ISSN 0012-9593

quatrième série - tome 52 fascicule 5 septembre-octobre 2019

ANNALES SCIENTIFIQUES de L'ÉCOLE NORMALE SUPÉRIEURE

P. SARNAK & P. ZHAO

The Quantum Variance of the Modular Surface

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

Annales Scientifiques de l'École Normale Supérieure

Publiées avec le concours du Centre National de la Recherche Scientifique

Rédaction / *Editor*

Annales Scientifiques de l'École Normale Supérieure,
45, rue d'Ulm, 75230 Paris Cedex 05, France.
Tél. : (33) 1 44 32 20 88. Fax : (33) 1 44 32 20 80.
annales@ens.fr

Édition et abonnements / *Publication and subscriptions*

Société Mathématique de France
Case 916 - Luminy
13288 Marseille Cedex 09
Tél. : (33) 04 91 26 74 64
Fax : (33) 04 91 41 17 51
email : abonnements@smf.emath.fr

Tarifs

Abonnement électronique : 420 euros.
Abonnement avec supplément papier :
Europe : 551 €. Hors Europe : 620 € ($ 930). Vente au numéro : 77 €.

ISSN 0012-9593 (print) 1873-2151 (electronic)

Directeur de la publication : Stéphane Seuret
Périodicité : 6 nos / an

Ann. Scient. Éc. Norm. Sup.
4^{e} série, t. 52, 2019, p. 1155 à 1200

THE QUANTUM VARIANCE OF THE MODULAR SURFACE

BY P. SARNAK AND P. ZHAO
WITH AN APPENDIX BY M. WOODBURY

ABSTRACT. – The variance of observables of quantum states of the Laplacian on the modular surface is calculated in the semiclassical limit. It is shown that this hermitian form is diagonalized by the irreducible representations of the modular quotient and on each of these it is equal to the classical variance of the geodesic flow after the insertion of a subtle arithmetical special value of the corresponding L-function.

RÉSUMÉ. – Nous calculons la variance des observables des états quantiques du Laplacien sur la surface modulaire dans la limite semiclassique. Nous montrons que cette forme hermitienne est diagonalisée par les représentations irréductibles du quotient modulaire et sur chacune de ces représentations, elle est égale à la variance classique du flot géodésique après insertion d'une subtile valeur spécifique de la fonction L correspondante.

1. Introduction

Let $G = PSL(2, \mathbb{R})$, $\Gamma = PSL(2, \mathbb{Z})$ and $X = \Gamma\backslash\mathbb{H}$ be the modular surface. X is a hyperbolic surface of finite area and it has a large discrete spectrum for the Laplacian (see [14] and [34]). The corresponding eigenfunctions can be diagonalized and we denote these Hecke-Maass forms by ϕ_j, $j = 1, 2, \ldots$. They are real valued and satisfy

$$\Delta\phi_j + \lambda_j\phi_j = 0, \quad T_n\phi_j = \lambda_j(n)\phi_j \tag{1}$$

and we normalize them by

$$\int_X \phi_j(z)^2 dA(z) = 1. \tag{2}$$

Here dA is the normalized hyperbolic area form and write $\lambda_j = \frac{1}{4} + t_j^2$. If $\lambda > 0$ then it is known that such a ϕ is a cusp form [14]. ϕ_j has a Fourier expansion,

$$\phi_j(z) = \sum_{n\neq 0} \frac{c_j(|n|)}{\sqrt{|n|}} W_{0,it_j}(4\pi|n|y)e(nx), \tag{3}$$

where W_{0,it_j} is the Whittaker function. X carries a further symmetry induced by the orientation reversing isometry $z \to -\bar{z}$ of $\mathbb{H}$ and our ϕ_j's are either even or odd with respect to this symmetry r

$$\phi_j(rz) = \epsilon_j \phi_j(z), \qquad \epsilon_j = \pm 1. \tag{4}$$

Correspondingly

$$c_j(n) = \epsilon_j c_j(-n). \tag{5}$$

The Iwasawa decomposition of $g \in G$ takes the form

$$g = n(x)a(y)k(\theta) \tag{6}$$

where

$$n(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, a(y) = \begin{pmatrix} y^{\frac{1}{2}} & 0 \\ 0 & y^{-\frac{1}{2}} \end{pmatrix}, k(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

$\mathbb{H}$ may be identified with G/K where $K = SO(2)/(\pm I)$ and then $\Gamma\backslash G$ is identified with the unit tangent space or phase space for the geodesic flow on X. The objects whose fluctuations we study in this paper are the Wigner distributions $d\omega_j$ on $\Gamma\backslash G$. These are quadratic functionals of the ϕ_j's and are given by (see the recent paper [1] for a detailed description of these distributions as well as their basic invariance properties),

$$d\omega_j = \phi_j(z) \sum_{k \in \mathbb{Z}} \phi_{j,k}(z) e^{-2ik\theta} d\omega \tag{7}$$

where

$$d\omega = \frac{dxdy}{y^2} \frac{d\theta}{2\pi}.$$

Here the $\phi_{j,k}$ are the shifted Maass cusp forms of weight k, normalized such that $\|\phi_{j,k}\|_2 = 1$ by raising and lowering operators, E_+ and E_- respectively, where [19]

$$E_+ = e^{-2i\theta}(2iy\frac{\partial}{\partial x} + 2y\frac{\partial}{\partial y} + i\frac{\partial}{\partial \theta}),$$

$$E_- = e^{2i\theta}(2iy\frac{\partial}{\partial x} - 2y\frac{\partial}{\partial y} + i\frac{\partial}{\partial \theta}).$$

They are eigenfunctions of the Casimir operator Ω, which acts on $C^\infty(\Gamma\backslash G)$.

The basic question concerning the ω_j's is their behavior in the semi-classical limit $t_j \to \infty$. Lindenstrauss [25] and Soundararajan [36] have shown that for an "observable" $\psi \in C(\Gamma\backslash G)$

$$\omega_j(\psi) \to \frac{1}{\mathrm{vol}(\Gamma\backslash G)} \int_{\Gamma\backslash G} \psi(g)dg, \quad \text{as} \quad j \to \infty \tag{8}$$

where dg is normalized Haar measure (i.e., a probability measure), this is the so called "QUE" property.

It is known after Watson [38] and Jakobson [19] that the generalized Lindelöf Hypothesis implies that if

$$\int_{\Gamma\backslash G} \psi(g)dg = 0 \tag{9}$$

then, for $\epsilon > 0$

$$\omega_j(\psi) \ll_\epsilon t_j^{-\frac{1}{2}+\epsilon}. \tag{10}$$

For the rest of the paper we will assume that the mean value of ψ is 0, i.e., (9) holds. The main result below is the determination of the quantum variance, namely the mean-square of the $\omega_j(\psi)$'s. These are computed for special observables (ones depending only on $z \in X$) in [30] where the ϕ_j's are replaced by holomorphic forms, and in [44] for the ω_j's at hand. The extension to the general observable that is carried out here is substantially more complicated and intricate. It comes with a reward in that the answer on the phase space is conceptually much more transparent and elegant.

The variance sums

$$S_\psi(T) := \sum_{t_j \le T} |\omega_j(\psi)|^2 \tag{11}$$

were introduced by Zelditch who showed (in much greater generality) that $S_\psi(T) = O(\frac{T^2}{\log T})$ [41]. Corresponding to (10) we expect that in our setting $S_\psi(T)$ will be at most $T^{1+\epsilon}$, since by Weyl's law [35], $\sum_{t_j \le T} 1 \sim \frac{T^2}{12}$. To each ϕ_j is associated its standard L-function $L(s, \phi_j)$ as well as its symmetric-square L-function, $L(s, \mathrm{sym}^2\phi_j)$. These and the other L-functions $L(s, \pi)$ that arise below have analytic continuations to $\mathbb{C}$ with a functional equation relating s to $1-s$. Our notation is that $L(s, \pi)$ is the finite part and $\Lambda(s.\pi)$ the completed L-function. While $L(1, \pi)$ is nonzero and depends mildly on π, $L(\frac{1}{2}, \pi)$ is a very subtle and much studied arithmetical invariant. For technical as well as arithmetical reasons it is natural to include weights in the variance sums (11). The "harmonic" weights $L(1, \mathrm{sym}^2\phi_j)$ satisfy

$$t_j^{-\epsilon} \ll_\epsilon L(1, \mathrm{sym}^2\phi_j) \ll_\epsilon t_j^\epsilon,$$

for $\epsilon > 0$ ([15], [17]) and they have a limiting distribution ([28]). In the end we can remove these harmonic weights as we do in Section 5 but for now we include them.

Theorem 1. – *Denote by $A_0(\Gamma\backslash G)$ the space of smooth right K-finite functions on $\Gamma\backslash G$ which are of mean 0 and of rapid decay. There is a sesquilinear form Q on $A_0(\Gamma\backslash G) \times A_0(\Gamma\backslash G)$ such that*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t_j \le T} L(1, \mathrm{sym}^2\phi_j)\omega_j(\psi_1)\overline{\omega}_j(\psi_2) = Q(\psi_1, \psi_2). \tag{12}$$

We call Q the quantum variance. The proof of Theorem 1 proceeds by proving the existence of the limit which comes with an explicit but formidable expression for Q, see (34) of Section 2. It involves infinite sums over arithmetic-geometric terms (twisted Kloosterman sums) and it appears very difficult to read any properties of Q directly from (34). For example even that Q is not identically zero (which is the case so that the exponent of T in the theorem is the correct one) is not clear. Using some a priori invariance properties of Q as well as some others that are derived from special cases of general versions of the daunting expression (34) allows us to eventually diagonalize Q.