The Weil bounds

Topics in Finite Fields (Fall 2013)
Rutgers University
Swastik Kopparty
Last modified: Thursday 16th February, 2017

1 The Statement

As we suggested earlier, the original form of the Weil bound deals with the number of solutions in $\mathbb{F}_q^2$ to a bivariate polynomial $H(X,Y) \in \mathbb{F}_q[X,Y]$.
For a field $\mathbb{F}$, a polynomial $H(X,Y) \in \mathbb{F}[X,Y]$ is called absolutely irreducible if it is irreducible in the polynomial ring $\overline{\mathbb{F}}[X,Y]$ (where $\overline{\mathbb{F}}$ is the algebraic closure of $\mathbb{F}$).
For example: if $\mathbb{F}$ is a field and α is a quadratic non-residue in $\mathbb{F}$, then:

- $Y^2 - X^2$ is reducible,
- $Y^2 - \alpha X^2$ is irreducible, but not absolutely irreducible,
- $Y^2 - X^2 + 1$ is absolutely irreducible.

To see the last item, note that any factorization of $H(X,Y) = Y^2 - X^2 + 1$ in $\overline{\mathbb{F}}[X,Y]$ is also a factorization of $H(X,Y)$ in $\mathbb{K}[Y]$, where $\mathbb{K}$ is the field $\overline{\mathbb{F}}(X)$, and thus the factorization must be of the form $(Y - a(X))(Y + a(X))$, where $a(X) \in \mathbb{K}$ satisfies $a(X)^2 = X^2 - 1$. But this cannot be.
We can now state the Weil bound.

Theorem 1. *Let $H(X,Y) \in \mathbb{F}_q[X,Y]$ be an absolutely irreducible polynomial of degree $\leq d$. Then:*

$$\left|\{(x,y) \in \mathbb{F}_q^2 \mid H(x,y) = 0\}\right| = q \pm O(d^2\sqrt{q}).$$

For small d, this is what we would expect for a random function $R : \mathbb{F}_q^2 \to \mathbb{F}_q$. Thus this is an expression of pseudorandomness.
The Weil bound above implies the character sum Weil bounds that we mentioned in an earlier lecture. This implication is not obvious at all, but certain special cases of it are (for example, the bound on $|\sum_{x \in \mathbb{F}_q} \chi(g(X))|$, when χ is the quadratic character).
We will not prove this Weil bound in full generality, but we will prove several important special cases, and in particular, we will see complete proofs of the character sum Weil bounds.
The absolute irreducibility assumption is important. Let us see 2 examples illustrating what could go wrong with reducibility and with irreducibility but absolute reducibility:

- $Y^2 - X^2$ has about $2q$ zeroes in $\mathbb{F}_q^2$ (which is the union of 2 lines).
- $Y^2 - \alpha X^2$ has only 1 zero in $\mathbb{F}_q^2$, namely $(0,0)$.

The first part of our proof of the character sum bounds is based on bounding the number of $\mathbb{F}_q$ solutions to some special equations. We give two different elementary approaches, one based on Stepanov's methods, and another based on Bombieri's methods. From this, we will be able to get some weaker versions of the character sum bounds we are trying to prove. These character sum bounds will be nontrivial only for certain special characters and over certain kinds of fields.
The second part of the proof is a dizzying display of generating function magic. It will enable us to deduce the theorem for all characters and all fields. This will involve relating the character sum over $\mathbb{F}_q$ with related character sums over $\mathbb{F}_{q^n}$ for all $n \geq 1$.

2 The Weil Bound for multiplicative character sums: Stepanov's proof

2.1 Strategy

We now prove an upper bound on the number of solutions to equations of a certain type.

Theorem 2. *Suppose $e \mid q-1$. Let $g(X) \in \mathbb{F}_q[X]$ be of degree d, such that $Y^e - g(X) \in \mathbb{F}_q[X,Y]$ is absolutely irreducible. Then*

$$|\{x \in \mathbb{F}_q \mid g(x) \text{ is a perfect } e\text{'th power }\}| \leq \frac{q}{e} + O(e^3 d\sqrt{q}).$$

The proof strategy is as follows: we will find a nonzero polynomial $R(X) \in \mathbb{F}_q[X]$ that vanishes whenever x is such that $g(x)$ is a perfect e'th power. Then the number of such x is bounded by the number of zeroes of R, which can be bounded by $\deg(R)$.
Here is one such polynomial: $R(X) = g(X)^{(q-1)/e} - 1$. We have that $R(x) = 0$ whenever $g(x)$ is a perfect e'th power. The number of such x is bounded by $\deg(R) \approx \frac{d}{e} \cdot q$, which doesn't suffice for us. This argument failed for good reason: we didn't use the fact that we are counting x which lie in $\mathbb{F}_q$!
The real argument is based on finding a nonzero polynomial $R(X)$ which vanishes whenever $g(x)$ is a perfect e'th power AND $x^q = x$. We will actually choose R to vanish with multiplicity M at each such x; and the resulting upper bound on the number of x is $\deg(R)/M$.

2.1.1 Multiplicity and Hasse Derivatives

We now discuss multiplicity of vanishing.
We say a polynomial $R(X) \in \mathbb{F}[X]$ vanishes at $\alpha \in \mathbb{F}$ with multiplicity at least M if $(X-\alpha)^M$ divides $R(X)$ (and we denote this by $\mathsf{mult}(R,\alpha) \geq M$.
Over the field $\mathbb{R}$, this can be expressed in terms of derivatives: $\mathsf{mult}(R,\alpha) \geq M$ iff for each $i < M$, we have $\frac{d^i R}{dX^r}(\alpha) = 0$.
In fields of positive characteristic, this is no longer true (essentially because iterated derivatives vanish too easily). Instead, we will use a slightly different notion of derivative, the Hasse derivative.
For $R(X) \in \mathbb{F}[X]$, we define its ith Hasse derivative $R^{(i)}(X) \in \mathbb{F}[X]$ by:

$$R^{(i)}(X) = \text{ coefficient of } Z^i \text{ in } R(X+Z) \ .$$

Note that we are defining the ith Hasse derivative directly, not as the i-fold repetition of the 1st Hasse derivative. In fact, the 1st Hasse derivative equals the 1st usual derivative.
From the definition, we have:

$$R(X+Z) = R^{(0)}(X)) + R^{(1)}(X)Z + R^{(2)}(X)Z^2 + \ldots .$$

This equation should remind you of Taylor's theorem, except it is missing some factorials. Indeed, the ith Hasse derivative is off from the ith usual derivative by a factor of $i!$.
The above equation also easily implies a criterion for vanishing multiplicity: $\mathsf{mult}(R,\alpha) \geq M$ if and only if for each $i < M$, we have $R^{(i)}(\alpha) = 0$.
Observe that the Hasse derivative map $R \mapsto R^{(i)}$ is $\mathbb{F}$-linear. Thus another way to define Hasse derivatives is by specifying it on the monomials:

$$(X^\ell)^{(i)} = \binom{\ell}{i} X^{\ell-i}.$$

The Leibniz product rule for Hasse derivatives takes the following form:

$$(R \cdot S)^{(i)}(X) = \sum_{j+k=i} R^{(j)}(X) S^{(k)}(X).$$

The chain rule (and the Faa di Bruno version) are a bit more complicated.
One can define multivariate Hasse derivatives and multivariate multiplicities analogously. Given $R(X_1, \ldots, X_m) \in \mathbb{F}[X_1, \ldots, X_m]$, the $\mathbf{i} = (i_1, \ldots, i_m)$ Hasse derivative of R, denoted $R^{(\mathbf{i})}(\mathbf{X})$, is defined by:

$$R^{(\mathbf{i})}(\mathbf{X}) = \text{ coefficient of } \mathbf{Z}^{\mathbf{i}} \text{ in } R(\mathbf{X} + \mathbf{Z}) \ .$$

They have many applications, but we will not have time to look at them.

2.2 Interpolating a polynomial

Let $S = \{x \in \mathbb{F}_q \mid g(x)^{(q-1)/e} = 1\}$. We are going to find a polynomial $R(X) \in \mathbb{F}_q[X]$, such that for every $x \in S$, we have $\mathsf{mult}(R, x) \geq M$.
For ease of notation, we only deal with the special case $e = 2$.
We will choose $R(X)$ to be of the form:

$$R(X) = A_0(X, X^q) + g(X)^{(q-1)/2} A_1(X, X^q),$$

where $A_0(X, Y), A_1(X, Y) \in \mathbb{F}_q[X, Y]$ are both polynomials from the set:

$$V = \{P(X, Y) \in \mathbb{F}_q[X, Y] \mid \deg_X(P) \leq A, \deg_Y(P) \leq B\},$$

(for some A, B to be specified later).
Note that V is a AB dimensional space over $\mathbb{F}_q$, and that $(A_0, A_1) \in V^2$. We will first show that if $(A_0, A_1) \neq (0, 0)$, then $R(X)$ is a nonzero polynomial. Next, we will find some homogeneous $\mathbb{F}_q$ linear constraints on $(A_0, A_1) \in V^2$ which imply that $R(X)$ vanishes at each point of S with multiplicity at least M The total number of these constraints is $< 2AB$: thus there is a nonzero $(A_0, A_1) \in V^2$ such that $R(X)$ has the desired vanishing property. Together, this implies that there is a nonzero $R(X)$ with the desired vanishing property, and we get an upper bound on $|S|$.

1. **Nonvanishing of $R(X)$:**

   Lemma 3. *If $A < \frac{q}{2} - d$, and if $(A_0(X, Y), A_1(X, Y)) \neq (0, 0)$, then $R(X) \neq 0$.*

   Proof. Suppose $R(X) = 0$. Then $A_0(X, X^q) + g(X)^{(q-1)/2} A_1(X, X^q) = 0$. So $A_0(X, X^q)^2 - g(X)^{q-1} A_1(X, X^q)^2 = 0$. This implies that $g(X) A_0(X, X^q)^2 - g(X)^q A_1(X, X^q)^2 = 0$. Since $g(X)$ has $\mathbb{F}_q$ coefficients, we get:

   $$g(X) A_0(X, X^q)^2 - g(X^q) A_1(X, X^q)^2 = 0.$$

   Taking this equation mod X^q, we get:

   $$g(X) A_0(X, 0)^2 - g(0) A_1(X, 0)^2 = 0 \mod X^q.$$

   Since the left hand side of this equation has degree $\leq d + (q - 2d) < q$, we get the genuine equality of polynomials (without mod):

   $$g(X) A_0(X, 0)^2 - g(0) A_1(X, 0)^2 = 0.$$

   So $g(X) = g(0) \frac{A_1(X,0)^2}{A_0(X,0)^2}$, which gives us a factorization of $Y^2 - g(X)$ in $\overline{\mathbb{F}}_q[X, Y]$, a contradiction. □

2. **High multiplicity vanishing of R on S:** We will now impose various $\mathbb{F}_q$-linear constraints on A_0, A_1 which will ensure that R vanishes with high multiplicity at each $x \in S$.

   - **Condition 0:** We ask that the following polynomial is identically 0:

     $$A_0(X, X) + 1 \cdot A_1(X, X).$$

     The number of $\mathbb{F}_q$-linear constraints that this imposes equals the maximum possible degree of $A_0(X, X) + A_1(X, X)$, which is $\leq A + B$. By imposing this condition, we have that for each $x \in S$:

     $$R(x) = A_0(x, x^q) + g(x)^{(q-1)/2} \cdot A_1(x, x^q) = A_0(x, x) + 1 \cdot A_1(x, x) = 0.$$

- **Condition 1:** We ask that the following polynomial is identically 0:

$$g(X) \cdot A_0^{(1,0)}(X, X) + \frac{q-1}{2} \cdot A_1(X, X) + g(X) \cdot A_1^{(1,0)}(X, X) = 0.$$

The number of $\mathbb{F}_q$-linear constraints that this imposes is $\leq A+B+d$. By imposing this condition, we have that for each $x \in S$:

$$\begin{aligned} g(x) \cdot R^{(1)}(x) &= g(x) \cdot A_0^{(1,0)}(x, x^q) + \frac{q-1}{2} g(x)^{(q-1)/2} \cdot A_1(x, x^q) + g(x) \cdot g(x)^{(q-1)/2} \cdot A_1^{(1,0)}(x, x^q) \\ &= g(x) \cdot A_0^{(1,0)}(x, x) + \frac{q-1}{2} \cdot A_1(x, x) + g(x) \cdot A_1^{(1,0)}(x, x) \\ &= 0, \end{aligned}$$

and thus $R^{(1)}(x) = 0$ (since $x \in S$ implies $g(x) \neq 0$).

- $\ddots$
- **Condition ℓ:** For general $\ell < M$, we observe that $g(X)^\ell \cdot R^{(\ell)}(X)$ is of the form:

$$g(X)^\ell \cdot R^{(\ell)}(X) = A_{0,\ell}(X, X^q) + g(X)^{(q-1)/2} \cdot A_{1,\ell}(X, X^q),$$

where $A_{0,\ell}(X, Y)$ and $A_{1,\ell}(X, Y)$ are polynomials with X-degree $\leq A + \ell d$ and Y-degre $\leq B$. Then Condition ℓ asks that the following polynomial is identically 0:

$$A_{0,\ell}(X, X) + A_{1,\ell}(X, X).$$

The number of $\mathbb{F}_q$-linear constraints that this imposes equals the maximum possible degree of $A_{0,\ell}(X, X) + A_{1,\ell}(X, X)$, which is $\leq A + \ell d + B$. By imposing this condition, we have that for each $x \in S$:

$$\begin{aligned} g(x)^\ell \cdot R^{(\ell)}(x) &= A_{0,\ell}(x, x^q) + g(x)^{(q-1)/2} A_{1,\ell}(x, x^q) \\ &= A_{0,\ell}(x, x) + A_{1,\ell}(x, x) \\ &= 0, \end{aligned}$$

and thus $R^{(\ell)}(x) = 0$ (since $x \in S$ implies $g(x) \neq 0$).

- $\ddots$

3. **Bounding $|S|$:** If $(A_0, A_1) \in V^2$ satisfy Conditions $0, 1, \ldots, M-1$, then we have $\mathsf{mult}(R, x) \geq M$ for each $x \in S$. The total number of homogeneous $\mathbb{F}_q$-linear constraints on $(A_0, A_1) \in V^2$ imposed by Conditions $0, 1, \ldots, M-1$ is at most:

$$(A+B) + (A+B+d) + (A+B+2d) + \ldots + (A+B+(M-1)d) \leq M(A+B) + d \cdot M^2.$$

If we can ensure that this quantity is less that $\dim(V^2) = 2AB$, then there is a nonzero $(A_0, A_1) \in V^2$ such that $\mathsf{mult}(R, x) \geq M$ for each $x \in S$. If we also have that $A < \frac{q}{2} - d$, then this implies that $R(X)$ is a nonzero polynomial. This implies that

$$|S| \leq \frac{\deg(R)}{M} \leq \frac{1}{M}\left(qB + A + d\frac{q-1}{2}\right) \leq \frac{1}{M} \cdot (q \cdot B + d).$$

Summarizing, if we have the following:

$$M(A+B) + d \cdot M^2 < 2AB,$$

$$A < \frac{q}{2} - d,$$

then we get:

$$|S| \leq \frac{qB + d}{M}.$$

Now we choose parameters A, B, M:

$$A = \frac{q}{2} - d - 1,$$

$$B = \frac{M(A+B) + dM^2}{2A} + 1 = M \cdot \frac{A + B + dM + \frac{2A}{M}}{2A},$$

which ensures that the above conditions hold, and so for every M we get:

$$|S| \leq \frac{qB + d}{M} \leq q \cdot \frac{A + B + dM + \frac{2A}{M}}{2A} + \frac{d}{M}.$$

Thus:

$$|S| \leq q \cdot \left(\frac{1}{2} + \frac{B}{2A} + \frac{dM}{2A} + \frac{1}{M} \right) + \frac{d}{M}.$$

Choosing $M = O(\sqrt{q})$, we get:

$$|S| \leq \frac{q}{2} + O(d\sqrt{q}).$$

2.3 A basic multiplicative character sum bound

Let q be odd and let χ be the quadratic character of $\mathbb{F}_q$. Let $g(X)$ be of degree d with $Y^2 - g(X)$ absolutely irreducible. (This is equivalent to $g(X) \neq \alpha \cdot g_1(X)^2$ for $\alpha \in \overline{\mathbb{F}}_q$, $g_1(X) \in \overline{\mathbb{F}}_q[X]$).
Applying the upper bound we just proved to $g(X)$ and to $\alpha g(X)$, where α is a quadratic nonresidue in $\mathbb{F}_q$, we get:

$$|\sum_{x \in \mathbb{F}_q} \chi(g(x))| \leq O(d\sqrt{q}).$$

This is the kind of bound that we wanted to show.
For general e, we get an upper bound of $q + O(e^2 d\sqrt{q})$ on the number of $(x, y) \in \mathbb{F}_q^2$ with $y^e = g(x)$, provided $Y^e - g(X)$ is absolutely irreducible. This in turn implies, for χ a character of order e, a bound of the form:

$$\left| \sum_{x \in \mathbb{F}_q} \chi(g(x)) \right| \leq O(e^3 d\sqrt{q}). \tag{1}$$

For small e, this is a good bound. For large e, this bound gets worse than the trivial bound of $O(q)$. The upcoming L-function argument will remove this poor dependence on e.

3 The Weil bound for additive character sums: Bombieri's proof

3.1 Strategy

Let p be prime. Let $q = p^n$. Let $\mathsf{Tr} : \mathbb{F}_q \to \mathbb{F}_p$ be the trace map. We say a polynomial $g(X)$ is p-free if it is a linear combination of monomials X^i where $p \not| i$.
The main step in proving the additive character sum Weil bounds is the following bound on the number of solutions to a bivariate equation.

Theorem 4. *For any p-free $g(X) \in \mathbb{F}_q[X]$ of degree $d < p^{n/2-1}$,*

$$|\{x \in \mathbb{F}_q \mid \mathsf{Tr}(g(x)) = 0\}| \leq \frac{q}{p} + O(d\sqrt{q}).$$

By Hilbert's Theorem 90, this can be reformulated as follows.
Let $V = \{(x, y) \in \mathbb{F}_q \times \mathbb{F}_q : P(x, y) = 0\}$, where $P(X, Y) \in \mathbb{F}_q[X, Y]$ is the polynomial $Y^p - Y - g(X)$. Hilbert Theorem 90 implies that $\mathsf{Tr}(g(x)) = 0$ if and only if there exists y such that $y^p - y - g(x) = 0$; note that if this happens then there are exactly p such y's. Therefore it suffices to show that

$$|V| \leq q + O(pd\sqrt{q}).$$

This is the form in which we will do the main argument.
At the very high level, our strategy is as follows. We will find a low-degree polynomial $Q(X, Y)$ relatively prime to $P(X, Y)$ such that for any $(x, y) \in V$, $Q(x, y)$ is 0. Thus the cardinality of V is at at most the number of points of intersection of $P(X, Y) = 0$ and $Q(X, Y) = 0$. We will then use a form of Bezout's Theorem to get an upper bound for this quantity.
These three requirements on $Q(X, Y)$ (low-degree, vanishing on V, and relatively prime to $P(X, Y)$) are in tension with each other. We can easily find $Q(X, Y)$ with the first two properties: any multiple of $P(X, Y)$ will do. To get the third property, we will take a multiple of $P(X, Y)$, reduce this multiple mod $\langle X^q - X, Y^q - Y\rangle$, and hope that this reduction is relatively prime to $P(X, Y)$. Note that reducing mod $\langle X^q - X, Y^q - Y\rangle$ does not affect the property of vanishing on V (since $V \subseteq \mathbb{F}_q^2$).
For simplicity of notation, we will prove the theorem only for $p = 2$. The case of general p is very similar, and we omit it.

3.1.1 A version of the Bezout Theorem

The (a, b)-*degree* of a monomial X^iY^j is defined to be $ai + bj$. The (a, b)-degree of a polynomial is the maximum of the (a, b)-degree of its monomials.
We will need the following version of Bezout's theorem:

Theorem 5. *Let $\mathbb{F}$ be a field. Let $P(X, Y) \in \mathbb{F}[X, Y]$ be a polynomial of the form $u(Y) - f(X)$, with X-degree equal to d_X and Y-degree equal to d_Y. Let $Q(X, Y) \in \mathbb{F}[X, Y]$ be relatively prime to $P(X, Y)$ and have (d_Y, d_X)-degree at most D. Then,*

$$|\{(x, y) \in \mathbb{F} \times \mathbb{F} : P(x, y) = Q(x, y) = 0\}| \leq D + d_X \cdot d_Y.$$

Sketch of proof: Since P is monic in Y, we can reduce $Q(X, Y)$ mod $P(X, Y)$ to get a polynomial $Q'(X, Y)$ which has Y-degree at most $d_Y - 1$. Observe that the reducing mod $P(X, Y)$ does not increase the (d_Y, d_X)-weighted degree. Thus $Q'(X, Y)$ has X-degree at most $\frac{D}{d_Y}$ and Y-degree at most $d_Y - 1$. Now by the standard proof of Bezout's theorem (using resultants), the number of common zeroes between $P(X, Y)$ and $Q'(X, Y)$ (and hence between $P(X, Y)$ and $Q(X, Y)$) is at most $d_X \cdot (d_Y - 1) + d_Y \cdot \frac{D}{d_Y} \leq D + d_X \cdot d_Y$.
We will apply this with $P(X, Y) = Y^2 - Y - g(X)$, and we will thus be interested in the $2, d$ weighted degree of polynomials $Q(X, Y)$.

3.2 A good basis

Let S be the set of all integers that can be written as either $2i$ or $2i + d$, for some nonnegative integer i. Let $S_j = \{s \in S : s \leq j\}$. For any $j \geq d$, we have $|S_j| = (j - (d - 1)/2)$. For any $s \in S$, let $M_s(X, Y)$ be the monomial $X^{s/2}$ if s is even, or $X^{(s-d)/2}Y$ if s is odd. Notice that the $(2, d)$-degree of $M_s(X, Y)$ is s.
First observe that any polynomial $R(X, Y) \in \mathbb{F}_q[X, Y]$ of $(2, d)$-degree j, is congruent modulo $P(X, Y)$ to exactly one polynomial of Y-degree at most 1 (repeatedly replacing every occurrence of Y^2 by $Y + g(X)$). We denote this polynomial $\overline{R(X, Y)}$. In fact, the same argument shows that $\overline{R(X, Y)}$ has $(2, d)$-degree at most j, and is in the $\mathbb{F}_q$ linear span of $\{M_s(X, Y) : s \in S_j\}$.

Claim 6. *$\overline{X^iY^j}$ has $(2, d)$-weighted degree exactly $2i + dj$.*

Proof. It suffices to consider $i = 0$. By induction one shows that:

1. If j is even:
$$\overline{Y^j} = g(X)^{j/2} + R_0(X,Y),$$
where $(2,d)$-weighted degree of $R_0(X,Y)$ is less than dj.

2. If j is odd:
$$\overline{Y^j} = Yg(X)^{(j-1)/2} + R_0(X,Y),$$
where $(2,d)$-weighted degree of $R_0(X,Y)$ is less than dj.

This completes the proof. □

Observe that the map $R(X,Y) \mapsto \overline{R(X,Y)}$ is $\mathbb{F}_q$-linear.

3.3 Interpolating a bivariate polynomial

Let $r = \lfloor n/2 \rfloor$. Let A, B be two integers (to be picked later).
Let $(a_{st})_{s \in S_A, t \in S_B}$ be formal variables over $\mathbb{F}_q$.
Consider the polynomial
$$\tilde{Q}(X,Y) := \sum_{s \in S_A, t \in S_B} a_{st} M_s(X,Y) M_t(X,Y)^{2^r}.$$

Its $(2,d)$-degree is at most $2^r B + A$. Thus $\overline{\tilde{Q}(X,Y)}$ is in the linear span of $\{M_u(X,Y) : u \in S_{2^r B + A}\}$. Thus the map sending $(a_{st})_{s \in S_A, t \in S_B}$ to $\overline{\tilde{Q}(X,Y)}$ is a linear map from a space of dimension $|S_A||S_B|$ to a space of dimension $|S_{2^r B + A}|$. Thus, if A, B satisfy $(A - (d-1)/2)(B - (d-1)/2) > 2^r B + A - (d-1)/2$, we know that there is a nonzero $\tilde{Q}(X,Y)$ of the above form such that $\overline{\tilde{Q}(X,Y)} = 0$ (i.e., $P(X,Y)$ divides $\tilde{Q}(X,Y)$). Take such a $\tilde{Q}(X,Y)$.
We will now see how to construct the polynomial $Q(X,Y)$ that we wanted earlier. Let
$$Q(X,Y) := \sum_{s \in S_A, t \in S_B} a_{st}^{2^{n-r}} M_s(X,Y)^{2^{n-r}} M_t(X,Y).$$

Note that the $(2,d)$-degree of $Q(X,Y)$ is at most $2^{n-r}A + B$.
Let us now check that for any $(x,y) \in V$, $Q(x,y) = 0$. The crucial observation is:
$$\begin{aligned} Q(X,Y) &= \sum_{s \in S_A, t \in S_B} a_{st}^{2^{n-r}} M_s(X,Y)^{2^{n-r}} M_t(X,Y)^{2^n} \mod \langle X^{2^n} - X, Y^{2^n} - Y \rangle \quad (2) \\ &= \tilde{Q}(X,Y)^{2^{n-r}} \mod \langle X^{2^n} - X, Y^{2^n} - Y \rangle. \quad (3) \end{aligned}$$

Take $(x,y) \in V$. As $P(x,y) = 0$ and $P(X,Y)$ divides $\tilde{Q}(X,Y)$, we conclude that $\tilde{Q}(x,y) = 0$. Furthermore, since $x, y \in \mathbb{F}_q$, we have $x^{2^n} - x = 0$ and $y^{2^n} - y = 0$. The crucial observation above now implies that $Q(x,y) = 0$.
Thus V is contained in the set of all common solutions (x,y) of $Q(x,y) = P(x,y) = 0$.
Finally, let us show that $Q(X,Y)$ is relatively prime to $P(X,Y)$. As $P(X,Y)$ is irreducible, it suffices to show that $\overline{Q(X,Y)}$ is nonzero. Note that:
$$\overline{Q(X,Y)} = \sum_{s \in S_A, t \in S_B} a_{st}^{2^{n-r}} \overline{M_s(X,Y)^{2^{n-r}} M_t(X,Y)}$$

By Claim 6, the $(2,d)$-degree of any single term $\overline{M_s(X,Y)^{2^{n-r}} M_t(X,Y)}$ is exactly $2^{n-r}s + t$. Now if $B < 2^{n-r}$, we see that all terms have distinct $(2,d)$-degrees, and hence any nonzero linear combination of them must be nonzero. Thus $\overline{Q(X,Y)} \neq 0$, and so $P(X,Y)$ is relatively prime to $Q(X,Y)$.
We may now apply Theorem 5, and conclude the number of common solutions (x,y) of $Q(x,y) = P(x,y) = 0$ is at most $2^{n-r}A + B + 2 \cdot d$. Thus $|V| \leq 2^{n-r}A + B + 2d$.
Summarizing, we showed that for any A, B satisfying

1. $A, B \geq d$,
2. $B < 2^{n-r}$,
3. $(A-(d-1)/2)(B-(d-1)/2) > (2^r B + A - (d-1)/2)$.

we have $|V| \leq 2^{n-r}A + B + 2d$.
Picking

$$A = 2^r + \frac{d-1}{2} + \frac{2^r \frac{d+1}{2}}{2^{n-r} - \frac{d+1}{2}}$$

and $B = 2^{n-r} - 1$, we get $|V| \leq 2^n + 2d2^{n-r} + 2d$, as desired.
The argument for general p is very similar.

3.4 A basic additive character sum bound

Let ψ_a be an additive character of $\mathbb{F}_q$. Applying the above upper bound to several polynomials of the form $ag(X) + \alpha$ (for suitable α), we get the following additive character sum bound:

$$\left| \sum_{x \in \mathbb{F}_q} \psi_a(g(X)) \right| \leq O(pd\sqrt{q}).$$

This bound is nontrivial only when q is much larger than p. The main step in deducing the full Weil bound via the L-function argument is to reduce to the case where q is much much bigger than p.

4 L-functions, the Riemann Hypothesis, and the Weil bounds

4.1 Multiplicative Character Sums

Let $g(T) \in \mathbb{F}_q[T]$. Let χ be a multiplicative character of $\mathbb{F}_q$ of order e. We want to bound $\left|\sum_{\alpha \in \mathbb{F}_q} \chi(g(\alpha))\right|$.

4.1.1 The associated L-function

Let $\Delta_\chi : \mathbb{F}_q[T] \to \mathbb{C}$ be given by:

$$\Delta_\chi(F(T)) = \chi(\prod_{i=1}^{\deg(F)} g(\alpha_i)),$$

where $F(T) = c \cdot \prod_{i=1}^{\deg(F)} (T - \alpha_i)$ and $\alpha_i \in \overline{\mathbb{F}}_q$. Note that (1) $\prod_{i=1}^{\deg(F)} g(\alpha_i)$ lies in $\mathbb{F}_q$, and so Δ_χ is properly defined, and (2) Δ_χ is multiplicative.
We will now see that Δ_χ as defined above is in fact a Dirichlet character. Let $g(T) = \prod_{j=1}^{k} (T - \beta_j)^{c_j}$, with

the β_j distinct. Then:

$$\begin{aligned}\Delta_\chi(F(T)) &= \chi(\prod_{i=1}^{\deg(F)} g(\alpha_i)) \\ &= \chi(\prod_{i=1}^{\deg(F)} \prod_{j=1}^{k} (\alpha_i - \beta_j)^{c_j}) \\ &= \chi((-1)^{k\deg(F)} \prod_{j=1}^{k} \prod_{i=1}^{\deg(F)} (\beta_j - \alpha_i)^{c_j}) \\ &= \chi((-1)^{k\deg(F)} \prod_{j=1}^{k} F(\beta_j)^{c_j}),\end{aligned}$$

which only depends on the values of F on the β_j, and hence only on $F \mod \tilde{g}(T)$, where $\tilde{g}(T) = \prod_{j=1}^{k}(T-\beta_j)$ is the squarefree part of $g(T)$. Thus Δ_χ is a Dirichlet character mod $\tilde{g}(T)$.
Consider the L-function $L(s, \Delta_\chi)$:

$$L(s, \Delta_\chi) = \sum_{F(T) \text{ monic}} \frac{\Delta_\chi(F)}{|F|^s},$$

where $|F| = q^{\deg(F)}$. By multiplicativity $L(s, \Delta_\chi)$ has an Euler factorization:

$$L(s, \Delta_\chi) = \prod_{P(T) \text{ monic irreducible}} \left(1 - \frac{\Delta_\chi(P)}{|P|^s}\right)^{-1}.$$

By a result from an earlier lecture, this L function is in fact a polynomial in q^{-s} of degree at most $\deg(\tilde{g})-1 = k-1$. Let

$$L(s, \Delta_\chi) = 1 + c_1 q^{-s} + \ldots + c_{k-1} q^{-(k-1)s}.$$

Note that

$$c_1 = \sum_{F(T) \text{ monic, degree } 1} \Delta_\chi(F) = \sum_{\alpha \in \mathbb{F}_q} \chi(g(\alpha)),$$

which is exactly the character sum that we are interested in!
Since $L(s, \Delta_\chi)$ is a polynomial, we can factorize $L(s, \Delta_\chi)$ as:

$$L(s, \Delta_\chi) = \prod_{j=1}^{t} (1 - \gamma_j q^{-s}).$$

Consequently, we have:

$$\sum_{\alpha \in \mathbb{F}_q} \chi(g(\alpha)) = c_1 = -\sum_{j=1}^{t} \gamma_j.$$

The Riemann Hypothesis for this L function, which we will prove shortly, states that all the zeroes of $L(s, \Delta_\chi)$ have $\Re(s) \leq \frac{1}{2}$. This is equivalent to the statement that $|\gamma_j| \leq \sqrt{q}$ for each j, and it is in this form that we will prove the Riemann Hypothesis.

4.1.2 Lifting to $\mathbb{F}_{q^n}$

We now consider a lifting of all the above to $\mathbb{F}_{q^n}$. Define $\chi_n : \mathbb{F}_{q^n} \to \mathbb{C}$ by:

$$\chi_n(\alpha) = \chi(\mathsf{Norm}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha)).$$

Observe that χ_n is a multiplicative character of $\mathbb{F}_{q^n}$ of order e.
Define $\Delta_{\chi_n} : \mathbb{F}_{q^n}[T] \to \mathbb{C}$ by:

$$\Delta_\chi(F(T)) = \chi(\prod_{i=1}^{\deg(F)} g(\alpha_i)),$$

where $F(T) = c \cdot \prod_{i=1}^{\deg(F)}(T - \alpha_i)$, and $\alpha_i \in \overline{\mathbb{F}}_q$. Again, Δ_{chi_n} is well-defined, multiplicative, and a Dirichlet character mod $\tilde{g}(T)$.
Consider the L-function $L_n(s, \Delta_{\chi_n})$:

$$L_n(s, \Delta_{\chi_n}) = \sum_{F(T) \text{ monic}} \frac{\Delta_{\chi_n}(F)}{|F|_n^s},$$

where $|F|_n = q^{n \deg(F)}$. We have that $L_n(s, \Delta_{\chi_n})$ is a polynomial in q^{-ns} of degree at most $k-1$. Let

$$L_n(s, \Delta_{\chi_n}) = c_{0,n} + c_{1,n}q^{-ns} + \ldots + c_{e-1,n}q^{-n(e-1)s}.$$

Note that

$$c_{1,n} = \sum_{\alpha \in \mathbb{F}_{q^n}} \chi_n(g(\alpha)). \tag{4}$$

In the next section, we prove that $L(s, \Delta_\chi)$ determines $L_n(s, \Delta_{\chi_n})$ via the formula:

$$L_n(s, \Delta_{\chi_n}) = \prod_{j=0}^{n-1} L(s + \frac{2\pi ij}{n \log q}, \Delta_\chi).$$

Consequently, we have:

$$\begin{aligned} L_n(s, \Delta_{\chi_n}) &= \prod_{j=0}^{n-1} \prod_{j'=1}^{t} \left(1 - \gamma_{j'} q^{-\left(s + \frac{2\pi ij}{n \log q}\right)}\right) \\ &= \prod_{j=0}^{n-1} \prod_{j'=1}^{t} \left(1 - \omega_n^j \gamma_{j'} q^{-s}\right) \\ &= \prod_{j'=1}^{t} \left(1 - \gamma_{j'}^n q^{-ns}\right). \end{aligned}$$

Thus:

$$c_{1,n} = -\sum_{j'=1}^{t} \gamma_{j'}^n. \tag{5}$$

4.1.3 Completing the proof of the Riemann Hypothesis

By Equation (4),

$$\sum_{j'=1}^{t} \gamma_{j'}^n = -\sum_{\alpha \in \mathbb{F}_{q^n}} \chi_n(g(\alpha)).$$

We can bound the right hand side of this equation using Equation (1), to get:

$$\left| \sum_{j'=1}^{t} \gamma_{j'}^n \right| \leq O(e^3 d \sqrt{q^n}).$$

The crucial thing to notice now is that we are applying our basic multiplicative character sum bound over a very large field $\mathbb{F}_{q^n}$, while keeping the parameter e, d fixed (this is when the basic bound is the strongest). We are now in the situation covered by the following lemma.

Lemma 7. *Suppose $\lambda_1, \ldots, \lambda_t \in \mathbb{C}$ and $a, b > 0$ are such that for all $n > 0$:*

$$\left| \sum_{j=1}^{t} \lambda_j^n \right| \leq a \cdot b^n.$$

Then $|\lambda_j| \leq b$ for each j.

Proof. For all $\epsilon > 0$, there exist infinitely many n such that for all j, $\arg(\lambda_j^n) \in (-\epsilon, \epsilon)$. For each such n, we have $\Re(\lambda_j^n) \geq |\lambda_j|^n \cos(\epsilon)$ for all j, and thus

$$|\sum_{j=1}^{t} \lambda_j^n| \geq \sum_{j=1}^{t} |\lambda_j|^n \cos(\epsilon)$$

for infinitely many n. This implies the result. □

Thus $|\gamma_j| \leq \sqrt{q}$ for all j, and the Riemann Hypothesis for $L(s, \Delta_\chi)$ follows.
Returning to our original character sum, we get:

$$\left| \sum_{\alpha \in \mathbb{F}_q} \chi(g(\alpha)) \right| = |c_1| = |\sum_{j=1}^{t} \gamma_j| \leq \sum_{j=1}^{t} |\gamma_j| \leq t\sqrt{q} \leq (k-1)\sqrt{q}.$$

4.1.4 Relating $L(s, \Delta_\chi)$ and $L_n(s, \Delta_{\chi_n})$

Lemma 8.

$$L_n(s, \Delta_{\chi_n}) = \prod_{j=0}^{n-1} L(s + \frac{2\pi i j}{n \log q}, \Delta_\chi).$$

Proof. The plan is to write both sides as Euler products and compare terms. Let $\sigma : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$ be the Frobenius automorphism $a \mapsto a^q$.
Let $P_0(T)$ be a monic irreducible polynomial in $\mathbb{F}_{q^n}[T]$ of degree ℓ. Let $P_0(T), \ldots, P_{k-1}(T)$ be the distinct Galois conjugates of $P_0(T)$ obtained as follows: $P_j(T)$ is obtained by replacing each coefficient a of $P_1(T)$ with $\sigma^j(a)$ (and $P_k(T) = P_0(T)$).
If S_j is the set of roots of $P_0(T)$, then $S_j = \sigma^j(S_0)$. Also note that $\prod_{j=0}^{k} P_j(T)$ is a monic irreducible polynomial in $\mathbb{F}_q[T]$ of degree $\ell \cdot k$.
Let $\alpha \in \overline{\mathbb{F}}_q$ be a root of $P_0(T)$. The set of roots of $P_j(T)$ is $\{\sigma^j(\alpha), \sigma^{j+n}(\alpha), \sigma^{j+2n}(\alpha), \ldots, \sigma^{j+(\ell-1)n}(\alpha)\}$, and $\sigma^{j+\ell n}(\alpha) = \sigma^j(\alpha)$. The set of roots of $P(T)$ is $\{\alpha, \sigma(\alpha), \ldots, \sigma^{\ell \cdot k - 1}(\alpha)\}$, and $\sigma^{\ell k}(\alpha) = \alpha$. Thus ℓk divides $n\ell$, $\ell = \frac{\ell \cdot k}{GCD(n, \ell \cdot k)}$ and $GCD(n, \ell \cdot k) = k$.

Now consider the following factors of $L_n(s, \Delta_{\chi_n})$:

$$\begin{aligned}
\prod_{j=0}^{k-1}\left(1-\frac{\Delta_{\chi_n}(P_j(T))}{|P_j|_n^s}\right)^{-1} &= \prod_{j=0}^{k-1}\left(1-\frac{\chi(\mathsf{Norm}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\prod_{j'=0}^{\ell-1} g(\sigma^{j+nj'}(\alpha))))}{|P_j|_n^s}\right)^{-1} \\
&= \prod_{j=0}^{k-1}\left(1-\frac{\chi(\prod_{j''=0}^{n-1}\sigma^{j''}(\prod_{j'=0}^{\ell-1} g(\sigma^{j+nj'}(\alpha))))}{|P_j|_n^s}\right)^{-1} \\
&= \prod_{j=0}^{k-1}\left(1-\frac{\chi(\prod_{j''=0}^{n-1}\prod_{j'=0}^{\ell-1} g(\sigma^{j''+j+nj'}(\alpha)))}{|P_j|_n^s}\right)^{-1} \\
&= \prod_{j=0}^{k-1}\left(1-\frac{\chi(\prod_{j''=0}^{n\ell-1} g(\sigma^{j''+j}(\alpha)))}{|P_j|_n^s}\right)^{-1} \\
&= \prod_{j=0}^{k-1}\left(1-\frac{\chi(\prod_{j''=0}^{n\ell-1} g(\sigma^{j''}(\alpha)))}{|P_j|_n^s}\right)^{-1} \\
&= \prod_{j=0}^{k-1}\left(1-\frac{\chi(\prod_{j''=0}^{k\ell-1} g(\sigma^{j''}(\alpha)))^{\frac{n}{k}}}{|P_j|_n^s}\right)^{-1} \\
&= \prod_{j=0}^{k-1}\left(1-\frac{\chi(\prod_{j''=0}^{k\ell-1} g(\sigma^{j''}(\alpha)))^{\frac{n}{k}}}{q^{n\ell s}}\right)^{-1} \\
&= \prod_{j=0}^{k-1}\left(1-\frac{\Delta_\chi(P(T))^{\frac{n}{k}}}{q^{n\ell s}}\right)^{-1} \\
&= \left(1-\frac{\Delta_\chi(P(T))^{\frac{n}{k}}}{q^{n\ell s}}\right)^{-k} \\
&= \left(1-\frac{\Delta_\chi(P(T))^{\frac{n}{k}}}{(q^{\ell k s})^{\frac{n}{k}}}\right)^{-k} \\
&= \prod_{j=0}^{n-1}\left(1-\omega_n^{\ell\cdot k\cdot j}\frac{\Delta_\chi(P(T))}{q^{\ell k s}}\right)^{-1} \\
&= \prod_{j=0}^{n-1}\left(1-\frac{\Delta_\chi(P(T))}{q^{\ell k s+\frac{2\pi i \ell k j}{n\log q}}}\right)^{-1} \\
&= \prod_{j=0}^{n-1}\left(1-\frac{\Delta_\chi(P(T))}{q^{\ell k\left(s+\frac{2\pi i j}{n\log q}\right)}}\right)^{-1} \\
&= \prod_{j=0}^{n-1}\left(1-\frac{\Delta_\chi(P(T))}{|P|^{\left(s+\frac{2\pi i j}{n\log q}\right)}}\right)^{-1}.
\end{aligned}$$

This proves the theorem: it gives us a way of identifying products of finite number of terms in the Euler product for the LHS with finite number of terms from the RHS. □

4.2 Additive character sums

Coming soon (hopefully).

Until then ... exercise!
The key differences:

1. Define $\Delta_\psi(F(T)) = \psi(\sum_{i=1}^{\deg(F)} g(\alpha_i))$.
2. Δ_ψ is not a Dirichlet character, but we still have $L(s, \Delta_\psi)$ being a low-degree polynomial in q^{-s}. This can be proved by using the Newton identities.
3. Define $\psi_n : \mathbb{F}_{q^n}[T]$ by $\psi_n(\alpha) = \psi(\mathsf{Tr}_{\mathbb{F}_{q^n}/\mathbb{F}_q}(\alpha))$.