u), \sigma''(u)) = \min(\sigma(u) + \sigma'(u), \sigma(u) + \sigma''(u)).$$

Because of the 1–1 correspondence between $\mathcal{K}$ and sum functions, $\wedge$ induces a binary operation on $\mathcal{K}$, also denoted by $\wedge$. In view of (3.4), $\mathcal{K}$ with the operations $\oplus, \wedge$ becomes a pseudolattice. We have

$$\underline{\underline{K}} \oplus \underline{\underline{U}} = \underline{\underline{U}}, \qquad \underline{\underline{K}} \wedge \underline{\underline{U}} = \underline{\underline{K}},$$
$$\underline{\underline{K}} \oplus \underline{\underline{0}} = \underline{\underline{K}}, \qquad \underline{\underline{K}} \wedge \underline{\underline{0}} = \underline{\underline{0}}.$$

Example. Let K be an infinite stalk. Let $r, s \in K$ with $\lambda(r) = \lambda$, $\lambda(s) = 2\lambda$. Set

$$\underline{\underline{K}} = K, \quad \underline{\underline{H}} = 2K(r), \quad \underline{\underline{F}} = K(s).$$

Then it is easily seen that for every $u \in U$, with u_K as defined in section 1,

$$\sigma_{\underline{\underline{K}}}(u) = \lambda(u_K), \quad \sigma_{\underline{\underline{H}}}(u) = \min(2\lambda(u_K), 2\lambda), \quad \sigma_{\underline{\underline{F}}}(u) = \min(\lambda(u_K), 2\lambda),$$

so that $(\sigma_{\underline{\underline{K}}} \wedge \sigma_{\underline{\underline{H}}})(u) = \min(\sigma_{\underline{\underline{K}}}(u), \sigma_{\underline{\underline{H}}}(u)) = \sigma_{\underline{\underline{F}}}(u)$, and therefore

$$\underline{\underline{K}} \wedge \underline{\underline{H}} = \underline{\underline{F}}.$$

4. Proof of Theorem A. Given $\underline{\underline{K}}$ as in (3.2), set

$$\underline{\underline{K}}^* = K_1 \cdots K_m. \tag{4.1}$$

Also set $\underline{\underline{0}}^* = L^0$, $\underline{\underline{U}}^* = U$.

Lemma 4.1. *$u \in \underline{\underline{K}}^*$ precisely when $\lambda(u) \leqq \sigma_{\underline{\underline{K}}}(u)$.*

Proof. Suppose $\underline{\underline{K}}$ is given by (3.2). Then by (1.5), $u \in \underline{\underline{K}}^*$ precisely when

$$\lambda(u) \leqq \sum_{i=1}^{m} \lambda(u_{K_i}) = \sum_{i=1}^{m} \sum_{\substack{s \in K_i \\ \{0\} < s \leqq u}} 1 = \sum_{\{0\} < s \leqq u} \tau_{\underline{\underline{K}}}(s) = \sigma_{\underline{\underline{K}}}(u).$$

□

Lemma 4.2. *For $\underline{\underline{H}}$, $\underline{\underline{K}}$ in $\mathcal{K}$,*

$$(\underline{\underline{H}} \oplus \underline{\underline{K}})^* = \underline{\underline{H}}^* \underline{\underline{K}}^*, \tag{4.2}$$

$$(\underline{\underline{H}} \wedge \underline{\underline{K}})^* = \underline{\underline{H}}^* \cap \underline{\underline{K}}^*. \tag{4.3}$$

Proof. We have $u \in \underline{\underline{H}}^* \underline{\underline{K}}^*$ precisely if there exist $h, k \leqq u$ with $h \in \underline{\underline{H}}^*$, $k \in \underline{\underline{K}}^*$ and $\lambda(u) \leqq \lambda(h) + \lambda(k)$. But then by the preceding lemma

$$\lambda(h) \leqq \sigma_{\underline{\underline{H}}}(h), \quad \lambda(k) \leqq \sigma_{\underline{\underline{K}}}(k),$$

so that

$$\lambda(u) \leqq \sigma_{\underline{\underline{H}}}(u) + \sigma_{\underline{\underline{K}}}(u). \tag{4.4}$$

Conversely, suppose (4.4) holds, and we have $\underline{\underline{H}} = H_1 \oplus \cdots \oplus H_\ell$ and (3.2). Then

$$\lambda(u) \leqq \sigma_{\underline{\underline{H}}}(u) + \sigma_{\underline{\underline{K}}}(u) = \sum_{i=1}^{\ell} \lambda(u_{H_i}) + \sum_{i=1}^{m} \lambda(u_{K_i})$$

and $u \in \underline{\underline{H}}^* \underline{\underline{K}}^*$, so that $u \in \underline{\underline{H}}^* \underline{\underline{K}}^*$ precisely when $\lambda(u) \leqq \sigma_{\underline{\underline{H}}}(u) + \sigma_{\underline{\underline{K}}}(u) = \sigma_{\underline{\underline{H}} \oplus \underline{\underline{K}}}(u)$, and (4.2) follows. This relation is obvious when $\underline{\underline{H}}$ or $\underline{\underline{K}}$ equals $\underline{\underline{U}}$.

Next, $u \in (\underline{\underline{H}} \wedge \underline{\underline{K}})^*$ precisely when $\lambda(u) \leqq \min(\sigma_{\underline{\underline{H}}}(u), \sigma_{\underline{\underline{K}}}(u))$, and this holds precisely when $u \in \underline{\underline{H}}^* \cap \underline{\underline{K}}^*$. □

Recall that $\mathcal{P} \subseteq \mathcal{T}$ as defined in the Introduction consisted of U and of products of stalks. The map $*$ from $\mathcal{K}$ into $\mathcal{T}$ is a map onto $\mathcal{P}$. In view of Lemma 4.2 we have

Theorem 4.3. *The map $*$ is a pseudolattice homomorphism from $\mathcal{K}$ onto $\mathcal{P}$.*

Corollary. *$\mathcal{P}$ is a subpseudolattice of $\mathcal{T}$.*

Recall that the operations on $\mathcal{T}$ were product and intersection. In particular, $\mathcal{P}$ is closed under intersection. Thus we have proved Theorem A in the context of universal trees more general than the trees $U(p)$ of the Introduction.

Let $\mathcal{P}' \subseteq \mathcal{T}$ consist of all trees of the type $L^n T$ with $n \geqq 0$ and $T \in \mathcal{P}$. Thus $\mathcal{P}'$ consists of U and products $L^n K_1 \cdots K_m$ where $K_1, \ldots, K_m$ are stalks. When U_1 is finite, i.e., when U contains only finitely many elements of level 1, then $L = L^1 \in \mathcal{P}$ and $\mathcal{P}' = \mathcal{P}$. But when U_1 is infinite, $\mathcal{P}'$ properly contains $\mathcal{P}$.

Theorem 4.4. *$\mathcal{P}'$ is a subpseudolattice of $\mathcal{T}$, i.e., it is closed under intersection.*

Proof. We have to show that $(L^\ell S) \cap (L^n T)$ lies in $\mathcal{P}'$ when S, T lie in $\mathcal{P}$. Say $\ell \leqq n$; then by the distributive law the above intersection is $L^\ell(S \cap L^q T)$ with $q = n - \ell \geqq 0$. We will show that $S \cap L^q T$ lies in $\mathcal{P}$. According to (1.4) we may write S as a product of factors $S(u) \in \mathcal{T}(u)$ $(u \in U_1)$, and since S is a finite product of stalks, $S(u) = \{0\}$ for u outside a finite set $M \subseteq U_1$. Similarly T is a product of factors $T(u) \in \mathcal{T}(u)$ $(u \in U_1)$, and $L = L^1$ may be considered a product (an infinite product when U_1 is infinite) $\prod_{u \in U_1} L(u)$, where $L(u)$ is the unique stalk of level 1 in $\mathcal{T}(u)$. Then

$$S \cap L^q T = \prod_{u \in M} (S(u) \cap L(u)^q T(u)).$$

Each of the factors in the product lies in $\mathcal{P}$ by the Corollary to Theorem 4.3, and therefore also the product. □

Now L as well as stalks are lean, so that

$$\mathcal{P} \subseteq \mathcal{P}' \subseteq \mathcal{L} \subseteq \mathcal{T}.$$

An element

$$\underline{\underline{K}} = \ell K \oplus n K^- \tag{4.5}$$

of $\mathcal{K}_0$, where K is a stalk of finite level $\lambda > 0$ and where $\ell \geqq 1$, $\ell + n \geqq 2$, will be called a *couple*. When $\lambda > 1$, $\underline{\underline{K}}$ is a sum of $\ell + n$ nonzero stalks, whereas for $\lambda = 1$, $\underline{\underline{K}} = \ell K \oplus n K^- = \ell K$ has ℓ summands. In particular when $\ell = 1$, $\underline{\underline{K}} = K$ will be called a *degenerate couple*; other couples are *nondegenerate*.

Say $K = K(t)$, so that K consists of $t = t_\lambda \to t_{\lambda-1} \to \cdots \to t_1 \to t_0 = \{0\}$. We have

$$\tau_{\underline{\underline{K}}}(u) = \begin{cases} 0 & \text{when} \quad u \not\leqq t, \\ \ell & \text{when} \quad u = t, \\ \ell + n & \text{when} \quad u < t, \end{cases}$$

and therefore

$$\sigma_{\underline{\underline{K}}}(u) = \begin{cases} (\ell + n)\lambda(u_K) \text{ when } \lambda(u_K) < \lambda(t), \\ (\ell + n)(\lambda - 1) + \ell \text{ when } \lambda(u_K) = \lambda(t). \end{cases}$$

$\underline{\underline{K}}^*$ is the union of the fans

$$F(t_\nu, (\ell + n)\nu) \quad \text{with} \quad 0 \leqq \nu < \lambda$$

and the fan

$$F(t_\lambda, (\ell + n)(\lambda - 1) + \ell).$$

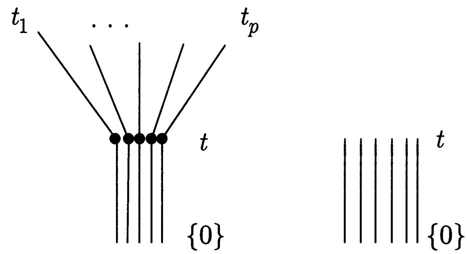

FIGURE 2. ($p = 5$)

5. Uniqueness of Factorization. In how far is the representation of a tree as a product of stalks unique? Put differently, in how far is the map $*$ from $\mathcal{K}$ onto $\mathcal{P}$ one-to-one? A partial answer is provided by the following

Proposition 5.1. *Suppose U is of valence $\geqq p$, where the integer p is not necessarily a prime. Suppose $K_1, \dots, K_m$ and $H_1, \dots, H_n$ are nonzero stalks with*

$$K_1 \cdots K_m = H_1 \cdots H_n. \tag{5.1}$$

Then unless $m \geqq p$, $n > p$ or $m > p$, $n \geqq p$, we have $m = n$ and $K_i = H_i$ $(i = 1, \dots, n)$ after suitable reordering. Put differently, we have $\underline{\underline{K}} = \underline{\underline{H}}$ with

$$\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m, \quad \underline{\underline{H}} = H_1 \oplus \cdots \oplus H_n. \tag{5.2}$$

In particular, products of fewer than p stalks have a unique factorization as a product of stalks. The bounds of the Proposition are best possible, as is seen from the following example. Let U be a tree of valence p (unique up to isomorphisms!). Let $t \in U_1$, so that $\lambda(t) = 1$, and let $t_1, \dots, t_p$ be the elements of level 2 having $t_i \to t$ $(i = 1, \dots, p)$. Then

$$K(t_1) \cdots K(t_p) = K(t)^{p+1}. \tag{5.3}$$

See Figure 2, where the underlying prime is $p = 5$.

For both sides of (5.3) consist of $\{0\}$ and the vertices $u \geqq t$ with $\lambda(u) \leqq p + 1$. Writing $\underline{\underline{K}} = K(t_1) \oplus \cdots \oplus K(t_p)$, $\underline{\underline{H}} = (p+1)K(t)$, we have $\sigma_{\underline{\underline{K}}}(u) = \sigma_{\underline{\underline{H}}}(u) = 0$ when $u \not\geqq t$, and

$$\sigma_{\underline{\underline{K}}}(u) = \begin{cases} p & \text{for} \quad u = t, \\ p+1 & \text{for} \quad u > t, \end{cases} \qquad \sigma_{\underline{\underline{H}}}(u) = p + 1 \quad \text{for } u \geqq t.$$

Proof of the Proposition. Let $\underline{\underline{K}}, \underline{\underline{H}}$ be given by (5.2). In view of Lemma 4.2, the relation (5.1) says that

$$\lambda(u) \leqq \sigma_{\underline{\underline{K}}}(u) \text{ precisely when } \lambda(u) \leqq \sigma_{\underline{\underline{H}}}(u). \tag{5.4}$$

We have to show that unless $m \geqq p$, $n > p$ or $m > p$, $n \geqq p$, we have identically

$$\sigma_{\underline{\underline{K}}}(u) = \sigma_{\underline{\underline{H}}}(u).$$

If not, choose u of minimal level with $\sigma_{\underline{\underline{K}}}(u) \neq \sigma_{\underline{\underline{H}}}(u)$, say $\sigma_{\underline{\underline{K}}}(u) < \sigma_{\underline{\underline{H}}}(u)$. By (5.4) we have

$$\sigma_{\underline{\underline{K}}}(u) < \sigma_{\underline{\underline{H}}}(u) < \lambda(u) \tag{5.5}$$

or

$$\lambda(u) \leqq \sigma_{\underline{\underline{K}}}(u) < \sigma_{\underline{\underline{H}}}(u). \tag{5.6}$$

In the first case, $u \notin H_1 \cup \cdots \cup H_n$, so that $\sigma_{\underline{\underline{H}}}(u^-) = \sigma_{\underline{\underline{H}}}(u)$, similarly $\sigma_{\underline{\underline{K}}}(u^-) = \sigma_{\underline{\underline{K}}}(u)$, contradicting the minimality of u. In the second case, suppose at first that at least p stalks K_i contain u. Then $p\lambda(u) \leqq \sigma_{\underline{\underline{K}}}(u) < \sigma_{\underline{\underline{H}}}(u)$. On the other hand, $\sigma_{\underline{\underline{K}}}(u) \leqq m\lambda(u)$, $\sigma_{\underline{\underline{H}}}(u) \leqq n\lambda(u)$, so that $m \geqq p$, $n > p$. We may therefore suppose that fewer than p stalks K_i contain u. Then there is a t with $t \to u$ which is constrained in no stalk K_i. Then every $s \geqq t$ has

$$\sigma_{\underline{\underline{K}}}(s) = \sigma_{\underline{\underline{K}}}(u) < \sigma_{\underline{\underline{H}}}(u).$$

But $\sigma_{\underline{\underline{H}}}(u) \geqq \sigma_{\underline{\underline{K}}}(u) + 1 \geqq \lambda(u) + 1 = \lambda(t)$ by (5.6). There is then an $s \geqq t$ with $\lambda(s) = \sigma_{\underline{\underline{H}}}(u)$; and this s has $\sigma_{\underline{\underline{H}}}(s) \geqq \sigma_{\underline{\underline{H}}}(u)$. Thus $\sigma_{\underline{\underline{K}}}(s) < \lambda(s) \leqq \sigma_{\underline{\underline{H}}}(s)$, contradicting (5.4).

Now suppose that U has infinite valence. Then the map $*$ onto $\mathcal{P}$ is injective, i.e., the presentation of a tree as a product of stalks (if possible at all) is unique. Also the presentation of elements of $\mathcal{P}'$ as $L^\ell K_1 \cdots K_m$ is unique. For if $L^\ell K_1 \cdots K_m = L^g H_1 \cdots H_n$ and if, say $\ell \leqq g$, we may cancel by L^ℓ, so that $K_1 \cdots K_m = L^{g-\ell} H_1 \cdots H_n$. Here the left hand side has only finitely many elements of level 1, hence so does the right hand side. Thus $\ell = g$ and $K_1 \cdots K_m = H_1 \cdots H_n$, so that again $m = n$ and $K_i = H_i$ $(i = 1, \ldots, n)$ after suitable ordering. □

6. Poincaré Series. Let $T \subseteq U$ be a tree, and $\lambda > 0$. We define $T(\lambda)$ to be the set of $u \in U$ such that $\alpha_T(u) \geqq \lambda$, but $\alpha_T(v) \geqq \lambda$ for no $v < u$. Then it is clear that the elements of $T(\lambda)$ are mutually incompatible. Since $\alpha_T(u^-) = \alpha_T(u)$ when $u \notin T$, it follows that $T(\lambda) \subseteq T$. When $t \in T_\lambda$, then because $\alpha_T(t) \geqq \lambda$, there is a $u \leqq t$ lying in $T(\lambda)$. But then $t \in F(u, \lambda)$; in fact $t \in F'(u, \lambda)$ where $F'(u, \lambda)$ consists of the "top" of $F(u, \lambda)$, i.e., its elements of level λ. Therefore

$$T_\lambda = \bigcup_{u \in T(\lambda)} F'(u, \lambda). \tag{6.1}$$

Since the elements of $T(\lambda)$ are incompatible, the union here is disjoint.

Suppose U is of finite valence p. Then $F'(u, \lambda)$ has cardinality $p^{\lambda - \lambda(u)}$. Therefore *the width of T at level* λ, defined as the cardinality $|T_\lambda|$ of T_λ, has

$$|T_\lambda| = \sum_{u \in T(\lambda)} p^{\lambda - \lambda(u)}. \tag{6.2}$$

The *Poincaré Series* of T is the formal series

$$\mathfrak{P}_T(Z) = \sum_{\lambda=0}^{\infty} |T_\lambda| Z^\lambda. \tag{6.3}$$

More generally, suppose $T \subseteq U$ is a tree such that $T(\lambda)$ is finite for every λ. Inspired by (6.2), we define polynomials

$$P_{T\lambda}(z) = \sum_{u \in T(\lambda)} z^{\lambda - \lambda(u)} \quad (\lambda \geqq 0), \tag{6.4}$$

and we define a General Poincaré Series

$$\mathfrak{P}_T(z, Z) = \sum_{\lambda=0}^{\infty} P_{T\lambda}(z) Z^\lambda; \tag{6.5}$$

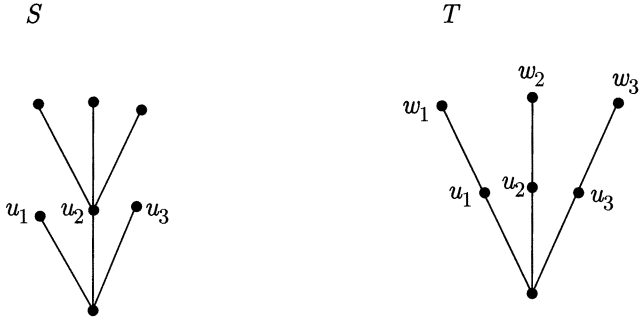

FIGURE 3. $(p = 3)$

this series lies in $\mathbb{Z}[z][[Z]]$. In the case when U is of finite valence p we have $\mathfrak{P}_T(Z) = \mathfrak{P}_T(p, Z)$. The example of Figure 3 with $p = 3$ shows that it may happen that $\mathfrak{P}_T(Z) = \mathfrak{P}_S(Z)$ yet $\mathfrak{P}_T(z, Z) \neq \mathfrak{P}_S(z, Z)$, so that the General Poincaré Series encodes more information than the ordinary Poincaré Series.

Here $\mathfrak{P}_S(Z) = \mathfrak{P}_T(Z) = 1 + 3Z + 3Z^2$. But $S(1) = T(1) = \{0\}$, $S(2) = \{u_2\}$, $T(2) = \{w_1, w_2, w_3\}$, so that $\mathfrak{P}_S(z, Z) = 1 + zZ + zZ^2$, $\mathfrak{P}_T(z, Z) = 1 + zZ + 3Z^2$. Note that both S, T are products of stalks.

Theorem 6.1. *Let $\underline{\underline{K}}$ be a stalk sum whose infinite stalks occur with multiplicities $c_1 > 0, \ldots, c_\ell > 0$, and let C be the set of distinct numbers among $c_1, \ldots, c_\ell$. Suppose*

$$T = \underline{\underline{K}}^* L^n \tag{6.6}$$

where $n \geqq 0$. Then the General Poincaré Series $\mathfrak{P}_T(z, Z)$ is rational, i.e., it lies in $\mathbb{Q}(z, Z)$, and its denominator divides

$$\prod_{c \in C} (1 - z^{c-1} Z^c).$$

Therefore when U is of valence p, the Poincaré Series $\mathfrak{P}_T(Z)$ is a rational function whose denominator divides

$$\prod_{c \in C} (1 - p^{c-1} Z^c).$$

Lemma 6.2. *Suppose $T = T_1 \cdots T_\ell$ and $S = T_1 \cup \cdots \cup T_\ell$ where $T_1, \ldots, T_\ell$ are trees. Then*

$$\alpha_T(u) = \alpha_T(u_S).$$

When $u \notin S$, or when $v \notin S$ for some $v \to u$, then

$$\alpha_T(u) = \sum_{i=1}^{\ell} \lambda(u_{T_i}). \tag{6.7}$$

Proof. Suppose $u \notin S$, so that $u \geqq v \to u_S$ with $v \notin S$. We have $u_S \in T$ and $\alpha_T(u_S) \geqq \sum_{i=1}^{\ell} \lambda((u_S)_{T_i}) = \sum_{i=1}^{\ell} \lambda(u_{T_i})$. On the other hand, when $t \geqq v$ we have $t \in T$ precisely when $\lambda(t) \leqq \sum_{i=1}^{\ell} \lambda(t_{T_i}) = \sum_{i=1}^{\ell} \lambda(u_{T_i})$. Therefore $\alpha_T(u) \leqq \alpha_T(u_S)$, and since the reversed inequality is trivial, our first assertion follows. Also (6.7) follows in this case.

When $u \in S$ but $v \to u$ with $v \notin S$, apply our arguments to v. We have $v_S = u$, therefore

$$\alpha_T(u) = \alpha_T(v_S) = \sum_{i=1}^{\ell} \lambda(v_{T_i}) = \sum_{i=1}^{\ell} \lambda(u_{T_i}).$$

□

Lemma 6.3. *Suppose $\underline{\underline{K}}$ is a sum of m stalks and T is given by (6.6). Then for every $u \in U$,*

$$\alpha_T(u) \leqq (\lambda(u) + 1)m + n.$$

Proof. Set $S = K_1 \cup \cdots \cup K_m \cup L^n$ where $K_1, \ldots, K_m$ are the summands of $\underline{\underline{K}}$. Since U is of valence $\geqq 2$, there is for every $u \in U$ a $u^{(1)} \to u$ such that $u^{(1)}$ lies in at most $m/2$ of these stalks $K_1, \ldots, K_m$. There is a $u^{(2)} \to u^{(1)}$ lying in at most $m/4$ of these stalks. Et cetera. Let ν be least such that $u^{(\nu)}$ lies in none of these m stalks. When u itself lies in none of the stalks, set $\nu = 0$, $u^{(0)} = u$. Pick $w \geqq u^{(\nu)}$ with $\lambda(w) > n$. Then $w \notin S$, so that by (6.7),

$$\alpha_T(w) = \sum_{i=1}^{m} \lambda(w_{K_i}) + \lambda(w_{L^n}).$$

Now at most $m/2$ of the stalks K_i reach $u^{(1)}$, at most $m/4$ reach $u^{(2)}$, and so on. Therefore

$$\sum_{i=1}^{m} \lambda(w_{K_i}) < m\lambda(u) + (m/2) + (m/4) + \cdots = (\lambda(u) + 1)m.$$

We obtain

$$\alpha_T(u) \leqq \alpha_T(w) < (\lambda(u) + 1)m + n.$$

□

Proof of Theorem 6.1. Suppose

$$\underline{\underline{K}} = c_1 K_1 \oplus \cdots \oplus c_\ell K_\ell \oplus d_1 H_1 \oplus \cdots \oplus d_q H_q$$

where $K_1, \ldots, K_\ell$ are distinct infinite stalks and $H_1, \ldots, H_q$ are distinct finite stalks. The total number of stalks counting multiplicities is $m = c_1 + \cdots + c_\ell + d_1 + \cdots + d_q$. Choose λ_1 so large that $\lambda_1 \geqq \lambda(H_i)$ for $1 \leqq i \leqq q$ and $\lambda_1 \geqq \lambda(K_i \cap K_j)$ for $1 \leqq i < j \leqq \ell$. Set

$$\lambda_2 = (\lambda_1 + 1)m + n.$$

Now when $u \in T(\lambda)$ where $\lambda > \lambda_2$, then $\lambda(u) > \lambda_1$ by Lemma 6.3, and u is contained in exactly one of $K_1, \ldots, K_\ell$, but not in $H_1, \ldots, H_q$ or L^n.

Let $u_{i,\mu}$ be the element of K_i of level μ where $\mu > \lambda_1$. Then since there is a $v \to u_{i,\mu}$ with v not in any of the K_i, the H_i, or L^n, there is a formula like (6.7) for $\alpha_T(u_{i,\mu})$, and similarly for $u_{i,\mu+1}$. Since $u_{i,\mu+1}$ lies only on the summand K_i of $\underline{\underline{K}}$, and since this summand occurs with multiplicity c_i, we have $\alpha_T(u_{i,\mu+1}) - \alpha_T(u_{i,\mu}) = c_i$. Therefore

$$\alpha_T(u_{i,\mu}) = \mu c_i + g_i \qquad (\mu > \lambda_1),$$

where g_i is a nonnegative integer. Given $\lambda > \lambda_2$, we have $\alpha_T(u_{i,\mu}) \geqq \lambda$ precisely when $\mu c_i + g_i \geqq \lambda$, and the smallest μ with this property is $\mu(i, \lambda) = \lceil (\lambda - g_i)/c_i \rceil$

where, for any real number x, $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. We may conclude that

$$T(\lambda) = \{u_{1,\mu(1,\lambda)}, \dots , u_{\ell,\mu(\ell,\lambda)}\}.$$

By (6.4)

$$P_{T\lambda}(z) = \sum_{i=1}^{\ell} z^{\lambda-\mu(i,\lambda)} \quad (\lambda > \lambda_2). \tag{6.8}$$

To finish the proof it will suffice to show that for each i, $1 \leqq i \leqq \ell$, the series

$$\sum_{\lambda>\lambda_2} z^{\lambda-\mu(i,\lambda)} Z^{\lambda}$$

is rational with denominator $1 - z^{c_i-1}Z^{c_i}$. We may as well do this with the sum over $\lambda > \lambda_2$ replaced by the sum over $\lambda > g_i$. Writing $\lambda = g_i + \rho + c_i\nu$ with $1 \leqq \rho \leqq c_i$ and $\nu \geqq 0$, we have $\mu(i, \lambda) = \nu + 1$, and our sum becomes

$$\sum_{\rho=1}^{c_i}\sum_{\nu=0}^{\infty} z^{g_i+\rho+c_i\nu-\nu-1} Z^{g_i+\rho+c_i\nu}$$

$$= z^{g_i} Z^{g_i+1}(1 + zZ + \dots + (zZ)^{c_i-1})(1 - z^{c_i-1}Z^{c_i})^{-1}. \quad \square$$

Given a natural number q written as $q = \sum_{i=0}^{k} c_i p^i$ with $p > 1$ and with digits c_i in $\{0, 1, \dots , p - 1\}$, set

$$\mathrm{dig}_p(q) = \sum_{i=0}^{k} c_i.$$

Proposition 6.4. *When U is of finite valence p, $\underline{\underline{K}}$ is a stalk sum with m summands and T is the tree given by (6.6), then*

$$dig_p|T_\lambda| \leqq m \qquad (\lambda = 0, 1, \dots).$$

Proof. We have $T = SL^n$ with $S = \underline{\underline{K}}^*$. It is easily seen that $\alpha_T(u) = \alpha_S(u) + n$. Therefore by Lemma 6.2, $\alpha_T(u) = \alpha_T(u^-)$ unless $u \in K_1 \cup \dots \cup K_m$. Therefore $T(\lambda) \subseteq K_1 \cup \dots \cup K_m$, so that $|T(\lambda)| \leqq m$. So the sum in (6.2) has at most m summands. The assertion now follows from the fact that if a number $q = \sum_{i=0}^{k} c_i p^i$ with $c_i \geqq 0$, then $\mathrm{dig}_p(q) \leqq \sum_{i=0}^{k} c_i$. $\square$

For later use we will discuss an example. Suppose $0 < a_1 < \dots < a_{m-1}$ are given integers. Let K_m be an infinite stalk and K_i for $1 \leqq i < m$ an infinite stalk with $\lambda(K_i \cap K_m) = a_i$. This is possible since U is of valence greater than 1. Set $\underline{\underline{K}} = K_1 \oplus \dots \oplus K_m$ and let T be the tree (6.6). When $u_{i,\mu} \in K_i$ of level $\mu > a_{m-1} + n$, then by Lemma 6.2,

$$\begin{aligned}\alpha_T(u_{i,\mu}) &= \sum_{j=1}^{m} \lambda((u_{i,\mu})_{K_j}) + \lambda((u_{i,\mu})_{L^n}) \\ &= a_1 + \dots + a_{i-1} + (m - i)a_i + \mu + n = \mu + b_i\end{aligned}$$

with $b_i = a_1 + \dots + a_{i-1} + (m - i)a_i + n$ (a term involving the non-existent a_m only *appears* for b_m). We have $b_1 < \dots < b_{m-1} = b_m$. Since each K_i occurs with multiplicity $c_i = 1$, the quantity $\mu(i,\lambda)$ of the proof of Theorem 6.1 has $\mu(i,\lambda) = \lambda - b_i$, so that $P_\lambda = P_{T\lambda}$ is given by

$$P_\lambda(z) = \sum_{i=1}^{m} z^{b_i} \tag{6.9}$$

for large λ.

Now suppose that

$$0 \leqq n < a_1 < \cdots < a_{m-1}. \tag{6.10}$$

It is then easily seen that $T(\lambda) = \{0\}$ when $\lambda \leqq n$, and $T(n+1) = \{u_1\}$, where u_1 of level 1 is on $K_1 \cap \cdots \cap K_m$. We used the fact that U is of valence > 1. Therefore $P_\lambda = z^\lambda$ when $\lambda \leqq n$, but $P_{n+1}(z) = z^n$. Therefore the number n can be detected from the Poincaré series $\mathfrak{P}_T(z,Z)$, and then by (6.9), also $b_1, \ldots, b_{m-1}$, hence also $a_1, \ldots, a_{m-1}$ can be detected. Put differently:

For different values of n, $a_1, \ldots, a_{m-1}$ with (6.10) *we obtain different General Poincaré Series $\mathfrak{P}_T(z,Z)$. When U is of finite valence $p > 1$, then in fact we obtain different Poincaré Series $\mathfrak{P}_T(Z)$.*

7. Discriminants and Resultants. Let $\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m$ be a stalk sum. Given such $\underline{\underline{K}}$, define its discriminant to be

$$\delta(\underline{\underline{K}}) = \sum_{\substack{1 \leqq i,j \leqq m \\ i \neq j}} \lambda(K_i \cap K_j),$$

and given another stalk sum $\underline{\underline{H}} = H_1 \oplus \cdots \oplus H_n$, define their resultant to be

$$\rho(\underline{\underline{K}}, \underline{\underline{H}}) = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda(K_i \cap H_j).$$

The discriminant is zero when $\underline{\underline{K}} = \underline{\underline{0}}$ or $m = 1$, and the resultant is zero when $\underline{\underline{K}} = \underline{\underline{0}}$ or $\underline{\underline{H}} = \underline{\underline{0}}$. The discriminant is finite precisely when every infinite stalk appears in the stalk sum $\underline{\underline{K}}$ at most once, and the resultant is finite precisely when $\underline{\underline{K}}, \underline{\underline{H}}$ have no infinite stalk as a common summand. Note that the discriminant (when finite) is always even. We have

$$\delta(\underline{\underline{K}} \oplus \underline{\underline{H}}) = \delta(\underline{\underline{K}}) + \delta(\underline{\underline{H}}) + 2\rho(\underline{\underline{K}}, \underline{\underline{H}}), \tag{7.1}$$

$$\rho(\underline{\underline{K}} \oplus \underline{\underline{K}}', \underline{\underline{H}}) = \rho(\underline{\underline{K}}, \underline{\underline{H}}) + \rho(\underline{\underline{K}}', \underline{\underline{H}}). \tag{7.2}$$

In the Introduction we defined isomorphisms of trees. Now two stalk sums $\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m$ and $\underline{\underline{H}} = H_1 \oplus \cdots \oplus H_m$ will be called isomorphic if after suitable ordering of the summands there is a bijection $\Psi : \bigcup_{i=1}^m K_i \to \bigcup_{i=1}^m H_i$ whose restriction to K_i is an isomorphism onto H_i $(i = 1, \ldots, m)$. Clearly when the universal tree has some fixed valence and $\underline{\underline{K}}, \underline{\underline{H}}$ are isomorphic, we have an isomorphism of trees: $\underline{\underline{K}}^* \sim \underline{\underline{H}}^*$.

A stalk sum will be called a *polynomial sum* (the name being justified in section 11) if it is a sum of infinite stalks and of couples as defined in section 4. When $\underline{\underline{K}}, \underline{\underline{H}}$ are isomorphic and one is a polynomial sum, then so is the other.

Theorem 7.1. *Given $m \geqq 1$, $\delta \geqq 1$, the number of isomorphism classes of polynomial sums with m nonzero summands and discriminant $\leqq \delta$ is*

$$\leqq c_3(m)\delta^{m-1}.$$

Except for the determination of $c_3(m)$ this bound is best possible.

Proof. Let

$$\underline{\underline{K}} = H_1 \oplus \cdots \oplus H_p \oplus J_1 \oplus \cdots \oplus J_q \oplus \underline{\underline{K}}_1 \oplus \cdots \oplus \underline{\underline{K}}_k$$

be a polynomial sum where H_i $(1 \leqq i \leqq p)$ is an infinite stalk, J_i $(1 \leqq i \leqq q)$ is a stalk of level 1 (i.e., a degenerate couple) and $\underline{\underline{K}}_i$ $(1 \leqq i \leqq k)$ is a nondegenerate couple, say $\underline{\underline{K}}_i = \ell_i K(a_i) + n_i K(a_i^-)$. Set

$$H_{p+i} = J_i \ (1 \leqq i \leqq q), \quad H_{p+q+i} = K(a_i) \ (1 \leqq i \leqq k).$$

The isomorphism class of $\underline{\underline{K}}$ depends only on p, q, k, on ℓ_i, n_i $(1 \leqq i \leqq k)$ and on the isomorphism class of

$$\underline{\underline{H}} = H_1 \oplus \cdots \oplus H_p \oplus H_{p+1} \oplus \cdots \oplus H_{p+q} \oplus H_{p+q+1} \oplus \cdots \oplus H_{p+q+k}.$$

The isomorphism class of a single stalk H depends only on $\lambda(H)$. When $H_1, \ldots, H_{i-1}$ are *given*, the isomorphism class of $H_1 \oplus \cdots \oplus H_{i-1} \oplus H_i$ depends only on $\lambda(H_i)$ and on $H_i \cap (H_1 \cup \cdots \cup H_{i-1})$. This intersection is in fact a stalk of the type $H_i \cap H_{j_i}$ $(1 \leqq j_i \leqq i-1)$ and is determined once we know j_i and the level $\lambda(H_i \cap H_{j_i})$. Therefore the class of $\underline{\underline{H}}$ depends only on j_i $(2 \leqq i \leqq p+q+k)$, on

(7.3) $$\lambda(H_i \cap H_{j_i}) \quad (2 \leqq i \leqq p+q+k)$$

and on $\lambda(H_{p+q+i}) = \lambda(K(a_i))$ $(1 \leqq i \leqq k)$. This last level will be written as

(7.4) $$\lambda(K(a_i) \cap K(a_i)) \quad \text{when } 1 \leqq i \leqq k \text{ and } \ell_i \geqq 2,$$

and as 1 plus

(7.5) $$\lambda(K(a_i) \cap K(a_i^-)) \quad \text{when } 1 \leqq i \leqq k \text{ and } \ell_i = 1.$$

When $\underline{\underline{K}}$ has m nonzero summands we have $p + q + 2k \leqq m$, since each $\underline{\underline{K}}_i$ has at least 2 summands. The number of choices for p, q, k, the ℓ_i, n_i $(1 \leqq i \leqq k)$ and j_i $(2 \leqq i \leqq p+q+k)$ is under a bound $c_4(m)$. The quantities (7.3) and (7.4) (when $\ell_i \geqq 2$) and (7.5) (when $\ell_i = 1$, so that $n_i \geqq 1$) are summands of $\delta(\underline{\underline{K}})$, hence are $\leqq \delta$. There are not more than $\delta + 1$ choices for each of them. The number of quantities (7.3), (7.4) and (7.5) is $(p+q+k-1)+k \leqq m-1$. Therefore the number of isomorphism classes is $\leqq c_4(m)(\delta+1)^{m-1} \leqq c_3(m)\delta^{m-1}$.

To show that our estimate is best possible, recall that the stalk sums $\underline{\underline{K}}$ constructed at the end of section 6 have distinct Poincaré series and therefore are non-isomorphic for different values of the parameters $0 < a_1 < \cdots < a_{m-1}$. We have

$$\delta(\underline{\underline{K}}) = 2(a_{m-1} + 2a_{m-2} + \cdots + (m-1)a_1).$$

For large δ, the number of possibilities for $a_1, \ldots, a_{m-1}$ with $\delta(\underline{\underline{K}}) \leqq \delta$ is at least $c_5(m)\delta^{m-1}$. □

We are going to finish this section with three lemmas which will be useful later on.

Lemma 7.2. *Let $\underline{\underline{K}}, \underline{\underline{H}}$ be stalk sums. Then*

$$\delta(\underline{\underline{K}}) = \sum_{u \neq \{0\}} \tau_{\underline{\underline{K}}}(u)(\tau_{\underline{\underline{K}}}(u) - 1),$$

$$\rho(\underline{\underline{K}}, \underline{\underline{H}}) = \sum_{u \neq \{0\}} \tau_{\underline{\underline{K}}}(u)\tau_{\underline{\underline{H}}}(u).$$

Proof. When $\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m$,

$$\delta(\underline{\underline{K}}) = \sum_{i \neq j} \lambda(K_i \cap K_j) = \sum_{i \neq j} \sum_{\substack{u \neq \{0\} \\ u \in K_i \cap K_j}} 1$$

$$= \sum_{u \neq \{0\}} \sum_{\substack{i \neq j \\ u \in K_i \cap K_j}} 1 = \sum_{u \neq \{0\}} \tau_{\underline{\underline{K}}}(u)(\tau_{\underline{\underline{K}}}(u) - 1).$$

The second assertion is shown similarly. □

Lemma 7.3. *Let $\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m \in \mathcal{K}_0$. Then for every vertex u,*

$$\sigma_{\underline{\underline{K}}}(u) - \lambda(u) \leqq \frac{1}{2}\, \delta(\underline{\underline{K}}).$$

Proof. Consider the vertices $u_{K_1}, \ldots, u_{K_m}$ (following the definition of u_T in section 1). Set $\lambda_j = \lambda(u_{K_j})$ $(j = 1, \ldots, m)$ and suppose that $\lambda_1 \geqq \cdots \geqq \lambda_m$. We observe that

$$\sum_{\{0\} < t \leqq u} \tau_{\underline{\underline{K}}}(t) = \sum_{j=1}^{m} \sum_{\substack{\{0\} < t \leqq u \\ t \in K_j}} 1 = \sum_{j=1}^{m} \lambda_j,$$

$$\sum_{\{0\} < t \leqq u} \tau^2_{\underline{\underline{K}}}(t) = \sum_{i=1}^{m}\sum_{j=1}^{m} \sum_{\substack{\{0\} < t \leqq u \\ t \in K_i \cap K_j}} 1 = \sum_{i=1}^{m}\sum_{j=1}^{m} \min(\lambda_i, \lambda_j),$$

so that by Lemma 7.2,

$$\delta(\underline{\underline{K}}) \geqq \sum_{\{0\} < t \leqq u} \tau_{\underline{\underline{K}}}(t)(\tau_{\underline{\underline{K}}}(t) - 1) = \sum_{i \neq j} \min(\lambda_i, \lambda_j)$$

$$= 2(\lambda_2 + 2\lambda_3 + \cdots + (m-1)\lambda_m) \geqq 2(\lambda_2 + \cdots + \lambda_m)$$

by our ordering of $\lambda_2, \ldots, \lambda_m$. On the other hand each $\lambda(u_{K_i}) \leqq \lambda(u)$, and therefore

$$\sigma_{\underline{\underline{K}}}(u) - \lambda(u) = \sum_{i=1}^{m} \lambda(u_{K_i}) - \lambda(u) \leqq \lambda_2 + \cdots + \lambda_m \leqq \frac{1}{2}\delta(\underline{\underline{K}}).$$

□

Lemma 7.4. *Again let $\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m$. Let $u_1, \ldots, u_\ell$ be mutually incompatible vertices of $K_1 \cup \cdots \cup K_m$. Then*

$$\sum_{i=1}^{\ell} (\sigma_{\underline{\underline{K}}}(u_i) - \lambda(u_i)) \leqq \delta(\underline{\underline{K}}).$$

Proof. Since u_i, u_j with $i \neq j$ are incompatible, they cannot lie in the same stalk. We may therefore suppose without loss of generality that $u_i \in K_i$ $(i = 1, \dots, \ell)$. We claim that

$$\sigma_{\underline{\underline{K}}}(u_i) - \lambda(u_i) \leqq \sum_{j \neq i} \lambda(K_i \cap K_j) \quad (i = 1, \dots, \ell),$$

and this will clearly imply the lemma. But

$$\begin{aligned} \sigma_{\underline{\underline{K}}}(u_i) &= \sum_{\{0\} < t \leqq u_i} \tau_{\underline{\underline{K}}}(t) = \sum_{\{0\} < t \leqq u_i} \sum_{K_j \ni t} 1 \\ &= \sum_{j=1}^{m} \sum_{\substack{\{0\} < t \leqq u_i \\ t \in K_j}} 1 \leqq \lambda(u_i) + \sum_{j \neq i} \lambda(K_i \cap K_j). \end{aligned}$$

□

8. The Width. We called the cardinality $|T_\lambda|$ the width of T at level λ. The *total width* $w(T)$ is the maximum of $|T_\lambda|$ over $\lambda = 0, 1, \dots$. These widths need not be finite.

Theorem 8.1. *Let U be a universal tree of finite valence p, where p is not necessarily a prime. Let $\underline{\underline{K}}$ be a stalk sum with $m \geqq 1$ summands and with finite discriminant δ. Then*

$$w(\underline{\underline{K}}^*) \leqq 2p^{\delta/2} + m - 2. \tag{8.1}$$

This estimate is best possible for every $p \geqq m$ and every even δ.

In the context of polynomials, the Theorem is due to Stewart [2, (38)]. In fact all the assertions of Stewart's Theorem 2 could be derived in the present context.

Proof. Set $T = \underline{\underline{K}}^*$ and $\sigma = \sigma_{\underline{\underline{K}}}$. Then by Lemma 4.1, $u \in T$ precisely if $\lambda(u) \leqq \sigma(u)$. Similarly we have $u \in T$ precisely when $\lambda(u) \leqq \alpha_T(u)$. In complete analogy to our procedure in section 6, we define $TT(\lambda)$ to be the set of $u \in U$ such that $\sigma(u) \geqq \lambda$ but $\sigma(v) \geqq \lambda$ for no $v < u$. In analogy to (6.1) we have

$$T_\lambda = \bigcup_{u \in TT(\lambda)} F'(u, \lambda) \tag{8.2}$$

where the union is disjoint. From this we obtain

$$|T_\lambda| = \sum_{u \in TT(\lambda)} p^{\lambda - \lambda(u)}, \tag{8.3}$$

which is analogous to (6.2). Thus if $TT(\lambda) = \{u_1, \dots, u_\ell\}$ and $\omega_i = \lambda - \lambda(u_i)$ $(i = 1, \dots, \ell)$, we have

$$|T_\lambda| = \sum_{i=1}^{\ell} p^{\omega_i}. \tag{8.4}$$

We observe that in view of $\omega_i \leqq \sigma(u_i) - \lambda(u_i)$,

(i) $\omega_i \leqq \delta/2$ by Lemma 7.3,
(ii) $\sum_{i=1}^{\ell} \omega_i \leqq \delta$ by Lemma 7.4,
(iii) $\ell \leqq m$, since $u_1, \dots, u_\ell$ are incompatible elements of $K_1 \cup \cdots \cup K_m$.

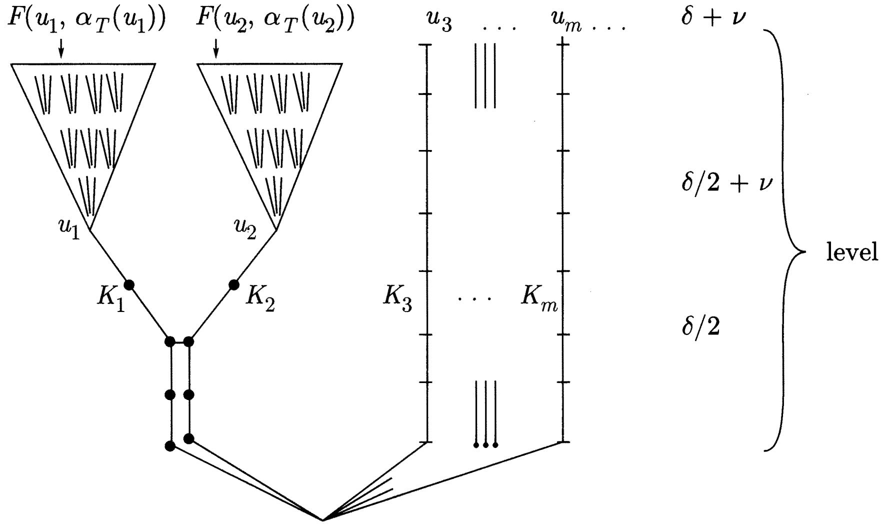

FIGURE 4

Since $p^{\omega+1} + p^{\psi-1} > p^{\omega} + p^{\psi}$ when $\omega \geq \psi$, it is clear that if, say, $\omega_1 \geqq \cdots \geqq \omega_\ell$, the maximum of the sum in (8.4) is taken when $\ell = m$ and $\omega_1 = \omega_2 = \delta/2$ and $\omega_3 = \cdots = \omega_\ell = 0$, so that indeed $|T_\lambda| \leqq 2p^{\delta/2} + m - 2$ for each λ, whence (8.1).

That (8.1) is best possible is seen from the example illustrated in Figure 4.

We have infinite stalks $K_1, \ldots, K_m$ with $\lambda(K_1 \cap K_2) = \delta/2$ and $\lambda(K_i \cap K_j) = 0$ when $i \neq j$ and $\{i,j\} \neq \{1,2\}$. For $\lambda = \delta + \nu$ the set $TT(\lambda)$ (with $T = K_1 \cdots K_m$) is given by $TT(\lambda) = \{u_1, \ldots, u_m\}$, where u_1, u_2 on K_1, K_2 have level $\delta/2 + \nu$ while $u_3, \ldots, u_m$ on $K_3, \ldots, K_m$ have level $\delta + \nu$. Therefore $\omega_1 = \omega_2 = \delta/2$ and $\omega_3 = \cdots = \omega_m = 0$. $\square$

A further theorem on widths will be given in section 15.

II. Polynomial Trees

9. Polynomial Trees in a Discrete Valuation Ring. The diagram of the Introduction could have been written as

$$\{0\} = \mathbb{Z}_p/p^0\mathbb{Z}_p \xleftarrow{\Phi_1} \mathbb{Z}_p/p\mathbb{Z}_p \xleftarrow{\Phi_2} \mathbb{Z}_p/p^2\mathbb{Z}_p \longleftarrow \cdots .$$

More generally, let R be a discrete valuation ring with maximal ideal $\mathfrak{p}$, and consider

$$\{0\} = R/\mathfrak{p}^0 R \xleftarrow{\Phi_1} R/\mathfrak{p}R \xleftarrow{\Phi_2} R/\mathfrak{p}^2 R \longleftarrow \cdots$$

where the Φ_λ are the natural homomorphisms. We obtain a tree $U = U(R)$ whose vertices are the elements of $R/\mathfrak{p}^\lambda R$ for $\lambda = 0, 1, \ldots,$ and the directed edges are $u \to v$ where $\Phi_\lambda(u) = v$ for some λ. Given $x \in R$, we write x_λ for the image of x under the natural homomorphism $R \to R/\mathfrak{p}^\lambda R$.

When $f \in R[X]$ we define $T(f)$ to consist of elements x_λ where $x \in R$, $\lambda \geqq 0$ and $(f(x))_\lambda = 0_\lambda$, i.e., $f(x) \equiv 0 \pmod{\mathfrak{p}^\lambda}$. Then $T(f)$ is a subtree of $U(R)$. Again (2) holds, i.e.,

$$T(fg) = T(f)T(g) \tag{9.1}$$

for polynomials f, g. Now if I is an ideal in $R[X]$, we let $T(I)$ be the intersection of the trees $T(f)$ with $f \in I$. Then $T(I)$ is again a tree. Clearly

$$T(I+J) = T(I) \cap T(J). \tag{9.2}$$

On the other hand, if $I = (f_1) + \cdots + (f_m)$ (a sum of principal ideals) and $J = (g_1) + \cdots + (g_\ell)$, so that $IJ = \sum_i \sum_j (f_i)(g_j)$, then

$$\begin{aligned} T(IJ) &= \bigcap_i \bigcap_j T(f_i g_j) = \bigcap_i \bigcap_j T(f_i)T(g_j) \\ &= (\bigcap_i T(f_i))(\bigcap_j T(g_j)) = T(I)T(J) \end{aligned}$$

by repeated application of the distributive law (1.3). Thus

$$T(IJ) = T(I)T(J). \tag{9.3}$$

The ideals I of $R[X]$ form a pseudolattice $\mathcal{J}$ under product $(\cdot)$ and sum $(+)$. In view of (9.2), (9.3), the map $I \to T(I)$ gives a homomorphism from $\mathcal{J}$ into the pseudolattice $\mathcal{T}$ of subtrees of $U(R)$, with $\mathcal{T}$ having the binary operations of product $(\cdot)$ and intersection $(\cap)$.

We will be most interested in two types of discrete valuation rings R. The first is $R = \mathbb{Z}_p$, already discussed in the Introduction. The other is the power series ring $R = k[[\mathcal{X}]]$ where k is a field. In this case $\mathfrak{p} = (\mathcal{X})$ and the residue class field is k. When k is of finite cardinality q (a prime power), then $U(R)$ is of finite valence q. When k is infinite, $U(R)$ is of infinite valence. $R/\mathfrak{p}^\lambda$ consists of polynomials $t(\mathcal{X}) \in k[\mathcal{X}]$ modulo $\mathfrak{p}^\lambda$, hence is a vector space of dimension λ over k. Its elements are uniquely represented by polynomials of degree $< \lambda$. Given $f(X) \in R[X] = k[[\mathcal{X}]][X]$, say $f(X) = f(\mathcal{X}, X)$, the elements of $T(f)$ of level λ are represented by polynomials $t(\mathcal{X})$ of degree $< \lambda$ with

$$f(\mathcal{X}, t(\mathcal{X})) \equiv 0 \pmod{\mathcal{X}^\lambda}.$$

10. A Generalized Theorem B. Let $|\cdot|$ be the absolute value on R given by $|x| = 2^{-\mathrm{ord}(x)}$ where ord is the valuation on R. *We will suppose from now on that R is complete under this absolute value*, i.e., that every Cauchy sequence has a limit in R. The following theorem is a generalization of Theorem B.

Theorem 10.1. *The set of trees $T(I)$ where I runs through the ideals of $R[X]$ is identical with $\mathcal{P}'$, i.e., the set of trees consisting of U and of products $L^n K_1 \cdots K_m$ where $n \geqq 0$ and $K_1, \ldots, K_m$ are stalks.*

In this section we will prove the easy half of this theorem, namely that U and products $L^n K_1 \cdots K_m$ are of the type $T(I)$. To begin with, $U = T(0)$ and $L^0 = T(1)$ where $0, 1$ stand for the constant polynomials equal to $0, 1$. By (9.3) it will now suffice to show that L and any stalk K is of the type $T(I)$. Now when π is a generator of the ideal $\mathfrak{p}$, then the constant polynomial $f(X) = \pi$ clearly has $T(\pi) = L$. Hence L is indeed of the type $T(I)$.

By our hypothesis that R be complete there is a 1–1 correspondence between elements y of R and infinite stalks, such that y corresponds to the stalk K with vertices y_λ ($\lambda \in \mathbb{Z}_{\geqq 0}$). The polynomial $f(X) = X - y$ has $(f(x))_\lambda = x_\lambda - y_\lambda = 0$ in $R/\mathfrak{p}^\lambda$ precisely when $x_\lambda = y_\lambda$, and therefore $T(f) = K$. Thus every infinite stalk is of the type $T(I)$ with a principal ideal I. Every finite stalk K may be expressed as $K = K_1 \cap K_2$ where K_1, K_2 are infinite stalks. (Here we used that $U(R)$ is of valence $\geqq 2$.) So if $K_1 = T(f_1)$, $K_2 = T(f_2)$ and $I = (f_1) + (f_2)$, we have

$$K = K_1 \cap K_2 = T(f_1) \cap T(f_2) = T(I)$$

by (9.2).

11. Proof of Theorem B. We will prove Theorem 10.1. It remains for us to show that every nonzero ideal I has $T(I)$ of the type $L^n K_1 \cdots K_m$. Since $T(I + J) = T(I) \cap T(J)$ and since $\mathcal{P}'$ is closed under intersection by Theorem 4.4, it will suffice to show that $T(f)$ is of the type $L^n K_1 \cdots K_m$ for every polynomial $f \neq 0$ in $R[X]$. Such a polynomial may be written as $f = \pi^n f_0$ with $n \geqq 0$ and f_0 a *primitive* polynomial, i.e,. a polynomial whose coefficients are not all divisible by π. Since $T(\pi) = L$, we may concentrate on primitive polynomials.

Proposition 11.1. *When f is a primitive polynomial, there is a polynomial sum $\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m$ such that*

$$T(f) = \underline{\underline{K}}^* = K_1 \cdots K_m. \tag{11.1}$$

When R has the property that there are irreducible polynomials of arbitrary degree in $\bar{F}[X]$ where $\bar{F}$ is the residue class field (e.g., when this field is finite, e.g., when $R = \mathbb{Z}_p$), then conversely for every polynomial sum $\underline{\underline{K}}$ there is a primitive polynomial f with (11.1).

The concept of a polynomial sum had been introduced in section 4; our present proposition justifies the terminology. A primitive polynomial may be written as a product of primitive polynomials which are irreducible over the quotient field F of R. Hence it will be enough to establish

Proposition 11.2. *When f is a primitive irreducible polynomial, then $T(f)$ is L^0, or an infinite stalk, or $\underline{\underline{K}}^*$ where $\underline{\underline{K}}$ is a couple.*

Conversely, there are primitive irreducible polynomials f with $T(f) = L^0$ or $T(f) = K$ where K is a given infinite stalk. When R has the property enunciated in Proposition 11.1 and when $\underline{\underline{K}}$ is a couple, there are primitive irreducible polynomials f with $T(f) = \underline{\underline{K}}^*$.

Proof. A primitive polynomial f of degree zero is a constant not divisible by π, and then $T(f) = L^0$. When f is primitive and irreducible of degree $d > 0$, then f factors as $c(X - \xi_1) \cdots (X - \xi_d)$ in its splitting field N. The absolute value $|\ |$ on F can uniquely be extended to N, and then $|\xi_1| = \cdots = |\xi_d|$ (see [1, Chapter XII, Proposition 2.5]).

Assume at first that this common absolute value is > 1. Then since $|x| \leqq 1$ for $x \in R$, we have $|x - \xi_i| = |\xi_i|$ and $|f(x)| = |c||\xi_1| \cdots |\xi_d| = |c'|$, where c' is the constant term of f. Since f is primitive, it may be deduced that $|c'| = 1$, so that $|f(x)| = 1$ and $(f(x))_\lambda \neq 0_\lambda$ for $\lambda > 0$. Therefore $T(f) = L^0$.

Next suppose that the common absolute value is $\leqq 1$. Let us first consider the case when f is of degree $d = 1$. We have $f(X) = c(X - y)$ with $y \in R$, $|c| = 1$. Then $T(f) = T(X - y)$ is the infinite stalk K with vertices y_λ where $\lambda \in \mathbb{Z}_{\geqq 0}$. Suppose, then, that $d > 1$. Again we write $f(X) = c(X - \xi_1) \cdots (X - \xi_d)$, where $\xi_1, \ldots, \xi_d$ lie in the splitting field N of f over F, with F the quotient field of R. We have $|c| = 1$ by primitivity. Let F' be the field $F' = F(\xi)$ with $\xi = \xi_1$. Then $[F' : F] = d = \underline{e}\underline{f}$, where $\underline{e}$ is the ramification index and $\underline{f}$ the degree of the residue class field extension.

Choose $y \in R$ with $|y-\xi|$ minimal — this is possible by the completeness of R. We have

$$(11.2) \qquad |x-\xi| \geqq |y-\xi|$$

for every $x \in R$. Say

$$(11.3) \qquad |y-\xi| = 2^{-h/\underline{e}}$$

with h a nonnegative integer. Pick $\lambda \in \mathbb{Z}$ with

$$(11.4) \qquad \lambda - 1 < h/\underline{e} \leqq \lambda,$$

and let K be the stalk $K = K(y_\lambda)$.

Again for $x \in R$, let x_μ be its image in $R/\mathfrak{p}^\mu$. Let ν be largest integer with $x_\nu = y_\nu$; then $|x-y| = 2^{-\nu}$. When $\nu < \lambda$, then $|x-y| > |y-\xi|$ and $|x-\xi| = 2^{-\nu}$. Since the absolute value is the same for conjugates over F, we have $|x-\xi_i| = 2^{-\nu}$ $(i = 1, \ldots, d)$ and $|f(x)| = 2^{-d\nu}$. Therefore $x_\mu \in T(f)$ precisely when $\mu \leqq \nu d$. Therefore when $0 < \nu < \lambda$, the fan $F(y_\nu, d\nu) \subseteq T(f)$. When $\nu \geqq \lambda$, then $|x-y| \leqq 2^{-\lambda} \leqq 2^{-h/\underline{e}}$; hence $|x-\xi| \leqq 2^{-h/\underline{e}}$, so that $|x-\xi| = 2^{-h/\underline{e}}$ by (11.2). Therefore $|f(x)| = 2^{-dh/\underline{e}} = 2^{-h\underline{f}}$. Thus $x_\mu \in T(f)$ precisely when $\mu \leqq h\underline{f}$. In particular, the fan $F(y_\lambda, h\underline{f}) \subseteq T(f)$. It is easily checked that

$$(11.5) \qquad h\underline{f} = (\lambda-1)d + \ell \quad \text{with} \quad 1 \leqq \ell \leqq d.$$

We may conclude that $T(f)$ is the union of the fans

$$F(y_\nu, d\nu) \qquad (0 \leqq \nu < \lambda)$$

and

$$F(y_\lambda, (\lambda-1)d + \ell).$$

When $\lambda = 0$, then $h = 0$, $\ell = d$ and $T(f) = L^0$. When $\lambda > 0$, the discussion at the end of section 4 shows that $T(f) = \underline{\underline{K}}^*$, where $\underline{\underline{K}}$ is the couple $\underline{\underline{K}} = \ell K(y_\lambda) + nK(y_\lambda^-)$ with $n = d - \ell$. When $\lambda = \ell = 1$, then $\underline{\underline{K}} = K(y_1)$ is degenerate, i.e., a single stalk of level 1.

Note that when $h/\underline{e} \in \mathbb{Z}$, so that $\lambda = h/\underline{e}$, we have $h\underline{f} = \lambda d$; therefore $\ell = d$, and $T(f)$ is $K(y_\lambda)^d$. In general, we should ideally represent $T(f)$ as the d-th power of a stalk of level $h/\underline{e}$, but this is not possible when $h/\underline{e} \notin \mathbb{Z}$. However, $T(f)$ is $K(y_\lambda)^\ell K(y_{\lambda-1})^n$, and here the "mean level" of the factors is

$$d^{-1}(\ell\lambda + n(\lambda-1)) = d^{-1}((\lambda-1)d + \ell) = d^{-1}h\underline{f} = h/\underline{e}$$

by (11.5).

For the converse we may clearly restrict ourselves to the assertion on couples. Say $\underline{\underline{K}} = \ell K(y_\lambda) + nK(y_\lambda^-)$ where $y \in R$. When $\underline{\underline{K}}$ is degenerate we set $n = 1$, so that always $d = \ell + n \geqq 2$. Suppose at first that $n > 0$. Set $\underline{e} = d$, $\underline{f} = 1$, $h = (\lambda-1)d + \ell$, so that (11.4), (11.5) hold, and $h/\underline{e} \notin \mathbb{Z}$. Let $F' = F(\pi')$ be the totally ramified extension of F with $(\pi')^{\underline{e}} = \pi$. We have $F' = F(\pi'^h + r\pi'^{h+1})$ for suitable $r \in R$ (see the argument in [1, chapter VII, §6]). Then $\xi = \pi'^h + r\pi'^{h+1} + y$ has $|\xi - y| = 2^{-h/\underline{e}}$, and every $x \in R$ has $|x-y| \neq |\xi - y|$; therefore $|\xi - x| = |\xi - y - (x-y)| \geqq 2^{-h/\underline{e}}$, so that (11.2), (11.3) hold. We take f to be the defining polynomial of ξ over F.

In the case when $n = 0$, let $d = \ell$ and set $\underline{e} = 1$, $\underline{f} = d$, $h = \lambda$, so that again (11.4), (11.5) hold. Let $\bar{g}(X)$ be a monic irreducible polynomial of degree d with coefficients in the residue class field $\bar{F}$ of R — such a polynomial exists by hypothesis. $\bar{g}$ lifts to an irreducible polynomial g in $R[X]$. Let $F' = F(\eta)$ where η

is a root of g, so that $F' \supset F$ is unramified and of degree d. Note that $|\eta| = 1$ and set $\xi = \pi^\lambda \eta + y$, so that $|\xi - y| = |\pi^\lambda \eta| = 2^{-\lambda}$. When $x \in R$, write $x - y = \pi^\rho \zeta$ with $|\zeta| = 1$. Then $|\xi - x| = |\pi^\lambda \eta - \pi^\rho \zeta| \geqq 2^{-\lambda}$, since either $\lambda \neq \rho$, or $\lambda = \rho$ and $\pi^\lambda \eta - \pi^\rho \zeta = \pi^\lambda(\eta - \zeta)$ and $|\eta - \zeta| = 1$, since η, ζ have different images in the residue class field $\overline{F(\eta)}$. Therefore (11.2), (11.3) hold. We take f to be the defining polynomial of ξ over F. □

We have given a canonical procedure to associate with every primitive irreducible polynomial f a polynomial sum $\underline{\underline{K}} = \underline{\underline{K}}(f)$ with $T(f) = \underline{\underline{K}}(f)^*$. Therefore we can associate in a canonical way a sum $\underline{\underline{K}} = \underline{\underline{K}}(f)$ with every primitive polynomial f in such a way that

$$\underline{\underline{K}}(f_1 f_2) = \underline{\underline{K}}(f_1) \oplus \underline{\underline{K}}(f_2) \tag{11.6}$$

and

$$T(f) = \underline{\underline{K}}(f)^* \tag{11.7}$$

by (9.1). The number m of summands of $\underline{\underline{K}}(f)$ has

$$m \leqq d \tag{11.8}$$

where $d = \deg f$.

12. Discriminants and Resultants of Polynomials. Given a nonconstant polynomial $f \in R[X]$, let $D(f)$ be its discriminant when $\deg f > 1$, and $D(f) = 1$ when $\deg f = 1$. Given nonconstant polynomials $f, g \in R[X]$, let $R(f, g)$ be their resultant. It is easily checked that for nonconstant polynomials f_1, f_2, g,

$$D(f_1 f_2) = D(f_1) D(f_2) R(f_1, f_2)^2, \qquad R(f_1 f_2, g) = R(f_1, g) R(f_2, g).$$

Thus the quantities $\delta(f) = \operatorname{ord} D(f)$, $\rho(f, g) = \operatorname{ord} R(f, g)$ have

$$\delta(f_1 f_2) = \delta(f_1) + \delta(f_2) + 2\rho(f_1, f_2), \tag{12.1}$$

$$\rho(f_1 f_2, g) = \rho(f_1, g) + \rho(f_2, g), \tag{12.2}$$

where some of these may be $+\infty$.

Proposition 12.1. *Let f, g in $R[X]$ be primitive and nonconstant. Then*

$$\delta(\underline{\underline{K}}(f)) \leqq \delta(f), \tag{12.3}$$

$$\rho(\underline{\underline{K}}(f), \underline{\underline{K}}(g)) \leqq \rho(f, g). \tag{12.4}$$

Proof. In view of the relations (7.1), (7.2) for stalk sums and the corresponding relations (12.1), (12.2) for polynomials, and in view of (11.6), we may restrict ourselves to polynomials f, g which are primitive and irreducible (over the quotient field F of R).

We begin with (12.3). When f is of degree 1, then $\underline{\underline{K}}(f)$ has at most one summand, so that $\delta(\underline{\underline{K}}(f)) = 0$ and (12.3) is true. When $\deg f > 1$ and the roots of f have absolute value > 1, then $\underline{\underline{K}} = \underline{\underline{0}}$ and again $\delta(\underline{\underline{K}}(f)) = 0$. We are left with the case when f is of degree $d > 1$ and its roots $\xi_1, \ldots, \xi_d$ have common absolute value $\leqq 1$. With suitable $y \in R$ we have (11.3), so that

$$|y - \xi_i| = 2^{-h/\underline{e}} \qquad (i = 1, \ldots, d). \tag{12.5}$$

Then $\underline{\underline{K}}(f) = \ell K(y_\lambda) + nK(y_\lambda^-)$ with $\ell + n = d$. Therefore

$$\begin{aligned}\delta(\underline{\underline{K}}(f)) &= \ell(\ell-1)\lambda + n(n-1)(\lambda-1) + 2\ln(\lambda-1)\\ &= (d-1)((\lambda-1)d+\ell) - \ell(d-\ell)\\ &\leqq (d-1)((\lambda-1)d+\ell) = (d-1)h\underline{f}\end{aligned}$$

by (11.5). On the other hand, $|\xi_i - \xi_j| \leqq 2^{-h/\underline{e}}$ by (12.5), and the leading coefficient of f has modulus 1, so that

$$\delta(f) = \operatorname{ord} D(f) \geqq d(d-1)h/\underline{e} = (d-1)h\underline{f}.$$

Now (12.3) follows.

As for (12.4), write

$$f(X) = c(X-\xi_1)\cdots(X-\xi_d), \quad g(X) = c'(X-\eta_1)\cdots(X-\eta_{d'})$$

in a suitable field extension. We note that $\underline{\underline{K}}(f) = \underline{\underline{0}}$ or $\underline{\underline{K}}(g) = \underline{\underline{0}}$ if the roots ξ_i of f or the roots η_j of g have absolute value > 1. Then $\rho(\underline{\underline{K}}(f), \underline{\underline{K}}(g)) = 0$ and we are done. We suppose, then, that the roots of f as well as of g have absolute values $\leqq 1$. We assume initially that $d > 1$, $d' > 1$. We have (12.5) and similarly $|y' - \eta_j| = 2^{-h'/\underline{e}'}$ $(1 \leqq j \leqq d')$ for suitable y, y' in R. Further, $\underline{\underline{K}}(f) = \ell K + nK^-$, $\underline{\underline{K}}(g) = \ell' K' + n' K'^-$ with $K = K(y_\lambda)$, $K' = K(y'_\mu)$ for suitable λ, μ.

Suppose at first that $K \cap K'$ is properly contained in K as well as in K'. Then $\lambda(K \cap K') = \lambda(K^- \cap K') = \lambda(K \cap K'^-) = \lambda(K^- \cap K'^-) = \nu$, say, where $\nu < \min(\lambda, \mu)$. We have $\rho(\underline{\underline{K}}(f), \underline{\underline{K}}(g)) = dd'\nu$. On the other hand $|y - y'| = 2^{-\nu}$, and since $\nu \leqq \lambda - 1 < h/\underline{e}$ by (11.4), similarly $\nu < h'/\underline{e}'$, we have $|\xi_i - \eta_j| = |y - y'| = 2^{-\nu}$ $(1 \leqq i \leqq d,\ 1 \leqq j \leqq d')$, so that $\rho(f, g) = dd'\nu$, and (12.4) holds with equality.

We now suppose that $K \cap K'$ is not properly contained in both K and K'; say $K \cap K' = K$, so that $K \subseteq K'$; therefore $\lambda \leqq \mu$. We have $\lambda(K \cap K') = \lambda$, $\lambda(K \cap K'^-) \leqq \lambda$, and $\lambda(K^- \cap K') = \lambda(K^- \cap K'^-) = \lambda - 1$; therefore

$$\rho(\underline{\underline{K}}(f), \underline{\underline{K}}(g)) \leqq \ell d' \lambda + nd'(\lambda - 1) = d'(d\lambda - n).$$

When $\lambda < \mu$ we have $h/\underline{e} \leqq \lambda \leqq \mu - 1 < h'/\underline{e}'$, and when $\lambda = \mu$, so that $K = K'$, we may suppose without loss of generality that $h/\underline{e} \leqq h'/\underline{e}'$. The infinite chains corresponding to y, y' both contain K, and therefore $|y - y'| \leqq 2^{-\lambda}$. Since $|y - \xi_i| = 2^{-h/\underline{e}}$, $|y' - \eta_j| = 2^{-h'/\underline{e}'}$, we obtain $|\xi_i - \eta_j| \leqq 2^{-h/\underline{e}}$, and therefore

$$\rho(f, g) \geqq dd'h/\underline{e} = d'h\underline{f} = d'(d\lambda - d + \ell) = d'(d\lambda - n)$$

by (11.5). So (12.4) follows.

We have assumed that $d > 1$, $d' > 1$. But when, e.g., $d = 1$, the construction of $\underline{\underline{K}}(f)$ follows the same pattern as before, if we set $\underline{e} = \underline{f} = 1$, $\ell = 1$, $n = 0$ and $h = \lambda = \infty$. We leave the details to the reader. □

13. Proof of Theorems C and C′. For Theorem C we will again work in the more general framework of complete discrete valuation rings R. When $f(X) = \pi^n f_0(X)$ with f_0 primitive, we have $T(f) = L^n \underline{\underline{K}}^*(f_0)$. Here $\underline{\underline{K}}(f_0)$ has $m \leqq d = \deg f$ summands by (11.8). We have $\delta(\underline{\underline{K}}(f_0)) \leqq \delta(f_0)$ by (12.3), and we have $\delta(f) = (2d-2)n + \delta(f_0)$, so that $d \geqq 2$ yields

$$n + \delta(\underline{\underline{K}}(f_0)) \leqq \delta(f).$$

Given n, the number of isomorphism classes of trees $T(f_0)$ where f_0 runs through primitive polynomials with $\delta(f_0) \leqq \delta$ is $\leqq c_6(d)\delta^{d-1}$ by Theorem 7.1 and (11.7). Taking the sum over $n \leqq \delta$, we obtain the bound $c_1(d)\delta^d$ of Theorem C.

When $n, a_1, \ldots, a_{d-1}$ satisfy (6.10) (but with d in place of m), the polynomial

$$f(X) = \pi^n(X - \pi^{a_1}) \cdots (X - \pi^{a_{d-1}})X$$

has $T(f) = L^n \underline{\underline{K}}^*$ with $\underline{\underline{K}}$ a stalk sum as described in section 6. Here

$$\delta(f) = (2d-2)n + 2((d-1)a_1 + (d-2)a_2 + \cdots + a_{d-1}). \tag{13.1}$$

The number of d-tuples $n, a_1, \ldots, a_{d-1}$ with $0 \leqq n < a_1 < \cdots < a_{d-1}$ and with the right hand side of (13.1) below δ is $\geqq c_7(d)\delta^d$ when δ is large. By what we have said in section 6, different d-tuples $n, a_1, \ldots, a_{d-1}$ give rise to different Poincaré series for f. Therefore as f ranges through polynomials of degree d (where $d \geqq 2$) and with $\delta(f) \leqq \delta$, we obtain at least $c_7(d)\delta^d$ different Poincaré series, and therefore at least that many nonisomorphic trees $T(f)$. This proves Theorem C.

We now turn to Theorem C$'$. In section 7 we defined isomorphisms of polynomial sums. We allowed reorderings of the summands. Now given *ordered* m-tuples $(K_1, \ldots, K_m)$ of stalks, we will say $(K_1, \ldots, K_m)$ and $(H_1, \ldots, H_m)$ are isomorphic if there is a map $\Psi : \bigcup K_i \to \bigcup H_i$ whose restriction to K_i is an isomorphism $K_i \to H_i$ $(i = 1, \ldots, m)$. Each stalk sum with m summands gives rise to at most $m!$ ordered m-tuples, so that as a variation on Theorem 7.1 we see that there are $\leqq c_8(m)\delta^{m-1}$ isomorphism classes of ordered m-tuples of stalks of discriminant $\leqq \delta$.

Now let f be a polynomial of degree $d > 0$ in $\mathbb{Z}[X]$, and write

$$f(X) = p^n f_1(X)^{e_1} \cdots f_s(X)^{e_s}, \tag{13.2}$$

where $f_1, \ldots, f_s$ are nonproportional polynomials which are irreducible over $\mathbb{Q}$, and whose coefficients lie in $\mathbb{Z}$ but are not all divisible by p. Say $\underline{\underline{K}}(f_i) = K_{i1} \oplus \cdots \oplus K_{im_i}$. The number of possibilities for $s, e_1, \ldots, e_s, m_1, \ldots, m_s$ is under a bound $c_9(d)$. From now on we will consider those quantities fixed. Set

$$g(X) = f_1(X) \cdots f_s(X).$$

Then the height $H(g)$ of g, i.e., the maximum modulus of its coefficients, has $H(g) \leqq c_{10}(d)H(f) \leqq c_{10}(d)H$. The discriminant $D(g)$ is $\neq 0$, and has $|D(g)| \leqq c_{11}(d)H^{2d-2}$, so that its p-adic order $\delta(g) \leqq c_{12}(d)((\log H / \log p) + 1)$. By Theorem 7.1, the number of isomorphism classes of stalk sums $\underline{\underline{K}}(g) = \underline{\underline{K}}(f_1) \oplus \cdots \oplus \underline{\underline{K}}(f_s)$ is

$$\leqq c_{13}(d)((\log H / \log p)^{d-1} + 1).$$

By the variation on Theorem 7.1 given above, the number of isomorphism classes of ordered tuples

$$(K_{11}, \ldots, K_{1m_1}, \ldots, K_{s1}, \ldots, K_{sm_s})$$

is still under this bound, provided $c_{13}(d)$ is replaced by a suitable larger constant. Given the class of this tuple, and given $e_1, \ldots, e_s$, the isomorphism class of

$$\underline{\underline{K}}(f_1^{e_1} \cdots f_s^{e_s}) = (K_{11} \cdots K_{1m_1})^{e_1} \cdots (K_{s1} \cdots K_{sm_s})^{e_s}$$

is determined. Since there are $\leqq (\log H / \log p) + 1$ possibilities for the extra parameter n in (13.2), the upper bound (4) of Theorem C$'$ is established. That this bound is essentially best possible is seen by considering polynomials

$$p^n(X - p^{a_1}) \cdots (X - p^{a_{d-1}})X.$$

14. The Poincaré Series of f. Suppose our ring R has a residue class field of finite order p (a prime power!). For example, $R = \mathbb{Z}_p$ where p is a prime, or $R = k[[\mathcal{X}]]$ where k is a field of cardinality p. The *Poincaré series of f* is the Poincaré series of $T(f)$, and is

$$\mathfrak{P}_f(Z) = \sum_{\lambda=0}^{\infty} |T(f)_\lambda| Z^\lambda.$$

By Theorems 6.1 and B, this series lies in $\mathbb{Q}(Z)$. As is well known by deep work of Igusa, and reproved by Denef, this is also true in the considerably more difficult case of polynomials in several variables.

Recall that $|T(f)_\lambda|$ is the number of solutions of $f(x) \equiv 0 \pmod{\mathfrak{p}^\lambda}$. Proposition 6.4 yields

$$\operatorname{dig}_p |T(f)_\lambda| \leqq \deg f \tag{14.1}$$

for nonconstant f. Theorem 8.1 gives

$$|T(f)_\lambda| \leqq 2p^{\delta/2} + d - 2, \tag{14.2}$$

where $\deg f = d > 0$ and δ is the p-adic order of the discriminant of f.

The Poincaré series has a different interpretation when $R = k[[\mathcal{X}]]$ where k is infinite. Recall from (6.1):

$$T_\lambda = \bigcup_{u \in T(\lambda)} F'(u, \lambda). \tag{14.3}$$

Here u is a polynomial $u = u(\mathcal{X}) \pmod{\mathcal{X}^{\lambda(u)}}$, and $F'(u, \lambda)$ consists of polynomials $t(\mathcal{X})$ (modulo $\mathcal{X}^\lambda$) having $t(\mathcal{X}) \equiv u(\mathcal{X}) \pmod{\mathcal{X}^{\lambda(u)}}$. Therefore $F'(u, \lambda)$ is a linear submanifold of U_λ of dimension $\lambda - \lambda(u)$. Since k is infinite, the decomposition of T_λ into a finite union of linear submanifolds is unique.

Theorem 14.1. *Suppose $R = k[[\mathcal{X}]]$ where k is infinite, and suppose $f \in R[X]$. Then for each λ, $T(f)_\lambda$ is a finite union of linear submanifolds $M_{\lambda_1}, \dots, M_{\lambda, n(\lambda)}$ of U_λ. Set $P_\lambda(z) = \sum_{i=1}^{n(\lambda)} z^{c_i}$, where $c_i = \dim M_{\lambda_i}$, and set $\mathfrak{P}(z, Z) = \sum_{\lambda=0}^{\infty} P_\lambda(z) Z^\lambda$. Then $\mathfrak{P}(z, Z)$ lies in $\mathbb{Q}(z, Z)$.*

The proof is obvious, since the M_{λ_i} are the $F'(u, \lambda)$ with $u \in T(f)(\lambda)$, and since by (6.4), (6.5), $\mathfrak{P}(z, Z)$ is the Poincaré series of $T(f)$, which is rational by Theorems 6.1 and B.

15. Width and Continued Fractions. Our main theorem here can be stated in the context of stalks:

Theorem 15.1. *Let U be a universal tree of finite valence p, which need not be a prime. Let*

$$T = K_1 \cdots K_m \tag{15.1}$$

be a product of $m \geqq 1$ stalks. Given $\lambda > 0$, write λ/m as a regular continued fraction:

$$\lambda/m = c_0 + \cfrac{1}{c_1 + \cfrac{}{\ddots + 1/c_n}}$$

with n odd. Then if $p \geqq m$, *we have*

$$(15.2) \qquad |T_\lambda| \leqq c_1 p^{\lambda-c_0-1} + c_3 p^{\lambda-c_0-c_2-1} + \cdots + c_n p^{\lambda-c_0-c_2-\cdots-c_{n-1}-1}.$$

This bound is best possible for every $\lambda > 0$, $m > 0$ and $p \geqq m$. But we will show this only when p is a prime, in section 16 in the context of polynomial trees.

The cardinality $|T_\lambda|$ can only decrease if some factors in (15.1) are removed, and therefore the bound (15.2) remains valid when T is a product of $\leqq m$ stalks. Since for a polynomial f we have $m = m(\underline{\underline{K}}(f)) \leqq d = \deg f$, Theorem D follows.

Corollary. *Define* m_1 *by* $\lambda = mc_0 + m_1$, *so that* $\lambda/m = c_0 + (m_1/m)$. *Note that* $1 \leqq m_1 \leqq m$. *Then*

$$(15.3) \qquad |T_\lambda| \leqq (m/m_1)p^{\lambda-c_0-1}.$$

In particular, $|T_\lambda| \leqq mp^{\lambda-c_0-1}$. This is the same bound as in Stewart [2, (44)] in the context of polynomial trees, since his exponent is $[\lambda(m-1)/m] = \lambda - c_0 + [-m_1/m] = \lambda - c_0 - 1$. Our Corollary yields the Corollary to Theorem D.

Deduction of the Corollary. Introduce the notation

$$[a_0] = a_0, \; [a_0, a_1, \ldots, a_\ell] = a_0 + \cfrac{1}{a_1 + \cfrac{\ddots}{+1/a_\ell}}$$

for reals a_j, with $a_1, a_2, \ldots$ positive. Our hypothesis then becomes

$$(15.4) \qquad \lambda/m = [c_0, c_1, \ldots, c_n].$$

Define $m_0, m_1, \ldots, m_n$ by $m_0 = m$, $\lambda = m_0 c_0 + m_1$,

$$(15.5) \qquad m_{i-2} = m_{i-1}c_{i-1} + m_i \qquad (2 \leqq i \leqq n).$$

Then each $m_i \in \mathbb{Z}$ and

$$\begin{aligned} \lambda/m = \lambda/m_0 = c_0 + 1/(m_0/m_1) &= [c_0, m_0/m_1] = \cdots \\ &= [c_0, \ldots, c_{i-1}, m_{i-1}/m_i] \\ &= [c_0, \ldots, c_{n-1}, m_{n-1}/m_n] \end{aligned}$$

for i in $1 \leqq i \leqq n$. Comparison with (15.4) gives

$$(15.6) \qquad m_{i-1}/m_i = [c_i, \ldots, c_n] \qquad (1 \leqq i \leqq n).$$

Therefore

$$(15.7) \qquad c_i = m_{i-1}/m_i - \frac{1}{[c_{i+1}, \ldots, c_n]} = \frac{(m_{i-1} - m_{i+1})}{m_i} \quad (1 \leqq i \leqq n-1),$$

but $c_n = m_{n-1}/m_n$.

We will show that for odd g in $1 \leqq g \leqq n$,

$$(15.8) \qquad |T_\lambda| \leqq c_1 p^{\lambda-c_0-1} + \cdots + c_{g-2}p^{\lambda-c_0-\cdots-c_{g-3}-1} + (m_{g-1}/m_g)p^{\lambda-c_0-\cdots-c_{g-1}-1},$$

which for $g = 1$ is to be interpreted as (15.3). When $g = n$, (15.8) is the same as (15.2), since $c_n = m_{n-1}/m_n$. To do induction down from g to $g-2$ (assuming

$g \geqq 3$) it will suffice to show that $\Delta_g \geqq 0$, where Δ_g is the difference of the right hand sides of (15.8) for $g-2$ and for g. But

$$\begin{aligned} p^{-\lambda+c_0+\cdots+c_{g-3}+1}\Delta_g &= (m_{g-3}/m_{g-2}) - c_{g-2} - (m_{g-1}/m_g)p^{-c_{g-1}} \\ &\geqq (m_{g-3}/m_{g-2}) - (m_{g-3}-m_{g-1})/m_{g-2} - (m_{g-1}/m_g)p^{-1} \geqq 0 \end{aligned}$$

by (15.7) and since $m_g p \geqq p \geqq m \geqq m_{g-2}$.

Proof of Theorem 15.1. Set $\underline{\underline{K}} = K_1 \oplus \cdots \oplus K_m$. Let $TT(\lambda)$ be the set of section 8; say $TT(\lambda) = \{u_1, \ldots, u_\ell\}$. Then in particular, $u_1, \ldots, u_\ell$ are incompatible vertices of $K_1 \cup \cdots \cup K_m$. We recall

$$\omega_i = \lambda - \lambda(u_i) \quad (i = 1, \ldots, \ell) \tag{15.9}$$

and formula (8.4):

$$|T_\lambda| = \sum_{i=1}^{\ell} p^{\omega_i}. \tag{15.10}$$

By relabeling the stalks of $\underline{\underline{K}}$ we may suppose that u_i lies precisely on the stalks $K_{i1}, \ldots, K_{iq_i}$. The stalks K_{ij} ($1 \leqq i \leqq \ell$, $1 \leqq j \leqq q_i$) are among the summands of $\underline{\underline{K}}$, and $K_{ij}, K_{i',j'}$ for $i \neq i'$ are distinct by the incompatibility of $u_i, u_{i'}$. Therefore

$$q_1 + \cdots + q_\ell \leqq m. \tag{15.11}$$

We may suppose that $q_1 \geqq \cdots \geqq q_\ell$.

Let $m_0, m_1, \ldots, m_n$ be as above. Then $m = m_0 > m_1 \cdots > m_{n-1} \geqq m_n$, with $m_{n-1} = m_n$ precisely if $c_n = 1$. *In what follows, g, h will be odd integers.* We have

$$\begin{aligned} \lambda &= m_0c_0 + m_1 = \cdots \\ &= m_0c_0 + m_2c_2 + \cdots + m_{g-1}c_{g-1} + m_g \qquad (1 \leqq g \leqq n), \end{aligned} \tag{15.12}$$

$$\begin{aligned} m_0 &= m_1c_1 + m_2 = \cdots \\ &= m_1c_1 + m_3c_3 + \cdots + m_gc_g + m_{g+1} \qquad (1 \leqq g \leqq n) \end{aligned} \tag{15.13}$$

where we set $m_{n+1} = 0$. Define

$$e_{-1} = 0, \quad e_g = c_1 + c_3 + \cdots + c_g \quad (1 \leqq g \leqq n), \tag{15.14}$$

$$f_{-2} = 0, \quad f_{g-1} = c_0 + c_2 + \cdots + c_{g-1} \quad (1 \leqq g \leqq n). \tag{15.15}$$

For $1 \leqq g \leqq n$, let $\mathcal{A}_g$ be the following assertion.

(a) *The estimate* (15.2) *is true if $\ell < e_g$.*

(b) *The estimate* (15.2) *is true if $\ell \geqq e_g$ and if $q_{e_g} < m_g$ or $\lambda(u_j) > f_{g-1}+1$ for some j in $e_{g-2} < j \leqq e_g$.*

(c) *When $\ell \geqq e_g$, then*

$$\sum_{i=1}^{e_g} p^{\omega_i} \leqq c_1 p^{\lambda - f_0 - 1} + \cdots + c_g p^{\lambda - f_{g-1} - 1}.$$

Now $\mathcal{A}_1, \mathcal{A}_3, \ldots, \mathcal{A}_n$ imply (15.2) as follows. By (a) we may suppose that $\ell \geqq e_n$. By (b) we may suppose that $q_{e_g} \geqq m_g$ ($1 \leqq g \leqq n$), so that

$$\begin{aligned} q_1 + q_2 + \cdots + q_{e_g} &\geqq e_1m_1 + (e_3 - e_1)m_3 + \cdots + (e_g - e_{g-2})m_g \\ &= c_1m_1 + c_3m_3 + \cdots + c_gm_g \\ &= m - m_{g+1} \end{aligned} \tag{15.16}$$

by the ordering $q_1 \geqq q_2 \geqq \cdots$ and by (15.13). Applying this with $g = n$, recalling that $m_{n+1} = 0$ and comparing with (15.11), we obtain $\ell = e_n$. Now (15.2) follows from (15.10) and part (c) of $\mathcal{A}_n$. □

Before proceeding further we insert the following. For any real number x let $\lceil x \rceil$ denote the smallest integer greater than or equal to x.

Lemma 15.2. *Given integers $m > 0, \lambda, b$, set*

$$\varphi(\nu) = \min(\lambda - \nu, \nu(m-1) + b).$$

Set $\nu_1 = \lceil (\lambda - b)/m \rceil$. Then for $\nu \in \mathbb{Z}$ we have $\varphi(\nu) \leqq \varphi(\nu_1) = \lambda - \nu_1$. When $\nu > \nu_1$ we have $\varphi(\nu) \leqq \lambda - \nu_1 - 1$.

Proof. The maximum of $\varphi(\nu)$ for $\nu \in \mathbb{R}$ is taken at $\nu = (\lambda - b)/m$, and the maximum for $\nu \in \mathbb{Z}$ is taken at ν_1 or $\nu_0 = [(\lambda - b)/m]$. Here $\varphi(\nu_1) = \lambda - \nu_1$, $\varphi(\nu_0) = \nu_0(m-1) + b$, and

$$\begin{aligned} \varphi(\nu_1) - \varphi(\nu_0) &= \lambda - b - m\nu_0 + \nu_0 - \nu_1 \\ &= \lambda - b - m[(\lambda - b)/m] + [(\lambda - b)/m] - \lceil (\lambda - b)/m \rceil \\ &\geqq 0 \end{aligned}$$

since $\lambda - b$ is an integer. Therefore the maximum of $\varphi(\nu)$ for $\nu \in \mathbb{Z}$ is taken at $\nu = \nu_1$, and is $\varphi(\nu_1) = \lambda - \nu_1$. When $\nu > \nu_1$, then $\varphi(\nu) = \lambda - \nu \leqq \lambda - \nu_1 - 1$. □

We will now in turn prove $\mathcal{A}_1, \mathcal{A}_3, \ldots, \mathcal{A}_n$. We will prove $\mathcal{A}_{g+2}$, assuming that either $g = -1$ or that $1 \leqq g \leqq n-2$ and $\mathcal{A}_1, \mathcal{A}_3, \ldots, \mathcal{A}_g$ have been established. In the latter case we may suppose that $\ell \geqq e_g$ and that (15.16) holds. Then in view of (15.11),

$$q_{e_g+1} + \cdots + q_\ell \leqq m_{g+1}. \tag{15.17}$$

When $g = -1$ this is the same as (15.11).

Lemma 15.3. *For $e_g < i \leqq \ell$ and ω_i given by (15.9),*

$$\omega_i \leqq \lambda - f_{g+1} - 1 - \lceil (m_{g+2} - q_i)/m_{g+1} \rceil.$$

Proof. Write $\tau = \tau_{\underline{\underline{K}}}$, $\sigma = \sigma_{\underline{\underline{K}}}$. Then

$$\sigma(u_i) = \sum_{\{0\} < t \leqq u_i} \tau(t) = \tau(u_i) + \sum_{h=-1}^{g} \Sigma_h \tag{15.18}$$

where

$$\Sigma_h = \sum_{\substack{\{0\} < t < u_i \\ f_{h-1} < \lambda(t) \leqq f_{h+1}}} \tau(t) \quad \text{when } -1 \leqq h \leqq g-2,$$

$$\Sigma_g = \sum_{\substack{\{0\} < t < u_i \\ f_{g-1} < \lambda(t)}} \tau(t).$$

Suppose $h \leqq g$ and $\{0\} < t < u_i$, $f_{h-1} < \lambda(t)$. Now if $1 \leqq j \leqq e_h$ (only possible if $g \geqq h \geqq 1$), say $e_{r-2} < j \leqq e_r$ where $1 \leqq r \leqq h$ with r odd, then we may suppose by part (b) of $\mathcal{A}_r$ that $\lambda(u_j) \leqq f_{r-1} + 1 \leqq f_{h-1} + 1 \leqq \lambda(t)$. Since $j \leqq e_h \leqq e_g < i$

and therefore u_j, u_i are incompatible, also u_j, t are incompatible, and the stalks K_{jv} $(1 \leqq v \leqq q_j)$ containing u_j do not contain t. Therefore when $f_{h-1} < \lambda(t)$,

$$\tau(t) \leqq m - (q_1 + \cdots + q_{e_h}) \leqq m_{h+1} \tag{15.19}$$

by (15.16), which is true by induction with g replaced by $h \leqq g$. Note that (15.19) is trivially true for $h = -1$. We may conclude that for $-1 \leqq h \leqq g-2$,

$$\Sigma_h \leqq (f_{h+1} - f_{h-1})m_{h+1} = c_{h+1}m_{h+1}. \tag{15.20}$$

Now suppose that

$$f_{g-1} \leqq \lambda(u_i) - 1. \tag{15.21}$$

Then by (15.19),

$$\Sigma_g \leqq (\lambda(u_i) - f_{g-1} - 1)m_{g+1}.$$

Since $\tau(u_i) = q_i$, (15.18), (15.20) and (15.12) yield

$$\begin{aligned}\sigma(u_i) &\leqq q_i + c_0m_0 + \cdots + c_{g-1}m_{g-1} + (\lambda(u_i) - f_{g-1} - 1)m_{g+1} \\ &= q_i + \lambda - m_g + (\lambda(u_i) - f_{g-1} - 1)m_{g+1}.\end{aligned}$$

The vertices u_i had $\lambda \leqq \sigma(u_i)$ by the definition of $TT(\lambda)$ (see section 8), and therefore

$$\omega_i = \lambda - \lambda(u_i) = \min(\lambda - \lambda(u_i), \lambda(u_i)(m_{g+1} - 1) + b_i)$$

with $b_i = q_i + \lambda - m_g - m_{g+1}(f_{g-1} + 1)$. By Lemma 15.2, $\omega_i \leqq \lambda - \nu(i)$, where $\nu(i) = \lceil (\lambda - b_i)/m_{g+1} \rceil$. But

$$\begin{aligned}\lambda - b_i &= m_{g+1}(f_{g-1} + 1) + m_g - q_i \\ &= m_{g+1}(f_{g-1} + 1 + c_{g+1}) + m_{g+2} - q_i \\ &= m_{g+1}(f_{g+1} + 1) + m_{g+2} - q_i,\end{aligned}$$

so that

$$\nu(i) = f_{g+1} + 1 + \lceil (m_{g+2} - q_i)/m_{g+1} \rceil \tag{15.22}$$

and

$$\omega_i \leqq \lambda - f_{g+1} - 1 - \lceil (m_{g+2} - q_i)/m_{g+1} \rceil,$$

as asserted in the Lemma. This holds when (15.21) is valid.

We will show that (15.21) is always valid, i.e., that $f_{g-1} \geqq \lambda(u_i)$ is impossible. For in that case $\Sigma_g = 0$, and in view of (15.19),

$$\begin{aligned}\Sigma_{g-2} &= \sum_{\substack{\{0\} < t < u_i \\ f_{g-3} < \lambda(t) \leqq f_{g-1}}} \tau(t) \\ &\leqq (f_{g-1} - f_{g-3})m_{g-1} = c_{g-1}m_{g-1},\end{aligned}$$

so that by (15.18), (15.20), (15.12),

$$\begin{aligned}\sigma(u_i) &\leqq q_i + c_0m_0 + \cdots + c_{g-3}m_{g-3} + c_{g-1}m_{g-1} \\ &= q_i + \lambda - m_g.\end{aligned}$$

By hypothesis $i > e_g$, so that $q_i \leqq m_{g+1} < m_g$ by (15.17), and therefore $\sigma(u_i) < \lambda$, contradicting the fact that we had $\sigma(u_i) \geqq \lambda$. The proof of the lemma is complete. □

When $e_g < i \leqq \ell$, we have $q_i \leq m_{g+1}$ from (15.17); therefore

$$\lceil (m_{g+2} - q_i)/m_{g+1} \rceil = \begin{cases} 0 \text{ when } q_i \geqq m_{g+2}, \\ 1 \text{ when } q_i < m_{g+2}. \end{cases} \tag{15.23}$$

Therefore by Lemma 15.3,

$$\omega_i \leqq \begin{cases} \lambda - f_{g+1} - 1 \text{ when } q_i \geqq m_{g+2}, \\ \lambda - f_{g+1} - 2 \text{ when } q_i < m_{g+2}. \end{cases}$$

Set $a = 0$ if $q_{e_g} < m_{g+2}$; otherwise let a be largest integer with $e_g + a \leqq \ell$ and $q_{e_g+a} \geqq m_{g+2}$. By (15.17), $am_{g+2} \leqq m_{g+1}$, so that $a \leqq m_{g+1}/m_{g+2} = [c_{g+2}, \dots, c_n]$ by (15.6), and therefore

$$a \leqq c_{g+2}. \tag{15.24}$$

(Note that $n = g + 2$ or $n \geqq g + 4$.) We obtain

$$\omega_i \leqq \begin{cases} \lambda - f_{g+1} - 1 \text{ when } e_g < i \leqq e_g + a, \\ \lambda - f_{g+1} - 2 \text{ when } e_g + a < i \leqq \ell. \end{cases}$$

Using the last assertion of Lemma 15.2 we can modify Lemma 15.3 to get

$$\omega_i \leqq \lambda - f_{g+1} - 2$$

also when $e_g < i \leqq e_g + a$ and $\lambda(u_i) \geqq \nu(i) + 1$, which by (15.22), (15.23) is the same as $\lambda(u_i) \geqq f_{g+1} + 2$.

By part (c) of $\mathcal{A}_g$ (when $g \geqq 1$) we see that

$$\sum_{i=1}^{\ell} p^{\omega_i} \leqq \sum_{h=1}^{e_g} c_h p^{\lambda - f_{h-1} - 1} + \sum_{i=e_g+1}^{\ell} p^{\omega_i}; \tag{15.25}$$

when $g = -1$ this is trivially true if we understand the first sum on the right hand side to be 0. The second summand on the right hand side is

$$\leqq (a - \varepsilon)p^{\lambda - f_{g+1} - 1} + (\ell - a + \varepsilon)p^{\lambda - f_{g+1} - 2},$$

where we set $\varepsilon = 1$ if $a > 0$ and $\lambda(u_i) \geqq f_{g+1} + 2$ for some i in $e_g < i \leqq e_g + a$, and $\varepsilon = 0$ otherwise. Now when

$$a - \varepsilon < c_{g+2}, \tag{15.26}$$

the second summand is

$$\begin{aligned} &\leqq c_{g+2}p^{\lambda - f_{g+1} - 1} - p^{\lambda - f_{g+1} - 1} + \ell p^{\lambda - f_{g+1} - 2} \\ &\leqq c_{g+2}p^{\lambda - f_{g+1} - 1}, \end{aligned}$$

since $p \geqq m \geqq \ell$, and then (15.2) follows from (15.10), (15.25).

When $\ell < e_{g+2}$, we have $e_g + a \leqq \ell < e_{g+2}$; therefore $a < c_{g+2}$, and (15.26) holds. This establishes part (a) of $\mathcal{A}_{g+2}$. When $\ell \geqq e_{g+2}$ and $q_{e_{g+2}} < m_{g+2}$, then $a < e_{g+2} - e_g = c_{g+2}$ and again (15.26) holds. When $\ell \geqq e_{g+2}$ and $a = c_{g+2}$ (note (15.24)!) and if $\lambda(u_i) \geqq f_{g+1} + 2$ for some i in $e_g < i \leqq e_g + a = e_{g+2}$, then $\varepsilon > 0$, whence (15.26) and (15.2) hold. This gives part (b) of assertion $\mathcal{A}_{g+2}$. Finally

$\omega_i \leqq \lambda - f_{g+1} - 1$ for $i > e_g$, so that by the truth of part (c) of $\mathcal{A}_g$ (or trivially when $g = -1$)

$$\sum_{i=1}^{e_{g+2}} p^{\omega_i} \leqq \sum_{i=1}^{e_g} p^{\omega_i} + \sum_{i=e_g+1}^{e_{g+2}} p^{\omega_i}$$
$$\leqq \sum_{h=1}^{g} c_h p^{\lambda - f_{h-1} - 1} + c_{g+2} p^{\lambda - f_{g+1} - 1},$$

which is part (c) of $\mathcal{A}_{g+2}$.

16. Congruences with Many Solutions. We still have to prove that the bound (7) in Theorem D is best possible. We will retain the notations introduced in the preceding section. From (15.5) we infer that for odd g, $1 \leqq g \leqq n-2$,

(16.1)
$$m_{g-1} = c_g m_g + m_{g+1} = \cdots = c_g m_g + c_{g+2} m_{g+2} + \cdots + c_{n-2} m_{n-2} + c_n m_n.$$

Set $a_1 = 0$ and, for odd g,

$$a_g = c_1 m_1 f_0 + c_3 m_3 f_2 + \cdots + c_{g-2} m_{g-2} f_{g-3} \quad (3 \leqq g \leqq n),$$
$$b_g = f_{g-1}(c_g m_g + c_{g+2} m_{g+2} + \cdots + c_n m_n) + m_g \quad (1 \leqq g \leqq n).$$

Lemma 16.1. $a_g + b_g = \lambda$ *$(1 \leqq g \leqq n)$.*

Proof. By (15.13) with $g = n$,

$$b_g = f_{g-1}(m - c_1 m_1 - \cdots - c_{g-2} m_{g-2}) + m_g,$$

with the interpretation that $b_1 = f_0 m + m_1$. Therefore $a_1 + b_1 = c_0 m + m_1 = \lambda$. When $g \geqq 3$, we note that $a_g - a_{g-2} = c_{g-2} m_{g-2} f_{g-3}$ and

$$b_g - b_{g-2} = (f_{g-3} + c_{g-1})(c_g m_g + \cdots + c_n m_n) + m_g$$
$$- f_{g-3}(c_{g-2} m_{g-2} + \cdots + c_n m_n) - m_{g-2}$$
$$= c_{g-1}(c_g m_g + \cdots + c_n m_n) - c_{g-2} m_{g-2} f_{g-3} + m_g - m_{g-2},$$

so that

$$a_g + b_g - a_{g-2} - b_{g-2} = c_{g-1} m_{g-1} - c_{g-1} m_{g-1} = 0$$

by (15.5), (16.1). The lemma follows by induction from $g-2$ to g.

To motivate what follows we wish to point out that the extremal case in the preceding section was when $\ell = e_n$ and when $\lambda(u_j) = f_{g-1} + 1$, $q_j = m_g$ for $e_{g-2} < j \leqq e_g$ (g odd, $1 \leqq g \leqq n$). It will be convenient to relabel the u_j as u_{gi} with $1 \leqq g \leqq n$ and $1 \leqq i \leqq c_g = e_g - e_{g-2}$, so that $\lambda(u_{gi}) = f_{g-1} + 1$, and m_g stalks of $\underline{\underline{K}}$ contain each u_{gi}.

We now begin our construction of a polynomial $f \in \mathbb{Z}[X]$. Pick $v_1 < v_3 < \cdots < v_n$ in $U(p)$ with $\lambda(v_g) = f_{g-1}$. Next, given g, pick u_{gi} ($1 \leqq i \leqq c_g$) in $U(p)$ having $u_{gi} > v_g$, $\lambda(u_{gi}) = f_{g-1} + 1 = \lambda(v_g) + 1$, and such that $u_{g1}, \ldots, u_{gc_g}, v_{g+2}$ are mutually incompatible when $g \leqq n-2$, and just $u_{g_1}, \ldots, u_{gc_g}$ are incompatible when $g = n$. (See Figure 5.)

Since $c_g + 1 \leqq m \leqq p$ when $g \leqq n-2$, and since $c_n \leqq p$ (with equality only when $n = 1$, $c_1 = p = m/m_1$, so that $m = p$, $m_1 = 1$), such a choice is possible. We claim that with this choice, any two vertices u_{gi}, u_{hj} with $(g,i) \neq (h,j)$ are incompatible. This is clear when $g = h$, and if, say, $g < h$, then u_{gi}, v_{g+2} are incompatible and $v_{g+2} \leqq v_h < u_{hj}$, so that indeed u_{gi}, u_{hj} are incompatible. Given h, i with h odd,

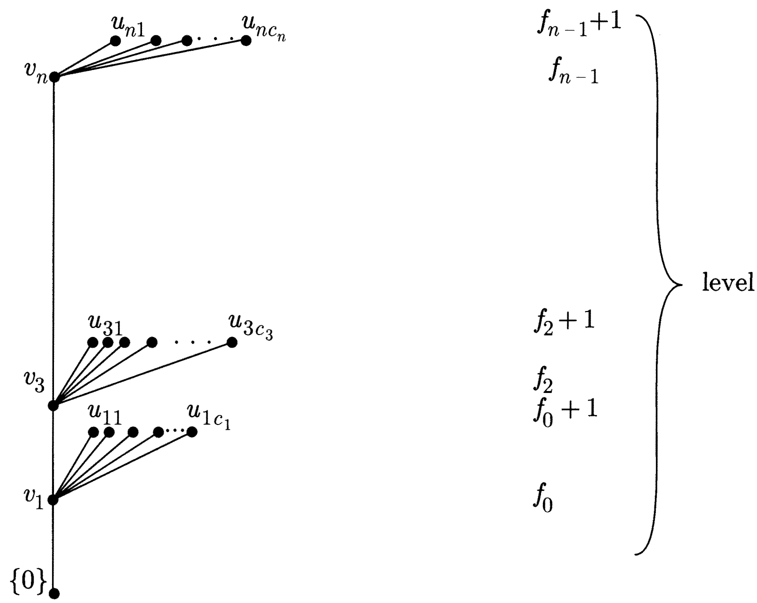

FIGURE 5

$1 \leqq h \leqq n$, $1 \leqq i \leqq c_h$, pick integers $x_{hiw} \in \mathbb{Z}$ $(1 \leqq w \leqq m_h)$ lying in the residue class u_{hi} $(\bmod p^{f_{h-1}+1})$. Set

$$f(X) = \prod (X - x_{hiw}),$$

where the product is over the set S of triples h, i, w with h odd, $1 \leqq h \leqq n$, $1 \leqq i \leqq c_h$, $1 \leqq w \leqq m_h$. Then, by the case $g = n$ of (15.13),

$$\deg f = \sum_{h=1}^{n} c_h m_h = m.$$

□

Lemma 16.2. *Suppose $x \in \mathbb{Z}$ lies in the residue class u_{gj} $(\bmod p^{f_{g-1}+1})$, where $1 \leqq g \leqq n$, $1 \leqq j \leqq c_g$. Then*

$$f(x) \equiv 0 \ (\bmod p^{\lambda}). \tag{16.2}$$

Proof. We claim that for $(h, i, w) \in S$ we have

$$\operatorname{ord}(x - x_{hiw}) \geqq \begin{cases} f_{h-1} \text{ when } h < g, \\ f_{g-1} \text{ when } h \geqq g, \\ f_{g-1} + 1 \text{ when } (h, i) = (g, j). \end{cases} \tag{16.3}$$

When $h < g$, we observe that $v_h < u_{hi}$, $v_h < v_g < u_{gj}$, so that (since $x \in u_{gj}, x_{hiw} \in u_{hi}$, and since $\lambda(v_h) = f_{h-1}$) indeed $\operatorname{ord}(x - x_{hiw}) \geqq f_{h-1}$. When $h \geqq g$, we similarly have $v_g < u_{gj}, v_g \leqq v_h < u_{hi}$, and therefore $\operatorname{ord}(x - x_{hiw}) \geqq f_{g-1}$. Finally, when $(h, i) = (g, j)$, we have both x, x_{hiw} in the class u_{gi} $(\bmod p^{f_{g-1}+1})$, and (16.3) follows.

Let $S = S_g^A \cup S_g^B$, with S_g^A, S_g^B respectively consisting of triples (h, i, w) with $h < g$ and with $h \geqq g$. By (16.3)

$$\operatorname{ord} \prod_{S_g^A} (x - x_{hiw}) \geqq c_1 m_1 f_0 + \cdots + c_{g-2} m_{g-2} f_{g-3} = a_g,$$

since for given h there are $c_h m_h$ choices for i, w. Similarly

$$\operatorname{ord} \prod_{S_g^B} (x - x_{hiw}) \geqq f_{g-1}(c_g m_g + \cdots + c_n m_n) + m_g = b_g.$$

We may conclude that $\operatorname{ord} f(x) \geqq a_g + b_g = \lambda$, and (16.2) follows. □

The number of $x \pmod{p^\lambda}$ which lie in the residue class $u_{gj} \pmod{p^{f_{g-1}+1}}$ is $p^{\lambda - f_{g-1}-1}$. Since the u_{gj} are pairwise incompatible, we obtain distinct residue classes $x \pmod{p^\lambda}$ for distinct pairs g, j. We may infer that the number of solutions of (16.2) is

$$\geqq \sum_{g=1}^{n} \sum_{j=1}^{c_g} p^{\lambda - f_{g-1} - 1} = c_1 p^{\lambda - f_0 - 1} + \cdots + c_n p^{\lambda - f_{n-1} - 1}.$$

By taking $m = d$ we see that the bound (7) is in fact best possible.

References

1. S. Lang, *Algebra*, 2nd ed., Addison–Wesley, 1984. MR **86j:**00003
2. C. L. Stewart, *On the number of solutions of polynomial congruences and Thue equations*, J. Amer. Math. Soc. 4 (1991), 793–835. MR **92j:**11032

Department of Mathematics, University of Colorado, Boulder, Colorado 80309
E-mail address: schmidt@euclid.Colorado.edu

Department of Pure Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
E-mail address: cstewart@watserv1.uwaterloo.ca