QUENCHED CLT FOR RANDOM TORAL AUTOMORPHISM

ARVIND AYYER, CARLANGELO LIVERANI, AND MIKKO STENLUND

ABSTRACT. We establish a quenched Central Limit Theorem (CLT) for a smooth observable of random sequences of iterated linear hyperbolic maps on the torus. To this end we also obtain an annealed CLT for the same system. We show that, almost surely, the variance of the quenched system is the same as for the annealed system. Our technique is the study of the transfer operator on an anisotropic Banach space specifically tailored to use the cone condition satisfied by the maps.

1. INTRODUCTION

The issue of limit laws in dynamical systems has been widely explored in the last decades and it has a clear relevance for physical applications. A prime example of a physically relevant system is the study of the statistical behavior of a Lorenz gas with randomly distributed obstacles. The case of periodic obstacles is known to be ergodic. This follows from the recurrence [19], which in turns follows from the CLT, [10, 20], which has been proved in [18] (see also [8] for more refined results on these issues). See [15] for more details and for the treatment of some (locally) aperiodic cases. On the contrary the random case (albeit one may naïvely think of it as an easier case) stands as a challenge.

If one considers the simplest possibility (the random position of the obstacles is a small i.i.d. perturbation of a periodic configuration) then, by Poincaré section, one is readily reduced to considering a random sequence of hyperbolic symplectic maps. Yet, such a sequence of maps is not i.i.d. due to the presence of recollisions. Recollisions are notoriously a source of serious problems in the study of gases but, quite surprisingly, even disregarding the recollision problem (i.e. for the i.i.d. case), the problem is poorly understood.

In this paper we address the easiest setting in which such a situation occurs: an i.i.d. sequence of smooth uniformly hyperbolic symplectic maps. To make the presentation as clear as possible we will steer away from the full generality in which the present results can be obtained (although we will comment on it) and we will consider an i.i.d. sequence of linear two dimensional toral automorphisms. Exponential decay of correlations has been shown in this setting in [1].

For such a model we will show that the time-N average of any smooth zero mean observable has Gaussian fluctuations of order $\sqrt{N}$ for almost every sequence of maps. Moreover, we identify the variance of such Gaussian fluctuations.

Similar, but less complete, results are obtained in [2] where the Gaussian nature of the fluctuations is established for each sequence of maps but neither the amplitude nor the variance is investigated.

The paper is self-contained and organized as following: Section 2 contains the precise description of the model we will discuss and states the main results of the paper. Such results depend on the understanding of the ergodic properties of sequences of maps. These are investigated in Section 3 where the needed ergodic properties are related to the spectral properties of transfer operators viewed on appropriate Banach spaces in the spirit of the line of research started with [6]. Next, in Section 4, we use the above results to establish a CLT averaged over the environment for a class of systems larger than the ones at hand but necessary to handle the quenched case. The latter is dealt with in Section 5 using an approach inspired by works on random walks in random environments; see [11] and references therein.

Convention 1.1. *In this paper we will use C to designate a generic constant depending only on the choice of the maps $\{T_i\}$ below. We will use instead $C_{a,b,c,\dots}$ for constants depending also on the parameters $a, b, c, \dots$. Consequently, the actual numerical value of such constants may vary from one occurrence to the next. On the contrary we will use $C_1, C_2, \dots$, to designate constants whose value is fixed throughout the paper.*

2. The model and the results

Let us consider two[1] matrices $\{A_i\}_{i=0}^1 \in SL(2, \mathbb{N})$ and define the toral automorphisms $T_i x = A_i x \mod 1$. Let $\wp \in [0, 1]$ and set $p_0 = \wp$, $p_1 = 1 - \wp$. We can then introduce the Markov operator $Q_\wp : L^\infty(\mathbb{T}^2, \mathbb{R}) \to L^\infty(\mathbb{T}^2, \mathbb{R})$ defined by

$$Q_\wp g(x) = \sum_{i=0}^1 p_i g(T_i(x)).$$

Such an operator defines a Markov Process. To describe it we consider the space of trajectories $\Omega_* := (\mathbb{T}^2)^{\mathbb{N}}$ endowed with the product topology and, letting $(x_0, x_i, \dots)$ be a general element in Ω_*, we have the obvious dynamics $\boldsymbol{\tau} : \Omega_* \to \Omega_*$ defined by $\boldsymbol{\tau}(x_0, x_1, \dots) = (x_1, \dots)$. For each initial measure μ on $\mathbb{T}^2$, the above Markov process defines a Borel probability measure P_μ on Ω_*. Let $\mathbb{E}_{P_\mu}$ be the expectation with respect to such a measure. Then

$$\mathbb{E}_{P_\mu}(g(x_0)) = \int g(\xi)\mu(d\xi)$$

$$\mathbb{E}_{P_\mu}(g(x_{i+1}) \mid x_i) = Q_\wp g(x_i).$$

---

[1] In fact, the following would hold almost verbatim also for any larger collection of matrices.

The measure P_μ is supported on a very small set of trajectories: P_μ-almost surely $x_{i+1} \in \{T_0 x_i, T_1 x_i\}$. Thus, if we consider $\Sigma := \{0,1\}^{\mathbb{N} \setminus \{0\}}$, we have

$$P_\mu(\cup_{\omega \in \Sigma}\{(x_0, T_{\omega_1} x_0, T_{\omega_2} T_{\omega_1} x_0, \dots)\}) = 1.$$

In other words we can define the probability space $\Omega = \Sigma \times \mathbb{T}^2$ (again equipped with the product topology), the map $F : \Omega \to \Omega$ defined by

$$F(\omega, x) = (\tau \omega, T_{\omega_1} x),$$

where $(\tau \omega)_i = \omega_{i+1}$, and the measure $\mathbf{P}_{\wp,\mu} = \mathbb{P}_\wp \times \mu$, where $\mathbb{P}_\wp$ is the Bernoulli measure with probability $\wp$ of having zero. We will denote by $\mathbb{E}_\wp$ the expectation with respect to $\mathbb{P}_\wp$. Note that if μ is simultaneously T_0 and T_1 invariant, then $\mathbf{P}_{\wp,\mu}$ is invariant for the map F. Since the maps are symplectic, this happens for the normalized Lebesgue measure $\mathfrak{m}$. Let us set $\mathbf{P}_\wp := \mathbf{P}_{\wp,\mathfrak{m}}$ and call $\mathbb{E}_{\mathbf{P}_\wp}$ the corresponding expectation. Finally, we define the map $\Psi : \Omega \to \Omega_*$ by

$$\Psi(\omega, x) := (x, T_{\omega_1} x, T_{\omega_2} T_{\omega_1} x, \dots).$$

It is then easy to verify that $\boldsymbol{\tau}^k(\Psi(\omega, x)) = \Psi(F^k(\omega, x))$ for all $(\omega, x) \in \Omega$, $k \in \mathbb{N}$, and $\mathbb{E}_{P_\mu}(h) = \mathbb{E}_{\mathbf{P}_{\wp,\mu}}(h \circ \Psi)$ for each continuous function $h : \Omega_* \to \mathbb{R}$,[2] that is the two Dynamical Systems $(\Omega_*, \boldsymbol{\tau}, P_\mathfrak{m})$ and $(\Omega, F, \mathbf{P}_\wp)$ are isomorphic and so are the σ-algebras $\mathcal{F}_k = \sigma\text{-}\{x, \omega_1, \dots, \omega_k\}$ and $\tilde{\mathcal{F}}_k = \sigma\text{-}\{x_0, \dots, x_k\}$. We will use the two processes above interchangeably as far as the study of measure theoretical properties is concerned.

For each function $f \in \mathcal{C}^\infty(\mathbb{T}^2, \mathbb{R})$, $f \not\equiv 0$, such that $\mathfrak{m}(f) = 0$ we can then define the random variables $X_k(\omega, x) := f(\pi \circ F^k(\omega, x))$, where $\pi(\omega, x) := x$. We are interested in studying the $\mathbb{P}_\wp$-almost sure asymptotic behavior, as $N \to \infty$, of the random variables

$$S_N(\omega) := \sum_{k=0}^{N-1} X_k(\omega, \cdot).$$

The first relevant fact lies in the following Lemma.

Lemma 2.1. *The dynamical system* $(\Omega, F, \mathbf{P}_\wp)$ *is ergodic.*

Proof. By the above discussion the ergodicity of $(\Omega, F, \mathbf{P}_\wp)$ is equivalent to the ergodicity of the stationary Markov process $P_\mathfrak{m}$. It is well known that the ergodicity of such a process is equivalent to the fact that $Q_\wp g = g$ implies $g = \text{constant}$ for each bounded measurable g. In section 3 we will see (Corollary 3.3) that there exists $p, q > 0$ such that, for each $f \in \mathcal{C}^{p+2d+1}, g \in \mathcal{C}^q$, holds

(2.1) $$\lim_{n \to \infty} \mathfrak{m}(f Q_\wp^n g) = \mathfrak{m}(f)\mathfrak{m}(g).$$

Taking $f \in \mathcal{C}^{p+2d+1}$ and $g \in L^\infty$, we can choose a sequence $(g_j)_{j=1}^\infty \subset \mathcal{C}^q$ that converges to g in L^1 and we obtain (2.1) also for such functions. But this means

---

[2]Indeed, if $h(x_0, x_1, \dots) = g(x_n)$, then $\mathbb{E}_{P_\mu}(h) = \mu(Q_\wp^n g) = \mathbb{E}_{\mathbf{P}_{\wp,\mu}}(h \circ \Psi)$. On the other hand if we have already the equality for functions depending on n variables, we can write $h(x_0, \dots, x_n, x_{n+1}) = g_{x_0, \dots, x_n}(x_{n+1})$ and, by induction,

$$\mathbb{E}_{P_\mu}(h) = \mathbb{E}_{P_\mu}(\mathbb{E}_{P_\mu}(h \mid x_1, \dots, x_n)) = \mathbb{E}_{P_\mu}(Q_\wp g_{x_1, \dots, x_n}(x_n))$$

$$= \mathbb{E}_{\mathbf{P}_{\wp,\mu}}\left(\sum_i p_i h(\Psi(\omega, x)_0, \dots, \Psi(\omega, x)_n, T_i \Psi(\omega, x)_n)\right)$$

$$= \mathbb{E}_{\mathbf{P}_{\wp,\mu}}(\mathbb{E}_{\mathbf{P}_{\wp,\mu}}(h \circ \Psi \mid x, \omega_1, \dots, \omega_n)) = \mathbb{E}_{\mathbf{P}_{\wp,\mu}}(h \circ \Psi).$$

The assertion follows then by the density of the local functions among the continuous ones.

that $Q_\wp g = g$ implies $\mathfrak{m}(fg) = \mathfrak{m}(f)\mathfrak{m}(g)$ for each $f \in \mathcal{C}^{p+2d+1}$ which readily implies that g is constant. $\qquad\square$

Since $\mathfrak{m}(S_N(\omega)) = 0$, thanks to the previous Lemma, we can apply the Birkhoff Ergodic Theorem and obtain

$$\lim_{N \to \infty} \frac{1}{N} S_N = 0 \quad \mathbf{P}_\wp - \text{almost surely.}$$

The next step is to investigate the variable $N^{-\frac{1}{2}} S_N$ and prove that it satisfies a (quenched) CLT.

Theorem 1. *For each $f \in \mathcal{C}^\infty(\mathbb{T}^2, \mathbb{R})$,[3] $f \not\equiv 0$, and for each $\wp \in [0,1]$ there exists $\Sigma^2_\wp \in \mathbb{R}_+$ such that, for $\mathbb{P}_\wp$-almost all sequences ω holds[4]*

$$N^{-\frac{1}{2}} S_N \Rightarrow \mathcal{N}\left(0, \Sigma^2_\wp\right) \quad under \ \mathfrak{m}.$$

In fact, $\Sigma^2_\wp$ depends analytically on $\wp$. Moreover, if f is not a simultaneous continuous coboundary[5] for each admissible[6] map T_i, then $\Sigma^2_\wp > 0$.

Remark 2.2. *Imposing $\Sigma^2_\wp > 0$ clearly excludes fewer observables in the random case than the deterministic one. From Theorem 1 and the classical Livschitz Theorem [14] easily follows that $\Sigma^2_\wp = 0$ if and only if $\sum_{j=0}^{k-1} f(x_j) = 0$ whenever $(x_j)_{j=0}^{k-1}$ is a closed orbit for some sequence of admissible maps $T_{\omega_1}, \ldots, T_{\omega_k}$. Unfortunately, to use such a criterion it may be necessary to check a very large number of trajectories. Yet, if $\wp \in]0,1[$, then the situation may be much simpler. Indeed, if $\Sigma^2_\wp = 0$, then it must be $f = g - g \circ T_0 = g - g \circ T_1$ and hence $g \circ T_1 \circ T_0^{-1} = g$. Thus if $T_1 \circ T_0^{-1}$ is ergodic,[7] then g must be constant and hence $f \equiv 0$ contrary to assumptions. That is: if $\{T_0, T_1\}$ are admissible and $T_1 \circ T_0^{-1}$ is ergodic, then $\Sigma^2_\wp > 0$.*

Remark 2.3. *Note that one cannot possibly extend our results to include all sequences; for instance, there exist sequences containing alternating, "deterministic", stretches of either T_0's or T_1's. If these stretches are of rapidly and ever increasing length, then the variance fails to exist.*

Before proving such a strong result we will obtain its averaged (annealed) version.

Lemma 2.4. *For each $\wp \in [0,1]$ there exists $\Sigma^2_\wp \in \mathbb{R}_+$ such that it holds*

$$N^{-\frac{1}{2}} S_N \Rightarrow \mathcal{N}\left(0, \Sigma^2_\wp\right) \quad under \ \mathbf{P}_\wp.$$

---

[3]In the proof we use $f \in \mathcal{C}^r(\mathbb{T}^2, \mathbb{R})$ for r large enough. Yet, since our bounds for r are far from optimal (nor do we strive to optimize them) we see no point in giving an explicit bound for r.

[4]By $\mathcal{N}\left(0, \Sigma^2\right)$ we mean the centered Gaussian random variable with variance Σ^2. The symbol $\Rightarrow$ stands for convergence in distribution. As usual, $\mathcal{N}(0,0)$ stands for the measure concentrated at zero.

[5]By a *simultaneous continuos coboundary* for a set of maps $\{T_i\}$ we mean that there exists a continuous function g such that $f = g - g \circ T_i$, for each map $\{T_i\}$.

[6]The map T_i is admissible if it appears with nonzero probability with respect to $\mathbb{P}_\wp$. In our case the admissible maps are $\{T_0, T_1\}$ unless $\wp \in \{0,1\}$.

[7]Note that this may easily fail even if $T_0 \neq T_1$. Indeed, consider the case

$$A_0 = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \qquad A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

Then $T_1(T_0^{-1}(x_1, x_2)) = (x_1, x_2 - x_1)$. In fact, all the functions $g(x_1, x_2) = \tilde{g}(x_1)$ are invariant. The identity $f = g - g \circ T_0 = g - g \circ T_1$ yields $\sum_{k=0}^{N-1} f \circ \pi \circ F^k(\omega, x) = g(x) - g \circ T_i \circ T_{\omega_N} \circ \cdots \circ T_{\omega_1}(x)$ for $i \in \{0,1\}$ and hence $\Sigma^2_\wp = 0$ if g is in L^2.

In turn such a result is based on a fine understanding of the dynamical properties of certain transfer operators associated to the process.

Remark 2.5. *Note that one could obtain similar results for any finite collection of smooth symplectic hyperbolic maps in any dimension or piecewise smooth maps in dimension two. This can be achieved at the price of using in the following section the functional setting of* [13, 5] *or* [7] *for the piecewise smooth case.*

Our first task will be to obtain some information on the spectral properties of such operators. To do so in a useful way it is necessary to introduce appropriate functional spaces. Instead of appealing to the general theory developed in [6, 4, 12, 5, 13] we will take advantage of the simplicity of the present setting and introduce explicitly a particularly simple version of such a theory. We will then see how it can be used to address the ergodic theoretical questions we are interested in.

3. Spectral properties of the Transfer operators

For further use (see section 5) we need to study more general automorphisms than the one introduced in the previous section, namely $\bigoplus_1^d T_i \ : \ \mathbb{T}^{2d} \to \mathbb{T}^{2d}$ defined by $\bigoplus_1^d T_i x = \bigoplus_1^d A_i \mod 1$ where $\bigoplus_1^d A_i$ is the d-fold direct sums of the matrices $A_0, A_1 \in SL(2, \mathbb{Z})$, that is

$$\bigoplus_1^d A_i = \begin{pmatrix} A_i & & \\ & \ddots & \\ & & A_i \end{pmatrix} \in SL(2d, \mathbb{Z}),$$

with $d \in \{1, 2\}$. These matrices are symplectic with respect to the symplectic form

$$J_{2d} = \begin{pmatrix} J_2 & & \\ & \ddots & \\ & & J_2 \end{pmatrix},$$

where $J_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$ is the standard symplectic form in two dimensions.[8]

Let us introduce the transfer operators $\mathcal{L}_{T_i}^{(d)} \varphi := \varphi \circ (\bigoplus_1^d T_i)^{-1}$ induced by the above automorphisms. We will consider the operator obtained from the latter by averaging over the Bernoulli measure, namely

$$\mathcal{L}_\wp^{(d)} := \wp \, \mathcal{L}_{T_0}^{(d)} + (1 - \wp) \mathcal{L}_{T_1}^{(d)}.$$

We also need to study perturbed operators of the form

$$\mathcal{L}_{g,\wp}^{(d)} \varphi := \mathcal{L}_\wp^{(d)} (e^g \varphi).$$

In order to avoid unnecessary proliferation of indices, we set

(3.1) $$\mathcal{L}_{T_i} := \mathcal{L}_{T_i}^{(1)}, \quad \mathcal{L}_\wp := \mathcal{L}_\wp^{(1)}, \quad \text{and} \quad \mathcal{L}_{g,\wp} := \mathcal{L}_{g,\wp}^{(1)}.$$

Finally, notice that the transfer operator $\mathcal{L}_\wp$ and the Markov operator $Q_\wp$ are dual:

(3.2) $$\mathfrak{m}(f Q_\wp g) = \mathfrak{m}(\mathcal{L}_\wp f \cdot g).$$

---

[8]In fact all the following can be easily generalized to any set of symplectic toral automorphisms which preserves the standard sector in any dimension, see [16, 17] for more details on the necessary machinery.

To study such operators it is necessary to introduce appropriate Banach spaces (see [13, 5, 12]). Here, given the simplicity of the situation, we can quickly introduce and use spaces inspired by [5, 12] whereby making the presentation self-consistent.[9]

Given $v := (v_1, \ldots, v_{2d}) \in \mathbb{R}^{2d}$, let us denote $\check{v} := (v_1, v_3, \ldots, v_{2d-1}) \in \mathbb{R}^d$ and $\hat{v} := (v_2, v_4, \ldots, v_{2d})$. Then the standard inner product in $\mathbb{R}^d$ of the latter two reads $\langle \check{v}, \hat{v} \rangle = \sum_{k=0}^{d-1} v_{2k+1} v_{2(k+1)}$. Consider now the cones $\mathcal{C}_- := \{v \in \mathbb{R}^{2d} : \langle \check{v}, \hat{v} \rangle \leq 0\}$, $\mathcal{C}_+ := \{v \in \mathbb{R}^{2d} : \langle \check{v}, \hat{v} \rangle \geq 0\}$. Then there exists $1 < \lambda \leq \Lambda$ and C_0, depending only on $\{A_0, A_1\}$, such that, for all $v \in \mathcal{C}_-$, $(i_1, \ldots, i_n) \in \{0, 1\}^n$, and $n \in \mathbb{N}$,

$$(3.3) \quad C_0^{-1} \lambda^n \leq \frac{\|(\bigoplus_1^d A_{i_1} \cdots \bigoplus_1^d A_{i_n})^{-1} v\|}{\|v\|}, \frac{\|(\bigoplus_1^d A_{i_1}^T \cdots \bigoplus_1^d A_{i_n}^T)^{-1} v\|}{\|v\|} \leq C_0 \Lambda^n;$$

and, for all $v \in \mathcal{C}_+$, $(i_1, \ldots, i_n) \in \{0, 1\}^n$, and $n \in \mathbb{N}$,

$$(3.4) \quad C_0^{-1} \lambda^n \leq \frac{\|(\bigoplus_1^d A_{i_1} \cdots \bigoplus_1^d A_{i_n}) v\|}{\|v\|}, \frac{\|(\bigoplus_1^d A_{i_1}^T \cdots \bigoplus_1^d A_{i_n}^T) v\|}{\|v\|} \leq C_0 \Lambda^n.$$

Moreover one can compute that there exists $\beta > 1$ such that $(\bigoplus_1^d A_i^{-1}) \mathcal{C}_- \subset \mathcal{C}_\beta$, where $\mathcal{C}_\beta := \{v \in \mathbb{R}^{2d} : \beta^{-1} \|\check{v}\|^2 \leq -\langle \check{v}, \hat{v} \rangle \leq \beta \|\check{v}\|^2\} \subset \mathrm{int}\ \mathcal{C}_-$.

Now, we proceed to define the norm for the Banach space we want to consider. Notice that the natural objects in these cones are not vectors but Lagrangian subspaces. Recall that, given a symplectic form J, a Lagrangian subspace $E \subset \mathbb{R}^{2d}$ is a d-dimensional subspace such that $\langle v, Jw \rangle = 0$ for all $v, w \in E$.

For our choice of symplectic form, every Lagrangian subspace can also be written as the set $E = \{v : \hat{v} = -U\check{v}\}$ for a specific symmetric $d \times d$ matrix U. Our convention is to write a minus sign in front of the U here, because then $E \subset \mathcal{C}_\beta$ if and only if $\beta^{-1} \mathbb{1} \leq U \leq \beta \mathbb{1}$.

Let us denote the set of Lagrangian subspaces as $\mathbb{L}$. For a Lagrangian subspace E and a vector k, we set $\langle E, k \rangle := \sup_{\substack{v \in E \\ \|v\|=1}} \langle v, k \rangle$. Then, for each $p, q \in \mathbb{R}_+$ and $f \in \mathcal{C}^\infty(\mathbb{R}^{2d}, \mathbb{R})$ we define the norm

$$\|f\|_{p,q} := \sup_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_-}} \sum_{k \in \mathbb{Z}^{2d} \setminus \{0\}} |f_k| \frac{|k|^p}{1 + |\langle E, k \rangle|^{p+q}} + |f_0|,$$

where f_k are the Fourier coefficients of f.

Notice that the nth power of $\mathcal{L}_\wp^{(d)}$ can be expanded

$$[\mathcal{L}_\wp^{(d)}]^n = \sum_{j=1}^n \sum_{i_j=0}^1 \wp^{\delta_{i_1,0} + \cdots + \delta_{i_n,0}} (1-\wp)^{\delta_{i_1,1} + \cdots + \delta_{i_n,1}} \mathcal{L}_{T_{i_n}}^{(d)} \cdots \mathcal{L}_{T_{i_1}}^{(d)}.$$

Because the Bernoulli weights above sum to unity,

$$\|[\mathcal{L}_\wp^{(d)}]^n f\|_{p,q} \leq \sup_{(i_1, \ldots, i_n) \in \{0,1\}^n} \|\mathcal{L}_{T_{i_n}}^{(d)} \cdots \mathcal{L}_{T_{i_1}}^{(d)} f\|_{p,q}.$$

---

[9]Actually our choice is more flexible than the one in [12], in the spirit of [5], and would allow to treat $\mathcal{C}^k$ maps, although it is not the goal here.

A straightforward computation shows that $(\mathcal{L}_{T_i}^{(d)} f)_k = f_{(\bigoplus_1^d A_i^T) k}$ for $k \in \bigoplus_1^d \mathbb{Z}^2 \cong \mathbb{Z}^{2d}$. Hence, for each $n \in \mathbb{N}$ and $(i_1, \ldots, i_n) \in \{0, 1\}^n$,

$$\|\mathcal{L}_{T_{i_n}}^{(d)} \cdots \mathcal{L}_{T_{i_1}}^{(d)} f\|_{p,q} = \sup_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_-}} \sum_{k \in \mathbb{Z}^{2d} \setminus \{0\}} |f_k| \frac{|(\bigoplus_1^d A_{i_1}^T \cdots \bigoplus_1^d A_{i_n}^T)^{-1} k|^p}{1 + |\langle E, (\bigoplus_1^d A_{i_1}^T \cdots \bigoplus_1^d A_{i_n}^T)^{-1} k \rangle|^{p+q}} + |f_0|.$$

We begin estimating the summand. Before that, we simplify notation and rename the matrix product. To avoid the problem of too many indices, we denote it only with the subscript n and assume the dimension and the sequence to be implicit:

$$\mathfrak{A}_n := \bigoplus_1^d A_{i_1}^T \cdots \bigoplus_1^d A_{i_n}^T.$$

Using (3.3) and the fact $E \subset \mathcal{C}_-$, we have

$$(3.5) \qquad C_0 \Lambda^n \langle E_n, k \rangle \geq \langle E, \mathfrak{A}_n^{-1} k \rangle \geq C_0^{-1} \lambda^n \langle E_n, k \rangle \qquad \text{where} \qquad E_n := (\mathfrak{A}_n^{-1})^T E$$

and consequently,

$$|f_k| \frac{|\mathfrak{A}_n^{-1} k|^p}{1 + |\langle E, \mathfrak{A}_n^{-1} k \rangle|^{p+q}} \leq C_0^{p+q} |f_k| \frac{|\mathfrak{A}_n^{-1} k|^p}{1 + \lambda^{n(p+q)} |\langle E_n, k \rangle|^{p+q}}.$$

Now there are two possible cases depending on where $\mathfrak{A}_n^{-1} k$ lies. If $\mathfrak{A}_{n-1}^{-1} k \notin \mathcal{C}_-$, then $|\mathfrak{A}_n^{-1} k| \leq C_0 \Lambda \lambda^{-n+1} |k|$ and we have

$$\frac{|\mathfrak{A}_n^{-1} k|^p}{1 + \lambda^{n(p+q)} |\langle E_n, k \rangle|^{p+q}} \leq C_0^p (\Lambda \lambda^{-1})^p \lambda^{-np} \frac{|k|^p}{1 + |\langle E_n, k \rangle|^{p+q}}.$$

While if $\mathfrak{A}_{n-1}^{-1} k \in \mathcal{C}_-$, then

$$(3.6) \qquad \langle E_n, k \rangle \geq \frac{\lambda \Lambda^{-2n+1} |k|}{2 C_0^3 \beta \sqrt{1 + \beta^2}} =: \frac{|k|}{B_{\beta,n}}.$$

Indeed, setting $k_n := \mathfrak{A}_{n-1}^{-1} k$,

$$\langle (\bigoplus_1^d A_{i_n}^{-1}) E, k_n \rangle \geq \inf_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_\beta}} \langle E, k_n \rangle = \inf_{\beta^{-1} \mathbb{1} \leq U \leq \beta \mathbb{1}} \sup_{\check{v} \in \mathbb{R}^d} \frac{\langle \check{v}, \check{k}_n \rangle - \langle \check{v}, U \hat{k}_n \rangle}{\sqrt{|\check{v}|^2 + |U\check{v}|^2}}$$

$$\geq \inf_{\beta^{-1} \mathbb{1} \leq U \leq \beta \mathbb{1}} \frac{-\langle \check{k}_n, \hat{k}_n \rangle + \langle \hat{k}_n, U \hat{k}_n \rangle}{\sqrt{|\hat{k}_n|^2 + |U \hat{k}_n|^2}}$$

$$\geq \frac{|\hat{k}_n|}{\beta \sqrt{1 + \beta^2}},$$

where we have chosen $\check{v} = -\hat{k}_n$ in the second line. On the other hand the choice $\check{v} = U^{-1} \check{k}_n$ yields

$$\langle (\bigoplus_1^d A_{i_n}^{-1}) E, k_n \rangle \geq \inf_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_\beta}} \langle E, k_n \rangle \geq \frac{|\check{k}_n|}{\beta \sqrt{1 + \beta^2}},$$

The inequality (3.6) follows from the above estimates, (3.5) and $C_0 \Lambda^{n-1} |k_n| \geq |k|$.

Next, we consider two subcases. If $|k| \geq B_{\beta,n}$, then $|\langle E_n, k \rangle| \geq 1$. Hence,

$$\frac{|\mathfrak{A}_n^{-1} k|^p}{1 + \lambda^{n(p+q)} |\langle E_n, k \rangle|^{p+q}} \leq C_0^p \frac{\Lambda^{pn} \lambda^{-(p+q)n} |k|^p}{|\langle E_n, k \rangle|^{p+q}} \leq 2 C_0^p \frac{\Lambda^{pn} \lambda^{-(p+q)n} |k|^p}{1 + |\langle E_n, k \rangle|^{p+q}},$$

which is a good estimate provided q is large enough so that $\Lambda^p \lambda^{-(p+q)} < 1$. The remainder is a finite sum which can be estimated because if $E \subset \mathcal{C}_-$, then $E_n \subset (\mathfrak{A}_n^{-1})^T \mathcal{C}_- \subset \mathcal{C}_-$. Thus, we have

$$\sup_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_-}} \sum_{\substack{k \in \mathbb{Z}^{2d} \setminus \{0\} \\ |k| < B_{\beta,n}}} \frac{|f_k| \, |\mathfrak{A}_n^{-1} k|^p}{1 + \lambda^{n(p+q)} |\langle E_n, k \rangle|^{p+q}} \le C_0^p \Lambda^{np} \sup_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_-}} \sum_{\substack{k \in \mathbb{Z}^{2d} \setminus \{0\} \\ |k| < B_{\beta,n}}} \frac{|f_k| \, |k|^p}{1 + |\langle E, k \rangle|^{p+q}}.$$

Accordingly, setting $\tilde{\mu} := \max\{\lambda^{-p}, \Lambda^p \lambda^{-p-q}\}$ we can collect all the above inequalities as

$$\|[\mathcal{L}_{\wp}^{(d)}]^n f\|_{p,q} \le C_1 \|f\|_{p,q},$$

$$(3.7) \quad \|[\mathcal{L}_{\wp}^{(d)}]^n f\|_{p,q} \le 2C_0^{q+2p} \tilde{\mu}^n \|f\|_{p,q} + \sup_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_-}} \sum_{\substack{k \in \mathbb{Z}^{2d} \setminus \{0\} \\ |k| < B_{\beta,n}}} \frac{C_0^{q+2p} \Lambda^{np} |f_k| \, |k|^p}{1 + |\langle E, k \rangle|^{p+q}} + |f_0|$$

$$\le C_2 \tilde{\mu}^n \|f\|_{p,q} + B_n \|f\|_{p-1,q+1},$$

where $B_n = C_0^{q+2p} B_{\beta,n} \Lambda^{np}$. Next, for each $\mu \in (\tilde{\mu}, 1)$ choose n_0 such that $C_2 \tilde{\mu}^{n_0} \le \mu^{n_0}$ and, for each $n \in \mathbb{N}$, write $n = kn_0 + m$ with $m \in \{0, \dots, n_0 - 1\}$. One can thus iterate the second of the (3.7) and obtain

$$\|[\mathcal{L}_{\wp}^{(d)}]^n f\|_{p,q} \le \mu^{kn_0} \|\mathcal{L}^m f\|_{p,q} + B_{n_0} \sum_{j=0}^{k-1} \mu^{jn_0} \|\mathcal{L}^{(k-1-j)n_0+m} f\|_{p-1,q+1},$$

which finally yields

$$(3.8) \quad \begin{aligned} \|[\mathcal{L}_{\wp}^{(d)}]^n f\|_{p,q} &\le C_1 \|f\|_{p,q}, \\ \|[\mathcal{L}_{\wp}^{(d)}]^n f\|_{p,q} &\le C_3 \mu^n \|f\|_{p,q} + B \|f\|_{p-1,q+1}, \end{aligned}$$

with $B = C_2 B_{n_0} (1 - \mu^{n_0})^{-1}$. We can then consider the closure, $\mathcal{B}^{p,q}$, of $\mathcal{C}^{\infty}$ in the space of distributions with respect to the norms $\|\cdot\|_{p,q}$. It is easy to prove the following:

Lemma 3.1. *The operators $\mathcal{L}_{\wp}^{(d)}$ are well defined bounded operators on $\mathcal{B}^{p,q}$, provided $\Lambda^p < \lambda^{p+q}$. In addition, the unit ball of $\mathcal{B}^{p,q}$ is relatively compact in $\mathcal{B}^{p-1,q+1}$.*

Theorem 3.2. *If $\Lambda^p < \lambda^{p+q}$, the operator $\mathcal{L}_{\wp}^{(d)}$ acting on $\mathcal{B}^{p,q}$ has an essential spectral radius smaller than μ. The rest of the spectrum consists of finitely many eigenvalues of finite multiplicity, all in the unit disk. The only eigenvalue of modulus one is one and the constant function equal to one is the corresponding eigenfunction.*

Proof. Lemma 3.1, the Lasota-Yorke type inequalities (3.8) imply the result (see [3, 6] for more details). □

Before proceeding, let us mention that for any $r, n \in \mathbb{N}$ we endow the space $\mathcal{C}^r(\mathbb{T}^n, \mathbb{R})$ with the norm $\|g\|_{\mathcal{C}^r} := \sum_{s=0}^r \|g^{(s)}\|_{\infty}$.

Moreover, a simple computation shows that $\mathcal{C}^r \subset \mathcal{B}^{p,q}$, provided $r > p + 2d$.

Corollary 3.3. *The equation (2.1) holds true.*

Proof. First of all notice that, for all $f \in \mathcal{B}^{p,q}$ and $g \in \mathcal{C}^q$ holds

$$|\mathfrak{m}(fg)| \leq \sum_{l \in \mathbb{Z}^2} |f_l| \, |g_{-l}| \leq \|g\|_{\mathcal{C}^q} \left(|f_0| + \sum_{l \in \mathbb{Z}^2 \setminus \{0\}} |f_l| \, |l|^{-q} \right)$$

$$\leq \|g\|_{\mathcal{C}^q} \|f\|_{p,q} \max \left(1, \sup_{\substack{E \in \mathbb{L} \\ l \in \mathbb{Z}^2 \setminus \{0\}}} \frac{1 + |\langle E, l \rangle|^{p+q}}{|l|^{p+q}} \right) \leq 2 \|g\|_{\mathcal{C}^q} \|f\|_{p,q}.$$

If p, q satisfy the hypothesis of Theorem 3.2, $\mathcal{L}_\wp$ has a spectral gap $\delta_\wp > 0$. Thus,

$$|\mathfrak{m}(f Q_\wp^n g) - \mathfrak{m}(f)\mathfrak{m}(g)| = |\mathfrak{m}(\mathcal{L}_\wp^n f \cdot g) - \mathfrak{m}(f)\mathfrak{m}(g)| \leq C(1 - \delta_\wp)^n \|f\|_{p,q} \|g\|_{\mathcal{C}^q},$$

because decomposing $\mathcal{L}_\wp := \mathcal{Q} + \mathcal{R}$ with $\mathcal{Q}f := \mathfrak{m}(f)$ we have $\mathcal{Q}\mathcal{R} = \mathcal{R}\mathcal{Q} = 0$ and therefore $\mathcal{L}_\wp^n f = \mathcal{R}^n f + \mathcal{Q}f$, where $\|R^n\|_{\mathcal{L}(\mathcal{B}^{p,q})} \leq C(1 - \delta_\wp)^n$ by the spectral radius formula. $\qquad \square$

Here is the last fact we need to know about the above functional analytic setting.

Lemma 3.4. *For each function* $g \in \mathcal{C}^{2p+q+2d+1}(\mathbb{T}^{2d}, \mathbb{R})$ *the multiplication operator* M_g *defined by* $M_g f = gf$, *is bounded in* $\mathcal{B}^{p,q}$ *by* $C \|g\|_{\mathcal{C}^{2p+q+2d+1}}$.

Proof. We define the norm[10]

$$\|g\|_r := \sup_{k \in \mathbb{Z}^{2d}} |g_k| (1 + |k|^r).$$

Clearly $g \in \mathcal{C}^r(\mathbb{T}^{2d}, \mathbb{R})$ implies $\|g\|_r \leq \|g\|_{\mathcal{C}^r}$. The $\mathcal{B}^{p,q}$-norm of the product then reads

$$(3.9) \qquad \|fg\|_{p,q} = \sup_{\substack{E \in \mathbb{L} \\ E \subset \mathcal{C}_-}} \sum_{\substack{k,l \in \mathbb{Z}^{2d} \\ k \neq 0}} |f_l| |g_{k-l}| \frac{|k|^p}{1 + |\langle E, k \rangle|^{p+q}} + \sum_{l \in \mathbb{Z}^{2d}} |f_l| |g_{-l}|.$$

Let us analyze the second term first:

$$\sum_{l \in \mathbb{Z}^{2d}} |f_l| |g_{-l}| \leq \|g\|_r \left(|f_0| + \sum_{l \in \mathbb{Z}^{2d} \setminus \{0\}} |f_l| |l|^{-r} \right).$$

The desired bound follows, if $r \geq q$, from

$$\frac{1}{2} |l|^{-q} \leq \frac{|l|^p}{1 + |\langle E, l \rangle|^{p+q}}.$$

Now, look at the summand in the first term of (3.9), ignoring the $l = 0$ case that can be taken care of separately.

$$|f_l| |g_{k-l}| \frac{|k|^p}{1 + |\langle E, k \rangle|^{p+q}} \leq \|g\|_r \left[|f_l| \frac{|l|^p}{1 + |\langle E, l \rangle|^{p+q}} \right]$$

$$\times \frac{|k|^p}{1 + |\langle E, k \rangle|^{p+q}} \frac{1 + |\langle E, l \rangle|^{p+q}}{|l|^p} \frac{1}{1 + |k - l|^r}.$$

---

[10]We are aware that our choices of norms and the subsequent estimates, are not the optimal ones. We are simply trying to simplify the arguments as much as possible even at the expense of some, not really relevant, optimality.

Notice that $|\langle E, l\rangle| \leq |\langle E, k\rangle| + |\langle E, k - l\rangle|$. Thus, on the one hand

$$\sum_k \frac{|k|^p}{1 + |\langle E, k\rangle|^{p+q}} \frac{1 + |\langle E, k\rangle|^{p+q}}{|l|^p} \frac{1}{1 + |k - l|^r} \leq \sum_k \frac{|k|^p}{|l|^p(1 + |k - l|^r)} \leq C.$$

One the other hand

$$\sum_k \frac{|k|^p}{1 + |\langle E, k\rangle|^{p+q}} \frac{1 + |\langle E, k - l\rangle|^{p+q}}{|l|^p} \frac{1}{1 + |k - l|^r} \leq 3 \sum_k \frac{|k|^p}{|l|^p(1 + |k - l|^{r-p-q})} \leq C,$$

provided $r > 2p + q + 2d$. The general term of $(|\langle E, k\rangle| + |\langle E, k - l\rangle|)^{p+q}$ is bounded similarly, using $|\langle E, k\rangle|^n |\langle E, k - l\rangle|^{p+q-n} \leq (1 + |\langle E, k\rangle|^n)(1 + |\langle E, k - l\rangle|^{p+q-n})$ for each $n = 0, \ldots, p + q$. □

4. Averaged CLT

To establish Lemma 2.4 it suffices to compute

$$\lim_{N \to \infty} \mathbb{E}_{\mathbf{P}_\wp} \left(e^{i\frac{\lambda}{\sqrt{N}} \sum_{k=0}^{N-1} f_{\omega,k}} \right),$$

where $f_{\omega,k} := X_k(\omega, \cdot) := f \circ T_{\omega_k} \circ \cdots \circ T_{\omega_1}$ with $f_{\omega,0} := f$, and show that this limit is the characteristic function $e^{-\frac{1}{2}\lambda^2 \Sigma_\wp^2}$ of the centered normal distribution with some variance $\Sigma_\wp^2$. Recalling the transfer operators in (3.1),

$$(4.1) \qquad\qquad \mathbb{E}_{\mathbf{P}_\wp} \left(e^{i\frac{\lambda}{\sqrt{N}} \sum_{k=0}^{N-1} f_{\omega,k}} \right) = \mathfrak{m}(\mathcal{L}_{i\lambda N^{-\frac{1}{2}} f, \wp}^N 1).$$

Hence Theorem 3.2 and Lemma 3.4 show that $\mathcal{L}_{i\lambda N^{-\frac{1}{2}} f, \wp}$ is a bounded operator on $\mathcal{B}^{p,q}$ provided $f \in \mathcal{C}^{2p+q+5}$ and depends analytically on λ. To continue it is necessary to study the leading eigenvalue of such an operator.

Note that, in general, given any positive operator $\mathcal{L}$ on the spaces $\mathcal{B}^{p,q}$ with maximal simple eigenvalue one, with a spectral gap and $\mathfrak{m}(\mathcal{L}\varphi) = \mathfrak{m}(\varphi)$ for each smooth φ, for any smooth complex valued function g we can define the family of operators $\mathcal{L}_\nu \varphi := \mathcal{L}(e^{\nu g}\varphi)$ and, thanks to Lemma 3.4, the standard perturbation theory applies. Thus there exists ϕ_ν, μ_ν, with $\mu_0 = 1$, such that

$$\mathcal{L}_\nu \phi_\nu = \mu_\nu \phi_\nu, \quad \mathfrak{m}(\phi_\nu) = 1.$$

Differentiating this relation with respect to ν and integrating one readily obtains

$$(4.2) \qquad\qquad \mu_\nu' = \mathfrak{m}\left(g e^{\nu g}\phi_\nu + e^{\nu g}\phi_\nu'\right)$$

and, setting $\nu = 0$, $\phi_0' = (\mathbb{1} - \mathcal{L})^{-1}[\mathcal{L}(g\phi_0) - \phi_0\mathfrak{m}(g\phi_0)] = \sum_{n=0}^\infty \mathcal{L}^n[\mathcal{L}(g\phi_0) - \phi_0\mathfrak{m}(g\phi_0)]$.[11] Finally, differentiating again yields

$$(4.3) \qquad\qquad \mu_0'' = \mathfrak{m}(g^2\phi_0) + 2\sum_{n=0}^\infty \mathfrak{m}(g\mathcal{L}^n\left[\mathcal{L}(g\phi_0) - \phi_0\mathfrak{m}(g\phi_0)\right]).$$

Thus, by standard perturbation theory and in view of Theorem 3.2 we can write $\mathcal{L}_\nu = \mu_\nu \mathcal{Q}_\nu + \mathcal{R}_\nu$ where $\mathcal{Q}_\nu^2 = \mathcal{Q}_\nu$, $\mathcal{R}_\nu \mathcal{Q}_\nu = \mathcal{Q}_\nu \mathcal{R}_\nu = 0$, the spectral radius of $\mathcal{R}_\nu$ is smaller than $\rho < 1$ for all $|\nu| \leq \nu_0$ for some $\nu_0 > 0$, and $|\mu_\nu - 1 - \mu_0'\nu - \frac{1}{2}\mu_0''\nu^2| \leq C|\nu|^3$ for some fixed constant $C > 0$ and $|\nu| \leq \nu_0$. In addition, $\mathcal{Q}_\nu$ is a rank one operator of the form $\phi_\nu \otimes m_\nu$ where m_ν belongs to the dual of the space, $m_0 = \mathfrak{m}$, and $|m_\nu(1) - 1| \leq C|\nu|$.

---

[11]The latter is well defined since $(\mathbb{1} - \mathcal{L})^{-1}$ is applied on a function from which the eigendirection of $\mathcal{L}$ corresponding to eigenvalue 1 has been projected out, and because of the spectral gap.

If we apply the above to the operator $\mathcal{L}_{i\lambda N^{-\frac{1}{2}}f,\wp}$, $\nu = i\lambda N^{-\frac{1}{2}}$ (hence $g = f$, and $\phi_0 \equiv 1$), then remembering equation (4.1) it follows that

(4.4)
$$\mathbb{E}_{\mathbf{P}_\wp}\left(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}}\right) = \mathfrak{m}\left(\mu_\nu^N \mathcal{Q}_\nu 1 + \mathcal{R}_\nu^N 1\right) = e^{N\ln(\mu_\nu)}(1 + \mathcal{O}(|\lambda|N^{-\frac{1}{2}})) + \mathcal{O}(\rho^N)$$
$$= e^{-\frac{1}{2}\mu_0''\lambda^2 + N\mathcal{O}\left((|\lambda|/N^{-\frac{1}{2}})^3\right)} + \mathcal{O}(|\lambda|N^{-\frac{1}{2}} + \rho^N),$$

because $\mathfrak{m}(\mathcal{Q}_\nu 1) = \mathfrak{m}(\phi_\nu m_\nu(1)) = m_\nu(1)$ and $|\mathfrak{m}(\mathcal{R}_\nu^N 1)| \leq \|\mathcal{R}_\nu^N\|_{\mathcal{L}(\mathcal{B}^{p,q})} \leq C\rho^N$. Hence, if $|\lambda|/\sqrt{N}$ is sufficiently small, we have

(4.5)
$$\left|\mathbb{E}_{\mathbf{P}_\wp}\left(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}}\right) - e^{-\frac{1}{2}\mu_0''\lambda^2}\right| \leq C\frac{1 + |\lambda|^3}{\sqrt{N}}.$$

Lemma 4.1. *The quantity* $\Sigma_\wp^2 \in \mathbb{R}$ *defined by*

$$\Sigma_\wp^2 := \lim_{N\to\infty} \frac{1}{N}\mathbb{E}_{\mathbf{P}_\wp}\left(\left[\sum_{k=0}^{N-1} X_k\right]^2\right)$$

is always nonnegative and given by

(4.6)
$$\mathfrak{m}(f^2) + 2\sum_{n=1}^{\infty} \mathfrak{m}(f\mathcal{L}_\wp^n f) = \mu_0''.$$

The map $[0,1] \to \mathbb{R}_+ : \wp \mapsto \Sigma_\wp^2$ *is analytic. Moreover, if f is not a $\mathcal{C}^0$ simultaneous coboundary for the admissible automorphisms T_i (see Remark 2.2), then $\Sigma_\wp^2 > 0$.*

Proof. A direct computation yields

$$\mathbb{E}_{\mathbf{P}_\wp}\left(\left[\sum_{k=0}^{N-1} X_k\right]^2\right) = \sum_{k=0}^{N-1}\mathbb{E}_{\mathbf{P}_\wp}(X_k^2) + 2\sum_{0\leq j<k\leq N-1}\mathbb{E}_{\mathbf{P}_\wp}(X_k X_j)$$
$$= N\mathfrak{m}(f^2) + 2\sum_{0\leq j<k\leq N-1}\mathfrak{m}(fQ_\wp^{k-j}f)$$
$$= N\left[\mathfrak{m}(f^2) + 2\sum_{n=1}^{N-1}\mathfrak{m}(f\mathcal{L}_\wp^n f)\right] - 2\sum_{n=1}^{N-1} n\,\mathfrak{m}(f\mathcal{L}_\wp^n f)$$

Using Corollary 3.3, the last sum converges exponentially fast in n. Hence $\Sigma_\wp^2$ exists and is nonnegative simply because it is the limit of a nonnegative quantity. To address this last issue, suppose $\Sigma_\wp^2 = 0$. Then

$$\left|\mathbb{E}_{\mathbf{P}_\wp}\left(\left[\sum_{k=0}^{N-1} X_k\right]^2\right)\right| \leq 2N\sum_{n=N}^{\infty}|\mathfrak{m}(f\mathcal{L}_\wp^n f)| + 2\sum_{n=1}^{N-1} n\,|\mathfrak{m}(f\mathcal{L}_\wp^n f)| \leq C$$

uniformly in N. This means that the random variables $Z_N := \sum_{k=0}^{N-1} X_k$ are uniformly bounded in L^2. By the Banach–Alaoglu Theorem, they form a weak-* relatively compact set. We can then extract a subsequence $(N_j)_{j=1}^{\infty}$ such that, for each $\varphi \in L^2(\Omega, \mathbf{P}_\wp)$,

$$\lim_{j\to\infty} \mathbb{E}_{\mathbf{P}_\wp}(\varphi Z_{N_j}) = \mathbb{E}_{\mathbf{P}_\wp}(\varphi Y)$$

for some L^2 random variable Y. If we choose φ to be a function of the x only, it follows that

$$\lim_{j \to \infty} \mathfrak{m}(\varphi \sum_{n=0}^{N_j - 1} Q_\wp^n f) = \mathfrak{m}(\varphi g)$$

where $g = \mathbb{E}_\wp(Y) \in L^2(\mathbb{T}^2, \mathfrak{m})$. On the other hand, for each smooth φ,

$$\mathfrak{m}(\varphi(f - g + Q_\wp g)) = \lim_{j \to \infty} \mathfrak{m}\left(\varphi\left(f - \sum_{n=0}^{N_j - 1} Q_\wp^n f + \sum_{n=0}^{N_j - 1} Q_\wp^{n+1} f\right)\right)$$

$$= \lim_{j \to \infty} \mathfrak{m}(\varphi Q_\wp^{N_j} f) = 0.$$

That is $f = g - Q_\wp g$. Next, consider the L^2 random variables $G_n := g \circ \pi \circ F^n$ and $M_{n+1} = \sum_{k=0}^n (X_k + G_{k+1} - G_k) = G_{n+1} - G_0 + \sum_{k=0}^n X_k$. For each $N \in \mathbb{N}$, we use Jensen's inequality to get

(4.7)
$$C \geq \mathbb{E}_{\mathbf{P}_\wp}\left(\left[\sum_{k=0}^{N-1} X_k\right]^2\right) = \mathbb{E}_{\mathbf{P}_\wp}\left([M_N - G_N + g]^2\right)$$

$$\geq \mathbb{E}_{\mathbf{P}_\wp}\left(M_N^2\right) - 2\sqrt{\mathbb{E}_{\mathbf{P}_\wp}\left(M_N^2\right)\mathbb{E}_{\mathbf{P}_\wp}\left([G_N - g]^2\right)}.$$

In fact, the process (M_n) is a martingale, since

$$\mathbb{E}_{\mathbf{P}_\wp}(G_{k+1} \mid x, \omega_1, \ldots, \omega_k) = \mathbb{E}_{P_{\mathfrak{m}}}(g(x_{k+1}) \mid x_k) = Q_\wp g(x_k) = (Q_\wp g) \circ \pi \circ F^k(x, \omega).$$

Thus,

$$\mathbb{E}_{\mathbf{P}_\wp}\left(M_N^2\right) = \sum_{k=0}^{N-1} \mathbb{E}_{\mathbf{P}_\wp}\left([X_k + G_{k+1} - G_k]^2\right)$$

$$= N\left\{\wp\, \mathfrak{m}([f + g \circ T_0 - g]^2) + (1 - \wp)\mathfrak{m}([f + g \circ T_1 - g]^2)\right\}.$$

The inequality (4.7) and the boundedness of $\mathbb{E}_{\mathbf{P}_\wp}\left([G_N - g]^2\right)$ imply that $\wp\, \mathfrak{m}([f + g \circ T_0 - g]^2) + (1 - \wp)\mathfrak{m}([f + g \circ T_1 - g]^2) = 0$, that is $f + g \circ T_i - g = 0$ for each admissible T_i.

The continuity of g follows from the usual Livschitz rigidity arguments. [12]

In order to prove analyticity of the variance $\Sigma_\wp^2$ with respect to $\wp$, first notice that there is a positive lower bound on the spectral gap $\delta_\wp$ appearing in the proof of Corollary 3.3 in a complex neighborhood of $[0, 1]$. Thus, the series in (4.6) converges

---

[12]In fact, in the present simple case one can provide the following direct proof: clearly $g = (\mathbb{1} - \mathcal{L}_{T_i})^{-1}\mathcal{L}_{T_i} f = \sum_{k=1}^\infty \mathcal{L}_{T_i}^k f$ for an admissible choice of T_i, convergence taking place in the $\|\cdot\|_{p,q}$ norm. Let $v^{u,s}$ be the unstable and stable vectors of T_i, respectively, and $\varphi \in \mathcal{C}^\infty$. Then

$$|\mathfrak{m}(\langle v^u, \nabla\varphi\rangle g)| \leq \sum_{k=1}^\infty |\mathfrak{m}(\varphi\langle v^u, \nabla\mathcal{L}_{T_i} f\rangle)| = \sum_{k=1}^\infty |\mathfrak{m}(\varphi\langle T_i^{-k} v^u, \mathcal{L}_{T_i}^k \nabla f\rangle)| \leq \sum_{k=1}^\infty \|\nabla f\|_\infty \lambda^{-k}\|\varphi\|_{L^1}.$$

On the other hand, $g(x) = \sum_{k=0}^n f \circ T_i^k + g \circ T_i^{n+1}$, and the mixing of T_i (proven exactly as in Corollary 3.3) implies

$$|\mathfrak{m}(\langle v^s, \nabla\varphi\rangle g)| \leq \sum_{k=0}^\infty |\mathfrak{m}(\langle v^s, \nabla\varphi\rangle f \circ T_i^k)| \leq \sum_{k=0}^\infty \|\nabla f\|_\infty \lambda^{-k}\|\varphi\|_{L^1}.$$

Taking the sup over $\{\varphi \in \mathcal{C}^\infty : \|\varphi\|_{L^1} = 1\}$, it follows that $\nabla g \in L^\infty$, which implies $g \in W^{1,2}$. Hence, by Morrey's inequality, $g \in \mathcal{C}^0$.

uniformly in $\wp$. The partial sums are polynomials of $\wp$, hence the limit $\Sigma_\wp^2$ is an analytic function of $\wp$. $\qquad \square$

We finish the section with two simple but important results.

Lemma 4.2. *Denoting $f_2(x,y) := f(x) - f(y)$, the operator $\mathcal{L}^{(2)}_{i\lambda N^{-\frac{1}{2}}f_2,\wp}$ satisfies*

$$\left| \mathfrak{m}_2\big([\mathcal{L}^{(2)}_{i\lambda N^{-\frac{1}{2}}f_2,\wp}]^N 1\big) - e^{-\lambda^2 \Sigma_\wp^2} \right| \le C \frac{1 + |\lambda|^3}{\sqrt{N}}.$$

Proof. The argument follows verbatim the previous discussion. Thus to prove the Lemma we only need to compute the second derivative of the leading eigenvalue, which we still denote μ_ν, and to show that $\mu_0'' = 2\Sigma_\wp^2$. In analogy with (4.3),

$$\mu_0'' = \mathfrak{m}_2(f_2^2) + 2 \sum_{n=1}^\infty \mathfrak{m}_2\left(f_2 [\mathcal{L}^{(2)}_\wp]^n f_2 \right),$$

where $\mathfrak{m}_2$ is the normalized Lebesgue measure on $\mathbb{T}^4$. Then $[\mathcal{L}^{(2)}_\wp]^n f_2(x,y) = \mathcal{L}^n_\wp f(x) - \mathcal{L}^n_\wp f(y)$, $\mathfrak{m}(1) = 1$, and $\mathfrak{m}(f) = 0$ yield $\mu_0'' = 2\Sigma_\wp^2$. $\qquad \square$

Lemma 4.3. *There exists $L_0 > 0$ such that, for all $L \in (0, L_0)$, the following estimate holds,*

$$(4.8) \qquad \mathbf{P}_\wp\left(\left\{ \left| \frac{1}{N} \sum_{k=0}^{N-1} f_{\omega,k} \right| \ge L \right\} \right) \le C e^{-CL^2 N}.$$

Proof. This is an averaged large deviation estimate and can be obtained exactly as the averaged CLT was obtained. Although the idea is standard we give here a sketch of the proof. For any random variable Y, for each $\beta > 0$,

$$\mathbf{P}_\wp(\{Y \ge L\}) \le \mathbb{E}_{\mathbf{P}_\wp}(\mathbb{1}_{\{Y \ge L\}} e^{\beta(Y-L)}) \le e^{-\beta L} \mathbb{E}_{\mathbf{P}_\wp}(e^{\beta Y}).$$

Moreover, $\mathbf{P}_\wp(\{|Y| \ge L\}) = \mathbf{P}_\wp(\{Y \ge L\}) + \mathbf{P}_\wp(\{Y \ge -L\})$. Applying such an inequality to the present situation we have

$$\mathbf{P}_\wp(\{Y \ge L\}) \le e^{-\beta L} \mathfrak{m}(\mathcal{L}^N_{\beta N^{-1}f,\wp} 1).$$

We again apply perturbation theory techniques at the beginning of this section to estimate the right-hand side. Using (4.4) with $\nu = \beta N^{-1}, g = f, \phi_0 = 1$, we have

$$\mathfrak{m}(\mathcal{L}^N_{\beta N^{-1}f,\wp} 1) = \mathfrak{m}\left(\mu_\nu^N \mathcal{Q}_\nu 1 + \mathcal{R}_\nu^N 1 \right) = e^{N \ln(\mu_\nu)}(1 + \mathcal{O}(|\beta| N^{-1})) + \mathcal{O}(\rho^N).$$

If we define the Legendre transform $I_C(L) = \sup_{|\nu| \le C} L\nu - \ln \mu_\nu$ and we call ν_* the value in which the sup is attained, then choosing $\beta = \nu_* N$ we have

$$\mathbf{P}_\wp(\{Y \ge L\}) \le (1 + C|\nu_*|) e^{-N \cdot I_C(L)} + \mathcal{O}(\rho^N).$$

To compute explicitly $I_C(L)$ we expand

$$L\nu - \ln \mu_\nu = L\nu - \frac{1}{2} \nu^2 \mu_0'' + \mathcal{O}(\nu^3).$$

Minimizing this quadratic expression leads to a value of $\nu_* = \frac{L}{\mu_0''}$ and gives (recalling $\mu_0'' = \Sigma_\wp^2$) the estimate,

$$\mathbf{P}_\wp(\{Y \ge L\}) \le 2 e^{-\frac{L^2 N}{2\Sigma_\wp^2}(1-\epsilon)}.$$

provided $L \le C\epsilon$ where $C\epsilon$ is small. $\qquad \square$

5. Quenched CLT

Now that we have the CLT in average we would like to establish it for a large class of sequences. Let $\Sigma_\wp^2$ be the variance of the average CLT with respect to the Bernoulli process with parameter $\wp$. We wish to show that for $\mathbb{P}_\wp$ almost all sequences ω we have the CLT with variance $\Sigma_\wp^2$.

To this end we start with an L^2 estimate: assuming that Y_N is a sequence of random variables such that $\bar{Y} := \lim_N \mathbb{E}_\wp(Y_N)$ exists and is real, we can compute

$$\mathbb{E}_\wp(|Y_N - \bar{Y}|^2) = \mathbb{E}_\wp(|Y_N|^2) - \bar{Y}^2 + 2\bar{Y}\,\Re(\bar{Y} - \mathbb{E}_\wp(Y_N)).$$

Thus, recalling the notation $f_{\omega,k} := f \circ T_{\omega_k} \circ \cdots \circ T_{\omega_1}$ and the bound (4.5),

$$(5.1) \quad \mathbb{E}_\wp\left(\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}}) - e^{-\frac{1}{2}\lambda^2\Sigma_\wp^2}\right|^2\right) = \mathbb{E}_\wp\left(|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}})|^2\right)$$
$$- e^{-\lambda^2\Sigma_\wp^2} + \mathcal{O}\left(\frac{1+|\lambda|^3}{\sqrt{N}}\right).$$

The first term on the right-hand side can be conveniently reinterpreted by introducing a product system. That is, consider the maps $T_{\omega_k} \oplus T_{\omega_k} : \mathbb{T}^4 \to \mathbb{T}^4$, which are represented by the block matrices $\begin{pmatrix} A_{\omega_k} & 0 \\ 0 & A_{\omega_k} \end{pmatrix} \in SL(4, \mathbb{N})$. Clearly they are hyperbolic toral automorphisms (although of a higher dimensional torus) which leave Lebesgue measure invariant. The stable and unstable directions are two dimensional. In perfect analogy with the averaged CLT one can define $f_2(x,y) := f(x) - f(y)$ and study the operator $\mathcal{L}^{(2)}_{i\lambda N^{-\frac{1}{2}}f_2,\wp}$ (see Section 3). A direct computation then shows that, calling $\mathfrak{m}_2$ the normalized Lebesgue measure on $\mathbb{T}^4$,

$$\mathbb{E}_\wp\left(|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}})|^2\right) = \mathbb{E}_\wp\left(\mathfrak{m}_2\left(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_2\circ(T_{\omega_k}\oplus T_{\omega_k})\circ\cdots\circ(T_{\omega_1}\oplus T_{\omega_1})}\right)\right)$$
$$= \mathfrak{m}_2\left(\left[\mathcal{L}^{(2)}_{i\lambda N^{-\frac{1}{2}}f_2,\wp}\right]^N 1\right).$$

By Lemma 4.2 and by (5.1),

$$\mathbb{E}_\wp\left(\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}}) - e^{-\frac{1}{2}\lambda^2\Sigma_\wp^2}\right|^2\right) \leq C\frac{1+|\lambda|^3}{\sqrt{N}}.$$

By Chebyshev inequality the above estimate implies

$$(5.2) \quad \mathbb{P}_\wp\left(\left\{\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}}) - e^{-\frac{1}{2}\lambda^2\Sigma_\wp^2}\right| \geq \varepsilon\right\}\right) \leq C\varepsilon^{-2}\frac{1+|\lambda|^3}{\sqrt{N}}.$$

One would then like to prove almost sure convergence by applying a Borel-Cantelli argument but two problems are in the way: on the one hand the sum over N of the above bound diverges, on the other hand one wants the limit to hold almost surely for all λ, that is one has potentially uncountably many sets to deal with. Both problems can be dealt with by applying Borel-Cantelli to subsequences and then showing that controlling the limit of such sequences one controls the limit for each N and λ. First of all, notice that

$$(5.3) \quad \left|e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}} - e^{i\frac{\lambda_1}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}}\right| \leq \frac{|\lambda - \lambda_1|}{\sqrt{N}}\left|\sum_{k=0}^{N-1}f_{\omega,k}\right|.$$

On the other hand, notice that the estimate (4.8) in Lemma 4.3 also implies

$$
\mathbf{P}_\wp\left(\left\{\left|\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1}f_{\omega,k}-\frac{1}{\sqrt{N+M}}\sum_{k=0}^{N+M-1}f_{\omega,k}\right|\geq\varepsilon\right\}\right)
$$

$$
=\mathbf{P}_\wp\left(\left\{\left|\frac{\sqrt{1+MN^{-1}}-1}{\sqrt{N+M}}\sum_{k=0}^{N+M-1}f_{\omega,k}-\frac{1}{\sqrt{N}}\sum_{k=N}^{N+M-1}f_{\omega,k}\right|\geq\varepsilon\right\}\right)
$$

(5.4)

$$
\leq\mathbf{P}_\wp\left(\left\{\left|\frac{1}{N+M}\sum_{k=0}^{N+M-1}f_{\omega,k}\right|\geq\frac{\varepsilon}{2\sqrt{N+M}[\sqrt{1+MN^{-1}}-1]}\right\}\right)
$$

$$
+\mathbf{P}_\wp\left(\left\{\left|\frac{1}{M}\sum_{k=0}^{M-1}f_{\omega,k}\right|\geq\frac{\varepsilon\sqrt{N}}{2M}\right\}\right)\leq Ce^{-CNM^{-1}\varepsilon^2}.
$$

Next, consider $b\in(\frac{1}{2},1)$ and the sets [13] $A_k:=\{2^k+[j2^{bk}]\}_{j\leq 2^{(1-b)k}}$, $\Lambda_k:=\{-k+jk^{-1}\}_{j\leq k^2}$ and $B_k:=A_k\times\Lambda_k$. For each $(N,\lambda)\in B_k$ let $\Delta_k(N,\lambda)=\{(N_1,\lambda_1)\in\mathbb{N}\times\mathbb{R}\ :\ |N-N_1|\leq 2^{bk}+1,\ |\lambda-\lambda_1|\leq k^{-1}\}$. Clearly

$$
\bigcup_{(N,\lambda)\in B_k}\Delta_k(N,\lambda)\supset\left\{(N,\lambda)\in\mathbb{N}\times\mathbb{R}\ :\ 2^k\leq N\leq 2^{k+1},\ |\lambda|\leq k\right\}=:J_k.
$$

We can then write

$$
\mathbb{P}_\wp\left(\left\{\sup_{(N,\lambda)\in J_k}\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_\wp^2}\right|\geq 4\varepsilon\right\}\right)
$$

$$
\leq\sum_{(N,\lambda)\in B_k}\mathbb{P}_\wp\left(\left\{\sup_{(N_1,\lambda_1)\in\Delta_k(N,\lambda)}\left|\mathfrak{m}(e^{i\frac{\lambda_1}{\sqrt{N_1}}\sum_{l=0}^{N_1-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda_1^2\Sigma_\wp^2}\right|\geq 4\varepsilon\right\}\right)
$$

$$
\leq\sum_{(N,\lambda)\in B_k}\left\{\mathbb{P}_\wp\left(\left\{\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_\wp^2}\right|\geq\varepsilon\right\}\right)\right.
$$

$$
+\mathbb{P}_\wp\left(\left\{\mathfrak{m}\left(\sup_{(N_1,\lambda_1)\in\Delta_k(N,\lambda)}\left|e^{i\frac{\lambda_1}{\sqrt{N_1}}\sum_{l=0}^{N_1-1}f_{\omega,l}}-e^{i\frac{\lambda_1}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}}\right|\right)\geq\varepsilon\right\}\right)
$$

$$
\left.+\mathbb{P}_\wp\left(\left\{\mathfrak{m}\left(\sup_{(N_1,\lambda_1)\in\Delta_k(N,\lambda)}\left|e^{i\frac{\lambda_1}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}}-e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}}\right|\right)\geq\varepsilon\right\}\right)\right\},
$$

where we have assumed $\Sigma_\wp k^{-1}\leq\varepsilon$ in order to deal with the difference $e^{-\frac{1}{2}\lambda^2\Sigma_\wp^2}-e^{-\frac{1}{2}\lambda_1^2\Sigma_\wp^2}$. For each bounded function $g\geq 0$ holds

$$
\mathbf{P}_\wp(\{g\geq A\})=\mathbb{E}_{\mathbf{P}_\wp}(\mathbb{1}_{\{g\geq A\}})\geq\mathbb{E}_\wp(\mathfrak{m}(\mathbb{1}_{\{g\geq A\}})\mathbb{1}_{\{\mathfrak{m}(g)\geq 2A\}}).
$$

But $\mathfrak{m}(g)\leq|g|_\infty\mathfrak{m}(\{g\geq A\})+A$, and $\mathfrak{m}(g)\geq 2A$ implies $\mathfrak{m}(\{g\geq A\})\geq A|g|_\infty^{-1}$. Thus,

$$
\mathbb{P}_\wp(\{\mathfrak{m}(g)\geq 2A\})\leq A^{-1}|g|_\infty\mathbf{P}_\wp(\{g\geq A\}).
$$

---

[13]Here $[x]$ stands for the integer closest to x.

We can estimate the above expression by

$$\mathbb{P}_{\wp}\left(\left\{\sup_{(N,\lambda)\in J_k}\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_{\wp}^2}\right|\geq 4\varepsilon\right\}\right)$$

$$\leq \sum_{(N,\lambda)\in B_k}\left[\mathbb{P}_{\wp}\left(\left\{\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_{\wp}^2}\right|\geq\varepsilon\right\}\right)\right.$$

$$+4\varepsilon^{-1}\mathbf{P}_{\wp}\left(\left\{\sup_{(N_1,\lambda_1)\in\Delta_k(N,\lambda)}\left|e^{i\frac{\lambda_1}{\sqrt{N_1}}\sum_{l=0}^{N_1-1}f_{\omega,l}}-e^{i\frac{\lambda_1}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}}\right|\geq\frac{\varepsilon}{2}\right\}\right)$$

$$\left.+4\varepsilon^{-1}\mathbf{P}_{\wp}\left(\left\{\sup_{(N_1,\lambda_1)\in\Delta_k(N,\lambda)}\left|e^{i\frac{\lambda_1}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}}-e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}}\right|\geq\frac{\varepsilon}{2}\right\}\right)\right],$$

Thus, remembering (5.3),

$$\mathbb{P}_{\wp}\left(\left\{\sup_{(N,\lambda)\in J_k}\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_{\wp}^2}\right|\geq 4\varepsilon\right\}\right)$$

$$\leq\sum_{(N,\lambda)\in B_k}\left[\mathbb{P}_{\wp}\left(\left\{\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_{\wp}^2}\right|\geq\varepsilon\right\}\right)\right.$$

$$+4\varepsilon^{-1}\mathbf{P}_{\wp}\left(\left\{\sup_{(N_1,\lambda_1)\in\Delta_k(N,\lambda)}|\lambda_1|\left|\frac{1}{\sqrt{N_1}}\sum_{l=0}^{N_1-1}f_{\omega,l}-\frac{1}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}\right|\geq\frac{\varepsilon}{2}\right\}\right)$$

$$\left.+4\varepsilon^{-1}\mathbf{P}_{\wp}\left(\left\{\sup_{(N_1,\lambda_1)\in\Delta_k(N,\lambda)}\frac{|\lambda_1-\lambda|}{\sqrt{N}}\left|\sum_{l=0}^{N-1}f_{\omega,l}\right|\geq\frac{\varepsilon}{2}\right\}\right)\right]$$

$$\leq\sum_{(N,\lambda)\in B_k}\left[\mathbb{P}_{\wp}\left(\left\{\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_{\wp}^2}\right|\geq\varepsilon\right\}\right)\right.$$

$$+4\varepsilon^{-1}\sum_{|N_1-N|\leq 2^{bk}+1}\mathbf{P}_{\wp}\left(\left\{\left|\frac{1}{\sqrt{N_1}}\sum_{l=0}^{N_1-1}f_{\omega,l}-\frac{1}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}\right|\geq\frac{\varepsilon}{4k}\right\}\right)$$

$$\left.+4\varepsilon^{-1}\mathbf{P}_{\wp}\left(\left\{\frac{1}{\sqrt{N}}\left|\sum_{l=0}^{N-1}f_{\omega,l}\right|\geq\frac{k\varepsilon}{2}\right\}\right)\right].$$

Then the estimates (4.8), (5.4) and (5.2) imply, for $k\geq\Sigma_{\wp}\varepsilon^{-1}$,

$$\mathbb{P}_{\wp}\left(\left\{\sup_{(N,\lambda)\in J_k}\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_{\wp}^2}\right|\geq 4\varepsilon\right\}\right)$$

$$\leq C\sum_{(N,\lambda)\in B_k}\left[\varepsilon^{-2}\frac{1+|\lambda|^3}{\sqrt{N}}+\varepsilon^{-1}\sum_{|N_1-N|\leq 2^{bk}+1}e^{-C\frac{N}{|N-N_1|}\varepsilon^2 k^{-2}}+\varepsilon^{-1}e^{-Ck^2\varepsilon^2}\right]$$

$$\leq Ck^2 2^{(1-b)k}\varepsilon^{-1}\left[\varepsilon^{-1}k^3 2^{-\frac{k}{2}}+2^{bk}e^{-C2^{k(1-b)}\varepsilon^2 k^{-2}}+e^{-C\varepsilon^2 k^2}\right],$$

for which it follows that the sum over k is finite. By Borel-Cantelli it follows that the above events $\left\{\sup_{(N,\lambda)\in J_k}\left|\mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{l=0}^{N-1}f_{\omega,l}})-e^{-\frac{1}{2}\lambda^2\Sigma_{\wp}^2}\right|\geq 4\varepsilon\right\}$ happen only finitely many times with probability one. That is, for each $\varepsilon>0$, there exists a

random variable $N_\varepsilon : \Omega \to \mathbb{N} \cup \{\infty\}$, $\mathbb{P}_\wp$-almost surely finite, such that

$$\sup_{|\lambda| \leq \log_2 N} \left| \mathfrak{m}(e^{i\frac{\lambda}{\sqrt{N}}\sum_{k=0}^{N-1} f_{\omega,k}}) - e^{-\frac{1}{2}\lambda^2 \Sigma_\wp^2} \right| \leq \varepsilon \quad \text{for} \quad N \geq N_\varepsilon.$$

Here we used the fact that, for each fixed N, $|\lambda| \leq \log_2 N$ implies $(N, \lambda) \in J_{\lfloor \log_2 N \rfloor}$. Let us call $\widetilde{\Omega}_\varepsilon$ the bad set of sequences, involving $N_\varepsilon = \infty$. It is an increasing set with decreasing ε, such that $\mathbb{P}_\wp(\bigcup_{\varepsilon>0} \widetilde{\Omega}_\varepsilon) = \lim_{\varepsilon \downarrow 0} \mathbb{P}_\wp(\widetilde{\Omega}_\varepsilon) = 0$; the bad set is independent of ε.

This concludes the proof and establishes the almost sure CLT where almost sure means that, fixing any Bernoulli measure, the set of the sequences for which we do not have CLT has zero measure. Note, however, that the limit (more precisely, the variance) is not constant but depends on $\wp$. This is natural since the deterministic limits $\wp = 0$ and $\wp = 1$ generically have different variances and as $\wp$ varies, the variance should interpolate smoothly between these two extremal values, which indeed is confirmed by Lemma 4.1.

REFERENCES

[1] A. Ayyer, M. Stenlund, *Exponential Decay of Correlations for Randomly Chosen Hyperbolic Toral Automorphisms*, Chaos, to appear.

[2] V. I. Bakhtin, *Random processes generated by a hyperbolic sequence of mappings. I*, Russian Acad. Sci. Izv. Math. **44** (1995), no. 2, 247–279,
Random processes generated by a hyperbolic sequence of mappings. II, Russian Acad. Sci. Izv. Math. **44** (1995), no. 3, 617–627.

[3] V. Baladi, *Positive transfer operators and decay of correlations*, Advanced Series in Nonlinear Dynamics, **16**, World Scientific (2000).

[4] V. Baladi, *Anisotropic Sobolev spaces and dynamical transfer operators: $\mathcal{C}^\infty$ foliations*, Preprint.

[5] V. Baladi, M. Tsujii, *Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms*, to appear in Ann. Inst. Fourier.

[6] M. Blank, G. Keller, C. Liverani, *Ruelle-Perron-Frobenius spectrum for Anosov maps*, Nonlinearity, **15**:6 (2001), 1905-1973.

[7] Mark F. Demers, C. Liverani, *Stability of Statistical Properties in Two-dimensional Piecewise Hyperbolic Maps*, to appear in Transactions of the American Mathematical Society (2007).

[8] D. Dolgopyat, D. Szasz, T. Varju, *Recurrence properties of Lorentz gas*, to appear in Duke Math. J..

[9] J.-R. Chazottes, S. Gouëzel, *On almost-sure versions of classical limit theorems for dynamical systems*, Probability Theory and Related Fields 138: 195–234, 2007.

[10] J.-P. Conze, *Sur un critére de rècurrence en dimension 2 pour les marches stationnaires, applications.* Ergodic Theory Dynam. Systems **19** (1999), no. 5, 1233–1245.

[11] D. Dolgopyat, G. Keller, C. Liverani, *Random Walk in Markovian Environment*, preprint.

[12] F. Faure, N. Roy, *Ruelle-Pollicott resonances for real analytic hyperbolic maps*, Nonlinearity **19** (2006), no. 6, 1233–1252.

[13] S. Gouëzel, C. Liverani, *Banach spaces adapted to Anosov systems*, Ergodic Theory and Dynamical Systems, 26, 1, 189–217, (2006).

[14] A. Katok, B.Hasselblatt, *Introduction to the modern theory of dynamical systems.* With a supplementary chapter by Katok and Leonardo Mendoza. Encyclopedia of Mathematics and its Applications, 54. Cambridge University Press, Cambridge (1995).

[15] Marco Lenci, *Aperiodic Lorentz gas: recurrence and ergodicity.* Ergodic Theory Dynam. Systems **23** (2003), no. 3, 869–883.

[16] C. Liverani, M. Wojtkowski, *Generalization of the Hilbert Metric to the Space of Positive Definite Matrices*, Pacific Journal of Mathematics, 166 n. 2, pp. 339-355, (1994).

[17] C. Liverani, M. Wojtkowski, *Ergodicity in Hamiltonian Systems*, Dynamics Reported, 4, pp. 130-202, (1995).

[18] L. A. Bunimovich, Ya. G. Sinai, N. I. Chernov, *Statistical properties of two-dimensional hyperbolic billiards*, Russ. Math. Surv., **46** (1991), no. 4, 47–106.

[19] N. Simanyi, *Towards a Proof of Recurrence for the Lorentz Process.* (Dynamical Systems and Ergodic Theory, Banach Center Publications, **23**). PWN, Polish Scientific Publishers, Warsaw, 1989.

[20] Klaus Schmidt, *On joint recurrence.* C. R. Acad. Sci. Paris Sr. I Math. **327** (1998), no. 9, 837–842.

ARVIND AYYER, DEPARTMENT OF PHYSICS, RUTGERS UNIVERSITY, 136 FRELINGHUYSEN ROAD, PISCATAWAY, NJ 08854, USA.
   E-mail address: `ayyer@physics.rutgers.edu`

CARLANGELO LIVERANI, DIPARTIMENTO DI MATEMATICA, II UNIVERSITÀ DI ROMA (TOR VERGATA), VIA DELLA RICERCA SCIENTIFICA, 00133 ROMA, ITALY.
   E-mail address: `liverani@mat.uniroma2.it`

MIKKO STENLUND, DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, 110 FRELINGHUYSEN ROAD, PISCATAWAY, NJ 08854, USA.
   E-mail address: `mstenlun@math.rutgers.edu`