EIGENVALUE SPACINGS FOR QUANTIZED CAT MAPS

ALEX GAMBURD, JOHN LAFFERTY, AND DAN ROCKMORE

Abstract. According to one of the basic conjectures in Quantum Chaos, the eigenvalues of a quantized chaotic Hamiltonian behave like the spectrum of the typical member of the appropriate ensemble of random matrices. We study one of the simplest examples of this phenomenon in the context of ergodic actions of groups generated by several linear toral automorphisms – "cat maps". Our numerical experiments indicate that for "generic" choices of cat maps, the unfolded consecutive spacings distribution in the irreducible components of the N-th quantization (given by the N-dimensional Weil representation) approaches the GOE/GSE law of Random Matrix Theory. For certain special "arithmetic" transformations, related to the Ramanujan graphs of Lubotzky, Phillips and Sarnak, the experiments indicate that the unfolded consecutive spacings distribution follows Poisson statistics; we provide a sharp estimate in that direction.

1. Introduction

Applications of Random Matrix Theory (RMT) in physics originated in Wigner's suggestion in the early fifties that the resonance lines of heavy nuclei, their determination by analytic means being intractable, might be modelled by the spectrum of a large random matrix [46]. While it was conceived of as a statistical approach to systems with many degrees of freedom, RMT also applies to systems with few degrees of freedom with chaotic classical dynamics; in fact, RMT lies at the heart of one of the basic conjectures in Quantum Chaos. Formulated by Bohigas, Giannoni, and Schmit in 1984 [8], it asserts that the eigenvalues of a quantized chaotic Hamiltonian (after suitable unfolding) behave like the spectrum of a typical member of the appropriate ensemble of random matrices. This conjecture complements an earlier conjecture of Berry and Tabor [6], asserting that the eigenvalues of quantized integrable systems follow the Poisson distribution; the Poisson distribution is also expected for *arithmetic* surfaces of constant negative curvature, following the pioneering numerical experiments in [42, 7, 10].

An important model for understanding quantization of classically chaotic systems is afforded by symplectic maps [47]. The simplest of these are the linear area-preserving transformations of the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$; that is, transformations $\binom{x_1}{x_2} \mapsto A\binom{x_1}{x_2}$ with $A \in SL(2,\mathbb{Z})$. These transformations,

The first author was supported in part by the NSF postdoctoral fellowship. The second author was supported in part by NSF grant CCR-0122581. The third author was supported in part by the Santa Fe Institute and AFOSR.

which have received a lot of attention in the physics and mathematics literature, go by the name "cat maps," which derives from the pictures in [2] which show a cartoon cat face and its images under a few iterates of A, displaying the chaotic features of $x \mapsto Ax$.

The quantization of such a linear transformation can be carried out by periodizing any one of the standard quantization procedures in $\mathbb{R}^2$. This has been carried out first by Hannay and Berry in [16] and has since been studied by many authors, see [3, 5, 12, 13, 22, 25] and references therein. We will adopt the quantization procedure given by Kurlberg and Rudnick in [27]. It yields for each integer $N \geq 1$ ("$N = 1/\hbar$") a unitary matrix $U_N(A)$ acting on $L^2(\mathbb{Z}/N\mathbb{Z})$. As we review in Section 2, and as detailed in [27], $U_N(A)$ is essentially the Weil or metaplectic representation of A reduced modulo N (first considered by Kloosterman [26]).

The behavior of the eigenstates of $U_N(A)$ has been the subject of intensive investigations in papers cited above, with important recent breakthroughs by Kurlberg and Rudnick [27]. The distribution of the eigenvalues of $U_N(A)$ is degenerate, and not what is expected for the quantization of a generic chaotic system, as shown by Keating [20]. Following an early attempt at restoring generic statistics by A. Lakshminarayan and N. L. Balazs in [31], several ways of recovering the predicted random matrix distribution for modified cat maps have been proposed. One approach, first considered by Basilio de Matos and Ozorio de Almeido in [4], and more recently by Keating and Mezzadri in [21], is to perturb a cat map by nonlinear shears; another, considered by Keppeler, Marklof, and Mezzadri in [23], is to couple a cat map with a two-spinor processing in a magnetic field.

In this paper we show how to recover the RMT predictions while staying within the framework of linear maps and representation theory. The basic idea, following [15], is to consider the ergodic action of the group generated by *several*[1] linear toral automorphisms, i.e., "several maps of a cat".

More precisely, let $\Gamma_{A_1,\ldots,A_k}$ be the group generated by the transformations $A_1, \ldots, A_k$ with $A_i \in SL(2,\mathbb{Z})$. The action of the group Γ is *strongly ergodic* if the associated element in the group ring of $SL(2,\mathbb{Z})$,

$$z_{A_1,\ldots,A_k} = A_1 + A_1^{-1} + \ldots + A_k + A_k^{-1} \tag{1}$$

has a spectral gap [43]. Let $\operatorname{supp}(z) = \{A_1, \ldots, A_k\}$ and let Γ_z be the group generated by $\operatorname{supp}(z)$. Numerical experiments [29, 30] indicate that a "generic" element z in the group ring of $SL(2,\mathbb{Z})$ has a spectral gap. In [14] it is proved that z has a spectral gap if the Hausdorff dimension of the limit set of Γ_z is large enough; see [44] for related results and [33, 34] for the discussion of this problem.

[1]The classical limit can be thought of as a random walk supported on the toral automorphisms in question, or, following Arnold and Krylov [1], as a dynamical system with noncommutative time.

We consider the quantizations $U_q(z)$,

$$U_q(z) = U_q(A_1) + U_q(A_1^{-1}) + \ldots + U_q(A_k) + U_q(A_k^{-1}).$$

For technical reasons, detailed below, we restrict ourselves to primes $q \equiv 1 \bmod 4$. The representation U_q is not irreducible but decomposes into two irreducible components U_q^- and U_q^+ of dimensions $\frac{1}{2}(q-1)$ and $\frac{1}{2}(q+1)$ respectively. With a suitable choice of basis, $U_q^+(z)$ lies in the space of real symmetric matrices, while $U_q^-(z)$ lies in the linear space of matrices H satisfying

$$(2) \quad H^* = H,\ J^tHJ = H^t,\ J = \begin{pmatrix} E & 0 & \ldots & 0 \\ 0 & E & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & E \end{pmatrix},\ E = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

In particular, in the latter case the eigenvalues of $U_q^-(z)$ are of the form $\lambda_1, \lambda_2, \ldots, \lambda_M$ where $M = (q-1)/4$ and each λ_j occurs with multiplicity 2.

The densities of eigenvalues are described by the sum of point masses

$$(3) \qquad \mu_q^+(z) = \frac{2}{q+1} \sum_{j=0}^{(q+1)/2} \delta_{\lambda_j(U_q^+(z))},$$

which is a probability measure supported in $[-2k, 2k]$ with a similar expression for $U_q^-(z)$:

$$(4) \qquad \mu_q^-(z) = \frac{4}{q-1} \sum_{j=0}^{(q-1)/4} \delta_{\lambda_j(U_q^-(z))}.$$

It is not difficult to show [38, 41, 15] that these converge to the measure $\nu_k(t)$, first considered by Kesten in [24], which is supported in the interval $[-2\sqrt{2k-1}, 2\sqrt{2k-1}]$ and given by

$$(5) \qquad d\nu_k(t) = \frac{\sqrt{2k-1-t^2/4}}{2\pi k(1-(t/2k)^2)}\ dt.$$

Our numerical experiments, described in Section 4 indicate that the unfolded consecutive spacing distribution for "generic" z (see the discussion in Section 4) follows the GSE law of Random Matrix Theory [37] for $U_q^-(z)$ and GOE law of Random Matrix Theory for $U_q^+(z)$.

We also consider certain arithmetic, or *Ramanujan* elements introduced by Lubotzky, Phillips and Sarnak [32], defined as follows. Let $H(\mathbb{Z})$ denote the ring of Hamilton quaternions $\alpha = x_0 + x_1 i + x_2 j + x_3 k, x_j \in \mathbb{Z}$. Let $\bar{\alpha} = x_0 - x_1 i - x_2 j - x_3 k$ and $N(\alpha) = \alpha\bar{\alpha}$. For $p \geq 3$ a prime number let $\tilde{g}_1, \tilde{g}_2, \ldots, \tilde{g}_k$ be a subset of $S = \{\alpha \in H(\mathbb{Z}) | N(\alpha) = p\}$ (it is well known [17] that S has $8(p+1)$ elements) satisfying

(1) $\tilde{g}_{j_1} \neq \varepsilon \tilde{g}_{j_2}$ for $j_1 \neq j_2$ and $\varepsilon \in \{\pm 1, \pm i, \pm j, \pm k\}$ a unit.
(2) $\tilde{g}_{j_1} \neq \varepsilon \overline{\tilde{g}_{j_2}}$ for any j_1, j_2 and ε a unit.

Among these there are $p+1$ elements with $x_0 > 0$ and odd and x_j even for $j = 1, 2, 3$.

Assume further that p and q are unequal primes, $p \equiv 1 \pmod 4$, $q \equiv 1 \pmod 4$. Let i be an integer satisfying $i \equiv -1 \pmod q$. To each element $\tilde{g}$ we can now associate the matrix g in $PGL_2(\mathbb{F}_q)$

$$\alpha \;\to\; \frac{1}{N(\alpha)} \begin{pmatrix} x_0 + x_1 i & x_2 + x_3 i \\ -x_2 + x_3 i & x_0 - x_1 i \end{pmatrix}$$

giving us the corresponding elements $g_1, g_2, \ldots, g_k \in PGL_2(\mathbb{F}_q)$. As detailed in [32], in the case $(\frac{p}{q}) = 1$, to which we restrict from now on, the elements g_j in fact lie in $PSL_2(\mathbb{F}_q)$. For example, for $p = 5$ we obtain the following matrices:

$$(6)\quad g_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1+2i & 0 \\ 0 & 1-2i \end{pmatrix}, \quad g_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}, \quad g_3 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2i \\ 2i & 1 \end{pmatrix}.$$

Let z_p denote the *Ramanujan* element; $\operatorname{supp}(z_p) = \{g_1, g_2, \ldots, g_k\}$ described above with $k = \frac{1}{2}(p+1)$. We find that the unfolded consecutive spacing distributions of $U_q^-(z_p)$ and $U_q^+(z_p)$ follow the Poisson distribution, in exact analogy to the arithmetic manifolds [42, 7, 10], and to the situation discussed in [15].

One way in which the difference between RMT and the Poisson distribution manifests itself is in the speed of convergence to the Kesten measure: the convergence is much faster in the RMT case [40, 15]. Our numerical experiments, detailed in Section 4, indicate that for a fixed z all irreducible representations of $SL_2(\mathbb{F}_q)$ behave in way similar to U_q^- and U_q^+ (depending on parity) with respect to spacings distributions; this is consistent with the observations made in [28, 29, 30] regarding the similarity of spectral properties of various irreducible representation of $SL_2(\mathbb{F}_q)$. As a result, the different behavior with respect to the speed of convergence is also apparent when we consider the image of z in the regular representation of $SL_2(\mathbb{F}_q)$. This is equivalent to considering the eigenvalues of the corresponding Cayley graphs of $SL_2(\mathbb{F}_q)$ with respect to the generators given by $\operatorname{supp}(z)$ reduced modulo q. We review the basic definitions in Section 5; see [28, 29, 30] for details.

Denote by $\mu_{X_q}(z)$ the empirical density for the Cayley graph of $SL_2(\mathbb{F}_q)$ associated with z. In Section 5 we prove the following result for a generic element z:

Theorem 1. *For q large enough*

$$D(\mu_{X_q(z)}, \nu_k) \;\ll_z\; \frac{1}{\log q},$$

where $k = |\operatorname{supp}(z)|$, and ν_k is the Kesten measure given by (5).

Here $D(\nu, \mu)$ is the discrepancy between the measures ν and μ; that is, $D(\nu, \mu) \;=\; \sup\{|\nu(I) - \mu(I)| : I = [a, b] \subset \mathbb{R}\}$.

In accordance with RMT predictions, supported by numerical experiments in Section 4, the discrepancy for the generic element should be $O((\log N)/N)$, where $N =\mid X_q \mid= O(q^3)$ (see Figure 3), so the result in Theorem 1 is probably very far from the truth.

For Ramanujan elements z_p, the spacings are Poisson and the discrepancy is not small (see Figure 3). We conclude section 5 by proving the following sharp lower bound, which is the analogue of the lower bounds for the remainder term in Weyl's law for arithmetic hyperbolic surfaces, see [18], [35], and [15].

Theorem 2. *Fix $p \geq 3$, let $X_{q,p}$ denote the Cayley graph of $SL_2(\mathbb{F}_q)$ associated with the Ramanujan element z_p. Let $k = \frac{1}{2}(p+1)$. Then*

$$D(\mu_{X_{q,p}}, \nu_k) \;\gg\; \frac{1}{q\log^2 q} = \frac{1}{\mid X_{p,q}\mid^{\frac{1}{3}} \log^2 \mid X_{p,q}\mid}.$$

2. Quantum Mechanics on the Torus.

In this section we briefly review the basics of quantum mechanics on a torus $\mathbb{T}^2$ viewed as a phase space; we follow [27], to which we refer for details. We will use abbreviations $e(z) = e^{2\pi i z}$ and $e_N(z) = e(z/N)$. Also, in this section p and q have their usual physical significance of momentum and position and carry no connotations of primality.

2.1. **Quantum States.** As the Hilbert space of states, we take distributions $\phi(q)$ on the line $\mathbb{R}$ which are periodic both in position and momentum representations: $\phi(q+1) = \phi(q)$, $[\mathcal{F}_h(\phi)](p+1) = [\mathcal{F}_h(\phi)(p)]$, where $[\mathcal{F}_h(\phi)](p) = h^{-1/2}\int \phi(q)e(-pq)/h)dq$. This restricts Planck's constant h to be an inverse integer. With $h = 1/N$ the Hilbert space of states $\mathcal{H}_N$ is of dimension N and consists of periodic point masses at the coordinates $q = Q/N$, $Q \in \mathbb{Z}$. We may then identify $\mathcal{H}_N$ with the N-dimensional vector space $L^2(\mathbb{Z}/N\mathbb{Z})$ with the inner product defined by

$$\langle \phi_1, \phi_2\rangle = \sum_{Q \bmod N} \frac{1}{N}\phi_1(Q)\overline{\phi_2(Q)}$$

2.2. **Observables.** Classical observables, i.e., functions $f \in C^\infty(\mathbb{T}^2)$, give rise to quantum observables, which are operators $\mathrm{Op}_N(f)$. To define these, start with translation operators

$$[t_1\phi](Q) = \phi(Q+1)$$

and

$$[t_2\phi](Q) = e_N(Q)\phi(Q)$$

which may be viewed as the analogues of differentiation and multiplication operators. In fact, in terms of the usual translation operators on the line

$\hat{q}\phi(q) = q\phi(q)$ and $\hat{p}\phi(q) = \frac{h}{2\pi i}\frac{d}{dq}\phi(q)$, they are given by $t_1 = e(\hat{p})$, $t_2 = e(\hat{q})$. In this context, the Heisenberg commutation relations read

$$t_1^a t_2^b = t_2^b t_1^a e_N(ab) \quad \forall a, b \in \mathbb{Z}$$

More generally, mixed translation operators are defined for $n = (n_1, n_2) \in \mathbb{Z}^2$ by

$$T_N(n) = e_N(2^{-1} n_1 n_2) t_2^{n_2} t_1^{n_1},$$

where 2^{-1} is the inverse of 2 in the field $\mathbb{F}_q$. These are unitary operators on $\mathcal{H}_N$, whose action on a wave-function $\phi \in \mathcal{H}_N$ is given by

$$T_N(n)\phi(Q) = e_N(2^{-1} n_1 n_2) e_N(n_2 Q)\phi(Q + n_1).$$

For any smooth function $f \in C^\infty(\mathbb{T}^2)$, define a quantum observable $\mathrm{Op}_N(f)$, called the *Weyl quantization of* f, by

$$\mathrm{Op}_N(f) = \sum_{n \in \mathbb{Z}^2} \hat{f}(n) T_N(n)$$

where $\hat{f}(n)$ are the Fourier coefficients of f.

2.3. **Cat maps.** A quantization of a smooth symplectic map A of the torus is a sequence of unitary maps $U_N : \mathcal{H}_N \to \mathcal{H}_N$ such that

$$U_N^* \mathrm{Op}_N(f) U_N - \mathrm{Op}_N(f \circ A) \to 0, \quad N \to \infty$$

The operator U_N is called the *quantum propagator*, its iterates give the evolution of the quantum system, and we require the quantum evolution to be asymptotic to classical evolution as $N \to \infty$; this is the analogue of Egorov's theorem.

In the case of a linear maps, $A \in SL(2, \mathbb{Z})$, one can construct a unitary operator $U_N(A)$ which satisfies an *exact* version of Egorov's theorem:

$$U_N(A)^* \mathrm{Op}_N(f) U_N(A) = \mathrm{Op}_N(f \circ A)$$

As detailed in [27] and as we review in the next section this quantum propagator is obtained by reducing A modulo N and considering the Weil representation.

3. Weil representation of $SL_2(\mathbb{F}_q)$

We now restrict N to be a prime q; we remark that consideration of Weil representations of $SL(2, \mathbb{Z}/2^k\mathbb{Z})$, $k > 1$, often necessary in physics, is technically more involved.

The group $SL_2(\mathbb{F}_q)$ is generated by the matrices of the form:

$$\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{7}$$

and so it suffices to specify U_q on such matrices.

We can choose as basis for the Weil representation the delta functions on $\mathbb{F}_q$,

$$\delta_x(y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases} \tag{8}$$

Using these, we get that the quantization of the generators (7) is given by the matrices of the following form. For $b \in \mathbb{F}_q$,

$$U_q\left(\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\right)\delta_y = e(by^2/q)\delta_y. \tag{9}$$

For $a \in \mathbb{F}_q^\times$,

$$U_q\left(\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}\right)\delta_y = \left(\frac{a}{q}\right)\delta_{a^{-1}y}, \tag{10}$$

and finally we have that

$$U_q\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)\delta_y(x) = \frac{\epsilon(q)}{\sqrt{q}}e(yx/q) \tag{11}$$

so that

$$U_q\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)\delta_y = \frac{\epsilon(q)}{\sqrt{q}}\sum_x e(yx/q)\delta_x \tag{12}$$

where

$$\epsilon(q) = \begin{cases} 1 & q \equiv 1 \pmod 4 \\ i & q \equiv 3 \pmod 4 \end{cases} \tag{13}$$

The representation U_q is not irreducible; it decomposes into two irreducible pieces of dimensions $(q-1)/2$ and $(q+1)/2$. One way to effect decomposition is the following. If we let

$$F(q) = U_q\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) \tag{14}$$

and let $S(q) = F(q)^2$ (so that $F^4 = I$ and $S^2 = I$), then the two irreducible pieces are

$$U_q^+ = U_q(I + S(q))/2 \tag{15}$$

and

$$U_q^- = U_q(I - S(q))/2. \tag{16}$$

4. Numerical Experiments

In this section we briefly review the relevant representation theory of $SL_2(\mathbb{F}_q)$, following the presentation of [28], and explain how the Weil representations can be recovered from specific irreducible representations. We then describe our numerical experiments, which compare the eigenvalue density and spacing distribution of the Weil representations for generic and Ramanujan elements.

The irreducible representations of $SL_2(\mathbb{F}_q)$ occur in two families, the discrete series representations and the principal series representations. The principal series representations are obtained as induced 1-dimensional representations from the upper triangular subgroup $B = \left\{ \begin{pmatrix} \alpha & b \\ 0 & \alpha^{-1} \end{pmatrix} \right\}$. More specifically, each character ψ of $\mathbb{F}_q^\times$ is associated with an induced representation ρ_ψ from B to $SL_2(\mathbb{F}_q)$. If $\psi^2 \neq 1$, then this representation is irreducible. However, if ψ is the trivial character or its unique nontrivial square root, denoted sgn, then ρ_ψ is reducible.

The U_q^+ component of the Weil representation can be realized as one of the two irreducible components of ρ_{sgn}. To explain, we recall from [28] that after fixing a generator α for $\mathbb{F}_q^\times$, a convenient basis for the principal series representations is given by ordering the basis elements as e_{α^j} with $j = 2, 4, q-3, 3, ..., 0, \infty$. Fixing a nontrivial additive character χ of $\mathbb{F}_q$, e.g., $\chi(n) = e^{2\pi i n/q}$, we set $d_\infty = e_\infty$ and for $x \in \mathbb{F}_q$ set $d_x = \sum_{t\in\mathbb{F}_q} \chi(xt)e_t$. The two irreducible pieces of ρ_{sgn}, of dimension $\frac{1}{2}(q+1)$, are the restrictions to the subspaces

(1) W_{sgn}^+ spanned by d_β where β is a square, and by $\epsilon(q)d_\infty + d_0$.
(2) W_{sgn}^- spanned by d_γ where γ is not a square and by $\epsilon(q)d_\infty - d_0$.

The U_q^+component of the Weil representation is the former and can be obtained by multiplying by a suitable projection matrix.

The U_q^- component of the Weil representation is realized from one of the discrete series representations. Each discrete series representation can be associated with a nondecomposable character of the quadratic extension $L = \mathbb{F}_q(\sqrt{\alpha})$; that is, a character having nontrivial restriction to the set of elements $C \subset L$ of norm one. The U_q^- component of the Weil representation is given in terms of the nondecomposable character ν for which ν^2 is trivial on C. In terms of the basis e_{α^j}, the two irreducible pieces of ρ_ν, of dimension $\frac{1}{2}(q-1)$, are the restrictions to subspaces

(1) V_ν^+ spanned by e_β where β is a square.
(2) V_ν^- spanned by e_γ where γ is not a square.

The Weil representation component U_q^- is the first one. It can be obtained by projection onto the even-labeled basis vectors, which in this case amounts to multiplying by a matrix whose diagonal is $(0, 1, 0, 1, \ldots, 0, 1)$.

There are several possible ways of introducing the notion of "generic" or "random" elements in $SL(2, \mathbb{Z})$; one is as follows: first pick an integer with

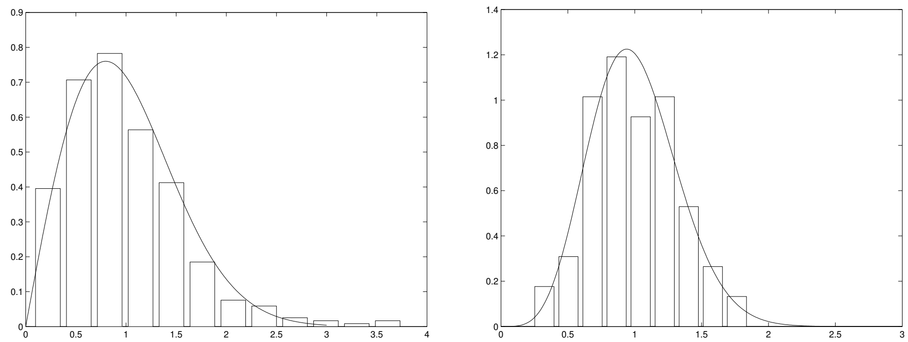

FIGURE 1. Spacing distribution for a random z with $|\operatorname{supp}(z)| = 3$ vs. the Wigner surmise for the GOE (left) for the irreducible component of the Weil representation U_{509} coming from the principal series representation induced from ψ_1, and the spacing distribution vs. the Wigner surmise for the GSE (right) for the irreducible component of the Weil representation coming from the discrete series representation with $\nu^2|_C = 1$.

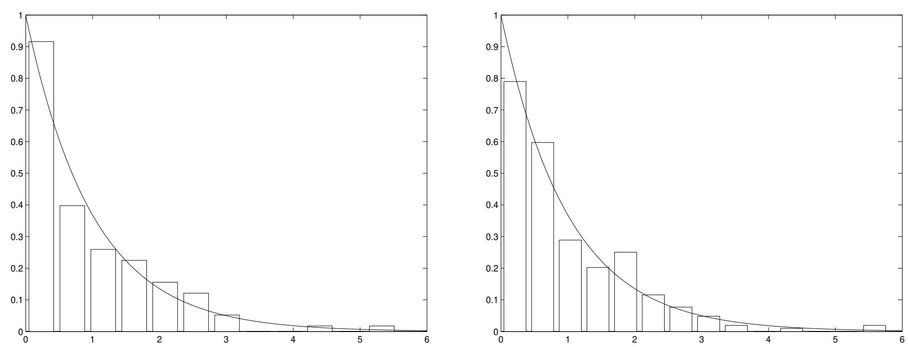

FIGURE 2. Spacing distribution for the Ramanujan element z_5 vs. the exponential distribution for the irreducible components of the Weil representation U_{509} coming from the principal series representation induced from ψ_1 (left), and from the discrete series representation with $\nu^2|_C = 1$.

Poisson distribution, then consider a random word of length chosen in the first step in the standard generators of $SL(2,\mathbb{Z})$. In Figures 1 and 2 we show the distribution of spacings between consecutive eigenvalues, where the eigenvalues are unfolded with respect to the Kesten measure to obtain average spacings equal to 1. This converges to GOE/GSE distributions for random generators and to the Poisson distribution for the Ramanujan generators. Plots comparing the empirical spacing distribution to the

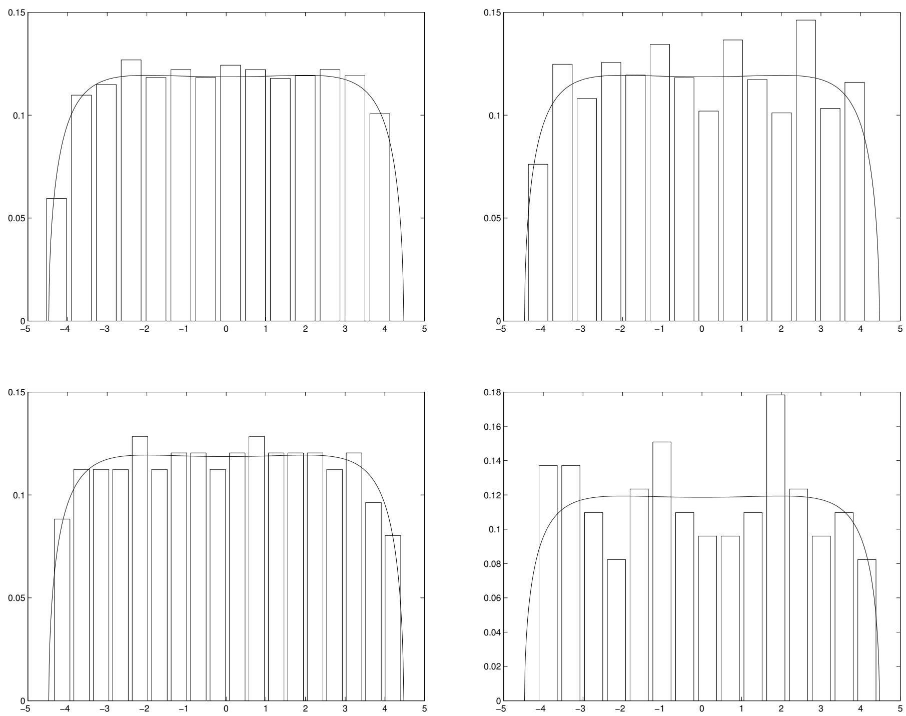

FIGURE 3. Eigenvalue density plots for empirical distributions versus the Kesten measure. The top row shows the distribution of the full spectrum of a random 6-regular Cayley graph (left) and the Ramanujan element z_5 for $SL_2(\mathbb{F}_{61})$ (right), corresponding to the spectrum of the regular representation. The bottom row shows the distribution of the spectrum for random and Ramanujan elements at an individual irreducible representation for $SL_2(\mathbb{F}_{509})$.

Wigner surmise for the appropriate ensemble are given in Figure 1 for random elements and compared to the exponential distribution in Figure 2 for Ramanujan elements.

Figure 3 shows the "density of states" versus the Kesten measure for random elements and for Ramanujan (Lubotzky-Phillips-Sarnak) elements for the regular representation, together with the distribution at an individual irreducible representation. The plots support the analysis presented in Theorems 1 and 2, with the Ramanujan elements having larger discrepancy than the random elements. We now turn to the proof of these results.

5. Discrepancy

5.1. **Trace formula for regular graphs.** We begin by reviewing the basic definitions, referring to [11, 45] for details. Let $X = (V, E)$ be a k-regular graph, that is a graph with each vertex having k neighbors. The *adjacency* *matrix* of X, $A(X)$ is the $\mid V \mid$ by $\mid V \mid$ matrix, with rows and columns indexed by vertices of X, such that the x, y entry is 1 if and only if x and y are adjacent and 0 otherwise. The *spectrum* of a graph is the spectrum of its adjacency matrix.

Let $N = \mid V \mid$. A is a symmetric matrix having N real eigenvalues which we can list in the decreasing order:

$$k = \lambda_0 > \lambda_1 \geq \ldots \geq \lambda_{N-1}.$$

To state the trace formula for a regular graph we need to recall a few definitions. A path *without back-tracking* of length r in X is a sequence of vertices in V, $x_0, x_1, \ldots, x_r$, such that x_i is adjacent to x_{i+1} for $i = 0, \ldots, r-1$ and $x_{i+1} \neq x_{i-1}$ for $i = 1, \ldots, r-1$. The origin of the path is x_0, the extremity is x_r. For $x \in V$, denote by $f_{l,x}$ the number of paths of length l in X without backtracking, with origin and extremity in x.

We are now ready to state the trace formula for k-regular graphs (which can be viewed as a discrete analogue of the Selberg's trace formula [18]). For every $m \in \mathbb{N}$ we have

$$\sum_{x \in V} \sum_{0 \leq r \leq \frac{m}{2}} f_{m-2r,x} = (k-1)^{\frac{m}{2}} \sum_{j=0}^{N-1} U_m\left(\frac{\lambda_j}{2\sqrt{k-1}}\right). \tag{17}$$

Here $U_m(x)$ are Chebyshev polynomials of the second kind defined as follows:

$$U_m(\cos\theta) = \frac{\sin(m+1)\theta}{\sin\theta} \quad (m \in \mathbb{N}). \tag{18}$$

We say that X is *vertex-transitive* if the group of automorphisms of X acts transitively on the vertex set V. Under this assumption, the number $f_{l,x}$ does not depend on the vertex x and we denote it simply by f_l. For vertex-transitive k-regular graphs on N vertices, the trace formula (17) takes the form

$$\sum_{0 \leq r \leq \frac{m}{2}} f_{m-2r} = \frac{1}{N}(k-1)^{\frac{m}{2}} \sum_{j=0}^{N-1} U_m\left(\frac{\lambda_j}{2\sqrt{k-1}}\right). \tag{19}$$

5.2. **Proof of Theorem 1.** Basic examples of vertex-transitive graphs are afforded by Cayley graphs. The *Cayley graph* for a group G with a generating set S, with $S = S^{-1}$, is the undirected graph with vertex set equal to G, such that there is an edge between a and b in X if and only if $as = b$ for some (necessarily unique) $s \in S$.

Now let $G_q = SL_2(\mathbb{F}_q)$, $N = \mid SL_2(\mathbb{F}_q) \mid = q(q^2-1)$. Let ϕ_q be the homomorphism of $SL(2,\mathbb{Z})$ onto $SL_2(\mathbb{F}_q)$ which associates with each matrix $B \in SL(2,\mathbb{Z})$ the matrix $\phi_q(B)$, obtained by reducing each element of X modulo q. Let z be a generic element, let $S(z) = \mathrm{supp}(z) = \{A_1, \ldots A_k\}$, let $S_q(z) = \phi_q(S(z))$ and consider the sequence of $2k$-regular graphs Cayley graphs of G_q with respect to the generators $S_q(z)$, $X_q(z) = X(G_q, S_q(z))$.

Proceeding exactly as in Section 2 of [14] we obtain that for for q large enough $\langle S_q(z)\rangle = SL_2(\mathbb{F}_q)$ and that, furthermore, the girth of $X_q(z)$, that is the length of its shortest circuit, satisfies the following estimate:

$$\text{girth}(X_q(z)) \geq 2\log_{\alpha(z)}(\frac{q}{2}) - 1, \tag{20}$$

where

$$\alpha(z) = \max_{L\in \text{supp}(z)} \|L\|.$$

Here the norm of a matrix L is defined by

$$\|L\| = \sup_{x\neq 0} \frac{\|Lx\|}{\|x\|},$$

and the norm of $x = (x_1, x_2)$ is given by $\|x\| = \sqrt{x_1^2 + x_2^2}$.

By Theorem 1.1. of McKay [36], the bound on girth (20) implies the convergence to Kesten measure.

Now let

$$F(X_q(z), t) = \int_{-2\sqrt{2k-1}}^{t} \mu_{X_q(z)}(x)dx$$

and

$$F_{2k}(t) = \int_{-2\sqrt{2k-1}}^{t} \nu_{2k}(x)dx.$$

By Theorem 4.4. of McKay [36], for every t we have

$$|F(X_q(z), t) - F_{2k}(t)| < \frac{24e\sqrt{2k-1}}{\pi^2(\text{girth}(X_q(z)) + 2)}.$$

This estimate, combined with the bound (20) completes the proof of Theorem 1.

We remark that numerical experiments in [19] indicate that the spacing distribution of eigenvalues of random regular graphs follows GOE distribution. Since random regular graphs asymptotically have logarithmic girth [9], the argument outlined above also applies to give similar discrepancy bound for the spectrum of random regular graphs.

5.3. **Proof of Theorem 2.** We now consider Ramanujan elements z_p; the associated Cayley graphs $X_{q,p} = X(G_q, S_q(z_p))$ are precisely the Ramanujan graphs considered in [32]. As shown in that paper, except for $\lambda_0 = 2k = p+1$, all the eigenvalues $\lambda_{j,N}$ of $X_{p,q}$ lie in the interval $[-2\sqrt{p}, 2\sqrt{p}]$. The graphs $X_{q,p}$ also satisfy the following sharp girth bound when $(\frac{p}{q}) = 1$ (the case to which we restrict ourselves):

$$\text{girth}(X_{p,q}) \geq 2\log_p q. \tag{21}$$

Our aim is to show that $D(\mu_{X_{q,p}}, \nu_{p+1})$ is large.

One checks easily that the discrepancy is invariant under continuous monotone changes of variable in the eigenvalue parameter[15]. Since $\lambda_{j,N} \in [-2\sqrt{p}, 2\sqrt{p}]$ it is convenient to use the variable $\theta_{j,N} \in [0,\pi]$, where

$$2\sqrt{p}\cos\theta_{j,N} \;=\; \lambda_{j,N}, \qquad 1 \le j \le N-1.$$

Set

$$\widetilde{\mu}_{X_{q,p}} \;=\; \frac{1}{N-1}\sum_{j=1}^{N-1} \delta_{\cos\theta_{j,N}}, \tag{22}$$

which is a probability measure on $[-1,1]$. Let $\widetilde{\nu}_{p+1}$ be the corresponding limit of the $\widetilde{\mu}_{X_{q,p}}$ as $q \to \infty$. Let

$$I_{n,q} \;=\; \int_{-1}^{1} \frac{\sin(n+1)t}{\sin t} d\widetilde{\mu}_{X_{q,p}}(t). \tag{23}$$

Now since for $n > 1$

$$\int_{-1}^{1} \frac{\sin(n+1)t}{\sin t} d\widetilde{\nu}_{p+1}(t) \;=\; 0, \tag{24}$$

integration by parts in (23) yields

$$|I_{n,q}| \;\le\; 2n^2\, D(\widetilde{\mu}_{X_{q,p}}, \widetilde{\nu}_{p+1}). \tag{25}$$

We will use (25) to give a lower bound for $D(\widetilde{\mu}_{X_{q,p}}, \widetilde{\nu}_{p+1})$.

To this end, we consider the trace formula (19) applied to the Ramanujan graphs $X_{q,p}$. Recalling the definition of Chebyshev polynomials (18), we obtain:

$$I_{n,q} = p^{-\frac{n}{2}} \left(\sum_{0 \le r \le \frac{n}{2}} f_{n-2r} - \frac{p^{\frac{n}{2}}}{N} U_n\!\left(\frac{\lambda_0}{2\sqrt{p}}\right) \right). \tag{26}$$

An easy computation shows that

$$U_n\!\left(\frac{\lambda_0}{2\sqrt{p}}\right) = U_n\!\left(\frac{p+1}{2\sqrt{p}}\right) = p^{-\frac{n}{2}} \frac{p^{n+1}-1}{p-1}. \tag{27}$$

As detailed in [32], for graphs $X_{q,p}$ the left-hand side of (19) has the following arithmetic interpretation:

$$\sum_{0 \le r \le \frac{m}{2}} f_{m-2r,x} = s_Q(p^m), \tag{28}$$

where $s_Q(p^m)$ is the number of integral representations of p^m by the quadratic form Q in 4 variables, defined by

$$Q(x_0,x_1,x_2,x_3) = x_0^2 + 4q^2(x_1^2 + x_2^2 + x_3^2).$$

The optimal estimate for $s_Q(p^n)$, is obtained by appealing to Ramanujan bounds proved by Eichler and Igusa (see [39]); the following bound, which suffices for our purposes, can be obtained by elementary means (see [11]):

$$s_Q(p^n) = O_\epsilon\left(\frac{p^{n(1+\epsilon)}}{q^3} + \frac{p^{n/2(1+2\epsilon)}}{q}\right) \tag{29}$$

Combining (26), (27), (28), and (29) and recalling that $N = \mid SL_2(\mathbb{F}_q)\mid = q(q^2-1)$, we have:

$$I_{n,q} = p^{\frac{-n}{2}}\left(O_\epsilon\left(\frac{p^{(n+\epsilon)}}{q^3} + \frac{p^{\frac{n}{2}(1+2\epsilon)}}{q}\right) - \frac{1}{q(q^2-1)}\frac{p^{n+1}-1}{p-1}\right). \tag{30}$$

Now keeping in mind the girth bound (21), we choose n such that

$$2\log_p q < n < 4\log_p q. \tag{31}$$

Substituting this choice of n into (30) and combining it with (25) completes the proof of Theorem 2.

References

[1] V. I. Arnold and A. L. Krylov, *Uniform distribution of points on a sphere and some ergodic properties of solutions of linear ordinary differential equations in the compex plane*, Soviet Math. Dokl., **4**, 1962, 1-5.

[2] V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, Benjamin, New York, 1968.

[3] Balian R. and Itzykson C. *Observations sur la mécanique quantique finie*, C. R. Acad. Sc. Paris **303**, (1986), 773-8.

[4] M. Basilio de Matos and A. Ozorio de Almeida, *Quantization of Anosov Maps*, Ann. Phys. **237** (1995), 46–65.

[5] De Bièvre S. , Degli Esposti M. and Giachetti R. *Quantization of a class of piecewise affine transformations on the torus*, Commun. Math. Phys., **176**, 1996, 73-94.

[6] M. Berry and M. Tabor Proc. Roy. Soc London A **356**, 1977, 375.

[7] E. Bogomolny, B. Georgeot, M.-J. Giannoni and C. Schmit. *Chaotic billiards generated by arithmetic groups* Phys. Rev. Lett. **69** (1992), no. 10, 1477–1480.

[8] O. Bohigas, M. Giannoni and C. Schmit. *Characterization of chaotic quantum spectra and universality of level fluctuation laws.* Phys. Rev. Lett. **52** (1984), no. 1, 1–4.

[9] Béla Bollobás, *Random graphs*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1985.

[10] J. Bolte, G. Steil and F. Steiner. *Arithmetical chaos and violation of universality in energy level statistics* Phys. Rev. Lett. 69 (1992), no. 15, 2188–2191.

[11] G. Davidoff, P. Sarnak, A. Valette *Elementary Number Theory, Group Theory, and Ramanujan Graphs*, Unpublished Manuscript.

[12] M. Degli Esposti, Quantization of the orientation preserving automorphism of a torus *Ann. Inst. Poincaré*, **58**, (1993), 323-341.

[13] M. Degli Espositi, S. Graffi and S. Isola, *Classical limit of the quantized hyperbolic toral automorphisms*, Comm. Math. Phys., **167**, (1995), 471-507.

[14] A. Gamburd, *Spectral gap for infinite index "congruence" subgroups of* $SL(2,\mathbb{Z})$, Israel Journal of Mathematics, **127**, (2002), 157-200.

[15] A. Gamburd, D. Jakobson, P. Sarnak, *Spectra of elements in the group ring of* SU(2), J. Eur. Math. Soc. **1**, 1999, 51-85.

[16] J. Hannay and M. Berry, *Quantization of linear maps on a torus – Fresnel diffraction by a periodic grating* Phys. D. **1D**, 1980, 267-209.

[17] G. Hardy and E. Wright. *An introduction to the theory of numbers,* fifth ed. Oxford Univ. Press, 1979.

[18] D. Hejhal. The Selberg trace formula for $PSL_2(\mathbb{R})$. Vol. I, LNM 548, Springer, 1976. Vol. II, LNM 1001, Springer, 1983.

[19] D. Jacobson, S. Miller, I. Rivin and Z. Rudnick, *Eigenvalue spacings for regular graphs*, in IMA Vol. Math. Appl. **109**, 1999, 317-327.

[20] J. Keating, *The cat map: quantum mechanics and classical motion*, Nonlinearity, **4**, (1991), 309-341.

[21] J. Keating, F. Mezzadri *Pseudo-symmetries of Anosov maps and spectral statistics*, Nonlinearity, **13** 2000, 747-775.

[22] J. Keating, F. Mezzadri, and J. Robbins *Quantum boundary conditions for torus maps*, Nonlinearity, **12**, (1999), 579-591.

[23] S. Keppeler, J. Marklof, F. Mezzadri *Quantum Cat Maps with spin* $1/2$, Nonlinearity, **14**, 2001, 719-738.

[24] H. Kesten. *Symmetric random walks on groups.* Trans. AMS 92 (1959), 336–354.

[25] S. Knabe *On the quantization of Arnold's cat*, J. Phys. A: Math. Gen., **23**, (1990), 2013-25.

[26] H. Kloosterman, *The behavior of general theta functions under the modular group and the characters of binary modular congruence groups*, Annals of Math., **47**, (1946), 317-375.

[27] P. Kurlberg and Z. Rudnick *Hecke theory and equidistribution for the quantization of linear maps of the torus* Duke Math. Jnl. **103**, No. 1, (2000), 47-77.

[28] J. Lafferty and D. Rockmore, *Fast Fourier analysis for* SL_2 *over a finite field and related numerical experiments*, Experimental Mathematics, **1**, 1992, 115-139.

[29] J. Lafferty and D. Rockmore, *Numerical investigation of the spectrum for certain families of Cayley graphs*, in DIMACS Series in Disc. Math. and Theor. Comp. Sci., Vol. 10, J. Friedman (ed.), 1993, 63-73.

[30] J. Lafferty and D. Rockmore, *Level spacings for Cayley graphs*, in IMA Vol. Math. Appl. **109**, 1999, 373-387.

[31] A. Lakshminarayan and N. L. Balazs, *On the quantum cat and sawtooth maps – Return to generic behaviour*, Chaos, Solitons and Fractals, **5**, 1995, 1169-1179.

[32] A. Lubotzky, R. Phillips, P. Sarnak, *Ramanujan Graphs*, Combinatorica **8** 1988, 261-277.

[33] A. Lubotzky and B. Weiss, *Groups and Expanders*, in DIMACS Series in Disc. Math. and Theor. Comp. Sci., Vol. 10, J. Friedman (ed.), 1993, 95-109.

[34] A. Lubotzky, *Cayley graphs: eigenvalues, expanders and random walks*, Surveys in Combinatorics, (P. Rowbinson ed.), London Math. Soc. Lecture Note Ser. 218, Cambridge Univ. Press, 1995, 155-189.

[35] W. Luo and P. Sarnak. *Number variance for arithmetic hyperbolic surfaces.* Comm. Math. Phys. 161 (1994), 419–432.

[36] B. McKay *The expected eigenvalue distribution of a large regular graph* Linear Algebra and its Applications **40** (1981) 203-216.

[37] M. Mehta. *Random matrices,* 2nd ed. Academic Press, 1991.

[38] P. Sarnak. *Statistical properties of eigenvalues of the Hecke operators,* in "Analytic number theory and Diophantine problems" (Stillwater, 1984), 321–331. Progr. Math. 70, Birkhauser, 1987.

[39] P. Sarnak, *Some Applications of Modular Forms*, Cambridge University Press, Vol. **99**, 1990.

[40] P. Sarnak, *Arithmetic Quantum Chaos*, Israel Math. Conf. Proceedings, **8**, (1995).

[41] J-P. Serre. *Repartition asymptotique des valeurs propres de l'operateur de Hecke* T_p. Jour. AMS 10 (1997), 75–102.

[42] C. Shmit, in *Chaos and Quantum Physics*, (M. Giannoni, A. Voros, J. Zinn-Justin, eds.), Elsevier Publ., 1991, pp. 333-369.

[43] K. Schmidt. *Amenability, Kazhdan's property T, strong ergodicity and invariant means for ergodic group-actions.* Ergod. Th. Dyn. Sys. 1 (1981), 223–236.

[44] Y. Shalom, *Expanding graphs and invariant means*, Combinatorica, **17**, 1997, 555-575.

[45] A. Terras *Fourier Analysis on Finite Groups and Applications* Cambridge Univeristy Press, 1999.

[46] E. Wigner *Random matrices in physics* SIAM Review, **9**, 1, 1967.

[47] S. Zelditch. *Index and dynamics of quantized contact transformations.* Ann. Inst. Fourier 47 (1997), 305–363.

Department of Mathematics, Stanford University, Stanford, CA 94305
E-mail address: `agamburd@math.stanford.edu`

School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213
E-mail address: `lafferty@cs.cmu.edu`

Department of Mathematics, Dartmouth College, Hanover, NH 03755
E-mail address: `rockmore@cs.dartmouth.edu`