# Exponential sums and the Chevalley-Warning theorem

C. Douglas Haessig

December 9, 2010

**Abstract**

These notes consist of an expanded version of a talk at the Michigan Number Theory Day on November 12, 2010.

## 1 Introduction

The Chevalley-Warning theorem has evolved into many different branches in number theory, all of which deserve their own lecture. I will only discuss a small part of only one of these branches. We begin by recalling the statement of Artin's conjecture, which deals with the following topic.

## 2 Solutions of polynomials over finite fields

Let $\mathbb{F}_q$ be the finite field with $q$ elements of characteristic $p$. Given polynomials $f_1, \dots, f_r \in \mathbb{F}_q[x_1, \dots, x_n]$ we may consider the number of simultaneous solutions over $\mathbb{F}_q$,

$$N(f_1, \dots, f_r) := \#\{x \in \mathbb{F}_q^n \mid f_1(x) = \cdots = f_r(x) = 0\}.$$

For notational convenience, we will often denote this simply by $N$.

In 1909 Dickson [4] conjectured that if $n > d_1 + \cdots + d_r$ and $N \geq 1$, then $N \geq 2$. From what I understand, around 1935, Emil Artin, while studying what is now known as quasi-algebraically closed fields (Lang-Tsen theory), suggested this problem to his student Ewald Warning. He also explained the problem to Claude Chevalley, who was visiting Artin at Göttingen at that time. In 1936, Chevalley proved the conjecture, with Warning's proof following also in 1936. However, Warning proved more by showing that $N \equiv 0$ modulo $p$, or equivalently, $p^\mu \mid N$ for some positive integer.[1] This result is now called the Chevalley-Warning theorem. It also suggests the following question.

**Question 2.1.** *How large is $\mu$?*

In 1964, using $p$-adic methods created by Dwork, Ax [3] showed that not only does $p$ divide $N$ but that $q^\mu \mid N$ with $\mu$ the least nonnegative integer greater than $(n - \sum d_i)/(\sum d_i)$. This lower bound was later improved by Katz [7] in 1971, also using Dwork's methods, to $(n - \sum d_i)/\max d_i$. This is best possible in the sense that given any positive integers satisfying $n > d_1 + \cdots + d_r$, there exist $r$ homogeneous polynomials in $n$ variables of degrees $d_i$ such that $N = q^\mu m$, with $p \nmid m$ and $\mu$ equal to Katz's lower bound. Recently, Moreno and Moreno have improved Katz's lower bound; see [9]. In another direction, Esnault [5] has explored replacing the hypothesis of the Chevalley-Warning theorem with a geometric condition.

We now wish to reinterpret this result in terms of exponential sums.

## 3 Character sum interpretation

For a fixed prime number $p$ and writing each rational number as $\frac{a}{b} = p^s \frac{m}{n}$ with $p \nmid m, n$, we may define the valuation $ord_p(a/b) := s$. If $q = p^r$, then $ord_q(\cdot) := \frac{1}{r} ord_p(\cdot)$. Thus, the above results of Ax-Katz say $ord_q N(f_1, \dots, f_r) \geq \mu$.

This valuation defines a metric $|\cdot|_p := \frac{1}{p^{ord_p(\cdot)}}$ on the rational numbers. Completing the rational numbers with this metric gives the $p$-adic numbers $\mathbb{Q}_p$. As you know, the usual absolute value $|\cdot|$ may be used to complete the rational numbers to obtain the real numbers $\mathbb{R}$, whose elements may all be describe by decimal expansions

[1] Warning proved more than this in his paper. With $r = 1$ and $n > d$, he showed $N \geq q^{n-d}$. See the recent paper by Heath-Brown [6] for an improvement.

$\sum_{i=-M}^{\infty} a_i 10^{-i}$ with $a_i \in \{0, 1, \ldots, 9\}$. The $p$-adics numbers have a similar description; the elements of $\mathbb{Q}_p$ have the "decimal expansion" $\sum_{i=-M}^{\infty} a_i p^i$ with $a_i \in \{0, 1, \ldots, p-1\}$. For a later application, we state the following result from $p$-adic analysis.

**Lemma 3.1.** *The series $\sum a_i T^i$ converges for $ord_q(T) > -\mu$ if and only if for every $\mu' < \mu$, $ord_q(a_i) \geq \mu' i$ for all $i$ sufficiently large.*

For a fixed primitive $p$-th root of unity, define the additive character $\psi(x) := \zeta_p^x$ on $\mathbb{F}_p$. Observe, for $x_0 \in \mathbb{F}_q$,

$$\sum_{x \in \mathbb{F}_q} \psi \circ Tr_{\mathbb{F}_q/\mathbb{F}_p}(x_0 x) = \begin{cases} q & \text{if } x_0 = 0 \\ 0 & \text{if } x_0 \neq 0. \end{cases}$$

This observation allows us to count zeros of a polynomial $f \in \mathbb{F}_q[x_1, \ldots, x_n]$ over $\mathbb{F}_q$:

$$\sum_{x_0, x_1 \ldots, x_n \in \mathbb{F}_q} \psi \circ Tr(x_0 f(x_1, \ldots, x_n)) = qN(f).$$

Denoting the left-hand side by $S(f)$, we see that

$$ord_q(N(f)) = -1 + ord_q(S(f)),$$

which means divisibility of $S(f)$ by $q$ translates directly into a Chevalley-Warning type divisibility of $N(f)$. It also allows us to generalize as follows.

Given any $f \in \mathbb{F}_q[x_1, \ldots, x_n]$, define

$$S(f) := \sum_{x_1 \ldots, x_n \in \mathbb{F}_q} \psi \circ Tr(f(x_1, \ldots, x_n)).$$

The Chevalley-Warning question is:

**Question 3.1.** *What is $ord_q\, S(f)$?*

This question was studied by Adolphson and Sperber [1] in 1987. To describe their result, we need the following notation. Writing the polynomial $f$ in terms of monomials $f(x_1, \ldots, x_n) = c_1 x_1^{r_1} \cdots x_n^{r_n} + \cdots$, we may define the Newton polytope $\Delta$ of $f$ as the convex closure of the points $(0, \ldots, 0), (r_1, \ldots, r_n), \ldots$. We will assume that $\Delta$ is $n$-dimensional. Associated to $\Delta$ is $w(\Delta)$, the smallest positive rational number such that the dilation $w(\Delta)\Delta$ intersects the strictly positive integers $\mathbb{Z}^n_{>0}$ nontrivially.

**Theorem 3.1** (Adolphson-Sperber [1])**.** *$ord_q(S(f)) \geq w(\Delta)$.*

For example, consider the Newton polytope of $f(x_1, x_2) = x_1^4 x_2 + x_1 x_2^3$. The line connecting the points (1,3) and (4,1) is $\frac{2}{11}x + \frac{3}{11}y = 11$. Since $(1, 1)$ belongs to $\Delta$, we may plug this into the linear form defining the line to obtain $w(\Delta) = \frac{2}{11} + \frac{3}{11} = \frac{5}{11}$. Hence, $ord_q\, S(f) \geq 5/11$.

# 4 $L$-functions

The Chevalley-Warning result, as well as the analogous question studied by Adolphson and Sperber, was concerned about solutions over $\mathbb{F}_q$. What happens if we ask the same question over an extension field $\mathbb{F}_{q^k}$? That is, given $f \in \mathbb{F}_q[x_1, \ldots, x_n]$, for each positive integer $k$ we may define

$$S_k(f) := \sum_{x_1 \ldots, x_n \in \mathbb{F}_{q^k}} \psi \circ Tr_{\mathbb{F}_{q^k}/\mathbb{F}_p}(f(x_1, \ldots, x_n)).$$

**Question 4.1.** *What is $ord_q\, S_k(f)$ as $k$ varies?*

Do you see that we have already given an answer to this? Observe that the lower bound $w(\Delta)$ has nothing to do with arithmetic, such as the coefficients of the polynomial $f$ or the field $\mathbb{F}_q$. It is a *geometric* lower bound, given in terms of geometric data, the degrees of the polynomial (and not all of the degrees either, only the ones that form the vertices of the Newton polytope $\Delta$). Thus, since the polynomial $f$ also belongs to $\mathbb{F}_{q^k}[x_1, \ldots, x_n]$, the theorem of Adolphson-Sperber tells us that $ord_{q^k}\, S_k(f) \geq w(\Delta)$. In other words,

$$ord_q\, S_k(f) \geq k w(\Delta). \tag{1}$$

By Lemma 3.1, this means the generating series $\sum_{k\geq 1} S_k(f)T^k$ converges $p$-adically on $ord_q(T) > -w(\Delta)$. We will come back to this in a moment.

To obtain more information about the sequence $\{S_k(f)\}_{k=1}^{\infty}$ we place it into a generating function that is different than the one above; its form suggests that the sequence may be defined by traces of a matrix (e.g. Grothendieck's trace formula; see blow). Define the $L$-function

$$L(f,T) := \exp\left(\sum_{k=1}^{\infty} S_k(f)\frac{T^k}{k}\right) = 1 + S_1(f)T + \cdots \tag{2}$$

The philosophy here is that any analytic information about the $L$-function, such as radius of convergence, zeros or poles, etc., may be translated into information about the sequence $S_k(f)$. Now, it is known that the $L$-function is a rational function with coefficients over $\mathbb{Z}[\zeta_p]$:

$$\begin{aligned} L(f,T) &= \frac{\prod(1-\alpha_i T)}{\prod(1-\beta_j T)} \\ &= \prod_{i=0}^{2n} det(1 - Frob_q T \mid H_c^i(\mathbb{A}^n_{\overline{\mathbb{F}}_q}, \mathcal{L}_{\psi(f)}))^{(-1)^{i+1}}, \end{aligned}$$

where the first equality is a theorem of Bombieri's, and the second is Grothendieck's trace formula. An immediate corollary of (1) is the following.

**Corollary 4.1.** *$ord_q(\alpha_i), ord_q(\beta_j) \geq w(f)$.*

*Proof.* Taking the logarithmic derivative $T\frac{d}{dT}\log$ of both sides of $L(f,T) = \prod(1-\alpha_i T)/\prod(1-\beta_j T)$ shows

$$\sum_{k\geq 1} S_k(f)T^k = \sum_j \sum_{k\geq 1} \beta_j^k T^k - \sum_i \sum_{k\geq 1} \alpha_i^k T^k.$$

As mentioned above, Adolphson-Sperber implies the left-hand side converges for $ord_q(T) > -w(f)$, and hence each series on the right has the same radius of convergence. Considering the geometric series of $\beta_j$ on the right-hand side, by Lemma 3.1, this means for every $\mu' < w(\Delta)$, $ord_q(\beta_j^k) \geq \mu' k$ for all $k$ sufficiently large. Hence, $ord_q(\beta_j) \geq \mu'$. The result follows. □

This also means that the reciprocal eigenvalues of the Frobenius $Frob_q$ acting on $H_c^i(\mathbb{A}^n_{\overline{\mathbb{F}}_q}, \mathcal{L}_{\psi(f)})$ also have $q$-adic order greater than or equal to $w(f)$, assuming no cancelation with other eigenvalues in the quotient.

Under certain regularity conditions on $f$, the associated $p$-adic cohomology becomes acyclic, which means the $L$-function is essentially a polynomial:

**Theorem 4.1** (Adolphson-Sperber [2])**.** *Suppose $f$ is $\Delta$-regular and $f$ contains a monomial of the form $ax_i^{r_i}$ for every $i$ (this is equivalent to the condition of $f$ being "commode"). Then,*

$$L(f,T)^{(-1)^{n+1}} = \sum_{i=0}^{nM} a_i T^i, \qquad a_i \in \mathbb{Z}[\zeta_p],$$

*where $M$ is described below.*

By (2), the Chevalley-Warning question asks about $ord_q(a_1)$, the coefficient of $T$ in the Taylor expansion about zero of $L(f,T)$. But what about the other coefficients?

**Question 4.2.** *What is $ord_q(a_i)$ for each $i$?*

This question is extremely difficult to answer in general. A more tractable problem is to estimate it by exhibiting a good lower bound, similar to the Ax-Katz bound. The following terminology originates from a conjecture of Katz, found in the same 1971 paper mentioned above (see also Mazur [8]). Associated to the polytope $\Delta$ is a Hodge polygon, defined as follows.

For each integer point $u \in \mathbb{Z}^n_{\geq 0}$, we may define a weight $w(u)$ as the smallest nonnegative rational number such that $u$ is in the dilation $w(u)\Delta$. Since the linear forms making up the boundary of the polytope $\Delta$ have rational coefficients, there exists a smallest positive integer $M$ such that $w(\mathbb{Z}^n_{\geq 0}) \subset \frac{1}{M}\mathbb{Z}_{\geq 0}$. Let $S := \{1, 2, \ldots, n\}$. For each subset $A \subset S$ and $i \in \mathbb{Z}_{\geq 0}$, define $W_A(i)$ as the number of point $u = (u_1, \ldots, u_n) \in \mathbb{Z}^n_{\geq 0}$ such that whose weight $w(u) = i/M$ and $u_i > 0$ for all $i \in A$. Define

$$h_\Delta(i) := \sum_{j=0}^{n} \left( \sum_{A\subset S, |A|=j} (-1)^{n-j} W_A(i-(n-j)M) \right).$$

The Hodge polygon $HP(\Delta)$ is defined as the lower convex hull of the points

$$\left(\sum_{i=0}^{k} h_\Delta(i), \frac{1}{M}\sum_{i=0}^{k} i h_\Delta(i)\right) \qquad k = 0, 1, \ldots, nM.$$

We define the $q$-adic Newton polygon $NP(f)$ of $L(f,T)$ as the lower convex hull of the points

$$(i, ord_q(a_i)) \qquad i = 0, 1, \ldots, nM.$$

**Theorem 4.2.** *(Adolphson-Sperber [2]) The q-adic Newton polygon $NP(f)$ lies on or above the Hodge polygon $HP(\Delta)$.*

**Question 4.3.** *When do we have equality between the Newton polygon and the Hodge polygon? (called* ordinary *when they conincide)*

Although there has been a lot of research into this question (see [11]), it is still very much an open question.

# 5 Unit root $L$-functions

In this last section we take a look at the Chavelley-Warning question for unit root $L$-functions. We will proceed via an example. Consider the one-parameter family of polynomials $f_t(x) := x^d + tx$ with $p \nmid d$. For $t \in \overline{\mathbb{F}}_q$, we define its degree by $deg(t) := [\mathbb{F}_q(t) : \mathbb{F}_q]$. It is well-known that for each $t$, the $L$-function is a polynomial of degree $d-1$:

$$L(f,t,T) = (1 - \alpha_1(t)T)\cdots(1 - \alpha_{d-1}T).$$

We now assume that $p \equiv 1 \bmod(d)$. In this case, the family is ordinary (NP = HP)

$$ord_{q^{deg(t)}}(\alpha_i(t)) = \frac{i}{d}.$$

We may now use this to form a new $L$-function, the so-called unit root $L$-function, which may be viewed as a sort of expectation value of the roots $\alpha_i(t)$ as $t$ varies. Define $\tilde{\alpha}_i(t) := \left(q^{-deg(t)}\right)^{i/d}\alpha_i(t)$ and note it is a $p$-adic unit. For each fixed $i$, we may define the unit root $L$-function

$$L_i(T) := \prod_{t\in\overline{\mathbb{F}}_q}\left(\frac{1}{1-\tilde{\alpha}_i T^{deg(t)}}\right)^{1/deg(t)}.$$

Wan's theorem [10] tells us that this is a $p$-adic meromorphic function.[2]

Writing $L_i(T) = 1 + S_1(i)T + \cdots$, where $S_i = \sum_{t\in\mathbb{F}_q}\tilde{\alpha}_i(t)$, a Chevalley-Warning type question asks for $ord_q\, S_1(i)$. This has been answered by some recent joint work with Steven Sperber:

**Theorem 5.1.** *(H. - Sperber) For each $i$, $ord_q\, S_1(i) \geq (d-1)/d$.*

Open questions abound:

**Question 5.1.** *When is this equality? Is there a Hodge theoretic interpretation of this lower bound? And what about the higher term coefficients? And other families?*

# References


[1] Alan Adolphson and Steven Sperber, *p-adic estimates for exponential sums and the theorem of Chevalley-Warning*, Ann. Sci. École Norm. Sup. (4) **20** (1987), no. 4, 545–556.

[2] Alan Adolphson and Steven Sperber, *Exponential Sums and Newton Polyhedra: Cohomolgy and Estimates*, Annals of Math. **130** (1989), no. 2, 367–406.

[3] James Ax, *Zeroes of polynomials over finite fields*, Amer. J. Math. **86** (1964), 255–261. MR 0160775 (28 #3986)


[2] When $d = 3$, a Tate-Deligne lifting of the Frobenius exists for this family which provides another proof of meromorphy using the Monsky-Reich trace formula. For general $d$ with $p \nmid d$ and $p \equiv 1 \bmod(d)$, it is likely such a lifting also exists; however, for general families, such liftings will rarely exist.

[4] L. E. Dickson, *On the representation of numbers by modular forms*, Bull. Amer. Math. Soc. **15** (1909), no. 7, 338–347.

[5] Hélène Esnault, *Varieties over a finite field with trivial Chow group of 0-cycles have a rational point*, Invent. Math. **151** (2003), no. 1, 187–191. MR 1943746 (2004e:14015)

[6] D.R. Heath-Brown, *A note on the chevalley–warning theorems*, arXiv:1009.3764v1 (2010).

[7] Nicholas M. Katz, *On a theorem of Ax*, Amer. J. Math. **93** (1971), 485–499. MR 0288099 (44 #5297)

[8] B. Mazur, *Frobenius and the Hodge filtration*, Bull. Amer. Math. Soc. **78** (1972), 653–667. MR 0330169 (48 #8507)

[9] O. Moreno and C. J. Moreno, *An elementary proof of a partial improvement to the Ax-Katz theorem*, Applied algebra, algebraic algorithms and error-correcting codes (San Juan, PR, 1993), Lecture Notes in Comput. Sci., vol. 673, Springer, Berlin, 1993, pp. 257–268. MR 1251983 (94k:11039)

[10] D. Wan, *Dwork's conjecture on unit root zeta functions*, Ann. Math. **150** (1999), 867–927.

[11] Daqing Wan, *Variation of p-adic Newton polygons for L-functions of exponential sums*, Asian J. Math. **8** (2004), no. 3, 427–471. MR 2129244 (2006b:11095)