# INTRODUCTION TO WEIL CONJECTURES, PART I: CURVES

JORI MERIKOSKI

ABSTRACT. This note is the first part of two of an introduction to the Weil conjectures and related exponential sum estimates. In this note I restrict to the case of curves. No background in algebraic geometry is assumed from the reader; the relevant notions from algebraic geometry are introduced as we go along.

The Weil conjectures concern the number of solutions to polynomial equations in finite fields. These have applications in other parts of number theory via exponential sum estimates; one can even consider the Weil conjectures as a far-reaching generalization of the theory of Gauss sums. I am by no means an expert in this subject, and one of the purposes of writing this note is to increase my own understanding. Another reason is to prepare the reader for more advanced texts on the topic, as much of the existing literature assumes already some background in algebraic geometry from the reader. At the end there are some further reading suggestions for the interested reader.

## CONTENTS



## 1. INTRODUCTION

The Weil conjectures (a theorem despite its name) concern the number of solutions to polynomial equations over finite fields. In this first part we will focus on the case of curves. To motivate the topic, consider the following basic problem: for any ring $R$, define

$$R[X,Y] := \left\{ \sum_{i,j\geq 0} a_{ij} X^i Y^j : \, a_{ij} \in R, \, a_{ij} = 0 \text{ for } i,j \gg 1 \right\},$$

the polynomial ring in two variables with coefficients in $R$. Let $p$ be a prime number and $\mathbb{F}_p = \{0,1,2,\ldots,p-1\}$ denote the finite field with $p$ elements. Fix a non-constant polynomial $f(X,Y) \in \mathbb{F}_p[X,Y]$. What can we then say about the set of solutions to the equation $f(x,y) = 0$ with $x,y \in \mathbb{F}_p$? In particular, what can we say about the number of solutions

$$N(f) := |\{(x,y) \in \mathbb{F}_p \times \mathbb{F}_p : \, f(x,y) = 0\}| \, .$$

As a trivial example, if $f(X,Y) = Y - X$, then clearly $N(f) = p$. In general, under some conditions on $f$, we might guess that $N(f) \sim p$ (we need some assumptions here since, for example, for $f = Y^2 - X^2 = (Y-X)(Y+X)$ we have $N(f) = 2p-1$). Why is this an interesting problem? We give two motivating problems:

**(1) Solutions to diophantine equations.** Let $f(X,Y) \in \mathbb{Z}[X,Y]$. Does there exist $x, y \in \mathbb{Z}$ such that $f(x,y) = 0$, and if so, what can be said about these solutions? A common strategy in studying this problem is to consider the reduction of $f$ modulo primes $p$; by taking the residue class of each coefficient of $f$ modulo $p$, we can consider $f$ as an element of $\mathbb{F}_p[X,Y]$. Suppose that for all primes $p$ we can find find a solution to $f(x,y) = 0 \,(\mathrm{mod}\, p)$. Then in some situations we can 'glue' these 'local solutions' modulo $p$ into a 'global solution' in $\mathbb{Z}$, by using the Chinese Remainder Theorem as follows: for large integers $n$, we expect that $\mathbb{Z}/n\mathbb{Z}$ approximates $\mathbb{Z}$ in some sense. If $n = p_1 \cdots p_k$ for some distinct primes $p_1, \ldots, p_k$, then the Chinese Remaider Theorem states that we have a ring isomorphism

$$\mathbb{Z}/n\mathbb{Z} \cong (\mathbb{Z}/p_1\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p_k\mathbb{Z})\,.$$

Then given a solution $(x_i, y_i)$ modulo $p_i$ for each $i = 1, \ldots, k$, we obtain a solution modulo $n$ by using the above ring isomorphism. This strategy is known as 'Hasse's principle' or 'local-to-global principle.' It can be made to work rigorously in some cases (e.g. for quadratic forms which have a root in $\mathbb{R}$). In any case, there is a converse to this, that is, if there exists a prime $p$ such that $f(x,y) = 0 \,(\mathrm{mod}\, p)$ has no solutions, then clearly there cannot be a solution in $\mathbb{Z}$.

**(2) Exponential sums.** In number theory one is often faced with exponential sums which have some algebraic structure. Consider the following: Let $f_1, f_2, g_1, g_2 \in \mathbb{F}_p[X]$, with $g_1, g_2 \neq 0$, and denote $U := \{x \in \mathbb{F}_p : f_1(x)g_1(x)g_2(x) \neq 0.\}$ Fix a Dirichlet character $\chi$ modulo $p$ (that is, a group homomorphism of multiplicative groups $\chi : \mathbb{F}_p^\times \to \mathbb{C}^\times$). We then want to understand the sums

$$\sum_{x \in U} \chi\left(\frac{f_1(x)}{g_1(x)}\right) e_p\left(\frac{f_2(x)}{g_2(x)}\right),$$

where $e_p(z) := e^{2\pi i z/p}$ (an additive character). Especially, one is often interested in showing that there is square-root cancellation in such sums, that is, a bound of the form $\ll \sqrt{p}$. A classical example of such sums is the Gauss sums

$$\sum_{x \in \mathbb{F}_p^\times} \chi(x) e_p(ax),$$

whose modulus is always exactly $\sqrt{p}$ if $a \in \mathbb{F}_p^\times$ and $\chi$ is non-trivial. One way to interpret the Weil conjectures is as a far-reaching generalization of the theory of Gauss sums.

In a very general situation (with some assumptions on $f_i, g_i$), there is a way to relate these sums to counting solutions to certain polynomial equations modulo $p$. The general case is quite involved, but to give a simple example of the argument, consider the case where the Dirichlet character is the Legendre symbol

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

Then

$$\sum_{x\in\mathbb{F}_p^\times}\left(\frac{x}{p}\right)=\sum_{x\in\mathbb{F}_p}\left(1+\left(\frac{x}{p}\right)\right)-p=\left|\left\{(x,y)\in\mathbb{F}_p\times\mathbb{F}_p:\ x=y^2\right\}\right|-p.$$

Although this is rather silly as the sum on the left hand side is trivially equal to 0, this example gives a flavour of the general argument.

As an example of a bad case consider an exponential sum over $\chi(f(x))$, where $f(x)=x^d$ and $\chi$ is a character of order dividing $d$, that is, $\chi^d\equiv 1$; then clearly there is no cancellation in the sum.

## 2. Algebraic geometry: bare necessities

To state the Weil conjecture of curves, we need to discuss some basic notions in algebraic geometry. We first recall some basic facts about field extensions and algebraically closed fields; the realm of classical algebraic geometry is in the solutions to polynomial equations over algebraically closed fields. This is best illustrated by the so-called Hilbert's Nullstellensatz ('theorem on the location of zeros'), which we will describe in the second section. Lastly, we need to discuss projective spaces and singular points of curves.

2.1. **Fields.** Recall that the characteristic of a field $k$ is defined to be $\min\{n\geq 1 : nx=0\ \forall x\in k\}$ if such an $n$ exists, and it is set to be 0 if no such $n$ exists. It is easy to show that for any field, the characteristic is always either 0 or some prime number.

Recall that a field extension $k\subseteq K$ is said to be algebraic if for any given $x\in K$ there exists a polynomial $f\in k[X]$ such that $f(x)=0$.

An algebraic field extension $k\subseteq K$ is said to be finite if $K$ is obtained from $k$ by adjoining a finite number of elements of $K$ (equivalently, $K$ is a finite $k$-vector space).

A field $K$ is said to be algebraically closed if for any $f\in K[X]$ there exists $x\in K$ such that $f(x)=0$. This is (by induction) equivalent to saying that every polynomial $f\in K[X]$ factors into a product of linear factors, and also equivalent to saying that $K$ admits no non-trivial algebraic extensions.

An algebraic closure of $k$ is an algebraic field extension $k\subseteq\overline{k}$ such that $\overline{k}$ is algebraically closed.

**Theorem 1.** *For every field $k$, there exists an algebraic closure $\overline{k}$. It is unique up to isomorphism.*

*Example.* Algebraic closure of $\mathbb{R}$ is $\mathbb{C}$ (this is the fundamental theorem of algebra).

*Example.* For a finite field $\mathbb{F}_p$, any finite algebraic extension is a field with $q=p^m$ elements for some integer $m\geq 1$. These are unique up to isomorphism, so we may speak of the finite field with $q$ elements, denoted as $\mathbb{F}_q$. The field $\mathbb{F}_q$ is the splitting field of the polynomial $X^q-X$, that is, the smallest algebraic extension of $\mathbb{F}_p$ such that $X^q-X$ splits in to linear factors. In fact we have

$$X^q-X=\prod_{a\in\mathbb{F}_q}(X-a).$$

In other words, $\mathbb{F}_q$ is obtained from $\mathbb{F}_p$ by adjoining the roots to the equation $X^q-X=0$.

We have $\mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^n}$ if and only if $m|n$. In particular,

$$\mathbb{F}_p \subseteq \mathbb{F}_{p^2} \subseteq \mathbb{F}_{p^{2\cdot 3}} \subseteq \mathbb{F}_{p^{4!}} \subseteq \cdots,$$

which lets us to construct an algebraic closure for $\mathbb{F}_p$ by setting

$$\overline{\mathbb{F}}_p = \bigcup_{m \geq 1} \mathbb{F}_{p^{m!}}.$$

## 2.2. **Hilbert's Nullstellensatz.**

**Theorem 2. (Hilbert's Nullstellensatz).** *Let $k$ be any field. If $\mathfrak{m}$ is a maximal ideal of the polynomial ring $k[X_1, \ldots, X_n]$, then the residue field*

$$k(\mathfrak{m}) := k[X_1, \ldots, X_n]/\mathfrak{m}$$

*is a finite extension of $k$.*

Recall that an ideal $I \subseteq R$ of a ring $R$ is said to be maximal if $I \neq R$ and for any ideal $I \subseteq J$ we have either $J = I$ or $J = R$. Equivalently, an ideal $I$ is maximal if the quotient ring $R/I$ is a field (exercise). Hilbert's Nullstellensatz has a myriad of different formulations in the literature; we have chosen the above since it is particularly well suited for our purposes.

The reason algebraic geometers love algebraically closed fields is that in that case Hilbert's Nullstellensatz implies

**Theorem 3.** *If $k$ is an algebraically closed field, then every maximal ideal of the polynomial ring $k[X_1, \ldots, X_n]$ is of the form $(X_1 - c_1, \ldots, X_n - c_n)$ for some constants $c_1, \ldots, c_n \in k$.*

*Exercise* 1. A field $k$ is algebraically closed if and only if every maximal ideal of $k[X]$ is of the form $(X - c)$.

To put it geometrically, Hilbert's Nullstellensatz says that if $k$ is an algebraically closed field, then there is a one-to-one correspondence between the maximal ideals of $k[X_1, \ldots, X_n]$ and the points of the affine space

$$\mathbb{A}^n_k := \{(c_1, \ldots, c_n) : c_1, \ldots, c_n \in k\}.$$

This establishes a dictionary between two fields of mathematics, commutative algebra and geometry. This interpretation of the Nullstellensatz generalizes as follows:

**Theorem 4.** *Let $k$ be an algebraically closed field, and let $\{f_1, \ldots, f_m\}$ be a finite set of polynomials from $k[X_1, \ldots, X_n]$. Define the 'affine algebraic variety defined by the polynomials $f_j$' as the set of common zeros*

$$V := \{c \in \mathbb{A}^n_k : f_j(c) = 0 \text{ for all } j = 1, 2, \ldots, m\}.$$

*Let $\mathcal{I}(V) \subseteq k[X_1, \ldots, X_n]$ be the ideal of polynomials $g$ such that $g(c) = 0$ for all $c \in V$ (this is called the ideal of definition of $V$, note that $f_j \in \mathcal{I}(V)$).*

*Then there is a one-to-one correspondence between the points of $V$ and the maximal ideals of the quotient ring (called the coordinate ring of $V$)*

$$\mathcal{P}(V) := k[X_1, \ldots, X_n]/\mathcal{I}(V),$$

*which is given by the map*

$$x \mapsto \mathfrak{m}_x := \{g \in \mathcal{P}(V) : \, g(x) = 0\}.$$

*Exercise* 2. Show that $\mathfrak{m}_x$ is a maximal ideal for any $x \in V$.

It is relatively easy to show that the map $x \mapsto \mathfrak{m}_x$ is injective; the hard part, which requires the Nullstellensatz, is to show that this map is surjective.

*Example.* Let $k = \mathbb{C}$, and consider the curve $V$ defined by $f(X, Y) = Y - X^2$, that is,

$$V := \left\{(x, y) \in \mathbb{C}^2 : y = x^2\right\}.$$

Then $\mathcal{I}(V) = (Y - X^2)$ so that the coordinate ring is $\mathcal{P}(V) = \mathbb{C}[X, Y]/(Y - X^2)$. The coordinate ring can be interpreted as the ring of polynomial functions on the curve $Y = X^2$; if $g, h \in \mathbb{C}[X, Y]$ are such that $g - h \in \mathcal{I}(V)$, then this means exactly that $g$ and $h$ agree on the curve $Y = X^2$. We can also compute the coordinate ring by substituting $Y = X^2$:

$$\mathcal{P}(V) = \mathbb{C}[X, Y]/(Y - X^2) \cong \mathbb{C}[X, X^2] = \mathbb{C}[X].$$

Clearly now the maximal ideal $(X - x) \subseteq \mathbb{C}[X]$ corresponds to the point $(x, x^2) \in V$.

In general, the ideal of definition is not necessarily the ideal $(f_1, \ldots, f_m)$ generated by the defining polynomials. It can be shown that $\mathcal{I}(V)$ is the radical ideal of $(f_1, \ldots, f_m)$, that is, the ideal consisting of polynomials $g$ such that for some positive $k$ we have $g^k \in (f_1, \ldots, f_m)$.

Modern algebraic geometry generalizes this notion even further by using the language of schemes; for any commutative ring $R$, we may define

$$\operatorname{Specm}(R) := \{\mathfrak{m} \subseteq R \text{ a maximal ideal}\} \quad \text{(the maximal spectrum)}$$
$$\operatorname{Spec}(R) := \{\mathfrak{p} \subseteq R \text{ a prime ideal}\} \quad \text{(the prime spectrum)}.$$

The idea then is to treat the prime spectrum of any commutative ring as a geometric object, even if the ring is not a coordinate ring of any variety over a field. This makes it possible to draw parallels between algebraic geometry and algebraic number theory. As an example, we have

$$\operatorname{Spec}(\mathbb{Z}) = \{(0)\} \cup \operatorname{Specm}(\mathbb{Z}) = \{(0)\} \cup \{(p) : p \in \mathbb{Z} \text{ a prime number}\}.$$

The prime spectrum of a ring (equipped with its so-called 'Zariski topology' and the 'structure sheaf') is referred to as an affine scheme. A general scheme is then an object which is locally an affine scheme (that is, an object which is made by gluing together affine schemes).

2.3. **Projective spaces.** The most natural framework for practising algebraic geometry is in projective spaces. To motivate why this might be so, consider two distinct complex lines $L_1, L_2 \subset \mathbb{C}^2$. Then there are two possibilities: either the lines are parallel, in which case $L_1 \cap L_2 = \emptyset$, or they are not parallel, in which case they meet at a single point in $\mathbb{C}^2$. This situation is remedied by agreeing that parallel lines meet at a point at infinity. The projective plane $\mathbb{P}^2_{\mathbb{C}}$ can be thought of as $\mathbb{C}^2$ with additional points at infinity (one for each set of parallel lines).

We construct the projective spaces as follows: Let $k$ be a field and define an equivalence relation $\sim$ on $k^{n+1} \setminus \{0\}$ by

$$(a_0, \ldots, a_n) \sim (b_0, \ldots, b_n)$$

if there exists some $\lambda \in k \setminus \{0\}$ such that

$$(a_0, \ldots, a_n) = (\lambda b_0, \ldots, \lambda b_n).$$

Then as a set we define the $n$-dimensional projective space over $k$ as the set of equivalence classes

$$\mathbb{P}_k^n := (k^{n+1} \setminus \{0\})/\sim .$$

We are mainly interested in curves in this note. An affine curve in $\mathbb{A}_k^2$ is the zero set of an absolutely irreducible polynomial $f(X,Y) \in k[X,Y]$, that is, a polynomial $f$ which is irreducible as an element of $\overline{k}[X,Y]$. For example, the zero set of $f = Y^2 - X^2$ is not a curve, but a union of two curves $Y - X$ and $Y + X$. If, for instance, $k = \mathbb{F}_p$ and $a \in \mathbb{F}_p$ is not a square, then the polynomial $f = Y^2 - aX^2$ is irreducible but not absolutely irreducible, so that it does not define a curve.

A projective curve of degree $n$ in $\mathbb{P}_k^2$ is the zero set of an absolutely irreducible homogeneous polynomial

$$f(X,Y,Z) = \sum_{i+j+l=n} a_{ijl} X^i Y^j Z^l.$$

Since $f$ is homogeneous, we have $f(\lambda X, \lambda Y, \lambda Z) = \lambda^n f(X,Y,Z)$ so that the zero set modulo the equivalence relation $\sim$ is well-defined. Any affine curve can be made into a projective curve by the process of homogenization:

*Example.* Let $f(X,Y) = Y^2 - X^3 - 1$. Then the homogenization of $f(X,Y)$ is given by $f(X,Y,Z) = Y^2Z - X^3 - Z^3$, that is, a homogeneous polynomial such that $f(X,Y) = f(X,Y,1)$. The projective curve given by $f$ is then just the set of equivalence classes defined by

$$\big\{(x,y,1) \in k^3 : f(x,y) = 0\big\} \cup \{(0,1,0)\} \subset \mathbb{P}_k^2,$$

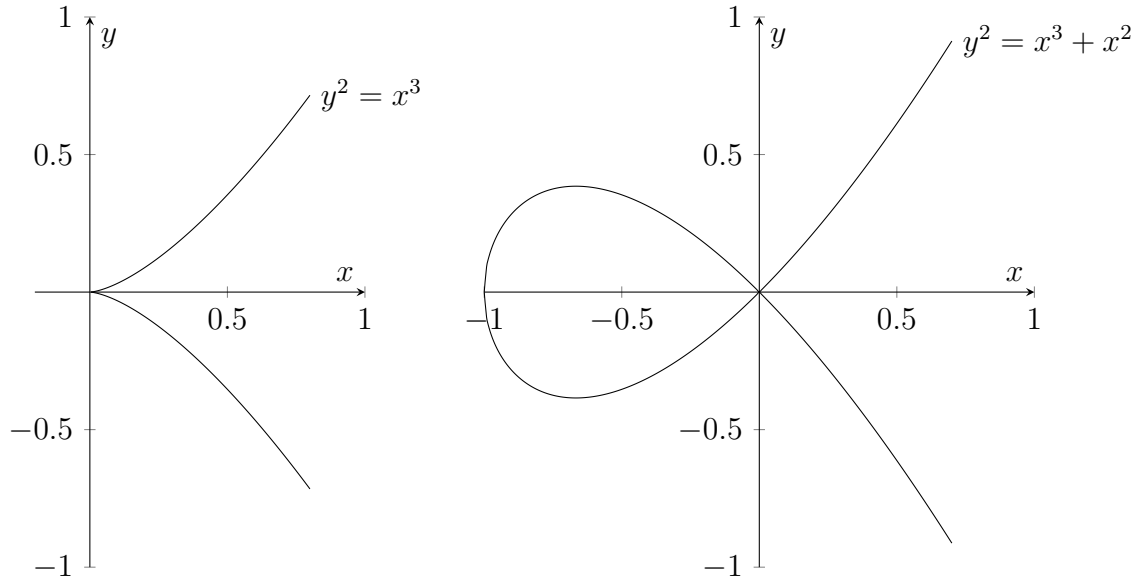
that is, there is just one point (the point at infinity) added to the set of affine points of the curve.

There is no better justification for the use of projective spaces than the classical theorem of Bézout on the number of intersection points of two curves.

**Theorem 5. (Bézout's Theorem)** *Let $k$ be an algebraically closed field, and let $V$ and $W$ be two distinct curves in $\mathbb{P}_k^2$, of degrees $d$ and $e$. Then the number of intersection points $|V \cap W|$, counted with multiplicities, is equal to $de$.*

We will not here specify what is meant by multiplicities, other than that it generalizes the notion of the multiplicity of a root of a polynomial. If $k = \mathbb{C}$ and we consider the curves defined by (homogenization of) the polynomial $f(X,Y) = Y - h(X)$, where $h$ is of degree $d$, and the polynomial $g(X,Y) = Y$, then we see that Bézout's Theorem is a generalization of the fundamental theorem of algebra. Bézout's theorem is remarkable in that just by agreeing that any two distinct lines always meet at a point (case $d = e = 1$), we get the expected number of intersection points for curves of any degrees.

2.4. **Non-singular curves.** To state the Weil conjecture in the case of curves, we need the notion of a non-singular curve. As the name suggests, this means that the curve is well behaved at every point. If $k = \mathbb{C}$, then a singular point could for example be a cusp or a double point (see the figure below).




(A) Cusp at (0,0) (B) Double point at (0,0)

We will not here give the general definition of a singular point. For an affine curve defined by $f(X,Y) = 0$, a point $(x,y)$ on the curve is singular if both of the partial derivatives vanish at that point, that is,

$$f(x,y) = \frac{\partial}{\partial x} f(x,y) = \frac{\partial}{\partial y} f(x,y) = 0$$

There is a completely algebraic way of defining a singular point, which generalizes to curves in general: a curve is non-singular if the tangent space at every point is one dimensional. For example, in the case of the curve $Y^2 = X^3 + X^2$, the tangent space at (0,0) is two dimensional, as the picture might suggest. This notion generalizes to higher dimensional varieties.

Similarly as with singularities, there is an algebraic way of defining a dimension of a variety. Curves are then varieties of dimension one. By a non-singular projective curve over a field $k$, we will mean a non-singular absolutely irreducible variety of dimension one in $\mathbb{P}^n_k$ for some $n \geq 1$. An example of this is the zero set of (homogenization of) an absolutely irreducible polynomial $f(X,Y)$ in $\mathbb{P}^2_k$, such that at all points on the curve either $\partial_x f \neq 0$ or $\partial_y f \neq 0$. Note that since $f$ is a polynomial, the derivatives are well-defined even if $k \neq \mathbb{C}$.

## 3. Statement of the Weil conjecture for curves

Throughout this section we fix a prime $p$ and a power of it $q = p^m$. We will consider the algebraic geometry of curves over the finite field $\mathbb{F}_q$.

### 3.1. **Riemann hypothesis for curves.**

We begin by stating perhaps the most important part of the Weil conjecture for curves, namely, the Riemann hypothesis for curves:

**Theorem 6. (Riemann hypothesis for curves).** *Let $\mathcal{C}$ be a non-singular projective curve over a finite field $\mathbb{F}_q$. Then*

$$||\mathcal{C}|-(q+1)| \leq 2g\sqrt{q},$$

*where $g = g(\mathcal{C})$ is the genus of the curve $\mathcal{C}$, which is a numerical invariant (a non-negative integer) depending only on the curve.*

*Remark* 1. If the curve is defined by a polynomial of degree $d$, then the genus satisfies $g \leq (d-1)(d-2)/2$. One can think of the genus as a measure of how wild the behaviour of the curve is.

*Remark* 2. Since we are working with projective curves, the expected number of points is $|\mathbb{P}^1_q| = q+1$. The genus of the projective line is 0.

*Example.* An elliptic curve $y^2 = x^3 + ax + b$ is non-singular if $4a^3 + 27b^2 \neq 0$. The genus of an elliptic curve is 1. By looking at the homogenization we see that there is exactly one point at infinity on the curve. Thus, we obtain

$$\left||\{(x,y) \in \mathbb{F}_p \times \mathbb{F}_p : y^2 = x^3 + ax + b\}| - p\right| \leq 2\sqrt{p}.$$

This bound for elliptic curves was obtained by Hasse before Weil gave a proof of the general case.

To understand why the above theorem is called the Riemann hypothesis for curves, we need to discuss the zeta function of a curve. The zeta function will also allow us to state the Weil conjectures for curves in its full form.

3.2. **Zeta function of a curve.** For any non-singular curve $\mathcal{C}$ over a finite field $\mathbb{F}_q$, we can consider $\mathcal{C}(\mathbb{F}_{q^n})$, the set of points on $\mathcal{C}$ over an extension field $\mathbb{F}_{q^n}$ of $\mathbb{F}_q$. For example, if $\mathcal{C}$ is defined by (the homogenization) a polynomial $f \in \mathbb{F}_q[X,Y]$ in $\mathbb{P}^2_{\mathbb{F}_q}$, we set

$$\mathcal{C}(\mathbb{F}_{q^n}) := \left\{ x \in \mathbb{P}^2_{\mathbb{F}_{q^n}} : f(x) = 0 \right\}.$$

Define also $\overline{\mathcal{C}} := \mathcal{C}(\overline{\mathbb{F}}_q)$, the points on $\mathcal{C}$ over the algebraic closure of $\mathbb{F}_q$; recall that $\mathbb{F}_{q^m} \subseteq \mathbb{F}_{q^n}$ iff $m|n$. Hence, $\mathcal{C}(\mathbb{F}_{q^m}) \subseteq \mathcal{C}(\mathbb{F}_{q^n})$ iff $m|n$, so that $\overline{\mathcal{C}}$ is given as the union of $\mathcal{C}(\mathbb{F}_{q^n})$.

Let $N_n := |\mathcal{C}(\mathbb{F}_{q^n})|$. Then we define the zeta function of $\mathcal{C}$ formally as

$$Z_{\mathcal{C}}(T) := \exp\left( \sum_{n=1}^{\infty} N_n \frac{T^n}{n} \right).$$

It is relatively easy to show that $N_n \ll q^n$, so that the sum converges to a complex valued function for $|T| < q^{-1}$, $T \in \mathbb{C}$.

*Exercise* 3. If $C = \mathbb{P}^1_{\mathbb{F}_q}$, show that $N_n = q^n + 1$ and

$$Z_{\mathbb{P}^1_{\mathbb{F}_q}}(T) = \exp\left( \sum_{n=1}^{\infty} (q^n+1) \frac{T^n}{n} \right) = \frac{1}{(1-T)(1-qT)}.$$

At first sight there is little resemblance of a zeta function. However, similarly as in the Nullstellensatz (in fact a projective version of Theorem 4), we can (vaguely speaking) consider the maximal ideals of the coordinate ring of $\mathcal{C}$. For any maximal ideal $\mathfrak{m}$ of the

coordinate ring, one can define the so-called residue field at $\mathfrak{m}$, denoted by $k(\mathfrak{m})$, which is always a finite extension of $\mathbb{F}_q$; write $\deg(\mathfrak{m}) = n$ if $k(\mathfrak{m}) = \mathbb{F}_{q^n}$. For a curve $\mathcal{C}$ over $\mathbb{F}_q$ (even though $\mathbb{F}_q$ is not algebraically closed), we have then the following Nullstellensatz-type result: any maximal ideal $\mathfrak{m}$ corresponds to exactly $n = \deg(\mathfrak{m})$ points on $\mathcal{C}(\mathbb{F}_{q^n})$. Furthermore, for any two distinct maximal ideals the corresponding sets of points are disjoint in $\overline{\mathcal{C}}$. In the other direction, for any point $x \in \overline{\mathcal{C}}$, if $n$ is the smallest integer such that $x \in \mathcal{C}(\mathbb{F}_{q^n})$, then there is a unique maximal ideal of degree $n$ which corresponds to a set of $n$ points containing $x$.

Combining the above discussion, we obtain (using the fact that $\mathcal{C}(\mathbb{F}_{q^m}) \subseteq \mathcal{C}(\mathbb{F}_{q^n})$ iff $m|n$)

$$N_n = \sum_{\mathfrak{m},\ \deg(\mathfrak{m})|n} \deg(\mathfrak{m})$$

Therefore, taking a logarithm of both sides of the definition of $Z_{\mathcal{C}}(T)$ gives

$$\log Z_{\mathcal{C}}(T) = \sum_{n=1}^{\infty} N_n \frac{T^n}{n} = \sum_{\mathfrak{m}} \sum_{n=1}^{\infty} \frac{T^{n\deg(\mathfrak{m})}}{n} = \sum_{\mathfrak{m}} -\log(1 - T^{\deg(\mathfrak{m})}).$$

Hence,

$$Z_{\mathcal{C}}(T) = \prod_{\mathfrak{m}} (1 - T^{\deg(\mathfrak{m})})^{-1}.$$

Letting $T = q^{-s}$, $s \in \mathbb{C}$, we define

$$\zeta_{\mathcal{C}}(s) := Z_{\mathcal{C}}(q^{-s}) = \prod_{\mathfrak{m}} (1 - |k(\mathfrak{m})|^{-s})^{-1},$$

which is gives the zeta function in the Euler product form.

*Remark* 3. If instead of a coordinate ring we consider the ring $\mathbb{Z}$, then the set of maximal ideals $\mathrm{Specm}(\mathbb{Z})$ consists of ideals $(p)$ for prime numbers $p$. The residue fields turn out to be $k((p)) = \mathbb{F}_p$. Thus, if we consider $\mathrm{Specm}(\mathbb{Z})$ to be a 'curve', we get a zeta function

$$\zeta_{\mathrm{Specm}(\mathbb{Z})}(s) = \prod_{p} (1 - |k((p))|^{-s})^{-1} = \prod_{p} (1 - p^{-s})^{-1} = \zeta(s),$$

which is precisely the Riemann zeta function. This justifies the terminology. However, there are some important differences; for instance, the zeta function of a curve $C$ over $\mathbb{F}_q$ is $2\pi i/\log q$-periodic.

We can now state the Weil Conjecture for curves.

**Theorem 7. (Weil).** *Let $\mathcal{C}$ be a non-singular projective curve over $\mathbb{F}_q$, of genus $g$. Then*

*(1) (Rationality) $Z_{\mathcal{C}}(T)$ is a rational function in $T$.*

*(2) (Functional equation) For $e = 2 - 2g$ (euler characteristic) we have*

$$Z_{\mathcal{C}}(q^{-1}T^{-1}) = q^{e/2}T^e Z_{\mathcal{C}}(T).$$

*(3) (Riemann hypothesis for curves) We have*

$$Z_{\mathcal{C}}(T) = \frac{P(T)}{(1-T)(1-qT)}$$

*with* $P(T) = \prod_{j=1}^{2g}(1-\omega_j T)$, *where* $\omega_j \in \mathbb{C}$ *satisfy* $|\omega_j| = \sqrt{q}$.

*Remark* 4. The tird part of the above theorem implies that if $\zeta_{\mathcal{C}}(s) = 0$, then $\Re\, s = 1/2$, in analoque with the usual Riemann hypothesis. Note that all of these properties are already verified for the projective line $\mathbb{P}^1_{\mathbb{F}_q}$ by the previous exercise.

As a corollary of the above theorem, we obtain a more general version of Theorem 6.

**Corollary 8.** *If* $N_n = |C(\mathbb{F}_{q^n})|$, *then*

$$|N_n - (q^n + 1)| \leq 2gq^{n/2}.$$

*Proof.* Taking the logarithmic derivative and using Taylor expansion, we find that

$$\begin{aligned}\sum_{n=1}^{\infty} N_n T^{n-1} &= \frac{d}{dT}\log Z_{\mathcal{C}}(T) \\ &= \frac{d}{dT}\left(-\log(1-T) - \log(1-qT) + \sum_{j=1}^{2g}\log(1-\omega_j T)\right) \\ &= \frac{d}{dT}\left(\sum_{n=1}^{\infty}\left(q^n + 1 - \sum_{j=1}^{2g}\omega_j^n\right)\frac{T^n}{n}\right) = \sum_{n=1}^{\infty}\left(q^n + 1 - \sum_{j=1}^{2g}\omega_j^n\right)T^{n-1}.\end{aligned}$$

Comparing the coefficients of the power series and using $|\omega_j| = \sqrt{q}$ we find that

$$|N_n - (q^n+1)| = \left|\sum_{j=1}^{2g}\omega_j^n\right| \leq 2gq^{n/2}.$$

□

*Remark* 5. It is possible to show that the above corollary is in fact equivalent to the third part of the Weil conjectures. This is analogous to the fact that the Riemann hypothesis for the Riemann zeta function is equivalent to $\pi(x) = \int_2^x \frac{dt}{\log t} + \mathcal{O}(x^{1/2+\epsilon})$ for any $\epsilon > 0$, where $\pi(x)$ is the number of primes up to $x$.

3.3. **Remarks on the proof.** Weil gave two proofs of his conjectures in the case of curves, both of which relied on his work of rewriting the foundations of algebraic geometry to accomodate fields in characteristic $> 0$. A new simpler proof using less algebraic geometry was given by Bombieri and Stepanov (aka Stepanov's method, which was generalized by Bombieri), although this proof is somewhat ad hoc.

We now briefly discuss one of Weil's proofs, which is perhaps the most geometrically intuitive. The rationality and the functional equation follow relatively easily from some standard theorems of algebraic geometry (mainly the Riemann-Roch Theorem). To prove the Riemann hypothesis for $\mathcal{C}$, we want show that

$$|N_n - (q^n+1)| \leq 2gq^{n/2}.$$

As with the other proofs, the key observation is this: define the Frobenius morphism

$$\mathrm{Fr}_q : \overline{\mathbb{F}}_q \to \overline{\mathbb{F}}_q, \quad x \mapsto x^q.$$

Then the set of fixed points of $\mathrm{Fr}_q^n : x \mapsto x^{q^n}$ is precisely $\mathbb{F}_{q^n} \subseteq \overline{\mathbb{F}}_q$, that is,

$$\left\{x \in \overline{\mathbb{F}}_q : x^{q^n} = x\right\} = \mathbb{F}_{q^n}.$$

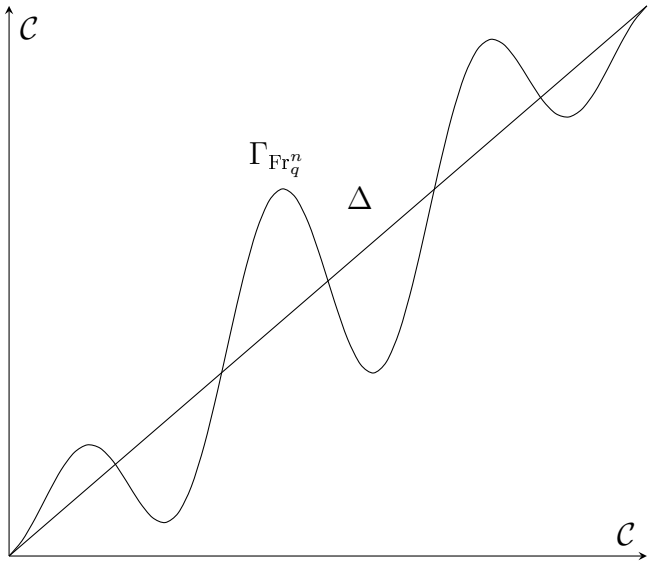



FIGURE 2. The diagonal and the graph of the Frobenius, a heuristic visualization (a similar picture is given in [11])

This is because $\mathbb{F}_{q^n}$ was obtained from $\mathbb{F}_q$ by adjoining the roots of the equation $X^{q^n} - X$. We can then define an action of the Frobenius morphism on the curve $\mathrm{Fr}_q : \overline{\mathcal{C}} \to \overline{\mathcal{C}}$ by letting the Frobenius map act on each coordinate separately.

*Exercise.* If $x, y \in \overline{\mathbb{F}}_q$, then $(x+y)^q = x^q + y^q$. Show that $\mathrm{Fr}_q(\overline{\mathcal{C}}) \subseteq \overline{\mathcal{C}}$ if $\mathcal{C}$ is defined by a polynomial $f \in \mathbb{F}_q[X, Y]$.

Hence, we obtain

$$\mathcal{C}(\mathbb{F}_{q^n}) = \left\{x \in \overline{\mathcal{C}} : \mathrm{Fr}_q^n(x) = x\right\},$$

that is, we can capture the $\mathbb{F}_{q^n}$-rational points by using the Frobenius morphism.

Now we are in position to use the machinery of algebraic geometry, since we have reduced problem on a curve over a finite field to a problem on a curve over an algebraically closed field, which is where algebraic geometry excels. Consider the product $\overline{\mathcal{C}} \times \overline{\mathcal{C}}$; in terms of algebraic geometry, this obtains the structure of a surface over $\overline{\mathbb{F}_q}$, but for us it is enough to just think of it as a cartesian product of sets. Define then two new curves which live on the surface $\overline{\mathcal{C}} \times \overline{\mathcal{C}}$, the diagonal

$$\Delta := \{(x, x) : x \in \overline{\mathcal{C}}\}$$

and the graph of the Frobenius

$$\Gamma_{\mathrm{Fr}_q^n} := \{(x, x^{q^n}) : x \in \overline{\mathcal{C}}\}.$$

Then $N_n = |\mathcal{C}(\mathbb{F}_{q^n})|$ is exactly the same as the number of intersection points of the diagonal and the graph of the Frobenius. The diagonal is a curve of degree one, while the graph of the Frobenius is a curve of degree $q^n$. Thus, by Bézout's Theorem we expect that the number of such intersection points should be $\approx q^n$; the tricky thing here is that our curves now live on a surface rather than in the projective plane. Weil's main hurdle in the proof was to create a theory of intersecions of curves on a surface (in positive

characteristic), a task too onerous for us to discuss here. Needless to say, Weil was successful in his venture, hence the theorem.

One of the key insights in Weil's proof is that algebraic geometry does not seem to care too much what the characteristic of the underlying field is, only that the field should be algebraically closed; this allows us to think about curves over $k = \overline{\mathbb{F}}_q$ as if they were geometric objects, as if $k$ were $\mathbb{C}$. This intuition is then made rigorous by writing the proofs in the language of commutative algebra ('think geometrically, prove algebraically').

## 4. Exponential sums

In this section we give two applications of Weil's bound for exponential sums. First we discuss multiplicative character sums.

**Theorem 9.** *Let $\chi$ be a non-trivial multiplicative character on $\mathbb{F}_q$ of order $d|q-1$. Let $g(X) \in \mathbb{F}_q[X]$ be a polynomial such that there is no polynomial $h(X) \in \overline{\mathbb{F}}_q[X]$ such that $g = h^d$. Let $m$ denote the number of distinct roots of $g$ in $\overline{\mathbb{F}}_q$. Then*

$$\left| \sum_{x \in \mathbb{F}_q^\times} \chi(g(x)) \right| \leq (m-1)\sqrt{q}.$$

By using averaging tricks, this problem can be reduced to obtaining a bound for

$$\left| |\{(x,y) \in \mathbb{F}_{q^n} \times \mathbb{F}_{q^n} : y^d = g(x)\}| - q^n \right|,$$

which we can do by using the Riemann hypothesis for the curve $y^d = g(x)$.

As another application, we describe Weil's bound for Kloosterman sums: for $a, b \in \mathbb{F}_p^\times$, define the Kloosterman sum as

$$S(a,b;p) := \sum_{x \in \mathbb{F}_p^\times} e\left( \frac{ax + b/x}{p} \right),$$

where the inverse $1/x$ is taken in $\mathbb{F}_p$.

**Theorem 10.** *We have*

$$|S(a,b;p)| \leq 2\sqrt{p}.$$

Again, after some averaging magic the proof is reduced to bounding

$$\left| |\{(x,y) \in \mathbb{F}_{p^n}^\times \times \mathbb{F}_{p^n} : y^p - y = ax - b/x\}| - p^n \right|,$$

which can further be reduced to counting solutions to polynomial equations.

In both of the above cases, one also needs to study the '$L$-function associated with the exponential sum', which is obtained similarly as the zeta function of a curve but in place of $N_n$ we have have an exponential sum over $\mathbb{F}_{q^n}$; these are analogous to the Diriclet $L$-functions $L(s,\chi) = \sum_{n \geq 1} \chi(n) n^{-s}$ of analytic number theory.

## Further reading suggestions and references

The zeta function for curves over finite fields was introduced in 1923 by Artin in his thesis [2],[3], where he also suggested the analogue of the Riemann hypothesis. In 1934 Hasse proved this in the case of elliptic curves [7]. The case of curves in general was solved by Weil in 1946 [13]. In 1949 Weil gave conjectures of a more general version for higher dimensional varieties over finite fields [12]; these became known as the Weil conjectures. This generalization will be the topic of the second part of these notes.

A useful book on commutative algebra (e.g. proof of Hilbert's Nullstellensatz) is *Introduction to Commutative Algebra* by Atiyah and McDonald [4]; it is a must-read if you want to study algebraic geometry. For a textbook on algebra in general (categories, rings, field extensions and homological algebra) I suggest *Algebra: Chapter 0* by Aluffi [1]; its size may be intimidating but it is big only because it contains a vast amount of examples and exercises which help to build an intuition. What is especially nice about this book is that it does not try to avoid the use category theory, which makes the book very coherent and neatly organized.

A standard textbook on algebraic geometry is *Algebraic Geometry* by Robin Hartshorne [6]. It can be quite a heavy book to read but it contains a lot of useful exercises, and if you are interested in studying the subject there is no substitute for working through the book. A new alternative for studying are the notes *Foundations of algebraic geometry* by Ravi Vakil [10]; this is especially useful as it begins with a thorough chapter on category theory, which may seem needlessly abstract at first but studying it pays off hugely in the long run.

As a softer introduction into the language of schemes, I suggest the book *The Geometry of schemes* by Eisenbud and Harris [5].

For exponential sum estimates I suggest the notes *Exponential sums over finite fields, I: elementary methods* by Kowalski [8]; it does not require background in algebraic geometry, and there is given the method of Bombieri and Stepanov.

For an exposition of Weil's proof (by intersection theory) of the Riemann hypothesis for curves (which is a guided exercise in the above-mentioned Hartshorne's *Albegraic Geometry*), I suggest the notes *Weil conjecture for curves* by Sam Raskin [9] (this requires a good knowledge of algebraic geometry).

If you have background in algebraic number theory, I suggest the book *Riemann hypothesis for function fields* by Frankenhuijsen [11]. It does not require much background at all, and it contains the proof by Bombieri and Stepanov. It also contains more discussion on the analogue between the Riemann hypothesis for curves and the usual Riemann hypothesis.

## References

[1] P. Aluffi. *Algebra: Chapter 0*, volume 104 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2009.

[2] E. Artin. Quadratische Körper im Gebiete der höheren Kongruenzen. I. *Math. Z.*, 19(1):153–206, 1924.

[3] E. Artin. Quadratische Körper im Gebiete der höheren Kongruenzen. II. *Math. Z.*, 19(1):207–246, 1924.

[4] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.

[5] D. Eisenbud and J. Harris. *The geometry of schemes*, volume 197 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.
[6] R. Hartshorne. *Algebraic geometry*. Springer-Verlag, New York-Heidelberg, 1977. Graduate Texts in Mathematics, No. 52.
[7] H. Hasse. Zur Theorie der abstrakten elliptischen Funktionenkörper III. Die Struktur des Meromorphismenrings. Die Riemannsche Vermutung. *J. Reine Angew. Math.*, 175:193–208, 1936.
[8] E. Kowalski. Exponential sums over finite fields, i: elementary methods. URL: `https://people.math.ethz.ch/~kowalski/exp-sums.pdf`.
[9] S. Raskin. Weil conjectures for curves, 2007. URL: `http://www.math.uchicago.edu/~mitya/beilinson/SamREU07.pdf`.
[10] R. Vakil. Math 216: Foundations of algebraic geometry, 2013. URL: `http://math.stanford.edu/~vakil/216blog/FOAGjun1113public.pdf`.
[11] M. van Frankenhuijsen. *The Riemann hypothesis for function fields*, volume 80 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2014. Frobenius flow and shift operators.
[12] A. Weil. Numbers of solutions of equations in finite fields. *Bull. Amer. Math. Soc.*, 55:497–508, 1949.
[13] A. Weil. *Foundations of algebraic geometry*. American Mathematical Society, Providence, R.I., 1962.

Department of Mathematics and Statistics, University of Turku, FI-20014 University of Turku, Finland
*E-mail address*: `jori.e.merikoski@utu.fi`