FUNCTION FIELDS AND RANDOM MATRICES

DOUGLAS ULMER

> ... le mathématicien qui étudie ces problèmes a l'impression de déchiffrer une inscription trilingue. Dans la première colonne se trouve la théorie riemannienne des fonctions algébriques au sens classique. La troisième colonne, c'est la théorie arithmétique des nombres algébriques. La colonne du milieu est celle dont la découverte est la plus récente; elle contient la théorie des fonctions algébriques sur un corps de Galois. Ces textes sont l'unique source de nos connaissances sur les langues dans lesquels ils sont écrits; de chaque colonne, nous n'avons bien entendu que des fragments; .... Nous savons qu'il y a des grandes différences de sens d'une colonne à l'autre, mais rien ne nous en avertit à l'avance.
>
> A. Weil, "De la métaphysique aux mathématiques" (1960)

The goal of this survey is to give some insight into how well-distributed sets of matrices in classical groups arise from families of L-functions in the context of the middle column of Weil's trilingual inscription, namely function fields of curves over finite fields. The exposition is informal and no proofs are given; rather, our aim is to illustrate what is true by considering key examples.

In the first section, we give the basic definitions and examples of function fields over finite fields and the connection with algebraic curves over function fields. The language is a throwback to Weil's Foundations, which is quite out of fashion but which gives good insight with a minimum of baggage. This part of the article should be accessible to anyone with even a modest acquaintance with the first and third columns of Weil's trilingual inscription, namely algebraic functions on Riemann surfaces and algebraic number fields.

The rest of the article requires somewhat more sophistication, although not much specific technical knowledge. In the second section, we introduce ζ- and L-functions over finite and function fields and their spectral interpretation. The cohomological apparatus is treated purely as a "black box." In the third section, we discuss families of L-functions over function fields, the main equidistribution theorems, and a small sample of applications to arithmetic. Although we do not give many details, we hope that this overview will illuminate the function field side of the beautiful Katz-Sarnak picture.

In the fourth section we give some pointers to the literature for those readers who would like to learn more of the sophisticated algebraic geometry needed to work in this area.

1. Function fields

In this first section we give a quick overview of function fields and their connection with curves over finite fields. The emphasis is on notions especially pertinent to function fields over finite fields (as opposed to function fields over algebraically closed fields), such as rational prime divisors on curves, places of function fields, and their behavior under extensions of fields and coverings of curves. The section ends with a Cebotarev equidistribution theorem which is a model for later more sophisticated equidistribution statements for matrices in Lie groups.

1.1. **Finite fields.**

If p is a prime number, then $\mathbf{Z}/p\mathbf{Z}$ with the usual operations of addition and multiplication modulo p is a field which we will also denote $\mathbf{F}_p$. If $\mathbf{F}$ is a finite field, then $\mathbf{F}$ contains a subfield isomorphic to $\mathbf{F}_p$ for a uniquely determined p, the *characteristic* of $\mathbf{F}$. (The subfield $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ is the image of the unique homomorphism of rings $\mathbf{Z} \to \mathbf{F}$ sending 1 to 1.) Since $\mathbf{F}$ is a finite dimensional vector space over its subfield $\mathbf{Z}/p\mathbf{Z}$, the cardinality of $\mathbf{F}$ must be p^f for some positive integer f. Conversely, for each prime p and positive integer f, there is a field with p^f elements, and any two such are (non-canonically) isomorphic. We may construct a field with $q = p^f$ elements by taking the splitting field of the polynomial $x^q - x$ over $\mathbf{F}_p$.

It is old-fashioned but convenient to fix a giant field Ω of characteristic p (say algebraically closed of infinite transcendence degree over $\mathbf{F}_p$) which will contain all fields under discussion. We won't mention Ω below, but all fields of characteristic p discussed are tacitly assumed to be subfields of Ω. Given Ω, we write $\overline{\mathbf{F}}_p$ for the algebraic closure of $\mathbf{F}_p$ in Ω (the set of elements of Ω which are algebraic over) and $\mathbf{F}_q$ for the unique subfield of $\overline{\mathbf{F}}_p$ with cardinality q. Its elements are precisely the q distinct solutions of the equation $x^q - x = 0$. With this notation, $\mathbf{F}_q \subset \mathbf{F}_{q'}$ if and only if q' is a power of q, say $q' = q^k$ in which case $\mathbf{F}_{q'}$ is a Galois extension of $\mathbf{F}_q$ with Galois group cyclic of order k generated by the q-power Frobenius map $\mathrm{Fr}_q(x) = x^q$.

1.2. **Function fields over finite fields.** We fix a prime p. A *function field F of characteristic p* is a finitely generated field extension of $\mathbf{F}_p$ of transcendence degree 1. The *field of constants* of F is the algebraic closure of $\mathbf{F}_p$ in F, i.e., the set of elements of F which are algebraic over . Since F is finitely generated, its field of constants is a finite field $\mathbf{F}_q$. When we say "F is a function field over $\mathbf{F}_q$" we always mean that $\mathbf{F}_q$ is the field of constants of F.

Examples:

(1) The most basic example is the rational function field $\mathbf{F}_q(x)$ where q is a power of p and x is an indeterminate. More explicitly, the elements of $\mathbf{F}_q(x)$ are ratios of polynomials in x with coefficients in $\mathbf{F}_q$. Its field of constants is $\mathbf{F}_q$.
(2) Let q be a power of p, and let F be the field extension of $\mathbf{F}_q$ generated by two elements x and y and satisfying the relation $y^2 = x^3 - 1$. More precisely, let F be the fraction field of $\mathbf{F}_q[x,y]/(y^2-x^3+1)$ or equivalently $F = \mathbf{F}_q(x)[y]/(y^2-x^3+1)$. The field of constants of F is $\mathbf{F}_q$. If $p > 3$, the field F is not isomorphic to the rational function field $\mathbf{F}_q(t)$. (This is a fun exercise. For hints, see [Sha77, p. 7]. Sadly, this point is missing from later editions of Shafarevitch's wonderful book.) The cases $p = 2$ and $p = 3$ are degenerate: F is isomorphic to the rational function field $\mathbf{F}_q(t)$. (If $p = 2$, let $t = (y+1)/x$ and note that $x = t^2$ and $y = t^3 - 1$. If $p = 3$, let $t = y/(x-1)$ and note that $x = t^2 + 1$ and $y = t^3$.)
(3) Similarly, if $p \neq 2, 5$ and q is a power of p, let F be the function field generated by x and y with relation $y^2 = x^5 - 1$. It can be shown that F has field of constants $\mathbf{F}_q$ and is not isomorphic to either of the examples above.
(4) Suppose that $p \equiv 3 \pmod 4$ so that -1 is not a square in $\mathbf{F}_p$. Let F be the function field generated over $\mathbf{F}_p$ by elements x_1, x_2, x_3 with relations $x_1x_2 = x_3$ and $x_2^2 + x_3^2 = 0$. It is not hard to see that the relations imply that $x_1^2 = -1$ and so $F \cong \mathbf{F}_{p^2}(x_2)$. The moral is that the field of constants of F is not always immediately visible from the defining generators and relations.

If F has constant field $\mathbf{F}_q$, then any element $x \in F \setminus \mathbf{F}_q$ is transcendental over $\mathbf{F}_q$ and so F contains a subfield $\mathbf{F}_q(x)$ isomorphic to the rational function field. Since F has transcendence degree 1, it is algebraic over the subfield $\mathbf{F}_q(x)$.

We can always choose the element $x \in F$ such that F is a finite *separable* extension of $\mathbf{F}_q(x)$. (It suffices to choose x which is not the p-th power of an element of F.) The theorem of the primitive element then guarantees that F is generated over $\mathbf{F}_q(x)$ by a single element y satisfying a separable polynomial over $\mathbf{F}_q(x)$:

$$f(y) = y^n + a_1(x)y^{n-1} + \cdots + a_0(x) = 0 \qquad \text{with } a_i(x) \in \mathbf{F}_q(x).$$

(Separable means that f has distinct roots, or equivalently, f and $\frac{df}{dy}$ are relatively prime in $\mathbf{F}_q(x)[y]$.) This shows that F is $\mathbf{F}_q(x)[y]/(f(y))$.

More symmetrically, we may clear the denominators in the a_i and express the relation between x and y via a two-variable polynomial over $\mathbf{F}_q$:

$$g(x,y) = \sum b_{ij}x^iy^j = 0 \qquad \text{with } b_{ij} \in \mathbf{F}_q.$$

This give us a presentation of F as the fraction field of $\mathbf{F}_q[x,y]/(g(x,y))$. Thus the general function field can be generated over its constant field by two elements satisfying a polynomial relation. Note that this representation is far from unique and it may be more natural in particular cases to give several generators and relations.

1.3. **Curves over finite fields.** Let $\overline{\mathbf{F}}_p$ be the algebraic closure of $\mathbf{F}_p$ and let $\mathbf{P}^n(\overline{\mathbf{F}}_p)$ denote the projective space of dimension n over $\overline{\mathbf{F}}_p$. Thus elements of $\mathbf{P}^n(\overline{\mathbf{F}}_p)$ are by definition the one-dimensional subspaces of the vector space $\overline{\mathbf{F}}_p^{n+1}$. If $(a_0,\dots,a_n) \in \overline{\mathbf{F}}_p^{n+1} \setminus (0,\dots,0)$, we write $[a_0 : \cdots : a_n]$ for the element of $\mathbf{P}^n(\overline{\mathbf{F}}_p)$ defined by the subspace spanned by $(a_0,\dots,a_n)$. We let $X_0,\dots,X_n$ denote the standard coordinates on $\overline{\mathbf{F}}_p^{n+1}$; of course the X_i do not give well-defined functions on $\mathbf{P}^n(\overline{\mathbf{F}}_p)$ but the ratio of two homogenous polynomials in the X_i of the same degree gives a well-defined function on the set where the denominator does not vanish. In particular, on the subset $X_0 \neq 0$, the functions $x_i = X_i/X_0$ $(i = 1,\dots,n)$ are a set of coordinates which give a bijection between the set where $X_0 \neq 0$ and the affine space $\mathbf{A}^n(\overline{\mathbf{F}}_p) = \overline{\mathbf{F}}_p^n$.

We put a topology on $\mathbf{P}^n(\overline{\mathbf{F}}_p)$ by declaring that a (Zariski) *closed* subset $Z \subset \mathbf{P}^n(\overline{\mathbf{F}}_p)$ is by definition the set of points where some collection of homogeneous polynomials vanishes. We may always take the set of polynomials to be finite and so a closed set has the form

$$Z = \left\{[a_0 : \cdots : a_n] \in \mathbf{P}^n(\overline{\mathbf{F}}_p) | f_1(a_0,\dots,a_n) = \cdots = f_k(a_0,\dots,a_n) = 0\right\}$$

where $f_1,\dots,f_k \in \overline{\mathbf{F}}_p[X_0,\dots,X_n]$ are homogeneous polynomials. A closed subset Z is said to be *defined over* $\mathbf{F}_q$ if we may take the f_i to have coefficients in $\mathbf{F}_q$.

We will work with the following definition, which is somewhat naive, but suitable for our purposes: A (smooth, projective) *curve* $\mathcal{C}$ *over* $\mathbf{F}_q$ is a closed subset $\mathcal{C} \subset \mathbf{P}^n(\overline{\mathbf{F}}_p)$ defined over $\mathbf{F}_q$, such that:

(1) $\mathcal{C}$ is infinite
(2) there exist homogeneous polynomials $f_1, \ldots, f_k$ vanishing identically on $\mathcal{C}$ such that for every $p \in \mathcal{C}$, the Jacobian matrix $(\frac{\partial f_i}{\partial X_j}(p))$ $(i = 1, \ldots, k$ and $j = 0, \ldots, n)$ has rank $n - 1$
(3) $\mathcal{C}$ is not the union of two proper closed subsets, i.e., if Z_1 and Z_2 are closed subsets and $\mathcal{C} = Z_1 \cup Z_2$ then $\mathcal{C} = Z_1$ or $\mathcal{C} = Z_2$

In the language of algebraic geometry, the first condition implies that $\mathcal{C}$ has positive dimension and the first two conditions imply that it is smooth and of dimension 1. The third condition says that $\mathcal{C}$ is *absolutely irreducible*. If in the third condition we insist that Z_1 and Z_2 be defined over $\mathbf{F}_q$ we arrive at the weaker condition that $\mathcal{C}$ is *irreducible*. Although there are sometimes good reasons to consider irreducible but not absolutely irreducible curves, for simplicity we will not do so except in one example below.

We equip $\mathcal{C}$ with the Zariski topology induced from $\mathbf{P}^n(\overline{\mathbf{F}}_p)$ so that its closed subsets are intersections of $\mathcal{C}$ with closed subsets $Z \subset \mathbf{P}^n(\overline{\mathbf{F}}_p)$.

Warning: in the current literature a curve $\mathcal{C}$ is usually defined in a more sophisticated way. The set we are considering here would be denoted $\mathcal{C}(\overline{\mathbf{F}}_p)$ and called the set of $\overline{\mathbf{F}}_p$-valued points of $\mathcal{C}$.

Examples:

(1) $\mathbf{P}^1 = \mathbf{P}^1(\overline{\mathbf{F}}_p)$ is the most basic example. It is defined by the zero polynomial on $\mathbf{P}^1$ (!) or, if that seems too tautological, by the equation $X_2 = 0$ in $\mathbf{P}^2(\overline{\mathbf{F}}_p)$. Either representation makes it clear that $\mathbf{P}^1$ is defined over $\mathbf{F}_p$.
(1′) For $p > 2$, let $\mathcal{C}_2$ be the curve in $\mathbf{P}^2(\overline{\mathbf{F}}_p)$ defined over $\mathbf{F}_p$ by the polynomial $X_1^2+X_2^2-X_0^2$. Note that restricted to $\mathcal{C}_2 \cap \{X_0 \neq 0\}$, the coordinate functions x_i satisfy $x_1^2 + x_2^2 = 1$.
(1″) For any p, let $\mathcal{C}_3$ be the curve in $\mathbf{P}^3(\overline{\mathbf{F}}_p)$ defined over $\mathbf{F}_p$ by the polynomials $X_0X_2 - X_1^2$, $X_0X_3 - X_1X_2$, and $X_1X_3 = X_2^2$. Note that restricted to $\mathcal{C}_3 \cap \{X_0 \neq 0\}$, the coordinate functions x_i satisfy $x_2 = x_1^2$ and $x_3 = x_1^3$.
(2) Assume that $p > 3$ and let $\mathcal{C}_3'$ be the curve in $\mathbf{P}^2(\overline{\mathbf{F}}_p)$ defined over $\mathbf{F}_p$ by the polynomial $X_0X_2^2 - X_1^3 + X_0^3$. (If $p = 2$ or 3 the second condition in the definition of a curve is not met: the Jacobian matrix is 0 at $[1, 0, 1]$ if $p = 2$ and at $[1, 1, 0]$ if $p = 3$.)

Note that restricted to $\mathcal{C}'_3 \cap \{X_0 \neq 0\}$, the coordinate functions x_i satisfy $x_2^2 = x_1^3 - 1$.

(3) Assume that $p \neq 2, 5$ and let $\mathcal{C}_5$ be the closed subset of $\mathbf{P}^3(\overline{\mathbf{F}}_p)$ defined over $\mathbf{F}_p$ by the equation polynomials $X_0X_2 - X_1^2$, $X_0X_3^2 - X_1X_2^2 + X_0^3$, and $X_1X_3 - X_2^3 + X_0^2X_1$. Note that restricted to $\mathcal{C}_5 \cap \{X_0 \neq 0\}$, the coordinate functions x_i satisfy $x_2 = x_1^2$ and $x_3^2 = x_1^5 - 1$.

(4) Assume that $p \equiv 3 \pmod 4$ so that $-1 \in \mathbf{F}_p$ is not a square. Let $\mathcal{C}'_2$ be defined over $\mathbf{F}_p$ by the three polynomials $X_0^2 + X_1^2$, $X_2^2 + X_3^2$, and $X_0X_3 - X_1X_2$. Then $\mathcal{C}'_2$ is irreducible, but it is not absolutely irreducible and so it is not a curve by our definition. Indeed, $\mathcal{C}'_2$ is the union of the two lines $\{X_0 = iX_1, X_2 = iX_3\}$ and $\{X_0 = -iX_1, X_2 = -iX_3\}$ defined over $\mathbf{F}_{p^2}$ where $i \in \mathbf{F}_{p^2}$ satisfies $i^2 = -1$. Note that restricted to $\mathcal{C}'_2 \cap \{X_0 \neq 0\}$, the coordinate functions x_i satisfy $x_1^2 = -1$, $x_2^2 + x_3^2 = 0$ and $x_3 = x_1x_2$.

1.4. **Morphisms and rational functions.** If $\mathcal{C} \subset \mathbf{P}^m(\overline{\mathbf{F}}_p)$ and $\mathcal{C}' \subset \mathbf{P}^n(\overline{\mathbf{F}}_p)$ are curves defined over $\mathbf{F}_q$, a *morphism* of curves is a map $\phi : \mathcal{C} \to \mathcal{C}'$ with the property that at each point $P \in \mathcal{C}$, ϕ is represented in an open neighborhood of P by homogenous polynomials. In other words, for each $P \in \mathcal{C}$ there should exist polynomials $f_0, \dots, f_n \in \overline{\mathbf{F}}_p[X_0, \dots, X_m]$, all homogeneous of the same degree, such that for all Q in some open neighborhood of P, not all of the f_i vanish at Q and $\phi(Q) = [f_0(Q) : \cdots : f_n(Q)]$. We say that ϕ is *defined over* $\mathbf{F}_q$ if it possible to choose the f_i with coefficients in $\mathbf{F}_q$. An *isomorphism* is a morphism which is bijective and whose inverse is a morphism.

Examples:

(1) If f_0 and f_1 are homogeneous polynomials in $\mathbf{F}_q[X_0, X_1]$ of the same degree, not both 0, and with no common factors, then

$$[a_0 : a_1] \mapsto [f_0(a_0, a_1) : f_1(a_0, a_1)]$$

defines a morphism $\mathbf{P}^1 \to \mathbf{P}^1$. Using that $\mathbf{F}_q[X_0, X_1]$ is a unique factorization domain, one checks that every morphism from $\mathbf{P}^1$ to itself defined over $\mathbf{F}_q$ is of this form.

(1′) For $p > 2$, the polynomials $f_0 = X_0^2 + X_1^2$, $f_1 = X_0^2 - X_1^2$, and $f_2 = 2X_0X_1$ define a morphism from $\mathbf{P}^1$ over $\mathbf{F}_p$ to the curve $\mathcal{C}_2$ in Example (1′) of Section 1.3. This morphism is an isomorphism with inverse defined by $f_0 = \frac{1}{2}(X_0 + X_1)$ and $f_1 = \frac{1}{2}(X_0 - X_1)$.

(1″) For any p, the polynomials $f_0 = X_0^3$, $f_1 = X_0^2X_1$, $f_2 = X_0X_1^2$, and $f_3 = X_1^3$ define a morphism ϕ from $\mathbf{P}^1$ over $\mathbf{F}_p$ to the curve $\mathcal{C}_3$ in Example (1″) of Section 1.3. This morphism is an isomorphism with inverse defined on $\{X_0 \neq 0\}$ by $f_0 = X_0$ and $f_1 = X_1$ and on $\{X_3 \neq 0\}$ by $f_0 = X_2$ and $f_1 = X_3$. In this example, it is not possible to define the inverse of ϕ by a single set of polynomials on all of $\mathcal{C}_3$. Note also that the polynomials defining a morphism are in general not at all unique. For example, on $\{X_0X_3 \neq 0\}$, the inverse of ϕ is defined both by $f_0 = X_0$ and $f_1 = X_1$ and by $f_0 = X_2$ and $f_1 = X_3$.

(2) Let $\mathcal{C}_3'$ be as in Example (2) of Section 1.3. We define a morphism $\phi : \mathcal{C}_3' \to \mathbf{P}^1$ by setting $\phi([a_0 : a_1 : a_2]) = [a_0 : a_1]$ on the open set where $a_0 \neq 0$ and $\phi([a_0 : a_1 : a_2]) = [a_1^2 : a_0^2 + a_2^2]$ on the open set where $a_0^2 + a_2^2 \neq 0$. These requirements are compatible since if $a_0 \neq 0$ and $a_0^2 + a_2^2 \neq 0$, then $a_1 \neq 0$ and

$$[a_0 : a_1] = [a_0a_1^2 : a_1^3] = [a_0a_1^2 : a_0^3 + a_0a_2^2] = [a_1^2 : a_0^2 + a_2^2].$$

If we think of $\mathbf{P}^1(\overline{\mathbf{F}}_p) \setminus \{[0, 1]\}$ as $\overline{\mathbf{F}}_p$ via $[a_0 : a_1] \mapsto a_1/a_0$, then the morphism ϕ extends the function $x_1 = X_1/X_0$, defined on $\mathcal{C}_3' \cap \{X_0 \neq 0\}$ to a morphism $\mathcal{C}_3' \to \mathbf{P}^1(\overline{\mathbf{F}}_p)$. Again, it is not possible to find a single pair of homogeneous polynomials representing ϕ at all points of $\mathcal{C}_3'$.

(2′) Let $\mathcal{C}_3'$ be as above. Choose a non-square element $a \in \mathbf{F}_p$ ($p > 3$) and define $\mathcal{C}_3'' \subset \mathbf{P}^2(\overline{\mathbf{F}}_p)$ by the equation $aX_0X_2^2 - X_1^3 + X_0^3 = 0$. Note that both $\mathcal{C}_3'$ and $\mathcal{C}_3''$ are defined over $\mathbf{F}_p$. Let $b \in \mathbf{F}_{p^2}$ be a square root of a and define a morphism $\phi : \mathcal{C}_3'' \to \mathcal{C}_3'$ by $\phi([a_0 : a_1 : a_2]) = [ba_o : a_1 : a_2]$. It is clear that ϕ is defined over $\mathbf{F}_{p^2}$ and is an isomorphism. On the other hand, one can show that $\mathcal{C}_3''$ and $\mathcal{C}_3'$ are not isomorphic over $\mathbf{F}_p$. This shows that two curves not isomorphic over their fields of definition may become isomorphic over a larger field. One says that $\mathcal{C}_3''$ is a *twist* of $\mathcal{C}_3'$.

(3) If $\mathcal{C} \subset \mathbf{P}^n(\overline{\mathbf{F}}_p)$ is a curve defined over $\mathbf{F}_q$, then there is an important morphism, the q-power Frobenius morphism $\mathrm{Fr}_q : \mathcal{C} \to \mathcal{C}$, defined by $\mathrm{Fr}_q([a_0 : \cdots : a_n]) = [a_0^q : \cdots : a_n^q]$. Note that the fixed points of Fr_q are precisely the points of $\mathcal{C}$ with coordinates in $\mathbf{F}_q$.

(4) If $\mathcal{C} \subset \mathbf{P}^n(\overline{\mathbf{F}}_p)$ is a curve and $f_0, \ldots, f_k$ are homogeneous polynomials in $X_0, \ldots X_n$ which do not all vanish identically on $\mathcal{C}$, then the map $\phi : \mathcal{C} \to \mathbf{P}^k(\overline{\mathbf{F}}_p)$ given by

$$\phi([a_0 : \cdots : a_n]) = [f_0(a_0, \ldots, a_n) : \cdots : f_k(a_0, \ldots, a_n)]$$

is well-defined on the non-empty open subset of $\mathcal{C}$ where not all of the f_i vanish. It is an important fact that ϕ can always be extended uniquely to a well-defined morphism on all of $\mathcal{C}$. (NB: This is false for higher dimensional varieties.) In particular, there are globally defined morphisms $x_i : \mathcal{C} \to \mathbf{P}^1$ extending the maps $[a_0 : \cdots : a_n] \mapsto [a_0 : a_i]$ which are *a priori* only defined on $\mathcal{C} \cap \{X_0 \neq 0\}$.

A *rational function* on a curve $\mathcal{C}$ over $\mathbf{F}_q$ is a morphism $\phi : \mathcal{C} \to \mathbf{P}^1$ defined over $\mathbf{F}_q$, except that we rule out the constant morphism with image $\infty = [0 : 1]$. (NB: This is a reasonable definition only for curves, not for higher dimensional varieties.) In a neighborhood of any $P \in \mathcal{C}$, ϕ can be represented by polynomials: $\phi(Q) = [f_0(Q) : f_1(Q)]$ where f_0 and f_1 are homogeneous of the same degree and f_0 does not vanish identically. It is useful to think of ϕ as an $\overline{\mathbf{F}}_p$-valued function (with poles) whose value at Q is $\frac{f_1(Q)}{f_0(Q)}$. We say that ϕ is *regular at* $P \in \mathcal{C}$ if $\phi(P) \neq \infty = [0 : 1]$. If we restrict to an open set where ϕ is regular, i.e., where f_0 does not vanish, then we get a well-defined $\overline{\mathbf{F}}_p$-valued function. If ϕ and ϕ' are two rational functions, we may restrict them to an open set where they both give well-defined $\overline{\mathbf{F}}_p$-valued functions, add or multiply them, and then extend back to rational functions on $\mathcal{C}$. More explicitly, if ϕ and ϕ' are represented on some open set $U \subset \mathcal{C}$ by $[f_0 : f_1]$ and $[f_0' : f_1']$, then $\phi + \phi'$ is represented by $[f_0 f_0' : f_0' f_1 + f_0 f_1']$ and $\phi\phi'$ is represented by $[f_0 f_0', f_1 f_1']$. This gives the set of rational functions the structure of a ring, in fact an algebra over $\mathbf{F}_q$. This algebra turns out to be a field extension of $\mathbf{F}_q$ of transcendence degree 1, i.e., a function field in the sense of the previous subsection. It is denoted $\mathbf{F}_q(\mathcal{C})$.

Note that the ratio f_1/f_0 can be written as a rational function (ratio of polynomials) in $x_1 = X_1/X_0, \dots, x_n = X_n/X_0$. This shows that if $\mathcal{C} \subset \mathbf{P}^n(\overline{\mathbf{F}}_p)$, then $\mathbf{F}_q(\mathcal{C})$ is generated over $\mathbf{F}_q$ by the rational functions $x_1, \dots, x_n$. To determine $\mathbf{F}_q(\mathcal{C})$, we need only determine the relations among the x_i.

Examples:

(1) As noted above, a rational function on $\mathbf{P}^1$ is given by two homogeneous polynomials f_0 and f_1 of the same degree, with $f_0 \neq 0$. Two rational functions $[f_0 : f_1]$ and $[f_0' : f_1']$ are equal if and only if $f_1/f_0 = f_1'/f_0'$. Thus we see that rational functions on $\mathbf{P}^1$ are equivalent to rational functions (ratios of polynomials) in $x = X_1/X_0$, i.e., $\mathbf{F}_p(\mathbf{P}^1) = \mathbf{F}_p(x)$ and more generally $\mathbf{F}_q(\mathbf{P}^1) = \mathbf{F}_q(x)$.

(1′) The function fields of the curves $\mathcal{C}_2$ and $\mathcal{C}_3$ in Examples (1′) and (1″) of Section 1.3 are also isomorphic to $\mathbf{F}_p(x)$. One can see this by using the relations among the x_i noted above, or by using the fact (to be explained below) that isomorphic curves have isomorphic function fields.

(2) Let $\mathcal{C}_3'$ be as in Example (2) of Section 1.3 and let x_1 be the rational function ϕ of that example (so $x_1([a_0 : a_1 : a_2]) = [a_0 : a_1]$ or $[a_1^2 : a_0^2 + a_2^2]$). Let x_2 be the rational function defined on all of $\mathcal{C}_3'$ by $x_2([a_0 : a_1 : a_2]) = [a_0 : a_2]$. Then x_1 and x_2 generate the field of rational functions on $\mathcal{C}_3'$ over $\mathbf{F}_q$ and they satisfy the relation $x_2^2 = x_1^3 - 1$. In other words, $\mathbf{F}_q(\mathcal{C}_3')$ is the field in Example (2) of Section 1.2.

(3) Let $\mathcal{C}_5$ be as in Example (3) of Section 1.3 and define rational functions x_1 and x_3 by

$$x_1([a_0 : a_1 : a_2 : a_3]) = \begin{cases} [a_0 : a_1] & \text{if } a_0 \neq 0 \\ [a_2^2 : a_0^2 + a_3^2] & \text{if } a_0^2 + a_3^2 \neq 0 \end{cases}$$

and

$$x_3([a_0 : a_1 : a_2 : a_3]) = [a_0 : a_3].$$

(We leave it to the reader to check that these formulas do indeed define rational functions on $\mathcal{C}_5$.) It is not hard to see that x_1 and x_3 generate $\mathbf{F}_q(\mathcal{C}_5)$. The equations defining $\mathcal{C}_5$ imply that $x_3^2 = x_1^5 - 1$ and that all relations among x_1 and x_3 are consequences of this one. Thus $\mathbf{F}_q(\mathcal{C}_5)$ is the function field of Example (3) of Section 1.2.

1.5. **The function field/curve dictionary.** The examples at the end of the last section illustrate the general fact that if $\mathcal{C}$ is a curve defined over $\mathbf{F}_q$, then the field of rational functions $\mathbf{F}_q(\mathcal{C})$ is a function field, i.e., a finitely generated extension of $\mathbf{F}_p$ of transcendence degree one, with field of constants $\mathbf{F}_q$.

Conversely, it turns out that every function field F with field of constants $\mathbf{F}_q$ is the field of rational functions of a curve defined over $\mathbf{F}_q$ which is uniquely determined up to $\mathbf{F}_q$-isomorphism. We sketch one construction of the curve corresponding to a function field F. As we pointed out above, F may be generated over $\mathbf{F}_q$ by two elements x and y satisfying a single relation

$$0 = g(x, y) = \sum b_{ij} x^i y^j \qquad \text{with } b_{ij} \in \mathbf{F}_q.$$

If g has degree d, we form

$$G(X_0, X_1, X_2) = X_0^d g(X_1/X_0, X_2/X_0) = \sum b_{ij} X_0^{d-i-j} X_1^i X_2^j$$

and consider the closed subset of $\mathbf{P}^2(\overline{\mathbf{F}}_p)$ defined by $G = 0$. This closed subset will be infinite and irreducible, but it will not in general be a curve under our definition, since it may not satisfy the Jacobian condition. If it does, we are finished. If not, the closed set $\{G = 0\}$ has singularities and the classical process of blowing up (see [Ful89, Chap. 7]) gives an algorithm to resolve the singularities and find a smooth curve in some high-dimensional projective space with function field F. By a suitable projection, the curve $\mathcal{C}$ can be embedded in $\mathbf{P}^3(\overline{\mathbf{F}}_p)$. In general we will not be able to find a *plane* curve with function field F. This is the case for example with the function field in Example (3) of Section 1.2 generated by x and y satisfying $y^2 = x^5 - 1$. The simplest curve with this function field is a curve in $\mathbf{P}^3(\overline{\mathbf{F}}_p)$ defined by three equations.

The dictionary between curves and function fields extends to morphisms and field extensions. More precisely, if $\mathcal{C}$ and $\mathcal{C}'$ are two curves defined over $\mathbf{F}_q$ and $\phi : \mathcal{C} \to \mathcal{C}'$ is a non-constant morphism defined over $\mathbf{F}_q$, then composition with ϕ induces a "pull-back" homomorphism of fields $\mathbf{F}_q(\mathcal{C}') \hookrightarrow \mathbf{F}_q(\mathcal{C})$ which is the identity on $\mathbf{F}_q$. Conversely, it can be shown that if F and F' are function fields over $\mathbf{F}_q$ with corresponding curves $\mathcal{C}$ and $\mathcal{C}'$, then a field inclusion $F' \hookrightarrow F$ which is the identity on $\mathbf{F}_q$ is induced by a unique non-constant morphism of curves $\mathcal{C} \to \mathcal{C}'$ which is defined over $\mathbf{F}_q$.

Examples:

(1) If $\mathcal{C}$ is a curve over $\mathbf{F}_q$ and x is a non-constant rational function on $\mathcal{C}$, then x is transcendental over $\mathbf{F}_q$. Thus the rational function field $F' = \mathbf{F}_q(x)$ is a subfield of $F = \mathbf{F}_q(\mathcal{C})$. The corresponding morphism $\mathcal{C} \to \mathbf{P}^1$ is the morphism x.

(2) Suppose F' is a function field with field of constants $\mathbf{F}_q$ and $\mathcal{C}'$ is the corresponding curve over $\mathbf{F}_q$. If r is a power of q so that $\mathbf{F}_r$ is a finite extension of $\mathbf{F}_q$, then the function field $F = \mathbf{F}_r F'$ corresponds to the same curve $\mathcal{C}'$ viewed over $\mathbf{F}_r$. (Here $\mathbf{F}_r F'$ is the compositum of $\mathbf{F}_r$ and F', i.e., the smallest field containing both $\mathbf{F}_r$ and F'.) In other words, $F = \mathbf{F}_r(\mathcal{C}')$.

(3) We say that an extension of function fields F/F' is *geometric* if it is separable and if the field of constants of F and F' are the same. If $n = [F : F']$ is the degree of the field extension, then the corresponding morphism of curves $\phi : \mathcal{C} \to \mathcal{C}'$ has degree n in the sense that for all but finitely many $P \in \mathcal{C}'$, $\phi^{-1}(P)$ consists of n points.

(4) If F/F' is a purely inseparable extension of function fields, say of degree p^m, then $F' = F^{p^m}$, the subfield of p^m-th powers. In terms of suitable equations, the morphism $\mathcal{C} \to \mathcal{C}'$ acts on points by raising their coordinates to the p^m-th power.

An arbitrary extension can be factored into three like these: Given F/F', let $\mathbf{F}_r$ be the field of constants of F and let F^{sep} be the separable closure of F' in F. Then $\mathbf{F}_r F'/F'$ is a constant field extension, $F^{sep}/\mathbf{F}_r F'$ is geometric, and F/F^{sep} is purely inseparable.

1.6. **Points, prime divisors, and places.** As we have defined it, a curve $\mathcal{C}$ over $\mathbf{F}_q$ is a set of points with coordinates in $\overline{\mathbf{F}}_p$. We would like to have a set which reflects the fact that the equations defining $\mathcal{C}$ have coefficients in $\mathbf{F}_q$. The naive thing to look at would be the set of $\mathbf{F}_q$-rational points of $\mathcal{C}$, i.e., those with coordinates in $\mathbf{F}_q$, but this set is too small to be useful—it may even be empty. The classical approach is to consider $\mathbf{F}_q$-rational prime divisors.

A *divisor* on $\mathcal{C}$ is a finite, formal, linear combination $\mathfrak{d} = \sum a_P P$ of points of $\mathcal{C}$ with integer coefficients. A divisor $\mathfrak{d}$ is called *effective* if $a_P \geq 0$ for all P. The *degree* of $\mathfrak{d}$ is $\deg(\mathfrak{d}) = \sum a_P$. The *support* of $\mathfrak{d}$, written $|\mathfrak{d}|$, is the set of points appearing in $\mathfrak{d}$ with non-zero coefficient.

If $\sigma \in \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q)$ and $P \in \mathcal{C}$, then P^σ is again in $\mathcal{C}$. (Here σ acts on the coordinates of P and the claim follows from the fact that the equations defining $\mathcal{C}$ have coefficients in $\mathbf{F}_q$.) We extend this action to divisors by linearity $((\sum a_P P)^\sigma = \sum a_P P^\sigma)$ and we say that a divisor $\mathfrak{d} = \sum a_P P$ is $\mathbf{F}_q$*-rational* if it is fixed by the Galois group, i.e., if $\mathfrak{d}^\sigma = \mathfrak{d}$ for all $\sigma \in \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q)$.

A *prime divisor* is an effective $\mathbf{F}_q$-rational divisor which is non-zero and cannot be written as the sum of two non-zero $\mathbf{F}_q$-rational effective divisors. (Note that whether or not a divisor is prime depends on the ground field over which we are considering our curve. A better terminology might be $\mathbf{F}_q$-prime, but we will stick with the traditional terminology.) It is not hard to see that the prime divisors of $\mathcal{C}$ are in bijection with the orbits of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q)$ acting on $\mathcal{C}$. If $\mathfrak{p}$ is a prime divisor, we define the residue field of $\mathfrak{p}$ to be the field generated over $\mathbf{F}_q$ by the coordinates of any point in the support of $\mathfrak{p}$. If $\mathfrak{p}$ is prime and has degree d, then the residue field at $\mathfrak{p}$ is $\mathbf{F}_{q^d}$.

If P is a point of $\mathcal{C}$ and $f \in \mathbf{F}_q(\mathcal{C})$ is a rational function on $\mathcal{C}$, then f has a well-defined order of vanishing or pole at P. One motivation for considering prime divisors is that the order of f at P is the same for all points P in the support of the prime divisor $\mathfrak{p}$ containing P. In other words, the various points in $|\mathfrak{p}|$ cannot be distinguished from one another by the vanishing of $\mathbf{F}_q$-rational functions.

Examples:

(1) Let $\mathcal{C} = \mathbf{P}^1$ over $\mathbf{F}_q$. The divisors of degree 1 are simply the points of $\mathbf{P}^1$ with coordinates in $\mathbf{F}_q$. The prime divisors of degree $d > 1$ are in bijection with the irreducible, monic polynomials in $\mathbf{F}_q[x]$ of degree d, a polynomial corresponding to the formal sum of its roots.

(2) Let $\mathcal{C}'_3$ be the curve in Example (2) of Section 1.3 over $\mathbf{F}_q$. If $a \in \mathbf{F}_q$ with $a^3 - 1 \neq 0$, consider the points $P = [a : b : 1]$ and $Q = [a : -b : 1]$ where $b \in \overline{\mathbf{F}}_p$ satisfies $b^2 = a$. The divisor $\mathfrak{d} = P + Q$ has degree two and it is prime if and only if $b \notin \mathbf{F}_q$. If $b \in \mathbf{F}_q$, then $\mathfrak{d}$ is the sum of two prime divisors, namely P and Q.

The set of prime divisors on $\mathcal{C}$ is more "arithmetical" than the full set of points on $\mathcal{C}$ (since it takes into account that $\mathcal{C}$ is defined over $\mathbf{F}_q$) and more convenient and flexible than the set of $\mathbf{F}_q$-rational points of $\mathcal{C}$.

Prime divisors play the role of the prime ideals of a number field. More precisely, if $\mathfrak{p} = \sum P_i$ is a prime divisor and if $f \in \mathbf{F}_q(\mathcal{C})$ we say f is regular (resp. vanishes) at $\mathfrak{p}$ if it is regular (resp. vanishes) at one and therefore all of the $P_i \in |\mathfrak{p}|$. The set of $f \in \mathbf{F}_q(\mathcal{C})$ which are regular at $\mathfrak{p}$ is a discrete valuation ring $R_\mathfrak{p}$ with fraction field $\mathbf{F}_q(\mathcal{C})$. The maximal ideal of $R_\mathfrak{p}$ is the set of f which vanish at $\mathfrak{p}$. The residue field at $\mathfrak{p}$ as we defined it above turns out to be $R_\mathfrak{p}$ modulo its maximal ideal. We get a valuation $\mathrm{ord}_\mathfrak{p} : \mathbf{F}_q(\mathcal{C})^\times \to \mathbf{Z}$ in the usual way. It turns out that every non-trivial valuation of $\mathbf{F}_q(\mathcal{C})$ is $\mathrm{ord}_\mathfrak{p}$ for a uniquely determined prime divisor $\mathfrak{p}$. (Therefore, it is possible, although not in my opinion advisable, to eliminate the geometry completely and study function fields via their valuations. What one gains in algebraic purity hardly seems to compensate for the loss of geometric intuition this approach entails.)

Here is one respect in which the analogy between function fields and number fields breaks down ("il y a des grandes différences de sens d'une colonne à l'autre"): in a number field F, there is a canonical Dedekind domain contained in F whose primes give the non-archimedean valuations of F, namely the ring of integers. In a function field, to get a Dedekind domain we fix a non-empty set of prime divisors S and then consider the ring R of functions regular at all primes not in S. The prime ideals of R are then in bijection with the prime divisors of $\mathbf{F}_q(\mathcal{C})$ except those in S, and with the valuations of $\mathbf{F}_q(\mathcal{C})$ except those arising from primes in S. One thinks of the primes in S as the "infinite primes", but there is no canonical choice for the set S.

1.7. **The Riemann-Roch theorem.** The Riemann-Roch theorem is true for curves over non-algebraically closed fields and the statement is essentially the same as for the case of curves over algebraically closed fields. We give the basics in our context.

Let $\mathcal{C}$ be a curve defined over $\mathbf{F}_q$ with function field $F = \mathbf{F}_q(\mathcal{C})$. For each $P \in \mathcal{C}$ and $0 \neq f \in F$, there is a well-defined order of vanishing or pole of f at P, denoted $\mathrm{ord}_P(f)$. The *divisor of* f is defined as the formal sum $(f) = \sum_P \mathrm{ord}_P(f)$ which is in fact a finite sum. It is not hard to see that (f) is $\mathbf{F}_q$-rational and a basic results says that it has degree 0: $\sum_P \mathrm{ord}_P(f) = 0$.

If $\mathfrak{d}$ is an $\mathbf{F}_q$-rational divisor, we define the Riemann-Roch space $L(\mathfrak{d})$ by

$$L(\mathfrak{d}) = \{f \in F^\times | (f) + \mathfrak{d} \text{ is effective}\} \cup \{0\}.$$

Roughly speaking, $L(\mathfrak{d})$ consists of rational functions whose poles are at worst given by $\mathfrak{d}$. It is clear that $L(\mathfrak{d})$ is an $\mathbf{F}_q$ vector space which turns out to be finite dimensional. Note that $L(\mathfrak{d})$ is obviously zero if $\mathfrak{d}$ has negative degree.

The Riemann-Roch theorem in its most basic form is a formula that often allows one to compute the dimension $l(\mathfrak{d})$ of $L(\mathfrak{d})$. The theorem says that there is a non-negative integer g, the *genus of* $\mathcal{C}$ and a divisor ω of degree $2g-2$ such that for all divisors $\mathfrak{d}$

$$l(\mathfrak{d}) - l(\omega - \mathfrak{d}) = \deg(\mathfrak{d}) - g + 1.$$

The divisor ω is not unique (if ω works, then so does $\omega + (f)$ for any non-zero f)). Despite this ambiguity, ω is called a *canonical divisor*. It turns out that ω can be calculated as the divisor of a rational 1-form (i.e., a 1-form possibly with poles) on $\mathcal{C}$.

It follows immediately that $l(\mathfrak{d}) \geq \deg(\mathfrak{d}) - g + 1$ with equality if $\deg \mathfrak{d} > 2g - 2$. This gives a large supply of functions with controlled poles.

As an example, note that on $\mathbf{P}^1$ over $\mathbf{F}_q$ the Riemann-Roch space $L(d\infty)$ is just the space of polynomials of degree d, which has dimension $d+1$. It follows that the genus of $\mathbf{P}^1$ is 0. One can check that the genus of the curve in Example (2) of Section 1.3 is 1 and the genus of the curve in Example (3) is 2.

As another application, which we leave as a simple exercise, the theorem implies a partial converse to the statement that $\mathbf{P}^1$ has genus zero: if $\mathcal{C}$ has genus zero and an $\mathbf{F}_q$-rational divisor of degree 1, then $\mathcal{C}$ is isomorphic to $\mathbf{P}^1$. It turns out that over a finite field $\mathbf{F}_q$ every curve has an $\mathbf{F}_q$-rational divisor of degree one, so this partial converse is in fact a complete converse.

The reader curious about what a number field analog of the Riemann-Roch theorem might be should consult Weil's "Basic Number Theory," [Wei95, Chap. VI].

1.8. **Extensions, coverings, and splitting.** Let $\mathcal{C}$ and $\mathcal{C}'$ be curves defined over $\mathbf{F}_q$ and let $\phi : \mathcal{C} \to \mathcal{C}'$ be a morphism of curves defined over $\mathbf{F}_q$. We say that ϕ has degree n if $n = [\mathbf{F}_q(\mathcal{C}) : \mathbf{F}_q(\mathcal{C}')]$. Given a point P in $\mathcal{C}$ or $\mathcal{C}'$ we write $\mathbf{F}_q(P)$ for the field generated over $\mathbf{F}_q$ by the coordinates of P. We define an inverse image mapping on divisors. If $P \in \mathcal{C}'$ and if set-theoretically the inverse image of P in $\mathcal{C}$ is $\{Q_1, \ldots, Q_k\}$, then we assign a multiplicity e_i to each Q_i by choosing a rational function f vanishing simply at P and setting $e_i =$ the order of vanishing of the pull-back of f at Q_i. We then define $\phi^{-1}(P)$ as $\sum e_i Q_i$ and extend to divisors by linearity. It turns out that if $\mathbf{F}_q(\mathcal{C})$ is separable over $\mathbf{F}_q(\mathcal{C}')$ (i.e., if we have a geometric extension of function fields), then for all but finitely many P, all the e_i are 1, and in general for all P, $\sum e_i = n$.

If $\mathfrak{p}$ is a prime divisor of $\mathcal{C}$, then we may decompose $\phi^{-1}(\mathfrak{p})$ into a sum of prime divisors $\mathfrak{q}_1, \ldots, \mathfrak{q}_g$. For each $\mathfrak{q}_i$ we may define the *residue degree* f_i as $\deg \mathfrak{q}_i / \deg \mathfrak{p}$ or equivalently, the degree of the field extension $\mathbf{F}_q(Q)/\mathbf{F}_q(P)$ where P is any point in the support of $\mathfrak{p}$ and Q is any point over P in the support of $\mathfrak{q}_i$. The *ramification index* e_i is the e_i defined above for any point P in the support of $\mathfrak{p}$ and any point Q over P in the support of $\mathfrak{q}_i$. It is a basic fact that for all $\mathfrak{p}$, $\sum_{i=1}^{g} e_i f_i = n$ where $n = [F : F']$.

Examples:

(1) Suppose that $p > 3$, q is a power of p, F is the fraction field of $\mathbf{F}_q[x, y]/(y^2 - x^3 + 1)$, and $\mathbf{F}' = \mathbf{F}_q(x)$, so that the corresponding morphism of curves $\phi : \mathcal{C} \to \mathcal{C}' = \mathbf{P}^1$ is as in Example (2) in Section 1.4. Suppose that $\mathfrak{p}$ is a prime divisor of degree one corresponding to a finite $\mathbf{F}_q$-rational point P with coordinate $x = a$. If $a^3 - 1 = 0$, then $\phi^{-1}(\mathfrak{p})$ is a single prime $\mathfrak{q}$ with $e = 2$ and $f = 1$; we say $\mathfrak{p}$ is ramified. If $a^3 - 1$ is a non-zero square of $\mathbf{F}_q$, then $\phi^{-1}(\mathfrak{p})$ consists of two primes $\mathfrak{q}_1$ and $\mathfrak{q}_2$, both with $e = 1$ and $f = 1$; we say that $\mathfrak{p}$ splits. Finally, if $a^3 - 1$ is a non-square in $\mathbf{F}_q$, then $\phi^{-1}(\mathfrak{p})$ consists of one prime $\mathfrak{q}$ with $e = 1$ and $f = 2$; we say that $\mathfrak{p}$ is inert.

(2) With notation as in the last example, if $\mathfrak{p}$ is a general prime, say $\mathfrak{p} = \sum P_i$, then the behavior of ϕ over each of the points P_i is the same (one ramified point, two points with the same field of coordinates as P_i, or two points with coordinates in a quadratic

extension of $\mathbf{F}_q(P_i)$) and so $\phi^{-1}(\mathfrak{p}) = 2\mathfrak{q}$ with $\deg \mathfrak{q} = \deg \mathfrak{p}$ ($\mathfrak{p}$ ramifies), $\phi^{-1}(\mathfrak{p}) = \mathfrak{q}_1 + \mathfrak{q}_2$ with $\deg \mathfrak{q}_i = \deg \mathfrak{p}$ ($\mathfrak{p}$ splits), or $\phi^{-1}(\mathfrak{p}) = \mathfrak{q}$ with $\deg \mathfrak{q} = 2 \deg \mathfrak{p}$ ($\mathfrak{p}$ is inert).

(3) If $\mathcal{C}$ is defined over $\mathbf{F}_q$ and $r = q^n$, then we may consider the splitting of $\mathbf{F}_q$-rational prime divisors into $\mathbf{F}_r$-rational prime divisors. This splitting is determined purely in terms of degrees: an $\mathbf{F}_q$-rational prime $\mathfrak{p}$ of degree d splits into $\gcd(n, d)$ $\mathbf{F}_r$-rational primes, each with $e = 1$ and $f = n/\gcd(d, n)$.

(4) If $\mathcal{C} \to \mathcal{C}'$ is a morphism of curves defined over $\mathbf{F}_q$ and is purely inseparable of degree p^m, then every prime $\mathfrak{p}$ of $\mathcal{C}'$ pulls back to a single prime $\mathfrak{q}$ of $\mathcal{C}$ with $e = p^m$ and $f = 1$.

In the case of a morphism $\mathcal{C} \to \mathcal{C}'$ corresponding to a geometric extension F/F' which is Galois, it is easy to see that for a fixed prime $\mathfrak{p}$ of $\mathcal{C}'$, the ramification and residue degrees e_i and f_i are all the same, in other words, $\mathfrak{p}$ splits into g primes, all with ramification index e and residue degree f, and we have $efg = n = [F : F']$. Only finitely many $\mathfrak{p}$ have $e > 1$ and one can make very precise statements about the distribution of primes having allowable values of f and g. See Section 1.10 below.

1.9. **Frobenius elements.** Let F' be a function field with constant field $\mathbf{F}_q$ and let F be a finite Galois extension of F' with Galois group G; for simplicity we assume the extension F/F' is geometric, i.e., the field of constants of F is $\mathbf{F}_q$. Let $\phi : \mathcal{C} \to \mathcal{C}'$ be the corresponding morphism of curves over $\mathbf{F}_q$. Fix a finite extension $\mathbf{F}_r$ of $\mathbf{F}_q$ and a point of $P \in \mathcal{C}'$ rational over $\mathbf{F}_r$. We may view P as an $\mathbf{F}_r$-rational prime divisor. Suppose that $\mathfrak{p}_1, \ldots, \mathfrak{p}_g$ are the $\mathbf{F}_r$-rational primes of $\mathcal{C}$ over P, so that as divisors $\phi^{-1}(P) = e\mathfrak{p}_1 + \cdots + e\mathfrak{p}_g$ where e is the ramification index. The Galois group G acts (transitively in fact) on the set of $\mathfrak{p}_i$ and we let $D_{\mathfrak{p}_i} \subset G$ denote the stabilizer of $\mathfrak{p}_i$, the *decomposition group at* $\mathfrak{p}_i$. Then $D_{\mathfrak{p}_i}$ acts on the residue field at $\mathfrak{p}_i$ and so we have a homomorphism $D_{\mathfrak{p}_i} \to \mathrm{Gal}(\mathbf{F}_{r'}/\mathbf{F}_r)$ where $\mathbf{F}_{r'} = \mathbf{F}_r(\mathfrak{p}_i) = \mathbf{F}_r(Q)$ for any $Q \in |\mathfrak{p}_i|$. This homomorphism is surjective with kernel denoted $I_{\mathfrak{p}_i}$, the *inertia group at* $\mathfrak{p}_i$. It turns out that the order of the inertia group is e, the ramification index of $\mathfrak{p}_i$. When $e = 1$, there is a distinguished element of $D_{\mathfrak{p}_i}$, namely the one that maps to the r-power Frobenius in $\mathrm{Gal}(\mathbf{F}_{r'}/\mathbf{F}_r)$. When $e > 1$ we get a distinguished coset of $I_{\mathfrak{p}_i}$ in $D_{\mathfrak{p}_i}$. Changing the choice of $\mathfrak{p}_i$ changes $D_{\mathfrak{p}_i}$, $I_{\mathfrak{p}_i}$ and the distinguished element or coset by conjugation by an element of G. Therefore, we get a well-defined conjugacy class in G depending only on $\mathbf{F}_r$ and P which we denote $\mathrm{Fr}_{\mathbf{F}_r,P}$. Similarly, we write $D_{\mathbf{F}_r,P}$ and $I_{\mathbf{F}_r,P}$ for the conjugacy classes of subgroups of G defined as above. It is not hard to check that $\mathrm{Fr}_{\mathbf{F}_{r^n},P} = \mathrm{Fr}^n_{\mathbf{F}_r,P}$.

One also associates decomposition and inertia subgroups and a Frobenius element to a prime $\mathfrak{p}$ of $\mathcal{C}$ as follows: we let $\mathbf{F}_r$ be the residue field at $\mathfrak{p}$ and choose $P \in |\mathfrak{p}|$ and then set $D_{\mathfrak{p}} = D_{\mathbf{F}_r,P}$, $I_{\mathfrak{p}} = I_{\mathbf{F}_r,P}$, and $\mathrm{Fr}_{\mathfrak{p}} = \mathrm{Fr}_{\mathbf{F}_r,P}$. The resulting conjugacy classes are well-defined independently of the choice of P. This Frobenius is more analogous to the Frobenius element considered over number fields.

Example: Let $\mathcal{C} \to \mathcal{C}' = \mathbf{P}^1$ be the morphism considered in Example (2) in Section 1.4 and again in Example (2) in Section 1.8. This is a Galois covering with group $G = \mathbf{Z}/2\mathbf{Z}$. If $a \in \mathbf{F}_r$ is such that $a^3-1 \neq 0$, and $P \in \mathbf{P}^1$ is the point $[1 : a]$, then the Frobenius class Fr_P is 1 if $a^3 - 1$ is a square in $\mathbf{F}_r$ and is -1 if it is not a square. If $\mathfrak{p}$ is an $\mathbf{F}_q$-rational prime divisor of $\mathbf{P}^1$, then $\mathrm{Fr}_{\mathfrak{p}}$ is 1 if $\mathfrak{p}$ splits and is -1 if $\mathfrak{p}$ is inert.

The definitions of decomposition and inertia subgroups and Frobenius elements extend to infinite Galois extensions in exactly the same way as in the number field context.

1.10. **Cebotarev equidistribution.** The classical Cebotarev density theorem says roughly that Frobenius elements are equidistributed in the Galois group of a Galois extension of number fields. To discuss a function field analogue, we keep the notations of the last section so that F/F' is a geometric Galois extension of function fields over $\mathbf{F}_q$, with corresponding morphism of curves $\mathcal{C} \to \mathcal{C}'$ defined over $\mathbf{F}_q$. We consider the distribution of Frobenius conjugacy classes $\mathrm{Fr}_{\mathbf{F}_r,P}$ as P varies over $\mathbf{F}_r$-rational points of $\mathcal{C}'$ for large r.

One analogue of the Cebotarev density theorem for function fields says that the Frobenius classes become equidistributed as r tends to infinity. More precisely, if $C \subset G$ is a conjugacy class, then

$$\lim_{r\to\infty} \frac{|\{P \in \mathcal{C}'(\mathbf{F}_r) | \, \mathrm{Fr}_{\mathbf{F}_r,P} \in C\}|}{|\{P \in \mathcal{C}'(\mathbf{F}_r)\}|} = \frac{|C|}{|G|}$$

where r tends to infinity through powers of q. A useful way to rephrase this is to consider conjugation invariant functions f on G. It make sense to evaluate such a function on a Frobenius conjugacy class and we have

$$\lim_{r\to\infty} \left| \frac{1}{|\mathcal{C}'(\mathbf{F}_r)|} \sum_{P\in\mathcal{C}'(\mathbf{F}_r)} f(\mathrm{Fr}_{\mathbf{F}_r,P}) - \frac{1}{|G|}\sum_{g\in G} f(g) \right| = 0$$

There is a more precise statement about the rate of convergence: given data as above, there exists a constant depending only on F/F'

and f such that for all powers r of q,

$$\left| \frac{1}{|\mathcal{C}'(\mathbf{F}_r)|} \sum_{P \in \mathcal{C}'(\mathbf{F}_r)} f(\mathrm{Fr}_{\mathbf{F}_r, P}) - \frac{1}{|G|} \sum_{g \in G} f(g) \right| \leq C r^{-1/2}.$$

The constant C can be made quite explicit in terms of the representation theory of G and the expansion of f in terms of characters. See [KS99b, 9.7.11-13] for details.

As a very simple example of what this means in down-to-earth terms, we return to Example (2) of Section 1.8. In that context, Cebotarev equidistribution says that for large r, for about $1/2$ of the elements $a \in \mathbf{F}_r$, $a^3 - 1$ is a square and for about $1/2$ of the a, it is not a square.

2. ζ-functions and L-functions

In this section we define ζ- and L-functions, give some examples, and discuss the spectral interpretation. Warning: we use a non-standard, radically simplified notation for certain cohomology groups. See Section 4 for references with a more complete treatment.

2.1. **The ζ-function of a curve.** Let F be a function field with field of constants $\mathbf{F}_q$. Let $\mathcal{C}$ be the corresponding curve and denote by $\mathcal{C}^0$ the set of $\mathbf{F}_q$-rational prime divisors of $\mathcal{C}$. We define the zeta-function of $\mathcal{C}$ in analogy with the Riemann zeta-function:

$$\zeta(\mathcal{C}, s) = \prod_{\mathfrak{p} \in \mathcal{C}^0} (1 - N\mathfrak{p}^{-s})^{-1}$$

where $N\mathfrak{p} = q^{\deg \mathfrak{p}}$ is the number of elements in the residue field at $\mathfrak{p}$. (This function depends not just on the curve $\mathcal{C}$ but also on the constant field $\mathbf{F}_q$ and when we want to make this dependence explicit, we write $\zeta(\mathcal{C}/\mathbf{F}_q, s)$.)

If C_m denotes the number of primes in $\mathcal{C}^0$ of degree m and N_n denotes the number of points of $\mathcal{C}$ defined over $\mathbf{F}_{q^n}$, then we have

$$N_n = \sum_{m|n} m C_m.$$

Rearranging formally, we find that

$$\zeta(\mathcal{C}, s) = \exp\left(\sum_{n=1}^{\infty} \frac{N_n}{n} q^{-ns} \right)$$

which makes the diophantine interest of ζ quite visible.

The product defining $\zeta(\mathcal{C}, s)$ and the rearranged sum converge absolutely in the region Re $s > 1$. Using the Riemann-Roch theorem, one can show that $\zeta(\mathcal{C},s)$ extends to a meromorphic function on all of $\mathbf{C}$, with simple poles at $s=1$ and $s=0$ and holomorphic elsewhere, and that it satisfies a functional equation relating s and $1-s$. (There are no Γ-factors because the product defining ζ is over all places of F.) More precisely,

$$q^{-s(1-g)}\zeta(\mathcal{C},s)=q^{(s-1)(1-g)}\zeta(\mathcal{C},1-s)$$

where g is the genus of $\mathcal{C}$.

Here are some examples: If F is the rational function field with constant field $\mathbf{F}_q$, so that $\mathcal{C}=\mathbf{P}^1$, then $N_n=q^n+1$ and so

$$\zeta(\mathcal{C},s)=\frac{1}{(1-q^{-s})(1-q^{1-s})}.$$

Let $\mathcal{C}$ be the curve with affine equation $y^2=x^3-x$ over $\mathbf{F}_p$ where $p\equiv 3 \pmod 4$ and $p>3$. Using the fact that -1 is not a square modulo p, it is easy to check that the number of points on E over $\mathbf{F}_p$ is $p+1$ and more generally, if f is odd, the number of points on E with coordinates in $\mathbf{F}_{p^f}$ is p^f+1. (One considers pairs $x=a$ and $x=-a$, excluding $x=0$ and ∞. Since -1 is not a square in $\mathbf{F}_q$, x^3-x is a square for exactly one of $x=a$ or $x=-a$; when it is a square there are two values of y with $y^2=x^3-x$ and none when it is not. Thus the number of solutions with finite non-zero a is $q-1$ and the total number of solutions is $q+1$.) A somewhat more elaborate argument using exponential sums allows one to show that for even f, the number of solutions over $\mathbf{F}_{p^f}$ is $p^f+1-2(-p)^{f/2}$. (See Koblitz [Kob93, II.2] or Ireland and Rosen [IR90, Chap. 18] for a nice exposition of this argument.) Using the expression for ζ in terms of the N_n, we conclude that

$$\zeta(\mathcal{C}/\mathbf{F}_p,s)=\frac{(1-\sqrt{-p}p^{-s})(1+\sqrt{-p}p^{-s})}{(1-p^{-s})(1-p^{1-s})}=\frac{1+p^{1-2s}}{(1-p^{-s})(1-p^{1-s})}.$$

As a third example we assume that $p>2$ and $q=p^f\equiv 1 \pmod 3$ and consider the curve $\mathcal{C}$ with affine equation $y^3=x^4-x^2$, or rather the smooth, projective curve obtained from this one by desingularization. (This curve is singular at $(x,y)=(0,0)$, but there is exactly one point over this one in the smooth curve, so for the purposes of counting points we may ignore this.) This curve has genus $g=2$.

Let $\lambda:{\mathbf{F}_q}^\times\to\mathbf{C}^\times$ be a character of order exactly 6 and for $a=1,2,4,5$ define

$$J_a=\sum_{\substack{x\in\mathbf{F}_q\\ x\neq 0,1}}\lambda^a(x(1-x)).$$

It is not hard to check that $|J_i| = q^{1/2}$ and $J_5 = \overline{J}_1$, $J_4 = \overline{J}_2$. Using arguments similar to those in Koblitz or Ireland and Rosen, one verifies that the number of points on $\mathcal{C}$ over $\mathbf{F}_{q^f}$ is $q^f + 1 - \sum_{a\in\{1,2,4,5\}} J_a^f$. This implies that

$$\zeta(\mathcal{C}/\mathbf{F}_q, s) = \frac{\prod_{a\in\{1,2,4,5\}}(1 - J_a q^{-s})}{(1-q^{-s})(1-q^{1-s})}.$$

In general, if $\mathcal{C}$ has genus g then $\zeta(\mathcal{C}, s)$ has the form

$$\frac{P(q^{-s})}{(1-q^{-s})(1-q^{1-s})}$$

where P is a polynomial of degree $2g$ with integer coefficients and constant term 1. Writing $P(T) = \prod_{i=1}^{2g}(1-\alpha_i T)$, the functional equation for ζ is equivalent to the fact that the set of inverse roots α_i is invariant under $\alpha_i \mapsto q/\alpha_i$. Moreover, ζ satisfies an analogue of the Riemann hypothesis: all of the inverse roots α_i have absolute value $q^{-1/2}$ and so the zeros of ζ lie on the line $\Re(s) = 1/2$. These results were proven in general by Weil in [Wei48].

More generally, one can define a zeta function for any variety defined over a finite field via a product or exponentiated sum as above. If X is smooth and complete of dimension d, then one knows that $\zeta(X, s)$ is a rational function in q^{-s} of a very special form. More precisely,

$$\zeta(X, s) = \frac{P_1(q^{-s})P_3(q^{-s})\cdots P_{2d-1}(q^{-s})}{P_0(q^{-s})P_2(q^{-s})\cdots P_{2d}(q^{-s})}$$

where each P_i is a polynomial with integer coefficients all of whose inverse roots have complex absolute value $q^{i/2}$ (an analogue of the Riemann hypothesis). Moreover, if the inverse roots of P_i are $\alpha_1, \ldots, \alpha_k$, then the inverse roots of P_{2d-i} are $q^d/\alpha_1, \ldots, q^d/\alpha_k$ and so $\zeta(X, s)$ extends to a meromorphic function in the plane and satisfies a functional equation for $s \to d - s$. These properties of the ζ-function were conjectured by Weil in 1949 and proved in full generality by Deligne in 1974.

2.2. **Spectral interpretation of ζ-functions.** Already at the time he made his famous conjectures, Weil envisioned a cohomological explanation for the conjectured properties of the zeta function. This was provided in important cases by Weil and later in full generality by Grothendieck, Deligne, and collaborators.

We fix an auxiliary prime ℓ not equal to the characteristic of $\mathbf{F}_q$. Attached to a curve $\mathcal{C}$ over a finite field $\mathbf{F}_q$ are finite-dimensional $\mathbf{Q}_\ell$-vector spaces $H^0(\mathcal{C})$, $H^1(\mathcal{C})$ and $H^2(\mathcal{C})$ each equipped with an action of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q)$. The ζ-function of $\mathcal{C}$ then has an interpretation in terms of the spectrum of the q-power Frobenius Fr_q, which is a generator of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q)$, namely

$$\zeta(\mathcal{C},s)=\frac{P_1(q^{-s})}{P_0(q^{-s})P_2(q^{-s})}$$

where

$$P_i(T)=\det\left(1-T\,\mathrm{Fr}_q\,|H^i(\mathcal{C})\right).$$

(It turns out that the eigenvalues of Fr_q are algebraic numbers, so that we may interpret them as complex numbers. In fact the coefficients of the reversed characteristic polynomials appearing here are integers, so there is no dependence on an embeddings of $\overline{\mathbf{Q}}$ into $\overline{\mathbf{Q}}_\ell$ and $\mathbf{C}$.)

It turns out that $H^0(\mathcal{C})$ is one-dimensional with trivial action of Fr_q, $H^2(\mathcal{C})$ is one-dimensional with Fr_q acting by multiplication by q and $H^1(\mathcal{C})$ is $2g$-dimensional, where g is the genus of g. This shows that $\zeta(\mathcal{C},s)$ is a rational function in q^{-s} of the form mentioned in the last section.

The functional equation is a manifestation of a Poincaré duality: there are pairings $H^i(\mathcal{C})\times H^{2-i}(\mathcal{C})\to H^2(\mathcal{C})$ compatible with the actions of Fr_q and this shows that the eigenvalues of Fr_q on H^i are q divided by the eigenvalues of Fr_q on H^{2-i}, which is the content of the functional equation.

The Riemann hypothesis, namely that the zeros of $\zeta(\mathcal{C},s)$ lie on the line $\Re(s)=1/2$, is equivalent to the statement that the eigenvalues of Fr_q on $H^1(\mathcal{C})$ have complex absolute value $q^{1/2}$.

All of the above generalizes to smooth proper varieties of any dimension over $\mathbf{F}_q$. For an X of dimension d, there are finite-dimensional $\mathbf{Q}_\ell$-vector spaces $H^0(X),\dots,H^{2d}(X)$ with an action of Fr_q; $H^0(X)$ is one-dimensional with trivial Fr_q action and $H^{2d}(X)$ is one-dimensional with Fr_q acting by multiplication by q^d. There is a Poincaré duality pairing $H^i(X)\times H^{2d-i}(X)\to H^{2d}(X)$ which is non-degenerate and compatible with the Frobenius actions. Finally, the eigenvalues of Fr_q on $H^i(X)$ are algebraic integers with absolute value $q^{i/2}$ in every complex embedding.

2.3. **Examples of L-functions.** Just as in the number field case, we can define L-functions associated to representations of the absolute Galois group of a function field. Before giving the general definitions, we consider three examples.

First, let F be a quadratic extension of $\mathbf{F}_q(t)$, corresponding to a branched cover $\mathcal{C}\to\mathbf{P}^1$ of degree 2. Since $F/\mathbf{F}_q(t)$ is a Galois extension with group $\{\pm 1\}$, we get a quadratic character

$$\chi : \mathrm{Gal}(\overline{\mathbf{F}_q(t)}/\mathbf{F}_q(t)) \to \mathrm{Gal}(F/\mathbf{F}_q(t)) \to \{\pm 1\}.$$

Let us define the L-function of χ as

$$L(\chi, s) = \prod_{\mathfrak{p}\in(\mathbf{P}^1)^0} (1 - \chi(\mathfrak{p}) N\mathfrak{p}^{-s})^{-1}$$

where for unramified $\mathfrak{p}$, $\chi(\mathfrak{p}) = \chi(\mathrm{Fr}_{\mathfrak{p}})$ is 1 if $\mathfrak{p}$ splits in F and -1 if $\mathfrak{p}$ is inert; we set $\chi(\mathfrak{p}) = 0$ if $\mathfrak{p}$ is ramified in F. An elementary (Euler-factor by Euler-factor) computation shows that

$$\zeta(\mathcal{C}, s) = \zeta(\mathbf{P}^1, s) L(\chi, s).$$

On the other hand,

$$\zeta(\mathcal{C}, s) = \frac{P(q^{-s})}{(1-q^{-s})(1-q^{1-s})}$$

and

$$\zeta(\mathbf{P}^1, s) = \frac{1}{(1-q^{-s})(1-q^{1-s})}$$

and so

$$L(\chi, s) = P(q^{-s}).$$

The functional equation for ζ is equivalent to

$$q^{gs} L(\chi, s) = q^{g(1-s)} L(\chi, 1-s).$$

This applies in particular to the curve $y^2 = x^3 - x$ considered above: we view it as a degree two cover of the t-line by $(x, y) \mapsto t = x$. It follows that

$$L(\chi, s) = (1 - \sqrt{-p}p^{-s})(1 + \sqrt{-p}p^{-s}) = 1 + p^{1-2s}.$$

For a second class of examples, consider a Galois extension $F/\mathbf{F}_q(t)$ with Galois group $\mathbf{Z}/d\mathbf{Z}$, corresponding to a degree d cyclic covering of curves $\mathcal{C} \to \mathbf{P}^1$. Let $\chi : \mathrm{Gal}(\overline{\mathbf{F}_q(t)}/\mathbf{F}_q(t)) \to \mathrm{Gal}(F/\mathbf{F}_q(t)) \to \mu_d \subset \overline{\mathbf{Q}}^\times$ be a complex valued character of order exactly d and for $i = 1, \ldots, d-1$ define

$$L(\chi^i, s) = \prod_{\mathfrak{p}\in(\mathbf{P}^1)^0} (1 - \chi^i(\mathfrak{p}) N\mathfrak{p}^{-s})^{-1}$$

where $\chi(\mathfrak{p}) = \chi(\mathrm{Fr}_{\mathfrak{p}})$ for unramified $\mathfrak{p}$ and $\chi(\mathfrak{p}) = 0$ if $\mathfrak{p}$ is ramified in F. Again an elementary calculation shows that

$$\zeta(\mathcal{C}, s) = \zeta(\mathbf{P}^1, s) L(\chi, s) L(\chi^2, s) \cdots L(\chi^{d-1}, s).$$

It turns out that each $L(\chi^i, s)$ for $i = 1, \ldots, d-1$ is a polynomial in q^{-s} and their product is the numerator $P(q^{-s})$ of $\zeta(\mathcal{C}, s)$.

For $d > 2$ a new phenomenon becomes apparent: the functional equation links two distinct L-functions. More precisely, we have

$$q^{N_i s/2} L(\chi^i, s) = \epsilon q^{N_i(1-s)/2} L(\chi^{-i}, 1-s)$$

where $N_i = N_{-i}$ is the degree of $L(\chi^i, s)$ as a polynomial in q^{-s} and ϵ is a complex number of absolute value 1. This will be important later when we discuss symmetry types.

As a specific example of this type, we consider the curve $\mathcal{C}$ defined by $y^3 = x^4 - x^2$, discussed above, viewed as a Galois cover of $\mathbf{P}^1$ of degree 3 via $(x, y) \mapsto t = x$. For a suitable choice of character $\chi : \mathrm{Gal}(F/\mathbf{F}_q(t)) \to \mu_3$, we have $L(\chi, s) = (1 - J_1 q^{-s})(1 - J_4 q^{-s})$ and $L(\chi^2, s) = (1 - J_2 q^{-s})(1 - J_5 q^{-s})$.

A third, more elaborate, class of examples comes from elliptic curves. Let E be an elliptic curve defined over F. This could be given, for example, by a Weierstrass equation

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6$$

where the a_i are in F. If E has good reduction at a place $\mathfrak{p}$ of F, we define a local Euler factor by

$$L_{\mathfrak{p}}(E, s) = (1 - a_{\mathfrak{p}} q_{\mathfrak{p}}^{-s} + q_{\mathfrak{p}}^{1-2s})$$

where $q_{\mathfrak{p}}$ is the cardinality of the residue field at $\mathfrak{p}$ and $q_{\mathfrak{p}} - a_{\mathfrak{p}} + 1$ is the number of points on the reduction of E at $\mathfrak{p}$. If E has bad reduction at $\mathfrak{p}$, we define a local factor by

$$L_{\mathfrak{p}}(E, s) = \begin{cases} 1 - q_{\mathfrak{p}}^{-s} & \text{if } E \text{ has split multiplicative reduction at } \mathfrak{p} \\ 1 + q_{\mathfrak{p}}^{-s} & \text{if } E \text{ has non-split multiplicative reduction at } \mathfrak{p} \\ 1 & \text{if } E \text{ has additive reduction at } \mathfrak{p}. \end{cases}$$

Then we define the global (Hasse-Weil) L-function of E as

$$L(E, s) = \prod_{\mathfrak{p}} L_{\mathfrak{p}}(E, s)^{-1}.$$

This L-function turns out to be a rational function in q^{-s} and it satisfies a functional equation for $s \to 2 - s$. More precisely, if E is not isomorphic to an elliptic curve defined over $\mathbf{F}_q$, then $L(E, s)$ is a polynomial in q^{-s} whose degree is determined by the genus of the curve corresponding to F and the places of bad reduction of E. In this case,

$$L(E, s) = \prod_{i=1}^{N} (1 - \alpha_i q^{-s})$$

where the set of inverse roots α_i is invariant under $\alpha_i \mapsto q^2/\alpha_i$ and each of them has complex absolute value q. In particular, the zeros of $L(E,s)$ lie on the line $\Re(s) = 1$.

2.4. **L-functions attached to Galois representations.** As in the number field case, over function fields there are two general classes of L-functions, automorphic L-functions attached to automorphic representations (generalizing Dirichlet characters, Hecke characters, etc.) and "motivic" L-functions attached to representations of Galois groups, and a Langlands philosophy which very roughly speaking says that the latter are the same as the former. In the function field setting there is a quite satisfactory understanding of the analytic properties of motivic L-functions which we sketch in this and the following section.

As usual, let $F = \mathbf{F}_q(\mathcal{C})$ be the function field of a curve over $\mathbf{F}_q$. We fix a prime ℓ and write E for a finite extension of $\mathbf{Q}_\ell$ which we may expand as necessary in the course of the discussion. The basic input data is a representation

$$\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_n(E)$$

which is continuous (for the Krull topology on $\mathrm{Gal}(\overline{F}/F)$ and the ℓ-adic topology on $\mathrm{GL}_n(E)$) and unramified outside a finite set of places of F. The latter means that for all but finitely many primes $\mathfrak{p}$, $\rho(I_\mathfrak{p}) = \{1\}$ where $I_\mathfrak{p}$ is the inertia subgroup at $\mathfrak{p}$. We assume that ρ is absolutely irreducible, i.e., is reducible even after extending scalars to $\overline{E}$. We also assume that ρ has a weight $w \in \mathbf{Z}$, which means that for every unramified prime $\mathfrak{p}$, all of the eigenvalues of $\rho(\mathrm{Fr}_\mathfrak{p})$ are algebraic integers and have absolute value $q^{w/2}$ in every complex embedding.

Given ρ, we define an L-function by

$$L(\rho, s) = \prod_\mathfrak{p} \det\left(1 - \rho(\mathrm{Fr}_\mathfrak{p}) N\mathfrak{p}^{-s} \Big| \, (E^n)^{I_\mathfrak{p}}\right)^{-1}.$$

Here $N\mathfrak{p}$ is the cardinality of the residue field at $\mathfrak{p}$, $I_\mathfrak{p}$ is the inertia group at $\mathfrak{p}$, and $(E^n)^{I_\mathfrak{p}}$ denotes the subspace of E^n where $I_\mathfrak{p}$ acts (via ρ) trivially; for almost all $\mathfrak{p}$ this will just be E^n itself. On the space of invariants $(E^n)^{I_\mathfrak{p}}$ there is a well-defined action of the Frobenius elements $\mathrm{Fr}_\mathfrak{p}$ and the local factors above are the reciprocals of the reversed characteristic polynomials of the action of $N\mathfrak{p}^{-s}$ times $\rho(\mathrm{Fr}_\mathfrak{p})$.

Easy estimates show that the product defining $L(\rho, s)$ converges absolutely in the region $\Re(s) > 1 + w/2$, uniformly on compact subsets, and so defines a holomorphic function there. As we will see in the next section, $L(\rho, s)$ has a meromorphic continuation to all of $\mathbf{C}$ which is entire if and only if ρ restricted to $\mathrm{Gal}(\overline{F}/\overline{\mathbf{F}}_p F)$ contains no copies of the trivial representation. In general, $L(\rho,s)$ satisfies a functional equation

$$L(\rho,s)=\epsilon(\rho,s)L(\rho^\vee,1-s)$$

where $\rho^\vee$ is the dual representation and $\epsilon(\rho,s)$ is an entire function with $\epsilon(\rho,1/2)$ a complex number of absolute value 1.

The attentive reader may be distressed by the apparent mixture of ℓ-adic and complex numbers in the definition of $L(\rho,s)$. To make things precise, we fix embeddings $\overline{\mathbf{Q}}\hookrightarrow\overline{\mathbf{Q}}_\ell$ and $\overline{\mathbf{Q}}\hookrightarrow\mathbf{C}$; since we assumed that the eigenvalues of $\rho(\mathrm{Fr}_\mathfrak{p})$ are algebraic numbers we may use the embeddings to regard the coefficients of the reversed characteristic polynomials as complex numbers.

The examples of the previous section can be fit into this general framework as follows. If K/F is a finite Galois extension and $\chi:\mathrm{Gal}(K/F)\to\mu_d\subset E=\mathbf{Q}_\ell(\mu_d)$ is a character, then composing with the natural projection $\mathrm{Gal}(\overline{F}/F)\to\mathrm{Gal}(K/F)$ gives a one-dimensional, absolutely irreducible ℓ-adic representation satisfying our hypotheses. It has weight $w=0$.

The elliptic curve example is somewhat more elaborate. In this case, we consider the ℓ-adic Tate module of E over F, namely $\varprojlim_m E(\overline{F})[\ell^m]$ which is isomorphic to $\mathbf{Z}_\ell^2$. There is an action of $\mathrm{Gal}(\overline{F}/F)$ on this Tate module and as ρ we take the dual of this representation. At a prime $\mathfrak{p}$ where E has good reduction, general ℓ-adic results show that the reversed characteristic polynomial of $\mathrm{Fr}_\mathfrak{p}$ is just the reversed characteristic polynomial of the $N\mathfrak{p}$-power Frobenius on the group $H^1(E \pmod{\mathfrak{p}})$ mentioned in the discussion of zeta functions. In particular, the coefficients of the local zeta function are given in terms of the number of points on the reduction of E at $\mathfrak{p}$ by the recipe mentioned in the previous section. Something similar, albeit more involved, happens at the places of bad reduction.

2.5. **Spectral interpretation of L-functions.** There is a spectral interpretation of L-functions which is quite parallel to that of ζ-functions—the key is to think of a representation ρ as providing coefficients for a cohomology theory. Of course we cannot explain the details here, but the idea is this: given ρ, we have cohomology groups $H^i(\mathcal{C},\rho)$ ($i=0,1,2$) which are finite-dimensional E-vector spaces with an action of $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q)$. (For experts, we are taking the lisse sheaf on an open subset of $\mathcal{C}$ associated to ρ, forming its middle extension on $\mathcal{C}$, and taking cohomology on $\mathcal{C}\times\operatorname{Spec}\overline{\mathbf{F}}_p$.) Then

$$L(\rho,s)=\frac{P_1(q^{-s})}{P_0(q^{-s})P_2(q^{-s})}$$

where

$$P_i(T) = \det\left(1 - T\,\mathrm{Fr}_q\,|H^i(\mathcal{C},\rho)\right).$$

If ρ has weight w then the eigenvalues of Fr_q on $H^i(\mathcal{C},\rho)$ are algebraic integers with absolute value $q^{(i+w)/2}$ in every complex embedding. Poincaré duality takes the form

$$H^i(\mathcal{C},\rho) \times H^{2-i}(\mathcal{C},\rho^\vee) \to H^2(\mathcal{C},\rho\otimes\rho^\vee) \to H^2(\mathcal{C}).$$

When ρ restricted to $\mathrm{Gal}(\overline{F}/\overline{\mathbf{F}}_pF)$ has no trivial factors, then $H^0(\mathcal{C},\rho)$ and $H^2(\mathcal{C},\rho)$ vanish and so the L-function is a polynomial in q^{-s} whose degree is just the dimension of $H^1(\mathcal{C},\rho)$. This dimension can be calculated in terms of the dimension and ramification properties of ρ and the genus of $\mathcal{C}$.

2.6. **Symmetries.** For many interesting representations ρ, there is additional structure coming from the fact that the space where ρ acts admits a Galois-equivariant pairing (at least up to a twist). More precisely, suppose given an absolutely irreducible $\rho : \mathrm{Gal}(\overline{F}/F) \to \mathrm{GL}_n(E)$. Naively we might ask for a pairing

$$\langle\cdot,\cdot\rangle : E^n \times E^n \to E$$

such that $\langle\rho(g)v,\rho(g)v'\rangle = \langle v,v'\rangle$ for all $g\in\mathrm{Gal}(\overline{F}/F)$, but this is not possible when the weight of ρ is non-zero. Instead we ask that

$$\langle\rho(g)v,\rho(g)v'\rangle = \chi_\ell(g)^w\langle v,v'\rangle$$

where $\chi_\ell(g)$ gives the action of g on ℓ-power roots of unity: $\zeta_{\ell^n}^g = \zeta_{\ell^n}^{\chi_\ell(g)}$ for all $\zeta_{\ell^n}\in\mu_{\ell^n}$. When a non-zero (and thus non-degenerate) such pairing exists, we say that ρ is self-dual of weight w. Moreover, the pairing must be either symmetric ($\langle v,v'\rangle = \langle v',v\rangle$) or skew symmetric ($\langle v,v'\rangle = -\langle v',v\rangle$); we say that ρ is orthogonally self-dual or symplectically self-dual respectively.

For example, a finite order character $\chi : \mathrm{Gal}(\overline{F}/F) \to \mu_d$ is self-dual if and only if it is of order 2, in which case it is orthogonally self-dual of weight 0. The representation of $\mathrm{Gal}(\overline{F}/F)$ on the dual of the Tate module of an elliptic curve over F is symplectically self-dual of weight 1.

When ρ self-dual, then so is $H^1(\mathcal{C},\rho)$, but with the opposite sign and weight $w+1$. In other words, when ρ is orthogonally (resp. symplectically) self-dual, then there is a skew-symmetric (resp. symmetric) pairing on $H^1(\mathcal{C},\rho)$ which satisfies $\langle\mathrm{Fr}_q\ v,\mathrm{Fr}_q\ v'\rangle = q^{w+1}\langle v,v'\rangle$.

Extending E if necessary, we may choose a basis of $H^1(\mathcal{C},\rho)$ in which the matrix of the form is the standard one times q^{w+1} and then the matrix of Frobenius in this basis will be $q^{(w+1)/2}$ times an orthogonal or symplectic matrix. Thus extra structure on ρ puts severe restrictions on the action of Frobenius.

At the level of L-functions, these restrictions are reflected in the functional equations: when ρ is symplectically self-dual, the sign in the functional equation is ± 1 (so that the sign sometimes forces vanishing at the central point) whereas when ρ is orthogonally self-dual, the sign in the functional equation is always $+1$ (so that the order of zero at the central point is even).

Note that when ρ is not self-dual, then the Frobenius matrix is *a priori* $q^{(w+1)/2}$ times a general matrix in GL and the functional equation relates two different L-functions and so cannot force zeros at the central point.

3. Families of L-functions

In this section, we come to the *raison d'être* of the article, namely an explanation of how families of L-functions over function fields give rise to well-distributed collections of matrices in classical groups. Rather than attempting to make precise general definitions, we consider several examples which we hope will make the key points clear.

3.1. **Arithmetic and geometric families.**

Let us fix a finite field $\mathbf{F}_q$ and consider all quadratic extensions of the rational function field $\mathbf{F}_q(t)$, or equivalently, all quadratic characters

$$\chi : \mathrm{Gal}(\overline{\mathbf{F}_q(t)}/\mathbf{F}_q(t)) \to \{\pm 1\}.$$

We exclude as trivial the unique character χ factoring through $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q)$ which corresponds to the extension $\mathbf{F}_{q^2}(t)$. We want to make statistical statements about the L-functions $L(\chi, s)$ and to do so, the most natural way to partially order them is by the genus of the corresponding field F or what amounts to the same thing, the degree of the conductor of χ.

To keep things as simple as possible, we assume that the characteristic p of $\mathbf{F}_q$ is > 2. In this case, the conductor of χ can be thought of as the set of $\mathfrak{p}$ where χ is ramified and the degree of the conductor of χ is just the sum of the degrees of the places $\mathfrak{p}$ in the conductor. The connection with the genus is given by the Riemann-Hurwitz formula: $g = (\deg(\mathrm{Cond}(\chi)) - 2)/2$.

There are finitely many χ with conductor $\leq N$ (the number is of the order q^N as $N \to \infty$) and so we may consider some quantity associated to $L(\chi, s)$, such as the height of its lowest zero or the spacings between zeros, average over those χ of conductor $\leq N$, and then take a limit as $N \to \infty$. This set-up is entirely analogous to the situation over $\mathbf{Q}$ or a number field, and apparently just as inaccessible. Katz and Sarnak [KS99a] have made several conjectures in this direction which are open. We call this family and ones like it *arithmetic*.

Considerably more can be done in the function field situation if we change the problem slightly. Namely, let us give ourselves the freedom to vary the constant field $\mathbf{F}_q$ as well: We consider quadratic extensions of $\mathbf{F}_{q^n}(t)$ or equivalently quadratic characters $\chi : \mathrm{Gal}(\overline{\mathbf{F}_{q^n}(t)}/\mathbf{F}_{q^n}(t)) \to \{\pm 1\}$, again excluding the character corresponding to $\mathbf{F}_{q^{2n}}(t)$. The number of such characters with conductor of degree $\leq N$ is of the order q^{nN}. We form the average over this set of some quantity associated to $L(\chi, s)$ and then take a limit as $n \to \infty$. This already gives interesting statements, but we may also take a second limit as $N \to \infty$. The advantage of first passing to the limit in n is that we get an infinite collection of L-functions *parameterized by a single algebraic variety.* For this reason we call such families *geometric*.

Let us explain how this parameterization comes about, still assuming for simplicity that $p > 2$. In this case, any quadratic extension F of $\mathbf{F}_{q^n}(t)$ can be obtained by adjoining the square root of a polynomial $f \in \mathbf{F}_{q^n}[t]$. If f is square free the degree of the conductor of χ is essentially the degree of f. (More precisely, it is $\deg(f)$ if $\deg(f)$ is even and $\deg(f) + 1$ if $\deg(f)$ is odd.) For simplicity we restrict to monic polynomials f; the set of monic polynomials of degree N is naturally an affine space of dimension N (using the coefficients of the polynomial as coordinates) and the set of square-free monic polynomials is a Zariski open subset $X \subset \mathbf{A}^N$. Thus we have a natural bijection between certain quadratic characters of conductor N of $\mathrm{Gal}(\overline{\mathbf{F}_{q^n}(t)}/\mathbf{F}_{q^n}(t))$ and $X(\mathbf{F}_{q^n})$, the points of X with coordinates in $\mathbf{F}_{q^n}$. We write χ_f for the character associated to $f \in X(\mathbf{F}_{q^n})$. This geometric structure allows one to bring the powerful tools of arithmetical algebraic geometry to bear, with decisive results.

3.2. **Variation of L-functions.** We continue with the example of L-functions attached to quadratic characters over $\mathbf{F}_{q^n}(t)$. As we explained in Section 2, $L(\chi_f, s)$ is the numerator of the zeta-function of the hyperelliptic curve $\mathcal{C} \to \mathbf{P}^1$ corresponding to the quadratic extension $F = \mathbf{F}_{q^n}(\sqrt{f})/\mathbf{F}_{q^n}(t)$ cut out by χ_f and it can be computed as the characteristic polynomial of Frobenius on a cohomology group. In particular, there is a symplectic matrix $A_f \in \mathrm{Sp}_{2g}(\mathbf{Q}_\ell)$, well-defined up to conjugacy, such that $L(\chi_f, s) = \det(1 - q^{n(1/2-s)}A_f)$. Thus we have a map from $X(\mathbf{F}_{q^n})$ to conjugacy classes of symplectic matrices.

(The reader uncomfortable with cohomology may proceed as follows: for each point in $f \in X(\mathbf{F}_{q^n})$ we may form the corresponding L-function

$L(\chi_f, s) = \prod(1 - \alpha_i q^{n(1/2-s)})$. The α_i are algebraic integers with absolute value 1 in any complex embedding and the collection of them is invariant under $\alpha_i \mapsto \alpha_i^{-1}$. There is thus a well-defined conjugacy class of symplectic matrices A_f so that the α_i are the eigenvalues of A_f. Of course the preceding sentence is equally true with "symplectic" replaced by "orthogonal" or "unitary"; the virtue of the cohomological approach is that it explains why symplectic matrices are the natural choice.)

The first main result is that in a suitable sense, these conjugacy classes become equidistributed as $n \to \infty$. To make this more precise, we use complex matrices and the compact unitary symplectic group USp_{2g}. Namely, we use the fixed embeddings $\overline{\mathbf{Q}} \hookrightarrow \mathbf{C}$ and $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_\ell$ to view ℓ-adic matrices as complex matrices. The Weyl unitarian trick and the Peter-Weyl theorem imply that the conjugacy class of A_f in $\mathrm{Sp}_{2g}(\mathbf{C})$ meets the maximal compact subgroup USp_{2g} in a unique USp_{2g}-conjugacy class. We write θ_f for any element of this class. The statement of equidistribution is that as $n \to \infty$, these classes become equidistributed with respect to Haar measure. More precisely, for any continuous, conjugation invariant function h on USp_{2g}, we have

$$\int_{\mathrm{USp}_{2g}} h\, d\mu_{Haar} = \lim_{n\to\infty} \frac{1}{|X(\mathbf{F}_{q^n})|} \sum_{f \in X(\mathbf{F}_{q^n})} h(\theta_f).$$

There is a more precise statement giving the rate of convergence:

$$\left| \int_{\mathrm{USp}_{2g}} h\, d\mu_{Haar} - \frac{1}{|X(\mathbf{F}_{q^n})|} \sum_{f \in X(\mathbf{F}_{q^n})} h(\theta_f) \right| < Cq^{-n/2}$$

where C is a constant depending only on X and h.

3.3. **Other families.** We consider two other examples of geometric families giving rise to general matrices and orthogonal matrices.

First we consider families of cubic L-series. More precisely, fix an integer d and consider the set of monic polynomials in x of degree d with coefficients in extensions of the finite field $\mathbf{F}_q$ where $q \equiv 1 \pmod 3$. The set of all such is naturally the affine space of dimension d, with coordinates given by the coefficients:

$$f = x^d + a_1 x^{d-1} + \cdots + a_{d-1}x + a_d \quad \leftrightarrow \quad (a_1, \ldots, a_d) \in \mathbf{A}^d(\mathbf{F}_{q^n}).$$

We let $X \subset \mathbf{A}^d$ be the Zariski open subset corresponding to polynomials with distinct roots, so that X is obtained from $\mathbf{A}^d$ by removing the zero set of the discriminant, a polynomial in $a_1, \ldots, a_d$. For each extension $\mathbf{F}_{q^n}$ of $\mathbf{F}_q$ and each $f \in X(\mathbf{F}_{q^n})$, the curve with affine equation $y^3 = f(x)$ is a cubic Galois covering of $\mathbf{P}^1$ corresponding to a cubic Galois extension of function fields $F/\mathbf{F}_{q^n}(t)$. There are two nontrivial characters of $\mathrm{Gal}(F/\mathbf{F}_{q^n}(t))$, which we denoted by χ_f and χ_f^{-1}. (We will not explain the details here, but there is a consistent way to choose which is χ_f and which is $\chi_{f^{-1}}$.) The character χ_f gives rise to an L-function $L(\chi_f, s)$ and, via the cohomological machinery discussed in the previous section, to a well-defined conjugacy class of matrices A_f in $\mathrm{GL}_N(\mathbf{Q}_\ell)$ where $N = d-2$ and, for convenience, $\ell \equiv 1 \pmod 3$. As we noted in Section 2.3, for cubic characters Poincaré duality and the functional equation link two distinct groups or L-functions and so there is no geometric reason for the Frobenius matrices to lie in a small group and in fact they do not. By results of Katz and the general machinery sketched below, for all sufficiently large d, the Frobenius conjugacy classes are equidistributed in an algebraic group containing the algebraic group SL_N over $\mathbf{Q}_\ell$ with finite index. As before, one makes this precise by using embeddings and Lie theory to deduce for each $f \in X(\mathbf{F}_{q^n})$ a well-defined conjugacy class θ_f in a compact Lie group G with $\mathrm{SU}_N \subset G \subset \mathrm{U}_N$ such that

$$L(\chi_f, s) = \det\left(1 - q^{n(1/2-s)}\theta_f\right) = \prod_{i=1}^{N}(1 - \alpha_i q^{n(1/2-s)})$$

where the α_i are the eigenvalues of θ_f. The equidistribution statement is then that

$$\left|\int_G h\, d\mu_{Haar} - \frac{1}{|X(\mathbf{F}_{q^n})|}\sum_{f \in X(\mathbf{F}_{q^n})} h(\theta_f)\right| < Cq^{-n/2}$$

for any continuous, conjugation-invariant function h on G.

For an example of an orthogonal family, we consider the family of quadratic twists of an elliptic curve. More precisely, assume that $p > 3$ and fix an elliptic curve E over $\mathbf{F}_q(t)$ defined by a Weierstrass equation

$$y^2 = x^3 + ax + b$$

with $a, b \in \mathbf{F}_q(t)$. We assume that the j-invariant of E is not in $\mathbf{F}_q$. Fix a degree d. For each monic square-free polynomial $f \in \mathbf{F}_{q^n}[x]$, we may form the quadratic twist E_f of E, with equation

$$fy^2 = x^3 + ax + b \tag{1}$$

and its L-function $L(E_f, s)$. If we assume that the zeros of f are disjoint from the points where E has bad reduction, then the degree of $L(E_f, s)$ as a polynomial in q^n is $N = 2d+c$ where c is a constant depending only on E. Let $X \subset \mathbf{A}^d$ be the Zariski open set whose points over $\mathbf{F}_{q^n}$ are the monic, square-free polynomials $f \in \mathbf{F}_q[x]$ with zeros disjoint from the primes dividing the discriminant of E. The cohomological machinery gives us, for each $f \in X(\mathbf{F}_{q^n})$, an orthogonal matrix $A_f \in \mathrm{O}_N(\mathbf{Q}_\ell)$, well-defined up to conjugacy, such that

$$L(E_f, s) = \det\left(1 - q^{n(1-s)}A_f\right).$$

As before, using the embeddings and Lie theory we deduce a conjugacy class θ_f in the compact group $\mathrm{O}_N(\mathbf{R})$. Under further hypotheses on E which we do not discuss one may conclude that in fact $\theta_f \in \mathrm{SO}_N(\mathbf{R})$. (We make these hypotheses only to simplify the equidistribution statement below.) Results of Katz and Deligne then say that the classes θ_f are equidistributed in the sense that

$$\left|\int_{\mathrm{SO}_N(\mathbf{R})} h\, d\mu_{Haar} - \frac{1}{|X(\mathbf{F}_{q^n})|}\sum_{f\in X(\mathbf{F}_{q^n})} h(\theta_f)\right| < Cq^{-n/2}$$

for any continuous, conjugation-invariant function h on $\mathrm{SO}_N(\mathbf{R})$.

3.4. **Idea of proofs.** We give a very brief sketch of the main ideas behind the proofs of the equidistribution statements above.

The first ingredient is monodromy. Let X be the variety parameterizing the family under study. Then we have the fundamental group $\pi_1(X)$, which is a quotient of the absolute Galois group of the function field of X over $\mathbf{F}_q$ and which gives automorphisms ("deck transformations") of unramified covers of X. There is a subgroup $\pi_1^{geom}(X) \subset \pi_1(X)$ such that

$$\pi_1(X)/\pi_1^{geom}(X) \cong \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_q).$$

The cohomological machinery gives rise to a representation $\rho : \pi_1(X) \to \mathrm{GL}_N(E)$ (here E is some finite extension of $\mathbf{Q}_\ell$) such that for each point $f \in X(\mathbf{F}_{q^n})$ with Frobenius conjugacy class $\mathrm{Fr}_f \in \pi_1(X)$, we have $\rho(\mathrm{Fr}_f) \in \mathrm{GL}_N(E)$ which is the conjugacy class associated to the L-function named by f. Attached to ρ are two monodromy groups $G^{geom} \subset G^{arith}$. These are defined as the Zariski closures of the images of ρ on $\pi_1^{geom}(X)$ and $\pi_1(X)$ respectively. A basic result of Deligne says that G^{geom} is a semi-simple algebraic group over E. When there is extra structure (i.e., a pairing), then we have an a priori containment $G^{arith} \subset \mathrm{Sp}$ or O. In favorable cases one can establish by geometric methods a lower bound Sp or O or SL $\subset G^{geom}$ and therefore equalities $G^{geom} = G^{arith} =$ Sp or O or SL. (Here we are glossing over several technicalities regarding the difference between G^{geom} and G^{arith} and between O and SO.) Part of Katz-Sarnak [KS99b, Chaps. 10-11], most of Katz [Kat02], and several other works of Katz are devoted to these kinds of calculations.

The second main ingredient is a very general equidistribution result of Deligne that says that whatever the arithmetic monodromy group is, the Frobenius classes are equidistributed in it. More precisely, forming classes θ_f in a compact Lie group G associated to G^{arith} and $f \in X(\mathbf{F}_{q^n})$, we have

$$\left| \int_G h \, d\mu_{Haar} - \frac{1}{|X(\mathbf{F}_{q^n})|} \sum_{f \in X(\mathbf{F}_{q^n})} h(\theta_f) \right| < Cq^{-n/2}$$

for all continuous, conjugation-invariant functions h on G. This equidistribution result was proven as a consequence of the Weil conjectures [Del80] and is explained in Katz-Sarnak [KS99b, Chap. 9].

3.5. **Large N limits.** Another part of the story, the part related to classical random matrix theory, relates to statistical measures of eigenvalues in the large N limit. More precisely, given an $N \times N$ unitary matrix with eigenvalues $e^{2\pi i \phi_j}$ with $0 \leq \phi_1 \leq \cdots \leq \phi_N < 1$ one forms a point measure on $\mathbf{R}$ with mass $1/N$ at each of the normalized spacings $N(\phi_2 - \phi_1), N(\phi_3 - \phi_2), \ldots, N(\phi_N - \phi_{N-1}), N(1 + \phi_1 - \phi_N)$. Averaging this measure over U_N (with respect to Haar measure) yields a measure on $\mathbf{R}$ and it turns out that one may take the limit as $N \to \infty$ and arrive at a measure on $\mathbf{R}$ which is absolutely continuous with respect to Lebesgue measure and has a real analytic density function. Similar results hold for other families of classical groups and it turns out that the measure obtained is the same for the symplectic groups Sp_{2N} and the orthogonal groups O_{2N} and O_{2N+1} (where in the latter case one ignores the forced eigenvalue 1).

Katz and Sarnak also consider other statistical measures of eigenvalues, for example the placement of the eigenvalue closest to 1. In this case there is again a scaling limit as $N \to \infty$ but now the resulting measure on $\mathbf{R}$ depends on the family of classical groups considered. For example, the density function for the symplectic family vanishes at 0, indicating that eigenvalues of symplectic matrices are "repelled" from 1, whereas this is not the case for the unitary and orthogonal families.

These results are purely Lie-theoretic and do not involve any algebraic geometry. We will not attempt to give any details here, but simply refer to Katz-Sarnak [KS99b].

For an example of the application of this in the function field context, we consider families X_g as in Section 3.2 parameterizing quadratic characters χ corresponding to curves $\mathcal{C} \to \mathbf{P}^1$ of genus g. Combining equidistribution results with theorems on large N limits, one sees that integrals with respect to the large N limit measure may be computed using Frobenius matrices. More precisely, suppose that ν_1 is the measure on $\mathbf{R}$ associated to the suitably normalized location of the eigenvalue nearest 1 for symplectic matrices. Then we have

$$\int_{\mathbf{R}} h\,d\nu_1 = \lim_{g\to\infty}\lim_{r\to\infty}\frac{1}{|X_g(\mathbf{F}_r)|}\sum_{f\in X_g(\mathbf{F}_r)} h(\phi_1(\theta_f))$$

for all continuous, compactly supported functions h on $\mathbf{R}$, where θ_f is the symplectic matrix associated to f, $\phi_1(\theta_f)$ is the normalized angle of its eigenvalue closest to 1, and r tends to ∞ through powers of q.

The only point we want to make here is that Katz and Sarnak conjecture that results like this should be true without taking the limit over large finite fields. In other words, one should have

$$\int_{\mathbf{R}} h\,d\nu_1 = \lim_{g\to\infty}\frac{1}{|X_g(\mathbf{F}_q)|}\sum_{f\in X_g(\mathbf{F}_q)} h(\phi_1(\theta_f))$$

This conjecture looks quite deep and will probably require new ideas going beyond the cohomological formalism.

3.6. **Applications.** We briefly mention three applications to arithmetic of the ideas around function fields and random matrices.

The first application is to guessing the symmetry type of a family of L-functions over a number field. The idea, roughly speaking, is to find a function field analogue of the given family and inspect the cohomology groups computing the L-functions to see whether there is extra symmetry present. If so, the symmetry group should be O, SO, or Sp; if not then it should contain SL. For example, if one looks at the family of quadratic Dirichlet characters over $\mathbf{Q}$, the function field analog is the family of quadratic characters considered in Section 3.1 and so one expects symplectic symmetries. Of course the symplectic group itself is nowhere in sight in the number field context, but one does find computationally that the statistics of low lying zeros obey the distributions associated with symplectic groups. See Katz-Sarnak [KS99a] for more on this and other examples.

The second application is to an analogue of the Goldfeld conjecture. Roughly speaking, this conjecture asserts that in the family of quadratic twists of an elliptic curve over $\mathbf{Q}$, 50% of the curves should have rank 0 and 50% should have rank 1. The most direct function field analogue would concern twists E_f of a given elliptic curve, as in

Equation 1 above, where $f \in \mathbf{F}_q[x]$ and it would assert that

$$\lim_{d\to\infty} \frac{|\{f \in \mathbf{F}_q[x] |\deg(f) \le d, \text{ ..., and } \operatorname{Rank} E_f(\mathbf{F}_q(t)) = 0\}|}{|\{f \in \mathbf{F}_q[x] |\deg(f) \le d, \text{ and ...}\}|} = \frac{1}{2}$$

where "..." stands for conditions on f, namely that f be square free and have zeros disjoint from the points where E has bad reduction. Similarly for rank 1. There are also conjectures where $\operatorname{Rank} E_f(\mathbf{F}_q(t))$ is replaced by $\operatorname{ord}_{s=1} L(E_f, s)$. These conjectures are completely open, although there are some recent nice examples of Chris Hall [Hal04]. But one can do more by allowing ground field extensions. More precisely, Katz proves in [Kat02] that for large d,

$$\lim_{n\to\infty} \frac{|\{f \in \mathbf{F}_{q^n}[x] |\deg(f) \le d, \text{ ..., and } \operatorname{ord}_{s=1} L(E_f, s) = 0\}|}{|\{f \in \mathbf{F}_{q^n}[x] |\deg(f) \le d, \text{ and ...}\}|} = \frac{1}{2}$$

under the assumption that E has at least one place of multiplicative reduction. (This hypothesis is needed to ensure that the monodromy group is the full orthogonal group O, rather than SO.) Similar results hold for analytic rank 1 and, with suitable modifications, for cases when the monodromy group is SO. One can deduce results for algebraic ranks by using the inequality $\operatorname{Rank} E_f(\mathbf{F}_{q^n}(t)) \le \operatorname{ord}_{s=1} L(E_f, s)$ which is known in the function field case.

The connection between equidistribution and these results is that with respect to Haar measure, 1/2 of the matrices in the orthogonal group have eigenvalue 1 with multiplicity 1 and 1/2 have eigenvalue 1 with multiplicity 0. Thus when the matrices computing the L-functions $L(E_f, s)$ are equidistributed in O, then we expect a simple zero at $s = 1$ for about 1/2 of the f and no zero about 1/2 of the f. See the introduction of [Kat02] for a lucid discussion of these results and the more general context, including cases where the monodromy is SO.

The third application is to non-vanishing results for twists. Given a function field F over $\mathbf{F}_q$, a Galois representation ρ of $\operatorname{Gal}(\overline{F}/F)$, and an integer $d > 1$, one expects to be able to find infinitely many characters $\chi : \operatorname{Gal}(\overline{F}/F) \to \mu_d$ of order d such that $L(\rho \otimes \chi, s)$ does not vanish at some given point $s = s_0$, for example the center of the functional equation. There are few general results in this direction, but if we modify the problem in the usual way then one can prove quite general theorems. Namely, one considers characters χ of $\operatorname{Gal}(\overline{F}/\mathbf{F}_{q^n}F)$ for varying n and with restrictions on the ramification of χ (for example, that the degree of the conductor of χ be less than some D and the ramification of χ be prime to the ramification of ρ). Then under mild hypotheses, one finds the existence of infinitely many characters χ (indeed a set of positive density in a suitable sense) with $L(\rho \otimes \chi, s)$ non-vanishing at a given point s_0. The precise statements involve both non-vanishing and simple vanishing because there may be vanishing forced by functional equations. The connection with equidistribution is that in any of the classical groups O, Sp, or SL, the set of matrices with a given number as eigenvalue has Haar measure zero (except of course for orthogonal matrices and eigenvalues ± 1, which are related to forced zeros). See [Ulm05] for this and more general non-vanishing results.

4. Further reading

In this section we give a personal and perhaps idiosyncratic overview of some of the literature covering the technology implicit in this article.

For an treatment of number theory in function fields very much parallel to classical algebraic number theory and requiring essentially no algebraic geometry, I recommend [Ros02].

For the basic theory of curves over an algebraically closed ground field, a standard reference in use for generations now is [Ful89]. This gives a student-friendly introduction, with all necessary algebraic background and complete details, of the basic theory of curves over an algebraically closed field. Weil's "Foundations" [Wei62] gives a complete and functional theory for algebraic geometry over arbitrary base fields, but it is quite difficult to read and the language has fallen into disuse—the much more powerful and flexible language of schemes is completely dominant. Various books on diophantine geometry and elliptic curves give short accounts, often incomplete or not entirely accurate, of algebraic geometry over general fields. For careful and complete expositions of the theory of curves over general fields, including the ζ-function and the Riemann hypothesis, two popular references are [Gol03] and [Sti93].

For the basics of general, higher dimensional algebraic geometry, there is no better reference than the first part of [Sha77]. This book gives a masterful exposition of the main themes and goals of the field with excellent taste. Part II of this work, on schemes and complex manifolds, is interesting but not sufficiently detailed to be of use as a primary reference.

One can get an excellent idea of some of the analogies between curves over finite fields and rings of integers in number fields, analogies which motivate many of the ideas in modern arithmetical algebraic geometry, from [Lor96]. Studying this work would be a good first step toward schemes, giving the student a valuable stock of examples and tools.

For an introduction to schemes from many points of view, in particular that of number theory, the best reference by far is a long typescript by Mumford and Lang which was meant to be a successor to "The Red

Book" (Springer Lecture Notes 1358) but which was never finished. These notes have excellent discussions of arithmetic schemes, Galois theory of schemes, the various flavors of Frobenius, flatness, issues of inseparability and imperfection, as well as a very down to earth introduction to coherent cohomology. (Some energetic young person would do the community a great service by cleaning up and TeXing these notes.) Some of this material was adapted by Eisenbud and Harris [EH00], including a nice discussion of the functor of points and moduli, but there is much more in the Mumford-Lang notes.

Another excellent and complete reference for the scheme-theoretic tools needed for arithmetical algebraic geometry is [Liu02] which has the virtue of truly being a textbook, with a systematic presentation and lots of exercises.

To my knowledge there is no simple entré into the jungle of étale cohomology. Katz's article [Kat94] in the Motives volume gives a clear and succinct statement of the basics, and Iwaniec and Kowalski [IK04, 11.11] give a short introduction to some basic notions with applications to exponential sums. To go deeper, I recommend SGA$4\frac{1}{2}$ for the main ideas and Milne's masterful text [Mil80], supplemented by the notes on his site (`http://jmilne.org`), for a systematic study.

For wonderful examples of this technology in action I suggest [KS99b] and the papers of Katz referred to there, including [Kat02] (which is the final version of the entry [K-BTBM] in the bibliography of [KS99b]).

Finally, for an in depth introduction to connections between random matrix theory and number theory, I recommend [MHS05], the proceedings of a Newton Institute school on the subject.

References

[Del80] P. Deligne. La conjecture de Weil. II. *Inst. Hautes Études Sci. Publ. Math.*, (52):137– 252, 1980.

[EH00] D. Eisenbud and J. Harris. *The geometry of schemes*, volume 197 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.

[Ful89] W. Fulton. *Algebraic curves*. Advanced Book Classics. Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA, 1989.

[Gol03] D. M. Goldschmidt. *Algebraic functions and projective curves*, volume 215 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2003.

[Hal04] C. Hall. L-functions of twisted Legendre curves. *preprint*, 2004.

[IR90] K. Ireland and M. Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1990.

[IK04] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Kat94] N. M. Katz. Review of l-adic cohomology. In *Motives (Seattle, WA, 1991)*, volume 55 of *Proc. Sympos. Pure Math.*, pages 21– 30. Amer. Math. Soc., Providence, RI, 1994.

[Kat02] N. M. Katz. *Twisted L-functions and monodromy*, volume 150 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2002.

[KS99b] N. M. Katz and P. Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1999.

[KS99a] N. M. Katz and P. Sarnak. Zeroes of zeta functions and symmetry. *Bull. Amer. Math. Soc. (N.S.)*, 36:1– 26, 1999.

[Kob93] N. Koblitz. *Introduction to elliptic curves and modular forms*, volume 97 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993.

[Liu02] Q. Liu. *Algebraic geometry and arithmetic curves*, volume 6 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2002.

[Lor96] D. Lorenzini. *An invitation to arithmetic geometry*, volume 9 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996.

[Mil80] J. S. Milne. *Étale cohomology*, volume 33 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J., 1980.

[Sha77] I. R. Shafarevich. *Basic algebraic geometry*. Springer-Verlag, Berlin, 1977.

[Sti93] H. Stichtenoth. *Algebraic function fields and codes*. Universitext. Springer-Verlag, Berlin, 1993.

[Tat94] J. T. Tate. *Conjectures on algebraic cycles in l-adic cohomology*, volume 55 of *Proc. Sympos. Pure Math.* Amer. Math. Soc., Providence, RI, 1994.

[Ulm05] D. L. Ulmer. Geometric non-vanishing. *Invent. Math.*, 159:133–186, 2005.

[Wei48] A. Weil. *Sur les courbes algébriques et les variétés qui s'en déduisent*. Actualités Sci. Ind., no. 1041 = Publ. Inst. Math. Univ. Strasbourg **7** (1945). Hermann et Cie., Paris, 1948.

[Wei49] A. Weil. Numbers of solutions of equations in finite fields. *Bull. Amer. Math. Soc.*, 55:497– 508, 1949.

[Wei62] A. Weil. *Foundations of algebraic geometry*. American Mathematical Society, Providence, R.I., 1962. MR0144898 (26 #2439).

[Wei95] A. Weil. *Basic number theory*. Classics in Mathematics. Springer-Verlag, Berlin, 1995.

Department of Mathematics

University of Arizona
Tucson, AZ
85721 USA