# On Quantum Unique Ergodicity for linear maps of the torus

Zeév Rudnick

**Abstract.** The problem of "quantum ergodicity" addresses the limiting distribution of eigenfunctions of classically chaotic systems. I survey recent progress on this question in the case of quantum maps of the torus. This example leads to analogues of traditional problems in number theory, such as the classical conjecture of Gauss and Artin that any (reasonable) integer is a primitive root for infinitely many primes, and to variants of the notion of Hecke operators.

## 1. Introduction

One of the few rigorous general results in the field of "Quantum Chaology" is *Quantum Ergodicity* [13, 2, 14]. To formulate this notion, recall that if the classical dynamics of a system are ergodic, then almost all trajectories of a particle cover the energy shell uniformly, that is to say that the time averages along the trajectory converge to the phase space average. The intuition afforded by the "Correspondence Principle" leads one to look for an analogous statement about the semiclassical limit of expectation values of observables in an energy eigenstate. As formulated by Schnirelman [13], the corresponding assertion is that when the classical dynamics is ergodic, for almost all eigenstates the expectation values of observables converge to the phase-space average.

A key question is: Under suitable assumptions on the system, can one say anything beyond "almost all"? For instance, when can one assert that *all* expectation values converge to the phase space average? Such behavior has sometimes been called quantum *unique* ergodicity [12].

Below is a survey some of recent attempts to understand question in the context of quantum maps of the torus. These are an important model for understanding the quantization of classically chaotic systems, first studied by Hannay and Berry [5]. We will devote special attention to linear hyperbolic automorphisms of the torus $\mathbf{T}^2$ - the so called "cat maps".

**Notation:** We will use the abbreviations $e(z) := e^{2\pi iz}$, $e_N(z) = e(z/N)$.

## 2. Quantum mechanics on the torus

We review the basics of quantum mechanics on the torus $\mathbf{T}^2$, viewed as a phase space [5, 7, 3, 4].

### 2.1. Quantum states

We start with a description of the Hilbert space of states of such a system. In brief, Planck's constant is restricted to be an inverse integer: $h = 1/N$, and the Hilbert space of states $\mathcal{H}_N$ is $N$-dimensional, in keeping with the intuition that each state occupies a Planck cell of volume $h = 1/N$ and the constraint that the total phase-space $\mathbf{T}^2$ has volume one. The "state vectors" are distributions on the line which are periodic in both momentum and position representations: $\psi(q+1) = \psi(q)$, $[\mathcal{F}_h\psi](p+1) = [\mathcal{F}_h\psi](p)$, where $[\mathcal{F}_h\psi](p) = h^{-1/2}\int \psi(q)\, e(-pq/h)\, dq$. The space of such distributions is finite dimensional, of dimension precisely $N = 1/h$, and consists of periodic point-masses at the coordinates $q = Q/N$, $Q \in \mathbf{Z}$. We may then identify $\mathcal{H}_N$ with the $N$-dimensional vector space $L^2(\mathbf{Z}/N\mathbf{Z})$, with the inner product $\langle \cdot, \cdot \rangle$ defined by

$$\langle \phi, \psi \rangle = \frac{1}{N} \sum_{Q \bmod N} \phi(Q)\, \overline{\psi}(Q)\ .$$

### 2.2. Observables

Classical observables (i.e. functions $f \in C^\infty(\mathbf{T}^2)$) give rise to quantum observables, that is operators $\mathrm{Op}_N(f)$ on $\mathcal{H}_N$. To define these, one starts with the translation operators

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

and

$$[t_2\psi](Q) = e_N(Q)\, \psi(Q),$$

which may be viewed as the analogues of differentiation and multiplication (respectively) operators. In fact in terms of the usual translation operators on the line $\hat{q}\psi(q) = q\psi(q)$ and $\hat{p}\psi(q) = \frac{h}{2\pi i}\frac{d}{dq}\psi(q)$, they are given by $t_1 = e(\hat{p})$, $t_2 = e(\hat{q})$. In this context, Heisenberg's commutation relations read

$$t_1^a t_2^b = t_2^b t_1^a e_N(ab) \qquad \forall a, b \in \mathbf{Z}.$$

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

$$T_N(n) = e_N(\frac{n_1 n_2}{2}) t_2^{n_2} t_1^{n_1}.$$

These are unitary operators on $\mathcal{H}_N$, whose action on a wave-function $\psi \in \mathcal{H}_N$ is given by:

$$T_N(n)\psi(Q) = e^{\frac{i\pi n_1 n_2}{N}} e(\frac{n_2 Q}{N})\psi(Q+n_1)\ .$$

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

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

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

The observables $\mathrm{Op}_N(f)$ satisfy: For any orthonormal basis of $\mathcal{H}_N$ we have

$$\frac{1}{N} \sum_{j=1}^{N} \langle \mathrm{Op}_N(f)\psi_j, \psi_j \rangle = \int_{\mathbf{T}^2} f + O_f(\frac{1}{N}) \,. \tag{1}$$

That is, the *mean* of the expectation values is asymptotic to the classical average of the the observable $f$.

### 2.3. Dynamics

To introduce dynamics, we consider a smooth, area-preserving (symplectic) map $A$ of the torus. Iterating $A$ we get a discrete dynamical system. For instance, if $A \in SL(2, \mathbf{Z})$ is a linear automorphism then the system is well-known to be chaotic if $A$ is hyperbolic, that is $|\operatorname{tr} A| > 2$ (such a map is called a "cat map" in the physics literature). Another example is the "Kronecker map",

$$\tau_\alpha : x \mapsto x + \alpha \mod 1, \quad \alpha = (\alpha_1, \alpha_2) \,. \tag{2}$$

If $1, \alpha_1, \alpha_2$ are linearly independent over the rationals then this map is *uniquely ergodic* , i.e. the only $\tau_\alpha$-invariant probability measure on the torus is Lebesgue measure.

**Definition 2.1.** *A quantization of $A$ is a sequence of unitary maps $U_N : \mathcal{H}_N \to \mathcal{H}_N$ such that*

$$U_N^* \, \mathrm{Op}_N(f) U_N - \mathrm{Op}_N(f \circ A) \to 0, \quad N \to \infty \,. \tag{3}$$

The operator $U_N$ is called the *quantum propagator*, whose iterates give the evolution of the quantum system, and we require the quantum evolution to be asymptotic to the classical evolution as $N \to \infty$ (this is an analogue of "Egorov's theorem"). In this case we say that the map $A$ is "quantizable". The eigenfunctions of $U_N$ play the rôle of energy eigenstates.

In the example of the linear map $A$, if we further assume $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ab \equiv cd \equiv 0 \mod 2$, then on can construct a unitary operator $U_N(A)$ which satisfies an *exact* version of Egorov's theorem:

$$U_N(A)^* \, \mathrm{Op}_N(f) U_N(A) = \mathrm{Op}_N(f \circ A) \,. \tag{4}$$

A quantization of the Kronecker map $\tau_\alpha$ (2) satisfying (3) was constructed in joint work with Jens Marklof (see [10] for the closely related case of a skew translation), by first doing so for rational $\alpha$, in which case we have an exact Egorov theorem (4), and then for the general case by approximating $\alpha$ by rationals.

## 3. Quantum ergodicity

For quantum maps, the form that quantum ergodicity assumes is the following:

**Theorem 3.1** ([1, 15, 16])**.** *Let $A$ be a quantizable area-preserving map of the torus. Assume $A$ is* ergodic*. Then for any orthonormal basis $\psi_j$ of $\mathcal{H}_N$ consisting of eigenfunctions of $U_N(A)$, there is a subset $J(N) \subset \{1, 2, \dots, N\}$, with $\frac{\#J(N)}{N} \to 1$, so that for $j \in J(N)$ we have:*

$$\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle \to \int_{\mathbf{T}^2} f, \qquad \text{as } N \to \infty$$

*for all observables $f \in C^\infty(\mathbf{T}^2)$.*

Theorem 3.1 is a consequence, using positivity and a standard diagonalization argument, of the following estimate for the variance due to Zelditch [15] (see the Appendix for the proof):

**Theorem 3.2** ([15])**.** *Let $A$ be a quantizable area-preserving, ergodic map of the torus. For any orthonormal basis $\psi_j$, $j = 1, \dots, N$ of of $\mathcal{H}_N$ consisting of eigenfunctions of $U_N(A)$, we have*

$$\frac{1}{N}\sum_{j=1}^{N}\left|\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle - \int_{\mathbf{T}^2} f\right|^2 \to 0$$

*for all observables $f \in C^\infty(\mathbf{T}^2)$.*

A key problem is:

**Problem 3.3.** Is it true that *all* eigenfunctions become equidistributed as $N \to \infty$?

For the Kronecker map $\tau_\alpha$, Jens Marklof and I gave an affirmative answer (see [10] for skew translations):

**Theorem 3.4.** *If $1, \alpha_1, \alpha_2$ linearly independent over the rationals, then for all eigenfunctions $\psi$ of $U_N(\tau_\alpha)$,*

$$\langle \mathrm{Op}_N(f)\psi, \psi\rangle \to \int_{\mathbf{T}^2} f, \quad N \to \infty\,.$$

This is a consequence of the quantization procedure for the map coupled with the fact that for such $\alpha$, the map is classicaly uniquely ergodic.

In the most interesting case of *hyperbolic* maps (e.g. "cat maps"), a basic problem is the existence of several invariant measures.

## 4. Beyond quantum ergodicity for cat maps

I will now describe some recent attempts, joint with Pär Kurlberg, to improve on quantum ergodicity (Theorem 3.1) for cat maps.

**4.1. Hecke operators** [8]

It transpires that there is a commutative group of unitary operators on the state-space $\mathcal{H}_N$ which commute with the quantized map and therefore act on its eigenspaces. We call these "Hecke operators", in analogy with the setting of the modular surface.

To understand their origin, one needs to note that it is possible to define $U_N(A)$ so that it only depends on the remainder of $A \mod 2N$ and satisfies (4). One thus gets a *projective* representation $A \mapsto U_N(A)$ of the subgroup of "quantizable" elements in the finite modular group $SL(2, \mathbf{Z}/2N\mathbf{Z})$. It turns out that it can be made into an *ordinary* representation if we further restrict to the subgroup $\Gamma(4, 2N)$ given by $g = I \mod 4$ for $N$ even, $g = I \mod 2$ for $N$ odd. Thus for $A, B \in \Gamma(4, 2N)$ we have $U_N(AB) = U_N(A)U_N(B)$. Consequently, if $AB = BA \mod 2N$ then their propagators commute. This is the basic principle that we use to form the Hecke operators (see [6] for another application of this idea).

**Remark 4.1.** *The congruence $AB = BA \mod 2N$ is much less restrictive than the equation $AB = BA$. The latter has as its solutions in $SL(2, \mathbf{Z})$ essentially only $\pm$ powers of $A$ (at least for $A$ "primitive").*

**4.2. Equidistribution of Hecke eigenfunctions**

Since the Hecke operators commute with $U_N(A)$, they act on its eigenspaces, and since they commute with each other there is a basis of $\mathcal{H}_N$ consisting of joint eigenfunctions of $U_N(A)$ and the Hecke operators, whose elements we call Hecke eigenfunctions. We show

**Theorem 4.2** ([8])**.** *Let $A \in SL(2, \mathbf{Z})$ be hyperbolic, $A = I \mod 4$, and $f \in C^\infty(\mathbf{T}^2)$ a smooth observable. Then for all normalized Hecke eigenfunctions $\phi \in \mathcal{H}_N$ of $U_N(A)$, the expectation values $\langle \mathrm{Op}_N(f)\phi, \phi\rangle$ converge to the phase-space average of $f$ as $N \to \infty$. Moreover, for all $\epsilon > 0$ we have*

$$\langle \mathrm{Op}_N(f)\phi, \phi\rangle = \int_{\mathbf{T}^2} f(x)dx + O_{f,\epsilon}(N^{-1/4+\epsilon}), \quad \text{as } N \to \infty\,.$$

The exponent of $1/4$ in our theorem is certainly not optimal, and more likely the correct exponent is $1/2$. What we in fact show is that if $\phi_i$, $i = 1, \dots, N$ is an orthonormal basis of $\mathcal{H}_N$ consisting of Hecke eigenfunctions then

$$\sum_{i=1}^{N} \left| \langle \mathrm{Op}_N(f)\phi_i, \phi_i\rangle - \int_{\mathbf{T}^2} f(x)dx \right|^4 \ll N^{-1+\epsilon} \tag{5}$$

(compare Theorem 3.2). We deduce Theorem 4.2 from (5) by taking an orthonormal basis with $\phi_1 = \phi$ and omitting all but one term on the LHS. If all terms on the LHS of (5) are of roughly the same size then we would expect this to give the exponent $1/2$.

**Remark 4.3.** *The Hecke eigenspaces have small dimension (at most $O(\log\log N)$), while the eigenspaces of $U_N(A)$ may have large dimension. In fact, the* mean *degeneracy is $N/\operatorname{ord}(A, N)$ where* $\operatorname{ord}(A, N)$ *the* order *(or period) of $A$ modulo $N$,*

*that is the least integer $k \geq 1$ for which $A^k = I \mod N$. It can be shown that the mean degeneracy can be as large as $N/\log N$ for arbitrarily large $N$.*

### 4.3. Arbitrary eigenfunctions

Since not all eigenfunctions of $U_N(A)$ are Hecke eigenfunctions, we have not completely solved Problem 3.3 - whether *all* eigenfunctions become equidistributed, that is if we have quantum *unique* ergodicity. In [9], we show equidistribution of *all* eigenfunctions of $U_N(A)$ for *almost all* integers $N$:

**Theorem 4.4.** *Let $A \in SL(2, \mathbf{Z})$ be hyperbolic. There is a set of integers $\mathcal{N}^*$ of density one so that all eigenfunctions of $U_N(A)$ are equidistributed, as $N \to \infty$, $N \in \mathcal{N}^*$.*

Previously, the only result giving an infinite set of $N$ for which all eigenfunctions of $U_N(A)$ become equidistributed is by Degli-Esposti, Graffi and Isola [4], which conditional on the Generalized Riemann Hypothesis give an infinite set of *primes*.

A key step in the proof of Theorem 4.4 is an estimate for the fourth power moment of the expectation values, involving the order of $A$ modulo $N$:

**Theorem 4.5** ([9])**.** *There is a sequence of integers of density one so that for all observables $f \in C^\infty(\mathbf{T}^2)$ and any orthonormal basis $\{\psi_j\}_{j=1}^N$ of $\mathcal{H}_N$ consisting of eigenfunctions of $U_N(A)$ we have:*

$$\sum_{j=1}^{N} |\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle - \int_{\mathbf{T}^2} f|^4 \ll \frac{N(\log N)^{14}}{\mathrm{ord}(A,N)^2}\,.$$

Thus for any subsequence of integers $N$ such that

$$\frac{\mathrm{ord}(A,N)}{N^{1/2}(\log N)^7} \to \infty \tag{6}$$

(and satisfying an additional "genericity" assumption) we find that for all eigenfunctions of $U_N(A)$, $\langle \mathrm{Op}_N(f)\psi, \psi\rangle \to \int_{\mathbf{T}^2} f$ as $N \to \infty$.

### 4.4. Controlling the order of $A$ modulo $N$

Theorem 4.5 reduces the problem of quantum ergodicity to that of finding sequences of integers satisfying (6), a problem closely related to the classical Gauss-Artin problem of showing that any integer, other than $\pm 1$ or a perfect square, is a primitive root modulo infinitely many primes; see [11] for a survey. We show

**Theorem 4.6.** *Let $A \in SL(2, \mathbf{Z})$ be hyperbolic. Then there exist $\delta > 0$ and a density one subset $S$ of the integers such that for all $N \in S$ we have*

$$\mathrm{ord}(A,N) \gg N^{1/2}\exp((\log N)^\delta)\,.$$

Combining Theorem 4.6 with Theorem 4.5 gives Theorem 4.4.

**Remark 4.7.** *We note that condition* (6) *fails infinitely often. In fact one can show that there are arbitrarily large integers so that* $\mathrm{ord}(A,N)$ *is smaller than const.$N/\log N$.*

## Appendix A. Proof of quantum ergodicity

We review the proof of Quantum Ergodicity (Theorem 3.1) as given in [15], that is we show that given $f \in C^\infty(\mathbf{T}^2)$, for any orthonormal basis $\psi_j$, $j = 1, \dots, N$ of of $\mathcal{H}_N$ consisting of eigenfunctions of $U_N(A)$, we have

$$\frac{1}{N}\sum_{j=1}^{N}\left|\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle - \int_{\mathbf{T}^2} f\right|^2 \to 0 \tag{7}$$

We do this to emphasize the difference between it and our results (Theorems 4.2, (5), 4.4 and 4.5), which while far stronger than what is given by (7), requires methods that are special to the arithmetic structure of the cat map. In contrast, the argument below uses nothing more than Egorov's theorem and the *ergodicity* of the map. Without loss of generality, we will in the sequel assume that $\int_{\mathbf{T}^2} f = 0$.

We first recall some basic properties of the quantized observables $\mathrm{Op}_N(f)$:

1. The adjoint is given by
$$\mathrm{Op}_N(f)^* = \mathrm{Op}_N(\bar{f}) \,. \tag{8}$$
2. The composition of operators satisfies:
$$\mathrm{Op}_N(f)\,\mathrm{Op}_N(g) = \mathrm{Op}_N(fg) + O_{f,g}(\frac{1}{N}) \tag{9}$$
for $f, g \in C^\infty(\mathbf{T}^2)$.

We fix $T \geq 1$. By Egorov's theorem (3), as $N \to \infty$ we have

$$\frac{1}{T}\sum_{j=1}^{T}(U_N(A)^t)^*\,\mathrm{Op}_N(f)U_N(A)^t \sim \frac{1}{T}\sum_{t=1}^{T}\mathrm{Op}_N(f\circ A^t) = \mathrm{Op}_N(f^T)$$

where $f^T := \frac{1}{T}\sum_{t=1}^{T} f\circ A^t$ is the ergodic average of $f$. Moreover, if $\psi_j$ is an eigenfunction: $U_N(A)\psi_j = e^{i\lambda_j}\psi_j$, then

$$\begin{aligned}\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle &= \langle \mathrm{Op}_N(f)U_N(A)\psi_j, U_N(A)\psi_j\rangle \\ &= \langle U_N(A)^*\,\mathrm{Op}_N(f)U_N(A)\psi_j, \psi_j\rangle \\ &\sim \langle \mathrm{Op}_N(f\circ A)\psi_j, \psi_j\rangle \,.\end{aligned}$$

Consequently, if $\psi_j$ is an eigenfunction then for all $T \geq 0$,

$$\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle = \langle \mathrm{Op}_N(f^T)\psi_j, \psi_j\rangle \,.$$

Now we look at the sum (recall $\int f = 0$)

$$S_2(f,N) := \frac{1}{N}\sum_{j=1}^{N}|\langle \mathrm{Op}_N(f)\psi_j, \psi_j\rangle|^2 \,.$$

We will show that $\lim_{N\to\infty} S_2(f,N) = 0$.

By Egorov (3), we have $S_2(f,N) \sim S_2(f^T,N)$ as $N \to \infty$, for all $T \geq 1$. By Cauchy-Schwartz, we have

$$|\langle \mathrm{Op}_N(f^T)\psi_j, \psi_j\rangle|^2 \leq \|\mathrm{Op}_N(f^T)\psi_j\|^2\|\psi_j\|^2 = \langle \mathrm{Op}_N(f^T)^*\,\mathrm{Op}_N(f^T)\psi_j, \psi_j\rangle \,.$$

Moreover, by (8), (9),

$$\mathrm{Op}_N(f^T)^* \mathrm{Op}_N(f^T) = \mathrm{Op}_N(|f^T|^2) + O_{f,T}(\frac{1}{N})$$

and so

$$S_2(f,N) \lesssim \frac{1}{N}\sum_{j=1}^{N} \langle \mathrm{Op}_N(|f^T|^2)\psi_j, \psi_j\rangle + O_{f,T}(\frac{1}{N}) \, .$$

By (1) we thus find that for fixed $T \geq 1$,

$$\limsup S_2(f,N) \leq \int_{\mathbf{T}^2} |f^T|^2 \, .$$

So far we have used nothing about the cat map except Egorov's theorem. Now we use the fact that it is *ergodic*, in particular the *mean ergodic theorem* holds: For $F \in L^2(\mathbf{T}^2)$, the ergodic averages $F^T$ converge to $\int_{\mathbf{T}^2} F$ in $L^2$. Thus we have $\int_{\mathbf{T}^2} |f^T|^2 \to 0$ as $T \to \infty$. Therefore given $\epsilon > 0$, we can find $T = T(f,\epsilon)$ for which $\int_{\mathbf{T}^2} |f^T|^2 < \epsilon$ and consequently

$$\limsup S_2(f,N) < \epsilon$$

which shows that $S_2(f,N) \to 0$ as required. □

## References


[1] A. Bouzouina and S. De Bièvre *Equipartition of the eigenfunctions of quantized ergodic maps on the torus*, Comm. Math. Phys. **178** (1996) 83-105.

[2] Y. Colin de Verdière, *Ergodicité et fonctions propres du laplacien*, Comm. Math. Phys. **102** (1985), 497-502.

[3] M. Degli Esposti *Quantization of the orientation preserving automorphisms of the torus*, Ann. Inst. Poincaré **58** (1993), 323–341.

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

[5] J.H. Hannay and M.V. Berry *Quantization of linear maps on a torus - Fresnel diffraction by a periodic grating*, Physica D **1** (1980), 267–291.

[6] J. Keating and F. Mezzadri *Pseudo-Symmetries of Anosov Maps and Spectral Statistics*, preprint.

[7] S. Knabe *On the quantisation of Arnold's cat*, J.Phys. A: Math. Gen. **23** (1990), 2013–2025.

[8] P. Kurlberg and Z. Rudnick *Hecke theory and equidistribution for the quantization of linear maps of the torus*, preprint chao-dyn/9901031, to appear in Duke Math. J.

[9] P. Kurlberg and Z. Rudnick *On quantum ergodicity for linear maps of the torus*, preprint math/9910145, to appear in Comm. Math. Phys.

[10] J. Marklof and Z. Rudnick *Quantum unique ergodicity for parabolic maps*, preprint math-ph/9901001, to appear in Geom. and Funct. Analysis.

[11] M. Murty *Artin's conjecture for primitive roots* Math. Intelligencer **10** (1988), no. 4, 59–67.

[12] Z. Rudnick and P. Sarnak *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math Phys. **161** (1994), 195–213.

[13] A. Schnirelman *Ergodic properties of eigenfunctions*, Usp. Math. Nauk **29** (1974), 181–182.

[14] S. Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*, Duke Math. J. **55** (1987), 919-941.

[15] S. Zelditch *Quantum ergodicity of $C^*$-dynamical systems*, Comm. Math.Phys. **177** (1996), 507–528.

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

Raymond and Beverly Sackler School of Mathematical Sciences,
Tel Aviv University,
Tel Aviv 69978, Israel
*E-mail address*: rudnick@math.tau.ac.il