CHAPTER II

ABSOLUTELY IRREDUCIBLE EQUATIONS $y^d=f(x)$ AND $y^q-y=f(x)$

2.1 Absolutely Irreducible Equations $y^d=f(x)$:

Let F_q be a finite field with q elements. Then the characteristic of F_q is p for some prime p and F_q contains F_p, the integers modulo p. We also have $q=p^k$, where $k=[F_q:F_p]$ and F_q is the splitting field of X^q-X over F_p.

Let f(X) be a polynomial with coefficients in F_q. In the following sections we prove the following

Theorem 2.1: Suppose $Y^d-f(X)$ is absolutely irreducible, (i.e. irreducible over F_q and every algebraic extension of F_q), and that $q > 100\ dm^2$, where m=deg f. If N is the number of zeros of the polynomial in F_q^2, then

$$\left| N - q \right| \leq 4d^{3/2}\, m\, q^{1/2} .$$

Lemma 2.1: Let C be a cyclic group of order h. For any integer $d>0$, let C^d be the subgroup of d^{th} powers of elements of C. Let $d'=(h,d)$, then $C^d = C^{d'}$ and consists precisely of those $x \varepsilon C$ with

$$x^{h/d'} = 1 . \tag{2.1}$$

For any $x \varepsilon C^d$, there are exactly d' elements $y \varepsilon C$ with $y^d=x$.

Proof: Write $C = \left\{ g , g^2, \ldots, g^h = 1 \right\}$. Suppose $x \varepsilon C^d$, then

$x = g^{id}$ for some i. Since $d' \mid d$,

$$x^{h/d'} = (g^{id/d' \, h}) = 1$$

Conversly, suppose $x^{h/d'} = 1$. Let $x = g^i$, then $g^{ih/d'} = 1$. Since order of $g^i = \frac{h}{(i,h)}$, we have $\frac{h}{(i,h)} \mid \frac{h}{d'} \Rightarrow d' \mid i$, i.e. $\frac{i}{d'}$ is an integer, say, $i = d'i_o$. We must show there is a $y \in C$, with $y^d = x$. If $y = g^j$, we need

$$g^{jd} = x = g^{i_o d'}$$

or

$$jd \equiv i_o d' \pmod{h} .$$

This congruence has a solution j since $(d,h)=d'$ divides $i_o d'$. Moreover, the number of solutions j (mod h) equals $(d,h)=d'$. Since (2.1) depends on d', we have $C^d = C^{d'}$. Hence the lemma.

Given an equation $y^d = f(x)$ in F_q. Let $N=N(d)$ be the number of solutions $(x,y) \in F_q \times F_q$ of $y^d = f(x)$. Let N_o be the number of solutions with $y = 0$; then N_o is the number of $x \in F_q$ with $f(x) = 0$.

For each solution (x,y) with $y \neq 0$, we have $f(x) \in (F_q^*)^d$, where F_q^* is $F_q - \{0\}$. Since F_q^* is a finite cyclic group of order $q-1$, by Lemma 2.1, we have

$$f(x)^{\frac{q-1}{d'}} = 1, \text{ where } d' = (q-1, d).$$

Let N_1 be the number of $x \in F_q$ with $f(x)^{\frac{q-1}{d'}} = 1$.

For such an x, there are d' elements y with $y^d = f(x)$. Hence $N = N_o + d' N_1$. This expression depends only on d', so $N(d) = N(d')$. Thus without loss of generality we may assume that $d \mid q-1$; then

$$N = N_o + dN_1 ,$$

where N_1 is the number of $x \in F_q$ with $f(x)^{\frac{q-1}{d}} = 1$.

Lemma 2.2: Let N be the number of solutions $(x,y) \in F_q \times F_q$ of $y^d = f(x)$, where $d \mid q-1$. Then $N = N_o + dN_1$, where N_o is the number of $x \in F_q$ with $f(x) = 0$, and N_1 is the number of $x \in F_q$ with $f(x)^{\frac{q-1}{d}} = 1$. Further, $N_o + N_1 + N_2 = q$, where N_2 is the number of x satisfying

$$(2.2) \quad (f(x)^{\frac{q-1}{d}})^{d-1} + (f(x)^{\frac{q-1}{d}})^{d-2} + \ldots + f(x)^{\frac{q-1}{d}} + 1 = 0.$$

Proof: We only have to show that $N_o + N_1 + N_2 = q$.

We have

$$Z^q - Z = Z\,(Z^{\frac{q-1}{d}} - 1)\,(Z^{\frac{(q-1)(d-1)}{d}} + Z^{\frac{(q-1)(d-2)}{d}} + \ldots\ldots + Z^{\frac{q-1}{d}} + 1).$$

Since every $z \in F_q$ satisfies, $z^q - z = 0$, and $Z^q - Z$ is a separable polynomial, every element of F_q is a root of one and only one of the factors of $Z^q - Z$, hence

$$q = N_o + N_1 + N_2 .$$

When m=1, then $|N-q| = |N_o + dN_1 - N_o - N_1 - N_2|$

$$= |(d-1)\, N_1 - N_2|$$

$$= \left|(d-1)\,\left(\frac{q-1}{d}\right) - (d-1)\,\frac{(q-1)}{d}\right| = 0.$$

and when d = 1, we have

$$|N - q| = |N_o + dN_1 - N_o - N_1 - N_2|$$

$$= |N_2| = 0.$$

Hence in order to prove Theorem 2.1, we may suppose that

(2.3) $m > 1$, $d > 1$ and $d \mid (q-1)$.

Temporarily we assume that $(d,m) = 1$ and $q = p$ or p^2, p prime.

Let

(2.4) $$g(X) = f(X)^{\frac{q-1}{d}}.$$

<u>Lemma 2.3:</u> Suppose $h_o(X), h_1(X), \ldots, h_{d-1}(X)$ are polynomials of the type

$$h_i(X) = k_{io}(X) + X^q k_{il}(X) + \ldots + X^{qK} k_{iK}(X)$$

for $0 \le i \le d-1$, and where $\deg k_{ij} \le \frac{q}{d} - m$. If

$$h_o(X) + g(X)h_1(X) + \ldots + g(X)^{d-1} h_{d-1}(X) = 0,$$

then each polynomial $k_{ij}(X) = 0$ $(0 \le i \le d-1,\ 0 \le j \le K)$.

<u>Proof:</u> A typical summand is of the form

$$\ell_{ij}(X) = g(X)^i X^{qj} k_{ij}(X).$$

It is sufficient to show that the degrees of non-zero summands are all distinct. We have

$$\deg \ell_{ij} = qj + \frac{(q-1)}{d} mi + \deg k_{ij}$$

$$= \frac{q}{d}(dj + mi) + \deg k_{ij} - \frac{i}{d} m,$$

hence $\frac{q}{d}(dj + mi) - m < \deg \ell_{ij} \leq \frac{q}{d}(dj + mi) + \frac{q}{d} - m.$

Thus we have to show that, if $(i,j) \neq (i',j')$, then

$$dj + mi \neq dj' + mi'.$$

If $dj + mi = dj' + mi'$,
then $mi \equiv mi' \pmod{d}$.

Since $(m,d) = 1$, we have $i \equiv i' \pmod{d}$. But $0 \leq i, i' \leq d-1$, we must have $i = i'$ and then $j = j'$. Hence the lemma.

Let $K[X]$ be the ring of polynomials over K. Let D be the differentiation operator defined by
$D(a_0 + a_1 X + \ldots + a_t X^t) = a_1 + 2a_2 X + \ldots + t\, a_t X^{t-1}$.

Lemma 2.4: Let K be a field of characteristic p, p a prime, and let M be an integer, $M \leq p$. Suppose $a(X) \in K[X]$ and for some $x \in K$,

$$0 = a(x) = Da(x) = D^2 a(x) = \ldots = D^{M-1} a(x).$$

Then $a(X)$ has a zero at x of order M; i.e. $(X-x)^M$ divides $a(X)$.

Proof: Write

$$a(X) = c_0 + c_1 (X-x) + c_2 (X-x)^2 + \ldots + c_t (X-x)^t .$$

Then,

$$D^{\ell}a(X) = \ell! \left[c_{\ell} + \binom{\ell+1}{\ell} c_{\ell+1} (X-x) + \ldots + \binom{t}{\ell} c_t (X-x)^{t-\ell}\right].$$

Substituting x, for $0 \leq \ell \leq M-1$, we have

$$0 = \ell!\, c_{\ell} .$$

But $\ell \leq M-1 < p$, so $\ell! \neq 0$ in K. Hence $c_{\ell} = 0$, $0 \leq \ell \leq M-1$. Then we get that $(X-x)^m$ divides $a(X)$.

Lemma 2.5: (Fundamental lemma): Let ε be an integer, $1 \leq \varepsilon \leq d-1$, and let $a(Z)$ be a polynomial of degree ε. Let G be the set of $x \in F_q$ with either $a(g(x)) = 0$ or $f(x) = 0$. Let $M \geq m+1$ be an integer with

$$(M+3)^2 \leq \frac{2q}{d} .$$

Then there exists a polynomial $r(X) \neq 0$, which has a zero of order $\geq M$ for every $x \in G$ and has

$$\deg r \leq \frac{\varepsilon}{d} qM + 4mq.$$

Proof: Let us try

$$r(X) = f(X)^M \sum_{i=0}^{d-1} \sum_{j=0}^{K} k_{ij}(X)\, g(X)^i\, X^{qj} ,$$

where the $k_{ij}(X)$ are polynomials with coefficients to be determined and $\deg k_{ij} \leq \frac{q}{d} - m$, and

$$K = \left[\frac{\varepsilon}{d}(M+m+1)\right],$$

"[ ]" denoting the integral part.

If D is the differentiation operator, then by induction on ℓ,

for $0 \le \ell \le M-1$,

$$D^{\ell} r(X) = f(X)^{M-\ell} \sum_{i=0}^{d-1} \sum_{j=0}^{K} k_{ij}^{(\ell)}(X)\, g(X)^i\, X^{qj},$$

where

$$k_{ij}^{(\ell+1)}(X) = f(X)\,(Dk_{ij}^{(\ell)}(X)) + (Df(X))\left(M-\ell+i\,\frac{q-1}{d}\right) k_{ij}^{(\ell)}(X)$$

Then $k_{ij}^{(\ell)}(X)$ is a polynomial, with

$$\deg k_{ij}^{(\ell+1)}(X) \le \deg k_{ij}^{(\ell)}(X) + m-1.$$

In particular

$$\begin{aligned} \deg k_{ij}^{(\ell)}(X) &\le \deg k_{ij}(X) + \ell(m-1) \\ &\le \frac{q}{d} - m + \ell(m-1) \\ &< \frac{q}{d} + \ell(m-1) - 1, \end{aligned}$$

by (2.3).

Since $(M+3)^2 \le \frac{2q}{d}$, we have $M < \sqrt{q}$, and since $q = p$ or p^2, we have $M < p$. Now to apply Lemma 2.4, we need for $x \in G$,

$$D^{\ell} r(x) = 0 \qquad (0 \le \ell \le M-1).$$

Since $a(Z)$ is of degree ε, for any $z \varepsilon F_q$ with $a(z) = 0$, we have

$$z^{\varepsilon} = c_0 + c_1 z + \ldots + c_{\varepsilon-1} z^{\varepsilon-1}.$$

Hence for $i \ge 0$,

$$z^i = c_0^{(i)} + c_1^{(i)} z + \ldots + c_{\varepsilon-1}^{(i)} z^{\varepsilon-1}.$$

In particular, for $x \in F_q$ with $a(g(x))=0$, we have $x^q = x$ and

$$g(x)^i = \sum_{t=0}^{\varepsilon-1} c_t^{(i)}\, g(x)^t .$$

Then for such an x,

$$D^{\ell} r(x) = f(x)^{M-\ell} \sum_{t=0}^{\varepsilon-1} s_t^{(\ell)}(x)\, g(x)^t,$$

where

$$s_t^{(\ell)}(X) = \sum_{i=0}^{d-1} \sum_{j=0}^{K} c_t^{(i)}\, k_{ij}^{(\ell)}(X)\, X^j .$$

So $D^{\ell} r(x) = 0$ for $x \in F_q$, $a(g(x)) = 0$, if the polynomials $s_t^{(\ell)}(X)$ $(0 \le t \le \varepsilon - 1)$ are all identically zero.

Note that

$$\deg s_t^{(\ell)} < \frac{q}{d} + \ell(m-1) - 1 + K.$$

Let B bethe number of coefficients of $s_t^{(\ell)}$, for $0 \le t \le \varepsilon-1$, $0 \le \ell \le M-1$, then

$$\begin{aligned} B &< \varepsilon \sum_{\ell=0}^{M-1} \left(\frac{q}{d} + \ell(m-1) + K \right) \\ &< \varepsilon M \left(\frac{q}{d} + K\right) + \frac{M^2}{2}(m-1)\,\varepsilon \\ &\le \varepsilon \frac{q}{d} M + \varepsilon M \left(\frac{\varepsilon}{d}(M+m+1) \right) + \frac{M^2}{2}(m-1)\,\varepsilon. \\ &\le \frac{\varepsilon q M}{d} + \varepsilon M^2 \left(\frac{m-1}{2} + \frac{\varepsilon}{d} \right) + \varepsilon M (m+1) \\ &< \varepsilon \frac{q}{d} M + \varepsilon M^2 \left(\frac{m+1}{2}\right) + \varepsilon M (m+1). \end{aligned}$$

Let A be the number of possible coefficients of all the k_{ij}, then

$$A \geq \left(\frac{q}{d} - m\right) d\,(K+1)$$

$$\geq (q - md)\ \frac{\varepsilon}{d}\,(M + m + 1)$$

$$= \varepsilon\,\frac{q}{d}\ M + \varepsilon\,\frac{q}{d}\,(m+1) - m\varepsilon(M+m+1)$$

$$\geq \varepsilon\,\frac{q}{d}\,M + \varepsilon\,\frac{q}{d}\,(m+1) - m\,\varepsilon\,(2M).$$

since $M \geq m+1$.

If $s_t^{(\ell)}$ are identically zero then the coefficients are homogeneous linear equations in the coefficients of k_{ij}. Then if $B < A$, we can obtain a non-trivial solution for these system of homogeneous linear equations. But $B < A$, if

$$M^{2\left(\frac{m+1}{2}\right)} + M\,(m+1) < \frac{q}{d}\,(m+1) - m\,(2M)$$

or $$M^{2\,\left(\frac{m+1}{2}\right)} + 3M\,(m+1) < \frac{q}{d}\,(m+1)$$

or $$M^2 + 6M < 2q/d\ .$$

This is guaranteed by the hypothesis that $(M+3)^2 \leq 2q/d$.

We constructed $r(X)$ such that it has a zero of order $\geq M$ for $x \in F_q$ with $a(g(x)) = 0$. Since $r(X)$ has a factor $f(X)^M$, it is clear that $r(X)$ has zero of order atleast M for each $x \in G$. By Lemma 2.3, $r(X) \neq 0$.

Finally

$$\deg r(X) \leq mM + \frac{q}{d} - m + (d-1)\, m\, \frac{(q-1)}{d} + qK$$

$$\leq Mm + \frac{q}{d} - m + (d-1)m\, \frac{(q-1)}{d} + q\left(\frac{\varepsilon}{d}(M+m+1)\right)$$

$$\leq \frac{\varepsilon}{d}\, qM + q\left(\frac{1}{d} + m + m+1\right) + mM$$

$$\leq \frac{\varepsilon}{d}\, qM + 4\, mq.$$

Hence the lemma is proved.

Proof of Theorem 2.1: In Lemma 2.5, the polynomial r(X) was constructed with a zero of order atleast M for every $x \in G$. But the number of zeros of r(X), counted with multiplicities, cannot exceed its degree; hence,

$$|G| \cdot M \leq \deg r \leq \frac{\varepsilon}{d}\, qM + 4q\, m,$$

or

$$|G| \leq \frac{\varepsilon}{d}\, q + 4q\, \frac{m}{M}.$$

Choose

$$M = \left[\sqrt{\frac{2q}{d}}\right] - 3 .$$

Since $q > 100\, dm^2$, we have

$$M \geq \sqrt{\frac{2q}{d}} - 4 \geq \sqrt{\frac{q}{d}} \geq m+1 .$$

Therefore

$$|G| \leq \frac{\varepsilon}{d}\, q + 4\, md^{\frac{1}{2}}\, q^{\frac{1}{2}} .$$

First choose $a(Z) = Z-1$; here $\varepsilon=1$. Then G is the set of $x \in F_q$ with either $g(x) = 1$ or $f(x) = 0$. Thus

$$|G| = N_1 + N_0 \leq \frac{q}{d} + 4\,md^{\frac{1}{2}}\,q^{\frac{1}{2}}.$$

whence

$$(2.5).\quad N = dN_1 + N_0 \leq d\,|G| \leq q + 4\,m\,d^{3/2}\,q^{1/2}.$$

Secondly, choose $a(Z) = Z^{d-1} + \ldots + Z + 1$. Here $\varepsilon = d-1$.

Now, $G = \left\{ x \in F_q : g(x)^{d-1} + \ldots + g(x) + 1 = 0 \text{ or } f(x) = 0 \right\}$.

Then,

$$|G| = N_2 + N_0 \leq \frac{d-1}{d}\,q + 4\,m\,d^{\frac{1}{2}}\,q^{\frac{1}{2}}$$

But $N_1 = q - N_0 - N_2 \geq \frac{q}{d} - 4\,md^{\frac{1}{2}}\,q^{\frac{1}{2}}$,

hence

$$(2.6)\quad N \geq dN_1 \geq q - 4\,md^{3/2}\,q^{1/2}.$$

Now combining (2.5) and (2.6) we get

$$|N - q| \leq 4\,m\,d^{3/2}\,q^{1/2}.$$

This does not complete the proof of Theorem 2.1 in its generality.

2.2. Removal of the Conditions $(m,d) = 1$.

The condition $(m,d) = 1$ was only required in the proof of Lemma 2.3. We prove this lemma under the condition that $Y^d - f(X)$ is absolutely irreducible.

Remark: Recall that $h_i(X)$ was a polynomial of the type

$$h_i(X) = k_{i0}(X) + X^q\,k_{i1}(X) + \ldots + X^{qK}\,k_{iK}(X),$$

where $\deg k_{ij} \leq \frac{q}{d} - m$.

It is easy to see that for $c \in F_q$, $h_i(X-c)$ is a polynomial of the same type. Since deg $f(X)$ = deg $f(X-c)$ and by Lemma 1.1 of Chapter I, $Y^d - f(X)$ is absolutely irreducible if and only if $Y^d - f(X-c)$ is absolutely irreducible and the number of roots of $Y^d - f(X)$ is equal to the number of roots of $Y^d - f(X-c)$, we may make a substitution $X \longrightarrow X - c$ and replace the polynomial $f(X)$ by $f(X-c)$. If $q > m$, we may choose $c \in F_q$ with $f(-c) \neq 0$. Therefore without loss of generality, we may assume that $f(0) \neq 0$.

First we consider the case $d=2$. Assume that $Y^2 - f(X)$ is absolutely irreducible and suppose

$$(2.7) \qquad h_0(X) + h_1(X)\, g(X) = 0,$$

$$\text{or} \qquad h_0(X) = -h_1(X)\, f(X)^{\frac{q-1}{2}}.$$

Then

$$h_0^2(X)\, f(X) = h_1^2(X)\, f(X)^q$$

$$\text{i.e. } (k_{00}(X) + X^q k_{01}(X) + \ldots + X^{qK} k_{0K}(X))^2 f(X) =$$

$$(k_{10}(X) + X^q k_{11}(X) + \ldots + X^{qK} k_{1K}(X))^2 f(X)^q .$$

Then, for some polynomial $\ell(X)$, we have

$$k_{00}^2(X)\, f(X) = k_{10}^2(X)\, f(0)^q + X^q \ell(X)$$

$$= k_{10}^2(X)\, f(0) + X^q \ell(X).$$

Here deg $k_{00}^2(X)\, f(X) \leq q - 2m + m = q - m < q$,

deg $k_{10}^2(X)\, f(0) \leq q - 2m < q$.

It follows that

$$k_{00}^2(X)\, f(X) = k_{10}^2(X)\, f(0) .$$

If $k_{00}(X) \neq 0$,

$$f(X) = \left(\sqrt{f(0)}\, \frac{k_{10}(X)}{k_{00}(X)} \right)^2$$

Then $Y^2 - f(X) = \left(Y - \sqrt{f(0)}\, \frac{k_{10}(X)}{k_{00}(X)}\right)\left(Y + \sqrt{f(0)}\, \frac{k_{10}(X)}{k_{00}(X)}\right).$

The factors of $Y^2-f(X)$ have coefficients in $\bar{F}_q(X)$ which is the quotient field of $\bar{F}_q[X]$ a unique factorization domain and $Y^2-f(X)$ is irreducible over $\bar{F}_q(X)$ if and only if it is irreducible over $\bar{F}_q[X]$. Thus we arrive at a contradiction since $Y^2-f(X)$ is absolutely irreducible. Therefore, $k_{00}(X)=0$ and $k_{10}(X) = 0$, since $f(0) \neq 0$. Then dividing (2.7) by X^q and repeating the argument, we conclude that $k_{01} = k_{11} = 0$. Continuing in this way we see that all the k_{ij} are 0. Hence Lemma 2.3 holds in the case $d = 2$.

Let $d > 2$. Form a polynomial

$$a(Y; H_0, H_1, \ldots, H_{d-1}) = H_0 + H_1 Y + \ldots + H_{d-1} Y^{d-1}.$$

Let $\zeta_1, \ldots, \zeta_d$ be elements of $\bar{F}_q$ with

$$X^d - 1 = (X-\zeta_1) \ldots (X - \zeta_d),$$

and put

$$b(Y; H_0, \ldots, H_{d-1}) = \prod_{i=1}^{d} a(\zeta_i Y; H_0, \ldots, H_{d-1}).$$

Then b is a polynomial symmetric in $\zeta_1 Y,..,\zeta_d Y$. Then by the fundamental theorem on symmetric polynomials, b must be a polynomial in the elementary symmetric functions $s_1,...,s_d$ of $\zeta_1 Y,...,\zeta_d Y$. But in this case $s_1 = ...= s_{d-1} = 0$ and $s_d = -Y^d$, so that

$$b(Y; H_o,...,H_{d-1}) = c(Y^d; H_o,...,H_{d-1}).$$

Here $c(W; H_o,...,H_{d-1})$ is a polynomial of degree d-1 in W, and of degree d in $H_o,...,H_{d-1}$. Now set

$$d(U,V; H_o,...,H_{d-1}) = V^{d-1}\, c(U/V; H_o,...,H_{d-1}).$$

Then d is a form of degree d-1 in U, V; and of degree d in $H_o,...,H_{d-1}$.

Now assume that $Y^d-f(X)$ is absolutely irreducible. Suppose

$$(2.8)\quad h_o(X) + h_1(X)\, g(X) + ...+ h_{d-1}(X)\, g(X)^{d-1} = 0.$$

With the above notation,

$$a\,(g(X);\, h_o(X),...,h_{d-1}(X)) = 0,$$

and then

$$c\,(g(X)^d\,;\, h_o(X),...,h_{d-1}(X)) = 0.$$

Since $g(X) = f(X)^{\frac{q-1}{d}}$, we obtain $g(X)^d = f(X)^q/f(X)$

and then

$$d(f(X)^q,\, f(X);\, h_o(X),...,h_{d-1}(X)) = 0.$$

Collecting all terms with no factor of X^q,

(2.9) $d(f(0),f(X); k_{00}(X),\ldots,k_{d-1,0}(X))+X^q\ell(X) = 0,$

for some polynomial ℓ. Now

(2.10) $d(f(0), f(X); k_{00}(X),\ldots,k_{d-1,0}(X))$

is of degree d-1 in f(o), f(X) and of degree d in $k_{00},..,k_{d-1,o}$.

But deg $k_{ij} \leq q/d - m$, so that degree of the polynomial (2.10) $\leq$ $(d-1)m + (\frac{q}{d} - m)\, d < q$. Hence by (2.9), we have

$$d(f(0), f(X); k_{00}(X),\ldots,k_{d-1,o}(X)) = 0.$$

Let η be the algebraic function with

$$\eta^d = \frac{f(X)}{f(0)}.$$

By Theorem 1.1 of Chapter I, $Y^d - \frac{1}{f(0)} f(X)$ is absolutely irreducible. Hence η is of degree d over $\bar{F}_q(X)$. Retracing the steps, we must have

$$c\,(f(0)/f(X), k_{00}(X),\ldots,k_{d-1,0}(X)) = 0,$$

or $$c(\frac{1}{\eta^d}; k_{00}(X),\ldots, k_{d-1,0}(X)) = 0$$

and $$b(\frac{1}{\eta}; k_{00}(X),\ldots, k_{d-1,0}(X)) = 0.$$

Therefore, some factor

$$a\,(\frac{\ell}{\eta}; k_{00}(X),\ldots, k_{d-1,0}(X)) = 0$$

or $$k_{00}(X) + \frac{\ell}{\eta}\, k_{10}(X) + \ldots + (\frac{\ell}{\eta})^{d-1} k_{d-1,0}(X) = 0.$$

But η is algebraic of degree d over $\bar{F}_q(X)$, so that

$$k_{oo}(X) = \ldots = k_{d-1,0}(X) = 0.$$

Now divide (2.8) by X^q and proceed similarly to obtain

$$k_{o1}(X) = \ldots = k_{d-1,1}(X) = 0.$$

Continuing in this way we see that all the $k_{ij}(X)$ are zero.

Thus we have proved Lemma 2.3 under the condition that $Y^d - f(X)$ is absolutely irreducible.

2.3. Removal of the Condition that $q = p$ or p^2.

Let K be a field. The polynomial ring K[X] is a vector space over K. Let $E^{(\ell)}$ ($\ell = 0,1,\ldots$) be the linear operator on K[X] with

$$E^{(\ell)}(X^t) = \binom{t}{\ell} X^{t-\ell} \quad (t=0,1,\ldots).$$

We call the operators $E^{(\ell)}$ hyperderivatives.

If D is the differential operator, then $D^{\ell}(X^t) = \ell! \binom{t}{\ell} X^{t-\ell}$ and hence $D^{\ell} = \ell!\, E^{(\ell)}$. Thus if K is of charactristic 0, then

$$E^{(\ell)} = \frac{1}{\ell!} D^{\ell}.$$

Lemma 2.6:

$$E^{(\ell)}(f_1(X),\ldots,f_t(X)) = \sum_{\substack{i_1 \geq 0,\ldots,i_t \geq 0 \\ i_1 + \ldots + i_t = \ell}} E^{(i_1)}(f_1(X)) \ldots E^{(i_t)}(f_t(X)).$$

Proof: We prove this by indiction on t. Let t = 2, then we have to show that

$$(2.11)\qquad E^{(\ell)}(f(X)g(X)) = \sum_{i=0}^{\ell} E^{(i)}(f(X))\, E^{(\ell-i)}(g(X)).$$

Since $E^{(j)}$ is a linear operator, we may suppose that f(X) and g(X) are monomials; say $f(X)=X^a$ and $g(X) = X^b$. Then (2.11) is equivalent to

$$\binom{a+b}{\ell} = \sum_{i=0}^{\ell} \binom{a}{i}\binom{b}{\ell-i}.$$

This is an immediate consequence of the definition of $\binom{a+b}{\ell}$ as the number of subsets with ℓ elements contained in a set of a+b elements.

Now assume the result for t-1, $t>2$, then

$$E^{(\ell)}(f_1(X)\ldots f_t(X)) = \sum_{\substack{i_1\geq 0, j\geq 0\\ i_1+j=\ell}} E^{(i_1)}(f_1(X))E^{(j)}(f_2(X)..f_t(X))$$

$$= \sum_{\substack{i_1\geq 0, j\geq 0\\ i_1+j=\ell}} E^{i_1}(f_1(X)) \sum_{\substack{i_2\geq 0,..,i_t\geq 0\\ i_2+..+i_t=j}} E^{(i_2)}(f_2(X))..E^{(i_t)}(f_t(X))$$

$$= \sum_{\substack{i_1\geq 0,..,i_t\geq 0\\ i_1+..+i_t=\ell}} E^{(i_1)}(f_1(X))\ldots E^{(i_t)}(f_t(X))$$

Corollary 2.1. $E^{(\ell)}(X-c)^t = \binom{t}{\ell}(X-c)^{t-\ell}$.

Proof: $$E^{(\ell)}(X-c)^t = \sum_{\substack{i_1 \geq 0,\ldots,i_t \geq 0 \\ i_1+\ldots+i_t=\ell}} E^{(i_1)}(X-c)..E^{(i_t)}(X-c)$$

But $E^{(1)}(X-c) = 1$ and $E^{(i)}(X-c) = 0$ if $i \geq 2$. Then in the above sum, we need only consider summands with each i_j either 0 or 1. The number of such summands is $\binom{t}{\ell}$, and each summand is $(X-c)^{t-\ell}$.

Corollary 2.2: If $0 \leq \ell \leq t$, then

(2.12) $$E^{(\ell)}(a(X)\, f(X)^t) = b(X)\, f(X)^{t-\ell},$$

where $b(X)$ is a polynomial with

$$\deg b = \deg a + \ell((\deg f) - 1).$$

Proof: We have

$$E^{(\ell)}(a(X)\, f(X)^t) = \sum_{\substack{i_0 \geq 0,\ldots,i_t \geq 0 \\ i_0+\ldots+i_t=\ell}} E^{i_0}(a(X))E^{(i_1)}(f(X))..E^{(i_t)}(f(X)).$$

Here every summand is divisible by $f(X)^{t-\ell}$. Hence a formula such as (2.12) holds. Furthermore

$$\begin{aligned}\deg b &= \deg E^{(\ell)}(af^t) - (t-\ell)\deg f \\ &= \deg a + t\deg f - \ell - (t-\ell)\deg f \\ &= \deg a + \ell((\deg f) - 1).\end{aligned}$$

Theorem 2.2: Suppose $E^{(\ell)}(f(x)) = 0$ for $\ell = 0,1,..,M-1$. Then $(X-x)^M$ divides $f(X)$.

Proof: We may write $f(X) = a_0 + a_1 (X-x) + \ldots + a_d (X-x)^d$.

Then by Corollary 2.1,

$$E^{(\ell)}(f(X)) = a_\ell + \binom{\ell+1}{\ell} a_{\ell+1} (X-x) + \ldots + \binom{d}{\ell} a_d (X-x)^{d-\ell}.$$

Substituting x, for $0 \leq \ell \leq M-1$, we see that $a_\ell = 0$ for $\ell = 0,1,..$ $\ldots M-1$ and hence $(X-x)^M$ divides $f(X)$.

Lemma 2.7: Suppose K is of characteristic $p > 0$. Let $r(X) = h(X,X^{p^\mu})$ for some polynomial $h(X,Y)$. Then for $\ell < p^\mu$,

$$E^{(\ell)} r(X) = E_X^{(\ell)} h(X,X^{p^\mu}),$$

where $E_X^{(\ell)}$ is the ''partial hyperderivative'' with respect to X of $h(X,Y)$.

Proof: By linearity, it suffices to take the case when $h(X,Y) = X^a Y^b$. Then by Lemma 2.6, we have

$$E^{(\ell)} r(X) = E^{(\ell)}(X^a X^{bp^\mu})$$

$$= \sum_{\substack{i_1 \geq 0, i_2 \geq 0 \\ i_1+i_2 = \ell}} E^{(i_1)}(X^a)\, E^{(i_2)}(X^{bp^\mu})$$

$$= \sum_{\substack{i_1 \geq 0, i_2 \geq 0 \\ i_1+i_2 = \ell}} E^{(i_1)}(X^a) \sum_{\substack{j_1 \geq 0,\ldots,j_b \geq 0 \\ j_1+\ldots+j_b = i_2}} E^{(j_1)}(X^{p^\mu}) \ldots E^{(j_b)}(X^{p^\mu})$$

and

$$E_X^{(\ell)} h(X,X^{p^\mu}) = X^{bp^\mu} E^{(\ell)}(X^a).$$

Hence it suffices to show that for $0 < \ell < p^\mu$,

$$E^{(\ell)}(X^{p^\mu}) = \binom{p^\mu}{\ell} X^{p^\mu-\ell} = 0 .$$

But since characteristic of K is p, we have for $0<\ell< p^\mu$,

$$\binom{p^\mu}{\ell} = \left(\frac{p^\mu}{\ell}\right) \binom{p^\mu-1}{\ell-1} = 0.$$

Hence the Lemma.

The condition that $q = p$ or p^2 was only required in the proof of Lemma 2.5. Now we prove Lemma 2.5 in general. Set

$$r(X) = h(X,X^q)$$

with

$$h(X,Y) = f(X)^M \sum_{i=0}^{d-1} \sum_{j=0}^{K} k_{ij}(X)\, g(X)^i\, Y^j$$

where the $k_{ij}(X)$ are polynomials with coefficients to be determined with $\deg k_{ij} \le \frac{q}{d} - m$, and

$$K = \left[\frac{\varepsilon}{d} (M + m+1) \right].$$

By corollary 2.2, since $g(X)$ is a power of $f(X)$ we have for $o \le \ell \le M-1$

$$E^{(\ell)}(f(X)^M k_{ij}(X)\, g(X)^i) = f(X)^{M-\ell} k_{ij}^{(\ell)}(X)\, g(X)^i,$$

where $\deg k_{ij}^{(\ell)} \le \deg k_{ij} + \ell(m-1)$.

By Lemma 2.7 we have, for $0 \leq \ell < M \leq q = p^k$,

$$E^{(\ell)} r(X) = f(Y)^{M-\ell} \sum_{i=0}^{d-1} \sum_{j=0}^{K} k_{ij}^{(\ell)}(X)\, g(X)^i\, X^{qj} .$$

Now we have to find $h(X,Y)$ such that for every $x \in G$,

$$E^{(\ell)} r(X) = 0 \qquad (0 \leq \ell \leq M-1).$$

For any $z \in F_q$ satisfying $a(z) = 0$, we have

$$z^{\varepsilon} = c_0 + c_1 z + \ldots + c_{\varepsilon-1} z^{\varepsilon-1} ,$$

since $a(z)$ is of degree ε. Hence for $i \geq 0$,

$$z^i = c_0^{(i)} + c_1^{(i)} z + \ldots + c_{\varepsilon-1}^{(i)} z^{\varepsilon-1} .$$

In particular, for $x \in F_q$ satisfying $a(g(x)) = 0$, we have $x^q = x$ and

$$g(x)^i = \sum_{t=0}^{\varepsilon-1} c_t^{(i)}\, g(x)^t .$$

Then for such an x,

$$E^{(\ell)} r(x) = f(x)^{M-\ell} \sum_{t=0}^{\varepsilon-1} s_t^{(\ell)}(x)\, g(x)^t ,$$

where

$$s_t^{(\ell)}(X) = \sum_{i=0}^{d-1} \sum_{j=0}^{K} c_t^{(i)}\, k_{ij}^{(\ell)}(X)\, X^j .$$

So certainly $E^{(\ell)} r(x) = 0$ for $x \in F_q$, $a(g(x)) = 0$, provided the polynomials

$$s_t^{(\ell)}(X) \qquad (0 \le t \le \varepsilon-1)$$

are all identically zero.

Note that

$$\deg\ s_t^{(\ell)} \le \frac{q}{d} + \ell(m-1) - m + K$$

$$< \frac{q}{d} + \ell(m-1) - 1 + K.$$

Let B be the number of coefficients of $s_t^{(\ell)}$ for $0 \le t \le \varepsilon - 1$, $0 \le \ell \le M-1$, then

$$B < \varepsilon \sum_{\ell=0}^{M-1} \left(\frac{q}{d} + \ell(m-1) + K \right)$$

$$< \varepsilon M \left(\frac{q}{d} + K \right) + \frac{M^2}{2} (m-1)\, \varepsilon$$

$$\le \varepsilon \frac{q}{d} M + \varepsilon M \left(\frac{\varepsilon}{d} (M + m + 1) \right) + \frac{M^2}{2}(m-1)\, \varepsilon .$$

$$\le \varepsilon \frac{q}{d} M + \varepsilon M^2 \left(\frac{m-1}{2} + \frac{\varepsilon}{d} \right) + \varepsilon M (m+1)$$

$$< \varepsilon \frac{q}{d} M + \varepsilon M^2 \left(\frac{m+1}{2} \right) + \varepsilon M (m+1).$$

Let A be the number of possible coefficients of all the k_{ij}, then

$$A \ge \left(\frac{q}{d} - m \right) d \ (K + 1)$$

$$\ge (q - md) \frac{\varepsilon}{d} (M + m + 1)$$

$$= \varepsilon \frac{q}{d} M + \varepsilon \frac{q}{d} (m+1) - m\, \varepsilon\, (M+m+1)$$

$$\geq \varepsilon \frac{q}{d} M + \varepsilon \frac{q}{d} (m+1) - m \varepsilon (2M) -$$

since $M \geq m + 1$.

If $s_t^{(\ell)}$ are identically zero, then the coefficients of $s_t^{(\ell)}$ are homogeneous linear equation in the coefficients of k_{ij}. Then if $B < A$, we can obtain a non-trivial solution for these system of homogeneous linear equations. But $B < A$ if

$$M^2 \left(\frac{m+1}{2} \right) + M (m+1) < \frac{q}{d} (m+1) - m (2M)$$

or $$M^2 \left(\frac{m+1}{2} \right) + 3M (m+1) < \frac{q}{d} (m+1)$$

or $$M^2 + 6M < 2 \frac{q}{d} .$$

This is guaranteed by the hypothesis that $(M+3)^2 \leq 2 \frac{q}{d}$.

So $E^{(\ell)}(r(x))=0$ for $x \varepsilon F_q$, $a(g(x))=0$, $0 \leq \ell \leq M-1$, then by Theorem 2.2, $(X-x)^M$ divides $r(X)$.

We constructed $r(X)$ such that it has a zero of order $\geq M$ for $x \varepsilon F_q$ with $a(g(x)) = 0$. Since $r(X)$ has a factor $f(X)^M$, it is clear that $r(X)$ has zero of order atleast M for each $x \varepsilon G$. By Lemma 2.3, $r(X) \neq 0$.

Finally,

$$\deg r(X) \leq mM + \frac{q}{d} - m + (d-1)m \frac{(q-1)}{d} + qK$$

$$\leq mM + \frac{q}{d} - m + (d-1)m \left(\frac{q-1}{d} \right) + q \left(\frac{\varepsilon}{d} (M+m+1) \right)$$

$$\leq \frac{\varepsilon}{d}\ qM+q\ (\frac{1}{d} + m + m + 1) + mM$$

$$\leq \frac{\varepsilon}{d}\ qM + 4\ mq$$

and the Lemma 2.5 is proved in general, and hence Theorem 2.1 is proved in general.

2.4 Absolutely irreducible equations $y^q - y = f(x)$.

Let $r = q^k$, then every automorphism of F_r over F_q is of the form ω^i ($1 \leq i \leq k-1$), where $\omega^i(x) = x^{q^i}$. The trace of an element x of F_r over F_q is

$$\mathcal{T}(x) = x + x^q + \ldots + x^{q^{k-1}}$$

Lemma 2.8: Let $x \in F_r$, with $F_q \subseteq F_r$, then the following are equivalent:

(i) $\mathcal{T}(x)=0$

(ii) There exists $y \in F_r$ with $x = y^q - y$

(iii) There exists precisely q elements $y \in F_r$ with $x = y^q - y$.

Proof: (i) $\Rightarrow$ (ii).

Since F_r is a finite separable extension of F_q, we have

$$\mathcal{T}: F_r \longrightarrow F_q$$

is a non-zero functional. Then there exists a $\theta \in F_r$, with $\mathcal{T}(\theta) \neq 0$. Now let

$$\alpha = \frac{1}{\mathcal{T}(\theta)} \left[x\,\theta^q + (x+x^q)\theta^{q^2} + \ldots + (x+x^q+\ldots+x^{q^{k-2}})\,\theta^{q^{k-1}}\right].$$

Then $x = \alpha - \alpha^q$.

Consider $-\alpha$, then $(-\alpha)^q - (-\alpha) = -\alpha^q + \alpha = x$. Take $y = -\alpha$, then $x = y^q - y$.

(ii) $\Rightarrow$ (iii).

We have $Y^q - Y - x = 0$ has at most q roots in F_r.

By (ii) there exists $y \in F_r$, such that $x = y^q - y$. Then for any $\alpha \in F_q$, we have

$$(y+\alpha)^q - (y+\alpha) = y^q + \alpha^q - y - \alpha = y^q - y = x$$

Therefore there are exactly q elements $y \in F_r$, with $x = y^q - y$.

(iii) $\Rightarrow$ (i).

By (iii), there exists an element $y \in F_r$, with $x = y^q - y$.
Then

$$\mathcal{T}(x) = y^q - y + y^{q^2} - y^q + \ldots + y^{q^k} - y^{q^{k-1}} = 0.$$

Let $f(x) \in F_q[X]$ and let N be the number of solutions $(x,y) \in F_r^2$ of $y^q - y = f(x)$.

For $w \in F_q$, let N_w be the number of $x \in F_r$ with

$$\mathcal{T}(f(x)) = w .$$

<u>Lemma 2.9:</u>

$$\sum_{w \in F_q} N_w = r \text{ and } N = qN_o.$$

<u>Proof:</u> The first statement is obvious. The fact that $N = qN_o$ follows from Lemma 2.8.

<u>Theorem 2.3:</u> Suppose $r = q^k$. Let $f(X) \in F_q[X]$, with $(q, \deg f) = 1$ and $\deg f < q$. If N is the number of solutions $(x,y) \in F_r^2$ of

$y^q - y = f(x)$, then

$$|N - r| < q^{[\frac{k}{2}]+4}$$

Note: This theorem yields no information when k=2; we get

$$|N - q^2| < q^5$$

but obviously

$$0 \leq N \leq |F_r^2| = q^4.$$

Thus we may assume that $k \geq 3$. Let $\gamma = [\frac{k}{2}]$, then $\gamma \geq 1$. Let

$$g(X) = f(X)^{q^\gamma} + f(X)^{q^{\gamma+1}} + \ldots + f(X)^{q^{k-1}} ,$$

$$h(X) = f(X) + f(X)^q + \ldots + f(X)^{q^{\gamma-1}} .$$

Lemma 2.10: Let $w \in F_q$ be fixed. Let M be divisible by q, and $0<M \leq q^{k-\gamma-1}$. Then there is a polynomial $u(X) \neq 0$, which has a zero of order $\geq M$ for every $x \in F_r$ with

$$\mathcal{T}(f(x)) = w$$

and deg $u(X) \leq M \frac{r}{q} + q^{k+1}$.

Proof: Let us try

$$u(X) = \sum_{i=0}^{q-1} \sum_{j=0}^{K} k_{ij}(X)\, g(X)^i\, X^{rj} ,$$

where $K = \frac{M}{q}$, and the polynomials $k_{ij}(X)$ have deg $k_{ij} < \frac{r}{q} = q^{k-1}$, and coefficients to be determined. Since $k \leq 2\gamma + 1$,

$M \leq q^{k-\gamma-1} \leq q^\gamma$.

Thus for $\ell < M \leq q^{\gamma}$ and $U(X) = a(X, X^{q^{\gamma}})$, then by Lemma 2.7 (with $\mu = \gamma\sigma$ if $q = p^{\sigma}$) for $\ell < p^{\gamma\sigma}$

$$E^{(\ell)} U(X) = E_X^{(\ell)} a(X, X^{q^{\gamma}}).$$

Since $X^r = X^{q^k}$ and since $g(X) = f(X^{q^{\gamma}}) + \ldots + f(X^{q^{k-1}})$, it follows that

$$E^{(\ell)} u(X) = \sum_{i=0}^{q-1} \sum_{j=0}^{K} k_{ij}^{(\ell)}(X)\, g(X)^i\, X^{rj}$$

with $k_{ij}^{(\ell)}(X) = E^{(\ell)} k_{ij}(X)$.

Let A be the total number of available coefficients of the polynomials $k_{ij}(X)$. Then

$$A = q^{k-1}\, q\,(K+1) = q^{k-1} M + q^k .$$

We have for $x \in F_r$ with $\mathcal{J}(f(x)) = w$, $x^r = x$ and $w = h(x) + g(x)$. So $E^{(\ell)} u(x) = s^{(\ell)}(x)$, where

$$s^{(\ell)}(X) = \sum_{i=0}^{q-1} \sum_{j=0}^{K} k_{ij}^{(\ell)}(X)\, (w - h(X))^i\, X^j .$$

In order that $u(X)$ has a zero of order M for $x \in F_r$, with $\mathcal{J}(f(x)) = w$, by Theorem 2.2, it is certainly sufficient that the polynomials $s^{(\ell)}(X)$ are identically zero for $0 \leq \ell \leq M-1$. Since $K \leq q^{\gamma-1}$

$$\begin{aligned} \deg s^{(\ell)}(X) &\leq q^{k-1} + (q-1)^2\, q^{\gamma-1} + K \\ &\leq q^{k-1} + (q^2 + 1 - 2q)\, q^{\gamma-1} + q^{\gamma-1} \end{aligned}$$

$$= q^{k-1} + q^{\gamma+1} + 2q^{\gamma-1}(1 - q)$$

$$\leq q^{k-1} + q^{\gamma+1} - 2 .$$

Let B be the number of coefficients of $s^{(\ell)}(X)$ for $0 \leq \ell \leq M-1$. Then

$$B = \sum_{\ell=0}^{M-1} (\deg s^{(\ell)}(X) + 1)$$

$$\leq \sum_{\ell=0}^{M-1} (q^{k-1} + q^{\gamma+1} - 1)$$

$$< M(q^{k-1} + q^{\gamma+1})$$

$$\leq Mq^{k-1} + q^{k-\gamma-1} q^{\gamma+1}$$

$$= Mq^{k-1} + q^{k} .$$

If $s^{(\ell)}(X) = 0$, then the coefficients of $s^{(\ell)}(X)$ are linear homogenous equations in the coefficients of the k_{ij}. Then if $B < A$, this system of linear homogeneous equations has a nontrivial solution. Hence we may choose the coefficients of $k_{ij}(X)$ not all zero, so that $u(X)$ has a zero of order atleast M for the elements $x \in F_r$ with $\mathcal{J}(f(x)) = w$. Moreover,

$$\deg u(X) \leq rK + (q-1)^2 q^{k-1} + q^{k-1}$$

$$\leq M\frac{r}{q} + q^{k+1} .$$

Finally, $u(X)$ does not vanish identically, because the non-zero summands

$$\ell_{ij}(X) = k_{ij}(X)\, g(X)^i\, X^{rj}$$

have degrees

$$\deg \ell_{ij}(X) = rj + iq^{k-1} \deg f + \deg k_{ij}$$

$$= q^{k-1}(qj + i \deg f) + \deg k_{ij},$$

which are distinct. For if

$$q^{k-1}(qj+i \deg f) + \deg k_{ij} = q^{k-1}(qj'+ i' \deg f) + \deg k_{i'j'}$$

$\Rightarrow$ $\deg k_{ij} \equiv \deg k_{i'j'} \pmod{q^{k-1}}$

$\Rightarrow$ $\deg k_{ij} = \deg k_{i'j'}$ $(\because \deg k_{ij} < q^{k-1})$

$\Rightarrow$ $qj + i \deg f = qj' + i' \deg f$

$\Rightarrow$ $i \deg f \equiv i' \deg f \pmod{q}$

$\Rightarrow$ $i \equiv i' \pmod{q}$ $(\because (\deg f, q) = 1)$

$\Rightarrow$ $i = i'$ $(\because 0 \leq i, i' \leq q-1)$

and $j = j'$.

<u>Proof of Theorem 2.3:</u> For fixed $w \in F_q$,

$$N_w M \leq \deg u \leq M \frac{r}{q} + q^{k+1},$$

or $$N_w \leq \frac{r}{q} + \frac{q^{k+1}}{M}.$$

Choose $M = q^{k-\gamma-1}$; then for $k \geq 3$, $q \mid M$. We obtain

$$N_w \leq \frac{r}{q} + q^{\gamma+2},$$

and by Lemma 2.9,

$$N_w = r - \sum_{v \neq w} N_v$$

$$\geq r - \left(\frac{r}{q} + q^{\gamma+2} \right) (q-1)$$

$$> \frac{r}{q} - q^{\gamma+3} \quad .$$

So $\left| N_w - \frac{r}{q} \right| < q^{\gamma+3}$

and in particular,

$$\left| N_o - \frac{r}{q} \right| < q^{\gamma+3} \quad .$$

By Lemma 2.9 again,

$$\left| N - r \right| < q^{\gamma+4} = q^{\left[\frac{k}{2} \right] + 4} \quad .$$