Mathematical Research Letters **5**, 273–279 (1998)

ZEROS OF SPARSE POLYNOMIALS OVER LOCAL FIELDS OF CHARACTERISTIC p

Bjorn Poonen

1. Statement of results

Let K be a field of characteristic $p > 0$ equipped with a valuation $v : K^* \to G$ taking values in an ordered abelian group G. Let $\mathcal{O}_K = \{\alpha \in K : v(\alpha) \geq 0\}$ and $\mathfrak{m}_K = \{\alpha \in K : v(\alpha) > 0\}$ be the valuation ring and maximal ideal, respectively, and suppose that the residue field $\mathcal{O}_K/\mathfrak{m}_K$ is finite, with q elements.

Theorem 1. *If $f(x) = a_0x^{n_0} + a_1x^{n_1} + \cdots + a_kx^{n_k}$ is a polynomial with $k+1$ nonzero coefficients $a_i \in K^*$, then f has at most q^k distinct zeros in K.*

This upper bound is sharp: if K is $\mathbb{F}_q((T))$ with the usual discrete valuation $v : K^* \to \mathbb{Z}$, if $V \subset K$ is an $\mathbb{F}_q$-subspace of dimension k, and if $c \in K$ is nonzero, then the polynomial $f(x) := c\prod_{\alpha\in V}(x-\alpha)$ has the form $a_0x + a_1x^q + \cdots + a_kx^{q^k}$ for some $a_0, a_1, \ldots, a_k \in K^*$.

Theorem 1 is the case $d = 1$ of the following generalization, which bounds the number of distinct zeros of *bounded degree*. Let $\mu(n)$ be the Möbius μ-function.

Theorem 2. *Fix $d \geq 1$. If $f(x) = a_0x^{n_0} + a_1x^{n_1} + \cdots + a_kx^{n_k}$ is a polynomial with $k+1$ nonzero coefficients $a_i \in K^*$, then the number of distinct zeros of f in $\overline{K}$ of degree at most d over K is at most $\sum_{j=1}^{d}\sum_{i|j} q^{ik}\mu(j/i)$.*

This upper bound is sharp as well, for every q, k, and d. Let $K = \mathbb{F}_q((T))$ and v be as before. Let $\mathbb{F}$ be a finite field containing $\mathbb{F}_{q^i}$ for $i \leq d$. Let $V \subset \mathbb{F}((T))$ be a k-dimensional $\mathbb{F}$-vector space that is $\mathrm{Gal}(\mathbb{F}/\mathbb{F}_q)$-stable (or equivalently, has an $\mathbb{F}$-basis of elements of K). Then equality is attained in Theorem 2 for $f(x) := c\prod_{\alpha\in V}(x-\alpha)$ for any $c \in K^*$. (The inner sum in Theorem 2 performs the inclusion-exclusion to count zeros of *exact* degree j.)

We make no claim that these are the only polynomials that attain equality; in fact there are many others. For example, if K, V, and f are as in the previous paragraph, and if the $\mathbb{F}$-basis of V consists of elements of K of distinct valuation, with all these valuations divisible by a single integer $e \geq 1$, then $f(x^e)$ also attains equality, as a short argument involving Hensel's lemma shows. Other examples can be constructed using the observation that if $f(x) \in K[x]$ has N zeros in a

Received February 25, 1998.

This research was supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship.

given field extension L of K, one of which is 0, then the same holds for $x^m f(1/x)$ when $m > \deg f$.

Remark. H. W. Lenstra, Jr. [Le1] proves related facts for finite extensions L of $\mathbb{Q}_p$, using very different methods. One of his results is that for any such L and any positive integer k, there exists a positive integer $B = B(k, L)$ with the following property: if $f \in L[x]$ is a nonzero polynomial with at most $k+1$ nonzero terms and $f(0) \neq 0$, then f has at most B zeros in L, *counted with multiplicities*. His bound $B(k, L)$ is explicit, but almost certainly not sharp. Finding a sharp bound seems difficult in general, although Lenstra does this for the case $k = 2$ and $L = \mathbb{Q}_2$ (the bound then is 6). He also applies his local result to bound uniformly the number of factors of given degree over number fields. In [Le2] he shows that if f is represented sparsely, then these factors can be found in polynomial time.

Remark. We cannot count multiplicities in either of *our* theorems and hope to obtain a bound depending only on k and K (and d, for Theorem 2), because of examples like $f(x) = (1+x)^{q^m}$ with $m \to \infty$. Requiring that f not be a p-th power would not eliminate the problem, because one could also take $f(x) = (1+x)^{q^m+1}$.

2. Proof of Theorem 1

By a *disk* in a valued field K, we mean either an "open disk" $D(x_0, g) := \{x \in K : v(x - x_0) > g\}$, or a "closed disk" $\overline{D}(x_0, g) := \{x \in K : v(x - x_0) \geq g\}$ where $x_0 \in K$ and $g \in G$.

Let $\sigma_1, \sigma_2, \ldots, \sigma_t$ be the non-vertical segments of the Newton polygon of f. Let $-g_j \in G \otimes \mathbb{Q}$ be the slope of σ_j. If $e_1, e_2, \ldots, e_r$ are the exponents of the monomials in f corresponding to points on a given σ_j, define N_j as the largest integer for which the images of $(1+x)^{e_1}$, $(1+x)^{e_2}$, $\ldots$, $(1+x)^{e_r}$ in $\mathbb{F}_p[x]/(x^{N_j})$ are linearly *dependent* over $\mathbb{F}_p$. We say that the σ_j are in a *proper order* if $N_1 \geq N_2 \geq \cdots \geq N_t$. This particular ordering is crucial to the proof, but it is hard to motivate its definition. It was discovered by analyzing proofs of many special cases of Theorem 1. For instance, if the Newton polygon of f has k non-vertical segments (each associated with exactly two exponents), then the segments are being ordered according to the p-adic absolute values of their horizontal lengths.

Lemma 3. *Let L be a field of characteristic $p > 0$ with a valuation $v : L^* \to G$. Suppose $f(x) = a_0 x^{n_0} + a_1 x^{n_1} + \cdots + a_k x^{n_k} \in L[x]$ with each a_i nonzero. List the segments of the Newton polygon of f in a proper order as above. Fix u and let $-g_u \in G \otimes \mathbb{Q}$ be the slope of the u-th segment σ_u. Suppose $r \in L$ is* not *a zero of f, and $v(r) = g_u$. Let S be the set of zeros of f in L lying inside $D(r, g_u)$. Then $\#\{v(\alpha - r) : \alpha \in S\} \leq k + 1 - u$.*

Proof. Replacing $f(x)$ by $cf(rx)$ for suitable $c \in L^*$, we may reduce to the case in which $r = 1$, $g_u = 0$, $f \in \mathcal{O}_L[x]$, and $f \bmod \mathfrak{m}_L$ is nonzero. Write $f(1+x) = \sum_{j=0}^{n_k} b_j x^j$, and let M be the smallest integer for which b_M is nonzero modulo $\mathfrak{m}_L$. By definition of N_u, $f(1+x) \not\equiv 0 \bmod (\mathfrak{m}_L, x^{N_u+1})$. Hence $M \leq N_u$.

For each $i \leq u$, we have $N_u \leq N_i$, so $M \leq N_i$, and there is some $\mathbb{F}_p$-linear relation in $\mathbb{F}_p[x]/(x^M)$ between the $(1+x)^e$ for the exponents e associated to σ_i. The subspace of $\mathbb{F}_p[x]/(x^M)$ spanned by the $(1+x)^e$, where e ranges over *all* the exponents in f, then has dimension at most $(k+1)-u$, since the u relations above are independent, the largest e involved in each relation being distinct from the others. It follows that the $\mathbb{F}_p$-subspace of K spanned by $b_0, b_1, \dots, b_{M-1}$ is at most $(k+1-u)$-dimensional. Then $\#\{v(b_i) : 0 \leq i < M \text{ and } b_i \neq 0\} \leq k+1-u$, because nonzero elements of distinct valuations are automatically $\mathbb{F}_p$-independent. The left endpoints of the negative slope segments of the Newton polygon of $f(1+x)$ correspond to b_i of distinct valuations for $i < M$, so there are at most $k+1-u$ such segments. Hence at most $k+1-u$ positive elements of G can be valuations of zeros of $f(1+x)$, which is what we needed to prove. $\square$

Remark. Note that there is no assumption on the residue field in Lemma 3; L could even be algebraically closed.

Let S be any finite subset of a field L with valuation v. We associate a tree $\mathcal{T}$ to S as follows. (See [St] for arboreal terminology.) Let $\mathcal{T}$ be the Hasse diagram of the finite poset (ordered by inclusion) of nonempty sets of the form $S \cap D$ where D is a disk. Clearly $\mathcal{T}$ is a tree, whose leaves are the singleton subsets of S. We would obtain the same tree if we required the disks D to be open (resp. closed), since S is finite.

Suppose r and S are as in Lemma 3. Let $T_0 > T_1 > \cdots > T_\ell$ be the longest chain in $\mathcal{T}$. Then T_ℓ is a leaf, and $\#T_\ell = 1$. Choose $r_0 \in D(r, g_u) \setminus S$ closer to the element of T_ℓ than to any other element of S. For various $g > g_u$, the set $S \cap D(r_0, g)$ can equal $T_0, T_1, \dots, T_\ell$, or $\emptyset$. Hence

$$\#\{v(\alpha - r_0) : \alpha \in S\} \geq \ell + 1.$$

On the other hand, Lemma 3 applied to r_0 yields

$$\#\{v(\alpha - r_0) : \alpha \in S\} \leq k + 1 - u.$$

Combining these, we have that *the length $\ell = \ell(\mathcal{T})$ of the tree satsifies $\ell \leq k - u$.*

Suppose $S_0 \in \mathcal{T}$ is not a leaf (i.e. $\#S_0 > 1$), and let $g = \min\{v(s-t) : s, t \in S_0\}$, so that for any $s \in S_0$, $\overline{D}(s, g)$ is the smallest disk containing S_0. Then the children of S_0 in the tree are nonempty sets of the form $S \cap D(x_0, g)$ for some $x_0 \in \overline{D}(s, g)$. In particular the number of children is at most the size of the residue field of L.

Proof of Theorem 1. Let notation be as in Lemma 3, but take $L = K$. By the theory of Newton polygons, each nonzero zero of f has valuation equal to g_u for some u. Let us now fix u and let Z_u be the number of zeros in K of valuation

g_u. We may assume $g_u \in G$, since otherwise $Z_u = 0$. Then $\{x \in K : v(x) = g_u\}$ is the union of $q-1$ open disks D_j of the form $D(x_j, g_u)$. As above, the tree corresponding to the set of zeros in D_j has length at most $k-u$, and each vertex has at most q children. Hence the tree has at most q^{k-u} leaves, and $Z_u \leq (q-1)q^{k-u}$. Allowing for the possibility that 0 also is a zero of f, we find that the number of zeros of f in K is at most

$$1 + \sum_{u=1}^{t} Z_u \leq 1 + \sum_{u=1}^{t} (q-1)q^{k-u} \leq 1 + \sum_{u=1}^{k} (q-1)q^{k-u} = q^k.$$

□

3. Valuation theory

Before proving Theorem 2, we will need to recall some facts from valuation theory. We write (K, v) for a field K with a valuation v. We say that (L, w) is an *extension* of (K, v) if $K \subseteq L$ and $w|_K = v$. In this case, when we say that L has the same value group (resp. residue field) as K, we mean that the inclusion of value groups (resp. residue fields) induced from the inclusion of (K, v) in (L, w) is an isomorphism. Recall that any valuation on a field K admits at least one extension to any field containing K. An abelian group G is *divisible* if for all $g \in G$ and $n \geq 1$, the equation $nx = g$ has a solution x in G.

Proposition 4. *Any valued field can be embedded in another valued field having the same residue field, but divisible value group.*

Proof. Let $v : K^* \to G$ be the original valuation. If G is not already divisible, then there exists $g \in G$ and a prime number n such that $nx = g$ has no solution in G. Pick $\alpha \in K^*$ with $v(\alpha) = g$, and extend v to a valuation on $L = K(\alpha^{1/n})$. Let e and f denote the ramification index and residue class degree for L/K. Then $e = n$, and the inequality $ef \leq n$ (Lemma 18 in Chapter 1 of [Sch]) forces $f = 1$. An easy Zorn's lemma argument now shows that v extends to a valuation $v : M^* \to G \otimes \mathbb{Q}$ where M is an extension with the same residue field as K, but with divisible value group. □

Recall that (L, w) is called an *immediate extension* of (K, v) if

(1) (L, w) is an extension of (K, v);
(2) (L, w) has the same value group as (K, v); and
(3) (L, w) has the same residue field as (K, v).

Also recall that (K, v) is called *maximally complete* if it has no nontrivial immediate extensions.

Proposition 5. *Every valued field has a maximally complete immediate extension.*

Proof. This is an old result of Krull: see Theorem 5 of Chapter 2 in [Sch]. □

Proposition 6. *Suppose that (K, v) is maximally complete of characteristic $p > 0$, and that $\mathbb{F}_q$ is contained in the residue field. Then $\mathbb{F}_q$ can be embedded in K.*

Proof. Apply a suitable version of Hensel's lemma (combine Theorems 6 and 7 of Chapter 2 of [Sch]) to the factorization of $x^q - x$ over $\mathbb{F}_q$. □

Proposition 7. *Suppose that (K, v) is maximally complete of characteristic $p > 0$, with divisible value group G and with residue field $\mathbb{F}_q$. If $L \subset \overline{K}$ is a finite extension of K of degree n, then L is the compositum of $\mathbb{F}_{q^n}$ and K in $\overline{K}$.*

Proof. Extend v to L. Theorem 11 in Chapter 2 of [Sch] shows that L is maximally complete, and that $ef = n$ holds for L/K. Since G is divisible, there are no ordered abelian groups G' with $1 < (G' : G) < \infty$. Hence $e = 1$, $f = n$, and the residue field of L is $\mathbb{F}_{q^n}$. Proposition 6 implies that the subfield $\mathbb{F}_{q^n}$ of $\overline{K}$ is contained in L. But the compositum of the linearly disjoint fields $\mathbb{F}_{q^n}$ and K in $\overline{K}$ is already n-dimensional over K, so the compositum must equal L. □

Remark. Lenstra notes that if one is interested in proving Theorem 2 only for polynomials over $K_0 = \mathbb{F}_q((T))$, then one can circumvent the theory of maximally complete fields by choosing $\sigma \in \operatorname{Gal}(\overline{K}_0/K_0)$ that acts as $x \mapsto x^q$ on $\overline{\mathbb{F}}_q$, and by taking K to be the fixed field of σ. This K contains K_0, still has residue field $\mathbb{F}_q$, and satisfies the conclusion of Proposition 7.

4. Proof of Theorem 2

In proving Theorem 2, we may first apply Propositions 4 and 5 to assume that the value group G is divisible and that (K, v) is maximally complete (still with residue field $\mathbb{F}_q$). Let $\mathbb{F} = \mathbb{F}_{q^{d!}} \subset \overline{K}$. Proposition 7 shows that all elements of $\overline{K}$ of degree at most d over K lie inside the compositum $L := \mathbb{F} \cdot K$ of fields in $\overline{K}$. Extend v to L.

For each $g \in G$, choose $\beta_g \in K$ with $v(\beta_g) = g$. Now suppose $\overline{D} := \overline{D}(x_0, g)$ is a closed ball in L. Let I be the subgroup of $\operatorname{Gal}(L/K) \cong \operatorname{Gal}(\mathbb{F}/\mathbb{F}_q)$ that maps $\overline{D}$ into $\overline{D}$. Division by β_g induces an isomorphism of I-modules $\overline{D}(0, g)/D(0, g) \cong \mathbb{F}$, so the cohomology group $H^1(I, \overline{D}(0, g)/D(0, g))$ is trivial. The long exact sequence associated with the exact sequence

$$0 \to \frac{\overline{D}(0,g)}{D(0,g)} \to \frac{L}{D(0,g)} \to \frac{L}{\overline{D}(0,g)} \to 0$$

of I-modules shows that $\overline{D}$ contains an open disk $D(x_1, g)$ mapped to itself by I. We then have a bijection of I-sets $\phi_{\overline{D}} : \overline{D}/D(0, g) \to \mathbb{F}$ that maps the coset $y + D(0, g)$ to the residue class of $(y - x_1)/\beta_g$. We assume that the elements β_g and the maps $\phi_{\overline{D}}$ are fixed once and for all.

Now let g_u, r, and S be as in Lemma 3, and let $\mathcal{T}$ be the tree associated to S as in Section 2, so that $\ell(\mathcal{T}) \le k - u$. We now describe a labelling of the vertices of $\mathcal{T}$ by elements of $\mathbb{F}$. Recall that if $S_0 \in \mathcal{T}$ is not a leaf, and if $\overline{D} = \overline{D}(s, g)$ is the smallest disk containing S_0, then the children of S_0 are nonempty sets of the form $S \cap D(x_0, g)$ for some $x_0 \in \overline{D}$. Label each child by $\phi_{\overline{D}}(D(x_0, g))$. Note that the children of S_0 are labelled with *distinct* elements of $\mathbb{F}$. Finally, label the root of $\mathcal{T}$ with the residue of r/β_{g_u} in $\mathbb{F}^*$.

Let R be the set of all roots of f in L, and let $\mathbb{F}[X]_{<k}$ denote the set of polynomials of the form $a_0 + a_1 X + \cdots + a_{k-1} X^{k-1}$ with $a_i \in \mathbb{F}$. We now define a map $\Phi : R \to \mathbb{F}[X]_{<k}$. First, if $0 \in R$, define $\Phi(0) = 0 \in \mathbb{F}[X]_{<k}$. If $z \in R$ is nonzero, then $v(z) = g_u$ for some u. Let $T_0 > T_1 > \cdots > T_n$ be the maximal chain ending at $T_n = \{z\}$ in the tree $\mathcal{T}$ associated to $S := R \cap D(z, g_u)$. Define $\Phi(z) = X^{u-1} \sum_{i=0}^{n} \operatorname{label}(T_i) X^i$. Since $n \le \ell(\mathcal{T}) \le k-u$, we have $\Phi(z) \in \mathbb{F}[X]_{<k}$.

Lemma 8.

(1) *The map* $\Phi : R \to \mathbb{F}[X]_{<k}$ *is injective.*
(2) *If* $z \in R$ *is of degree* j *over* K, *then* $\Phi(z) \in \mathbb{F}_{q^j}[X]$.

Proof. To prove injectivity, we describe how to reconstruct z from $\Phi(z)$. If $\Phi(z) = 0$, then z must be 0. Otherwise its lowest degree monomial involves X^{u-1} where $v(z) = g_u$. Hence, assuming from now on that $z \neq 0$, we can reconstruct $v(z)$ from $\Phi(z)$. Next, the *coefficient* of X^{u-1} determines which (nontrivial) coset of $D(0, g_u)$ in $\overline{D}(0, g_u)$ z belongs to. The other coefficients uniquely determine a path ending at the leaf $\{z\}$ in the tree associated to this coset. Thus $\Phi(z)$ determines z.

For the second part, it suffices to show that if H is the subgroup of $\operatorname{Gal}(L/K)$ fixing $z \in R$, then H (or equivalently the isomorphic subgroup of $\operatorname{Gal}(\mathbb{F}/\mathbb{F}_q)$) fixes the coefficients of $\Phi(z)$ also. We may assume $z \neq 0$. Let $g_u = v(z)$, and let $T_0 > T_1 > \cdots > T_n = \{z\}$ be the maximal chain in the tree $\mathcal{T}$ associated to the coset $z + D(0, g_u)$ in which z lies. Since H preserves the coset $z + D(0, g_u)$, H fixes the label of T_0. Now suppose $1 \le i \le n$. The smallest disk containing T_{i-1} is of the form $\overline{D} := \overline{D}(z, g)$ for some $g > g_u$, so H is contained in the subgroup $I \subseteq \operatorname{Gal}(L/K)$ preserving this disk. The label of the child T_i is $\phi_{\overline{D}}(z + D(0, g))$, and $\phi_{\overline{D}}$ respects the action of $H \subseteq I$, so H fixes this label. This holds for all i, so H fixes all coefficients of $\Phi(z)$. $\square$

Lemma 8 shows that the number of zeros of f in $\overline{K}$ of degree at most d is less than or equal to the number of polynomials in $\mathbb{F}[X]_{<k}$ that are defined over $\mathbb{F}_{q^j}$ for some $j \le k$. The number of such polynomials defined over $\mathbb{F}_{q^j}$ but no subfield is $\sum_{i|j} q^{ik} \mu(j/i)$, by Möbius inversion. Theorem 2 follows upon summing over j.

Acknowledgements

I thank Hendrik Lenstra for bringing the problems to my attention, and for suggesting that q^k might be the correct bound for Theorem 1.

References

[Le1] H. W. Lenstra, Jr., *On the factorization of lacunary polynomials*, Number Theory in Progress, De Gruyter, 1998, (to appear).

[Le2] ———, *Finding small degree factors of lacunary polynomials*, Number Theory in Progress, De Gruyter, 1998, (to appear).

[Sch] O. Schilling, *The theory of valuations*, Mathematical Surveys, No. 4, Amer. Math. Soc., New York, 1950.

[St] R. Stanley, *Enumerative Combinatorics, Volume* 1, Cambridge University Press, 1997.

Dept. of Math., University of California, Berkeley, CA 94720-3840 USA
E-mail address: poonen@math.berkeley.edu