Zeroes of Polynomials Over Finite Fields

James Ax


Stable URL:
http://links.jstor.org/sici?sici=0002-9327%28196404%2986%3A2%3C255%3AZOPOFF%3E2.0.CO%3B2-B

*American Journal of Mathematics* is currently published by The Johns Hopkins University Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/jhup.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR
ZEROES OF POLYNOMIALS OVER FINITE FIELDS.*

By James Ax.

1. Introduction. Let \( F = F(X_1, \cdots, X_n) \) be a polynomial of (total) degree \( d \) over a finite field \( k \) with \( q \) elements. In section 3, making use of some ideas of B. Dwork in [2], we prove the following theorem:

If \( b \) is the largest integer (strictly) less than \( n/d \) then \( q^b \) divides the number of zeroes of \( F \).

E. Artin had conjectured that if \( F \) is homogeneous and \( n > d \) then \( F \) has a non-trivial zero. C. Chevalley proved this in [1] and even showed the hypothesis of homogeneity could be replaced by the weaker assumption of no constant term. E. Warning in [4], using a lemma of Chevalley, showed that even without this last assumption the characteristic \( p \) of \( k \) divides \( N(F) \), the number of zeroes of \( F \) (counting the trivial zero if \( F \) has no constant term). In Section 2 we give a quick proof of the Chevalley-Warning theorem independent of the Chevalley lemma. Nevertheless, there does not seem to be any simple proof of the fact that \( q \) divides \( N(F) \) if \( n > d \).

In Section 4 we exhibit, for each \( n \) and \( d \), a polynomial of degree \( d \) in \( n \) variables such that the highest power of \( p \) dividing the number of its zeroes is precisely \( q^b \) if \( b \) is the largest integer less than \( n/d \). While our result is the best possible divisibility relation in this sense, E. Warning in [4] showed that if \( n > d \) and if \( F \) has at least one zero then \( N(F) \) is at least \( q^{n-d} \).

The zeta function \( Z(H; t) \) of the hypersurface \( H \) defined by \( F \) over \( k \) is defined by

\[
Z(H; t) = \exp \left( \sum_{s=1}^{\infty} N_s t^s / s \right)
\]

where \( N_s \) is the number of zeroes of \( F \) in the field with \( q^s \) elements. Let \( \Omega \) denote the completion of the algebraic closure of the \( p \)-adic completion of the rationals, and let \( \mid \mid \) be the valuation on \( \Omega \) normed so that \( \mid p \mid = 1/p \).

The referee has shown how our result may be reformulated as the following statement, using the above notation.

Received February 15, 1963.
* Written with the partial support of NSF Grant G23834.
THEOREM. Each pole and each zero of $Z(H; t)$ has $p$-adic valuation at least $q^b$.

Indeed, it follows from the rationality of $Z(H; t)$ [2], that

$$\exp \left( \sum_{s=1}^{\infty} N_s t^s / s \right) = \prod_{i} (1 - \alpha_i t) / \prod_{j} (1 - \beta_j t)$$

where the $\alpha_i$, $\beta_j$ are algebraic integers, $\alpha_i \neq \beta_j$, for $i$ and $j$ ranging over finite sets. By logarithmic differentiations we obtain

$$\sum_{s=1}^{\infty} N_s t^{s-1} = \sum_{j} \beta_j (1 - \beta_j t)^{-1} - \sum_{i} \alpha_i (1 - \alpha_i t)^{-1}.$$ 

If we now assume our result, $|N_s| \leq |q^{bs}| = q^{-b s}$, so that the left side converges in $\Omega$ for $|t| < q^b$, then $|\alpha_i|, |\beta_j| \leq q^{-b}$ which verifies the referee’s statement. The converse follows from

$$N_t = \sum_{j} \beta_j - \sum_{i} \alpha_i.$$

Throughout this paper, $F, N(F), n, d, b, k, q,$ and $p$ are above. $Z$ denotes the integers, $Z_+$ the nonnegative integers. If $u = (u_1, \ldots, u_r) \in (Z_+)^r$, $X^u$ denotes the monomial $\prod_{i=1}^{r} X_i^{u_i}$ and we define height $u = \text{degree} X^u = \sum_{i=1}^{r} u_i$.

2. Quick proof of the Chevalley-Warning theorem. Since each element of $k$ is a $q-1$ root of unity or zero, we have for each $x \in k^n$ that $1 - F(x)^{q-1} = 1$ if $F(x) = 0$, zero otherwise. Summing over $x \in k^n$, we have (in $k$)

$$N(F) = \sum (1 - F(x)^{q-1}) = -\sum F(x)^{q-1}.$$ 

Now $F^{q-1}$, being of degree $d(q-1)$ is a $k$-linear combination of monomials of degree at most $d(q-1)$. If $X^u$ is such a monomial, we compute

$$\sum_{x \in k^n} x^u = \prod_{i=1}^{n} \sum_{x_i \in k} x_i^{u_i} = \prod_{i=1}^{n} X(u_i)$$

where $Y(u_i) = q-1 = -1$ if $u_i$ is positive multiple of $q-1$, zero otherwise. If $d < n$, then height $u \leq d(q-1) < n(q-1)$ which implies that the sum in (2) is zero. Hence, the sum in (1) is zero, i.e., $N(F) \equiv 0 \text{ mod } p$.

3. Proof of the theorem. Let $q = p^f$, $Q_p$ be the $p$-adic completion of the rationals, and $K$ the unique unramified extension of $Q_p$ of degree $f$. Then

---

\[2\] This equation is the essential fact in our proof as in Warning’s. We then proceed directly to the result in a way suggestive of certain manipulations in the sequel.
the residue class field of \( K \) is \( k \). Let \( T \) denote the set of Teichmüller representatives of \( k \) in \( K \); let \( T^* = T - \{0\} \), the \( q - 1 \) roots of unity. Let \( \xi \) be a primitive \( p \)-th root of unity. If \( \alpha \) is an integer of \( \mathbb{Q}_p \), \( \xi^\alpha \) is defined to be the \( \xi^\alpha \) if \( \alpha \in \mathbb{Z}_p \) is congruent to \( \alpha \) modulo \( p \). Letting \( S \) denote the trace of \( K \) over \( \mathbb{Q}_p \) we define

\[
C = \sum_{m=0}^{q-1} c(m) U^m
\]

to be the unique polynomial of degree \( q - 1 \) with coefficients in \( K(\xi) \) such that \( C(t) = \xi^{g(t)} \) for all \( t \in T \). Summing \( C(t) t^{-j} \) over \( t \in T^* \) we find

\[
(3) \quad c(j)(q-1) = g(j) \quad \text{for} \quad 0 < j < q - 1 \quad \text{where the Gauss sum} \quad g(j) \quad \text{is defined for} \quad 0 \leq j < q - 1 \quad \text{by}
\]

\[
g(j) = \sum t^{-j} \xi^{g(t)} \quad (t \in T^*).
\]

Summing \( C(t) \) over \( t \in T^* \) and using that the trace function is not identically zero on a finite field, we find

\[-1 = g(0) = (q-1)(c(0) + c(q-1)).\]

Since

\[(3') \quad c(0) = 1,\]

we have

\[(3'') \quad c(q-1)(q-1) = -q.\]

If \( 0 \leq j \leq q - 1 \), let \( j_i \) for \( i = 0, \ldots, f - 1 \) be such that \( 0 \leq j_i \leq p - 1 \) and \( j = \sum_{i=0}^{f-1} j_i p^i \). We set \( \sigma(j) = \sum_{i=0}^{f-1} j_i \), \( \rho(j) = \prod_{i=0}^{f-1} j_i! \) and \( \lambda = \xi - 1 \). Then Stickelberger's congruence [3] (and [2] for further reference),

\[
g(j) \rho(j) / \lambda^{\sigma(j)} \equiv -1 \mod \lambda \quad \text{for} \quad 0 \leq j < q - 1
\]

together with (3), (3'), and (3'') certainly imply

\[(4) \quad c(j) \equiv 0 \mod \lambda^{\sigma(j)} \quad \text{for} \quad 0 \leq j \leq q - 1.\]

The map \( x \to \xi^{g(x)} \) is a non-trivial character of the additive group of the integers of \( K \), trivial on the maximal ideal of the integers of \( K \). Thus the map \( \beta \) from \( k \) to the \( p \)-th roots of unity defined by \( \beta(x) = C(t) \) for \( x \in k \) and \( t \) the Teichmüller representative of \( x \) is a non-trivial character of the additive group of \( k \). If \( u \in k \), then \( \sum \beta(x_0u) = q \) if \( u = 0 \), zero otherwise where the sum if over \( x_0 \in k \). It follows that

\[
qN(F) = \sum \beta(x_0 F(x_1, \ldots, x_n)) \quad ((x_0, \ldots, x_n) \in k^{n+1}).
\]
Let

\[ F = \sum w a(w) x^w \quad (w \in W) \]

where \( W \) is the set of \( w \in (\mathbb{Z}_+)^n \) such that height \( w \leq d \). We have

\[ qN(F) = \sum_{w \in W} \prod u \beta(a(w)x^w) \quad (x = (x_0, \ldots, x_n) \in k^{n+1}) \]

where if \( w = (w_1, \ldots, w_n) \in (\mathbb{Z}_+)^n \) then \( w' = (1, w_1, \ldots, w_n) \in (\mathbb{Z}_+)^{n+1} \). If \( A(w) \) is the Teichmüller representative of \( a(w) \) for each \( w \in W \), then

\[ qN(F) = \sum_{w \in W} \prod (A(w)t^w) \quad (t = (t_0, \ldots, t_n) \in T^{n+1}) \]

and so

\[
qN(F) = \sum_{m \in M} \sum_{w \in W} \sum_{m=0}^{q-1} c(m) A(w)^m t^m w'
\]

where \( M \) is the set of functions on \( W \) with values from the integers \( 0, 1, \ldots, q-1 \). Setting \( a(m) = \prod_{w \in W} A(w)^m w \in T, \ e(m) = \sum_{w \in W} m(w) w' \) for \( m \in M \), we may rewrite (5) as

\[(5') \quad qN(F) = \sum_{m \in M} a(m) \prod_{w \in W} c(m(w)) \sum_{t \in T^{n+1}} t^{e(m)}.
\]

If \( v \in (\mathbb{Z}_+)^r \) we write \( q-1 \mid v \) if there exists \( u \in (\mathbb{Z}_+)^r \) such that \( v = (q-1)u \) and \( q-1 \nmid v \) otherwise. Let \( m \) be an arbitrary element of \( M \). Then we easily compute

\[(6) \quad \sum_{t \in T^{n+1}} t^{e(m)} = 0 \quad \text{if} \quad q-1 \mid e(m) \]

and

\[(6') \quad \sum_{t \in T^{n+1}} t^{e(m)} = q^{n+1} \quad \text{if} \quad e(m) = (0, \ldots, 0).
\]

We now assume \( q-1 \mid e(m) \) and \( e(m) \neq (0, \ldots, 0) \), i.e., \( m(w) \neq 0 \) for some \( w \in W \). Let \( e(m) = \sum_{w \in W} m(w) w \) and let \( s \) be the number of non-zero entries in \( e(m) \), \( 0 \leq s \leq n \). We have

\[(6'') \quad \sum_{t \in T^{n+1}} t^{e(m)} = (q-1)^{s+1} q^{n-s} \quad \text{if} \quad q-1 \mid e(m) \neq (0, \ldots, 0) \]

taking into account that the first entry of \( e(m) \), \( \sum_{w \in W} m(w) \), is a non-zero multiple of \( q-1 \). For each \( w \in W \), let \( m_i(w) \) for \( i = 0, \ldots, f-1 \) be such that \( m(w) = \sum_{i=0}^{f-1} m_i(w) p^i \) and \( 0 \leq m_i(w) \leq p-1 \). We extend the definition of \( m_z(w) \) to all \( z \in Z \) by letting \( m_z(w) = m_r(w) \) if \( r \) is the least non-negative
residue of $z$ modulo $f$ and define for each $j = 0, \ldots, f - 1$ the function $m^{(j)} \in M$ by

$$m^{(j)}(w) = \sum_{t=0}^{f-1} m_{t-j}(w) p^t.$$

Using $t^a = t$ for all $t \in T$, we readily compute that the effect of substituting $m^{(j)}$ for $m$ in the sum in (6'') is the same as if we formally substitute $t^{p^j}$ for $t$, i.e., no change since $t \mapsto t^p$ is a permutation of $T$. We deduce from the mutually exclusive (6), (6') and (6'') that $q - 1 | e(m^{(j)})'$, and the number of non-zero entries of $e(m^{(j)})$ is again $s$ for each $j = 0, \ldots, f - 1$. This yields the inequalities

$$s(q - 1) \leq \text{height } e(m^{(j)}) = \text{height } \sum_{w \in W} m^{(j)}(w) w \leq d \sum_{w \in W} m^{(j)}(w).$$

Since $\sum_{w \in W} m^{(j)}(w)$, the first entry of $e(m^{(j)})'$, is a multiple of $q - 1$ we conclude

$$(s/d)^*(q - 1) \leq \sum_{w \in W} m^{(j)}(w),$$

where $(y)^*$ means the smallest integer not less than $y$. Summing this relation over $j = 0, \ldots, f - 1$, using (7) and interchanging order of summation twice we obtain

$$f(s/d)^*(q - 1) \leq \sum_{w \in W} \sum_{t=0}^{f-1} p^t \sum_{j=0}^{f-1} m_{t-j}(w).$$

Thus with $\sigma$ as used in (4)

$$f(s/d)^*(q - 1) \leq \sum_{w \in W} \sum_{t=0}^{f-1} p^t \sigma(m(w)).$$

So

$$f(p - 1)(s/d)^* \leq \sum_{w \in W} \sigma(m(w))$$

which in view of (4) and the fact that $p$ divides $\lambda^{p-1}$ implies the exponent of the highest power of $q$ dividing $\prod_{w \in W} c(m(w))$ is at least $(s/d)^*$. Combining this with (6), (6'), and (6'') we see from (5') that

$$q^r \mid q(N(F)) \text{ if } r = \min r(s)$$

where

$$r(s) = (s/d)^* + n - s, \quad s = 0, 1, \ldots, n.$$

Now

$$h \geq (s + h)/d - (s/d)^*, \quad h \in \mathbb{Z}_+$$

since in going from $h$ to $h + 1$ the left side increases by one while the right side increases by at most one. Substituting $h = n - s$ in the relation and
using \( b + 1 = (n/d)^* \) we see from (9) that \( r(s) \geq b + 1 \) for \( s = 0, \cdots, n \).

By (8) \( q^b \) divides \( N(F) \).

**Corollary.** If \( F_i \) is a polynomial in \( n \) variables of degree \( d_i \) for \( i = 1, \cdots, j \) then the number \( N \) of common zeroes of the \( F_i \) is divisible by \( q^b \) if \( n > b \sum_{i=0}^j d_i \).

**Proof.** A standard combinatorial argument shows

\[
N = -\sum_S (-1)^{#S} N( \prod_{i \in S} F_i )
\]

where the sum is over all non-empty subsets \( S \) of the set of integers \( 1, \cdots, j \) and where \( #S \) = number of elements of \( S \). The corollary follows from the theorem since for each \( S \),

\[
\text{degree } \prod_{i \in S} F_i \leq \sum_{i=1} \sum_{i=1}^j d_i.
\]

4. **Examples.** If \( a \in \mathbb{Z}_+, a > 0 \) we define

\[
G_{a,d}(X_1, \cdots, X_{ad}) = X_1 \cdots X_d + \cdots + X_{(a-1)d+1} \cdots X_{ad}
\]

and assert that the highest power of \( p \) dividing \( N(G_{a,d}) \) is \( q^{a-1} \). Now \( N(G_{1,d}) = q^d - (q - 1)^d \). \( N(G_{a+1,d}) = N(G_{a,d}) \) times the number of zeroes of \( X_{(a-1)d+1} \cdots X_{ad} \) (in \( k^d \)) plus the number of non-zeroes of \( G_{a,d} \) (in \( k^{ad} \)) times the (constant) number of representations of a non-zero element of \( k \) by \( X_{(a-1)d+1} \cdots X_{ad} \) (in \( k^d \)), i.e.,

\[
N(G_{a+1,d}) = N(G_{a,d})N(G_{1,d}) + \left( q^{ad} - N(G_{a,d}) \right)(q-1)^{d-1}
\]

(10)

\[
= qN(G_{a,d})\left(q^{d-1} - (q-1)^{d-1}\right) + q^{ad}(q-1)^{d-1}.
\]

For \( d > 1 \) this yields our assertion recursively; for \( d = 1 \) our assertion is immediate. If \( n = bd + h \) with \( 0 < h \leq d \) (so that \( b \) is largest integer less than \( n/d \)) we set

\[
F(X_1, \cdots, X_n) = G_{b,d}(X_1, \cdots, X_{bd}) \text{ if } h = 1,
\]

\[
F(X_1, \cdots, X_n) = G_{b,d}(X_1, \cdots, X_{bd}) + X_{bd+1} \cdots X_n \text{ if } h > 1.
\]

We assert that the highest power of \( p \) dividing \( N(F) \) is \( q^b \).

If \( h = 1 \) this follows from our previous assertion since in this case \( N(F) = qN(G_{b,d}) \). If \( h > 1 \) our previous assertion still yields the desired result since by reasoning similar to that used in establishing (10) we have

\[
N(F) = qN(G_{b,d})\left(q^{h-1} - (q-1)^{h-1}\right) + q^{bd}(q-1)^{h-1}.
\]
REFERENCES.


