Exponential Sums

Lior Bary-Soroker

last update June 22, 2023

1 Introduction

The goal of the introduction is to give some examples for exponential sums, to suggest interesting questions about them, to highlight some intriguing phenomena, and to touch briefly some applications.

Let p be a prime number and

$$\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} = \{0, \ldots, p-1\}$$

be the finite field with p elements. Given two polynomials $f, g \in \mathbb{F}_p[x]$ with $g \neq 0$, we are interested in sums of the form

$$S = \sum_{\substack{x\in\mathbb{F}_p \\ g(x)\neq 0}} \exp\left(\frac{2\pi i}{p}\cdot\frac{f(x)}{g(x)}\right) = \sum_{\substack{x\in\mathbb{F}_p \\ g(x)\neq 0}} e_p(f(x)/g(x))) \in \mathbb{C}.$$

As $e(x) := \exp(2\pi i x)$ is 1-periodic, the function $e_p(x) := e(x/p)$ is well-defined modulo $p\mathbb{Z}$. Hence, S is also well-defined.

Maybe another remark is in place, just to clarify notation: we invert modulo p, so if, for example $p = 7$, and $g(x) = 3$, then $1/g(x) = 5$.

Sums of the form S arise frequently in number theory.

Here are some typical questions and goals we want to study about such sums.

1. Is there a closed expression for S?

2. S has a trivial upper bound: $|S| \leq p$; can we do better $|S| \leq \theta^{-1}p$, where $\theta > 1$?

3. Can one obtain lower bounds? This is more challenging, and luckily, appears less in applications.

4. Assume we have a family of exponential sums S_ℓ; how does it varies as a function of ℓ?

5. $e_p \colon \mathbb{Z}/p\mathbb{Z} \to \mathbb{C}$ is an additive character of $\mathbb{F}_p$. Can we take multiplicative character? Mixed sums?

6. What happens for higher dimensions?

1.1 Linear sums

The simplest example if a linear sum, that is to say, $f(x) = ax$ and $g(x) = 1$. So

$$S_a = \sum_{x \in \mathbb{F}_p} e_p(ax).$$

Put $w = e_p(a)$. Then, $S_a = 1 + w + \cdots + w^{p-1}$. So $S_a = \begin{cases} p, & a \equiv 0(p) \\ 0, & a \not\equiv 0(p). \end{cases}$ The same argument work for any modulus (and not only prime):

Theorem 1.1.

$$\sum_{x=0}^{m-1} e_m(ax) = \begin{cases} m, & a \equiv 0(m) \\ 0, & a \not\equiv 0(m). \end{cases}$$

So in this simple case, we have an explicit expression.

Let see a fun application of the trivial Theorem 1.1.

Theorem 1.2. *Let X be any finite set of integers and let $f\colon X \to \mathbb{F}_p$ be a function. Denote by $N_k(a)$ the number of solutions of*

$$f(x_1) + \cdots + f(x_k) \equiv f(x_{k+1}) + \cdots + f(x_{2k}) + a \mod p$$

with $x_i \in X$. Then $N_k(a) \leq N_k(0)$.

Proof. By Theorem 1.1,

$$\begin{aligned} N_k(a) &= \sum_{x_1,\ldots,x_{2k} \in X} \frac{1}{p} \sum_{c=0}^{p-1} e_p(c(f(x_1) + \cdots + f(x_k) - f(x_{k+1}) - \cdots - f(x_{2k}) - a)) \\ &= \sum_{x_1,\ldots,x_{2k} \in X} \frac{1}{p} \sum_{c=0}^{p-1} e_p(-ca) \prod_{i=1}^{k} e_p(cf(x_i)) \prod_{i=1}^{k} e_p(-cf(x_{k+i})) \\ &= \frac{1}{p} \sum_{c=0}^{p-1} e_p(-ca) \left(\sum_{x \in X} e_p(cf(x)) \right)^k \left(\sum_{x \in X} e_p(-cf(x)) \right)^k \end{aligned}$$

Since $f(x)$ is real and since $N_k(a)$ is a nonnegative integer, we conclude that

$$N_k(a) = |N_k(a)| = \left| \frac{1}{p} \sum_{c=0}^{p-1} e_p(-ca) \left| \sum_{x \in X} e_p(cf(x)) \right|^{2k} \right| \leq \frac{1}{p} \sum_{c=0}^{p-1} \left| \sum_{x \in X} e_p(cf(x)) \right|^{2k} = N_k(0),$$

as needed. □

Exercise 1.1. Conclude that $N_k(0) \geq \frac{\#X^{2k}}{p}$. (Hint: Compute the expected value of $N_k(a)$ as a is chosen uniformly from $\mathbb{F}_p$.)

We may show that the values $N_k(a)$ are close to their expected value, if we can bound exponential sums. Indeed, in the $N_k(a) = \frac{1}{p}\sum_{c=0}^{p-1} e_p(-ca)\left|\sum_{x\in X} e_p(cf(x))\right|^{2k}$, we notice that $c = 0$ contributes the expected value $\#X^{2k}/p$. Hence

$$|N_k(a) - \#X^2k/p| \leq \frac{1}{p}\sum_{c=1}^{p-1}\left|\sum_{x\in X} e_p(cf(x))\right|^{2k}.$$

Hence, had we had a nontrivial bound on the exponential sums $\left|\sum_{x\in X} e_p(cf(x))\right| \leq \theta^{-1}\#X$, $c \neq 0(p)$, $\theta > 1$, we could have a concentration result: $|N_k(a) - \#X^{2k}/p| \leq \theta^{-2k}\#X^{2k}$.

1.2 Gauss quadratic sums

Gauss sums are the simplest example apart from the linear sums. It is one of the oldest sums considered, and one of the few interesting ones with explicit expression. In this case, we take $f(x) = ax^2$ and $g(x) = 1$. We define the Gauss sum to be

$$G_a = S = \sum_{x\in\mathbb{F}_p} e_p(ax^2).$$

If $a = 0(p)$, then $G_0 = p$. So from now on assume $a \neq 0(p)$. If $p = 2$, then $G_1 = 0$. So assume also that $p \neq 2$.

Let $\omega = e_p(1)$, let

$$\left(\frac{y}{p}\right) = \begin{cases} 0, & \text{if } y = 0 \\ 1, & \text{if } y \in \mathbb{F}_p^{\times 2} \\ -1, & \text{otherwise.} \end{cases}$$

be the Legendre symbol, and let

$$T_a = \sum_{y\in\mathbb{F}_p}\left(\frac{y}{p}\right)\omega^{ay}.$$

Note that T_a is a mixed sum.

We claim that $G_a = T_a$ if $a \neq 0(p)$. Indeed, since

$$1 + \left(\frac{y}{p}\right) = \#\{x \in \mathbb{F}_p : x^2 = y\},$$

we have that

$$\begin{aligned} G_a &= \sum_{x\in\mathbb{F}_p} e_p(ax^2) = \sum_{y\in\mathbb{F}_p}\sum_{x^2=y} e(ay) = \sum_{y\in\mathbb{F}_p}\#\{x\in\mathbb{F}_p : x^2 = y\}\omega^{ay} \\ &= \sum_{y\in\mathbb{F}_p}\left(1 + \left(\frac{y}{p}\right)\right)\omega^{ay} = \sum_{y\in\mathbb{F}_p}\omega^{ay} + T_a = T_a, \end{aligned}$$

since $a \neq 0$.

Lemma 1.3. *Assume $a \neq 0(p)$ and $p > 2$. Then*

1. $T_a = \left(\frac{a}{p}\right) T_1$.

2. $G_a^2 = T_a^2 = (-1)^{\frac{p-1}{2}} p$.

Proof. We have

$$\left(\frac{a}{p}\right) T_a = \sum_{y \in \mathbb{F}_p} \left(\frac{ay}{p}\right) \omega^{ay} = \sum_{x} \left(\frac{x}{p}\right) \omega^x = T_1,$$

as needed for 1.

By 1., it suffices to prove 2. for $a = 1$ since $T_a^2 = T_1^2$. For any $b \neq 0(p)$, by 1., we have $T_b T_{-b} = \left(\frac{-1}{p}\right) T_1^2$. Hence

$$\sum_{b \in \mathbb{F}_p^\times} T_b T_{-b} = \left(\frac{-1}{p}\right)(p-1) T_1^2.$$

On the other hand,

$$\begin{aligned} \sum_{b \in \mathbb{F}_p^\times} T_b T_{-b} &= \sum_{b \in \mathbb{F}_p^\times} \left(\sum_{x \in \mathbb{F}_p} \left(\frac{x}{p}\right) \omega^{bx} \right) \left(\sum_{y \in \mathbb{F}_p} \left(\frac{y}{p}\right) \omega^{-by} \right) \\ &= \sum_{x,y \in \mathbb{F}_p} \left(\frac{xy}{p}\right) \sum_{b \in \mathbb{F}_p^\times} \omega^{b(x-y)} = p \sum_{x \in \mathbb{F}_p} \left(\frac{x^2}{p}\right) = p(p-1). \end{aligned}$$

So we conclude that $\left(\frac{-1}{p}\right) T_1^2 = p$, as needed. □

Corollary 1.4. *We have* $G_1 = T_1 = \begin{cases} \epsilon_p \sqrt{p} & \text{if } p = 1(4) \\ \epsilon_p i \sqrt{p} & \text{if } p = 3(4) \end{cases}$, *where* $\epsilon_p \in \{\pm 1\}$.

The sign ϵ_p is always 1. We will not prove it here. Here is a quote from Wikipedia:

> *"In fact, the identity $g(1;p)^2 = \left(\frac{-1}{p}\right) p$ was easy to prove and led to one of Gauss's proofs of quadratic reciprocity. However, the determination of the sign of the Gauss sum turned out to be considerably more difficult: Gauss could only establish it after several years' work. Later, Dirichlet, Kronecker, Schur and other mathematicians found different proofs."*

Exercise 1.2. The corollary remains true if we replace the odd prime p by any odd integer.

Gauss sums are used to prove the following

Theorem 1.5 (The quadratic reciprocity law). *Let $p \neq q$ be odd primes. Then* $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$.

Quick proof using the explicit expression for the quadratic Gauss sum. For a, n, let

$$G_{a,n} = \sum_{x \in \mathbb{Z}/n\mathbb{Z}} e(ax^2/n).$$

By the Chinese Remainder Theorem, each $x \in \mathbb{Z}/pq\mathbb{Z}$ may be written uniquely as $x = px_1 + qx_2$, where $x_1 \in \mathbb{F}_q$ and $x_2 \in \mathbb{F}_p$. Therefore,

$$\begin{aligned} G_{1,pq} &= \sum_{x_1 \in \mathbb{F}_q} \sum_{x_2 \in \mathbb{F}_p} e\left(\frac{(px_1 + qx_2)^2}{pq} \right) = \sum_{x_1 \in \mathbb{F}_q} \sum_{x_2 \in \mathbb{F}_p} e_q(px_1^2) e_p(qx_2^2) e(2x_1x_2) \\ &= \sum_{x_1 \in \mathbb{F}_q} e_q(px_1^2) \sum_{x_2 \in \mathbb{F}_p} e_p(qx_2^2) = G_{p,q} G_{q,p} = \left(\frac{p}{q} \right) \left(\frac{q}{p} \right) G_{1,p} G_{1,q}. \end{aligned}$$

This finishes the proof if we use the theorem

$$G_{1,n} = \begin{cases} \sqrt{n} & n = 1(4) \\ i\sqrt{n} & n = 3(4). \end{cases}$$

□

We remark, that Gauss only used that $G_a^2 = (-1)^{\frac{p-1}{2}}$ to deduce the QRL (see Exercise 1.4).

1.3 Kloosterman sums

In this sum, we take $f(x) = ax^2 + b$ and $g(x) = x$; so that

$$K(a, b; p) = \sum_{x=1}^{p-1} e_p(ax + bx^{-1}).$$

If $ab = 0$, we have a linear sum.

1.3.1 Upper bounds

Getting nontrivial upper bounds is the most crucial ingredient in the proof of the following beautiful

Theorem 1.6 (Kloosterman)**.** *Let $a_1, \ldots, a_4$ be positive integers. Then, for any N sufficiently large, there exists a solution $(x_1, \ldots, x_4) \in \mathbb{Z}^4$ for the Diophantine equation*

$$a_1x_1^2 + \cdots + a_4x^2 = N,$$

provided there are no local obstructions.

Today the best bound comes from Weil's resolution of the Riemann hypothesis for curves which gives $|K(a,b;p)| \leq 2\sqrt{p}$ for $a,b \not\equiv 0(p)$.

This is close to the best possible, as Kloosterman proved that there exists a,b with $|K(a,b;p)| \geq \sqrt{2p-2}$ (Exercise 1.5).

We now present a weaker bound due to Kloosterman.

Theorem 1.7. *Let $a,b \in \mathbb{F}_p^\times$, then $|K(a,b;p)| \leq 2p^{3/4}$.*

Proof. We try to understand $K(a,b;p)$ as a family and exploit its symmetries. Let

$$M_k = \sum_{a,b\in\mathbb{F}_p^\times} |K(a,b;p)|^{2k} \tag{1}$$

be the $2k$ moment of the Kloosterman sums. We have the symmetry $K(a,b;p) = K(ac,bc^{-1};p)$, $c \not\equiv 0(p)$. We get that

$$(p-1)|K(a,b;p)|^{2k} = \sum_{c=1}^{p-1} |K(ac,bc^{-1};p)|^{2k} \leq M_k,$$

so

$$|K(a,b;p)| \leq \left(\frac{M_k}{p-1}\right)^{1/2k}. \tag{2}$$

Next, we estimate M_k for $k \geq 1$. Since $K(0,b;p) = K(b,0;p)$, we have

$$M_k = \sum_{a,b\in\mathbb{F}_p} |K(a,b;p)|^{2k} - 2\sum_{a\in\mathbb{F}_p^\times} K(a,0;p)^{2k} - |K(0,0;p)|^{2k}.$$

If $a \not\equiv 0$, then $K(a,0;p) = -1$, by Theorem 1.1. Even more trivially, $K(0,0;p) = p-1$.

We expand the Kloosterman sum, to get

$$M_k = \sum_{a,b\in\mathbb{F}_p} \sum_{x,\mathbf{y}\in\mathbb{F}_p^{\times k}} e_p\left(a\sum_{i=1}^{k}(x_i - y_i)\right) e_p\left(b\sum_{i=1}^{k}(x_i^{-1} - y_i^{-1})\right) - 2(p-1) - (p-1)^{2k}.$$

(Here $x = (x_1,\ldots,x_k)$ and $\mathbf{y} = (y_1,\ldots,y_k)$.)

Changing order of summation in the first sum, and using Theorem 1.1, we get that

$$M_k = p^2N_k - 2(p-1) - (p-1)^{2k},$$

where N_k is the number of solutions $x,\mathbf{y} \in \mathbb{F}_q^{\times k}$ to

$$\begin{cases} x_1 + \cdots + x_k = y_1 + \cdots + y_k \\ x_1^{-1} + \cdots + x_k^{-1} = y_1^{-1} + \cdots + y_k^{-1}. \end{cases}$$

Obviously, $N_1 = p-1$. So,

$$M_1 = p^2(p-1) - 2(p-1) - (p-1)^2 = p^3 - 2p^2 + 1.$$

If we plug this in (2), we are worse than the trivial bound.

For $k = 2$, there are $2(p-1)^2 - (p-1)$ obvious solutions $\mathbf{y} = (x_1, x_2)$ or $\mathbf{y} = (x_2, x_1)$. Moreover, there are $(p-1)^2$ with $x_1 + x_2 = 0$ or $y_1 + y_2 = 0$, out of which $2(p-1)$ have already been counted.

If (x_1, x_2, y_1, y_2) is any other solution, then

$$\begin{cases} x_1 + x_2 = y_1 + y_2 \\ y_1 y_2 (x_1 + x_2) = x_1 x_2 (y_1 + y_2). \end{cases}$$

Since $x_1 + x_2 \neq 0$, we get that $y_1 y_2 = x_1 x_2$. Hence, both the pairs x_1, x_2 and y_1, y_2 are solutions of the quadratic equation $X^2 - (x_1 + x_2)T + x_1 x_2 = 0$. Hence either $(x_1, x_2) = (y_1, y_2)$ or $(x_1, x_2) = (y_2, y_1)$, contradiction.

From this we get that $N_2 = 2(p-1)^2 - (p-1) + (p-1)^2 - 2(p-1) = 3(p-1)(p-2)$ and so $\frac{M_2}{p-1} = 3p^2(p-2) - 2 - (p-1)^3 = 2p^3 - 3p^2 - 3p - 1$. Plugging this into (2), we get

$$|K(a, b; p)| \leq 2p^{3/4},$$

as needed. □

1.3.2 Lower bounds

We already stated a strong lower bound for $\max |K(a, b; p)$ in 2. We aim for a uniform lower bound.

Let's derive some more properties of Kloosterman sums. First of, obviously we have $\overline{K(a, b; p)} = K(a, b; p)$, so $K(a, b; p) \in \mathbb{R}$.

Lemma 1.8. *For any $a, b \in \mathbb{F}_p^\times$ we have $K(a, b; p) \neq 0$.*

Proof. The proof uses the algebraic properties of cyclotomic extensions: Let $\zeta = e_p(1)$ be the primitive p-th root of unity, let $K = \mathbb{Q}(\zeta)$ be the cyclotomic field. Then $[K : \mathbb{Q}] = p - 1$, $K/\mathbb{Q}$ is Galois, $\mathrm{Gal}(K/\mathbb{Q}) = (\mathbb{Z}/p\mathbb{Z})^*$ where $(a, p) = 1$ defined the automorphism given by $\zeta \mapsto \zeta^a$ (i.e., $\sum a_i \zeta^i \mapsto \sum a_i \zeta^{ai}$), the ring of integers is $O_K = \mathbb{Z}[\zeta]$. The prime p is totally ramified in K, that means that there exists a prime ideal $\mathfrak{p}$ of O_K such that $pO_K = \mathfrak{p}^{p-1}$. Moreover, $\mathfrak{p} = (1 - \zeta, p) = \{x(1 - \zeta) + yp : x, y \in O_K\}$.

We reduce modulo $\mathfrak{p}$ to get that $1 \equiv \zeta \mod \mathfrak{p}$. Hence,

$$K(a, b; p) = \sum_{x \in \mathbb{F}_p^\times} \zeta^{ax + bx^{-1}} = \sum_{x \in \mathbb{F}_p^\times} 1 = -1 \mod \mathfrak{p}.$$

In particular, $K(a, b; p) \neq 0$. □

We are now ready to give a uniform lower bound on Kloosterman sums.

Theorem 1.9 (Fouvry)**.** *For any $a, b \in \mathbb{F}_p^\times$,*

$$|K(a, b; p)| \geq \left(\frac{1}{2p^{2/3}} \right)^{p-2}.$$

Proof. We adopt the notation of the proof the the lemma. For any $1 \leq m \leq p-1$, the conjugate of $K(a,b;p)$ by the corresponding automorphism of $K/\mathbb{Q}$ is $K(ma,mb;p)$. On the one hand, $N_{K/\mathbb{Q}}K(a,b;p) := \prod_{1\leq m\leq p-1} K(am,bm;p) \in \mathbb{Q}$ as it in invariant to the Galois action; on the other hand, $N_{K/\mathbb{Q}}K(a,b;p) \in O_K$ as a product of algebraic integers. By Gauss Lemma, $O_K \cap \mathbb{Q} = \mathbb{Z}$, so $N_{K/\mathbb{Q}}K(a,b;p) \in \mathbb{Z}$. In particular,

$$1 \leq |N_{K/\mathbb{Q}}K(a,b;p)| = K(a,b;p)(2p^{3/4})^{p-2},$$

where we applied the Kloosterman upper bound to all terms with $m \neq 1$. This finishes the proof. □

Remark 1. It is not clear how to do better, after replacing the Kloosterman bound by the Weil bound. Kloosterman sums illustrate Question 2 and Question 3.

Remark 2. It is open whether, for example, $K(1,1;p) > 0$ infinitely often.

The last remark leads us to equidistribution questions.

1.4 Equidistribution

Let's discuss the questions here through the following example: Let $f(x) = ax^3 + bx$, $a \neq 0$ and $g(x) = 1$, so that

$$S(a,b;p) = \sum_{x\in\mathbb{F}_p} e_p(ax^3 + bx).$$

Then $S(a,b;p) \in \mathbb{R}$. The Weil bounds give that $|S(a,b;p)| \leq 2\sqrt{p}$, so it is natural to ask how the numbers

$$\theta_p^{a,b} = \frac{S(a,b;p)}{2\sqrt{p}}$$

distribute on $[-1,1]$. Two natural limits:

1. Horizontal distribution: Fix $a,b \in \mathbb{Z}$, $a \neq 0$ and vary p.
2. Vertical distribution: Fix large p and vary $a,b \in \mathbb{F}_p^\times \times \mathbb{F}_p$.

In both cases, it is conjectured that $\theta_p^{a,b}$ become equidistributed for the Sato-Tate measure, see Figure 1. The vertical case is a Theorem by Livne:

Theorem 1.10 (Livne). *For every interval $[\alpha,\beta] \subseteq [-1,1]$, we have*

$$\lim_{p\to\infty} \frac{1}{p(p-1)} \#\{a,b \in \mathbb{F}_p^\times \times \mathbb{F}_p : \theta_p^{a,b} \in [\alpha,\beta]\} = \frac{2}{\pi}\int_\alpha^\beta \sqrt{1-x^2}dx.$$

Let $f = \mathbb{1}_{[\alpha,\beta]}$ be the indicator function for the interval, then the theorem states that

$$\lim_{p\to\infty} \frac{1}{p(p-1)} \sum_{a,b\in\mathbb{F}_p^\times\times\mathbb{F}_p} f(\theta_p^{a,b}) = \frac{2}{\pi}\int_{-1}^1 f(x)\sqrt{1-x^2}dx. \tag{3}$$

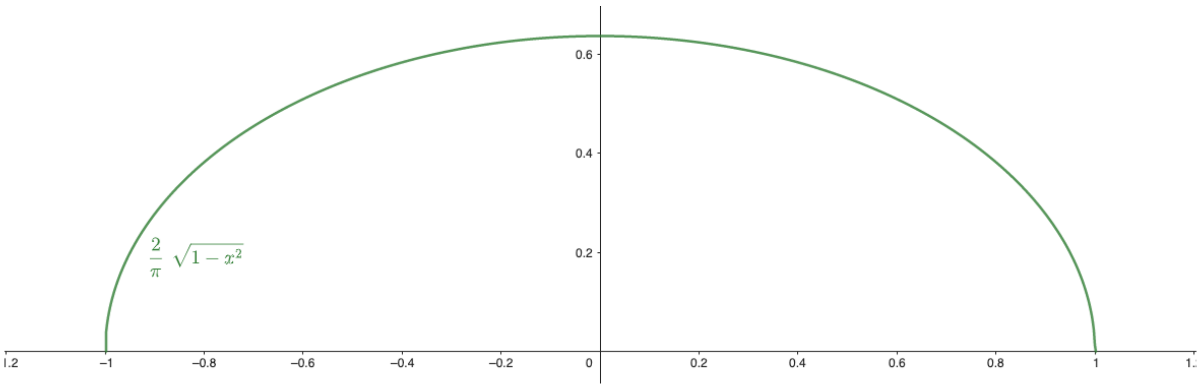

Figure 1: The Sato-Tate measure

Indicator functions are hard to work with using harmonic analysis, because they are not smooth. So one tries to prove (3) for smooth functions f. If (3) holds true for all smooth f, then equidistribution holds (that is (3) holds for indicator functions of intervals).

Furthermore, since we may approximate smooth functions by polynomials, and since both sides of(3) are linear in f, we conclude that it suffices to consider $f(x) = x^k$. Let us state this general principle in our case.

Proposition 1.11. *To prove Theorem 1.10 it suffices to prove*

$$\lim_{p\to\infty} \frac{1}{p(p-1)} \sum_{a,b\in\mathbb{F}_p^\times\times\mathbb{F}_p} (\theta_p^{a,b})^k = \frac{2}{\pi}\int_{-1}^{1} x^k\sqrt{1-x^2}dx,$$

for every $k \geq 0$.

The left-hand side of the equation in the proposition is easy to compute:

Exercise 1.3. Calculate $c_k = \frac{2}{\pi}\int_{-1}^{1} x^k\sqrt{1-x^2}dx$ and show that $c_k = 0$ if k is odd and $c_k = \frac{1}{2^k(k/2+1)}\binom{k}{k/2}$ if k is even.

Denote the right-hand side by

$$VM_k = \frac{1}{p(p-1)} \sum_{a,b\in\mathbb{F}_p^\times\times\mathbb{F}_p} (\theta_p^{a,b})^k \tag{4}$$

so by expanding and rearranging we get

$$
\begin{aligned}
VM_k &= \frac{1}{2^k p^{k/2+1}(p-1)} \sum_{a\neq 0}\sum_{b} \big(\sum_{x} e_p(ax^3+b)\big)^k \\
&= \frac{1}{2^k p^{k/2+1}(p-1)} \sum_{a\neq 0}\sum_{b} \sum_{x_1,\dots,x_k} \prod_{i} e_p(ax_i^3+bx_i) \\
&= \frac{1}{2^k p^{k/2+1}(p-1)} \sum_{a\neq 0}\sum_{b} \sum_{x_1,\dots,x_k} e_p(a\sum_{i} x_i^3) e_p(b\sum_{i} x_i) \\
&= \frac{1}{2^k p^{k/2+1}(p-1)} \sum_{x_1,\dots,x_k} \Big(\sum_{a\neq 0} e_p(a\sum_{i} x_i^3)\Big)\Big(\sum_{b} e_p(b\sum_{i} x_i)\Big)
\end{aligned}
$$

By the linear sum (Theorem 1.1), we conclude that

$$
\sum_{a\neq 0} e_p(a\sum_{i} x_i^3) = \begin{cases} -1, & \sum x_i^3 \neq 0 \\ p-1, & \sum x_i^3 = 0 \end{cases} \quad \text{and} \quad \sum_{b} e_p(b\sum_{i} x_i) = \begin{cases} 0 & ,\sum x_i \neq 0 \\ p, & \sum x_i = 0 \end{cases}
$$

Hence,

$$
VM_k = \frac{1}{2^k p^{k/2+1}(p-1)} p\delta_{\sum x_i=0}(p\delta_{\sum x_i^3=0} - 1) = \frac{1}{2^k p^{k/2+1}(p-1)}(p^2 X_k(\mathbb{F}_p) - p^k).
$$

Here $\#X_k(\mathbb{F}_p)$ is the number of solutions of $\sum x_i^3 = \sum x_i = 0$. Thus, equidistribution questions leads us naturally to seek precise estimates for the number of points on higher dimensional varieties. (Questions 4, 6).

Exercise 1.4. Compute VM_k for $k = 2, 3$ and show that $VM_k \to c_k$ in cases.

1.5 Exercises

1. In the notation of Theorem 1.2, conclude that $N_k(0) \geq \frac{\#X^{2k}}{p}$. (Hint: Compute the expected value of $N_k(a)$ as a is chosen uniformly from $\mathbb{F}_p$.)

2. Prove Theorem 1.2 using Cauchy-Schwartz (hint: think of the x_i-s as iid random variables, conclude that $X = f(x_1) + \cdots + f(x_k)$ and $X' = f(x_{k+1}) + \cdots + f(x_{2k})$ have the same distribution, and compute both the relevant probabilities $N_k(a)/p$ explicitly).

3. Show that Corollary 1.4 remains true if we replace the odd prime p by any odd integer.

4. In this exercise, we will prove the QRL (Theorem 1.5) using Lemma 1.3. Let $g_p = G_{1,p} = \sum_{x\in\mathbb{F}_p}\left(\frac{x}{p}\right)\omega^x$, where $\omega = e_p(1)$. We also let $p^* = \left(\frac{-1}{p}\right)p = (-1)^{\frac{p-1}{2}}p$.

   a) Show that $\mathbb{Q}(g_p)$ is a quadratic extension of $\mathbb{Q}$ contained in the cyclotomic field $\mathbb{Q}(\omega)$.

   b) Use $g_p^2 = p^*$ to prove that $g_p^{q-1} = \left(\frac{p^*}{q}\right) \mod q$. (You may want to use modular arithmetic on the ring of integers $\mathbb{Z}[g_p]$ modulo some prime lying above q.)

   c) Use that raising to q-th power is an automorphism on fields of characteristic q to compute that $g_p^q = \left(\frac{q}{p}\right)g_p \mod q$. (You may want to use item 1 in Lemma 1.3.)

   d) Deduce the QRL.

5. In the notation (1), note that $M_2 \leq \max_{a,b\in\mathbb{F}_p^\times}|K(a,b;p)|^2 M_1$, and deduce that there exist $a, b \in \mathbb{F}_p^\times$ such that $|K(a,b;p)| \geq \sqrt{2p-2}$.

6. Calculate $c_k = \frac{2}{\pi}\int_{-1}^{1} x^k\sqrt{1-x^2}dx$ and show that $c_k = 0$ if k is odd and $c_k = \frac{1}{2^k(k/2+1)}\binom{k}{k/2}$ if k is even.

7. Compute VM_k (see (4)) and show that $VM_k \to c_k$ for $k = 2, 3$ (where c_k is the constant from the previous exercise).

2 Finite Fields

2.1 The additive structure

Let F be a field with unit 1_F. Then we have a unique ring homomorphism $\lambda\mathbb{Z} \to F$ defined by $\lambda(n) = 1_F + \cdots + 1_F$. Then $\ker(\lambda) = (n\mathbb{Z})$, for some $n \geq 0$. We call $n = \mathrm{Char}(F)$ the characteristic of F.

If $n = 0$, then $\mathbb{Z} \subseteq F$ (via the canonical embedding λ) hence $\mathbb{Q} \subseteq F$.

Otherwise, $n = p$ is a prime number since the image is a domain (Exercise 2.1) and by the isomorphism theorem $F \supset \lambda(\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z} =: \mathbb{F}_p$.

If F is finite of cardinality q, then $\mathrm{Char}(F) = p$, and $\mathbb{F}_p \subseteq F$. Since F is a vector space over $\mathbb{F}_p$ of finite dimension $f = [F : \mathbb{F}_p] := \dim_{\mathbb{F}_p} F$, we conclude that $q = p^f$.

Theorem 2.1. *Let $\mathbb{F}_q$ be a finite field with q elements. Then $q = p^f$, for p prime and $f \in \mathbb{N}$. Moreover, for every such q there exists a unique field with q elements in an algebraic closure of $\mathbb{F}_p$, and it is defined as the set of solutions of $X^q - X = 0$.*

Proof. We already showed the first claim. For the second, note that as the multiplicative group $\mathbb{F}_q^\times$ is of order $q-1$, by Lagrange theorem $x^{q-1} - 1 = 0$ for all $x \in \mathbb{F}_q^\times$. Hence, $x^q - x = 0$ for all $x \in \mathbb{F}_q$. This proves the uniqueness of $\mathbb{F}_q$.

For the existence, write $g = X^q - X$. So $g' = -1$, hence $\gcd(g, g') = 1$. This implies that g has exactly q solutions in the algebraic closure. So it remains to show that the set of solution is a field. And indeed, since $x \mapsto x^p$ is a homomorphism in characteristic p, so is the compositum $x \mapsto x^q$. So if $g(x) = g(y) = 0$, then $g(x + y) = x^q + y^q - x - y = 0$ and $g(xy) = x^q y^q - xy = xy - xy = 0$. (We do not have to check closeness for inversion as it is a finite set, but we may easily do so). □

Definition 2.2. We call the field automorphism $x \mapsto x^q$ of any field of characteristic p the q-Frobenius and denote it by ϕ_q. Then, $\mathbb{F}_q$ equals the set of fixed points of ϕ_q in the algebraic closure.

Theorem 2.3. *$\mathbb{F}_{p^n} \subseteq \mathbb{F}_{p^m}$ if and only if $n \mid m$. In particular, $\mathbb{F}_q$ has a unique extension of any degree.*

Proof. If $\mathbb{F}_{p^n} \subseteq \mathbb{F}_{p^m}$, then $p^m = (p^n)^{[\mathbb{F}_{p^m}:\mathbb{F}_{p^n}]}$. So, $n \mid m$.

Vice-versa, assume $m = kn$, so if we put $q = p^n$, we have $q^k = p^m$. Let $x \in \mathbb{F}_{p^n}$, then $\phi_q(x) = x$. Since, $\phi_{p^m} = \phi_q \circ \cdots \circ \phi_q$, we get that $\phi_{p^m}(x) = x$. Hence, $x \in \mathbb{F}_{p^m}$. □

The extension $\mathbb{F}_q \subseteq \mathbb{F}_{q^n}$ is of degree n and Galois (as the splitting field of the separable polynomial $X^{q^n} - X$). In particular, $\phi_q = \phi_q|_{\mathbb{F}_{q^n}} \in \mathrm{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q)$. Also, since $\phi_q(x) = x$ if and only if $x \in \mathbb{F}_q$, we get that $\mathbb{F}_{q^n}^{\langle \phi_q \rangle} = \mathbb{F}_q$. Hence,

Corollary 2.4. *The extension $\mathbb{F}_q \subseteq \mathbb{F}_{q^n}$ is Galois with cyclic Galois group having a distinguished generator: $\mathrm{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q) = \langle \phi_q|_{\mathbb{F}_{q^n}} \rangle \cong \mathbb{Z}/n\mathbb{Z}$.*

2.2 The multiplicative structure

The goal is to prove the following

Theorem 2.5. *The multiplicative group $\mathbb{F}_q^\times$ is cyclic.*

We first prove two lemmas.

Lemma 2.6. *Let $\varphi(n) = \#\{1 \leq m \leq n : \gcd(n, m) = 1\}$ be the Euler totient function. Then $n = \sum_{d|n} \phi(d)$.*

Sketch of Proof. Use that $h(n) = \sum_{d|n} \phi(d)$ is multiplicative (Exercise 2.2) and check the equality on prime powers:

$$h(p^k) = \sum_{1\leq \ell \leq k} p^k(1-p^{-1}) + 1 = \frac{p^{k+1}-p}{p-1}(1-p^{-1}) + 1 = p^k.$$

A different proof is to use the partition $\{1 \leq m \leq n\} = \bigcup_{d|n} X_d$, where $X_d = \{1 \leq m \leq n : \gcd(m,n) = d\}$ and noting that X_d is in bijection with $\{1 \leq m \leq n/d : \gcd(m,n/d) = 1\}$ (by $m \in X_d \mapsto m/d$). □

Lemma 2.7. *Let H be a finite group of order n. Assume that for $\#\{x \in H : x^d = 1\} \leq d$, for all $d \mid n$. Then H is cyclic.*

Proof. Let $d \mid n$. If there exists x of order d, then every of the d elements x^i also satisfies $(x^i)^d = 1$. By assumption, there are no others, so if $y^d = 1$, then $y \in \langle x \rangle$. In particular, the number of elements n_d of order d, is $\phi(d)$ or 0.

We conclude that $n = |H| = \sum_{d|n} n_d \leq \sum_{d|n} \phi(d) = n$. If one of the $n_d = 0$, then we have contradiction. Hence, $n_d > 0$ for all d, in particular $n_n > 0$, so there exists an element of order n, hence H is cyclic. □

Proof of Theorem 2.5. Since $\mathbb{F}_q$ is a field, $\#\{x^m - 1 = 0\} \leq m$. So by the lemma $\mathbb{F}_q^\times$ is cyclic. □

2.3 Chevalley-Warning Theorem

A key ingredient in understanding exponential sums, is to study solutions to equations over finite fields. For example, the equation $X^2 + Y^2 = 0$ has a nontrivial solution in $\mathbb{F}_q$ if and only if $q \equiv 1(4)$. But the equation $X^2 + Y^2 + Z^2 = 0$ always has a nontrivial solution. It turns out that once the number of variables is large enough, there is a solution over any finite field:

Theorem 2.8. *Let $f_1, \dots, f_k \in \mathbb{F}_q[X_1, \dots, X_n]$ be polynomials such that $\sum \deg f_i < n$. Then*

$$N = \#\{x \in \mathbb{F}_q^n : f_1(x) = \cdots = f_k(x) = 0\} \equiv 0 \mod p.$$

In particular, if the f_i are homogenous of positive degree (or if the free coefficient is zero), then there exists a nontrivial solution.

We present an easy proof by Ax that only uses what we've learned so far on $\mathbb{F}_q$.

Proof. Let $P = \prod_{i=1}^k (1 - f_i^{q-1})$ and $x \in \mathbb{F}_q^n$. If $f_1(x) = \cdots = f_k(x) = 0$, then $P(x) = 1$. Otherwise, for some i, $f_i(x) \neq 0$, so $f_i(x)^{q-1} = 1$, and $P(x) = 0$. We conclude that

$$N = \sum_{x \in \mathbb{F}_{q^n}} P(x) \mod p.$$

By hypothesis, $\deg P < n(q-1)$. That means that the monomials appearing in P are of the form $X^{\mathbf{a}} = X_1^{a_1} \cdots X_n^{a_n}$, with $\sum a_i < n(q-1)$. Hence, it suffices to prove that $\sum_{x \in \mathbb{F}_q^n} x^{\mathbf{a}} = 0$, for any such $\mathbf{a}$. Suppose w.l.o.g. that $a_1 < q-1$. So it suffices to prove that $\sum_{a_1 \in \mathbb{F}_q} x_1^{a_1} = 0$.

If $a_1 = 0$, then $\sum_{a_1 \in \mathbb{F}_q} x_1^{a_1} = 0$, and hence $\sum_{x \in \mathbb{F}_q^n} x^{\mathbf{a}} = 0$, and we are done.

Otherwise, $1 \le a_1 < q-1$. So, there exists $y \in \mathbb{F}_q^\times$ such that $y^{a_1} \ne 1$ (Theorem 2.5). So,

$$y^{a_1} \sum_{x_1 \in \mathbb{F}_q} x_1^{a_1} = \sum_{x_1 \in \mathbb{F}_q} (yx_1)^{a_1} = \sum_{x_1 \in \mathbb{F}_q} x_1^{a_1}.$$

Thus $\sum_{x_1 \in \mathbb{F}_q} x_1^{a_1} = 0$. □

2.4 The trace and norm maps

Let $\mathbb{F}_{q^n}/\mathbb{F}_q$. We define to maps: *The trace map*

$$\mathrm{Tr} = \mathrm{Tr}_{\mathbb{F}_q^n/\mathbb{F}_q} \colon \mathbb{F}_{q^n} \to \mathbb{F}_q, \qquad \mathrm{Tr}(x) = x + x^q + \cdots + x^{q^{n-1}},$$

and norm map

$$\mathrm{Norm} = \mathrm{Norm}_{\mathbb{F}_{q^n}/\mathbb{F}_q} \colon \mathbb{F}_{q^n} \to \mathbb{F}_q, \qquad \mathrm{Norm}(x) = xx^q \cdots x^{q^{n-1}} = x^{\frac{q^n-1}{q-1}}.$$

Often, one restricts the domain of Norm to $\mathbb{F}_{q^n}^\times$ and the range to $\mathbb{F}_q^\times$ and view it as a map between the multiplicative groups of the fields. We leave the proof of the following basic properties as an exercise:

1. Tr is $\mathbb{F}_q$-linear map.
2. $\mathrm{Norm}|_{\mathbb{F}_{q^n}^\times}$ is a group homomorphism.
3. Both maps are $\mathrm{Gal}(\mathbb{F}_q^n/\mathbb{F}_q)$ invariant, that is to say, for any $\sigma \in \mathrm{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q)$ we have $\mathrm{Tr}(\sigma(x)) = \mathrm{Tr}(x)$ and $\mathrm{Norm}(\sigma(x)) = \mathrm{Norm}(x)$.
4. Given $x \in \mathbb{F}_{q^n}$, consider the linear map given by multiplication by x: $m_x \colon \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$, $m_x(y) = xy$. Prove that $\mathrm{Tr}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(x) = \mathrm{Tr}(m_x)$ and that $\mathrm{Norm}(x) = \det(m_x)$.

In our setting, Hilbert's Theorem 90 gives the kernels of these maps.

Lemma 2.9 (H90). *Let* $\mathbb{F}_{q^n}/\mathbb{F}_q$ *and* $\mathrm{Tr} \colon \mathbb{F}_{q^n} \to \mathbb{F}_q$ *and* $\mathrm{Norm} \colon \mathbb{F}_{q^n}^\times \to \mathbb{F}_q^\times$ *the respective trace and norm maps.*

1. *The trace map is surjective and* $\ker \mathrm{Tr} = \{x \in \mathbb{F}_{q^n} : x = y^q - y,\ y \in \mathbb{F}_{q^n}\}$.
2. *The norm map is surjective and* $\ker \mathrm{Norm} = \{x \in \mathbb{F}_{q^n}^\times : x = y^q/y = y^{q-1}, y \in \mathbb{F}_{q^n}\}$.

Proof. Consider $\delta\colon \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ given by $\delta(y) = y^q - y$. Then δ is $\mathbb{F}_q$-linear, $\ker\delta = \{y^q - y = 0\} = \mathbb{F}_q$, so $\dim(\operatorname{Im}\delta) = n-1$. It suffices to prove that $\operatorname{Im}\delta = \ker\operatorname{Tr}$, since then $\dim\operatorname{Im}\operatorname{Tr} = n-(n-1) = 1$, hence Tr would be surjective.

Clearly, $\operatorname{Tr}\circ\delta = 0$, so $\operatorname{Im}\delta \subseteq \ker\operatorname{Tr}$. On the other hand, $\ker\operatorname{Tr} = \{x \in \mathbb{F}_{q^n} : P(x) = 0\}$ where $P(x) = x + \cdots + x^{q^{n-1}}$ is a polynomial of degree $\leq q^{n-1}$. So $\dim\ker\operatorname{Tr} \leq n-1$. This finishes the proof of 1.

The proof of 2. is similar: Let $\Delta\colon \mathbb{F}_{q^n}^\times \to \mathbb{F}_q^\times$ be the group homomorphism given by $\Delta(y) = y^{q-1}$. Similarly to the previous case, $\ker\Delta = \mathbb{F}_q^\times$, $\operatorname{Im}\Delta \subseteq \ker\operatorname{Norm}$, hence it suffices to prove that $\#\ker\operatorname{Norm} \leq \frac{q^n-1}{q-1}$, which is true since $\mathbb{F}_{q^n}^\times$ is cyclic. □

Let us summarize in diagrams. We have the following two exact sequences:

$$\begin{tikzcd}
0 \arrow[r] & \mathbb{F}_q \arrow[r] & \mathbb{F}_{q^n} \arrow[r, "\delta"] & \mathbb{F}_{q^n} \arrow[r, "\mathrm{Tr}"] & \mathbb{F}_q \arrow[r] & 0
\end{tikzcd}$$

$$\begin{tikzcd}
0 \arrow[r] & \mathbb{F}_q^\times \arrow[r] & \mathbb{F}_{q^n}^\times \arrow[r, "\Delta"] & \mathbb{F}_{q^n}^\times \arrow[r, "\mathrm{Norm}"] & \mathbb{F}_q^\times \arrow[r] & 0
\end{tikzcd}$$

2.5 Exercises

1. Let R, S be two rings with 1 and $\phi\colon R \to S$ be a ring homomorphism (which, by assumption, satisfies $\phi(1_R) = 1_S$). Prove that the image of $\phi(R)$ is a domain (i.e., has no zero divisors) if and only if $\ker\phi$ is a prime ideal (that is, $xy \in \ker\phi$ implies that $x \in \ker\phi$ or $y \in \ker\phi$).

2. Recall that an arithmetic function is a function $f\colon \mathbb{N} \to \mathbb{C}$. We say that f is *multiplicative* if $f(mn) = f(m)f(n)$ whenever $\gcd(m,n) = 1$.

   a) Let f, g be multiplicative arithmetic functions such that $f(p^k) = g(p^k)$ for every prime power p^k. Prove that $f = g$.

   b) Show that if f, g are multiplicative, then $f * g$ is also multiplicative, where $f * g(n) = \sum_{d|n} f(d)g(n/d)$.

   c) Prove that $\mathbb{1}(n) = 1$ and $\phi(n) = \#\{1 \leq a \leq n : \gcd(n,a) = 1\}$ are multiplicative.

   d) Deduce that $h = \phi * \mathbb{1}$ is multiplicative.

3. Prove the basic properties of the trace and norm map:

   a) Tr is $\mathbb{F}_q$-linear map.

   b) $\operatorname{Norm}|_{\mathbb{F}_{q^n}^\times}$ is a group homomorphism.

   c) Both maps are $\operatorname{Gal}(\mathbb{F}_q^n/\mathbb{F}_q)$ invariant, that is to say, for any $\sigma \in \operatorname{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q)$ we have $\operatorname{Tr}(\sigma(x)) = \operatorname{Tr}(x)$ and $\operatorname{Norm}(\sigma(x)) = \operatorname{Norm}(x)$.

   d) Given $x \in \mathbb{F}_{q^n}$, consider the linear map given by multiplication by x: $m_x\colon \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$, $m_x(y) = xy$. Prove that $\operatorname{Tr}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(x) = \operatorname{Tr}(m_x)$ and that $\operatorname{Norm}(x) = \det(m_x)$.

3 Characters of finite fields

3.1 The general theory

Let G be an abstract group. A character of G is a group homomorphism $\chi\colon G \to \mathbb{C}^\times$. The set of characters of G form a group $\widehat{G} = \{\chi\colon G \to \mathbb{C}^\times\}$ which is called the *dual group*. The operations are given by

$$(\chi_1 \cdot \chi_2)(x) = \chi_1(x)\chi_2(x) \qquad \text{and} \qquad \chi^{-1}(x) = \chi(x)^{-1}(= \overline{\chi}(x) := \overline{\chi(x)},$$

and the unit element is defined by $\chi_0(x) = 1$.

If G is a topological group, we tactically assume that χ is continuous.

If G is a finite group of order n, then $1 = \chi(1) = \chi(g^n) = \chi(g)^n$, for every $\chi \in \widehat{G}$. Hence, the values are in $\mu_n = \{\zeta \in \mathbb{C} : \zeta^n = 1\} = \{e_n(a) : a \in \mathbb{Z}/n\mathbb{Z}\}$.

Example 1. $e_p \in \widehat{\mathbb{F}_p}$.

We want to prove the following

Theorem 3.1. *Let G be a finite abelian group, then $\hat{G} \cong G$.*

We break the proof into a few lemmas.

Lemma 3.2. *Let C_n be a cyclic group of order n with a generator g (written multiplicatively). For each $a \in \mathbb{Z}/n\mathbb{Z}$, we have that $\chi_a \in \widehat{G}$, where χ_a is given by $\chi_a(g^t) = e_n(at)$. Moreover, the map $X\colon \mathbb{Z}/n\mathbb{Z} \to \widehat{G}$ defined by $a \mapsto \chi_a$ is an isomorphism.*

In particular, $\widehat{C_n} = C_n$.

Proof. It is obvious that $\chi_a \in \widehat{G}$. Also, it is clear that X is injective. So, it remains to show that for every $\chi \in \widehat{G}$ there exists $a \in \mathbb{Z}/n\mathbb{Z}$ with $\chi = \chi_a$.

Since we must have $\chi(g) \in \mu_n$, we have that $\chi(g) = e_n(a) = \chi_a(n)$ for some $a \in \mathbb{Z}/n\mathbb{Z}$. But as C_n is cyclic, we get that $\chi = \chi_a$. □

Remark 3. The isomorphism in the lemma is not canonical, as it depends on the generator of C_n, and there is no canonical generator. For example, we proved that $\mathbb{F}_p^\times$ is cyclic using a counting argument, but there is no canonical generator.

For example, is 2 a generator of $\mathbb{F}_p^\times$ infinitely often? This is open. The Artin conjecture says that if $a \in \mathbb{Z}$ is not a square or -1, then it generates $\mathbb{F}_p^\times$ infinitely often (we say that a is primitive root modulo p).

- Hooley proved that GRH implies the Artin conjecture (using exponentials sums...)
- Ram Murty showed that there exists infinitely many a-s for whihc the Artin conjecture holds true.
- Heath-Brown showed that the set of primes for which the conjecture is *not* true has at most two elements. So $2, 3$ or 5 are primitive roots infinitely often.

- There is no number that is known to be primitive infinitely often.

Lemma 3.3. *Let $G = G_1 \times G_2$ be a direct product of abelian groups. Then $\widehat{G} \cong \widehat{G}_1 \times \widehat{G}_2$.*

Proof. Let $\psi \colon \widehat{G}_1 \times \widehat{G}_2 \to \widehat{G}$ be given by $\chi := \psi(\chi_1, \chi_2)(g_1, g_2) = \chi_1(g_1)\chi_2(g_2)$. Then, readily we see that χ is a character. Moreover, if we write $\chi_1(g_1) = \chi(g_1, 1)$ and $\chi_2(g_2) = \chi(1, g_2)$, then the inverse of ψ is given by $\Phi(\chi) = (\chi_1, \chi_2)$. □

Remark 4. Here the isomorphism is canonical.

Proof of Theorem 3.1. By the structure theorem of finite abelian groups $G = \prod_i G_i$, where G_i are cyclic groups. So, by the lemmas

$$\widehat{G} \cong \prod_i \widehat{G}_i \cong \prod_i G_i = G,$$

as needed. □

Corollary 3.4. *The map $g \mapsto (\chi \mapsto \chi(g))$ induces an isomorphism $\widehat{\widehat{G}} \cong G$.*

The next key property that we need is orthogonality. Recall that we denote by χ_0 the trivial character.

Theorem 3.5 (Orthogonality relations)**.** *Let G be a finite abelian group. Then*

$$\frac{1}{|G|}\sum_{g\in G}\chi(g) = \begin{cases} 1, & \chi = \chi_0 \\ 0, & \chi \neq \chi_0 \end{cases}$$

$$\frac{1}{|\widehat{G}|}\sum_{\chi\in\hat{G}}\chi(g) = \begin{cases} 1, & g = 1 \\ 0, & g \neq 1. \end{cases}$$

Proof. By the corollary, the two orthogonality relations are equivalent. Hence, we will prove only the first. Put $X = \sum_{g\in G}\chi(g)$. If $\chi = \chi_0$, then $\chi(g) = 1$ for all g, and hence $X = |G|$, as needed. If $\chi \neq \chi_0$, then there exists $h \in G$ such that $\chi(h) \neq 1$. So,

$$\chi(h)X = \sum_{g\in G}\chi(hg) = \sum_{g'\in G}\chi(g') = X.$$

This implies that $X = 0$. □

Exercise 3.1. Deduce from the orthogonality relation the following more general orthogonality relations

$$\frac{1}{|G|}\sum_{g\in G}\chi_1(g)\bar{\chi}_2(g) = \begin{cases} 1, & \chi_1 = \chi_2 \\ 0, & \chi_1 \neq \chi_2 \end{cases}$$

$$\frac{1}{|\widehat{G}|}\sum_{\chi\in\hat{G}}\chi(g_1)\chi(g_2^{-1}) = \begin{cases} 1, & g_1 = g_2 \\ 0, & g_1 \neq g_2. \end{cases}$$

3.2 Additive characters of finite fields

It is easy to give a satisfactory description of the additive characters of a finite field. Let $\mathbb{F}_q$, $q = p^n$ be a finite field. First, we construct one non-trivial character:

$$\psi \colon \mathbb{F}_q \to \mathbb{C}, \qquad \psi(x) = e_p(\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(x)).$$

Proposition 3.6. *If ψ is any nontrivial character of $\mathbb{F}_q$, then the map that sends $a \in \mathbb{F}_q$ to $\psi_a \in \widehat{\mathbb{F}_q}$ given by $\psi_a(x) = \psi(ax) = e_p(\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}(ax))$ is an isomorphism.*

Proof. It is obvious that $a \mapsto \psi_a$ is an isomorphism. Since $\mathbb{F}_q \cong \hat{\mathbb{F}}_q$, it only remains to show that it is injective. And indeed, if $a \neq 0$, and if $x \in \mathbb{F}_q$ is such that $\psi(x) \neq 1$, then $\psi_a(a^{-1}x) \neq 1$, so $\psi_a \neq \psi_0 = \mathrm{id}$, as needed. □

3.3 Multiplicative characters

Let us start by denoting by $\chi_0 \in \widehat{\mathbb{F}_q^\times}$ the principal character: $\chi_0(x) = 1$ for all $x \in \mathbb{F}_q^\times$. For any other $\chi \in \widehat{\mathbb{F}_q^\times}$, we say that the *order of χ is d* if $\min(k \geq 1 : \chi^k = \chi_0) = d$. Since $\mathbb{F}_q^\times$ is cyclic, $d \mid q-1$ and there is a character χ of order $q-1$ so that all other characters are powers of χ. Abstractly, if $g \in \mathbb{F}_q^\times$ is a primitive element, then $\chi(g^a) = e_{q-1}(a)$, is a character of order $q-1$.

However, as we have no canonical generator for $\mathbb{F}_q^\times$, we do not have a uniform way to construct the principal character.

We first start with pulling up characters via the norm map:

Lemma 3.7. *Let $\chi \in \widehat{\mathbb{F}_q^\times}$ be a character of order d, the $x \mapsto \chi(\mathrm{Norm}_{\mathbb{F}_{q^n},\mathbb{F}_q}(x))$ is a character of $\mathbb{F}_{q^n}^\times$ of order d.*

Proof. The first part follows from the multiplicativity of the norm map, and the second part from the surjectivity of the norm map. □

It is convenient to extend the definition of multiplicative characters to all of $\mathbb{F}_q$ by setting:

$$\chi(0) = \begin{cases} 0, & \chi \neq \chi_0 \\ 1, & \chi = \chi_0. \end{cases}$$

3.4 Exercises

1. Deduce from the orthogonality relation the following more general orthogonality relations

$$\frac{1}{|G|}\sum_{g\in G} \chi_1(g)\bar{\chi}_2(g) = \begin{cases} 1, & \chi_1 = \chi_2 \\ 0, & \chi_1 \neq \chi_2 \end{cases}$$

$$\frac{1}{|\widehat{G}|}\sum_{\chi\in\hat{G}} \chi(g_1)\chi(g_2^{-1}) = \begin{cases} 1, & g_1 = g_2 \\ 0, & g_1 \neq g_2. \end{cases}$$

2. Let $\mathbb{F}_q$ be a finite field with q elements, $d \mid q-1$, and $x \in \mathbb{F}_q$. Prove that

$$\sum_{\chi^d=\chi_0} \chi(x) = \#\{y \in \mathbb{F}_q : y^d = x\}.$$

Here the sum runs over all characters of order dividing d.

4 Example of general exponential sums

We keep the notation that $\mathbb{F}_q$ is a finite field, ψ, χ denoting arbitrary additive and multiplicative characters, respectively, and ψ_0, χ_0 are the principal (aka trivial) characters.

4.1 Gauss Sums

We generalize the Gauss sums appeared before:

$$g(\chi, \psi) = \sum_{x\in\mathbb{F}_q} \chi(x)\psi(x).$$

If one of the characters are principal, we readily compute the sum:

$$\begin{cases} g(\chi_0, \psi) = 0, & \psi \neq \psi_0 \\ g(\chi, \psi_0) = 0, & \chi \neq \chi_0 \\ g(\chi_0, \psi_0) = q. \end{cases}$$

Assume q is odd, and $\chi = \chi_2$ is the unique character of order 2, and for $\psi \neq \psi_0$. By Exercise 3.2, $\chi_2(x) = \#\{y^2 = x\} - 1$.

$$g(\chi_2, \psi) = \sum_{x\in\mathbb{F}_q} \Big(\sum_{y^2=x} 1 - 1\Big)\psi(x) = \sum_y \psi(y^2) - \sum_x \psi(x) = \sum_y \psi(y^2).$$

Thus, if also $q = p$, we retrieve the classical Gauss sum, introduced before.

We now fix a nontrivial character ψ so that $\psi_a(x) := \psi(ax)$, $a \in \mathbb{F}_q$ are all the additive characters. Then

$$g(\chi, \psi_a) = \sum_x \chi(x)\psi(ax) = \chi(a^{-1})\sum_x \chi(x)\psi(x) = \overline{\chi(a)}g(\chi, \psi). \tag{5}$$

Theorem 4.1. *If $\chi \neq \chi_0$ and $\psi \neq \psi_0$, $|g(\chi, \psi)| = q^{1/2}$.*

Proof. Let $g = g(\chi, \psi)$. Then

$$|g|^2 = \sum_x \sum_y \chi(x)\psi(x)\overline{\chi(y)\psi(y)}.$$

As $\chi \neq \chi_0$, we have that $\chi(0) = 0$, hence we may restrict the sum to nonzero y-s and we get

$$\begin{aligned}|g|^2 &= \sum_x \sum_{y\neq 0} \chi(xy^{-1})\psi(x-y) = \sum_u \sum_{y\neq 0} \chi(u)\psi(y(u-1)) \\ &= \sum_u \chi(u) \sum_{y\neq 0} \psi(y(u-1)) = q - \sum_u \chi(u) = q,\end{aligned}$$

as needed. □

Let pause to note the remarkable property of Gauss sums: They are algebraic integers of modulus exactly $q^{1/2}$. Moreover, for any embedding $\sigma\colon \mathbb{Q}(g) \to \mathbb{C}$, the $|\sigma(g)| = |g(\sigma\chi, \sigma\psi)| = q^{1/2}$.

Definition 4.2. Let q be prime number and $m \in \mathbb{Z}$ an integer. A *q-Weil number of weight m* is an algebraic integer α such that for any embedding $\sigma\colon \mathbb{Q}(\alpha) \to \mathbb{C}$ we have $|\sigma(\alpha)| = q^{m/2}$.

So Gauss sums are Weil numbers of weight 1 and roots of unity of weight 0.

4.2 Jacobi Sums

Let χ, λ be multiplicative characters. The Jacobi sum associated to them is

$$J(\chi, \lambda) = \sum_{x+y=1} \chi(x)\lambda(y)$$

They appear in counting solutions, for example:

Proposition 4.3. *Let N_p be the number of solutions to $X^2 + Y^2 = 1$ in $\mathbb{F}_p$, $p > 2$. Then*

$$N_p = p - \left(\frac{-1}{p}\right).$$

Proof. Let χ_2 be the character of order 2, aka the Legendre symbol. We have

$$N_p = \sum_{a+b=1} (1 + \chi_2(a))(1 + \chi_2(b)) = p + J(\chi_2, \chi_2).$$

So the proof follows from the following lemma. □

Lemma 4.4. *For any non-prinicipal χ, we have that $J(\chi, \chi^{-1}) = -\chi(-1)$.*

Proof. Since $\chi(0) = 0$, we get that $J(\chi, \chi^{-1}) = \sum_{x\neq 1} \chi(\frac{x}{1-x}) = \sum_{z\neq -1} \chi(z) = -\chi(-1)$, where we applied the change of variables $z = x/(1-x)$ that maps $\mathbb{F}_q - \{1\}$ to $\mathbb{F}_q - \{-1\}$ (note that $x = z/(z+1)$ if $z \neq -1$). □

We may express Jacobi sums via Gauss sums:

Theorem 4.5. *Let χ, λ be non-principal with $\chi\lambda$ non-principal. Then, for any $\psi \neq \psi_0$, we have*

$$J(\chi, \lambda) = \frac{g(\chi, \psi)g(\lambda, \psi)}{g(\chi\lambda, \psi)}.$$

Proof. Since $\chi(0) = \lambda(0) = 0$, we have

$$J(\chi, \lambda)g(\chi\lambda, \psi) = \sum_{x\neq 0,1} \sum_{y\neq 0} \chi(xy)\lambda(y(1-x)\psi(y).$$

The assignment $u = xy$ and $v = y(1-x)$, so (u, v), gives a bijection of the pairs (x, y) as in the sum, and the pairs (u, v) with $uv \neq 0$ and $u + v \neq 0$ (indeed the invertible assignment is $y = u + v$ and $x = \frac{u}{u+v}$). Hence,

$$\begin{aligned} J(\chi, \lambda)g(\chi\lambda, \psi) &= \sum_{\substack{u,v\in\mathbb{F}_q^\times \\ u+v\neq 0}} \chi(u)\lambda(v)\psi(u+v) \\ &= \sum_{\substack{u,v\in\mathbb{F}_q \\ u+v\neq 0}} \chi(u)\lambda(v)\psi(u+v) \\ &= g(\chi, \psi)g(\lambda, \psi) - \sum_{u\in\mathbb{F}_q} \chi(u)\lambda(-u) = g(\chi, \psi)g(\lambda, \psi) \end{aligned}$$

where the last equality follows from the orthogonality relations since $\chi \neq \lambda$. □

Corollary 4.6. *$J(\chi, \lambda)$ is a q-Weil number of weight* 1.

Theorem 4.7. *(Fermat) Let $p \equiv 1(4)$ be a prime. Then there exists $a, b \in \mathbb{Z}$ such that $p = a^2 + b^2$.*

Proof. Since $4 \mid p - 1$, there exists a character χ of order 4. Let $J = J(\chi, \chi_2)$. Then $|J|^2 = p$. On the other hand, $J = a + bi$, as the values of χ and χ_2 are in $\{\pm 1, \pm i\}$. so $|J|^2 = a^2 + b^2$. □

5 Salié sums

One generalization of Kloosterman sums is

$$K(\psi, \eta) = \sum_{x\in\mathbb{F}_q^\times} \psi(x)\eta(x^{-1}),$$

where ψ, η are additive characters. A related, but easier, sum is the Salié sum

$$T(\psi, \eta) = \sum_{x\in\mathbb{F}_q^\times} \chi_2(x)\psi(x)\eta(x^{-1}),$$

where ψ, η are additive characters and χ_2 is multiplicative character of order 2.

Theorem 5.1. *Assume $\psi, \eta \neq \psi_0$. Then*

$$T(\psi, \eta) = g(\chi_2, \psi) \sum_{y^2=4a} \psi(y),$$

where $a \in \mathbb{F}_q^\times$ is such that $\eta = \psi_a$.

Proof. The idea is to study the variation of the function

$$\phi(b) = T(\psi_b, \eta) = \sum_{x \neq 0} \chi_2(x)\psi(bx + ax^{-1}), \qquad b \in \mathbb{F}_q^\times.$$

By the orthogonality relations (Fourier transform)

$$\phi(b) = \sum_{\chi} \hat{\phi}(\chi)\chi(b),$$

where

$$\hat{\phi}(\chi) = \frac{1}{q-1} \sum_{b \neq 0} \phi(b)\bar{\chi}(b).$$

We compute the Fourier coefficients:

$$\hat{\phi}(\chi) = \frac{1}{q-1} \sum_{b \neq 0} \bar{\chi}(b) \sum_{x \neq 0} \chi_2(x)\psi(bx + ax^{-1}) = \frac{1}{q-1} \sum_{x \neq 0} \chi_2(x)\psi(ax^{-1}) \sum_{b \neq 0} \bar{\chi}(b)\psi(bx).$$

Applying (5), we get that the right-hand sum equals $\chi(x)g(\bar{\chi}, \psi)$, so

$$\hat{\phi}(\chi) = \frac{g(\bar{\chi}, \psi)}{q-1} \sum_{x \in \mathbb{F}_q^\times} \chi_2(x)\chi(x)\psi(ax^{-1}) = \frac{g(\bar{\chi}, \psi)g(\chi_2\bar{\chi}, \psi_a)}{q-1} = \frac{g(\bar{\chi}, \psi)g(\chi_2\bar{\chi}, \psi)\chi_2(a)\chi(a)}{q-1}$$

By Exercise 5.2, we get that $g(\bar{\chi}, \psi)g(\chi_2\bar{\chi}, \psi) = \chi(4)g(\bar{\chi}^2, \psi)g(\chi_2, \psi)$ if $\chi \neq \chi_2, \chi_0$. If $\chi = \chi_2, \chi_0$, then $g(\bar{\chi}, \psi)g(\chi_2\bar{\chi}, \psi) = 0$. Plugging this in, we get

$$\hat{\phi}(\chi) = \frac{\chi_2(a)\chi(4)g(\chi_2, \psi)g(\bar{\chi}^2, \psi)}{q-1}.$$

Thus,

$$T(\psi, \eta) = \phi(1) = \sum_{\chi} \hat{\phi}(\chi) = \frac{\chi_2(a)g(\chi_2, \psi)}{q-1} \sum_{\chi} \chi(4a)g(\bar{\chi}^2, \psi).$$

Opening the inner sum and using orthogonality,

$$T(\psi, \eta) = \frac{\chi_2(a)g(\chi_2, \psi)}{q-1} \psi(x) \sum_{x \neq 0} \sum_{\chi} \chi(4ax^{-2}) = \frac{\chi_2(a)g(\chi_2, \psi)}{q-1} \sum_{x=4a^2} \psi(x).$$

Finally, we may remove the factor $\chi_2(a)$, because if it equals -1, then the inner sum is 0. □

Since $y^2 = 4a$ has either 0 or 2 solution, $T(\psi, \eta)$ is a sum of at most two q-Weil numbers of weight 1, hence:

Corollary 5.2. $|T(\psi, \eta)| \leq 2\sqrt{p}$.

5.1 Exercises

1. Prove that if $p \equiv 3(4)$, then $q \neq a^2 + b^2$.
2. Let q be odd prime power. Let χ_2 be the character of order 2, and χ, ψ characters so that $\chi \neq \chi_0, \chi_2$ and $\psi \neq \psi_0$. Prove

$$g(\chi^2, \psi)g(\chi_2, \psi) = \chi(4)g(\chi, \psi)g(\chi\chi_2, \psi).$$

6 Equations over finite fields

Given a field F, we write

$$\mathbb{A}^n(F) = \{(x_1, \dots, x_n) : x_i \in F\} = F^n$$

for the affine n-space over F. It is a vector space over F.

On $\mathbb{A}^{n+1} \smallsetminus \{0\}$ define an equivalence relation

$$(x_0, \dots, x_n) \sim (y_0, \dots, y_n) \iff \exists \lambda \in F^\times \forall i : x_i = \lambda y_i.$$

We denote the equivalence classes by $[x_0 : \cdots : x_n]$, and we define the n-projective space over F to be

$$\mathbb{P}^n(F) = (\mathbb{A}^{n+1}(F) \smallsetminus \{0\}) / \sim = \{[x_0 : \cdots : x_n] : x_i \in F\}.$$

We call points $[x] = [x_0 : \cdots : x_n]$ with $x_0 \neq 0$ finite points, and those with $x_0 = 0$ points at infinity.

Lemma 6.1. *There is a bijection between the finite points and $\mathbb{A}^n(F)$ and the points at infinity and $\mathbb{P}^{n-1}(F)$.*

Proof. Indeed, every finite point is equivalent to exactly one point of the form $[1 : x_1 : \cdots : x_n]$, hence the bijection with $\mathbb{A}^n(F)$ is trivial. Similarly, if $[0 : x_1 : \cdots : x_n] \in \mathbb{P}^n(F)$, then not all x_i are zero, and hence it defines a point on $\mathbb{P}^{n-1}(F)$, and this is obviously a bijection. □

We usually abbreviate and write the above lemma as

$$\mathbb{P}^n(F) = \mathbb{A}^n(F) \coprod \mathbb{P}^{n-1}(F).$$

If F is finite, we readily compute the sizes of the affine and projective spaces:

Proposition 6.2. $\#\mathbb{A}^n(\mathbb{F}_q) = q^n$ *and* $\#\mathbb{P}^n(\mathbb{F}_q) = q^n + \cdots + 1 = \frac{q^{n+1}-1}{q-1}$.

For a vector of nonnegative integers $i = (i_1, \dots, i_n)$ we write $|i| = \sum_{j=1}^{n} i_j$, and given a vector $X = (X_1, \dots, X_n)$ of elements of a ring, we write $X^i = X_1^{i_1} \cdots X_n^{i_n}$. Given

$$f(X) = f(X_1, \dots, X_n) = \sum_{|i| \leq d} a_i X^i$$

with $a_i \in \mathbb{F}_q$, and degree $\leq d$, we are interested in

$$N_f = \#\{x \in \mathbb{A}^n(\mathbb{F}_q) : f(x) = 0\}.$$

The equation $f(X) = 0$ defines a hypersurface in $\mathbb{A}^n$.

In the projective setting, we are given a homogenous polynomial of degree d

$$F(X_0, \dots, X_n) = \sum_{|i|=d} a_i X^i, \qquad a_i \in \mathbb{F}_q.$$

Now the value $F(x)$ is not well-defined for $x \in \mathbb{P}^n(F)$, but the solutions to $F = 0$ in $\mathbb{P}^n$ are well-defined. We are interested in

$$N_F^* = \#\{[x] \in \mathbb{P}^n(\mathbb{F}_q) : F(x) = 0\},$$

and we say that $F = 0$ defines a hypersurface in $\mathbb{P}^n$.

6.1 Crude bounds

Let $f(X) = f(X_1, \dots, X_n) \in \mathbb{F}_q[X]$ be a nonzero polynomial of degree d.

Proposition 6.3. $N_f \leq dq^{n-1}$.

Proof. If $d = 0$, then $N_f = 0$ and we are done. If $d = 1$, then $f = 0$ defines a translation of a linear subspace of codimension 1, so $N_f = q^{n-1}$. If $n = 1$, then f is univariate, hence $N_f \leq d$.

We proceed by double induction on n, d.

Case 1: Assume there exists $a \in \mathbb{F}_q$ such that $(X_1 - a) \mid f$. Then, $f(X) = (X_1 - a)g(X)$, with $\deg g = d - 1$. So by induction

$$N_f \leq q^{n-1} + N_g \leq q^{n-1} + (d-1)q^{n-1} = dq^{n-1}.$$

Case 2: For any $a \in \mathbb{F}_q$, $g_a(X_2, \dots, X_n) = f(a, X_2, \dots, X_n)$ is nonzero and of degree $\leq d$. So by induction,

$$N_f \leq q \max_a(N(g_a)) \leq qdq^{n-2} = dq^{n-1},$$

as needed for the proof. □

Given $f(X_1, \dots, X_n)$ of degree d, we homogenize it:

$$f^*(X_0, \dots, X_n) = X_0^d f(X_1/X_0, \dots, X_n/X_0).$$

So f^* is a degree-d form.

Lemma 6.4. $N_f \le N^*_{f^*} \le N_f + dq^{n-2}(1-q^{-1})$.

Proof. Let $F(X_1,\dots,X_n) = f^*(0,X_1,\dots,X_n)$. Then, F is a nonzero form of degree d, and

$$N^*_{f^*} = N_f + N^*_F.$$

It remains to show that $N^*_F \le dq^{n-2}(1-q^{-1})$. And indeed, $N^*_F = N_F/(q-1) \le dq^{n-1}/(q-1)$, by Lemma 6.3. □

We try to understand the statistics of N_f. For this, we write

$$\Omega_d = \{f \in \mathbb{F}_q[X_1,\dots,X_n] : \deg f \le d\}.$$

Since there are $\omega_d := \binom{n+d}{d}$-many $i = (i_1,\dots,i_n)$ with $|i| = d$, we get that

$$\#\Omega_d = q^{\omega_d}.$$

Theorem 6.5. *Let f be uniform in Ω_d, then*

$$\mathbb{E}(N_f) = q^{n-1}.$$

Proof. By direct computation and the fact that given an $x \in \mathbb{F}_q^n$, the condition $f(x) = 0$ is linear in the coefficients of f. □

Exercise 6.1. For f uniform in Ω_d, we have that $\mathrm{Var}(N_f) = q^{n-1} - q^{n-2}$.

These two exercises may be expressed that typically

$$N_f = q^{n-1} + O(q^{\frac{n-1}{2}}),$$

that is, we have square root cancelation.

We proceed into some cases where such an estimate can be achieved.

6.2 Quadratic hypersurfaces

Assume q is odd. A quadratic form over $\mathbb{F}_q$ is a form of degree 2 and is given by equation

$$Q(X) = Q(X_1,\dots,X_n) = \sum_{1\le i,j\le n} a_{ij}X_iX_j$$

with $a_{ij} = a_{ji} \in \mathbb{F}_q$. In other words,

$$Q(X) = X^TAX,$$

with $A = (a_{ij})$ a symmetric matrix. Define

$$\det Q = \det A.$$

Two quadratic forms are equivalent, written as $Q_1 \sim Q_2$, if there is a nonsingular matrix M such that $Q_1(X) = Q_2(MX)$, or equivalently, $A_1 = M^TA_2M$. Hence, if

$Q_1 \sim Q_2$, we have $\det Q_1 = \det Q_2 m^2$, with $m \in \mathbb{F}_q^\times$. For $d \in \mathbb{F}_q^\times$, we use the Legendre symbol: For $d \in \mathbb{F}_q^\times$, we use the Legendre symbol $\left(\frac{d}{q}\right) = \begin{cases} +1, & d = \square \\ -1, & \text{otherwise.} \end{cases}$ So, $Q_1 \sim Q_2$ implies that $\left(\frac{\det Q_1}{q}\right) = \left(\frac{\det Q_2}{q}\right)$.

We say that Q represent $a \in \mathbb{F}_q^\times$, if there exists $x \in \mathbb{F}_q^n$ such that $Q(x) = a$. We say that Q represents zero if in addition $x \neq 0$.

Lemma 6.6. *Suppose Q represents $a \in \mathbb{F}_q^\times$. Then $Q(X) \sim aX_1^2 + P(X_2, \ldots, X_n)$, for some quadratic form P over $\mathbb{F}_q$.*

Proof. Exercise. □

Corollary 6.7. *Every Q is equivalent to a diagonal quadratic form.*

Lemma 6.8. *If a nonsingular quadratic form Q represents 0, then it represents any $a \in \mathbb{F}_q^\times$.*

Proof. Since equivalent quadratics represent the same elements, we may assume that $Q(X) = \sum_i a_i X_i^2$, $a_i \neq 0$. Let $0 \neq x \in \mathbb{F}_q^n$ be such that $Q(x) = 0$. W.l.o.g., we may assume $x_1 \neq 0$. Let $y_1 = x_1(1+t)$ and $y_i = x_i(1-t)$, $i = 2, \ldots, n$. Then, $Q(y) = \sum_i a_i x_i^2 + t^2 \sum_i a_i x_i^2 + 4a_1 x_1 t - 2t \sum_i a_i x_i^2 = 4a_1 x_1 t$. So when we vary t, we get all elements in $\mathbb{F}_q$. □

Let $Q(X)$ be a nonsingular quadratic form in $n \geq 3$ variables. By Chevalley-Warning Theorem, Q represents 0, hence it represents 1. So $Q(X) \sim X_1^2 + P(X_2, \ldots, X_n)$. We conclude that there exists a nonzero solution $x \in \mathbb{F}_p^n$ to $x_1^2 = P(x_2, \ldots, x_n)$. If $x_1 \neq 0$, then P represents -1. If $x_1 = 0$, then P represents 0, hence -1. So,

$$Q(X) \sim X_1^2 - X_2^2 + R(X_3, \ldots, X_n) \sim X_1 X_2 + R(X_3, \ldots, X_n)$$

for some nonsingular quadratic form R in $n-2$ variables.

We shall use the above elementary consideration to estimate

$$N_Q = \#\{x \in \mathbb{F}_q^n : Q(x) = 0\}.$$

In fact, we give an exact formula:

Theorem 6.9. *Let $Q(X) \in \mathbb{F}_q[X_1, \ldots, X_n]$ be a nonsingular quadratic form, $n \geq 1$, and write $\Delta = \det Q$. Then*

$$N_Q = \begin{cases} q^{n-1}, & 2 \nmid n \\ q^{n-1} + (q-1) q^{\frac{n-2}{2}} \left(\frac{(-1)^{n/2}\Delta}{q}\right), & 2 \mid n. \end{cases}$$

Proof. We prove the statement by induction on n. If $n = 1$, then $N_Q = 1$, as needed. If $n = 2$, then $Q = aX_1^2 + bX_2^2 = a(X_1^2 + b/aX_2^2)$, for nonzero a, b. Moreover, $\left(\frac{-\Delta}{q}\right) = \left(\frac{-ab}{q}\right)$. If $\left(\frac{-ab}{=}\right) - 1$, then we only have the trivial solution. So $N_Q = 1$, as needed. If $\left(\frac{-ab}{=}\right)1$, then there are $1 + 2(q-1) = 2q - 1$, solutions, as needed.

For the induction step, assume $n \geq 3$. We may assume w.l.o.g. that $Q = X_1X_2 + R(X_3, \dots, X_n)$, with R nonsingular quadratic form. There are $2q-1$ pairs (x_1, x_2) with x_1x_2, hence the number of x counted by N_Q with $R = 0$ is $(2q-1)N_R$.

Likewise, given $x_3, \dots, x_n$ such that $R \neq 0$, there exist $q-1$ choices of (x_1, x_2) such that $x_1x_2 = -R(x_3, \dots, x_n)$, hence this contributes to the count in N_Q exactly $(q-1)(q^{n-2} - N_R)$. So

$$N_Q = q^{n-1} - q^{n-2} + qN_R. \tag{6}$$

By induction,

$$N_R = \begin{cases} q^{n-3}, & 2 \nmid n \\ q^{n-3} + (q-1)q^{\frac{n-4}{2}}\left(\frac{(-1)^{n/2}\Delta}{q}\right), & 2 \mid n. \end{cases}$$

Plugging this in (6), immediately gives the assertion. □

6.3 Diagonal hypersurfaces

Given any polynomial $f \in \mathbb{F}_q[X_1, \dots, X_n]$, by orthogonality of characters, we have

$$N_f = \frac{1}{q} \sum_{a \in \mathbb{F}_q} \sum_{x \in \mathbb{F}_q^n} \psi(af(x)), \tag{7}$$

where $\psi \neq \psi_0$ is a non-principal character. We (7) to give a formula for a diagonal hypersurface of degree d.

Theorem 6.10. *Assume $d \mid q-1$, $f(X) = a_1X_1^d + \cdots + a_nX_n^d$, with nonzero $a_i \in \mathbb{F}_q$, $i = 1, \dots, n$. Then*

$$N_f = q^{n-1} + (1 - q^{-1}) \sum_{\chi_1, \dots, \chi_n} \prod_i \bar{\chi}_i(a_i) g(\chi_i, \psi), \tag{8}$$

where the sum runs over multiplicative characters $\chi_1, \dots, \chi_n$ of $\mathbb{F}_q$ satisfying $\chi_i \neq \chi_0$, $\prod_i \chi_i \neq \chi_0$ and $\chi_i^d = \chi_0$ (for all $i = 1, \dots, n$) and $g(\chi, \psi)$ is the Gauss sum. In particular,

$$|N_f - q^{n-1}| \leq A(d)(1 - q^{-1})q^{n/2},$$

where $A(d)$ is the number of $b = (b_1, \dots, b_n) \in \mathbb{Z}^n$ with $0 < b_i < d$ and $b_1 + \cdots + b_n \equiv 0 \mod d$.

Proof. Applying (6), we get that

$$\begin{aligned} qN_f &= \sum_{a \in \mathbb{F}_q} \sum_{x \in \mathbb{F}_q^n} \psi(a \sum_i a_i x_i^d) = \sum_{a \in \mathbb{F}_q} \sum_{x \in \mathbb{F}_q^n} \prod_{i=1}^n \psi(aa_ix_i^d) = \sum_{a \in \mathbb{F}_q} \prod_{i=1}^n \sum_{x_i \in \mathbb{F}_q} \psi(aa_ix_i^d) \\ &= q^n + \sum_{a \neq 0} \prod_i \sum_{x_i \in \mathbb{F}_q} \psi(aa_ix_i^d) \end{aligned}$$

Using $\sum_{\chi^d=\chi_0}\chi(y) = \#\{x : y = x^d\}$, we may rewrite the inner sums in the right-hand side as

$$\sum_{x_i\in\mathbb{F}_q}\psi(aa_ix_i^d) = \sum_{y_i\in\mathbb{F}_q}\psi(aa_iy_i)\sum_{\chi_i^d=\chi_0}\chi_i(y_i).$$

Since $a \neq 0$ in these sums, we may change order of summation and make the change of variables $y_i \mapsto y_i/(aa_i)$, to get

$$\sum_{x_i\in\mathbb{F}_q}\psi(aa_ix_i^d) = \sum_{\chi_i^d=\chi_0}\bar{\chi}_i(aa_i)\sum_{y_i\in\mathbb{F}_q}\psi(y_i)\chi_i(y_i) = \sum_{\chi_i^d=\chi_0}\bar{\chi}_i(aa_i)g(\chi_i,\psi).$$

Collecting everything together, we get

$$qN_f - q^n = \sum_{\substack{\chi_1,\cdots\chi_n\\ \chi_i^d=\chi_0}}\prod\bar{\chi}_i(a_i)g(\chi_i,\psi)\sum_{a\neq 0}\prod_i\bar{\chi}_i(a).$$

If $\chi_1\cdots\chi_n \neq \chi_0$, then $\sum_{a\neq0}\prod_i\bar{\chi}_i(a) = 0$ by orthogonality. Otherwise, $\sum_{a\neq0}\prod_i\bar{\chi}_i(a) = q-1$. If $\chi_i = \chi_0$, then $g(\chi_i,\psi) = 0$, hence we may remove these terms from the sums. This finishes the proof of the formula for N_f.

The second part follows immediately from the formula, after noting that error term is a sum of $A(d)$ q-Weil numbers of weight n. □

Exercise 6.2. Prove by induction on n that $A(d) = \frac{d-1}{d}[(d-1)^n - (-1)^{n-1}] < (d-1)^n$, and deduce

$$|N_f - q^n| < (1-q^{-1})(d-1)^nq^{n/2},$$

for diagonal form f of degree d and n variables.

Next, we want to study the dependence of the number of solutions on the field of coordinates. For an integer $\nu \geq 1$, let

$$N_f(\nu) = \#\{x \in \mathbb{F}_{q^\nu} : f(x) = 0\}.$$

(Here, as before, $f = \sum_{i=1}^n a_iX_i^d$, $d \mid q-1$, and $a_i \in \mathbb{F}_q^\times$.)

Since $d \mid q-1$, then multiplicative characters of order d of $\mathbb{F}_{q^\nu}$ are exactly

$$\chi_\nu := \chi\circ\mathrm{Norm}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q},$$

as χ runs over characters of order d of $\mathbb{F}_q$. Since $a_i \in \mathbb{F}_q$, $\mathrm{Norm}_{(\mathbb{F}_{q^\nu}/\mathbb{F}_q}(a_i) = a_i^\nu$, so $\bar{\chi}_\nu(a_i) = (\bar{\chi}(a_i))^\nu$ Moreover,

$$\psi_\nu = \psi\circ\mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}$$

is a non-principal additive character of $\mathbb{F}_{q^\nu}$. Let $g_\nu(\chi,\psi) = g(\chi_\nu,\psi_\nu)$.We will prove below the Hasse-Davenport relations that shows that $-g_\nu = (-g)^\nu$. Plugging all of these into (8) gives

$$N_f(\nu) = q^{\nu(n-1)} + (-1)^{(\nu-1)n}(1-q^{-\nu})\sum_{\chi_1,\dots,\chi_n}\Big(\prod\bar{\chi}_ig(\chi_i,\psi)\Big)^\nu, \tag{9}$$

where the sum is over non-principal multiplicative characters $\chi_1, \dots, \chi_n$ of $\mathbb{F}_q$ with $\chi_i^d = \chi_0$ and $\prod_i \chi_i = \chi_0$. Thus,

$$N_f(\nu) = \sum_i \alpha_i^\nu - \sum_j \beta_j^\nu,$$

where α_i and β_j are q-Weil numbers of various weights.

In 1949, Weil conjectured that such formula *always* exists for the counting function $N_f(\nu)$, associated to any $f \in \mathbb{F}_q[X_1, \dots, X_n]$. We will return to this later on.

6.4 The Hasse-Davenport relations

Let χ be a multiplicative character if $\mathbb{F}_q$ and ψ an additive character. As before, for an extension $\mathbb{F}_{q^\nu}/\mathbb{F}_q$ we write $\chi_\nu = \chi \circ \mathrm{Norm}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}$ and $\psi_\nu = \psi \circ \mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}$ for the associated characters of $\mathbb{F}_{q^\nu}$. Let $g = g(\chi, \psi)$ and $g_\nu = g(\chi_\nu, \psi_\nu)$. Our goal is to prove that $-g_\nu = (-g)^\nu$.

For a polynomial,

$$f(X) = X^d - c_1 X^{d-1} + \cdots + (-1)^d c_d \in \mathbb{F}_q[X]$$

we define $\lambda(f) = \chi(c_d)\psi(c_1)$ and $\lambda(1) = 1$.

Lemma 6.11. *For any $f, g \in \mathbb{F}_q[X]$ we have $\lambda(fg) = \lambda(f)\lambda(g)$.*

Proof. Obvious. □

Lemma 6.12. *Let $\alpha \in \mathbb{F}_{q^\nu}$ and let f be the minimal polynomial of α over $\mathbb{F}_q$, and put $d = \deg f$. Then*

$$\lambda(f)^{\nu/d} = \chi_\nu(\alpha)\psi_\nu(\alpha).$$

Proof. We have

$$f(X) = X^d - c_1 X^{d-1} + \cdots + (-1)^d c_d = \prod_{i=1}^{d} (X - \alpha^{q^i}).$$

So $\mathrm{Tr}(\alpha) = \sum_{i=0}^{\nu-1} \alpha^{q^i} = \frac{\nu}{d}(\sum_{i=0}^{d-1} \alpha^{q^i}) = \frac{\nu}{d} c_1$, and similarly, $\mathrm{Norm}(\alpha) = \prod_{i=0}^{\nu-1} \alpha^{q^i} = \left(\prod_{i=0}^{d-1} \alpha^{q^i}\right)^{\nu/d} = c_d^{\nu/d}$. So, $\lambda(f)^{\nu/d} = \chi(c_d)^{\nu/d}\psi(c_1)^{\nu/d} = \chi(c_d^{\nu/d})\psi(\frac{\nu}{d}c_1) = \chi_\nu(c_d)\psi_\nu(c_1)$. □

We apply the lemma to rewrite g_ν: Let f_α be the minimal polynomial of $\alpha \in \mathbb{F}_{q^\nu}$, then $\deg f \mid \nu$, so by the lemma we have

$$g_\nu = \sum_{\alpha \in \mathbb{F}_{q^\nu}} \lambda(f_\alpha)^{\nu/\deg f_\alpha}.$$

There is a d-to-1 correspondence between elements of degree d and irreducible polynomials of degree d, hence

$$g_\nu = \sum_{d \mid \nu} \sum_{\deg P = d} d\lambda(P)^{\nu/d}, \tag{10}$$

where P denotes irreducible monic polynomial.

We add a piece of notation: let $\mathscr{M} \subseteq \mathbb{F}_q[X]$ be the subset of monic polynomials, and $\mathscr{M}_d = \{f \in \mathscr{M} : \deg f = d\}$.

Theorem 6.13 (Hasse-Davenport relation)**.** *Let $\chi \neq \chi_0$ be a multiplicative character and $\psi \neq \psi_0$ an additive character, and $\nu \geq 1$ an integer. The $g(\chi_\nu, \psi_\nu) = (-1)^{\nu-1}g(\chi, \psi)$.*

Proof. Expand $(1 - \lambda(P)u^{\deg P})^{-1}$ as a geometric series im $u$and applying unique factorization in $\mathbb{F}_q[X]$, we have

$$\prod_P (1 - \lambda(P)u^{\deg P})^{-1} = \sum_{f \in \mathscr{M}} \lambda(f)u^{\deg f} = \sum_{d=0}^{\infty} \sum_{f \in \mathscr{M}_d} \lambda(f)u^d$$

We compute the coefficient of u^d. For $d = 1$, the coefficient is 1. For $d = 1$, the coefficient is $\sum_{a \in \mathbb{F}_q} \lambda(X - a) = \sum a\chi(a)\psi(a) = g(\chi, \psi)$. For $d > 1$, the coefficient is

$$\sum_{f \in \mathscr{M}_d} \lambda(f) = q^{d-2} \sum_{c_1, c_d \in \mathbb{F}_q} \chi(c_d)\psi(c_1) = 0.$$

Hence,

$$\prod_P (1 - \lambda(P)u^{\deg P})^{-1} = 1 + g(\chi, \psi)u.$$

Now we apply logarithmic derivative and multiplication by u to both sides. The left-hand side transformsin to

$$\begin{aligned} \sum_P \frac{\deg P \cdot \lambda(P)u^{\deg P}}{1 - \lambda(P)u^{\deg P}} &= \sum d \sum_{\deg P = d} d \sum_{r=1}^{\infty} \lambda(P)^r u^{r \deg P} \\ &= \sum_{\nu=1}^{\infty} u^\nu \sum_{d|\nu} \sum_{\deg P = d} d\lambda(P)^{\nu/d} = \sum_{\nu \geq 1} g_\nu u^\nu, \end{aligned}$$

where the last equility follows from (10). The right-hand side transforms to

$$\frac{gu}{1 + gu} = \sum_{\nu \geq 1} (-1)^{\nu-1} g^\nu u^\nu.$$

Comparing the coefficients, finishes the proof. □

6.5 The Zeta function of a hypersurface

Let $F \in \mathbb{F}_q[X_0, \ldots, X_n]$ be a form of degree d. For any ν, let

$$N_F^*(\nu) = \#\{[x] \in \mathbb{P}^n(\mathbb{F}_{q^\nu}) : F(x) = 0\}.$$

We define the zeta function to be

$$Z_F(u) = \exp\left(\sum_{\nu\geq 1} \frac{N_F^*(\nu)}{\nu} u^\nu\right). \tag{11}$$

We view it as a formal power series with rational coefficients, that is, $Z_F(u) \in \mathbb{Q}[[u]]$, or as a complex function that is analytic in the disc $\{|u| \leq q^{-n}\}$.

Example 2. Let us consider $\mathbb{P}^n_{\mathbb{F}_q}$ (e.g., take $F = X_{n+1} \in \mathbb{F}_q[X_0, \dots, X_{n+1}]$). Then,

$$Z_{\mathbb{P}^n_{\mathbb{F}_q}} = \exp\left(\sum_{\nu\geq 1} \frac{\sum_{i=0}^n q^{i\nu}}{\nu} u^\nu\right) = \prod_{i=0}^n \exp\left(\sum_{\nu\geq 1} \frac{(q^i u)^\nu}{\nu}\right) = \frac{1}{(1-u)(1-qu)\cdots(1-q^n u)}.$$

In particular, $Z_{\mathbb{P}^n_{\mathbb{F}_q}}$ is a rational function.

Example 3. We compute the zeta function of $F(X_0, X_1, X_2) = X_0^3 + x_1^3 + X_2^3 = 0$ in $\mathbb{P}^2_{\mathbb{F}_q}$.

We have

$$N_F^*(\nu) = \frac{N_F(\nu) - 1}{q^\nu - 1},$$

where $N_F(\nu) = \{x \in \mathbb{F}_q^3 : F(x) = 0\}$ is the number of affine solutions. By (9),

$$N_F(\nu) = q^{2\nu} + (-1)^{\nu-1}(1 - q^{-\nu})(g(\chi,\psi)^{3\nu} + g(\chi^2,\psi)^{3\nu}),$$

where $\chi \neq \chi_0$ is a character of order 3 and $\psi \neq \psi_0$. Combining the these equations together yields

$$N_F^*(\nu) = q^\nu + 1 + \frac{(-1)^{\nu-1}(g(\chi,\psi)^{3\nu} + g(\chi^2,\psi)^{3\nu})}{q^\nu}.$$

Since

$$\exp\left(-\sum_{\nu\geq 1} \frac{(-q^{-1}g(\chi,\psi)^3 u)^\nu}{\nu}\right) = 1 + q^{-1}g(\chi,\psi)^3 u$$

and similarly, for χ^2, we have

$$Z_F(u) = \frac{(1 + q^{-1}g(\chi,\psi)^3 u)(1 + q^{-1}g(\chi^2,\psi)^3 u)}{(1-u)(1-qu)}.$$

As $\chi^2 = \bar{\chi}$, we have,

$$g(\chi,\psi)g(\chi^2,\psi) = g(\chi,\psi)g(\bar{\chi},\psi) = g(\chi,\psi)\chi(-1)g(\bar{\chi},\bar{\psi}) = \chi(-1)g(\chi,\psi)\overline{g(\chi,\psi)} = \chi(-1)q.$$

Note that, as χ is of order 3, it follows that $\chi(-1) = \chi((-1)^3) = \chi^3(-1) = 1$. Applying Theorem 4.5 ($g(\chi,\psi)^2 = \pi g(\chi^2,\psi)$, $\pi = J(\chi,\chi)$) we get

$$g(\chi,\psi)^3 = J(\chi,\chi)g(\chi,\psi)g(\chi^2,\psi) = \pi q.$$

Similarly, $g(\chi^2,\psi) = \bar{\pi}q$. To conclude, we get that

$$\boxed{Z_F(u) = \frac{(1+\pi u)(1+\bar{\pi}u)}{(1-u)(1-qu)} \in \mathbb{Q}(u)}$$

where π is a q-Weil number.

The Zeta function is a priori in $\mathbb{Q}((u))$, and we saw two examples in which the Zeta function is in fact rational in $\mathbb{Q}(u)$. Moreover, in these examples, the zeros and poles of the zeta functions were q-Weil numbers.

This is an incident of a very general theme. Given a nonzero form $F \in \mathbb{F}_q[X_0, X_1, X_2]$ of degree $\deg F = d$ which is non-singular over any algebraic extension of $\mathbb{F}_q$, Weil (1948) proved that

$$Z_F(u) = \frac{P(T)}{(1-T)(1-qu)},$$

where $p(T) \in \mathbb{Z}[T]$ of degree $(d-1)(d-2) = 2g$, g is the genus of the curve defined by $F = 0$. Furthermore, he proved that the roots of P are q-Weil numbers of weight -1, i.e. $P = \prod_{i=1}^{2g}(1-\pi_i u)$ and $|\pi_i| = q^{1/2}$.

The last statement is called *The Riemann Hypothesis for Curves*. To see the analogy with the classical Riemann Hypothesis, change variables to s with $u = q^{-s}$. Then

$$\zeta_F(s) = Z_F(q^{-s}) = (1-q^{-s})^{-1}(1-q^{1-s})^{-1}P(q^{-s}).$$

So ζ_F has a simple pole at $s = 1$, and the roots of ζ are on the critical line $\Re(s) = 1/2$ if and only if the roots of $Z_F(u)$ are q-Weil numbers of weight -1.

For higher dimensions, namely nonzero forms $F \in \mathbb{F}_q[X_0, \ldots, X_n]$, Dwork (1959) proved that

$$Z_F(u) = \frac{\prod_i (1-\alpha_i u)}{\prod_j (1-\beta_j u)}, \qquad \alpha_i, \beta_j \in \mathbb{C}. \tag{12}$$

The fact that the constant term is 1 is easy, just note that $Z_F(0) = 1$ by its definition.

Lemma 6.14. *Assume $Z_F(u)$ has the form as in* (12). *Then*

$$N_F^*(\nu) = \sum_j \beta_j^\nu - \sum_i \alpha_j^\nu.$$

Proof. As usual, we apply the operator $u\frac{d\log}{du}$ to (12) and using the definition (11) and we get

$$\sum_{\nu\geq 1} N_F^*(\nu)u^\nu = \sum_i \frac{-\alpha_i u}{1-\alpha_i u} - \sum_j \frac{-\beta_j u}{1-\beta_j u}.$$

Expanding to geometric series and comparing coefficients finish the proof. □

Remark 5. The converse of the lemma is easily seen to be true.

Let us end the section with a brief cohomological interpretation of the zeta function.

Let V be a non-singular projective hypersurface defined by a form $F \in \mathbb{F}_q[X_0, \ldots, X_n]$. Let $\Omega\colon V \to V$ be given by $\Omega([x_0, \ldots, x_n]) = [x_0^q, \ldots, x_n^q]$ be the $(n+1)$-fold Frobenius operator. We have $[x] \in \mathbb{P}^n(\mathbb{F}_{q^\nu})$ if and only if $\Omega^\nu([x]) = [x]$. So

$$V(\mathbb{F}_{q^\nu}) = \{[x] \in \mathbb{P}^n(\mathbb{F}_{q^\nu}) : F(x) = 0\} = V(\bar{\mathbb{F}}_q)^{\Omega^\nu}.$$

In topology, we have the Lefschetz fixed-point theorem, which counts the number of fix points of an operator by means of traces on the induced mapping on the homology. The analog in arithmetic is the *Grothendieck-Lefschetz fixed-point formula* stating that for any prime $\ell \neq p = \operatorname{Char}\mathbb{F}_q$, we have

$$N_F^*(\nu) = \sum_{i=0}^{2(n-1)} (-1)^i \mathrm{Tr}(\Omega^{*\nu}; H^i(V, \mathbb{Q}_\ell)),$$

where H^i denotes the i-th étale cohomology group, $\mathbb{Q}_\ell$ the ℓ-adic numbers, and Ω^* the induced mapping on the cohomology group. In face $H^i = H^i(V, \mathbb{Q}_\ell)$ is a finite dimensional vector space over $\mathbb{Q}_\ell$, with $B_i := \dim H^i$ the ith Betti number. The cohomology groups vanish for $i > 2\dim V = 2(n-1)$. By linear algebra, we deduce that

$$\begin{aligned} Z_F(u) &= \prod_{i=0}^{2(n-1)} \exp\left(\sum_{j=1}^{\infty} \mathrm{Tr}(\Omega^{*\nu}, H^i)\frac{u^\nu}{\nu}\right) = \prod_{i=0}^{2(n-1)} \det(I - u\Omega^*; H^i)^{(-1)^{i+1}} \\ &= \frac{P_1(u)\cdots P_{2n-3}(u)}{P_0(u)P_2(u)\cdots P_{2n-2}(u)}, \end{aligned}$$

where $P_i(u) = \det(I - u\Omega^*; H^i) \in \mathbb{Q}[T]$ and $\deg P_i = B_i$. Write $P_i(u) = \prod_j (1 - \alpha_{ij}u)$. The values α_{ij} are called *the characteristic values of the zeta function* and they are the eigenvalues of the Frobenius morphism Ω^*. If V is non-singular, most cohomologies vanish:

Theorem 6.15 (Deligne). *Let $e = n - 1 = \dim V$.*

1. *$P_i(u) \in \mathbb{Z}[u]$, $P_0 = 1 - u$ and $P_{2e} = 1 - q^e u$.*

2. *(RH) For any $0 \leq i \leq 2e$ the characteristic values α_{ij} are q-Weil numbers of Weight i*

3. *(Functional equation) Let $\chi(V) = \sum_{i=1}^{2\ell}(-1)^i B_i$ and $\epsilon = 1$ if e is odd and $(-1)^N$ if e is even and N is the multiplicity of the eigenvalue $q^{e/2}$ of Ω^* acting on H^e. Then*
$$Z_F(q^{-e}u^{-1}) = \epsilon q^{\frac{e\chi(V)}{2}} u^{\chi(V)} Z_F(u).$$

4. *$Z_F(u) = P_e^*(u)^{(-1)^{e-1}} \prod_{j=0}^{e}(1 - q^j u)^{-1}$, where $P_e^* = \begin{cases} P_e(u), & 2 \nmid e \\ P_e(u)(1 - q^{2/e}u), & 2 \mid e. \end{cases}$*

We will not prove this theorem here! Grothendieck (1972) has calculated the Betti numbers:

$$B_e = \frac{d-1}{d}[(d-1)^n + (-1)^{n+1}] + \begin{cases} 0, & 2 \mid e \\ 1, & 2 \mid e, \end{cases}$$

where $d = \deg F$. Hence, the using the Lemma 6.14 we deduce the following general results, which generalizes the diagonal case:

Corollary 6.16. *Assume $V = \{F = 0\}$ is nonsingular projective hypersurface of dimension e defined over $\mathbb{F}_q$, then*

$$|\#V(\mathbb{F}_q) - \#\mathbb{P}^e(\mathbb{F}_q)| \leq B_e q^{e/2}.$$

6.6 Exercises

1. For f uniform in Ω_d, we have that $\mathrm{Var}(N_f) = q^{n-1} - q^{n-2}$.

2. Prove Lemma 6.6.

3. Let $A(d) = A_n(d)$ be the number of $b = (b_1, \dots, b_n) \in \mathbb{Z}^n$ with $0 < b_i < d$ and $b_1 + \cdots + b_n \equiv 0 \mod d$. Prove by induction on n that $A(d) = \frac{d-1}{d}[(d-1)^n - (-1)^{n-1}] < (d-1)^n$, and deduce

$$|N_f - q^n| < (1 - q^{-1})(d-1)^n q^{n/2},$$

for diagonal form f of degree d and n variables.

7 Riemann Hypothesis for sums in one variable

In this section, we formulate the Weil estimates for exponential sums in one variable and provide an elementary self-contained proof for hyperelliptic curves $Y^2 = f(X)$.

7.1 Statements

Let $\mathbb{F}_q$ be a finite field and let χ and ψ be multiplicative and additive characters of $\mathbb{F}_q$, respectively. Let $f \in \mathbb{F}_q[X]$ be a polynomial of degree d. The Weil bounds are given by:

Theorem 7.1. *Assume that χ is of order $e > 1$, $e \mid q-1$. Let $1 \leq m \leq d$ be the number of distinct roots of f in $\bar{\mathbb{F}}_q$. Assume further that $f \neq h^e$ for any $h \in \mathbb{F}_q[X]$. Then*

$$\left|\sum_{a \in \mathbb{F}_q} \chi(f(a))\right| \leq (m-1)\sqrt{q}. \tag{13}$$

Theorem 7.2. *Assume that $\psi \neq \psi_0$, that $d < q$ and that $(d, q) = 1$. Then*

$$\sum_{a \in \mathbb{F}_q} \psi(f(a)) \leq (d-1)\sqrt{q}. \tag{14}$$

Let us introduce the companion sums

$$S_\nu^{(1)}(f) = \sum_{a \in \mathbb{F}_{q^\nu}} \chi(\mathrm{Norm}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(f(a)))$$

$$S_\nu^{(2)}(f) = \sum_{a \in \mathbb{F}_{q^\nu}} \psi(\mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(f(a)))$$

Then we can form the associated zeta functions:

$$Z_f^{(i)}(u) = \exp\left(\sum \frac{S_\nu^{(i)}(f)u^\nu}{\nu}\right), \qquad i = 1, 2. \tag{15}$$

As before, those are rational functions and one may observe the analogy with the Riemann Hypothesis of ordinary zeta functions. More generally, we have:

Theorem 7.3 (Dwork)**.** *Let $f, g \in \mathbb{F}_q[X]$, and χ and ψ be multiplicative and additive characters of $\mathbb{F}_q$, respectively. Let*

$$S_\nu = \sum_{a \in \mathbb{F}_{q^\nu}} \chi(\mathrm{Norm}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(f(a)))\psi(\mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(g(a))).$$

Then there exist coprime polynomials $P, Q \in \mathbb{C}[T]$ with $P(0) = Q(0) = 1$ such that

$$Z(u) := \exp\left(\sum \frac{S_\nu u^\nu}{\nu}\right) = \frac{P(u)}{Q(u)}.$$

We will not prove Dwork's theorem in general in the course, but we will prove the special case when $g = 0$.

Recall that the Hasse-Davenport relation is equivalent to

$$Z(u) = \exp\left(\sum_{\nu=1}^{\infty} \frac{g_\nu(\chi, \psi)u^\nu}{\nu}\right) = 1 + g(\chi, \psi)u,$$

which is a special case of Dwork's theorem.

Another special is for Kloosterman sums: Let $K = K(\psi, \eta) = \sum_{x \in \mathbb{F}_p^\times} \psi(x)\eta(x^{-1})$, and consider the companion sums

$$K_\nu = \sum_{x \in \mathbb{F}_{q^\nu}^\times} \psi(\mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(x))\eta(\mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(x^{-1})).$$

Also in this case we may compute the zeta function explicitly, similarly:

Exercise 7.1. Prove that $Z(u) = \exp(\sum_{\nu \geq 1} K_\nu u^\nu/\nu) = \frac{1}{1+Ku+qu^2}$.

Lemma 7.4. *Let $f \in \mathbb{F}_q[X]$ be a non-constant monic polynomial, and let $e \mid q-1$. Then, for all $\nu \geq 1$ we have:*

$$\#\{(x, y) \in \mathbb{F}_{q^\nu}^2 : y^e = f(x)\} - q^\nu = \sum_\chi \sum_{x \in \mathbb{F}_{q^\nu}} \chi(\mathrm{Norm}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(f(x))),$$

where χ runs over all multiplicative characters of $\mathbb{F}_q$ satisfying $\chi \neq \chi_0$ and $\chi^e = \chi_0$.

Proof. Using the fact that characters of $\mathbb{F}_{q^\nu}$ of order dividing $q-1$ are coming from $\mathbb{F}_q$ via the norm map (Lemma 3.7) and the fact that $\#\{y \in \mathbb{F}_{q^\nu} : y^e = a\} = \sum_{\chi^e = \chi_0} \chi(a)$ we conclude that

$$\#\{(x, y) \in \mathbb{F}_{q^\nu}^2 : y^e = f(x)\} = \sum_{x \in \mathbb{F}_{q^\nu}} \sum_{\chi^e = \chi_0} \chi(\mathrm{Norm}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(f(x))).$$

Moving the contribution of the principal character to the left-hand side, finishes the proof. □

7.2 Proof the Riemann Hypothesis on number of points bound

In this section we assume the following proposition, which we will prove later, and deduce from it the Riemann Hytpothesis from multiplicative characters (13).

Proposition 7.5. *Let $\nu \geq 1$, let $f(X) \in \mathbb{F}_q[X]$ be a monic and non-constant polynomial of degree d, and let $e \mid q-1$ with $(e,d) = 1$. Then there exists $C = C(d,e) > 0$ such that*

$$\left|\#\{(x,y) \in \mathbb{F}_{q^\nu}^2 : y^e = f(x)\} - q^\nu\right| \leq Cq^{\nu/2}.$$

An auxiliary result that we will need is:

Lemma 7.6. *Let $w_1, \ldots, w_r \in \mathbb{C}$, let $A, B > 0$, and assume $|\sum w_i^\nu| \leq AB^\nu$ for all $\nu \gg 1$. Then $|w_i| \leq B$ for all i.*

Proof. Consider the complex power series

$$D(z) = \sum_{\nu\geq 1}\sum_{i=1}^{r} w_i^\nu z^\nu = \sum_{i=1}^{r} \frac{1}{1-w_i z}.$$

By hypothesis D converges absolutely in the disc $|z| < B^{-1}$, so D is analytic there, hence its poles w_i are outside the region. That is to say $1/|w_i| \geq 1/B$. □

The rest of the subsection is devoted to the proof of (13).

Let χ be multiplicative character of $\mathbb{F}_q$ of order $e \mid q-1$, let $f \in \mathbb{F}_q[X]$ be non-constant of degree d, let $1 \leq m \leq d$ be the number of distinct roots of f in $\bar{\mathbb{F}}_q$, and assume $f \neq h^e$. We need to prove the inequality (13).

We assume Dwork's theorem, which in fact gives in our setting[1] that

$$Z(u) = \exp\left(\sum_{\nu\geq 1} \frac{S_\nu}{\nu} u^\nu\right) = \prod_{1\leq j\leq m-1} (1 - w_j u),$$

with $w_j \in \mathbb{C}$ and where $S_\nu = \sum_{a\in\mathbb{F}_{q^\nu}} \chi(\mathrm{Norm}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(f(a)))$. Then, by Lemma 6.14, we have

$$S_\nu = -(w_1^\nu + \cdots + w_{m-1}^\nu)$$

so by Lemma 7.4,

$$N_\nu - q^\nu = -\sum_{\chi}(w_1(\chi)^\nu + \cdots + w_{m-1}(\chi)^\nu),$$

where the sum runs on characters $\chi^e = \chi_0$ and $\chi \neq \chi_0$. Assuming Proposition 7.5 and Lemma 7.6, we get that $|w_i(\chi)| \leq q^{1/2}$, so $|N_\nu - q^\nu| \leq (m-1)q^{\nu/2}$, and this finishes the proof when $\nu = 1$. □

[1] If time permits, we will prove this.

7.3 The Stephanov method

The goal now is to prove Proposition 7.5. So we fix $f \in \mathbb{F}_q[X]$ monic and non-constant of degree d and we fix $e \mid q-1$ with $\gcd(e,d) = 1$. Write

$$a(f) = \#\{(x,y) \in \mathbb{F}_q^2 : y^e = f(x)\} - q$$

Lemma 7.7. *We have $a(f) \geq -(e-1)\max_{\epsilon \in \mathbb{F}_q^\times} |a(\epsilon f)|$.*

Proof. Fix representatives $\{\epsilon_1 = 1, \epsilon_2, \dots, \epsilon_e\}$ for $\mathbb{F}_q^\times / \mathbb{F}_q^{e\times}$. Let $f_\epsilon = \epsilon^{-1} f$ so that $\deg f_\epsilon = d$. Let

$$C_\epsilon^* = \{(x,y) \in \mathbb{F}_q \times \mathbb{F}_q^\times : y^e = f_\epsilon(x)\}.$$

Then,

$$\#C_\epsilon^* = q + a(f_\epsilon) - N_f, \tag{16}$$

where $N_f = \#\{x \in \mathbb{F}_q : f(x) = 0\} \leq d$. Since, for each x with $f(x) \neq 0$, there is unique ϵ_i and e values of y such that $f(x) = \epsilon_i y^e$, we get that

$$\sum_{i=1}^{e} C_{\epsilon_i}^\times = \sum_{i=1}^{e} \sum_{\substack{y^e = \epsilon_i^{-1} f(x) \\ f(x) \neq 0}} 1 = \sum_{x, f(x) \neq 0} e = e(q - N_f).$$

Plugging in (16) and rearranging, we get that

$$0 = \sum_{\epsilon_i} a(f_{\epsilon_i}),$$

hence $a(f) = a(f_1) \geq -(e-1)\max_{i=2,\dots,n} |a(f_\epsilon)|$. □

In view of the lemma, to prove Proposition 7.5, it suffices to establish an upper bound of the form

$$N_f \leq q + O(q^{1/2}), \qquad q \gg 1, \tag{17}$$

where the implied constant depends only on d, e. We assume that $\gcd(e,d) = 1$ and $e \mid q-1$. In particular, this implies that the polynomial $Y^e - f(X)$ is absolutely irreducible (we leave it as an exercise).

The plan to prove (17) is surprisingly simple, called the polynomial method. Namely, we will construct an auxiliary polynomial $A \in \mathbb{F}_q[X]$ and a parameter $m \geq 1$ such that

1. $A \neq 0$, and
2. if $(x,y) \in \mathbb{F}_{q^\nu}$ is counted by N_f (that is, $y^e = f(x)$), then $(X-x)^m \mid A$.

This will imply that

$$N_f \leq e \frac{\deg A}{m}.$$

So, the proof would be reduced to bounding $\deg A$.

Let

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

Lemma 7.8. *Let $h_i(X) = k_{i0}(X) + X^q k_{i1}(X) + \cdots + X^{qK} k_{iK}$, $0 \leq i \leq e-1$ be polynomials with $\deg k_{ij} \leq \frac{q}{e} - d$. Suppose that $\sum_{i=0}^{e-1} h_i(X)g(X)^i = 0$. Then $k_{ij} = 0$ for all i, j.*

Proof. A typical summand is of the form

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

It suffices to show that the degree of the non-zero summands are all distinct, because then the sum cannot be zero. And indeed,

$$\deg \ell_{ij} = qj + \frac{id(q-1)}{e} + \deg k_{ij} = \frac{q}{e}(ej + id) + \deg k_{ij} - \frac{id}{e}.$$

By the assumption on $\deg k_{ij}$, we get

$$\deg \ell_{ij} \leq \frac{q}{e}(ej + id) + \frac{q}{e} - d$$

while since $i/e < 1$ we get that

$$\deg \ell_{ij} > \frac{q}{e}(ej + id) - d.$$

Hence it suffices to show that for $(i, j) \neq (i', j')$ we have $ej + id \neq ej' + i'd$. And indeed, if $ej + id \neq ej' + i'd$, then $id \equiv i'd \mod e$, so $i \equiv i \mod e$, as $\gcd(d, e) = 1$ and thus $i = i'$, as $0 \leq i, i' < e$. Thus, also $j = j'$. □

To produce the polynomial A with zeros of high multiplicity, it is natural to use derivatives. However, in positive characteristic there are some a-normalities, coming from the fact that $f^{(i)}(X) = 0$ for all $i \geq p$. So the Taylor expansion

$$f(X + T) = \sum_{i=0}^{\deg f} \frac{f^{(i)}(T)}{i!} X^i,$$

which holds true in characteristic 0, is not well-defined by derivatives in positive characteristic (as $f^{(i)} = 0$ and $i! = 0$ for $i \geq p$).There is an easy formal fix to that:

Definition 7.9. Let K be a field, and $f \in K[X]$ a polynomial. We define the i-th Hasse-Schmidt derivative $f^{[i]}(T) \in \mathbb{F}_q[T]$ of f to be the coefficient of X^i in the equation

$$f(X + T) = \sum_i f^{[i]}(T) X^i. \tag{18}$$

In particular, $f^{[0]}(X) = f(X)$, $f^{[1]}(X) = f'(X)$. If the characteristic of K is p, then we have $f^{[i]}(X) = f^{(i)}(X)/i!$ for all $i < p$.

Exercise 7.2. Prove that $(X^\alpha)^{[k]} = \binom{\alpha}{k} X^{\alpha - k}$ for any integers $\alpha, k \geq 0$ and deduce a close formula for $f^{[k]}$.

Lemma 7.10. $(\prod_{i=1}^r f_i)^{[k]} = \sum\limits_{k_1+\cdots+k_r=k} \prod_{i=1}^r f_i^{[k_i]}$.

Proof. The general case follows from the case $r = 2$ by induction. So to this end assume $r = 2$. By linearity, we may assume that $f_i = X^{\alpha_i}$, $i = 1, 2$. By the Exercise 7.4, the left-hand side equals to

$$\binom{\alpha_1 + \alpha_2}{k} X^{\alpha_1+\alpha_2-k},$$

and the right-hand side equals to

$$\sum_{i=0}^{k} \binom{\alpha_1}{i} \binom{\alpha_2}{i-k} X^{\alpha_1+\alpha_2-k}.$$

Since $\binom{\alpha_1+\alpha_2}{k} = \sum_{i=0}^{k} \binom{\alpha_1}{i}\binom{\alpha_2}{i-k}$, the proof is done. □

Lemma 7.11. $((X - c)^\alpha)^{[k]} = \binom{\alpha}{k}(X - c)^{\alpha-i}$.

Proof. By the previous lemma,

$$((X-c)^\alpha)^{[k]} = \sum_{k_1+\cdots+k_\alpha=k} \prod_{i=1}^{\alpha} (X-c)^{[k_i]} = \binom{\alpha}{k}(X-c)^{\alpha-i},$$

where in the last equality we used that the derivative is non-zero if and only if $k_i = 0, 1$. □

Exercise 7.3. Prove that for any $0 \leq \ell \leq t$, $a, f \in K[X]$, we have

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

for some $b \in K[X]$ with $\deg b = \deg a + \ell(\deg f - 1)$.

We now state formally the property that Hasse-Schmidt derivatives detect high order zeros.

Proposition 7.12. *For $x \in K$ and $f(X) \in K[X]$, we have that $f^{[\ell]}(x) = 0$ for all $0 \leq \ell \leq M - 1$ if and only if $(X - x)^M \mid f(X)$.*

Proof. Substituting x for T and $X - x$ for X in Definition 18 gives that $f^{[\ell]}(x)$ for all $0 \leq \ell \leq M - 1$ if and only if $(X - x)^M \mid f$. □

Lemma 7.13. *Let $\epsilon \in [1, e - 1] \cap \mathbb{Z}$ and $a \in \mathbb{F}_q[X]$ of degree ϵ. Let*

$$S = \{x \in \mathbb{F}_q : a(g(x)) = 0 \text{ or } f(x) = 0\}.$$

Let $M \geq d+1$ be an integer such that $(M+3)^2 \leq \frac{2q}{e}$. Then there exists nonzero $r \in \mathbb{F}_q[X]$ which has zero of order $\geq M$ for every $x \in S$ and satisfies $\deg r \leq \frac{\epsilon}{e}qM + 4dq$.

Proof. We consider $r(X) = h(X, X^q)$, where

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

and $\deg k_{ij} \le \frac{q}{e} - d$ with coefficients to be determined, and $K = \lfloor \frac{\epsilon}{e}(M + d + 1) \rfloor$ (recall that $g = f^{q-1}e$).

First, the zeros of f appear with multiplicity $\ge M$. Second,

$$\begin{aligned} \deg r &\le dM + \frac{q}{e} - d + \frac{(e-1)(q-1)d}{e} + \frac{\epsilon}{e} q(M + d + 1) \\ &\le \frac{\epsilon M q}{e} + M^2 + \frac{qd}{e}(1/d + (e-1) + \epsilon + \epsilon/d) \\ &\le \frac{\epsilon M q}{e} + 4dq. \end{aligned}$$

Hence, it remains to show we can choose the k_{ij} so that the other $x \in S$ will have multiplicity $\ge M$ and such that $r \ne 0$.

We start by computing derivatives: Since $g = f^{\frac{q-1}{e}}$, by Exercise 7.7, we have

$$(f(X)^M \sum_{j=0}^{K} k_{ij}(X) g(X)^i)^{[\ell]} = f(X)^{M-\ell} k_{ij\ell}(X) g(X)^i,$$

where $k_{ij\ell}(X)$ has degree $\deg k_{ij} + \ell(d-1)$. Hence, by Exercise 7.6, whenever $0 \le \ell < M \le q$, we have

$$r^{[\ell]}(X) = f(X)^{M-\ell} \sum_{i=0}^{e-1} \sum_{j=0}^{K} k_{ij\ell}(X) g(X)^i X^{qj}. \tag{19}$$

Fix x with $a(g(x)) = 0$. By Proposition 7.12, it suffices to prove that $r^{[\ell]}(x) = 0$, $0 \le \ell < M$. Let $y \in \mathbb{F}_q$ be such that $a(y) = 0$. Then $y^\epsilon = \sum_{t=0}^{\epsilon-1} c_t y^t$. Therefore, for any $i \ge 0$ we may write

$$y^i = \sum_{t=0}^{\epsilon-1} c_{ti} y^t.$$

In particular, we apply this to $y = g(x)$ to get that $g(x)^i = \sum_{t=0}^{\epsilon-1} c_{ti} g(x)^t$, and thus, as $x^q = x$, may plug this in (19), to get

$$r^{[\ell]}(x) = f(x)^{M-\ell} \sum_{t=0}^{\epsilon-1} g(x)^t s_{t\ell},$$

where

$$s_{t\ell}(X) = \sum_{i=0}^{e-1} \sum_{j=0}^{K} c_{ti} k_{ij\ell}(X) X^j.$$

In particular, if all $s_{t\ell}$ are the zero polynomial, but $r \neq 0$, we are done.

We have

$$\deg s_{t\ell}(X) \leq \max_{i,j} \deg k_{ij\ell} + K \leq \max_{i,j} \deg k_{ij} + \ell(d-1) + K \leq \frac{q}{e} + \ell(d-1) + K - 1.$$

Let B be the total number of coefficients of all the $s_{t\ell}$. Then

$$\begin{aligned} B &\leq \sum_{t=0}^{\epsilon-1} \sum_{\ell=0}^{M-1} \deg s_{t\ell} \leq \sum_{t=0}^{\epsilon-1} \sum_{\ell=0}^{M-1} (\frac{q}{e} + \ell(d-1) + K - 1) \\ &= \epsilon M(q/e + K - 1) + \epsilon \sum_{\ell=0}^{M-1} \ell(d-1) < \epsilon M(q/e + K) + \frac{M^2}{2}(d - \\ &\leq \frac{\epsilon q}{e} M + \frac{\epsilon^2}{e} M(M + d + 1) + \frac{M^2}{2}(d-1)\epsilon \\ &\leq \frac{\epsilon q}{e} M + \epsilon M(d+1) + \epsilon M^2(\frac{d-1}{2} + \frac{\epsilon}{e}) \end{aligned}$$

$$\leq \frac{\epsilon q}{e} M + \epsilon M(d+1) + \epsilon M^2(\frac{d+1}{2}).$$

Let C be the total number of coefficients of all k_{ij}, then

$$C \geq (\frac{q}{e} - d)e(K+1) \geq \frac{q\epsilon}{e} M + \frac{q\epsilon}{e}(d+1) - 2Md\epsilon.$$

If $C > A$, then $s_{t\ell} = 0$ defines a homogenous system of linear equations in the coefficients of k_{ij} with more variable than equations, hence it admits a non-trivial solution. By Lemma 7.8, $r \neq 0$ in this case, and we are done.

And indeed, as $(M+3)^2 \leq \frac{2q}{e}$, we have that $M^2 + 6M < 2q/e$. Hence $M^2(d+1)/2 + 3M(d+1) < \frac{q}{e}(d+1)$ which implies that $B < C$. □

Corollary 7.14. *For every M as in the lemma we have $\#S \leq \frac{\deg r}{M} \leq \frac{\epsilon}{e} q + 4\frac{dq}{M}$.*

To conclude the proof of Proposition 7.5, we apply the corollary with $M = \lfloor \sqrt{\frac{2q}{e}} - 3 \rfloor$, so that $M \geq d+1$ if q is sufficiently large, and thus $\#S \leq \frac{\epsilon}{e} q + 4e^{1/2} q^{1/2}$.

Take $a(Y) = Y - 1$ so that $\epsilon = 1$. Then $S = \{x \in \mathbb{F}_q : g(x) = 1 \text{ or } f(x) = 0\}$. Hence

$$N \leq e \cdot \#S \leq q + O(\sqrt{q}).$$

□

7.4 Exercises

1. Prove that
$$Z(u) = \exp\left(\sum_{\nu \geq 1} \frac{K_\nu u^\nu}{\nu}\right) = \frac{1}{1 + Ku + qu^2},$$

where $K_\nu = \sum_{x \in \mathbb{F}_{q^\nu}^\times} \psi(\mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(x))\eta(\mathrm{Tr}_{\mathbb{F}_{q^\nu}/\mathbb{F}_q}(x^{-1}))$ and $K = K_1$ is the Kloosterman sum.

Hint: Mimic the proof of the Hasse-Davenport relation with the character $\mu(X^d + a_1X^{d-1} + \cdots + a_d) = \psi(a_1)\eta(a_{d-1}/a_d)$ extended to $G = \{f/g : f, g \in \mathscr{M}, f(0)g(0) \neq 0\}$.

2. Assume a bound of the form $|\#\{(x, y) \in \mathbb{F}_{q^\nu}^2 : y^q - y = f(x)\} - q^\nu| \leq Cq^{\nu/2}$, and prove (14). Hint follow similar steps as in the proof of (14).

3. Assume that $\gcd(e, d) = 1$ and $e \mid q - 1$ and $f(X) \in \mathbb{F}_q[X]$ is of degree d. Then $Y^e - f(X)$ is absolutely irreducible (that is to say, $Y^e - f(X)$ is irreducible in the ring $\bar{\mathbb{F}}_q[Y, X]$).

4. Prove that $(X^\alpha)^{[k]} = \binom{\alpha}{k}X^{\alpha-k}$ for any integers $\alpha, k \geq 0$ and deduce a close formula for $f^{[k]}$.

5. Open brackets in $f_1(T + X)f_2(T + X)$ to give an alternative proof to Lemma 7.10.

6. Show that if $r(X) = h(X, X^q) \in \mathbb{F}_q[X]$, where $h(X, Y) \in \mathbb{F}_q[X, Y]$, then, for any $0 \leq \ell < q$, we have $r^{[\ell]}(X) = h_X^{[\ell]}(X, X^q)$, where $h_X^{[\ell]}$ is the Hasse-Schmidt derivative with respect to X (i.e., as an element of $\mathbb{F}_q(Y)[X]$).

   Hint: Use that $(X^q)^{[\ell]} = 0$ for $\ell < q$.

7. Prove that for any $0 \leq \ell \leq t$, $a, f \in K[X]$, we have

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

for some $b \in K[X]$ with $\deg b = \deg a + \ell(\deg f - 1)$.