Annales Henri Poincaré

On the Rate of Quantum Ergodicity for Quantised Maps

Roman Schubert

Abstract. We study the distribution of expectation values and transition amplitudes for quantised maps on the torus. If the classical map is ergodic then the variance of the distribution of expectation values will tend to zero in the semiclassical limit by the quantum ergodicity theorem. Similarly the variance of transition amplitude goes to zero if the map is weak mixing. In this paper we derive estimates on the rate by which these variances tend to zero. For a class of hyperbolic maps we derive a rate which is logarithmic in the semiclassical parameter, and then show that this bound is sharp for cat maps. For a parabolic map we get an algebraic rate which again is sharp.

1. Introduction

The behaviour of eigenfunctions and eigenvalues of Schrödinger operators in the semiclassical limit depends strongly on the dynamical properties of the underlying classical Hamiltonian system. For a generic chaotic system it is conjectured that the eigenvalue and eigenfunction statistics behave universally. By the Bohigas Giannoni Schmidt conjecture, [5], the eigenvalue statistics can be modelled by random matrix theory, and Berry's random wave model, [4], predicts that the eigenfunctions behave locally like random superposition of plane waves.

One way of characterising the behaviour of eigenfunctions is by studying expectation values of observables and a fundamental result about them is the quantum ergodicity theorem. It states that if the underlying classical system is ergodic, then for sufficiently nice observables and almost all eigenfunction the expectation values tend to the classical mean value of the observable. This result goes back to Shnirelman, Zelditch and Colin de Verdière, [9, 31, 36], for the eigenfunctions of the Laplacian on compact manifolds, and has since then been generalised to many other situations.

But the question on the rate by which the expectation values approach the classical mean has proved to be very difficult. For the eigenfunctions of the Laplacian on a compact manifold of negative curvature, Zelditch, [38], proved an upper

bound on the mean behaviour of the expectation values which decays logarithmically in the semiclassical parameter. This bound has been recently generalised in [30], but the order could not be improved. It is believed that this bound is far from the optimal one, and this view is supported by conjectures, numerical studies and physical heuristics, see [1, 3, 16, 18, 29, 35]. Only for the case of the modular surface Luo and Sarnak were recently able to compute the variance of the matrix elements, [25], and so obtained the optimal bound.

Since for Schrödinger operators estimating the rate of quantum ergodicity seems to be a very hard problem it is interesting to look for simpler classes of system where this question might be attacked with more hope of success. This is the reason why we consider quantised maps on the two-dimensional torus in this article. For these maps we sometimes have better control on the semiclassical approximations, which will allow us to understand the rate of quantum ergodicity better.

The first class of systems we study are perturbed cat maps. These are hyperbolic maps and we obtain for their quantisations the same logarithmic upper bound on the mean rate of quantum ergodicity as in [30, 38]. For the unperturbed cat maps Kurlberg and Rudnick, [23], showed that the variance of matrix elements decays much quicker for a special choice of basis of eigenfunctions, a so called Hecke eigenbasis. But since the eigenvalues of the cat map can have high multiplicities, there are many choices of a eigenbasis. We will show that there exists an eigenbasis for which the variance of expectation values is of logarithmic order in the semiclassical parameter. As will be explained below this is due to very particular properties of cat maps, and is probably not true for a generic perturbation of the cat-map.

In our methods the hyperbolicity of the classical system is the source of large error terms in semiclassical approximations of time evolution, so one might wonder if we get better bounds if the system is ergodic, but not hyperbolic. The second class of quantised maps we study are exactly of this type. These are parabolic maps whose quantisations have been studied by Marklov and Rudnick, [27], who have derived sharp bounds on individual eigenfunctions. So they are interesting test cases for us, and we show that our general method produces sharp results in this case, too.

Quantum ergodicity is a quite general and stable result, it relies only on a few properties, like the validity of some correspondence principle between quantum and classical mechanics, and it therefore has been proved in many different situations. The main general insight to be drawn from the results presented here is that in this general framework the logarithmic bound from Zelditch [30, 38] is sharp, as we show for the cat map, and in order to improve the bounds one needs additional assumptions.

This paper is organised as follows. In the first two sections we describe the general setup. In the first section we recall some notions on ergodic maps on the torus. In the second section we recall the quantisation of maps on the torus and state the quantum ergodicity theorems relating variances of expectation values

and transition amplitudes to ergodicity and weak mixing of the classical map. We then state a general result which relates the variance of expectation values and transition amplitudes to an average of the classical autocorrelation function, plus semiclassical error terms. This result shows that the quality of bounds we have on the error terms for the semiclassical time evolution determines the rate of quantum ergodicity, and it will be our main tool in the following sections.

In the last two sections we then study explicit classes of maps and apply the method of Section 3 to them. In the first part of Section 4 we obtain a logarithmic upper bound on the rate for perturbed cat maps. In the second part we show that this bound is sharp for unperturbed cat maps. In the last section we get a bound on the rate of quantum ergodicity for parabolic maps, which is algebraic in the semiclassical parameter. By the results of Marklof and Rudnick, [27], this bound is sharp, too.

Notation. We will use the notation $e(x) := e^{2\pi i x}$ and $e_N(x) := e(x/N)$. Furthermore we will write $e(*)$ for a phase factor, i.e., a number with $|e(*)| = 1$, the value of $e(*)$ can change from line to line.

2. Maps on the torus

Let $T^2 := \mathbb{R}^2/\mathbb{Z}^2$ be the two-torus and dx the normalised Lebesgue measure on T^2. We will study some classes of smooth symplectic maps on the torus

$$\Phi : T^2 \mapsto T^2 \,. \tag{1}$$

In this section we will recall ergodic notions we will need later on, see [34] for more details.

For $a \in L^2(T^2)$ we define the average

$$\bar{a} := \int_{T^2} a(x) \, dx \tag{2}$$

and the autocorrelation function as

$$C[a](t) := \int_{T^2} a^*(x) a\big(\Phi^t(x)\big) \, dx - |\bar{a}|^2 \,, \tag{3}$$

for $t \in \mathbb{Z}$. Properties like ergodicity, weak mixing and mixing can be expressed in terms of the autocorrelation function. The map is ergodic if for every $a \in L^2(T^2)$ one has

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} C[a](t) = 0 \,, \tag{4}$$

it is weak mixing if for every $\theta \in \mathbb{R}$ and $a \in L^2(T^2)$ one has

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} e(\theta t) C[a](t) = 0 \,, \tag{5}$$

and it is mixing if for every $a \in L^2(T^2)$

$$\lim_{t \to \infty} C[a](t) = 0. \tag{6}$$

The map is called exponentially mixing if there is a $\gamma > 0$ such that for any $a \in C^\infty(T^2)$

$$C[a](t) = O_a(e^{-\gamma|t|}). \tag{7}$$

If a map Φ is mixing and the correlation functions for smooth observables decay at least like $|C[a](t)| = O(t^{-1-\delta})$ for some $\delta > 0$, then we can define for $a \in C^\infty(T^2)$ the quantity

$$V[a] := \sum_{t=-\infty}^{\infty} C[a](t). \tag{8}$$

One can show that under the same condition on the decay of $C[a](t)$ one has

$$V[a] = \lim_{T \to \infty} \frac{1}{T} \int_{T^2} \left| \sum_{t=0}^{T-1} \left(a(\Phi^t(x)) - \bar{a} \right) \right|^2 \, \mathrm{d}x, \tag{9}$$

i.e., $V[a]$ is actually the variance of the fluctuations of the time mean around the space mean. This representation shows as well that $V[a] \geq 0$. Since we will only work with (8) and not with (9) we skip the easy proof of their equivalence.

3. Quantum maps and quantum ergodicity

Let us first quickly review the setup for quantised maps on the 2-torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$, see [10, 13, 26] for some recent and more complete treatments, we follow here the presentation in [7].

The Hilbert space is constructed from generalised functions on the line by requiring periodicity in position and in momentum space. For $x = (p, q) \in \mathbb{R}^2$ let

$$T(p, q) := \exp\left(\frac{\mathrm{i}}{\hbar}(pQ - qP) \right), \tag{10}$$

be the phase space translation operator acting on functions on $\mathbb{R}$, where $Q\psi(q) := q\psi(q)$ and $P\psi(q) := \frac{\hbar}{\mathrm{i}} \frac{\mathrm{d}\psi(q)}{\mathrm{d}q}$ are the position and momentum operators, respectively. The simultaneous periodicity condition reads

$$T(1, 0)\psi = \mathrm{e}(\kappa_1)\psi, \quad T(0, 1)\psi = \mathrm{e}(\kappa_2)\psi, \tag{11}$$

where two constant phase factors $\kappa = (\kappa_1, \kappa_2) \in \mathbb{R}^2/\mathbb{Z}^2$ are allowed because the quantum mechanical state represented by a wave function does not depend on an overall phase. This set of equations has non-trivial solutions only when

$$\frac{1}{2\pi\hbar} = N \in \mathbb{N} \tag{12}$$

and then the vector space $\mathcal{H}_N(\kappa)$ of solutions has dimension N and consists of distributions of the form

$$\psi(q) = \sum_{k \in \mathbb{Z}} \Psi(k)\delta\big(q - (k + \kappa_1)/N\big), \quad \text{with} \quad \Psi(k + N) = \mathrm{e}(-\kappa_2)\Psi(k). \quad (13)$$

The translation operators

$$T_N(n) := T(n_1/N, n_2/N), \quad (14)$$

with $2\pi\hbar = 1/N$ and where $n = (n_1, n_1) \in \mathbb{Z}^2$, leave the space $\mathcal{H}_N(\kappa)$ invariant and if one equips $\mathcal{H}_N(\kappa)$ with the inner product

$$\langle \psi, \phi \rangle = \frac{1}{N}\sum_{q=1}^{N} \Psi^*(q)\Phi(q), \quad (15)$$

then the restrictions of $T_N(n)$ to $\mathcal{H}_N(\kappa)$ become unitary. By slight abuse of notation we will denote the restrictions of $T_N(n)$ to $\mathcal{H}_N(\kappa)$ as well by $T_N(n)$.

So instead of one fixed Hilbert space, as in quantum mechanics on $\mathbb{R}^n$, the quantisation of the 2-torus leads for each $\hbar$ with $1/\hbar = N \in \mathbb{N}$ to a family of Hilbert spaces $\mathcal{H}_N(\kappa)$ depending on an additional parameter κ. But for each class of maps we study we will fix for each N a particular choice of κ, κ_N, and therefore we will most of the time discard it from the notation.

We will need some further properties of the translation operators $T_N(n)$, they satisfy

$$T_N(m)T_N(n) = \mathrm{e}_N\big(\omega(m,n)/2\big)T_N(m+n), \quad (16)$$

where $\omega(m,n) = m_1 n_2 - m_2 n_1$, and $T_N^*(n) = T_N(-n)$. Furthermore their trace satisfies $\operatorname{Tr} T_N(n) \neq 0$ only if $n = 0 \mod N$, and

$$\frac{1}{N}\operatorname{Tr} T_N^*(n)T_N(m) = \begin{cases} \mathrm{e}(*) & \text{if} \quad n = m \mod N \\ 0 & \text{otherwise} \end{cases}, \quad (17)$$

and in the particular case that $n = m$ the phase factor is $\mathrm{e}(*) = 1$.

The quantisation of observables is now defined by a Weyl quantisation prescription using the translation operators. For $a \in C^\infty(T^2)$ one defines the Weyl quantisation as

$$\operatorname{Op}_N[a] := \sum_{n \in \mathbb{Z}^2} \hat{a}(n)T_N(n) \quad (18)$$

with the Fourier coefficients

$$\hat{a}(n) = \int_{T^2} a(x)\mathrm{e}\big(-\omega(n,x)\big)\,\mathrm{d}x. \quad (19)$$

The function a is called the symbol of the operator $\operatorname{Op}_N[a]$. This quantisation prescription has natural properties, e.g., real valued symbols correspond to selfadjoint operators. Note that in particular $T_N(n) = \operatorname{Op}_N[\mathrm{e}(\omega(n,x))]$.

Traces are related to the phase space average of the symbols, using (18) one finds

$$\operatorname{Tr} \operatorname{Op}_N[a] = \int_{T^2} a(x) \, \mathrm{d}x + O_a(1/N^\infty) \,, \tag{20}$$

for $a \in C^\infty(T^2)$, where the remainder depends on derivatives of a. One related more precise result we will need later on is

Lemma 1. *Let $a, b \in C^\infty(T^2)$, then for all $L \geq 3$ we have*

$$\frac{1}{N} \operatorname{Tr} \operatorname{Op}_N[a] \operatorname{Op}_N[b] = \int_{T^2} a(x)b(x) \, \mathrm{d}x + O\left(\frac{|a|_L |b|_{3+L}}{N^L}\right). \tag{21}$$

where $|a|_L := \sum_{|\alpha| \leq L} \sup_{x \in T^2} |\partial^\alpha a(x)|$.

Proof. Using the definition (18) and (17) we obtain

$$
\begin{aligned}
\frac{1}{N} \operatorname{Tr} \operatorname{Op}_N[a] \operatorname{Op}_N[b] &= \sum_{n \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}^2} \hat{a}(n)\hat{b}(-n+mN)\mathrm{e}(*) \\
&= \sum_{n \in \mathbb{Z}^2} \hat{a}(n)\hat{b}(-n) \\
&\quad + \sum_{n \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}^2 \setminus \{0\}} \hat{a}(n)\hat{b}(-n+mN)\mathrm{e}(*) \,.
\end{aligned}
\tag{22}
$$

Now we have for the first term

$$\int_{T^2} a(x)b(x) \, \mathrm{d}x = \sum_{n \in \mathbb{Z}^2} \hat{a}(n)\hat{b}(-n) \tag{23}$$

and by partial integration for $m \neq 0$

$$|\hat{a}(n)| \leq C|a|_k(1+|n|)^{-k} \,, \quad |\hat{b}(-n+mN)| \leq C|b|_L(1+|n|)^L(N|m|)^{-L} \,, \tag{24}$$

so with $k = L + 3$ and $L \geq 3$ the remainder term converges and the result follows. □

Finally we look at the quantisation of maps. Let $\Phi : T^2 \to T^2$ be a symplectic diffeomorphism on the torus, we say that a sequence of unitary operators $\{U_N\}_{N \in \mathbb{N}}$, $U_N : \mathcal{H}_N \to \mathcal{H}_N$, is a quantisation of Φ if for all $a \in C^\infty(T^2)$ we have

$$\lim_{N \to \infty} \|U_N^* \operatorname{Op}_N[a]U_N - \operatorname{Op}_N[a \circ \Phi]\| = 0 \,. \tag{25}$$

This property is a manifestation of the correspondence principle, it means that for large N the quantum system reproduces the classical system. In the mathematical literature this is usually referred to as the validity of an Egorov theorem. A more quantitative version is usually true,

$$U_N^{-t} \operatorname{Op}_N[a]U_N^t = \operatorname{Op}_N[a \circ \Phi^t] + \frac{1}{N} R_N(t) \,. \tag{26}$$

with $\|R_N(t)\| \leq C$ for t in a finite interval. If we let t increase then the available bounds for $\|R_N(t)\|$ increase exponentially for generic hyperbolic maps.

The concrete procedure by which the unitary operators U_N are constructed depends on the map and we will discuss this for particular classes of examples below.

In case that the classical map is ergodic the Egorov property (25) allows to prove a quantum ergodicity theorem. We will recall the statements now. In their formulation we will use the notation

$$|\theta|_{S^1} := \min_{k \in \mathbb{Z}} |\theta + k| \,, \tag{27}$$

for $\theta \in \mathbb{R}$.

Theorem 1. *Assume the map $\Phi : T^2 \to T^2$ is ergodic, and U_N is a quantisation of Φ. Let for every $N \in \mathbb{N}$, ψ_j^N, $j = 1, \ldots, N$, be an orthonormal basis of eigenfunctions with eigenvalues $\mathrm{e}(\theta_j^N)$ of U_N and let $a \in C^\infty(T^2)$, then we have*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^N \left| \langle \psi_j^N, \mathrm{Op}_N[a]\psi_j^N \rangle - \overline{a} \right|^2 = 0 \,. \tag{28}$$

Furthermore, Φ is ergodic if, and only if, for every $a \in C^\infty(T^2)$

$$\lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N} \sum_{|\theta_i^N - \theta_j^N|_{S^1} \leq \delta} \left| \langle \psi_i^N, \mathrm{Op}_N[a]\psi_j^N \rangle - \delta_{ij}\overline{a} \right|^2 = 0 \,. \tag{29}$$

The result (28) was proved in [7], and (29) is a version for maps of a result in [37], see as well [32]. We will give below a simple proof in order to prepare for the estimates on the rate later on.

The estimate (28) means that the variance of the distribution of the expectation values around its mean is going to 0. This implies that almost all of the expectation values tend to the mean value, which is the usual way in which quantum ergodicity is stated. The second estimate (29) implies of course the first one, but it is stronger and implies in addition that almost all of the near diagonal transition amplitudes have to tend to 0.

Theorem 2. *Assume the map $\Phi : T^2 \to T^2$ is weak mixing, and U_N is a quantisation of Φ. Let for every $N \in \mathbb{N}$, ψ_j^N, $j = 1, \ldots, N$, be an orthonormal basis of eigenfunctions of U_N, with eigenvalues $\mathrm{e}(\theta_i^N)$, and let $a \in C^\infty(T^2)$, then we have for any $\theta \in \mathbb{R}$*

$$\lim_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N} \sum_{|\theta_i^N - \theta_j^N - \theta|_{S^1} \leq \delta} \left| \langle \psi_i^N, \mathrm{Op}_N[a]\psi_j^N - \delta_{ij}\overline{a} \rangle \right|^2 = 0 \,. \tag{30}$$

Conversely, if (30) holds for every $a \in C^\infty(T^2)$ and $\theta \in \mathbb{R}$, then Φ is weak mixing.

The analogue of this result for eigenfunctions of the Laplacian on compact Riemannian manifolds was given in [37, 39].

Our main objective in this paper is to obtain bounds on the rate by which the averages over expectation values and transition amplitudes in Theorems 1 and 2

tend to zero. Our main tool to do so will be the next proposition. In its statement and in the following we will fix a function $f \in \mathcal{S}(\mathbb{R})$ which satisfies

$$f, \hat{f} \geq 0, \quad f(0) = 1 \quad \text{and} \quad \text{supp} \, \hat{f} \in [-1, 1] \tag{31}$$

and then define for $T > 0$

$$\hat{f}_T(t) := \frac{1}{T} \hat{f}\left(\frac{t}{T}\right) \quad \text{and} \quad F_T(\theta) := \sum_{n \in \mathbb{Z}} f\big(T(\theta - n)\big). \tag{32}$$

Such a function can be easily constructed by choosing a function $g > 0$ with support in $[-1/2, 1/2]$ and taking $\hat{f} = g * g$, and then normalising g such that $f(0) = 1$.

Note that we can always set $\bar{a} = 0$ without loss of generality. This will shorten some formulas.

Proposition 1. *Let* $U_N : \mathcal{H}_N \to \mathcal{H}_N$ *be a quantisation of the map* $\Phi : T^2 \to T^2$, *and let* ψ_j^N, $j = 1, \ldots, N$, *be an orthonormal basis of eigenfunctions of* U_N *with eigenvalues* $\mathrm{e}(\theta_j^N)$. *Then for every* $a \in C^\infty(T^2)$ *with* $\bar{a} = 0$ *we have*

$$\frac{1}{N} \sum_{j=1}^{N} \big|\langle \psi_j^N, \mathrm{Op}_N[a]\psi_j^N \rangle\big|^2 \leq \sum_{t \in \mathbb{Z}} \hat{f}_T(t) C[a](t)$$

$$+ \frac{1}{N} \sum_{t \in \mathbb{Z}} \hat{f}_T(t) \|R_N(t)\| \tag{33}$$

$$+ \frac{C_{a,L}}{N^L} \sum_{t \in \mathbb{Z}} \hat{f}_T(t) |a \circ \Phi^t|_L$$

for $T > 0$, $L \geq 3$, *where* $R_N(t)$ *is the remainder term in* (26) *and* $\hat{f}_T$ *is as in* (32). *Furthermore let* $\theta \in \mathbb{R}$ *then for every* $\delta > 0$ *there is a* $C > 0$ *such that*

$$\frac{1}{N} \sum_{|\theta_i^N - \theta_j^N - \theta|_{S^1} \leq \delta/T} \big|\langle \psi_i^N, \mathrm{Op}_N[a]\psi_j^N \rangle\big|^2 \leq C \sum_{t \in \mathbb{Z}} \mathrm{e}(\theta t) \hat{f}_T(t) C[a](t)$$

$$+ C\frac{1}{N} \sum_{t \in \mathbb{Z}} \mathrm{e}(\theta t) \hat{f}_T(t) \|R_N(t)\| \tag{34}$$

$$+ C\frac{C_{a,L}}{N^L} \sum_{t \in \mathbb{Z}} \mathrm{e}(\theta t) \hat{f}_T(t) |a \circ \Phi^t|_L.$$

This rather technical looking result will be the main tool in following investigations on the rate of quantum ergodicity. In (33) and (34) we have estimates on the variances of expectation values and transition amplitudes in terms of three quantities. The first term on the right hand side of (33) and (34) does not depend on N it is a purely classical quantity whose behaviour for large T is determined by the ergodic properties of the map Φ, as was discussed in Section 2. The remaining two terms on the right hand side of (33) and (34) depend on N and T, and they will determine how we can couple the two limits $N \to \infty$ and $T \to \infty$ in an optimal

way. As can already be seen by the appearance of $|a \circ \Phi^t|_L$ the behaviour of these terms will differ sharply between hyperbolic and non-hyperbolic maps.

The proof of Proposition 1 will rely on

Lemma 2. *Assume U_N satisfy the assumptions of Proposition* 1 *and f be as in* (31), *then for $a \in C^\infty(T^2)$ with $\bar{a} = 0$ and any $\theta \in \mathbb{R}$*

$$
\frac{1}{N} \sum_{i,j=1}^{N} F_T(\theta_i^N - \theta_j^N - \theta) \big| \langle \psi_i^N, \mathrm{Op}_N[a]\psi_j^N \rangle \big|^2
$$

$$
= \sum_{t \in \mathbb{Z}} \hat{f}_T(t) \, \mathrm{e}(\theta t) \frac{1}{N} \, \mathrm{Tr} \, \mathrm{Op}_N[a]^* U_N^{-t} \, \mathrm{Op}_N[a] U_N^t, \quad (35)
$$

with $\hat{f}_T$ and F_T given in (32), *and in particular*

$$
\frac{1}{N} \sum_{j=1}^{N} \big| \langle \psi_j^N, \mathrm{Op}_N[a]\psi_j^N \rangle \big|^2 \leq \sum_{t \in \mathbb{Z}} \hat{f}_T(t) \frac{1}{N} \, \mathrm{Tr} \, \mathrm{Op}_N[a]^* U_N^{-t} \, \mathrm{Op}_N[a] U_N^t. \quad (36)
$$

Proof. Let ψ_j^N, $\mathrm{e}(\theta_j^N)$, $j = 1, \ldots, N$, be the eigenfunctions and eigenvalues of U_N, then we have

$$
\mathrm{Tr} \, \mathrm{Op}_N[a]^* U_N^{-t} \, \mathrm{Op}_N[a] U_N^t = \sum_{i,j=1}^{N} \big| \langle \psi_j^N, \mathrm{Op}[a]\psi_i^N \rangle \big|^2 \mathrm{e}\big(t(\theta_j^N - \theta_i^N)\big). \quad (37)
$$

Now we have by the Poisson summation formula

$$
\sum_{t \in \mathbb{Z}} \hat{f}_T(t) \mathrm{e}\big(t(\theta_j^N - \theta_i^N - \theta)\big) = \sum_{n \in \mathbb{Z}} f\big(T(\theta_j^N - \theta_i^N - \theta - n)\big) = F_T(\theta_j^N - \theta_i^N - \theta) \quad (38)
$$

for any $T > 0$. So if we multiply (37) with $\hat{f}_T(t) \, \mathrm{e}(t\theta)$ and sum over t we obtain (35). If we use then furthermore that $F_T \geq 0$ and $F_T(0) \geq 1$ we obtain the estimate (36) by setting $\theta = 0$, restricting the sum in (35) to the terms $i = j$ and using that the remaining terms with $i \neq j$ are positive. $\qquad \square$

Proof of Proposition 1. By Lemma 2 we have to estimate $\frac{1}{N} \, \mathrm{Tr} \, \mathrm{Op}_N[a] U_N^{-t} \, \mathrm{Op}_N[a] U_N^t$. Using (26) we obtain

$$
\frac{1}{N} \, \mathrm{Tr} \, \mathrm{Op}_N[a] U_N^{-t} \, \mathrm{Op}_N[a] U_N^t = \frac{1}{N} \, \mathrm{Tr} \, \mathrm{Op}_N[a] \, \mathrm{Op}_N[a \circ \Phi^t] + \frac{1}{N^2} \, \mathrm{Tr} \, \mathrm{Op}_N[a] R_N(t), \quad (39)
$$

but $\mathrm{Op}_N[a]$ is bounded and $\mathrm{Tr} \, I_N = N$, so

$$
\frac{1}{N^2} |\mathrm{Tr} \, \mathrm{Op}_N[a] R_N(t)| \leq \frac{\|\mathrm{Op}_N[a]\|}{N} \|R_N(t)\|. \quad (40)
$$

With Lemma 1 we then get

$$
\frac{1}{N} \, \mathrm{Tr} \, \mathrm{Op}_N[a] \, \mathrm{Op}_N[a \circ \Phi^t] = C[a](t) + O\left(|a|_{L+3} \frac{|a \circ \Phi^t|_L}{N^L} \right) \quad (41)
$$

for $L \geq 3$, and so we have

$$\frac{1}{N} \operatorname{Tr} \operatorname{Op}_N[a] U_N^{-t} \operatorname{Op}_N[a] U_N^t$$
$$= C[a](t) + O_a\left(\frac{\|R_N(t)\|}{N}\right) + O\left(|a|_{L+3} \frac{|a \circ \Phi^t|_L}{N^L}\right). \quad (42)$$

Together with (36) this gives (33). In order to prove (34) we use (35), since F_T is positive and $F_T(0) \geq 1$ for any $\delta > 0$ there is a $C > 0$ such that

$$\frac{1}{N} \sum_{i,j=1}^N F_T(\theta_i^N - \theta_j^N - \theta) \left|\langle \psi_i^N, \operatorname{Op}_N[a] \psi_j^N \rangle\right|^2$$
$$\geq \frac{1}{C} \frac{1}{N} \sum_{|\theta_i^N - \theta_j^N - \theta|_{S^1} \leq \delta/T} \left|\langle \psi_i^N, \operatorname{Op}_N[a] \psi_j^N \rangle\right|^2 \quad (43)$$

and now using (35) and (42) gives (34). $\qquad \square$

Proof of Theorem 1. Let us first prove (28). Taking the limit $N \to \infty$ in (33) gives

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{j=1}^N \left|\langle \psi_j^N, \operatorname{Op}_N[a] \psi_j^N \rangle - \overline{a}\right|^2 \leq \sum_{t \in \mathbb{Z}} \hat{f}_T(t) C[a](t). \quad (44)$$

Now we take the limit $T \to \infty$ and by ergodicity, see (4), the right hand side tends to zero. The proof of (29) follows the same line, but using (34) with $\theta = 0$ instead. Taking the limit $N \to \infty$ gives

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{|\theta_i - \theta_j|_{S^1} \leq \delta/T} \left|\langle \psi_i^N, \operatorname{Op}_N[a] \psi_j^N \rangle\right|^2 \leq C \sum_{t \in \mathbb{Z}} \hat{f}_T(t) C[a](t) \quad (45)$$

and taking $T \to \infty$ implies that the right hand side tends to 0. But on the left hand side $T \to \infty$ is equivalent to $\delta \to 0$.

To prove the converse part, we use that from the proof of Proposition 1 we have

$$\lim_{N \to \infty} \frac{1}{N} \operatorname{Tr} \operatorname{Op}_N[a] U_N^{-t} \operatorname{Op}_N[a] U_N^t = C[a](t) \quad (46)$$

and so by (35)

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i,j=1}^N F_T(\theta_i^N - \theta_j^N - \theta) \left|\langle \psi_i^N, \operatorname{Op}_N[a] \psi_j^N \rangle\right|^2 = \sum_{t \in \mathbb{Z}} \hat{f}_T(t) \, \mathrm{e}(\theta t) C[a](t). \quad (47)$$

For $\theta = 0$ this reads

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i,j=1}^N F_T(\theta_i^N - \theta_j^N) \left|\langle \psi_i^N, \operatorname{Op}_N[a] \psi_j^N \rangle\right|^2 = \sum_{t \in \mathbb{Z}} \hat{f}_T(t) C[a](t). \quad (48)$$

Now if (29) holds, then the left hand side of (48) has to tend to 0 for $T \to \infty$, so the right hand side has to go to 0, too. This is valid for all $a \in C^\infty(T^2)$ with $\overline{a} = 0$

and since $C^\infty(T^2)$ is dense in $L^2(T^2)$ and $\hat{f} \geq 0$ this means the map Φ is ergodic by (4). $\qquad \square$

We will skip the proof of Theorem 2, as it follows the same line as the proof of Theorem 1, one uses (47) for all $\theta \in \mathbb{R}$ and weak mixing, (5), instead of ergodicity.

4. Hyperbolic maps

4.1. Cat maps

A cat map, or a hyperbolic toral automorphism, on the 2-torus $T^2 = \mathbb{R}^2/\mathbb{Z}^2$ is given by a matrix $A \in SL(2, \mathbb{Z})$ with $|\operatorname{Tr} A| > 2$ acting on T^2 as

$$\begin{pmatrix} p \\ q \end{pmatrix} \mapsto A \begin{pmatrix} p \\ q \end{pmatrix} \mod \mathbb{Z}^2 \, .$$

The condition $|\operatorname{Tr} A| > 2$ ensures that the map is hyperbolic. These maps are ergodic and rapidly mixing, i.e., for $a \in C^\infty(T^2)$ $C[a](t) = O_k(e^{-k\gamma|t|})$ for all $k > 0$, where $\gamma > 0$ is the Liapunov exponent of A.

These maps have been quantised on $\mathcal{H}_N(\kappa)$ with $\kappa = (0, 0)$ in [19] for even N and for odd N if the off-diagonal terms of A are even. The problems for odd N can be avoided by choosing $\mathcal{H}_N(\kappa)$ with $\kappa = (1/2, 1/2)$ for these N, [7,14]. So therefore we will fix our choice of κ in this section from now on, $\mathcal{H}_N = \mathcal{H}_N((0, 0))$ for even N and $\mathcal{H}_N = \mathcal{H}_N((1/2, 1/2))$ for odd N. The map A can then be quantised using the metaplectic representation and the resulting sequence of unitary operators U_N satisfies

$$U_N^* \operatorname{Op}_N[a] U_N = \operatorname{Op}_N[a \circ A] \, , \tag{49}$$

see [7,10,13,21]. So here the Egorov theorem holds exactly.

4.2. Perturbed cat maps

We will now consider a class of Anosov maps, the perturbed cat maps introduced in [2]. We will rely mainly on the recent study by Bouclet and De Bièvre in [8] of these maps. Let $A \in SL(2, \mathbb{Z})$ be a cat map and $g \in C^\infty(T^2)$ a real valued function, and consider the Hamiltonian flow $\phi^t : T^2 \to T^2$ generated by g. One can define then

$$\Phi_\varepsilon := \phi^\varepsilon \circ A : T^2 \to T^2 \tag{50}$$

which for small ε is a small perturbation of the Anosov map A, and hence by structural stability Φ_ε will be an Anosov map, too. The quantisation of Φ_ε is now defined as

$$U_{\varepsilon,N} := e^{-2\pi i N\varepsilon \operatorname{Op}_N[g]} U_N \, , \tag{51}$$

where U_N is the quantisation of A. In [8] it is then shown that there is a constant $\Gamma > 0$ such that for $t \in \mathbb{Z}$

$$\|U_{\varepsilon,N}^{-t} \operatorname{Op}_N[a] U_{\varepsilon,N}^t - \operatorname{Op}_N[a \circ \Phi_\varepsilon^t]\| \leq C_a \frac{1}{N} e^{\Gamma|t|} \, . \tag{52}$$

In fact the estimates in [8] are more precise, and Γ is estimated quite explicitly, but the estimate (52) is sufficient for our purpose.

Theorem 3. *Let $U_{\varepsilon,N}$ be the sequence of quantum maps (51) and ψ_j^N, $j = 1, \ldots, N$ an orthonormal basis of eigenfunctions of $U_{\varepsilon,N}$ for every $N \in \mathbb{N}$. Then for every $a \in C^\infty(T^2)$ there is a constant C_a such that*

$$\frac{1}{N} \sum_{j=1}^N \left| \langle \psi_j^N, \mathrm{Op}_N[a]\psi_j^N \rangle - \bar{a} \right|^2 \leq C_a \frac{1}{\ln N} \,. \tag{53}$$

The same result has been recently proved for the baker's map too, see [15]. For cat maps much stronger results are known for a special choice of a basis of eigenfunctions, due to their arithmetic nature, see [21, 23] and Section 4.3 below.

Proof. We will use (33). The Egorov estimate (52) gives $\|R_N(t)\| \leq C\mathrm{e}^{\Gamma|t|}$ and since the map is hyperbolic there is a constant $\Gamma' > 0$ such that $|a \circ \Phi_\varepsilon^t|_3 \leq C\mathrm{e}^{3\Gamma'|t|}$. Furthermore Φ_ε is exponentially mixing, i.e., $C[a](t) \leq C\mathrm{e}^{-\gamma|t|}$, and so

$$\sum_{t \in \mathbb{Z}} \hat{f}_T(t)C[a](t) = O(1/T) \,, \quad \sum_{t \in \mathbb{Z}} \hat{f}_T(t)\|R_N(t)\| = O(\mathrm{e}^{\Gamma T}) \,, \tag{54}$$

and

$$\sum_{t \in \mathbb{Z}} \hat{f}_T(t)|a \circ \Phi_\varepsilon^t|_3 = O(\mathrm{e}^{3\Gamma'T}) \,. \tag{55}$$

Therefore we obtain

$$\frac{1}{N} \sum_{j=1}^N \left| \langle \psi_j^N, \mathrm{Op}[a]\psi_j^N \rangle - \bar{a} \right|^2 \leq O(1/T) + O(\mathrm{e}^{\Gamma T}/N) + O(\mathrm{e}^{3\Gamma'T}/N^3) \,, \tag{56}$$

and the right hand side becomes minimal for the choice of $T = \delta \ln N$, with $\delta < \min\{\Gamma, \Gamma'\}$, which gives then (53). $\qquad\square$

By the same methods one can derive an upper bound on the transition amplitudes, strengthening the estimates (29) and Theorem 2.

Theorem 4. *Let $U_{\varepsilon,N}$ be the sequence of quantum maps (51) and ψ_j^N, $j = 1, \ldots, N$, an orthonormal basis of eigenfunctions of $U_{\varepsilon,N}$ for every $N \in \mathbb{N}$. Then for any $\delta > 0$ and every $a \in C^\infty(T^2)$ there is a constant C_a such that*

$$\frac{1}{N} \sum_{|\theta_i^N - \theta_j^N - \theta|_{S^1} \leq \delta/\ln N} \left| \langle \psi_i^N, \mathrm{Op}_N[a]\psi_j^N \rangle - \delta_{ij}\bar{a} \right|^2 \leq C_a \frac{1}{\ln N} \,, \tag{57}$$

for every $\theta \in \mathbb{R}$.

We skip the proof since it is identical to the one of Theorem 3, only using (34) instead of (33).

4.3. Cat maps: The logarithmic bound is sharp

In this section we want to show that the upper bound in Theorem 3 is sharp for unperturbed cat maps. The unperturbed cat map has very special properties, in particular for certain values of N the quantum map has very large multiplicities of the eigenvalues. This means that there are many bases of eigenfunctions and quantum ergodic properties can, and do, depend on the choice of a basis of eigenfunctions. In this section we will show that for a given observable $\mathrm{Op}_N[a]$ one can choose the basis of eigenfunctions such that the upper bound in Theorem 3 becomes sharp.

The special properties of cat maps are due to the fact that they are periodic, [19]. Let $A \in SL(2, \mathbb{Z})$ be a cat map and $U_N : \mathcal{H}_N \to \mathcal{H}_N$ its quantisation, the quantum period $P(N) \in \mathbb{N}$ is defined to be the smallest positive integer such that there is a $\phi_N \in \mathbb{R}/\mathbb{Z}$ with

$$U_N^{P(N)} = \mathrm{e}(\phi_N) I_N \tag{58}$$

with I_N the identity operator. The origin of these quantum periods are periods of the action of A on $\mathbb{Z}^2$ modulo N, for a given $A \in SL(2, \mathbb{Z})$ one defines $T(N)$ to be the smallest integer such that $A^{T(N)} \equiv I \mod N$, i.e., such that there exists a matrix A_N with integer entries such that

$$A^{T(N)} = I + N A_N . \tag{59}$$

Then one has either $P(N) = 2T(N)$ or $P(N) = T(N)$, depending on N and A, as was shown in [19]. $P(N)$ is on average of order N, [20], but there exist N where it is of order $\ln N$. More precisely, there exists a sequence N_k, $k \in \mathbb{N}$, such that

$$P(N_k) = \frac{2}{\gamma} \ln N_k + O(1) \tag{60}$$

where γ is the Liapunov exponent of A, [6, 22].

Now the relation (58) implies that the eigenvalues of U_N are $P(N)$'th roots of unity shifted by ϕ_N and since the dimension of $\mathcal{H}_N$ is N, a quantum period much shorter than N implies large multiplicities of the eigenvalues and therefore there are many different choices for an orthonormal basis of eigenfunctions of U_N.

For an observable $a \in C^\infty(T^2)$ we define the quantum average over one period as

$$\overline{\mathrm{Op}_N[a]}^P := \frac{1}{P(N)} \sum_{t=0}^{P(N)-1} U_N^{-t} \mathrm{Op}_N[a] U_N^t . \tag{61}$$

Then we have $U_N^{-1} \overline{\mathrm{Op}_N[a]}^P U_N = \overline{\mathrm{Op}_N[a]}^P$ and so the averaged operator commutes with the quantised cat map,

$$\overline{\mathrm{Op}_N[a]}^P U_N = U_N \overline{\mathrm{Op}_N[a]}^P . \tag{62}$$

If a is real valued, then $\mathrm{Op}_N[a]$ and $\overline{\mathrm{Op}_N[a]}^P$ are selfadjoint, and so by (62) we can choose an orthonormal basis of joint eigenfunctions of $\overline{\mathrm{Op}_N[a]}^P$ and U_N.

We will say that a function $a \in C^\infty(T^2)$ is a trigonometric polynomial of degree R if

$$a(x) = \sum_{n \in \mathbb{Z}^2,\, |n| \le R} \hat{a}(n)\, \mathrm{e}\left(\omega(n, x)\right). \tag{63}$$

We can now give the main result of this section.

Theorem 5. *Let $A \in SL(2, \mathbb{Z})$ be hyperbolic and denote its quantisation by U_N: $\mathcal{H}_N \to \mathcal{H}_N$. Let N_k, $k \in \mathbb{N}$, be a sequence with $P(N_k) = \frac{2}{\gamma} \ln N_k + O(1)$, then for every $R > 0$ there exists a $K_R > 0$ such that if $a(x)$ is a real valued trigonometric polynomial of degree R and $\psi_j^{N_k}$, $j = 1, \dots N_k$, an orthonormal basis of joint eigenfunctions of $\overline{\mathrm{Op}_{N_k}[a]}^P$ and U_{N_k}, then we have*

$$\frac{1}{N_k} \sum_{j=1}^{N_k} \left| \langle \psi_j^{N_k}, \mathrm{Op}_{N_k}[a] \psi_j^{N_k} \rangle - \bar{a} \right|^2 = \frac{1}{P(N_k)} V[a], \tag{64}$$

for $k \ge K_R$.

Recall that $V[a] = 0$ if, and only if, there is a $b \in C^\infty(T^2)$ such that $a = b - b \circ A$, see, e.g., [33, Proposition 4.14]. So there are many function with $V[a] > 0$, and therefore the bound from Theorem 3 is sharp. P. Kurlberg has as well obtained logarithmic lower bounds on the variance of expectation values for cat maps, [24].

The variance of expectation values depends strongly on the choice of the basis of eigenfunctions, in [23] it is shown that for a different choice than ours, a so called Hecke basis, the variance of expectation values is of order $1/N$. That is the behaviour which is expected for a generic Anosov system. But there is a point of view from which the logarithmic decay looks natural. In [12, 16] it is argued that the decay of the variance is of order $1/T_H(N)$, where $T_H(N)$ is the so called Heisenberg-time. Now the Heisenberg-time is inversely proportional to the mean level spacing, and so $T_H(N) \sim P(N)$, if we measure the level spacing without multiplicities, which means that (64) coincides with the expressions in [12, 16].

The basic idea we use here, namely the averaging over one quantum period, is the same that was used in [17] to prove the existence of scarred eigenstates of cat maps.

The proof of Theorem 5 proceeds in several steps. We first reduce it to the computation of a trace. Without loss of generality we will assume in the following that $\bar{a} = 0$.

Lemma 3. *Let ψ_j^N, $j = 1, \dots N$, be an orthonormal basis of joint eigenfunctions of $\overline{\mathrm{Op}_N[a]}^P$ and U_N, then*

$$\frac{1}{N} \sum_{j=1}^{N} \left| \langle \psi_j^N, \mathrm{Op}_N[a] \psi_j^N \rangle \right|^2 = \frac{1}{N} \mathrm{Tr}\left(\overline{\mathrm{Op}_N[a]}^P \mathrm{Op}_N[a]^* \right). \tag{65}$$

In general for an arbitrary basis the right hand side of (65) is an upper bound of the left hand side, so our choice of basis maximises the variance.

Proof. Since ψ_j^N are eigenfunctions of $\overline{\mathrm{Op}_N[a]}^P$ and orthonormal we have $\langle\psi_j^N,$ $\overline{\mathrm{Op}_N[a]}^P\psi_k^N\rangle = 0$ if $j \neq k$. For $j = k$ we use that ψ_j^N are eigenfunctions of U_N to get $\langle\psi_j^N, \overline{\mathrm{Op}_N[a]}^P\psi_j^N\rangle = \langle\psi_j^N, \mathrm{Op}_N[a]\psi_j^N\rangle$. Using these two relations gives

$$\mathrm{Tr}\,\overline{\mathrm{Op}_N[a]}^P\overline{\mathrm{Op}_N[a]^*}^P = \sum_{j,k=1}^N \left|\langle\psi_j^N, \overline{\mathrm{Op}_N[a]}^P\psi_k^N\rangle\right|^2$$
$$= \sum_{j=1}^N \left|\langle\psi_j^N, \overline{\mathrm{Op}_N[a]}^P\psi_j^N\rangle\right|^2 \tag{66}$$
$$= \sum_{j=1}^N \left|\langle\psi_j^N, \mathrm{Op}_N[a]\psi_j^N\rangle\right|^2.$$

Furthermore, by using that $\overline{\mathrm{Op}_N[a]}^P$ commutes with U_N, we obtain

$$\mathrm{Tr}\,\overline{\mathrm{Op}_N[a]}^P\overline{\mathrm{Op}_N[a]^*}^P = \frac{1}{P(N)}\sum_{t=0}^{P(N)-1}\mathrm{Tr}\,\overline{\mathrm{Op}_N[a]}^P U_N^{-t}\,\mathrm{Op}_N[a]^* U_N^t$$
$$= \mathrm{Tr}\,\overline{\mathrm{Op}_N[a]}^P\,\mathrm{Op}_N[a]^*. \qquad \square \tag{67}$$

Now we turn our attention to short periods (60). In the definition of the average $\overline{\mathrm{Op}_N[a]}^P$, equation (61), we averaged from $t = 0$ to $t = P(N) - 1$, but since U_N and $\overline{\mathrm{Op}_N[a]}^P$ commute we have $\overline{\mathrm{Op}_N[a]}^P = U_N^T\overline{\mathrm{Op}_N[a]}^P U_N^{-T}$ for any $T \in \mathbb{Z}$, which means we can shift the average to $t \in [-T, P(N) - 1 - T] \cap \mathbb{Z}$ for any T, and we will do this for the choice

$$T = [P(N)/2], \tag{68}$$

which is the integer part of $P(N)/2$.

Lemma 4. *Let N_k, $k \in \mathbb{N}$, be a sequence satisfying (60). Then for every $R > 0$ there exist a $K_R > 0$ such that for $k \geq K_R$ whenever there are two $m, n \in \mathbb{Z}^2$ with $|m|, |n| \leq R$ and*

$$A^t n = m \mod N_k, \tag{69}$$

for some $t \in [-[P(N_k)/2], P(N_k) - 1 - [P(N_k)/2]] \cap \mathbb{Z}$, then

$$A^t n = m. \tag{70}$$

Proof. The relation $A^t n = m \mod N_k$ means that there is a $\alpha_k \in \mathbb{Z}^2$ such that

$$A^t n - m = N_k\alpha_k \tag{71}$$

and we start by deriving a bound on the size of α_k. Since A is hyperbolic and $|t| \leq P(N_k)/2 = \frac{1}{\gamma}\ln N_k + O(1)$ we find

$$|A^t n - m| \leq R(1 + Ce^{\gamma|t|}) \leq R(1 + C'N_k), \tag{72}$$

and since by (71) $|\alpha_k| = |A^t n - m|/N_k$ there is a $\tilde{R} > 0$, independent of k, such that

$$|\alpha_k| \leq \tilde{R}. \tag{73}$$

So the α_k can only have finite size, and now we want to show that actually only $\alpha_k = 0$ can occur for k large enough, that will prove the Lemma. Let $W_{u/s}$ be the stable and unstable manifolds of the fixed-point at 0, these are straight lines through the origin with irrational slopes, see [17], and by definition they satisfy for $t \geq 0$

$$d(W_u, A^t n) \leq C\mathrm{e}^{-\gamma t}, \quad d(W_s, A^{-t} n) \leq C\mathrm{e}^{-\gamma t}, \tag{74}$$

where $d(W_{u/s}, A^{\pm t} n)$ denotes the Euclidean distance between the manifolds $W_{s/u}$ and the points $A^{\pm t} n$.

On the other hand, since the slopes of $W_{u/s}$ are irrational and the slopes of the lines $\mathbb{R} \ni \tau \mapsto m + \tau \alpha$, for $\alpha \neq 0$, $\alpha, m \in \mathbb{Z}^2$, are rational, the distance between $W_{u/s}$ and the points $m + N\alpha$ grows linearly in N. So for N large enough there is a constant C depending on $R, \tilde{R}$ such that for all $m, \alpha \in \mathbb{Z}^2$ with $|m| \leq R$, $|\alpha| \leq \tilde{R}$ and $\alpha \neq 0$ we have

$$d(W_{u/s}, m + N\alpha) \geq CN. \tag{75}$$

But since $A^t n = m + \alpha_k N_k$ the only way the two inequalities (74) and (75) can hold simultaneously is if

$$N_k \ll \mathrm{e}^{-\gamma |t|} \tag{76}$$

and so N_k has to be bounded. So the only case remaining allowing an infinite series of N_k is the case with $\alpha_k = 0$ and hence $A^t n = m$. $\qquad\square$

Lemma 5. *Let us write $n \sim_A m$ if n, m are in the same orbit of A on $\mathbb{Z}^2$. Assume N_k, $k \in \mathbb{N}$, is a sequence for which the quantum periods satisfy* (60), *then for every $R > 0$ there exist a $K_R > 0$ such that for $m, n \in \mathbb{Z}^2 \backslash \{0\}$ with $|m|, |n| \leq R$ and $k \geq K_R$, we have*

$$\frac{1}{N_k} \operatorname{Tr} T_{N_k}(n)^* \overline{T_{N_k}(m)}^P = \begin{cases} \frac{1}{P(N_k)} & \text{if} \quad n \sim_A m, \\ 0 & \text{otherwise}. \end{cases} \tag{77}$$

Proof. Since $T_N(n) = \operatorname{Op}_N[\mathrm{e}(\omega(n, x))]$ we have by Egorov's Theorem $U_N T_N(n) U_N^* = \operatorname{Op}_N[\mathrm{e}(\omega(n, A^{-1} x))]$, and because A is symplectic $\omega(n, A^{-1} x) = \omega(An, x)$, so $U_N T_N(n) U_N^* = T_N(An)$. Iterating this result gives

$$U_N^t T_N(n) U_N^{-t} = T_N(A^t n), \tag{78}$$

for all $t \in \mathbb{Z}$. Using this and the discussion before Lemma 4 about the average gives

$$
\begin{aligned}
\overline{T_N(m)}^P &= \frac{1}{P(N)} \sum_{t=-[P(N)/2]}^{P(N)-1-[P(N)/2]} U_N^t T_N(m) U_N^{-t} \\
&= \frac{1}{P(N)} \sum_{t=-[P(N)/2]}^{P(N)-1-[P(N)/2]} T_N(A^t m)
\end{aligned}
\tag{79}
$$

and so

$$
\begin{aligned}
&\frac{1}{N_k} \operatorname{Tr} T_{N_k}(n)^* \overline{T_{N_k}(m)}^P \\
&\qquad = \frac{1}{P(N_k)} \sum_{t=-[P(N_k)/2]}^{P(N_k)-1-[P(N_k)/2]} \frac{1}{N_k} \operatorname{Tr} T_{N_k}(n)^* T_{N_k}(A^t m) .
\end{aligned}
\tag{80}
$$

But $\operatorname{Tr} T_{N_k}(n)^* T_{N_k}(A^t m) \neq 0$ only if $n \equiv A^t m \mod N_k$, see (17), and by Lemma 4 this is for the t range we sum over and $k \geq K_R$ only the case if $n = A^t m$, i.e., $n \sim_A m$. In that case we have $\operatorname{Tr} T_{N_k}(n)^* T_{N_k}(A^t m) = N_k$ and so the result follows. $\qquad \square$

Proof of Theorem 5. Let a be a real valued trigonometric polynomial of order R, then by Lemma 3 we have to compute $\frac{1}{N_k} \operatorname{Tr} \operatorname{Op}_{N_k}[a]^* \overline{\operatorname{Op}_{N_k}[a]}^P$. By the definition of $\operatorname{Op}_N[a]$ we have

$$
\begin{aligned}
&\frac{1}{N_k} \operatorname{Tr} \operatorname{Op}_{N_k}[a]^* \overline{\operatorname{Op}_{N_k}[a]}^P \\
&\qquad = \sum_{n,m \in \mathbb{Z}^2 \setminus \{0\}} \hat{a}(n)^* \hat{a}(m) \frac{1}{N_k} \operatorname{Tr} T_{N_k}(n)^* \overline{T_{N_k}(m)}^P ,
\end{aligned}
\tag{81}
$$

where we used that $\hat{a}(0) = \bar{a} = 0$ by assumption, and then Lemma 5 gives

$$
\frac{1}{N_k} \operatorname{Tr} \operatorname{Op}_{N_k}[a]^* \overline{\operatorname{Op}_{N_k}[a]}^P = \frac{1}{P(N_k)} \sum_{n \in \mathbb{Z}^2} \sum_{m : m \sim_A n} \hat{a}(n)^* \hat{a}(m)
\tag{82}
$$

for $k \geq K_R$. But

$$
\sum_{m : m \sim_A n} \hat{a}(n)^* \hat{a}(m) = \sum_{t \in \mathbb{Z}} \hat{a}(n)^* \hat{a}(A^t n) ,
\tag{83}
$$

where the sum over t is finite since $\hat{a}(m) = 0$ for $|m| > R$, and so

$$
\frac{1}{N_k} \operatorname{Tr} \operatorname{Op}_{N_k}[a]^* \overline{\operatorname{Op}_{N_k}[a]}^P = \frac{1}{P(N_k)} \sum_{t \in \mathbb{Z}} \sum_{n \in \mathbb{Z}^2} \hat{a}(n)^* \hat{a}(A^t n) .
\tag{84}
$$

On the other hand side, using $a(x) = \sum_n \hat{a}(n) \operatorname{e}(\omega(n, x))$ one finds

$$
C[a](t) = \sum_{n \in \mathbb{Z}^2} \hat{a}(n)^* \hat{a}(A^t n)
\tag{85}
$$

and therefore

$$\frac{1}{N_k} \operatorname{Tr} \operatorname{Op}_{N_k}[a]^* \overline{\operatorname{Op}_{N_k}[a]}^P = \frac{1}{P(N_k)} \sum_{t \in \mathbb{Z}} C[a](t) = \frac{1}{P(N_k)} V[a] . \qquad (86)$$

$\square$

5. Parabolic maps

In this section we will study quantisations of parabolic maps. The interest in parabolic maps from our point of view is that they can be ergodic, but are not hyperbolic. Since hyperbolicity is the main obstacle in the control of the quantum time evolution for large times, we expect that for parabolic maps we get much stronger results. This is indeed the case, and as we will show below one again gets optimal results in some cases.

Our example will be the parabolic map studied by Marklof and Rudnick in [27], see as well [7, 11]. Let $\alpha \in \mathbb{R}$, then the map $\Psi_\alpha : T^2 \to T^2$ is defined by

$$\Psi_\alpha : \begin{pmatrix} p \\ q \end{pmatrix} \mapsto \begin{pmatrix} p + \alpha \\ q + 2p \end{pmatrix} \mod 1 . \qquad (87)$$

If α is irrational this map is uniquely ergodic but not weak mixing and not hyperbolic. This map is quantised in [27] and it is shown that their quantisation U_N satisfies the Egorov estimate

$$\|U_N^{-t} \operatorname{Op}_N[a] U_N^t - \operatorname{Op}_N[a \circ \Psi_\alpha^t]\| \le C_a \frac{|t|}{N} \qquad (88)$$

for $t \in \mathbb{Z}$.

In order to study the rate of quantum ergodicity, we need an estimate on the rate of classical ergodicity.

Lemma 6. *Let $a \in C^\infty(T^2)$ and $C[a](t)$ be the autocorrelation function of the map* (87) *and assume that α satisfies a Diophantine condition, i.e., there are $C, \gamma > 0$ such that $|k\alpha - l| \ge C/|k|^\gamma$ for all $k, l \in \mathbb{Z}\backslash\{0\}$. Then we have for $f \in \mathcal{S}(\mathbb{R})$*

$$\sum_{t \in \mathbb{Z}} \frac{1}{T} \hat{f}\left(\frac{t}{T}\right) C[a](t) = O\left(\frac{1}{T}\right) , \qquad (89)$$

where $\hat{f}$ denotes the Fourier-transform of f. Furthermore, if $a(p, q)$ depends only on p then

$$\sum_{t \in \mathbb{Z}} \frac{1}{T} \hat{f}\left(\frac{t}{T}\right) C[a](t) = O_M\left(\frac{1}{T^M}\right) , \quad for \ all \quad M \in \mathbb{N} . \qquad (90)$$

Proof. We have

$$\Psi_\alpha^t : \begin{pmatrix} p \\ q \end{pmatrix} \mapsto \begin{pmatrix} p + t\alpha \\ q + 2tp + \alpha t(t-1) \end{pmatrix} , \qquad (91)$$

and with $a(x) = \sum_{n \in \mathbb{Z}^2} \hat{a}(n)\mathrm{e}(nx)$ (notice that we use here the standard Fourier series, not the one twisted with ω used in the quantisation procedure) we get

$$C[a](t) = \sum_{n,m \in \mathbb{Z}^2 \setminus \{0\}} \hat{a}(n)\hat{a}(m) \int_{T^2} \mathrm{e}(nx)\mathrm{e}\big(m\Psi_\alpha^t(x)\big) \, \mathrm{d}x \,. \tag{92}$$

Then we find

$$\int_{T^2} \mathrm{e}(nx)\mathrm{e}\big(m\Psi_\alpha^t(x)\big) \, \mathrm{d}x$$
$$= \delta(-n_1, m_1 + 2tm_2)\delta(-n_2, m_2)\mathrm{e}\big(m_1\alpha t + m_2\alpha t(t-1)\big) \,, \tag{93}$$

where $\delta(m, n)$ denotes the Kronecker delta, and therefore

$$C[a](t)$$
$$= \sum_{(m_1, m_2) \in \mathbb{Z}^2 \setminus \{0\}} \hat{a}(-m_1 - 2tm_2, -m_2)\hat{a}(m_1, m_2)\mathrm{e}\big(m_1\alpha t + m_2\alpha t(t-1)\big) \,. \tag{94}$$

Now we split $C[a](t)$ into two parts, $C[a](t) = C^0[a](t) + C^1[a](t)$, such that $C^0[a](t)$ contains only the terms with $m_2 = 0$

$$C^0[a](t) = \sum_{m \in \mathbb{Z} \setminus \{0\}} \hat{a}(-m, 0)\hat{a}(m, 0)\mathrm{e}(m\alpha t) \,. \tag{95}$$

The second term satisfies

$$|C^1[a](t)| \le C_K(1 + |t|)^{-K} \tag{96}$$

for all $K \in \mathbb{N}$ since the Fourier-coefficients $\hat{a}(n)$ are quickly decreasing and therefore

$$\sum_{t \in \mathbb{Z}} \frac{1}{T}\hat{f}\left(\frac{t}{T}\right) C^1[a](t) = O\left(\frac{1}{T}\right) \,. \tag{97}$$

For the first term we find

$$\sum_{t \in \mathbb{Z}} \frac{1}{T}\hat{f}\left(\frac{t}{T}\right) C^0[a](t) = \sum_{m \in \mathbb{Z} \setminus \{0\}} |\hat{a}(m, 0)|^2 \sum_{t \in \mathbb{Z}} \frac{1}{T}\hat{f}\left(\frac{t}{T}\right) \mathrm{e}(m\alpha t) \tag{98}$$

and by the Poisson summation formula we obtain

$$\sum_{t \in \mathbb{Z}} \frac{1}{T}\hat{f}\left(\frac{t}{T}\right) \mathrm{e}(m\alpha t) = \sum_{n \in \mathbb{Z}} f\big(T(m\alpha - n)\big) = O_M(|m|^{\gamma M} T^{-M}) \tag{99}$$

since $f \in \mathcal{S}(\mathbb{R})$ and by the Diophantine condition on α. And since the Fourier-coefficients $\hat{a}(n)$ are quickly decreasing we find

$$\sum_{t \in \mathbb{Z}} \frac{1}{T}\hat{f}\left(\frac{t}{T}\right) C^0[a](t) = O_M(T^{-M}) \tag{100}$$

Combining the two estimates for $C^0[a](t)$ and $C^1[a](t)$ gives the lemma. $\square$

Combining the Egorov estimate and this lemma we then obtain from Proposition 1.

Theorem 6. *Let U_N be the quantisation of the map* (87) *due to* [27] *with a Diophantine α, and ψ_j^N, $j = 1, \ldots N$, a orthonormal basis of eigenfunctions. Then we have*

$$\frac{1}{N} \sum_{j=1}^{N} \left| \langle \psi_j^N, \mathrm{Op}[a] \psi_j^N \rangle - \bar{a} \right|^2 \leq C_a \frac{1}{N^{1/2}} \,, \tag{101}$$

and if $a(p, q)$ depends on p only then we have the stronger estimate

$$\frac{1}{N} \sum_{j=1}^{N} \left| \langle \psi_j^N, \mathrm{Op}[a] \psi_j^N \rangle - \bar{a} \right|^2 \leq C_{a,\varepsilon} \frac{1}{N^{1-\varepsilon}} \,, \tag{102}$$

for every $\varepsilon > 0$.

Proof. By the Egorov estimate (88) and since $|a \circ \Psi_\alpha^t|_k \leq C_k |t|^k$,

$$\sum_{t \in \mathbb{Z}} \hat{f}_T(t) \|R_N\| = O(T) \,, \quad \sum_{t \in \mathbb{Z}} \hat{f}_T(t) |a \circ \Psi_\alpha^t|_k = O(T^k) \,. \tag{103}$$

If we use then (89) we obtain by Proposition 1

$$\frac{1}{N} \sum_{j=1}^{N} \left| \langle \psi_j^N, \mathrm{Op}[a] \psi_j^N \rangle - \bar{a} \right|^2 = O\left(\frac{1}{T}\right) + O\left(\frac{T^3}{N^3}\right) + O\left(\frac{T}{N}\right) \,, \tag{104}$$

and so the choice $T = N^{1/2}$ gives (101). If we have instead the faster decay (90) we get

$$\frac{1}{N} \sum_{j=1}^{N} \left| \langle \psi_j^N, \mathrm{Op}[a] \psi_j^N \rangle - \bar{a} \right|^2 = O_M\left(\frac{1}{T^M}\right) + O\left(\frac{T^3}{N^3}\right) + O\left(\frac{T}{N}\right) \,, \tag{105}$$

for every $M \in \mathbb{N}$ and so by choosing $T = N^{\varepsilon'}$, with ε' small enough, and M large enough we obtain (102). $\square$

The results in [27] show that the estimate (101) is optimal, so we again obtain a sharp estimate. The analysis in [27] is much more detailed and provides sharp estimates for the rate of quantum ergodicity for individual eigenfunctions. But Theorem 6 might still be of some interest because the proof is of a more dynamical nature, and therefore may be easier to extend to other cases.

One further class of systems where one could apply the same methods is given by perturbed Kronecker maps, which were recently studied by Rosenzweig, [28]. Here the proof would be very similar to the one of (102), and we would get the same rate $O_\varepsilon(1/N^{1-\varepsilon})$. But in [28] a stronger bound on individual eigenfunctions is given, so our method does not give an optimal result.

The parabolic maps Ψ_α are not weak mixing, so by Theorem 2 the off-diagonal matrix elements $\langle \psi_i^N, \mathrm{Op}_N[a] \psi_j^N \rangle$ are not going to 0. In fact, using the same methods as in the proof of Lemma 6, one finds for Ψ_α^t that for $\theta = k\alpha$, $k \in \mathbb{Z} \backslash \{0\}$,

$$\sum_{t \in \mathbb{Z}} \frac{1}{T} \hat{f}\left(\frac{t}{T}\right) C[a](t) \mathrm{e}(\theta t) = |\hat{a}(k, 0)|^2 + O\left(\frac{1}{T}\right) \,. \tag{106}$$

From this result together with the techniques used in the proof of Theorem 6 one can derive

$$\lim_{N\to\infty} \frac{1}{N} \sum_{|\theta_i^N - \theta_j^N - \theta/N|_{S^1} \leq \delta/N^{1/2}} \left| \langle \psi_i^N, \mathrm{Op}_N[a] \psi_j^N \rangle \right|^2 = |\hat{a}(k,0)|^2, \qquad (107)$$

for any $\delta > 0$, where ψ_i^N, $\mathrm{e}(\theta_i^N)$, $i = 1, \ldots, N$ are the eigenvectors and eigenvalues of U_N, and $\theta = k\alpha$.

Acknowledgements

References

[1] R. Aurich and M. Taglieber, *On the rate of quantum ergodicity on hyperbolic surfaces and for billiards*, Physica D **118** (1998), 84–102.

[2] M. Basilio de Matos and A. M. Ozorio de Almeida, *Quantization of Anosov maps*, Ann. Physics **237** (1995), no. 1, 46–65.

[3] A. Bäcker, R. Schubert, and P. Stifter, *Rate of quantum ergodicity in euclidean billiards*, Phys. Rev E **57** (1998), 5425–5447, erratum: ibid. **58** (1998), 5192.

[4] M. V. Berry, *Regular and irregular semiclassical wavefunctions*, J. Phys. A **10** (1977), no. 12, 2083–2091.

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

[6] F. Bonechi and S. De Bièvre, *Exponential mixing and $|\ln \hbar|$ time scales in quantized hyperbolic maps on the torus*, Comm. Math. Phys. **211** (2000), no. 3, 659–686.

[7] A. Bouzouina and S. De Bièvre, *Equipartition of the eigenfunctions of quantized ergodic maps on the torus*, Comm. Math. Phys. **178** (1996), no. 1, 83–105.

[8] J. Bouclet and S. De Bièvre, *Long time propagation and control on scarring for perturbed quantized hyperbolic toral automorphisms*, Ann. Henri Poincaré **6** (2005), no. 5, 885–913.

[9] Y. Colin de Verdière, *Ergodicité et fonctions propres du laplacien*, Comm. Math. Phys. **102** (1985), no. 3, 497–502.

[10] S. De Bièvre, *Quantum chaos: a brief first visit*, Second Summer School in Analysis and Mathematical Physics (Cuernavaca, 2000), Contemp. Math., vol. 289, Amer. Math. Soc., Providence, RI, 2001, 161–218.

[11] S. De Bièvre and M. Degli Esposti, *Egorov theorems and equidistribution of eigenfunctions for the quantized sawtooth and baker maps*, Ann. Inst. H. Poincaré Phys. Théor. **69** (1998), no. 1, 1–30.

[12] T. O. de Carvalho, J.P. Keating, and J.M. Robbins, *Fluctuations in quantum expectation values for chaotic systems with broken time reversal symmetry*, J. Phys. A (1998), 5631–5640.

[13] M. Degli Esposti and S. Graffi, *Quantum maps*, The Mathematical Aspects of Quantum Maps, Lecture Notes in Physics, vol. 618, Springer, 2003, 49–90.

[14] M. Degli Esposti, S. Graffi, and S. Isola, *Classical limit of the quantized hyperbolic toral automorphisms*, Comm. Math. Phys. **167** (1995), no. 3, 471–507.

[15] M. Degli Esposti, S. Nonnenmacher, and B. Winn, *Quantum variance and ergodicity for the baker's map*, Comm. Math. Phys. **263** (2006), no. 2, 325–352.

[16] B. Eckhardt, S. Fishman, J. Keating, O. Agam, J. Main, and K. Müller, *Approach to ergodicity in quantum wave functions*, Phys. Rev. E **52** (1995), no. 6, 5893–5903.

[17] F. Faure, S. Nonnenmacher, and S. De Bièvre, *Scarred eigenstates for quantum cat maps of minimal periods*, Comm. Math. Phys. **239** (2003), no. 3, 449–492.

[18] M. Feingold and A. Peres, *Distribution of matrix elements of chaotic systems*, Phys. Rev. A (3) **34** (1986), no. 1, 591–595.

[19] J. H. Hannay and M. V. Berry, *Quantization of linear maps on a torus-Fresnel diffraction by a periodic grating*, Phys. D **1** (1980), no. 3, 267–290.

[20] J. P. Keating, *Asymptotic properties of the periodic orbits of the cat maps*, Nonlinearity **4** (1991), no. 2, 277–307.

[21] P. Kurlberg and Z. Rudnick, *Hecke theory and equidistribution for the quantization of linear maps of the torus*, Duke Math. J. **103** (2000), no. 1, 47–77.

[22] P. Kurlberg and Z. Rudnick, *On quantum ergodicity for linear maps of the torus*, Comm. Math. Phys. **222** (2001), no. 1, 201–227.

[23] P. Kurlberg and Z. Rudnick, *On the distribution of matrix elements for the quantum cat map*, Ann. of Math. **161** (2005), no. 1, 489–507.

[24] P Kurlberg, private communication, 2005.

[25] W. Luo and P. Sarnak, *Quantum variance for Hecke eigenforms*, Ann. Sci. École Norm. Sup. (4) **37** (2004), no. 5, 769–799.

[26] J. Marklof and S. O'Keefe, *Weyl's law and quantum ergodicity for maps with divided phase space*, Nonlinearity **18** (2005), no. 1, 277–304, With an appendix "Converse quantum ergodicity" by Steve Zelditch.

[27] J. Marklof and Z. Rudnick, *Quantum unique ergodicity for parabolic maps*, Geom. Funct. Anal. **10** (2000), no. 6, 1554–1578.

[28] L. Rosenzweig, *Quantum unique ergodicity for maps on the torus*, Ann. Henri Poincaré **7** (2006), no. 3, 447–469.

[29] Z. Rudnick and P. Sarnak, *The behaviour of eigenstates of arithmetic hyperbolic manifolds*, Comm. Math. Phys. **161** (1994), no. 1, 195–213.

[30] R. Schubert, *Upper bounds on the rate of quantum ergodicity*, Ann. Henri Poincaré **7** (2006), no. 6, 1085–1098.

[31] A. I. Šnirel'man, *Ergodic properties of eigenfunctions*, Uspehi Mat. Nauk **29** (1974), no. 6 (180), 181–182.

[32] T. Sunada, *Quantum ergodicity*, Progress in inverse spectral geometry, Trends Math., Birkhäuser, Basel, 1997, 175–196.

[33] M. Viana, *Stochastic dynamics of deterministic systems*, Brazillian Math. Colloqium, IMPA, 1997.

[34] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer Verlag, New York, 1982.

[35] M. Wilkinson, *A semiclassical sum rule for matrix elements of classically chaotic systems*, J. Phys. A **20** (1987), no. 9, 2415–2423.

[36] S. Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*, Duke Math. J. **55** (1987), no. 4, 919–941.

[37] S. Zelditch, *Quantum transition amplitudes for ergodic and for completely integrable systems*, J. Funct. Anal. **94** (1990), no. 2, 415–436.

[38] S. Zelditch, *On the rate of quantum ergodicity. I. Upper bounds*, Comm. Math. Phys. **160** (1994), no. 1, 81–92.

[39] S. Zelditch, *Quantum mixing*, J. Funct. Anal. **140** (1996), no. 1, 68–86.

Roman Schubert
School of Mathematics
University of Bristol
Bristol BS8 1TW
United Kingdom
e-mail: `roman.schubert@bristol.ac.uk`