MSc Project Report

Exponential Sums and Weil Bounds: A Survey of Elementary Methods and Decoding Applications

May, 2026

Soham Chatterjee

Supervisor: **Mrinal Kumar**

School of Technology and Computer Science
Tata Institute of Fundamental Research, Mumbai

soham.chatterjee@tifr.res.in
Website: sohamch08.github.io

Preface

Character sums of polynomials over finite fields give rise to error-correcting codes of remarkable parameters. For a multiplicative character ψ of $\mathbb{F}_q$ and polynomials $f \in \mathbb{F}_q[X]$ of small degree, the codeword $(\psi(f(\alpha)))_{\alpha \in \mathbb{F}_q}$ has length q and, by the Weil bound, relative distance approaching $1/2$. Two specific families exhibit this behavior: the quadratic residue character codes obtained by applying $\eta(\beta) = \beta^{(q-1)/2}$ to squarefree monic polynomials of degree $O(\sqrt{q})$, and the dual BCH codes obtained over $\mathbb{F}_{2^b}$ by composing the absolute trace with polynomials supported on odd-degree monomials. For both families, the distance analysis reduces entirely to the Weil bound; in the case of dual BCH codes, to such an extent that, absent the Weil bound, the existence of binary codes with these parameters is not known by any other means. Until recently the decoding problem for these codes was open even for non-constant degree; the situation was resolved by [Kop26], whose algorithm combines the Berlekamp–Welch like framework with the Stepanov polynomial method and introduces a class of high-degree polynomials, pseudopolynomials, as the right algebraic object for capturing the structure of character evaluations. The Stepanov method is also the elementary route to the Weil bound itself; the same machinery that proves the bound thus also decodes the codes it gives rise to. This report develops both threads in a unified, self-contained way.

The polynomial method of Stepanov and Bombieri works directly via auxiliary polynomials on two specific curves: $Y^d = f(X)$ for multiplicative character sums and $Y^q - Y = f(X)$ for additive character sums. The primary references are [Sch76, Chapter I-II] and [LN96, Chapter 2,5,6]. The report is organized as follows. Chapter 1 develops the theory of finite fields: field extensions, the Frobenius automorphism, and trace and norm maps. Chapter 2 treats characters of finite abelian groups and finite fields, Gaussian and Jacobi sums, and develops the analytic infrastructure of L-functions: the Riemann zeta function and its function-field analogue over $\mathbb{F}_q[X]$, the Dirichlet L-function $L(s, \chi)$, which is then used to prove the Davenport–Hasse lifting theorem relating Gaussian and Jacobi sums over $\mathbb{F}_q$ to those over its extensions. Chapter 3 addresses the problem of counting solutions to equations over finite fields and derives Weil-type bounds for additive, multiplicative, and quadratic character sums, together with the corresponding estimates for quadratic forms and diagonal equations. Chapter 4 gives the two principal elementary proofs of the Weil bound: Stepanov's proof for multiplicative character sums via the curve $Y^d = f(X)$, and Bombieri's proof for additive character sums via the Artin–Schreier curve $Y^q - Y = f(X)$, together with an alternate argument for the latter using Bézout's theorem. Chapter 5 develops the theory of pseudopolynomials and the polynomial-time decoding algorithms for the quadratic character and dual BCH code families, following Kopparty. Appendices A–D collect supporting material: algebraic and analytic prerequisites, modulus bounds from power sum estimates, continued fractions over polynomial rings, and Hasse derivatives.

The report assumes familiarity with algebra at the level of groups, rings, and fields, and with linear algebra. No prior exposure to algebraic geometry or analytic number theory is required.

Soham Chatterjee
School of Technology and Computer Science
Tata Institute of Fundamental Research, Mumbai
May 2026

Acknowledgements

I would like to thank my guide Prof. Mrinal Kumar for guiding me throughout the project. Over the course of numerous meetings, he has cleared concepts, doubts, and difficulties that I encountered, not only in this project but also during different stages of my academic life for which he has my warmest thanks. I would also like to sincerely thank Prof. Prahladh Harsha for his guidance and support throughout this period. The discussions and meetings with Prof. Prahladh Harsha and Prof. Mrinal Kumar have greatly shaped my understanding and research. They consistently suggested interesting topics and papers to read, encouraged me to learn new areas, and guided me whenever I struggled with unfamiliar concepts even outside academic life. I am deeply grateful for his kindness, patience, and support in many aspects throughout this journey.

I would also like to thank Prof. Ramprasad Saptharishi, Prof. Arkadev Chattopadhyay, Prof. Umang Bhaskar, and Prof. T. Kavitha for their courses at TIFR, which introduced me to several beautiful ideas in mathematics and theoretical computer science. Equally, I am grateful to Prof. Partha Mukhopadhyay, Prof. Amit Kumar Sinhababu, and Prof. Sumanta Ghosh for the foundations they built during my time at CMI. Their teaching shaped my love for the subject more than they perhaps know.

I would like to express my deepest gratitude to my parents for their unconditional love, sacrifices, and constant support throughout my academic journey. Despite the many hardships and medical difficulties faced by both my mother and father, they always encouraged and supported my love for mathematics and theoretical computer science. I especially want to acknowledge my mother's strength and resilience in enduring serious medical hardships while continuing to care for and encourage me at every stage. Their unwavering belief in me, even during difficult times, has been a constant source of strength, and whatever I have achieved would not have been possible without them.

Nishant Das, Shubham Bhardwaj, Spandan Podder, and Vivek Karunakaran have made my time at TIFR genuinely special. From arguments about computer science that drifted into everything else, to the kind of easy honesty that only comes with real friendship, they have been a constant presence, the sort of people whose company makes a difficult period feel manageable and, often, quite good. I did not expect to find friends like these here, and I am glad I did. I also want to thank the postdoctoral researchers at TIFR, Susobhan, Sagnik, Dipan, and Anirban, for their easy company and for advice about research life that was always as honest as it was welcome.

Saptarshi Sahoo, Sougata Panda, and Soumil Baksi are among the closest people in my life. What we had at CMI. The hours we lost track of, the arguments that had no end, the shared chaos of the S24 Room is something I carry with me, and they remain my dearest friends to this day. I am grateful for them in a way that goes well beyond what these pages can hold.

The CMI Zone deserves a mention of its own. That little corner where we spent countless hours, and where so many of the best moments quietly happened. Finally, I want to specially thank Somnath Bhattacharjee and Saswata Mukherjee, whose enthusiasm for the algebraic side of theoretical computer science was infectious and did a great deal to shape where my interests eventually led.

I also acknowledge the use of LLM (Claude) as a tool during the preparation of this document, specifically for writing LaTeX macros, setting up preamble and other structural elements, and for editorial assistance with the typesetting.

List of Notations

The following table lists the notation used throughout this report, organized by topic.

Symbol	Meaning
General	
$\mathbb{N}$	Set of all natural numbers i.e., positive integers
$\mathbb{Z}$	Set of all integers
$\mathbb{Z}_0$	Set of all non-negative integers
$\mathbb{R}$	Set of real numbers
$\mathbb{C}$	Set of complex numbers
$[n]$	The set $\{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$
i	Imaginary unit: $i^2 = -1$
S^1	$\{z \in \mathbb{C} : \lvert z\rvert = 1\}$, the multiplicative group of unit complex numbers
ζ_m	A primitive m-th root of unity
$[x]$	Floor function: largest integer $\leq x$; also written $\lfloor x \rfloor$
Finite Fields	
$\mathbb{F}_q$	Finite field with q elements
$\mathbb{F}_q^*$	Multiplicative group of nonzero elements of $\mathbb{F}_q$
$\mathbb{F}_p$	Prime field with p elements, $p = \mathrm{char}(\mathbb{F}_q)$
$\mathbb{F}_q^{*(d)}$	Subgroup of d-th powers in $\mathbb{F}_q^*$
E, F, L, K	Generic notation for finite fields or finite extensions of $\mathbb{F}_q$
$[E : \mathbb{F}_q]$	Degree of the extension $E/\mathbb{F}_q$
$\mathbb{F}_q[X_1, \ldots, X_n]$	Polynomial ring in n variables over $\mathbb{F}_q$
$\mathbb{F}_q^{\leq d}[X_1, \ldots, X_n]$	Polynomials in $\mathbb{F}_q[X_1, \ldots, X_n]$ of total degree at most d
Characters of Finite Abelian Groups	
G	A finite Abelian group with identity element 1_G
Z_k	Cyclic multiplicative group of order k
χ_0	Trivial character of G: $\chi_0(g) = 1$ for all $g \in G$
$\overline{\chi}$	Conjugate (and inverse) character of χ: $\overline{\chi}(g) = \overline{\chi(g)}$
$\mathrm{Hom}(A, B)$	Set of all group homomorphisms from A to B
$\widehat{G} = \mathrm{Hom}(G, S^1)$	Dual group of G; the group of all characters of G
$\mathrm{Ann}(\chi)$	Annihilator of character χ: $\{g \in G : \chi(g) = 1\}$, a subgroup of G
$\mathrm{Ann}(H)$	Annihilator of subgroup $H \leq G$: $\{\chi \in \widehat{G} : \chi(h) = 1 \ \forall h \in H\}$

continued on next page

(*continued from previous page*)

Symbol	Meaning
ev_g	Evaluation character of $\widehat{G}$ defined by $\mathrm{ev}_g(\chi) = \chi(g)$
Maps on Finite Fields	
$\mathrm{Tr}\colon \mathbb{F}_q \to \mathbb{F}_p$	Absolute trace function
$N\colon \mathbb{F}_q \to \mathbb{F}_p$	Absolute norm function
$\mathrm{Tr}_{E/\mathbb{F}_q}$	Trace function from E to $\mathbb{F}_q$
$\mathrm{N}_{E/\mathbb{F}_q}$	Norm function from E to $\mathbb{F}_q$
$\sigma\colon \mathbb{F}_q \to \mathbb{F}_q$	Frobenius automorphism: $\sigma(\alpha) = \alpha^p$
Additive Characters of $\mathbb{F}_q$	
χ, τ	Generic notation for additive characters
χ_1	Canonical additive character of $\mathbb{F}_q$: $\chi_1(\alpha) = e^{2\pi i\,\mathrm{Tr}(\alpha)/p}$
χ_β	Additive character defined by $\chi_\beta(\alpha) = \chi_1(\beta\alpha)$, for $\beta \in \mathbb{F}_q$
$\boldsymbol{\chi}_q$	Set of all additive characters of $\mathbb{F}_q$
$\chi^{(r)}$	Additive character of $\mathbb{F}_q$ lifted to E via $\chi^{(r)} = \chi \circ \mathrm{Tr}_{E/\mathbb{F}_q}$, $[E : \mathbb{F}_q] = r$
Multiplicative Characters of $\mathbb{F}_q$	
ψ, λ	Generic notation for multiplicative characters
ψ_0	Trivial multiplicative character of $\mathbb{F}_q$
η	Quadratic character of $\mathbb{F}_q$ (for odd q): $\eta(\alpha) = \pm 1$ according as α is a square
$\boldsymbol{\mathcal{M}}_q$	Set of all multiplicative characters of $\mathbb{F}_q^*$
$\boldsymbol{\mathcal{M}}_q^{(e)}$	Set of multiplicative characters of $\mathbb{F}_q$ of exponent e
$\mathrm{ord}(\psi)$	Order of ψ: smallest positive integer d with $\psi^d = \psi_0$
$\left(\frac{\alpha}{q}\right)$	Legendre symbol; equals $\eta(\alpha)$ for $\alpha \in \mathbb{F}_q$
$\psi^{(r)}$	Multiplicative character of $\mathbb{F}_q$ lifted to E via $\psi^{(r)} = \psi \circ \mathrm{N}_{E/\mathbb{F}_q}$, $[E\colon \mathbb{F}_q] = r$
Gaussian and Jacobi Sums	
$G(\psi, \chi)$	Gaussian sum: $\sum_{\alpha \in \mathbb{F}_q^*} \psi(\alpha)\,\chi(\alpha)$, for $\psi \in \boldsymbol{\mathcal{M}}_q$, $\chi \in \boldsymbol{\chi}_q$
$J_\alpha(\lambda_1, \ldots, \lambda_k)$	Generalized Jacobi sum: $\displaystyle\sum_{\substack{\alpha_1,\ldots,\alpha_k \in \mathbb{F}_q \\ \alpha_1+\cdots+\alpha_k=\alpha}} \lambda_1(\alpha_1)\cdots\lambda_k(\alpha_k)$
$J(\lambda_1, \ldots, \lambda_k)$	Jacobi sum; special case $\alpha = 1$ of J_α
$J_0(\lambda_1, \ldots, \lambda_k)$	Jacobi sum at zero; special case $\alpha = 0$ of J_α
Kloosterman Sums	
$K(\chi; a, b)$	Kloosterman sum: $\sum_{\gamma \in \mathbb{F}_q^*} \chi(a\gamma + b\gamma^{-1})$, for $\chi \in \boldsymbol{\chi}_q$, $a, b \in \mathbb{F}_q$
$K(\chi^{(r)}; a, b)$	Kloosterman sum lifted to the degree-r extension of $\mathbb{F}_q$
w_1, w_2	Reciprocal roots of the L-function associated to a Kloosterman sum
Salié and Jacobsthal Sums	
$S(\chi, \tau)$	Salié sum: $\sum_{\gamma \in \mathbb{F}_q^*} \eta(\gamma)\,\chi(\gamma)\,\tau(\gamma^{-1})$, for $\chi, \tau \in \boldsymbol{\chi}_q$, q odd
$H_n(\alpha)$	Jacobsthal sum: $\sum_{\gamma \in \mathbb{F}_q} \eta(\gamma^{n+1} + \alpha\gamma)$, for $\alpha \in \mathbb{F}_q^*$

continued on next page

(*continued from previous page*)

Symbol	Meaning
$I_n(\alpha)$	Associated Jacobsthal sum: $\sum_{\gamma\in\mathbb{F}_q}\eta(\gamma^n+\alpha)$, for $\alpha\in\mathbb{F}_q^*$
Classical Riemann Zeta Function	
$\zeta(s)$	Riemann zeta function: $\sum_{n=1}^{\infty} n^{-s}$, for $\Re(s)>1$
$\xi(s)$	Completed zeta function: $\frac{1}{2}s(s-1)\,\Gamma(s/2)\,\pi^{-s/2}\,\zeta(s)$
$\Gamma(s)$	Gamma function: $\int_0^\infty e^{-t}t^{s-1}\,dt$
$\Lambda(n)$	Von Mangoldt function: $\log p$ if $n=p^k$, else 0
ρ	Non-trivial zero of $\xi(s)$
Dirichlet L-functions over $\mathbb{F}_q[X]$	
Φ	Set of all monic polynomials in $\mathbb{F}_q[X]$
Φ_d	Monic polynomials in $\mathbb{F}_q[X]$ of degree exactly d
$\mathrm{Irr}(\Phi)$	Set of irreducible monic polynomials in Φ
$\mathfrak{n}(f)=q^{\deg(f)}$	Norm of the polynomial $f\in\mathbb{F}_q[X]$
G	Group of rational functions $r=h_1/h_2$ in $\mathbb{F}_q(X)$ with $h_1,h_2\in\Phi$ monic
$\overline{G}$	Subgroup of G on which a given character is defined; extended to G by zero
$L(s,\chi)$	Dirichlet L-function: $\sum_{h\in\Phi}\chi(h)\,\mathfrak{n}(h)^{-s}$, for a multiplicative character χ
$\overline{L}(z,\chi)$	Power series satisfying $\overline{L}(z,\chi)=\sum_{h\in\Phi}\chi(h)\,z^{\deg(h)}$ and $\overline{L}(q^{-s},\chi)=L(s,\chi)$
L_k	k-th log-derivative coefficient: $\displaystyle\sum_{\substack{h\in\mathrm{Irr}(\Phi)\\ \deg(h)\mid k}}\deg(h)\,\chi(h)^{k/\deg(h)}$
Equations over $\mathbb{F}_q$ *— Solution Sets*	
$\mathcal{V}(f)$	Solution set of $f(X_1,\dots,X_n)=0$ in $\mathbb{F}_q^n$
$Z(f)$	Number of solutions: $\lvert\mathcal{V}(f)\rvert$
$\overline{\mathcal{V}}(f)$	Set of *non-trivial* solutions (excluding $(0,\dots,0)$ if it's a zero) of $f=0$
$\overline{Z}(f)$	Number of non-trivial solutions: $\lvert\overline{\mathcal{V}}(f)\rvert$
$\mathcal{V}(f_1,\dots,f_k)$	Common solution set: $\mathcal{V}(f_1)\cap\cdots\cap\mathcal{V}(f_k)$
$Z(f_1,\dots,f_k)$	Count of common solutions: $\lvert\mathcal{V}(f_1,\dots,f_k)\rvert$
$Z_E(P=0)$	Number of solutions of $P(X_1,\dots,X_n)=0$ in E where $P\in\mathbb{F}_q[X_1,\dots,X_n]$, $[E\colon\mathbb{F}_q]>1$
Ω_d	Space $\mathbb{F}_q^{\le d}[X_1,\dots,X_n]$ of polynomials of total degree at most d
$\omega(d)$	Set of n-tuples $(i_1,\dots,i_n)\in\mathbb{Z}_0^n$ with $i_1+\cdots+i_n\le d$
Quadratic Forms	
C_f	Coefficient matrix of a quadratic form f: $(C_f)_{ij}=a_{ij}$ where $f=X^\top C_f X$
$\det(f)$	Determinant of the coefficient matrix C_f
$\vartheta:\mathbb{F}_q\to\mathbb{Z}$	Auxiliary function: $\vartheta(0)=q-1$ and $\vartheta(\alpha)=-1$ for $\alpha\neq 0$
Diagonal Equations	
$d_i=\gcd(k_i,q-1)$	Exponent reduced modulo the order of $\mathbb{F}_q^*$, for $a_iX_i^{k_i}$ in a diagonal equation
$M(d_1,\dots,d_n)$	$\#\left\{(t_1,\dots,t_n)\in\prod_i[d_i-1]:\frac{t_1}{d_1}+\cdots+\frac{t_n}{d_n}\in\mathbb{Z}\right\}$
$D=\mathrm{lcm}(d_1,\dots,d_n)$	LCM of exponents; appears in the closed-form formula for $M(d_1,\dots,d_n)$
Character Sums with Polynomials (Weil Sums)	

continued on next page

(*continued from previous page*)

Symbol	Meaning
$W(\chi; f)$	Additive Weil sum: $\sum_{\alpha\in\mathbb{F}_q} \chi(f(\alpha))$, for $\chi \in \boldsymbol{\mathcal{X}}_q$, $f \in \mathbb{F}_q[X]$
$W(\psi; f)$	Multiplicative Weil sum: $\sum_{\alpha\in\mathbb{F}_q} \psi(f(\alpha))$, for $\psi \in \boldsymbol{\mathcal{M}}_q$, $f \in \mathbb{F}_q[X]$
$W(\psi, \chi; f, g)$	Mixed Weil sum: $\sum_{\alpha\in\mathbb{F}_q} \psi(f(\alpha))\,\chi(g(\alpha))$, for $\psi \in \boldsymbol{\mathcal{M}}_q$, $\chi \in \boldsymbol{\mathcal{X}}_q$, $f, g \in \mathbb{F}_q[X]$
Lifted Weil Sums (over Extension Fields)	
$E = \mathbb{F}_{q^k}$	Degree-k extension of $\mathbb{F}_q$; $[E : \mathbb{F}_q] = k$
$W^{(k)}(\psi^{(k)}; f)$	Lifted multiplicative Weil sum: $\sum_{\alpha\in E} \psi^{(k)}(f(\alpha))$
$W^{(k)}(\chi^{(k)}; g)$	Lifted additive Weil sum: $\sum_{\alpha\in E} \chi^{(k)}(g(\alpha))$
$W^{(k)}(\psi^{(k)}, \chi^{(k)}; f, g)$	Lifted mixed Weil sum: $\sum_{\alpha\in E} \psi^{(k)}(f(\alpha))\,\chi^{(k)}(g(\alpha))$
Special L-functions (Weil chapter)	
$\overline{G}$ (Weil context)	The specific subgroup $\{r \in G : h_1(\gamma_i)\cdot h_2(\gamma_i) \neq 0\ \forall i\}$ for roots γ_i of a fixed f
$\varsigma : \overline{G} \to \mathbb{F}_q$	Multiplicative map $\varsigma(r) = \prod_i f(\alpha_i)\cdot\prod_j f(\beta_j)^{-1}$ for $r = \prod(X-\alpha_i)/\prod(X-\beta_j)$
$\varrho : G \to \mathbb{F}_q$	Additive map $\varrho(r) = \sum_i g(\alpha_i) - \sum_j g(\beta_j)$ for a fixed polynomial g
$\xi : G \to \mathbb{C}$	Combined character: $\xi(r) = (\psi\circ\varsigma)(r)\cdot(\chi\circ\varrho)(r)$
$L(s, \xi)$	L-function on ξ: $\sum_{h\in\Phi} \xi(h)\,\mathfrak{n}(h)^{-s}$
$\overline{L}(z, \xi)$	Power series: $\sum_{h\in\Phi} \xi(h)\,z^{\deg(h)}$, related by $\overline{L}(q^{-s}, \xi) = L(s, \xi)$
$L^{(k)}(s, \xi^{(k)})$	L-function on the lifted character $\xi^{(k)}$ over $\mathbb{F}_{q^k}$
$\overline{L}^{(k)}(z, \xi^{(k)})$	Power series form of the lifted L-function
$\Phi^{(k)}$	Set of monic polynomials over $E = \mathbb{F}_{q^k}$
$\mathfrak{n}^{(k)}(h) = q^{k\cdot\deg(h)}$	Norm of $h \in \Phi^{(k)}$ over E; analog of $\mathfrak{n}(h) = q^{\deg(h)}$ for $\mathbb{F}_q$
Absolute Varieties	
$\mathcal{A}$	Variety defined over $\mathbb{F}_q$
$Z_k(\mathcal{A})$	Number of $\mathbb{F}_{q^k}$-rational points of $\mathcal{A}$: $\lvert\{(\alpha_1,\ldots,\alpha_n) \in \mathcal{A} : \alpha_i \in \mathbb{F}_{q^k}\}\rvert$
Hasse Derivatives	
$F^{(\ell)}(X)$	ℓ-th Hasse derivative of $F(X)$: coefficient of Z^ℓ in $F(X+Z)$
$\mathrm{mult}(F, \alpha)$	Multiplicity of vanishing of F at α: largest r with $(X-\alpha)^r \mid F(X)$
Pseudopolynomials	
$\Lambda(X)$	The polynomial $X^q - X$ over $\mathbb{F}_q$
$F_{\langle\ell\rangle}(X)$	ℓ-th pseudoderivative of F: unique polynomial of degree $< q$ agreeing with $F^{(\ell)}$ on $\mathbb{F}_q$
$\mathrm{pdeg}(F)$	Pseudodegree of F: $\max_{\ell\geq 0} \deg(F_{\langle\ell\rangle})$
$F_i(X)$	Coefficient polynomials in the base-Λ expansion $F = \sum_i F_i\Lambda^i$, with $\deg(F_i) < q$
Decoding from Character Evaluations	
$r : \mathbb{F}_q \to \{0, \pm 1\}$	Received (noisy) word; approximates $\psi\circ g$ or $\mathrm{Tr}\circ g$

continued on next page

(*continued from previous page*)

Symbol	**Meaning**
e	Number of errors: $\lvert\{\alpha \in \mathbb{F}_q : r(\alpha) \neq (\psi \circ g)(\alpha)\}\rvert$
S	Error set: $\{\alpha \in \mathbb{F}_q : r(\alpha) \neq (\psi \circ g)(\alpha)\}$
$Z_S(X)$	Error locator polynomial: $\prod_{\alpha \in S}(X - \alpha)$
$\Delta(r_1, r_2)$	Hamming distance: $\lvert\{\alpha \in \mathbb{F}_q : r_1(\alpha) \neq r_2(\alpha)\}\rvert$

Contents

Chapter 1

Finite Fields

Fields are the most fundamental algebraic structures in which one can add, subtract, multiply, and divide (by nonzero elements) without leaving the set. Commonly known examples of fields are set of rationals $\mathbb{Q}$, or real numbers $\mathbb{R}$, or complex numbers $\mathbb{C}$. But the set of integers $\mathbb{Z}$ form a ring but not a field, since most elements have no multiplicative inverse. A *finite field* is a field that contains only finitely many elements. We assume the reader is familiar with basic ring theory; for a thorough treatment see [Art11]. Before proceeding, we recall the definitions of rings and fields.

Definition 1.1: Rings and Fields

(i) *A ring R is an abelian group with a multiplication operation such that the ring is closed under the multiplication operation. We assume the ring has a multiplicative identity $e \in R$ such that for all element $a \in R$, $ea = ae = a$.*

(ii) *A ring is called* commutative *if the multiplication operation is commutative.*

(iii) *A ring is called an* integral domain *if it is a commutative ring with identity $e \neq 0$ in which $ab = 0$ implies $a = 0$ or $b = 0$.*

(iv) *A ring is called a* division ring *(or* skew field*) if the nonzero elements of R form a group under the multiplication operation.*

(v) *A commutative division ring is called a* field.

For any ring R with identity, the *characteristic*, $\operatorname{char}(R)$ is the smallest positive integer n such that $n \cdot 1 = 0$, or 0 if no such n exists. One can show that the characteristic of an integral domain is always either 0 or a prime; see [Art11] for a proof. Since every field is an integral domain, the characteristic of a field is always 0 or a prime p.

§ 1.1 Field Extensions

The study of field theory is primarily the study of field extensions. The classical problems of ruler-and-compass constructions and the solvability of polynomial equations by radicals were settled by analyzing appropriate field extensions. The central idea is to associate a the *Galois group* to a field extension, converting field-theoretic questions into group-theoretic ones. Since the Galois group of a finite extension is finite, the rich structure theory of finite groups becomes available. In this section we build the foundational vocabulary: degrees, algebraic elements, minimal polynomials, and the key structural results about simple and algebraic extensions.

1.1.1 Degrees and Generated Subfields

Let $F \subseteq K$ are fields. Then K is called a *field extension* of F. We denote it by K/F. Here F is called the *base field*. We can think of K as a vector space over F where the addition of two vectors follows the addition operation of K and the multiplication of any $\alpha \in K$ and $c \in F$ follows the multiplication operation of K, $c \cdot \alpha = c\alpha$. Then we denote the dimension of K over F by

$$[K : F] := \dim_F(K)$$

This dimension is called the *degree of field extension* of F or simply *degree of* $[K : F]$. If $[K : F] < \infty$ we call it by a special name, *finite extension*. Finite extensions of finite fields are very important to us and in this report we will work on finite fields and their finite extensions. Let $\alpha \in K$. Then we denote $F(\alpha)$ to be the field generated by F and α following the multiplication and addition operations of K. We will define what it means to be the field generated by F and α in Definition 1.1.1. It is very easy to see $F(\alpha)$ is also a field, and it contains F but also contained in K.

Let F be any field and suppose X be a formal variable. Then $F[X]$ denotes the polynomial ring over F. We define the rational function field $F(X)$ to be the *fraction ring* of $F[X]$ which consists of all quotients f/g where $f, g \in F[X]$ with $g(X) \not\equiv 0$. Similarly, we can define the polynomial ring $F[X_1, \ldots, X_n]$ over multiple formal variables and their fraction field $F(X_1, \ldots, X_n)$.

Look at the following example. $\mathbb{R} \subseteq \mathbb{C}$ are both fields. But $\mathbb{C}$ is a 2-dimensional vector space over $\mathbb{R}$ as we can take the basis $\{1, i\}$. So $\mathbb{C}/\mathbb{R}$ is a finite extension of degree 2. But $\mathbb{Q}/\mathbb{R}$ is an infinite field extension. On the other hand let $\alpha \in \mathbb{R}$. Then consider the field $\mathbb{Q}(\alpha)$. Now degree of $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ depends on what kind of element α is. If $\alpha = \sqrt{2}$ then $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$ but if $\alpha = \pi$ then $[\mathbb{Q}(\pi) : \mathbb{Q}] = \infty$.

Definition 1.1.1: Generated Subfields

Let K/F be a field extension and $S \subseteq K$.

(i) The ring generated by F and S, *written $F[S]$, is the intersection of all subrings of K containing F and S.*

(ii) The field generated by F and S, *written $F(S)$, is the intersection of all subfields of K containing F and S.*

If $S = \{a_1, \ldots, a_n\}$ is finite we write $F[a_1, \ldots, a_n]$ and $F(a_1, \ldots, a_n)$. We call $F(\alpha_1, \ldots, \alpha_n)$ finitely generated *field. For any field extension K/F if there exists an element $\alpha \in K$ such that $K = F(\alpha)$ then such extension is called* simple extension.

From this definition it easily follows that for any $S \subseteq K$, $F(S)$ is the smallest subfield of K containing F and α.

Let K/F be a field extension. For any $\alpha \in K$ consider the map $\mathrm{ev}_\alpha : F[X] \to K$ defined by $f(X) \mapsto f(\alpha)$. Hence, ev_α is a ring homomorphism.

Observation 1.1. *For any $\alpha \in K$, $\mathrm{Im}(\mathrm{ev}_\alpha)$ is a subring of K and ev_α is F-vector space homomorphism.*

Now we can use this map ev_α to give descriptions of the elements of $F[\alpha]$ and $F(\alpha)$.

Lemma 1.1.1

Let K be a field extension of F and let $a \in K$. Then

$$F[a] = \{f(a) : f(x) \in F[x]\} \quad \text{and} \quad F(a) = \left\{ \frac{f(a)}{g(a)} : f, g \in F[x],\ g(a) \neq 0 \right\}.$$

Moreover, $F(a)$ is the quotient field of $F[a]$.

Proof: The evaluation map ev_a is a ring homomorphism with image $\{f(a) : f \in F[x]\}$. So this set is a subring of K. Any subring of K containing F and a must contain $f(a)$, for all $f \in F[X]$. Therefore, $F[a] = \{f(a) : f \in F[x]\}$. The fraction field of $F[a]$ is $\{f(a)/g(a) : g(a) \neq 0\}$; it is contained in any subfield of K containing $F[a]$, hence equals $F(a)$. ■

We call this procedure *adjoining* α to F. Using similar arguments we can describe the ring $F[a_1,\dots,a_n]$ and the field $F(a_1,\dots,a_n)$ where K/F is a field extension and $a_1,\dots,a_n \in K$. So we have the following theorem.

Theorem 1.1.2

Let K be a field extension of F and let $a_1,\dots,a_n \in K$. Then

$$F[a_1,\dots,a_n] = \{f(a_1,\dots,a_n) : f \in F[x_1,\dots,x_n]\},$$
$$F(a_1,\dots,a_n) = \left\{\frac{f(a_1,\dots,a_n)}{g(a_1,\dots,a_n)} : f,g \in F[x_1,\dots,x_n],\ g(a_1,\dots,a_n) \neq 0\right\}.$$

Moreover, $F(a_1,\dots,a_n)$ is the quotient field of $F[a_1,\dots,a_n]$.

Now we can also define fields generated by adjoining arbitrary number of elements. In this case we can give the description of the new field in terms of all finite extensions of fields created by adjoining finitely many elements.

Theorem 1.1.3

Let K be a field extension of F and let $S \subseteq K$ be any subset. If $\alpha \in F(S)$, then $\alpha \in F(a_1,\dots,a_n)$ for some $a_1,\dots,a_n \in S$. Therefore,

$$F(S) = \bigcup \{F(a_1,\dots,a_n) : a_1,\dots,a_n \in S\},$$

where the union is over all finite subsets of S.

Proof: Let $U := \cup\{F(a_1,\dots,a_n) : a_1,\dots,a_n \in S\}$. Each $F(a_1,\dots,a_n)$ contains F and lies in $F(S)$ where $a_1,\dots,a_n \in S$. So the union U satisfies $U \subseteq F(S)$. The set U contains F and S. If $\alpha \in F(a_1,\dots,a_n)$ and $\beta \in F(b_1,\dots,b_m)$, then $\alpha \pm \beta$, $\alpha\beta$, α/β (if $\beta \neq 0$) are all in $F(a_1,\dots,a_n,b_1,\dots,b_m) \subseteq U$. Therefore, U is a field. Being the smallest subfield containing F and S, it equals $F(S)$. ∎

1.1.2 Algebraic Elements and Minimal Polynomials

Definition 1.1.2: Algebraic and Transcendental Elements

Let K be a field extension of F. An element $\alpha \in K$ is algebraic over F *if there is a nonzero polynomial $f(X) \in F[X]$ with $f(\alpha) = 0$. If α is not algebraic over F, then α is called* transcendental over F. *If every element of K is algebraic over F, then K is said to be* algebraic over F, *and K/F is called an* algebraic extension.

Hermite (1873) proved that e is transcendental over $\mathbb{Q}$ and Lindemann (1882) proved the same for π. On the other hand, $\sqrt[n]{r} \in \mathbb{R}$ is algebraic over $\mathbb{Q}$ for any $r \in \mathbb{Q}$ (root of $x^n - r$), as is $i = \sqrt{-1}$ (root of $x^2 + 1$).

Definition 1.1.3: Minimal Polynomial

If α is algebraic over a field F, the minimal polynomial *of α over F is the monic polynomial $p(X)$ of least degree in $F[X]$ for which $p(\alpha) = 0$; it is denoted* $\min(F,\alpha)$. *Equivalently,* $\min(F,\alpha)$ *is the monic generator of the kernel of the evaluation homomorphism* ev_α.

The minimal polynomial depends on the base field. For instance, $\min(\mathbb{Q}, i) = x^2 + 1$ while $\min(\mathbb{C}, i) = x - i$, since $i \in \mathbb{C}$ is already in the base field.

The minimal polynomial of an element and the degree of a field extension are two of the most basic tools we shall use. The following proposition gives a relation between these objects.

Theorem 1.1.4

Let K/F be a field extension and let $\alpha \in K$ be algebraic over F. If $p(X) = \min(F, \alpha)$, then:

(i) *$p(X)$ is irreducible over F.*

(ii) *If $g(X) \in F[X]$, then $g(\alpha) = 0$ if and only if $p(X) \mid g(X)$.*

(iii) *If $n = \deg(\min(F, \alpha))$, then the elements $1, \alpha, \ldots, \alpha^{n-1}$ form a basis for $F(\alpha)$ over F. So $[F(\alpha) : F] = \deg(\min(F, \alpha)) < \infty$. Moreover, $F(\alpha) = F[\alpha]$.*

(iv) *$F(\alpha) \cong F[X] / \langle p(X) \rangle$.*

Proof: The *first isomorphism theorem* gives $F[\alpha] \cong F[X] / \langle p(X) \rangle$. Since $F[\alpha] \subseteq K$ is an integral domain, $\langle p(X) \rangle$ is a prime ideal. Now in an PID every nonzero prime is maximal, so $F[X] / \langle p(X) \rangle$ is a field and $F[\alpha] = F(\alpha)$. This gives (i) and (iv). Part (ii) follows from $\ker(\mathrm{ev}_\alpha) = \langle p(X) \rangle$.

For (iii): any $b \in F(\alpha) = F[\alpha]$ satisfies $b = g(\alpha)$. Dividing gives $g = qp + r$ with $\deg r < n$ and $q, r \in F[X]$. So $b = r(\alpha)$ lies in the span of $\{1, \alpha, \ldots, \alpha^{n-1}\}$. Therefore, the set $\{1, \alpha, \ldots, \alpha^{n-1}\}$ spans $F(\alpha)$ as a vector space. If $a_0 + a_1 \cdot \alpha + \cdots + a_{n-1} \cdot \alpha^{n-1} = 0$ where $a_0, \ldots, a_{n-1} \in F$ then the polynomial $f(X) = \sum_{i=0}^{n-1} a_i X^i$ is in $\ker(\mathrm{ev}_\alpha)$. Hence, either $f \equiv 0$ or $p \mid f$ by part (ii). But $\deg(f) < n$. So $f \equiv 0$. Thus, $1, \alpha, \ldots, \alpha^{n-1}$ forms a basis of $F(\alpha)$ over F. ∎

The part (iii) of above theorem gives us a basis of $F(\alpha)$ over F using only α. So we get the following observation:

Observation 1.2. *If K/F is a field extension and $\alpha \in K$ is algebraic over F where $p = \min(F, \alpha)$ then $[F(\alpha) : F] = \deg(p)$.*

Notice that if $\alpha, \beta \in K$ are both algebraic over F and $\min(F, \alpha) = \min(F, \beta) = p$ i.e., their minimal polynomials are same then by the above theorem

$$F(\alpha) \cong F[X] \Big/ \langle p \rangle \cong F(\beta)$$

Therefore we have the observation:

Observation 1.3. *If K/F is a field extension and $\alpha, \beta \in K$ are both algebraic over F, and they have the same minimal polynomials then $F(\alpha) \cong F(\beta)$ with the isomorphism $\varphi : \alpha \mapsto \beta$.*

1.1.3 Finite and Algebraic Extensions

Lemma 1.1.5

If $[K : F] < \infty$, then K is algebraic and finitely generated over F.

Proof: Let $\{\alpha_1, \ldots, \alpha_n\}$ be an F-basis of K. Then for any element $\alpha \in K$, there exists $a_1, \ldots, a_n$ such that $\alpha = a_1 \cdot \alpha_1 + \cdots + a_n \cdot \alpha_n$. So certainly we have $K = F(\alpha_1, \ldots, \alpha_n)$ and hence K is finitely generated over F. Since $[K : F] = n$, for any $\alpha \in K$, the $n + 1$ elements $1, \alpha, \ldots, \alpha^n$ span only an n-dimensional space. So they are F-linearly dependent. Thus, there exists $\beta_0, \ldots, \beta_n \in F$ such that $\sum_{i=0}^{n} \beta_i \cdot \alpha^i = 0$. So we get a polynomial $f(X) = \sum_{i=0}^{n} \beta_i \cdot X^i$ and $f(\alpha) = 0$. Therefore, α is algebraic over F. ∎

Let K/F be a finite extension and $K = F(\alpha_1, \ldots, \alpha_n)$. Then we can break up the extension K/F into collection of subextensions. Let $L_i = F(\alpha_1, \ldots, \alpha_i)$. Set $L_0 = F$. Then basically $L_{i+1} = L_i(\alpha_{i+1})$. And we have

$$F = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_{n-1} \subseteq L_n = K$$

We will show a theorem which relates the degree of the extension K/F and each subextensions.

Theorem 1.1.6

Let $F \subseteq L \subseteq K$ be fields. Then $[K : F] = [K : L] \cdot [L : F]$.

Proof: Let $\{a_i\}_{i\in I}$ be an F-basis of L and $\{b_j\}_{j\in J}$ an L-basis of K. We show $\{a_i \cdot b_j\}_{I\times J}$ is an F-basis of K. For any $\gamma \in K$, write $\gamma = \sum_j \alpha_j b_j$ with $\alpha_j \in L$. Now each $\alpha_j = \sum_i \beta_{ij} a_i$ with $\beta_{ij} \in F$; then $\gamma = \sum_{i,j} \beta_{ij} a_i b_j$. Therefore, the set $\{a_i \cdot b_j\}_{I\times J}$ spans K over F.

If $\sum_{i,j} \beta_{ij} a_i b_j = 0$ for some non-trivial linear combinations where $\beta_{i,j} \in F$ for all $i \in I$, $j \in J$. Now

$$\sum_{i,j} \beta_{ij} a_i b_j = \sum_j \Big(\sum_i \beta_{ij} a_i\Big) b_j = 0.$$

Since $\{b_j\}_{j\in J}$ is a basis of K over L, each $\sum_i \beta_{ij} a_i = 0$. And by F-independence of $\{a_i\}$, each $\beta_{ij} = 0$. So $\{a_i \cdot b_j\}_{I\times J}$ form an F-basis of K. Therefore,

$$[K : F] = |\{a_i \cdot b_j\}_{I\times J}| = |\{a_i : i \in I\}| \cdot |\{b_j : j \in J\}| = [K : L] \cdot [L : F]$$

So we have the theorem. ∎

Lemma 1.1.7

Let $K = F(X)$ be the field of rational functions in X over F. If $u \in K$ with $u \notin F$ let $u = f/g$ where $f, g \in F[X]$ and $\gcd(f, g) = 1$. Let $L = F(u)$. Then

$$[K : L] = \max\{\deg(f), \deg(g)\}$$

Proof: By Theorem 1.1.4 we need to find the minimal polynomial of X over L to determine $[K : L]$. Consider the polynomial $p(Y) = u \cdot g(Y) - f(Y) \in L[Y]$. Then X is a root of p. Therefore, X is algebraic over L and so $[K : L] < \infty$ and $K = L(X)$.

Notice that $\deg(p) = \max\{\deg(f), \deg(g)\}$. Let $f = \sum_{i=0}^{n} a_i X^i$ and $g = \sum_{j=0}^{m} b_i \cdot X^i$. We will show that p is irreducible over L. Now the element u cannot be algebraic over k otherwise by Theorem 1.1.6 we have $[K : F] = [K : L] \cdot [L : F] < \infty$, and we know $F(X)$ is an infinite extension of F. So u is transcendental over F and hence $F[u] \cong F[X]$. Now observe that the polynomial $p \in F[Y][u] \subseteq F(Y)[u]$ which is of degree 1. Hence, p is irreducible over $k(Y)$. Now $p \in k[Y][u] = k[u][Y]$ and therefore p is irreducible over $k(u)$. Hence, p is the minimal polynomial of X over F. So we have the lemma. ∎

Note:-

In the proof of Theorem 1.1.6 the proof does not assume that the field extensions needs to be finite.

We will now prove a converse theorem of Lemma 1.1.5.

Theorem 1.1.8

Let K/F be a field extension. If $\alpha_1, \ldots, \alpha_n \in K$ are algebraic over F, then

$$[F(\alpha_1, \ldots, \alpha_n) : F] \leq \prod_{i=1}^{n} [F(\alpha_i) : F].$$

Proof: We will prove this by induction on n. The base case, $n = 1$ follows from the Theorem 1.1.6. So suppose this is true for $n - 1$. Take $L = F(\alpha_1, \ldots, \alpha_{n-1})$. Then by induction hypothesis $[L : F] \leq \prod_{i=1}^{n-1} [F(\alpha_i) : F]$. Since α_n is

algebraic over F and hence over L, with $\min(L, \alpha_n) \mid \min(F, \alpha_n)$ in $L[X]$, we have $[L(\alpha_n) : L] \leq [F(\alpha_n) : F]$. Therefore, $[F(\alpha_1, \ldots, \alpha_n) : L] \leq [F(\alpha_n) : F]$. So by Theorem 1.1.6 we have

$$[F(\alpha_1, \ldots, \alpha_n) : F] = [L(\alpha_n) : L] \cdot [L : F] \leq [F(\alpha_n) : F] \cdot \prod_{i=1}^{n-1} [F(\alpha_i) : F].$$

So we have the theorem. ∎

The inequality above can be strict in some cases. For example let $a = \sqrt[4]{2}$ and $b = \sqrt[4]{18}$. Then both $[\mathbb{Q}(a) : \mathbb{Q}] = [\mathbb{Q}(b) : \mathbb{Q}] = 4$ as $X^4 - 2$ and $X^4 - 18$ are both irreducible polynomials over $\mathbb{Q}$. But $[\mathbb{Q}(\sqrt[4]{2}, \sqrt{3}) : \mathbb{Q}] = 8$ and $\sqrt[4]{18} \in \mathbb{Q}(\sqrt[4]{2}, \sqrt{3})$. Hence, in this the inequality doesn't follow.

So from Theorem 1.1.8 we have the following criterion for an element being algebraic over a field.

Corollary 1.1.9

Let K/F be a field extension and $\alpha \in K$. Then α is algebraic over F if and only if $[F(\alpha) : F] < \infty$. Moreover, K is algebraic over F if $[K : F] < \infty$.

Now we can extend Theorem 1.1.8 to the case of fields generated by arbitrary number of elements or extending field.

Theorem 1.1.10

Let K be a field extension of F and $S \subseteq K$ such that each element of S is algebraic over F. Then $F(S)/F$ is algebraic. If $|S| < \infty$, then $[F(S) : F] < \infty$.

Proof: For any $\alpha \in F(S)$. Theorem 1.1.3 gives $\alpha \in F(a_1, \ldots, a_m)$ for some $a_1, \ldots, a_m \in S$. By Theorem 1.1.8, $F(a_1, \ldots, a_m)$ is algebraic over F. So α is algebraic over F by Theorem 1.1.5. Since α is any arbitrary element of $F(S)$, $F(S)$ is algebraic over F. If S is finite, then $[F(S) : F] < \infty$ by Theorem 1.1.8. ∎

We will now show that the property of being algebraic is a transitive property. In the finite extensions case this property easily follows from Theorem 1.1.6 and Corollary 1.1.9. We will show this for general extensions below.

Theorem 1.1.11 Transitivity of Algebraic Extensions

Let $F \subseteq L \subseteq K$ be fields. If L/F and K/L are algebraic, then K/F is algebraic.

Proof: Let $\alpha \in K$. Since K/L is algebraic, let $f(X) = \sum_{i=0}^{d} a_i X^i$ where each $a_i \in L$ is the minimal polynomial of α over L. Since L/F is algebraic and $f \in L[X]$, each a_i is algebraic over F. Consider the field $L' = F(a_0, \ldots, a_d)$. So by Theorem 1.1.8 the extension L'/F is finite i.e., $[L' : F] < \infty$. Since $f \in L'[X]$ and $p(\alpha) = 0$, the element α is algebraic over L' and $[L'(\alpha) : L'] < \infty$. By the tower law,

$$[L'(\alpha) : F] = [L'(\alpha) : L'] \cdot [L' : F] < \infty,$$

Since $F(\alpha) \subseteq L'(\alpha)$, α is algebraic over F. Since α is any arbitrary element of K, K/F is algebraic. ∎

Definition 1.1.4: Algebraic Closure in K

Let K be a field extension of F. The set

$$\{\, a \in K : a \text{ is algebraic over } F \,\}$$

is called the algebraic closure of F in K.

Corollary 1.1.12

Let K be a field extension of F and let L be the algebraic closure of F in K. Then L is a field, and therefore is the largest algebraic extension of F contained in K.

Proof: Let $a, b \in L$. Then both are algebraic over F by Theorem 1.1.8 and hence $F(a, b)$ is algebraic over F and $F(a, b) \subseteq L$. Since $a \pm b, a \cdot b, a/b$ (if $b \neq 0$) all are in $F(a, b) \subseteq L$, L is closed under the field operation. So L is a subfield of K. Since each element of K which is algebraic over F is contained in L, L is the largest algebraic extension of F contained in K. ■

1.1.4 Splitting Field

Definition 1.1.5: Splits over K

If K is an extension field of F and if $f(X) \in F[X]$, then f splits over K *if $f(X) = a\prod_i (X - \alpha_i) \in K[X]$ for some $\alpha_1, \ldots, \alpha_n \in K$ and $a \in F$. In other words, f splits over K if f factors completely into linear factors in $K[X]$.*

In order to talk about roots of a given polynomial, we need to have extension fields that contain the roots of the polynomial. The next theorem shows that for any $f \in F[X]$, there is a finite extension of F over which f splits.

Theorem 1.1.13

Let $f(x) \in F[x]$ have degree n. There is an extension field K of F with $[K : F] \leq n$ such that K contains a root of f. In addition, there is a field L containing F with $[L : F] \leq n!$ such that f splits over L.

Proof: Let $p(X)$ is an irreducible factor of $f(X)$ in $F[X]$ and let $K = F[X] / \langle p(X) \rangle$. Thus, K contains a root of $p(X)$ say α. Therefore, α is also a root of $f(X)$. Since $[K\colon F] = \deg(p) \leq n$ we have the first part.

For the second part we will use induction on n. From the first part now we have a field $K \supseteq F$ such that K contains a root α of f. So f factors into $f(X) = (X - \alpha)g(X)$ in $K[X]$ where $g(X) \in K[X]$. Now $\deg(g) = n - 1$. So by induction hypothesis there exists a finite extension L of K such that $[L\colon K] \leq (n-1)!$ and g splits over L. That means f splits over L and $[L\colon F] = [L\colon K] \cdot [K\colon F] \leq n!$. So we have the theorem. ■

Definition 1.1.6: Splitting Field

Let K be an extension of F and let $f \in F[X]$.

(i) K is called the splitting field of f over F *if and only if f splits over K and $K = F(\alpha_1, \ldots, \alpha_n)$ where $\alpha_1, \ldots, \alpha_n$ are exactly the roots of f.*

(ii) If S is a set of non-constant polynomials over F then K is the splitting field of S over F *if and only if each $f \in S$ splits over K and $K = F(S_R)$ where S_R is the set of all roots of all $f \in S$.*

So Theorem 1.1.13 yields immediately the existence of splitting fields for a finite set of polynomials. We have proved the existence of splitting fields for finite sets of polynomials. What about infinite sets? Suppose that K is a splitting field over F of the set of all non-constant polynomials over F. We do not know yet that such a field exists or not.

Lemma 1.1.14

If K is a field, then the following statements are equivalent:

(i) There are no algebraic extensions of K other than K itself.

(ii) There are no finite extensions of K other than K itself.

(iii) If L is a field extension of K, then $K = \{a \in L : a \text{ is algebraic over } K\}$.

(iv) Every $f(x) \in K[x]$ splits over K.

(v) Every $f(x) \in K[x]$ has a root in K.

(vi) Every irreducible polynomial over K has degree 1.

Proof: We will show the implications in a circular order, (i) $\implies$ (ii) $\implies$ (iii) $\implies$ (iv) $\implies$ (v) $\implies$ (vi) $\implies$ (i).

(i) $\implies$ (ii): This is clear since any finite extension is algebraic.

(ii) $\implies$ (iii): Let $a \in L$ be algebraic over K. Then $K(a)$ is a finite extension of K. But there are no finite extensions of K other than itself. So $K(a) = K$. Thus, $a \in K$.

(iii) $\implies$ (iv): Let $f \in K[X]$ and L be the splitting field of f over K. Since L is algebraic over K we have $L = K$ i.e., f splits over K.

(iv) $\implies$ (v): This is obviously true.

(v) $\implies$ (vi): Let $f \in K[X]$ be irreducible. Therefore, f has a root in K, so f has a linear factor. But f is irreducible. Hence, $\deg(f) = 1$.

(vi) $\implies$ (i): Let L be an algebraic extension of K. Let $a \in L$ and $p = \min(K, a)$. Then $\deg(p) = 1$. So $[K(a) : K] = 1$ and hence $a \in K$. So $L = K$. ∎

Definition 1.1.7: Algebraically Closed Field

If any field K satisfies the equivalent conditions of Lemma 1.1.14 then K is said to be algebraically closed. *If K is an algebraic extension of F then K is said to be* algebraic closure *of F.*

We haven't shown the existence of such a theorem or even that every field has an algebraic closure. Consider the field $\mathbb{C}$. This field is algebraically closed. But we will not prove if every field has an algebraic closure. We will just state the theorem and move on.

Theorem 1.1.15

Let F be a field. Then F has an algebraic closure.

The existence of an algebraic closure yields immediately the existence of a splitting field for an arbitrary set of non-constant polynomials.

Corollary 1.1.16

Let F be a field and S be a set of non-constant polynomials over F. Then S has a splitting field over F.

Proof: Let K be an algebraic closure of F. Then each $f(x) \in S$ splits over K. Let X be the set of roots of all $f \in S$. Then $F(X) \subseteq K$ is a splitting field for S over F, since each f splits over $F(X)$ and this field is generated by the roots of all the polynomials from S. ■

So using the corollary we can immediately relate algebraic closure with splitting fields.

Corollary 1.1.17

If F is a field then the splitting field of the set of all non-constant polynomials over F is an algebraic closure of F.

But now comes another question. Can one field have multiple algebraic closures of same field? Or even can we have more than one splitting fields for same set of polynomials over a field? The answer is splitting fields are unique, and hence algebraic closures are also unique. Below we share the theorem which proves the above. We will not prove it because its beyond our scope of study. But for further reading and understanding the reader can look [Mor96].

Theorem 1.1.18 Isomorphism Extension Theorem, [Mor96, Chapter 1]

Let $\sigma : F \to F'$ be a field isomorphism. Let $S = \{f_i(x)\}$ be a set of polynomials over F, and let $S' = \{\sigma(f_i)\}$ be the corresponding set over F'. Let K be a splitting field for S over F, and let K' be a splitting field for S' over F'. Then there is an isomorphism $\tau : K \to K'$ with $\tau|_F = \sigma$. Furthermore, if $\alpha \in K$ and α' is any root of $\sigma(\min(F, \alpha))$ in K', then τ can be chosen so that $\tau(\alpha) = \alpha'$.

The isomorphism extension theorem immediately settles the question of uniqueness of splitting fields and algebraic closures: any two splitting fields of the same set of polynomials over F must be isomorphic, since the identity on F extends to an isomorphism between them. As a special case, algebraic closures are unique up to F-isomorphism.

Corollary 1.1.19

Let F be a field, and let S be a subset of $F[x]$. Any two splitting fields of S over F are F-isomorphic. In particular, any two algebraic closures of F are F-isomorphic.

Proof: For the proof of the first statement, the isomorphism extension theorem gives an isomorphism extending id on F between any two splitting fields of S. The second statement follows from the first, since any algebraic closure of F is a splitting field of the set of all non-constant polynomials in $F[x]$. ■

As a corollary to the existence and uniqueness of algebraic closures, we can prove that any algebraic extension of a field F can be viewed as living inside a fixed algebraic closure of F.

Corollary 1.1.20

Let F be a field, and let N be an algebraic closure of F. If K is an algebraic extension of F, then K is isomorphic to a subfield of N.

Proof: Let M be an algebraic closure of K. By Theorem 1.1.11, M is algebraic over F; hence, M is also an algebraic closure of F. Therefore, by the previous corollary, $M \cong N$. If $f : M \to N$ is an F-isomorphism, then $f(K)$ is a subfield of N isomorphic to K. ∎

§ 1.2 Some Properties of Finite Fields

In this section we describe fields of finite order. The characteristic of a finite field cannot be zero so it is a prime number. Therefore, if K is a finite field and $\text{char}(K) = p$ then $\mathbb{F}_p \subseteq K$. Since K is finite it will be a finite dimensional vector space over $\mathbb{F}_p$. So K is isomorphic to $\mathbb{F}_p^r$ where $r = [K\colon \mathbb{F}_p]$ as a $\mathbb{F}_p$-vector space which contains p^r elements. So order of a finite field is always a prime power. We write $\mathbb{F}_q := K$ where $q = p^r$. Below we list the main facts of finite fields without proof. If the reader is interested he can look into the books [Art11, LN96].

Theorem 1.2.1 Properties of Finite Field

Let p be a prime integer, and let $q = p^r$ be a positive power of p.

(i) Let K be a field of order q. The elements of K are roots of the polynomial $X^q - X$.

(ii) The irreducible factors of the polynomial $X^q - X$ over the field $\mathbb{F}_p$ are the irreducible polynomials in $\mathbb{F}_p[X]$ whose degrees divide r.

(iii) Let K be a field of order q. The multiplicative group K^* of nonzero elements of K is a cyclic group of order $q - 1$.

(iv) There exists a field of order q, and all fields of order q are isomorphic.

(v) A field of order p^r contains a subfield of order p^k if and only if k divides r.

Observation 1.4. *Since for every $r \in \mathbb{N}$ there exists an element $\alpha \in \mathbb{F}_{q^r}$ such that $\mathbb{F}_{q^r}^*$ is generated by α then $\mathbb{F}_{q^r} = \mathbb{F}_q(\alpha)$. Hence, every finite extensions of finite fields are simple extensions.*

Since finite fields are finite extensions of the base prime field $\mathbb{F}_p$ we can obtain irreducible polynomials for every positive integer degree.

Corollary 1.2.2

For every $d \in \mathbb{N}$ there exists an irreducible polynomial of degree d over the finite field $\mathbb{F}_p$.

Proof: By part (iv) of the above theorem there is a field K of order p^d. Its degree $[K\colon \mathbb{F}_p] = d$. By part (iii) there is an element $\alpha \in K$ such that α generates the cyclic group K^*. So $K = \mathbb{F}_p(\alpha)$. Therefore, α is a root of a d-degree irreducible polynomial over $\mathbb{F}_p$. ∎

Definition 1.2.1: Primitive Element

Let K be finite field over $\mathbb{F}_p$. Then K has an element α which generates K/F i.e., $K = \mathbb{F}_p(\alpha)$. Then α is called the primitive element *of K.*

Let f be an irreducible polynomial over the finite field $\mathbb{F}_q$ and let α be a root of f in the splitting field of f over $\mathbb{F}_q$. Therefore, $f = \min(\mathbb{F}_q, \alpha)$. Now suppose $h \in \mathbb{F}_q[X]$ such that $h(\alpha) = 0$. So by Theorem 1.1.4, $f \mid h$. Like Theorem 1.2.1(ii) we will show that we can similar result for any general finite field.

Lemma 1.2.3

Let $f \in \mathbb{F}_q[X]$ be an irreducible polynomial of degree d. Then $f(X)$ divides $X^{q^n} - X$ if and only if $d \mid n$.

Proof: Suppose $f \mid X^{q^n} - X$. Let α be a root of f in the splitting field. Then $\alpha^{q^n} = \alpha$ and so $\alpha \in \mathbb{F}_{q^n}$. Now $\mathbb{F}_q \subseteq \mathbb{F}_q(\alpha) \subseteq \mathbb{F}_{q^n}$. Since $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = d$ and $[\mathbb{F}_{q^n} : \mathbb{F}_q] = n$. So by Theorem 1.1.6, $d \mid n$.

Let $d \mid n$. Then $\mathbb{F}_{q^d} \subseteq \mathbb{F}_{q^n}$. If α is a root of f in the splitting field of f over $\mathbb{F}_q$ then $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^d} \subseteq \mathbb{F}_{q^n}$. Therefore, α is a root of $X^{q^n} - X$. Since f is the minimal polynomial of α and $X^{q^n} - X \in \mathbb{F}_q$, $f \mid X^{q^n} - X$. ■

Theorem 1.2.4

If f is an irreducible polynomial of $\mathbb{F}_q[X]$ of degree d then f has a root $\alpha \in \mathbb{F}_{q^d}$. Furthermore, all roots of f are given by d distinct elements $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{d-1}}$ of $\mathbb{F}_{q^d}$.

Proof: Let $f(X) = \sum_{i=0}^{d} a_i X^i$. Then

$$f\left(\alpha^{q^j}\right) = \sum_{i=0}^{d} a_i \cdot \left(\alpha^{q^j}\right)^i = \sum_{i=0}^{d} a_i^{q^j} \cdot \left(\alpha^i\right)^{q^j} = \left(\sum_{i=0}^{d} a_i \cdot \alpha^i\right)^{q^j} = f^{q^j}(\alpha) = 0$$

Since $\alpha^{q^d} = \alpha$ the roots of f are $\alpha, \alpha^q, \ldots, \alpha^{q^{d-1}}$. ■

This leads us to look at roots and field automorphisms in a new way. Let F, F' be two fields and $\varphi : F \to F'$ a homomorphism. Then there is an induced homomorphism $\varphi : F[X] \to F'[X]$ given by

$$\varphi\left(\sum_{i=0}^{n} a_i \cdot X^i\right) = \sum_{i=0}^{n} \varphi(a_i) \cdot X^i$$

So if $f = \sum_{i=0}^{n} a_i X^i$ factors into $a(X - \alpha_1) \cdots (X - \alpha_n)$ then $\varphi(f) = \varphi(a)(X - \varphi(\alpha_1)) \cdots (X - \varphi(\alpha_n))$. The *Frobenius map*, $\sigma : \alpha \mapsto \alpha^q$ for any element of the extended field $\mathbb{F}_{q^d}$ over $\mathbb{F}_q$ is a $\mathbb{F}_q$-automorphism meaning that automorphism fixes the elements of $\mathbb{F}_q$. The above theorem tells us if α is a root of f which is an irreducible polynomial over $\mathbb{F}_q$ then $\sigma(\alpha)$ is also a root.

Now by the above theorem if f is irreducible over $\mathbb{F}_q$ of degree d then $\mathbb{F}_{q^d}$ contains a root α of f. Then $\alpha, \alpha^q, \ldots, \alpha^{q^{d-1}}$ all are in $\mathbb{F}_{q^d}$. Therefore, f splits in $\mathbb{F}_{q^d}$. Since f is the minimal polynomial of α, $\mathbb{F}_{q^d} = \mathbb{F}_q[X] / \langle f \rangle$. So $\mathbb{F}_{q^d}$ is the splitting field of f. Since all finite fields of size q^d are isomorphic for any other degree d irreducible polynomial $g \in \mathbb{F}_q[X]$, if β is a root of g then

$$\mathbb{F}_q(\alpha) \cong \mathbb{F}_q[X] \Big/ \langle f \rangle \cong \mathbb{F}_{q^d} \cong \mathbb{F}_q[X] \Big/ \langle g \rangle \cong \mathbb{F}_q(\beta)$$

So the splitting fields of f and g are isomorphic. So we have the following corollary.

Corollary 1.2.5

Let f is an irreducible polynomial of $\mathbb{F}_q[X]$ of degree d. Then the splitting field of f over $\mathbb{F}_q$ is given by $\mathbb{F}_{q^d}$. Moreover, any two irreducible polynomials $f, g \in \mathbb{F}_q[X]$ of same degree has isomorphic splitting fields.

Now we have heard the term conjugate of an element in complex numbers. The notion of conjugates is more general than just in complex numbers.

Definition 1.2.2

Let $\mathbb{F}_{q^d}$ be an extension of $\mathbb{F}_q$ and $\alpha \in \mathbb{F}_{q^d}$. Then the elements $\alpha, \alpha^q, \ldots, \alpha^{q^{d-1}}$ are called the conjugates *of α.*

The conjugates of $\alpha \in \mathbb{F}_{q^m}$ with respect to $\mathbb{F}_q$ are distinct if and only if the minimal polynomial of α over $\mathbb{F}_q$ has degree m. Otherwise, the degree d of this minimal polynomial is a proper divisor of m, and then the conjugates of α with respect to $\mathbb{F}_q$ are the distinct elements $\alpha, \alpha^q, \ldots, \alpha^{q^{d-1}}$, each repeated m/d times.

Lemma 1.2.6

For any $\alpha \in \mathbb{F}_q^*$ the conjugates with respect to any subfield of $\mathbb{F}_q$ have the same order in the group $\mathbb{F}_q^*$.

Proof: Let $\mathbb{F}_{q'}$ be a subfield of $\mathbb{F}_q$ where $q'^r = q$. Since $\mathbb{F}_q^*$ is a cyclic group of order $q-1$ and $\gcd(q'^i, q-1) = 1$ for all $0 \le i \le r-1$. Therefore, the order of the conjugates of α has the same order as α in the group $\mathbb{F}_q^*$. ■

§ 1.3 Trace and Norm

Let V be an n-dimensional vector space over K where K is a field. Let $\varphi : V \to V$ be any linear operator. For any basis of V we can generate the matrix of φ with respect to that basis. This is called the *matrix representation* of φ and denoted by $[\varphi]$.

We now turn to field extensions. For a finite extension of field L/K we associate each element α of L the K-linear transformation $m_\alpha : L \to L$ where $m_\alpha(x) = \alpha \cdot x$ for any $x \in L$. Now m_α is indeed a linear transformation from L to L as

$$m_\alpha(x+y) = \alpha(x+y) = \alpha \cdot x + \alpha \cdot y = m_\alpha(x) + m_\alpha(y) \qquad m_\alpha(c \cdot x) = \alpha(c \cdot x) = c(\alpha \cdot x) = c \cdot m_\alpha(x)$$

for all $x, y \in L$ and $c \in K$. Now by choosing a K-basis of L we can create a matrix representation for m_α which is denoted by $[m_\alpha]$. So for any $c \in K$ the matrix $[m_c]$ is cI_k where $[L : K] = n$.

Definition 1.3.1: Trace and Norm of α

For any $\alpha \in L$ where L/K be a finite extension the trace and norm of α from L to K are the trace and determinant of a matrix representation of the K-linear map m_α and denoted as

$$\mathrm{Tr}_{L/K}(\alpha) = \mathrm{Tr}([m_\alpha]) \in K, \qquad \mathrm{N}_{L/K}(\alpha) = \det([m_\alpha])$$

For any finite field $\mathbb{F}_q$ if $\mathbb{F}_p$ is the base prime field of $\mathbb{F}_q$ then for any $\alpha \in \mathbb{F}_q$ we denote the trace and norm of α from $\mathbb{F}_q$ to $\mathbb{F}_p$ by $\mathrm{Tr}_{\mathbb{F}_q}(\alpha)$ and $\mathrm{N}_{\mathbb{F}_q}(\alpha)$, and we call it the *absolute trace* and *absolute norm* of α respectively.

Observation 1.5. *For any $\alpha \in K$, where $[L : K] = n$, $\mathrm{Tr}_{L/K}(\alpha) = n \cdot \alpha$ and $\mathrm{N}_{L/K}(\alpha) = \alpha^n$. In particular $\mathrm{Tr}_{L/K}(1) = [L : K]$ and $\mathrm{N}_{L/K}(1) = 1$.*

Note:-

There is an alternate definition of trace and norm in terms of sum and product of field embeddings of L into an algebraic closure of K. But those definitions is hard to work with and also beyond our scope. So interested readers can look into [Lan02, Chapter 1]

Lemma 1.3.1

Let $\alpha, \beta \in L$.

(i) If $\alpha \neq \beta$ then $m_\alpha \neq m_\beta$.

(ii) As functions from $L \to L$

$$m_{\alpha+\beta} = m_\alpha + m_\beta \quad \text{and} \quad m_{\alpha\beta} = m_\alpha \circ m_\beta$$

and m_1 is the identity map $L \to L$.

Proof: Since $m_\alpha(1) = \alpha$ we can recover the element α uniquely from the mapping m_α. Therefore, the map, $\varphi : L \to \{m_\alpha : \alpha \in L\}$ where $\varphi(\alpha) = m_\alpha$ is a bijection. And hence if $\alpha, \beta \in L$ and $\alpha \neq \beta$ then $m_\alpha \neq m_\beta$.

Now for any $x \in L$ we have

$$m_{\alpha+\beta}(x) = (\alpha+\beta)x = \alpha \cdot x + \beta \cdot x = m_\alpha(x) + m_\beta(x)$$

and

$$m_\alpha \circ m_\beta(x) = m_\alpha(\beta \cdot x) = (\alpha\beta)x = m_{\alpha\beta}(x)$$

Since x is an arbitrary element of L we have $m_{\alpha+\beta} = m_\alpha + m_\beta$ and $m_{\alpha\beta} = m_\alpha \circ m_\beta$. Easily m_1 is the identity mapping since $1 \cdot \alpha = \alpha$ and therefore $m_1 \circ m_\alpha = m_\alpha$. ■

Theorem 1.3.2

The trace $\mathrm{Tr}_{L/K} : L \to K$ *is a* K*-linear and the norm* $\mathrm{N}_{L/K} : L \to K$ *is multiplicative. Moreover,* $\mathrm{N}_{L/K}(L^*) \subseteq K^*$.

Proof: By Lemma 1.3.1 for any $\alpha, \beta \in L$ we have $m_{\alpha+\beta} = m_\alpha + m_\beta$ and $m_{\alpha\beta} = m_\alpha \circ m_\beta$. Therefore, if we look into the matrix representation of m_α and m_β then we get

$$[m_{\alpha+\beta}] = [m_\alpha] + [m_\beta], \qquad [m_{\alpha\beta}] = [m_\alpha \circ m_\beta] = [m_\alpha] \cdot [m_\beta]$$

So

$$\mathrm{Tr}_{L/K}(\alpha+\beta) = \mathrm{Tr}([m_{\alpha+\beta}]) = \mathrm{Tr}([m_\alpha]) + \mathrm{Tr}([m_\beta]) = \mathrm{Tr}_{L/K}(\alpha) + \mathrm{Tr}_{L/K}(\beta)$$

and

$$\mathrm{N}_{L/K}(\alpha\beta) = \det([m_{\alpha\beta}]) = \det([m_\alpha]) \cdot \det([m_\beta]) = \mathrm{N}_{L/K}(\alpha) \cdot \mathrm{N}_{L/K}(\beta)$$

So $\mathrm{Tr}_{L/K}$ is additive and $\mathrm{N}_{L/K}$ is multiplicative.

Now to show $\mathrm{Tr}_{L/K}$ is K-linear we need to show for any $c \in K$ and $\alpha \in L$ we have $\mathrm{Tr}_{L/K}(c \cdot \alpha) = c \cdot \mathrm{Tr}_{L/K}(\alpha)$ which is true since

$$\mathrm{Tr}_{L/K}(c \cdot \alpha) = \mathrm{Tr}([m_{c\alpha}]) = \mathrm{Tr}(c[m_\alpha]) = c \cdot \mathrm{Tr}([m_\alpha]) = c \cdot \mathrm{Tr}_{L/K}(\alpha)$$

Now $\mathrm{N}_{L/K}(1) = 1$. So for any non-zero $\alpha \in L^*$, $\mathrm{N}_{L/K}(\alpha) \cdot \mathrm{N}_{L/K}(\alpha^{-1}) = 1$ and hence $\mathrm{N}_{L/K}(\alpha) \neq 0$. Therefore, $\mathrm{N}_{L/K}(L^*) \subseteq K^*$. ■

Now the elements $\mathrm{Tr}_{L/K}(\alpha)$ and $\mathrm{N}_{L/K}(\alpha)$ can be expressed in terms of the coefficients or the roots of the minimal polynomial of α over K. To explain this we need another polynomial in $K[X]$ which is very similar to the characteristic polynomial of any square matrix.

Definition 1.3.2: Characteristic Polynomial of $\alpha \in L$

For $\alpha \in L$ where L/K is a finite extension the characteristic polynomial *of α relative to the extension L/K is the characteristic polynomial of the matrix representation $[m_\alpha]$:*

$$P_{\alpha,L/K}(X) = \det(X \cdot I_n - [m_\alpha]) \in K[X]$$

where $n = [L : K]$. Therefore, $\deg(P_{\alpha,L/K}) = [L : K]$.

Note that here we defined the characteristic polynomial to be of the form $\det(X \cdot I_n - A)$ instead of $\det(A - X \cdot I_n)$ which we typically do in linear algebra. This is because we want our polynomials to monic.

If A is a square matrix then the trace of A and the determinant of A appears as signed coefficients in its characteristic polynomial.

$$\det(X \cdot I_n - A) = X^n - \mathrm{Tr}(A)X^{n-1} + \cdots + (-1)^n \det(A)$$

Therefore

$$P_{\alpha,L/K}(X) = X^n - \mathrm{Tr}_{L/K}(\alpha)X^{n-1} + \cdots + (-1)^n \mathrm{N}_{L/K}(\alpha) \tag{1.1}$$

Below we show a Cayley-Hamilton like theorem for characteristic polynomial of field elements.

Theorem 1.3.3

For every $\alpha \in L$, $P_{\alpha,L/K}(\alpha) = 0$.

Proof: By the consequence of Cayley-Hamilton theorem we have $P_{\alpha,L/K}([m_\alpha]) = O$. Now for any polynomial $f \in K[X]$ where $f(X) = X^n + \sum_{i=0}^{n-1} a_i X^i$ then

$$f([m_\alpha]) = [m_\alpha]^n + \sum_{i=0}^{n-1} a_i [m_\alpha]^i = [m_{\alpha^n}] + \sum_{i=0}^{n-1} [m_{a_i\alpha^i}] = m_{\alpha^n + a_{n-1}\alpha^{n-1} + \cdots + c_0} = m_{f(\alpha)}$$

Therefore $P_{\alpha,L/K}([m_\alpha]) = m_{P_{\alpha,L/K}(\alpha)}$. Since $P_{\alpha,L/K}([m_\alpha]) = O$ and by Lemma 1.3.1 all m_α maps are distinct we have $P_{\alpha,L/K}(\alpha) = 0$. ∎

Now for any $\alpha \in L$ let $f = \min(K, \alpha)$. Let $\deg(f) = d$. Hence, $[K(\alpha) : K] = d$. Therefore, $P_{\alpha,K(\alpha)/K}$ is a monic degree d polynomial over K. Since both f and $P_{\alpha,K(\alpha)/K}$ are monic and have same degree and shares a common root we have $f = P_{\alpha,K(\alpha)/K}$. So we have the following theorem.

Theorem 1.3.4

For any $\alpha \in L$, where L/K is a finite extension. Let $f = \min(K, \alpha)$ is the minimal polynomial of α in $K[X]$. If $L = K(\alpha)$ then $P_{\alpha,L/K}(X) = f(X)$.

More generally, $P_{\alpha,L/K}(X) = f^{n/d}$ where $d = \deg(f) = [K(\alpha) : K]$.

Proof: We already prove the first part in the discussion before the theorem. So we will show the second part here. Let $[L : K(\alpha)] = m$ and $\{\beta_1, \ldots, \beta_m\}$ is a basis of L over $K(\alpha)$. $1, \alpha, \ldots, \alpha^{d-1}$ is a basis of $K(\alpha)$ over K. Therefore, $\{\alpha^i\beta_j : 0 \le i \le d-1, j \in [n]\}$ is a basis of L over K. So for all $i \in \{0, 1, \ldots, d-1\}$ there exists $c_{i,j} \in \mathbb{F}_q$ such that $\alpha \cdot \alpha^i = \sum_{j=0}^{d-1} c_{ij}\alpha^j$. Therefore for any $t \in [n]$ and $i \in \{0, 1, \ldots, d-1\}$, $\alpha \cdot \alpha^i\beta_t = \sum_{j=0}^{d-1} c_{ij}\alpha^j \cdot \beta_t$. Hence, the matrix $[m_\alpha]$ is a repeated diagonal $d \times d$ block matrix of m_α for $K(\alpha)$ over K created by c_{ij}. Therefore, $P_{\alpha,L/K} = (P_{\alpha,K(\alpha)/K})^{n/d}$. ∎

Corollary 1.3.5

Let the minimal polynomial of α over K be $X^d + c_{d-1}X^{d-1} + \cdots c_1X + c_0$. Then

$$\mathrm{Tr}_{K(\alpha)/K}(\alpha) = -c_{d-1} \quad \textit{and} \quad \mathrm{N}_{K(\alpha)/k}(\alpha) = (-1)^d c_0$$

and more generally when $K \subseteq K(\alpha) \subseteq L$ and $[L : K] = n$,

$$\mathrm{Tr}_{L/K}(\alpha) = -\frac{n}{d}c_{d-1} = -[L : K(\alpha)] \cdot c_{d-1}, \quad \textit{and} \quad \mathrm{N}_{L/K}(\alpha) = (-1)^n \cdot c_0^{n/d} = (-1)^n \cdot c_0^{[L:K(\alpha)]}$$

Proof: Let f is the minimal polynomial of α over K. Hence by Theorem 1.3.4, $P_{\alpha,K(\alpha)/K} = f(X)$. So we have the first part. Now

$$P_{\alpha,L/K}(X) = f^{n/d}(X) = \left(X^d + c_{d-1}X^{d-1} + \cdots + c_1X + c_0\right)^{n/d} = X^n + \frac{n}{d}c_{d-1}X^{n-1} + \cdots + c_o^{n/d}$$

Therefore by comparing with (1.1) we have the result. ∎

Observation 1.6. *We can rewrite the above formulas for* $\mathrm{Tr}_{L/K}(\alpha)$ *and* $\mathrm{N}_{L/K}(\alpha)$ *in terms of* $\mathrm{Tr}_{K(\alpha)/K}$ *and* $\mathrm{N}_{K(\alpha)/K}$ *in the following way*

$$\mathrm{Tr}_{L/K}(\alpha) = [L\colon K(\alpha)]\,\mathrm{Tr}_{K(\alpha)/K}(\alpha), \qquad \mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{K(\alpha)\colon K}(\alpha)^{[L\colon K(\alpha)]}$$

From Corollary 1.3.5 we can draw a relation between the coefficients of the trace and the norm with the roots of the minimum polynomial of α. For any $\alpha \in L$ if f is the minimum polynomial of α over K then $P_{\alpha,K(\alpha)/K} = f(X)$. Now let $f(X) = X^d + c_{d-1}X^{d-1} + \cdots c_1X + c_0$ and f splits into linear factors $f(X) = (X-\alpha_1)\cdots(X-\alpha_d)$ in the splitting field of f over K. Then $-c_{d-1} = \alpha_1 + \cdots + \alpha_d$ and $(-1)^d c_0 = \alpha_1 \cdots \alpha_d$. Therefore, by rewriting Corollary 1.3.5 we get the following corollary.

Corollary 1.3.6

Let the minimal polynomial for α *over* K *splits completely over a large enough field extension* L *of* K *as* $(X-\alpha_1)\cdots(X-\alpha_d)$*. Then*

$$\mathrm{Tr}_{K(\alpha)/K}(\alpha) = \alpha_1 + \cdots + \alpha_d \quad \textit{and} \quad \mathrm{N}_{K(\alpha)/K}(\alpha) = \alpha_1 \cdots \alpha_d$$

More generally, if $K \subseteq K(\alpha) \subseteq L$*, then*

$$\mathrm{Tr}_{L/K}(\alpha) = \frac{n}{d}(\alpha_1 + \cdots + \alpha_d) \quad \textit{and} \quad \mathrm{N}_{L/K}(\alpha) = (\alpha_1 \cdot \alpha_d)^{n/d}$$

where $[L\colon K] = n$ *and* $[K(\alpha)\colon K] = d$*, so* $n/d = [L\colon K(\alpha)]$*.*

Now for finite fields for any $\alpha \in \mathbb{F}_{q^d}$ let f is the minimal polynomial of α over $\mathbb{F}_q$. Then by Theorem 1.2.4 we know the roots of f are $\alpha, \alpha^q, \ldots, \alpha^{\deg(f)-1}$. Hence, rewriting Corollary 1.3.6 in terms of conjugates of α we get the following corollary.

Corollary 1.3.7 Formula of Trace and Norm in Finite Fields

Let $K/\mathbb{F}_q$ *be finite extension with* $[K\colon \mathbb{F}_q] = n$*. Then for any* $\alpha \in K$*,*

$$\mathrm{Tr}_{K/\mathbb{F}_q}(\alpha) = \alpha + \alpha^q + \cdots + \alpha^{q^{n-1}} \quad \textit{and} \quad \mathrm{N}_{K/\mathbb{F}_q}(\alpha) = \alpha \cdot \alpha^q \cdots \alpha^{q^{n-1}}$$

In particular, the trace and norm are invariant under the Frobenius automorphism $\alpha \mapsto \alpha^q$: replacing α by α^q merely cycles the summands (respectively, the factors) in the expressions above, leaving the total unchanged.

Theorem 1.3.8

Let K *be a finite extension of* $F = \mathbb{F}_q$*. The* F*-linear transformations from* K *into* F *are exactly the mappings* L_β*,* $\beta \in K$*, where* $L_\beta(\alpha) = \mathrm{Tr}_{K/F}(\beta\alpha)$ *for all* $\alpha \in K$*. Furthermore,* $L_\beta \neq L_\gamma$ *whenever* β *and* γ *are distinct elements of* K*.*

Proof: For any $\alpha_1, \alpha_2 \in K$ and $c \in F$,

$$L_\beta(\alpha_1 + c\alpha_2) = \mathrm{Tr}_{K/F}(\beta(\alpha_1 + c\alpha_2)) = \mathrm{Tr}_{K/F}(\beta\alpha_1) + c\ \mathrm{Tr}_{K/F}(\beta\alpha_2) = L_\beta(\alpha_1) + c\,L_\beta(\alpha_2),$$

using the F-linearity of $\mathrm{Tr}_{K/F}$ from Theorem 1.3.2.

Let $\beta, \gamma \in K$ with $\beta \neq \gamma$. For any $\alpha \in K$,

$$L_\beta(\alpha) - L_\gamma(\alpha) = \mathrm{Tr}_{K/F}(\beta\alpha) - \mathrm{Tr}_{K/F}(\gamma\alpha) = \mathrm{Tr}_{K/F}((\beta-\gamma)\alpha).$$

Since $\beta - \gamma \neq 0$, the map $\alpha \mapsto (\beta-\gamma)\alpha$ is a K-linear bijection on F, so the image of $\alpha \mapsto \mathrm{Tr}_{F/K}((\beta-\gamma)\alpha)$ equals the image of $\mathrm{Tr}_{F/K}$ itself. As $\mathrm{Tr}_{F/K} : F \to K$ is a non-zero K-linear map onto K, since $\mathrm{Tr}_{K/F}(1) = [K\colon F] > 0$, there exists $\alpha \in F$ such that $\mathrm{Tr}_{F/K}((\beta-\gamma)\alpha) \neq 0$. Hence, $L_\beta \neq L_\gamma$.

Write $F = \mathbb{F}_q$ and $K = \mathbb{F}_{q^m}$, so $[K\colon F] = m$. Since the L_β are pairwise distinct as β ranges over $K = \mathbb{F}_{q^m}$, they yield q^m distinct F-linear maps from K to F. On the other hand, any F-linear map $K \to F$ is determined by its values on a fixed F-basis of K, which has m elements; each value can be assigned arbitrarily in $F = \mathbb{F}_q$. So the total number of F-linear maps $K \to F$ is q^m. Therefore, the maps $\{L_\beta : \beta \in K\}$ exhaust all F-linear transformations from K into F.

■

Theorem 1.3.9

Let K be a finite extension of $F = \mathbb{F}_q$. Then for $\alpha \in K$ we have $\mathrm{Tr}_{K/F}(\alpha) = 0$ if and only if $\alpha = \beta^q - \beta$ for some $\beta \in K$.

Proof: The sufficiency of the condition is obvious. To prove the necessity, suppose $\alpha \in K = \mathbb{F}_{q^m}$ with $\mathrm{Tr}_{K/F}(\alpha) = 0$ and let β be a root of $X^q - X - \alpha$ in some extension field of K. Then $\beta^q - \beta = \alpha$ and

$$\begin{aligned} 0 = \mathrm{Tr}_{K/F}(\alpha) &= \alpha + \alpha^q + \cdots + \alpha^{q^{m-1}} \\ &= (\beta^q - \beta) + (\beta^q - \beta)^q + \cdots + (\beta^q - \beta)^{q^{m-1}} \\ &= (\beta^q - \beta) + (\beta^{q^2} - \beta^q) + \cdots + (\beta^{q^m} - \beta^{q^{m-1}}) \\ &= \beta^{q^m} - \beta, \end{aligned}$$

so that $\beta \in K$. ■

Theorem 1.3.10 Transitive Property of Trace and Norm

Let $L/K/F$ be finite extensions. Then for any $\alpha \in L$

$$\mathrm{Tr}_{L/F}(\alpha) = \mathrm{Tr}_{K/F} \circ \mathrm{Tr}_{L/K}(\alpha) \quad \textit{and} \quad \mathrm{N}_{L/F}(\alpha) = \mathrm{N}_{K/F} \circ \mathrm{N}_{L/K}(\alpha)$$

Proof: Let $[L\colon K] = m$ and $[K\colon F] = n$. Suppose $\{a_1, \ldots, a_n\}$ is a F-basis of K and $\{b_1, \ldots, b_m\}$ is a K-basis of L. Then $\{a_i \cdot b_j : i \in [n], j \in [m]\}$ is a F-basis of L.

For any $\alpha \in L$ let

$$\alpha \cdot b_j = \sum_{i=1}^{m} c_{ij} \cdot b_i, \quad c_{ij} a_s = \sum_{r=1}^{n} d_{ijrs} \cdot a_r \tag{1.2}$$

where $c_{ij} \in K$ and $d_{ijrs} \in F$. Therefore,

$$\alpha(a_s b_j) = \sum_{i=1}^{m} c_{ij} \cdot a_s \cdot b_i = \sum_{i=1}^{m} \sum_{r=1}^{n} b_{ijrs} \cdot a_r b_j$$

So now if we look into the matrix representation of $[m_\alpha]_{L/K}$ and $[m_\alpha]_{L/F}$ then we have

$$[m_\alpha]_{L/K} = (c_{ij})_{1 \leq i,j \leq n}, \quad [m_{c_{ij}}]_{K/F} = (b_{ijrs})_{1 \leq r,s \leq n}, \quad [m_\alpha]_{L/F} = \left([m_{c_{ij}}]_{K/F}\right)_{1 \leq i,j \leq n}$$

where the field extension in the subscript indicates what extension is being used for that matrix. The last matrix is a block matrix. So using all these three matrices we get

$$\mathrm{Tr}_{K/F} \circ \mathrm{Tr}_{L/K}(\alpha) = \mathrm{Tr}_{K/F}\left(\sum_{i=1}^{m} c_{ii}\right) = \sum_{i=1}^{m} \mathrm{Tr}_{K/F}(c_{ii}) = \sum_{i=1}^{m} \mathrm{Tr}\left([m_{c_{ii}}]_{K/F}\right) = \sum_{i=1}^{m}\sum_{r=1}^{n} b_{iirr} = \mathrm{Tr}_{L/F}(\alpha)$$

So we have the transitivity of the trace function. Now we will show for the norm function. For norm it is very delicate and harder to show. We will show a proof from [Sch05]. Again assume $[L\colon K] = m$ and $[K\colon F] = n$. We will prove this by dividing it into 3 cases.

Case I $\alpha \in K$**:** Suppose the transitivity formula already holds. Then $\mathrm{N}_{L/K}(\alpha) = \alpha^{[L\colon K]}$ and hence $\mathrm{N}_{K/F} \circ \mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{K/F}(\alpha^{[L\colon K]}) = \mathrm{N}_{K/F}(\alpha)^{[L\colon K]}$. So in this case we will try to show that this expression is true.

In this case $F \subseteq F(\alpha) \subseteq K \subseteq L$. Let $f - \min(F, \alpha)$ and $\deg(f) = d$. Therefore, by Theorem 1.3.4 we have $P_{\alpha,F(\alpha)/F}(X) = f(X)$ and $P_{\alpha,K/F}(X) = P_{\alpha,F(\alpha)/F}(X)^{[K\colon F(\alpha)]}$ and $P_{\alpha,L/F}(X) = P_{\alpha,F(\alpha)/F}(X)^{[L\colon F(\alpha)]}$. Since $F \subseteq F(\alpha) \subseteq K \subseteq L$ we have by Theorem 1.1.6, $[L\colon F(\alpha)] = [L\colon K] \cdot [K\colon F(\alpha)]$. Therefore,

$$P_{\alpha,L/F}(X) = P_{\alpha,F(\alpha)/F}(X)^{[L\colon F(\alpha)]} = \left(P_{\alpha,F(\alpha)/F}(X)^{[K\colon F(\alpha)]}\right)^{[L\colon K]} = P_{\alpha,K/F}(X)^{[L\colon K]}$$

So if we compare the constant terms we get

$$(-1)^{[L\colon F]}\,\mathrm{N}_{L/F}(\alpha) = \left((-1)^{[K\colon F]}\,\mathrm{N}_{K/F}(\alpha)\right)^{[L\colon K]} \implies \mathrm{N}_{L/F}(\alpha) = \mathrm{N}_{K/F}(\alpha)^{[L\colon K]}$$

Hence we have settled this case.

Case II $L = K(\alpha)$**:** Let f is the minimal polynomial of α over K. Then $\deg(f) = m$. Let $f(X) = X^m + c_{m-1}X^{m-1} + \cdots + c_0$. Then

$$\mathrm{N}_{K/F} \circ \mathrm{N}_{K(\alpha)/K}(\alpha) = \mathrm{N}_{K/F}((-1)^m c_0) = (-1)^{mn}\,\mathrm{N}_{K/F}(c_0)$$

So now we will compute $\mathrm{N}_{K/F}(c_0)$.

Let $\{b_1, \ldots, b_n\}$ is an F-basis of K. So an F-basis of $K(\alpha)$ is $\{\alpha^i \cdot b_j\colon 0 \le i \le m-1, 1 \le j \le n\}$. By definition $\mathrm{N}_{K(\alpha)/F}(\alpha)$ is the determinant of the matrix for multiplication by α on $K(\alpha)$ as an F-linear map. Now for all $j \in [n]$, $\alpha \cdot (\alpha^i b_j) = \alpha^{i+1} b_j$ where $0 \le i \le m-2$ but

$$\alpha \cdot (\alpha^{m-1} b_j) = \alpha^m b_j = -\alpha^{m-1}(c_{m-1} b_j) - \cdots - \alpha(c_1 b_j) - c_0 b_j$$

Let C_i be the matrix of expressing $c_i b_j$ for all $j \in [n]$ as F-linear combinations of $b_1, \ldots, b_n$. Hence, using the above basis we have the matrix to be

$$[m_\alpha]_{K(\alpha)/F} = \begin{bmatrix} O & O & \cdots & O & -C_0 \\ I_n & O & \cdots & O & -C_1 \\ O & I_n & \cdots & O & -C_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \cdots & I_n & -C_{m=1} \end{bmatrix} = \begin{bmatrix} O & -C_0 \\ I_{(m-1)n} & B \end{bmatrix}$$

where O is the zero matrix of dimension $n \times (m-1)n$ and B is the $(m-1)n \times n$ matrix gathering all the $C_1, \ldots, C_{m-1}$. Therefore,

$$\det \begin{bmatrix} O & -C_0 \\ I_{(m-1)n} & B \end{bmatrix} = (-1)^{(m-1)n^2} \det(-C_0) = (-1)^{(m-1)n} \cdot (-1)^n \det(C_0) = (-1)^{mn} \det(C_0)$$

Now C_0 is the matrix for multiplication by c_0 and therefore, $\det(c_0) = \mathrm{N}_{K/F}(c_0)$. So, $\mathrm{N}_{K(\alpha)/F}(\alpha) = (-1)^{mn}\,\mathrm{N}_{K/F}(c_0)$ and previously we prove that $(-1)^{mn}\,\mathrm{N}_{K/F}(c_0) = \mathrm{N}_{K/F} \circ \mathrm{N}_{K(\alpha)/K}(\alpha)$. So we have settled case II.

Case III General Situation: Therefore in this case $F \subseteq K \subseteq F(\alpha) \subseteq L$. Then we have

$$\begin{aligned}
\mathrm{N}_{L/F}(\alpha) &= \mathrm{N}_{F(\alpha)/F}(\alpha)^{[L:\,F(\alpha)]} && [\text{By Case I for } L/F(\alpha)/F] \\
&= \left(\mathrm{N}_{K/F} \circ \mathrm{N}_{F(\alpha)/K}(\alpha)\right)^{[L:\,F(\alpha)]} && [\text{By Case II for } F(\alpha)/K/F] \\
&= \mathrm{N}_{K/F}\left(\mathrm{N}_{F(\alpha)/K}(\alpha)^{[L:\,F(\alpha)]}\right) \\
&= \mathrm{N}_{K/F} \circ \mathrm{N}_{L/K}(\alpha) && [\text{By Case I for } L/F(\alpha)/F]
\end{aligned}$$

This settles case III too. And hence the proof is complete. ■

This concludes our discussion on finite fields and their basic theory, including field extensions, the structure of $\mathbb{F}_q$, and the trace and norm maps. In the next chapter we turn to the theory of characters of finite fields, which will be the primary tool in our study of equations over finite fields. For a thorough and self-contained treatment of finite fields beyond what is covered here, the reader is referred to [LN96].

Chapter 2

Exponential Sums over Finite Fields

An *exponential sum* over $\mathbb{F}_q$ is a finite sum of the form $\sum_{\alpha\in\mathbb{F}_q} f(\alpha)$, where f takes values on the unit circle in $\mathbb{C}$. Because $\mathbb{F}_q$ carries two group structures simultaneously, the additive group $(\mathbb{F}_q, +)$ and the multiplicative group $(\mathbb{F}_q^*, \times)$, there are two natural families of unit-circle-valued homomorphisms, called *additive* and *multiplicative characters*, and exponential sums arise by combining them in various ways. These sums encode deep arithmetic about $\mathbb{F}_q$: the absolute value $|\sum_\alpha f(\alpha)|$ measures how uniformly distributed the values of f are, and proving that it is small (typically $O(\sqrt{q})$) has direct consequences for counting solutions to polynomial equations, the distribution of quadratic residues, and the structure of Gauss and Kloosterman sums.

Two Types of Characters. A finite field $\mathbb{F}_q$ has two natural families of characters. *Additive characters* $\chi : \mathbb{F}_q \to S^1$ are group homomorphisms from $(\mathbb{F}_q, +)$; every one is of the form $\chi_\beta(\alpha) = e^{2\pi i\,\mathrm{Tr}(\beta\alpha)/p}$ for some $\beta \in \mathbb{F}_q$, where $\mathrm{Tr} : \mathbb{F}_q \to \mathbb{F}_p$ is the absolute trace. *Multiplicative characters* $\psi : \mathbb{F}_q^* \to S^1$ are group homomorphisms from the cyclic group $(\mathbb{F}_q^*, \times)$ of order $q-1$; a particularly important special case is the *quadratic character* η, the finite-field analog of the Legendre symbol. The interplay between these two families, their orthogonality, their products, and their behavior under field extensions, is the engine driving all the results in this chapter.

§ 2.1 Characters of Finite Abelian Groups

Let G be a finite abelian group with the identity element 1_G. Let S^1 be the set of complex numbers with absolute value 1. Then consider the multiplicative group S^1 with the operation of multiplication.

Definition 2.1.1: Character χ of G

A character χ *of* G *is a group homomorphism* $\chi : G \to S^1$ *to the multiplicative group* S^1 *i.e. for all* $g_1, g_2 \in G$

$$\chi(g_1 \cdot g_2) = \chi(g_1) \cdot \chi(g_2)$$

So if χ is any character of G then $\chi(1_G) = 1$. Since by definition for all $g \in G$,

$$\chi(g)^{|G|} = \chi(g^{|G|}) = \chi(1_G) = 1$$

the values of χ are $|G|^{th}$ roots of unity. Also we have for any $g \in G$,

$$\chi(g)\chi(g^{-1}) = \chi(g \cdot g^{-1}) = \chi(1_G) = 1$$

so $\chi(g^{-1}) = \chi(g)^{-1} = \overline{\chi(g)}$.

There are some special characters of G which have specific name. The character χ_0 defined by $\chi_0(g) = 1$ for all $g \in G$ is called the *trivial character*. For any character χ of G, the character $\overline{\chi}$ defined by $\overline{\chi}(g) = \overline{\chi(g)}$ for all $g \in G$ is called the *conjugate character* of χ.

Given any finitely many characters of G, $\chi_1, \ldots, \chi_n$ we can define the product character $\chi = \chi_1 \cdot \chi_2 \cdots \chi_n$ such that for all $g \in G$,

$$\chi(g) = \chi_1(g) \cdot \chi_2(g) \cdots \chi_n(g)$$

Hence for all $g \in G$,

$$\chi(g) \cdot \overline{\chi}(g) = |\chi(g)|^2 = 1$$

Hence the $\overline{\chi}$ is also the inverse character of χ. Hence the set of all characters of G forms a group, $\mathrm{Hom}(G, S^1)$ under multiplication whose identity element is the trivial character χ_0.

Definition 2.1.2: Dual Group to G

For a finite abelian group G, the group of all characters of G with the identity element χ_0 is called the dual group *to G and is denoted by $\widehat{G}$ i.e. $\widehat{G} := \mathrm{Hom}(G, S^1)$.*

Observation 2.1. *The dual group $\widehat{G}$ is a finite abelian group. So we can talk about characters of $\widehat{G}$ too.*

Since the values of characters of G can only be $|G|^{th}$ roots of unity, the dual group $\widehat{G}$ is a finite abelian group.

Lemma 2.1.1 Characters of Cyclic Group

If G is a finite cyclic group of order n and let g be the generator of G, then for all $j \in \{0, 1, \ldots, n-1\}$ define the function

$$\chi_j(g^k) = e^{2\pi i \frac{j \cdot k}{n}} \quad \forall\, k \in \{0, 1, \ldots, n-1\}$$

Then $\widehat{G} = \{\chi_j : j \in \{0, 1, \ldots, n-1\}$.

Proof: Certainly for all $j \in \{0, 1, \ldots, n-1\}$ defines a character of G. On the other hand suppose, $\chi \in \widehat{G}$. Then $\chi(g)$ must be an n^{th} root of unity. Hence there exists $j \in \{0, 1, \ldots, n-1\}$ such that $\chi(g) = e(j/n)$. Therefore $\chi = \chi_j$. Hence $\widehat{G} = \{\chi_j : 0 \le j \le n-1\}$. ∎

Observation 2.2. *The characters defined in above lemma forms a cyclic group with χ_1 being the generator of $\widehat{G}$.*

Note:-

When $G = (\mathbb{Z}/n\mathbb{Z})^*$, the characters of G described by $\chi_j(\alpha) = e^{2\pi i \frac{\alpha \cdot j}{n}}$ like in Lemma 2.1.1 are precisely the classical *Dirichlet characters* modulo n, which are of fundamental importance in analytic number theory.

Corollary 2.1.2 Isomorphism of a Cyclic Group with its Dual

If G is a finite cyclic group of order n with g being the generator then $G \cong \widehat{G}$ by the isomorphism $\varphi : g^j \mapsto \chi_j$ for all $j \in \{0, 1, \ldots, n-1\}$.

Now we will show that we can extend a character of any subgroup to a character of the group.

Theorem 2.1.3 Extending Characters from Subgroup to Group

Let H be a subgroup of the finite abelian group G and let χ_H be a character of H. Then χ_H can be extended to a character of G that is there exists a character $\chi \in \widehat{G}$ with $\chi(h) = \chi_H(h)$ for all $h \in H$.

Proof: Let H be a proper subgroup of G. Let $\alpha \in G \setminus H$. Then let H_1 be the group generated by H and α i.e. $H_1 = \langle H, a \rangle$. Let m be the smallest positive integer such that $\alpha^m \in H$. Then for all $g \in H_1$, g can be written uniquely as $g = \alpha^j \cdot h$ where $0 \leq j < m$ and $h \in H$.

Let ω be a complex number such that $\omega^m = \chi_H(\alpha^m)$. Then for all $g \in H_1$ where $g = \alpha^j \cdot h$, $h \in H$ define the function $\chi_1(g) = \omega^j \chi_H(h)$. Then for any $g_1, g_2 \in H_1$ where $g_1 = \alpha^i \cdot h_1$, $g_2 = \alpha^j \cdot h_2$, we have $g_1 \cdot g_2 = \alpha^{j+k} \cdot h_1 \cdot h_2 = \alpha^{j+k \bmod m} \cdot h_1 \cdot h_2$. Hence

$$\chi_1(g_1 \cdot \chi_2) = \omega^{j+k \bmod m} \chi_H(h_1) \cdot \chi_H(h_2) = \alpha^j \chi_H(h_1) \cdot \omega^k \chi_H(h_2) = \chi_1(g_1) \cdot \chi_1(g_2)$$

Hence χ_1 is a character of H_1. And also for all $h \in H$ we have $\chi_1(h) = \chi_H(h)$.

So if $H_1 = G$ we are done. Otherwise we can continue this process above until we get G. This process we have to do finitely many times as G is a finite group. ∎

Therefore if we have two distinct elements $g_1, g_2 \in G$, $g_1 \neq g_2$ then take the subgroup generated by the element $h = g_1 \cdot g_2^{-1} \neq 1$. Since this subgroup is a cyclic subgroup generated by h, by Lemma 2.1.1 we get a character χ' of the cyclic subgroup such that $\chi'(h) \neq 1$. Then by Theorem 2.1.3 we get a character χ of G such that $\chi(h) = \chi'(h) \neq 1$ and hence $\chi(g_1) \neq \chi(g_2)$. Therefore we have the following observation

Observation 2.3. *For any two distinct elements $g_1, g_2 \in G$, $g_1 \neq g_2$ we can get a character $\chi \in \widehat{G}$ such that $\chi(g_1) \neq \chi(g_2)$.*

Now consider any character χ of G. If $g \in G$ such that $\chi(g) = 1$ then $\chi(g^{-1}) = 1$. Hence the set of all elements $g \in G$ such that $\chi(g)$ forms a group. This group is called the *annihilator* of χ and is denoted by $\mathrm{Ann}(\chi)$.

Similarly suppose H is a subgroup of G. Then if $\chi \in \widehat{G}$ is a character of G such that $\chi(h) = 1$ for all $h \in H$ then $\overline{\chi}(h) = 1$ for all $h \in H$. Hence the set of all characters $\chi \in \widehat{G}$ such that $\chi(h) = 1$ for all $h \in H$ forms a group. This group is called the *annihilator* of H and is denoted by $\mathrm{Ann}(H)$. Since in both cases the notation Ann is used usually the context will make it clear whether we are talking about the annihilator of a character or the annihilator of a subgroup.

Definition 2.1.3: Annihilator

Let G be a finite abelian group. For any character $\chi \in \widehat{G}$, the annihilator *of χ, $\mathrm{Ann}(\chi) = \{g \in G : \chi(g) = 1\}$ which is a subgroup of G.*

For any subgroup H of G, the annihilator *of H, $\mathrm{Ann}(H) = \{\chi \in \widehat{G} : \chi(h) = 1 \text{ for all } h \in H\}$ which is a subgroup of $\widehat{G}$.*

Now every element of $\widehat{G}$ is a function χ which takes an element of G and maps to a complex number in S^1. Similarly we can think of every element of as a function from $\widehat{G}$ to S^1. In fact for every element $g \in G$ define the homomorphism $\mathrm{ev}_g : \widehat{G} \to S^1$ where $\mathrm{ev}_g(\chi) = \chi(g)$. Hence ev_g is a character of $\widehat{G}$ for all $g \in G$.

Theorem 2.1.4

Let G be a finite abelian group.

(i) If $\chi \in \widehat{G}$ be any nontrivial character then

$$\sum_{g \in G} \chi(g) = 0 \tag{2.1}$$

(ii) If $g \in G$ with $g \neq 1_G$ then

$$\sum_{\chi \in \widehat{G}} \chi(g) = 0 \tag{2.2}$$

Proof: Since χ is nontrivial there exists $h \in G$ such that $\chi(h) \neq 1$. Then

$$\chi(h)\sum_{g\in G}\chi(g) = \sum_{g\in G}\chi(h\cdot g) = \sum_{g\in G}\chi(g)$$

Hence we have

$$(\chi(h)-1)\sum_{g\in G}\chi(g) = 0$$

Since we know $\chi(h) \neq 1$ we have $\sum_{g\in G}\chi(g) = 0$.

Now from the discussion above for any $g \in G$ and $\chi \in hG$ we have $\chi(g) = \mathrm{ev}_g(\chi)$. Hence

$$\sum_{\chi\in\widehat{G}}\chi(g) = \sum_{\chi\in\widehat{G}}\mathrm{ev}_g(\chi)$$

Since $g \neq 1_G$, there is a $\chi \in \widehat{G}$ such that $\chi(g) \neq 1$. Hence ev_g is a nontrivial character of $\widehat{G}$. Hence by the first part of the theorem we have $\sum_{\chi\in\widehat{G}}\mathrm{ev}_g(\chi) = 0$. ■

The above theorem leaves the case when χ is the trivial character. Then for all $g \in G$, $\chi(g) = 1$. Hence we have $\sum_{g\in G}\chi(g) = |G|$. Hence we have the following corollary

Observation 2.4. *For any finite abelian group G. Then*

(i) Let χ is a character of G.

$$\sum_{g\in G}\chi(g) = \begin{cases} |G| & \text{if } \chi = \chi_0 \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

(ii) $|G| = \sum_{\chi\in\widehat{G}}\sum_{g\in G}\chi(g)$

So now by counting in two ways for the sum $\sum_{\chi\in\widehat{G}}\sum_{g\in G}\chi(g)$ we get the following theorem:

Theorem 2.1.5

For any finite abelian group G, we have $|G| = |\widehat{G}|$.

Proof: From Observation 2.4(ii) we have

$$|G| = \sum_{\chi\in\widehat{G}}\sum_{g\in G}\chi(g) = \sum_{g\in G}\sum_{\chi\in\widehat{G}}\chi(g) = |\widehat{G}|$$

Hence we have the result. ■

From above theorem and Observation 2.4(i) we get the *orthogonality relations for characters.*

Observation 2.5. *Let G be a finite abelian group.*

(i) Let χ, τ are characters of G. Then

$$\frac{1}{|G|}\sum_{g\in G}\chi(g)\cdot\overline{\tau(g)} = \begin{cases} 1 & \text{if } \chi = \tau \\ 0 & \text{otherwise} \end{cases}$$

(ii) If $g, h \in G$ are any two elements then

$$\frac{1}{|G|}\sum_{\chi\in\widehat{G}}\chi(g)\cdot\overline{\chi(h)} = \begin{cases} 1 & \text{if } g = h \\ 0 & \text{otherwise} \end{cases}$$

Since we already showed that for every element $g \in G$, ev_g is a character of $\widehat{G}$. Theorem 2.1.5 gives us the possibility that G and $\widehat{G}$ are isomorphic. In Lemma 2.1.1 we already showed that for cyclic groups G and $\widehat{G}$ are isomorphic. In the next few results we will show that this is the case for all finite abelian groups.

By Fundamental Theory of Finite Abelian Groups every finite abelian group G can be written as a direct product of finitely many cyclic groups. Hence it is enough to show that if G_1 and G_2 are two finite abelian groups and $G = G_1 \times G_2$ then $\widehat{G}$ is isomorphic to the direct product of $\widehat{G}_1$ and $\widehat{G}_2$.

Theorem 2.1.6 Isomorphism of Product of Dual Groups with Dual of Their Product

Let G_1 and G_2 be two finite abelian groups and let $G = G_1 \times G_2$. Then $\widehat{G} \cong \widehat{G}_1 \times \widehat{G}_2$.

Proof: Since G is the direct product of G_1 and G_2, G consists of pairs (g_1, g_2) where $g_1 \in G_1$ and $g_2 \in G_2$. So consider the map $\varphi : \widehat{G}_1 \times \widehat{G}_2 \to \widehat{G}$ where for any $\chi_1 \in \widehat{G}_1$ and $\chi_2 \in \widehat{G}_2$, $\chi := \varphi(\chi_1, \chi_2)$ is the character of G defined by the following:

$$\text{for all } (g_1, g_2) \in G \qquad \chi(g_1, g_2) = \chi(g_1, 1) \cdot \chi(1, g_2) = \chi_1(g_1) \cdot \chi_2(g_2)$$

Hence φ is a homomorphism. Now we will show that φ is an isomorphism.

Let $\chi \in \widehat{G}$. Then define $\chi_1(g_1) := \chi(g_1, 1)$ for all $g_1 \in G_1$ and similarly define $\chi_2(g_2) := \chi(1, g_2)$ for all $g_2 \in G_2$. Then χ_1 is a character of G_1 and χ_2 is a character of G_2. And also for any $(g_1, g_2) \in G$ we have

$$\chi_1(g_1) \cdot \chi_2(g_2) = \chi(g_1, 1) \cdot \chi(1, g_2) = \chi(g_1, g_2)$$

Hence $\varphi(\chi_1, \chi_2) = \chi$. Hence φ is an isomorphism. Therefore $\widehat{G} \cong \widehat{G}_1 \times \widehat{G}_2$. ∎

Theorem 2.1.7 Isomorphism of Group with its Dual

For any finite abelian group G, G is isomorphic to $\widehat{G}$.

Proof: By Fundamental Theorem of Finite Abelian Groups, G can be written as a direct product of finitely many cyclic groups i.e. there exists cyclic groups $Z_{n_1}, Z_{n_2}, \ldots, Z_{n_k}$ such that

$$G \cong Z_{n_1} \times Z_{n_2} \times \cdots \times Z_{n_k}$$

Hence by repeated use of Theorem 2.1.6 we have $\widehat{G} \cong \widehat{Z}_{n_1} \times \widehat{Z}_{n_2} \times \cdots \times \widehat{Z}_{n_k}$. Since by Lemma 2.1.1 for all $i \in \{1, 2, \ldots, k\}$, $\widehat{Z}_{n_i} \cong Z_{n_i}$ we have $\widehat{G} \cong Z_{n_1} \times Z_{n_2} \times \cdots \times Z_{n_k} \cong G$ Hence we have the result. ∎

§ 2.2 Characters of Finite Fields

In a finite field $\mathbb{F}_q$ there are two inherent group structures–namely, the additive group and the multiplicative group. The nonzero elements of $\mathbb{F}_q$ forms a cyclic group of $q-1$ elements under multiplication and the whole $\mathbb{F}_q$ forms a group of q elements under addition.

2.2.1 Additive Characters of $\mathbb{F}_q$

We first turn to the additive group of $\mathbb{F}_q$. Characters of the additive group $\mathbb{F}_q$ are called *additive characters* of $\mathbb{F}_q$. We denote the set of all additive characters of $\mathbb{F}_q$ by $\boldsymbol{\mathcal{X}}_q$.

Suppose p be the characteristic of $\mathbb{F}_q$. Then the prime field contained in $\mathbb{F}_q$ is $\mathbb{F}_p$. Let $\mathrm{Tr} : \mathbb{F}_q \to \mathbb{F}_p$ be the absolute *trace function* from $\mathbb{F}_q$ to $\mathbb{F}_p$. Then define the function χ_1 by

$$\chi_1(\alpha) = e^{2\pi i \frac{\mathrm{Tr}(\alpha)}{p}} \quad \forall\, \alpha \in \mathbb{F}_q$$

Then χ_1 is an additive character of $\mathbb{F}_q$ since for all $\alpha_1, \alpha_2 \in \mathbb{F}_q$, $\mathrm{Tr}(\alpha_1 + \alpha_2) = \mathrm{Tr}(\alpha_1) + \mathrm{Tr}(\alpha_2)$ and hence $\chi_1(\alpha_1 + \alpha_2) = \chi_1(\alpha_1) \cdot \chi_1(\alpha_2)$. This character χ_1 is called the *canonical additive character* of $\mathbb{F}_q$.

Theorem 2.2.1 Structure of Additive Characters

For $\beta \in \mathbb{F}_q$, define χ_β with $\chi_\beta(\alpha) = \chi_1(\beta \cdot \alpha)$ for all $\alpha \in \mathbb{F}_q$. Then χ_β is an additive character of $\mathbb{F}_q$ and every additive character of $\mathbb{F}_q$ is obtained this way i.e. $\forall\, \chi \in \boldsymbol{\mathcal{X}}_q$, there exists $\beta \in \mathbb{F}_q$ such that $\forall\, \alpha \in \mathbb{F}_q$, $\chi(\alpha) = \chi_\beta(\alpha)$.

Proof: Now for any $\alpha_1, \alpha_2 \in \mathbb{F}_q$, we have

$$\chi_\beta(\alpha_1 + \alpha_2) = \chi_1(\beta(\alpha_1 + \alpha_2)) = \chi_1(\beta \cdot \alpha_1) \cdot \chi_1(\beta \cdot \alpha_2) = \chi_\beta(\alpha_1) \cdot \chi_\beta(\alpha_2)$$

Hence χ_β is an additive character of $\mathbb{F}_q$.

Now we will show that for any two distinct $\beta_1, \beta_2 \in \mathbb{F}_q$ with $\beta_1 \neq \beta_2$ there exists an $\alpha \in \mathbb{F}_q$ such that $\chi_{\beta_1}(\alpha) \neq \chi_{\beta_2}(\alpha)$. Suppose that is not the case. Then for all $\alpha \in \mathbb{F}_q$,

$$\chi_{\beta_2}(\alpha) = \chi_{\beta_2}(\alpha) \iff \chi_1(\beta_1 \cdot \alpha) = \chi_1(\beta_2 \cdot \alpha) \iff \chi_1((\beta_1 - \beta_2)\alpha) = 1.$$

Now $\chi_1((\beta_1 - \beta_2)\alpha) = 1$ for all $\alpha \in \mathbb{F}_q$ if and only if $\mathrm{Tr}((\beta_1 - \beta_2)\alpha) = 0$ for all $\alpha \in \mathbb{F}_q$. Since $\mathrm{Tr}(X) = X + X^p + \cdots + X^{p^{r-1}}$ where $q = p^r$. Since Tr is $\mathbb{F}_p$–linear, $\mathrm{Tr}((\beta_1 - \beta_2)X)$ has p^r roots in $\mathbb{F}_q$ which is more than the degree p^{r-1}. Hence $\mathrm{Tr}((\beta_1 - \beta_2)X)$ is the zero polynomial. Therefore $\beta_1 = \beta_2$. So for any two distinct $\beta_1, \beta_2 \in \mathbb{F}_q$ with $\beta_1 \neq \beta_2$ there exists an $\alpha \in \mathbb{F}_q$ such that $\chi_{\beta_1}(\alpha) \neq \chi_{\beta_2}(\alpha)$.

Hence for every $\beta \in \mathbb{F}_q$ we have a distinct additive character. Therefore we have q additive characters. Since there are q additive characters of $\mathbb{F}_q$ by Theorem 2.1.5 every additive character of $\mathbb{F}_q$ can be obtained the above way. ■

The orthogonality relations in Observation 2.5 when applied to additive characters and along with Theorem 2.2.1 gives the following result.

Theorem 2.2.2 Orthogonality of Additive Characters

(i) For any $a \in \mathbb{F}_q$, the corresponding additive character is χ_a. Then

$$\sum_{\gamma \in \mathbb{F}_q} \chi_a(\gamma) = \begin{cases} 0 & \text{if } a \neq 0 \\ q & \text{if } a = 0 \end{cases}$$

(ii) For any two additive characters χ_a, χ_b where $a, b \in \mathbb{F}_q$,

$$\sum_{\gamma \in \mathbb{F}_q} \chi_a(\gamma) \cdot \overline{\chi_b}(\gamma) = \begin{cases} 0 & \text{if } a \neq b \\ q & \text{if } a = b \end{cases}$$

(iii) For any two elements $\alpha, \beta \in \mathbb{F}_q$,

$$\sum_{c \in \mathbb{F}_q} \chi_c(\alpha) \cdot \overline{\chi_c(\beta)} = \begin{cases} 0 & \text{if } \alpha \neq \beta \\ q & \text{if } \alpha = \beta \end{cases}$$

2.2.2 Multiplicative Characters of $\mathbb{F}_q$

Characters of the multiplicate group $\mathbb{F}_q^*$ are called *multiplicative characters* of $\mathbb{F}_q$. We denote the set of all multiplicative characters of $\mathbb{F}_q$ by $\boldsymbol{\mathcal{M}}_q$. Since $\mathbb{F}_q^*$ is a cyclic group of order $q-1$ by Lemma 2.1.1 we have the following:

Lemma 2.2.3 Structure of Multiplicative Characters

The group of multiplicative characters of $\mathbb{F}_q$ is a cyclic group of order $q-1$ and are given by $\{\psi_j : 0 \le j \le q-2\}$ where if g is the generator of $\mathbb{F}_q^*$ then for all $k \in \{0, 1, \ldots, q-2\}$ we have $\psi_j(g^k) = e^{2\pi i \frac{jk}{q-1}}$ where ψ_0 is the identity element of the group and the trivial character.

Every multiplicative character ψ of $\mathbb{F}_q$ has the property that $\psi^{q-1} = \psi_0$. Let ψ be the generator of the cyclic group of multiplicative characters. Then ψ is called the *principal character*. For any multiplicative character ψ of $\mathbb{F}_q$ we say ψ is of *order* d if $\psi^d = \psi_0$ and d is the smallest positive integer with this property. We use $\operatorname{ord}(\psi)$ to denote the order of ψ. Hence $d \mid q-1$. We say ψ is of *exponent* e if $\psi^e = \psi_0$. Clearly $d \mid e$ where d is the order of ψ. Let $\boldsymbol{\mathcal{M}}_q^{(e)}$ denote the set of all multiplicative characters of $\boldsymbol{\mathcal{M}}_q$ of exponent e.

Note:-

For odd q a special multiplicative character of interest is the *quadratic character*. It is denoted by η. It only takes the value ± 1. So $\operatorname{ord}\eta = 2$. For all $\alpha \in \mathbb{F}_q^*$, $\eta(\alpha) = 1$ if α is a *quadratic residue* and -1 if α is a non-quadratic residue. Therefore for all $\alpha \in \mathbb{F}_q$, we have $\eta(\alpha) = \alpha^{(q-1)/2}$. The *Legendre symbol* $\left(\frac{\alpha}{q}\right)$ for any $\alpha \in \mathbb{F}_q$ is defined by $\eta(\alpha) = \left(\frac{\alpha}{q}\right)$

For any multiplicative character ψ of $\mathbb{F}_q$, the value of $\psi(-1)$ is of special interest. We obviously have $\psi(-1) = \pm 1$. Let $\operatorname{ord}(\psi) = m$. Then $m \mid q-1$. Hence values of ψ are m^{th} roots of unity. Therefore $\psi(-1)$ can be -1 if m is even. Let g be a primitive element of $\mathbb{F}_q$ i.e. $\mathbb{F}_q^*$ is generated by g. Hence $\psi(g) = \zeta_m$ where ζ_m is the m^{th} root of unity. If m is even

$$\psi(-1) = \psi\left(g^{(q-1)/2}\right) = \zeta_m^{(q-1)/2}$$

Now

$$\zeta_m^{(q-1)/2} = -1 \iff \frac{q-1}{2} \equiv \frac{m}{2} \bmod m \iff \frac{q-1}{m} \equiv 1 \bmod 2$$

So m should be even and $\frac{q-1}{m}$ should be odd. Therefore we have the following two observations.

Observation 2.6. *For any multiplicative character $\psi \in \boldsymbol{\mathcal{M}}_q$, $\psi(-1) = -1$ if and only if $\frac{q-1}{m}$ is odd and m is even.*

Observation 2.7. *If $\psi \in \boldsymbol{\mathcal{M}}_q$ is the principal multiplicative character of $\mathbb{F}_q$ then $\psi(-1) = -1$.*

Suppose $d \mid q-1$. Then for every multiplicative character ψ of exponent d and for every $\alpha \in \mathbb{F}_q^*$, we have $\psi(\alpha^d) = \psi^d(\alpha) = 1$. Hence $\psi(\alpha) = 1$ if $\alpha \in (\mathbb{F}_q^*)^{(d)}$. Conversely if $\psi(\alpha) = 1$ for every $\alpha \in (\mathbb{F}_q^*)^{(d)}$ then $\psi^d = \psi_0$. Hence ψ is a character of exponent d then $\psi(\alpha)$ depends only on the cosets of $\mathbb{F}_q^* / (\mathbb{F}_q^*)^{(d)}$. Therefore there are precisely d such characters of exponent d.

Theorem 2.2.4 Character Sum over dth-Power Characters

Suppose $d \mid q-1$. Suppose $\alpha \in \mathbb{F}_q^*$. Then

$$\sum_{\psi \in \boldsymbol{\mathcal{M}}_q^{(d)}} \chi(\alpha) = \begin{cases} d & \text{if } \alpha \in (\mathbb{F}_q^*)^{(d)} \\ 0 & \text{otherwise} \end{cases}$$

Proof: The characters of exponent d are characters of the group $\mathbb{F}_q^* / (\mathbb{F}_q^*)^{(d)}$. Hence this follows from Observation 2.4. ■

Again, the orthogonality relations in Observation 2.5 when applied to multiplicative characters gives the following:

Theorem 2.2.5 Orthogonality of Multiplicative Characters

(i) For any multiplicative character ψ of $\mathbb{F}_q$

$$\sum_{\alpha\in\mathbb{F}_q^*}\psi(\alpha)=\begin{cases}0 & \text{if } \psi\neq\psi_0\\ q-1 & \text{if } \psi=\psi_0\end{cases}$$

(ii) For any two multiplicative characters ψ and λ of $\mathbb{F}_q$,

$$\sum_{\alpha\in\mathbb{F}_q^*}\psi(\alpha)\cdot\overline{\lambda(\alpha)}=\begin{cases}0 & \text{if } \psi\neq\lambda\\ q-1 & \text{if } \psi=\lambda\end{cases}$$

(iii) For any two elements $\alpha,\beta\in\mathbb{F}_q^*$,

$$\sum_{\psi\in\boldsymbol{\mathcal{M}}_q}\psi(\alpha)\cdot\overline{\psi(\beta)}=\begin{cases}0 & \text{if } \alpha\neq\beta\\ q-1 & \text{if } \alpha=\beta\end{cases}$$

Therefore the multiplicative characters form an orthonormal basis on all complex valued functions from $\mathbb{F}_q^*$ to $\mathbb{C}$. So we have the observation

Observation 2.8. *obs:mult-char-fourier-basis The set of multiplicative characters $\boldsymbol{\mathcal{M}}_q$ of $\mathbb{F}_q$ forms an orthonormal basis for the vector space W of all complex valued functions $\varphi\colon\mathbb{F}_q^*\to\mathbb{C}$.*

Usually we extend the definition of multiplicative characters to the whole $\mathbb{F}_q$ so that we can define special character sums over the whole field. Let ψ is a multiplicative character then $\psi(0)$ is defined to be 0 if ψ is nontrivial multiplicative character and $\psi(0)=1$ if ψ is the trivial character.

2.2.3 Lifting of Characters to Finite Extension

Let E be a finite extension field of $\mathbb{F}_q$. Let χ_1 and μ_1 be the canonical additive characters of $\mathbb{F}_q$ and E respectively. Let $\mathrm{Tr}_{E/\mathbb{F}_q}$ be the trace function from E to $\mathbb{F}_q$.

Then χ_1 and μ_1 are related by the identity

$$\chi_1(Tr_{E/\mathbb{F}_q}(\alpha))=\mu_1(\alpha)\quad\forall\,\alpha\in E$$

Since the trace function satisfies transitivity property we have $\mathrm{Tr}_E=\mathrm{Tr}_{\mathbb{F}_q}\circ\mathrm{Tr}_{E/\mathbb{F}_q}$. Hence for all $\alpha\in E$, $\mu_1(\alpha)=\chi_1(\mathrm{Tr}_{E/\mathbb{F}_q}(\alpha))$.

Similarly if ψ_1 and λ_1 are principal multiplicative characters of $\mathbb{F}_q$ and E respectively then,

$$\lambda_1(\alpha)=\psi_1(\mathrm{N}_{E/\mathbb{F}_q}(\alpha))\quad\forall\,\alpha\in E^*$$

where $\mathrm{N}_{E/\mathbb{F}_q}$ is the *norm function* from E to $\mathbb{F}_q$.

In the opposite way for any additive character χ of $\mathbb{F}_q$ or multiplicative character ψ of $\mathbb{F}_q$ we can lift the characters to an additive and multiplicative character of E respectively by taking $\chi\circ\mathrm{Tr}_{E/\mathbb{F}_q}$ and $\psi\circ\mathrm{N}_{E/\mathbb{F}_q}$ respectively. We denote these by $\chi^{(r)}:=\chi\circ\mathrm{Tr}_{E/\mathbb{F}_q}$ and $\psi^{(r)}:=\psi\circ\mathrm{N}_{E/\mathbb{F}_q}$ where $r=[E:\mathbb{F}_q]$.

In the following sections we will be working with hybrid character sums which are formed by taking product of multiple characters of both kind. Hence we will be using χ,τ to denote additive characters and ψ,λ to denote multiplicative characters.

§ 2.3 Gaussian Sums

Having studied additive and multiplicative characters of $\mathbb{F}_q$ separately, we now turn to hybrid sums that interleave both types. The first and most fundamental such object is the *Gaussian sum*. A Gaussian sum is formed by taking the product of a multiplicative character and an additive character and summing over all nonzero elements of $\mathbb{F}_q$. The interplay between the two kinds of characters gives remarkable algebraic and analytic properties, and they will serve as the key tool throughout this chapter. We define them as follows.

Definition 2.3.1: Gaussian Sum

Let ψ be a multiplicative character and χ be an additive character of $\mathbb{F}_q$. Then the Gaussian Sum *associated to ψ and χ is defined as*

$$G(\psi, \chi) = \sum_{\alpha \in \mathbb{F}_q^*} \psi(\alpha) \cdot \chi(\alpha)$$

The absolute value of the Gaussian sum is at most $q-1$ by the triangle inequality. However, we will see below that when both ψ and χ are non-trivial, $|G(\psi, \chi)| = \sqrt{q}$ — far smaller than this naive bound.

Theorem 2.3.1 Evaluation of Gaussian Sums

Let ψ be a multiplicative character and χ be an additive character of $\mathbb{F}_q$. Then

$$G(\psi, \chi) = \begin{cases} q-1 & \text{if } \psi = \psi_0, \chi = \chi_0 \\ -1 & \text{if } \psi = \psi_0, \chi \neq \chi_0 \\ 0 & \text{if } \psi \neq \psi_0, \chi = \chi_0 \end{cases} \tag{2.4}$$

If ψ is non-trivial multiplicative character and χ is non-trivial additive character then $|G(\psi, \chi)| = \sqrt{q}$.

Proof: If $\psi = \psi_0$ and $\chi = \chi_0$ then for all $\alpha \in \mathbb{F}_q^*$ we have $\psi(\alpha) = 1$ and $\chi(\alpha) = 1$. Hence $G(\psi_0, \chi_0) = q - 1$.

Now suppose $\psi = \psi_0$ and χ is non-trivial additive character. Then by Theorem 2.2.2(i)

$$G(\psi_0, \chi) = \sum_{\alpha \in \mathbb{F}_q^*} \chi(\alpha) = \sum_{\alpha \in \mathbb{F}_q} \chi(\alpha) - \chi(0) = -1$$

Let ψ be a non-trivial multiplicative character and $\chi = \chi_0$. Then for all $\alpha \in \mathbb{F}_q^*$ we have $\chi(\alpha) = 1$. Hence by Theorem 2.2.5(i)

$$G(\psi, \chi_0) = \sum_{\alpha \in \mathbb{F}_q^*} \psi(\alpha) = 0$$

So now assume both ψ and χ are non-trivial multiplicative and additive characters of $\mathbb{F}_q$ respectively. Then we

have

$$\begin{aligned}
|G(\psi,\chi)|^2 &= \overline{G(\psi,\chi)} \cdot G(\psi,\chi) \\
&= \left(\sum_{\alpha\in\mathbb{F}_q^*} \overline{\psi(\alpha)} \cdot \overline{\chi(\alpha)}\right) \cdot \left(\sum_{\beta\in\mathbb{F}_q^*} \psi(\beta)\cdot\chi(\beta)\right) \\
&= \sum_{\alpha\in\mathbb{F}_q^*}\sum_{\beta\in\mathbb{F}_q^*} \overline{\psi(\alpha)}\cdot\overline{\chi(\alpha)}\cdot\psi(\beta)\cdot\chi(\beta) \\
&= \sum_{\alpha\in\mathbb{F}_q^*}\sum_{\beta\in\mathbb{F}_q^*} \psi(\beta\cdot\alpha^{-1})\cdot\chi(\beta-\alpha) \\
&= \sum_{\alpha\in\mathbb{F}_q^*}\sum_{\gamma\in\mathbb{F}_q^*} \psi(\gamma)\cdot\chi(\alpha(\gamma-1)) && [\text{Take } \gamma = \beta\cdot\alpha^{-1}] \\
&= \sum_{\gamma\in\mathbb{F}_q^*} \psi(\gamma)\cdot\left(\sum_{\alpha\in\mathbb{F}_q^*}\chi(\alpha(\gamma-1))\right) \\
&= \sum_{\gamma\in\mathbb{F}_q^*} \psi(\gamma)\cdot\left(\sum_{\alpha\in\mathbb{F}_q}\chi(\alpha(\gamma-1))\right) - \chi(0)\sum_{\gamma\in\mathbb{F}_q^*}\psi(\gamma) \\
&= \sum_{\gamma\in\mathbb{F}_q^*} \psi(\gamma)\cdot\left(\sum_{\alpha\in\mathbb{F}_q}\chi(\alpha(\gamma-1))\right) && [\text{By Theorem 2.2.5(i)}]
\end{aligned}$$

Hence by Theorem 2.2.2(i)

$$\sum_{\alpha\in\mathbb{F}_q}\chi(\alpha(\gamma-1)) = \begin{cases} 0 & \text{if } \gamma \neq 1 \\ q & \text{if } \gamma = 1 \end{cases}$$

Therefore $|G(\psi,\chi)|^2 = \psi(1)q = q$ and so we have the required result. ■

Now under various transformations of the additive or multiplicative characters Gaussian sums gives some useful identities. We will see some of these below.

Lemma 2.3.2 Gaussian Sum Identities

Gaussian sums for the finite field $\mathbb{F}_q$ satisfy the following identities. For any $a \in \mathbb{F}_q$ let χ_a be the additive character defined as in Theorem 2.2.1. Let $\psi \in \boldsymbol{\mathcal{M}}_q$ be any multiplicative character and $\chi \in \boldsymbol{\mathcal{X}}_q$ be any additive character of $\mathbb{F}_q$. Then we have the following identities

(i) $G(\psi,\chi_{ab}) = \overline{\psi(a)}\cdot G(\psi,\chi_b)$ *for any* $a \in \mathbb{F}_q^*$ *and* $b \in \mathbb{F}_q$.

(ii) $G(\psi,\overline{\chi}) = \psi(-1)\cdot G(\psi,\chi)$.

(iii) $G(\overline{\psi},\chi) = \psi(-1)\overline{G(\psi,\chi)}$.

(iv) $G(\psi,\chi)\cdot G(\overline{\psi},\chi) = \psi(-1)q$ *if* ψ *and* χ *are non-trivial multiplicative and additive characters of* $\mathbb{F}_q$ *respectively.*

(v) $G(\psi^p,\chi_b) = G(\psi,\chi_{\sigma(b)})$ *for any* $b \in \mathbb{F}_q$ *where* $p = \text{char}(\mathbb{F}_q)$ *and* $\sigma(X) = X^p$ *be the Frobenius map of* $\mathbb{F}_q$.

Proof:

(i) For any $\gamma \in \mathbb{F}_q$ we have $\chi_{ab}(\gamma) = \chi_1(ab\cdot\gamma) = \chi_b(a\cdot\gamma)$. Hence

$$G(\psi,\chi_{ab}) = \sum_{\gamma\in\mathbb{F}_q^*}\psi(\gamma)\cdot\chi_{ab}(\gamma) = \sum_{\gamma\in\mathbb{F}_q^*}\psi(\gamma)\cdot\chi_b(a\cdot\gamma)$$

Now take $d = a \cdot \gamma$. Then $\gamma = a^{-1} \cdot d$. Hence

$$G(\psi, \chi_{ab}) = \sum_{d \in \mathbb{F}_q^*} \psi(a^{-1} \cdot d) \cdot \chi_b(d) = \psi(a^{-1}) \sum_{d \in \mathbb{F}_q^*} \psi(d) \cdot \chi_b(d) = \overline{\psi(a)} \cdot G(\psi, \chi_b)$$

Hence we have the first identity.

(ii) Now we have $\chi = \chi_b$ for some suitable $b \in \mathbb{F}_q$. Therefore $\overline{\chi} = \chi_{(-1)\cdot b}$. So by the first identity we have $G(\psi, \overline{\chi}) = \overline{\psi(-1)} \cdot G(\psi, \chi_b) = \psi(-1) \cdot G(\psi, \chi)$. The last equality holds because $\psi(-1) = \pm 1$.

(iii) Now again we have $\overline{\chi} = \chi_b$ for some suitable $b \in \mathbb{F}_q$. Hence $\chi = \chi_{(-1)\cdot b}$. Again by using the first identity we have

$$G(\overline{\psi}, \chi) = G\left(\overline{\psi}, \chi_{(-1)\cdot b}\right) = \overline{\psi(-1)} \cdot G(\overline{\psi}, \chi_b) = \psi(-1) \cdot G(\overline{\psi}, \overline{\chi})$$

Now from the definition of Gaussian sums we have $G(\overline{\psi}, \overline{\chi}) = \overline{G(\psi, \chi)}$. Hence we have the required identity.

(iv) Using part (ii) and (iii) we have

$$G(\psi, \chi) \cdot G(\overline{\psi}, \chi) = G(\psi, \chi) \cdot \psi(-1)\overline{G(\psi, \chi)} = \psi(-1) \cdot |G(\psi, \chi)|^2$$

Since ψ and χ are non-trivial multiplicative and additive characters of $\mathbb{F}_q$ respectively by Theorem 2.3.1 we have $|G(\psi, \chi)|^2 = q$. Hence we have the required identity.

(v) If Tr is the absolute trace function from $\mathbb{F}_q$ to $\mathbb{F}_p$ then for any $\alpha \in \mathbb{F}_q$ we have $\mathrm{Tr}(\alpha) = \mathrm{Tr}(\alpha^p)$. Hence $\chi_1(\alpha) = \chi_1(\alpha^p)$. Hence for any $\alpha \in \mathbb{F}_q$, we get $\chi_b(\alpha) = \chi_1(b \cdot \alpha) = \chi_1(b^p \cdot \alpha^p) = \chi_{\sigma(b)}(\alpha^p)$. Therefore

$$G(\psi^p, \chi_b) = \sum_{\alpha \in \mathbb{F}_q^*} \psi^p(\alpha) \cdot \chi_b(\alpha) = \sum_{\alpha \in \mathbb{F}_q^*} \psi(\alpha^p)\chi_{\sigma(b)}(\alpha^p) = G(\psi, \chi_{\sigma(b)})$$

Therefore we have the required result.

■

Now we can use Gaussian Sums in a variety of context. For any multiplicative character ψ we can express the value of ψ at any nonzero $\alpha \in \mathbb{F}_q$ using Gaussian sums and additive characters. And similarly for any additive character χ of $\mathbb{F}_q$ we can express the value of χ at any nonzero $\alpha \in \mathbb{F}_q$ using Gaussian sums and multiplicative characters.

Now by Theorem 2.2.2(iii) for any $\alpha \in \mathbb{F}_q^*$ we have

$$\begin{aligned}
\psi(\alpha) &= \frac{1}{q} \sum_{\beta \in \mathbb{F}_q^*} \psi(\beta) \sum_{c \in \mathbb{F}_q} \chi_c(\alpha)\overline{\chi_c(\beta)} \\
&= \frac{1}{q} \sum_{c \in \mathbb{F}_q} \chi_c(\alpha) \sum_{\beta \in \mathbb{F}_q^*} \psi(\beta) \cdot \overline{\chi}_c(\beta) \\
&= \frac{1}{q} \sum_{c \in \mathbb{F}_q} \chi_c(\alpha) \cdot G(\psi, \overline{\chi}_c) \\
&= \frac{1}{q} \sum_{\chi \in \boldsymbol{\mathcal{X}}_q} G(\psi, \overline{\chi})\chi(\alpha)
\end{aligned}$$

This may be thought of as the *Fourier expansion* of ψ in terms of the additive characters of $\mathbb{F}_q$, with Gaussian sums appearing as Fourier coefficients.

We can do similar thing for multiplicative characters of $\mathbb{F}_q$. By using Theorem 2.2.5(iii) we get

$$\chi(\alpha) = \frac{1}{q-1} \sum_{\psi \in \boldsymbol{\mathcal{M}}_q} G(\overline{\psi}, \chi)\chi(\alpha)$$

This can be interpreted as the Fourier expansion of the multiplicative characters of $\mathbb{F}_q$ with the Gaussian sums as Fourier coefficients. Hence we have the observations:

Observation 2.9. *For any multiplicative character ψ and additive character χ of $\mathbb{F}_q$ and any $\alpha \in \mathbb{F}_q^*$ we have*

(i) $\psi(\alpha) = \frac{1}{q} \sum_{\chi \in \boldsymbol{X}_q} G(\psi, \overline{\chi})\chi(\alpha)$

(ii) $\chi(\alpha) = \frac{1}{q-1} \sum_{\psi \in \boldsymbol{M}_q} G(\overline{\psi}, \chi)\psi(\alpha)$

We now give another special formula for Gaussian sums for 2-degree extension of finite fields and it has a lot of application. But this has a restriction on the underlying field.

Theorem 2.3.3 Stickelberger's Theorem

Let q be a prime power and $\psi \in \boldsymbol{M}_q$ be a non-trivial multiplicative character of $\mathbb{F}_{q^2}$ of order m such that $m \mid q-1$. Let χ_1 be the canonical additive character of $\mathbb{F}_{q^2}$. Then

$$G(\psi, \chi_1) = \begin{cases} q & \text{if } m \text{ is odd or } \frac{q+1}{m} \text{ is even} \\ -q & \text{if } m \text{ is even and } \frac{q+1}{m} \text{ is odd} \end{cases}$$

Proof: For brevity of notations we write $E = \mathbb{F}_{q^2}$. Let γ be a primitive element of E. So set $g = \gamma^{q+1}$. Therefore $g^{q-1} = 1$ and so we get g as primitive element of $\mathbb{F}_q$. Therefore for all $\alpha \in E^*$ there exists unique $j \in \{0, 1, \ldots, q-1$ and $k \in \{0, \ldots, q+1\}$ such that $\alpha = g^j \cdot \gamma^k$. Hence $\psi(g) = \psi^{q+1}(\gamma) = 1$ as $\mathrm{ord}(\psi) = m \mid q+1$. Therefore we have

$$\begin{aligned} G(\psi, \chi_1) &= \sum_{j=0}^{q-2} \sum_{k=0}^{q} \psi(g^j \cdot \gamma^k) \cdot \chi_1(g^j \cdot \gamma^k) \\ &= \sum_{k=0}^{q} \psi^k(\gamma) \sum_{j=0}^{q-2} \chi_1(g^j \cdot \chi^k) \\ &= \sum_{k=0}^{q} \psi^k(\gamma) \sum_{\alpha \in \mathbb{F}_q^*} \chi_1(\alpha \cdot \chi^k) \end{aligned}$$

Let τ_1 is the canonical additive character of $\mathbb{F}_q$ then for any $\alpha \in \mathbb{F}_q^*$ and $k \in \{0, 1, \ldots, q\}$ we get

$$\chi_1(\alpha \cdot \gamma^k) = \tau_1\left(\mathrm{Tr}_{E/\mathbb{F}_q}(\alpha \cdot \gamma^k)\right) = \tau_1\left(\alpha \cdot \mathrm{Tr}_{E/\mathbb{F}_q}(\gamma^k)\right)$$

Hence we get

$$\sum_{\alpha \in \mathbb{F}_q^*} \chi_1(\alpha \cdot \chi^k) = \sum_{\alpha \in \mathbb{F}_q^*} \tau_1\left(\alpha \cdot \mathrm{Tr}_{E/\mathbb{F}_q}(\gamma^k)\right) = \begin{cases} -1 & \text{if } \mathrm{Tr}_{E/\mathbb{F}_q}(\gamma^k) \neq 0 \\ q-1 & \text{if } \mathrm{Tr}_{E/\mathbb{F}_q}(\gamma^k) = 0 \end{cases}$$

Now since $[E : \mathbb{F}_q] = 2$ we have $\mathrm{Tr}_{E/\mathbb{F}_q}(\gamma^k) = \gamma^k + \gamma^{k \cdot q}$. Therefore

$$\mathrm{Tr}_{E/\mathbb{F}_q}(\gamma^k) = 0 \iff \gamma^k = -\gamma^{k \cdot q} \iff \gamma^{k(q-1)} = -1$$

So if q is odd then $\gamma^{k(q-1)} = -1 \iff k = \frac{q+1}{2}$ and therefore we have

$$\sum_{\alpha \in \mathbb{F}_q^*} \chi_1(\alpha \cdot \chi^k) = \begin{cases} -1 & \text{if } k \neq \frac{q+1}{2} \\ q-1 & \text{if } k = \frac{q+1}{2} \end{cases}$$

And by using Theorem 2.2.2

$$G(\psi, \chi_1) = q \cdot \psi^{q+1/2}(\gamma) - \sum_{k=0}^{q} \psi^k(\gamma) = q \cdot \psi^{(q+1)/2}$$

Recall that ψ has order m and $m \mid q+1$. Therefore $\psi^{(q+1)/2}(\gamma) = \pm 1$. Now if $\frac{q+1}{m}$ is even then we can rewrite

$$\psi^{(q+1)/2}(\gamma) = (\psi^m(\gamma))^{(q+1)/2m} = 1$$

So suppose $\frac{q+1}{m}$ is odd, say $\frac{q+1}{m} = 2\ell + 1$ for some $\ell \geq 0$. Since q is odd, $q+1$ is even, and as $(q+1)/m$ is odd the factor of 2 in $q+1$ must lie entirely in m; in particular m is even. Writing $\frac{q+1}{2} = m\ell + \frac{m}{2}$, we get

$$\psi^{(q+1)/2}(\gamma) = (\psi^m(\gamma))^\ell \cdot \psi^{m/2}(\gamma) = \psi^{m/2}(\gamma).$$

Now $\psi^{m/2}(\gamma)$ is a square root of unity since order of $\psi = m$ and γ is primitive element of E we have $\psi^{(q+1)/2}(\gamma) = -1$. Therefore $G(\psi, \chi_1) = (-1)^{(q+1)/m} \cdot q$. So we have the theorem. ∎

We can give a relation between Gaussian sums a multiplicative and additive character and Gaussian sum of lifted characters. To build that relation we need the theory of L-functions in section 2.5. Later we proved the Davenport-Hasse Theorem which gives this relation.

Before going into this in the next section we will define another hybrid sum called *Jacobi Sums*. Like Gaussian sum this also has many important applications. In the later sections we will use both Gaussian sums and Jacobi Sums to prove many interesting results

§ 2.4 Jacobi Sums

Jacobi sums form another class of exponential sums which are of great importance in algebraic number theory. Jacobi sums are defined on product of many multiplicative characters.

Let $\psi \in \boldsymbol{M}_q$ is a multiplicative character of $\mathbb{F}_q$. So we extend the definition of ψ following subsection 2.2.2 by setting $\psi(0) = 1$ if ψ is the trivial character and $\psi(0) = 0$ if ψ is a non-trivial character. From now on by default we will always use this extended definition of multiplicative characters.

So now let $\lambda_1, \lambda_2, \ldots, \lambda_k \in \boldsymbol{M}_q$ be k multiplicative characters of $\mathbb{F}_q$. Let $\alpha \in \mathbb{F}_q$ be fixed. Then define the sum

$$J_\alpha(\lambda_1, \ldots, \lambda_k) = \sum_{\substack{\alpha_1, \ldots, \alpha_k \in \mathbb{F}_q \\ \alpha_1 + \cdots + \alpha_k = \alpha}} \lambda_1(\alpha_1) \cdot \lambda_2(\alpha_2) \cdots \lambda_k(\alpha_k)$$

Thus in the above sum there are q^{k-1} terms. Now if $\alpha \neq 0$ then define $\beta_i = \alpha^{-1} \cdot \alpha_i$. Then $\beta_1 + \cdots + \beta_k = 1$ and the above summations gives us

$$\begin{aligned} J_\alpha(\lambda_1, \ldots, \lambda_k) &= \sum_{\substack{\beta_1, \ldots, \beta_k \in \mathbb{F}_q \\ \beta_1 + \cdots + \beta_k = 1}} \lambda_1(\alpha \cdot \beta_1) \cdots \lambda_k(\alpha \cdot \beta_2) \\ &= (\lambda_1 \cdots \lambda_k)(\alpha) \sum_{\substack{\beta_1, \ldots, \beta_k \in \mathbb{F}_q \\ \beta_1 + \cdots + \beta_k = 1}} \lambda_1(\beta_1) \cdots \lambda_k(\beta_2) \\ &= (\lambda_1 \cdots \lambda_k)(\alpha) \cdot J_1(\lambda_1, \ldots, \lambda_k) \end{aligned}$$

Because of this relation it is enough to consider only the sums $J_0(\lambda_1, \ldots, \lambda_k)$ and $J_1(\lambda_1, \ldots, \lambda_k)$. The second sum is called the *Jacobi sum* and it has more important applications and so we use a slightly simpler notation.

Definition 2.4.1: Jacobi Sum

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{M}_q$ be multiplicative characters of $\mathbb{F}_q$. Then the Jacobi sum is defined by

$$J(\lambda_1, \ldots, \lambda_k) = \sum_{\substack{\alpha_1, \ldots, \alpha_k \in \mathbb{F}_q \\ \alpha_1 + \cdots + \alpha_k = 1}} \lambda_1(\alpha_1) \cdots \lambda_k(\alpha_k)$$

Now we want to find the value of $J(\lambda_1, \ldots, \lambda_k)$ and $J_0(\lambda_1, \ldots, \lambda_k)$. If $k = 1$ then $J(\lambda) = \lambda(1) = 1$ for any $\lambda \in \boldsymbol{\mathcal{M}}_q$ of $\mathbb{F}_q$. And $J_0(\lambda) = 0$ if λ is non-trivial and its 1 if λ is trivial multiplicative character. Therefore, Jacobi sums are only of interest when $k \geq 2$. Now first we have the following observation

Observation 2.10. *By the definition of the Jacobi sums $J(\lambda_1, \ldots, \lambda_k)$ and $J_0(\lambda_1, \ldots, \lambda_k)$ are independent of the order in which the characters λ_i are listed.*

2.4.1 Evaluating Jacobi Sums

We begin with the degenerate cases where at least one character is trivial, then build up to the general absolute-value formula via the key connection to Gaussian sums.

Lemma 2.4.1 Jacobi Sums with Trivial Characters

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{\mathcal{M}}_q$ are multiplicative characters of $\mathbb{F}_q$. Then

(i) If $\lambda_1 = \cdots = \lambda_k = \psi_0$ where ψ_0 is the trivial multiplicative character then

$$J(\lambda_1, \ldots, \lambda_k) = J_0(\lambda_1, \ldots, \lambda_k) = q^{k-1}$$

(ii) If some but not all λ_i's are trivial then

$$J(\lambda_1, \ldots, \lambda_k) = J_0(\lambda_1, \ldots, \lambda_k) = 0$$

Proof:

(i) Since all λ_i's are trivial for all $\alpha \in \mathbb{F}_q$, $\lambda(\alpha) = 1$. Hence each term of both the Jacobi sums give 1. Now the sum is over q^{k-1} choices of $(\alpha_1, \ldots, \alpha_k) \in \mathbb{F}_q^k$. Hence $J(\lambda_1, \ldots, \lambda_k) = J_0(\lambda_1, \ldots, \lambda_k) = q^{k-1}$.

(ii) WLOG suppose $\lambda_1, \ldots, \lambda_t$ are non-trivial multiplicative characters where $1 \leq t < k$ and $\lambda_{t+1}, \ldots, \lambda_k$ are all trivial. Then we have

$$\begin{aligned} J(\lambda_1, \ldots, \lambda_k) &= \sum_{\substack{\alpha_1, \ldots, \alpha_k \in \mathbb{F}_q \\ \alpha_1 + \cdots + \alpha_k = 1}} \lambda_1(\alpha_1) \cdots \lambda_k(\alpha_k) \\ &= \sum_{\substack{\alpha_1, \ldots, \alpha_k \in \mathbb{F}_q \\ \alpha_1 + \cdots + \alpha_k = 1}} \lambda_1(\alpha_1) \cdots \lambda_t(\alpha_t) \end{aligned}$$

Now for any fixed $\alpha_1, \ldots, \alpha_t \in \mathbb{F}_q$ there are q^{k-t-1} solutions of $(\alpha_{t+1}, \ldots, \alpha_k)$ such that $\alpha_{t+1} + \cdots + \alpha_k = 1 - (\alpha_1 + \cdots + \alpha_t)$. Therefore we get

$$\begin{aligned} J(\lambda_1, \ldots, \lambda_k) &= q^{k-h-1} \sum_{\alpha_1, \ldots, \alpha_t \in \mathbb{F}_q} \lambda_1(\alpha_1) \cdots \lambda_t(\alpha_t) \\ &= q^{k-h-1} \prod_{i=1}^{t} \left(\sum_{\alpha_i \in \mathbb{F}_q} \lambda_i(\alpha_i) \right) \\ &= 0 \qquad \text{[By Theorem 2.2.5(i)]} \end{aligned}$$

With a very similar argument for $J_0(\lambda_1, \ldots, \lambda_k)$ we get the identity.

■

Before going to the case where all λ_i's are non-trivial we will show that a relation between $J_0(\lambda_1, \ldots, \lambda_k)$ and $J(\lambda_1, \ldots, \lambda_k)$. This will show that studying $J(\lambda_1, \ldots, \lambda_k)$ suffices to study all kinds of Jacobi sums.

Lemma 2.4.2 Relation between J_0 and J

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{\mathcal{M}}_q$ are multiplicative characters of $\mathbb{F}_q$ with λ_k is non-trivial. Then

$$J_0(\lambda_1, \ldots, \lambda_k) = \begin{cases} 0 & \text{if } \prod_{i=1}^k \lambda_i \text{ is non-trivial} \\ \lambda_k(-1) \cdot (q-1) \cdot J(\lambda_1, \ldots, \lambda_{k-1}) & \text{if } \prod_{i=1}^k \lambda_i \text{ is trivial} \end{cases}$$

Proof: For $k = 1$, we already showed $J_0(\lambda_1) = \lambda_1(0) = 0$ as λ_1 is non-trivial by hypothesis. So suppose $k \geq 2$. Then we have

$$\begin{aligned} J_0(\lambda_1, \ldots, \lambda_k) &= \sum_{\alpha \in \mathbb{F}_q} J_{-\alpha}(\lambda_1, \ldots, \lambda_{k-1}) \cdot \lambda_k(\alpha) \\ &= \sum_{\alpha \in \mathbb{F}_q} \left(\prod_{i=1}^{k-1} \lambda_i \right)(-\alpha) \cdot J(\lambda_1, \ldots, \lambda_{k-1}) \cdot \lambda_k(\alpha) \\ &= J(\lambda_1, \ldots, \lambda_{k-1}) \sum_{\alpha \in \mathbb{F}_q} \left(\prod_{i=1}^{k-1} \lambda_i \right)(-\alpha) \cdot \lambda_k(\alpha) \\ &= \left(\prod_{i=1}^{k-1} \lambda_i \right)(-1) \cdot J(\lambda_1, \ldots, \lambda_{k-1}) \sum_{\alpha \in \mathbb{F}_q} \left(\prod_{i=1}^{k} \lambda_i \right)(\alpha) \end{aligned}$$

Now by Theorem 2.2.5(i) if $\prod_{i=1}^k \lambda_i$ is nontrivial then the sum 0, so the first case is shown and if $\prod_{i=1}^k \lambda_i$ is trivial then the last sum is $q-1$. Now since $\prod_{i=1}^k \lambda_i$ is trivial $\prod_{i=1}^{k-1} \lambda_i = \overline{\lambda_k}(-1) = \lambda_k(-1)$. So we have

$$J_0(\lambda_1, \ldots, \lambda_k) = \lambda_k(-1) \cdot (q-1) \cdot J(\lambda_1, \ldots, \lambda_{k-1})$$

Hence we have the lemma. ∎

Hence the only case for $J_0(\lambda_1, \ldots, \lambda_k)$ for which we still don't know the absolute value is the case when all λ_i's are non-trivial and $\prod_{i=1}^k \lambda_i$ is trivial. Since we understand Gaussian sums very well we will first build an important connection between Jacobi sums and Gaussian sums that will allow us to calculate the absolute value of Jacobi sums.

Theorem 2.4.3 Jacobi–Gaussian Sum Relation

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{\mathcal{M}}_q$ are non-trivial multiplicative characters of $\mathbb{F}_q$ and $\chi \in \boldsymbol{\mathcal{X}}_q$ is a non-trivial additive character of $\mathbb{F}_q$.

(i) If $\prod_{i=1}^k$ is non-trivial then

$$J(\lambda_1, \ldots, \lambda_k) = \frac{G(\lambda_1, \chi) \cdot G(\lambda_2, \chi) \cdots G(\lambda_k, \chi)}{G(\lambda_1 \cdots \lambda_k, \chi)}$$

(ii) If $\prod_{i=1}^k$ is trivial then

$$\begin{aligned} J(\lambda_1, \ldots, \lambda_k) &= -\lambda_k(-1) \cdot J(\lambda_1, \ldots, \lambda_{k-1}) \\ &= -\frac{1}{q} \prod_{i=1}^{k} G(\lambda_i, \chi) \end{aligned}$$

Proof: Since each λ_i's are non-trivial we have $\lambda_i(0) = 0$. By definition of

$$G(\lambda_i, \chi) = \sum_{\alpha_i \in \mathbb{F}_q^*} \lambda_i \alpha \cdot \chi(\alpha) = \sum_{\alpha_i \in \mathbb{F}_q} \lambda_i \alpha \cdot \chi(\alpha)$$

So now we will try to express the product of Gaussian sums as product of a Jacobi and a Gaussian sum.

$$\begin{aligned} G(\lambda_1,\chi)\cdots G(\lambda_k,\chi) &= \prod_{i=1}^{k}\left(\sum_{\alpha_i\in\mathbb{F}_q}\lambda_i(\alpha_i)\cdot\chi(\alpha_i)\right)\\ &= \sum_{\alpha_1,\ldots,\alpha_k\in\mathbb{F}_q}\left(\prod_{i=1}^{k}\lambda_i(\alpha_i)\right)\cdot\chi\left(\sum_{i=1}^{k}\alpha_i\right)\\ &= \sum_{\alpha\in\mathbb{F}_q}\chi(\alpha)\sum_{\substack{\alpha_1,\ldots,\alpha_k\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_k=\alpha}}\left(\prod_{i=1}^{k}\lambda_i(\alpha_i)\right)\\ &= \sum_{\alpha\in\mathbb{F}_q}\chi(\alpha)\cdot J_\alpha(\lambda_1,\ldots,\lambda_k) \end{aligned} \tag{2.5}$$

Now if $\prod_{i=1}^{k}\lambda_i$ is non-trivial then by Lemma 2.4.2 we get $J_0(\lambda_1,\ldots,\lambda_k)=0$. Now since for all $\alpha\in\mathbb{F}_q^*$ we have $J_\alpha(\lambda_1,\ldots,\lambda_k)=\left(\prod_{i=1}^{k}\lambda_i\right)(\alpha)\cdot J(\lambda_1,\ldots,\lambda_k)$. Therefore we have

$$\prod_{i=1}^{k}G(\lambda_i,\chi)=J(\lambda_1,\ldots,\lambda_k)\sum_{\alpha\in\mathbb{F}_q^*}\left(\prod_{i=1}^{k}\lambda_i\right)(\alpha)\cdot\chi(\alpha)=J(\lambda_1,\ldots,\lambda_k)\cdot G\left(\prod_{i=1}^{k}\lambda_i,\chi\right)$$

Since χ is non-trivial and $\prod_{i=1}^{k}\lambda_i$ is non-trivial we have $G(\lambda_1\cdots\lambda_k,\chi)\neq 0$ by Theorem 2.3.1. So we have the first result.

Now if $\prod_{i=1}^{k}\lambda_i$ is trivial, then for all $\alpha\in\mathbb{F}_q^*$ we have $J_\alpha(\lambda_1,\ldots,\lambda_k)=J(\lambda_1,\ldots,\lambda_k)$. Hence rewriting the equation in (2.5) we get

$$\begin{aligned} \prod_{i=1}^{k}G(\lambda_i,\chi) &= \sum_{\alpha\in\mathbb{F}_q}\chi(\alpha)\cdot J_\alpha(\lambda_1,\ldots,\lambda_k)\\ &= J_0(\lambda_1,\ldots,\lambda_k)+J(\lambda_1,\ldots,\lambda_k)\sum_{\alpha\in\mathbb{F}_q}\chi(\alpha)\\ &= J_0(\lambda_1,\ldots,\lambda_k)+(q-1)J(\lambda_1,\ldots,\lambda_k) && \text{[By Theorem 2.2.2(i)]}\\ &= \sum_{\alpha\in\mathbb{F}_q}J(\lambda_1,\ldots,\lambda_k)\\ &= \sum_{\alpha_1,\ldots,\alpha_k\in\mathbb{F}_q}\prod_{i=1}^{k}\lambda_i(\alpha_i)\\ &= \prod_{i=1}^{k}\left(\sum_{\alpha\in\mathbb{F}_q}\lambda_i(\alpha_i)\right)=0 && \text{[By Theorem 2.2.5(i)]} \end{aligned}$$

Since λ_k is non-trivial and $\prod_{i=1}^{k}\lambda_i$ is trivial using Lemma 2.4.2 we have

$$\lambda_k(-1)\cdot(q-1)\cdot J(\lambda_1,\ldots,\lambda_{k-1})+(q-1)J(\lambda_1,\ldots,\lambda_k)=0\iff J(\lambda_1,\ldots,\lambda_k)=-\lambda_k(-1)\cdot J(\lambda_1,\ldots,\lambda_{k-1})$$

So we already got the first expression for the second case. Since $\prod_{i=1}^{k-1}\lambda_i$ is non-trivial by induction hypothesis we get

$$\begin{aligned} \lambda_k(-1)\cdot J(\lambda_1,\ldots,\lambda_{k-1}) &= \lambda_k(-1)\frac{G(\lambda_1,\chi)\cdots G(\lambda_{k-1},\chi)}{G(\lambda_1\cdots\lambda_{k-1},\chi)}\\ &= \lambda_k(-1)\frac{G(\lambda_1,\chi)\cdots G(\lambda_k,\chi)}{G(\overline{\lambda_k},\chi)\cdot G(\lambda_k,\chi)} && \text{[Since } \lambda_1\cdots\lambda_k \text{ trivial, } \lambda_1\cdots\lambda_{k-1}=\overline{\lambda_k}\text{]}\\ &= \lambda_k(-1)\frac{G(\lambda_1,\chi)\cdots G(\lambda_k,\chi)}{\lambda_k(-1)q} && \text{[By Lemma 2.3.2(iv)]}\\ &= \frac{1}{q}\cdot G(\lambda_1,\chi)\cdots G(\lambda_k,\chi) \end{aligned}$$

Hence we obtain the second expression. Therefore we have the theorem. ■

Now we have all the results to calculate the absolute value of the Jacobi sum for non-trivial multiplicative characters. So we have the following theorem.

Theorem 2.4.4 Absolute Value of Non-trivial Jacobi Sums

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{\mathcal{M}}_q$ be non-trivial multiplicative characters of $\mathbb{F}_q$. Then

$$|J(\lambda_1, \ldots, \lambda_k)| = \begin{cases} q^{(k-1)/2} & \text{If } \prod_{i=1}^{k} \lambda_i \text{ is non-trivial} \\ q^{(k-2)/2} & \text{If } \prod_{i=1}^{k} \lambda_i \text{ is trivial} \end{cases}$$

Proof: Suppose $\prod_{i=1}^{k} \lambda_i$ is non-trivial. Then for any non-trivial additive character $\chi \in \boldsymbol{\mathcal{X}}_q$ of $\mathbb{F}_q$ by Theorem 2.4.3 we get

$$J(\lambda_1, \ldots, \lambda_k) = \frac{G(\lambda_1, \chi) \cdot G(\lambda_2, \chi) \cdots G(\lambda_k, \chi)}{G(\lambda_1 \cdots \lambda_k, \chi)}$$

Since each λ_i non-trivial by Theorem 2.3.1 we get $|G(\lambda_i, \chi)| = q^{1/2}$ for all $i \in [k]$ and $|G(\lambda_1 \cdots \lambda_k, \chi)| = q^{1/2}$. Hence we get

$$|J(\lambda_1, \ldots, \lambda_k)| = \frac{\prod_{i=1}^{k} |G(\lambda_i, \chi)|}{|G(\lambda_1 \cdots \lambda_k, \chi)|} = q^{(q-1)/2}$$

Now suppose $\prod_{i=1}^{k} \lambda_i$ is trivial. Then again by Theorem 2.4.3 we have

$$J(\lambda_1, \ldots, \lambda_k) = -\lambda_k(-1) \cdot J(\lambda_1, \ldots, \lambda_{k-1})$$

Now since $\prod_{i=1}^{k} \lambda_i$ is trivial and λ_k is non-trivial we have $\prod_{i=1}^{k-1} \lambda_i$ is non-trivial. Hence by using the first case we have $|J(\lambda_1, \ldots, \lambda_{k-1})| = q^{(k-2)/2}$. Hence $J(\lambda_1, \ldots, \lambda_k)| = q^{(k-2)/2}$. Therefore we have the result. ■

Using the value of $J(\lambda_1, \ldots, \lambda_k)$ we can also get the absolute value of $J_0(\lambda_1, \ldots, \lambda_k)$ when all λ_i's are non-trivial and $\prod_{i=1}^{k} \lambda_i$ is trivial.

Corollary 2.4.5 Absolute Value of J_0 for Non-trivial Characters

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{\mathcal{M}}_q$ be non-trivial multiplicative characters of $\mathbb{F}_q$ and $\prod_{i=1}^{k} \lambda_i$ is trivial. Then

$$|J_0(\lambda_1, \ldots, \lambda_k)| = (q-1)q^{(k-2)/2}$$

Proof: Using Lemma 2.4.2 we have $J_0(\lambda_1, \ldots, \lambda_k) = \lambda_k(-1) \cdot (q-1) \cdot J(\lambda_1, \ldots, \lambda_{k-1})$. Now since $\prod_{i=1}^{k} \lambda_i$ is trivial and λ_k is non-trivial we have $\prod_{i=1}^{k-1} \lambda_i$ is non-trivial. Therefore by Theorem 2.4.4 we have $|J(\lambda_1, \ldots, \lambda_{k-1})| = q^{(k-2)/2}$. Therefore together we have the corollary. ■

Combining Lemma 2.4.1, Lemma 2.4.2, Theorem 2.4.4, and the corollary above, we obtain a complete description of the absolute values of Jacobi sums in all cases.

Theorem 2.4.6 Complete Evaluation of Jacobi Sums

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{\mathcal{M}}_q$ be multiplicative characters of $\mathbb{F}_q$. Then

(i) If λ_i is trivial for all $i \in [k]$ then $J(\lambda_1, \ldots, \lambda_k) = J_0(\lambda_1, \ldots, \lambda_k) = q^{k-1}$.

(ii) If some but not all λ_i's are trivial then $J(\lambda_1, \ldots, \lambda_k) = J_0(\lambda_1, \ldots, \lambda_k) = 0$.

(iii) If all λ_i's are non-trivial and $\prod_{i=1}^{k} \lambda_i$ is non-trivial then $|J(\lambda_1, \ldots, \lambda_k)| = q^{(k-1)/2}$ and $J_0(\lambda_1, \ldots, \lambda_k) = 0$.

(iv) If all λ_i's are non-trivial and $\prod_{i=1}^{k} \lambda_i$ is trivial then $|J(\lambda_1, \ldots, \lambda_k)| = q^{(k-2)/2}$ and $|J_0(\lambda_1, \ldots, \lambda_k)| = (q-1)q^{(k-2)/2}$.

2.4.2 The Davenport–Hasse Relations

The case $k = 2$ is the one that appears in most arithmetic applications. The central result here is the Davenport–Hasse product formula, which expresses a high power of a Gaussian sum as a product of Jacobi sums. We first establish a useful preliminary identity.

Lemma 2.4.7 Gaussian Sum Power via Jacobi Products

Let λ be a multiplicative character of $\mathbb{F}_q$ of order $m \geq 2$ and let χ be a non-trivial additive character of $\mathbb{F}_q$. Then

$$G(\lambda, \chi)^m = \lambda(-1) \cdot q \cdot J(\lambda, \lambda) \cdot J(\lambda, \lambda^2) \cdots J(\lambda, \lambda^{m-2})$$

Proof: If $m = 2$ then $\overline{\lambda} = \lambda$ and so $G(\lambda, \chi)^2 = G(\lambda, \chi) \cdot G(\overline{\lambda}, \chi)$ and the right hand side becomes $\lambda(-1) \cdot q$ which we already have from Lemma 2.3.2(iv). So now suppose $m \geq 3$.

Then from Theorem 2.4.3 we have for all $i \in [m-2]$

$$J(\lambda, \lambda^i) = \frac{G(\lambda, \chi) \cdot G(\lambda^i, \chi)}{G(\lambda^{i+1}, \chi)}$$

Therefore taking product of all $J(\lambda, \lambda^i)$ for all $i \in [m-2]$ we get

$$\begin{aligned} \prod_{i=1}^{m-2} J(\lambda, \lambda^i) &= \prod_{i=1}^{m-2} \frac{G(\lambda, \chi) \cdot G(\lambda^i, \chi)}{G(\lambda^{i+1}, \chi)} \\ &= G(\lambda, \chi)^{m-2} \, \frac{G(\lambda, \chi)}{G(\lambda^{m-1}, \chi)} \\ &= \frac{G^{m-1}(\lambda, \chi)}{G(\lambda^{m-1}, \chi)} \end{aligned}$$

Since $\mathrm{ord}(\lambda) = m$ we have $\lambda^{m-1} = \overline{\lambda}$. Therefore we have

$$\prod_{i=1}^{m-2} J(\lambda, \lambda^i) = \frac{G^{m-1}(\lambda, \chi)}{G(\lambda^{m-1}, \chi)} = \frac{G^m(\lambda, \chi)}{G(\overline{\lambda}, \chi) \cdot G(\lambda, \chi)} = \frac{G^m(\lambda, \chi)}{\lambda(-1) \cdot q}$$

Therefore we have the lemma. ∎

Now we give another result of relation between Gaussian sums for the $k = 2$ case. This gives relation between two Gaussian sums on non-trivial multiplicative characters of different orders using Jacobi sums. We will not prove this theorem as it is far too complicated to be included in here.

Theorem 2.4.8 Davenport–Hasse Product Formula

Let $\psi, \lambda \in \boldsymbol{\mathcal{M}}_q$ be multiplicative characters of $\mathbb{F}_q$ such that $\mathrm{ord}(\lambda) = m \geq 2$ and ψ^m is non-trivial. Let $\chi_b \in \boldsymbol{\mathcal{X}}_q$ be a non-trivial additive character of $\mathbb{F}_q$ as defined in Theorem 2.2.1. Then

$$\frac{G(\psi, \chi_b)^m}{G(\psi^m, \chi_{mb})} = \prod_{j=1}^{m-1} J(\psi, \lambda^j).$$

Over the prime field $\mathbb{F}_p$ the Davenport–Hasse relation admits a more explicit form, in which the root of unity appearing on the right hand side is written in terms of a primitive root modulo p. This form appears as Theorem 3 in the paper of Huard, Spearman and Williams.

Theorem 2.4.9 Davenport–Hasse Relation over $\mathbb{F}_p$, [HSW95, Theorem 3]

Let $m \geq 2$ and $n \geq 1$ be integers, let p be an odd prime with $p \equiv 1 \bmod mn$, and set $f = (p-1)/mn$. Let g be a primitive root modulo p, and for any integer $l \not\equiv 0 \bmod p$ let $\mathrm{ind}_g(l)$ denote the least nonnegative integer with $g^{\mathrm{ind}_g(l)} \equiv l \bmod p$. Set $\beta_{mn} = e^{2\pi i/mn}$ and let $\lambda \in \boldsymbol{\mathcal{M}}_p$ be the multiplicative character of order mn determined by $\lambda(g) = \beta_{mn}$. Let $\chi \in \boldsymbol{\mathcal{X}}_p$ be a non-trivial additive character of $\mathbb{F}_p$. Then for every $t = 1, \ldots, m-1$,

$$\frac{G(\lambda^{tn}, \chi) \cdot \prod_{j=1}^{n-1} G(\lambda^{mj}, \chi)}{\prod_{j=0}^{n-1} G(\lambda^{mj+t}, \chi)} = \beta_{mn}^{nt \cdot \mathrm{ind}_g(n)},$$

equivalently,

$$\prod_{j=1}^{n-1} \frac{J(\lambda^{mj}, \lambda^t)}{J(\lambda^{tj}, \lambda^t)} = \beta_{mn}^{nt \cdot \mathrm{ind}_g(n)}.$$

As an application of the Davenport–Hasse product formula, together with the evaluation of Jacobi sums in Theorem 2.4.6, we obtain the following closed form for the product of Gaussian sums on the m characters $\psi \cdot \lambda^i$ twisting ψ by the powers of λ. The two cases reflect the parity of m, with the even case picking up an extra quadratic Gaussian sum $G(\eta, \chi_b)$ where η is the quadratic character of $\mathbb{F}_q$.

Corollary 2.4.10 Product of Twisted Gaussian Sums

Let $\psi, \lambda \in \boldsymbol{\mathcal{M}}_q$ be multiplicative characters of $\mathbb{F}_q$ such that $\mathrm{ord}(\lambda) = m \geq 2$ and ψ^m is non-trivial. Let $\chi_b \in \boldsymbol{\mathcal{X}}_q$ be a non-trivial additive character of $\mathbb{F}_q$ as defined in Theorem 2.2.1.

$$\prod_{j=0}^{m} G(\psi \cdot \lambda^j, \chi_b) = \begin{cases} q^{(m-1)/2} G(\psi^m, \chi_{mb}) & \text{if } m \text{ is odd} \\ (-1)^{(q-1)(m-2)/8} \cdot q^{(m-2)/2} \cdot G(\eta, \chi_b) \cdot G(\psi^m, \chi_{mb}) & \text{if } m \text{ is even} \end{cases}$$

Proof: Since $\mathrm{ord}(\lambda) \geq 2$ and ψ^m is non-trivial we have $\psi \cdot \lambda^j$ is non-trivial for all $j \in [m-1]$. By Davenport-Hasse Product Formula we have

$$\frac{G(\psi, \chi_b)^m}{G(\psi^m, \chi_{mb})} = \prod_{j=1}^{m-1} J(\psi, \lambda^j)$$

Now by using Theorem 2.4.3 on the right hand side of this expression we get

$$\frac{G(\psi, \chi_b)^m}{G(\psi^m, \chi_{mb})} = \prod_{j=1}^{m-1} \frac{G(\psi, \chi_b) \cdot G(\lambda^j, \chi_b)}{G(\psi \cdot \lambda^i, \chi_b)} = G(\psi, \chi_b)^{m-1} \prod_{i=1}^{j} \frac{G(\lambda^j, \chi_b)}{G(\psi \cdot \lambda^i, \chi_b)}$$

So together we have

$$\prod_{j=0}^{m-1} G(\psi \cdot \lambda^j, \chi_b) = G(\psi^m, \chi_{mb}) \prod_{j=1}^{m-1} G(\lambda^j, \chi_b) \tag{2.6}$$

Notice we already obtained the left hand side and one term in the right hand side correctly. So now the value at right only depends on the product of $G(\lambda^j, \chi_b)$'s.

Now if m is odd then we have

$$\begin{aligned}\prod_{j=1}^{m-1} G(\lambda^j, \chi_b) &= \prod_{j=1}^{(m-1)/2} G(\lambda^j, \chi_b) G(\lambda^{m-j}, \chi_b) \\ &= \prod_{j=1}^{(m-1)/2} G(\lambda^j, \chi_b) \cdot G(\overline{\lambda^j}, \chi_b) && [\text{As } \mathrm{ord}(\lambda) = m = \text{odd}] \\ &= \prod_{j=1}^{(m-1)/2} \lambda^j(-1) \cdot q && [\text{By Lemma 2.3.2(iv)}] \\ &= q^{(m-1)/2} \prod_{j=1}^{(m-1)/2} \lambda^j(-1)\end{aligned}$$

Since m is odd by Observation 2.6 we have $\lambda(-1) = 1$. Therefore

$$\prod_{j=1}^{m-1} G(\lambda^j, \chi_b) = q^{(m-1)/2}$$

So we have the first case together with (2.6). Now for the remaining case of m being even we have

$$\begin{aligned}\prod_{j=1}^{m-1} G(\lambda^j, \chi_b) &= G(\lambda^{m/2}, \chi_b) \prod_{j=1}^{(m-2)/2} G(\lambda^j, \chi_b) G(\lambda^{m-j}, \chi_b) \\ &= G(\eta, \chi_b) \prod_{j=1}^{(m-2)/2} G(\lambda^j, \chi_b) \cdot G(\overline{\lambda^j}, \chi_b) && [\text{As } \mathrm{ord}(\lambda) = m = \text{even}] \\ &= G(\eta, \chi_b) \prod_{j=1}^{(m-2)/2} \lambda^j(-1) \cdot q && [\text{By Lemma 2.3.2(iv)}] \\ &= q^{(m-2)/2} \cdot G(\eta, \chi_b) \prod_{j=1}^{(m-2)/2} \lambda^j(-1)\end{aligned}$$

Again by Observation 2.6 since m is even if $\frac{q-1}{m}$ is odd then $\lambda(-1) = 1$ and if $\frac{q-1}{m}$ is even then $\lambda(-1) = 1$. Hence we can write $\lambda(-1) = (-1)^{(q-1)/m}$. Therefore

$$\prod_{j=1}^{m-1} G(\lambda^j, \chi_b) = (-1)^{(q-1)(m-2)/8} \cdot q^{(m-2)/2} \cdot G(\eta, \chi_b)$$

So together with (2.6) we have the result. ∎

Corollary 2.4.11

Let $\psi, \lambda \in \boldsymbol{M}_q$ be multiplicative characters of $\mathbb{F}_q$ such that $\mathrm{ord}(\lambda) = m \geq 2$ and ψ^m is non-trivial. Let $\chi \in \boldsymbol{X}_q$ be any non-trivial additive character of $\mathbb{F}_q$. Then

$$\prod_{j=0}^{m-1} G(\psi \cdot \lambda^j) = \psi^{-m}(m) \cdot G(\psi^m, \chi) \prod_{j=1}^{m-1} G(\lambda^j, \chi)$$

where for any $\gamma \in \mathbb{F}_q$, $\chi_m(\gamma) = \chi(m \cdot \gamma)$ similar to as defined in Theorem 2.2.1.

Proof: By Davenport-Hasse Product Formula we have

$$\frac{G(\psi, \chi)^m}{G(\psi^m, \chi)} = \prod_{j=1}^{m} J(\psi, \lambda^j)$$

Now using Theorem 2.4.3 for any $j \in \{m-1\}$ we have

$$J(\psi, \lambda^j) = \frac{G(\psi, \chi) \cdot G(\lambda^j, \chi}{G(\psi \cdot \lambda^j, \chi)}$$

So replacing $J(\psi, \lambda^j)$ we get

$$\frac{G(\psi, \chi)^m}{G(\psi^m, \chi)} = \prod_{j=1}^{m-1} \frac{G(\psi, \chi) \cdot G(\lambda^j, \chi}{G(\psi \cdot \lambda^j, \chi)} = G(\psi, \chi)^{m-1} \prod_{j=1}^{m-1} \frac{G(\lambda^j, \chi}{G(\psi \cdot \lambda^j, \chi)}$$

So we get the final relation

$$\prod_{j=0}^{m-1} G(\psi \cdot \lambda^j) = G(\psi^m, \chi_m) \prod_{j=1}^{m-1} G(\lambda^j, \chi) \tag{2.7}$$

Now by Lemma 2.3.2(i) we have $G(\psi^m, \chi_m) = \psi^{-m}(m) G(\psi^m, \chi)$. So replacing $G(\psi^m, \chi_m)$ we get the theorem. ∎

Since for the trivial multiplicative character $\psi_0 \in \boldsymbol{\mathcal{M}}_q$ we have $\lambda^0 = \psi_0$. Therefore $\prod_{j=0}^{m-1} G(\lambda^j, \chi) = -\prod_{j=1}^{m-1} G(\lambda^j, \chi)$ by Theorem 2.3.1. Therefore from (2.7) we have

$$\prod_{j=0}^{m-1} G(\psi \cdot \lambda^j) = -G(\psi^m, \chi_m) \prod_{j=0}^{m-1} G(\lambda^j, \chi)$$

So we get the following observation:

Observation 2.11. *Let $\psi, \lambda \in \boldsymbol{\mathcal{M}}_q$ be multiplicative characters of $\mathbb{F}_q$ such that $\mathrm{ord}(\lambda) = m \geq 2$ and ψ^m is non-trivial. Let $\chi \in \boldsymbol{\mathcal{X}}_q$ be any non-trivial additive character of $\mathbb{F}_q$. Then*

$$\prod_{j=0}^{m-1} G(\psi \cdot \lambda^j) = -\psi^{-m}(m) \cdot G(\psi^m, \chi) \prod_{j=0}^{m-1} G(\lambda^j, \chi)$$

2.4.3 Application: Fermat's Two-Square Theorem

A classical application of the arithmetic of Jacobi sums is a short representation-theoretic proof of a celebrated theorem of Fermat: every prime $p \equiv 1 \bmod 4$ is a sum of two integer squares. The key point is that when λ is a multiplicative character of $\mathbb{F}_p$ of order 4, the Jacobi sum $J(\lambda, \eta)$ lies in the ring of Gaussian integers $\mathbb{Z}[i]$ and has absolute value exactly $\sqrt{p}$.

Theorem 2.4.12 Fermat's Two-Square Theorem

Let p be an odd prime with $p \equiv 1 \bmod 4$. Then there exist integers $A, B \in \mathbb{Z}$ such that $p = A^2 + B^2$.

Proof: Let $\lambda, \eta \in \boldsymbol{\mathcal{M}}_q$ are multiplicative characters such that $\mathrm{ord}(\lambda) = 4$ and η is the quadratic character. Since $\mathrm{ord}(\lambda) = 4$, λ takes values in $\{\pm 1, \pm i\} \subseteq \mathbb{Z}[i]$, and since η takes values in $\{\pm 1\} \subseteq \mathbb{Z}$, the Jacobi sum

$$J(\lambda, \eta) = \sum_{\alpha \in \mathbb{F}_p} \lambda(\alpha) \cdot \eta(1-\alpha) \in \mathbb{Z}[i].$$

Write $J(\lambda, \eta) = A + iB$ with $A, B \in \mathbb{Z}$. Since λ and η are both non-trivial and $\lambda \cdot \eta = \lambda^3$ is also non-trivial (as $\mathrm{ord}(\lambda) = 4$), Theorem 2.4.4 gives $|J(\lambda, \eta)| = p^{1/2}$; equivalently $A^2 + B^2 = p$. ∎

We give another proof of this theorem in subsection 2.9.2 during the discussion of Jacobsthal Sums in section 2.9.

§ 2.5 L-functions

The characters of cyclic groups defined in Lemma 2.1.1 are precisely the classical *Dirichlet characters*. Dirichlet characters are of paramount importance in analytic number theory: they appear in the study of primes in arithmetic progressions, the distribution of prime ideals, and the analytic continuation of L-functions. In the finite-field setting, the multiplicative characters of $\mathbb{F}_q$ play an analogous role, and the associated L-functions encode deep arithmetic information about character sums over $\mathbb{F}_q$. The treatment here focuses on the aspects directly relevant to the Weil bounds; readers interested in the general theory of L-functions in its full depth are encouraged to consult Iwaniec and Kowalski [IK12] for the analytic theory, and Miller and Takloo-Bighash [Mil06] for a broader introduction to modern number theory and the role of L-functions within it.

Before specialising to the function-field setting, it is helpful to pause and describe the general analytic framework into which all of these objects fit. The central analytic device is a *Dirichlet series*

Definition 2.5.1: Dirichlet Series

Let $a(n)$ be a sequence of complex numbers and $s \in \mathbb{C}$ be any complex number. Then a series of the form

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s}, \qquad s \in \mathbb{C},$$

is called a Dirichlet series.

Two very familiar Dirichlet series organize a vast section of analytic number theory:

1. **Zeta functions** arise when the coefficients are the trivial ones, $a_n \equiv 1$, so that the Dirichlet series simply counts the size of some arithmetic object. The prototype is the *Riemann zeta function* $\zeta(s) = \sum_n n^{-s}$, which encodes the multiplicative structure of $\mathbb{Z}$.

2. L**-functions** arise when the coefficients are *twisted* by a character Λ, giving $L(s, \Lambda) = \sum_{h:\text{ monic}} \Lambda(h)/\mathfrak{n}(h)^s$. The character Λ is what allows one to isolate arithmetic information that a zeta function averages over. Since L-functions are defined using polynomials over a finite field and then we can talk about the prime analogue in polynomial rings i.e. irreducible polynomials. Taking the trivial character $\Lambda \equiv 1$ recovers ζ up to finitely many Euler factors, so zeta functions are the "character-free" special case of L-functions.

The rest of this section unfolds this dictionary in two parallel worlds. First we recall the classical story over $\mathbb{Q}$ — the Riemann zeta function, its zeros, and its L-function generalisations — so that the relevant analytic phenomena are in view. Like in the case we go to $\mathbb{Q}$ the fraction field of $\mathbb{Z}$, we develop the function-field $\mathbb{F}_q(X)$ analog: the zeta function associated to $\mathbb{F}_q[X]$ and its twists by multiplicative characters of $\mathbb{F}_q$. The zeta function turns out to be a rational function of q^{-s}, and the analog of the Riemann Hypothesis is a theorem (due to Weil) rather than an open problem.

2.5.1 The Classical Riemann Zeta Function

The prototype of everything that follows is the *Riemann zeta function*.

Definition 2.5.2: Riemann Zeta Function

For any complex number $s \in \mathbb{C}$, the zeta function at s is

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

which converges absolutely in the right half-plane $\Re(s) > 1$. In fact for $s \in \mathbb{C}$ with $\Re(s) > 1$ the product $\prod_{n=2}^{\infty} \zeta(n)$ also converges. Most proofs in this subsection are omitted, especially those relying on contour integration or the Mellin transform; the emphasis is on statements together with their arithmetic meaning.

Theorem 2.5.1 Euler Product for $\zeta(s)$

For $s \in \mathbb{C}$ with $\Re(s) > 1$,

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}.$$

Proof: Since $\mathbb{Z}$ is a UPD (unique factorization domain) every positive integer n can be written uniquely as a product of powers of positive prime numbers.

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p\text{: prime}} \left(\sum_{i=0}^{\infty} \frac{1}{p^{i \cdot s}} \right) = \prod_{p\text{: prime}} (1 - p^{-s})^{-1}$$

So we have the lemma. ∎

The identity is the analytic translation of *unique factorisation*: expanding each geometric factor $(1 - p^{-s})^{-1} = \sum_{k \geq 0} p^{-ks}$ and multiplying reproduces the sum over positive integers. Since $\zeta(s)$ has a pole at $s = 1$ (the harmonic series diverges), the product must have infinitely many factors — giving an analytic proof of the infinitude of primes. Dirichlet series evaluated on certain "natural points" encode interesting algebraic and arithmetic information.

The Euler product is the first of many important properties of $\zeta(s)$. The next is that $\zeta(s)$ has a meromorphic continuation to the entire plane; we recall the relevant definitions (zero, pole, meromorphic function) in Appendix A.2. The special values of ζ encode deep arithmetic: $\zeta(2n)$ for $n \geq 1$ is a rational multiple of π^{2n}, expressible via Bernoulli numbers. Odd values are much harder — Apéry (1979) proved $\zeta(3)$ is irrational, while the arithmetic nature of $\zeta(5), \zeta(7), \ldots$ remains unknown. The negative-integer values $\zeta(0), \zeta(-1), \ldots$ make sense only *after* analytic continuation: naively $\zeta(-1) = 1 + 2 + 3 + \cdots$, yet the continuation assigns it the value $-\frac{1}{12}$. To study $\zeta(s)$ and its analytic continuation we use the Gamma function

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^{s-1} dt$$

The analytic continuation is most cleanly stated for the *completed zeta function.*

Theorem 2.5.2 Analytic Continuation of the Completed Zeta Function

Define the completed zeta function by

$$\xi(s) = \tfrac{1}{2} s(s-1)\, \Gamma\left(\tfrac{s}{2}\right) \pi^{-s/2} \zeta(s).$$

*Then $\xi(s)$, originally defined for $\Re(s) > 1$, has an analytic continuation to an entire function and satisfies the **functional equation** $\xi(s) = \xi(1 - s)$.*

From now on whenever we say zeta function we often mean the analytic continuation of zeta function by default. Now we will study the zeros of zeta function and why they are important. The Gamma function has poles at all negative integers. From the functional equation of the zeta function it has zeros at all the poles of $\Gamma(s/2)$ i.e. at all negative even integers. These zeros are called *trivial zeros* of ζ. The $\zeta(s)$ has positive value for all $s \in \mathbb{C}$ with $\Re(s) > 1$. Therefore all other zeros lie in the *critical strip* $0 \leq \Re(s) \leq 1$; the functional equation forces them to be symmetric under $s \mapsto 1 - s$, and conjugation $s \mapsto \bar{s}$ forces reflection symmetry across the real axis. Hence

Theorem 2.5.3 Trivial Zeros and Symmetry of Non-trivial Zeros

The only zeros of $\zeta(s)$ where $s \in \mathbb{C}$ for $\Re(s) < 0$ are at $s = -2n$ for $n \in \mathbb{N}$. In the critical strip the zeros lie

symmetrically around the critical line $\Re(s) = 1/2$ i.e. if $a + ib$ is a zero so is $1 - a + ib$.

regarding the zeros of the zeta function. The line $\Re(s) = 1/2$ is the *critical line*, and the most famous open problem in mathematics is:

***Riemann Hypothesis**: Every non-trivial zero of $\zeta(s)$ satisfies $\Re(s) = 1/2$.*

While Hardy in 1914 showed that there are infinitely many zeros on the critical line the problem of every non-trivial zero being on the critical line is a major open problem in mathematics.

Theorem 2.5.4 Hardy, 1914

Infinitely many non-trivial zeros of $\zeta(s)$ lie on the critical line $\Re(s) = 1/2$.

Another interesting property, first noted in Riemann's original paper is the product expansion of zeta function using the non-trivial zeros.

Theorem 2.5.5 Hadamard Product

There exist constants $A, B \in \mathbb{C}$ such that

$$\xi(s) = e^{A+Bs} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho},$$

where the product runs over all non-trivial zeros ρ of ζ.

Note:-

The Hadamard Product gives the "zero-side" factorisation of ξ, to be contrasted with the Euler product as the "prime-side" factorisation. Together they are the source of the famous duality between primes and zeros: equating the two factorisations — typically after taking logarithmic derivatives — converts statements about primes into statements about zeros and vice versa.

Zeros ↔ primes. Taking the logarithmic derivative of the Euler product we obtain,

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s},$$

where $\Lambda(n) = \log p$ if $n = p^k$ and 0 otherwise (the *von Mangoldt function*). Therefore we can see how the non-trivial zeros of $\zeta(s)$ and the primes are very related.

2.5.2 Dirichlet L-function

The function-field world has its own zeta function, obtained by replacing $\mathbb{Z}$ with the polynomial ring $\mathbb{F}_q[X]$ and the primes with monic irreducible polynomials. The notational parallel with the classical case is exact, but the analytic behavior is dramatically simpler: the resulting zeta function is a *rational* function of q^{-s}, with no mystery in its analytic continuation, and (as will be developed in later sections) a fully-proven analog of the Riemann Hypothesis.

Let $\Phi \subseteq \mathbb{F}_q[X]$ denote the set of all monic polynomials, and let $\Phi_d \subseteq \Phi$ be the subset of polynomials of degree exactly d, for each $d \in \mathbb{N}$. A basic count gives $|\Phi_d| = q^d$: a degree-d monic polynomial is determined by its d lower-order coefficients, each of which is a free element of $\mathbb{F}_q$. This simple combinatorial fact is what makes the function-field zeta function so much more tractable than its classical counterpart.

For any polynomial $f \in \mathbb{F}_q[X]$, the quotient ring $\mathbb{F}_q[X] \,/\, \langle f \rangle$ is a finite $\mathbb{F}_q$-algebra of size $q^{\deg(f)}$. We define the *norm* of f to be

$$\mathfrak{n}(f) = \left| \mathbb{F}_q[X] \Big/ \langle f \rangle \right| = q^{\deg(f)},$$

by direct analogy with the norm of an ideal in the ring of integers of a number field. This norm is the direct replacement for the size of $\mathbb{Z}/n\mathbb{Z}$ (namely n) that appears in the classical Riemann zeta function.

To define L-functions in the function field setting we first need to extend multiplicative characters to the polynomial ring. Let $G \subseteq \mathbb{F}_q(X)$ be the group of rational functions $h_1(X)/h_2(X)$ where $h_2 \neq 0$ and $h_1, h_2 \in \Phi$ where $\Phi \subseteq G$ is the set of all monic polynomials over $\mathbb{F}_q$. From now on we use $h(X) \in G$ to denote monic polynomials. Let $\overline{G}$ be the subgroup of G such that $h_1 \cdot h_2 \in \overline{G}$ whenever $h_1, h_2 \in \overline{G}$. For a character Λ on $\overline{G}$ we extend Λ to all of G by setting $\Lambda(r) = 0$ for all $r \notin \overline{G}$, after which Λ remains multiplicative on G. This allows us to sum Λ over all monic polynomials.

Definition 2.5.3: Dirichlet L-function

Let Λ be a character on $\overline{G}$ extended to G as above. For a complex number $s = a + ib \in \mathbb{C}$ with $a > 1$, the Dirichlet L-function *associated to Λ is defined by*

$$L(s, \Lambda) = \sum_{h \in \Phi} \Lambda(h) \cdot \mathfrak{n}(h)^{-s}$$

where the sum runs over all monic polynomials $h \in \Phi \subseteq G$ and $\mathfrak{n}(h) = q^{\deg(h)}$ is the norm of h.

Again this sum is absolutely convergent for $\Re(s) > 1$. Like in the case of Riemann-Zeta functions we can express $L(s, \Lambda)$ in a Euler product form using only irreducible polynomials.

Theorem 2.5.6 Euler Product for $\zeta(s)$

For $s \in \mathbb{C}$ with $\Re(s) > 1$ we have

$$L(s, \Lambda) = \prod_{\substack{h \in \Phi \\ h \text{ irreducible}}} (1 - \Lambda(h) \cdot \mathfrak{n}(h)^{-s})^{-1}$$

Proof: Since $\Re(s) > 1$ we have

$$L(s, \Lambda) = \sum_{h \in \Phi} \Lambda(h) \cdot \mathfrak{n}(h)^{-s} = \prod_{\substack{h \in \Phi \\ h \text{ irreducible}}} \left(\sum_{i=0}^{\infty} \Lambda(h^i) \cdot \mathfrak{n}(h^i)^{-s} \right) = \prod_{\substack{h \in \Phi \\ h \text{ irreducible}}} (1 - \Lambda(h) \cdot \mathfrak{n}(h)^{-s})^{-1}$$

Hence we have the lemma. ∎

So let $\Lambda : G \to \mathbb{C}$ be a multiplicative complex valued function over the G where G is as defined above which also satisfies $|\Lambda(r)| \leq 1$ for all $r \in G$ and $\Lambda(1) = 1$. Then for any $s \in \mathbb{C}$ with $\Re(s) > 1$ we can rewrite

$$L(s, \Lambda) = \sum_{h \in \Phi} \Lambda(h) \cdot \mathfrak{n}(h)^{-s} = \sum_{d=0}^{\infty} \left(\sum_{h \in \Phi_d} \Lambda(h) \right) q^{-d \cdot s}$$

Since for any $h \in \Phi$ we have $\mathfrak{n}(q)^s = (q^s)^{\deg(h)}$ we can consider the corresponding power series

$$\overline{L}(z, \Lambda) = \sum_{h \in \Phi} \Lambda(h) \cdot z^{\deg(h)}$$

hence $\overline{L}(q^{-s}, \Lambda) = L(s, \Lambda)$. Now using Theorem 2.5.6 for $\overline{L}(z, \Lambda)$ we have

$$\overline{L}(z, \Lambda) = \prod_{\substack{h \in \Phi \\ h \text{ irreducible}}} \left(1 - \Lambda(h) \cdot z^{\deg(h)} \right)^{-1}$$

Now we do some operations on the power series $\overline{L}(z,\Lambda)$ to get some expressions which help us later in proving some properties of exponential character sums. So for brevity of notations we use $\mathrm{Irr}(\Phi)$ to denote the set of irreducible polynomials of Φ. We have $\log \overline{L}(z,\Lambda) = -\sum_{h\in \mathrm{Irr}(\Phi)} \log\left(1-\Lambda(h)z^{\deg(h)}\right)$. So by differentiating we get

$$\frac{d}{dz}\log \overline{L}(z,\Lambda) = -\sum_{h\in \mathrm{Irr}(\Phi)} \frac{-\Lambda(h)\cdot \deg(h)\cdot z^{\deg(h)-1}}{1-\Lambda(h)\cdot z^{\deg(h)}} = \sum_{h\in \mathrm{Irr}(\Phi)} \frac{\Lambda(h)\cdot \deg(h)\cdot z^{\deg(h)-1}}{1-\Lambda(h)\cdot z^{\deg(h)}}$$

So

$$\begin{aligned} z\cdot \frac{d}{dz}\log \overline{L}(z,\Lambda) &= \sum_{h\in \mathrm{Irr}(\Phi)} \frac{\Lambda(h)\cdot \deg(h)\cdot z^{\deg(h)-1}}{1-\Lambda(h)\cdot z^{\deg(h)}} \\ &= \sum_{h\in \mathrm{Irr}(\Phi)} \Lambda(h)\cdot \deg(h)\cdot z^{\deg(h)} \sum_{k=0}^{\infty} \Lambda^k(h)\cdot z^{k\cdot \deg(h)} \\ &= \sum_{h\in \mathrm{Irr}(\Phi)} \deg(h) \sum_{k=1}^{\infty} \Lambda^k(h)\cdot z^{k\cdot \deg(h)} \\ &= \sum_{k=1}^{\infty} \left(\sum_{\substack{h\in \mathrm{Irr}(\Phi),\\ \deg(h)\mid k}} \deg(h)\cdot \Lambda(h)^{k/\deg(h)} \right) z^k \end{aligned} \tag{2.8}$$

So define $L_k := \sum_{h\in \mathrm{Irr}(\Phi),\ \deg(h)\mid k} \deg(h)\cdot \Lambda(h)^{k/\deg(h)}$. Hence we have $z\cdot \frac{d}{dz}\log \overline{L}(z,\Lambda) = \sum_{k=1}^{\infty} L_k\cdot z^k$.

Now suppose there exists a positive integer $N\in\mathbb{N}$ such that for all $d>N$ we have $\sum_{h\in\Phi_d}\Lambda(g)=0$. Then we have

$$L(s,\Lambda) = \sum_{d=0}^{N}\left(\sum_{h\in\Phi_d}\Lambda(d)\right)q^{-d\cdot s}, \qquad \overline{L}(z,\Lambda) = \sum_{d=0}^{N}\left(\sum_{h\in\Phi_d}\Lambda(d)\right)z^d$$

Since now $\overline{L}(z,\Lambda)$ is not an infinite power series anymore, instead a degree N polynomial over $\mathbb{C}$ there exists complex numbers $w_1,\cdots,w_N\in\mathbb{C}$ such that $\overline{L}(z,\Lambda) = (1-w_1z)(1-w_2z)\cdots(1-w_Nz)$. Now applying $z\cdot\frac{d}{dz}\log\overline{L}(z,\Lambda)$ on this expression we get

$$\begin{aligned} z\cdot\frac{d}{dz}\log\overline{L}(z,\Lambda) &= -\sum_{i=1}^{N}\frac{w_iz}{1-w_iz} \\ &= -\sum_{i=1}^{N} w_iz\left(\sum_{j=0}^{\infty} w_i^jz^j\right) \\ &= -\sum_{i=1}^{N}\left(\sum_{j=1}^{\infty} w_i^j\right)z^i \end{aligned} \tag{2.9}$$

Therefore by comparing coefficients of (2.9) with (2.8) we get $L_k = -\sum_{i=1}^{N} w_i^k$. for all $k\in\mathbb{N}$. So we have the following theorem

Theorem 2.5.7

Let $\Lambda : G\to\mathbb{C}$ be a multiplicative function with $|\Lambda(r)|\le 1$ for all $r\in G$ and $\Lambda(1)=1$, and suppose there exists a positive integer $N\in\mathbb{N}$ such that $\sum_{h\in\Phi_d}\Lambda(h)=0$ for every $d>N$. Then $\overline{L}(z,\Lambda)$ is a polynomial of degree at most N over $\mathbb{C}$, and there exist complex numbers $w_1,\ldots,w_N\in\mathbb{C}$ such that

$$\overline{L}(z,\Lambda) = \prod_{i=1}^{N}(1-w_iz), \qquad \text{and} \qquad L(s,\Lambda) = \prod_{i=1}^{N}\left(1-w_iq^{-s}\right).$$

Moreover, for every $k \in \mathbb{N}$,

$$L_k = \sum_{\substack{h \in \mathrm{Irr}(\Phi) \\ \deg(h) \mid k}} \deg(h) \cdot \Lambda(h)^{k/\deg(h)} = -\sum_{i=1}^{N} w_i^k.$$

§ 2.6 Lifting of Gaussian and Jacobi Sums

So using the discussion on *Dirichlet L-functions* we can relate Gaussian sums in $\mathbb{F}_q$ with Gaussian sum of lifted characters to any finite extension of $\mathbb{F}_q$.

Theorem 2.6.1 Davenport-Hasse Theorem

Let $\chi \in \boldsymbol{\mathcal{X}}_q$ be an additive character and $\psi \in \boldsymbol{\mathcal{M}}_q$ be a multiplicative character of $\mathbb{F}_q$, not both of them trivial. Suppose χ and ψ are lifted to characters $\chi^{(r)}$ and $\psi^{(r)}$ respectively of the finite extension field E of $\mathbb{F}_q$ with $[E\colon \mathbb{F}_q] = r$. Then

$$G(\psi^{(r)}, \chi^{(r)}) = (-1)^{r-1} G(\psi, \chi)^r$$

Proof: Like previous section G is the set of all rational functions $r(X) = {}^{h_1}/_{h_2}$ where $h_1, h_2 \in \Phi$, h_2 non-zero. Let $\overline{G}$ be the subgroup of G consisting of rational functions $r(X) = \frac{h_1(X)}{h_2(X)}$ having $h_1(0), h_2(0) \neq 0$. Then define the multiplicative function $\Lambda : G \to \mathbb{C}$ such that for all $r \in G$, $|\Lambda(r)| \leq 1$ and $\Lambda(1) = 1$ in the following way:

If $r(X) \in G$ where $r(X) = \frac{h_1(X)}{h_2(X)}$ and in the common splitting field E of h_1, h_2, they factor as:

$$h_1(X) = \prod_{i=1}^{d_1} (X - \alpha_i) \qquad \text{and} \qquad h_1(X) = \prod_{j=1}^{d_2} (X - \beta_j)$$

where $\alpha_i, \beta_j \in E$ for all $i \in [d_1]$, $j \in [d_2]$ then

$$\Lambda(r) = \psi\left(\prod_{i=1}^{d_1} \alpha_i \cdot \prod_{j=1}^{d_2} \beta_j^{-1}\right) \cdot \chi\left(\sum_{i=1}^{d_1} \alpha_i - \sum_{j=1}^{d_2} \beta_j\right)$$

Now for any monic polynomial $h \in \Phi$ let $h(X) = X^n - c_1 \cdot X^{n-1} + \cdots + (-1)^n c_n$ where $c_i \in \mathbb{F}_q$ for all $i \in [n]$. Then we have $\Lambda(h) = \psi(c_n) \cdot \chi(c_1)$. Clearly Λ is a multiplicative map on G. Now for a given (c_1, c_k) pair there are q^{n-2} polynomials where same (c_1, c_n) occurs. Hence

$$\sum_{h \in \Phi_k} \Lambda(h) = q^{n-2} \sum_{c_1 \in \mathbb{F}_q} \sum_{c_n \in \mathbb{F}_q^*} \psi(c_n) \cdot \chi(c_1) = q^{n-2} \left(\sum_{c_n \in \mathbb{F}_q}\right)\left(\sum_{c_1 \in \mathbb{F}_q} \chi(c_1)\right)$$

Since ψ and χ are non-trivial by Theorem 2.2.2 and Theorem 2.2.5 we have $\sum_{h \in \Phi_k} \Lambda(h) = 0$ for $n > 1$.

Now $\Phi(0) = \{1\}$ and $\Phi(1) = \{x - \alpha\colon \alpha \in \mathbb{F}_q^*\}$. Hence

$$\sum_{h \in \Phi_1} \Lambda(h) = \sum_{\alpha \in \mathbb{F}_q^*} \psi(\alpha) \cdot \chi(\alpha) = G(\psi, \chi)$$

Therefore

$$\overline{L}(z, \Lambda) = 1 + G(\psi, \chi) \cdot z \qquad \text{and} \qquad L(s, \chi) = 1 + q^{-s} \cdot G(\psi, \chi)$$

Therefore root of $\overline{L}(z, \Lambda)$ is $-G(\psi, \chi)$.

Now we will calculate L_r where $r = [E\colon \mathbb{F}_q]$. From the definition of L_r in previous section we have

$$L_r = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) \mid r}} \deg(h) \cdot \chi(h)^{r/\deg(h)} = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) \mid r}} \deg(h) \cdot \Lambda\left(h^{r/\deg(h)}\right)$$

Now for any irreducible h let γ be a root of h in the splitting field of f. Let

$$h^{r/\deg(f)} = x^r - c_1 \cdot x^{r-1} + \cdots + (-1)^k c_r = \prod_{i=o}^{r-1} \left(X - \gamma^{q^i}\right)$$

Hence $c_1 = \mathrm{Tr}_{E/\mathbb{F}_q}(\gamma)$ and $c_r = \mathrm{N}_{E/\mathbb{F}_q}(\gamma)$. Hence we have

$$\Lambda\left(h^{r/\deg(h)}\right) = \psi(c_r) \cdot \chi(c_1) = \psi\left(\mathrm{N}_{E/\mathbb{F}_q}(\gamma)\right) \cdot \chi\left(\mathrm{Tr}_{E/\mathbb{F}_q}(\gamma)\right) = \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma)$$

Now we can express $\deg(h) \cdot \Lambda\left(h^{r/\deg(h)}\right) = \sum_{\gamma \in E, h(\gamma)=0} \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma)$. Therefore

$$L_r = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) | k}} \deg(h) \cdot \Lambda\left(h^{k/\deg(h)}\right) = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) | k}} \sum_{\gamma \in E, h(\gamma)=0} \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma)$$

Now all irreducible polynomials of $\mathbb{F}_q$ whose degree divides r with non-zero constant their roots in E are disjoint sets and partition E^*. Therefore we have

$$L_r = \sum_{\gamma \in E^*} \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma) = G(\psi^{(r)}, \chi^{(r)})$$

Since by Theorem 2.5.7 we have $L_r = -(-G(\psi, \chi))^r$. So we have the theorem. ∎

An analogous lifting phenomenon holds for Jacobi sums. Using the relation between Jacobi and Gaussian sums established in Theorem 2.4.3 together with the Davenport-Hasse Theorem, one can compare the Jacobi sum of a tuple of characters over $\mathbb{F}_q$ with the Jacobi sum of their lifts to a finite extension. The following theorem makes this precise.

Theorem 2.6.2 Lifting of Jacobi Sums

Let $\lambda_1, \ldots, \lambda_k \in \boldsymbol{\mathcal{M}}_q$ be multiplicative characters of $\mathbb{F}_q$, not all of which are trivial. Suppose $\lambda_1, \ldots, \lambda_k$ are lifted to characters $\lambda_1(r), \ldots, \lambda_k^{(r)}$ respectively, of the finite extension field E of $\mathbb{F}_q$ with $[E : \mathbb{F}_q] = r$. Then

$$J\left(\lambda_1^{(r)}, \ldots, \lambda_k^{(r)}\right) = (-1)^{(r-1)(k-1)} J(\lambda_1, \ldots, \lambda_k)^r$$

Proof: Now if any λ_i is trivial then the lifted character $\lambda^{(r)}$ is also trivial over E. If Then by Theorem 2.4.6(ii) we get both sides to be 0. So suppose $\lambda_1, \ldots, \lambda_k$ are non-trivial. Now we will use Theorem 2.4.3. Let $\chi \in \boldsymbol{\mathcal{X}}_q$ be any non-trivial additive character of $\mathbb{F}_q$. So let $\chi^{(r)}$ be the lifted character of χ to E. If $\prod_{i=1}^k \lambda_i$ is non-trivial then we have

$$\begin{aligned}
J\left(\lambda_1^{(r)}, \ldots, \lambda_k^{(r)}\right) &= \frac{\prod_{i=1}^k G\left(\lambda_i^{(r)}, \chi^{(r)}\right)}{G\left(\prod_{i=1}^k \lambda_i^{(r)}, \chi^{(r)}\right)} \\
&= \frac{\prod_{i=1}^k (-1)^{r-1} G(\lambda_i, \chi)^r}{(-1)^{r-1} G\left(\prod_{i=1}^k \lambda_i, \chi\right)^r} && \text{[By Davenport-Hasse Theorem]} \\
&= (-1)^{(s-1)(r-1)} \left[\frac{\prod_{i=1}^k G(\lambda_i, \chi)}{G\left(\prod_{i=1}^k \lambda_i, \chi\right)}\right]^r \\
&= (-1)^{(s-1)(r-1)} J(\lambda_1, \ldots, \lambda_k)^r
\end{aligned}$$

Now the only case remaining is when $\prod_{i=1}^{k} \lambda_i$ is trivial. In that case we have

$$\begin{aligned} J\left(\lambda_1^{(r)}, \ldots, \lambda_k^{(r)}\right) &= -\frac{1}{q^r} \prod_{i=1}^{k} G\left(\lambda_i^{(r)}, \chi^{(r)}\right) \\ &= -\frac{1}{q^s} \cdot (-1)^{(r-1)\cdot k} \prod_{i=1}^{k} G(\lambda_i, \chi)^r \qquad \text{[By Davenport-Hasse Theorem]} \\ &= (-1)^{(r-1)(k-1)} \cdot J(\lambda_1, \ldots, \lambda_k) \end{aligned}$$

So we have the theorem. ∎

§ 2.7 Kloosterman Sums

We now turn to a genuinely different family of exponential sums, introduced by Hendrik Kloosterman in 1926 in his refinement of the Hardy–Littlewood circle method for the representation of integers by positive definite quaternary quadratic forms. Instead of twisting two characters of different types, a Kloosterman sum twists a single additive character χ with the *inversion* map $\gamma \mapsto \gamma^{-1}$ on $\mathbb{F}_q^*$, producing a sum of the form $\sum_{\gamma} \chi(a\gamma + b\gamma^{-1})$.

Definition 2.7.1: Kloosterman Sums

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and let $a, b \in \mathbb{F}_q$. Then the Kloosterman Sum *associated to χ, a, b is defined as*

$$K(\chi; a, b) = \sum_{\gamma \in \mathbb{F}_q^*} \chi(a \cdot \gamma + b \cdot \gamma^{-1})$$

When one or both of the parameters a, b vanish, the Kloosterman sum collapses to an additive character sum on $\mathbb{F}_q^*$, which we have already computed in Theorem 2.2.2. If both $a = b = 0$ then for all $\gamma \in \mathbb{F}_q^*$, $\chi(a \cdot \gamma + b \cdot \gamma^{-1}) = \chi(0) = 1$. Hence $K(\chi; 0, 0) = q - 1$.

If exactly one of a, b is zero then we have either

$$K(\chi; a, b) = \sum_{\gamma \in \mathbb{F}_q^*} \chi(a \cdot \gamma) = -1 \text{ or } K(\chi; a, b) = \sum_{\gamma \in \mathbb{F}_q^*} \chi(b \cdot \gamma^{-1}) = -1$$

So in both cases $K(\chi; a, b) = -1$. So we have the following observations:

Observation 2.12. *Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and $a, b \in \mathbb{F}_q$ such that $ab = 0$. Then*

(i) $K(\chi; 0, 0) = q - 1$ if both $a = b = 0$;

(ii) if exactly one of a, b is 0 then $K(\chi; a, b) = -1$;

Now another property to notice is Kloosterman sum is always real since

$$\overline{K(\chi; a, b)} = \sum_{\gamma \in \mathbb{F}_q^*} \overline{\chi}(a \cdot \gamma + b \cdot \gamma^{-1}) = \sum_{\gamma \in \mathbb{F}_q^*} \chi(-(a \cdot \gamma + b \cdot \gamma^{-1})) = \sum_{\gamma \in \mathbb{F}_q^*} \chi(a \cdot (-\gamma) + b \cdot (-\gamma)^{-1}) = K(\chi; a, b)$$

Therefore we have another observation:

Observation 2.13. *For all additive characters $\chi \in \boldsymbol{\chi}_q$ and $a, b \in \mathbb{F}_p$, the Kloosterman Sum is real i.e. $\overline{K(\chi; a, b)} = K(\overline{\chi}; a, b) = K(\chi; -a, -b) = K(\chi; a, b)$.*

Now we can express Kloosterman sums as product of two additive characters like Jacobi sums is for product of multiplicative characters. For any $\chi \in \boldsymbol{\chi}_q$ let $\chi = \chi_c$ for some $c \in \mathbb{F}_q$ as defined in Theorem 2.2.1. Then

$$K\left(\chi_c; a, b\right) = \sum_{\gamma \in \mathbb{F}_q^*} \chi_c\left(a \cdot \gamma + b \cdot \gamma^{-1}\right) = \sum_{\gamma \in \mathbb{F}_q^*} \chi_{a \cdot c}(\gamma) \cdot \chi_{-b \cdot c}(\gamma) =: K\left(\chi_{ac}, \overline{\chi}_{bc}\right)$$

Observation 2.14. *The Kloosterman sum can equivalently be viewed as two-additive-character sum: for any $a, b, c \in \mathbb{F}_q$ and $\chi_a, \chi_b, \chi_c \in \boldsymbol{\chi}_q$ (in the notation of Theorem 2.2.1),*

$$K\left(\chi_a, \chi_b\right) \coloneqq K\left(\chi_1; a, -b\right) \qquad \text{and} \qquad K\left(\chi_c; a, -b\right) = K\left(\chi_{ac}, \chi_{bc}\right)$$

When the ground field has odd characteristic there is a useful alternative expression for $K(\chi; a, b)$ as a sum of a single additive character against the quadratic character η evaluated on a quadratic polynomial. This representation is often the starting point for asymptotic bounds on Kloosterman sums in arithmetic contexts.

Lemma 2.7.1

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ with q odd, let η be the quadratic character of $\mathbb{F}_q$ (extended by $\eta(0) = 0$), and let $a, b \in \mathbb{F}_q$ be not both zero. Then

$$K(\chi; a, b) = \sum_{d \in \mathbb{F}_q} \chi(d) \cdot \eta(d^2 - 4ab)$$

Proof: If one of a, b is zero then by Observation 2.12(ii) we have $K(\chi; a, b) = -1$. And then at the right hand side we have the sum over $\eta(\alpha^2 - 4ab) = \eta(\alpha^2) = 1$ and therefore the sum is of $\chi(\alpha)$ over all non-zero alpha which by Theorem 2.2.2(i) gives $q - 1 = -1$. Therefore in this case the lemma is satisfied. So now assume $a \cdot b \neq 0$.

Let for any $\gamma \in \mathbb{F}_q^*$, denote $d = a \cdot \gamma + b \cdot \gamma^{-1}$. Now for any $d \in \mathbb{F}_q$ let $Z(d)$ is the number of $\gamma \in \mathbb{F}_q^*$ such that $d = a \cdot \gamma + b \cdot \gamma^{-1}$. Now for any $d \in \mathbb{F}_q$,

$$a \cdot \gamma + b \cdot \gamma^{-1} = d \iff a \cdot \gamma^2 - d \cdot \gamma + b = 0$$

Therefore if $P_d(X) = a \cdot X^2 - d\dot{X} + b \in \mathbb{F}_q[X]$ be a polynomial then $Z(d) =$ number of zeros of $P_d(X)$. Hence $Z(d) \leq 2$. Let Δ_d is the discriminant of P_d i.e. $\Delta_d = d^2 - 4ab$. Hence If $\Delta_d = 0$ then $Z(d) = 1$. If a non-zero square root of Δ_d exists in $\mathbb{F}_q$ then Δ_d is a quadratic residue i.e. $\eta(\Delta_d) = 1$ then $Z(d) = 2$ and if it doesn't exists then Δ_d is a quadratic non-residue i.e. $\eta(\Delta_d) = -1$ then $Z(d) = 0$. Therefore together we can write $Z(d) = \eta(\Delta_d) + 1$ with the extended definition of η by setting $\eta(0) = 0$. Therefore we have

$$K(\chi; a, b) = \sum_{d \in \mathbb{F}_q} \chi(d) \cdot Z(d) = \sum_{d \in \mathbb{F}_q} \chi(d) \cdot \cdot \eta(\Delta_d) = \sum_{d \in \mathbb{F}_q} \chi(d) \cdot \cdot \eta(d^2 - 4ab)$$

Hence we have the lemma. ∎

The following theorem is the Kloosterman analogue of the Davenport-Hasse Theorem: the entire sequence of lifted sums $K(\chi^{(s)}; a, b)$ is controlled by just two algebraic numbers w_1, w_2 — the reciprocal roots of the associated L-function. The proof uses the Dirichlet L-function theory from section 2.5.

Theorem 2.7.2 Lifting of Kloosterman Sum

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and let $a, b \in \mathbb{F}_q$ with $ab \neq 0$. Suppose χ is lifted to character $\chi^{(r)}$ of the finite extension field E of $\mathbb{F}_q$ with $[E\colon \mathbb{F}_q] = r$. Then there exist complex numbers w_1, w_2 (depending only

on χ, a, b), which are either complex conjugates of each other or both real, such that for every positive integer r

$$K(\chi^{(r)}; a, b) = \sum_{\gamma \in \mathbb{F}_{q^r}^*} \chi^{(r)}(a\gamma + b\gamma^{-1}) = -w_1^r - w_2^r,$$

Proof: Let G be the set of all rational functions $r(X) \in \mathbb{F}_q(X)$ such that for any $r \in G$, $r(X) = {}^{h_1(X)}/_{h_2(X)}$ where $h_1, h_2 \in \mathbb{F}_q[x]$ and both h_1, h_2 are monic. Let $\overline{G}$ the subgroup of G such that for any $r \in \overline{G}$, $r(X) = {}^{h_1(X)}/_{h_2(X)}$ where h_1, h_2 have non-zero constant term. We will define a multiplicative function $\Lambda : G \to \mathbb{C}$ such that $|\Lambda(r)| \leq 1$ for all $r \in G$ and $\Lambda(1) = 1$. Let $r \in \overline{G}$ where $r(X) = \frac{h_1(X)}{h_2(X)}$ and in common splitting field E of h_1, h_2 they factor as

$$h_1(X) = \prod_{i=1}^{d_1}(X - \alpha_i) \qquad \text{and} \qquad h_2(X) = \prod_{j=1}^{d_2}(X - \beta_j)$$

Then we will define Λ using χ in the following way:

$$\Lambda(r) = \chi\left[a\left(\sum_{i=1}^{d_1}\alpha_i - \sum_{j=1}^{d_2}\beta_j\right) + b\left(\sum_{i=1}^{d_1}\frac{1}{\alpha_i} - \sum_{j=1}^{d_2}\frac{1}{\beta_j}\right)\right] \tag{2.10}$$

Then if $h \in \Phi_n$ such that $h(X) = X^n - c_1 \cdot X^{n-1} + \cdots + (-1)^n \cdot c_n$ then $\Lambda(h) = \chi\left(a \cdot c_1 + b \cdot c_{k-1} \cdot c_k^{-1}\right)$. Now it is enough to show that the corresponding coefficients are additive if two polynomials are multiplied since χ is an additive character. Let $h_1 \in \Phi_{n_1}$ and $g_2 \in \Phi_{n_2}$ and

$$g_1(X) = X^{n_1} + \sum_{i=1}^{n_1}(-1)^i \cdot c_{1,i} \cdot X^{n_1 - i}, \quad g_2(X) = X^{n_2} + \sum_{j=2}^{n_2}(-2)^j \cdot c_{2,j} \cdot X^{n_2 - j} \text{ and } g_1 \cdot g_2 = X^{n_1+n_2} + \sum_{t=1}^{n_1+n_2}(-1)^t \cdot c_j \cdot X^{n_1+n_2-t}$$

Then by comparing coeffieicnets we have

$$c_1 = c_{1,1} + c_{2,1}, \quad c_{n_1+n_2} = c_{1,n_1} \cdot c_{2,n_2}, \quad \text{and} \quad c_{n_1+n_2-1} = c_{1,n_1} \cdot c_{2,n_2-1} + c_{1,n_1-1} \cdot c_{2,n_2}$$

Hence Λ is indeed a multiplicative function. Now we extend Λ to G by setting $\Lambda(r) = 0$ for all $h \notin \overline{G}$.

Now for any $k \in \mathbb{Z}_0$,if (c_1, c_{k-1}, c_k) is fixed where $c_k \neq 0$, there are q^{k-3} possible choices for $c_2, \ldots, c_{k-2}$ such that all polynomials with coefficients c_i's give same value for $\chi(a \cdot c_1 + b \cdot c_{k-1} \cdot c_k^{-1}$. So we will show that for $k \geq 3$, $\sum_{h \in \Phi_k} \Lambda(h) = 0$. So

$$\begin{aligned}
\sum_{g \in \Phi_k} \Lambda(g) &= \sum_{c_1, \ldots, c_{k-1}} \sum_{c_k \in \mathbb{F}_q^*} \chi(a \cdot c_1 + b \cdot c_{k-1} \cdot c_k^{-1}) \\
&= q^{k-3} \sum_{c_1, c_{k-1} \in \mathbb{F}_q} \sum_{c_k \in \mathbb{F}_q^*} \chi(a \cdot c_1 + b \cdot c_{k-1} \cdot c_k^{-1}) \\
&= q^{k-3}\left(\sum_{c_1 \in \mathbb{F}_q} \chi(a \cdot c_1)\right)\left(\sum_{c_{k-1} \in \mathbb{F}_q} \sum_{c_k \in \mathbb{F}_q^*} \chi(b \cdot c_{k-1} \cdot c_k^{-1})\right) = 0 \qquad \text{[By Theorem 2.2.2(i)]}
\end{aligned}$$

Hence we have $\overline{L}(z, \Lambda) = 1 + z\sum_{h \in \Phi_1}\Lambda(h) + z^2\sum_{h \in \Phi_2}\Lambda(h)$. Now $\Phi_1 = \{X - \gamma \mid \gamma \in \mathbb{F}_q^*\}$. Therefore we have

$$\sum_{h \in \Phi_1} \Lambda(h) = \sum_{\gamma \in \mathbb{F}_q^*} \chi(a \cdot \gamma + b \cdot \gamma^{-1}) = K(\chi; a, b)$$

And furthermore for Φ_2 we get

$$\begin{aligned}
\sum_{h \in \Phi_2} \Lambda(h) &= \sum_{c_1 \in \mathbb{F}_q} \sum_{c_2 \in \mathbb{F}_q^*} \Lambda(X^2 - c_1 \cdot X + c_2) \\
&= \sum_{c_1 \in \mathbb{F}_q} \sum_{c_2 \in \mathbb{F}_q^*} \chi(a \cdot c_1 + b \cdot c_1 \cdot c_2^{-1}) \\
&= \sum_{c_1 \in \mathbb{F}_q} \sum_{c_2 \in \mathbb{F}_q^*} \chi(c_1(a + b \cdot c_2^{-1}))
\end{aligned}$$

So if $c_2 = -a^{-1}b$ then $a + b \cdot c_2^{-1} = 0$ and hence $\sum_{c_1 \in \mathbb{F}_q} \chi(0) = q$ and if $c_2 \neq -a^{-1}b$ then $\sum_{c_1 \in \mathbb{F}_q} \chi(c_1(a + b \cdot c_2^{-1})) = 0$ by Theorem 2.2.2(i). So therefore $\sum_{h \in \Phi_2} \Lambda(h) = q$ and thus we get $\overline{L}(z, \Lambda) = 1 + k(\chi; a, b) \cdot z + q \cdot z^2$. Let $w_1, w_2 \in \mathbb{C}$ which are either complex conjugates or both real such that

$$\overline{L}(z, \Lambda) = (1 - w_1 \cdot z)(1 - w_2 \cdot z)$$

Now we will calculate L_r where $r = [E : \mathbb{F}_q]$. From the definition of L_r in previous section we have

$$L_r = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) | r}} \deg(h) \cdot \chi(h)^{r/\deg(h)} = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) | r}} \deg(h) \cdot \Lambda\left(h^{r/\deg(h)}\right)$$

Now for any irreducible h let γ be a root of h in the splitting field of f. Let

$$h^{r/\deg(f)} = x^r - c_1 \cdot x^{r-1} + \cdots + (-1)^k c_r = \prod_{i=o}^{r-1} \left(X - \gamma^{q^i}\right)$$

Hence $c_1 = \mathrm{Tr}_{E/\mathbb{F}_q}(\gamma)$ and $c_r = \mathrm{N}_{E/\mathbb{F}_q}(\gamma)$. Hence we have

$$\Lambda\left(h^{r/\deg(h)}\right) = \psi(c_r) \cdot \chi(c_1) = \psi\left(\mathrm{N}_{E/\mathbb{F}_q}(\gamma)\right) \cdot \chi\left(\mathrm{Tr}_{E/\mathbb{F}_q}(\gamma)\right) = \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma)$$

Now we can express $\deg(h) \cdot \Lambda\left(h^{r/\deg(h)}\right) = \sum_{\gamma \in E, h(\gamma) = 0} \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma)$. Therefore

$$L_r = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) | k}} \deg(h) \cdot \Lambda\left(h^{k/\deg(h)}\right) = \sum_{\substack{h \in \mathrm{Irr}(\Phi), \\ \deg(h) | k}} \sum_{\gamma \in E, h(\gamma) = 0} \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma)$$

Now all irreducible polynomials of $\mathbb{F}_q$ whose degree divides r with non-zero constant their roots in E are disjoint sets and partition E^*. Therefore we have

$$L_r = \sum_{\gamma \in E^*} \psi^{(r)}(\gamma) \cdot \chi^{(r)}(\gamma) = G(\psi^{(r)}, \chi^{(r)})$$

Since by Theorem 2.5.7 we have $L_r = -w_1^r - w_2^r$. So we have the theorem. ∎

Assuming w_1, w_2 as above, every lifted Kloosterman sum is a symmetric function of w_1, w_2, which by Waring's identities can be rewritten purely in terms of $K = K(\chi; a, b)$ and q. This gives the following closed-form reduction of $K(\chi^{(s)}; a, b)$ to the ground-field sum.

Lemma 2.7.3 Reduction of Lifted Kloosterman Sums

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and let $a, b \in \mathbb{F}_q$ with $ab \neq 0$. Then for every positive integer s,

$$K\left(\chi^{(s)}; a, b\right) = \sum_{j=0}^{\lfloor s/2 \rfloor} (-1)^{s-j-1} \cdot \frac{s}{s-j} \cdot \binom{s-j}{j} \cdot q^j \cdot K(\chi; a, b)^{s-2j}.$$

Proof: Consider the polynomial $X_1^r + X_2^r$. Now by Waring's Formula we can express this polynomial in terms of elementary symmetric polynomials $\mathrm{ESym}_1, \ldots, \mathrm{ESym}_r$ which gives

$$\begin{aligned} X_1^r + X_2^r &= \sum_{\substack{d_1, d_2 \in \mathbb{Z}_0 \\ d_1 + 2d_2 = r}} (-1)^{d_2} \frac{r(d_1 + d_2 - 1)!}{d_1! d_2!} \mathrm{ESym}_1(X_1, X_2)^{d_1} \cdot \mathrm{ESym}_2(X_1, X_2)^{d_2} \\ &= \sum_{j=0}^{\lfloor r/2 \rfloor} (-1)^j \frac{(r-j-1)! r}{j!(s-2j)!} \mathrm{ESym}_1(X_1, X_2)^{s-2j} \cdot \mathrm{ESym}_2(X_1, X_2)^j \\ &= \sum_{j=0}^{\lfloor r/2 \rfloor} (-1)^j \binom{r-j}{j} \frac{r}{s-j} \mathrm{ESym}_1(X_1, X_2)^{s-2j} \cdot \mathrm{ESym}_2(X_1, X_2)^j \end{aligned}$$

Now we take $X_1 = w_1$ and $X_2 = w_2$ obtained from Theorem 2.7.2 we get $w_1^k + w_2^k = -K(\chi^{(k)}; a, b)$ for any $k \in \mathbb{N}$ and $w_1 \cdot w_2 = q$. So by this replacing in the equation we get the result. ∎

This lemma gives a relation between powers of Kloosterman sum and lifted Kloosterman sums. We can obtain a recursive relation between different lifts of Kloosterman sums very easily using this identiy

$$w_1^r + w_2^r = (w_1^{r-1} + w_2^{r-1})(w_1 + w_2) - (w_1^{r-2} + w_2^{r-2})w_1w_2$$

which gives us

$$K(\chi^{(r)}; a, b) = -K(\chi^{(r-1)}; a, b) \cdot K(\chi; a, b) - q \cdot K(\chi^{(r-2)}; a, b)$$

and so we have the corollary.

Corollary 2.7.4 Kloosterman Sum Recurrence Relation

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and let $a, b \in \mathbb{F}_q$ with $ab \neq 0$. Then for every positive integer s, the lifted Kloosterman sums satisfy the recurrence

$$K\left(\chi^{(s)}; a, b\right) = -K\left(\chi^{(s-1)}; a, b\right) \cdot K(\chi; a, b) - K\left(\chi^{(s-2)}; a, b\right) \cdot q \qquad \text{for all } s \geq 2,$$

with the convention $K(\chi^{(0)}; a, b) := -2$ (and noting $K\left(\chi^{(1)}; a, b\right) = K(\chi; a, b)$).

§ 2.8 Character Sums via the Quadratic Character

For certain special characters, the associated Gaussian sums can be evaluated explicitly. A very special multiplicative character to study is the quadratic character, η. We get a very celebrated formula for the Gaussian sum of η and the canonical additive character.

Theorem 2.8.1 Gaussian Sum of the Quadratic Character

Let $\mathbb{F}_q$ be a finite field with $q = p^r$ where p is an odd prime and $r \in \mathbb{N}$. Let η be the quadratic character of $\mathbb{F}_q$ and let χ_1 be the canonical additive character of $\mathbb{F}_q$. Then

$$G(\eta, \chi_1) = \begin{cases} (-1)^{r-1} \cdot q^{1/2} & \text{if } p \equiv 1 \bmod 4 \\ (-1)^{r-1} \cdot i^r \cdot q^{1/2} & \text{if } p \equiv 3 \bmod 4 \end{cases}$$

In other words,

$$G(\eta, \chi_1) = (-1)^{r-1} \cdot i^{\frac{(p-1)^2}{4} r} \cdot q^{1/2}$$

Proof: Since η always takes values in ± 1 we have $\eta = \overline{\eta}$. Hence by using Lemma 2.3.2(iv) we have $G(\eta, \chi_1)^2 = \eta(-1)q$. Now if $q \equiv 1 \bmod 4$ then $\eta(-1) = 1$ and if $q \equiv 3 \bmod 4$ then $\eta(-1) = -1$. Hence

$$G(\eta, \chi_1) = \begin{cases} \pm\, q^{1/2} & \text{if } q \equiv 1 \bmod 4 \\ \pm\, i \cdot q^{1/2} & \text{if } q \equiv 3 \bmod 4 \end{cases} \tag{2.11}$$

Since $i^{(p-1)^2/4} = \begin{cases} 1 & \text{if } q \equiv 1 \bmod 4 \\ 0 & \text{if } q \equiv 3 \bmod 4 \end{cases}$

Observation 2.15. *For any finite field $\mathbb{F}_q$, $G(\eta, \chi_1) = \pm\, i^{(q-1)^2/4} \cdot q^{1/2}$*

It remains to resolve the $\pm$ sign. We first prove this for $r = 1$ then extend to $r > 1$.

Step 1: Setting up the linear operator T. Let W be the vector space of all complex valued functions on $\mathbb{F}_p^*$, a $(p-1)$-dimensional vector space over $\mathbb{C}$. Let $\mathcal{B} \coloneqq \{f_1, \ldots, f_{p-1}\}$ where $f_j(k) = 1$ if $j = k$ and otherwise 0 be the standard basis of W. By the orthogonality of multiplicative characters in Theorem 2.2.5 the set $\mathcal{B}_{\mathcal{M}_q} \coloneqq \{\psi_0, \ldots, \psi_{p-2}\}$ as defined in Lemma 2.2.3 also forms a basis of W.

Let $\zeta_p = e^{2\pi i/p}$ be the p^{th} root of unity. Hence if $\chi_k \in \boldsymbol{\mathcal{X}}_p$ is as defined in Theorem 2.2.1, then $\chi_k(j) = \zeta_p^{jk}$ for all $j \in \mathbb{F}_p$. Define the linear operator $T : W \to W$ by

$$\forall\, h \in W,\ \forall\, k \in [p-1] \qquad (Th)(k) = \sum_{j=1}^{p-1} \zeta_p^{jk} h(j). \tag{2.12}$$

For each basis element ψ_i,

$$(T\psi_i)(k) = \sum_{j=1}^{p-1} \chi_k(j) \cdot \psi_i(j) = G(\psi_i, \chi_k) = \overline{\psi_i}(k) \cdot G(\psi_i, \chi_1) \quad \forall\, k \in \mathbb{F}_p^*,$$

where the last equality uses Lemma 2.3.2(i). Hence $T\psi_i = G(\psi_i, \chi_1) \cdot \overline{\psi_i}$ for every multiplicative character of $\mathbb{F}_p$.

Step 2: Computing $\det(T)$ via the character basis. Since $\overline{\psi} = \psi$ only for ψ_0 and $\eta = \psi_{p-1/2}$, the matrix of T in $\mathcal{B}_{\mathcal{M}_q}$ has two diagonal entries $G(\psi_0, \chi_1) = -1$ and $G(\eta, \chi_1)$; all other characters pair up as conjugates $(\psi, \overline{\psi})$, each contributing the block $\begin{pmatrix} 0 & G(\overline{\psi}, \chi_1) \\ G(\psi, \chi_1) & 0 \end{pmatrix}$ with determinant contribution $-\psi(-1)p$ by Lemma 2.3.2(iv). Using $\psi_j(-1) = (-1)^j$ (since $\psi(-1) = -1$ for the principal character),

$$\begin{aligned} \det(T) &= -G(\eta, \chi_1) \cdot (-p)^{(p-3)/2} \prod_{j=1}^{\frac{p-3}{2}} (-1)^j \\ &= (-1)^{(p-1)/2} \cdot G(\eta, \chi_1) \cdot p^{(p-3)/2} \cdot (-1)^{(p-1)(p-3)/8} \\ &= \pm(-1)^{(p-1)/2} \cdot i^{(p-1)(p-2)/2} \cdot p^{(p-2)/2}, \end{aligned}$$

where in the last step we substitute Observation 2.15 and simplify the exponent. Thus:

$$\det(T) = \pm(-1)^{(p-1)/2} \cdot i^{(p-1)(p-2)/2} \cdot p^{(p-2)/2}. \tag{2.13}$$

Step 3: Computing $\det(T)$ via the standard basis. From (2.12), the (j,k) entry of T in $\mathcal{B}$ is ζ_p^{jk}. Factoring and reducing modulo p (since $p \mid 1 + 2 + \cdots + (p-1)$), the determinant equals that of the Vandermonde matrix $V(\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1})$:

$$\det(T) = \prod_{1 \le j < k \le p-1} \left(\zeta_p^k - \zeta_p^j\right).$$

Setting $\zeta_{2p} = e^{\pi i/p}$ so that $\zeta_{2p}^2 = \zeta_p$ and using $\zeta_{2p}^t - \zeta_{2p}^{-t} = 2i \sin \frac{t\pi}{p}$, we expand the Vandermonde product:

$$\begin{aligned} \det(T) &= \prod_{1 \le j < k \le p-1} \zeta_{2p}^{j+k} \left(\zeta_{2p}^{k-j} - \zeta_{2p}^{-(k-j)}\right) \\ &= \zeta_{2p}^{\sum_{j<k}(j+k)} \cdot i^{(p-1)(p-2)/2} \prod_{1 \le j < k \le p-1} \left(2 \sin \frac{\pi(k-j)}{p}\right). \end{aligned}$$

Computing the exponent sum $\sum_{1 \le j < k \le p-1} (j+k) = \frac{p(p-1)(p-2)}{2}$ gives $\zeta_{2p}^{\sum(j+k)} = (-1)^{(p-1)/2}$, so:

$$\det(T) = (-1)^{(p-1)/2} \cdot i^{(p-1)(p-2)/2} \cdot \prod_{1 \le j < k \le p-1} \left(2 \sin \frac{\pi(k-j)}{p}\right). \tag{2.14}$$

Step 4: Resolving the sign for $r = 1$. Comparing (2.13) and (2.14), the first two factors coincide. Since $k - j \in [p-1]$ for all $1 \le j < k \le p-1$, we have $\sin \frac{\pi(k-j)}{p} > 0$, so $\prod_{j<k} 2 \sin \frac{\pi(k-j)}{p} > 0$. Therefore the product on the right of (2.14) equals $p^{(p-2)/2}$, and matching with (2.13) forces the $\pm$ to be $+$:

$$G(\eta, \chi_1) = i^{(p-1)^2/4} \cdot p^{1/2}.$$

Step 5: Extension to $r > 1$ via character lifting. For $r > 1$, lift η and χ_1 from $\mathbb{F}_p$ to $\mathbb{F}_q$ ($q = p^r$) via $\eta^{(r)} = \eta \circ N$ and $\chi_1^{(r)} = \chi_1 \circ \mathrm{Tr}$, where N and Tr are the absolute norm and trace from $\mathbb{F}_q$ to $\mathbb{F}_p$ respectively. By Davenport-Hasse Theorem,

$$G(\eta^{(r)}, \chi_1^{(r)}) = (-1)^{r-1} G(\eta, \chi_1)^r = (-1)^{r-1} \left(i^{(p-1)^2/4} \cdot p^{1/2} \right)^r = (-1)^r \cdot i^{\frac{(p-1)^2 r}{4}} \cdot q^{1/2}.$$

Since $\eta^{(r)}$ is the quadratic character of $\mathbb{F}_q$, this is the stated formula. ∎

The proof of the above theorem gives us another important result.

Lemma 2.8.2 Vandermonde Determinant with Roots of Unity

If p is a odd prime and ζ_p is the p^{th} root of unity. Let $V(\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1})$ is the $p \times p$ Vandermonde Matrix. Then

$$\det\left(V(\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1})\right) = (-1)^{(p-1)/2} \cdot i^{(p-1)(p-2)/2} \cdot p^{(p-2)/2}$$

As a striking application of the machinery developed around the quadratic character, we can now give a proof of one of the classical result of number theory – *Law of Quadratic Reciprocity*.

Theorem 2.8.3 Law of Quadratic Reciprocity

For any two distinct odd primes p, q we have

$$\left(\frac{p}{q}\right) \cdot \left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}$$

Proof: Let χ_1 be the canonical additive character of $\mathbb{F}_p$ and η is the quadratic character of $\mathbb{F}_p$. Then by Lemma 2.3.2(iv) we gave $G^2(\eta, \chi_1) = (-1)^{(p-1)/2} \cdot p$. So

$$G^q(\eta, \chi_1) = G(\eta, \chi_1) \cdot \left(G^2(\eta, \chi_1)\right)^{(q-1)/2} = (-1)^{(p-1)(q-1)/4} \cdot p^{(q-1)/2} \cdot G(\eta, \chi_1) \tag{2.15}$$

Now we calculate $G^q(\eta, \chi_1)$ in different way.

$$\begin{aligned} G^q(\eta, \chi_1) &= \left(\sum_{\alpha \in \mathbb{F}_p^*} \eta(\alpha) \cdot \chi_1(\alpha) \right)^q \\ &\equiv \sum_{\alpha \in \mathbb{F}_p^*} \eta^q(\alpha) \cdot \chi_q^q(\alpha) \bmod q \\ &\equiv \sum_{\alpha \in \mathbb{F}_p^*} \eta(\alpha) \cdot \chi_q(\alpha) \bmod q \\ &\equiv G(\eta, \chi_q) \bmod q \\ &\equiv \eta(q) G(\eta, \chi_1) \bmod q \qquad \text{[By Lemma 2.3.2(i)]} \end{aligned} \tag{2.16}$$

Now we already obtained $\left(\frac{q}{p}\right)$ in the right hand side above and $(-1)^{(p-1)(q-1)/4}$ in (2.15). Since η and χ_1 are both non-trivial characters of $\mathbb{F}_p$ we have $|G(\eta, \chi_1)| = p^{1/2}$ by Theorem 2.3.1 and so $G(\eta, \chi_1) \neq 0$. So now comparing (2.16) with (2.15) we

get

$$(-1)^{(p-1)(q-1)/4} \cdot p^{(q-1)/2} \cdot G(\eta, \chi_1) \equiv \eta(q) \cdot G(\eta, \chi_1) \bmod q \iff (-1)^{(p-1)(q-1)/4} \cdot p^{(q-1)/2} \equiv \eta(q) \bmod q$$

By Fermat's Little Theorem we have $p^{q-1} \equiv 1 \bmod q$ as p and q are two distinct primes. So we have

$$\begin{aligned} & (-1)^{(p-1)(q-1)/4} \cdot p^{(q-1)/2} \equiv \eta(q) \bmod q \\ \implies & (-1)^{(p-1)(q-1)/4} \cdot p^{(q-1)/2} \cdot p^{(q-1)/2} \equiv \eta(q) \cdot p^{(q-1)/2} \bmod q \\ \implies & (-1)^{(p-1)(q-1)/4} \equiv \eta(q) \cdot p^{(q-1)/2} \bmod q \\ \implies & (-1)^{(p-1)(q-1)/4} \equiv \left(\frac{q}{p}\right) \cdot \left(\frac{p}{q}\right) \bmod q \end{aligned}$$

Since both sides are ± 1 we can remove the modulo q. And therefore we have the theorem. ∎

We close the section with an elementary but remarkably clean evaluation of the quadratic character summed against a quadratic polynomial. Its value depends only on the leading coefficient and the discriminant, and it will play a basic role in many later arguments – most notably in extracting the quadratic character of the discriminant when counting solutions of quadratic equations over $\mathbb{F}_q$, and in the reformulation of Kloosterman sums in section 2.7.

Theorem 2.8.4 Quadratic Character Sum via Discriminant

Let $f(X) = a_2X^2 + a_1X + a_0 \in \mathbb{F}_q[X]$ with q odd and $a_2 \neq 0$, let $\Delta = a_1^2 - 4a_0a_2$ denote the discriminant of f, and let η be the quadratic character of $\mathbb{F}_q$ (extended by $\eta(0) = 0$). Then

$$\sum_{\alpha \in \mathbb{F}_q} \eta(f(\alpha)) = \begin{cases} -\eta(a_2) & \text{if } \Delta \neq 0, \\ (q-1) \cdot \eta(a_2) & \text{if } \Delta = 0. \end{cases}$$

We will not prove this theorem now. We will discuss this again in section 3.2.

§ 2.9 Jacobsthal Sums

Building further on the theme of character sums twisted by the quadratic character, we now introduce a classical family of sums first studied by Ernst Jacobsthal in 1907 in connection with the representation of primes $p \equiv 1 \bmod 4$ as sums of two squares. Unlike Gaussian and Jacobi sums, which are defined purely multiplicatively, Jacobsthal sums mix an additive shift with the quadratic character η of $\mathbb{F}_q$, which makes them particularly well suited to problems involving quadratic residues and the arithmetic of $\mathbb{F}_q^*$.

Definition 2.9.1: Jacobsthal Sums

For $\alpha \in \mathbb{F}_q^*$ where q is odd and for any $n \in \mathbb{N}$ the Jacobsthal Sum *is defined as*

$$H_n(\alpha) = \sum_{\gamma \in \mathbb{F}_q} \eta(\gamma^{n+1} + \alpha \cdot \gamma) = \sum_{\gamma \in \mathbb{F}_q} \eta(\gamma) \cdot \eta(\gamma^n + \alpha)$$

Another associated sum we define is

$$I_n(\alpha) = \sum_{\gamma \in \mathbb{F}_q} \eta(\gamma^n + \alpha)$$

for all $\alpha \in \mathbb{F}_q^*$.

Now for $n = 1$ we get $H_1(\alpha) = 1$ by Theorem 3.2.7 and by Theorem 2.2.2(i), $I_1(\alpha) = 0$ for all $\alpha \in \mathbb{F}_q^*$

The sums H_n and I_n are closely related: in the following theorem we give a recursion relation between I_n and H_m for any $n, m \in \mathbb{N}$.

Lemma 2.9.1 Relation between H_n and I_m

For every $n \in \mathbb{N}$ and every $\alpha \in \mathbb{F}_q^*$,

$$I_{2n}(\alpha) = I_n(\alpha) + H_n(\alpha)$$

Proof: For any $\beta \in \mathbb{F}_q$ let $Z(\beta)$ is the number of $\gamma \in \mathbb{F}_q$ such that $\gamma^2 = \beta$. Hence $Z(\beta) = 1 + \eta(\beta)$. Hence we have

$$\begin{aligned} I_{2n}(\alpha) &= \sum_{\gamma \in \mathbb{F}_q} \eta(\gamma^{2n} + \alpha) \\ &= \sum_{\beta \in \mathbb{F}_q} \eta(\beta^n + \alpha) \cdot Z(\beta) \\ &= \sum_{\beta \in \mathbb{F}_q} \eta(\beta^n + \alpha)(1 + \eta(\beta)) \\ &= \sum_{\beta \in \mathbb{F}_q} \eta(\beta^n + \alpha) + \sum_{\beta \in \mathbb{F}_q} \eta(\beta) \cdot \eta(\beta^n + \alpha) \\ &= I_n(\alpha) + H_n(\alpha) \end{aligned}$$

Hence we have the theorem. ■

2.9.1 Relations between Jacobsthal and Jacobi Sum

Although H_n and I_n are defined using only the quadratic character, they can be evaluated in terms of Jacobi sums for multiplicative characters of higher order. The key observation is that the count of n^{th}-power roots in $\mathbb{F}_q^*$ is a sum of multiplicative characters of order dividing $\gcd(n, q-1)$, and regrouping the definition of I_n accordingly expresses it as a linear combination of Jacobi sums $J(\lambda^j, \eta)$.

Theorem 2.9.2 Jacobi Sum Representation of I_n

Let $n \in \mathbb{N}$ and $\alpha \in \mathbb{F}_q^*$, and set $d = \gcd(n, q-1)$. Let $\lambda \in \boldsymbol{M}_q$ be a multiplicative character of $\mathbb{F}_q$ of order d. Then

$$I_n(\alpha) = \eta(\alpha) \sum_{j=1}^{d-1} \lambda^j(-\alpha) \cdot J(\lambda^j, \eta).$$

Proof: For any $\beta \in \mathbb{F}_q$ let $Z(\beta)$ is the number of $\gamma \in \mathbb{F}_q$ such that $\gamma^n = \beta$. Then we have

$$I_n(\alpha) = \sum_{\gamma \in \mathbb{F}_q} \eta(\gamma^n + \alpha) = \sum_{\beta \in \mathbb{F}_q} \eta(\beta + \alpha) \cdot Z(\beta)$$

Suppose $\beta \neq 0$. Then by Theorem 2.2.5(iii) we have

$$Z(\beta) = \frac{1}{q-1} \sum_{\gamma \in \mathbb{F}_q^*} \sum_{\psi \in \boldsymbol{M}_q} \psi(\gamma^n)\overline{\psi}(\beta) = \frac{1}{q-1} \sum_{\psi \in \boldsymbol{M}_q} \overline{\psi}(\beta) \sum_{\gamma \in \mathbb{F}_q^*} \psi^n(\gamma)$$

Now by Theorem 2.2.5(i), ψ^n is trivial then $\sum_{\gamma \in \mathbb{F}_q^*} \psi^n(\gamma) = q-1$ and otherwise its 0. Now ψ^n is trivial if and only if $\psi = \overline{\lambda}^j$ for all $j \in \{0, 1, \ldots, d-1\}$. Then we have

$$Z(\beta) = \sum_{j=0}^{d-1} \overline{\lambda}^j(\beta)$$

Suppose $\beta = 0$ then $Z(\beta) = 1$ as 0 is the only solution. Therefore this also holds for $\beta = 0$. Hence we have

$$\begin{aligned} I_n(\alpha) &= \sum_{\beta\in\mathbb{F}_q} \eta(\beta+\alpha) \sum_{j=0}^{d-1} \lambda^j(\beta) \\ &= \sum_{j=0}^{d-1} \sum_{\beta\in\mathbb{F}_q} \eta(\beta+\alpha)\lambda^j(\beta) \\ &= \eta(-1) \sum_{j=0}^{d-1} \sum_{\beta\in\mathbb{F}_q} \eta(-(\beta+\alpha))\lambda^j(\beta) \\ &= \eta(-1) \sum_{j=0}^{d-1} J_{-\alpha}(\lambda^j, \eta) \\ &= \eta(-1) \sum_{j=0}^{d-1} (\lambda^j \cdot \eta)(-\alpha) J(\lambda^j, \eta) \\ &= \eta(\alpha) \sum_{j=0}^{d-1} \lambda^j(-\alpha) J(\lambda^j, \eta) \end{aligned}$$

Hence we have the theorem. ■

The Jacobsthal sum H_n satisfies a similar Jacobi-sum representation, but now involving a character λ of order $2d$ rather than d, and with the parity of the power of 2 dividing $q-1$ relative to n playing a decisive role: when all the 2-power content of $q-1$ already divides n, the sum H_n vanishes identically.

Lemma 2.9.3 Jacobi Sum Representation of H_n

Let $n \in \mathbb{N}$ and $\alpha \in \mathbb{F}_q^*$, and set $d = \gcd(n, q-1)$. If the largest power of 2 dividing $q-1$ also divides n, then $H_n(\alpha) = 0$. Otherwise, let $\lambda \in \mathcal{M}_q$ be a multiplicative character of $\mathbb{F}_q$ of order $2d$; then

$$H_n(\alpha) = \eta(\alpha) \cdot \lambda(-1) \sum_{j=0}^{d-1} \lambda^{2j+1}(\alpha) \cdot J(\lambda^{2j+1}, \eta).$$

Proof: Since the largest power of 2 dividing $q-1$ also divides n we have $d = \gcd(n, q-1) = \gcd(2n, q-1)$. Therefore by Theorem 2.9.2, $I_n(\alpha) = I_{2n}(\alpha)$. Therefore by Lemma 2.9.1 we get $H_n(\alpha) = 0$.

Now suppose that is not the case. Then $\gcd(2n, q-1) = 2d$. Then by Theorem 2.9.2 we get

$$I_{2n}(\alpha) = \eta(\alpha) \sum_{j=1}^{2d-1} \lambda^j(-\alpha) J(\lambda^j, \eta) \tag{2.17}$$

Since $\mathrm{ord}(\lambda^2) = d$ using the same theorem for $I_n(\alpha)$ and λ^2 we get

$$I_n(\alpha) = \eta(\alpha) \sum_{j=0}^{d-1} \lambda^{2j}(-\alpha) J(\lambda^{2j}, \eta) \tag{2.18}$$

Therefore from Lemma 2.9.1 we have $H_n(\alpha) = I_{2n}(\alpha) - I_n(\alpha)$. So subtracting (2.18) from (2.17) we get the expression. ■

Note:-

By Theorem 2.4.6, $|J(\lambda^j,\eta)| = q^{1/2}$ for each $j \in \{1,\dots,d-1\}$ since λ^j and η are non-trivial and $\lambda^j\eta$ is non-trivial for $j \neq d/2$. Summing the $d-1$ terms in Theorem 2.9.2 therefore gives $|I_n(\alpha)| \leq (d-1)q^{1/2}$ for all $\alpha \in \mathbb{F}_q^*$.

2.9.2 Yet Another Proof of Fermat's Two Square Theorem

We close this section with the classical application of Jacobsthal sum. Using H_2 one can compute the two integers A, B in the representation $p = A^2 + B^2$ (guaranteed by Theorem 2.4.12) directly, and moreover with a canonical sign choice determined by a congruence modulo 4. Here $p \equiv 1 \bmod 4$ is prime, η is the quadratic character of $\mathbb{F}_p$, and $\lambda \in \boldsymbol{\mathcal{M}}_p$ is a multiplicative character of $\mathbb{F}_p$ of order 4, so that $\lambda^2 = \eta$ and $\lambda^3 = \overline{\lambda}$.

By Theorem 2.9.3 applied with $n = 2$ and hence $d = \gcd(2, p-1) = 2$,

$$H_2(1) = \lambda(-1)\sum_{j=0}^{1}\lambda^{2j+1}(1)\cdot J(\lambda^{2j+1},\eta) = \lambda(-1)\left(J(\lambda,\eta) + J(\lambda^3,\eta)\right).$$

Since $\lambda^3 = \overline{\lambda}$ and η is real, $J(\lambda^3,\eta) = \overline{J(\lambda,\eta)}$, so $H_2(1) = 2\lambda(-1)\cdot\Re J(\lambda,\eta)$. In the notation of Theorem 2.4.12, $J(\lambda,\eta) = A + iB$ with $A, B \in \mathbb{Z}$ and $p = A^2 + B^2$, so

$$\Re J(\lambda,\eta) = A = \tfrac{1}{2}\lambda(-1)\cdot H_2(1).$$

A similar computation starting from an $\alpha \in \mathbb{F}_p$ with $\eta(\alpha) = -1$ uses $\lambda(\alpha) = \pm i$ and $\lambda^3(\alpha) = -\lambda(\alpha)$ to give

$$H_2(\alpha) = \eta(\alpha)\lambda(-1)\left(\lambda(\alpha)J(\lambda,\eta) - \lambda(\alpha)\overline{J(\lambda,\eta)}\right) = -\lambda(-1)\lambda(\alpha)\cdot 2i\cdot\Im J(\lambda,\eta),$$

which identifies $\Im J(\lambda,\eta) = B$ up to the unit factor $-\lambda(-1)\lambda(\alpha) \in \{\pm 1, \pm i\}$; absorbing this factor gives $B = \frac{1}{2}H_2(\alpha)$ (up to sign).

To pin down the sign of A, we expand $H_2(1)$ directly. Pairing γ with $-\gamma$ and using $\eta(-1) = 1$ (since $p \equiv 1 \bmod 4$) gives

$$H_2(1) = \sum_{\gamma=1}^{p-1}\eta(\gamma)\cdot\eta(\gamma^2+1) = 2\sum_{\gamma=1}^{(p-1)/2}\eta(\gamma)\cdot\eta(\gamma^2+1),$$

so $\frac{1}{2}H_2(1) = \sum_{\gamma=1}^{(p-1)/2}\eta(\gamma)\eta(\gamma^2+1)$. On the other hand Theorem 3.2.7 applied to $f(X) = X^2+1$ (discriminant $-4 \neq 0$ and $\eta(1) = 1$) gives $\sum_{\gamma\in\mathbb{F}_p}\eta(\gamma^2+1) = -1$, so $\sum_{\gamma=1}^{(p-1)/2}\eta(\gamma^2+1) = -1$ (excluding the zero term). Subtracting and reducing modulo 4, using the fact that $(\eta(\gamma)-1)(\eta(\gamma^2+1)-1) \equiv 0 \bmod 4$ whenever $\eta(\gamma^2+1) \neq 0$, one obtains after a short computation

$$\tfrac{1}{2}H_2(1) + 1 \equiv \tfrac{3-p}{2} - \lambda(-1) \bmod 4.$$

Since $\lambda(-1) = 1$ when $p \equiv 1 \bmod 8$ and $\lambda(-1) = -1$ when $p \equiv 5 \bmod 8$, in both cases $\frac{1}{2}H_2(1) + 1 \equiv 0 \bmod 4$, i.e. $\frac{1}{2}H_2(1) \equiv -1 \bmod 4$.

Hence setting $A = \frac{1}{2}\lambda(-1)H_2(1)$ and $B = \frac{1}{2}H_2(\alpha)$ (for any $\alpha \in \mathbb{F}_p$ with $\eta(\alpha) = -1$), we obtain integers with $A^2 + B^2 = p$, A odd, B even, and the normalization $A \equiv -1 \bmod 4$. With this normalization A is uniquely determined: if $p = A^2 + B^2 = C^2 + D^2$ with A, C odd and $A \equiv C \equiv -1 \bmod 4$, then choosing $h, k \in \mathbb{Z}$ with $A \equiv hB$, $C \equiv kD \bmod p$ gives $h^2 + 1 \equiv k^2 + 1 \equiv 0 \bmod p$, hence $C \equiv \pm hD \bmod p$; writing $C_1 \equiv hD \bmod p$ and $C = \pm C_1$, we compute

$$p^2 = (A^2+B^2)(C_1^2+D^2) = (AC_1 + BD)^2 + (AD - BC_1)^2,$$

and both $AC_1 + BD \equiv (h^2+1)BD \equiv 0 \bmod p$ and $AD - BC_1 \equiv 0 \bmod p$. Dividing by p^2 gives $1 = (\pm 1)^2 + 0^2$, forcing $AD - BC_1 = 0$ and $|AC_1 + BD| = p$. Since $\gcd(A,B) = \gcd(C_1, D) = 1$, it follows that $A = \pm C_1$ and hence $A = \pm C$; the congruence $A \equiv C \equiv -1 \bmod 4$ then forces $A = C$.

Thus the Jacobsthal sum $\frac{1}{2}\lambda(-1)H_2(1)$ produces the canonical integer A in Fermat's representation $p = A^2 + B^2$, and $\frac{1}{2}H_2(\alpha)$ produces B up to sign.

§ 2.10 Salié Sums

The Kloosterman sums of the previous section are notoriously difficult to evaluate in closed form – even their size is controlled only by the deep theorem of Weil. Remarkably, a very mild modification of the Kloosterman sum, obtained by inserting the quadratic character η as a weight, produces a family of sums for which exact evaluations *are* possible. These are the *Salié sums*, introduced by Hans Salié in 1931.

Definition 2.10.1: Salié Sum

Let $\mathbb{F}_q$ be a finite field with q odd, and let $\chi, \tau \in \boldsymbol{\mathcal{X}}_q$ be non-trivial additive characters of $\mathbb{F}_q$. Then the Salié sum *associated to χ and τ is defined as*

$$S(\chi,\tau)=\sum_{\gamma\in\mathbb{F}_q^*}\eta(\gamma)\cdot\chi(\gamma)\cdot\tau(\gamma^{-1})$$

Theorem 2.10.1 Evaluation of Salié Sums

Let $\mathbb{F}_q$ be a finite elements with q being odd and $\chi, \tau \in \boldsymbol{\mathcal{X}}_q$ be non-trivial additive character of $\mathbb{F}_q$. Let $a \in \mathbb{F}_q$ such that for all $\gamma \in \mathbb{F}_q$, $\tau(\gamma) = \chi(a \cdot \gamma)$. Then we have

$$S(\chi,\tau)=G(\eta,\chi)\sum_{b\,:\,b^2=4a}\chi(b)$$

Proof: Since $\tau(\gamma) = \chi(a\cdot\gamma)$ for all $\gamma \in \mathbb{F}_q$ we have $\chi(\gamma)\cdot\tau(\gamma^{-1}) = \chi(\gamma)\cdot\chi(a\cdot\gamma^{-1}) = \chi(\gamma + a\gamma^{-1})$, so $S(\chi,\tau) = \sum_{\gamma\in\mathbb{F}_q^*}\eta(\gamma)\chi(\gamma+a\gamma^{-1})$.

Step 1: Multiplicative Fourier expansion. For $b \in \mathbb{F}_q^*$ let $\chi_b \in \boldsymbol{\mathcal{X}}_q$ denote the additive character $\chi_b(\gamma) = \chi(b\gamma)$, similar to as given by Theorem 2.2.1. Define $\varphi : \mathbb{F}_q^* \to \mathbb{C}$ by

$$\varphi(b)=S(\chi_b,\tau)=\sum_{\gamma\in\mathbb{F}_q^*}\eta(\gamma)\chi(b\gamma+a\gamma^{-1}).$$

By Theorem 2.2.5, the multiplicative characters form an orthonormal basis for functions on $\mathbb{F}_q^*$, so we may expand

$$\varphi(b)=\sum_{\lambda\in\boldsymbol{\mathcal{M}}_q}\hat{\varphi}(\lambda)\lambda(b),\quad\text{where}\quad\hat{\varphi}(\lambda)=\frac{1}{q-1}\sum_{b\in\mathbb{F}_q^*}\varphi(b)\overline{\lambda}(b).$$

Step 2: Computing the Fourier coefficients. Substituting the definition of φ and interchanging the order of summation:

$$\begin{aligned}\hat{\varphi}(\lambda)&=\frac{1}{q-1}\sum_{b\in\mathbb{F}_q^*}\overline{\lambda}(b)\sum_{\gamma\in\mathbb{F}_q^*}\eta(\gamma)\chi(b\gamma+a\gamma^{-1})\\&=\frac{1}{q-1}\sum_{\gamma\in\mathbb{F}_q^*}\eta(\gamma)\chi(a\gamma^{-1})\sum_{b\in\mathbb{F}_q^*}\overline{\lambda}(b)\chi(b\gamma).\end{aligned}$$

Substituting $b \mapsto b\gamma^{-1}$ in the inner sum gives $\sum_{b\in\mathbb{F}_q^*}\overline{\lambda}(b)\chi(b\gamma) = \lambda(\gamma)\sum_{b\in\mathbb{F}_q^*}\overline{\lambda}(b)\chi(b) = \lambda(\gamma)\cdot G(\overline{\lambda},\chi)$. Hence

$$\hat{\varphi}(\lambda)=\frac{G(\overline{\lambda},\chi)}{q-1}\sum_{\gamma\in\mathbb{F}_q^*}\eta(\gamma)\lambda(\gamma)\chi(a\gamma^{-1}).$$

Substituting $\gamma \mapsto \gamma^{-1}$ in the remaining sum:

$$\sum_{\gamma\in\mathbb{F}_q^*}\eta(\gamma)\lambda(\gamma)\chi(a\gamma^{-1})=\sum_{\gamma\in\mathbb{F}_q^*}\overline{\eta}(\gamma)\overline{\lambda}(\gamma)\chi(a\gamma)=G(\overline{\eta}\cdot\overline{\lambda},\chi_a)$$

where $\chi_a(\gamma) = \chi(a\gamma)$. Since η is real-valued, $\overline{\eta} = \eta$. By Lemma 2.3.2(i), $G(\lambda, \chi_a) = \overline{\lambda}(a)G(\lambda, \chi)$, so

$$G(\overline{\eta} \cdot \overline{\lambda}, \chi_a) = \eta(a)\lambda(a) \cdot G(\overline{\eta} \cdot \overline{\lambda}, \chi).$$

Altogether:

$$\hat{\varphi}(\lambda) = \frac{\eta(a)\lambda(a) \cdot G(\overline{\lambda}, \chi) \cdot G(\overline{\eta} \cdot \overline{\lambda}, \chi)}{q-1}. \tag{2.19}$$

Step 3: Applying the Hasse-Davenport product formula. By Corollary 2.4.11 applied with $m = 2$, the quadratic character η, and $\overline{\lambda}$ in place of ψ, we have

$$G(\overline{\lambda}, \chi) \cdot G(\overline{\lambda}\eta, \chi) = \lambda(4) \cdot G(\overline{\lambda}^2, \chi) \cdot G(\eta, \chi).$$

Since $\overline{\eta} \cdot \overline{\lambda} = \overline{\lambda}\eta$, substituting into (2.19):

$$\hat{\varphi}(\lambda) = \frac{\eta(a)G(\eta, \chi)}{q-1} \cdot \lambda(a) \cdot \lambda(4) \cdot G(\overline{\lambda}^2, \chi) = \frac{\eta(a)G(\eta, \chi)}{q-1} \cdot \lambda(4a) \cdot G(\overline{\lambda}^2, \chi).$$

Step 4: Final Step. Since $\lambda(1) = 1$ for every $\lambda \in \boldsymbol{\mathcal{M}}_q$:

$$\begin{aligned} S(\chi, \tau) = \varphi(1) = \sum_{\lambda \in \boldsymbol{\mathcal{M}}_q} \hat{\varphi}(\lambda) &= \frac{\eta(a)G(\eta, \chi)}{q-1} \sum_{\lambda \in \boldsymbol{\mathcal{M}}_q} \lambda(4a) \cdot G(\overline{\lambda}^2, \chi) \\ &= \frac{\eta(a)G(\eta, \chi)}{q-1} \sum_{\lambda \in \boldsymbol{\mathcal{M}}_q} \lambda(4a) \sum_{\gamma \in \mathbb{F}_q^*} \overline{\lambda}^2(\gamma)\chi(\gamma) \\ &= \frac{\eta(a)G(\eta, \chi)}{q-1} \sum_{\gamma \in \mathbb{F}_q^*} \chi(\gamma) \sum_{\lambda \in \boldsymbol{\mathcal{M}}_q} \lambda(4a\gamma^{-2}) \end{aligned}$$

By the orthogonality of multiplicative characters (Theorem 2.2.5(ii)):

$$\sum_{\lambda \in \boldsymbol{\mathcal{M}}_q} \lambda(4a\gamma^{-2}) = \begin{cases} q-1 & \text{if } \gamma^2 = 4a, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $S(\chi, \tau) = \eta(a) \cdot G(\eta, \chi) \sum_{b: b^2 = 4a} \chi(b)$. If $\eta(a) = -1$ then $4a$ is a non-square so $b^2 = 4a$ has no solution and the sum is 0; if $\eta(a) = 1$ the factor is 1. In both cases:

$$S(\chi, \tau) = G(\eta, \chi) \sum_{b:\ b^2 = 4a} \chi(b).$$

Therefore we have the theorem. ∎

Lemma 2.10.2 Bound on Salié Sums

Let $\mathbb{F}_q$ be a finite elements with q being odd and $\chi, \tau \in \boldsymbol{\mathcal{X}}_q$ be non-trivial additive character of $\mathbb{F}_q$. Let $a \in \mathbb{F}_q$ such that for all $\gamma \in \mathbb{F}_q$, $\tau(\gamma) = \chi(a \cdot \gamma)$. Then $|S(\chi, \tau)| \leq 2 \cdot q^{1/2}$.

Proof: Since $a \neq 0$, the equation $b^2 = 4a$ has either 0 or 2 solutions in $\mathbb{F}_q$. From Theorem 2.10.1 if 0 solutions, $S(\chi, \tau) = 0$ and if $\pm b_0$ are the two solutions, $S(\chi, \tau) = G(\eta, \chi)\ (\chi(b_0) + \chi(-b_0))$. Each term $G(\eta, \chi)\chi(\pm b_0)$ is a product of the Gauss sum with a root of unity. By Theorem 2.3.1, $|G(\eta, \chi)| = q^{1/2}$, so $|S(\chi, \tau)| \leq 2q^{1/2}$. ∎

Chapter 3

Equations over Finite Fields

The central question of this chapter is: given a polynomial $f \in \mathbb{F}_q[X_1, \ldots, X_n]$, how many solutions does $f(X_1, \ldots, X_n) = 0$ have in $\mathbb{F}_q^n$? We write $\mathcal{V}(f)$ for the solution set and $Z(f) = |\mathcal{V}(f)|$ for its size; for a system $f_1 = \cdots = f_k = 0$ we write $\mathcal{V}(f_1, \ldots, f_k) = \mathcal{V}(f_1) \cap \cdots \cap \mathcal{V}(f_k)$ and $Z(f_1, \ldots, f_k)$ for the common solution count. Variants $\overline{\mathcal{V}}$ and $\overline{Z}$ track the non-trivial (nonzero) solutions. The notation mirrors that of varieties of algebraic geometry: $\mathcal{V}(f)$ is the affine variety cut out by f.

The chapter develops two complementary strategies. The first is *direct combinatorics*: for special polynomial families (univariate polynomials, systems of low total degree, quadratic forms, diagonal equations) one can write exact formulas or prove sharp divisibility constraints on $Z(f)$. The second is *character sum methods*: the indicator of whether α is a solution can be encoded as an additive character sum, turning the problem of counting solutions into the problem of bounding $\sum_\alpha \chi(f(\alpha))$. These two strategies reinforce each other throughout the chapter.

§ 3.1 Some Elementary Results

In this section we give some elementary results on the number of solutions. We present some classical theorems such as those of König-Rados, Chevalley, and Warning. We also establish elementary upper bounds for the number of solutions and results on the expected order of magnitude.

3.1.1 Univariate Polynomials

We start of with univariate polynomials. Let $f \in \mathbb{F}_q[X]$ be a univariate polynomial Naturally the number of solutions in $\mathbb{F}_q$ for the equation $f(X) = 0$ is basically the degree of $\gcd(f(X), X^q - X)$. Since checking if 0 is a solution or not is very easy to check we'll only consider non-zero solutions. So the number of non-zero solutions of $f(X) = 0$ is basically the degree of $\gcd(f(X), X^{q-1} - 1)$. Therefore we may always assume without loss of generality that $\deg(f) \leq q - 1$. Suppose

$$f(X) = a_{q-1} \cdot X^{q-1} + a_{q-2} \cdot X^{q-2} + \cdots + a_1 \cdot X + a_0$$

where $a_i \in \mathbb{F}_q$ for all $0 \leq i \leq q - 1$. Since for all $\alpha \in \mathbb{F}_q^*$ we have $\alpha^{q-1} = 1$ the variable X^{q-1} will always evaluate to be 1 for all non-zero values of X. Therefore the number of solutions of f is same as of the equation $a_{q-2} \cdot X^{q-2} + \cdots + a_1 \cdot X + (a_0 + a_{q-1}) = 0$. Therefore we can always assume $\deg(f) \leq q - 2$. So suppose we have a polynomial $f \in \mathbb{F}_q[X]$ such that

$$f(X) = a_{q-2} \cdot X^{q-2} + \cdots + a_1 \cdot X + a_0$$

where $a_i \in \mathbb{F}_q$ for all $0 \leq i \leq q - 2$.

Now we can also analyze the number of solutions of f using matrix methods. From f we can define the following matrix

$$M_f = \begin{bmatrix} a_0 & a_1 & \cdots & a_{q-2} \\ a_1 & a_2 & \cdots & a_0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{q-2} & a_0 & \cdots & a_{q-3} \end{bmatrix}$$

This matrix is a *left circulant matrix* i.e. each row is obtained from the preceding row by a cyclic shift of the entries to the left. The number of non-zero solutions of $f(X) = 0$ and the rank of the matrix M_f are very related.

Theorem 3.1.1 König-Rados Theorem

Let $f(X) = \sum_{i=0}^{q-2} a_i \cdot X^i \in \mathbb{F}_q[X]$ be a univariate polynomial. Then

$$\overline{Z}(f) = q - 1 - \text{rank}(M_f)$$

Proof: Since $\deg(f) \leq q-2$ we have $\overline{Z}(f) \leq q-2$. $\overline{Z}(f) = q-1$ rises only when f is a zero polynomial i.e. M_f is a zero matrix and in that case $\text{rank}(M_f) = 0$ which satisfies the theorem. Let γ be the primitive element of $\mathbb{F}_q$. Then consider the *Vandermonde matrix*, $V(\gamma, \gamma^2, \ldots, \gamma^{q-1})$:

$$V = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \gamma & \gamma^2 & \cdots & \gamma^{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{q-2} & \gamma^{2(q-2)} & \ldots & \gamma^{(q-1)(q-2)} \end{bmatrix}$$

Then using $\gamma^{i\cdot(q-1)} = 1$ for all $i \in [q-1]$ we obtain

$$M_f \cdot V = \begin{bmatrix} f(\gamma) & f(\gamma^2) & \cdots & f(\gamma^{q-1}) \\ \gamma^{-1}f(\gamma) & \gamma^{-2}f(\gamma^2) & \cdots & \gamma^{-(q-1)}f(\gamma^{q-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{-(q-2)}f(\gamma) & \gamma^{-2(q-2)}f(\gamma^2) & \cdots & \gamma^{-(q-1)(q-2)}f(\gamma^{q-1}) \end{bmatrix}$$

Now for any $i \in [q-1]$ we have $f(\gamma^i) = 0$ if and only if the i^{th} row becomes zero. Therefore the $\text{rank}(M_f \cdot V) = q - 1 - \overline{Z}(f)$. Since V is the Vandermonde matrix V and $M_f \cdot V$ have the same rank. So the theorem follows. ■

In the next subsections we will give simple upper and lower bounds for solutions multivariate polynomial equations or system of polynomial equations.

3.1.2 Chevalley–Warning Theorem

We begin with two elementary lemmas about power sums that will serve as the key tools.

Lemma 3.1.2

Let k be a non-negative integer. Suppose for $k = 0$, $0^0 = 1$. Then

$$\sum_{\alpha \in \mathbb{F}_q} \alpha^k = \begin{cases} 0 & \text{if } k = 0 \text{ or } q-1 \nmid k \\ -1 & \text{if } q-1 \mid k \end{cases}$$

Proof: For $k = 0$ the sum becomes $q \equiv 0 \bmod q$. So suppose $k > 0$. Let γ be the primitive element of $\mathbb{F}_q$. Then we have

$$\sum_{\alpha \in \mathbb{F}_q} \alpha^k = \sum_{\alpha \in \mathbb{F}_q^*} \alpha^k = \sum_{j=0}^{q-2} \gamma^{j \cdot k} = \sum_{j=0}^{q-2} (\gamma^k)^j$$

Now if $q - 1 \mid k$ then $\gamma^k = 1$ so the sum evaluates to be -1. If $q - 1 \nmid k$ then $\sum_{j=0}^{q-2} (\gamma^k)^j = \frac{\gamma^{k(q-1)} - 1}{\gamma^k - 1} = 0$. Hence we have the lemma. ∎

So using this lemma we can prove if we take the sum over all evaluations of a polynomial then the sum evaluates to be zero.

Lemma 3.1.3

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ with $\deg(f) < n(q-1)$. Then

$$\sum_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q} f(\alpha_1, \ldots, \alpha_k) = 0$$

Proof: It suffices to prove this property for every monomial of f. So suppose $X_1^{k_1} \cdots X_n^{k_n}$ be a monomial of f where $k_1 + \cdots + k_n < n(q-1)$ and $k_i \in \mathbb{Z}_0$ for all $i \in [n]$. Since $k_1 + \cdots + k_n < n(q-1)$ then by Pigeon Hole Principle there exists $j \in [n]$ such that $k_j < q - 1$. Then

$$\sum_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q} \alpha_1^{k_1} \cdots \alpha_n^{k_n} = \prod_{i=1}^{n} \left(\sum_{\alpha_i \in \mathbb{F}_q} \alpha^{k_i} \right)$$

Then for j, by Lemma 3.1.2 $\sum_{\alpha_j \in \mathbb{F}_q} \alpha^{k_j} = 0$. So we have the lemma. ∎

The two lemmas above are the principal tools behind the Chevalley–Warning theorem. It asserts that whenever the total degree of a system is small relative to the number of variables, the number of common zeros is divisible by p — forcing, in particular, that a system with one solution must have a second.

Theorem 3.1.4 Chevalley–Warning Theorem

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ with $\deg(f) < n$. Then the number of solutions of the equation $f(X_1, \ldots, X_n) = 0$ in $\mathbb{F}_q^n$ is divisible by the characteristic p of $\mathbb{F}_q$.

Moreover, if $f(0, \ldots, 0) = 0$ then there exists a nontrivial solution in $\mathcal{V}(f)$ i.e. $\overline{\mathcal{V}}(f) \neq \emptyset$ or $\overline{Z}(f) \geq 1$.

Proof: We first construct a polynomial and then will use Lemma 3.1.3. So take the polynomial $F(X_1, \ldots, X_n) = 1 - f^{q-1}(X_1, \ldots, X_n)$. By construction F always takes vale 0 or 1. Therefore for any $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ we have

$$F(\alpha_1, \ldots, \alpha_n) = 1 \iff f(\alpha_1, \ldots, \alpha_n) = 0$$

So

$$\sum_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q} F(\alpha_1, \ldots, \alpha_k) = Z(f) \bmod p$$

Now by Lemma 3.1.3 we have

$$\sum_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q} F(\alpha_1, \ldots, \alpha_k) = 0$$

Therefore $p \mid Z(f)$. So we have the first part.

Now since $f(0,\dots,0) = 0$ we have $Z(f) \geq 1$. Since p is a prime $p \geq 2$. So by first part $p \mid Z(f)$. Hence $Z(f) \geq 2$. So we have the second part too. ■

Note:-

The bound $\deg(f) < n$ in Chevalley–Warning Theorem is the best possible: the norm polynomial $f(X_1,\dots,X_n) = \prod_{j=0}^{n-1}\left(\sum_{i=1}^{n} \alpha_i^{q^j} X_i\right)$, where $\{\alpha_1,\dots,\alpha_n\}$ is a basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$, has degree n and satisfies $f(\gamma_1,\dots,\gamma_n) = N_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\sum_i \alpha_i\gamma_i)$, which is zero only when $\gamma_j = 0$ for all j. Thus $Z(f) = 1$, so divisibility by $p \geq 2$ fails.

Now we can extend the Chevalley–Warning Theorem to a system of polynomial equations $f_1 = 0,\dots,f_k = 0$ with $f_1,\dots,f_k \in \mathbb{F}_q[X_1,\dots,X_n]$. In this case we are interested in the number of common solutions.

Theorem 3.1.5 Chevalley–Warning Theorem for Systems of Equations

Let $f_1,\dots,f_k \in \mathbb{F}_q[X_1,\dots,X_n]$ with $\deg(f_1\cdots f_k) = \deg(f_1)+\cdots+\deg(f_k) < n$. Then $p \mid Z(f_1,\dots,f_k)$ where $p = \mathrm{char}(\mathbb{F}_q)$.

Moreover, if $(0,\dots,0) \in \mathcal{V}(f_1,\dots,f_k)$ then $\overline{\mathcal{V}}(f_1,\dots,f_k) \neq \emptyset$ i.e. $\overline{Z}(f_1,\dots,f_k) \geq 1$.

Proof: Again we construct a polynomial. Consider the polynomial $F \in \mathbb{F}_q[X_1,\dots,X_n]$ such that $F = (1-f_1^{q_1})\cdots(1-f_k^{q-1})$. Therefore for any $\alpha_1,\dots,\alpha_n \in \mathbb{F}_q$ we have $F(\alpha_1,\dots,\alpha_n) = 1$ if and only if $f_i(\alpha_1,\dots,\alpha_n) = 0$ for all $i \in [k]$. Now $\deg(F) \leq \deg(f_1)+\cdots+\deg(f_k) < n(q-1)$. Now by applying Chevalley–Warning Theorem for F we get the theorem. ■

3.1.3 Lower Bounds on the Number of Solutions

In the next theorem we will show if we have a system of polynomial equations $f_1 = 0,\dots,f_k = 0$ with $f_1,\dots,f_k \in \mathbb{F}_q[X_1,\dots,X_n]$ then for any affine subspace W all its parallel shifts have same number of solutions modulo the characteristic of $\mathbb{F}_q$. If W_1 and W_2 are two affine subspaces of $\mathbb{F}_q^n$ of same dimension then we call they are parallel affine subspaces if they are obtained by translation from the same linear subspace.

Lemma 3.1.6 Parallel Shift Invariance of Solution Counts

Let $f_1,\dots,f_k \in \mathbb{F}_q[X_1,\dots,X_n]$. If W_1 and W_2 are two parallel affine subspaces of $\mathbb{F}_q^n$ of dimension d where $d = \deg(f_1)+\cdots+\deg(f_k)$ then

$$|W_1 \cap \mathcal{V}(f_1,\dots,f_k)| \equiv |W_2 \cap \mathcal{V}(f_1,\dots,f_k)| \bmod p$$

where $p = \mathrm{char}(\mathbb{F}_q)$.

Proof: Now if $W_1 = W_2$, we get $|W_1 \cap \mathcal{V}(f_1,\dots,f_k)| = |W_2 \cap \mathcal{V}(f_1,\dots,f_k)|$. So we get the theorem. So we can assume $W_1 \neq W_2$. Now by change of coordinates we can take

$$W_1 = \{(\alpha_1,\dots,\alpha_n) \in \mathbb{F}_q^n \colon \alpha_i = 0,\ \forall\, i \in [n]\}$$
$$W_2 = \{(\alpha_1,\dots,\alpha_n) \in \mathbb{F}_q^n \colon \alpha_1 = 1 \text{ and } \alpha_i = 0,\ \forall\, 2 \leq i \leq n\}$$

Again we construct a polynomial $G \in \mathbb{F}_q[X_1,\dots,X_n]$ where

$$G(X_1,\dots,X_n) = (-1)^{n-d}(X_1^{q-2}+\cdots X_1+1)(X_2^{q-1}-1)\cdots(X_n^{q-1}-1)$$

So $\deg(G) = (n-d)(q-1)-1$.

Observation 3.1. *With the above construction of G for any $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_q^n$*

(i) if $(\alpha_1, \ldots, \alpha_n) \in W_1$ then $G(\alpha_1, \ldots, \alpha_n) = -1$

(ii) if $(\alpha_1, \ldots, \alpha_n) \in W_2$ then $G(\alpha_1, \ldots, \alpha_n) = 1$

(iii) if $(\alpha_1, \ldots, \alpha_n) \notin W_1 \cup W_2$ then $G(\alpha_1, \ldots, \alpha_n) = 0$

Then we construct the following polynomial $F \in \mathbb{F}_q[X_1, \ldots, X_n]$ such that

$$F = (1 - f_1^{q-1}) \cdots (1 - f_k^{q-1})G$$

So now we have $\deg(F) \leq d(q-1) + (n-d)(q-1) - 1 = n(q-1) - 1$. Therefore for any $(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}_q^n$, if $(\alpha_1, \ldots, \alpha_n) \in W_1 \cap \mathcal{V}(f_1, \ldots, f_k)$ then $F(\alpha_1, \ldots, \alpha_n) = 1$. If $(\alpha_1, \ldots, \alpha_n) \in W_2 \cap \mathcal{V}(f_1, \ldots, f_k)$ then $F(\alpha_1, \ldots, \alpha_n) = -1$ and 0 elsewhere. So

$$\sum_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q} F(\alpha_1, \ldots, \alpha_n) = |W_1 \cap \mathcal{V}(f_1, \ldots, f_k)| - |W_2 \cap \mathcal{V}(f_1, \ldots, f_k)|$$

Now by using Chevalley–Warning Theorem on F we have the lemma. ∎

The Lemma 3.1.6 above can be used to obtain a lower bound on $Z(f_1, \ldots, f_k)$: the Chevalley–Warning Theorem guarantees that $p \mid Z(f_1, \ldots, f_k)$ whenever $d < n$, but by counting how many parallel affine hyperplanes each must contain a solution we can say much more.

Theorem 3.1.7 Lower Bound via Parallel Shifts

Let $f_1, \ldots, f_k \in \mathbb{F}_q[X_1, \ldots, X_n]$ with $d = \deg(f_1) + \cdots + \deg(f_k) < n$. If $Z(f_1, \ldots, f_k) \geq 1$ then $Z(f_1, \ldots, f_k) \geq q^{n-d}$

Proof: We will prove this by analyzing two cases.

Case I: Let there exists an affine subspace W_1 of $\mathbb{F}_q^n$ with dimension d such that $|W_1 \cap \mathcal{V}(f_1, \ldots, f_k)| \not\equiv 0 \bmod p$. Then by Lemma 3.1.6 for all parallel affine spaces W_2 of W_1 we have $|W_2 \cap \mathcal{V}(f_1, \ldots, f_k)| \not\equiv 0 \bmod p$. In particular $|W_2 \cap \mathcal{V}(f_1, \ldots, f_k)| \geq 1$. Since W_1 is of dimension d there are q^{n-d} distinct affine subspaces parallel to W_1. Therefore

$$\mathcal{V}(f_1, \ldots, f_k) = \bigsqcup_{\substack{W : \text{ affine subspace} \\ \text{parallel to } W_1}} \Big(W \cap \mathcal{V}(f_1, \ldots, f_k) \Big)$$

For each such W we have $|W \cap \mathcal{V}(f_1, \ldots, f_k)| \geq 1$. Therefore

$$Z(f_1, \ldots, f_k) = \sum_{\substack{W : \text{ affine subspace} \\ \text{parallel to } W_1}} |W \cap \mathcal{V}(f_1, \ldots, f_k)| \geq q^{n-d}$$

Therefore we have the theorem in this case.

Case II: We have $|W \cap \mathcal{V}(f_1, \ldots, f_k)| \equiv 0 \bmod p$. for all affine subspaces W of $\mathbb{F}_q^n$ of dimension d. Since $Z(f_1, \ldots, f_k) \geq 1$ there exists $k \in [d]$ such that for any affine subspace W' of dimension k, $|W' \cap \mathcal{V}(f_1, \ldots, f_k)| \equiv 0 \bmod p$ but there is an affine subspace W'' of dimension $(k-1)$ such that $|W'' \cap \mathcal{V}(f_1, \ldots, f_k)| \not\equiv 0 \bmod p$. Now consider all affine subspaces W' of dimension k containing W''. There are exactly $\frac{q^{n-(k-1)}-1}{q-1}$ many such affine subspaces. For each such affine subspace W' we have

$$|(W' \setminus W'') \cap \mathcal{V}(f_1, \ldots, f_k)| = |W' \cap \mathcal{V}(f_1, \ldots, f_k)| - |W'' \cap \mathcal{V}(f_1, \ldots, f_k)| \not\equiv 0 \bmod p$$

So $|(W' \setminus W'') \cap \mathcal{V}(f_1, \ldots, f_k)| \geq 1$. Therefore we have

$$\mathcal{V}(f_1, \ldots, f_k) = \Big(W'' \cap \mathcal{V}(f_1, \ldots, f_k)\Big) \sqcup \bigsqcup_{\substack{W' : \text{affine subspace} \\ \text{of dimension } k \\ \text{contains } W''}} \Big((W' \setminus W'') \cap \mathcal{V}(f_1, \ldots, f_k)\Big)$$

Hence we have

$$Z(f_1, \ldots, f_k) = |W'' \cap \mathcal{V}(f_1, \ldots, f_k)| + \sum_{\substack{W' : \text{affine subspace} \\ \text{of dimension } k \\ \text{contains } W''}} |(W' \setminus W'') \cap \mathcal{V}(f_1, \ldots, f_k)| \geq 1 + \frac{q^{n-(k-1)} - 1}{q - 1} > q^{n-d}$$

So we have the theorem. ∎

The lower bound on the number of solutions above is best possible even for $k = 1$ i.e. for any positive integers d and n with $d < n$ there is a polynomial $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ of degree d such that $f(X_1, \ldots, X_n) = 0$ has exactly q^{n-d} solutions in $\mathbb{F}_q^n$. Consider the polynomial defined in the support Chevalley–Warning Theorem, $\prod_{i=0}^{n-1} \sum_{j=1}^{n} \alpha_j^{q^i} X_j$ where $\{\alpha_j : j \in [n]\}$ was a basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. But in this example we take the same polynomial but on d variables but the solution space is over $\mathbb{F}_{q^n}$. Hence our polynomial in this case is $f(X_1, \ldots, X_n) = g(X_1, \ldots, X_d) = \prod_{i=0}^{n-1} \sum_{j=1}^{d} \alpha_j^{q^i} X_j$. Since f had only one solution, g has a solution at $(\gamma_1, \ldots, \gamma_n)$ if $\gamma_i = 0$ for all $i \in [d]$. Therefore g has exactly q^{n-d} solutions in $\mathbb{F}_q^n$.

3.1.4 Upper Bounds on the Number of Solutions

We now give upper bounds on the number of solutions of a polynomial equation. The first is the celebrated Schwartz–Zippel lemma, which bounds the probability of a random point being a root in terms of total degree.

Lemma 3.1.8 Schwartz-Zippel Lemma

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ with $\deg(f) = d$ be a non-zero polynomial. If $S \subseteq \mathbb{F}_q$ be any subset then

$$\mathop{\mathbb{P}}_{\alpha_1, \ldots, \alpha_n \in S}[f(\alpha_1, \ldots, \alpha_n) = 0] \leq \frac{d}{|S|}$$

This lemma immediately gives an upper bound of $d \cdot q^{n-1}$ solutions for the equations $f(X_1, \ldots, X_n) = 0$ in $\mathbb{F}_q^n$. Hence

Corollary 3.1.9 Total Degree Solution Bound

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ with $\deg(f) = d$ be a non-zero polynomial. Then the equation $f(X_1, \ldots, X_n) = 0$ has at most $d \cdot q^{n-1}$ solutions in $\mathbb{F}_q^n$.

We can give another probabilistic upper bound when we control the degree of f in each variable separately rather than just the total degree.

Theorem 3.1.10 Individual Degree Bound

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ be a non-zero polynomial with $\deg_{X_i}(f) \leq d$ for all $i \in [n]$. Then for $(\alpha_1, \ldots, \alpha_n)$ chosen uniformly at random from $\mathbb{F}_q^n$,

$$\mathop{\mathbb{P}}_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q}[f(\alpha_1, \ldots, \alpha_n) = 0] \leq 1 - \left(1 - \frac{d}{q}\right)^n$$

Proof: We proceed by induction on n. For the base case $n = 1$, since $f \in \mathbb{F}_q[X_1]$ is non-zero with $\deg_{X_1}(f) \le d$, it has at most d roots in $\mathbb{F}_q$. Hence $\mathbb{P}_{x_1 \in \mathbb{F}_q}[f(x_1) = 0] \le d/q = 1 - (1 - d/q)^1$.

Assume the bound holds for $n-1$ variables. Let $t = \deg_{X_n}(f)$. Then write f as a polynomial in X_n with coefficients in $\mathbb{F}_q[X_1, \ldots, X_{n-1}]$:

$$f(X_1, \ldots, X_n) = \sum_{i=0}^{t} f_i(X_1, \ldots, X_{n-1}) \cdot X_n^k.$$

Since $\deg_{X_j}(f) \le d$ for each $j \in [n-1]$, every f_i satisfies $\deg_{X_j}(f_i) \le d$.

Let $\alpha_1, \ldots, \alpha_{n-1}$ be uniformly random points on $\mathbb{F}_q$ and α_n be uniform on $\mathbb{F}_q$, independently. Then by induction hypothesis

$$\mathbb{P}_{\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_q}[f_t(\alpha_1, \ldots, \alpha_{n-1}) = 0] \le 1 - \left(1 - \frac{d}{q}\right)^{n-1}$$

Now we have

$$\begin{aligned}
\mathbb{P}_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q}[f(\alpha_1, \ldots, \alpha_n) = 0] &\le \mathbb{P}_{\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_q}[f_t(\alpha_1, \ldots, \alpha_{n-1}) = 0] + \mathbb{P}_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q}[f(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n) = 0 \wedge f_t(\alpha_1, \ldots, \alpha_{n-1}) \neq 0] \\
&\le \mathbb{P}_{\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_q}[f_t(\alpha_1, \ldots, \alpha_{n-1}) = 0] + \\
&\qquad \mathbb{P}_{\alpha_n \in \mathbb{F}_q}[f(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n) = 0 \mid f_t(\alpha_1, \ldots, \alpha_{n-1}) \neq 0] \mathbb{P}_{\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_q}[f_t(\alpha_1, \ldots, \alpha_{n-1}) \neq 0]
\end{aligned}$$

Let $p := \mathbb{P}_{\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_q}[f_t(\alpha_1, \ldots, \alpha_{n-1}) = 0]$. For the second term after fixing $\alpha_1, \ldots, \alpha_{n-1}$ we have $\mathbb{P}_{\alpha_n \in \mathbb{F}_q}[f(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n) = 0 \mid f_t(\alpha_1, \ldots, \alpha_{n-1}) \neq 0] \le \frac{d}{q}$. Now

$$\mathbb{P}_{\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_q}[f_t(\alpha_1, \ldots, \alpha_{n-1}) \neq 0] = 1 - \mathbb{P}_{\alpha_1, \ldots, \alpha_{n-1} \in \mathbb{F}_q}[f_t(\alpha_1, \ldots, \alpha_{n-1}) = 0]$$

Therefore together we get

$$\mathbb{P}_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q}[f(\alpha_1, \ldots, \alpha_n) = 0] \le p + (1-p)\frac{d}{q} = \frac{d}{q} + \left(1 - \frac{d}{q}\right)p$$

By induction hypothesis $p \le 1 - \left(1 - \frac{d}{q}\right)^{n-1}$. Therefore we have

$$\mathbb{P}_{\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q}[f(\alpha_1, \ldots, \alpha_n) = 0] \le \frac{d}{q} + \left(1 - \frac{d}{q}\right)\left[1 - \left(1 - \frac{d}{q}\right)^{n-1}\right] = 1 - \left(1 - \frac{d}{q}\right)^n$$

So we have the lemma. ∎

We can give another upper bound based on individual degrees of each variables. If degree of each variable is at most d then there can be at most $d(q^{n-1} - 1)$ many non-trivial solutions. For this we consider f to be *homogeneous* i.e. total degree of each monomial is same.

Theorem 3.1.11 Nontrivial Solution Bound for Homogeneous Polynomials

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ be a homogeneous with $\deg(f) = d \ge 1$. Then the equation $f(X_1, \ldots, X_n) = 0$ has at most $d(q^{n-1} - 1)$ non-trivial solutions.

Proof: We use double induction on n and d. For the base case $n = 1$ then $f = aX_1^d$ for some $a \neq 0$, so the only root is $X_1 = 0$, giving $\overline{Z}(f) = 0 \le d(q^0 - 1) = 0$. If $d = 1$ then $f = a_1X_1 + \cdots + a_nX_n$ is a non-zero linear form whose zero set is an $(n-1)$-dimensional hyperplane, so $\overline{Z}(f) = q^{n-1} - 1 = d(q^{n-1} - 1)$.

Suppose $n > 1$, $d > 1$, and the bound holds for non-constant homogeneous polynomials in at most n variables of degree less than d and for those in fewer than n variables of degree at most d. We consider two cases.

Case 1: $X_1 \mid f$. Write $f = X_1 \cdot g$ where $g \in \mathbb{F}_q[X_1, \ldots, X_n]$ is homogeneous of degree $d-1$. Every nontrivial solution $(\alpha_1, \ldots, \alpha_n) \in \overline{Z}(f)$ with $\alpha_i \in \mathbb{F}_q$ for all $i \in [n]$, satisfies $\alpha_1 = 0$ or $g(\alpha_1, \ldots, \alpha_n) = 0$. Solutions with $\alpha_1 = 0$: any $(0, \alpha_2, \ldots, \alpha_n) \neq 0$ works, giving $q^{n-1} - 1$ solutions. Solutions with $\alpha_1 \neq 0$: these are nontrivial zeros of g, and by the induction hypothesis on d there are at most $(d-1)(q^{n-1}-1)$. Hence

$$\overline{Z}(f) \leq (q^{n-1}-1) + (d-1)(q^{n-1}-1) = d(q^{n-1}-1).$$

Case 2: $X_1 \nmid f$. Write $f = \sum_{k=0}^{d} f_k(X_2, \ldots, X_n) \cdot X_1^k$ with f_k homogeneous of degree $d-k$. Since $X_1 \nmid f$ we have $f_0 \neq 0$, so for every $\alpha \in \mathbb{F}_q$ such that $f(\alpha, X_2, \ldots, X_n)$ is a non-zero polynomial of degree d in $n-1$ variables we have by Schwartz–Zippel Lemma it has at most $d \cdot q^{n-2}$ zeros in $\mathbb{F}_q^{n-1}$. Now summing over $\alpha \in \mathbb{F}_q^*$ gives at most $(q-1) \cdot d \cdot q^{n-2}$ nontrivial solutions with $\alpha_1 \neq 0$. For solutions with $\alpha_1 = 0$: $f(0, X_2, \ldots, X_n) = f_0(X_2, \ldots, X_n)$ is a nonzero homogeneous polynomial in $n-1$ variables of degree d, so by the induction hypothesis on n it has at most $d(q^{n-2}-1)$ nontrivial zeros. Altogether, $\overline{Z}(f) \leq (q-1)d \cdot q^{n-2} + d(q^{n-2}-1) = d\left(q^{n-1} - q^{n-2} + q^{n-2} - 1\right) = d(q^{n-1}-1)$. ∎

§ 3.2 Character Sums with Polynomials

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and let $f \in \mathbb{F}_q[X]$ be a non-constant polynomial. In this section we will look sums of the form

$$\sum_{\alpha \in \mathbb{F}_q} \chi(f(\alpha))$$

Such sums are called *Weil Sums.*

Definition 3.2.1: Weil Sum

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and let $f \in \mathbb{F}_q[X]$ be a non-constant polynomial. The Weil sum *associated to χ and f is*

$$W(\chi; f) \coloneqq \sum_{\alpha \in \mathbb{F}_q} \chi(f(\alpha)).$$

More generally, for a multiplicative character $\psi \in \boldsymbol{\mathcal{M}}_q$ and a additive character $\chi \in \boldsymbol{\chi}_q$ and a polynomials $f, g \in \mathbb{F}_q(X)$, the Weil sum *is*

$$W(\psi; f) \coloneqq \sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))$$

$$W(\psi, \chi; f, g) \coloneqq \sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))\, \chi(g(\alpha)).$$

Weil sums are extremely difficult to evaluate. One usually has to be satisfied with estimates for the absolute value of the sum. In certain cases it is easier to find the character sums exactly.

3.2.1 Additive Character Sums

The first family of Weil sums we evaluate exactly are those of the form $\chi(ac^n + b)$, where the polynomial is a shifted monomial. The key is to factor the sum over c using multiplicative characters of order $d = \gcd(n, q-1)$.

Theorem 3.2.1 Weil Sum of a Shifted Monomial

Let $\chi \in \boldsymbol{\mathcal{X}}_q$ be a non-trivial additive character of $\mathbb{F}_q$, $n \in \mathbb{N}$, and let $\lambda \in \boldsymbol{\mathcal{M}}_q$ be a multiplicative character of order $d = \gcd(n, q-1)$. Then for any $a, b \in \mathbb{F}_q$ with $a \neq 0$,

$$\sum_{\gamma \in \mathbb{F}_q} \chi(a \cdot \gamma^n + b) = \chi(b) \sum_{j=1}^{d-1} \overline{\lambda}^j(a)\, G(\lambda^j, \chi).$$

Proof: Let $\tau \in \boldsymbol{\mathcal{X}}_q$ be a non-trivial additive character of $\mathbb{F}_q$ defined by $\tau(\gamma) = \chi(a \cdot \gamma)$ for all $\gamma \in \mathbb{F}_q$. Then we get

$$\sum_{\gamma \in \mathbb{F}_q} \chi(a \cdot \gamma^n + b) = \chi(b) \sum_{\gamma \in \mathbb{F}_q} \chi(a \cdot \gamma^n) = \chi(b) \sum_{\gamma \in \mathbb{F}_q} \tau(\gamma^n)$$

Now by Observation 2.9(ii) we get $\tau(\gamma^n) = \frac{1}{q-1} \sum_{\psi \in \boldsymbol{\mathcal{M}}_q} G(\overline{\psi}, \tau) \cdot \psi(\gamma^n)$ for all $\gamma \in \mathbb{F}_q$. Therefore we get

$$\begin{aligned}\sum_{\gamma \in \mathbb{F}_q} \tau(\gamma^n) &= \tau(0) + \sum_{\gamma \in \mathbb{F}_q^*} \frac{1}{q-1} \sum_{\psi \in \boldsymbol{\mathcal{M}}_q} G(\overline{\psi}, \tau) \cdot \psi(\gamma^n) \\ &= \tau(0) + \frac{1}{q-1} \sum_{\psi \in \boldsymbol{\mathcal{M}}_q} G(\overline{\psi}, \tau) \sum_{\gamma \in \mathbb{F}_q^*} \psi^n(\gamma)\end{aligned}$$

Now if ψ^n is trivial then $\sum_{\gamma \in \mathbb{F}_q^*} \psi^n(\gamma) = q - 1$ and if ψ^n is non-trivial then $\sum_{\psi \in \boldsymbol{\mathcal{M}}_q} \psi^n(\gamma) = 0$ by Theorem 2.2.5(i). Now ψ^n is trivial if and only if $\mathrm{ord}(\psi) \mid d$. Since $\mathrm{ord}(\overline{lm}) = d$, the characters ψ with order dividing d are exactly λ^j for all $0 \leq j \leq d-1$. Therefore

$$\sum_{\gamma \in \mathbb{F}_q} \tau(\gamma^n) = 1 + \sum_{j=0}^{d-1} G(\lambda^j, \tau) = \sum_{j=1}^{d-1} G(\lambda^j, \tau) = \sum_{j=1}^{d-1} \overline{\lambda}^j(a) \cdot G(\lambda^j, \chi)$$

Hence we have the theorem. ∎

From Theorem 2.3.1 we have the absolute value of $G(\psi, \chi)$ to be $q^{1/2}$ if χ, ψ are non-trivial additive and multiplicative characters respectively. So we get the absolute value of the above character sum to be at most $(d-1)q^{1/2}$. Therefore as an immediate consequence we get the following bound on the absolute value of such sums.

Theorem 3.2.2 Bound on Weil Sum of a Shifted Monomial

Let $\chi \in \boldsymbol{\mathcal{X}}_q$ be a non-trivial additive character of $\mathbb{F}_q$, $n \in \mathbb{N}$, and $d = \gcd(n, q-1)$. Then for any $a, b \in \mathbb{F}_q$ with $a \neq 0$,

$$\left| \sum_{c \in \mathbb{F}_q} \chi(ac^n + b) \right| \leq (d-1)\sqrt{q}.$$

So if $\gcd(n, q-1) = 1$ then the sum at right hand side is over no elements. So the sum evaluates to be zero. So we have the following corollary:

Corollary 3.2.3 Vanishing for Coprime Exponent

Let $\chi \in \boldsymbol{\mathcal{X}}_q$ be a non-trivial additive character of $\mathbb{F}_q$ and $n \in \mathbb{N}$. If $\gcd(n, q-1) = 1$ then for any $a, b \in \mathbb{F}_q$ with $a \neq 0$,

$$\sum_{c \in \mathbb{F}_q} \chi(ac^n + b) = 0.$$

For quadratic polynomials in odd characteristic the Weil sum can be evaluated exactly in terms of a Gaussian sum and the quadratic character.

Theorem 3.2.4 Weil Sum of a Quadratic Polynomial

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character of $\mathbb{F}_q$ with q odd, and let $f(X) = a_2X^2 + a_1X + a_0 \in \mathbb{F}_q[X]$ with $a_2 \neq 0$. Then

$$\sum_{\gamma \in \mathbb{F}_q} \chi(f(\gamma)) = \chi\left(a_0 - a_1^2(4a_2)^{-1}\right)\, \eta(a_2)\, G(\eta, \chi),$$

where $\eta \in \boldsymbol{\mathcal{M}}_q$ is the quadratic character of $\mathbb{F}_q$.

Proof: Since f is a quadratic polynomial we can express f as in the form of Theorem 3.2.1 for degree 2.

$$f(X) = a_2X^2 + a_1X + a_0 = a_2\left(X + a_1 \cdot (2a_2)^{-1}\right)^2 + \left(a_0 - a_1^2(4a_2)^{-1}\right)$$

Let $\overline{\gamma} = \gamma + a_1(2a_2)^{-1}$ and $\beta = a_0 - a_1^2(4a_2)^{-1}$. Then we have

$$\sum_{\gamma \in \mathbb{F}_q} \chi(f(\gamma)) = \sum_{\overline{\gamma} \in \mathbb{F}_q} \chi(a_2 \cdot \overline{\gamma}^2 + \beta)$$

So by Theorem 3.2.1

$$\sum_{\gamma \in \mathbb{F}_q} \chi(f(\gamma)) = \chi(\beta) \cdot \eta(a_2) \cdot G(\eta, \chi)$$

Hence we have the result. ∎

For polynomials whose monomials are all *affine p-polynomials* i.e. powers of $p = \mathrm{char}(\mathbb{F}_q)$, the sum collapses to either 0 or a full power of q depending on a linear condition on the parameter b.

Theorem 3.2.5 Weil Sum of an Affine p-Polynomial

Let $p = \mathrm{char}(\mathbb{F}_q)$ and let

$$f(X) = a_rX^{p^r} + a_{r-1}X^{p^{r-1}} + \cdots + a_1X^p + a_0X + a \in \mathbb{F}_q[X]$$

be an affine p-polynomial over $\mathbb{F}_q$. Let $\chi_b \in \boldsymbol{\chi}_q$, $b \in \mathbb{F}_q^*$, be a non-trivial additive character as defined in Theorem 2.2.1. Then

$$\sum_{\gamma \in \mathbb{F}_q} \chi_b(f(\gamma)) = \begin{cases} \chi_b(a) \cdot q & \text{if } ba_r + b^pa_{r-1} + \cdots + b^{p^{r-1}}a_1^{p^{r-1}} + b^{p^r}a_0^{p^r} = 0, \\ 0 & \text{otherwise} \end{cases}$$

Proof: We can bring out the $\chi_b(a)$ factor from $\chi_b(f(\gamma))$ as a is the constant term. Consider the following polynomial $F(X) = b \cdot f(X)$. So we have

$$\sum_{\gamma \in \mathbb{F}_q} \chi_b(f(\gamma)) = \chi_b(a) \sum_{\gamma \in \mathbb{F}_q} \chi_1(F(\gamma))$$

where χ_1 is the canonical additive character of $\mathbb{F}_q$. Let τ be the additive character defined as $\tau(\gamma) = \chi_1(F(\gamma))$ for all $\gamma \in \mathbb{F}_q$. Then by Theorem 2.2.2(i) we have

$$\sum_{\gamma} \tau(\gamma) = \begin{cases} q & \text{if } \tau \text{ is trivial} \\ 0 & \text{otherwise} \end{cases}$$

So now it suffices to characterize those affine p-polynomials $F(X)$ for which τ is trivial. Let $q = p^s$ and let Tr is the absolute trace function from $\mathbb{F}_q$ to $\mathbb{F}_p$. Then τ is trivial if and only if $\mathrm{Tr}(F(\gamma)) = 0$ for all $\gamma \in \mathbb{F}_q$. Therefore we get

$$\tau \text{ is trivial} \iff \sum_{j=0}^{s-1} F^{p^j}(X) \equiv 0 \bmod \langle X^q - X \rangle$$

Now we have

$$\sum_{j=0}^{s-1} F^{p^j}(X) = \sum_{j=0}^{s-1} \left(\sum_{i=0}^{r} b \cdot a_i \cdot X^{p^i} \right)^{p^j} = \sum_{j=0}^{s-1} \sum_{i=0}^{r} b^{p^j} \cdot a_i^{p^j} \cdot X^{p^{i+j}} \equiv \sum_{k=0}^{s-1} \left(\sum_{i=0}^{r} b^{p^{k-i}} a_i^{p^{k-i}} \right) X^{p^k} \bmod \langle X^q - X \rangle$$

Therefore we get

$$\sum_{j=0}^{s-1} F^{p^j}(X) \equiv 0 \bmod \langle X^q - X \rangle \iff \sum_{i=0}^{r} b^{p^{k-i}} a_i^{p^{k-i}} = 0, \ \forall\, 0 \le k \le s-1$$

Now for all $j \in \{0, \ldots, s-1\}$ we have

$$\sum_{i=0}^{r} b^{p^{k-i}} a_i^{p^{k-i}} = 0 \iff \left(\sum_{i=0}^{r} b^{p^{k-i}} a_i^{p^{k-i}} \right)^{p^{r-k}} = 0 \iff \sum_{i=0}^{r} b^{p^{r-i}} a_i^{p^{r-i}} = 0$$

So we get the Theorem. ∎

The above theorem covers both odd and even characteristic, whereas Theorem 3.2.4 was restricted to odd q. Specializing to the case where f is a quadratic polynomial $f(X) = a_2X^2 + a_1X + a_0$ and applying the p-polynomial criterion, we can fill this gap and obtain the analogue of Theorem 3.2.4 for even characteristic:

Corollary 3.2.6 Weil Sum of Quadratic Polynomial in Even Characteristic

Let $f(X) = a_2X^2 + a_1X + a_0 \in \mathbb{F}_q[X]$ with q even, and let $\chi_b \in \boldsymbol{\chi}_q$, $b \in \mathbb{F}_q^*$, be as defined in Theorem 3.2.5. Then

$$\sum_{c\in\mathbb{F}_q} \chi_b(f(c)) = \begin{cases} \chi_b(a_0)\cdot q & \text{if } a_2 = ba_1^2, \\ 0 & \text{otherwise.} \end{cases}$$

3.2.2 Quadratic Character Sums

When the character is the quadratic character η, the sums $\sum_c \eta(f(c))$ can be evaluated by completing the square (for quadratic f) or by the continued fraction algorithm (for polynomials with no roots in $\mathbb{F}_q$).

Theorem 3.2.7 Quadratic Character Sum of a Quadratic Polynomial

Let $f(X) = a_2X^2 + a_1X + a_0 \in \mathbb{F}_q[X]$ with q odd and $a_2 \neq 0$, let $\Delta = a_1^2 - 4a_0a_2$ denote the discriminant of f, and let η be the quadratic character of $\mathbb{F}_q$ (extended by $\eta(0) = 0$). Then

$$\sum_{\gamma\in\mathbb{F}_q} \eta(f(\gamma)) = \begin{cases} -\eta(a_2) & \text{if } \Delta \neq 0, \\ (q-1)\cdot\eta(a_2) & \text{if } \Delta = 0. \end{cases}$$

Proof: Since $4a_2^2 = (2a_2)^2$ we have $\eta(4a_2) = 1$. So we multiply the sum by $\eta(4a_2)^2$ and the sum stays the same. So we get for any $\gamma \in \mathbb{F}_q$

$$\eta(f(\gamma))\eta(4a_2^2\cdot f(\gamma)) = \eta(a_2)\cdot\eta(4a_2^2\cdot\gamma^2 + 4a_1a_2\cdot\gamma + 4a_2a_1) = \eta(a_2)\cdot\eta\left((2a_2\cdot\gamma + a_1)^2 - \Delta\right)$$

This gives $\sum_{\gamma\in\mathbb{F}_q}\eta(f(\gamma)) = \eta(a_2)\sum_{\overline{\gamma}\in\mathbb{F}_q}\eta(\overline{\gamma}^2 - \Delta)$ where for any $\gamma \in \mathbb{F}_q$ define $\overline{\gamma} = 2a_2\cdot\gamma + a_1$. Now if $\Delta = 0$ then $\eta(\overline{\gamma}^2) = 1$ and hence $\sum_{\gamma\in\mathbb{F}_q}\eta(f(\gamma)) = -\eta(a_2)$.

So suppose $\Delta \neq 0$. Let for any $\beta \in \mathbb{F}_q$, $Z(\beta)$ is the number of solutions of $\overline{\gamma} \in \mathbb{F}_q$ such that $\overline{\gamma}^2 = \beta^2 - \Delta$. Then $\sum_{\beta\in\mathbb{F}_q} Z(\beta)$ is the total number of ordered pairs $(\overline{\gamma}, \beta) \in \mathbb{F}_q^2$ such that $\overline{\gamma}^2 - \beta^2 = \Delta$. Now

$$\overline{\gamma}^2 - \beta^2 = \Delta \iff (\overline{\gamma} + \beta)(\overline{\gamma} - \beta) = \Delta$$

Since $\Delta \neq 0$, both $\overline{\gamma} + \beta$ and $\overline{\gamma} - \beta$ are nonzero. Take $u = \overline{\gamma} + \beta$ and $v = \overline{\gamma} - \beta$. Then there is a one-one correspondence for each $(\overline{\gamma}, \beta) \in \mathbb{F}_q^2$ such that $\overline{\gamma}^2 - \beta^2 = \Delta$ and each $(u, v) \in \mathbb{F}_q^2$ such that $u\cdot v = \Delta$. For any $u \in \mathbb{F}_q^*$, we automatically get the value v hence number of such $(u, v) \in \mathbb{F}_q^2$ such that $u\cdot v = \Delta$ is $q-1$. Therefore $\sum_{\beta\in\mathbb{F}_q} Z(\beta) = q-1$. So we have the result. ■

3.2.3 Weil Sums of Continued Fractions

There is a remarkable connection between Weil sums for the quadratic character η and the *continued fraction algorithm* for rational functions over $\mathbb{F}_q$. We collect the necessary algebraic facts about continued fractions of rational functions in Appendix C; here we set up the specific expansion needed for the application to character sums and state the main results. Recall from Appendix C that every rational function $r_0/r_1 \in \mathbb{F}_q(X)$ with $r_1 \neq 0$ admits a unique *continued fraction expansion*

$$\frac{r_0}{r_1} = A_0 + \cfrac{1}{A_1 + \cfrac{1}{A_2 + \cdots}} =: [A_0, A_1, \ldots, A_s],$$

where $A_0, A_1, \ldots, A_s \in \mathbb{F}_q[X]$ are the *partial quotients* produced by the Euclidean algorithm applied to r_0 and r_1, with $A_1, \ldots, A_s$ all of positive degree. Associated to this expansion are the *convergents* $P_i/Q_i := [A_0, A_1, \ldots, A_i]$, where the numerator and denominator sequences are defined by the recurrences

$$P_{-1} = 1, \quad P_0 = A_0, \quad P_i = A_i P_{i-1} + P_{i-2}, \qquad Q_{-1} = 0, \quad Q_0 = 1, \quad Q_i = A_i Q_{i-1} + Q_{i-2}.$$

Each P_i/Q_i is a best rational approximation to r_0/r_1, and the denominators Q_i are pairwise coprime with strictly increasing degrees.

For q odd, put $G(X) = X^q - X$, let $f \in \mathbb{F}_q[X]$ be a polynomial of positive degree with no roots in $\mathbb{F}_q$, and set $F(X) = f(X)^{(q-1)/2}$. Consider the continued fraction expansions

$$\frac{F(X)-1}{G(X)} = [A_0, A_1, \ldots, A_s] \qquad \text{and} \qquad \frac{F(X)+1}{G(X)} = [a_0, a_1, \ldots, a_t]$$

with P_i, Q_i and p_i, q_i the corresponding convergent sequences. It is clear that $A_0 = a_0$. Define n_f to be the largest integer m such that $A_i = a_i$ for $i = 0, 1, \ldots, m$.

Lemma 3.2.8

For q odd, put $G(X) = X^q - X$, let $f \in \mathbb{F}_q[X]$ be a polynomial of positive degree with no roots in $\mathbb{F}_q$, and set $F(X) = f(X)^{(q-1)/2}$. Let

$$\frac{F(X)-1}{G(X)} = [A_0, A_1, \ldots, A_s] \qquad \textit{and} \qquad \frac{F(X)+1}{G(X)} = [a_0, a_1, \ldots, a_t]$$

Then either $n_f = s = t-1$ or $n_f = t = s-1$.

Proof: Let P_i, Q_i are defined as above from $[A_0, \ldots, A_s]$ and p_i, q_i are defined from $[a_0, \ldots, a_t]$. Then using Corollary C.2 and Corollary C.5 we have

$$\frac{P_s}{Q_s} = \frac{F(X)-1}{G(X)} \qquad \text{and} \qquad \frac{p_t}{q_t} = \frac{F(X)+1}{G(X)}$$

So we get

$$Q_s(X) = \frac{\alpha_1 \cdot G(X)}{\gcd(F(X)-1, G(X))}, \quad q_t(X) = \frac{\alpha_2 \cdot G(X)}{\gcd(F(X)+1, G(X))} \quad \text{for some } \alpha_1, \alpha_2 \in \mathbb{F}_q^* \tag{3.1}$$

Since f has no roots in $\mathbb{F}_q$, for all $\alpha \in \mathbb{F}_q$, $f^{q-1}(\alpha) = 1$. So $G(X) \mid f^{q-1}(X) - 1$ i.e. $G(X) \mid (F(X)-1)(F(X)+1)$. So

$$\gcd(G(X), F(X)-1) \cdot \gcd(G(X), F(X)+1) = G(X)$$

Therefore $Q_s(X) \cdot q_t(X) = \alpha \cdot G(X)$ with $\alpha \in \mathbb{F}_q^*$. Let $n = n_f < s$ and $n < t$. Therefore by Lemma C.3 we get

$$\frac{F(X)-1}{G(X)} = \frac{P_{n+1} + \beta_{n+1} \cdot P_n}{Q_{n+1} + \beta_{n+1} \cdot Q_n}, \quad \text{and} \quad \frac{F(X)+1}{G(X)} = \frac{p_{n+1} + \gamma_{n+1} \cdot p_n}{q_{n+1} + \gamma_{n+1} \cdot q_n}$$

with $\beta_{n+1}, \gamma_{n+1}$ being rational functions with negative degree. Therefore

$$\frac{2}{G(X)} = \frac{F(X)+1}{G(X)} - \frac{F(X)-1}{G(X)} = \frac{H(X)}{(Q_{n+1} + \beta_{n+1} \cdot Q_n)(q_{n+1} + \gamma_{n+1} \cdot q_n)} \tag{3.2}$$

where

$$H = \underbrace{p_{n+1} \cdot Q_{n+1} - P_{n+1} \cdot q_{n+1}}_{H_0} + \beta_{n+1} \underbrace{(p_{n+1} \cdot Q_n - P_n \cdot q_{n+1})}_{H_1} + \gamma_{n+1} \underbrace{(p_n \cdot Q_{n+1} - P_{n+1} \cdot q_n)}_{H_2} + \beta_{n+1} \cdot \gamma_{n+1} \underbrace{(p_n \cdot Q_n - P_n \cdot q_n)}_{H_3} \tag{3.3}$$

Now by definition of n we have $P_i = p_i$ and $Q_i = q_i$ for all $-1 \le i \le n$. By Lemma C.4 we have $H_1 = (-1)^n$ and $H_2 = (-1)^{n+1}$ and $H_3 = 0$. So now we have to calculate H_0.

$$\begin{aligned} H_0 &= p_{n+1} \cdot Q_{n+1} - P_{n+1} \cdot q_{n+1} \\ &= (A_{n+1} \cdot p_n + p_{n-1})(A_{n+1} \cdot Q_n + Q_{n-1}) - (A_{n+1} \cdot P_n + P_{n-1})(a_{n+1} \cdot q_n + q_{n-1}) \\ &= a_{n+1}(p_n \cdot Q_{n-1} - P_n \cdot q_{n-1}) + A_{n+1}(p_{n-1} \cdot Q_n - q_{n-1} \cdot P_n) \\ &= (-1)^{n-1} a_n - (-1)^{n-1} A_{n+1} \\ &= (-1)^n (A_{n+1} - a_{n+1}) \end{aligned}$$

Therefore $H = (-1)^n(A_{n+1} - a_{n+1} + \beta_{n+1} - \gamma_{n+1})$. Now by definition of n, $A_{n+1} \neq a_{n+1}$. So $\deg(H) \ge 0$. Therefore from the equation (3.2) we get $G(X) \cdot H(X) = 2(Q_{n+1} + \beta_{n+1} \cdot Q_n)(q_{n+1} + \gamma_{n+1} \cdot q_n)$. By the definition of G we have $\deg(G) + \deg(H) \ge q$. On the other hand, $\deg(G) + \deg(H) = \deg(Q_{n+1}) + \deg(q_{n+1}) \le \deg(Q_s) + \deg(q_t) \le q$. Therefore we have $\deg(H) = 0$. So $\deg(A_{n+1}) = \deg(a_{n+1})$. And hence $\deg(Q_{n+1}) = \deg(q_{n+1}) = q/2$. But we have q to be odd. Hence contradiction ↯ So we have either $n = s$ or $n = t$. Suppose $n = s$. Then $t > s$ and therefore we can write

$$\frac{F(X) + 1}{G(X)} = \frac{p_{s+1} + \gamma_{s+1} \cdot p_s}{q_{s+1} + \gamma_{s+1} \cdot q_s}$$

Therefore we get new equality for (3.2)

$$\begin{aligned} \frac{2}{G(X)} &= \frac{F(X) + 1}{G(X)} - \frac{F(X) - 1}{G(X)} \\ &= \frac{p_{s+1} + \gamma_{s+1} \cdot p_s}{q_{s+1} + \gamma_{s+1} \cdot q_s} - \frac{P_s}{Q_s} \\ &= \frac{Q_s \cdot p_{s+1} - P_s \cdot q_{s+1} + \gamma_{s+1}(Q_s \cdot p_s - P_s \cdot q_s)}{Q_s(q_{s+1} + \gamma_{s+1} \cdot q_s)} \\ &= \frac{(-1)^s}{Q_s(q_{s+1} + \gamma_{s+1} \cdot q_s)} \qquad \text{[By Lemma C.4]} \end{aligned}$$

So we have $(-1)^s \cdot G = 2 \cdot Q_s(q_{s+1} + \gamma_{s+1} \cdot q_s)$. Again we will compare the degrees. And since $\deg(G) = q$ we have $\deg(Q_s) + \deg_{q_{s+1}} = q$. On the other hand we have $\deg(Q_s) + \deg_{q_{s+1}} \le \deg(Q_s) + \deg(q_t) = q$. Therefore $\deg(q_{s+1}) = \deg(q_t)$. Since degrees of q_i increases strictly we have $t = s + 1$. Similarly, if $n = t$ then we get $t + 1 = s$. ∎

So now we can derive the Weil sum of quadratic character for any polynomial f which has no rooots in $\mathbb{F}_q$ where q is odd.

Theorem 3.2.9 Weil Sum via Continued Fractions

Let $\eta \in \boldsymbol{\mathcal{M}}_q$ be the quadratic character of $\mathbb{F}_q$, q odd, and let $f \in \mathbb{F}_q[X]$ be a polynomial of positive degree with no roots in $\mathbb{F}_q$. Then

$$\sum_{\gamma \in \mathbb{F}_q} \eta(f(\gamma)) = \begin{cases} \deg(a_s) & \text{if } n_f = s, \\ -\deg(A_s) & \text{if } n_f = t, \end{cases}$$

where A_s and a_s are obtained from the continued fraction expansions above.

Proof: We define the following two sets $Z_1 := |\{\gamma \in \mathbb{F}_q : \eta(f(\gamma)) = 1\}|$ and $Z_{-1} := |\{\gamma \in \mathbb{F}_q : \eta(f(\gamma)) = -1\}|$. Then basically we have $\sum_{\gamma \in \mathbb{F}_q} \eta(f(\gamma)) = Z_1 - Z_{-1}$. Since $\eta(f(\gamma)) = 1$ if and only if $F(\gamma) = f^{(q-1)/2}(\gamma) = 1$ we have $Z_1 = \deg(\gcd(F(X) - 1, X^q - X))$ and hence by (3.1) we have $Z_1 = q - \deg(Q_s)$.

Similarly $\eta(f(\gamma)) = -1$ if and only if $F(\gamma) = -1$. Hence $Z_{-1} = \deg(\gcd(F(X) + 1, X^q - X)) = q - \deg(q_t)$. So if $n_f = t = s - 1$ then $q_t = Q_{s-1}$. And therefore $q_t = Q_{s-1}$. So $Z_1 - Z_{-1} = -\deg(Q_s) + \deg(Q_{s-1}) = \deg(A_s)$. In case of $n_f = s = t - 1$ we get $Z_1 - Z_{-1} = \deg(q_t) - \deg(q_{t-1}) = \deg(a_t)$. So the result follows. ∎

Note:-

The hypothesis that f has no roots in $\mathbb{F}_q$ is essential. For example, take $f(X) = X$: then $\sum_{\gamma\in\mathbb{F}_q} \eta(f(\gamma)) = 0$ by Theorem 2.2.5(i), yet $\frac{F(X)-1}{X^q-X} = [0, X^{(q+1)/2} + X]$ and $\frac{F(X)+1}{X^q-X} = [0, X^{(q+1)/2} - X]$, giving $s = t = 1$ and $\deg(A_s) = \deg(a_t) = {}^{q+1}/_2 \neq 0$. The continued fraction formula does not reproduce the correct value in this case.

In the next section we turn to a statistical perspective and study the average number of solutions of polynomial equations over $\mathbb{F}_q$. We will show that the average value of $Z(f)$ over all polynomials f of bounded degree is exactly q^{n-1}, and then compute the variance of the solution count around this mean.

§ 3.3 Average Number of Solutions

We now investigate the *average number of solutions* of a polynomial equations. Let $d \in \mathbb{N}$. Then define $\Omega_d := \mathbb{F}_q^{\leq d}[X_1, \ldots, X_n]$ and $\omega(d)$ be the set of $n-$tuples $(i_1, \ldots, i_n) \in \mathbb{Z}_0$ such that $i_1 + \cdots + i_n \leq d$. Then

Observation 3.2. *With the definitions of Ω_d and $\omega(d)$ as above we have $|\Omega_d| = q^{|\omega(d)|}$.*

Lemma 3.3.1 Average of Solution Count

With the notations of Ω_d and $\omega(d)$ as defined above we have

$$\frac{1}{|\Omega_d|}\sum_{f\in\Omega_d} Z(f) = q^{n-1}$$

Proof: We have

$$\sum_{f\in\Omega_d} Z(f) = \sum_{f\in\Omega_d} \sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q \\ f(\alpha_1,\ldots,\alpha_n)=0}} 1 = \sum_{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q} \sum_{\substack{f\in\Omega_d \\ f(\alpha_1,\ldots,\alpha_n)=0}} 1$$

Let for any $f \in \Omega_d$

$$f(X_1, \ldots, X_n) = \sum_{(i_1,\ldots,i_n)\in\omega(d)} f_{i_1,\ldots,i_n} X_1^{i_1} \cdots X_n^{i_n} \quad f_{i_1,\ldots,i_n} \in \mathbb{F}_q,\ \forall\ (i_1, \ldots, i_n) \in \omega(d)$$

Then for a fixed $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$ we get to choose the coefficients $f_{i_1,\ldots,i_n}$ for all $(i_1, \ldots, i_n) \in \omega(d) \setminus \{(0, \ldots, 0)\}$ arbitrarily and then determining the constant term $f_{0,\ldots,o}$ to make $f(\alpha_1, \ldots, \alpha_n) = 0$. Thus the number of $f \in \omega_d$ with $f(\alpha_1, \ldots, \alpha_n) = 0$ is equal to $q^{|\omega(d)|-1}$. Therefore $\sum_{f\in\Omega_d} Z(f) = q^n \cdot q^{|\omega(d)|-1} = |\Omega_d| \cdot q^{n-1}$. So we have the lemma. ∎

The proof of the above lemma gives us the following observation:

Observation 3.3. *For any $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_q$, the number of polynomials $f \in \Omega_d$ such that $f(\alpha_1, \ldots, \alpha_n) = 0$ is $q^{|\omega(d)|-1}$.*

The above lemma implies that a polynomial equation in n variables has on average q^{n-1} solutions in $\mathbb{F}_q^n$. So we now consider average deviation from the average number of solutions, q^{n-1}.

Theorem 3.3.2 Variance of Solution Count

With the notations of Ω_d and $\omega(d)$ as defined above we have

$$\frac{1}{|\Omega_d|}\sum_{f\in\Omega_d} \left(Z(f) - q^{n-1}\right)^2 = q^{n-1} - q^{n-2}$$

Proof: Let $\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in \mathbb{F}_q^n$ be any two points. Then we get

$$\sum_{f\in\Omega_d} Z(f)^2 = \sum_{f\in\Omega_d} \left(\sum_{\substack{\alpha\in\mathbb{F}_q^n \\ f(\alpha)=0}} 1 \right)^2 = \sum_{f\in\Omega_d} \sum_{\substack{\alpha\in\mathbb{F}_q^n \\ f(\alpha)=0}} \sum_{\substack{\beta\in\mathbb{F}_q^n \\ f(\beta)=0}} 1 = \sum_{\alpha,\beta\in\mathbb{F}_q^n} \sum_{\substack{f\in\Omega_d \\ f(\alpha)=f(\beta)=0}} 1$$

Now if $\alpha = \beta$ then by Observation 3.3 we have $\sum_{f\in\Omega_d,\, f(\alpha)=0} 1 = q^{|\omega(d)|-1}$. Now suppose $\alpha \neq \beta$. This gives two linear equations for the coefficients of f. Therefore there are $q^{|\omega(d)|-2}$ polynomials $f \in \Omega_d$ such that $f(\alpha) = f(\beta) = 0$. Hence we obtain

$$\begin{aligned}
\sum_{f\in\Omega_d} Z(f)^2 = &= \sum_{\alpha\in\mathbb{F}_q^n} |\Omega_d| \cdot q^{-1} + \sum_{\alpha,\beta\in\mathbb{F}_q^n, \alpha\neq\beta} |\Omega_d| \cdot q^{-2} && \text{[By Observation 3.2]} \\
&= q^n \cdot |\Omega_d| \cdot q^{-1} + q^n(q^n - 1)|\Omega_d| \cdot q^{-2} \\
&= |\Omega_d| \left(q^{n-1} + q^{2n-2} - q^{n-2} \right)
\end{aligned}$$

So now we get

$$\begin{aligned}
\sum_{f\in\Omega_d} (Z(f) - q^{n-1})^2 &= \sum_{f\in\Omega_d} Z(f)^2 + 2 \cdot q^{n-1} \sum_{f\in\Omega_d} Z(f) + q^{2n-2} \sum_{f\in\Omega_d} 1 \\
&= |\Omega_d| \left(q^{2n-2} + q^{n-1} - q^{n-2} \right) - 2 \cdot q^{n-1} \cdot |\Omega_d| \cdot q^{n-1} + q^{2n-2} \cdot |\Omega_d| \\
&= |\Omega_d| (q^{n-1} - q^{n-2})
\end{aligned}$$

So we have the result. ∎

From this result we can expect that $|Z(f) - q^{n-1}|$ is often $O\left(q^{(q-1)/2}\right)$. Later we will see instances of such expected behavior for various polynomial equations. In the next sections we will talk about character sums over polynomial evaluations.

§ 3.4 Quadratic Forms

In this section we will study about quadratic forms and its equivalent forms which are easier to study than a general quadratic form. We will give bounds on the deviation of number of solutions from average number of solutions. Before all of that let's see what is a quadratic form.

Definition 3.4.1: Quadratic Forms

A quadratic form over $\mathbb{F}_q$ is a degree-2 homogeneous *polynomial or the zero polynomial.*

If $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ is a quadratic form. Now if q is odd then for any $i, j \in [n]$ we write $a_{i,j}X_iX_j = \frac{1}{2}a_{i,j}X_iX_j + \frac{1}{2}a_{ij}X_jX_i$ which leads to the representation

$$f(X_1, \ldots, X_n) = \sum_{i,j\in[n]} a_{i,j}X_iX_j \quad \text{for all } a_{i,j} \in \mathbb{F}_q \text{ with } a_{i,j} = a_{j,i}.$$

From this formulation we can represent f by a matrix denoted by C_f where $(i, j)^{th}$ entry of C_f is $a_{i,j}$. This matrix is called the *coefficient matrix*.

Observation 3.4. *Since $a_{i,j} = a_{j,i}$ we have $C_f^\top = C_f$.*

So if X denotes the column vector of all variables then $f = X^\top C_f X$. With this representation now we can talk about equivalence of two quadratic forms.

Definition 3.4.2: Equivalent Quadratic Forms

Let $f, g \in \mathbb{F}_q[X_1, \ldots, X_n]$ be two quadratic forms over $\mathbb{F}_q$. Then f and g are said to be equivalent *if f can be transformed into g by a non-singular linear transformation of the variables.*

So let $f, g \in \mathbb{F}_q[X_1, \ldots, X_n]$ are two equivalent quadratic forms. Then Let X be the variable vector for f and $\tilde{X}$ be the variable vector for g. Then $f = X^\top C_f X$ and $g = \tilde{X}^\top C_g \tilde{X}$. Suppose M is the non-singular linear transformation of variables i.e. $\tilde{X} = MX$. For odd q we have

$$\tilde{X}^\top C_g \tilde{X} = g = (MX)^\top C_f (MX) = X^\top (M^\top C_f M) X,$$

so $C_g = M^\top C_f M$.

Observation 3.5. *If $f, g \in \mathbb{F}_q[X_1, \ldots, X_n]$ (q odd) are two quadratic forms then f and g are equivalent if and only if there exists an invertible matrix $M \in \mathbb{F}_q^{n \times n}$ such that $C_g = M^\top C_f M$ where M is the non-singular linear transformation.*

Since $\tilde{X} = MX$. So M gives the one-one correspondence between $\mathcal{V}(f(X_1, \ldots, X_n) = \alpha)$ and $\mathcal{V}(g(X_1, \ldots, X_n) = \alpha)$ for any $\alpha \in \mathbb{F}_q$.

Observation 3.6. *If $f, g \in \mathbb{F}_q[X_1, \ldots, X_n]$ (q odd) are two quadratic forms then f and g are equivalent by the non-singular linear transformation $M \in \mathbb{F}_q^{n \times n}$. Then $\gamma \mapsto M\gamma$ where $\gamma \in \mathbb{F}_q^n$ gives an one-one correspondence between $\mathcal{V}(f(X_1, \ldots, X_n) = \alpha)$ and $\mathcal{V}(g(X_1, \ldots, X_n) = \alpha)$ for any $\alpha \in \mathbb{F}_q$.*

We will use another terminology to make our life easier. If $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ is a quadratic form then for any $\alpha \in \mathbb{F}_q$ we will say f *represents* α if $f(X_1, \ldots, X_n) = \alpha$ has a solution.

To calculate and understand the of solutions of quadratic forms we define the integer-valued function $\vartheta : \mathbb{F}_q \to \mathbb{Z}$ by $\vartheta(\alpha) = -1$ for all $\alpha \in \mathbb{F}_q^*$ and $\vartheta(0) = q - 1$. Then we have the following properties of ϑ function.

Lemma 3.4.1

For any finite field $\mathbb{F}_q$ we have

$$\sum_{\alpha \in \mathbb{F}_q} \vartheta(\alpha) = 0,$$

and for any $\alpha \in \mathbb{F}_q$ and integers $1 \le k \le m$,

$$\sum_{\substack{\alpha_1, \ldots, \alpha_m \in \mathbb{F}_q \\ \alpha_1 + \cdots + \alpha_m = \alpha}} \vartheta(\alpha_1) \cdots \vartheta(\alpha_k) = \begin{cases} 0 & \text{if } 1 \le k < m \\ \vartheta(\alpha) \cdot q^{m-1} & \text{if } k = m. \end{cases}$$

Proof: Since for all $\alpha \in \mathbb{F}_q^*$, $\vartheta(\alpha) = -1$ we have $\sum_{\alpha \in \mathbb{F}_q^*} \vartheta(\alpha) = -(q-1)$. Therefore $\vartheta(0) + \sum_{\alpha \in \mathbb{F}_q^*} \vartheta(\alpha) = 0$. So we have the first result.

Now suppose $k < m$. Then we have

$$\begin{aligned} \sum_{\substack{\alpha_1, \ldots, \alpha_m \in \mathbb{F}_q \\ \alpha_1 + \cdots + \alpha_m = \alpha}} \prod_{j=1}^{k} \vartheta(\alpha_j) &= \sum_{\alpha_1, \ldots, \alpha_k \in \mathbb{F}_q} \left(\prod_{j=1}^{k} \vartheta(\alpha_j) \right) \sum_{\substack{\alpha_{k+1}, \ldots, \alpha_m \in \mathbb{F}_q \\ \alpha_{k+1} + \cdots + \alpha_m = \alpha - \alpha_1 - \cdots - \alpha_k}} 1 \\ &= q^{m-k-1} \sum_{\alpha_1, \ldots, \alpha_k \in \mathbb{F}_q} \prod_{j=1}^{k} \vartheta(\alpha_j) \\ &= q^{m-k-q} \prod_{j=1}^{k} \sum_{\alpha_j \in \mathbb{F}_q} \vartheta(\alpha_j) = 0 \end{aligned}$$

Now assume $k = m$. Here we will use induction. The case $m = 1$ holds trivially. Suppose the formula is shown for $m \geq 1$. Then

$$\begin{aligned}
\sum_{\substack{\alpha_1,\ldots,\alpha_{m+1}\in\mathbb{F}_q \\ \alpha_1+\cdots+\alpha_{m+1}=\alpha}} \prod_{j=1}^{m+1} \vartheta(\alpha_j) &= \sum_{\substack{\alpha_1,\ldots,\alpha_{m+1}\in\mathbb{F}_q \\ \alpha_1+\cdots+\alpha_{m+1}=\alpha}} \prod_{j=1}^{m+1} \vartheta(\alpha_j) + \sum_{\substack{\alpha_1,\ldots,\alpha_{m+1}\in\mathbb{F}_q \\ \alpha_1+\cdots+\alpha_{m+1}=\alpha}} \prod_{j=1}^{m} \vartheta(\alpha_j) \\
&= \sum_{\substack{\alpha_1,\ldots,\alpha_{m+1}\in\mathbb{F}_q \\ \alpha_1+\cdots+\alpha_{m+1}=\alpha}} \left[\prod_{j=1}^{m} \vartheta(\alpha_j)\right] \cdot [\vartheta(\alpha_{m+1}) + 1] \\
&= \sum_{\alpha_1,\ldots,\alpha_m\in\mathbb{F}_q} \left[\prod_{j=1}^{m} \vartheta(\alpha_j)\right] \cdot \left[\vartheta\left(\alpha - \sum_{j=1}^{m} \alpha_j\right) + 1\right] \\
&= q \sum_{\substack{\alpha_1,\ldots,\alpha_m\in\mathbb{F}_q \\ \alpha_1+\cdots+\alpha_m=\alpha}} \prod_{j=1}^{m} \vartheta(\alpha_j) && \text{[If } \alpha_1+\cdots+\alpha_m \neq \alpha \text{ then the term is 0]} \\
&= q \cdot q^{m-2} \cdot \vartheta(\alpha) && \text{[By Induction Hypothesis]} \\
&= q^{m-1} \cdot \vartheta(\alpha)
\end{aligned}$$

so now we have the second result too. So we have the lemma. ∎

3.4.1 Odd Characteristic Quadratic Forms

We now resume the study of quadratic forms $f \in \mathbb{F}_q[X_1,\ldots,X_n]$ where q is odd. We will first show that every quadratic form $f \in \mathbb{F}_q[X_1,\ldots,X_n]$ where q is odd is equivalent to a *diagonal quadratic form* i.e. a quadratic form of type $a_1X_q^2 + \cdots + a_nX_n^2$ where $a_i \in \mathbb{F}_q$ for all $i \in [n]$.

Lemma 3.4.2 Reduction by Separating a Variable

Let q be odd and $f \in \mathbb{F}_q[X_1,\ldots,X_n]$ be a quadratic form with $n \geq 2$. If f represents $\alpha \in \mathbb{F}_q^*$ then f is equivalent to $\alpha \cdot X_1^2 + g(X_2,\ldots,X_n)$ where g is a quadratic form in $n-1$ variables.

Proof: Since f represents α, $\exists\ (\gamma_1,\ldots,\gamma_n) \in \mathbb{F}_q^n$ such that $f(\gamma_1,\ldots,\gamma_n) = \alpha$. Since $\alpha \neq 0$ not all γ_i are zero. So we can find an invertible matrix $M \in \mathbb{F}_q^{n\times n}$ such that in the first column of M the entries are $\gamma_1,\ldots,\gamma_n$. Therefore if we apply the non-singular linear transformation by M to f, we obtain a quadratic form in $Y_1,\ldots,Y_n$ for which the coefficient of Y_1^2 is $f(\gamma_1,\ldots,\gamma_n) = \alpha$. Thus f is equivalent to a quadratic form of the type

$$\alpha Y_1^2 + 2b_2Y_1Y_2 + \cdots + 2bY_1Y_n + h(Y_2,\ldots,Y_n) = \alpha(Y_1 + b_2\cdot\alpha^{-1}\cdot Y_2 + \cdots b_n\cdot\alpha^{-1}\cdot Y_n)^2 + g(Y_2,\ldots,Y_n)$$

for some $b_2,\ldots,b_n \in \mathbb{F}_q$ where $g, h \in \mathbb{F}_q[X_1,\ldots,X_n]$ are two quadratic forms. So now take the non-singular linear transformation

$$\begin{aligned}
Z_1 &= Y_1 + b_2\cdot\alpha^{-1}\cdot Y_2 + \cdots + b_n\cdot\alpha^{-1}\cdot Y_n \\
Z_i &= Y_i\ \forall\ i \in \{2,\ldots,n\}
\end{aligned}$$

Then we get a quadratic form of the desired type and this is equivalent to f. ∎

Applying Lemma 3.4.2 repeatedly separates variables one by one, eventually reaching a diagonal form. The only subtlety is ensuring at each step that the remaining form represents a non-zero element of $\mathbb{F}_q$; this is handled in the proof of Theorem 3.4.3 below.

Theorem 3.4.3 Diagonal Form of Quadratic Forms over $\mathbb{F}_q$

For any quadratic form $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ where q is odd, there exists $a_1, \ldots, a_n \in \mathbb{F}_q$ such that f is equivalent to the diagonal quadratic form $a_1X_1^2 + \cdots + a_nX_n^2$.

Proof: We will prove this using induction on the number of variables. If $n = 1$ then $f(X_1) = a_{1,1}X_1^2$. Then its already diagonal. So suppose $n \geq 2$ and the result holds for quadratic forms in $(n-1)$ variables.

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ be a quadratic form in n variables. Now the theorem is trivial if f is a zero polynomial. So suppose f is a non-zero polynomial. If there exists $i \in [n]$, $a_{i,i} \neq 0$ then f represents $a_{i,i}$ since at the point $X_i = 1$ and $X_j = 0$ for all $j \in [n] \setminus \{i\}$, f evaluates to be $a_{i,i}$. If for all $i \in [n]$, $a_{i,i} = 0$ then $\exists\, i, j \in [n]$, $i \neq j$ such that $a_{i,j} = a_{j,i} \neq 0$. Then f represents $2a_{i,j}$ since at the point $X_i = X_j = 1$ and $X_t = 0$ for all $t \in [n] \setminus \{i, j\}$, f evaluates to be $2a_{i,j}$.

So f represents some element $a_1 \in \mathbb{F}_q^*$ and hence by Lemma 3.4.2 we have f is equivalent to a quadratic form $a_1X_1^2 + g(X_2, \ldots, X_n)$ where g is a quadratic form in $(n-1)$ variables. Now by induction hypothesis g is equivalent to a diagonal quadratic form $a_2X_2^2 + \cdots a_nX_n^2$. Therefore f is equivalent to the diagonal quadratic form $a_1X_1^2 + \cdots + a_nX_n^2$. ■

So now we can only talk about the solutions of diagonal quadratic forms since all general quadratic forms in $\mathbb{F}_q[X_1, \ldots, X_n]$ are equivalent to a diagonal quadratic form. Now let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ is equivalent to $a_1X_1^2 + \cdots + a_nX_n^2$. Then some a_i's may be zero. Since non-singular linear transformation preserves ranks, equivalent quadratic forms have coefficient matrices with same rank. Therefore non-singular linear transformation also preserves the non-zero ness of determinant of the coefficient matrix.

Observation 3.7. *Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ is a quadratic form is equivalent to $a_1X_1^2 + \cdots + a_nX_n^2$. Then* $\operatorname{rank}(C_f) = n - |\{i \in [n] : a_i = 0\}|$. *Similarly,* $\det(C_f) = 0$ *if and only if* $\prod_{j=1}^{n} a_i = 0$.

Observation 3.8. *Let $f, g \in \mathbb{F}_q[X_1, \ldots, X_n]$ are two equivalent quadratic forms by the non-singular linear transformation $M \in \mathbb{F}_q^{n \times n}$. Since by Observation 3.5 we have $C_g = M^\top C_f M$ and therefore* $\det(g) = \det(f) \cdot \det(M)^2$.

If the coefficient matrix C_f for any quadratic form $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ has rank n then we call f to be *non-degenerate*. We also define *determinant of f*, $\det(f)$ to be the determinant of the coefficient matrix C_f.

The two-variable lemma below is the base case for the induction used in Theorem 3.4.5. Once we know $Z(a_1X_1^2 + a_2X_2^2 = \alpha)$ exactly, the convolution property Lemma 3.4.1 allows us to build up the solution count for n variables.

Lemma 3.4.4 Solutions of a Two-Variable Diagonal Quadratic Form

For odd q, let $\alpha \in \mathbb{F}_q$, $a_1, a_2 \in \mathbb{F}_q^*$, and let $\eta \in \boldsymbol{\mathcal{M}}_q$ be the quadratic character of $\mathbb{F}_q$. Then

$$Z\left(a_1X_1^2 + a_2X_2^2 = \alpha\right) = q + \vartheta(\alpha) \cdot \eta(-a_1a_2).$$

Proof: Now we can break $Z(a_1X_1^2 + a_2X_2^2 = \alpha)$ into sum over all $\alpha_1, \alpha_2 \in \mathbb{F}_q$ product of $Z(a_1X_1^2 = \alpha_1)$ and $Z(a_2X_2^2 = \alpha_2$.

So we have

$$\begin{aligned}
Z(a_1X_1^2 + a_2X_2^2 = \alpha) &= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}} Z(a_1X_1^2 = \alpha_1)\cdot Z(a_2X_2^2 = \alpha_2)\\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}} (1+\eta(a_1^{-1}\alpha_1))\cdot(1+\eta(a_2^{-1}\alpha_2))\\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}} \left(1+\eta(a_1^{-1}\alpha_1)+\eta(a_2^{-1}\alpha_2)+\eta(a_1^{-1}a_2^{-1}\alpha_1\alpha_2)\right)\\
&= q + \eta(a_1)\sum_{\alpha_1\in\mathbb{F}_q}\eta(\alpha_1) + \eta(a_2)\sum_{\alpha_2\in\mathbb{F}_q}\eta(\alpha_2) + \eta(a_1a_2)\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}}\eta(\alpha_1\alpha_2)\\
&= q + \eta(a_1a_2)\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}}\eta(\alpha_1\alpha_2) && \text{[By Theorem 2.2.5(i)]}\\
&= q + \eta(a_1a_2)\sum_{\gamma\in\mathbb{F}_q}\eta(\gamma(\alpha-\gamma))
\end{aligned}$$

Now consider the polynomial $g(X) = \alpha X - X^2$. Then by Theorem 3.2.7 using the function ϑ we have $\sum_{\gamma\in\mathbb{F}_q}\eta(g(\gamma)) = \vartheta(\alpha)\eta(-1)$. Therefore we have the lemma. ∎

Theorem 3.4.5 Solution Count for Non-degenerate Diagonal Quadratic Forms

Let f be a non-degenerate quadratic form in $\mathbb{F}_q[X_1,\ldots,X_n]$ with n even and q odd. Then for any $\alpha\in\mathbb{F}_q$,

(i) If n is even $Z\left(f(X_1,\ldots,X_n)=\alpha\right) = q^{n-1} + \vartheta(\alpha)\cdot q^{(n-2)/2}\cdot\eta\left((-1)^{n/2}\cdot\det(f)\right)$.

(ii) if n is odd $Z\left(f(X_1,\ldots,X_n)=\alpha\right) = q^{n-1} + q^{(n-1)/2}\cdot\eta\left((-1)^{(n-1)/2}\cdot\alpha\cdot\det(f)\right)$.

Proof: By Theorem 3.4.3 f is equivalent to a diagonal quadratic form $g(X_1,\ldots,X_n) = a_1X_1^2 + \cdots a_nX_n^2$ by the non-singular linear transformation $M\in\mathbb{F}_q^{n\times n}$. By Observation 3.8, we have $\eta(\det(f)) = \eta(\det(g))$. Hence it suffices to prove the theorem for $g(X_1,\ldots,X_n)=\alpha$. Since f is non-degenerate by Observation 3.7, $a_i\neq 0$ for all $i\in[n]$.

(i) Let n is even. So take $m = n/2$. Then

$$\begin{aligned}
Z(g(X_1,\ldots,X_n)=\alpha) &= \sum_{\substack{\alpha_1,\ldots,\alpha_m\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_m=\alpha}}\prod_{j=1}^{m} Z\left(a_{2j-1}X_{2j-1}^2 + a_{2j}X_{2j}^2 = \alpha_j\right)\\
&= \sum_{\substack{\alpha_1,\ldots,\alpha_m\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_m=\alpha}}\prod_{j=1}^{m}\left(q+\vartheta(\alpha_j)\cdot\eta(-a_{2j-1}a_{2j})\right) && \text{[By Lemma 3.4.4]}\\
&= \sum_{\substack{\alpha_1,\ldots,\alpha_m\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_m=\alpha}} q^m + \sum_{\substack{\alpha_1,\ldots,\alpha_m\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_m=\alpha}}\prod_{j=1}^{m}\vartheta(\alpha_j)\cdot\eta(-a_{2j-1}a_{2j})\\
&= q^{m-1}\cdot q^m + \eta\left((-1)^m\det(g)\right)\sum_{\substack{\alpha_1,\ldots,\alpha_m\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_m=\alpha}}\prod_{j=1}^{m}\vartheta(\alpha_j)\\
&= q^{2m-1} + \eta\left((-1)^m\det(g)\right)\cdot\vartheta(\alpha)\cdot q^{m-1} && \text{[By Lemma 3.4.1]}\\
&= q^{n-1} + \eta\left((-1)^{n/2}\det(f)\right)\cdot\vartheta(\alpha)\cdot q^{(n-2)/2}
\end{aligned}$$

(ii) Now suppose n is odd. We will use induction on n. Now this formula is valid for $n = 1$. So suppose $n \geq 3$. Then

$$\begin{aligned}
Z(g(X_1,\ldots,X_n)=\alpha) &= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}} Z(a_1X_1^2=\alpha_1)\cdot Z(a_2X_2^2+\cdots+a_nX_n^2=\alpha_2)\\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}} \left(1+\eta(a_1^{-1}\alpha_1)\right)\cdot\left(q^{n-2}+\eta\left((-1)^{n-1/2}a_2\cdots a_n\right)\cdot\vartheta(\alpha)\cdot q^{(n-3)/2}\right) \qquad \text{[By first part]}\\
&= \sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}} q^{n-1}+q^{n-2}\cdot\eta(a_1)\sum_{\alpha_1\in\mathbb{F}_q}\eta(\alpha_1)\\
&\qquad + q^{(q-3)/2}\cdot\eta\left((-1)^{n-1/2}a_2\cdots a_n\right)\sum_{\alpha_2\in\mathbb{F}_q}\vartheta(\alpha_2)\\
&\qquad + q^{(q-3)/2}\cdot\eta\left((-1)^{n-1/2}a_2\cdots a_n\right)\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}}\vartheta(\alpha_2)\eta(a_1^{-1}\alpha_1)\\
&= q^{n-1}+q^{(q-3)/2}\cdot\eta\left((-1)^{n-1/2}a_1\cdots a_n\right)\sum_{\substack{\alpha_1,\alpha_2\in\mathbb{F}_q\\ \alpha_1+\alpha_2=\alpha}}\vartheta(\alpha_2)\eta(\alpha_1) \qquad \text{[By Lemma 3.4.1]}\\
&= q^{n-1}+q^{(q-3)/2}\cdot\eta\left((-1)^{n-1/2}\cdot\det(f)\right)\sum_{\gamma\in\mathbb{F}_q}\eta(\gamma)\cdot\vartheta(\alpha-\gamma)
\end{aligned}$$

So now we only have to calculate $\sum_{\gamma\in\mathbb{F}_q}\eta(\gamma)\cdot\vartheta(\alpha-\gamma)$. Now

$$\sum_{\gamma\in\mathbb{F}_q}\eta(\gamma)\cdot\vartheta(\alpha-\gamma)=\sum_{\gamma\in\mathbb{F}_q}\eta(\gamma)\cdot[\vartheta(\alpha-\gamma)+1]=q\cdot\eta(\alpha)$$

So we have the results for both cases. ∎

The following alternate proof uses Jacobi sums directly and illustrates how the machinery of the previous chapter applies to count solutions of polynomial equations.

Alternate Proof: Again it suffices to prove this theorem for the diagonal quadratic form $g(X_1,\ldots,X_n)=a_1X_1^2+\cdots a_nX_n^2$ where $a_i\in\mathbb{F}_q^*$ for all $i\in[n]$. Here we will work with the extended definition of η. Let λ_0 is the trivial multiplicative character and $\lambda_1=\eta$. Then

$$\begin{aligned}
Z(g(X_1,\ldots,X_n)=\alpha) &= \sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\prod_{j=1}^{n}Z\left(a_jX_j^2=\alpha_j\right)\\
&= \sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\left(\lambda_1(a_j\alpha_j)+\lambda_0(a_j\alpha_j)\right)\\
&= \sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\sum_{i_1,\ldots,i_n=0}^{1}\lambda_{i_1}(a_1\alpha_1)\cdots\lambda_{i_n}(a_n\alpha_n)\\
&= \sum_{i_1,\ldots,i_n=0}^{1}\left(\prod_{j=1}^{n}\lambda_{i_j}(a_j)\right)\sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\prod_{j=1}^{n}\lambda_{i_j}(\alpha_j)\\
&= \sum_{i_1,\ldots,i_n=0}^{1}\left(\prod_{j=1}^{n}\lambda_{i_j}(a_j)\right)\cdot J_\alpha\left(\lambda_{i_1},\ldots,\lambda_{i_n}\right)
\end{aligned}$$

Now if not all but some λ_{i_j}'s are trivial then $J_\alpha\left(\lambda_{i_1},\ldots,\lambda_{i_n}\right)=0$ by Theorem 2.4.6(ii). So this remains the cases when all are $\lambda_{i,j}$'s are trivial or non-trivial. Now if $i_j=0$ for all $j\in[n]$ then by Theorem 2.4.6(i), $J_\alpha\left(\lambda_{i_1},\ldots,\lambda_{i_n}\right)=q^{n-1}$. Therefore

by

$$Z(g(X_1,\dots,X_n)=\alpha)=q^{n-1}+\eta(\det(f))\cdot J_\alpha\underbrace{(\eta,\dots,\eta)}_{n-\text{times}} \tag{3.4}$$

If $\alpha\neq 0$ then above expression gives us

$$Z(g(X_1,\dots,X_n)=\alpha)=q^{n-1}+\eta(\det(f))\cdot\eta^n(\alpha)\cdot J(\eta,\dots,\eta) \tag{3.5}$$

(i) Now suppose n is even. Let $\alpha\neq 0$. Since n is even η^n is trivial character. So from the equation (3.5) we have

$$Z(g(X_1,\dots,X_n)=\alpha)=q^{n-1}+\eta(\det(f))\cdot J(\eta,\dots,\eta)$$

Let $\chi\in\boldsymbol{\chi}_q$ is a non-trivial additive character of $\mathbb{F}_q$. Then by Theorem 2.4.3(ii) we have

$$J(\eta,\dots,\eta)=-\frac{1}{q}G(\eta,\chi)^n=-\frac{1}{q}\left(G^2(\eta,\chi)\right)^{n/2}=-\frac{1}{q}(\eta(-1)q)^{n/2}=-q^{(n-2)/2}\cdot\eta\left((-1)^{n/2}\right)$$

Hence if $\alpha\neq 0$ then we have

$$Z(g(X_1,\dots,X_n)=\alpha)=q^{n-1}+\eta(\det(f))\cdot\left(-q^{(n-2)/2}\cdot\eta\left((-1)^{n/2}\right)\right)=q^{n-1}-q^{(q-2)/2}\eta\left((-1)^{q/2}\cdot\det(f)\right)$$

Now suppose $\alpha=0$. Then (3.4) gives

$$Z(g(X_1,\dots,X_n)=\alpha)=q^{n-1}+\eta(\det(f))\cdot J_0(\eta,\dots,\eta)$$

Since n is even η^n is trivial. Therefore by Lemma 2.4.2(ii) we have

$$J_0\underbrace{(\eta,\dots,\eta)}_{n-\text{times}}=\eta(-1)\cdot(q-1)\cdot J\underbrace{(\eta,\dots\eta)}_{(n-1)-\text{times}}$$

So by Theorem 2.4.3(i) we have

$$J\underbrace{(\eta,\dots\eta)}_{(n-1)-\text{times}}=\frac{G(\eta,\chi)^{n-1}}{G(\eta^{n-1},\chi)}=G^{n-2}(\eta,\chi)=\left(G^2(\eta,\chi)\right)^{(n-2)/2}=q^{(n-2)/2}\cdot\eta\left((-1)^{(n-2)/2}\right)$$

Therefore from (3.4) we get

$$Z(g(X_1,\dots,X_n)=\alpha)=q^{n-1}+(q-1)\cdot q^{(n-2)/2}\cdot\eta\left((-1)^{n/2}\cdot\det(f)\right)$$

So combining the result of $\alpha\neq 0$ and $\alpha=0$ we get the theorem.

(ii) Now suppose n is odd. Then $\eta^n=\eta$ which is non-trivial. If $\alpha=0$ then by Lemma 2.4.2(i) we have $J_0(\eta,\dots,\eta)=0$. So assume $\alpha\neq 0$. Then in the equation (3.5) by using Theorem 2.4.3(i) we have

$$J(\eta,\dots,\eta)=\frac{G(\eta,\chi)^n}{G(\eta^n,\chi)}=G^{n-1}(\eta,\chi)=\left(G^2(\eta,\chi)\right)^{(n-1)/2}=q^{(n-1)/2}\cdot\eta\left((-1)^{(n-1)/2}\right)$$

So in both cases

$$Z(g(X_1,\dots,X_n)=\alpha)=q^{n-1}+q^{(n-1)/2}\cdot\eta\left((-1)^{(n-1)/2}\cdot\alpha\cdot\det(f)\right)$$

This expression works for $\alpha=0$ because then the second term becomes zero.

Hence we have theorem. ∎

Note:-

For any general degree-2 polynomial $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ there exists $g, h \in \mathbb{F}_q[X_1, \ldots, X_n]$ where g is a quadratic form and $\deg(h) \leq 1$. Then by some non-singular linear transformation we can transform f into an equivalent diagonal quadratic form. Then we obtain an equation of following form:

$$\tilde{f}(X_1, \ldots, X_n) = a_1 X_1^2 + \cdots a_k X_k^2 + b_1 X_1 + \cdots b_n X_n$$

where $a_i \in \mathbb{F}_q^*$ for all $i \in [k]$, $k \leq n$ and $b_j \in \mathbb{F}_q$ for all $j \in [n]$. Hence for any $\alpha \in \mathbb{F}_q$, the number of solutions of $f(X_1, \ldots, X_n) = \alpha$ is same as number of solutions of $\tilde{f}(X_1, \ldots, X_n) = \alpha$.
If $k < n$ then without loss of generality we can assume $b_n \neq 0$. Then the number of solutions is q^{n-1} as we can arbitrarily substitute elements for $X_1, \ldots, X_{n-1}$ and then value of X_n is uniquely determined.
If $k = n$ then takes the non-singular linear transformation $X_i = Y_i - b_i(2a_i)^{-1}$ for all $i \in [n]$ gives an equivalent diagonal quadratic form $a_1 Y_1^2 + \cdots + a_n Y_n^2 = c$ for some $c \in \mathbb{F}_q$. Now we can use Theorem 3.4.5 to find the number of solutions.

3.4.2 Even Characteristic Quadratic Forms

We will now study quadratic forms and its number of solutions over finite fields $\mathbb{F}_q$ with $\text{char}(\mathbb{F}_q) = 2$ i.e. q is even. Again like in the case of odd characteristic we will first reduce any general quadratic form to canonical forms and then we will try to find the number of solutions of such canonical forms.

For even characteristic fields we define *non-degenerate* quadratic forms differently. Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ is a quadratic form in n-variables. Then f is called non-degenerate if f is not equivalent to a quadratic form in fewer than n variables.

Note:-

This definition is consistent with the odd-characteristic case: there, if f is equivalent to a form in fewer than n variables, Theorem 3.4.3 and Observation 3.7 show that $\text{rank}(C_f) < n$, i.e. $\det(f) = 0$, which is the standard criterion for degeneracy. So in both characteristics one can reduce to the non-degenerate case by passing to the equivalent form in fewer variables.

Lemma 3.4.6 Reduction for Quadratic Forms in even characteristic

A non-degenerate quadratic form $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ with q even and $n \geq 3$ is equivalent to $X_1X_2 + g(X_3, \ldots, X_n)$ where g is a quadratic form in $n-2$ variables.

Proof: First we will show that f is equivalent to a quadratic form where coefficient of X_1^2 is 0. So let

$$f(X_1, \ldots, X_n) = \sum_{i,j \in [n]} a_{i,j} X_i X_j$$

where $a_{i,j} = a_{j,i} \in \mathbb{F}_q$. If there exists $i \in [n]$ such that $a_{i,i} = 0$ then we can permute the variables so that X_i goes to $X1$ and thus $a_{1,1} = 0$. So assume $a_{i,i} \neq 0$ for all $i \in [n]$. Now if $a_{i,j} = 0$ for all $i \neq j$ then

$$f(X_1, \ldots, X_n) = a_{1,1} X_1^2 + \cdots + a_{n,n} X_n^2 = \left(a_{1,1}^{q/2} X_1 + \cdots + a_{n,n}^{q/2} X_n \right)^2$$

then f is equivalent to Y_1^2 by the non-singular linear transformation

$$Y_1 = a_{1,1}^{q/2} X_1 + \cdots + a_{n,n}^{q/2} X_n$$
$$Y_i = X_i \; \forall \; i \in \{2, \ldots, n\}$$

This contradicts the non-degenerate property of f ↯ So there exists $i, j \in [n]$, $i \neq j$ such that $a_{i,j} \neq 0$. Without loss of generality suppose $a_{2,3} \neq 0$. Now

$$f(X_1, \ldots, X_n) = a_{2,2} X_2 + X_2(a_{1,2} X_1 + a_{2,3} X_3 + \cdots + a_{2,n} X_n) + g_1(X_1, X_3 \ldots, X_n)$$

Then take the non-singular linear transformation

$$Y_3 = a_{2,3}^{-1}(a_{1,2}X_1 + a_{2,3}X_3 + \cdots + a_{2,n}X_n)$$
$$Y_i = X_i \; \forall \; i \in [n] \setminus \{3\}$$

This gives an equivalent quadratic form $a_{22}Y_2^2 + Y_2Y_3 + g_2(Y_1, Y_3, \ldots, Y_n)$. Let b is the coefficient of X_1^2 in g_2. So now take the non-singular linear transformation

$$Z_2 = \left(a_{2,2}^{-1}b\right)^{q/2} Y_1 + Y_2$$
$$Z_i = Y_i \; \forall \; i \in [n] \setminus \{2\}$$

This transformation makes the coefficient of Z_1^2 to be 0. Now we'll use f to denote this final quadratic form with variables $X_1, \ldots, X_n$. Since f is non-degenerate not all $a_{i,j}$'s are 0 where $i \neq j$. Without loss of generality suppose $a_{1,2} \neq 0$. Then take the non-singular linear transformation

$$Y_2 = a_{1,2}^{-1}(Y_2 + a_{1,3}Y_3 + \cdots + a_{1,n}Y_n)$$
$$Y_i = X_i \; \forall \; i \in [n] \setminus \{2\}$$

Then we get a quadratic form of the type

$$Y_1Y_2 + \sum_{2 \leq i,j \leq n} c_{i,j}Y_iY_j$$

where $c_{i,j} \in \mathbb{F}_q$ for all $2 \leq i, j \leq n$. So take the non-singular linear transformation

$$Z_1 = Y_1 + c_{2,2}Y_2 + \cdots + c_{2,n}Y_n$$
$$Z_i = Y_i \; \forall \; i \in [n] \setminus \{1\}$$

This transformation gives a quadratic form of the type $Z_1Z_2 + g_3(Z_3, \ldots, Z_n)$. So we have the lemma. ∎

So now we will use this theorem to find proper canonical forms for any general quadratic forms. But we have to treat the cases of odd and even number of variables separately.

Theorem 3.4.7 Canonical Forms for even characteristic

Let $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ be a non-degenerate quadratic form with q even.

(i) If n is odd then f is equivalent to $X_1X_2 + X_3X_4 + \cdots + X_{n-2}X_{n-1} + X_n^2$.

(ii) If n is even then f is

(a) either equivalent to $X_1X_2 + \cdots + X_{n-1}X_n$,

(b) or to $X_1X_2 + \cdots + X_{n-3}X_{n-2} + X_{n-1}^2 + a \cdot X_n^2$ where $a \in \mathbb{F}_q$ satisfies $\mathrm{Tr}(a) = 1$ where Tr is the absolute trace function from $\mathbb{F}_q$ to $\mathbb{F}_2$.

***Proof*:**

(i) Suppose n is odd. By applying Lemma 3.4.6 again and again on f we get an equivalent quadratic form of the type $X_1X_2 + X_3X_4 + \cdots + X_{n-2}X_{n-1} + aX_n^2$ for some $a \in \mathbb{F}_q^*$. Then take the non-singular linear transformation

$$Y_n = a^{-q/2}X_n$$
$$Y_i = X_i \; \forall \; i \in [n-1]$$

Then we get the desired quadratic form.

(ii) So now assume n is even. Again by applying Lemma 3.4.6 again and again on f we get an equivalent quadratic form of the type

$$X_1X_2 + \cdots + X_{n-3}X_{n-2} + bX_{n-1}^2 + cX_{n-1}X_n + dX_n^2$$

with $b, c, d \in \mathbb{F}_q$. If $c = 0$ then take the non-singular linear transformation

$$\begin{aligned} Y_{n-1} &= b^{q/2}X_{n-1} + d^{q/2}X_n \\ Y_i &= X_i \ \forall \ i \in [n] \setminus \{n-1\} \end{aligned}$$

Then we get a quadratic form with $(n-1)$-variables. But f is non-degenerate. Hence contradiction ↯ Therefore $c \in \mathbb{F}_q^*$.

If $b = 0$ then take the non-singular linear transformation

$$\begin{aligned} Y_{n-1} &= cX_{n-1} + dX_n \\ Y_i &= X_i \ \forall \ i \in [n] \setminus \{n-1\} \end{aligned}$$

With this transformation we get the first form.

If $b \neq 0$ then take the non-singular linear transformation

$$\begin{aligned} Y_{n-1} &= b^{-q/2}X_{n-1} \\ Y_n &= b^{q/2}c^{-1}X_n \\ Y_i &= X_i \ \forall \ i \in [n-2] \end{aligned}$$

Then we get a quadratic form of the type

$$Y_1Y_2 + \cdots + Y_{n-3}Y_{n-2} + Y_{n-1}^2 + Y_{n-1}Y_n + aY_n^2$$

for some $a \in \mathbb{F}_q$. If $X^2 + X + a$ is reducible in $\mathbb{F}_q$ then there exists $c_1, c_2 \in \mathbb{F}_q$ such that

$$X^2 + X + a = (X + c_1)(X + c_2) \implies Y_{n-1}^2 + Y_{n-1}Y_n + aY_n^2 = (Y_{n-1} + c_1Y_n)(Y_{n-1} + c_2Y_n)$$

Therefore if we take the non-singular linear transformation

$$\begin{aligned} Z_{n-1} &= Y_{n-1} + c_1Y_n \\ Z_n &= Y_{n-1} + c_2Y_n \\ Z_i &= Y_i \ \forall \ i \in [n-2] \end{aligned}$$

then we get a quadratic form of the first form. Now $X^2 + X + a$ is irreducible in $\mathbb{F}_q[X]$ if $\mathrm{Tr}(a) = 1$. Then we have the second form.

Therefore we have the canonical forms for all cases. ∎

Now by the above theorem it suffices to only find the number of solutions for canonical non-degenerate quadratic forms. We will first show two results on two variable version then we will use it to show the general version.

Lemma 3.4.8 Solutions of a Two-Variable Quadratic Form in even characteristic

For even q, let $a, \alpha \in \mathbb{F}_q$ with $\mathrm{Tr}(a) = 1$. Then

$$Z\left(X_1^2 + X_1X_2 + aX_2^2 = \alpha\right) = q - \vartheta(\alpha).$$

where Tr is the absolute trace function from $\mathbb{F}_q$ to $\mathbb{F}_2$.

Proof: Since $\mathrm{Tr}(a) = 1$ the polynomial $X^2 + X + a$ is irrducible in $\mathbb{F}_q[X]$. So there exists $\gamma \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$ such that $X^2 + X + a = (X + \gamma)(X + \gamma^q)$. So

$$f(X_1, X_2) = X_1^2 + X_1X_2 + aX_2^2 = (X_1 + \gamma X_2)(X_1 + \gamma^q X_2)$$

Therefore for any $(\beta_1, \beta_2) \in \mathbb{F}_q^2$ we have

$$f(\beta_1, \beta_2) = (\beta_1 + \gamma\beta_2)(\beta_1 + \gamma^q\beta_2) = (\beta_1 + \gamma\beta_2)(\beta_1 + \gamma\beta_2)^q = (\beta_1 + \gamma\beta_2)^{q+1}$$

So we get

$$Z(f(X_1, X_2) = \alpha) = |\{\gamma \in \mathbb{F}_{q^2} : \gamma^{q+1} = \alpha\}|$$

Hence if $\alpha = 0$ then $Z\left(X_1^2 + X_1X_2 + aX_2^2 = 0\right) = 1 = q - \vartheta(0)$. If $\alpha \neq 0$ since $\mathbb{F}_{q^2}^*$ is cyclic we have $\alpha^{(q^2-1)/q+1} = \alpha^{q-1} = 1$ and hence there are $q + 1$ elements in $\mathbb{F}_{q^2}$ such that $\gamma^{q+1} = \alpha$. Hence $Z\left(X_1^2 + X_1X_2 + aX_2^2 = \alpha\right) = q + 1 = q + \vartheta(\alpha)$. ■

Lemma 3.4.9

For even q, let $\alpha \in \mathbb{F}_q$. Then $Z(X_1X_2 = \alpha) = q + \vartheta(\alpha)$.

Proof: If $\alpha \neq 0$ then both X_1 and X_2 has to be assigned with non-zero values. After assigning X_1 arbitrarily from $\mathbb{F}_q^*$ the value of X_2 gets uniquely determined. Therefore $Z(X_1X_2 = \alpha) = q - 1 = q + \vartheta(\alpha)$.

Suppose $\alpha = 0$. Then at least one of X_1 or X_2 gets assigned 0. Then the other variable can take any value from $\mathbb{F}_q^*$. Therefore there are $2q - 1$ such solutions. Hence $Z(X_1X_2 = \alpha) = 2q - 1 = q + \vartheta(\alpha)$. So we have the lemma. ■

Theorem 3.4.10 Solution Counts for Canonical Quadratic Forms in even characteristic

Let $\mathbb{F}_q$ be a finite field with q even. Then for any $\alpha \in \mathbb{F}_q$:

(i) If n is odd, $Z\left(X_1X_2 + \cdots + X_{n-2}X_{n-1} + X_n^2 = \alpha\right) = q^{n-1} + \vartheta(\alpha) \cdot q^{(n-1)/2}$.

(ii) If n is even:

(a) $Z\left(X_1X_2 + \cdots + X_{n-1}X_n = \alpha\right) = q^{n-1} + \vartheta(\alpha) \cdot q^{(n-2)/2}$;

(b) for $a \in \mathbb{F}_q$ with $\mathrm{Tr}(a) = 1$, then $Z\left(X_1X_2 + \cdots + X_{n-1}X_n + X_{n-1}^2 + a \cdot X_n^2 = \alpha\right) = q^{n-1} - \vartheta(\alpha) \cdot q^{(n-2)/2}$, where Tr is the absolute trace function from $\mathbb{F}_q$ to $\mathbb{F}_2$.

Proof:

(i) Suppose n is odd. Since q is even the equation $X^2 = \alpha$ for any $\alpha \in \mathbb{F}_q$ has an unique solution in $\mathbb{F}_q$. Therefore $Z\left(X_1X_2 + \cdots + X_{n-2}X_{n-1} + X_n^2 = \alpha\right) = q^{n-1} + \vartheta(\alpha) \cdot q^{(n-1)/2}$ because after assigning the variables $X_1, \ldots, X_{n-1}$ arbitrarily the value of X_n is uniquely determined.

(ii) So let n is even. Let $m = n/2$. Then we have

$$\begin{aligned} Z\left(X_1X_2 + \cdots + X_{n-1}X_n = \alpha\right) &= \sum_{\substack{\alpha_1,\ldots,\alpha_m \in \mathbb{F}_q \\ \alpha_1+\cdots+\alpha_m=\alpha}} \prod_{j=1}^{n} Z(X_{2j-1}X_{2j} = \alpha_j) \\ &= \sum_{\substack{\alpha_1,\ldots,\alpha_m \in \mathbb{F}_q \\ \alpha_1+\cdots+\alpha_m=\alpha}} \prod_{j=1}^{m} (q + \vartheta(\alpha_j)) && \text{[By Lemma 3.4.9]} \\ &= q^{2m-1} + \sum_{\substack{\alpha_1,\ldots,\alpha_m \in \mathbb{F}_q \\ \alpha_1+\cdots+\alpha_m=\alpha}} \prod_{j=1}^{m} \vartheta(\alpha_j) \\ &= q^{2m-1} + \vartheta(\alpha) \cdot q^{m-1} = q^{n-1} + \vartheta(\alpha) \cdot q^{(n-2)/2} && \text{[By Lemma 3.4.1]} \end{aligned}$$

So we have the first part. Now for the second part we already showed the $n = 2$ case in Lemma 3.4.8. So assume $n \geq 4$.

$$\begin{aligned} & Z\left(X_1X_2 + \cdots + X_{n-3}X_{n-2} + X_{n-1}^2 + a \cdot X_n^2 = \alpha\right) \\ = & \sum_{\substack{\alpha_1,\alpha_2 \in \mathbb{F}_q \\ \alpha_1+\alpha_2=\alpha}} Z\left(X_1X_2 + \cdots + X_{n-3}X_{n-2} = \alpha_1\right) \cdot Z\left(X_{n-1}X_n + X_{n-1}^2 + a \cdot X_n^2 = \alpha_2\right) \\ = & \sum_{\substack{\alpha_1,\alpha_2 \in \mathbb{F}_q \\ \alpha_1+\alpha_2=\alpha}} \left(q^{n-3} + \vartheta(\alpha_1) \cdot q^{(n-4)/2}\right)(q - \vartheta(\alpha_2)) && \text{[By first part and Lemma 3.4.8]} \\ = & \; q^{n-1} - q^{(n-4)/2} \sum_{\substack{\alpha_1,\alpha_2 \in \mathbb{F}_q \\ \alpha_1+\alpha_2=\alpha}} \vartheta(\alpha_1) \cdot \vartheta(\alpha_2) && \text{[By Lemma 3.4.1]} \\ = & \; q^{n-1} - q^{(n-4)/2} \cdot q \cdot \vartheta(\alpha) \\ = & \; q^{n-1} - \vartheta(\alpha) \cdot q^{(n-2)/2} \end{aligned}$$

So now we have the result for second part

Therefore we get the values of number of solutions of quadratic forms in even characteristic finite fields. ∎

§ 3.5 Diagonal Equations

In the previous section we talked about equations involving diagonal quadratic forms. These are a special case of much general family of equations called *diagonal equations*

Definition 3.5.1: Diagonall Equations

A diagonal equation over $\mathbb{F}_q$ is an equation of the type

$$a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = \alpha$$

for some positive integers $k_1, \ldots, k_n \in \mathbb{N}$ with coefficients $a_1, \ldots, a_n \in \mathbb{F}_q^*$ and $\alpha \in \mathbb{F}_q$.

3.5.1 Counting Solutions using Jacobi Sums

Now we express the number of solutions $Z\left(a_1X_1^{k_1}+\cdots+a_nX_n^{k_n}=\alpha\right)$ in terms of Jacobi sums.

$$Z\left(a_1X_1^{k_1}+\cdots+a_nX_n^{k_n}=\alpha\right)=\sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\prod_{j=1}^{n}Z\left(a_jX_j^{k_j}=\alpha_j\right)=\sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\prod_{j=1}^{n}Z\left(X_j^{k_j}=a_j^{-1}\alpha_j\right)$$

Let λ is a multiplicative character of $\mathbb{F}_q$ with $\operatorname{ord}(\lambda)=d=\gcd(k,q-1)$. Then for any $k\in\mathbb{N}$ and $\gamma\in\mathbb{F}_q$ we have $Z(X^k=\gamma)=\sum_{j=0}^{d-1}\lambda^j$. So here let for $j\in[n]$ let $d_i=\gcd(k_i,q-1)$ and λ_i is the multiplicative character of $\mathbb{F}_q$ of order d_i. Then we have

$$\begin{aligned}
Z\left(a_jX_j^{k_j}=\alpha_j\right)&=\sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\prod_{j=1}^{n}Z\left(X_j^{k_j}=a_j^{-1}\alpha_j\right)\\
&=\sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\left(\sum_{t_1=0}^{d_1-1}\lambda_1^{t_1}(a_1^{-1}\alpha_1)\right)\cdots\left(\sum_{t_n=0}^{d_n-1}\lambda_n^{t_n}(a_n^{-1}\alpha_n)\right)\\
&=\sum_{t_1=0}^{d_1-1}\cdots\sum_{t_n=0}^{d_n-1}\lambda_1^{t_1}(a_1^{-1})\cdots\lambda_n^{t_n}(a_n^{-1})\sum_{\substack{\alpha_1,\ldots,\alpha_n\in\mathbb{F}_q\\ \alpha_1+\cdots+\alpha_n=\alpha}}\lambda_1^{t_1}(\alpha_1)\cdots\lambda_n^{t_n}(\alpha_n)\\
&=\sum_{t_1=0}^{d_1-1}\cdots\sum_{t_n=0}^{d_n-1}\overline{\lambda}_1^{t_1}(a_1)\cdots\overline{\lambda}_n^{t_n}(a_n)\cdot J_\alpha\left(\lambda_1^{t_1},\ldots,\lambda_n^{t_n}\right)
\end{aligned}$$

If $(t_1,\ldots,t_n)=(0,\ldots,o)$ then by Theorem 2.4.6(i) we have $J_\alpha(\lambda_1^{t_1},\ldots,\lambda_n^{t_n})=q^{n-1}$. Also by Theorem 2.4.6(ii) if at least one of t_j is zero then $J_\alpha(\lambda_1^{t_1},\ldots,\lambda_n^{t_n})=0$. So we have

$$Z\left(a_jX_j^{k_j}=\alpha_j\right)=q^{n-1}+\sum_{t_1=1}^{d_1-1}\cdots\sum_{t_n=1}^{d_n-1}\overline{\lambda}_1^{t_1}(a_1)\cdots\overline{\lambda}_n^{t_n}(a_n)\cdot J_\alpha\left(\lambda_1^{t_1},\ldots,\lambda_n^{t_n}\right)$$

So we have the following observation:

Observation 3.9. *For any $\alpha\in\mathbb{F}_q$ let $a_1X_1^{k_1}+\cdots+a_nX_n^{k_n}=\alpha$ is a diagonal equation with $k_1,\ldots,k_n\in\mathbb{N}$ and $a_1,\ldots,a_n\in\mathbb{F}_q^*$. Suppose for all $i\in[n]$, $d_i=\gcd(k_i,q-1)$ and $\lambda_i\in\boldsymbol{\mathcal{M}}_q$ be a multiplicative character of $\mathbb{F}_q$ of order d_i, then*

$$Z\left(a_jX_j^{k_j}=\alpha_j\right)=q^{n-1}+\sum_{t_1=1}^{d_1-1}\cdots\sum_{t_n=1}^{d_n-1}\overline{\lambda}_1^{t_1}(a_1)\cdots\overline{\lambda}_n^{t_n}(a_n)\cdot J_\alpha\left(\lambda_1^{t_1},\ldots,\lambda_n^{t_n}\right)\tag{3.6}$$

Now we distinguish the cases when $\alpha=0$ and when $\alpha\neq0$. If $\alpha=0$ then by Lemma 2.4.2(i) for any $(t_1,\ldots,t_n)$ if $\lambda_1^{t_1}\cdots\lambda_n^{t_n}$ is non-trivial then $J_0\left(\lambda_1^{t_1},\ldots,\lambda_n^{t_n}\right)=0$. So let T denote the set of all tuples $(t_1,\ldots,t_n)\in\bigtimes_{i=1}^n[d_i-1]$ such that $\lambda_1^{t_1}\cdots\lambda_n^{t_n}$ is trivial. Then we have the following result

Theorem 3.5.1 Solution Count for Diagonal Equations at Zero

For any diagonal equation $a_1X_1^{k_1}+\cdots+a_nX_n^{k_n}=0$ with $k_1,\ldots,k_n\in\mathbb{N}$ and $a_1,\ldots,a_n\in\mathbb{F}_q^*$, the number of solutions is

$$Z\left(a_1X_1^{k_1}+\cdots+a_nX_n^{k_n}=0\right)=q^{n-1}+\sum_{(t_1,\ldots,t_n)\in T}\overline{\lambda}_1^{t_1}(a_1)\cdots\overline{\lambda}_n^{t_n}(a_n)\cdot J_0\left(\lambda_1^{t_1},\ldots,\lambda_n^{t_n}\right),$$

where $d_i=\gcd(k_i,q-1)$, λ_i has order d_i, and T is the set of tuples $(t_1,\ldots,t_n)$ with each $t_i\geq1$ such that $\lambda_1^{t_1}\cdots\lambda_n^{t_n}$ is trivial.

Now we can a trivial upper bound on the set T. Since $T\subseteq\bigtimes_{i=1}^n[d_i-1]$ we have $|T|\leq\prod_{i=1}^n(d_i-1)$. So we get the following upper bound from the above theorem.

Corollary 3.5.2 Bound on Solutions of Diagonal Equation at Zero

For any diagonal equation $a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = 0$ *with* $k_1, \ldots, k_n \in \mathbb{N}$ *and* $a_1, \ldots, a_n \in \mathbb{F}_q^*$,

$$\left| Z\left(a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = 0\right) - q^{n-1} \right| \leq (q-1) \cdot q^{n/2-1} \prod_{i=1}^{n} (d_i - 1),$$

where $d_i = \gcd(k_i, q-1)$ *for all* $i \in [n]$.

Proof: For all $(t_1, \ldots, t_n) \in T$ we have $\lambda_1^{t_1} \cdots \lambda_n^{t_n}$ is trivial. Therefore $|J_0(\lambda_1^{t_1}, \ldots, \lambda_n^{t_n})| = (q-1) \cdot q^{(n-2)/2}$ by Theorem 2.4.6(iv) Therefore using the trivial bound of T and absolute value of $|J_0(\lambda_1^{t_1}, \ldots, \lambda_n^{t_n})|$ we get the above result. ■

Let λ is the generator of $\boldsymbol{M}_q$. Then for all $i \in [n]$, $\lambda_i = \lambda^{(q-1)/d_i}$. Hence

$$\lambda_1^{t_1} \cdots \lambda_n^{t_n} = \lambda^{t_1 \frac{q-1}{d_1} + \cdots + t_n \frac{q-1}{d_n}} = \lambda^{(q-1)\left(\frac{t_1}{d_1} \cdots + \frac{t_n}{d_n}\right)}$$

Therefore $\lambda_1^{t_1} \cdots \lambda_n^{t_n}$ is trivial if and only if $\sum_{i=1}^n \frac{t_i}{d_i} \in \mathbb{Z}$. So let $M(d_1, \ldots, d_n)$ is the number of all tuples $(t_1, \ldots, t_n) \in \bigtimes_{i=1}^n [d_i - 1]$ such that $\sum_{i=1}^n \frac{t_i}{d_i} \in \mathbb{Z}$. Therefore $|T| = M(d_1, \ldots, d_n)$. So we have the following theorem.

Corollary 3.5.3 Bound on Solutions of Diagonal Equation at Zero (via M)

For any diagonal equation $a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = 0$ *with* $k_1, \ldots, k_n \in \mathbb{N}$ *and* $a_1, \ldots, a_n \in \mathbb{F}_q^*$,

$$\left| Z\left(a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = 0\right) - q^{n-1} \right| \leq (q-1) \cdot q^{(n-2)/2} \cdot M(d_1, \ldots, d_n),$$

where $d_i = \gcd(k_i, q-1)$ *for all* $i \in [n]$.

So now assume $\alpha \neq 0$ then we have $J_\alpha(\lambda_1^{t_1}, \ldots, \lambda_n^{t_n}) = \left(\lambda_1^{t_1} \cdots \lambda_n^{t_n}\right)(\alpha) J(\lambda_1^{t_1}, \ldots, \lambda_n^{t_n})$. So we get the following formulation of number of solutions for $\alpha \neq 0$.

Theorem 3.5.4 Solution Count for Diagonal Equations at Nonzero Values

Let $\alpha \in \mathbb{F}_q^*$. *For any diagonal equation* $a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = \alpha$ *with* $k_1, \ldots, k_n \in \mathbb{N}$ *and* $a_1, \ldots, a_n \in \mathbb{F}_q^*$,

$$Z\left(a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = \alpha\right) = q^{n-1} + \sum_{t_1=1}^{d_1-1} \cdots \sum_{t_n=1}^{d_n-1} \lambda_1^{t_1}(a_1^{-1}\alpha) \cdots \lambda_n^{t_n}(a_n^{-1}\alpha) \cdot J\left(\lambda_1^{t_1}, \ldots, \lambda_n^{t_n}\right),$$

where $d_i = \gcd(k_i, q-1)$ *and* $\lambda_i \in \boldsymbol{M}_q$ *has order* d_i *for all* $i \in [n]$.

Now we will show another way to find the number of solutions for $\alpha \neq 0$ using the notation $M(d_1, \ldots, d_n)$ we used in the case of $\alpha = 0$.

Theorem 3.5.5 Bound on Solutions of Diagonal Equation at Nonzero Values

Let $\alpha \in \mathbb{F}_q^*$. *For any diagonal equation* $a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = \alpha$ *with* $k_1, \ldots, k_n \in \mathbb{N}$ *and* $a_1, \ldots, a_n \in \mathbb{F}_q^*$,

$$\left| Z\left(a_1X_1^{k_1} + \cdots + a_nX_n^{k_n} = \alpha\right) - q^{n-1} \right| \leq \left[\prod_{i=1}^{n} (d_i - 1) - \left(1 - q^{-1/2}\right) \cdot M(d_1, \ldots, d_n) \right] q^{(n-2)/2}.$$

Proof: By Theorem 2.4.6(iii-iv) if $\alpha \neq 0$ then $|J_\alpha(\lambda_1^{t_1}, \ldots, \lambda_n^{t_n})| = q^{(n-1)/2}$. So we have

$$\left| Z\left(a_1X_1^{k_1} + \cdots a_nX_n^{k_n} = \alpha\right) - q^{n-1} \right| \leq \left[\prod_{i=1}^{n} (d_i - 1) - |T| \right] q^{(n-1)/2} + |T| \cdot q^{(q-2)/2} = \left[\prod_{i=1}^{n} (d_i - 1) - \left(1 - q^{-1/2}\right) |T| \right] q^{(n-1)/2}$$

Since $|T| = M(d_1, \ldots, d_n)$ we have the theorem. ∎

3.5.2 A General Formula of $M(d_1, \ldots, d_n)$

We follow the notation from the previous subsection. Now suppose one of the d_i is relatively prime to all other d_j's where $j \neq i$ i.e. $\gcd(d_i, d_j) = 1$ for all $j \in [n]$, $j \neq i$. If for some $(t_1, \ldots, t_n) \in \bigtimes_{i=1}^n [d_i - 1]$ we have $\sum_{i=1}^n \frac{t_i}{d_i} = m \in \mathbb{Z}$ then we have

$$t_1 \cdot (d_2 \cdots d_n) + k \cdot d_1 = m \cdot (d_1, c \ldots d_n)$$

Thus $t_1 \cdot (d_2 \cdots d_n) \equiv 0 \bmod d_1$. Since d_1 is relatively prime with all other d_j for all $j \in \{2, \ldots, n\}$ we have $d_1 \mid t_1$. Hence contradiction ↯ Hence such tuple $(t_1, \ldots, t_n)$ does not exist. Therefore $M(d_1, \ldots, d_n) = 0$.

Now we already have a trivial upper bound on $M(d_1, \ldots, d_n) \leq \prod_{i=1}^n (d_i - 1)$. We will give a general formula for $M(d_1, \ldots, d_n)$. Suppose $D = \operatorname{lcm}(d_1, \ldots, d_n)$. Then for any tuple $(t_1, \ldots, t_n) \in \bigtimes_{i=1}^n [d_i - 1]$ observe that

$$\frac{1}{D} \sum_{h=0}^{D-1} \exp\left[2\pi i h \left(\frac{t_1}{d_1} + \cdots + \frac{t_n}{d_n}\right)\right] = \begin{cases} 1 & \text{if } \frac{t_1}{d_1} + \cdots + \frac{t_n}{d_n} \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

Then using this expression we have

$$\begin{aligned}
M(d_1, \ldots, d_n) &= \sum_{t_1=1}^{d_1-1} \cdots \sum_{t_n=1}^{d_n-1} \frac{1}{D} \sum_{h=0}^{D-1} \exp\left[2\pi i h \sum_{i=1}^n \frac{t_i}{d_i}\right] \\
&= \frac{1}{D} \sum_{h=0}^{D-1} \left(\sum_{t_1=1}^{d_1-1} \exp\left[t_1 \cdot \frac{h}{d_1}\right] \right) \cdots \left(\sum_{t_n=1}^{d_n-1} \exp\left[t_n \cdot \frac{h}{d_n}\right] \right) \\
&= \frac{1}{D} \sum_{h=0}^{D-1} \prod_{i=1}^n \left(\sum_{t_i=1}^{d_i-1} \exp\left[t_i \cdot \frac{h}{d_i}\right] \right) \\
&= \frac{1}{D} \sum_{h=0}^{D-1} \prod_{i=1}^n \left(\sum_{t_i=0}^{d_i-1} \exp\left[t_i \cdot \frac{h}{d_i}\right] - 1 \right) \\
&= \frac{1}{D} \sum_{h=0}^{D-1} \left[(-1)^n + \sum_{k=1}^n (-1)^{n-k} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \left(\sum_{t_{i_1}=0}^{d_{i_1}-1} \exp\left[t_{i_1} \frac{h}{d_{i_1}}\right] \right) \cdots \left(\sum_{t_{i_n}=0}^{d_{i_n}-1} \exp\left[t_{i_n} \frac{h}{d_{i_n}}\right] \right) \right] \\
&= (-1)^n + \frac{1}{D} \sum_{h=0}^{D-1} \sum_{k=1}^n (-1)^{n-k} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \prod_{j=1}^k \left(\sum_{t_{i_j}=0}^{d_{i_j}-1} \exp\left[t_{i_j} \frac{h}{d_{i_j}}\right] \right)
\end{aligned}$$

Now for any $d \in \mathbb{N}$ we have

$$\sum_{j=0}^{d} \exp\left[j \frac{h}{d}\right] = \begin{cases} d & \text{if } h \equiv 0 \bmod d \\ 0 & \text{otherwise} \end{cases}$$

So in the product term at the end of the last expression survives if $d_{i_j} \mid h$ for all $j \in [k]$ i.e. $\operatorname{lcm}(d_{i_1}, \cdots, d_{i_k}) \mid h$. So in the product at the end of last expression we obtain

$$\prod_{j=1}^k \left(\sum_{t_{i_j}=0}^{d_{i_j}-1} \exp\left[t_{i_j} \frac{h}{d_{i_j}}\right] \right) = \begin{cases} d_{i_1} \cdots d_{i_k} & \text{if } d_{i_j} \mid h \text{ for all } j \in [k] \\ 0 & \text{otherwise} \end{cases}$$

Hence we have

$$\begin{aligned} M(d_1,\dots,d_n) &= (-1)^n + \frac{1}{D}\sum_{h=0}^{D-1}\sum_{k=1}^{n}(-1)^{n-k}\sum_{\substack{1\le i_1<\cdots<i_k\le n\\ \operatorname{lcm}(d_{i_1},\dots,d_{i_k})\mid h}}\left(d_{i_1}\cdots d_{i_k}\right)\\ &= (-1)^n + \sum_{h=0}^{D-1}\sum_{k=1}^{n}(-1)^{n-k}\sum_{1\le i_1<\cdots<i_k\le n}\left(d_{i_1}\cdots d_{i_k}\right)\frac{1}{D}\sum_{\substack{h=0\\ \operatorname{lcm}(d_{i_1},\dots,d_{i_n})\mid h}}^{D-1} 1\\ &= (-1)^n + \sum_{h=0}^{D-1}\sum_{k=1}^{n}(-1)^{n-k}\sum_{1\le i_1<\cdots<i_k\le n}\frac{d_{i_1}\cdots d_{i_k}}{\operatorname{lcm}\left(d_{i_1},\dots,d_{i_n}\right)} \end{aligned}$$

So we have the following theorem summarizing the formula.

Theorem 3.5.6 Closed Form for $M(d_1,\dots,d_n)$

Let $d_1,\dots,d_n \in \mathbb{N}$ such that $d_i \mid q-1$ for all $i \in [n]$ and set $D = \operatorname{lcm}(d_1,\dots,d_n)$. Then

$$M(d_1,\dots,d_n) = (-1)^n + \sum_{k=1}^{n}(-1)^{n-k}\sum_{1\le i_1<\cdots<i_k\le n}\frac{d_{i_1}\cdots d_{i_k}}{\operatorname{lcm}\left(d_{i_1},\dots,d_{i_k}\right)}.$$

Chapter 4

Weil Bounds on Special Cases

A central question in arithmetic geometry is: *how many $\mathbb{F}_q$-rational points does an algebraic curve have?* André Weil established in a series of papers [Wei41, Wei48, Wei49] that for any smooth projective absolutely irreducible curve of genus g over $\mathbb{F}_q$, the number of rational points N satisfies $|N - (q+1)| \leq 2gq^{1/2}$. In [Wei41] Weil first proved the Riemann hypothesis for function fields and deduced point-count bounds for hyperelliptic curves. The subsequent papers [Wei48, Wei49] refined the approach and applied it to exponential sums and equations over finite fields. Weil's original proof used the full machinery of algebraic geometry over function fields. This chapter develops the Weil bounds from first principles, working with two families of curves that are central to analytic number theory: the *superelliptic* $Y^d = f(X)$ and the *Artin–Schreier equation* $Y^q - Y = f(X)$.

Character Sums and Rational Points. The number of $\mathbb{F}_q$-rational points on $Y^d = f(X)$ is not just a geometric count; it encodes the multiplicative character sum $\sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))$. Precisely, if $d \mid q-1$ and ψ is a non-trivial multiplicative character of order d, then (by Theorem 4.1.9) bounding $Z(Y^d = f(X))$ is enough to bounding $\sum_\alpha \psi(f(\alpha))$. Analogously, for $Y^q - Y = f(X)$, the solution count encodes the additive character sum $\sum_\alpha \chi(f(\alpha))$. Proving sharp bounds on these character sums is the overarching goal of the chapter.

Two Elementary Proofs. The first two sections establish bounds via the *polynomial method*: constructing an auxiliary polynomial that vanishes with high multiplicity on the solution set, then comparing degree with zero-count.

§ Section 4.1 (*Stepanov Method: $Y^d - f(X)$*) treats the multiplicative case. Many years after Weil's original proof, Stepanov [Ste69] discovered an elementary method to count points on hyperelliptic curves, which he subsequently extended in [Ste70] to prove congruences for a prime modulus and then presented in a unified form at the 1974 ICM [Ste74]. This approach, now known as the *Stepanov polynomial method* or *polynomial method with multiplicities*, works by interpolating an auxiliary polynomial Q that vanishes with high multiplicity on the solution set, then comparing degree with zero-count. It gives $|Z(Y^d = f(X)) - q| \leq 4d^{3/2}q^{1/2}$ (Stepanov's Theorem). The proof is presented in two ways: first under the simplifying assumptions $\gcd(m, d) = 1$ and $q = p$ or p^2; then in full generality, removing the coprimality assumption via a symmetrization argument with the roots of unity. A connection to genus is drawn: H. M. Stark sharpened the constant $4d^{3/2}$ to $2g$ for the hyperelliptic case $d = 2$, recovering the Weil bound.

§ Section 4.2 (*Bombieri Method: $Y^q - Y - f(X)$*) treats the additive case. Bombieri [Bom] adapted and generalized Stepanov's method in his 1972–73 Bourbaki seminar, giving a streamlined treatment that also covers additive character sums via the Artin–Schreier curve. An auxiliary polynomial is constructed using the trace map on field extensions, yielding $|Z_E(Y^q - Y = f(X)) - q^k| < q^{[k/2]+4}$ (Bombieri's Theorem). The strategy parallels Stepanov but exploits the Frobenius structure of the Artin–Schreier curve.

L-function Machinery. The bounds from the polynomial method are sharp in the solution count, but to pass to sharp bounds on character sums one needs to control the individual *reciprocal roots* w_j of an associated L-function.

§ Section 4.3 (*Special L-functions*) constructs the multiplicative function $\xi = (\psi \circ \varsigma) \cdot (\chi \circ \varrho)$ on the group G of rational functions over $\mathbb{F}_q(X)$. Here ς and ϱ are the "multiplicative" and "additive" valuations at the roots of f, respectively. Two subgroups $\tilde{H}$ (where $\psi \circ \varsigma = 1$) and H (where $\chi \circ \varrho = 1$) are identified, and their intersection $\overline{H}$ is where $\xi = 1$. This structure is the function-field analog of the Gaussian period decomposition.

§ Section 4.4 (*Character Sums of $\psi(f(X))\chi(g(X))$*) uses the ξ-function to build the L-function $\overline{L}(z, \xi)$ and shows it is a polynomial of degree $n + m - 1$ (Theorem 4.4.3). The leading coefficient is $\sum_\alpha \psi(f(\alpha))\chi(g(\alpha))$, and the polynomial factors as $\prod_j (1 - w_j z)$. Lifting ξ to a finite extension field $E = \mathbb{F}_{q^k}$ and using the factorization of the lifted L-function (Theorem 4.4.6) shows that the sums over E are $-\sum_j w_j^k$, so all information about the character sum over every extension field is encoded in the finitely many reciprocal roots $w_1, \ldots, w_{n+m-1}$.

Weil Bounds from the Roots. § Section 4.5 (*Weil Bounds via the Lifting Method*) closes the argument. The polynomial-method bounds of § Section 4.1 and § Section 4.2 show $|Z_E(\cdot) - q^k| = O(q^{k/2})$ for large k, which translates (via the power-sum bound) into $|w_j| \le q^{1/2}$ for every reciprocal root. Summing over j then gives:

- $\left|\sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))\right| \le (m-1)q^{1/2}$ when $Y^d - f(X)$ is absolutely irreducible,
- $\left|\sum_{\alpha \in \mathbb{F}_q} \chi(g(\alpha))\right| \le (n-1)q^{1/2}$ under appropriate conditions on $Y^q - Y - g(X)$,
- $\left|\sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))\chi(g(\alpha))\right| \le (n+m-1)q^{1/2}$ for mixed sums.

The chapter closes with the general Weil Bound (Weil Bound Theorem): for any absolutely irreducible polynomial of total degree d, $|Z(f(X,Y) = 0) - q| < \sqrt{2}\, d^{5/2} q^{1/2}$, whose proof rests on the Riemann Hypothesis for curves.

§ 4.1 Stepanov Method: $Y^d - f(X)$

A curve of the form $Y^d = f(X)$ over $\mathbb{F}_q$, where $f \in \mathbb{F}_q[X]$ and $d \ge 2$, is called a *superelliptic curve*. When $d = 2$ it is a *hyperelliptic curve*. The main result of this section gives a sharp bound on the number of $\mathbb{F}_q$-rational points on such a curve, proved via an auxiliary polynomial that vanishes with high multiplicity on the solution set. The theorem we will be proving in this section is the following:

Theorem 4.1.1 Stepanov's Theorem

Suppose that $Y^d - f(X) \in \mathbb{F}_q[X, Y]$ is an absolutely irreducible polynomial where $\deg(f) = m$. Suppose $q > 100dm^2$. Then

$$|Z(Y^d = f(X)) - q| \le 4d^{3/2} q^{1/2}.$$

4.1.1 Absolute Irreducibility Criteria

The hypothesis that $Y^d - f(X)$ is absolutely irreducible is central to Stepanov's Theorem. The following theorem gives a concrete criterion in terms of the factorization of f over $\overline{\mathbb{F}_q}$.

Theorem 4.1.2 Characterisation of Absolute Irreducibility of $Y^d - f(X)$

Suppose $Y^d - f(X) \in \mathbb{F}_q[X, Y]$. Then the following are equivalent:

(i) $Y^d - f(X)$ is absolutely irreducible.

(ii) $Y^d - c \cdot f(X)$ is absolutely irreducible for any $c \in \mathbb{F}_q^*$.

(iii) If f factors into $f(X) = a\prod_{i=1}^{m}(X-\alpha_i)^{k_i}$ over $\overline{\mathbb{F}_q}$ where $\alpha_i \in \overline{\mathbb{F}_q}$ for some positive integers $k_1, \ldots, k_m$, then $r = \gcd(d, k_1, \ldots, k_m) = 1$.

Proof: We will prove this lemma in three steps:

(i) $\implies$ (ii): Suppose $Y^d - c\cdot f(X)$ is reducible over $\overline{\mathbb{F}_q}[X]$. Since $c \in \mathbb{F}_q^*$, we can write $(Y/\sqrt[d]{c})^d - f(X)$. Therefore, $(Y/\sqrt[d]{c})^d - f(X)$ is also reducible over $\overline{\mathbb{F}_q}[X]$. But we have $Y^d - f(X)$ irreducible. Hence, contradiction ↯ So $Y^d - c\cdot f(X)$ is absolutely irreducible.

(ii) $\implies$ (iii): Assume the contrary i.e, $r > 1$. Then consider the polynomial

$$g(X) = \prod_{i=1}^{m}(X-\alpha_i)^{k_i/r}$$

Then we have

$$Y^d - {}^1\!/_a f(X) = Y^d - g(X)^r = (Y^{d/r} - g(X))(g(X)^{r-1} + \cdots g(X) + 1)$$

Therefore, $Y^d - {}^1\!/_a f(X)$ is reducible over $\overline{\mathbb{F}_q}[X]$. But we are given that $Y^d - c\cdot f(X)$ is absolutely irreducible for any $c \in \mathbb{F}_q^*$. Hence, contradiction ↯ We have $r = 1$.

(iii) $\implies$ (i): Let E be the function field $\overline{\mathbb{F}_q}(X)$. Then we can consider $Y^d - f(X)$ as a polynomial of $E[Y] = \overline{\mathbb{F}_q}(X)[Y]$. Suppose $\gamma \in \overline{E}$ be a root of $Y^d - f(X)$ and ζ_d be the d^{th} root of unity. Then notice that

$$Y^d - f(X) = (Y-\gamma)(Y-\zeta_d\cdot\gamma)\cdots(Y-\zeta_d^{d-1}\cdot\gamma)$$

Therefore, $\zeta_d^i\cdot\gamma$ for all $0 \le i < d$ are the roots of $Y^d - f(X)$ over $\overline{E}$. Now suppose for the sake of contradiction assume $Y^d - f(X)$ is reducible over $\overline{\mathbb{F}_q}[X]$. Then there exists a $t \in [d]$ such that

$$\left(Y-\zeta_d^{i_1}\cdot\gamma\right)\cdots\left(Y-\zeta_d^{i_t}\cdot\gamma\right) \in \overline{\mathbb{F}_q}[X,Y]$$

for some $0 \le i_1 < \cdots < i_t \le d-1$. Now the constant term of this polynomial is $\gamma^t\prod_{j=1}^{t}\zeta_d^{ij}$. Therefore, $\gamma^t \in \overline{\mathbb{F}_q}[X]$. Let $l \in \mathbb{Z}_0$ be the smallest non-negative integer such that $\gamma^l \in \overline{\mathbb{F}_q}[X]$. Then $l \mid d$. So consider the polynomial $h(X) = \gamma^l$.

Now since γ is a root of $Y^d - f(X)$ we have $\gamma^d = f(X)$. Therefore, $g^{d/l}(X) = f(X)$. So take $k = {}^d\!/_l$. Then $k \mid k_i$ for all $i \in [n]$. Therefore, we obtained $k > 1$ such that $k \mid \gcd(d, k_1, \ldots, k_m)$. Hence, contradiction ↯ So $Y^d - f(X)$ is absolutely irreducible. ∎

The most useful special case of Theorem 4.1.2 is when $\gcd(\deg f, d) = 1$, which is the criterion used in the first proof of Stepanov's Theorem.

Corollary 4.1.3 Absolute Irreducibility from Coprime Degree

Suppose $f \in \mathbb{F}_q[X]$ with $\deg(f) = m$. Then $Y^d - f(X)$ is absolutely irreducible if $\gcd(m, d) = 1$.

4.1.2 Proof of Stepanov's Theorem

The proof reduces the solution count $Z(Y^d = f(X))$ to the count of $\alpha \in \mathbb{F}_q$ for which $f(\alpha)$ is a perfect d-th power; the following lemma makes this reduction explicit.

Lemma 4.1.4 Solution Count for $Y^d = f(X)$

Let $d \mid q-1$ and $f \in \mathbb{F}_q[X]$ be a polynomial. Let $F(X) = f^{(q-1)/d'}(X)$ where $d' = \gcd(q-1, d)$. Then the number of solutions of $Y^d = f(X)$ in $\mathbb{F}_q^2$ is given by

$$Z(Y^d = f(X)) = Z_0 + d' \cdot Z_1$$

where $Z_0 = |\{\alpha \in \mathbb{F}_q : f(\alpha) = 0\}|$ and $Z_1 = |\{\alpha \in \mathbb{F}_q : F(\alpha) = 1\}|$.

Furthermore, if $Z_2 = |\{\alpha \in \mathbb{F}_q : F^{d'-1}(\alpha) + \cdots + F(\alpha) + 1 = 0\}|$ then $Z_0 + Z_1 + Z_2 = q$.

Proof: Now for any solution $(X, Y) = (\alpha, \beta)$ such that $\beta^d = f(\alpha)$ either $\beta - 0$ or $\beta \neq 0$. So we consider these two cases separately.

Case I: $\beta = 0$ In this case the notice that the set of all $\alpha \in \mathbb{F}_q$ such that $f(\alpha) = 0$ is exactly the set of solutions such that $\beta = 0$. Therefore, the number of solutions for which $\beta = 0$ is Z_0.

Case II: $\beta \neq 0$ If for any solution (α, β) if $\beta \neq 0$ then $f(\alpha) \in \mathbb{F}_q^{*\,(d)}$. Therefore, for all such α's we have

$$f^{(q-1)/d'}(\alpha) = F(\alpha) = 1, \quad \text{where } d' = \gcd(q-1, d).$$

Now if Z_1 is the number of $\alpha \in \mathbb{F}_q$ such that $f(X)^{(q-1)/d'})(\alpha) = 1$. This is exactly the set Z_1. Now for each such $\alpha \in \mathbb{F}_q^*$ there are d' solutions for β in $\mathbb{F}_q$ such that $\beta^d = f(\alpha)$. Hence, there are $d' \cdot Z_1$ such solutions.

Thus, we get the total number of solutions is $Z_0 + d' \cdot Z_1$. Hence, $Z(Y^d = f(X)) = Z_0 + d' \cdot Z_1$. So we have the first part

Now for every $\alpha \in \mathbb{F}_q$, $f(\alpha) \in \mathbb{F}_q$. Therefore, $f^q(\alpha) = f(\alpha)$. Now we get the following factorization.

$$H(X) = f^q(X) - f(X) = f(X) \cdot \left(f^{(q-1)/d'} - 1\right)\left(F(X)^{d'-1} + \cdots + F(X) + 1\right).$$

So $\deg H = q$ and for all $\alpha \in \mathbb{F}_q$, $H(\alpha) = 0$. Now if $H(\alpha) = 0$ then either $f(\alpha) = 0$ or $F(\alpha) = 0$ or $F^{d-1}(\alpha) + \cdots + F(\alpha) + 1 = 0$. The number of zeros in these three cases are Z_0, Z_1 and Z_2 respectively. Therefore, we get $Z_0 + Z_1 + Z_2 = q$. ∎

4.1.2.1 Stepanov's Theorem with Restricted Conditions

To prove Stepanov's Theorem we will assume $d \mid q-1$ and $\gcd(d, m) = 1$, so $Y^d - f(X)$ is absolutely irreducible over $\mathbb{F}_q$. We also temporarily assume that $q = p$ or $q = p^2$ where p is a prime. The key linear-independence property below ensures that a vanishing relation among $h_0, \ldots, h_{d-1}$ forces each coefficient h_{ij} to vanish — it is what makes the polynomial method work.

Lemma 4.1.5

Let $m, d \in \mathbb{N}$ such that $\gcd(d, m) = 1$. Suppose $h_0, h_1, \ldots, h_{d-1}(X) \in \mathbb{F}_q[X]$ are polynomials of the type

$$h_i(X) = \sum_{j=0}^{t} h_{ij}(X) \cdot X^{qj}, \quad \forall\, 0 \leq i \leq d-1$$

where $h_{ij} \in \mathbb{F}_q[X]$ with $\deg(h_{ij}) \leq \frac{q}{d} - m$. If

$$h(X) = h_0(X) + h_1(X) \cdot F(X) + \cdots + h_{d-1}(X) \cdot F(X)^{d-1} \equiv 0$$

then for all $i \in \{0, 1, \ldots, d-1\}$ and $j \in \{0, \ldots, t\}$ we have $h_{ij} \equiv 0$.

Proof: A summand of $h(X)$ is of the form $h_{ij}(X) \cdot X^{qj} \cdot F(X)^i$.

Want: We want to show that for any $i \in \{0, 1, \ldots, d-1\}$ and $j \in \{0, \ldots, t\}$ the degree of term $q_{i,j} \coloneqq h_{ij}(X) \cdot X^{qj} \cdot F(X)^i$ are distinct. Then $h(X) \equiv 0$ directly implies $h_{ij} \equiv 0$ for all $i \in \{0, 1, \ldots, d-1\}$ and $j \in \{0, \ldots, t\}$.

Now for any $i \in \{0, 1, \ldots, d-1\}$ and $j \in \{0, \ldots, t\}$ we have

$$\deg(q_{ij}) = qj + i\,\frac{q-1}{d}m + \deg(h_{ij}) = \frac{q}{d}(dj + im) - \frac{i}{d}m + \deg(h_{ij})$$

Therefore,

$$\frac{q}{d}(dj + im) - m < \deg(q_{ij}) \leq \frac{q}{d}(dj + im) + \frac{q}{d} - m$$

Enough to Show: For pairs $(i, j) \neq (i', j')$ we have $dj + im \neq dj' + i'm$

Suppose there exists $(i, j) \neq (i', j')$ such that $dj + im = dj' + i'm$. Then we have

$$im \equiv i'm \bmod d \implies i \equiv i' \bmod d$$

where the implication comes from $\gcd(m, d) = 1$. But $i, i' \in \{0, 1, \ldots, d-1\}$ so we have $i = i'$. So now we have $j = j'$. Hence, contradiction ↯ So the degrees of each summand in $h(X)$ are distinct. Therefore, $h_{ij} \equiv 0$ for all $i \in \{0, 1, \ldots, d-1\}$ and $j \in \{0, \ldots, t\}$. ∎

The heart of the proof is the construction of a nonzero polynomial Q of controlled degree that vanishes with high multiplicity on the zero set of f and on the preimage of g under F.

Lemma 4.1.6 Existence of Auxiliary Polynomial Q

Let $1 \leq r \leq d-1$ and let $g \in \mathbb{F}_q[X]$ be a polynomial of degree r. Let $M \in \mathbb{N}$ be a positive integer satisfying $M \geq m+1$ and $(M+3)^2 \leq 2q/d$. Then there exists a non-zero polynomial $Q(X) \in \mathbb{F}_q[X]$ with $\deg(Q) \leq r/d \cdot qM + 4mq$ such that for every $\alpha \in \mathbb{F}_q$:

(i) if $g(F(\alpha)) = 0$, then $Q(\alpha) = 0$ with multiplicity at least M; and

(ii) if $f(\alpha) = 0$, then $Q(\alpha) = 0$ with multiplicity at least M.

Proof: We will construct Q of the form

$$Q(X) = f(X)^M \sum_{i=0}^{d-1} h_i(X) \cdot F(X)^i$$

where $h_i(X) = \sum_{j=0}^{t} h_{ij}(X) \cdot X^{qj}$ with $\deg(h_{ij}) \leq \frac{q}{d} - m$ for all $i \in \{0, 1, \ldots, d-1\}$ and $j \in \{0, \ldots, t\}$ like Lemma 4.1.5. Take $t = \lfloor (M + m + 1)^{r/d} \rfloor$.

Computing the Hasse Derivatives. We will calculate $H^{(n)}(Q)$ for $0 \leq n \leq M-1$ for the multiplicity condition. Now for every term $f^M \cdot h_{ij} \cdot F^i$ we have $H^{(n)}(f^M \cdot h_{ij} \cdot F^i) = f^{M-n} h_{ij,n} \cdot F^i$ where

$$\deg(h_{ij,n}) \leq \deg(h_{ij}) + n(k-1) \leq \frac{q}{d} - m + n(k-1) \leq \frac{q}{m} + n(k-1) - 1 \tag{4.1}$$

Condition for Multiplicity at zeros of $g(F(\alpha))$. Now we can write $Q(X) = \tilde{Q}(X, X^q)$ for some $\tilde{Q} \in \mathbb{F}_q[X, Y]$ where

$$\tilde{Q}(X, Y) = f(X)^M \sum_{i=0}^{d-1} \sum_{j=0}^{t} h_{ij}(X) \cdot Y^j \cdot F(X)^i$$

Since $M \leq q$ we have

$$H^{(n)}(Q(X)) = f^{M-n}(X) \sum_{i=0}^{d-1} \sum_{j=0}^{t} h_{ij,n}(X) \cdot F(X)^i \cdot X^{qj}.$$

Let $g(X) = g_0 + g_1 \cdot X + \cdots g_r \cdot X^r$ where $g_i \in \mathbb{F}_q$ Then from these g_i's we get s_{ij}'s such that for all $\alpha \in \mathbb{F}_q$ with $g(\alpha) = 0$ we have

$$\alpha^i = \sum_{i=0}^{r-1} s_{ij} \cdot \alpha^j$$

for all $0 \leq i, j \leq r$. So if for any $\alpha \in \mathbb{F}_q$ we have $g(F(\alpha)) = 0$ we get $F^i(\alpha) = \sum_{i=0}^{r-1} s_{ij} \cdot F(\alpha)^j$. So now replacing with $s_{i,j}$'s we get for any $n \in \{0, 1, \ldots, M-1\}$ we get for any $\alpha \in \mathbb{F}_q$ such that $g(F(\alpha)) = 0$:

$$\begin{aligned}
(H^{(n)}Q)(\alpha) &= f^{M-n}(\alpha) \sum_{i=0}^{d-1} \sum_{j=0}^{t} h_{ij,n}(\alpha) \cdot F^i(\alpha) \cdot \alpha^{qj} \\
&= f^{M-n}(\alpha) \sum_{i=0}^{d-1} \sum_{j=0}^{t} h_{ij,n}(\alpha) \cdot F^i(\alpha) \cdot \alpha^{j} && [\alpha^q = \alpha \; \forall \; \alpha \in \mathbb{F}_q] \\
&= f^{M-n}(\alpha) \sum_{i=0}^{d-1} \sum_{j=0}^{t} \sum_{l=0}^{r-1} h_{ij,n}(\alpha) \cdot F^i(\alpha) \cdot s_{j,l} \cdot \alpha^{l} \\
&= f^{M-n}(\alpha) \sum_{l=0}^{r-1} q_{l,n}(\alpha) \cdot F^i(\alpha) && \left[\text{where } q_{i,n}(X) = \sum_{i=0}^{d-1} \sum_{j=0}^{t} h_{ij,n}(X) \cdot X^j\right]
\end{aligned}$$

Counting Number of Equations. From the degree bound of $h_{ij,n}$ in (4.1) we have $\deg(q_{l,n}) \leq q/d + n(k-1) - 1 + t$. We want $(H^{(n)}Q)(\alpha) = 0$ for all $n \in \{0, \ldots, M-1\}$ for all $\alpha \in \mathbb{F}_q$ such that $g(F(\alpha)) = 0$. So it suffices to have the polynomials $q_{l,n}$ be identically zero. So if s is the total number of equations s then we have

$$\begin{aligned}
s < \sum_{n=0}^{M-1} r\left(\frac{q}{d} + n(k-1) + t\right) = rM\left(\frac{q}{d} + t\right) + \frac{1}{2}r(m-1)M(M-1) \leq \frac{rq}{d}M + rM\frac{r}{d}(M+m+1) + \frac{1}{2}r(m-1)M^2 \\
< \frac{rq}{d}M + \frac{1}{2}rM^2(m+1) + rM(m+1) \quad (4.2)
\end{aligned}$$

Here the last inequality follows from the fact that $r < d$.

Counting Number of Variables. Since we want a non-zero solution for the polynomials h_i which satisfies the above equations, the coefficients of h_{ij} are the variables. Let v be the total number of possible coefficients of $h_{ij,n}$ for all $0 \leq i \leq d-1$ and $0 \leq j \leq t$. Then we have

$$v \geq \left(\frac{q}{d} - m\right) d(t+1) \geq (q - md)\frac{r}{d}(M+m+1) = \frac{rq}{d}M + \frac{rq}{d}(m+1) - rm(M+k+1) \geq \frac{rq}{d}M + \frac{rq}{d}(m+1) - 2rmM \quad (4.3)$$

Existence of non-zero solution. So now if number of variables is more than the number of equations then we get a non-zero solution for the polynomials $h_{i,j}$ and henceforth for h_i. So we want $v > s$. From (4.2) and (4.3) it suffices to have

$$\begin{aligned}
&\frac{rq}{d}M + \frac{1}{2}rM^2(m+1) + rM(m+1) \leq \frac{rq}{d}M + \frac{rq}{d}(m+1) - 2rmM \\
\iff &\frac{1}{2}rM^2(m+1) + rM(m+1) \leq \frac{rq}{d}(m+1) - 2rmM \\
\iff &\frac{1}{2}M^2(m+1) + M(m+1) \leq \frac{q}{d}(m+1) - 2mM \\
\impliedby &\frac{1}{2}M^2(m+1) + 3M(m+1) \leq \frac{q}{d}(m+1) \\
\iff &\frac{1}{2}M^2 + 3M \leq \frac{q}{d}
\end{aligned}$$

The last inequality holds from the assumption that $(M+3)^2 \leq \frac{2q}{d}$. So we get a non-zero solution for the polynomials h_{ij} for all $0 \leq i \leq d-1$ and $0 \leq j \leq t$. Therefore, Q is a non-zero polynomial otherwise Lemma 4.1.5 will show every h_{ij} is a zero polynomial. So Q vanishes with multiplicity at least M for all $\alpha \in \mathbb{F}_q$ such that $g(F(\alpha)) = 0$.

Multiplicity at Zeros of f. Since Q has a factor $f^M(X)$, every $\alpha \in \mathbb{F}_q$ with $f(\alpha) = 0$ is a zero of Q with multiplicity at least M.

Degree Bound. Since $\deg(h_{ij}) \leq \frac{q}{d} - m$ we have $\deg(h_i) \leq \frac{q}{d} - m + tq$. Therefore,

$$\begin{aligned}
\deg(Q) &\leq \underbrace{mM}_{\deg(f^M)} + \underbrace{\frac{q}{d} - m + qt}_{\deg(h_i)} + \underbrace{(d-1)\,\frac{q-1}{d}\,m}_{\deg(F^i)} \\
&< mM + \frac{q}{d} + qm + \frac{qr}{d}(M + m + 1) \\
&< mq^{1/2} + \frac{r}{d}qM + q\left(\frac{1}{d} + 2m + 1\right) < \frac{r}{d}qM + 4mq
\end{aligned}$$

So we have the required degree bound for Q. Hence, we have the lemma. ∎

Combining Lemma 4.1.4, Lemma 4.1.5, and Lemma 4.1.6, we now prove Stepanov's Theorem under the temporary assumptions $\gcd(d, m) = 1$ and $q = p$ or $q = p^2$; these are removed in the following subsubsection.

Theorem 4.1.7 Stepanov's Theorem under Simplifying Assumptions

Let $d \mid q-1$ and $f \in \mathbb{F}_q[X]$ be a polynomial of degree m. Assume that

(i) $\gcd(d, m) = 1$,

(ii) $q \geq 100dm^2$

then

$$|Z(Y^d = f(X)) - q| < 4md^{3/2}q^{1/2}.$$

Proof: For any $g \in \mathbb{F}_q[X]$ of degree r in Lemma 4.1.6 we constructed a polynomial $Q \in \mathbb{F}_q[X]$ where for all $\alpha \in \mathbb{F}_q$ such that $g(F(\alpha)) = 0$ or $f(\alpha) = 0$, α is a zero of Q with multiplicity at least M. Let T_g be the set of all $\alpha \in \mathbb{F}_q$ such that $g(F(\alpha)) = 0$ or $f(\alpha) = 0$. Therefore,

$$|T_g| \cdot M \leq \deg(Q) \leq {}^{r}/_{d} \cdot qM + 4qm \implies |T_g| \leq \frac{r}{d} + 4q\frac{m}{M}$$

So now we choose $M = \left\lfloor \sqrt{2q/d} \right\rfloor - 3$. Since $q > 100dm^2$ we have $M \geq \sqrt{2q/d} - 4 \geq \sqrt{q/d} \geq m + 1$. Therefore, we get

$$|T_g| \leq \frac{r}{d}q + 4md^{1/2}q^{1/2}$$

Now take $g(X) = X - 1$. Then $r- = 1$. Observe that in this case T_g is the set of all $\alpha \in \mathbb{F}_q$ such that either $F(\alpha) = 1$ or $f(\alpha) = 0$. Then $|T_g| = Z_0 + Z_1$ where Z_0 and Z_1 are defined in Lemma 4.1.4. Thus,

$$|T_g| = Z_0 + Z_1 \leq \frac{q}{d} + 4md^{1/2}q^{1/2}$$

Therefore,

$$Z(Y^d = f(X)) = Z_0 + Z_1 \leq d \cdot |T_g| \leq q + 4md^{3/2}q^{1/2} \implies Z(Y^d = f(X)) - q \leq 4md^{3/2}q^{1/2}$$

Now we have to show the lower bound.

Now select $g(X) = X^{d-1} + \cdots + X + 1$. Then $r = d - 1$. Notice that in this case $T_g = Z_0 + Z_2$. Therefore, in this case we have

$$|T_g| = Z_0 + Z_2 \leq \frac{d-1}{d}q + 4md^{1/2}q^{1/2} \implies Z_1 = q - Z_0 - Z_2 \geq \frac{q}{d} - 4md^{1/2}q^{1/2}$$

Therefore we have

$$Z(Y^d = f(X)) \geq d \cdot Z_1 \geq q - 4md^{3/2}q^{1/2}$$

Thus, we have the theorem. ∎

4.1.2.2 Removal of the assumption $\gcd(m, d) = 1$

In this part we will remove the assumption that $\gcd(m, d) = 1$. So now we can not draw conclusion from $\gcd(d, m) = 1$ that $Y^d - f(X)$ is absolutely irreducible. We have to prove Theorem 4.1.7 just using the assumption that $Y^d - f(X)$ is absolutely irreducible.

Note:-

Now observe that this condition was only required in the proof of Lemma 4.1.5.

Reduction to $f(0) \neq 0$. We can assume $f(0) \neq 0$ without loss of generality. Since f is a non-zero polynomial there exists $c \in \mathbb{F}_q$ such that $f(c) \neq 0$. Then we can replace f by $f(X + c)$ and the number of solutions of $Y^d = f(X)$ will not change as well as the polynomials $h_i(X + C)$ in Lemma 4.1.5 and Lemma 4.1.6 doesn't change. So from now on we will assume that $f(0) \neq 0$.

Constructing Auxiliary polynomials. Let $\zeta_1, \ldots, \zeta_d \in \overline{\mathbb{F}_q}$ be the roots of $X^d - 1$, so $X^d - 1 = (X - \zeta_1) \cdots (X - \zeta_d)$. Working in $\overline{\mathbb{F}_q}[T_0, \ldots, T_{d-1}, Y]$ with formal variables $T_0, \ldots, T_{d-1}, Y$, consider the polynomial

$$\widehat{H}(Y, T_0, \ldots, T_{d-1}) = T_0 + T_1 Y + \cdots + T_{d-1}Y^{d-1}$$

and define

$$G = \prod_{i=1}^{d} \widehat{H}(\zeta_i Y, T_0, \ldots, T_{d-1}) \in \overline{\mathbb{F}_q}[T_0, \ldots, T_{d-1}, Y].$$

Since permuting the ζ_i's only permutes the factors of the product, G is symmetric in $\zeta_1 Y, \ldots, \zeta_d Y$. By Theorem A.5.2, G is a polynomial in elementary symmetric polynomials in $\zeta_1 Y, \ldots, \zeta_d Y$. Since $X^d - 1 = (X - \zeta_1) \cdots (X - \zeta_d)$ we have

$$\mathrm{ESym}_j(\zeta_1, \ldots, \zeta_d) \equiv 0 \text{ for all } j \in [d-1], \quad \text{and} \quad \mathrm{ESym}_d(\zeta_1, \ldots, \zeta_d) \equiv= -Y^{d+1}.$$

Therefore, every monomial of G with non-zero Y-degree as Y-degree multiples of d. Hence, there exists $\widehat{G} \in \overline{\mathbb{F}_q}[T_0, \dots, T_{d-1}, Y]$ such that $G \equiv \widehat{G}(Y^d, T_0, \dots, T_{d-1})$ where $\deg_Y(\widehat{G}) = d-1$ and for all $j \in \{0, \dots, d-1\}$, $\deg_{T_j}(\widehat{G}) = d$. So now we construct the polynomial $P \in \mathbb{F}_q[Y_1, Y_2, T_0, \dots, T_{d-1}]$ as

$$P(Y_1, Y_2, T_0, \dots, T_{d-1}) = Y_2^{d-1} \cdot \widehat{G}\left(\frac{Y_1}{Y_2}, T_0, \dots, T_{d-1}\right)$$

then $\deg(Y_1)P = d-1$ and $\deg_{Y_2}(P) = d-1$ and $\deg_{T_j} = d$ for all $j \in \{0, \dots, d-1\}$.
Want: We want to show that Lemma 4.1.5 holds even with the only assumption that $Y^d - f(X)$ is absolutely irreducible. After that we can apply Lemma 4.1.6 and therefore we have Theorem 4.1.7 without the assumption $\gcd(m, d) = 1$.

Getting to the Case of P applied on h_{i0}'s. Suppose $h_0, \dots, h_{d-1} \in \mathbb{F}_q[X]$ with $h_i(X) = \sum_{j=0}^{t} h_{ij}(X) \cdot X^{qj}$ for all $0 \le i \le d-1$ with degree bounds as defined in Lemma 4.1.5 such that

$$h_0(X) + h_1(X) \cdot F(X) + \cdots + h_{d-1}(X) \cdot F(X)^{d-1} \equiv 0$$

Therefore, we have $\widehat{H}(F, h_0, \dots, h_{d-1}) \equiv 0$. Hence, by construction we have $\widehat{G}(F^d, h_0, \dots, h_{d-1}) \equiv 0$. Now $F(X) = f^{(q-1)/d}(X)$. So $F(X)^d = f^{q-1}(X) = {}^{f^q(X)}/_{f(X)}$. Therefore,

$$P(f^q, f, h_0, \dots, h_{d-1}) = f^{d-1}(X)\widehat{G}\left(\frac{f^q(X)}{f(X)}, h_0, \dots, h_{d-1}\right) = f^{(d-1)}\widehat{G}(F^d, h_0, \dots, h_{d-1}) \equiv 0$$

Gathering all terms with no factor of X^q gives

$$P(f(0), f(X), h_{00}, \cdots, h_{d-1,0}) + X^q \cdot P'(X) \equiv 0 \tag{4.4}$$

Now $\deg(P(f(0), f(X), h_{00}, \cdots, h_{d-1,0})) \le (d-1)m + d\,({}^q/_d - m) < q$. Hence, $P(f(0), f(X), h_{00}, \cdots, h_{d-1,0}) \equiv 0$ by (4.4).
Enough to Show: It suffices to show that $h_{00}, \dots, h_{d-1,0}$ are identically zero. If that happens then we can divide every thing by X^q and run the same argument to show that $h_{ij} \equiv 0$ for all $i \in \{0, \dots, d-1\}$ and $j \in \{0, \dots, t\}$.

Concluding $h_{i,0} \equiv 0$. Let $\gamma \in \overline{\mathbb{F}_q(X)}$ such that $\gamma^d = \frac{f(X)}{f(0)}$. Since $Y^d - f(X)$ is absolutely irreducible by Theorem 4.1.2 $Y^d - \frac{1}{f(0)} f(X)$ is also absolutely irreducible. Now we already have $P(f(0), f(X), h_{00}, \cdots, h_{d-1,0}) \equiv 0$. And

$$P(f(0), f(X), h_{00}, \cdots, h_{d-1,0}) = f^{d-1}(X) \cdot \widehat{G}\left(\frac{f(0)}{f(X)}, h_{00}, \dots, h_{d-1,0}\right) \equiv 0$$

Since f is a non-zero polynomial we have $\widehat{G}\left(\frac{f(0)}{f(X)}, h_{00}, \dots, h_{d-1,0}\right) \equiv 0$. Since $\gamma^d = {}^{f(X)}/_{f(0)}$ we have

$$0 \equiv \widehat{G}\left(\frac{f(0)}{f(X)}, h_{00}, \dots, h_{d-1,0}\right) = \widehat{G}\left(\frac{1}{\gamma^d}, h_{00}, \dots, h_{d-1,0}\right) = G\left({}^1/_\gamma, h_{00}, \dots, h_{d-1,0}\right) \tag{4.5}$$

Therefore, there exists $i \in [d]$ such that $\widehat{H}\left({}^{\zeta_i}/_\gamma, h_{00}, \dots, h_{d-1,0}\right) \equiv 0$. Expanding it out we obtain

$$h_{00}(X) + \frac{\zeta_i}{\gamma} h_{10}(X) + \cdots + \left(\frac{\zeta_i}{\gamma}\right)^{d-1} h_{d-1,0}(X) \equiv 0 \iff h_{00} \cdot \gamma^{d-1} + h_{10} \cdot \gamma^{d-2} \cdot \zeta_i + \cdots + h_{d-1,0} \cdot \zeta_i^{d-1} \equiv 0$$

Let $R(Y) = h_{00} \cdot Y^{d-1} + h_{10} \cdot Y^{d-2} \cdot \zeta_i + \cdots + h_{d-1,0} \cdot \zeta_i^{d-1} \in (X)[Y]$. Now since $Y^d - f(X)$ is absolutely irreducible, $[\overline{\mathbb{F}_q}(X, \gamma) \colon \overline{\mathbb{F}_q}(X)] = d$. But we obtained a polynomial $R(Y)$ of degree $d-1$ over $\overline{\mathbb{F}_q}(X)$ such that $R(\gamma) \equiv 0$. Therefore, $R(Y)$ is a zero polynomial. So $h_{i,0} \equiv 0$ for all $i \in \{0, \dots, d-1\}$. Hence, we are done.

So now we have finally proved the Stepanov's Theorem.

4.1.3 Connection between the Bound and Genus of Curve

The constant $4md^{3/2}$ in Stepanov's Theorem, while explicit, is not tight as a function of the curve's geometry. For the hyperelliptic case $d = 2$, the correct geometric measure of complexity is the *genus* g, and H. M. Stark [Sta73] showed that the Stepanov argument can be refined to replace $4md^{3/2}$ with a constant of size $2g$, recovering the sharp Weil bound.

Genus of a Hyperelliptic Curve. Let $f \in \mathbb{F}_q[X]$ have degree m with m distinct roots, so $Y^2 - f(X)$ is absolutely irreducible. The smooth projective model C of the affine curve $Y^2 = f(X)$ has genus

$$g = \begin{cases} \dfrac{m-1}{2} & \text{if } m \text{ is odd,} \\ \dfrac{m-2}{2} & \text{if } m \text{ is even.} \end{cases} \tag{4.6}$$

Equivalently, $g = \lfloor (m-1)/2 \rfloor$. The difference between the two cases has a geometric explanation: when m is odd the projective closure of $Y^2 = f(X)$ is already smooth at the point at infinity, so the genus is $\frac{m-1}{2}$; when m is even there is a singularity at infinity, and resolving it lowers the genus by $\frac{1}{2}$, giving $\frac{m-2}{2}$. In both cases

$$2g \le m - 1, \quad \text{with equality iff } m \text{ is odd.}$$

The genus is a birational invariant that captures the arithmetic complexity of C more faithfully than the raw degree m: it counts the number of independent holomorphic differentials on C, and it governs both the structure of the Jacobian and the growth rate of the number of rational points.

Weil's Theorem in Terms of Genus. For any smooth projective absolutely irreducible curve of genus g over $\mathbb{F}_q$, Weil's theorem gives

$$|N - (q+1)| \le 2g\, q^{1/2}$$

where N is the number of $\mathbb{F}_q$-rational points (including those at infinity). For the hyperelliptic curve $C\colon Y^2 = f(X)$ with $N = Z(Y^2 = f(X))$, writing this out using (4.6) yields

$$|N - q| \le 2g\, q^{1/2} \ (m \text{ odd}), \qquad |N - q + 1| \le 2g\, q^{1/2} \ (m \text{ even}). \tag{4.7}$$

The constant $2g$ is sharp: it cannot be replaced by any smaller constant independent of q. Comparing with the Stepanov bound $|N - q| < 4m \cdot 2^{3/2}\, q^{1/2} \approx 11.3\, m\, q^{1/2}$, the Weil constant $2g \le m - 1$ is smaller by roughly a factor of 5.6.

Stark's Sharpening. For $q = p$ prime and f with m distinct roots, Stark [Sta73] proved the sharper bound

$$|N - q| \le (m-1)\, q^{1/2}.$$

Translating via (4.6):

$$|N - q| \le 2g\, q^{1/2} \ (m \text{ odd}), \qquad |N - q| \le (2g+1)\, q^{1/2} \ (m \text{ even}). \tag{4.8}$$

For odd m this matches the Weil bound (4.7) exactly. For even m it falls short by $q^{1/2}$, which is explained by the same infinity-point singularity that reduces the genus by $\frac{1}{2}$.

4.1.4 Bound for Multiplicative Character Sum of $f(X)$

To bound $\sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))$ via Stepanov's Theorem, we first need to understand when $f(X)$ is a d-th power in $\mathbb{F}_q[X]$, since that case must be handled separately.

Lemma 4.1.8 Characterisation of d-th Powers in $\mathbb{F}_q[X]$

Let $f(X) \in \mathbb{F}_q[X]$ and let $d \mid q-1$. Then the following are equivalent:

(i) $f(X) = ak^d(X)$ *with* $a \in \mathbb{F}_q$ *and* $k(X) \in \mathbb{F}_q[X]$.

(ii) $f(X) = h^d(X)$ *with* $h(X) \in \overline{\mathbb{F}_q}[X]$.

(iii) $f(X) = a(X-\alpha_1)^{k_1} \cdots (X-\alpha_m)^{k_m}$ *with* $\alpha_i \in \overline{\mathbb{F}_q}$ *and* $d \mid k_i$ *for all* $i \in [m]$.

Proof: We will prove this in three steps.

(i) $\implies$ (ii): Let $\gamma \in \overline{\mathbb{F}_q}$ be such that $\gamma^d = a$. Then take $h(X) = \gamma \cdot k(X)$. Then we have $f(X) = h^d(X)$.

(ii) $\implies$ (iii): This direction comes directly.

(iii) $\implies$ (i): Consider the polynomial $k(X) = \prod_{i=1}^{m}(X-\alpha)i)^{k_i/m}$. Then $f(X) = a \cdot k^d(X)$.
Enough to Show: $k(X) \in \mathbb{F}_q[X]$.

Since $a \in \mathbb{F}_q$ we have $k^d(X) \in \mathbb{F}_q[X]$. Let $k(X) = X^u + c_1X^{u-1} + \cdots + c_u$ where $c_i \in \overline{\mathbb{F}_q}$ for all $i \in [u]$. Now we look at the coefficients of $k^d(X)$. We will prove via induction that each $c_i \in \mathbb{F}_q$. The coefficient of X^{du-1} is $d \cdot c_1$. As $d \neq 0$ and $d \in \mathbb{F}_q$, we have $c_1 \in \mathbb{F}_q$. So suppose $c_1, \ldots, c_{i-1} \in \mathbb{F}_q$. Now the coefficient of X^{du-i} is $dc_i +$ some polynomial over $c_1, \ldots, c_{i-1}$. Since $c_1, \ldots, c_{i-1} \in \mathbb{F}_q$ and coefficient of X^{du-i} is in $\mathbb{F}_q$ we have $dc_i \in \mathbb{F}_q \implies c_i \in \mathbb{F}_q$. Therefore, $k(X) \in \mathbb{F}_q[X]$. Hence, we have the lemma. ∎

When $Y^d - f(X)$ is absolutely irreducible, the Stepanov solution-count bound translates directly into a Weil-type bound on the character sum $\sum_\alpha \psi(f(\alpha))$.

Theorem 4.1.9 Bound for $W(\psi; f)$: Absolutely Irreducible Case

Suppose $d \mid q-1$ and let $\psi \in \boldsymbol{M}_q^{(d)}$ be a non-trivial multiplicative character of exponent d. Let $f(X) \in \mathbb{F}_q[X]$ be a polynomial of degree m such that $Y^d - f(X)$ is absolutely irreducible. Suppose $q > 100dm^3$. Then

$$\left| \sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha)) \right| < 5md^{3/2}q^{1/2}.$$

Proof: Let g be a primitive element of $\mathbb{F}_q$. Let N_k be the number of $\alpha \in \mathbb{F}_q$ such that $f(\alpha)$ is in the coset $g^k \cdot \mathbb{F}_q^{*(d)}$.in the group $\mathbb{F}_q^* / \mathbb{F}_q^{*(d)}$. Then we have

$$W(\psi; f) = \sum_{k=0}^{d-1} N_k \cdot \psi(g^k)$$

Let Z_k be the number of $(\alpha, \beta) \in \mathbb{F}_q^2$ such that $\beta^d = f(\alpha) \cdot g^{-k}$. Since $Y^d - f(X)$ is absolutely irreducible, by Theorem 4.1.2 we have $Y^d - f(X) \cdot g^{-k}$ is also absolutely irreducible. So by Stepanov's Theorem we get

$$|Z_k - q| < 4md^{3/2}q^{1/2}$$

Let $\overline{Z}_k$ be the number of solutions $(\alpha, \beta) \in \mathbb{F}_q^2$ such that $\beta^d = f(\alpha) \cdot g^{-k}$ and $\beta \neq 0$. Then $|Z_k - \overline{Z}_k| \leq m$. So we have

$$|\overline{Z}_k - q| < 5md^{3/2}q^{1/2}$$

Now notice that $N_k = \overline{Z}_k/d$. We can write $Z_k = q/d + R_k$. Therefore, we have

$$|R_k| < 5md^{1/2}q^{1/2}$$

So now

$$W(\psi; f) = \sum_{k=0}^{d-1} N_k \cdot \psi(g^k) = \sum_{k=0}^{d-1} \left(\frac{q}{d} + R_k\right) psi(g^k) = \sum_{k=0}^{d-1} R_k \cdot \psi^k(g)$$

So we get the bound on the absolute value of $W(\psi; f)$ as

$$|W(\psi; f)| \leq \sum_{k=0}^{d-1} |R_k| < 5md^{3/2}q^{1/2}$$

Thus, we have the theorem. ∎

The absolute irreducibility hypothesis in Theorem 4.1.9 can be weakened to the condition that f is not a d-th power, which covers all remaining non-trivial cases.

Theorem 4.1.10 Bound for $W(\psi; f)$: Non-d-th-Power Case

Suppose $d \mid q-1$ and let $\psi \in \mathcal{M}_q^{(d)}$ be a non-trivial multiplicative character with $\operatorname{ord}(\psi) = d > 1$. Let $f(X) \in \mathbb{F}_q[X]$ be a polynomial of degree m which is not a d-th power. Suppose $q > 100dm^3$. Then

$$\left|\sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))\right| < 5md^{3/2}q^{1/2}.$$

Proof: Let $f(X)$ factors as $f(X) = a(X-\alpha_1)^{k_1} \cdots (X-\alpha_t)^{k_t}$ where α_i's are distinct elements of $\overline{\mathbb{F}_q}$. Let $e = \gcd(d, k_1, \ldots, k_t)$. Suppose e is a proper divisor of d. Then consider the polynomial

$$k(X) = \prod_{i=1}^{t} (X - \alpha_i)^{k_i/e}$$

then by Lemma 4.1.8 $f(X) = a \cdot k(X)^e$. Now in $k(X)$, $\gcd(d/e, k_1/e, \ldots, k_t/e) = 1$. So by Theorem 4.1.2 we have $Y^{d/e} - k(X)$ is absolutely irreducible. Now the character ψ^e is also of exponent d/e. So for any $\alpha \in \mathbb{F}_q$ we have

$$\psi(f(\alpha)) = \psi(a) \cdot \psi^e(k(\alpha))$$

So by Theorem 4.1.9 we have

$$|W(\psi; f)| = \left|\sum_{\alpha \in \mathbb{F}_q} \psi(a) \cdot \psi^r(k(\alpha)\right| = \left|\sum_{\alpha \in \mathbb{F}_q} \psi^e(k(\alpha))\right| < 5\left(\frac{m}{e}\right)\left(\frac{d}{e}\right)^{3/2} q^{1/2} \leq 5md^{3/2}q^{1/2}$$

Thus, we have the theorem. ∎

§ 4.2 Bombieri Method: $Y^q - Y - f(X)$

A curve of the form $Y^q - Y = f(X)$ over $\mathbb{F}_q$, where $f \in \mathbb{F}_q[X]$, is called an *Artin–Schreier curve*. Let E be the finite extension of $\mathbb{F}_q$ with $[E\colon \mathbb{F}_q] = k$. The main result of this section is:

Theorem 4.2.1 Bombieri's Theorem

Let $f \in \mathbb{F}_q[X]$ with $(q, \deg f) = 1$ and $\deg f < q$. Then

$$\left|Z_E(Y^q - Y = f(X)) - q^k\right| < q^{[k/2]+4}$$

Note:-

The bound is meaningful when k is large. For small k like $k = 2$ it gives $\left|Z_E(Y^q - Y = f(X)) - q^k\right| < q^5$, whereas the solution space $E^2 = \mathbb{F}_{q^2}^2$ has size only q^4.

4.2.1 Proof of Bombieri's Theorem via the Polynomial Method

The proof mirrors the Stepanov method: reduce the solution count to a count by trace value, then construct an auxiliary univariate polynomial Q that vanishes with high multiplicity at each point with a given trace value, and bound its degree.

Lemma 4.2.2 Trace Decomposition of Solutions

Let for any $\alpha \in \mathbb{F}_q$, Z_α denote the number of $\gamma \in E$ such that $\mathrm{Tr}_{E/\mathbb{F}_q}(\gamma) = \alpha$.

$$\sum_{\alpha \in \mathbb{F}_q} Z_\alpha = q^k \quad \text{and} \quad Z_E(Y^q - Y = f(X)) = q \cdot Z_0.$$

Proof: The first statement is obvious and the later follows from Theorem 1.3.9 ■

Set $r = [k/2]$; we may assume $k \geq 3$ so $r \geq 1$. Split $\mathrm{Tr}_{E/\mathbb{F}_q}(f(X)) = g_0(X) + g_r(X)$ by defining

$$g_r(X) = f(X)^{q^r} + f(X)^{q^{r+1}} + \cdots + f(X)^{q^{k-1}},$$
$$g_0(X) = f(X) + f(X)^q + \cdots + f(X)^{q^{r-1}}.$$

The key lemma constructs a polynomial Q that vanishes with high multiplicity on every $\gamma \in E$ such that $\mathrm{Tr}_{E/\mathbb{F}_q}(f(\gamma)) = \alpha$.

Lemma 4.2.3 Existence of Auxiliary Polynomial Q

Let $\alpha \in \mathbb{F}_q$ and let $M \in \mathbb{N}$ satisfy $q \mid M$ and $0 < M \leq q^{k-r-1}$ where $r = [k/2]$. Then there exists a non-zero polynomial $Q \in \mathbb{F}_q[X]$ with $\deg(Q) \leq M \cdot q^{k-1} + q^{k+1}$ such that every $\gamma \in E$ with $\mathrm{Tr}_{E/\mathbb{F}_q}(f(\gamma)) = \alpha$ is a zero of Q with multiplicity at least M.

Proof: We follow the same approach as Stepanov's Theorem. Construct polynomials $h_0, \ldots, h_{q-1}$ where $h_i(X) = \sum_{j=0}^{t} h_{ij}(X) \cdot X^{q^k \cdot j}$ with $\deg(h_{ij}) < q^{k-1}$, and set

$$Q(X) = \sum_{i=0}^{q-1} h_i(X) \cdot g_r(X)^i.$$

The goal is to choose the h_{ij} so that Q vanishes at every $\gamma \in E$ with $\mathrm{Tr}_{E/\mathbb{F}_q}(f(\gamma)) = \alpha$ with multiplicity at least M. Take $t = \lfloor M/q \rfloor$. Since $k \leq 2r + 1$ we have $M \leq q^r$.

Computing the Hasse Derivatives. Since $g_r(X) = \sum_{i=r}^{k-1} f^{q^i}(X) = \sum_{i=0}^{k-r-1} f\left(X^{q^r}\right)^{q^i}$, define $\tilde{g}_r \in \mathbb{F}_q[Y]$ by $\tilde{g}_r(Y) = \sum_{i=0}^{k-r-1} f(Y)^{q^i}$, so $\tilde{g}_r\left(X^{q^r}\right) = g_r(X)$. Then $Q(X) = \tilde{Q}\left(X, X^{q^r}\right)$ where

$$\tilde{Q}(X, Y) = \sum_{i=0}^{q-1} \sum_{j=0}^{t} h_{ij}(X) \cdot \tilde{g}_r(Y)^i \cdot Y^{q^{k-r}}.$$

Since $M \le q^r$, for any $n < M$ the Hasse derivative satisfies

$$H^{(n)}Q(X) = \sum_{i=0}^{q-1}\sum_{j=0}^{t} h_{ij}^{(n)}(X)\cdot g_r(X)^i \cdot X^{q^r\cdot j}$$

where $h_{ij}^{(n)}(X) = H^{(n)}h_{ij}(X)$.

Imposing the Multiplicity Condition. Since $\mathrm{Tr}_{E/\mathbb{F}_q}(f(X)) = g_0(X) + g_r(X)$, for any $\gamma \in E$ with $\mathrm{Tr}_{E/\mathbb{F}_q}(f(\gamma)) = \alpha$ we have $g_r(\gamma) = \alpha - g_0(\gamma)$ and $\gamma^{q^k} = \gamma$. Substituting, the condition $H^{(n)}(Q)(\gamma) = 0$ for all such γ and all $0 \le n < M$ reduces to requiring that the polynomials

$$Q_n(X) = \sum_{i=0}^{q-1}\sum_{j=0}^{t} h_{ij}^{(n)}(X)\cdot(\alpha - g_0(X))^i \cdot X^j$$

be identically zero for all $0 \le n < M$.

Counting Equations. From the degree bound of h_{ij} we have $\deg(Q_n) \le q^{k-1} + (q-1)^2q^{r-1} + t \le q^{k-1} + q^{r+1} - 2$ for all $0 \le n < M$. Hence the total number of linear equations s on the coefficients of h_{ij} satisfies

$$s \le \sum_{n=0}^{M-1} \deg(Q_n) + 1 < M(q^{k-1} + q^{r+1}) \le M\cdot q^{k-1} + q^k. \tag{4.9}$$

Counting Variables. The unknowns are the coefficients of h_{ij} for all $0 \le i \le q-1$ and $0 \le j \le t$. Their total count is

$$v = \sum_{i=0}^{q-1}\sum_{j=0}^{t}(\deg(h_{ij}) + 1) = q(t+1)q^{k-1} = M\cdot q^{k-1} + q^k. \tag{4.10}$$

Existence of a Non-Zero Solution. Comparing (4.9) and (4.10) gives $v > s$, so there exists a non-zero solution for the h_{ij}, and hence a non-zero Q vanishing with multiplicity at least M at every $\gamma \in E$ with $\mathrm{Tr}_{E/\mathbb{F}_q}(f(\gamma)) = \alpha$. Now

$$\deg(Q) \le q^{k-1}\cdot t + (q-1)^2q^{k-1} + q^{k-1} \le M\cdot q^{k-1} + q^{k+1}$$

Non-Vanishing of Q. Each summand of Q has the form $h_{ij}(X)\cdot g_r(X)^i \cdot X^{q^k\cdot j}$, of degree $q^{k-1}(qj + i\cdot\deg f) + \deg(h_{ij})$. By the same degree-separation argument as in Lemma 4.1.5, distinct pairs (i,j) give summands of distinct degree. Hence, $Q \not\equiv 0$, so we have the lemma. ∎

We now deduce Bombieri's Theorem from Lemma 4.2.3 and Lemma 4.2.2. Fix $\alpha \in \mathbb{F}_q$. Applying Lemma 4.2.3 with parameter M gives a polynomial Q of degree at most $M\cdot q^{k-1} + q^{k+1}$ that vanishes with multiplicity $\ge M$ at every $\gamma \in E$ with $\mathrm{Tr}_{E/\mathbb{F}_q}(f(\gamma)) = \alpha$. Counting zeros gives

$$Z_\alpha \cdot M \le \deg Q \le M\cdot q^{k-1} + q^{k+1} \implies Z_\alpha \le q^{k-1} + \frac{q^{k+1}}{M}.$$

Choose $M = q^{k-r-1}$; then $q \mid M$ for $k \ge 3$, and the hypothesis $M \le q^{k-r-1}$ is satisfied, giving $Z_\alpha \le q^{k-1} + q^{r+2}$. Since $\sum_{\alpha\in\mathbb{F}_q} Z_\alpha = q^k$ by Lemma 4.2.2, we also obtain a lower bound:

$$Z_\alpha = q^k - \sum_{\beta\ne\alpha} Z_\beta \ge q^k - (q-1)(q^{k-1} + q^{r+2}) > q^{k-1} - q^{k+3}.$$

Hence $|Z_\alpha - q^{k-1}| < q^{r+3}$ for every $\alpha \in \mathbb{F}_q$, and in particular $|Z_0 - q^{k-1}| < q^{r+3}$. Since $Z_E(Y^q - Y = f(X)) = q\cdot Z_0$ by Lemma 4.2.2,

$$\left|Z_E(Y^q - Y = f(X)) - q^k\right| = q\cdot|Z_0 - q^{k-1}| < q^{r+4} = q^{[k/2]+4},$$

which is Bombieri's Theorem.

4.2.2 Alternate Proof via Bézout's Theorem

We now give an alternate approach, following [Kop13] and [Kum24]. The argument is more direct but yields a weaker bound. Let E be the finite extension field of $\mathbb{F}_q$ with $[E\colon \mathbb{F}_q] = k$.

Theorem 4.2.4 Bombieri's Bound via Bézout

Let $[E\colon \mathbb{F}_q] = k$ so $|E| = q^k$. Let $f \in \mathbb{F}_q[X]$ be a polynomial of degree d such that $d < q$. Then

$$Z_E(Y^q - Y = f(X)) \leq q^k + O(q \cdot d \cdot q^{k/2}).$$

The final bound on $Z_E(Y^q - Y = f(X))$ comes from bounding the number of common zeros of P and the auxiliary polynomial Q we will construct. The standard Bézout theorem counts common zeros by the product of degrees, but $P = Y^q - Y - f(X)$ has a very unbalanced structure (high Y-degree, low X-degree). The following weighted version exploits this imbalance by assigning different weights to the two variables, producing a bound of the form $D + qd$ rather than Dq.

Lemma 4.2.5 Weighted Bézout's Theorem

For integers $a, b > 0$, the (a,b)-degree of a monomial X^iY^j is $ai + bj$ and of a polynomial is the maximum over its monomials. Let $P(X,Y) = u(Y) - f(X) \in E[X,Y]$ with $\deg_X P = d_X$ and $\deg_Y P = d_Y$. If $Q(X,Y)$ is relatively prime to P and has (d_Y, d_X)-degree at most D, then

$$Z(P,Q) \leq D + d_X \cdot d_Y.$$

Proof: Since P is monic in Y, reduce Q modulo P to get Q_0 with $\deg_Y Q_0 \leq d_Y - 1$, without increasing the (d_Y, d_X)-degree. Then Q_0 has X-degree $\leq D/d_Y$ and Y-degree $\leq d_Y - 1$. By the resultant bound, $Z(P,Q) = Z(P,Q_0) \leq d_X(d_Y - 1) + d_Y \cdot {}^{D}/_{d_Y} = D + d_X d_Y - d_X \leq D + d_X d_Y$. ∎

The dimension argument that produces the auxiliary polynomial $\tilde{Q}$ rests on comparing the number of free coefficient variables against the number of linear constraints forced by the vanishing condition. Both counts are expressed in terms of $|S_N|$, the number of monomials $M_{i,j} = X^iY^j$ whose (q,d)-degree does not exceed a threshold N. The following lemma gives the precise asymptotic.

Lemma 4.2.6 Size of S_N for (q,d)-Degrees

Let $S_N = \{(i,j) : i \geq 0,\ j \in \{0,\ldots,q-1\},\ qi + dj \leq N\}$. For $N \geq (q-1)d$ we have $|S_N| = N - {}^{(q-1)d}/_2 + O(q)$.

Proof: We count $|S_N|$ by summing over the q possible values of j. Fix $j \in \{0,\ldots,q-1\}$. The condition $qi + dj \leq N$ with $i \geq 0$ is equivalent to $0 \leq i \leq {}^{(N-dj)}/_q$. Since $N \geq (q-1)d \geq dj$, this range is non-empty for every j, and the number of valid non-negative integers i is

$$\lfloor (N - dj)/q \rfloor + 1 = \frac{N - dj}{q} + O(1),$$

where the $O(1)$ accounts for the fractional part discarded by the floor. Summing over all $j \in \{0,\ldots,q-1\}$:

$$\begin{aligned} |S_N| &= \sum_{j=0}^{q-1} (\lfloor (N - dj)/q \rfloor + 1) = \sum_{j=0}^{q-1} \frac{N - dj}{q} + O(q) \\ &= \frac{1}{q}\left(qN - d\sum_{j=0}^{q-1} j\right) + O(q) = \frac{1}{q}\left(qN - d \cdot \frac{q(q-1)}{2}\right) + O(q) = N - \frac{(q-1)d}{2} + O(q). \end{aligned}$$

∎

Once Q is constructed, we must show it is not a multiple of P — otherwise the Bézout bound is vacuous. The argument reduces to showing that the reduced monomials appearing in Q all have distinct (q,d)-degrees, so no cancellation can occur. This distinctness follows from the injectivity of the map $(i,j) \mapsto qi + dj$, which in turn relies on $q \nmid d$, a consequence of f being q-free.

Lemma 4.2.7 Injectivity of (q,d)-Degrees

Since f is q-free, $q \nmid d$, and the map $(i,j) \mapsto qi + dj$ is injective on $\{(i,j) : i \geq 0,\ j \in \{0,\dots,q-1\}\}$.

Proof: If $qi + dj = qi' + dj'$ set $e = \gcd(q,d)$; then $(q/e)(i-i') = (d/e)(j'-j)$ and $\gcd(q/e, d/e) = 1$ force $q/e \mid (j'-j)$. Since $q \nmid d$ we have $q/e \geq 2$, so $|j'-j| < q/e$ forces $j = j'$ and then $i = i'$. ∎

Let $M_{i,j}(X,Y) = X^i Y^j$ for $i \geq 0$ and $j \in \{0,\dots,q-1\}$. Set $d_Y = q$ and $d_X = d$.

Reduction Modulo P and Degree Preservation. Since P is monic of degree q in Y, every $R(X,Y)$ reduces mod P to $\bar{R}$ of Y-degree $\leq q-1$ by replacing $Y^q \mapsto Y + f(X)$. A single step on $X^a Y^b$ $(b \geq q)$ gives

$$X^a Y^b \longmapsto X^a Y^{b-q+1} + f(X) \cdot X^a Y^{b-q}.$$

The first term has (q,d)-degree $qa + db - d(q-1) < qa + db$, and each monomial $X^{a+\ell} Y^{b-q}$ $(0 \leq \ell \leq d)$ in the second has degree $q(a+\ell) + d(b-q) \leq qa + db$. By induction, reduction preserves (q,d)-degree, and if R has degree $\leq N$ then $\bar{R} \in E\text{-span}\{M_{i,j} : (i,j) \in S_N\}$. By Lemma 4.2.7 the monomials in S_N have distinct (q,d)-degrees.

Irreducibility of P. Since f is q-free, $f \notin \wp(E(X))$ where $\wp(h) = h^q - h$, so P is irreducible over $E[X]$ by Artin–Schreier theory.

Constructing $\tilde{Q}$: Dimension Argument. Set $r = \lceil k/2 \rceil$. Let $A, B > 0$ with $B < q^{k-r}$. Define

$$\tilde{Q}(X,Y) := \sum_{\substack{(i,j) \in S_A \\ (s,t) \in S_B}} a_{(i,j),(s,t)} \cdot M_{i,j}(X,Y) \cdot \overline{M_{s,t}(X,Y)^{q^r}}$$

with formal variables $a_{(i,j),(s,t)}$. Its (q,d)-degree is at most $A + q^r B$, so $\tilde{Q} \in E\text{-span}\{M_{u,v} : (u,v) \in S_{A+q^r B}\}$. By Lemma 4.2.6 the condition

$$\left(A - \tfrac{(q-1)d}{2}\right)\left(B - \tfrac{(q-1)d}{2}\right) > A + q^r B - \tfrac{(q-1)d}{2} \tag{4.11}$$

ensures $|S_A| \cdot |S_B| > |S_{A+q^r B}|$, giving a nonzero tuple with $P \mid \tilde{Q}$. Fix such a $\tilde{Q}$.

Constructing Q and the Frobenius Identity. Define

$$Q(X,Y) := \sum_{\substack{(i,j) \in S_A \\ (s,t) \in S_B}} a_{(i,j),(s,t)}^{q^{k-r}} \cdot M_{i,j}(X,Y)^{q^{k-r}} \cdot M_{s,t}(X,Y).$$

Lemma 4.2.8

$$Q(X,Y) \equiv \tilde{Q}(X,Y)^{q^{k-r}} \bmod \langle X^{q^k} - X,\ Y^{q^k} - Y \rangle$$

Proof: Taking the q^{k-r} power distributes over the sums:

$$\tilde{Q}^{q^{k-r}} = \sum_{\substack{(i,j) \in S_A \\ (s,t) \in S_B}} a_{(i,j),(s,t)}^{q^{k-r}} \cdot M_{i,j}^{q^{k-r}} \cdot M_{s,t}^{q^r \cdot q^{k-r}} = \sum_{\substack{(i,j) \in S_A \\ (s,t) \in S_B}} a_{(i,j),(s,t)}^{q^{k-r}} \cdot M_{i,j}^{q^{k-r}} \cdot M_{s,t}^{q^k}.$$

For $\alpha, \beta \in E$ we have $\alpha^{q^k} = \alpha$ and $\beta^{q^k} = \beta$, so $M_{s,t}(\alpha,\beta)^{q^k} = \alpha^{q^k s}\beta^{q^k t} = \alpha^s\beta^t = M_{s,t}(\alpha,\beta)$, giving the identity. ∎

Verifying the Three Properties.

(i) Vanishing on $Z_E(Y^q - Y = f(X))$: For $(\alpha,\beta) \in V \subseteq E^2$: $P \mid \tilde{Q}$ gives $\tilde{Q}(\alpha,\beta) = 0$, hence $Q(\alpha,\beta) = \tilde{Q}(\alpha,\beta)^{q^{k-r}} = 0$ by Lemma 4.2.8.

(ii) Relative primality to P.

> **Lemma 4.2.9** Distinct (q,d)-Degrees
>
> *All monomials of Q have distinct (q,d)-degrees; hence $\bar{Q} \neq 0$ and $\gcd(P,Q) = 1$.*

Proof: The degree of $M_{i,j}^{q^{k-r}} \cdot M_{s,t}$ is $q^{k-r}(qi+dj) + (qs+dt)$. If two pairs give the same value then $q^{k-r}\left[(qi+dj) - (qi'+dj')\right] = (qs'+dt') - (qs+dt)$. Since $|(qs+dt) - (qs'+dt')| \leq qB + d(q-1) < q \cdot q^{k-r}$ (using $B < q^{k-r}$ and $d < q^{k-r}$), the left side (a multiple of q^{k-r}) must equal zero. By Lemma 4.2.7 applied twice, both pairs are equal. Hence, degrees are distinct, $\bar{Q} \neq 0$, and since P is irreducible, $\gcd(P,Q) = 1$. ∎

(iii) Applying Lemma 4.2.5 with $D = q^{k-r}A + B$:

$$Z_E(Y^q - Y = f(X)) \leq D + d_X \cdot d_Y = q^{k-r}A + B + qd.$$

Choosing Parameters and the Final Bound. Set $B = q^{k-r} - 1$ and $A = q^r + {}^{(q-1)d+2}/_2$. One verifies (4.11) holds. Substituting:

$$Z_E(Y^q - Y = f(X)) \leq q^{k-r}\left(q^r + \tfrac{(q-1)d+2}{2}\right) + (q^{k-r} - 1) + qd = q^k + \tfrac{(q-1)d+2}{2} \cdot q^{k-r} + q^{k-r} - 1 + qd.$$

Since $r = [{}^k/_2]$, $q^{k-r} \leq q\sqrt{q^k}$, so all error terms are $O(qd\sqrt{q^k})$:

$$Z_E(Y^q - Y = f(X)) \leq q^k + O\left(q \cdot d \cdot \sqrt{q^k}\right),$$

which is the required bound. So we have Theorem 4.2.4.

§ 4.3 Special L-functions

In this section we will construct special L-functions which will help us study polynomials like $Y^d - f(X)$ or $Y^q - Y - f(X)$ where they are absolutely irreducible and character sums related to such polynomials. Remember the notions we introduced in section 2.5.

Let G be the group of all rational functions $r(X) \in \mathbb{F}_q(X)$ such that for any $r(X) = {}^{h_1(X)}/_{h_2(X)}$ where $h_1, h_2 \in \Phi$.
Let $f(X) \in \mathbb{F}_q[X]$ be a monic polynomial in $\mathbb{F}_q[X]$. Let $f(X)$ factors into

$$f(X) = (X - \gamma_1)^{a_1} \cdots (X - \gamma_m)^{a_m}$$

in the field $\overline{\mathbb{F}}_q[X]$ i.e., f has m distinct roots. Then define the subgroup $\overline{G}$ such that $r = {}^{h_1}/_{h_2} \in \overline{G}$ then

$$h_1(\gamma_i) \cdot h_2(\gamma_i) \neq 0 \quad \forall\, i \in [m]$$

Then $\overline{G}$ has the property that if $r_1 \cdot r_2 \in \overline{G}$ then $r_1, r_2 \in \overline{G}$.

Now we define a multiplicative function $\varsigma : \overline{G} \to \mathbb{F}_q$. Let $r = {}^{h_1}/_{h_2} \in \overline{G}$ factors into

$$r(X) = \prod_{i=1}^{d_1}(X - \alpha_i) \cdot \prod_{j=1}^{d_2}(X - \beta_j)^{-1}$$

in $\overline{\mathbb{F}_q}(X)$. Then define

$$\varsigma(r) = \prod_{i=1}^{d_1} f(\alpha_i) \cdot \prod_{j=1}^{d_2} f(\beta_j)^{-1} = \prod_{t=1}^{m} r(\gamma_t)^{a_t}$$

Take $\varsigma(r) = 1$ if $f(X) \equiv 1$. Then $\varsigma(r) \in \mathbb{F}_q$. Then we have the following observation

Observation 4.1. *For all $r_1, r_2 \in \overline{G}$ we have $\varsigma(r_1 \cdot r_2) = \varsigma(r_1) \cdot \varsigma(r_2)$.*

If $\psi \in \boldsymbol{\mathcal{M}}_q$ is a multiplicative character of $\mathbb{F}_q$ then for any $r_1, r_2 \in \overline{G}$ we have

$$\psi(\varsigma(r_1 \cdot r_2)) = \psi(\varsigma(r_1)) \cdot \psi(\varsigma(r_2))$$

Therefore $\psi \circ \varsigma$ is a character on $\overline{G}$. Now we extend $\psi \circ \varsigma$ to G by setting $\psi \circ \varsigma(r) = 0$ for all $r \in G \setminus \overline{G}$. So $\psi \circ \varsigma$ is a character on G.

Let $\tilde{H}$ be a subgroup of $\overline{G}$ consisting only $r = {}^{h_1}/_{h_2}$ with $h_1(\gamma_i) = h_2(\gamma_i) \neq 0$ for all $i \in [m]$. Therefore, by definition of ς we have the following lemma

Lemma 4.3.1

For all $r \in \tilde{H}$, $\psi \circ \varsigma(r) = 1$.

Now let $g \in \mathbb{F}_q[X]$ be any fixed polynomial of degree n and constant term is 0. Again for any $r = {}^{h_1}/_{h_2} \in G$ let r factors like above. Then define a function $\varrho : G \to \mathbb{F}_q$.

$$\varrho(r) = \sum_{i=1}^{d_1} g(\alpha_i) - \sum_{j=1}^{d_2} g(\beta_j)$$

Take $\varrho(r) = 0$ if $r(X) \equiv 1$. Then for all $r \in \mathbb{F}_q$ therefore $\varrho(r) \in \mathbb{F}_q$. Then we have the following observation

Observation 4.2. *For all $r_1, r_2 \in G$ we have $\varrho(r_1 \cdot r_2) = \varrho(r_1) + \varrho(r_2)$.*

Now if $\chi \in \boldsymbol{\mathcal{X}}_q$ is an additive character of $\mathbb{F}_q$ then for all $r_1, r_2 \in G$ we have

$$\chi(\varrho(r_1 \cdot r_2)) = \chi(\varrho(r_1) + \varrho(r_2)) = \chi(\varrho(r_1)) \cdot \chi(\varrho(r_2))$$

Therefore $\chi \circ \varrho$ is a character on G.

Let H be a subset of G consisting of rational functions $r = {}^{h_1}/_{h_2}$ where

$$h_1(X) = X^{d_1} + \sum_{i=1}^{d_1}(-1)^i \cdot a_i \cdot X^{d_1 - i}, \qquad h_2(X) = X^{d_2} + \sum_{i=1}^{d_2}(-1)^i \cdot b_i \cdot X^{d_2 - i}$$

such that $a_i = b_i$ for all $i \in [n]$. Therefore, H is a subgroup of G. Now notice that suppose $h(X)$ is a monic polynomial and $r = {}^{h_1}/_{h_2} \in G$ be a rational function. Now we will show another way of writing this condition which is more algebraic to work with.

$$h_1(X) = X^{d_1}\left(1 + \sum_{i=1}^{d_1}(-1)^i \cdot a_i \cdot \frac{1}{X^i}\right), \qquad h_2(X) = X^{d_2}\left(1 + \sum_{j=1}^{d_2}(-1)^i \cdot b_j \cdot \frac{1}{X^j}\right)$$

So take the polynomials $\widehat{h_1}(Y) = 1 + \sum_{i=1}^{d_1}(-1)^i \cdot a_i \cdot Y^i$ and $\widehat{h_2}(Y) = 1 + \sum_{j=1}^{d_2}(-1)^i \cdot b_j \cdot Y^j$. Now since $a_i = b_i$ for all $i \in [n]$ we have $\widehat{h_1} \equiv \widehat{h_2} \bmod \langle X \rangle^{n+1}$ So we have the following observations

Observation 4.3. *If $r = \frac{h_1}{h_2} \in G$ where $h_1 = X^{d_1} + \sum_{i=1}^{d_1}(-1)^i \cdot a_i \cdot X^{d_1-i}$ and $h_1 = X^{d_2} + \sum_{j=1}^{d_2}(-1)^j \cdot a_j \cdot X^{d_2-j}$ then consider the polynomial $\widehat{h_1}(Y) = \sum_{i=1}^{d_1}(-1)^i \cdot a_i \cdot Y^i$ and $\widehat{h_2}(Y) = \sum_{j=1}^{d_2}(-1)^j \cdot a_j \cdot Y^j$ that is we have*

$$\widehat{h_1} = Y^{d_1} h_1(1/Y), \quad \text{and} \quad \widehat{h_2} = Y^{d_2} h(1/Y)$$

Then $r \in H \iff \widehat{h_1} \equiv \widehat{h_2} \bmod \langle X \rangle^{n+1}$.

Lemma 4.3.2

Let $r = {}^{h_1}/_{h_2} \in G$ be a rational function with $h_1, h_2 \in \Phi$ with degrees d_1, d_2 respectively and let h be a monic polynomial of degree d. Suppose ${}^{h}/_{r} \in H$. Then

$$\widehat{h} \equiv \hat{r} \bmod \langle X \rangle^{n+1}$$

Proof: From h_1, h_2, h we construct the corresponding polynomials $\widehat{h_1}(Y), \widehat{h_2}(Y), \widehat{h}(Y)$. Since $\frac{h}{r} = \frac{h \cdot h_2}{h_1}$ we have ${}^{(h \cdot h_2)}/_{h_1} \in H$. Now we construct the polynomial $\widehat{h \cdot h_2}(Y)$ from $h \cdot h_2$. Then we have

$$\widehat{h \cdot h_2}(Y) = Y^{d+d_2}(h(1/Y) \cdot h_2(1/Y)) = \left(Y^d h(1/Y)\right) \cdot \left(Y^{d_2} h(1/Y)\right) = \widehat{h} \cdot \widehat{h_2}$$

Since ${}^{h}/_{r} \in H$ we have $\widehat{h_1} \equiv \widehat{h_2} \cdot \widehat{h} \bmod \langle X \rangle^{n+1}$. Now we compute $\widehat{r}$. Now $r = {}^{h_1}/_{h_2}$ so $h_2 \cdot r = h_1$ and therefore $\widehat{h_2 \cdot r} = \widehat{h_2} \cdot \widehat{r}$. Hence,

$$\widehat{h_2} \cdot \widehat{r} \equiv \widehat{h_1} \equiv \widehat{h_2} \cdot \widehat{h}$$

Since $\widehat{h_2}$ is a non-zero polynomial we have $\widehat{r} \equiv \widehat{h} \bmod \langle X \rangle^{n+1}$. ∎

Lemma 4.3.3

For all $r \in H$, $\chi \circ \varrho(r) = 1$

Proof: Let $r \in H$ and $r = {}^{h_1}/_{h_2}$ where $h_1, h_2 \in \Phi$ and

$$h_1(X) = X^{d_1} + \sum_{i=1}^{d_1}(-1)^i \cdot a_i \cdot X^{d_1-i}, \qquad h_1(X) = X^{d_2} + \sum_{i=1}^{d_2}(-1)^i \cdot b_i \cdot X^{d_2-i}$$

such that $a_i = b_i$ for all $i \in [n]$. The polynomial $\sum_{i=1}^{k} g(Y_i)$ is a symmetric polynomial for any $k \in \mathbb{N}$ of degree N in variables $Y_1, \ldots, Y_k$. Hence, it is a polynomial in the first n elementary symmetric polynomials $\mathrm{ESym}_j(Y_1, \ldots, Y_k)$ for all $j \in [n]$. Hence, there exists a polynomial $P \in \mathbb{F}_q[X_1, \ldots, X_n]$ such that for any $k \in \mathbb{N}$,

$$\sum_{i=1}^{k} g(Y_i) = P\left(\mathrm{ESym}_1, \ldots, \mathrm{ESym}_n\right)$$

Now if we take $k = d_1$ and plugin $\alpha_1, \ldots, \alpha_{d_1}$ in $Y_1, \ldots, Y_{d_1}$ we obtain

$$\sum_{i=1}^{d_1} g(\alpha_i) = P\left(\mathrm{ESym}_1(\alpha_1, \ldots, \alpha_{d_1}), \ldots, \mathrm{ESym}_n(\alpha_1, \ldots, \alpha_{d_1})\right) = P(a_1, \ldots, a_n)$$

And similarly for $k = d_2$ and assign $\beta_1, \ldots, \beta_{d_2}$ for $Y_1, \ldots, Y_{d_2}$ we obtain

$$\sum_{j=1}^{d_2} g(\beta_j) = P\left(\mathrm{ESym}_1(\beta_1, \ldots, \beta_{d_2}), \ldots, \mathrm{ESym}_n(\beta_1, \ldots, \beta_{d_2})\right) = P(b_1, \ldots, b_n)$$

Since $r \in H$ we have $a_i = b_i$ for all $i \in [n]$. Therefore, $\sum_{i=1}^{d_1} g(\alpha_i) = \sum_{j=1}^{d_2} g(\beta_j)$. Therefore, for all $r \in H$ we have $\varrho(r) = 0$ and hence $\chi \circ \varrho(r) = 1$ for all $r \in H$. ∎

So now we define the multiplicative function $\xi : G \to \mathbb{C}$ where

$$\xi(r) = (\psi \circ \varsigma)(r) \cdot (\chi \circ \varrho)(r)$$

Note:-

Recall the proof of Davenport-Hasse Theorem. We defined similar multiplicative function in the proof. There the corresponding f and g was $f(X) = g(X) = X$.

Let $\overline{H} = H \cap \tilde{H}$. Then from Lemma 4.3.1 and Lemma 4.3.3 we have the following:

Corollary 4.3.4

For all $r \in \overline{H}$ we have $\xi(r) = 1$.

Lemma 4.3.5

Suppose $t \in \mathbb{Z}_0$. Then every coset of $\overline{H}$ in $\overline{G}$ in the group $\overline{G} / \overline{H}$ contains precisely q^t polynomials of degree $n + m + t$.

Proof: Suppose $r(X) = {}^{h_1}/_{h_2} \in \overline{G}$. Then it suffices to show that there are precisely q^t polynomials $h(X) = X^{n+m+t} + \sum_{i=1}^{n+m+t} (-1)^i \cdot b_i \cdot X^{n+m+t-i}$ with ${}^{h}/_{r} \in \overline{H}$. By Lemma 4.3.2 this means there are q^t polynomials $h(X) \in \Phi$ with $\deg(h) = n + m + t$ such that

$$\widehat{h} \equiv \widehat{r} \bmod \langle X \rangle^{n+1} \tag{4.12}$$

$$k(\gamma_i) = r(\gamma_i) \quad \forall i \in [m] \tag{4.13}$$

Now the coefficients $b_1, \ldots, b_n$ are already determined by the condition (4.12). Now the equations in (4.13) gives m linear equations. The matrix of this equation is an $m \times m$ Vandermonde matrix over γ_i^j for all $i \in [m]$ and $j \in \{0, 1, \ldots, m-1\}$. Hence, by the system of linear equations m coefficients $\{b_{n+t+1}, \ldots, b_{n+t+m}\}$ gets uniquely determined. Hence, we can pick $b_{n+1}, \ldots, b_{n+t}$ arbitrarily. Therefore, there are q^t many choices for $h(X)$. ∎

§ 4.4 Characters sums of $\psi(f(X))\chi(g(X))$

In this section we will use the multiplicative function $\xi : G \to \mathbb{C}$ in the previous section and prove bounds on character sums like $\sum_{\alpha \in \mathbb{F}_{q^k}} \psi^{(k)}(f(X))\chi^{(k)}(g(X))$ where $\chi \in \boldsymbol{\chi}_q$ and $\psi \in \boldsymbol{\mathcal{M}}_q$, and they are lifted to $\chi^{(k)}$ and $\psi^{(k)}$ for the field $\mathbb{F}_{q^k}$.

So suppose we have an additive character $\chi \in \boldsymbol{\chi}_q$ and a multiplicative character $\psi \in \boldsymbol{\mathcal{M}}_q$ of $\mathbb{F}_q$. Let $f(X) \in \mathbb{F}_q[X]$ be a monic polynomial in $\mathbb{F}_q[X]$. Let $f(X)$ factors into

$$f(X) = (X - \gamma_1)^{a_1} \cdots (X - \gamma_m)^{a_m}$$

in the field $\overline{\mathbb{F}}_q[X]$ i.e., f has m distinct roots. And let $g \in \mathbb{F}_q[X]$ be any fixed polynomial of degree n and constant term is 0. Define the $G, \overline{G}, \tilde{H}, H$ and $\overline{H}$ like previous section. Then we define the maps $\varsigma : G \to \mathbb{F}_q$ and $\varrho : G \to \mathbb{F}_q$ and the corresponding multiplicative map $\xi : G \to \mathbb{C}$.

4.4.1 L-function on ξ

We want to create the L-function on ξ and use it to calculate the character sums and get bounds on absolute value of the character sum. We will show that the L-function $\overline{L}(z, \xi)$ is of finite degree. First we will show ξ is indeed non-trivial.

Theorem 4.4.1 Non-triviality of ξ

Suppose either $\psi \in \boldsymbol{\mathcal{M}}_q$ is a nontrivial multiplicative character of exponent d and $Y^d - f(X)$ is absolutely irreducible. If the map ξ is non-trivial i.e. $\exists\, r \in \overline{G}$ such that $\chi(r) \neq 1$.

Proof: Suppose $\xi(r) = 1$ for all $r \in \overline{G}$. Then $(\psi \circ \varsigma)(r) \cdot (\chi \circ \varrho)(r) = 1$. Since $\psi \circ \varsigma(r)$ is a d^{th} root of unity and $\chi \circ \varrho(r)$ is a p^{th} roots of unity with $\gcd(d, p) = 1$ we have

$$\psi \circ \varsigma(r) = \chi \circ \varrho(r) = 1$$

Therefore it suffices to find a $k \in \overline{G}$ such that $\psi \circ \varsigma(k) \neq 1$ in the first case and in the second case it will suffice to find a $k \in \overline{G}$ such that $\chi \circ \varrho(k) \neq 1$.

ψ is non-trivial: Suppose $\text{ord}(\psi) = e$ where $e \mid d$. Since $Y^d - f(X)$ is absolutely irreducible not all a_i are multiplies of e by Theorem 4.1.2. Without loss of generality suppose $e \nmid a_1$. For any $c_2, \ldots, c_m \in \mathbb{F}_q^*$ we can pick $c_1 \in \mathbb{F}_q^*$ such that $c_1^{a_1} \cdots c_m^{a_m} \notin \left(\mathbb{F}_q^*\right)^{(e)}$ and hence $\psi\left(c_1^{a_1} \cdots c_m^{a_m}\right) \neq 1$. Now by the arguments of Lemma 4.3.5 there exists a polynomial $k(X) \in \overline{G}$ such that $k(\gamma_i) = c_i$ for all $i \in [m]$. Therefore $\varsigma(k) = c_1^{a_1} \cdots c_m^{a_m}$ and hence $\psi \circ \varsigma(k) \neq 1$.

χ is non-trivial:

(i) Suppose first that $\gcd(n, q) = 1$. Then suppose $g(X) = aX^n + g_1(X)$ where $\deg(g_1) < n$ and $g_1(0) = 0$. If $k(X) = X^n + \alpha = \prod_{j=1}^{n}(X - \alpha_j)$ where $\alpha_j \in \overline{\mathbb{F}_q}$. Then $g_1(\alpha_1) + \cdots g_n(\alpha_n) = 0$ since g_1 is a polynomial with constant term 0 and it is a polynomial in the first $n-1$ elementary symmetric polynomials in $\alpha_1, \ldots, \alpha_n$. Therefore

$$\sum_{j=1}^{n} g(\alpha_j) = a \sum_{j=1}^{n} \alpha_j^n = a \cdot n \cdot (-1)^{n+1} \cdot \alpha$$

Therefore $\varrho(k) = a \cdot n \cdot (-1)^{n+1} \cdot \alpha$ and thus $\chi \circ \varrho(k) = \chi\left(a \cdot n \cdot (-1)^{n+1} \cdot \alpha\right)$. So for a proper choice of α we get $\chi \circ \varrho(k) \neq 1$ since n is not divisible by the $\text{char}(\mathbb{F}_q)$.

(ii) Suppose $Y^q - Y - f(X)$ is absolutely irreducible. Now for any $\alpha \in \mathbb{F}_q^*$ we have

$$\alpha \cdot Y^q - \alpha \cdot Y - g(X) = (\alpha \cdot Y)^q - (\alpha \cdot Y) - f(X)$$

therefore $\alpha \cdot Y^q - \alpha \cdot Y - g(X)$ is absolutely irreducible. Hence $Y^p - Y - \alpha \cdot g(X)$ is also absolutely irreducible where $p = \text{char}(\mathbb{F}_q)$. Let a is such that $\chi(\gamma) = \chi_a(\gamma)$ for all $\gamma \in \mathbb{F}_q$ as defined in Theorem 2.2.1. Let $Z_r(Y^q - Y - a \cdot g(X))$ is the number of solutions of $Y^q - Y - a \cdot g(X)$ over $\mathbb{F}_{q^r}$. Now by Weil Bound Theorem we have $Z_r(Y^q - Y = g(X)) = q^r + O(q^{r/2})$. Hence if r is large $Z_r(Y^q - Y = g(X)) < p \cdot q^r$. Let E denote the field $\mathbb{F}_{q^r}$. Now suppose $c \in \mathbb{F}_{q^r}$. If $\text{Tr}_E(a \cdot g(c)) = 0$ then there are p values of $b \in$ such that $b^q - b - a \cdot g(c) = 0$. If $\text{Tr}_E(a \cdot g(c)) \neq 0$ there are no such $b \in \mathbb{F}_{q^r}$. Since $Z_r(Y^q - Y = g(X)) < p \cdot q^r$ there is a $b \in \mathbb{F}_{q^r}$ with $\text{Tr}_E(a \cdot g(c)) \neq 0$. So take $k(X) = \prod_{i=0}^{r-1}(X - c^{q^i})$. Then

$$\varrho(k) = g(c) + g(c^q) + \cdots g(c^{q^{r-1}}) = \text{Tr}_E(g(c))$$

So $\chi \circ \varrho(k) \neq 1$.

Therefore we have the theorem. ∎

We eventually want to construct the L-function using the multiplicative function ξ. The above theorem helps us to invoke Theorem 2.5.7.

Lemma 4.4.2 Vanishing of Character Sum on ξ

Suppose either $\psi \in \boldsymbol{\mathcal{M}}_q$ is a nontrivial multiplicative character of exponent d and $Y^d - f(X)$ is absolutely irreducible. Or, $\chi \in \boldsymbol{\mathcal{X}}_q$ is a non-trivial additive character and

(i) either, $\gcd(n, q) = 1$

(ii) or, more generally $Y^q - Y - g(X)$ is absolutely irreducible.

Suppose $t \geq 0$. Then

$$\sum_{\substack{h \in \Phi \\ \deg(h) = n+m+t}} \xi(g) = 0$$

Proof: By Corollary 4.3.4 and Theorem 4.4.1, ξ induces a non-trivial character on the finite group $\overline{G}\,/\,H$. As h runs through all polynomials of $\overline{G}$ of degree $n + m + t$ then by Lemma 4.3.5 it will lie precisely land q^l times in every coset of $\overline{G}\,/\,H$. So by Theorem 2.2.2(i) and Theorem 2.2.5(i) we have the lemma. ■

As we mentioned in the previous section. We can extend ξ to G by setting $\xi(h) = 0$ if h is a polynomial not in $\overline{G}$. So now we form the L-functions, $L(s, \xi)$ and $\overline{L}(Z, \xi)$ on ξ. And using the above lemma we have the following theorem:

Theorem 4.4.3 Finite Degree of $\overline{L}(z, \xi)$

Suppose either $\psi \in \boldsymbol{\mathcal{M}}_q$ is a nontrivial multiplicative character of exponent d and $Y^d - f(X)$ is absolutely irreducible. Or, $\chi \in \boldsymbol{\mathcal{X}}_q$ is a non-trivial additive character and

(i) either, $\gcd(n, q) = 1$

(ii) or, more generally $Y^q - Y - g(X)$ is absolutely irreducible.

Then

$$\overline{L}(z, \xi) = 1 + c_1 z + \cdots + c_{n+m-1} z^{n+m-1}, \quad L(s, \xi) = 1 + c_1 u + \cdots + c_{n+m-1} u^{n+m-1}$$

where $u = q^{-s}$.

So we can use Theorem 2.5.7 and factorize $\overline{L}(z, \xi)$ and get complex numbers $w_1, \ldots, w_{n+m-1}$ such that

$$\overline{L}(z, \xi) = \prod_{j=1}^{n+m-1} (1 - w_i z).$$

We can say a lot about the coefficients, specifically c_1.

Theorem 4.4.4 Leading Coefficient c_1 of $\overline{L}(z, \xi)$

If ψ is nontrivial or if ψ is trivial and $f(X) = 1$ then

$$c_1 = \sum_{\gamma \in \mathbb{F}_q} \psi(f(\gamma)) \chi(g(\gamma)).$$

Proof: We have

$$\begin{aligned} c_1 &= \sum_{\substack{h\in\Phi+1\\ \deg(g)=t}} \xi(h) \\ &= \sum_{\substack{\alpha\in\mathbb{F}_q\\ \alpha\neq\gamma_i\ \forall i\in[m]}} \xi(X-\alpha) \\ &= \sum_{\substack{\alpha\in\mathbb{F}_q\\ \alpha\neq\gamma_i\ \forall i\in[m]}} \psi\circ\varsigma(X-\alpha)\cdot\chi\circ\varrho(X-\alpha) \\ &= \sum_{\substack{\alpha\in\mathbb{F}_q\\ f(\alpha)\neq 0}} \psi(f(\alpha))\cdot\chi(g(\alpha)) \\ &= \sum_{\alpha\in\mathbb{F}_q} \psi(f(\alpha))\cdot\chi(g(\alpha)) \end{aligned}$$

So we have the theorem. ∎

4.4.2 Field Extensions

Suppose we aer given a finite field extension E of $\mathbb{F}_q$ where $[E\colon \mathbb{F}_q] = r$. So like subsection 2.2.3 we can lift ψ and χ to $\psi^{(r)}$ and $\chi^{(r)}$ for E. And using these two characters we can define the lifted multiplicative function $\xi^{(r)}$ on the group $G_r, \overline{G}_r$ with rational functions defined similarly as $G, \overline{G}$ but on E. So for any $r \in G_r$ we define

$$\xi^{(r)} = \psi^{(r)}\circ\varsigma(r)\cdot\chi^{(r)}\circ\varrho(r)$$

We define $\tilde{H}_r, H_r, \overline{H}_r$ for $G_r, \overline{G}_r$ similarly as $\tilde{H}, H, \overline{H}$. Now using $\xi^{(r)}$ we also create the L-functions, $L^{(r)}(s,\xi^{(r)})$ and $\overline{L}^{(r)}(z,\xi^{(r)})$. The norm of a polynomial $h \in \Phi^{(r)}$ over E is $\mathfrak{n}^{(r)}(h) \coloneqq q^{r\cdot\deg(h)} = |E|^{\deg(h)}$, the analog of $\mathfrak{n}(h) = q^{\deg(h)}$ for the base field. So we have

$$L^{(r)}(s,\xi^{(r)}) = \sum_{h\in\Phi^{(r)}} \xi^{(r)}(h)\cdot\mathfrak{n}^{(r)}(h)^{-s}, \qquad \overline{L}^{(r)}(z,\xi^{(r)}) = \sum_{h\in\Phi^{(r)}} \xi^{(r)}(h)\cdot z^{\deg(h)}$$

where $\Phi^{(r)}$ is the set of all monic polynomials over E. We similarly define Φ_d for any $d \in \mathbb{Z}_0$. We also use the Weil sum notation from section 3.2:

$$W(\psi,\chi;f,g) \coloneqq \sum_{\alpha\in\mathbb{F}_q} \psi(f(\alpha))\cdot\chi(g(\alpha)), \qquad W^{(r)}(\psi^{(r)},\chi^{(r)};f,g) \coloneqq \sum_{\alpha\in E} \psi^{(r)}(f(\alpha))\cdot\chi^{(r)}(g(\alpha))$$

where the superscript (r) indicates the sum is over the extension field $E = \mathbb{F}_{q^r}$ using the lifted characters.

Lemma 4.4.5 Splitting of Irreducible Polynomials over $\mathbb{F}_{q^k}$

Let $K = \mathbb{F}_{q^k}$ and $E = \mathbb{F}_{q^r}$. Suppose $h(X) = X^d - a_1X^{d-1} + \cdots + (-1)^d a_d$ be an irreducible polynomial in $\mathbb{F}_q[X]$. Then in $\mathbb{F}_{q^k}[X]$, h splits into $r = \gcd(k,d)$ many irreducible polynomials of degree d/r:

$$h(X) = h_1(X)\cdots h_r(X)$$

If we normalize $h_i(X)$ such that h_i is monic then $h_i(X) \in \mathbb{F}_{q^r}[X]$ for all $i \in [r]$. Then

(i) $\mathfrak{n}^{(k)}(h_i) = \mathfrak{n}(h)^{k/r}$

(ii) $\xi^{(k)}(h_i) = \xi^{k/r}(h)$

Proof:

(i) Now by definition $\mathfrak{n}^{(k)}(h_i) = q^{k(d/r)} = \mathfrak{n}(h)^{k/r}$. So we get the first result.

(ii) Since $h = h_1 \cdots h_r$ we have $\varsigma(h) = \prod_{i=1}^{r} \varsigma(h_i)$. Now by Theorem A.1.1 we have $\varsigma(h_i) \in E$ for all $i \in [r]$ and they are conjugates of each other over $\mathbb{F}_q$. Therefore $\varsigma(h) = N_{E/\mathbb{F}_q}(\varsigma(h_i))$ for any $i \in [r]$. Therefore

$$N_{K/\mathbb{F}_q}(\varsigma(h_i)) = \left(N_{E/\mathbb{F}_q}(h_i)\right)^{k/r} = \varsigma(h)^{k/r} \quad \forall\, i \in [r]$$

On the other hand we have $\varrho(h) = \sum_{i=1}^{r} \varsigma(h_i)$ for any $i \in [r]$. Therefore $\varrho(h) = \mathrm{Tr}_{E/\mathbb{F}_q}$. Thus

$$\mathrm{Tr}_{K/\mathbb{F}_q}(\varsigma(h_i)) = \frac{k}{r}\,\mathrm{Tr}_{E/\mathbb{F}_q}(\varrho(h_i)) = \frac{k}{r}\varrho(r)$$

So together we have

$$\xi^{(k)}(h_i) = \psi^{(k)} \circ \varsigma(h_i) \cdot \chi^{(k)} \circ \varrho(h) = \psi\left(N_{K/\mathbb{F}_q}(\varsigma(h_i))\right) \cdot \chi\left(\mathrm{Tr}_{K/\mathbb{F}_q}(\varrho(h_i))\right) = \psi^{k/r}(\varsigma(h)) \cdot \chi^{k/r}(\varsigma(h)) = \xi^{k/r}(h)$$

So we have the theorem.

■

Now we can write $L^{(k)}(s, \xi^{(k)})$ and $\overline{L}^{(k)}(z, \xi^{(k)})$ in terms of L-functions on the field $\mathbb{F}_q$ using ξ and factorize the lifted L-function as product of L-function on the base field.

Theorem 4.4.6 Factorisation of the Lifted L-function

Let $K = \mathbb{F}_{q^k}$ be a finite extension of $\mathbb{F}_q$ and let $\zeta_k = e^{2\pi i/k}$ be a primitive k-th root of unity. Then

$$L^{(k)}\left(s, \xi^{(k)}\right) = \prod_{t=1}^{k} L\left(s - \frac{2\pi i t}{k \log q}, \xi\right).$$

Proof: By Euler Product for $L(s, \chi)$ we have

$$L^{(k)}\left(s, \xi^{(k)}\right) = \prod_{l(X) \in \mathrm{Irr}(\Phi^{(k)})} \left(1 - \xi^{(k)}\mathfrak{n}^{-s}(k)(l)\right)^{-1}$$

Since h is an irreducible monic polynomial in $\mathbb{F}_q[X]$ of degree d it splits intro r polynomials over K where $r = \gcd(d, k)$ i.e. $h = h_1 \cdot h_r$. Now by Lemma 4.4.5

$$\xi^{(k)}(h_i)\mathfrak{n}^{-s}(k)(h_i) = (\xi(h)\mathfrak{n}^{-s}(h))^{k/r}$$

On the other hand, every monic irreducible polynomial $l(X) \in k[X]$ is the factor of an unique monic irreducible $h(X) \in \mathbb{F}_q[X]$. Therefore

$$L^{(k)}\left(s, \xi^{(k)}\right) = \prod_{h \in \mathrm{Irr}(\Phi)} \left(1 - (\xi(h)\mathfrak{n}^{-s}(h))^{\frac{k}{\gcd(k, \deg(h))}}\right)^{-\gcd(k, \deg(h))}$$

Then by Lemma A.4.1 we have

$$\begin{aligned}
L^{(k)}\left(s,\xi^{(k)}\right) &= \prod_{h\in \mathrm{Irr}(\Phi)}\left(1-(\xi(h)\mathfrak{n}^{-s}(h))^{\frac{k}{\gcd(k,\deg(h))}}\right)^{-\gcd(k,\deg(h))}\\
&= \prod_{h\in \mathrm{Irr}(\Phi)}\prod_{t=1}^{k}\left(1-e^{2\pi i\frac{t\cdot\deg(h)}{k}}\xi(h)\mathfrak{n}^{-s}(h)\right)^{-1}\\
&= \prod_{t=1}^{k}\prod_{h\in \mathrm{Irr}(\Phi)}\left(1-q^{2\pi i\frac{t\cdot\deg(h)}{k}\log_q e}\xi(h)\mathfrak{n}^{-s}(h)\right)^{-1}\\
&= \prod_{t=1}^{k}\prod_{h\in \mathrm{Irr}(\Phi)}\left(1-\xi(h)\mathfrak{n}^{-\left(s-2\pi i\frac{t}{k\log q}\right)}(h)\right)^{-1}\\
&= \prod_{t=0}^{k-1}L\left(s-\frac{2\pi i t}{k\log q},\xi\right)
\end{aligned}$$

So we have the theorem. ■

Now by Theorem 4.4.3, $L(s,\xi)$ is a polynomial. So there exist complex numbers $w_1,\ldots,w_l\in\mathbb{C}$ such that $L(s,\xi)=(1-w_1u)\cdots(1-w_lu)$ for some $l\in\mathbb{N}$ where $u=q^{-s}$. By Theorem 4.4.6, $\overline{L}^{(k)}\left(z,\xi^{(k)}\right)$ or $L^{(k)}\left(s,\xi^{(k)}\right)$ is also a polynomial (a product of finitely many polynomials). We will prove that the reciprocal roots of $L^{(k)}\left(s,\xi^{(k)}\right)$ are w_i^k for all $i\in[l]$.

Corollary 4.4.7 Reciprocal Roots of the Lifted L-function

If $L(s,\xi)=(1-w_1u)\cdots(1-w_lu)$ *with* $u=q^{-s}$ *then*

$$L^{(k)}\left(s,\xi^{(k)}\right)=(1-w_1^k\cdot u_k)\cdots(1-w_l^k\cdot u_k)$$

where $u_k=q^{-ks}$.

Proof: By Theorem 4.4.6 we have

$$L^{(k)}\left(s,\xi^{(k)}\right)=\prod_{t=1}^{k}L\left(s-\frac{2\pi i t}{k\log q},\xi\right)$$

Now for any $t\in[k]$ we have $q^{-\left(s-2\pi i\frac{t}{k\log q}\right)}=e^{2\pi i\frac{t}{k}}\cdot q^{-s}=e^{2\pi i\frac{t}{k}}\cdot u$. So we get

$$\begin{aligned}
L^{(k)}\left(s,\xi^{(k)}\right) &= \prod_{t=1}^{k}L\left(s-\frac{2\pi i t}{k\log q},\xi\right)\\
&= \prod_{t=1}^{k}\prod_{j=1}^{l}\left(1-w_je^{2\pi i\frac{t}{k}}u\right)\\
&= \prod_{j=1}^{l}\prod_{t=1}^{k}\left(1-w_je^{2\pi i\frac{t}{k}}u\right)\\
&= \prod_{j=1}^{l}\left(1-w_j^ku^k\right)\\
&= \prod_{j=1}^{l}\left(1-w_j^k\cdot u_k\right)
\end{aligned}$$

So we have the result. ■

So now we can give an analog of Theorem 4.4.4 for the lifted sum and give a relation with $W^{(k)}\left(\psi^{(k)}, \chi^{(k)}; f, g\right)$ with the coefficients of $L^{(k)}\left(s, \xi^{(k)}\right)$.

Corollary 4.4.8 Character Sum via Reciprocal Roots

Suppose either $\psi \in \boldsymbol{\mathcal{M}}_q$ is a nontrivial multiplicative character of exponent d and $Y^d - f(X)$ is absolutely irreducible. If $\chi \in \boldsymbol{\mathcal{X}}_q$ is a additive character either non-trivial or trivial and $f(X) = 1$ then

$$W^{(k)}\left(\psi^{(k)}, \chi^{(k)}; f, g\right) = -\sum_{i=1}^{n+m-1} w_i^k.$$

Proof: If we apply Theorem 4.4.4 to $\mathbb{F}_{q^k}$ instead of $\mathbb{F}_q$ then using Corollary 4.4.7 we have

$$L^{(k)}\left(s, \xi^{(k)}\right) = 1 + c_{k,1} \cdot u_k + \cdots + c_{k,n+m-1} \cdot u_k^{n+m-1} = \prod_{j=1}^{n+m-1} (1 - w_j^k \cdot u_k)$$

Thereore $c_{k,1} = W^{(k)}\left(\psi^{(k)}, \chi^{(k)}; f, g\right)$. And on the other hand we have that $c_{k,1} = -\sum_{j=1}^{n+m-1} w_j^k$. ∎

§ 4.5 Weil Bounds via the Lifting Method

In this section we will try to give bounds on the absolute value of deviation of number of zeros from the average number of zeros. And then we will give upper bounds on the absolute values of $W^{(k)}\left(\psi^{(k)}; f\right)$, $W^{(k)}\left(\chi^{(k)}; g\right)$ and $W^{(k)}\left(\psi^{(k)}, \chi^{(k)}; f, g\right)$. But here we will count the solutions or give the bounds on absolute value on lifted characters in a finite extension field $E = \mathbb{F}_{q^k}$ of $\mathbb{F}_q$. For any polynomial $P(X, Y) \in \mathbb{F}_q[X, Y]$ we denote the number of solutions of $P(X, Y) = 0$ in the field E by $Z_E(P(X, Y) = 0)$.

4.5.1 Bound for Multiplicative Character Sums

Let $\psi \in \boldsymbol{\mathcal{M}}_q$ be a non-trivial multiplicative character. And let $\chi_0 \in \boldsymbol{\mathcal{X}}_q$ be the trivial additive character. Let $f(X) \in \mathbb{F}_q[X]$ is a monic polynomial with precisely m distinct roots. Let $g(X) = 0$. Then $W\left(\psi, \chi_0; f, 0\right) = W\left(\psi; f\right)$ and $W^{(k)}\left(\psi^{(k)}, \chi_0^{(k)}; f, 0\right) = W^{(k)}\left(\psi^{(k)}; f\right)$. So by Corollary 4.4.7 there exists complex numbers $w_1, \ldots, w_{m-1} \in \mathbb{C}$ such that

$$W\left(\psi; f\right) = \sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha)) = -\sum_{j=1}^{m-1} w_j, \qquad W^{(k)}\left(\psi^{(k)}; f\right) = \sum_{\alpha \in E} \psi^{(k)}(f(\alpha)) = -\sum_{j=1}^{m-1} w_j^k \tag{4.14}$$

Suppose $\psi \in \boldsymbol{\mathcal{M}}_q^{(d)}$ where $d > 1$. Then there are d characters of exponent d. For each such non-trivial character ψ we define complex numbers $w_{\psi,j} \in \mathbb{C}$ for all $j \in [m-1]$ such that

$$W\left(\psi; f\right) = -\sum_{j=1}^{m-1} w_{\psi,j}, \qquad W^{(k)}\left(\psi^{(k)}; f\right) = -\sum_{j=1}^{m-1} w_{\psi,j}^k$$

So taking the sum over all non-trivial $\psi \in \boldsymbol{\mathcal{M}}_q^{(d)}$ we get

$$\sum_{\psi \in \boldsymbol{\mathcal{M}}_q^{(d)}, \psi \neq \psi_0} W^{(k)}\left(\psi^{(k)}; f\right) = -\sum_{\psi \in \boldsymbol{\mathcal{M}}_q^{(d)}, \psi \neq \psi_0} \sum_{j=1}^{m-1} w_{\psi,j}^k$$

On the other hand if ψ is trivial then we have $W^{(k)}\left(\psi_0^{(k)}; f\right) = q^k$. Then we have the following lemma

Lemma 4.5.1 Counting Solutions of $\beta^d = \alpha$ via Additive Characters

For a given $\alpha \in E$ the number of $\beta \in E$ with $\beta^d = \alpha$ equals

$$\sum_{\psi \in \boldsymbol{M}_q^{(d)}} \psi^{(k)}(\alpha) = \sum_{\psi \in \boldsymbol{M}_q^{(d)}} \psi(N_{E/\mathbb{F}_q}(\alpha)).$$

Proof: Since $\alpha \to N_{E/\mathbb{F}_q}(\alpha)$ is a group homomorphism from $E^* \to \mathbb{F}_q^*$ which is an onto homomorphism. Now for any $\gamma \in \mathbb{F}_q^*$ the number of $\beta \in E^*$ such that $N_{E/\mathbb{F}_q}(\beta) = \gamma$ is exactly $1 + q + \cdots + q^{k-1} = {}^{E^*}/_{|\mathbb{F}_q^*|}$. Then restriction $N_{E/\mathbb{F}_q}\big|_{E^{*(d)}} : E^{*(d)} \to \mathbb{F}_q^{*(d)}$ is also onto. So for any $\alpha \in E$, $N_{E/\mathbb{F}_q}(\alpha) \in \mathbb{F}_q^{*(d)} \iff \alpha \in E^{*(d)}$.

Now by Theorem 2.2.4 we have $\sum_{\psi \in \boldsymbol{M}_q^{(d)}} \psi^{(k)}(\alpha) \in \{0, 1, d\}$. If $\alpha = 0$ then the sum is 1 and otherwise if $N_{E/\mathbb{F}_q} \in \mathbb{F}_q^{*(d)}$ then the sum is d and if $N_{E/\mathbb{F}_q} \notin \mathbb{F}_q^{*(d)}$ then the sum is 0.

(i) If $\alpha = 0$ and $\alpha \notin E^{*(d)}$ then there is single solution for $\beta = 0$ such that $\beta^d = \alpha$.
(ii) If $\alpha \neq 0$ and $\alpha \notin E^{*(d)}$ then there are no solution for β such that $\beta^d = \alpha$.
(iii) If $\alpha \in E^{*(d)}$ then there are d solutions for $\beta \in E$ such that $\beta^d = \alpha$.

So the sum in the lemma statement exactly gives the number of solutions. ∎

Now we apply the above lemma for the solutions of $Y^d - f(X)$ and count the number of solutions using multiplicative character sums. So we take the above sum over all possible $\alpha \in E$, and we get the following lemma:

Lemma 4.5.2 Solutions of $Y^d = f(X)$ via Character Sums

Suppose $f(X) \in \mathbb{F}_q[X]$ has precisely m distinct roots. Then

$$Z_E\left(Y^d - f(X)\right) = \sum_{\psi \in \boldsymbol{M}_q^{(d)}} \sum_{\alpha \in \mathbb{F}_q} \psi^{(k)}(f(\alpha)) = \sum_{\psi \in \boldsymbol{M}_q^{(d)}} W^{(k)}\left(\psi^{(k)}; f\right).$$

We prove the Weil bound for polynomials like $Y^d - f(X)$ in three steps below. And we also show bound of the absolute value of the complex numbers $w_{\psi,j}$ for all non-trivial $\psi \in \boldsymbol{M}_q^{(d)}$ and $j \in [m-1]$.

Lemma 4.5.3 Reciprocal Roots of $W^{(k)}\left(\psi^{(k)}; f\right)$ have Modulus $\leq q^{1/2}$

Let $f(X) \in \mathbb{F}_q[X]$ be a polynomial with precisely m distinct roots, and let $\psi \in \boldsymbol{M}_q$ be a non-trivial multiplicative character of exponent d. If $Y^d - f(X)$ is absolutely irreducible or $f(X)$ is not a d^{th} power, then for every $j \in [m-1]$ we have $|w_{\psi,j}| \leq q^{1/2}$.

Proof: We can assume f to be monic since $\psi(a \cdot f(\alpha)) = \psi(a) \cdot \psi(f(\alpha)$ and hence ψ evaluated on the leading coefficient gets out of the sum which doesn't change the absolute value. So we will assume f is monic. Now by Stepanov's Theorem, $|Z_E(Y^d = f(X)) - q^k| = O(q^{k/2})$. Since $W^{(k)}\left(\psi_0^{(k)}; f\right) = q^k$, Lemma 4.5.2 gives

$$\left| \sum_{\psi \in \boldsymbol{M}_q^{(d)}, \psi \neq \psi_0} W^{(k)}\left(\psi^{(k)}; f\right) \right| = O(q^{k/2}).$$

Therefore, $|W^{(k)}\left(\psi^{(k)}; f\right)| = O(q^{k/2})$ for each non-trivial ψ. Since $W^{(k)}\left(\psi^{(k)}; f\right) = -\sum_{j=1}^{m-1} w_{\psi,j}^k$, Lemma B.1 gives $|w_{\psi,j}| \leq q^{1/2}$ for all j. ∎

Now by Lemma 4.5.3, $|w_{\psi,j}| \leq q^{1/2}$ for all $j \in [m-1]$. Hence, $|W(\psi; f)| \leq (m-1)q^{1/2}$ and $|W^{(k)}\left(\psi^{(k)}; f\right)| \leq (m-1)q^{k/2}$. So we have the following theorem.

Theorem 4.5.4 Weil Bound for $W^{(k)}\left(\psi^{(k)}; f\right)$: Absolutely Irreducible Case

Let $\psi \in \boldsymbol{\mathcal{M}}_q$ be a non-trivial multiplicative character of exponent d. Suppose $f(X) \in \mathbb{F}_q[X]$ has precisely m distinct roots and $Y^d - f(X)$ is absolutely irreducible. Then

$$|W(\psi; f)| \leq (m-1)q^{1/2} \quad \text{and} \quad \left|W^{(k)}\left(\psi^{(k)}; f\right)\right| \leq (m-1)q^{k/2}.$$

We can give the same bound for the case where f is not a d^{th} power instead of the absolute irreducibility of $Y^d - f(X)$.

Theorem 4.5.5 Weil Bound for $W^{(k)}\left(\psi^{(k)}; f\right)$: Non-$d$-th-Power Case

Let $\psi \in \boldsymbol{\mathcal{M}}_q$ be a non-trivial multiplicative character of exponent d. Suppose $f(X) \in \mathbb{F}_q[X]$ has precisely m distinct roots and $f(X)$ is not a d-th power in $\mathbb{F}_q[X]$. Then

$$|W(\psi; f)| \leq (m-1)q^{1/2} \quad \text{and} \quad \left|W^{(k)}\left(\psi^{(k)}; f\right)\right| \leq (m-1)q^{k/2}.$$

Proof: Let $f(X) = a(X-\alpha_1)^{k_1} \cdots (X-\alpha_m)^{k_m}$ where $\alpha_1, \ldots, \alpha_m \in \overline{\mathbb{F}_q}$ are the m distinct roots of f. Let $e = \gcd(d, k_1, \ldots, k_m)$. Since f is not a d-th power in $\mathbb{F}_q[X]$, by Lemma 4.1.8 we have $e < d$, i.e., e is a proper divisor of d.

Define the polynomial $h(X) = \prod_{i=1}^{m}(X-\alpha_i)^{k_i/e}$, so that $f(X) = a \cdot h(X)^e$. In $h(X)$ we have $\gcd\left({}^{d}/_{e}, {}^{k_1}/_{e}, \ldots, {}^{k_m}/_{e}\right) = 1$, so by Theorem 4.1.2 the polynomial $Y^{d/e} - h(X)$ is absolutely irreducible. The polynomial $h(X)$ has precisely the same m distinct roots $\alpha_1, \ldots, \alpha_m$ as f.

Since $f(\alpha) = a \cdot h(\alpha)^e$ for every $\alpha \in \mathbb{F}_q$, the character ψ satisfies $\psi^{(k)}(f(\alpha)) = \psi^{(k)}(a) \cdot \left(\psi^{(k)}\right)^e (h(\alpha))$. The character $\left(\psi^{(k)}\right)^e$ has exponent ${}^{d}/_{e}$, and $|\psi(a)| = 1$. Applying Theorem 4.5.4 to $\left(\psi^{(k)}\right)^e$ and $h(X)$ (which has m distinct roots and $Y^{d/e} - h(X)$ is absolutely irreducible) gives

$$\left|W^{(k)}\left(\psi^{(k)}; f\right)\right| = \left|\sum_{\alpha \in \mathbb{F}_q} \psi^{(k)}(a) \cdot \left(\psi^{(k)}\right)^e (h(\alpha))\right| = \left|W^{(k)}\left(\left(\psi^{(k)}\right)^e; h\right)\right| \leq (m-1)q^{1/2}.$$

■

Corollary 4.5.6 Bound on the Number of Solutions of $Y^d = f(X)$ over E

Let $\psi \in \boldsymbol{\mathcal{M}}_q$ be a non-trivial multiplicative character of exponent d. Suppose $f(X) \in \mathbb{F}_q[X]$ has precisely m distinct roots and either $Y^d - f(X)$ is absolutely irreducible, or $f(X)$ is not a d-th power. Then

$$\left|Z_E\left(Y^d = f(X)\right) - q^k\right| \leq (d-1)(m-1)q^{k/2}, \qquad \left|Z\left(Y^d = f(X)\right) - q\right| \leq (d-1)(m-1)q^{1/2}.$$

Proof: By Lemma 4.5.2, $Z_E(Y^d = f(X)) = q^k + \sum_{\psi \neq \psi_0} W^{(k)}\left(\psi^{(k)}; f\right)$. Since there are $d-1$ non-trivial characters in $\boldsymbol{\mathcal{M}}_q^{(d)}$ and Lemma 4.5.3 gives $|W^{(k)}\left(\psi^{(k)}; f\right)| \leq (m-1)q^{k/2}$ for each, the triangle inequality yields the bound. Similarly, we get $\left|Z\left(Y^d = f(X)\right) - q\right| \leq (d-1)(m-1)q^{1/2}$ ■

4.5.2 Bound for Additive Character Sums

Let $\chi \in \boldsymbol{\mathcal{X}}_q$ be a non-trivial additive character. Let $\psi_0 \in \boldsymbol{\mathcal{M}}_q$ be the trivial multiplicative character. Let $g \in \mathbb{F}_q$ be a polynomial of degree n. Take $f(X) = 1$. Then $W(\psi_0, \chi; 1, g) = W(\chi; g)$ and $W^{(k)}\left(\psi_0^{(k)}, \chi^{(k)}; 1, g\right) = W^{(k)}\left(\psi^{(k)}; f\right)$. So by

Corollary 4.4.7 there exists complex numbers $w_1, \ldots, w_{n-1}$ such that

$$W(\chi; g) = \sum_{\alpha \in \mathbb{F}_q} \chi(g(\alpha)) = -\sum_{j=1}^{n-1} w_j, \qquad W^{(k)}\left(\chi^{(k)}; g\right) = \sum_{\alpha \in E} \chi^{(k)}(g(\alpha)) = -\sum_{j=1}^{n-1} w_j^k \tag{4.15}$$

There are q additive characters $\chi \in \boldsymbol{\chi}_q$ of $\mathbb{F}_q$. For each non-trivial character χ we define complex numbers $w_{\chi,j} \in \mathbb{C}$ for all $j \in [n-1]$ such that

$$W(\chi; g) = -\sum_{j=1}^{n-1} w_{\chi,j}, \qquad W^{(k)}\left(\chi^{(k)}; g\right) = -\sum_{j=1}^{n-1} w_{\chi,j}^k$$

So if we take the sum over all non-trivial additive character we get

$$\sum_{\chi \in \boldsymbol{\chi}_q, \chi \neq \chi_0} W^{(k)}\left(\chi^{(k)}; g\right) = -\sum_{\chi \in \boldsymbol{\chi}_q, \chi \neq \chi_0} \sum_{j=1}^{n-1} w_{\chi,j}^k$$

On the other hand if $\chi = \chi_0$ i.e. if χ is trivial then $W^{(k)}\left(\chi^{(k)}; g\right) = q^k$. Then we have the following lemma

Lemma 4.5.7 Counting Solutions of $\beta^q - \beta = \alpha$ via Additive Characters

For a given $\alpha \in E$ the number of $\beta \in E$ with $\beta^q - \beta = \alpha$ equals

$$\sum_{\chi \in \boldsymbol{\chi}_q} \chi^{(k)}(\alpha) = \sum_{\chi \in \boldsymbol{\chi}_q} \chi\left(\mathrm{Tr}_{E/\mathbb{F}_q}(\alpha)\right).$$

Proof: Now we know if $\mathrm{Tr}_{E/\mathbb{F}_q}(\alpha) = 0$ then we have q solutions for $\beta \in E$ such that $\beta^q - \beta = \alpha$ and by Theorem 2.2.2(i) we have

$$\sum_{\chi \in \boldsymbol{\chi}_q} \chi^{(k)}(\alpha) = q$$

If $\mathrm{Tr}_{E/\mathbb{F}_q}(\alpha) \neq 0$ then there is not $\beta \in E$ such that $\beta^q - \beta = \alpha$. And by Theorem 2.2.2(i) we get $\sum_{\chi \in \boldsymbol{\chi}_q} \chi^{(k)}(\alpha) = 0$. So we have the lemma. ∎

Now we apply the above lemma for the solutions of $Y^q - Y - g(X)$ and count the number of solutions using multiplicative character sums. So we take the above sum over all possible $\alpha \in E$, and we get the following lemma:

Lemma 4.5.8 Solutions of $\beta^q - \beta = g(\alpha)$ via Character Sums

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character and $g(X) \in \mathbb{F}_q[X]$ of degree n. The number of $(\alpha, \beta) \in E^2$ with $\beta^q - \beta = g(\alpha)$ equals

$$\sum_{\chi \in \boldsymbol{\chi}_q} \sum_{\alpha \in \mathbb{F}_q} \chi^{(k)}(g(\alpha)) = \sum_{\chi \in \boldsymbol{\chi}_q} W^{(k)}\left(\chi^{(k)}; g\right).$$

Now we will prove the Weil bound for the polynomial $Y^q - Y - g(X)$ below. We will also show bound on the absolute value of the complex numbers $w_{\chi,j}$ for all non-trivial $\chi \in \boldsymbol{\chi}_q$ and $j \in [n-1]$.

Lemma 4.5.9 Reciprocal Roots of $W^{(k)}\left(\chi^{(k)}; g\right)$ have Modulus $\leq q^{1/2}$

Let $\chi \in \boldsymbol{\chi}_q$ be a non-trivial additive character and $g(X) \in \mathbb{F}_q[X]$ of degree n. Suppose that

(i) either $n < q$ and $\gcd(n, q) = 1$,

(ii) or $Y^q - Y - g(X)$ is absolutely irreducible.

Then for every $j \in [n-1]$ we have $|w_{\chi,j}| \leq q^{1/2}$.

Proof: We can assume $g(X)$ has constant term to be 0 since $\chi(g(\alpha)+a) = \chi(g(\alpha)) \cdot chi(a)$ and so χ evaluated on the constant term gets out of the sum, and it doesn't change the absolute value. So we will assume g has constant term to be 0. Now suppose we have

$$\left|Z(Y^q - y - g(X)) - q^k\right| = O(q^{k/2}) \tag{4.16}$$

Since $W^{(k)}\left(\chi_0^{(k)}; g\right) = q^k$, by Lemma 4.5.8 we get

$$\left|\sum_{\chi \in \boldsymbol{X}_q, \chi \neq \chi_0} W^{(k)}\left(\chi^{(k)}; g\right)\right| = O(q^{k/2})$$

So Lemma B.1 gives $|w_{\chi,j}| \leq q^{1/2}$.

(i) By Bombieri's Theorem we have the bound in (4.16). So the lemma is true for this case.

(ii) By Weil Bound Theorem the bound in (4.16) is true. So the lemma holds.

Therefore, we have the lemma. ■

Now by the above lemma we have $|w_{\chi,j}| \leq q^{1/2}$ for all non-trivial $\chi \boldsymbol{X}_q$ and for all $j \in [n-1]$. Therefore $|W(\chi; g)| \leq (n-1)q^{1/2}$ and $\left|W^{(k)}\left(\chi^{(k)}; g\right)\right| \leq (n-1)q^{k/2}$. So we have the following theorem.

Theorem 4.5.10 Weil Bound for $W^{(k)}\left(\chi^{(k)}; g\right)$

Let $\chi \in \boldsymbol{X}_q$ be a non-trivial additive character and $g(X) \in \mathbb{F}_q[X]$ of degree n. Suppose that

(i) either $n < q$ and $\gcd(n, q) = 1$,

(ii) or $Y^q - Y - g(X)$ is absolutely irreducible.

Then

$$|W(\chi; g)| \leq (n-1)q^{1/2} \quad \text{and} \quad \left|W^{(k)}\left(\chi^{(k)}; g\right)\right| \leq (n-1)q^{k/2}.$$

Corollary 4.5.11 Bound on the Number of Solutions of $Y^q - Y = g(X)$ over E

Let $\chi \in \boldsymbol{X}_q$ be a non-trivial additive character and $g(X) \in \mathbb{F}_q[X]$ of degree n. Suppose that

(i) either $n < q$ and $\gcd(n, q) = 1$,

(ii) or $Y^q - Y - g(X)$ is absolutely irreducible.

Then

$$\left|Z_E(Y^q - Y = g(X)) - q^k\right| \leq (q-1)(n-1)q^{k/2} \qquad |Z(Y^q - Y = g(X)) - q| \leq (q-1)(n-1)q^{1/2}.$$

Proof: Now by Lemma 4.5.9 we have $Z_E(Y^q - Y = g(X)) = q^k + \sum_{\chi \neq \chi_0} W^{(k)}\left(\chi^{(k)}; g\right)$. Since there are $q-1$ non-trivial characters in $\boldsymbol{X}_q$ and Theorem 4.5.10 we have $\left|Z_E(Y^q - Y = g(X)) - q^k\right| \leq (q-1)(d-1)q^{k/2}$ and $|Z(Y^q - Y = g(X)) - q| \leq (q-1)(d-1)q^{1/2}$. ■

4.5.3 Bound for Mixed Character Sums

Suppose $f, g \in \mathbb{F}_q[X]$ such that f is monic and f has precisely m distinct roots in $\overline{\mathbb{F}_q}$ and g has degree n with constant term 0. Let $\psi \in \boldsymbol{M}_q^{(d)}$ be non-trivial multiplicative character of exponent f and $\chi \in \boldsymbol{X}_q$ be a non-trivial additive character

of $\mathbb{F}_q$. So by Corollary 4.4.7 there exists complex numbers $w_1, \ldots, w_{n+m-1}$ such that

$$W(\psi, \chi; f, g) = \sum_{\alpha \in \mathbb{F}_q} \psi(f(\alpha))\chi(g(\alpha)) = -\sum_{j=1}^{n+m-1} w_j, \quad W^{(k)}\left(\psi^{(k)}, \chi^{(k)}; f, g\right) = \sum_{\alpha \in E} \psi^{(k)}(f(\alpha)) \cdot \chi^{(k)}(g(\alpha)) = -\sum_{j=1}^{n+m-1} w_j^k$$

Now for every non-trivial multiplicative character $\psi \in \boldsymbol{\mathcal{M}}_q^{(q)}$ or exponent d and every non-trivial additive character $\chi \in Xq$ we define the complex numbers $w_{\psi,\chi,j} \in \mathbb{C}$ for all $j \in [n+m-1]$ such that

$$W(\psi, \chi; f, g) = -\sum_{j=1}^{n+m-1} w_{\psi,\chi,j}, \qquad W^{(k)}\left(\psi^{(k)}, \chi^{(k)}; f, g\right) = -\sum_{j=1}^{n+m-1} w_{\psi,\chi,j}^k$$

On the other hand if at least one of ψ, χ becomes trivial then we have

(i) If ψ is trivial i.e. $\psi = \psi_0$ but χ is non-trivial then $W^{(k)}\left(\psi_0^{(k)}, \chi^{(k)}; f, g\right) = W^{(k)}\left(\chi^{(k)}; g\right)$ since for all α, $\psi_0(f(\alpha)) = 1$. and so by the equation (4.15) in subsection 4.5.2 we have

$$W^{(k)}\left(\psi_0^{(k)}, \chi^{(k)}; f, g\right) = \sum_{\alpha \in E} \chi^{(k)}(g(\alpha)) = -\sum_{j=1}^{n-1} w_{\chi,j}^k$$

and similarly we have $W(\psi_0, \chi; f, g) = W^{(k)}\left(\chi^{(k)}; g\right) = -\sum_{j=1}^{n-1} w_{\chi,j}$.

(ii) If χ is trivial i.e. $\chi = \chi_0$ but ψ is non-trivial then $W^{(k)}\left(\psi^{(k)}, \chi_0^{(k)}; f, g\right) = W^{(k)}\left(\psi^{(k)}; f\right)$ since for all α, $\chi_0(g(\alpha)) = 1$. and so by the equation (4.14) in subsection 4.5.1 we have

$$W^{(k)}\left(\psi^{(k)}, \chi_0^{(k)}; f, g\right) = \sum_{\alpha \in E} \psi^{(k)}(f(\alpha)) = -\sum_{j=1}^{n-1} w_{\psi,j}^k$$

and similarly we have $W\left(\psi, \chi_0; f, g\right) = W^{(k)}\left(\psi^{(k)}; f\right) = -\sum_{j=1}^{n-1} w_{\psi,j}$.

(iii) Now if both ψ, χ are trivial i.e. $\psi = \psi_0$ and $\chi = \chi_0$ then $W^{(k)}\left(\psi_0^{(k)}, \chi_0^{(k)}; f, g\right) = q^k$.

So now we will count for any $\alpha_1, \alpha_2, \in E$ the number of $(\beta, \gamma) \in E^2$ such that both

$$\beta^d = \alpha_1, \quad \text{and} \quad \gamma^q - \gamma = \alpha_2$$

is satisfied

Lemma 4.5.12 Joint Counting via Multiplicative and Additive Characters

For $\alpha_1, \alpha_2 \in E$ the number of $(\beta, \gamma) \in E^2$ with $\beta^d = \alpha_1$ and $\gamma^q - \gamma = \alpha_2$ equals

$$\sum_{\psi \in \boldsymbol{\mathcal{M}}_q^{(d)}} \sum_{\chi \in \boldsymbol{X}_q} \psi^{(k)}(\alpha_1)\, \chi^{(k)}(\alpha_2).$$

Proof: Now

$$|\{(\beta, \gamma) \in E^2 \mid \alpha_1 = \beta^d,\ \alpha_2 = \gamma^q - \gamma\}| = |\{\beta \in E \mid \alpha_1 = \beta^d\}| \cdot |\{\gamma \in E \mid \alpha_2 = \gamma^q - \gamma\}|$$

So therefore by Lemma 4.5.1 and Lemma 4.5.7 we have

$$|\{\beta \in E \mid \alpha_1 = \beta^d\}| = \sum_{\psi \in \boldsymbol{\mathcal{M}}_q^{(d)}} \psi^{(k)}(\alpha_1)$$

$$|\{\gamma \in E \mid \alpha_2 = \gamma^q - \gamma\}| = \sum_{\chi \in \boldsymbol{X}_q} \chi^{(k)}(\alpha_2)$$

Therefore we have

$$|\{(\beta,\gamma)\in E^2 \mid \alpha_1=\beta^d,\ \alpha_2=\gamma^q-\gamma\}| = \left(\sum_{\psi\in\boldsymbol{M}_q^{(d)}}\psi^{(k)}(\alpha_1)\right)\left(\sum_{\chi\in\boldsymbol{X}_q}\chi^{(k)}(\alpha_2)\right) = \sum_{\psi\in\boldsymbol{M}_q^{(d)}}\sum_{\chi\in\boldsymbol{X}_q}\psi^{(k)}(\alpha_1)\chi^{(k)}(\alpha_2)$$

So we have the lemma. ∎

Now we apply the above lemma for the solutions of $Y^d-f(X)$ and $Z^q-Z-g(X)$ and count the number of common solutions using multiplicative and additive character sums. So we take the sum over all possible $\alpha\in E$ for f and g and we get the following lemma:

Lemma 4.5.13 Solutions of $\beta^d=f(\alpha)$, $\gamma^q-\gamma=g(\alpha)$ via Character Sums

The number of $(\alpha,\beta,\gamma)\in E^3$ with $\beta^d=f(\alpha)$ and $\gamma^q-\gamma=g(\alpha)$ is

$$\sum_{\psi\in\boldsymbol{M}_q^{(d)}}\sum_{\chi\in\boldsymbol{X}_q}W^{(k)}\left(\psi^{(k)},\chi^{(k)};f,g\right)$$

So now we give a general Weil Bound for the number of common solutions of $Y^d-f(X)$ and $Z^q-Z-g(X)$. We will have to use a very non-trivial theorem Theorem 4.5.15 from algebraic geometry. We will assume that theorem. And then give a bound on the absolute value of mixed weil sum.

Theorem 4.5.14 Weil Bound for Mixed Character Sums

Let $\psi\in\boldsymbol{M}_q$ and $\chi\in\boldsymbol{X}_q$ be non-trivial multiplicative and additive characters with $\mathrm{ord}(\psi)=d$, $d\mid q-1$. Let $f(X)\in\mathbb{F}_q[X]$ have precisely m distinct roots and $g(X)\in\mathbb{F}_q[X]$ with $\deg(g)=n$. Suppose that

(i) either $\gcd(d,\deg(f))=\gcd(n,q)=1$,

(ii) or both $Y^d-f(X)$ and $Z^q-Z-g(X)$ are absolutely irreducible.

Then

$$\left|\sum_{\alpha\in\mathbb{F}_q}\psi(f(\alpha))\chi(g(\alpha))\right|\le(m+n-1)q^{1/2}.$$

Proof: Like in the previous two sections we can assume f to be monic and g has constant term zero since $\psi(af(\alpha))=\psi(a)\cdot\psi(f(\alpha))$ and $\chi(b+g(\alpha))=\chi(b)\cdot\chi(g(\alpha))$ and so ψ, χ evaluated a and b respectively gets out of the sum and it doesn't change the absolute value. So we will assume f monic and g has constant term zero. Now suppose we already have

$$\left|Z_E\left(Y^d=f(X),Z^q-Z=g(X)\right)-q^k\right|=O(q^{k/2})$$

We have $W^{(k)}\left(\psi_0^{(k)},\chi_0^{(k)};f,g\right)=q^k$ by Lemma 4.5.13 we have

$$\left|\sum_{\psi\in\boldsymbol{M}_q^{(d)},\psi\neq\psi_0}\sum_{\chi\in\boldsymbol{X}_q,\chi\neq\chi_0}W^{(k)}\left(\psi^{(k)},\chi^{(k)};f,g\right)+\sum_{\psi\in\boldsymbol{M}_q^{(d)},\psi\neq\psi_0}W^{(k)}\left(\psi^{(k)};f\right)+\sum_{\chi\in\boldsymbol{X}_q,\chi\neq\chi_0}W^{(k)}\left(\chi^{(k)};g\right)\right|=O(q^{k/2})$$

Now by Lemma 4.5.9 and Lemma 4.5.3, $|w_{\psi,i}|\le q^{1/2}$, $|w_{\chi,j}|\le q^{1/2}$ for all $i\in[m-1]$, $\psi\in\boldsymbol{M}_q^{(d)}$ where $\psi\neq\psi_0$ and $j\in[n-1]$, $\chi\in\boldsymbol{X}_q$ where $\chi\neq\chi_0$. Therefore by Lemma B.1 we have $|w_{\psi,\chi,j}|\le q^{1/2}$ for all $j\in[n+m-1]$ and $\psi\in\boldsymbol{M}_q^{(d)}$ with $\psi\neq\psi_0$ and $\chi\in\boldsymbol{X}_q$ with $\chi\neq\chi_0$. Therefore we get $\left|W^{(k)}\left(\psi^{(k)},\chi^{(k)};f,g\right)\right|\le(n+m-1)q^{k/2}$ and $|W(\psi,\chi;f,g)|\le(n+m-1)q^{1/2}$.

Now all we need to prove $\left|Z_E\left(Y^d = f(X), Z^q - Z = g(X)\right) - q^k\right| = O(q^{\sqrt{k}2})$. Now under the conditions of the theorem $Y^d = f(X)$ and $Z^q - Z = g(X)$ defines a special curve called absolute variety. So by a theorem from algebraic geometry Theorem 4.5.15 we get the bound. ■

The following theorem helps us to count number of solutions in something called absolute variety in algebraic geometry. The proof of this theorem is beyond our scope so we ommit the proof of the theorem.

Theorem 4.5.15

Let $\mathcal{A}$ be an absolute variety of dimension d defined over $\mathbb{F}_q$. Let $Z_k(\mathcal{A})$ be the number of points $(\alpha_1, \ldots, \alpha_n) \in \mathcal{A}$ with each coordinate in $\mathbb{F}_{q^k}$. Then as $k \to \infty$ we have

$$Z_k(\mathcal{A}) = q^{kd} + O(q^{k(d-1/2)})$$

4.5.4 General Weil Bounds

André Weil (1940) generalized the above bounds for any absolutely irreducible polynomial over a finite field and showed that the deviation of number of solutions from average number of solution in any finite extended field is in the order of square root of the size of the extended field. This estimate follows from the Riemann Hypothesis for curves over finite fields.

Theorem 4.5.16 Weil Bound Theorem

Let $f(X, Y) \in \mathbb{F}_q[X, Y]$ is an absolutely irreducible polynomial and of total degree $d > 0$. If $q > 250d^5$ then

$$|Z(f(X, Y) = 0) - q| < \sqrt{2}d^{5/2}q^{1/2}$$

The proof of this theorem use non-trivial amount of algebraic geometry and is covered in the Chapter 3 of [Sch76]. In fact Riemann Hypothesis gives a stronger estimate

$$|Z(f(X, Y) = 0) - q| \leq (d-1)(d-2)q^{1/2} + c(d)$$

for some constant $c(d)$ which depends on $d = \deg(f)$.

Chapter 5

Decoding from Character Evaluations

Character sums of polynomials over finite fields carry deep arithmetic information, and one of the most striking manifestations of this is that they give rise to good error-correcting codes. The mechanism is simple: two polynomials whose character evaluations are close must be nearly indistinguishable by the character, and the Weil bounds make this precise. For a multiplicative character ψ of $\mathbb{F}_q$ and two polynomials $f, g \in \mathbb{F}_q[X]$ of small degree, the Weil bound controls gives a lower bound on the Hamming distance between the codewords $(\psi(f(\alpha)))_{\alpha\in\mathbb{F}_q}$ and $(\psi(g(\alpha)))_{\alpha\in\mathbb{F}_q}$. The resulting codes turn out to have distance approaching $q/2$ while the message length grows polynomially in q.

Two families are of particular interest. For q odd, applying the quadratic residue character $\psi(\beta) = \beta^{(q-1)/2}$ to squarefree monic polynomials $g(X)$ of degree $d = o(\sqrt{q})$ defines a nonlinear code over $\{0, \pm 1\}$ of length q. For $q = 2^b$, taking $g(X)$ with only odd-degree monomials and forming $\mathrm{Tr} \circ g : \mathbb{F}_q \to \mathbb{F}_2$ gives the classical *dual BCH codes*, introduced independently by Bose–Ray-Chaudhuri [BRC60] and Hocquenghem [Hoc59]. These codes are linear over $\mathbb{F}_2$, have length q and dimension $\frac{d+1}{2} \log_2 q$, and their distance analysis relies entirely on the Weil bound for additive character sums. Dual BCH codes are remarkable: for small d they are *unmatched* codes in the sense that, without the Weil bounds, there is no known proof that binary codes of their parameters even exist.

Despite this theoretical richness, decoding these codes from a constant fraction of errors was a long-standing open problem. The essential difficulty is that each evaluation $\psi(g(\alpha))$ carries very little information about g. The quadratic character discards the magnitude and retains only whether $g(\alpha)$ is a quadratic residue. So the system of equations one must solve is both under determined and nonlinear. Even the zero-error interpolation problem, of recovering g exactly from the values $\psi(g(\alpha))$ for all $\alpha \in \mathbb{F}_q$, had no polynomial-time solution for non-constant d prior to the work presented here. Recently Swastik Kopparty in [Kop26] gave polynomial time algorithm to decode from $1/8^{th}$ of minimum distance fraction of errors.

The goal of this chapter is to present the first polynomial-time algorithms for decoding both code families from a constant fraction of errors, following Kopparty. The algorithms combine two classical ideas: the *Berlekamp–Welch* framework, which uses a high-multiplicity error-locator polynomial to zero out error positions and reduces decoding to a linear system, and the *Stepanov polynomial method*, the elementary approach to the Weil bounds that works by interpolating auxiliary polynomials vanishing with high multiplicity at points of interest. The fusion of these two ideas gives rise to what we call *pseudopolynomials*: high-degree polynomials all of whose Hasse derivatives, when evaluated on $\mathbb{F}_q$, agree with polynomials of much smaller degree. These objects turn out to be the correct algebraic framework for capturing the special structure of character evaluations of polynomials.

§ 5.1 Theory of Pseudopolynomials

5.1.1 Pseudopolynomials

Definition 5.1.1: Pseudoderivative

Let $F(X) \in \mathbb{F}_q[X]$. Define the ℓ-th pseudoderivative *of F, denoted $F_{\langle \ell \rangle}(X)$, to be the unique polynomial of degree at most $q-1$ whose evaluations on $\mathbb{F}_q$ agree with the evaluations of $F^{(\ell)}(X)$. Explicitly:*

$$F_{\langle \ell \rangle}(X) = \left(F^{(\ell)}(X) \bmod \Lambda(X)\right),$$

where $\Lambda(X) = X^q - X$.

Abusing notation, we will sometimes use $F_{\langle \ell \rangle}$ to denote the function $F_{\langle \ell \rangle} : \mathbb{F}_q \to \mathbb{F}_q$, defined by $F_{\langle \ell \rangle}(\alpha) = F^{(\ell)}(\alpha)$.

Definition 5.1.2: Pseudodegree

Let $F(X) \in \mathbb{F}_q[X]$. We define the pseudodegree *of F, denoted* $\mathrm{pdeg}(F)$*, to be:*

$$\max_{\ell \geq 0} \deg(F_{\langle \ell \rangle}).$$

*We say $F(X)$ is a k-*pseudopolynomial *if* $\mathrm{pdeg}(F) \leq k$*. When $\deg(F) = d$ we refer to F as a k-*pseudopolynomial of degree d.

By definition, we have same basic observations on how pdeg behaves on addition and multiplication of polynomials.

Observation 5.1. *Like normal degree of polynomials* pdeg *behave the similar way, but we have inequality instead of equality. For any $F, G \in \mathbb{F}_q[X]$*

- $\mathrm{pdeg}(F+G) \leq \max(\mathrm{pdeg}(F), \mathrm{pdeg}(G))$.
- $\mathrm{pdeg}(F \cdot G) \leq \mathrm{pdeg}(F) + \mathrm{pdeg}(G)$.

5.1.1.1 Algebraic Characterization

Lemma 5.1.1 Algebraic Characterization of Pseudodegree

Let $F(X) \in \mathbb{F}_q[X]$, and let $F(X) = \sum_{i \geq 0} F_i(X)\,\Lambda(X)^i$, with each $\deg(F_i) < q$, be the base-$\Lambda(X)$ expansion of $F(X)$. Then $\mathrm{pdeg}(F) = \max_{i \geq 0} \deg(F_i)$.

Proof: We expand $F^{(\ell)}$ by applying the product rule for Hasse derivatives to each term of the base-Λ expansion:

$$F^{(\ell)}(X) = \sum_{i \geq 0} \left(F_i \cdot \Lambda^i\right)^{(\ell)}(X) = \sum_{i \geq 0} \left(\sum_{\substack{\ell_0, \ell_1, \ldots, \ell_i \\ \sum \ell_j = \ell}} F_i^{(\ell_0)}(X) \cdot \prod_{j=1}^{i} \Lambda^{(\ell_j)}(X) \right).$$

When we reduce mod $\Lambda(X)$, recall that $\Lambda^{(0)} = \Lambda \equiv 0$, and $\Lambda^{(j)} = 0$ for $j \notin \{0, 1, q\}$. So the only terms surviving mod Λ

are those with $\ell_1, \ldots, \ell_i \in \{1, q\}$, which forces $\ell \geq i$ and $\ell_0 = \ell - (\ell_1 + \cdots + \ell_i) \in [0, \ell - i]$. Collecting these terms:

$$F^{(\ell)}(X) \bmod \Lambda(X) = \sum_{i=0}^{\ell} \sum_{\ell_0=0}^{\ell-i} F_i^{(\ell_0)}(X) \cdot \underbrace{\left(\sum_{\substack{\ell_1,\ldots,\ell_i \in \{1,q\} \\ \ell_1+\cdots+\ell_i = \ell-\ell_0}} \prod_{j=1}^{i} \Lambda^{(\ell_j)}(X) \right)}_{=: C_{i,\ell,\ell_0,q} \in \mathbb{Z}}$$

$$= (-1)^{\ell} F_{\ell}(X) + \sum_{i=0}^{\ell-1} \sum_{\ell_0=0}^{\ell-i} F_i^{(\ell_0)}(X) \cdot C_{i,\ell,\ell_0,q},$$

where the leading term isolates $i = \ell$, $\ell_0 = 0$, the unique case with $\ell_1 = \cdots = \ell_i = 1$, which contributes the factor $(-1)^{\ell}$.

Forward direction. When $\deg(F_i) \leq k$ for all i, each $F_i^{(\ell_0)}$ has degree at most k, so the entire expression $F^{(\ell)} \bmod \Lambda(X) = F_{\langle \ell \rangle}(X)$ has degree at most k. Since ℓ was arbitrary, $\operatorname{pdeg}(F) \leq k$.

Reverse direction. Rearranging the identity above gives

$$(-1)^{\ell} F_{\ell}(X) = F_{\langle \ell \rangle}(X) - \sum_{i=0}^{\ell-1} \sum_{\ell_0=0}^{\ell-i} F_i^{(\ell_0)}(X) \cdot C_{i,\ell,\ell_0,q}.$$

We argue by induction on ℓ. The base case $\ell = 0$ gives $F_0 = F_{\langle 0 \rangle}$ directly, so $\deg(F_0) \leq \operatorname{pdeg}(F)$. In the inductive step, assuming $\deg(F_i) \leq \operatorname{pdeg}(F)$ for all $i < \ell$, every term $F_i^{(\ell_0)}$ on the right has degree at most $\operatorname{pdeg}(F)$, so $\deg(F_\ell) \leq \operatorname{pdeg}(F)$ as well. Hence, $\max_{i \geq 0} \deg(F_i) \leq \operatorname{pdeg}(F)$. The two directions together give $\operatorname{pdeg}(F) = \max_{i \geq 0} \deg(F_i)$. ■

5.1.1.2 Multiplicities

Now the pseudopolynomials are special because even though we may have a very large degree pseudopolynomials the multiplicities of an irreducible factor is small. The following lemma gives a bound on the multiplicity modulo q.

Lemma 5.1.2 Multiplicity Bound for Pseudopolynomials

Let $k < q$, and let $F(X)$ be a nonzero k-pseudopolynomial of degree D over $\mathbb{F}_q$. For an irreducible polynomial $H(X) \in \mathbb{F}_q[X]$, let μ be the highest power of $H(X)$ that divides $F(X)$. Then:

$$\mu \mod q \in \left[0,\ k + \left\lfloor \frac{D}{q} \right\rfloor\right].$$

Proof: Pick a root $\alpha \in \overline{\mathbb{F}_q}$ of $H(X)$, so that $\mu = \operatorname{mult}(F, \alpha)$ and $F^{(\mu)}(\alpha) \neq 0$. By Lemma 5.1.1, the base-Λ expansion of F has the form $F(X) = \sum_{i=0}^{t-1} F_i(X)\Lambda(X)^i$ with $t = \lfloor D/q \rfloor + 1$ and $\deg(F_i) \leq k$.

It suffices to show $F^{(\ell)}(X) = 0$ for every ℓ with $\ell \bmod q \in [k + t, q - 1]$, since this would confine $\mu \bmod q$ to $[0, k + t - 1] = [0, k + \lfloor D/q \rfloor]$.

Expanding via the product rule:

$$F^{(\ell)}(X) = \sum_{i=0}^{t-1} \sum_{\substack{\ell_0,\ell_1,\ldots,\ell_i \\ \sum \ell_j = \ell}} F_i^{(\ell_0)}(X) \cdot \prod_{j=1}^{i} \Lambda^{(\ell_j)}(X).$$

A summand vanishes if $\ell_0 > k$, since $\deg(F_i) \leq k$ forces $F_i^{(\ell_0)} = 0$. It vanishes if some $\ell_j \notin \{0, 1, q\}$, since $\Lambda^{(\ell_j)} = 0$ for such ℓ_j. It also vanishes mod Λ if some $\ell_j = 0$, since that contributes a factor $\Lambda^{(0)} = \Lambda$. Thus, every nonzero summand must have $\ell_0 \in [0, k]$ and $\ell_1, \ldots, \ell_i \in \{1, q\}$.

When $\ell \bmod q \in [k+t, q-1]$, no such decomposition is possible: a sum of one integer from $[0,k]$ and at most $t-1$ integers from $\{1,q\}$ has residue mod q at most $k+(t-1) < k+t$. So every summand vanishes, giving $F^{(\ell)}(X) = 0$. ■

From this lemma we can get a relation of size of a set of zero and the degree of then interpolated pseudopolynomials which becomes zero at the set. For any $S \subseteq \mathbb{F}_q$, and any c, k, M satisfying

$$|S| < \frac{c}{M} \cdot k,$$

there exists a nonzero k-pseudopolynomial $F(X)$ of degree $D < cq$ such that for all $\alpha \in S$, $\mathrm{mult}(F, \alpha) \geq M$. This follows from dimension and constraint counting.

In the reverse direction, the following theorem shows that when q is prime, every nonzero k-pseudopolynomial of degree $D < cq$ has at most $2\frac{c}{M}k$ roots of multiplicity at least M.

Theorem 5.1.3 High Multiplicity Zeroes Bound, [Kop26]

Let q be prime, and let $F(X)$ be a nonzero k-pseudopolynomial of degree $D < cq$ over $\mathbb{F}_q$. Suppose M satisfies $c < M < q$. Then:

$$\left|\{\alpha \in \mathbb{F}_q \mid \mathrm{mult}(F, \alpha) \geq M\}\right| \leq \min\left(\frac{c}{M-c+1} \cdot k + c, \quad k\right).$$

We will not prove this theorem. The proof uses some arguments from [GK14] on wronskian ranks. The full proof of this theorem is in [Kop26, Section 5.7].

5.1.1.3 Codes From Pseudopolynomials

Now using Pseudopolynomials we can create codes similar to univariate multiplicity codes. Let $\Sigma = \mathbb{F}_q^M$. The codewords lie in $\Sigma^{\mathbb{F}_q}$: for each k-pseudopolynomial $F(X)$ of degree $D < cq$, the codeword $\mathrm{Enc}(F) : \mathbb{F}_q \to \Sigma^{\mathbb{F}_q}$ is defined:

$$\mathrm{Enc}(F)(\alpha) = \left(F^{(0)}(\alpha), F^{(1)}(\alpha), \ldots, F^{(M-1)}(\alpha)\right).$$

This code has cardinality $|\Sigma|^{ck/M}$, block length q, rate $R = \frac{c}{M} \cdot \frac{k}{q}$, and minimum distance at least

$$\left(1 - \min\left(\frac{k}{q}, \quad \frac{c}{M-c+1} \cdot \frac{k}{q} + \frac{c}{q}, \quad \frac{c}{M}\right)\right) \cdot q,$$

which is always at least $(1-2R) \cdot q$.

5.1.2 Twisted Pseudopolynomials

Definition 5.1.3: r-Twisted Pseudopolynomial

Suppose $r : \mathbb{F}_q \to \mathbb{F}_q$ is a function. Let $k < q$. We say $F(X) \in \mathbb{F}_q[X]$ is an r-twisted (h, M)-pseudopolynomial *if for all $\ell < M$, there is some $U_\ell(X) \in \mathbb{F}_q[X]$ of degree at most h such that:*

$$F_{\langle \ell \rangle} = r \cdot U_\ell.$$

The following lemma shows that any two r-twisted (h, M)-pseudopolynomials with degree at most cq are very closely related, provided h and c are small enough.

Lemma 5.1.4

Let c, h, M be parameters, with $c < M/2$. Let $r : \mathbb{F}_q \to \mathbb{F}_q$. Suppose $F(X), G(X) \in \mathbb{F}_q[X]$ with $\deg(F), \deg(G) < cq$ are both r-twisted (h, M)-pseudopolynomials and $k > \frac{M}{M-2c} \cdot (h+1)$. Then there exist nonzero k-pseudopolynomials $A(X), B(X)$, with $\deg(A), \deg(B) < Mq$, such that:

$$A(X) \cdot F(X) = B(X) \cdot G(X).$$

Proof: Now if $k \geq q$ then the result is immediate. So assume $k < q$. Since F, G both are r-twisted (h, M)-pseudopolynomial for each $\ell < M$, we have:

$$F_{\langle \ell \rangle} = r \cdot U_\ell, \qquad G_{\langle \ell \rangle} = r \cdot V_\ell,$$

where $U_\ell(X), V_\ell(X) \in \mathbb{F}_q[X]$ have degrees at most h.

The goal is to find A and B in the base-Λ form

$$A(X) = \sum_{i=0}^{M-c-1} A_i(X)\, \Lambda(X)^i, \qquad B(X) = \sum_{i=0}^{M-c-1} B_i(X)\, \Lambda(X)^i,$$

with $\deg(A_i), \deg(B_i) \leq k$, so that $AF = BG$. Treating the coefficients of the A_i and B_i as unknowns, their total count exceeds $2(M-c) \cdot k$.

The condition $AF = BG$ is captured by requiring $(A \cdot F)_{\langle \ell \rangle} = (B \cdot G)_{\langle \ell \rangle}$ for each $\ell < M$. Expanding each pseudo-derivative using the twisted conditions on F and G, these equalities can be packaged into a single identity in $\mathbb{F}_q[X, T]$:

$$\left(\sum_{\ell_1 < M} A_{\langle \ell_1 \rangle}(X) \cdot T^{\ell_1} \right) \cdot \left(\sum_{\ell_2 < M} U_{\ell_2}(X) \cdot T^{\ell_2} \right) = \left(\sum_{\ell_3 < M} B_{\langle \ell_3 \rangle}(X) \cdot T^{\ell_3} \right) \cdot \left(\sum_{\ell_4 < M} V_{\ell_4}(X) \cdot T^{\ell_4} \right) \pmod{T^M}.$$

Each of the M coefficient equalities (in $\mathbb{F}_q[X]$, of degree at most $h+k$) is a homogeneous linear condition on the unknowns, giving at most $(h+k+1) \cdot M$ constraints in total. Since by assumption

$$(h+k+1) \cdot M < 2k \cdot (M-c),$$

the number of constraints is strictly less than the number of unknowns, guaranteeing a nonzero solution.

It remains to verify that any such solution satisfies $AF = BG$ as a polynomial identity, not merely on $\mathbb{F}_q$. The construction ensures $(A \cdot F - B \cdot G)_{\langle \ell \rangle} = 0$ for every $\ell < M$, so every point of $\mathbb{F}_q$ is a root of $AF - BG$ of multiplicity at least M. But

$$\deg(A \cdot F - B \cdot G) < cq + (M-c)q = Mq,$$

so the polynomial $AF - BG$ has more roots (counting multiplicity) than its degree, forcing $AF - BG = 0$. ∎

The next lemma shows that if we have many r-twisted (h, M)-pseudopolynomials with degree at most cq and c, h small, then they are nontrivially related. The smallness requirement on c is weaker, but the deduced relation is also weaker.

Lemma 5.1.5

Let m, c, h, M be parameters, with $c < \frac{m-1}{m} \cdot M$. Let $r : \mathbb{F}_q \to \mathbb{F}_q$. Suppose $F_1(X), \ldots, F_m(X) \in \mathbb{F}_q[X]$ with $\deg(F_i) < cq$ are all r-twisted (h, M)-pseudopolynomials. Let

$$k > \frac{M}{(m-1)M - mc} \cdot (h+1).$$

Then there exist k-pseudopolynomials $A_1(X), \ldots, A_m(X)$ with $\deg(A_i) < Mq$ and not all zero, such that, $\sum_i A_i(X)\, F_i(X) = 0$.

The proof is exactly like that of Lemma 5.1.4, which is the $m = 2$ case. The exact same argument also gives robust version.

Theorem 5.1.6 Robust Version for Two

Let m, c, h, γ, M be parameters, with $c < \frac{1}{2}(1-2\gamma)M$. Let $r_1, r_2 : \mathbb{F}_q \to \mathbb{F}_q$, with: $\Delta(r_1, r_2) \leq \gamma q$. Suppose $F(X), G(X) \in \mathbb{F}_q[X]$ with $\deg(F), \deg(G) < cq$ are r_1-twisted and r_2-twisted (h, M)-pseudopolynomials respectively. Let

$$k > \frac{M}{(1-2\gamma)M - 2c} \cdot (h+1).$$

Then there exist k-pseudopolynomials $A(X), B(X)$, with $\deg(A), \deg(B) < Mq$, not all zero, such that, $A(X) \cdot F(X) = B(X) \cdot G(X)$.

Theorem 5.1.7 Robust Version for Many

Let m, c, h, γ, M be parameters, with $c < \frac{1}{m}((m-1) - m\gamma)M$. Let $r_1, \ldots, r_m : \mathbb{F}_q \to \mathbb{F}_q$, with:

$$\left|\{\alpha \in \mathbb{F}_q \mid r_i(\alpha) \neq r_j(\alpha) \text{ for some } i, j\}\right| \leq \gamma q.$$

Suppose $F_1(X), \ldots, F_m(X) \in \mathbb{F}_q[X]$ with $\deg(F_i) < cq$ are r_i-twisted (h, M)-pseudopolynomials. Let

$$k > \frac{M}{(m-1-m\gamma)M - mc} \cdot (h+1).$$

Then there exist k-pseudopolynomials $A_1(X), \ldots, A_m(X)$, with $\deg(A_i) < Mq$, not all zero, such that, $\sum_i A_i(X)\, F_i(X) = 0$.

§ 5.2 Decoding Multiplicative Character Evaluations

Let $\psi \in \boldsymbol{\chi}_q$ is a non-trivial multiplicative character of $\mathbb{F}_q$ where $\mathrm{ord}(\psi) = d$. Then the polynomial based code we can construct is $(\psi \circ f(\alpha))_{\alpha \in \mathbb{F}_q}$ where $f \in \mathbb{F}_q[X]$. Now suppose g is another polynomial. Then for any $\alpha \in \mathbb{F}_q$,

$$\psi((f \cdot g^d)(\alpha)) = \psi(f(\alpha)) \cdot \psi^d(g(\alpha)) = \psi(f(\alpha))$$

So we assume f doesn't have a power d factor. Similarly, for any α, β where $\beta \neq 0$

$$\psi((\beta^d \cdot f)(\alpha)) = \psi^d(\beta) \cdot \psi(f(\alpha)) = \psi(f(\alpha))$$

Since different leading coefficient can lead to same character evaluation we assume f to be monic.

5.2.1 Quadratic Character

Now we take the quadratic character $\eta \in \boldsymbol{\mathcal{M}}_q$. Hence, we assume the polynomials are monic and squarefree. Then by Theorem 4.5.5 we have

$$\left|\sum_{\alpha} \eta((g_1 \cdot g_2)(\alpha))\right| < O(d\sqrt{q})$$

Therefore the vectors $\eta \circ g_1$ and $\eta \circ g_2$ differs in at least $\left(q - O(d\sqrt{q})\right) = \left(\frac{1}{2} - \frac{d}{\sqrt{q}}\right) q$ many locations. So their hamming distance is at least $\left(\frac{1}{2} - \frac{d}{\sqrt{q}}\right) q$ which gives the distance of the code.

Below we give an algorithm to decode these codes from 1/4th of the minimum distance of the code. The decoding algorithm follows the *Berlekamp-Welch* of *Reed-Solomon codes* to some extent. Let g is the closest codeword. So suppose $G(X) \coloneqq g^{(q-1)/2}(X)$.

If there had been zero error in the received word then we know the polynomial $\tilde{G}$ is $\psi \circ g$-twisted $(O(dM), M)$-pseudopolynomial. But we have the received word very close to the original codeword $\psi \circ g$. So we will try to find a r-twisted pseudopolynomial F with good parameters such that we can use Lemma 5.1.4 or Theorem 5.1.6 to relate them. So first we have to show such F exists.

5.2.1.1 Constructing r-twisted Pseudopolynomial

Let $S = \{\alpha \in \mathbb{F}_q : \psi \circ g(\alpha) \neq r(\alpha)\}$ be the *error set*. Then consider the *error locator polynomial* $Z_S(X) = \prod_{\alpha \in S}(X - \alpha)^M$. Then all elements of S are roots of Z_S with multiplicity at M. So we take any multiple of Z_S, $E(X)$ which is a h-pseudopolynomial of degree $< cq$ where h, c, M are parameters to be chosen later. So by construction $E(X) = \sum_{i=0}^{c-1} E_i(X) \cdot \Lambda^i(X)$ by Lemma 5.1.1. So $\deg(E_i) \leq h$. Therefore the number of coefficients is $c(h+1)$. But number of constraints is $e \cdot M$. Hence we want $c(h+1) > e \cdot M$. So it is enough to have $ch = eM$.

So now we have a nonzero multiple of $Z_S(X)$ such that for all $\alpha \in S$, $\text{mult}(E, \alpha) \geq M$. Therefore we take the polynomial $G(X) \cdot E(X)$ where $G(X) = g^{\frac{q-1}{m}+M}(X)$. We will show that this F is a r-twisted $(h + dM, M)$-pseudopolynomial.

First we will calculate $E_{\langle \ell \rangle}$ for all $0 \leq \ell < M$. By Lemma D.2 we have

$$E^{(\ell)}(X) = \sum_{i=0}^{c-1} (-1)^i E_i^{(\ell-i)}(X) \bmod \Lambda(X) \implies E_{\langle \ell \rangle}(X) = \sum_{i=0}^{c-1} (-1)^i E_i^{(\ell-i)}(X)$$

Therefore, for any $\alpha \in \mathbb{F}_q$ we have

$$\begin{aligned}
(E \cdot G)^{(\ell)}(\alpha) &= \sum_{\ell_1+\ell_2=\ell} E^{(\ell_1)}(\alpha) \cdot G^{(\ell_2)}(\alpha) \\
&= \sum_{\ell_1+\ell_2=\ell} E_{\langle \ell_1 \rangle}(\alpha) \cdot G_{g,w,\ell_2}(\alpha) \cdot g^{M-\ell_2}(\alpha) \cdot (\psi \circ g(\alpha)) && \text{[By Theorem D.4]} \\
&= \psi \circ g(\alpha) \sum_{\ell_1+\ell_2=\ell} E_{\langle \ell_1 \rangle}(\alpha) \cdot G_{g,w,\ell_2}(\alpha) \cdot g^{M-\ell_2}(\alpha)
\end{aligned}$$

So define

$$V_\ell(X) = \sum_{\ell_1+\ell_2=\ell} E_{\langle \ell_1 \rangle}(X) \cdot G_{g,w,\ell_2}(X) \cdot g^{M-\ell_2}(X).$$

Then we have $(E \cdot G)^{(\ell)}(\alpha) = r(\alpha) \cdot V_\ell(\alpha)$. Since at error locations we have $E^{(i)}(\alpha)$ for all $0 \leq i < M$ we have $V_\ell(\alpha) = 0$. So this equality holds. Now $\deg(V_\ell) \leq d\left(\frac{q-1}{2} + M\right) cq =: D$ and $\deg(V_\ell) \leq h + dM$. So we have F to be a r-twisted $(h + dM, M)$-pseudopolynomial of degree D. So if we find any r-twisted $(h + dM, M)$-pseudopolynomial F of degree D via interpolation then the $E \cdot G$ and F, are both r-twisted. So by Lemma 5.1.4 there exists $A, B \in \mathbb{F}_q[X]$ such that $A \cdot (EG) = B \cdot F$. We can rewrite and say there exists $A, B \in \mathbb{F}_q[X]$ such that $A \cdot G = B \cdot F(X)$. So we have the following algorithm:

Algorithm 1: Decoding from $1/8$−th of distance of Quadratic Character Evaluation

Input: Noisy received word, $r : \mathbb{F}_q \to \{0, 1, -1\}$, with degree $d \leq O(\epsilon\sqrt{q})$ and error bound $e \leq (1/8 - \epsilon)\, q$

Output: Find the closest vector $\eta \circ f$ where $f \in \mathbb{F}_q[X]$ of degree at most d

begin

- Set $M = 16/\epsilon \cdot d$, $c = M/2$, $h = 2 \cdot e$, $D = d((q-1)/2 + M) + cq$
- Use system of linear equations to find r-twisted $(h + dM, M)$-pseudopolynomial $F(X)$ of degree at most D.
- Complete factorization of $F(X)$ into distinct monic irreducible factors $\{h_1, \ldots, h_k\}$ with $\text{mult}(h_j, F) = k_j$.
- Compute the set $J = \left\{j \in [t] : k_j \in [3q/8, 7q/8] \bmod q\right\}$. Then compute $f(X) = \prod_{j \in J} h_j$.
- **return** $f(X)$

5.2.1.2 Relating F, G and Factor Multiplicities

To make sense of the last step of separating out the high multiplicity factors we first have to look into A, B.

Step 1: Relating F and G We now show that any polynomial $F(X)$ obtained in the algorithm must satisfy a nontrivial algebraic relation with $G(X)$. Let $U_0, \ldots, U_{M-1}$ be the polynomials such that $F^{(\ell)}(\alpha) = r(\alpha) \cdot U_\ell(\alpha)$ for all $0 \le \ell < M$, $\alpha \in \mathbb{F}_q$. So now we set parameters $t = 3/8M$ and $k = e + 4dM$. We claim that there exist nonzero polynomials $A(X)$ and $B(X)$ of the form

$$A(X) = \sum_{i=0}^{t-1} A_i(X)\Lambda^i(X), \qquad B(X) = \sum_{i=0}^{t+c-1} B_i(X)\Lambda^i(X),$$

with $\deg(A_i) \le k$, $\deg(B_i) \le k + 2e$, such that $A(X)F(X) = B(X)G(X)$. We again argue by dimension counting. The total number of coefficients of the A_i and B_i equals $N = t(k+1) + (t+c)(k+2e+1)$. Substituting the values of t and k, we get

$$N > \frac{3}{8}M(e + 4dM) + \frac{7}{8}M(3e + 4dM) = 3Me + 5dM^2. \tag{5.1}$$

We regard these coefficients as variables and impose homogeneous linear conditions so that for every $0 \le \ell < M$ and every $\alpha \in \mathbb{F}_q \setminus S$,

$$(A \cdot F)^{(\ell)}(\alpha) = (B \cdot G)^{(\ell)}(\alpha). \tag{5.2}$$

Consequently, the polynomial $A(X)F(X) - B(X)G(X)$ vanishes to multiplicity at least M at every point of $\mathbb{F}_q \setminus S$. Hence it has at least

$$M \cdot |\mathbb{F}_q \setminus S| \ge M(q - e) > \left(\frac{7}{8} + \varepsilon\right) Mq$$

roots counted with multiplicity. On the other hand,

$$\deg(A \cdot F - B \cdot G) \le \max\{\deg(A \cdot F), \deg(B \cdot G)\} < \left(\frac{7}{8} + \frac{\varepsilon}{2}\right) Mq.$$

Thus, $A(X)F(X) = B(X)G(X)$. It remains to show that Equation (5.2) can indeed be enforced using fewer than N linear constraints. For each ℓ, define $A_{\langle \ell \rangle}(X) = \sum_{i=0}^{t-1} (-1)^i A_i^{(\ell-i)}(X)$, $B_{\langle \ell \rangle}(X) = \sum_{i=0}^{t+c-1} (-1)^i B_i^{(\ell-i)}(X)$. By Lemma D.2 we get $A^{(\ell)}(\alpha) = A_{\langle \ell \rangle}(\alpha)$ and $B^{(\ell)}(\alpha) = B_{\langle \ell \rangle}(\alpha)$ for every $\alpha \in \mathbb{F}_q$. Now recall that $F^{(\ell)}(\alpha) = r(\alpha)U_\ell(\alpha)$. Hence,

$$(A \cdot F)^{(\ell)}(\alpha) = \sum_{\ell_1 + \ell_2 = \ell} A^{(\ell_1)}(\alpha) F^{(\ell_2)}(\alpha) \implies (A \cdot F)^{(\ell)}(\alpha) = r(\alpha) \cdot \sum_{\ell_1 + \ell_2 = \ell} A_{\langle \ell_1 \rangle}(\alpha) U_{\ell_2}(\alpha). \tag{5.3}$$

Similarly, we get

$$G^{(\ell)}(\alpha) = (\psi \circ g)(\alpha)\, H_{g,w,\ell}(\alpha)\, g^{M-\ell}(\alpha), \implies (B \cdot G)^{(\ell)}(\alpha) = (\psi \circ g)(\alpha) \cdot \sum_{\ell_1 + \ell_2 = \ell} B_{\langle \ell_1 \rangle}(\alpha) G_{g,w,\ell_2}(\alpha) g^{M-\ell_2}(\alpha). \tag{5.4}$$

We therefore impose, for each $0 \le \ell < M$, the polynomial identity

$$\sum_{\ell_1 + \ell_2 = \ell} A_{\langle \ell_1 \rangle}(X) U_{\ell_2}(X) = \sum_{\ell_1 + \ell_2 = \ell} B_{\langle \ell_1 \rangle}(X) G_{g,w,\ell_2}(X) g^{M-\ell_2}(X). \tag{5.5}$$

Both sides have degree at most $3e + dM$. Hence, the total number of linear conditions arising from all these identities is at most $M(3e + dM + 1)$, which is strictly smaller than N by Equation (5.1). Therefore there exists a nontrivial choice of $A(X), B(X)$ satisfying Equation (5.5). Finally, combining Equations (5.3), (5.4), and (5.5), and using the fact that $r(\alpha) = (\psi \circ g)(\alpha)$ for all $\alpha \in \mathbb{F}_q \setminus S$, we obtain Equation (5.2), and hence $A(X)F(X) = B(X)G(X)$.

Step 2: Recovering the factors of $g(X)$ We now use the identity $A(X)F(X) = B(X)G(X)$ to compare the irreducible factors of $F(X)$ and $G(X)$. The essential point is that the special form of $A(X)$ and $B(X)$ forces all their irreducible factors to occur with multiplicity very close to a multiple of q by Lemma 5.1.2. In contrast, the factors arising from $g(X)$ occur inside $G(X)$ with multiplicity approximately $q/2$ modulo q. Since $g(X)$ is squarefree, this allows us to identify precisely the irreducible factors of $g(X)$.

Now let $H(X)$ be an irreducible polynomial in $\mathbb{F}_q[X]$. Denote by $\mu_F, \mu_A, \mu_B, \mu_G, \mu_g$ the multiplicities of $H(X)$ in $F(X), A(X), B(X), G(X), g(X)$ respectively. Since $g(X)$ is squarefree, $\mu_g \in \{0, 1\}$. From the identity $AF = BG$, we obtain $\mu_F = \mu_G + \mu_B - \mu_A$. Applying the lemma to A and B, we have

$$\mu_B \in q\mathbb{Z} + [0, 3e + 4dM + \tfrac{7}{8}M] \quad \text{and} \quad \mu_A \in q\mathbb{Z} + [0, e + 4dM + \tfrac{3}{8}M].$$

Therefore $\mu_F \in \mu_G + q\mathbb{Z} + [-e - 5dM,\ 3e + 5dM]$. Now $\mu_G = \left(\frac{q-1}{2} + M\right)\mu_g$.

If $\mu_g = 1$, then $\mu_G = \frac{q}{2} + \frac{M-1}{2}$, and hence $\mu_F \in q\mathbb{Z} + \frac{q}{2} + [-e - 5dM,\ 3e + 6dM]$. Using $e \le \left(\frac{1}{8} - \varepsilon\right) q$ and $6dM < \varepsilon q$, we obtain $\mu_F \in q\mathbb{Z} + (3q/8, 7q/8)$. On the other hand, if $\mu_g = 0$, then $\mu_G = 0$, and thus

$$\mu_F \in q\mathbb{Z} + [-e - 5dM,\ 3e + 5dM] \subseteq q\mathbb{Z} + (-q/8, 3q/8).$$

Hence, the set J selected in Step 5 of the algorithm consists precisely of those irreducible factors which divide $g(X)$. Therefore, the polynomial produced in Step 5 is exactly $f(X) = g(X)$.

5.2.2 General Multiplicative Character of order m

Now we take the multiplicative character $\psi \in \boldsymbol{\mathcal{M}}_q$or order $m > 1$. Hence, we assume the polynomials are monic and don't have a factor of power m. Then by Theorem 4.5.5 we have $|\sum_\alpha \psi((g_1 \cdot g_2)(\alpha))| < O(m\sqrt{q})$. Therefore the hamming distance is at least $\left(1 - \frac{1}{m}\right) q$ which gives the distance of the code. So the unique decoding radius is $\frac{1}{2}\left(1 - \frac{1}{m}\right) q$.

Now we can also think of the character as $\psi : \mathbb{F}_q \to \{0\} \cup \mathbb{F}_q^{*^{((q-1)/m)}}$ where $\alpha \mapsto \alpha^{(q-1)/m}$. So here we will consider the multiplicative as an map from $\mathbb{F}_q \to \{0\} \cup \mathbb{F}_q^{*^{((q-1)/m)}}$. Like for the case of quadratic character we will not prove the correctness of the algorithm instead we just give the pseudocode of the algorithm below. The details of the correctness of the algorithm is given in [Kop26].

5.2.3 Alternate Proof Weil Bound on Multiplicative Character using Pseudopolynomials

The pseudopolynomial framework developed in §5.1 yields clean alternative proofs of the Weil bounds for multiplicative character sums. The key input is Lemma 5.1.2, which constrains the multiplicity of any irreducible factor of a pseudopolynomial modulo q.

Weil bound for the quadratic residue character. We prove Theorem 4.5.5 for η. Let $f(X), g(X) \in \mathbb{F}_q[X]$ be distinct monic squarefree polynomials of degree at most $d \le O(\epsilon\sqrt{q})$. We claim

$$\Delta(\eta \circ f,\ \eta \circ g) \ \ge\ \left(\tfrac{1}{2} - \epsilon\right) q.$$

Set $M = \frac{2}{\epsilon} d$ and $c = d$, and define

$$F(X) = f(X)^{\frac{q-1}{2} + M}, \qquad G(X) = g(X)^{\frac{q-1}{2} + M}.$$

Since $f^{(q-1)/2}(\alpha) = \eta(f(\alpha))$ for every $\alpha \in \mathbb{F}_q$, F is an $(\eta \circ f)$-twisted (dM, M)-pseudopolynomial and G is an $(\eta \circ g)$-twisted (dM, M)-pseudopolynomial, both with $h = dM$.

Suppose for contradiction that $\Delta(\eta \circ f, \eta \circ g) \le \gamma q$ for $\gamma = \frac{1}{2} - \epsilon$. With

$$k \ =\ \frac{1}{\epsilon} \cdot (h + 1) \ \le\ O\left(\frac{d^2}{\epsilon^2}\right) \ <\ \frac{q}{4},$$

Algorithm 2: Decoding from $1/12$−th of distance of m^{th}-order Character Evaluation

Input: Noisy received word, $r : \mathbb{F}_q \to \{0, 1, -1\}$, with degree $d \le O(\frac{\epsilon}{m}\sqrt{q})$ and error bound $e \le (1/12 - \epsilon)\, q$

Output: Find the closest vector $\psi \circ f$ where $f \in \mathbb{F}_q[X]$ of degree at most d

begin

  Set $M = 16/\epsilon \cdot dm$, $c = M/2$, $h = 2 \cdot e$, $D = d((q-1) \cdot (m-1)/m + M) + cq$

  for $k \in [m-1]$ **do**

    Use system of linear equations to find r^k-twisted $(h + (m-1)dM, M)$-pseudopolynomial $F_k(X)$ of degree at most D.

    Complete factorization of $F(X)$ into distinct monic irreducible factors $\{h_{1,k}, \ldots, h_{t_k,k}\}$ with $\mathrm{mult}(h_{j,k}, F_k) = d_{j,k}$.

  Let the irreducible factors are $H_1, \ldots, H_t$ with $\mathrm{mult}(H_j, F_k) = d_{j,k}$.

  for $j \in [t]$ **do**

    Let $\mu_j \in \{0, 1, \ldots, m-1\}$ be the unique number such the for all $k \in [m-1]$

$$d_{k,j} - \frac{k \cdot \mu_j}{m} \cdot (q-1) \in \left(-\left(\frac{1}{12} - \epsilon\right) q, \left(\frac{3}{12} - \epsilon\right) q\right) \bmod q$$

  Compute $f(X) = \prod_{j=1}^{t} h_j^{\mu_j}(X)$.

  return $f(X)$

Theorem 5.1.6 produces nonzero k-pseudopolynomials $A(X), B(X)$ of degree less than Mq satisfying $A(X) \cdot F(X) = B(X) \cdot G(X)$.

Let $H(X)$ be an irreducible polynomial that appears with distinct factor multiplicity in $f(X)$ and $g(X)$. Such H exists since $f \neq g$ are squarefree. Without loss of generality $H \mid f$ but $H \nmid g$, so $\mathrm{mult}(H, F) \equiv \frac{q-1}{2} + M \pmod q$ while $\mathrm{mult}(H, G) = 0$. From $AF = BG$:

$$\mathrm{mult}(H, A) + \frac{q-1}{2} + M \;\equiv\; \mathrm{mult}(H, B) \pmod q.$$

By Lemma 5.1.2, both $\mathrm{mult}(H, A) \bmod q$ and $\mathrm{mult}(H, B) \bmod q$ lie in $[0, k + O(d)] \subseteq [0, q/4 + O(d)]$. But the identity forces $\mathrm{mult}(H, B)$ to sit $\approx q/2$ away from $\mathrm{mult}(H, A)$ modulo q, which is outside this range — a contradiction.

Weil bound for the m-th power residue character. We prove the Weil bound for ψ with $\mathrm{ord}(\psi) = m$, m prime. Suppose $f, g \in \mathbb{F}_q[X]$ are distinct monic polynomials of degree $d \le O_m(\epsilon\sqrt{q})$, each irreducible factor appearing with multiplicity in $\{1, 2, \ldots, m-1\}$. We claim

$$\Delta(\psi \circ f,\, \psi \circ g) \;\ge\; \left(1 - \tfrac{1}{m} - \epsilon\right) q.$$

Set $M = O_m(d/\epsilon)$, $c = O_m(d)$, $F(X) = f(X)^{(q-1)/m+M}$, $G(X) = g(X)^{(q-1)/m+M}$. For $i \in [m]$ let

$$r_i(\alpha) = \left(\psi^{i-1} \circ f(\alpha)\right) \cdot \left(\psi^{m-i} \circ g(\alpha)\right), \qquad F_i(X) = F(X)^{i-1} \cdot G(X)^{m-i}.$$

Each F_i is an r_i-twisted $(O_m(dM), M)$-pseudopolynomial.

Suppose $\Delta(\psi \circ f, \psi \circ g) \le \gamma q$ for $\gamma = 1 - \frac{1}{m} - \epsilon$, so $|\{\alpha \mid r_i(\alpha) \neq r_j(\alpha) \text{ for some } i, j\}| \le \gamma q$. With $h = O_m(dM)$ and

$$k = \frac{1}{\epsilon} \cdot (h+1) \le O\left(\frac{d^2}{\epsilon^2}\right) < \frac{q}{2m},$$

Theorem 5.1.7 gives k-pseudopolynomials $A_1(X), \ldots, A_m(X)$, not all zero, with $\deg(A_i) < Mq$ and

$$\sum_{i=1}^{m} A_i(X) \cdot F(X)^{i-1} \cdot G(X)^{m-i} = 0.$$

Let $H(X)$ be an irreducible polynomial that appears with distinct factor multiplicity in $f(X)$ and $g(X)$. By Lemma 5.1.2, all the nonzero terms in the above sum have distinct factor multiplicity of $H(X)$ modulo q (because the contribution $(i-1)\ \mathrm{mult}(H,F) + (m-i)\ \mathrm{mult}(H,G)$ takes m pairwise distinct values mod q as i varies, and $k < q/(2m)$ prevents the A_i terms from closing the gap). Therefore the sum cannot be zero, which is a contradiction.

§ 5.3 Decoding Dual BCH Codes (Additive Character Evaluations)

In this section we basically decode the famous Dual BCH codes. We first discuss the code construction. Let q be a power of 2, say $q = 2^b$. Let Tr be the absolute trace function from $\mathbb{F}_q \to \mathbb{F}_2$. Let $\mathrm{Tr}(X)$ be the polynomial $\mathrm{Tr}(X) = \sum_{i=0}^{b-1} X^{2^i}$. Now suppose $\deg(g) = d$. Then we can write g as

$$g(X) = a + \tilde{g}(X) + h(X) + h^2(X)$$

where $a \in \mathbb{F}_q, \tilde{g}, h \in \mathbb{F}_q[X]$ which satisfies the following:

(i) $\tilde{g}$ has only odd degree monomials

(ii) $h(X)$ has 0 constant term

Now for any $\alpha \in \mathbb{F}_q$, we have $\mathrm{Tr} \circ g(\alpha) = \mathrm{Tr}(a) + \mathrm{Tr} \circ \tilde{g}(\alpha)$. Therefore, either $\mathrm{Tr} \circ g = \mathrm{Tr} \circ \tilde{g}$ or $\mathrm{Tr} \circ g = 1 + \mathrm{Tr} \circ \tilde{g}$. So we assume g only has odd degree monomials. So now take the $\chi_1 \in \boldsymbol{\chi}_q$ canonical additive character of $\mathbb{F}_q$. Then for any α, $\chi_1(\alpha) = (-1)^{\mathrm{Tr}(\alpha)}$. So for a polynomial $g \in \mathbb{F}_q[X]$ of degree at most d, the encoding of the dual BCH code is

$$\mathrm{Enc}(g) = \left(\chi_1 \circ g(\alpha)\right)_{\alpha \in \mathbb{F}_q}$$

by Theorem 4.5.10 we have

$$\left| \sum_{\alpha \in \mathbb{F}_q} \chi_1 \circ g(\alpha) \right| \leq (d-1)\sqrt{q}$$

Therefore for any two distinct polynomials $g_1, g_2 \in \mathbb{F}_q[X]$ of degree d we have

$$\left| \sum_{\alpha \in \mathbb{F}_q} \chi_1 \circ (g_1 - g_2)(\alpha) \right| \leq O(d\sqrt{q})$$

So if $d = o(\sqrt{q})$ then we have $\Delta(\mathrm{Tr} \circ g_1, \mathrm{Tr} \circ g_2) \in \left(\frac{1}{2} \pm O\left(\frac{d}{\sqrt{q}}\right)\right) q$. This gives the distance of the dual BCH code. Since we are taking power of (-1) we will assume the codeword is $\mathrm{Tr} \circ g(\alpha)$ instead of $(-1)^{\mathrm{Tr} \circ g(\alpha)}$ since taking the power is just mapping it to the $\mathbb{F}_q$ field. Below we give the algorithm of decoding from $1/8^{th}$ of the distance of dual BCH codes from [Kop26]

5.3.1 Algorithm

Like in the case of decoding from multiplicative character evaluations we proceed with Berlekamp-Welch like arguments. The approach is very similar to that of previous section but because we are working with additive characters there will be some differences.

If we were in a zero error setting then we know the polynomial $G = \mathrm{Tr} \circ g$ is the original codeword. But we have the received word $r : \mathbb{F}_q \to \{0,1\}$ very close to the original codeword. Let $S = \{\alpha \in \mathbb{F}_q : \ \mathrm{Tr} \circ g(\alpha) \neq r(\alpha)\}$. So $|S| = e$. Let $Z_S(X) = \prod_{\alpha \in S}(X - \alpha)^M$. Hence Z_S has zero at every point of S with multiplicity M. Again we take a non-zero multiple of $Z_S(X)$, $E(X)$ which is a h-pseudopolynomial of degree $\leq cq$. So it suffices to have $ch = em$ to get such polynomial. But the degree of $E(X)$ is of the form $iq + j$ for some $i \in \{0, \ldots, c\}$ and $j \in \{0, \ldots, h\}$. We multiply $E(X)$ with Λ^{c-i} so that the

degree become $cq + j$. So let $h^* = j$. Like the case of multiplicative characters for all $0 \le \ell < M-1$, $\deg(E_{\langle \ell \rangle}) \le h$. So we take the polynomial $E(X) \cdot G(X)$. So now we compute derivatives of $E \cdot G$. For any $\alpha \in \mathbb{F}_q$ and $0 \le \ell < M$ we have

$$\begin{aligned}
(E \cdot G)^{(\ell)}(\alpha) &= \sum_{\ell_1+\ell_2=\ell} E^{(\ell_1)}(\alpha) \cdot G^{(\ell_2)}(\alpha) \\
&= E^{(\ell)}(\alpha) \cdot G(\alpha) + \sum_{\ell_1+\ell_2=\ell, \ell_1<\ell} E^{(\ell_1)}(\alpha) \cdot G^{(\ell_2)}(\alpha) \\
&= E^{(\ell)}(\alpha) \cdot G(\alpha) + \left(\sum_{\ell_1+\ell_2=\ell, \ell_1<\ell} E_{\langle \ell_1 \rangle}(\alpha) \cdot G_{g,\ell_2}(\alpha) \right) && \text{[By Corollary D.3]} \\
&= E^{(\ell)}(\alpha) \cdot \mathrm{Tr} \circ g(\alpha) + \left(\sum_{\ell_1+\ell_2=\ell, \ell_1<\ell} E_{\langle \ell_1 \rangle}(\alpha) \cdot G_{g,\ell_2}(\alpha) \right)
\end{aligned}$$

So define $V_\ell(X) = \sum_{\ell_1+\ell_2=\ell, \ell_1<\ell} E_{\langle \ell_1 \rangle}(X) \cdot G_{g,\ell_2}(X)$. Then we have

$$(E \cdot G)^{(\ell)}(\alpha) = r(\alpha) \cdot E^{(\ell)}(\alpha) + V_\ell(\alpha)$$

where $\deg(V_\ell) \le h + dM$. The change to $r(\alpha)$ from $\mathrm{Tr} \circ g(\alpha)$ is valid because $E^{(\ell)}(\alpha) = 0$ for all $0 \le \ell < M-1$ at all error points i.e., for all $\alpha \in S$. The degree bound for $E \cdot G$ we get is $\deg(E \cdot G) = (cq + h^*) + d \cdot \frac{q}{2}$. So we get $E \cdot G, V_0, \ldots, V_{M-1}$ which satisfies the above equations with proper degree bounds. So we have the algorithm:

Algorithm 3: Decoding from $1/8^{th}$ of distance of Dual BCH Codes

Input: Noisy received word $r : \mathbb{F}_q \to \{0,1\}$ with parameters degree $d \le O(\epsilon\sqrt{q})$, error-bound $e \le (1/8 - \epsilon)\, q$.

Output: Find the closest codeword $\mathrm{Tr} \circ g$ where $g \in \mathbb{F}_q[X]$ of degree at most d.

begin
- $P = 0$
- **while** $d > 0$ **do**
  - Set $M = 16/\epsilon d$, $c = M/2$, $h = 2e$
  - **for** $h^* \in \{0, \ldots, h\}$ **do**
    - Solve linear system to find $F(X), E(X), V_0, \ldots, V_{M-1} \in \mathbb{F}_q[X]$ with $\deg(F) = (d/2 + c)q + h^*$ and $\deg(E) = cq + h^*$ where E is a h-pseudopolynomial, $\deg(V_i) \le h + dM$ for all $0 \le i < M$ such that
    $$F^{(\ell)}(\alpha) = r(\alpha) \cdot E^{(\ell)}(\alpha) + V_\ell(\alpha)$$
    for all $\alpha \in \mathbb{F}_q$ and $0 \le \ell < M$.
  - **if** *no such* h^* *exists* **then**
    - Continue
  - Let $aX^{\deg(F)}$ and $bX^{\deg(E)}$ are leading monomials of F, E respectively
  - $a_d \longleftarrow a/b$
  - $P \longleftarrow P + a_d X^d$
  - **for** $\alpha \in \mathbb{F}_q$ **do**
    - $r(\alpha) \longleftarrow r(\alpha) - \mathrm{Tr}(a_d \alpha^d)$
  - $d \longleftarrow d - 2$
- **return** P

5.3.2 Relating F, E and G and Their Leading Coefficients

Suppose the algorithm was able to find the polynomials $F, E, V_0, \ldots, V_{M-1}$ such that for all $\alpha \in \mathbb{F}_q$ and $0 \le \ell < M$,

$$F^{(\ell)}(\alpha) = r(\alpha) \cdot E^{(\ell)}(\alpha) + V_\ell(\alpha)$$

Here we will anlayze the last few steps of the algorithm correctly determines the monomials of G and their coefficients.

Step 1: Relating F, E and G: Like in the case of Multiplicative characters we want pseudopolynomials $A, B \in \mathbb{F}_q[X]$ such that

$$A(X)\big(F(X) - E(X) \cdot G(X)\big) = B(X) \tag{5.6}$$

So we set the parameters $t, k \in \mathbb{N}$ such that $t = 3/8 \cdot M$ and $k = e + 4dM$. We claim there exists pseudopolynomials $A, B \in \mathbb{F}_q[X]$ of the form

$$A(X) = \sum_{i=0}^{t-1} A_i(X)\Lambda^i(X), \qquad B(X) = \sum_{i=0}^{t+c+\frac{d}{2}-1} B_i(X)\Lambda^i(X)$$

where $\deg(A_i) \le k$ and $\deg(B_j) \le k + h + dM$ for all $i \in \{0, \ldots, t-1\}$ and $j \in \{0, \ldots, t + c + \frac{d}{2} - 1\}$ such that (5.6) holds. The total number of coefficients of A, B is $N = t(k+1) + (t + c + d/2) \cdot (k + h + dM + 1)$. Now Substituting the values of t, k we get

$$N > \frac{3}{8}M(e + 4dM) + \frac{7}{8}M(3e + 5dM) > 3eM + 5dM^2 \tag{5.7}$$

Using the (5.6) we impose the linear constraints using the coefficients that for all $0 \le \ell < M$ and for $\alpha \in \mathbb{F}_q \setminus S$

$$(A \cdot (F - E \cdot G))^{(\ell)}(\alpha) = B^{(\ell)}(\alpha) \tag{5.8}$$

So this says the polynomial equation $A(X) \cdot (F(X) - E(X) \cdot G(X)) - B(X)$ vanishes at every $\alpha \in \mathbb{F}_q \setminus S$ with multiplicity at least M. So this has at least

$$M \cdot |\mathbb{F}_q \setminus S| geq M(q - e) > \left(\frac{7}{8} + \epsilon\right) M \tag{5.9}$$

many roots with counting multiplicity. On the other hand the degree bounds of $A(X) \cdot (F(X) - E(X) \cdot G(X)) - B(X)$ gives

$$\deg(A \cdot (F - E \cdot G) - B) \le \max(\deg(A) + \deg(F - EG), \deg(B)) < \left(\frac{7}{8} + \frac{\epsilon}{2}\right) MQ \tag{5.10}$$

So comparing (5.10) with (5.9) we get $A(X) \cdot (F(X) - E(X) \cdot G(X)) \equiv B(X)$. Now it remains to show that from (5.8) can be enforced using fewer than N linear equations. Now for each $\ell \in \{0, \ldots, M-1\}$ we have

$$A_{\langle \ell \rangle}(X) = \sum_{i=0}^{t-1} (-1)^i A_i^{(\ell-i)}(X), \quad B_{\langle \ell \rangle}(X) = \sum_{i=0}^{t+c+\frac{d}{2}-1} (-1)^i B_i^{(\ell-i)}(X), \quad E_{\langle \ell \rangle}(X) = \sum_{i=0}^{c} (-1)^i E_i^{(\ell-i)}(X)$$

such that for all $\alpha \in \mathbb{F}_q$ we have $A^{(\ell)}(\alpha) = A_{\langle \ell \rangle}(\alpha)$, $B^{(\ell)}(\alpha) = B_{\langle \ell \rangle}(\alpha)$ and $E^{(\ell)}(\alpha) = E_{\langle \ell \rangle}(\alpha)$. Thus, for any $\alpha \in \mathbb{F}_q \setminus S$ we have

$$\begin{aligned}
(F - E \cdot G)^{(\ell)}(\alpha) &= \Big(r(\alpha)E^{(\ell)}(\alpha) + V_\ell(\alpha)\Big) - \sum_{\ell_1 + \ell_2 = \ell} E^{(\ell_1)}(\alpha) \cdot G^{(\ell_2)}(\alpha) \\
&= \big(r(\alpha)E_{\langle \ell \rangle}(\alpha) + V_\ell(\alpha)\big) - \left(E_{\langle \ell \rangle}(\alpha)G(\alpha) + \sum_{\ell_1 + \ell_2 = \ell, \ell_1 < \ell} E_{\langle \ell_1 \rangle}(\alpha) \cdot G_{g,\ell_2}(\alpha)\right) \\
&= (r(\alpha) - G(\alpha))E_{\langle \ell \rangle}(\alpha) + \left(V_\ell(\alpha) - \sum_{\ell_1 + \ell_2 = \ell, \ell_1 < \ell} E_{\langle \ell_1 \rangle}(\alpha)G_{g,\ell_2}(\alpha)\right) \\
&= 0 + W_\ell(\alpha) = W_\ell(\alpha)
\end{aligned}$$

where $W_\ell(X)$ is the polynomial of degree at most $h + dM$ given by

$$W_\ell(X) = V_\ell(X) - \sum_{\ell_1+\ell_2=\ell, \ell_1<\ell} E_{\langle \ell_1 \rangle}(X) G_{g,\ell_2}(X)$$

So for any $\alpha \in \mathbb{F}_q \setminus S$ we have

$$(A \cdot (F - E \cdot G))^{(\ell)}(\alpha) = \sum_{\ell_1+\ell_2=\ell} A^{(\ell_1)}(\alpha) \cdot (F - E \cdot G)^{(\ell_2)}(\alpha) = \sum_{\ell_1+\ell_2=\ell} A_{\langle \ell_1 \rangle}(\alpha) \cdot W_{\ell_2}(\alpha)$$

SO we have the linear constraints for every $0 \le \ell < M$ the following polynomial equations

$$\sum_{\ell_1+\ell_2=\ell} A_{\langle \ell_1 \rangle}(X) \cdot W_{\ell_2}(X) = B_\ell(X)$$

Both left and right of the quality degrees are equal, $3e + dM$. So the number of constrains imposed is at most $M \cdot (3e + dM + 1) < 3eM + 5dM^2$. Combining this with (5.7) we get a non-zero solution for A, B such that (5.6) is satisfied.

Step 2: Relating the Leading Coefficients: We now have the identity $A(X)\big(F(X) - E(X) \cdot G(X)\big) = B(X)$. So $\deg(A) + \deg(F - E \cdot G) = \deg(B)$. Since A is a k-pseudopolynomial and B is a $k + h + dM$-pseudopolynomial we have

$$\deg(A) \in q\mathbb{Z} + [0, k] \subseteq q\mathbb{Z} + \left[0, \frac{1}{8}q\right), \qquad \deg(B) \in q\mathbb{Z} + [0, k + h + dM] \subseteq q\mathbb{Z} + \left[0, \frac{3}{8}q\right)$$

Therefore $\deg(F - E \cdot G) \in q\mathbb{Z} + \left(-\frac{1}{8}q, \frac{3}{8}q\right)$. Now by construction $\deg(F) = \deg(E \cdot G) = (c + d/2)q + h^*$. So we get

$$\deg(F) = \deg(E \cdot G) \in q\mathbb{Z} + \left[\frac{q}{2}, \frac{q}{2} + 2e\right] \subseteq q\mathbb{Z} + \left[\frac{1}{2}q, \frac{3}{4}q\right).$$

Since $\left(-\frac{1}{8}q, \frac{3}{8}q\right)$ and $\left[\frac{1}{2}q, \frac{3}{4}q\right)$ are disjoint sets modulo q. So $\deg(F - E \cdot G) \neq \max\{\deg(F), \deg(E \cdot G)\}$. Therefore, many monomials of high degree in F and $E \cdot G$ got canceled in $F - E \cdot G$. So F and $E \cdot G$ have the same leading monomial. Let the leading coefficients of F, E are a, b. Then the leading coefficient of G which must equal a/b. So the algorithm in each iteration of the while loop correctly computes the leading coefficient of the rest of the polynomial. Hence, the algorithm correctly outputs $\mathrm{Tr} \circ G$.

Chapter 6

Bibliography

[Art11] Michael Artin. *Algebra*. Pearson Education, Prentice Hall, Boston, Mass., 2. ed. edition, 2011.

[Bom] Enrico Bombieri. *Counting points on curves over finite fields*, pages 234–241. Springer-Verlag. doi:10.1007/bfb0057311.

[BRC60] R.C. Bose and D.K. Ray-Chaudhuri. *On a class of error correcting binary group codes*. Information and Control, 3(1):68–79, March 1960. doi:10.1016/s0019-9958(60)90287-4.

[GK14] Venkatesan Guruswami and Swastik Kopparty. *Explicit subspace designs*. Combinatorica, 36(2):161–185, October 2014. doi:10.1007/s00493-014-3169-1.

[Hoc59] Alexis Hocquenghem. *Codes correcteurs d'erreurs*. Chiffres, 2:147–156, 1959.

[HSW95] James G. Huard, Blair K. Spearman, and Kenneth S. Williams. *An arithmetic approach to the Davenport-Hasse relation over GF(p)*. Rocky Mountain Journal of Mathematics, 25(4):1341–1350, 1995. doi:10.1216/rmjm/1181072149.

[IK12] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*. AMS ebook collection. American Mathematical Society, Providence, R.I, online-ausg. edition, 2012. Includes bibliographical references (p. 599-610) and index. - Electronic reproduction; Providence, Rhode Island; American Mathematical Society; 2012. - Description based on print version record.

[Kop13] Swastik Kopparty. *The Weil bound*. Lecture notes, Rutgers University, Finite Fields course, Fall 2013, 2013.

[Kop26] ———. *Recovering polynomials over finite fields from noisy character values*. January 2026. arXiv:2601.07137.

[Kum24] Mrinal Kumar. *Polynomial Methods in Combinatorics*. Course, TIFR Mumbai, Fall 2024, 2024.

[Lan02] Serge Lang. *Algebra*. Springer New York, 2002. doi:10.1007/978-1-4613-0041-0.

[LN96] Rudolf Lidl and Harald Niederreiter. *Finite Fields*. Cambridge University Press, October 1996. doi:10.1017/cbo9780511525926.

[Mil06] Steven J. Miller. *An Invitation to Modern Number Theory*. Princeton University Press, Princeton, 2006. Description based on publisher supplied metadata and other sources.

[Mor96] Patrick Morandi. *Field and Galois Theory*. Springer New York, 1996. doi:10.1007/978-1-4612-4040-2.

[Sch76] Wolfgang M. Schmidt. *Equations over Finite Fields: An Elementary Approach*, volume 536 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 1976. doi:10.1007/BFb0080437.

[Sch05] A. J. Scholl. *Transitivity of trace and norm*, 2005. Lecture notes for Galois Theory (Michaelmas 2005).

[Sta73] Harold M. Stark. *On the Riemann hypothesis in hyperelliptic function fields.* In *Analytic Number Theory*, volume 24 of *Proceedings of Symposia in Pure Mathematics*, pages 285–302. American Mathematical Society, Providence, Rhode Island, 1973.

[Ste69] S A Stepanov. *On the number of points of a hyperelliptic curve over a finite field.* Mathematics of the USSR-Izvestiya, 3(5):1103–1114, October 1969. doi:10.1070/im1969v003n05abeh000834.

[Ste70] S. A. Stepanov. *Elementary method in the theory of congruences for a prime modulus.* Acta Arithmetica, 17(3):231–247, 1970.

[Ste74] Serguei A. Stepanov. *An elementary method in the theory of equations over finite fields.* In *Proceedings of the International Congress of Mathematicians*, volume 1, pages 383–391. Vancouver, Canada, 1974.

[Wei41] André Weil. *On the riemann hypothesis in function-fields.* Proceedings of the National Academy of Sciences, 27(7):345–347, 1941. doi:10.1073/pnas.27.7.345.

[Wei48] ———. *On some exponential sums.* Proceedings of the National Academy of Sciences, 34(5):204–207, May 1948. doi:10.1073/pnas.34.5.204.

[Wei49] ———. *Numbers of solutions of equations in finite fields.* Bulletin of the American Mathematical Society, 55(5):497–508, 1949. doi:10.1090/s0002-9904-1949-09219-4.

Chapter A

Algebraic and Analytic Background

This appendix collects several results from algebra and analysis that are used at various points throughout the report. The topics are diverse: field theory, complex analysis, elementary number theory, roots of unity, and symmetric polynomials, but each result is self-contained and included for ease of reference.

§ A.1 Fields

For any algebraic extension $E/\mathbb{F}_q$, the *Galois group* $\mathrm{Gal}(E/\mathbb{F}_q)$ is the group of all field automorphisms $\sigma : E \to E$ that fix every element of $\mathbb{F}_q$. For the extension $\mathbb{F}_{q^r}/\mathbb{F}_q$, this group is cyclic of order r, generated by the Frobenius automorphism $\phi : \alpha \mapsto \alpha^q$. Recall from the chapter on finite fields that, one of the key fact connecting automorphisms to irreducible polynomials and it's roots is the Isomorphism Extension Theorem from the subsection 1.1.4: any isomorphism between two subfields of an algebraic closure extends to an automorphism of the whole closure, and in particular, any two roots of an irreducible polynomial over $\mathbb{F}_q$ are conjugates under the action of $\mathrm{Gal}(\mathbb{F}_{q^r}/\mathbb{F}_q)$.

The following theorem makes this concrete for irreducible polynomials over $\mathbb{F}_q$. It describes precisely how an irreducible polynomial $h \in \mathbb{F}_q[X]$ of degree d behaves when viewed over a larger field $\mathbb{F}_{q^k}$: it breaks into $\gcd(k,d)$ irreducible factors, all of the same degree, and the Galois group of $\mathbb{F}_{q^r}/\mathbb{F}_q$ (where $r = \gcd(k,d)$) acts transitively on these factors, exactly the statement of the Isomorphism Extension Theorem applied in this finite-field setting.

Theorem A.1.1

Let $K = \mathbb{F}_{q^k}$ and $E = \mathbb{F}_{q^r}$. Suppose $h(X) = X^d - a_1X^{d-1} + \cdots + (-1)^d a_d$ be an irreducible polynomial in $\mathbb{F}_q[X]$. Then in $\mathbb{F}_{q^k}[X]$, h splits into $r = \gcd(k,d)$ many irreducible polynomials of degree d/r:

$$h(X) = h_1(X) \cdots h_r(X)$$

If we normalize $h_i(X)$ such that h_i is monic then $h_i(X) \in \mathbb{F}_{q^r}[X]$ for all $i \in [r]$. Then for any $i, j \in [r]$ with $i \neq j$ there exists $\sigma \in \mathrm{Gal}(\mathbb{F}_{q^r}/\mathbb{F}_q)$ such that $\sigma(h_i) = h_j$.

§ A.2 Meromorphic and Analytic Continuation

The analytic continuation of the Riemann zeta function, and more generally of Dirichlet L-functions, requires the notion of a meromorphic function. We record the relevant definitions and a motivating example here.

Definition A.2.1: Zero, Pole, Order, Residue

Assume f has a convergent Laurent series expansion about z_0:

$$f(z) = a_n(z-z_0)^n + a_{n+1}(z-z_0)^{n+1} + \cdots = \sum_{m=n}^{\infty} a_m(z-z_0)^m,$$

with $a_n \neq 0$. If $n > 0$ we say f has a zero of order n *at z_0; if $n < 0$ we say f has a* pole of order $-n$ *at z_0. If $n = -1$ we say f has a* simple pole with residue a_{-1}. *We denote the order of f at z_0 by* $\operatorname{ord}_f(z_0) = n$.

Definition A.2.2: Meromorphic Function

We say f is meromorphic *at z_0 if there exists a disk about z_0 of radius r and an integer n_0 such that for all z with $|z-z_0| < r$,*

$$f(z) = \sum_{n \geq n_0} a_n(z-z_0)^n.$$

If f is meromorphic at each point in a region, we say f is meromorphic *on that region. If $n_0 \geq 0$, we say f is* ***analytic****.*

A canonical example illustrates the notion of continuation. Consider the geometric series $G(r) = \sum_{n=1}^{\infty} r^n$, which converges for $|r| < 1$ and equals $\frac{1}{1-r}$ there. Define $H(r) = \frac{1}{1-r}$; then H is well-defined for all $r \neq 1$ and agrees with G wherever G converges. Since H is defined on a strictly larger domain and has a simple pole at $r = 1$ (with residue -1), we call H a **meromorphic continuation** of G. Had H no poles, we would call it an **analytic continuation**. A function that is analytic at every point of $\mathbb{C}$, equivalently one whose Taylor expansion converges everywhere so that it has no poles anywhere, is said to be **entire**.

§ A.3 Fermat's Little Theorem

Fermat's Little Theorem is one of the oldest results in elementary number theory, recorded in a 1640 letter from Pierre de Fermat to Frénicle de Bessy. It underlies essentially every computation we perform in the multiplicative group $\mathbb{F}_p^*$ – from the very existence of the canonical additive character of $\mathbb{F}_q$ (where $q = p^r$), through the character congruence $\psi(l) \equiv l^f \bmod P$ used in the arithmetic proof of the Davenport–Hasse relation, down to the mod-q reduction of Gauss sums appearing in the proof of the Law of Quadratic Reciprocity. We record it here for completeness.

Theorem A.3.1 Fermat's Little Theorem

Let p be a prime. Then for every integer a we have

$$a^p \equiv a \pmod{p},$$

and in particular, if $p \nmid a$ then $a^{p-1} \equiv 1 \pmod{p}$.

Proof: The non-zero residues modulo p form a group $\mathbb{F}_p^* = \{1, 2, \ldots, p-1\}$ under multiplication of order $p-1$. By Lagrange's theorem every element of a finite group satisfies $g^{|G|} = e$, so $a^{p-1} \equiv 1 \pmod{p}$ for every $a \in \mathbb{F}_p^*$, i.e. every a with $p \nmid a$. Multiplying both sides by a gives $a^p \equiv a \pmod{p}$ for such a; and the congruence $a^p \equiv a \pmod{p}$ is trivially true when $p \mid a$, so it holds for every integer a. ∎

§ A.4 A Product Identity for Roots of Unity

The following identity, used in the factorisation of lifted L-functions, expresses a product of linear factors twisted by roots of unity as a perfect power.

Lemma A.4.1 Roots-of-Unity Product Identity

Let $n, m \in \mathbb{N}$ and set $d = \gcd(m, n)$. Then

$$\prod_{t=1}^{n} \left(1 - e^{2\pi i m t/n} X\right) = \left(1 - X^{n/d}\right)^d .$$

Proof: Let $\omega = e^{2\pi i/n}$ be a primitive n-th root of unity, so the factors are $1 - \omega^{mt}X$ for $t = 1, \dots, n$. Since $t = n$ gives $\omega^{mn} = 1 = \omega^0$, the product is the same as $\prod_{t=0}^{n-1}(1 - \omega^{mt}X)$.

Because $d = \gcd(m, n)$, the element ω^m is a primitive (n/d)-th root of unity. As t ranges over $\{0, 1, \dots, n-1\}$, the values ω^{mt} cycle through the (n/d)-th roots of unity exactly d times each. Hence,

$$\prod_{t=0}^{n-1} \left(1 - \omega^{mt}X\right) = \left(\prod_{k=0}^{n/d-1} \left(1 - (\omega^m)^k X\right)\right)^d .$$

Since ω^m is a primitive (n/d)-th root of unity, the inner product equals $1 - X^{n/d}$ by the standard factorisation of $1 - Y^N = \prod_{k=0}^{N-1}(1 - \zeta^k Y)$. Therefore, the whole product equals $(1 - X^{n/d})^d$. ■ ■

§ A.5 Elementary Symmetric Polynomials

Waring's Formula expresses each power-sum symmetric polynomial $s_k = x_1^k + x_2^k + \cdots + x_n^k$ as an explicit polynomial in the elementary symmetric polynomials $\sigma_1, \dots, \sigma_n$. It is the source of the closed-form reduction of lifted Kloosterman sums $K(\chi^{(s)}; a, b)$ to the ground-field sum $K(\chi; a, b)$ in Lemma 2.7.3: applying Waring's formula to $\omega_1^s + \omega_2^s$ with $\sigma_1 = \omega_1 + \omega_2 = -K(\chi; a, b)$ and $\sigma_2 = \omega_1\omega_2 = q$ produces the binomial-coefficient expansion. We state the general identity here for reference.

Theorem A.5.1 Waring's Formula

Let $x_1, \dots, x_n$ be indeterminates, let

$$\sigma_j = \sum_{1 \le i_1 < i_2 < \cdots < i_j \le n} x_{i_1} x_{i_2} \cdots x_{i_j} \qquad (j = 1, \dots, n)$$

be the elementary symmetric polynomials in $x_1, \dots, x_n$, and let

$$s_k = x_1^k + x_2^k + \cdots + x_n^k$$

be the k^{th} power-sum. Then for every integer $k \ge 1$,

$$s_k = \sum (-1)^{i_2 + i_4 + i_6 + \cdots} \cdot \frac{(i_1 + i_2 + \cdots + i_n - 1)! \cdot k}{i_1! \cdot i_2! \cdots i_n!} \cdot \sigma_1^{i_1} \sigma_2^{i_2} \cdots \sigma_n^{i_n},$$

where the summation runs over all n-tuples $(i_1, \dots, i_n)$ of non-negative integers satisfying $i_1 + 2i_2 + \cdots + ni_n = k$. Each coefficient of $\sigma_1^{i_1}\sigma_2^{i_2}\cdots\sigma_n^{i_n}$ is an integer.

Theorem A.5.2

Suppose $f \in \mathbb{F}_q[X_1, \ldots, X_n]$ *is a symmetric polynomial. Let* $\overline{X} = (X_1, \ldots, X_n)$. *Then there exists a polynomial* $g \in \mathbb{F}_q[Y_1, \ldots, Y_n]$ *such that*

$$f(X_1, \ldots, X_n) = g(\mathrm{ESym}_1(\overline{X}), \ldots, \mathrm{ESym}_n(\overline{X})).$$

Moreover,

(i) if $\deg_{X_i}(f) = d$ *for all* $i \in [n]$ *then* $\deg G = d$.

(ii) if $\deg f = D$ *then for every monomial* $Y_1^{d_1} \cdots Y_n^{d_n}$ *of* g *with non-zero coefficient then* $d_1 + 2 \cdot d_2 + \cdots + n \cdot d_n = D$.

Chapter B

Modulus Bounds from Power Sum Estimates

A recurring step in the proof of the Weil bounds is the following: one shows that the power sums $\omega_1^k + \cdots + \omega_l^k$ grow at most like B^k, and then concludes that each reciprocal root ω_j of the associated L-function satisfies $|\omega_j| \le B = q^{1/2}$. This appendix collects the analytic and number-theoretic lemmas that make this deduction rigorous. The logical order is: simultaneous Diophantine approximation (Lemma B.3) feeds into a lower bound on the real parts of power sums (Lemma B.4), which in turn implies the modulus bounds (Lemma B.1 and Lemma B.2).

The first lemma deduces modulus bounds directly from a growth hypothesis on the absolute value of the power sum, using an analytic function argument.

Lemma B.1

Let $\omega_1, \dots, \omega_l$ be complex numbers, and let $B > 0$. If

$$|\omega_1^k + \cdots + \omega_l^k| = O(B^k) \quad \text{for all } k \in \mathbb{N}$$

then $|\omega_j| \le B$ for all $j = 1, \dots, l$.

Proof: For $|z| < 1$ we have the following series for $\log(1-z)$:

$$-\log(1-z) = z + \frac{1}{2}z^2 + \frac{1}{3}z^3 + \cdots = \sum_{k=1}^{\infty} \frac{1}{k} z^k.$$

So for $|z| < \frac{1}{|\omega|}$ we have $-\log(1-\omega z) = \sum_{k=1}^{\infty} \frac{1}{k}\omega^k z^k$, and therefore

$$-\sum_{j=1}^{l} \log(1-\omega_j z) = \sum_{k=1}^{\infty} \frac{1}{k}\left(\omega_1^k + \cdots + \omega_l^k\right) z^k.$$

By hypothesis, $|\omega_1^k + \cdots + \omega_l^k| = O(B^k)$, so this series is convergent for $|z| < B^{-1}$. Hence the function on the left is analytic for $|z| < B^{-1}$, which means $1 - \omega_j z \neq 0$ whenever $|z| < B^{-1}$.

Now if $|\omega_j| > B$ for some j, take $z = \frac{1}{\omega_j}$. Then $|z| = \frac{1}{|\omega_j|} < B^{-1}$, but $1 - \omega_j \cdot \frac{1}{\omega_j} = 0$, a contradiction ↯. Hence $|\omega_j| \le B$ for all $j \in [l]$. ∎

The second lemma is a strengthening: it suffices to bound only the real part of the power sum rather than its absolute value.

This is the form that actually appears in the Weil bound argument, where the power sums are shown to be dominated by B^k after taking real parts via a suitable averaging.

Lemma B.2

Let $\omega_1, \dots, \omega_l$ be complex numbers, and let $B > 0$. If

$$\Re(\omega_1^k + \cdots + \omega_l^k) \le B^k \quad \text{for } k = 1, 2, \dots,$$

then $|\omega_j| \le B$ for all $j = 1, \dots, l$.

Proof: The proof follows the same argument as Lemma B.1. Since $\Re(\omega_1^k + \cdots + \omega_l^k) \le |\omega_1^k + \cdots + \omega_l^k|$, the series

$$-\sum_{j=1}^{l} \log(1 - \omega_j z) = \sum_{k=1}^{\infty} \frac{1}{k}\left(\omega_1^k + \cdots + \omega_l^k\right) z^k$$

has real part bounded by $\sum_{k=1}^{\infty} \frac{1}{k} B^k |z|^k < \infty$ for $|z| < B^{-1}$. Hence the left side is analytic for $|z| < B^{-1}$, so $1 - \omega_j z \neq 0$ for such z. If $|\omega_j| > B$, taking $z = 1/\omega_j$ gives $|z| < B^{-1}$ but $1 - \omega_j z = 0$, a contradiction $\lightning$. Hence $|\omega_j| \le B$ for all $j \in [l]$. ∎

The next lemma is Dirichlet's theorem on simultaneous Diophantine approximation. It is the number-theoretic engine behind the final lemma: it produces integers k for which $e^{2\pi i k\theta_j}$ is simultaneously close to 1 for each j, making the real parts of ω_j^k large.

Lemma B.3

Let $\theta_1, \dots, \theta_l$ be real numbers. Then there exist $(l+1)$-tuples of integers $(k, m_1, \dots, m_l)$ with arbitrarily large $k > 0$ such that

$$\left|\theta_i - \frac{m_i}{k}\right| < k^{-1-(1/l)} \quad (i = 1, \dots, l).$$

Proof: For any $x \in \mathbb{R}$, write $x = [x] + \{x\}$ where $[x]$ is the integer part and $\{x\} \in [0, 1)$ is the fractional part. Let $N > 0$ be an integer. Consider the $N^l + 1$ points

$$\big(\{u\theta_1\}, \dots, \{u\theta_l\}\big), \quad u = 0, 1, \dots, N^l.$$

These are $N^l + 1$ points in the half-open unit cube $0 \le x_j < 1$. This unit cube can be decomposed into N^l half-open small cubes of side N^{-1}. Hence, there will be two points lying in the same small cube. Let these two points have parameters u' and u with $u' < u$. Then

$$|\{u\theta_j\} - \{u'\theta_j\}| < N^{-1} \quad \forall\, j \in [l],$$

or equivalently $|u\theta_j - u'\theta_j - m_j| < N^{-1}$ for certain integers $m_1, \dots, m_l$. Let $k = u - u'$. Then $k < N^l$ and

$$|k\theta_j - m_j| < N^{-1} \quad \forall\, j \in [l].$$

Since $k \le N^l$, we have $N \ge k^{1/l}$, and so $kN \ge k^{1+1/l}$. Therefore,

$$\left|\theta_j - \frac{m_j}{k}\right| < \frac{1}{kN} \le k^{-1-1/l} \quad \forall\, j \in [l].$$

If any θ_j is irrational, as $N \to \infty$ the inequality $|k\theta_j - m_j| < N^{-1}$ cannot hold for bounded k, so there exist tuples with arbitrarily large k. If all $\theta_j = a_j/b$ are rational (with $b > 0$), set $k = tb$, $m_j = ta_j$ for $t = 1, 2, \dots$ ∎

The final lemma is the strongest of the four. It says that for any collection of nonzero complex numbers, there are infinitely many k at which the real part of their k-th power sum is essentially as large as the sum of the k-th powers of their moduli. Together with the preceding lemma, this completes the chain of implications used to bound the reciprocal roots of L-functions.

Lemma B.4

Let $\omega_1, \dots, \omega_l$ be non-zero complex numbers. Then there exist infinitely many positive integers k such that

$$\Re(\omega_1^k + \cdots + \omega_l^k) > \left(1 - 2\pi k^{-1/l}\right)\left(|\omega_1|^k + \cdots + |\omega_l|^k\right).$$

Hence, for any given $\varepsilon > 0$, there exist infinitely many k such that

$$\Re(\omega_1^k + \cdots + \omega_l^k) > (1 - \varepsilon)\left(|\omega_1|^k + \cdots + |\omega_l|^k\right).$$

Proof: Write $\omega_j = |\omega_j|\, e(\theta_j)$ for real θ_j, where $e(\theta) := e^{2\pi i\theta}$. By Lemma B.3 applied to $\theta_1, \dots, \theta_l$, there are infinitely many positive integers k and integers $m_1, \dots, m_l$ such that

$$|k\theta_j - m_j| < k^{-1/l} \quad \forall\, j \in [l].$$

For any such k, using $|e(\alpha) - e(\beta)| \le 2\pi|\alpha - \beta|$ for real α, β:

$$|e(k\theta_j) - 1| = |e(k\theta_j) - e(m_j)| \le 2\pi|k\theta_j - m_j| < 2\pi k^{-1/l}.$$

Hence, $\Re(e(k\theta_j)) > 1 - 2\pi k^{-1/l}$, and so

$$\Re(\omega_j^k) = |\omega_j|^k\, \Re(e(k\theta_j)) > \left(1 - 2\pi k^{-1/l}\right)|\omega_j|^k \quad \forall\, j \in [l].$$

Summing over j gives the stated inequality. The ε-version follows by taking k large enough that $2\pi k^{-1/l} < \varepsilon$. ∎

Chapter C

Continued Fractions over Polynomial Rings

Just as every real number has a continued fraction expansion obtained by iterating the floor function, every rational function $r_0/r_1 \in \mathbb{F}_q(X)$ has an analogous expansion obtained by iterating polynomial division. The role of the floor function is played by the leading term of the quotient in the Euclidean algorithm, and the resulting *partial quotients* $A_0, A_1, \ldots, A_s$ are polynomials over $\mathbb{F}_q$.

Concretely, we write

$$\frac{r_0}{r_1} = A_0 + \cfrac{1}{A_1 + \cfrac{1}{A_2 + \cdots}} =: [A_0, A_1, A_2, \ldots, A_s].$$

The *convergents* of this expansion are the rational functions $[A_0, A_1, \ldots, A_i] = P_i/Q_i$, where (P_i) and (Q_i) are sequences of polynomials defined by the recurrences

$$P_i = A_i P_{i-1} + P_{i-2}, \qquad Q_i = A_i Q_{i-1} + Q_{i-2},$$

starting from $P_{-1} = 1$, $P_0 = A_0$, $Q_{-1} = 0$, $Q_0 = 1$.

Let $r_0, r_1 \in \mathbb{F}_q[X]$ with $r_1 \neq 0$. Applying the Euclidean algorithm yields a sequence

$$r_i = A_i r_{i+1} + r_{i+2}, \quad i = 0, 1, \ldots, s,$$

where $0 \leq \deg(r_{i+1}) < \deg(r_i)$ for $i = 1, \ldots, s$ and $r_{s+2} = 0$. Here $A_0, A_1, \ldots, A_s \in \mathbb{F}_q[X]$, with $A_1, \ldots, A_s$ of positive degree. This gives the *continued fraction expansion*

$$\frac{r_0}{r_1} = A_0 + \cfrac{1}{A_1 + \cfrac{1}{A_2 + \cdots}} =: [A_0, A_1, A_2, \ldots, A_s].$$

Define polynomials $P_i, Q_i \in \mathbb{F}_q[X]$ for $i = -1, 0, 1, \ldots, s$ by the recurrences

$$P_{-1} = 1, \quad P_0 = A_0, \quad P_i = A_i P_{i-1} + P_{i-2} \quad (i = 1, \ldots, s),$$
$$Q_{-1} = 0, \quad Q_0 = 1, \quad Q_i = A_i Q_{i-1} + Q_{i-2} \quad (i = 1, \ldots, s).$$

It is clear that $\deg(P_{i-1}) < \deg(P_i)$ and $\deg(Q_{i-1}) < \deg(Q_i)$ for $i = 1, \ldots, s$. We extend the definition of degree by $\deg(f/g) = \deg(f) - \deg(g)$ for a rational function f/g.

Lemma C.1 Rational Approximants via Convergents

For any rational function ρ of nonnegative degree,

$$[A_0, A_1, \ldots, A_{i-1}, \rho] = \frac{\rho P_{i-1} + P_{i-2}}{\rho Q_{i-1} + Q_{i-2}} \quad \text{for } i = 1, \ldots, s+1.$$

Proof: We prove by induction on i. For $i = 1$:

$$[A_0, \rho] = A_0 + \frac{1}{\rho} \qquad \text{and} \qquad \frac{\rho P_0 + P_{-1}}{\rho Q_0 + Q_{-1}} = \frac{\rho A_0 + 1}{\rho} = A_0 + \frac{1}{\rho}.$$

So the base case holds. Now suppose the identity holds for some i with $1 \le i \le s$. Since $A_i + \rho^{-1}$ is of positive degree, we can apply the inductive hypothesis with $A_i + \rho^{-1}$ in place of ρ:

$$\begin{aligned}
[A_0, \ldots, A_i, \rho] &= [A_0, \ldots, A_{i-1},\ A_i + \rho^{-1}] \\
&= \frac{(A_i + \rho^{-1})\, P_{i-1} + P_{i-2}}{(A_i + \rho^{-1})\, Q_{i-1} + Q_{i-2}} \\
&= \frac{\rho^{-1} P_{i-1} + A_i P_{i-1} + P_{i-2}}{\rho^{-1} Q_{i-1} + A_i Q_{i-1} + Q_{i-2}} \\
&= \frac{\rho^{-1} P_{i-1} + P_i}{\rho^{-1} Q_{i-1} + Q_i} \\
&= \frac{P_{i-1} + \rho P_i}{Q_{i-1} + \rho Q_i},
\end{aligned}$$

where we used the recurrences $P_i = A_i P_{i-1} + P_{i-2}$ and $Q_i = A_i Q_{i-1} + Q_{i-2}$. This is precisely the formula for index $i+1$, completing the induction. ∎

Corollary C.2 Convergents Formula

For $i = 0, 1, \ldots, s$ we have $[A_0, A_1, \ldots, A_i] = P_i / Q_i$.

Proof: Apply Lemma C.1 with $\rho = A_i$ (and index i): $[A_0, \ldots, A_{i-1}, A_i] = (A_i P_{i-1} + P_{i-2})/(A_i Q_{i-1} + Q_{i-2}) = P_i / Q_i$. ∎

Lemma C.3 Representation of r_0/r_1 via Convergents

For $i = 0, 1, \ldots, s$ we have

$$\frac{r_0}{r_1} = \frac{P_i + \beta_i P_{i-1}}{Q_i + \beta_i Q_{i-1}},$$

where $\beta_i = r_{i+2}/r_{i+1}$ is a rational function of negative degree.

Proof: We proceed by induction on i. For $i = 0$:

$$\frac{P_0 + \beta_0 P_{-1}}{Q_0 + \beta_0 Q_{-1}} = \frac{A_0 + \beta_0}{1} = A_0 + \frac{r_2}{r_1} = \frac{A_0 r_1 + r_2}{r_1} = \frac{r_0}{r_1}.$$

Now suppose the identity holds for some $0 \le i < s$. We show it for $i+1$. Using the recurrences $P_{i+1} = A_{i+1} P_i + P_{i-1}$ and $Q_{i+1} = A_{i+1} Q_i + Q_{i-1}$:

$$\frac{P_{i+1} + \beta_{i+1} P_i}{Q_{i+1} + \beta_{i+1} Q_i} = \frac{(A_{i+1} + \beta_{i+1})\, P_i + P_{i-1}}{(A_{i+1} + \beta_{i+1})\, Q_i + Q_{i-1}}.$$

Now observe that

$$A_{i+1} + \beta_{i+1} = A_{i+1} + \frac{r_{i+3}}{r_{i+2}} = \frac{A_{i+1}\, r_{i+2} + r_{i+3}}{r_{i+2}} = \frac{r_{i+1}}{r_{i+2}} = \beta_i^{-1},$$

where we used the Euclidean division $r_{i+1} = A_{i+1} r_{i+2} + r_{i+3}$. Substituting:

$$= \frac{\beta_i^{-1}\, P_i + P_{i-1}}{\beta_i^{-1}\, Q_i + Q_{i-1}} = \frac{P_i + \beta_i P_{i-1}}{Q_i + \beta_i Q_{i-1}} = \frac{r_0}{r_1},$$

where the last equality is the induction hypothesis. ■

Lemma C.4 Determinant Identity for Convergents

For $i = 0, 1, \ldots, s$ we have $P_i Q_{i-1} - P_{i-1} Q_i = (-1)^{i-1}$.

Proof: We prove by induction. For $i = 0$:

$$P_0 Q_{-1} - P_{-1} Q_0 = A_0 \cdot 0 - 1 \cdot 1 = -1 = (-1)^{-1}.$$

Suppose the identity holds for some i. Then:

$$\begin{aligned} P_{i+1} Q_i - P_i Q_{i+1} &= (A_{i+1} P_i + P_{i-1})\, Q_i - P_i\, (A_{i+1} Q_i + Q_{i-1}) \\ &= A_{i+1} P_i Q_i + P_{i-1} Q_i - A_{i+1} P_i Q_i - P_i Q_{i-1} \\ &= P_{i-1} Q_i - P_i Q_{i-1} \\ &= -(P_i Q_{i-1} - P_{i-1} Q_i) \\ &= -(-1)^{i-1} = (-1)^i. \end{aligned}$$

■

Corollary C.5 Coprimality of Convergents

For $i = 0, 1, \ldots, s$ we have $\gcd(P_i, Q_i) = 1$.

Proof: Let $d_i = \gcd(P_i, Q_i)$. By Lemma C.4, d_i divides $P_i Q_{i-1} - P_{i-1} Q_i = (-1)^{i-1}$, so $d_i = 1$. ■

Chapter D

Hasse Derivatives

Definition D.1: Hasse Derivative

Let $\mathbb{F}$ be any field. Let $P \in \mathbb{F}[X_1, \ldots, X_n]$. Then $\bar{i} \in \mathbb{Z}_0^n$ Hasse derivative of P is denoted by $P^{(\bar{i})}(X_1, \ldots, X_n)$ or $H^{(\bar{i})}(P)$ which is the coefficient of $Z^{i_1} \cdots Z_n^{i_n} =: \overline{Z}^{\bar{i}}$ in the polynomial $P(X_1 + Z_1, \ldots, X_n + Z_n)$. Thus

$$P(X_1 + Z_1, \ldots, X_n + Z_n) = \sum_{\bar{i}} P^{(\bar{i})}(X_1, \ldots, X_n) \cdot \overline{Z}^{\bar{i}}$$

For univariate polynomials we immediately get a closed form formula of k^{th} hasse derivative. Let $f \in \mathbb{F}[X]$ where $f(X) = \sum_{i=0}^{d} a_i X^i$. Then

$$f^{(k)}(X) = \sum_{i=0}^{d} \binom{k}{i} \cdot a_i X^{i-k}$$

From the definition we have the following observation

Observation D.1. *For any $f, f_1, f_2 \in \mathbb{F}_q[X_1, \ldots, X_n]$ and $\alpha \in \mathbb{F}$ and for any $\bar{k} \in \mathbb{Z}^n$ we have*

(i) $(f_1 + f_2)^{(\bar{k})}(X_1, \ldots, X_n) = f_1^{(\bar{k})}(X_1, \ldots, X_n) + f_2^{(\bar{k})}(X_1, \ldots, X_n)$

(ii) $(\alpha \cdot f)^{(\bar{k})}(X_1, \ldots, X_n) = \alpha \cdot f^{(\bar{k})}(X_1, \ldots, X_n)$.

Since throughout this report we only use hasse derivatives in the context of univariate polynomials we will only focus on univariate polynomials and see some properties of hasse derivatives. Here we take $\mathbb{F}$ to be any field.

Lemma D.1

(i) For $f_1, \ldots, f_t \in \mathbb{F}[X]$ then we have

$$(f_1 \cdot f_t)^{(n)}(X) = \sum_{\substack{n_1, \ldots, n_t \geq 0 \\ n_1 + \cdots n_t = n}} \prod_{i=1}^{t} f_i^{(n_i)}(X)$$

(ii) For any $\alpha \in \mathbb{F}$ we have

$$\left((X - c)^t\right)^{(n)} = \binom{t}{n} \cdot (X - c)^{t-n}.$$

(iii) For $0 \leq n \leq$ and $f, g \in \mathbb{F}[X]$ we have

$$\left(gf^t\right)^{(n)}(X) = \tilde{g} \cdot f^{t-n}(X)$$

where $\tilde{g} \in \mathbb{F}[X]$ *with* $\deg(\tilde{g}) \leq \deg(g) + n(\deg(f) - 1)$.

Proof:

(i) Since the hasse derivative operator is linear it is enough to show this property for monomials. So suppose $f_j(X) = X^{k_j}$ for all $j \in [t]$. Then $f_1 \cdots f_t = X^{\sum_{j=1}^{t} k_j}$. Therefore,

$$(f_1 \cdots f_t)^{(n)}(X) = \binom{\sum_{j=1}^{t} k_j}{t}, \qquad \prod_{j=1}^{t} f_j^{n_j}(X) = \prod_{j=1}^{t} \binom{k_j}{n_j} \text{ for all } n_j \in \mathbb{Z}_0, j \in [t]$$

So it is enough to show that

$$\binom{\sum_{j=1}^{t} k_j}{t} = \sum_{\substack{n_1,\ldots,n_t \geq 0 \\ n_1+\cdots n_t = n}} \prod_{j=1}^{t} \binom{k_j}{n_j}$$

But this is true by comparing the coefficients of X^n in $(X+1)^{\sum_{j=1}^{t} k_j} = \prod_{j=1}^{t} (X+1)^{k_j}$. So we have the result.

(ii) Now take $f_j(X) = X - c$ for all $j \in [t]$. Then $f_j^{(n)}(X) = 1$ if $n = 1$ and 0 if $n > 1$. Therefore, if we use the first part on $f_1, \ldots, f_j$ in the sum on right-hand side the only terms which survives are when each n_j are either 0 or 1. The number of such terms are $\binom{t}{n}$ and each term is $(X-c)^{t-n}$.

(iii) Again by using the first part we get

$$(gf^t)^{(n)}(X) = \sum_{\substack{n_0,n_1,\ldots,n_t \geq 0 \\ n_0+n_1+\cdots+n_t = n}} g^{(n_0)}(X) \prod_{j=1}^{t} f^{(n_j)}(X)$$

Now each summand is divisible by f^{t-n}. Hence, there exists some $\tilde{g} \in \mathbb{F}[X]$ such that $(gf^t)^{(n)}(X) = \tilde{g}(X) \cdot f^{t-n}(X)$. Furthermore,

$$\deg(\tilde{g}) = \deg\left((gf^t)^{(n)}\right) - (t-n)\deg(f) \leq \deg(g) + t\deg(f) - n - (t-n)\deg(f) = \deg(g) + (\deg(f) - 1)$$

So we have the Lemma. ∎

Now for our use-case in chapter 5 we use the polynomial $\Lambda(X) = X^q - X$ and its powers. And we compute its derivatives to do polynomial arithmetic.

Observation D.2. *For* $\Lambda(X)$ *and for any* $\ell \in \mathbb{Z}_0$

$$\Lambda^{(\ell)}(X) = \begin{cases} \Lambda(X) & \ell = 0 \\ -1 & \ell = 1 \\ 1 & \ell = q \\ 0 & \textit{otherwise} \end{cases}$$

Lemma D.2

Suppose $U(X), V(X) \in \mathbb{F}_q[X]$, $\deg(V) < q$, $i \geq 0$, *and* $U(X) = V(X) \cdot \Lambda^i(X)$. *Then for* $0 \leq \ell < q$,

$$U^{(\ell)}(X) \equiv (-1)^i \cdot V^{(\ell-i)}(X) \pmod{\Lambda(X)}.$$

Proof: Now using the Lemma D.1(i) we get

$$U^{(\ell)}(X) = \sum_{\substack{\ell_0,\ldots,\ell_i \geq 0 \\ \ell_0+\cdots+\ell_i = \ell}} V^{(\ell_0)}(X) \prod_{j=1}^{i} \Lambda^{(\ell_j)}(X)$$

Now if $\ell < i$ then in every summand one of $\ell_1, \ldots, \ell_i$ must equal 0. Therefore, all the terms are multiples of $\Lambda(X)$. So $U^{(\ell)}(X)$ is a multiple of $\Lambda(X)$.

Now from the above observation if $\ell \geq i$ then all but at most one of these terms is a multiple of $\Lambda(X)$. Since $\ell < q$ every term has all the $\ell_j < q$. Therefore, $\ell_1, \ldots, \ell_i \in \{0, 1\}$. If at least one of $\ell_1, \ldots, \ell_i$ equals to 0 then $\Lambda(X)$ survives as a factor of the term. So the only remaining term is $\ell_1 = \cdots = \ell_i = 1$. In this case $\ell_0 = \ell - i$ and therefore this term equals $V^{(\ell-i)}(X) \cdot (-1)^i$. Hence we have the Lemma. ∎

As a corollary of the Lemma D.1(i,iii) and the above Lemma we get the following result which is very usefull to us in chapter 5.

Corollary D.3

Let $f \in \mathbb{F}_q[X]$ (q odd) be a polynomial with $\deg(g) \leq d$. Then the ℓ^{th} derivative of $g^j(X)$ is of the form

$$(g^j)^{(\ell)}(X) = G_{g,j,\ell}(X) \cdot g(X)^{j-\ell}$$

where $G_{g,j,\ell}(X) \in \mathbb{F}_q[X]$ and $\deg(G_{g,j,\ell}) \leq d\ell - \ell$.

Moreover, for q even case if we take the polynomial $\mathrm{Tr} \circ g(X)$ where Tr is the absolute trace function from $\mathbb{F}_q \to \mathbb{F}_2$ then for any ℓ,

$$G^{(\ell)}(X) = \begin{cases} G(X) & \ell = 0 \\ G_{g,\ell}(X) & \ell > 0 \end{cases}$$

where $G_{g,\ell} \in \mathbb{F}_q[X]$ with $\deg(G_{g,\ell}) \leq d\ell$.

Theorem D.4

Suppose $\mathrm{char}(\mathbb{F}) = p$ where p is a prime. Let $Q(X, Y) \in \mathbb{F}[X, Y]$ be a polynomial and define $h(X) = Q\left(X, X^{p^s}\right)$ for some $s \in \mathbb{N}$. Then for all $0 \leq n < p^s$

$$h^{(n)}(X) = Q^{(n,0)}\left(X, X^{p^s}\right)$$

where $Q^{(i,0)}$ is the i^{th} partial derivative of Q with respect to X where we consider Q as a polynomial of $\mathbb{F}[Y][X]$.

Proof: Since hasse derivative is a linear operator it suffices to prove the above for monomials of the form $X^i Y^j$. So suppose $Q(X, Y) = X^i Y^j$. Then $h(X) = Q(X, X^{p^s}) = X^i \cdot \left(X^{p^s}\right)^j$. Now take $f_1(X) = X^i$ and $f_2 = \cdots = f_{k+1} = X^{p^s}$. Then $h = f_1 \cdots f_{k+1}$. So by Lemma D.1(i)

$$h^{(n)}(X) = \sum_{\substack{n_1, \ldots, n_{k+1} \geq 0 \\ n_1 + \cdots n_{k+1} = n}} \prod_{i=1}^{k+1} f_i^{(n_i)}(X)$$

Since $p \mid \binom{p^s}{n}$ for all $0 \leq n < p^s$ we have

$$(X^{p^s})^{(n)} = \binom{p^s}{X}^{p^s - n} \equiv 0$$

Therefore the only term which survives is when $n_1 = \cdots n_{k+1} = 0$ and $n_0 = n$. That means $h^{(n)} = Q^{(n,0)}(X, X^{p^s})$. ∎

Index