

# THE CHEVALLEY-WARNING THEOREM (FEATURING... THE ERDÖS-GINZBURG-ZIV THEOREM)

PETE L. CLARK

## 1. THE CHEVALLEY-WARNING THEOREM

In this handout we shall discuss a result that was conjectured by Emil Artin in 1935 and proved shortly thereafter by Claude Chevalley. A refinement was given by Artin's graduate student Ewald Warning, who, as the story goes, was the one whom Artin had intended to prove the theorem before Chevalley came visiting Göttingen and got Artin to say a little too much about the mathematics his student was working on.

One of the charms of the Chevalley-Warning theorem is that it can be stated and appreciated without much motivational preamble. So let's just jump right in.

### 1.1. Statement of the theorem(s).

Let  $q = p^a$  be a prime power, and let  $\mathbb{F}_q$  be a finite field of order  $q$ . We saw earlier in the course that there exists a finite field of each prime power cardinality.<sup>1</sup> For the reader who is unfamiliar with finite fields, it may be a good idea to just replace  $\mathbb{F}_q$  with  $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$  on a first reading, and then afterwards look back and see that the assumption of an arbitrary finite field changes nothing.

**Theorem 1.** (*Chevalley's Theorem*) *Let  $n, d_1, \dots, r$  be positive integers such that  $d_1 + \dots + d_r < n$ . For each  $1 \leq i \leq r$ , let  $P_i(t_1, \dots, t_n) \in \mathbb{F}_q[t_1, \dots, t_n]$  be a polynomial of total degree  $d_i$  with zero constant term:  $P_i(0, \dots, 0) = 0$ . Then there exists  $0 \neq x = (x_1, \dots, x_n) \in \mathbb{F}_q^n$  such that*

$$P_1(x) = \dots = P_r(x) = 0.$$

Exercise 1: Suppose we are given any system of polynomials  $P_1(t), \dots, P_r(t)$  in  $n$  variables  $t_1, \dots, t_n$  with  $\sum_i \deg(P_i) < n$ . Deduce from Chevalley's that if there exists at least one  $x \in \mathbb{F}_q^n$  such that  $P_1(x) = \dots = P_r(x)$ , then there exists  $y \neq x$  such that  $P_1(y) = \dots = P_r(y)$ . (Hint: Make a change of variables to reduce to Chevalley's theorem.)

In other words, Exercise 1 asserts that a system of polynomials in  $n$  variables over  $\mathbb{F}_q$  cannot have exactly one common solution, provided the sum of the degrees is less than  $n$ . Warning's theorem gives a generalization:

---

<sup>1</sup>It can be shown that any two finite fields of the same order are isomorphic; indeed this is (literally) a textbook application of the uniqueness of splitting fields of polynomials and can be found in any graduate level algebra text treating field theory. But we don't need this uniqueness statement here.

**Theorem 2.** (*Warning's Theorem*) Let  $n, d_1, \dots, r$  be positive integers such that  $d_1 + \dots + d_r < n$ . For each  $1 \leq i \leq r$ , let  $P_i(t_1, \dots, t_n) \in \mathbb{F}_q[t_1, \dots, t_n]$  be a polynomial of total degree  $d_i$ . Let

$$Z = \#\{(x_1, \dots, x_n) \in \mathbb{F}_q^n \mid P_1(x_1, \dots, x_n) = \dots = P_r(x_1, \dots, x_n) = 0.\}$$

Then  $Z \equiv 0 \pmod{p}$ .

Arin had conjectured the following special case:

**Corollary 3.** Let  $P(t_1, \dots, t_n) \in \mathbb{F}[t_1, \dots, t_n]$  be a homogeneous polynomial of degree  $d$  in  $n$  variables over a finite field  $\mathbb{F}$ . If  $n > d$  then there exists  $(0, \dots, 0) \neq (x_1, \dots, x_n) \in \mathbb{F}^n$  such that  $P(x_1, \dots, x_n) = 0$ .

In the sequel, we will refer to any of Theorem 1, Theorem 2 or Corollary 3 as the **Chevalley-Warning theorem**.

## 1.2. Applications to Quadratic Forms.

Taking  $d = 1$  Corollary 3 asserts that any homogeneous linear equation  $at_1 + bt_2 = 0$  (with  $a$  and  $b$  not both 0) over a finite field has a nonzero solution. Of course linear algebra tells us that the solution set to such an equation is a one-dimensional vector space, and this holds over *any* field, infinite or otherwise. So this is a trivial case.

Already the case  $d = 2$  is much more interesting. A homogeneous polynomial of degree 2 is called a **quadratic form**. For simplicity, we shall for the most part consider here only nondegenerate diagonal forms over a field  $F$ ,<sup>2</sup> i.e.,

$$q(x) = q(x_1, \dots, x_n) = a_1x_1^2 + \dots + a_nx_n^2, \quad x_1, \dots, x_n \in F, \quad x_1 \cdots x_n \neq 0.$$

For some fields  $F$ , no matter how large we take  $n$  to be, we can still choose the coefficients so that  $q(x) = 0$  has no nontrivial solution. For instance, consider the sum of squares form

$$q_n(x) = x_1^2 + \dots + x_n^2.$$

This has no nontrivial solution over the real numbers, or over any subfield of  $\mathbb{R}$ .

**Proposition 4.** Let  $F$  be any field, and consider the form  $q_a(x) = x_1^2 + ax_2^2$ .

- The form  $q_{2,a}$  has a nontrivial solution iff there exists  $\alpha \in F$  such that  $\alpha^2 = -a$ .
- Therefore if  $F = \mathbb{F}_q$ ,  $q = p^a$ , then  $q_{2,1}(x) = x_1^2 + x_2^2$  has a nontrivial solution iff  $p = 2$ ,  $p \equiv 1 \pmod{4}$  or  $a$  is even.

Exercise 2: Prove Proposition 4.

Exercise 3: a) Suppose  $F$  is a field (of characteristic different from 2) which admits a quadratic field extension  $K = F(\sqrt{\alpha})$ . Deduce that there exists a binary quadratic form  $q_{2,a}(x)$  over  $F$  which has no nontrivial solution.

b) For any odd  $q = p^a$ , show that there exists a binary quadratic form  $q$  over  $\mathbb{F}_q$  with only the trivial solution. Can you write one down explicitly?

c)\* Show that the conclusion of part b) still holds when  $q$  is even, although in this case one has to take a nondiagonal form  $q(x, y) = ax^2 + bxy + cy^2 = 0$ .

According to Corollary 3, any quadratic form in at least three variables over a

<sup>2</sup>When the characteristic of  $F$  is not 2, one can diagonalize every quadratic form by making a linear change of variables, so no generality is lost by restricting to the diagonal case.

finite field has a nontrivial solution. This is quite different from the situation for  $F = \mathbb{R}$  or  $F = \mathbb{Q}$ . And it has many useful consequences, e.g.:

**Proposition 5.** *Let  $F$  be a field of characteristic different from 2 in which each quadratic form in three variables has a nontrivial solution. Then, for any  $a, b, c \in F^\times$ , there exist  $x, y \in F$  such that*

$$ax^2 + by^2 = c.$$

*Proof.* In other words, we are claiming that the **inhomogeneous** equation  $ax^2 + by^2 = c$  has a solution over  $F$ . To see this, we **homogenize**: introduce a third variable  $z$  and consider the equation:  $ax^2 + by^2 - cz^2 = 0$ . By Corollary 3 there are  $x_0, y_0, z_0 \in K$ , not all zero, such that  $ax_0^2 + by_0^2 = cz_0^2$ . If  $z_0 \neq 0$ , then we can divide through, getting

$$a \left( \frac{x_0}{z_0} \right)^2 + b \left( \frac{y_0}{z_0} \right)^2 = c.$$

If  $z_0 = 0$ , this doesn't work; rather, we get a nontrivial solution  $(x_0, y_0)$  to  $ax_0^2 + by_0^2 = 0$ . Dividing by  $a$  we get  $x_0^2 + (\frac{b}{a})y_0^2 = 0$ . But as above this can only happen if  $-\frac{b}{a} = t^2$  is a square in  $K$ , and then we can factor  $q(x) = x^2 - t^2y^2 = (x+ty)(x-ty)$ . This gives us a lot more leeway in solving the equation. For instance, we could factor  $c$  as  $c \cdot 1$  and give ourselves the linear system

$$x + ty = c$$

$$x - ty = 1$$

which has a solution  $(x, y) = (\frac{c+1}{2}, \frac{c-1}{2})$ . Note that it is here that we use the hypothesis that the characteristic of  $K$  is not 2.  $\square$

In particular this gives an alternate (much more sophisticated!) proof of [Minkowski's Theorem Handout, Lemma 15].

## 2. TWO PROOFS OF WARNING'S THEOREM

### 2.1. Polynomials and polynomial functions.

We begin with a discussion about polynomials as ring elements versus polynomials as functions which is of interest in its own right. (In fact, it is because of the interest of these auxiliary results that we have chosen to include this proof.)

Let  $R$  be an integral domain and  $R[t_1, \dots, t_n]$  be the polynomial ring in  $n$  indeterminates over  $R$ . An element  $P(t) = P(t_1, \dots, t_n) \in R[t_1, \dots, t_n]$  is a purely formal object: it is a finite  $R$ -linear combination of monomial terms, which are added and multiplied according to simple formal rules.

Note that this is not the perspective on polynomials one encounters in calculus and analysis. For instance a univariate polynomial  $P(t) = a_n t^n + \dots + a_1 t + a_0 \in \mathbb{R}[t]$  is regarded as a **function** from  $\mathbb{R}$  to  $\mathbb{R}$ , given of course by  $x \mapsto P(x)$ . Similarly for multivariable polynomials:  $P(t_1, \dots, t_n) \in \mathbb{R}[t_1, \dots, t_n]$  may be defined by the same formal  $\mathbb{R}$ -linear combination of monomial terms as above but that is just notation: what matters is the function  $\mathbb{R}^n \rightarrow \mathbb{R}$  given by  $(x_1, \dots, x_n) \mapsto P(x_1, \dots, x_n)$ . In other words, a polynomial in  $n$  variables can be **evaluated** at any point of  $\mathbb{R}^n$ .

Can these two perspectives be reconciled? A moment's thought makes it clear that the "evaluation" of a polynomial is a perfectly algebraic operation: in other words, given any domain  $R$  and element  $P(t)$  of the polynomial ring  $R[t_1, \dots, t_n]$ , we can evaluate  $P$  at any point  $(x_1, \dots, x_n) \in R^n$ , getting an element  $P(x_1, \dots, x_n)$ . To be formal about it, we have an **evaluation map**:

$$\Phi : R[t_1, \dots, t_n] \mapsto \text{Map}(R^n, R),$$

where by  $\text{Map}(R^n, R)$  we just mean the set of all functions  $f : R^n \rightarrow R$ . In fact this map  $\Phi$  has some nice algebraic structure. The set  $\text{Map}(R^n, R)$  of all functions from  $R^n$  to  $R$  can be made into a commutative ring in which addition and multiplication are just defined "pointwise":

$$\begin{aligned} (f + g)(x_1, \dots, x_n) &= f(x_1, \dots, x_n) + g(x_1, \dots, x_n), \\ (fg)(x_1, \dots, x_n) &= f(x_1, \dots, x_n) \cdot g(x_1, \dots, x_n). \end{aligned}$$

It is straightforward to see that the evaluation map  $\Phi$  is then a homomorphism of rings. Let us put

$$\mathcal{P}_n := \Phi(R[t_1, \dots, t_n]) \subset \text{Map}(R^n, R),$$

so that  $\mathcal{P}_n$  is the ring of polynomial functions in  $n$  variables on  $R$ .

We are interested in the following question: if  $P(t), Q(t) \in R[t_1, \dots, t_n]$  are such that for all  $(x_1, \dots, x_n) \in R^n$  we have  $P(x_1, \dots, x_n) = Q(x_1, \dots, x_n)$  – so that  $P$  and  $Q$  give the same function from  $R^n$  to  $R$  – must  $P(t) = Q(t)$  as elements of  $R[t_1, \dots, t_n]$ ? In other words, is  $\Phi$  injective?

I hope you know that in the familiar case of  $n = 1$ ,  $R = \mathbb{R}$  the answer is "yes": two real univariate polynomials which give the same function are term-by-term equal. The proof is as follows: define  $R(t) := P(t) - Q(t)$ . We are given that  $R(x) = 0$  for all  $x \in R$ . But if  $R(x)$  were not the zero polynomial, it would have some degree  $d \geq 0$  and basic (high school!) algebra shows that a polynomial over a field of degree  $d$  cannot have more than  $d$  roots. But  $R(x)$  has infinitely many roots, so it must be the identically zero polynomial.

Evidently this argument works for univariate polynomials over any infinite field. The following is a stronger result:

**Proposition 6.** *Let  $R$  be an infinite integral domain and  $n \in \mathbb{Z}^+$ . Then the evaluation map*

$$\Phi : R[t_1, \dots, t_n] \rightarrow \mathcal{P}_n \subset \text{Map}(R^n, R)$$

*is a homomorphism of rings.*

*a) Moreover  $\Phi$  is injective.*

*b) However,  $\Phi$  is not surjective: not every function  $f : R^n \rightarrow R$  is given by a polynomial.*

*Proof.* a) Again it suffices to show that if  $\Phi(P(t)) = 0$ , then  $P(t)$  is the zero polynomial. If  $n = 1$ , we just showed this when  $R$  was a field. But that argument easily carries over, since every integral domain  $R$  can be embedded into a field  $F$  (namely its field of fractions). If there existed a nonzero polynomial  $P(t) \in R[t]$  such that there were infinitely  $x \in R$  such that  $P(x) = 0$ , then since  $R \subset F$ , there are also infinitely many  $x \in F$  such that  $P(x) = 0$ , contradiction. Assume now that  $n > 1$ . In general the theory of polynomials of several variables can be significantly

more complicated than that of univariate polynomials, but here we can use a dirty trick:

$$R[t_1, \dots, t_{n-1}, t_n] = (R[t_1, \dots, t_{n-1}])[t_n].$$

In other words, a polynomial  $P(t_1, \dots, t_n)$  in  $n$  variables over the integral domain  $R$  may be viewed as a polynomial  $Q(t_n) := (P(t_1, \dots, t_{n-1}))(t_n)$  in one variable over the integral domain  $R_{n-1} := R[t_1, \dots, t_{n-1}]$ . If  $P(x_1, \dots, x_n) = 0$  for all  $(x_1, \dots, x_n) \in R^n$  then, since  $R \subset R_{n-1}$ , the univariate polynomial  $Q(t_n)$  has infinitely many roots in  $R_{n-1}$  and thus is identically zero by the above argument.

As for part b), for instance the function  $\mathbf{1}_0 : R^n \rightarrow R$  which maps  $0 \in R^n$  to 1 and every other element of  $R^n$  to 0 is not a polynomial function. You are asked to show this in Exercise 4 below. Another argument is by counting: for infinite  $R$ , the cardinality of  $R[t_1, \dots, t_n]$  is equal to the cardinality of  $R$ , whereas the total number of functions from  $R^n$  to  $R$  has cardinality  $|R|^{|R^n|} = 2^{|R|} > |R|$ , so “most” functions are not polynomials.  $\square$

Exercise 4: Let  $R$  be an infinite integral domain and  $n \in \mathbb{Z}^+$ . Show that the characteristic function  $\mathbf{1}_0$  of the origin – i.e., the function which maps  $0 = (0, \dots, 0)$  to 1 and every other element of  $R^n$  to zero – is not a polynomial function. (Hint: restrict the function  $\mathbf{1}_0$  to a line passing through the origin, and thereby reduce to the case  $n = 1$ .)

We shall not need Proposition 6 in our work on the Chevalley-Warning theorem, but it is interesting to contrast the infinite case with the finite case. First of all:

**Lemma 7.** *Let  $R = \mathbb{F}_q$  be a finite integral domain (necessarily a field). Then every function  $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$  is given by a polynomial.*

*Proof.* We first express the function  $\mathbf{1}_0$ , which takes 0 to 1 and every other element to 0, as a polynomial. Indeed, since  $x^{q-1} = 1$  for  $x \in \mathbb{F}_q^\times$  and  $0^{q-1} = 0$ , we have for all  $x = (x_1, \dots, x_n) \in \mathbb{F}_q^n$  that

$$\mathbf{1}_0(\mathbf{x}) = \prod_{i=1}^n (1 - x_i^{q-1}).$$

For an arbitrary function  $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$ , define

$$(1) \quad P_f(t) := \sum_{y \in \mathbb{F}_q^n} f(y) \prod_{i=1}^n (1 - (t_i - y_i)^{q-1}).$$

Then every term in the sum of  $P_f(x)$  with  $y \neq x$  yields zero whereas the term  $y = x$  yields  $f(x)$ .  $\square$

On the other hand, over a finite field  $\mathbb{F}_q$ , a nonzero polynomial may evaluate to the zero function: indeed  $t^q - t$  is a basic one variable example. There is no contradiction here because a nonzero polynomial over a domain cannot have more roots than its degree, but  $t^q - t = \prod_{a \in \mathbb{F}_q} (t - a)$  has exactly as many roots as its degree. Moreover, no nonzero polynomial of degree less than  $q$  could lie in the kernel of the evaluation map, so  $t^q - t$  is a minimal degree nonzero element of  $\text{Ker}(\Phi)$ . But, since  $\mathbb{F}_q[t]$  is a PID, every nonzero ideal  $I$  is generated by its unique monic element of least degree, so  $\text{Ker}(\Phi) = \langle t^q - t \rangle$ .

We would like to compute  $\text{Ker}(\Phi)$  in the multivariable case. Reasoning as above it is clear that for all  $1 \leq i \leq n$  the polynomials  $t_i^q - t_i$  must lie in the kernel of the evaluation map, so at least we have  $J = \langle t_1^q - t_1, \dots, t_n^q - t_n \rangle \subset \text{Ker}(\Phi)$ . We will see that in fact  $J = \text{Ker}(\Phi)$ . We can even do better: for each polynomial  $P(t)$  we can find a canonical element  $\tilde{P}$  of the coset  $P(t) + \text{Ker}(\Phi)$ .

The key idea is that of a reduced polynomial. We say that a monomial  $ct_1^{a_1} \cdots t_n^{a_n}$  is **reduced** if  $a_i < q$  for all  $i$ . A polynomial  $P \in \mathbb{F}_q[t]$  is **reduced** if each of its nonzero monomial terms is reduced. Equivalently, a reduced polynomial is one for which the total degree in each variable is less than  $q$ .

Example: The polynomial  $P_f(t)$  above is a sum of polynomials each having degree  $q - 1$  in each variable, so is reduced.

Exercise 5: The reduced polynomials form an  $\mathbb{F}_q$ -subspace of  $\mathbb{F}_q[t_1, \dots, t_n]$ , with a basis being given by the reduced monomials.

The idea behind the definition is that if in a monomial term we had an exponent  $t_i^{a_i}$  with  $a_i \geq q$ , then from the perspective of the associated function this is just wasteful: we have

$$x_i^{a_i} = x_i^{q+(a_i-q)} = x_i^q x_i^{a_i-q} = x_i x_i^{a_i-q} = x_i^{a_i-(q-1)}.$$

Thus by a sequence of “elementary reductions” of this type we can convert any polynomial  $P$  into a reduced polynomial  $\tilde{P}$ . Moreover, a little reflection makes clear that  $P - \tilde{P} \in J$ .

Is it possible for a given polynomial  $P$  to be congruent modulo  $\text{Ker}(\Phi)$  to more than one reduced polynomial? Well, the reduced polynomials form an  $\mathbb{F}_q$ -vector subspace of the space of all polynomials with basis given by the reduced monomials, of which there are  $q^n$ , so the total number of reduced polynomials is  $q^{q^n}$ . In fact this is also the total number of functions from  $\mathbb{F}_q^n$  to  $\mathbb{F}_q$ . Since we know that every function is given by some reduced polynomial, it must be that evaluation map restricted to reduced polynomials is a bijection. Finally, since we showed that every polynomial was equivalent modulo  $J$  to a reduced polynomial, so that  $\#\mathbb{F}_q[t]/J \leq q^{q^n}$ . By surjectivity of  $\Phi$  we know  $\#\mathbb{F}_q[t]/\text{Ker}(\Phi) = \#\text{Map}(\mathbb{F}_q^n, \mathbb{F}_q) = q^{q^n}$ . Therefore the quotient map  $\mathbb{F}_q[t]/J \rightarrow \mathbb{F}_q[t]/\text{Ker}(\Phi)$  is a bijection and hence  $J = \text{Ker}(\Phi)$ .

Remark: More standard is to prove that a nonzero reduced polynomial does not induce the zero function by induction on the number of variables. Then the surjectivity of  $\Phi$  can be deduced from the injectivity on reduced polynomials by noticing, as we did, that the domain and codomain are finite sets with the same cardinality. Our treatment here is undeniably more complicated than this, but also seems more interesting. It will also smooth the way for our first proof of Warning’s theorem.

Let us summarize all the preceding results:<sup>3</sup>

---

<sup>3</sup>Although all parts of this theorem must be well-known, I have not seen the full statement in the literature.

**Theorem 8.** (*Polynomial evaluation theorem*) Let  $R$  be an integral domain and  $n \in \mathbb{Z}^+$ . Let  $\Phi : R[t] = R[t_1, \dots, t_n] \rightarrow \text{Map}(R^n, R)$  be the homomorphism of rings obtained by associating to each polynomial the corresponding polynomial function  $x = (x_1, \dots, x_n) \mapsto P(x)$ .

a) If  $R$  is infinite, then  $\Phi$  is injective but not surjective: every function  $f : R^n \rightarrow R$  is represented by at most one polynomial, and there exist functions not represented by any polynomial.

b) If  $R$  is finite, then  $\Phi$  is surjective but not injective: its kernel is the ideal  $\langle t_1^q - t_1, \dots, t_n^q - t_n \rangle$ . Thus every function  $f : R^n \rightarrow R$  is represented by infinitely many polynomials. Moreover, for each  $f$  there exists a unique **reduced** polynomial representative, given explicitly as the polynomial  $P_f(t)$  of (1) above.

If  $f : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$  is any function, we define its **reduced degree** in  $t_i$  to be the degree in  $t_i$  of the associated reduced polynomial, and similarly its **reduced total degree** to be the total degree of the associated reduced polynomial.

Exercise 6: Show that if  $P$  is any polynomial, the total degree  $\deg(\tilde{P})$  of  $\tilde{P}$  is less than or equal to the total degree  $\deg(P)$  of  $P$ .

### 2.2. First proof of Warning's Theorem.

We have polynomials  $P_1(t), \dots, P_r(t)$  in  $n$  variables with  $\sum_{i=1}^r \deg(P_i) < n$ . Put

$$(2) \quad Z = \{(x_1, \dots, x_n) \in \mathbb{F}_q^n \mid P_1(x) = \dots = P_r(x) = 0.\}$$

We want to show that  $\#Z \equiv 0 \pmod{p}$ . Let  $\mathbf{1}_Z : \mathbb{F}_q^n \rightarrow \mathbb{F}_q$  be the ( $\mathbb{F}_q$ -valued) "characteristic function" of the subset  $Z$ , i.e., the function which maps  $x$  to 1 if  $x \in Z$  and  $x$  to 0 otherwise. Now one polynomial representative  $\mathbf{1}_Z$  is

$$(3) \quad P(t) := \prod_{i=1}^r (1 - P_i(t)^{q-1});$$

whereas – essentially by (1) above – the reduced polynomial representative is

$$Q_Z(t) = \sum_{x \in Z} \prod_{i=1}^n (1 - (t_i - x_i)^{q-1}).$$

Now comes the devilry: the total degree of  $P(t)$  is  $(q-1) \sum_i d_i < (q-1)n$ .

On the other hand, consider the coefficient of the monomial  $t_1^{q-1} \dots t_n^{q-1}$  in  $Q_Z(t)$ : it is  $(-1)^n \#Z$ . If we **assume** that  $\#Z$  is not divisible by  $p$ , then this term is nonzero and  $Q_Z(t)$  has total degree at least  $(q-1)n$ . By Exercise X.X, we have

$$\deg(\tilde{P}) \leq \deg(P) < (q-1)n \leq \deg(Q_Z).$$

Therefore  $\tilde{P} \neq Q_Z$ , whereas we ought to have  $\tilde{P} = Q_Z$ , since each is the reduced polynomial representative of  $\mathbf{1}_Z$ . Evidently we assumed something we shouldn't have: rather, we must have  $p \mid \#Z$ , qed.

### 2.3. Ax's proof of Warning's theorem.

We maintain the notation of the previous section, especially the polynomial  $P(t)$  of (3) and the subset  $Z$  of (2). Because  $P(x) = \mathbf{1}_Z(x)$  for all  $x \in \mathbb{F}_q^n$ , we have

$$\#Z \equiv \sum_{x \in \mathbb{F}_q^n} P(x) \pmod{p}.$$

So we just need to evaluate the sum. Since every polynomial is an  $\mathbb{F}_q$ -linear combination of monomial terms, it is reasonable to start by looking for a formula for  $\sum_{x \in \mathbb{F}_q^n} x_1^{a_1} \cdots x_n^{a_n}$  for non-negative integers  $a_1, \dots, a_n$ . It is often the case that if  $f : G \rightarrow \mathbb{C}$  is a “nice” function from an abelian group to the complex numbers, then the complete sum  $\sum_{x \in G} f(x)$  has a simple expression. Case in point:

**Lemma 9.** *Let  $a_1, \dots, a_n$  be non-negative integers.*

- a) *If for  $1 \leq i \leq n$ ,  $a_i$  is a positive multiple of  $q-1$ , then  $\sum_{x \in \mathbb{F}_q^n} x_1^{a_1} \cdots x_n^{a_n} = (-1)^n$ .*  
b) *In every other case – i.e., for at least one  $i$ ,  $1 \leq i \leq n$ ,  $a_i$  is not a positive integer multiple of  $q-1$  – we have  $\sum_{x \in \mathbb{F}_q^n} x_1^{a_1} \cdots x_n^{a_n} = 0$ .*

*Proof.* Part a) is not needed in the sequel and is just stated for completeness; we leave the proof as an exercise.

As for part b), we have

$$\sum_{x \in \mathbb{F}_q^n} x_1^{a_1} \cdots x_n^{a_n} = \prod_{i=1}^n \left( \sum_{x_i \in \mathbb{F}_q} x_i^{a_i} \right).$$

By assumption there exists at least one  $i$ ,  $1 \leq i \leq n$ , such that  $a_i$  is either 0 or is positive but not a multiple of  $q-1$ . If  $a_i = 0$ , then  $\sum_{x_i \in \mathbb{F}_q} x_i^{a_i} = \sum_{x_i \in \mathbb{F}_q} 1 = q \equiv 0 \in \mathbb{F}_q$ , so assume that  $a_i$  is positive but not divisible by  $q-1$ . Let  $\alpha$  be a generator for the cyclic group  $\mathbb{F}_q^\times$ , and put  $\beta = \alpha^{a_i}$ . Then

$$\sum_{x_i \in \mathbb{F}_q} x_i^{a_i} = 0^{a_i} + \sum_{x_i \in \mathbb{F}_q^\times} x_i^{a_i} = 0 + \sum_{N=0}^{q-2} (\alpha^N)^{a_i} = \sum_{N=0}^{q-2} \beta^N = \frac{1 - \beta^{q-1}}{1 - \beta} = \frac{1 - 1}{1 - \beta} = 0.$$

□

Finally, the polynomial  $P(t)$  has degree  $\sum_{i=1}^r d_i(q-1) = (q-1) \sum_{i=1}^r d_i < (q-1)n$ . Thus in each monomial term  $ct_1^{a_1} \cdots t_n^{a_n}$  in  $P(t)$  must have  $a_1 + \dots + a_r < (q-1)n$ , so it can't be the case that each  $a_i \geq q-1$ . Therefore Lemma 9 applies, and  $\sum_{x \in \mathbb{F}_q^n}$  is an  $\mathbb{F}_q$ -linear combination of sums each of which evaluates to 0 in  $\mathbb{F}_q$  and therefore the entire sum is 0.

### 3. SOME LATER WORK

Under the hypotheses of Warning's theorem we can certainly have 0 solutions. For instance, we could take  $P_1(t)$  to be any polynomial with  $\deg(P_1) < \frac{n}{2}$  and  $P_2(t) = P_1(t) + 1$ . Or, when  $q$  is odd, let  $a \in \mathbb{F}_q$  be a quadratic nonresidue, let  $P_1(t)$  be a polynomial of degree less than  $\frac{n}{2}$  and put  $P(t) = P_1(t)^2 - a$ .

On the other hand, it is natural to wonder: in Warning's theorem, we might actually have  $\#Z \equiv 0 \pmod{q}$ ? The answer is now known, but it took 46 years.

Let us first consider the case of  $r = 1$ , i.e., a single polynomial  $P$  of degree less than  $n$ . In his original 1926 paper, E. Warning proved that  $\#Z$ , if positive, is at

least  $q^{n-d}$ . And in the same 1964 paper containing the quick proof of Warning's theorem, J. Ax showed that  $q^b \mid \#Z$  for all  $b < \frac{n}{d}$ . By hypothesis we can take  $b = 1$ , so the aforementioned question has an affirmative answer in this case.

For the case of multiple polynomials  $P_1, \dots, P_r$  of degrees  $d_1, \dots, d_r$ , in a celebrated 1971 paper N. Katz showed that  $q^b \mid \#Z$  for all positive integers  $b$  satisfying

$$b < \frac{n - (d_1 + \dots + d_r)}{d_1} + 1.$$

Since the above fraction is by hypothesis strictly positive, we can take  $b = 1$  getting indeed  $\#Z \equiv 0 \pmod{q}$  in all cases.

These divisibilities are called estimates of **Ax-Katz type**. It is known that there are examples in which the Ax-Katz divisibilities are best possible, but refining these estimates in various cases is a topic of active research: for instance there is a 2007 paper by W. Cao and Q. Sun, *Improvements upon the Chevalley-Warning-Ax-Katz-type estimates*, J. Number Theory 122 (2007), no. 1, 135–141.

Notice that the work since Warning has focused on the problem of getting best possible  $p$ -adic estimates for the number of solutions: that is, instead of bounds of the form  $\#Z \geq N$ , we look for bounds of the form  $\text{ord}_p(\#Z) \geq N$ . Such estimates are closely linked to the  **$p$ -adic cohomology** of algebraic varieties, a beautiful (if technically difficult) field founded by Pierre Deligne in his landmark paper "Weil II."

The hypotheses of the Chevalley-Warning theorem are also immediately suggestive to algebraic geometers: (quite) roughly speaking there is a geometric division of algebraic varieties into three classes: Fano, Calabi-Yau, and general type. The degree conditions in Warning's theorem are precisely those which give, among the class of algebraic varieties represented nicely by  $r$  equations in  $n$  variables ("smooth complete intersections"), the Fano varieties. A recent result of H el ene Esnault gives the geometrically natural generalization: any Fano variety over  $\mathbb{F}_q$  has a rational point. There are similar results for other Fano-like varieties.

#### 4. THE ERDÖS-GINZBURG-ZIV THEOREM

##### 4.1. A Mathematical Card Game.

Consider the following game. One starts with a deck of one hundred cards (or  $N$  cards, for some arbitrary positive integer  $N$ ). Any number of players may play; one of them is the dealer. The dealer shuffles the deck, and the player to the dealer's left selects a card ("any card") from the deck and shows it to everyone. The player to the dealer's right writes down the numerical value of the card, say  $n$ , and keeps this in a place where everyone can see it. The card numbered  $n$  is reinserted into the deck, which is reshuffled. The dealer then deals cards face up on the table, one at a time, at one minute intervals, or sooner by unanimous consent (i.e., if everyone wants the next card, including the dealer, then it is dealt; otherwise the dealer waits for a full minute). A player wins this round of the game by correctly selecting any  $k > 0$  of the cards on the table such that the sum of their numerical values is divisible by  $n$ . When all the cards are dealt, the players have as much time as they wish.

For example, suppose that  $n = 5$  and the first card dealt is 16. 16 is not divisible by 5, so the players all immediately ask for another card: suppose it is 92. 92 is not divisible by 5 and neither is  $92 + 16 = 118$ , so if the players are good, they will swiftly ask for the next card. (Of course, if they are any good at all, they will not be thinking of the numbers as 16 and 92 but rather as  $1 \pmod{5}$  and  $2 \pmod{5}$ , but let's present things literally so as to understand the game mechanics.) Suppose the next card is 64. Then someone can win by collecting the 64 and the 16 and calling attention to the fact that  $64 + 16 = 80$  is divisible by 5.

Here's the question: is it always possible to win the game, or can all the cards be dealt with no solution?

We claim that it is never necessary to deal more than  $n$  cards before a solution exists. Moreover, so long as the total number  $N$  of cards in the deck is sufficiently large compared to the selected modulus  $n$ , it is possible for fewer than  $n$  cards to be insufficient.

To see the latter, note that if  $n = 1$  we obviously need  $n$  cards, and if  $n = 2$  we will need  $n$  cards iff the first card dealt is odd. If  $n = 3$  we may need  $n$  cards iff  $N \geq 4$ , since if 1 and 4 are the first two cards dealt there is no solution. In general, if the cards dealt are  $1, 1+n, 1+2n, \dots, 1+(n-2)n$ , then these are  $n-1$  cards which are all  $1 \pmod{n}$  and clearly we cannot obtain  $0 \pmod{n}$  by adding up the values of any  $0 < k \leq n-1$  of these. This is possible provided  $N \geq n^2 - 2n + 1 = (n-1)^2$ .<sup>4</sup>

But why are  $n$  cards always sufficient? We can give an explicit algorithm for finding a solution: for each  $1 \leq k \leq n$ , let  $S_k = a_1 + \dots + a_k$  be the sum of the values of the first  $k$  cards. If for some  $k$ ,  $S_k$  is divisible by  $n$ , we are done: we can at some point select all the cards. Otherwise, we have a sequence  $S_1, \dots, S_n$  of elements in  $\mathbb{Z}/n\mathbb{Z}$ , none of which are  $0 \pmod{n}$ . By the pigeonhole principle, there must exist  $k_1 < k_2$  such that  $S_{k_1} \equiv S_{k_2} \pmod{n}$ , and therefore

$$0 \equiv S_{k_2} - S_{k_1} = a_{k_1+1} + \dots + a_{k_2} \pmod{n}.$$

In other words, not only does a solution exist, for some  $k \leq n$  a solution exists which we can scoop up quite efficiently, by picking up a consecutive run of cards from right to left starting with the rightmost card.

Notice that this is not always the only way to win the game, so if this is the only pattern you look for you will often lose to more skillful players. For instance, in our example of  $n = 5$ , the sequence (which we will now reduce mod 5) 1, 2, 4 already has a solution but no consecutively numbered solution.

An interesting question that we will leave the reader with is the following: fix  $n$  and assume that  $N$  is much larger than  $n$ : this is effectively the same as drawing with replacement (because after we draw any one card  $a_i$ , the change in the proportion of the cards in the deck which are congruent to  $a_i \pmod{n}$  is negligible

---

<sup>4</sup>We neglect the issue of figuring out exactly how many cards are necessary if  $n$  is moderately large compared to  $N$ . It seems interesting but does not segue into our ultimate goal.

if  $N$  is sufficiently large, and we will never deal more than  $n$  cards). Suppose then that we deal  $1 \leq k \leq n$  cards. What is the probability that a solution exists?

Anyway, we have proven the following amusing mathematical fact:

**Theorem 10.** *Let  $a_1, \dots, a_n$  be any integers. There exists a nonempty subset  $I \subset \{1, \dots, n\}$  such that  $\sum_{i \in I} a_i \equiv 0 \pmod{n}$ .*

**4.2. The Erdos-Ginzburg-Ziv Theorem.**

After a while it is tempting to change the rules of any game. Suppose we “make things more interesting” by imposing the following additional requirement: we deal cards in sequence as before with a predetermined “modulus”  $n \in \mathbb{Z}^+$ . But this time, instead of winning by picking up any (positive!) number of cards which sum to 0 modulo  $n$ , we must select precisely  $n$  cards  $a_{i_1}, \dots, a_{i_n}$  such that  $a_{i_1} + \dots + a_{i_n} \equiv 0 \pmod{n}$ . Now (again assuming that  $N \gg n$ , or equivalently, dealing with replacement), is it always possible to win eventually? If so, how many cards must be dealt?

Well, certainly at least  $n$ : since the problem is more stringent than before, again if the first  $n - 1$  congruence classes are all  $1 \pmod{n}$  then no solution exists. If we have at least  $n$  instances of  $1 \pmod{n}$  then we can take them and win. On the other hand, if the first  $n - 1$  cards are all 1’s, then by adding up any  $k \leq n - 1$  of them we will get something strictly less than  $n$ , so if the next few cards all come out to be  $0 \pmod{n}$ , then we will not be able to succeed either. More precisely, if in the first  $2n - 2$  cards we get  $n - 1$  instances of  $1 \pmod{n}$  and  $n - 1$  instances of  $0 \pmod{n}$ , then there is no way to select precisely  $n$  of them that add up to  $0 \pmod{n}$ . Thus at least  $2n - 1$  cards may be required. Conversely:

**Theorem 11.** *(Erdős-Ginzburg-Ziv, 1961) Let  $n \in \mathbb{Z}^+$  and  $a_1, \dots, a_{2n-1} \in \mathbb{Z}$ . There exists a subset  $I \subset \{1, \dots, 2n - 1\}$  such that:*

- (i)  $\#I = n$ .
- (ii)  $\sum_{i \in I} a_i \equiv 0 \pmod{n}$ .

*Proof.* (C. Bailey and R.B. Richter) The first step is to deduce the theorem for  $n = p$  a prime using Chevalley-Warning. The second step is to show that if the theorem holds for  $n_1$  and for  $n_2$ , it holds also for  $n_1 n_2$ .

Step 1: Suppose  $n = p$  is a prime number. Let  $a_1, \dots, a_{2p-1} \in \mathbb{Z}$ . Consider the following elements of the polynomial ring  $\mathbb{F}_p[t_1, \dots, t_{2p-1}]$ :

$$P_1(t_1, \dots, t_{2p-1}) = \sum_{i=1}^{2p-1} a_i t_i^{p-1},$$

$$P_2(t_1, \dots, t_{2p-1}) = \sum_{i=1}^{2p-1} t_i^{p-1}.$$

Since  $P_1(0) = P_2(0) = 0$  and  $\deg(P_1) + \deg(P_2) = 2p - 2 < 2p - 1$ , by Chevalley-Warning there exists  $0 \neq x = (x_1, \dots, x_{2p-1}) \in \mathbb{F}_p^{2p-1}$  such that

$$(4) \quad \sum_{i=1}^{2p-1} a_i x_i^{p-1} = 0,$$

$$(5) \quad \sum_{i=1}^{2p-1} x_i^{p-1} = 0.$$

Put

$$I = \{1 \leq i \leq 2p-1 \mid x_i \neq 0\}.$$

Since (as usual!)  $x^{p-1}$  is equal to 1 if  $x \neq 0$  and 0 if  $x = 0$ , (4) and (5) yield:

$$\begin{aligned} \sum_{i \in I} a_i &\equiv 0 \pmod{p}, \\ \sum_{i \in I} 1 &\equiv 0 \pmod{p}. \end{aligned}$$

But we have  $0 < \#I < 2p$ , and therefore  $\#I = p$ , completing the proof of Step 1.

Step 2: Because we know the theorem is true for all primes  $n$ , by induction we may assume that  $n = km$  for  $1 < k$ ,  $m < n$  (i.e.,  $n$  is composite) and, by induction, that the theorem holds for  $k$  and  $m$ .

By an easy induction on  $r$ , one sees that if for any  $r \geq 2$  we have  $rk - 1$  integers  $a_1, \dots, a_{rk-1}$ , then there are  $r - 1$  pairwise disjoint subsets of  $I_1, \dots, I_{r-1}$  of  $\{1, \dots, rk - 1\}$ , each of size  $k$ , such that for all  $1 \leq j \leq r - 1$  we have  $\sum_{i \in I_j} a_i \equiv 0 \pmod{k}$ . Apply this with  $r = 2m$  to our given set of  $2n - 1 = (2mk) - 1$  integers: this gives  $2m - 1$  pairwise disjoint subsets  $I_1, \dots, I_{2m-1} \subset \{1, \dots, 2n - 1\}$ , each of size  $k$ , such that for all  $1 \leq j \leq 2m - 1$  we have

$$\sum_{i \in I_j} a_i \equiv 0 \pmod{k}.$$

Now, for each  $j$  as above, put

$$b_j = \sum_{i \in I_j} a_i, \quad b'_j = \frac{b_j}{k}.$$

We thus have  $2m - 1$  integers  $b'_1, \dots, b'_{2m-1}$ . Again using our inductive hypothesis, there exists  $J \subset \{1, \dots, 2m - 1\}$  such that  $\#J = m$  and  $\sum_{j \in J} b'_j \equiv 0 \pmod{m}$ . Let  $I = \bigcup_{j \in J} I_j$ . Then  $\#I = km = n$  and

$$\sum_{i \in I} a_i \equiv \sum_{j \in J} \sum_{i \in I_j} a_i \equiv \sum_{j \in J} kb'_j \equiv 0 \pmod{km}.$$

□

### 4.3. EGZ theorems in finite groups.

This application of Chevalley-Waring – one which makes good use of our ability to choose multiple polynomials – is apparently well-known to combinatorial number theorists. But I didn't know about it until Patrick Corn brought it to my attention.

As with the Chevalley-Waring theorem itself, the EGZ theorem is sort of a prototype for a whole class of problems in combinatorial algebra. In any group  $G$  (which, somewhat unusually, we will write additively even if it is not commutative) a **zero sum sequence** is a finite sequence  $x_1, \dots, x_n$  of elements of  $G$  such that (guess what?)  $x_1 + \dots + x_n = 0$ . By a **zero sum subsequence** we shall mean the sequence  $x_{i_1}, \dots, x_{i_k}$  associated to a nonempty subset  $I \subset \{1, \dots, n\}$ . In this

language, our Theorem 10 says that any sequence of  $n$  elements in  $\mathbb{Z}/n\mathbb{Z}$  has a zero sum subsequence. The same argument proves the following result:

**Theorem 12.** *Let  $G$  be a finite group (not necessarily commutative), of order  $n$ . Then any sequence  $x_1, \dots, x_n$  in  $G$  has a zero sum subsequence.*

Some EGZ-type theorems in this context are collected in the following result.

**Theorem 13.** *(EGZ for finite groups)*

*a) (Erdős-Ginzburg-Ziv, 1961) Let  $G$  be a finite solvable group of order  $n$  and  $x_1, \dots, x_{2n-1} \in G$ . Then there exist distinct indices  $i_1, \dots, i_n$  (not necessarily in increasing order) such that  $x_{i_1} + \dots + x_{i_n} = 0$ .*

*b) (Olson, 1976) Same as part a) but for any finite group.*

*c) (Sury, 1999) Same as part a) but the indices can be chosen in increasing order:  $i_1 < \dots < i_n$ .*

*d) (Sury, 1999) The conclusion of part c) holds for a finite group  $G$  provided it holds for all of its Jordan-Hölder factors.*

We draw the reader's attention to the distinction between the results of parts a) and b) and those of c) and d): in the first two parts, we are allowed to reorder the terms of the subsequence, whereas in the latter two we are not. In a commutative group it makes no difference – thus, the generalization to all finite abelian groups is already contained in the original paper of EGZ – but in a noncommutative group the desire to preserve the order makes the problem significantly harder.

The inductive argument in Step 2 of Theorem 11 is common to all the proofs, and is most cleanly expressed in Sury's paper as the fact that the class of finite groups for which EGZ holds is closed under extensions. Thus the case in which  $G$  is cyclic of prime order is seen to be crucial. In 1961 Erdős, Ginzburg and Ziv gave an “elementary” proof avoiding Chevalley-Warning. Nowadays there are several proofs available; a 1993 paper of Alon and Dubiner presented at Erdős' 80th birthday conference gives five different proofs. Olson's proof also uses only elementary group theory, but is not easy. In contrast, Sury's paper makes full use of Chevalley-Warning and is the simplest to read: it is only three pages long.

Sury's result has the intriguing implication that it would suffice to prove the EGZ theorem for all finite simple groups (which are now completely classified...). To my knowledge no one has followed up on this.

There is another possible generalization of the EGZ theorem to finite abelian, but non-cyclic, groups. Consider for instance  $G(n, 2) := \mathbb{Z}_n \times \mathbb{Z}_n$ , which of course has order  $n^2$ . Rather than asking for the maximal length of a sequence without an  $n^2$ -term zero sum subsequence, one might ask for the maximal length of a sequence without an  $n$ -term zero sum subsequence. (One might ask many other such questions, of course, but in some sense this is the most reasonable “vector-valued analogue” of the EGZ situation.) A bit of thought shows that the analogous lower bound is given by the sequence consisting of  $n - 1$  instances each of  $(0, 0)$ ,  $(0, 1)$ ,  $(1, 0)$  and  $(1, 1)$ : in other words, this is the “obvious” sequence with no  $n$ -term zero-sum subsequence, of length  $4(n - 1)$ . It was conjectured by A. Kemnitz in 1983 that indeed any sequence in  $G(n, 2)$  of length at least  $4n - 3$  has an  $n$ -term zero sum subsequence. Kemnitz's conjecture was proved in 2003 independently by C.

Reiher (an undergraduate!) and C. de Fiore (a high school student!!). Both proofs use the Chevalley-Waring theorem, but in quite intricate and ingenious ways.

For any positive integer  $k$ , define  $G(n, d) = (Z_n)^d$ , the product of  $d$  copies of the cyclic group of order  $n$ , and consider lengths of sequences without an  $n$ -term zero sum subsequence: let us put  $f(n, d)$  for the maximal length of such a sequence. Analogues of the above sequences with  $\{0, 1\}$ -coordinates give

$$f(n, d) \geq 2^d(n - 1).$$

In 1973 Heiko Harborth established the (much larger) upper bound

$$f(n, d) \leq n^d(n - 1).$$

Harborth also computed  $G(3, 3) = 18 > 2^3(3 - 1)$ : i.e., in this case the “obvious” examples do not have maximal length! It seems that the computation of  $G(n, 3)$  for all  $n$  – or still more, of  $G(n, d)$  for all  $d$  – would be a significant achievement.