



**HAL**  
open science

## On the CLT for rotations and BV functions

Jean-Pierre Conze, Stéphane Le Borgne

► **To cite this version:**

Jean-Pierre Conze, Stéphane Le Borgne. On the CLT for rotations and BV functions. *Comptes Rendus. Mathématique*, 2019, 357 (2), pp.212-215. 10.1016/j.crma.2019.01.008 . hal-01777584v3

**HAL Id: hal-01777584**

**<https://hal.science/hal-01777584v3>**

Submitted on 8 Jun 2021 (v3), last revised 10 Jan 2022 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ON THE CLT FOR ROTATIONS AND BV FUNCTIONS

JEAN-PIERRE CONZE AND STÉPHANE LE BORGNE

IRMAR - UMR 6625, F-35000 Rennes, France

ABSTRACT. Let  $x \mapsto x + \alpha \pmod{1}$  be a rotation on the circle and let  $\varphi$  be a step function. We denote by  $\varphi_n(x)$  the corresponding ergodic sums  $\sum_{j=0}^{n-1} \varphi(x + j\alpha)$ . For a class of irrational rotations, containing the class with bounded partial quotients, and under a Diophantine condition on the discontinuity points of  $\varphi$ , we show that  $\varphi_n/\|\varphi_n\|_2$  is asymptotically Gaussian for  $n$  in a set of density 1. The proof is based on decorrelation inequalities for the ergodic sums taken at times  $q_k$ , where  $(q_k)$  is the sequence of denominators of  $\alpha$ . Another important point is the control of the variance  $\|\varphi_n\|_2^2$  for  $n$  belonging to a large set of integers. When  $\alpha$  is a quadratic irrational, the size of this set can be precisely estimated.

## CONTENTS

|  |    |
|--|----|
| <b>Introduction</b>  | 1  |
| <b>1. Variance of the ergodic sums</b>                             | 3  |
| 1.1. Reminders on continued fractions, Ostrowski's expansion       | 3  |
| 1.2. Bounds for the variance                                       | 5  |
| 1.3. A counter-example   | 8  |
| 1.4. A special case: quadratic numbers                             | 9  |
| <b>2. A central limit theorem and its application to rotations</b> | 17 |
| 2.1. Decorrelation and CLT   | 17 |
| 2.2. Application to step functions, examples                       | 20 |
| <b>3. Proof of Proposition 2.1 (CLT)</b>                           | 23 |
| <b>4. Proof of Proposition 2.3 (decorrelation)</b>                 | 30 |
| References   | 36 |

## Introduction

For a dynamical system  $(X, \mu, T)$  and an observable  $\varphi$  on  $X$ , a general question is the asymptotic behaviour in distribution of the ergodic sums  $\sum_0^{L-1} \varphi \circ T^k$  after normalisation.

---

*Date:* June 8, 2021.

For a large class of observables and chaotic systems, many results of convergence toward a Gaussian distribution are known.

When the dynamical system has zero entropy, in particular for a rotation, the situation is different. Nevertheless one can ask if, at least, there are observables satisfying a non degenerate Central Limit Theorem. In this direction there are positive answers: M. Denker and R. Burton (1987), then M. Weber, M. Lacey and other authors proved for rotations the existence of functions, whose ergodic sums satisfy a CLT after self-normalization. In general for a measure preserving aperiodic system, further results by D. Volný and P. Liardet (1997), J.-P. Thouvenot and B. Weiss (2012) showed that any distribution can appear as a limiting distribution of the ergodic sums of some functions after normalisation.

A different question is to ask if, for smooth systems, there is a CLT for explicit functions in a certain class of regularity. Here we consider step functions on  $X = \mathbb{R}/\mathbb{Z}$  and their ergodic sums  $\varphi_n(x) := \sum_0^{n-1} \varphi(x + j\alpha)$  over an irrational rotation  $x \mapsto x + \alpha \pmod{1}$ .

By the Denjoy-Koksma inequality, if  $\varphi$  is a centered function with bounded variation, the sequence  $(\varphi_n)$  is uniformly bounded along the sub-sequence of denominators of  $\alpha$ . But, besides, a stochastic behaviour at a certain scale can occur along other sub-sequences  $(n_k)$ . We propose a quantitative analysis of this phenomenon.

Let us mention the following related papers. For  $\psi := 1_{[0, \frac{1}{2}[} - 1_{[\frac{1}{2}, 0]}$ , F. Huveneers [9] studied the existence of a sequence  $(n_k)_{k \in \mathbb{N}}$  such that  $(\psi_{n_k})$  after normalization is asymptotically normally distributed. In [4] it was shown that, when  $\alpha$  has unbounded partial quotients, along some subsequences the ergodic sums of  $\varphi$  in a class of step functions can be approximated by a Brownian motion.

Here we will use as in [9] a method based on decorrelation inequalities which applies in particular when the sequence of partial quotients of  $\alpha$  is bounded or under a slightly more general Diophantine assumption. It relies on an abstract central limit theorem valid under some suitable decorrelation conditions. If  $\varphi$  is a step function, we give conditions which insure that for  $n$  in a set of integers of density 1, the distribution of  $\varphi_n / \|\varphi_n\|_2$  is asymptotically Gaussian (Theorem 2.4). Beside the remarkable recent “temporal” limit theorems for rotations (see [1], [6], [2]), this shows that a “spatial” asymptotic normal distribution can also be observed for  $n$  in a large set of integers.

An important point is the control of the variance  $\|\varphi_n\|_2^2$ . We will see that, at least in the case of  $\alpha$  with bounded partial quotients, we can expect that, for  $n$  belonging to a set of density 1,  $\|\varphi_n\|_2^2$  should be approximately of order  $\ln n$ .

In Section 1, we study the set of integers for which the variance of the ergodic sums is big. Our most precise information is obtained in the special case where  $\alpha$  is a quadratic irrational.

The central limit theorem is presented in Section 2. It is based on the decorrelation between the ergodic sums at times  $q_k$ , the denominators of  $\alpha$ , and on an abstract central limit theorem. To apply the results to a step function, a Diophantine condition is needed on the discontinuities of  $\varphi$  which holds generically.

The proofs of the CLT and of the decorrelation are given in Sections 3 and 4.

The results of this paper have been announced in [3].

## 1. Variance of the ergodic sums

**Notation** The uniform measure on  $\mathbb{T}^1$  identified with  $X = [0, 1[$  is denoted by  $\mu$ . A function  $\varphi$  on  $\mathbb{T}^1$  is identified with a 1-periodic function of a real variable. We denote by  $V(\varphi)$  the *variation* of the restriction of  $\varphi$  to  $[0, 1]$  and write BV for “bounded variation”.

The class of real centered BV functions on  $\mathbb{T}^1$  is denoted by  $\mathcal{C}$ . It contains the 1-periodic step functions with a finite number of discontinuities. The Fourier coefficients  $\hat{\varphi}(r)$  of a function  $\varphi \in \mathcal{C}$  satisfy:

$$(1) \quad \hat{\varphi}(r) = \frac{\gamma_r(\varphi)}{r}, r \neq 0, \text{ with } K(\varphi) := \sup_{r \neq 0} |\gamma_r(\varphi)| \leq \frac{V(\varphi)}{2\pi} < +\infty.$$

Let  $\alpha = [0; a_1, a_2, \dots]$  be an irrational number in  $]0, 1[$ , with partial quotients  $a_n = a_n(\alpha)$ , numerators  $p_n$  and denominators  $q_n$ ,  $n \geq 1$ .

The ergodic sums  $\sum_{j=0}^{n-1} \varphi(x + j\alpha)$  of a 1-periodic function  $\varphi$  for the rotation by  $\alpha$  are denoted by  $\varphi_n(x)$ . Their Fourier expansion is  $\varphi_n(x) = \sum_{r \neq 0} \frac{\gamma_r(\varphi)}{r} e^{\pi i(n-1)r\alpha} \frac{\sin \pi n r \alpha}{\sin \pi r \alpha} e^{2\pi i r x}$ .

If  $\varphi \in \mathcal{C}$ , then  $V(\varphi_n) \leq nV(\varphi)$  and  $|\widehat{\varphi_n}(r)| = \frac{|\gamma_r(\varphi)|}{|r|} \frac{|\sin \pi n r \alpha|}{|\sin \pi r \alpha|} \leq \frac{nK(\varphi)}{|r|}, r \neq 0$ .

### 1.1. Reminders on continued fractions, Ostrowski's expansion.

For  $u \in \mathbb{R}$ ,  $\{u\}$  denotes its fractional part and  $\|u\| := \inf_{n \in \mathbb{Z}} |u - n| = \min(\{u\}, 1 - \{u\})$  its distance to  $\mathbb{Z}$ . Recall that  $2\|x\| \leq |\sin \pi x| \leq \pi\|x\|$ ,  $\forall x \in \mathbb{R}$ .

For  $n \geq 1$ , writing  $\alpha = \frac{p_n}{q_n} + \frac{\theta_n}{q_n}$ , we have

$$(2) \quad \frac{1}{a_{n+1} + 2} \leq \frac{q_n}{q_{n+1} + q_n} \leq q_n \|q_n \alpha\| = q_n |\theta_n| \leq \frac{q_n}{q_{n+1}} = \frac{q_n}{a_{n+1} q_n + q_{n-1}} \leq \frac{1}{a_{n+1}},$$

$$(3) \quad \theta_n = (-1)^n \|q_n \alpha\|, \quad \alpha = \frac{p_n}{q_n} + (-1)^n \frac{\|q_n \alpha\|}{q_n}, \quad \frac{1}{2} q_{n+1}^{-1} \leq \theta_n \leq q_{n+1}^{-1},$$

$$(4) \quad q_{n+1}/q_{n+k} \leq C \rho^k, \quad \forall n, k \geq 1, \text{ with } C = \frac{5 + \sqrt{5}}{2}, \rho = \frac{\sqrt{5} - 1}{2} < 1.$$

Let us show the last inequality: for  $n \geq 1$  fixed, putting  $r_0 = q_n$ ,  $r_1 = q_{n+1}$ ,  $r_{k+1} = r_k + r_{k-1}$ , for  $k \geq 1$ , we have  $q_{n+k} \geq r_k$ ,  $\forall k \geq 0$  by induction and (4) follows easily.

For  $n \geq 1$ , we denote by  $m(n)$  the integer such that  $n \in [q_{m(n)}, q_{m(n)+1}[$ .

If  $\alpha$  has bounded partial quotients (i.e.,  $B := \sup a_n < \infty$ ), then  $m(n)$  is of order  $\ln n$ : there is a constant  $\lambda > 0$  such that  $\frac{\ln n}{m(n)} \in [\lambda, \ln(B + 1)]$ .

Every integer  $n \geq 1$  can be represented as follows ( $\alpha$ -Ostrowski's expansion):

$$(5) \quad \text{if } n < q_{m+1}, n = \sum_{k=0}^m b_k q_k, \text{ with } 0 \leq b_0 \leq a_1 - 1, 0 \leq b_k \leq a_{k+1} \text{ for } 1 \leq k \leq m.$$

Indeed, if  $n \in [q_0, q_1 = a_1[$ , then (5) is satisfied, and if  $n \in [q_m, q_{m+1}[$  with  $m \geq 1$ , we write  $n = b_m q_m + r$ , with  $1 \leq b_m \leq a_{m+1}$ ,  $0 \leq r < q_m$ . By iteration, we get either  $r = 0$  at some point and the algorithm stops, or  $n \in [q_0, q_1[$ . In either cases we obtain (5).

In this way, we can code every  $n < q_{m+1}$  by a word  $b_0 \dots b_m$ , with  $b_0 \in \{0, 1, \dots, a_1 - 1\}$  and  $b_j \in \{0, 1, \dots, a_{j+1}\}$ ,  $j = 1, \dots, m$ .

In this representation,  $b_{m(n)} \neq 0$  and  $b_j = 0$  for  $m(n) < j \leq m$  when  $m > m(n)$ . In the latter case, there are  $m - m(n)$  zero's at the right end. For a given  $m$  and  $n < q_{m+1}$ , this Ostrowski's expansion is "proper" (without zeros at the end) if  $m = m(n)$ .

For  $m \geq 0$ , we call *admissible of length  $m + 1$*  a finite word  $b_0 \dots b_m$  such that  $b_0 \in \{0, 1, \dots, a_1 - 1\}$ ,  $b_j \in \{0, 1, \dots, a_{j+1}\}$ , for  $j = 1, \dots, m$  and such that, for two consecutive letters  $b_{j-1}, b_j$ , if  $b_j = a_{j+1}$  then  $b_{j-1} = 0$ .

Remark that if  $b_0 \dots b_m$  is admissible,  $m \geq 1$ , then  $b_0 \dots b_{m-1}$  is admissible. Let us show by induction that the Ostrowski's expansion of an integer  $n$  is admissible. Let  $n$  be in  $[q_m, q_{m+1}[$ . We start the construction of the expansion of  $n$  as above. Now the following steps of the algorithm yield the Ostrowski's expansion of  $n - b_m q_m$ . Since  $n - b_m q_m \in [0, q_m[$ , the Ostrowski's expansion of  $n - b_m q_m$  is admissible. It remains to check that, if  $b_m = a_{m+1}$ , then  $b_{m-1} = 0$ . But if  $b_{m-1} \neq 0$ , we would have  $n \geq a_{m+1} q_m + q_{m-1} = q_{m+1}$ , a contradiction.

Conversely, if  $b_0 \dots b_m$  is admissible, one shows by induction that  $b_0 + b_1 q_1 + \dots + b_m q_m < q_{m+1}$ . This holds if  $m = 0$ , since  $b_1 < q_1 = a_1$ . Assume that this is true for the length  $m$ . Let  $b_0 \dots b_m b_{m+1}$  be admissible of length  $m + 1$ .

If  $b_{m+1} = a_{m+2}$ , then  $b_m = 0$  and  $b_0 + b_1 q_1 + \dots + b_m q_m = b_0 + b_1 q_1 + \dots + b_{m-1} q_{m-1} < q_m$ , so that  $b_0 + b_1 q_1 + \dots + b_{m+1} q_{m+1} < q_m + a_{m+2} q_{m+1} = q_{m+2}$ .

If  $b_{m+1} \leq a_{m+2} - 1$ , then  $b_0 + b_1 q_1 + \dots + b_{m+1} q_{m+1} < q_{m+1} + (a_{m+2} - 1) q_{m+1} < q_{m+2}$ .

Therefore, if we associate to an admissible word the integer  $n = b_0 + b_1 q_1 + \dots + b_m q_m$ , there is a bijection between the Ostrowski's expansions of integers  $n < q_{m+1}$  and the set of admissible words of length  $m + 1$ . The number of admissible words of length  $m$  is  $q_m - 1$ .

For  $n$  given by (5), putting  $n_0 = b_0$ ,  $n_k = \sum_{t=0}^k b_t q_t$ , for  $k \leq m(n)$ , we have

$$(6) \quad \begin{aligned} \varphi_n(x) &= \sum_{k=0}^{m(n)} \sum_{j=n_{k-1}}^{n_k-1} \varphi(x + j\alpha) = \sum_{k=0}^{m(n)} \sum_{j=0}^{b_k q_k - 1} \varphi(x + n_{k-1}\alpha + j\alpha) \\ &= \sum_{k=0}^{m(n)} \sum_{i=0}^{b_k-1} \varphi_{q_k}(x + (n_{k-1} + i q_k)\alpha) = \sum_{k=0}^{m(n)} f_k(x). \end{aligned}$$

$$(7) \quad \text{with } f_k(x) := \sum_{i=0}^{b_k-1} \varphi_{q_k}(x + (n_{k-1} + i q_k)\alpha) = \varphi_{b_k q_k}(x + n_{k-1}\alpha),$$

By convention, we put  $\sum_{i=0}^{b_k-1} \varphi_{q_k}(x + (n_{k-1} + iq_k)\alpha) = 0$ , if  $b_k = 0$ .

If  $\varphi$  is a BV centered function, then it holds (*Denjoy-Koksma inequality*):

$$(8) \quad \|\varphi_q\|_\infty = \sup_x \left| \sum_{i=0}^{q-1} \varphi(x + i\alpha) \right| \leq V(\varphi), \text{ if } q \text{ is a denominator of } \alpha.$$

One can also show that if  $\varphi$  satisfies (1) then  $\|\varphi_{q_n}\|_2 \leq 2\pi K(\varphi)$ . By (8), we have for  $f_k$  defined by (7):  $\|f_k\|_\infty \leq b_k V(\varphi) \leq a_{k+1} V(\varphi)$ .

## 1.2. Bounds for the variance.

Let  $\varphi \in \mathcal{C}$  and  $n \in [q_{\ell-1}, q_\ell[$ . The variance is bounded from below as follows:

$$\|\varphi_n\|_2^2 = 2 \sum_{k>1} |\widehat{\varphi}(k)|^2 \frac{(\sin \pi n k \alpha)^2}{(\sin \pi k \alpha)^2} \geq 2 \sum_{j=1}^{\ell} |\widehat{\varphi}(q_j)|^2 \frac{(\sin \pi n q_j \alpha)^2}{(\sin \pi q_j \alpha)^2} \geq c_0 \sum_{j=1}^{\ell} |\widehat{\varphi}(q_j)|^2 \frac{\|n q_j \alpha\|^2}{\|q_j \alpha\|^2},$$

with  $c_0 = \frac{8}{\pi^2}$ . Therefore, by (2) we have, for  $0 < \delta < \frac{1}{2}$ ,

$$(9) \quad \|\varphi_n\|_2^2 \geq c_0 \sum_{j=1}^{\ell} |\gamma_{q_j}(\varphi)|^2 a_{j+1}^2 \|n q_j \alpha\|^2 \geq c_0 \delta^2 \sum_{j=1}^{\ell} |\gamma_{q_j}(\varphi)|^2 a_{j+1}^2 \mathbf{1}_{\|n q_j \alpha\| \geq \delta}.$$

An upper bound for the variance and a lower bound for the mean of the variance are shown in [4]: there are constants  $C, c > 0$  such that

$$(10) \quad \|\varphi_n\|_2^2 \leq C K(\varphi)^2 \sum_{j=0}^{\ell} a_{j+1}^2, \quad \forall n \in [q_\ell, q_{\ell+1}[,$$

$$(11) \quad \frac{1}{n} \sum_{k=0}^{n-1} \|\varphi_k\|_2^2 \geq c \sum_{j=0}^{\ell-1} |\gamma_{q_j}(\varphi)|^2 a_{j+1}^2, \quad \forall n \in [q_\ell, q_{\ell+1}[.$$

Now our first goal will be to bound from below the variance  $\|\varphi_n\|_2$  for  $n$  in a large set.

### Bounds for the variance for $n$ in a large set of integers

According to (9), a lower bound for  $\|\varphi_n\|_2$  depends on two separate conditions:

First we need the following condition on the Fourier coefficients of  $\varphi$ :

$$(12) \quad \exists M, \eta, \theta > 0 \text{ such that } \text{Card} \{j \leq N : a_{j+1} |\gamma_{q_j}(\varphi)| \geq \eta\} \geq \theta N, \quad \forall N \geq M.$$

This condition clearly holds when  $\varphi$  is the function  $\varphi^0(x) = \{x\} - \frac{1}{2}$ , since in this case  $|\gamma_{q_j}(\varphi^0)| = \frac{1}{2\pi}$ ,  $\forall j$ . Its validity, related to Diophantine conditions on the points of discontinuities, will be discussed for some step functions in Subsection 2.2.

For  $j < \ell$ , we will estimate how many times  $\{n q_j \alpha\} \in I_\delta := [0, \delta] \cup [1 - \delta, 1]$  for  $n \leq q_\ell$  and deduce from this estimation that  $\sum_{j=1}^{\ell} \mathbf{1}_{I_\delta}(\{n q_j \alpha\}) = \sum_{j=1}^{\ell} \mathbf{1}_{\|n q_j \alpha\| \leq \delta}$  is small for a large set of values of  $n$ .

**Lemma 1.1.** *For every  $\delta \in ]0, \frac{1}{2}[$  and every interval of integers  $I = [N_1, N_1 + L[$ , we have*

$$(13) \quad \sum_{n=N_1}^{N_1+L-1} 1_{I_\delta}(\{n q_j \alpha\}) \leq 20(\delta + q_{j+1}^{-1})L, \forall j \text{ such that } q_{j+1} \leq 2L.$$

*Proof.* For a fixed  $j$  and  $0 \leq N_1 < N_1 + L$ , let us describe the behaviour of the sequence  $(\|n q_j \alpha\|, n = N_1, \dots, N_1 + L - 1)$ .

Recall that (modulo 1) we have  $q_j \alpha = \theta_j$ , with  $\theta_j = (-1)^j \|q_j \alpha\|$  (see (3)). We treat the case  $j$  even (hence  $\theta_j > 0$ ). The case  $j$  odd is analogous.

We are going to count how many times, for  $j$  even, we have  $\{n \theta_j\} < \delta$  or  $1 - \delta < \{n \theta_j\}$ .

We start with  $n_1 := N_1$ . Putting  $w(j, 1) := \{n_1 \theta_j\}$ , we have  $\{n \theta_j\} = w(j, 1) + (n - n_1)\theta_j$ , for  $n = n_1, n_1 + 1, \dots, n_2 - 1$ , where  $n_2$  is such that  $w(j, 1) + (n_2 - 1 - n_1)\theta_j < 1 < w(j, 1) + (n_2 - n_1)\theta_j$ .

Putting  $w(j, 2) := \{n_2 \theta_j\}$ , we have  $w(j, 2) = w(j, 1) + (n_2 - n_1)\theta_j - 1 < \theta_j$ . Starting now from  $n_2$ , we have  $\{n \theta_j\} = w(j, 2) + (n - n_2)\theta_j$  for  $n = n_2, n_2 + 1, \dots, n_3 - 1$ , where  $n_3$  is such that  $w(j, 2) + (n_3 - 1 - n_2)\theta_j < 1 < w(j, 2) + (n_3 - n_2)\theta_j$ .

We iterate up to  $R(j)$ , where  $n_{R(j)-1} < N_1 + L \leq n_{R(j)}$ . This construction yields a sequence  $n_1 < n_2 < \dots < n_{R(j)}$  such that  $\{n \theta_j\} = w(j, i) + (n - n_i)\theta_j$ ,  $\forall n \in [n_i, n_{i+1}[$ , and

$$w(j, i) + (n_{i+1} - 1 - n_i)\theta_j < 1 < w(j, i) + (n_{i+1} - n_i)\theta_j,$$

with  $w(j, i)$  defined recursively by  $w(j, i + 1) = \{w(j, i) + (n_{i+1} - n_i)\theta_j\}$  and satisfying  $w(j, i) < \theta_j$ , for  $i = 1, \dots, R(j)$ .

Since  $(n_{i+1} - n_i + 1)\theta_j \geq w(j, i) + (n_{i+1} - n_i)\theta_j > 1$  for  $i \neq 1$  and  $i \neq R(j)$ , we have  $n_{i+1} - n_i \geq \theta_j^{-1} - 1$ , for each  $i \neq 1, R(j)$ . This implies  $R(j) \leq \frac{L}{\theta_j^{-1} - 1} + 2$ .

For each  $i$ , the number of integers  $n \in [n_i, n_{i+1} - 1[$  such that  $\{n \theta_j\} \in [0, \delta[ \cup ]1 - \delta, 1[$  is bounded by  $2(1 + \delta \theta_j^{-1})$ . (This number is less than 2 if  $\delta < \theta_j$ .)

Altogether, using (3) and the assumption  $2L \geq q_{j+1}$ , the number of integers  $n \in I$  such that  $\{n \theta_j\} \in [0, \delta[ \cup ]1 - \delta, 1[$  is bounded by

$$\begin{aligned} 2R(j)(1 + \delta \theta_j^{-1}) &\leq \left(\frac{2L}{\theta_j^{-1} - 1} + 4\right)(1 + \delta \theta_j^{-1}) \leq 4(L + \theta_j^{-1})(\delta + \theta_j) \\ &\leq 4(L + 2q_{j+1})(\delta + q_{j+1}^{-1}) \leq 20(\delta + q_{j+1}^{-1})L. \quad \square \end{aligned}$$

**Remark 1.** For every  $\delta \in ]0, \frac{1}{2r}[$  and every interval  $[N_1, N_1 + L]$ , we have by a slight extension of Lemma 1.1:

$$(14) \quad \text{Card} \{n \in [N_1, N_1 + L] : d(nq_j \alpha, \mathbb{Z}/r) \leq \delta\} \leq 20r(\delta + q_{j+1}^{-1})L, \text{ if } q_{j+1} \leq 2L.$$

**Lemma 1.2.** *Let  $I = [N_1, N_1 + L]$  be an interval and  $\ell$  such that  $q_\ell \leq 2L$ .*

*a) For all  $\delta \in ]0, \frac{1}{2}[$  and  $\zeta \in ]0, 1[$ , the density of set*

$$(15) \quad A := \{n \in I : \text{Card}(j < \ell : d(nq_j \alpha, \mathbb{Z}) \leq \delta) \leq \zeta \ell\}$$

satisfies

$$(16) \quad \text{Card}(A) \geq (1 - 20 \zeta^{-1} (\delta + C \ell^{-1})) L.$$

b) Under Condition (12) on  $\varphi$ , there are positive constants  $\eta_0, c$  (not depending on  $\delta$ ) such that, for every  $\delta \in ]0, \frac{1}{2}[$ , the subset  $V(I, \delta, \ell) := \{n \in I : \|\varphi_n\|_2 \geq \eta_0 \delta \sqrt{\ell}\}$  satisfies:

$$(17) \quad \text{Card}(V(I, \delta, \ell)) \geq (1 - c(\delta + \ell^{-1})) L.$$

*Proof.* a) Let  $A^c = I \setminus A$  be the complementary of  $A$ . We will find an upper bound of the density  $L^{-1} \text{Card}(A^c)$  by counting the number of values of  $n$  in  $I$  such that  $\|nq_j\alpha\| < \delta$  in an array indexed by  $(j, n)$ .

By summing (13) from  $j = 0$  to  $j = \ell - 1$  and using the definition of  $A$ , we get:

$$\begin{aligned} 20(\delta\ell + \sum_{0 \leq j \leq \ell-1} q_{j+1}^{-1})L &\geq \sum_{0 \leq j \leq \ell-1} \sum_{n \in I} 1_{I_\delta}(\{nq_j\alpha\}) \\ &\geq \sum_{n \in A^c} \sum_{0 \leq j \leq \ell-1} 1_{I_\delta}(\{nq_j\alpha\}) \geq \sum_{n \in A^c} \zeta \ell = \zeta \ell \text{Card}(A^c). \end{aligned}$$

With  $C := \sum_{j=0}^{\infty} q_j^{-1}$ , we have  $\text{Card}(A^c) \leq 20 \zeta^{-1} (\delta + C \ell^{-1}) L$ , so (16) is shown.

b) With  $\zeta = \frac{1}{2}\theta$ , where  $\theta$  is the constant in (12), in view of the definition of  $A$  and (12), we have, for  $n \in A$ :

$$\text{Card}(\{j \leq \ell - 1 : \|nq_j\alpha\| \geq \delta\} \cap \{j : |\gamma_{q_j}(\varphi)| \geq \eta\}) \geq (1 - (\zeta + 1 - \theta)) \ell = \frac{1}{2}\theta \ell.$$

Putting  $c := 20 \zeta^{-1} \max(1, C)$  and  $\eta_0 = (\frac{1}{2}c_0 \eta^2 \theta)^{\frac{1}{2}}$ , this implies by (9) and (16):

$$(18) \quad \|\varphi_n\|_2^2 \geq \frac{1}{2}c_0 \delta^2 \eta^2 \theta \ell = \eta_0^2 \delta^2 \ell, \quad \forall n \in A, \quad \text{and} \quad \text{Card}(A) \geq 1 - c(\delta + \ell^{-1})L;$$

hence  $A \subset V(I, \delta, \ell)$  and therefore  $V(I, \delta, \ell)$  satisfies (17).  $\square$

The constants  $c$  and  $\eta_0$  below are those of Lemma 1.2.

**Theorem 1.3.** *Under Condition (12) on  $\varphi$ , the density of the subset*

$$W := \{n \in \mathbb{N} : \|\varphi_n\|_2 \geq \eta_0 \left(\frac{m(n)}{\ln m(n)}\right)^{\frac{1}{2}}\}$$

satisfies for every  $N \geq 1$ :

$$\frac{\text{Card}(W \cap [0, N[)}{N} \geq 1 - 2c(\ln m(N))^{-\frac{1}{2}}.$$

*Proof.* Since  $t/\ln t$  is increasing for  $t \geq e$ , except for the first terms  $n$ , we have for  $n$  in  $W^c \cap [0, N[$ :

$$\|\varphi_n\|_2 < \eta_0 \left(\frac{m(n)}{\ln m(n)}\right)^{\frac{1}{2}} \leq \eta_0 \left(\frac{m(N)}{\ln m(N)}\right)^{\frac{1}{2}}.$$

Therefore, by Lemma 1.2 with  $\delta = (\ln m(N))^{-\frac{1}{2}}$  and  $\ell = m(N)$ , it follows

$$\frac{\text{Card}(W^c \cap [0, N[)}{N} \leq c(\ln m(N))^{-\frac{1}{2}} + \eta m(N)^{-1} \leq 2c(\ln m(N))^{-\frac{1}{2}}. \quad \square$$



### 1.3. A counter-example.

In the next sections we will show that, under a Diophantine condition on  $\alpha$ , for a big set of  $n$ , the distribution of  $\varphi_n/\|\varphi_n\|_2$  is approximately Gaussian. In particular, by (11), if  $n_\ell \leq q_{\ell+1}$  is an integer such that  $\|\varphi_{n_\ell}\|_2 = \max_{k < q_{\ell+1}} \|\varphi_k\|_2$ , then we have the lower bound  $\|\varphi_{n_\ell}\|_2^2 \geq c \sum_{j=0}^{\ell-1} |\gamma_{q_j}|^2 a_{j+1}^2$ . Under Condition (12), it can be shown that, for these indices  $n_\ell$  giving the record variances, when the partial quotients of  $\alpha$  are bounded, the distribution of  $\varphi_{n_\ell}/\|\varphi_{n_\ell}\|_2$  is asymptotically Gaussian.

Let us show by a counter-exemple that this is not true without a condition on  $\alpha$ .

For a parameter  $\gamma > 0$ , let the sequence  $(a_n)_{n \geq 1}$  be defined by

$$a_n = \lfloor n^\gamma \rfloor \text{ if } n \in \{2^k : k \geq 0\}, = 1 \text{ if } n \notin \{2^k : k \geq 0\}.$$

Let  $\alpha$  be the number which has  $(a_n)_{n \geq 1}$  for sequence of partial quotients. For simplicity, let us take for  $\varphi$  the sawtooth function  $\varphi^0$  defined above for which  $\gamma_k = \frac{-1}{2\pi i}, \forall k$ .

For  $\ell > 1$ , let  $n_\ell := \max\{n < q_{\ell+1} : \|\varphi_n\|_2 = \max_{k < q_{\ell+1}} \|\varphi_k\|_2\}$ . We have

$$\|\varphi_{n_\ell}\|_2^2 \geq c \sum_{j=0}^{\ell-1} |\gamma_{q_j}|^2 a_{j+1}^2 \geq c \sum_{s=0}^{\lfloor \log_2(\ell) \rfloor} 2^{2\gamma s} \geq c\ell^{2\gamma}.$$

In the sum  $\varphi_{n_\ell}(x) = \sum_{k=0}^{\ell} \sum_{j=0}^{b_k q_k - 1} \varphi(x + N_{k-1}\alpha + j\alpha)$  defined in (6), we can isolate the indices  $k$  for which  $k+1$  is a power of 2 (for the other indices  $a_{k+1} = 1$ ) and write  $\varphi_{n_\ell}(x) = U_\ell + V_\ell$  with

$$U_\ell = \sum_{p=1}^{\lfloor \log_2(\ell) \rfloor} \sum_{j=0}^{b_{2^p} q_{2^p} - 1} \varphi(x + N_{2^p-1}\alpha + j\alpha),$$

$$V_\ell = \sum_{k \in [0, \ell] \cap \{a_{k+1}=1\}} \sum_{j=0}^{b_k q_k - 1} \varphi(x + N_{k-1}\alpha + j\alpha).$$

We will see in (65) that the variance of a sum where  $b_k$  equals 0 or 1 is bounded as follows

$$\left\| \sum_{k \in [0, \ell] \cap \{a_{k+1}=1\}} \sum_{j=0}^{b_k q_k - 1} \varphi(x + N_{k-1}\alpha + j\alpha) \right\|_2^2 \leq C\ell \log(\ell).$$

On the other side, we also have

$$\left| \sum_{p=1}^{\lfloor \log_2(\ell) \rfloor} \sum_{j=0}^{b_{2^p-1} q_{2^p-1} - 1} \varphi(x + N_{2^p-2}\alpha + j\alpha) \right| \leq \sum_{p=1}^{\lfloor \log_2(\ell) \rfloor} a_{2^p} V(\varphi) \leq C \sum_{p=1}^{\lfloor \log_2(\ell) \rfloor} 2^{\gamma p} \leq C\ell^\gamma.$$

The previous bounds imply  $\frac{\varphi_{n_\ell}}{\|\varphi_{n_\ell}\|_2} = \frac{U_\ell}{\|\varphi_{n_\ell}\|_2} + \frac{V_\ell}{\|\varphi_{n_\ell}\|_2}$  with  $\|\varphi_{n_\ell}\|_2 \geq c\ell^\gamma$ ,  $\|U_\ell\|_\infty \leq C\ell^\gamma$ ,  $\|V_\ell\|_2 \leq (C\ell \log(\ell))^{1/2}$ .

Thus, if  $\gamma > 1/2$ , one has  $\left\| \frac{U_\ell}{\|\varphi_{n_\ell}\|_2} \right\|_\infty \leq \frac{C}{c}$ ,  $\left\| \frac{V_\ell}{\|\varphi_{n_\ell}\|_2} \right\|_2 \rightarrow 0$  and the limit points of the distributions of  $\frac{\varphi_{n_\ell}}{\|\varphi_{n_\ell}\|_2}$  have all their supports included in  $[-\frac{C}{c}, \frac{C}{c}]$ , hence are not Gaussian.

#### 1.4. A special case: quadratic numbers.

In this subsection we show that, if  $\alpha$  is a quadratic number, the variance  $\|\varphi_n\|_2^2$  of the ergodic sums of  $\varphi$  under the rotation by  $\alpha$  is of order  $\ln n$  for  $n$  in a big set of integers whose size is precisely estimated.

The main step in the proof is the following proposition showing that, for most of the integers  $n$ ,  $\|nq_j\alpha\|$  is far from 0 for a big proportion of  $j$ 's:

**Proposition 1.4.** *If  $\alpha$  is a quadratic number, for every  $\varepsilon_0 \in ]0, \frac{1}{2}[$ , there are  $\delta \in ]0, \frac{1}{2}[$  and positive constants  $C$  and  $\xi$  such that for every  $\ell \geq 1$ :*

$$(19) \quad \text{Card} \{n < q_{\ell+1} : \text{Card} (j < \ell : d(nq_j\alpha, \mathbb{Z}) \geq \delta) \geq (1 - \varepsilon_0)\ell\} \geq (1 - Cq_{\ell+1}^{-\xi})q_{\ell+1}.$$

The proposition implies:

**Theorem 1.5.** *If  $\alpha$  is a quadratic number and if  $\varphi$  satisfies Condition (12), there are positive constants  $\eta_1, \eta_2, R$  and  $\xi \in ]0, 1[$  such that, for  $N$  big enough, it holds :*

$$(20) \quad \text{Card} \{n \leq N : \eta_1 \ln n \leq \|\varphi_n\|_2^2 \leq \eta_2 \ln n\} \geq N(1 - RN^{-\xi}).$$

*Proof.* There is  $\eta_2 > 0$  such that the upper bound in (20) holds for every  $n \geq 1$ : indeed, when  $\alpha$  is quadratic, as  $(q_k)$  is equivalent to a geometric sequence,  $m(n)$  is equivalent to  $\ln(n)$  up to a multiplicative constant factor. Therefore, for  $n \in [q_\ell, q_{\ell+1}[$  (i.e.,  $m(n) = \ell$ ), (10) implies  $\|\varphi_n\|_2^2 \leq CK(\varphi)^2 \sum_{j=0}^{\ell} a_{j+1}^2 \leq \eta_2 \ln(n)$ , for some positive constant  $\eta_2$ .

For the lower bound, by (9) we have  $\|\varphi_n\|_2^2 \geq c_0 \delta^2 \sum_{j=1}^{\ell} |\gamma_{q_j}(\varphi)|^2 a_{j+1}^2 \mathbf{1}_{\|nq_j\alpha\| \geq \delta}$ . Let  $\varphi$  in  $\mathcal{C}$  be such that (12) is satisfied and, for  $\varepsilon_0 = \frac{1}{2}\theta$ , let  $\delta = \delta(\varepsilon_0)$  be given by Proposition 1.4. According to (12) and (19), for  $\ell$  big enough, the set of integers  $n < q_{\ell+1}$  such that simultaneously  $\|nq_j\alpha\| \geq \delta$  and  $|\gamma_{q_j}(\varphi)| \geq \eta$ , for at least  $\frac{1}{2}\theta\ell$  different indices  $j$ , has a cardinal bigger than  $q_{\ell+1}(1 - Cq_{\ell+1}^{-\xi})$  for some constants  $C > 0$ ,  $\xi \in ]0, 1[$ .

Therefore we have  $\|\varphi_n\|_2^2 \geq \frac{c_0}{2}\eta^2 \delta^2 \theta \ell = \eta_1 \ell$  for more than  $q_{\ell+1}(1 - Cq_{\ell+1}^{-\xi})$  values of  $n$  between 1 and  $q_{\ell+1}$ .

This shows that, for  $N \in [q_\ell, q_{\ell+1}[$ , the cardinal of the set  $\{n < N : \|\varphi_n\|_2^2 \leq \eta_1 \ell\}$  is less than  $Cq_{\ell+1}^{1-\xi} \leq C'N^{1-\xi}$  (because for a quadratic number  $\sup_\ell q_{\ell+1}/q_\ell < +\infty$ ). Hence, the result.  $\square$

#### Proof of Proposition 1.4

The proof consists in several steps. To bound from below  $d(nq_j\alpha, \mathbb{Z})$ , successively we code  $n$  as an admissible word (Ostrowski's coding), reduce long words to short words, then interpret cardinals in terms of cylinders and invariant measure for a subshift. Finally we use a result of large deviations recalled in Lemma 1.6.

For the reader's convenience, at each step we will consider first the simpler special case of the golden mean  $\alpha = \frac{\sqrt{5}+1}{2}$  (the corresponding rotation number is  $\frac{\sqrt{5}-1}{2} \in ]0, 1[$ ). Then the general case is treated between the signs “ $\diamond$ ” and “ $\triangle$ ” and may be skipped if  $\alpha$  is the golden mean.

When  $\alpha$  is the golden mean, its partial quotients are equal to 1 and  $(q_n)$  (the Fibonacci sequence with  $q_{-1} = 0, q_0 = 1, q_1 = 1, \dots$ ) is almost a geometric sequence of ratio  $\alpha$ . We have

$$(21) \quad q_n = \frac{1}{5}[(2 + \alpha)\alpha^n + (-1)^n(3 - \alpha)\alpha^{-n}], \quad n \geq 0,$$

$$(22) \quad \alpha^n + (-\alpha)^{-n} \in \mathbb{Z}, \quad d(\alpha^n, \mathbb{Z}) = \alpha^{-n}, \quad n \geq 1.$$

◇ For a general quadratic number  $\alpha$ , the sequence  $(a_n)$  is ultimately periodic: there are integers  $n_0, p$  such that  $a_{n+p} = a_n, \forall n \geq n_0$ .

Let  $A_1 := \begin{pmatrix} 0 & 1 \\ 1 & a_{n_0+1} \end{pmatrix}$ ,  $A_i := \begin{pmatrix} 0 & 1 \\ 1 & a_{n_0+i} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_{n_0+i-1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & a_{n_0+1} \end{pmatrix}$ , for  $i > 1$ .

From the recursive relation  $\begin{pmatrix} q_n \\ q_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & a_{n+1} \end{pmatrix} \begin{pmatrix} q_{n-1} \\ q_n \end{pmatrix}$ , between the denominators  $(q_n)$  of  $\alpha$ , it follows,  $\forall k \geq 1$ ,  $\begin{pmatrix} q_{n_0+kp+m-1} \\ q_{n_0+kp+m} \end{pmatrix} = A_m A_p^k \begin{pmatrix} q_{n_0-1} \\ q_{n_0} \end{pmatrix}$ ,  $m = 1, \dots, p$ .

The matrix  $A_p$  is a  $2 \times 2$  matrix with determinant  $(-1)^p$  and non negative integer coefficients (positive if  $p > 1$ ). It has two distinct eigenvalues  $\lambda > 1$  and  $-\lambda^{-1}$  (where  $\lambda$  is a quadratic number) and it is diagonal in a basis of  $\mathbb{R}^2$  with coordinates in  $\mathbb{Q}[\lambda]$ . We have  $\lambda^p + (-\lambda^{-1})^p \in \mathbb{Z}$ .

Without loss of generality we may suppose that  $p$  is even (otherwise, we replace it by  $2p$ ). Therefore there are integers  $r, s_\ell, t_\ell, u_\ell, v_\ell$  for  $\ell \in \{0, \dots, p-1\}$  such that

$$(23) \quad q_{n_0+kp+\ell} = \frac{1}{r} [(s_\ell + t_\ell \lambda) \lambda^k + (u_\ell + v_\ell \lambda) \lambda^{-k}], \quad \forall k \geq 0.$$

For every  $\ell$ ,  $(q_{n_0+kp+\ell})_{k \geq 1}$  behaves like a geometric progression with ratio  $\lambda$ .

For the golden mean, (23) corresponds to (21) for  $n$  even and  $r = 5$ .  $\triangle$

### 1) Ostrowski's coding, invariant measure for a subshift of finite type and counting

As recalled in Subsection 1.1, every  $n < q_{m+1}$  is coded by an "admissible" word  $b_0 \dots b_m$ , with  $b_0 \in \{0, 1, \dots, a_1 - 1\}$ ,  $b_j \in \{0, 1, \dots, a_{j+1}\}$ ,  $j = 1, \dots, m$ , where for two consecutive letters  $b_{j-1}, b_j$ , if  $b_j = a_{j+1}$ , then  $b_{j-1} = 0$ .

For  $\alpha$  the golden mean, a finite word  $b_0 \dots b_m$  is admissible if it is composed of 0's and 1's and two consecutive letters  $b_i, b_{i+1}$  cannot be both 1. We denote by  $X$  the space of infinite admissible words, that is sequence of 0, 1 without two consecutive 1's. For simplicity the letter  $b$  will denote an admissible word, either finite or infinite. The context will make clear if  $b$  is finite or not.

If  $b = b_0 \dots b_\ell$  is an admissible word, we put  $n_{b,\ell} := \sum_{i=0}^{\ell} b_i q_i$ .

When  $\alpha$  is the golden mean, we use the sub-shift  $(X, \sigma)$ , where  $\sigma = \sigma_X$  is the shift on  $X$ . Let  $\mu$  be the  $\sigma$ -invariant probability measure on  $X$  of maximal entropy. Let  $C_{x_0 \dots x_n}$  denote the cylinder composed of sequences starting with  $x_0 \dots x_n$ . For  $n \geq 1$ , depending

whether  $x_0$  et  $x_n$  are both equal to 1, or only one of them, or none, we have

$$\mu(C_{x_0\dots x_n}) = \frac{1}{\alpha + 2}\alpha^{-n}, \frac{\alpha}{\alpha + 2}\alpha^{-n} \text{ or } \frac{\alpha^2}{\alpha + 2}\alpha^{-n}.$$

If  $E \subset X$  is a union of cylinders of length  $n$ , its measure can be compared to the number of cylinders which compose it:

$$(24) \quad \frac{\alpha + 2}{\alpha + 1}\mu(E) \leq \alpha^{-n} \text{Card} \{\text{cylinder } \mathcal{W} \text{ of length } n : \mathcal{W} \subset E\} \leq (\alpha + 2)\mu(E).$$

◇ In the general case, let us consider the set of infinite admissible sequences corresponding to the Ostrowski expansions for the periodic part of the sequence  $(a_n)$ :

$$X := \{x = (x_i)_{i \in \mathbb{N}} \text{ such that } \forall i \ x_{i-1}x_i \neq ua_{n_0+i+1} \text{ with } u \neq 0\}.$$

The space  $X$  is invariant under the action of  $\sigma_X^p$  (because  $(a_n)$  is  $p$ -periodic for  $n \geq n_0$ ). We define an irreducible aperiodic sub-shift of finite type as follows: the state space of  $Y$  is the set of words  $x_0 \dots x_{p-1}$  of  $X$ , a transition between two such words  $w_1$  and  $w_2$  is allowed if the concatenation  $w_1w_2$  is the beginning of length  $2p$  of a sequence in  $X$ .

From (23) we see that the exponential growth rate of the number of Ostrowski expansions of length at most  $n_0 + pk$  is  $\ln \lambda$  (with respect to  $k$ ). It is also the growth rate of the number of words of length  $pk$  of  $X$ . As these words correspond to the words of length  $k$  in  $Y$ , the topological entropy of  $(Y, \sigma)$  is  $\ln \lambda$  (where  $\sigma = \sigma_Y$  is the shift to the left on  $Y$ ). There is a unique invariant probability measure  $\mu$  on  $(Y, \sigma)$  with entropy  $\ln \lambda$ . This measure can be constructed as follows. Let  $B$  be the matrix with entries 0 and 1 that gives the allowed transitions between elements of the alphabet of  $Y$ . As the topological entropy of  $Y$  is the logarithm of the spectral radius of  $B$ , this spectral radius is  $\lambda$ . Let  $U$  and  $V$  be two positive vectors such that  $BU = \lambda U$ ,  ${}^tBV = \lambda V$ ,  ${}^tUV = 1$ . The measure  $\mu$  is the Markovian measure determined by its values on cylinders given by

$$\mu(C_{y_0y_1\dots y_n}) = V_{y_0}U_{y_n}\lambda^{-n},$$

when  $y_0y_1\dots y_n$  is an admissible word (see [10] pp.21-23 and p.166 for more details on this classical construction). As there are only finitely many products  $V_{y_0}U_{y_n}$ , there exists a constant  $c' > 0$  such that, if a subset  $E$  of  $Y$  is a union of cylinders of length  $n$ , then

$$(25) \quad \frac{1}{c'}\mu(E) \leq \text{Card} \{\mathcal{W} \text{ cylinders of length } n : \mathcal{W} \subset E\} \lambda^{-n} \leq c'\mu(E). \quad \triangle$$

#### *Lemma of large deviations*

We will use the following inequality of large deviations for irreducible Markov chains with finite state space (see [11], Theorem 3.3):

**Lemma 1.6.** *Let  $A$  be finite union of cylinders. For every  $\varepsilon \in ]0, 1[$ , there are two positive constants  $R(\varepsilon)$ ,  $\xi(\varepsilon)$  depending on  $A$  such that*

$$(26) \quad \mu\{x \in X : \frac{1}{L} \sum_{k=0}^{L-1} \mathbf{1}_A(\sigma^k x) \leq \mu(A)(1 - \varepsilon)\} \leq R(\varepsilon) e^{-\xi(\varepsilon)L}, \quad \forall L \geq 1.$$

2) *Reduction of the Ostrowski expansion to a “window”*

By (2) and (4) we have, for a constant  $\rho < 1$ ,  $\|q_i q_j \alpha\| \leq C \rho^{|j-i|}$ . Hence, for  $0 \leq j \leq \ell$ , if  $\kappa$  is such that  $0 \leq j - \kappa \leq j + \kappa \leq \ell$ :

$$(27) \quad \left\| \sum_{i=0}^{\ell} b_i q_i q_j \alpha \right\| - \left\| \sum_{i=j-\kappa}^{j+\kappa-1} b_i q_i q_j \alpha \right\| \leq \left\| \sum_{i=j+\kappa}^{\ell} b_i q_i q_j \alpha \right\| + \left\| \sum_{i=0}^{j-\kappa-1} b_i q_i q_j \alpha \right\| \leq C \rho^{-\kappa}.$$

It means that  $\left\| \sum_{i=0}^{\ell} b_i q_i q_j \alpha \right\| = \|n_{b,\ell} q_j \alpha\|$  is well approximated by  $\left\| \sum_{i=j-\kappa}^{j+\kappa-1} b_i q_i q_j \alpha \right\|$  which depends on a word with indices belonging to a window around  $j$ , with a precision depending on the size of the window. This is valid for any irrational  $\alpha$ .

The quantity introduced in the next definition can be viewed as a function of an infinite word  $b$  or of a finite word  $b_{j-\kappa_0}, \dots, b_{j+\kappa_0}$ . We put

$$(28) \quad \Gamma(b, j) := \frac{1}{5} \sum_{i=j-\kappa_0}^{j+\kappa_0} (-1)^i b_i (\alpha^{j-i} + (-\alpha)^{i-j}) \alpha.$$

A simple computation shows that  $\Gamma(b, j+1) = -\Gamma(\sigma b, j)$ . Therefore we have:

$$(29) \quad \Gamma(\sigma^k b, \kappa_0) = (-1)^k \Gamma(b, k + \kappa_0).$$

**Lemma 1.7.** *Let  $\alpha$  be the golden mean. For every  $\delta > 0$ , there is  $\kappa_0 = \kappa_0(\delta)$  such that*

$$(30) \quad d(n_{b,\ell} q_j \alpha - \Gamma(b, j), \mathbb{Z}/5) \leq \delta, \text{ if } j \geq \kappa_0.$$

*Proof.* We can restrict the sum  $n_{b,\ell} q_j \alpha = \sum_{i=0}^{\ell} b_i q_i q_j \alpha$  to the sum  $\sum_{i=j-\kappa_0}^{j+\kappa_0} b_i q_i q_j \alpha$ , since their distance modulo 1 is small for  $\kappa_0$  big enough by (27).

By (21), we have  $q_i q_j = \frac{1+\alpha}{5} \alpha^{i+j} + \frac{2-\alpha}{5} (-\alpha)^{-(i+j)} + \frac{(-1)^i}{5} (\alpha^{j-i} + (-\alpha)^{i-j})$ ;

$$\text{hence: } \sum_{i=j-\kappa_0}^{j+\kappa_0} b_i q_i q_j \alpha = \frac{1}{5} \sum_{i=j-\kappa_0}^{j+\kappa_0} [b_i (1+\alpha) \alpha^{i+j+1} + b_i (-1)^{i+j} (2-\alpha) \alpha^{1-(i+j)}] + \Gamma(b, j).$$

The distance to  $\mathbb{Z}$  of the first sum above at right is small by (22).  $\square$

The lemma shows that for the golden mean the distance to  $\mathbb{Z}/5$  of  $\sum_{j=0}^{\ell} b_i q_i q_j \alpha$  is almost the distance to  $\mathbb{Z}/5$  of  $\Gamma(b, j)$ , which depends on the “short” word  $b_{j-\kappa_0} \dots b_{j+\kappa_0}$  (reduction to a window of width  $2\kappa_0$  of the “long” word  $b_0 \dots b_{\ell}$ ) in such a way that its values, when  $j$  varies, are the values of a fixed function computed for shifted words.

$\diamond$  The lemma extends to a general quadratic number. We need some notation.

For an integer  $i$ , we write  $i = \underline{i} + p \eta_i + n_0$ , where  $\underline{i}$  is the class of  $i - n_0$  modulo  $p$  and  $\eta_i$  the integer part of  $(i - n_0)/p$ . The classes mod  $p$  are identified with the integers  $0, \dots, p-1$ . With the notation introduced in (23), we put

$$T(\underline{i}, \underline{j}) := \frac{\alpha}{r^2} (s_{\underline{i}} + t_{\underline{i}} \lambda) (u_{\underline{j}} + v_{\underline{j}} \lambda), \quad U(\underline{i}, \underline{j}) := \frac{\alpha}{r^2} (u_{\underline{i}} + v_{\underline{i}} \lambda) (s_{\underline{j}} + t_{\underline{j}} \lambda).$$

**Lemma 1.8.** *Let  $\delta \in ]0, \frac{1}{2r}[$ . There is  $\kappa_0 = \kappa_0(\delta)$  such that, if  $j \geq n_0 + \kappa_0 p$ ,*

$$(31) \quad d(n_{b,\ell} q_j \alpha - \sum_{i=n_0+(\eta_j-\kappa_0)p}^{n_0+(\eta_j+\kappa_0)p-1} b_i [T(\underline{i}, \underline{j}) \lambda^{\eta_i-\eta_j} + U(\underline{i}, \underline{j}) \lambda^{\eta_j-\eta_i}], \mathbb{Z}/r) \leq \delta.$$

*Proof.* Recall that  $(a_n)$  is  $p$ -periodic for  $n \geq n_0$ . We consider indices  $j \geq n_0$  and take sums on windows union of blocks of length  $p$ , hence of the form  $n_0 + mp, \dots, n_0 + qp - 1$ . Using (23), the product  $q_{n_0+kp+m} q_{n_0+k'p+m'}$  is equal to

$$\begin{aligned} & \frac{1}{r^2} \left[ (s_m + t_m \lambda)(s_{m'} + t_{m'} \lambda) \lambda^{k+k'} + (u_m + v_m \lambda)(u_{m'} + v_{m'} \lambda) \lambda^{-(k+k')} \right] \\ & + \frac{1}{r^2} \left[ (s_m + t_m \lambda)(u_{m'} + v_{m'} \lambda) \lambda^{k-k'} + (u_m + v_m \lambda)(s_{m'} + t_{m'} \lambda) \lambda^{k'-k} \right]. \end{aligned}$$

Still using (23), we have

$$\frac{s_m}{r} \frac{(s_{m'} + t_{m'} \lambda)}{r} \lambda^{k+k'} = \frac{s_m}{r} (q_{n_0+(k'+k)p+m'} - \frac{1}{r} (u_{m'} + v_{m'} \lambda) \lambda^{-(k'+k)}).$$

From this (and a similar equality) we obtain

$$\begin{aligned} & q_{n_0+kp+m} q_{n_0+k'p+m'} \alpha \\ & - \frac{1}{r^2} \left[ (s_m + t_m \lambda)(u_{m'} + v_{m'} \lambda) \lambda^{k-k'} + (u_m + v_m \lambda)(s_{m'} + t_{m'} \lambda) \lambda^{k'-k} \right] \alpha \\ & = \frac{s_m}{r} q_{n_0+(k'+k)p+m'} \alpha + \frac{t_m}{r} q_{n_0+(k'+k+1)p+m'} \alpha \\ & - \left[ \frac{s_m}{r^2} (u_{m'} + \frac{v_{m'}}{r} \lambda) \lambda^{-(k'+k)} \alpha + \frac{t_m}{r^2} (u_{m'} + v_{m'} \lambda) \lambda^{-(k'+k+1)} \alpha \right]. \end{aligned}$$

Since  $d(q_{n_0+(k'+k)p+m'} \alpha, \mathbb{Z}/r) \leq d(q_{n_0+(k'+k)p+m'} \alpha, \mathbb{Z}) \leq C \lambda^{-(k'+k)}$  by (2), the distance of the left side term above to  $\mathbb{Z}/r$  is bounded by  $C \lambda^{-(k'+k)}$ . It follows:

$$d(q_i q_j \alpha - [T(\underline{i}, \underline{j}) \lambda^{\eta_i-\eta_j} + U(\underline{i}, \underline{j}) \lambda^{\eta_j-\eta_i}], \mathbb{Z}/r) \leq C \lambda^{-p(\eta_j+\eta_i)}, \forall i, j \geq n_0.$$

Thus, using (27), for  $\kappa_0$  large enough and if  $j \geq n_0 + \kappa_0 p$ , we have:

$$\begin{aligned} & d(n_{b,\ell} q_j \alpha - \sum_{i=n_0+(\eta_j-\kappa_0)p}^{n_0+(\eta_j+\kappa_0)p-1} b_i [T(\underline{i}, \underline{j}) \lambda^{\eta_i-\eta_j} + U(\underline{i}, \underline{j}) \lambda^{\eta_j-\eta_i}], \mathbb{Z}/r) \\ & \leq C \lambda^{-\kappa_0 p} + C \sum_{i=n_0+(\eta_j-\kappa_0)p}^{n_0+(\eta_j+\kappa_0)p-1} \lambda^{-\eta_j-\eta_i} \leq C \lambda^{-\kappa_0 p} + 2C \kappa_0 \lambda^{\kappa_0-2\eta_j} \leq \delta. \quad \square \end{aligned}$$

△

3) From long words to short words

**Lemma 1.9.** *Let  $1 \leq \ell_0 \leq \ell_1 \leq \ell$  be three integers and let  $\Lambda : b_0 \dots b_\ell \mapsto b_{\ell_0} \dots b_{\ell_1}$  be the “restriction” map from the set  $\mathcal{L}$  of admissible words to shortened words. There is a constant  $c > 0$  such that, if  $\mathcal{S}$  is the image of  $\Lambda$ , for any subset  $\mathcal{P}$  of  $\mathcal{S}$ , we have*

$$\frac{\text{Card}(\mathcal{P})}{\text{Card}(\mathcal{S})} \leq c \frac{\text{Card}\{w \in \mathcal{L} : \Lambda(w) \in \mathcal{P}\}}{\text{Card}(\mathcal{L})}.$$

We can take  $c = 4$  for the golden mean,  $c = u_0^{-2}$ , with  $u_0 = \inf_{k>1} \frac{q_{k-1}-1}{q_k-1}$ , in the general case.

*Proof.* The proof is given for the golden mean. The general case is analogous.

The ways of completing a short word into a long one depend only on the first letter  $b_{\ell_0}$  and the last letter  $b_{\ell_1}$ : if  $b_{\ell_0} \neq 1$ , any admissible beginning fits; if  $b_{\ell_0} = 1$ , then only the admissible beginnings finishing by 0 fit; if  $b_{\ell_1} = 0$  then any admissible ending fits; if  $b_{\ell_1} = 1$ , only endings with 0 as first letter fit.

The number of admissible words of length  $r$  is  $q_{r+1}$ , the number of admissible words of length  $r$  beginning (or ending) by 0 is  $q_r$ .

Let denote  $\mathcal{S}_i$ ,  $i = 1, \dots, 4$ , the set of short words  $b_{\ell_0} \dots b_{\ell_1}$  such that  $b_{\ell_0} = b_{\ell_1} = 0$ ,  $b_{\ell_0} = 0$  and  $b_{\ell_1} = 1$ ,  $b_{\ell_0} = 1$  and  $b_{\ell_1} = 0$ ,  $b_{\ell_0} = b_{\ell_1} = 1$ , respectively.

Depending on the set  $\mathcal{S}_i$  to which  $\Lambda(w)$  belongs, the cardinal of  $\text{Card} \Lambda^{-1}(\Lambda(w))$  is  $D_1 = q_{\ell_0} q_{\ell-\ell_1+1}$ ,  $D_2 = q_{\ell_0} q_{\ell-\ell_1}$ ,  $D_3 = q_{\ell_0-1} q_{\ell-\ell_1+1}$ , or  $D_4 = q_{\ell_0-1} q_{\ell-\ell_1}$  respectively.

Since,  $\frac{1}{2} \leq q_r/q_{r+1} \leq 1$ , for all  $r$ , we have  $D_1 = \max_i D_i$ ,  $D_4 = \min_i D_i$ ,  $D_4 \leq D_1 \leq 4D_4$  and finally

$$\begin{aligned} \text{Card}(\mathcal{P}) &= \sum_{i=1}^4 \text{Card}(\mathcal{P} \cap \mathcal{S}_i) = \sum_{i=1}^4 \frac{1}{D_i} \text{Card}\{w \in \mathcal{L} : \Lambda(w) \in \mathcal{P} \cap \mathcal{S}_i\} \\ &\leq \frac{1}{D_4} \sum_{i=1}^4 \text{Card}\{w \in \mathcal{L} : \Lambda(w) \in \mathcal{P} \cap \mathcal{S}_i\} = \frac{1}{D_4} \text{Card}\{w \in \mathcal{L} : \Lambda(w) \in \mathcal{P}\}, \\ \text{Card}(\mathcal{S}) &= \sum_{i=1}^4 \text{Card}(\mathcal{S}_i) = \sum_{i=1}^4 \frac{1}{D_i} \text{Card}\{w \in \mathcal{L} : \Lambda(w) \in \mathcal{S}_i\} \geq \frac{1}{D_1} \text{Card}(\mathcal{L}). \quad \square \end{aligned}$$

4a) End of the proof of Proposition 1.4 when  $\alpha$  is the golden mean

Let  $\delta$  be a small positive number. Its value will be chosen later. It follows from (14) for  $\ell$  big enough that, if  $q_{j+1}^{-1} < \delta$ :

$$(32) \quad \text{Card}\{n \in [1, q_{\ell+1}[ : d(nq_j\alpha, \mathbb{Z}/5) \leq 3\delta\} \leq C_1 \delta q_{\ell+1}, \forall j \leq \ell.$$

If  $\kappa_0$  is big enough, from (30) in Lemma 1.7, we have with  $\Gamma(b, j)$  defined in (28):

$$(33) \quad d(n_{b,\ell} q_j \alpha, \mathbb{Z}/5) \geq 3\delta \Rightarrow d(\Gamma(b, j), \mathbb{Z}/5) \geq 2\delta \Rightarrow d(n_{b,\ell} q_j \alpha, \mathbb{Z}/5) \geq \delta.$$

By taking  $\kappa_0$  large enough, we can suppose  $q_{\kappa_0+1}^{-1} < \delta$ . By (32) (translated in terms of words) for each  $j \in [\kappa_0, \ell]$ , the proportion of words  $b = b_0 \dots b_\ell$  of length  $\ell + 1$  for which

$d(n_{b,\ell}q_j\alpha, \mathbb{Z}/5) \geq 3\delta$ , is smaller than  $C_1\delta$ . Therefore, if  $\ell \geq \kappa_0$ , we get

$$(34) \quad \text{Card}\{b_0 \dots b_\ell : d(\Gamma(b, j), \mathbb{Z}/5) \leq 2\delta\} \leq C_1 \delta q_\ell, \forall j \in [\kappa_0, \ell].$$

But  $\Gamma(b, j)$  depends only on the short word  $b_{j-\kappa_0} \dots b_{j+\kappa_0}$ , part of the long word  $b = b_0 \dots b_\ell$ . It follows, using Lemma 1.9 that

$$(35) \quad \text{Card}\{b_{j-\kappa_0} \dots b_{j+\kappa_0} : d(\Gamma(b, j), \mathbb{Z}/5) \leq 2\delta\} \leq C_2 \delta q_{2\kappa_0+2}, \forall j \in [\kappa_0, \ell].$$

Putting  $A_\delta := \{b : d(\Gamma(b, \kappa_0), \mathbb{Z}/5) \geq 2\delta\}$ , it follows from (35) and (24):

$$\mu(A_\delta^c) \leq \alpha^{-2\kappa_0-2} \text{Card}\{b_{j-\kappa_0} \dots b_{j+\kappa_0} : d(\Gamma(b, j), \mathbb{Z}/5) \leq 2\delta\} \leq C_2 \delta q_{2\kappa_0+2} \alpha^{-2\kappa_0-2} \leq C_3 \delta.$$

Let  $C_4$  be a constant  $> C_3$  and  $\varepsilon = C_4\delta$ . Observe that we can chose  $\ell$  large enough so that  $\mu(A_\delta)(\ell - \kappa_0) \geq (1 - \varepsilon)\ell$ : indeed, we have  $\mu(A_\delta) - (1 - \varepsilon) > 0$  and by taking  $\ell > \mu(A_\delta)\kappa_0/(\mu(A_\delta) - (1 - \varepsilon))$  we obtain the required inequality.

Now we use  $\sum_{k=0}^{L-1} \mathbf{1}_{A_\delta}(\sigma^k b) = \text{Card}\{k < L : d(\Gamma(\sigma^k b, \kappa_0), \mathbb{Z}/5) \geq 2\delta\}$  and (29). According (33) with  $j = k + \kappa_0$  and Lemma 1.6 with  $A = A_\delta$  and  $\varepsilon = C_4\delta$  (we assume  $\delta < C_4^{-1}$ ), there are two positive constants  $R = R(\varepsilon)$ ,  $\xi = \xi(\varepsilon)$  such that

$$\mu\{b \in X : \text{Card}\{j \in [\kappa_0, L + \kappa_0[: d(\Gamma(b, j), \mathbb{Z}/5) \geq 2\delta\} \leq \mu(A_\delta)(1 - \varepsilon)L\} \leq R(\varepsilon) e^{-\xi L}.$$

Using “ $\Rightarrow$ ” in (33), we have therefore, taking  $L = \ell - \kappa_0$ , for  $\ell - \kappa_0 \geq j \geq \kappa_0$ ,

$$\mu\{b \in X : \text{Card}\{j \in [\kappa_0, \ell[: d(n_{b,\ell}q_j\alpha, \mathbb{Z}/5) \geq \delta\} \leq \mu(A_\delta)(1 - \varepsilon)(\ell - \kappa_0)\} \leq R e^{-\xi(\ell - \kappa_0)}.$$

By (24), the previous inequality translated in terms of cardinal yields for a constant  $C_5$ :

$$\text{Card}\{b_0 \dots b_\ell : \text{Card}\{j < \ell : d(n_{b,\ell}q_j\alpha, \mathbb{Z}/5) \geq \delta\} \leq (1 - C_4\delta)^2 \ell\} \leq C_5 e^{-\xi\ell} q_{\ell+1}.$$

If  $\delta$  is taken small enough to get  $(1 - \varepsilon_0) \leq (1 - C_4\delta)^2$  and using that  $e^{\xi\ell}$  is equivalent to a power of  $q_{\ell+1}$  (because  $(q_\ell)_\ell$  is equivalent to a geometric progression), the previous inequality shows (19) of Proposition 1.4.  $\square$

$\diamond$  4b) *End of the proof of Proposition 1.4 for a general quadratic number*

As for the golden number, we take a positive number  $\delta$  whose value will be fixed later. By (14), if  $\kappa_0$  is large enough, we have for some  $C_1 > 0$

$$\text{Card}\{n \in [1, q_{\ell+1}[: d(nq_j\alpha, \mathbb{Z}/r) \leq 3\delta\} \leq C_1 \delta q_{\ell+1}, \forall j \in [n_0 + \kappa_0, \ell];$$

hence, in terms of admissible words  $b_0 \dots b_\ell$ , if  $j \in [n_0 + \kappa_0, \ell]$ ,

$$(36) \quad \text{Card}\{b_0 \dots b_\ell : d(n_{b,\ell}q_j\alpha, \mathbb{Z}/r) \leq 3\delta\} \leq C_1 \delta q_{\ell+1}.$$

Let  $\Gamma_j, \Gamma_j^0$  be the functions on  $Y$

$$\begin{aligned} \Gamma_j(b) &:= \sum_{i=n_0+(\eta_j-\kappa_0)p}^{n_0+(\eta_j+\kappa_0)p-1} b_i [T(\underline{i}, \underline{j})\lambda^{\eta_i-\eta_j} + U(\underline{i}, \underline{j})\lambda^{\eta_i-\eta_j}], \\ \Gamma_j^0(b) &:= \sum_{i=n_0}^{n_0+2\kappa_0p-1} b_i [T(\underline{i}, \underline{j})\lambda^{\eta_i-\kappa_0} + U(\underline{i}, \underline{j})\lambda^{\kappa_0-\eta_i}]. \end{aligned}$$



Remark that the sums on the right can be viewed as functions of  $y$  through the  $b_i$ 's. Letting  $y_k := b_{n_0+kp} \cdots b_{n_0+kp+p-1}$ , we see that the sum inside the definition of  $\Gamma_j$  is a function of  $y_{\eta_j - \kappa_0} \cdots y_{\eta_j + \kappa_0 - 1}$ .

Let  $A_\delta$  be the subset of  $Y$  defined by

$$A_\delta := \{y : d(\Gamma_{\underline{j}}^0(y), \mathbb{Z}/r) \geq 2\delta, \text{ for } \underline{j} = 0, \dots, p-1\}.$$

By (31) in Lemma 1.8, if  $\kappa_0$  is sufficiently large, we have the implication

$$(37) \quad d(n_{b,\ell} q_j \alpha, \mathbb{Z}/r) \geq 3\delta \Rightarrow d(\Gamma_j(b), \mathbb{Z}/r) \geq 2\delta.$$

As  $\underline{j} = \underline{j-p}$ ,  $\eta_{j+p} = \eta_j + 1$  and  $\eta_{i+p} = \eta_i + 1$ , we obtain by  $\eta_j - \kappa_0$  iterations:

$$\begin{aligned} & \sum_{i=n_0+(\eta_j-\kappa_0)p}^{n_0+(\eta_j+\kappa_0)p-1} b_i T(\underline{i}, \underline{j}) \lambda^{\eta_i - \eta_j} = \sum_{i=n_0+(\eta_j-1-\kappa_0)p}^{n_0+(\eta_j-1+\kappa_0)p-1} b_{i+p} T(\underline{i+p}, \underline{j}) \lambda^{\eta_{i+p} - \eta_j} \\ & = \sum_{i=n_0+(\eta_{j-p}-\kappa_0)p}^{n_0+(\eta_{j-p}+\kappa_0)p-1} b_{i+p} T(\underline{i}, \underline{j}) \lambda^{\eta_{i+1} - \eta_j} = \sum_{i=n_0+(\eta_{j-p}-\kappa_0)p}^{n_0+(\eta_{j-p}+\kappa_0)p-1} b_{i+p} T(\underline{i}, \underline{j}) \lambda^{\eta_i - \eta_{j-p}} \\ & = \dots = \sum_{i=n_0+(\eta_{j-(\eta_j-\kappa_0)p-\kappa_0}p-\kappa_0)p}^{n_0+\eta_{j-(\eta_j-\kappa_0)p-\kappa_0}p-1} b_{i+(\eta_j-\kappa_0)p} T(\underline{i}, \underline{j}) \lambda^{\eta_i - \eta_{j-(\eta_j-\kappa_0)p}}. \end{aligned}$$

Since  $\eta_{j-(\eta_j-\kappa_0)p} = \kappa_0$ , the last quantity reduces to  $\sum_{i=n_0}^{n_0+2\kappa_0p-1} b_{i+(\eta_j-\kappa_0)p} T(\underline{i}, \underline{j}) \lambda^{\eta_i - \kappa_0}$ .

The same computation can be done for  $\sum_{i=n_0+(\eta_j-\kappa_0)p}^{n_0+(\eta_j+\kappa_0)p-1} b_i U(\underline{i}, \underline{j}) \lambda^{\eta_i - \eta_j}$ . Taking the sum for the  $T$ 's and  $U$ 's, we get

$$(38) \quad \Gamma_j(y) = \Gamma_{\underline{j}}^0(\sigma^{\eta_j - \kappa_0} y).$$

From (37), (36) and (38), it follows, if  $\ell \geq n_0 + 2\kappa_0 p$  and  $j \geq n_0 + \kappa_0$ ,

$$\text{Card}\{b_0 \dots b_\ell : d(\Gamma_{\underline{j}}(\sigma^{\eta_j - \kappa_0} y), \mathbb{Z}/r) \leq 2\delta\} \leq C_1 \delta q_{\ell+1}.$$

But  $\Gamma_{\underline{j}}^0(\sigma^{\eta_j - \kappa_0} y)$  depends only on the short word  $b_{n_0+(\eta_j-\kappa_0)p} \cdots b_{n_0+(\eta_j+\kappa_0)p-1}$ , which is a sub-word of the ‘‘long’’ word  $b_0 \dots b_\ell$ . By Lemma 1.9 we obtain for constants  $C_2, C_3 > 0$ :

$$(39) \quad \text{Card}\{b_{n_0+(\eta_j-\kappa_0)p} \cdots b_{n_0+(\eta_j+\kappa_0)p-1} : d(\Gamma_{\underline{j}}^0(\sigma^{\eta_j - \kappa_0} y), \mathbb{Z}/r) \leq 2\delta\} \leq C_2 \delta \lambda^{2\kappa_0}.$$

Then, (39) and (25) imply that

$$(40) \quad \mu(A_\delta^c) = \mu\{y \in Y : d(\Gamma_m^0(y), \mathbb{Z}/r) < 2\delta, m = 0, \dots, p-1\} \leq C_3 \delta.$$

Now, we have

$$\sum_{k=0}^{n-1} \mathbf{1}_{A_\delta}(\sigma^k y) = \text{Card}\{k < n : d(\Gamma_m^0(\sigma^k y), \mathbb{Z}/r) \geq 2\delta, m = 0, \dots, p-1\},$$

$$\Gamma_m^0(\sigma^k y) = \sum_{i=n_0+kp}^{n_0+(k+2\kappa_0)p-1} b_i [T(\underline{i}, m)\lambda^{\eta_i-k-\kappa_0} + U(\underline{i}, m)\lambda^{k+\kappa_0-\eta_i}],$$

and, if  $j = (k + \kappa_0)p + m \in [n_0 + \kappa_0, \ell]$  (i.e.,  $\eta_j = k + \kappa_0$ ,  $\underline{j} = m$ ),

$$d(\Gamma_m^0(\sigma^k y), \mathbb{Z}/r) \geq 2\delta \Rightarrow d(n_{b,\ell}q_j\alpha, \mathbb{Z}/r) \geq \delta.$$

In particular:

$$\begin{aligned} & p \text{Card}\{k < \eta_\ell - \kappa_0 : d(\Gamma_m^0(\sigma^k y), \mathbb{Z}/r) \geq 2\delta, m = 0, \dots, p-1\} \\ & \leq \text{Card}\{j < (\eta_\ell - \kappa_0)p : d(n_{b,\ell}q_j\alpha, \mathbb{Z}/r) \geq \delta\}. \end{aligned}$$

By Lemma 1.6, for the Markov chain deduced from  $Y$  with state space the set of words of length  $2\kappa_0$  in  $Y$ , we get from (26):

$$\mu\{y : \text{Card}\{j < (\eta_\ell - \kappa_0)p : d(n_{b,\ell}q_j\alpha, \mathbb{Z}/r) \geq \delta\} \leq \mu(A_\delta)(1-\varepsilon)p(\eta_\ell - \kappa_0)\} \leq Re^{-\xi(\eta_\ell - \kappa_0)}.$$

This can be translated in terms of cardinal using (25):

$$\text{Card}\{y_0 \dots y_{\eta_\ell - \kappa_0} : \text{Card}\{j < (\eta_\ell - \kappa_0)p : d(n_{b,\ell}q_j\alpha, \mathbb{Z}/r) \geq \delta\} \leq \mu(A_\delta)(1-\varepsilon)p(\eta_\ell - \kappa_0)\}$$

is smaller than  $C_4 e^{-\xi(\eta_\ell - \kappa_0)} \lambda^{\eta_\ell - \kappa_0}$ . It implies

$$\text{Card}\{b_0 \dots b_\ell : \text{Card}\{j < \ell : d(n_{b,\ell}q_j\alpha, \mathbb{Z}/r) \geq \delta\} \leq \mu(A_\delta)(1-\varepsilon)p(\eta_\ell - \kappa_0)\} \leq C_5 e^{-\xi(\eta_\ell - \kappa_0)} \lambda^{\eta_\ell - \kappa_0}.$$

If  $\eta_\ell > (\kappa_0 + 1)/\varepsilon$  (that is  $\ell \geq p(\kappa_0 + 2)/\varepsilon$ ), then  $p(\eta_\ell - \kappa_0) \geq (1 - \varepsilon)\ell$  and, for some  $C_6 > 0$ , there are less than  $C_6 e^{-\xi\eta_\ell} \lambda^{\eta_\ell}$  words  $b$  of length  $\ell$  such that

$$(41) \quad \text{Card}\{j \leq \ell : d(\sum_{i=0}^{\ell} b_i q_i q_j \alpha, \mathbb{Z}/r) \geq \delta\} \leq \mu(A_\delta)(1-\varepsilon)^2 \ell.$$

By (40), for  $\varepsilon_0 > 0$ , we can choose  $\varepsilon$  and  $\delta$  such that  $\mu(A_\delta)(1-\varepsilon)^2 = 1 - \varepsilon_0$ . On the other hand, since  $c^{-1}q_{\ell+1} \leq \lambda^{\eta_\ell} \leq cq_{\ell+1}$  for some  $c > 0$ ,  $C_6 e^{-\xi\eta_\ell} \lambda^{\eta_\ell} \leq C_7 q_{\ell+1}^{1-\zeta}$ , for some positive constants  $\zeta, C_7$ . Finally, we have obtained (19) (in terms of number of admissible words):

$$\text{Card}\{b_0 \dots b_\ell : \text{Card}\{j < \ell : d(n_{b,\ell}q_j\alpha, \mathbb{Z}/r) \geq \delta\} \leq (1 - \varepsilon_0)\ell\} \leq C_7 q_{\ell+1}^{1-\zeta}. \quad \square$$

$\triangle$

## 2. A central limit theorem and its application to rotations

### 2.1. Decorrelation and CLT.

*An abstract CLT under a decorrelation property*

Below  $Y_1$  denotes a r.v. with a normal distribution  $\mathcal{N}(0, 1)$ . Recall that, if  $X, Y$  are two real random variables, their mutual (Kolmogorov) distance in distribution is defined by:  $d(X, Y) = \sup_{x \in \mathbb{R}} |\mathbb{P}(X \leq x) - \mathbb{P}(Y \leq x)|$ .

The notation  $C$  denotes an absolute constant whose value may change from a line to the other.

**Proposition 2.1.** *Let  $N$  be a positive integer. Let  $(q_k)_{1 \leq k \leq N}$  be an increasing sequence of positive integers such that for a constant  $\rho \in ]0, 1[$*

$$(42) \quad q_k/q_m \leq C \rho^{m-k}, \quad 1 \leq k < m \leq N.$$

*Let  $(f_k)_{1 \leq k \leq N}$  be real centered BV functions such that for constants  $u_k$*

$$(43) \quad \|f_k\|_\infty \leq u_k, \quad V(f_k) \leq C u_k q_k, \quad 1 \leq k \leq N.$$

*Moreover assume that there is a constant  $\theta$  such that*

$$(44) \quad \left| \int_X \psi f_k d\mu \right| \leq C V(\psi) u_k \frac{k^\theta}{q_k}, \quad 1 \leq k \leq N, \forall \psi \text{ BV},$$

$$(45) \quad \left| \int_X \psi f_k f_m d\mu \right| \leq C V(\psi) u_k u_m \frac{m^\theta}{q_k}, \quad 1 \leq k \leq m \leq N, \forall \psi \text{ BV centered},$$

$$(46) \quad \left| \int_X \psi f_k f_m f_t d\mu \right| \leq C V(\psi) u_k u_m u_t \frac{t^\theta}{q_k}, \quad 1 \leq k \leq m \leq t \leq N, \forall \psi \text{ BV centered}.$$

*Then, putting  $w_N := \max_{j=1}^N u_j$ ,  $S_N := f_1 + \dots + f_N$ , there is for every  $\delta > 0$  a constant  $C(\delta) > 0$  (depending only on  $\delta$ ) such that, if  $\frac{w_N}{\|S_N\|_2} \leq N^{p-\frac{1}{2}}$  with  $p \in [0, \frac{1}{8}]$ ,*

$$(47) \quad d\left(\frac{S_N}{\|S_N\|_2}, Y_1\right) \leq C(\delta) N^{-\frac{1-8p}{12} + \delta}.$$

The proposition is proved in Section 3. We apply it to an irrational rotation by taking for  $q_k$ 's the denominators of  $\alpha$  (they satisfy (42)) and for  $f_k$  the ergodic sums  $\varphi_{b_k q_k}$  of a function  $\varphi$  (composed by a translation), where the  $b_k$ 's ( $b_k \leq a_{k+1}$ ) are given by the Ostrowski's expansion described above.

#### *Decorrelation between partial ergodic sums*

In order to apply the previous proposition we will prove decorrelation properties between the ergodic sums of  $\varphi \in \mathcal{C}$  at time  $q_n$  under the following assumption on  $\alpha$ :

**Hypothesis 1.** *There are two constants  $A \geq 1$ ,  $p \geq 0$  such that*

$$(48) \quad a_n \leq A n^p, \quad \forall n \geq 1.$$

The case  $\alpha$  of bounded type (bpq), i.e., when the sequence  $(a_k)$  of partial quotients of  $\alpha$  is bounded, corresponds in the previous condition to  $p = 0$ .

**Lemma 2.2.** *a) For every  $p > 1$ , for a.e.  $\alpha$ , there is a finite constant  $A(\alpha, p)$  such that*

$$(49) \quad a_n \leq A(\alpha, p) n^p, \quad \forall n \geq 1.$$

*b) If  $\alpha$  satisfies (48), then there is  $c > 0$  such that*

$$(50) \quad \|k\alpha\| \geq \frac{c}{|k| (\log k)^p}, \quad \forall k > 1.$$

*Proof.* a) We have  $a_{n+1}(\alpha) = \lfloor 1/\theta^n(\alpha) \rfloor$  where  $\theta$  is the Gauss map. Let  $\gamma > 1$ . Since  $\alpha \rightarrow (a_1(\alpha))^{\frac{1}{\gamma}}$  is integrable for the  $\theta$ -invariant measure  $\frac{dx}{1+x}$  on  $]0, 1]$ , we have, for a constant  $A(\gamma)$ :  $\mu\{\alpha : a_n(\alpha) > n^s\} \leq A(\gamma) n^{-\frac{s}{\gamma}}$ .

By the Borel-Cantelli lemma, it follows that for a.e  $\alpha$  there is  $C(\alpha, \gamma)$  such that, if  $s > \gamma$ ,  $a_n(\alpha) \leq C(\alpha, \gamma) n^s, \forall n \geq 1$ .

b) For every irrational  $\alpha$ , there are  $C > 0$  and  $\lambda > 1$  such that the denominators of  $\alpha$  satisfy  $q_\ell \geq C\lambda^\ell$ , for every  $\ell \geq 1$ . For  $k \geq 2$ , let  $n$  be such that  $q_{n-1} \leq k < q_n$ . Since  $C\lambda^{n-1} \leq q_{n-1} \leq k$ , it follows that  $n \leq C' \log k$ , for some constant  $C'$ . By (48), we have  $a_n \leq An^p \leq A(C' \log k)^p$ .

As  $\|k\alpha\| > \|q_{n-1}\alpha\| \geq \frac{1}{2q_n} \geq \frac{1}{4a_n q_{n-1}} \geq \frac{1}{4a_n k}$ , this implies (50).  $\square$

**Proposition 2.3.** *Let  $\psi$  and  $\varphi$  be BV centered functions. Suppose that  $\alpha$  satisfies Hypothesis 1. Then there are constants  $C, \theta_1, \theta_2, \theta_3$  such that, for every  $1 \leq k \leq m \leq \ell$ :*

$$(51) \quad \left| \int_X \psi \varphi_{b_k q_k} d\mu \right| \leq C V(\psi) V(\varphi) \frac{k^{\theta_1}}{q_k} b_k,$$

$$(52) \quad \left| \int_X \psi \varphi_{b_k q_k} \varphi_{b_m q_m} d\mu \right| \leq C V(\psi) V(\varphi)^2 \frac{m^{\theta_2}}{q_k} b_k b_m,$$

$$(53) \quad \left| \int_X \psi \varphi_{b_k q_k} \varphi_{b_m q_m} \varphi_{b_\ell q_\ell} d\mu \right| \leq C V(\psi) V(\varphi)^3 \frac{\ell^{\theta_3}}{q_k} b_k b_m b_\ell.$$

The proposition is proved in Section 4. From the propositions 2.1 and 2.3 we will deduce a convergence toward a Gaussian distribution under a variance condition, by bounding the distance to the normal distribution.

**Theorem 2.4.** *Let  $\varphi$  be in  $\mathcal{C}$  satisfying (12).*

1) *With the notation of Theorem 1.3, the set*

$$(54) \quad W := \{n \in \mathbb{N} : \|\varphi_n\|_2 \geq \eta_0 (\log m(n))^{-\frac{1}{2}} m(n)^{\frac{1}{2}}\}$$

*has density 1 in  $\mathbb{N}$ .*

*Suppose that  $\alpha$  satisfies Hypothesis 1 (i.e., for constants  $A \geq 1, p \geq 0, a_n \leq An^p, \forall n \geq 1$ ) with  $p < \frac{1}{8}$ . Then, for  $\delta \in ]0, \frac{1-8p}{12}[$ , there is a constant  $C(\delta)$  such that for  $n$  big enough in  $W$ :*

$$(55) \quad d\left(\frac{\varphi_n}{\|\varphi_n\|_2}, Y_1\right) \leq C(\delta) m(n)^{-\frac{1-8p}{12} + \delta} \xrightarrow[n \in W, n \rightarrow \infty]{} 0.$$

*In particular if  $\alpha$  is bpq,  $p = 0$  and  $m(n)$  can be replaced by  $\log n$ .*

2) *Suppose that  $\alpha$  is a quadratic irrational. With the notation of Theorem 1.5, let*

$$V := \{n \geq 1 : \eta_1 \sqrt{\log n} \leq \|\varphi_n\|_2 \leq \eta_2 \sqrt{\log n}\}.$$

*Then, there are two constants  $R, \xi > 0$  such that*

*- the density of  $V$  satisfies:*

$$(56) \quad \text{Card}(V \cap [1, N]) \geq N(1 - RN^{-\xi}), \text{ for } N \geq N_0;$$

- for  $\delta \in ]0, \frac{1}{12}[$ , there is a constant  $C(\delta)$  such that, for  $n \in V$ :

$$(57) \quad d\left(\frac{\varphi_n}{\|\varphi_n\|_2}, Y_1\right) \leq C(\delta) (\log n)^{-\frac{1}{12}+\delta} \xrightarrow{n \in V, n \rightarrow \infty} 0.$$

*Proof.* 1) The result on the density of the set  $W$  follows from Theorem 1.3.

For (55), we use Proposition 2.1 with  $N = m(n)$  such that  $n \in [q_{m(n)}, q_{m(n)+1}[$ ,  $f_k$  defined by (7) and the decomposition of the ergodic sums given by (6), i.e.,

$$\varphi_n(x) = \sum_{k=0}^{m(n)} f_k(x), \text{ where } f_k(x) := \sum_{i=0}^{b_k-1} \varphi_{q_k}(x + (n_{k-1} + iq_k)\alpha) = \varphi_{b_k q_k}(x + n_{k-1}\alpha).$$

The decorrelation inequalities in Proposition 2.3 are obtained for functions of the form  $\varphi_{b_k q_k}$ . But in the proof of the decorrelation inequalities, one sees that they remain valid for  $f_k$ , since translations on the variable do not change the modulus of the Fourier coefficients.

As  $\|f_k\|_\infty \leq b_k V(\varphi) \leq a_{k+1} V(\varphi)$ , up to a fixed factor the constant  $u_k$  in the statement of Proposition 2.3 can be taken to be  $a_{k+1} \leq k^p$ , for some constant  $p > 0$ , by Hypothesis 1.

With the notation of Proposition 2.1, we have  $w_N := \max_{j=1}^N b_j$ ,  $\varphi_n = S_N = f_1 + \dots + f_N$ . For  $n \in Z_b$  and under Hypothesis 1, we have  $\frac{w_N}{\|S_N\|_2} \leq CN^{p-\frac{1}{2}}(\log N)^{\frac{1}{2}}$  (the factor  $(\log N)^{\frac{1}{2}}$  can be absorbed in the factor  $N^{p-\frac{1}{2}}$  by taking  $p$  larger). By (47) we obtain the following inequality:

$$d\left(\frac{\varphi_n}{\|\varphi_n\|_2}, Y_1\right) \leq C(\delta) m(n)^{-\frac{1-8p}{12}+\delta}.$$

2) In the quadratic case,  $p = 0$  and the property of the set  $V$  is given by Theorem 1.5.  $\square$

**Remark 2.** The previous result is written with a self-normalisation. If  $\alpha$  is quadratic, let us consider the ergodic sums normalised by  $\sqrt{\ln n}$ :  $(\varphi_n/\sqrt{\ln n})_{n \geq 1}$ . Then, for  $n \in V$ , the accumulation points of the sequence of distributions are Gaussian non degenerated with a variance belonging to a compact interval.

## 2.2. Application to step functions, examples.

If  $\varphi$  belongs to the class  $\mathcal{C}$  of centered BV functions, with Fourier series  $\sum_{r \neq 0} \frac{\gamma_r(\varphi)}{r} e^{2\pi i r x}$ , to apply Theorem 2.4 we have to check Condition (12) on the coefficients  $\gamma_{q_k}(\varphi)$ , i.e.:

$$\exists M, \eta, \theta > 0 \text{ such that } \frac{1}{N} \text{Card} \{j \leq N : a_{j+1} |\gamma_{q_j}(\varphi)| \geq \eta\} \geq \theta, \forall N \geq M.$$

The functions  $\{x\} - \frac{1}{2} = \frac{-1}{2\pi i} \sum_{r \neq 0} \frac{1}{r} e^{2\pi i r x}$  and  $1_{[0, \frac{1}{2}[} - 1_{[\frac{1}{2}, 1[} = \sum_r \frac{2}{\pi i(2r+1)} e^{2\pi i(2r+1)x}$  are immediate examples where (12) are satisfied. In the second case, one observes that  $\gamma_{q_k} = 0$  if  $q_k$  is even,  $= \frac{2}{\pi i}$  if  $q_k$  is odd. Clearly, (12) is satisfied, because two consecutive  $q_k$ 's are relatively prime and therefore cannot be both even.

In general, for a step function, Condition (12) (and therefore a lower bound for the variance  $\|\varphi_n\|_2^2$  for a large set of integers  $n$ ) is related to the Diophantine properties of its discontinuities with respect to  $\alpha$ . We discuss now this point.

Let us consider a centered step function  $\varphi$  on  $[0, 1[$  taking a non null constant value  $v_j \in \mathbb{R}$  on the interval  $[u_j, u_{j+1}[$ ,  $j = 0, 1, \dots, s$ , with  $u_0 = 0 < u_1 < \dots < u_s < u_{s+1} = 1$ :

$$(58) \quad \varphi = \sum_{j=0}^s v_j 1_{[u_j, u_{j+1}[} - c.$$

The constant  $c$  above is such that  $\varphi$  is centered, but it plays no role below.

**Lemma 2.5.** *If  $\varphi$  is given by (58), there is a continuous function  $H_\varphi(u_1, \dots, u_s) \geq 0$  such that*

$$(59) \quad |\gamma_r(\varphi)|^2 = \pi^{-2} H_\varphi(ru_1, \dots, ru_s).$$

*Proof.* Since  $\widehat{\varphi}(r) = \sum_{j=0}^s \frac{v_j}{\pi r} e^{-\pi i r(u_j + u_{j+1})} \sin \pi r(u_{j+1} - u_j)$ ,  $r \neq 0$ ,  $H_\varphi(u_1, \dots, u_s)$  is

$$\left[ \sum_{j=0}^s v_j \cos \pi(u_j + u_{j+1}) \sin \pi(u_{j+1} - u_j) \right]^2 + \left[ \sum_{j=0}^s v_j \sin \pi(u_j + u_{j+1}) \sin \pi(u_{j+1} - u_j) \right]^2. \quad \square$$

*Examples:* 1)  $\varphi = \varphi(u, \cdot) = 1_{[0, u[} - u$ ,  $H_\varphi(u) = \sin^2(\pi u)$ .

2)  $\varphi = \varphi(w, u, \cdot) = 1_{[0, u[} - 1_{[w, u+w]}$ ,  $H(\varphi) = 4 \sin^2(\pi u) \sin^2(\pi w)$ .

**Corollary 2.6.** *Condition (12) is satisfied if the parameter  $(u_1, \dots, u_s)$  is such that the sequence  $(q_k u_1, \dots, q_k u_s)_{k \geq 1}$  is uniformly distributed in  $\mathbb{T}^s$ .*

Since  $(q_k)$  is a strictly increasing sequence of integers, for almost every  $(u_1, \dots, u_s)$  in  $\mathbb{T}^s$ , the sequence  $(q_k u_1, \dots, q_k u_s)_{k \geq 1}$  is uniformly distributed in  $\mathbb{T}^s$ . Hence:

$$\lim_n \frac{1}{n} \sum_{k=1}^n H_\varphi(q_k u_1, \dots, q_k u_s) = \int_{\mathbb{T}^s} H(x_1, \dots, x_s) dx_1 \dots dx_s > 0, \text{ for a.e. } (u_1, \dots, u_s) \in \mathbb{T}^s.$$

Therefore (12) is a generic property: it is satisfied for a.e. value of  $(u_1, \dots, u_s)$  in  $\mathbb{T}^s$ .

Besides a generic result, there are also specific values of the parameter  $(u_1, \dots, u_s)$  for which (12) holds. A simple example is:

*Example 3:*  $\varphi(\frac{r}{s}, \cdot) = 1_{[0, \frac{r}{s}[} - \frac{r}{s}$ , for  $r, s \in \mathbb{N}$ ,  $0 < r < s$ .

### Remarks about example 1)

1) It is known that if  $\alpha$  is bpq and if  $\lim_k |\sin(\pi q_k u)| = 0$  for the denominators  $q_k$  of  $\alpha$ , then  $u \in \mathbb{Z}\alpha + \mathbb{Z}$ . But there is an uncountable set of  $u$ 's such that  $\lim_N \frac{1}{N} \sum_{k=1}^N \sin^2(\pi q_k u) = 0$ .

Observe also that, if  $\alpha$  is not bpq, there are many  $u$ 's which do not satisfy the previous equidistribution property. Let  $u = \sum_{n \geq 0} b_n q_n \alpha \pmod{1}$ ,  $b_n \in \mathbb{Z}$ , be the so-called Ostrowski expansion of  $u$ , where  $q_n$  are the denominators of  $\alpha$ . It can be shown that, if  $\sum_{n \geq 0} \frac{|b_n|}{a_{n+1}} < \infty$ , then  $\lim_k \|q_k u\| = 0$ . There is an uncountable set of  $u$ 's satisfying this condition if  $\alpha$  is not bpq. Therefore we have degeneracy of the variance for these  $u$ 's, although ergodicity of the cocycle holds if  $u$  is not in the countable set  $\mathbb{Z}\alpha + \mathbb{Z}$ .

2) Again in example 1, for the case 1) of Theorem 2.4, we need a slight improvement of the equidistribution argument mentioned previously, namely:

If  $1 \leq a_n \leq n^\gamma$ , with  $\gamma < \frac{1}{6}$ , then:  $\lim_N \frac{\sum_{k=1}^N a_k^2 \sin^2(\pi q_k u)}{\sum_{k=1}^N a_k^2} = \frac{1}{2}$ , for almost every  $u$ .

This is an easy consequence of the following lemma:

**Lemma 2.7.** *Let  $(u_n)$  be a sequence of real numbers such that  $1 \leq u_n \leq n^\gamma$ , with  $0 \leq \gamma < \frac{1}{2}$ . If  $(X_n)$  is a sequence of bounded orthogonal random variables on a probability space  $(\Omega, \mu)$ , then  $\lim_N \frac{\sum_{k=1}^N u_k X_k}{\sum_{k=1}^N u_k} = 0$ , for a.e.  $\omega$ .*

### Vectorial case

For simplicity, we consider the case of two components. Let be given a vectorial function  $\Phi = (\varphi^1, \varphi^2)$ , where  $\varphi^1, \varphi^2$  are two centered step functions with respectively  $s_1, s_2$  discontinuities:  $\varphi^i = \sum_{j=0}^{s_i} v_j^i 1_{[u_j^i, u_{j+1}^i[} - c_i$ , for  $i = 1, 2$ .

Let the matrix  $\Gamma_n$  be defined by  $\Gamma_n(a, b) := (\log n)^{-1} \|a\varphi_n^1 + b\varphi_n^2\|_2^2$  and denote by  $I_2$  the 2-dimensional identity matrix.

**Theorem 2.8.** *If  $\alpha$  is bpq and if the condition (12) is satisfied uniformly with respect to  $(a, b)$  in the unit sphere, there are  $0 < r_1, r_2 < +\infty$  two constants such that for a “large” set of integers  $n$  as in Theorem 2.4:*

- $\Gamma_n$  satisfies inequalities of the form  $r_1 I_2 \leq \Gamma_n(a, b) \leq r_2 I_2$ ;
- the distribution of  $\Gamma_n^{-1} \Phi_n$  is close to the standard 2-dimensional normal law.

*Proof.* We only sketch the proof. The classical method of proof of a CLT for a vectorial function is to show a scalar CLT for all linear combinations of the components of the function. So the proof is like that of Theorem 2.1, but taking care of the bound from below of the variance for  $a\varphi_n^1 + b\varphi_n^2$ : (12) should be uniform for  $(a, b)$  on the unit sphere. This is done in the next proposition.  $\square$

**Proposition 2.9.** *Let  $\Lambda$  be a compact space and  $(F_\lambda, \lambda \in \Lambda)$  be a family of nonnegative non identically null continuous functions on  $\mathbb{T}^d$  depending continuously on  $\lambda$ . If a sequence  $(z_n)$  is equidistributed in  $\mathbb{T}^d$ , then*

$$(60) \quad \exists N_0, \eta > 0 \text{ such that } \text{Card}\{n \leq N : F_\lambda(z_n) \geq \eta\} \geq \theta N, \forall N \geq N_0, \forall \lambda \in \Lambda.$$

*Proof.* For  $\lambda \in \Lambda$ , let  $u_\lambda \in \mathbb{T}^d$  be such that  $F_\lambda(u_\lambda) = \sup_{u \in \mathbb{T}^d} F_\lambda(u)$ . We have  $F_\lambda(u_\lambda) > 0$  and there is  $\eta_\lambda > 0$  and an open neighborhood  $U_\lambda$  of  $u_\lambda$  such that  $F_\lambda(u) > 2\eta_\lambda$  for  $u \in U_\lambda$ . Using the continuity of  $F_\lambda$  with respect to the parameter  $\lambda$ , the inequality  $F_\zeta(u) > \eta_\lambda$  holds for  $u \in U_\lambda$  and  $\zeta$  in an open neighborhood  $V_\lambda$  of  $\lambda$ . By compactness of  $\Lambda$ , there is a finite set  $(\lambda_j, j \in J)$  such that  $(V_{\lambda_j}, j \in J)$  is an open covering of  $\Lambda$ . Let  $\theta := \frac{1}{2} \inf_{j \in J} \text{Leb}(U_{\lambda_j})$ .

By equidistribution of  $(z_n)$ , there is  $N_0$  such that  $\frac{1}{N} \sum_{n=1}^N 1_{U_{\lambda_j}}(z_n) \geq \theta, \forall N \geq N_0, \forall j \in J$ .

Let  $\eta := \inf_{j \in J} \eta_{\lambda_j}$ . For every  $\lambda \in \Lambda$ , there is  $j \in J$  such that  $\lambda \in V_{\lambda_j}$  and therefore  $F_\lambda(z_n) \geq \eta_{\lambda_j} \geq \eta$ , if  $z_n \in U_{\lambda_j}$ . This implies:

$$\text{Card}\{n \leq N : F_\lambda(z_n) \geq \eta\} \geq \text{Card}\{n \leq N : z_n \in U_{\lambda_j}\} \geq \theta N, \forall N \geq N_0. \quad \square$$

*A generic result*

By Proposition 2.9 applied for  $(a, b)$  in the unit sphere, for a.e. values of the parameter  $(u_1^1, \dots, u_{s_1}^1, u_1^2, \dots, u_{s_2}^2)$ , the functions  $a\varphi^1 + b\varphi^2$  satisfy Condition (12) uniformly in  $(a, b)$  in the unit sphere. Hence Theorem 2.8 applies generically with respect to the discontinuities.

*Special values: an application to the rectangular billiard in the plane*

*Example 4* Now, for an application to the periodic billiard, we consider the vectorial function  $\psi = (\varphi^1, \varphi^2)$  with

$$\begin{aligned}\varphi^1 &= 1_{[0, \frac{\alpha}{2}]} - 1_{[\frac{1}{2}, \frac{1}{2} + \frac{\alpha}{2}]} = \frac{2}{\pi} \sum_{r \in \mathbb{Z}} e^{-\pi i(2r+1)\frac{\alpha}{2}} \frac{\sin(\pi(2r+1)\frac{\alpha}{2})}{2r+1} e^{2\pi i(2r+1)}, \\ \varphi^2 &= 1_{[0, \frac{1}{2} - \frac{\alpha}{2}]} - 1_{[\frac{1}{2}, 1 - \frac{\alpha}{2}]} = \frac{-2i}{\pi} \sum_{r \in \mathbb{Z}} e^{\pi i(2r+1)\frac{\alpha}{2}} \frac{\cos(\pi(2r+1)\frac{\alpha}{2})}{2r+1} e^{2\pi i(2r+1)}.\end{aligned}$$

The Fourier coefficients of  $\varphi^1$  and  $\varphi^2$  of order  $r$  are null for  $r$  even.

Let us consider a linear combination  $\varphi_{a,b} = a\varphi^1 + b\varphi^2$ . For  $r = 2t + 1$  odd, we have:

$$c_{2t+1}(a\varphi^1 + b\varphi^2) = \frac{2}{\pi} \frac{1}{2t+1} e^{-\pi i(2t+1)\frac{\alpha}{2}} [a \sin(\pi(2t+1)\frac{\alpha}{2}) - ib \cos(\pi(2t+1)\frac{\alpha}{2})].$$

If  $q_j$  is even,  $\gamma_{q_j}(\varphi_{a,b})$  is null. If  $q_j$  is odd, we have

$$|\gamma_{q_j}(\varphi_{a,b})|^2 = |a \sin(\pi q_j \frac{\alpha}{2}) - ib \sin(\pi(\frac{1}{2} + q_j \frac{\alpha}{2}))|^2,$$

For  $q_j$  odd, we have by (2),

$$\begin{aligned}\|q_j \frac{\alpha}{2}\| &= \|\frac{p_j}{2} + \frac{\theta_j}{2}\|, \text{ hence } \left| \|q_j \frac{\alpha}{2}\| - \|\frac{p_j}{2}\| \right| \leq \left| \frac{\theta_j}{2} \right| \leq \frac{1}{2q_{n+1}}, \\ \|\frac{1}{2} + q_j \frac{\alpha}{2}\| &= \|\frac{q_j}{2} - \frac{p_j}{2} - \frac{\theta_j}{2}\|, \text{ hence } \left| \|q_j \beta_2\| - \left( \|\frac{1}{2} + \frac{p_j}{2}\| \right) \right| \leq \left| \frac{\theta_j}{2} \right| \leq \frac{1}{2q_{n+1}},\end{aligned}$$

This implies, for  $q_j$  odd:  $\gamma_{q_j}(\varphi_{a,b}) = a(1 + O(\frac{1}{q_{j+1}}))$ , if  $p_j$  is odd,  $= b(1 + O(\frac{1}{q_{j+1}}))$ , if  $p_j$  is even.

The computation shows that, if  $\alpha$  is such that, in average, there is a positive proportion of pairs  $(p_j, q_j)$  which are (even, odd) and a positive proportion of pairs  $(p_j, q_j)$  which are (odd, odd), then the condition of Theorem 2.8 is fulfilled by the vectorial step function  $\psi = (\varphi^1, \varphi^2)$ .

For an application to the model of rectangular periodic billiard in the plane described in [5], we refer to [4].

### 3. Proof of Proposition 2.1 (CLT)

The difference  $H_{X,Y}(\lambda) := |\mathbb{E}(e^{i\lambda X}) - \mathbb{E}(e^{i\lambda Y})|$  can be used to get an upper bound of the distance  $d(X, Y)$  thanks to the following inequality ([7], Chapter XVI, Inequality (3.13)):



if  $X$  has a vanishing expectation, then, for every  $U > 0$ ,

$$(61) \quad d(X, Y) \leq \frac{1}{\pi} \int_{-U}^U H_{X,Y}(\lambda) \frac{d\lambda}{\lambda} + \frac{24}{\pi} \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{U}.$$

Using (61), we get an upper bound of the distance between the distribution of  $X$  and the normal law by bounding  $|\mathbb{E}(e^{i\lambda X}) - e^{-\frac{1}{2}\sigma^2\lambda^2}|$ .

We will use the following remarks:

$$(62) \quad V(fg) \leq \|f\|_\infty V(g) + \|g\|_\infty V(f), \quad \forall f, g \text{ BV},$$

$$(63) \quad \text{if } g \in \mathcal{C}^1(\mathbb{R}, \mathbb{R}) \text{ and } u \text{ is BV, then } V(g \circ u) \leq \|g'\|_\infty V(u).$$

Let  $w_k := \max_{j=1}^k u_j$ , where  $u_j$  is larger than  $\|f_j\|_\infty$  (see Proposition 2.1).

Since  $V(f_k) \leq C u_k q_k$ , (44) implies

$$(64) \quad \left| \int_X f_k f_m d\mu \right| \leq C \frac{q_k}{q_m} m^\theta w_m^2, \quad \text{for } k \leq m.$$

### Bounding the moments

**Lemma 3.1.** *Under the assumption of Proposition 2.1, there is  $C_1$  such that*

$$(65) \quad \int_X \left| \sum_{k=m}^{m+\ell} f_k \right|^2 \leq C_1 \ln(m+\ell) \sum_{j \in [m, m+\ell]} u_j^2 \leq C_1 \ell \ln(m+\ell) w_{m+\ell}^2.$$

$$(66) \quad \int_X \left| \sum_{k=m}^{m+\ell} f_k \right|^3 \leq C_1 \ell \ln^2(m+\ell) w_{m+\ell}^3, \quad \forall m, \ell \geq 1,$$

$$(67) \quad \int_X \left| \sum_{k=m}^{m+\ell} f_k \right|^4 \leq C_1 \ell^2 \ln^2(m+\ell) w_{m+\ell}^3, \quad \forall m, \ell \geq 1.$$

*Proof.* We show (66) and (67). The proof of (65) is the same.

1) For (66), it suffices to bound the sums  $\sum_{m \leq s \leq t \leq u \leq m+\ell} \left| \int_X f_s f_t f_u d\mu \right|$ .

Replacing  $f_k$  by  $w_{m+\ell}^{-1} f_k$ , we will deduce the bound (66) from the inequalities (43), (44), (45) when  $w_k \leq 1$ , for  $1 \leq k \leq m+\ell$ . By (43) and (62), we have

$$\left| \int_X f_s f_t f_u d\mu \right| \leq C \quad \text{and} \quad V(f_s f_t) \leq C(q_s + q_t) \leq 3Cq_t.$$

From (44) and (42), then from (45) and (42), we obtain

$$\begin{aligned} \left| \int_X (f_s f_t) \cdot f_u d\mu \right| &\leq C \frac{V(f_s f_t)}{q_u} u^\theta \leq C \frac{q_t}{q_u} u^\theta \leq C \rho^{(u-t)} u^\theta, \\ \left| \int_X f_s \cdot (f_t f_u) d\mu \right| &\leq C \frac{V(f_s)}{q_t} u^\theta \leq C \frac{q_s}{q_t} u^\theta \leq C \rho^{(t-s)} u^\theta. \end{aligned}$$

Set  $\kappa = \frac{\theta+3}{\ln(1/\rho)} \ln(m+\ell)$ . If  $t-s$  or  $u-t \geq \kappa$ , the previous inequalities imply:

$$\left| \int_X f_s f_t f_u d\mu \right| \leq C \rho^\kappa u^\theta \leq C(m+\ell)^{-\theta-3} u^\theta \leq C(m+\ell)^{-3}.$$

It implies:

$$\sum_{m \leq s \leq t \leq u \leq m+\ell: \max(t-s, u-t) > \kappa} \left| \int_X f_s f_t f_u d\mu \right| \leq C \ell^3 (\ell+m)^{-3} \leq C.$$

Now the result follows from:

$$\sum_{m \leq s \leq t \leq u \leq m+\ell: \max(t-s, u-t) \leq \kappa} \left| \int_X f_s f_t f_u d\mu \right| \leq C \ell \kappa^2.$$

2) For (67), we bound the sums  $\sum_{m \leq s \leq t \leq u \leq v \leq m+\ell} \left| \int_X f_s f_t f_u f_v d\mu \right|$  using (42) and successively (44), (45), (46).

We obtain (because  $f_v$  is centered for the first inequality):

$$\begin{aligned} \left| \int_X (f_s f_t f_u) \cdot f_v d\mu \right| &\leq C \frac{V(f_s f_t f_u)}{q_v} v^\theta \leq C \frac{q_u}{q_v} v^\theta \leq C \rho^{(v-u)} v^\theta, \\ \left| \int_X [f_s f_t - \mathbb{E}(f_s f_t)] f_u f_v d\mu \right| &\leq C \frac{V(f_s f_t)}{q_u} v^\theta \leq C \frac{q_t}{q_u} v^\theta \leq C \rho^{(u-t)} v^\theta, \\ \left| \int_X f_s f_t f_u f_v d\mu \right| &\leq C \frac{V(f_s)}{q_t} v^\theta \leq C \frac{q_s}{q_t} v^\theta \leq C \rho^{(t-s)} v^\theta. \end{aligned}$$

Putting  $\kappa = \frac{\theta+4}{\ln(1/\rho)} \log(m+\ell)$ , we get by the previous inequalities, for constants  $C, C_2, C_3$ :

$$\begin{aligned} &\sum_{m \leq s \leq t \leq u \leq v \leq m+\ell: \max(t-s, u-t, v-u) > \kappa} \left| \int_X f_s f_t f_u f_v d\mu \right| \\ &\leq C \ell^4 (\ell+m)^{-4} + \sum_{m \leq s \leq t \leq u \leq v \leq m+\ell} \left| \int_X f_s f_t d\mu \right| \left| \int_X f_u f_v d\mu \right| \\ &\leq C + \left( \sum_{m \leq s \leq t \leq m+\ell} \left| \int_X f_s f_t d\mu \right| \right)^2 \leq C + C_2 \ell^2 (\ln(m+\ell))^2. \end{aligned}$$

The remaining terms give a bound which can be absorbed in the previous one, namely:

$$\sum_{m \leq s \leq t \leq u \leq m+\ell: \max(t-s, u-t, v-u) \leq \kappa} \left| \int_X f_s f_t f_u f_v d\mu \right| \leq C \ell \kappa^3 \leq C_3 \ell \log(m+\ell)^3. \quad \square$$

## Proof of Proposition 2.1

The proof is given in several steps.

*Defining blocks*

We split the sum  $S_n := f_1 + \dots + f_n$  into small and large blocks. The small ones will be removed, providing gaps and allowing to take advantage of the decorrelation properties assumed in the statement of the proposition.

Let  $\tau, \delta$  be parameters ( $\delta$  close to 0) such that  $0 < \delta < \frac{1}{2}$  and  $\delta < \tau$ . We set for  $n \geq 1$ :

$$\begin{aligned} n_1 = n_1(n) &:= \lfloor n^\tau \rfloor, n_2 = n_2(n) := \lfloor n^\delta \rfloor, \\ \nu = \nu(n) &:= n_1 + n_2, p(n) := \lfloor n/\nu(n) \rfloor + 1 = n^{1-\tau} + h_n \sim n^{1-\tau}, \text{ with } |h_n| \text{ bounded.} \end{aligned}$$

For  $0 \leq k < p(n)$ , we put (with  $f_j = 0$ , if  $n < j \leq n + \nu$ )

$$F_{n,k} = f_{k\nu(n)+1} + \dots + f_{k\nu(n)+n_1(n)}, \quad G_{n,k} = f_{k\nu(n)+n_1(n)+1} + \dots + f_{(k+1)\nu(n)}.$$

The sums  $F_{n,k}, G_{n,k}$  have respectively  $n_1 \sim n^\tau, n_2 \sim n^\delta$  terms and  $S_n$  reads

$$S_n = \sum_{k=0}^{p(n)-1} (F_{n,k} + G_{n,k}).$$

We put  $S'_n := \sum_{k=0}^{p(n)-1} F_{n,k}$ ,  $v_k = v_{n,k} := (\int_X F_{n,k}^2 d\mu)^{\frac{1}{2}}$ .

The following inequalities are implied by (65):

$$(68) \quad v_k^2 = v_{n,k}^2 = \|F_{n,k}\|_2^2 \leq Cn^\tau \ln n w_n^2, \quad \|G_{n,k}\|_2^2 \leq Cn^\delta \ln n w_n^2, \quad 0 \leq k < p(n).$$

Since  $q_1 + q_2 + \dots + q_n \leq Cq_{n+1}, \forall n \geq 1$ , by (42), it follows by (63) and hypothesis (43):

$$(69) \quad \mathbb{V}(e^{i\zeta(F_{n,0} + \dots + F_{n,k-1})}) \leq C|\zeta| w_n q_{(k-1)\nu+n_1}.$$

**Lemma 3.2.**

$$(70) \quad \left| \|S_n\|_2^2 - \sum_{k=0}^{p(n)-1} v_k^2 \right| = \left| \|S_n\|_2^2 - \sum_{k=0}^{p(n)-1} \|F_{n,k}\|_2^2 \right| \leq Cn^{1-\frac{\tau-\delta}{2}} \ln n w_n^2,$$

$$(71) \quad \|S_n - S'_n\|_2^2 = \left\| \sum_{k=0}^{p(n)-1} G_k \right\|^2 \leq Cn^{1-\tau+\delta} \ln n w_n^2.$$

*Proof.* It follows from (64) and (42), with  $C_0 = \frac{C\rho}{(1-\rho)^2}$ ,

$$\left| \int_X \left( \sum_{u=a}^b f_u \right) \left( \sum_{t=c}^d f_t \right) d\mu \right| \leq C_0 \rho^{c-b} d^\theta w_d^2, \quad \forall a \leq b < c \leq d.$$

Therefore, we have, with  $C_1 = C_0 \sum_{i \geq 0} \rho^{i\nu}$ , writing simply  $F_k, G_k$  instead of  $F_{n,k}, G_{n,k}$ ,

$$\begin{aligned} \sum_{0 \leq j < k < p(n)} \left| \int F_j F_k d\mu \right| &\leq C_0 n^\theta w_n^2 \sum_{0 \leq j < k < p(n)} \rho^{k\nu+1-(j\nu+n_1)} \\ &\leq C_0 n^\theta w_n^2 \rho^{n_2} \sum_{0 \leq j < k < p(n)} \rho^{(k-1)\nu-j\nu} \leq C_0 \rho^{n_2} n^\theta w_n^2 p(n) \sum_{i \geq 0} \rho^{i\nu} \leq C_1 n^{\frac{1}{2}-\delta+\theta} w_n^2 \rho^{n^\delta}. \end{aligned}$$

The LHS of (70) is less than the sum for  $k = 0$  to  $p(n) - 1$  of

$$\begin{aligned} & \int G_k^2 d\mu + \left| \int G_k F_k d\mu \right| + \left| \int G_k F_{k+1} d\mu \right| \\ & + 2 \sum_{0 \leq j < k} \left[ \left| \int F_j (F_k + G_k) d\mu \right| + \left| \int G_j G_k d\mu \right| \right] + 2 \sum_{0 \leq j < k-1} \left| \int G_j F_k d\mu \right|. \end{aligned}$$

The first term is bounded by  $C n^\delta \ln n w_n^2$ , the second one and the third one bounded by  $C n^{\frac{\delta+\tau}{2}} \ln n w_n^2$  are the biggest. The other terms are negligible as shown by the preliminary computation:  $n^\theta \rho^{n^\delta}$  is small compared to a power of  $n$ , for  $n$  big.

Therefore the LHS of (70) is less than:  $C_1 n^{1-\tau} n^{\frac{\delta+\tau}{2}} \ln n w_n^2 = C_1 n^{1-\frac{\tau-\delta}{2}} \ln n w_n^2$ .

An analogous computation shows that the LHS of (71) behaves like  $\sum_{k=0}^{p(n)-1} \int G_k^2 d\mu$  which gives the bound  $C n^{1-\tau+\delta} \ln n w_n^2$  of (71).  $\square$

*Approximation of the characteristic function of the sum  $S'_n$  by a product*

For  $\zeta \in \mathbb{R}$ , let  $I_{n,-1}(\zeta) := 1$ ,  $I_{n,k}(\zeta) := \int_X e^{i\zeta(F_{n,0}+\dots+F_{n,k})} d\mu$ ,  $0 \leq k < p(n)$ .

**Lemma 3.3.** *For  $0 \leq k < p(n)$ , we have*

$$(72) \quad |I_{n,k}(\zeta) - (1 - \frac{\zeta^2}{2} v_{n,k}^2) I_{n,k-1}(\zeta)| \leq C (|\zeta|^3 w_n^3 n^\tau \ln^2(n) + \zeta^4 w_n^4 n^{2\tau} \ln^2(n)).$$

*Proof.* We use  $e^{iu} = 1 + iu - \frac{1}{2}u^2 - \frac{i}{6}u^3 + u^4 r(u)$ , with  $|r(u)| \leq \frac{1}{24}$ , for  $u \in \mathbb{R}$ . Let  $k \geq 1$ .

We have  $I_{n,k}(\zeta) = \int_X e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})} [1 + i\zeta F_{n,k} - \frac{\zeta^2}{2} F_{n,k}^2 - \frac{i}{6}\zeta^3 F_{n,k}^3 + \zeta^4 F_{n,k}^4 r(\zeta F_{n,k})] d\mu$ .

For the first term, using (69), we have:

$$\begin{aligned} & \left| \int_X e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})} F_{n,k} d\mu \right| \leq \sum_{j=1}^{n_1} \left| \int_X e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})} f_{k\nu+j} d\mu \right| \\ & \leq C |\zeta| w_n \sum_{j=1}^{n_1} \frac{q_{(k-1)\nu+n_1}}{q_{k\nu+j}} (k\nu+j)^\theta w_{(k-1)\nu+n_1} \leq C |\zeta| w_n^2 n^\theta \sum_{j=1}^{n_1} \rho^{\nu+j-n_1} \\ (73) \quad & \leq \frac{C\rho}{1-\rho} |\zeta| w_n^2 n^\theta \rho^{n_2}. \end{aligned}$$

Similarly, for the second term we apply (45) and (69) and we get:

$$\begin{aligned} & \left| \int_X e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})} F_{n,k}^2 d\mu - I_{n,k-1} \int_X F_{n,k}^2 d\mu \right| \\ (74) \quad & \leq CV (e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})}) w_n^2 \sum_{j'=1}^{n_1} \sum_{j=1}^{j'} \frac{((k-1)\nu+j')^\theta}{q_{(k-1)\nu+j}} \leq C |\zeta| w_n^3 n^{\theta+\tau} \rho^{n_2}. \end{aligned}$$

Likewise (46) and Lemma 3.1 imply:  $|\int_X e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})} F_{n,k}^3 d\mu|$

$$\begin{aligned} &\leq |\int_X (e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})} - \mathbb{E}(e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})}) F_{n,k}^3 d\mu| + |\int_X F_{n,k}^3 d\mu| \\ (75) \quad &\leq C |\zeta| n^{1+2\tau+\theta} w_n^3 \rho^{n_2} + C n^\tau w_n^3 \ln^2(n). \end{aligned}$$

At last, by (67) we have

$$(76) \quad |\int_X e^{i\zeta(F_{n,0}+\dots+F_{n,k-1})} F_{n,k}^4 r(\zeta F_{n,k}) d\mu| \leq \int_X F_{n,k}^4 d\mu \leq w_n^4 n^{2\tau} \ln(n)^2.$$

From (73), (74), (75) and (76), we deduce that  $|I_{n,k}(\zeta) - (1 - \frac{\zeta^2}{2} v_{n,k}^2) I_{n,k-1}(\zeta)|$  is bounded up to a constant factor  $C$  by

$$|\zeta|^2 w_n^2 n^\theta \rho^{n_2} + |\zeta|^3 w_n^3 n^{\theta+\tau} \rho^{n_2} + |\zeta|^4 n^{1+2\tau+\theta} w_n^4 \rho^{n_2} + |\zeta|^3 w_n^3 n^\tau \ln^2(n) + |\zeta|^4 w_n^4 n^{2\tau} \ln^2(n).$$

In the sum above, for  $n$  big, we keep only the last two terms, since for  $n$  big enough the first terms are smaller than the last ones.  $\square$

If  $X$  and  $Y$  are two real square integrable random variables, then  $|\mathbb{E}(e^{iX}) - \mathbb{E}(e^{iY})| \leq \|X - Y\|_2$ . Therefore, using (71), we have for  $J_n(\zeta) := \int_X e^{i\zeta S_n} d\mu$ :

$$(77) \quad |J_n(\zeta) - I_{n,p(n)}(\zeta)| \leq |\zeta| \|S_n - S'_n\|_2 \leq C |\zeta| w_n n^{\frac{1-\tau+\delta}{2}} (\ln n)^{\frac{1}{2}},$$

then, by (77) and (72) of Lemma 3.3, we get

$$\begin{aligned} |J_n(\zeta) - \prod_{k=1}^{p(n)} (1 - \frac{1}{2} \zeta^2 v_k^2)| &\leq |J_n(\zeta) - I_{n,p(n)}(\zeta)| + \sum_{k=0}^{p(n)-1} |I_{n,k}(\zeta) - (1 - \frac{\zeta^2}{2} v_k^2) I_{n,k-1}(\zeta)| \\ &\leq C [|\zeta| w_n n^{\frac{1-\tau+\delta}{2}} (\ln n)^{1/2} + n^{1-\tau} |\zeta|^3 w_n^3 n^\tau \ln^2(n) + n^{1-\tau} \zeta^4 w_n^4 n^{2\tau} \ln(n)^2] \\ (78) \quad &\leq C [|\zeta| w_n n^{\frac{1-\tau+\delta}{2}} (\ln n)^{\frac{1}{2}} + |\zeta|^3 w_n^3 n (\ln n)^2 + \zeta^4 w_n^4 n^{1+\tau} (\ln n)^2]. \end{aligned}$$

*Approximation of the exponential by a product*

Below,  $\zeta$  will be such that  $|\zeta| v_{n,k} \leq 1$ . This is satisfied if

$$(79) \quad |\zeta| n^{\frac{\tau}{2}} w_n (\log n)^{\frac{1}{2}} \leq 1.$$

**Lemma 3.4.** *If  $(\rho_k)_{k \in J}$  is a finite family of real numbers in  $[0, 1]$ , then*

$$(80) \quad 0 \leq e^{-\sum_{k \in J} \rho_k} - \prod_{k \in J} (1 - \rho_k) \leq \sum_{k \in J} \rho_k^2, \text{ if } 0 \leq \rho_k \leq \frac{1}{2}, \forall k.$$

*Proof.* We have  $\ln(1 - u) = -u - u^2 v(u)$ , with  $\frac{1}{2} \leq v(u) \leq 1$ , for  $0 \leq u \leq \frac{1}{2}$  and  $1 - e^{-\sum \varepsilon_k} \leq \sum \varepsilon_k$ , if  $\sum_k \varepsilon_k \geq 0$ .

Writing  $1 - \rho_k = e^{-\rho_k - \varepsilon_k}$ , with  $\varepsilon_k = -(\ln(1 - \rho_k) + \rho_k)$ , the previous inequality implies  $0 \leq \varepsilon_k \leq \rho_k^2$ , if  $0 \leq \rho_k \leq \frac{1}{2}$ . Therefore, under this condition, we have:

$$0 \leq e^{-\sum_J \rho_k} - \prod_J (1 - \rho_k) = e^{-\sum_J \rho_k} (1 - e^{-\sum_J \varepsilon_k}) \leq e^{-\sum_J \rho_k} \sum \varepsilon_k \leq \sum \varepsilon_k \leq \sum \rho_k^2. \quad \square$$

We apply (80) with  $\rho_k = \frac{1}{2}\zeta^2 v_{n,k}^2$ , under Condition (79). In view of (68) it follows:

$$(81) \quad |e^{\frac{1}{2}\zeta^2 \sum v_k^2} - \prod_{k=0}^{p(n)-1} (1 - \frac{1}{2}\zeta^2 v_k^2)| \leq \frac{1}{4}\zeta^4 \sum_{k=0}^{p(n)-1} v_k^4 \leq C\zeta^4 w_n^4 n^{1+\tau} \ln^2(n).$$

The bound is like the last term in (78).

### Conclusion

From (78) and (81), it follows:

$$|J_n(\zeta) - e^{-\frac{1}{2}\zeta^2 \sum_{k=0}^{p(n)-1} v_k^2}| \leq C [|\zeta| w_n n^{\frac{1-\tau+\delta}{2}} (\ln n)^{\frac{1}{2}} + |\zeta|^3 w_n^3 n (\ln n)^2 + \zeta^4 w_n^4 n^{1+\tau} (\ln n)^2].$$

We replace  $\zeta$  by  $\frac{\lambda}{\|S_n\|_2}$ ; hence Condition (79) becomes

$$(82) \quad \frac{\lambda}{\|S_n\|_2} n^{\frac{\tau}{2}} w_n (\log n)^{\frac{1}{2}} \leq 1.$$

We get:  $|\int_X e^{i\lambda \frac{S_n(x)}{\|S_n\|_2}} d\mu(x) - e^{-\frac{1}{2} \frac{\lambda^2}{\|S_n\|_2^2} \sum_k v_k^2}|$

$$\leq C [|\lambda| \frac{w_n}{\|S_n\|_2} n^{1-\tau+\delta} + |\lambda|^3 \frac{w_n^3}{\|S_n\|_2^3} n \ln^2(n) + \lambda^4 \frac{w_n^4}{\|S_n\|_2^4} n^{1+\tau} \ln^2(n)].$$

Since  $|e^{-a} - e^{-b}| \leq |a - b|$ , for any  $a, b \geq 0$ , we have, by (70):

$$|e^{-\frac{1}{2}\lambda^2} - e^{-\frac{1}{2} \frac{\lambda^2}{\|S_n\|_2^2} \sum_{k=0}^{p(n)-1} v_k^2}| \leq \frac{1}{2} \frac{\lambda^2}{\|S_n\|_2^2} \left| \|S_n\|_2^2 - \sum_{k=0}^{p(n)-1} v_k^2 \right| \leq C\lambda^2 \frac{w_n^2}{\|S_n\|_2^2} n^{1-\frac{\tau-\delta}{2}} \ln n.$$

Let us call respectively  $E_1$  the error in neglecting the sums on the small blocks,  $E_2$  the error in the replacement of  $e^{-\frac{1}{2}\lambda^2}$  by  $\exp(-\frac{1}{2}\lambda^2 \frac{\sum v_k^2}{\|S_n\|_2^2})$ ,  $E_3$  the error of order 3 in the expansion,  $E_4$  the approximation error of the exponential by the product.

Finally we get the bound  $|\int_X e^{i\lambda \frac{S_n(x)}{\|S_n\|_2}} d\mu(x) - e^{-\frac{1}{2}\lambda^2}| \leq E_1 + E_2 + E_3 + E_4 \leq$

$$C [|\lambda| \frac{w_n}{\|S_n\|_2} n^{\frac{1-\tau+\delta}{2}} \ln^{\frac{1}{2}}(n) + \lambda^2 \frac{w_n^2}{\|S_n\|_2^2} n^{1-\frac{\tau-\delta}{2}} \ln n + |\lambda|^3 \frac{w_n^3}{\|S_n\|_2^3} n \ln^2(n) + \lambda^4 \frac{w_n^4}{\|S_n\|_2^4} n^{1+\tau+\delta} \ln^2(n)].$$

Denote by  $Y_1$  a r.v. with  $\mathcal{N}(0, 1)$ -distribution. Putting  $R_n := \frac{w_n}{\|S_n\|_2}$ , the bound reads:

$$(83) \quad C [|\lambda| R_n n^{\frac{1-\tau+\delta}{2}} + \lambda^2 R_n^2 n^{1-\frac{\tau-\delta}{2}} \ln n + |\lambda|^3 R_n^3 n \ln^2(n) + \lambda^4 R_n^4 n^{1+\tau+\delta} \ln^2(n)].$$

Notice that  $\delta$  can be taken arbitrary small. A change of its value modifies the generic constant  $C$  in the previous inequalities. Therefore we take  $\delta = 0$  in the optimisation below, keeping in mind that the constant factor in the inequalities depends on  $\delta$ . Likewise the  $\ln n$  factors can be neglected.

We have an inequality of the form  $H_{\frac{S_n}{\|S_n\|_2}, Y_1}(\lambda) \leq C \sum_{i=1}^4 |\lambda|^{\alpha_i} R_n^{\alpha_i} n^{\gamma_i}$ , where the exponents are given by the previous inequality. In view of (61), it follows that, up to a constant

factor,  $d(\frac{S_n}{\|S_n\|_2}, Y_1) \leq$

$$\frac{1}{U_n} + \sum_{i=1}^4 \alpha_i^{-1} U_n^{\alpha_i} R_n^{\alpha_i} n^{\gamma_i} \leq \frac{1}{U_n} + U_n R_n n^{\frac{1-\tau}{2}} + \frac{1}{2} U_n^2 R_n^2 n^{1-\tau} + \frac{1}{3} U_n^3 R_n^3 n + \frac{1}{4} U_n^4 R_n^4 n^{1+\tau}.$$

Now, we optimize the choice of  $U = U_n$ . As  $R_n$  is less than  $n^{-\beta}$  for some  $\beta > 0$ , if we take  $U_n = n^\gamma$  with  $\gamma > 0$ , then the previous inequality gives inside the bracket the bound:

$$n^{\frac{1-\tau}{2}-\beta+\gamma} + \frac{1}{2} n^{1-\frac{\tau}{2}-2\beta+2\gamma} + \frac{1}{3} n^{1-3\beta+3\gamma} + \frac{1}{4} n^{1+\tau-4\beta+4\gamma}.$$

We choose  $U_n$  such that  $1/U_n$  is of the same order as the second term, i.e., we take  $n^{-\gamma} = n^{1-\frac{\tau}{2}-2\beta+2\gamma}$ , i.e.,  $\gamma = \frac{\frac{\tau}{2}+2\beta-1}{3}$ . If  $\tau = \frac{1}{2}$  and if  $\beta = \frac{1}{2} - p$  with  $p > 0$ , then it gives:

$$\gamma = \frac{1-8p}{12} > 0 \text{ if } p < \frac{1}{8}.$$

The four terms in the bound are respectively:

$$(A) = -\frac{1}{6} + \frac{p}{3}, \quad (B) = -\gamma = -\frac{1}{12} + \frac{2p}{3}, \quad (C) = -\frac{1}{4} + p, \quad (D) = -\frac{1}{6} + \frac{4p}{3}.$$

We check that  $(B)$  is the biggest term:  $(B) - (A) = \frac{1}{12} + \frac{p}{3} > 0$ ,  $(B) - (C) = \frac{1}{6} - \frac{p}{3} > 0$ , if  $p < \frac{1}{2}$ ,  $(B) - (D) = \frac{1}{12} - \frac{2p}{3} > 0$ , if  $p < \frac{1}{8}$ .

This gives the bound stated in Proposition 2.1 for the distance to the normal law:

For every  $\delta > 0$ , for  $N$  big enough, there is a constant  $C(\delta) > 0$  (depending only on  $\delta$ ) such that, if  $\frac{w_N}{\|S_N\|_2} \leq N^{p-\frac{1}{2}}$  with  $p \in [0, \frac{1}{8}[$ , then

$$d(\frac{S_N}{\|S_N\|_2}, Y_1) \leq C(\delta) N^{-\frac{1-8p}{12}+\delta} \ll 1.$$

To conclude, observe that, if  $w_n \leq n^p$  with  $p < \frac{1}{8}$  and  $\|S_n\|_2 \geq Cn^{\frac{1}{2}}/(\log n)^{\frac{1}{2}}$ , (82) is satisfied for  $|\lambda| \leq U_n = n^\gamma$ , since

$$n^\gamma n^{\frac{\tau}{2}} \frac{w_n (\log n)^{\frac{1}{2}}}{\|S_n\|_2} \leq Cn^{\frac{1-8p}{12}} n^{p-\frac{1}{4}} (\log n)^{\frac{1}{2}} = Cn^{-\frac{1}{6}+\frac{p}{3}} (\log n)^{\frac{1}{2}} \leq 1, \text{ for } n \text{ big. } \square$$

#### 4. Proof of Proposition 2.3 (decorrelation)

For the proof of Proposition 2.3, by homogeneity, we may assume that  $\psi$  and  $\varphi$  are BV centered functions with variation  $\leq 1$ . Moreover, we may also assume  $b_i = 1, \forall i$ . Indeed, the decorrelation inequalities will follow from bounds on sums of products of quantities like  $|\widehat{\varphi_{b_n q_n}}(j)| \leq b_n |\widehat{\varphi_{q_n}}(j)|$  or  $\|\varphi_{b_n q_n}\|_2 \leq b_n \|\varphi_{q_n}\|_2$ .

First we truncate the Fourier series of the ergodic sums  $\varphi_q$ . For functions in  $\mathcal{C}$ , the remainders are easily controlled and it suffices to treat the case of trigonometric polynomials.

For  $\varphi \in \mathcal{C}$ , the Fourier coefficients of order  $j \neq 0$  of the ergodic sum  $\varphi_n$  satisfy:

$$(84) \quad |\widehat{\varphi_n}(j)| = \frac{|\gamma_j(\varphi)|}{|j|} \frac{|\sin \pi n j \alpha|}{|\sin \pi j \alpha|} \leq \frac{\pi}{2} \frac{K(\varphi)}{|j|} \frac{\|n j \alpha\|}{\|j \alpha\|}.$$

Recall also (cf. (8)) that, if  $q$  is a denominator of  $\alpha$ , then

$$\|\varphi_q\|_\infty = \sup_x \left| \sum_{\ell=0}^{q-1} \varphi(x + \ell\alpha) \right| \leq V(\varphi) \text{ and } \|\varphi_q\|_2 \leq 2\pi K(\varphi).$$

We will use the notations:  $S_L f$  for the partial sum of order  $L \geq 1$  of the Fourier series of  $f \in L^2(\mathbb{T})$ ,  $R_L f := f - S_L f$  for the remainder and,  $q_n$  denoting the denominators of  $\alpha$ ,

$$(85) \quad a'_n := \frac{q_{n+1}}{q_n} \leq a_{n+1} + 1, \quad c_n := \frac{q_{n+1}}{q_n} \ln q_{n+1}.$$

### Preliminary inequalities and truncation

We begin by some inequalities which are valid for any irrational number  $\alpha$ .

**Lemma 4.1.** *There is a constant  $C$  such that, if  $q$  is a denominator of  $\alpha$ ,*

$$(86) \quad \sum_{|j| \geq q} \frac{1}{j^2} \frac{\|Lj\alpha\|^2}{\|j\alpha\|^2} \leq C \frac{L}{q}, \quad \forall L \in [1, q].$$

*Proof.* If  $f$  is a non negative BV function with integral  $\mu(f)$ , by Denjoy-Koksma inequality applied to  $f - \mu(f)$ , we have

$$\sum_{j=q}^{\infty} \frac{f(j\alpha)}{j^2} \leq \sum_{i=1}^{\infty} \frac{1}{(iq)^2} \sum_{r=0}^{q-1} f((iq+r)\alpha) \leq \frac{1}{q^2} \left( \sum_{i=1}^{\infty} \frac{1}{i^2} \right) (q\mu(f) + V(f)) = \frac{\pi^2}{6} \left( \frac{\mu(f)}{q} + \frac{V(f)}{q^2} \right).$$

Taking for  $f(x)$  respectively  $1_{[0, \frac{1}{L}]}(|x|)$  and  $\frac{1}{x^2} 1_{[\frac{1}{L}, \frac{1}{2}]}(|x|)$ , we obtain:

$$\sum_{j: \|j\alpha\| \leq \frac{1}{L}, j \geq q} \frac{1}{j^2} \leq C \left( \frac{1}{q^2} + \frac{1}{Lq} \right), \quad \sum_{j: \|j\alpha\| \geq \frac{1}{L}, j \geq q} \frac{1}{j^2} \frac{1}{\|j\alpha\|^2} \leq C \left( \frac{L^2}{q^2} + \frac{L}{q} \right).$$

This implies (86), since for  $L \leq q$ :

$$\frac{1}{2} \sum_{|j| \geq q} \frac{1}{j^2} \frac{\|Lj\alpha\|^2}{\|j\alpha\|^2} \leq L^2 \sum_{\|j\alpha\| \leq \frac{1}{L}, j \geq q} \frac{1}{j^2} + \sum_{\|j\alpha\| > \frac{1}{L}, j \geq q} \frac{1}{j^2} \frac{1}{\|j\alpha\|^2} \leq 2C \left( \frac{L}{q} + \frac{L^2}{q^2} \right) \leq 4C \frac{L}{q}. \quad \square$$

We will use the good repartition of the numbers  $\|k\alpha\|$  when  $k$  varies between 1 and  $q_n$  through two inequalities given in the following lemma, which will be used several times.

**Lemma 4.2.** *We have*

$$(87) \quad \sum_{j=q_t}^{q_{t+1}-1} \frac{1}{\|j\alpha\|} \leq \sum_{j=1}^{q_{t+1}-1} \frac{1}{\|j\alpha\|} \leq C q_{t+1} \ln q_{t+1}, \quad \forall t \geq 0,$$

$$(88) \quad \sum_{1 \leq j < q_{r+1}} \frac{1}{j \|j\alpha\|} \leq C \sum_{t=0}^r \frac{q_{t+1}}{q_t} \ln q_{t+1} = C \sum_{t=0}^r c_t, \quad \forall r \geq 0.$$

*Proof.* There is exactly one element of the set  $\{j\alpha \bmod 1, j = 1, \dots, q_{t+1} - 1\}$  in each interval  $[\frac{\ell}{q_{t+1}}, \frac{\ell+1}{q_{t+1}}[$ ,  $\ell = 1, \dots, q_{t+1} - 1$ . Moreover, for  $1 \leq j < q_{t+1}$ , one has  $\|j\alpha\| \geq \frac{1}{2q_{t+1}}$ .



This implies: 
$$\sum_{j=1}^{q_{t+1}-1} \frac{1}{\|j\alpha\|} \leq 2q_{t+1} + \sum_{\ell=1}^{q_{t+1}-1} \frac{1}{\ell/q_{t+1}} \leq Cq_{t+1} \ln q_{t+1}.$$

From (87) applied for  $t = 1, \dots, r$ , we deduce (88):

$$\sum_{1 \leq j < q_{r+1}} \frac{1}{j \|j\alpha\|} = \sum_{t=0}^r \sum_{q_t \leq j < q_{t+1}} \frac{1}{j \|j\alpha\|} \leq \sum_{t=0}^r \frac{1}{q_t} \sum_{q_t \leq j < q_{t+1}} \frac{1}{\|j\alpha\|} \leq C \sum_{t=0}^r \frac{q_{t+1}}{q_t} \ln q_{t+1}. \quad \square$$

**Lemma 4.3.** *For  $\varphi \in \mathcal{C}$ , it holds*

$$(89) \quad \|S_{q_r} \varphi_{q_n}\|_{\infty} \leq C V(\varphi) \ln(q_r).$$

*Proof.* Using the Fejer kernel, we get

$$\|S_{q_r} \varphi_{q_n}\|_{\infty} \leq \|\varphi_{q_n}\|_{\infty} + \frac{1}{q_r} \sum_{|j| < q_r} |j \widehat{\varphi_{q_n}}(j)| \leq \|\varphi_{q_n}\|_{\infty} + CK(\varphi) \frac{1}{q_r} \sum_{j=1}^{q_r-1} \frac{1}{\|j\alpha\|}.$$

(89) follows by (8) and (87). □

### Truncation

Now we bound the truncation error for the Fourier series of the ergodic sums  $\varphi_{b_n q_n}$ .

**Lemma 4.4.** *If  $\psi$  is bounded and  $\varphi \in \mathcal{C}$ , with  $C_1 = V(\varphi)^2 \|\psi\|_{\infty}$ ,  $C_2 = V(\varphi)^3 \|\psi\|_{\infty}$ , up to a numerical factor, we have, for  $q_n \leq q_m \leq q_r \leq q_\ell$ :*

$$(90) \quad \left| \int \psi [\varphi_{q_n} \varphi_{q_m} - S_{q_\ell} \varphi_{q_n} S_{q_\ell} \varphi_{q_m}] d\mu \right| \leq C_1 \left(\frac{q_m}{q_\ell}\right)^{\frac{1}{2}},$$

$$(91) \quad \left| \int \psi [\varphi_{q_n} \varphi_{q_m} \varphi_{q_r} - S_{q_\ell} \varphi_{q_n} S_{q_\ell} \varphi_{q_m} S_{q_\ell} \varphi_{q_r}] d\mu \right| \leq C_2 \left(\frac{q_r}{q_\ell}\right)^{\frac{1}{2}} \ln^2(q_\ell).$$

*Proof.* We use the bound (86) which gives, for  $q_n \leq q_\ell$ ,

$$\|R_L \varphi_{q_n}\|_2^2 = \sum_{|j| \geq q_\ell} |\widehat{\varphi_{q_n}}(j)|^2 = \sum_{|j| \geq q_\ell} \frac{|\gamma_j(\varphi)|^2}{j^2} \frac{\|q_n j \alpha\|^2}{\|j \alpha\|^2} \leq C^2 K(\varphi)^2 \frac{q_n}{q_\ell}.$$

For  $\psi$  bounded, as  $\|\varphi_{q_n}\|_2 \leq CK(\varphi)$ , this implies:  $\left| \int \psi [\varphi_{q_n} \varphi_{q_m} - S_{q_\ell} \varphi_{q_n} S_{q_\ell} \varphi_{q_m}] d\mu \right| \leq$

$$\|\psi\|_{\infty} [\|\varphi_{q_n}\|_2 \|R_{q_\ell} \varphi_{q_m}\|_2 + \|R_{q_\ell} \varphi_{q_n}\|_2 \|\varphi_{q_m}\|_2] \leq C V(\varphi)^2 \|\psi\|_{\infty} \left[ \left(\frac{q_n}{q_\ell}\right)^{\frac{1}{2}} + \left(\frac{q_m}{q_\ell}\right)^{\frac{1}{2}} \right].$$

This proves (90). For (91), in each term of the expansion of  $(S_{q_\ell} \varphi_{q_n} + R_{q_\ell} \varphi_{q_n})(S_{q_\ell} \varphi_{q_m} + R_{q_\ell} \varphi_{q_m})(S_{q_\ell} + R_{q_\ell}) - S_{q_\ell} \varphi_{q_n} S_{q_\ell} \varphi_{q_m} S_{q_\ell} \varphi_{q_r}$ , we bound one factor in  $L^2$ -norm and the others in uniform norm using (89). □

### Inequalities under Hypothesis 1

Recall that the decorrelation inequalities of Lemma 4.5 are based on Hypothesis 1 on  $\alpha$ .

From (48) in Hypothesis 1, one deduces: for constants  $B, C$ , the coefficients in Ostrowski's expansion satisfy  $b_n \leq B n^p$  and, since  $q_n \leq B^n (n!)^p$ ,

$$(92) \quad \ln q_n \leq C n \ln n, \quad c_n \leq C n^{p+1} \ln n.$$

The case when  $\alpha$  is bpq corresponds to  $p = 0$  and we have then  $\ln q_n \leq Cn$ .

Let us mention that Hardy and Littlewood in [8] considered quantities similar to that below for  $\alpha$  bpq. One of their motivations was to study the asymptotic of the number of integral points contained in homothetic of triangles.

**Lemma 4.5.** *If  $a_{k+1} \leq Ak^p, \forall k \geq 1$  and  $n \leq m \leq \ell$ , we have for every  $\Lambda \geq 1$ :*

$$(93) \quad \sum_{j=1}^{\infty} \frac{\|q_n j \alpha\|}{j^2 \|j \alpha\|} \leq C \frac{n^{p+2} \ln n}{q_{n+1}},$$

$$(94) \quad \sum_{1 \leq j, k < q_\Lambda, j \neq k} \frac{\|q_n j \alpha\| \|q_m k \alpha\|}{|k-j| k j \|j \alpha\| \|k \alpha\|} \leq \frac{C}{q_{n+1}} \Lambda^{2p+4} (\ln \Lambda)^2,$$

$$(95) \quad \sum_{-q_\Lambda < i, j, k < q_\Lambda, i+j+k \neq 0} \frac{\|q_n i \alpha\| \|q_m j \alpha\| \|q_\ell k \alpha\|}{|i+j+k| i j k \|i \alpha\| \|j \alpha\| \|k \alpha\|} \leq \frac{C}{q_{n+1}} \Lambda^{3p+8}.$$

*Proof.* 1) *Proof of (93)*

We use the inequalities:  $\frac{\|j q_k \alpha\|}{j} \leq \|q_k \alpha\| \leq \frac{1}{q_{k+1}}$  for  $j < q_{k+1}$ ,  $\|j q_k \alpha\| \leq 1$  for  $j \geq q_{k+1}$ . For  $\ell > n$ , we write  $\sum_{j=1}^{q_\ell-1} \frac{\|q_n j \alpha\|}{j^2 \|j \alpha\|} = (A) + (B)$ , with

$$\begin{aligned} (A) &:= \sum_{j=1}^{q_{n+1}-1} \frac{1}{j} \frac{\|q_n j \alpha\|}{j \|j \alpha\|} \leq \frac{1}{q_{n+1}} \sum_{j=1}^{q_{n+1}-1} \frac{1}{j} \frac{1}{\|j \alpha\|} \\ &\leq \frac{1}{q_{n+1}} \sum_{k=0}^n \frac{1}{q_k} \sum_{j=q_k}^{q_{k+1}-1} \frac{1}{\|j \alpha\|} \leq C \frac{1}{q_{n+1}} \sum_{k=0}^n \frac{q_{k+1}}{q_k} \ln q_{k+1}, \text{ by (88);} \end{aligned}$$

$$\begin{aligned} (B) &:= \sum_{j=q_{n+1}}^{q_\ell-1} \frac{\|q_n j \alpha\|}{j^2 \|j \alpha\|} \leq \sum_{k=n+1}^{\ell-1} \sum_{j=q_k}^{q_{k+1}-1} \frac{1}{j^2 \|j \alpha\|} \\ &\leq \sum_{k=n+1}^{\ell-1} \frac{1}{q_k^2} \sum_{j=q_k}^{q_{k+1}-1} \frac{1}{\|j \alpha\|} \leq C \sum_{k=n+1}^{\ell-1} \frac{1}{q_k} \frac{q_{k+1}}{q_k} \ln q_{k+1}, \text{ by (87).} \end{aligned}$$

By (4), we know that  $\frac{q_{n+1}}{q_k} \leq C \rho^{k-n}$ , with  $\rho < 1$ , for  $k \geq n+1$ . By hypothesis,  $a_{k+1} \leq Ak^p$ .

It follows with the notation (85):  $(A) \leq \frac{C}{q_{n+1}} \sum_{k=0}^n c_k \leq \frac{C n^{p+2} \ln n}{q_{n+1}}$  and for  $(B)$ , with a bound which doesn't depend on  $\ell \geq n$ :

$$\frac{1}{q_{n+1}} \sum_{k=n+1}^{\ell-1} \frac{q_{n+1}}{q_k} \frac{q_{k+1}}{q_k} \ln q_{k+1} \leq C \frac{1}{q_{n+1}} \sum_{j=0}^{\infty} \rho^j (j+n+1)^{p+1} \ln(j+n+1) \leq \frac{C n^{p+1} \ln n}{q_{n+1}}. \quad \square$$

2) *Proof of (94)*

To bound the sum in (94), we cover the square  $[1, q_\Lambda[ \times [1, q_\Lambda[$  in  $\mathbb{N} \times \mathbb{N}$  by rectangles  $R_{r,s} = [q_r, q_{r+1}[ \times [q_s, q_{s+1}[$  for  $r$  and  $s$  varying between 0 and  $\Lambda - 1$  and then we bound the sum on each of these rectangles (minus the diagonal if  $r = s$ ).

Distinguishing different cases according to the positions of  $r$  and  $s$  with respect to  $n + 1$  and  $m + 1$ , we have, for  $j \in [q_r, q_{r+1}[$ ,  $k \in [q_s, q_{s+1}[$ ,  $j \neq k$ .

$$\frac{\|q_n j \alpha\| \|q_m k \alpha\|}{|k - j| j k \|j \alpha\| \|k \alpha\|} \leq \frac{1}{q_{\max(r, n+1)} q_{\max(s, m+1)}} \frac{1}{|k - j| \|j \alpha\| \|k \alpha\|}.$$

By (87) and (88), using  $\|(k - j)\alpha\| \leq \|j\alpha\| + \|k\alpha\|$ , we have

$$\begin{aligned} \sum_{(j,k) \in R_{r,s}} \frac{1}{|k - j| \|j \alpha\| \|k \alpha\|} &\leq \sum_{(j,k) \in R_{r,s}} \left( \frac{1}{|k - j| \|(k - j)\alpha\| \|j \alpha\|} + \frac{1}{|k - j| \|(k - j)\alpha\| \|k \alpha\|} \right) \\ &\leq q_{\max(r,s)+1} \ln(q_{\max(r,s)+1}) \sum_{t=0}^{\max(r,s)} c_t. \end{aligned}$$

It follows

$$\begin{aligned} \sum_{(j,k) \in R_{r,s}, j \neq k} \frac{\|q_n j \alpha\| \|q_m k \alpha\|}{|k - j| j k \|j \alpha\| \|k \alpha\|} &\leq \frac{q_{\max(r,s)+1}}{q_{\max(r, n+1)} q_{\max(s, m+1)}} \ln(q_{\max(r,s)+1}) \sum_{t=0}^{\max(r,s)} c_t \\ &\leq \frac{1}{q_{n+1}} \ln(q_{\Lambda+1}) \sum_{t=0}^{\Lambda} c_t \max_{k=1, \dots, \Lambda} a'_k. \end{aligned}$$

The square  $[1, q_\Lambda[ \times [1, q_\Lambda[$  is covered by  $\Lambda^2$  rectangles  $R_{r,s}$  and the sums on these rectangles are bounded by the same quantity. It follows, with Hypothesis 1,

$$\sum_{1 \leq j, k < q_\Lambda, j \neq k} \frac{\|q_n j \alpha\| \|q_m k \alpha\|}{|k - j| j k \|j \alpha\| \|k \alpha\|} \leq \Lambda^2 \frac{C}{q_{n+1}} \ln(q_{\Lambda+1}) \sum_{t=0}^{\Lambda} c_t \max_{k=1, \dots, \Lambda} a'_k \leq \frac{C}{q_{n+1}} \Lambda^{2p+5} \ln(\Lambda)^2. \quad \square$$

### 3) Proof of (95)

Here we consider sums with three indices  $i, j, k$ . Though we do not write it explicitly, these sums are to be understood to be taken on non zero indices  $i, j, k$  such that  $i + j + k \neq 0$ . We cover the set of indices by sets of the form

$$R_{\pm r, \pm s, \pm t} = \{(i, j, k) : \pm i \in [q_r, q_{r+1}[, \pm j \in [q_s, q_{s+1}[, \pm k \in [q_t, q_{t+1}[ \}$$

Distinguishing different cases according to the positions of  $r, s$  and  $t$  with respect to  $n + 1$ , we get: if  $(i, j, k) \in R_{\pm r, \pm s, \pm t}$  and  $n \leq m \leq \ell$ ,

$$(96) \quad \frac{\|q_n i \alpha\| \|q_m j \alpha\| \|q_\ell k \alpha\|}{|i| |j| |k|} \leq \frac{1}{q_{\max(r, n+1)} q_{\max(s, n+1)} q_{\max(t, n+1)}}.$$

$$\text{We have } \frac{1}{\|i \alpha\| \|j \alpha\| \|k \alpha\|} \leq \frac{1}{\|(i + j + k)\alpha\|} \left[ \frac{1}{\|j \alpha\| \|k \alpha\|} + \frac{1}{\|i \alpha\| \|k \alpha\|} + \frac{1}{\|i \alpha\| \|j \alpha\|} \right].$$

We then use (87) and (88) three times, sum over  $R_{\pm r, \pm s, \pm t}$  and get:

$$\begin{aligned} & \sum_{(i,j,k) \in R_{\pm r, \pm s, \pm t}} \frac{1}{|i+j+k| \|i\alpha\| \|j\alpha\| \|k\alpha\|} \\ & \leq \left( \sum_{v=0}^{3 \max(r,s,t)} c_v \right) \ln^2(q_{\max(r,s,t)+1}) (q_{s+1}q_{t+1} + q_{r+1}q_{t+1} + q_{r+1}q_{s+1}). \end{aligned}$$

By (96) we then have:

$$\begin{aligned} & \sum_{R_{\pm r, \pm s, \pm t}} \frac{\|q_n i \alpha\| \|q_m j \alpha\| \|q_\ell k \alpha\|}{|i+j+k| |i| |j| |k| \|i\alpha\| \|j\alpha\| \|k\alpha\|} \\ & \leq C \frac{(\sum_{v=0}^{3 \max(r,s,t)} c_v) \ln^2(q_{\max(r,s,t)})}{q_{\max(r,n+1)} q_{\max(s,n+1)} q_{\max(t,n+1)}} (q_{s+1}q_{t+1} + q_{r+1}q_{t+1} + q_{r+1}q_{s+1}) \\ & \leq \frac{C}{q_{n+1}} \left( \sum_{v=0}^{3\Lambda} c_v \right) \ln^2(q_{\Lambda+1}) \left( \max_{k=1, \dots, \Lambda} a'_k \right)^2. \end{aligned}$$

One needs  $8\Lambda^3$  boxes  $R_{\pm r, \pm s, \pm t}$  to cover the set  $\{-q_\Lambda < i, j, k < q_\Lambda, i+j+k \neq 0\}$ . This implies for a constant  $C$ :

$$\sum_{-q_\Lambda < i, j, k < q_\Lambda, i+j+k \neq 0} \frac{\|q_n i \alpha\| \|q_m j \alpha\| \|q_\ell k \alpha\|}{|i+j+k| |i| |j| |k| \|i\alpha\| \|j\alpha\| \|k\alpha\|} \leq \frac{C}{q_{n+1}} \Lambda^{3p+8}. \quad \square$$

### Proof of Proposition 2.3

By (84) we have  $|\int \psi \varphi_{q_n} d\mu| \leq \sum_{j \neq 0} |\widehat{\varphi_{q_n}}(j)| |\widehat{\psi}(-j)| \leq K \sum_{j \geq 1} \frac{\|q_n j \alpha\|}{j^2 \|j\alpha\|}$  and (51) follows from (93):  $\sum_{j \geq 1} \frac{\|q_n j \alpha\|}{j^2 \|j\alpha\|} \leq C \frac{n^{p+2} \ln n}{q_{n+1}}$ .

We prove now (52). With  $L = q_\Lambda$ , we have:

$$\int \psi S_L \varphi_{q_n} S_L \varphi_{q_m} d\mu = \sum_{|j|, |k| \leq L, j \neq k} \widehat{\varphi_{q_n}}(j) \widehat{\varphi_{q_m}}(k) \widehat{\psi}(j-k).$$

In what follows, the constant  $C$  is equal to  $V(\psi)V(\varphi)^2$  (up to a factor not depending on  $\psi$  and  $\varphi$ ) which may change.

Recall that, by (4), there is a constant  $B$  such that  $m \leq B \ln q_m, \forall m \geq 1$ .

The functions  $\psi, \varphi$  are real valued. By (94), it holds  $|\int \psi S_L \varphi_{q_n} S_L \varphi_{q_m} d\mu| \leq$

$$\sum_{|j|, |k| \leq L, j \neq k} |\widehat{\varphi_{q_n}}(j)| |\widehat{\varphi_{q_m}}(k)| |\widehat{\psi}(j-k)| \leq C \sum_{1 \leq j, k \leq L} \frac{\|q_n j \alpha\| \|q_m k \alpha\|}{|k-j| j k \|j\alpha\| \|k\alpha\|} \leq \frac{C}{q_{n+1}} \Lambda^{2p+4} (\ln \Lambda)^2.$$

Putting it together with the truncation error term (90) and replacing  $q_{n+1}$  by  $q_n$ , we get

$$(97) \quad \left| \int_X \psi \varphi_{q_n} \varphi_{q_m} d\mu \right| \leq C \left[ \frac{\Lambda^{2p+4} (\ln \Lambda)^2}{q_n} + \left( \frac{q_m}{q_\Lambda} \right)^{\frac{1}{2}} \right], \text{ for } n \leq m \leq \Lambda.$$

Recall that  $\left( \frac{q_m}{q_\Lambda} \right)^{\frac{1}{2}} \leq \rho^{\frac{\Lambda-m}{2}}$ . Let us take  $\Lambda - m$  of order  $2(\ln \frac{1}{\rho})^{-1} \ln q_n$ , i.e., such that the second term in the bracket of the RHS of (97) is of order  $1/q_n$ . We have then  $\Lambda \leq \max(m, C_1 \log q_n)$  and with Hypothesis 1 the first term in the bracket is less than

$$\frac{C_1}{q_n} \max((\ln q_n)^{2p+5}, m^{2p+5}) \leq \frac{C_2}{q_n} \max((\ln q_n)^{2p+5}, (\ln q_m)^{2p+5}) \leq \frac{C_2}{q_n} (\ln q_m)^{2p+5} \leq C_3 \frac{m^{2p+5}}{q_n}.$$

This shows (52) with  $\theta_2 = 2p + 5$ .

In the same way, (53) follows from (91) and (95).  $\square$

## REFERENCES

- [1] Beck (J.): Randomness of the square root of 2 and the giant leap, part 1, 2: *Periodica Mathematica Hungarica* 60(2) (2010) p.137-242 and 62(2) (2011) p. 127-246.
- [2] Bromberg (M.) and Ulcigrai (C.): A temporal central limit theorem for real-valued cocycles over rotations, preprint ArXiv:1705.06484
- [3] Conze (J.-P.), Le Borgne (S.): On the CLT for rotations and BV functions: note aux C.R.A.S., t. 357 (2019), no. 2, 212-215. [arXiv:1804.09929]
- [4] Conze (J.-P.), Isola (S.) and Le Borgne (S.): Diffusive behaviour of ergodic sums over rotations, *Stochastics and Dynamics*, vol. 19, no. 02 (2019). [ArXiv:1705.10550]
- [5] Conze (J.-P.), Gutkin (E.): On recurrence and ergodicity for geodesic flows on non-compact periodic polygonal surfaces. *Ergodic Theory Dynam. Systems*, 32(2): 491-515, 2012.
- [6] Dolgopyat (D.) and Sarig (O.): Temporal distributional limit theorems for dynamical systems. *Journal of Statistical Physics* (2016), p. 1-34.
- [7] Feller (W.): *An introduction to probability theory and its application, Vol II, second edition.*
- [8] Hardy (G.H.) and Littlewood (J.E.): Some problems of diophantine approximation: a series of cosecants, *Bulletin of the Calcutta Mathematical Society*, no. 20 (1930), p. 251-266.
- [9] Huveneers (F.): Subdiffusive behavior generated by irrational rotations, *Ergodic Theory Dynam. Systems* 29 (2009), no. 4, p. 1217-1233.
- [10] Kitchens (B.): *Symbolic dynamics. One-sided, two-sided and countable state Markov shifts* Universitext. Springer-Verlag, Berlin, 1998.
- [11] Lezaud (P.): Chernoff-type bound for finite Markov chains, *Ann. Appl. Probab.* 8 (1998), no. 3, p. 849-867.

UNIV RENNES, CNRS, IRMAR - UMR 6625, F-35000 RENNES, FRANCE

*E-mail address:* conze@univ-rennes1.fr

*E-mail address:* stephane.leborgne@univ-rennes1.fr