1. Week One

What is an exponential sum?
We often see things like

$$S = \sum_{n=1}^{N} \exp(2\pi i \alpha_n), \text{ where } \alpha_n \in \mathbb{R}$$

Recall that

$$\begin{aligned}\exp(2\pi i\alpha) &= \sum_{n=0}^{\infty} \frac{(2\pi i\alpha)^n}{n!} \\ &= \sum_{n=even} (-1)^{n/2} \frac{(2\pi\alpha)^n}{n!} + i \sum_{n=odd} (-1)^{(n-1)/2} \frac{(2\pi\alpha)^n}{n!} \\ &= \cos(2\pi\alpha) + i \sin(2\pi\alpha)\end{aligned}$$

It follows that $|\exp(2\pi i\alpha)| = 1$, for all $\alpha \in \mathbb{R}$.
For $\alpha = 1/2$, we get Euler's formula $\exp(i\pi) + 1 = 0$.
$S \in \mathbb{C}$ and $|S| \leq N$ because each summand has absolute value 1.

Exercise 1. 1. $|S| = N \Leftrightarrow \exp(2\pi i\alpha_1) = \exp(2\pi i\alpha_n), \forall n \geq 1$
2. $\sum_{n=1}^{N} \exp(2\pi i n/N) = 0$(This follows from the geometric series).

The subject of exponential sums seeks to give improved estimates on S from knowledge of α_n. for the most part we will be interested in upper bounds but we will also have conditions so that $|S| \neq 0$ or even lower bounds.

Definition 1. If $\alpha_i \in \mathbb{Q}$ then S is called a rational exponential sum.

We will mostly be looking at rational exponential sums i.e $\alpha_n = a_n/M$ where $a_n, M \in \mathbb{Z}$. Since $\exp(2\pi i a_n/M)^M = 1$ we have that $\exp(2\pi i a_n/M)$ is a root of unity. So $S = \sum \exp(2\pi i a_n/M) \in \mathbb{Z}[\exp(2\pi i/M)]$
Now we state a theorem due to Weil.

Theorem 1 (Weil)**.** Let $f(x) \in \mathbb{Z}[x]$ of degree $d < p$ and p a prime number such that $f(x) \bmod p$ is not a constant. then

$$|\sum_{n=0}^{p-1} \exp(2\pi i f(n)/p)| \leq (d-1)\sqrt{p}$$

This theorem is closely related to the Riemann Hypothesis for curves over finite fields.

Definition 2. The quadratic Gauss sum is defined as follows $G = \sum_{n=1}^{p} \exp(2\pi i n^2/p)$.

Theorem 2 (Gauss)**.** Let $\sqrt{p}$ denote the positive square root of the prime number p, then

$$G = \begin{cases} \sqrt{p}, & p \equiv 1 \bmod 4 \\ i\sqrt{p}, & p \equiv 3 \bmod 4 \\ 0, & p = 2. \end{cases}$$

We now define the Legendre symbol.

Definition 3.

$$\left(\frac{n}{p}\right) = \begin{cases} 0, & n \equiv 0 \bmod p \\ 1, & \exists\, m \not\equiv 0 \bmod p \text{ and } m^2 \equiv \bmod p \\ -1, & \text{otherwise.} \end{cases}$$

Next, let

$$\begin{aligned} G' &= \sum_{n=1}^{p} \left(\frac{n}{p}\right) \exp\left(2\pi i n/p\right) \\ &= \sum_{n=1}^{p} \left(\frac{n}{p}\right) \exp\left(2\pi i n/p\right) + \sum_{n=1}^{p} \exp\left(2\pi i n/p\right) \\ &= \sum_{n=1}^{p} \left(\frac{n}{p}\right) \exp\left(2\pi i n^2/p\right) \\ &= G. \end{aligned}$$

From Gauss's Theorem we know for an odd prime p, $G^2 = (-1)^{(p-1)/2}p$. Now $G \in \mathbb{Z}[\exp(2\pi i/p)]$ For an odd prime $q \neq p$ let us define $R = \mathbb{Z}[\exp(2\pi i/p)]/(q)$ and in this ring we have $(x+y)^q = x^q + y^q$. Let $\overline{G}$ be the image of G in R. Then,

$$\overline{G}^q = \sum_{n=1}^{p} \left(\left(\frac{n}{p}\right)\right)^q \zeta^{nq}$$

where ζ is the image of $\exp(2\pi i/p)$ in R. Note also that as n varies from 1 to p, $nq \bmod p$ also varies from 1 to p. We have

$$\begin{aligned}
\overline{G}^q &= \sum_{n=1}^{p} \left(\frac{nq^{-1}}{p}\right) \zeta^n \\
&= \sum_{n=1}^{p} \left(\frac{q^{-1}}{p}\right)\left(\frac{n}{p}\right) \zeta^n \\
&= \left(\frac{q}{p}\right)^{-1} \sum_{n=1}^{p} \left(\frac{n}{p}\right) \zeta^n \\
&= \left(\frac{q}{p}\right) \sum_{n=1}^{p} \left(\frac{n}{p}\right) \zeta^n \\
&= \left(\frac{q}{p}\right) \overline{G}
\end{aligned}$$

and hence we have $(\overline{G}^2)^{(q-1)/2} = \left(\frac{q}{p}\right)$ which, combined with Gauss's theorem, gives $(-1)^{(p-1)(q-1)/4}\, p^{(q-1)/2} = \left(\frac{q}{p}\right) \bmod q$.

We also have Euler's theorem $p^{(p-1)/2} \equiv \left(\frac{p}{q}\right) \bmod q$. Hence we have the quadratic reciprocity law.

$$(-1)^{(p-1)(q-1)/4} \left(\frac{p}{q}\right) = \left(\frac{q}{p}\right).$$

Now we state and prove Euler's theorem.

Theorem 3 (Euler)**.** $\left(\frac{n}{p}\right) \equiv n^{(p-1)/2} \bmod p$.

Proof. It is clear in the case $n \equiv 0 \bmod p$. If $n \equiv m^2 \bmod p$ then

$$n^{(p-1)/2} \equiv m^{(p-1)} \equiv 1 \bmod p$$

Again, $(n^{(p-1)/2})^2 \equiv 1 \bmod p$. Since $x^{(p-1)/2} - 1 = 0$ has at most $(p-1)/2$ roots in $\mathbb{Z}/(p)$ all the $(p-1)/2$ squares are roots. So there is no room for other roots. □

We shall now generalise the fact

$$\begin{aligned}
\sum_{m=0}^{n} \exp(2\pi m(b-a)/n) &= n, \quad a \equiv b \bmod n \\
&= 0 \text{ otherwise.}
\end{aligned}$$

Let $\mathbb{C}^* = \mathbb{C} - \{0\}$ be considered as a group under multiplication. Let $T = \{z \in \mathbb{C} : |z| = 1\}$ and define $\mu_n = \{z \in \mathbb{C} : z^n = 1\}$. Then $\mu = \bigcup_{n \geq 1} \mu_n \subset T$ and μ_n, μ, T are all subgroups of $\mathbb{C}^*$.

Definition 4. A character of an abelian group G is a homomorphism $\chi : G \to \mathbb{C}^*$

Note that $\chi(G) \subseteq \mu_{|G|}$. The character χ_0 given by $\chi_0(g) = 1$, $\forall g \in G$ is called the principal or trivial character.

If $G = (\mathbb{Z}/n\mathbb{Z}, +)$, for any $a \in \mathbb{Z}$ we define $\chi_a(m) = \exp(2\pi iam/n)$ which is a character since $\chi_a(m_1 + m_2) = \chi_a(m_1)\chi_a(m_2)$.

Definition 5. We define the dual $\widehat{G}$ of G as the set of all characters of G.

Exercise 2. $\widehat{G}$ is an abelian group under the multiplication $\chi_1\chi_2(g) = \chi_1(g)\chi_2(g)$ with identity χ_0.

Theorem 4. $\widehat{G}$ is isomorphic to G for any finite abelian group G.

Corollary 5. The dual of $\widehat{G}$ is isomorphic to G.

We can give a simple direct proof of the corollary since, unlike the theorem, the isomorphism is canonical. For any $g \in G$ define a character ψ_g of $\widehat{G}$ i.e a map $\psi_g : \widehat{G} \to \mathbb{C}^*$ by $\psi_g(\chi) = \chi(g)$. It follows that $g \mapsto \psi_g$ is an isomorphism from G to the dual of $\widehat{G}$.

The theorem follows by the classification of finite abelian groups once we prove the following two lemmas.

Lemma 6. $\widehat{G \times H} \cong \widehat{G} \times \widehat{H}$.

Proof. Define a map $\widehat{G \times H} \to \widehat{G} \times \widehat{H}$ by $\chi \to (\chi|G, \chi|H)$ and define a map $\widehat{G} \times \widehat{H} \to \widehat{G \times H}$ by $(\chi_1, \chi_2) \mapsto ((g, h) \mapsto \chi_1(g)\chi_2(h))$. It is easy to check that these are homomorphisms and inverses of each other. Hence the lemma follows. □

Lemma 7. If G is a finite cyclic group then $\widehat{G}$ is isomorphic to G .

Proof. If G is finite cyclic then $G \cong \mathbb{Z}/n\mathbb{Z}$. Define a map G to $\widehat{G}$ by $a \mapsto \chi_a$, defined above. Since $\chi_a\chi_b = \chi_{(a+b)}$ it follows that the map is a homomorphism. The kernel of this map is $n\mathbb{Z}$. So $\mathbb{Z}/n\mathbb{Z} \hookrightarrow \widehat{G}$. Now to show this map is surjective, let $\chi \in \widehat{G}$. Look at $\chi(1) = \xi$. We have

$$\begin{aligned} \xi^n &= (\chi(1))^n \\ &= \chi(n) \\ &= \chi(0) \\ &= 1 \end{aligned}$$

This implies $\xi = \exp(2\pi ia/n)$ for some $a \in \mathbb{Z}$. So we have,

$$\begin{aligned} \chi(m) &= (\chi(1))^m \\ &= \xi^m \\ &= \exp(2\pi iam/n) \\ &= \chi_a(m) \end{aligned}$$

So $\chi = \chi_a$. Hence the map is surjective. □

Theorem 8. The following identities hold

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

b) $\dfrac{1}{|\widehat{G}|}\displaystyle\sum_{g\in\widehat{G}}\chi(g)=\begin{cases}0 & g\neq 1,\ g\in G\\ 1 & g=1\end{cases}$

c) $\dfrac{1}{|G|}\displaystyle\sum_{g\in G}\chi_1(g)\overline{\chi_2(g)}=\begin{cases}0 & \chi_1\neq\chi_2,\ \chi_1,\ \chi_2\in\widehat{G}\\ 1 & \chi_1=\chi_2,\ \chi_1,\ \chi_2\in\widehat{G}\end{cases}$

d) $\dfrac{1}{|\widehat{G}|}\displaystyle\sum_{g\in\widehat{G}}\chi(g_1)\overline{\chi(g_2)}=\begin{cases}0 & g_1\neq g_2,\ g_1,g_2\in G\\ 1 & g_1=g_2,\ g_1,g_2\in G\end{cases}$

Proof. If $|z|=1$ then we have that $\bar{z}=z^{-1}$. Our first goal will be to reduce b), c) and d) to a). Now consider the left hand side of c).

$$\begin{aligned}\frac{1}{|G|}\sum_{g\in G}\chi_1(g)\overline{\chi_2(g)}&=\frac{1}{|G|}\sum_{g\in G}\chi_1(g)\chi_2(g)^{-1}\\ &=\frac{1}{|G|}\sum_{g\in G}\chi_1(g)(\chi_2)^{-1}(g)\\ &=\frac{1}{|G|}\sum_{g\in G}\chi_1(\chi_2)^{-1}(g)\end{aligned}$$

Now we observe that c) follows from a) applied to $\chi=\chi_1(\chi_2)^{-1}$. Likewise, d)follows from b) applied to $g=g_1(g_2)^{-1}$, and b) is a) applied to $\widehat{G}$. The two groups have the same cardinality since they are isomorphic. So it remains to prove a).
If $\chi=\chi_0$, then it is clear that

$$\frac{1}{|G|}\sum_{g\in G}\chi_0(g)=1$$

If $\chi\neq\chi_0$ then $\exists g_0\in G,\ \chi(g_0)\neq 1$

$$\begin{aligned}\frac{1}{|G|}\sum_{g\in G}\chi(gg_0)&=\frac{1}{|G|}\sum_{g\in G}\chi(g)\chi(g_0)\\ &=\frac{\chi(g_0)}{|G|}\sum_{g\in G}\chi(g)\end{aligned}\tag{1}$$

On the other hand, $\frac{1}{|G|}\sum_{g\in G}\chi(gg_0)=\frac{1}{|G|}\sum_{g\in G}\chi(g)$ as $g\mapsto gg_0$ is a permutation of G. It follows that

$$(\chi(g_0)-1)\left(\frac{1}{|G|}\sum_{g\in G}\chi(g)\right)=0$$

and, since $\chi(g_0)\neq 1$, we finally get

$$\frac{1}{|G|}\sum_{g\in G}\chi(g)=0.$$

□

2. Week two

Definition 6. $\mathbb{C}[G] = \{F : G \to \mathbb{C} : F \text{ a function}\}$

Let $G = \{g_1, g_2, \ldots, g_n\}$ with $n = |G|$. Then F is characterized by its values $F(g_1), F(g_2), \ldots, F(g_n)$. So $F \mapsto (F(g_1), F(g_2), \ldots, F(g_n))$ defines a bijection from $\mathbb{C}[G]$ to $\mathbb{C}^n$. We will consider $\mathbb{C}[G]$ as a vector space over $\mathbb{C}$. The operations in this vector space are defined as follows:

Let $F_1, F_2 \in \mathbb{C}[G]$ and $\lambda \in \mathbb{C}$,

- $(F_1 + F_2)(g) = F_1(g) + F_2(g)$
- $(\lambda F_1)(g) = \lambda F_1(g)$

We define an inner product on this vector space:

Definition 7. If F_1, f_2 be in $\mathbb{C}[G]$. Then,

$$\langle F_1, F_2 \rangle = \frac{1}{|G|} \sum_{g \in G} F_1(g)\overline{F_2(g)}$$

This is an Hermitian scalar product on $\mathbb{C}[G]$. An Hermitian scalar product on a complex vector space V is a function $<,>: V \times V \to \mathbb{C}$ such that $\forall v, w, z \in V$ and $\forall \lambda \in \mathbb{C}$ then,

- $< v + w, z >=< v, z > + < w, z >$
- $\lambda < v, z >=< \lambda v, z >$
- $< v, z >= \overline{< z, v >}$
- $< z, z > \geq 0$, and, $< z, z >= 0$ if and only if $z = 0$.

.

Theorem 9. $\widehat{G}$ is an orthonormal basis for $(\mathbb{C}[G], <, >)$.

Proof. First of all, $\widehat{G}$ is an orthonormal subset of $(\mathbb{C}[G], <, >)$ by the results of last week. Also, any orthonormal set is linearly independent. Finally $\#\widehat{G} = \#G = \dim \mathbb{C}[G]$, so $\widehat{G}$ is a basis. □

Corollary 10. $\forall F \in \mathbb{C}[G]$ we have $F = \sum_{\chi \in \widehat{G}} < F, \chi > \chi$.

This expresion is called the *Fourier expansion of* F. For all $g \in G$, we have $F(g) = \sum_{\chi \in \widehat{G}} < F, \chi > \chi(g)$. The map $\chi \mapsto < F, \chi >$ is a function $\widehat{G} \mapsto \mathbb{C}$. This is called the *finite Fourier transform of* F.

Let R be a finite commutative ring. We have two groups associated to it:

- $(R, +)$; the additive group of R, which is a finite abelian group.
- $(R^*, \cdot)$; the group of units of R under multiplication, which is also a finite abelian group.

Example 1. Let p be a prime number. Then, for $\mathbb{Z}/p\mathbb{Z}$,

- $a \mapsto e^{2\pi i a/p} \in \widehat{(\mathbb{Z}/p\mathbb{Z}, +)}$
- $a \mapsto \left(\frac{a}{p}\right) \in \widehat{\mathbb{Z}/p\mathbb{Z}^*}$

Let $\widehat{(R,+)} = \{\chi : R \to \mathbb{C}^* | \chi(x+y) = \chi(x)\chi(y)\}$. We can make $\widehat{(R,+)}$ into an R-module as follows: given $a \in R$ and $\chi \in \widehat{(R,+)}$, $a\chi(x) = \chi(ax)$. Then $a\chi(x+y) = \chi(a(x+y)) = \chi(ax+ay) = \chi(ax)\chi(ay) = a\chi(x)a\chi(y)$. So $a\chi$ is a character.

We know $\widehat{(R,+)} \cong (R,+)$ as groups. We can ask: is $\widehat{(R,+)}$ isomorphic to $(R,+)$ as an R-module? In other words, does there exist a character χ of $(R,+)$ such that $\{a\chi | a \in R\} = \widehat{(R,+)}$?

Example 2. Let $R = \mathbb{Z}/n\mathbb{Z}$. All characters are of the form $\chi_a : x \mapsto e^{2\pi i a x/n}$ (so $\chi_a = a\chi_1$).

Let's look at finite fields. Let $\mathbb{F}_q = \mathbb{F}_{p^n}$ with p a prime. Recall that $\mathrm{Tr}_{\mathbb{F}_{p^n}/\mathbb{F}_p}(x) = x + x^p + \cdots + x^{p^{n-1}}$.

Theorem 11. $\widehat{(\mathbb{F}_q,+)}$ is isomorphic to $\mathbb{F}_q$ as an $\mathbb{F}_q$ vector space, with generator $\chi_1 : x \mapsto e^{2\pi i Tr_{\mathbb{F}_q/\mathbb{F}_p}(x)/p}$ That is, all characters of $\widehat{(\mathbb{F}_q,+)}$ are of the form $x \mapsto e^{2\pi i Tr_{\mathbb{F}_q/\mathbb{F}_p}(ax)/p}$ for $a \in \mathbb{F}_q$.

Proof. $\widehat{(\mathbb{F}_q,+)}$ is isomorphic to $(\mathbb{F}_q,+)$ as abelian groups. So $\widehat{(\mathbb{F}_q,+)}$ has q elements. Also $\widehat{(\mathbb{F}_q,+)}$ is an $\mathbb{F}_q$ vector space. Therefore, it has to be 1-dimensional. Any non-zero element of $\widehat{(\mathbb{F}_q,+)}$ is a basis for it as an $\mathbb{F}_q$ vector space. So I need to show that $\chi_1 \neq \chi_0$, i.e. we need show that $Tr_{\mathbb{F}_q/\mathbb{F}_p}(x)$ is not always zero. $Tr_{\mathbb{F}_q/\mathbb{F}_p}(x)$ is a polynomial of degree p^{n-1} so it has at most p^{n-1} roots. On the other hand $\#\mathbb{F}_q = p^n > p^{n-1}$. So there exists $x \in \mathbb{F}_q$ such that $Tr_{\mathbb{F}_q/\mathbb{F}_p}(x) \neq 0$. □

But in general the answer to the question "is $\widehat{(R,+)}$ isomorphic to $(R,+)$ as an R-module?" is no. Rings that answer "yes" to this question are called *Frobenius rings*.

Since in what follows, we will refer to the inner product on both $\mathbb{C}[R^*]$ and $\mathbb{C}[R]$, we will distinguish them by an appropriate subscript.

Let R be a finite commutative ring. Let $\chi \in \widehat{(R,+)}$. Then χ, when restricted to R^*, gives us a function for which we can calculate its Fourier expansion in terms of the characters of R^*: $\chi|_{R^*} = \frac{1}{|R^*|}\sum_{\psi \in \widehat{R^*}} \langle \chi, \psi \rangle_{\mathbb{C}[R^*]}\, \psi$ (remember that $\chi|_{R^*}$ is just a function on R^*, not a character in $\widehat{(R^*,\cdot)}$).

Definition 8. Given an additive character $\chi \in \widehat{(R,+)}$ and a multiplicative chararcter $\psi \in \widehat{R^*}$ we call $G(\chi,\psi) = \langle \chi, \psi \rangle_{\mathbb{C}[R^*]} = \sum_{x \in R^*} \chi(x)\overline{\psi(x)} \in \mathbb{C}$ the *Gauss sum of* χ *and* ψ.

We would like to extend $\psi \in \widehat{R^*}$ to all of R. We can do this by defining $\psi(x) = 0$ whenever $x \notin R^*$. In particular, if R is a field then $R - R^* = \{0\}$. Then all we are doing is defining $\psi(0) = 0$ and the identity $\psi(xy) = \psi(x)\psi(y)$ will remain true $\forall x, y \in R$.

Now, we expand this extended ψ in terms of the characters $\chi \in \widehat{(R,+)}$. We get $\psi = \frac{1}{|R|}\sum_{\chi \in \widehat{(R,+)}} \langle \psi, \chi \rangle_{\mathbb{C}[R]}\, \chi$ where $\langle \psi, \chi \rangle_{\mathbb{C}[R]}$ is the inner product on $\mathbb{C}[R]$. Then $\langle \psi, \chi \rangle_{\mathbb{C}[R]} = \sum_{x \in R} \psi(x)\overline{\chi(x)} = \sum_{x \in R^*} \psi(x)\overline{\chi(x)} = \overline{G(\chi,\psi)}$, the last equality being true since $\psi(x) = 0$ if $x \in R - R^*$.

Theorem 12. Let $\mathbb{F}_q$ be a finite field and $\chi \in \widehat{(\mathbb{F}_q, +)}$, $\chi \neq \chi_0$. Then $\forall \chi \in \widehat{\mathbb{F}_q^*}$, $|G(\chi, \psi)| = q^{1/2}$.

Proof.

$$\begin{aligned}
|G(\chi,\psi)|^2 &= G(\chi,\psi)\overline{G(\chi,\psi)} \\
&= \left(\sum_{x\in\mathbb{F}_q^*} \chi(x)\overline{\psi(x)}\right)\left(\sum_{x\in\mathbb{F}_q^*} \overline{\chi(x)}\psi(x)\right) \\
&= \sum_{x,y\in\mathbb{F}_q^*} \psi(x)\chi(-x)\psi(y^{-1})\chi(y) \\
&= \sum_{x,y\in\mathbb{F}_q^*} \psi(xy^{-1})\chi(y-x) \qquad (\text{take } xy^{-1}=u, \text{ then } x=yu) \\
&= \sum_{u,y\in\mathbb{F}_q^*} \psi(u)\chi(y-uy) \\
&= \sum_{u\in\mathbb{F}_q^*} \psi(u)\sum_{y\in\mathbb{F}_q^*} \chi(y(1-u))
\end{aligned}$$

$y \mapsto \chi((1-u)y)$ is an additive character, which is not χ_0 if $u \neq 1$. Then,

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

Then, $|G(\chi,\psi)|^2 = \sum_{u\in\mathbb{F}_q^*, u\neq 1} \psi(u)(-1) + \psi(1)(q-1) = q - \sum_{u\in\mathbb{F}_q^*} \psi(u) = q$. □

Exercise 3. Work out $G(\chi,\psi)$ when $\psi = \psi_0$ or $\chi = \chi_0$.

Let $S_n = \sum_{x\in\mathbb{F}_p} e^{2\pi i x^n/p}$, $n|(p-1)$. Let $\chi_1(x) = e^{2\pi i x/p}$.

Theorem 13.

$$S_n = \sum_{\psi\in\widehat{\mathbb{F}_q^*}, \psi^n=1} G(\chi_1, \psi)$$

Corollary 14.

$$|S_n| \leq (n-1)p^{1/2}$$

Proof. (of the theorem) The map $x \mapsto x^n$ is a homomorphism with kernel $\Gamma = \{x \in \mathbb{F}_p^* | x^n = 1\}$, $|\Gamma| = n$. Then $S_n = n\sum_{c\in(\mathbb{F}_p^*)^n} e^{2\pi i c/p}$, where $(\mathbb{F}_p^*)^n$ are the n-th powers in $\mathbb{F}_p^*$ since (we claim)

$$\sum_{\psi\in\widehat{\mathbb{F}_p^*}, \psi^n=1} \psi(c) = \begin{cases} n & \text{if } c \in (\mathbb{F}_p^*)^n \\ 0 & \text{if } c \notin (\mathbb{F}_p^*)^n \end{cases}$$

We have the following sequence $\mathbb{F}_p^* \xrightarrow{\overline{\pi}} \mathbb{F}_p^*/\Gamma \xrightarrow{\overline{\psi}} \mathbb{C}^*$, where $\overline{\pi}$ is the projection over the quotient $\mathbb{F}_p^*/\Gamma$. Then $\widehat{\mathbb{F}_p^*/\Gamma} \cong \{\psi \in \widehat{\mathbb{F}_p^*} | \text{trivial on } \Gamma\} \cong \{\psi \in \widehat{\mathbb{F}_p^*} | \psi^n = 1\}$. Then, $(*) = \sum_{c\in\mathbb{F}_p^*} \sum_{\psi^n=1} e^{2\pi ic/p} = \sum_{\psi^n=1} \sum_{c\in\mathbb{F}_p^*} \chi_1(c)\overline{\psi(c)} = \sum_{\psi^n=1} G(\chi_1, \psi)$. □

Notice that this corollary is a special case of Weil's theorem. We give an outline of yet another alternative proof and leave the details as an exercise for the reader .

Exercise 4. i. $S_n = \sum_{x\in\mathbb{F}_p} e^{2\pi i(cx)^n/p}$ for any $c \in \mathbb{F}_p^*$.

ii. $\frac{p-1}{n}|S_n|^2 \leq \sum_{t\in\mathbb{F}_p^*} \left|\sum_x e^{2\pi itx^n/p}\right|^2$

iii. Evaluate the RHS.

3. Week Three

For a prime p, a quadratic Gauss sum is defined by

$$G = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) e^{2\pi ia/p}$$

where

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a square mod } p, a \neq 0 \\ -1 & \text{if } a \text{ is not a square mod } p, a \neq 0 \\ 0 & \text{if } a = 0 \end{cases}$$

First day of class: we outlined the proof of quadratic reciprocity, using two additional facts which were only mentioned. We prove these here:

$$G^2 = (-1)^{(p-1)/2} p \tag{2}$$

$$G = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \bmod 4 \\ i\sqrt{p} & \text{if } p \equiv 3 \bmod 4 \end{cases} \tag{3}$$

Define

$$\eta(a) = \left(\frac{a}{p}\right), \eta \in \widehat{\mathbb{F}_p^*}$$
$$\chi_1(a) = e^{2\pi ia/p}, \chi_1 \in \widehat{\mathbb{F}_p}$$

Then

$$G = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) e^{2\pi ia/p} = G(\eta, \chi_1)$$

Conjugating,

$$\overline{G} = \sum_{a=1}^{p-1} \left(\frac{a}{p}\right) e^{-2\pi ia/p} = \sum_{a=1}^{p-1} \left(\frac{-a}{p}\right) e^{2\pi ia/p} = \sum_{a=1}^{p-1} \left(\frac{-1}{p}\right)\left(\frac{a}{p}\right) e^{2\pi ia/p} = \left(\frac{-1}{p}\right) G$$

Since $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \bmod p$, we see that $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}$, so

$$\overline{G} = (-1)^{(p-1)/2} G$$

From last week, we know $|G| = \sqrt{p}$, so:

$$p \equiv 1 \pmod 4 \Rightarrow G = \overline{G} \Rightarrow G \in \mathbb{R} \Rightarrow G = \pm\sqrt{p}$$
$$p \equiv 3 \pmod 4 \Rightarrow G = -\overline{G} \Rightarrow G \in i\mathbb{R} \Rightarrow G = \pm i\sqrt{p}$$

Note: $p = |G|^2 = G\overline{G} = (-1)^{(p-1)/2} G^2$, proving (1) above.

Now we simply need $G = +\sqrt{p}$ or $G = +i\sqrt{p}$

The Fourier Transform is a function from $\mathbb{C}[\mathbb{F}_p] \to \mathbb{C}\left[\widehat{\mathbb{F}_p}\right]$, where the map $F \to \hat{F}$ is given by

$$\hat{F}(\chi) = \frac{1}{p} \sum_{x \in \mathbb{F}_p} F(x)\overline{\chi(x)}$$

Our strategy will be to compute the determinant of this transformation in two ways. But what does determinant mean in this context?

Suppose we choose bases for $\mathbb{C}[\mathbb{F}_p]$ and $\mathbb{C}\left[\widehat{\mathbb{F}_p}\right]$. Then consider the natural isomorphisms:

$\mathbb{C}[\mathbb{F}_p] \cong \mathbb{C}^p$ by $F \to (F(0), F(1), \ldots, F(p-1))$, and

$\mathbb{C}\left[\widehat{\mathbb{F}_p}\right] \cong \mathbb{C}^p$ by $\varphi \to (\varphi(\chi_0), \varphi(\chi_1), \ldots, \varphi(\chi_{p-1}))$, where $\chi_a(x) = e^{2\pi i a x/p}$

By composing these two maps and the Fourier transform, we get a map $L : \mathbb{C}^p \to \mathbb{C}^p$, and this map has a determinant.

3.1. **Calculating Determinant, Method 1.** Consider the canonical basis of $\mathbb{C}^p : e_0, e_1, \ldots, e_{p-1}$ where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ has a 1 in the jth place, 0 everywhere else.

Then

$$\hat{e}_j(\chi) = \frac{1}{p} \sum_{x \in \mathbb{F}_p} e_j(x)\overline{\chi(x)} = \frac{1}{p}\overline{\chi(j)}$$

because $e_j(x) = 0$ for $j \neq x$.

Thus

$$\hat{e}_j(\chi_a) = \frac{1}{p}\overline{\chi_a(j)} = \frac{1}{p} e^{2\pi i a j/p}$$

Let $\zeta = e^{2\pi i/p}$. Then

$$\hat{e}_j = \frac{1}{p}\left(\zeta^{0j}, \zeta^{1j}, \zeta^{2j}, \ldots, \zeta^{(p-1)j}\right)$$

so

$$L = \frac{1}{p}\left(\zeta^{ja}\right)_{0 \leq j,a \leq p-1}$$

on the basis $e_0, \ldots, e_{p-1}$.

Then this is a Vandermonde matrix, so

$$\det(L) = \frac{1}{p^p}\det(\zeta^{ja}) = \frac{1}{p^p}\prod_{0\le m<n\le p-1}(\zeta^m - \zeta^n)$$

Let $\delta = e^{\pi i/p}$, so $\delta^2 = \zeta$. Then

$$\det(L) = \frac{1}{p^p}\prod_{0\le m<n\le p-1}\left(\delta^{2m} - \delta^{2n}\right) = \frac{1}{p^p}\prod_{0\le m<n\le p-1}\delta^{m+n}\left(\delta^{m-n} - \delta^{-(m-n)}\right)$$

Now

$$\prod_{0\le m<n\le p-1}\delta^{m+n} = \delta^{\sum(m+n)} = \delta^{\frac{p(p-1)(p-2)}{2}} = (-1)^{\frac{p-1}{2}}$$

because $\delta^p = -1$ and $p-2$ is odd so $(-1)^{p-2} = -1$.
Also,

$$\left(\delta^{m-n} - \delta^{-(m-n)}\right) = 2i\sin\left(\frac{\pi(m-n)}{p}\right)$$

which (because $m < n$) is always a negative number times i. Thus

$$\prod_{0\le m<n\le p-1}\left(\delta^{m-n} - \delta^{-(m-n)}\right)$$

is the product of $\frac{p(p-1)}{2}$ negative numbers multiplied by $i^{\frac{p(p-1)}{2}}$.

Therefore, if $p \equiv 1 \bmod 8$, then: $(-1)^{\frac{p-1}{2}} = 1$ and $\frac{p(p-1)}{2} \equiv 0 \bmod 4$, and thus $\det(L) \in \mathbb{R}^+$.

Similarly: if $p \equiv 5 \bmod 8$, then: $(-1)^{\frac{p-1}{2}} = 1$ and $\frac{p(p-1)}{2} \equiv 2 \bmod 4$, and thus $\det(L) \in \mathbb{R}^-$.

If $p \equiv 3 \bmod 8$, then $(-1)^{\frac{p-1}{2}} = -1$ and $\frac{p(p-1)}{2} \equiv 3 \bmod 4$, and thus $\det(L) \in i\mathbb{R}^-$.

Finally, if $p \equiv 7 \bmod 8$, then $(-1)^{\frac{p-1}{2}} = -1$ and $\frac{p(p-1)}{2} \equiv 1 \bmod 4$, and thus $\det(L) \in i\mathbb{R}^+$.

3.2. **Calculating Determinant, Method 2.** Consider some $\psi \in \widehat{\mathbb{F}_p^*}$, so $\psi : \mathbb{F}_p^* \to \mathbb{C}$. Then extend ψ to $\mathbb{F}_p$ by defining $\psi(0) = 0$. Now $\psi : \mathbb{F}_p \to \mathbb{C}$, so $\psi \in \mathbb{C}[\mathbb{F}_p]$. This gives us an embedding $\widehat{\mathbb{F}_p^*} \subset \mathbb{C}[\mathbb{F}_p] \cong \mathbb{C}^p$.

This means we can think of $\widehat{\mathbb{F}_p^*}$ as a basis for the subspace of $\mathbb{C}^p$ with 0th coordinate 0. Thus $\widehat{\mathbb{F}_p^*} \cup \{e_0\}$ is a basis for $\mathbb{C}^p$.

Then take the Fourier Transform of ψ: (assume $a \neq 0$)

$$\begin{aligned}\hat{\psi}(\chi_a) &= \frac{1}{p}\sum_{x\in\mathbb{F}_p}\psi(x)\chi_a(x)\\ &= \frac{1}{p}\sum_{x\in\mathbb{F}_p}\psi(x)\chi_1(ax)\\ &= \frac{1}{p}\sum_{y\in\mathbb{F}_p}\psi(a^{-1}y)\chi_1(y)\\ &= \frac{1}{p}\psi(a^{-1})G(\psi,\chi_1)\\ &= \frac{1}{p}\overline{\psi(a)}G(\psi,\chi_1)\end{aligned}$$

because under the substitution $y = ax$, y ranges over all of $\mathbb{F}_p$, and we use the definition of G given earlier.

Thus

$$\hat{\psi} = \frac{G(\psi,\chi_1)}{p}\overline{\psi}$$

Consider $\psi \neq \psi_0$:

If $\overline{\psi} \neq \psi$, then by the formula above,

$$\begin{pmatrix} \hat{\psi} \\ \hat{\overline{\psi}} \end{pmatrix} = \begin{pmatrix} 0 & \frac{G(\psi,\chi_1)}{p} \\ \frac{G(\overline{\psi},\chi_1)}{p} & 0 \end{pmatrix}\begin{pmatrix} \psi \\ \overline{\psi} \end{pmatrix}$$

This will give us 2×2 submatrices in L. If $\overline{\psi} = \psi$, then this simplifies to

$$\hat{\psi} = \frac{G(\psi,\chi_1)}{p}\psi$$

giving us 1×1 submatrices instead.

Finally, we must consider ψ_0 and e_0. Now we know that $e_0 = (1,0,\ldots,0)$ and $\psi_0 = (0,1,\ldots,1)$ by extension. Thus $\hat{e}_0 = (1,1,\ldots,1) = e_0 + \psi_0$ and

$$\hat{\psi}_0(\chi_a) = \frac{1}{p}\sum_{x\in\mathbb{F}_p}\psi_0(x)\overline{\chi_a(x)} = \frac{1}{p}\sum_{x\in\mathbb{F}_p^*}\overline{\chi_a(x)} = \begin{cases} \frac{p-1}{p} & a = 0 \\ \frac{-1}{p} & a \neq 0 \end{cases}$$

Thus

$$\hat{\psi}_0 = \left(\frac{p-1}{p}, \frac{-1}{p}, \ldots, \frac{-1}{p}\right) = \frac{p-1}{p}e_0 - \frac{1}{p}\psi_0$$

Thus L looks like

$$\begin{pmatrix} (A) & & & & & & \\ & (B_1) & & & & & \\ & & \ddots & & & & \\ & & & (B_k) & & & \\ & & & & (C_1) & & \\ & & & & & \ddots & \\ & & & & & & (C_m) \end{pmatrix}$$

where A is the 2×2 matrix corresponding to e_0 and ψ_0, $B_1, \ldots, B_k$, are 1x1 matrices corresponding to when $\overline{\psi} = \psi$, and $C_1, \ldots, C_m$, are 2×2 matrices corresponding to when $\overline{\psi} \neq \psi$.

Now we know exactly what A looks like. Consider the B's: Suppose $\overline{\psi} = \psi$. Then ψ takes only real values, so it must take values in ± 1. Thus $\psi^2 = \psi_0$. Since $\mathbb{F}_p^*$ is cyclic, $\hat{\mathbb{F}}_p^*$ is cyclic, so either $\psi = \psi_0$ or ψ is the unique element of order 2, i.e. $\psi = \eta$. Thus this gives us the only B matrix, so L must be of the form:

$$L = \begin{pmatrix} \begin{pmatrix} 1 & 1 \\ \frac{p-1}{p} & \frac{-1}{p} \end{pmatrix} & & & & \\ & \frac{G(\eta,\chi_1)}{p} & & & \\ & & \begin{pmatrix} 0 & \frac{G(\psi,\chi_1)}{p} \\ \frac{G(\overline{\psi},\chi_1)}{p} & 0 \end{pmatrix} & & \\ & & & \ddots & \\ & & & & \begin{pmatrix} 0 & \frac{G(\psi,\chi_1)}{p} \\ \frac{G(\overline{\psi},\chi_1)}{p} & 0 \end{pmatrix} \end{pmatrix}$$

Now we calculate $\det(L)$. The determinant of A is clearly -1, and the determinant of B is itself. Consider the $C_i's$:

Now

$$\overline{G(\psi,\chi_1)} = \overline{\sum \psi(x)\overline{\chi_1(x)}} = \sum \overline{\psi(x)}\chi_1(x) = \sum \overline{\psi(-y)\chi_1(y)} = \psi(-1)G(\overline{\psi},\chi_1)$$

Thus

$$G(\psi,\chi_1)G(\overline{\psi},\chi_1) = \psi(-1)G(\psi,\chi_1)\overline{G(\psi,\chi_1)} = \psi(-1)|G(\psi,\chi_1)|^2 = \psi(-1)p$$

Thus

$$\det(L) = \det(A)\det(B)\prod \det(C_i) = (-1)\frac{G(\eta,\chi_1)}{p} \prod_{\psi,\bar{\psi},\psi\neq\bar{\psi}} \frac{-\psi(-1)p}{p^2}$$

Now L is a $p \times p$ matrix, so the product is over $p-3$ characters (all except e_0, ψ_0, η) and thus over $\frac{p-3}{2}$ pairs of characters.

Since all we care about is the sign, we can write

$$\det(L) = G(\eta,\chi_1)(-1)(-1)^{\frac{p-3}{2}} \prod \psi(-1)(\text{something positive})$$

We need to know whether the product of the $\psi(-1)$ is positive or negative. Let g be a generator of $\mathbb{F}_p^*$. We know all the ψ's can be written as ψ_k, where $\psi_k(g) = e^{2\pi ik/(p-1)}$, for $k = 1, 2, \ldots, p-1$. However we want to exclude ψ_0 and η, because these are not part of the product, and we know η is $\psi_{(p-1)/2}$. Thus we are left with ψ_k for $k = 1, \ldots, (p-3)/2, (p+1)/2, \ldots, p-2$.

Notice that $\psi_k(-1) = \psi_k(g^{(p-1)/2}) = e^{\pi ik}$. Thus this is (-1) when k is odd, and $(+1)$ when k is even.

Now $\overline{\psi_k} = \psi_{-k}$. We are only taking the product of $\psi(-1)$ over pairs $\psi, \overline{\psi}$, and we see that $\psi_k(-1) = \overline{\psi_k}(-1)$ because its real, so it doesn't matter which choice (ψ_k or $\overline{\psi_k}$) we make. Thus we can choose $k = 1, 2, \ldots, (p-3)/2$. Since $\psi_k(-1) = -1$ exactly when k is odd, all we care about is how many odd choices of k we have.

So $p \equiv 1 \pmod 8$: Then $(p-3)/2 \equiv 3 \pmod 4$, giving an even number of odd choices of k.

So $p \equiv 3 \pmod 8$: Then $(p-3)/2 \equiv 0 \pmod 4$, giving an even number of odd choices of k.

So $p \equiv 5 \pmod 8$: Then $(p-3)/2 \equiv 1 \pmod 4$, giving an odd number of odd choices of k.

So $p \equiv 7 \pmod 8$: Then $(p-3)/2 \equiv 2 \pmod 4$, giving an odd number of odd choices of k.

Therefore,

$$\prod \psi(-1) = \begin{cases} 1 & \text{if } p \equiv 1 \text{ or } 3 \pmod 8 \\ -1 & \text{if } p \equiv 5 \text{ or } 7 \pmod 8 \end{cases}$$

Now we combine this together with part 1: Suppose $p \equiv 1 \pmod 8$: Then

$$\begin{aligned} \det(L) &= G(\eta, \chi_1)(-1)^{\frac{p-1}{2}} \prod \psi(-1)(\text{something positive}) \\ &= G(\eta, \chi_1)(+1)(+1)(\text{something positive}) \end{aligned}$$

Since we know $\det(L) \in \mathbb{R}^+$, we see that $G(\eta, \chi_1) \in \mathbb{R}^+$.

Similarly, suppose $p \equiv 3 \pmod 8$: Then

$$\begin{aligned} \det(L) &= G(\eta, \chi_1)(-1)^{\frac{p-1}{2}} \prod \psi(-1)(\text{something positive}) \\ &= G(\eta, \chi_1)(-1)(+1)(\text{something positive}) \end{aligned}$$

Since we know $\det(L) \in \mathbb{C}^-$, we see that $G(\eta, \chi_1) \in \mathbb{R}^+$.

Suppose $p \equiv 5 \pmod 8$: Then

$$\begin{aligned} \det(L) &= G(\eta, \chi_1)(-1)^{\frac{p-1}{2}} \prod \psi(-1)(\text{something positive}) \\ &= G(\eta, \chi_1)(+1)(-1)(\text{something positive}) \end{aligned}$$

Since we know $\det(L) \in \mathbb{R}^-$, we see that $G(\eta, \chi_1) \in \mathbb{R}^+$.

Suppose $p \equiv 7 \pmod 8$: Then

$$\begin{aligned} \det(L) &= G(\eta, \chi_1)(-1)^{\frac{p-1}{2}} \prod \psi(-1)(\text{something positive}) \\ &= G(\eta, \chi_1)(-1)(-1)(\text{something positive}) \end{aligned}$$

Since we know $\det(L) \in i\mathbb{R}^+$, we see that $G(\eta, \chi_1) \in i\mathbb{R}^+$.

Since we already know that $|G|^2 = p$, we have shown that

$$G = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \bmod 4 \\ i\sqrt{p} & \text{if } p \equiv 3 \bmod 4 \end{cases}$$

as desired.

We start by comparing the rings $\mathbb{Z}$ and $\mathbb{F}_q[x]$. They are very similar - both are Euclidean, their quotient fields are finite. Consider the parallels:

$$\begin{array}{ccc} \mathbb{Z} & & \mathbb{F}_q[x] \\ \text{primes} & \longleftrightarrow & \text{monic irred. polys} \\ \mathbb{Z}^+ = \mathbb{N} & \longleftrightarrow & \text{monic polys} \\ \mathbb{Q} & \longleftrightarrow & \mathbb{F}_q(x) \\ \sum_{n=1}^{\infty} \frac{a_n}{n^s} & \longleftrightarrow & \sum_{h \text{ monic}} \frac{a_n}{(Nh)^s} \end{array}$$

The last line gives the Dirichlet series analog in $\mathbb{F}_q[x]$. Our analog definition comes from noting that $n = \#\mathbb{Z}/n\mathbb{Z}$, so we define $Nh := \#\mathbb{F}_q[x]/(h)$.

But this definition really gives $Nh = q^{\deg h}$, so the Dirichlet sum becomes

$$\sum_{h \text{ monic}} a_h (q^{-s})^{\deg h} = \sum_{h \text{ monic}} a_h T^{\deg h}$$

by letting $T = q^{-s}$.

We're not going to worry too much about the convergence of these series.

Definition 3.1. A Dirichlet character is a character $\chi : \mathbb{Z} \to \mathbb{C}$ that is:
1) Periodic: $\exists q$ s.t. $\chi(n+q) = \chi(n) \forall n$
2) Multiplicative: $\chi(mn) = \chi(m)\chi(n)$

3.3. **L-functions.** Let $M = \{\text{monic polynomials in } \mathbb{F}_q[x]\}$, and let $\lambda : M \to \mathbb{C}$ such that λ is multiplicative. Assume $\lambda(1) = 1$.

Definition 3.2. $L(\lambda, T) := \sum_{h \in M} \lambda(h) T^{\deg h}$, and consider this as an element of $\mathbb{C}[[T]]$, i.e. all formal power series in T. Thus we don't need to worry about convergence.

Proposition 3.3.

$$L(\lambda, T) = \prod_{\substack{p \in M \\ p \text{ irred}}} (1 - \lambda(p) T^{\deg p})^{-1}$$

Proof. $1 - \lambda(p) T^{\deg p} \in \mathbb{C}[[T]]$. Also we note that it's a unit:

$$(1 - \lambda(p) T^{\deg p})^{-1} = \sum_{n=0}^{\infty} (\lambda(p) T^{\deg p})^n = \sum_{n=0}^{\infty} \lambda(p^n) T^{n \deg p}$$

using the geometric series expansion. So we take this over all polynomials:

$$\prod_{\substack{p \in M \\ p \text{ irred}}} (1 - \lambda(p) T^{\deg p})^{-1} = \prod_{\substack{p \in M \\ p \text{ irred}}} \sum_{n=0}^{\infty} \lambda(p^n) T^{n \deg p}$$

To handle the fact that this is an infinite product, we take this over polynomials with degree <= D, and let $D \to \infty$. Formally,

$$\text{ord}_{T=0}\left(\prod_{\substack{p\in M\\ p\ irred\\ \deg p<=D}} (1-\lambda(p)T^{\deg p})^{-1} - L(\lambda,T)\right) \to \infty \text{ as } D \to \infty$$

□

Now, consider the expansion of

$$\prod_{\substack{p\in M\\ p\ irred}} \sum_{n=0}^{\infty} \lambda(p^n)T^{n\deg p}$$

Clearly the constant term (degree 0) is 1. The term of degree 1 is $\sum_{\deg p=1}\lambda(p)T^1$, which is equal to $\sum_{\alpha\in\mathbb{F}_q}\lambda(x-\alpha)T$. Similarly, the term of degree 2 is of the form $(\sum\lambda(x-\alpha)\lambda(x-\beta)+\sum\lambda(x^2-ax-b))T^2$. We notice that to get terms with small degrees in the expansion, we only need to look at polynomials p with small degree.

This leads to the conclusion that the product is equal to

$$\sum_{p_1,p_2,\ldots}\lambda(p_1^{n_1}p_2^{n_2}\ldots p_r^{n_r})T^{n_1\deg p_1+\ldots+n_r\deg p_r} = \sum_{h\in M}\lambda(h)T^{\deg h}$$

and this includes every term exactly once because we have unique factorization in $\mathbb{F}_q[x]$.

Now suppose $\lambda(h)=1$ for all h. Then

$$\sum_{h\in M}\lambda(h)T^{\deg h} = \sum_{h\in M}T^{\deg h} = \sum_{d=0}^{\infty}\left(\sum_{\substack{h\in M\\ deg\ h=d}}1\right)T^d = \sum_{d=0}^{\infty}q^dT^d = \frac{1}{1-qT}$$

We can also evaluate this sum using the equality from the last section:

$$\sum_{h\in M}\lambda(h)T^{\deg h} = \sum_{h\in M}T^{\deg h} = \prod_{\substack{p\in M\\ p\ irred}}(1-T^{\deg p})^{-1} = \prod_{d=1}^{\infty}(1-T^d)^{-N_d}$$

where N_d is the number of monic irreducible polynomials of degree d. Combining these results, we get

$$\prod_{d=1}^{\infty}(1-T^d)^{-N_d} = \frac{1}{1-qT}$$

By taking the logarithmic derivatives of both sides and simplifying, we get $\sum_{d|n}dN_d = q^n$. We can then use the Mobius inversion formula to show that $N_n = \frac{1}{n}\sum_{d|n}\mu(d)q^{n/d} \sim \frac{q^n}{n}$. This shows that the number of prime polynomials of degree n is like $\frac{q^n}{n}$, which is the prime number theorem for polynomials.

Let $f(x)\in\mathbb{F}_p[x]$ for prime p. If h is irreducible and monic, and $\alpha\in\overline{\mathbb{F}_p}$ is a root of h, then define

$$\lambda_f(h) := e^{2\pi i Tr_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p}(f(\alpha))/p}$$

and $\lambda_f(1) := 1$.

Note: This is well-defined. Since the trace is the sum of the conjugates, this is constant over all roots, and thus is independent of the choice of α.

For general $h = \prod h_i^{n_i}$, with h_i irreducible, define $\lambda_f(h) := \prod \lambda_f(h_i)^{n_i}$. This is the unique extension to $\lambda_f : M \to \mathbb{C}$ that is multiplicative.

Our next goal is to prove the following:

Theorem 3.4. *If $p \nmid f(x)$, then $L(\lambda, T)$ is a polynomial of degree at most* $\deg f(x)$.

From before,

$$L(\lambda_f, T) = 1 + \sum_{\alpha \in \mathbb{F}_p} \lambda_f(x - \alpha)T + \dots$$

From the above theorem, we know that $L(\lambda_f, T) = \prod(1 - w_i T)$. On the other hand, the coefficient in T of the above expansion is $\sum_{\alpha \in \mathbb{F}_p} e^{2\pi i f(\alpha)/p}$ by the definition of λ_f. Thus $\sum w_i = \sum_{\alpha \in \mathbb{F}_p} e^{2\pi i f(\alpha)/p}$.

Previously, we discussed the following: There exists $w_1, \dots, w_n \in \mathbb{C}$ such that for all $m \geq 1$

$$\sum_{\alpha \in \mathbb{F}_{p^m}} e^{2\pi i \operatorname{Tr}_{\mathbb{F}_{p^m}/\mathbb{F}_p}(f(\alpha))/p} = -(w_1^m + \dots w_n^m)$$

and $|w_i| = p^{1/2}$

This statement is essentially equivalent to the last theorem.

4. Week Four

Let $f(x) \in \mathbb{F}_p[x]$ of degree d , $p \nmid d$. Given $h(x) \in \mathbb{F}_p[x]$ a monic, irreducible polynomial, define $\lambda(h) := e^{2\pi i \operatorname{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p}(f(\alpha))/p}$ for any root $\alpha \in \mathbb{F}_p$ of $h(x)$, and $\lambda(a) := 1$, for $a \in \mathbb{F}_p^*$.

If $h(x) \in \mathbb{F}_p[x]$, and $h(x) \neq 0$. $h(x) = a \prod_i h_i$ where h_i are monic and irreducible.

Let $\lambda(h) = \prod_i \lambda(h_i)$ then $\lambda(h_1 h_2) = \lambda(h_1)\lambda(h_2)$ is true for all h_1, $h_2 \in \mathbb{F}_p[x]$. Let $M = \{h(x) \in \mathbb{F}_q[x], \text{ monic, and not } 0\}$.

Finally, let $L(\lambda, T) = \sum_{h \in M} \lambda(h) T^{\deg h} \in \mathbb{C}[[T]]$, $L(\lambda, 0) = 1$.

Theorem 15. $L(\lambda, T) = \prod_{h \in M,\, h \text{ irr}} (1 - \lambda(h) T^{\deg h})^{-1}$.

Theorem 16. $L(\lambda, T)$ is a polynomial in T of degree at most d.

Corollary 17. There exists $w_1, \dots, w_d \in \mathbb{C}^*$ such that for all $m \geqslant 1$ we have $\sum_{\alpha \in \mathbb{F}_{p^m}} e^{2\pi i \operatorname{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p}(f(\alpha))/p} = -(w_1^m + \dots + w_d^m)$.

Proof. (assuming Theorems 1 and 2). There exists $w_1, \dots, w_d \in \mathbb{C}^*$ such that $L(\lambda, T) = \prod_{i=1}^d (1 - w_i T)$. Then,

$$\frac{L'(\lambda, T)}{L(\lambda, T)} = \sum_{i=1}^{d} \frac{-w_i}{1 - w_i T} = \sum_{m=0}^{\infty} (-\sum_{i=1}^{d} w_i^{m+1}) T^m \quad (*)$$

where $' = \frac{d}{dT}$. Now by Theorem 1,

$$\frac{L'(\lambda,T)}{L(\lambda,T)} = \sum_{h\in M,\ h\ \mathtt{irr}} \frac{\lambda(h)\deg h T^{\deg h-1}}{1-\lambda(h)T^{\deg h}}$$

$$= \sum_{h\in M,\ h\ \mathtt{irr}} \sum_{m=0}^{\infty} \lambda(h)\deg h T^{\deg h-1}\lambda(h)^m T^{m\deg h}$$

$$= \sum_{h\in M,\ h\ \mathtt{irr}} \sum_{m=0}^{\infty} \lambda(h)^{m+1}\deg h T^{(m+1)\deg h-1}$$

$$= \sum_{n=0}^{\infty}\Big(\sum_{h\in M,\ h\ \mathtt{irr}} \lambda(h)^{\frac{n+1}{\deg h}}\deg h\Big)T^n$$

Then,

$$\frac{L'(\lambda,T)}{L(\lambda,T)} = \sum_{n=0}^{\infty}\Big(\sum_{\substack{h\in M,\ h\ \mathtt{irr}\\ \deg h\mid(n+1)}} \lambda(h)^{\frac{n+1}{\deg h}}\deg h\Big)T^n \quad (\bigstar)$$

For $m \geqslant 1$, if $\alpha \in \mathbb{F}_{p^m}$, then α has a minimal polynomial over $\mathbb{F}_p$, $h(x) \in \mathbb{F}_{p^m}[x]$, monic, irreducible and of degree $\delta \mid m$; such that every root of $h(x)$ is in $\mathbb{F}_{p^m}$.

Conversely, any monic, irreducible polynomial $h(x) \in \mathbb{F}_p[x]$ of degree $\delta \mid m$ has roots in $\mathbb{F}_{p^m}$.

Now,

$$(\bigstar) = \sum_{\delta\mid(n+1)}\ \sum_{\substack{h\in M,\ h\ \mathtt{irr}\\ \deg h=\delta}}\ \sum_{\alpha,\ h(\alpha)=0} e^{2\pi i(\mathrm{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p} f(\alpha))\frac{n+1}{\delta}/p}$$

$$= \sum_{\alpha\in\mathbb{F}_{p^{n+1}}} e^{2\pi i(\mathrm{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p} f(\alpha))\frac{n+1}{\delta}/p}$$

Now, $\mathrm{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p} f(\alpha))\frac{n+1}{\delta} = \mathrm{Tr}_{\mathbb{F}_{p^m}/\mathbb{F}_p} f(\alpha)$ by the transitivity of the trace. Indeed, if $n+1=m$,

$$\mathrm{Tr}_{\mathbb{F}_{p^m}/\mathbb{F}_p} f(\alpha) = \mathrm{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p}(\mathrm{Tr}_{\mathbb{F}_{p^m}/\mathbb{F}_{p(\alpha)}} f(\alpha))$$
$$= \mathrm{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p}(\frac{m}{\delta}f(\alpha)) = \frac{m}{\delta}\,\mathrm{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p}(f(\alpha))$$

□

Definition 9. Let $h(x) \in \mathbb{F}_p[x]$ be a monic, irreducible polynomial, then $\partial h = \mathrm{Tr}_{\mathbb{F}_{p(\alpha)}/\mathbb{F}_p}(f(\alpha))$, and $\lambda(h) = e^{2\pi i\partial h/p}$.

Now, $h(x)$ can be written as $h(x) = \prod_{i=1}^{m}(x-\alpha_i)$, $\alpha_i \in \mathbb{F}_{p^m}$, then

$$\partial h = \sum_{i=1}^{m} f(\alpha_i) = \sum_{i=1}^{m} f(\alpha_i)\mathtt{res}_{x=\alpha_i}(\frac{1}{x-a_i} + \sum_{j\neq i}\frac{1}{x-\alpha_i})$$

$$= \sum_{i=1}^{m} f(\alpha_i)\mathtt{res}_{x=x_i}\frac{h'(x)}{h(x)}$$

Also,

$$\partial h = \sum_{i=1}^{m} \mathtt{res}_{x=\alpha_i} f(x)\frac{h'(x)}{h(x)} = -\mathtt{res}_{\infty} f(x)\frac{h'(x)}{h(x)}$$

Let K be an algebraically closed field, and $K(x)$ be a field of rational functions with coeficients in K. Let $R(x) \in K(x)$ and $\mathbb{P}^1(K) = K \cup \{\infty\}$.

Definition 10. The *residue* of $R(x)$ at $\alpha \in \mathbb{P}^1(K)$, $\mathtt{res}_\alpha R(x)dx$, is defined in the following way:

- For $\alpha \in K$, write $R(x) = \sum_{n \geqslant n_0} a_n(x-\alpha)^n$, then $\mathtt{res}_\alpha R(x)dx = a_{-1}$
- For $\alpha = \infty$ define $\mathtt{res}_\infty R(x)dx = \mathtt{res}_0 - \frac{1}{x^2}R(\frac{1}{x})dx$

Now, let us suppose that $R(x) = \frac{A(x)}{B(x)}$, where $A(x), B(x) \in K[x]$, and $(x - \alpha) \mid B(x)$. Then, $B(x) = (x-\alpha)^k C(x)$, where $C(\alpha) \neq 0$, and $C(x) \in K[x]$ similarly, $A(x) = U(x)(x-\alpha)^k + V(x)$, where $U(x), V(x) \in K[x]$. So,

$$R(x) = (U(x) + \frac{V(x)}{(x-\alpha)^k})\frac{1}{C(x)}$$

Or equivalently,

$$R(x) = (U(x) + \frac{v_0 + v_1(x-\alpha) + \ldots + v_m(x-\alpha)^m)}{(x-\alpha)^k})\frac{1}{C(x)}$$

Then,

$$\mathtt{res}_\alpha R(x)dx = \frac{v_{k-1}}{C(\alpha)}$$

Theorem 18. $\mathtt{res}_\alpha R(x)dx \neq 0$ for only a finite number of $\alpha \in \mathbb{P}^1(K)$ and $\sum_{\alpha \in \mathbb{P}^1(K)} \mathtt{res}_\alpha R(x)dx = 0$.

Denote by $\rho(R(x)dx) := \sum_{\alpha \in \mathbb{P}^1(K)} \mathtt{res}_\alpha R(x)dx$.

Lemma 19. $\rho((R(x) + S(x))dx) = \rho(R(x)dx) + \rho(S(x)dx)$. In fact, $\forall\ \alpha \in \mathbb{P}^1(K)$ $\mathtt{res}_\alpha((R(x) + S(x))dx) = \mathtt{res}_\alpha(R(x)dx) + \mathtt{res}_\alpha(S(x)dx)$.

Proof. Let $R(x) = \frac{A_1(x)}{B(x)}$ and $S(x) = \frac{A_2(x)}{B(x)}$, with $B(x) = (x-\alpha)^k C(x)$, where $C(\alpha) \neq 0$, and $A_i(x) = U_i(x)(x-\alpha)^k + V_i(x)$.
Then,

$$R(x) = \frac{1}{C(x)}(U_1(x) + \frac{V_1(x)}{(x-\alpha)^k})$$
$$S(x) = \frac{1}{C(x)}(U_2(x) + \frac{V_2(x)}{(x-\alpha)^k})$$

So,

$$R(x) + S(x) = \frac{(A_1(x) + A_2(x)}{B(x)}$$
$$= \frac{1}{C(x)}(U_1(x) + U_2(x) + \frac{V_1(x) + V_2(x)}{(x-\alpha)^k})$$

Then, the lemma follows by definition of residues. □

Lemma 20. *Partial Fractions Decomposition.* Let $R(x) = \frac{A(x)}{\prod_{i=1}^{m}(x-\alpha_i)^{n_i}} \in K(x)$, $A(x) \in K[x]$ and $\alpha_1, \ldots, \alpha_m$ distinct, then there exist $A_1(x), \ldots, A_m(x) \in K[x]$ such that $R(x) = \sum_{i=1}^{m} \frac{A_i(x)}{(x-\alpha_i)^{n_i}}$.

Proof. By induction on m. The case $m = 1$ is obvious. Suppose true for $m < k$. Now, let us prove it for $m = k$. We have that, $\gcd(\prod_{i=1}^{k-1}(x-\alpha_i)^{n_i}, (x-\alpha_k)^{n_k}) = 1$, then there exists $U(x), V(x) \in K(x)$ such that

$$U(x)\prod_{i=1}^{k-1}(x-\alpha_i)^{n_i} + V(x)(x-\alpha_k)^{n_k} = 1.$$

So,

$$R(x) = \frac{U(x)A(x)}{(x-\alpha_k)^{n_k}} + \frac{V(x)A(x)}{\prod_{i=1}^{k-1}(x-\alpha_i)^{n_i}}$$

Then, let $A_m(x) = U(x)A(x)$, and by induction we have that $\frac{V(x)A(x)}{\prod_{i=1}^{k-1}(x-\alpha_i)^{n_i}} = \sum_{i=1}^{m-1} \frac{A_i(x)}{(x-\alpha_i)^{n_i}}$ □

Proof. (of Theorem 0.4) it is enough by lemmas 1 and 2, to prove that $\rho(\frac{A(x)}{(x-\alpha)^n}dx) = 0$. Now $A(x) = a_0 + a_1(x-\alpha) + \ldots + a_d(x-\alpha)^d$, so

$$\mathtt{res}_\alpha R(x)dx = a_{n-1}$$

and,

$$\mathtt{res}_\infty R(x)dx = \mathtt{res}_0 \frac{-1}{x^2} R(1/x)dx$$

$$= \mathtt{res}_0 \frac{\frac{-1}{x^2}(a_0 + a_1(\frac{1}{x}-\alpha) + \ldots + a_d(\frac{1}{x}-\alpha)^d}{(\frac{1}{x}-\alpha)^n}$$

$$= \sum \mathtt{res}_0 \frac{-a_i(1-x\alpha)}{x^{2+i-m}(1-x\alpha)^n}dx = -a_{n-1}$$

□

For all $h(x) \in \mathbb{F}_p$ monic and irreducible.We have $\partial(h) = -\mathtt{res}_\infty f(x)\frac{h'(x)}{h(x)}dx$. Also, $\lambda(h) = e^{2\pi i \partial(h)/p}$. Then $\lambda(h_1h_2) = \lambda(h_1)\lambda(h_2)$, and therefore $\partial(h_1h_2) = \partial(h_1)\partial(h_2)$. So,

$$\frac{(h_1h_2)'}{h_1h_2} = \frac{h_1'}{h_1} + \frac{h_2'}{h_2}$$

Then,

$$= \sum_{m=0}^{\infty}\Big(\sum_{\substack{h\in M \\ \deg h = m}} \lambda(h)T^m\Big)$$

We want to prove that $L(\lambda, T) = 0$, for $m > d = \deg f$.

Definition 11. $h_1 \sim h_2$ if $\frac{h_1}{h_2} - 1$ has a zero of multiplicity $\geqslant d+1$ at ∞ (i.e. degree of denominator − degree of numerator $\geqslant d+1$).

Lemma 21. If $h_1 \sim h_2$, then $\partial(h_1) = \partial(h_2)$.

Proof.

$$\partial(h_1) - \partial(h_2) = \mathtt{res}_\infty(f\frac{h_2'}{h_2} - f\frac{h_1'}{h_1})dx$$
$$= \mathtt{res}_\infty(f\frac{(\frac{h_2}{h_1})'}{(\frac{h_2}{h_1})})dx = 0$$

because f has a pole of order d at ∞. $(\frac{h_2}{h_1})'$ has a zero of order $\geqslant d$ at ∞ and $(\frac{h_2}{h_1}) = 1$ at ∞.

□

5. Week Five

Recall that $f(x) \in \mathbb{F}_p[x]$ with $\deg(f) = d$ where $\gcd(d, p) = 1$, $\partial(h) = -\mathtt{res}_\infty(f\frac{dh}{h})$, $\lambda(h) = e^{2\pi i\partial(h)/p}$ for $h \in M$, where $M = \{h(x) \in \mathbb{F}_p[x], \text{ monic }\}$. Let also, for $m \geq 1$, $M_m = \{h \in M \mid \deg h = m\}$.

We have seen that if $\sum_{h\in M_m} \lambda(h) = 0$ for $m \geq d+1$, $L(\lambda, T)$ is a polynomial and from this it follows that

$$\sum_{\alpha\in\mathbb{F}_{p^m}} e^{2\pi i Tr_{\mathbb{F}_{p^m}/\mathbb{F}_p}(f(\alpha))/p} = -(\omega_1^m + \cdots + \omega_d^m)$$

for some $\omega_1, \ldots, \omega_d \in \mathbb{C}^*$.

Definition 12. Let $u = 1/x$. Then $\mathrm{ord}_\infty(h-1) = m$ if $h(u) = 1 + a_m u^m + a_{m+1}u^{m+1} + \cdots, a_m \neq 0$ as a power series in u.

Lemma 22. If $\mathrm{ord}_\infty(h-1) \geq d+1$, then $\mathtt{res}_\infty(fh'/hdx) = 0$.

Proof: One can check that $\frac{1}{h}\frac{dh}{dx}dx = \frac{1}{h}\frac{dh}{du}du$ and that $\frac{dh}{du}$ has a zero of order at least d at ∞. Hence, at ∞, f has a pole of order d, $\frac{1}{h}$ is regular, and $\frac{dh}{du}$ has a zero of order at least d. Then $\mathrm{ord}_\infty(f\frac{1}{h}\frac{dh}{du}) \geq 0$ so $\mathtt{res}_\infty(f\frac{1}{h}\frac{dh}{du}du) = \mathtt{res}_\infty(f\frac{h'}{h}dx) = 0$.

Note that $\{h \in \mathbb{F}_p(x)^* | \,\mathrm{ord}_\infty(h-1) > 0\}$ is a group under multiplication and that $\Gamma_n = \{h \in \mathbb{F}_p(x)^* | \,\mathrm{ord}_\infty(h-1) \geq n\}$ is a subgroup. We define $G = \Gamma_1/\Gamma_d$.

Lemma 23. G is a finite group of order p^d under multiplication. Further, λ induces a non-trivial character on G.

Proof:

Sublemma 1. $\Gamma_n/\Gamma_{n+1} \cong \mathbb{F}_p$.

Proof: For $h \in \Gamma_n$, $h = 1 + \alpha u^n + \cdots$. Define $\varphi_n : \Gamma_n \to \mathbb{F}_p$ by $h \mapsto \alpha$. Then

$$\varphi_n[(1 + \alpha u^n + \cdots)(1 + \beta u^n + \cdots)] = \varphi_n[1 + (\alpha + \beta)u^n + \cdots] = \alpha + \beta + \cdots$$

so that φ_n is a homomorphism. φ_n is clearly surjective with $\ker \varphi_n = \Gamma_{n+1}$. Hence, $\Gamma_n/\Gamma_{n+1} \cong \mathbb{F}_p$.

Then by the sublemma, G is a finite group of order p^d. By a previous lemma, $\lambda : \Gamma_1 \to \mathbb{C}^*$ is a homomorphism with $\lambda|_{\Gamma_{d+1}} \equiv 1$.

We want to show that λ is non-trivial on G.
If $h = \frac{1+x^d}{x^d} = 1 + u^d$, $h \in \Gamma_d \backslash \Gamma_{d+1}$, we have

$$f\frac{dh}{h} = \frac{a_0u^d + \cdots + a_d}{u^d}\frac{du^{d-1}}{1+u^d} =$$

$$d\frac{a_0u^d + \cdots + a_d}{u}\frac{1}{1+u^d} = d\frac{a_0u^d + \cdots + a_d}{u}(1 - u^d + u^{2d} - \cdots) = \frac{da_d}{u} + \cdots .$$

Then $-\mathtt{res}_\infty(f\frac{dh}{h}) = \partial(h) = -da_d \neq 0 \pmod p$. Therefore da_d/p is not in $\mathbb{Z}$ so $\lambda(h) \neq 1$.

Since λ is non-trivial on G, $\sum_{h\in G} \lambda(h) = 0$. Now define $\psi_m : M_m \to G$ by $\psi_m(h) = \frac{h}{x^m}$ for $m \geq d+1$.

Lemma 24. For $g \in G$, $\#\psi_m^{-1}(g) = p^{m-d}$.

Assuming the lemma,

$$\sum_{h\in M_m} \lambda(h) = \sum_{h\in M_m} \lambda(\frac{h}{x^m}\cdot x^m) = \lambda(x^m)\sum_{g\in G}\sum_{\psi_m(h)=g, h\in M_m} \lambda(g) = \lambda(x^m)p^{m-d}\sum_{g\in G}\lambda(g) = 0.$$

Proof of Lemma: If $h(x) = x^m + \cdots + b_0$, $\frac{h(x)}{x^m} = 1 + \cdots + b_0u^m \in \Gamma_1$. Then $\psi_m(h)$ is the image of $1+\cdots+b_0u^m$ in $G = \Gamma_1/\Gamma_{d+1}$. In $G, \psi_m(h) = 1+\cdots+b_{m-d}u^d$. Therefore $\psi_m(h) = \psi_m(g)$ if and only if $h(x) = x^m + b_{m-1}x^{m-1} + \cdots + b_0, g(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_0$ with $b_{m-1} = a_{m-1}, \cdots b_{m-d} = a_{m-d}$. It follows that $\#\psi_m(g) = p^{m-d}$.

6. Week Six

We have $L(T, \lambda)$ is a polynomial.
We have proven:

Theorem 25. Given $f(x) \in \mathbb{F}_p[x]$, of degree d, $1 \leq d < p$. Then there exists $\omega_1, \ldots, \omega_d \in \mathbf{C}$ such that

$$S_m = \sum_{x\in\mathbb{F}_{p^m}} e^{2\pi i Tr_{\mathbb{F}_{p^m}/\mathbb{F}_p}(f(x))/p} = -(\omega_1^m + \cdots + \omega_d^m).$$

Now we can say

Lemma 26. a) If $|\omega_i| \leq B, i = 1, \ldots, d$, then $|S_m| \leq dB^m \ \forall m \geq 1$
b) If $\exists c > 0$ such that $|S_m| \leq cB^m \ \forall m \geq 1$, then $|\omega_i| \leq B \ \forall i$.

Proof. a). Obvious.
b). Consider $\varphi(z) = \sum_{m=1}^{\infty} -S_m z^m$.
By the comparison test, using the hypothesis, then $\varphi(z)$ converges for $|z| < \frac{1}{B}$.
So it is analytic on $|z| < \frac{1}{B}$ (has no poles).
But $\varphi(z) = \sum_{m=1}^{\infty} \sum_{j=1}^{d} \omega_j^m z^m = \sum_{j=1}^{d} \frac{\omega_j z}{1-\omega_j z}$. (geometric series)
This has poles at $z = \frac{1}{\omega_j}, j = 1, \ldots, d$. These poles must be outside the disk of radius $\frac{1}{B}$, since we said no poles in $|z| < \frac{1}{B}$.
Therefore $|\frac{1}{\omega_j}| \geq \frac{1}{B}$.
$\Rightarrow |\omega_j| \leq B$. □

Note: In fact, later we will prove that $|\omega_j| = q^{\frac{1}{2}}$. This corresponds under $T \mapsto q^{-s}$ to $\operatorname{Re} s = \frac{1}{2}$. This is a theorem of Weil; corresponds to the Riemann hypothesis of number fields. (So it's been proven in function fields, but not number fields.)

Theorem 27 (Hardy-Littlewood-Weyl)**.** Let $1 \leq d < p$. Then

$$|S_m| \leq 2q^{1-\frac{1}{2^{d-1}}}.$$

Note: Here $B = q^{1-\frac{1}{2^{d-1}}}$.
This is an improvement over $|S_m| < q$, since this is $< q$ for q's big compared to d.

Proof. By induction on d.
When $d = 1$, we have $|S_m| \leq 2$. In fact, when $d = 1$, $S_m = 0 \ \forall m \geq 1$. This is because $\sum e^{2\pi i Tr(a+bx)/p} = e^{2\pi i Tr(a)/p} \sum e^{2\pi i Tr(bx)/p} = 0$.
Note: When $d = 2$ we get the Quadratic Gauss Sum.
Assume true for $\deg f < d$. Then

$$\begin{aligned}
|S_m|^2 &= S_m \overline{S_m} \\
&= \sum_{x \in \mathbb{F}_{p^m}} e^{2\pi i Tr(f(x))/p} \sum_{y \in \mathbb{F}_{p^m}} e^{2\pi i Tr(-f(y))/p} \\
&= \sum_{x,y \in \mathbb{F}_{p^m}} e^{2\pi i Tr(f(x)-f(y))/p} \quad \textit{Let } y = x + z. \\
&= \sum_{x,z \in \mathbb{F}_{p^m}} e^{2\pi i Tr(f(x)-f(x+z))/p} \\
&= \sum_{z \in \mathbb{F}_{p^m}} \sum_{x \in \mathbb{F}_{p^m}} e^{2\pi i Tr(f(x)-f(x+z))/p}
\end{aligned}$$

Call this second sum T_z.
Now for fixed $z \in \mathbb{F}_{p^m}$, $f(x) - f(x+z)$ is a polynomial in x of degree strictly smaller than d.
Assume $f(x) - f(x+z)$ is nonconstant for $z \neq 0$. (See below for justification.)
Then $|T_z| \leq 2q^{1-\frac{1}{2^{d-2}}}$ by the induction hypothesis. (Note: if $z = 0$, then $T_0 = q$.)

So

$$\begin{aligned}|S_m|^2 &\leq q + (q-1)2q^{1-\frac{1}{2^{d-2}}} = 2(q-1)q^{1-\frac{1}{2^{d-2}}} + q \\ &\leq 2qq^{1-\frac{1}{2^{d-2}}} + 2q^{1-\frac{1}{2^{d-2}}}q \\ &= 4q^{2-\frac{1}{2^{d-2}}} \\ \Rightarrow |S_m| &\leq 2q^{1-\frac{1}{2^{d-1}}}.\end{aligned}$$

Now, can $f(x) - f(x+z)$ be constant for $z \neq 0$?
Let $f(x) = x^d + a_1x^{d-1} + \cdots + a_d$.
Then $f(x) - f(x+z) = x^d - (x+z)^d + a_1(x^{d-1} - (x+z)^{d-1}) + \cdots = -dzx^{d-1} + \cdots$
The leading coefficient of $f(x) - f(x+z)$ as a polynomial in x is $-dz$.
This is nonzero if $1 \leq d < p$, $z \neq 0$, which is the hypothesis in the theorem.
So our assumption holds, and we're done. □

(Note: The estimate $|S_m| \leq 2q^{1-\frac{1}{2^{d-1}}}$ is nontrivial for q big compared to d.
If $\frac{2}{q^{\frac{1}{2^{d-1}}}} < 1$, then $q > 2^{2^{d-1}}$, so we must get q really big compared to d for this estimate to make a difference.)

Now let's modify our notation. Let $S = \sum_{x \in \mathbb{F}_q} e^{2\pi i Tr(f(x))/p}$.

Theorem 28 (Mordell)**.** Let deg $f = d < p$. Then

$$|S| \leq 2(dd!)^{\frac{1}{2d}} q^{1-\frac{1}{2d}}$$

(Note: The goal is to get $|S| \leq q^{1/2}$, so we're getting closer...)

Proof.

$$\begin{aligned}S &= \sum e^{2\pi i Tr(f(x))/p} \\ &= \sum e^{2\pi i Tr(f(x)-f(0)+f(0))/p} \\ &= e^{2\pi i Tr(f(0))/p} \sum e^{2\pi i Tr(f(x)-f(0))/p}\end{aligned}$$

So when we take absolute values, then since $|e^{2\pi i Tr(f(0))/p}| = 1$, we can say

$$|\sum e^{2\pi i Tr(f(x))/p}| = |\sum e^{2\pi i Tr(f(x)-f(0))/p}|,$$

So without loss of generality we will assume $f(0) = 0$ in the rest of the proof.

For all $\lambda \in \mathbb{F}_q, \lambda \neq 0$, $x \mapsto \lambda x$ is a bijection on $\mathbb{F}_q$. So

$$|S_m| = |\sum_{x \in \mathbb{F}_q} e^{2\pi i Tr(f(\lambda x))/p}|$$

$$\Rightarrow |S|^{2d} = |\sum_{x\in\mathbb{F}_q} e^{2\pi i Tr(f(\lambda x))/p}|^{2d}.$$

Let Λ be a subset of $\mathbb{F}_q$ of cardinality at least $(q-1)/d$ such that the $\lambda^d, \lambda \in \Lambda$ are distinct. (See Lemma 2 below for existence of such a set.)

Then

$$\frac{q-1}{d}|S|^{2d} \leq \sum_{\lambda\in\Lambda} |S|^{2d} = \sum_{\lambda\in\Lambda} |\sum_{x\in\mathbb{F}_q} e^{2\pi i Tr(f(\lambda x))/p}|^{2d}.$$

(Note: We are just summing over polynomials of degree d and with f(0)=0; and all of the polynomials are different because of our choice of Λ.)

$$\leq \sum_{c_1,\ldots,c_d\in\mathbb{F}_q} |\sum_{x\in\mathbb{F}_q} e^{2\pi i Tr(\sum_{j=1}^d c_j x^j)/p}|^{2d}$$

(This is sum over all polynomials of degree d.)

(Remark: In examining the usefulness of this technique, we must keep in mind that $\lim_{p\to\infty}(\sum_{i=1}^m |x_i|^p)^{1/p} = \max_{1\leq i\leq m} |x_i|$ so in our sum, the biggest term will stand out.)

Now call $S_c = \sum e^{2\pi i Tr(\sum c_j x^j)/p}$, where $c = (c_1, \ldots, c_d)$

Then

$$\begin{aligned}
\sum_c |S_c|^{2d} &= \sum_c S_c^d \overline{S_c}^d \\
&= \sum_c (\sum_{x\in\mathbb{F}_q} e^{2\pi i Tr(\sum_j c_j x^j)/p})^d (\sum_{y\in\mathbb{F}_q} e^{2\pi i Tr(\sum_j -c_j y^j)/p})^d \\
&= \sum_c \sum_{\substack{x_1,\ldots,x_d\in\mathbb{F}_q \\ y_1,\ldots,y_d\in\mathbb{F}_q}} e^{2\pi i Tr(\sum_{j,k=1}^d c_j x_k^j - \sum_{j,k=1}^d c_j y_k^j)/p} \\
&= \sum_{\substack{x_1,\ldots,x_d\in\mathbb{F}_q \\ y_1,\ldots,y_d\in\mathbb{F}_q}} \sum_c e^{2\pi i Tr(\sum_{j,k=1}^d c_j (x_k^j - y_k^j))/p} \\
&= \sum_{\substack{x_1,\ldots,x_d\in\mathbb{F}_q \\ y_1,\ldots,y_d\in\mathbb{F}_q}} \prod_{j=1}^d (\sum_{c_j\in\mathbb{F}_q} e^{2\pi i Tr(c_j \sum_{k=1}^d (x_k^j - y_k^j))/p})
\end{aligned}$$

Now,

$$\sum_{c_j\in\mathbb{F}_q} e^{2\pi i Tr(c_j \sum_{k=1}^d (x_k^j - y_k^j))/p} = 0$$

unless $\sum_{k=1}^d (x_k^j - y_k^j) = 0$, in which case it equals q.

So

$$\prod_{j=1}^d (\sum_{c_j\in\mathbb{F}_q} e^{2\pi i Tr(c_j \sum_{k=1}^d (x_k^j - y_k^j))/p}) = \sum_{\substack{x_1,\ldots,x_d\in\mathbb{F}_q \\ y_1,\ldots,y_d\in\mathbb{F}_q \\ \sum_{k=1}^d (x_k^j - y_k^j)=0 \\ j=1,\ldots,d}} q^d$$

Now, $\sum_{k=1}^{d}(x_k^j - y_k^j) = 0$ for $j = 1, \ldots, d$ means $\sum_{k=1}^{d} x_k^j = \sum_{k=1}^{d} y_k^j$ for $j = 1, \ldots, d$.
But this is true if and only if $\{y_1, \ldots, y_d\}$ is a permutation of $\{x_1, \ldots, x_d\}$, $since$d¡p, $asfollowsfromtheNewt$

So we have

$$\sum_c |S_c|^{2d} \leq q^d q^d d!$$

(The second q^d is for the d x's $\in \mathbb{F}_q$, and $d!$ is for the permutation of the d y's.)

So putting everything together, we get:
$((q-1)/d)|S|^{2d} \leq d!q^{2d}$, so $|S|^{2d} \leq dd!q^{2d-1}q/(q-1)$. But $q/(q-1) \leq 2$, so $|S|^{2d} \leq 2dd!q^{2d-1}$.

Therefore, taking the $2d$ th root, we get

$$|S| \leq (2dd!)^{\frac{1}{2d}} q^{1-\frac{1}{2d}}$$

as was desired. □

Lemma 29. $\exists \Lambda \subseteq \mathbb{F}_q^*, |\Lambda| \geq \frac{q-1}{d}$ such that $\forall \lambda_1, \lambda_2 \in \Lambda, \lambda_1 \neq \lambda_2$, we have $\lambda_1^d \neq \lambda_2^d$.

Proof. Take Λ maximal with the property that $\forall \lambda_1, \lambda_2 \in \Lambda, \lambda_1 \neq \lambda_2$.
Now $\sharp\{x \in \mathbb{F}_q^* | \exists \lambda \in \Lambda, x^d = \lambda^d\} \leq d|\Lambda|$, because for each value of λ, we can only find d x's such that the equality holds. If $|\Lambda| < \frac{q-1}{d}$, then $\exists x \in \mathbb{F}_q^*$ such that $x^d \neq \lambda^d \ \forall x \in \Lambda$. Then $\Lambda \cup \{x\}$ has this same property. □

7. Week Seven

Theorem 7.1. *Given $f(x) \in \mathbb{F}_p[x]$ with $1 \leq \deg f < p$, then there exists $C > 0$ such that for all $m \geq 1$*
(1)

$$S_{f,m} := |\sum_{x \in \mathbb{F}_{p^m}} e^{2\pi i Tr_{\mathbb{F}_{p^m}/\mathbb{F}_p}(f(x))/p}| \leq Cp^{m/2}$$

What we will prove is:
Given d, $1 \leq d < p$ there exists $C = C(d) > 1$ such that for all $m \geq 1$ and for all $f(x) \in \mathbb{F}_{p^m}[x]$ with $\deg f = d$, we have
(2) the cardinality of $\{(x,y) \in \mathbb{F}_{p^m}^2 | y^p - y = f(x)\} \leq p^m + Cp^{m/2}$.
Today we will prove that (2) $\Rightarrow$ (1).
Step 1: (2) $\Rightarrow$ # $\{(x,y)|y^p - y = f(x)\} \geq p^m - C'p^{m/2}$ with $C' = C'(d)$.

Proof. Recall that given $\alpha \in \mathbb{F}_{p^m}$, there exists $\beta \in \mathbb{F}_{p^m}, \beta^p - \beta = \alpha$ if and only if $Tr_{\mathbb{F}_{p^m}/\mathbb{F}_p}(\alpha) = 0$ and that $Tr_{\mathbb{F}_{p^m}/\mathbb{F}_p} : \mathbb{F}_{p^m} \to \mathbb{F}_p$ is surjective.
Choose $\alpha_j \in \mathbb{F}_{p^m}$ with $Tr(\alpha_j) = j, j = 0, \ldots, p-1$ and $\alpha_0 = 0$. Then,

$$\sum_{j=0}^{p-1} \#\{(x,y) \in \mathbb{F}_{p^m}^2 | y^p - y = f(x) + \alpha_j\} = p^{m+1}.$$

Indeed, given $x_0 \in \mathbb{F}_{p^m}$, look at $Tr(f(x_0)) \in \mathbb{F}_p$. Then the equation $y^p - y = f(x_0) + \alpha_j$ where $j = -Tr(f(x_0))$ has p solutions, and the equations $y^p - y = f(x_0) + \alpha_j$, $j \neq -Tr(f(x_0))$ have 0 solutions. Summing over $x_0 \in \mathbb{F}_{p^m}$ gives the result.

Now apply (2) to $y^p - y = f(x) + \alpha_j$, $j = 1, \ldots, p-1$

$$\#\{y^p - y = f(x) + \alpha_j\} \leq p^m + Cp^{m/2} \ \forall j \geq 1$$

So

$$\#\{(x,y)|y^p - y = f(x)\} = p^{m+1} - \sum_{j=1}^{p-1} \#\{y^p - y = f(x) + \alpha_j\}$$

$$\geq p^{m+1} - \sum_{j=1}^{p-1} (p^m + Cp^{m/2})$$

$$= p^{m+1} - (p-1)p^m - (p-1)Cp^{m/2}$$

$$= p^m - (p-1)Cp^{m/2}$$

So $C' = (p-1)C$ and

$$= p^m - C'p^{m/2}$$

and we have Step 1.

□

Step 2 How does this connect with exponential sums?

$$\sum_{c=0}^{p-1} S_{cf,m} = \#\{(x,y) \in \mathbb{F}_{p^m} | y^p - y = f(x)\}$$

Proof.

$$\sum_{c=0}^{p-1} S_{cf,m} = \sum_{c \in \mathbb{F}_p} \sum_{x \in \mathbb{F}_{p^m}} e^{2\pi i \operatorname{Tr}(cf(x))/p}$$

$$= \sum_x \sum_c e^{2\pi i c \operatorname{Tr}(f(x))/p}$$

and

$$\sum_c e^{2\pi i c \operatorname{Tr}(f(x))/p} = \begin{cases} 0 & \text{if } Tr(f(x)) \neq 0 \\ p & \text{if } Tr(f(x)) = 0 \end{cases}$$

So

$$\sum_c S_{cf,m} = \sum_{x \in \mathbb{F}_{p^m}} \#\{(x,y) \in \mathbb{F}_{p^m} | y^p - y = f(x)\}$$

as needed. Now

$$\sum_{c=0}^{p-1} S_{cf,m} = \sum_{c=1}^{p-1} S_{cf,m} + p^m$$

So

$$\sum_{c=1}^{p-1} S_{cf,m} = \#\{(x,y) \in \mathbb{F}_{p^m} | y^p - y = f(x)\} - p^m$$

and now (2) $\Rightarrow$

$$|\sum_{c=1}^{p-1} S_{cf,m}| \leq C' p^{m/2}$$

We know from the fact that the L-function is a polynomial that there exists $w_1(c), \ldots, w_d(c) \in \mathbb{C}$ such that for all $m \geq 1$

$$S_{cf,m} = -(w_1(c)^m + \cdots + w_d(c)^m).$$

So we get

$$|\sum_{c=1}^{p-1} \sum_{j=1}^{d} w_j(c)^m| \leq C' p^{m/2} \quad \forall m \geq 1.$$

$$\Rightarrow |w_j(c)| \leq p^{1/2} \quad \forall j, c$$

from prior results.

$$\Rightarrow |S_{f,m}| = |-\sum_{j=1}^{d} w_j(1)^m| \leq d p^{m/2}.$$

□

Step 3 We want to investigate

$$y^p - y = f(x) \quad f(x) \in \mathbb{F}_{p^m}[x], \deg f = d, 1 \leq d < p$$

Rough idea: We will construct a polynomial $W_k(x) \in \mathbb{F}_q[x]$ with $q = p^m$ (with smallest possible degree) and $W_k(x) \neq 0$ which will have a zero of multiplicity k at all $x \in \mathbb{F}_q$ with $Tr(f(x)) = 0$. Then we will get

$$\#\{y^p - y = f(x)\} \leq p \frac{\deg W_k}{k}$$

Example:

$$Tr(f(x)) = f(x) + f(x)^p + \cdots + f(x)^{p^{m-1}} = W_1(x)$$

$\deg W_1(x) = p^{m-1} d$ and $\#\{y^p - y = f(x)\} \leq p^m d$.

$$W_2(x) = \text{“}y^q - y - \frac{dy}{dx}(x^q - x)\text{”}.$$

But

$$\begin{aligned} y^q - y &= y^{p^m} - y^{p^{m-1}} + y^{p^{m-1}} - y^{p^{m-2}} + y^{p^{m-2}} - \cdots - y^{p^2} + y^{p^2} - y^p + y^p - y \\ &= (y^p - y)^{p^{m-1}} + (y^p - y)^{p^{m-2}} + \cdots + (y^p - y)^p + (y^p - y) \\ &= W_1(x) \end{aligned}$$

and

$$y^p - y = f(x) \Rightarrow -\frac{dy}{dx} = f'(x)$$

So $W_2(x) = W_1(x) + f'(x)(x^q - x)$.

Claim 7.2. *$W_2(x)$ has double zeros at roots of $W_1(x)$.*

Proof. $W_2'(x) = f'(x) + f'(x)(-1) + f''(x)(x^q - x) = f''(x)(x^q - x)$, which vanishes for all $x \in \mathbb{F}_q$. □

So,

$$\#\{y^p - y = f(x)\} \leq p\frac{\deg W_2}{2}$$
$$\leq p\frac{\max\{(p^{m-1}d), (p^m + d - 1)\}}{2}$$
$$= \frac{p(p^m + d - 1)}{2} \text{ since } d < p.$$

Back to investigating $y^p - y = f(x)$ over $\mathbb{F}_q$, a finite field,$q = p^m$, $f(x) \in \mathbb{F}_q[x], \deg f = d, 1 \leq d < p$.

We want to eventually consider $\#\{(x, y) \in \mathbb{F}_q^2 | y^p - y = f(x)\}$. Let's think of y as an algebraic function of x:

$K = \mathbb{F}_q(x, y)$

$|$

$\mathbb{F}_q(x)$

(Adjoin a root of $y^p - y - f(x) = 0$). This is a cyclic Galois extension of degree p. Let $a \in \mathbb{F}_q$, $t = x - a$ and form $\mathbb{F}_q((t))$. (field of fractions of the ring of formal power series in t, the ring of formal power series in t is denoted $\mathbb{F}_q[[t]]$) Now $\mathbb{F}_q(x) \to \mathbb{F}_q((t))$ is an injection.

Lemma 7.3. *If there exists $b \in \mathbb{F}_q$ such that $b^p - b = f(a)$, then there exists an embedding*

$$l : K \to \mathbb{F}_q((t))$$

extending the embedding

$$\mathbb{F}_q(x) \to \mathbb{F}_q((t))$$

where $x \to t + a$. such that $l(y) = b + \cdots \in \mathbb{F}_q[[t]]$.

Proof. By Galois theory, we just need to produce some element y_1 of $\mathbb{F}_q[[t]]$ of the form $y_1 = b + \ldots$ satisfying $y_1^p - y_1 = f(t + a)$. Let $g(t) = f(t + a) - f(a)$, then we want to find $z \in \mathbb{F}_q[[t]]$ such that $z = a_1 t + \ldots$ (with no constant term) such that $z^p - z = g(t)$. Now $y_1 = z + b$ and

$$y_1^p - y_1 = z^p - z + b^p - b = g(t) + f(a) = f(t + a).$$

So if we find such a z, we're done. Consider

$$z = -\sum_{j=1}^{\infty} g(t)^{p^j} = g(t) + g(t)^p + g(t)^{p^2} + \ldots$$

Note that $g(t)$ has degree d and $g(0) = 0$. Well,

$$z^p = -(g(t)^p + g(t)^{p^2} + \ldots)$$

so $z^p - z = g(t)$ as required. □

Informally, consider the following

$$\begin{array}{ccc} \mathbb{F}_q(x,y) & \hookrightarrow & \mathbb{F}_q((t)) \\ \Big| & \nearrow & \\ \mathbb{F}_q(x) & & \end{array}$$

$\frac{d}{dx} : \mathbb{F}_q(x) \to \mathbb{F}_q(x)$ is such that

1) $\mathbb{F}_q$-linear
2) $\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}$
3) $\frac{dx}{dx} = 1$

There is a unique extension of this operator to $\mathbb{F}_q(x,y)$.

$$y^p - y = f(x)$$

So we need $-\frac{dy}{dx} = \frac{df}{dx}$ Also, there is a unique continuous extension of this operator in $\mathbb{F}_q((t))$ and it turns out to be $\frac{d}{dt}$.

Let's define higher derivatives (for looking at the order of vanishing). Note: x^p has a multiplicity p zero at zero but $\frac{d^i x^p}{dx^i} = 0$ for all $i \geq 1$. So we need to modify:

$$\frac{d^p x^n}{dx^p} = n(n-1)\dots(n-p+1)x^{n-p}$$

but one of those coefficients is zero which we don't want. So let's use

$$\frac{1}{p!}\frac{d^p x^n}{dx^p} = \frac{n(n-1)\dots(n-p+1)x^{n-p}}{p!} = \binom{n}{p}x^{n-p}$$

i.e. for all $m \geq 1$ define $D^{(m)} : \mathbb{F}_q(x) \to \mathbb{F}_q(x)$ as the unique operator such that $D^{(m)}$ is

1) $\mathbb{F}_q$-linear
2) $D^{(m)}(uv) = \sum_{j=0}^{m} D^{(j)}uD^{(m-j)}v$
3) $D^{(m)}x^n = \binom{n}{p}x^{n-m} \quad \forall n$

"Morally" $D^{(m)} = \frac{1}{m!}\frac{d^m}{dx^m}$.

Theorem 7.4. *a) $D^{(m)}$ has a unique extension to K satisfying 1), 2) and 3)*
b) $D^{(m)}$ has a unique continuous extension to $\mathbb{F}_q((t))$ satisfying 1), 2), and 3).

Moreover, if $\phi : K \to \mathbb{F}_q((t))$ is injective with $\phi(x) = t + a$ then $D^{(m)}\phi(u) = \phi(D^{(m)}u)$ for all $u \in K$.

The proof will be given next week.

8. Week Eight

Definition 13 (Hasse Derivation)**.** A Hasse Derivation on a field F is a collection of operators $D^{(m)} : F \to F$ satisfying:

(1) $D^{(m)}(u+v) = D^{(m)}u + D^{(m)}v$,
(2) $D^{(m)}uv = \sum_{j=0}^{m} D^{(j)}u\, D^{(m-j)}v$,
(3) $D^{(n)} \circ D^{(m)} = \binom{n+m}{m}D^{(n+m)}$.

Morally, in the context of finite fields, we may think of it as "$D^{(m)} = \frac{1}{m!}\left(D^{(1)}\right)^m$".

Theorem 30. (1) There exists a unique Hasse derivation on $\mathbb{F}_q(x)$ satisfying $D^{(m)}x^n = \binom{n}{m}x^{n-m}$ for all $n, m \geq 1$.

(2) There exists a unique continuous Hasse derivation on $\mathbb{F}_q((x))$, satisfying $D^{(m)}x^n = \binom{n}{m}x^{n-m}$ for all $n, m \geq 1$.

(3) If $K = \mathbb{F}_q(x, y)$, where $y^p - y = f(x)$, $f(x) \in \mathbb{F}_q(x)$, $\deg(f) = d$, $(d, q) = 1$, there exists a unique Hasse derivation extending the one in 1. Moreover, under the embedding $K \hookrightarrow \mathbb{F}_q((x))$ the derivations from 2, 3 coincide.

Proof. We first prove 2:
Recall the definition of $\mathbb{F}_q((x))$:

$$\mathbb{F}_q((x)) = \left\{ \sum_{j=n}^{\infty} \alpha_j x^j \,|\, n \in \mathbb{Z},\, \alpha_j \in \mathbb{F}_q \right\}.$$

We see immediately that the uniqueness of the derivation in 2 is clear: there can only be one way to satisfy the constraints of the Hasse derivation, and that of the theorem. We must check existence by showing that the conditions are satisfied for the operator defined by $D^{(m)} \sum \alpha_j x^j = \sum \binom{j}{m} \alpha_j x^{j-m}$.

- $D^{(m)}(u + v) = D^{(m)}u + D^{(m)}v$: This is clear, as power series are added term by term.
- $D^{(m)}uv = \sum_{j=0}^{m} D^{(j)}u\, D^{(m-j)}v$: We write uv as a sum of products, and compare terms to reduce to the case:

$$D^{(m)}(uv) = \sum_{j=0}^{m} D^{(j)}u\, D^{(m-j)}v, \quad u = x^n,\, v = x^k.$$

Now $D^{(m)}x^n x^k = \binom{n+k}{m}x^{n+k-m}$, and

$$\begin{aligned} \sum_{j=0}^{m} D^{(j)}x^n\, D^{(m-j)}x^k &= \sum_{j=0}^{m} \binom{n}{j} x^{n-j}x^{k-(m-j)} \\ &= \sum_{j=0}^{m} \binom{n}{j}\binom{k}{m-j} x^{n+k-m}. \end{aligned}$$

The equality now follows from the following identity[1]:

$$\binom{n+k}{m} = \sum_{j=0}^{m} \binom{n}{j}\binom{k}{m-j}.$$

[1]This may be proved, for instance, by noting that a Hasse derivative may be defined over $\mathbb{C}(x)$ by $D^{(m)} = \frac{1}{m!}\frac{d^m}{dx^m}$ and computing in this field.

- $D^{(n)} \circ D^{(m)} u = \binom{n+m}{m} D^{(n+m)}$: Again we may reduce to checking for x^k:

$$\begin{aligned} D^{(n)} \circ D^{(m)} x^k &= \binom{k}{m}\binom{k-m}{n} x^{k-m-n}, \\ D^{(n+m)} x^k &= \binom{k}{n+m} x^{k-n-m}. \end{aligned}$$

The result then follows from the following identity[2]:

$$\binom{k}{m}\binom{k-m}{n} = \binom{n+m}{m}\binom{k}{n+m}.$$

1 follows from 2 by restriction.

To show the existence of 3, one needs to show that the derivative of y in $\mathbb{F}_q((x))$ remains in K (viewed inside $\mathbb{F}_q((x))$). Uniqueness in 3 follows the usual argument. So, regard $y^p - y = f(x)$ as an equation in $\mathbb{F}_q((x))$. Taking its derivative (in $\mathbb{F}_q((x))$) then gives:

$$D^{(m)} y^p - D^{(m)} y = D^{(m)} f(x).$$

On the other hand, using induction and property 2 of the Hasse derivative, one obtains

$$D^{(m)} y^p = \sum_{m_1+\ldots+m_p=m} D^{(m_1)} y \ldots D^{(m_p)} y.$$

Now, the m'th derivative of y only appears on the right as $y^{p-1} D^{(m)} y$, and occurs p times, so that we may write $D^{(m)} y^p$ as

$$D^{(m)} y^p = \sum_{\substack{m_1+\ldots+m_p=m \\ 0<m_1,\ldots,m_p<m}} D^{(m_1)} y \ldots D^{(m_p)} y.$$

Now, using induction we may define $D^{(m)} y^p$ in terms of $D^{(m)} f(x)$ and lower derivatives of y, which finally allows us to get $D^{(m)} y$. In particular, we have:

$$\begin{aligned} D^{(m)} y &= -D^{(m)} f(x), \quad m < p, \\ D^{(p)} y &= -D^{(p)} f(x) - (Df(x))^p. \end{aligned}$$

□

Theorem 31. For all $u \in \mathbb{F}_q[[x]]$ we have $u^q - u = \sum_{m=1}^{\infty} D^{(m)} u (x^q - x)^m$.

Proof. Note that this series converges, as the lowest term of $(x^q - x)^m$ is x^m, so there are only finitely many terms for each x^j.

[2]A triviality:

$$\begin{aligned} \binom{k}{m}\binom{k-m}{n} &= \frac{k!}{m!(k-m)!}\frac{(k-m)!}{n!(k-m-n)!} = \frac{k!}{m!n!(k-m-n)!}, \\ \binom{n+m}{m}\binom{k}{n+m} &= \frac{(n+m)!}{m!n!}\frac{k!}{(n+m)!(k-m-n)!} = \frac{k!}{m!n!(k-m-n)!}. \end{aligned}$$

Both sides are linear, so it is enought to prove the proposition for $u = x^n$, $n \geq 0$:

$$\begin{aligned}
\text{LHS} &= u^q - u \\
&= x^{nq} - x^n, \\
\text{RHS} &= \sum_{m=1}^{\infty} \binom{n}{m} x^{n-m}(x^q - x)^m \\
&= \sum_{m=1}^{n} \binom{n}{m} x^{n-m}(x^q - x)^m \\
&= [x + (x^q - x)]^n - \binom{n}{0} x^n \\
&= x^{qn} - x^n.
\end{aligned}$$

□

Recall, we want a bound on

$$N = |\mathcal{C}|, \quad \mathcal{C} = \{(x, y) \in \mathbb{F}_q | y^p - y = f(x)\}.$$

To this end, we wish to construct a non-zero polynomial $W_k(x)$ which has a zero of multiplicity at least k at each $x \in P_1\mathcal{C}$, where P_1 is the canonical projection onto the first coordinate. Recall that the existence of such a polynomial shows

$$N \leq p\frac{\text{deg}W_k(x)}{k}.$$

One possible candidate for $W_k(x)$ is given by

$$W_k(x) = y^q - y - \sum_{m=1}^{k-1} D^{(m)}y(x^q - x)^m. \tag{4}$$

We have already shown that this is a polynomial in x, and it obviously has the correct multiplicities at its zeroes (if it is non-zero), but we still need to show that it is, in fact, non-zero.

Writing $q = p^r$, we get

$$\begin{aligned}
W_k(x) &= y^q - y - \sum_{m=1}^{k-1} D^{(m)}y(x^q - x)^m \\
&= \sum_{j=1}^{r-1} (f(x))^p - \sum_{m=1}^{k-1} D^{(m)}y(x^q - x)^m.
\end{aligned}$$

Now

$$y = -\sum_{j=0}^{\infty} f(x)^{p^j},$$

so

$$D^{(m)}y = \sum_{p^j|m} \left(D^{m/p^j} f(x)\right)^{p^j}, \quad m \geq p$$

so that we see that, in fact, $D^{(m)}y \in \mathbb{F}_q[x]$, for $m \geq 1$. Now

$$\deg D^{(m)}y \leq \max_{p^j|m} p^j \left(d - \frac{m}{p^j}\right) \leq m(d-1),$$

so

$$\deg D^{(m)}y(x^q - x)^m \leq m(d-1) + mq$$

which then shows

$$\deg W_k(x) \leq \max\left\{\frac{dq}{p}, (k-1)(d-1+q)\right\}.$$

This finally gives a bound for the number of points on the curve:

$$N \leq \frac{p}{k}\max\left\{\frac{dq}{p}, (k-1)(d-1+q)\right\}.$$

What would a good k be? Increasing k decreases the importance of $\frac{dq}{p}$ relative to $(k-1)(d-1+q)$. So the best k would be exactly that k where $\frac{dq}{p}$ becomes less than $(k-1)(d-1+q)$. If k were a real parameter, this would be at

$$k = \frac{dq}{p(d-1+q)} + 1.$$

We must thus compare $\left\lfloor \frac{dq}{p(d-1+q)} \right\rfloor + 1$, $\left\lceil \frac{dq}{p(d-1+q)} \right\rceil + 1$. We may, however, allow k to be real if we simply wish to understand the behaviour of our bound. We see

$$\begin{aligned} N &\lesssim \frac{p}{\frac{dq}{p(d+q-1)} + 1}\frac{dq}{p} \\ &= \frac{dqp(d+q-1)}{dq + p(d+q-1)} = B. \end{aligned}$$

When is this good? If q is large, we see that

$$B \sim \frac{dpq}{d+p},$$

which is not at all good (recall that we have bounds of pq, dq for free). However, if $d \sim p$, we have

$$B \sim \frac{dq}{2}.$$

We still need to check that the $W_k(x) \neq 0$. From the expression for $W_k(x)$ (Eq. 4), and using

$$D\left(D^{(m)}y(x^q - x)^m\right) = -D^{(m)}ym(x^q - x)^{m-1} + (x+1)D^{(m+1)}y(x^q - x)^m,$$

we see that

$$\begin{aligned} DW_k &= -Dy - \sum_{m=1}^{k-1} \left(mD^{(m)}y(x^q - x)^{m-1} - (m+1)D^{(m+1)}y(x^q - x)^m \right) \\ &= -Dy + Dy - kD^{(k)}y(x^q - x)^{k-1} \\ &= -kD^{(k)}y(x^q - x)^{k-1}. \end{aligned}$$

Of course, $W_k(x) \neq 0$ if $DW_k(x) \neq 0$. We certainly have $(x^q - x)^m \neq 0$. Thus, if we can determine when $-kD^{(k)}y \neq 0$ we may find a sufficient condition for $W_k(x)$ being non-zero. Of course, we need $k \neq 0 \mod p$. This is no problem, as we have two almost optimal values for k, and at least one of those will be non-zero mod p. We now make the computation:

$$\begin{aligned} D^{(k)}y &= D^{(k)} \left(-\sum_{p^j | m} \left(D^{(m/p^j)} f(x) \right)^{p^j} \right) \\ &= -D^{(k)} f(x), \end{aligned}$$

so, if p does not devide k, and $D^{(k)}f(x) \neq 0$, then $W_k(x) \neq 0$. But $f(x) = a_0 x^d + \ldots$, so that

$$D^{(k)}f(x) = \binom{d}{k} a_0 x^{d-k} + \ldots.$$

Thus it is enough for $\binom{d}{k} \neq 0 \mod p$ ($k < d$ if k is optimal).

8.1. **An improvement on $W_k(x)$.** Let us consider

$$n = \#\{ap + bd | a, b \geq 0,\ (a,b) \neq (0,0),\ b < p,\ ap + bd \leq M\}$$

One may show

$$n = M - \frac{(p-1)(d-1)}{2}$$

if $M > (p-1)(d-1)$. Now define

$$\{z_1, \ldots, z_n\} = \{x^a y^b | a, b \geq 0,\ (a,b) \neq (0,0),\ b < p,\ ap + bd \leq M\}.$$

Then the z_i are $\mathbb{F}_q$-linearly independent as elements of $\mathbb{F}_q(x,y)$. We consider the determinant:

$$W = \det \begin{pmatrix} z_1^q - z_1 & z_2^q - z_2 & \cdots & z_n^q - z_n \\ Dz_1 & Dz_2 & \cdots & Dz_n \\ \vdots & \ddots & & \vdots \\ D^{(n-1)}z_1 & \cdots & & D^{(n-1)}z_n \end{pmatrix} \tag{5}$$

We shall see subsequently that this has all the properties required of a candidate W. We may relate it to the previous W by noting that, if $M = d$, then

$$\begin{aligned} ap + bd &\leq M \\ \Leftrightarrow ap + bd &\leq d \end{aligned}$$

which is true iff $b = 1$, $a = 0$ or $b = 0$, $a \leq \frac{d}{p}$. Thus

$$\{z_1, \ldots, z_n\} = \{y, x, x^2, \ldots, x^{(d/p)}\}$$

and we recover the original $W_k(x)$ (exercise).

9. Week Nine

Lemma 32. (1) If $z_1, ..., z_n \in \mathbb{F}_q[[x]]$, then W has a zero of multiplicity at least n at $x = 0$.

(2) If z_j has a zero of order exactly j at $x = 0$ for $j = 1, ..., n$ then W has a zero of exactly order n at $x = 0$ and, in particular, W is not identically zero.

Proof: Use $u^q - u = \sum_{m=1}^{\infty} \mathrm{D}^{(m)} u (x^q - x)^m$ from last class to substitute for $z_j^q - z_j$ in W. Thus

$$W = \det \begin{pmatrix} \sum_{m=1}^{\infty} \mathrm{D}^{(m)} z_1 (x^q - x)^m & \ldots & \sum_{m=1}^{\infty} \mathrm{D}^{(m)} z_n (x^q - x)^m \\ \mathrm{D}^{(1)} z_1 & \ldots & \mathrm{D}^{(1)} z_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} z_1 & \ldots & \mathrm{D}^{(n-1)} z_n \end{pmatrix}.$$

Now multiply the j^{th} row for $j = 2, ..., n-1$ by $-(x^q - x)^{(j-1)}$ and add it to the first row. This cancels the first $n - 1$ terms of each infinite sum. Thus,

$$W = \det \begin{pmatrix} \sum_{m=n}^{\infty} \mathrm{D}^{(m)} z_1 (x^q - x)^m & \ldots & \sum_{m=n}^{\infty} \mathrm{D}^{(m)} z_n (x^q - x)^m \\ \mathrm{D}^{(1)} z_1 & \ldots & \mathrm{D}^{(1)} z_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} z_1 & \ldots & \mathrm{D}^{(n-1)} z_n \end{pmatrix}.$$

We can then factor out $(x^q - x)^n$ out of the first row and the determinant giving that

$$W = (x^q - x)^n \det \begin{pmatrix} \sum_{m=n}^{\infty} \mathrm{D}^{(m)} z_1 (x^q - x)^{m-n} & \ldots & \sum_{m=n}^{\infty} \mathrm{D}^{(m)} z_n (x^q - x)^{m-n} \\ \mathrm{D}^{(1)} z_1 & \ldots & \mathrm{D}^{(1)} z_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} z_1 & \ldots & \mathrm{D}^{(n-1)} z_n \end{pmatrix}.$$

Note that this shows that W has a zero of multiplicity at least n at $x = 0$ proving (1).

Now $(x^q - x)^n = x^n (x^{q-1} - 1)^n$ has a zero of order exactly n, so if the determinant of the last matrix above does not vanish at $x = 0$ then we will have proven (2). We can now use the additional hypothesis that $x_j = a_j x^j + ...$ for $j = 1, ..., n$ with $a_j \neq 0$. Thus, $\mathrm{D}^{(m)} z_j = a_j \binom{j}{m} x^{j-m} + ...$ The last term of the first row is $\sum_{m=n}^{\infty} \mathrm{D}^{(m)} z_n (x^q - x)^{m-n}$. When $x = 0$, this becomes $\mathrm{D}^{(n)} z_n(0) = a_n$. On the second row, $\mathrm{D}^{(1)} z_1(0) = a_1$ when $x = 0$ and all the other entries are zero. On the third row, the second entry $\mathrm{D}^{(2)} z_2(0) = a_2$ is

nonzero when $x = 0$; all later entries are 0. We can see the pattern of $\mathrm{D}^{(j)} z_j(0) = a_j$ and $\mathrm{D}^{(k)} z_j = 0$ for $j > k$ when $x = 0$. Thus by reordering the rows,

$$\widetilde{W} = \pm \det \begin{pmatrix} a_1 & 0 & 0 & ... & 0 \\ * & a_2 & 0 & ... & 0 \\ * & * & a_3 & ... & 0 \\ \vdots & & & & \vdots \\ * & * & * & ... & a_n \end{pmatrix}.$$

$\widetilde{W} = \pm a_1 a_2 a_3 ... a_n \neq 0$ by hypothesis. Thus, W has a zero at $x = 0$ of order exactly n, proving (2).

If $\bar{z_j} = \sum_{k=1}^{n} a_{jk} z_k$ with $a_{jk} \in \mathrm{F}_q$ and you make the determinant with the $\bar{z_j}$, then $\bar{W} = \det(a_{jk}) W$. If $z_1, ..., z_n$ all vanish at $x = 0$, $\mathrm{ord}(z_1) = \min \mathrm{ord}\ z_j$, then replace $z_2, ..., z_n$ with $z_2 - a_{21} z_1, ..., z_n - a_{n1} z_1$. Now the orders will be increasing but not necessarily consecutive. The lemma (2) does not necessarily hold.

For example, let $n = 2, p \geq 3$, $z_1 = x$ and $z_2 = x^p$. Then, $W = \det \begin{pmatrix} x^q - x & x^{pq} - x^p \\ 1 & 0 \end{pmatrix}$ has order p, not $n = 2$! For another example, let $n = 3, z_1 = x, z_2 = x^p$, and $z_3 = x^{p^2}$. Then,

$$W = \det \begin{pmatrix} x^q - x & x^{pq} - x^p & x^{p^2 q} - x^{p^2} \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \equiv 0.$$

Exercise: If $z_j = a_j x^{\epsilon_j} + ...$ where $\epsilon_1 < \epsilon_2 < ...$ then ord $W \geq n + \sum_{j=1}^{n} (\epsilon_j - j)$.

Again, $y^p - y = f(x)$ in F_q, deg $f = d$ and $(d, q) = 1$. We want an upper bound for the number of solutions (x, y). We need only look at when q is square so assume $q = p^m$ where m is even. Consider the set $\{x^a y^b \mid 0 \leq a, b, (a, b) \neq (0, 0), b \leq p-1, ap + bd \leq \sqrt{q}\}$. This is a set of $\sqrt{q} - \frac{(p-1)(d-1)}{2}$ functions. Any subset of these functions will work fine for part (1) of the lemma, but applying part (2) is difficult. However, at $x = \infty \ldots$

Theorem 33. Every integer $k \geq (p-1)(d-1)$ has a unique representation as $k = ap + bd$ where $0 \leq a, b$ and $b \leq p - 1$.

Elementary number theory gives this theorem. If we have two representations for $k = a_0 p + b_0 d = a_1 p + b_1 d$ then $a_1 = a_0 + sd$ and $b_1 = b_0 - sp$ for some s. The fact that $k \geq (p-1)(d-1)$ gives that there is at least one solution in the first quadrant of the a, b plane. The restriction that $b \leq p - 1$ ensures the uniqueness.

Thus, for $j = 1, ..., \sqrt{q} - (p-1)(d-1)$ we can find (a, b) satisfying $0 \leq a, b$ and $b \leq p-1$ such that $ap + bd = \sqrt{q} - j$ and a corresponding $z_j = x^{a_j} y^{b_j}$. Let $n = \sqrt{q} - (p-1)(d-1)$

and $z_1, ..., z_n$ have been chosen. Again, $W = \det \begin{pmatrix} z_1^q - z_1 & ... & z_n^q - z_n \\ \mathrm{D}^{(1)} z_1 & ... & \mathrm{D}^{(1)} z_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} z_1 & ... & \mathrm{D}^{(n-1)} z_n \end{pmatrix}$. Note, as we limit j, we are only using the higher powers of x and y. As we are throwing out some functions from W, it is no longer a polynomial of just x. Now $W \in \mathrm{F}_q(x, y) = K$ where y is a solution to $y^p - y = f(x)$.

$\mathrm{F}_q(x) \subset K$ and $\mathrm{F}_q(x)$ embeds into $\mathrm{F}_q((t))$ with the embedding given by $t = x - a$ with $a \in \mathrm{F}_q$. If there exists b such that $b^p - b = f(a)$, we can extend the embedding to K. We want the embedding such that the image of $x, y \notin \mathrm{F}_q[[t]]$.

Consider $x = \alpha t^{-r} + ...$ and $y = \beta t^{-s} + ...$ with $r, s > 0$ (x, y with poles). Then $y^p - y = \beta^p t^{-ps} + ...$ and $f(x) = a_0 \alpha^d t^{-rd} + ...$ so we need $ps = rd$. Since $(p, d) = 1$, $p|r$ and $d|s$. The simplest case would be $p = r, d = s$.

Lemma 34. There is an embedding of K into $\mathrm{F}_q((t))$ such that the image of x and y satisfy $x = \alpha t^{-p} + ...$ and $y = \beta t^{-d} + ...$ where $\alpha\beta \neq 0$.

Proof: Postponed.

10. Week Ten

Lemma 10.1.

$$W := \det \begin{pmatrix} z_1^q - z_1 & \cdots & z_n^q - z_n \\ \mathrm{D}^{(1)} z_1 & \cdots & \mathrm{D}^{(1)} z_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} z_1 & \cdots & \mathrm{D}^{(n-1)} z_n \end{pmatrix} = \pm \det \begin{pmatrix} 1 & z_1^q & \cdots & z_n^q \\ 1 & z_1 & \cdots & z_n \\ 0 & \mathrm{D}^{(1)} z_1 & \cdots & \mathrm{D}^{(1)} z_n \\ \vdots & \vdots & & \vdots \\ 0 & \mathrm{D}^{(n-1)} z_1 & \cdots & \mathrm{D}^{(n-1)} z_n \end{pmatrix}$$

Proof. On the second determinant, subtract the second row from the first, then eliminate the first column and second row. □

Lemma 10.2. *For any functions $z_0, ..., z_n, h \neq 0$:*

$$h^{q+n} \det \begin{pmatrix} z_0^q & \cdots & z_n^q \\ z_0 & \cdots & z_n \\ \mathrm{D}^{(1)} z_0 & \cdots & \mathrm{D}^{(1)} z_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} z_0 & \cdots & \mathrm{D}^{(n-1)} z_n \end{pmatrix} = \det \begin{pmatrix} (hz_0)^q & \cdots & (hz_n)^q \\ hz_0 & \cdots & hz_n \\ \mathrm{D}^{(1)}(hz_0) & \cdots & \mathrm{D}^{(1)}(hz_n) \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)}(hz_0) & \cdots & \mathrm{D}^{(n-1)}(hz_n) \end{pmatrix}$$

Proof.
- $(hz_j)^q = h^q z_j^q$ so factor h^q from first row.
- $hz_j = hz_j$ (duh!) so factor h from second row.
- $\mathrm{D}(hz_j) = h\mathrm{D}z_j + (\mathrm{D}h)z_j$ so subtract from third row the second row times $\mathrm{D}h/h$ and then factor out h.
  In general, $\mathrm{D}^{(m)}(hz_j) = h\mathrm{D}^{(m)} z_j +$ linear combination of $\mathrm{D}^{(r)} z_j, r < m$, so a linear combination of rows 2 to m can be added to this row to change it to $h\mathrm{D}^{(m)} z_j$.

□

Taking $h = z_1^{-1}, z_0 = 1$,

$$W = \pm z_1^{q+n} \det \begin{pmatrix} (1/z_1)^q & 1 & (z_2/z_1)^q & \cdots & (z_n/z_1)^q \\ 1/z_1 & 1 & z_2/z_1 & \cdots & z_n/z_1 \\ \mathrm{D}^{(1)}(1/z_1) & 0 & \mathrm{D}^{(1)}(z_2/z_1) & \cdots & \mathrm{D}^{(1)}(z_n/z_1) \\ \vdots & \vdots & \vdots & & \vdots \\ \mathrm{D}^{(n-1)}(1/z_1) & 0 & \mathrm{D}^{(n-1)}(z_2/z_1) & \cdots & \mathrm{D}^{(n-1)}(z_n/z_1) \end{pmatrix}$$

Or with $u_j = z_{j+1}/z_1$ $(1 \le j \le n-1)$, $u_n = 1/z_1$,

$$W = \pm z_1^{q+n} \det \begin{pmatrix} u_1^q - u_1 & \cdots & u_n^q - u_n \\ \mathrm{D}^{(1)} u_1 & \cdots & \mathrm{D}^{(1)} u_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} u_1 & \cdots & \mathrm{D}^{(n-1)} u_n \end{pmatrix}.$$

With $z_j = \gamma_j t^{j-\sqrt{q}} + \cdots$ we get

$$u_j = \frac{\gamma_{j+1} t^{j-\sqrt{q}+1} + \cdots}{\gamma_1 t^{1-\sqrt{q}} + \cdots} = \frac{\gamma_{j+1}}{\gamma_1} t^j + \cdots$$

and $u_n = \gamma_1^{-1} t^{\sqrt{q}-1} + \cdots$.
Using the identity $u_j^q - u_j = \sum_{n=1}^{\infty} \mathrm{D}^{(m)} u_j (t^q - t)^m$,

$$W = \pm z_1^{q+n} (t^q - t)^n \det \begin{pmatrix} \sum_{m=n}^{\infty} \mathrm{D}^{(m)} u_1 (t^q - t)^{m-n} & \cdots & \sum_{m=n}^{\infty} \mathrm{D}^{(m)} u_n (t^q - t)^{m-n} \\ \mathrm{D}^{(1)} u_1 & \cdots & \mathrm{D}^{(1)} u_n \\ \vdots & & \vdots \\ \mathrm{D}^{(n-1)} u_1 & \cdots & \mathrm{D}^{(n-1)} u_n \end{pmatrix} = \pm z_1^{q+n} (t^q - t)^n F.$$

Using

$$\mathrm{D}^{(m)} u_n = \begin{cases} \binom{\sqrt{q}-1}{m} \gamma_1^{-1} t^{\sqrt{q}-1-m} + \cdots & m < \sqrt{q} - 1 \\ \text{holomorphic in t} & m \ge \sqrt{q} \end{cases},$$

the upper right entry of F is

$$\sum_{m=n}^{\infty} \mathrm{D}^{(m)} u_n (t^q - t)^{n-m} = \left(\sum_{m=n}^{\sqrt{q}-1} \binom{\sqrt{q}-1}{m} \right) \gamma_1^{-1} t^{\sqrt{q}-1-n} + \cdots$$

If $m < n$, $\mathrm{D}^{(m)} u_n = \binom{\sqrt{q}-1}{m} \gamma_1^{-1} t^{\sqrt{q}-1-m} + \cdots$ so

$$\mathrm{ord}_{t=0} D^{(m)} u_n \ge \sqrt{q} - 1 - m > \sqrt{q} - 1 - n.$$

So any term in the expansion of the determinant F involving a $\mathrm{D}^{(m)} u_n$, $m < n$, will have order of vanishing at $t = 0$ strictly bigger than $\sqrt{q} - 1 - n$.

We want to evaluate $F/t^{\sqrt{q}-1-n}$ at $t=0$:

$$Du_1 = \frac{\gamma_2}{\gamma_1} + \cdots \neq 0 \text{ at } t=0.$$
$$Du_j \text{ vanishes at } t=0, j>1.$$
$$D^{(m)}u_j(0) = \begin{cases} 0 & \text{if } m<j, \\ \gamma_{j+1}/\gamma_1 \neq 0 & m=j \\ * & m>j \end{cases}$$

so $F/t^{\sqrt{q}-1-n}$ at $t=0$ is

$$\det \begin{pmatrix} * & * & * & \dots & * & \sum_{m=n}^{\sqrt{q}-1}\binom{\sqrt{q}-1}{m} \\ \gamma_2/\gamma_1 & 0 & 0 & \dots & 0 & 0 \\ * & \gamma_3/\gamma_1 & 0 & \dots & 0 & 0 \\ * & * & \gamma_4/\gamma_1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & * & \dots & \gamma_{n-1}/\gamma_1 & 0 \end{pmatrix}.$$

This immediately gives the following theorem.

Theorem 35. If $\sum_{m=n}^{\sqrt{q}-1}\binom{\sqrt{q}-1}{m} \not\equiv 0 \pmod p$ then

$$\begin{aligned} \operatorname{ord}_{t=0} W &= \sqrt{q}-1-n+n-(\sqrt{q}-1)(q+n) \\ &= -(\sqrt{q}-1)(q+n-1) \end{aligned}$$

and in particular, W is not identically 0.

If $\sum_{m=n}^{\sqrt{q}-1}\binom{\sqrt{q}-1}{m} \equiv 0 \pmod p$ then we trade n for $n-1$. Then we have to worry about $\sum_{m=n-1}^{\sqrt{q}-1}\binom{\sqrt{q}-1}{m} = \binom{\sqrt{q}-1}{n-1} \not\equiv 0$ which is true by

Lemma 36. $\binom{\sqrt{q}-1}{k} \not\equiv 0 \pmod p$ if $0 \le k \le \sqrt{q}-1$.

Proof. $\sqrt{q} = p^r$ for some r; $\sqrt{q}-1 = p^r-1 = (p-1)+p(p-1)+\cdots+p^{r-1}(p-1)$, and

$$\begin{aligned} \sum_{k=0}^{\sqrt{q}-1}\binom{\sqrt{q}-1}{k}x^k &= (x+1)^{\sqrt{q}-1} \\ &= \prod_{j=0}^{r-1}(x+1)^{(p-1)p^j} && \text{(by above comment)} \\ &\equiv \prod (x^{p^j}+1)^{p-1} && \pmod p \\ &= \prod_j \left(\sum_{b=0}^{p-1}\binom{p-1}{b}x^{p^j b}\right) \\ &= \sum_{b_0,\dots,b_{r-1}=0}^{p-1}\left(\prod_{j=0}^{r-1}\binom{p-1}{b_j}\right)x^{b_0+b_1p+\cdots+b_{r-1}p^{r-1}} \end{aligned}$$

but $b_0 + b_1 p + \cdots + b_{r-1}p^{r-1}$ is the base p expansion of some $k < p^r$ so by uniqueness $\prod_{j=0}^{r-1} \binom{p-1}{b_j}$ is the coefficient of x^k. Therefore $\binom{\sqrt{q}-1}{k} \equiv \prod_{j=0}^{r-1} \binom{p-1}{b_j} \not\equiv 0 \pmod{p}$. $\square$

Let $K = \mathbb{F}_q(x, y)$, $y^p - y = f(x)$.

Last week: $K \hookrightarrow \mathbb{F}_q((t)), t = x - a$

Tuesday: $K \hookrightarrow \mathbb{F}_q((t)), x = \alpha t^{-p} + \cdots, y = \beta t^{-d} + \cdots$

All the analysis so far used derivatives with respect to t for the appropriate t. Denote the derivative with respect to t by $\mathrm{D}_t^{(m)}$ and the derivative with respect to x by $\mathrm{D}_x^{(m)}$.

If $t = x - a$ then $\mathrm{D}_t^{(m)} = \mathrm{D}_x^{(m)}, \forall m$. What if $t \neq x - a$?

Chain rule: $\mathrm{D}_t^{(1)} = \mathrm{D}_t^{(1)} x \cdot \mathrm{D}_x^{(1)}$ (same as $\frac{d}{dt} = \frac{dx}{dt}\frac{d}{dx}$). What about higher derivatives?

Let T be a new variable. Define α_{nk} by

$$\left(\sum_{m=1}^{\infty} \mathrm{D}_t^{(m)} x \cdot T^m\right)^n = \sum_{k=1}^{\infty} \alpha_{nk} T^k \in \mathbb{F}_q((t))[\![T]\!], \alpha_{nk} \in \mathbb{F}_q((t))$$

It follows that

$$\alpha_{nk} = \begin{cases} 0 & k < n \\ (\mathrm{D}_t^{(1)} x)^n & k = n \\ \text{some horrible polynomial in } \mathrm{D}_t^{(1)} x, \ldots, \mathrm{D}_t^{(k)} x & k > n \end{cases}$$

Theorem 10.3. *If $t \in K$, then for all $u \in K$,*

$$\mathrm{D}_t^{(k)} u = \sum_{n=1}^{k} \alpha_{nk} \mathrm{D}_x^{(n)} u.$$

We postpone the proof. Here are a couple of examples:

$$k = 1 : \mathrm{D}_t^{(1)} u = \alpha_{11} \mathrm{D}_x^{(1)} u = \mathrm{D}_t^{(1)} x \cdot \mathrm{D}_x^{(1)} u$$
$$k = 2 : \mathrm{D}_t^{(2)} u = \alpha_{12} \mathrm{D}_x^{(1)} u + \alpha_{22} \mathrm{D}_x^{(2)} u = \mathrm{D}_t^{(2)} x \cdot \mathrm{D}_x^{(1)} u + (\mathrm{D}_t^{(1)} x)^2 \cdot \mathrm{D}_x^{(2)} u$$

or in conventional notation, by differentiating $\frac{du}{dt} = \frac{dx}{dt}\frac{du}{dx}$,

$$\frac{1}{2}\frac{d^2u}{dt^2} = \frac{1}{2}\frac{d^2x}{dt^2} + \left(\frac{dx}{dt}\right)^2 \frac{1}{2}\frac{d^2u}{dx^2}.$$

Using the above formula we can return to our calculation of W.

Let W_t = determinant for W with $\mathrm{D} = \mathrm{D}_t$. Let W_x = likewise with $\mathrm{D} = \mathrm{D}_x$.

$$W_t = \prod_{k=1}^{n-1} \alpha_{kk} \det \begin{pmatrix} z_1^q - z_1 & \cdots & z_n^q - z_n \\ \frac{1}{\alpha_{1,1}} \mathrm{D}_t^{(1)} z_1 & \cdots & \frac{1}{\alpha_{1,1}} \mathrm{D}_t^{(1)} z_n \\ \vdots & & \vdots \\ \frac{1}{\alpha_{n-1,n-1}} \mathrm{D}_t^{(n-1)} z_1 & \cdots & \frac{1}{\alpha_{n-1,n-1}} \mathrm{D}_t^{(n-1)} z_n \end{pmatrix} = \prod_{k=1}^{n-1} \alpha_{kk} \cdot W_x = (\mathrm{D}_t^{(1)} x)^{n(n-1)/2} \cdot W_x$$

because the k-th row of of the displayed determinant is equal to the k-th row of W_x plus a linear combination of rows 2 through $k-1$ for $k > 2$. We summarize:

Theorem 37. $W_t = \left(\frac{dx}{dt}\right)^{n(n-1)/2} W_x$.

Now we will prove the formula for the D_t in terms of the D_x.

The reader can easily check, using the defining properties of the Hasse derivatives, that the map

$$\mathcal{D}_t : K \to K((T)), u \mapsto \sum_{m=0}^{\infty} \mathrm{D}_t^{(m)} u \, T^m \text{ (so } t \mapsto t+T)$$

is a field homomorphism.

(**Exercise:** prove $\mathcal{D}_t$ is the unique injective homomorphism $K \to K((T))$ mapping t to $t+T$.)

Similarly define

$$\mathcal{D}_x : K \to K((X)), u \mapsto \sum \mathrm{D}_x^{(m)} u \, X^m, x \mapsto x + X$$

I want to define an isomorphism of K-algebras $\Phi : K((T)) \to K((X))$ such that $\Phi \circ \mathcal{D}_t$ is an injective homomorphism from K to $K((X))$ sending x to $x+T$. Once Φ is constructed, by uniqueness of $\mathcal{D}_x$, we get $\Phi \circ \mathcal{D}_t = \mathcal{D}_x$.

We know that $\Phi(\mathcal{D}_t t) = \Phi(t+T) = t + \Phi(T)$ and that $\mathcal{D}_x t = \sum_{m=0}^{\infty} \mathrm{D}_x^{(m)} t \cdot X^m$. Equating these two expressions yield $\Phi(T) = \sum_{m=1}^{\infty} \mathrm{D}_x^{(m)} t \cdot X^m$ and this determines Φ.

Now we use $\Phi \circ \mathcal{D}_t = \mathcal{D}_x$ evaluated at an arbitrary u:

$$\sum_{m=0}^{\infty} \mathrm{D}_x^{(m)} u \cdot X^m = \mathcal{D}_x u = \Phi(\mathcal{D}_t u) = \Phi\left(\sum_{n=0}^{\infty} \mathrm{D}_t^{(n)} u \cdot T^n\right) = \sum_{n=0}^{\infty} \mathrm{D}_t^{(n)} u \cdot \Phi(T)^n =$$

$$\sum_{n=0}^{\infty} \mathrm{D}_t^{(n)} u \cdot \left(\sum_{m=1}^{\infty} \mathrm{D}_x^{(m)} t \cdot X^m\right)^n = \sum_{n=0}^{\infty} \mathrm{D}_t^{(n)} u \sum_{k=0}^{\infty} \alpha_{nk} X^k = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \mathrm{D}_t^{(n)} u \cdot \alpha_{nk} X^k$$

and so

$$\mathrm{D}_x^{(m)} u = \sum_{n=0}^{\infty} \alpha_{nm} \mathrm{D}_t^{(m)} u = \sum_{n=1}^{m} \alpha_{nm} \mathrm{D}_t^{(m)} u$$

because $\alpha_{0m} = 0, m > 0$ and $\alpha_{nm} = 0$ for $n > m$. □

11. Week Eleven

Again, $y^p - y = f(x)$ in F_q, $\deg f = d$ and $(d, q) = 1$. We want an upper bound for the number of solutions (x, y). We need only look at when q is square so assume $q = p^m$ where m is even. Consider the set $\{x^a y^b \mid 0 \le a, b, (a,b) \ne (0,0), b \le p-1, ap+bd \le \sqrt{q}\}$. This is a set of $\sqrt{q} - \frac{(p-1)(d-1)}{2}$ functions.

Thus, for $j = 1, ..., \sqrt{q} - (p-1)(d-1)$ we can find (a, b) satisfying $0 \le a, b$ and $b \le p-1$ such that $ap + bd = \sqrt{q} - j$ and a corresponding $z_j = x^{a_j} y^{b_j}$. Let $n = \sqrt{q} - (p-1)(d-1)$ and $z_1, ..., z_n$ have been chosen. We will work with this n or, if necessary replace n by $n-1$ as above.

With our choice of the $z's$, the determinant W_x will not necessarily be a polynomial in x. Now $W_x \in \mathbb{F}_q(x, y) = K$ where y is a solution to $y^p - y = f(x)$. We will use below

that W_x has as many zeros as poles, with multiplicity. Alternatively we can count zeros and poles of the polynomial in x given by the $K/\mathbb{F}_q(x)$-norm of W_x.

To complete the analysis of the poles of W_x we need to consider what t is.

Lemma 38. There is an embedding of K into $\mathbb{F}_q((t))$, where $t = x^u y^v$ with $up + vd = -1$ such that the image of x and y satisfy $x = \alpha t^{-p} + ...$ and $y = \beta t^{-d} + ...$ where $\alpha\beta \neq 0$ and $\operatorname{ord}_{t=0} dx/dt = (p-1)(d-1)$.

Proof. Let $t = x^u y^v$ where $up + vd = -1$. By writing the equations $y^p - y = f(x)$ and $t = x^u y^v$ in terms of $t, z = t^p x, w = t^d y$ we can recursively find coefficients $z_n, w_n, n \geq 0$, such that $z = \sum_{n\geq 0} z_n t^n, w = \sum_{n\geq 0} w_n t^n$. To compute $\operatorname{ord}_{t=0} dx/dt$, look at $dt/dx = x^u y^v (u/x - v f'(x)/y)$. Then u/x does not have a pole at $t = 0$, while $v f'(x)/y$ has a pole of order $(d-1)p - d$. The result follows.

Now, $W_t = \pm z_1^{q+n}(t^q - t)^n F$ and F has a zero of order $\sqrt{q} - 1 - n$ at $t = 0$, so $\operatorname{ord}_{t=0} W_t = \sqrt{q} - 1 - n + n - (q+n)(\sqrt{q} - 1) = -(\sqrt{q} - 1)(q + n - 1)$. So $\operatorname{ord}_{t=0} W_x = \operatorname{ord}_{t=0}((dt/dx)^{n(n-1)/2} W_t) = -(\sqrt{q} - 1)(q + n - 1) - (p-1)(d-1)n(n-1)/2$. We also proved earlier that each pair $(a, b) \in \mathbb{F}_q^2$ with $y^p - y = f(x)$ contributes a zero of order n to W_x. Thus the total number of such pairs, N say, satisfies $nN \leq (\sqrt{q} - 1)(q + n - 1) + (p-1)(d-1)n(n-1)/2$ and taking into account that n is either $\sqrt{q} - (p-1)(d-1)$ or $\sqrt{q} - (p-1)(d-1) - 1$, yields $N \leq q + O(\sqrt{q})$, which completes the proof of our main theorem. $\square$