BAGCHI'S THEOREM FOR FAMILIES OF AUTOMORPHIC FORMS

E. KOWALSKI

Abstract. We prove a version of Bagchi's Theorem and of Voronin's Universality Theorem for the family of primitive cusp forms of weight 2 and prime level, and discuss under which conditions the argument will apply to a general reasonable family of automorphic L-functions.

1. Introduction

The first "universality theorem" for Dirichlet series is Voronin's Theorem [18] for the Riemann zeta function, which states that for any $r < 1/4$, any continuous function φ defined and non-vanishing on the disc $|s - 3/4| \leqslant r$ in $\mathbf{C}$, which is holomorphic in the interior, and any $\varepsilon > 0$, there exists $t \in \mathbf{R}$ such that

$$\max_{|s-3/4|\leqslant r} |\zeta(s + it) - \varphi(s)| < \varepsilon.$$

In other words, up to arbitrary precision, any function φ can be approximated by some vertical translate of the Riemann zeta function.

Bagchi, in his thesis [1], provided a clear conceptual explanation of this result, as the combination of two independent statements:

– Viewing translates of the Riemann zeta function by $t \in [-T, T]$ as random variables with values in a space of holomorphic function on the disc, Bagchi proves that these random variables converge in law, as $T \to +\infty$, to a natural random Dirichlet series, which is also expressed as a random Euler product;
– Computing the support of the limiting random Dirichlet series, and checking that it contains the space of nowhere vanishing holomorphic functions on the disc, the universality theorem follows easily.

The key step, from our point of view, is the first part, which we call *Bagchi's Theorem*. Indeed, once the convergence in law is known, it follows that there is "some" universality statement, with respect to the functions in the support of the limiting random Dirichlet series. The second step makes this support explicit. (This might be compared with Deligne's Equidistribution Theorem, as applied to families of exponential sums for instance: Deligne's Theorem shows that there is always *some* equidistribution of these sums.)

The goal of this note is to give a first example of a genuinely higher-degree statement of this type, and to deduce the corresponding universality statement. We will also indicate a general principle that should apply in many more cases.

Key words and phrases. Modular forms, L-functions, Bagchi's Theorem, Voronin's Theorem, random Dirichlet series.

Partially supported by a DFG-SNF lead agency program grant (grant 200021L_153647).

Theorem 1.1 (Universality in level aspect)**.** *For q prime $\geqslant 17$, let $S_2(q)^*$ be the non-empty*[1] *finite set of primitive cusp forms for $\Gamma_0(q)$ with weight 2 and trivial nebentypus. For $f \in S_2(q)^*$, let $L(f,s)$ denote its Hecke L-function*

$$L(f,s) = \sum_{n \geqslant 1} \lambda_f(n) n^{-s},$$

normalized so that the critical line is $\mathrm{Re}(s) = \frac{1}{2}$.

For any real number $r < \frac{1}{4}$, let D be the open disc centered at $3/4$ with radius r. Then for any continuous function $\varphi : \bar{D} \to \mathbf{C}$ which is holomorphic and non-vanishing in D and satisfies

$$\varphi(\sigma) > 0 \text{ for } \sigma \in D \cap \mathbf{R}, \tag{1.1}$$

we have

$$\liminf_{q \to +\infty} \frac{1}{|S_2(q)^*|} |\{f \in S_2(q)^* \mid \|L(f,\cdot) - \varphi\|_\infty < \varepsilon\}| > 0$$

for any $\varepsilon > 0$, where the L^∞ norm is the norm on $\bar{D}$.

The main difference with previous results involving cusp forms (the first one being due to Laurinčikas and Matsumoto [13]) is that we do not fix one such L-function $L(f,s)$ and consider shifts (or twists) $L(f, s+it)$ or $L(f \times \chi, s)$, but rather we average over the discrete family of primitive forms in $S_2(q)^*$. It is also important to remark that the condition (1.1) is necessary for a function on D to be approximated by L-functions $L(f,s)$ with $f \in S_2(q)^*$. (We will give more general statements where the discs D are replaced with more general compact sets in the strip $\frac{1}{2} < \sigma < 1$).

We will prove this Theorem in Section 2, after stating the results generalizing the two steps of Bagchi's strategy for the zeta function. The proof of Bagchi's Theorem for this family is an analogue of a proof for the Riemann zeta function that is simpler than Bagchi's proof (it avoids both the use of the ergodic theorem and any tightness or weak-compactness argument).

In Section 5, we discuss very briefly how this strategy can in principle be applied to very general families of L-functions, as defined in [8].

Acknowledgements. The simple proof of Bagchi's Theorem for the Riemann zeta funtion, that we generalize here to modular forms, was found during a course on probabilistic number theory at ETH Zürich in the Fall Semester 2015; thanks are due to the students who attended this course for their interest and remarks, and to B. Löffel for assisting with the exercises (see [9] for a draft of the lecture notes for this course, especially Chapter 3).

Thanks to the referee for carefully reading the text, and in particular for pointing out a number of confusing mistakes in references to the literature.

Notation. As usual, $|X|$ denotes the cardinality of a set. By $f \ll g$ for $x \in X$, or $f = O(g)$ for $x \in X$, where X is an arbitrary set on which f is defined, we mean synonymously that there exists a constant $C \geqslant 0$ such that $|f(x)| \leqslant Cg(x)$ for all $x \in X$. The "implied constant" is any admissible value of C. It may depend on the set X which is always specified or clear in context. We write $f \asymp g$ if $f \ll g$ and $g \ll f$ are both true.

[1] We assume $q \geqslant 17$ to ensure this property; it also holds for $q = 11$.

We use standard probabilistic terminology: a probability space $(\Omega, \Sigma, \mathbf{P})$ is a triple made of a set Ω with a σ-algebra Σ and a measure $\mathbf{P}$ on Σ with $\mathbf{P}(\Omega) = 1$. We denote by $\mathbf{E}(X)$ the expectation on Ω. The law of a random variable X is the measure ν on the target space of X defined by $\nu(A) = \mathbf{P}(X \in A)$. If $A \subset \Omega$, then $\mathbf{1}_A$ is the characteristic function of A.

2. Equidistribution and universality for modular forms in the level aspect

We will prove Theorem 1.1 by combining the results of the following two steps, each of which will be proved in a forthcoming section. Throughout, we assume that q is a prime number $\geqslant 17$.

Step 1. (Equidistribution; Bagchi's Theorem) For q prime, we view the finite set $S_2(q)^*$ as a probability space with the probability measure proportional to the "harmonic" measure where $f \in S_2(q)^*$ has weight

$$\frac{1}{\langle f, f \rangle}$$

in terms of the Petersson inner product. We write $\mathbf{E}_q(\cdot)$ or $\mathbf{P}_q(\cdot)$ for the corresponding expectation and probability. Hence there exists a constant $c_q > 0$ such that

$$\mathbf{E}_q(\varphi(f)) = \sum_{f \in S_2(q)^*} \frac{c_q}{\langle f, f \rangle} \varphi(f)$$

for any $\varphi\colon S_2(q)^* \to \mathbf{C}$. From the Petersson formula, it is known that $c_q \to 1/(4\pi)$ as $q \to +\infty$ (see, e.g., Iwaniec and Kowalski [7, Ch. 14] or Cogdell and Michel [3]).

Let D be a relatively compact open set in $\mathbf{C}$ such that D is invariant under complex conjugation. Define $\mathcal{H}(D)$ to be the Banach space of functions holomorphic on D and continuous and bounded on $\bar{D}$, with the norm

$$\|\varphi\|_\infty = \sup_{s \in \bar{D}} |\varphi(s)|.$$

This is a separable complex Banach space. Define also $\mathcal{H}_{\mathbf{R}}(D)$ to be the set of $\varphi \in \mathcal{H}(D)$ such that $f(\bar{s}) = \overline{f(s)}$ for all $s \in D$ (this is well-defined since C is assumed to be invariant under conjugation). Note that the L-function of f, restricted to D, is an element of $\mathcal{H}_{\mathbf{R}}(D)$ since the Hecke eigenvalues $\lambda_f(n)$ are real for all $n \geqslant 1$.

We define L_q to be the random variable $S_2(q)^* \to \mathcal{H}(D)$ mapping $f \in S_2(q)^*$ to the restriction of $L(f, s)$ to D. (This depends on D, but the choice of D will always be clear in the context.)

If $\bar{D}$ is a compact subset of the strip $\frac{1}{2} < \mathrm{Re}(s) < 1$, then we will show that L_q converges in law to a random Dirichlet series. To define the limit, let $(X_p)_p$ be a sequence of independent random variables indexed by primes, taking values in the matrix group $\mathrm{SU}_2(\mathbf{C})$ and distributed according to the probability Haar measure on $\mathrm{SU}_2(\mathbf{C})$.

Theorem 2.1 (Bagchi's Theorem for modular forms)**.** *Assume that $\bar{D}$ is a compact subset of the strip $\frac{1}{2} < \mathrm{Re}(s) < 1$. Then, as $q \to +\infty$, the random variables L_q converge in law to the random Euler product*

$$L_D(s) = \prod_p \det(1 - X_p p^{-s})^{-1} = \prod_p (1 - \mathrm{Tr}(X_p) p^{-s} + p^{-2s})^{-1}$$

which is almost surely convergent in $\mathcal{H}(D)$, and belongs almost surely to $\mathcal{H}_{\mathbf{R}}(D)$.

Step 2. (Support of the random Euler product)

To deduce Theorem 1.1 from Theorem 2.1, we need the following computation of the support of the limiting measure.

Theorem 2.2. *Suppose that D is a disc with positive radius and diameter a segment of the real axis, always with $\bar{D}$ contained in $\frac{1}{2} < \mathrm{Re}(s) < 1$. The support of the law of the random Euler product L_D contains the set of functions $\varphi \in \mathcal{H}(D)$ such that $\varphi(x) > 0$ for $x \in D \cap \mathbf{R}$.*

Note that since $D \cap \mathbf{R}$ is an interval of positive length in $\mathbf{R}$, the condition $\varphi(x) > 0$ for all $x \in D \cap \mathbf{R}$ implies by analytic continuation that $\varphi \in \mathcal{H}_{\mathbf{R}}(D)$, which by Bagchi's Theorem 2.1 is a necessary condition to be in the support of L_D.

Step 3. (Conclusion) The elementary Lemma 2.4 below, combined with Theorems 2.1 and 2.2, implies Theorem 1.1 in the form

$$\liminf_{q\to+\infty} \mathbf{P}_q(\|L(f,\cdot)-\varphi\|_\infty < \varepsilon) = \liminf_{q\to+\infty} \sum_{\substack{f\in S_2(q)^* \\ \|L(f,\cdot)-\varphi\|_\infty<\varepsilon}} \frac{c_q}{\langle f,f\rangle} > 0 \tag{2.1}$$

for any function φ as in Theorem 1.1 and any $\varepsilon > 0$. We can easily deduce the "natural density" version from this: let A be the set of those $f \in S_2(q)^*$ such that $\|L(f,\cdot)-\varphi\|_\infty < \varepsilon$; then for any parameter $\eta > 0$, the definition of the harmonic measure on $S_2(q)^*$ gives

$$\frac{1}{|S_2(q)^*|}\sum_{f\in A} 1 = \mathbf{E}_q\Big(\mathbf{1}_A(f)\frac{\langle f,f\rangle}{c_q|S_2(q)^*|}\Big) \geqslant \eta\Big(\mathbf{P}_q(A) - \mathbf{P}_q\Big(\frac{\langle f,f\rangle}{c_q|S_2(q)|^*} < \eta\Big)\Big).$$

There exists $\delta > 0$ such that the first term is $\geqslant \delta > 0$ for all q large enough by (2.1); on the other hand, a result of Cogdell and Michel [3, Cor. 1.16] and the classical relation between the Petersson norm and the symmetric square L-function at $s = 1$ (see, e.g., [7, (5.101)]) imply that we can find $\eta > 0$ such that

$$\lim_{q\to+\infty} \mathbf{P}_q\Big(\frac{\langle f,f\rangle}{c_q|S_2(q)|^*} < \eta\Big) < \frac{\delta}{2}.$$

For this value of η, we obtain

$$\liminf_{q\to+\infty} \frac{1}{|S_2(q)^*|}\sum_{f\in A} 1 \geqslant \frac{\eta\delta}{2} > 0.$$

More precisely, the result of Cogdell-Michel is that for any $\eta > 0$, we have

$$\lim_{q\to+\infty} \mathbf{P}_q(L(\mathrm{Sym}^2 f, 1) \leqslant \eta) = F(\log\eta)$$

where F is the limiting distribution function for the special value at 1 of the symmetric square L-function of $f \in S_2(q)^*$. Since $F(x) \to 0$ when $x \to -\infty$, we obtain the result.

Remark 2.3. It would also be possible to argue throughout with the uniform probability measure on $S_2(q)^*$; the only change would be a slightly different form of Theorem 2.1, where the random variables (X_p) would not be identically distributed (compare with the equidistribution theorems of Serre [16] and Conrey–Duke–Farmer [4]).

Lemma 2.4. *Let M be a separable complete metric space and (X_n) a sequence of random variables with values in M that converge in law to X. Let S be the support of the law of X. Then for any $x \in S$, and any open neighborhood U of x, we have*

$$\liminf_{n\to+\infty} \mathbf{P}(X_n \in U) > 0.$$

Proof. By classical criteria for convergence in law, we have

$$\liminf_{n\to+\infty} \mathbf{P}(X_n \in U) \geqslant \mathbf{P}(X \in U) \tag{2.2}$$

for any open set $U \subset X$ (see, e.g., [2, Th. 2.1 (iv)]). Since $x \in S$, we have $\mathbf{P}(X \in U) > 0$, hence the result. ☐

3. Proof of Theorem 2.1

We begin with some preliminaries concerning the random Euler product L_D. In fact, if will be convenient to view it as a holomorphic function on larger domains then D, in a way that will be clear below. For this purpose, we fix a real number σ_0 such that $\frac{1}{2} < \sigma_0 < 1$, and such that the compact set $\bar{D}$ is contained in the half-plane S defined by $\mathrm{Re}(s) > \sigma_0$.

We recall that for $\nu \geqslant 0$, the d-th Chebychev polynomial is defined by

$$U_\nu(2\cos(x)) = \frac{\sin((\nu+1)x)}{\sin(x)}.$$

The importance of these polynomials for us lies in their relation with the representation theory of $\mathrm{SU}_2(\mathbf{C})$, namely

$$U_\nu(2\cos(x)) = \mathrm{Tr}\Big(\mathrm{Sym}^\nu \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix}\Big)$$

for any $x \in \mathbf{R}$, where Sym^d is the d-th symmetric power of the standard 2-dimensional representation of $\mathrm{SU}_2(\mathbf{C})$.

We define a sequence of random variables $(Y_n)_{n\geqslant 1}$ by

$$Y_n = \prod_{p^\nu || n} U_\nu(\mathrm{Tr}(X_p)).$$

In particular, we have $Y_n Y_m = Y_{nm}$ if n and m are coprime, and $Y_p = \mathrm{Tr}(X_p)$ if p is prime. The sequence (Y_p) is independent and Sato-Tate distributed. Moreover, since $|U_\nu(t)| \leqslant \nu+1$ for all $\nu \geqslant 0$ and all $t \in \mathbf{R}$, we have

$$|Y_n| \leqslant \prod_{p^\nu || n} (\nu+1) = d(n) \ll n^\varepsilon$$

for $n \geqslant 1$ and $\varepsilon > 0$, where the implied constant depends only on ε.

Lemma 3.1. (1) *Almost surely, the random Euler product*

$$\prod_p \det(1 - X_p p^{-s})^{-1}$$

converges and defines a holomorphic function on S. In particular, L_D converges almost surely to define an $\mathcal{H}(D)$-valued random variable.

(2) *Almost surely, we have*

$$\prod_p \det(1 - X_p p^{-s})^{-1} = \sum_{n\geqslant 1} Y_n n^{-s}$$

for all $s \in S$, and in particular L_D coincides with the random Dirichlet series on the right.

(3) *For $\sigma > 1/2$ and $u \geqslant 2$, we have*

$$\mathbf{E}\Big(\Big|\sum_{n\leqslant u} Y_n n^{-\sigma}\Big|^2\Big) \ll 1,$$

where the implied constant depends only on σ.

Proof. (1) Let σ be a fixed real number such that $\frac{1}{2} < \sigma < \sigma_0$. By expanding, we can write

$$-\log\det(1 - X_p p^{-s}) = Y_p p^{-s} + g_p(s)$$

where the random series

$$\sum_p g_p(s)$$

converges absolutely (and surely) for $\mathrm{Re}(s) > 1/2$. Since $\mathbf{E}(Y_p p^{-\sigma}) = 0$ and $\mathbf{E}(Y_p^2 p^{-2\sigma}) = p^{-2\sigma}$, Kolmogorov's Three Series Theorem (see, e.g., [14, Th. 0.III.2]) implies that the random series

$$\sum_p Y_p p^{-\sigma}$$

converges almost surely. By well-known results on Dirichlet series, this means that the random series

$$\sum_p Y_p p^{-s}$$

converges almost surely to a holomorphic function on the half-plane S. This implies the first statement by taking the exponential. The second follows by restricting to D since $\bar{D}$ is contained in S.

(2) We first show that almost surely the random Dirichlet series

$$\tilde{L}(s) = \sum_{n\geqslant 1} Y_n n^{-s}$$

converges and defines a function holomorphic on S. The key point is that the variables Y_n for n squarefree form an orthonormal system: we have

$$\mathbf{E}(Y_n Y_m) = \delta(n, m)$$

if n and m are squarefree numbers. Indeed, if $n \neq m$, there is a prime p dividing only one of n and m, say $p \mid n$, and then by independence we get $\mathbf{E}(Y_n Y_m) = \mathbf{E}(Y_p)\mathbf{E}(Y_{n/p} Y_m) = 0$; and if $n = m$ is squarefree then we have

$$\mathbf{E}(Y_n^2) = \prod_{p\mid n} \mathbf{E}(Y_p^2) = 1.$$

Fix again σ such that $\frac{1}{2} < \sigma < \sigma_0$. By the Rademacher–Menchov Theorem (see, e.g. [9, Th. B.8.4]), the random series

$$\sum_n{}^{\flat}\, Y_n n^{-\sigma}$$

over squarefree numbers converges almost surely. By elementary factorization and properties of products of Dirichlet series (the product of an absolutely convergent Dirichlet series and a convergent one is convergent, see e.g. [6, Th. 54]) the same holds for

$$\sum_{n\geqslant 1} Y_n n^{-\sigma}.$$

As in (1), this gives the almost sure convergence of the series defining $\tilde{L}(s)$ to a holomorphic function in S. Restricting gives the $\mathcal{H}(D)$-valued random variable $\tilde{L}_D$.

Finally, almost surely both the random Euler product and the random Dirichlet series converge and are holomorphic for $\mathrm{Re}(s) > \sigma_0$. For $\mathrm{Re}(s) > 3/2$, they converge absolutely, and coincide by a well-known formal Euler product computation: for any prime p and any $x \in \mathbf{R}$, denoting

$$t(x) = \begin{pmatrix} e^{ix} & 0 \\ 0 & e^{-ix} \end{pmatrix}$$

we have

$$\begin{aligned} \det(1 - t(x)p^{-s})^{-1} &= (1 - e^{ix}p^{-s})^{-1}(1 - e^{-ix}p^{-s})^{-1} \\ &= \sum_{\nu\geqslant 0} \mathrm{Tr}\,\mathrm{Sym}^{\nu}(t(x))p^{-\nu s} = \sum_{\nu\geqslant 0} U_{\nu}(x)p^{-\nu s} \end{aligned}$$

(compare the discussion of Cogdell and Michel in [3, §2]). By analytic continuation, we deduce that $L_D = \tilde{L}_D$ almost surely as $\mathcal{H}(D)$-valued random variables.

(3) Since the random varibles Y_n are real-valued, we have

$$\mathbf{E}\Big(\Big|\sum_{n\leqslant u} Y_n n^{-\sigma}\Big|^2\Big) = \sum_{n,m\leqslant u} \frac{1}{(nm)^{\sigma}} \mathbf{E}(Y_n Y_m).$$

For given n and m, let $d = (n, m)$ and $n' = n/d$, $m' = m/d$. Then by multiplicativity and independence of the variables $(Y_p)_p$, we have

$$\mathbf{E}(Y_n Y_m) = \mathbf{E}(Y_d^2)\mathbf{E}(Y_{n'})\mathbf{E}(Y_{m'}).$$

By the definition of Y_p, we have $\mathbf{E}(Y_{n'}) = 0$ if n' is divisible by a prime p with odd exponent, and similarly for $\mathbf{E}(Y_{m'})$. Hence we have $\mathbf{E}(Y_n Y_m) = 0$ unless both n' and m' are squares. Therefore

$$\mathbf{E}\Big(\Big|\sum_{n\leqslant u} Y_n n^{-\sigma}\Big|^2\Big) \leqslant \sum_{d\geqslant 1} \frac{\mathbf{E}(Y_d^2)}{d^{2\sigma}} \sum_{m,n\geqslant 1} \frac{1}{(mn)^{2\sigma}} \mathbf{E}(Y_m Y_n) < +\infty$$

since $\sigma > 1/2$ and $\mathbf{E}(Y_n) \ll n^{\varepsilon}$ for any $\varepsilon > 0$. □

The key arithmetic properties of the family $S_2(q)^*$ of modular forms that are required in the proof of Theorem 2.1 are the following:

Proposition 3.2 (Local spectral equidistribution)**.** *As $q \to +\infty$, the sequence $(\lambda_f(p))_p$ of Fourier coefficients of $f \in S_2(q)^*$ converges in law to the sequence $(Y_p)_p$.*

Proof. This is a well-known consequence of the Petersson formula, see e.g. [10, Prop. 8], [12, Appendix] or [3, Prop. 1.9]; here restricting to prime level q and weight 2 also simplifies matters since this ensures that the old space of $S_2(q)$ is zero. □

Proposition 3.3 (First moment estimate)**.** *There exists an absolute constant $A \geqslant 1$ such that for any real number $\delta > 0$ with $\delta < 1/2$, and for any $s \in \mathbf{C}$ such that $\frac{1}{2} + \delta \leqslant \mathrm{Re}(s)$, we have*

$$\mathbf{E}_q(|L(f,s)|) \ll (1+|s|)^A,$$

where the implied constant depends only on δ.

Proof. This follows easily, using the Cauchy-Schwarz inequality, from the second moment estimate [11, Prop. 5] of Kowalski and Michel (with $\Delta = 0$); although this statement is not formally the same, it is in fact a more difficult average (it operates closer to the critical line). ☐

We now prove some additional lemmas.

Lemma 3.4 (Polynomial growth)**.** *For any real number $\sigma > \sigma_0$, we have*

$$\mathbf{E}\Big(\Big|\sum_{n\geqslant 1} Y_n n^{-s}\Big|\Big) \ll 1+|s|$$

uniformly for all s such that $\mathrm{Re}(s) \geqslant \sigma > \sigma_0$.

Proof. We write

$$L(s) = \sum_{n\geqslant 1} Y_n n^{-s}.$$

This is almost surely a function holomorphic on the half-plane S. The series

$$\sum_{n\geqslant 1} \frac{Y_n}{n^{\sigma_0}}$$

converges almost surely. Therefore the partial sums

$$S_u = \sum_{n\leqslant u} \frac{Y_n}{n^{\sigma_0}}$$

are bounded almost surely. By summation by parts, it follows from the convergence of the series $L(s)$ that for any s with real part $\mathrm{Re}(s) \geqslant \sigma > \sigma_0$, we have

$$L(s) = (s-\sigma_0)\int_1^{+\infty} \frac{S_u}{u^{s-\sigma_0+1}}du,$$

where the integral converges almost surely. Hence almost surely

$$|L(s)| \leqslant (1+|s|)\int_1^{+\infty} \frac{|S_u|}{u^{\sigma-\sigma_1+1}}du.$$

Fubini's Theorem and the Cauchy-Schwarz inequality then imply

$$\begin{aligned}\mathbf{E}(|L(s)|) &\leqslant (1+|s|)\int_1^{+\infty} \mathbf{E}(|S_u|)\frac{du}{u^{\sigma-\sigma_0+1}} \\ &\leqslant (1+|s|)\int_1^{+\infty} \mathbf{E}(|S_u|^2)^{1/2}\frac{du}{u^{\sigma-\sigma_0+1}} \ll 1+|s|\end{aligned}$$

by Lemma 3.1 (3). ☐

We now consider some elementary approximation statements of the L-functions and of the random Dirichlet series by smoothed partial sums. For this, we fix once and for all a smooth function $\varphi\colon [0,+\infty[\to [0,1]$ with compact support such that $\varphi(0)=1$, and we denote $\hat{\varphi}$ its Mellin transform.

We also fix $T\geqslant 1$ and a compact interval I in $]1/2,1[$ such that the compact rectangle $R=I\times[-T,T]\subset \mathbf{C}$ is contained in S and contains D in its interior. We then finally define $\delta>0$ so that

$$\min\{\mathrm{Re}(s)\mid s\in R\}=\frac{1}{2}+2\delta.$$

Lemma 3.5. *For $N\geqslant 1$, define the $\mathcal{H}(D)$-valued random variable*

$$L_D^{(N)}=\sum_{n\geqslant 1}Y_n\varphi\Bigl(\frac{n}{N}\Bigr)n^{-s}.$$

We then have

$$\mathbf{E}(\|L_D-L_D^{(N)}\|_\infty)\ll N^{-\delta}$$

for $N\geqslant 1$, where the implied constant depends on D.

Proof. We again write

$$L(s)=\sum_{n\geqslant 1}Y_n n^{-s}$$

when we wish to view the Dirichlet series as defined and holomorphic (almost surely) on S.

For any s in the rectangle R, we have almost surely the representation

$$L(s)-L^{(N)}(s)=-\frac{1}{2i\pi}\int_{(-\delta)}L(s+w)\hat{\varphi}(w)N^w dw \tag{3.1}$$

by standard contour integration.[2]

We also have almost surely for any v in D the Cauchy formula

$$L_D(v)-L_D^{(N)}(v)=\frac{1}{2i\pi}\int_{\partial R}(L(s)-L^{(N)}(s))\frac{ds}{s-v},$$

where the boundary of R is oriented counterclockwise. The definition of the rectangle R ensures that $|s-v|^{-1}\gg 1$ for $v\in D$ and $s\in\partial R$, and therefore

$$\|L_D-L_D^{(N)}\|_\infty\ll\int_{\partial R}|L(s)-L^{(N)}(s)|\,|ds|.$$

Using (3.1) and writing $w=-\delta+iu$ with $u\in\mathbf{R}$, we obtain

$$\|L_D-L_D^{(N)}\|_\infty\ll N^{-\delta}\int_{\partial R}\int_{\mathbf{R}}|L(-\delta+iu+s)|\,|\hat{\varphi}(-\delta+iu)||ds|du.$$

Therefore, taking the expectation, and using Fubini's Theorem, we get

$$\begin{aligned}\mathbf{E}(\|L_D-L_D^{(N)}\|_\infty)&\ll N^{-\delta}\int_{\partial R}\int_{\mathbf{R}}\mathbf{E}\bigl(|L(-\delta+iu+s)|\bigr)\,|\hat{\varphi}(-\delta+iu)||ds|du\\&\ll N^{-\delta}\sup_{s=\sigma+it\in R}\int_{\mathbf{R}}\mathbf{E}\bigl(|L(-\delta+iu+\sigma+it)|\bigr)\,|\hat{\varphi}(-\delta+iu)|du.\end{aligned}$$

[2] Here and below, it is important that the "almost surely" property holds for *all* s, which is the case because we work with random holomorphic functions, and not with particular evaluations of these random functions at specific points s.

We therefore need to bound

$$\int_{\mathbf{R}} \mathbf{E}\big(|L(-\delta+iu+\sigma+it)|\big)\ |\hat{\varphi}(-\delta+iu)|du.$$

for some fixed $\sigma+it$ in the compact rectangle R. The real part of the argument $-\delta+iu+\sigma+it$ is $\sigma-\delta \geqslant \frac{1}{2}+\delta$ by definition of δ, and hence

$$\mathbf{E}(|L(-\delta+iu+\sigma+it)|) \ll 1+|-\delta+iu+\sigma+it| \ll 1+|u|$$

uniformly for $\sigma+it \in R$ and $u \in \mathbf{R}$ by Lemma 3.4. Since $\hat{\varphi}$ decays faster than any polynomial at infinity, we conclude that

$$\int_{\mathbf{R}} \mathbf{E}\big(|L(-\delta+iu+\sigma+it)|\big)\ |\hat{\varphi}(-\delta+iu)|du \ll 1$$

uniformly for $s=\sigma+it \in R$, and the result follows. □

We proceed similarly for the L-functions.

Lemma 3.6. *For $N \geqslant 1$ and $f \in S_2(q)^*$, define*

$$L^{(N)}(f,s) = \sum_{n\geqslant 1} \lambda_f(n)\varphi\Big(\frac{n}{N}\Big)n^{-s},$$

and define $\mathsf{L}_q^{(N)}$ to be the $\mathcal{H}(D)$-valued random variable mapping f to $L^{(N)}(f,s)$ restricted to D. We then have

$$\mathbf{E}_q(\|\mathsf{L}_q-\mathsf{L}_q^{(N)}\|_\infty) \ll N^{-\delta}$$

for $N \geqslant 1$ and all q.

Proof. For any $s \in R$, we have the representation

$$L(f,s)-L^{(N)}(f,s) = -\frac{1}{2i\pi}\int_{(-\delta)} L(f,s+w)\hat{\varphi}(w)N^w dw. \tag{3.2}$$

and for any v with $\mathrm{Re}(v) > 1/2$, Cauchy's theorem gives

$$L(f,v)-L^{(N)}(f,v) = \frac{1}{2i\pi}\int_{\partial R}(L(f,s)-L^{(N)}(f,s))\frac{ds}{s-v},$$

where the boundary of R is oriented counterclockwise. As in the previous argument, we deduce that

$$\|\mathsf{L}_q-\mathsf{L}_q^{(N)}\|_\infty \ll \int_{\partial R}|L(f,s)-L^{(N)}(f,s)||ds|$$

for $f \in S_2(q)^*$. Taking the expectation with respect to f and changing the order of summation and integration leads to

$$\begin{aligned}\mathbf{E}_q\Big(\|\mathsf{L}_q-\mathsf{L}_q^{(N)}\|_\infty\Big) &\ll \int_{\partial R}\mathbf{E}_q\big(|L(f,s)-L^{(N)}(f,s)|\big)|ds| \\ &\ll \sup_{s\in\partial R}\mathbf{E}_q\big(|L(f,s)-L^{(N)}(f,s)|\big). \end{aligned}\tag{3.3}$$

Applying (3.2) and using again Fubini's Theorem, we obtain

$$\mathbf{E}_q\big(|L(f,s)-L^{(N)}(f,s)|\big) \ll N^{-\delta}\int_{\mathbf{R}}|\hat{\varphi}(-\delta+iu)|\mathbf{E}_q\big(|L(f,-\delta+iu+\sigma+it)|\big)du$$

for $s \in \partial R$. Since $\sigma - \delta \geqslant \frac{1}{2} + \delta$, we get

$$\mathbf{E}_q\big(|L(f, -\delta + iu + \sigma + it)|\big) \ll (1 + |u|)^A \tag{3.4}$$

by Proposition 3.3, where A is an absolute constant. Hence

$$\mathbf{E}_q\big(|L(f,s) - L^{(N)}(f,s)|\big) \ll N^{-\delta} \int_{\mathbf{R}} |\hat{\varphi}(-\delta + iu)|(1 + |u|)^A du \ll N^{-\delta}. \tag{3.5}$$

We conclude from (3.3) that

$$\mathbf{E}_q\Big(\|\mathsf{L}_q - \mathsf{L}_q^{(N)}\|_\infty\Big) \ll N^{-\delta},$$

as claimed. □

Proof of Theorem 2.1. A simple consequence of the definition of convergence in law shows that it is enough to prove that for any bounded and Lipschitz function $f : \mathcal{H}(D) \longrightarrow \mathbf{C}$, we have

$$\mathbf{E}_q(f(\mathsf{L}_q)) \longrightarrow \mathbf{E}(f(L_D))$$

as $q \to +\infty$ (see [2, p. 16, (ii)⇒ (iii) and (1.1), p. 8]). To prove this, we use the Dirichlet series expansion of L_D given by Lemma 3.1 (2).

Let $N \geqslant 1$ be some integer to be chosen later. Let

$$\mathsf{L}_q^{(N)} = \sum_{n \geqslant 1} \lambda_f(n) n^{-s} \varphi\Big(\frac{n}{N}\Big)$$

(viewed as random variable defined on $S_2(q)^*$) and

$$L_N = \sum_{n \geqslant 1} Y_n n^{-s} \varphi\Big(\frac{n}{N}\Big)$$

be the smoothed partial sums of the Dirichlet series, as in Lemmas 3.6 and 3.5.

We then write

$$|\mathbf{E}_q(f(\mathsf{L}_q)) - \mathbf{E}(f(L))| \leqslant |\mathbf{E}_q(f(\mathsf{L}_q) - f(\mathsf{L}_q^{(N)}))| + \\ |\mathbf{E}_q(f(\mathsf{L}_q^{(N)})) - \mathbf{E}(f(L^{(N)}))| + |\mathbf{E}(f(L^{(N)}) - f(L))|.$$

Since f is a Lipschitz function on $\mathcal{H}(D)$, there exists a constant $C \geqslant 0$ such that

$$|f(x) - f(y)| \leqslant C\|x - y\|_\infty$$

for all x, $y \in \mathcal{H}(D)$. Hence we have

$$|\mathbf{E}_q(f(\mathsf{L}_q)) - \mathbf{E}(f(L))| \leqslant C\mathbf{E}_q(\|\mathsf{L}_q - \mathsf{L}_q^{(N)}\|_\infty) + \\ |\mathbf{E}_q(f(\mathsf{L}_q^{(N)})) - \mathbf{E}(f(L^{(N)}))| + C\mathbf{E}(\|L^{(N)} - L\|_\infty).$$

Fix $\varepsilon > 0$. Lemmas 3.6 and 3.5 together show that there exists some $N \geqslant 1$ such that

$$\mathbf{E}_q(\|\mathsf{L}_q - \mathsf{L}_q^{(N)}\|_\infty) < \varepsilon$$

for all $q \geqslant 2$ and

$$\mathbf{E}(\|L^{(N)} - L\|_\infty) < \varepsilon.$$

We *fix* such a value of N. By Proposition 3.2 (and composition with a continuous function), the random variables $\mathsf{L}_q^{(N)}$ (which are Dirichlet polynomials) converge in law to $L^{(N)}$ as $q \to +\infty$. We deduce that we have

$$|\mathbf{E}_q(f(\mathsf{L}_q)) - \mathbf{E}(f(L))| < 4\varepsilon$$

for all q large enough. This finishes the proof. □

4. Proof of Theorem 2.2

For the computation of the support of the random Dirichlet series $L(s)$, we apply a trick to exploit the analogous result known for the case of the Riemann zeta function. We denote $\widehat{\mathrm{SU}}_2$ the product of copies of the unit circle indexed by primes, so an element (x_p) of $\widehat{\mathrm{SU}}_2$ is a family of matrices in $\mathrm{SU}_2(\mathbf{C})$ indexed by p.

The assumptions on D in Theorem 2.2[3] imply that there exists τ such that $1/2 < \tau < 1$ and $r > 0$ such that

$$D = \{s \in \mathbf{C} \mid |s - \tau| \leqslant r\} \subset \{s \in \mathbf{C} \mid 1/2 < \mathrm{Re}(s) < 1\}.$$

Lemma 4.1. *Let N be an arbitrary positive real number. The set of all series*

$$\sum_{p>N} \frac{\mathrm{Tr}(x_p)}{p^s}, \qquad (x_p) \in \widehat{\mathrm{SU}}_2$$

which converge in $\mathcal{H}(D)$ is dense in the subspace $\mathcal{H}_{\mathbf{R}}(D)$.

In the proof and the next, we allow ourselves the luxury of writing sometimes $\|\varphi(s)\|_\infty$ instead of $\|\varphi\|_\infty$.

Proof. Bagchi [1, Lemma 5.2.10] proves (using results of complex analysis due to Bernstein, Polyá and others) that the set of series

$$\sum_{p>N} \frac{e^{i\theta_p}}{p^s}, \qquad \theta_p \in \mathbf{R}$$

that converge in $\mathcal{H}(D)$ is dense in $\mathcal{H}(D)$ (precisely, he proves this for $N = 1$, but the same proof applies to any value of N). If $\varphi \in \mathcal{H}_{\mathbf{R}}(D)$ and $\varepsilon > 0$, we can therefore find real numbers (θ_p) such that

$$\Big\|\frac{\varphi(s)}{2} - \sum_{p>N} \frac{e^{i\theta_p}}{p^s}\Big\|_\infty < \frac{\varepsilon}{2}.$$

It follows then that

$$\Big\|\frac{\overline{\varphi(\bar{s})}}{2} - \sum_{p>N} \frac{e^{-i\theta_p}}{p^s},\Big\|_\infty < \frac{\varepsilon}{2},$$

hence (since $\varphi \in \mathcal{H}_{\mathbf{R}}(D)$) that

$$\Big\|\varphi(s) - \sum_{p>N} \frac{e^{i\theta_p} + e^{-i\theta_p}}{p^s}\Big\|_\infty < \varepsilon,$$

[3] These assumptions could be easily weakened, as has been done for Voronin's Theorem.

which gives the result since

$$e^{i\theta_p}+e^{-i\theta_p}=\mathrm{Tr}\begin{pmatrix} e^{i\theta_p} & 0\\ 0 & e^{-i\theta_p}\end{pmatrix}$$

is the trace of a matrix in $\mathrm{SU}_2(\mathbf{C})$. □

We will use this to prove:

Proposition 4.2. *The support of the law of*

$$\log L_D(s)=-\sum_p \log\det(1-X_p p^{-s})$$

in $\mathcal{H}(D)$ *is* $\mathcal{H}_{\mathbf{R}}(D)$.

Proof. Since $X_p\in \mathrm{SU}_2(\mathbf{C})$, the function $\log L(s)$ is almost surely in the space $\mathcal{H}_{\mathbf{R}}(D)$. Since the summands are independent, a well-known result concerning the support of random series (see, e.g., [9, Prop. B.8.7]) shows that it suffices to prove that the set of convergent series

$$-\sum_p \log\det(1-x_p p^{-s}),\qquad (x_p)\in\widehat{\mathrm{SU}}_2,$$

is dense in $\mathcal{H}_{\mathbf{R}}(D)$. Denote $L(s;(x_p))$ this series, when it converges in $\mathcal{H}(D)$.

We can write

$$-\sum_p \log\det(1-x_p p^{-s})=\sum_p\frac{\mathrm{Tr}(x_p)}{p^s}+g(s;(x_p))$$

where $s\mapsto g(s;(x_p))$ is holomorphic in the region $\mathrm{Re}(s)>1/2$. Indeed

$$g(s;(x_p))=\sum_p\log\Bigl(1+\sum_{k\geqslant 0}\mathrm{Tr}(x_p)^k p^{-(k+2)s}\Bigr).$$

Fix $\varphi\in\mathcal{H}_{\mathbf{R}}(D)$ and let $\varepsilon>0$ be fixed. There exists $N\geqslant 1$ such that

$$\Bigl\|s\mapsto\sum_{p>N}\log\Bigl(1+\sum_{k\geqslant 0}\mathrm{Tr}(x_p)^k p^{-(k+2)s}\Bigr)\Bigr\|_\infty<\varepsilon \tag{4.1}$$

for *any* $(x_p)\in\widehat{\mathrm{SU}}_2$. Now take $x_p=1\in\mathrm{SU}_2(\mathbf{C})$ for $p\leqslant N$ and define

$$\varphi_1=\varphi+\sum_{p\leqslant N}\frac{\mathrm{Tr}(x_p)}{p^s}=\varphi+2\sum_{p\leqslant N}\frac{1}{p^s},$$

which belongs to $\mathcal{H}_{\mathbf{R}}(D)$. By Lemma 4.1, there exist x_p for $p>N$ in $\mathrm{SU}_2(\mathbf{C})$ such that

$$\Bigl\|\sum_{p>N}\frac{\mathrm{Tr}(x_p)}{p^s}-\varphi_1(s)\Bigr\|_\infty<\varepsilon.$$

The left-hand side is the norm of

$$\log L(s;(x_p))-g(s;(x_p))-\sum_{p\leqslant N}\frac{\mathrm{Tr}(x_p)}{p^s}-\varphi_1(s)=\log L(s;(x_p))-\varphi(s)-g(x;(x_p)),$$

and by (4.1), we obtain

$$\|\log L(s;(x_p))-\varphi(s)\|_\infty<2\varepsilon.$$

This implies the lemma. □

Using composition with the exponential function and a lemma of Hurwitz (see, e.g., [17, 3.45]) on zeros of limits of holomorphic functions, we see that the support of the limiting Dirichlet series L_D in $\mathcal{H}(D)$ is the union of the zero function and the set of functions $\varphi \in \mathcal{H}_{\mathbf{R}}(D)$ such that $\varphi(\sigma) > 0$ for $\sigma \in D \cap \mathbf{R}$. In particular, this proves Theorem 2.2.

5. Generalizations

It is clear from the proof that Bagchi's Theorem should hold in considerable generality for any family of L-functions. Indeed, the crucial ingredients are the local spectral equidistribution (Proposition 3.2), and the first moment estimate (Proposition 3.3).

The first result is a qualitative statement that is understood to be at the core of any definition of "family" of L-functions (this is explained in [8], but also appears, with a different terminology, for the families of Conrey–Farmer–Keating–Rubinstein–Snaith [5] and Sarnak–Shin–Templier [15]); it is now known in many circumstances (indeed, often in quantitative form).

The moment estimate is typically derived from a second-moment bound, and is also definitely expected to hold for a reasonable family of L-functions, but it has only been proved in much more restricted circumstances than local spectral equidistribution. However, it is very often the case that one can at least prove (using local spectral equidistribution) a weaker statement: for some σ_1 such that $1/2 < \sigma_1 < 1$, the second moment of the L-functions satisfies the analogue of Proposition 3.3; an analogue of Bagchi's Theorem then follows at least for compact discs in the region $\sigma_1 < \mathrm{Re}(s) < 1$.

As far as universality (i.e., Theorem 2.2) is concerned, one may expect that (using tricks similar to the proof of Theorem 2.2) only two different cases really occur, depending on whether the coefficients of the L-functions are real (as in our case) or complex (as in the case of vertical translates of a fixed L-function).

References

[1] B. Bagchi: *Statistical behaviour and universality properties of the Riemann zeta function and other allied Dirichlet series*, PhD thesis, Indian Statistical Institute, Kolkata, 1981; available at library.isical.ac.in/jspui/handle/10263/4256

[2] P. Billingsley: *Convergence of probability measures*, Wiley (1968).

[3] J. Cogdell and P. Michel: *On the complex moments of symmetric power L-functions at $s = 1$*, Internat. Math. Res. Notices (2004), 1561–1617.

[4] B. Conrey, W. Duke and D. Farmer: *The distribution of the eigenvalues of Hecke operators*, Acta Arith. 78 (1997), 405–409.

[5] J.B. Conrey, D. Farmer, J. Keating, M. Rubinstein and N. Snaith: *Integral moments of L-functions*, Proc. Lond. Math. Soc. 91 (2005) 33–104.

[6] G.H. Hardy and M. Riesz: *The general theory of Dirichlet's series*, Cambridge Tracts in Math. 18, C.U.P. 1915.

[7] H. Iwaniec and E. Kowalski: *Analytic number theory*, Colloquium Publ. 53, American Math. Soc. 2004.

[8] E. Kowalski: *Families of cusp forms*, Pub. Math. Besançon, 2013, 5–40.

[9] E. Kowalski: *Arithmetic randonnée: an introduction to probabilistic number theory*, ETH Zürich Lecture Notes, www.math.ethz.ch/~kowalski/probabilistic-number-theory.pdf.

[10] E. Kowalski, Y-K. Lau, K. Soundararajan and J. Wu: *On modular signs*, Math. Proc. Cambridge Phil. Soc. 149 (2010), 389–411 doi:10.1017/S030500411000040X

[11] E. Kowalski and Ph. Michel: *The analytic rank of $J_0(q)$ and zeros fo automorphic L-functions*, Duke Math. J. 100 (1999), 503–542.

[12] E. Kowalski, A. Saha and J. Tsimerman: *Local spectral equidistribution for Siegel modular forms and applications*, Compositio Math. 148 (2012), 335–384.
[13] A. Laurinčikas and K. Matsumoto: *The universality of zeta-functions attached to certain cusp forms*, Acta Artih. 98 (2001), 345–359.
[14] D. Li and H. Queffélec: *Introduction à l'étude des espaces de Banach; Analyse et probabilités*, Cours Spécialisés 12, S.M.F, 2004.
[15] P. Sarnak, S-W. Shin and N. Templier: *Families of L-functions and their symmetry*, in "Families of automorphic forms and the trace formula", Simons Symposia, Springer 2016, 531–578.
[16] J-P. Serre: *Répartition asymptotique des valeurs propres de l'opérateur de Hecke T_p*, J. American Math. Soc. 10 (1997), 75–102.
[17] E.C. Titchmarsh: *The theory of functions*, 2nd edition, Oxford Univ. Press, 1939.
[18] S.M. Voronin: *Theorem on the 'universality' of the Riemann zeta function*, Izv. Akad. Nauk SSSR, 39 (1975); 475–486; translation in Math. USSR Izv. 9 (1975), 443–445.

ETH Zürich – D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland
E-mail address: kowalski@math.ethz.ch