# KOLMOGOROV–SMIRNOV STATISTICS OF SATAKE PARAMETERS AND THE ANALYTIC PROPERTIES OF DIRICHLET SERIES ASSOCIATED TO ELLIPTIC CURVES

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

KOLMOGOROV–SMIRNOV STATISTICS OF SATAKE PARAMETERS AND THE ANALYTIC PROPERTIES OF DIRICHLET SERIES ASSOCIATED TO ELLIPTIC CURVES

Daniel Miller, Ph.D.

Cornell University 2017

Abstract here.

## BIOGRAPHICAL SKETCH

Daniel Miller was born in St. Paul, Minnesota.

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

## ACKNOWLEDGEMENTS

For starters, I'd like to thank my parents Jay and Cindy for noticing and fostering my mathematical interests early on, and for being loving and supportive the whole way through. I'd also like to thank my undergraduate thesis advisor, Griffith Elder, without whose encouragement and inspiration I'd probably never 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 to my graduate student friends Sasha Patotski and Balázs Elek for sharing my early love of algebraic geometry, for laughing with me at the absurdities of academic life, and listening to my ramblings about number theory long after they'd stopped being interesting.

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

I am especially thankful to my advisor Ravi. He kindled my first love for number theory, and stayed supportive as my research bounced all over the place, and helped focus and ground my thesis when I needed concrete results.

Lastly, 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

Let's start with something basic, an elliptic curve $E_{/\mathbf{Q}}$. For any prime $l$, we have the Tate module of $E$, written $\mathrm{T}_l E$. This is a rank-2 $\mathbf{Z}_l$-module with continuous $G_{\mathbf{Q}}$-action, so it induces a continuous representation

$$\rho_{E,l}\colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l).$$

It is known (citation?) that the quantities $a_p(E) = \operatorname{tr} \rho_l(\mathrm{fr}_p)$ lie in $\mathbf{Z}$ and satisfy the Hasse bound

$$|a_p(E)| \leqslant 2\sqrt{p}.$$

Thus we can define, for each prime $p$, the corresponding Satake parameter for $E$.

$$\theta_p(E) = \cos^{-1}\left(\frac{a_p(E)}{2\sqrt{p}}\right) \in [0,\pi).$$

The Satake parameters are packaged into an $L$-function as follows:

$$L^{\mathrm{an}}(E,s) = \prod_p \frac{1}{(1-e^{i\theta_p(E)}p^{-s})(1-e^{-i\theta_p(E)}p^{-s})}.$$

More generally we have, for each $k \geqslant 1$, the $k$-th symmetric power $L$-function

$$L^{\mathrm{an}}(\operatorname{sym}^k E,s) = \prod_p \prod_{j=0}^{k} \frac{1}{1-e^{i(k-2j)\theta_p(E)}p^{-s}}.$$

Numerical experiments suggest that the Satake parameters are distributed with respect to the Sato–Tate distribution $\mathrm{ST} = \frac{2}{\pi}\sin^2\theta\,\mathrm{d}\theta$. The "goodness of fit" of the Satake parameters to the Sato–Tate distribution is quantified by the *discrepancy*:

$$\mathrm{D}^\star(\{\theta_p(E)\}_{p\leqslant X},\mathrm{ST}) = \sup_{x\in[0,\pi]}\left|\frac{\#\{p\leqslant X : \theta_p(E)\in[0,x)\}}{\pi(X)} - \int_0^x \mathrm{dST}\right|.$$

The decay of the discrepancy is closely related to the analytic properties of the $L(\operatorname{sym}^k E,s)$. First, here is the famous Sato–Tate conjecture (now a theorem) in the language we have defined.

**Theorem 1.0.1** (Sato–Tate conjecture)**.** $\mathrm{D}^{\star}(\{\theta_p(E)\}_{p\leqslant X}, \mathrm{ST}) \to 0$.

**Theorem 1.0.2.** *The Sato–Tate conjecture for $E$ holds if and only if each of the functions $L(\operatorname{sym}^k E, s)$ have analytic continuation past $\Re s = 1$.*

The stunning recent proof of the Sato–Tate conjecture (citation) in fact showed that the functions $L(\operatorname{sym}^k E, s)$ were potentially automorphic, which gives analytic continuation.

There is an analogy between the above equivalence and classical analytic number theory. Let $K/\mathbf{Q}$ be a finite Galois extension, and $\rho\colon \operatorname{Gal}(K/\mathbf{Q}) \to \mathrm{GL}_n(\mathbf{C})$ an irreducible representation. Recall the Artin $L$-function is

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

Let $\operatorname{Gal}(K/\mathbf{Q})^{\natural}$ be the set of conjugacy classes in $\operatorname{Gal}(K/\mathbf{Q})$. The analogue of discrepancy here is:

$$\mathrm{D}(\{\mathrm{fr}_p\}_{p\leqslant X}) = \sup_{c\in\operatorname{Gal}(K/\mathbf{Q})^{\natural}} \left| \frac{\#\{p \leqslant X : \rho(\mathrm{fr}_p) \in c\}}{\pi(X)} - \frac{1}{\#\operatorname{Gal}(K/\mathbf{Q})^{\natural}} \right|.$$

**Theorem 1.0.3.** *The "discrepancy" $\mathrm{D}(\{\mathrm{fr}_p\}_{p\leqslant X}) \to 0$ if and only if $L(\rho, s)$ has analytic continuation past $\Re s = 1$ for all non-trivial irreducible representations $\rho$ of $\operatorname{Gal}(K/\mathbf{Q})$.*

In the case of Artin $L$-functions, we know moreover that

**Theorem 1.0.4.** *The "discrepancy" satisfies the bound $\mathrm{D}(\{\mathrm{fr}_p\}_{p\leqslant X}) \ll X^{-1/2+\epsilon}$ if and only if $L(\rho, s)$ satisfies the Riemann Hypothesis for all non-trivial irreducible representation $\rho$ of $\operatorname{Gal}(K/\mathbf{Q})$.*

In this context, the "Riemann Hypothesis" for $L(\rho, s)$ means exactly that $\log L(\rho, s)$ has analytic continuation to $\Re s = 1/2$.

The connection between the Riemann Hypothesis and "strong Sato–Tate" generalizes to elliptic curves and more general motives. For the moment, we stick to elliptic curves. In this case, "strong Sato–Tate" was conjectured by Akiyama–Tanigawa. More precisely,

Conjecture:

Let $E_{/\mathbf{Q}}$ be a non-CM elliptic curve. Then $\mathrm{D}^\star(\{\theta_p(E)\}_{p\leqslant X}, \mathrm{ST}) \ll X^{-1/2+\epsilon}$.

Moreover, one side of the equivalence "Riemann Hypothesis $\Leftrightarrow$ strong Sato–Tate" is known.

**Theorem 1.0.5.** *Let $E_{/\mathbf{Q}}$ be an elliptic curve. If the Akiyama–Tanigawa conjecture for $E$ holds, then all $L(\operatorname{sym}^k E, s)$ satisfy the Riemann Hypothesis.*

It is natural to assume that the converse to this theorem holds. However (and that is the main point of this thesis) it does not! In this thesis, I construct a range of counterexamples to the implication "strong Sato–Tate implies Riemann," and explore why the two are equivalent for Artin $L$-functions.

I also provide computational evidence for the Akiyama–Tanigawa conjecture (for elliptic curves and also generic abelian 2-folds).

Similar work: [Pan11].

Claim: the Riemann Hypothesis is equivalent to the following description of the distribution of prime numbers. For a real number $x$, let

$$P_x = \frac{1}{\pi(x)} \sum_{p\leqslant x} \delta_{p/x}.$$

This is a discrete probability measure supported on $[0,1]$. Moreover, let $L_x$ be the continuous probability measure with cdf

$$L_x[0,t] = \frac{\mathrm{Li}(tx)}{\mathrm{Li}(x)}.$$

Then $\mathrm{D}(P_x, L_x) \ll x^{-\frac{1}{2}+\epsilon}$. I'm pretty sure that this statement is equivalent to $|\pi(x) - \mathrm{Li}(x)| \ll x^{\frac{1}{2}+\epsilon}$, which is already known to be equivalent to RH.

CHAPTER 2

# DISCREPANCY

## 2.1 Definitions and first results

The 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 known properties, as well as how discrepancy changes when sequences are tweaked and/or combined.

Discrepancy will be defined for measures on the $d$-dimensional half-open box $[0,\infty)^d$. For vectors $x, y \in [0,\infty)^d$, we say $x < y$ if $x_1 < y_1, \ldots, x_d < y_d$, and in that case write $[x, y)$ for the half-open box $[x_1, y_1) \times \cdots \times [x_d, y_d)$.

**Definition 2.1.1.** *Let $\mu, \nu$ be probability measures on $[0,\infty)^d$. The* discrepancy *of $\mu$ with respect to $\nu$ is*

$$\mathrm{D}(\mu, \nu) = \sup_{x<y} \left| \mu[x, y) - \nu[x, y) \right|,$$

*where $x < y$ range over $[0,\infty)^d$.*

*The* star discrepancy *of $\mu$ with respect to $\nu$ is*

$$\mathrm{D}^\star(\mu, \nu) = \sup_{0<y} \left| \mu[0, y) - \nu[0, y) \right|,$$

*where $y$ ranges over $[0,\infty)^d$.*

**Lemma 2.1.2.** *Let $\mu, \nu$ 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 the discrepancy is taken over a larger set than that defining star discrepancy. To prove the second inequality, let $x < y$ be in $[0,\infty)^d$. For $S \subset \{1,\dots,d\}$, let

$$I_S = \{t \in [0,y) : t_i < x_i \text{ for all } i \in S\}.$$

The inclusion-exclusion principle for measures tells us that:

$$\mu[x,y) = \sum_{S\subset\{1,\dots,d\}} (-1)^{\#S}\mu(I_S),$$

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

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

□

We are usually interested in comparing empirical measures and their conjectured distribution. Namely, let $\boldsymbol{x} = \{x_p\}$ be a sequence in $[0,\infty)^d$ indexed by the prime numbers, and $\mu$ a Borel measure on $[0\infty)^d$. For any real number $N \geqslant 2$, we write $\boldsymbol{x}^N$ for the empirical measure given by

$$\boldsymbol{x}^N(S) = \frac{1}{\pi(N)}\sum_{p\leqslant N}\delta_{x_p}(S) = \frac{\#\{p \leqslant N : x_p \in S\}}{\pi(N)}.$$

Also, we write $\boldsymbol{x}_{\geqslant N}$ for the truncated sequence $(x_p)_{p\geqslant N}$, and similarly for $\boldsymbol{x}_{\leqslant N}$, etc. In this context,

$$\mathrm{D}^\star(\boldsymbol{x}^N,\nu) = \sup_{y\in[0,\infty)^d}\left|\frac{\#\{p \leqslant N : x_p \in [0,y)\}}{\pi(N)} - \int_{[0,y)} \mathrm{d}\nu\right|.$$

If the measure $\nu$ is only defined on a subset of $[0,\infty)^d$, we will tacitly extend it by zero. Moreover, if the sequence $\boldsymbol{x}$ actually lies in a torus $(\mathbf{R}/a\mathbf{Z})^d$, we identify that torus with the $[0,a)^d \subset [0,\infty)^d$. If $\nu$ is the Lebesgue measure (on $[0,\infty)^d$) or the normalized Haar measure on the torus, we write $\mathrm{D}^\star(\boldsymbol{x}^N)$ in place of $\mathrm{D}^\star(\boldsymbol{x}^N,\nu)$.

Sometimes the sequence $\boldsymbol{x}$ will not be indexed by the prime numbers, but rather by some other discrete subset of $\mathbf{R}^+$. In that case we will still use the notations $\boldsymbol{x}^N$, $\boldsymbol{x}_{\geqslant N}$, etc., keeping in mind that $\pi(N)$ is replaced by $\#\{\text{indices } \leqslant N\}$.

## 2.2 The Koksma–Hlawka inequality

Basically just summarize the paper [Ö99].

**Theorem 2.2.1** (Koksma–Hlawka)**.** *Todo.*

## 2.3 Comparing sequences

**Lemma 2.3.1.** *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be sequences in $[0,\infty)$. Suppose $\nu = f\cdot\lambda$ for $f$ a bounded continuous function and $\lambda$ the Lebesgue measure. Then*

$$\left|\mathrm{D}^\star(\boldsymbol{x}^N,\nu) - \mathrm{D}^\star(\boldsymbol{y}^N,\nu)\right| \leqslant \|f\|_\infty \epsilon + \mathrm{D}^\star(\boldsymbol{x}^N,\nu) + \frac{\#\{p\leqslant N : \|x_p - y_p\|_\infty \geqslant \epsilon\}}{\pi(N)}.$$

*Proof.* Let $\epsilon > 0$ and $t\in[0,\infty)$ be arbitrary. For all $p\leqslant N$ such that $y_p < t$, either $x_p < t+\epsilon$ or $\|x_p - y_p\|_\infty \geqslant \epsilon$. It follows that

$$\boldsymbol{y}^N[0,t) \leqslant \boldsymbol{x}^N[0,t+\epsilon) + \frac{\#\{p\leqslant N : \|x_p - y_p\|_\infty \geqslant \epsilon\}}{\pi(N)}.$$

Moreover, we trivially have

$$\left|\boldsymbol{x}^N[0,t+\epsilon) - \nu[0,t+\epsilon)\right| \leqslant \mathrm{D}^\star(\boldsymbol{x}^N,\nu).$$

Putting these together, we get:

$$\begin{aligned}\boldsymbol{y}^N[0,t) - \nu[0,t) &\leqslant \boldsymbol{x}^N[0,t+\epsilon) - \nu[0,t) + \frac{\#\{p \leqslant N : \|x_p - y_p\|_\infty \geqslant \epsilon\}}{\pi(N)} \\ &\leqslant \nu[t,t+\epsilon) + \mathrm{D}^\star(\boldsymbol{x}^N,\nu) + \frac{\#\{p \leqslant N : \|x_p - y_p\|_\infty \geqslant \epsilon\}}{\pi(N)} \\ &\leqslant \|f\|_\infty \epsilon + \mathrm{D}^\star(\boldsymbol{x}^N,\nu) + \frac{\#\{p \leqslant N : \|x_p - y_p\|_\infty \geqslant \epsilon\}}{\pi(N)}\end{aligned}$$

as desired. □

**Lemma 2.3.2.** *Let $\sigma$ 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 $\nu$ be an absolutely continuous measure on $[0,\infty)$ such that $\sigma_*\nu$ is also supported on $[0,\infty)$. Then*

$$\left|\mathrm{D}(\boldsymbol{x}^N,\nu) - \mathrm{D}(\sigma_*\boldsymbol{x}^N,\sigma_*\nu)\right| \leqslant \frac{2}{\pi(N)}.$$

*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 $\nu$ is absolutely continuous, $\nu\{t\} = 0$ for all $t \geqslant 0$. In particular, $\nu[s,t) = \nu(s,t]$. In contrast, $\boldsymbol{x}^N\{t\} \leqslant \pi(N)^{-1}$. For any interval $[s,t)$ in $[0,\infty)$, we know that

$$\left|\boldsymbol{x}^N[s,t) - \boldsymbol{x}^N(s,t]\right| \leqslant \frac{2}{\pi(N)},$$

hence

$$\left|\boldsymbol{x}^N[s,t) - \nu[s,t) - (\sigma_*\boldsymbol{x}^N)[a-t,a-s) - (\sigma_*\nu)[a-t,a-s)\right| \leqslant \frac{2}{\pi(N)}.$$

This proves the result. □

## 2.4 Combining sequences

**Definition 2.4.1.** *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be sequences in $[0,\infty)^d$. We write $\boldsymbol{x} \wr \boldsymbol{y}$ for the interleaved sequence*

$$(x_2, y_2, x_3, y_3, x_5, y_5, \dots, x_p, y_p, \dots).$$

For the interleaved sequence $\boldsymbol{x} \wr \boldsymbol{y}$, we write $(\boldsymbol{x} \wr \boldsymbol{y})^N$ for the empirical measure

$$(\boldsymbol{x} \wr \boldsymbol{y})^N = \frac{1}{2\pi(N)} \sum_{p \leqslant N} \delta_{x_p} + \delta_{y_p}.$$

**Theorem 2.4.2.** *Let $I$ and $J$ be disjoint open boxes in $[0,\infty)^d$, and let $\mu$, $\nu$ 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*

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

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

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

This yields the second inequality in the statement of the theorem. To see the first, assume that the maximum discrepancy is $\mathrm{D}(\boldsymbol{x}^N, \mu)$, and let $[s,t)$ be a half-open box such that $|\boldsymbol{x}^N[s,t) - \mu[s,t)|$ is within an arbitrary $\epsilon$ of $\mathrm{D}(\boldsymbol{x}^N, \mu)$. We can assume that $[s,t)$ does not intersect $J$, and thus

$$\big|(\boldsymbol{x} \wr \boldsymbol{y})^N[s,t) - (\mu+\nu)[s,t)\big| = |\boldsymbol{x}^N[s,t) - \mu[s,t)|,$$

which yields the result. $\square$

CHAPTER 3

# STRANGE DIRICHLET SERIES

## 3.1 Definitions

We start by considering a very general class of Dirichlet series. In fact, they are all Dirichlet series that admit a product formula with degree-1 factors, but in this thesis they will be called strange Dirichlet series. The motivating example was suggested by 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 "know" the rank of $E$?

**Definition 3.1.1.** *Let $\boldsymbol{z}=(z_2,z_3,z_5,\dots)$ be a sequence of complex numbers indexed by the primes. The associated* strange Dirichlet series *is*

$$L(\boldsymbol{z},s)=\prod_p \frac{1}{1-z_p p^{-s}}.$$

If $z_p$ is only defined for all but finitely many primes, then we tacitly set $\boldsymbol{z}_p=0$ for all primes for which $z_p$ is not defined.

**Lemma 3.1.2.** *Let $\boldsymbol{z}$ be a sequence with $\|\boldsymbol{z}\|_\infty \leqslant 1$. Then $L(\boldsymbol{z},s)$ defines a holomorphic function on the region $\{\Re s>1\}$. Moreover, on that region,*

$$\log L(\boldsymbol{z},s)=\sum_{p^r}\frac{z_p^n}{np^{ns}}.$$

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

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

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

**Lemma 3.1.3** (Abel summation)**.** *Let $\boldsymbol{z} = (z_2, z_3, z_5, \dots)$ be a sequence of complex numbers, $f$ a smooth complex-valued function on $\mathbf{R}$. Then*

$$\sum_{p\leqslant N} f(p) z_p = f(N) \sum_{p\leqslant N} z_p - \int_2^N f'(x) \sum_{p\leqslant x} z_p \,\mathrm{d}x.$$

*Proof.* Simply note that if $p_1, \dots, p_n$ is an enumeration of the primes $\leqslant N$, we have

$$\begin{aligned}\int_2^N f'(x) \sum_{p\leqslant x} z_p \,\mathrm{d}x &= \sum_{p\leqslant N} z_p \int_{p_n}^N f' + \sum_{i=1}^{n-1} \sum_{p\leqslant p_{i+1}} z_p \int_{p_i}^{p_{i+1}} f' \\ &= (f(N) - f(p_n)) \sum_{p\leqslant N} z_p + \sum_{i=1}^{n-1} (f(p_{i+1}) - f(p_i)) \sum_{p\leqslant p_{i+1}} z_p \\ &= f(N) \sum_{p\leqslant N} z_p - \sum_{p\leqslant N} f(p) z_p,\end{aligned}$$

as desired. $\square$

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

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

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

For each $p$, we have

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

Elementary analysis gives

$$1 \leqslant \frac{1}{1 - p^{-\Re s}} \leqslant 2 + 2\sqrt{2},$$

so the second piece of $\log L(\boldsymbol{z}, s)$ converges absolutely when $\Re s > \frac{1}{2}$. We could simply cite [Ten95, II.1 Th. 10]; instead we prove directly that $\sum_p \frac{z_p}{p^s}$ converges absolutely to a holomorphic function on the region $\{\Re s > \alpha\}$.

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

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

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

**Theorem 3.1.5.** *Let $\boldsymbol{z} = (z_2, z_3, \ldots)$ be a sequence with $\|\boldsymbol{z}\|_\infty \leqslant 1$, and assume $\log L(\boldsymbol{z}, s)$ has analytic continuation to $\{\Re s > \alpha\}$ for some $\alpha \in \frac{1}{2}, 1]$, and that for $\sigma > \alpha$, we have $|\log L(\boldsymbol{z}, \sigma + it)| \ll |t|^{1-\epsilon}$ (implied constant independent of $\sigma$.) Then $|\sum_{p \leqslant N} z_p| \ll N^{\alpha+\epsilon}$.*

*Proof.* Recall that we can write

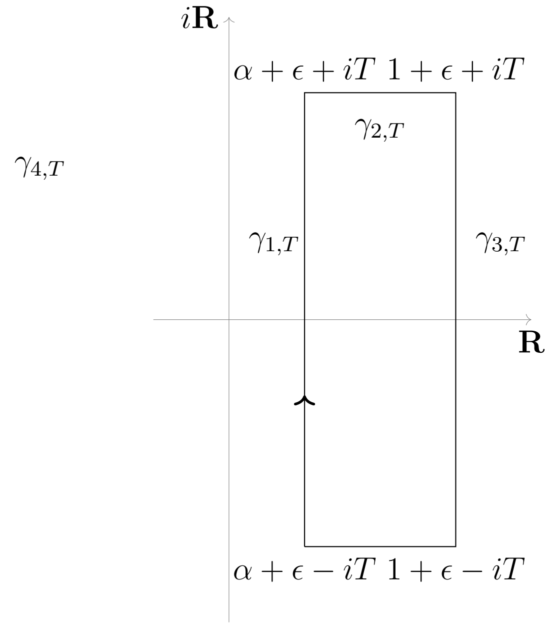
$$\log L(\boldsymbol{z}, s) = \sum_p \frac{z_p}{p^s} + \sum_p \sum_{r \geqslant 2} \frac{z_p^r}{r p^{rs}} = \sum_p \frac{z_p}{p^s} + O(\zeta(2\Re s)).$$

Thus, for any $\epsilon > 0$, analytic continuation and the bound on $|\log L(\boldsymbol{z}, \sigma + it)|$ implies the same analytic continuation and bound for $\sum \frac{z_p}{p^s}$ on $\{\Re s > \alpha + \epsilon\}$.

For any $T > 0$, let $\gamma_T = \gamma_{1,T} + \gamma_{2,T} + \gamma_{3,T} + \gamma_{4,T}$ be the following contour:

$$\begin{aligned}\gamma_{1,T}(t) &= (\alpha + \epsilon) + it \qquad t \in [-T, T] \\ \gamma_{2,T}(t) &= t + iT \qquad t \in [\alpha + \epsilon, 1 + \epsilon] \\ \gamma_{3,T}(t) &= (1 + \epsilon) + it \qquad t \in [T, -T] \\ \gamma_{4,T}(t) &= t - iT \qquad t \in [1 + \epsilon, \alpha + \epsilon].\end{aligned}$$

Graphically, the contour looks like this:




By Perron's formula [Apo76, Th. 11.18],

$$\lim_{T\to\infty} \frac{1}{2\pi i} \int_{-\gamma_{3,T}} \sum_p \frac{z_p}{p^s} N^z \frac{\mathrm{d}z}{z} = \frac{1}{2} \sum_{p\leqslant N} z_p.$$

for $N \in \mathbf{Z}$, and the same without the $\frac{1}{2}$ on the right-hand side when $N \notin \mathbf{Z}$.

Let $h(s)$ be the analytic continuation of $\sum z_p p^{-s}$ to $\{\Re s > \alpha\}$. Since $\int_{\gamma_T} h(s) \frac{\mathrm{d}s}{s} = 0$, we obtain

$$\left|\sum_{p\leqslant N} z_p\right| \ll \lim_{T\to\infty} \left( \left|\int_{\gamma_{1,T}} h(s) N^s \frac{\mathrm{d}s}{s}\right| + \left|\int_{\gamma_{2,T}} h(s) N^s \frac{\mathrm{d}s}{s}\right| + \left|\int_{\gamma_{4,T}} h(s) N^s \frac{\mathrm{d}s}{s}\right| \right).$$

We know that $|h(\sigma+it)| \ll |t|^{1-\epsilon}$, so we can bound

$$\left|\int_{\gamma_{2,T}} h(s) N^s \frac{\mathrm{d}s}{s}\right| = \left|\int_{\alpha+\epsilon}^{1+\epsilon} \frac{h(t+iT) N^{t+iT}}{t+iT} \,\mathrm{d}t\right| \ll \frac{N^{1+\alpha}}{T^\epsilon},$$

and similarly for $\gamma_{4,T}$. Finally, note that

$$\left|\int_{\gamma_{1,T}} h(s) N^s \frac{\mathrm{d}s}{s}\right| \ll \int_{-T}^{T} |t|^{1-\epsilon} \frac{N^{\alpha+\epsilon}}{(\alpha+\epsilon)^2+t^2} \,\mathrm{d}t \ll N^{\alpha+\epsilon}.$$

Letting $T \to \infty$ we obtain the desired result. □

In this thesis, we are interested in the following sort of strange Dirichlet series. 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 strange Dirichlet series.

## 3.2 Relation to automorphic and motivic $L$-functions

## 3.3 The Riemann Hypothesis

## 3.4 Discrepancy of sequences and the Riemann Hypothesis

## 3.5 Strange Dirichlet series over function fields

CHAPTER 4

# IRRATIONALITY EXPONENTS

## 4.1 Definitions and first results

We follow the notation of [Lau09]. Let $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$ be such that the $x_i$ are $\mathbf{Q}$-linearly independent.

**Definition 4.1.1.** *Let $\omega_0(x)$ (resp. $\omega_{d-1}(x)$) be the supremum of the set of real numbers $\omega$ for which there exist infinitely many $m = (m_0, \ldots, m_d) \in \mathbf{Z}^{r+1}$ such that*

$$\max\{|m_0 x_i - m_i|\} \leqslant \|m\|_\infty^{-\omega} \qquad \textit{(resp.}$$
$$|m_0 + m_1 x_1 + \cdots + m_r x_r| \leqslant \|m\|_\infty^{-\omega}\textit{)}.$$

These two quantities are related by Khintchine's Transference Principle, namely

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

Moreover, these inequalities are sharp in a very strong sense.

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

*Proof.* Do this. □

**Theorem 4.1.3.** *When $d = 1$, relate $\omega_0(x)$ to the irrationality measure.*

*Proof.* Recall that the irrationality measure $\mu(x)$ is the infimum of the set of positive reals $\mu$ such that

$$0 < \left| x - \frac{p}{q} \right| < q^{-\mu}$$

has only finitely many solutions $p/q$ with $p, q$ integers. □

Mention Roth's theorem. . . generalize to higher dimension?

Now given $x \in \mathbf{R}^d$, we write $d(x, \mathbf{Z}^d) = \min_{m \in \mathbf{Z}^d} |x - m|$, where $|\cdot|$ is any fixed norm on $\mathbf{R}^d$. Note that $d(x, \mathbf{Z}^d) = 0$ if and only if $x \in \mathbf{Z}^d$.

**Lemma 4.1.4.** *Let $x \in \mathbf{R}^d$ with $\|x\|_\infty \leqslant 1$ and $\omega_0(x)$ (resp. $\omega_{d-1}(x)$) finite. Then*

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

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

$$\max\{|m_0 x_i - m_i|\} \geqslant C_0 \|m\|_\infty^{-\omega_0(x)-\epsilon},$$
$$|m_0 + m_1 x_1 + \cdots + m_d x_d| \geqslant C_{d-1} \|m\|_\infty^{-\omega_{d-1}(x)-\epsilon}$$

for all $m \neq 0$.

Start with the first inequality in the statement of the result, where up to constant, we may assume that $|\cdot| = \|\cdot\|_\infty$ in the definition of $d(nx, \mathbf{Z}^d)$. Let $m = (m_1, \ldots, m_d)$ be the lattice point achieving the minimum $|nx - m|$. Then we know that

$$d(nx, \mathbf{Z}^d) \geqslant C_0 \|(m_1, \ldots, m_d)\|_\infty^{-\omega_0(x)-\epsilon}.$$

Moreover, since $|nx - m| < 1$, there exists a constant $C_0'$ such that

$$d(nx, \mathbf{Z}^d) \geqslant C_0' |n|^{-\omega_0(x)-\epsilon}.$$

It follows that

$$\frac{1}{d(nx, \mathbf{Z}^d)} \ll |n|^{\omega_0(x)+\epsilon},$$

the implied constant depending on $x$, $\epsilon$, and the choice of norm $|\cdot|$.

Now let's consider the second inequality in the statement of the result. Note that $d(m_1x_1+\cdots+m_dx_d,\mathbf{Z}) = |m_0+m_1x_1+\cdots+m_dx_d$ for some $m_0$ with $|m_0| \leqslant \|(m_1,\dots,m_d)\|_2\|x\|_2+1$. Thus $\|(m_1,\dots,m_d)\|_\infty \ll \|x\|_2\|(m_1,\dots,m_d)\|_2$, which gives us

$$d(m_1x_1+\cdots+m_dx_d,\mathbf{Z}) \geqslant C_{d-1}\|(m_1,\dots,m_d)\|_2^{-\omega_{d-1}(x)-\epsilon}.$$

This implies

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

the implied constant depending on $x$, $\epsilon$, and the choice of $|\cdot|$. □

## 4.2 Irrationality exponents and discrepancy

Let $x \in \mathbf{R}^d$ with $x_1,\dots,x_d$ linearly independent over $\mathbf{Q}$. We wish to control the discrepancy of the sequence $\{x, 2x, 3x, \dots\}$ in $(\mathbf{R}/\mathbf{Z})^d$.

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

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

*where the first sum ranges over $m \in \mathbf{Z}^d$, $r(m) = \prod \max\{1,|m_i|\}$, 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*

$$\left|\sum_{n\leqslant N} e^{2\pi inx}\right| \ll \frac{1}{d(x,\mathbf{Z})}.$$

*Proof.* We begin with an easy bound:

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

Since $|e^{2\pi i m x}-1| = \sqrt{2-2\cos(2\pi x)}$ 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 $|\sin(\pi x)| \geqslant d(x, \mathbf{Z})$, whence the result. $\square$

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

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

*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. $\square$

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

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

*Proof.* Choose $\delta > 0$ such that $\frac{1}{\omega_{d-1}(x)+1+\delta} = \frac{1}{\omega_{d-1}(x)+1} - \epsilon$.

By Corollary 4.2.3, we know that

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

and by Lemma 4.1.4, we know that $d(\langle m, x\rangle, \mathbf{Z})^{-1} \ll |m|^{\omega_{d-1}(x)+\delta}$. It follows that

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

The only tricky part is bounding the sum.

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

It follows that

$$\mathrm{D}(\boldsymbol{x}^N)\ll\frac{1}{h}+\frac{1}{N}(\log h)^{d-1}h^{\omega_{d-1}(x)+\delta}.$$

Setting $h\approx N^{\frac{1}{1+\omega_{d-1}(x)+\delta}}$, we see that

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

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 $x\in\mathbf{R}$ be such that $x_1,\ldots,x_d$ are linearly independent over $\mathbf{Q}$, and let $\boldsymbol{x}=(x,2x,3x,\ldots)$ in $(\mathbf{R}/\mathbf{Z})^d$. Then*

$$\mathrm{D}(\boldsymbol{x}^N)=\Omega\left(N^{-\frac{d}{\omega_0(x)}-\epsilon}\right).$$

*Proof.* Here $f=\Omega(g)$ in the sense of Hardy, namely that $\limsup\frac{f}{g}>0$. We follow the proof of [KN74, Ch. 2, Th. 3.3]. Given $\epsilon>0$, there exists $\delta>0$ such that $\frac{d}{\omega_0(x)-\delta}=\frac{d}{\omega_0(x)}+\epsilon$.

By the definition of $\omega_0(x)$, there exist infinitely many $(q,m_1,\ldots,m_d)$ with $q>0$ such that

$$\|qx-m\|_\infty\leqslant\|(q,m_1,\ldots,m_d)\|_\infty^{-\omega_0(x)+\delta/2}.$$

Since $\|(q,m_1,\ldots,m_d)\|_\infty\geqslant q$, we derive the stronger statement that for infinitely many $q\to\infty$, there exists $m=(m_1,\ldots,m_d)\in\mathbf{Z}^d$ such that $\|qx-m\|_\infty\leqslant$

$q^{-\omega_0(x)+\delta/2}$ or, equivalently, $|x - \frac{m}{q}| \leqslant q^{-1-\omega_0(x)+\delta/2}$. Pick such a $q$, and let $N = \lfloor q^{\omega_0(x)-\delta} \rfloor$. Then for each $n \leqslant N$, we have $\|nx - \frac{n}{q}m\|_\infty \leqslant q^{-1-\delta/2}$. Thus, for each $n \leqslant N$, each $nx$ 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 of $\boldsymbol{x}^N$ is bounded below by $2^{-d}q^{-d}$. Since $q^{\omega_0(x)-\delta} \leqslant 2N$, the discrepancy at $N$ is bounded below by

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

$\square$

CHAPTER 5

# DEFORMATION THEORY

## 5.1 Category of test objects

The following is an exposition and explication of the theory outlined in [SGA $3_1$, $\mathrm{VII_B}$, §0–1]. In particular, we will heavily use the notions of a pseudocompact ring, pseudocompact modules, etc. Let $\Lambda$ be a pseudocompact ring. Write $\mathsf{C}_\Lambda$ for the opposite of the category of $\Lambda$-algebras which have finite length as $\Lambda$-modules. Given such a $\Lambda$-algebra $A$, write $X = \mathrm{Spf}(A)$ for the corresponding object of $\mathsf{C}_\Lambda$, and we put $A = \mathscr{O}(X)$.

**Lemma 5.1.1.** *Let $\Lambda$ be a pseudocompact ring, $\mathsf{C}_\Lambda$ as above. Then $\mathsf{C}_\Lambda$ is closed under finite limits and colimits.*

**Lemma 5.1.2.** *Let $\Lambda$ be a pseudocompact local ring. Then $\Lambda$ is henselian, in any of the following senses:*

*1. d*

*Proof.* [EGA $4_4$, 18.5.?] □

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, so for $X, Y \in \mathcal{C}$, we 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 5.1.3.** *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 $\mathsf{C}_\Lambda$ together with a natural isomorphism $\mathcal{X}(\cdot) \simeq \varinjlim X_i(\cdot)$. 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.*

*Proof.* This follows from the results of [KS06, 6.1]. $\square$

If $R$ is a pseudocompact $\Lambda$-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 $\Lambda$-algebra homomorphisms.

**Lemma 5.1.4.** *The funtor* Spf *induces an (anti-)equivalence between the category of pseudocompact* $\Lambda$*-algebras and* $\mathsf{Ind}(\mathsf{C}_\Lambda)$*.*

*Proof.* This is [SGA $3_1$, $\mathrm{VII_B}$ 0.4.2 Prop.]. $\square$

So $\mathsf{Ind}(\mathsf{C}_\Lambda)$ is the category of pro-representable functors on finite length $\Lambda$-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 *noetherian*. We do not make this restriction.

**Lemma 5.1.5.** *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, this is easy to prove, see e.g. [Joh02, 4.2.3]. $\square$

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

**Corollary 5.1.6.** *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* $\Gamma$ *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 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 latter claim is just a special case. $\square$

## 5.2 Quotients in the flat topology

If $\Lambda$ is a pseudocompact ring, the category $\mathsf{Ind}(\mathsf{C}_\Lambda)$ has nice "geometric" properties. However, for operations like taking quotients, we will 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 $\mathsf{C}_\Lambda$ 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 $\mathsf{C}_\Lambda$. We call it the *flat topology*, even though finite presentation comes for free because all the rings are finite length.

**Lemma 5.2.1.** *Let* $\mathsf{Sh}_{\mathrm{fl}}(\mathsf{C}_\Lambda)$ *be the category of sheaves (of sets) on* $\mathsf{C}_\Lambda$ *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 moreover, 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 5.1.3, 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 $\mathsf{C}_\Lambda$, hence $\mathcal{X}(X) \to \mathcal{X}(U) \rightrightarrows \mathcal{X}(U \times_X U)$ is an equalizer. $\square$

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 $\mathsf{C}_\Lambda$ as follows. If $P$ is a property of maps in $\mathsf{C}_\Lambda$ (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 $\mathsf{C}_\Lambda$, 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 $\Lambda$-algebra maps $R \to A$, the induced map $A \to R' \otimes_R A$ has $P$.

**Theorem 5.2.2.** *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.*

*Proof.* This is [SGA $3_1$, $\mathrm{VII_B}$ 1.4]. $\square$

By [Mat89, 29.7], if $k$ is a field and $R$ is a complete regular local $k$-algebra, then $R \simeq k[\![t_1, \dots, t_n]\!]$. In particular, $R$ admits an augmentation $\epsilon\colon R \to k$. 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 $3_1$, $\mathrm{VII}_{\mathrm{B}}$ 0.1.1], it suffices to prove the result when $\mathcal{X} = \mathrm{Spf}(R')$, $\mathcal{Y} = \mathrm{Spf}(R)$, for local $\Lambda$-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

$$\begin{tikzcd} R' \arrow[r, dotted] \arrow[dr, two heads] & R \arrow[d, two heads] \\ R \arrow[u] \arrow[r, two heads] \arrow[ur, equal] & k, \end{tikzcd}$$

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 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 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

$$\begin{tikzcd}
\mathcal{R} \arrow[dr, equal] \arrow[drr, bend left] \arrow[ddr, bend right] & & \\
& \mathcal{X} \times_{\mathcal{X}/\mathcal{R}} \mathcal{X} \arrow[r] \arrow[d] & \mathcal{X} \arrow[d] \\
& \mathcal{X} \arrow[r] & \mathcal{X}/\mathcal{R}
\end{tikzcd}$$

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. $\square$

**Example 5.2.6.** The hypothesis on residue fields in 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 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 $\boldsymbol{\alpha}_p$ defined by

$$\boldsymbol{\alpha}_p(S) = \{s \in \mathscr{O}(S) \colon s^p = 0\}.$$

The quotient $\widehat{\mathbf{G}_\mathrm{a}}/\boldsymbol{\alpha}_p$ has as 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 hypothese of 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 Groupoids and quotient stacks

**Lemma 5.3.1.** *Let $\mathcal{G} \in \mathsf{Ind}(\mathsf{C}_\Lambda)$ be a smooth connected group. Then every $\mathcal{G}$-torsor is trivial.*

*Proof.* Let $\mathcal{P} \to \mathcal{B}$ be a $\mathcal{G}$-torsor in $\mathsf{Ind}(\mathsf{C}_\Lambda)$. That is, $\mathcal{P}$ has an action of $\mathcal{G}_\mathcal{S}$ for which $\mathcal{P} \times_\mathcal{B} \mathcal{P} \simeq \mathcal{G} \times \mathcal{P}$ as $\mathcal{G}$-spaces. [...not done...] □

**Theorem 5.3.2.** *Let $\mathcal{G}$ be a smooth connected group in $\mathsf{Ind}(\mathsf{C}_\Lambda)$, and $\mathcal{X} \in \mathsf{Ind}(\mathsf{C}_\Lambda)$ a $\mathcal{G}$-object. Then the quotient stack $[\mathcal{X}/\mathcal{G}](S)$ has as objects $\mathcal{X}(S)/\mathcal{G}(S)$, but with extra automorphisms?*

*Proof.* Use triviality of torsors. □

## 5.4 Deformations of group representations

Let $\Gamma \in \mathsf{Gp}(\widehat{\mathsf{C}_\Lambda})$ and $\mathcal{G} \in \mathsf{Ind}(\mathsf{C}_\Lambda)$. By 5.1.6, the functor

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

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].

Assume $\Lambda$ is local, with maximal ideal $\mathfrak{m}$ and residue field $\mathbf{k}$. Fix $\bar\rho \in \mathrm{Rep}^\square(\Gamma, \mathcal{G})(\mathbf{k})$, i.e. a residual representation $\bar\rho\colon \Gamma \to \mathcal{G}(\mathbf{k})$. Let $\mathrm{Rep}^\square(\Gamma, \mathcal{G})_{\bar\rho}$ be the connected component of $\bar\rho$ in $\mathrm{Rep}^\square(\Gamma, \mathcal{G})$. Assume that $\mathcal{G}$ and $\mathrm{Z}(\mathcal{G})$ are smooth; then the quotient $\mathcal{G}^{\mathrm{ad}} = \mathcal{G}/\mathrm{Z}(\mathcal{G})$ is also smooth. Let $\mathcal{G}^{\mathrm{ad},\circ}$ be the connected component of 1 in $\mathcal{G}^{\mathrm{ad}}$.

**Theorem 5.4.1.** *Suppose $(\Lambda, \mathfrak{m}, \mathbf{k})$ is local. If $\mathcal{X}, \mathcal{Y} \in \mathsf{Ind}(\mathsf{C}_\Lambda)$ are connected and $\mathcal{X}(\mathbf{k}) \neq \varnothing$, then $\mathcal{X} \times_\Lambda \mathcal{Y}$ is connected.*

*Proof.* We are reduced to proving the following result from commutative algebra: if $R, S$ are local pro-artinian $\Lambda$-algebras and $R$ has residue field $\mathbf{k}$, 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. $\mathfrak{s}$) ranges 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 $\mathcal{G}^{\mathrm{ad},\circ}$ on $\mathrm{Rep}^\square(\Gamma, \mathcal{G})$ preserves $\mathrm{Rep}^\square(\Gamma, \mathcal{G})_{\bar\rho}$. Thus we may put

$$\mathrm{Rep}(\Gamma, \mathcal{G})_{\bar\rho} = [\mathrm{Rep}^\square(\Gamma, \mathcal{G})_{\bar\rho}/\mathcal{G}^{\mathrm{ad},\circ}].$$

If $\mathcal{G}^{\mathrm{ad},\circ}$ acts faithfully on $\mathrm{Rep}^\square(\Gamma, \mathcal{G})_{\bar\rho}$, then we recover the classical notion of the deformation functor.

**Theorem 5.4.2.** *Let $\Gamma$ be a profinite group, $\bar\rho\colon \Gamma \to \mathcal{G}(\mathbf{k})$ a representation with $\mathrm{H}^0(\Gamma, \mathrm{Ad}\,\bar\rho) = 0$. Then $\mathrm{Rep}(\Gamma, \mathcal{G})_{\bar\rho}$ exists and is what you expect.*

*Proof.* Need assumptions on $\mathrm{Z}(\mathcal{G})$, $\mathcal{G}$ should be smooth.

Need $\mathrm{Z}(\mathcal{G}) = \ker(\mathcal{G} \to \mathrm{GL}(\mathfrak{g}))$ in connected case. This should use $\mathfrak{g} = \mathrm{Lie}(\mathrm{Aut}\,\mathcal{G})$, via deviations in [SGA $3_1$]. $\square$

[...local conditions]

## 5.5 Tangent spaces and obstruction theory

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

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

This gives us a class $\exp \in \mathrm{H}^2(\mathcal{G}(S_0), \mathfrak{g}(I))$. For $\rho_0 \colon \Gamma \to \mathcal{G}(S_0)$, 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 $\rho_0$ lifts to $\rho$. So obstruction theory naturally for $\mathrm{Rep}^{\square}(\Gamma, \mathcal{G})$.

[Use [Wei94, 6.6.4]. Given setting as above, $\rho_0^*(\exp)$ is the pullback by $\rho_0$:

$$\begin{tikzcd}
0 \arrow[r] & \mathfrak{g}(I) \arrow[r] \arrow[d, equal] & \mathcal{G}(S) \times_{\mathcal{G}(S_0)} \Gamma \arrow[r] \arrow[d] & \Gamma \arrow[r] \arrow[d, "\rho_0"] & 1 \\
0 \arrow[r] & \mathfrak{g}(I) \arrow[r] & \mathcal{G}(S) \arrow[r] & \mathcal{G}(S_0) \arrow[r] & 1
\end{tikzcd}$$

Computing explicitly, we see the result. ]

**Proposition 5.5.1.** *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:*

$$\begin{tikzcd}
 & \widetilde{H} \arrow[d] \\
G \arrow[ur, "\widetilde{f}"] \arrow[r, "f"'] & H.
\end{tikzcd}$$

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

$$\begin{tikzcd}
0 \arrow[r] & M \arrow[r] \arrow[d, equal] & G \times_H \widetilde{H} \arrow[r] \arrow[d] & G \arrow[r] \arrow[d, "f"] & 1 \\
0 \arrow[r] & M \arrow[r] & \widetilde{H} \arrow[r] & H \arrow[r] & 1.
\end{tikzcd}$$

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

---

Let $\mathcal{X} \in \mathsf{Ind}(\mathsf{C}_{/\Lambda})$ be smooth, and $\mathrm{L}_{\mathcal{X}/\Lambda} \simeq \Omega^1_{\mathcal{X}/\Lambda}[0]$ be its cotangent complex. Fix $x_0 \in \mathcal{X}(S_0)$. From the chain $S_0 \xrightarrow{x_0} \mathcal{X} \to \mathrm{Spf}(\Lambda)$, we get a distinguished triangle [Ill71, II 2.1.5.6]

$$x_0^*\mathrm{L}_{\mathcal{X}/\Lambda} \longrightarrow \mathrm{L}_{S_0/\Lambda} \longrightarrow \mathrm{L}_{S_0/\mathcal{X}} \longrightarrow .$$

If $I$ is a coherent sheaf on $S_0$, we get a long exact sequence:

$$\mathrm{Ext}^0(\mathrm{L}_{S_0/\Lambda}, M) \to \mathrm{Ext}^0(x_0^*\mathrm{L}_{\mathcal{X}/\Lambda}, M) \to \mathrm{Ext}^1(\mathrm{L}_{S_0/\mathcal{X}}, M) \to \mathrm{Ext}^1(\mathrm{L}_{S_0/\Lambda}, M) \to \mathrm{Ext}^1(x_0^*\mathrm{L}_{\mathcal{X}/\Lambda},$$

If $\mathcal{X}_{/\Lambda}$ is smooth, then $\mathrm{Ext}^1(x_0^*\mathrm{L}_{\mathcal{X}/\Lambda}, M) = 0$ and $\mathrm{L}_{\mathcal{X}/\Lambda} = \Omega^1_{\mathcal{X}/\Lambda}$. This gives us an exact sequence

$$\mathrm{Ext}^0(\mathrm{L}_{S_0/\Lambda}, M) \to \hom(\Omega^1_{\mathcal{X}/\Lambda}, M) \to \mathrm{Ext}^1(\mathrm{L}_{S_0/\mathcal{X}}, M) \to \mathrm{Ext}^1(\mathrm{L}_{S_0/\Lambda}, M) \to 0.$$

The result [Ill71, III 2.1.7] tells us that the choice of $S \in \mathsf{Ex}_{S_0}(M)$ gives us an element of $\mathrm{Ext}^1(\mathrm{L}_{S_0/\Lambda}, M)$. Its fiber admits an action of $\hom(\Omega^1_{\mathcal{X}/\Lambda}, M)$. The only thing remaining is: we need $\mathrm{Ext}^0(\mathrm{L}_{S_0/\Lambda}, M) = 0$, which doesn't hold in complete generality.

CHAPTER 6

# CONSTRUCTING GALOIS REPRESENTATIONS

## 6.1 Main idea

Basic ideas is as follows. Start with $\rho_1 \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_l)$. At each stage, we have $\rho_n \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}/l^n)$. At that stage, we're allowed to choose the (integral) characteristic polynomial for Frobenius at an arbitrarily large set of primes $R_n$. Then some (but density zero) extra primes ramify, and then we get $\rho_{n+1}$ that agrees with our choices for $R_1 \cup \cdots \cup R_n$. Then we have to choose characteristic polynomials for $R_{n+1}$, but these are already determined modulo $l^{n+1}$.

Basic idea is, we can choose the $R_n$ to be so huge that the primes involved are way way bigger than $l^n$. For example, we could have $R_n = \{p \leqslant l^{l^n}\}$. Thus the set of possible $a_p$'s is very big (big enough) so that we can get discrepancy to behave as we like (both decaying slowly and decaying quickly) to any measure $\mu$ such that $\mathrm{ST}/\mu$ is bounded away from zero. (By this, we mean: if $\mathrm{ST} = f \cdot \lambda$ and $\mu = g \cdot \lambda$, where $\lambda$ is Lebesgue, then $f/g$ is bounded away from zero.)

CHAPTER 7

# COUNTEREXAMPLE VIA DIOPHANTINE APPROXIMATION

## 7.1 Supporting results

Give $(\mathbf{R}/\mathbf{Z})^d$ the natural 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 $m \in \mathbf{Z}^d$

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

where $\langle m, x\rangle = m_1x_1 + \cdots + m_dx_d$ is the usual inner product.

**Theorem 7.1.1.** *Fix $x \in (\mathbf{R}/\mathbf{Z})^d$ with $\omega_{d-1}(x)$ finite. Then*

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

*as $m$ ranges over $\mathbf{Z}^r \smallsetminus 0$.*

*Proof.* From Lemma 4.2.2 we know that

$$\left|\sum_{n\leqslant N} e^{2\pi i\langle m,nx\rangle}\right| \ll \frac{1}{d(\langle m,x\rangle,\mathbf{Z})},$$

and from Lemma 4.1.4, we know that $d(\langle m,x\rangle,\mathbf{Z})^{-1} \ll |m|^{\omega_{d-1}(x)+\epsilon}$. The result follows. □

**Theorem 7.1.2.** *Let $x \in \mathbf{R}^d$ with $\omega_{d-1}(x)$ finite. Then let $f \in L^1((\mathbf{R}/\mathbf{Z})^d)$ with $\widehat{f}(0) = 0$ and suppose the Fourier coefficients of $f$ satisfy the bound $|\widehat{f}(m)| \ll |m|^{-\frac{1}{d-1}-\omega_{d-1}(x)-\epsilon}$. Then*

$$\left|\sum_{n\leqslant N} f(nx)\right| \ll 1.$$

*Proof.* Write $f$ as a Fourier series:

$$f(x) = \sum_{m\in\mathbf{Z}^r} \widehat{f}(m)e^{2\pi pi\langle m,x\rangle}.$$

Since $\widehat{f}(0)=0$, we can compute:

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

The sum converges since the exponent is less than $-\frac{1}{d-1}$, and it doesn't depend on $N$, hence the result. □

## 7.2 Pathological Satake parameters

Let $p_1=2, p_2=3, p_3=5,\dots$ be an enumeration of the prime numbers. Let $y\in\mathbf{R}^d$ with $y_1,\dots,y_d$ linearly independent over $\mathbf{Q}$. The associated sequence of "fake Satake parameters" is

$$\boldsymbol{x}=(y,2y,3y,4y,\dots),$$

where we put $x_{p_n}=ny \mod \mathbf{Z}^d$. By Theorem 4.1.2, we can arrange for $\omega_0(y)=w$ and $\omega_{d-1}(y)=dw+d-1$.

**Theorem 7.2.1.** *The sequence $\boldsymbol{x}$ is equidistributed in $(\mathbf{R}/\mathbf{Z})^d$, with discrepancy decaying as*

$$\mathrm{D}(\boldsymbol{x}^N)\ll N^{-\frac{1}{dw+d}+\epsilon}$$

*and for which*

$$\mathrm{D}(\boldsymbol{x}^N)=\Omega\left(N^{-\frac{d}{w}-\epsilon}\right).$$

*However, for any $f\in C^\infty((\mathbf{R}/\mathbf{Z})^d)$ with $\widehat{f}(0)=0$, the strange Dirichlet series $L_f(\boldsymbol{x},s)$ satisfies the Riemann Hypothesis.*

### 7.3 Some remarks on isotropic discrepancy

d

CHAPTER 8

**DIRECT COUNTEREXAMPLE**

## 8.1 Main ideas

The goal for this chapter is to construct a Galois representation $\rho_l \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$, ramified at a density-zero set of primes, such that

1. We have $a_p = \operatorname{tr} \rho(\mathrm{fr}_p) \in \mathbf{Z}$ whenever $\rho_l$ is unramified at $p$.
2. $|a_p| \leqslant 2\sqrt{p}$. ($\rho_l$ satisfies the Hasse bound.)
3. For $\theta_p = \cos^{-1} a_p/2\sqrt{p})$, the discrepancy $\mathrm{D}^\star(\boldsymbol{\theta}^N, \mathrm{ST}) \to 0$, but slower than any $N^{-\epsilon}$.
4. For any smooth $f \in C^\infty(\mathbf{R}/2\pi\mathbf{Z})$ such that $f(\pi - \theta) = -f(\theta)$, the associated $L$-function $L_f(\boldsymbol{\theta}, s)$ satisfies the Riemann Hypothesis.

These points together imply that $\rho_l$ satisfies the Sato–Tate conjecture, and $L(\operatorname{sym}^k \rho_l, s)$ satisfy the Riemann Hypothesis for odd $k$, but that $\rho_l$ does not satisfy the strong Sato–Tate conjecture.

## 8.2 Construction

Fix, for the remainder of this section, a continuous representation

$$\bar{\rho}_l \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{F}_l).$$

For each $p$ at which $\bar{\rho}_l$ is unramified, we write

$$\Theta_p(\bar{\rho}_l) = \left\{ \cos^{-1}\left(\frac{a}{2\sqrt{p}}\right) : a \in \mathbf{Z},\ |a| \leqslant 2\sqrt{p}, \text{ and } a \equiv \operatorname{tr} \bar{\rho}_l(\mathrm{fr}_p) \pmod{l} \right\}.$$

For the finitely many primes $p$ for which $\Theta_p(\bar{\rho}_l)$ is empty, redefine $\Theta_p(\bar{\rho}_l)$ to include some elements for which $|a| > 2\sqrt{p}$. We have a sequence of $\Theta_p(\bar{\rho}_l)$ for which at most finitely many do not satisfy the Hasse bound.

**Theorem 8.2.1.** *There exists a choice of $\theta_p \in \Theta_p(\bar{\rho}_l)$ for odd-indexed primes $\{2, 5, 11, \dots\}$ such that*

1. *$\theta_p \in [0, \pi/2)$ for all but finitely many $p$.*
2. *$\mathrm{D}\left(\boldsymbol{\theta}^N_{\mathrm{odd}}, \mathrm{ST}|_{[0,\pi/2)}\right) \to 0$, but is not $\ll N^{-\epsilon}$ for any $\epsilon > 0$.*

*Proof.* This is intuitively obvious, but a bit tricky to prove rigorously.

Two key ideas:

1. If we're given a "bad" finite distribution $\nu$, we can choose "good" $\theta_p$'s to make the combined distribution close enough (discrepancy-wise) to ST.

2. If we're given a "good" finite distribution $\nu$, we can choose "bad" $\theta_p \sim \pi/2$ to make the combined distribution far away (discrepancy-wise) from ST. □

Claim: let $\mu, \nu$ be two absolutely continuous distributions. Suppose there is a sequence $\{T_p\}$ of $\mu$-distributed sets, such that $\mathrm{D}(T_p, \mu) \ll p^{-1/2}$. Suppose moreover that $\mu/\nu$ is bounded away from zero (at the pdf side). Then we can choose $t_p \in T_p$ so that $\{t_p\}$ is $\nu$-equidistributed with good discrepancy.

## 8.3 Associated Galois representation

Let $\mu$ be an absolutely continuous measure on $[0, \pi]$ such that the pushforward $\cos_* \mu$ is bounded (this is true for the Sato–Tate measure). Fix a prime $l \geqslant 5$ and a constant $\alpha \in (0, 1/2]$. We want to construct a weight-2 Galois representation $\rho_l \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}_l)$, ramified at a density zero set of primes, such that

1. If $\rho_l$ is unramified at $p$, then $a_p = \operatorname{tr} \rho_l(\mathrm{fr}_p) \in \mathbf{Z}$ and satisfies the Hasse bound $|a_p| \leqslant 2\sqrt{p}$.

2. If we write $\theta_p = \cos^{-1}(a_p/2\sqrt{p})$ for the Satake parameters at unramified primes, then $\mathrm{D}(\boldsymbol{\theta}^N, \mu) \ll N^{-\alpha+\epsilon}$ and $\mathrm{D}(\boldsymbol{\theta}^N, \mu) = \Omega(N^{-\alpha-\epsilon})$.

Recall the van der Corput sequence $\{x_p\}$ satisfies $\mathrm{D}(\boldsymbol{x}^N) \ll N^{-1+\epsilon}$. Let $\nu = \cos_* \mu$; this is an absolutely continuous measure supported on $[-1,1]$. By transforming the van der Corput sequence by a continuous map, we may assume that in fact $\mathrm{D}(\boldsymbol{x}^N, \nu) \ll N^{-1+\epsilon}$. In fact, by alternating between "van der Corput elements" and "bad elements" we can ensure that not only does $\mathrm{D}(\boldsymbol{x}^N, \nu) \ll N^{-\alpha+\epsilon}$, but also $\mathrm{D}(\boldsymbol{x}^N, \nu) = \Omega(N^{-\alpha-\epsilon})$.

We start by choosing a modular mod-$l$ representation $\rho_1 \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}/l)$, which is ramified at a finite set of primes $S_1$. Let $R_1 = \{p \leqslant r_1 : p \notin S_1\}$. For $p \in R_1$, we can choose $a_p \in \mathbf{Z}$ subject only to the condition $a_p \equiv \operatorname{tr} \rho_1(\mathrm{fr}_p) \pmod{l}$. For any $p \in R_1$, the set

$$T_p(l) = \left\{ \frac{a}{2\sqrt{p}} : |a| \leqslant 2\sqrt{p} \text{ and } a \equiv \operatorname{tr} \rho_1(\mathrm{fr}_p) \pmod{l} \right\}$$

has an element within $lp^{-1/2}$ of any element of $[-1,1]$. Choose $a_p \in T_p(l)$ so that $\left| \frac{a_p}{2\sqrt{p}} - x_p \right| \leqslant lp^{-1/2}$. It follows that for $p \in R_1$, we have

$$|\,\mathrm{D}(\{a_p/2\sqrt{p}\}_{p \leqslant N}, \nu) - \mathrm{D}(\boldsymbol{x}^N, \nu)| \ll lN^{-1/2}$$

We get a lift of $\rho_1$ to $\rho_2 \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}/l^2)$ respecting our choices of the $a_p$ for $p \in R_1$, which is ramified at one (perhaps two) extra primes.

What happens next is in stages. We'll already have a mod-$l^{n+1}$ representation $\rho_{n+1} \colon G_{\mathbf{Q}} \to \mathrm{GL}_2(\mathbf{Z}/l^n)$, together with choices of $a_p$ for $p \in R_1 \cup \cdots \cup R_n$ that ensure $|\,\mathrm{D}(\{a_p/2\sqrt{p}\}_{p \leqslant N}, \nu) - \mathrm{D}(\boldsymbol{x}^N, \nu)| \ll ?$

The main question is: how do we choose $r_1$, and the later $r_n$? We ensure that a) the set $T_p(l^n)$ are non-empty, and that b) $l^n < \log(r_n)$. This gives us that for $N \leqslant r_n$, we have

$$|\,\mathrm{D}(\{a_p/2\sqrt{p}\}_{p\leqslant N}, \nu) - \mathrm{D}(\boldsymbol{x}^N, \nu)| \ll N^{-\frac{1}{2}+\epsilon}.$$

CHAPTER 9

# COMPUTATIONAL EVIDENCE FOR THE AKIYAMA–TANIGAWA CONJECTURE

CHAPTER 10

# CONCLUDING REMARKS AND FUTURE DIRECTIONS

## BIBLIOGRAPHY


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

[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.

[Ill71] Luc Illusie. *Complexe cotangent et déformations. I*. Vol. 239. Lecture Notes in Mathematics. Springer-Verlag, 1971.

[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.

[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.

[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.

[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.