COUNTEREXAMPLES RELATED TO THE SATO–TATE CONJECTURE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Daniel Miller

May 2017

COUNTEREXAMPLES RELATED TO THE SATO–TATE CONJECTURE

Daniel Miller, Ph.D.

Cornell University 2017

Let $E_{/\mathbf{Q}}$ be an elliptic curve. The Sato–Tate conjecture (now a theorem) tells us that the angles $\theta_p = \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$ are equidistributed in $[0,\pi]$ with respect to the measure $\frac{2}{\pi}\sin^2\theta$ if E is non-CM (resp. $\frac{1}{2\pi}\mathrm{d}\theta + \frac{1}{2}\delta_{\pi/2}$ if E is CM). Call μ the measure in question. Akiyama and Tanigawa conjecture that the discrepancy

$$D_N = \sup_{x\in[0,\pi]}\left|\frac{1}{\pi(N)}\sum_{p\leqslant N} 1_{[0,x]}(\theta_p) - \int 1_{[0,x]}\,\mathrm{d}\mu\right|$$

asymptotically decays like $N^{-\frac{1}{2}+\epsilon}$, as is suggested by computational evidence and certain reasonable heuristics on the Kolmogorov–Smirnov statistic. This conjecture implies the Riemann Hypothesis for all L-functions associated with E. It is natural to assume that the converse ("Riemann Hypothesis implies discrepancy estimate") holds, as is suggested by analogy with Artin L-functions. We show that when E has CM, there is no reason to believe that the converse holds, as there are "fake Satake parameters" yielding L-functions which satisfy the Generalized Riemann Hypothesis, but for which the discrepancy decays like $N^{-\epsilon}$.

We also show that there are Galois representations $\rho\colon \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{Z}_l)$, ramified at an arbitrarily thin set of primes, whose Satake parameters can be made to converge at any specified rate to any fixed measure μ on $[0,\pi]$ for which $\cos_*\mu$ is absolutely continuous.

BIOGRAPHICAL SKETCH

Daniel Miller was born in St. Paul, Minnesota. He completed his Bachelor of Science at the University of Nebraska–Omaha, during which he played the piano competitively and attended Cornell's Summer Mathematics Institute. He started his Ph.D. at Cornell planning on a career in academia. Halfway through he had a change of heart, and will be joining Microsoft's Analysis and Experimentation team as a data scientist after graduation. He is happily married to Ivy Lai Miller, and has a cute but grumpy cat named Socrates.

This thesis is dedicated to my undergraduate adviser, Griff Elder. He is the reason I ever considered a career in mathematics, and his infectious enthusiasm for number theory has inspired me more than I can say.

ACKNOWLEDGEMENTS

I'll like to begin by thanking my parents Jay and Cindy for noticing and fostering my mathematical interests early on, and for being unfailingly loving and supportive. I'd also like to thank my undergraduate thesis advisor, Griffith Elder, without whose encouragement and inspiration I probably never would have considered a career in math.

I'd like to thank Tara Holm for organizing Cornell's Summer Mathematics Institute in 2011, Jason Boynton for teaching a fantastic algebra class, and Anthony Weston for introducing me to the world of nonlinear functional analysis.

Thanks go to fellow graduate students Sasha Patotski, Balázs Elek, and Sergio Da Silva for sharing my early love of algebraic geometry, laughing with me at the absurdities of academic life, and listening to my ramblings about number theory.

I owe a big debt of gratitude to the mathematics department at Cornell, where many professors were generous with their time and ideas. I appreciate Yuri Berest, John Hubbard, Farbod Shokrieh, Birget Speh, and David Zywina for letting me bounce ideas off them, helping me add rigor to half-baked ideas, and pointing me in new and interesting directions of research.

I am especially thankful to my adviser Ravi Ramakrishna. He kindled my first love for number theory, stayed supportive as my research bounced all over the place, and kept me focused, grounded, and concrete when I needed to be.

Most importantly, I thank my loving wife Ivy for being there for me through the highs and the lows, both when I (prematurely) thought my thesis was complete, and when I thought my results were completely in shambles. I couldn't have done it without her.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Motivation from classical analytic number theory

Start with an old problem central to number theory—counting prime numbers. As usual, let $\pi(x)$ be the prime counting function and $\mathrm{Li}(x) = \int_2^x \frac{\mathrm{d}t}{\log t}$ be the Eulerian logarithmic integral. There is a normalized empirical measure $P_x = \frac{1}{\pi(x)} \sum_{p \leqslant x} \delta_{p/x}$, capturing the distribution of the set of primes $\leqslant x$. The prime number theorem tells us that as $x \to \infty$, this empirical measure converges weakly to the "true" measure $L_x = \frac{\mathrm{Li}(tx)}{\mathrm{Li}(x)}\,\mathrm{d}t$. The standard approach to proving the prime number theorem is by showing that the Riemann ζ-function has non-vanishing meromorphic continuation past $\Re = 1$.

Theorem 1.1.1. *The function $\zeta(s)$ admits a non-vanishing meromorphic continuation past $\Re = 1$ with at most a simple pole at $s = 1$, if and only if $P_x \to L_x$ weakly.*

Since $\zeta(s)$ does have the desired properties, the prime number theorem holds. It is natural to try to quantify the rate of converge of P_x to L_x. One way to do this is via the (star) discrepancy

$$\mathrm{D}^\star(P_x, L_x) = \sup_{t \in [0,1]} |P_x[0,t] - L_x[0,t]| = \sup_{t \in [0,1]} \left| \frac{\pi(tx)}{\pi(x)} - \frac{\int_2^{tx} \frac{\mathrm{d}s}{\log s}}{\int_2^x \frac{\mathrm{d}s}{\log s}} \right|.$$

Numerical experiments suggest that $\mathrm{D}^\star(P_x, L_x) \ll x^{-\frac{1}{2}+\epsilon}$, and in fact we have the following result.

Theorem 1.1.2. *The Riemann Hypothesis is true if and only if $\mathrm{D}^\star(P_x, L_x) \ll x^{-\frac{1}{2}+\epsilon}$.*

Neither side of this equivalence is known for certain to be true!

The above discussion finds a natural generalization in Artin L-functions. Let $K/\mathbf{Q}$ be a finite Galois extension with group $G = \mathrm{Gal}(K/\mathbf{Q})$. For any rational prime p at which K is unramified, let fr_p be the conjugacy class of the Frobenius at p in G. For any irreducible representation $\rho\colon G \to \mathrm{GL}_n(\mathbf{C})$, there is a corresponding L-function defined as

$$L(\rho, s) = \prod_p \frac{1}{\det(1 - \rho(\mathrm{fr}_p)p^{-s})},$$

where here (and for the remainder of this thesis) we tacitly omit from the product those primes at which ρ is ramified. Given a cutoff x, there is a natural empirical measure $P_x = \frac{1}{\pi(x)} \sum_{p \leqslant x} \delta_{\mathrm{fr}_p}$ on $G^\natural$, the set of conjugacy classes in G. Let $\mathrm{D}(P_x) = \sup_{S \subset G^\natural} \left| P_x(S) - \frac{\#S}{\#G^\natural} \right|$. Then P_x converges weakly to the uniform measure on $G^\natural$ if and only if $\mathrm{D}(P_x) \to 0$.

Theorem 1.1.3. *The measure P_x converge weakly to the uniform measure on $G^\natural$ if and only if the function $L(\rho, s)$ admits a nonvanishing analytic continuation past $\Re = 1$ for all nontrivial ρ.*

Both sides of this equivalence are true, and known as the Chebotarev density theorem. Moreover, there is a version of the strong prime number theorem in this context.

Theorem 1.1.4. *The bound $\mathrm{D}(P_x) \ll x^{-\frac{1}{2}+\epsilon}$ holds if and only if each $L(\rho, s)$, ρ nontrivial, satisfies the Riemann Hypothesis.*

This whole discussion generalizes to a more complicated set of Galois representations—those arising from elliptic curves and more general motives.

1.2 Discrepancy and Riemann Hypothesis for elliptic curves

Let $E_{/\mathbf{Q}}$ be an elliptic curve. For any prime l, the l-adic Tate module of E induces a continuous representation $\rho_l \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$. It is known that the quantities $a_p = \operatorname{tr} \rho_l(\mathrm{fr}_p)$ lie in $\mathbf{Z}$ and satisfy the Hasse bound $|a_p| \leqslant 2\sqrt{p}$. For each unramified prime p, the corresponding Satake parameter for E is $\theta_p = \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right) \in [0, \pi]$. These parameters are packaged into an L-function as follows:

$$L(E,s) = \prod_p \frac{1}{(1-e^{i\theta_p}p^{-s})(1-e^{-i\theta_p}p^{-s})} = \prod_p \frac{1}{\det\left(1 - \begin{pmatrix} e^{i\theta_p} & \\ & e^{-i\theta_p} \end{pmatrix} p^{-s}\right)}.$$

More generally we have, for each irreducible representation of $\mathrm{SU}(2)$, which will be sym^k for some $k \geqslant 1$, the k-th symmetric power L-function:

$$L(\mathrm{sym}^k E, s) = \prod_p \prod_{j=0}^{k} \frac{1}{1 - e^{i(k-2j)\theta_p}p^{-s}} = \prod_p \frac{1}{\det\left(1 - \mathrm{sym}^k \begin{pmatrix} e^{i\theta_p} & \\ & e^{-i\theta_p} \end{pmatrix} p^{-s}\right)}.$$

Numerical experiments suggest that the Satake parameters are equidistributed with respect to the Sato–Tate distribution $\mathrm{ST} = \frac{2}{\pi}\sin^2\theta\,\mathrm{d}\theta$. Indeed, for any cutoff x, let P_x be the empirical measure $P_x = \frac{1}{\pi(x)}\sum_{p \leqslant x} \delta_{\theta_p}$. The convergence of the P_x to the Sato–Tate measure is closely related to the analytic properties of the $L(\mathrm{sym}^k E, s)$. First, here is the famous Sato–Tate Conjecture (now a theorem) in our notation.

Theorem 1.2.1 (Taylor et. al.)**.** *If E is non-CM, the measures P_x converge weakly to* ST*.*

Theorem 1.2.2 (Serre)**.** *Let Sato–Tate conjecture holds for (a non-CM) E if and only if each of the functions $L(\mathrm{sym}^k E, s)$ have analytic continuation past $\Re = 1$.*

The stunning recent proof of the Sato–Tate conjecture [CHT08; Tay08; HSBT10] in fact showed that the functions $L(\mathrm{sym}^k E, s)$ were potentially automorphic, which gives the desired analytic continuation.

The Riemann Hypothesis, and its analogue for Artin L-functions, has a natural generalization to elliptic curves. In this context, the discrepancy of the set $\{\theta_p\}_{p\leqslant x}$ is

$$\mathrm{D}\left(\{\theta_p\}_{p\leqslant x}, \mathrm{ST}\right) = \sup_{t\in[0,\pi]} \left|P_x[0,t] - \mathrm{ST}[0,t]\right|.$$

The following conjecture is first made in [AT99]: for $E_{/\mathbf{Q}}$ a non-CM elliptic curve, the estimate $\mathrm{D}\left(\{\theta_p\}_{p\leqslant x}, \mathrm{ST}\right) \ll x^{-\frac{1}{2}+\epsilon}$ holds. The authors go on to prove a special case of the following theorem, proved in full generality in [Maz08].

Theorem 1.2.3 (Mazur)**.** *If* $\mathrm{D}\left(\{\theta_p\}_{p\leqslant x}, \mathrm{ST}\right) \ll x^{-\frac{1}{2}+\epsilon}$*, then all the functions* $L(\mathrm{sym}^k E, s)$ *satisfy the Riemann Hypothesis.*

This discussion also makes sense when E has complex multiplication (for simplicity, we consider $E_{/F}$ where F is the field of definition of the complex multiplication). The Sato–Tate measure for such E is the Haar measure on SO(2), i.e. the uniform measure on $[0,\pi]$. Instead of symmetric power L-functions, there is an L-function for each character of SO(2). Once again, "Akiyama–Tanigawa implies Riemann Hypothesis" holds.

It is natural to assume that the converse to the implication "Akiyama–Tanigawa implies General Riemann Hypothesis" holds. David Zywina first suggested to the author that it might not. In this thesis, we construct a range of counterexamples to the implication "Strong Sato–Tate implies Riemann Hypothesis" for the case of CM abelian varieties Moreover, we generalize the results of [Pan11] to show that

there can be no purely Galois-theoretic proof of the Sato–Tate conjecture, for the are Galois representations with arbitrary Sato–Tate distributions!

1.3 Notation conventions

Whenever l is mentioned it is a rational prime $\geqslant 7$.

Write $f \ll g$ if $f = O(g)$, i.e. there is a constant $C > 0$ such that $f \leqslant Cg$.

Write $f = \Omega(g)$ (in the convention of Hardy–Littlewood) if $\limsup \frac{f}{g} > 0$.

The symbol $f = \Theta(g)$ means there exist constants $0 < C_1 < C_2$ such that $C_1 g \leqslant f \leqslant C_2 g$. Equivalently, $g \ll f$ and $f \ll g$.

If μ is a measure on $\mathbf{R}$, then write $\mu[a, b]$ for $\mu([a, b])$, and similarly for $[a, b)$, $(a, b]$, etc. In general, whenever it simplifies the notation, we will write μS for $\mu(S)$ if μ is a measure and S is a measurable set.

If μ is a measure on $\mathbf{R}$, then *cumulative distribution function (cdf) of* μ is given by $\mathrm{cdf}_\mu(x) = \mu[-\infty, x]$.

If $z \in \mathbf{C}$, write $\Re z$ for the real part of z.

If $\alpha \in \mathbf{R}$, we write $\Re > \alpha$ for the half-plane of complex numbers with real part $> \alpha$. So a function has analytic continuation to the half-plane $\{z \in \mathbf{C} : \Re z > \alpha\}$ if and only if the function extends to $\Re > \alpha$.

We write $\boldsymbol{x} = (x_1, x_2, \dots)$ for infinite sequences and $\vec{x} = (x_1, \dots, x_d)$ for vectors. Sometimes we will have a sequence of vectors, written as $\boldsymbol{x} = (\vec{x}_1, \vec{x}_2, \dots)$.

If $\boldsymbol{x} = (x_1, x_2, \dots)$ is a sequence, write $P_{\boldsymbol{x},N} = \frac{1}{N}\sum_{n \leqslant N} \delta_{x_n}$ for the corresponding empirical measure. If $\boldsymbol{x} = (x_\alpha)$ is instead indexed by some other (discrete) subset of $\mathbf{R}^+$, write

$$P_{\boldsymbol{x},N} = \frac{1}{\#\{\text{indices } \leqslant N\}} \sum_{\alpha \leqslant N} \delta_{x_\alpha}.$$

Omitted entries in matrices are zero, i.e. $\left(\begin{smallmatrix} a & \\ & b \end{smallmatrix}\right)$ means $\left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)$.

CHAPTER 2

DISCREPANCY

2.1 Equidistribution

Discrepancy (also known as the Kolmogorov–Smirnov statistic) is a way of measuring how closely sample data fits a predicted distribution. It has many applications in computer science and statistics, but here we will focus on only the basic properties, such as how discrepancy changes when sequences are "tweaked" and combined.

First, recall that discrepancy provides a way of sharpening the soft convergence results of [Ser89, A.1]. Let X be a compact topological space, $\boldsymbol{x} = (x_2, x_3, x_5, \dots)$ a sequence of points in X indexed by the rational primes.

Definition 2.1.1. *Let μ be a continuous probability measure on X. The sequence $\boldsymbol{x}$ is* equidistributed *with respect to μ if for all $f \in C(X)$, we have*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\leqslant N} f(x_p) = \int f \,\mathrm{d}\mu.$$

In other words, $\boldsymbol{x}$ is μ-equidistributed if the empirical measures $P_{\boldsymbol{x},N} = \frac{1}{\pi(N)} \sum_{p\leqslant N} \delta_{x_p}$ converge to μ in the weak topology. It is easy to see that $\boldsymbol{x}$ is μ-equidistributed if and only if $\left|\sum_{p\leqslant N} f(x_p)\right| = o(\pi(N))$ for all f having $\int f \,\mathrm{d}\mu = 0$. In fact, one can restrict to a set of f which generate a dense subpace of $C(X)^{\mu=0}$.

In the discussion in [Ser89, A.1], X is the space of conjugacy classes in a compact Lie group, and f is allowed to range over the characters of irreducible,

nontrivial, unitary representations of the group. Serre's results can be generalized to a much broader class of Dirichlet series, which are of the form

$$L_f(\boldsymbol{x}, s) = \prod_p \frac{1}{1 - f(x_p)p^{-s}}.$$

In fact, in light of the following theorem, we can consider functions f which are only only continuous almost everywhere. This allows us to consider step functions like $-1_{[0,\pi/2)} + 1_{(\pi/2,\pi]}$ on $[0, \pi]$.

Theorem 2.1.2. *Let X be a compact separable metric space with no isolated points. Let μ be a Borel measure on X and let $f\colon X \to \mathbf{C}$ be bounded and measurable. Then f is continuous almost everywhere if and only if*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\leqslant N} f(x_p) = \int f \,\mathrm{d}\mu$$

for all μ-equidistributed sequences $\boldsymbol{x}$.

Proof. This follows immediately from the proof of [Maz95, Th. 1]. □

2.2 Definitions and first results

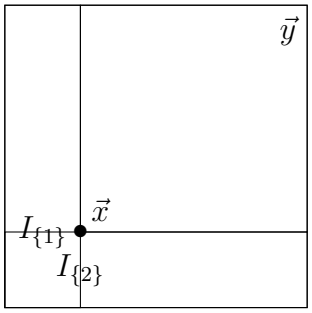
We will define discrepancy for measures on the d-dimensional half-open box $[0, \vec{\infty}) = [0, \infty)^d \subset \mathbf{R}^d$. For vectors $\vec{x}, \vec{y} \in [0, \vec{\infty})$, we say $\vec{x} < \vec{y}$ if $x_i < y_i \,\forall i$, and in that case write $[\vec{x}, \vec{y})$ for the half-open box $[x_1, y_1) \times \cdots \times [x_d, y_d)$.

Definition 2.2.1. *Let μ, ν be probability measures on $[0, \vec{\infty})$. The* discrepancy *of μ with respect to ν is $\mathrm{D}(\mu, \nu) = \sup_{\vec{x}<\vec{y}} |\mu[\vec{x}, \vec{y}) - \nu[\vec{x}, \vec{y})|$, where $\vec{x} < \vec{y}$ range over $[0, \vec{\infty})$. The* star discrepancy *of μ with respect to ν is $\mathrm{D}^\star(\mu, \nu) = \sup_{0\leqslant\vec{y}} |\mu[0, \vec{y}) - \nu[0, \vec{y})|$, where $\vec{y}$ ranges over $[0, \vec{\infty})$.*

Lemma 2.2.2. *Let μ, ν be Borel measures on $\mathbf{R}^d$. Then $\mathrm{D}^\star(\mu,\nu) \leqslant \mathrm{D}(\mu,\nu) \leqslant 2^d\,\mathrm{D}^\star(\mu,\nu)$.*

Proof. The first inequality holds because the supremum defining discrepancy is taken over a larger set than that defining star discrepancy. To prove the second inequality, let $\vec{x} < \vec{y}$ be in $[0,\vec{\infty})$. For $S \subset \{1,\dots,d\}$, let $I_S = \{\vec{t} \in [0,\vec{y}) : t_i < x_i \,\forall\, i \in S\}$. Inclusion-exclusion tells us that $\mu[\vec{x},\vec{y}) = \sum_{S\subset\{1,\dots,d\}} (-1)^{\#S} \mu(I_S)$, and

Figure 2.1: The sets $I_{\{1\}}$ and $I_{\{2\}}$ when $d = 2$. $I_\varnothing$ is the large square and $I_{\{1,2\}}$ is the small square.



similarly for ν. Since each of the I_S are half-open boxes intersecting the origin, we know that $|\mu(I_S) - \nu(I_S)| \leqslant \mathrm{D}^\star(\mu,\nu)$. It follows that

$$|\mu[\vec{x},\vec{y}) - \nu[\vec{x},\vec{y})| \leqslant \sum_{S\subset\{1,\dots,d\}} |\mu(I_S) - \nu(I_S)| \leqslant 2^d\,\mathrm{D}^\star(\mu,\nu).$$

For a discussion and related context, see [KN74, Ch. 2 Ex. 1.2]. □

Since we are only interested in the asymptotics of discrepancy, we will sometimes gloss over the distinction between discrepancy and star discrepancy, using whichever makes a proof easier to follow.

We are usually interested in comparing empirical measures and their conjectured asymptotic distribution. Let $\boldsymbol{x} = (\vec{x}_1, \vec{x}_2, \dots)$ be a sequence in $[0,\vec{\infty})$, and μ

a probability measure on $[0, \vec{\infty})$. For any real number $N \geqslant 1$, the empirical measure associated to the truncated sequence $\boldsymbol{x}_{\leqslant N} = (\vec{x}_n)_{n \leqslant N}$ is $P_{\boldsymbol{x},N} = \frac{1}{N} \sum_{n \leqslant N} \delta_{\vec{x}_n}$. Write $\mathrm{D}_N(\boldsymbol{x}, \mu) = \mathrm{D}_N(P_{\boldsymbol{x},N}, \mu)$, and likewise for the star discrepancy. In this context,

$$\mathrm{D}_N^\star(\boldsymbol{x}, \mu) = \sup_{\vec{y} \in [0, \vec{\infty})} \left| \frac{\#\{n \leqslant N : \vec{x}_n \in [0, \vec{y})\}}{N} - \int_{[0, \vec{y})} \mathrm{d}\mu \right|.$$

If the measure μ is only defined on a Borel subset of $[0, \vec{\infty})^d$, we tacitly extend it by zero to $\mathbf{R}^d$. If the sequence $\boldsymbol{x}$ lies in a torus $(\mathbf{R}/a\mathbf{Z})^d$, we identify that torus with $[0, a)^d \subset [0, \infty)^d$. If λ is normalized Haar measure on the torus, we write $\mathrm{D}_N(\boldsymbol{x})$ in place of $\mathrm{D}_N(\boldsymbol{x}, \lambda)$.

If the sequence $\boldsymbol{x}$ lies in the space $G^\natural$ of conjugacy classes in a compact Lie group G, choose a maximal torus $T \subset G$, and recall that $G^\natural = T/W$, where W is the Weyl group of T. There is a half-open box in $\mathfrak{t} = \mathrm{Lie}(T)$ which maps bijectively to T under the exponential map. Choose a "half-open" polyhedral set Q that maps bijectively to T/W. Then $Q \subset \mathfrak{t}$ and, if we choose a basis for $\mathfrak{t}$ mapping to zero in T, then it makes sense to talk about the discrepancy of a sequence in $G^\natural$ with respect to the Haar measure. The paper [Ros13] has a different definition of discrepancy which only works for semisimple simply-connected groups, but also proves an Erdös–Turán inequality in that context. It is likely that a reasonable application of isotropic discrepancy would render these definitions equivalent, at least for asymptotic purposes, but as the two definitions coincide for SU(2), we do not explore this further.

Sometimes the sequence $\boldsymbol{x}$ will not be indexed by the natural numbers, but by the rational primes, or some other discrete subset of $\mathbf{R}^+$. In that case we will

still use the notations $\mathrm{D}_N(\boldsymbol{x}, \mu)$, $\boldsymbol{x}_{\leqslant N} = (\vec{x}_1, \ldots, \vec{x}_N)$, $\boldsymbol{x}_{\geqslant N} = (\vec{x}_N, \vec{x}_{N+1}, \ldots)$, etc., keeping in mind that the set $\{\vec{x}_\alpha : \alpha \leqslant N\}$ is involved, and that in formulas $\frac{1}{N}$ is replaced by $\#\{\text{indices } \leqslant N\}^{-1}$.

Why half-open boxes? The choice of sets of the form $[\vec{x}, \vec{y})$ in the definition of discrepancy seems rather arbitrary. It is. Discrepancy can easily be defined as a supremum over all open (or closed) balls—and in fact those definitions generalize to arbitrary metric spaces. There are also more subtle definitions involving suprema over open or closed convex sets (isotropic discrepancy). See [KN74] for a discussion and comparison of these differing definitions. In this thesis, we restrict to half-open boxes because they are computationally tractable, fit well with Diophantine approximation on tori, and the theory is most developed for this definition.

2.3 Statistical heuristics

Replace the Satake parameters θ_p of an elliptic curve with a sequence $\{\theta_p\}$ of iid random variables with continuous common distribution μ, which we may take to be $\frac{2}{\pi}\sin^2\theta\,\mathrm{d}\theta$, supported on $[0, \pi]$. Then the discrepancy (known as the Kolmogorov–Smirnov statistic in this context) is the function-valued random variable

$$D_N = \sup_{x \in [0,\pi]} \left| \frac{1}{\pi(N)} \sum_{p \leqslant N} 1_{[0,x]} \circ \theta_p - \int 1_{[0,x]}\,\mathrm{d}\mu \right|.$$

Kolmogorov and Smirnov proved that the inside of the absolute value converges to zero. The Glivenko–Cantelli Theorem says that $D_N \to 0$ almost surely, and even better, the normalized discrepancy $\sqrt{\pi(N)}D_N$ approaches a limiting distribution K (supremum of the Brownian Bridge) which does not depend on μ. The rate of convergence of $\sqrt{\pi(N)}D_N$ to that distribution is quantified by the Dvoretzky–

Kiefer–Wolfowitz inequality, which tells us that $\mathrm{P}\left(\sqrt{\pi(N)}D_N > t\right) \leqslant 2e^{-2t^2}$. These theorems suggest that for $E_{/\mathbf{Q}}$ a non-CM elliptic curve, the "true" discrepancy $\mathrm{D}_N(\boldsymbol{\theta}, \mathrm{ST})$ should decay like $\pi(N)^{-\frac{1}{2}}$, or at least $N^{-\frac{1}{2}+\epsilon}$. Similarly for E a CM elliptic curve and the Sato–Tate measure. Ideally, the normalized discrepancy $\sqrt{\pi(N)}\,\mathrm{D}_N^\star(\boldsymbol{\theta}, \mathrm{ST})$ would also be equidistributed, but sadly, numerical experiments suggest this is not the case.

2.4 The Koksma–Hlawka inequality

Here we summarize the results of the paper [Ö99], generalizing them as needed for our context. Recall that a function f on $[0, \vec{\infty}) \subset \mathbf{R}^d$ is said to be of *bounded variation* (in the measure-theoretic sense) if there is a finite Radon measure ν such that $f(\vec{x}) - f(0) = \nu[0, \vec{x}]$. In such a case we write $\mathrm{Var}(f) = |\nu|$. If the appropriate differentiability conditions are satisfied, then $\mathrm{Var}(f) = \int_{[0,\vec{\infty})} \left|\frac{\mathrm{d}^d f}{\mathrm{d}t_1 \ldots \mathrm{d}t_d}\right|$.

Theorem 2.4.1 (Koksma–Hlawka)**.** *Let μ be a probability measure on $[0, \vec{\infty})$, f of bounded variation. For any sequence $\boldsymbol{x} = (\vec{x}_1, \vec{x}_2, \dots)$ in $[0, \vec{\infty})$, we have*

$$\left|\frac{1}{N}\sum_{n \leqslant N} f(\vec{x}_n) - \int f \,\mathrm{d}\mu\right| \leqslant \mathrm{Var}(f)\,\mathrm{D}_N(\boldsymbol{x}, \mu).$$

Proof. By assumption, there is a finite Radon measure ν such that $f(\vec{y}) - f(0) = \nu[0, \vec{y}]$. What follows relies on the fact that $1_{[0,\vec{x}]}(\vec{y}) = 1_{[\vec{y},\vec{\infty})}(\vec{x})$.

$$\begin{aligned}
\frac{1}{N}\sum_{n \leqslant N} f(\vec{x}_n) - \int f \,\mathrm{d}\mu &= \frac{1}{N}\sum_{n \leqslant N} \left(f(\vec{x}_n) - f(0)\right) - \int \left(f(\vec{y}) - f(0)\right) \mathrm{d}\mu(\vec{y}) \\
&= \frac{1}{N}\sum_{n \leqslant N} \int 1_{[\vec{y},\vec{\infty})}(\vec{x}_n) \,\mathrm{d}\nu(\vec{y}) - \int\int 1_{[0,\vec{y}]} \,\mathrm{d}\nu \,\mathrm{d}\mu(\vec{y}) \\
&= \int \left(\frac{1}{N}\sum_{n \leqslant N} 1_{[\vec{y},\vec{\infty})}(\vec{x}_n) - \int 1_{[\vec{y},\vec{\infty})} \,\mathrm{d}\mu\right) \mathrm{d}\nu(\vec{y}).
\end{aligned}$$

It follows that

$$\left|\frac{1}{N}\sum_{n\leqslant N} f(\vec{x}_n) - \int f \,\mathrm{d}\mu\right| \leqslant \sup_{\vec{y}\in[0,\vec{\infty})} \left|\frac{1}{N}\sum_{n\leqslant N} 1_{[\vec{y},\vec{\infty})}(\vec{x}_n) - \int 1_{[\vec{y},\vec{\infty})}\,\mathrm{d}\mu\right| \cdot |\nu|.$$

The supremum is bounded above by $\mathrm{D}_N(\boldsymbol{x},\mu)$, so the proof is complete. □

This theorem is proved in a somewhat restrictive setting, and can be generalized. For f a function on $\mathbf{R}^+$ that is bounded variation in the traditional sense (for example, piecewise continuous) and μ a continuous probability measure, the inequality

$$\left|\frac{1}{N}\sum_{n\leqslant N} f(x_n) - \int f \,\mathrm{d}\mu\right| \leqslant \mathrm{Var}(f)\,\mathrm{D}_N^\star(\boldsymbol{x},\mu)$$

holds [KN74, Ch. 2, Th. 5.1]. In particular, when μ is the Sato–Tate measure and f is piecewise continuous, we can apply this inequality.

2.5 Comparing and combining sequences

Throughout this section, λ is the Lebesgue measure on $\mathbf{R}$.

Lemma 2.5.1. *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be sequences in $[0,\infty)$. Suppose μ is an absolutely continuous probability measure on $[0,\infty)$ with bounded Radon–Nikodym derivative $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}$. Let $\epsilon > 0$ be arbitrary. Then*

$$|\mathrm{D}_N^\star(\boldsymbol{x},\nu) - \mathrm{D}_N^\star(\boldsymbol{y},\nu)| \leqslant \left\|\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}\right\|_\infty \epsilon + \frac{\#\{n\leqslant N : |x_n - y_n| \geqslant \epsilon\}}{N}.$$

Proof. Fix $\epsilon > 0$, and let $t \in [0,\infty)$ be arbitrary. For all $n \leqslant N$ such that $y_n < t$, either $x_n < t+\epsilon$ or $|x_n - y_n| \geqslant \epsilon$. It follows that

$$P_{\boldsymbol{y},N}[0,t) \leqslant P_{\boldsymbol{x},N}[0,t+\epsilon) + \frac{\#\{n\leqslant N : |x_n - y_n| \geqslant \epsilon\}}{N}.$$

Moreover, we have $|P_{\boldsymbol{x},N}[0,t+\epsilon) - \nu[0,t+\epsilon)| \leqslant \mathrm{D}_N^\star(\boldsymbol{x},\nu)$. Putting these together, we get:

$$\begin{aligned} P_{\boldsymbol{y},N}[0,t) - \nu[0,t) &\leqslant P_{\boldsymbol{x},N}[0,t+\epsilon) - \nu[0,t) + \frac{\#\{n \leqslant N : |x_n - y_n| \geqslant \epsilon\}}{N} \\ &\leqslant \nu[t,t+\epsilon) + \mathrm{D}_N^\star(\boldsymbol{x},\nu) + \frac{\#\{n \leqslant N : |x_n - y_n| \geqslant \epsilon\}}{N} \\ &\leqslant \left\| \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \right\|_\infty \epsilon + \mathrm{D}_N^\star(\boldsymbol{x},\nu) + \frac{\#\{n \leqslant N : |x_n - y_n| \geqslant \epsilon\}}{N} \end{aligned}$$

This tells us that $\mathrm{D}_N^\star(\boldsymbol{y},\nu) \leqslant \left\| \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \right\|_\infty \epsilon + \mathrm{D}_N^\star(\boldsymbol{x},\nu) + \frac{\#\{n \leqslant N : |x_n - y_n| \geqslant \epsilon\}}{N}$. Reversing the roles of $\boldsymbol{x}$ and $\boldsymbol{y}$, we obtain the desired result. □

Lemma 2.5.2. *Let σ be an isometry of $\mathbf{R}$, and $\boldsymbol{x}$ a sequence in $[0,\infty)$ such that $\sigma(\boldsymbol{x})$ is also in $[0,\infty)$. Let ν be an absolutely continuous measure on $[0,\infty)$ such that $\sigma_*\nu$ is contained in $[0,\infty)$. Then $|\mathrm{D}_N(\boldsymbol{x},\nu) - \mathrm{D}_N(\sigma_*\boldsymbol{x},\sigma_*\nu)| \leqslant \frac{2}{N}$.*

Proof. Every isometry of $\mathbf{R}$ is a combination of translations and reflections. The statement is clear with translations (the two discrepancies are equal). So, suppose $\sigma(t) = a - t$ for some $a > 0$. Since ν is absolutely continuous, $\nu\{t\} = 0$ for all $t \geqslant 0$. In particular, $\nu[s,t) = \nu(s,t]$. By definition, $P_{\boldsymbol{x},N}\{t\} \leqslant N^{-1}$. For any interval $[s,t)$ in $[0,\infty)$, we know that $|P_{\boldsymbol{x},N}[s,t) - P_{\boldsymbol{x},N}(s,t]| \leqslant \frac{2}{N}$, hence

$$|P_{\boldsymbol{x},N}[s,t) - \nu[s,t) - P_{\sigma_*\boldsymbol{x},N}[a-t,a-s) - \sigma_*\nu[a-t,a-s)| \leqslant \frac{2}{N}.$$

This proves the result. □

A technique we will use throughout this thesis involves comparing the discrepancy of a sequence with the discrepancy of a pushforward sequence, with respect to the pushforward measure.

Lemma 2.5.3. *Let I, J be closed connected intervals and $f\colon I \twoheadrightarrow J$ a continuous monotonic map. If $\boldsymbol{x}$ is a sequence in I and μ is an absolutely continuous probability measure on I, then $|\mathrm{D}_N(\boldsymbol{x}, \mu) - \mathrm{D}_N(f_*\boldsymbol{x}, f_*\mu)| \leqslant \frac{4}{N}$.*

Proof. Because f is continuous and monotonic, given $[u, v) \subset I$, there exists $[x, y) \subset J$ such that $f[u, v)$ differs from $[x, y)$ by at most two elements. Similarly, if $[x, y) \subset J$, there exists $[u, v) \subset I$ such that $f^{-1}[x, y)$ differs from $[u, v)$ by at most two elements. Writing $P_{\boldsymbol{x},N}$ as usual for the empirical measure, we compute (in the first case):

$$|P_{f_*\boldsymbol{x},N}[x, y) - f_*\mu[x, y) - (P_{\boldsymbol{x},N}[u, v) - \mu[u, v))| \leqslant \frac{4}{N},$$

equality $f_*\mu[u, v) = \mu[x, y)$ following from the continuity of μ. It follows that $\mathrm{D}_N(\boldsymbol{x}, \mu) \leqslant \mathrm{D}_N(f_*\boldsymbol{x}, f_*\mu) + \frac{4}{N}$. Reversing the roles of $\boldsymbol{x}$ and $f_*\boldsymbol{x}$, the result follows. □

Now we show that the discrepancy behaves as expected when two sequences are interleaved.

Definition 2.5.4. *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be sequences in $[0, \vec{\infty}) \subset \mathbf{R}^d$. We write $\boldsymbol{x} \wr \boldsymbol{y}$ for the interleaved sequence $(x_1, y_1, x_2, y_2, x_3, y_3, \dots)$.*

Write $P_{\boldsymbol{x}\wr\boldsymbol{y},N} = \frac{1}{2}(P_{\boldsymbol{x},N} + P_{\boldsymbol{y},N})$ for the combined empirical measure of the interleaved sequence $\boldsymbol{x} \wr \boldsymbol{y}$.

Theorem 2.5.5. *Let I and J be disjoint open boxes in $[0, \vec{\infty})$, and let μ, ν be absolutely continuous probability measures on I and J, respectively. Let $\boldsymbol{x}$ be a sequence in I and $\boldsymbol{y}$ be a sequence in J. Then*

$$\max\{\mathrm{D}_N(\boldsymbol{x}, \mu), \mathrm{D}_N(\boldsymbol{y}, \nu)\} \leqslant \mathrm{D}_N(\boldsymbol{x} \wr \boldsymbol{y}, \mu + \nu) \leqslant \mathrm{D}_N(\boldsymbol{x}, \mu) + \mathrm{D}(\boldsymbol{y}, \nu)$$

Proof. Any half-open box in $[0, \vec{\infty})$ can be split by a coordinate hyperplane into two disjoint half-open boxes $[\vec{a}, \vec{b}) \sqcup [\vec{s}, \vec{t})$, each of which intersects at most one of I and J. We may assume that $[\vec{a}, \vec{b}) \cap J = \varnothing$ and $[\vec{s}, \vec{t}) \cap I = \varnothing$. Then

$$\begin{aligned}\left|P_{\boldsymbol{x}\wr\boldsymbol{y},N}([\vec{a}, \vec{b}) \sqcup [\vec{s}, \vec{t})) - (\mu + \nu)([\vec{a}, \vec{b}) \sqcup [\vec{s}, \vec{t}))\right| &\leqslant |P_{\boldsymbol{x},N}[\vec{a}, \vec{b}) - \mu[\vec{a}, \vec{b})| + |P_{\boldsymbol{y},N}[\vec{s}, \vec{t}) - \nu[\vec{s}, \vec{t})| \\ &\leqslant \mathrm{D}_N(\boldsymbol{x}, \mu) + \mathrm{D}_N(\boldsymbol{y}, \nu).\end{aligned}$$

This yields the second inequality in the statement of the theorem. To see the first, assume that the maximum discrepancy is $\mathrm{D}_N(\boldsymbol{x}, \mu)$, and let $[\vec{s}, \vec{t})$ be a half-open box such that $|P_{\boldsymbol{x},N}[\vec{s}, \vec{t}) - \mu[\vec{s}, \vec{t})|$ is within some arbitrary ϵ of $\mathrm{D}_N(\boldsymbol{x}, \mu)$. We can replace $[\vec{s}, \vec{t})$ with a smaller box to ensure it does not intersect J. Assuming $[\vec{s}, \vec{t}) \cap J = \varnothing$, we have $\left|P_{\boldsymbol{x}\wr\boldsymbol{y},N}[\vec{s}, \vec{t}) - (\mu + \nu)[\vec{s}, \vec{t})\right| = |P_{\boldsymbol{x},N}[\vec{s}, \vec{t}) - \mu[\vec{s}, \vec{t})|$, which yields the result. □

2.6 Examples

Historically, one of the first interesting examples of an equidistributed sequence is the set of translates of an irrational number modulo one.

Theorem 2.6.1. *Let $a \in \mathbf{R}$ be irrational. Then the sequence $\boldsymbol{x} = (a \mod 1, 2a \mod 1, 3a \mod 1, \dots)$ is equidistributed in $[0, 1]$.*

We will prove this result (originally due to Weyl) in Chapter 4. It is known, and we will prove, that sequences of this form have discrepancy which decays like $N^{-\alpha\pm\epsilon}$, for $\alpha \in \left(0, \frac{1}{2}\right)$. It can be useful to have a sequence whose discrepancy decays faster. The best known rate of decay is achieved by the following example.

Definition 2.6.2. *The* van der Corput sequence *is* $\boldsymbol{v} = \left(\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \dots\right)$. *More precisely, write* n *in base* 2 *as* $n = \sum a_i 2^i$*; then* $v_n = \sum a_i 2^{-(i+1)}$.

The van der Corput sequence has generalizations to other bases and higher dimensions, but we will not use these. Its discrepancy has extremely fast convergence to zero.

Lemma 2.6.3 ([KN74, Ch. 2 Th. 3.5])**.** $\mathrm{D}_N(\boldsymbol{v}) \leqslant \frac{\log(N+1)}{N \log 2}$.

The van der Corput sequence is uniformly distributed (equidistributed with respect to the Lebesgue measure). We can use the results of the previous section to construct sequences equidistributed with respect to more general measures.

Theorem 2.6.4. *Let* μ *be an absolutely continuous probability measure on an interval* I *whose cdf is strictly increasing on* I*. Then there exists a sequence* $\boldsymbol{x} = (x_1, x_2, \dots)$ *in* I *such that* $\mathrm{D}_N(\boldsymbol{x}, \mu) \ll \frac{\log(N)}{N}$.

Proof. Since $(\mathrm{cdf}_\mu)_* \mu$ is the uniform measure, Lemma 2.5.3 tells us that $\left|\mathrm{D}_N(\mathrm{cdf}_\mu^{-1}(\boldsymbol{v}), \mu) - \mathrm{D}_N(\boldsymbol{v})\right| \ll N^{-1}$, which gives us the desired result with $\boldsymbol{x} = \mathrm{cdf}_\mu^{-1}(\boldsymbol{v})$. □

Theorem 2.6.5. *Let* μ *be an absolutely continuous probability measure, supported on* I*, whose cdf is strictly increasing on* I*. Fix* $\alpha \in (0, 1)$*. Then there exists a sequence* $\boldsymbol{x} = (x_1, x_2, \dots)$ *such that* $\mathrm{D}_N(\boldsymbol{x}, \mu) = \Theta(N^{-\alpha})$.

Proof. Let $I = [a, b]$. If $\boldsymbol{x}_{\leqslant N}$ is a sequence of length N, let $\boldsymbol{x}_{\leqslant N} :^M a$ be the sequence $(x_1, \dots, x_N, a, \dots, a)$ (M copies of a). Then

$$\mathrm{D}(\boldsymbol{x}_{\leqslant N} :^M a, \mu) \geqslant \left| \frac{\#\{n \leqslant N + M : x_n = a\}}{N + M} - \mu\{a\} \right| \geqslant \frac{M}{N + M}.$$

On the other hand for $J = [s, t) \subset I$,

$$\begin{aligned}\left|P_{\boldsymbol{x}_{\leqslant N}:^M a}(J) - P_{\boldsymbol{x}_{\leqslant N}}(J)\right| &\leqslant \frac{\left|\#\{n \leqslant N : x_n \in J\} + M - \frac{M+N}{N}\#\{n \leqslant N : x_n \in J\}\right|}{M+N} \\ &\leqslant \frac{2M}{M+N},\end{aligned}$$

which implies that $\mathrm{D}\left(\boldsymbol{x}_{\leqslant N} :^M a, \mu\right) \leqslant \mathrm{D}\left(\boldsymbol{x}^N, \mu\right) + \frac{2M}{M+N}$. Let $\boldsymbol{v}$ be the μ-equidistributed van der Corput sequence of Theorem 2.6.4, possibly transformed linearly to lie in $[a, b]$. We know that $\mathrm{D}(\boldsymbol{v}^N, \mu) \ll \frac{\log N}{N}$, which converges to zero faster than $N^{-\alpha}$.

We construct the sequence $\boldsymbol{x}$ via the following recipe. Start with $(x_1 = v_1, x_2 = v_2, \dots)$ until, for some N_1, $\mathrm{D}_{N_1}(\boldsymbol{x}, \mu) < N_1^{-\alpha}$. Then set $x_{N_1+1} = a$, $x_{N_1+2} = a$, …, until $\mathrm{D}_{N_1+M_1}(\boldsymbol{x}, \mu) > (N_1 + M_1)^{-\alpha}$. Then set $x_{N_1+M_1+1} = v_{N_1+1}$, $x_{N_1+M_1+2} = v_{N_1+2}$, …, until once again $\mathrm{D}_{N_1+M_1+N_2}(\boldsymbol{x}, \mu) < (N_1 + M_1 + N_2)^{-\alpha}$. Repeat indefinitely. We will show first, that the two steps are possible, and that nowhere does $\mathrm{D}_N(\boldsymbol{x}, \mu)$ differ by too much from $N^{-\alpha}$.

Note that $\frac{M+1}{N+M+1} - \frac{M}{N+M} \leqslant N^{-1}$. This tells us that when we are adding a's at the end of $\boldsymbol{x}_{\leqslant N}$, the discrepancy of $\boldsymbol{x}_{\leqslant N} :^M a$ increases by at most N^{-1} at each step. So if $\mathrm{D}(\boldsymbol{x}_{\leqslant N}, \mu) < N^{-\alpha}$, we can ensure that $\mathrm{D}(\boldsymbol{x}_{\leqslant N} :^M a, \mu)$ is at most N^{-1} greater than $N^{-\alpha}$. Moreover, we know that $\mathrm{D}(\boldsymbol{x}_{\leqslant N} : a, \mu)$ is at most $\frac{2}{N+1}$ away from $\mathrm{D}(\boldsymbol{x}_{\leqslant N}, \mu)$. So when adding van der Corput elements to the end of the sequence, its discrepancy cannot decay any faster than by $\frac{2}{N+1}$ per a added. This yields

$$\left|\mathrm{D}_N(\boldsymbol{x}, \mu) - N^{-\alpha}\right| \ll N^{-1},$$

which is even stronger than we need. □

CHAPTER 3

DIRICHLET SERIES WITH EULER PRODUCT

3.1 Definitions

We start by considering a very general class of Dirichlet series, namely all Dirichlet series that admit a product formula with degree 1 factors. The motivating example was suggested to the author by Ravi Ramakrishna. Let $E_{/\mathbf{Q}}$ be an elliptic curve and let

$$L_{\mathrm{sgn}}(E,s) = \prod_p \frac{1}{1-\mathrm{sgn}(a_p)p^{-s}}.$$

How much can we say about the behavior of $L_{\mathrm{sgn}}(E,s)$? For example, does it admit analytic continuation past $\Re = 1$? (Yes!) Can the rank of E be found from $L_{\mathrm{sgn}}(E,s)$? (Not so clear.)

Definition 3.1.1. *Let $\boldsymbol{x} = (x_2, x_3, x_5, \dots)$ be a sequence of complex numbers indexed by the primes. The associated Dirichlet series is $L(\boldsymbol{x},s) = \prod_p (1-x_p p^{-s})^{-1}$.*

If x_p is defined only for a subset of the primes, we tacitly set $x_p = 0$ (so the Euler factor is 1) at all primes for which x_p is not defined.

Lemma 3.1.2. *Let $\boldsymbol{x}$ be a sequence with $|\boldsymbol{x}|_\infty \leqslant 1$. Then $L(\boldsymbol{x},s)$ defines a holomorphic function on the region $\Re > 1$. On that region, $\log L(\boldsymbol{x},s) = \sum_{p^r} \frac{x_p^r}{rp^{rs}}$.*

Proof. Expanding the product for $L(\boldsymbol{x},s)$ formally, we have

$$L(\boldsymbol{x},s) = \sum_{n\geqslant 1} \frac{\prod_p x_p^{v_p(n)}}{n^s}.$$

An easy comparison with the Riemann zeta function tells us that this sum is holomorphic on $\Re > 1$. By [Apo76, Th. 11.7], the product formula holds in the same region. The formula for $\log L(\boldsymbol{x}, s)$ comes from [Apo76, 11.9 Ex. 2]. □

Lemma 3.1.3 (Abel summation)**.** *Let $\boldsymbol{x} = (x_2, x_3, x_5, \dots)$ be a sequence of complex numbers, f a smooth complex-valued function on $\mathbf{R}$. Then*

$$\sum_{p \leqslant N} f(p) x_p = f(N) \sum_{p \leqslant N} x_p - \int_2^N f'(t) \sum_{p \leqslant t} x_p \, \mathrm{d}t.$$

Proof. If $p_1, \dots, p_n$ is an enumeration of the primes $\leqslant N$, we have

$$\begin{aligned}
\int_2^N f'(t) \sum_{p \leqslant t} x_p \, \mathrm{d}t &= \sum_{p \leqslant N} x_p \int_{p_n}^N f'(t) \, \mathrm{d}t + \sum_{i=1}^{n-1} \sum_{p \leqslant p_i} x_p \int_{p_i}^{p_{i+1}} f'(t) \, \mathrm{d}t \\
&= (f(N) - f(p_n)) \sum_{p \leqslant N} x_p + \sum_{i=1}^{n-1} (f(p_{i+1}) - f(p_i)) \sum_{p \leqslant p_i} x_p \\
&= f(N) \sum_{p \leqslant N} x_p - \sum_{p \leqslant N} f(p) x_p,
\end{aligned}$$

as desired. □

Theorem 3.1.4. *Assume $|\sum_{p \leqslant N} x_p| \ll N^{\alpha + \epsilon}$ for some $\alpha \in [\frac{1}{2}, 1]$. Then the series for $\log L(\boldsymbol{x}, s)$ converges conditionally to a holomorphic function on $\Re > \alpha$.*

Proof. Formally split the sum for $\log L(\boldsymbol{x}, s)$ into two pieces:

$$\log L(\boldsymbol{x}, s) = \sum_p \frac{x_p}{p^s} + \sum_p \sum_{r \geqslant 2} \frac{x_p^r}{r p^{rs}}.$$

For each p, we have

$$\left| \sum_{r \geqslant 2} \frac{x_p^r}{r p^{rs}} \right| \leqslant \sum_{r \geqslant 2} p^{-r \Re s} = p^{-2\Re s} \frac{1}{1 - p^{-\Re s}}.$$

Elementary analysis gives $1 \leqslant \frac{1}{1 - p^{-\Re s}} \leqslant 2 + 2\sqrt{2}$, so the second piece of $\log L(\boldsymbol{x}, s)$ converges absolutely on $\Re > \frac{1}{2}$. We could simply cite [Ten95, II.1 Th. 10] to

finish the proof; instead we prove directly that $\sum \frac{x_p}{p^s}$ converges absolutely to a holomorphic function on the region $\Re > \alpha$.

By Lemma 3.1.3 (Abel summation) with $f(t) = t^{-s}$, we have

$$\begin{aligned}\sum_{p \leqslant N} \frac{x_p}{p^s} &= N^{-s} \sum_{p \leqslant N} x_p + s \int_2^N \sum_{p \leqslant t} x_p \frac{\mathrm{d}t}{t^{s+1}} \\ &\ll N^{-\Re s + \alpha + \epsilon} + |s| \int_2^N t^{\alpha+\epsilon} \frac{\mathrm{d}t}{t^{\Re s + 1}}.\end{aligned}$$

Since $\alpha - \Re s < 0$, the first term is bounded. Since $\Re s + 1 - \alpha > 1$ and ϵ is arbitrary, the integral converges absolutely, and the proof is complete. □

The proof of Theorem 3.1.4 actually gives an absolutely convergent expression for $\log L(\boldsymbol{x}, s)$ on the region $\Re > \alpha$. Namely,

$$\log L(\boldsymbol{x}, s) = s \int_2^\infty t^{-s-1} \left(\sum_{p \leqslant t} x_p \right) \mathrm{d}t + \sum_p \sum_{r \geqslant 2} \frac{x_p^r}{r p^{rs}}.$$

Let X be a space, $f \colon X \to \mathbf{C}$ a function with $|f|_\infty \leqslant 1$, and $\boldsymbol{x} = (x_2, x_3, \dots)$ a sequence in X. Write

$$L_f(\boldsymbol{x}, s) = \prod_p \frac{1}{1 - f(x_p) p^{-s}},$$

for the associated Dirichlet series. In the remainder, we will exclusively focus on Dirichlet series of this type.

3.2 Relation to automorphic and motivic L-functions

Suppose G is a compact group, $G^\natural$ the space of conjugacy classes in G. If $\boldsymbol{x} = (x_2, x_3, x_5, \dots)$ is a sequence in $G^\natural$ and ρ is a finite-dimensional representation of

G, put

$$L(\rho(\boldsymbol{x}), s) = \prod_p \frac{1}{\det(1 - \rho(x_p)p^{-s})}.$$

Clearly $L((\rho_1 \oplus \rho_2)(\boldsymbol{x}), s) = L(\rho_1(\boldsymbol{x}), s)L(\rho_2(\boldsymbol{x}), s)$. Now, let $T \subset G$ be a maximal torus, and recall that $T \twoheadrightarrow G^\natural$. The representation $\rho|_T$ decomposes as $\bigoplus \chi^{\oplus m_\chi}$, where χ ranges over characters of T and the entire expression is W-invariant. We may regard the x_p as lying in T/W, so we have

$$L(\rho(\boldsymbol{x}), s) = \prod_\chi L(\chi(\boldsymbol{x}), s)^{m_\chi}.$$

If the trivial representation appears in $\rho|_T$, this product formula will include a copy (possibly several) of $\zeta(s)$. Since $\chi(x_p) \in S^1$, the above formula decomposes $L(\rho(\boldsymbol{x}), s)$ into a product of Dirichlet series of the type considered above.

3.3 Discrepancy of sequences and the Riemann Hypothesis

Definition 3.3.1. *We say the* Riemann Hypothesis *for $L(\boldsymbol{x}, s)$ holds if the function $\log L(\boldsymbol{x}, s)$ admits analytic continuation to $\Re > \frac{1}{2}$.*

Under reasonable analytic hypotheses, namely conditional convergence of $\log L(\boldsymbol{x}, s)$ on $\Re > \frac{1}{2}$, [Ten95, II.1 Th. 10] gives an estimate $|\sum_{p \leqslant N} x_p| \ll N^{\frac{1}{2}+\epsilon}$.

Theorem 3.3.2. *Let (X, μ) be a probability space in which discrepancy and Koksma–Hlawka make sense, and let $\boldsymbol{x} = (x_2, x_3, x_5, \dots)$ be a sequence in X with $\mathrm{D}_N(\boldsymbol{x}, \mu) \ll N^{-\frac{1}{2}+\epsilon}$. For any function f on X of bounded variation with $\int f \, \mathrm{d}\mu = 0$, $L_f(\boldsymbol{x}, s)$ satisfies the Riemann Hypothesis.*

Proof. By the Koksma–Hlawka inequality, the bound on discrepancy yields the

estimate $\left|\sum_{p\leqslant N} f(x_p)\right| \ll N^{\frac{1}{2}+\epsilon}$. By Theorem 3.1.4, the Riemann Hypothesis holds for $L_f(\boldsymbol{x}, s)$. □

The same proof shows that if $\mathrm{D}_N(\boldsymbol{x}, \mu) \ll N^{-\alpha+\epsilon}$, then $\log L_f(\boldsymbol{x}, s)$ conditionally converges to a holomorphic function on $\Re > 1 - \alpha$.

Let $F = \mathbf{F}_q(t)$ be a function field, $E_{/F}$ a generic elliptic curve. There is, for every prime $\mathfrak{p}$ of F, a Satake parameter $\theta_\mathfrak{p} \in [0, \pi]$, defined in the usual way. It is known [Kat88, Ch. 3] that

$$\left|\sum_{\mathrm{N}(\mathfrak{p})\leqslant x} \operatorname{tr} \operatorname{sym}^k \begin{pmatrix} e^{i\theta_\mathfrak{p}} & \\ & e^{-i\theta_\mathfrak{p}} \end{pmatrix}\right| \ll k\sqrt{x}.$$

This tells us that for any $f \in C(\mathrm{SU}(2)^\natural)$ with $\sum_{k\geqslant 1} |\widehat{f}(\operatorname{sym}^k)| < \infty$ and $\widehat{f}(1) = 0$, the strange Dirichlet series $L_f(\boldsymbol{\theta}, s)$ satisfies the Riemann Hypothesis.

The best estimate on discrepancy is found in [Nie91], where it is shown that $D_N \ll N^{-\frac{1}{4}}$ by applying a generalization of the Koksma–Hlawka inequality to $\mathrm{SU}(2)^\natural$. Namely, for any odd r, we have

$$\mathrm{D}_x(\boldsymbol{\theta}, \mathrm{ST}) \ll \frac{1}{r} + \sum_{k=1}^{2r-1} \frac{1}{k} \left| \frac{1}{\pi_F(x)} \sum_{\mathrm{N}(\mathfrak{p})\leqslant x} \operatorname{tr} \operatorname{sym}^k \begin{pmatrix} e^{i\theta_\mathfrak{p}} & \\ & e^{-i\theta_\mathfrak{p}} \end{pmatrix}\right|.$$

Using the above estimate on character sums, Niederreiter is able to derive $\mathrm{D}_x \ll x^{-\frac{1}{4}}$. This fits the results of [BK15; RT16], both of which derive estimates of the form $\mathrm{D}_N \ll N^{-\frac{1}{4}+\epsilon}$ under GRH + functional equation for the (non-CM) elliptic curve in question.

CHAPTER 4

IRRATIONALITY EXPONENTS

4.1 Definitions and first results

We follow the notation of [Lau09]. Fix a dimension $d \geqslant 1$, and let $\vec{x} = (x_1, \ldots, x_d) \in \mathbf{R}^d$ be such that the x_i are linearly independent over $\mathbf{Q}$.

Definition 4.1.1. *Let $\omega_0(\vec{x})$ (resp. $\omega_{d-1}(\vec{x})$) be the supremum of the set of real numbers w for which there exist infinitely many $(n, \vec{m}) \in \mathbf{Z} \times \mathbf{Z}^d$ such that*

$$|n\vec{x} - \vec{m}|_\infty \leqslant |(n, \vec{m})|_\infty^{-w} \qquad \textit{(resp.}$$
$$|n + \langle \vec{m}, \vec{x} \rangle| \leqslant |(n, \vec{m})|_\infty^{-w}\textit{).}$$

These two quantities are related by Khintchine's transference principle, namely

$$\frac{\omega_{d-1}(\vec{x})}{(d-1)\omega_{d-1}(\vec{x}) + d} \leqslant \omega_0(\vec{x}) \leqslant \frac{\omega_{d-1}(\vec{x}) - d + 1}{d}.$$

Moreover, the second of these inequalities is sharp in a very strong sense.

Theorem 4.1.2 (Jarník)**.** *Let $w \geqslant 1/d$. Then there exists $\vec{x} \in \mathbf{R}^d$ such that $\omega_0(\vec{x}) = w$ and $\omega_{d-1}(\vec{x}) = dw + d - 1$.*

Theorem 4.1.3. *If $d = 1$, then $\omega_0(x) = \mu - 1$, where μ is the traditional irrationality measure of x.*

Theorem 4.1.4 (Roth)**.** *Let $x \in (\overline{\mathbf{Q}} \cap \mathbf{R}) \smallsetminus \mathbf{Q}$. Then $\omega_0(x) = 1$.*

Given $\vec{x} \in \mathbf{R}^d$, write $\mathrm{d}(\vec{x}, \mathbf{Z}^d) = \min_{\vec{m} \in \mathbf{Z}^d} |\vec{x} - \vec{m}|_\infty$. Note that $\mathrm{d}(\vec{x}, \mathbf{Z}^d) = 0$ if and only if $\vec{x} \in \mathbf{Z}^d$. Moreover, $\mathrm{d}(-, \mathbf{Z}^d)$ is well-defined for elements of $(\mathbf{R}/\mathbf{Z})^d$.

Lemma 4.1.5. *Let $\vec{x} \in \mathbf{R}^d$ with $|\vec{x}|_\infty < 1$ and $\omega_0(\vec{x})$ (resp. $\omega_{d-1}(\vec{x})$) finite. Then*

$$\frac{1}{\mathrm{d}(n\vec{x}, \mathbf{Z}^d)} \ll |n|^{\omega_0(\vec{x})+\epsilon} \qquad \textit{for } n \in \mathbf{Z} \textit{ (resp.}$$
$$\frac{1}{\mathrm{d}\left(\langle \vec{m}, \vec{x}\rangle, \mathbf{Z}\right)} \ll |\vec{m}|_\infty^{\omega_{d-1}(\vec{x})+\epsilon} \qquad \textit{for } \vec{m} \in \mathbf{Z}^d\textit{).}$$

Proof. Let $\epsilon > 0$. Then there are only finitely many $n \in \mathbf{Z}$ (resp. $\vec{m} \in \mathbf{Z}^d$) such that the inequalities in Definition 4.1.1 hold with $w = \omega_0(x)+\epsilon$ (resp. $w = \omega_{d-1}(\vec{x})+\epsilon$). In other words, there exist constants $C_0, C_{d-1} > 0$ such that

$$|n\vec{x} - \vec{m}|_\infty \geqslant C_0 |(n, \vec{m})|_\infty^{-\omega_0(\vec{x})-\epsilon},$$
$$|n + \langle \vec{m}, \vec{x}\rangle| \geqslant C_{d-1} |(n, \vec{m})|_\infty^{-\omega_{d-1}(\vec{x})-\epsilon}$$

for all $(n, \vec{m}) \neq 0$ in $\mathbf{Z} \times \mathbf{Z}^d$.

Start with the first inequality. Fix n, and let $\vec{m}$ be the lattice point achieving the minimum $|n\vec{x}-\vec{m}|_\infty$; then $\mathrm{d}(n\vec{x}, \mathbf{Z}^d) \geqslant C_0 |(n, \vec{m})|_\infty^{-\omega_0(\vec{x})-\epsilon}$. Since $|n\vec{x}-\vec{m}|_\infty < 1$, $|n| \geqslant \frac{|\vec{m}|_\infty}{|\vec{x}|_\infty} - 1$, which gives $\mathrm{d}(n\vec{x}, \mathbf{Z}^d) \geqslant C_0' |n|^{-\omega_0(\vec{x})-\epsilon}$ for C_0' depending on $\vec{x}$. It follows that $\frac{1}{\mathrm{d}(n\vec{x}, \mathbf{Z}^d)} \ll |n|^{\omega_0(\vec{x})+\epsilon}$, the implied constant depending on x and ϵ.

Now we consider the second inequality. Note that $\mathrm{d}(\langle \vec{m}, \vec{x}\rangle, \mathbf{Z}) = |n + \langle \vec{m}, \vec{x}\rangle|$ for some n with $|n| \leqslant \|\vec{m}\|_2 \|\vec{x}\|_2 + 1$. Thus $|(n, \vec{m})|_\infty \ll |\vec{m}|_2 \ll |\vec{m}|_\infty$ (any two norms on a finite-dimensional Banach space are equivalent), which gives us $\mathrm{d}(\langle \vec{m}, \vec{x}\rangle, \mathbf{Z}) \geqslant C_{d-1}' |\vec{m}|^{-\omega_{d-1}(\vec{x})-\epsilon}$, for some constant C_{d-1}'. This implies

$$\frac{1}{d(\langle \vec{m}, \vec{x}\rangle, \mathbf{Z})} \ll |\vec{m}|^{\omega_{d-1}(\vec{x})+\epsilon},$$

the implied constant depending on $\vec{x}$ and ϵ. □

4.2 Irrationality exponents and discrepancy

Let $\vec{x} \in (\mathbf{R}/\mathbf{Z})^d$ with $x_1, \dots, x_d$ linearly independent over $\mathbf{Q}$ (this condition makes sense for elements of $(\mathbf{R}/\mathbf{Z})^d$). We wish to control the discrepancy of the sequence $(x, 2x, 3x, \dots)$ with respect to the Haar measure of $(\mathbf{R}/\mathbf{Z})^d$.

Theorem 4.2.1 (Erdös–Turán–Koksma)**.** *Let $\boldsymbol{x}$ be a sequence in $(\mathbf{R}/\mathbf{Z})^d$ and h an arbitrary integer. Then*

$$\mathrm{D}_N(\boldsymbol{x}) \ll \frac{1}{h} + \sum_{0 \leqslant |\vec{m}|_\infty \leqslant h} \frac{1}{r(\vec{m})} \left| \frac{1}{N} \sum_{n \leqslant N} e^{2\pi i \langle \vec{m}, \vec{x}_n \rangle} \right|,$$

where the first sum ranges over $\vec{m} \in \mathbf{Z}^d$, $r(\vec{m}) = \prod \max\{1, |m_i|\}$, and the implied constant depends only on d.

Proof. This is [DT97, Th. 1.21]. □

Lemma 4.2.2. *Let $x \in \mathbf{R}$. Then*

$$\left| \sum_{n \leqslant N} e^{2\pi i n x} \right| \leqslant \frac{2}{\mathrm{d}(x, \mathbf{Z})}.$$

Proof. We begin with an easy bound:

$$\left| \sum_{n \leqslant N} e^{2\pi i n x} \right| = \frac{|e^{2\pi i (N+1)x} - e^{2\pi i x}|}{|e^{2\pi i x} - 1|} \leqslant \frac{2}{|e^{2\pi i x} - 1|}.$$

Since $|e^{2\pi i x} - 1| = \sqrt{2 - 2\cos(2\pi x)}$ and $\cos(2\theta) = 1 - 2\sin^2\theta$, we obtain

$$\left| \sum_{n \leqslant N} e^{2\pi i n x} \right| \leqslant \frac{1}{|\sin(\pi x)|}.$$

It is easy to check that $|\sin(\pi x)| \geqslant \mathrm{d}(x, \mathbf{Z})$, whence the result. □

Corollary 4.2.3. *Let $\vec{x} \in (\mathbf{R}/\mathbf{Z})^d$ with $(x_1, \dots, x_d)$ linearly independent over $\mathbf{Q}$. Then for $\boldsymbol{x} = (\vec{x}, 2\vec{x}, 3\vec{x}, \dots)$, we have*

$$\mathrm{D}_N(\boldsymbol{x}) \ll \frac{1}{h} + \frac{1}{N} \sum_{0<|\vec{m}|_\infty \leqslant h} \frac{1}{r(\vec{m})\mathrm{d}(\langle \vec{m}, \vec{x} \rangle, \mathbf{Z})}$$

for any integer h, with the implied constant depending only on d.

Proof. Apply the Erdös–Turán–Koksma inequality, and bound the exponential sums using Lemma 4.2.2. □

Theorem 4.2.4. *Let $\boldsymbol{x} = (\vec{x}, 2\vec{x}, 3\vec{x}, \dots)$ in $(\mathbf{R}/\mathbf{Z})^d$. Then*

$$\mathrm{D}_N(\boldsymbol{x}) \ll N^{-\frac{1}{\omega_{d-1}(\vec{x})+1}+\epsilon}.$$

Proof. Fix $\epsilon > 0$, and choose $\delta > 0$ such that $\frac{1}{\omega_{d-1}(\vec{x})+1+\delta} = \frac{1}{\omega_{d-1}(\vec{x})+1} - \epsilon$. By Corollary 4.2.3, we know that

$$\mathrm{D}_N(\boldsymbol{x}) \ll \frac{1}{h} + \frac{1}{N} \sum_{0<|\vec{m}|_\infty \leqslant h} \frac{1}{r(\vec{m})\mathrm{d}(\langle \vec{m}, \vec{x} \rangle, \mathbf{Z})},$$

and by Lemma 4.1.5, we know that $\mathrm{d}(\langle \vec{m}, \vec{x} \rangle, \mathbf{Z})^{-1} \ll |\vec{m}|^{\omega_{d-1}(\vec{x})+\delta}$. It follows that

$$\mathrm{D}_N(\boldsymbol{x}) \ll \frac{1}{h} + \frac{1}{N} \sum_{0<|\vec{m}|_\infty \leqslant h} \frac{|\vec{m}|_\infty^{\omega_{d-1}(\vec{x})+\delta}}{r(\vec{m})}.$$

All that remains is to bound the sum.

$$\begin{aligned}
\sum_{0<|\vec{m}|_\infty \leqslant h} \frac{|\vec{m}|_\infty^{\omega_{d-1}(\vec{x})+\delta}}{r(m)} &\ll \int_1^h \int_1^{t_d} \cdots \int_1^{t_2} \frac{t_d^{\omega_{d-1}(\vec{x})+\delta}}{t_1 \dots t_d} \,\mathrm{d}t_1 \dots \mathrm{d}t_d \\
&\ll \int_1^h t^{\omega_{d-1}(\vec{x})+\delta-1} \,\mathrm{d}t \prod_{j=1}^{d-1} \int_1^h \frac{\mathrm{d}t}{t} \\
&\ll (\log h)^{d-1} h^{\omega_{d-1}(\vec{x})+\delta}.
\end{aligned}$$

It follows that $\mathrm{D}_N(\boldsymbol{x}) \ll \frac{1}{h} + \frac{1}{N}(\log h)^{d-1} h^{\omega_{d-1}(\vec{x})+\delta}$. Setting $h \approx N^{\frac{1}{1+\omega_{d-1}(\vec{x})+\delta}}$, we see that $D_N(\boldsymbol{x}) \ll N^{-\frac{1}{\omega_{d-1}(\vec{x})+1+\delta}} = N^{-\frac{1}{\omega_{d-1}(\vec{x})+1}+\epsilon}$.

For a slightly different proof of a similar result, given as a sequence of exercises, see [KN74, Ch. 2, Ex. 3.15, 16, 17]. □

Theorem 4.2.5. *Let $\vec{x} \in \mathbf{R}^d$ be such that $x_1, \dots, x_d$ are linearly independent over $\mathbf{Q}$, and let $\boldsymbol{x} = (\vec{x}, 2\vec{x}, 3\vec{x}, \dots)$ in $(\mathbf{R}/\mathbf{Z})^d$. Then $\mathrm{D}_N(\boldsymbol{x}) = \Omega\left(N^{-\frac{d}{\omega_0(\vec{x})}-\epsilon}\right)$.*

Proof. We follow the proof of [KN74, Ch. 2, Th. 3.3], modifying it as needed for our context. Given $\epsilon > 0$, there exists $\delta > 0$ such that $\frac{d}{\omega_0(\vec{x})-\delta} = \frac{d}{\omega_0(\vec{x})} + \epsilon$.

By the definition of $\omega_0(\vec{x})$, there exist infinitely many $(q, \vec{m})$ with $q > 0$ such that $|q\vec{x} - \vec{m}|_\infty \leqslant |(q, \vec{m})|_\infty^{-\omega_0(\vec{x})+\delta/2}$. Since $|(q, \vec{m})|_\infty \geqslant q$, we derive the stronger statement that for infinitely many $q \to \infty$, there exists $\vec{m} \in \mathbf{Z}^d$ such that $|q\vec{x} - \vec{m}|_\infty \leqslant q^{-\omega_0(\vec{x})+\delta/2}$ or, equivalently, $|\vec{x} - q^{-1}\vec{m}| \leqslant q^{-1-\omega_0(\vec{x})+\delta/2}$. Fix one such q, and let $N = \lfloor q^{\omega_0(\vec{x})-\delta} \rfloor$. For each $n \leqslant N$, we have $|n\vec{x} - \frac{n}{q}\vec{m}|_\infty \leqslant q^{-1-\delta/2}$. For each $n \leqslant N$, $n\vec{x}$ is within $q^{-1-\delta/2}$ of the grid $\frac{1}{q}\mathbf{Z}^d \subset (\mathbf{R}/\mathbf{Z})^d$. Thus, they miss a box with side lengths $q^{-1} - 2q^{-1-\delta/2}$. For q sufficiently large, $q^{-1} - 2q^{-1-\delta/2} \geqslant 1/2q$, so the discrepancy is bounded below by $2^{-d}q^{-d}$. Since $q^{\omega_0(\vec{x})-\delta} \leqslant 2N$, the discrepancy at N is bounded below by

$$2^{-d}\left((2N)^{\frac{1}{\omega_0(\vec{x})+\delta}}\right)^{-d} = 2^{-d-\frac{d}{\omega_0(\vec{x})+\delta}} N^{-\frac{d}{\omega_0(\vec{x})+\delta}} = 2^{-d\left(1+\frac{1}{\omega_0(\vec{x})}\right)-\epsilon} N^{-\frac{d}{\omega_0(\vec{x})}-\epsilon}.$$

Since $\mathrm{D}_N(\boldsymbol{x})$ can, for $N \to \infty$, be bounded below by a constant multiple of $N^{-\frac{d}{\omega_0(\vec{x})}}$, the proof is complete. □

4.3 Pathological Satake parameters for CM abelian varieties

Give $(\mathbf{R}/\mathbf{Z})^d$ the Haar measure normalized to have total mass one. Recall that for any $f \in L^1((\mathbf{R}/\mathbf{Z})^d)$, the Fourier coefficients of f are, for $\vec{m} \in \mathbf{Z}^d$:

$$\widehat{f}(\vec{m}) = \int_{(\mathbf{R}/\mathbf{Z})^d} e^{2\pi i \langle \vec{m}, \vec{x} \rangle} \, \mathrm{d}\vec{x},$$

where $\langle \vec{m}, \vec{x} \rangle = m_1 x_1 + \cdots + m_d x_d$ is the usual inner product.

Theorem 4.3.1. *Fix $\vec{x} \in (\mathbf{R}/\mathbf{Z})^d$ with $\omega_{d-1}(\vec{x})$ finite. Then*

$$\left| \sum_{n \leqslant N} e^{2\pi i \langle \vec{m}, n\vec{x} \rangle} \right| \ll |\vec{m}|_\infty^{\omega_{d-1}(\vec{x})+\epsilon}$$

as $\vec{m}$ ranges over $\mathbf{Z}^d \smallsetminus 0$.

Proof. From Lemma 4.2.2 we know that $\left|\sum_{n \leqslant N} e^{2\pi i \langle \vec{m}, n\vec{x} \rangle}\right| \ll \mathrm{d}(\langle \vec{m}, \vec{x} \rangle, \mathbf{Z})^{-1}$, and from Lemma 4.1.5, we know that $\mathrm{d}(\langle \vec{m}, \vec{x} \rangle, \mathbf{Z})^{-1} \ll |\vec{m}|_\infty^{\omega_{d-1}(x)+\epsilon}$. The result follows.

□

Theorem 4.3.2. *Let $\vec{x} \in \mathbf{R}^d$ with $\omega_{d-1}(\vec{x})$ finite. Fix $f \in L^1((\mathbf{R}/\mathbf{Z})^d)$ with $\widehat{f}(0) = 0$ and suppose the Fourier coefficients of f satisfy the bound $|\widehat{f}(\vec{m})| \ll |\vec{m}|_\infty^{-\frac{1}{d-1} - \omega_{d-1}(\vec{x}) - \epsilon}$. Then $\left|\sum_{n \leqslant N} f(n\vec{x})\right| \ll 1$.*

Proof. Write f as a Fourier series: $f(\vec{x}) = \sum_{\vec{m} \in \mathbf{Z}^d} \widehat{f}(\vec{m}) e^{2\pi i \langle \vec{m}, \vec{x} \rangle}$. Since $\widehat{f}(0) = 0$,

we can compute:

$$\begin{aligned}\left|\sum_{n\leqslant N} f(n\vec{x})\right| &= \left|\sum_{n\leqslant N}\sum_{\vec{m}\in\mathbf{Z}^d\smallsetminus 0} \widehat{f}(\vec{m})e^{2\pi i n\langle\vec{m},\vec{x}\rangle}\right| \\ &\leqslant \sum_{\vec{m}\in\mathbf{Z}^d\smallsetminus 0} |\widehat{f}(\vec{m})|\left|\sum_{n\leqslant N} e^{2\pi i n\langle\vec{m},\vec{x}\rangle}\right| \\ &\ll \sum_{\vec{m}\in\mathbf{Z}^d\smallsetminus 0} |\vec{m}|_\infty^{-\frac{1}{d-1}-\omega_{d-1}(\vec{x})-\epsilon}|m|_\infty^{\omega_{d-1}(\vec{x})+\epsilon/2} \\ &\ll \sum_{\vec{m}\in\mathbf{Z}^d\smallsetminus 0} |\vec{m}|_\infty^{-\frac{1}{d-1}-\epsilon/2}.\end{aligned}$$

The sum converges since the exponent is less than $-\frac{1}{d-1}$, hence the result. □

Let $p_1 = 2, p_2 = 3, p_3 = 5, \dots$ be an enumeration of the rational primes. Let $\vec{y} \in \mathbf{R}^d$ with $y_1, \dots, y_d$ linearly independent over $\mathbf{Q}$. The associated sequence of "fake Satake parameters" is $\boldsymbol{x} = (\vec{y}, 2\vec{y}, 3\vec{y}, 4\vec{y}, \dots)$, where we put $\vec{x}_{p_n} = n\vec{y}$ mod $\mathbf{Z}^d$. By Theorem 4.1.2, we can arrange for $\omega_0(\vec{y}) = w$ and $\omega_{d-1}(\vec{y}) = dw+d-1$.

Theorem 4.3.3. *The sequence $\boldsymbol{x}$ is equidistributed in $(\mathbf{R}/\mathbf{Z})^d$, with discrepancy decaying as $\mathrm{D}_N(\boldsymbol{x}) \ll N^{-\frac{1}{dw+d}+\epsilon}$, and for which $\mathrm{D}_N(\boldsymbol{x}) = \Omega\left(N^{-\frac{d}{w}-\epsilon}\right)$. However, for any $f \in C^\infty((\mathbf{R}/\mathbf{Z})^d)$ with $\widehat{f}(0) = 0$, the Dirichlet series $L_f(\boldsymbol{x}, s)$ satisfies the Riemann Hypothesis.*

Proof. The upper bound on discrepancy is Theorem 4.2.4, and the lower bound is Theorem 4.2.5. For the functions f in question, Theorem 4.3.2 gives an estimate (stronger than) $\left|\sum_{p\leqslant N} f(\vec{x}_p)\right| \ll N^{\frac{1}{2}}$, and Theorem 3.1.4 tells us this estimate implies the Riemann Hypothesis. □

Let's apply this theorem to abelian varieties with complex multiplication. Let $K/\mathbf{Q}$ be a finite Galois extension, $A_{/K}$ an abelian variety with complex multipli-

cation by F, defined over K. Let $\mathfrak{a} = \mathrm{Lie}(A)$ be the Lie algebra of A; this is a K-vector space with natural F-action. The determinant of this action gives a homomorphism $\det_{\mathfrak{a}}\colon \mathrm{R}_{K/\mathbf{Q}}\,\mathbf{G}_{\mathrm{m}} \to \mathrm{R}_{F/\mathbf{Q}}\,\mathbf{G}_{\mathrm{m}}$, whose image is the motivic Galois group G_A [Yu15]. Let $\mathrm{N}_{F/\mathbf{Q}}\colon \mathrm{R}_{F/\mathbf{Q}}\,\mathbf{G}_{\mathrm{m}} \to \mathbf{G}_{\mathrm{m}}$ be the norm map; then in the notation of [Ser94], $G_A^1 = \mathrm{im}(\det_{\mathfrak{a}})^{\mathrm{N}_{F/\mathbf{Q}}=1}$, and $\mathrm{ST}(A)$ is the maximal compact subgroup of $G_A^1(\mathbf{C})$. Every unitary representation of $\mathrm{ST}(A)$ is the restriction of a character of G_A. Since characters always extend from subtori, any unitary representation of $\mathrm{ST}(A)$ is the restriction of a character of $\mathrm{R}_{F/\mathbf{Q}}\,\mathbf{G}_{\mathrm{m}}$.

Fix a rational prime l. The Galois representation associated to A is $\rho_l\colon G_{\mathbf{Q}} \to G_A(\mathbf{Q}_l) \subset (F \otimes \mathbf{Q}_l)^\times$. For $\sigma\colon F \hookrightarrow \mathbf{C}$, also write σ for the corresponding character of $\mathrm{R}_{F/\mathbf{Q}}\,\mathbf{G}_{\mathrm{m}}$. Shimura and Taniyama proved [ST68] that there exists a Hecke character $\chi_\sigma\colon \mathbf{A}_K^\times/K^\times \to \mathbf{C}^\times$ such that $L^{\mathrm{alg}}(\sigma \circ \rho_l, s) = L(s, \chi_\sigma)$. The analytic L-function of A is normalized as $L(A,s) = \prod_\sigma L^{\mathrm{alg}}(\sigma \circ \rho_l, s + 1/2)$, so for $\omega_\sigma = \chi_\sigma \| \cdot \|^{-1/2}$, we have

$$L(A,s) = \prod_{\sigma\colon F \hookrightarrow \mathbf{C}} L(s, \omega_\sigma).$$

A character r of $\mathrm{R}_{F/\mathbf{Q}}\,\mathbf{G}_{\mathrm{m}}$ is of the form $r = \sum a_\sigma \sigma$ with $a \in \mathbf{Z}$. Write $\omega_r = \prod \omega_\sigma^{a_\sigma}$. Then

$$L(r_* \rho_l, s) = L(s, \omega_r).$$

The Riemann Hypothesis holds for all $L(r_*\rho_l, s)$ if and only if it holds for each $L(\sigma \circ \rho_l, s) = L(s, \omega_\sigma)$. Since we already know analytic continuation of $\log L(s, \omega_\sigma)$ past $\Re = 1$, the Sato–Tate conjecture holds for A. However, the above theorem shows that even if each $L(r_*\rho_l, s)$ satisfies the Riemann Hypothesis, we may not immediately conclude that the Akiyama–Tanigawa conjecture holds for A.

CHAPTER 5

DEFORMATION THEORY

5.1 Category of test objects

This section summarizes the theory in [SGA 3_1, $\mathrm{VII_B}$, §0–1], adapting it to the deformation theory of Galois representations. All rings are commutative with unit.

Definition 5.1.1. *Let Λ be a topological ring. A topological Λ-module M is* pseudocompact *if it is a filtered inverse limit of discrete finite-length Λ-modules. The ring Λ is pseudocompact if it is pseudocompact as a module over itself.*

Let Λ be a pseudocompact ring, and write C_Λ for the opposite of the category of Λ-algebras which have finite length as Λ-modules. Given such a Λ-algebra A, write $X = \mathrm{Spf}(A)$ for the corresponding object of C_Λ, and put $A = \mathscr{O}(X)$.

Lemma 5.1.2. *Let Λ be a pseudocompact ring. Then C_Λ is closed under finite limits and colimits.*

Proof. That C_Λ is closed under finite colimits follows from the fact that finite-length Λ-algebras are closed under finite limits (the underlying modules are closed under finite limits). Moreover, since the tensor product of finite length modules also has finite length, and quotients of length modules have finite length, C_Λ is closed under finite limits. □

Lemma 5.1.3. *Let Λ be a pseudocompact local ring. Then Λ is henselian, in any of the following senses:*

1. *Every finite Λ-algebra is a product of local Λ-algebras.*

2. *The first condition is satisfied for Λ-algebras of the form $\Lambda[t]/f$, where f is monic.*

3. *Let $\mathfrak{m}$ be the maximal ideal of Λ. Then $A \mapsto A/\mathfrak{m}$ is an equivalence of categories from finite étale Λ-algebras to finite étale $\Lambda/\mathfrak{m}$-algebras.*

Proof. The conditions are equivalent by [EGA 4_4, 18.5.11], so we only prove that the first holds. Recall that $\Lambda = \varprojlim \Lambda/\mathfrak{a}$, where $\mathfrak{a}$ ranges over closed ideals of finite index. Let A be a pseudocompact Λ-algebra. For any ideal $\mathfrak{a} \subset \Lambda$, the ring $\Lambda/\mathfrak{a}$ is henselian by [EGA 4_4, 18.5.14], so $A/\mathfrak{a}$ is a product of local $\Lambda/\mathfrak{a}$-algebras. Moreover, by [EGA 4_4, 18.5.4], the map $A/\mathfrak{a} \to A/\mathfrak{m}$ is a bijection on idempotents. The inverse limit of these compatible systems of idempotents decompose A into a product of local Λ-algebras. □

Following Grothendieck, if $\mathcal{C}$ is an arbitrary category, we write $\widehat{\mathcal{C}} = \hom(\mathcal{C}^\circ, \mathsf{Set})$ for the category of contravariant functors $\mathcal{C} \to \mathsf{Set}$. We regard $\mathcal{C}$ as a full subcategory of $\widehat{\mathcal{C}}$ via the Yoneda embedding: for $X, Y \in \mathcal{C}$, write $X(Y) = \hom_{\mathcal{C}}(Y, X)$. With this notation, the Yoneda Lemma states that $\hom_{\widehat{\mathcal{C}}}(X, P) = P(X)$ for all $X \in \mathcal{C}$.

Lemma 5.1.4 ([KS06, 6.1])**.** *Let $\mathcal{X} \in \widehat{\mathsf{C}_\Lambda}$. Then $\mathcal{X}$ is left exact if and only if there exists a filtered system $\{X_i\}_{i \in I}$ in $\mathcal{C}_\Lambda$ together with a natural isomorphism $\mathcal{X} \simeq \varinjlim X_i$. Write $\mathsf{Ind}(\mathsf{C}_\Lambda)$ for the category of such functors. Then $\mathsf{Ind}(\mathsf{C}_\Lambda)$ is closed under colimits, and the Yoneda embedding $\mathsf{C}_\Lambda \hookrightarrow \mathsf{Ind}(\mathsf{C}_\Lambda)$ preserves filtered colimits.*

Lemma 5.1.5 ([KS06, 6.1.17])**.** *The functors* $\mathsf{C}_\Lambda \to \mathsf{Ind}(\mathsf{C}_\Lambda) \to \widehat{\mathsf{C}_\Lambda}$ *are left exact.*

If R is a pseudocompact Λ-algebra, write $\mathrm{Spf}(R)$ for the object of $\widehat{\mathsf{C}_\Lambda}$ defined by $\mathrm{Spf}(R)(A) = \hom_{\mathrm{cts}/\Lambda}(R, A)$, the set of continuous Λ-algebra homomorphisms.

Lemma 5.1.6 ([SGA 3_1, $\mathrm{VII_B}$ 0.4.2 Prop.])**.** *The funtor* Spf *induces an (anti-) equivalence between the category of pseudocompact* Λ*-algebras and* $\mathsf{Ind}(\mathsf{C}_\Lambda)$*.*

So $\mathsf{Ind}(\mathsf{C}_\Lambda)$ is the category of pro-representable functors on finite length Λ-algebras. *Warning*: in many papers, for example the foundational [Maz97], one reserves the term *pro-representable* for functors of the form $\mathrm{Spf}(R)$, where R is required to be noetherian. We do not make this restriction.

Lemma 5.1.7. *The category* $\mathsf{Ind}(\mathsf{C}_\Lambda)$ *is an exponential ideal in* $\widehat{\mathsf{C}_\Lambda}$*.*

Proof. By this we mean the following. Let $\mathcal{X} \in \mathsf{Ind}(\mathsf{C}_\Lambda)$, $P \in \widehat{\mathsf{C}_\Lambda}$. Then the functor $\mathcal{X}^P$ defined by $\mathcal{X}^P(S) = \hom_{\widehat{\mathsf{C}_\Lambda}_{/S}}(P_{/S}, \mathcal{X}_{/S})$ is also in $\mathsf{Ind}(\mathsf{C}_\Lambda)$. Given the characterization of $\mathsf{Ind}(\mathsf{C}_\Lambda)$ as left exact functors, the lemma follows directly from [Joh02, 4.2.3]. □

If $\mathcal{C}$ is a category, we write $\mathsf{Gp}(\mathcal{C})$ for the category of group objects in $\mathcal{C}$.

Corollary 5.1.8. *Let* $\Gamma \in \mathsf{Gp}(\widehat{\mathsf{C}_\Lambda})$ *and* $\mathcal{G} \in \mathsf{Gp}(\mathsf{Ind}(\mathsf{C}_\Lambda))$*, then the functor* $[\Gamma, \mathcal{G}]$ *defined by* $[\Gamma, \mathcal{G}](S) = \hom_{\mathsf{Gp}_{/S}}(\Gamma_{/S}, \mathcal{G}_{/S})$ *is in* $\mathsf{Ind}(\mathsf{C}_\Lambda)$*. In particular, if* Γ *is a profinite group, then the functor* $[\Gamma, \mathcal{G}](S) = \hom_{\mathrm{cts}/\,\mathsf{Gp}}(\Gamma, \mathcal{G}(S))$ *is in* $\mathsf{Ind}(\mathsf{C}_\Lambda)$*.*

Proof. The first claim follows easily from Lemma 5.1.7 and Lemma 5.1.5. Just note that $[\Gamma, \mathcal{G}]$ is the equalizer:

$$[\Gamma, \mathcal{G}] \longrightarrow \mathcal{G}^\Gamma \underset{m_{\mathcal{G}*}}{\overset{m_\Gamma^*}{\rightrightarrows}} \mathcal{G}^{\Gamma\times\Gamma},$$

that is, those $f\colon \Gamma \to \mathcal{G}$ such that $f \circ m_\Gamma = m_\mathcal{G} \circ (f \times f)$. The second claim is a special case of the first one. □

5.2 Quotients in the flat topology

If Λ is a pseudocompact ring, the category $\mathsf{Ind}(\mathsf{C}_\Lambda)$ has nice geometric properties. However, for operations like taking quotients, it is convenient to embed it into the larger category $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$ of flat sheaves. We call a collection $\{U_i \to X\}$ of morphisms in C_Λ a *flat cover* if each ring map $\mathscr{O}(X) \to \mathscr{O}(U_i)$ is flat, and moreover $\mathscr{O}(X) \to \prod \mathscr{O}(U_i)$ is faithfully flat. By [SGA 3_1, IV 6.3.1], this is a subcanonical Grothendieck topology on C_Λ. We call it the *flat topology*, even though finite presentation comes for free because all the rings involved are finite length over Λ.

Lemma 5.2.1. *Let $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$ be the category of sheaves (of sets) on C_Λ with respect to the flat topology. Then a presheaf $P \in \widehat{\mathsf{C}_\Lambda}$ lies in $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$ if and only if $P(\coprod U_i) = \prod P(U_i)$ and whenever $U \to X$ is a flat cover where $\mathscr{O}(U)$ and $\mathscr{O}(X)$ are local rings, the sequence*

$$P(X) \longrightarrow P(U) \rightrightarrows P(U \times_X U).$$

is exact. Moreover, $\mathsf{Ind}(\mathsf{C}_\Lambda) \subset \mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$.

Proof. The first claim is the content of [SGA 3_1, IV 6.3.1(ii)]. For the second, note that any $\mathcal{X} \in \mathsf{Ind}(\mathsf{C}_\Lambda)$ will, by Lemma 5.1.4, convert arbitrary colimits into limits. Thus $\mathcal{X}(\coprod U_i) = \prod \mathcal{X}(U_i)$. If $U \to X$ is a flat cover, then by (loc. cit.), $U \times_X U \rightrightarrows U \to X$ is a coequalizer diagram in C_Λ, hence $\mathcal{X}(X) \to \mathcal{X}(U) \rightrightarrows \mathcal{X}(U \times_X U)$ is an equalizer. □

Our main reason for introducing the category $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$ is that, as a Grothendieck topos, it is closed under arbitrary colimits. Recall that in an *equivalence relation* in $\widehat{\mathsf{C}_\Lambda}$ is a morphism $R \to X \times X$ such that, for all S, the map $R(S) \to X(S) \times X(S)$ is an injection whose image is an equivalence relation on $X(S)$. We define the quotient X/R to be the coequalizer

$$R \rightrightarrows X \longrightarrow X/R.$$

By Giraud's Theorem [MLM94, App.], for any $S \in \mathsf{C}_\Lambda$, the natural map $X(S)/R(S) \to (X/R)(S)$ is injective. It will not be surjective in general.

We let $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$ inherit definitions from C_Λ as follows. If P is a property of maps in C_Λ (for example, "flat," or "smooth,") and $f\colon X \to Y$ is a morphism in $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$, we say that f has P if for all $S \in \mathsf{C}_\Lambda$ and $y \in Y(S)$, the pullback $X_S = X \times_Y S$ lies in C_Λ, and the pullback map $X_S \to S$ has property P. For example, if $X = \mathrm{Spf}(R')$ and $Y = \mathrm{Spf}(R)$, then $X \to Y$ has property P if and only if for all finite length A and continuous Λ-algebra maps $R \to A$, the induced map $A \to R' \otimes_R A$ has P.

Theorem 5.2.2 ([SGA 3_1, $\mathrm{VII_B}$ 1.4])**.** *Let $\mathcal{R} \to \mathcal{X} \times \mathcal{X}$ be an equivalence relation in $\mathsf{Ind}(\mathsf{C}_\Lambda)$ such that one of the maps $\mathcal{R} \to \mathcal{X}$ is flat. Then the quotient $\mathcal{X}/\mathcal{R}$ lies in $\mathsf{Ind}(\mathsf{C}_\Lambda)$, and $\mathcal{X} \to \mathcal{X}/\mathcal{R}$ is a flat cover.*

By [Mat89, 29.7], if k is a field and R is a complete regular local k-algebra, then $R \simeq k[\![t_1, \ldots, t_n]\!]$. In particular, R admits an augmentation $R \to k$. There is a general analogue of this result, but first we need a definition.

Definition 5.2.3. *A map $f\colon \mathcal{X} \to \mathcal{Y}$ in $\mathsf{Ind}(\mathsf{C}_\Lambda)$ is a* residual isomorphism *if for all $S = \mathrm{Spf}(k) \in \mathsf{C}_\Lambda$ where k is a field, the map $f\colon \mathcal{X}(S) \to \mathcal{Y}(S)$ is a bijection.*

Lemma 5.2.4. *Let $f\colon \mathcal{X} \to \mathcal{Y}$ be a smooth map in $\mathsf{Ind}(\mathsf{C}_\Lambda)$ that is a residual isomorphism. Then f admits a section.*

Proof. By [SGA 3_1, VII_B 0.1.1], it suffices to prove the result when $\mathcal{X} = \mathrm{Spf}(R')$, $\mathcal{Y} = \mathrm{Spf}(R)$, for local Λ-algebras $R \to R'$ with the same residue field. Let $k = R/\mathfrak{m}_R \xrightarrow{\sim} R'/\mathfrak{m}_{R'}$ be their common residue field. From the diagram

$$\begin{array}{ccc} R' & \dashrightarrow & R \\ \uparrow & \searrow\!\!\searrow & \downarrow\!\!\downarrow \\ R & \twoheadrightarrow & k, \end{array}$$

the definition of formal smoothness, and a limiting argument involving the finite length quotients $R/\mathfrak{a}$, we obtain the result. □

Corollary 5.2.5. *Let $\mathcal{R} \to \mathcal{X} \times \mathcal{X}$ be an equivalence relation satisfying the hypotheses of Theorem 5.2.2. Suppose further that*

1. *One of the maps $\mathcal{R} \to \mathcal{X}$ is smooth, and*
2. *The projection $\mathcal{X} \to \mathcal{X}/\mathcal{R}$ is a residual isomorphism.*

Then $\mathcal{X} \to \mathcal{X}/\mathcal{R}$ admits a section, so $\mathcal{X}(S)/\mathcal{R}(S) \xrightarrow{\sim} (\mathcal{X}/\mathcal{R})(S)$ for all $S \in \mathsf{C}_\Lambda$.

Proof. By Lemma 5.2.4, it suffices to prove that $\mathcal{X} \to \mathcal{X}/\mathcal{R}$ is smooth. By [EGA 4_4, 17.7.3(ii)], smoothness can be detected after flat descent. So base-change with respect to the projection $\mathcal{X} \to \mathcal{X}/\mathcal{R}$. In the following commutative diagram

$$\begin{array}{ccc} \mathcal{R} & \rightrightarrows & \\ & \searrow & \mathcal{X} \times_{\mathcal{X}/\mathcal{R}} \mathcal{X} \longrightarrow \mathcal{X} \\ & & \downarrow \qquad\quad \downarrow \\ & \searrow & \mathcal{X} \longrightarrow \mathcal{X}/\mathcal{R} \end{array}$$

we can ensure the smoothness of $\mathcal{R} \to \mathcal{X}$ by our hypotheses. Since $\mathcal{X} \to \mathcal{X}/\mathcal{R}$ is smooth after flat base-change, the original map is smooth. □

Example 5.2.6. The hypothesis on residue fields in Corollary 5.2.5 is necessary. To see this, let $\Lambda = k$ be a field, $k \hookrightarrow K$ a finite Galois extension with Galois group G. Then $G \times \mathrm{Spf}(K) \rightrightarrows \mathrm{Spf}(K)$ has quotient $\mathrm{Spf}(k)$, but the map $\mathrm{Spf}(K)(S) \to \mathrm{Spf}(k)(S)$ is *not* surjective for all $S \in \mathsf{C}_k$, e.g. it is not for $S = \mathrm{Spf}(k)$.

Example 5.2.7. The hypothesis of smoothness in Theorem 5.2.5 is necessary. To see this, let k be a field of characteristic $p > 0$. Then the formal additive group $\widehat{\mathbf{G}_{\mathrm{a}}} = \mathrm{Spf}(k[\![t]\!])$ has a subgroup $\boldsymbol{\alpha}_p(S) = \{s \in \mathscr{O}(S) \colon s^p = 0\}$. The quotient $\widehat{\mathbf{G}_{\mathrm{a}}}/\boldsymbol{\alpha}_p$ has affine coordinate ring $k[\![t^p]\!]$. In particular, the following sequence is exact in the flat topology:

$$0 \longrightarrow \boldsymbol{\alpha}_p \longrightarrow \widehat{\mathbf{G}_{\mathrm{a}}} \xrightarrow{(\cdot)^p} \widehat{\mathbf{G}_{\mathrm{a}}} \longrightarrow 0.$$

It follows that $\boldsymbol{\alpha}_p \times \widehat{\mathbf{G}_{\mathrm{a}}} \rightrightarrows \widehat{\mathbf{G}_{\mathrm{a}}} \xrightarrow{(\cdot)^p} \widehat{\mathbf{G}_{\mathrm{a}}}$ is a coequalizer in $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_k)$ satisfying all the hypotheses of Corollary 5.2.5 except smoothness. And indeed, as one sees by letting $S = \mathrm{Spf}(A)$ for any non-perfect k-algebra A, the map $(\cdot)^p \colon \widehat{\mathbf{G}_{\mathrm{a}}}(S) \to \widehat{\mathbf{G}_{\mathrm{a}}}(S)$ is *not* surjective for all S.

5.3 Deformations of group representations

Here we elaborate on (and correct some mistakes in) the arguments in [BÏ13, §2.1].

Let $\Gamma \in \mathsf{Gp}(\widehat{\mathsf{C}_\Lambda})$ and $G_{/\Lambda}$ be a smooth group scheme of finite type. Write $\widehat{G}$ for the group object in $\mathsf{Ind}(\mathsf{C}_\Lambda)$ given by $\widehat{G}(\mathrm{Spf}\, A) = G(\mathrm{Spec}\, A)$. By Corollary 5.1.8,

the functor

$$\mathrm{Rep}^{\square}(\Gamma, \widehat{G})(S) = \hom_{\mathsf{Gp}/S}(\Gamma_S, \widehat{G}_S) = \hom_{\mathsf{Gp}}(\Gamma(S), G(S))$$

is in $\mathsf{Ind}(\mathsf{C}_\Lambda)$. We would like to define an ind-scheme $\mathrm{Rep}(\Gamma, \mathcal{G})$ as "$\mathrm{Rep}^{\square}(\Gamma, \mathcal{G})$ modulo conjugation," but this requires some care. The conjugation action of $\mathcal{G}$ on $\mathrm{Rep}^{\square}(\Gamma, \mathcal{G})$ will have fixed points, so the quotient will be badly behaved. We loosely follow [Til96, Ch. 2–3] in dealing with this potential problem.

Assume Λ is local, with maximal ideal $\mathfrak{m}$ and residue field k. Fix $\bar{\rho} \in \mathrm{Rep}^{\square}(\Gamma, \widehat{G})(k)$, i.e. a residual representation $\bar{\rho}\colon \Gamma \to G(k)$. Let $\mathrm{Rep}^{\square}(\Gamma, \widehat{G})_{\bar{\rho}}$ be the connected component of $\bar{\rho}$ in $\mathrm{Rep}^{\square}(\Gamma, \widehat{G})$. Assume that G and $\mathrm{Z}(G)$ are smooth; then the quotient $\widehat{G}^{\circ,\mathrm{ad}} = \widehat{G}^{\circ}/\mathrm{Z}(\widehat{G}^{\circ})$ is also smooth, where $(-)^{\circ}$ denotes "connected component of identity." Since $\mathrm{Z}(\widehat{G}^{\circ})$ is smooth, the quotient sheaf $\widehat{G}^{\circ,\mathrm{ad}}$ is the same as the quotient presheaf.

Theorem 5.3.1. *Suppose $(\Lambda, \mathfrak{m}, k)$ is local. If $\mathcal{X}, \mathcal{Y} \in \mathsf{Ind}(\mathsf{C}_\Lambda)$ are connected and $\mathcal{X}(k) \neq \varnothing$, then $\mathcal{X} \times_\Lambda \mathcal{Y}$ is connected.*

Proof. The proof reduces to proving the following result from commutative algebra: if R, S are local pro-artinian Λ-algebras and R has residue field k, then $R\widehat{\otimes}_\Lambda S$ is local. Since $R\widehat{\otimes}_\Lambda S = \varprojlim (R/\mathfrak{r}) \otimes_\Lambda (S/\mathfrak{s})$, $\mathfrak{r}$ (resp. $\mathfrak{s}$) ranging over all open ideals in R (resp. S), we may assume that both R and S are artinian. The rings R and S are henselian, so $R \otimes S$ is local if and only if $(R/\mathfrak{m}_R) \otimes (S/\mathfrak{m}_S) = S/\mathfrak{m}_S$ is local, which it is. $\square$

We conclude that the action of $\widehat{G}^{\circ,\mathrm{ad}}$ on $\mathrm{Rep}^{\square}(\Gamma, \widehat{G})$ preserves $\mathrm{Rep}^{\square}(\Gamma, \mathcal{G})_{\bar{\rho}}$. Thus we may put $\mathrm{Rep}(\Gamma, G)_{\bar{\rho}} = \mathrm{Rep}^{\square}(\Gamma, \widehat{G})_{\bar{\rho}}/\widehat{G}^{\circ,\mathrm{ad}}$. This quotient will only be

well-behaved if $\widehat{G}^{\circ,\mathrm{ad}}$ acts faithfully on $\mathrm{Rep}^{\square}(\Gamma, \widehat{G})$. That faithfulness (or lack thereof) is governed by the action of Γ on $\mathfrak{g}(k)$. Let $\mathfrak{z} = \mathrm{Lie}(\mathrm{Z}(G))$.

Lemma 5.3.2. *Let $A_1 \twoheadrightarrow A_0$ be a square-zero extension of artinian Λ-algebras with kernel I. Let $\rho \in \mathrm{Rep}^{\square}(\Gamma, G)$. Then $\mathrm{Stab}_{\ker(G(A_1)\to G(A_0))}(\rho) = \exp(\mathrm{H}^0(\Gamma, \mathfrak{g}(I))$.*

Proof. First, note that $\ker(G(A_1) \to G(A_0)) = \exp(\mathfrak{g}(I))$. For $X \in \mathfrak{g}(I)$, $X \in \mathrm{Stab}_{\ker(G(A_1)\to G(A_0))}(\rho)$ if and only if $\exp(X)\rho\exp(-X) = \rho$, i.e. $\rho^{-1}\exp(X)\rho = \exp(X)$, which is clearly equivalent to $X \in \mathrm{H}^0(\Gamma, \mathfrak{g}(I))$. □

Theorem 5.3.3. *Let $\Gamma, G, \bar{\rho}$ be as above. Assume $\mathrm{H}^0(\Gamma, \mathfrak{g}(k)) = \mathfrak{z}(k)$. Then $\mathrm{Rep}(\Gamma, G)_{\bar{\rho}}$ exists and, for A a local Λ-algebra, consists of $\widehat{G}^{\circ}(A)$-conjugacy classes of maps $\Gamma \to G(A)$ that reduce to $\bar{\rho}$ modulo $\mathfrak{m}_A$.*

Proof. We begin by proving that if $\mathrm{H}^0(\Gamma, \mathfrak{g}(k)) = \mathfrak{z}(k)$, then $\widehat{G}^{\circ,\mathrm{ad}}$ acts faithfully on $\mathrm{Rep}^{\square}(\Gamma, \widehat{G})$. First, note that if the given $\mathrm{H}^0 = \mathfrak{z}$, then $\mathrm{H}^0(\Gamma, \mathfrak{g}^{\mathrm{ad}}(k)) = 0$, where $\mathfrak{g}^{\mathrm{ad}} = \mathrm{Lie}(\widehat{G}^{\circ,\mathrm{ad}}) = \mathfrak{g}/\mathfrak{z}$. The action of $\widehat{G}^{\circ,\mathrm{ad}}$ is faithful if and only if, whenever A is a local Artinian Λ-algebra and $\rho\colon \Gamma \to G(A)$ agrees with $\bar{\rho}$ modulo $\mathfrak{m}_A$, and for $g \in G(A)$ with $g \equiv 1$ modulo $\mathfrak{m}_A$, then if $g\rho g^{-1} = \rho$, we have $g = 1$. Let $I \subset A$ be a square-zero ideal that is one-dimensional over k. By induction on the length of A, we may assume that $g \equiv 1$ modulo I. Applying Lemma 5.3.2, we see that $g = 1$.

Since the action of $\widehat{G}^{\circ,\mathrm{ad}}$ on $\mathrm{Rep}^{\square}(\Gamma, \widehat{G})_{\bar{\rho}}$ is faithful, we can apply Theorem 5.2.5 to see that $\mathrm{Rep}(\Gamma, G)_{\bar{\rho}}$ is the presheaf quotient, which is exactly what is described in the statement of the theorem. □

5.4 Tangent spaces and obstruction theory

A routine argument shows that if $I \subset A$ is a square-zero ideal, and $\rho\colon \Gamma \to G(A/I)$ admits some lift to $G(A)$, then the set of such lifts form a $\mathrm{H}^1(\Gamma, \mathfrak{g}(I))$-torsor. We show that obstruction theory also works in this more general context.

For $S_0 \in \mathsf{C}_\Lambda$, let Ex_{S_0} be the category of square-zero thickenings of S_0. An object of Ex_{S_0} is a closed embedding $S_0 \hookrightarrow S$ whose ideal of definition has square zero. For any such object, there is an "exponential exact sequence"

$$0 \longrightarrow \mathfrak{g}(I) \longrightarrow G(S) \longrightarrow G(S_0) \longrightarrow 1$$

This gives us a class $\exp \in \mathrm{H}^2(G(S_0), \mathfrak{g}(I))$. For $\rho_0\colon \Gamma \to G(S_0)$, the obstruction class is $o(\rho_0, I) = \rho_0^*(\exp) \in \mathrm{H}^2(\Gamma, \mathfrak{g}(I))$. It's easy to check that $o(\rho_0, I) = 0$ if and only if ρ_0 lifts to ρ. More formally, we use [Wei94, 6.6.4]. Given setting as above, $\rho_0^*(\exp)$ is the pullback by ρ_0:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathfrak{g}(I) & \longrightarrow & \mathcal{G}(S) \times_{\mathcal{G}(S_0)} \Gamma & \longrightarrow & \Gamma & \longrightarrow & 1 \\
 & & \| & & \downarrow & & \downarrow{\scriptstyle \rho_0} & & \\
0 & \longrightarrow & \mathfrak{g}(I) & \longrightarrow & \mathcal{G}(S) & \longrightarrow & \mathcal{G}(S_0) & \longrightarrow & 1.
\end{array}$$

Proposition 5.4.1. *Let $f\colon G \to H$ be a morphism of profinite groups. Suppose M is a discrete H-module and $c \in \mathrm{H}^2(H, M)$ corresponds to the extension*

$$0 \longrightarrow M \longrightarrow \widetilde{H} \longrightarrow H \longrightarrow 1.$$

Then $f^*c = 0$ in $\mathrm{H}^2(G, M)$ if and only if there is a map $\widetilde{f}\colon G \to \widetilde{H}$ making the following diagram commute:

$$\begin{array}{ccc}
 & & \widetilde{H} \\
 & \overset{\widetilde{f}}{\nearrow} & \downarrow \\
G & & \\
 & \underset{f}{\searrow} & \\
 & & H.
\end{array}$$

Proof. By [Wei94, 6.6.4], the class f^*c corresponds to the pullback diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M & \longrightarrow & G \times_H \widetilde{H} & \longrightarrow & G & \longrightarrow & 1 \\ & & \| & & \downarrow & & \downarrow{\scriptstyle f} & & \\ 0 & \longrightarrow & M & \longrightarrow & \widetilde{H} & \longrightarrow & H & \longrightarrow & 1. \end{array}$$

Writing explicitly what it means for $G \times_H \widetilde{H} \to G$ to split yields the result. □

Corollary 5.4.2. *Let $\rho \in \mathrm{Rep}(\Gamma, G)_{\bar{\rho}}(A/I)$, where I is a square-zero ideal. Then ρ lifts to A if and only if $\rho^*(\exp_I) = 0$ in the group $\mathrm{H}^2(\Gamma, \mathfrak{g}(I))$.*

CHAPTER 6

CONSTRUCTING GALOIS REPRESENTATIONS

6.1 Notation and necessary results

In this section we loosely summarize and adapt the results of [KLR05; Pan11]. Throughout, if F is a field and M a G_F-module, we write $\mathrm{H}^\bullet(F, M)$ in place of $\mathrm{H}^\bullet(G_F, M)$. All Galois representations will take values in $\mathrm{GL}_2(\mathbf{Z}/l^n)$ or $\mathrm{GL}_2(\mathbf{Z}_l)$ for l a (fixed) rational prime, and all deformations will have fixed determinant. So we consider the cohomology of $\mathrm{Ad}^0\,\bar{\rho}$, the induced representation on trace-zero matrices by conjugation.

If S is a set of rational primes, $\mathbf{Q}_S$ denotes the largest extension of $\mathbf{Q}$ unramified outside S. So $\mathrm{H}^i(\mathbf{Q}_S, -)$ is what is usually written as $\mathrm{H}^1(G_{\mathbf{Q},S}, -)$. If M is a $G_{\mathbf{Q}}$-module and S a finite set of primes, write

$$Ш^i_S(M) = \ker\left(\mathrm{H}^i(\mathbf{Q}_S, M) \to \prod_{p\in S} \mathrm{H}^i(\mathbf{Q}_p, M)\right).$$

If l is a rational prime and S a finite set of primes containing l, then for any $\mathbf{F}_l[G_{\mathbf{Q}_S}]$-module M, write $M^\vee = \hom_{\mathbf{F}_l}(M, \mathbf{F}_l)$ with the obvious $G_{\mathbf{Q}_S}$-action, and write $M^* = M^\vee(1)$ for the Cartier dual of $M^\vee$. By [NSW08, Th. 8.6.7], there is an isomorphism $Ш^1_S(M^*) = Ш^2_S(M)^\vee$. As a result, if $Ш^1_S(M) = Ш^2_S(M) = 0$ and $S \subset T$, then $Ш^1_T(M) = Ш^2_T(M) = 0$.

Definition 6.1.1. *A* good residual representation *is an odd, absolutely irreducible, weight-2 representation $\bar{\rho}\colon G_{\mathbf{Q}_S} \to \mathrm{GL}_2(\mathbf{F}_l)$, where $l \geqslant 7$ is a rational prime.*

Roughly, "good residual representations" have enough properties that we can

prove meaningful theorems about their lifts without assuming the modularity results of Khare–Wintenberger.

Theorem 6.1.2. *Let $\bar{\rho}\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_l)$ be a good residual representation. Then there exists a weight-2 lift of $\bar{\rho}$ to $\mathbf{Z}_l$, ramified at the same set of primes as $\bar{\rho}$.*

Proof. This is [Ram02, Th. 1], taking into account that the paper in question allows for arbitrary fixed determinants. □

Definition 6.1.3. *Let $\bar{\rho}\colon G_{\mathbf{Q}_S} \to \mathrm{GL}_2(\mathbf{F}_l)$ be a good residual representation. A prime $p \not\equiv \pm 1 \pmod{l}$ is* nice *if $\mathrm{Ad}^0\,\bar{\rho} \simeq \mathbf{F}_l \oplus \mathbf{F}_l(1) \oplus \mathbf{F}_l(-1)$, i.e. if the eigenvalues of $\bar{\rho}(\mathrm{fr}_p)$ have ratio p.*

Theorem 6.1.4 (Ramakrishna)**.** *Let $\bar{\rho}$ be a good residual representation and p a nice prime. Then any deformation of $\bar{\rho}|_{G_{\mathbf{Q}_p}}$ is induced by $G_{\mathbf{Q}_p} \to \mathrm{GL}_2(\mathbf{Z}_l[\![a,b]\!]/\langle ab\rangle)$, sending*

$$\mathrm{fr}_p \mapsto \begin{pmatrix} p(1+a) & \\ & (1+a)^{-1} \end{pmatrix} \qquad \tau_p \mapsto \begin{pmatrix} 1 & b \\ & 1 \end{pmatrix},$$

where $\tau_p \in G_{\mathbf{Q}_p}$ is a generator for tame inertia.

We close this section by introducing some new terminology and notation to condense the lifting process used in [KLR05].

Fix a good residual representation $\bar{\rho}$. We will consider weight-2 deformations of $\bar{\rho}$ to $\mathbf{Z}/l^n$ and $\mathbf{Z}_l$. Call such a deformation a "lift of $\bar{\rho}$ to $\mathbf{Z}/l^n$ (resp. $\mathbf{Z}_l$)." We will often restrict the local behavior of such lifts, i.e. the restrictions of a lift to $G_{\mathbf{Q}_p}$ for p in some set of primes. The necessary constraints are captured in the following definition.

Definition 6.1.5. *Let $\bar{\rho}$ be a good residual representation, $h\colon \mathbf{R}^+ \to \mathbf{R}^+$ a function decreasing to zero. An* h-bounded lifting datum *is a tuple $(\rho_n, R, U, \{\rho_p\}_{p\in R\cup U})$, where*

1. *$\rho_n\colon G_{\mathbf{Q}_R} \to \mathrm{GL}_2(\mathbf{Z}/l^n)$ is a lift of $\bar{\rho}$.*
2. *R and U are finite sets of primes, R containing l and all primes at which ρ_n ramifies.*
3. *$\pi_R(x) \leqslant h(x)\pi(x)$ for all x.*
4. *$Ш^1_R(\mathrm{Ad}^0\,\bar{\rho}) = Ш^2_R(\mathrm{Ad}^0\,\bar{\rho}) = 0$.*
5. *For all $p \in R\cup U$, $\rho_p \equiv \rho_n|_{G_{\mathbf{Q}_p}} \pmod{l^n}$.*
6. *For all $p \in R$, ρ_p is ramified.*
7. *ρ_n admits a lift to $\mathbf{Z}/l^{n+1}$.*

If $(\rho_n, R, U, \{\rho_p\})$ is an h-bounded lifting datum, we call another h-bounded lifting datum $(\rho_{n+1}, R', U', \{\rho_p\})$ a *lift of* $(\rho_n, R, U, \{\rho_p\})$ if $U \subset U'$, $R \subset R'$, and for all $p \in R\cup U$, the two possible ρ_p agree.

Theorem 6.1.6. *Let $\bar{\rho}$ be a good residual representation, $h\colon \mathbf{R}^+ \to \mathbf{R}^+$ decreasing to zero. If $(\rho_n, R, U, \{\rho_p\})$ is an h-bounded lifting datum, $U' \supset U$ is a finite set of primes disjoint from R, and $\{\rho_p\}_{p\in U'}$ extends $\{\rho_p\}_{p\in U}$, then there exists an h-bounded lift $(\rho_{n+1}, R', U', \{\rho_p\})$ of $(\rho_n, R, U, \{\rho_p\})$.*

Proof. By [KLR05, Lem. 8], there exists a finite set N of nice primes such that the map

$$\mathrm{H}^1(\mathbf{Q}_{R\cup N}, \mathrm{Ad}^0\,\bar{\rho}) \to \prod_{p\in R} \mathrm{H}^1(\mathbf{Q}_p, \mathrm{Ad}^0\,\bar{\rho}) \times \prod_{p\in U'} \mathrm{H}^1_{\mathrm{nr}}(\mathbf{Q}_p, \mathrm{Ad}^0\,\bar{\rho}) \tag{6.1}$$

is an isomorphism. In fact, $\#N = \dim \mathrm{H}^1(\mathbf{Q}_{R\cup N}, \mathrm{Ad}^0\,\bar{\rho}^*)$, and the primes in N are chosen, one at a time, from Chebotarev sets. This means we can force them to be large enough to ensure that the bound $\pi_{R\cup N}(x) \leqslant h(x)\pi(x)$ continues to hold. We also choose the primes in N to be larger than any prime in U'.

By our hypothesis, ρ_n admits a lift to $\mathbf{Z}/l^{n+1}$; call one such lift ρ^*. For each $p \in R \cup U'$, $\mathrm{H}^1(\mathbf{Q}_p, \mathrm{Ad}^0\,\bar{\rho})$ acts simply transitively on lifts of $\rho_n|_{G_{\mathbf{Q}_p}}$ to $\mathbf{Z}/l^{n+1}$. In particular, there are cohomology classes $f_p \in \mathrm{H}^1(\mathbf{Q}_p, \mathrm{Ad}^0\,\bar{\rho})$ such that $f_p \cdot \rho^* \equiv \rho_p \pmod{l^{n+1}}$ for all $p \in R \cup U'$. Moreover, for all $p \in U'$, the class f_p is unramified. Since the map (6.1) is an isomorphism, there exists $f \in \mathrm{H}^1(\mathbf{Q}_{R\cup N}, \mathrm{Ad}^0\,\bar{\rho})$ such that $f \cdot \rho^*|_{G_{\mathbf{Q}_p}} \equiv \rho_p \pmod{l^{n+1}}$ for all $p \in R \cup U'$.

Clearly $f \cdot \rho^*|_{G_{\mathbf{Q}_p}}$ admits a lift to $\mathbf{Z}_l$ for all $p \in R\cup U'$, but it does not necessarily admit such a lift for $p \in N$. By repeated applications of [Pan11, Prop. 3.10], there exists a set $N' \supset N$, with $\#N' \leqslant 2\#N$, of nice primes and $g \in \mathrm{H}^1(\mathbf{Q}_{R\cup N'}, \mathrm{Ad}^0\,\bar{\rho})$ such that $(g+f)\cdot\rho^*$ still agrees with ρ_p for $p \in R \cup U'$, and $(g+f)\cdot\rho^*$ is nice for all $p \in N'$. As above, the primes in N' are chosen one at a time from Chebotarev sets, so we can continue to ensure the bound $\pi_{R\cup N'}(x) \leqslant h(x)\pi(x)$ and also that all primes in N' are larger than those in U'. Let $\rho_{n+1} = (g+f)\cdot\rho^*$. Let $R' = R \cup \{p \in N' : \rho_{n+1} \text{ is ramified at } p\}$. For each $p \in R' \smallsetminus R$, choose a lift ρ_p of $\rho_{n+1}|_{G_{\mathbf{Q}_p}}$ to $\mathbf{Z}_l$.

Since $\rho_{n+1}|_{G_{\mathbf{Q}_p}}$ admits a lift to $\mathbf{Z}/l^{n+2}$ (in fact, it admits a lift to $\mathbf{Z}_l$) for each p, and $Ш^1_{R'}(\mathrm{Ad}^0\,\bar{\rho}) = Ш^2_{R'}(\mathrm{Ad}^0\,\bar{\rho}) = 0$, the deformation ρ_{n+1} admits a lift to $\mathbf{Z}/l^{n+2}$. Thus $(\rho_{n+1}, R', U', \{\rho_p\})$ is the desired lift of $(\rho_n, R, U, \{\rho_p\})$. □

6.2 Galois representations with specified Satake parameters

Fix a good residual representation $\bar{\rho}$, and consider weight-2 deformations of $\bar{\rho}$. The final deformation, $\rho\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$, will be constructed as the inverse limit of a compatible collection of lifts $\rho_n\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}/l^n)$. At any given stage, we will be concerned with making sure that there exists a lift to the next stage, and that there is a lift with the necessary properties. Fix a sequence $\boldsymbol{x} = (x_1, x_2, \dots)$ in $[-1, 1]$. The set of unramified primes of ρ is not determined at the beginning, but at each stage there will be a large finite set U of primes which we know will remain unramified. Reindexing $\boldsymbol{x}$ by these unramified primes, we will construct ρ so that for all unramified primes p, $\operatorname{tr}\rho(\mathrm{fr}_p) \in \mathbf{Z}$, satisfies the Hasse bound, and has $\operatorname{tr}\rho(\mathrm{fr}_p) \approx x_p$. Moreover, we can ensure that the set of ramified primes has density zero in a very strong sense (controlled by a parameter function h) and that our trace of Frobenii are very close to specified values (the "closeness" again controlled by a parameter function b).

Given any deformation ρ, write $\pi_{\mathrm{ram}(\rho)}(x)$ for the function which counts ρ_n-ramified primes $\leqslant x$.

Theorem 6.2.1. *Let l, $\bar{\rho}$, $\boldsymbol{x}$ be as above. Fix functions $h\colon \mathbf{R}^+ \to \mathbf{R}^+$ (resp. $b\colon \mathbf{R}^+ \to \mathbf{R}_{\geqslant 1}$) which decrease to zero (resp. increase to infinity). Then there exists a weight-2 deformation ρ of $\bar{\rho}$, such that:*

1. *$\pi_{\mathrm{ram}(\rho)}(x) \ll h(x)\pi(x)$.*

2. *For each unramified prime p, $a_p = \operatorname{tr}\rho(\mathrm{fr}_p) \in \mathbf{Z}$ and satisfies the Hasse*

bound.

3. *For each unramified prime p,* $\left|\frac{a_p}{2\sqrt{p}} - x_p\right| \leqslant \frac{lb(p)}{2\sqrt{p}}$.

Proof. Begin with $\rho_1 = \bar{\rho}$. By [KLR05, Lem. 6], there exists a finite set R, containing the set of primes at which $\bar{\rho}$ ramifies, such that $Ш^1_R(\mathrm{Ad}^0\,\bar{\rho}) = Ш^2_R(\mathrm{Ad}^0\,\bar{\rho}) = 0$. Let R_1 be the union of R and all primes p with $\frac{l}{2\sqrt{p}} > 2$. For all $p \notin R_1$ and any $a \in \mathbf{F}_l$, there exists $a_p \in \mathbf{Z}$ satisfying the Hasse bound with $a_p \equiv a \pmod{l}$. In fact, given any $x_p \in [-1, 1]$, there exists $a_p \in \mathbf{Z}$ satisfying the Hasse bound such that $\left|\frac{a_p}{2\sqrt{p}} - x_p\right| \leqslant \frac{l}{2\sqrt{p}}$. Choose, for all primes $p \in R_1$, a ramified lift ρ_p of $\rho_1|_{G_{\mathbf{Q}_p}}$. Let U_1 be the set of primes not in R_1 such that $\frac{l^2}{2\sqrt{p}} > \min\left(2, \frac{lb(p)}{2\sqrt{p}}\right)$. For each $p \in U_1$, there exists $a_p \in \mathbf{Z}$, satisfying the Hasse bound, such that

$$\left|\frac{a_p}{2\sqrt{p}} - x_p\right| \leqslant \frac{l}{2\sqrt{p}} \leqslant \frac{lb(p)}{2\sqrt{p}},$$

and moreover $a_p \equiv \mathrm{tr}\,\bar{\rho}(\mathrm{fr}_p) \pmod{l}$. For each $p \in U_1$, let ρ_p be an unramified lift of $\bar{\rho}|_{G_{\mathbf{Q}_p}}$ with $\mathrm{tr}\,\rho_p$ being the desired a_p. It may not be that $\pi_{R_1}(x) \leqslant h(x)\pi(x)$ for all x, but there is a scalar multiple h^* of h so that $\pi_{R_1}(x) \leqslant h^*(x)\pi(x)$ for all x.

We have constructed our first h^*-bounded lifting datum $(\rho_1, R_1, U_1, \{\rho_p\})$. We proceed to construct $\rho = \varprojlim \rho_n$ inductively, by constructing a new h^*-bounded lifting datum for each n. We ensure that U_n contains all primes for which $\frac{l^{n+1}}{2\sqrt{p}} > \min\left(2, \frac{lb(p)}{2\sqrt{p}}\right)$, so there are always integral a_p satisfying the Hasse bound which satisfy any mod-l^{n+1} constraint, and that can always choose these a_p so as to preserve statement 2 in the theorem.

The base case is already complete, so suppose we are given $(\rho_{n-1}, R_{n-1}, U_{n-1}, \{\rho_p\})$. We may assume that U_{n-1} contains all primes for which $\frac{l^n}{2\sqrt{p}} > \min\left(2, \frac{lb(p)}{2\sqrt{p}}\right)$. Let

U_n be the set of all primes not in R_{n-1} such that $\frac{l^{n+1}}{2\sqrt{p}} > \min\left(2, \frac{lb(p)}{2\sqrt{p}}\right)$. For each $p \in U_n \smallsetminus U_{n-1}$, there is an integer a_p, satisfying the Hasse bound, such that $a_p \equiv \rho_n(\mathrm{fr}_p) \pmod{l^n}$, and moreover $\left|\frac{a_p}{2\sqrt{p}} - x_p\right| \leqslant \frac{lb(p)}{2\sqrt{p}}$. For such p, let ρ_p be an unramified lift of $\rho_n|_{G_{\mathbf{Q}_p}}$ such that $\operatorname{tr}\rho_n(\mathrm{fr}_p)$ is the desired a_p. By Theorem 6.1.6, there exists an h^*-bounded lifting datum $(\rho_n, R_n, U_n, \{\rho_p\})$ extending and lifting $(\rho_{n-1}, R_{n-1}, U_{n-1}, \{\rho_p\})$. This completes the inductive step. □

We will apply this theorem to construct Galois representations with specified Sato–Tate distributions in the next section, but for now here is a small consequence, which addresses the results in [Sar07]. Sarnak remarks that for $E_{/\mathbf{Q}}$ a non-CM elliptic curve with rank r, the partial sums $\frac{\log x}{\sqrt{x}} \sum_{p \leqslant x} \frac{a_p}{\sqrt{p}}$ approach a limiting distribution with mean $1 - 2r$.

Corollary 6.2.2. *Let $L \in [-\infty, \infty]$ and $\epsilon > 0$ be given. Then there exists a weight 2 Galois representation $\rho\colon G \to \mathrm{GL}_2(\mathbf{Z}_l)$, such that each $a_p = \operatorname{tr}\rho(\mathrm{fr}_p) \in \mathbf{Z}$ satisfies the Hasse bound,*

$$L = \lim_{N\to\infty} \frac{\log N}{\sqrt{N}} \sum_p \frac{a_p}{\sqrt{p}}$$

and $\pi_{\mathrm{ram}(\rho)}(x) \ll e^{-x}$.

Proof. Begin with a sequence (x_p) in $[-1, 1]$ such that $\lim_{N\to\infty} \frac{\log N}{\sqrt{N}} \sum_{p \leqslant N} x_p = L$. If $L = \pm\infty$, we can choose $x_p = \pm 1$. By Theorem 6.2.1, there exists $\rho\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$ with $\pi_{\mathrm{ram}(\rho)}(x) \ll e^{-x}$, and such that for each unramified p, $a_p = \operatorname{tr}\rho(\mathrm{fr}_p) \in \mathbf{Z}$, satisfies the Hasse bound, has $\left|\frac{a_p}{2\sqrt{p}} - b_p\right| < \frac{l \log p}{\sqrt{p}}$, and by looking at the proof, we can even ensure that $\sum \left(\frac{a_p}{2\sqrt{p}} - b_p\right)$ converges conditionally.

Note that

$$\left|\frac{\log N}{\sqrt{N}}\sum_{p\leqslant N}\frac{a_p}{2\sqrt{p}}-\frac{\log N}{\sqrt{N}}\sum_{p\leqslant N}x_p\right|\leqslant\frac{\log N}{\sqrt{N}}\left(\pi_{\mathrm{ram}(\rho)}(N)+\sum_{p\leqslant N}\left(\frac{a_p}{2\sqrt{p}}-x_p\right)\right)\to 0.$$

When $L\neq\pm\infty$, this shows that the limit in question exists and is L. When $L=\pm\infty$, this shows that the the sums in question diverge to L. □

6.3 Arbitrary Sato–Tate distributions

For $k\geqslant 1$, let

$$U_k(\theta)=\operatorname{tr}\operatorname{sym}^k\left(\begin{smallmatrix}e^{i\theta}&\\&e^{-i\theta}\end{smallmatrix}\right)=\frac{\sin((k+1)\theta)}{\sin\theta}.$$

Then $U_k(\cos^{-1}t)$ is the k-th Chebyshev polynomial of the 2nd kind. Moreover, $\{1\}\cup\{U_k\}$ forms an orthonormal basis for $L^2([0,\pi],\mathrm{ST})=L^2(\mathrm{SU}(2)^\natural)$.

This section has two parts. First, for any reasonable measure μ on $[0,\pi]$ invariant under the same "flip" automorphism as the Sato–Tate measure, there is a sequence (a_p) of integers satisfying the Hasse bound $|a_p|\leqslant 2\sqrt{p}$, such that for $\theta_p=\cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$, the discrepancy $\mathrm{D}_N(\boldsymbol{\theta},\mu)$ behaves like $\pi(N)^{-\alpha}$ for predetermined $\alpha\in\left(0,\frac{1}{2}\right)$, while for any odd k, the strange Dirichlet series $L_{U_k}(\boldsymbol{\theta},s)$, which we will write as $L(\operatorname{sym}^k\boldsymbol{\theta},s)$, satisfies the Riemann Hypothesis. In the second part of this section, we associate Galois representations to these fake Satake parameters.

Definition 6.3.1. *Let $\mu=f(t)\,\mathrm{d}t$ be an absolutely continuous measure with continuous cdf on $[0,\pi]$. If $f(t)\ll\sin(t)$, then μ is a* Sato–Tate compatible measure.

The key facts about Sato–Tate compatible measures are that $\cos_*\mu$ satisfies the hypotheses of Theorem 2.6.5, so there are "$N^{-\alpha}$-decaying van der Corput se-

quences" for $\cos_* \mu$, and also that since $\cos\colon [0,\pi] \to [-1,1]$ is an strictly decreasing, we know that for any sequence $\boldsymbol{x}$ on $[-1,1]$, $\mathrm{D}_N(\boldsymbol{x}, \cos_* \mu) \approx \mathrm{D}_N(\cos^{-1} \boldsymbol{x}, \mu)$.

Theorem 6.3.2. *Let μ be a Sato–Tate compatible measure, and fix $\alpha \in \left(0, \frac{1}{2}\right)$. Then there exists a sequence of integers a_p satisfying the Hasse bound, such that if we set $\theta_p = \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$, then $\mathrm{D}_N(\boldsymbol{\theta}, \mu) = \Theta(\pi(N)^{-\alpha})$.*

Proof. Apply Theorem 2.6.5 to find a sequence $\boldsymbol{x}$ such that $\mathrm{D}_N(\boldsymbol{x}, \cos_* \mu) = \Theta(\pi(N)^{-\alpha})$. For each prime p, there exists an integer a_p such that $|a_p| \leqslant 2\sqrt{p}$ and $\left|\frac{a_p}{2\sqrt{p}} - x_p\right| \leqslant p^{-1/2}$. Let $y_p = \frac{a_p}{2\sqrt{p}}$, and apply Lemma 2.5.1 with $\epsilon = N^{-1/2}$. We obtain

$$|\mathrm{D}_N(\boldsymbol{x}, \cos_* \mu) - \mathrm{D}_N(\boldsymbol{y}, \cos_* \mu)| \ll N^{-1/2} + \frac{\pi(N^{1/2})}{\pi(N)},$$

which tells us that $\mathrm{D}_N(\boldsymbol{y}, \cos_* \mu) = \Theta(\pi(N)^{-\alpha})$. Now let $\boldsymbol{\theta} = \cos^{-1}(\boldsymbol{y})$. Apply Lemma 2.5.3 to $\boldsymbol{\theta} = \cos^{-1}(\boldsymbol{y})$, and we see that $\mathrm{D}_N(\boldsymbol{\theta}, \mu) = \Theta(\pi(N)^{-\alpha})$. □

We can improve this example by controlling the behavior of the sums $\sum_{p \leqslant N} U_k(\theta_p)$ for odd k. Let σ be the involution of $[0,\pi]$ given by $\sigma(\theta) = \pi - \theta$. Note that $\sigma_* \mathrm{ST} = \mathrm{ST}$. Moreover, note that for any odd k, $U_k \circ \sigma = -U_k$, so $\int U_k \,\mathrm{dST} = 0$. Of course, $\int U_k \,\mathrm{dST} = 0$ for the reason that U_k is the trace of a non-trivial unitary representation, but we will directly use the "oddness" of U_k in what follows.

Theorem 6.3.3. *Let μ be a σ-invariant Sato–Tate compatible measure. Fix $\alpha \in \left(0, \frac{1}{2}\right)$. Then there is a sequence of integers a_p, satisfying the Hasse bound, such that for $\theta_p = \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$, we have*

1. $\mathrm{D}_N(\boldsymbol{\theta}, \mu) = \Theta(\pi(N)^{-\alpha})$.

2. *For all odd k, $\left|\sum_{k\leqslant N} U_k(\theta_p)\right| \ll \pi(N)^{1/2}$.*

Proof. The basic ideas is as follows. Enumerate the primes

$$p_1 = 2, q_1 = 3, p_2 = 5, q_2 = 7, p_3 = 11, q_3 = 13, \ldots.$$

Consider the measure $\mu|_{[0,\pi/2)}$. An argument nearly identical to the proof of Theorem 6.3.2 shows that we can choose a_{p_i} satisfying the Hasse bound so that

$$\mathrm{D}_N\left(\{\theta_{p_i}\}, \mu|_{[0,\pi/2)}\right) = \Theta(N^{-\alpha}).$$

We can also choose the $a_{q_i} \in [\pi/2, \pi]$ so that $\left|\frac{a_{p_i}}{2\sqrt{p_i}} + \frac{a_{q_i}}{2\sqrt{q_i}}\right| \ll \frac{1}{\sqrt{p_i}}$. If $\boldsymbol{x}$ is the sequence of the $\frac{a_{p_i}}{2\sqrt{p_i}}$ and $\boldsymbol{y}$ is the corresponding sequence with the q_i-s, then Lemma 2.5.2, Lemma 2.5.1, and Theorem 2.5.5 tell us that $\mathrm{D}_N(\boldsymbol{x} \wr \boldsymbol{y}, \mu) = \Theta(N^{-\alpha})$.

Moreover, $U_k(\cos^{-1} t)$ is an odd polynomial in t, so if $|x_i - (-y_i)| \ll p_i^{-1/2}$, then $|U_k(\theta_{p_i}) + U_k(\theta_{q_i})| \ll p_i^{-1/2}$. We can then bound

$$\left|\sum_{i\leqslant N} (U_k(\theta_{p_i}) + U_k(\theta_{q_i}))\right| \ll \sum_{p\leqslant N} p^{-1/2} \ll \pi(N)^{1/2}.$$

□

Theorem 6.3.4. *Let μ be a Sato–Tate compatible σ-invariant measure on $[0, \pi]$. Fix $\alpha \in \left(0, \frac{1}{2}\right)$ and a good residual representation $\bar{\rho}\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_l)$. Then there exists a weight-2 lift $\rho\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$ of $\bar{\rho}$ such that*

1. *$\pi_{\mathrm{ram}(\rho)}(x) \ll e^{-x}\pi(x)$.*
2. *For each unramified prime p, $a_p = \operatorname{tr}\rho(\mathrm{fr}_p) \in \mathbf{Z}$ and satisfies the Hasse bound.*
3. *If, for unramified p we set $\theta_p = \cos^{-1}\left(\frac{a_p}{2\sqrt{p}}\right)$, then $\mathrm{D}_N(\boldsymbol{\theta}, \mu) = \Theta(\pi(N)^{-\alpha})$.*

4. For each odd k, the function $L(\mathrm{sym}^k \rho, s)$ satisfies the Riemann Hypothesis.

Proof. Let $\boldsymbol{x}$ be an $N^{-\alpha}$-decay van der Corput sequence for $\cos_* \mu|_{[0,\pi/2)}$. Let $\boldsymbol{y} = -\boldsymbol{x}$. Then $\mathrm{D}_N(\boldsymbol{x} \wr \boldsymbol{y}, \cos_* \mu) = \Theta(N^{-\alpha})$. Set $h(x) = e^{-x}$ and $b(x) = \log(x)$. By Theorem 6.2.1, there is a $\rho\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$ lifting $\bar{\rho}$ such that parts 1 and 2 of the theorem hold. The discrepancy estimate comes from Lemma 2.5.2, Lemma 2.5.1, and Theorem 2.5.5 as above, while the Riemann Hypothesis for odd symmetric powers follows from the proof of Theorem 6.3.3. □

This entire discussion also works with absolutely continuous measure μ, supported on a proper subinterval of $[0, \pi]$, so long as their cdf is strictly increasing on that interval. For example, fix $\epsilon > 0$, and let $\mu_\epsilon = B(t)\,\mathrm{d}t$, where B is a bump function on $\left[\frac{\pi}{2} - \epsilon, \frac{\pi}{2} + \epsilon\right]$, normalized to have total mass one. Then Theorem 6.3.4 gives Galois representations with empirical Sato–Tate distribution converging at any specified rate to μ_ϵ. This is a strictly stronger result than [Pan11, Th. 5.2].

CHAPTER 7

CONCLUDING REMARKS AND FUTURE DIRECTIONS

7.1 Fake modular forms

The Galois representations of Theorem 6.3.4 have "fake modular forms" associated to them. Namely, there is a representation of $\mathrm{GL}_2(\mathbf{A})$ with the specified Satake parameters at each prime (for now, set $\theta_p = 0$ at ramified primes). It is natural to ask if these "fake modular forms" have any interesting properties. For example, we know that all their odd symmetric powers satisfy the Riemann Hypothesis. The author is unaware of any further results (say about analytic continuation or functional equation) concerning these fake modular forms.

7.2 Dense free subgroups of compact semisimple groups

Let G be a compact semisimple Lie group, for example SU(2). By [BG03], G contains a dense free subgroup $\Gamma = \langle \gamma_1, \gamma_2 \rangle$. We will now follow the argument of [AK63] to hint at how Γ may yield equidistributed sequences with "bad" discrepancy and small character sums.

Given an integer N, let B_N be the "closed ball of size N" in Γ, that is the set of products $\gamma_{\sigma(1)} \dots \gamma_{\sigma(n)}$, where $n \leqslant N$ and $\sigma\colon \{1, \dots, n\} \to \{1, 2\}$ is a function. We will write $\sigma\colon [n] \to [2]$ in this case. Given an irreducible unitary representation $\rho \in \widehat{G}$, we wish to control the behavior of $\sum_{\gamma \in B_N} \operatorname{tr} \rho(\gamma)$, ideally to show an

estimate of the form

$$\left|\sum_{\gamma\in B_N} \operatorname{tr}\rho(\gamma)\right| \ll (\#B_N)^{\frac{1}{2}+\epsilon}.$$

In fact, $\#B_N = \sum_{n=0}^{N} 2^n = 2^{N+1}-1$. We can encode these sums in terms of convolutions of a measure as follows. Let μ be the measure $\delta_{\gamma_1^{-1}} + \delta_{\gamma_2^{-1}}$ on G. If ρ is any unitary representation (not necessarily irreducible or even finite-dimensional) then μ acts on ρ via $\rho(\mu)\int \rho\,\mathrm{d}\mu$. So, if $\rho = L^2(G)$ via the left regular representation, then $(\mu\cdot f)(x) = f(\gamma_1 x) + f(\gamma_2 x)$, while if $\rho\in\widehat{G}$ and $v\in\rho$, then $\mu\cdot v = \rho(\gamma_1)v + \rho(\gamma_2)v$. Note that

$$\mu^{*n} = \sum_{\sigma\colon [n]\to[2]} \delta_{\gamma_{\sigma(1)}\cdots\gamma_{\sigma(n)}}.$$

This tells us that $\sum_{\gamma\in B_N} f(\gamma) = \sum_{n\leqslant N}\mu^{*n}(f)$. So we really only need to study how μ and its powers act on the functions $\operatorname{tr}\rho$, $\rho\in\widehat{G}$.

First note that $\operatorname{tr}\rho$ generates a subrepresentation of $L^2(G)$ which is isomorphic to ρ. On that representation, we claim that μ is invertible, hence $\sum_{n=0}^{N}\mu^{*n} = (\mu^{*(N+1)}-1)(\mu-1)^{-1}$.It follows that $\|\sum_{n=0}^{N}\mu^{*n}\| \leqslant \frac{\|\mu\|^{N+1}}{\|\mu-1\|}$,

Note that $\|\mu\|^{N+1}\leqslant 2^{(N+1)\alpha}$ if and only if $\|\mu\|\leqslant 2^{\alpha}$. In other words, to get the Riemann Hypothesis for L-functions coming from Γ, we need $\|\mu\|\leqslant\sqrt{2}$. If $v\in\rho$ has norm 1, then

$$\begin{aligned}\|\rho(\mu)v\|^2 &= \langle\rho(\gamma_1^{-1})v + \rho(\gamma_2^{-1})v, \rho(\gamma_1^{-1})v + \rho(\gamma_2^{-1})v\rangle\\ &= 2\|v\|^2 + 2\Re\langle\rho(\gamma_2\gamma_1^{-1})v, v\rangle.\end{aligned}$$

So, we want $\Re\langle\rho(\gamma_2\gamma_1^{-1})v, v\rangle\leqslant 0$ for all irreducible ρ. Sadly, even for SU(2), this is not possible.

Write $\gamma = \gamma_2\gamma_1^{-1}$, then the identity $\langle\rho(\gamma)\rho(\delta)v, \rho(\delta)v\rangle = \langle\rho(\delta^{-1}\gamma\delta)v, v\rangle$ tells us

that we can restrict our search to γ of the form $\left(\begin{smallmatrix} a & \\ & \overline{a} \end{smallmatrix}\right)$ with $|a| = 1$. Now

$$\left\langle \left(\begin{smallmatrix} a & \\ & \overline{a} \end{smallmatrix}\right) \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right), \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) \right\rangle = \Re(a),$$

which appears to be promising. But a similar computation with sym^2 shows that one can always get $\langle \mathrm{sym}^2 \gamma v, v \rangle = 1$, so the above approach fails.

There may be alternative ways of bounding the sums $\sum \mu^{*n}(\operatorname{tr} \rho)$, but we do not investigate them here.

BIBLIOGRAPHY

[AT99] Shigeki Akiyama and Yoshio Tanigawa. "Calculation of values of L-functions associated to elliptic curves". In: *Math. Comp.* 68.227 (1999), pp. 1201–1231.

[Apo76] Tom M. Apostol. *Introduction to analytic number theory*. Undergraduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1976.

[AK63] V. I. Arnol'd and A. L. Krylov. "Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain". In: *Dokl. Akad. Nauk SSSR* 148 (1963), pp. 9–12.

[BÏ3] Gebhard Böckle. "Deformations of Galois representations". In: *Elliptic curves, Hilbert modular forms and Galois deformations*. Adv. Courses Math. CRM Barcelona. Birkhäuser/Springer, Basel, 2013, pp. 21–115.

[BG03] E. Breuillard and T. Gelander. "On dense free subgroups of Lie groups". In: *J. Algebra* 261.2 (2003), pp. 448–467.

[BK15] Alina Bucar and Kiran Kedlaya. *An application of the effective Sato–Tate conjecture*. 2015. eprint: `arXiv:1301.0139`.

[CHT08] Laurent Clozel, Michael Harris, and Richard Taylor. "Automorphy for some l-adic lifts of automorphic mod l Galois representations". In: *Publ. Math. Inst. Hautes Études Sci.* 108 (2008). With Appendix A, summarizing unpublished work of Russ Mann, and Appendix B by Marie-France Vignéras, pp. 1–181.

[DT97] Michael Drmota and Robert F. Tichy. *Sequences, discrepancies and applications*. Vol. 1651. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1997.

[EGA 4_4] Alexandre Grothendieck. *Éléments de gómétrie algébrique. IV. Étude locale des schémas et des morphismes de schémas IV*. 32. 1967.

[SGA 3_1] Alexandre Grothendieck and Michel Demazure, eds. *Schémas en groupes (SGA 3). Tome I. Propriétés générales des schémas en groupes*. Vol. 151. Lecture Notes in Mathematics. Springer-Verlag, 1970.

[HSBT10] Michael Harris, Nick Shepherd-Barron, and Richard Taylor. "A family of Calabi-Yau varieties and potential automorphy". In: *Ann. of Math. (2)* 171.2 (2010), pp. 779–813.

[Joh02] Peter Johnstone. *Sketches of an elephant: a topos theory compendium*. Vol. 44, 45. Oxford Logic Guides. Oxford University Press, 2002.

[KS06] Massaki Kashiwara and Pierre Schapira. *Categories and sheaves*. Vol. 332. Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, 2006.

[Kat88] Nicholas M. Katz. *Gauss sums, Kloosterman sums, and monodromy groups*. Vol. 116. Annals of Mathematics Studies. Princeton University Press, Princeton, NJ, 1988.

[KLR05] Chandrashekhar Khare, Michael Larsen, and Ravi Ramakrishna. "Constructing semisimple p-adic Galois representations with prescribed properties". In: *Amer. J. Math.* 127.4 (2005), pp. 709–734.

[KN74] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Pure and Applied Mathematics. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974.

[Lau09] Michel Laurent. "On transfer inequalities in Diophantine approximation". In: *Analytic number theory*. Cambridge Univ. Press, Cambridge, 2009, pp. 306–314.

[MLM94] Saunders Mac Lane and Ieke Moerdijk. *Sheaves in geometry and logic*. Second. Universitext. A first introduction to topos theory. Springer-Verlag, 1994.

[Mat89] Hideyuki Matsumura. *Commutative ring theory*. Second. Vol. 8. Cambridge Studies in Advanced Mathematics. Translated from the Japanese by M. Reid. Cambridge University Press, 1989.

[Maz97] Barry Mazur. "An introduction to the deformation theory of Galois representations". In: *Modular forms and Fermat's last theorem (Bostom, MA, 1995)*. New York: Springer, 1997, pp. 243–311.

[Maz08] Barry Mazur. "Finding meaning in error terms". In: *Bull. Amer. Math. Soc. (N.S.)* 45.2 (2008), pp. 185–228.

[Maz95] Fernando Mazzone. "A characterization of almost everywhere continuous functions". In: *Real Anal. Exchange* 21.1 (1995/96), pp. 317–319.

[NSW08] Jürgen Neukirch, Alexander Schmidt, and Kay Wingberg. *Cohomology of number fields*. Second. Vol. 323. Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, 2008.

[Nie91] Harald Niederreiter. "The distribution of values of Kloosterman sums". In: *Arch. Math. (Basel)* 56.3 (1991), pp. 270–277.

[Ö99] G. Ökten. *Error reduction techniques in quasi-Monte Carlo integration*. Vol. 30. 7-8. 1999, pp. 61–69.

[Pan11] Aftab Pande. "Deformations of Galois representations and the theorems of Sato–Tate and Lang–Trotter". In: *Int. J. Number Theory* 7.8 (2011), pp. 2065–2079.

[Ram02] Ravi Ramakrishna. "Deforming Galois representations and the conjectures of Serre and Fontaine–Mazur". In: *Ann. of Math. (2)* 156.1 (2002), pp. 115–154.

[Ros13] Zev Rosengarten. *An Erdös–Turán Inequality For Compact Simply-Connected Semisimple Lie Groups*. 2013. eprint: `arXiv:1305.2458`.

[RT16] Jeremy Rouse and Jesse Thorner. *The explicit Sato–Tate conjecture and densities pertaining to Lehmer-type questions*. 2016. eprint: `arXiv:1305.5283`.

[Sar07] Peter Sarnak. *Letter to: Barry Mazur on "Chebyshev's bias" for $\tau(p)$*. 2007.

[Ser89] Jean-Pierre Serre. *Abelian l-adic representations and elliptic curves*. Second. Advanced Book Classics. With the collaboration of Willem Kuyk and John Labute. Addison-Wesley Publishing Company, 1989.

[Ser94] Jean-Pierre Serre. "Propriétés conjecturales des groupes de Galois motiviques et des représentations l-adiques". In: *Motives (Seattle, WA, 1991)*. Vol. 55. Proc. Sympos. Pure Math. Amer. Math. Soc., Providence, RI, 1994, pp. 377–400.

[ST68] Jean-Pierre Serre and John Tate. "Good reduction of abelian varieties". In: *Ann. of Math. (2)* 88 (1968), pp. 492–517.

[Tay08] Richard Taylor. "Automorphy for some l-adic lifts of automorphic mod l Galois representations. II". In: *Publ. Math. Inst. Hautes Études Sci.* 108 (2008), pp. 183–239.

[Ten95] Gérald Tenenbaum. *Introduction to analytic and probabilistic number theory*. Vol. 46. Cambridge Studies in Advanced Mathematics. Translated from the second French edition (1995) by C. B. Thomas. Cambridge University Press, Cambridge, 1995.

[Til96] Jacques Tilouine. *Deformations of Galois representations and Hecke algebras*. Mehta Research Institute of Mathematics, 1996.

[Wei94] Charles Weibel. *An introduction to homological algebra*. Vol. 38. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1994.

[Yu15] Chia-Fu Yu. "A note on the Mumford–Tate conjecture for CM abelian varieties". In: *Taiwanese J. Math.* 19.4 (2015), pp. 1073–1084.