# Contents

# Introduction

In a remarkable numerical experiment, Odlyzko [**Odl**] found that the distribution of the (suitably normalized) spacings between successive zeroes of the Riemann zeta function is (empirically) the same as the so-called GUE measure, a certain probability measure on $\mathbb{R}$ arising in random matrix theory. His experiment was inspired by work of Montgomery [**Mon**], who determined the pair correlation distribution between zeroes (in a restricted range), and who noted the compatibility of what he found with the GUE prediction. Recent results of Rudnick and Sarnak [**Ru-Sar**] are also compatible with the belief that the distribution of the spacings between zeroes, not only of the Riemann zeta function, but also of quite general automorphic $L$-functions over $\mathbb{Q}$, are all given by the GUE measure, or, as we shall say, all satisfy the Montgomery-Odlyzko Law. Unfortunately, proving this seems well beyond range of existing techniques, and we have no results to offer in this direction.

However, it is a long established principle that problems which seem inaccessible in the number field case often have finite field analogues which are accessible. In this book we establish the Montgomery-Odlyzko Law for wide classes of zeta and $L$-functions over finite fields.

To fix ideas, let us consider a special case, which none the less contains all the essential phenomena, the case of curves over finite fields. Thus we consider a finite field $\mathbb{F}_q$, and a proper, smooth, geometrically connected curve $C/\mathbb{F}_q$ of genus $g$. [For example, if we take a homogeneous form $F(X, Y, Z)$ over $\mathbb{F}_q$ of degree $d$ in three variables such that $F$ and its first partial derivatives have no common zeroes in $\overline{\mathbb{F}}_q$, then the projective plane curve of equation $F = 0$ in $\mathbb{P}^2$ is such a curve, of genus $g = (d-1)(d-1)/2$.] The zeta function of $C/\mathbb{F}_q$, denoted $Z(C/\mathbb{F}_q, T)$, was first introduced by Artin [**Artin**] in his thesis. It is the basic diophantine invariant of $C/\mathbb{F}_q$, constructed out of the numbers $N_n := \mathrm{Card}(C(\mathbb{F}_{q^n}))$ of points on $C$ with coordinates in the unique field extension $\mathbb{F}_{q^n}$ of $\mathbb{F}_q$ of each degree $n \geq 1$. $T$ As a formal series over $\mathbb{Q}$ in one variable $T$, $Z(C/\mathbb{F}_q, T)$ is defined as the generating series

$$Z(C/\mathbb{F}_q, T) := \exp\left(\sum_{n\geq 1} N_n T^n/n\right).$$

One knows that in fact $Z(C/\mathbb{F}_q, T)$ is a rational function of $T$, of the form

$$P(T)/(1-T)(1-qT),$$

where $P(T)$ is a polynomial of degree $2g$ with $\mathbb{Z}$-coefficients. By the Riemann Hypothesis for curves over finite fields [**Weil-CA**], one knows that the reciprocal zeroes of $P(T)$ all have complex absolute value $\mathrm{Sqrt}(q)$, i.e., we have

$$P(T) = \prod_{j=1}^{2g}(1 - \alpha_j T), \quad \text{with } |\alpha_j|_{\mathbb{C}} = \mathrm{Sqrt}(q) \text{ for all } j.$$

We write

$$\alpha_j = \mathrm{Sqrt}(q)e^{i\varphi_j}, \qquad 0 \le \varphi_j < 2\pi.$$

Renumbering, we may assume that

$$0 \le \varphi_1 \le \varphi_2 \le \cdots \le \varphi_{2g} < 2\pi.$$

The normalized spacings between the (reciprocal) zeroes of the zeta function of $C/\mathbb{F}_q$ are the following $2g$ real numbers. The first $2g-1$ are

$$(g/\pi)(\varphi_2 - \varphi_1), (g/\pi)(\varphi_3 - \varphi_2), \ldots, (g/\pi)(\varphi_{2g} - \varphi_{2g-1}),$$

and the last is the "wraparound" spacing

$$(g/\pi)(\varphi_1 + 2\pi - \varphi_{2g}).$$

The spacing measure $\mu = \mu(C/\mathbb{F}_q)$ attached to $C/\mathbb{F}_q$ is the probability measure on $\mathbb{R}$, supported in $\mathbb{R}_{\ge 0}$, which gives mass $1/2g$ to each of the $2g$ normalized spacings.

Before going on, we must first say what is the GUE measure on $\mathbb{R}$, cf. 1.0–2. For this, we first pick an integer $N \ge 1$, and consider the unitary group $U(N)$ of size $N$. Given an element $A$ in $U(N)$, its $N$ eigenvalues lie on the unit circle, and we form the $N$ normalized (to have mean 1) spacings between pairs of adjacent eigenvalues, and out of these $N$ spacings we form the probability measure on $\mathbb{R}$ which gives mass $1/N$ to each of the $N$ normalized spacings. This measure we call $\mu(A, U(N))$, the spacing measure attached to an element $A$ in $U(N)$. We view $A \mapsto \mu(A, U(N))$ as a measure-valued function on $U(N)$. One can make sense of the integral of this function over $U(N)$ against the total mass one Haar measure $dA$ on $U(N)$: the result makes sense as a probability measure on $\mathbb{R}$, denoted

$$\mu(U(N)) := \int_{U(N)} \mu(A, U(N))\, dA.$$

One then shows that as $N$ grows, the measures $\mu(U(N))$ on $\mathbb{R}$ have a limit which is again a probability measure on $\mathbb{R}$, which we denote $\mu(\mathrm{univ})$, and call the GUE measure.[1] One shows that its cumulative distribution function

$$\mathrm{CDF}_{\mu(\mathrm{univ})}(x) := \int_{[-\infty, x]} d\mu(\mathrm{univ})$$

is continuous on $\mathbb{R}$. [In fact, this measure has a density, which vanishes outside $\mathbb{R}_{>0}$, and is real analytic on $\mathbb{R}_{\ge 0}$, cf. Appendix: Graphs for a picture.]

For the application to curves that we have in mind, we need to know that we can obtain the GUE measure not just from the series of unitary groups $U(N)$, but also from any of the series of compact classical groups. Indeed, suppose we are given any compact subgroup $K$ of a given unitary group $U(N)$. We can, for each element $A$ in $K$, form the spacing measure attached to $A$ thought of as an element of $U(N)$. To remind ourselves that we do this only for elements of $K$, we denote this measure $\mu(A, K)$. Then we view $A \mapsto \mu(A, K)$ as a measure-valued function on $K$, and we integrate this function against the total mass one Haar measure $dA$

[1] In the physics literature, this measure often carries Wigner's name

on the compact group $K$. The result, denoted $\mu(K) := \int_K \mu(A,K)\,dA$, is itself a probability measure on $\mathbb{R}$.

We can perform this construction with $K$ any of the compact classical groups, $U(N)$ or $SU(N)$ or $USp(2N)$ or $SO(2N+1)$ or $SO(2N)$ or $O(2N+1)$ or $O(2N)$ in their standard representations. We show that for $G(N)$ running over any of these series of compact classical groups, the sequence of probability measures $\mu(G(N))$ on $\mathbb{R}$ converges, as $N$ grows, to the **same** measure $\mu(\text{univ})$, the GUE measure, that we obtained as the large $N$ limit of the $\mu(U(N))$ measures. [The case which will be relevant to curves over finite fields will turn out to be the compact symplectic groups $USp(2N)$.]

Now let us return to a curve $C/\mathbb{F}_q$ over a finite field, of some genus $g$. Since the spacing measure $\mu(C/\mathbb{F}_q)$ gives each of $2g$ points mass $1/2g$, its CDF is a step function, with $2g$ jumps. So it cannot possibly be the case that $\mu(C/\mathbb{F}_q)$ is equal to the GUE measure, whose CDF is continuous. Moreover, as we shall see later in this Introduction, over any finite field there are sequences of curves of increasing genus whose spacing measures are arbitrarily close to the delta measure $\delta_0$ supported at the origin. So it is simply **not true** that the spacing measures of **all** curves of large genus are close to the GUE measure. What we show is that "most" curves of large genus over a large finite field have their spacing measure quite close to the GUE measure, or in other words that "most" curves of sufficiently large genus over a sufficiently large finite field satisfy the Montgomery-Odlyzko Law to an arbitrary degree of precision.

To make this more precise, we need a numerical measure of how close two probability measures on $\mathbb{R}$, say $\mu$ and $\nu$, are. For this, we use the Kolmogoroff-Smirnov discrepancy, defined as the greatest vertical distance between the graphs of their CDF's:

$$\text{discrep}(\mu,\nu) := \operatorname*{Sup}_{s \text{ in } \mathbb{R}} |\,\text{CDF}_\mu(s) - \text{CDF}_\nu(s)|.$$

Notice that $\text{discrep}(\mu,\nu)$ is a number which always lies in the closed interval $[0,1]$, just because CDF's of probability measures take values in $[0,1]$.

Now let us denote by $\mathcal{M}_g(\mathbb{F}_q)$ the set, known to be finite, consisting of all $\mathbb{F}_q$-isomorphism classes of genus $g$ curves over $\mathbb{F}_q$. Our essential result about the spacing measures $\mu(C/\mathbb{F}_q)$ attached to curves over finite fields, and their relation to the GUE measure $\mu(\text{univ})$, is this:

**Theorem** (cf. 12.2.3)**.** *We have the double limit formula*

$$\lim_{g\to\infty}\lim_{q\to\infty}(1/|\mathcal{M}_g(\mathbb{F}_q)|)\sum_{C \textit{ in } \mathcal{M}_g(\mathbb{F}_q)} \text{discrep}(\mu(\textit{univ}),\mu(C/\mathbb{F}_q)) = 0.$$

*More precisely, for any real $\varepsilon > 0$, there exists an integer $N(\varepsilon)$ such that for any genus $g > N(\varepsilon)$, we have the inequality*

$$\lim_{q\to\infty}(1/|\mathcal{M}_g(\mathbb{F}_q)|)\sum_{C \textit{ in } \mathcal{M}_g(\mathbb{F}_q)} \text{discrep}(\mu(\textit{univ}),\mu(C/\mathbb{F}_q)) \le g^{\varepsilon-1/6}.$$

To see what this says concretely, pick a small $\varepsilon > 0$, and fix a genus $g > N(\varepsilon)$. Then for $q$ sufficiently large, say $q > M(\varepsilon, g)$, we will have

$$(**) \qquad (1/|\mathcal{M}_g(\mathbb{F}_q)|)\sum_{C \text{ in } \mathcal{M}_g(\mathbb{F}_q)} \text{discrep}(\mu(\text{univ}),\mu(C/\mathbb{F}_q)) \le 2g^{\varepsilon-1/6}.$$

To see what this last inequality means about "most" curves, pick any two positive real numbers $\alpha$ and $\beta$ with $\alpha + \beta = 1/6 - \varepsilon$. Denote by

$$\mathcal{M}_g(\mathbb{F}_q)(\text{discrep} > g^{-\alpha}) \subset \mathcal{M}_g(\mathbb{F}_q)$$

the set of those $C$ in $\mathcal{M}_g(\mathbb{F}_q)$ for which

$$\text{discrep}(\mu(\text{univ}), \mu(C/\mathbb{F}_q)) > g^{-\alpha}.$$

Then we easily infer from $(**)$ above that

$$|\mathcal{M}_g(\mathbb{F}_q)(\text{discrep} > g^{-\alpha})|/|\mathcal{M}_g(\mathbb{F}_q)| \leq 2g^{-\beta},$$

i.e., the fraction of curves in $\mathcal{M}_g(\mathbb{F}_q)$ whose discrepancy exceeds $g^{-\alpha}$ is at most $2g^{-\beta}$, provided that $g > N(\varepsilon)$ and provided that $q > M(\varepsilon, g)$. In other words, if $g$ and then $q$ are sufficiently large, then the probability is at least $1 - 2g^{-\beta}$ that a randomly chosen curve in $\mathcal{M}_g(\mathbb{F}_q)$ has discrepancy $\leq g^{-\alpha}$. This is the sense in which most curves of sufficiently large genus over a sufficiently large finite field have a spacing measure which is arbitrarily close to the GUE measure.

To explain how one proves such results, we must now return to a discussion of the GUE measure $\mu(\text{univ})$ and its genesis from compact classical groups $G(N)$. Suppose we take a particular $G(N)$, and an element $A$ in $G(N)$. How close is the spacing measure $\mu(A, G(N))$ to the GUE measure? The answer is that "most" elements $A$ of a large $G(N)$ have their spacing measures quite close to the GUE measure, as the following "law of large numbers" shows.

**Theorem** (cf. 1.2.6)**.** *In any of the series of compact classical groups $G(N) = U(N)$ or $SU(N)$ or $USp(2N)$ or $SO(2N+1)$ or $SO(2N)$ or $O(2N+1)$ or $O(2N)$, we have*

$$\lim_{N\to\infty} \int_{G(N)} \text{discrep}(\mu(A, G(N)), \mu(\mathit{univ}))\, dA = 0.$$

*More precisely, given $\varepsilon > 0$, there exists an integer $N(\varepsilon)$ such that for any $N > N(\varepsilon)$, we have*

$$\int_{G(N)} \text{discrep}(\mu(A, G(N)), \mu(\mathit{univ}))\, dA \leq N^{\varepsilon - 1/6}.$$

We also remark that the integrand above,

$$A \mapsto \text{discrep}(\mu(A, G(N)), \mu(\text{univ})),$$

is a continuous (cf. 1.0.12) central function on $G(N)$. This remark will allow us below to apply Deligne's equidistribution theorem (cf. 9.2.6, 9.6.10, 9.7.10) in a completely straightforward way.

The connection between Theorems 12.2.3 and 1.2.6 comes about through monodromy, and Deligne's equidistribution theorem (9.6.10). Recall that the zeta function of a genus $g$ curve $C/\mathbb{F}_q$ is of the form $P(T)/(1-T)(1-qT)$, for $P$ a polynomial of degree $2g$ with the property that the auxiliary polynomial $P(T/\operatorname{Sqrt}(q))$ has all its roots on the unit circle. However, the polynomial $P(T/\operatorname{Sqrt}(q))$ has a bit more structure; namely, its $2g$ roots on the unit circle can be partitioned into $g$ pairs of inverses $(\xi, 1/\xi)$ on the unit circle. One interpretation of this fact is that there exists a conjugacy class $\vartheta(C/\mathbb{F}_q)$ in the compact group $USp(2g)$ such that

$$P(T/\operatorname{Sqrt}(q)) = \det(1 - T\vartheta(C/\mathbb{F}_q)).$$

Because conjugacy classes in $USp(2g)$ are uniquely determined by their characteristic polynomials, there is a unique such conjugacy class $\vartheta(C/\mathbb{F}_q)$ in $USp(2g)$, which we call the unitarized Frobenius conjugacy class attached to $C/\mathbb{F}_q$.

Now fix an integer $N \geq 1$ and a genus $g \geq 1$. Consider a proper smooth family $\pi : \mathcal{C} \to S$ of genus $g$ curves, parameterized by a scheme $S$ which, for simplicity, we assume to be smooth and surjective over $\mathrm{Spec}(\mathbb{Z}[1/N])$ with geometrically connected fibres. We further assume that for every prime number $p$ which does not divide $N$, the geometric monodromy group of the family of curves

$$\pi \otimes \mathbb{F}_p : \mathcal{C} \otimes \mathbb{F}_p \to S \otimes \mathbb{F}_p$$

in characteristic $p$ is the full symplectic group $Sp(2g)$. Once we have made this assumption about the monodromy of the family, Deligne's equidistribution theorem (cf. 9.6.10) says the following. For each finite field $\mathbb{F}_q$ of characteristic not dividing $N$, and each point $s$ in the finite set $S(\mathbb{F}_q)$ of $\mathbb{F}_q$-valued points of $S$, look at the curve $\mathcal{C}_s/\mathbb{F}_q$ named by the point $s$, and look at its unitarized Frobenius conjugacy classes $\vartheta(\mathcal{C}_s/\mathbb{F}_q)$ in $USp(2g)$. Then these unitarized Frobenius conjugacy classes are equidistributed in the space $USp(2g)^{\#}$ of conjugacy classes in $USp(2g)$ for the probability measure $\mu_{\mathrm{Haar}}$ on $USp(2g)^{\#}$ which is the direct image from $USp(2g)$ of its normalized Haar measure, in the following sense: for any $\mathbb{C}$-valued continuous central function $f$ on $USp(2g)$, we have the limit formula

$$\lim_{q\to\infty, q \text{ prime to } N} (1/|S(\mathbb{F}_q)|) \sum_{s \text{ in } S(\mathbb{F}_q)} f(\vartheta(\mathcal{C}_s/\mathbb{F}_q))$$
$$= \int_{USp(2g)^{\#}} f \, d\mu_{\mathrm{Haar}}.$$

In order to apply this to study the discrepancy for the curves over finite fields $\mathcal{C}_s/\mathbb{F}_q$ which occur in our family, we have only to apply Deligne's equidistribution theorem above to the continuous central function $f$ on $USp(2g)$ given by

$$A \mapsto \mathrm{discrep}(\mu(A, USp(2g)), \mu(\mathrm{univ})).$$

We know from Theorem 1.2.6 quoted above that, given $\varepsilon > 0$, there is an $N(\varepsilon)$ such that for $g \geq N(\varepsilon)$, we have, for this $f$, the estimate

$$\int_{USp(2g)^{\#}} f \, d\mu_{\mathrm{Haar}} \leq g^{\varepsilon - 1/6}.$$

So if our family $\mathcal{C}/S$ has $g > N(\varepsilon)$, and we use Deligne's equidistribution theorem to calculate this integral, we find the estimate

$$\lim_{q\to\infty, \text{prime to } N} (1/|S(\mathbb{F}_q)|) \sum_{s \text{ in } S(\mathbb{F}_q)} \mathrm{discrep}(\mu(\mathrm{univ}), \mu(\mathcal{C}_s/\mathbb{F}_q)) \leq g^{\varepsilon - 1/6}.$$

In particular, if $q$ is prime to $N$ and sufficiently large, we will have

$$(**) \qquad (1/|S(\mathbb{F}_q)|) \sum_{s \text{ in } S(\mathbb{F}_q)} \mathrm{discrep}(\mu(\mathrm{univ}), \mu(\mathcal{C}_s/\mathbb{F}_q)) \leq 2g^{\varepsilon - 1/6}.$$

To see what this means about the discrepancy of "most" curves in the family $\mathcal{C}/S$, pick a pair of positive real numbers $\alpha, \beta$ with $\alpha + \beta = 1/6 - \varepsilon$. Denote by

$$S(\mathbb{F}_q)(\mathrm{discrep} > g^{-\alpha}) \subset S(\mathbb{F}_q)$$

the set of those $s$ in $S(\mathbb{F}_q)$ for which

$$\mathrm{discrep}(\mu(\mathrm{univ}), \mu(\mathcal{C}_s/\mathbb{F}_q)) > g^{-\alpha}.$$

Then we easily infer from $(**)$ above that

$$|S(\mathbb{F}_q)(\text{discrep} > g^{-\alpha})|/|S(\mathbb{F}_q)| \leq 2g^{-\beta},$$

provided that $g \geq N(\varepsilon)$ and that $q$ is prime to $N$ and sufficiently large (how large depends on the particular family $\mathcal{C}/S$).

To obtain Theorem 12.2.3 stated above about $\mathcal{M}_g$, we need only replace $S(\mathbb{F}_q)$ in the above formulas by $\mathcal{M}_g(\mathbb{F}_q)$. There are some technical difficulties to be overcome in justifying this formal replacement; cf. 10.6 and 10.7 for an exhaustive discussion of these difficulties and their resolution.

Once we know that "most" curves over finite fields have their spacing measure close to the GUE measure, it is natural to ask if, given a finite field $\mathbb{F}_q$, we can exhibit a single **explicit** sequence of curves $\{C_g/\mathbb{F}_q\}_g$ with $C_g$ of genus $g$, whose spacing measures $\mu(C_g/\mathbb{F}_q)$ approach the GUE measure $\mu(\text{univ})$ as $g \to \infty$, in the sense that $\lim_{g\to\infty} \text{discrep}(\mu(C_g/\mathbb{F}_q), \mu(\text{univ})) = 0$. We do not know how to do this at present.

To the extent that we can write down families (of varieties, of exponential sums, of ...) over finite fields whose geometric monodromy groups are large classical groups, we will get results similar to those for curves for the behavior of the discrepancy in these families as well. We work this out explicitly for universal families of abelian varieties (where the group is again $Sp$), of smooth hypersurfaces in projective space (where the group is either $Sp$ or $O$), and for multi-variable Kloosterman sums (where the group is either $Sp, SL$ or $SO(odd)$). Again in these more general cases we do not know how to write down explicit sequences of objects of the type considered whose spacing measures approach the GUE measure.

In the case of the Kloosterman sums $\text{Kl}_n(\psi, a \text{ in } \mathbb{F}_q^\times)$ there is a plausible candidate for such a sequence.

**Conjecture.** *Fix a finite field $\mathbb{F}_q$, fix any choice of $a$ in $\mathbb{F}_q^\times$ and fix any choice of the nontrivial additive character $\psi$ of $\mathbb{F}_q$. Then the spacing measure*

$$\mu(\text{Kl}_n(\psi, a \textit{ in } \mathbb{F}_q^\times))$$

*attached to $\text{Kl}_n(\psi, a \text{ in } \mathbb{F}_q^\times)$, or more precisely to its $L$-function, tends to the GUE measure as $n \to \infty$ in the sense that*

$$\lim_{n\to\infty} \text{discrep}(\mu(\text{Kl}_n(\psi, a \textit{ in } \mathbb{F}_q^\times)), \mu(\textit{univ})) = 0.$$

Suppose now that we fix an integer $N \geq 1$, and a large integer $g$. Suppose that we are given a curve $C/\mathbb{Z}[1/N]$ which is proper and smooth with geometrically connected fibres of genus $g$. For any prime $p$ not dividing $N$, the reduction $\text{mod}\, p$ of our curve $C/\mathbb{Z}[1/N]$ is a curve $C \otimes \mathbb{F}_p/\mathbb{F}_p$ of genus $g$, which has a spacing measure $\mu(C \otimes \mathbb{F}_p/\mathbb{F}_p)$. When is it reasonable to expect that for most primes $p$ which are prime to $N$, the spacing measure $\mu(C \otimes \mathbb{F}_p/\mathbb{F}_p)$ is close to the GUE measure $\mu(\text{univ})$? When should we expect some other behaviour?

Given $C/\mathbb{Z}[1/N]$ as above, for every prime $p$ not dividing $N$, we obtain a unitarized Frobenius conjugacy class $\vartheta(C \otimes \mathbb{F}_p/\mathbb{F}_p)$ in $USp(2g)$. When is it reasonable to expect that these classes $\vartheta(C \otimes \mathbb{F}_p/\mathbb{F}_p)$ are equidistributed in $USp(2g)^{\#}$, in the sense that for any $\mathbb{C}$-valued continuous central function $f$ on $USp(2g)$ we have

$$\int_{USp(2g)} f(A)\, dA = \lim_{X\to\infty} (1/\pi(X)) \sum_{p \leq X \text{ prime to } N} f(\vartheta(C \otimes \mathbb{F}_p/\mathbb{F}_p))?$$

The **Generalized Sato-Tate Conjecture** is that this equidistribution (of the classes $\vartheta(C \otimes \mathbb{F}_p/\mathbb{F}_p)$ in $USp(2g)^{\#}$) holds whenever $C/\mathbb{Z}[1/N]$ has big arithmetic monodromy, in the sense that for every prime $l$, the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on $H^1(C \otimes \overline{\mathbb{Q}}, \mathbb{Q}_l)$ has image which is open in the group $GSp(2g, \mathbb{Q}_l)$ of symplectic similitudes. As an immediate application of Theorem 12.2.3, we find:

**Theorem.** *In the notations of Theorem 1.2.6, let $\varepsilon > 0$ and $g > N(\varepsilon)$. Suppose $C/\mathbb{Z}[1/N]$ is a curve of genus $g$ as above, which has big arithmetic monodromy. Suppose the generalized Sato-Tate conjecture holds. Then we have the inequality*

$$\lim_{X \to \infty} (1/\pi(X)) \sum_{p \leq X \text{ prime to } N} \mathrm{discrep}(\mu(\vartheta(C \otimes \mathbb{F}_p/\mathbb{F}_p)), \mu(\mathit{univ})) \leq g^{\varepsilon - 1/6}.$$

**Corollary.** *Suppose for each integer $g_i$ in an infinite subset $\Gamma$ of $\mathbb{Z}_{\geq 1}$, we are given an integer $N_i \geq 1$ and a curve $C_{g_i}/\mathbb{Z}[1/N_i]$ of genus $g_i$ which has big arithmetic monodromy. Suppose the generalized Sato-Tate conjecture holds. Then the double limit $\lim_{i \to \infty} \lim_{X \to \infty}$ of*

$$(1/\pi(X)) \sum_{p \leq X \text{ prime to } N_i} \mathrm{discrep}(\mu(\vartheta(C_{g_i} \otimes \mathbb{F}_p/\mathbb{F}_p)), \mu(\mathit{univ}))$$

*vanishes.*

**Question.** Notations and hypotheses as in the corollary, suppose that all $N_i$ have a common value $N$, cf. pages 12–13 of this Introduction for examples of such situations. What is the density of the set of primes $p$ not dividing $N$ for which

$$\lim_{g \to \infty \text{ in } \Gamma} \mathrm{discrep}(\mu(\vartheta(C_g \otimes \mathbb{F}_p/\mathbb{F}_p)), \mu(\mathrm{univ})) = 0?$$

Presumably this need not hold for **every** prime $p$ not dividing $N$.

Now let us turn to the opposite extreme, cases in which either we can prove or we expect that the spacing measure is far from the GUE measure.

We first give, for every odd prime $p$, a sequence of curves over the prime field $\mathbb{F}_p$ whose genera go to infinity and whose spacing measures converge to the delta measure $\delta_0$ supported at the origin. For each power $q = p^f$ of $p$, we consider the hyperelliptic curve $C_q/\mathbb{F}_p$ of equation

$$C_q : Y^2 = X^q - X.$$

This curve has genus $g$ given by $2g = q - 1$. Over $\mathbb{F}_q$, this curve admits the Artin-Schreier action $X \mapsto X + \alpha, Y \mapsto Y$ of the additive group of $\mathbb{F}_q$. If we pick any prime $l \neq p$, and decompose the cohomology group $H^1(C_q \otimes \overline{\mathbb{F}_q}, \overline{\mathbb{Q}_l})$ under this action, each of the $q - 1$ nontrivial additive characters $\psi$ of $\mathbb{F}_q$ occurs with multiplicity one. On the corresponding one-dimensional eigenspace, the Frobenius with respect to $\mathbb{F}_q$, $\mathrm{Frob}_q$, necessarily acts as a scalar, and that scalar is none other than minus the quadratic Gauss sum over $\mathbb{F}_q$:

$$-G_q(\psi, \chi_2) := -\sum_{x \text{ in } \mathbb{F}_q^\times} \psi(x)\chi_2(x),$$

where we have written $\chi_2$ for the quadratic character of $\mathbb{F}_q^\times$. Now it is elementary that the quadratic Gauss sum $G_q(\psi, \chi_2)$ with any nontrivial $\psi$ satisfies

$$(G_q(\psi, \chi_2))^2 = \chi_2(-1)q.$$

Moreover, both square roots occur as $\psi$ varies. Thus $\mathrm{Frob}_q$ has precisely 2 distinct eigenvalues on $H^1$. To see what this means for $\mathrm{Frob}_p$, write $q = p^f$. Then $\mathrm{Frob}_q$ is the $f$'th power of $\mathrm{Frob}_p$, and hence $\mathrm{Frob}_p$ has at most $2f$ distinct eigenvalues on $H^1$. This means that among the $q-1 = p^f - 1$ normalized spacings between the reciprocal zeroes of the zeta function of $C_q/\mathbb{F}_p$, all but at most $2f$ of the spacings are equal to zero. Since $2f/(p^f-1) \to 0$ as $f \to \infty$, we see that the spacing measures $\mu(C_1/\mathbb{F}_p)$ approach $\delta_0$ as $f \to \infty$. Since the GUE measure is absolutely continuous with respect to Lebesgue measure, it gives the origin mass zero. Hence we have

$$\mathrm{discrep}(\mu(C_q/\mathbb{F}_p), \mu(\mathrm{univ})) \geq 1 - 2f/(p^f - 1),$$

a crude quantification of the statement that $\mu(C_q/\mathbb{F}_p)$ is far from the GUE measure.

Here is another example, valid for any prime $p > 0$, of a sequence of curves over $\mathbb{F}_p$ whose genera go to infinity and whose spacing measures converge to $\delta_0$. For each power $q := p^f$ of $p$, consider the degree $q+1$ Fermat curve over $\mathbb{F}_p$, say, $\mathrm{Fermat}(q+1)/\mathbb{F}_p$, of homogeneous equation

$$X^{q+1} + Y^{q+1} = Z^{q+1}.$$

This curve has genus $g$ given by $2g = q(q-1)$. It is elementary that over $\mathbb{F}_{q^2}$, this curve has $1+q^3$ points. [Hint: for $x$ in $\mathbb{F}_{q^2}$, $x^{q+1}$ is its norm to $\mathbb{F}_q$, and the norm map is surjective.] But we readily compute that

$$1 + q^3 = 1 + q^2 + 2gq = 1 + q^2 - 2g(-q).$$

Thus the Weil bound is attained, and hence every eigenvalue of $\mathrm{Frob}_{q^2}$ on $H^1$ is $-q$. Therefore $\mathrm{Frob}_p$ has at most $2f$ distinct eigenvalues on $H^1$, while $\dim H^1 = p^f(p^f-1)$, and we conclude as in the previous example that the spacing measure $\mu(\mathrm{Fermat}(q+1)/\mathbb{F}_p)$ tends to $\delta_0$ as $f \to \infty$, and that

$$\mathrm{discrep}(\mu(\mathrm{Fermat}(q+1)/\mathbb{F}_p), \mu(\mathrm{univ})) \geq 1 - 2f/(p^f(p^f-1)).$$

We now turn to a case in which we **expect** the spacing measure to be far from the GUE measure. For each prime $l$, consider the modular curve $X_0(l)/\mathbb{Z}[1/l]$, whose genus $g_l$ is approximately $(l-1)/12$. Choose any prime $l'$. When we decompose $H^1 := H^1(X_0(l) \otimes \overline{\mathbb{Q}}, \overline{\mathbb{Q}}_{l'})$ under the Hecke operators, we find a direct sum of $g_l$ two-dimensional subspaces, corresponding to the $g_l$ different weight two normalized ($a_1(f) = 1$) Hecke-eigenforms $f = \sum_{n \geq 1} a_n(f) q^n$ on $X_0(l)$. For each such eigenform $f$, and each prime $p$ with $p \neq l$ and $p \neq l'$, the characteristic polynomial of $\mathrm{Frob}_p$ on the two-dimensional Hecke eigenspace in $H^1$ named by $f$ is $X^2 - a_p(f)X + p$. We know by Deligne that $|a_p(f)| \leq 2\,\mathrm{Sqrt}(p)$, so the two eigenvalues of $\mathrm{Frob}_p$ here are $\mathrm{Sqrt}(p)e^{\pm i\vartheta_p(f)}$, where $\vartheta_p(f)$ is the unique angle in $[0, \pi]$ such that $a_p(f) = 2\,\mathrm{Sqrt}(p)\cos(\vartheta_p(f))$. We denote by $\vartheta_p$ in $[0,\pi]^{g_l}$ the $g_l$-tuple of angles $\vartheta_p(f)$ indexed by eigenforms $f$, and we view $\vartheta_p$ as a conjugacy class in the product group $USp(2)^{g_l} = SU(2)^{g_l}$.

This Hecke-eigenvalue decomposition of $H^1$ is respected by $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and forces the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ to land in the subgroup of the $g_l$-fold self product $GL(2, \overline{\mathbb{Q}}_{l'})^{g_l}$ of $GL(2, \overline{\mathbb{Q}}_{l'})$ with itself consisting of elements $(A_1, \ldots, A_{g_l})$ all of which have equal determinants. According to Ribet [**Rib**, 7.18], the image of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is Zariski dense in this group. In view of Ribet's result, the natural Sato-Tate conjecture for $X_0(l)$ is the assertion that the conjugacy classes $\{\vartheta_p\}_{p \neq l}$ in the product group $SU(2)^{g_l}$ are equidistributed with respect to Haar measure in the sense that for any $\mathbb{C}$-valued continuous central function $h$ on $SU(2)^{g_l}$, the integral

$$\int_{SU(2)^{g_l}} h(A)\,dA = \int_{[0,\pi]^{g_l}} h(x_1,\ldots,x_{g_l}) \prod_{i=1}^{g_l} (2/\pi)\sin^2(x_i)\,dx_i$$

can be computed as the limit

$$\lim_{X\to\infty} (1/\pi(X)) \sum_{p\le X, p\ne l} h(\vartheta_p).$$

Let us admit the truth of this Sato-Tate conjecture for $X_0(l)$. We denote by

$$F(x) := (2/\pi)\int_{[0,x]} \sin^2(t)\,dt = (2x - \sin(2x))/2\pi$$

the bijection from $[0,\pi]$ to $[0,1]$ which carries the measure $(2/\pi)\sin^2(t)\,dt = dF$ on $[0,\pi]$ to uniform measure $dx$ on $[0,1]$. [We call $F$ the straightening function for the measure in question.] Given an element $\vartheta$ in $[0,\pi]^{g_l}$, we denote by $F(\vartheta)$ in $[0,1]^{g_l}$ the result of applying $F$ component by component. The Sato-Tate conjecture for $X_0(l)$ says precisely that the $g_l$-tuples $\{F(\vartheta_p)\}_{p\ne l}$ in $[0,1]^{g_l}$ are equidistributed in $[0,1]^{g_l}$ for uniform measure. Arrange the components of $F(\vartheta_p)$ in increasing numerical order, say

$$0 \le F(\vartheta_p)_1 \le F(\vartheta_p)_2 \le \cdots \le F(\vartheta_p)_{g_l} \le 1.$$

The $F$-straightened spacing measure $\mu_F(X_0(l)\otimes\mathbb{F}_p/\mathbb{F}_p)$ attached to $X_0(l)\otimes\mathbb{F}_p/\mathbb{F}_p$ is the measure of total mass $1-1/g_l$ on $\mathbb{R}$ which gives each of the $g_l - 1$ normalized spacings $s_i := g_l(F(\vartheta_p)_{i+1} - F(\vartheta_p)_i)$ the mass $1/g_l$. An elementary analysis of spacings between points in an interval which are randomly chosen for Lebesgue measure shows that in this kind of question, the limiting answer is not the rather exotic GUE, but rather the much more elementary and familiar exponential distribution $\mu(\text{Poisson})$, the measure on $\mathbb{R}$ supported in $\mathbb{R}_{\ge 0}$ and given there by $e^{-x}\,dx$.

**Theorem.** *Assume the Sato-Tate conjecture above for all the modular curves $X_0(l)/\mathbb{Z}[1/l], l$ any prime. Then the $F$-straightened spacing measures*

$$\mu_F(X_0(l)\otimes\mathbb{F}_p/\mathbb{F}_p)$$

*are, for large $l$, very near the Poisson measure $\mu(\text{Poisson})$ for most primes $p\ne l$. More precisely, the double limit $\lim_{l\to\infty}\lim_{X\to\infty}$ of*

$$(1/\pi(X)) \sum_{p\le X, p\ne l} \operatorname{discrep}(\mu_F(X_0(l)\otimes\mathbb{F}_p/\mathbb{F}_p), \mu(\text{Poisson}))$$

*vanishes.*

**Question.** Is it true that for each prime $p$, we have

$$\lim_{l\to\infty, l\ne p} \operatorname{discrep}(\mu_F(X_0(l)\otimes\mathbb{F}_p/\mathbb{F}_p), \mu(\text{Poisson})) = 0?$$

We now discuss another aspect of our work, the distribution in families of "low-lying zeroes". [This terminology "low-lying zeroes" is inspired by the number field picture, where we expect all the nontrivial zeroes to lie on a single vertical line, and we measure height from the real axis. In the finite field case, where the normalized zeroes lie on the unit circle, it would be more accurate to speak of "zeroes near 1".] For simplicity, we will discuss only the case of curves. Recall that the zeta function of a genus $g$ curve $C/\mathbb{F}_q$ is of the form $P(T)/(1-T)(1-qT)$, for $P$ a polynomial of degree $2g$ with the property that the auxiliary polynomial $P(T/\operatorname{Sqrt}(q))$ has all its $2g$ roots on the unit circle, and its $2g$ roots can be partitioned into $g$ pairs of inverses $(\xi, 1/\xi)$. So we may write $P(T)$ as

$$P(T) = \prod_{j=1}^{g} (1 - \alpha_j T)(1 - \alpha_{-j} T),$$

with

$$\alpha_{-j} = \overline{\alpha}_j, \qquad \alpha_j \alpha_{-j} = q.$$

If we pick the $\alpha_j$ (rather than the $\alpha_{-j}$) to lie in the upper half plane, we have

$$\alpha_j = \operatorname{Sqrt}(q) e^{i\varphi_j}, \qquad 0 \leq \varphi_j \leq \pi,$$

and with suitable renumbering the $g$ angles $\varphi_j$ in $[0, \pi]$ may be assumed to be in increasing order:

$$0 \leq \varphi_1 \leq \varphi_2 \leq \cdots \leq \varphi_g \leq \pi.$$

With this numbering, we refer to $(g/\pi)\varphi_j$ as the $j$'th normalized angle attached to the curve $C/\mathbb{F}_q$, or, if we like, attached to the unitarized Frobenius conjugacy class $\vartheta(C/\mathbb{F}_q)$.

More generally, given any element $A$ in $USp(2g)$, we have

$$\det(1 - TA) = \prod_{j=1}^{g} (1 - Te^{i\varphi_j})(1 - Te^{-i\varphi_j})$$

for a unique sequence of angles $0 \leq \varphi_1 \leq \varphi_2 \leq \cdots \leq \varphi_g \leq \pi$. For each integer $1 \leq j \leq g$, the function on $USp(2g)$ defined by

$$A \mapsto \varphi_j := \varphi_j(A)$$

is a continuous central function from $USp(2g)$ to $[0, \pi]$. We refer to $(g/\pi)\varphi_j(A)$ as the $j$'th normalized angle attached to the conjugacy class of the element $A$. The function

$$A \mapsto (g/\pi)\varphi_j(A)$$

is a continuous central function from $USp(2g)$ to $\mathbb{R}_{\geq 0}$.

If we take the direct image of normalized Haar measure $\mu_{\text{Haar}}$ on $USp(2g)$ by the map $A \mapsto (g/\pi)\varphi_j(A)$, we obtain a probability measure on $\mathbb{R}$ supported in $\mathbb{R}_{\geq 0}$, which we denote $\nu(j, USp(2g))$.

There are analogous constructions for the other classical groups, cf. 6.9 for the details, which give rise to probability measures $\nu(j, G(N))$ for $1 \leq j \leq N$ and $G(N)$ any of $U(N), USp(2N), SO(2N+1), SO(2N), O_-(2N+1), O_-(2N+2)$, all on the real line and all supported in $\mathbb{R}_{\geq 0}$. Now unlike the spacing measures, which were "universal" in the sense that the large $N$ limit existed and was independent of which sequence of $G(N)$'s we ran through, these $\nu$ measures do depend on the sequence of $G(N)$'s chosen.

**Theorem** (7.5.5, 7.5.6)**.** *For each integer $j \geq 1$, there exist*[2] *three probability measures $\nu(j), \nu(-, j)$, and $\nu(+, j)$ on $\mathbb{R}$, supported in $\mathbb{R}_{\geq 0}$ and having continuous*

[2]In fact, these measures all have densities, which for $j = 1$ are shown in Appendix: Graphs.

*CDF's, such that we have the following large $N$ limit statements for convergence in the sense of uniform convergence of CDF's*:

$$\lim_{N\to\infty} \nu(j, G(N)) = \begin{cases} \nu(j), & \text{if } G(N) = U(N), \\ \nu(+, j), & \text{if } G(N) = SO(2N) \text{ or } O_-(2N+1), \\ \nu(-, j), & \text{if } G(N) = USp(2N), SO(2N+1), O_-(2N+2). \end{cases}$$

Here is a convenient mnemonic to remember which sign in $\nu(\pm, j)$ is given by which orthogonal group series: the sign is the common sign of $\det(-A)$ for $A$ in either $SO(2N)$ or $O_-(2N+1)$ or $SO(2N+1)$ or $O_-(2N+2)$. Experts will recognize this sign as being the **sign in the functional equation** of $P(T) := \det(1 - TA)$ under $T \mapsto 1/T$, which for orthogonal $A$ is

$$T^{\deg(P)} P(1/T) = \det(-A)P(T).$$

To see what this means concretely for curves, fix an integer $N \geq 1$, and suppose for each genus $g \geq 1$ we are given a proper smooth family $\pi : \mathcal{C}_g \to S_g$ of genus $g$ curves, parameterized by a scheme $S_g$ which is smooth and surjective over $\mathrm{Spec}(\mathbb{Z}[1/N])$ with geometrically connected fibres. We further assume that for every prime number $p$ which does not divide $N$, the geometric monodromy group of the family of curves

$$\pi \otimes \mathbb{F}_p : \mathcal{C}_g \otimes \mathbb{F}_p \to S_g \otimes \mathbb{F}_p$$

in characteristic $p$ is the full symplectic group $Sp(2g)$. For example, we might take $N = 2$, and for $\mathcal{C}_g/S_g$ the universal family of hyperelliptic curves $Y^2 = f_{2g+1}(X)$ parameterized by the space $S_g := \mathcal{H}_{2g+1}$ of monic polynomials $f_{2g+1}$ of degree $2g+1$ with invertible discriminant.

Let now $h(x)$ be any $\mathbb{C}$-valued continuous function on $\mathbb{R}$. By Deligne's equidistribution theorem, for each integer $j \geq 1$, and each genus $g \geq j$, we can compute the integral

$$\int_{\mathbb{R}} h \, d\nu(j, USp(2g)) := \int_{USp(2g)} h((g/\pi)\varphi_j(A)) \, dA$$

as the limit

$$\lim_{q\to\infty,\text{ prime to } N} (1/|S_g(\mathbb{F}_q)|) \sum_{s \text{ in } S(\mathbb{F}_q)} h((g/\pi)\varphi_j(\vartheta(\mathcal{C}_{g,s}/\mathbb{F}_q))).$$

Using the theorem above, about the large $N$ limit of the measures $\nu(j, USp(2g))$ being $\nu(-, j)$, we find that, for $h(x)$ any bounded $\mathbb{C}$-valued continuous function on $\mathbb{R}$, we can compute the integral

$$\int_{\mathbb{R}} h \, d(\nu)(-, j) = \lim_{g\to\infty} \int_{USp(2g)} h((g/\pi)\varphi_j(A)) \, dA$$

as the double limit

$$\lim_{g\to\infty} \lim_{q\to\infty,\text{ prime to } N} (1/|S_g(\mathbb{F}_q)|) \sum_{s \text{ in } S(\mathbb{F}_q)} h((g/\pi)\varphi_j(\vartheta(\mathcal{C}_{g,s}/\mathbb{F}_q))).$$

If we look instead at universal families of hypersurfaces of even dimension, and average over those whose functional equation (for the factor of its zeta function corresponding to the primitive part of middle dimensional cohomology) has a fixed sign $\varepsilon = \pm 1$, we get double limit formulas for $\int_{\mathbb{R}} h \, d\nu(\varepsilon, j)$. [Universal families of odd dimensional hypersurfaces have monodromy group $Sp$, so will lead only to double limit formulas for $\int_{\mathbb{R}} h\, d\nu(-,j)$.]

We now leave the realm of what is proven, and discuss what might be true if, in the double limit formulas above, we were to omit the inner limit over $q$. Again to fix ideas, we return to the case of curves. Fix an integer $N \geq 1$, and an infinite subset $\Gamma$ of $\mathbb{Z}_{\geq 1}$. Suppose for each genus $g$ in $\Gamma$ we are given a proper smooth family $\pi : \mathcal{C}_g \to S_g$ of genus $g$ curves, parameterized by a scheme $S_g$ which is smooth and surjective over $\mathrm{Spec}(\mathbb{Z}[1/N])$ with geometrically connected fibres. We further assume that for every prime number $p$ which does not divide $N$, and for every $g$ in $\Gamma$, the geometric monodromy group of the family of curves

$$\pi \otimes \mathbb{F}_p : \mathcal{C}_g \otimes \mathbb{F}_p \to S_g \otimes \mathbb{F}_p$$

in characteristic $p$ is the full symplectic group $Sp(2g)$. To further simplify matters, we assume also that $S_g(\mathbb{F}_p)$ is nonempty for every prime $p$ not dividing $N$, and for every genus $g$ in $\Gamma$.

We say that this collection of families $\{\mathcal{C}_g/S_g\}_{g \text{ in } \Gamma}$ weakly calculates the measure $\nu(-,j)$ if, for every finite field $\mathbb{F}_q$ of characteristic prime to $N$, and for every bounded continuous function $h$ on $\mathbb{R}$, we can calculate $\int_{\mathbb{R}} h\, d\nu(-,j)$ as

$$\lim_{n\to\infty} \left(1/\sum_{g\leq n \text{ in } \Gamma} |S_g(\mathbb{F}_q)|\right) \sum_{g\leq n \text{ in } \Gamma} \sum_{s \text{ in } S_g(\mathbb{F}_q)} h((g/\pi)\varphi_j(\vartheta(\mathcal{C}_{g,s}/\mathbb{F}_q))).$$

We say that this collection of families $\{\mathcal{C}_g, S_g\}_{g \text{ in } \Gamma}$ strongly calculates the measure $\nu(-,j)$ if, for every finite field $\mathbb{F}_q$ of characteristic prime to $N$, and for every bounded continuous function $h$ on $\mathbb{R}$, we can calculate $\int_{\mathbb{R}} h\, d\nu(-,j)$ as

$$\lim_{g\to\infty \text{ in } \Gamma} (1/|S_g(\mathbb{F}_q)|) \sum_{s \text{ in } S_g(\mathbb{F}_q)} h((g/\pi)\varphi_j(\vartheta(\mathcal{C}_{g,s}/\mathbb{F}_q))).$$

It is elementary that if $\{\mathcal{C}_g/S_g\}_{g \text{ in } \Gamma}$ strongly calculates the measure $\nu(-,j)$, then $\{\mathcal{C}_g/S_g\}_{g \text{ in } \Gamma}$ weakly calculates it as well. If for every finite field $\mathbb{F}_q$ of characteristic prime to $N$, the cardinalities $|S_g(\mathbb{F}_q)|$ grow fast enough that the ratios

$$(1/|S_g(\mathbb{F}_q)|) \sum_{\gamma\leq g \text{ in } \Gamma} |S_\gamma(\mathbb{F}_q)|$$

stay bounded (as $g$ varies over $\Gamma$, $q$ fixed: the bound can vary with $q$), then the two notions, strong and weak calculation of $\nu(-,j)$, are equivalent.

**Conjecture.** *Fix an integer $N \geq 1$, and an infinite subset $\Gamma$ of $\mathbb{Z}_{\geq 1}$. Suppose for each genus $g$ in $\Gamma$ we are given a proper smooth family $\pi : \mathcal{C}_g \to S_g$ of genus $g$ curves, parameterized by a scheme $S_g$ which is smooth and surjective over* $\mathrm{Spec}(\mathbb{Z}[1/N])$ *with geometrically connected fibres. Suppose that for every $g$ in $\Gamma$ and for every prime number $p$ which does not divide $N$, the geometric monodromy group of the family of curves*

$$\pi \otimes \mathbb{F}_p : \mathcal{C}_g \otimes \mathbb{F}_p \to S_g \otimes \mathbb{F}_p$$

*in characteristic $p$ is the full symplectic group $Sp(2g)$. Suppose also that $S_g(\mathbb{F}_p)$ is nonempty for every prime $p$ not dividing $N$, and for every genus $g$ in $\Gamma$. Then for every integer $j \geq 1$, the collection of families $\{\mathcal{C}_g/S_g\}_{g \text{ in } \Gamma}$ weakly calculates the measure $\nu(-,j)$. Moreover, if in addition we have* $\lim_{g\to\infty \text{ in } \Gamma} |S_g(\mathbb{F}_q)| = \infty$

*for every finite field $\mathbb{F}_q$ of characteristic prime to $N$, then the collection of families $\{\mathcal{C}_g/S_g\}_{g \text{ in } \Gamma}$ strongly calculates the measure $\nu(-,j)$.*[3]

Let us give some examples of situations $N, \Gamma, \{\mathcal{C}_g/S_g\}_{g \text{ in } \Gamma}$ which satisfy all of the hypotheses imposed in the statement of the conjecture.

1) $N = 1, \Gamma = \mathbb{Z}_{\geq 1}, \mathcal{M}_{g,3K}$, universal family of curves with $3K$ structure (cf. 10.6).

2) $N = 2, \Gamma = \mathbb{Z}_{\geq 1}, \mathcal{H}_{2g+1}$ or $\widetilde{\mathcal{H}}_{2g+1}$, family $Y^2 = f_{2g+1}(X)$ (cf. 10.1.18.3–4).

2bis) $N = 2, \Gamma = \mathbb{Z}_{\geq 1}, \mathcal{H}_{2g+2}$ or $\widetilde{\mathcal{H}}_{2g+2}$, family $Y^2 = f_{2g+2}(X)$ (cf. 10.1.18.4–5).

3) $N = 2, \Gamma = \mathbb{Z}_{\geq 1}$: write $2g$ to the base 2 as $\sum_{\text{some } \alpha \geq 1} 2^\alpha$, and define $F_{2,2g}(X)$ as the corresponding product of cyclotomic polynomials $\Phi_{2^{\alpha+1}}(X) = (X^{2^\alpha} + 1)$:

$$F_{2,2g}(X) := \prod_{\alpha \text{ in base 2 expansion of } 2g} (X^{2^\alpha} + 1).$$

Take for $\mathcal{C}_g/S_g$ the one-parameter ("$T$") family of hyperelliptic curves of equation $Y^2 = (X - T)F_{2,2g}(X)$.

4) $N = 2l$ for $l$ a prime, $\Gamma =$ those integers $g \geq 1$ such that in the base $l$ expression of $2g$, all the digits are either 0 or $l - 1$: write $2g$ to the base $l$ as $\sum_{\text{some } \alpha \geq 0} (l-1)l^\alpha$, and define $F_{l,2g}(X)$ as the corresponding product of cyclotomic polynomials $\Phi_{l^{\alpha+1}}(X)$. Take for $\mathcal{C}_g/S_g$ the one-parameter ("$T$") family of hyperelliptic curves of equation $Y^2 = (X - T)F_{l,2g}(X)$. [Of course, if we take $l = 2$ in this example, we find example 3).]

5) $N = 2l$ for $l$ a prime, $\Gamma =$ those $g \geq 1$ such that $2g = (l-1)l^\alpha$ for some integer $\alpha \geq 0$. Take for $\mathcal{C}_g/S_g$ the one-parameter ("$T$") family of hyperelliptic curves of equation $Y^2 = (X - T)\Phi_{l^{\alpha+1}}(X)$.

Notice that in examples 1), 2), and 2bis), we do have

$$\lim_{g \to \infty \text{ in } \Gamma} |S_g(\mathbb{F}_q)| = \infty.$$

But in examples 3) through 5), the parameter space $S_g$ is always a Zariski open set in the affine line $\mathbb{A}^1$, so $|S_g(\mathbb{F}_q)| \leq q$ is uniformly bounded. The relevant sets $S_g(\mathbb{F}_q)$ are always nonempty, since both 0 and $\pm 1$ are always allowed parameter values.

The conjecture in the examples 2) and 2bis) ($\widetilde{\mathcal{H}}$ version) can be viewed as a statement about the low-lying zeroes of the $L$-functions of all quadratic extensions of $\mathbb{F}_q(X)$. So seen, it has an analogue for Dirichlet $L$-series with quadratic character, which we will now formulate. Thus we take a quadratic extension $K/\mathbb{Q}$, of discriminant $D_K$, corresponding to the quadratic Dirichlet character $\chi_K$. We assume that $L(s, \chi_K)$ satisfies the Riemann Hypothesis. We write the nontrivial zeroes of $L(s, \chi_K)$ (which by the (even!) functional equation occur in conjugate pairs) as $1/2 \pm i\gamma_{K,j}$ with $0 \leq \gamma_{K,1} \leq \gamma_{K,2} \leq \gamma_{K,3} \leq \cdots$.

**Conjecture.** *The low-lying zeroes of Dirichlet $L$-functions with quadratic character weakly calculate the measure $\nu(-,j)$, in the following sense. For any integer $j \geq 1$, and for any compactly supported continuous $\mathbb{C}$-valued function $h$ on $\mathbb{R}$, we can calculate the integral $\int_{\mathbb{R}} h\, d\nu(-,j)$ as*

$$\lim_{X \to \infty} (1/|\{K \textit{ with } D_K \leq X\}|) \sum_{K \textit{ with } D_K \leq X} h(\gamma_{K,j} \log(D_K)/2\pi).$$

[3]This second part of the conjecture, about strong calculation, seems to us more speculative than the first part.

The measure $\nu(-,1)$ has a density which has the remarkable property of vanishing to second order at the origin, see Appendix: Graphs. Thus our conjecture implies that $L(s,\chi_K)$ rarely has a zero at or even near the point $1/2$. This sort of behaviour was observed at a crude level by Hazelgrove, who seems to have been the first to experiment numerically with zeroes of Dirichlet $L$-functions. The above conjecture offers a precise version, and hints at the existence, in the global case, of some remarkable analogue, yet to be discovered, of the symplectic monodromy which in the function field case binds together the $L$-functions of quadratic extensions of $\mathbb{F}_q(X)$.

For a global situation in which both measures $\nu(-,1)$ and $\nu(+,j)$ arise, one has elliptic curves over, say, $\mathbb{Q}$, where some analogue of orthogonal monodromy seems to enter. Let us grant that all elliptic curves over $\mathbb{Q}$ are modular, so that their $L$-functions are entire, and let us assume further that these $L$-functions have all their (nontrivial) zeroes on the line $\mathrm{Re}(s) = 1$. For each integer $n \geq 1$, denote by $\mathcal{E}_n$ the set of $\mathbb{Q}$-isogeny classes of elliptic curves over $\mathbb{Q}$ with conductor $N_E = n$, and by $\mathcal{E}_{n,+}$ and $\mathcal{E}_{n,-}$ the subsets of $\mathcal{E}_n$ consisting of those curves whose $L$-functions have an even or odd functional equation, respectively.

If $E/\mathbb{Q}$ has an even functional equation, the nontrivial zeroes of $L(s,E/\mathbb{Q})$ occur in conjugate pairs, and we write them as

$$1 \pm i\gamma_{K,j} \text{ with } 0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq \cdots .$$

If $E/\mathbb{Q}$ has an odd functional equation, then 1 is a zero of $L(s,E/\mathbb{Q})$, and the remaining nontrivial zeroes of $L(s,E/\mathbb{Q})$ occur in conjugate pairs: we write the remaining zeroes as $1 \pm i\gamma_{E,j}$ with $0 \leq \gamma_{E,1} \leq \gamma_{E,2} \leq \gamma_{E,3} \leq \cdots$.

**Conjecture.** *The low-lying zeroes of $L$-functions of elliptic curves over $\mathbb{Q}$ weakly calculate the measure $\nu(\pm j)$, in the following sense. For any integer $j \geq 1$, and for any compactly supported continuous $\mathbb{C}$-valued function $h$ on $\mathbb{R}$, we can calculate the integrals $\int_{\mathbb{R}} h\, d\nu(\pm,j)$ as follows:*

$$\int_{\mathbb{R}} h\, d\nu(-,j) = \lim_{X\to\infty} \left(1/\sum_{n\leq X} |\mathcal{E}_{n,-}|\right) \sum_{n\leq X, E \text{ in } \mathcal{E}_{n,-}} h(\gamma_{E,j}\log(N_E)/2\pi),$$

*and*

$$\int_{\mathbb{R}} h\, d(+,j) = \lim_{X\to\infty} \left(1/\sum_{n\leq X} |\mathcal{E}_{n,+}|\right) \sum_{n\leq X, E \text{ in } \mathcal{E}_{n,+}} h(\gamma_{E,j}\log(N_E)/2\pi).$$

As already remarked above, the measure $\nu(-,1)$ has a density which vanishes to second order at the origin. So the conjecture for $j = 1$ predicts that among $L$-functions of elliptic curves $E/\mathbb{Q}$ with odd functional equation, most should have only a simple zero at $s = 1$ and no zeroes very near to $s = 1$. In particular, just using the absolute continuity of $\nu(-,1)$ with respect to Lebesgue measure, the conjecture implies that

$$\lim_{X\to\infty} \left(1/\sum_{n\leq X} |\mathcal{E}_{n,-}|\right) \sum_{n\leq X} |\{E \text{ in } \mathcal{E}_{n,-} \text{ with } \gamma_{E,1} = 0\}| = 0.$$

The measure $\nu(+,1)$ also has a density. Its density, unlike that of $\nu(-,1)$, is nonzero at the origin. Nonetheless, $\nu(+,1)$ is absolutely continuous with respect to

Lebesgue measure on $\mathbb{R}$. So our conjecture for $j = 1$ implies that among $E/\mathbb{Q}$ with even functional equation, most should have no zero at $s = 1$, in the sense that

$$\lim_{X\to\infty}\left(1/\sum_{n\le X}|\mathcal{E}_{n,+}|\right)\sum_{n\le X}|\{E \text{ in } \mathcal{E}_{n,+} \text{ with } \gamma_{E,1}=0\}| = 0.$$

If we also admit the conjecture of Birch and Swinnerton-Dyer that for $E/Q$ the rank of the Mordell-Weil group $E(\mathbb{Q})$ is equal to the order of vanishing of $L(s, E/\mathbb{Q})$ at $s = 1$, then the above consequences of the conjecture imply in turn that

$$\lim_{X\to\infty}\left(1/\sum_{n\le X}|\mathcal{E}_{n,-}|\right)\sum_{n\le X}|\{E \text{ in } \mathcal{E}_{n,-} \text{ with rank } (E(\mathbb{Q}))>1\}| = 0,$$

$$\lim_{X\to\infty}\left(1/\sum_{n\le X}|\mathcal{E}_{n,+}|\right)\sum_{n\le X}|\{E \text{ in } \mathcal{E}_{n,+} \text{ with rank } (E(\mathbb{Q}))>0\}| = 0.$$

These last two statements together imply in turn that

$$\lim_{X\to\infty}\left(1/\sum_{n\le X}|\mathcal{E}_{n}|\right)\sum_{n\le X}|\{E \text{ in } \mathcal{E}_{n} \text{ with rank } (E(\mathbb{Q}))\ge 2\}| = 0,$$

or, in words, zero percent of elliptic curves over $\mathbb{Q}$ have rank 2 or more. The truth of this has recently been called into question; cf. [**Kra-Zag**], [**Fer**], [**Sil**]. However, if our conjecture is correct, then the "contradictory" data is simply an artifact of too restricted a range of computation.

This ends our venture into speculation and conjecture. We now turn to a summary of the contents of this book. The book falls naturally into three parts. The first part, which consists of Chapters 1 through 8, is devoted to the theory of spacing measures on large classical groups. Chapter 1 is devoted to defining the spacing measures which are our main object of study, and to stating the main results about them. In Chapter 2, we define "naive" versions of the spacing measures which we find more amenable to combinatorial analysis. We then formulate versions of our main results for these "naive" spacing measures, and show that they imply the main results stated in the first chapter. The remainder of Chapter 2, along with all of Chapters 3 and 4, is devoted to successive reduction steps (2.9.1, 3.0.1, 3.1.9, 4.2.2–4) in proving the main results. By the end of Chapter 4 we are reduced to proving the three estimates of 4.2.2 and the "tail estimate" 3.1.9, iv). In Chapter 5, we first recall Weyl's explicit formulas for Haar measure on the classical groups. We then combine Weyl's formulas with a method of orthogonal polynomials (5.1.3) which goes back to Gaudin [**Gaudin**]. Thus armed, we establish (in 5.8.3, 5.10.3 and 5.11.2) the three estimates of 4.2.2. Chapter 6 is devoted to the proof of the tail estimate of 3.1.9, iv). To prove it, we introduce "eigenvalue location measures" in 6.9, and in 6.10.5 we give a tail estimate for the first of these measures. We then (6.11, 6.12) relate these measures to spacing measures, which allows us in 6.13 to prove the required tail estimate 3.1.9, iv). At this point, all (save 1.7.6) of the results announced in Chapters 1 and 2 have been proven. The remainder of Chapter 6 explores multi-variable versions of the eigenvalue location measures. In Chapter 7 we form generating series out of the limit spacing measures, we prove 1.7.6, and we relate these generating series, in the case of one variable, to certain infinite-dimensional Fredholm determinants, which are themselves large $N$ limits of finite-dimensional Fredholm determinants. We use this theory to construct large $N$ limits of the (one-variable) eigenvalue location measures, and establish the relations between the limit spacing measures and the limit eigenvalue location measures in one variable. Chapter 8 is devoted to a discussion of these same questions in several variables.

The second part of the book, which consists of Chapters 9 through 11, is devoted to algebro-geometric situations over finite fields which, by means of Deligne's equidistribution theorem and the determination of monodromy groups, provide us with "Frobenius conjugacy classes" in large compact classical groups which are suitably equidistributed for Haar measure. In Chapter 9, we give various "abstract" versions of Deligne's equidistribution theorem, in the language of pure lisse sheaves. Roughly speaking, these theorems assert an equidistribution (for Haar measure) of Frobenius conjugacy classes in the space of conjugacy classes of a maximal compact subgroup of the geometric monodromy group attached to the situation at hand. In Chapters 10 and 11, we prove various families (of curves, of abelian varieties, of hypersurfaces, of Kloosterman sums) to have geometric monodromy groups which are large classical groups.

The third part of the book, which consists of Chapters 12 and 13, applies the theory of the first part of the book to the families proven to have big monodromy in the second part of the book. Chapter 12 looks specifically at GUE discrepancies in these families, and Chapter 13 looks at the distribution of low-lying eigenvalues in these same families.

The book concludes with two appendices. The first appendix, Densities, develops an approach to eigenvalue location measures through densities, determines their large $N$ limits, and presents a result of Harold Widom, that the large $N$ limits of the densities, and hence of the eigenvalue location measures, for the groups $SU(N)$ exist and are equal to those for $U(N)$. The second appendix, Graphs, shows the densities of the GUE measure and of a few of the one-variable eigenvalue location measures.

# Notation Index

# Subject Index