UNIVERSITY OF PUERTO RICO
RIO PIEDRAS CAMPUS
FACULTY OF NATURAL SCIENCES
DEPARTMENT OF MATHEMATICS

# Solvability of systems of polynomial equations with multivariate polynomials as coefficients

By

Carlos Emanuel Seda Damiani

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN MATHEMATICS
AT THE UNIVERSITY OF PUERTO RICO

June 17th, 2020

APPROVED BY THE MASTER OF SCIENCE ADVISORY COMMITTEE
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN MATHEMATICS
AT THE UNIVERSITY OF PUERTO RICO

ADVISOR:

Ivelisse Rubio, Ph.D.
University of Puerto Rico, Río Piedras
Department of Computer Science

READERS:

Luis A. Medina, Ph.D.
University of Puerto Rico, Río Piedras
Department of Mathematics

Raúl Figueroa, Ph.D.
University of Puerto Rico, Río Piedras
Department of Mathematics

**Abstract**

In [3] Castro, Moreno and Rubio generalize the results of Moreno-Moreno's theorem that gives a bound for the power of a prime $p$ to divide the number of common zeros of the multivariate polynomials $F_1, ..., F_t$ over a finite field $\mathbb{F}_q$. This generalization regarded the coefficients of the polynomials to be uni-variate polynomials over a finite field instead of plain elements of the finite field. The result led to improve a theorem of Carlitz, for the estimation of the number of variables needed so that a system of polynomial equations with coefficients in $\mathbb{F}_q[X]$ can have non-trivial zeros. We generalize the results of Castro, Moreno and Rubio to polynomials whose coefficients are multivariate polynomials over finite fields.

# Acknowledgments

I want to begin by thanking Dr. Ivelisse Rubio whose mentoring, feedback and lessons where crucial in seeing this work through to the end. I also want to dedicate this work to my parents Mayra Damiani and Jaime Seda, who have been very supportive throughout all of my academic career and it is thanks to their hard work that I have been able to pursue my academical goals.

I also want to give my most sincere thanks to Dr. Javier Figueroa and Dr. Eduardo Nicolau who gave me the opportunity and privilege to be a participant of the *PRLSAMP Bridge to the Doctorate* fellowship program.

Finally, I want to thank Dr. Luis A. Medina and Dr. Raúl Figueroa for being part of my thesis committee and for their feedback on this work.

# Contents

# Chapter 1

# Introduction

We look all the way back to 1936, when Claude Chevalley [1] proved a conjecture of Artin showing the existence of non-trivial roots of a system of polynomials over a finite field $\mathbb{F}_q$.

*If $F_1, \ldots, F_t$ are polynomials in $n$ variables, with coefficients in $\mathbb{F}_q$, the finite field with $q = p^f$ elements, which have the trivial zero and $\deg(F_1) + \cdots + \deg(F_t) < n$, then $F_1, \ldots, F_t$ has a non-trivial zero in $(\mathbb{F}_q)^n$.*

This result has seen several improvements throughout the years, some of these include a theorem by Katz [8] which provides a bound on the power of $q$ that divides the number of common zeros of the system of polynomials. In 1995, Carlos Moreno and Oscar Moreno [9], instead of the degree of the polynomials, consider the $p$-weight degrees of the polynomials $F_1, \ldots, F_t$ and present a result that improves Katz' result in many cases. It also works for some cases where Katz' result does not give any information.

Another result that stems from Chevalley's work came in 1952 when Leonard Carlitz [7] applies the result of Chevalley to present a sufficient condition on the number of variables $Y_1, \ldots, Y_n$ that are needed for a system of polynomials with coefficients in $\mathbb{F}_q[X]$ to have a non-trivial solution in $(\mathbb{F}_q[X])^n$. The approach used by Carlitz combined with the method of restriction of scalars is used in the work of Castro, Moreno and Rubio [3] to obtain a generalization of Moreno-Moreno's Theorem 1 in [9] for systems of polynomials $F_1, \ldots, F_t$ in the variables $Y_1, \ldots, Y_n$ which have coefficients that are polynomials in the variable $X$ (that is, $F_i \in (\mathbb{F}_q[X])[Y_1, \ldots, Y_n]$ for every $i$). A significant improvement to the result of Carlitz was obtained as a consequence of this generalization. In this work we will follow the methods used in [3] in order to further generalize the results of Castro, Moreno and Rubio to systems of polynomials $F_1, \ldots, F_t$ that have multivariate polynomials as coefficients, $F_i \in (\mathbb{F}_q[X_1, \ldots, X_w])[Y_1, \ldots, Y_n]$, and, as well, generalize the improvement to Carlitz' result.

# Chapter 2

# Preliminaries

In this chapter we provide the definitions and results needed to understand our work. We omit the proof of the results that are well known.

## 2.1 Fields and Vector Spaces

First we provide definitions and results from the theory of fields and vector spaces taken from [2] and [10]. These results and definitions will be used in Section 2.2 to define *finite fields* and to prove some of their properties.

**Definition 2.1.** A ring $(F, +, \cdot)$ is a **field** if and only if all of the following hold

1. $(F, +)$ is a commutative group.

2. $(F^{\times}, \cdot)$ is a commutative group.

3. $a(b + c) = ab + ac$ for all $a, b, c \in R$

**Notation.** We denote the multiplicative and additive identities of a **field** as 1 and 0 respectively.

**Definition 2.2.** The **characteristic** of a ring $R$, denoted by $\text{char}(R)$, is the least positive integer $n$ such that $nr = 0$ for all $r \in R$. If no such integer exists, we say that $R$ has characteristic 0.

**Proposition 2.3.** The characteristic of a field $F$ is either 0 or a prime $p$.

*Proof.* Let $F$ be a field and suppose that $\text{char}(F)$ is not 0. Suppose that $\text{char}(F) = n$, where $n = rs$, $r, s \in \mathbb{Z}$ and $r > 1$, $s > 1$. Then $(r1)(s1) = (rs)(1) = n1 = 0$, which implies that $r1 = 0$ or $s1 = 0$ since $F$ has no proper zero divisors. Therefore, $rk = r1k = 0$ or $sk = s1k = 0$ for all $k \in F$. This is a contradiction to the minimality of $n$, since both $r$ and $s$ are smaller than $n$. Therefore, $r = 1$ or $s = 1$ and the characteristic of $F$ must be a prime $p$. ■

**Definition 2.4.** A set $V$ is said to be a **vector space** over a field $F$ if $V$ is an abelian group under addition and, for each $a, b$ in $F$ and $u, v$ in $V$ the following are true.

1. $a(v + u) = av + au$
2. $(a + b)v = av + bv$
3. $a(bv) = (ab)v$
4. $1v = v$

**Definition 2.5.** A subset $S = \{v_1, v_2 \ldots, \}$ of a vector space $V$ is said to be **linearly independent** if $a_1v_1 + a_2v_2 + \cdots = 0$ with $a_i \in F$ implies that $a_1 = a_2 = \ldots = 0$. If this is not the case, the elements of $S$ are called **linearly dependent**.

**Definition 2.6.** Let $V$ be a vector space over $F$. A subset $B$ of $V$ is called a **basis** of $V$ if $B$ is linearly independent over $F$ and every element of $V$ is a linear combination of elements of $B$.

**Definition 2.7.** A vector space that has a basis consisting of a finite amount of elements $n$ is said to have **dimension** $n$.

**Definition 2.8.** A field $E$ is an **extension** of a field $F$ if $F \subseteq E$ and the operations of $F$ are those of $E$ restricted to $F$.

**Definition 2.9.** If $E$, considered as a vector space over $F$, has finite dimension, then E is called a **finite extension** of $F$.

**Definition 2.10.** Let $E$ be a finite extension of a field $F$. It is said that $E$ has **degree** $[E : F]$ over $F$ if $E$ has dimension $[E : F]$ as a vector space over $F$.

**Definition 2.11.** Let $E$ be extension of a field $F$. Given a set of elements $A$ in the larger field $E$ we denote by $F(A)$ the smallest sub-extension field of $F$ that contains the elements of $A$. We say $F(A)$ is constructed by **adjointing** of the elements of $A$ to $F$ or *generated* by $A$.

**Definition 2.12.** Let $E$ be an extension of a field $F$ and let $f(X) \in F[X]$, where $F[X]$ is the *ring of polynomials with coefficients in* $F$, where $f(X)$ a non-constant polynomial. We say that $f(X)$ *splits* in $E$ if there are elements $a \in F$ and $a_1, a_2, \ldots, a_n \in E$ such that

$$f(X) = a(X - a_1)(X - a_2) \cdots (X - a_n).$$

$E$ is called a **splitting field** of $f(X)$ over $F$ if $E = F(a_1, a_2, \ldots, a_n)$.

**Theorem 2.13.** Let $F$ be a field and let $f(X) \in F[X]$. Then any two splitting fields of $f(X)$ over $F$ are isomorphic.

## 2.2 Finite Fields

The results that we present on Chapters 3 and 4 are of system of polynomials that have coefficients and take values over finite fields, we make use of properties of finite fields like the characteristic, the size and degrees of their extensions. Now, we provide definitions and results from number theory and finite field theory taken from [4] and [6].

**Definition 2.14.** A **finite field** $\mathbb{F}_q$ is a *field* that contains a finite number of elements $q$.

**Theorem 2.15.** If $\mathbb{F}_q$ is a finite field with $q$ elements, then $q = p^f$.

*Proof.* Let $\mathbb{F}_q$ be a finite field with $q$ elements. Since every field has characteristic either 0 or a prime $p$ we have that the characteristic of $\mathbb{F}_q$ is $p$ because it is finite, thus there is only a finite number $p$ such that $1 + 1 + \cdots + 1 = 0$. Thus $\mathbb{F}_q$ has copy of $\mathbb{F}_p$, since it has characteristic $p$. This means that $\mathbb{F}_q$ is a field extension of $\mathbb{F}_p$, and we may see $\mathbb{F}_q$ as a vector space over $\mathbb{F}_p$. If the dimension is $f$, then let $\mu = \{\mu_1, \mu_2, \ldots, \mu_f\}$ be a basis for $\mathbb{F}_q$. Every $x \in \mathbb{F}_q$ can be written as a linear combination

$$x = x_1\mu_1 + \cdots + x_f\mu_f$$

and there are $p$ choices for each of these $x_i$, thus a total of $p^f$ different elements in

$\mathbb{F}_q$. ■

From now on, $p$ denotes a prime and $q = p^f$.

**Proposition 2.16.** If $\mathbb{F}_q$ is a finite field with $q = p^f$ elements, then for every $x \in \mathbb{F}_q$ it is true that $x^q = x$.

*Proof.* The case where $x = 0$ is trivial. Now notice that the other $q - 1$ nonzero elements of $\mathbb{F}_q$ form a group under the multiplication operation of the finite field. Therefore $x^{q-1} = 1$, since the order of $\mathbb{F}_q - \{0\}$ is $q - 1$, which implies that $x^q = x$ for all $x \in \mathbb{F}_q$. ■

**Notation.** We will denote the group of order $q - 1$ of non-zero elements of $\mathbb{F}_q$ under its multiplication operation as $\mathbb{F}_q^*$.

**Theorem 2.17.** The multiplicative group of a finite field $\mathbb{F}_q^*$ is cyclic.

The following definition for a particular element of $\mathbb{F}_q$ is used in the example given in Section 2.25.

**Definition 2.18.** An element $\alpha \in \mathbb{F}_q$ is called a **primitive root of** $\mathbb{F}_q$ if it is a generator for $\mathbb{F}_q^*$.

**Theorem 2.19.** There exists a unique finite field $\mathbb{F}_q$ for every $q = p^f$.

*Proof.* We have that $\mathbb{F}_q$ contains a copy of $\mathbb{F}_p$, and, since the nonzero elements of $\mathbb{F}_q$ form a multiplicative group of order $q - 1$, every element of $\mathbb{F}_q$ is a zero of the polynomial $f(X) = X^q - X$ by Proposition 2.16. So, $\mathbb{F}_q$ must be a splitting field for $f(X) = X^q - X$ over the subfield of $\mathbb{F}_q$ that is isomorphic to $\mathbb{F}_p$. By Theorem 2.13,

we have that splitting fields are unique up to isomorphisms. Therefore, the finite field $\mathbb{F}_q$ is unique, up to isomorphisms. ■

The following property of finite fields $\mathbb{F}_q$ will be an important part of the method of *restriction of scalars*, discussed in Section 2.4.

**Proposition 2.20.** For all $a, b \in \mathbb{F}_q$, $q = p^f$, we have that $(a+b)^{p^f} = a^{p^f} + b^{p^f}$.

*Proof.* Let $f = 1$ and notice that $(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} a^k b^{p-k} = \binom{p}{0} b^p + \sum_{k=1}^{p-1} \binom{p}{k} a^k b^{p-k} + \binom{p}{p} a^p = b^p + \sum_{k=1}^{p-1} \binom{p}{k} a^k b^{p-k} + a^p$. For $1 \leq k \leq p-1$ we have that

$$\binom{p}{k} = \frac{p(p-1)!}{k!(p-k)!}$$

and since $p$ is prime, no factor in the denominator is a common factor of $p$. Therefore $p \mid \binom{p}{k}$, then $\binom{p}{k} a^k b^{p-k}$ is divisible by $p$ and hence the whole sum is divisible by $p$ as well. Since $\mathbb{F}_q$ has characteristic $p$, we get that $\sum_{k=1}^{p-1} \binom{p}{k} a^k b^{p-k} = 0$. Assume for $f = k$ that $(a+b)^{p^k} = a^{p^k} + b^{p^k}$ is true. Now $(a+b)^{p^{k+1}} = \left((a+b)^{p^k}\right)^p = \left(a^{p^k} + b^{p^k}\right)^p$ because of the induction hypothesis. If $x = a^{p^k}$ and $y = b^{p^k}$, then $= \left(a^{p^k} + b^{p^k}\right)^p = (x+y)^p = x^p + y^p$. Therefore, we have that $(a+b)^{p^{k+1}} = x^p + y^p = a^{p^{k+1}} + b^{p^{k+1}}$. By the principle of mathematical induction we have shown that $(a+b)^{p^f} = a^{p^f} + b^{p^f}$. ■

## 2.3 $p$-weight degrees

Throughout this work, one of the main focuses are going to be the degrees and the *p-weight degrees* of the polynomials in the system. The results we are going to

show will make use of both of these concepts. In this section, we provide definitions and some properties related to the notion of $p$-weight degree.

**Definition 2.21.** Let $d$ be a non-negative integer, $d = a_0 + a_1 p + \cdots + a_r p^r$, where $0 \leq a_i < p$. We define and denote the **$p$-weight of** $d$ as $\sigma_p(d) = a_0 + a_1 + \cdots + a_r$. We call $a_0 + a_1 p + \cdots + a_r p^r$ the *base $p$ expansion of $d$*.

**Example.** For $p = 3$, consider the base 3 expansion of 5 and 8,

$$5 = 2 + 1 \cdot 3 \text{ and } 8 = 2 + 2 \cdot 3.$$

Then $\sigma_3(5) = 3$ and $\sigma_3(8) = 4$

**Proposition 2.22.** Let $d \geq 0$ be an integer. Then $\sigma_p(d) \leq d$.

*Proof.* Let $d = a_0 + a_1 p + \cdots + a_r p^r$ be the base $p$ expansion of the integer $d$. If $d < p$, then $d = d \cdot p^0$, and $\sigma_p(d) = d$. If $d \geq p$, then $\sigma_p(d) = a_0 + \cdots + a_r < a_0 + a_1 p + \cdots + a_r p^r = d$. ■

**Definition 2.23.** The **$p$-weight degree of a monomial** $\mathbf{Y^d} = Y_1^{d_1} Y_2^{d_2} \cdots Y_n^{d_n}$ is $\omega_p(\mathbf{Y^d}) = \sigma_p(d_1) + \cdots + \sigma_p(d_n)$. The **$p$-weight degree of a polynomial** $F(Y_1, \ldots, Y_n) = \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{Y^d}$ is $\omega_p(F) = \max\{\omega_{\mathbf{p}}(\mathbf{Y^d})\}$.

**Example 1.** Consider the polynomial

$$F(Y_1, \ldots, Y_5) = Y_1 Y_2 Y_3^2 + Y_1 Y_4 + Y_5^6$$

in $n = 5$ variables, $\deg(F) = 6$ and taking values over $\mathbb{F}_3$. The 3-weight degree of

the monomials of $F$ are $\omega_3(Y_1Y_2Y_3^2) = 2\sigma_3(1) + \sigma_3(2) = 4$, $\omega_3(Y_1Y_4) = 2\sigma_3(1) = 2$ and $\omega_3(Y_5^6) = \sigma_3(6) = 2$. Therefore, the 3-weight degree of $F$ is $\omega_3(F) = 4 < 6 = \deg(F)$.

**Example 2.** Consider the polynomial

$$F(Y_1, \ldots, Y_5) = Y_1Y_2Y_3Y_4 + Y_1Y_4 + Y_5$$

in $n = 5$ variables, $\deg(F) = 4$ and taking values over $\mathbb{F}_3$. The 3-weight for the monomials of $F$ are $\omega_3(Y_1Y_2Y_3Y_4) = 4\sigma_3(1) = 4$, $\omega_3(Y_1Y_4) = 2\sigma_3(1) = 2$ and $\omega_3(Y_5) = 1$. Therefore, the 3-weight degree of $F$ is $\omega_3(F) = 4 = \deg(F)$.

From these examples we notice that $\omega_p(F) \leq \deg(F)$. We provide a proof of this property.

**Proposition 2.24.** Let $F(Y_1, \ldots, Y_n) = \sum_{\mathbf{d}} c_{\mathbf{d}} \mathbf{Y}^{\mathbf{d}}$ be a polynomial, then $\omega_p(F) \leq \deg(F)$.

*Proof.* Let $\mathbf{Y}^{\mathbf{d}} = Y_1^{d_1}Y_2^{d_2}\cdots Y_n^{d_n}$ be a monomial, we know that $\deg(\mathbf{Y}^{\mathbf{d}}) = d_1 + \cdots + d_n$. Now, by Proposition 2.22, we have that $\omega_p(\mathbf{Y}^{\mathbf{d}}) = \sigma_p(d_1) + \cdots + \sigma_p(d_n) \leq d_1 + \cdots + d_n = \deg(\mathbf{Y}^{\mathbf{d}})$. Therefore, for each of the monomials $\mathbf{Y}^{\mathbf{d}}$ of the polynomial $F$, we have that $\omega_p(\mathbf{Y}^{\mathbf{d}}) \leq \deg(\mathbf{Y}^{\mathbf{d}})$. Now let $\mathbf{Y}^{\mathbf{d_1}}$ and $\mathbf{Y}^{\mathbf{d_2}}$ be the monomials of $F$ such that $\omega_p(F) = \max\{\omega_p(\mathbf{Y}^{\mathbf{d}})\} = \omega_p(\mathbf{Y}^{\mathbf{d_1}})$ and $\deg(F) = \max\{\deg(\mathbf{Y}^{\mathbf{d}})\} = \deg(\mathbf{Y}^{\mathbf{d_2}})$. Now $\omega_p(F) = \omega_p(\mathbf{Y}^{\mathbf{d_1}}) \leq \deg(\mathbf{Y}^{\mathbf{d_1}}) \leq \deg(\mathbf{Y}^{\mathbf{d_2}}) = \deg(F)$. ■

**Remark.** It is very tempting to conjecture that given two polynomials $F$ and $G$ with $\deg(F) < \deg(G)$, then $\omega_p(F) < \omega_p(G)$. It turns out that this statement is

false. Let $p = 3$, $F(Y) = Y^2$, and $G(Y) = Y^9$. It is clear that $\deg(F) < \deg(G)$ but $\omega_3(F) = 2 > 1 = \omega_3(G)$.

## 2.4 Restriction of Scalars

The method of restriction of scalars is one of the main tools used in our research and we present it now with details.

What the method of restriction of scalars does is that it changes a polynomial in $n$ variables, taking values in a field $\mathbb{F}_q$, $q = p^f$, to a new polynomial in $nf$ variables taking values in $\mathbb{F}_p$.

We start with a polynomial $F(Y_1, \ldots, Y_n)$ with coefficients in and taking values in $\mathbb{F}_q$. Since $F(Y_1, \ldots, Y_n)$ is a sum of terms, we only need to consider a generic term

$$\alpha Y_1^{d_1} Y_2^{d_2} \ldots Y_n^{d_n} \tag{2.1}$$

of degree $d = d_1 + d_2 + \cdots + d_n$. Since we can look at $\mathbb{F}_{p^f}$ as a vector space of dimension $f$ over $\mathbb{F}_p$ there is a a basis

$$\mu = \{\mu_1, \ldots, \mu_f\} \tag{2.2}$$

of $\mathbb{F}_{p^f}$ over $\mathbb{F}_p$. Then, each variable $Y_j \in \mathbb{F}_{p^f}$ of the monomial (2.1) can be expressed as a linear combination of the basis elements and new variables $Z_k^{(j)}$ that takes values from $\mathbb{F}_p$, $Y_j = Z_1^{(j)}\mu_1 + \cdots + Z_f^{(j)}\mu_f$, thus

$$\alpha Y_1^{d_1} Y_2^{d_2} \ldots Y_n^{d_n} = \alpha \left( \sum_{k=1}^{f} Z_k^{(1)} \mu_k \right)^{d_1} \cdots \left( \sum_{k=1}^{f} Z_k^{(n)} \mu_k \right)^{d_n}. \tag{2.3}$$

Rewrite each $d_j$ using its base $p$ expansion, so $d_j = a_{j0} + a_{j1}p + \cdots + a_{jr}p^r$ and since for all $a, b \in \mathbb{F}_q$ we have that $(a+b)^{p^f} = a^{p^f} + b^{p^f}$ and $a^q = a$ we get

$$\alpha Y_1^{d_1} Y_2^{d_2} \dots Y_n^{d_n} = \alpha \left( \sum_{k=1}^{f} Z_k^{(1)} \mu_k \right)^{a_{10}+a_{11}p+\cdots+a_{1r}p^r} \cdots \left( \sum_{k=1}^{f} Z_k^{(n)} \mu_k \right)^{a_{n0}+a_{n1}p+\cdots+a_{nh}p^h}$$

$$= \alpha \left[ \left( \sum_{k=1}^{f} Z_k^{(1)} \mu_k \right)^{a_{10}} \left( \sum_{k=1}^{f} Z_k^{(1)} \mu_k^p \right)^{a_{11}} \cdots \left( \sum_{k=1}^{f} Z_k^{(1)} \mu_k^{p^r} \right)^{a_{1r}} \right] \cdots \tag{2.4}$$

$$\left[ \left( \sum_{k=1}^{f} Z_k^{(n)} \mu_k \right)^{a_{n0}} \left( \sum_{k=1}^{f} Z_k^{(n)} \mu_k^p \right)^{a_{n1}} \cdots \left( \sum_{k=1}^{f} Z_k^{(n)} \mu_k^{p^h} \right)^{a_{nh}} \right],$$

where the degree of this expression is the $p$-weight degree of our original term. Doing the same for each of the monomials in $F$ and combining terms for each basis element $\mu_k$ we obtain

$$F(Y_1, \dots, Y_n) = \sum_{k=1}^{f} G_k(Z_1^{(1)}, \dots, Z_f^{(1)}, \dots, Z_1^{(n)}, \dots, Z_f^{(n)}) \mu_k, \tag{2.5}$$

and since $\mu$ is a basis of $\mathbb{F}_{p^f}$ over $\mathbb{F}_p$ we have that $F(Y_1, \dots, Y_n) = 0$ if and only if $G_k = 0$ for each $1 \leq k \leq f$. Hence, looking for the zeros of a polynomial $F$ in $n$ variables $Y_1, \dots, Y_n$, and of degree $d$ is the same as finding the common zeros of a system of $f$ polynomials $G_1, \dots, G_f$ in $fn$ variables, with coefficients in $\mathbb{F}_p$ and of degree $\deg(G_k) \leq \omega_p(F)$.

We now have the following proposition that summarizes the result obtained after applying the method of restriction of scalars to a system of polynomials.

**Proposition 2.25** (Moreno-Moreno, Lemma 1, [9])**.** [Restriction of Scalars] Let

$q = p^f$ and $F_1, \ldots, F_t$ be polynomials in $n$ variables $Y_1, \ldots, Y_n$ defined over $\mathbb{F}_q$. Let $N(\{F_i\}, \mathbb{F}_q)$ be the number of solutions to $F_1 = F_2 = \ldots = F_t = 0$. Then, there exists a system $\{{}_iG_k\}$, where $1 \leq i \leq t$ and $1 \leq k \leq f$, of $ft$ polynomials in $nf$ variables $Z_1^{(1)}, \ldots, Z_f^{(1)}, \ldots, Z_1^{(n)}, \ldots, Z_f^{(n)}$ over $\mathbb{F}_p$, $\deg({}_iG_k) \leq \omega_p(F_i)$, such that $N(\{F_i\}, \mathbb{F}_q) = N(\{{}_iG_k\}, \mathbb{F}_p)$. Where $N(\{{}_iG_k\}, \mathbb{F}_p)$ is the number of solutions to the system $\{{}_iG_k = 0\}$ over $\mathbb{F}_p$.

We now take a polynomial in 3 variables taking values over $\mathbb{F}_9$ and transform it into a polynomial in 6 variables that takes values over $\mathbb{F}_3$.

**Example.** Consider the finite field $\mathbb{F}_3$ and the polynomial $G(X) = X^2 + X + 2$. It can be checked that this polynomial is irreducible over $\mathbb{F}_3$. Let $\alpha$ be such that $\alpha^2 + \alpha + 2 = 0$. The splitting field for the polynomial $G(X)$ is

$$\mathbb{F}_3[X] \Big/ (X^2 + X + 2) \cong \mathbb{F}_3(\alpha) \cong \mathbb{F}_{3^2}.$$

Notice that $\alpha^0 = 1$, $\alpha^1 = \alpha$, $\alpha^2 = 2\alpha+1$, $\alpha^3 = 2\alpha+2$, $\alpha^4 = 2$, $\alpha^5 = 2\alpha$, $\alpha^6 = \alpha+2$ and $\alpha^7 = \alpha + 1$ and therefore $\alpha$ is a primitive root of $\mathbb{F}_9$. Also the finite field $\mathbb{F}_9$ is a vector space of dimension 2 over $\mathbb{F}_3$ with the basis $\{1, \alpha\}$.

Consider the polynomial

$$F(Y_1, Y_2, Y_3) = \alpha Y_1^{17} Y_2^5 + Y_3^4 \tag{2.6}$$

in $\mathbb{F}_9[Y_1, Y_2, Y_3]$. Note that $\omega_3(F) = 8$ and $\deg(F) = 22$. Using the method of restriction of scalars with this basis we have that, the variable $Y_j$ taking values from

$\mathbb{F}_9$ can be written as a linear combination of the basis elements 1 and $\alpha$ such that $Y_j = Z_1^{(j)} + \alpha Z_2^{(j)}$ where the new variables $Z_i^{(j)}$ takes values from $\mathbb{F}_3$. The first monomial of polynomial (2.6), can be rewritten as

$$\alpha Y_1^{17} Y_2^5 = \alpha \left( Z_1^{(1)} + \alpha Z_2^{(1)} \right)^{17} \left( Z_1^{(2)} + \alpha Z_2^{(2)} \right)^5 .$$

Rewrite the exponents with respect to their base 3 expansion:

$$\alpha Y_1^{17} Y_2^5 = \alpha \left( Z_1^{(1)} + \alpha Z_2^{(1)} \right)^{2+2\cdot 3+1\cdot 3^2} \left( Z_1^{(2)} + \alpha Z_2^{(2)} \right)^{2+1\cdot 3}$$

Proposition 2.20 implies that

$$\alpha Y_1^{17} Y_2^5 = \alpha \left( Z_1^{(1)} + \alpha Z_2^{(1)} \right)^2 \left( [Z_1^{(1)}]^3 + \alpha^3 [Z_2^{(1)}]^3 \right)^2 \left( [Z_1^{(1)}]^{3^2} + \alpha^{3^2} [Z_2^{(1)}]^{3^2} \right)$$

$$\left( Z_1^{(2)} + \alpha Z_2^{(2)} \right)^2 \left( [Z_1^{(2)}]^3 + \alpha^3 [Z_2^{(2)}]^3 \right).$$

By Proposition 2.16 we get

$$\alpha Y_1^{17} Y_2^5 = \alpha \left( Z_1^{(1)} + \alpha Z_2^{(1)} \right)^2 \left( Z_1^{(1)} + \alpha^3 Z_2^{(1)} \right)^2 \left( Z_1^{(1)} + \alpha Z_2^{(1)} \right)$$

$$\left( Z_1^{(2)} + \alpha Z_2^{(2)} \right)^2 \left( Z_1^{(2)} + \alpha^3 Z_2^{(2)} \right)$$

$$= \alpha \left( Z_1^{(1)} + \alpha^3 Z_2^{(1)} \right) \left( Z_1^{(1)} + \alpha^3 Z_2^{(1)} \right)^2 \left( Z_1^{(2)} + \alpha Z_2^{(2)} \right)^2 \left( Z_1^{(2)} + \alpha^3 Z_2^{(2)} \right)$$

$$= \alpha \left( Z_1^{(1)} + \alpha^3 Z_2^{(1)} \right)^3 \left( Z_1^{(2)} + \alpha Z_2^{(2)} \right)^2 \left( Z_1^{(2)} + \alpha^3 Z_2^{(2)} \right).$$

Again, by Proposition 2.16

$$\alpha Y_1^{17} Y_2^5 = \alpha \left( Z_1^{(1)} + \alpha Z_2^{(1)} \right) \left( Z_1^{(2)} + \alpha Z_2^{(2)} \right)^2 \left( Z_1^{(2)} + \alpha^3 Z_2^{(2)} \right)$$

expanding the square on the second term yields

$$\alpha Y_1^{17} Y_2^5 = \alpha \left( Z_1^{(1)} + \alpha Z_2^{(1)} \right) \left( [Z_1^{(2)}]^2 + 2\alpha Z_1^{(2)} Z_2^{(2)} + \alpha^2 [Z_2^{(2)}]^2 \right) \left( Z_1^{(2)} + \alpha^3 Z_2^{(2)} \right).$$

Multiplying these factors out and grouping them with respect to the elements of the basis we obtain

$$\alpha Y_1^{17} Y_2^5 = 1 \cdot \left( Z_1^{(1)} [Z_1^{(2)}]^2 Z_2^{(2)} + 2 Z_1^{(1)} Z_1^{(2)} [Z_2^{(2)}]^2 + 2 Z_1^{(1)} Z_2^{(2)} + Z_2^{(1)} Z_1^{(2)} + 2 Z_2^{(1)} Z_2^{(2)} \right) +$$

$$\alpha \cdot \left( Z_1^{(1)} [Z_1^{(2)}]^2 Z_2^{(2)} + Z_1^{(1)} Z_1^{(2)} + Z_1^{(1)} Z_2^{(2)} + Z_2^{(1)} Z_1^{(2)} + Z_2^{(1)} Z_2^{(2)} \right). \tag{2.7}$$

For the second monomial of polynomial (2.6) we have

$$Y_3^4 = \left( Z_1^{(3)} + \alpha Z_2^{(3)} \right)^4 = \left( Z_1^{(3)} + \alpha Z_2^{(3)} \right) \left( Z_1^{(3)} + \alpha Z_2^{(3)} \right)^3.$$

By Proposition 2.16

$$Y_3^4 = \left( Z_1^{(3)} + \alpha Z_2^{(3)} \right) \left( Z_1^{(3)} + \alpha^3 Z_2^{(3)} \right) = [Z_1^{(3)}]^2 + 2 Z_1^{(3)} Z_2^{(3)} + 2 [Z_2^{(3)}]^2. \tag{2.8}$$

Adding (2.7) and (2.8) and grouping up the similar terms with respect to the basis

elements, the polynomial (2.6) is

$$F = 1 \cdot \Big( Z_1^{(1)}[Z_1^{(2)}]^2 Z_2^{(2)} + 2Z_1^{(1)} Z_1^{(2)}[Z_2^{(2)}]^2 + 2Z_1^{(1)} Z_2^{(2)} + Z_2^{(1)} Z_1^{(2)} +$$

$$2Z_2^{(1)} Z_2^{(2)} + [Z_1^{(3)}]^2 + 2Z_1^{(3)} Z_2^{(3)} + 2[Z_2^{(3)}]^2 \Big) + \tag{2.9}$$

$$\alpha \cdot \Big( Z_1^{(1)}[Z_1^{(2)}]^2 Z_2^{(2)} + Z_1^{(1)} Z_1^{(2)} + Z_1^{(1)} Z_2^{(2)} + Z_2^{(1)} Z_1^{(2)} + Z_2^{(1)} Z_2^{(2)} \Big).$$

Using the method of restriction of scalars $F(Y_1, Y_2, Y_3)$ is turned into a new polynomial in 6 variables, where each of the coefficients of the basis elements is a polynomial of degree less than $\omega_3(F) = 8$. Also, by Proposition 2.25, $F(Y_1, Y_2, Y_3) = 0$ has the same number of solutions than the system of two equations that is produced by the coefficients of the basis elements.

$$G_1(Z_1^{(1)}, \ldots, Z_2^{(3)}) = Z_1^{(1)}[Z_1^{(2)}]^2 Z_2^{(2)} + 2Z_1^{(1)} Z_1^{(2)}[Z_2^{(2)}]^2 + 2Z_1^{(1)} Z_2^{(2)}$$
$$+ Z_2^{(1)} Z_1^{(2)} + 2Z_2^{(1)} Z_2^{(2)} + [Z_1^{(3)}]^2 + 2Z_1^{(3)} Z_2^{(3)} + 2[Z_2^{(3)}]^2 = 0$$

$$G_2(Z_1^{(1)}, \ldots, Z_2^{(3)}) = Z_1^{(1)}[Z_1^{(2)}]^2 Z_2^{(2)} + Z_1^{(1)} Z_1^{(2)} + Z_1^{(1)} Z_2^{(2)} + Z_2^{(1)} Z_1^{(2)} + Z_2^{(1)} Z_2^{(2)} = 0$$

# Chapter 3

# Previous Results

## 3.1 Chevalley-Warning's Theorem

As stated in the introduction, the starting point of this work is a conjecture by Artin, later proven by Chevalley in 1936, regarding solutions of polynomials with coefficients over a finite field and for systems of polynomials as well.

**Theorem 3.1** (Chevalley, Theorem 1, [1]). *Let $F$ be a polynomial in $n$ variables, with coefficients in a finite field $\mathbb{F}_q$, which has the trivial zero. If $n > \deg(F)$, then $F$ also has a non-trivial zero in $(\mathbb{F}_q)^n$.*

**Theorem 3.2** (Chevalley, Theorem 2, [1]). *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables, with coefficients in a finite field $\mathbb{F}_q$, which have the trivial zero. If $n > \deg(F_1) + \cdots + \deg(F_t)$, then $F_1, \ldots, F_t$ also have a common non-trivial zero in $(\mathbb{F}_q)^n$.*

Notice that these results give us information on the solvability of the system in question. The refinement of Chevalley's Theorem by Warning is one that provides an

estimation of the amount of solutions that a system of polynomials has.

**Theorem 3.3** (Warning, Theorem 3 [11]). *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables and $N$ be the number of their common zeros in $(\mathbb{F}_q)^n$. If $n > \deg(F_1) + \cdots + \deg(F_t)$, then $p \mid N$.*

This theorem is also known as the Chevalley-Warning theorem. Warning's result only provides information on the number of solutions. If $N = 0$ then $p \mid N$ is still true. Adding to the hypothesis the assumption of $F_1, \ldots, F_t$ having the trivial solution (just like in Chevalley's theorem) gives us information on the solvability of the system. This is,

**Corollary 3.3.1.** *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables which have the trivial zero, and $N$ be the number of their common zeros in $(\mathbb{F}_q)^n$. If $n > \deg(F_1) + \cdots + \deg(F_t)$, then there are also non-trivial solutions and $p \mid N$.*

## 3.2 Katz and Moreno-Moreno's theorems

The Chevalley-Warning's theorem 3.3 has seen various improvements, one of the best known was given by Katz in 1971. Katz improved the results of Ax in [5] using tools such as cohomology theory and $p$-adic theory from algebraic topology and algebraic number theory.

**Theorem 3.4** (Katz, Theorem 1.0, [8]). *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables,*

*$N$ be the number of common zeros in $(\mathbb{F}_q)^n$ of $F_1, \ldots, F_t$, and*

$$\mu = \left\lceil \left( \frac{n - \sum_{i=1}^{t} \deg(F_i)}{\max_{1 \le i \le t}\{\deg(F_i)\}} \right) \right\rceil .$$

*If $\mu \geq 0$, then $q^\mu \mid N$.*

Another improvement to Chevalley-Warning's theorem was obtained by C. Moreno and O. Moreno in 1995. Using the restriction of scalars method, Moreno-Moreno transformed the system of $t$ polynomials $F_1, \ldots, F_t$ in $n$ variables and degrees $\deg(F_i)$ to another system of $ft$ polynomials in $fn$ variables and degrees at most $\omega_p(F_i)$ with the same number of solutions. Then they applied Katz' theorem 3.4 to obtain,

**Theorem 3.5** (Moreno-Moreno, Theorem 1, [9])**.** *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables, $q = p^f$, $N$ be the number of common zeros of $F_1, \ldots, F_t$ in $(\mathbb{F}_q)^n$ and*

$$\mu = \left\lceil \left( \frac{n - \sum_{i=1}^{t} \omega_p(F_i)}{\max_{1 \le i \le t}\{\omega_p(F_i)\}} \right) f \right\rceil .$$

*If $\mu \geq 0$, then $p^\mu \mid N$.*

This theorem improves Katz' theorem 3.4 in many cases. Its proof uses the method of restriction of scalars introduced in Section 2.4.

Combining Katz' and Moreno-Moreno's theorem one obtains the following result which provides the best power of $p$ that divides the number the number of common zeros in all cases where both results apply.

**Theorem 3.6** (Moreno-Moreno, Theorem 0-1, [9])**.** *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables with coefficients in $\mathbb{F}_q$, $q = p^f$. Let $\omega_p(F_i)$ be the $p$-weight degree of $F_i$ and*

*$N$ the number of common zeros of the system. If*

$$\mu = \max\left(\left\lceil \frac{n - \sum_{i=1}^{t} \deg(F_i)}{\max_{1\leq i\leq t}\{\deg(F_i)\}} \right\rceil f, \left\lceil \left( \frac{n - \sum_{i=1}^{t} \omega_p(F_i)}{\max_{1\leq i\leq t}\{\omega_p(F_i)\}} \right) f \right\rceil\right)$$

*then $p^{\mu} \mid N$.*

In the following example, we present a polynomial for which the Moreno-Moreno theorem provides more useful information than Katz' theorem.

**Example 1.** Consider the polynomial

$$F(Y_1, \ldots, Y_5) = Y_1 Y_2 Y_3^2 + Y_1 Y_4 + Y_5^6$$

in $n = 5$ variables over $\mathbb{F}_3$ and $N$ be the number of zeros of $F(Y_1, \ldots, Y_5) = 0$. Here $\deg(F) = 6 > 5$ so Katz' result gives us

$$\mu_1 = \left\lceil \frac{5-6}{5} \right\rceil = 0$$

which implies that $1 \mid N$. Since $\omega_3(F) = 4 < 5$, Moreno-Moreno's theorem give us

$$\mu_2 = \left\lceil \frac{5-4}{4} \right\rceil = 1,$$

which implies that $3 \mid N$.

In this next example, we present a polynomial for which Moreno-Moreno's theorem improves the bound for the number of solutions obtained by Katz' theorem.

**Example 2.** Let

$$F(Y_1, \ldots, Y_5) = Y_1Y_2 + Y_3 + Y_1Y_4 + Y_5^4$$

be a polynomial in $n = 5$ variables over $\mathbb{F}_3$, and $N$ be the number of solutions of $F(Y_1, \ldots, Y_5) = 0$. Here $\deg(F) = 4 < 5$, so from Katz' theorem we get

$$\mu_1 = \left\lceil \frac{5-4}{4} \right\rceil = 1$$

which implies that $3 \mid N$.

Since $\omega_3(F) = 2 < 5$, from Moreno-Moreno's theorem we get

$$\mu_2 = \left\lceil \frac{5-2}{2} \right\rceil = 2$$

which implies that $3^2 \mid N$.

Finally, we consider a polynomial for which Katz' theorem gives a better bound for the number of solutions than Moreno-Moreno's theorem.

**Example 3.** Let

$$F(Y_1, \ldots, Y_5) = Y_1Y_2Y_3Y_4 + Y_1Y_4 + Y_5$$

be a polynomial in $n = 5$ variables over $\mathbb{F}_{3^2}$ and $N$ be the number of solutions of $F(Y_1, \ldots, Y_5) = 0$. Here we have that $\deg(F) = \omega_3(F) = 4 < 5$.

Since

$$\mu_1 = \left\lceil 2\left(\frac{5-4}{4}\right) \right\rceil = 1,$$

Moreno-Moreno's theorem implies that $3 \mid N$ for Moreno-Moreno's theorem but,

Katz' theorem gives

$$\mu_2 = \left\lceil \frac{5-4}{4} \right\rceil = 1,$$

which implies that $9 \mid N$.

## 3.3 Carlitz' generalized problem

In 1952, Carlitz studied a generalization of the problem that was tackled by Chevalley. He considered a system of polynomials $F_1, \ldots, F_t$ whose coefficients were not just elements of a finite field $\mathbb{F}_q$, instead, they are polynomials in $\mathbb{F}_q[X]$.

To illustrate the approach used by Carlitz to prove his theorem we consider the case of a quadratic equation of the form

$$F(Y_1, \ldots, Y_n) = \sum_{1}^{n} A_{ij}(X) Y_i Y_j + \sum_{1}^{n} A_i(X) Y_i = 0, \tag{3.1}$$

where $A_{ij}(X)$, $A_i(X) \in \mathbb{F}_q[X]$, $\deg(A_{ij}) \leq a$, and $\deg(A_i) \leq a$. Observe that $F \in \mathbb{F}_q[X][Y_1, \ldots, Y_n]$. We need polynomials $y_1, y_2, \ldots, y_n \in \mathbb{F}_q[X]$ where $\deg(y_i) \leq m$, for some fixed large enough $m$, such that $F(y_1, \ldots, y_n) = 0$ as a polynomial in the variable $X$.

Consider a generic solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X])^n$ for Equation (3.1) with

$$y_i = c_{im} X^m + c_{i(m-1)} X^{m-1} + \cdots + c_{i1} X + c_{i0} \tag{3.2}$$

and substitute the solution $(y_1, \ldots, y_n) \in (3.1)$ to get

$$F(y_1, \ldots, y_n) = \sum_1^n A_{ij}(X)(c_{im}X^m + \cdots + c_{i0})(c_{jm}X^m + \cdots + c_{j0})$$

$$+ \sum_1^n A_i(X)(c_{im}X^m + \cdots + c_{i0}) = 0. \tag{3.3}$$

We want $F(y_1, \ldots, y_n) = 0$ in $\mathbb{F}_q[X]$. That is, we want $F(y_1, \ldots, y_n)$ to be the polynomial 0 in the variable $X$. Note that $\deg_X(F(y_1, \ldots, y_n)) \leq a + 2m$. Grouping the similar terms in the variable $X$, we get.

$$F(y_1, \ldots, y_n) = P_{a+2m}(c_{1m}, \ldots, c_{nm})X^{a+2m} + \cdots + P_0(c_{10}, \ldots, c_{n0}) = 0. \tag{3.4}$$

This means that we want the coefficients $P_{a+2m}, \ldots, P_1, P_0$ of each monomial $X^{a+2m}, X^{a+2m-1}, \ldots, X, 1$ to be equal to 0. This produces a system

$$\begin{gathered} P_{a+2m}(c_{1m}, \ldots, c_{nm}) = 0 \\ \vdots \\ P_1(c_{10}, c_{11} \ldots, c_{n0}, c_{n1}) = 0 \\ P_0(c_{10}, \ldots, c_{n0}) = 0 \end{gathered} \tag{3.5}$$

of $a + 2m + 1$ equations in $n(m + 1)$ variables $c_{10}, \ldots, c_{1m}, \ldots, c_{n0}, \ldots, c_{nm}$ which are the unknown coefficients of $y_1, \ldots, y_n$. So $F(Y_1, \ldots, Y_n) = 0$ in (3.1) has a solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X])^n$ if and only if (3.5) has a solution in $(\mathbb{F}_q)^{n(m+1)}$. Note that none of the equations in (3.5) has a constant term, therefore (3.5) has the trivial

solution. Since the polynomials $y_i = c_{im}X^m + c_{i(m-1)}X^{m-1} + \cdots + c_{i1}X + c_{i0}$ were substituted in the monomials $Y_iY_j$, $Y_i$ of (3.5), the degree of each equation in (3.5) is at most 2. Chevalley's theorem 3.2 guarantees that the system has solutions if $n(m+1) > 2(a+2m+1)$. We now see that this condition is satisfied when $n \geq 5$ and $m \geq 2a-2$. Note that $m \geq 2a-2$ implies that $m+4 \geq 2a+2$, and hence

$$m + 5 > 2a + 2.$$

Since $n \geq 5$, we have $m + n \geq m + 5 > 2a + 2$. Also $n \geq 5$ implies that $n - 4 > 0$ and, therefore, $m(n-4) \geq m$. Thus $m(n-4) + n \geq m + n > 2a + 2$, and therefore,

$$mn - 4m + n > 2a + 2.$$

We have shown that $n(m+1) > 2(a+2m+1)$. Therefore a quadratic polynomial equation with polynomial coefficients in $n$ variables that has the trivial zero, will have a non-trivial zero when $n \geq 5$. We now show Carlitz' results for a single polynomial of degree $\deg(F)$.

**Theorem 3.7** (Carlitz, Theorem 1, [7])**.** *Let $F$ be a polynomial in $n$ variables, with coefficients in $\mathbb{F}_q[X]$, which has the trivial zero. If $n > \deg(F)^2$, then $F$ also has a non-trivial zero in $(\mathbb{F}_q[X])^n$. Moreover, there are polynomials of this form which have only the trivial zero if $n = [\deg(F)]^2$.*

*Proof.* Consider the case where $F(Y_1, \ldots, Y_n)$ is a polynomial with degree $\deg(F) = k$

and coefficients $A_{h_k}$ in $\mathbb{F}_q[X]$ with degree at most $a$.

$$F(Y_1, \ldots, Y_n) = \sum A_{h_k} \boldsymbol{Y}^{h_k} + \cdots + \sum A_{h_1} \boldsymbol{Y}^{h_1} = 0 \tag{3.6}$$

where $\boldsymbol{Y}^{h_l} = Y_1^{h_{l_1}} \ldots Y_n^{h_{l_n}}$, $h_l = (h_{l_1}, \ldots, h_{l_n})$ and $0 < \sum_{j=1}^{n} h_{l_j} = l \leq \deg(F) = k$.

This means we need polynomials $y_1, y_2, \ldots, y_n \in \mathbb{F}_q[X]$, with $\deg(y_i) \leq m$, for some fixed large enough $m$, such that $F(y_1, \ldots, y_n)$ is the polynomial 0 in the variable $X$.

We consider a generic solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X])^n$ with $y_i$ as in (3.2). Substituting the solution into $F$ and looking at $F(y_1, \ldots, y_n)$ as a polynomial in the variable $X$, we get that its degree is $\deg_X(F(y_1, \ldots, y_n)) \leq a + \deg(F)m$. Grouping the similar terms in the variable $X$ we get

$$F(y_1, \ldots, y_n) = P_{a+\deg(F)m}(c_{1m}, \ldots, c_{nm}) X^{a+\deg(F)m} + \cdots + P_0(c_{10}, \ldots, c_{n0}) = 0. \tag{3.7}$$

This means that we want the coefficients $P_{a+\deg(F)m}$, $P_{a+\deg(F)m-1}$, ..., $P_0$ of each term $X^{a+\deg(F)m}, X^{a+\deg(F)m-1}, \ldots, X, 1$ to be equal to 0. This provides a system

$$\begin{gathered} P_{a+\deg(F)m}(c_{1m}, \ldots, c_{nm}) = 0 \\ \vdots \\ P_1(c_{10}, c_{11}, \ldots, c_{n0}, c_{n1}) = 0 \\ P_0(c_{10}, \ldots, c_{n0}) = 0 \end{gathered} \tag{3.8}$$

of $a + \deg(F)m + 1$ equations in $n(m+1)$ variables $c_{10}, \ldots, c_{1m}, \ldots, c_{n0}, \ldots, c_{nm}$,

which are the unknown coefficients of $y_1, \ldots, y_n$. So $F(Y_1, \ldots, Y_n) = 0$ in (3.6) has a solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X])^n$ if and only if (3.8) has a solution in $(\mathbb{F}_q)^{n(m+1)}$. Note that none of the equations in (3.8) has a constant term, therefore (3.8) has the trivial solution. Since the polynomials $y_i = c_{im}X^m + c_{i(m-1)}X^{m-1} + \cdots + c_{i1}X + c_{i0}$ were substituted in the monomials of $F(Y_1, \ldots, Y_n)$, the degree of each equation in (3.8) is at most $\deg(F)$.

Chevalley's theorem 3.2 guarantees that the system has solutions if $n(m+1) > \deg(F)(a + \deg(F)m + 1)$. We now see that this condition is satisfied when $n \geq [\deg(F)]^2 + 1$ and $m \geq \deg(F)(a - \deg(F) + 1)$. Since $n \geq [\deg(F)]^2 + 1$, we get

$$m + n \geq m + [\deg(F)]^2 + 1 \geq a\deg(F) - [\deg(F)]^2 + \deg(F) + [\deg(F)]^2 + 1$$

and $m+n > a\deg(F)+\deg(F)$. Also $n \geq [\deg(F)]^2+1$ implies that $n-[\deg(F)]^2 > 0$ so $m(n - [\deg(F)]^2) \geq m$. Notice that

$$m\left(n - [\deg(F)]^2\right) + n \geq m + n > a\deg(F) + \deg(F),$$

therefore,

$$mn + n > a\deg(F) + [\deg(F)]^2 m + \deg(F).$$

Finally, we obtain that $n(m+1) > \deg(F)(a + \deg(F)m + 1)$. ■

The next result, extending Theorem 3.7 to systems of polynomial equations is not proven in [7]. We provide a complete proof for this result.

**Theorem 3.8** (Carlitz, [7]). *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables, with coefficients in $\mathbb{F}_q[X]$, which have the trivial zero. If $n > \sum_{i=1}^{t} \deg(F_i)^2$, then $F_1, \ldots, F_t$ also have a non-trivial zero in $(\mathbb{F}_q[X])^n$. Moreover, there are equations of this form which have only the trivial solution if $n = \sum_{i=1}^{t} \deg(F_i)^2$.*

*Proof.* As in the proof of Theorem 3.7, for each polynomial in the system, $F_1, \ldots F_t$, consider a polynomial equation with degree $\deg(F) = k$ and coefficients $A_{h_k}$ in $\mathbb{F}_q[X]$ with degree at most $a$.

$$F(Y_1, \ldots, Y_n) = \sum A_{h_k} \boldsymbol{Y}^{h_k} + \cdots + \sum A_{h_1} \boldsymbol{Y}^{h_1} = 0 \tag{3.9}$$

where $\boldsymbol{Y}^{h_l} = Y_1^{h_{l_1}} \ldots Y_n^{h_{l_n}}$, $h_l = (h_{l_1}, \ldots, h_{l_n})$, and $0 < \sum_{j=1}^{n} h_{l_j} = l \leq \deg(F) = k$

So we need polynomials $y_1, y_2, \ldots, y_n \in \mathbb{F}_q[X]$ where $\deg(y_j) \leq m$, for some fixed large enough $m$, such that for each F in the system $F_1, \ldots F_t$, $F(y_1, \ldots, y_n)$ is the polynomial 0 in the variable $X$.

For each $F_1, \ldots F_t$, we get a system

$$\begin{gathered} P_{a+\deg(F_i)m}(c_{1m}, \ldots, c_{nm}) = 0 \\ \vdots \\ P_1(c_{10}, c_{11}, \ldots, c_{n0}, c_{n1}) = 0 \\ P_0(c_{10}, \ldots, c_{n0}) = 0 \end{gathered} \tag{3.10}$$

of $a + \deg(F_i)m + 1$ equations in $n(m+1)$ variables $c_{10}, \ldots, c_{1m}, \ldots, c_{n0}, \ldots, c_{nm}$, which are the unknown coefficients of $y_1, \ldots, y_n$.

This gives a system of $\sum_{i=1}^{t}(a+\deg(F_i)m+1)$ equations in $n(m+1)$ variables. Each $F_i$ gives a block of $a+\deg(F_i)m+1$ equations of degree $\deg(F_i)$. Hence the sum of the degrees of all the equations is $\sum_{i=1}^{t}\deg(F_i)(a+\deg(F_i)m+1)$. Chevalley's theorem 3.2 guarantees a solution if

$$n(m+1) > \sum_{i=1}^{t}\deg(F_i)(a+\deg(F_i)m+1).$$

We now show that this condition is satisfied when $n \geq \sum_{i=1}^{t}[\deg(F_i)]^2+1$ and $m \geq \sum_{i=1}^{t}\deg(F_i)(a-\deg(F_i)+1)$. Since $n \geq \sum_{i=1}^{t}[\deg(F_i)]^2+1$, we get

$$m+n \geq m+\sum_{i=1}^{t}[\deg(F_i)]^2+1 \geq \sum_{i=1}^{t}a\deg(F_i)-\sum_{i=1}^{t}[\deg(F_i)]^2+\sum_{i=1}^{t}\deg(F_i)+\sum_{i=1}^{t}[\deg(F_i)]^2+1,$$

and $m+n > \sum_{i=1}^{t}a\deg(F_i)+\sum_{i=1}^{t}\deg(F_i)$. Also $n \geq \sum_{i=1}^{t}[\deg(F_i)]^2+1$ implies that $n-\sum_{i=1}^{t}[\deg(F_i)]^2 > 0$ so

$$m\left(n-\sum_{i=1}^{t}[\deg(F_i)]^2\right) \geq m.$$

Notice that

$$m\left(n-\sum_{i=1}^{t}[\deg(F_i)]^2\right)+n \geq m+n > \sum_{i=1}^{t}a\deg(F_i)+\sum_{i=1}^{t}\deg(F_i),$$

therefore,

$$mn + n > \sum_{i=1}^{t} a \deg(F_i) + \sum_{i=1}^{t} [\deg(F_i)]^2 m + \sum_{i=1}^{t} \deg(F_i).$$

We have shown that $n(m+1) > \sum_{i=1}^{t} \deg(F_i)(a + \deg(F_i)m + 1)$. Therefore, by Chevalley's theorem, $F_1 = F_2 = \cdots = F_t = 0$ has a non-trivial solution in $(\mathbb{F}_q[X])^n$.

■

Now we show a polynomial equation where $n = [\deg(F)]^2$ that only has the trivial solution proving that the bound in the theorem is tight.

**Example.** Let $q$ be odd and consider the equation

$$F(Y_1, \ldots, Y_4) = Y_1^2 - \alpha Y_2^2 + X^2(Y_3^2 - \alpha Y_4^2) = 0, \tag{3.11}$$

where $F \in (\mathbb{F}_q[X])[Y_1, \ldots, Y_4]$, $n = [\deg(F)]^2 = 4$ and $\alpha$ is a non-square in $\mathbb{F}_q$. Now suppose that $(p_1(X), p_2(X), p_3(X), p_4(X)) \in (\mathbb{F}_q[X])^4$ is a non-trivial solution to $F(Y_1, \ldots, Y_4) = 0$ where $\deg(p_1(X)) + \cdots + \deg(p_4(X))$ is minimal. Evaluating this solution yields

$$F(p_1(X), \ldots, p_4(X)) = p_1(X)^2 - \alpha p_2(X)^2 + X^2(p_3(X)^2 - \alpha p_4(X)^2) = 0. \tag{3.12}$$

We want $F(p_1(X), \ldots, p_4(X)) = 0$ as a polynomial in the variable $X$, thus each of the coefficients of the powers of $X$ has to be 0. Since each $p_i(X)$ is like (3.2), we get

the constant term of $F(p_1(X), \ldots, p_4(X))$ from

$$p_1(X)^2 - \alpha p_2(X)^2 = (c_{10} + c_{11}X + \cdots + c_{1m}X^m)^2$$

$$-\alpha(c_{20} + c_{21}X + \cdots + c_{2m}X^m)^2 = 0. \tag{3.13}$$

Note that, the constant term in (3.13), $c_{10}^2 - \alpha c_{20}^2 = 0$ implies that $\alpha = \left(c_{10}c_{20}^{-1}\right)^2$ which contradicts $\alpha$ being a non-square. Therefore $c_{10} = c_{20} = 0$ and Equation (3.13) can be written as

$$p_1(X)^2 - \alpha p_2(X)^2 = X^2(c_{11} + \cdots + c_{1m}X^{m-1})^2$$

$$-\alpha X^2(c_{21} + \cdots + c_{2m}X^{m-1})^2 = 0. \tag{3.14}$$

Hence $p_i(X) = X^2 p_i'(X)$ for $i = 1, 2$ and thus Equation (3.12) can be written as

$$F(p_1(X), \ldots, p_4(X)) = X^4(p_1'(X)^2 - \alpha p_2'(X)^2) + X^2(p_3(X)^2 - \alpha p_4(X)^2) = 0$$

and this holds if and only if

$$X^2(p_1'(X)^2 - \alpha p_2'(X)^2) + (p_3(X)^2 - \alpha p_4(X)^2) = 0.$$

Hence

$$F(p_3(X), p_4(X), p_1'(X), p_2'(X)) = p_3(X)^2 - \alpha p_4(X)^2 + X^2(p_1'(X)^2 - \alpha p_2'(X)^2) = 0,$$

which means that $(p_3(X), p_4(X), p_1'(X), p_2'(X))$ is a solution to the equation $F(Y_1, \ldots, Y_4) = 0$, but $\deg(p_1'(X)) < \deg(p_1(X))$ and $\deg(p_2'(X)) < \deg(p_2(X))$, which is a contradiction to the minimality condition. Therefore $F(Y_1, \ldots, Y_4)$ only has the trivial solution.

## 3.4 The Castro-Moreno-Rubio generalization and improvement

Using Carlitz' approach of Theorem 3.8 and the method of restriction of scalars presented in Section 2.4, Castro-Moreno-Rubio obtain a generalization of Moreno-Moreno's Theorem 3.5.

The approach consists of considering polynomial equations

$$F(Y_1, \ldots, Y_n) = \sum A_{h_k} \boldsymbol{Y}^{h_k} + \cdots + \sum A_{h_1} \boldsymbol{Y}^{h_1} = 0$$

of degree $\deg(F) = k$, with coefficients $A_{h_k} \in \mathbb{F}_q[X]$ of $\deg(A_{h_k}) \leq a$ associated to a system of polynomials $F_1, \ldots, F_t$. Then we consider and evaluate a generic solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X])^n$ with

$$y_j = c_{jm} X^m + c_{j(m-1)} X^{m-1} + \cdots + c_{j1} X + c_{j0} \tag{3.15}$$

to obtain a polynomial $F(y_1, \ldots, y_n)$ of degree $\deg(F)$ in the variables $Y_1, \ldots, Y_n$ and degree at most $a + \deg(F)m$ in the variable $X$ for each $F_i$ in the system. Since the

generic solution was evaluated into the variables $Y_1, \ldots, Y_n$ we get that its degree will still be $\deg(F)$ if we consider $F$ as polynomial of the unknown coefficients $c_{10}, \ldots, c_{nm}$. Continuing by using the method of restriction of scalars on $c_{10}, \ldots, c_{nm}$ and rewriting them as a combination of the basis elements $\mu_1, \ldots, \mu_f$ of $\mathbb{F}_{p^f}$ over $\mathbb{F}_p$,

$$c_{j\alpha} = a_1^{(j\alpha)}\mu_1 + a_2^{(j\alpha)}\mu_2 + \cdots + a_f^{(j\alpha)}\mu_f$$

where $a_k^{(j\alpha)}$ are new variables taking values in $\mathbb{F}_p$

This means that every monomial in the variables $c_{10}, \ldots, c_{nm}$ turns into a polynomial in the new variables $a_1^{(j\alpha)}, \ldots, a_f^{(j\alpha)}$ of degree $\omega_p(F)$. Thus turning the polynomial $F$ into a polynomial with coefficients of degrees at most $a + \deg(F)m$ in the variable $X$, which also depends on the basis elements $\mu_1, \ldots, \mu_f$ and variables $a_1^{(j\alpha)}, \ldots, a_f^{(j\alpha)}$.

Rewrite the new polynomial grouping up terms of the same degree in the variable $X$. Since we want this polynomial to be the 0 polynomial in $\mathbb{F}_p[X]$, then each of the coefficients of $X^0, \ldots, X^{a+\deg(F)m}$ has to be 0. Do the same for each polynomial $F_i$ of the system $F_1, \ldots, F_t$.

We get then that for each $F_i$ we have a system of $a + \deg(F_i)m + 1$ equations of degrees at most $\omega_p(F_i)$ in the $fn(m+1)$ variables $a_1^{(j\alpha)}, \ldots, a_f^{(j\alpha)}$.

So we get a system of $\sum_{i=1}^{t}(a + \deg(F_i)m + 1)$ equations which involve the basis elements $\mu_1, \ldots, \mu_f$ as well. Each of the equations in this system can be written as a linear combination of the basis elements, thus each of the $\sum_{i=1}^{t}(a + \deg(F_i)m + 1)$ equations gives $f$ more equations so we get a system of $(a + \deg(F_1)m + 1)f + \cdots + (a + \deg(F_t)m + 1)f$ equations in $fn(m+1)$ variables where the degrees of the $i$-th

block of equations is at most $\omega_p(F_i)$.

**Theorem 3.9** (Castro-Moreno-Rubio, Theorem 6, [3]). *Let $q = p^f$, $F_1, \ldots, F_t$ be polynomials in $n$ variables, with coefficients in $\mathbb{F}_q[X]$ that have degree at most $a$. If*

$$n > \sum_{i=1}^{t} \omega_p(F_i) \deg(F_i) \text{ and } m \geq \sum_{i=1}^{t} \omega_p(F_i)(a + 1 - \deg(F_i)),$$

*then the number of solutions $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X])^n$ of $F_1 = 0, \ldots, F_t = 0$ with $\deg(y_j) \leq m$ is divisible by $p^\mu$ where*

$$\mu = \left\lceil \left( \frac{n(m+1) - \sum_{i=1}^{t} \omega_p(F_i)(a + \deg(F_i)m + 1)}{\max_{1 \leq i \leq t}\{\omega_p(F_i)\}} \right) f \right\rceil.$$

*Proof.* As in the proofs of Theorems 3.7 and 3.8, for each polynomial in the system $F_1, \ldots, F_t$, consider a polynomial equation with degree $\deg(F) = k$ and coefficients $A_{h_k}$ in $\mathbb{F}_q[X]$ with degree at most $a$.

$$F(Y_1, \ldots, Y_n) = \sum A_{h_k} \boldsymbol{Y}^{h_k} + \cdots + \sum A_{h_1} \boldsymbol{Y}^{h_1} = 0 \tag{3.16}$$

where $\boldsymbol{Y}^{h_l} = Y_1^{h_{l_1}} \ldots Y_n^{h_{l_n}}$, $h_l = (h_{l_1}, \ldots, h_{l_n})$, and $0 < \sum_{j=1}^{n} h_{l_j} = l \leq \deg(F) = k$. We need polynomials $y_1, y_2, \ldots, y_n \in \mathbb{F}_q[X]$ with $\deg(y_j) \leq m$, for some fixed large enough $m$, such that $F_i(y_1, \ldots, y_n)$ is the polynomial 0 in the variable $X$. Consider a generic solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X])^n$ with $y_j$ as in (3.2). For each $F$ in the system $F_1, \ldots, F_t$, substitute the solution and look at $F(y_1, \ldots, y_n)$ as a polynomial in the variable $X$. The degree in the variable $X$ is $\deg_X(F(y_1, \ldots, y_n)) \leq a + \deg(F)m$. Then

$F(y_1, \ldots, y_n)$ is now a polynomial with $n(m+1)$ variables $c_{10}, \ldots, c_{1m}, \ldots, c_{n0}, \ldots, c_{nm}$ which are the coefficients of the $y_j$. Note that the degree of $F$ is also $\deg(F)$ when $F$ is seen as a polynomial in these new variables.

Using the method of restriction of scalars, detailed in Section 2.4, we represent each variable $c_{10}, \ldots, c_{1m}, \ldots, c_{n0}, \ldots, c_{nm}$ taking values over $\mathbb{F}_q$ as a linear combination of the $f$ basis elements of $\mathbb{F}_{p^f}$ over $\mathbb{F}_p$. Now each $c_{j\alpha}$ is rewritten as

$$c_{j\alpha} = a_1^{(j\alpha)}\mu_1 + a_2^{(j\alpha)}\mu_2 + \cdots + a_f^{(j\alpha)}\mu_f, \tag{3.17}$$

where for each $1 \leq j \leq n$ and $0 \leq \alpha \leq m$ we have that $a_1^{(j\alpha)}, \ldots, a_f^{(j\alpha)}$ are new variables that take values over $\mathbb{F}_p$ and $\mu_1, \ldots, \mu_f$ are the basis elements of $\mathbb{F}_q$ over $\mathbb{F}_p$. Each monomial of $F$ in the variables $c_{10}, \ldots, c_{nm}$ becomes a polynomial in the variables $a_1^{(j\alpha)}, \ldots, a_f^{(j\alpha)}$ which has degree $\omega_p(F)$.

Looking at $F(y_1, \ldots, y_n)$ after using the method of restriction of scalars and grouping the similar terms in the variable $X$ we get that Equation (3.16) is now like

$$F(y_1, \ldots, y_n) = P_{a+\deg(F)m}(a_1^{(1m)}, \ldots, a_f^{(nm)})X^{a+\deg(F)m} + \cdots$$

$$+P_1(a_1^{(10)}, a_1^{(11)}, \ldots, a_f^{(10)}, a_f^{(11)})X + P_0(a_1^{(10)}, \ldots, a_f^{(n0)}) = 0. \tag{3.18}$$

Where $F$ is a polynomial in $fn(m+1)$ variables

$$a_1^{(10)}, \ldots, a_f^{(1m)}, \ldots, a_1^{(n0)}, \ldots, a_f^{(nm)}$$

taking values over $\mathbb{F}_p$. We want this polynomial $F$ to be the 0 polynomial in $\mathbb{F}_p[X]$, for each $F$ in the system $F_1, \ldots, F_t$. This means that each coefficient of $X^0, \ldots, X^{a+deg(F)m}$ has to be 0. This provides a system

$$\begin{gathered} P_{a+\deg(F)m}(a_1^{(1m}, \ldots, a_f^{(nm)}) = 0 \\ \vdots \\ P_1 a_1^{(10)}, a_1^{(11)}, \ldots, a_f^{(10)}, a_f^{(11)}) = 0 \\ P_0(a_1^{(10)}, \ldots, a_f^{(n0)}) = 0 \end{gathered} \tag{3.19}$$

of $a+\deg(F)m+1$ equations in $fn(m+1)$ variables $a_1^{(10)}, \ldots, a_f^{(1m)}, \ldots, a_1^{(n0)}, \ldots, a_f^{(nm)}$. Note that the degree of the equations in these variables is $\omega_p(F)$ and each of these equations involves the basis elements $\mu_1, \ldots, \mu_f$ as well.

So each $F_i$ produces a system of $a + \deg(F_i)m + 1$ equations of degrees $\omega_p(F_i)$ in the variables $a_1^{(10)}, \ldots, a_f^{(1m)}, \ldots, a_1^{(n0)}, \ldots, a_f^{(nm)}$ and so we have a system of $\sum_{i=1}^{t}(a + \deg(F_i)m+1)$ equations. Combining the terms with the same basis elements we have that each equation in (3.19) can be written as

$$P_r(a_1^{(10)}, \ldots, a_f^{(nm)}, \mu_1, \ldots, \mu_f) = G_{1,r}(a_1^{(10)}, \ldots, a_f^{(nm)})\mu_1 + \cdots$$

$$+G_{f,r}(a_1^{(10)}, \ldots, a_f^{(nm)})\mu_f$$

for $0 \leq r \leq a + \deg(F)m$. This is, the equation $F(y_1, \ldots, y_n) = 0$ is written as a linear combination of the basis elements. Therefore $F = 0$ if and only if $G_{k,r} = 0$ for $1 \leq k \leq f$. We want that each coefficients of the basis elements to be equal to 0, this

provides a system

$$\begin{gathered} G_{f,a+\deg(F)m}\left(a_1^{(10)},\ldots,a_f^{(1m)},\ldots,a_1^{(n0)},\ldots,a_f^{(nm)}\right)=0 \\ \vdots \\ G_{1,a+deg(F)m}\left(a_1^{(10)},\ldots,a_f^{(1m)},\ldots,a_1^{(n0)},\ldots,a_f^{(nm)}\right)=0 \\ \vdots \\ G_{1,0}\left(a_1^{(10)},\ldots,a_f^{(1m)},\ldots,a_1^{(n0)},\ldots,a_f^{(nm)}\right)=0 \end{gathered} \tag{3.20}$$

of $(a+\deg(F)m+1)f$ equations, for each $F$ in the system $F_1,\ldots,F_t$. Each one of the $\sum_{i=1}^{t}(a+\deg(F_i)+1)$ equations gives $f$ more equations and so we end up with a system of $(a+\deg(F_1)m+1)f+\cdots+(a+\deg(F_t)m+1)f$ equations in $fn(m+1)$ variables where the degrees of the $i$-th block of equations is at most $\omega_p(F_i)$.

We have a system of $\sum_{i=1}^{t}(a+\deg(F_i)m+1)f$ equations in $fn(m+1)$ over $\mathbb{F}_p$. Proposition 2.25 gives us that the sum of the degrees of the polynomial in this system is at most $f\sum_{i=1}^{t}\omega_p(F_i)(a+\deg(F_i)m+1)$ and that the original system has the same number of solutions as this one. Katz' Theorem 3.4 says that $p^{\mu'}$ divides the number of common zeros of the system of polynomials if $\mu'\geq 0$ and

$$\mu'=\left\lceil\frac{fn(m+1)-\sum_\gamma\deg(G_\gamma)}{\max\{\deg(G_\gamma)\}}\right\rceil \tag{3.21}$$

where the $G_\gamma$ are the polynomials in the system of $\sum_{i=1}^{t}(a+\deg(F_i)m+1)f$ equations.

Since $\deg(G_\gamma) \leq \omega_p(F_i)$ for some $1 \leq i \leq t$, we have that

$$fn(m+1) - \sum_{\gamma} \deg(G_\gamma) \geq fn(m+1) - f\sum_{i=1}^{t} \omega_p(F_i)(a + \deg(F_i)m + 1).$$

Also, note that

$$\frac{1}{\max\{\deg(G_\gamma)\}} \geq \frac{1}{\max_{1\leq i\leq t}\{\omega_p(F_i)\}} \geq 0.$$

Therefore

$$\mu' = \left\lceil \frac{fn(m+1) - \sum_{\gamma} \deg(G_\gamma)}{\max\{\deg(G_\gamma)\}} \right\rceil \geq \left\lceil \frac{fn(m+1) - f\sum_{i=1}^{t} \omega_p(F_i)(a + \deg(F_i)m + 1)}{\max_{1\leq i\leq t}\{\omega_p(F_i)\}} \right\rceil = \mu.$$

We show now that $fn(m+1) - f\sum_{i=1}^{t} \omega_p(F_i)(a + \deg(F_i)m + 1) \geq 0$. When $n \geq \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i)+1$ and $m \geq \sum_{i=1}^{t} \omega_p(F_i)(a-\deg(F_i)+1)$. If $n \geq \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i)+1$, we get

$$m + n \geq m + \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i) + 1$$

$$\geq \sum_{i=1}^{t} a\omega_p(F_i) - \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i) + \sum_{i=1}^{t} \omega_p(F_i) + \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i) + 1$$

and $m+n > \sum_{i=1}^{t} a\omega_p(F_i) + \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i)$. Also $n \geq \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i)+1$ implies that $n - \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i) > 0$ so

$$m\left(n - \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i)\right) \geq m.$$

Notice that

$$m\left(n-\sum_{i=1}^{t}\omega_p(F_i)\deg(F_i)\right)+n\geq m+n>\sum_{i=1}^{t}a\omega_p(F_i)+\sum_{i=1}^{t}\omega_p(F_i).$$

Therefore

$$mn+n>\sum_{i=1}^{t}a\omega_p(F_i)+\sum_{i=1}^{t}\omega_p(F_i)\deg(F_i)m+\sum_{i=1}^{t}\omega_p(F_i).$$

We have shown that $n(m+1)>\sum_{i=1}^{t}\omega_p(F_i)(a+\deg(F_i)m+1)$, thus we get that the numerator of $\mu$,

$$\left(n(m+1)-\sum_{i=1}^{t}\omega_p(F_i)(a+\deg(F_i)m+1)\right)f>0$$

which means that

$$\mu\geq 0.$$

Finally, if $\mu\geq 0$, then $\mu'\geq 0$. Let $N$ be the number of solutions to the system of polynomials, because of Katz' Theorem we have that $p^{\mu'}\mid N$. Since $0\leq\mu\leq\mu'$, we get that $p^{\mu}\mid N$ where

$$\mu=\left\lceil\frac{fn(m+1)-f\sum_{i=1}^{t}\omega_p(F_i)(a+\deg(F_i)m+1)}{\max_{1\leq i\leq t}\{\omega_p(F_i)\}}\right\rceil.$$

■

By Chevalley's Theorem 3.2 we get that the system has solutions if $fn(m+1) > f\sum_{i=1}^{t}\omega_p(F_i)(a+\deg(F_i)m+1)$. Which was proved above. Therefore, as an immediate consequence to this Theorem 3.9, the following improvement to Carlitz' Theorem 3.8 is obtained.

**Corollary 3.9.1** (Castro-Moreno-Rubio, Corollary 8, [3]). *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables, with coefficients in $\mathbb{F}_q[X]$, such that they have the trivial zero. If*

$$n > \sum_{i=1}^{t}\omega_p(F_i)deg(F_i),$$

*then $F_1, \ldots, F_t$ also have a non-trivial zero in $(\mathbb{F}_q[X])^n$. Moreover if $n = \sum_{i=1}^{t}\omega_p(F_i)deg(F_i)$, there are systems of this form which only have the trivial zero.*

We now illustrate an example in order to show that the bound given for the number of variables is tight.

**Example.** Let $q = p^f$ be odd and $\alpha$ be a non-square in $\mathbb{F}_q$ and consider the polynomial

$$F(Y_1, \ldots, Y_n) = Y_1^{q+1} - \alpha Y_2^{q+1} + X^2(Y_3^{q+1} - \alpha Y_4^{q+1}) + \cdots + X^{2q}(Y_{n-1}^{q+1} - \alpha Y_n^{q+1})$$

with coefficients over $\mathbb{F}_q[X]$ and $n = \omega_p(F)\deg(F) = 2(q+1)$. Suppose that $(p_1(X), \ldots, p_n(X)) \in (\mathbb{F}_q[X])^n$ is a non-trivial solution to $F(Y_1, \ldots, Y_n) = 0$ where $\deg(p_1(X)) + \cdots + \deg(p_n(X))$ is minimal. Evaluating this solution yields

$$F(p_1(X), \ldots, p_n(X)) = p_1(X)^{q+1} - \alpha p_2(X)^{q+1} + X^2(p_3(X)^{q+1} - \alpha p_4(X)^{q+1}) +$$

$$\cdots + X^{2q}(p_{n-1}(X)^{q+1} - \alpha p_n(X)^{q+1}) = 0. \tag{3.22}$$

We want $F(p_1(X), \ldots, p_n(X)) = 0$ in the variable $X$, thus each of the coefficients of the powers of $X$ has to be 0. Since each $p_i(X)$ is like (3.15), we get the constant term of $F(p_1(X), \ldots, p_n(X))$ from

$$p_1(X)^{q+1} - \alpha p_2(X)^{q+1} = (c_{10} + c_{11}X + \cdots + c_{1m}X^m)^{q+1}$$

$$-\alpha(c_{20} + c_{21}X + \cdots + c_{2m}X^m)^{q+1} = 0. \tag{3.23}$$

Note that, the constant term in (3.23), $c_{10}^{q+1} - \alpha c_{20}^{q+1} = 0$ implies that $\alpha = \left(c_{10}c_{20}^{-1}\right)^{q+1} = \left(\left(c_{10}c_{20}^{-1}\right)^k\right)^2$ which contradicts $\alpha$ being a non-square. Therefore $c_{10} = c_{20} = 0$ and Equation (3.23) is

$$p_1(X)^{q+1} - \alpha p_2(X)^{q+1} = X^{q+1}(c_{11} + \cdots + c_{1m}X^m)^{q+1}$$

$$-\alpha X^{q+1}(c_{21} + \cdots + c_{2m}X^m)^{q+1} = 0. \tag{3.24}$$

Hence $p_i(X) = X^2 p_i'(X)$ for $i = 1, 2$ and thus Equation (3.22) can be written as

$$F(p_1(X), \ldots, p_n(X)) = X^{2(q+1)}(p_1'(X)^{q+1} - \alpha p_2'(X)^{q+1}) + X^2(p_3(X)^{q+1} - \alpha p_4(X)^{q+1}) +$$

$$\cdots + X^{2q}(p_{n-1}(X)^{q+1} - \alpha p_n(X)^{q+1}) = 0$$

and this holds if and only if

$$X^{2q}(p_1'(X)^{q+1} - \alpha p_2'(X)^{q+1}) + (p_3(X)^{q+1} - \alpha p_4(X)^{q+1}) + X^2(p_5(X)^{q+1} - \alpha p_6(X)^{q+1}) +$$

$$\cdots + X^{2q-2}(p_{n-1}(X)^{q+1} - \alpha p_n(X)^{q+1}) = 0.$$

Hence

$$F(p_3(X), \ldots, p_n(X), p_1'(X), p_2'(X)) = (p_3(X)^{q+1} - \alpha p_4(X)^{q+1}) + X^2(p_5(X)^{q+1} -$$

$$\alpha p_6(X)^{q+1}) + \cdots + X^{2q-2}(p_{n-1}(X)^{q+1} - \alpha p_n(X)^{q+1}) + X^{2q}(p_1'(X)^{q+1} - \alpha p_2'(X)^{q+1}) = 0,$$

which means that $(p_3(X), \ldots, p_n(X), p_1'(X), p_2'(X))$ is a solution to the equation $F(Y_1, \ldots, Y_n) = 0$, but $\deg(p_1'(X)) < \deg(p_1(X))$ and $\deg(p_2'(X)) < \deg(p_2(X))$ a contradiction to the minimality condition of $\deg(p_1(X)) + \cdots + \deg(p_n(X))$. Therefore $F(Y_1, \ldots, Y_n)$ only has the trivial solution.

# Chapter 4

# Systems of polynomials with multivariate polynomial coefficients

## 4.1 Carlitz' results for systems with multivariate polynomial coefficients

In [7] Carlitz provides a result that gives a bound on the number of variables needed so that systems of polynomials that have multivariate polynomial coefficients and the trivial zero have non-trivial zeros as well. The bound obtained for the numbers of variables needed is not tight.

In the proofs for Theorems 3.7, 3.8 and 3.9 we are able to count the number of equations in the system of polynomials by grouping the polynomial coefficients with respect to the powers of $X$. Now the polynomial coefficients are multivariate so, to extend the results to multivariate polynomial coefficients, we make use of the

following proposition:

**Lemma 4.1** (Carlitz, [7]). *If $F$ is a polynomial of degree $m$ in $w$ variables, then the amount of terms of $F$ is at most $\binom{m+w}{w}$.*

For the next proof, we will construct a monomial of degree at most $m$ in $w$ variables, and counting in how many ways this can be done.

*Proof.* Consider a list of $m+w$ spaces. We will fill the list of spaces with 1's and 0's. Choose $w$ spaces in the list to fill them with 0's and the remaining $m$ spaces with 1's. After the list has been filled out, we read how many 1's there are after each 0. Starting from the left, each 0 represents a new variable and the amount of 1's after the 0 is its corresponding exponent. Hence the amount of monomials, including the constant term, in a polynomial of degree $m$ of $w$ variables is at most, the number of ways in which one can choose the variables and their respective exponents, $\binom{m+w}{w}$. ■

We now illustrate this method of constructing the monomials for a multivariable polynomial.

**Example.** Let $F(Y_1, Y_2, Y_3)$ be a polynomial of degree 2. By Lemma 4.1 $F(Y_1, Y_2, Y_3)$ has at most $\binom{3+2}{2} = \binom{5}{2} = 10$ terms. Following the description in the proof, some of the vectors and their corresponding monomials are

$$(11000) \to Y_1^0 Y_2^0 Y_3^0 = 1$$

$$(10100) \to Y_1^1 Y_2^0 Y_3^0 = Y_1$$

$$(01100) \to Y_1^2 Y_2^0 Y_3^0 = Y_1^2$$

The following extension of Carlitz' theorem 3.8 can be found in [7], but no proof is given. This extension gives a bound to the number of variables needed so that systems of polynomial equations that have multivariate coefficients and the trivial zero have non-trivial solutions as well. We present a complete proof of the result.

**Theorem 4.2** (Carlitz, [7]). *Let $q = p^f$ and $F_1, \ldots, F_t$ be polynomials in $n$ variables, with coefficients in $\mathbb{F}_q[X_1, \ldots, X_w]$ which have the trivial zero. If*

$$n > \sum_{i=1}^{t} \deg(F_i)^{w+1},$$

*then $F_1, \ldots, F_t$ also has a non-trivial zero in $(\mathbb{F}_q[X_1, \ldots, X_w])^n$.*

*Proof.* As in the proofs of Theorems 3.7, 3.8 and 3.9, for each of the polynomials in the system $F_1, \ldots F_t$, consider a polynomial equation with degree $\deg(F)$ and coefficients $A_{h_k}$ in $\mathbb{F}_q[X_1, \ldots, X_w]$ with degree at most $a$.

$$F(Y_1, \ldots, Y_n) = \sum A_{h_k} \boldsymbol{Y}^{h_k} + \cdots + \sum A_{h_1} \boldsymbol{Y}^{h_1} = 0, \tag{4.1}$$

where $\boldsymbol{Y}^{h_l} = Y_1^{h_{l_1}} \ldots Y_n^{h_{l_n}}$, $h_l = (h_{l_1}, \ldots, h_{l_n})$ and $\sum_{j=1}^{n} h_{l_j} = l \leq \deg(F)$.

We need polynomials $y_1, y_2, \ldots, y_n \in \mathbb{F}_q[X_1, \ldots, X_w]$ with $\deg(y_j) \leq m$, for some fixed large enough $m$ such that for each $F$ in the system $F_1, \ldots, F_t$, $F(y_1, \ldots, y_n)$ is the polynomial 0 in the variables $X_1, \ldots, X_w$.

We consider a generic solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X_1, \ldots, X_w])^n$. By Lemma 4.1 we know that each $y_j = c_{j\binom{m+w}{w}-1}\boldsymbol{X}^{\binom{m+w}{w}-1} + \cdots + c_{j1}\boldsymbol{X} + c_{j0}$ has at most $\binom{m+w}{w}$ terms in the variables $X_1, \ldots, X_w$. Substituting the solution into $F$ and looking at $F(y_1, \ldots, y_n)$, as a polynomial in the variables $X_1, \ldots, X_w$ we get that its degree is $\deg_{X_1,\ldots,X_w}(F(y_1, \ldots, y_n)) \leq a + \deg(F)m$.

By Lemma (4.1), $F$ has at most $\binom{a+\deg(F)m+w}{w}$ terms. Now we want the polynomial coefficients of each term on $X_1, \ldots, X_w$ to be equal to 0. This provides a system of $\binom{a+\deg(F)m+w}{w}$ equations, for each $F$ in the system $F_1, \ldots, F_t$, in the $n\binom{m+w}{w}$ variables $c_{10}, \ldots, c_{n\binom{m+w}{w}-1}$ taking values from $\mathbb{F}_q$. Note that the degree of the new systems in the variables $c_{10}, \ldots, c_{n\binom{m+w}{w}-1}$ is still $\deg(F_i)$. Each $F(Y_1, \ldots, Y_n) = 0$ in Equation (4.1) has a solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X_1, \ldots, X_w])^n$ if and only if, all the $t$ systems of $\binom{a+\deg(F)m+w}{w}$ equations in $n\binom{m+w}{w}$ variables have a common solution in $(\mathbb{F}_q)^{n\binom{m+w}{w}}$. Note that the equations in these systems have the trivial solution.

Each $F_i$ gives a block of $\binom{a+\deg(F_i)m+w}{w}$ equations of degree $\deg(F_i)$ in the variables $c_{10}, \ldots, c_{n\binom{m+w}{w}-1}$. Hence the sum of the degrees of all the equations is $\sum_{i=1}^{t} \deg(F_i)\binom{a+\deg(F_i)m+w}{w}$. This gives a system of $\sum_{i=1}^{t}\binom{a+\deg(F_i)m+w}{w}$ equations in $n\binom{m+w}{w}$ variables. Applying Chevalley's theorem (3.2) we get that the

system has solutions if

$$n\binom{m+w}{w} > \sum_{i=1}^{t} \deg(F_i)\binom{a+\deg(F_i)m+w}{w}.$$

We now see that this condition is satisfied when $n \geq \sum_{i=1}^{t}[\deg(F_i)]^{w+1}+1$. We take $m > \sum_{i=1}^{t} \deg(F_i)H$ for some $H$. Since $n \geq \sum_{i=1}^{t}[\deg(F_i)]^{w+1}+1$ then $n-\sum_{i=1}^{t}[\deg(F_i)]^{w+1} \geq 1$ and

$$m(n - \sum_{i=1}^{t}[\deg(F_i)]^{w+1}) + n \geq m + n > \sum_{i=1}^{t} \deg(F_i)H + \sum_{i=1}^{t}[\deg(F_i)]^{w+1}.$$

Thus, we have

$$n(m+1) > \sum_{i=1}^{t} \deg(F_i)H + m\sum_{i=1}^{t}[\deg(F_i)]^{w+1} + \sum_{i=1}^{t}[\deg(F_i)]^{w+1}$$

$$= \sum_{i=1}^{t} \deg(F_i)\left(H + m[\deg(F_i)]^{w} + [\deg(F_i)]^{w}\right)$$

which implies that

$$n(m+1)\cdots(m+w) > \sum_{i=1}^{t} \deg(F_i)\left(H + m[\deg(F_i)]^{w} + [\deg(F_i)]^{w}\right)(m+2)\cdots(m+w).$$

Note that, by factoring out the powers of $m$ we get that

$$\sum_{i=1}^{t} \deg(F_i)\left(H + m[\deg(F_i)]^{w} + [\deg(F_i)]^{w}\right)(m+2)\cdots(m+w)$$

$$= m^w \left( \sum_{i=1}^{t} [\deg(F_i)]^{w+1} \right) + m^{w-1} \left( \sum_{i=1}^{t} \deg(F_i) H + \sum_{i=1}^{t} [\deg(F_i)]^{w+1} J_1 \right) + \cdots$$

$$+ \sum_{i=1}^{t} \deg(F_i) H K_w + \sum_{i=1}^{t} \deg(F_i)^{w+1} 2 \cdots w.$$

Since $H$ is chosen as large as necessary then

$$= m^w \left( \sum_{i=1}^{t} [\deg(F_i)]^{w+1} \right) + m^{w-1} \left( \sum_{i=1}^{t} \deg(F_i) H + \sum_{i=1}^{t} [\deg(F_i)]^{w+1} J_1 \right) + \cdots$$

$$+ \sum_{i=1}^{t} \deg(F_i) H K_w + \sum_{i=1}^{t} \deg(F_i)^{w+1} 2 \cdots w$$

$$> m^w \left( \sum_{i=1}^{t} [\deg(F_i)]^{w+1} \right) + m^{w-1} \left( \sum_{i=1}^{t} \deg(F_i) a + \sum_{i=1}^{t} [\deg(F_i)]^{w+1} J_1' \right) + \cdots$$

$$+ \sum_{i=1}^{t} \deg(F_i) a + \sum_{i=1}^{t} \deg(F_i) w!$$

$$= \sum_{i=1}^{t} \deg(F_i)(a + \deg(F_i)m + w) \cdots (a + \deg(F_i)m + 1).$$

Therefore

$$n(m+1) \cdots (m+w) > \sum_{i=1}^{t} \deg(F_i)(a + \deg(F_i)m + w) \cdots (a + \deg(F_i)m + 1).$$

Hence

$$n \frac{(m+w)!}{m!w!} > \sum_{i=1}^{t} \deg(F_i) \frac{(a + \deg(F_i)m + w)!}{(a + \deg(F_i)m)!w!}.$$

We have shown that

$$n\binom{m+w}{w} > \sum_{i=1}^{t} \deg(F_i)\binom{a+\deg(F_i)m+w}{w}$$

and therefore, by Chevalley's theorem, $F_1 = F_2 = \cdots = F_t = 0$ has a non-trivial solution in $(\mathbb{F}_q[X_1, \ldots, X_w])^n$.

■

## 4.2 Improving Carlitz' result and generalizing Castro-Moreno-Rubio

Combining Castro-Moreno-Rubio's and Carlitz approaches in (3.9) and (4.2) respectively, we obtain the following generalization of Theorem 3.9.

**Theorem 4.3.** *Let $q = p^f$ and $F_1, \ldots, F_t$ be polynomials in the variables $Y_1, \ldots, Y_n$, with coefficients in $\mathbb{F}_q[X_1, \ldots, X_w]$ that have degree at most $a$. For a large enough $m$, if $n > \sum_{i=1}^{t} \omega_p(F_i) \deg(F_i)^w$, then the number of $n$-tuples $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X_1, \ldots, X_w])^n$ with degree at most $m$ that are common zeros to the system of polynomials $F_1, \ldots, F_t$ is divisible by $p^\mu$ where*

$$\mu = \left\lceil \left( \frac{n\binom{m+w}{w} - \sum_{i=1}^{t} \omega_p(F_i)\binom{a+\deg(F_i)m+w}{w}}{\max_{1\le i\le t}\{\omega_p(F_i)\}} \right) f \right\rceil.$$

*Proof.* As in the proofs of Theorems 3.7, 3.8, 3.9 and 4.2, for each polynomial in the system $F_1, \ldots F_t$, consider a polynomial equation with degree $\deg(F)$ and coefficients

$A_{h_k}$ in $\mathbb{F}_q[X_1, \ldots, X_w]$ with degree at most $a$.

$$F(Y_1, \ldots, Y_n) = \sum A_{h_k} \boldsymbol{Y}^{h_k} + \cdots + \sum A_{h_1} \boldsymbol{Y}^{h_1} = 0, \tag{4.2}$$

where $h_l = (h_{l_1}, \ldots, h_{l_n})$, $\sum_{j=1}^{n} h_{l_j} = l \leq \deg(F)$ and $\boldsymbol{Y}^{h_l} = Y_1^{h_{l_1}} \ldots Y_n^{h_{l_n}}$.

We need polynomials $y_1, y_2, \ldots, y_n \in \mathbb{F}_q[X_1, \ldots, X_w]$ with $\deg(y_j) \leq m$, for some fixed large enough $m$ such that $F_i(y_1, \ldots, y_n)$ is the polynomial 0 in the variables $X_1, \ldots, X_w$.

We consider a generic solution $(y_1, \ldots, y_n) \in (\mathbb{F}_q[X_1, \ldots, X_w])^n$. By Lemma 4.1 we know that each $y_j = c_{j\binom{m+w}{w}-1)} \boldsymbol{X}^{\binom{m+w}{w}-1} + \cdots + c_{j1}\boldsymbol{X} + c_{j0}$ has at most $\binom{m+w}{w}$ terms in the variables $X_1, \ldots, X_w$. Substituting the solution into each $F$ in the system $F_1, \ldots, F_t$ and looking at $F(y_1, \ldots, y_n)$ as a polynomial in the variables $X_1, \ldots, X_w$ we get that its degree is $\deg_{X_1,\ldots,X_w}(F(y_1, \ldots, y_n)) \leq a + \deg(F)m$. Observe that $F(y_1, \ldots, y_n)$ still has degree $\deg(F)$ when looked as a polynomial in the variables $c_{10}, \ldots, c_{1\binom{m+w}{w}-1)}, \ldots, c_{n0}, \ldots, c_{n\binom{m+w}{w}-1)}$.

Using now the method of restriction of scalars, detailed in Section 2.4, we represent each $c_{10}, \ldots, c_{1\binom{m+w}{w}-1)}, \ldots, c_{n0}, \ldots, c_{n\binom{m+w}{w}-1)}$ as a linear combination of the $f$ basis elements of $\mathbb{F}_q$ over $\mathbb{F}_p$. Now each $c_{j\alpha}$ is rewritten as

$$c_{j\alpha} = a_1^{(j\alpha)}\mu_1 + a_2^{(j\alpha)}\mu_2 + \cdots + a_f^{(j\alpha)}\mu_f, \tag{4.3}$$

where for each $1 \leq j \leq n$ and $0 \leq \alpha \leq \binom{m+w}{w} - 1$ we have new variables $a_1^{(j\alpha)}, \ldots, a_f^{(j\alpha)}$ are our new variable taking values over $\mathbb{F}_p$, and $\mu_1, \ldots, \mu_f$ are the basis elements of $\mathbb{F}_q$

over $\mathbb{F}_p$. Each monomial of $F$ in the variables $c_{10}, \ldots, c_{1\left(\binom{m+w}{w}-1\right)}, \ldots, c_{n0}, \ldots, c_{n\left(\binom{m+w}{w}-1\right)}$ becomes a polynomial in the variables $a_1^{(j\alpha)}, \ldots, a_f^{(j\alpha)}$ which has degree at most $\omega_p(F)$. So $F(y_1, \ldots, Y_n)$ is a polynomial in the variables $X_1, \ldots, X_w, a_1^{10}, \ldots, a_f^{n\left(\binom{m+w}{w}-1\right)}$ and the basis elements $\mu_1, \ldots, \mu_f$ with coefficients in $\mathbb{F}_p$.

$F(y_1, \ldots, y_n)$ as a polynomial in the variables $X_1, \ldots, X_w$, it has degree at most $a + \deg(F)m$. So, by Lemma 4.1, it has at most $\binom{a+\deg(F)m+w}{w}$ terms in $X_1, \ldots, X_w$. Now we want to group the terms for each monomial $X_1^{e_1} \cdots X_w^{e_w}$ and the polynomial coefficients of each monomial in these variables to be equal to 0. This provides a system

$$\begin{aligned} P_{\binom{a+\deg(F)m+w}{w}-1}\left(a_1^{(10)}, \ldots, a_f^{\left(1\left(\binom{m+w}{w}-1\right)\right)}, \ldots, a_1^{(n0)}, \ldots, a_f^{\left(n\left(\binom{m+w}{w}-1\right)\right)}\right) &= 0 \\ \vdots & \\ P_1\left(a_1^{(10)}, \ldots, a_f^{\left(1\left(\binom{m+w}{w}-1\right)\right)}, \ldots, a_1^{(n0)}, \ldots, a_f^{\left(n\left(\binom{m+w}{w}-1\right)\right)}\right) &= 0 \\ P_0\left(a_1^{(10)}, \ldots, a_f^{\left(1\left(\binom{m+w}{w}-1\right)\right)}, \ldots, a_1^{(n0)}, \ldots, a_f^{\left(n\left(\binom{m+w}{w}-1\right)\right)}\right) &= 0 \end{aligned} \tag{4.4}$$

of $\binom{a+\deg(F)m+w}{w}$ equations in $fn\binom{m+w}{w}$ variables

$$a_1^{(10)}, \ldots, a_f^{\left(1\left(\binom{m+w}{w}-1\right)\right)}, \ldots, a_1^{(n0)}, \ldots, a_f^{\left(n\left(\binom{m+w}{w}-1\right)\right)}$$

which are the unknown coefficients of $y_1, \ldots, y_n$ after being rewritten using the method of restriction of scalars. Note that the degree of the equations in the system is at most $\omega_p(F)$ and each of these equations involves the basis elements $\mu_1, \ldots, \mu_f$ as well.

So each $F_i$ produces a system of $\binom{a+\deg(F_i)m+w}{w}$ equations of degrees at most $\omega_p(F_i)$

in the variables $a_1^{(10)}, \ldots, a_f^{\binom{n\left(\binom{m+w}{w}-1\right)}{}}$ and so we have a system of

$$\sum_{i=1}^{t} \binom{a + \deg(F_i)m + w}{w}$$

equations. In each of these equations, we combine the terms with the same basis elements to write the polynomials $P_0, \ldots, P_{\binom{a+deg(F)m+w}{w}-1}$ in the System 4.4 as linear combinations of the basis elements. Each of the equations in 4.4 is equal to 0 if and only if the coefficients of $\mu_1, \ldots, \mu_f$ are 0. We obtain a a system

$$\begin{aligned}
G_{f,\binom{a+deg(F)m}{w}-1}\left(a_1^{(10)}, \ldots, a_f^{\binom{1\left(\binom{m+w}{w}-1\right)}{}}, \ldots, a_1^{(n0)}, \ldots, a_f^{\binom{n\left(\binom{m+w}{w}-1\right)}{}}\right) &= 0 \\
&\vdots \\
G_{1,\binom{a+deg(F)m}{w}-1}\left(a_1^{(10)}, \ldots, a_f^{\binom{1\left(\binom{m+w}{w}-1\right)}{}}, \ldots, a_1^{(n0)}, \ldots, a_f^{\binom{n\left(\binom{m+w}{w}-1\right)}{}}\right) &= 0 \qquad (4.5) \\
&\vdots \\
G_{1,0}\left(a_1^{(10)}, \ldots, a_f^{\binom{1\left(\binom{m+w}{w}-1\right)}{}}, \ldots, a_1^{(n0)}, \ldots, a_f^{\binom{n\left(\binom{m+w}{w}-1\right)}{}}\right) &= 0
\end{aligned}$$

of $\binom{a+\deg(F)m+w}{w} f$ equations for each $F$ in the system $F_1, \ldots, F_t$. We end up with a system of $\binom{a+\deg(F_1)m+w}{w} f + \cdots + \binom{a+\deg(F_t)m+w}{w} f$ equations in $fn\binom{m+w}{w}$ variables over $\mathbb{F}_p$ where the degrees of the $i$-th block of equations is at most $\omega_p(F_i)$.

Note that Proposition 2.25 gives us that the sum of the degrees of the polynomial in this system is at most $f \sum_{i=1}^{t} \omega_p(F_i) \binom{a + \deg(F_i)m + w}{w}$ and that the original system has the same number of solutions as this one. Katz' theorem 3.4 says that $p^{\mu'}$ divides

$N$, the number of common zeros of the system of polynomials, if $\mu' \geq 0$ and

$$\mu' = \left\lceil \frac{fn\binom{m+w}{w} - \sum_{\gamma} \deg(H_\gamma)}{\max\{\deg(H_\gamma)\}} \right\rceil \tag{4.6}$$

where the $H_\gamma$ are the polynomials in the system of $\sum_{i=1}^{t} \binom{a + \deg(F_i)m + w}{w} f$ equations. Since $\deg(H_\gamma) \leq \omega_p(F_i)$ for some $1 \leq i \leq t$, we have that

$$fn\binom{m+w}{w} - \sum_{\gamma} \deg(H_\gamma) \geq fn\binom{m+w}{w} - f\sum_{i=1}^{t} \omega_p(F_i)\binom{a + \deg(F_i)m + w}{w}.$$

Also, note that

$$\frac{1}{\max\{\deg(H_\gamma)\}} \geq \frac{1}{\max_{1\leq i\leq t}\{\omega_p(F_i)\}} \geq 0.$$

Therefore

$$\mu' = \left\lceil \frac{fn\binom{m+w}{w} - \sum_{\gamma} \deg(H_\gamma)}{\max\{\deg(H_\gamma)\}} \right\rceil \geq \left\lceil \frac{fn\binom{m+w}{w} - f\sum_{i=1}^{t} \omega_p(F_i)\binom{a + \deg(F_i)m + w}{w}}{\max_{1\leq i\leq t}\{\omega_p(F_i)\}} \right\rceil = \mu.$$

We show now that $fn\binom{m+w}{w} - f\sum_{i=1}^{t} \omega_p(F_i)\binom{a + \deg(F_i)m + w}{w} \geq 0$ if $n \geq \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w + 1$. We take $m > \sum_{i=1}^{t} \omega_p(F_i)H$. Since $n \geq \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w + 1$ then $n - \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w \geq 1$ and

$$m(n - \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w) + n \geq m + n > \sum_{i=1}^{t} \omega_p(F_i)H + \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w.$$

Thus, we have that

$$n(m+1) > \sum_{i=1}^{t} \omega_p(F_i)H + m\sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w + \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w$$

$$= \sum_{i=1}^{t} \omega_p(F_i)\left(H + m[\deg(F_i)]^w + [\deg(F_i)]^w\right),$$

which implies that

$$n(m+1)\cdots(m+w) > \sum_{i=1}^{t} \omega_p(F_i)\left(H + m[\deg(F_i)]^w + [\deg(F_i)]^w\right)(m+2)\cdots(m+w).$$

Note that, by factoring out the powers of $m$ we get that

$$\sum_{i=1}^{t} \omega_p(F_i)\left(H + m[\deg(F_i)]^w + [\deg(F_i)]^w\right)(m+2)\cdots(m+w)$$

$$= m^w\left(\sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w\right) + m^{w-1}\left(\sum_{i=1}^{t} \omega_p(F_i)H + \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w J_1\right) + \cdots$$

$$+ \sum_{i=1}^{t} \omega_p(F_i)HK_w + \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i)^w 2\cdots w$$

and since $H$ is chosen as large as necessary then

$$m^w\left(\sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w\right) + m^{w-1}\left(\sum_{i=1}^{t} \omega_p(F_i)H + \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w J_1\right) + \cdots$$

$$+ \sum_{i=1}^{t} \omega_p(F_i)HK_w + \sum_{i=1}^{t} \omega_p(F_i)\deg(F_i)^w 2\cdots w$$

$$> m^w \left( \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w \right) + m^{w-1} \left( \sum_{i=1}^{t} \omega_p(F_i)a + \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w J_1' \right) + \cdots$$

$$+ \sum_{i=1}^{t} \omega_p(F_i)a + \sum_{i=1}^{t} \omega_p(F_i)w!$$

$$= \sum_{i=1}^{t} \omega_p(F_i)(a + \deg(F_i)m + w) \cdots (a + \deg(F_i)m + 1).$$

Therefore

$$n(m+1)\cdots(m+w) > \sum_{i=1}^{t} \omega_p(F_i)(a + \deg(F_i)m + w) \cdots (a + \deg(F_i)m + 1).$$

Hence

$$n\frac{(m+w)!}{m!w!} > \sum_{i=1}^{t} \omega_p(F_i)\frac{(a + \deg(F_i)m + w)!}{(a + \deg(F_i)m)!w!}.$$

Therefore, we have that

$$n\binom{m+w}{w} > \sum_{i=1}^{t} \omega_p(F_i)\binom{a + \deg(F_i)m + w}{w}$$

and

$$\left( n\binom{m+w}{w} - \sum_{i=1}^{t} \omega_p(F_i)\binom{a + \deg(F_i)m + w}{w} \right) f > 0$$

which means that $\mu \geq 0$.

Finally, since $\mu \geq 0$, then $\mu' \geq 0$. By Katz' theorem we have that $p^{\mu'} \mid N$. Since

$0 \leq \mu \leq \mu'$, we get that $p^\mu \mid N$ where

$$\mu = \left\lceil \frac{fn\binom{m+w}{w} - f\sum_{i=1}^{t} \omega_p(F_i)\binom{a+\deg(F_i)m+w}{w}}{\max_{1\leq i\leq t}\{\omega_p(F_i)\}} \right\rceil .$$

■

Applying Chevalley's theorem 3.2 we get that the system has solutions in $(\mathbb{F}_q[X_1, \ldots, X_W])^n$ if

$$fn\binom{m+w}{w} > \sum_{i=1}^{t} f\omega_p(F_i)\binom{a+\deg(F_i)m+w}{w}.$$

We saw in the proof of Theorem 4.3 that this condition is satisfied when $n \geq \sum_{i=1}^{t} \omega_p(F_i)[\deg(F_i)]^w + 1$. An immediate consequence of Theorem 4.3 is an improvement to Carlitz' theorem 4.2.

**Corollary 4.3.1.** *Let $q = p^f$ and $F_1, \ldots, F_t$ be polynomials in $n$ variables, with coefficients in $\mathbb{F}_q[X_1, \ldots, X_w]$ that have the trivial zero. If*

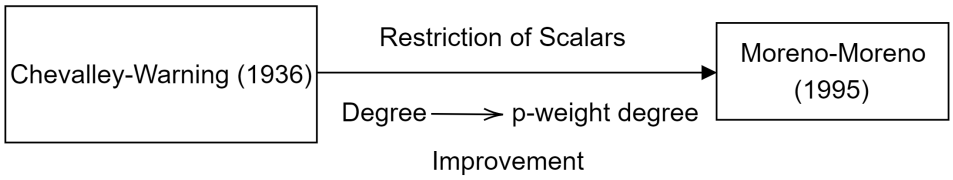
$$n > \sum_{i=1}^{t} \omega_p(F_i) \deg(F_i)^w,$$

*then the system also has a non-trivial zero in $(\mathbb{F}_q[X_1, \ldots, X_w])^n$.*

# Chapter 5

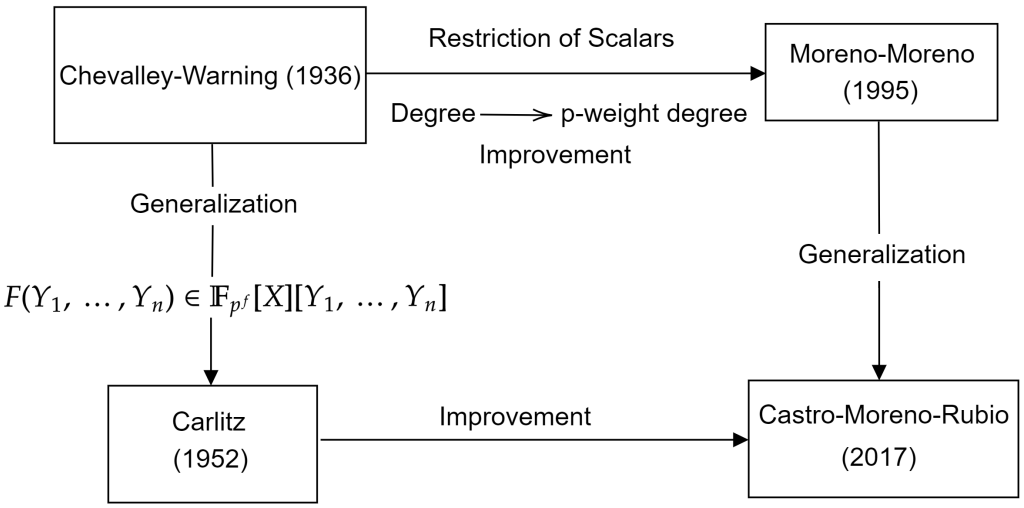
# Conclusions

This work set out to generalize the results of Castro-Moreno-Rubio in [3] and, in doing so, improve Carlitz' result in [7]. We began from the results of Chevalley-Warning with Theorems 3.2 and 3.3 and saw an improvement provided by Katz in Theorem 3.4. Using the method of restriction of scalars, detailed in Section 2.4, we presented Moreno-Moreno's results in Theorem 3.5. All of these results considered systems of polynomials with coefficients that are elements of the finite field $\mathbb{F}_q$.

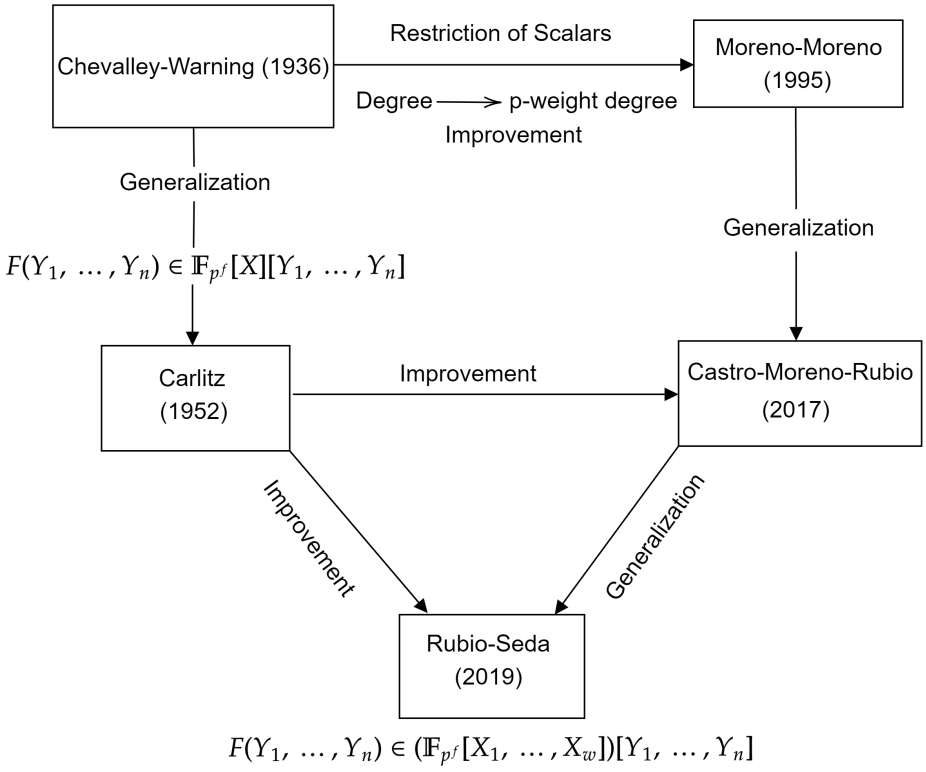



Then, we went over Carlitz' Theorems 3.7 and 3.8 where the coefficients are now polynomials. Combining the methods of Moreno-Moreno and Carlitz we then discussed the results of Castro-Moreno-Rubio in Theorem 3.9 and Corollary 3.9.1. In these results Castro-Moreno-Rubio generalized Moreno-Moreno's result and consequently

obtained an improvement for Carlitz' result.




The next step was to generalize the results of Castro-Moreno-Rubio to multivariate polynomials as coefficients. Combining the methods used in the proof for Theorem 3.9 and Lemma 4.1 we obtain this generalization in Theorem 4.3 and consequently obtain an improvement for Carlitz' Theorem 4.2 in Corollary 4.3.1.


Chevalley-Warning (1936)
Restriction of Scalars
Moreno-Moreno (1995)
Degree ⟶ p-weight degree
Improvement
Generalization
$F(Y_1, \ldots, Y_n) \in \mathbb{F}_{p^f}[X][Y_1, \ldots, Y_n]$
Generalization
Carlitz (1952)
Improvement
Castro-Moreno-Rubio (2017)
Improvement
Generalization
Rubio-Seda (2019)
$F(Y_1, \ldots, Y_n) \in (\mathbb{F}_{p^f}[X_1, \ldots, X_w])[Y_1, \ldots, Y_n]$

# Bibliography


[1] Chevalley C. Démonstration d'une hypothese de M. Artin. In *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, volume 11, pages 73–75. Springer, 1935.

[2] Foote R. M. Dummit, D.S. Abstract algebra. John Wile & Sons. *Inc., Hoboken, NJ*, 2004.

[3] Moreno I. Rubio F. Castro, O. An improvement of a theorem of Carlitz. *Journal of Pure and Applied Algebra*, **224**:106246, 2020.

[4] C. Mummert GL., Mullen. Finite fields and applications (Student mathematical library volume 41). *Amer Mathematical Society*, 2007.

[5] Ax J. Zeroes of polynomials over finite fields. *American Journal of Mathematics*, 86(2):255–261, 1964.

[6] M. Rosen K, Ireland. *A classical introduction to modern number theory*, volume 84. Springer Science & Business Media, 2013.

[7] Carlitz L. et al. Some applications of a theorem of Chevalley. *Duke Mathematical Journal*, 18(4):811–819, 1951.

[8] Katz NM. On a theorem of Ax. *American Journal of Mathematics*, 93(2):485–499, 1971.

[9] Moreno O. and CJ. Moreno. Improvements of the Chevalley-Warning and the Ax-Katz theorems. *American Journal of Mathematics*, 117(1):241–244, 1995.

[10] Lang S. *Introduction to linear algebra*. Springer Science & Business Media, 2012.

[11] E. Warning. Bemerkung zur vorstehenden Arbeit von Herrn Chevalley. *Abh. Math. Semin. Univ. Hamb.*, 11:76–83, 1935.