

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Mathematics

2016

A Tannakian Category and a Horizontal Equidistribution Conjecture for Exponential Sums

Sawin, Will

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp01dz010s47x>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

A TANNAKIAN CATEGORY AND A
HORIZONTAL EQUIDISTRIBUTION
CONJECTURE FOR EXPONENTIAL SUMS

WILLIAM FRENCH SAWIN

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
MATHEMATICS
ADVISER: NICHOLAS M. KATZ

MAY 2016

© Copyright by William French Sawin, 2016.

All rights reserved.

Abstract

We construct a Tannakian category of perverse sheaves on the additive group over the integers, with sheaf convolution as the tensor operation. A class of number-theoretic equidistribution problems, including the equidistribution of Gauss sums and Kloosterman sums and the equidistribution of roots of polynomial congruences, can be viewed as problems about the equidistribution of conjugacy classes in the Tannakian group of this category. Using this correspondence, we conjecture answers to these equidistribution questions.

Acknowledgements

Nick Katz taught me almost all of the techniques I used in the thesis, showed me the motivating problems, and pored over several drafts, fixing innumerable mathematical mistakes and weaknesses of exposition.

Peter Sarnak had many helpful conversations with me about the background and context for this work. Chapter 8 and most of Chapter 7 are direct results of his suggestions.

Pierre Deligne suggested the use of nearby cycles, which turned out to be critical.

J.-P. Serre and S.J. Patterson read drafts of this work and made many helpful suggestions.

Eric Naslund told me not to use the word “so” so much.

I had clarifying conversations on Mathoverflow with users David Speyer, Marty, and Qiaochu Yuan.

The contributions of Lisa Sawin, Steve Sawin, and Yihe Dong included, but were not limited to, helping me manage my time so I could actually finish this work.

A huge number of graduate students and others at Princeton helped me distract myself from my thesis with activities such as Hanabi, Dungeons and Dragons, both board game clubs, soccer, and just hanging out in the common room. Marie Bjerrum, Olivier Debarre, Chris Hall, Zhi Jiang, Peter Johnstone, Martí Lahoz, Tom Leinster, Philippe Michel, Paul Nelson, Doron Puder, Emmanuel Kowalski, Ashwath Rabindranath, Zeev Rudnick, and even Nick Katz helped me distract myself from my thesis by giving me other problems to work on. The help of both these groups was invaluable.

I would like to thank my readers, my committee, and anyone else who chooses to read this text, as well as all those mentioned previously for their contributions, and anyone I may have forgotten to mention.

Contents

Abstract	iii
Acknowledgements	iv
1 Introduction	1
2 Examples	5
3 The Tannakian Category	15
4 Properties of the Tannakian Category	62
5 An Equidistribution Conjecture	97
6 Complete Analysis of the Examples	108
7 Further Examples	124
8 Automorphic Forms and L-functions	155
9 Numerical Experiments	162
10 Small-Number Bias	169
Bibliography	182

Chapter 1

Introduction

A sequence x_n of points in a topological space X is equidistributed according to a probability measure μ on X if for each continuous bounded function f from X to $\mathbb{C}$,

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n f(x_k)}{n} = \int_X f \mu$$

and a sequence $x_p \in X$ indexed by the primes is equidistributed according to μ if for each continuous bounded function f from X to $\mathbb{C}$

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n, p \text{ prime}} f(x_p)}{\pi(n)} = \int_X f \mu$$

where $\pi(n)$ is the number of primes less than or equal to n .

Given a sequence x_p of arithmetic interest, we would like to know if it is equidistributed, and if so, according to what measure.

We would also like to have a general framework in which to consider many equidistribution questions. For instance, Dirichlet's theorem on primes in arithmetic progressions, Chebotarev's density theorem, Hecke's equidistribution theorem, and the Sato-Tate conjecture can all be viewed as statements about the equidistribution of traces of Frobenius conjugacy classes in Galois representations. Extending all these

statements, the generalized Sato-Tate conjecture predicts the distribution of the conjugacy classes of Frobenius elements in an arbitrary pure Galois representation of geometric origin [Serre, 2011, 8.2.2(A_2)]. This connection deepens our understanding of these questions and the relationships between them. Because these problems arise from Galois representations associated to motives, Serre called these problems motivated.

In addition, there are many equidistribution questions that cannot be expressed in terms of Frobenius conjugacy classes in Galois representations. Among these are the distribution of the roots of a polynomial modulo primes, the distribution of Gauss and Kummer sums, and the distribution of Kloosterman sums. Because these problems do not come from motives, Serre called them unmotivated. In what follows, we will propose a general framework that includes these cases, using ℓ -adic étale sheaves:

Let ℓ be a prime and let us fix an isomorphism between $\overline{\mathbb{Q}}_\ell$ and $\mathbb{C}$, allowing us to use them interchangeably in algebraic equations. Let $\mathcal{F}$ be a constructible $\overline{\mathbb{Q}}_\ell$ -sheaf on $\mathbb{A}_{\mathbb{Z}[\frac{1}{\ell}]}^1$. For each prime p and each element $x \in \mathbb{F}_p$, let $\text{tr}(\text{Frob}_p, \mathcal{F}, x)$ be the trace of the geometric Frobenius element in the Galois group of $\mathbb{F}_p$ acting on the stalk of $\mathcal{F}$ at the point (p, x) . Form the sequence $\{a_p(\mathcal{F})\}_p$ of complex numbers indexed by the primes:

$$a_p(\mathcal{F}) = \sum_{x=0}^{p-1} \text{tr}(\text{Frob}_p, \mathcal{F}, x) e^{2\pi i x/p}$$

We can ask the general equidistribution question: Is the sequence $\{a_p(\mathcal{F})\}_p$ once normalized through division by an appropriate power of $\sqrt{p}$, equidistributed according to some measure on $\mathbb{C}$? If so, according to what measure?

The structure of our paper is as follows:

In Chapter 2, we describe the aforementioned equidistribution questions (Gauss sums, Kloosterman sums, etc.) and show how to view them in terms of the distribution of the sequence $a_p(\mathcal{F})$ for appropriately chosen $\mathcal{F}$.

In Chapter 3, for each finite set of primes S including ℓ , we define a pro-algebraic group G_S over $\mathbb{C}$. The group G_S is an analogue of the Galois group of the maximal extension of $\mathbb{Q}$ unramified outside of S . For each prime p not in S , we define a conjugacy class Frob_p in $G_S(\mathbb{C})$. Then we prove the following theorem:

Theorem 1.0.1. *Let $\mathcal{F}$ be a constructible $\overline{\mathbb{Q}}_\ell$ -sheaf on $\mathbb{A}_{\mathbb{Z}[\frac{1}{\ell}]}^1$. There is a finite set of primes S and two representations V and W of G_S such that for all primes $p \notin S$;*

$$a_p(\mathcal{F}) = \text{tr}(\text{Frob}_p, V) - \text{tr}(\text{Frob}_p, W)$$

where $\text{tr}(\text{Frob}_p, V)$ is the trace of Frob_p acting on V .

This allows us to view the distributions of these exponential sums as distributions of traces of Frobenius conjugacy classes in a representation of a group analogous to the Galois group. Hopefully this makes these problems a little more motivated. (The groups G_S are all quotients of a single universal group, but this group does not contain Frobenius conjugacy classes, just as $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ does not contain Frobenius conjugacy classes.)

We will construct the group G_S and the classes Frob_p by constructing an appropriate Tannakian category.

In Chapter 4, we describe some useful properties of G_S and the associated Tannakian category.

In Chapter 5, we use our description of exponential sums in terms of representations of G_S to conjecture a distribution of $a_p(\mathcal{F})$, following along the lines of the generalized Sato-Tate conjecture.

In Chapter 6, we demonstrate that this conjecture agrees with the existing conjecture in the cases discussed in Chapter 2 and, in the cases where the answer is known, agrees with the known result.

In Chapter 7, we discuss further cases of the conjecture. We chose these cases to demonstrate the variety of different distributions that (conjecturally) appear as the distribution of $a_p(\mathcal{F})$.

In Chapter 8, we discuss the (seemingly grim) prospects for proving new cases of this conjecture.

In Chapter 9, we present numerical evidence for some cases of the conjecture.

In Chapter 10, we attempt to predict the rate of convergence of the distribution to its conjectural limit.

The technical arguments used in Chapter 2 to construct the Tannakian category were adapted from those used by Gabber and Loeser [1996, A3] to study multiplicative convolution. The idea of using a Tannakian category to answer an equidistribution question comes from Katz [2012, Theorem 1.1], who applied the Tannakian construction of Gabber and Loeser [1996] to answer a question about multiplicative character sums. The use of nearby cycles to specialize to characteristic p was suggested to me by Pierre Deligne. It is similar to the technique of nearby cycles specialization used by Krämer and Weissauer [2013, Lemma 14.1] to study sheaf convolution.

I would like to thank Nick Katz, Peter Sarnak, J.P. Serre, and S.J. Patterson for their helpful comments.

Chapter 2

Examples

In this chapter we will show how, knowing the distribution of $a_p(\mathcal{F})$ for appropriately chosen $\mathcal{F}$, one could determine the distribution of roots of a polynomial modulo primes, the distribution of Gauss and Kummer sums, and the distribution of Kloosterman sums. We will also show how to determine the distribution of a generalization of a_p involving a sheaf on an arbitrary scheme and an exponential of an arbitrary function, as well as answer a question about taking the roots of a polynomial modulo one prime and looking at them modulo a different prime.

At times we will also use knowledge of the joint distribution of $a_p(\mathcal{F}_1), \dots, a_p(\mathcal{F}_n)$ for multiple sheaves $\mathcal{F}_1, \dots, \mathcal{F}_n$ to answer these natural equidistribution questions.

Example 2.0.2. Let $f(x) \in \mathbb{Z}[x]$ be a polynomial. The Chebotarev density theorem tells us the distribution of the number of roots of f in $\mathbb{F}_p$ as p varies, but it does not tell us where those roots are located in $\mathbb{F}_p$. We can lift each root in $\mathbb{F}_p$ to a unique integer x in the interval $[0, p)$, then divide x by p to get a point in the unit interval $[0, 1)$. Alternatively, we can identify the ends of the interval to view x/p as a point in the unit circle $\mathbb{R}/\mathbb{Z}$.

Are the roots of f in $\mathbb{F}_p$ as p varies equidistributed in that interval according to the uniform measure, or perhaps another measure? (To ask this question precisely,

we need to define an order on the set of pairs $(p, x \in \mathbb{F}_p)$ with $f(x) = 0$. Order them lexicographically, with p first and x second.) This question was studied by Duke et al. [1995, pp. 2], taking up a theme of Hooley [1963, pp. 113], and their result was generalized by Tóth [2000, Theorem 1.2]. Estimates for the rate of convergence were obtained by Homma [2008, pp. 502] and Duke et al. [2012, Theorem 1.3].

We can fit this equidistribution question into our framework, by constructing a sequence of sheaves $\mathcal{F}_k$ for each $k \in \mathbb{Z}$ such that

$$a_p(\mathcal{F}_k) = \sum_{x \in \mathbb{F}_p, f(x)=0} e^{2\pi i k x/p}$$

Viewing the unit circle $\mathbb{R}/\mathbb{Z}$ as a group, Weyl's equidistribution criterion says that the sequence $\{x/p\}_{(p,x)}$ indexed by pairs $(p, x \in \mathbb{F}_p)$ with $f(x) = 0$ is equidistributed according to a measure μ if, for each character $e^{2\pi i k x}$ of $\mathbb{R}/\mathbb{Z}$, the average value of $x \mapsto e^{2\pi i k x}$ over the sequence is equal to $\int_{\mathbb{R}/\mathbb{Z}} e^{2\pi i k x} \mu(x)$. In other words, the sequence $\{x/p\}_{(p,x)}$ is equidistributed if for each $k \in \mathbb{Z}$

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} a_p(\mathcal{F}_k)}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}|} = \int_{\mathbb{R}/\mathbb{Z}} e^{2\pi i k x/p} \mu(x)$$

Let m be the number of irreducible factors of m . As $\lim_{n \rightarrow \infty} \sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}|/\pi(n) = m$ by Chebotarev's density theorem,

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} a_p(\mathcal{F}_k)}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}|} = \frac{1}{m} \lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} a_p(\mathcal{F}_k)}{\pi(n)}$$

so this limit is determined by the distribution of $a_p(\mathcal{F}_k)$. If we fix for each k a measure μ_k on $\mathbb{C}$ such that $a_p(\mathcal{F}_k)$ is equidistributed according to μ_k , then $\{x/p\}_{(p,x)}$ is equidistributed according to the unique measure μ , if it exists, such that $\int_{\mathbb{R}/\mathbb{Z}} e^{2\pi i k x/p} \mu(x) = \int_{\mathbb{C}} y \mu_k(y)$.

We now construct the $\mathcal{F}_k$. Let d be the degree of f and let

$$f_k(x) = f\left(\frac{x}{k}\right) k^d$$

Then the roots of f_k are k times the roots of f , so for primes not dividing k it is sufficient to find a sheaf $\mathcal{F}_k$ such that

$$a_p(\mathcal{F}_k) = \sum_{x \in \mathbb{F}_p, f_k(x)=0} e^{2\pi i x/p}$$

An $\mathcal{F}_k$ satisfying the following will do:

$$\mathrm{tr}(\mathrm{Frob}_p, \mathcal{F}_k, x) = \begin{cases} 1 & \text{if } f_k(x) \equiv 0 \pmod{p} \\ 0 & \text{if } f_k(x) \not\equiv 0 \pmod{p} \end{cases}$$

This is easy to do. Let i_k be the inclusion from $\mathrm{Spec}(\mathbb{Z}[\frac{1}{\ell}, x]/f_k(x))$ to $\mathrm{Spec}(\mathbb{Z}[\frac{1}{\ell}, x]) = \mathbb{A}_{\mathbb{Z}[\frac{1}{\ell}]}^1$ and let $\mathcal{F}_k = i_{k*} \overline{\mathbb{Q}}_\ell$. Then at each point (p, x) such that $f_k(x) \not\equiv 0 \pmod{p}$, the point (p, x) is not in the image of i_k , so the stalk of $\mathcal{F}_k$ is 0, for a trace of 0. For each point such that $f_k(x) \equiv 0 \pmod{p}$, (p, x) is in the image of just one point, so the stalk is the 1-dimensional vector space $\overline{\mathbb{Q}}_\ell$ with trivial Frobenius action, for a trace of 1. Thus $\mathcal{F}_k$ has exactly the desired property, hence the distribution of $\{a_p(\mathcal{F}_k)\}_p$ for all k determines the distribution of the roots of f .

Definition 2.0.3. Let n be a natural number. Let $\chi : \mu_n \rightarrow \overline{\mathbb{Q}}_\ell^\times$ be a faithful character. The Kummer sheaf $\mathcal{L}_\chi$ on $\mathbb{A}_{\mathbb{Z}[1/n, \mu_n]}^1 - \{0\}$ can be defined as follows:

The cover $\mathbb{A}_{\mathbb{Z}[1/n, \mu_n]}^1 - \{0\} \rightarrow \mathbb{A}_{\mathbb{Z}[1/n, \mu_n]}^1 - \{0\}$ which sends x to x^n is finite étale Galois with automorphism group μ_n . By the definition of the étale fundamental group, any such cover determines a homomorphism $\pi_1(\mathbb{A}_{\mathbb{Z}[1/n, \mu_n]}^1 - \{0\}) \rightarrow \mu_n$. Composing with χ gives a homomorphism $\pi_1(\mathbb{A}_{\mathbb{Z}[1/n, \mu_n]}^1 - \{0\}) \rightarrow \overline{\mathbb{Q}}_\ell^\times$, which defines a rank 1 lisse ℓ -adic sheaf $\mathcal{L}_\chi$.

For any $\mathbb{Z}[1/n, \mu_n]$ -algebra R , we define a Kummer sheaf on $\mathbb{A}_R^1 - \{0\}$ by pullback.

Let $q \equiv 1$ modulo n be a prime power. If $R = \mathbb{F}_q$ then for $x \in \mathbb{F}_q^\times$, the action of Frob_q on $\mathcal{L}_\chi$ at x is determined by its action on the n th roots of x . The arithmetic Frobenius sends $\sqrt[n]{x}$ to

$$\left(\sqrt[n]{x}\right)^q = x^{\frac{q-1}{n}} \sqrt[n]{x}$$

Hence the action of the arithmetic Frobenius is multiplication by the n th root of unity $x^{\frac{q-1}{n}}$, and the geometric Frobenius, its inverse, acts as multiplication by the inverse root of unity, thus

$$\text{tr}(\mathcal{L}_\chi, \text{Frob}_q, x) = \chi(x^{-\frac{q-1}{n}})$$

The map $x \mapsto \chi(x^{-\frac{q-1}{n}})$ is a character of $\mathbb{F}_q^\times$ of order n . Every character of order n can be expressed in this way; $\mathbb{F}_q^\times$ is a cyclic group of order $q-1$ so every character of order n is divisible by $\frac{q-1}{n}$.

Because $\mathbb{F}_q^\times$ is cyclic, as χ runs through all the faithful characters of μ_n , then $\chi(x^{-\frac{q-1}{n}})$ runs through all the order n characters of $\mathbb{F}_q^\times$. For $q = p$ a prime these are the order n Dirichlet characters mod p .

Example 2.0.4. Recall the Gauss sum:

$$G(\chi, \psi) = \sum_{x=1}^{p-1} \chi(x) \psi(x)$$

where χ is a character of the multiplicative group of $\mathbb{F}_p$ and ψ is a character of the additive group. We can consider the distribution as p varies and ψ remains the standard exponential character $\psi(x) = e^{2\pi i \frac{x}{p}}$. However there is no obvious nice way to choose χ . For simplicity we can just sum over all characters χ of some exact order n . These only exist when $p \equiv 1$ modulo n , so the sum will be 0 for other p . We will have

$$a_p = \sum_{\chi} \sum_{x=1}^{p-1} \chi(x) e^{2\pi i x/p} = \sum_{x=0}^{p-1} \left(\sum_{\chi} \chi(x) \right) e^{2\pi i x/p}$$

As $\sum_{y=0}^{p-1} e^{2\pi i y^n/p}$ is $\sum_{x=0}^{p-1} \left(\sum_{\chi|\chi^n=1} \chi(x) \right) e^{2\pi i x/p}$, by inclusion-exclusion we have (with μ the Möbius function)

$$a_p = \sum_{d|n} \mu(d) \sum_{y=0}^{p-1} e^{2\pi i y^{n/d}/p}$$

We need a sheaf $\mathcal{F}$ such that:

$$\mathrm{tr}(\mathcal{F}, \mathrm{Frob}_p, x) = \begin{cases} \sum_{\chi} \chi(x) & \text{if } p \equiv 1 \pmod{n} \\ 0 & \text{if } p \not\equiv 1 \pmod{n} \end{cases}$$

where the sum is over χ of order exactly n .

This sheaf can be constructed from the étale cover of $\mathbb{G}_{m, \mathbb{Z}[\frac{1}{n}]}$ defined by adjoining the n th roots of unity μ_n and the n th root $\sqrt[n]{x}$ of the coordinate x . This cover has automorphism group the semidirect product $\mu_n \rtimes (\mathbb{Z}/n)^\times$, with the action of $a \in (\mathbb{Z}/n)^\times$ on $\zeta \in \mu_n$ sending ζ to ζ^a . The group acts on the cover where $(\zeta, a) \in \mu_n \rtimes (\mathbb{Z}/n)^\times$ sends generators $\sqrt[n]{x}$ and ζ' of $\mathbb{Z}[x, x^{-1}, \mu_n, \sqrt[n]{x}]$ to $\zeta \sqrt[n]{x}$ and ζ'^a .

This cover defines a surjective homomorphism $\pi_1 \left(\mathbb{G}_{m, \mathbb{Z}[\frac{1}{n}]} \right) \rightarrow \mu_n \rtimes (\mathbb{Z}/n)^\times$. By composing with the induced representation from μ_n to $\mu_n \rtimes (\mathbb{Z}/n)^\times$ of a faithful character of μ_n , we obtain a lisse sheaf on $\mathbb{G}_{m, \mathbb{Z}[\frac{1}{n}]}$, which we extend by zero to all of $\mathbb{A}_{\mathbb{Z}[\frac{1}{\ell}]}^1$, obtaining $\mathcal{F}$.

Thus $\mathcal{F}$ has the desired trace: Because the representation is induced from a normal subgroup, the trace of any element not in this normal subgroup is 0. If $p \not\equiv 1 \pmod{n}$, then Frob_p does not fix μ_n , thus it is not in that subgroup, so the trace is 0.

For $p \equiv 1 \pmod{n}$, the cover of $\mathbb{G}_{m, \mathbb{F}_p}$ defined by adjoining μ_n is trivial. Thus the representation $\pi_1 \left(\mathbb{G}_{m, \mathbb{F}_p} \right) \rightarrow \mu_n \rtimes (\mathbb{Z}/n)^\times$ has image contained in μ_n . When restricted to the μ_n subgroup, the induced representation is the sum of all the conjugates of a faithful one-dimensional representation of μ_n , hence the sum of all faithful 1-dimensional representations of μ_n .

Hence $\mathcal{F}$ is the sum of the Kummer sheaves for each faithful 1-dimensional character of μ_n . Thus the trace of Frobenius is the sum of the traces of these 1-dimensional sheaves, which is the sum of $\chi(x)$ for each Dirichlet character χ of order exactly n .

An important special case is when $n = 3$. For $p \equiv 1 \pmod{3}$, there are just two characters of $\mathbb{F}_p^\times$ of order exactly 3, χ and $\bar{\chi}$, and our a_p is the Kummer sum

$$K(p) = G\left(\chi, e^{2\pi i \frac{x}{p}}\right) + G\left(\bar{\chi}, e^{2\pi i \frac{x}{p}}\right)$$

It is always real and lies in the interval $[-2\sqrt{p}, 2\sqrt{p}]$. Kummer divided the sum by $\sqrt{p}$ and counted for how many primes congruent to 1 mod 3 it lies in the intervals $[-2, -1]$, $[-1, 1]$, and $[1, 2]$. He conjectured, based on numerical evidence, that the primes might be distributed in the three intervals according to the ratio 1 : 2 : 3 [Kummer, 1846, pp. 353]. This was disproved by [Heath-Brown and Patterson, 1979, Theorem 1], who showed that the ratio was 1 : 1 : 1 by exactly computing the distribution of a_p . They also explained Kummer's misleading evidence - this particular sum suffers from small-number bias, which means the average of a_p over $p \leq n$ is significantly larger for small n than in the limit as n tends to infinity. We will discuss this more in Chapter 10.

Rather than adding up all the Gauss sums corresponding to characters of order n , we can take the set of all $\varphi(n)$ Gauss sums, divide each by $\sqrt{p}$, and view it as a set of $\varphi(n)$ points on the unit circle. (Since a Gauss sum is a complex number of absolute value $\sqrt{p}$, dividing it by $\sqrt{p}$ produces a point in the unit circle.) Equivalently, a set of $\varphi(n)$ points on S^1 is a point $x_p \in \text{Sym}^{\varphi(n)}(S^1)$. The distribution of these x_p in that space will also be a special case of our general equidistribution conjecture for suitable sheaves $\mathcal{F}$. This will be explained in Lemma 6.0.97.

Example 2.0.5. The Kloosterman sum is defined as

$$\mathcal{Kl}(a, b; p) = \sum_{y=1}^{p-1} e^{2\pi i \frac{ay+by^{-1}}{p}}$$

where y^{-1} is understood to be the inverse of y modulo p . The distribution of this sum for fixed a, b and varying p translates in a straightforward way into this theory.

We can rewrite this sum as

$$\sum_{y=1}^{p-1} e^{2\pi i \frac{ay+by^{-1}}{p}} = \sum_{x=0}^{p-1} |\{y \in \mathbb{F}_p^\times \mid ay + by^{-1} = x\}| e^{2\pi i \frac{x}{p}}$$

To obtain these coefficients, we need a sheaf $\mathcal{F}$ on $\mathbb{A}_{\mathbb{Z}[\frac{1}{\ell}]}^1$ whose trace of Frobenius at a point (p, x) is equal to the number of solutions of the equation $ay + by^{-1} = x$ in $\mathbb{F}_p^\times$. By the Lefschetz trace formula, the pushforward of the constant sheaf $\overline{\mathbb{Q}}_\ell$ along the natural map from $\mathbb{A}_{\mathbb{Z}}^1 - \{0\}$ to $\mathbb{A}_{\mathbb{Z}}^1$ defined by $x = ay + by^{-1}$ has exactly the desired trace.

Katz [1988, 13.5.3] computed the distribution of these sums as a or b varies and p goes to ∞ , and raised the question of the distribution when a and b are fixed and p varies. This question remains unanswered, but there has been partial progress on the equidistribution of Kloosterman sums with varying composite modulus [Xi, 2015, Theorem 1.1]

Example 2.0.6. Example 2.0.5 is a special case of a general construction. Let X be any scheme of finite type over $\mathbb{Z}[\frac{1}{\ell}]$, $\mathcal{F}$ any constructible $\overline{\mathbb{Q}}_\ell$ -sheaf on X , and $f \in \Gamma(X, \mathcal{O}_X)$ any regular function on X . Define

$$a_p(X, \mathcal{F}, f) = \sum_{y \in X(\mathbb{F}_p)} \text{tr}(\text{Frob}_p, \mathcal{F}, y) e^{2\pi i \frac{f(y)}{p}}$$

Then we can rewrite the sum:

$$a_p(X, \mathcal{F}, f) = \sum_{x \in \mathbb{F}_p} \left(\sum_{y \in X(\mathbb{F}_p), f(y)=x} \text{tr}(\text{Frob}_p, \mathcal{F}, y) \right) e^{2\pi i \frac{x}{p}}$$

Using the Lefschetz trace formula, this is:

$$\sum_{x \in \mathbb{F}_p} \left(\sum_{i=0}^{2 \dim X} (-1)^i \text{tr}(\text{Frob}_p, R^i f_! \mathcal{F}, x) \right) e^{2\pi i \frac{x}{p}} = \sum_{i=0}^{2 \dim X} (-1)^i a_p(R^i f_! \mathcal{F})$$

Therefore summing over an arbitrary scheme, and taking the exponential of an arbitrary function, does not make our equidistribution question more general - the new sum can be computed from the old.

Example 2.0.7. One simple special case of Example 2.0.6 is summing the exponential of a polynomial function on $\mathbb{A}^1$. Setting $X = \mathbb{A}^1$, $\mathcal{F} = \overline{\mathbb{Q}}_\ell$, and $f(x) \in \mathbb{Z}[x]$ a polynomial, we obtain the exponential sum:

$$S(f(x), p) = \sum_{x \in \mathbb{F}_p} e^{2\pi i \frac{f(x)}{p}}$$

The distribution of this as p varies is closely related to the problem studied in [Patterson, 2003], which is about the distribution of the analogous sum over $\mathbb{Z}/n$ as n varies.

Example 2.0.8. Let D be an integer that is not a perfect square and let m be a positive integer. For p a prime such that D is a square modulo p , let a be the unique integer satisfying $0 < a < \frac{p}{2}$ and $a^2 \equiv D \pmod{p}$. Rather than considering the distribution of a/p in the interval $(0, 1/2)$, we could consider the distribution of the congruence class of a modulo m as p varies. (A version of this question was suggested in a lecture by Serre [2012].)

Perhaps surprisingly, this question fits into our framework as well.

Suppose that the sequence $\left(\frac{x}{p}, p \bmod m\right)$ in $S^1 \times \mathbb{Z}/m$ indexed by pairs of a prime p and a root x of the equation $(mx)^2 - D \cong 0 \bmod p$ is equidistributed on $S^1 \times \mathbb{Z}/m$ according to the product of the uniform measure on S^1 and the uniform measure on the set of invertible residue classes modulo m . Then the congruence classes of $a \bmod m$ is equidistributed according to the uniform measure on $\mathbb{Z}/m$ as p varies.

(We will then show, by an argument similar to Example 2.0.2, that this equidistribution problem of pairs follows from equidistribution of suitable sheafs $a_p(\mathcal{F}_n, \chi)$.)

To see this, let b be a root of $(mb)^2 - D \equiv 0 \bmod p$ between 0 and p and suppose that p does not divide m . We claim that which of the $2m$ intervals $\left[0, \frac{p}{2m}\right], \left[\frac{p}{2m}, \frac{2p}{2m}\right], \dots, \left[\frac{(2m-1)p}{2m}, p\right]$ contains b , combined with the residue class of p modulo m , determines the residue class of a modulo m :

To prove the claim, we use the fact that $a \cong \pm b \bmod p$, so we can write $a = \epsilon mb + kp$ for some sign $\epsilon = \pm 1$ and integer k . To determine ϵ and k , we use the fact that $0 < a < \frac{p}{2}$, and therefore:

$$\frac{-2kp}{2m} < \epsilon b < \frac{(1-2k)p}{2m}$$

Thus if b is in an interval of the form $\left[\frac{2np}{2m}, \frac{(2n+1)p}{2m}\right]$, then $\epsilon = 1$ and $k = -n$, and if b is in an interval of the form $\left[\frac{(2n+1)p}{2m}, \frac{(2n+2)p}{2m}\right]$ then $\epsilon = -1$ and $k = n + 1$.

Because $a \equiv kp \bmod m$, so $a \bmod m$ is determined by $k \bmod m$ and by $p \bmod m$. Because b/p is uniformly distributed, it is equally likely to be in any of the intervals, hence k is equally likely to be any residue class mod m . (Each residue class corresponds to exactly two intervals.) As b is independent of $p \bmod m$, this is true for each fixed invertible residue class of p modulo m , thus for each invertible residue class a is uniformly distributed in $(\mathbb{Z}/m)$, hence it is uniformly distributed overall.

Here it is crucial that the uniform measure on S^1 places zero measure on the boundary of the interval, so we can conclude from equidistribution that the fraction of b lying in the interval is equal to the measure of the interval.

Now observe that the equidistribution of this set of pairs follows from equidistribution conjectures on $a_p(\mathcal{F}_{n,\chi})$ for suitably chosen n, χ . Ignoring the finitely many p dividing m , the sequence $\left(\frac{x}{p}, p \bmod m\right)$ lies in $S^1 \times (\mathbb{Z}/m)^\times$, which is a compact group. Hence by Weyl's equidistribution theorem, the sequence is equidistributed if for each nontrivial character of the group, the integral of the character vanishes. This happens if for each $k \in \mathbb{Z}$ and $\chi(\mathbb{Z}/m)^\times \rightarrow \mathbb{C}^\times$

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n, \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p) e^{2\pi i \frac{kx}{p}}}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | (mx)^2 = D\}|} = 0$$

unless $k = 0$ and $\chi = 1$. Let $\mathcal{F}_{k,\chi}$ be the sheaf $\mathcal{F}_k$ from Example 2.0.2 such that $a_p(\mathcal{F}_k) = \sum_{x \in \mathbb{F}_p | (mx)^2 = D} e^{2\pi i \frac{kx}{p}}$ tensored with the 1-dimensional Galois representation associated by the Dirichlet character χ . Then $a_p(\mathcal{F}_{k,\chi}) = \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p) e^{2\pi i \frac{kx}{p}}$. Since we can estimate the denominator using Dirichlet's theorem, the distribution of our original sequence is determined by the distribution of $a_p(\mathcal{F}_{k,\chi})$.

We have claimed that these examples do not fall under the purview of the generalized Sato-Tate conjecture. Indeed, there is no obvious formula relating these exponential sums and the traces of Frobenius of Galois representations. But might there be a non-obvious formula? The answer is probably no. Any Galois representation arising from geometry should have its traces of Frobenius in a fixed number field, but it is easy to check that these exponential sums a_p do not lie in any fixed number field. We will discuss this issue in more detail, as well as the issue of whether the a_p could be the Hecke eigenvalues of an automorphic form, in Chapter 8.

Chapter 3

The Tannakian Category

In this section, we prove Theorem 1.0.1. We will make free use of the standard theorems about étale cohomology as in [Milne, 1980].

As a first step to Theorem 1.0.1, we can indeed write our exponential sums in terms of traces of Frobenius conjugacy classes. To do this, we need to define the Artin-Schreier sheaf.

Definition 3.0.9. The Artin-Schreier cover of $\mathbb{A}_{\mathbb{F}_p}^1$ is the map $\mathbb{A}_{\mathbb{F}_p}^1 \rightarrow \mathbb{A}_{\mathbb{F}_p}^1$ defined by $x \rightarrow x - x^p$. This is a finite étale Galois cover with automorphism group $\mathbb{Z}/p$, so it defines a homomorphism $\pi_1(\mathbb{A}_{\mathbb{F}_p}^1) \rightarrow \mathbb{Z}/p$. By composing this with a character ψ of $\mathbb{Z}/p$ defined over $\overline{\mathbb{Q}_\ell}$, we obtain a character of $\pi_1(\mathbb{A}_{\mathbb{F}_p}^1)$, and hence a rank one lisse $\overline{\mathbb{Q}_\ell}$ -sheaf on $\mathbb{A}_{\mathbb{F}_p}^1$. This is called the Artin-Schreier sheaf, $\mathcal{L}_\psi$.

If $x \in \mathbb{F}_p$ and $y - y^p = x$, then the arithmetic Frobenius acts on y by

$$y^p = y - x$$

so the geometric Frobenius, its inverse, sends y to $y + x$, so Frob_p at x acts on the covering by adding x , thus after composing with ψ we get $\text{tr}(\text{Frob}_p, \mathcal{L}_\psi, x) = \psi(x)$.

Let ψ be the character of $\mathbb{Z}/p$ that takes x to $e^{2\pi i \frac{x}{p}}$, using the isomorphism between $\overline{\mathbb{Q}}_\ell$ and $\mathbb{C}$ that we fixed in the introduction. Then $\text{tr}(\text{Frob}_p, \mathcal{L}_\psi, x) = e^{2\pi i \frac{x}{p}}$, so for any constructible ℓ -adic sheaf $\mathcal{F}$ on $\mathbb{A}_{\mathbb{F}_p}^1$

$$\sum_{x=0}^{p-1} e^{2\pi i \frac{x}{p}} \text{tr}(\text{Frob}_p, \mathcal{F}, x) = \sum_{i=0}^2 (-1)^i \text{tr}(\text{Frob}_p, H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, \mathcal{F} \otimes \mathcal{L}_\psi)) \quad (3.1)$$

where H_c^i is étale cohomology with compact supports. This is an immediate consequence of the Lefschetz trace formula and the fact that the compactly supported cohomology of any sheaf on $\mathbb{A}^1$ vanishes outside of degrees 0, 1, and 2.

However, given a sheaf $\mathcal{F}$ on $\mathbb{A}_{\mathbb{Z}}^1$, the Frobenius conjugacy classes we have just defined for the restriction of $\mathcal{F}_{\mathbb{A}_{\mathbb{F}_p}^1}$ are not obviously conjugacy classes in the same group on all p . These conjugacy classes each act on the cohomology group $H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, \mathcal{F}_{\mathbb{A}_{\mathbb{F}_p}^1} \otimes \mathcal{L}_\psi)$. There is no obvious relationship between these groups, because there is usually no a sheaf on $\mathbb{A}_{\mathbb{Z}}^1$ whose pullback to $\mathbb{A}_{\mathbb{F}_p}^1$ is $\mathcal{F}_{\mathbb{A}_{\mathbb{F}_p}^1} \otimes \mathcal{L}_\psi$ for every prime p . (The only exception is when $\mathcal{F}$ is supported on the zero section. Then $\mathcal{F}_{\mathbb{A}_{\mathbb{F}_p}^1} \otimes \mathcal{L}_\psi = \mathcal{F}_{\mathbb{A}_{\mathbb{F}_p}^1}$.)

We can construct such a group, though, through the machinery of Tannakian categories. For a reference on Tannakian categories, see [Deligne and Milne, 1982]. If we can construct a Tannakian category of sheaves on $\mathbb{A}_{\mathbb{Z}}^1$ such that for a given p , pulling back to $\mathbb{A}_{\mathbb{F}_p}^1$, tensoring with $\mathcal{L}_\psi$, and taking cohomology is a fibre functor, then Frobenius will be an automorphism of that functor, and hence will naturally give rise to a conjugacy class in the Tannakian group. If we can do this for all but finitely many primes p , then we will have Frobenius conjugacy classes for all those primes in our Tannakian group.

For the tensor product in our Tannakian category, we will use sheaf convolution:

Definition 3.0.10. Let R be a Noetherian ring with ℓ invertible. Let $D_c^b(\mathbb{A}_R^1, \overline{\mathbb{Q}}_\ell)$ be the bounded derived category of complexes of $\overline{\mathbb{Q}}_\ell$ -sheaves with constructible cohomology on $\mathbb{A}_R^1$.

We will define two functors $*, *! : D_c^b(\mathbb{A}_R^1, \overline{\mathbb{Q}}_\ell) \times D_c^b(\mathbb{A}_R^1, \overline{\mathbb{Q}}_\ell) \rightarrow D_c^b(\mathbb{A}_R^1, \overline{\mathbb{Q}}_\ell)$.

Let A and B be two objects of $D_c^b(\mathbb{A}_R^1, \overline{\mathbb{Q}}_\ell)$.

There are three natural maps $\mathbb{A}^1 \times \mathbb{A}^1 \rightarrow \mathbb{A}^1$ we will use. We have projection onto the first factor pr_1 and projection onto the second factor pr_2 . Viewing $\mathbb{A}^1$ as $\mathbb{G}_a$, we have the group operation $sum : \mathbb{G}_a \times \mathbb{G}_a \rightarrow \mathbb{G}_a$. Using these, we define convolution:

$$A * B = Rsum_*(pr_1^* A \otimes pr_2^* B)$$

and convolution with compact supports:

$$A *_! B = Rsum_!(pr_1^* A \otimes pr_2^* B)$$

We will write $pr_1^* A \otimes pr_2^* B$ as $A \boxtimes B$.

Then we have the identity:

$$H_c^i(\mathbb{A}_{\mathbb{F}_p}^1 (A *_! B) \otimes \mathcal{L}_\psi)$$

(by definition of $*_!$)

$$= H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, Rsum_!(A \boxtimes B) \otimes \mathcal{L}_\psi)$$

(by the Leray spectral sequence [Artin et al., 1972, XVII, Proposition 5.2.9])

$$= H_c^i(\mathbb{A}_{\mathbb{F}_p}^2, (A \boxtimes B) \otimes sum^* \mathcal{L}_\psi)$$

(by the fact that $sum^* \mathcal{L}_\psi = pr_1^* \mathcal{L}_\psi \otimes pr_2^* \mathcal{L}_\psi$ [Katz, 1988, Section 4.3 (3)])

$$= H_c^i(\mathbb{A}_{\mathbb{F}_p}^2, pr_1^* A \otimes pr_2^* B \otimes pr_1^* \mathcal{L}_\psi \otimes pr_2^* \mathcal{L}_\psi)$$

(by the Künneth formula [Artin et al., 1972, XVII, Theorem 5.4.3])

$$= \sum_{a+b=i} H_c^a(\mathbb{A}_{\mathbb{F}_p}^1, A \otimes \mathcal{L}_\psi) \otimes H_c^b(\mathbb{A}_{\mathbb{F}_p}^1 B \otimes \mathcal{L}_\psi)$$

By proper base change, performing $*_!$ over $\mathbb{Z}$ and then reducing modulo p is the same as reducing modulo p and then performing $*_!$, so convolving our sheaves on $\mathbb{A}_{\mathbb{Z}}^1$ will indeed give us a good notion of tensor product.

To get a Tannakian category, we will need to make some technical modifications. The objects of our category will not quite be sheaves on $\mathbb{A}_{\mathbb{Z}}^1$. Because convolution is naturally a derived category operation, the convolution of two sheaves need not be a sheaf. Instead we will use perverse sheaves on $\mathbb{A}_{\mathbb{Q}}^1$, impose some conditions of “good reduction” at primes p not in S , and use localization to set certain negligible sheaves to 0.

We will be able to show that this category is a Tannakian category, and that the Tannakian group we get from it includes all the Frobenius elements of Equation (3.1). To that end we will do the following:

1. First show that convolution makes the category of perverse sheaves on $\mathbb{A}_{\mathbb{Q}}^1$, modulo the negligible sheaves, into a rigid symmetric monoidal abelian category.
2. (a) Next define a notion of the specialization of the sheaf mod p via the nearby cycles functor.
 (b) Using this notion of specialization mod p , define the notion of “good reduction” at p .
3. Show that in the category of perverse sheaves with good reduction at p , specialization mod p followed by tensoring with $\mathcal{L}_\psi$ and taking compact cohomology is a fibre functor, making that subcategory into a neutral Tannakian category.

4. Show that starting with a constructible sheaf on $\mathbb{A}_{\mathbb{Z}}^1$, taking the fiber on $\mathbb{A}_{\mathbb{Q}}^1$, and specializing mod p in this way agrees with the usual notion of specialization at all but finitely many primes, so that the conjugacy classes of the Frobenius elements in Equation (3.1) correspond to conjugacy classes of the Tannakian group.

These steps will culminate in a proof of Theorem 1.0.1.

Recall that a complex K of constructible $\overline{\mathbb{Q}}_\ell$ -sheaves on a curve over a field is perverse if and only if its cohomology objects $\mathcal{H}^i(K)$ vanish for $i \neq -1, 0$, are punctual for $i = 0$, and have no nonzero punctual sections for $i = -1$ [Katz, 1995, 2.3.6].

Definition 3.0.11. Let $\mathcal{Perv}$ be the category of perverse sheaves on $\mathbb{A}_{\mathbb{Q}}^1$ with coefficients in $\overline{\mathbb{Q}}_\ell$. Let $\mathcal{Neg} \subset \mathcal{Perv}$ be the category of geometrically constant sheaves placed in degree -1 . Let $D^b(\mathcal{Neg}) \subset D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_\ell)$ be the category of complexes of sheaves with geometrically constant perverse cohomology sheaves, equivalently, geometrically constant cohomology sheaves [Katz, 1995, Lemma 2.3.2.1].

We will construct a Tannakian category structure on an appropriate subcategory of $\mathcal{Perv}/\mathcal{Neg}$.

Remark 3.0.12. We first describe a key technique we will use repeatedly when studying convolution. Over any ring R , let d be the map $\mathbb{G}_a \times \mathbb{G}_a \rightarrow \mathbb{G}_a$ given by the function $d(x, y) = x - y$, i.e. subtraction under the group law. Then:

$$A * B = Rsum_*(pr_1^* A \otimes pr_2^* B) = Rpr_{1*}(d^* A \otimes pr_2^* B)$$

$$A *_! B = Rsum_!(pr_1^* A \otimes pr_2^* B) = Rpr_{1!}(d^* A \otimes pr_2^* B)$$

This equivalence is simply a change of variables on $\mathbb{A}_R^2$. View $\mathbb{A}_R^2$ as a plane with coordinate x, y . Then the function pr_1 is just x , pr_2 is y , and sum is $x + y$. If we

replace (x, y) with (x', y') where $x' = x + y$ and $y' = y$, then in the new coordinates pr_1 is $x' - y'$, pr_2 is y' , and sum is x' . So the maps pr_1, pr_2 and sum in the old coordinates become d, pr_2 , and pr_1 in the new coordinates.

To compute compactly supported cohomology along pr_1 , we use the most natural compactification. Let $u : \mathbb{A}^1 \times \mathbb{A}^1 \rightarrow \mathbb{A}^1 \times \mathbb{P}^1$ be the open immersion given by taking the product of the identity from $\mathbb{A}^1$ to $\mathbb{A}^1$ and the immersion $j : \mathbb{A}^1 \rightarrow \mathbb{P}^1$. Let $\pi : \mathbb{A}^1 \times \mathbb{P}^1 \rightarrow \mathbb{A}^1$ be the projection onto the first factor. Then π is a compactification of pr_1 , and so we can write:

$$A *_! B = R\pi_* u_!(d^* A \otimes pr_2^* B)$$

We will sometimes also use the compactification to compute the usual cohomology:

$$A * B = R\pi_* Ru_*(d^* A \otimes pr_2^* B)$$

We will work out a number of basic facts about sheaf convolution on $\mathbb{A}_{\mathbb{Q}}^1$:

Lemma 3.0.13. *$\mathcal{N}eg$ is a Serre subcategory of $\mathcal{P}erv$ and $D^b(\mathcal{N}eg)$ is a thick subcategory of $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_{\ell})$. (The notion of a thick subcategory of a triangulated category is defined in [Gabber and Loeser, 1996, 3.6].)*

Proof. By definition, $\mathcal{N}eg$ is a Serre subcategory if and only if for an exact sequence $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ of elements of $\mathcal{P}erv$, $B \in \mathcal{N}eg$ if and only if $A \in \mathcal{N}eg$ and $C \in \mathcal{N}eg$.

To prove if, note that

$$\mathcal{E}xt_{\mathbb{A}_{\mathbb{Q}}^1}^1(\overline{\mathbb{Q}}_{\ell}, \overline{\mathbb{Q}}_{\ell}) = H^1(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_{\ell}) = 0$$

so any extension of constant sheaves on $\mathbb{A}_{\mathbb{Q}}^1$ is constant, so any extension of geometrically constant sheaves on $\mathbb{A}_{\mathbb{Q}}^1$ is geometrically constant.

To prove only if, geometrically B is a sum of copies of $\overline{\mathbb{Q}}_\ell[1]$. Because $\overline{\mathbb{Q}}_\ell$ is an irreducible lisse sheaf, $\overline{\mathbb{Q}}_\ell[1]$ is an irreducible perverse sheaf, so because B is geometrically semisimple in $\mathcal{Perv}$, A and C are both geometrically sums of copies of $\overline{\mathbb{Q}}_\ell$ - i.e. geometrically constant.

Since $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_\ell) = D^b(\mathcal{Perv})$, $D^b(\mathcal{Neg})$ is a thick subcategory of $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_\ell)$, by [Gabber and Loeser, 1996, Proposition 3.6.1]. $\square$

Lemma 3.0.14. *Let A and B be two perverse sheaves on $\mathbb{A}_{\mathbb{Q}}^1$. Then the natural forget supports map $A *_! B \rightarrow A * B$ is an isomorphism modulo $D^b(\mathcal{Neg})$.*

Proof. We can fit the forget supports map $A *_! B \rightarrow A * B$ into a distinguished triangle $A *_! B \rightarrow A * B \rightarrow C$ using the mapping cone C . To show the map is an isomorphism modulo $D^b(\mathcal{Neg})$, we must show that C is in $D^b(\mathcal{Neg})$.

To compute this mapping cone, we use the compactification of Remark 3.0.12:

$$A *_! B = Rsum_!(A \boxtimes B) = R\pi_* u_!(d^* A \otimes pr_2^* B)$$

$$A * B = Rsum_*(A \boxtimes B) = R\pi_* Ru_*(d^* A \otimes pr_2^* B)$$

The following commutative diagram should be helpful:

$$\begin{array}{ccccc} \mathbb{A}^1 \times \infty & \xrightarrow{i} & \mathbb{A}^1 \times \mathbb{P}^1 & \xleftarrow{u} & \mathbb{A}^1 \times \mathbb{A}^1 & \xrightarrow{pr_2} & \mathbb{A}^1 \\ & & \downarrow \pi & & \downarrow d & & \\ & & \mathbb{A}^1 & & \mathbb{A}^1 & & \end{array}$$

Let V be the mapping cone that completes the distinguished triangle $u_!(d^* A \otimes pr_2^* B) \rightarrow Ru_*(d^* A \otimes pr_2^* B) \rightarrow V$. Then because $R\pi_*$ preserves distinguished triangles, $C = R\pi_* V$.

The map $u_!(d^* A \otimes pr_2^* B) \rightarrow Ru_*(d^* A \otimes pr_2^* B)$ is an isomorphism on the image of u , which is $\mathbb{A}^1 \times \mathbb{A}^1$ inside $\mathbb{A}^1 \times \mathbb{P}^1$, thus V vanishes on $\mathbb{A}^1 \times \mathbb{A}^1$, so V is supported on the complement, $\mathbb{A}^1 \times \infty$. We will compute V étale-locally on the line $\mathbb{A}^1 \times \infty$ by

working over the étale local ring of a point (d, ∞) on the line $\mathbb{A}^1 \times \infty$. By definition $u_!(d^*A \otimes pr_2^*B)$ vanishes on that line, so this is the same as computing $Ru_*(d^*A \otimes pr_2^*B)$ on that line.

$d^*A \otimes pr_2^*B$ is lisse away from the inverse image along d of the singular points of A and the inverse image along pr_2 of the singular points of B . These are given by lines with equations of the form $x' - y' = c$ or $y' = c$.

The key point is this: Let $\mathcal{F}$ be a sheaf on $\mathbb{A}^1 \times \mathbb{A}^1$ that is lisse away from a closed set Z . Suppose that Z remains closed when viewed as a subset of $\mathbb{A}^1 \times \mathbb{P}^1$. Then $Ru_*\mathcal{F}$ is locally constant on $\mathbb{A}^1 \times \infty$ in $\mathbb{A}^1 \times \mathbb{P}^1$.

This is an étale-local question, so we may work over the étale-local ring R of some point (d, ∞) of $\mathbb{A}^1 \times \infty$. When pulled back to $\text{Spec} R$, $\mathcal{F}$ is lisse away from the inverse image of the line $\mathbb{A}^1 \times \infty$.

Let x be the coordinate of the second $\mathbb{A}^1$. Let $t = x^{-1}$. Then t is an element of R and the vanishing set of t is exactly the inverse image of $\mathbb{A}^1 \times \infty$. So $\mathcal{F}$ is a lisse sheaf on $\text{Spec } R[t^{-1}]$ for R a strict Henselian local ring, and corresponds to a representation of the fundamental group of that ring. By Abyankhar's lemma [Grothendieck and Raynaud, 1960, Exposé XIII, Proposition 5.2], since all primes are invertible in R , that fundamental group is $\hat{\mathbb{Z}}(1)$. There is a map $\text{Spec } R[t^{-1}] \rightarrow \mathbb{G}_m$ given by the coordinate t^{-1} . The fundamental group of $\mathbb{G}_m$ is also $\hat{\mathbb{Z}}(1)$, and we can see that this map induces an isomorphism of fundamental groups because in each case the cover corresponding to the unique index n subgroup is the cover defined by adjoining the n th root of the coordinate, so the map induces an isomorphism mod n for each n .

Hence locally, $\mathcal{F}$ is isomorphic to a sheaf defined by pulling back a lisse sheaf on $\mathbb{G}_m$ along the coordinate t^{-1} . Recall that t^{-1} is just x . Equivalently, $\mathcal{F}$ is locally isomorphic to $pr_2^*\mathcal{G}$ for some sheaf $\mathcal{G}$ on $\mathbb{A}^1$, lisse on $\mathbb{G}_m$. So it is sufficient to prove

that $Ru_*pr_2^*\mathcal{G}$ is locally constant on $\mathbb{A}^1 \times \infty$. But by smooth base change,

$$Ru_*pr_2^*\mathcal{G} = pr_2^*Ru_*\mathcal{G},$$

which is constant on the fibers of pr_2 , including $\mathbb{A}^1 \times \infty$.

Having verified the key point, we can see that $Ru_*(d^*A \otimes pr_2^*B)$ is locally constant on $\mathbb{A}^1 \times \infty$, so V is locally constant on the divisor at ∞ , thus $C = R\pi_*V$ is locally constant. But locally constant sheaves on $\mathbb{A}_{\mathbb{Q}}^1$ are geometrically constant, because the geometric fundamental group of $\mathbb{A}_{\mathbb{Q}}^1$ is trivial, so C is geometrically constant. $\square$

When working modulo $D^b(\mathcal{N}eg)$ in the future, we will use whichever of $*$ or $*_!$ is most convenient.

Lemma 3.0.15. *Let A, B be perverse sheaves on $\mathbb{A}_{\mathbb{Q}}^1$. The convolution $A *_! B \cong A * B$ lies in $\mathcal{P}erv/\mathcal{N}eg \subset D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_{\ell})/D^b(\mathcal{N}eg)$.*

Proof. Recall that $D^b(\mathcal{P}erv) = D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_{\ell})$ [Beilinson, 1987, Main Theorem 1.3]. Hence we can write any element of $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_{\ell})$ as a complex in the category of perverse sheaves. By definition ${}^p\mathcal{H}^i$ refers to the i th cohomology object of this complex.

$A * B$ is semiperverse [Katz, 1990, 8.1.9] and thus ${}^p\mathcal{H}^i(A * B) = 0$ for $i > 0$. Let D represent Verdier duality, then by [Katz, 1990, 8.1.9]:

$$D(A *_! B) = DA * DB$$

Because DA and DB are perverse, $DA * DB$ is semiperverse. Because the dual of $A *_! B$ is semiperverse, ${}^p\mathcal{H}^i(A *_! B) = 0$ for $i < 0$. There is a map ${}^p\mathcal{H}^i(A *_! B) \rightarrow {}^p\mathcal{H}^i(A * B)$ in each degree i . In each degree other than 0, either the source or target is 0, and since the mapping cone is geometrically constant, the kernel or cokernel is geometrically constant, so the source and target are geometrically constant. Hence $A *_! B$ and $A * B$ are each geometrically constant in each degree other than 0, and in

degree 0 they are isomorphic up to geometrically constant sheaves. So:

$$A *_! B \cong {}^p\mathcal{H}^0(A *_! B) \cong {}^p\mathcal{H}^0(A * B) \cong A * B$$

modulo $D^b(\mathcal{N}eg)$. □

Lemma 3.0.16. *Let A, B be elements of $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_\ell)$. If at least one of A or B is in $D^b(\mathcal{N}eg)$, then $A * B \in D^b(\mathcal{N}eg)$.*

Proof. Without loss of generality, $A \in D^b(\mathcal{N}eg)$. A sheaf is in $D^b(\mathcal{N}eg)$ if and only if, after base changing to $\overline{\mathbb{Q}}$, it is a complex of constant sheaves. So we will prove over $\overline{\mathbb{Q}}$ that if A is a complex of constant sheaves, then $A *_! B$ is a complex of constant sheaves. Since A is a sum of copies of the constant sheaf $\overline{\mathbb{Q}}_\ell$ placed in different degrees, it is sufficient to deal with the case $A = \overline{\mathbb{Q}}_\ell$.

Recall that:

$$\overline{\mathbb{Q}}_\ell * B = \overline{\mathbb{Q}}_\ell *_! B = Rsum_!(\overline{\mathbb{Q}}_\ell \boxtimes B) = Rpr_{1!}(d^*\overline{\mathbb{Q}}_\ell \otimes pr_2^*B)$$

But $d^*\overline{\mathbb{Q}}_\ell = \overline{\mathbb{Q}}_\ell$, and $\overline{\mathbb{Q}}_\ell \otimes pr_2^*B = pr_2^*B$, so we have:

$$\overline{\mathbb{Q}}_\ell * B = Rpr_{1!}pr_2^*B$$

This is the base change of the compactly supported cohomology of B along the morphism $\pi : \mathbb{A}_{\mathbb{Q}}^1 \rightarrow \text{Spec } \overline{\mathbb{Q}}$ by the identical morphism $\pi' : \mathbb{A}_{\mathbb{Q}}^1 \rightarrow \text{Spec } \overline{\mathbb{Q}}$. By the proper base change theorem, the cohomology of the base change is just the pullback of the cohomology, hence is geometrically constant. □

Corollary 3.0.17. *The two sheaf convolution operations $*$ and $*_!$ are well-defined functors $\mathcal{P}erv/\mathcal{N}eg \times \mathcal{P}erv/\mathcal{N}eg \rightarrow \mathcal{P}erv/\mathcal{N}eg$ and these two are equivalent.*

Proof. By Lemma 3.0.15, both are functors $\mathcal{P}erv \times \mathcal{P}erv \rightarrow \mathcal{P}erv/\mathcal{N}eg$. By Lemma 3.0.16, they send elements of $\mathcal{N}eg$ to the 0 object. Because they preserve distinguished triangles, they are exact, and hence factor through $\mathcal{P}erv/\mathcal{N}eg$, giving functors $\mathcal{P}erv/\mathcal{N}eg \times \mathcal{P}erv/\mathcal{N}eg \rightarrow \mathcal{P}erv/\mathcal{N}eg$. These are equivalent by Lemma 3.0.14. $\square$

Proposition 3.0.18. *$(\mathcal{P}erv/\mathcal{N}eg, *)$ is a symmetric monoidal abelian category.*

Proof. It is not too hard to write down the unit, symmetry, and associativity isomorphisms using the Künneth formula and the Leray spectral sequence. To prove that the category is symmetrical monoidal, there are also several diagrams involving these isomorphisms whose commutativity must be checked. These commutative diagrams can be deduced from the definitions of the isomorphisms, but the argument to do so involve repeatedly manipulating complicated expressions. We give instead a conceptual argument. However, this requires some unusual concepts.

First, note that it is sufficient to find a symmetric monoidal structure on the derived category where $*$ is the tensor product, because if the isomorphisms exist and diagrams commute in the derived category then the same isomorphisms exist and diagrams commute in $\mathcal{P}erv/\mathcal{N}eg$ because it is a quotient category of a full subcategory of $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_{\ell})$.

Second, we will use a slightly unusual definition of symmetric monoidal category. This is the balanced definition in which we define, not just $A \otimes B$, but $\otimes_I(A_i)_{i \in I}$ for all finite sets I . A definition, including all the diagrams that must be shown to commute, is given in [Deligne and Milne, 1982, Proposition 1.5] and later they note that it is equivalent to the usual one. (We will mention later where we take advantage of this definition in our argument.)

To define the convolution of an arbitrary number of sheaves first note that when we take the exterior product of a finite set of sheaves $\boxtimes_{i \in I} A_i$ by applying $\boxtimes$ pairwise in any order, it does not depend on the order we apply $\boxtimes$ because $\boxtimes$ is defined using tensor product and the tensor product defines a symmetric monoidal structure on

the category of sheaves. There is a Künneth formula for $\boxtimes_{i \in I} A_i$, which is exactly like the usual Künneth formula but involves tensoring an arbitrary finite set of sheaves, and can be proved by iteratively applying the usual Künneth formula. The fact that no matter which order we apply it in we get the same isomorphism follows from the definition, but even ignoring that our argument will work equally well no matter what order the Künneth formula is applied in.

Now we are ready to define the structure. For a finite set I , define π_I to be the map $(\mathbb{A}^1)^I \rightarrow \mathbb{A}^1$ given by the equation $(x_i)_{i \in I} \mapsto \sum_{i \in I} x_i$. Define $*_I : D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_\ell)^I \rightarrow D_c^b(\mathbb{A}_{\mathbb{Q}}^1)$ by $(A_i)_{i \in I} \mapsto R\pi_{I*} \boxtimes_{i \in I} A_i$.

Then for any $\alpha : I \rightarrow J$, we have the isomorphism $\chi(\alpha) : *_I(A_i)_{i \in I} \rightarrow *_J(*_{\alpha^{-1}(j)}(A_i)_{i \in \alpha^{-1}(j)})_{j \in J}$ defined by composing three isomorphisms:

$$R\pi_{J*} \boxtimes_{j \in J} (R\pi_{\alpha^{-1}(j)*} \boxtimes_{i \in \alpha^{-1}(j)} A_i)$$

(by the Künneth formula)

$$= R\pi_{J*} R \left(\prod_{j \in J} \pi_{\alpha^{-1}(j)} \right)_* \boxtimes_{j \in J} (\boxtimes_{i \in \alpha^{-1}(j)} A_i)$$

(by the Leray spectral sequence and the identity $\pi_J \circ \prod_{j \in J} \pi_{\alpha^{-1}(j)} = \pi_I$)

$$= R\pi_{I*} \boxtimes_{j \in J} (\boxtimes_{i \in \alpha^{-1}(j)} A_i)$$

(by the symmetric monoidality of $\boxtimes$)

$$= R\pi_{I*} \boxtimes_{i \in I} A_i$$

It remains to check two the conditions (a) and (b) Deligne and Milne. Condition (a) is that when I is a single element then $*_I$ is the identity functor and that when α

is the map from the one-element map to itself then $\chi(\alpha)$ is the identity. The first part holds because here $(\mathbb{A}^1)^I$ is $\mathbb{A}^1$, π_I is the identity morphism, and $\boxtimes$ is the identity functor, so $*_I$ is the identity functor. The second part holds because when there is only one factor in the tensor product and all morphisms we are pushing forward along are identities, the Künneth, Leray, and symmetric monoidality isomorphisms are the identity, so the composition of them is the identity.

Condition (b) is a commutative square. This is more difficult and we prove it by a series of reductions. First, note that to prove a commutative square involving derived functors it is sufficient to prove the commutative square on the original functors, because the derived functor is computed from the original functor using an injective resolution. However, this reduction requires us to pass from the derived category of constructible sheaves to the category of arbitrary sheaves as the injective resolution may involve non-constructible sheaves. The Leray spectral sequence and the symmetric monoidality of $\boxtimes$ are still identities for constructible sheaves, but the Künneth formula becomes only a morphism - for the product of a pair, this morphism is $f_{1*}A_1 \boxtimes f_{2*}A_2 \rightarrow (f_1 \times f_2)_*(A_1 \boxtimes A_2)$ and for the product of an arbitrary set it is $\boxtimes_{j \in J} (f_{j*}A_j) \rightarrow \left(\prod_{j \in J} f_j \right)_* \boxtimes_{j \in J} A_j$. So we only obtain a morphism

$$\pi_{J*} \boxtimes_{j \in J} (\pi_{\alpha^{-1}(j)*} \boxtimes_{i \in \alpha^{-1}(j)} A_i) \rightarrow \pi_{I*} \boxtimes_{i \in I} A_i$$

and not an isomorphism. However, to check that a commutative square of isomorphisms commutes it is sufficient to check that the two ways of getting from any one corner to any other corner agree. Because our arrows point the opposite way as the arrows in Deligne and Milne's diagram, the two ways of composing morphisms to get from the bottom-right corner to the top-left corner still exist for arbitrary sheaves, so to show that the square of isomorphism commutes it is sufficient to check that

those two compositions are equal, for which we can reduce to the case of non-derived functors.

Next we handle the case of skyscraper sheaves (and then we will deduce the general sheaf case from the skyscraper sheaf case). For skyscraper sheaves δ_{x_i} , their exterior tensor product $\boxtimes_{i \in I} \delta_{x_i}$ is the skyscraper sheaf $\delta_{(x_i)_{i \in I}}$, and the pushforward of that $\pi_{I*} \delta_{(x_i)_{i \in I}}$ is simply the skyscraper sheaf $\delta_{\sum_i x_i}$. So the convolution of any number of skyscraper sheaves is a skyscraper sheaf. Moreover all the intermediate steps in the morphisms $\chi(\alpha)$ are equal, by repeating those two isomorphisms, to the same skyscraper sheaf. So to check that this square commutes it is sufficient to check that each of the three morphisms that make up $\chi(\alpha)$ is simply the identity map from this skyscraper sheaf to itself. This is just saying that the Künneth isomorphism, the Leray spectral sequence, and the symmetric monoidality of $\boxtimes$, when applied in a situation where the input sheaves are skyscraper sheaves with fixed generators, preserve the canonical generators of the output skyscraper sheaves. This is obvious in each case from the definition - we never multiply any section by a number in any of these definitions, so the section giving the canonical generator is never sent to a different section.

Now let $(A_i)_{i \in I}$ be a tuple of sheaves. For each tuple of skyscraper sheaves δ_{x_i} and maps $f_i : A_i \rightarrow \delta_{x_i}$, we know the diagram commutes for δ_{x_i} , so the diagram commutes modulo the kernel of the map induced by functoriality $\pi_{I*} \boxtimes_{i \in I} A_i \rightarrow \pi_{I*} \boxtimes_{i \in I} \delta_{x_i}$. Thus to show the diagram commutes it is sufficient to show that the intersection of the kernels of all these maps for all possible tuples is empty. Take any nonzero section of $\pi_{I*} \boxtimes_{i \in I} A_i$ over an open set. It comes from a section of $\boxtimes_{i \in I} A_i$ on the pullback of that open set to $(\mathbb{A}^1)^I$. That nonzero section must have a nonzero stalk at some point $(x_i)_{i \in I}$, which is an element of the stalk $\otimes_i (A_i)_{x_i}$ of $\boxtimes_{i \in I} A_i$ at $(x_i)_{i \in I}$. Hence there exists a tuple of linear maps $(A_i)_{x_i} \rightarrow \overline{\mathbb{Q}}_\ell$ (or $\mathbb{F}_\ell$ or $\mathbb{Z}/\ell^n$ or any other coefficient ring) that are nonzero on this nonzero element, giving a tuple of maps $A_i \rightarrow \delta_{x_i}$. Then this

section is sent to a nonzero section by the natural map $\boxtimes_{i \in I} A_I \rightarrow \boxtimes_{i \in I} \delta_{x_i}$, so so the corresponding section of $\pi_{I*} \boxtimes_{I \in I} A_i$ is sent to a nonzero section of $\pi_{I*} \boxtimes_{i \in I} \delta_{x_i}$ and we win. $\square$

It will sometimes be helpful to have a single concrete representative of each equivalence class of $\mathcal{Perv}/\mathcal{Neg}$. For this we turn to a special case of a definition of Katz:

Definition 3.0.19. [Katz, 1995, 2.6.1] For A perverse sheaves on $\mathbb{A}_{\mathbb{Q}}^1$, we say that A has property $\mathcal{P}$ if, for all $B \in \mathcal{Perv}$, both $A * B$ and $A *_! B$ are perverse.

Lemma 3.0.20. *For A a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$, the following are equivalent:*

- 1) A has property $\mathcal{P}$.
- 2) A has no subobject or quotient that is a geometrically constant sheaf.
- 3) A , base changed to any algebraically closed field, has no subobject or quotient that is a constant sheaf.
- 4) $H_c^1(\mathbb{A}_{\mathbb{Q}}^1, A)$ and $H^{-1}(\mathbb{A}_{\mathbb{Q}}^1, A) = 0$
- 5) $H_c^i(\mathbb{A}_{\mathbb{Q}}^1, A)$ and $H^i(\mathbb{A}_{\mathbb{Q}}^1, A) = 0$ for all $i \neq 0$.

Proof. 4) is equivalent to 5) because, for a perverse sheaf on an affine curve, the only possible nonvanishing cohomology groups are H^{-1} and H^0 , and the only possible nonvanishing compactly supported cohomology groups are H^0 and H^1 .

3) is equivalent to 4) because:

$$\mathrm{Hom}_{\mathbb{A}_{\mathbb{Q}}^1}(\overline{\mathbb{Q}}_{\ell}[1], A) = H^{-1}(\mathbb{A}_{\mathbb{Q}}^1, A)$$

$$\mathrm{Hom}_{\mathbb{A}_{\mathbb{Q}}^1}(A, \overline{\mathbb{Q}}_{\ell}[1]) = H_c^1(\mathbb{A}_{\mathbb{Q}}^1, A)^{\vee}$$

The equivalence of 1) and 3) follows from [Katz, 1995, Lemmas 2.6.13, 2.6.14 and 2.6.15].

Finally, 2) is equivalent to 3) because if A has a subobject over $\mathbb{A}_{\mathbb{Q}}^1$ that is constant, then we can write $\mathrm{Hom}_{\mathbb{A}_{\mathbb{Q}}^1}(\overline{\mathbb{Q}}_\ell[1], A)$ as a nontrivial Galois representation V , and $\overline{\mathbb{Q}}_\ell[1] \otimes V$ will be a geometrically constant subobject of A . The same argument, dually, applies to quotient objects. $\square$

Lemma 3.0.21. *1. The functor from the full subcategory of $\mathcal{P}erv$ consisting of elements with property $\mathcal{P}$, to $\mathcal{P}erv/\mathcal{N}eg$, is an equivalence of categories. In particular, every equivalence class in $\mathcal{P}erv/\mathcal{N}eg$ has a unique-up-to-unique isomorphism representative in $\mathcal{P}erv$ with property $\mathcal{P}$.*

Call this the property $\mathcal{P}$ representative of this equivalence class.

2. If A is an element of $\mathcal{P}erv$, then there is a negligible subobject N of A and a negligible quotient object M of A/N such that the kernel of the quotient map from A/N to M is the property $\mathcal{P}$ representative of the class of A .

Proof. Let A be an object of $\mathcal{P}erv$. Let N be a maximal negligible subobject of A . Then A/N has no negligible subobjects as if it had a negligible subobject N' , the kernel of the composition $A \rightarrow A/N \rightarrow (A/N)/N'$ would be an extension of N by N' , hence negligible as $\mathcal{N}eg$ is a Serre subcategory. Let M be the maximal negligible quotient of A/N . By an identical argument, the kernel of the natural map from A/N to M has no negligible quotient objects. It also has no negligible subobjects because any subobject of it is already a subobject of A/N . Hence it has property $\mathcal{P}$, and by construction it is equivalent to A in $\mathcal{P}erv/\mathcal{N}eg$. (These maximal subobjects and quotient objects necessarily exist because $\mathcal{P}erv$ is an Artinian category [Beilinson et al., 1982, Theorem 4.3.1(i)], so there are no infinite ascending chains of subobjects and quotient objects.)

This shows that the functor from the full subcategory of $\mathcal{P}erv$ with property $\mathcal{P}$ to $\mathcal{P}erv/\mathcal{N}eg$ is essentially surjective, and that a property $\mathcal{P}$ representative of each equivalence class is constructed as stated. It remains to check that this functor is

fully faithful. This implies that there is a unique property $\mathcal{P}$ representative of each equivalence class, up to isomorphism, and hence the unique property $\mathcal{P}$ representative is produced by a suitable sequence. The formula for homomorphisms in a quotient by a Serre subcategory is [Gabriel, 1962, pp. 365]:

$$\mathrm{Hom}_{\mathcal{P}erv/\mathcal{N}eg}(X, Y) = \varinjlim \mathrm{Hom}_{\mathcal{P}erv}(X', Y/N)$$

where the limit is taken over subobjects $X' \subseteq X, N \subseteq Y$ where X/X' and N are in $\mathcal{N}eg$. Hence if neither X nor Y has any quotients or subobjects in $\mathcal{N}eg$, then homomorphisms between the objects in the quotient category are just the usual homomorphisms, and thus the map is faithful and full. $\square$

Proposition 3.0.22. *Let $[-1] : \mathbb{A}^1 \rightarrow \mathbb{A}^1$ be the map that sends x to $-x$. Let D be Verdier duality.*

$(\mathcal{P}erv/\mathcal{N}eg, *)$ is a rigid category, and for a perverse sheaf A , the dual object is given by the formula:

$$A^\vee = [-1]^* DA$$

Proof. A category is rigid if every object has a dual. If A is perverse then $[-1]^* DA$ is perverse, so it is sufficient to prove that $[-1]^* DA$ is actually a dual object.

Let $A^\vee = [-1]^* DA$. By a criterion of Deligne [1990, 2.2], to show that $A^\vee$ is the dual of A , it is sufficient to show that $\otimes A$ and $\otimes A^\vee$ are adjoint functors.

This will follow from two simple identities:

$$\mathrm{Hom}(A^\vee, B) = \mathrm{Hom}(1, A * B)$$

$$(A * B)^\vee = A^\vee * B^\vee$$

To prove the first identity, note that the unit object in our monoidal category is just the skyscraper sheaf $\delta(0)$, so we compute:

$$\begin{aligned}\mathrm{Hom}(\delta(0), A * B) &= \mathrm{Hom}(\delta(0), Rsum_*(A \boxtimes B)) \\ &= \mathrm{Hom}(sum^*\delta(0), A \boxtimes B)\end{aligned}$$

Let i be the inclusion of the anti-diagonal $x + y = 0$ into $\mathbb{A}_R^2$. Then we continue:

$$\begin{aligned}&= \mathrm{Hom}(i_!\overline{\mathbb{Q}}_\ell, A \boxtimes B) \\ &= \mathrm{Hom}(\overline{\mathbb{Q}}_\ell, Ri^!(A \boxtimes B))\end{aligned}$$

Let Δ be the diagonal map, then:

$$= \mathrm{Hom}(\overline{\mathbb{Q}}_\ell, R\Delta^!([-1]^*A) \boxtimes B)$$

and by [Grothendieck, 1977, Exp. III (3.1.1) and (3.2.1)]:

$$= \mathrm{Hom}([-1]^*DA, B)$$

This gives the first identity in $D^b(\mathcal{P}erv)$. To get the same identity in $\mathcal{P}erv/\mathcal{N}eg$, we may assume that A and B have property $\mathcal{P}$. In this case by Lemma 3.0.20(2), $[-1]^*DA$ has property $\mathcal{P}$, so by Lemma 3.0.21:

$$\mathrm{Hom}_{\mathcal{P}erv/\mathcal{N}eg}([-1]^*DA, B) = \mathrm{Hom}_{\mathcal{P}erv}([-1]^*DA, B)$$

To obtain

$$\mathrm{Hom}_{\mathcal{P}erv/\mathcal{N}eg}(\delta(0), A * B) = \mathrm{Hom}_{\mathcal{P}erv}(\delta(0), A * B)$$

we again use the formula for homomorphisms in a Serre subcategory [Gabriel, 1962, pp. 365]:

$$\mathrm{Hom}_{\mathcal{P}_{erv}/\mathcal{N}eg}(X, Y) = \lim_{\rightarrow} \mathrm{Hom}_{\mathcal{P}_{erv}}(X', Y/N)$$

where the limit is taken over subobjects $X' \subseteq X, N \subseteq Y$ where X/X' and N are in $\mathcal{N}eg$. So it is sufficient to check that $A * B$ has no subobjects in $\mathcal{N}eg$ and $\delta(0)$ has no quotient objects in $\mathcal{N}eg$. Because A and B have property $\mathcal{P}$, $H^i(\mathbb{A}^1, A)$ and $H^i(\mathbb{A}^1, B)$ vanish for $i < 0$, so $H^i(\mathbb{A}^1, A * B)$ vanishes for $i < 0$ by the Kunneth formula, thus $A * B$ has no geometrically constant subobjects. On the other hand $\delta(0)$ has property $\mathcal{P}$ and thus has no quotient objects in $\mathcal{N}eg$, so we are done, verifying the first identity.

To check the second identity, observe that $[-1]^*$ commutes with convolution because $[-1]$ is an automorphism of the group $\mathbb{G}_a$. Duality does not usually commute with convolution, as

$$D(A * B) = DA *_! DB$$

[Katz, 1990, 8.1.9] but we know that usual and compactly supported convolution are the same modulo $D^b(\mathcal{N}eg)$, so this verifies the second identity. Then we have:

$$\mathrm{Hom}(B * A, C) = \mathrm{Hom}(1, (B * A)^\vee, C) = \mathrm{Hom}(1, B^\vee * A^\vee \otimes C) = \mathrm{Hom}(B, A^\vee * C)$$

so $*A^\vee$ is the right adjoint to $*A$. Then by [Deligne, 1990, 2.2] we are done. (This argument was suggested by Johnson-Freyd [2015] on Mathoverflow.) $\square$

Let $p \neq \ell$ be a prime. Let $R\Psi_p$ be the nearby cycles functor [Deligne, 1977, Th. Finitude, 3.1] relative to $\mathbb{Z}_p$ that takes objects of $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_\ell)$ to objects of $D_c^b(\mathbb{A}_{\mathbb{F}_p}^1, \overline{\mathbb{Q}}_\ell)$ with an action of $\mathrm{Gal}(\overline{\mathbb{Q}}_p | \mathbb{Q}_p)$ and is defined as:

Definition 3.0.23. Let $\mathbb{Z}_p^{ur}$ be the integral closure of $\mathbb{Z}_p$ in the maximal unramified extension of $\mathbb{Q}_p$, and let $\overline{\mathbb{Q}}_p$ and $\overline{\mathbb{F}}_p$ be the algebraic closures of $\mathbb{Q}_p$ and $\mathbb{F}_p$ respectively. For a scheme X over $\mathbb{Z}_p$ (e.g. $\mathbb{A}^1$), let i be the map $X_{\overline{\mathbb{F}}_p} \rightarrow X_{\mathbb{Z}_p^{ur}}$, let $\bar{j}$ be the map

$X_{\overline{\mathbb{Q}}_p} \rightarrow X\mathbb{Z}_p^{ur}$, and let $\bar{k}$ be the map $X_{\overline{\mathbb{Q}}_p} \rightarrow X_{\mathbb{Q}_p}$ induced by the natural maps $\mathbb{Z}_p^{ur} \rightarrow \overline{\mathbb{F}}_p$, $\mathbb{Z}_p^{ur} \rightarrow \overline{\mathbb{Q}}_p$, $\mathbb{Z}_p \rightarrow \overline{\mathbb{Q}}_p$. Then for a complex of sheaves A on $X_{\mathbb{Q}_p}$, we define:

$$R\Psi_p A = i^* R\bar{j}_* k^* A$$

with the natural action of $\text{Gal}(\overline{\mathbb{Q}}_p|\mathbb{Q}_p)$ coming from k^* and preserved by functoriality of i^* and $R\bar{j}_*$. Later, we will also use vanishing cycles. Let A be a sheaf on $\mathbb{A}_{\mathbb{Z}_p}^1$, and let $o : \mathbb{A}_{\mathbb{Z}_p^{ur} \rightarrow \mathbb{Z}_p}^1$ be the natural map, then $\bar{k}^* A_{\mathbb{Q}_p} = \bar{j}^* o^* A$. By adjunction there is a natural map:

$$i^* A \rightarrow i^* R\bar{j}_* j^* A = R\Psi_p A_{\mathbb{Q}_p}$$

and $R\Phi_p A$ is defined as the mapping cone of this map.

If A is perverse on $\mathbb{A}_{\mathbb{Q}_p}^1$, then the nearby cycles sheaf $R\Psi_p A$ is perverse on $\mathbb{A}_{\overline{\mathbb{F}}_p}^1$ [Illusie, 1994, Corollary 4.5].

Definition 3.0.24. Let F be a field in which ℓ is invertible, and let A be a perverse sheaf on $\mathbb{A}_F^1$. We say that the generic rank of A , or $\text{rank}(A)$, is the dimension of the stalk of $\mathcal{H}^{-1}(A)$ at the generic point of $\mathbb{A}_F^1$. (There is always a nonempty open set over which A is lisse of this rank.)

The definition of good reduction is as follows:

Definition 3.0.25. A perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ has good reduction at p if it satisfies the following three conditions:

1.

$$\text{rank}(R\Psi_p A) = \text{rank}(A)$$

2. The singularities of A are p -adic integers. That is, when viewed as elements of $\overline{\mathbb{Q}}_p$, they in fact lie in $\overline{\mathbb{Z}}_p$.

3. The inertia group in $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$ acts trivially on $R\Psi_p A$.

This definition is motivated as follows: These conditions, we will later see (Lemma 3.0.36), ensure that

$$R\Psi_p(A * B) = R\Psi_p A * R\Psi_p B \quad (3.2)$$

as long as A and B have good reduction.

The necessity of the first condition for the identity (3.2) is seen by taking A to be the skyscraper sheaf supported at $1/p$ and B the skyscraper sheaf supported at $-1/p$. Then $R\Psi_p A$ and $R\Psi_p B$ both vanish, but $A * B$ is the skyscraper sheaf supported at 0, and thus $R\Psi_p(A * B)$ is also a skyscraper sheaf and is nontrivial.

The necessity of the second condition for the identity (3.2) is seen by taking A to be the Kummer sheaf $\mathcal{L}_\chi$ for χ a nontrivial character of μ_2 , and B to be $\mathcal{L}_{\chi^{-1}}$. Then $R\Psi_2 A$ and $R\Psi_2 B$ again both vanish, but $A * B$ is equivalent mod $\mathcal{N}eg$ to a Tate twist of a skyscraper sheaf at 0, hence $R\Psi_2(A * B)$ is again nontrivial. (This phenomenon does not occur only at 2. Similar but more complicated examples can be constructed for other primes using the pushforward of the constant sheaf along the p th power map.)

The third condition may not be necessary to ensure $R\Psi_p(A * B) = R\Psi_p A * R\Psi_p B$, but we will want a well-defined action of Frob_p on $R\Psi_p A$, for which we need the inertia group to act trivially.

The main goal of the remainder of this chapter will be to establish the identity (3.2) and show that good reduction is stable under convolution. Then it will not be too difficult to construct the Tannakian category.

We will need to control the behavior at ∞ of sheaves with good reduction at p . For that reason we make the following definitions:

Definition 3.0.26. Let S be any base. A reduced closed subscheme of $\mathbb{A}_S^1$ is bounded if it remains closed when viewed as a closed subset of $\mathbb{P}_S^1$.

Definition 3.0.27. Let S be a scheme over $\mathbb{Z}_p$. An element of $D_c^b(\mathbb{A}_{S_{\mathbb{Q}_p}}^1, \overline{\mathbb{Q}_\ell})$ is extensible if there is an open set U of $\mathbb{A}_S^1$ and a complex of lisse sheaves $L \in D_c^b(U, \overline{\mathbb{Q}_\ell})$, such that :

- The complement in $\mathbb{A}_S^1$ of U is bounded
- Let $u : U_{\mathbb{Q}_p} \rightarrow \mathbb{A}_{S_{\mathbb{Q}_p}}^1$ and $v : U_{\mathbb{Q}_p} \rightarrow U$ be the natural open immersions. Then:

$$u^* A \cong v^* L$$

If A is extensible, then we may control the nearby cycles of A by comparing them with the lisse sheaf L . In the next three lemmas we will show that sheaves with good reduction are extensible. This will require some notation and concepts that are not used in the remainder of the paper.

Let η_0 be the geometric generic point of $\mathbb{A}_{\mathbb{Q}}^1$. Let η_p be the geometric generic point of $\mathbb{A}_{\mathbb{F}_p}^1$. We can also view them as geometric points of $\mathbb{A}_{\mathbb{Z}_p}^1$.

Let I_p be the inertia group of $\mathbb{Q}_p$ at p . Let I_{η_p} be the inertia group of $\mathbb{Q}_p(x)$ for the unique valuation which makes elements of $\mathbb{Z}_p[x]$ that are nonzero mod p have order 0 and makes p have order 1. This valuation is what Dwork called the Gauss norm.

There is a natural map $I_{\eta_p} \rightarrow I_p$ because $\mathbb{Q}_p(x)$ is an extension of $\mathbb{Q}_p$, and this valuation, restricted to $\mathbb{Q}_p$, is the standard valuation on $\mathbb{Q}_p$. This map is surjective as both valuations are integer-valued, so any ramified extension of $\mathbb{Q}_p$ remains ramified as an extension of $\mathbb{Q}_p(x)$. Let P be the kernel of this map. Then $I_p = I_{\eta_p}/P$.

Lemma 3.0.28. *Let A be a complex in $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}_\ell})$. Then we have an isomorphism of complexes of I_p -representations*

$$(R\Psi_p A)_{\eta_p} = (A_{\eta_0})^P$$

Here I_{η_p} acts on A_{η_0} because I_{η_p} is a subgroup of the absolute Galois group of $\mathbb{Q}(x)$. Thus P , a subgroup of I_{η_p} , acts on A_{η_0} , and I_p , the quotient by that subgroup, acts on $(I_{\eta_0})^P$.

Proof. To compute $(R\Psi_p A)_{\eta_p}$, when taking vanishing cycles in $\mathbb{A}_{\mathbb{Z}_p}^1$, we may work in any étale neighborhood of η_p . So we may work in the limit of all étale neighborhoods of η_p . Call this S' . It is the spectrum of the strict Henselian local ring of $\mathbb{A}_{\mathbb{Z}_p}^1$ at η_p . That ring is a discrete valuation ring. The geometric generic point of S' is η_0 and its geometric special point is η_p .

We want to compute the nearby cycles of A on S' relative to the base $\text{Spec } \mathbb{Z}_p$. The nearby cycles of A on S' relative to the base S' are much easier, because S' is zero-dimensional over S' , and so

$$(R\Psi_{p,S'} A)_{\eta_p} = A_{\eta_0}$$

By [Deligne, 1977, Th. finitude, Lemma 3.4], the vanishing cycles relative to $\mathbb{Z}_p$ are just the P -invariants of the vanishing cycles relative to S' , and the isomorphism, being functorial, is I_p -equivariant, so we obtain the stated formula.

□

Corollary 3.0.29. *Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$.*

$$\text{rank}(R\Psi_p A) \leq \text{rank}(A)$$

with equality only if P acts trivially on A_{η_0} .

Proof. By Lemma 3.0.28

$$\text{rank}(R\Psi_p A) = \dim(\mathcal{H}^{-1}(R\Psi_p A))_{\eta_p} = \dim(\mathcal{H}^{-1}(A)_{\eta_0})^P \leq \dim \mathcal{H}^{-1}(A)_{\eta_0} = \text{rank}(A)$$

The middle inequality is an equality if and only if P acts trivially on $\mathcal{H}^{-1}(A)_{\eta_0}$. Because A is perverse, A_{η_0} is supported in degree -1 , so P acts trivially on $H^{-1}(A_{\eta_0})$ if and only if it acts trivially on A_{η_0} . $\square$

Lemma 3.0.30. *Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$. If A has good reduction at p then A is extensible when pulled back to $\mathbb{A}_{\mathbb{Q}_p}^1$.*

Moreover, we may take U to be the complement of the closure in $\mathbb{A}_{\mathbb{Z}_p}^1$ of the set of singularities of A in $\mathbb{A}_{\mathbb{Q}_p}^1$.

Proof. Because $\text{rank}(A) = \text{rank}(R\Psi_p A)$, Corollary 3.0.29 shows that P acts trivially on A_{η_0} . Because I_p acts trivially on $R\Psi_p A$, it acts trivially on $(R\Psi_p A)_{\eta_p}$, and so I_{η_p} acts trivially on A_{η_0} .

A is lisse away from some finite set of singularities. These singularities lie in $\overline{\mathbb{Z}_p}$, so their closure in $\mathbb{A}_{\mathbb{Z}_p}^1$ is a bounded closed set. Let U be its complement. We will construct a lisse sheaf on U which is equal to A when pulled back to $U_{\mathbb{Q}_p}$.

A pulled back to $U_{\mathbb{Q}_p}$ is a lisse sheaf in degree -1 , hence a representation ρ of $\pi_1(U_{\mathbb{Q}_p})$. The pullback of ρ to $\text{Gal}(\overline{\mathbb{Q}_p(x)}|\mathbb{Q}_p(x))$ is just $\mathcal{H}^{-1}(A)_{\eta_0}$, thus I_{η_p} acts trivially on this representation under the composition:

$$I_{\eta_p} \rightarrow \text{Gal}(\overline{\mathbb{Q}_p(x)}|\mathbb{Q}_p(x)) \rightarrow \pi_1(U_{\mathbb{Q}_p})$$

We will show that the quotient of $\pi_1(U_{\mathbb{Q}_p})$ by the minimal normal subgroup containing I_{η_p} is $\pi_1(U)$. Then the representation ρ will factor through $\pi_1(U)$, giving the desired lisse sheaf on U , and we will be done.

To see this, we use the Galois correspondence. Consider a finite étale Galois cover of $U_{\mathbb{Q}_p}$ which, when extended by normalization to U , is unramified at the generic point. By purity of ramification, it extends to a finite étale cover of U . Hence all quotients of $\pi_1(U_{\mathbb{Q}_p})$ on which I_{η_p} acts trivially factor through $\pi_1(U)$, as desired. $\square$

To show that the conditions of good reduction are preserved by convolution, we will need computations involving the rank and Euler characteristic:

Lemma 3.0.31. *Let K be a separably closed field in which ℓ is invertible, and let A be a perverse sheaf on $\mathbb{A}_K^1$.*

$$\chi(\mathbb{A}_K^1, A) = -\text{rank}(A) + \text{swan}(\mathcal{H}^{-1}(A), \infty) + \sum_{x \in \mathbb{A}^1(K)} (\text{drop}(\mathcal{H}^{-1}(A), x) + \dim(\mathcal{H}^0(A), x) + \text{swan}(\mathcal{H}^{-1}(A), x))$$

where $\text{drop}(\mathcal{H}^{-1}(A), x)$ is the drop in the rank of $\mathcal{H}^{-1}(A)$ at x , $\dim(\mathcal{H}^0(A), x)$ is the rank of the stalk of $\mathcal{H}^0(A)$ at x , and $\text{swan}(\mathcal{H}^{-1}(A), x)$ is the Swan conductor of the local monodromy of $\mathcal{H}^{-1}(A)$ at x .

Proof. This is almost a special case of the Euler-Poincare formula [Grothendieck, 1977, Theorem 7.1], which is stated for a complex of sheaves on a proper curve. A perverse sheaf is a complex of sheaves, and the local terms in this formula are the same as the local terms described in [Grothendieck, 1977], but $\mathbb{A}^1$ is not proper. However we may extend the sheaf by 0 from $\mathbb{A}^1$ to $\mathbb{P}^1$, which does not affect the Euler characteristic. Then the global term for a sheaf on a curve of Euler characteristic 2, such as $\mathbb{P}^1$, is $-2\text{rank}(A)$, and the local contribution at ∞ is $\text{rank}(A) + \text{Swan}(A)$, giving exactly the stated formula.

□

Lemma 3.0.32. *Let K be a separably closed field and let A and B be two perverse sheaves on $\mathbb{A}_K^1$ that are tamely ramified at ∞ and lisse away from the points $x_1, \dots, x_n$ and $y_1, \dots, y_m$ respectively. Then $A *_! B$ is lisse away from the points $x_i + y_j$.*

Proof. We will show this by applying Deligne's semicontinuity theorem [Laumon, 1981, Corollary 2.1.2] to $R\pi_* j_!((d^*A) \otimes (pr_2^*B))$. Let $U = \mathbb{A}^1 - \{x_i + y_j | 1 \leq i \leq n, 1 \leq j \leq m\}$. We will show that on $\pi^{-1}(U)$, the sheaf $j_!((d^*A) \otimes (pr_2^*B))$ is lisse away from a smooth divisor, with constant Swan conductor along each component of that

divisor, and lisse when pulled back to that divisor. It will follow that the cohomology is lisse.

Let D be the divisor where $j_!((d^*A) \otimes (pr_2^*B))$ is not lisse. D consists of the lines $x = x_i$ and $y = y_j$, which after the change of variables become $x' - y' = x_i$ and $y' = y_j$, and the divisor at ∞ with equation $y' = \infty$. Thus D is a union of lines, and hence is smooth away from the points where the lines intersect. In the original coordinates, the points where the lines intersect are (x_i, y_j) . In the sheared coordinates, these intersection points are $(x_i + y_j, y_j)$ and map under π to the points $x_i + y_j$. Hence the intersection of D with $\pi^{-1}(U)$ is smooth.

d^*A is lisse away from the lines of the form $x' - y' = x_i$ and pr_2^*B is lisse away from the lines of the form $y' = y_j$. Let us fix a line L with equation $x' - y' = x_i$. Because the lines don't intersect, we can choose a neighborhood of $L \cap \pi^{-1}(U)$ which does not intersect any of the other lines, so that pr_2^*B will be lisse on that neighborhood. We will check that $j_!((d^*A) \otimes (pr_2^*B))$ has constant Swan conductor around $L \cap \pi^{-1}(U)$, and is lisse when pulled back to $L \cap \pi^{-1}(U)$, working locally in that neighborhood.

It is sufficient to show that d^*A has constant Swan conductor along L and is lisse when pulled back to L , because tensoring with the lisse sheaf pr_2^*B will multiply the Swan conductor by a constant value and preserve lisseness, and because taking $j_!$ will not change anything in a neighborhood of L . But this is clear: d^*A must have the same Swan conductor at L in each fiber of π because d maps each fiber isomorphically to $\mathbb{A}_K^1$, and d^*A must be lisse when pulled back to L because L is the inverse image along d of a point of $\mathbb{A}_K^1$, so d^*A is constant along L .

By the same argument with A and B reversed, we can see that for L a line of the form $y' = y_j$, $j_!((d^*A) \otimes (pr_2^*B))$ has constant Swan conductor around $L \cap \pi^{-1}(U)$, and is lisse when pulled back to $L \cap \pi^{-1}(U)$.

Finally, $j_!((d^*A) \otimes (pr_2^*B))$ is tamely ramified along the line at ∞ because both A and B are tamely ramified at ∞ , and it is zero, hence lisse, on the line at ∞ .

So by Deligne's semicontinuity theorem the cohomology complex $R\pi_* j_!((d^*A) \otimes (pr_2^*B))$ must be lisse [Laumon, 1981, Corollary 2.1.2]. The theorem states that $R\pi_* u_! \mathcal{F}$ is lisse for a lisse sheaf $\mathcal{F}$ on an open set U of a scheme X with a proper map $\pi : X \rightarrow S$, such that the complement Y of U is a smooth divisor over S and such that $\mathcal{F}$ has constant Swan conductor along Y . Here we have a slightly different situation: Taking $S = \mathbb{A}_S^1 - \{x_i + y_j\}$, $X = \mathbb{P}_S^1$, and U the complement of the lines $x' - y' = x_i$, $y' = y_j$, and $y' = \infty$, instead of $u_! \mathcal{F}$ we have $j_!((d^*A) \otimes (pr_2^*B))$, a complex of sheaves lisse on U and lisse on the complement of U . But it follows that $R\pi_* j_!((d^*A) \otimes (pr_2^*B))$ is lisse in our case as well by a simple reduction. First, by the Leray spectral sequence, it suffices to show $R\pi_* \mathcal{H}^i(j_!((d^*A) \otimes (pr_2^*B)))$ is lisse. Second, for i the inclusion of Y there is an exact sequence $u_! u^*(j_!((d^*A) \otimes (pr_2^*B))) \rightarrow (j_!((d^*A) \otimes (pr_2^*B))) \rightarrow i_* i^*(j_!((d^*A) \otimes (pr_2^*B)))$, so it suffices to show that $R\pi_* u_! u^*(j_!((d^*A) \otimes (pr_2^*B))) \rightarrow (j_!((d^*A) \otimes (pr_2^*B)))$ and $R\pi_* i_* i^*(j_!((d^*A) \otimes (pr_2^*B)))$ are lisse. The first is lisse by [Laumon, 1981, Corollary 2.1.2] and the second one is lisse because $\pi \circ i$ is finite étale and so pushing forward along it preserves lisseness.

□

Lemma 3.0.33. *Let K be a separably closed field, and let A and B be two perverse sheaves on $\mathbb{A}_K^1$ that are tamely ramified at ∞ , one of which has property $\mathcal{P}$. Let $\chi(A)$ and $\chi(B)$ be the Euler characteristics of A and B on $\mathbb{A}_K^1$. Then:*

$$\text{rank}(A *_! B) = \text{rank}(A)\chi(B) + \chi(A)\text{rank}(B) + \text{rank}(A)\text{rank}(B)$$

Proof. Since convolution commutes with étale base change, we may base change to the separable closure of K , and so without loss of generality we may assume that K is separably closed.

Because one of A or B has property $\mathcal{P}$, $A *_! B$ is perverse. Over an open set, it is lisse of rank $\text{rank}(A *_! B)$ in degree -1 , and so the Euler characteristic of its stalk at

any point in this open set is $-\text{rank}(A *_! B)$. We will compute the Euler characteristic and hence compute the rank.

Let $x_1, \dots, x_n$ be the coordinates of the non-lisse points of A , and let $y_1, \dots, y_m$ be the coordinates of the non-lisse points of B . Fix a point z such that $z \neq x_i + y_j$ for any $1 \leq i \leq n, 1 \leq j \leq m$. By Lemma 3.0.32, $A *_! B$ sheaf is lisse at z .

Let r be the map $\mathbb{A}_K^1 \rightarrow \mathbb{A}_K^1$ that sends x to $z - x$. By proper base change, we see that $(A *_! B)_z$ is just $H_c^*(\mathbb{A}_K^1, A \otimes r^*B)$.

Because $x_1, \dots, x_n, z - y_1, \dots, z - y_m$ are all distinct, in a neighborhood of each point of $\mathbb{A}^1$ either A or r^*B is lisse in degree -1 , which means that locally $A \otimes r^*B$ is a perverse sheaf shifted by one degree. Since perversity is a local condition, $A \otimes r^*B$ is a perverse sheaf shifted by one degree. We compute its Euler characteristic using the Euler-Poincare formula, Lemma 3.0.31, with an additional minus sign.

Let:

$$a_i = \text{drop}(\mathcal{H}^{-1}(A), x_i) + \dim(\mathcal{H}^0(A), x_i) + \text{swan}(\mathcal{H}^{-1}(A), x_i)$$

$$b_j = \text{drop}(\mathcal{H}^{-1}(B), y_j) + \dim(\mathcal{H}^0(N), y_j) + \text{swan}(\mathcal{H}^{-1}(B), y_j)$$

Then a_i is the local contribution to the Euler characteristic of A from the point x_i , and b_j is the local contribution to the Euler characteristic of B from the point y_j . By Lemma 3.0.31,

$$\begin{aligned} \chi(A) &= -\text{rank}(A) + \sum_{i=1}^n a_i \\ \chi(B) &= -\text{rank}(B) + \sum_{j=1}^m b_j \end{aligned}$$

Since r^*B is lisse at x_i and its rank is $\text{rank}(B)$, the local contribution to the Euler characteristic of $A \otimes r^*B$ at x_i is $-\text{rank}(B)a_i$, and since A is lisse at $z - y_j$ and its

rank is $\text{rank}(A)$, the local contribution to the Euler characteristic of $A \otimes r^*B$ at $z - y_j$ is $-\text{rank}(A)b_j$. The generic rank of $A \otimes r^*B$ is $\text{rank}(A)\text{rank}(B)$. So we have:

$$\begin{aligned}
\chi(\mathbb{A}^1, A \otimes r^*B) &= \text{rank}(A)\text{rank}(B) - \sum_{i=1}^n \text{rank}(B)a_i - \sum_{j=1}^m \text{rank}(A)b_j \\
&= -\text{rank}(B) \left(-\text{rank}(A) + \sum_{i=1}^n a_i \right) - \text{rank}(A) \left(-\text{rank}(B) + \sum_{j=1}^m b_j \right) - \text{rank}(A)\text{rank}(B) \\
&= -(\chi(A)\text{rank}(B) + \text{rank}(A)\chi(B) + \text{rank}(A)\text{rank}(B))
\end{aligned}$$

Because the Euler characteristic is $-(\chi(A)\text{rank}(B) + \text{rank}(A)\chi(B) + \text{rank}(A)\text{rank}(B))$, and the sheaf is lisse in degree -1 , its rank must be $-(\chi(A)\text{rank}(B) + \text{rank}(A)\chi(B) + \text{rank}(A)\text{rank}(B))$.

□

Lemma 3.0.34. *Let S be a smooth scheme over $\mathbb{Z}_p$. Let j be the open immersion $\mathbb{A}_S^1 \rightarrow \mathbb{P}_S^1$. Let $A \in D_c^b(\mathbb{A}_{S_{\mathbb{Q}_p}}^1, \overline{\mathbb{Q}}_\ell)$ be a complex. If A is extensible, then:*

1.

$$R\Psi_p j_! A = j_! R\Psi_p A$$

2. $R\Psi_p A$ is tamely ramified at ∞ .

Proof. Because nearby cycles are defined locally, the first statement is certainly true everywhere on $\mathbb{A}_{\mathbb{F}_p}^1$. It remains to verify the statement on the divisor at ∞ of $\mathbb{P}_{S_{\mathbb{F}_p}}^1$. Because both statements are local questions on the divisor at ∞ , we may work in the neighborhood of a point on that divisor.

Let L be a complex of lisse sheaves on an open set U of $\mathbb{A}_S^1$, as described in the definition of extensibility. Let Z be its bounded closed complement. Rather than working with $j_! A$ on all of $\mathbb{P}_S^1$, we will work with $j_! L$ on $\mathbb{P}_S^1 - Z$.

$j_!L$ is lisse on $\mathbb{P}_S^1 - Z$ away from the divisor at ∞ . Because the divisor at ∞ is a smooth divisor which is generically in characteristic 0, L is tamely ramified at the generic point of that divisor, and hence is tamely ramified around that divisor. Thus the vanishing cycles $R\Phi_p j_!L$ of $j_!L$ relative to $\mathbb{Z}_p$ are 0 by [Deligne and Katz, 1973, XIII Lemma 2.1.11].

Because $R\Phi_p j_!L$ vanishes, $R\Psi_p j_!L_{\mathbb{Q}_p} = j_!L_{\overline{\mathbb{F}}_p}$.

Away from Z , $L_{\mathbb{Q}_p}$ is isomorphic to A , so $R\Psi_p j_!A = j_!L_{\overline{\mathbb{F}}_p}$. Hence $R\Psi_p j_!A$ is zero at ∞ , verifying the first statement, and $R\Psi_p j_!A = j_!R\Psi_p A$ is tamely ramified around ∞ , verifying the second statement. $\square$

Lemma 3.0.35. *If $A \in \mathcal{Perv}$ has good reduction at p , then:*

$$\chi(\mathbb{A}_{\overline{\mathbb{F}}_p}^1, R\Psi_p A) = \chi(\mathbb{A}_{\mathbb{Q}}^1, A)$$

Proof. The Euler characteristic is preserved by extension by zero, so:

$$\chi(\mathbb{A}_{\mathbb{Q}}^1, A) = \chi(\mathbb{P}_{\mathbb{Q}}^1, j_!A) = \chi(\mathbb{P}_{\overline{\mathbb{F}}_p}^1, R\Psi_p j_!A) = \chi(\mathbb{P}_{\overline{\mathbb{F}}_p}^1, j_!R\Psi_p A) = \chi(\mathbb{A}_{\overline{\mathbb{F}}_p}^1, R\Psi_p A)$$

The middle two equalities follow, respectively, from the fact that the cohomology of the nearby cycles sheaf on the special fiber is the cohomology of the original sheaf on the geometric generic fiber and from Lemma 3.0.34. $\square$

Lemma 3.0.36. *If A and B have good reduction at p , then:*

$$R\Psi_p A *_! R\Psi_p B = R\Psi_p (A *_! B)$$

Proof. We apply the compactification of Remark 3.0.12 to compute $R\Psi_p A *_! R\Psi_p B$:

$$(R\Psi_p A) *_! (R\Psi_p B) = R\pi_* u_! ((d^* R\Psi_p A) \otimes (pr_2^* R\Psi_p B))$$

We will verify the sequence of identities:

$$\begin{aligned}
& R\pi_* j_!((d^* R\Psi_p A) \otimes (pr_2^* R\Psi_p B)) \\
&= R\pi_* j_! R\Psi_p((d^* A) \otimes (pr_2^* B)) \\
&= R\pi_* R\Psi_p j_!((d^* A) \otimes (pr_2^* B)) \\
&= R\Psi_p R\pi_* j_!((d^* A) \otimes (pr_2^* B)) \\
&= R\Psi_p(A *! B)
\end{aligned}$$

The first identity is valid because $d^* A \otimes pr_2^* B$ is the external tensor product $A \boxtimes B$ composed with a change of coordinates. Nearby cycles compute with external tensor products by [Beilinson and Bernstein, 1993, Lemma 5.1.1], and they clearly commute with change of coordinates.

The second identity will follow from Lemma 3.0.34 if we can show that $d^* A \otimes pr_2^* B$ is extensible as a complex on $\mathbb{A}^2 = \mathbb{A}_{\mathbb{A}^1}^1$. The perverse sheaves A and B are both extensible, hence isomorphic to lisse complexes on open sets U and V with bounded complements Z_1 and Z_2 . The tensor product $d^* A \otimes pr_2^* B$ is isomorphic to a lisse complex on $d^* U \cap pr_2^* V$, so it is sufficient to show that its complement $d^* Z_1 \cup pr_2^* Z_2$ is bounded. It is sufficient to show that the maps: $d, pr_2 : \mathbb{A}^1 \times \mathbb{A}^1 \rightarrow \mathbb{A}^1$ can be extended to maps $\mathbb{A}^1 \times \mathbb{P}^1 \rightarrow \mathbb{P}^1$, so that the inverse image of a set in $\mathbb{A}^1$ closed in $\mathbb{P}^1$ will be a set in $\mathbb{A}^1 \times \mathbb{A}^1$ closed in $\mathbb{A}^1 \times \mathbb{P}^1$, hence bounded. For pr_2 , the extension is still pr_2 , and for d it extends because the map that sends y to $x - y$ is a well-defined automorphism of $\mathbb{P}^1$ for each finite x .

The third identity follows because $R\Psi_p$ commutes with proper pushforward.

The fourth identity is a tautology. □

Lemma 3.0.37. *Let A and B be two perverse sheaves, at least one of which has property $\mathcal{P}$. If A and B have good reduction at p , then $A *_! B$ has good reduction at p .*

Proof. For the first condition for good reduction, we can see, using Lemma 3.0.33, that

$$\text{rank}(A *_! B) = \text{rank}(A)\chi(B) + \chi(A)\text{rank}(B) + \text{rank}(A)\text{rank}(B).$$

(by definition and Lemma 3.0.35)

$$= \text{rank}(R\Psi_p A)\chi(R\Psi_p B) + \chi(R\Psi_p A)\text{rank}(R\Psi_p B) + \text{rank}(R\Psi_p A)\text{rank}(R\Psi_p B)$$

(by Lemma 3.0.33)

$$= \text{rank}(R\Psi_p A *_! R\Psi_p B)$$

(by Lemma 3.0.36)

$$= \text{rank}(R\Psi_p(A *_! B))$$

For the second condition, observe that the singularities of $A *_! B$ are all sums of the singularities of A and B by Lemma 3.0.32. Because sums of p -adic integers are p -adic integers, the singularities of $A *_! B$ are p -adic integers.

For the third, note that if I_p acts trivially on $R\Psi_p A$ and $R\Psi_p B$, it acts trivially on $R\Psi_p A *_! R\Psi_p B$ by functoriality, which is the same thing as $R\Psi_p(A *_! B)$ by Lemma 3.0.36. □

Lemma 3.0.38. *Let*

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$$

be an exact sequence of perverse sheaves. If B has good reduction at p , then A and C do as well.

In other words, subobjects and quotients of perverse sheaves with good reduction have good reduction.

Proof. We check for A and C the three conditions of good reduction.

Observe that:

$$0 \rightarrow R\Psi_p A \rightarrow R\Psi_p B \rightarrow R\Psi_p C \rightarrow 0$$

is also a short exact sequence.

For generic rank, , we have:

$$\text{rank}(A) + \text{rank}(C) = \text{rank}(B) = \text{rank}(R\Psi_p B) = \text{rank}(R\Psi_p A) + \text{rank}(R\Psi_p C) \leq \text{rank}(A) + \text{rank}(C)$$

because rank is additive in exact sequences and because B has good reduction. Thus the final inequality, due to Corollary 3.0.29, is an equality, so $\text{rank}(R\Psi_p A) = \text{rank}(A)$ and $\text{rank}(R\Psi_p C) = \text{rank}(C)$.

The singularities of A and C must occur at the singularities of B , since a perverse sheaf lisse at a point can never be decomposed into two perverse sheaves not lisse at that point (Given a lisse perverse sheaf A , it is sufficient to check that, in some neighborhood of that point, all the Jordan-Holder factors of A in the category of perverse sheaves are lisse. But in a neighborhood of that point A is lisse, hence a successive extension of irreducible lisse sheaves, and irreducible lisse sheaves are irreducible perverse sheaves, so those are the Jordan-Holder factors of A .)

For the inertia action, subobjects and quotients of objects with trivial group action have trivial group action, hence because I_p acts trivially on $R\Psi_p B$, it must also act trivially on $R\Psi_p A$ and $R\Psi_p C$.

□

Corollary 3.0.39. *An equivalence class in $\mathcal{Perv}/\mathcal{Neg}$ contains an element with good reduction at p if and only if its property $\mathcal{P}$ representative has good reduction at p .*

Proof. By Lemma 3.0.21, the property $\mathcal{P}$ representative can be constructed from any given element by repeated subobjects and quotients. By Lemma 3.0.38, subobjects and quotients preserve good reduction. So if there is any element with good reduction, by repeatedly taking subobjects and quotients, the property $\mathcal{P}$ representative has good reduction. The converse is obvious. $\square$

Definition 3.0.40. We say that an equivalence class in $\mathcal{Perv}/\mathcal{Neg}$ has good reduction at p if and only if its property $\mathcal{P}$ representative has good reduction at p . (Equivalently, by Corollary 3.0.39, if any element of the class has good reduction at p .)

Lemma 3.0.41. *The full subcategory of $\mathcal{Perv}/\mathcal{Neg}$ consisting of objects with good reduction at p is closed under direct sums, kernels, cokernels, convolution, and duality.*

Proof. A kernel or cokernel is defined for an arbitrary map between two elements of the equivalence class. Using Lemma 3.0.21, we see that we can always take the map to be between two elements with property $\mathcal{P}$. The kernel and cokernel may not have property $\mathcal{P}$, but they certainly have good reduction by Lemma 3.0.38, so their equivalence classes have good reduction.

Similarly for direct sum, we just need to check that the direct sum of perverse sheaves with good reduction has good reduction. For all three of the conditions, this is obvious. The singularities of $A + B$ are the singularities of A and the singularities of B , and $R\Psi_p(A + B) = R\Psi_p A + R\Psi_p B$.

For convolution, we may assume A and B have property $\mathcal{P}$, and then $A *_! B$ has good reduction by Lemma 3.0.37.

For duality, we know that by Lemma 3.0.22, duals in the $\mathcal{Perv}/\mathcal{Neg}$ are given by $A^\vee = [-1]^* DA$. Good reduction is clearly preserved by $[-1]^*$ so we focus on Verdier duality. Verdier duality preserves generic rank, and commutes with nearby cycles [Illusie, 1994, Theorem 4.2], so the condition about ranks is satisfied. Verdier duality

preserves the singular set so the condition about singularities is satisfied. Duality is functorial, so preserves the inertia action, so that condition is satisfied as well. $\square$

The following proposition is not needed for other results in this paper, but is useful:

Proposition 3.0.42. *Let $A \in \mathcal{Perv}$ have property $\mathcal{P}$ and good reduction at p . Then $R\Psi_p A$ also has property $\mathcal{P}$.*

Proof. By [Katz, 1995, Lemmas 2.6.13, 2.6.14 and 2.6.15], $R\Psi_p A$ has property $\mathcal{P}$ if and only if it has no subobjects or quotients that are Artin-Schreier sheaves. By Lemma 3.0.34, $R\Psi_p A$ is tamely ramified at ∞ , so the only Artin-Schreier sheaf that can be a subobject or a quotient is the trivial sheaf. Just as in Lemma 3.0.20, we have:

$$\mathrm{Hom}_{\mathbb{A}_{\mathbb{F}_p}^1}(\overline{\mathbb{Q}_\ell}[1], R\Psi_p A) = H^{-1}(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A)$$

$$\mathrm{Hom}_{\mathbb{A}_{\mathbb{F}_p}^1}(R\Psi_p A, \overline{\mathbb{Q}_\ell}[1]) = H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A)^\vee$$

So it is sufficient to show that $H^{-1}(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A)$ and $H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A)$ vanish. Let j be the open immersion $\mathbb{A}^1 \rightarrow \mathbb{P}^1$, then by Lemma 3.0.34 and Lemma 3.0.20

$$H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) = H^1(\mathbb{P}_{\mathbb{F}_p}^1, j_! R\Psi_p A) = H^1(\mathbb{P}_{\mathbb{F}_p}^1, R\Psi_p j_! A) = H^1(\mathbb{P}_{\mathbb{Q}}^1, j_! A) = H_c^1(\mathbb{A}_{\mathbb{Q}}^1, A) = 0$$

and because DA also has property $\mathcal{P}$ and good reduction at p (Lemma 3.0.41)

$$H^{-1}(\mathbb{A}_{\mathbb{F}_p}^1, A) = H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, DR\Psi_p A)^\vee = H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p DA)^\vee = 0$$

Here we also use the fact that nearby cycles commute with duality [Illusie, 1994, Theorem 4.2]. $\square$

Lemma 3.0.43. *Let ψ be a nontrivial additive character of $\mathbb{F}_p$. Let $\mathcal{L}_\psi$ be the corresponding character sheaf. Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ with good reduction at p .*

(1) *$H^i(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ vanishes for $i \neq 0$.*

(2) *The natural forget supports map*

$$H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) \rightarrow H^i(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$$

is an isomorphism.

Proof. Let u be the open immersion $\mathbb{A}^1 \rightarrow \mathbb{P}^1$. We will show that the natural forget supports map $u_!(R\Psi_p A \otimes \mathcal{L}_\psi) \rightarrow Ru_*(R\Psi_p A \otimes \mathcal{L}_\psi)$ is an isomorphism.

It is sufficient to show that all the stalks of $Ru_*(R\Psi_p A \otimes \mathcal{L}_\psi)$ at ∞ vanish. This is true because A has good reduction, so by Lemma 3.0.34 $R\Psi_p A$ is tamely ramified at ∞ , so $R\Psi_p A \otimes \mathcal{L}_\psi$ is totally wildly ramified at ∞ , and in particular has no local monodromy invariants or coinvariants. The stalk of $Ru_* R\Psi_p(A \otimes \mathcal{L}_\psi)$ at ∞ can be computed by a spectral sequence from the monodromy invariants and coinvariants and so vanishes.

Hence the natural map

$$H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) = H^i(\mathbb{P}_{\mathbb{F}_p}^1, u_! R\Psi_p A \otimes \mathcal{L}_\psi) \rightarrow H^i(\mathbb{P}_{\mathbb{F}_p}^1, Ru_* R\Psi_p A \otimes \mathcal{L}_\psi) = H^i(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$$

is an isomorphism.

To prove the first part, note that $R\Psi_p A$ is a perverse sheaf, so $R\Psi_p A \otimes \mathcal{L}_\psi$ is as well, and H^i of an affine variety with coefficients in a perverse sheaf always vanishes for $i > 0$. Similarly the compactly supported cohomology vanishes for $i < 0$. Because they are equal, they both vanish away from 0. $\square$

Lemma 3.0.44. $A \mapsto H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is an exact tensor functor from the full subcategory of $\mathcal{Perv}/\mathcal{Neg}$ consisting of objects with good reduction at p to the category of vector spaces.

Proof. First note that there is no ambiguity regarding elements of $\mathcal{Neg}$: The functor sends them to 0, because constant sheaves on $\mathbb{A}_{\mathbb{Q}_p}^1$ are sent to constant sheaves on $\mathbb{A}_{\mathbb{F}_p}^1$ by $R\Psi_p$, which when tensored with $\mathcal{L}_\psi$ have no cohomology.

The functors $R\Psi_p$, $\otimes \mathcal{L}_\psi$, and cohomology all preserve distinguished triangles, and by Lemma 3.0.43 the cohomology groups above and below 0 vanish, so the long exact sequences become short exact sequences and hence the functor is exact.

Finally we check that this is a tensor functor.

$$H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p(A * B) \otimes \mathcal{L}_\psi)$$

(by the dual of Lemma 3.0.36)

$$= H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, (R\Psi_p A * R\Psi_p B) \otimes \mathcal{L}_\psi\right)$$

(by definition of $*$)

$$= H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, Rsum_*(R\Psi_p A \boxtimes R\Psi_p B) \otimes \mathcal{L}_\psi\right)$$

(by the Leray spectral sequence)

$$= H^0\left(\mathbb{A}_{\mathbb{F}_p}^2, (R\Psi_p A \boxtimes R\Psi_p B) \otimes sum^* \mathcal{L}_\psi\right)$$

(by [Katz, 1988, Section 4.3 (3)])

$$= H^0\left(\mathbb{A}_{\mathbb{F}_p}^2, (R\Psi_p A \otimes \mathcal{L}_\psi) \boxtimes (R\Psi_p B \otimes \mathcal{L}_\psi)\right)$$

(because of the Künneth formula and Lemma 3.0.43)

$$= H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi\right) \otimes H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p B \otimes \mathcal{L}_\psi\right)$$

We also need to check that this tensor isomorphism satisfies all the necessary commutative diagrams. As in Proposition 3.0.18 we will use a balanced definition to do this. We define for any finite set I a morphism $\gamma_I : H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p(*_I A_i) \otimes \mathcal{L}_\psi\right) \rightarrow \otimes_I H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A_i \otimes \mathcal{L}_\psi\right)$ by:

$$H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p(*_I A_i) \otimes \mathcal{L}_\psi\right)$$

(by definition)

$$= H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\psi_p(R\pi_{I*} \boxtimes_I A_I) \otimes \mathcal{L}_\psi\right)$$

(by base change)

$$\rightarrow H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\pi_{I*} R\psi_p \boxtimes_I A_I) \otimes \mathcal{L}_\psi\right)$$

([Beilinson and Bernstein, 1993, Lemma 5.1.1])

$$= H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, R\pi_{I*} \boxtimes_I R\psi_p A_I \otimes \mathcal{L}_\psi\right)$$

(by the Leray spectral sequence)

$$= H^0\left((\mathbb{A}^1)_{\mathbb{F}_p}^I, \boxtimes_I R\psi_p A_I \otimes \pi^* \mathcal{L}_\psi\right)$$

(by iteratively applying [Katz, 1988, Section 4.3 (3)])

$$= \left((\mathbb{A}^1)_{\mathbb{F}_p}^I, \boxtimes_I (R\psi_p A_I \otimes \mathcal{L}_\psi)\right)$$

(by the Künneth formula and Lemma 3.0.43)

$$= \otimes_I H^0 \left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A_i \otimes \mathcal{L}_\psi \right)$$

In order to check that $A \mapsto H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is a tensor functor, it suffices to check that:

- 1) The map γ_I is an isomorphism.
- 2) γ_I agrees with the identity in the case where I has cardinality 0 or 1 and agrees with the previously described isomorphism when I has cardinality 2.
- 3) For any map $\alpha : I \rightarrow J$ of finite sets, the following diagram commutes:

$$\begin{array}{ccccc} H^0 \left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p(*_I A_i) \otimes \mathcal{L}_\psi \right) & \xrightarrow{\gamma_I} & & \xrightarrow{\quad} & \otimes_I H^0 \left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A_i \otimes \mathcal{L}_\psi \right) \\ \downarrow \chi(\alpha) & & & & \downarrow \chi(\alpha) \\ H^0 \left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p(*_J *_{\alpha^{-1}(J)} A_i) \otimes \mathcal{L}_\psi \right) & \xrightarrow{\gamma_J} & \otimes_J H^0 \left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p(*_{\alpha^{-1}(J)} A_i) \otimes \mathcal{L}_\psi \right) & \xrightarrow{\prod_J \gamma_{\alpha^{-1}(J)}} & \otimes_J \otimes_{\alpha^{-1}(J)} H^0 \left(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A_i \otimes \mathcal{L}_\psi \right) \end{array}$$

Indeed, because the associativity isomorphism is the composition of two uses of $\chi(\alpha)$, that the associativity isomorphism commutes with γ follows from combining two copies of this diagram. The symmetry isomorphism and unit isomorphism follows similarly.

Furthermore 1) follows inductively from 2) and 3). Part 2) establishes it for I of cardinality at most 2. Suppose it is known for cardinality at most n and let I be a set with cardinality at most $n+1$. Let α be a function from this set to a two-element set that is not a constant function. Then in this commutative diagram, the γ s on the bottom row all are for sets of size at most n , so are isomorphisms, and the column maps are isomorphisms, so the top row γ_I is an isomorphism.

2) is not too hard. In the $|I| = 0$ case, simply observe that all the objects in the chain of morphisms defining γ_I are the same one-dimensional vector space and all the arrows are the identity. Similarly, in the $|I| = 1$ case π_I is the identity,

$\boxtimes$ is trivial, and all the isomorphisms and the base change map are used in their trivial cases where they are the identity. Comparing our definition of $|I| = 2$ with the one we gave earlier specific to that case, we see the only difference is the first two lines, where we use a base change map and the compatibility of $R\Psi$ and $\boxtimes$ instead of using Proposition 3.0.36. However, examining Proposition 3.0.36, it uses the compatibility of $\boxtimes$ and $R\Psi$ and takes a compactification $\pi \circ j$ of sum , then uses the map $R\pi_* j_! R\psi_p K \rightarrow R\pi_* R\Psi_p j_! K = R\Psi_p R\pi_* j_! K$, whose dual is the map $R\Psi_p R\pi_* Rj_* DK = R\pi_* R\Psi_p j_* DK \rightarrow R\pi_* Rj_* R\psi_p K$ coming from base change for π , which since base change is defined locally is the same as the base change map for $\pi \circ j$, which is the base change map of sum , matching our second morphism exactly.

3) is the interesting one. We divide it into two parts. The first part, more difficult, handles just $R\Psi_p$, and the second part handles the rest of the functor. The first part is to check that the commutative diagram using fragments of γ for the horizontal arrows

$$\begin{array}{ccc} R\Psi_p(*_I A_i) & \xrightarrow{\quad\quad\quad} & *_I R\Psi_p A_i \\ \downarrow \chi(\alpha) & & \downarrow \chi(\alpha) \\ R\Psi_p(*_J *_{\alpha^{-1}(j)} A_i) & \xrightarrow{\quad\quad\quad} *_J R\Psi_p(*_{\alpha^{-1}(j)} A_i) \xrightarrow{\quad\quad\quad} *_J *_{\alpha^{-1}(j)} R\Psi_p A_i \end{array}$$

To prove this we must check that the base change and compatibility-with- $\boxtimes$ morphisms for $R\Psi$ commute with the three isomorphisms we used to define $\chi(\alpha)$ - the Künneth formula, the Leray spectral sequence, and the monoidality of $\boxtimes$. We need to prove one smaller diagram commutative. For each of them we note only the arrows that are valid for arbitrary sheaves so that we can remove the derived functor and commute directly to verify that the diagram commutes.

Compatibility of $R\Psi$ and Künneth:

$$\begin{array}{ccccc} \boxtimes_I R\Psi_p Rf_{i*} A_i & \longrightarrow & \boxtimes_I Rf_{i*} R\Psi_p A_i & \longrightarrow & (\prod_{i \in I} Rf_i)_* \boxtimes_I R\Psi_p A_i \\ \downarrow & & & & \downarrow \\ R\Psi_p \boxtimes_I Rf_{i*} A_i & \longrightarrow & R\Psi_p (\prod_{i \in I} Rf_i)_* \boxtimes_I A_i & \longrightarrow & (\prod_{i \in I} Rf_i)_* R\Psi_p \boxtimes_I A_i \end{array}$$

Compatibility of $R\Psi$ and Leray:

$$\begin{array}{ccccc}
R\Psi_p Rf_* Rg_* A & \longrightarrow & Rf_* R\Psi_p Rg_* A & \longrightarrow & Rf_* Rg_* R\Psi_p A \\
\parallel & & & & \parallel \\
R\Psi_p R(f \circ g)_* A & \longrightarrow & & \longrightarrow & R(f \circ g)_* R\Psi_p A
\end{array}$$

Compatibility of $R\Psi$ and monoidality of $\boxtimes$:

$$\begin{array}{ccccc}
R\Psi_p \boxtimes_I A_i & \longleftarrow & & \boxtimes_I R\Psi_p A_i & \\
\downarrow & & & \downarrow & \\
R\Psi_p \boxtimes_J \boxtimes_{\alpha^{-1}(j)} A_i & \longleftarrow & \boxtimes_J R\Psi_p \boxtimes_{\alpha^{-1}(j)} A_i & \longleftarrow & \boxtimes_J \boxtimes_{\alpha^{-1}(j)} R\Psi_p A_i
\end{array}$$

For each diagram, we reduced to the analogous version for non-derived functors, check that it holds for skyscraper sheaves, and then check the reduction to skyscraper sheaves, following Lemma 3.0.18. The functors Rf_* and $\boxtimes$ preserve skyscraper sheaves and $R\Psi_p$ sends a skyscraper sheaf to either a skyscraper sheaf or the zero sheaf, so all the objects in these diagrams are either skyscraper sheaves or the zero sheaf.

First we check that for each diagram, if any object is the zero sheaf, then enough are that all maps are zero and the diagram commutes. This happens when $R\Psi_p$ is supported at a skyscraper sheaf whose reduction modulo p is not in the variety. But if this happens for Rf_* of a sheaf, then it must already happen for that sheaf, since if the reduction mod p of a point is in the variety then f of that reduction is in the variety. Similarly if this happens for one sheaf then it happens for the $\boxtimes$ of many sheaves, because if one point has bad reduction then its product must have bad reduction. So if any object in the first diagram is zero, the bottom-right object is, and if any object in the second diagram is zero, one on the right side is, and if any object in the third diagram is zero, one on the left side is. In either case the diagram is trivial.

Second we check, in the case that each object in each diagram is a skyscraper sheaf, that each morphism is the identity, preserving the canonical section. The

morphisms are base change for $R\Psi$, the isomorphism $R\Psi(A \boxtimes B) = R\Psi A \boxtimes R\Psi B$, as well as Künneth, Leray, and monoidality of $\boxtimes$. But none of these can possibly affect the canonical generator of a skyscraper sheaf because at no point is the section multiplied by anything.

Finally, we must check the case for a general sheaf. It is sufficient to show that, for each section of one of the final objects of the diagram, there are some maps from A_i to skyscraper sheaves that are nontrivial on that section. For each diagram, the final object (bottom-right in the first and bottom-left in the last two) comes from applying some combination of Ψ_p , $\boxtimes$, and f_* . Each nontrivial section of one of these functors applied to a sheaf comes from one or more nontrivial sections of the original sheaf. Following it backwards, from a section of the final object we obtain a tuple of sections of the A_i s. For each A_i , we can find a map to a skyscraper sheaf which sends that section to something nontrivial by taking a stalk. Applying functoriality for each functor in the composition, we see that the section of the final object is sent to a nontrivial section of the final object in the case where each A_i is replaced with a skyscraper sheaf, and we win.

The remainder is much simpler. We can handle it all with a single commutative diagram, where the horizontal arrows are the remainder of γ_I . The remainder of γ_I is an isomorphism, but only the leftmost arrow is defined for an arbitrary sheaf.

$$\begin{array}{ccccc}
H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, (*_I A_i) \otimes \mathcal{L}_\psi\right) & \longleftarrow & & \longleftarrow & \otimes_I H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, A_i \otimes \mathcal{L}_\psi\right) \\
\downarrow & & & & \downarrow \\
H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, (*_J *_{\alpha^{-1}(j)} A_i) \otimes \mathcal{L}_\psi\right) & \longleftarrow & \otimes_J H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, (*_{\alpha^{-1}(j)} A_i) \otimes \mathcal{L}_\psi\right) & \longleftarrow & \otimes_J \otimes_{\alpha^{-1}(j)} H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, A_i \otimes \mathcal{L}_\psi\right)
\end{array}$$

To check this diagram commutes it is sufficient to do so for the non-derived functors. First let us check this for skyscraper sheaves. Convolution, tensoring with $\mathcal{L}_\psi$, and taking cohomology all send skyscraper sheaves to skyscraper sheaves, but ten-

soring with $\mathcal{L}_\psi$ does not fix the canonical generator. Instead choosing a generator for $\delta_x \otimes \mathcal{L}_\psi$ is the same as choosing a generator of $\mathcal{L}_\psi(x)$. Similarly all the isomorphisms involved but [Katz, 1988, Section 4.3 (3)] already preserve generators. So it remains to check that the action of the iterated isomorphism of [Katz, 1988, Section 4.3 (3)] on stalks does not depend on the order the isomorphism is used in. In other words, the isomorphism $(x+y)^*\mathcal{L}_\psi = x^*\mathcal{L}_\psi \otimes y^*\mathcal{L}_\psi$ allows us to write $(x+y+z)^*\mathcal{L}_\psi$ as $x^*\mathcal{L}_\psi \otimes y^*\mathcal{L}_\psi \otimes z^*\mathcal{L}_\psi$ in two different ways, the first passing through $(x+y)^*\mathcal{L}_\psi \otimes z^*\mathcal{L}_\psi$ and the second passing through $x^*\mathcal{L}_\psi \otimes (y+z)^*\mathcal{L}_\psi$. Are these two isomorphisms the same? Yes. They are irreducible lisse sheaves on $\mathbb{A}^3$, so it is sufficient to check this on the stalks at any given point. We choose the point $(0,0,0)$ where the stalk of every involved sheaf has a canonical generator and the map is the isomorphism.

To reduce from all sheaves to skyscraper sheaves it is sufficient to show that every section of $H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, (*_J *_{\alpha^{-1}(j)} A_i) \otimes \mathcal{L}_\psi\right)$ maps nontrivially to a section of $H^0\left(\mathbb{A}_{\mathbb{F}_p}^1, (*_J *_{\alpha^{-1}(j)} \delta_{x_i}) \otimes \mathcal{L}_\psi\right)$ under some tuple of maps $A_i \rightarrow \delta_{x_i}$. But such a section is a global section of $*_J *_{\alpha^{-1}(j)} A_i$ tensored with $\mathcal{L}_\psi$, which comes from some section of $\boxtimes_J *_{\alpha^{-1}(j)} A_i$ tensored with $\mathcal{L}_\psi$, which comes locally from a product of sections of $*_{\alpha^{-1}(j)} A_i$ tensored with $\mathcal{L}_\psi$, each of which arises locally from a product of sections of A_i , which must have nonzero stalks at some point. Then those stalks give nontrivial maps to the skyscraper sheaf.

□

Theorem 3.0.45. *Let S be a finite set of primes including ℓ . Let $\mathcal{CC}_S$ the full subcategory of $\mathcal{Perv}/\mathcal{Neg}$ consisting of objects with good reduction at all primes p not in S with the sheaf convolution monoidal structure.*

$\mathcal{CC}_S$ is a neutral Tannakian category.

For any prime p not in S and ψ any nontrivial exponential character, a fibre functor is $A \mapsto H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$. There is a natural action of Frob_p as an automorphism of this functor.

Proof. We must check that $\mathcal{CC}_S$ is a rigid symmetric monoidal abelian category with an exact tensor functor to the category of vector spaces. Since $\mathcal{CC}_S$ is defined as a full subcategory of a rigid symmetric monoidal abelian category, it is sufficient to check it is closed under direct sums, kernels, cokernels, duals, and tensors, and contains the unit. This follows from Lemma 3.0.41, together with the fact that it contains the unit object, the skyscraper sheaf of rank 1 at 0.

The fact that $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is an exact tensor functor is Lemma 3.0.44. Thus $\mathcal{CC}_S$ category is Tannakian.

To obtain the Frob_p action, recall that $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$ acts on $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$. The inertia subgroup acts trivially on $R\Psi_p A$ by the definition of good reduction, hence trivially on $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$, so the quotient of the Galois group by the inertia subgroup, which contains Frob_p , has a well-defined action. $\square$

Proposition 3.0.46. *Let N be a positive integer multiple of ℓ . Let A be an element of $D_c^b(\mathbb{A}_{\mathbb{Z}[\frac{1}{N}]}^1, \overline{\mathbb{Q}_\ell})$. Then there is a finite set of primes S including all the primes dividing N such that, for each p not in S , for each i , ${}^p\mathcal{H}^i(A_{\mathbb{Q}})$ has good reduction at p and $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p {}^p\mathcal{H}^i(A_{\mathbb{Q}}) \otimes \mathcal{L}_\psi)$ is isomorphic to $H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, A_{\mathbb{F}_p} \otimes \mathcal{L}_\psi)$ as a vector space with an action of Frob_p .*

We can take S to be the set of primes p that either divide N or such that one of the following conditions fails:

- *A is lisse in a neighborhood of the generic point of $\mathbb{A}_{\mathbb{F}_p}^1$.*
- *The closed set $Z \subset \mathbb{A}_{\mathbb{Z}[\frac{1}{N}]}^1$ where A is not lisse is a finite étale cover of $\text{Spec } \mathbb{Z}[\frac{1}{N}]$ over a neighborhood of p .*
- *When pulled back to Z , A is lisse again.*

Proof. Let p be a prime not in S . Because p satisfies the stated conditions, Deligne's semicontinuity theorem implies that the Swan conductor around Z is semicontinuous

in a neighborhood of p [Laumon, 1981, Theorem 2.1.1(i)]. Since the Swan conductor is generically 0 in characteristic 0 where the sheaf is tame, the Swan conductor around Z is constant in that neighborhood.

Let U be the complement of Z , j the inclusion of U , and i the inclusion of Z . As j^*A satisfies the conditions of [Laumon, 1981, Theorem 2.1.1(ii)] in a neighborhood of p , $R\Phi_p j_! j^*A$ vanishes. Furthermore $R\Phi_p i^*A$ vanishes because it is complex of lisse sheaves on a smooth scheme, thus $R\Phi_p i_* i^*A = i_* R\Phi_p i^*A$ vanishes because i_* being finite, commutes with $R\Phi_p$, so by the distinguished triangle $j_! j^*A \rightarrow A \rightarrow i_* i^*A$, $R\Phi_p A = 0$.

Therefore the natural map $A_{\mathbb{F}_p} \rightarrow R\Psi_p A_{\mathbb{Q}}$ is an isomorphism.

Using this isomorphism, we will show that ${}^p\mathcal{H}^i(A_{\mathbb{Q}})$ has good reduction at p . The key point is that because $R\Psi_p$ preserves perversity and distinguished triangles, it is an exact functor from the category of perverse sheaves on $\mathbb{A}_{\mathbb{Q}_p}^1$ to the category of perverse sheaves on $\mathbb{A}_{\mathbb{F}_p}^1$ and thus

$${}^p\mathcal{H}^i \circ R\Psi_p = R\Psi_p \circ {}^p\mathcal{H}^i$$

Note that A is lisse in a neighborhood of the generic point of $\mathbb{A}_{\mathbb{F}_p}^1$. On an open set where it is lisse, ${}^p\mathcal{H}^i$ is the same as H^{i-1} , shifted by 1 [Katz, 1995, Lemma 2.3.2.1(c)]. So $\text{rank}({}^p\mathcal{H}^i(A_{\mathbb{Q}}))$ is the rank of $\mathcal{H}^{i-1}(A)$ on the open set where it is lisse, which includes the special fiber at p , so it is also $\text{rank}({}^p\mathcal{H}^i(A_{\mathbb{F}_p}))$. Then:

$${}^p\mathcal{H}^i(A_{\mathbb{F}_p}) = {}^p\mathcal{H}^i(R\Psi_p A_{\mathbb{Q}}) = R\Psi_p {}^p\mathcal{H}^i(A_{\mathbb{Q}})$$

hence

$$\text{rank}({}^p\mathcal{H}^i(A_{\mathbb{Q}})) = \text{rank}({}^p\mathcal{H}^i(A_{\mathbb{F}_p})) = \text{rank}(R\Psi_p {}^p\mathcal{H}^i(A_{\mathbb{Q}}))$$

so ${}^p\mathcal{H}^i(A_{\mathbb{Q}})$ satisfies the first condition of good reduction at p .

Since the set of singularities of A is finite étale at p , the singular points of $A_{\mathbb{Q}}$, and thus of ${}^p\mathcal{H}^i(A_{\mathbb{Q}})$, are located at p -adic integers, so ${}^p\mathcal{H}^i(A_{\mathbb{Q}})$ satisfies the second condition.

Finally, the inertia action on ${}^p\mathcal{H}^i(A_{\mathbb{F}_p})$ is trivial. Since this is equal to $R\Psi_p {}^p\mathcal{H}^i(A_{\mathbb{Q}})$, the inertia action on $R\Psi_p {}^p\mathcal{H}^i(A_{\mathbb{Q}})$ is trivial as well, so ${}^p\mathcal{H}^i(A_{\mathbb{Q}})$ satisfies the third condition.

Hence ${}^p\mathcal{H}^i(A_{\mathbb{Q}})$ has good reduction at p for all p not in S .

Tensoring with $\mathcal{L}_\psi$ and taking cohomology, we have the $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$ -equivariant isomorphism

$$H_c^i(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, A_{\mathbb{F}_p} \otimes \mathcal{L}_\psi) \cong H_c^i(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, R\Psi_p A_{\mathbb{Q}} \otimes \mathcal{L}_\psi).$$

$H_c^i(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, R\Psi_p A_{\mathbb{Q}} \otimes \mathcal{L}_\psi)$ is given by a spectral sequence in terms of $H_c^a(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, R\Psi_p {}^p\mathcal{H}^b(A_{\mathbb{Q}}) \otimes \mathcal{L}_\psi)$. Because ${}^p\mathcal{H}^b(A_{\mathbb{Q}})$ has good reduction at p , $H_c^a(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, R\Psi_p {}^p\mathcal{H}^b(A_{\mathbb{Q}}) \otimes \mathcal{L}_\psi)$ vanishes for $a \neq 0$ by Lemma 3.0.43, and so the spectral sequence degenerates and we have the $\text{Gal}(\overline{\mathbb{F}_p}|\mathbb{F}_p)$ -equivariant isomorphism:

$$H_c^i(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, A_{\mathbb{F}_p} \otimes \mathcal{L}_\psi) = H_c^0(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, R\Psi_p {}^p\mathcal{H}^i(A_{\mathbb{Q}}) \otimes \mathcal{L}_\psi)$$

which gives the desired equality of Frobenius actions. H_c^0 is the same as H^0 here by Lemma 3.0.43. $\square$

Definition 3.0.47. Let G_S be the Tannakian fundamental group of $\mathcal{CC}_S$.

Definition 3.0.48. Let Frob_p be the conjugacy class in G_S given by the action of Frob_p on the fibre functor of $\mathcal{CC}_S$ given by $A \mapsto H^0(\mathbb{A}_{\overline{\mathbb{F}_p}}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ for ψ the standard additive character $\psi(x) = e^{2\pi i \frac{x}{p}}$.

We now prove Theorem 1.0.1.

Proof. Let $\mathcal{F}$ be a constructible sheaf on $\mathbb{A}_{\mathbb{Z}[\frac{1}{\ell}]}^1$. We apply Proposition 3.0.46 to $\mathcal{F}$, and let V be the sum of ${}^p\mathcal{H}^i(\mathcal{F}_{\mathbb{Q}})$ for i even and W be the sum for i odd. (Actually

$i = 0, 1$ are the only ones that appear, but this is not needed for our argument.)

$$\sum_{x=0}^{p-1} e^{2\pi i \frac{x}{p}} \operatorname{tr}(\mathcal{F}, \operatorname{Frob}_p, x) = \operatorname{tr}(V, \operatorname{Frob}_p) - \operatorname{tr}(W, \operatorname{Frob}_p)$$

By Equation 3.1, the left side is equal to:

$$\begin{aligned} & \sum_{i=0}^2 (-1)^i \operatorname{tr}(H_c^i(\mathbb{A}_{\overline{\mathbb{F}}_p}^1, \mathcal{F}_{\mathbb{F}_p} \otimes \mathcal{L}_\psi), \operatorname{Frob}_p) \\ &= \sum_{i=0}^2 (-1)^i \operatorname{tr}(H_c^i(\mathbb{A}_{\overline{\mathbb{F}}_p}^1, R\Psi_p({}^p\mathcal{H}^i(\mathcal{F}_{\mathbb{Q}}) \otimes \mathcal{L}_\psi), \operatorname{Frob}_p) \\ &= \sum_{i=0}^2 (-1)^i \operatorname{tr}({}^p\mathcal{H}^i(\mathcal{F}_{\mathbb{Q}}), \operatorname{Frob}_p) \\ &= \operatorname{tr}(V, \operatorname{Frob}_p) - \operatorname{tr}(W, \operatorname{Frob}_p) \end{aligned}$$

as desired. □

Remark 3.0.49. It is necessary to use perverse sheaves for any study of sheaf convolution because the category of usual ℓ -adic sheaves is not Artinian and so can never be given the structure of a Tannakian category. Because we do not know of any good theory of perverse sheaves over a base other than a field, we need to use perverse sheaves over $\mathbb{Q}$. Specifically, the fact that perverse sheaves form an abelian category is key. Nearby cycles are the best way to get from sheaves over $\mathbb{Q}$ to sheaves over $\overline{\mathbb{F}}_p$, because they preserve information, such as the Euler characteristic, about the original sheaf. This idea was suggested to me by Pierre Deligne.

As these theorems show, for most primes, specializing a sheaf with nearby cycles is the same as taking the usual pullback, so at most primes the formalism of nearby cycles is not required to specialize. However, to show that the set of bad primes does not increase requires careful analysis of nearby and vanishing cycles.

Chapter 4

Properties of the Tannakian Category

Let S be a finite set of primes including ℓ . In this chapter, we prove several useful lemmas about G_S and its category of representations $\mathcal{CC}_S$.

First we will check that some standard notions like irreducibility and weights behave as expected in this category:

Lemma 4.0.50. *The property $\mathcal{P}$ representative of an irreducible object of $\mathcal{CC}_S$ is an irreducible perverse sheaf.*

Proof. Let A be the property $\mathcal{P}$ representative of an equivalence class in $\mathcal{Perv}/\mathcal{Neg}$ corresponding to an irreducible representation. We will show that A is irreducible. Suppose there is an exact sequence $0 \rightarrow B \rightarrow A \rightarrow C \rightarrow 0$ in $\mathcal{Perv}$. Then by Lemma 3.0.38, B and C also have good reduction away from S . This sequence remains exact in $\mathcal{Perv}/\mathcal{Neg}$ by construction of quotient categories and so remains exact in $\mathcal{CC}_S$. However as A maps to an irreducible object in $\mathcal{CC}_S$, either B or C must map to the 0 object in $\mathcal{CC}_S$, and thus be in $\mathcal{Neg}$. But A has Property $\mathcal{P}$ and cannot have any nontrivial quotients or subobjects in $\mathcal{Neg}$ by Lemma 5.0.89, so B or C is trivial. $\square$

Lemma 4.0.51. *Let N be a natural number that is a multiple of ℓ . Let A be an element of $D_c^b(\mathbb{A}_{\mathbb{Z}[1/N]}^1, \overline{\mathbb{Q}}_\ell)$ which is perverse when pulled back to $\mathbb{A}_{\mathbb{Q}}^1$.*

If A is a pure complex of weight w then for all but finitely many primes p , we have for all n , and for all nontrivial additive characters ψ of $\mathbb{F}_{p^n}$, $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A_{\mathbb{Q}} \otimes \mathcal{L}_\psi)$ is pure of weight w .

Proof. Proposition 3.0.46 states that $R\Psi_p A$ is isomorphic to $A_{\mathbb{F}_p}$ at all but finitely many primes p . We will prove the purity of $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A_{\mathbb{Q}} \otimes \mathcal{L}_\psi)$ at primes where $R\Psi_p A$ is isomorphic to $A_{\mathbb{F}_p}$. Because $A_{\mathbb{F}_p}$ is pure of weight w , and $\mathcal{L}_\psi$ is pure of weight 0, $R\Psi_p A \otimes \mathcal{L}_\psi$ is pure of weight w . Thus by Deligne's Weil II theorem [Deligne, 1980, Variant 6.2.3], $H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is mixed of weight $\leq w$. Furthermore $D(R\Psi_p A \otimes \mathcal{L}_\psi)$ is pure of weight $-w$, thus by Weil II $H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, D(R\Psi_p A \otimes \mathcal{L}_\psi))$ is mixed of weight $\leq -w$. We have

$$DH_c^0(\mathbb{A}_{\mathbb{F}_p}^1, D(R\Psi_p A \otimes \mathcal{L}_\psi)) = H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$$

so $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is mixed of weight $\geq w$.

By Lemma 3.0.43 the “forget supports” map $H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) \rightarrow H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is an isomorphism. As the left side is mixed of weight $\leq w$ and the right is mixed of weight $\geq w$, both are pure of weight w . $\square$

Definition 4.0.52. We say that $A \in \mathcal{Perv}/\mathcal{Neg}$ with good reduction outside S is pure of weight w if, for all $p \notin S$, we have: for all prime powers p^n , and all nontrivial additive characters ψ of $\mathbb{F}_{p^n}$, $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is pure of weight w .

Remark 4.0.53. We expect but are unable to prove to prove that Frob_p is pure at all the primes of good reduction and not just at these primes p where $R\Psi_p A$ agrees with the usual specialization and at which the usual specialization is pure. This is difficult already in the special case where A is a Galois representation $H^i(X, \overline{\mathbb{Q}}_\ell)$ for X a smooth proper algebraic variety, placed in degree 0 and at the point 0. Then we

would need to prove, not only that Frob_p is pure of weight i at primes of good reduction of X but, that Frob_p is pure of weight i at all unramified primes of $H^i(X, \overline{\mathbb{Q}}_\ell)$. This is a special case of Deligne's weight-monodromy conjecture [Deligne, 1970, pp.430], which is open. In Lemma 4.0.51 we avoided this difficulty by throwing away finitely many primes.

We have an explicit formula for the dimension of the G_S -representation associated to an object:

Lemma 4.0.54. *Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ with good reduction at p .*

The dimension of A viewed as a representation of G_S is equal to

$$\sum_{x \in \mathbb{A}^1(\overline{\mathbb{Q}})} (\text{drop}(\mathcal{H}^{-1}(A), x) + \dim(\mathcal{H}^0(A), x))$$

where $\text{drop}(\mathcal{H}^{-1}(A), x)$ is the drop in the rank of $\mathcal{H}^{-1}(A)$ at x and $\dim(\mathcal{H}^0(A), x)$ is the rank of the stalk of $\mathcal{H}^0(A)$ at x .

Proof. The dimension of A viewed as a representation of G_S is equal to the dimension of $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ because $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is a fiber functor.

We compute the dimension of $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ by computing the Euler characteristic $\chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$. By Lemma 3.0.43, the cohomology groups in degrees other than 0 vanish, so the dimension equals the Euler characteristic. To compute the Euler characteristic we will first compute

$$\chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) - \chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A)$$

By Lemma 3.0.31, each of these Euler characteristics is a sum of local terms minus the rank of each sheaf. We will show that most of these terms cancel and the difference is $\text{rank}(A)$. Indeed, as $\mathcal{L}_\psi$ is lisse of rank 1 on $\mathbb{A}_{\mathbb{F}_p}^1$, tensoring with it does not change

the local invariants anywhere on $\mathbb{A}_{\mathbb{F}_p}^1$, so we have

$$\text{rank}(R\Psi_p A \otimes \mathcal{L}_\psi) = \text{rank}(R\Psi_p A)$$

and for all $x \in \mathbb{A}_{\mathbb{F}_p}^1$

$$\text{drop}(\mathcal{H}^{-1}(R\Psi_p A \otimes \mathcal{L}_\psi), x) = \text{drop}(\mathcal{H}^{-1}(R\Psi_p A), x)$$

$$\dim(\mathcal{H}^0(R\Psi_p A \otimes \mathcal{L}_\psi), x) = \dim(\mathcal{H}^0(R\Psi_p A), x)$$

$$\text{Swan}(\mathcal{H}^{-1}(R\Psi_p A \otimes \mathcal{L}_\psi), x) = \text{Swan}(\mathcal{H}^{-1}(R\Psi_p A), x)$$

So all the local terms cancel except the Swan conductor at ∞ . To calculate that term, observe that, by Lemma 3.0.34, $R\Psi_p A$ is tame at ∞ and so has a Swan conductor of 0. Furthermore $\mathcal{L}_\psi$ is wildly ramified with break 1 at ∞ , so tensoring it with a tamely ramified sheaf $R\Psi_p A$ gives a sheaf that is wildly ramified with a break of 1. The Swan conductor is the break times the dimension, which is $\text{rank}(R\Psi_p A)$, so the Swan conductor at ∞ of $R\Psi_p A \otimes \mathcal{L}_\psi$ is $\text{rank}(R\Psi_p A)$. So we have

$$\chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) - \chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) = \text{rank}(R\Psi_p A) = \text{rank}(A)$$

$$\chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) = \chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) + \text{rank}(A)$$

By Lemma 3.0.35, the Euler characteristic of $R\Psi A$ is the Euler characteristic of A , so

$$\chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) = \chi(\mathbb{A}_{\mathbb{Q}}^1, A) + \text{rank}(A)$$

Now we use Lemma 3.0.31 to compute the Euler characteristic of A . The $-\text{rank}(A)$ term from Lemma 3.0.31 cancels with the $\text{rank}(A)$ term here, and the Swan conductors in Lemma 3.0.31 all vanish as A is a sheaf over a field of characteristic zero and hence

is tamely ramified. Removing those terms from the formula of Lemma 3.0.31, we get

$$\chi(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi) = \sum_{x \in \mathbb{A}^1(\overline{\mathbb{Q}})} (\text{drop}(\mathcal{H}^{-1}(A), x) + \dim(\mathcal{H}^0(A), x))$$

□

Using this, we may classify the one-dimensional objects. These will be important in Chapter 5 because we will pay special attention to abelian quotients of G_S , and irreducible representations of abelian quotients of G_S are one-dimensional. The classification will be in terms of skyscraper sheaves and Kummer sheaves. First we will prove some preliminary lemmas about these sheaves.

Definition 4.0.55. Let V be a representation of $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ over $\overline{\mathbb{Q}}_\ell$, and let r be a rational number. Define δ_r to be the skyscraper sheaf with stalk $\overline{\mathbb{Q}}_\ell$ at the point r of $\mathbb{A}_{\mathbb{Q}}^1$.

Thus $\delta_r \otimes V$, where we view V as a lisse sheaf on $\mathbb{A}^1$, is the unique sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ supported at the point r whose stalk at r is V .

Lemma 4.0.56. Let A be an object of $D_c^b(\mathbb{A}_{\mathbb{Q}}^1, \overline{\mathbb{Q}}_\ell)$. Let V be a Galois representation, r a rational number, and $[+r]$ the map $\mathbb{A}_{\mathbb{Q}}^1 \rightarrow \mathbb{A}_{\mathbb{Q}}^1$ that sends x to $x + r$. Then:

$$A * \delta_r \otimes V = A *! \delta_r \otimes V = ([+r]_* A) \otimes V$$

In particular in the special case $r = 0$ we have the identity $A * \delta_0 \otimes V = A \otimes V$, and in the special case $V = \mathbb{1}$, we have the identity $A * \delta_r = [+r]_* A$.

Proof. $A \boxtimes \delta_r \otimes V$ is supported on $\mathbb{A}^1 \times \{r\}$, which is isomorphic to $\mathbb{A}^1$. On that copy of $\mathbb{A}^1$, it is isomorphic to $A \otimes V$. Thus it is the pushforward of $A \otimes V$ from $\mathbb{A}^1 \times \{r\}$ to all of $\mathbb{A}^1 \times \mathbb{A}^1$. Then $A * \delta_r \otimes V$ is the pushforward along *sum*, which by the Leray spectral sequence is the pushforward of $A \otimes V$ along the composed map $[+r]_*$.

For compactly supported convolution, the same argument using the Leray spectral sequence with compact supports shows that $A * \delta_r \otimes V$ is $R[+r]_! A \otimes V$, but $[+r]$ is an isomorphism, hence proper, so $[+r]_!$ is the same as $[+r]_*$. $\square$

Lemma 4.0.57. *Let χ be the nontrivial character of μ_2 . Then the Kummer sheaf $\mathcal{L}_\chi[1]$ has good reduction at p for all $p \notin \{2, \ell\}$ but $R\Psi_2 \mathcal{L}_\chi[1] = 0$ as long as $\ell \neq 2$.*

Proof. The first statement follows easily from the criterion of Proposition 3.0.46 describing primes of good reduction. For the second statement, let π be the squaring map from $\mathbb{A}^1$ to $\mathbb{A}^1$. As π is finite, it commutes with formation of nearby cycles. Over $\mathbb{Z}[1/2\ell]$, $R\pi_* \overline{\mathbb{Q}}_\ell = \overline{\mathbb{Q}}_\ell \oplus \mathcal{L}_\chi$, but π is radicial in characteristic 2, hence $R\pi_* \overline{\mathbb{Q}}_\ell = \overline{\mathbb{Q}}_\ell$ in characteristic 2.

$$R\Psi_2 \overline{\mathbb{Q}}_\ell \oplus R\Psi_2 \mathcal{L}_\chi = R\Psi_2 R\pi_* \overline{\mathbb{Q}}_\ell = R\pi_* R\Psi_2 \overline{\mathbb{Q}}_\ell = R\pi_* \overline{\mathbb{Q}}_\ell = \overline{\mathbb{Q}}_\ell$$

As $R\Psi_2 \overline{\mathbb{Q}}_\ell = \overline{\mathbb{Q}}_\ell$ by local acyclicity, $R\Psi_2 \mathcal{L}_\chi$ vanishes. $\square$

Now we can show:

Lemma 4.0.58. *Let N be the product of all primes in S . Every one-dimensional representation of G_S corresponds to a perverse sheaf that is of one of the following two forms*

- $\delta_r \otimes V$ with V a 1-dimensional Galois representations $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow \mathbb{G}_m(\overline{\mathbb{Q}}_\ell)$ unramified outside S and with $r \in \mathbb{Z}[1/N] \subset \mathbb{Q}$.
- If $2 \in S$, $\mathcal{L}_\chi[1] * \delta_r \otimes V$ with $\mathcal{L}_\chi$ the Kummer sheaf corresponding to χ the non-trivial character of μ_2 , V a 1-dimensional Galois representation $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow \mathbb{G}_m(\overline{\mathbb{Q}}_\ell)$ unramified outside S , and $r \in \mathbb{Z}[1/N] \subset \mathbb{Q}$.

Proof. Let A be the property $\mathcal{P}$ representative of the one-dimensional representation. By Lemma 4.0.50, A is irreducible in the category of perverse sheaves over $\mathbb{Q}$, hence

a sum of Galois conjugate irreducible perverse sheaves over $\overline{\mathbb{Q}}$. Using the characterization of [Katz, 1995, Ch.2, pp. 11], each of these is either a middle extension sheaf placed in degree -1 or a punctual sheaf in degree 0 . Perverse sheaves of the two different kinds cannot be conjugate, so A is either a middle extension sheaf placed in degree -1 or a punctual sheaf in degree 0 .

Observe that the dimension can be expressed using Lemma 4.0.54 as a sum of local terms.

In the case where A is punctual in degree 0 , the local term at a point of $\mathbb{A}^1(\overline{\mathbb{Q}})$ is the rank at the point. Since the dimension is 1 , A is supported at a single point, of rank 1 . Because A is defined over $\mathbb{Q}$, this unique point must be a rational point. The stalk at the point r is some one-dimensional Galois representation V , and we may write

$$A = \delta_r \otimes V$$

This rational number r must be a p -adic integer for all p not in S , by the definition of good reduction, so $r \in \mathbb{Z}[1/N]$. The representation V must be unramified at p not in S because the $\text{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$ action on the stalk on $R\Psi_p A$ at the point r is just V restricted to $\text{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$. By the definition of good reduction this action of $\text{Gal}(\overline{\mathbb{Q}_p}/\mathbb{Q}_p)$ is unramified, so V is unramified at p .

In the case where A is an irreducible middle extension sheaf in degree -1 , the local term at a point of $\mathbb{A}^1(\overline{\mathbb{Q}})$ is the drop at that point. Thus there is exactly one non-lisse point on the sheaf, with a drop of 1 . As before, the point must be a rational point r .

As A is an irreducible middle extension sheaf in degree -1 , it is geometrically semisimple, hence a sum of geometrically irreducible middle extension sheaves on $\mathbb{A}_{\overline{\mathbb{Q}}}^1 - \{r\}$. These are middle extensions of irreducible representations of $\pi_1(\mathbb{A}_{\overline{\mathbb{Q}}}^1 - \{r\}) = \hat{\mathbb{Z}}(1)$. Because that group is abelian, the irreducible representations are 1 -dimensional characters. The trivial character corresponds to the trivial sheaf and the

other characters correspond to translated Kummer sheaves $r_*\mathcal{L}_\chi[1]$. Because A has property $\mathcal{P}$, no trivial sheaves can appear as summands of A , hence A is geometrically a sum of Kummer sheaves. Each Kummer sheaf has a drop of 1 so exactly one Kummer sheaf appears in the sum and thus A is geometrically a Kummer sheaf.

There is a Galois action on the set of isomorphism classes of Kummer sheaves, and as A is defined over $\mathbb{Q}$ it must be a fixed point. The only nontrivial Kummer sheaf fixed by the Galois action corresponds to the only nontrivial character fixed by the Galois action, which is the unique nontrivial quadratic character. Hence A is geometrically isomorphic to $\mathcal{L}_\chi(x-r)[1]$ for χ the nontrivial character of μ_2 . Therefore A is isomorphic to $L_\chi(x-r)[1]$ tensored with a one-dimensional Galois representation V . Then we have:

$$A = \mathcal{L}_\chi(x-r)[1] \otimes V = \mathcal{L}_\chi[1] * \delta_r \otimes V$$

by Lemma 4.0.56

By the definition of good reduction, all the singularities of A must be p -adic integers for $p \notin S$, so r is a p -adic integer for all $p \notin S$, and thus $r \in \mathbb{Z}[1/N]$.

Next we show that $2 \in S$ if $A = \mathcal{L}_\chi[1] * \delta_r \otimes V$ has good reduction away from S . First note that if $\ell = 2$, then $2 \in S$, so we may assume that $\ell \neq 2$. By Lemma 4.0.57, $R\Psi_2\mathcal{L}_\chi = 0$. After shifting $\mathcal{L}_\chi$ by 1, translating it by a 2-adic integer r , and tensoring with a Galois representation V , the nearby cycles $R\Psi_2(\mathcal{L}_\chi[1] * \delta_r \otimes V)$ will still vanish. This contradicts good reduction at 2 as then $\text{rank}(R\Psi_2 A) = 0$ but $\text{rank}(A) = 1$, so $2 \in S$.

Last we show V is unramified at p for all $p \notin S$. Observe that as $2 \in S$ and $\ell \in S$, $\mathcal{L}_\chi[1]$ has good reduction at all primes not in S by Lemma ???. Thus $(\mathcal{L}_\chi[1])^\vee$ has good reduction at all $p \notin S$ by Lemma 4.0.57, so $A * (\mathcal{L}_\chi[1])^\vee = \delta_r \otimes V$ has good reduction, thus by passing to the previous case V is unramified outside S .

□

We will find it very helpful to study the G_S -representation associated to a sheaf using the Fourier transform of the sheaf. This definition and lemma allow us to do that:

Definition 4.0.59. Let $j : \mathbb{A}_{\mathbb{F}_p}^1 - \{0\} \rightarrow \mathbb{A}_{\mathbb{F}_p}^1$ be the inclusion. Let ψ be the standard exponential character $\psi : \mathbb{Z}/p \rightarrow \mathbb{C}^\times$, $\psi(x) = e^{2\pi i \frac{x}{p}}$.

Let p_1 and p_2 be the two coordinate projections $\mathbb{A}_{\mathbb{F}_p}^2 \rightarrow \mathbb{A}_{\mathbb{F}_p}^1$ and let $\mathcal{L}_\psi(xy)$ be the Artin-Schrier sheaf $\mathcal{L}_\psi$ pulled back along the map $xy : \mathbb{A}_{\mathbb{F}_p}^2 \rightarrow \mathbb{A}_{\mathbb{F}_p}^1$.

Let FT_ψ be the ℓ -adic Fourier transform functor $FT_\psi(A) = Rp_{2!}(p_1^*A \otimes \mathcal{L}_\psi(xy))$.

Define FTS_p as the composition of functors $j^* \circ FT_\psi \circ R\Psi_p$ from perverse sheaves on $\mathbb{A}_{\mathbb{Q}}^1$ to complexes of sheaves on $\mathbb{A}_{\mathbb{F}_p}^1 - \{0\}$. (We will see shortly that for A a perverse sheaf good reduction at p , $FTS_p(A)$ is a sheaf in degree 0. This will be convenient, which is why we have not shifted the Fourier transform by one degree, as is usually done.

Lemma 4.0.60. *Let $\rho : G_S \rightarrow GL_n(\overline{\mathbb{Q}_\ell})$ be a representation, corresponding to a perverse sheaf A on $\mathbb{A}_{\mathbb{Q}}^1$. Let p be a prime not in S , then:*

1. $FTS_p(A)$ is a lisse sheaf of rank n concentrated in degree 0.
2. $FTS_p(A)$ is tamely ramified at 0 and has all breaks ≤ 1 at ∞ .
3. The geometric monodromy group of $FTS_p(A)$ is a subgroup of the image of ρ inside $GL_n(\overline{\mathbb{Q}_\ell})$.

Moreover, FTS_p is an exact symmetric monoidal functor from G_S -representations to lisse sheaves on $\mathbb{G}_{m, \mathbb{F}_p}$.

Proof. The lisseness and the facts about the local monodromies follow from Laumon's theory of the local Fourier transform, which describes how the local monodromies of the Fourier transform of a sheaf are determined by the local monodromies of the sheaf.

By Lemma 3.0.34, $R\Psi_p A$ has tame local monodromy at ∞ . It follows that $FT_\psi R\Psi_p A$ is lisse away from 0 and tamely ramified at 0 [Katz, 1990, Corollary 7.4.5]. We also have that $FT_\psi R\Psi_p A$ is ramified with all breaks ≤ 1 at ∞ [Katz, 1990, Theorem 7.4.1]. So $FTS_p(A) = j^* FT_\psi R\Psi_p A$ is lisse, tamely ramified at 0, and has all breaks ≤ 1 at ∞ . Katz proves these only for middle extension sheaves, but they are obvious for skyscraper sheaves, as the Fourier transform of a skyscraper sheaf is an Artin-Schreier sheaf. As any perverse sheaf with tame local monodromy at infinity is an iterated extension of middle sheaves with tame local monodromy at ∞ and skyscraper sheaves, and Fourier transform preserves distinguished triangles, these facts hold also for general perverse sheaves.

By proper base change, the stalk of $FT_\psi R\Psi_p A$ at a point λ is $H_c^*(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p \otimes \mathcal{L}_{\psi \circ \lambda})$. For $\lambda \neq 0$, $\psi \circ \lambda$ is a nontrivial additive character, and thus by Lemma 3.0.43 $H_c^*(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p \otimes \mathcal{L}_{\psi \circ \lambda})$ is a complex concentrated in degree 0. The stalk has rank n because $H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p \otimes \mathcal{L}_{\psi \circ \lambda})$ is a fiber functor and so its dimension is exactly equal to the Tannakian dimension, which is n . Hence $FTS_p(A) = j^* FT_\psi R\Psi_p A$ is concentrated in degree 0, and because it is lisse, it is lisse of rank n .

Next we will show that FTS_p is an exact tensor functor from the category of representations of G_S to the category of lisse sheaves on $\mathbb{G}_{m, \mathbb{F}_p}$. Afterwards we will deduce the statement about monodromy groups.

First we will show that FTS_p is well-defined on $\mathcal{Perv}/\mathcal{Neg}$, i.e. that $FTS_p(A) = 0$ if A is in $\mathcal{Neg}$. Such an A is geometrically constant, and the first step in the definition of nearby cycles is pullback to an algebraically closed field, so we may assume that A is constant. Then $R\Psi_p A$ is constant by the local acyclicity of smooth morphisms. The Fourier transform of a constant sheaf is just a skyscraper sheaf supported at the point 0, which becomes the zero sheaf when pulled back along j .

Tensor product of G_S -reps corresponds to (compactly-supported) convolution of sheaves A , and

$$FTS(A *_! B)$$

(by definition)

$$= j^* FT_\psi R\Psi_p(A *_! B)$$

(by Lemma 3.0.36)

$$= j^* FT_\psi((R\Psi_p A) *_! (R\Psi_p B))$$

(because of the general fact that $FT_\psi(C *_! D) = FT_\psi(C) \otimes FT_\psi(D)$, [Brylinski, 1986, Corollary 9.6]. Their FT_ψ differs from ours by a shift of 1, the dimension, so their identity also differs from ours by a shift.)

$$= j^*((FT_\psi R\Psi_p A) \otimes (FT_\psi R\Psi_p B))$$

(because pullback commutes with tensor product)

$$= (j^* FT_\psi R\Psi_p A) \otimes (j^* FT_\psi R\Psi_p B)$$

(by definition)

$$= FTS(A) \otimes FTS(B)$$

with the first and last identity by definition, the second by Lemma 3.0.36, and the third by [Brylinski, 1986, Corollary 9.6]. Thus FTS is a tensor functor - the diagrams commute by the same argument with more sheaves. Note that FTS preserves distinguished triangles as it is a composition of nearby cycles, Fourier transform, and pullback, and nearby cycles and Fourier transform are compositions of pullbacks, pushforwards, and tensor products, and pullback, pushforward, and tensor product preserves distinguished triangles. Because it sends perverse sheaves to complexes sup-

ported in degree 0 and preserves distinguished triangles, it is an exact functor from perverse sheaves to lisse sheaves.

Finally, we need to prove that the monodromy group is contained in $\rho(G_S)$. Let H be the Tannakian fundamental group of the category of lisse sheaves on $\mathbb{G}_{m, \overline{\mathbb{F}}_p}$. This is equivalent to the category of ℓ -adic representations of $\pi_1(\mathbb{A}_{\overline{\mathbb{F}}_p}^1 - \{0\})$. By Tannaka-Krein duality, this exact tensor functor defines a homomorphism $f : H \rightarrow G_S$, up to conjugation, such that $FTS_p(A)$ is the lisse sheaf corresponding to the representation $\rho \circ f$ of H . By definition, the monodromy group of a lisse sheaf is the Zariski closure of the image of π_1 in the associated representation. This is the same as the image of H in that representation. Hence the monodromy group of $FTS_p(A)$ is the image of $\rho \circ f$, which is a subgroup of the image of ρ . $\square$

Using these facts about the Fourier transform, we can prove three results that help us compute the determinant of a G_S representation in terms of the sheaf.

There are a few intermediate steps. Let $\det^\tau(A)$ be the determinant of A in the Tannakian category (to distinguish from the determinant of A on the open subset where it is lisse). Then $\det^\tau(A)$ is always a one-dimensional object. Our strategy is to compute some invariant of $FTS_p(\det^\tau(A))$, then compute the same invariant of FTS_p applied to the objects described in Lemma 4.0.58, and compare to identify $\det^\tau(A)$. First observe:

Lemma 4.0.61.

$$FTS_p(\det^\tau(A)) = \det(FTS_p(A))$$

Proof. By Lemma 4.0.60, FTS_p is a tensor functor, hence it commutes with Tannakian category constructions such as the determinant. $\square$

Lemma 4.0.62. *Let χ be the nontrivial quadratic character. Let W be the one-dimensional ℓ -adic $\text{Gal}(\overline{\mathbb{F}}_p|\mathbb{F}_p)$ representation on which Frob_p acts by minus the Gauss sum $G(\chi, \psi)$.*

Let V be a one-dimensional ℓ -adic Galois representation and r a rational number.

If p is a prime of good reduction of $\delta_r \otimes V$:

$$FTS_p(\delta_r \otimes V) = V \otimes \mathcal{L}_\psi(rx)$$

If p is a prime of good reduction of $\delta_V(r) * \mathcal{L}_\chi[1]$:

$$FTS_p(\delta_r \otimes V * \mathcal{L}_\chi[1]) = V \otimes \mathcal{L}_\psi(rx) \otimes FTS_p \mathcal{L}_\chi[1] = V \otimes W \otimes \mathcal{L}_\psi(rx) \otimes \mathcal{L}_\chi$$

Proof. By Lemma 4.0.58, V is unramified at p , r is a p -adic integer, and in the second case $p \neq 2$.

First observe that $R\Psi_p \delta_r \otimes V = \delta_r \otimes V$. This is because we may form a sheaf on $\mathbb{A}_{\mathbb{Z}_p}^1$ that is the pushforward of the constant sheaf V from $\text{Spec } \mathbb{Z}_p$ to $\mathbb{A}_{\mathbb{Z}_p}^1$ along the map that sends x to r . The general and special fibers of this sheaf are both equal to $\delta_r \otimes V$, and it is locally acyclic because it is the pushforward of a locally acyclic constant sheaf along a finite morphism. Hence $R\Psi_p$ of the generic fiber is the special fiber. It follows that

$$FTS_p \delta_r \otimes V = j^* FT_\psi(R\Psi_p \delta_r \otimes V) = j^* FT_\psi(\delta_r \otimes V) = V \otimes j^* FT_\psi(\delta_r) = V \otimes \mathcal{L}_\psi(rx)$$

Second note that $R\Psi_p \mathcal{L}_\chi[1] = \mathcal{L}_\chi[1]$, as $\mathcal{L}_\chi[1]$ is locally acyclic away from 2 by [Deligne and Katz, 1973, XIII Lemma 2.1.11], so the nearby cycles of the generic fiber is the special fiber. Hence

$$FTS_p \mathcal{L}_\chi[1] = j^* FT_\psi \mathcal{L}_\chi[1] = \mathcal{L}_\chi \otimes W$$

as by Lemma 4.0.60 it is a lisse sheaf of rank 1 on $\mathbb{G}_m$, hence determined by its trace of Frobenius at each $\mathbb{F}_{p^n}$ -point, and the trace of Frob_{p^n} at $x \in \mathbb{F}_{p^n}$ is

$$-\sum_{y \in \mathbb{F}_{p^n}} \chi(y) \psi \left(\text{tr}_{\mathbb{F}_{p^n}}^{\mathbb{F}_p} xy \right) = -\chi(x) \sum_{y \in \mathbb{F}_{p^n}} \chi(y) \psi \left(\text{tr}_{\mathbb{F}_{p^n}}^{\mathbb{F}_p} y \right) = \chi(x) (-G(\chi, \psi))^n$$

which is equal to the eigenvalue of Frob_{p^n} acting on $W \otimes \mathcal{L}_\chi$ at x .

Finally by Lemma 3.0.46 FTS_p is a tensor functor hence

$$FTS_p(\delta_r \otimes V * \mathcal{L}_\chi[1]) = FTS_p(\delta_r \otimes V) \otimes FTS_p(\mathcal{L}_\chi[1]) = V \otimes \mathcal{L}_\psi(rx) \otimes \mathcal{L}_\chi \otimes W$$

□

Lemma 4.0.63. *Let S be a smooth scheme over $\mathbb{Z}_p$. Let A be an object of $D_c^b(\mathbb{A}_{S_{\mathbb{Q}_p}}^1, \overline{\mathbb{Q}_\ell})$ that is extensible. Let U be an open subset of S and L a lisse sheaf on U satisfying the conditions from the definition of extensibility. Then $R\Psi_p A$, restricted to $U_{\overline{\mathbb{F}_p}}$, is isomorphic to L restricted to $U_{\overline{\mathbb{F}_p}}$.*

Proof. The proof is by a sequence of isomorphisms:

$$(R\Psi_p A)_{U_{\overline{\mathbb{F}_p}}}$$

(because nearby cycles commute with restriction to an open subset)

$$= R\Psi_p A_{U_{\mathbb{Q}_p}}$$

(by the definition of extensibility)

$$= R\Psi_p L_{U_{\mathbb{Q}_p}}$$

(by local acyclicity of smooth morphisms)

$$= L_{U_{\mathbb{F}_p}}$$

□

Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ with good reduction at a prime p . The next lemma allows us to compute the local monodromy representation of $R\Psi_p A$ at ∞ , which we will need to compute properties of $FTS_p(A)$. Recall (Lemma 3.0.34) that $R\Psi_p A$ is tamely ramified at ∞ . Furthermore, recall that $R\Psi_p A$ has a natural action of $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$, so this local monodromy representation carries an action of $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$. That is, it is a representation of $\prod_{l \neq p} \mathbb{Z}_l(1) \rtimes \text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$.

Lemma 4.0.64. *Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ with good reduction at a prime p . The local monodromy representation of $R\Psi_p A$ at ∞ , as a representation of $\prod_{l \neq p} \mathbb{Z}_l \rtimes \text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$, is obtained from the decomposition group representation of A at ∞ by restricting from the decomposition group $\text{Gal}(\overline{\mathbb{Q}((x^{-1}))}|\mathbb{Q}((x^{-1}))) = \prod_l \mathbb{Z}_l(1) \rtimes \text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ to the subgroup $\prod_{l \neq p} \mathbb{Z}_l(1) \rtimes \text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$.*

Furthermore the $\mathbb{Z}_p(1)$ factor acts trivially in the decomposition group representation of A at ∞ .

Proof. By Lemma 3.0.30, there is an open set U of $\mathbb{A}_{\mathbb{Z}_p}^1$ whose complement is bounded and a lisse sheaf L on U such that the restriction of $L[1]$ to $U_{\mathbb{Q}_p}$ is isomorphic to the restriction of A to $U_{\mathbb{Q}_p}$. Hence the decomposition group representation of A at ∞ is the pullback of L to $\text{Spec}(\mathbb{Q}_p((x^{-1})))$. By Lemma 4.0.63, the restriction of $R\Psi_p A$ to $U_{\mathbb{F}_p}$ is isomorphic to the restriction of L to $U_{\mathbb{F}_p}$, so the local monodromy representation of $R\Psi_p A$ at ∞ is the pullback of L to $\text{Spec}(\mathbb{F}_p((x^{-1})))$.

Because the complement of U is bounded, it does not intersect the divisor at ∞ . Hence the map $\text{Spec}(\mathbb{Z}_p((x^{-1}))) \rightarrow \text{Spec}(\mathbb{Z}[x]) = \mathbb{A}_{\mathbb{Z}_p}^1$ factors through U . As L remains lisse when pulled back to $\text{Spec}(\mathbb{Z}_p((x^{-1})))$, it corresponds to a representation of $\pi_1(\text{Spec}(\mathbb{Z}_p((x^{-1}))))$. The pullback of L to $\text{Spec}(\mathbb{Q}_p((x^{-1})))$ corre-

sponds to the restriction of this representation to $\text{Gal}(\overline{\mathbb{Q}_p((x^{-1}))}|\mathbb{Q}_p((x^{-1})))$. Similarly, the pullback of L to $\text{Spec}(\mathbb{F}_p((x^{-1})))$ is the restriction of this representation to $\text{Gal}(\overline{\mathbb{F}_p((x^{-1}))}|\mathbb{F}_p((x^{-1})))$.

By [Katz, 1980, pp. 121], $\pi_1(\text{Spec}(\mathbb{Z}_p((x^{-1})))) = \prod_{l \neq p} \mathbb{Z}_l(1) \rtimes \pi_1(\text{Spec} \mathbb{Z}_p)$.

Taking a representation of $\prod_{l \neq p} \mathbb{Z}_l(1) \rtimes \pi_1(\text{Spec} \mathbb{Z}_p)$, pulling back to $\prod_l \mathbb{Z}_l(1) \rtimes \text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$ by the homomorphism that sends the $\mathbb{Z}_p$ factor and the inertia group in $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$ to 0, and then restricting to $\prod_{l \neq p} \mathbb{Z}_l(1) \rtimes \text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$ is the same as pulling back directly to $\prod_{l \neq p} \mathbb{Z}_l(1) \rtimes \text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$. This is precisely the desired relationship between the decomposition group representation of A and the local monodromy representation of $R\Psi_p A$.

Furthermore as the decomposition group representation of A at ∞ is obtained by pullback from $\prod_{l \neq p} \mathbb{Z}_l(1) \rtimes \pi_1(\text{Spec} \mathbb{Z}_p)$ to $\prod_l \mathbb{Z}_l(1) \rtimes \text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$, so the $\mathbb{Z}_p(1)$ factor, which is in the kernel of that homomorphism, acts trivially.

□

Lemma 4.0.65. *Let p be a prime. Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ that is a middle extension sheaf placed in degree -1 , with good reduction at p , and whose finite singularities $\alpha_1, \dots, \alpha_n$, viewed as elements of $\overline{\mathbb{Z}_p}$, are distinct modulo the maximal ideal (i.e. their differences are all p -adic units).*

Then $R\Psi_p A$ is a middle extension sheaf placed in degree -1 whose finite singularities are the reductions of α_i mod the maximal ideal of $\overline{\mathbb{Z}_p}$, with the drop of $R\Psi_p A$ at the reduction of α_i equal to the drop of A at α_i . Moreover $R\Psi_p A$ is tamely ramified at each singularity.

Proof. By Lemma 3.0.30, there is an open set U on $\mathbb{A}_{\mathbb{Z}_p}^1$ and a lisse sheaf L on U with $L_{U_{\mathbb{Q}_p}}[1] \cong A_{U_{\mathbb{Q}_p}}$. Furthermore we may take U to be the complement of the closure of the singular locus of A , which is a smooth divisor because the reductions of the α_i are distinct. Around each component of this divisor, by Abhyankhar's lemma L has tame ramification, so its local monodromy representation is the same when restricted

to $\mathbb{A}_{\mathbb{Q}_p}^1$ or $\mathbb{A}_{\mathbb{F}_p}^1$. The first one is the local monodromy of A , and the second is the local monodromy of $R\Psi_p A$ by Lemma 4.0.63. Moreover the drop of $R\Psi_p A$ at a point is at least the drop of its local monodromy representation, which is equal to the drop of the local monodromy representation of A at the corresponding point, which is equal to the drop of A at that point because A is middle extension. Hence the drop of $R\Psi_p A$ at the reduction of $\alpha_i \bmod p$ is at least the drop of A at α_i .

By Lemma 3.0.35, A and $R\Psi_p A$ have the same Euler characteristic. Observing the formula from Lemma 3.0.31, each term is either equal for A and $R\Psi_p A$, or the $R\Psi_p A$ contribution is higher:

- $\text{rank}(A) = \text{rank}(R\Psi_p A)$ because A has good reduction.
- The rank of $\mathcal{H}^0(A)$ vanishes at each point because A is a middle extension sheaf, and the rank of $\mathcal{H}^0(R\Psi_p A)$ is nonnegative.
- We have just shown that the drop of $R\Psi_p A$ at the reduction of each finite singular point is at least the drop of A at that finite singular point.
- The Swan conductor of A vanishes at each point because it is a sheaf in characteristic 0, and the Swan conductor of $R\Psi_p A$ is nonnegative.

Because the Euler characteristics of A and $R\Psi_p A$ are equal, this implies that each individual term is equal.

- The rank of $\mathcal{H}^0(R\Psi_p A)$ vanishes at each point, so it is concentrated in degree -1 .
- The drop of $R\Psi_p A$ at the reduction of each singular point equals the drop of $R\Psi_p A$ at that singular point, which we showed was equal to the drop of the local monodromy representation of $R\Psi_p A$ at that point, so $R\Psi_p A$ is a middle extension sheaf in degree -1 .

- The Swan conductor of $R\Psi_p A$ vanishes at each point, so it is tamely ramified.

These are all the desired statements. □

Lemma 4.0.66. *Write $\det^\tau(A)$ as either $\delta_r \otimes V$ or $\delta_r \otimes V * \mathcal{L}_\chi[1]$, per Lemma 4.0.58. The second case occurs if and only if the usual determinant of the local monodromy representation of A at ∞ is nontrivial.*

Proof. By Laumon's local Fourier transform, the local monodromy at 0 of $FTS_p(A)$ can be obtained by applying a particular functor to the local monodromy at ∞ of $R\Psi_p A$. For p a prime of good reduction, $R\Psi_p A$ is tamely ramified at ∞ (Lemma 3.0.30). The functor sends tamely ramified representations to their own dual representation [Katz, 1990, Theorem 7.4.4(2,3)]. So the determinant of $FTS_p(A)$ at 0 has the same local monodromy representation as the dual of the determinant of $R\Psi_p A$ at ∞ . (Again, that only handles middle extension sheaves in degree -1 , but the Fourier transforms of skyscraper sheaves are Artin-Schreier sheave, which have trivial local monodromy at 0 and thus do not contribute to the determinant.) Hence by Lemma 4.0.64, the determinant of $FTS_p(A)$ at 0 is also the same as the dual of the determinant of the local monodromy representation of A at ∞ .

From the calculations of $FTS_p(\delta_r \otimes V)$ and $FTS_p(\delta_r \otimes V * \mathcal{L}_\chi[1])$ in Lemma 4.0.62 we see that $FTS_p(\delta_r \otimes V)$ always has trivial local monodromy at 0 and $FTS_p(\delta_r \otimes V * \mathcal{L}_\chi[1])$ one always has nontrivial local monodromy at 0. Hence if the determinant of the local monodromy representation of A at ∞ is trivial, $\det^\tau(A)$ has the first form, and if nontrivial, the second. □

Lemma 4.0.67. *Write $\det^\tau(A)$ as either $\delta_r \otimes V$ or $\delta_r \otimes V * \mathcal{L}_\chi[1]$, as per Lemma 4.0.58. Then (compare to Lemma 4.0.54):*

$$r = \sum_{x \in \mathbb{A}^1(\mathbb{Q})} x \cdot (\text{drop}(\mathcal{H}^{-1}(A), x) + \dim(\mathcal{H}^0(A), x))$$

i.e. r is the sum, with suitable multiplicity, of the singular points.

Proof. Let

$$s = \sum_{x \in \mathbb{A}^1(\mathbb{Q})} x \cdot (\text{drop}(\mathcal{H}^{-1}(A), x) + \dim(\mathcal{H}^0(A), x))$$

We will show that, for sufficiently large p , the local monodromy representation of $\det(FTS_p(A))$ at ∞ restricted to the wild inertia group is isomorphic to the local monodromy representation of $\mathcal{L}_\psi(sx)$ at ∞ restricted to the wild inertia group. On the other hand by Lemma 4.0.62 the wild inertia representation of $FTS_p(\delta_r \otimes V)$ and $FTS_p(\delta_r \otimes V * \mathcal{L}_\chi)$ at ∞ is the same as that of $\mathcal{L}_\psi(rx)$, and these are only isomorphic if $r \equiv s \pmod{p}$. Since we have this congruence for all sufficiently large p , we obtain $r = s$.

The function s is additive in extensions of sheaves, so we may assume A is irreducible, hence either supported only in degree 0 or a middle extension sheaf in degree -1 . In the first case, we do not need to use local Fourier transform - geometrically A is a sum of skyscraper sheaves supported at points $\alpha_1, \dots, \alpha_n$, so $R\Psi_p A$ is a sum of skyscraper sheaves supported at the reductions of the α_i modulo p , and its Fourier transform is geometrically a sum of Artin-Schreier sheaves $\mathcal{L}_\psi(\alpha_i x)$, so its determinant is $\mathcal{L}_\psi((\sum_{i=1}^n \alpha_i) x)$.

In the case when A is a middle extension sheaf in degree -1 , first observe that by Lemma 4.0.65 for all p sufficiently large, $R\Psi_p A$ is supported in degree -1 , and the singularities of $R\Psi_p A$ are reductions of the singularities of A , tamely ramified and with the same drop. Hence the sum of the singularities α_i of p with multiplicity their drop is the reduction of $s \pmod{p}$. For these p , Laumon's formula writes the local monodromy at ∞ as a sum of contributions from the singular points α_i . Each one is a tame representation tensored with $\mathcal{L}_\psi(\alpha_i x)$. The rank of the tame representation is exactly the drop. Thus the wild inertia representation of $FTS(R\Psi_p A)$ at ∞ it is a

sum of $\mathcal{L}_\psi(\alpha_i x)$ with multiplicity the drop at α_i . The determinant of this wild inertia representation is $\mathcal{L}_\psi(sx)$, as desired. $\square$

Lemma 4.0.68. *Let A and B be two perverse sheaves that are isomorphic when pulled back to the local field $\mathbb{Q}((x^{-1}))$.*

Write $\det^\tau(A)$ as either $\delta_{r_a} \otimes V_A$ or $\delta_{r_a} \otimes V_A * \mathcal{L}_\chi[1]$, per Lemma 4.0.58.

Write $\det^\tau(B)$ as either $\delta_{r_b} \otimes V_B$ or $\delta_{r_b} \otimes V_B * \mathcal{L}_\chi[1]$, per Lemma 4.0.58.

Note that by assumption and Lemma 4.0.66, if $\det^\tau(A)$ has the first form then $\det^\tau(B)$ does as well, and if $\det^\tau(A)$ has the second form then $\det^\tau(B)$ does as well.

Then

$$V_A \otimes \det H_c^1(\mathbb{A}_{\mathbb{Q}}^1, A) \otimes \det H_c^0(\mathbb{A}_{\mathbb{Q}}^1, A)^{-1} = V_B \otimes \det H_c^1(\mathbb{A}_{\mathbb{Q}}^1, B) \otimes \det H_c^0(\mathbb{A}_{\mathbb{Q}}^1, B)^{-1}$$

Proof. To show these two Galois representations are equal, it is sufficient to show the actions of Frobenius are equal at all sufficiently large p . We will show it at all primes of good reduction. Fix such a prime p .

Consider the exact sequence of [Laumon, 1987, second equation on pp.160] for $K' =$ the perverse sheaf $FT_\psi R\Psi_p A[1]$ (and $K = R\Psi_p A$, since Laumon defines $K' = FT_\psi K$, which in our notation is $FT_\psi K[1]$). The first and fourth terms are the stalks of $FT_\psi R\Psi_p A$ in degree 0 and 1, which by proper base change are $H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A)$ and $H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A)$. The second term is the local monodromy representation of $FT_\psi(R\Psi_p A)$ at 0. The third term is, according to [Laumon, 1987, Proposition 2.3.2.1], equal to $R^{-1}\Phi_{\eta_0}(p^*\alpha_! R\Psi_p A \otimes \mathcal{L}_\psi(xx'))_{\infty,0}$, which according to [Laumon, 1987, Lemma 2.4.2.1] is what in [Laumon, 1987, Definition 2.4.2.3] is defined to be $\mathcal{F}_\psi^{\infty,0'}(R\Psi_p A)_{\eta_\infty}$. The precise definition is not important - what matters for our purposes is that it comes from applying some functor to the local monodromy representation of $R\Psi_p A$ at ∞ .

$$0 \rightarrow H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) \rightarrow FT_\psi(R\Psi_p A)_{\eta_0} \rightarrow \mathcal{F}_\psi^{\infty, 0'}(R\Psi_p A)_{\eta_\infty} \rightarrow H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) \rightarrow 0$$

As $R\Psi_p A$ has an action of $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$, this whole exact sequence has an action of $\text{Gal}(\overline{\mathbb{Q}_p}|\mathbb{Q}_p)$. So we obtain

$$\det \left(\mathcal{F}_\psi^{\infty, 0'}(R\Psi_p A)_{\eta_\infty} \right) = \det (FT_\psi(R\Psi_p A)_{\eta_0}) \otimes \det \left(H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) \right) \otimes \det \left(H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) \right)^{-1}$$

By Lemma 4.0.61,

$$\det (FT_\psi(R\Psi_p A)_{\eta_0}) = FTS_p(\det^\tau(A))(0)$$

Assume $\det^\tau(A)$ has the form $\delta_{r_A} \otimes V_A$ (and $\det^\tau(B)$ has a similar form). Then by Lemma 4.0.62, $FTS_p(\det^\tau(A))(0)$ is precisely V_A .

By Lemma 3.0.34 and proper base change,

$$H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A) = H_c^i(\mathbb{A}_{\mathbb{Q}}^1, A)$$

Hence

$$\det \left(\mathcal{F}_\psi^{\infty, 0'}(R\Psi_p A)_{\eta_\infty} \right) = V_A \otimes \det H_c^1(\mathbb{A}_{\mathbb{Q}}^1, A) \otimes \det H_c^0(\mathbb{A}_{\mathbb{Q}}^1, A)^{-1}$$

By Lemma 4.0.64 and the assumption on A and B , $(R\Psi_p A)_{\eta_\infty}$ is isomorphic to $(R\Psi_p B)_{\eta_\infty}$. Thus

$$V_A \otimes \det H_c^1(\mathbb{A}_{\mathbb{Q}}^1, A) \otimes \det H_c^0(\mathbb{A}_{\mathbb{Q}}^1, A)^{-1} = V_B \otimes \det H_c^1(\mathbb{A}_{\mathbb{Q}}^1, B) \otimes \det H_c^0(\mathbb{A}_{\mathbb{Q}}^1, B)^{-1}$$

If $\det^\tau(A)$ has the form $\delta_{r_A} \otimes V_A * \mathcal{L}_\chi[1]$ (and $\det^\tau(B)$ has a similar form). Then the only thing that changes is that $FTS_p(\det^\tau(A))(0)$ is now $V_A \otimes W \otimes \mathcal{L}_\chi$, and the same form $FTS_p(\det^\tau(B))(0)$. We simply cancel the $W \otimes \mathcal{L}_\chi$ from both sides and obtain the same equation. $\square$

We can write the Galois group of the maximal extension of $\mathbb{Q}$ unramified outside S as a quotient of G_S . In fact we can write a slightly larger group as a quotient of G_S :

Definition 4.0.69. Let H_S be the Tannakian fundamental group of the category of continuous representations $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow GL_n(\overline{\mathbb{Q}}_\ell)$ unramified outside S .

Let $g_S : G_S \rightarrow H_S$ be the natural homomorphism defined using Tannaka-Krein duality on the exact tensor functor that sends a representation V to $\delta_0 \otimes V$.

Let G_S^{geom} be the kernel of g_S .

Lemma 4.0.70. *The homomorphism g_S is surjective.*

Proof. By Tannaka-Krein duality, it is sufficient to show that the functor Inc that sends V to $\delta_0 \otimes V$ is fully faithful, and that a subobject of $\text{Inc}(V)$ is Inc of a subobject of V . [Deligne and Milne, 1982, Proposition 2.21].

To show that Inc is fully faithful, we need to check that the group of homomorphisms between two sheaves supported at the point 0 is isomorphic to the group of homomorphisms between their stalks. This follows from the definitions.

To show that subobjects of the image are in the image, first note that by Lemma 3.0.21, if A and B are the property $\mathcal{P}$ representatives of two equivalence classes with A a subobject of B in $\mathcal{CC}_S$, then A maps to B . The kernel of this map is zero in $\mathcal{Perv}/\mathcal{Neg}$, hence an element of $\mathcal{Neg}$, but A has no negligible subobjects so the kernel vanishes. Thus A is a subobject of B in $\mathcal{Perv}$. Geometrically $\delta_0 \otimes V$ is a sum of copies of δ_0 , an irreducible perverse sheaf, hence any subobject is a sum of copies of δ_0 and thus of the form $\delta_W(0)$ for some representation W . The kernel of the map

$\delta_0 \otimes W \rightarrow \delta_0 \otimes V$ is the kernel of the map $W \rightarrow V$, and hence if $\delta_0 \otimes W$ is a subobject of $\delta_0 \otimes V$ then W is a subobject of V . $\square$

So H_S is the quotient of G_S by G_S^{geom} .

We will sometimes need to work with perverse sheaves of geometric origin. We will give an inductive argument that sheaves of geometric origin in the sense of [Beilinson et al., 1982] also satisfy another condition, that we refer to as being of quasi-geometric origin. First note the following lemma that allows us to upgrade an apparently weaker geometric-origin-type condition to a stronger one:

Lemma 4.0.71. *Let X be a scheme of finite type over a field k of characteristic 0 and let K be an irreducible perverse sheaf on X . Suppose K is a Jordan-Holder factor of a perverse homology sheaf of $R\pi_! \overline{\mathbb{Q}}_\ell$ for a separated map of finite type $\pi : Y \rightarrow X$. Then K is a Jordan-Holder factor of a perverse homology sheaf of $R\pi'_* \overline{\mathbb{Q}}_\ell$ for a proper map $\pi' : Y' \rightarrow X$ from a smooth scheme Y' .*

Proof. We prove this by induction on the dimension. Assume the statement holds for all Y of dimension $< d$.

First we prove it when Y is a smooth scheme over k of dimension d . Take a relative compactification of π , $\bar{\pi} : \bar{Y} \rightarrow X$ and then take a resolution of singularities $\tilde{Y}$ of $\bar{Y}$, possible over k by [Bierstone and Milman, 1991, Theorem and discussion on pp.14]. Then $\tilde{Y}$ is a relative compactification of Y over X . Let $Z = \tilde{Y} - Y$. There is an excision long exact sequence relating the compactly supported cohomology relative to X of Y , $\tilde{Y}$, and Z , so every Jordan-Holder factor of the cohomology of X must be a Jordan-Holder factor of the cohomology of $\tilde{Y}$ or Z . In the first case we may take $Y' = \tilde{Y}$, in the second case use the induction hypothesis.

For a general separated Y of finite type over k of dimension $\leq d$, let U be the smooth locus of Y and Z its complement. There is an excision long exact sequence relating the compactly supported cohomology relative to X of Y , Z , and U . Any

Jordan-Holder factor of the relative compactly supported cohomology of Y must be a factor of the cohomology of U or Z . If it is a factor of the cohomology of U , as U smooth of dimension d , we reduce to the previous case. If it is a factor of Z , as Z is separated, of finite type, and of dimension $< d$, we apply the induction hypothesis.

Then it holds for all Y of dimension $\leq d$, verifying the induction step. $\square$

Definition 4.0.72. A perverse sheaf K is of quasisgeometric origin if every Jordan-Holder factor of K is also a Tate twist of a Jordan-Holder factor of $R\pi_*\overline{\mathbb{Q}}_\ell$ for a proper map $\pi : Y \rightarrow X$ from a smooth scheme Y .

An element of $D_c^b(X, \overline{\mathbb{Q}}_\ell)$ is of quasi-geometric origin if every perverse homology sheaf is of geometric origin.

Proposition 4.0.73. *Let k be a field of characteristic 0. Let X be a scheme of finite type over k and let K be a complex of geometric origin on X .*

1. *The Verdier dual of X is of geometric origin.*
2. *Let $f : X \rightarrow Z$ be a separated morphism of finite type to a scheme Z of finite type. Then $Rf_!K$ is of geometric origin.*
3. *Let $f : Z \rightarrow X$ be a separated morphism of finite type. Then f^*K is of geometric origin.*
4. *Let Z be a scheme of finite type over k and let L be a complex of geometric origin on Y . Then $K \boxtimes L$ is of geometric origin.*

Proof. For each statement we may reduce to the case of an irreducible perverse sheaf: An element $K \in D_c^b(X, \overline{\mathbb{Q}}_\ell)$ of quasi-geometric origin has a filtration whose associated graded objects $K_1, \dots, K_n$ are irreducible perverse sheaves of quasi-geometric origin, so by applying $Rf_!$ and taking the spectral sequence of a filtered complex we obtain a spectral sequence computing $Rf_!K$ from $Rf_!K_1, \dots, Rf_!K_n$, so any Jordan-Holder

factor of $Rf_!K$ must be a Jordan-Holder factor of $Rf_!K_i$ for some i . The same argument works for D , f^* , or $\boxtimes L$.

Thus, after taking a Tate twist, we may assume that K is a Jordan-Holder factor of $R\pi_*\overline{\mathbb{Q}}_\ell$ along a proper map π from a smooth scheme Y .

For part 1, because D sends perverse sheaves to perverse sheaves and acts as an exact contravariant functor on perverse sheaves, it follows that DK is a Jordan-Holder factor of $DR\pi_*\overline{\mathbb{Q}}_\ell = R\pi_!D\overline{\mathbb{Q}}_\ell = R\pi_*D\overline{\mathbb{Q}}_\ell = R\pi_*\overline{\mathbb{Q}}_\ell(d)[2d]$, hence is of geometric origin.

For part 4, we may likewise assume that L is a Jordan-Holder factor of $R\pi'_*\overline{\mathbb{Q}}_\ell$. Then because $\boxtimes$ is an exact functor from pairs of perverse sheaves to perverse sheaves, $K \boxtimes L$ is a Jordan-Holder factor of $R(\pi \times \pi')_*\overline{\mathbb{Q}}_\ell$, hence of geometric origin.

Parts 2 and 3 are more subtle. Take a filtration of $R\pi_*\overline{\mathbb{Q}}_\ell$ whose associated graded objects are irreducible perverse sheaves. By the decomposition theorem [Beilinson et al., 1982, Theorem 6.2.5], this filtration splits when pulled back to the algebraic closure of k . We obtain a filtration of $Rf_!R\pi_*\overline{\mathbb{Q}}_\ell$ or $f^*R\pi_*\overline{\mathbb{Q}}_\ell$ whose associated graded objects are Rf^* or $Rf_!$ of the irreducible perverse sheaves. Because the previous filtration splits when pulled back to $\overline{k}$, this filtration splits when pulled back to $\overline{k}$ as well, hence the spectral sequence of the filtration degenerates on the second page, thus every Jordan-Holder factor of the associated graded objects remains a Jordan-Holder factor of the whole complex.

For part 2, as $Rf_!R\pi_*\overline{\mathbb{Q}}_\ell = R(f \circ \pi)_!\overline{\mathbb{Q}}_\ell$, every Jordan-Holder factor of $Rf_!K$ is a Jordan-Holder factor of the compactly supported pushforward of the constant sheaf along a separated morphism of finite type.

For part 3, as $f^*R\pi_!\overline{\mathbb{Q}}_\ell$ is the compactly supported pushforward of $\overline{\mathbb{Q}}_\ell$ along the base change of π , every Jordan-Holder factor of f^*K is a Jordan-Holder factor of the compactly supported pushforward of the constant sheaf along a separated morphism of finite type.

In either case we conclude using Lemma 4.0.71 that $Rf_!K$ or f^*K is of quasi-geometric origin.

□

Definition 4.0.74. Let k be the field. Define the set of complexes of geometric origin on schemes of finite type over k to be the smallest set of isomorphism classes of objects in $D_c^b(X, \overline{\mathbb{Q}}_\ell)$ on schemes X of finite type over k that

- Contains the sheaf $\overline{\mathbb{Q}}_\ell$ on $\text{Spec } k$
- For every separated morphisms $f : X \rightarrow Y$ between schemes of finite type, whenever it contains a complex A on X , contains Rf_*A and $Rf_!A$, and whenever it contains a complex B on Y , contains f^*B and $f^!B$.
- Whenever it contains two complexes A and B on the same scheme, contains $A + B$, $A \otimes B$ and $\text{RHom}(A, B)$
- Whenever it contains a complex A , contains every subquotient of ${}^p\mathcal{H}^i(A)$ for all integers i .
- Whenever it contains a complex A , contains every shift and Tate twist of A .

This definition is identical to the definition of [Beilinson et al., 1982] except our base field is not necessarily $\mathbb{C}$.

Every perverse sheaf we consider will have geometric origin in this sense, and the proof that they do will follow from their construction. Note that Galois representations of finite image, viewed as sheaves on $\text{Spec } k$, have geometric origin: Every representation of finite image factors through $|Gal(L/k)|$ for some finite extension L of k , and then is a subrepresentation of the pushforward from $\text{Spec } L$ to $\text{Spec } k$ of the constant sheaf.

Our use of similar terminology for these two definitions is justified by the following result:

Corollary 4.0.75. *A perverse sheaf of geometric origin is also of quasi-geometric origin.*

Proof. The set of complexes of quasi-geometric origin contains the sheaf $\overline{\mathbb{Q}}_\ell$ on $\mathrm{Spec} k$. It is sufficient to check that it is closed under all the listed operations. It follows immediately from the definition that the set of complexes of quasi-geometric origin is closed under direct sums and taking perverse homology sheaves, subobjects, quotient objects, and Tate twists. The fact that it is closed under $Rf_!$ and f^* is part of Proposition 3.0.36. For all other listed operations, it is a corollary of Proposition 3.0.36, as we may write the remaining operations in terms of the operations described in Proposition 3.0.36 using standard identities among Grothendieck's six operations:

$$Rf_* = D \circ Rf_! \circ D$$

$$f^! = D \circ f^* \circ D$$

For $\Delta : X \rightarrow X \times X$ the diagonal:

$$A \otimes B = \Delta^*(A \boxtimes B)$$

$$RHom(A, B) = \Delta^!(DA \boxtimes B)$$

□

Corollary 4.0.76. *Say that an element of $\mathcal{Perv}/\mathcal{Neg}$ is of geometric origin if its property $\mathcal{P}$ representative is. The full category of objects in $\mathcal{CC}_S$ of geometric origin is a Tannakian subcategory.*

Proof. We must show it contains the identity is closed under sums, kernels, cokernels, convolution, and duals. It contains the identity because the skyscraper sheaf at the origin is a pushforward of the constant sheaf from a point. It is closed under sums

because perverse sheaves of geometric origin are closed under sums. It is closed under kernels and cokernels because the set of perverse sheaves of geometric origin is closed under subobjects and quotients. By definition, $A * B = Rsum_*(pr_1^*A \otimes pr_2^*B)$ so if A and B have geometric origin then $A * B$ does as well, and by Lemma 3.0.22 the Tannakian dual $A^\vee$ is equal to $[-1]^*DA$ so if A has geometric origin then $A^\vee$ does as well. $\square$

Lemma 4.0.77. *Let A be a complex of geometric origin on a scheme X of finite type over k and let x be a k -point. The semisimplification of the stalk $\mathcal{H}^i(A)_x$ of A at x is a Hodge-Tate representation.*

Proof. Because the set of complexes of geometric origin is preserved by pullback, we may assume that X consists of a single point. Then sheaves on X are equivalent to Galois representations. It is sufficient to prove that the irreducible components of Galois representations of geometric origin are Hodge-Tate. By Corollary 4.0.75, it is sufficient to prove that irreducible components of $H^i(Y, \overline{\mathbb{Q}}_\ell)$ for a smooth proper variety Y are Hodge-Tate. The cohomology of a smooth, proper variety Y is Hodge-Tate by [Faltings, 1988, III, Theorem 4.1], and irreducible components remain Hodge-Tate because subrepresentations and quotient representations of a Hodge-Tate representation remain Hodge-Tate [Brinon and Conrad, 2009, Proposition 2.4.4]. $\square$

We will prove a few results that help us characterize when finite groups occur as quotients of G_S and G_S^{geom} (the latter of which was defined in Definition 4.0.55).

Lemma 4.0.78. *Let A be a perverse sheaf with property $\mathcal{P}$ that corresponds to a representation of $\rho : G_S \rightarrow GL_n(\mathbb{C})$ such that $\rho(G_S^{geom})$ is finite. Then A is geometrically isomorphic to a sum of Kummer sheaves and skyscraper sheaves supported at the origin.*

Proof. Recall from Lemma 4.0.60 that the geometric monodromy group of $FTS_p(A)$ is contained in the image of ρ . We will show that the geometric monodromy group

is also contained in $\rho(G_S^{geom})$. The exact tensor functor $FTS_p(A)$ gives a map $f : \pi_1(\mathbb{A}_{\mathbb{F}_p}^1 - \{0\}) \rightarrow G_S$. We will show that the image of f lies in G_S^{geom} - equivalently, that the composition of f with g_S is trivial. This composition $g_S \circ f$ corresponds to the exact tensor functor where we take a Galois representation V unramified along S , place it at 0, and then apply FTS_p . We need to show that we always get a constant sheaf. But for any sheaf B supported at 0, $R\Psi_p B$ is supported at 0, so $FT_\psi R\Psi_p B$ is constant, as desired.

The representation associated to A has finite image and thus is semisimple. Because A is the property $\mathcal{P}$ representative of a semisimple representation, it is a semisimple perverse sheaf (Lemma 4.0.50), thus it is geometrically semisimple, hence a sum of a middle extension sheaf in degree -1 and a punctual sheaf in degree 0 [Katz, 1995, Ch.2, pp.11]. Because each summand of A has good reduction away from S , and has nearby cycles lisse on the same open set that A does, we may reduce to the case when A is either punctual or a middle extension sheaf.

Let N be the order of $\rho(G_S^{geom})$. Because the geometric monodromy of $FTS_p(A)$ is contained in $\rho(G_S^{geom})$, which has order N , the geometric monodromy of $FTS_p(A)$ has order dividing N . For any prime $p > N$ with $p \notin S$, this implies that $FTS_p(A)$ is tamely ramified, because its geometric monodromy group has order prime to p . By Laumon's local Fourier transform, if $R\Psi_p A$ had a nonzero finite singular point, $FTS_p(A)$ would be wildly ramified at ∞ . More precisely, the local monodromy representation at ∞ would have a summand with break exactly 1 by [Katz, 1990, Theorem 7.4.1(3)] in the case where A is a middle extension sheaf, and because the Fourier transform of a skyscraper sheaf is an Artin-Schreier sheaf otherwise. Thus $R\Psi_p A$ is lisse away from 0 for all $p > N$.

Using this, we will show that A is lisse away from 0 on $\mathbb{A}_{\mathbb{Q}}^1$.

When A is punctual, its support must be a finite subset of $\overline{\mathbb{Q}}$. At a prime p at which A has good reduction, the support must consist of p -adic integers in $\overline{\mathbb{Z}}_p$. Then

$R\Psi_p A$ is a punctual sheaf supported at the reduction of these numbers modulo the maximal ideal. As $R\Psi_p A$ is lisse away from zero for p sufficiently large, the reductions must be 0, hence the elements of the support are all contained in an ideal lying over p for all sufficiently large p , and thus must all be 0. So A is lisse away from zero.

When A is middle extension, let $\alpha_1, \dots, \alpha_n$ be the singularities of A . For all p large enough the α_i are distinct modulo p and A has good reduction at p , thus by Lemma 4.0.65, their reductions mod p are also singularities of $R\Psi_p A$. Thus for all p large enough the reductions of the α_i mod p are 0, so they must be 0.

Every semisimple perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$, lisse away from 0, is geometrically isomorphic to a direct sum of Kummer sheaves and skyscraper sheaves supported at 0. This is because the geometric fundamental group of $\mathbb{G}_{m, \mathbb{Q}}$ is $\hat{\mathbb{Z}}(1)$, so semisimple representations of the abelian group $\hat{\mathbb{Z}}(1)$ are sums of one-dimensional characters of $\hat{\mathbb{Z}}(1)$. Each such character must have finite order because a character of infinite order would have infinitely many Galois conjugates and hence the representation would have to be infinite-dimensional to be defined over $\mathbb{Q}$. The finite order characters factor through μ_n and thus correspond to Kummer sheaves. $\square$

The strongest possible result about finite monodromy does not hold for all sheaves, but it does hold for sheaves of geometric origin

Lemma 4.0.79. *Let A be a perverse sheaf with property $\mathcal{P}$ that corresponds to a representation $\rho : G_S \rightarrow GL_n(\mathbb{C})$ such that $\rho(G_S)$ is finite.*

Assume A is of geometric origin.

Then A is a Galois representation with finite image placed at the origin.

Proof. As A has finite image, it is a sum of irreducible representations. It is sufficient to prove this for irreducible representations. Using Lemma 4.0.78, we can take A to be geometrically isomorphic to a sum of Kummer sheaves and skyscraper sheaves supported at the origin. A sheaf geometrically isomorphic to a sum of skyscraper

sheaves supported at the origin is just a Galois representation supported at the origin. In this case, because g_S is surjective (Lemma 4.0.70), the Galois representation must have finite image, and we are done.

So we may assume that A is geometrically isomorphic to a sum of Kummer sheaves. We will then derive a contradiction. Over some number field K , A is isomorphic to a sum of sheaves of the form $V \otimes \mathcal{L}_\chi$ where V is a Galois representation and $\mathcal{L}_\chi$ is a Kummer sheaf corresponding to a faithful character χ of $\mu_n \subset K$. For p a prime of $\mathbb{Q}$ that is split in K , $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p(V \otimes \mathcal{L}_\chi) \otimes \mathcal{L}_\psi)$ is a summand of $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$. The action of Frob_p on $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ factors through a finite group, so has eigenvalues roots of unity of order $\leq C$ for some constant C , thus the action of Frob_p on $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi(V \otimes \mathcal{L}_\chi) \otimes \mathcal{L}_\psi)$ has eigenvalues roots of unity of order $\leq C$. This action is $-G(\chi, \psi)$ times the action of Frob_p on V .

Therefore we have a representation V of $\text{Gal}(\overline{\mathbb{Q}}|K)$ such that for all p split in K , Frob_p acting on V has eigenvalues $G(\chi, \psi)^{-1}$ times a root of unity of order $\leq C$. If V is not semisimple we may take the semisimplification and it will still have this property. As A is of geometric origin, its pullback to the point 1 is of geometric origin by Proposition 4.0.73, so the semisimplification of its stalk at 1 is of geometric origin by Lemma 4.0.77. Hence V , which is a semisimplification of a summand of that stalk, is Hodge-Tate. Let $a \in \mathbb{Z}$ be a Hodge-Tate weight of V , then an is a Hodge-Tate weight of $V^{\otimes n}$. We will construct a one-dimensional Galois representation W such that for all p split in K , Frob_p acts on W by $G(\chi, \psi)^n$. We will show that the Hodge-Tate weight k of W is in the interval $[1, n-1]$. Then $V^{\otimes n} \otimes W$ is a Galois representation such that for all split primes, Frob_p has eigenvalues roots of unity of order $\leq C$, hence a representation where all elements of the Galois group have eigenvalues roots of unity of order $\leq C$, hence an Artin representation. So its Hodge-Tate weights are all 0. But $an + k$ is a Hodge-Tate weight of $V^{\otimes n} \otimes W$, which is a contradiction as $k \in [1, n-1]$ and cannot be a multiple of n .

To construct this geometric Galois representation, use the Jacobi sums:

$$J(\chi_1, \chi_2) = -\frac{G(\chi_1, \psi)G(\chi_2, \psi)}{G(\chi_1\chi_2, \psi)}$$

These are normally defined only for $\chi_1, \chi_2, \chi_1\chi_2$ trivial, but we extend them to all characters using the convention that $G(1, \psi) = -1$. Hence we have the identity

$$(-G(\chi, \psi))^n = \prod_{i=0}^{n-1} (J(\chi^i, \chi))$$

If χ_1, χ_2 , and $\chi_1\chi_2$ are all nontrivial then $J(\chi_1, \chi_2)$ is the action of Frobenius on a one-dimensional Galois representation appearing in the cohomology of the Fermat curve with equation $x^n + y^n = z^n$. Hence it has Hodge-Tate weight either 0 or 1. If χ_1 or χ_2 is trivial then $J(\chi_1, \chi_2)$ is 1, so is the Frobenius action of a trivial Galois representation, which has Hodge-Tate weight 0. Finally, if χ_1 and χ_2 are nontrivial but $\chi_1\chi_2$ is trivial, then $J(\chi_1, \chi_2)$ is p , so corresponds to the Frobenius action of the cyclotomic character, with Hodge-Tate weight 1, as it appears in $H^{1,1}(\mathbb{P}^1)$.

Tensoring for each i the Galois representation corresponding to $J(\chi^i, \chi)$, we get a Galois representation where Frobenius acts by $(-G(\chi, \psi))^n$. Its Hodge-Tate weight is the sum of weights $\sum_{i=0}^{n-1} w_i$ where $w_0 = 0$, $w_{n-1} = 1$, and w_i is either 0 or 1 for all other values of i , meaning it is somewhere between 1 and $n - 1$. This gives the desired contradiction. $\square$

Remark 4.0.80. There exist counterexamples if we do not have the condition about the representations being Hodge-Tate, or another condition constraining the Galois representations somehow. Let ℓ be an odd prime and consider the approximate square root map $s : \mathbb{Z}_\ell^\times \rightarrow \mathbb{Z}_\ell^\times$ that sends any element congruent to 1 mod ℓ to its unique square root in $\mathbb{Z}_\ell^\times$ and every element of $\mu_{\ell-1}$ to itself. Let $\rho : \text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow \mathbb{Z}_\ell^\times$ be the cyclotomic character. Then the composition $s \circ \rho$ defines a one-dimensional Galois

representation V with $V \otimes V = \mathbb{Q}_\ell(-1)$ tensor a Dirichlet character of order $\ell - 1$. It follows that $\mathcal{L}_\chi[1] \otimes V^{-1}$ has finite Tannakian monodromy group, because its square under the convolution operation is a Galois representation of finite order placed at 0.

These next lemmas let us establish that a representation is Lie-irreducible, a useful property.

Definition 4.0.81. We say that a representation of an algebraic group or Lie group is Lie-irreducible if it remains irreducible when restricted to the identity component.

Lemma 4.0.82. *Let V be an irreducible representation of an algebraic group G over an algebraically closed field of characteristic 0. Assume that for each nontrivial irreducible representation W of G that factors through a finite group, $\mathrm{Hom}(V \otimes W, V) = 0$.*

Then V is Lie-irreducible.

Proof. Let G_0 be the identity component of G and consider $\mathrm{Hom}_{G_0}(V, V)$. This is naturally a representation of the component group G/G_0 . The component group is a finite group, and thus this representation decomposes into a sum of irreducible representations. The multiplicity of the trivial representation is the dimension of $\mathrm{Hom}_G(V, V)$, which is 1 as V is irreducible. Let W be a nontrivial irreducible representation of G/G_0 . The multiplicity of W in $\mathrm{Hom}_{G_0}(V, V)$ is the dimension of

$$\mathrm{Hom}_G(W, \mathrm{Hom}_{G_0}(V, V)) = \mathrm{Hom}_G(W, V \otimes V^\vee) = \mathrm{Hom}_G(V \otimes W, V)$$

and by assumption $\dim \mathrm{Hom}_G(V \otimes W, V)$ is zero. Hence $\mathrm{Hom}_{G_0}(V, V)$ is a one-dimensional representation.

V is irreducible, hence when restricted to G_0 it is semisimple. Because $\mathrm{Hom}_{G_0}(V, V)$ is one-dimensional, the restriction of V to G_0 is indecomposable. Being also simple, the restriction is irreducible. $\square$

Also note that if a representation of an algebraic group is Lie-irreducible, then as the identity component has a faithful irreducible representation, it is reductive,

hence equal to its (semisimple) commutator subgroup times its center. Because the representation is irreducible, the center consists only of scalars, so the commutator subgroup of the identity component acts irreducibly. This is very useful, because we can often enumerate the semisimple groups with an irreducible representation of a given dimension and handle them case by case.

Lemma 4.0.83. *Let A be a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$ with property $\mathcal{P}$ and good reduction away from S .*

Suppose that A is of geometric origin.

If A is geometrically irreducible, then A corresponds to a Lie-irreducible representation of G_S .

Proof. Since A is an irreducible perverse sheaf, it corresponds to an irreducible representation of G_S . So by Lemma 4.0.82, if $\text{Hom}(V \otimes W, V) = 0$ for each irreducible nontrivial G_S -representations W with finite image then A is Lie-irreducible. $\text{Hom}(V \otimes W, V)$ vanishes as long as W is not a subrepresentation of $V \otimes V^\vee$, so we may assume that W is a subrepresentation of $V \otimes V^\vee$. Then W corresponds to a subobject of $A * A^\vee$. As A is of geometric origin, $A * A^\vee$ is of geometric origin by Corollary 4.0.76.

So by Lemma 4.0.79, any (irreducible, nontrivial) finite representation of G_S that appears in $A * A^\vee$ is of the form $\delta_0 \otimes W$ for some irreducible, nontrivial finite Galois representation W . As $A * \delta_0 \otimes W = A \otimes W$ (Lemma 4.0.56), we must show that $\text{Hom}(A \otimes W, A) = 0$. This follows from the fact that A is geometrically irreducible. The dimension of $\text{Hom}(A \otimes W, A)$ is the dimension of $\text{Hom}_{\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})}(W, \text{Hom}_{\mathbb{A}_{\mathbb{Q}}^1}(A, A))$ which is the multiplicity of W in $\text{Hom}_{\mathbb{A}_{\mathbb{Q}}^1}(A, A)$, which is one-dimensional by the assumption that A is geometrically irreducible, and has a trivial action of $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ because it is generated by the identity element which is fixed by the Galois group. Hence the dimension is 0. □

In this theorem geometric irreducibility is certainly a necessary condition, but unlike the previous lemma, we do not know if there exist counterexamples without the hypothesis of geometric origin.

Chapter 5

An Equidistribution Conjecture

In this chapter, we will state a conjecture about the equidistribution of the Frobenius conjugacy classes in G_S . We begin by reviewing the generalized Sato-Tate conjecture and using it as a model for our conjecture.

To state the generalized Sato-Tate conjecture, we first need to set up some notation.

We say that a representation of $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ over $\overline{\mathbb{Q}}_\ell$ is of geometric origin if it arises as a subquotient of an étale cohomology group of an algebraic variety over $\mathbb{Q}$ with coefficients in $\overline{\mathbb{Q}}_\ell$.

Let $\rho : \text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow GL_n(\overline{\mathbb{Q}}_\ell)$ be a Galois representation that is of geometric origin and pure of weight w . Because it has geometric origin, it is unramified away from a finite set S of primes.

Let G be the Zariski closure of the image of ρ inside $GL_n(\overline{\mathbb{Q}}_\ell)$. It is an algebraic group over $\overline{\mathbb{Q}}_\ell$, so using an embedding $\iota : \overline{\mathbb{Q}}_\ell \rightarrow \mathbb{C}$ we can view it as an algebraic group over $\mathbb{C}$, and then its group $G(\mathbb{C})$ of complex points form a Lie group. Let K be a maximal compact subgroup of $G(\mathbb{C})$. Let X be the space of conjugacy classes of K .

Let G_0 be the intersection of the kernels of all homomorphisms $G \rightarrow \mathbb{G}_m$ such that the composition $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow G \rightarrow \mathbb{G}_m$ is a power of the cyclotomic character. Let K_0 be a maximal compact subgroup of G_0 .

For each prime p which is unramified for ρ , Frob_p is a conjugacy class of G whose eigenvalues are complex numbers of norm $p^{w/2}$, so $\text{Frob}_p/p^{w/2}$ is a conjugacy class of matrices whose eigenvalues have norm 1. We would like it to be a conjugacy class in G , so we will show that if $w \neq 0$ then the scalar $p^{w/2}$ is in G .

We use an argument of Deligne that shows that if ρ is pure of weight $w \neq 0$ then all scalar matrices are in G . This is because G contains the Zariski closure of the subgroup generated by (the semisimplification of) Frob_p . Diagonalize Frob_p so its entries are the eigenvalues $\lambda_1 \cdots \lambda_n$. The Zariski closure of the subgroup generated by Frob_p is defined by all relations of the form $\lambda_1^{e_1} \cdots \lambda_n^{e_n} = 1$ that are satisfied by $\lambda_1, \dots, \lambda_n$. Taking absolute values, we get $|\lambda_1|^{e_1} \cdots |\lambda_n|^{e_n} = 1$, so $p^{(\sum_{i=1}^n e_i) \times w/2} = 1$, so $\sum_{i=1}^n e_i = 0$. Hence all scalar matrices satisfy this relation and therefore are in G .

Thus $\text{Frob}_p/p^{w/2}$ is indeed a conjugacy class in G , whose eigenvalues have norm 1. Assuming $\text{Frob}_p/p^{w/2}$ is semisimple, it is contained in some compact torus, and hence conjugate to an element of K . If not, we can semisimplify it and achieve the same effect. Let $\theta_p \in X$ be the conjugacy class of the semisimplification of $\text{Frob}_p/p^{w/2}$ in K .

We might guess this sequence is equidistributed according to the Haar measure of K . But this is not so, because by dividing by $p^{w/2}$ we have ensured that Frob_p is contained in the smaller subgroup G_0 : If we take a homomorphism $\chi : G \rightarrow \mathbb{G}_m$ such that the composition $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow G \rightarrow \mathbb{G}_m$ is the n th power of the cyclotomic character, then $\chi(\text{Frob}_p)$ is the action of Frob_p on the n th power of the cyclotomic character, which is p^{-n} . The weight of the n th power of the cyclotomic character is $-2n$, so the element of G that acts on representations of weight w by $p^{w/2}$ acts on χ as p^{-n} and thus $\chi(\text{Frob}_p/p^{w/2}) = 1$, hence $\theta_p \in G_0$.

So instead we guess that θ_p is equidistributed according to the Haar measure of K_0 . Thus the generalized Sato-Tate conjecture is:

Generalized Sato-Tate Conjecture. For any Galois representation ρ of geometric origin that is pure of weight w , θ_p is equidistributed in X according to the pushforward of the Haar measure of K_0 from K_0 to X .¹ [Serre, 2011, 8.2.2]

This provides our model for an equidistribution conjecture. Also helpful is Deligne's equidistribution theorem [Deligne, 1980, Theorem 3.5.3], which is a similar statement in a different context, but is a theorem and not a conjecture. We review a slight variant of it, closely related to the version [Katz and Sarnak, 1999, Theorem 9.2.6], that will provide a closer analogue to the conjecture we will soon make.

Let X be a smooth geometrically connected variety of dimension d over a finite field $\mathbb{F}_q$, ℓ a prime number invertible in $\mathbb{F}_q$, $\iota : \overline{\mathbb{Q}}_\ell \rightarrow \mathbb{C}$ an embedding, and $\mathcal{F}$ a lisse $\overline{\mathbb{Q}}_\ell$ -sheaf on X of rank $r \geq 1$ corresponding to an r -dimensional continuous representation of $\pi_1(X)$. Denote by G_{arith} the Zariski closure of $\pi_1(X)$ in GL_r and by G_{geom} the Zariski closure of $\pi_1^{geom}(X)$. Assume that $\mathcal{F}$ is pure of weight 0. Let K be a maximal compact subgroup of G_{arith} , and let K_0 be a maximal compact subgroup of G_{geom} contained in K .

For each field extension $\mathbb{F}_{q^n}$ and point $x \in X(\mathbb{F}_{q^n})$, we have a Frobenius conjugacy class $\iota(\text{Frob}_{q^n, x})$ of $G_{arith}(\mathbb{C})$. By assumption, all its eigenvalues have absolute value 1, so its semisimplification $\iota(\text{Frob}_{q^n, x})^{ss}$ is conjugate to an element of K . Let $\theta_{q^n, x}$ be the conjugacy class in K of this semisimplification.

¹We have made a conjecture about the distribution of the conjugacy classes in K . We could have instead made a conjecture about the distribution of the conjugacy classes in K_0 , using again the Haar measure on K_0 . These conjectures are equivalent because the map from conjugacy classes of K_0 to conjugacy classes of K is injective: For each character χ of G such that the composition $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q}) \rightarrow G \rightarrow \mathbb{G}_m$ is the n th power of the cyclotomic character for some $n \neq 0$, the scalar $p^{w/2}$ in K acts by p^{-n} on χ , which generates a dense subgroup of the image of χ , so the center of K maps surjectively to the image of χ , so it maps surjectively to K/K_0 , so the action of K/K_0 on conjugacy classes of K_0 is trivial.

Observe that the homomorphism $\pi_1(X) \rightarrow G_{arith}/G_{geom}$ factors through $\text{Gal}(\overline{\mathbb{F}}_q|\mathbb{F}_q)$. Let σ be the image of a generator Frob_q of $\text{Gal}(\overline{\mathbb{F}}_q|\mathbb{F}_q)$ under this homomorphism. Then $\iota(\sigma)^{ss}$ is conjugate to an element of the maximal compact subgroup K/K_0 of G_{arith}/G_{geom} . Let α be a coset representative.

Let μ be the Haar measure on K_0 .

Theorem 5.0.84. *Let f be a continuous, conjugacy-invariant function on K . Then*

$$\lim_{n \rightarrow \infty} \left(\frac{\sum_{x \in X(\mathbb{F}_{q^n})} f(\theta_{q^n, x})}{q^{nd}} - \int_{g \in K_0} f(\alpha^n g) d\mu(g) \right) = 0$$

Proof. First note that $|X(\mathbb{F}_{q^n})| = q^{nd} + o(q^{nd})$. This implies that $\left| \frac{\sum_{x \in X(\mathbb{F}_{q^n})} f(\theta_{q^n, x})}{q^{nd}} \right| \leq (1 + o(1)) \sup |f|$. Because the linear combinations of characters are dense in the space of continuous, conjugacy-invariant functions in the sup-norm topology, it suffices to prove that this limit vanishes when f is a linear combination of characters. By linearity, it suffices to prove this when f is the character of an irreducible representation V . Then

$$f(\theta_{q^n, x}) = \text{tr}(\iota(\text{Frob}_{q^n, x})^{ss}, V) = \iota(\text{tr}(\text{Frob}_{q^n, x}, V))$$

so

$$\frac{\sum_{x \in X(\mathbb{F}_{q^n})} f(\theta_{q^n, x})}{q^{nd}} = \frac{\sum_{i=0}^{2d} \iota(\text{tr}(\text{Frob}_q^n, H_c^i(X_{\overline{\mathbb{F}}_q}, \mathcal{F})))}{q^{nd}}.$$

The next term

$$\int_{g \in K_0} \text{tr}(g\alpha^n, V) d\mu(g) = \text{tr} \left(\left(\int_{g \in K_0} g\mu(g) \right) \alpha^n, V \right).$$

$\int_{g \in K_0} g d\mu$ is the projection onto the K_0 -invariant subspace of V , which vanishes unless V factors through K/K_0 , in which case it is the trace of α^n on V , equivalently, the trace of $\iota(\text{Frob}_{q^n})$ on V .

But $\text{tr}(\text{Frob}_q^n, H_c^{2d}(X_{\overline{\mathbb{F}}_q}, \mathcal{F}))$ vanishes unless V is an irreducible representation of G_{arith}/G_{geom} , in which case it is exactly q^{nd} times the trace of Frob_{q^n} on V by Poincare

duality. So

$$\iota(\mathrm{tr}(\mathrm{Frob}_q^n, H_c^{2d}(X_{\mathbb{F}_q}, \mathcal{F})) = q^{nd} \int_{g \in K_0} f(g\alpha^n) d\mu$$

and those terms cancel. It remains to show that

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=0}^{2d-1} \iota(\mathrm{tr}(\mathrm{Frob}_q^n, H_c^i(X_{\mathbb{F}_q}, \mathcal{F})))}{q^{nd}} = 0$$

but this follows immediately from Deligne's Riemann hypothesis, as $\left| \iota(\mathrm{tr}(\mathrm{Frob}_{q^n}, H_c^i(X_{\mathbb{F}_q}, \mathcal{F}))) \right| \leq q^{in/2} \dim(H_c^i(X_{\mathbb{F}_q}, \mathcal{F})) = o(q^{nd})$ if $i < 2d$. $\square$

This statement shows “equidistribution” according to the uniform measure on a coset $\alpha^n K_0$ of K_0 inside K . However, because this measure is a function of n , we have modified the typical equidistribution statement by subtracting the integral against of f this measure from the average of f over points before, not after, taking the limit.

In our equidistribution conjecture, the measure will not vary with n , but it will also not be a Haar measure. Instead, our conjectural measure will be a linear combination of uniform measures on cosets of subgroups. As in Deligne's theorem, these will be normal subgroups with abelian quotients.

We now begin to state our conjecture. Let S be a finite set of primes, and let $A \in \mathcal{Perv}$ be pure of weight w , of geometric origin, and have good reduction outside S , corresponding to a representation $\rho : G_S \rightarrow GL_n(\mathbb{C})$. Let G be the image of ρ and let K be a maximal compact subgroup of G . Let θ_p be the conjugacy class of $\mathrm{Frob}_p^{ss}/p^{w/2}$ in K . We would like to conjecture that Frob_p is equidistributed, based on the examples of the Sato-Tate conjecture and Deligne's equidistribution theorem, it will not be according to the Haar measure. These examples suggest that we should apply special scrutiny to abelian quotients of K , equivalently, apply special scrutiny to the one-dimensional sheaves of geometric origin. By Lemma 4.0.58, we can completely characterize these. Each one-dimensional sheaf is a convolution of objects of the following three forms:

- A 1-dimensional $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ -representation V of geometric origin. Here a_p is the action of Frobenius on V . Any such Galois representation is Hodge-Tate by Lemma 4.0.77, hence a power of the cyclotomic character tensored with a 1-dimensional Artin representation [Serre, 1968, III-50 Corollary and III-11 Proposition 2]. The action of Frobenius on V will then be a power of p times a Dirichlet character. Once we normalize by dividing by the power of p , Dirichlet's theorem on primes in arithmetic progressions shows the sequence is equidistributed according to the Haar measure of the n th roots of unity, where n is the order of the Dirichlet character.
- A rank 1 skyscraper sheaf placed at the point $r \in \mathbb{Q}$ in degree 0. Let $r = a/b$ for $a, b \in \mathbb{Z}$ relatively prime. Then $a_p(\delta_{a/b}) = e^{2\pi i \frac{x}{p}}$, where $bx \equiv a$ modulo p . Because $bx \equiv a$ modulo p , we have $bx = a + kp$ for some integer k , so $a_p = e^{2\pi i (\frac{a}{bp} + \frac{k}{b})}$. As p goes to ∞ , $\frac{a}{bp}$ goes to 0, so a_p has the same distribution as $e^{2\pi i \frac{k}{b}}$. To compute the distribution of a_p , we must compute k modulo b . We know $a + kp = bx$, hence $a + kp \equiv 0 \pmod{b}$, so $k \equiv -ap^{-1} \pmod{b}$. By Dirichlet's theorem on primes in arithmetic progressions, the sequence $p \mapsto a_p$ is equidistributed according to a measure that assigns mass $1/\varphi(b)$ to each of the primitive b th roots of unity. This is not the Haar measure of any subgroup of $\mathbb{G}_m$ unless $b = 1$.
- The sheaf $\mathcal{L}_\chi[1]$ for χ the nontrivial character of μ_2 , placed in degree -1 . Here $a_p = -g(\chi)$, where $g(\chi)$ is the quadratic Gauss sum $\sum_{x=1}^{p-1} e^{2\pi i \frac{x^2}{p}} \left(\frac{x}{p}\right)$. The Gauss sum $g(\chi)$ is $\sqrt{p}$ or $i\sqrt{p}$ depending on whether $p \equiv 1$ or 3 modulo 4 . So $a_p/\sqrt{p} = -1$ or $-i$ depending on whether $p \equiv 1$ or 3 modulo 4 . So by Dirichlet's theorem on primes in arithmetic progressions, $a_p/\sqrt{p}$ is equidistributed according to the measure which assigns probability $1/2$ to -1 and $1/2$ to $-i$. This is not the Haar measure of any subgroup of $\mathbb{G}_m$.

In each of the 3 cases just discussed, once a_p is normalized, it is equal to a number that depends only on the congruence class of p modulo some suitable modulus plus an error term that is $O(1/p)$.

We will make the minimal possible adjustment to the Haar measure that predicts the correct outcome for abelian quotients. This means the measure should be uniform with respect to the largest possible subgroup that maps trivially to every abelian quotient of the group. That is, the measure should be uniform with respect to the commutator subgroup.

Let S be a finite set of primes, and let $A \in \mathcal{Perv}$ be pure of weight w , of geometric origin, and have good reduction outside S , corresponding to a representation $\rho : G_S \rightarrow GL_n(\mathbb{C})$. Let $G = \rho(G_S)$ and let K be a maximal compact subgroup of $G(\mathbb{C})$.

Definition 5.0.85. Let the Dirichlet measure on the abelianization $K^{ab} = K/[K, K]$ be the measure defined as follows: Choose a faithful set of characters $\chi_1, \dots, \chi_n$ of $G^{ab} = G/[G, G]$. The abelianization of K is a maximal compact subgroup of the abelianization of G , so each element of K is uniquely determined by its action on $\chi_1, \dots, \chi_n$. By composition, each χ_i gives a representation of G_S and hence a one-dimensional sheaf of the form described by Lemma 4.0.58

Let m be the smallest natural number such that, for each character χ_i , the corresponding one-dimensional sheaf is either of the form $\delta_{a/m} \otimes V$ or $\mathcal{L}_\chi * \delta_{a/m} \otimes V$ for $a \in \mathbb{Z}$, V a one-dimensional Galois representation with conductor dividing m , and with $\mathcal{L}_\chi$ appearing only if 4 divides m . Define a function d from the set of invertible congruence classes mod m to K^{ab} that sends $r \in (\mathbb{Z}/m)^\times$ to the unique element of K^{ab} that:

1. acts on a one-dimensional Galois representation corresponding to a Dirichlet character χ times any power of the cyclotomic character as $\chi(r)$.

2. acts on the skyscraper sheaf $\delta_{a/m}$ as $e^{-2\pi i \frac{ar^{-1}}{m}}$, where r^{-1} is the inverse modulo m .

3. acts on $\mathcal{L}_\chi[1]$ as -1 if $r \equiv 1$ modulo 4 and $-i$ if $r \equiv 3$ modulo 4.

and acts on convolutions of these as a product of those numbers.

Then the Dirichlet measure is the pushforward of the uniform measure on the set of invertible congruence class mod m to K^{ab} along d .

In each of our three earlier computations, the conjugacy class of Frob_p in K^{ab} is equidistributed according to the Dirichlet measure. So we adjust the Haar measure on K to account for this:

Definition 5.0.86. Let the adjusted Haar measure on K be the unique measure that is invariant under left translation by $[K, K]$ and whose projection to $K/[K, K] = K^{ab}$ is the Dirichlet measure.

We conjecture that no further adjustment is needed for the non-abelian case, and so the adjusted Haar measure is correct:

Conjecture 1. Let S be a finite set of primes, and let $A \in \mathcal{P}erv$ be pure of weight w , of geometric origin, and have good reduction outside S , corresponding to a representation $\rho : G_S \rightarrow GL_n(\mathbb{C})$. Let G be the image of ρ and let K be a maximal compact subgroup of G . Let X be the space of conjugacy classes of K . For each prime p not in S , $\text{Frob}_p^{ss}/p^{w/2}$ defines a conjugacy class in K , and hence a point $\theta_p \in X$. The sequence θ_p is equidistributed according to the pushforward of the adjusted Haar measure on K from K to X .

There are two things to check to make sure this conjecture is well-defined and plausible:

Proposition 5.0.87. *In the notation of Conjecture 1, θ_p is in fact a conjugacy class of G for any prime p not in S .*

Proof. Frob_p is a conjugacy class of G by definition. Because G is an algebraic group, Frob_p^{ss} is in G . So it is sufficient to show that $p^{w/2}$ is in G . There is an automorphism of the fiber functor of the category of representations of G_S that for each λ acts on the generalized eigenspace of Frob_p with eigenvalue λ by $|\lambda|$. This corresponds by Tannaka-Krieger duality to a conjugacy class of G_S . Because ρ is pure of weight w , each eigenvalue λ of Frob_p on ρ has $|\lambda| = p^{w/2}$, so this class is sent by ρ to the scalar $p^{w/2}$. $\square$

Proposition 5.0.88. *The image of θ_p in K^{ab} is equidistributed according to the Dirichlet measure on K^{ab} .*

Proof. By Weyl's equidistribution theorem and the definition of the Dirichlet measure, it is sufficient to prove that, for each character χ of G^{ab} ,

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \frac{\chi(\text{Frob}_p)}{p^{w(\chi)/2}}}{\pi(n)} = \frac{\sum_{r \in (\mathbb{Z}/m)^\times} \chi(d(r))}{\phi(m)}$$

χ gives by composition a one-dimensional representation of G_S , which we know has the form $\delta_{a/m} \otimes V$ or $\mathcal{L}_\chi[1] * \delta_{a/m} \otimes V$, where V has conductor dividing m and where, in the second case, m is multiple of 4. Our three earlier computations about the action of Frob_p on $\delta_0 \otimes V$, and $\delta_{a/m}$, and $\mathcal{L}_\chi$ imply that, for a prime p not dividing m , $\text{Frob}_p/p^{w/2}$ and $d(p)$ have the same eigenvalue acting on $\delta_0 \otimes V$, their eigenvalues acting on $\delta_{a/m}$ differ by $O(1/p)$, and they have the same eigenvalue acting on $\mathcal{L}_\chi$. Hence by convolving objects of those types to obtain χ , we get:

$$\frac{\chi(\text{Frob}_p)}{p^{w(\chi)/2}} = \chi(d(p)) + O(1/p)$$

Hence

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \frac{\chi(\text{Frob}_p)}{p^{w(\chi)/2}}}{\pi(n)} = \lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} (\chi(d(p)) + O(1/p))}{\pi(n)} = \lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \chi(d(p))}{\pi(n)} = \frac{\sum_{r \in (\mathbb{Z}/m)^\times} \chi(d(r))}{\phi(m)}$$

by Dirichlet's theorems on primes in arithmetic progressions, as desired. $\square$

An alternate form of the conjecture is:

Conjecture 2. Let S be a finite set of primes, and let $A \in \mathcal{P}erv$ be irreducible, pure of weight w , of geometric origin, with good reduction outside S , and corresponding to a representation of G_S of dimension > 1 . Let a_p be the trace of Frob_p acting on $H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi A \otimes \mathcal{L}_\psi)$. Then:

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n, p \text{ prime}} \frac{a_p}{p^{w/2}}}{\pi(n)} = 0$$

Proposition 5.0.89. *Conjecture 1 and Conjecture 2 are equivalent.*

Proof. μ be the adjusted Haar measure. By Weyl's equidistribution criterion, the sequence of conjugacy classes θ_p is equidistributed according to μ if and only if, for each irreducible representation V of G_S , as n goes to ∞ the average over $p \leq n$ of the trace of θ_p converges to $\int_K \text{tr}(g, V) d\mu(g)$.

If V is a one-dimensional representation, then $K = K^{ab}$ and the adjusted Haar measure is the Dirichlet measure by definition. The average of $\text{tr}(\theta_p, V)$ over primes converges to the integral of $\text{tr}(g, V)$ against the Dirichlet measure by Proposition 5.0.88. So the condition that the average over $p \leq n$ of the trace of θ_p converges to $\int_K \text{tr}(g, V) d\mu(g)$ is automatically satisfied for one-dimensional representations.

To check that the two conjectures are equivalent for higher-dimensional representations, it is sufficient to show that $\int_K \text{tr}(V) d\mu = 0$ for all irreducible representations V of dimension greater than 1. The adjusted Haar measure μ is a finite sum of Haar measures on cosets of $[K, K]$, so it is sufficient to check this statement for the Haar measure of a coset of $[K, K]$. Let μ_h be the Haar measure of $[K, K]$ and let $\gamma[K, K]$ be a coset. We must check that

$$\int_{[\gamma K, \gamma K]} \text{tr}(\gamma g, V) d\mu_h(g) = 0$$

By linearity,

$$\int_{[K,K]} \text{tr}(\gamma g, V) d\mu_h(g) = \text{tr} \left(\gamma \int_{[K,K]} g d\mu_h(g), V \right)$$

$\int_{[K,K]} g d\mu_h(g)$ is the projection onto the $[K, K]$ -invariant subspace of V . Since $[K, K]$ is a normal subgroup of K and V is an irreducible representation of K , that subspace is either 0 or V . Because V is an irreducible representation of dimension > 1 , the action of K on V doesn't factor through an abelian quotient, so $[K, K]$ acts nontrivially on V , thus the invariant subspace must be 0, hence

$$\int_{[K,K]} g d\mu_h(g) = 0$$

and

$$\int_{[K,K]} \text{tr}(\gamma g, V) d\mu_h(g) = 0$$

□

Chapter 6

Complete Analysis of the Examples

This chapter finishes the analysis of the examples from Chapter 2. In Chapter 2, we showed how some number-theoretic problems could be interpreted as questions about the distribution of our exponential sums $a_p(\mathcal{F})$. In Chapter 3, we interpreted those exponential sums as the traces of Frobenius elements in representations of the Tannakian group G_S . We developed some tools to help compute the action of G_S in Chapter 4. We conjectured a distribution for the conjugacy classes of these Frobenius elements in terms of the action of G_S in Chapter 5.

Now we will see which G_S -representations correspond to the exponential sums of Chapter 2 and which distributions are implied by Conjecture 1. In all the problems we study, our conjecture agrees with the previously existing conjecture or, when there was no previous conjecture, with a natural guess. Furthermore, two cases of our conjecture are already theorems, which have been proved using techniques from analytic number theory (sieve methods and metaplectic automorphic forms).

In Chapter 7, we will look at some special cases of Conjecture 1 that have not been previously studied by number theorists. In those cases we instead justify our conjecture with numerical evidence, which we discuss in Chapter 9.

In our analysis of each example, the first step will be to compute the image of G_S in the representation corresponding to the sheaf we are studying.

Definition 6.0.90. Let $\mathcal{C}$ be a Tannakian category with a fiber functor. The Tannakian monodromy group of an object in $\mathcal{C}$, corresponding to a representation V of the Tannakian group, is the image of the Tannakian group in $GL(V)$.

The first example concerns the distribution of roots of polynomials, as in Example 2.0.2.

Proposition 6.0.91. *Let $f(x) \in \mathbb{Z}[x]$ be a separable polynomial with no rational roots. Then assuming Conjecture 1, the sequence $(x/p) \in S^1$ parameterized by ordered pairs of a prime p and a root x of f modulo p are equidistributed in the circle according to the uniform measure, as in Example 2.0.2*

Proof. Let f_k be the polynomial whose roots are k times the roots of f , let i_k be the inclusion from $\text{Spec}(\mathbb{Q}[x]/(f_k(x)))$ into $\mathbb{A}_{\mathbb{Q}}^1$, and let $\mathcal{F}_k = i_{k*}\overline{\mathbb{Q}}_{\ell}$. Then by Weyl's equidistribution criterion, the roots of f are equidistributed if, for all $k \neq 0$.

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p, f_k(x)=0} e^{2\pi i k x/p}}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}|} = 0$$

Note that $\sum_{x \in \mathbb{F}_p, f_k(x)=0} e^{2\pi i k x/p} = a_p(\mathcal{F}_k)$ and the denominator is $\pi(n)$ times the number of irreducible factors of f times $1 + o(1)$, so this limit is zero as long as:

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} a_p(\mathcal{F}_k)}{\pi(n)} = 0$$

By Proposition 5.0.89, Conjecture 1 is equivalent to Conjecture 2. The limit of $\frac{\sum_{p \leq n} a_p(\mathcal{F}_k)}{\pi(n)}$ is zero as long as $\mathcal{F}_k$ is a sum of irreducible perverse sheaves of geometric origin, each pure of weight 0, each with Tannakian dimension greater than 1. (By Proposition 3.0.46, being of geometric origin, they automatically have good reduction outside a finite set of primes.)

$\mathcal{F}_k$ is punctual in degree 0, hence perverse, and is the pushforward of the constant sheaf from the smooth scheme $\text{Spec}(\mathbb{Q}[x]/(f_k(x)))$ along the proper morphism f_k , hence of geometric origin and pure of weight 0. So it is sufficient to show that it is a sum of irreducible perverse sheaves of Tannakian dimension greater than 1.

As $\text{Spec}(\mathbb{Q}[x]/(f_k(x)))$ is the disjoint union of $\text{Spec}(\mathbb{Q}[x]/(g_i(x)))$ over irreducible factors g_i of f_k , it is sufficient to show that the pushforward from $\text{Spec}(\mathbb{Q}[x]/(g_i(x)))$ to $\mathbb{A}_{\mathbb{Q}}^1$ of the constant sheaf is an irreducible perverse sheaves of Tannakian dimension greater than 1.

This sheaf is irreducible as geometrically it is the sum of the skyscraper sheaves at the roots of g_i , so any nontrivial proper subsheaf is the sum of the skyscraper sheaves at some nontrivial proper subset of the roots, and hence cannot be defined over $\mathbb{Q}$ as no such subset is defined over $\mathbb{Q}$.

The dimension is, by Lemma 4.0.54, the number of roots of g_i , which is equal to the degree of g_i . As f has no rational roots, f_k has no rational roots, so f_k has no factors of degree 1, hence the degree of g_i is greater than 1. $\square$

Proposition 6.0.91 is known unconditionally when f is quadratic and irreducible. Duke et al. [1995, Theorem on pp. 2] handled the case where $f(x) = ax^2 + bx + c$ is quadratic, b is even, and the discriminant $b^2 - 4ac$ is negative. Tóth [2000, Theorem 1.2] extended the result to all quadratic polynomials.

The case when the degree of f is greater than 2 has not been proven. Duke et al. [1995, pp. 1] guessed that uniform equidistribution is still obtained.

Next we can study the problem of Example 2.0.8, about the congruence class mod m of small roots of $x^2 - D \equiv 0 \pmod{p}$.

Proposition 6.0.92. *Let D be an integer that is not a perfect square and m a natural number such that the conductor of the quadratic character $\mathbb{Q}(\sqrt{D})$ does not divide m . For a prime p modulo which D is a perfect square, let x_p be the smallest integer root*

of $x^2 - D \equiv 0 \pmod{p}$. Assuming Conjecture 1, the congruence class of x_p modulo m is equidistributed in $\mathbb{Z}/m$ according to the uniform measure.

Proof. We will continue to use the notation of Example 2.0.8. In that example, we showed that to prove equidistribution, it is sufficient to show that for each $k \in \mathbb{Z}$ and $\chi : (\mathbb{Z}/m)^\times \rightarrow \mathbb{C}^\times$

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n, \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p) e^{2\pi i \frac{kx}{p}}}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | (mx)^2 = D\}|} = 0$$

unless $k = 0$ and $\chi = 1$. Since the denominator is $(1 + o(1))\pi(n)$ by Dirichlet's theorem, it is sufficient to show that

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n, \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p) e^{2\pi i \frac{kx}{p}}}{\pi(n)} = 0$$

First assume $k \neq 0$. In Example 2.0.2 we constructed an object $\mathcal{F}_k$ such that $a_p(\mathcal{F}_k) = \sum_{x \in \mathbb{F}_p | (mx)^2 = D} e^{2\pi i \frac{kx}{p}}$. By tensoring with the Galois representation V_χ associated to χ , we have an object $\mathcal{F}_k \otimes V_\chi$ with $a_p(\mathcal{F}_k \otimes V_\chi) = \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p) e^{2\pi i \frac{kx}{p}}$. Hence it is sufficient to show that:

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} a_p(\mathcal{F}_k \otimes V_\chi)}{\pi(n)} = 0$$

By Lemma 4.0.54, as $\mathcal{F}_k \otimes V_\chi$ is geometrically a sum of skyscraper sheaves at two points, it is a two-dimensional object in the Tannakian category. To be reducible it would have to be an extension of two sheaves, each supported at one of the points, which is impossible because the points are not defined over $\mathbb{Q}$. Thus it is an irreducible two-dimensional representation.

Furthermore $\mathcal{F}_k \otimes V_\chi$ is a summand of the pushforward from $\text{Spec } \mathbb{Q}(\mu_m)[x]/((mx)^2 - Dk^2)$ to $\mathbb{A}_{\mathbb{Q}}^1$, so it is of geometric origin and pure of weight 0. Thus $\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} a_p(\mathcal{F}_k \otimes V_\chi)}{\pi(n)}$ is 0 assuming Conjecture 2, which is equivalent to Conjecture 1 by Proposition 5.0.89.

If $k = 0$ and $\chi \neq 1$, then ignoring the finitely many p dividing m , the limit reduces to

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p)}{\pi(n)} = \lim_{n \rightarrow \infty} \frac{\sum_{p \leq n, \left(1 + \left(\frac{D}{p}\right)\right) \chi(p)} \chi(p)}{\pi(n)}$$

By Dirichlet's theorem, this converges to 0 as long as χ is not equal to the Dirichlet character $p \mapsto \left(\frac{D}{p}\right)$. As that character is the quadratic character of $\mathbb{Q}(\sqrt{D})$, by assumption its conductor does not divide m , χ does not equal $\left(\frac{D}{p}\right)$, so the sum indeed converges to 0. □

In fact, this is a theorem without assuming Conjecture 1. Tóth [2000, Theorem 1.2] proves that for each modulus m and each invertible residue class $r \bmod m$, the sequence $\frac{x}{p}$ indexed by pairs of a prime p congruent to $r \bmod m$ and a root $x \in \mathbb{F}_p$ of $x^2 - D = 0$ is equidistributed according to the Haar measure on $\mathbb{R}/\mathbb{Z}$. This implies the formula

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p) e^{2\pi i \frac{kx}{p}}}{\pi(n)} = 0$$

for $k \neq 0$, by breaking up the sum $\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (mx)^2 = D} \chi(p) e^{2\pi i \frac{kx}{p}}$ into sums over each residue class. Then we can show that the limit also is zero for $k = 0$ and χ nontrivial by the same argument as in the proof of Proposition 6.0.92. As discussed in Example 2.0.8, this suffices to prove equidistribution of the residue class of the root mod m . (The assumption that the conductor of $\mathbb{Q}(\sqrt{D})$ does not divide m is inessential, and is just used to avoid a small technical difficulty.)

Next we study the distribution of Kloosterman sums:

Lemma 6.0.93. *Let a, b be nonzero integers. Let π be the map from $\mathbb{G}_{m, \mathbb{Z}}$ to $\mathbb{A}_{\mathbb{Z}}^1$ defined by $x = ay + by^{-1}$. Let $\mathcal{F} = \pi_* \mathbb{Q}_l$.*

Then $\mathcal{F}[1]$, pulled back to $\mathbb{A}_{\mathbb{Q}}^1$, is perverse, of geometric origin, and pure of weight 1. It has good reduction away from the set S of primes dividing $2lab$, and corresponds

to a 2-dimensional representation of G_S with image group GL_2 . The determinant is $\delta_0 \otimes \overline{\mathbb{Q}}_\ell(-1)$.

Proof. $\mathcal{F}$ is the pushforward of a lisse sheaf along a dominant map of smooth curves, which is a middle extension sheaf by [Katz, 1990, Lemma 7.3.2], and a middle extension sheaf on a curve is perverse when placed in degree -1 . By definition $\mathcal{F}$ is the pushforward of a sheaf pure of weight 0 along a finite cover, hence is pure of weight 0, so placing it in degree -1 makes it pure of weight 1. Because it is the pushforward of the constant sheaf along π , it is of geometric origin.

To show $\mathcal{F}[1]$ has good reduction, we observe that $\mathcal{F}$ is lisse of rank 2 away from the critical values of $ay + by^{-1}$, which are the roots of $x^2 - 4ab$, and lisse of rank 1 restricted to the locus where $x^2 = 4ab$. Thus by Proposition 3.0.46, it has good reduction at primes other than ℓ where the curve $x^2 - 4ab$ is finite étale, which are exactly the primes not dividing $2\ell ab$.

By [Katz, 1988, Theorem 11.1], the Fourier transform of the reduction of the sheaf mod p for p not dividing a or b has geometric monodromy group SL_2 , so by Lemma 4.0.60 the Tannakian monodromy group of $\mathcal{F}$ contains SL_2 .

The determinant is some 1-dimensional object in our Tannakian category. As f can be extended to a map $\mathbb{P}^1 \rightarrow \mathbb{P}^1$ that is finite étale over ∞ , the local monodromy of $\mathcal{F}$ is 0, so by Lemma 4.0.66 the determinant has the form $\delta_r \otimes V$. As the singularities are $\pm 2\sqrt{ab}$, each with drop 1, we know from Lemma 4.0.67 that $r = 0$. Hence the determinant is of the form $\delta_0 \otimes V$ for some Galois representation V .

To show that V is the cyclotomic character, it is sufficient to show that Frob_p acts on it by multiplication by p for all but finitely many p . As $A \mapsto H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p A \otimes \mathcal{L}_\psi)$ is a fiber functor, the eigenvalue of Frob_p on $\det H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p \mathcal{F}[1] \otimes \mathcal{L}_\psi)$ is the same as the action of Frob_p on

$$H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p \delta_0 \otimes V \otimes \mathcal{L}_\psi) = H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, \delta_0 \otimes V \otimes \mathcal{L}_\psi) = V$$

The fact that the determinant of the action of Frob_p on $H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, \mathcal{F} \otimes \mathcal{L}_\psi)$ is p follows from the fact that the two-dimensional representation $H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, \mathcal{F} \otimes \mathcal{L}_\psi)$ is symplectically self-dual up to a Tate twist, which is proved in [Katz, 1988, Proposition 5.5.1]. Then because by Lemma 3.0.46, the usual specialization agrees with the nearby cycles specialization at all but finitely many primes, it follows that Frob_p acts on V by multiplication by p at all but finitely many primes, and hence V is the cyclotomic character.

Since the determinant is the cyclotomic character, and the Tannakian monodromy group of the cyclotomic character is $\mathbb{G}_m$, the Tannakian monodromy group of $\mathcal{F}$ maps surjectively onto $\mathbb{G}_m$ under the determinant map. Because we already know the Tannakian monodromy group of $\mathcal{F}$ contains SL_2 , we can conclude that it is GL_2 .

□

Remark 6.0.94. In fact if ℓ , a , and b are all odd then $\mathcal{F}$ actually has good reduction at 2, though this is not stated in Lemma 6.0.93. One can see this by bypassing Lemma 3.0.46 and computing $R\Psi_2\mathcal{F}$ directly.

Proposition 6.0.95. *If Conjecture 1 holds, then the normalized Kloosterman sum $\mathcal{Kl}(a, b; p)/\sqrt{p}$ is equidistributed according to the Sato-Tate measure.*

Proof. By the Lefschetz trace formula, the Kloosterman sum $\mathcal{Kl}(a, b; p)$ is minus the trace of Frob_p acting on $\mathcal{F}[1] \in \mathcal{Perv}/\mathcal{Neg}$, where $\mathcal{F} = \pi_*\overline{\mathbb{Q}}_\ell$ is the sheaf described in Lemma 6.0.93. So it is sufficient to estimate the distribution of the Frobenius conjugacy classes of this sheaf.

In the notation of Conjecture 1, $G = GL_2$ so $K = U(2)$. The commutator $[K, K] = SU(2)$ and $K^{ab} = U(1)$. To compute the adjusted Haar measure on K , we must compute the Dirichlet measure on K^{ab} . Because K^{ab} is generated by just one character, which is the cyclotomic character, the Dirichlet measure is supported entirely at the identity. Thus the adjusted Haar measure is exactly the Haar measure

on $SU(2)$. To find the distribution of the normalized Kloosterman sum, assuming Conjecture 1, we push forward on the trace map. The pushforward of the Haar measure of $SU(2)$ along the trace map is, by definition, the Sato-Tate measure. Since this measure is unchanged by switching x and $-x$, the Kloosterman sum is also distributed according to the Sato-Tate measure. $\square$

Next we study sums of the form, for $a \in \mathbb{Q}$, $a \neq 0$:

$$\sum_{x=0}^{p-1} e^{2\pi i \frac{ax^3+x}{p}}$$

This corresponds to the sheaf $f_*\overline{\mathbb{Q}}_\ell$, where f is the map $\mathbb{A}^1 \rightarrow \mathbb{A}^1$ with equation ax^3+x . The distribution of this was studied by Livné and Patterson [2002], and they predicted distribution according to the Sato-Tate measure [Livné and Patterson, 2002, pp. 2]. This is what Conjecture 1 predicts as well:

Lemma 6.0.96. *Let $a \in \mathbb{Q}$, $a \neq 0$, and let $f : \mathbb{A}_{\mathbb{Q}}^1 \rightarrow \mathbb{A}_{\mathbb{Q}}^1$ be defined by the equation $f(x) = ax^3 + x$. Then $f_*\overline{\mathbb{Q}}_\ell[1]$ is a perverse sheaf, of geometric origin, and pure of weight 0 which has good reduction away from some finite set S of primes, and is a two-dimensional G_S -representation, whose image is the full GL_2 . The determinant character is $\delta_0 \otimes \overline{\mathbb{Q}}_\ell(-1)$.*

Proof. $f_*\overline{\mathbb{Q}}_\ell$ is a pushforward of a lisse sheaf along a finite morphism, hence by [Katz, 1990, Lemma 7.4.2] is a middle extension sheaf, so when placed in degree -1 it is perverse. Also $f_*\overline{\mathbb{Q}}_\ell$ is the pushforward of the constant sheaf along a proper morphism from a smooth scheme, hence is of geometric origin and pure of weight 0, and thus $f_*\overline{\mathbb{Q}}_\ell[1]$ is pure of weight 1. Since $f_*\overline{\mathbb{Q}}_\ell$ extends to a sheaf over $\mathbb{A}_{\mathbb{Z}[1/N]}^1$ for N the denominator of a , by Proposition 3.0.46 it has good reduction at all but finitely many primes.

$f_*\overline{\mathbb{Q}}_\ell$ is lisse away from the branch points of f , which are $f(x)$ for x solutions of $f'(x) = 0$. These solutions are the roots of $3ax^2 + 1$, so $x = \pm\sqrt{\frac{-1}{3a}}$ and $f(x) = \pm\frac{2}{3}\sqrt{\frac{-1}{3a}}$. There are two such points. Since the derivative vanishes to first order at each point, they are points of ramification order 2, so the drop is 1 at both points. Thus by Lemma 4.0.54, this is a two-dimensional object in our Tannakian category. Since f has (geometric) monodromy group S_3 , as a sheaf $f_*\overline{\mathbb{Q}}_\ell$ factors through the permutation representation of S_3 , which splits into a sum of a trivial representation and a nontrivial (geometrically) irreducible representation. Removing the trivial summand, as it is an element of $\mathcal{N}eg$, we obtain a perverse sheaf which is geometrically irreducible on an open subset of $\mathbb{A}_{\mathbb{Q}}^1$. Since it is a pushforward of a lisse sheaf along a finite morphism, it's a middle extension sheaf by [Katz, 1990, Lemma 7.3.2] (and still is a middle extension after removing the trivial summand), hence a geometrically irreducible perverse sheaf. Then by Lemma 4.0.83 it is a Lie-irreducible G_S representation. Every Lie-irreducible subgroup of GL_2 contains SL_2 , so the Tannakian monodromy group contains SL_2 .

We will compute the determinant by showing that the sheaf is self-dual up to a Tate twist. We can compute the dual using Lemma 3.0.22. Because f is finite, pushforward along f commutes with Verdier duality, and because f is odd, it commutes with $[-1]$, so:

$$(f_*\overline{\mathbb{Q}}_\ell[1])^\vee = [-1]^*D(f_*\overline{\mathbb{Q}}_\ell[1]) = [-1]^*f_*(D\overline{\mathbb{Q}}_\ell) = [-1]^*f_*\overline{\mathbb{Q}}_\ell[1](1) = f_*[-1]^*\overline{\mathbb{Q}}_\ell[1](1) = f_*\overline{\mathbb{Q}}_\ell[1](1)$$

Hence there is a perfect pairing:

$$f_*\overline{\mathbb{Q}}_\ell[1] * f_*\overline{\mathbb{Q}}_\ell[1] \rightarrow \delta_0 \otimes \overline{\mathbb{Q}}_\ell(-1)$$

This pairing is either symplectic or orthogonal. Were it orthogonal, the Tannakian group would be contained in GO_2 , which it is not, so the pairing is symplectic. Then the determinant is equal to the symplectic similitude character, which is $\delta_0 \otimes \overline{\mathbb{Q}_\ell}(-1)$.

The determinant of $f_*\overline{\mathbb{Q}_\ell}$ is the cyclotomic character, so the Tannakian monodromy group of the determinant is $\mathbb{G}_m$, hence the Tannakian monodromy group of $f_*\overline{\mathbb{Q}_\ell}$ maps surjectively onto $\mathbb{G}_m$ under the determinant map. Since we already know the Tannakian monodromy group of $\overline{\mathbb{Q}_\ell}$ contains SL_2 , we can conclude that it is GL_2 . $\square$

As in Proposition 6.0.95, this means the adjusted Haar measure is the Haar measure of SU_2 . So assuming Conjecture 1, the sum is equidistributed according to the Sato-Tate measure.

Now we study the distribution of Gauss and Kummer sums:

Lemma 6.0.97. *Let $\mathcal{F}$ be the sheaf described in Example 2.0.4. Let ψ be a fixed additive character. Then for $p \equiv 1$ modulo n , the eigenvalues of Frob_p on $H^1(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p\mathcal{F} \otimes \mathcal{L}_\psi)$ are minus the Gauss sums $G(\chi, \psi)$ for χ varying over multiplicative characters of order n .*

Proof. Recall that $\mathcal{F}$ is constructed using a representation of the automorphism group of the cover defined by adjoining the n th roots of unity and the n th roots of the coordinate x . Since $p \equiv 1$ modulo n , $\mathbb{Q}_p$ contains the n th roots of unity, so the cover pulled back to $\mathbb{A}_{\mathbb{Q}_p}^1$ just involves adjoining the n th roots of x . This cover has automorphism group μ_n , which is the μ_n subgroup of the original automorphism group $\mu_n \rtimes (\mathbb{Z}/n)^\times$. So the representation is the original representation restricted to $\mathbb{Z}/n$, which is the sum of all the faithful characters of μ_n .

Thus $\mathcal{F}$, pulled back to $\mathbb{A}_{\mathbb{Q}_p}^1$, is the direct sum of sheaves of generic rank 1 corresponding to faithful characters of the automorphism group of the cover defined by adjoining the n th roots of x . By Definition 2.0.3, these are the Kummer sheaves,

which correspond to order n Dirichlet characters, and all order n Dirichlet characters arise this way.

Hence $\mathcal{F}_{\mathbb{Q}_p}$ is a sum of 1-dimensional Kummer sheaves $\mathcal{L}_\chi$ for each order n Dirichlet character χ . By construction, the trace of Frobenius on the stalk of the Kummer sheaf at a point is given by the formula $\text{tr}(\text{Frob}_p, \mathcal{L}_\chi, x) = \chi(x)$.

The Kummer sheaf on $\mathbb{A}_{\mathbb{Z}_p}^1$ is locally acyclic by [Deligne and Katz, 1973, XIII Lemma 2.1.11], so $R\psi_p(\mathcal{L}_\chi)_{\mathbb{Q}_p} = (\mathcal{L}_\chi)_{\overline{\mathbb{F}}_p}$. Thus $R\Psi_p \mathcal{L}_\chi \otimes \mathcal{L}_\psi = \mathcal{L}_\chi \otimes \mathcal{L}_\psi$. This sheaf is lisse of rank 1 away from 0 and ∞ with a drop of 1 at 0 and ∞ , with a Swan conductor of 0 at 0 and 1 at ∞ , so its Euler characteristic is $-1 + 1 + 1 = 1$ by Lemma 3.0.31. Thus its étale cohomology has a single eigenvalue of Frobenius. This eigenvalue is equal to the trace of Frobenius on that cohomology group, which by the Lefschetz fixed point formula is minus the sum over all the points of the trace of Frobenius on $\mathcal{L}_\chi \otimes \mathcal{L}_\psi$, which is

$$-\sum_{x \in \mathbb{F}_p} \psi(x)\chi(x) = -G(\chi, \psi)$$

Thus $H_c^0(R\Psi_p \mathcal{F}[1])$ is a sum of 1-dimensional étale cohomology groups, each of which has an eigenvalue of Frobenius equal to minus a Gauss sum corresponding to a different character. Therefore the eigenvalues of Frobenius on the whole group are minus the asserted Gauss sums. $\square$

Lemma 6.0.98. *Let $\mathcal{F}$ be the sheaf described in Example 2.0.4 for $n = 3$. Placed in degree -1 , it is a perverse sheaf with good reduction away from 3 and ℓ . The Tannakian monodromy group is the wreath product $\mathbb{G}_m \wr \{\pm 1\}$, with its standard two-dimensional representation in which each copy of $\mathbb{G}_m$ acts on a one-dimensional subspace.*

The determinant of this representation is the cyclotomic character and the natural projection to $\{\pm 1\}$ is the quadratic character of $\mathbb{Q}(\sqrt{-3})$.

It is of geometric origin and pure of weight 1.

Proof. At primes other than 3 and ℓ , the sheaf is lisse away from 0 and vanishes at 0, so by Proposition 3.0.46 it has good reduction. Lemma 4.0.54 immediately implies it is two-dimensional.

The usual monodromy of $\mathcal{F}$ factors through the two-dimensional irreducible representation of S_3 . Tensoring this representation with the sign representation gives the same representation again. Thus tensoring $\mathcal{F}$ with the quadratic character χ of $\mathbb{Q}(\mu_3) = \mathbb{Q}(\sqrt{-3})$ gives the same sheaf. Tensoring with χ is the same as convolving with $\delta_0 \otimes \chi$ by Lemma 4.0.56. Moreover $\delta_0 \otimes \chi$ is a quadratic character of $G_{\{3,\ell\}}$ because

$$\delta_0 \otimes \chi * \delta_0 \otimes \chi = \delta_0 \otimes \chi \otimes \chi = \delta_0 = 1$$

Hence tensoring our $G_{\{3,\ell\}}$ -representation with a quadratic character yields the same representation.

Let V is a two-dimensional representation such that tensoring it with the quadratic character χ yields the same representation. Let t be the matrix that gives the isomorphism $V \cong V \otimes \chi$. Then by assumption for g in the monodromy group of V , $tg = g\chi(g)t$, so $g^{-1}tg = \chi(g)t$. Since $\chi(g)$ attains the values ± 1 , t is conjugate to $-t$. Because t is invertible, it has some nonzero eigenvalue λ , and thus $-\lambda$ is also an eigenvalue of t . Because these are distinct, they are all the eigenvalues, and after a change of basis we may write $t = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$. Then if $\chi(g) = 1$, we have $tg = gt$, so

g has the form $g = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$ and if $\chi(g) = -1$, we have $tg = -gt$, so g has the form $g = \begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix}$.

The group of all matrices of these two forms is the wreath product $\mathbb{G}_m \wr \{\pm 1\}$, acting by its monomial representation. Indeed, it is the semidirect product of the

identity component consisting of matrices of the form $\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$, which is isomorphic to $\mathbb{G}_m \times \mathbb{G}_m$, with the group generated by the matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which is isomorphic to ± 1 , where the conjugation action of ± 1 on $\mathbb{G}_m \times \mathbb{G}_m$ is by switching the two factors. This is precisely the definition of the wreath product $\mathbb{G}_m \wr \{\pm 1\}$. Furthermore, the identity component of this group consists of matrices with $\chi(g) = 1$, and the other component consist of matrices with $\chi(g) = -1$.

Therefore the Tannakian monodromy group is a subgroup of $\mathbb{G}_m \wr \{\pm 1\}$. Since the quadratic character $\delta_0 \otimes \chi$ is nontrivial, the Tannakian monodromy group contains elements of both components. To check that the Tannakian monodromy group is all of $\mathbb{G}_m \wr \{\pm 1\}$, we must check that it contains the identity component. It is sufficient to check that any Frobenius element generates a rank two torus, or in other words that the eigenvalues of Frobenius satisfy no multiplicative relations.

If $p \equiv 1 \pmod{3}$, the eigenvalues of Frobenius are minus the two Gauss sums $G(\chi, \psi)$ and $G(\bar{\chi}, \psi)$. We must check that they satisfy no multiplicative relations. Both $G(\chi, \psi)$ and $G(\bar{\chi}, \psi)$ are complex numbers of absolute value $\sqrt{p}$. Thus any relation between them would have to have the form $G(\chi, \psi)^n = G(\bar{\chi}, \psi)^n$. As $G(\chi, \psi)$ and $G(\bar{\chi}, \psi)$ are both elements of the field $\mathbb{Q}(\mu_3, \mu_p)$, this would imply that $G(\chi, \psi)$ and $G(\bar{\chi}, \psi)$ have the same valuation at each prime of that field. Because p is split in $\mathbb{Q}(\mu_3)$ and totally ramified in $\mathbb{Q}(\mu_p)$, there are two primes lying over p . We will compute the p -adic valuations of $G(\chi, \psi)$ and $G(\bar{\chi}, \psi)$ at one of these primes using Stickelberger's congruence and observe that they are not equal. To compute the valuations, we may view the characters χ and ψ as taking values in $\overline{\mathbb{Q}_p}^\times$. Each character $\chi : \mathbb{F}_p^\times \rightarrow \overline{\mathbb{Q}_p}^\times$ can be written as $\chi_a : x \mapsto [x]^a$ where $[x]$ is the Teichmuller lift for some $a \in \{0, \dots, p-1\}$, and each character $\psi : \mathbb{F}_p \rightarrow \overline{\mathbb{Q}_p}^\times$ can be written as $\psi_\zeta : x \mapsto \zeta^x$ for some p th root of

unity ζ . Stickelberger's congruence [Stickelberger, 1890, 6.15a]

$$G(\chi_a, \psi_\zeta) \equiv \frac{(\zeta - 1)^a}{a!} \pmod{(\zeta - 1)^{a+1}}$$

implies that

$$v_p(G(\chi_a, \psi_\zeta)) = \frac{a}{p-1}$$

The only two values of a that give characters of order 3 are $(p-1)/3$ and $2(p-1)/3$. Hence χ is either $\chi_{(p-1)/3}$ or $\chi_{2(p-1)/3}$ and $\bar{\chi}$ is the other character. Thus the p -adic valuation of $G(\chi, \psi)$ is $\frac{1}{3}$ or $\frac{2}{3}$ and the p -adic valuation of $G(\bar{\chi}, \psi)$ is the other fraction. Because those two valuations are distinct, there are no multiplicative relations between $G(\chi, \psi)$ and $G(\bar{\chi}, \psi)$.

Now we compute the determinant. Observe that $f_*\bar{\mathbb{Q}}_\ell$ for $f : \mathbb{A}^1 \rightarrow \mathbb{A}^1$ the map that sends x to x^3 is the sum of a rank 1 constant sheaf and a rank 2 sheaf corresponding to the standard 2-dimensional representation of the monodromy group S_3 of the cover defined by adjoining μ_3 and the 3rd roots of x . Because $\mathcal{F}$ is also the sheaf corresponding to that representation of that cover, we have $f_*\bar{\mathbb{Q}}_\ell = \bar{\mathbb{Q}}_\ell + \mathcal{F}$. So it is sufficient to compute the determinant of $f_*\bar{\mathbb{Q}}_\ell$.

Its local monodromy at ∞ is the sum of all characters of order 3, so the determinant of the local monodromy at ∞ is the product of all these characters, which is trivial. Hence by Lemma 4.0.66 the Tannakian determinant has the form $\delta_r \otimes V$. By Lemma 4.0.67, as the only singular point is 0, we have $r = 0$. Finally we will show that $V = \mathbb{Q}_\ell(-1)$ by applying Lemma 4.0.68 and comparing to $g_*\bar{\mathbb{Q}}_\ell$ for $g = x^3 + x$. As f and g are finite maps from $\mathbb{A}^1$,

$$H_c^0(\mathbb{A}_{\mathbb{Q}}^1, f_*\bar{\mathbb{Q}}_\ell[1]) = H_c^0(\mathbb{A}_{\mathbb{Q}}^1, g_*\bar{\mathbb{Q}}_\ell[1]) = H_c^0(\mathbb{A}_{\mathbb{Q}}^1, \bar{\mathbb{Q}}_\ell[1]) = 0$$

$$H_c^1(\mathbb{A}_{\mathbb{Q}}^1, f_*\bar{\mathbb{Q}}_\ell[1]) = H_c^1(\mathbb{A}_{\mathbb{Q}}^1, g_*\bar{\mathbb{Q}}_\ell[1]) = H_c^1(\mathbb{A}_{\mathbb{Q}}^1, \bar{\mathbb{Q}}_\ell[1]) = \bar{\mathbb{Q}}_\ell(-1)$$

so the other terms in Lemma 4.0.68 cancel and these sheaves have the same determinant. In Lemma 6.0.96 we saw that the determinant of $g_*\overline{\mathbb{Q}}_\ell[1]$ was $\delta_0 \otimes \mathbb{Q}_\ell(-1)$, so the determinant of $\mathcal{F}[1]$ is also $\delta_0 \otimes \mathbb{Q}_\ell(-1)$.

Finially, because $\mathcal{F}$ is a summand of $f_*\overline{\mathbb{Q}}_\ell$, it is of geometric origin and pure of weight 0, hence pure of weight 1 when placed in degree 0 $\square$

Proposition 6.0.99. *If Conjecture 1 holds, the normalized Kummer sum $K(1;p)/\sqrt{p}$ for primes $\equiv 1 \pmod{3}$ is distributed as $2 \cos \theta$ for θ uniformly distributed on S^1 .*

Proof. The abelianization of $\mathbb{G}_m \wr \{\pm 1\}$ is $\mathbb{G}_m \times \{\pm 1\}$, with the isomorphism given by the determinant character and the natural projection to $\{\pm 1\}$. The determinant character is the cyclotomic character and the projection is the quadratic character of $\mathbb{Q}(\sqrt{-3})$. Hence the Dirichlet measure on the abelianization puts mass $1/2$ on the identity and mass $1/2$ on $(1, 1)$. The first corresponds to primes $1 \pmod{3}$ and the second to primes $2 \pmod{3}$. Since we are interested in primes that are $1 \pmod{3}$, we are concerned with the conditional probability measure that is conditioned on lying over the identity. This is the Haar measure of the maximal compact subgroup of the derived subgroup $[\mathbb{G}_m \wr \mathbb{Z}/2, \mathbb{G}_m \wr \mathbb{Z}/2] = \mathbb{G}_m$. The maximal compact subgroup of $\mathbb{G}_m$ is the circle group S^1 , and the Haar measure on the circle group is the uniform measure. This $\mathbb{G}_m$ acts by the standard character on one one-dimensional subspace and the inverse character on the other subspace. The standard and inverse characters of S^1 are $e^{i\theta}$ and $e^{-i\theta}$, hence the trace is $2 \cos \theta$, for θ uniformly distributed. $\square$

This distribution is obtained, by [Heath-Brown and Patterson, 1979, Theorem 2]. However Patterson conjectured that the empirical distribution converges very slowly to this limit [Patterson, 1978, Conjecture on pp. 2]. The average over primes $< n$ of a function of the Kummer sum has not just the main term coming from the average against the limiting measure and the error term but a secondary term. This secondary term declines at the rate $n^{-1/6+o(1)}$, not at the rate $n^{-1/2+o(1)}$ that

occurs when the distribution is controlled by an L-function satisfying the Riemann hypothesis. This bias was still substantial in the range Kummer looked at, explaining his misleading numerical data. We will discuss how this behavior might generalize to other exponential sums of this type in Chapter 10.

Chapter 7

Further Examples

This chapter discusses some further cases of Conjecture 1. In most of these cases, the distribution of a_p is a problem we do not believe has been considered in the literature before. In some of these cases, our conjecture coincides with a naïve guess at the distribution of a_p . In other cases we were unable to guess the distribution before computing the Tannakian monodromy group. In each case the results of numerical tests of our predictions are reported in Chapters 9 and 10.

First consider a question about the joint distributions of two different exponential sums. If there is not a clear relationship between the sums, one might reasonably expect independence between the distributions. In one very simple case, Conjecture 1 agrees with this guess:

Proposition 7.0.100. *Let f be a polynomial in $\mathbb{Z}[x]$ with no rational roots, let K be the number field generated by the roots of f , and let $L/\mathbb{Q}$ be a Galois field extension disjoint with K . For each pair $p, x \in \mathbb{F}_p$ such that $f(x) = 0 \bmod p$ and p is unramified in L , consider the ratio $x/p \in \mathbb{R}/\mathbb{Z}$ and the conjugacy class Frob_p of $\text{Gal}(L/\mathbb{Q})$.*

If Conjecture 1 holds, the sequence of pairs $(\frac{x}{p}, \text{Frob}_p)$, parameterized by the sequence of pairs $(p, x \in \mathbb{F}_p)$ with $f(x) = 0 \bmod p$ and p unramified in L , is equidistributed according to the Haar measure on $\mathbb{R}/\mathbb{Z} \times \text{Gal}(L/\mathbb{Q})$.

Proof. By Weyl's equidistribution theorem, it is sufficient to show that for $k \in \mathbb{Z}$ and V an irreducible representation of $\text{Gal}(L/\mathbb{Q})$

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (f(x)=0)} \text{tr}(\text{Frob}_p, V) e^{2\pi i \frac{kx}{p}}}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}|} = 0$$

unless $k = 0$ and V is the trivial representation. By Chebotarev's density theorem, $\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x)=0\}|}{\pi(n)}$ is equal to the number of irreducible factors of f . Hence is sufficient to show that

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (f(x)=0)} \text{tr}(\text{Frob}_p, V) e^{2\pi i \frac{kx}{p}}}{\pi(n)} = 0$$

because this will imply that:

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (f(x)=0)} \text{tr}(\text{Frob}_p, V) e^{2\pi i \frac{kx}{p}}}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}|} = \lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (f(x)=0)} \text{tr}(\text{Frob}_p, V) e^{2\pi i \frac{kx}{p}} \pi(n)}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}| / \pi(n)}$$

is 0 divided by the number of irreducible factors of f and thus is 0. We will show that

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \sum_{x \in \mathbb{F}_p | (f(x)=0)} \text{tr}(\text{Frob}_p, V) e^{2\pi i \frac{kx}{p}}}{\sum_{p \leq n} |\{x \in \mathbb{F}_p | f(x) = 0\}|}$$

First assume $k \neq 0$. In Example 2.0.2 we constructed an object $\mathcal{F}_k$ such that

$$a_p(\mathcal{F}_k) = \sum_{x \in \mathbb{F}_p | f(x)=0} e^{2\pi i \frac{kx}{p}}$$

We may view V as a lisse geometrically constant sheaf on $\mathbb{A}^1$, so that the trace of Frobenius acting on V at every point is $\text{tr}(\text{Frob}_p, V)$. Then the tensor product $\mathcal{F}_k \otimes V$ satisfies

$$a_p(\mathcal{F}_k \otimes V) = \sum_{x \in \mathbb{F}_p | f(x)=0} \text{tr}(\text{Frob}_p, V) e^{2\pi i \frac{kx}{p}}$$

The sheaf $\mathcal{F}_k \otimes V$ is perverse because it is punctual, is pure of weight 0, and is of geometric origin because $\mathcal{F}_k$ and V are. The sheaf $\mathcal{F}_k \otimes V$ has good reduction away from ℓ , the ramified primes of L , the primes dividing the leading term of f , the primes dividing the discriminant of f , and the primes dividing k by Proposition 3.0.46. Thus by Conjecture 2, this limit is 0 as long as $\mathcal{F}_k \otimes V$ has no one-dimensional Jordan-Holder factors. But any one-dimensional factor would, by Lemma 4.0.58, have to be supported on all of $\mathbb{A}^1$ or a single rational point. The sheaf $\mathcal{F}_k$ and hence the tensor product $\mathcal{F}_k \otimes V$, are punctual, so a one-dimensional factor would have to be supported on a single rational point. Both $\mathcal{F}_k$ and $\mathcal{F}_k \otimes V$ are supported on k times the roots of f , and no rational number is k times a root at f as k is nonzero and f has no rational roots. Hence there are no one-dimensional Jordan-Holder factors of $\mathcal{F}_k \otimes V$.

It remains to treat the case when $k = 0$ and V is irreducible and nontrivial. The limit simplifies to

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \text{tr}(\text{Frob}_p, V) |\{x \in \mathbb{F}_p | f(x) = 0\}|}{\pi(n)}$$

The cardinality $|\{x \in \mathbb{F}_p | f(x) = 0\}|$ is, for all but finitely many primes, equal to the trace of Frob_p acting on a $\text{Gal}(K/\mathbb{Q})$ -representation W , the linearization of the permutation representation acting on the roots of f .

$$\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \text{tr}(\text{Frob}_p, V) |\{x \in \mathbb{F}_p | f(x) = 0\}|}{\pi(n)} = \lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \text{tr}(\text{Frob}_p, V \otimes W)}{\pi(n)}$$

By Chebotarev's density theorem, $\lim_{n \rightarrow \infty} \frac{\sum_{p \leq n} \text{tr}(\text{Frob}_p, V \otimes W)}{\pi(n)}$ is the dimension of the $\text{Gal}(KL/\mathbb{Q})$ -invariant subspace of $V \otimes W$. Since V is a representation of $\text{Gal}(L/\mathbb{Q})$ and W is a representation of $\text{Gal}(K/\mathbb{Q})$, and L and K are disjoint number fields so that $\text{Gal}(KL/\mathbb{Q}) = \text{Gal}(K/\mathbb{Q}) \times \text{Gal}(L/\mathbb{Q})$, this is the tensor product of the $\text{Gal}(L/\mathbb{Q})$ -

invariant subspace of V and the $\text{Gal}(K/\mathbb{Q})$ -invariant subspace of W . The representation V is irreducible and nontrivial, so this invariant subspace of V vanishes and the limit is zero. $\square$

Rather than starting with an equidistribution question and finding an appropriate $\mathcal{F}$ to study it with, we may start with a sheaf and find the corresponding equidistribution question. For instance, let $f(x)$ be a polynomial with rational coefficients and let $\mathcal{L}_\chi$ be the Kummer sheaf of the quadratic character. Then for $\mathcal{F} = f^*\mathcal{L}_\chi$, the exponential sum $a_p(\mathcal{F})$ is the following Jacobi symbol sum:

$$a_p(\mathcal{F}) = \sum_{x=0}^{p-1} \left(\frac{f(x)}{p} \right) e^{2\pi i \frac{x}{p}}$$

For $f(x) = x$ this is a quadratic Gauss sum, and for $f(x) = x^2 - 4ab$ this is the Kloosterman sum. In this chapter we will study this question for $f(x)$ of degree at least 3. One simple special case is the “twisted elliptic curve” sum:

$$a_p = \sum_{x=0}^{p-1} \left(\frac{4x^3 - g_2x - g_3}{p} \right) e^{2\pi i \frac{x}{p}}$$

Lemma 7.0.101. *Let $f \in \mathbb{Q}[x]$ be a polynomial of degree n without repeated roots. Then $f^*\mathcal{L}_\chi$ is a middle extension sheaf. Moreover, $f^*\mathcal{L}_\chi[1]$ is a perverse sheaf of geometric origin and pure of weight 1, with good reduction outside some finite set S of primes, that corresponds to an n -dimensional representation of G_S .*

Proof. Consider the smooth hyperelliptic curve $y^2 = f(x)$ with projection π to $\mathbb{A}^1$ along the x coordinate. Then $\pi_*\overline{\mathbb{Q}}_\ell = \overline{\mathbb{Q}}_\ell + f^*\mathcal{L}_\chi$. As π is a dominant map from a smooth curve, $\pi_*\overline{\mathbb{Q}}_\ell$ is a middle extension sheaf, hence $f^*\mathcal{L}_\chi$ is a middle extension sheaf and $f^*\mathcal{L}_\chi[1]$ is perverse. As π is a proper map from a smooth curve, $\pi_*\overline{\mathbb{Q}}_\ell$ is of geometric origin and pure of weight 0, and so $f^*\mathcal{L}_\chi$ is of geometric origin and pure of weight 0 as well. Because $f^*\mathcal{L}_\chi$ is pure of weight 0, $f^*\mathcal{L}_\chi[1]$ is pure of weight 1.

Since $f^*\mathcal{L}_\chi[1]$ has rank 1 away from the roots of f and rank 0 at the roots of f , the drop is 1 at each root. There are n distinct roots, so by Lemma 4.0.54, $f^*\mathcal{L}_\chi[1]$ is an n -dimensional object in $\mathcal{CC}_S$. $\square$

Lemma 7.0.102. *Let f be a polynomial with no repeated roots. Then $f^*\mathcal{L}_\chi[1]$ corresponds to a Lie-irreducible representation of G_S .*

Proof. Since $f^*\mathcal{L}_\chi[1]$ is lisse of rank 1, it is geometrically irreducible on the open set where it is lisse, which is the set where $f \neq 0$. By Lemma 7.0.101 $f^*\mathcal{L}_\chi[1]$ is a middle extension sheaf. Any geometrically irreducible middle extension sheaf is a geometrically irreducible perverse sheaf. Thus by Lemma 4.0.83, $f^*\mathcal{L}_\chi[1]$ corresponds to a Lie-irreducible G_S -representation. $\square$

Definition 7.0.103. For $D \in \mathbb{Q}$, let χ_D be the quadratic character of D : the 1-dimensional Galois representation that sends every element of $\text{Gal}(\overline{\mathbb{Q}}|\mathbb{Q})$ that fixes $\sqrt{D}$ to 1 and every element that does not fix $\sqrt{D}$ to -1 .

Lemma 7.0.104. *Let f be a polynomial with no repeated roots.*

If $f(x) = f(a-x)$ for some $a \in \mathbb{Q}$, then there is a perfect pairing of G_S -representations

$$f^*\mathcal{L}_\chi[1] * f^*\mathcal{L}_\chi[1] \rightarrow \delta_1(-1)$$

If $f(x) = -f(a-x)$ for some $a \in \mathbb{Q}$, then there is a perfect pairing of G_S -representations:

$$f^*\mathcal{L}_\chi[1] * f^*\mathcal{L}_\chi[1] \rightarrow \delta_a(-1) \otimes \chi_a$$

Otherwise, there is no perfect pairing $f^*\mathcal{L}_\chi[1] * f^*\mathcal{L}_\chi[1] \rightarrow V$ for any one-dimensional G_S representation V .

Proof. First assume that one of these two functional equations holds. Because $f^*\mathcal{L}_\chi$ is an irreducible middle extension sheaf, $f^*\mathcal{L}_\chi[1]$ has property $\mathcal{P}$, and hence $f^*\mathcal{L}_\chi[1] *!$

$f^*\mathcal{L}_\chi[1]$ is perverse. Let us compute the stalk of $\mathcal{H}^0(f^*\mathcal{L}_\chi[1] * f^*\mathcal{L}_\chi[1])$ at a . By the proper base change theorem, the stalk is

$$H_c^0(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi[1] \otimes s^*f^*\mathcal{L}_\chi[1]) = H_c^2(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi \otimes s^*f^*\mathcal{L}_\chi)$$

where $s(x) = a - x$.

In the case $f(x) = f(a - x)$, we have $f \circ s = f$, hence this is the same as

$$H_c^2(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi \otimes f^*\mathcal{L}_\chi)$$

By Poincare duality, $H_c^2(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi \otimes f^*\mathcal{L}_\chi)$ is the Tate twist of the geometric monodromy coinvariants of $f^*\mathcal{L}_\chi \otimes f^*\mathcal{L}_\chi$. Because $f^*\mathcal{L}_\chi \otimes f^*\mathcal{L}_\chi$ has trivial arithmetic monodromy, its top cohomology group is precisely $\overline{\mathbb{Q}}_\ell(-1)$. So the stalk at a is $\overline{\mathbb{Q}}_\ell(-1)$ and hence we have a nontrivial map $f^*\mathcal{L}_\chi[1] * f^*\mathcal{L}_\chi[-1] \rightarrow \delta_a(-1)$. By Lemma 7.0.102, $f^*\mathcal{L}_\chi[1]$ is an irreducible representation of G_S , and hence this map is a perfect pairing - otherwise the kernel would give a nontrivial subrepresentation.

In the case $f(x) = -f(a - x)$, the argument is almost the same. Now $f \circ s = [-1] \circ f$ where $[-1]$ is the map $x \rightarrow -x$. So we have

$$H_c^2(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi \otimes s^*f^*\mathcal{L}_\chi) = H_c^2(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi \otimes f^*[-1]^*\mathcal{L}_\chi)$$

(as $[-1]^*\mathcal{L}_\chi = \chi_{-1} \otimes \mathcal{L}_\chi$)

$$= H_c^2(\mathbb{A}_{\overline{\mathbb{Q}}}^1, \chi_{-1} \otimes f^*\mathcal{L}_\chi \otimes f^*\mathcal{L}_\chi)$$

The geometric monodromy of $\chi_{-1} \otimes f^*\mathcal{L}_\chi \otimes f^*\mathcal{L}_\chi$ is trivial. The geometric monodromy invariants of $\chi_{-1} \otimes f^*\mathcal{L}_\chi \otimes f^*\mathcal{L}_\chi$, viewed as a Galois representation, equal χ_{-1} . Thus the top cohomology group of that sheaf is $\overline{\mathbb{Q}}_\ell(-1) \otimes \chi_{-1}$. As before, this

gives a nontrivial map $f^*\mathcal{L}_\chi[1] * f^*\mathcal{L}_\chi[1] \rightarrow \delta_a(-1) \otimes \chi_{-1}$, which must be a perfect pairing as $f^*\mathcal{L}_\chi[1]$ is irreducible.

Now assume that there exists a perfect pairing $f^*\mathcal{L}_\chi[1] * f^*\mathcal{L}_\chi[1] \rightarrow V$ for a one-dimensional G_S -representation V . It follows that $f^*\mathcal{L}_\chi[1] = V * (f^*\mathcal{L}_\chi[1])^\vee$. By Proposition 3.0.22,

$$f^*\mathcal{L}_\chi[1] = V * (f^*\mathcal{L}_\chi[1])^\vee = V * D([-1]^* f^*\mathcal{L}_\chi[1]),$$

Let $\alpha_1, \dots, \alpha_n$ be the singularities of $f^*\mathcal{L}_\chi$: equivalently, the roots of f . The singularities of $D([-1]^* f^*\mathcal{L}_\chi[1])$ are the same as the singularities of $[-1]^* f^*\mathcal{L}_\chi$, as by double-duality a sheaf is lisse at a point if and only if its Verdier dual is lisse at that point. So the singularities of $D([-1]^* f^*\mathcal{L}_\chi[1])$ are $-\alpha_1, \dots, -\alpha_n$. By Lemma 4.0.58, V has a unique singular point. Let a be the singular point of V . Then by Lemma 3.0.32 the only possible singularities of $V * D([-1]^* f^*\mathcal{L}_\chi[1])$ are $a - \alpha_1, \dots, a - \alpha_n$. By assumption, $V * D([-1]^* f^*\mathcal{L}_\chi[1])$ is isomorphic to $f^*\mathcal{L}_\chi[1]$ whose singularities are $\alpha_1, \dots, \alpha_n$, hence

$$\{\alpha_1, \dots, \alpha_n\} \subseteq \{a - \alpha_1, \dots, a - \alpha_n\}.$$

Because both these sets have n elements and thus

$$\{\alpha_1, \dots, \alpha_n\} = \{a - \alpha_1, \dots, a - \alpha_n\}.$$

Therefore $f(x)$ and $f(a - x)$ have the same roots. Since $f(x)$ and $f(a - x)$ have no repeated roots, they must be equal up to scalar multiplication, say $f(x) = \epsilon f(a - x)$. Because the leading coefficient of $f(a - x)$ is $(-1)^n$ times the leading coefficient of $f(x)$, $\epsilon = (-1)^n = \pm 1$, hence either $f(x) = a - x$ or $f(x) = -f(a - x)$. $\square$

Lemma 7.0.105. *Let f be a polynomial of degree n with no repeated roots, $f(x) = ax^n + bx^{n-1} + \dots$. Then the determinant of $f^*\mathcal{L}_\chi$ as a G_S -representation is: (n even)*

$$\delta_{-b/a} \left(-\frac{n}{2} \right)$$

(n odd)

$$\mathcal{L}_\chi[1] * \delta_{-b/a} \left(-\frac{n-1}{2} \right) \otimes \chi_a$$

Proof. By Lemma 4.0.66, $\mathcal{L}_\chi$ appears in the determinant of $f^*\mathcal{L}_\chi[1]$ if and only if the determinant of the local monodromy representation of $f^*\mathcal{L}_\chi$ at ∞ is nontrivial. The local monodromy representation is one-dimensional, so its determinant is nontrivial if and only if it is nontrivial. The local monodromy comes from the cover defined by adjoining a square root of $ax^n + \dots$, which is trivial for n even and nontrivial for n odd.

$f^*\mathcal{L}_\chi$ is a middle extension sheaf and each singular point appears with a drop of 1. So by Lemma 4.0.67, the translation factor appearing in the determinant of $f^*\mathcal{L}_\chi[1]$ is the sum of the singular points of $f^*\mathcal{L}_\chi[1]$. The singular points are the roots of f , and their sum is $-\frac{b}{a}$.

Finally, we apply Lemma 4.0.68 and compare $f^*\mathcal{L}_\chi$ with another sheaf with the same decomposition representation at ∞ . Split into two cases, n even or n odd:

n even: We have already seen that the determinant of $f^*\mathcal{L}_\chi[1]$ is of the form $\delta_{-\frac{b}{a}} \otimes V$ for some one-dimensional Galois representation V . The sheaf $f^*\mathcal{L}_\chi$ is unramified at ∞ , and the underlying Galois representation is χ_a . So it has the same decomposition group representation at ∞ as the constant sheaf $\overline{\mathbb{Q}}_\ell \otimes \chi_a$. The constant sheaf $\overline{\mathbb{Q}}_\ell \otimes \chi_a$ has Tannakian dimension 0, so its determinant in the Tannakian category is 1. By Lemma 4.0.68:

$$V \otimes \det H_c^1 \left(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi[1] \right) \otimes \det H_c^0 \left(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^*\mathcal{L}_\chi[1] \right)^{-1}$$

$$= \det H_c^1 \left(\mathbb{A}_{\overline{\mathbb{Q}}}^1, \overline{\mathbb{Q}}_{\ell} \otimes \chi_a[1] \right) \otimes \det H_c^0 \left(\mathbb{A}_{\overline{\mathbb{Q}}}^1, \overline{\mathbb{Q}}_{\ell} \otimes \chi_a[1] \right)^{-1}$$

We will compute all the cohomology groups appearing in the above formula.

$H_c^1(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^* \mathcal{L}_{\chi}[1])$ vanishes since $f^* \mathcal{L}_{\chi}$ has nontrivial local monodromy at the roots of f , hence nontrivial global monodromy.

$H_c^0(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^* \mathcal{L}_{\chi}[1]) = H_c^1(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^* \mathcal{L}_{\chi})$ can be computed by taking the middle extension of $j_* f^* \mathcal{L}_{\chi}$ along the inclusion $j : \mathbb{A}^1 \rightarrow \mathbb{P}^1$. The middle extension has stalk χ_a at ∞ , so $H_c^1(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^* \mathcal{L}_{\chi})$ is the extension of $H_c^1(\mathbb{P}_{\overline{\mathbb{Q}}}^1, j_* f^* \mathcal{L}_{\chi})$ by χ_a . Furthermore $j_* f^* \mathcal{L}_{\chi}$ is orthogonally self-dual so $H_c^1(\mathbb{P}_{\overline{\mathbb{Q}}}^1, j_* f^* \mathcal{L}_{\chi})$ is symplectically self-dual up to a Tate twist. Its dimension is $n - 2$ so its determinant is $\mathbb{Q}_{\ell}(-\frac{n-2}{2})$, and $\det H_c^0(\mathbb{A}_{\overline{\mathbb{Q}}}^1, f^* \mathcal{L}_{\chi}[1]) = \mathbb{Q}_{\ell}(-\frac{n-2}{2}) \otimes \chi_a$.

$$H_c^1(\mathbb{A}_{\overline{\mathbb{Q}}}^1, \mathbb{Q}_{\ell} \otimes \chi_a[1]) = H_c^2(\mathbb{A}_{\overline{\mathbb{Q}}}^1, \mathbb{Q}_{\ell} \otimes \chi_a) = \mathbb{Q}_{\ell}(-1) \otimes \chi_a.$$

$$H_c^0(\mathbb{A}_{\overline{\mathbb{Q}}}^1, \mathbb{Q}_{\ell} \otimes \chi_a[1]) = H_c^1(\mathbb{A}_{\overline{\mathbb{Q}}}^1, \mathbb{Q}_{\ell} \otimes \chi_a) \text{ vanishes.}$$

Multiplying it all together, we see the Galois representation appearing in the determinant satisfies

$$V \otimes \left(\mathbb{Q}_{\ell} \left(-\frac{n-2}{2} \right) \otimes \chi_a \right)^{-1} = \mathbb{Q}_{\ell}(-1) \otimes \chi_a$$

The unique solution is $V = \mathbb{Q}_{\ell}(-n)$, as desired.

n odd: $f(x) = ax^n + \dots$, and n is odd, so in the local field $\mathbb{Q}((x^{-1}))$, $f(x)$ is equal to ax times a perfect square. Hence the decomposition group representation of $f^* \mathcal{L}_{\chi}$ is the same as the decomposition group representation of $\mathcal{L}_{\chi}(ax) = \mathcal{L}_{\chi} \otimes \chi_a$. Let $A = f^* \mathcal{L}_{\chi}[1]$ and $B = \mathcal{L}_{\chi} \otimes \chi_a$. As in Lemma 4.0.68, let V_A and V_B be the Galois representations appearing in the Tannakian determinants of A and B . Then B is a rank 1 object in the Tannakian category, so the Tannakian determinant of B is B

and $V_B = \chi_a$. By Lemma 4.0.68:

$$\begin{aligned} & V_A \otimes \det H_c^1 \left(\mathbb{A}_{\mathbb{Q}}^1, f^* \mathcal{L}_\chi[1] \right) \otimes \det H_c^0 \left(\mathbb{A}_{\mathbb{Q}}^1, f^* \mathcal{L}_\chi[1] \right)^{-1} \\ &= \chi_a \otimes \det H_c^1 \left(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{L}_\chi[1] \otimes \chi_a \right) \otimes \det H_c^0 \left(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{L}_\chi[1] \otimes \chi_a \right)^{-1} \end{aligned}$$

$H_c^1(\mathbb{A}_{\mathbb{Q}}^1, f^* \mathcal{L}_\chi[1])$ vanishes as $f^* \mathcal{L}_\chi$ has nontrivial local monodromy at the roots of f , hence nontrivial global monodromy.

Let j be the inclusion $\mathbb{A}^1 \rightarrow \mathbb{P}^1$. Then $H_c^0(\mathbb{A}_{\mathbb{Q}}^1, f^* \mathcal{L}_\chi[1]) = H_c^1(\mathbb{A}_{\mathbb{Q}}^1, f^* \mathcal{L}_\chi) = H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_! f^* \mathcal{L}_\chi)$. Observe that $j_! f^* \mathcal{L}_\chi$ is a middle extension sheaf, because $f^* \mathcal{L}_\chi$ is a middle extension on $\mathbb{A}^1$ and $f^* \mathcal{L}_\chi$ has no local monodromy invariants at ∞ . As $f^* \mathcal{L}_\chi$ is an orthogonally self-dual middle extension sheaf, its cohomology is symplectically self-dual up to a Tate twist. As the rank of the cohomology of $f^* \mathcal{L}_\chi$ is $n - 1$, the determinant is $\mathbb{Q}_\ell \left(-\frac{n-1}{2} \right)$.

$H_c^1(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{L}_\chi[1] \otimes \chi_a) = H_c^2(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{L}_\chi \otimes \chi_a)$ and $H_c^0(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{L}_\chi[1] \otimes \chi_a[1]) = H_c^1(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{L}_\chi \otimes \chi_a)$ both vanish, the first because $\mathcal{L}_\chi$ has nontrivial monodromy and the second because the Euler characteristic vanishes.

Hence the equation becomes $V_A \otimes \mathbb{Q}_\ell \left(-\frac{n-1}{2} \right) = \chi_a$ or $V_A = \mathbb{Q}_\ell \left(\frac{n-1}{2} \right) \otimes \chi_a$ as desired.

□

Proposition 7.0.106. *Let $f(x) = 4x^3 - g_2x - g_3$ be a cubic polynomial with no repeated roots. Then $f^* \mathcal{L}_\chi[1]$ has Tannakian monodromy group GL_3 unless $g_3 = 0$, in which case it has Tannakian monodromy group GO_3 . Let $V = \mathcal{L}_\chi[1] \otimes \chi_{-1}$. Then Tannakian monodromy group of $f^* \mathcal{L}_\chi[1] * V^{-1}$ is SL_3 unless $g_3 = 0$, in which case it is SO_3 . Thus, assuming Conjecture 1, the exponential sum:*

$$a_p = \sum_{x=0}^{p-1} \left(\frac{4x^3 - g_2x - g_3}{p} \right) e^{2\pi i \frac{x}{p}}$$

divided by $\sqrt{p}$ for $p \equiv 1 \pmod{4}$ and $-i\sqrt{p}$ for $p \equiv 3 \pmod{4}$, is equidistributed according to the trace of the Haar measure on SU_3 or SO_3 respectively.

Proof. By Lemma 7.0.101, $f^*\mathcal{L}_\chi[1]$ is a 3-dimensional representation of G_S . By Lemma 7.0.102, this representation is Lie-irreducible. Any Lie-irreducible subgroup of GL_3 contains as a maximal semisimple subgroup either SO_3 or SL_3 . If the maximal semisimple subgroup is SO_3 then the representation is self-dual, possibly with a similitude character, and if the maximal semisimple subgroup is SL_3 then it is not self-dual with any similitude character.

By Lemma 7.0.104, $f^*\mathcal{L}_\chi[1]$ is self-dual up to a twist if and only if f satisfies an identity of the form $f(x) = \pm f(a - x)$. Since $f(x) = 4x^3 - g_2x - g_3$, $f(a - x) = -4x^3 + 12ax^2 + \dots$, so such an identity can only occur if $a = 0$ and the identity takes the form $f(x) = -f(-x)$. We have $f(x) = -f(-x)$ if and only if $g_3 = 0$. If $g_3 = 0$ then $f^*\mathcal{L}_\chi[1]$ is self-dual after twisting by $\delta_0(-1) \otimes \chi_{-1}$ by Lemma 7.0.104.

By Lemma 7.0.105, the determinant character of $f^*\mathcal{L}_\chi[1]$ is $\mathcal{L}_\chi[1](1)$ and thus the determinant has Tannakian monodromy group $\mathbb{G}_m$.

In the case $g_3 \neq 0$, the Tannakian monodromy group contains SL_3 . As the determinant has Tannakian monodromy group $\mathbb{G}_m$, $f^*\mathcal{L}_\chi[1]$ has monodromy group GL_3 .

Because the action of G_S on $V = \mathcal{L}_\chi[1] \otimes \chi_{-1}$ factors through the abelianization of G_S , the Tannakian monodromy group of $f^*\mathcal{L}_\chi[1] * V^{-1}$ contains SL_3 and is contained in GL_3 . We will show that the Tannakian monodromy group of $f^*\mathcal{L}_\chi[1] * V^{-1}$ is actually SL_3 . It is sufficient to show that the determinant of $f^*\mathcal{L}_\chi[1] * V^{-1}$ is trivial

As the determinant character of $f^*\mathcal{L}_\chi[1]$ is $\delta_0(-1) * \mathcal{L}_\chi[1]$,

$$\det(f^*\mathcal{L}_\chi[1] * V^{-1}) = \delta_0(-1) * \mathcal{L}_\chi[1] * V^{-1} * V^{-1} * V^{-1}$$

$$V * V = \mathcal{L}_\chi[1] * \mathcal{L}_\chi[1] = \delta_0(-1) \otimes \chi_{-1}$$

$$V * V * V = \delta_0(-1) \otimes \chi_{-1} * V = \delta_0(-1) \otimes \chi_{-1} * \mathcal{L}_\chi[1] * \delta_0 \otimes \chi_{-1} = \delta_0(-1) * \mathcal{L}_\chi[1]$$

Thus $\det(f^*\mathcal{L}_\chi[1] * V^{-1})$ is indeed trivial.

In the case $g_3 = 0$, SO_3 is contained in the monodromy group and the monodromy group is contained in GO_3 . Because the determinant character is not finite, the monodromy group is GO_3 .

Because the action of G_S on $V = \mathcal{L}_\chi[1] * \chi_{-1}$ factors through the abelianization of G_S , the Tannakian monodromy group of $f^*\mathcal{L}_\chi[1] * V^{-1}$ contains SO_3 and is contained in GO_3 . We will show that the Tannakian monodromy group of $f^*\mathcal{L}_\chi[1] * V^{-1}$ is actually SO_3 .

To do this it is sufficient to show that the determinant and similitude characters of $f^*\mathcal{L}_\chi[1] * V^{-1}$ are both trivial. We have already shown that the determinant character is trivial. The similitude character of $f^*\mathcal{L}_\chi[1]$ is $\delta_0(-1) \otimes \chi_{-1}$ by Lemma 7.0.104, and therefore the similitude character of $f^*\mathcal{L}_\chi[1] * V^{-1}$ is $\delta_0(-1) \otimes \chi(-1) * V^{-1} * V^{-1}$. As

$$V * V = \mathcal{L}_\chi[1] * \mathcal{L}_\chi[1] = \delta_0(-1) \otimes \chi(-1)$$

this is trivial.

Hence the monodromy group of $f^*\mathcal{L}_\chi[1] * V^{-1}$ is either SL_3 or SO_3 , the first when $g_3 \neq 0$ and the second when $g_3 = 0$.

The trace of Frob_p acting on $f^*\mathcal{L}_\chi[1] * V^{-1}$ is the exponential sum a_p divided by the eigenvalue of Frob_p acting on V , which is $-\sqrt{p}$ if $p \equiv 1 \pmod{4}$ and $i\sqrt{p}$ if $p \equiv 3 \pmod{4}$.

Because $f^*\mathcal{L}_\chi[1]$ is of pure of weight 1, and V is pure of weight 1, $f^*\mathcal{L}_\chi[1] * V^{-1}$ is pure of weight 0. Because $f^*\mathcal{L}_\chi[1]$ is of geometric origin, and V is of geometric origin, by Corollary 4.0.76, $f^*\mathcal{L}_\chi[1] * V^{-1}$ is of geometric origin. Hence assuming Conjecture 1, the trace of Frob_p on $f^*\mathcal{L}_\chi[1] * V^{-1}$ is equidistributed according to the adjusted Haar measure on SL_3 or SO_3 respectively. Because those groups are perfect,

the adjusted Haar measure is equal to the Haar measure, and we have the desired equidistribution statement. $\square$

Another natural sum of this type has the form

$$S(f(x), p) = \sum_{x \in \mathbb{F}_p} e^{2\pi i \frac{f(x)}{p}}$$

For $f(x)$ linear, this sum is 0. For $f(x)$ quadratic, this is a quadratic Gauss sum. For cubics and beyond, the distribution is not so simple. As we discussed in Lemma 6.0.96, for $f(x)$ cubic and normalized to take the form $ax^3 + x$ we expect Sato-Tate equidistribution. Now consider the quartic and higher degree cases. Recall from Example 2.0.7 that this exponential sum corresponds to the sheaf $\mathcal{F} = f_* \overline{\mathbb{Q}}_\ell$, so the first step is to compute the Tannakian monodromy group of $\mathcal{F}[1]$.

Lemma 7.0.107. *Let $f(x) \in \mathbb{Q}[x]$ be a polynomial of degree n , and thus a map $f : \mathbb{A}_{\mathbb{Q}}^1 \rightarrow \mathbb{A}_{\mathbb{Q}}^1$. Let $\mathcal{F} = f_* \overline{\mathbb{Q}}_\ell$. Then $\mathcal{F}[1]$ is a perverse sheaf of geometric origin, pure of weight 1 with good reduction away from some finite set S of primes. The perverse sheaf $\mathcal{F}[1]$ corresponds to an $n - 1$ -dimensional representation of G_S .*

Proof. By [Katz, 1990, Lemma 7.3.2], $\mathcal{F}$ is a middle extension sheaf, so $\mathcal{F}[1]$ is perverse. The sheaf $\mathcal{F}$ is the restriction to $\mathbb{A}_{\mathbb{Q}}^1$ of a sheaf on $\mathbb{A}_{\mathbb{Z}[1/N]}^1$, so by Proposition 3.0.46 $\mathcal{F}[1]$ has good reduction away from finitely many primes. Thus $\mathcal{F}[1]$ is a representation of G_S for some S . The sheaf $\mathcal{F}[1]$ is of geometric origin and pure of weight 0 as f is a proper morphism from a smooth variety. Hence $\mathcal{F}[1]$ is pure of weight 1.

The Euler characteristic of this sheaf is $\chi(\mathbb{A}^1, f_* \overline{\mathbb{Q}}_\ell[1]) = \chi(\mathbb{A}^1, \overline{\mathbb{Q}}_\ell[1]) = -1$. Consider the equation of Lemma 3.0.31. The left, side, the Euler characteristic, is -1 , so the right side must sum to -1 . As $\mathcal{F}[1]$ is a tamely ramified sheaf concentrated in degree -1 , the Swan conductor and $\mathcal{H}^0(A)$ terms in the formula vanish. The only remaining terms are minus the rank of $\mathcal{F}$ and the sum over all singular points of the

drop of $\mathcal{F}[1]$ at that point. The rank of $\mathcal{F}$ is n so the sum over all singular points of the drop at that point must be $n - 1$. Hence by Lemma 4.0.54, the Tannakian dimension of $\mathcal{F}[1]$ is $n - 1$. $\square$

To compute the arithmetic part of the determinant of $\mathcal{F}[1]$, we have the following lemma computing the action of Frobenius on this determinant:

Lemma 7.0.108. *Let $f(x) \in \mathbb{F}_q[x]$ be a polynomial of degree n and thus a map $f : \mathbb{A}_{\mathbb{F}_q}^1 \rightarrow \mathbb{A}_{\mathbb{F}_q}^1$. Assume that the characteristic of $\mathbb{F}_q$ does not divide n . Let a be the leading coefficient of f and let r be the sum of $f(\alpha_i)$ over α_i the roots of $\frac{df}{dx}$, with multiplicity.*

Then if n is odd:

$$\det(\text{Frob}_q, H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi)) = \psi(r) q^{(n-1)/2}$$

If n is even:

$$\det(\text{Frob}_q, H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi)) = -\psi(r) q^{(n-2)/2} G(\chi_2, \psi) \chi_2 \left(\frac{an}{2} \right)$$

where χ_2 is the quadratic character of $\mathbb{F}_q^\times$.

Proof. We apply the Hasse-Davenport method (compare e.g. [Katz, 2012, Theorem 21.1]. The L-function of the Galois representation $f^* \mathcal{L}_\psi$ over the field $\mathbb{F}_q(t)$ is a polynomial $\det(1 - T \text{Frob}_q, H^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi))$. Since $H^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi)$ has dimension $n - 1$, its determinant is $(-1)^{n-1}$ times the degree $n - 1$ term of that L -function. The degree $n - 1$ term of that L function can be written as the sum over monic polynomials g of degree $n - 1$ of ψ of the sum of $f(\alpha)$ over the roots α of g . The sum of $f(\alpha)$ over the roots α of g can be expressed, using Newton's identities, as a polynomial in the coefficients of g . Let $s_1, \dots, s_{n-1}$ be the coefficients of g , equal to the elementary symmetric polynomials of the roots of g . Then the degree $n - 1$ term of the L-function

is

$$\sum_{s_1, \dots, s_{n-1} \in \mathbb{F}_q} \psi(P(s_1, \dots, s_{n-1}))$$

for a suitable polynomial P . In the usual grading on the ring of symmetric functions, s_i has degree i . In this grading P is of degree $\leq n$. Hence if $i > \frac{n}{2}$, P is a linear function of the variable s_i , with coefficient of degree $\leq n - i$. Thus the sum over just the variable s_i , is zero unless $\partial P / \partial s_i = 0$, in which case it is p times the value at any fixed value of s_i . Applying this to s_i for all $i > \frac{n}{2}$ gives the identity

$$\sum_{s_1, \dots, s_{n-1} \in \mathbb{F}_q} \psi(P(s_1, \dots, s_{n-1})) = \sum_{s_1, \dots, s_{\lfloor \frac{n}{2} \rfloor} \mid \frac{\partial P}{\partial s_{\lfloor \frac{n}{2} \rfloor + 1}} = 0, \dots, \frac{\partial P}{\partial s_{n-1}} = 0} q^{\lfloor \frac{n-1}{2} \rfloor} \psi\left(P(s_1, \dots, s_{\lfloor \frac{n}{2} \rfloor}, *, \dots, *)\right)$$

where $*$ denotes that we may plug any element of $\mathbb{F}_q$ into the formula and get the same result.

This is a sum over the solutions of the system of polynomial equations $\left(\frac{\partial P}{\partial s_{n-i}} = 0\right)_{i > \frac{n}{2}}$ in the variables $s_1, \dots, s_{\lfloor \frac{n-1}{2} \rfloor}$, with an additional sum in the variable $s_{\frac{n}{2}}$ if n is even. As we mentioned, $\frac{\partial P}{\partial s_i}$ is of degree $\leq n - i$. Hence $\frac{\partial P}{\partial s_i}$ is a polynomial in $s_1, \dots, s_{n-i-1}$ plus a constant times s_{n-i} . If these constants are nonzero for $i \in \{\lfloor \frac{n}{2} \rfloor + 1, \dots, n\}$ then these equations have a unique solution, as inductively s_{n-i} is uniquely determined by s_{n-j} for $j > i$ and the condition that $\frac{\partial P}{\partial s_i} = 0$.

In that case our sum reduces to a single value of $\psi\left(P(s_1, \dots, s_{\lfloor \frac{n}{2} \rfloor}, *, \dots, *)\right)$ (n odd) or a sum in one variable (n even).

Indeed, we will show that the coefficient of s_{n-i} in $\frac{\partial P}{\partial s_i}$ is nonzero. The constant is equal to $\frac{\partial^2 P}{\partial s_i \partial s_{n-i}}$ as no other terms of the polynomial contribute to that partial derivative, so it is sufficient to show that partial derivative is nonzero.

P is a linear combination of the first n power sum symmetric functions. The k th power sum symmetric function is homogenous of degree k , hence its derivative vanishes unless $k = n$. The coefficient of the n th power sum symmetric function is

the leading term a , which is nonzero. So it is sufficient to show, for N_k the k th power sum symmetric function, that $\frac{\partial^2 N_n}{\partial s_i \partial s_{n-i}} \neq 0$.

By Newton's identities (e.g. [Mead, 1992, pp. 749])

$$N_n = \sum_{k=1}^{n-1} (-1)^{n-1-k} s_{n-k} N_k + (-1)^{n-1} n s_n$$

we have:

$$\begin{aligned} \frac{\partial^2 N_n}{\partial s_i \partial s_{n-i}} &= \frac{\partial^2 \sum_{k=1}^{n-1} (-1)^{n-1-k} s_{n-k} N_k}{\partial s_i \partial s_{n-i}} \\ &= \sum_{k=1}^{n-1} (-1)^{n-1-k} \frac{\partial^2 s_{n-k} N_k}{\partial s_i \partial s_{n-i}} = \sum_{k=i, n-i} (-1)^{n-1-k} \frac{\partial^2 s_{n-k} N_k}{\partial s_i \partial s_{n-i}} \\ &= (-1)^{n-1-i} \frac{\partial N_i}{\partial s_i} + (-1)^{i-1} \frac{\partial N_{n-i}}{\partial s_{n-i}} = (-1)^{n-1-i} (-1)^{i-1} i + (-1)^{i-1} (-1)^{n-i-1} (n-i) = (-1)^n n \neq 0 \end{aligned}$$

For n odd, the degree $n-1$ coefficient of the L-function has been reduced to a sum over a single point. For n even, it has been reduced to a sum in one variable $s_{n/2}$.

In the odd case, that point is a solution of the equations $\frac{\partial P}{\partial s_i} = 0$ for $i > n/2$. In fact, we will take it to be a point that solves that equation for all i . The normalized derivative $g_0 = \frac{1}{an} \frac{df(x)}{dx}$ is such a solution. To prove this, first assume that $\frac{df(x)}{dx}$ is squarefree. Let α_1, α_{n-1} be the roots of g . We may write $s_1, \dots, s_{n-1}$ as polynomials in $\alpha_1, \dots, \alpha_{n-1}$. Then by the chain rule

$$\frac{\partial P(s_1, \dots, s_{n-1})}{\partial \alpha_j} = \sum_{i=1}^{n-1} \frac{\partial s_i}{\partial \alpha_j} \frac{\partial P(s_1, \dots, s_{n-1})}{\partial s_i}$$

but

$$\frac{\partial P(s_1, \dots, s_{n-1})}{\partial \alpha_j} = \frac{\partial \sum_{i=1}^{n-1} f(\alpha_i)}{\partial \alpha_j} = \frac{df(x)}{dx}(\alpha_i) = 0$$

when the α_i are roots of g_0 , and the Jacobian matrix $M_{ij} = \frac{\partial s_i}{\partial \alpha_j}$ is invertible at g_0 because its determinant is the discriminant of g_0 , which by assumption is nonzero.

So $\frac{\partial P(s_1, \dots, s_{n-1})}{\partial s_i} = 0$ at g_0 for all i . This handles the case when $\frac{df(x)}{dx}$ is squarefree. For arbitrary f , the derivatives $\frac{\partial P}{\partial s_i}(\frac{1}{an} \frac{df}{dx}) = 0$ must vanish because they are polynomials in the coefficients of f that we have just seen vanish generically.

At this point the value of $P(s_1, \dots, s_{n-1})$ is $\psi(r)$, so our original sum is $q^{\frac{n-1}{2}} \psi(r)$, as claimed.

In the even case, P is a quadratic function of the remaining variable $s_{n/2}$. So the sum $\sum_{s_1, \dots, s_{\lfloor \frac{n}{2} \rfloor} | \frac{\partial P}{\partial s_{\lfloor \frac{n}{2} \rfloor + 1}} = 0, \dots, \frac{\partial P}{\partial s_{n-1}} = 0} \psi \left(P(s_1, \dots, s_{\lfloor \frac{n}{2} \rfloor}, *, \dots, *) \right)$ is a Gauss sum. We can choose the $*$ s so as to set $x^{n-1} + \sum_{i=1}^{n-1} (-1)^i s_i x^{n-1-i} = \frac{1}{an} \frac{df}{dx} + tx^{n/2}$ for a parameter t , because as we saw in the previous case $\frac{1}{an} \frac{df}{dx}$ satisfies the equations $\frac{\partial P}{\partial s_i} = 0$. Hence P is a quadratic function in t , whose value at 0 is r , and whose derivative at 0 is 0, so

$$P = r + \frac{\partial^2 P}{\partial s_{\frac{n}{2}}^2} \frac{t^2}{2} = r + ant^2/2$$

by our earlier calculations of the derivative $\frac{\partial^2 P}{\partial s_{\frac{n}{2}}^2}$. The Gauss sum $\sum_{t \in \mathbb{F}_q} \psi(r + ant^2/2)$ is equal to $G(\chi, \psi) \psi(r) \chi_2\left(\frac{an}{2}\right)$ and hence the total sum is $q^{\frac{n-2}{2}} G(\chi, \psi) \psi(r) \chi_2\left(\frac{an}{2}\right)$. Adding the minus sign from converting between the L-function and the determinant gives the stated formula. $\square$

Lemma 7.0.109. *Let $f(x) \in \mathbb{Z}[x]$ be a polynomial of degree n , and thus a map $f : \mathbb{A}_{\mathbb{Q}}^1 \rightarrow \mathbb{A}_{\mathbb{Q}}^1$. Let a be the leading coefficient of f and let r be the sum of $f(\alpha_i)$ over α_i the roots, with multiplicity of $\frac{df}{dx}$. Let $\mathcal{F} = f_* \overline{\mathbb{Q}}_{\ell}$. Then the Tannakian determinant of $\mathcal{F}[1]$ is:*

$$\delta_r \left(\frac{n-1}{2} \right)$$

(n odd)

$$\delta_r \left(\frac{n-2}{2} \right) \otimes \chi_{\frac{an}{2}} * \mathcal{L}_{\chi}$$

(n even)

Proof. By [Katz, 1990, Lemma 7.3.2], $\mathcal{F}$ is a middle extension sheaf, so $\mathcal{F}[1]$ is perverse. The sheaf $\mathcal{F}$ is the restriction to $\mathbb{A}_{\mathbb{Q}}^1$ of a sheaf on $\mathbb{A}_{\mathbb{Z}}^1$, so by Proposition 3.0.46, $\mathcal{F}[1]$ has good reduction away from finitely many primes. Thus $\mathcal{F}[1]$ is a representation of G_S for some S . Furthermore $\mathcal{F}[1]$ is of geometric origin and pure of weight 0 as f is a proper morphism from a smooth variety. Thus $\mathcal{F}[1]$ is pure of weight 1.

As a representation of the local monodromy group at ∞ , $f_*\overline{\mathbb{Q}}_\ell$ is the linearization of the monodromy representation of the fibers of f . There is a single branch point of f lying over ∞ , so the local monodromy at ∞ is an n -cycle, so the determinant of the local monodromy is the sign of an n -cycle, which is 1, trivial, for n odd and, -1 , nontrivial, for n even. Hence by Lemma 4.0.66, $\mathcal{L}_\chi$ appears in the determinant for n even but not n odd.

The sheaf has singularities at the critical values of $f(\alpha_i)$. Let y be a critical value. The drop at y is the degree of f minus the number of roots of $f(x) - y$. Hence the drop at y is equal to the sum over the roots α of $f(x) - y$ of the multiplicity of α in $f(x) - y$ minus 1. Hence the drop at y equal to the sum over α satisfying both $f(\alpha) - y = 0$ and $\frac{df}{dx}(\alpha) = 0$ of the multiplicity of α in $\frac{df}{dx}$. Hence the sum over y of y times the drop at y is equal to the sum, with multiplicity, of $f(\alpha)$ over the roots α of $\frac{df}{dx}$, which is r . Therefore by Lemma 4.0.67 the skyscraper sheaf appearing in the determinant is δ_r .

So the determinant of $\mathcal{F}[1]$ is of the form $\delta_r \otimes V$ (n odd) or $\delta_r \otimes V * \mathcal{L}_\chi[1]$ (n even) for some 1-dimensional Galois representation V . Thus it is of the form $\delta_r \left(\frac{n-1}{2}\right) \otimes W$ (n odd) or $\delta_r \left(\frac{n-2}{2}\right) \otimes \chi_{\frac{an}{2}} \otimes W * \mathcal{L}_\chi[1]$ (n even) for some 1-dimensional Galois representation W . It remains to show that W is trivial.

We will do this by directly finding the action of $\text{Frob}_{p,\psi}$ on $\det \mathcal{F}[1]$, which is equal to the eigenvalue of Frobenius acting on $\det H^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p \mathcal{F}[1] \otimes \mathcal{L}_\psi)$.

$$H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p \mathcal{F}[1] \otimes \mathcal{L}_\psi) = H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, R\Psi_p f_* \overline{\mathbb{Q}}_\ell[1] \otimes \mathcal{L}_\psi)$$

(as $R\Psi_p$ commutes with finite pushforward)

$$= H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, f_* R\Psi_p \overline{\mathbb{Q}}_\ell[1] \otimes \mathcal{L}_\psi)$$

(by local acyclicity of smooth morphisms)

$$= H_c^0(\mathbb{A}_{\mathbb{F}_p}^1, f_* \overline{\mathbb{Q}}_\ell[1] \otimes \mathcal{L}_\psi) = H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, f_* \overline{\mathbb{Q}}_\ell \otimes \mathcal{L}_\psi)$$

(by the Leray spectral sequence)

$$= H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi)$$

The eigenvalue of Frobenius acting on $\det \left(H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi) \right)$ is computed in Lemma 7.0.108 as Then if n is odd:

$$\det(\text{Frob}_p, H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi)) = \psi(r) p^{(n-1)/2}$$

If n is even:

$$\det(\text{Frob}_q, H_c^1(\mathbb{A}_{\mathbb{F}_p}^1, f^* \mathcal{L}_\psi)) = -\psi(r) p^{(n-2)/2} G(\chi_2, \psi) \chi_2 \left(\frac{an}{2} \right)$$

If n is odd this determinant is equal to the eigenvalue of $\text{Frob}_{p,\psi}$ acting on $\delta_r \left(\frac{n-1}{2} \right)$ and if n is even the determinate is equal to the eigenvalue of Frob_p acting on $\delta_r \left(\frac{n-2}{2} \right) \otimes \chi_{\frac{an}{2}} * \mathcal{L}_\chi[1]$, because the terms match up perfectly - Tate twists with powers of p , δ_r with $\psi(r)$, $\mathcal{L}_\chi$ with minus the Gauss sum, and $\chi_{\frac{an}{2}}$ with $\chi_2 \left(\frac{an}{2} \right)$.

Since the eigenvalue is the same before and after tensoring with W , it follows that the eigenvalue of Frob_p acting on W is 1. Because this is true for all sufficiently large primes, W is trivial, and our claimed determinant is correct.

□

For a sufficiently generic polynomial, the Tannakian monodromy group is GL_{n-1} :

Lemma 7.0.110. *Let $f(x) \in \mathbb{Z}[x]$ be a polynomial of degree n , and thus a map $f : \mathbb{A}_{\mathbb{Q}}^1 \rightarrow \mathbb{A}_{\mathbb{Q}}^1$. Assume that f satisfies the following two conditions:*

1. $\frac{df}{dx}$ has no multiple roots in $\overline{\mathbb{Q}}$.
2. Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be four roots of $\frac{df}{dx}$ in $\overline{\mathbb{Q}}$, not necessarily distinct. If $f(\alpha_1) + f(\alpha_2) = f(\alpha_3) + f(\alpha_4)$, then $\{\alpha_1, \alpha_2\} = \{\alpha_3, \alpha_4\}$.

Then $f_*\overline{\mathbb{Q}}_{\ell}$ corresponds to a G_S -representation with Tannakian monodromy group GL_{n-1} .

Proof. Observe that if α_1 and α_2 have no distinct roots, then $f(\alpha_1) + f(\alpha_1) \neq f(\alpha_2) + f(\alpha_2)$, so $f(\alpha_1) \neq f(\alpha_2)$. For p sufficiently large, $\frac{df}{dx}$ will still have no multiple roots mod p and $f(\alpha_1)$ will still not equal $f(\alpha_2)$ for distinct roots α_1, α_2 , so f will be a supermorse function in the sense of [Katz, 1990, 7.10.2.2]. Thus by [Katz, 1990, Lemma 7.10.2.3], $f_*\overline{\mathbb{Q}}_{\ell}/\overline{\mathbb{Q}}_{\ell}$ is an irreducible tame reflection sheaf. The singularities of $f_*\overline{\mathbb{Q}}_{\ell}/\overline{\mathbb{Q}}_{\ell}$ are exactly the $f(\alpha_i)$ for α_i roots of $\frac{df}{dx}$. For p sufficiently large, no relations of the form $f(\alpha_1) + f(\alpha_2) = f(\alpha_3) + f(\alpha_4)$ among these singular points hold modulo p that do not hold in characteristic 0. Hence $f_*\overline{\mathbb{Q}}_{\ell}/\overline{\mathbb{Q}}_{\ell}$ satisfies the conditions of [Katz, 1990, Theorem 7.9.6] for p sufficiently large, so its Fourier transform has geometric monodromy group containing SL_{n-1} . Hence by Lemma 4.0.60 the Tannakian monodromy group contains SL_{n-1} . Finally, by Lemma 7.0.109 the the determinant of this representation has Tannakian monodromy group $\mathbb{G}_m$, so the whole representation must have Tannakian monodromy group GL_{n-1} . $\square$

Both of these conditions are satisfied when the coefficients of f lie in some open subset of $\mathbb{A}^{n+1}$, so this implies, assuming Conjecture 1, that $S(f(x), p)$ for a generic polynomial p has a distribution given by the Haar measure on SU_{n-1} twisted by the

determinant character of Lemma 7.0.109. However, special polynomials that fail to satisfy one of these conditions can have wildly different distributions.

One simple case of this is when $n = 4$ and f is the composition of two quadratic polynomials but is not equal to a fourth power plus a constant.

Lemma 7.0.111. *Let $f = g \circ h$ be the composition of two quadratic polynomials but not be of the form $a(x + b)^4 + c$. Then $f_*\overline{\mathbb{Q}}_\ell$ is a sum of 3 geometrically irreducible middle extension sheaves: the constant sheaf $\overline{\mathbb{Q}}_\ell$, $g_*\overline{\mathbb{Q}}_\ell/\overline{\mathbb{Q}}_\ell$, and $f_*\overline{\mathbb{Q}}_\ell/g_*\overline{\mathbb{Q}}_\ell$. When placed in degree -1 , these sheaves have good reduction outside a set of primes S and they correspond to 0, 1, and 2-dimensional representations of G_S respectively.*

Proof. The geometric and arithmetic monodromy groups of the cover $f : \mathbb{A}^1 \rightarrow \mathbb{A}^1$ are subgroups of S_4 , have order a power of 2 because f is a composition of degree 2 covers, and contain a 4-cycle because the local monodromy of ∞ of the cover is a 4-cycle. Hence both monodromy groups are either D_4 or the cyclic group of order 4. But the geometric monodromy group cannot be the cyclic group of order 4 as then f would have the form $a(x + b)^4 + c$, so the geometric and arithmetic monodromy groups are both D_4 .

$f_*\overline{\mathbb{Q}}_\ell$ is a middle extension sheaf [Katz, 1990, Lemma 7.3.2] and has finite monodromy, so it is a sum of irreducible middle extension sheaves. Hence $f_*\overline{\mathbb{Q}}_\ell$ splits as a sum $f_*\overline{\mathbb{Q}}_\ell/g_*\overline{\mathbb{Q}}_\ell + g_*\overline{\mathbb{Q}}_\ell/\overline{\mathbb{Q}}_\ell + \overline{\mathbb{Q}}_\ell$. Because the 4-dimensional permutation representation of D_4 has three irreducible components, these three summands are all geometrically irreducible. By Proposition 3.0.46, these sheaves have good reduction away from some finite set of primes.

By Lemma 7.0.109, $f_*\overline{\mathbb{Q}}_\ell$ is a three-dimensional G_S -representation and $g_*\overline{\mathbb{Q}}_\ell$ is one-dimensional, so $f_*\overline{\mathbb{Q}}_\ell/g_*\overline{\mathbb{Q}}_\ell$ is two-dimensional as a representation of G_S . Furthermore $\overline{\mathbb{Q}}_\ell$ is a 0-dimensional representation of G_S as it is negligible, so $g_*\overline{\mathbb{Q}}_\ell/\overline{\mathbb{Q}}_\ell$ is a one-dimensional representation of G_S . □

Lemma 7.0.112. *Let $f = g \circ h$ be the composition of two quadratic polynomials but not of the form $a(x+b)^4 + c$. Let a be the leading coefficient of g . Let r_1 be the critical value of g and let r_2 be g applied to the critical value of h .*

Then the Tannakian monodromy group of $f_*\overline{\mathbb{Q}}_\ell[1]$ is contained in $GL_1 \times GL_2$, acting by sum of the standard representations of GL_1 and GL_2 , and contains $1 \times SL_2$.

The determinant character of the one-dimensional subrepresentation is $\delta_{r_1}(\chi_a) * \mathcal{L}_\chi[1]$ and the determinant character of the two-dimensional subrepresentation is $\delta_{r_1+r_2}(-1) \otimes \chi_{2a}$.

Proof. The object $f_*\overline{\mathbb{Q}}_\ell[1]$ is a sum of a one-dimensional representation and a two-dimensional representation, so its monodromy group is contained in $GL_1 \times GL_2$. Because $f_*\overline{\mathbb{Q}}_\ell/g_*\overline{\mathbb{Q}}_\ell$ is geometrically irreducible, the two-dimensional representation is Lie-irreducible by Lemma 4.0.83, so its Tannakian monodromy group contains SL_2 . Hence the Tannakian monodromy group contains a subgroup of $GL_1 \times GL_2$ that maps subjectively onto SL_2 under the second projection, whose commutator subgroup must be $1 \times SL_2$.

We can compute the determinants using Lemma 7.0.109. By the chain rule, $\frac{df}{dx} = \frac{dh}{dx}(x) \frac{dg}{dx}(h(x))$. So the three roots of $\frac{df}{dx}$ are the root of dh/dx and the inverse images under h of the root of dg/dx . After applying $f = g \circ h$, the first is sent to r_2 and the last two are sent to r_1 . Let b be the leading coefficient of h , then ab^2 is the leading coefficient of f . Hence by Lemma 7.0.109:

$$\det^\tau(f_*\overline{\mathbb{Q}}_\ell[1]) = \delta_{2r_1+r_2}(-1) \otimes \chi_{2ab^2} * \mathcal{L}_\chi[1] = \delta_{2r_1+r_2}(-1) \otimes \chi_{2a} * \mathcal{L}_\chi[1]$$

$$\det^\tau(g_*\overline{\mathbb{Q}}_\ell[1]) = \delta_{r_1} \otimes \chi_a * \mathcal{L}_\chi[1]$$

Therefore:

$$\det^\tau(f_*\overline{\mathbb{Q}}_\ell[1]/g_*\overline{\mathbb{Q}}_\ell[1]) = \delta_{r_1+r_2}(-1) \otimes \chi_{2a}$$

Because $g_*\overline{\mathbb{Q}}_\ell[1]$ and $f_*\overline{\mathbb{Q}}_\ell[1]/g_*\overline{\mathbb{Q}}_\ell[1]$ are the summands of $f_*\overline{\mathbb{Q}}_\ell$ corresponding to the one- and two-dimensional representations respectively, these last two equations give the formula for the determinants of those two representations. $\square$

Lemma 7.0.112 allows one to compute the conjectural distribution of $\frac{a_p(f_*\overline{\mathbb{Q}}_\ell)}{\sqrt{p}}$ by following the definition of the adjusted Haar measure. First compute the map d described in the definition of the Dirichlet measure, choosing $\det(g_*\overline{\mathbb{Q}}_\ell[1])$ and $\det(f_*\overline{\mathbb{Q}}_\ell[1]/g_*\overline{\mathbb{Q}}_\ell[1])$ as the generating set of characters. The map d is a map from congruence classes mod m for a suitable m to $U(1) \times U(1)$. Write $d(r) = (d_1(r), d_2(r)) \in U(1) \times U(1)$. For each r , take the distribution of the sum of the traces of a random pair of matrices in $U(1) \times U(2)$ conditional on the determinant of the first matrix being $d_1(r)$ and the determinant of the second matrix being $d_2(r)$. Since the first matrix is precisely $d_1(r)$ and the second matrix is $\sqrt{d_2(r)}$ times a matrix in $SU(2)$, this distribution of the trace is the Sato-Tate measure rotated in the complex plane by $\sqrt{d_2(r)}$ and translated by $d_1(r)$. Then sum over all invertible r mod m and divide by $\varphi(m)$.

For instance, take $g(x) = x^2 + x$, $h(x) = x^2$, $f(x) = x^4 + x^2$. Then $a = 1$, $r_1 = -1/4$, $r_2 = 0$. Then the determinant objects of $g_*\overline{\mathbb{Q}}_\ell$ and $f_*\overline{\mathbb{Q}}_\ell/g_*\overline{\mathbb{Q}}_\ell$ are $\delta_{-1/4} * \mathcal{L}_\chi[1]$ and $\delta_{-1/4}(-1) \otimes Q(2)$. Hence we may take $m = 8$. By the definition of d :

- $r \equiv 1 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as -1 , on $\delta_{-1/4}$ as i , and on χ_2 as 1 , so $d_1 = -i$, $d_2 = i$.
- $r \equiv 3 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as $-i$, on $\delta_{-1/4}$ as $-i$, and on χ_2 as -1 , so $d_1 = -1$, $d_2 = i$.

- $r \equiv 5 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as -1 , on $\delta_{-1/4}$ as i , and on χ_2 as -1 , so $d_1 = -i$, $d_2 = -i$.
- $r \equiv 7 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as $-i$, on $\delta_{-1/4}$ as $-i$, and on χ_2 as 1 , so $d_1 = -1$, $d_2 = -i$.

So we expect the distribution of the trace of Frobenius to be equal to the distribution of $x + yz$ where $x \in \{-1, -i\}$, $y \in \{\sqrt{i}, \sqrt{-i}\}$, and $z \in [-2, 2]$ are independently distributed: x and y uniformly distributed and z according to the Sato-Tate measure. The exponential sum is minus that.

Or take $g(x) = 2x^2 + 4x + 1$, $h(x) = x^2 + x$, $f(x) = x^4 + (x + 1)^4$. Here $a = 2$, $r_1 = -1$, $r_2 = 1/8$. Now the determinant objects are $\delta_{-1} \otimes \chi_2 * \mathcal{L}_\chi[1]$ and $\delta_{-7/8}(-1) \otimes Q(2)$, so we may again take $m = 8$. Recall that $d(r)$ always acts trivially on δ_{-1} , so:

- $r \equiv 1 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as -1 , on $\delta_{-7/8}$ as $\frac{1-i}{\sqrt{2}}$, and on χ_2 as 1 , so $d_1 = -1$, $d_2 = \frac{1-i}{\sqrt{2}}$.
- $r \equiv 3 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as $-i$, on $\delta_{-7/8}$ as $\frac{-1-i}{\sqrt{2}}$, and on χ_2 as -1 , so $d_1 = i$, $d_2 = \frac{1+i}{\sqrt{2}}$.
- $r \equiv 5 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as -1 , on $\delta_{-7/8}$ as $\frac{-1+i}{\sqrt{2}}$, and on χ_2 as -1 , so $d_1 = 1$, $d_2 = \frac{1-i}{\sqrt{2}}$.
- $r \equiv 7 \pmod{8}$: $d(r)$ acts on $\mathcal{L}_\chi$ as $-i$, on $\delta_{-7/8}$ as $\frac{1+i}{\sqrt{2}}$, and on χ_2 as 1 , so $d_1 = -i$, $d_2 = \frac{1+i}{\sqrt{2}}$.

Thus the distribution is a superposition of four translated and rotated copies of the Sato-Tate distribution - we do not see a simpler description.

These two examples, as well as another case with D_4 monodromy that we do not repeat here, occur in [Patterson, 2003, pp. 141], which is concerned with the

distribution of exponential sums over $\mathbb{Z}/n\mathbb{Z}$ rather than just $\mathbb{Z}/p\mathbb{Z}$. Patterson finds empirically that the sums $\sum_{n < N} \sum_{x \in \mathbb{Z}/n} e^{2\pi i f(x)/n}$ associated to these three polynomials with D_4 monodromy show no more cancellation than that implied by the Weil bound $|\sum_{x \in \mathbb{Z}/n} e^{2\pi i f(x)/n}| = O(n^{1/2+\epsilon})$. This makes sense, as sums of this type do not always cancel when averaged over the primes either. It is a consequence of Conjecture 2 that the average of the normalized exponential sum over the primes is equal to the average value of d_1 , which can be nonzero - for the first sum it is $\frac{-1-i}{2}$ - though it can also be zero, as in our second case. On the other hand, for general polynomials of degree 4, Patterson expects the average over composites to be 0, and Conjecture 2 predicts that the average over the primes is 0. So there seems to be some relationship between the average over primes and the average over composites, though not a direct one - e.g. the averages need not always be equal.

As a final example of an interesting sheaf, consider a sheaf defined by an elliptic surface. Let $\pi : \mathcal{E} \rightarrow \mathbb{A}^1$ be an elliptic surface. Then we will consider the sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell(1)$. When we put the surface in Weierstrass form $y^2 = 4x^3 - g_2(t)x - g_3(t)$, this corresponds to the Legendre symbol sum:

$$a_p = \frac{1}{p} \sum_{t=0}^{p-1} \left(\sum_{x=0}^{p-1} \left(\frac{4x^3 - g_2(t)x - g_3(t)}{p} \right) \right) e^{2\pi i \frac{t}{p}}$$

We compute the Tannakian monodromy group in one instance. Let $\pi : \mathcal{E} \rightarrow \mathbb{A}^1$ be the “squared Dwork family” elliptic surface in $\mathbb{P}_{\mathbb{A}_{\mathbb{Z}[1/3\ell]}^2}^2$ defined by the equation $3xyz = t^2(x^3 + y^3 + z^3)$ in projective coordinates $(x : y : z)$ with parameter t . Let $\mathcal{F} = R^1\pi_*\overline{\mathbb{Q}}_\ell(1)$ be a sheaf on $\mathbb{A}_{\mathbb{Z}[1/3\ell]}^1$.

Lemma 7.0.113. *For a prime $p \notin \{3, \ell\}$, $\mathcal{F}_{\mathbb{F}_p}$ is geometrically isomorphic to the middle extension from $\mathbb{G}_m$ to $\mathbb{A}^1$ of the pullback along the map $[6] : t \mapsto t^6$ of the hypergeometric sheaf $\mathcal{H}(\psi; 1, 1; \chi_3, \overline{\chi}_3)$ (in the notation of [Katz, 2007, pp. 3], where χ_3 and $\overline{\chi}_3$ are the two nontrivial cubic characters).*

Furthermore, $\mathcal{F}_{\mathbb{Q}}[1]$ is perverse of geometric origin and for each $p \notin \{3, \ell\}$, $\mathcal{F}_{\mathbb{F}_p}$ is pure of weight 1.

Proof. $\mathcal{E}$ is the pullback along the map $t \mapsto t^{-2}$ of the elliptic surface with equation

$$(x^3 + y^3 + z^3) = 3\lambda xyz$$

(using the new parameter $\lambda = t^2$). This surface is the variety X_λ described in [Katz, 2009, pp. 2] in the special case $n = 3, d = 3, W = (1, 1, 1)$. Our sheaf $R^1\pi_*\overline{\mathbb{Q}}_\ell$ is the same as what Katz refers to as $Prim^{n-2}$. He decomposes $Prim^{n-2}$ into summands $Prim_{V/W}^{n-2}$ indexed by classes in $(\mathbb{Z}/d\mathbb{Z})_0^n/W$. Here $(\mathbb{Z}/d\mathbb{Z})_0^n$ is the group of triples of elements of $\mathbb{Z}/3$ that sum to 0, and $(\mathbb{Z}/d\mathbb{Z})_0^n/W$ is a cyclic group of three elements. The only elements of $(\mathbb{Z}/d\mathbb{Z})_0^n$ whose entries are all nonzero are $(1, 1, 1)$ and $(2, 2, 2)$, which both lie in the identity class in $(\mathbb{Z}/d\mathbb{Z})_0^n/W$, so by [Katz, 2009, Lemma 3.1] only $Prim_{0/W}^{n-2}$ is nonzero.

To compute $Prim_{0/W}^{n-2}$, note the following (in the notation of [Katz, 2009, pp. 12]): Because $V = 0$ is trivial and all the w are 1, $List(V, W) = \{1, 1, 1\}$. Furthermore $List(all\ d) = \{1, \chi_3, \overline{\chi}_3\}$. Thus $\mathcal{H}_{V,W} = \mathcal{H}(\chi_3, \overline{\chi}_3; 1, 1)$. Then by [Katz, 2009, Theorem 5.3], $Prim_{0/W}^{n-2}$ is geometrically isomorphic over the open subset where it is lisse to $[3]^*\mathcal{H}(\chi_3, \overline{\chi}_3; 1, 1)$. Pulling back along the map $t \mapsto t^{-2}$ is equivalent to inverting the coordinate of the hypergeometric sheaf, which is the same as switching the two sets of characters, and then pulling back along $[2]$, resulting in $[6]^*\mathcal{H}(1, 1; \chi_3, \overline{\chi}_3)$.

It remains to check that $\mathcal{F}$ is actually a middle extension sheaf, so that it is isomorphic to the middle extension of $[6]^*\mathcal{H}(1, 1; \chi_3, \overline{\chi}_3)$ everywhere, not just on the open set where it is lisse. At each singular point of $\mathcal{E}$, the special fiber of $\mathcal{E}$ is the union of three lines intersecting transversely, and hence by proper base change the stalk of $\mathcal{F}$ has rank 1. By the description of $\mathcal{F}$ as a hypergeometric sheaf, at each singular points its local monodromy is unipotent of rank 2, so the local monodromy invariants

of $\mathcal{F}$ have rank 1. Thus it is sufficient to show that the map from stalks to local monodromy invariants is injective. By the vanishing cycles long exact sequence, the kernel of the map from the stalk of $R^1\pi_*\overline{\mathbb{Q}}_\ell$ at the special point to the stalk of $R^1\pi_*\overline{\mathbb{Q}}_\ell$ at the general point is a quotient of $H^0(\mathcal{E}_s, R\Phi_s\overline{\mathbb{Q}}_\ell)$. But $H^0(\mathcal{E}_s, R\Phi_s\overline{\mathbb{Q}}_\ell)$ vanishes as $R\Phi_s\overline{\mathbb{Q}}_\ell$ vanishes at smooth points and is supported in degree 1 at singular points by the Picard-Lefschetz formula [Deligne and Katz, 1973, XV, Proposition 3.1.2(i)]

Now the other two claims:

$\mathcal{F}_{\mathbb{Q}}$ is a middle extension sheaf and is irreducible on the open set where it is lisse because its j invariant is nonconstant. Hence $\mathcal{F}_{\mathbb{Q}}[1]$ is an irreducible perverse sheaf. Thus it appears in the Jordan-Holder decomposition of ${}^p\mathcal{H}^2(R\pi_*\overline{\mathbb{Q}}_\ell)$ and so is of geometric origin.

The stalks of $\mathcal{F}_{\mathbb{F}_p}$ on the open set where it is lisse are Tate modules, hence pure of weight -1 , so $\mathcal{F}_{\mathbb{F}_p}[1]$ is pure of weight 0 on the open set where it is lisse. Because $\mathcal{F}_{\mathbb{F}_p}$ is a middle extension sheaf, $\mathcal{F}_{\mathbb{F}_p}[1]$ is pure of weight 0 everywhere. $\square$

Lemma 7.0.114. *Let $p \neq 3, \ell$ be a prime. The Fourier transform of $\mathcal{F}_{\mathbb{F}_p}$ has geometric monodromy group G_2 and the Tannakian monodromy group of $\mathcal{F}_{\mathbb{Q}}[1]$ contains G_2 .*

Proof. By Lemma 7.0.113, $\mathcal{F}_{\overline{\mathbb{F}_p}} = [6]^*\mathcal{H}(\psi; 1, 1; \chi_3, \overline{\chi}_3)$. Let $(1, \chi_2, \chi_3, \overline{\chi}_3, \chi_6, \overline{\chi}_6)$ be the characters of μ_6 . Then by [Katz, 1990, Theorem 9.3.2] the Fourier transform of $[6]^*\mathcal{H}(\psi; 1, 1; \chi_3, \overline{\chi}_3)$ is

$$[6]^* \text{Cancel}(\mathcal{H}(\psi; \chi_3, \overline{\chi}_3, 1, \chi_2, \chi_3, \overline{\chi}_3, \chi_6, \overline{\chi}_6; 1, 1))$$

(canceling duplicate characters)

$$= [6]^*\mathcal{H}(\psi; \chi_3, \overline{\chi}_3, \chi_2, \chi_3, \overline{\chi}_3, \chi_6, \overline{\chi}_6; 1)$$

(tensoring with $\mathcal{L}_{\chi_2}$ does not change the pullback, and is equivalent to multiplying each involved character by χ_2 [Katz, 1990, (8.3.3)])

$$= [6]^* \mathcal{H}(\psi; \chi_6, \bar{\chi}_6, 1, \chi_6, \bar{\chi}_6, \chi_3, \bar{\chi}_3; \chi_2)$$

Let $\alpha = \beta = \chi_6$. Then

$$\mathcal{H}(\psi; \chi_6, \bar{\chi}_6, 1, \chi_6, \bar{\chi}_6, \chi_3, \bar{\chi}_3; \chi_2) = \mathcal{H}(\psi; 1, \alpha, \beta, \alpha\beta, \bar{\alpha}, \bar{\beta}, \overline{\alpha\beta}; \chi_2)$$

Hence by [Katz, 2007, Theorem 3.1], this hypergeometric sheaf has geometric monodromy group G_2 , so its pullback by $[6]$ also has geometric monodromy group G_2 . Thus by Lemma 4.0.60 the Tannakian monodromy group contains G_2 . $\square$

The computations in Lemma 7.0.114 were shown to me by Nick Katz.

Lemma 7.0.115. *The Tannakian monodromy group of $\mathcal{F}_{\mathbb{Q}}[1]$ is contained in $\mathbb{G}_m \times G_2$.*

Proof. Consider the complex $\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1]$ on $\mathbb{A}_{\mathbb{Z}[1/3\ell]}^1$.

The Fourier transform of its restriction to $\mathbb{A}_{\mathbb{F}_p}^1$, $FT_{\psi}((\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1])_{\mathbb{F}_p})$ is equal to $(FT_{\psi}(\mathcal{F}_{\mathbb{F}_p}[1]))^{\otimes 3}$ [Brylinski, 1986, Corollary 9.6]. As the Fourier transform of $\mathcal{F}$ has geometric monodromy group G_2 on $\mathbb{G}_m$, its third tensor power has a constant summand when restricted to the open set $\mathbb{G}_m$.

The stalk of at the point 0 of $FT_{\psi}((\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1])_{\mathbb{F}_p})$ is equal to $H_c^*(\mathbb{A}_{\mathbb{F}_p}^1, \mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1]) = H_c^*(\mathbb{A}_{\mathbb{F}_p}^1, \mathcal{F}[1])^{\otimes 3}$ which is concentrated in degree 0 because $H_c^i(\mathbb{A}_{\mathbb{F}_p}^1, \mathcal{F})$ vanishes for $i \neq 1$. Hence $FT_{\psi}((\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1])_{\mathbb{F}_p})$ is concentrated in degree 0.

We have just seen that $FT_{\psi}((\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1])_{\mathbb{F}_p})$ has a nontrivial map to the constant sheaf $\overline{\mathbb{Q}}_{\ell}$ on the open set $\mathbb{G}_m$, hence it has a nontrivial map to the constant sheaf $\overline{\mathbb{Q}}_{\ell}$ everywhere by the adjunction $\mathrm{Hom}_{\mathbb{A}^1}(FT_{\psi}((\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1])_{\mathbb{F}_p}), j_* \overline{\mathbb{Q}}_{\ell}) = \mathrm{Hom}_{\mathbb{G}_m}(j^* FT_{\psi}((\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1])_{\mathbb{F}_p}), \overline{\mathbb{Q}}_{\ell})$ because $j_* \overline{\mathbb{Q}}_{\ell} = \overline{\mathbb{Q}}_{\ell}$. Thus $(\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1])_{\mathbb{F}_p}$ has a nontrivial map to the skyscraper sheaf at 0.

Because $\mathcal{F}_{\mathbb{Q}}[1]$ is irreducible and nonconstant, by Lemma 3.0.20 it has property $\mathcal{P}$, hence $\mathcal{F}_{\mathbb{Q}}[1] *! \mathcal{F}_{\mathbb{Q}}[1] *! \mathcal{F}_{\mathbb{Q}}[1]$ a perverse sheaf on $\mathbb{A}_{\mathbb{Q}}^1$. So the generic fiber of $\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1]$ is perverse. Therefore the fiber over all but finitely many primes is perverse. At these primes p , $\mathcal{F}_{\overline{\mathbb{F}}_p}[1] *! \mathcal{F}_{\overline{\mathbb{F}}_p}[1] *! \mathcal{F}_{\overline{\mathbb{F}}_p}[1]$ is a perverse sheaf with a nontrivial map to the skyscraper sheaf at 0, hence has a nonzero stalk in degree 0 at the point 0. Since this occurs for infinitely many primes, $\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1]$ has a nonzero stalk in degree 0 at the point 0 over $\mathbb{Q}$, hence maps nontrivially to the skyscraper sheaf $\delta_0 \otimes V$ at 0, where V is the stalk of $\mathcal{F}[1] *! \mathcal{F}[1] *! \mathcal{F}[1]$ at 0, viewed as a Galois representation.

Thus the geometric Tannakian monodromy group of $\mathcal{F}$ fixes some trilinear form, so since the geometric monodromy group contains G_2 it must fix up to scaling the unique form fixed by G_2 , which means the geometric monodromy group is G_2 , and the arithmetic monodromy group is contained in the normalizer of G_2 , which is $\mathbb{G}_m \times G_2$. $\square$

The argument in Lemma 7.0.115 is adapted from [Katz, 1990, Theorem 14.6].

Lemma 7.0.116. *The Tannakian dual of $\mathcal{F}_{\mathbb{Q}}[1]$ is equal to $\mathcal{F}_{\mathbb{Q}}[1]$.*

Proof. By Proposition 3.0.22, $(\mathcal{F}_{\mathbb{Q}}[1])^{\vee} = [1]^* D\mathcal{F}_{\mathbb{Q}}[1]$. By the Weil pairing, $D\mathcal{F}_{\mathbb{Q}}[1] = \mathcal{F}_{\mathbb{Q}}[1]$. By the symmetry of $\mathcal{E}$, $[-1]^* \mathcal{F}_{\mathbb{Q}}[1] = \mathcal{F}_{\mathbb{Q}}[1]$. $\square$

Lemma 7.0.117. *Let $j : \mathbb{A}^1 \rightarrow \mathbb{P}^1$ be the open immersion. The determinant of the Galois representation $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_* \mathcal{F})$ is trivial.*

Proof. It is sufficient to show that, for p sufficiently large, the determinant of the action of Frob_p on $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_* \mathcal{F})$ is 1. The characteristic polynomial of Frob_p is the L -function of the elliptic curve $\mathcal{E}$ over $\mathbb{F}_p(t)$. The dimension of $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_* \mathcal{F})$ is minus the Euler characteristic of $j_* \mathcal{F}$. Because $j_* \mathcal{F}$ is a sheaf of rank 2 on $\mathbb{P}^1$ with a drop of 1 at 7 singular points, its Euler characteristic is $2 \times 2 - 7 \times 1 = -3$, so $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_* \mathcal{F})$

has dimension 3 and the L-function of $\mathcal{E}$ has degree 3. Therefore the determinant of the action of Frob_p on $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_*\mathcal{F})$ is minus the epsilon factor of this elliptic curve. So it is sufficient to show that for large p , the epsilon factor is -1 .

The epsilon factor is a product of the local epsilon factors over the singular fibers, where fibers over closed points, not geometric points, are considered. For a semistable fiber, the epsilon factor is -1 for split curves and $+1$ for non-split curves [Kramer and Tunnel, 1982, 2.1.7], as the Weil-Deligne representation has the form $sp(2) \otimes \chi$ for χ trivial if the fiber is split and χ unramified quadratic if the fiber is non-split. All the singular fibers are semistable. There is an involution $t \rightarrow -t$ of the elliptic surface. This involution sends split fibers to split and non-split to non-split, so all the epsilon factors of fibers not fixed by this involution cancel. The only closed singular point fixed by the involution is 0, as no 6th root of unity is Galois conjugate to its own negation. The fiber at 0, which is a curve with equation $xyz = 0$, is split, so the local epsilon factor is -1 , and the global factor is -1 as desired. $\square$

Lemma 7.0.118. *The Tannakian determinant of $\mathcal{F}_{\mathbb{Q}}[1]$ is δ_0 .*

Proof. First, because $\mathcal{E}$ has good reduction at ∞ , the local monodromy of $\mathcal{F}_{\mathbb{Q}}$ at ∞ is trivial, so by Lemma 4.0.66 the Tannakian determinant of $\mathcal{F}_{\mathbb{Q}}[1]$ has the form $\delta_r \otimes V$. Second, $\mathcal{E}$ has semistable reduction at the 6th roots of unity and 0, so $\mathcal{F}$ has drop 1 at those 7 points. Their sum is 0, so by Lemma 4.0.67, r is 0.

Finally, we show that V is trivial using Lemma 4.0.68. Let $A = \mathcal{F}[1]$. Let W be the Tate module of the elliptic curve with equation $x^3 + y^3 + z^3 = 0$, and let B be the constant sheaf $W[1]$. Then A and B have the same, trivial, local monodromy at ∞ and the same Galois action on that local monodromy representation. Hence we can apply Lemma 4.0.68. Observe that $H_c^2(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{F}) = 0$, $H_c^1(\mathbb{A}_{\mathbb{Q}}^1, W) = 0$, and $\det^{\tau}(B) = 1$ so $V_B = 1$. This Lemma 4.0.68 simplifies to the formula

$$V_A = \det \left(H_c^1(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{F}) \right) \otimes \left(H_c^2(\mathbb{A}_{\mathbb{Q}}^1, W) \right).$$

Observe that $H_c^2(\mathbb{A}_{\mathbb{Q}}^1, W) = W(-1)$ and its determinant, by the Weil pairing, is $\mathbb{Q}_\ell(-1)$. Observe also that for the inclusion $j : \mathbb{A}^1 \rightarrow \mathbb{P}^1$, $H_c^1(\mathbb{A}_{\mathbb{Q}}^1, \mathcal{F})$ is the extension of $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_*\mathcal{F})$ by W and so its determinant is the determinant of $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_*\mathcal{F})$ times the determinant of W . By Lemma 7.0.117 the determinant of $H_c^1(\mathbb{P}_{\mathbb{Q}}^1, j_*\mathcal{F})$ is trivial and the determinant of W is $\mathbb{Q}_\ell(1)$. Hence $V = V_A$ is trivial. $\square$

Lemma 7.0.119. *The Tannakian monodromy group of $\mathcal{F}_{\mathbb{Q}}[1]$ is equal to G_2 .*

Proof. The Tannkian monodromy group is a subgroup of $\mathbb{G}_m \times G_2$ (Lemma 7.0.115), containing G_2 (Lemma 7.0.114), whose standard representation is self-dual (Lemma 7.0.116) and whose determinant is trivial (Lemma 7.0.118). Because the standard representation is self-dual, it must fix the unique G_2 -invariant bilinear form, and hence lie in $\pm 1 \times G_2$. As the determinant of elements in $-1 \times G_2$ is -1 , the Tannakian monodromy group must be G_2 . $\square$

Chapter 8

Automorphic Forms and L-functions

In this chapter we discuss a few broad questions related to our main conjecture. What strategies are available for proving new cases of our conjecture? Can suitable L-functions be used to attack it? Might some cases of Conjecture 1 actually follow trivially from cases of the generalized Sato-Tate conjecture? Might cases of Conjecture 1 follow from equidistribution conjectures on automorphic forms?

It is known that the generalized Sato-Tate conjecture would follow from the existence of analytic continuations of the L-functions of all Galois representations of geometric origin, which itself would follow from the Langlands program. The Sato-Tate conjecture for non-CM elliptic curves over totally real fields was proved in [Harris et al., 2011, Corollary 8.9] by proving potential automorphy, and thus meromorphic continuation, for the necessary Galois representations, and further advances in the Langlands program might resolve more cases of the generalized Sato-Tate conjecture.

Two nontrivial, unmotivated cases of Conjecture 1 have been resolved - the distribution of Kummer sums by Heath-Brown and Patterson [1979, Theorem 2], and the distribution of roots of a quadratic congruence by Duke et al. [1995, Theorem on pp.

2]. Both proofs rely critically on viewing the exponential sum over $\mathbb{F}_p$ as a special case of a more general sum over $\mathbb{Z}/n$ - this is noted $g(c)$ in [Heath-Brown and Patterson, 1979] and $\rho_h(n)$ in [Duke et al., 1995]. One then evaluates the sum of this function where n varies in an arithmetic progression, and more intricate sums involving bilinear or trilinear forms [Heath-Brown and Patterson, 1979, pp. 117] [Duke et al., 1995, pp. 425]. These are combined with Vaughan's identity in [Heath-Brown and Patterson, 1979] and a combinatorial identity in [Duke et al., 1995] to sieve out the primes and get the desired equidistribution statement. In both cases, the sums $g(c)$ and $\rho_h(n)$ are the Fourier coefficients of metaplectic automorphic forms [Heath-Brown and Patterson, 1979, pp. 125].

One might try to apply a similar strategy to more general cases of Conjecture 1. There are natural ways to extend many of the functions $p \mapsto a_p(\mathcal{F})$ from primes to general natural numbers, but as far as we know, there is no reason to expect general exponential sums $a_p(\mathcal{F})$ to be connected to automorphic forms.

Next we will consider the following question:

For which sheaves $\mathcal{F}$ does there exist a Galois representation V such that $a_p(\mathcal{F}) = \text{tr}(\text{Frob}_p, V)$ for all but finitely many primes p ?

For such sheaves our equidistribution question was really a motivated equidistribution question all along, and the distinctions we have drawn between that question and the generalized Sato-Tate conjecture would disappear.

When $\mathcal{F}$ is a skyscraper sheaf $\delta_0 \otimes V$, or is equivalent to a skyscraper sheaf modulo constant sheaves, we have $a_p(\mathcal{F}) = \text{tr}(\text{Frob}_p, V)$ for all p . We suspect that these are the only examples of $\mathcal{F}$ with $a_p(\mathcal{F}) = \text{tr}(\text{Frob}_p, V)$.

We will provide some evidence for this. Our strategy will be to show that for a Galois representation V , the function $p \mapsto \text{tr}(\text{Frob}_p, V)$ has a certain property that $p \mapsto a_p(\mathcal{F})$ does not have for more general $\mathcal{F}$. We will show this property, namely that there exists a single number field containing $\text{tr}(\text{Frob}_p, V)$ for all p , only assuming

V is a Galois representation of geometric origin and conditional on standard number-theoretic conjectures.

Lemma 8.0.120. *Let V be a ℓ -adic Galois representation of geometric origin over $\mathbb{Q}$. Assume the Tate conjecture and the semisimplicity conjecture for ℓ -adic cohomology. Then there is a number field K such that for all but finitely many primes p , $\text{tr}(\text{Frob}_p, V)$ is contained in K .*

Proof. By the definition of geometric origin, V is a subquotient of $H^*(X_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell)$ for some smooth projective variety X over $\mathbb{Q}$. Assuming the semisimplicity conjecture, V is in fact a summand of $H^*(X, \overline{\mathbb{Q}}_\ell)$. So there is some Galois-equivariant endomorphism e of $H^*(X_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell)$ that projects idempotently onto V . Assuming the Tate conjecture, that projector is a $\overline{\mathbb{Q}}_\ell$ -linear combination of the classes in $\text{Hom}(H^*(X_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell), H^*(X_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell))$ of algebraic cycles on $X \times X$ of dimension $\dim X$.

Let R be the subalgebra of $\text{Hom}(H^*(X_{\mathbb{C}}, \mathbb{Q}), H^*(X_{\mathbb{C}}, \mathbb{Q}))$ generated by algebraic cycle classes. Then because ℓ -adic etale cohomology is isomorphic to singular cohomology tensored with $\overline{\mathbb{Q}}_\ell$, $R \otimes \overline{\mathbb{Q}}_\ell$ is the subalgebra of $\text{Hom}(H^*(X_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell), H^*(X_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell))$ generated by algebraic cycle classes. By the Tate conjecture, $R \otimes \overline{\mathbb{Q}}_\ell$ is the algebra of Galois-equivariant endomorphisms of $H^*(X_{\overline{\mathbb{Q}}}, \overline{\mathbb{Q}}_\ell)$ which by the semisimplicity conjecture is a semisimple algebra. The algebra R is also finite-dimensional over $\mathbb{Q}$, and hence a finite-dimensional semisimple algebra. Thus $R \otimes \overline{\mathbb{Q}}_\ell$ is isomorphic to a product of matrix algebras $\prod_i M_{n_i}$ and this isomorphism sends the idempotent projector e to a tuple of idempotent matrices (e_i) . Any element e' of $\prod_i M_{n_i}$ whose i th component e'_i is idempotent of the same rank as e_i for all i is a projector onto an isomorphic representation: Since all idempotent matrices of the same rank are conjugate, for each i there is a matrix that conjugates e'_i to e_i , hence there is a tuple of matrices that conjugates e' to e , and this automorphism of $H^*(X, \overline{\mathbb{Q}}_\ell)$ restricts to an isomorphism between the image of e' and the image of e .

Let K be a number field such that $R \otimes K$ is a product of matrix algebras. Then there is an element e' of $R \otimes K$ whose i th component has the same rank as e_i for all i . Thus e' is a K -linear combination of algebraic cycle classes such that e' forms an idempotent projector onto a representation isomorphic to V .

Thus $\mathrm{tr}(\mathrm{Frob}_p, V) = \mathrm{tr}(e' \mathrm{Frob}_p, H^*(X, \overline{\mathbb{Q}}_\ell))$. For all primes p where X has good reduction, this is equal to the trace of the same operator on $H^*(X_{\overline{\mathbb{F}}_p}, \overline{\mathbb{Q}}_\ell)$. Over $X_{\overline{\mathbb{F}}_p}$, Frob_p is an algebraic cycle class so $e' \mathrm{Frob}_p$ is a K -linear combination of algebraic cycle classes and thus $\mathrm{tr}(e' \mathrm{Frob}_p, H^*(X_{\overline{\mathbb{F}}_p}, \overline{\mathbb{Q}}_\ell))$ is a K -linear combination of traces of algebraic cycles on cohomology. The trace of an algebraic cycle on cohomology is an intersection number, hence an integer. Therefore the trace of Frobenius lies in K . $\square$

We can show under mild assumptions on $\mathcal{F}$ that there is no fixed number field containing $a_p(\mathcal{F})$ for all primes p . For such an $\mathcal{F}$, this suggests that $a_p(\mathcal{F})$ is not the trace of a Galois representation of geometric origin:

Proposition 8.0.121. *Let $\mathcal{F}$ be a constructible ℓ -adic sheaf on $\mathbb{A}_{\mathbb{Z}[1/N]}^1$ such that $\mathcal{F}_{\mathbb{Q}}$ is a middle extension sheaf. Suppose that there is a number field K such that for all primes p not dividing N and all points $x \in \mathbb{F}_p$, $\mathrm{tr}(\mathrm{Frob}_p, \mathcal{F}, x) \in K$. Suppose also that $\mathcal{F}_{\mathbb{Q}}$ on $\mathbb{A}_{\mathbb{Q}}^1$ is not the middle extension of a lisse sheaf on $\mathbb{G}_{m, \mathbb{Q}}$ with unipotent local monodromy at 0.*

Then there is no number field L such that the sum $a_p(\mathcal{F})$ is contained in L for all but finitely many primes p .

Proof. Assume that the $a_p(\mathcal{F})$ are contained in some number field L for all but finitely many p . By increasing N if necessary, we may assume that $a_p(\mathcal{F})$ is contained in L for all p not dividing N . Take p , not dividing N , large enough that $\mathbb{Q}(\mu_p)$ is disjoint from KL , e.g. if p is unramified in KL . The Galois group $\mathrm{Gal}(\mathbb{Q}(\mu_p)|\mathbb{Q}) =$

$\text{Gal}(KL(\mu_p)|KL)$ is isomorphic to $(\mathbb{Z}/p)^\times$. An element $\sigma \in (\mathbb{Z}/p)^\times$ acts on

$$a_p = \sum_{x=0}^{p-1} e^{2\pi i \frac{x}{p}} \text{tr}(\text{Frob}_p, \mathcal{F}, x)$$

by fixing $\text{tr}(\text{Frob}_p, \mathcal{F}, x)$ and replacing $e^{2\pi i \frac{x}{p}}$ by $e^{2\pi i \frac{\sigma x}{p}}$. Because $a_p \in L$, a_p is fixed by this action, so for all $\sigma \in \mathbb{F}_p^\times$,

$$a_p = \sum_{x=0}^{p-1} e^{2\pi i \frac{\sigma x}{p}} \text{tr}(\text{Frob}_p, \mathcal{F}, x)$$

In other words the finite Fourier transform of the function $x \mapsto \text{tr}(\text{Frob}_p, \mathcal{F}, x)$ is constant away from 0. It follows by applying the inverse Fourier transform that $x \mapsto \text{tr}(\text{Frob}_p, \mathcal{F}, x)$ is constant away from 0.

Let x and y be two nonzero algebraic numbers. For all but finitely many of the primes p that split in $\mathbb{Q}(x, y)$, $\text{tr}(\text{Frob}_p, \mathcal{F}_{\mathbb{F}_p}, x) = \text{tr}(\text{Frob}_p, \mathcal{F}_{\mathbb{F}_p}, y)$. Let $\mathcal{F}_x$ and $\mathcal{F}_y$ be the stalks of $\mathcal{F}_{\mathbb{Q}}$ at x and y . Because $\mathcal{F}$ is a constructible sheaf, for all but finitely many p that split in $\mathbb{Q}(x, y)$, $\mathcal{F}_x$ and $\mathcal{F}_y$ are unramified at p and $\text{tr}(\text{Frob}_p, \mathcal{F}_x) = \text{tr}(\text{Frob}_p, \mathcal{F}_y)$. Apply this to a sequence of p such that their conjugacy classes of Frobenius acting on the stalks $\mathcal{F}_x$ and $\mathcal{F}_y$ converge ℓ -adically to the identity, so that their traces converge ℓ -adically to the rank. It follows that $\text{rank}(\mathcal{F}_x) = \text{rank}(\mathcal{F}_y)$. Hence $\mathcal{F}$ has constant rank away from 0.

Because $\mathcal{F}_{\mathbb{Q}}$ is a middle extension sheaf, it is lisse away from 0.

Therefore, after base change to some cyclotomic field, $\mathcal{F}$ is the middle extension from $\mathbb{G}_m$ to $\mathbb{A}^1$ of an iterated extension of Kummer sheaves $\mathcal{L}_{\chi_i}$ tensored with Galois representations V_i . So for p sufficiently large and split in that cyclotomic field:

$$\text{tr}(\text{Frob}_p, \mathcal{F}, x) = \sum_{i=1}^n \text{tr}(\text{Frob}_p, V_i) \text{tr}(\text{Frob}_p, \mathcal{L}_{\chi_i}, x) = \sum_{i=1}^n \text{tr}(\text{Frob}_p, V_i) \chi_i(x)$$

In this formula we may assume that no χ_i appears twice by combining those terms and summing the representations V_i . For p sufficiently large, $\text{tr}(\text{Frob}_p, \mathcal{F}, x)$ is constant away from 0. By linear independences of characters, this implies that $\text{tr}(\text{Frob}_p, V_i) = 0$ whenever χ_i is nontrivial. Because this holds for all sufficiently large p split in the cyclotomic field, V_i must be zero whenever χ_i is nontrivial. So the only $\mathcal{L}_\chi$ that appears in the decomposition of $\mathcal{F}$ is the trivial sheaf.

Hence $\mathcal{F}$ is a middle extension from $\mathbb{G}_m$ to $\mathbb{A}^1$ of an iterated extension of constant sheaves tensored with Galois representations, thus a middle extension of a sheaf with unipotent geometric monodromy. $\square$

The two hypotheses on $\mathcal{F}$ this theorem can be easily checked in practice. For instance the sheaf $f^*\mathcal{L}_\chi$ for f a nonconstant polynomial with no repeated roots and χ the standard quadratic character has its trace function contained in $\mathbb{Q}$ and does not have unipotent geometric monodromy.

Instead of asking whether $a_p(\mathcal{F})$ are traces of Frobenius of a Galois representation, we could ask whether they are the Hecke eigenvalues of an automorphic form on GL_n , so that $a_p(\mathcal{F})$ would have an automorphic L -function. For Kloosterman sums, this question was asked by Katz [1980, Question 1.2.5.3].

One reason to hope that the $a_p(\mathcal{F})$ are Hecke eigenvalues is that it might allow one to find a connection between our group G_S and the conjectural group $L_\mathbb{Q}$ whose representations should classify automorphic forms over $\mathbb{Q}$, as described by Arthur [2002]. Most optimistically, G_S would be a quotient of $L_\mathbb{Q}$. However, we do not expect that $a_p(\mathcal{F})$ are the Hecke eigenvalues of an automorphic form (except when they are equal to the Frobenius traces of a Galois representation), and so we do not see any relationship between the two groups.

We can't completely rule out the possibility of describing $a_p(\mathcal{F})$ in this way, but we can prove some results that make it seem unlikely. First, observe that any automor-

phic form that contributes to the cohomology of a locally symmetric space must have its Hecke eigenvalues all contained in some fixed number field as each Hecke eigenvector is an eigenvector of an algebra acting on a finite-dimensional $\mathbb{Q}$ -vector space and so is defined over a finite extension of $\mathbb{Q}$. So whenever $\mathcal{F}$ satisfies the hypotheses of Proposition 8.0.121, $a_p(\mathcal{F})$ cannot be Hecke eigenvalues of an automorphic form appearing in cohomology.

We do not know any method to definitely rule out the existence of a non-cohomological automorphic form having $a_p(\mathcal{F})$ as its eigenvalues. However, the methods of Booker [2000, Section 3] can rule out the existence of automorphic forms on a given group with Hecke eigenvalues $a_p(\mathcal{F})$ if one imposes an upper bound on both the level and the eigenvalue of the Laplacian. For instance, Booker [2000, Theorem 1.1] showed that the Kloosterman sum $\mathcal{K}\ell(1, 1; p)$ does not equal the Hecke eigenvalues of a Maass form of level N and eigenvalue λ unless $N(\lambda + 3) > 18.3 \times 10^6$. One could perform the same computation on any other exponential sum and hope to get a similar bound, but never actually rule out automorphy.

Chapter 9

Numerical Experiments

In this chapter we collect some experimental data on cases of Conjecture 1 that, as far as we know, have not been considered or tested before. The goal is to demonstrate the plausibility of the conjecture beyond the previously-studied cases of it. In particular, we would like to find evidence to support the predictions of Chapter 7.

For each class of sheaves discussed in Chapter 7, we took some of the simplest nontrivial cases of it, computed the value of a_p for the first several primes, and compared those values to the predicted distribution. The comparison we used was a Kolmogorov-Smirnov test [Gnedenko, 1967, pp. 450-451] [Press et al., 1988, pp. 490-492]. This statistical test measures the difference between the predicted and observed distribution of a real-valued random variable by computing the supremum of the distance of the cumulative distribution functions of the two distribution. Because this test only allows for a single real variable, when studying a complex-valued distribution we perform the test on the real and imaginary parts separately.

We can describe the output of the Kolmogorov-Smirnov test in at least three ways. First, we can simply state the “discrepancy”, the supremum of the difference between the cumulative distribution functions. Second, we can take the Kolmogorov-Smirnov statistic, which is the product of the discrepancy and the square root of the number

of data points. This square root was chosen because the expected discrepancy were the data actually a random sample from the distribution would be proportional to the inverse of the square root of the number of data points. This product is called the Kolmogorov-Smirnov statistic. Third, we can compute the p-value, which is the probability that a larger Kolmogorov-Smirnov statistic would occur in a random sample. When the discrepancy and Kolmogorov-Smirnov statistic are small and the p-value is large, the data closely matches the distribution.

We also include some histograms to give a picture of how close the predicted and actual distributions are, but this does not represent the best statistical test because it involves arbitrarily splitting the data into bins.

If the Kolmogorov-Smirnov statistic is large this does not necessarily imply that the data is not converging to the distribution. It only means that the discrepancy between the observed data and the predicted distribution is greater than what one would expect for random data. The data arising from a number-theoretic computation is not random, and the error could converge to zero at some rate slower than the rate at which the error of random data converges to zero. This would represent a sequence with a substantial small-number bias: The distribution of a_p for a random large prime p would be similar to the predicted distribution, but the distribution of a_p for a random small prime is biased away from it. An example of this is the Kummer sum, which is believed to have small-number bias [Patterson, 1978, Conjecture on pp. 2]. This (conjectural) bias was very significant in the numerical data computed by Kummer [Kummer, 1846, pp. 353], causing him to make an inaccurate conjecture on the distribution of his sum. In most of the cases we studied, we observed no small-number bias. However, in a few of the cases, we expected and observed substantial small-number bias. These sums were of the form:

$$\sum_{x=0}^{p-1} e^{2\pi i f(x)/p}$$

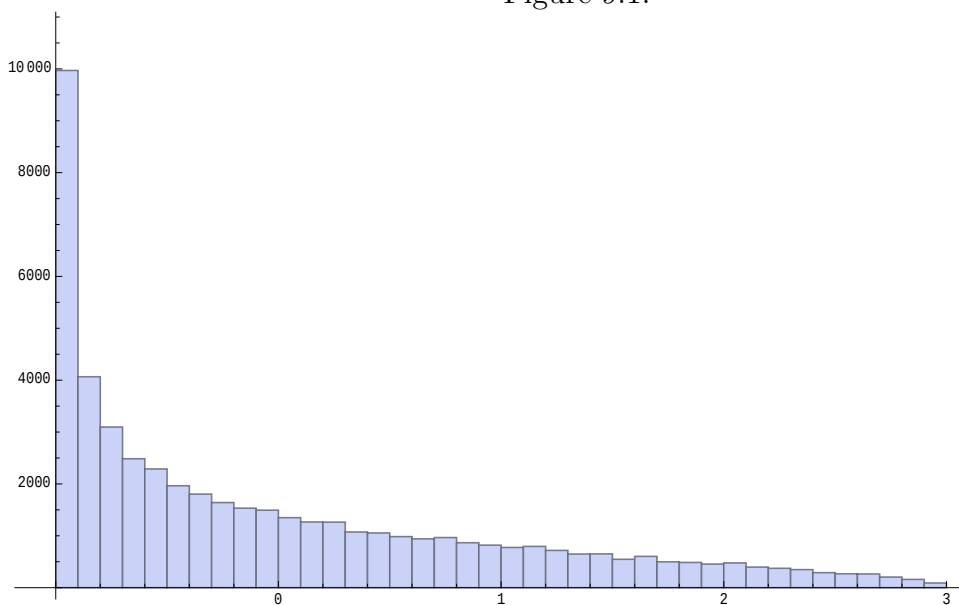
for a quartic polynomial f , and hence are similar to the Kummer sum $\sum_{x=0}^{p-1} e^{2\pi i x^3/p}$. We will describe the heuristic we used to estimate the size of this bias, and the more forgiving numerical tests we used to test the conjecture in those cases, in Chapter 10.

In this chapter, we will describe all the better-behaved numerical tests:

First, we considered Proposition 7.0.100. We checked whether the locations of the roots of a monic irreducible polynomial mod p are independent of the splitting behavior of p in a number field. For a quadratic polynomial and a number field abelian over $\mathbb{Q}$, this was already known by [Tóth, 2000, Theorem 1.2]. Thus we chose to study the case of a quadratic polynomial and a cubic number field, and we used the quadratic polynomials and cubic number fields with the smallest possible discriminants: polynomials $x^2 + 1$, $x^2 + x + 1$, $x^2 - x - 1$ and number fields $\mathbb{Q}(x)/(x^3 + x^2 - 1)$, $\mathbb{Q}(x)/(x^3 + x - 1)$, $\mathbb{Q}(x)/(x^3 - x^2 + x + 1)$.

For each of these three cubic fields and for each conjugacy class in S_3 , we found all the primes among the first 50,000 primes whose Frobenius element was in that conjugacy class. For each prime, we found the smallest positive integer root modulo p of each of these three quadratic polynomials that has a root modulo p and divided it by p . The predicted distribution from Proposition 7.0.100 is that, for each cubic field, each conjugacy class, and each quadratic polynomial, the normalized first root of that polynomial modulo primes in that conjugacy class should be distributed according to the uniform distribution on $[0, 1/2]$. For each triple we performed a Kolmogorov-Smirnov test of the data against the distribution, so that there were 27 tests in total. Each test used a subset of the primes - between 4,098 and 12,575 primes. The lowest p-value we obtained was .09. Since we performed 27 different tests, this is clear evidence that our conjecture fits the data, at least in this range. The largest normalized Kolmogorv-Smirnov statistic was 1.243 and the largest discrepancy was .018494.

Figure 9.1:



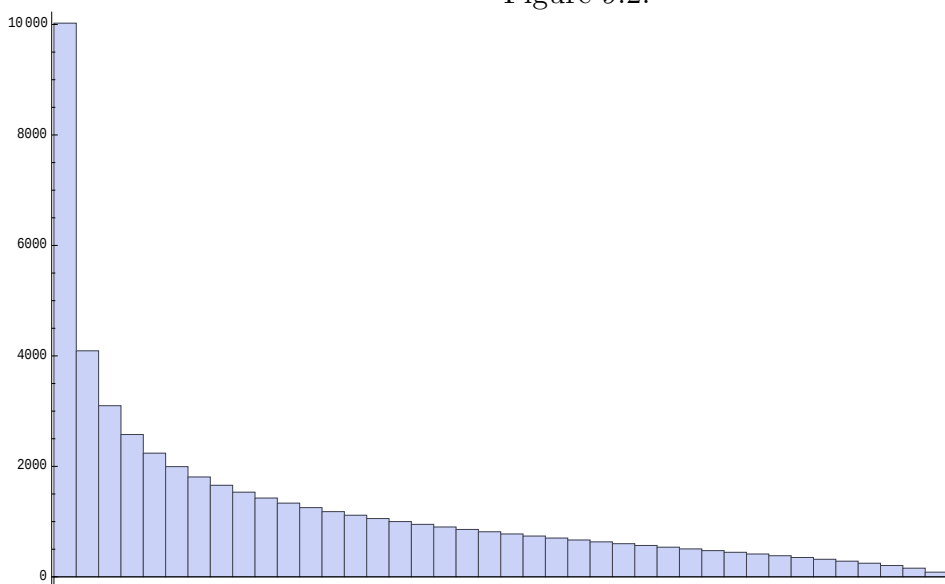
Next we considered the “twisted elliptic curve” sum:

$$\sum_{x=0}^{p-1} \left(\frac{4x^3 - g_2x - g_3}{p} \right) e^{2\pi i \frac{x}{p}}$$

The distribution of this sum, suitably normalized, was predicted in Proposition 7.0.106 to follow the Haar measure on SO_3 or SU_3 , depending on whether $g_3 = 0$. We sought examples of both distributions. For $g_3 = 0$, we chose the elliptic curve $y^2 = x^3 - x$, computed this sum for the first 50,000 primes, divided it by the normalizing factor ($\sqrt{p}$ if $p \equiv 1$ modulo 4 and $-i\sqrt{p}$ if $p \equiv 3$ modulo 4), and compared it to the Haar measure on SO_3 . The Kolmogorov-Smirnov statistic gave a p -value of .38, well within the range of random variation. In terms of histograms, one can see that this represents an almost perfect match:

The distribution function of the image of the Haar measure of $SO(3)$ under the trace map is $\frac{1}{2\pi} \sqrt{\frac{3-x}{x+1}}$, which has a pole at $x = -1$, explaining why the distribution is so lopsided.

Figure 9.2:



For $g_3 \neq 0$, we chose 9 simple curves. The method for determining which elliptic curves are simple was somewhat arbitrary - we chose curves with the smallest coefficients from the isogeny classes of the smallest conductor. These were $y^2 - y = x^3 - x^2$, $y^2 - y = x^3 + x$, $y^2 - xy - x = x^3 - x^2$, $y^2 - xy - y = x^3 + x^2 + x$, $y^2 - y = x^3 - x^2 - x$, $y^2 = x^3 - x^2 + x$, $y^2 - xy = x^3 - x$, $y^2 = x^3 + x^2 - x$, and $y^2 - xy - y = x^3$. We then performed the standard change of variables to put them in the form $y^2 = 4x^3 - g_2x - g_3$, and computed the sum a_p for the first 10,000 primes. For each curve, we compared the distribution of the real and imaginary parts of the normalized sum against the distribution of the real and imaginary parts of a random element of SU_3 . Since no exact formula is available, we used a Monte Carlo method to estimate these distributions [Mezzadri, 2007, pp. 597]. Thus there were 18 Kolmogorov-Smirnov tests in total. The lowest two p -values were .024, for the imaginary part of the $y^2 - xy - y = x^3 + x^2 + x$ sum, and .027, for the real part of the $y^2 - y = x^3 - x^2 - x$ sum. This number of false positives is reasonable for 18 tests. By the definition of p -value, were the a_p totally random, the probability that we would get a p -value less than .05 is .05. With 18 tests, we expect .9 p -values less than .5, and getting 2 instead is not a large deviation.

The next-smallest were a p -value of .059 for the real part of $y^2 - y = x^3 + x$ and a p -value of .105 for the real part of $y^2 - y = x^3 - x^2$. All of these reflect a discrepancy of at most .015 and a Kolmogorov-Smirnov statistic of at most 1.5.

We also studied a sum related to an elliptic surface that conjecturally gives G_2 equidistribution (Lemma 7.0.119). The relevant sum is easier to compute if the surface is in Weierstrass form. One Weierstrass form is:

$$3y^2 = x^3 - (3 + 24t^6)x + 2 - 40t^6 - 16t^{12}$$

(One can check that this is the same elliptic surface easily by comparing the singular fibers, and then observing that this elliptic surface is determined by its singular fibers up to a quadratic twist, and comparing the fibers at 0 to check that the quadratic twist is trivial.) We computed the sum for the first 1000 primes and compared it to the Haar measure on G_2 , which we estimated by using a Monte Carlo method to approximate the Weyl integration formula. The p -value obtained was .85 - the data fit the distribution almost exactly, given the small amount of data. In terms of histograms, compare the data for the first 1000 primes and a Monte Carlo sample of 100,000 random matrices:

Figure 9.3:

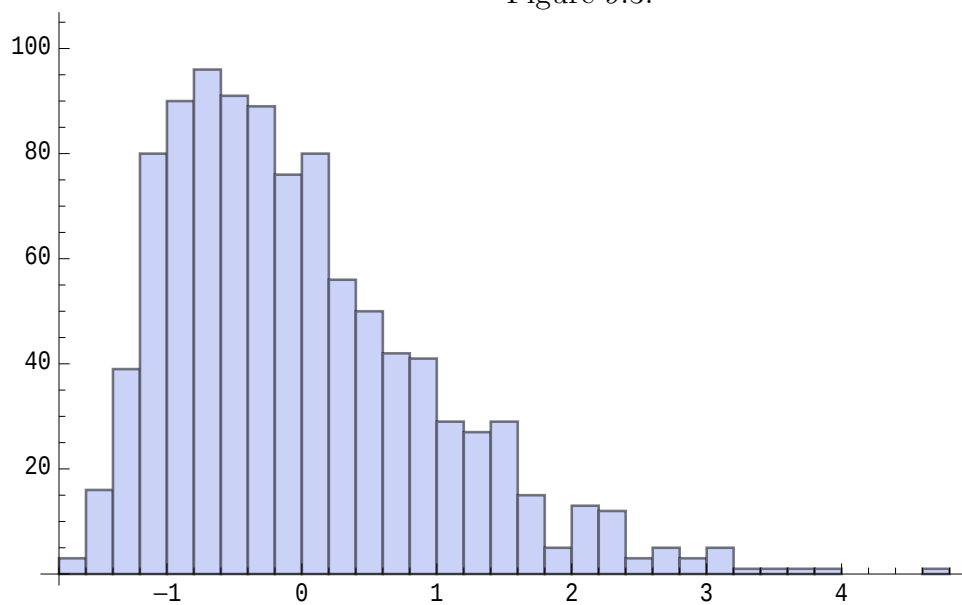
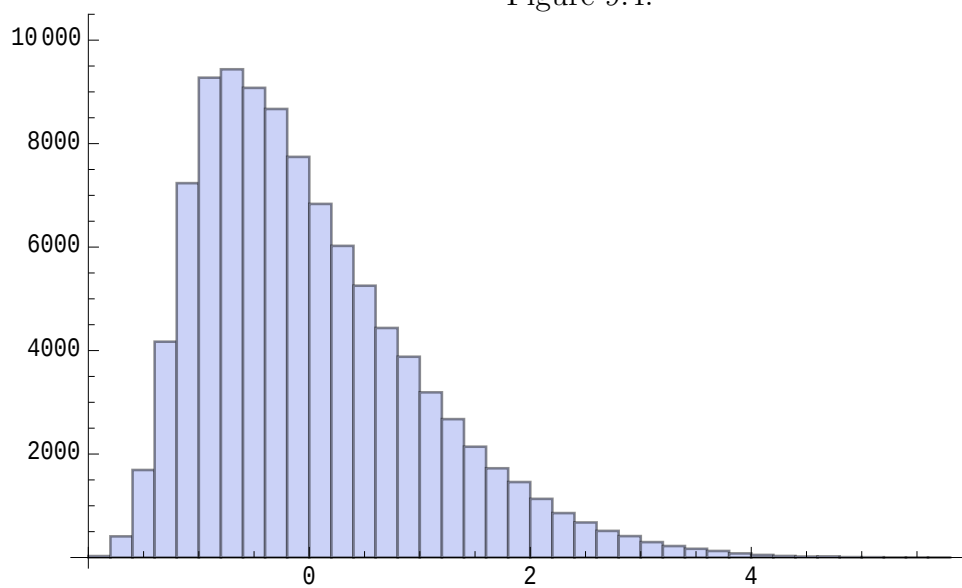


Figure 9.4:



Chapter 10

Small-Number Bias

Let $\mathcal{F}$ be a constructible ℓ -adic sheaf on $\mathbb{A}_{\mathbb{Z}[\frac{1}{N}]}^1$ that is pure of weight w . Suppose that Conjecture 2 implies that

$$\lim_{N \rightarrow \infty} \frac{\sum_{p < N} \frac{a_p(\mathcal{F})}{p^{w/2}}}{\pi(N)} = 0$$

and suppose that we believe Conjecture 2. How quickly do we expect this fraction goes to 0 as a function of N ? The most ambitious guess would be that it satisfies square-root cancellation, so that we have a bound of the form $\frac{\sum_{p < N} a_p(\mathcal{F})/p^{w/2}}{\pi(N)} = O(N^{-1/2+\epsilon})$ for each $\epsilon > 0$. If for some $\mathcal{F}$ the sum converges to 0, but does not satisfy a bound of that form, then we say that $\mathcal{F}$ has small-number bias.

The question of which sheaves have square-root cancellation and which have small-number bias is interesting not just for reasons of curiosity, but because it tells us how to properly interpret a Kolmogorov-Smirnov test. Let $\mathcal{F}$ be a sheaf, pure of weight w , and of Tannakian dimension n . Furthermore assume for simplicity that $a_p(\mathcal{F})$ is always real. Then $a_p(\mathcal{F})/p^{w/2}$ is a real number in the interval $[-n, n]$. Let μ be a measure on $[-n, n]$. Consider the discrepancy d_N of the uniform measure on the set

$\{a_p(\mathcal{F})/p^{w/2} | p < N\}$ and the measure μ . Then by integration by parts:

$$\int_{-n}^n \left(\int_x^n d\mu \right) dx = \int_{-n}^n \left(\int_{-n}^y dx \right) d\mu_y = \int_{-n}^n (y+n) d\mu_y = \int_{-n}^n y d\mu_y + n$$

Applying the same formula to the measure that puts weight $\frac{1}{\pi(N)}$ on $a_p(\mathcal{F})$ for each $p < N$, we obtain

$$\int_{-n}^n \left(\frac{|\{p | p < N, a_p(\mathcal{F})/p^{w/2} < x\}|}{\pi(N)} \right) dx = \frac{\sum_{p < N} a_p(\mathcal{F})/p^{w/2}}{\pi(N)} + n$$

Therefore:

$$\begin{aligned} \left| \frac{\sum_{p < N} a_p(\mathcal{F})/p^{w/2}}{\pi(N)} - \int_{-n}^n y d\mu_y \right| &= \left| \int_{-n}^n \left(\frac{|\{p | p < N, a_p(\mathcal{F})/p^{w/2} < x\}|}{\pi(N)} - \int_{-n}^x d\mu \right) dx \right| \\ &\leq \int_{-n}^n \left| \frac{|\{p | p < N, a_p(\mathcal{F})/p^{w/2} < x\}|}{\pi(N)} - \int_{-n}^x d\mu \right| dx \leq \int_{-n}^n d_N dx \leq 2nd_N \end{aligned}$$

Let $K_N = \sqrt{\pi(N)}d_N$ be the Kolmogorov-Smirnov statistic. Then we have

$$\left| \frac{\sum_{p < N} a_p(\mathcal{F})/p^{w/2}}{\pi(N)} - \int_y d\mu_y \right| \leq \frac{2n}{\sqrt{\pi(N)}} K_N$$

Hence if the left side is not bounded by $N^{-1/2+\epsilon}$, the Kolmogorov-Smirnov statistic must grow faster than N^ϵ . Hence the sheaf will fail the Kolmogorov-Smirnov test by an arbitrarily large amount, even if $a_p(\mathcal{F})$ is equidistributed according to μ .

For most of the sums we considered, there is no reason to expect small number bias, and indeed our predictions for those sheaves passed the Kolmogorov-Smirnov test. However there is one family of sums for which we have good reason to expect significant bias.

The exponential sums in question are the sums $\sum_{x \in \mathbb{F}_p} e^{2\pi i f(x)/p}$ for polynomials $f(x)$, corresponding to the sheaves $f_* \overline{\mathbb{Q}}_\ell$. For the same sum over $\mathbb{Z}/n$ instead of over

$\mathbb{F}_p$, [Patterson, 2003, Conjecture 4.1] gave a conjectural formula for the average of $\sum_{x \in \mathbb{Z}/n} e^{2\pi i f(x)/n}$ over $n < N$. The cubic and quartic cases of this conjecture are not known, but a modified version where the sum is taken over $\mathcal{O}_K/c$ for $\mathcal{O}_K$ the ring of integers of a fixed number field K and c an element of $\mathcal{O}_K$ - by [Livné and Patterson, 2002, Theorem 1.2] for cubic polynomials with $K = \mathbb{Q}(\mu_3)$ and by [Herman, 2013, Theorem 1.3] for quartic polynomials with $K = \mathbb{Q}(\mu_8)$.

A heuristic argument suggests that an analogue of Patterson's conjecture holds in our situation. Patterson predicts that, for f a polynomial of degree d , not satisfying the equation $f(x) = f(a - x)$ for any rational a , there is a constant C_f such that:

$$\frac{\sum_{c < N} \sum_{x \in \mathbb{Z}/c} e^{2\pi i f(x)/c}}{N} = C_f N^{1/d} + o(N^{1/d})$$

and gives a formula for the constant C_f .

The heuristic implies that f a polynomial of degree d , there is a constant C_f such that

$$\frac{\sum_{p < N} \frac{\sum_{x \in \mathbb{F}_p} e^{2\pi i f(x)/p}}{\sqrt{p}}}{\pi(N)} = C_f N^{1/d-1/2} + o(N^{1/d-1/2})$$

These formulas are very similar except that our exponent is reduced by $1/2$ because we have divided our sum by $\sqrt{p}$. Also, our heuristic suggests that the same formula holds for all polynomials - we will use additional information to restrict its application. As a simple test of this heuristic, note that, in the case $d = 2$ this heuristic suggests that the average over all primes

$$\lim_{N \rightarrow \infty} \frac{\sum_{p < N} \frac{\sum_{x \in \mathbb{F}_p} e^{2\pi i f(x)/p}}{\sqrt{p}}}{\pi(N)} = C_f$$

may be nonzero. This suggestion is correct: For f a quadratic polynomial, $e^{2\pi i f(x)/p}$ is a Gauss sum, and the average over primes of normalized Gauss sums is nonzero.

We do not conjecture that this heuristic is precisely correct, and certainly not for all polynomials. After we derive the heuristic we will describe why it is incomplete.

Now we explain the heuristic. Recall that we are estimating the sum

$$\frac{\sum_{p < N} \frac{\sum_{x=0}^{p-1} e^{2\pi i \frac{f(x)}{p}}}{\sqrt{p}}}{\pi(N)}$$

The heuristic is based on the idea that we can divide the inner sum $\sum_{x=0}^{p-1} e^{2\pi i \frac{f(x)}{p}}$ into sums over two regions. We can evaluate the sum over the first region explicitly and show that it does not cancel when summed over p . We do not have a simple formula for the sum over the second region and so we optimistically predict that the sum on that region behaves randomly and cancels when summed over p . Under that assumption, the contribution from the first region will be larger than the second and so will dominate the size of the sum.

The first region consist of all x in the interval $[-p^{1/d+\delta}, p^{1/d+\delta}]$ for any δ with $0 < \delta < 1/d^2$, and the second region is the remainder.

Lemma 10.0.122. *Let $f(x) = a_d x^d + a_{d-1} x^{d-1} + \cdots + a_0$. Assume that for all i , $|a_i| < A$. Then for any positive real C ,*

$$\left| \sum_{x \in \mathbb{Z}, |x| < Cp^{1/d}} e^{2\pi i \frac{f(x)}{p}} - p^{1/d} \int_{-C}^C e^{2\pi i a_d x^d} dx \right| \leq 2 + 8\pi AC^d$$

Proof. If $|x| < Cp^{1/d}$ then

$$\left| \frac{f(x)}{p} - \frac{a_d x^d}{p} \right| = \left| \sum_{i=0}^{d-1} \frac{a_i x^i}{p} \right| \leq \sum_{i=0}^{d-1} \left| \frac{a_i x^i}{p} \right| \leq \sum_{i=0}^{d-1} AC^i p^{i/d-p} \leq dAC^{d-1} p^{-1/d}$$

hence

$$\left| e^{2\pi i \frac{f(x)}{p}} - e^{2\pi i \frac{a_d x^d}{p}} \right| \leq 2\pi dAC^{d-1} p^{-1/d}$$

so

$$\left| \sum_{x \in \mathbb{Z}, |x| < Cp^{1/d}} e^{2\pi i \frac{f(x)}{p}} - \sum_{x \in \mathbb{Z}, |x| < Cp^{1/d}} e^{2\pi i \frac{a_d x^d}{p}} \right| \leq 4\pi d A C^d$$

Moreover if $|x| < Cp^{1/d}$ then

$$\left| \frac{d}{dx} e^{2\pi i \frac{a_d x^d}{p}} \right| = \left| 2\pi i \frac{da_d x^{d-1}}{p} e^{2\pi i \frac{a_d x^d}{p}} \right| = \left| \frac{da_d x^{d-1}}{p} \right| < 2\pi d A C^{d-1} p^{-1/d}$$

Hence by the bound for the error term in Riemann integration:

$$\left| \sum_{x \in \mathbb{Z}, |x| < Cp^{1/d}} e^{2\pi i \frac{a_d x^d}{p}} - \int_{-Cp^{1/d}}^{Cp^{1/d}} e^{2\pi i \frac{a_d x^d}{p}} dx \right| \leq 2 + 2Cp^{1/d} (2\pi d A C^{d-1} p^{-1/d}) = 2 + 4\pi C^d A$$

By change of variables

$$\int_{-Cp^{1/d}}^{Cp^{1/d}} e^{2\pi i \frac{a_d x^d}{p}} dx = p^{1/d} \int_{-C}^C e^{2\pi i a_d x^d} dx$$

so:

$$\left| \sum_{x \in \mathbb{Z}, |x| < Cp^{1/d}} e^{2\pi i \frac{f(x)}{p}} - p^{1/d} \int_{-C}^C e^{2\pi i a_d x^d} dx \right| \leq 2 + 8\pi A C^d$$

□

Lemma 10.0.123. *Let C, D be real numbers with $D > C > 0$.*

$$\left| \int_{-C}^C e^{2\pi i a_d x^d} dx - \int_{-D}^D e^{2\pi i a_d x^d} dx \right| \leq \frac{2}{\pi d a_d C^{d-1}}$$

Proof. This difference is equal to $\int_C^D e^{2\pi i a_d x^d} dx + \int_{-D}^{-C} e^{2\pi i a_d x^d} dx$. Because the two terms are conjugate to each other, it suffices to prove the bound

$$\left| \int_C^D e^{2\pi i a_d x^d} dx \right| \leq \frac{1}{\pi d a_d C^{d-1}}$$

By integration by parts with $u = \frac{1}{2\pi i da_d x^{d-1}}$, $v = e^{2\pi i a_d x^d}$, we have $udv = e^{2\pi i a_d x^d}$

so

$$\begin{aligned} \left| \int_C^D e^{2\pi i a_d x^d} dx \right| &= \left| \frac{e^{2\pi i a_d x^d}}{2\pi i da_d x^{d-1}} \Big|_C^D - \int_C^D \frac{-(d-1)e^{2\pi i a_d x^d}}{2\pi i da_d x^d} dx \right| \\ &\leq \frac{1}{2\pi da_d C^{d-1}} + \frac{1}{2\pi da_d D^{d-1}} + \int_C^D \frac{d-1}{2\pi da_d x^d} dx = \frac{1}{2\pi da_d C^{d-1}} + \frac{1}{2\pi da_d C^{d-1}} - \frac{1}{2\pi da_d D^{d-1}} = \frac{1}{\pi da_d C^{d-1}} \end{aligned}$$

as desired. $\square$

It immediately follows that the limit $\lim_{C \rightarrow \infty} \int_{-C}^C e^{2\pi i a_d x^d} dx$ exists. Let $M_{a,d}$ be this limit, we taking D to ∞ we have

$$\left| \int_{-C}^C e^{2\pi i a_d x^d} dx - M_{a,d} \right| \leq \frac{2}{\pi da_d C^{d-1}}$$

Corollary 10.0.124. *Let δ satisfy $0 < \delta < 1/d^2$. Then we have*

$$\sum_{x \in \mathbb{Z}, |x| < p^{1/d+\delta}} e^{2\pi i \frac{f(x)}{p}} = p^{1/d} M_{a,d} + o(p^{1/d})$$

Proof. By Lemma 10.0.122:

$$\left| \sum_{x \in \mathbb{Z}, |x| < p^{1/d+\delta}} e^{2\pi i \frac{f(x)}{p}} - p^{1/d} \int_{-p^\delta}^{p^\delta} e^{2\pi i \frac{a_d x^d}{p}} dx \right| \leq 2 + 8\pi A p^{d\delta}$$

By Lemma 10.0.123:

$$\left| p^{1/d} \int_{-p^\delta}^{p^\delta} e^{2\pi i a_d x^d} dx - p^{1/d} M_{a,d} \right| \leq \frac{2p^{1/d-(d-1)\delta}}{\pi da_d}$$

By assumption $\delta < 1/d^2$ so $2 + 6\pi A p^{d\delta} = o(p^{1/d})$ and $\delta > 0$ so $\frac{2p^{1/d-(d-1)\delta}}{\pi da_d} = o(p^{1/d})$.

The claim follows. $\square$

Now we can split our sum:

$$\sum_{x=0}^{p-1} e^{2\pi i \frac{f(x)}{p}} = \sum_{x \in \mathbb{Z}, |x| < p^{1/d+\delta}} e^{2\pi i \frac{f(x)}{p}} + \sum_{x \in \mathbb{Z}, p^{1/d+\delta} \leq x \leq p-p^{1/d+\delta}} e^{2\pi i \frac{f(x)}{p}}$$

The first term is what we have an asymptotic for, from Corollary 10.0.124. For the second term we only have a bound - it is of size $O(\sqrt{p})$, because the complete sum is $O(\sqrt{p})$ by the Weil bound and the first part is $o(\sqrt{p})$, so their difference remains $O(\sqrt{p})$. The second term, however, appears to behave randomly. Our heuristic is that, summing over $p < N$, the second term has square root cancellation. This implies:

$$\sum_{p < N} \frac{\sum_{x=0}^{p-1} e^{2\pi i \frac{f(x)}{p}}}{\sqrt{p}} = \sum_{p < N} \frac{\sum_{x \in \mathbb{Z}, |x| < p^{1/d+\delta}} e^{2\pi i \frac{f(x)}{p}}}{\sqrt{p}} + \sum_{p < N} \frac{\sum_{x \in \mathbb{Z}, p^{1/d+\delta} \leq x \leq p-p^{1/d+\delta}} e^{2\pi i \frac{f(x)}{p}}}{\sqrt{p}}$$

(by Corollary 10.0.124 and by heuristic assumption)

$$= \sum_{p < N} (p^{1/d-1/2} M_{a,d} + o(p^{1/d-1/2})) + O(N^{1/2+\epsilon})$$

$$= M_{a,d} \sum_{p < N} p^{1/d-1/2} + \sum_{p < N} o(p^{1/d-1/2}) + O(N^{1/2+\epsilon})$$

(using the prime number theorem to show that $\sum_{p < N} p^{1/d-1/2} = \frac{\pi(N)N^{1/d-1/2}}{1/d+1/2} + o(\pi(N)N^{1/d-1/2})$)

$$= M_{a,d} \left(\frac{\pi(N)N^{1/d-1/2}}{1/d+1/2} + o(\pi(N)N^{1/d-1/2}) \right) + o(\pi(N)N^{1/d-1/2}) + O(N^{1/2+\epsilon})$$

(taking $\epsilon < 1/d$, we may group the error terms)

$$= \frac{M_{a,d}}{1/d+1/2} \pi(N)N^{1/d-1/2} + o(\pi(N)N^{1/d-1/2})$$

So our heuristic estimate is:

$$\frac{\sum_{p < N} \frac{\sum_{x=0}^{p-1} e^{2\pi i \frac{f(x)}{p}}}{\sqrt{p}}}{\pi(N)} = \frac{M_{a,d}}{1/d + 1/2} N^{1/d-1/2} + o(N^{1/d-1/2})$$

This gives a simple picture of the decline rate we might expect for these sums. However, there are some additional complexities that this picture does not account for, which suggest ways the heuristic could be changed.

First, we could modify the heuristic so that the first term is not just the sum over the interval $|x| < p^{1/d+\delta}$, but is the sum over a disjoint union of intervals of the form $|x - ab^{-1}| < p^{1/d+1/\delta}$ for all pairs of integers (a, b) , satisfying $0 \leq a < b < p^\delta$ (analogous to the major arcs in the circle method). This short sum can be estimated as an adelic integral, and by evaluating it we would obtain a more refined heuristic formula for the constant in the leading term. The leading term would again be a constant times $N^{1/d-1/2}$, but the constant would be an oscillatory adelic integral, similar to the formula of [Patterson, 2005, Conjecture 2.2, connected to integral on pp. 93]. Because the exact constant term is not important here, we will not go into detail about this more difficult argument.

Second, the heuristic does not tell us what to do when $f_*\overline{\mathbb{Q}_\ell}$ splits into a direct sum of sheaves. For instance, suppose f is equal to a composition $f = g \circ h$. Then $g_*\overline{\mathbb{Q}_\ell}$ is a summand of $f_*\overline{\mathbb{Q}_\ell}$. Because the degree of g is less than the degree of f , the heuristic predicts that the average of $a_p(g_*\overline{\mathbb{Q}_\ell})$ over primes $p < N$ will be much larger than the average of $a_p(f_*\overline{\mathbb{Q}_\ell})$. This is only possible if the average of $a_p(f_*\overline{\mathbb{Q}_\ell}/g_*\overline{\mathbb{Q}_\ell})$ over primes $p < N$ is almost exactly the same size as the average of $a_p(f_*\overline{\mathbb{Q}_\ell})$, but of the opposite sign, so that the two contributions cancel. If this seems unlikely then we might doubt the heuristic for these f .

Finally, the heuristic does not tell us anything about sheaves not of the form $f_*\overline{\mathbb{Q}_\ell}$ for some polynomial f . The empirical evidence from the previous chapter seems to

suggest that many sheaves satisfy full square-root cancellation, but we don't know why or which ones don't. It would be very interesting to expand the heuristic into a plausible conjecture that describes the order of cancellation associated to each sheaf.

Remark 10.0.125. Perhaps the heuristic applies to some f but not others. We might guess, by analogy to Patterson's conjecture [Patterson, 2003, Conjecture 4.1], that the estimate

$$\frac{\sum_{p < N} \frac{\sum_{x \in \mathbb{F}_p} e^{2\pi i f(x)/p}}{\sqrt{p}}}{\pi(N)} = C_f N^{1/d-1/2} + o(N^{1/d-1/2})$$

coming from our heuristic holds whenever $f(x) \neq f(a-x)$ for any a .

Were this true, it would contradict Conjecture 2.

To see this, take $f(x) = g(x)^2$ where g is a cubic polynomial that does not satisfy $g(x) = -g(a-x)$ for any a . Then $f(x) \neq f(a-x)$ for any a . Let $a_p = \sum_x e^{2\pi i f(x)/p} / \sqrt{p}$. Then this heuristic predicts that the average value of a_p over the first N primes is asymptotic to $N^{1/6-1/2}$, and in particular converges to 0. However, $f_* \overline{\mathbb{Q}}_\ell$ contains as a summand $[2]_* \overline{\mathbb{Q}}_\ell$ where $[2]$ is the squaring map and contains no other subobjects that are 1-dimensional in $\mathcal{CC}_S$. The sum associated to $[2]_* \overline{\mathbb{Q}}_\ell$ is the Gauss sum $\sum_x e^{2\pi i x^2/p}$ which is $\sqrt{p}$ if p is congruent to 1 mod 4 and $i\sqrt{p}$ if p is congruent to 3 mod 4. Therefore the average of $\sum_x e^{2\pi i x^2/p} / \sqrt{p}$ over the prime $< N$ converges to $\frac{1+i}{2}$. Conjecture 2 predicts the averages of the other terms converge to 0 which implies that

$$\lim_{N \rightarrow \infty} \frac{\sum_{p < N} \frac{\sum_{x \in \mathbb{F}_p} e^{2\pi i f(x)/p}}{\sqrt{p}}}{\pi(N)} = \frac{1+i}{2}$$

and does not converge to 0..

Numerical experiments suggest that for $f(x) = (x^3 + x + 1)^2$, Conjecture 2 is accurate. We computed the empirical average of the exponential sum $\sum_{x \in \mathbb{F}_p} e^{2\pi i (x^3 + x + 1)^2/p} / \sqrt{p}$ over the first 50,000 primes, and it was $.523 + .502i$, which is not far from $\frac{1+i}{2}$. We also computed the average over the first N primes for $N < 50,000$, and the error seemed to decrease as a function of N and not

oscillate wildly. In particular, the average of the exponential sum showed no signs of converging to 0. Hence this bold version of our heuristic, analogous to Patterson's conjecture appears to be inaccurate. We stress that at no point did Patterson actually conjecture this analogue.

Regardless of these details, we may still use the heuristic for one key purpose - it tells us what tests to perform on the sums $a_p(\mathcal{F})$ to see if their distributions are converging. If the average of the sum converges to 0 at a rate similar to $N^{1/d-1/2}$, then we could also expect the discrepancy between the distribution of the sum and the limiting distribution to converge to 0 at a rate similar to $N^{1/d-1/2}$. By plotting the discrepancy as a function of N , if it appears to be converging to 0 at that rate, we could hope it continues to converge at that rate as N goes to ∞ . We did this for sums of the form

$$\sum_{x=0}^{p-1} e^{2\pi i \frac{f(x)}{p}}$$

for $f(x)$ a quartic polynomial. The quartic polynomials were selected so that we could observe the distinction between the exponential sum associated to a generic quartic polynomial, which we expect from Lemma 7.0.110 is equidistributed according to SU_3 after suitable normalization, and sums associated to special quartic polynomials, which we expect from Lemma 7.0.112 will follow a different distribution. We looked at three of the polynomials mentioned in Patterson [2003]: $f(x) = x^4 + x$, $f(x) = x^4 + x^2$, and $f(x) = x^4 + (x + 1)^4$. The first one satisfies the conditions of Lemma 7.0.110 while the last two satisfy the conditions of Lemma 7.0.112.

For $f(x) = x^4 + x^2$ or $x^4 + (x + 1)^4$, we have computed in Chapter 7 the expected distribution as a convex combination of four rotated and translated copies of the Sato-Tate measure. This can be visualized using four line segments in the complex plane such that the sum, divided by $\sqrt{p}$, is expected to fall on each line segment with equal probability, and on each line segment it is expected to follow the Sato-Tate measure. The fact that the sum lies on the line segment follows from Proposition

5.0.88 so our numerical experiments focused on the position of the sum on the line segment. We subtracted the translation and rotation factors from a_p and compared against the Sato-Tate measure. We considered each of the congruence classes of p modulo 8 separately, because for each congruence class we have two choices about how to project onto the line segment, but as long as we consider them separately, this only represents a global change of sign for the random variable, which does not affect the discrepancy.

We performed a dyadic discrepancy test. For each polynomial and each congruence class mod 8, we divided the first 2^{15} primes in that congruence class into a sequence of disjoint intervals, with the n th interval containing exactly 2^n primes. For each such interval, we computed the discrepancy between the observed and predicted distribution, as in the Kolmogorov-Smirnov test. We then took these logs and plotted them as a function of n . If the discrepancy is a constant times $N^{-1/4} \approx 2^{-n/4}$, we should expect this log plot to be approximately linear of slope $-(\log 2)/4$. If the plot does not appear linear, but instead appears to decrease for a time and then stabilize, that would be evidence against Conjecture 1. The plots all appeared roughly linear, although the slope was not always clearly equal to $-(\log 2)/4$ - sometimes it was larger. So this give some evidence in favor of Conjecture 1. Here are two of these plots, with both polynomials summed over the first 2^{15} primes congruent to 1 modulo 8:

For $f(x) = x^4 + x$, we apply Lemma 7.0.110 to compute the Tannakian monodromy group. We check the conditions: The derivative of $f(x)$ is $4x^3 - 1$ and has three distinct roots. The critical values of $f(x)$ are of the form $(x^3 + 1)x$ for $x^3 = -1/4$, so they are all of the form $(3/4) \cdot (-1/4)^{1/3}$. These three numbers satisfy no nontrivial relations of the form $f(\alpha_1) + f(\alpha_2) = f(\alpha_3) + f(\alpha_4)$. So its Tannakian monodromy group is GL_3 . Using Lemma 7.0.109, we compute that its determinant is $\mathcal{L}_\chi[1](1) \otimes \chi_2$. So we can divide by a third root of the determinant, such as $\mathcal{L}_\chi[1] \otimes \chi_{-2}$, to obtain a sum

Figure 10.1:

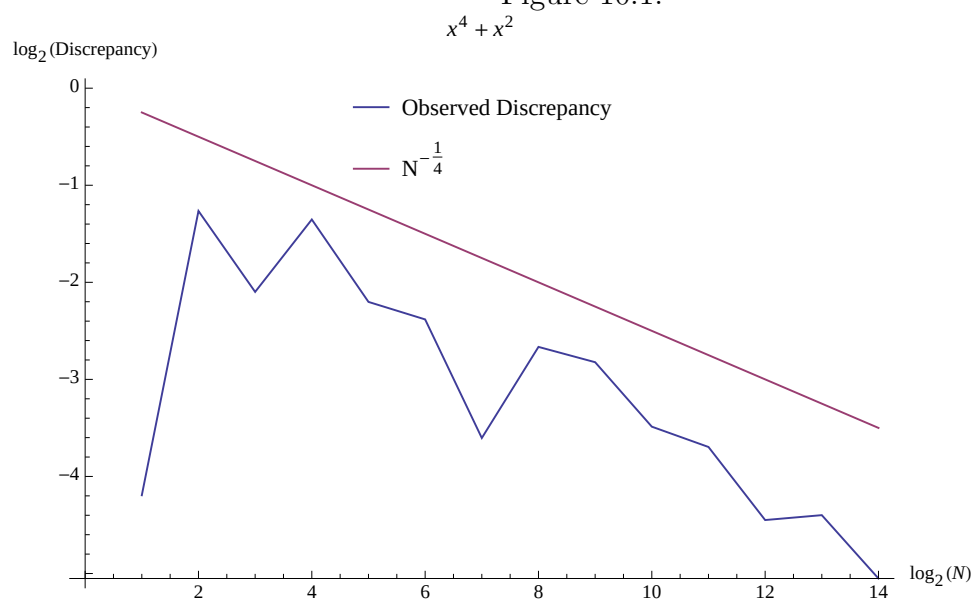


Figure 10.2:

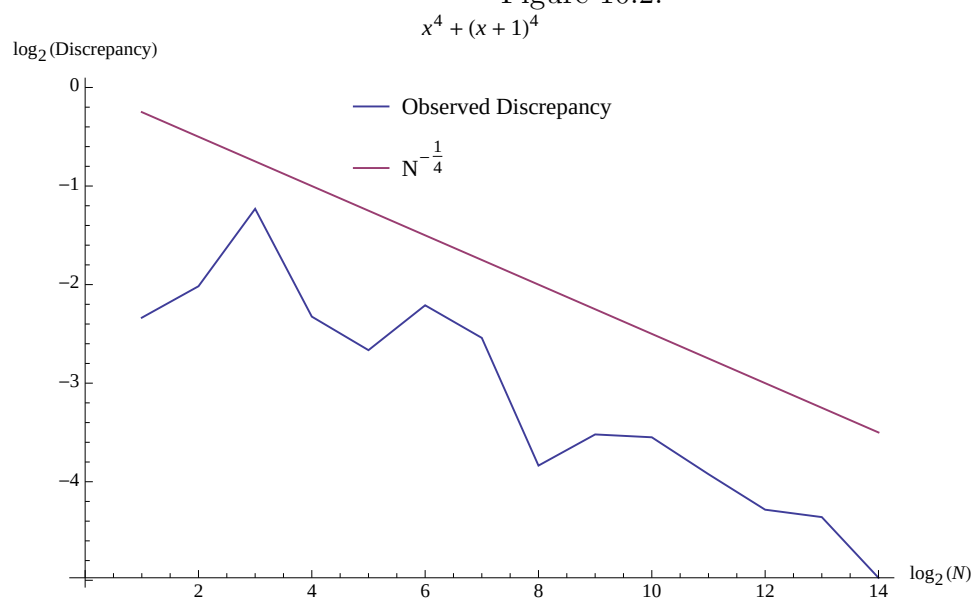
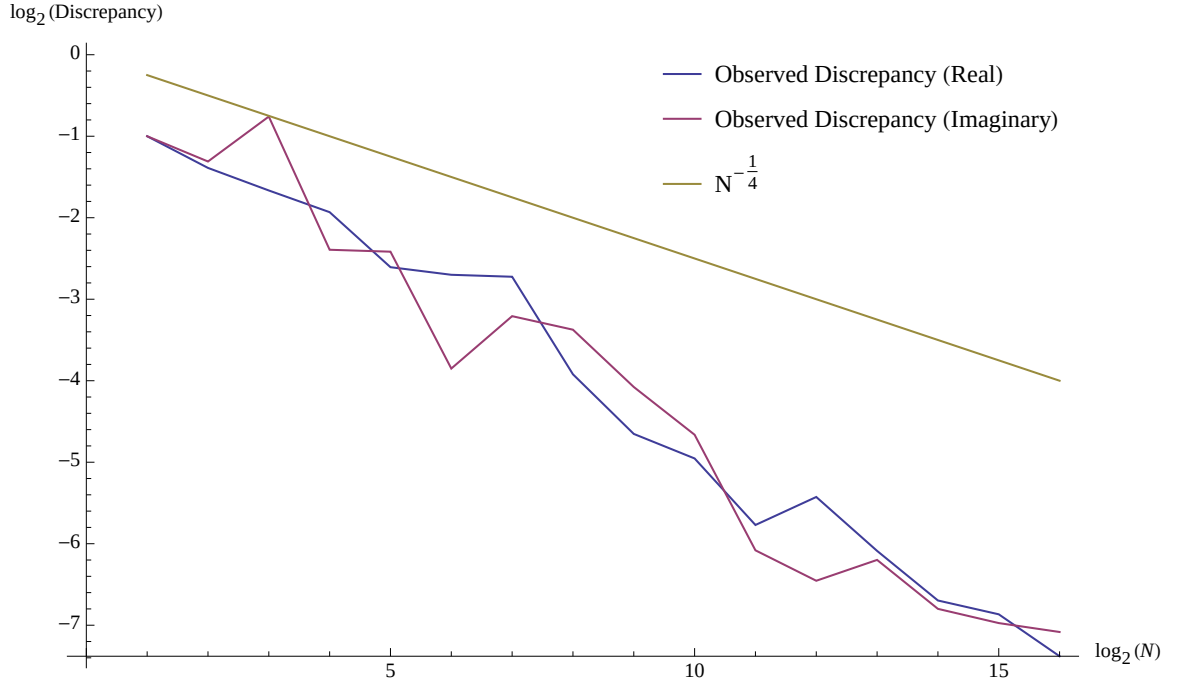


Figure 10.3:
 $x^4 + x$



whose conjectural distribution is the Haar measure of SU_3 . We performed dyadic discrepancy tests to test the conjecture that this ratio is equidistributed according to the image under the trace of the Haar measure of SU_3 . Because the discrepancy is only defined for real random variables, we considered the real and imaginary parts separately. We computed a_p for the first 2^{17} primes, and divided it by the Gauss sum and the quadratic character. We estimated the Haar measure of SU_3 using Weyl's integration formula and Monte Carlo integration. The graph appears linear, so again Conjecture 1 seems to hold up, although the slope appears somewhat larger than the predicted slope of $1/4$.

Bibliography

- Arthur, J. (2002). A note on the automorphic Langlands group. *Canadian Mathematical Bulletin*, 45(4):466–482.
- Artin, M., Grothendieck, A., and Verdier, J.-L. (1972). *Séminaire de Géométrie Algébrique du Bois Marie - 1963-64 - Théorie des topos et cohomologie étale des schémas - vol .3*, volume 305 of *Lecture Notes in Mathematics*. Springer-Verlag.
- Beilinson, A. and Bernstein, J. (1993). A proof of Jantzen conjectures. *Advances in Soviet Mathematics*, 16(Part 1):1–50.
- Beilinson, A., Bernstein, J., and Deligne, P. (1982). Faisceaux pervers. In *Analysis and topology on singular spaces, I (Luminy, 1981)*, volume 100 of *Asterisque*, pages 5–171. Société Mathématique de France.
- Beilinson, A. A. (1987). On the derived category of perverse sheaves. In Manin, Y., editor, *K-Theory, Arithmetic, and Geometry*, volume 1289 of *Lecture Notes in Mathematics*, pages 27–41. Springer-Verlag.
- Bierstone, E. and Milman, P. (1991). A simple constructive proof of canonical resolution of singularities. In Mora, T. and Traverso, C., editors, *Effective Methods in algebraic geometry*, volume 94 of *Progress in Mathematics*, pages 11–30. Birkhäuser, Boston.
- Booker, A. R. (2000). A test for identifying Fourier coefficients of automorphic forms and application to Kloosterman sums. *Experimental Mathematics*, 9(4):571–581.
- Brinon, O. and Conrad, B. (2009). CMI summer school notes on p-adic Hodge theory. <http://math.stanford.edu/~conrad/papers/notes.pdf>.
- Brylinski, J. (1986). Transformations canoniques, dualité projective, théorie de Lefschetz, transformation de Fourier et sommes trigonométriques. In *Géométrie et Analyse Microlocales*, volume 140-141 of *Astérisque*, pages 3–134. Société Mathématique de France.
- Deligne, P. (1970). Théorie de Hodge 1. *Actes Congrès intern. math.*, 1:425–430.
- Deligne, P. (1977). *SGA4 1/2 Cohomologie étale*, volume 569 of *Lecture Notes in Mathematics*. Springer.

- Deligne, P. (1980). La conjecture de Weil. II. *Publications mathématiques de l'I.H.É.S.*, 52:137–252.
- Deligne, P. (1990). Catégories tannakiennes. In Cartier, P., Illusie, L., Katz, N. M., Laumon, G., Manin, Y. I., and Ribet, K. A., editors, *The Grothendieck Festschrift vol. II Progress in Mathematics*, volume 87, pages 111–195. Birkhäuser.
- Deligne, P. and Katz, N. (1973). *Séminaire de Géométrie Algébrique du Bois Marie - 1967-79 - Groupes de monodromie en géométrie algébrique (SGA7) - vol. 2*. Number 340 in Lecture Notes in Mathematics. Springer-Verlag.
- Deligne, P. and Milne, J. S. (1982). Tannakian categories. In *Hodge Cycles, Motives, and Shimura Varieties*, volume 900 of *Lecture Notes in Mathematics*. Springer-Verlag.
- Duke, W., Friedlander, J. B., and Iwaniec, H. (1995). Equidistribution of roots of a quadratic congruence to prime moduli. *Annals of Mathematics*, 141(2):423–441.
- Duke, W., Friedlander, J. B., and Iwaniec, H. (2012). Weyl sums for quadratic roots. *International Mathematics Research Notices*, 11:2493–2594.
- Faltings, G. (1988). p -adic Hodge theory. *Journal of the American Mathematical Society*, 1(1):255–299.
- Gabber, O. and Loeser, F. (1996). Faisceaux pervers ℓ -adiques sur un tore. *Duke Mathematical Journal*, 83(3):501–606.
- Gabriel, P. (1962). Des catégories abéliennes. *Bulletin de la Société Mathématique de France*, 90:323–448.
- Gnedenko, B. V. (1967). *The Theory of Probability*. Chelsea Publishing Company.
- Grothendieck, A. (1977). *Cohomologie l -adique et Fonctions L (SGA 5, Séminaire de Géométrie Algébrique du Bois-Marie 1965/1966)*, volume 589 of *Lecture Notes in Mathematics*. Springer-Verlag.
- Grothendieck, A. and Raynaud, M. (1960). *Revêtements étales et groupe fondamental (SGA 1, Séminaire de Géométrie Algébrique du Bois-Marie 1960/1961)*, volume 224 of *Lecture Notes in Mathematics*. Springer-Verlag.
- Harris, T. B.-L. M., Shepherd-Barron, N., and Taylor, R. (2011). A family of Calabi-Yau varieties and potential automorphy ii. *P.R.I.M.S*, 47:29–98.
- Heath-Brown, D. R. and Patterson, S. (1979). The distribution of Kummer sums at prime arguments. *J. Reine Angew. Math*, 310:11–130.
- Herman, P. E. (2013). On Patterson’s conjecture: Sums of quartic exponential sums. <http://arxiv.org/abs/1305.1243>.

- Homma, K. (2008). On the discrepancy of uniformly distributed roots of quadratic congruences. *Journal of Number Theory*, 128(3):500 – 508.
- Hooley, C. (1963). On the number of divisors of quadratic polynomials. *Acta Mathematica*, 110:97–114.
- Illusie, L. (1994). Autour du théorème de monodromie locale. In *Périodes p -adiques*, volume 223 of *Asterisque*, pages 9–57. Société Mathématique de France.
- Johnson-Freyd, T. (2015). Establishing duality in Tannakian categories. MathOverflow. URL:<http://mathoverflow.net/q/199081> (version: 2015-03-04).
- Katz, N. M. (1980). *Sommes Exponentielles*, volume 79 of *Astérisque*. Société Mathématique de France.
- Katz, N. M. (1988). *Gauss Sums, Kloosterman Sums, and Monodromy*, volume 116 of *Annals of Mathematics Studies*. Princeton University Press.
- Katz, N. M. (1990). *Exponential Sums and Differential Equations*, volume 124 of *Annals of Mathematics Studies*. Princeton University Press.
- Katz, N. M. (1995). *Rigid Local Systems*, volume 139 of *Annals of Mathematics Studies*. Princeton University Press.
- Katz, N. M. (2007). G_2 and hypergeometric sheaves. *Finite Fields and their Applications*, 13(2):175–223.
- Katz, N. M. (2009). Another look at the dwork family. In Tschinkel, Y. and Zarhin, Y., editors, *Algebra, Arithmetic, and Geometry*, volume 270 of *Progress in Mathematics*, pages 89–126. Springer.
- Katz, N. M. (2012). *Convolution and Equidistribution: A Sato-Tate Theorem for Finite Field Mellin Transforms*, volume 180 of *Annals of Mathematics Studies*. Princeton University Press.
- Katz, N. M. and Sarnak, P. (1999). *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45 of *Colloquium Publications*. American Mathematical Society.
- Kramer, K. and Tunnel, J. (1982). Elliptic curves and local ϵ -factors. *Compositio Mathematica*, 46:307–352.
- Krämer, T. and Weissauer, R. (2013). On the Tannaka group attached to the theta divisor of a generic principally polarized abelian variety. <http://arxiv.org/abs/1309.3754>.
- Kummer, E. (1846). De residuis cubicis disquisitiones nonnullae analyticae. *J. reine. angew. Math.*, 32:341–359.

- Laumon, G. (1981). Semi-continuité du conducteur de swan (d'après p. Deligne). In *Caractéristique d'Euler-Poincaré*, volume 82-83 of *Asterisque*, pages 173–219. Société Mathématique de France.
- Laumon, G. (1987). Transformation de fourier, constantes d'Équations fonctionnelles et conjecture de Weil. *Publications mathématiques de l'I.H.É.S.*, 65:131–210.
- Livné, R. and Patterson, S. (2002). The first moment of cubic exponential sums. *Inventiones mathematicae*, 148:79–116.
- Mead, D. G. (1992). Newton's identities. *The American Mathematical Monthly*, 99(8):749–751.
- Mezzadri, F. (2007). How to generate random matrices from the classical compact groups. *Notices of the American Mathematical Society*, 54(5):592–604.
- Milne, J. S. (1980). *Étale Cohomology*, volume 33 of *Princeton Mathematical Series*. Princeton University Press, Princeton, N.J.
- Patterson, S. (1978). On the distribution of Kummer sums. *Journal für die reine und angewandte Mathematik*, 1978:126–143.
- Patterson, S. (2003). The asymptotic distribution of exponential sums, I. *Experimental Mathematics*, 12:129–255.
- Patterson, S. (2005). The asymptotic distribution of exponential sums, II. *Experimental Mathematics*, 14:87–98.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1988). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press.
- Serre, J.-P. (1968). *Abelian ℓ -adic representations*. Number 7 in Research Notes in Mathematics. CRC Press.
- Serre, J.-P. (2011). *Lectures on $N_X(p)$* . Research Notes in Mathematics. A K Peters/CRC Press.
- Serre, J.-P. (2012). Minerva lectures 2012 - J.P. Serre talk 1: Equidistribution. <https://www.youtube.com/watch?v=Rxl3BemTjfk#t=3049>.
- Stickelberger, L. (1890). Über eine verallgemeinerung der kreistheilung. *Mathematische Annalen*, 37:321–367.
- Tóth, A. (2000). Roots of quadratic congruences. *International Mathematics Research Notices*, 14:719–739.
- Xi, P. (2015). Sign changes of Kloosterman sums with almost prime moduli. *Monatshefte für Mathematik*, 77:141–163.