Exponential sums over finite fields, II: introduction to cohomological methods

E. Kowalski

ETH Zürich – D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland

E-mail address: kowalski@math.ethz.ch

Contents

CHAPTER 1

Introduction

1. Why is the one-variable theory not sufficient?

In these notes, we give a motivated introduction to the methods first introduced by Grothendieck and his school for the study of exponential sums over finite fields. These were developped most crucially by Deligne (who proved the Riemann Hypothesis over finite fields in an extraordinarily general setting and established a powerful and flexible formalism to exploit it), and they were, and are still, applied extremely successfully by N. Katz in a number of deep works. Besides its own inner motivation and applications to other fields, this approach based on algebraic geometry and wide-ranging applications in analytic number theory. It has led to spectacular progress in extremely concrete arithmetic problems.

These notes are a follow-up to the first part ([**21**]) and will refer to the latter when giving references to the "elementary" theory contained there.

As a basic orientation, the first chapters will be concerned with defining a certain type of exponential sums, which we call "algebraic exponential sums", and which generalize quite extensively the one-variable character sums of the type

$$S = \sum_{x \in \mathbf{F}_q} \chi(g(x))\psi(f(x))$$

considered in [**21**], where χ is a multiplicative character of $\mathbf{F}_q$, ψ is an additive character and f, g are polynomials in $\mathbf{F}_q[X]$. Algebraic exponential sums will be of the form

$$\sum_{x \in V(\mathbf{F}_q)} \Lambda(x),$$

where both the possible summation sets and the summands will require a certain amount of preliminary work and background setup to be defined.

2. Outline of the rest of the book

In the next chapters we will provide background material concerning the following necessary tools for the development of the theory:

- Algebraic geometry;
- The definition and idea of the construction of the étale fundamental group;
- The p-adic fields and their basic properties;
- Representation theory (language and elementary results).

We do not assume prior expose to any of these; we hope however that the simple accounts we give will induce readers to read some more material. The discussion will also include some basic examples and references to exponential sums.

Once this is done, we can start the discussion of exponential sums with basic definitions of "algebraic" exponential sums in many variables, with summands which may be more general than character values. The heart of the text is the chapter concerning the statement and the formalism of the general form of the Riemann Hypothesis over finite fields which was proved by P. Deligne. The final chapters then explain some basic applications and illustrate its versatility. In particular, we will emphasize Deligne's Equidistribution Theorem.

CHAPTER 2

Background material: algebraic geometry

In this first background section, we present a very concrete introduction to the language of modern algebraic geometry. The goal is to make it possible to give full rigorous statements in this book which use the language of the original literature, without requiring the readers to be familiar with complete treatments of the foundational material (such as those in [**8**] or [**22**]). Of course, readers who *are* already knowledgeable about (even quite basic) scheme theory and arithmetic geometry may skip this chapter (and refer to its contents later only if needed, say when an example is referred-to later). The statements that we insert to give hints of the relation of algebraic geometry to our purpose of exponential sums, will be repeated later in more precise form.

1. Affine algebraic varieties

Except for isolated examples which will be independent of the main course of the book, we will restrict our attention, for simplicity, to so-called *affine* algebraic varieties. On the other hand, we want a clean and correct treatment of rationality questions over arbitrary fields and even rings, and therefore we can not (as is customary) restrict our attention to points with coordinates in algebraically closed fields.

We start with an informal definition based on "systems of equations". Let A be an arbitrary commutative ring, with unit $1 \in A$. An algebraic variety X over A is supposed to be related to "solutions of polynomial equations with coefficients in A". Consider therefore polynomials

$$f_1, f_2, \ldots, \tag{2.1}$$

in some polynomial ring

$$A[X_1, X_2, \ldots].$$

The set of $x = (x_i)$ in A with

$$f_1(x) = f_2(x) = \cdots = 0$$

is of course well-defined, but is often too "small" to be thought of as a real geometric object. In particular, very different equations, that we want to think as defining different geometric objects, may have the same set of solutions in A. For instance, think of $A = \mathbf{Q}$, and the two equations $x^n + y^n = z^n$ and $xyzt = 1$, in four variables. For all $n \geqslant 3$, there are no solutions (!), and yet clearly, the geometric objects are not the same, as one intuitively graphs by graphing the plot of (say) z as a function of *real* x and y. This indicates the way out: we can look at the solutions in "bigger" rings than A, where the equations still make sense, and hope to recover the geometric object from such sets of points.

Precisely, we can define a set of solutions $X(B)$ of the equations for any ring B given with a ring-homomorphism $f : A \to B$, or in other words, for any A-algebra. (Note that the homomorphism f, although it is often omitted, is part of the structure).

EXAMPLE 2.1. Let $A = \mathbf{Z}$, let $f(X, Y) = X^2 + Y^2 + 1 \in \mathbf{Z}[X, Y]$, and let X be the corresponding algebraic variety. We have therefore $X(\mathbf{Z}) = \emptyset$, and even $X(\mathbf{R}) = \emptyset$. However, $X(\mathbf{Q}(i))$ is not empty, as it contains $(\pm i, 0)$. Here, the rings $\mathbf{R}$ and $\mathbf{Q}(i)$ are given their (only) $\mathbf{Z}$-algebra structure, the inclusion $\mathbf{Z} \to \mathbf{Q}(i)$ or $\mathbf{Z} \to \mathbf{R}$. But we can also look at the $\mathbf{Z}$-algebras $\mathbf{Z}/p\mathbf{Z}$ for p prime, where $\mathbf{Z} \to \mathbf{Z}/p\mathbf{Z}$ is (of course) the reduction modulo p. Then, for instance, one checks that

$$X(\mathbf{Z}/5\mathbf{Z}) = \{(0, \pm 2), (\pm 2, 0)\} \subset (\mathbf{Z}/5\mathbf{Z})^2.$$

As it turns out, the collection of sets $(X(B))_B$, B running over all A-algebras, enscapsulates all geometric information one might need, provided one remembers the following obvious piece of extra data: for any A-algebras B and C given with a map

$$g : B \to C$$

of A-algebras, there is an associated way, denoted $g_* = g_{*,B,C}$, to map solutions of the equations with coefficients in B to those with coefficients in C, i.e., a (set) map

$$g_* : X(B) \to X(C).$$

These maps have some obvious properties:

– If $B = C$ and g is the identity, then g_* is also the identity on the set of solutions.

– If B, C, D are A-algebras with maps

$$B \xrightarrow{g} C \xrightarrow{h} D, \qquad B \xrightarrow{h \circ g} C,$$

mapping the solutions "from B to D" can be done by passing through C, or in other words the map

$$(h \circ g)_* : X(B) \to X(D)$$

is the composite

$$(h \circ g)_* = h_* \circ g_* : X(B) \xrightarrow{g_*} X(C) \xrightarrow{h_*} X(D). \tag{2.2}$$

Then one can say that the geometric object defined by the equations f_1, f_2, ..., is entirely characterized by the all collection of data

$$(X(B), g_* : X(B) \to X(C))_{B, g : B \to C}.$$

And, more importantly maybe, if A is a ring of "arithmetic" nature (e.g, $\mathbf{Z}$ or $\mathbf{Z}/p\mathbf{Z}$), the whole arithmetic of the object is contained in this data.

This language of solution sets is very convenient. Its major conceptual difficulty is that it is quite tricky to give conditions which ensure that an arbitrarily given such collection of data (i.e., sets $X(B)$ for every A-algebra B, and maps $g_* : X(B) \to X(C)$ for all $B \to C$, satisfying the two conditions above) amounts to giving sets of solutions of some (fixed) equations. For our purposes this is not very problematic because we will typically have equations at our disposal for applications. We give an example however to indicate that the maps g_* are certainly of importance.

EXAMPLE 2.2 (Not an algebraic variety). Consider the base ring $A = \mathbf{Z}$ again. A $\mathbf{Z}$-algebra B is just a ring B, since there is a unique map $\mathbf{Z} \to B$. Define

$$N(B) = \{b \in B \mid \text{there does not exist } c \in B \text{ with } c^2 = b\},$$

the set of non-squares in B. This might sound algebraic enough to be an algebraic variety. However that is not the case: there is no way to define the associated maps g_*! Indeed, assume those existed; consider then the composite maps

$$\mathbf{R} \xrightarrow{g} \mathbf{C} \xrightarrow{h} \mathbf{C}[T],$$

and the hypothetical induced maps: these would be

$$]-\infty, 0[= N(\mathbf{R}) \longrightarrow \emptyset = N(\mathbf{C}) \longrightarrow N(\mathbf{C}[T]) \neq \emptyset$$

(the last because, e.g., $T \in N(\mathbf{C}[T])$). But this is absurd, because there is no map from a non-empty set to $\emptyset$...

The related example

$$\square(B) = \{b \in B \mid \text{there does exist } c \in B \text{ with } c^2 = b\},$$

which is the complement of $N(B)$ in B, is more subtle, because there are obvious maps g_* in the case: given $g : B \to C$ is a ring-homomorphism, we can define $g_*(b) = g(b)$, and this is a map

$$g_* : \square(B) \to \square(C).$$

Moreover, the "functoriality" condition (2.2) is obviously valid! However, we will explain quickly that the "complement" of an algebraic variety is also one, and hence if there *existed* (an arbitrary system of) equations defining exactly the squares in a ring, there would also exist one defining the sets $N(B)$, which we have checked is impossible.

We have not yet given a proper definition. We will do so in a *third* way, bypassing equations and sets of solutions but easily related to both. The point is that we want to be able to define functions on our algebraic varieties (e.g., to provide arguments for exponential sums, or coordinates). Given equations (2.1), there are obvious functions on the set(s) of solutions, namely, the restriction to $X(B)$ of all the polynomials in $A[X_1, X_2, \ldots]$, mapped to $B[X_1, X_2, \ldots]$ using the given morphism $A \to B$. Two polynomial functions differing by any polynomial combination of the equations f_i obviously induce the same maps on all sets $X(B)$. In other words, there is an obvious ring of functions, given by the quotient ring

$$\mathcal{O}(X) = A[X_1, X_2, \ldots]/(f_1, f_2, \ldots).$$

Note that this ring is itself an A-algebra. The third definition is then based on the fact that ring of functions $\mathcal{O}(X)$, as an abstract A-algebra, characterizes completely the geometric (and arithmetic) object X. Hence, because it is very easy to define and play around with A-algebras, one can simply say that an affine algebraic variety over A "is" a A-algebra.

To check this claim of characterization, we need only observe the following lemma:

LEMMA 2.3. *Let $\mathcal{O}(X)$ be defined as above for given equations f_i. Then, for any A-algebra B, we have a bijection*

$$X(B) \simeq \mathrm{Hom}_A(\mathcal{O}(X), B)$$

given by mapping x to the A-algebra homomorphism

$$\varphi_x : \left\{ \begin{array}{ccc} \mathcal{O}(X) & \longrightarrow & B \\ f & \mapsto & f(x) \end{array} \right. ,$$

and for any A-algebra morphism $g : B \to C$, the map

$$g_* : X(B) \to X(C)$$

can be defined by

$$g_*(\varphi) = g \circ \varphi : \mathcal{O}(X) \to C, \quad \text{for all } \varphi \in \mathrm{Hom}_A(\mathcal{O}(X), B) \simeq X(B).$$

Intuitively, the map $\varphi_x : \mathcal{O}(X) \to B$ associated to a point $x \in X(B)$ is the map of *evaluation of a function defined on X at the point x*. One recovers the point x from the evaluation map by noting that its coordinates are, simply, the results of evaluating at x (using φ_x) the "coordinate functions"... We now formalize this.

PROOF. Consider

$$\mathrm{Hom}_A(\mathcal{O}(X), B) = \mathrm{Hom}_A(A[X_1, X_2, \ldots]/(f_1, f_2, \ldots), B).$$

By the very definition of quotient rings, to give an element φ in this set is the same as giving the elements b_i in B to which φ maps, and these can be chosen arbitrarily, provided all the relations that the X_i satisfy are also satisfied by the b_i. But this is the same as saying that $f_j(b_1, b_2, \ldots) = 0$ for all j, i.e., that $b = (b_1, b_2, \ldots)$ is in $X(B)$. This shows that we have our bijection. The description of g_* is then also clear: since $g_*(b) = (g(b))$, the corresponding mapping $\psi \in \mathrm{Hom}_A(\mathcal{O}(X), C)$ satisfies

$$\psi(X_i) = g(b_i) = g(\varphi(X_i)),$$

for all i, and hence $\psi = g \circ \varphi$. □

REMARK 2.4. Of course, the ring $\mathcal{O}(X)$ does not allow us to recover the exact *equations* – whatever they where – that were used to define X. But that is as it should be, since specific equations are insubstantial things, and any number of changes of variables, substitutions, etc, can change them, whereas the intrinsic geometric nature of the object is not, and should not be, altered.

We can summarize as follows:

DEFINITION 2.5 (Affine algebraic variety, regular functions). Let A be a ring, commutative with unit. An *affine algebraic variety X defined over A* (the base ring is often incorporated in the notation by writing X/A) is the equivalent data of any of the following three descriptions (where only the third is entirely unambiguous):

– The data of a number of polynomial equations with coefficients in A;

– The data of the sets of points $X(B)$, for B any A-algebra, together with all induced maps $X(B) \xrightarrow{g_*} X(C)$ for any A-algebra morphism $B \xrightarrow{g} C$.

– The data of the ring $\mathcal{O}(X)$, which is an arbitrary A-algebra.

The A-algebra $\mathcal{O}(X)$ is called the *ring of (regular) functions* on X.

In essence – and in greater generality, rigor, and context – this definition is due to Grothendieck. The next section will give some indications of how remarkably flexible it is, compared with more naïve approaches. One may already note that there is no condition on the base ring, and no condition on the ring $\mathcal{O}(X)$.

REMARK 2.6. To adhere with standard notation, although we do not attempt to motivate the terminology, we will write $X = \mathrm{Spec}(R)$ for the algebraic variety which has $\mathcal{O}(X) = R$, which is called the "spectrum" of R. By definition, its points "with coordinates in B" are given by

$$\mathrm{Spec}(R)(B) = \mathrm{Hom}_A(R, B).$$

2. First examples

3. Computing with algebraic varieties

CHAPTER 3

Summands for algebraic exponential sums

The target of this chapter is the definition of a suitably large collection of summands for algebraic exponential sums. The final definition is in Section 5. Readers who wish to go straight to the heart of the Riemann Hypothesis in the next chapter (and who are familiar already with the fields of ℓ-adic numbers) can skip to that section with little loss of continuity, going back to Section 4 as (and if) needed later on. The goal of the first sections, however, is to attempt to motivate the definition and to link it with other basic ideas of number theory.

1. From Dirichlet characters to Galois characters

In the first part of this book, we succeeded in expressing exponential sums in one variable, either additive or multiplicative, as sums related to Dirichlet (or generalized) characters of the field $\mathbf{F}_q(T)$ of rational functions (see Sections 4.1 and 5.3 in [**21**]).

However, this approach does not readily extend to more than one variable (or to more complicated summands). The cohomological methods are based, instead, on an alternate representation, which involves *Galois* characters of some kind, instead of Dirichlet characters. There is nothing outlandish in this shift; in classical algebraic number theory, it has been a fundamental theme, which is intimately related to the topic of *reciprocity laws*. Indeed, one version of the classical Kronecker-Weber theorem takes the following form: there is a bijection, preserving L-functions, between primitive Dirichlet characters (of $\mathbf{Z}$) and Galois characters of $\mathbf{Q}$, i.e., homomorphisms

$$\rho : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathbf{C}^\times,$$

with finite image, where the L-function of such a character is defined as the analytic continuation (to the maximal extent possible) of the Euler product

$$L(\rho, s) = \prod_{p \text{ unramified}} (1 - \rho(\mathrm{Fr}_p)p^{-s})^{-1},$$

which is absolutely convergent for $\mathrm{Re}(s) > 1$ (because ρ has finite image, hence image contained in the roots of unity of some order). The factors $\rho(\mathrm{Fr}_p)$, and the set of p being used, are defined as follows: first of all, since ρ has finite image, its kernel has finite index, and if k is the fixed field of the latter, ρ can be considered as a homomorphism

$$\mathrm{Gal}(k/\mathbf{Q}) \longrightarrow \mathbf{C}^\times,$$

where the left-hand Galois group is abelian. Let $\mathbf{Z}_k$ be the ring of integers in k. As usual, for every prime p which is *unramified* in $\mathbf{Z}_k$, after fixing a

prime ideal $\mathfrak{p}$ dividing $p\mathbf{Z}_k$, and defining

$$q = |\mathbf{Z}_k/\mathfrak{p}|,$$

there exists a unique element $\mathrm{Fr}_p \in \mathrm{Gal}(k/\mathbf{Q})$ such that

$$\sigma(x) \equiv x^q \,(\mathrm{mod}\,\mathfrak{p})$$

for all $x \in \mathbf{Z}_K$. (This is unique as element of the group, and depends only on p, because the Galois group is abelian here.) Thus $\rho(\mathrm{Fr}_p)$ is well-defined, and so is the L-function $L(\rho, s)$.

In this case, and much of the time for number fields, the statement that each such Artin L-function $L(\rho, s)$ is also an L-function of a Dirichlet character $\chi(\rho)$, is considered mostly as information concerning ρ. For instance, this is the only known way to prove that these L-functions have analytic continuation – using the corresponding fact for Dirichlet L-functions, which is fairly elementary (using the Poisson formula).

Over finite fields, *the situation is completely changed!* It turns out – this will be taken up in the next chapter – that it is possible to prove the analytic continuation (in fact, rationality, and many other properties) of all L-functions of Galois origin,[1] over (almost) arbitrary algebraic varieties. Thus, they become extremely powerful tools. And especially, they become useful for the study of exponential sums, because those can be – quite easily and generally – represented as sums related to Galois-like objects.

The "right" notion will be introduced in Section 3.[2] Here we will first check directly, for additive character sums (of arbitrarily many variables), that there is such a representation in terms of a Galois character.

2. From Galois groups to fundamental groups

As a second preliminary step towards the general case, we will explain here the definition of the étale algebraic fundamental group of a variety, which will be used (instead of the Galois group of the function field, which is usually too big and unwieldy to be of use) in the next section.

3. Lisse ℓ-adic sheaves

Example 3.1 (Tate sheaves). We now give a very basic example for varieties over a finite field $\mathbf{F}_q$, the *Tate sheaves*. Although they are extremely simple, they play a fundamental supporting role in the theory.

4. Formalism of ℓ-adic sheaves

A fundamental feature of ℓ-adic sheaves is that *they are objects of linear algebra*, in some sense. Hence, they are subject to all the classical operations of linear algebra, and this gives a rich formalism that can be used to great effect. We summarize this quickly.

Example 3.2 (Homomorphisms). We first discuss the analogue of linear maps between vector spaces. Given two ℓ-adic sheaves ρ_1, ρ_2 (on the

[1] Whereas, over $\mathbf{Q}$, most Artin l-functions remain completely shrouded in mystery.

[2] Though for this book we will not in fact treat the most general possible cases.

same variety V, and with the same ℓ), acting on E_1 and E_2, respectively, a *homomorphism*

$$\Phi \, : \, \rho_1 \longrightarrow \rho_2$$

is a $\overline{\mathbf{Q}}_\ell$-linear map

$$\Phi \, : \, E_1 \to E_2,$$

such that the commutation rule

$$\Phi \circ \rho_1(\sigma) = \rho_2(\sigma) \circ \Phi,$$

holds for all $\sigma \in \pi_1(V)$. In other words, writing linear actions on E_1 and E_2 with a simple dot, we have

$$\Phi(\sigma \cdot v) = \sigma \cdot \Phi(v)$$

for all $v \in E_1$.

Of particular importance are injective and surjective morphisms, and isomorphisms: those are defined by asking that Φ, as a linear map, has the corresponding property.

Examples of injective morphisms are of course inclusions of subsheaves (i.e., subrepresentations of ρ): if $E_2 \subset E_1$ is a $\overline{\mathbf{Q}}_\ell$-linear subspace stable under the action of $\pi_1(V)$, then the restriction of $\rho_1(\sigma)$ ot E_2 defines a lisse sheaf ρ_2 acting on E_2, and the inclusion defines an injective morphism $\rho_2 \to \rho_1$. Also, there is then an induced action on the quotient space $E_3 = E_1/E_2$, and this gives a lisse ℓ-adic sheaf ρ_3 with a surjective morphism $\rho_1 \to \rho_3$.

In particular, if $\rho_1 \xrightarrow{\Phi} \rho_2$ is given, one checks immediately that $\operatorname{Ker}\Phi \subset E_1$ is stable under $\pi_1(V)$, and hence gives a subsheaf, and that $\operatorname{Im}(\Phi) \subset E_2$ is also a subsheaf of ρ_2. Moreover, the classical isomorphism

$$\operatorname{Im}(\Phi) \simeq E_1/\operatorname{Ker}(\Phi)$$

is an isomorphism of $\pi_1(V)$-representations. We see that one can speak of morphisms, kernels, cokernels, exact sequences, etc, of ℓ-adic sheaves on V (for a given ℓ). It is in fact an abelian category.

As a further example, given a lisse sheaf ρ acting on E, one can define the invariant subsheaf $\rho^{\pi_1(V)}$ by the (trivial) action of $\pi_1(V)$ on the subspace

$$E^{\pi_1(V)} = \{v \in E \mid \rho(\sigma)v = v \text{ for all } \sigma \in \pi_1(V)\} \subset E.$$

One can also define the *coinvariant* space

$$E_{\pi_1(V)} = E/E_1 \tag{3.1}$$

where E_1 is the space spanned by vectors of the form

$$\rho(\sigma)v - v \tag{3.2}$$

for $\sigma \in \pi_1(V)$, $v \in E$ (it is easy to check that this is a subsheaf of ρ). From the definition of the induced action on E/E_1, note that $E_{\pi_1(V)}$ carries also a trivial action of $\pi_1(V)$. Intuitively, the invariant space is the largest subspace of ρ on which $\pi_1(V)$ acts trivially, while the coinvariant space is the largest *quotient* on which the group acts trivially.

Example 3.3 (Direct sums). For instance, one can define easily direct sums of sheaves: given $\rho_1, \ldots, \rho_k$ which are all ℓ-adic sheaves on V/A, with the same ℓ as usual, one can form

$$\rho = \rho_1 \oplus \cdots \oplus \rho_k$$

which is defined by the obvious action on the direct sum

$$E = E_1 \oplus \cdots \oplus E_k$$

of the spaces E_i on which ρ_i acts:

$$\rho(\sigma)(v_1 + \cdots + v_k) = \sum_{i=1}^{k} \rho_i(\sigma) v_i$$

for all $v_i \in E_i$. Note, for each i, the obvious morphisms of ℓ-adic sheaves

$$\rho \to \rho_i, \qquad \rho_i \to \rho,$$

where the first is surjective and the second injective.

Example 3.4 (Tensors, dual, hom-spaces). An additional structure, also usual for vector spaces, is the tensor product. Given two ℓ-adic sheaves ρ_1 and ρ_2, acting on E_1 and E_2 respectively, one can form

$$\rho_3 = \rho_1 \otimes \rho_2,$$

acting on the tensor product $E_1 \otimes E_2$ by

$$\rho_3(\sigma) = \rho_1(\sigma) \otimes \rho_2(\sigma) \, : \, E_1 \otimes E_2 \to E_1 \otimes E_2.$$

This can be repeated with multiple factors, and the subspaces of tensor powers giving the symmetric powers and alternating powers also exist and are (sub)sheaves on V. The usual decompositions, such as

$$\rho \otimes \rho \simeq \mathrm{Sym}^2(\rho) \oplus \wedge^2(\rho),$$

where Sym^2 is the symmetric square and $\wedge^2$ is the alternating square, hold.

There is also a notion of dual, or contragredient, sheaf:

Definition 3.5. Let V/A be an algebraic variety, let ℓ be a prime number invertible in $\mathcal{O}(V)$ and let ρ be a lisse ℓ-adic sheaf on V acting on E. The *dual* of ρ, denoted $\check{\rho}$, is the $\pi_1(V)$-action on the space $\check{E} = \mathrm{Hom}(E, \overline{\mathbf{Q}}_\ell)$ of linear forms on E defined by

$$\langle \check{\rho})(\sigma)\lambda, v\rangle = \langle \lambda, \rho(\sigma)v\rangle$$

for $\sigma \in \pi_1(V)$, $\lambda \in \check{E}$, $v \in E$, in terms of the duality bracket.

The classical isomorphism of vector spaces

$$\mathrm{Hom}(E_1, E_2) \simeq E_2 \otimes \check{E}_1$$

(where the pure tensors $w \otimes \lambda$ correspond to the rank-1 linear maps $\Psi(v) = \lambda(v)w$) shows that one can also give the space of linear maps the structure of an ℓ-adic sheaf. It is easy to check that the corresponding action is described by

$$(\sigma \cdot \Psi)(v) = \sigma \cdot \Psi(\sigma^{-1} \cdot v)$$

for $\Psi \in \mathrm{Hom}(E_1, E_2)$ a linear map and $v \in E_1$. In particular, note that the invariant sheaf of $\mathrm{Hom}(E_1, E_2)$ is given by

$$\{\Psi \, : \, E_1 \to E_2 \mid \sigma \cdot \Psi(\sigma^{-1} \cdot v) = v \text{ for all } \sigma \in \pi_1(V)\},$$

which is the group of sheaf-homomorphisms (not only $\overline{\mathbf{Q}}_\ell$-linear) between E_1 and E_2.

Another useful fact is the isomorphism

$$(\check{\rho})^{\pi_1(V)} \simeq (\rho_{\pi_1(V)})^{\vee}, \tag{3.3}$$

(or in other words, the dual of the invariant space is the coinvariant of the dual.) Indeed, by definition of the coinvariant quotient space (3.1) as quotient of E, its dual is the subspace of $\check{E}$ of linear forms λ such that all vectors (3.2) are in the kernel of λ. But since

$$\langle \lambda, \sigma \cdot v - v\rangle = \langle \lambda, \sigma \cdot v\rangle - \langle \lambda, v\rangle = \langle \sigma^{-1} \cdot \lambda, v\rangle - \langle \lambda, v\rangle,$$

we see that this is equivalent with λ being invariant in the dual sheaf.

Example 3.6 (Irreducibility, semisimplicity). Proceeding as in the representation theory of finite groups, one defines irreducible and semisimple sheaves:

Definition 3.7. Let V/A be an algebraic variety, let ℓ be a prime number invertible in $\mathcal{O}(V)$.

(1) A lisse ℓ-adic sheaf ρ on V is *irreducible* if the only subsheaves of ρ are the zero space $0 \subset \rho$ and ρ itself.

(2) A lisse ℓ-adic sheaf ρ on V is *semisimple* if there exist irreducible sheaves $\rho_1, \ldots, \rho_k$ on V such that

$$\rho \simeq \rho_1 \oplus \cdots \oplus \rho_k,$$

i.e., ρ is a direct sum of irreducible sheaves.

For instance, any sheaf of rank 1 is necessarily irreducible. The most important property of irreducible sheaves, in general, is the famous Schur lemma:

Lemma 3.8 (Schur's lemma). *Let ρ be an irreducible lisse ℓ-adic sheaf on V/A, and τ another lisse ℓ-adic sheaf. A homomorphism $\Phi : \rho \to \tau$ is either zero or injective. In the second case, one says that ρ* occurs in *τ.*

Proof. This is the same proof as the standard case: the kernel $\operatorname{Ker}\Phi$ is a subsheaf of ρ, hence – by definition of irreducibility – either it is 0, in which case Φ is injective, or it is ρ itself, in which case $\Phi = 0$. □

However, one feature one is used to from the case of representations of finite groups does *not* extend to general ℓ-adic sheaves: not all of them are semisimple. Concretely, this means there exist examples of sheaves ρ with a subsheaf ρ_1 which is neither 0 nor ρ itself, for which there does not exist ρ_2 such that

$$\rho \simeq \rho_1 \oplus \rho_2.$$

We will give an example in the next section.

We have now given examples essentially similar to linear algebra. In these, the base variety, and the prime ℓ, were fixed (indeed, only $\pi_1(V)$ played a role, and any other group would have done as well!). We now discuss what can be said about changing V, or changing ℓ.

Example 3.9 (Varying the base variety). We consider here only the simplest cases of relations between ℓ-adic sheaves on two varieties, when we are given W and V over the base ring A, ℓ invertible in A (hence both in

$\mathcal{O}(W)$ and $\mathcal{O}(V)$), and a morphism $W \to V$ of algebraic varieties, namely we assume that $W \to V$ is a finite étale covering. In that case, we know that $\pi_1(W)$ can be identified with a finite-index subgroup of $\pi_1(V)$, the index being the degree $[W : V]$ of the covering.

It follows immediately that we have a restriction operation Res_V^W that associates, to any ℓ-adic sheaf ρ on V, a sheaf $\mathrm{Res}_V^W(\rho)$ on W, which is simply the restriction to this finite index subgroup. Note that the rank of the restriction is the same as that of the original sheaf.

In the opposite direction, there is a well-known construction in representation theory, namely the *induction* Ind_W^V: starting from a lisse ℓ-adic sheaf ρ on W, say

$$\rho : \pi_1(W) \to \mathrm{GL}(E),$$

with $\dim(E) = r$, we construct as follows a sheaf $\tilde{\rho}$ of rank $r[W : V]$ on V: let

$$F = \{\varphi : \pi_1(V) \to E \mid \varphi \text{ is continuous and } \varphi(\sigma_1\sigma_2) = \rho(\sigma_1)(\varphi(\sigma_2)) \text{ for } \sigma_1 \in \pi_1(W),\ \sigma_2 \in \pi_1(V)\} \tag{3.4}$$

(where the elements φ are just arbitrary continuous functions) and define

$$\tilde{\rho}(\sigma)\varphi(x) = \varphi(x\sigma)$$

for all $\varphi \in F$ (the so-called regular representation of $\pi_1(V)$).

It is essentially immediate that $\tilde{\rho}$ does indeed defined a representation on F. We need to check that $\dim F = r[W : V]$ and that $\tilde{\rho}$ is continuous. Both are quite easy, the point being that the transformation property of functions in F implies that if C is any fixed set of coset representatives of $\pi_1(V)/\pi_1(V)$, the restriction map

$$\begin{cases} F & \longrightarrow & E^{[W:V]} \\ \varphi & \mapsto & (\varphi(x))_{x\in C} \end{cases}$$

is first obviously a $\overline{\mathbf{Q}}_\ell$-linear injection (because $\varphi(\sigma)$ is determined, for every σ, by the element $x \in C$ to which it is equivalent under $\pi_1(W)$), and in fact bijective because one can fix any $(\alpha_x)_{x\in C}$ in E^C and define a function unambiguously by

$$\varphi(\sigma x) = \rho(\sigma)(\alpha_x)$$

for $\sigma \in \pi_1(W)$, $x \in C$. This function is easily checked to be in F: we have

$$\varphi(\sigma_1\sigma_2) = \varphi(\sigma_1\sigma' x) = \rho(\sigma_1\sigma')(\alpha_x) = \rho(\sigma_1)\varphi(\sigma_2)$$

for $\sigma_1 \in \pi_1(W)$, $\sigma_2 = \sigma' x \in \pi_1(V)$.

Example 3.10 (Variation of ℓ). This is the least understood phenomenon. Although, a priori, the definition of a lisse ℓ-adic sheaf certainly depends on ℓ, it turns out that to a large extent, in applications, one ends up with results which are independent of ℓ, in some sense.

Finally, in the case of most interest for us, when the base ring A is a finite field $\mathbf{F}_q$ of characteristic p, a final piece of formalism relates the "arithmetic" theory with its "geometric" counterpart, when V is replaced by $\bar{V}$ over an algebraic closure of $\mathbf{F}_q$.

EXAMPLE 3.11 (Arithmetic versus geometric). Recall that when $V/\mathbf{F}_q$ is an algebraic variety, we have a base change

$$\bar{V} = V \times \overline{\mathbf{F}}_q \longrightarrow V$$

and an associated short exact sequence of fundamental groups

$$1 \to \pi_1(\bar{V}) \to \pi_1(V) \to \mathrm{Gal}(\overline{\mathbf{F}}_q/\mathbf{F}_q) \to 1.$$

In particular (this could be considered as an instance of Example 3.9), for any lisse ℓ-adic sheaf ρ on $V/\mathbf{F}_q$, we obtain one by restriction on $\bar{V}$:

$$\bar{\rho} = \rho\,|_{\pi_1(\bar{V})}\ .$$

For any property $\mathcal{P}$ of sheaves, it is customary to say that ρ has $\mathcal{P}$ geometrically if $\bar{\rho}$ satisfies the property. To emphasize this, one also sometimes says that ρ has $\mathcal{P}$ arithmetically when it holds for ρ itself.

For instance, one may speak of a *geometrically irreducible* sheaf, or of a *geometrically semisimple* sheaf, or one might say that ρ_1 and ρ_2 are *geometrically isomorphic*. The Tate sheaves $\overline{\mathbf{Q}}_\ell(n)$ on $V/\mathbf{F}_q$ (Example 3.1), to give a concrete example, are all geometrically isomorphic, and geometrically trivial.

5. Algebraic exponential sums and examples

To summarize the discussion in this chapter and the previous one, we have defined a certain type of summation sets (rational points of algebraic varieties) and a certain type of summands (trace functions of lisse ℓ-adic sheaves, ℓ prime to the characteristic) over such sets. This allows us to give a complete definition of what is meant by an *algebraic exponential sum* in this book.

DEFINITION 3.12 (Algebraic exponential sum). Let $V/\mathbf{F}_q$ be an algebraic variety over a finite field $\mathbf{F}_q$ of characteristic p, with $\mathcal{O}(V)$ an integral domain. Let $\ell \neq p$ be a prime number and

$$\rho\ :\ \pi_1(V,\eta) \longrightarrow \mathrm{GL}(E)$$

be a lisse ℓ-adic sheaf of rank $r \geqslant 0$ on V, where E is a $\overline{\mathbf{Q}}_\ell$-vector space.

(1) The *algebraic exponential sum* associated to (V,ρ) is the sequence of sums $(\mathcal{S}_\nu(V;\rho))_{\nu\geqslant 1}$ given by

$$\mathcal{S}_\nu(V;\rho) = \sum_{x\in V(\mathbf{F}_{q^\nu})} \mathrm{Tr}(\rho(\mathrm{Fr}_{x,q^\nu})),$$

which are elements of $\overline{\mathbf{Q}}_\ell$. We will often denote

$$\Lambda_\rho(x) = \mathrm{Tr}(\rho(\mathrm{Fr}_{x,q^\nu}))$$

for $x \in V(\mathbf{F}_{q^\nu})$.

(2) The L-function of the algebraic exponential sum associated to (V,ρ) is the formal power series with coefficients in $\overline{\mathbf{Q}}_\ell$ defined by

$$L(V;\rho) = \exp\Big(\sum_{\nu\geqslant 1} \frac{\mathcal{S}_\nu(V;\rho)}{\nu} T^\nu\Big) \in \overline{\mathbf{Q}}_\ell[[T]].$$

Note that, a priori, these sums are here defined to take values in the ℓ-adic field $\overline{\mathbf{Q}}_\ell$, and not in the field of complex numbers. However, in all applications of these sums to analytic number theory (to the author's knowledge), this apparent behavior is to a large extent illusory. To be precise, in these applications, the starting point is some finite sum

$$S = \sum_{x \in V(\mathbf{F}_p)} \Lambda(x)$$

of complex numbers $\Lambda(x)$ over points of an algebraic variety over a finite field, typically $\mathbf{Z}/p\mathbf{Z}$ (these numbers are algebraic, but not necessarily roots of unity, as the example of average behavior of Kloosterman sums shows), which one wants to understand. Then, by some means or other (often by quoting a general theorem to that effect), it is known that for any $\ell \neq p$ (or possibly just for some of them), there exists a lisse ℓ-adic sheaf ρ over the underlying variety V such that

$$\mathcal{S}_1(V;\rho) \text{=} S,$$

where the[3] "fat equal" sign means that for *some* field map $\overline{\mathbf{Q}}_\ell \xrightarrow{\iota} \mathbf{C}$, we have

$$S = \iota(\mathcal{S}_1(V;\rho)).$$

Example 3.13 (Point counting). The most elementary example (and the only one to which the original conjectures of Weil explicitly referred!) is to take $\rho = \overline{\mathbf{Q}}_\ell$; we then have, obviously, the formula

$$\mathcal{S}_\nu(V;\overline{\mathbf{Q}}_\ell) = |V(\mathbf{F}_{q^\nu})|,$$

where the right-hand side, as an integer, is well and unambiguously defined independently of it being in $\overline{\mathbf{Q}}_\ell$ or any other ring containing $\mathbf{Z}$.

The next examples explain how this works for the basic examples of additive and multiplicative character sums.

Example 3.14 (Additive character sums). Let $V/\mathbf{F}_q$ be as in the definition, and let $f \in \mathcal{O}(V)$ be a non-constant function on V. In Section 2, we saw how to express the character sums

$$\sum_{x \in V(\mathbf{F}_{q^\nu})} \psi(\mathrm{Tr}(f(x)))$$

as sums of local traces

$$\sum_{x \in V(\mathbf{F}_{q^\nu})} \mathrm{Tr}(\mathcal{L}_{\psi(f)}(\mathrm{Fr}_{x,q^\nu}))$$

for the corresponding complex-valued character

$$\mathcal{L}_{\psi(f)} : \pi_1(V) \longrightarrow \mathbf{C}^\times$$

(we insert the trace here, although it is superfluous, in order to facilitate comparison).

We can bring this to the form in Definition 3.12 quite easily, because we now that this character has finite image, in fact it takes values in the subgroup $\boldsymbol{\mu}_p$ of p-th roots of unity in $\mathbf{C}^\times$. This group "exists" in $\overline{\mathbf{Q}}_\ell$ for any

[3] Non-standard.

prime $\ell \neq p$ (indeed, in a finite extension of $\mathbf{Q}_\ell$, and even in $\mathbf{Q}_\ell$ itself if ℓ is chosen so that $\ell \equiv 1 \pmod{p}$), and is isomorphic to $\boldsymbol{\mu}_p$ (since they are both cyclic of order p), although this isomorphism is not canonical if $p \neq 2$.

Thus let $\boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell)$ be the group of p-th roots of unity in $\overline{\mathbf{Q}}_\ell$. There exists an isomorphism

$$\boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell) \xrightarrow{\iota} \boldsymbol{\mu}_p$$

and an additive character

$$\psi_\ell : \mathbf{F}_q \to \boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell) \subset \overline{\mathbf{Q}}_\ell^\times$$

of $\mathbf{F}_q$ with values in $\boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell)$ such that

$$\psi(\operatorname{Tr} f(x)) = \iota\big(\psi_\ell(\operatorname{Tr} f(x))\big)$$

for all $x \in V(\mathbf{F}_{q^\nu})$. Composing with the map

$$\pi_1(V) \to \operatorname{Gal}(V_f/V) \to \mathbf{F}_q$$

already considered in Section 2, we obtain a lisse ℓ-adic sheaf $\mathcal{L}_{\psi_\ell(f)}$ of rank 1 on V (it is continuous, as before, because the image is finite). The group isomorphism ι extends uniquely to an isomorphism of the corresponding cyclotomic fields

$$\mathbf{Q}_\ell(\boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell)) \simeq \mathbf{Q}(\boldsymbol{\mu}_p) \subset \mathbf{C},$$

and we find that

$$\sum_{x \in V(\mathbf{F}_{q^\nu})} \psi(\operatorname{Tr}(f(x))) = \iota\big(\mathcal{S}_\nu(V; \mathcal{L}_{\psi_\ell(f)})\big)$$

for every $\nu \geqslant 1$.

We will now use the formalism of ℓ-adic sheaves to "reconstruct" these $\mathcal{L}_{\psi_\ell(f)}$ from scratch; this will provide an example of the induction operation from Example 3.9. Fix a prime $\ell \neq p$ and a character

$$\psi : \mathbf{F}_q \to \overline{\mathbf{Q}}_\ell^\times,$$

which of course takes values in $\boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell)$ (we have changed notation a bit).

Now consider, as in Section 2, the Artin-Schreier covering

$$W \to V$$

where W is given by the equation

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

over V. We have the homomorphism

$$\mathcal{L}_f : \pi_1(W) \to \operatorname{Gal}(W/V) \to \mathbf{F}_q$$

given by sending σ to $\sigma(y_0) - y_0$ for any fixed solution y_0 of the equation. Consider then the trivial ℓ-adic sheaf $\overline{\mathbf{Q}}_\ell$ on W, and form the induced sheaf

$$\rho = \operatorname{Ind}_W^V(\overline{\mathbf{Q}}_\ell)$$

on V, which is of rank $q = [W : V]$ according to Example 3.9. We can then look, in the description (3.4), at the $(\psi \circ \mathcal{L}_f)$-isotypic component:

$$\rho_\psi = \{\varphi : \pi_1(V) \to \overline{\mathbf{Q}}_\ell \mid \varphi(\sigma_1\sigma_2) = \psi(\mathcal{L}_f(\sigma_1))\varphi(\sigma_2) \text{ for all } \sigma_1 \in \pi_1(W)\}.$$

It is again easy to see that ρ_ψ is a subsheaf of ρ, and indeed that

$$\mathrm{Ind}_W^V(\overline{\mathbf{Q}}_\ell) = \rho = \bigoplus_\psi \rho_\psi \tag{3.5}$$

where each ρ_ψ is of rank 1 (and hence irreducible). Given a field isomorphism $\mathbf{Q}_\ell(\boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell)) \xrightarrow{\iota} \mathbf{Q}(\boldsymbol{\mu}_p)$, we have

$$\iota(\mathcal{S}_\nu(V;\rho_\psi)) = \sum_{x\in V(\mathbf{F}_q^\nu)} \iota\circ\mathrm{Tr}(\psi(f(x)))$$

(where again the trace could be dispensed with), and the $\iota\circ\psi$ are all the complex-valued additive characters of $\mathbf{F}_q$. In other words, the complex-valued characters $\mathcal{L}_{\psi(f)}$ are the same as the $\iota\circ\rho_\psi$.

Example 3.15 (Multiplicative character sums). The same argument applies of course to multiplicative character sums. Thus for $V/\mathbf{F}_q$ as in the definition, $f\in\mathcal{O}(V)^\times$ an invertible function on V, χ a multiplicative character of order d of $\mathbf{F}_q^\times$, we find that there exists, for each prime $\ell\neq p$, a lisse ℓ-adic sheaf of rank 1 on V, denoted $\mathcal{L}_{\chi\ell(g)}$, and a field isomorphism

$$\iota\,:\,\mathbf{Q}_\ell(\boldsymbol{\mu}_d(\overline{\mathbf{Q}}_\ell))\simeq\mathbf{Q}(\boldsymbol{\mu}_d)\subset\mathbf{C}$$

such that

$$\sum_{x\in V(\mathbf{F}_{q^\nu})}\chi(Ng(x)) = \iota\big(\mathcal{S}_\nu(V;\mathcal{L}_{\chi\ell(g)})\big)$$

for every $\nu\geqslant 1$. We also leave as an exercise for the reader to find a description of the corresponding sheaves as subsheaves of one constructed by induction from the trivial sheaf on the covering with equation $y^d=g(x)$.

Example 3.16 (Operations on algebraic exponential sums). Because of the extended formalism of ℓ-adic sheaves, and the elementary properties of the trace operation, there are many operations on exponential sums which reflect corresponding operations at the level of the associated sheaves. For instance:

– If $\rho=\rho_1\oplus\rho_2$, we have

$$\mathcal{S}_\nu(V;\rho_1\oplus\rho_2)=\mathcal{S}_\nu(V;\rho_1)+\mathcal{S}_\nu(V;\rho_2)$$

for all $\nu\geqslant 1$. In terms of L-functions, this amounts to

$$L(V;\rho_1\oplus\rho_2)=L(V;\rho_1)L(V;\rho_2)$$

(product of formal power series).

Note that a moment's thought shows that one does not need a *direct sum* decomposition: if instead we have a subsheaf ρ_1 of ρ, and denote by ρ_2 the quotient sheaf, so that there is a short exact sequence

$$0\to\rho_1\to\rho\to\rho_2\to 0,$$

the trace under ρ of any $\sigma\in\pi_1(V)$ remains the sum

$$\mathrm{Tr}(\rho_1(\sigma))+\mathrm{Tr}(\rho_2(\sigma)),$$

and therefore we still have

$$\mathcal{S}_\nu(V;\rho)=\mathcal{S}_\nu(V;\rho_1)+\mathcal{S}_\nu(V;\rho_2),\qquad L(V;\rho)=L(V;\rho_1)L(V;\rho_2).$$

In other words, the algebraic exponential sums, as invariants of the ℓ-adic sheaves, do not "see" the difference between a semisimple sheaf and one built by non-trivial "extensions".

– If $\rho = \rho_1 \otimes \rho_2$, we have

$$\mathcal{S}_\nu(V; \rho_1 \otimes \rho_2) = \sum_{x \in V(\mathbf{F}_{q^\nu})} \mathrm{Tr}(\rho_1(\mathrm{Fr}_{x,q^\nu}))\,\mathrm{Tr}(\rho_2(\mathrm{Fr}_{x,q^\nu})),$$

for all $\nu \geqslant 1$. Thus the summands are multiplied. There is no simple relation between the L-functions in that case.

– Some other relations are somewhat deeper; for instance, (3.5) corresponds to the formula that expresses the point counting on an Artin-Schreier covering as function of the additive characters of $\mathbf{F}_q$: given $V/\mathbf{F}_q$, $f \in \mathcal{O}(V)$ and W the covering

$$y^q - y = f(x),$$

we have

$$|W(\mathbf{F}_{q^\nu})| = \mathcal{S}_\nu(W, \overline{\mathbf{Q}}_\ell) = \sum_\psi \mathcal{S}_\nu(V; \rho_\psi),$$

where the sum extends over all characters $\psi : \mathbf{F}_q \to \boldsymbol{\mu}_p(\overline{\mathbf{Q}}_\ell)$. This is exactly the same as the formula in Lemma 5.1 in [**21**]! And note that although we can obtain it from the description (3.5) of the induced representation, its elementary direct proof can be taken as motivation for guessing that the latter is correct.

– To give another example, one may wonder about creating an algebraic exponential sum corresponding to the product of sums

$$\mathcal{S}_\nu(V; \rho_1)\mathcal{S}_\nu(V; \rho_2)$$

(instead of multiplying the summands, as was done using $\rho_1 \otimes \rho_2$).

If we expand the product, we find

$$\mathcal{S}_\nu(V; \rho_1)\mathcal{S}_\nu(V; \rho_2) = \sum_{x,y \in V(\mathbf{F}_{q^\nu})} \sum \Lambda_{\rho_1}(x)\Lambda_{\rho_2}(y),$$

which looks reasonably like a sum over the $\mathbf{F}_{q^\nu}$-rational points of the product $V \times V$.

Thus the natural question is: does there exist, on the variety $W = V \times V$, a lisse ℓ-adic sheaf $\rho_1 \boxtimes \rho_2$, such that

$$\Lambda_{\rho_1 \boxtimes \rho_2}(x, y) = \Lambda_{\rho_1}(x)\Lambda_{\rho_2}(y)$$

for all $(x, y) \in W(\mathbf{F}_{q^\nu})$, $\nu \geqslant 1$? The answer is, unsurprisingly, yes. Indeed, more generally, let ρ_1 and ρ_2 be lisse ℓ-adic sheaves on $V_1/\mathbf{F}_q$ and $V_2/\mathbf{F}_q$, respectively. There is then a natural homomorphism

$$\pi : \pi_1(V_1 \times V_2) \to \pi_1(V_1) \times \pi_1(V_2), \tag{3.6}$$

and it is elementary to check that the composite

$$\rho_1 \boxtimes \rho_2 : \pi_1(V_1 \times V_2) \to \pi_1(V_1) \times \pi_1(V_2) \xrightarrow{\rho_1 \otimes \rho_2} \mathrm{GL}(E_1 \otimes E_2)$$

has the desired property

$$\Lambda_{\rho_1 \boxtimes \rho_2}(x, y) = \Lambda_{\rho_1}(x)\Lambda_{\rho_2}(y),$$

using the fact that the Frobenius conjugacy class $\mathrm{Fr}_{(x,y),q^\nu}$ maps under π to the pair

$$(\mathrm{Fr}_{x,q^\nu}, \mathrm{Fr}_{y,q^\nu}).$$

This operation $\boxtimes$ is called the *external tensor product*; note that the rank of $\rho_1 \boxtimes \rho_2$ is the product of the ranks of the factors. We emphasize that it is not the same as the tensor product itself!

(As a final remark, the projection (3.6) can be shown to be surjective; however – and this is in contrast with the usual topological fundamental group – it is not always an isomorphism for V_1, $V_2/\mathbf{F}_q$: there are typically many more finite étale coverings of $V_1 \times V_2$ than those which can be constructed by the analogue of the external tensor product on the factors, the basic examples being given by the – many – Artin-Schreier coverings of $V_1 \times V_2$ which are not reducible to this type.)

Example 3.17 (Non-semisimple sheaf). We give here a very simple example of a non-semisimple lisse sheaf over a finite field. Indeed, our base variety will be $V = \mathrm{Spec}(\mathbf{F}_p)$, a single point with coefficients in a prime field with $p \geqslant 3$. We know then that $\pi_1(V) \simeq \mathrm{Gal}(\overline{\mathbf{F}}_q/\mathbf{F}_q)$. Now fix a prime $\ell \neq p$ and consider the splitting field k_ℓ in $\overline{\mathbf{F}}_p$ of the polynomial

$$X^\ell - 2 \in \mathbf{F}_p[X].$$

This is a Kummer equation, and hence there exists a homomorphism

$$\mathrm{Gal}(k_\ell/\mathbf{F}_p) \longrightarrow \mathbf{F}_\ell^\times \rtimes \mathbf{F}_\ell$$

where the semi-direct product on the right can be injected in $\mathrm{GL}_2(\mathbf{F}_\ell)$ by the homomorphism

$$(\xi, m) \mapsto \begin{pmatrix} \xi & m \\ 0 & 1 \end{pmatrix}.$$

Hence we obtain a homomorphism

$$\rho : \mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p) \longrightarrow \mathrm{GL}_2(\mathbf{F}_\ell).$$

As an $\mathbf{F}_\ell$-adic lisse sheaf acting on an $\mathbf{F}_\ell$-vector space E of rank 2 on $\mathrm{Spec}(\mathbf{F}_p)$, we claim that ρ is not semisimple whenever

$$\ell \nmid p - 1 \text{ and } X^\ell - 2 \text{ has no root in } \mathbf{F}_p.$$

Indeed, from the form of the matrices above, we have an exact sequence

$$0 \to \mathbf{F}_\ell(1) \to E \to \mathbf{F}_\ell \to 0$$

where the first map is the injection of the span of the first basis vector, the Tate twist corresponding to the fact that the arithmetic Frobenius $\mathrm{Gal}(\overline{\mathbf{F}}_p/\mathbf{F}_p)$ acts on it (as on ℓ-th roots of unity) by multiplication by q, while it acts trivially on the quotient. Indeed, if E were semisimple, there would be a fixed vector (corresponding to the trivial quotient $\mathbf{F}_\ell$). However, this is not possible under the conditions above (e.g., if (a, b) is the fixed vector, we would need

$$\xi a + mb = a$$

for all matrices in the image of ρ; since the ℓ-th roots of unity are not in $\mathbf{F}_p$, this is not possible with $b = 0$, and since $X^\ell - 2$ has not root in $\mathbf{F}_p$, the matrices with $\xi = 1$ and $m \in \mathbf{F}_\ell$ arbitrary are in the image, and exclude the possibility $b \neq 0$).

Although this is not an example at the level of $\overline{\mathbf{Q}}_\ell$-sheaves, it is possible to "boost" it to such a situation by considering the equations

$$X^{\ell^m} - 2 = 0$$

for $m \geqslant 1$ and putting them all together (in the spirit of the Tate modules of elliptic curves).

CHAPTER 4

The Riemann Hypothesis over finite fields

Bibliography

[1] E. Bombieri: *Counting points on curves over finite fields (d'après S. A. Stepanov)*, Séminaire N. Bourbaki, exposé no. 430, Lecture Notes in Math. 383 (1974), 234–241.

[2] E. Bombieri and J. Bourgain: *On Kahane's ultraflat polynomials*, J. Eur. Math. Soc. 11 (2009), 627–703.

[3] Z. Chatzidakis, L. van den Dries and A. Macintyre: *Definable sets over finite fields*, J. reine angew. Math. 427 (1992), 107–135

[4] P. Deligne: *Cohomologie étale*, S.G.A $4\frac{1}{2}$, L.N.M 569, Springer Verlag (1977).

[5] P. Deligne: *La conjecture de Weil : I*, Publ. Math. IHÉS 43 (1974), 273–307

[6] P. Deligne: *La conjecture de Weil, II*, Publ. Math. IHÉS 52 (1980), 137–252.

[7] É. Fouvry and N. Katz: *A general stratification theorem for exponential sums, and applications*, J. reine angew. Math. 540 (2001), 115–166.

[8] R. Hartshorne: *Algebraic geometry*, Grad. Texts in Math. 52, Springer-Verlag (1977).

[9] G.H. Hardy and E.M. Wright: *An introduction to the theory of numbers*, 5th Edition, Oxford Univ. Press, 1979.

[10] K. Ireland and M. Rosen: *A Classical Introduction to Modern Number Theory*, 2nd Edition, GTM 84, Springer-Verlag (1990).

[11] H. Iwaniec and E. Kowalski: *Analytic Number Theory*, A.M.S Colloq. Publ. 53, A.M.S (2004).

[12] N. Katz: *Moments, monodromy and perversity: a diophantine perspective*, Annals of Math. Studies 159, Princeton Univ. Press 2005.

[13] N. Katz: *Larsen's alternative, moments, and the monodromy of Lefschetz pencils*, Contributions to automorphic forms, geometry, and number theory, 521–560, Johns Hopkins Univ. Press, Baltimore, MD, 2004.

[14] N. Katz: *Gauss sums, Kloosterman sums and monodromy*, Annals of Math. Studies, 116, Princeton Univ. Press, 1988.

[15] N. Katz: *Exponential sums over finite fields and differential equations over the complex numbers: some interactions*, Bull. A.M.S 23 (1990), 269–309.

[16] N. Katz: *Sums of Betti numbers in arbitrary characteristic*, Finite Fields Appl. 7 (2001), no. 1, 29–44.

[17] N. Katz: *Twisted L-functions and monodromy*, Annals of Math. Studies 150, Princeton Univ. Press 2002.

[18] N. Katz and P. Sarnak: *Random matrices, Frobenius eigenvalues and monodromy*, A.M.S Colloquium Publ. 45, 1999.

[19] N. Katz and G. Laumon: *Transformation de Fourier et majoration de sommes exponentielles*, Publ. Math. I.H.É.S 62 (1985), 145–202.

[20] E. Kowalski: lecture notes and other documents related to this course, `http://www.math.ethz.ch/~kowalski/exp-sums.html`

[21] E. Kowalski: *Exponential sums over finite fields, I: elementary methods*, lectures notes from course at ETH Zürich, Spring Semester 2010, `http://www.math.ethz.ch/~kowalski/exp-sums.pdf`

[22] Q. Liu: *Algebraic geometry and arithmetic curves*, Oxford Grad. Texts in Math., 2002.

[23] W. Schmidt: *Equations over finite fields: an elementary approach*, Lecture Notes in Math. 536, Springer Verlag 1974.

[24] J. Silverman: *The arithmetic of elliptic curves*, Grad. Texts in Math. 106, Springer Verlag 1986.
[25] A. Weil: *Numbers of solutions of equations in finite fields*, Bull. A.M.S 55 (1949), 497–508.
[26] A. Weil: comments on [**25**], Collected Works, vol. I, 568–569, Springer 1979.
[27] H.B. Yu: *Note on Heath-Brown's estimate for Heilbronn's exponential sum*, Proc. American Math. Soc. 127 (1999), 1995–1998.