ROOTS OF SPARSE POLYNOMIALS OVER A FINITE FIELD

An Undergraduate Research Scholars Thesis

by

ALEXANDER KELLEY

Submitted to the Undergraduate Research Scholars program
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by
Research Advisor: Dr. J. Maurice Rojas

May 2016

Major: Computer Science
Applied Mathematics

TABLE OF CONTENTS

ABSTRACT

Roots of Sparse Polynomials over a Finite Field

Alexander Kelley
Department of Computer Science
Department of Mathematics
Texas A&M University

Research Advisor: Dr. J. Maurice Rojas
Department of Mathematics

For a t-nomial $f(x) = \sum_{i=1}^{t} c_i x^{a_i} \in \mathbb{F}_q[x]$, we show that the number of distinct, nonzero roots of f is bounded above by $2(q-1)^{1-\varepsilon}C^{\varepsilon}$, where $\varepsilon = 1/(t-1)$ and C is the size of the largest coset in $\mathbb{F}_q^*$ on which f vanishes completely. Additionally, we describe a number-theoretic parameter depending only on q and the exponents a_i which provides a general and easily-computable upper bound for C. We thus obtain a strict improvement over an earlier bound of Canetti et al. which is related to the uniformity of the Diffie-Hellman distribution.

CHAPTER I

INTRODUCTION

Over the real numbers, the classical Descartes' Rule implies that the number of distinct, real roots of a t-nomial $f(x) = c_1 x^{a_1} + \cdots + c_t x^{a_t} \in \mathbb{R}[x]$ is less than $2t$, regardless of its degree. It is a natural algebraic problem to look for analogous sparsity-dependent bounds over other fields that are not algebraically closed. In [3], Canetti et al. derive the following analogue of Descartes' Rule for polynomials in $\mathbb{F}_q[x]$.

Theorem I.1. *([3], Lemma 7) For $f(x) = c_1 x^{a_1} + c_2 x^{a_2} + \cdots + c_t x^{a_t} \in \mathbb{F}_q[x]$ (with c_i nonzero), if $R(f)$ denotes the number of distinct, nonzero roots of f in $\mathbb{F}_q$, then*

$$R(f) \leq 2(q-1)^{1-1/(t-1)} D^{1/(t-1)} + O\left((q-1)^{1-2/(t-1)} D^{2/(t-1)}\right),$$

where

$$D(f) = \min_i \max_{j \neq i} \{\gcd(a_i - a_j, q-1)\}.$$

For $\vartheta \in \mathbb{F}_p^*$, the associated Diffie-Hellman distribution is defined by the random variable $(\vartheta^x, \vartheta^y, \vartheta^{xy})$ where x and y are uniformly random over $\{1, \ldots, p-1\}$. The Diffie-Hellman cryptosystem relies on the assumption that an attacker cannot easily determine ϑ^{xy} given the values of ϑ, ϑ^x, and ϑ^y. In [3], Canetti et al. showed that Diffie-Hellman distributions are very nearly uniform (which is an important property for the security of the cryptosystem), and the bound in Theorem I.1 was the central tool which powered their arguments.

Since then, the bound has been a useful tool for studying various number-theoretic problems; in [1] it was used to study the solutions of certain exponential congruences, and in [4] it was used to study the correlation of linear recurring sequences over $\mathbb{F}_2$. The main result of this paper is is a new bound (Theorem II.3 of Section 2 below) improving Theorem I.1 by removing the asymptotic term and replacing D by a smaller, intrinsic parameter.

CHAPTER II

STATEMENT OF RESULTS

More recently in [2], Bi, Cheng, and Rojas studied the computational complexity of deciding whether a t-nomial f has a root in $\mathbb{F}_q^*$. Along the way, they derive the following characterization of the roots of a sparse polynomial in $\mathbb{F}_q[x]$.

Theorem II.1. *([2], Theorem 1.1) For $f(x) = c_1x^{a_1} + c_2x^{a_2} + \cdots + c_tx^{a_t} \in \mathbb{F}_q[x]$, define $\delta(f) = \gcd(a_1, a_2, \ldots, a_t, q-1)$. The set of nonzero roots of f in $\mathbb{F}_q$ is the union of no more than*

$$2\left(\frac{q-1}{\delta}\right)^{1-1/(t-1)}$$

cosets of two subgroups $H_1 \subseteq H_2$ of $\mathbb{F}_q^*$, where

$$|H_1| = \delta,$$
$$|H_2| \geq \delta^{1-1/(t-1)}(q-1)^{1/(t-1)}.$$

This result does not immediately yield any bound on the number of roots $R(f)$ since there is no upper bound given for the size of the H_2-cosets. However, if for some reason we were assured that the set of roots was a union of only H_1-cosets, we could conclude

$$R(f) \leq \delta \cdot 2\left(\frac{q-1}{\delta}\right)^{1-1/(t-1)} = 2(q-1)^{1-1/(t-1)}\delta^{1/(t-1)},$$

which is an improvement on Theorem I.1 since it can be easily checked that $\delta(f) \leq D(f)$ always.

Theorem II.2. *For $f(x) = c_1x^{a_1} + \cdots + c_tx^{a_t} \in \mathbb{F}_q[x]$, define*

$$S(f) := \{k \,|\, (q-1) \,:\, \text{for all } i, \text{ there is a } j \neq i \text{ with } a_i \equiv a_j \bmod k\}.$$

If f vanishes completely on a coset of size k, then $k \in S(f)$.

Proof. For some generator g of $\mathbb{F}_q^*$, let $\alpha\langle g^{\frac{q-1}{k}}\rangle$ denote a coset of the unique subgroup of order k in $\mathbb{F}_q^*$, and let $\beta = \alpha^k$. The members of this coset are exactly the roots of the binomial $x^k - \beta$. So, f vanishes completely on this coset if and only if $(x^k - \beta) \mid f$, or equivalently if $f \equiv 0 \bmod (x^k - \beta)$.

To see when this happens, we view f in the ring $\mathbb{F}_q[x]/\langle x^k - \beta\rangle$. In this ring, we have the relation $x^k \equiv \beta$, so if each a_i has remainder r_i mod k, then

$$f \equiv c_1\beta^{\lfloor a_1/k\rfloor}x^{r_1} + \cdots + c_t\beta^{\lfloor a_t/k\rfloor}x^{r_t} \mod (x^k - \beta).$$

Now f might be identically zero (in this ring) since the r_i's are not necessarily distinct. However, there is one obvious barrier to this: if just one r_i is unique, then f in particular contains the nonzero monomial $(c_i\beta^{\lfloor a_i/k\rfloor})x^{r_i}$. $f \equiv 0$ requires that each remainder r_i has at least one "partner" $r_j = r_i$ so that monomials can cancel. Therefore $(x^k - \beta) \mid f$ implies that, for each $i \in \{1, 2, \ldots, t\}$, there is some $j \neq i$ with $a_i \equiv a_j \bmod k$. □

Thus $S(f)$ lists the sizes of cosets on which f might possibly vanish completely. For example, if $a_1 = 0$ and the other exponents $a_{i>1}$ are all prime to $q-1$ then $S(f) = \{1\}$, and so it is structurally impossible for f to vanish completely on any nontrivial coset, regardless of choice of coefficients $c_i \in \mathbb{F}_q^*$. On the other hand whenever $k \in S(f)$, there is some choice of $c_i \in \mathbb{F}_q^*$ so that f does indeed vanish completely on a given coset of size k.

When $\max(S) < \delta^{1-1/(t-1)}(q-1)^{1/(t-1)}$, Theorem II.2 can be combined with Theorem II.1 to get a bound on $R(f)$ by ruling out the possibility of H_2-cosets. If $\max(S)$ is any larger, Theorem II.1 is no longer helpful; the most we can conclude is that $R(f) \leq |H_2|2\left(\frac{q-1}{\delta}\right)^{1-1/(t-1)} \leq \max(S)\cdot 2\left(\frac{q-1}{\delta}\right)^{1-1/(t-1)}$, which is worse than trivial ($R \leq q-1$). However, $S(f)$ turns out also to be independently useful for deriving sparsity-dependent bounds.

Theorem II.3. *Let $f(x) = c_1 x^{a_1} + \cdots + c_t x^{a_t} \in \mathbb{F}_q[x]$ (with c_i nonzero), and let $\delta(f)$ be defined as above, and let C denote the size of the largest coset in $\mathbb{F}_q^*$ on which f vanishes completely. If $R(f)$ denotes the number of distinct, nonzero roots of f in $\mathbb{F}_q^*$, then we have*

$$R(f) \leq 2(q-1)^{1-1/(t-1)} C^{1/(t-1)},$$

and furthermore if $C < \delta^{1-1/(t-1)}(q-1)^{1/(t-1)}$, then

$$R(f) \leq 2(q-1)^{1-1/(t-1)} \delta^{1/(t-1)}.$$

This result is a strict improvement on Theorem I.1, since $D(f)$ is in particular an upper bound for $S(f)$ and therefore also for $C(f)$. In fact, we can get another easily computable upper bound for $S(f)$ that is in general tighter than $D(f)$.

Proposition II.4. *For $f(x) = c_1 x^{a_1} + \cdots + c_t x^{a_t} \in \mathbb{F}_q[x]$ define the parameters*

$$\begin{aligned}
\delta(f) &= \gcd(a_1, a_2, \ldots, a_t, q-1) \\
D(f) &= \min_i \max_{j \neq i} \{\gcd(a_i - a_j, q-1)\} \\
Q(f) &= \gcd_i \operatorname*{lcm}_{j \neq i} (\gcd(a_i - a_j, q-1)) \\
K(f) &= \min_i \max_{j \neq i} \{\gcd(a_i - a_j, Q)\}
\end{aligned}$$

These parameters relate to $S(f)$ as follows.

- *$\delta \in S$.*
- *For all $k \in S$, $k \mid Q$.*
- *D, Q, and K are all upper bounds for S, and $K \leq \min(D, Q)$.*

CHAPTER III

PROOFS

The general strategy employed here (and in both [2] and [3]) for obtaining sparsity-dependent bounds on $R(f)$ can be loosely sketched as follows. Consider integers e prime to $q-1$, which have the property that the map $x \mapsto x^e$ is a bijection on $\mathbb{F}_q^*$. This means that $x \mapsto x^e$ simply permutes the nonzero roots of a polynomial, so $R(f(x)) = R(f(x^e))$. Furthermore, $f(x^e)$ is equivalent (as a mapping on $\mathbb{F}_q^*$) to any $g(x) = c_1 x^{b_1} + \cdots + c_t x^{b_t}$ with $b_i \equiv ea_i \bmod (q-1)$. Thus the basic idea is to find some e so that the remainders of $ea_i \bmod (q-1)$ are all small, yielding a g of small degree, and so $R(f) = R(g) \leq deg(g)$.

The following lemma, a fact about the geometry of numbers, will be our main tool for achieving the desired degree reduction.

Lemma III.1. *Fix the natural numbers $a_1, a_2, \ldots, a_t, N$. If $n \leq N/\gcd(a_1, a_2, \ldots, a_t, N)$, there is an $e \in \{1, 2, \ldots, n-1\}$ and a $v \in N\mathbb{Z}^t$ so that*

$$0 < \max_{1 \leq i \leq t} |ea_i + v_i| \leq N/n^{1/t}.$$

Proof. Consider the vectors $l_i = i(a_1, \ldots, a_t) = (ia_1, \ldots, ia_t) \in (\mathbb{R}/N\mathbb{Z})^t$ for $i \in \{1, 2, \ldots n\}$. Let $\|\cdot\|_\infty$ denote the standard infinity norm on $\mathbb{R}^t$. We wish to view these vectors geometrically as points in $\mathbb{R}^t$, but they are only defined up to equivalence in $(\mathbb{R}/N\mathbb{Z})^t$, so define

$$\|l\|_N = \min_{v \in N\mathbb{Z}^t} \|l + v\|_\infty,$$

which gives the smallest norm of any representative of the equivalence class $l + N\mathbb{Z}^t$ viewed as a point in $\mathbb{R}^t$ (equivalently, $\|l\|_N$ gives the distance from l to the nearest lattice point in $N\mathbb{Z}^t$). Suppose that

$$d = \min_{i \neq j} \|l_j - l_i\|_N.$$

Since the vectors are all at least d apart, the sets

$$B_i = \{x \in (\mathbb{R}/N\mathbb{Z})^t \,:\, \|x - l_i\|_N < d/2\}$$

are disjoint, so each l_i sits in its own personal box of volume d^t. We may choose to represent these n disjoint sets uniquely in the fundamental domain $[0,N)^t$, which has volume N^t. Therefore we have a total volume of $n \cdot d^t$ sitting in a volume of N^t; we conclude that $d \leq N/n^{1/t}$.

Note that the modular definition of distance is crucial here; consider instead n points in $[0,N]^t$ that are d-separated only in the standard l_∞ metric. A volume-packing argument becomes much more complicated in this case because the box around a point near the boundary may lie partly outside $[0,N]^t$ (it doesn't "wrap around"), and so some points do not absorb a full d^t worth of volume from $[0,N]^t$.

To finish, we find i, j (with $1 \leq i < j \leq n$) so that $\|l_j - l_i\|_N = d$ and set $l_e = l_{(j-i)} = (j-i)(a_1, \ldots, a_t) = l_j - l_i$. We have

$$\|l_e\|_N = \min_{v \in N\mathbb{Z}^t} \|(ea_1, \ldots, ea_t) + v\|_\infty \leq N/n^{1/t},$$

and e satisfies $1 \leq e \leq n-1$. The subgroup of $(\mathbb{Z}/N\mathbb{Z})^t$ generated by $(a_1, \ldots, a_n)$ has order $N/\gcd(a_1, \ldots, a_t, N) \geq n$. Since $0 < e < n$, $e(a_1, \ldots, a_n) \not\equiv (0, \ldots, 0) \in (\mathbb{Z}/N\mathbb{Z})^t$, which verifies that $\|l_e\|_N > 0$.

□

Lemma III.1 and its proof are extremely similar in spirit to the argument used by Canetti et al. in [3]. They also viewed the n vectors as points in $[0,N)^t$, but to find a pair of nearby points they partitioned the hypercube into $< n$ equally-sized sub-cubes and appealed to the pigeonhole principle. Here we were able to avoid this discretization of space which lead to the small asymptotic term appearing in Theorem I.1, which turns out to be unnecessary.

Proof of Theorem II.3. The second claim is immediate from Theorem II.1, since there can be no H_2-cosets of roots. We now prove the first claim.

Let $f(x) = c_1 x^{a_1} + c_2 x^{a_2} + \cdots + c_t x^{a_t} \in \mathbb{F}_q[x]$ with c_i nonzero, and let C denote the size of the largest coset in $\mathbb{F}_q^*$ on which f vanishes completely. For our purposes, we may assume that $a_1 = 0$, since otherwise we can write

$$\begin{aligned} f(x) &= x^{a_1} \tilde{f}(x) \\ \tilde{f}(x) &= c_1 + c_2 x^{a_2 - a_1} + \cdots + c_t x^{a_t - a_1}, \end{aligned}$$

showing that f has a root at zero, but its nonzero roots are just the roots of $\tilde{f}$, so $R(f) = R(\tilde{f})$ and $C(f) = C(\tilde{f})$. Therefore we continue assuming that $a_1 = 0$.

Consider $\delta(f) = \gcd(a_2, \ldots, a_t, q-1)$. The nonzero roots of $f(x) = c_1 + c_2 x^{a_2} + \cdots + c_t x^{a_t}$ are in one-to-one correspondence with the solutions of the system

$$\begin{aligned} c_1 + c_2 y^{a_2/\delta} + \cdots + c_t y^{a_t/\delta} &= 0 & y &\in \langle g^\delta \rangle \\ x^\delta &= y & x &\in \mathbb{F}_q^* \end{aligned}$$

If f has no roots in $\mathbb{F}_q^*$ then our bound is of course true, so suppose this system has at least one solution (y_0, x_0). Then in fact the system has at least δ solutions and f vanishes on the coset $\{x : x^\delta = y_0\}$. This allows us to conclude that $C \geq \delta$, and so $\left(\frac{q-1}{C}\right) \leq (q-1)/\gcd(a_2, \ldots, a_t, q-1)$. Therefore we can apply Lemma III.1 to find an $e \in \{1, 2, \ldots, \frac{q-1}{C} - 1\}$ and a $v \in (q-1)\mathbb{Z}^{t-1}$ so that

$$0 < \|(ea_2, \ldots, ea_t) + v\|_\infty \leq (q-1) \Big/ \left(\frac{q-1}{C}\right)^{1/(t-1)}.$$

Suppose $k = \gcd(e, q-1) = 1$. Then the mapping $x \mapsto x^e$ is a bijection on $\mathbb{F}_q^*$ that simply permutes the roots of f, thus $R(f(x)) = R(f(x^e))$. We are interested in $f(x^e)$ only as a function on $\mathbb{F}_q^*$ (rather than as a formal object in $\mathbb{F}_q[x]$), and since $\mathbb{F}_q^*$ is a group of order $(q-1)$, this function is not changed by shifting its exponents by $v_i \in (q-1)\mathbb{Z}$. Thus we may represent the function $f(x^e)$ as the (possibly Laurent) polynomial

$$f(x^e) = c_1 + c_2 x^{ea_2 + v_2} + \cdots + c_t x^{ea_t + v_t},$$

which satisfies

$$0 < M = \max_{1 \le i \le t} |ea_i + v_i| \le (q-1)^{1-1/(t-1)} C^{1/(t-1)}.$$

Again we are only interested in nonzero roots; note that $R(f(x^e)) = R(x^M f(x^e))$. Since $x^M f(x^e)$ is an honest polynomial in $\mathbb{F}_q[x]$ with non-negative exponents, we have $R(f) = R(x^M f(x^e)) \le \deg(x^M f(x^e)) \le 2M$ and we are done.

However, we might have $k = \gcd(e, q-1) > 1$. In this case $x \mapsto x^e$ is not a bijection - it takes $\mathbb{F}_q^* = \langle g \rangle$ to a smaller subgroup $\langle g^e \rangle = \langle g^k \rangle$ of size $\left(\frac{q-1}{k}\right)$. However, we can still cover $\mathbb{F}_q^*$ by k cosets of this subgroup. We have

$$R(f(x^e)) = \sum_{i=0}^{k-1} \frac{1}{k} R(f(g^i x^e)),$$

since $\mathbb{F}_q^* = \bigcup_{i=0}^{k-1} g^i \langle g^e \rangle$, and $x^e = y$ has k solutions for each $y \in \langle g^e \rangle$. Now we repeat our earlier tricks and arrive at

$$R(f) \le \sum_{i=0}^{k-1} \frac{1}{k} \deg(x^M f(g^i x^e)) \le 2M,$$

except that we must be careful that no $f(g^i x^e)$ is identically zero, preventing us from using degree to bound root number. If $f(g^i x^e)$ is identically zero then f vanishes completely on the coset $g^i \langle g^e \rangle = g^i \langle g^k \rangle$ of size $\left(\frac{q-1}{k}\right)$. However, since $k = \gcd(e, q-1) \le e < \left(\frac{q-1}{C}\right)$, we have

$$\frac{q-1}{k} > \frac{q-1}{\left(\frac{q-1}{C}\right)} = C,$$

so this is impossible by the definition of C; the cosets are too large for f to vanish on completely.

□

Proof of Proposition II.4. For $f(x) = c_1 x^{a_1} + \cdots + c_t x^{a_t} \in \mathbb{F}_q[x]$, we have the following equivalent definitions for S:

$$\begin{aligned} S(f) &= \{k \,|\, (q-1) \,:\, \forall i, \exists j \neq i \text{ such that } a_i \equiv a_j \bmod k\} \\ &= \{k \,|\, (q-1) \,:\, \forall i, \exists j \neq i \text{ such that } k \,|\, (a_i - a_j)\} \\ &= \{k \in \mathbb{N} \,:\, \forall i, \exists j \neq i \text{ such that } k \,|\, \gcd(a_i - a_j, q-1)\} \\ &= \bigcap_{i=1}^{t} \bigcup_{j \neq i} \{k \in \mathbb{N} \,:\, k \,|\, \gcd(a_i - a_j, q-1)\}. \end{aligned}$$

Clearly by the second definition we have $\delta(f) = \gcd(a_1, a_2, \ldots, a_t, q-1) \in S$. From the fourth definition we can get an upper bound for S by passing to the superset

$$\bigcap_{i=1}^{t} \bigcup_{j \neq i} \{k \in \mathbb{N} \,:\, k \leq \gcd(a_i - a_j, q-1)\} \supseteq S,$$

which has maximal element

$$D = \min_i \max_{j \neq i} \{\gcd(a_i - a_j, q-1)\}.$$

Alternatively, by considering a different lattice structure on the integers, we can pass to the superset

$$\bigcap_{i=1}^{t} \{k \in \mathbb{N} \,:\, k \,|\, \gcd(L_j, q-1))\} = \{k \in \mathbb{N} \,:\, k \,|\, Q\} \supseteq S,$$

where

$$\begin{aligned} L_i &= \operatorname{lcm}(a_i - a_1, \ldots, a_i - a_{i-1}, a_i - a_{i+1}, \ldots, a_i - a_t), \\ Q &= \gcd(L_1, \ldots, L_t, q-1) = \gcd_i \operatorname*{lcm}_{j \neq i} (\gcd(a_i - a_j, q-1)). \end{aligned}$$

Since we now know that, in the end, any member of S must be a divisor of Q, we can redefine S (equivalently) using a smaller ambient space:

$$\begin{aligned} S(f) &= \{k \,|\, Q : \forall i, \exists j \neq i \text{ such that } k \,|\, (a_i - a_j)\} \\ &= \bigcap_{i=1}^{t} \bigcup_{j \neq i} \{k \in \mathbb{N} : k \,|\, \gcd(a_i - a_j, Q)\} \\ &\subseteq \bigcap_{i=1}^{t} \bigcup_{j \neq i} \{k \in \mathbb{N} : k \leq \gcd(a_i - a_j, Q)\}. \end{aligned}$$

Considering the maximal element of this last superset of S gives the final upper bound

$$K = \min_i \max_{j \neq i} \{\gcd(a_i - a_j, Q)\},$$

which is obviously no larger than either D or Q. □

REFERENCES

[1] Balog, Antal; Broughan, Kevin A.; and Shparlinski, Igor E. "On the number of solutions of exponential congruences". Acta Arithmetica, 148 (1). pp. 93-103. ISSN 0065-1036 (print), 1730-6264 (online)

[2] Bi, Jingguo; Cheng, Qi; and Rojas, J. Maurice. "Sub-Linear Root Detection, and New Hardness Results, for Sparse Polynomials Over Finite Fields." proceedings of ISSAC (International Symposium on Symbolic and Algebraic Computation, June 26-29, Boston, MA), pp. 61-68, ACM Press, 2013.

[3] Canetti, Ran; Friedlander, John B.; Konyagin, Sergei; Larsen, Michael; Lieman, Daniel; and Shparlinski, Igor E. "On the statistical properties of Diffie-Hellman distributions." Israel J. Math. 120 (2000), pp. 2346.

[4] Friedlander, John, et al. "On the correlation of binary M-sequences." Designs, Codes and Cryptography 16.3 (1999): 249-256.